

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

BRIGHT DATA LTD.,

Petitioner

v.

OXYLABS, UAB,

Patent Owner,

IPR2023-00840

U.S. Patent No. 10,965,770

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 claims 1-26 of U.S. Patent No. 10,965,770 to Vilcinskas, et *al.* (“the ’770 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 ’770 patent, its prosecution history, and the documents listed in the Petition’s Patent Owner’s List of Exhibits. 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 ’770 patent as of its filing date of September 11, 2020. I have been asked to provide my technical review, analysis, insights, and opinions regarding the ’770 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 '770 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 the investment 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 '770 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 AND INVALIDITY

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. Regarding the claim term "proxy rotator, the '770 patent defines a proxy rotator "is a part of the Service Provider chosen for a particular scraping query, to search and collect Infrastructure and is coupled with separate external components implementing specific assisting functionalities and is responsible for proxy control, rotation, maintenance, collecting statistical data, and

reporting.” EX.1001, 7:24-28. Accordingly, in my opinion, a “proxy rotator” means “is a part of the Service Provider Infrastructure and is coupled with separate external components implementing specific assisting functionalities and is responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting. A POSA would understand “proxy rotator” in the context of the ’770 Patent means “a system coupled with separate external components implementing a specific functionality and responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting.” ¶

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.

MY UNDERSTANDING OF ANTICIPATION

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 have been 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 '770 patent would have been an individual who, as of September 11, 2020, the filing date of the '770 patent, had either: a Master's Degree or higher in the field of electrical engineering, computer engineering, or computer science with a knowledge of Internet Communications and standards, or as of that time had a Bachelor's Degree in the same fields and two or more years of experience in Internet Communications and Standards. I exceeded that level of skill as of September 11, 2020.

31. Based on the foregoing, I believe that I am qualified to provide reliable opinions in the technical field of the '770 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 (September 11, 2020).

OVERVIEW OF THE '770 PATENT AND CHALLENGED CLAIMS

Background of the '770 Patent

32. The '770 patent relates to a “web crawling process” including “checking the parameters of a request received from a User's Device, adjusting the request parameters according to pre-established Scraping logic, selecting a Proxy according to the criteria of the pre-established Scraping logic, sending the adjusted request to the Target through the selected Proxy, checking metadata received from the Target, and forwarding the data to the User's device.” *See* Abstract. Fig. 1 of the '770 Patent (reproduced below) shows a system with a service provider

infrastructure 114 positioned between a user device 102 and target servers 116 and 118. Ex. 1001 at 8:59-9:16.

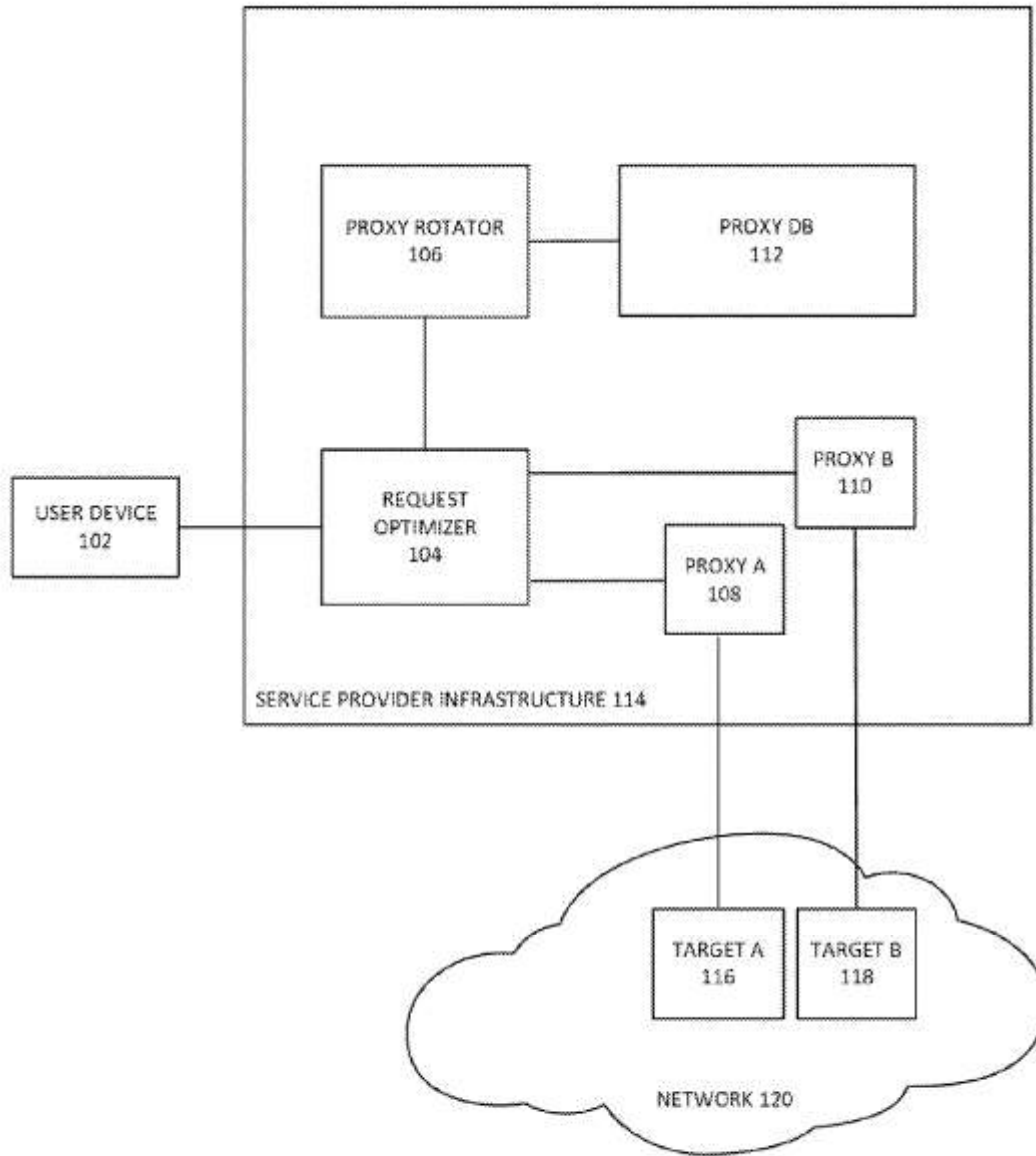


FIGURE 1

The user device 102 is disclosed as a device outside of the service provider infrastructure 114 implementing a browser 110 which may communicate a hypertext transfer protocol (HTTP) request for a web page stored on target 116 or target 118. *Id.* A request optimizer 104 is shown to accept requests from the User Device and implements the requests through the system. Using an analysis algorithm, the Request Optimizer can read the metadata of the request and adjust the request using different pre-established Scraping strategies to better scrape the Target. *Id.* at 7:12-23.

33. Optimizing requests may include having the request subjected to analysis and adjustment according to the selected scraping strategy. *Id.* at 12:5-9. The parameters of a strategy may also be adjusted if a reply received from the target is negative, where the parameters may include a URL address, a header, cookies, or the body of the request. *Id.* at 17:10-20.

