

UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE PATENT TRIAL AND APPEAL BOARD

BRIGHT DATA LTD.,

Petitioner

v.

METACLUSTER LT, UAB,

Patent Owner,

Case No. IPR2023-01442

U.S. Patent No. 7,020,667

DECLARATION OF DR. TIM ARTHUR WILLIAMS

Mail Stop "PATENT BOARD"
Patent Trial and Appeal Board
U.S. Patent & Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

I, Dr. Tim Arthur Williams, declare the following:

1. I have been retained as an independent expert in this matter by counsel for Bright Data Ltd. (“Bright Data”). I have been asked to provide my opinions on certain issues in connection with at least claims 1-20 of U.S. Patent No. 7,020,667 to Guest, *et al.* (“the ’667 Patent”).

2. I am being paid for my work in preparing this declaration at my normal consulting rate plus reimbursement of direct expenses. My compensation is not tied to the outcome of this matter and is not based on the substance of the opinions that I provide.

3. In preparing my opinions expressed herein, I have reviewed the ’667 Patent, its prosecution history (EX. 1002), and the documents listed in the Petitioner’s List of Exhibits. I have also reviewed Patent Owner’s infringement contentions and Petitioner’s invalidity contentions in the related litigation, *Bright Data Ltd. v. Oxylabs, UAB*, Case No. 2:23-cv-00171 (E.D. Tex.). I may rely upon these materials, my knowledge and experience, and/or additional materials to rebut arguments raised by the patent owner. Further, I may also consider additional documents and information in forming any necessary opinions, including any relevant documents that have not yet have been made available to me.

4. I am familiar with the technology described in the ’667 Patent as of its filing date of July 18, 2002. I have been asked to provide my technical review,

analysis, insights, and opinions regarding the '667 Patent. I have used my experience and insight along with my study of the above-identified references as the basis for my opinions, which support the grounds of unpatentability set forth in the Petition for *Inter Partes* Review of the '667 Patent.

MY QUALIFICATIONS

5. My background and qualifications as a technical expert for this matter are summarized below. A copy of my CV is attached as Exhibit A.

6. I hold degrees from Michigan Technological University (Bachelor of Science in Electrical Engineering, 1976) and the University of Texas at Austin (Master of Science in Electrical Engineering, 1982; Ph.D. in Electrical Engineering, 1985; and Master of Business Administration, 1991). I have also been a Registered Patent Agent with the U.S. Patent and Trademark Office (the "USPTO") since 2002.

7. I am an industry professional with over 45 years of experience in wireless communications, computer networking and telecommunications technology.

8. I am currently active as Chief Executive Officer at Beach Technologies, LLC (Danville, CA) a company related to intellectual property consulting.

9. I am also currently active as a Member at Calumet Venture Management (Madison, WI) a company related to investments into start-up companies.

10. Beginning in 2004, I was the Founder and Chairman at DoceoTech Inc. (Danville, CA) which provides training for engineers in wireless, computer networking, and telephony technologies.

11. From 2008 to 2010, I was Founder and Board Member of BitRail Networks, Inc (Miami, FL). This company designed and produced computer networking equipment. One market the company served was edge devices for residential and community access.

12. From 2006 to 2015, I was Founder and Board Member of BEEcube, Inc. (Freemont, CA). This company built high speed computing and computer networking equipment. One market the company served was networking equipment for backhaul networks used in 5G cellular networks.

13. From 2004 to 2008, I was Founder and CEO of SiBEAM, Inc. This company designed and produced wireless networking IC and equipment.

14. From 1999 to 2000, I was Interim CEO and Advisory Board Member of Atheros Communications, Inc. (Palo Alto, CA). This company designed and produced wireless networking IC and equipment.

15. From 1998 to 2000, I was CTO of Picazo Communications, Inc. (San Jose, CA). This company-built computer networking equipment to provide VoIP PBX functionality.

16. From 1991 to 1998, I was Co-Founder, CTO, VP Engineering of Wireless Access, Inc. (Santa Clara, CA). This company developed over the air communication protocols for communication between the subscriber device and the network.

17. From 1979 to 1991, I was a Member of the Technical Staff at Motorola, Inc. (Schaumburg, IL and Austin, TX). In IL, I designed protocols for Digital voice communications. In TX, I designed ICs for communications including Telecom, Wireless, Cellular and Computer Networking.

18. I believe that I am qualified to provide reliable opinions in the field of the '667 Patent, including what a person of ordinary skill in the art (a "POSA") would have understood from the prior art in this field at that time and his or her view of anticipation and obviousness based on that prior art.

MY UNDERSTANDING OF CLAIM CONSTRUCTION

19. I understand that, before the PTAB, claims are to be given their plain and ordinary meaning in light of the specification and prosecution history as would have been read by a POSA.

MY UNDERSTANDING OF ANTICIPATION

20. I understand that for purposes of an *inter partes* review (IPR) proceeding, the only basis for invalidity that may be considered are patents and printed publications, *i.e.*, whether the claims in question are anticipated under 35

U.S.C. § 102 or rendered obvious under 35 U.S.C. § 103, by patents or printed publications.

21. I understand that a claim is considered anticipated (under 35 U.S.C. § 102) if each and every limitation of the claim may be found, expressly or inherently, within the disclosure of a single prior art reference (patent or printed publication), with the features arranged in the reference in the manner claimed.

22. My understanding of inherent disclosure of a claimed feature by a prior art reference is that it means the prior art reference must necessarily reflect the presence or operation of the feature (and that a possibility or even a strong likelihood that the feature is present is not sufficient disclosure).

MY UNDERSTANDING OF OBVIOUSNESS

23. I understand a patent claim is invalid if the claimed invention would have been obvious to a POSA at the time the invention was made. This means that even if all of the requirements of the claim cannot be found in a single prior art reference that would anticipate the claim, the claim can still be invalid based on a single prior art reference and the knowledge of a POSA or that a POSA would combine multiple references to arrive at the claimed invention, including showing that a POSA would be motivated to do so.

24. I also understand that an obviousness analysis requires an understanding of the scope and content of the prior art, any differences between the

alleged invention and the prior art, and the level of ordinary skill in evaluating the pertinent art. I further understand that to prove prior art, or a combination of prior art renders a patent claim obvious, it is necessary to: identify the particular references, singly or in combination; identify which limitations of the patent claim appear in the references; and explain how the references would have been combined.

25. I also understand that exemplary rationales that may support a conclusion of obviousness include: combining prior art elements according to known methods to yield predictable results; a simple substitution of one known element for another to obtain predictable results; use of a known technique to improve similar devices (methods, or products) in the same way; applying a known technique to a known device (method, or product) ready for improvement to yield predictable results; choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success; a known work in one field of endeavor prompting variations of it for use in either the same field or a different one based on design incentives or other market forces if the variations are predictable to a POSA; as well as some teaching, suggestion, or motivation in the prior art that would have led a POSA to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed invention.

26. I further understand that the above list of rationales provided is not all inclusive. Other rationales may also be sufficient, so long as there is a link between

the factual underpinnings relied on and an articulated rationale for why those factors render the invention obvious.