34. FIG. 3A and 3B, reproduced below illustrate the sequence of a scraping request.

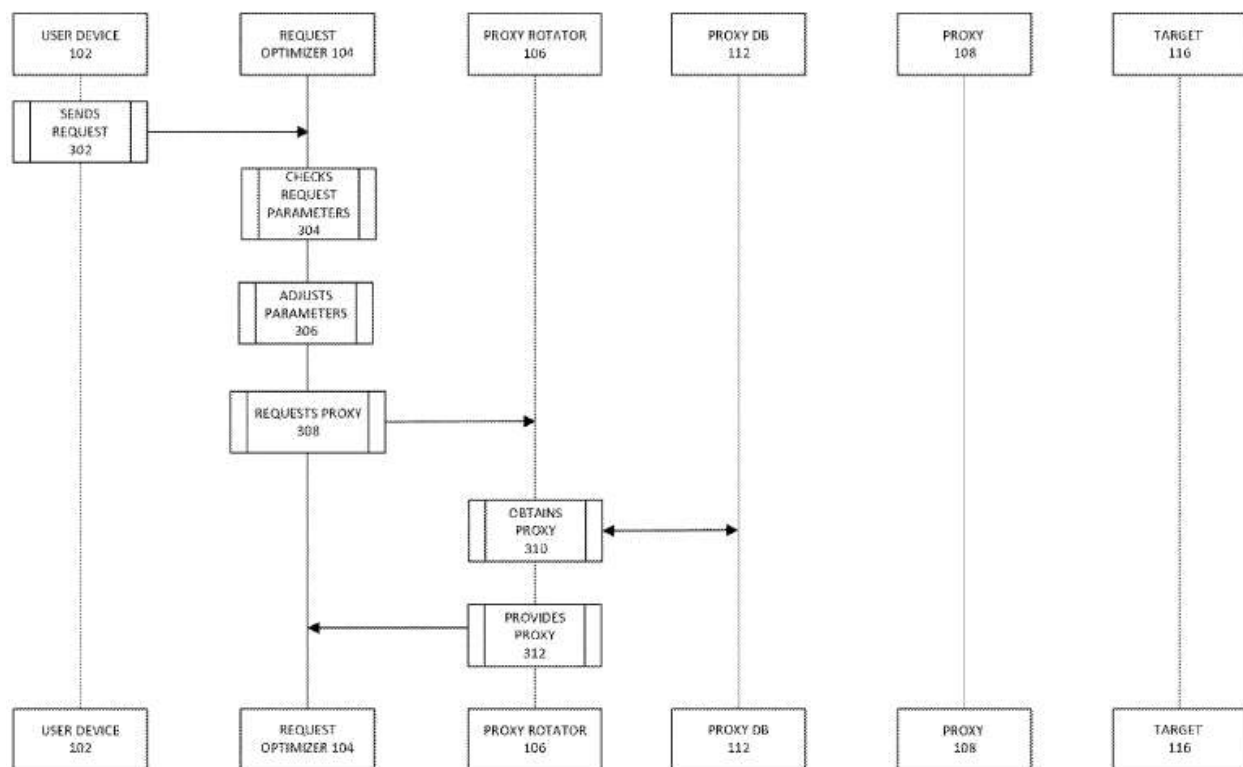


FIGURE 3 A

Ex. 1001 FIG. 3A

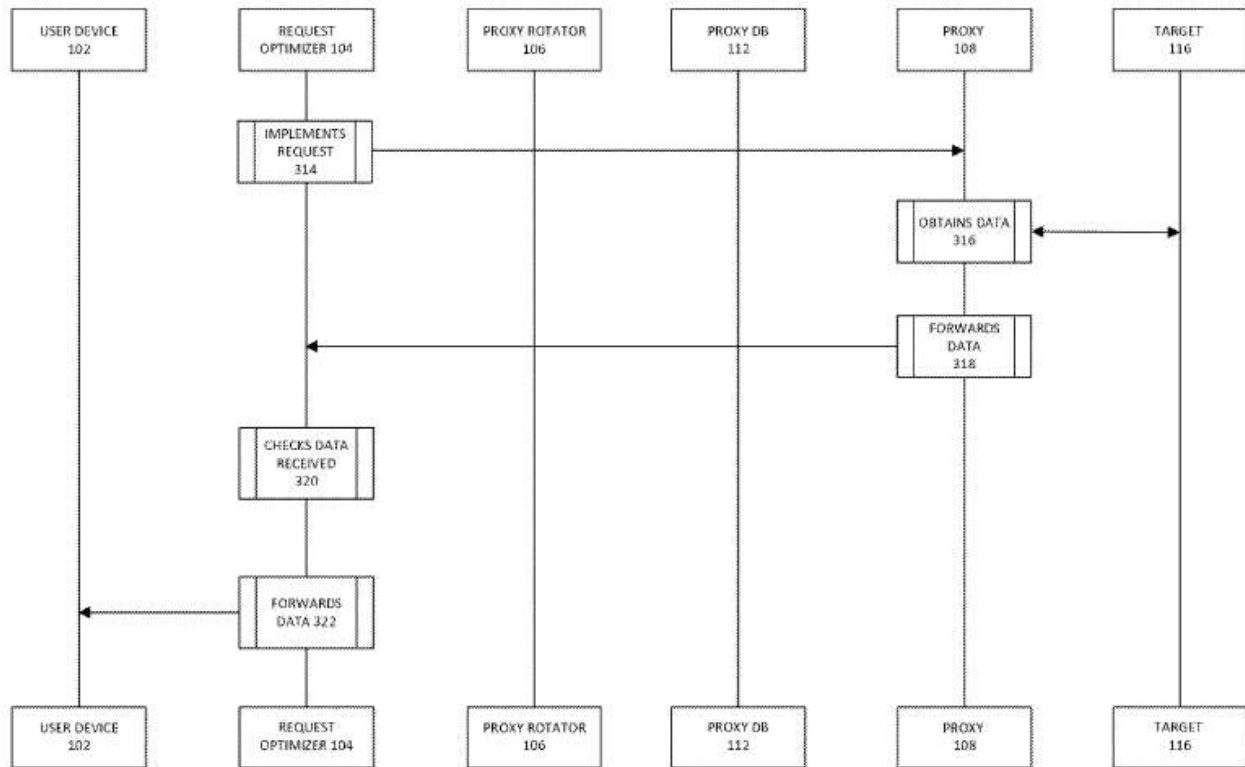


FIGURE 3 B

Ex. 1001 FIG. 3B

35. The '770 patent discloses selecting a scraping strategy and then identifying any discrepancies between the request from the user device and the selected scraping strategy. *Id.* at 16:62-65. The '770 further discloses an adjusting of parameters of the request according to the scraping strategy where the request optimizer reads metadata of the request and adjusts the request using different pre-established scraping strategies to better scrape the target. *Id.* at 7:16-19.

36. The adjusting may include a modification of the query received from the user device. *Id.* at 10:34-36. The request optimizer may also adjust parameters to increase efficiency where if a URL address is incorrect the request optimizer may correct the URL address. *Id.* at 11:17-23. FIG. 5 of the '770 patent (reproduced below) illustrates the request being received by the request optimizer 104 that includes a variety of scraping strategies that may be used to adjust the request to achieve better scraping results. Such strategies may include introducing or adjusting custom headers, identifying geographical locations, time zones, or language used. *Id.* at 15:56-60.

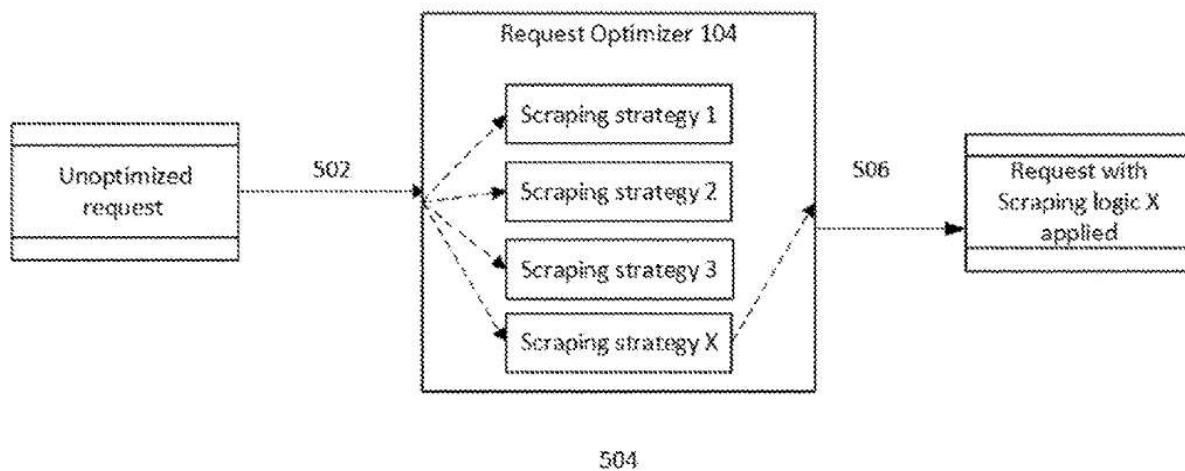


Figure 5

37. The request optimizer includes a scraping strategy 504 where the parameters or policies may be optimized for a particular type of request or for a particular target. *Id.* at 8:27-29. The request optimizer may apply the optimizations

based on the original request. *Id.* at 8:27-31. Further, scraping strategies may be developed by analyzing a target server's parameters and the target server's response data. *Id.* at 9:21-25. Such strategies may include selecting a proxy based on an appropriate geographic location. *Id.* at 9:43-44.

38. Once the request parameters are adjusted the request optimizer may ask the proxy rotator for a particular proxy that is suitable to implement the user's request. *Id.* at 12:30-36. The '770 patent discloses that the sending of the adjusted request to a target includes selecting the proxy according to the scraping strategy. *Id.* at 18:2-4. Further, as the proxy is selected according to the scraping strategy the selection includes geolocation settings, proxy type, proxy availability, or any such combination. *Id.* at 17:25-27. The '770 patent then discloses that a reply from the target may be received. *Id.* at 18:3-4.

39. The underlying application for the '770 patent was filed on September 11, 2020, with 26 claims. Ex. 1002 at 94-97. Applicant concurrently submitted a Track One request for prioritized examination, which was granted on October 8, 2020.

40. The United States Patent and Trademark Office ("USPTO" or "Office") issued an Office action on December 9, 2020 objecting to the claims due to various informalities but indicated that all claims would be allowable if rewritten or amended to overcome the informalities objections.

41. On December 28, 2020 Applicant submitted a response amending claims to overcome the informalities objections.

42. On December 31, 2020 the Office held an interview with Applicant to discuss further amendments to claim 18 to overcome an informalities objection. In the interview the Applicant agreed to an Examiner's Amendment.

43. The Office issued a Notice of Allowance on January 12, 2021.

44. The '770 Patent issued on March 30, 2021.

The Claims

45. The '770 patent has twenty-six claims, including three independent claims: Nos. 1, 16, and 19. Claim 1 recites a method as described below:

1. A method for data extraction and gathering from a web server comprising:
comparing parameters of a request received from a user device with a list of scraping strategies;
selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request;
identifying discrepancies between the request and the scraping strategy;
adjusting the parameters of the request according to the scraping strategy to form an adjusted request;
sending the adjusted request to a target through a proxy selected according to the scraping strategy;
receiving a reply from the target.

46. Dependent claims 2-15 depend from claim 1 and recite additional limitations.

47. Claim 16 recites a system as described below:

16. A system for data extraction and gathering from a network comprising at least:
a proxy rotator operable to provide a plurality of proxies over the network; and
a request optimizer operable to compare parameters of a request received from a user device with a list of scraping strategies, to select a scraping strategy from the list of scraping strategies a scraping strategy that corresponds to the parameters of the request, to identify discrepancies between the request and the scraping strategy, to adjust the parameters of the request according to the scraping strategy to form an adjusted request, and to send the adjusted request to a target through a proxy selected from a proxy database by the proxy rotator according to the scraping strategy.

48. Dependent claims 17 and 18 depend from claim 16 and recite additional limitations.

49. Claim 19 recites a non-transitory computer readable medium as described below:

19. A non-transitory computer readable medium for data extraction and gathering from a web server comprising instructions which, when executed by a computing device within a service provider infrastructure, cause the computing device to perform:
comparing parameters of a request received from a user device with a list of scraping strategies;
selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request;
identifying discrepancies between the request and the scraping strategy;
adjusting the parameters of the request according to the scraping strategy to form an adjusted request;
sending the adjusted request to a target through a proxy selected according to the scraping strategy;
receiving a reply from the target.

50. Dependent claims 20-26 depend from claim 19 and recite additional limitations.

51. The particular claims at issue are claims 1-26.

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

Applicant Admitted Prior Art

52. The “Background” section of the ’770 patent describes how a proxy, also called a server or gateway, functions. For example, a proxy receives a request

from a user and creates a proxy tunnel toward a remote server, forwarding the request to the remote server through the tunnel, receiving the response from the remote server and forwarding the response back to the user. Ex. 1001 at 1:27-37.

53. The '770 patent further describes a proxy server as:

A proxy server is basically a computer on the internet with its own IP address that the client's computer "knows". When a client sends a web request, the request is sent to the proxy server first. The proxy server then makes the web request on the client's behalf, collects the response from the web server, and forwards the web page data so that the client can see the page in the browser or in another manner e.g. as data written in a file on the disk for bulk processing by data aggregation and analysis tools. When the proxy server forwards the web requests, the proxy server can make changes to the data but yet provide the data requested. For example, a proxy server can change the client's IP address, so the web server is not provided the geographical location of the client. Also, a proxy server can block access to certain web pages, based on IP address or domain name. *Id.* at 1:58-2:5.

54. The '770 cites state-of-the-art features in "modern proxy servers" where proxy servers can change the IP address and other identifying information that the web request may contain. Proxy servers can also be used to control internet usage, improve browsing speeds using cache and send the saved copy to the client. *Id.* at 2:6-24.

55. The background section further describes using proxies for web scraping, data mining and to bypass certain Internet restrictions by enabling a user

to request the content through a proxy server instead of accessing the content directly. *Id.* at 2:25-36.

56. The '770 patent describes use of a “web crawler” which may be used for “browsing the web for the requested info and retrieving raw Hypertext Markup Language (HTML) code of the accessed web sites” as well as “indexing of web content by search engines”. *Id.* at 2:52-56. The '770 Patent also describes “web scraping” used for fetching the accessed web pages and parsing the content for the requested data. *Id.* at 2:57-59.

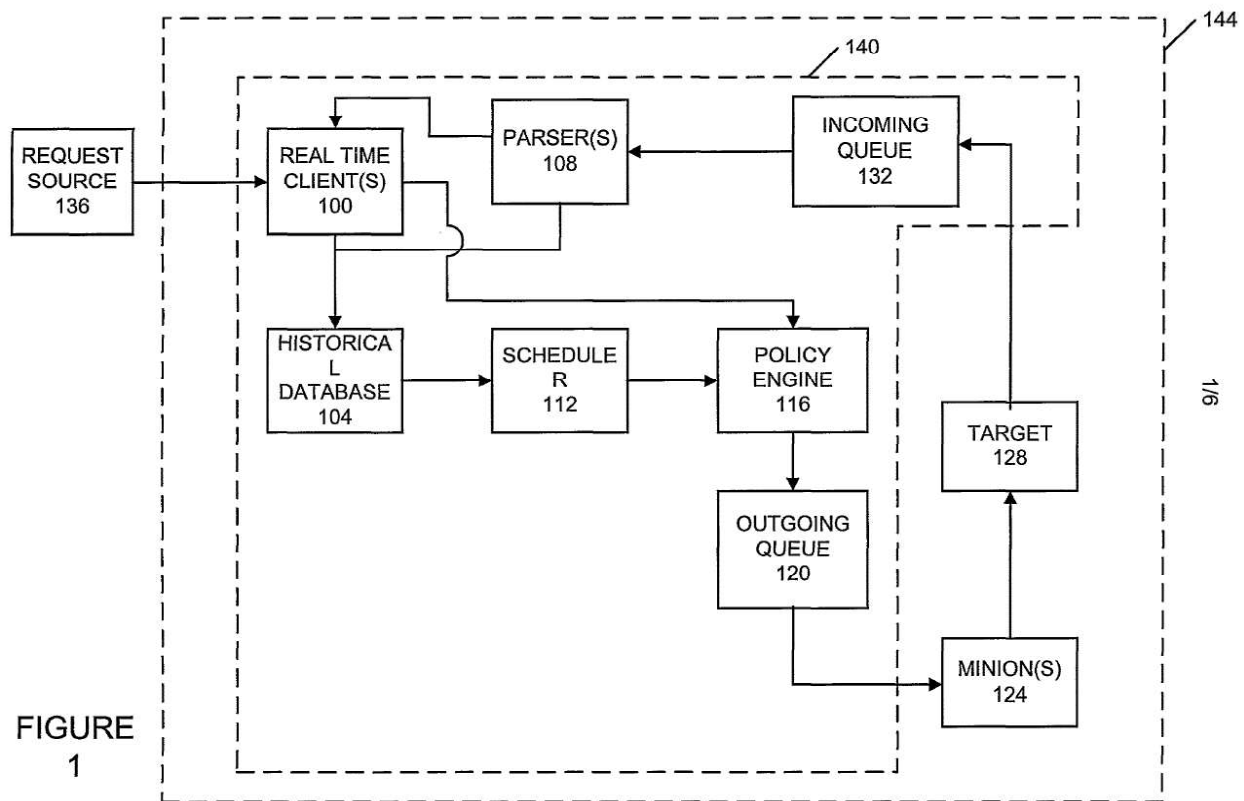
Cousins (Ex. 1004)

57. International Publication Number WO 2001/50320 to Cousins, et al. (“Cousins,” Ex. 1004) was published on July 12, 2001, more than nineteen years before the September 11, 2020, filing date of the '770 Patent. Thus, I understand that Cousins qualifies as prior art under 35 U.S.C. § 102(a)(1).

58. Additionally, in my opinion, Cousins is analogous art to the '770 Patent at least because Cousins and the '770 Patent both relate to the crawling and scraping of information on the web using a distributed network.

59. Cousins is titled “Minimal Impact Crawler.” Ex. 1004 at code (54). Cousins describes a system and method for a minimal impact crawler for searching and retrieving information based on policies for a target site. The crawler employs

a plurality of minions to retrieve crawl requests and crawl the target web sites. *See id.* at Abstract, code (57). *See* also Fig. 1 reproduced below.



60. Cousins discloses the use of a policy engine that receives a user request and assembles policies for the target site. Ex. 1004 at 3:13-15. The requests may be ordered by their priority to optimize data gathering in which the policy engine examines the crawl request in light of priority-setting rules. *Id.* at 4:17-22. For example, Cousins discloses an efficiency scraping policy to consolidate duplicate requests to the same target, where “two client engines 100 ... could desire the same URL, but desire different information from the URL” and that “the policy engine 116 combines the multiple requests 400 into a single crawl request

400 comprising one URL field 408, and two client and parser ID fields 312, 404.”

Id. at 11:18-24.

61. Cousins also discloses a set of policies that include:

- “Do not crawl if web site is shut down
- Only crawl if latency of connection is less than a predetermined value
- Crawl only on off-peak times
- Crawl anytime (for a ‘friendly’ web site)
- Never crawl
- Crawl only after a predetermined time period has elapsed”

Id. at 10:25-32.

62. Cousins further discloses the use of minions, *e.g.*, proxies, such that “[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data.” *Id.* at 4:26-27. Further, the minions may be “located topologically as close as possible to the target site’s server.” *Id.* at 15:1-2. Furthermore, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies. *Id.* at 16:28-29.

63. In addition, “[b]y providing unique minions 124 to crawl web sites for which they are custom designed, the crawls can be much more efficient than a generic crawl, thus minimizing the impact of the crawl on the target web site” and

“[i]n a further embodiment, a minion 124 may be instructed to use cookies or ignore them based upon the categorization policy to make the crawler's data gathering as benign as possible.” *Id.* at 16:10-14.

RFC 7231 (Ex. 1005)

64. Request for Comments 7231, titled “Hypertext Transfer Protocol (HTTP/1.1): Semantics and Content,” was published by the Internet Engineering Task Force (“IETF”) Standards Track under Editor R. Fielding in June 2014. (“RFC 7231,” Ex. 1005). As discussed below, in my opinion, RFC 7231 qualifies as a printed publication that was publicly accessible before the September 11, 2020, filing date of the ’770 Patent. *See* Ex. 1001 at code (22). Thus, I understand that RFC 7231 qualifies as prior art under 35 U.S.C. § 102(a)(1).

65. The IETF is an international non-profit standards organization that develops technical standards for Internet protocols. *See About the IETF*, <https://www.ietf.org/about/introduction/#mission> (last visited April 11, 2023). The IETF is a volunteer organization that develops and publishes Internet standards that are widely published and used throughout the industry. *See id.* The IETF makes the majority of its developed standards publicly available. *See Open records*, IETF, <https://www.ietf.org/about/open-records/> (last visited April 11, 2023). RFCs are the basic publication series for the IETF and are developed by Working Groups

within the IETF. In my opinion, A POSA would have been very familiar with RFC 7231 and the Internet standards explained within.

66. RFC 7231 was published in June 2014 and approved for publication by the Internet Engineering Steering Group of the IETF. RFC 7231 is freely available for download at <https://www.rfc-editor.org/rfc/rfc7231>. RFC 7231 is intentionally disseminated per the IETF's open records policy.

67. In my opinion, RFC 7231 is analogous art to the '770 patent at least because both RFC 7231 and the '770 patent relate to accessing web pages utilizing the Hypertext Transfer Protocol. As further discussed below, RFC 7231 and the '770 patent both involve using HTTP and metadata. *See, e.g.*, Ex. 1001 at 3:11-22; 7:18-20; 7:48-50; 12:56-59; 13:60-67.

68. RFC 7231 explains that “[r]epresentation header fields provide metadata about the representation” and “[w]hen a message includes a payload body, the representation header fields describe how to interpret the representation data enclosed in the payload body.” Ex. 1005 at 8.

Taylor (Ex. 1006)

69. U.S. Publication No. 2017/0180382 to Taylor et al. (“Taylor,” Ex. 1006) was published on June 22, 2017, more than one year before the filing date of

the '770 Patent.¹ Ex. 1001 at code (22). Thus, I understand that Taylor qualifies as prior art under 35 U.S.C. § 102(a)(1).

70. In my opinion, Taylor is analogous art to the '770 Patent at least because Taylor and the '770 Patent both relate to management of proxy servers within a service provider infrastructure. Further, Taylor and the '770 Patent both relate to the transmission of digital information as both Taylor and the '770 Patent are similarly classified in Intl. Cl. "H04L 29" transmission of digital information relating to arrangements, apparatus, circuits, or systems. Ex. 1001 at code (51); Ex. 1006 at code (51); Ex. 1007 at code (51). *See* also Fig. 1 reproduced below.

¹ Taylor issued as U.S. Patent No. 10,348,687 on July 9, 2019. Ex. 1007.

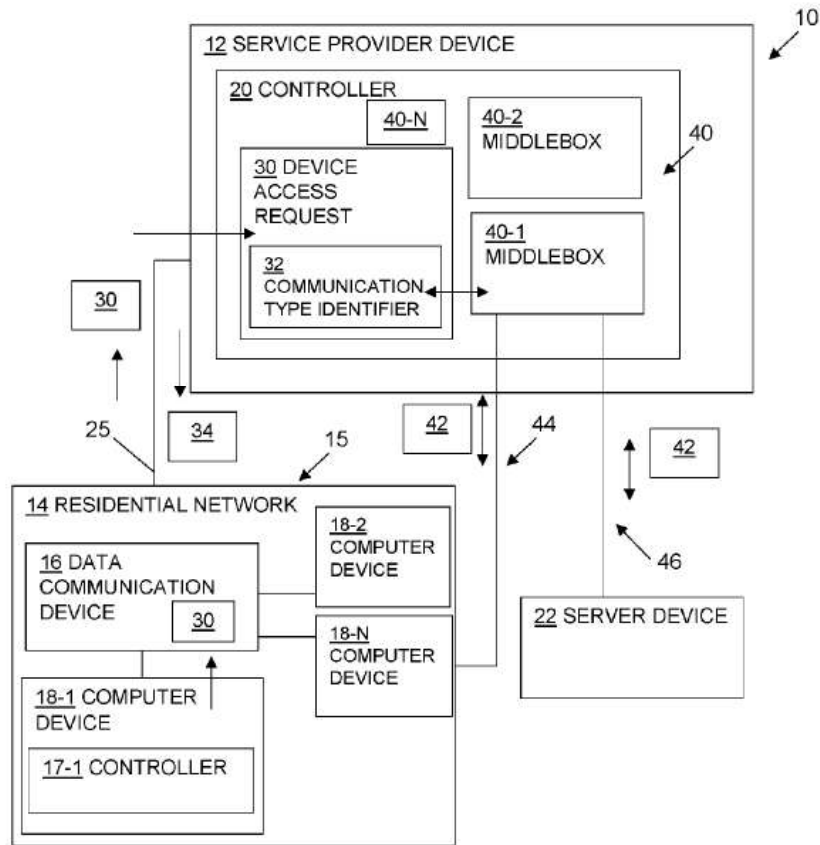


FIG. 1

71. Taylor is titled “Method and Apparatus for Using Software Defined Networking and Network Function Virtualization to Secure Residential Networks.” Ex. 1006 at code (54). Taylor discloses a service provider device that includes a controller configured to receive a request from a client device and establishing a secure communication channel between the client device and a middlebox. *Id.* at Abstract, code (57). Taylor further discloses selecting a middlebox proxy based on a correspondence of the destination IP address and a domain name server database. *Id.* at [0044].