27. I understand certain objective indicia can also be important evidence regarding whether a patent is obvious or nonobvious. Such indicia include: commercial success of products covered by the patent claims; a long-felt need for the invention; failed attempts by others to make the invention; copying of the invention by others in the field; unexpected results achieved by the invention as compared to the closest prior art; praise of the invention by the infringer or others in the field; the taking of licenses under the patent by others; expressions of surprise by experts and those skilled in the art at the making of the invention; and patentee proceeding contrary to accepted wisdom of the prior art. I understand that to have legal weight, there must be a nexus, *i.e.*, a connection, between any such secondary considerations and the alleged invention. I have not received any evidence of any secondary considerations in this matter, but reserve the right to review additional information as it becomes available and to supplement or amend my declaration accordingly.

28. I also understand that “obviousness” is a legal conclusion based on the underlying factual issues of the scope and content of the prior art, the differences between the claimed invention and the prior art, the level of ordinary skill in the field

of the prior art, and any objective indicia of non-obviousness. For that reason, I am not rendering a legal opinion on the ultimate legal question of obviousness.

29. Rather, my opinions as set forth herein address the underlying facts and factual analysis that would support a legal conclusion of obviousness or nonobviousness. When I use the term “obvious,” therefore, it is with reference to the perspective of a POSA at the time of the claimed invention.

PERSON OF ORDINARY SKILL IN THE ART

30. I believe the qualifications of a POSA in the field of the '667 Patent would be 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. I exceeded that level of skill as of July 18, 2002.

31. Based on the foregoing, I believe that I am qualified to provide reliable opinions in the technical field of the '667 Patent, including opinions on what a POSA would have understood from the patent itself and the prior art in the field at the time of the alleged invention (July 18, 2002).

OVERVIEW OF THE '667 PATENT AND CHALLENGED CLAIMS

Background of the '667 Patent

32. 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 at Abstract. 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. *Id.*

33. 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.” *Id.* at 2:19-24. 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. *Id.* at 2:38-40.

34. “[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.” *Id.* at 4:16-18. 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.” *Id.* at 4:45-50.

35. 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.” *Id.* at 4:57-62. 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.” *Id.* at 5:12-16.

36. 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.” *Id.* at 5:17-25.

Prosecution History of the '667 Patent

37. 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 at 154.

38. 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. *Id.* at 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”). *Id.* at 127. 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 unpatentable over Niazi in view of U.S. Patent No. 6,532,473 to Bright (“Bright”). *Id.* at 130-133.

39. 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. *Id.* at 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. *Id.* at 106-118.

40. 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. *Id.* at 81-85. The examiner further rejected claim 16 as being obvious over Niazi in view of M2 Presswire and further in view of Chau. *Id.* at 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. *Id.* at 85-86.

41. 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 at 59. Applicant also further cancelled claims 26-30 and 33-41. EX. 1002 at 58-76.

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

The Claims

43. The '667 Patent has 20 claims, including three independent claims: Nos. 1, 16, and 20. 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.

44. Dependent claims 2-20 depend from claim 1 and require additional subject matter.

45. Independent claim 16 is a computer program product claim with elements that mirror the requirements of claim 1. Claim 16 specifies a first and second set of instruction codes.

46. Dependent claims 17-19 depend from claim 16 and require additional subject matter.

47. Independent claim 20 recites a computer implemented system comprising “means for” elements to perform the same steps recited in claim 1.

48. The particular claims at issue are claims 1-20.

MY UNDERSTANDING OF CLAIM CONSTRUCTION AND INVALIDITY

49. I understand that, before the PTAB, claims are to be given their plain and ordinary meaning in light of the specification and prosecution history as would have been read by a POSA. I also note that the ‘667 Patent provided general definitions for certain terms used in the specification to facilitate the understanding of the alleged invention without limiting its scope. EX. 1001 at 3:4-54. In my opinion, these definitions are consistent with the plain and ordinary meanings as would have been understood by a POSA.

50. I understand that for purposes of an *inter partes* review (IPR) proceeding, the only basis for invalidity that may be considered are patents and printed publications, i.e., whether the claims in question are anticipated under 35 U.S.C. § 102 or rendered obvious under 35 U.S.C. § 103, by patents or printed publications.

OVERVIEW OF THE STATE OF THE ART AT THE TIME OF FILING

Applicant Admitted Prior Art

51. 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 at 1:34-36. But, the ’667 patent then states that “[t]he user can always manually transfer the data viewed on the Internet into a database.” *Id.* at 1:36-37.

52. 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.” *Id.* at 1:45-52.

53. 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.” *Id.* at 1:54-60.

DaCosta (EX. 1004)

54. 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, I understand that DaCosta qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

55. In my opinion, 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. 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 at codes (51) and (52); EX. 1004 at codes (51) and (52).

56. DaCosta is titled “System for Providing Database Functions for Multiple Internet Sources.” EX. 1004 at code (54). DaCosta presents a system for automatically extracting data from at least one electronic document accessible through the Internet or other computer network. *Id.* at 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.” *Id.* at 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.” *Id.* at 18:18-21.

57. DaCosta discloses “an architecture for a smart web server to scour data from the Internet for deposit in a data repository.” *Id.* at 5:1-3. “[D]ata feeds Web Folders that comprise a knowledge repository accessible by smart clients.” *Id.* at 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.” *Id.* at 11:20-22.

58. 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.” *Id.* at 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.” *Id.* at 18:17-23.

Kraft (EX. 1005)

59. 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, I understand that Kraft qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

60. In my opinion, 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. 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 at codes (51) and (52); EX. 1005 at codes (51) and (52).

61. Kraft is titled “System and Method for Associating a Destination Document to a Source Document During a Save Process.” EX. 1005 at 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. *Id.* at Abstract, code (57). Kraft further discloses a crawler that is run periodically to update previously stored data. *Id.* at ¶[0003]. 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.” *Id.* at ¶[0001].

ANALYSIS OF THE GROUNDS

62. As set forth below, my opinions relate to the following grounds:

Ground 1	Claims 1-20	Anticipation by DaCosta
Ground 2	Claims 1-11, and 13-20	Anticipation by Kraft

GROUND 1: ANTICIPATION OF CLAIMS 1-20 BY DaCOSTA (EX. 1004)

63. In my opinion, DaCosta discloses each of the limitations of claims 1-20 of the '667 patent.

Claim 1

64. **“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: ...”** In my opinion, to the extent the preamble is limiting, DaCosta discloses this claim limitation. DaCosta presents a system for automatically extracting data from at least one electronic document accessible through the Internet or other computer networks. EX. 1004 at 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.” *Id.* at 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.” *Id.* at 18:18-21.

65. “... **initializing a data retrieval stage by: ...**” In my opinion, DaCosta discloses this claim limitation. 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.” *Id.* at 11:5-11.

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

67. “... **selecting the source file; ...**” In my opinion, DaCosta discloses this claim limitation.

68. 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.” *Id.* at 2:31-38.

69. 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.” *Id.* at 3:21-24.

70. 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.” *Id.* at 2:55-64.

71. **“... parsing selected data elements in the source file into a list of data elements; ...”** In my opinion, DaCosta discloses this claim limitation. 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.” *Id.* at 11:53-55.