RFC 7540 (Ex. 1008)

72. Request for Comments 7540, titled “Hypertext Transfer Protocol – HTTP/2” was published by the Internet Engineering Task Force (“IETF”) Standards Track under Editor M. Thomson in May 2015. (“RFC 7540,” Ex. 1008). As discussed below, in my opinion, RFC 7540 qualifies as a printed publication that was publicly accessible before the September 11, 2020, filing date of the ’770 Patent. *See* Ex. 1001 at code (22). Thus, I understand that RFC 7540 qualifies as prior art under 35 U.S.C. § 102(a)(1).

73. As previously discussed, the IETF is an international non-profit standards organization that develops technical standards for Internet protocols. RFCs are the basic publication series for the IETF and are developed by Working Groups within the IETF. In my opinion, A POSA would have been very familiar with RFC 7540 and the Internet standards explained within.

74. RFC 7540 was published in May 2015 and approved for publication by the Internet Engineering Steering Group of the IETF. RFC 7540 is freely available for download at <https://www.rfc-editor.org/rfc/rfc7540>. RFC 7540 is intentionally disseminated per the IETF’s open records policy.

75. In my opinion, RFC 7540 is analogous art to the ’770 patent at least because both RFC 7540 and the ’770 patent relate to accessing web pages utilizing the Hypertext Transfer Protocol. As further discussed below, RFC 7540 and the

'770 patent both involve using HTTP and maintaining connections. *See, e.g.*, Ex. 1001 at 3:11-22; 3:32-34; 15:19-22; 15:36-38.

76. RFC 7540 explains that “HTTP/2 connections are persistent” and that “[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or until the server closes the connection.” Ex. 1008 at 65.

SQUID – THE DEFINITIVE GUIDE (Ex. 1009)

77. The open source Squid proxy is documented in Wessels, Ex. 1009 hereto (“Squid”). The book was published by a major publisher in the field and has the usual conventional markers (copyright, ISBN, etc.) indicating a reliable publication date in 2004.

78. Squid is an “open source caching proxy for the Web.” Ex. 1009 at 1. Squid further discloses “both a proxy and a cache” where “[a]s a proxy, Squid is an intermediary in a web transaction” where “[i]t accepts a request from a client, processes that request, and then forwards the request to the origin server.” *Id.* And, as “a cache, Squid stores recently retrieved web content for possible reuse later” and where “[s]ubsequent requests for the same content may be served from the cache, rather than contacting the origin server again.” *Id.*

SQUID – WEBSITE (Ex. 1010)

79. The Squid website from a web.archive.org capture on September 20, 2018 discusses “Rotating the Squid outbounds IPs” and presents sample code for “Splitting uploads and downloads between two ISP links.” Ex. 1010 at 1.

GROUND 1: ANTICIPATION OF CLAIMS

1, 4-6, 9-14, 16, 17, 19, and 22-26

BY COUSINS (Ex. 1004)

80. In my opinion, Cousins discloses each of the limitations of independent claims 1, 16, and 19 and dependent claims 4, 5, 6, 9, 10, 11, 12, 13, 14, 17, 22, 23, 24, 25, and 26, of the ’770 patent. Thus, in my opinion, claims 1, 4-6, 9-14, 16, 17, 19, and 22-26 of the ’770 are anticipated by Cousins.

Claim 1

81. **“A method for data extraction and gathering from a web server comprising: ...”** In my opinion, to the extent the preamble is limiting, Cousins discloses this claim limitation. . Cousins discloses “[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network.” Ex. 1004 at Abstract. Fig. 1 of Cousins shows a web server, *i.e.*, target 128, where “[t]he crawler (144) applies the policies to schedule crawls (144) of the target site (128) and stores data retrieved from the crawl (144).” *Id.*

82. Further, “[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention.” *Id.* at 6:8-10. “As crawlers are computer-driven, they can access large amounts of data very quickly, and they can perform crawls continuously.” *Id.* at 1:20-21.

83. **“... comparing parameters of a request received from a user device with a list of scraping strategies; ...”** In my opinion, Cousins discloses this claim limitation. Cousins discloses “[a] request source [that] is any computer process that requests information from a data address” where “a user will access the crawler’s web site through the user’s Internet connection, and then input a request for a specific item into a form provided by the crawler’s web site” *Id.* at 6:15-24. Cousins discloses both “scheduled” and “user-demanded crawl[s].” *Id.* at 3:26. In my opinion Cousins disclosure of a user’s access through an Internet connection, inputting a request for a specific item into a form, as well as the creation of a user-demanded crawl discloses a user request from a user device.

84. Cousins states that the parameters of a request include a request for a specific item, for example, that “a policy engine is provided that receives a request for a specific item and assembles policies for the target site containing information about the specific item.” Ex. 1004 at 3:13-15. In addition, Cousins discloses comparing the request parameters with lists of policies, *e.g.*, scraping strategies,

that includes where “[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc.” *Id.* at 3:13-16. Cousins further discusses additional possible scraping strategies for a target site that may include best practices and/or efficiency based criteria such as “[d]o not crawl between 9:00 a.m. and 5:00 p.m.” or “[d]o not crawl if web site is shut down”, or “[n]ever crawl.” *Id.* at 3:17-18. Further, Cousins discloses that such “policies allow the crawler to tailor its crawl to the requirements and preferences of the crawled site.” *Id.* at 3:18-19.

85. **“... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request; ...”** In my opinion, Cousins discloses this claim limitation. Cousins describes one of the strategies as creating a more efficient scrape where “in a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL” and “prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128” such that “[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404.” *Id.* at 11:17-24. Further, “by coalescing multiple requests from different request sources 136 for

similar data into a single crawl, the crawler 144 of the present invention provides lower overhead, lower latency and globally higher throughput. It also decreases the crawling burden on the target site 128.” *Id.* at 11:27-30.

86. Another strategy is time based so that the most current information is obtained, where “[f]or example, a policy for a target site 128 may be ‘crawl if the data is older than ten hours’” where “[i]n this example, the client engine 100 passes a request to the policy engine 116, as discussed below, which then checks a policy stored for the target site 128 and the date/time stamp of the data in the historical database 104 that would satisfy the request to determine whether the data in the database 104 may be used to satisfy the request.” *Id.* at 8:16-21.

87. Cousins also describes a general selecting of a scraping strategy where “a database may be created for time of day, or target site, or connectivity loss, or site closure” where the policy engine “will examine the time of day database to determine if a policy is applicable to the crawl request based on the current time of day” and that “[t]he policy engine 116 will examine all of the databases to build a list of all policies that apply to a particular request for data.” *Id.* at 10:9-12.

88. **“... identifying discrepancies between the request and the scraping strategy; ...”** In my opinion, Cousins discloses this claim limitation. The ’770 patent does not define what may be classified as a discrepancy. The Merriam-Webster dictionary defines a discrepancy as “1: the quality or state of disagreeing

or being at variance” and “2: an instance of disagreeing or being at variance.”² In the context of the ’770 Patent, a POSA would have understood identifying a discrepancy as identifying a disagreement or some difference between the parameters of the received request and the parameters of a request according to the selected scraping strategy.

89. An example of a strategy including a discrepancy was discussed above, where there are two separate received requests with both requests for the same target site. Instead of executing the two separate requests the strategy “combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404.” *Id.* at 11:17-24.

90. Cousins describes another example discrepancy pertaining to the time when data is to be crawled from a target server where a request is made to retrieve “a current event news story on a web site,” but a scraping strategy dictates that a recently executed crawl already has retrieved the requested data making a new crawl inefficient. For example, “a request is for a current event news story on a web site, the window may require that the data have been crawled within the last four hours,” but “[f]or data that changes less frequently, such as a description of an archived news story, the window may permit data to be retrieved from the database

² <https://www.merriam-webster.com/dictionary/discrepancy>, accessed on April 11, 2023. The same website from May 16, 2020 on web.archive.org contains the same definition.

104 if it has been crawled within the last seven days or longer” and further “[i]n an embodiment in which the crawler 144 verifies the validity of data in the database 104, the policy engine 116 preferably stores the policies governing when a site 128 should be crawled.” *Id.* at 8:10-16.

91. Cousins further explains that “[f]or example, item information and seller information do not change during the life of an auction” and “[t]herefore, the initial crawl can provide this information, and then no further crawl is necessary unless a user request for price updates is received.” *Id.* at 4:9-11.

92. Cousins also describes a discrepancy regarding the priority of competing requests. For example, two requests arrive both claiming priority, where “for example, if a first request to a target site 128 is determined to be appropriate for transmission, and is a user-demand request, and a second request to a different target site 128 is also determined to be acceptable but is from an automated discovery crawl, the user demand request is given a higher priority than the request from the automated discovery crawl” where “[t]he use of this dynamic priority setting greatly improves the efficiency and responsiveness of the crawler 144.” *Id.* at 12:21-26.

93. **“... adjusting the parameters of the request according to the scraping strategy to form an adjusted request; ...”** In my opinion, Cousins discloses this claim limitation. Cousins describes adjusting a variety of request

parameters to form an adjusted request. For example, Cousins also describes adjusting parameters by consolidating multiple requests as discussed above where “two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL” and “prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128” such that “[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404.” *Id.* at 11:17-24.

94. Cousins also describes adjusting the use policy parameters of a request. For example, “[i]f a crawl request 400 is rescheduled by the scheduler 112, upon being received again by the policy engine 116, the policy engine 116 builds new policies for the rescheduled request to account for any new policies that may apply to the crawl request 400.” *Id.* at 11:6-9. In addition, “[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID[logical user id³] 300 to the policy engine 116 to generate policies for the request for data.” *Id.* at 8:22-23.

³ “An LUID is a data structure used to store and organize data used by the crawler 144 and to uniquely identify each auction monitored by the crawler 144 of the

95. **“... sending the adjusted request to a target through a proxy selected according to the scraping strategy; ...”** In my opinion, Cousins discloses this claim limitation. As previously discussed, Cousins discloses scraping strategies for a target site that include best practices and/or efficiency-based criteria. Cousins discloses, due to network congestion or a network failure, the selection of a proxy is determined by the scraping strategy where “a target site can be accessed from multiple different pathways” where “if one minion is unable to crawl a target site, because of congestion or network failures, a different minion located elsewhere in the Internet can still access the target site and provide the requested crawl.” *Id.* at 4:33-5:3. Further, Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” *Id.* at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. Squid discloses the uses of proxies, referred to as redirectors, that is “an external process that rewrites URIs from client requests” where “Squid fetches the new URI automatically, as though the client originally requested it.” Ex. 1009 at 185.

present invention. ... The LUID also has a unique identification number 312 used by the crawler 144 to internally differentiate different auctions, or other collections of data. The use of LUIDs 300 to track data stored on target sites 128 provides a high degree of flexibility to the crawler 144.” See page 7 at 1:12.

96. “... **receiving a reply from the target.**” In my opinion, Cousins discloses this claim limitation. As discussed in the applicant admitted prior art section, the ’770 patent describes how the then current proxy systems functioned, including that “a proxy operates in the following manner: i) receives a request from a user in a public network, or private network separated by a firewall, to a remote server that contains the desired content; ii) creates a proxy tunnel toward the remote server; iii) forwards the request to the remote server through the established tunnel; iv) receives the response from the remote server; and v) forwards the response back to the user.” Ex. 1001 at 1:30-37.

97. Further, Cousins discloses receiving a reply from the target where “[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data.” Ex. 1004 at 4:26-27. Also, Cousins describes the retrieval of data from the target where “[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention” and that “[a] crawl is defined as the act of fetching a collection of data, typically a web page, from the target site 128, or crawler.” *Id.* at 6:8-11. Further, Cousins also discloses that “policies are generated by an operator of the crawler 144 based on data received from the target site 128.” *Id.* at 10:1-2.

98. In addition, Cousins discloses receiving a reply from the target where “[t]he client engine 100 associates the LUID with a client engine identification and a parser identification” and “[t]his allows the crawler 144 to identify which parser 108 to use when the collection of data is retrieved from the target site 128, and to which client engine to route the requested data.” Ex. 1004 at 7:16-19.

99. Thus, for the reasons discussed above, claim 1 is anticipated by Cousins.

Claim 16

100. **“A system for data extraction and gathering from a network comprising at least: ...”** In my opinion, to the extent the preamble is limiting, Cousins discloses this claim limitation. Fig. 1 of Cousins shows a web server, *i.e.*, target 128, and in Fig. 7, *i.e.*, target site server 708 in which “crawlers have been developed ... that provide data gathering, searching, and retrieval services to allow users or Internet-related applications to locate desired data from the millions of web sites.” *Id.* at 1:14-16. Further, Cousins discloses “[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network.” *Id.* at Abstract.

101. Further, “[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in accordance with the present invention.” *Id.* at 6:8-10. “As crawlers are computer-

driven, they can access large amounts of data very quickly, and they can perform crawls continuously.” *Id.* at 1:20-21.

102. “... a proxy rotator operable to provide a plurality of proxies over the network; ...” In my opinion, Cousins discloses this claim limitation. As discussed above, the “proxy rotator” is a system coupled with separate external components implementing a specific functionality and responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting.. This meaning is based on the ’770 patent description that a proxy rotator “is a part of the Service Provider Infrastructure and is coupled with separate external components implementing specific assisting functionalities and is responsible for proxy control, rotation, maintenance, collecting statistical data, and reporting.” Ex. 1001 at 7:24-28. Cousins discloses a policy engine 116 that is responsible for the control of minions, *e.g.*, the proxies, where “[a] policy engine (116) is provided that receives a request for a specific item and assembles policies for the target site ... [where] [t]he policies are rules that determine the crawl (144) of a target site (128) ... [where] [t]he crawler (144) employs a plurality of minions (144) to retrieve crawl (144) requests and crawl (144) the target web sites (128) to obtain the necessary data.” Ex. 1004 at Abstract. Further, “the policy engine 116 decides not only whether a request may be transmitted at that time for a given target site 128 but

also decides the priority between competing requests for different target sites 128.”

Id. at 12:18-20.

103. Further, Cousins describes that the proxies, *e.g.*, minions, may be strategically located, for example, “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” *Id.* at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses.” *Id.* at 16:28-29. The policy engine may also rotate proxy requests through the use of multiple minions where “a target site can be accessed from multiple different pathways” where “a different minion located elsewhere in the Internet can still access the target site and provide the requested crawl.” *Id.* at 4:33-5:3.

104. Cousins also discloses where the policy engine can utilize multiple minions such that “a target site can be accessed from multiple different pathways ... if one minion is unable to crawl a target site ... a different minion located elsewhere in the Internet can still access the target site.” *Id.* at 4:33-5:2.

105. **“... a request optimizer operable to compare parameters of a request received from a user device with a list of scraping strategies, ...”** As discussed above, the ’770 patent describes a request optimizer as “a next generation proxy” *accepting requests* from the User Device and *implements* the requests through the system. Using an analysis algorithm, the Request Optimizer

can read the metadata of the request and adjust the request using different pre-established Scraping strategies to better scrape the Target. Ex. 1001 at 7:12-23.

106. Cousins discloses that the policy engine 116 *accepts requests and implements* the requests through the use of a crawler and minions where the “policy engine (116) is provided that receives a request for a specific item and assembles policies for the target site ... [where] policies are rules ... [where] the crawler (144) applies the policies to schedule crawls (144) ... [where] [t]he crawler (144) employs a plurality of minions (144) to retrieve crawl (144) requests and crawl (144) the target web sites (128) to obtain the necessary data.” *Id.* at Abstract, emphasis added.

107. Cousins discloses that the policy engine 116 receives a request where the “policy engine is provided that receives a request for a specific item and assembles polices for the target site containing information about the specific item.” Ex. 1004 at 3:13-15.

108. Cousins discloses comparing the request parameters with lists of policies, *e.g.*, scraping strategies, that includes where “[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc.” *Id.* at 3:13-16. Cousins further discusses additional possible scraping strategies for a target site that may include best practices and/or efficiency based criteria such as “[d]o not crawl between 9:00 a.m. and 5:00 p.m.” or “[d]o

not crawl if web site is shut down”, or “[n]ever crawl.” *Id.* at 3:17-18. Further, Cousins discloses that such “policies allow the crawler to tailor its crawl to the requirements and preferences of the crawled site.” *Id.* at 3:18-19.

109. Thus, in my opinion, for these reasons and the reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

110. **“... to select a scraping strategy from the list of scraping strategies a scraping strategy that corresponds to the parameters of the request, ...”** This claim limitation mirrors a limitation present in claim 1. In my opinion, all of the analysis/steps of claim 1 are performed by the policy engine. As discussed, Cousins discloses that the policy engine 116 “that receives a request for a specific item and assembles policies for the target site that contain information about the specific item.” *Id.* at 3:13-15. Thus, Cousins discloses this claim limitation.

111. **“... to identify discrepancies between the request and the scraping strategy, ...”** This claim limitation mirrors a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

112. **“... to adjust the parameters of the request according to the scraping strategy to form an adjusted request, ...”** This claim limitation mirrors

a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

113. **“... to send the adjusted request to a target through a proxy selected from a proxy database by the proxy rotator according to the scraping strategy.** In my opinion, Cousins discloses this claim limitation. As previously discussed, Cousins discloses scraping strategies for a target site that include best practices and/or efficiency-based criteria. Cousins discloses, due to network congestion or a network failure, the selection of a proxy is determined by the scraping strategy where “a target site can be accessed from multiple different pathways” where “if one minion is unable to crawl a target site, because of congestion or network failures, a different minion located elsewhere in the Internet can still access the target site and provide the requested crawl.” *Id.* at 4:33-5:3. Further, Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” *Id.* at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29.

114. Further, Cousins discloses that the data of a particular crawl of a selected target website is stored in a database, for example, “after a crawl, either scheduled or by request, the crawler stores data retrieved from the crawl into a historical database” and “when a scheduled or user-demanded crawl is to occur, the

crawler first checks the historical database to determine whether the request can be satisfied from the data stored in the database.” Ex. 1004 at 3:24-26. If the request cannot be satisfied from the stored data, then a new crawl is requested based on the minion selection in the database. *Id.* at 4:1-3.

115. Thus, for the reasons discussed above, claim 16 is anticipated by Cousins.

Claim 19

116. **“A non-transitory computer readable medium for data extraction and gathering from a web server comprising instructions which, when executed by a computing within a service provider infrastructure, cause the computing device to perform: ...”** In my opinion, to the extent the preamble is limiting, Cousins discloses this claim limitation. Fig. 1 of Cousins shows a web server, *i.e.*, target 128, and in Fig. 7, *i.e.*, target site server 708 in which “crawlers have been developed ... that provide data gathering, searching, and retrieval services to allow users or Internet-related applications to locate desired data from the millions of web sites.” *Id.* at 1:14-16. Further, Cousins discloses “[a] system and method are provided that provide a minimal impact crawler (144) for searching and retrieving information on a distributed network.” *Id.* at Abstract.

117. Further, “[t]he crawler 144 is the software, hardware, or software/hardware combination that performs a crawl on a target site 128 in

accordance with the present invention.” *Id.* at 6:8-10. “As crawlers are computer-driven, they can access large amounts of data very quickly, and they can perform crawls continuously.” *Id.* at 1:20-21. Cousins discloses minions where “[t]o service the crawl requests, a plurality of minions retrieve the crawl requests and crawl the target web sites to obtain the necessary data.” *Id.* at 4:26-27. Cousins further discloses that “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. Further, Cousins discloses that “[i]n another embodiment, minions 124 use accounts at different Internet Service Providers.” *Id.* at 16:27-28.

118. **“... comparing parameters of a request received from a user device with a list of scraping strategies, ...”** This claim limitation mirrors a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

119. **“... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request, ...”** This claim limitation mirrors a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

120. **“... identifying discrepancies between the request and the scraping strategy, ...”** This claim limitation mirrors a limitation present in claim

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

121. “... **adjusting the parameters of the request according to the scraping strategy to form an adjusted request, ...**” This claim limitation mirrors a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

122. “... **sending the adjusted request to a target through a proxy selected according to the scraping strategy.**” This claim limitation mirrors a limitation present in claim 1. In my opinion, for the same reasons discussed above regarding claim 1, Cousins discloses this claim limitation.