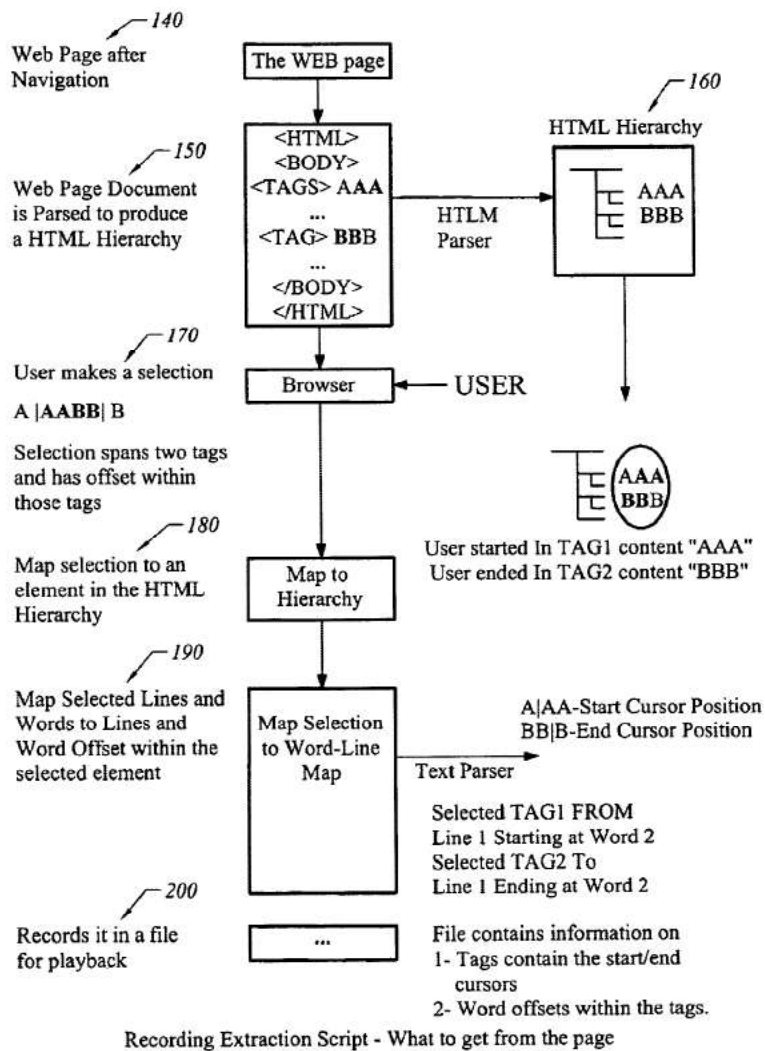


FIG. 4

72. 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.” *Id.* at 7:2-10.

73. **“... selecting at least one of the data elements from the list to be stored in the target database; ...”** In my opinion, DaCosta discloses this claim limitation. 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.’” *Id.* at 9:56-65.

74. 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.” *Id.* at 11:12-17.

75. 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.” *Id.* at 11:25-31.

76. **“... identifying source locations of the selected at least one of the data elements; ...”** In my opinion, DaCosta discloses this claim limitation. 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.” *Id.* at 5:51-55.

77. DaCosta states that “an object of the present invention is to provide a data location and extraction tool capable of automated operation.” *Id.* at 2:2-3. 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.” *Id.* at 5:12-20.

78. **“... 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; ...”** In my opinion, DaCosta discloses this claim limitation.

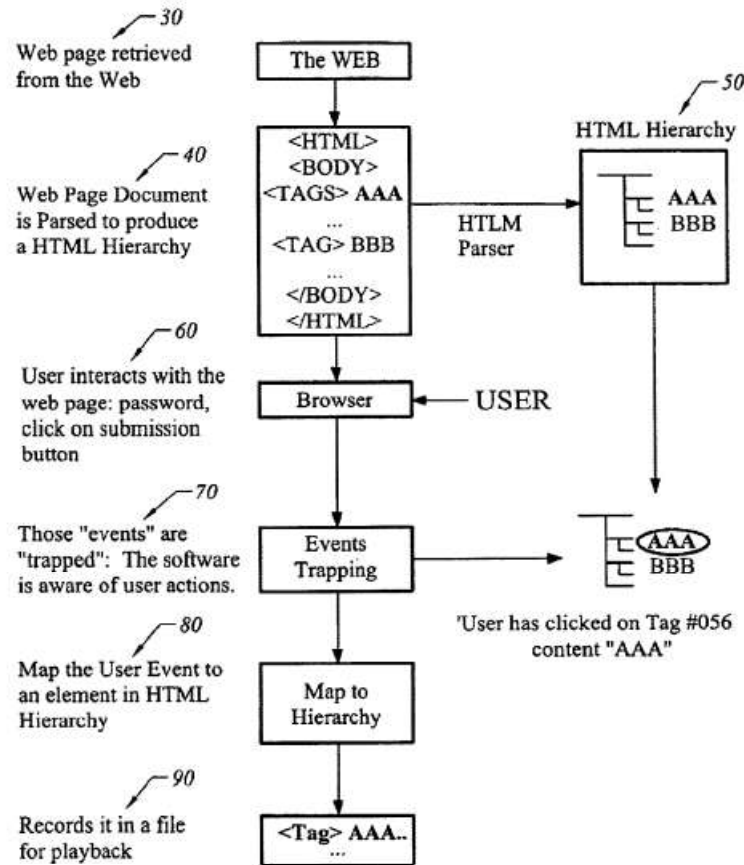
DaCosta discloses “an architecture for a smart web server to scour data from the Internet for deposit in a data repository.” *Id.* at 5:1-3. DaCosta discloses a repository, *e.g.*, a target database, in which acquired data is stored. For example, “extracted data can be stored in a conventional repository, 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.” *Id.* at 18:18-23.

79. 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.” *Id.* at 18:33-38.

80. 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.” *Id.* at 18:15-21.

81. 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.” *Id.* at 5:12-20.

82. 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.



Recording Navigation Script - Getting to a page

FIG. 2

83. "... 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; ..." In my opinion, 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.

84. 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.” *Id.* at 11:19-24.

85. 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.” *Id.* at FIG. 10. DaCosta further discloses that “[d]ata may also be transferred to a data base for SQL processing.” *Id.* at FIG. 10.

86. 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,” *Id.* at 18:34-38.

87. 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.” *Id.* at 18:17-23.

88. **“... 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 ...”** In my opinion, DaCosta discloses this claim limitation. 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.” *Id.* at 5:12-20.

89. 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.” *Id.* at 7:62-65. 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.” *Id.* at 18:23-25.

90. **“... 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.”** In my opinion, DaCosta discloses this limitation. 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.” *Id.* at 7:57-65.

91. 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.” *Id.* at 5:12-20.

92. 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.” *Id.* at 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.” *Id.* at 18:23-25.

93. Thus, in my opinion DaCosta closes all the limitations of claim 1.

Claim 2

94. Claim 2 depends from claim 1 and adds the limitation **“further comprising processing the selectively parsed new data elements stored in the target database.”** I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

95. 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.” *Id.* at 5:12-20.