123. Thus, for the reasons discussed above, claim 19 is anticipated by Cousins.

124. I will now address the dependent claims under ground 1 as being anticipated by Cousins.

Claim 4

125. Claim 4 depends from claim 1 and adds the limitation “**wherein if the reply received from the target is negative, the method repeats by adjusting the parameters of the request according to a different scraping strategy from the list of scraping strategies to form a different adjusted request that is then sent to the target.**” I incorporate my discussion of claim 1 discussed above. In my

opinion, Cousins discloses this claim limitation. Cousins discloses conducting a scrape, retrieving information from a target, and then storing that information in a database. However, “[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID 300 to the policy engine 116 to generate policies for the request for data,” thus the new policies are used to form a new adjusted request for data from the target. *Id.* at 8:22-23.

126. In another example, if the reply from the target is negative, for example, “if servers at the target site’s main address are not accessible, for example, as a result of a third party attacking a target site 128 by sending an overly burdensome number of requests to the site 128,” the parameters of the request may be further adjusted by retrieving an alternate exact URL such that “the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104” which “allows the crawler 144 to bypass the main path to the target site 128 and access the target site’s direct address for the item of interest.” *Id.* at 8:33-9:5. “Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning.” *Id.* at 9:5-7.

Claim 5

127. Claim 5 depends from claim 4 and adds the limitation “**wherein the method is repeated with alternative scraping strategies selected from the list**

of scraping strategies until a positive response is obtained or until a maximum number of attempts is reached.” I incorporate my discussion of claim 4 above and of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed in claim 4, Cousins discloses conducting an alternative scraping strategy to obtain a positive response.

Claim 6

128. Claim 6 depends from claim 1 and adds the limitation “**wherein at least one of the parameters within the following components of the request is adjusted: a. URL address; b. headers; c. cookies; d. request body.**” I incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed in claims 4 and 5, Cousins discloses adjusting a URL address. For example, “if servers at the target site’s main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128,” the parameters of the request may be further adjusted by retrieving an alternate exact URL such that “the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104” which “allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest.” *Id.* at 8:33-9:5. “Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning.” *Id.* at 9:5-7.

Claim 9

129. Claim 9 depends from claim 1 and adds the limitation “**wherein the proxy remains available for multiple user device requests.**” I incorporate my

discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, Cousins discloses a queue of multiple user requests where “a minion 124 will pull the request from the queue 120 and fetch the page from the target site 128 that contains the requested data” and “[i]n a preferred embodiment, once a request is fed to a minion 124, the output queue 120 automatically schedules a second crawl request 400 to be made in the near future on the expectation that the first crawl request 400 will fail.” *Id.* at 8:13:4-8. In another example, Cousins discloses a dedicated minion for a target web site where “[b]y providing unique minions 124 to crawl web site for which they are custom[ed] designed, the crawls can be much more efficient.” *Id.* at 16:10-11.

130. Cousins discloses the management of multiple requests, for example, “[i]n a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL” and “[a]ccordingly, prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128” and “[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404.” *Id.* at 11:17-24.

Claim 10

131. Claim 10 depends from claim 1 and adds the limitation “**wherein the parameters of the request are not adjusted if the request received from the user device is aligned to one of the scraping strategies from the list.**” I

incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. Cousins discloses that a parameter of a request may include “a URI that identifies a specific item of data ... [f]or example, a user may request the crawler 144 to track a particular auction for a camera that appears on an auction site.” *Id.* at 6:26-28. Cousins further discloses that a scraping strategy based on a time stamp, for example, “if a request is for a current event news story on a web site, the window may require that the data have been crawled within the last four hours.” *Id.* at 8:10-11. Thus, if the stored crawl data is older than four hours, the requested crawl will proceed without any adjustment to the request parameters.

Claim 11

132. Claim 11 depends from claim 1 and adds the limitation “**wherein request optimizer operational settings include at least one of the following: scraping strategies, session ID numbers, servers accepted as request originators, operation timeframes, scheduled maintenance jobs, connectivity sessions timeouts, request processing logs, or any combination thereof and**

such settings are stored in a database on disk, in a flat file on disk, in a database in memory, or other storage media.” I incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, scraping strategies, *e.g.*, policies, are stored in a database, such as “[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc.” *Id.* at 3:15-16.

133. Another example, Cousins discloses an operational setting including at least a session ID number, for example, “[i]n a preferred embodiment, the real time client engine 100 converts 204 the request information into an LUID (logical user id)” and where “[a]n LUID is a data structure used to store and organize data used by the crawler 144 and to uniquely identify each auction monitored by the crawler 144 of the present invention.” *Id.* at 6:33-7:3. In addition, the “LUIDD also has a unique identification number 312.” *Id.* at 7:9.

134. In addition, Cousins discloses connectivity timeouts, as discussed above where “[i]f the data is invalid, or the data is nonresponsive, the client engine 100 passes 218 the LUID 300 to the policy engine 116 to generate policies for the request for data,” thus the new policies are used to form a new adjusted request for data from the target. *Id.* at 8:22-23.

Claim 12

135. Claim 12 depends from claim 1 and adds the limitation “**wherein the target is a web server serving the reply accessible by HTTP/HTTPS protocols and is attached to a network.**” I incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, “crawlers gather this information by using the Internet in a process known as crawling” where “[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” *Id.* at 1:17-20.

Claim 13

136. Claim 13 depends from claim 1 and adds the limitation “**wherein the proxy is a proxy server integrated into a service provider infrastructure.**” I incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies.” *Id.* at 16:28-29. In addition, “[m]inions 124 are preferably implemented to execute on servers 700 or other computing devices that reside anywhere.” *Id.* at 14:27-28.

137. Cousins discloses “in one embodiment, the crawler 144 collocates servers throughout the world, and minions 124 are stored on the collocated

servers” and “[i]n another embodiment, minions 124 use accounts at different Internet Service Providers.” *Id.* at 16:25-28. Accordingly, Cousins discloses the use of minions and proxies in both scenarios, integrated into a service provider infrastructure and separate from a service provider infrastructure.

Claim 14

138. Claim 14 depends from claim 1 and adds the limitation “**wherein the proxy is a proxy server separate from a service provider infrastructure.**” I incorporate my discussion of claim 1 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies.” *Id.* at 16:28-29.

139. Cousins discloses “in one embodiment, the crawler 144 collocates servers throughout the world, and minions 124 are stored on the collocated servers” and “[i]n another embodiment, minions 124 use accounts at different Internet Service Providers.” *Id.* at 16:25-28. Accordingly, Cousins discloses the use of minions and proxies in both scenarios, integrated into a service provider infrastructure and separate from a service provider infrastructure.

Claim 17

140. Claim 17 depends from claim 16 and adds the limitation “**wherein the request optimizer can adjust the request received from the user device by modifying at least one of the parameters within the following components: a. URL address; b. headers; c. cookies; d. request body.**” I incorporate my discussion of claim 16 in ground 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, as discussed above in claims 4, 5, and 6, Cousins discloses adjusting a URL address. For example, “if servers at the target site’s main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128,” the parameters of the request may be further adjusted by retrieving an alternate exact URL such that “the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104” which “allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest.” *Id.* at 8:33-9:5. “Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning.” *Id.* at 9:5-7.

Claim 22

141. Claim 22 depends from claim 19 and adds the limitation “**wherein if the reply received from the target is negative, the method repeats by adjusting the parameters of the request according to a different scraping strategy from the list of scraping strategies to form a different adjusted request that is then sent to the target.**” I incorporate my discussion of claims 1, 4, and 19 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed in claim 4, Cousins discloses adjusting the parameters of the request to form a different adjusted request.

Claim 23

142. Claim 23 depends from claim 22 and adds the limitation “**wherein the method is repeated with alternative scraping strategies selected from the list of scraping strategies until a positive response is obtained or until a maximum number of attempts is reached.**” I incorporate my discussion of claims 1, 5, 19, and 22 discussed above. In my opinion, Cousins discloses this claim limitation. As discussed in claim 4, Cousins discloses conducting an alternate scraping strategy to obtain a positive response.

Claim 24

143. Claim 24 depends from claim 19 and adds the limitation “**wherein the proxy remains available for multiple user device requests.**” I incorporate my

discussion of claim 19 in ground 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, as discussed above in claim 9, Cousins discloses a queue of multiple user requests where “a minion 124 will pull the request from the queue 120 and fetch the page from the target site 128 that contains the requested data” and “[i]n a preferred embodiment, once a request is fed to a minion 124, the output queue 120 automatically schedules a second crawl request 400 to be made in the near future on the expectation that the first crawl request 400 will fail.” *Id.* at 8:13:4-8.

144. Cousins discloses the management of multiple requests, for example, “[i]n a more complex case, two client engines 100 (or client engines 100 and any scheduled automatic crawls) could desire the same URL, but desire different information from the URL” and “[a]ccordingly, prior to placing a request for data in the queue 120, the policy engine 116 checks the requests for data currently in the queue 120 to determine if any request for data in the queue 120 is for the same target site 128” and “[i]f there are, the policy engine 116 combines the multiple requests 400 into a single crawl request 400 comprising one URL field 408, and two client and parser ID fields 312, 404.” *Id.* at 11:17-24.

Claim 25

145. Claim 25 depends from claim 19 and adds the limitation “**wherein at least one of the following parameters of the request is adjusted: a. URL**

address; b. headers; c. cookies; d. request body.” I incorporate my discussion of claim 19 in ground 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, as discussed above in claims 4, 5, and 6, Cousins discloses adjusting a URL address. For example, “if servers at the target site’s main address are not accessible, for example, as a result of a third party attaching a target site 128 by sending an overly burdensome number of requests to the site 128,” the parameters of the request may be further adjusted by retrieving an alternate exact URL such that “the crawler 144 of the present invention is still able to satisfy requests because the exact URL for a given web page is stored in the database 104” which “allows the crawler 144 to bypass the main path to the target site 128 and access the target site's direct address for the item of interest.” *Id.* at 8:33-9:5. “Thus, the crawler 144 will provide connectivity to a target site 128 even in circumstances when main path of access to the target site 128 is malfunctioning.” *Id.* at 9:5-7.

Claim 26

146. Claim 26 depends from claim 19 and adds the limitation “**wherein the parameters of the request are not adjusted if the request received from the user device is aligned to one of the scraping strategies from the list.**” I incorporate my discussion of claim 19 in ground 1 discussed above. In my opinion, Cousins discloses this claim limitation. For example, as discussed above in claim

10, Cousins discloses that a parameter of a request may include “a URI that identifies a specific item of data ... [f]or example, a user may request the crawler 144 to track a particular auction for a camera that appears on an auction site.” *Id.* at 6:26-28. Cousins further discloses that a scraping strategy based on a time stamp, for example, “if a request is for a current event news story on a web site, the window may require that the data have been crawled within the last four hours.” *Id.* at 8:10-11. Thus, if the stored crawl data is older than four hours, the requested crawl will proceed without any adjustment to the request parameters.

**GROUND 2: OBVIOUSNESS OF CLAIMS 1, 2, 3, and 16-21 BY COUSINS
(Ex. 1004) IN VIEW OF THE COMMON KNOWLEDGE OF A POSA**

147. In my opinion, Cousins discloses, or at minimum, Cousins in view of the common knowledge of a POSA teaches each of the limitations of claims 1, 2, 3, and 16-21 of the '770 patent. I will first address the independent claims.

Claim 1

148. **“A method for data extraction and gathering from a web server comprising: ...”** In my opinion, to the extent the preamble is limiting, and for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

149. **“... comparing parameters of a request received from a user device with a list of scraping strategies; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

150. **“... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

151. **“... identifying discrepancies between the request and the scraping strategy; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

152. “... **adjusting the parameters of the request according to the scraping strategy to form an adjusted request; ...**” In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation

153. “... **sending the adjusted request to a target through a proxy selected according to the scraping strategy; ...**” I incorporate my discussion of this element as presented above in Ground 1 regarding claim 1. As discussed, Cousins discloses minions as proxies. Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” *Id.* at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. As would have been known by a POSA, Squid is “a popular open source caching proxy for the Web.” Ex. 1009 at 1.

154. In my opinion, it would have been obvious to a POSA to utilize minions as proxies or to have minions use proxies as Cousins specifically states such a configuration. A POSA would also have known of the advantages of using proxies, for example as stated in applicant’s own admitted prior art in the ’770 patent, “[p]roxies can be used to bypass certain Internet restrictions (*e.g.* firewalls) by enabling a user to request the content through a (remote) proxy server instead of

accessing the content directly” where “[p]roxy servers are often used to get around geo-IP based content restrictions.” *Id.* at 2:25-29.

155. Thus, it would have been obvious to combine Cousins’ use of minions to include the use of minions as proxies, or the use of minions that use proxies as such an adaptation was specifically disclosed in Cousins. Further, use of minions as proxies or the use of minions and proxies would have been an obvious implementation for a POSA, including the use of proxies as was extensively discussed in the background section of the ’770 patent as applicant admitted prior art. Such a combination represents combining prior art elements (proxies and web scrapers) according to known methods (servicing requests for data from target web sites) to yield predictable results (bypass certain Internet restrictions).

156. **“... receiving a reply from the target.”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

157. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches the limitations of claim 1.

Claim 16

158. **“A system for data extraction and gathering from a network comprising at least: ...”** In my opinion, to the extent the preamble is limiting, and

for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation.

159. “... **a proxy rotator operable to provide a plurality of proxies over the network; ...**” In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “a proxy rotator” then such a limitation would have been common knowledge to a POSA.

160. As previously discussed, Cousins explicitly states that the “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. Cousins explicit reference to the Squid proxy would have been known by a POSA to be a proxy rotator as evidenced by the September 2018 web.archive capture of the Squid proxy website explaining the use of a proxy rotator entitled “Rotating the Squid outbound IPs.” Ex. 1010 at pp. 1-2. Thus, Squid’s disclosure of “Rotating the Squid outbound IPs” would have been known by a POSA to a separate piece of software that could exist on any computing device, for example within policy engine 116, server 700, or outgoing queue 120. Squid also discloses proxy rotating where “Squid iterates between all addresses (a.k.a. round-robin) for each cache miss” and “to use a redirector (see Chapter 11) to select the backend server.” Ex. 1009 at 309.

161. Further, the '770 patent, in its background section, discusses the use of “rotating proxies” with a web crawler where “[t]he task of attempting data retrieval further includes managing Hypertext Transfer Protocol (HTTP) and Hypertext Transfer Protocol Secure (HTTPS) requests, rotating proxies, and headers.” Ex. 1001 at 3:7-10.

162. In my opinion, it would have been obvious to a POSA to adapt the minions in Cousins to include the use of a proxy rotator as described in the '770 patent. The use of proxy rotators with a web scraper would have been an obvious implementation for a POSA and would have also been done expecting success as doing so was discussed in both Cousins and in the background section of the '770 patent.

163. Thus, it would have been obvious to combine Cousins' use of minions to scrape the web to include the use of a proxy rotator as such an adaptation was specifically disclosed in Cousins. Further, use of minions and a proxy rotator for web scraping would have been an obvious implementation for a POSA, including the use of rotating proxies as was discussed in the background section of the '770 patent as applicant admitted prior art and was also disclosed in the Squid documentation referenced by Cousins. Such a combination represents combining prior art elements (web scrapers and rotating proxies) according to known methods

(servicing requests for data from target web sites) to yield predictable results (bypass certain Internet restrictions).

164. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

165. **“... a request optimizer operable to compare parameters of a request received from a user device with a list of scraping strategies, ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation. The '770 patent, in an embodiment, states that request optimizer 104 “is a next generation proxy that is a part of the service provider infrastructure that accepts requests from the User Device and implements the requests through the system.” Ex. 1004 at 7:12-15.

166. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

167. **“... to select a scraping strategy from the list of scraping strategies a scraping strategy that corresponds to the parameters of the request, ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation.

168. **“... to identify discrepancies between the request and the scraping strategy, ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation.

169. **“... to adjust the parameters of the request according to the scraping strategy to form an adjusted request, ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation.

170. **“... to send the adjusted request to a target through a proxy selected from a proxy database by the proxy rotator according to the scraping strategy.** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation.

171. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches the limitations of claim 16.

Claim 19

172. **“A non-transitory computer readable medium for data extraction and gathering from a web server comprising instructions which, when executed by a computing within a service provider infrastructure, cause the computing device to perform: ...”** In my opinion, to the extent the preamble is limiting, and for at least the reasons discussed above in Ground 1 regarding claim 19, Cousins discloses this claim limitation.

173. **“... comparing parameters of a request received from a user device with a list of scraping strategies; ...”** In my opinion, for at least the

reasons discussed above in Ground 1 regarding claim 19, Cousins discloses this claim limitation.

174. **“... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 19, Cousins discloses this claim limitation.

175. **“... identifying discrepancies between the request and the scraping strategy; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 19, Cousins discloses this claim limitation.

176. **“... adjusting the parameters of the request according to the scraping strategy to form an adjusted request; ...”** In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 19, Cousins discloses this claim limitation.

177. **“... sending the adjusted request to a target through a proxy selected according to the scraping strategy.”** I incorporate my discussion of this element as presented above in Ground 1 regarding claim 19. However, if the Board does not agree that Cousins discloses “a proxy” then such a limitation would have been common knowledge to a POSA.

178. As previously discussed, Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.”

Ex. 1004 at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. As would have been known by a POSA, Squid is “a popular open source caching proxy for the Web.” Ex. 1009 at 1.

179. In my opinion, it would have been obvious to a POSA to utilize minions as proxies or to have minions use proxies as Cousins specifically states such a configuration. A POSA would also have known of the advantages of using proxies, for example as stated in applicant’s own admitted prior art in the ’770 patent, “[p]roxies can be used to bypass certain Internet restrictions (*e.g.* firewalls) by enabling a user to request the content through a (remote) proxy server instead of accessing the content directly” where “[p]roxy servers are often used to get around geo-IP based content restrictions.” *Id.* at 2:25-29.

180. Thus, it would have been obvious to combine Cousins’ use of minions to include the use of minions as proxies, or the use of minions that use proxies as such an adaptation was specifically disclosed in Cousins. Further, use of minions as proxies or the use of minions and proxies would have been an obvious implementation for a POSA, including the use of proxies as was extensively discussed in the background section of the ’770 patent as applicant admitted prior art. Such a combination represents combining prior art elements (proxies and web

scrapers) according to known methods (servicing requests for data from target web sites) to yield predictable results (bypass certain Internet restrictions).

181. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches the limitations of claim 19.

182. I will now address the dependent claims 2 and 3 that depend from claim 1 and then claim 18 that depends from claim 16, and then claims 20 and 21 that depend from claim 19.

Claim 2

183. Claim 2 depends from claim 1 and adds the limitation “**further comprising: checking and analyzing the metadata of the reply received from the target.**” I incorporate my discussion of claim 1 in ground 2 discussed above. In my opinion, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “checking and analyzing the metadata” then such a limitation would have been common knowledge to a POSA.

184. Cousins discloses that “[c]rawlers gather this information by using the Internet in a process known as crawling” where “[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20.

185. The '770 patent states that “[i]t should be noted that the metadata of the response is always visible to the Request Optimizer 104, e.g. URL and HTTP headers.” Ex. 1001 at 7:18-20.

186. A POSA would have understood that the use of standard web protocols includes checking and analyzing associated metadata found in URL and HTTP headers. A POSA would have known that, as stated in RFC 7231, “when a network “message includes a payload body, the representation header fields describe how to interpret the representation data enclosed in the payload body.” Ex. 1005 at 5.

187. In addition, a POSA would have known that, as stated in RFC 7231, “HTTP semantics include the intentions defined by each request method ... and the meaning of other control data and resource metadata that might be given in response header fields.” Ex. 1005 at 6.

188. Thus, it would have been obvious to adapt Cousins’ use of a crawler using standard web protocols (TCP/IP, HTTP, etc.) to include the analyzing of metadata. Further, the analyzing of metadata as contained in standard web protocols would have been an obvious implementation for a POSA, as discussed in RFC 7231. Such a combination represents combining prior art elements (standard web protocols and metadata) according to known methods (the use of HTTP

protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites).

189. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

Claim 3

190. Claim 3 depends from claim 2 and adds the limitation “**further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata.**” I incorporate my discussion of claim 2 above and of claim 1 in ground 2 discussed above. In my opinion, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “forwarding the reply to the user device pursuant to the analysis of the metadata” then such a limitation would have been common knowledge to a POSA.

191. As discussed, Cousins discloses “[a] request source [that] is any computer process that requests information from a data address” where “a user will access the crawler’s web site through the user’s Internet connection, and then input a request for a specific item into a form provided by the crawler’s web site” *Id.* at 6:15-24. Cousins discloses both “scheduled” and “user-demanded crawl[s].” *Id.* at 3:26. In my opinion, as would have been obvious to a POSA, Cousins disclosure of a user’s access through an Internet connection, and to input a request, as well as the

creation of a user-demanded crawl discloses requests and replies to and from a user device.

192. As discussed in claim 2, Cousins discloses that “[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20. And, that the ’770 patent states that “[i]t should be noted that the metadata of the response is always visible to the Request Optimizer 104, e.g. URL and HTTP headers.” Ex. 1001 at 7:18-20.

193. A POSA would have understood that the use of standard web protocols includes forwarding content such as the reply to the user device based on the analysis of the metadata. For example, a POSA would have known that, as stated in RFC 7231, the HTTP protocol allows for “the recipient [to] establish a tunnel to the destination origin server identified by the request-target and, if successful, thereafter restrict its behavior to blind forwarding of packets, in both directions, until the tunnel is closed.” Ex. 1005 at 30.