96. Thus, in my opinion, DaCosta discloses all the limitations of claim 2.

Claim 3

97. Claim 3 depends from claim 2 and adds the limitation “**wherein processing the selectively parsed new data elements is executed automatically.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

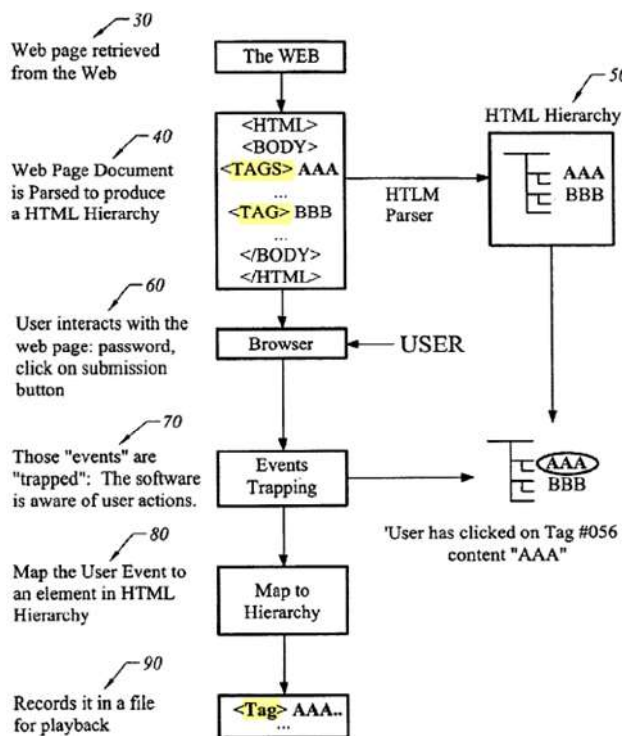
98. DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, 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.” *Id.* at 5:12-20.

99. Thus, in my opinion, DaCosta discloses all the limitations of claim 3.

Claim 4

100. Claim 4 depends from claim 1 and adds the limitation “**wherein the source file is formatted with tags.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in

DaCosta. 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.” *Id.* at 2:52-54. Highlighted FIG. 2 is reproduced below illustrating the use of tags.



Recording Navigation Script - Getting to a page

FIG. 2

101. 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).” *Id.* at 13:48-62.

102. Thus, in my opinion, DaCosta discloses all the limitations of claim 4.

Claim 5

103. Claim 5 depends from claim 1 and adds the limitation “**wherein the at least one of the data elements from the sources locations are derived from an application output file.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

104. 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.” *Id.* at 5:12-20.

105. Thus, in my opinion, DaCosta discloses all the limitations of claim 5.

Claim 6

106. Claim 6 depends from claim 1 and adds the limitation “**wherein the at least one of the data elements from the sources locations are derived from an output of a server program.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. 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 at 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 at 3:53-50; *see also, e.g., id.* at claims 1-4.

107. Thus, in my opinion, DaCosta discloses all the limitations of claim 6.

Claim 7

108. Claim 7 depends from claim 1 and adds the limitation “**wherein the target database is a relational database.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. 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." *Id.* at 5:10-20. DaCosta further discloses a user can "access the information in a database format." *Id.* at 13:64-14:12. 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." *Id.* at 18:8-25.

109. Thus, in my opinion, DaCosta discloses all the limitations of claim 7.

Claim 8

110. Claim 8 depends from claim 1 and adds the limitation "**wherein initializing the data retrieval stage further comprises presenting selection options to a user.**" I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. For example, 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).” *Id.* at 9:25-51.

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

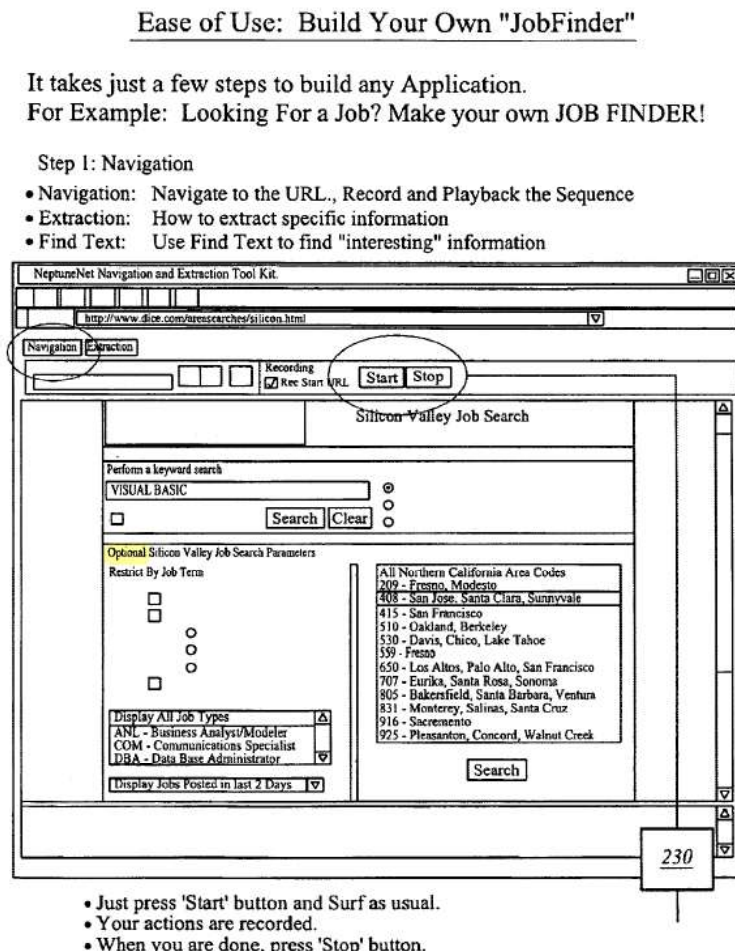


FIG. 8

112. Thus, in my opinion, DaCosta discloses all the limitations of claim 8.

Claim 9

113. Claim 9 depends from claim 1 and adds the limitation “**wherein the at least one of the data elements is dynamic.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. 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. *Id.* at 18:12-14. DaCosta discloses “[s]mart servers provide distributed, multiple dynamic virtual data feeds.” *Id.* at 18:14-15.

114. Further, “[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.” *Id.* at 6:21-26.

115. Thus, in my opinion, DaCosta discloses all the limitations of claim 9.

Claim 10

116. Claim 10 depends from claim 9 and adds the limitation “**wherein initializing the data retrieval stage further comprises analyzing the new data elements.**” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

117. 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.” *Id.* at 7:57-65.

118. Thus, in my opinion, DaCosta discloses all the limitations of claim 10.

Claim 11

119. Claim 11 depends from claim 10 and adds the limitation **“wherein the new data elements comprise data that changes periodically.”** I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

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

121. Further, “[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.” *Id.* at 6:21-26.

122. Thus, in my opinion, DaCosta discloses all the limitations of claim 11.

Claim 12

123. Claim 12 depends from claim 1 and adds the limitation **“wherein processing the new data elements is executed manually.”** I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

124. 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).” *Id.* at 16:14-23.

125. Thus, in my opinion, DaCosta discloses all the limitations of claim 12.

Claim 13

126. Claim 13 depends from claim 1 and adds the limitation **“further comprising automatically updating data elements stored in the target**

database.” I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta.

127. For example, “[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.” *Id.* at 6:21-26.

128. 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.” *Id.* at 5:12-20.

129. Thus, in my opinion, DaCosta discloses all the limitations of claim 13.

Claim 14

130. Claim 14 depends from claim 13 and adds the limitation **“wherein automatically updating the data elements stored in the target database comprises periodically retrieving new data elements and updating the data elements stored in the target database with the retrieved new data elements.”** I

incorporate my above discussion of claims 1 and 13 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. 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.” *Id.* at 7:57-65.

131. 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.” *Id.* at 18:23-25.

132. Thus, in my opinion, DaCosta discloses all the limitations of claim 14.

Claim 15

133. Claim 15 depends from claim 1 and adds the limitation **“further comprising reinitializing the data retrieval stage.”** I incorporate my above discussion of claim 1 in Ground 1. In my opinion, this additional claim limitation is disclosed in DaCosta. 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.” *Id.* at 5:15-20.

134. Thus, in my opinion, DaCosta discloses all the limitations of claim 15.

Claim 16

135. **“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: ...”**

In my opinion, to the extent the preamble is limiting, DaCosta discloses this claim limitation. DaCosta presents a method and system for automatically extracting data from at least one electronic document accessible through the Internet or other computer networks. EX. 1004 at 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.” *Id.* at 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.” *Id.* at 18:18-21.

136. “... a first set of instruction codes for initializing a data retrieval stage by: ...” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation. 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.

137. “... selecting the source file; ...” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

138. “... parsing selected data elements in the source file into a list of data elements; ...” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

139. “... selecting at least one of the data elements from the list to be stored in the target database; ...” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

140. **“... identifying source locations of the selected at least one of the data elements; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

141. **“... 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; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

142. **“... 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; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

143. **“... 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 ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

144. **“... 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.”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation. 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.

Claim 17

145. Claim 17 depends from claim 16 and adds the limitation **“wherein the source file is formatted with tags.”** This claim limitation closely matches a limitation present in claim 4. In my opinion, for the same reasons discussed above regarding claim 4, DaCosta discloses this claim limitation. Thus, in my opinion, DaCosta discloses all the limitations of claim 17.

Claim 18

146. Claim 18 depends from claim 16 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived**

from an application output file.” This claim limitation closely matches a limitation present in claim 5. In my opinion, for the same reasons discussed above regarding claim 5, DaCosta discloses this claim limitation. Thus, in my opinion, DaCosta discloses all the limitations of claim 18.

Claim 19

147. Claim 19 depends from claim 16 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived from an output of a server program.”** This claim limitation closely matches a limitation present in claim 6. In my opinion, for the same reasons discussed above regarding claim 6, DaCosta discloses this claim limitation. Thus, in my opinion, DaCosta discloses all the limitations of claim 19.

Claim 20

148. **“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: ...”** In my opinion, to the extent the preamble is limiting, DaCosta discloses this claim limitation. DaCosta presents a method and system for automatically extracting data from at least one electronic document accessible through the Internet or other computer networks. EX. 1004 at 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.” *Id.* at 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.” *Id.* at 18:18-21.

149. “... **means for initializing a data retrieval stage by: ...**” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation. 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.

150. “... **selecting the source file; ...**” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

151. “... **parsing selected data elements in the source file into a list of data elements; ...**” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

152. **“... selecting at least one of the data elements from the list to be stored in the target database; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

153. **“... identifying source locations of the selected at least one of the data elements; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

154. **“... 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; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

155. **“... 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; ...”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

156. **“... 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 ...” This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation.

157. **“... 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.”** This claim limitation closely matches a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, DaCosta discloses this claim limitation. 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.

GROUND 2: ANTICIPATION OF CLAIMS 1-11, AND 13-20 BY KRAFT (EX. 1005)

158. In my opinion, Kraft discloses each of the limitations of claims 1-11, and 13-20 of the '667 patent.

Claim 1

159. **“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: ...” In my opinion, to the extent the preamble is limiting, Kraft discloses this claim limitation. 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 at Abstract, code (57). Kraft further discloses a crawler that is run periodically to update previously stored data. *Id.* at ¶[0003]. 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.” *Id.* at ¶[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.” *Id.* at ¶[0060].

160. **“... initializing a data retrieval stage by: ...”** In my opinion, Kraft discloses this claim limitation. 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.” *Id.* at ¶[0057]; *see also*, Figure 3.

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

162. “... selecting the source file; ...” In my opinion, Kraft discloses this claim limitation. 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.” *Id.* at ¶[0014]. 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.” *Id.* at ¶[0033].

163. 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).” *Id.* at ¶[0058]. 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).” *Id.* at ¶[0059]. FIG. 1 of Kraft is reproduced below.

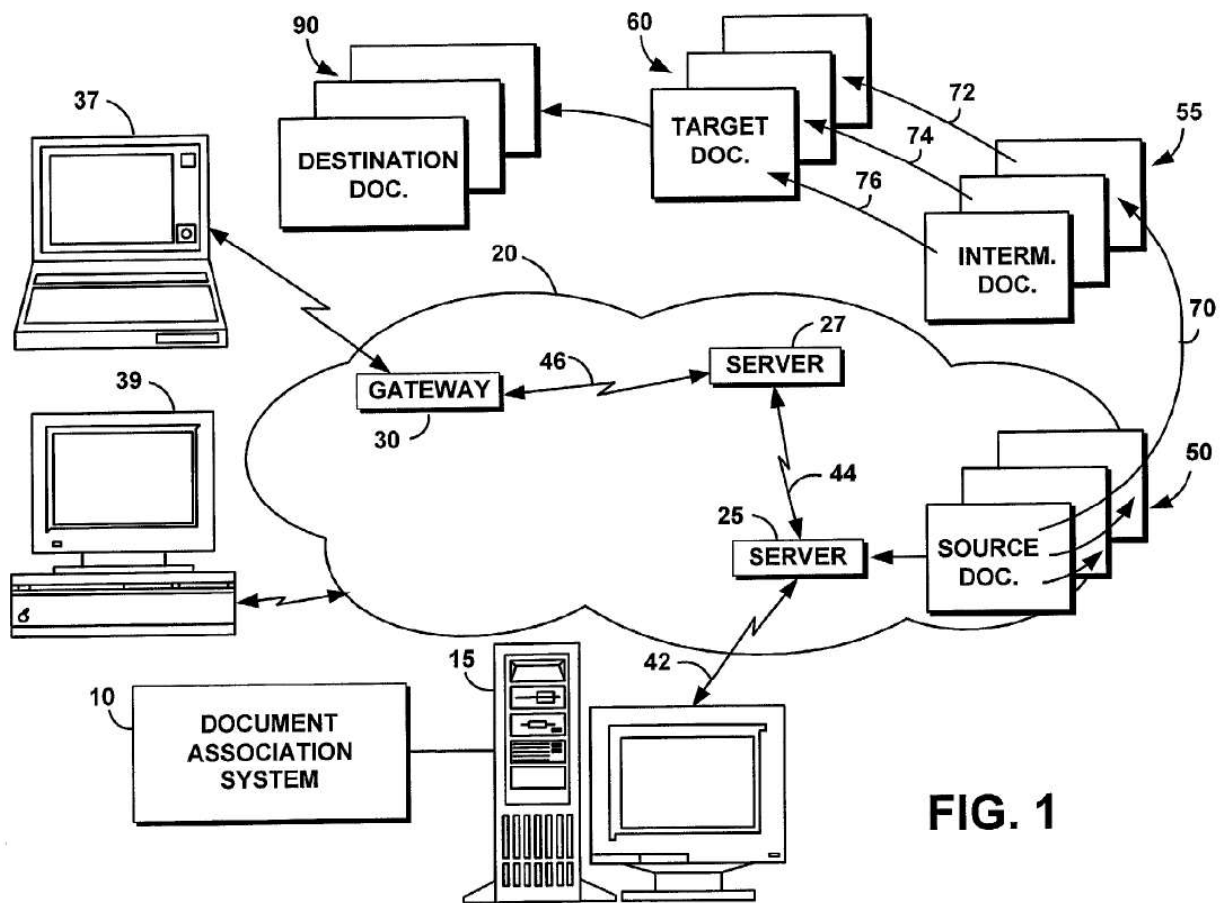


FIG. 1

164. 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>.” *Id.* at ¶¶ [0061]-[0062]. 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. *Id.* at ¶¶ [0063]-[0064].

165. **“... parsing selected data elements in the source file into a list of data elements; ...”** In my opinion, Kraft discloses this claim limitation. 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.” *Id.* at ¶[0028].

166. 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).” *Id.* at ¶[0058].

167. **“... selecting at least one of the data elements from the list to be stored in the target database; ...”** In my opinion, Kraft discloses this claim limitation. 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.” *Id.* at ¶[0059].

168. “... **identifying source locations of the selected at least one of the data elements; ...**” In my opinion, Kraft discloses this claim limitation. 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).” *Id.* at ¶[0058].

169. 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.” *Id.* at ¶[0060].

170. 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.” *Id.* at ¶[0009]. 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.” *Id.* at ¶¶[0010]-[0013].

171. 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.” *Id.* at ¶[0064].

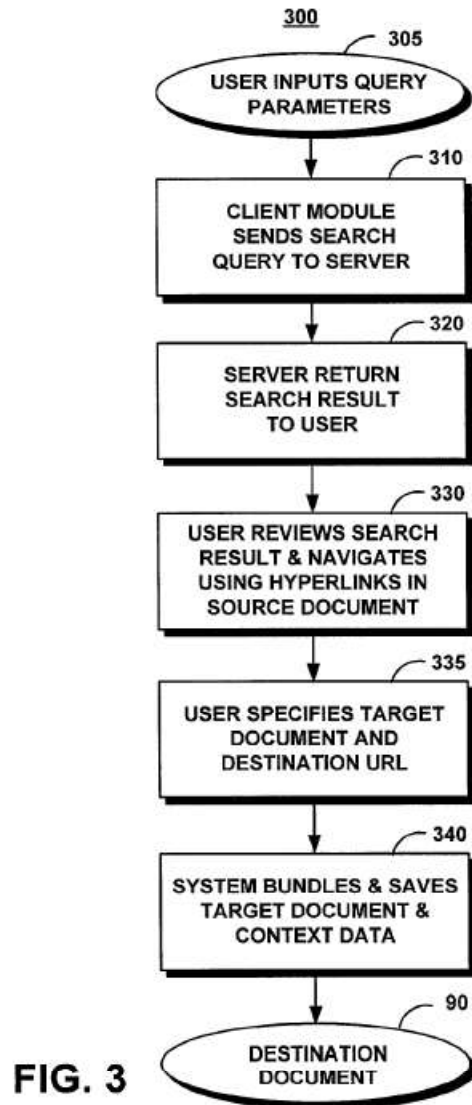
172. **“... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”** In my opinion, Kraft discloses this claim limitation. Kraft discloses the target database as the “destination documents repository 170 where destination documents 90 (FIG, 1) are stored.” *Id.* at ¶[0045].

173. 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).” *Id.* at ¶[0059].

174. 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.” *Id.* at ¶[0051].

175. **“... 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; ...”** In my opinion, Kraft discloses this limitation. 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.” *Id.* at ¶[0009]. The bundled context data and destination document are saved at the destination URL, as shown in FIG. 3, reproduced below. Kraft further discloses destination documents 90 are saved in a destination documents repository 170, *e.g.*, target database. *Id.* at ¶[0045], ¶[0050]; *see also*, Figure 2.



176. 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.” *Id.* at ¶¶[0010]-[0013].

177. 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). *Id.* at ¶¶[0062]-[0063].

178. **“... 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 ...”** In my opinion, Kraft discloses this claim limitation. 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.” *Id.* at ¶[0009].

179. 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.” *Id.* at ¶[0016].

180. Kraft also discloses that the “system 10, transparently to the user, continuously or periodically operates in the background.” *Id.* at ¶[0044]. 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.” *Id.* at ¶[0003].

181. **“... 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.”** In my opinion, Kraft discloses this limitation. 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.” *Id.* at ¶[0016].

182. 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).” *Id.* at ¶¶[0046]-[0047].

183. Thus, in my opinion Kraft discloses all the limitations of claim 1.

Claim 2

184. Claim 2 depends from claim 1 and adds the limitation “**further comprising processing the selectively parsed new data elements stored in the target database.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

185. 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.” *Id.* at ¶¶[0050]-[0051].

186. Kraft further discloses automatically updating the destination document, for example, if the target document was deleted from its original location or relocated. *Id.* at ¶[0016].

187. Thus, in my opinion, Kraft discloses all the limitations of claim 2.

Claim 3

188. Claim 3 depends from claim 2 and adds the limitation “**wherein processing the selectively parsed new data elements is executed automatically.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

189. Kraft further discloses that “the service provider 100 is generally comprised of a web crawler 200.” *Id.* at ¶[0052]. 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.” *Id.* at ¶[0026].

190. Kraft further discloses automatically updating the destination document, for example, if the target document was deleted from its original location or relocated. *Id.* at ¶[0016].

191. Thus, in my opinion, Kraft discloses all the limitations of claim 3.

Claim 4

192. Claim 4 depends from claim 1 and adds the limitation “**wherein the source file is formatted with tags.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.
Kraft

193. 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.” *Id.* at ¶[0028].

194. Thus, in my opinion, Kraft discloses all the limitations of claim 4.

Claim 5

195. Claim 5 depends from claim 1 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived from an application output file.”** I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

196. 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.” *Id.* at ¶[0063].

197. Thus, in my opinion, Kraft discloses all the limitations of claim 5.

Claim 6

198. Claim 6 depends from claim 1 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived from an output of a server program.”** I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft. 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. *Id.* at ¶ [0061]. 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.” *Id.* at ¶[0028].

199. Thus, in my opinion, Kraft discloses all the limitations of claim 6.

Claim 7

200. Claim 7 depends from claim 1 and adds the limitation **“wherein the target database is a relational database.”** I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

201. Kraft discloses the use of databases, for example, “[t]he information compiled by the crawler is stored in a Metadata repository or database.” *Id.* at ¶[0003]. 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.” *Id.* at ¶[0040]. Kraft also discloses the use of documents repository 170, or any other data storage system, for later use by the document storage manager 150. *Id.* at ¶[0050].

202. Kraft further discloses that the disclosed invention is applicable to database “and other tables with indexed entries.” *Id.* at ¶[0056]. 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.” *Id.* at ¶[0054].

203. Thus, in my opinion, Kraft discloses all the limitations of claim 7.

Claim 8

204. Claim 8 depends from claim 1 and adds the limitation “**wherein initializing the data retrieval stage further comprises presenting selection options to a user.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft. 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.” *Id.* at ¶[0057].

205. Thus, in my opinion, Kraft discloses all the limitations of claim 8.

Claim 9

206. Claim 9 depends from claim 1 and adds the limitation “**wherein the at least one of the data elements is dynamic.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft. For example, Kraft discloses that a web page can be dynamic or static. *Id.* at ¶[0037]. 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.” *Id.* at ¶[0045].

207. 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).” *Id.* at ¶¶[0046]-[0047].

208. Thus, in my opinion, Kraft discloses all the limitations of claim 9.

Claim 10

209. Claim 10 depends from claim 9 and adds the limitation “**wherein initializing the data retrieval stage further comprises analyzing the new data elements.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

210. Kraft discloses where “[t]he crawler is run periodically to update previously stored data and to append information about newly created web pages.” *Id.* at ¶[0003].

211. Thus, in my opinion, Kraft discloses all the limitations of claim 10.

Claim 11

212. Claim 11 depends from claim 10 and adds the limitation “**wherein the new data elements comprise data that changes periodically.**” I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

213. As stated above in 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.” *Id.* at ¶[0003]. 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).” *Id.* at ¶¶[0046]-[0047].

214. Thus, in my opinion, Kraft discloses all the limitations of claim 11.

Claim 13

215. Claim 13 depends from claim 1 and adds the limitation **“further comprising automatically updating data elements stored in the target database.”** I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft.

216. 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.” *Id.* at ¶[0009]. 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.” *Id.* at ¶[0016].

217. Thus, in my opinion, Kraft discloses all the limitations of claim 13.

Claim 14

218. Claim 14 depends from claim 13 and adds the limitation **“wherein automatically updating the data elements stored in the target database comprises periodically retrieving new data elements and updating the data elements stored in the target database with the retrieved new data elements.”** I incorporate my above discussion of claims 1 and 13 in Ground 2. In my opinion,

this additional claim limitation is disclosed in Kraft. 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.” *Id.* at ¶[0003].

219. Thus, in my opinion, Kraft discloses all the limitations of claim 14.

Claim 15

220. Claim 15 depends from claim 1 and adds the limitation **“further comprising reinitializing the data retrieval stage.”** I incorporate my above discussion of claim 1 in Ground 2. In my opinion, this additional claim limitation is disclosed in Kraft. 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.” *Id.* at ¶[0026].

221. Thus, in my opinion, Kraft discloses all the limitations of claim 15.

Claim 16

222. **“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: ...”**

In my opinion, to the extent the preamble is limiting, Kraft discloses this claim

limitation. Kraft presents a system and method for use with a web browser and the Internet to save the location and other context along with the content of the web page when being stored. EX. 1005 at Abstract, code (57). Kraft further discloses a crawler that is run periodically to update previously stored data. *Id.* at ¶[0003]. 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.” *Id.* at ¶[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.” *Id.* at ¶[0060].

223. “... a first set of instruction codes for initializing a data retrieval stage by: ...” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

224. “... selecting the source file; ...” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

225. “... parsing selected data elements in the source file into a list of data elements; ...” This claim limitation closely matches a limitation present in

ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

226. **“... selecting at least one of the data elements from the list to be stored in the target database; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

227. **“... identifying source locations of the selected at least one of the data elements; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

228. **“... 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; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

229. **“... 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; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

230. **“... 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 ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

231. **“... 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.”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

Claim 17

232. Claim 17 depends from claim 16 and adds the limitation **“wherein the source file is formatted with tags** This claim limitation closely matches a limitation present in claim 4, ground 2. In my opinion, for the same reasons discussed above regarding claim 4, Kraft discloses this claim limitation. Thus, in my opinion, Kraft discloses all the limitations of claim 17.

Claim 18

233. Claim 18 depends from claim 16 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived from an application output file.”** This claim limitation closely matches a limitation present in claim 5, ground 2. In my opinion, for the same reasons discussed above regarding claim 4, Kraft discloses this claim limitation. Thus, in my opinion, Kraft discloses all the limitations of claim 18.

Claim 19

234. Claim 19 depends from claim 16 and adds the limitation **“wherein the at least one of the data elements from the sources locations are derived from an output of a server program.”** This claim limitation closely matches a limitation present in claim 6, ground 2. In my opinion, for the same reasons discussed above regarding claim 4, Kraft discloses this claim limitation. Thus, in my opinion, Kraft discloses all the limitations of claim 19.

Claim 20

235. **“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: ...”** In my opinion, to the extent the preamble is limiting, Kraft discloses this claim limitation. Kraft presents a system and method for use with a web browser and the Internet to save the location

and other context along with the content of the web page when being stored. EX. 1005 at Abstract, code (57). Kraft further discloses a crawler that is run periodically to update previously stored data. *Id.* at ¶[0003]. 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.” *Id.* at ¶[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.” *Id.* at ¶[0060].

236. “... **means for initializing a data retrieval stage by: ...**” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

237. “... **selecting the source file; ...**” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

238. “... **parsing selected data elements in the source file into a list of data elements; ...**” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

239. **“... selecting at least one of the data elements from the list to be stored in the target database; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

240. **“... identifying source locations of the selected at least one of the data elements; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

241. **“... 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; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

242. **“... 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; ...”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

243. **“... 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 ...” This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

244. **“... 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.”** This claim limitation closely matches a limitation present in ground 2, claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Kraft discloses this claim limitation.

.....

In signing this declaration, I recognize that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office, and that I may accordingly be subject to cross-examination. If such cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

This declaration represents the opinions I have formed to date. I reserve the right to revise, supplement, and/or amend my opinions stated herein based on new information and on my continuing analysis of the materials produced in this proceeding.

I declare that the foregoing is true and correct under penalty of perjury of the laws of the United States of America.

Signed in Danville, CA on 3 Oct, 2023.



Dr. Tim Arthur Williams

EXHIBIT A

Tim Arthur Williams, Ph.D.
Curriculum Vitae

Dr. Williams has 45 years of professional experience in wireless communications, networking and telecom technology. He is an entrepreneur who has participated in the organization and operation of start up companies that brought wireless LAN, software VoIP PBX, and 2-way paging technology to the marketplace. Dr. Williams holds numerous patents in wireless and signal processing technology. He is an experienced litigation support consultant with experience in patent infringement matters. Dr. Williams is also a registered Patent Agent.

- | | |
|----------------------------------|--------------------------------|
| ▪Wireless LAN | ▪Telecommunications Technology |
| ▪Cellular and PCS Standards | ▪VoIP Technology |
| ▪Cellular Telephone Architecture | ▪Computer Networking |
| ▪Digital Signal Processing | ▪Wireless Networks & Protocols |

Year	University	Degree
1991	University of Texas at Austin	MBA
1985	University of Texas at Austin	Ph.D., Dissertation: “Digital Signal Processing Techniques for Acoustic Log Data”
1982	University of Texas at Austin	MSEE, Thesis: “Cepstral Processing of Speech Signals”
1976	Michigan Technological University	BSEE

Professional Experience

From: 2008
To: 2010
Organization: Expressume, Inc / Montage Inc. – Milwaukee, WI
Title: Board Member
Summary: This company sells software for human resource recruiting. This company was sold in June 2019.

From: 2008
To: 2014
Organization: Faculte, Inc. – San Jose, CA
Title: Board Member
Summary: This company provided SaaS (Software as a Service) web video based communication products.

From: 2008
To: 2010
Organization: BitRail Networks Inc. – Miami, FL
Title: Founder, Board Member
Summary: This company sold computer networking solutions.

From: 2008
To: Present
Organization: Calumet Venture Management – Madison, WI
Title: Member
Summary: This company provides seed capital and management expertise to small companies.

From: 2006
To: 2015
Organization: BEEcube Inc. – Fremont, CA
Title: Founder, Board Member, Board Advisor
Summary: This company built high speed processing solutions. This company was sold to National Instruments, Inc. in Feb 2015.

Tim Arthur Williams, Ph.D.
Curriculum Vitae

From: 2006
To: 2015
Organization: Topaz Equity, LLC – Danville, CA
Title: Founder, Board Member
Summary: This is a private equity investment company. It owned AtomAMPD which develops, markets and sells software based network solutions.

From: 2004
To: Present
Organization: DoceoTech Inc. – Danville, CA
Title: Founder, Chairman
Summary: This was a training company that provides training for engineers in Wireless, Networking, and Telephony technologies. It is currently owned by Beach Technologies, LLC.

From: 2004
To: 2006
Organization: SiBEAM, Inc. – Sunnyvale, CA
Title: Founder, Chief Executive Officer
Summary: This is a fabless semiconductor company that develops high-speed wireless networking ICs. This company was sold to Silicon Image, Inc. in Apr 2011.

From: 2001
To: 2004
Organization: JetQue, Inc. – Danville, CA
Title: Founder, Chief Executive Officer
Summary: This company created messaging solutions for the mobile professional.

From: 1999
To: 2000
Organization: Atheros Communications, Palo Alto, CA
Title: Interim CEO, Advisory Board Member
Summary: This company builds wireless LAN ICs. Atheros became a public company in May 2004. (ATHR) This company was sold to QCOM in Jan 2011.

Tim Arthur Williams, Ph.D.
Curriculum Vitae

From: 1998
To: 2000
Organization: Picazo Communications, Inc. – San Jose, CA
Title: Chief Technology Officer, Advisory Board Member
Summary: This company built and sold software PBXs Telephony equipment using VoIP and Circuit Switched Technologies. The company was purchased by Intel.

From: 1996
To: Present
Organization: Beach Technologies, LLC – Danville, CA
Title: Chief Executive Officer
Summary: This is a consulting company that provides IP services. It owns DoceoTech LLC and Streaming Knowledge LLC, which perform the same services.

From: 1991
To: 1998
Organization: Wireless Access, Inc. – Santa Clara, CA
Title: Co-Founder, Chief Technical Officer, Vice President of Engineering, Vice President of Business Strategy
Summary: This was a startup company focusing on the Narrow Band PCS equipment market. The company developed the over the air protocols, the subscriber equipment and the ICs to deploy 2-way paging services. The company was sold to Glenarye Electronics.

From: 2014
To: 2021
Organization: Through Technology, LLC. – Chicago, IL
Title: Partner
Summary: This is a private equity investment company. It owns Through Technology Group, PTE LTD, which is registered in Singapore.

From: 1979
To: 1991
Organization: Motorola, Inc. – Austin, TX – Semiconductor Sector

Tim Arthur Williams, Ph.D.
Curriculum Vitae

Title: Sr. Engineer, Member Technical Staff, Sr. MTS
Summary: Business manager, project leader, and senior technical member of the teams which were responsible for product development of the following systems:

- ADPCM transcoder,
- ISDN U-reference point transceiver,
- CT-2 voice codec and channel modem,
- GSM voice codec and channel modem,
- TDMA voice codec and channel modem
- CDMA voice codec and channel modem, and
- Japanese Digital Cellular voice codec and channel modem.

From: 1976

To: 1979

Organization: Motorola Inc. - Chicago, IL - Communications Sector - Digital Voice Privacy Group

Title: Engineer

Summary: This group built the first commercial digitally encrypted two-way FM land mobile radio system.

Professional Certifications

▪ Patent Agent – U.S. Patent and Trademark Office #50,790 (Jan 2002)

Issued Patents

Patent	Date	Description
9,787,471	2017	Data Enciphering or Deciphering using a Hierarchical Assignment System
7,904,117	2011	Wireless Communication Device using Adaptive Beamforming
6,781,962	2004	Apparatus and Method for Stored Voice Message Control
6,600,481	2003	Data entry apparatus and method
6,088,457	2000	Method and apparatus for over the air programming a communication device
5,854,595	1998	Communications apparatus and method with a computer interchangeable integrated circuit card
5,557,642	1996	Direct conversion receiver for multiple protocols

Tim Arthur Williams, Ph.D.
Curriculum Vitae

5,428,638	1995	Method and apparatus for reducing power consumption in digital communications devices
5,345,406	1994	Bandpass sigma delta converter suitable for multiple protocols
5,101,344	1992	Data processor having split level control store
5,001,661	1991	Data processor with combined adaptive LMS and general multiplication functions
4,989,169	1991	Digital tone detector using a ratio of two demodulators of differing frequency
4,972,356	1990	Systolic IIR decimation filter
4,965,762	1990	Mixed size radix recoded multiplier
4,947,363	1990	Pipelined processor for implementing the least-mean-squares algorithm
4,876,542	1989	Multiple output oversampling A/D converter with each output containing data and noise
4,862,169	1989	Oversampled A/D converter using filtered, cascaded noise shaping modulators
4,843,585	1989	Pipelineable structure for efficient multiplication and accumulation operations
4,843,390	1989	Oversampled A/D converter having digital error correction
4,796,219	1989	Serial two's complement multiplier
4,737,925	1988	Method and apparatus for minimizing a memory table for use with nonlinear monotonic arithmetic functions
4,734,876	1988	Circuit for selecting one of a plurality of exponential values to a predetermined base to provide a maximum value
4,727,508	1988	Circuit for adding and/or subtracting numbers in logarithmic representation
4,722,067	1988	Method and apparatus for implementing modulo arithmetic calculations
4,682,302	1987	Logarithmic arithmetic logic unit
4,618,946	1986	Dual page memory system having storage elements which are selectively swapped between the pages
4,406,010	1983	Receiver for CVSD modulation with integral filtering
4,398,262	1983	Time multiplexed n-ordered digital filter

Patent Applications and Continuation Applications

Appl. Num.	Pub. Date	Description
20070037528	2007	Wireless Communication Device using Adaptive Beamforming
20040252679	2004	Stored Voice message Control Extensions
15/711,590	2017	Data Enciphering or Deciphering using a Hierarchical Assignment

Tim Arthur Williams, Ph.D.
Curriculum Vitae

Published Papers

6 Sept 2016 Putnam, Jonathan D. and Williams, Tim A., The Smallest Salable Patent-Practicing Unit (SSPPU): Theory and Evidence. Available at SSRN: <https://ssrn.com/abstract=2835617>

Litigation Experience

See attached.