194. Thus, it would have been obvious to combine Cousins’ use of a crawler to forward retrieved data from a target web site to a user device based on an analysis of the metadata with the use of standard web protocols that include the analyzing of metadata. The analyzing of metadata as contained in standard web protocols would have been an obvious implementation for a POSA, as discussed in

RFC 7231. Such a combination represents combining prior art elements (standard web protocols and metadata) according to known methods (the use of HTTP protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites and sending the data back to the user device).

195. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

Claim 17

196. Claim 17 depends from claim 16 and adds the limitation “**wherein at least one of the parameters within the following components of the request is adjusted: a. URL address; b. headers; c. cookies; d. request body.**” I incorporate my discussion of claim 16 in ground 2 and claims 6 and 13 of ground 1 discussed above.

Claim 18

197. Claim 18 depends from claim 16 and adds the limitation “**wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device.**” I incorporate my discussion of claim 16 in ground 2 discussed above. In my opinion, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “wherein the request optimizer utilizes the same proxy for an unlimited number of requests received

from the same user device” then such a limitation would have been common knowledge to a POSA.

198. A POSA would have understood that the use of standard web protocols includes utilizing the same proxy for an unlimited number of requests received from the same user device, also known as a persistent connection. For example, a POSA would have known that, as stated in RFC 7540, the HTTP specifies that “HTTP/2 connections are persistent” and that “[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or until the server closes the connection.” Ex. 1008 at 65.

199. Thus, it would have been obvious to adapt Cousins’ use of a crawler using standard web protocols (TCP/IP, HTTP, etc.) to include the use of a persistent connect to the same proxy. Further, the use of persistent web connections to the same proxy would have been an obvious implementation for a POSA, as discussed in RFC 7540. Such a combination represents combining prior art elements (standard web protocols and connection methods) according to known methods (the use of HTTP protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites).

200. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

Claim 20

201. Claim 20 depends from claim 19 and adds the limitation “**further comprising checking and analyzing the metadata of the reply received from the target.**” I incorporate my discussion of claims 2 and 19 in ground 2 discussed above. In my opinion, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “further comprising checking and analyzing the metadata of the reply received from the target” then such a limitation would have been common knowledge to a POSA as discussed above regarding claim 2.

202. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

Claim 21

203. Claim 21 depends from claim 20 and adds the limitation “**further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata.**” I incorporate my discussion of claims 3, 19, and 20 above in ground 2 discussed above. In my opinion, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata”

then such a limitation would have been common knowledge to a POSA as discussed above regarding claim 3.

204. Thus, for the reasons discussed above, Cousins in view of the common knowledge of a POSA teaches this limitation.

**GROUND 3: OBVIOUSNESS OF CLAIMS 1, 7, 8, 13, and 16-19 BY
COUSINS (Ex. 1004) IN VIEW OF TAYLOR (Ex. 1006)**

The Combination of Cousins and Taylor

205. In my opinion, it would have been obvious to combine Cousins with Taylor.

206. Cousins is titled “Minimal Impact Crawler” and discloses the use of minions to retrieve crawl requests and crawl the target web sites. *See* Ex. 1004 at Abstract, code (57). Cousins further discloses that the minions “use accounts at different Internet Service Providers” and that minions “may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies.” *Id.* at 16:27-29.

207. In my opinion, a POSA would recognize the use of proxies from Internet Service Providers allows for a large network to more effectively crawl and retrieve information from target web sites. Thus, in my opinion, it would have been obvious to adapt and expand Cousins to incorporate the service provider proxies and infrastructure taught by Taylor.

208. Taylor discloses a “service provider device [that] is further configured to direct network communication to a proxy device which, in turn, exchanges communication with targeted server devices over a wide area network, such as the Internet.” Ex. 1006 at ¶[0008]. Taylor further discloses that the service provider device is configured to “detect the communication type identifier” associated with

the request and based on a correspondence of the destination IP address and a DNS database the service provider device “can select an appropriate middlebox 40” to proxy the network communication between a communication device and a server device. *Id.* at ¶[0044].

209. The use of a service provider proxies and databases as suggested by Taylor into the Cousins’ web crawler architecture would have been an obvious implementation for a POSA and would have also been done expecting success.

210. In my opinion, Cousins and Taylor provide a teaching, suggestion, or motivation to combine Taylor’s service provider middlebox proxy architecture with Cousins web crawler architecture that also uses a database to select the appropriate proxies associated with a target server.

211. Such a combination represents combining prior art elements (service provider proxies) according to known methods (web page crawler requests and responses) to yield predictable results (increased amounts of proxies and performance).

212. In my opinion, Cousins in view of Taylor discloses each of the limitations of claims 1, 7, 8, 13, and 16-19 of the ’770 Patent as discussed below. These claims are obvious by the disclosures in Cousins and Taylor.

213. I will first address the independent claims 1, 16, and 19.

Claim 1

214. **“A method for data extraction and gathering from a web server comprising: ...”** In my opinion, to the extent the preamble is limiting, and for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses this claim limitation.

215. **“... comparing parameters of a request received from a user device with a list of scraping strategies; ...”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses this claim limitation. Cousins discloses a user’s access through an Internet connection, inputting a request for a specific item into a form, as well as the creation of a user-demanded crawl discloses a user request from a user device.

216. However, if the Board does not agree that Cousins discloses a “user device,” then Taylor in the same field discloses such a limitation.

217. Taylor discloses where, “[f]or example, assume the case where a user of a first computer device 18-1 seeks to establish a video conferencing connection (e.g., SKYPE) with server device 22” and [i]n such a case, the computer device 18-1 generates a device access request 30 identifying the server device 22 and forwards the request within the residential network 14 to the data communication device 16.” Ex. 1006 at [0027].

218. It would have been obvious to combine Cousins' user's access through an Internet connection and the inputting of a request for a specific item into a form to include the use of a device or a computer device as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (inputting data into a computing device) according to known methods (networks and web pages) to yield predictable results.

219. **"... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request; ..."** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses this claim limitation.

220. **"... identifying discrepancies between the request and the scraping strategy; ..."** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses this claim limitation.

221. **"... adjusting the parameters of the request according to the scraping strategy to form an adjusted request; ..."** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses this claim limitation

222. **"... sending the adjusted request to a target through a proxy selected according to the scraping strategy; ..."** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 1, Cousins discloses

this claim limitation. However, if the Board does not agree that Cousins discloses “a proxy” then Taylor, in the same field, discloses such a limitation.

223. As previously discussed, Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” *Id.* at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29. Further, Squid, as known by a POSA, is a proxy rotator. Ex. 1010 at pp. 1-2.

224. Taylor discloses selecting a proxy based on the strategy of the type of communication detected, for example, “[b]ased upon the detected communication type identifier 32, the service provider device 12 can select an appropriate, associated middlebox 40 configured to proxy and monitor the application, workflow, or protocol traffic from the client device 15.” Ex. 1006 at ¶[0029].

225. In another example, Taylor discloses selecting a proxy based upon the type of application, workflow, or protocol being used, for example, each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis.” Ex. 1006 at ¶[0024].

226. It would have been obvious to adapt Cousins’ use of minions to include the use of minions as proxies, or the use of minions that use proxies as such an adaptation was specifically disclosed in Cousins. It would have also been obvious to adapt Cousins’ use of minions to include the functionality of a proxy as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (web scraping and proxies) to yield predictable results.

227. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

228. “... **receiving a reply from the target.**” In my opinion, for at least the reasons discussed above in Ground 1 regarding claim 1, Cousins discloses this claim limitation.

229. Thus, for the reasons discussed above, Cousins in view of Taylor teaches the limitations of claim 1.

Claim 16

230. “**A system for data extraction and gathering from a network comprising at least: ...**” In my opinion, to the extent the preamble is limiting, and for at least the reasons discussed above in Ground 1 regarding claim 16, Cousins discloses this claim limitation. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

231. **“... a proxy rotator operable to provide a plurality of proxies over the network; ...”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 16. Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

232. **“... a request optimizer operable to compare parameters of a request received from a user device with a list of scraping strategies, ...”** In my opinion, for at least the reasons discussed above in Ground 1 and Ground 2 regarding claim 16, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation. The '770 patent, in an embodiment, states that request optimizer 104 “is a next generation proxy that is a part of the service provider infrastructure that accepts requests from the User Device and implements the requests through the system.” Ex. 1001 at 7:12-15. As discussed in claim 16, Cousins discloses the functions of the request optimizer in the policy engine 116, which accepts requests and implements the requests through the use of a crawler and minions that obtain the requested data.

233. However, if the Board does not agree that Cousins discloses a request optimizer that is a part of the service provider infrastructure then Taylor, in the same field, discloses such a limitation.

234. Taylor discloses a “proxy that is a part of the service provider infrastructure that accepts requests from the User Device and implements the requests through the system.” For example, Taylor discloses “a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel” including a “virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox.” Ex. 1006 at ¶[0007]. Taylor also discloses that the “service provider device is further configured to direct network communication to a proxy device which, in turn, exchanges communication with targeted server devices over a wide area network, such as the Internet.” *Id.* at ¶[0008]. Taylor further discloses a method that “includes receiving, by the service provider device, a device access request from the client device ... [and] establishing, by the service provider device, a secure communication channel between the client device and a middlebox associated with the detected communication type identifier. *Id.* at ¶[0009].

235. It would have been obvious to adapt Cousins’ policy engine, *e.g.*, request optimizer, that is a part of a service provider infrastructure as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (service provider control) to yield predictable results.

236. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

237. **“... to select a scraping strategy from the list of scraping strategies a scraping strategy that corresponds to the parameters of the request, ...”** In my opinion, for at least the reasons discussed above in Ground 1 and Ground 2 regarding claim 16, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

238. **“... to identify discrepancies between the request and the scraping strategy, ...”** In my opinion, for at least the reasons discussed above in Ground 1 and Ground 2 regarding claim 16, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

239. **“... to adjust the parameters of the request according to the scraping strategy to form an adjusted request, ...”** In my opinion, for at least the reasons discussed above in Ground 1 and Ground 2 regarding claim 16, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

240. “... to send the adjusted request to a target through a proxy selected from a proxy database by the proxy rotator according to the scraping strategy. In my opinion, for at least the reasons discussed above in Ground 1 and Ground 2 regarding claim 16, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation.

241. However, if the Board does not agree that Cousins, or Cousins in view of the common knowledge of a POSA, discloses “a proxy selected from a proxy database” then Taylor, in the same field, discloses such a limitation. As discussed, Taylor discloses that the service provider device, *e.g.*, the service provider infrastructure, is configured to “detect the communication type identifier” associated with the request and based on a correspondence of the destination IP address and a DNS database the service provider device “can select an appropriate middlebox 40 ... to proxy and monitor the network communication between the data communication device 16 and the server device 22.” *Id.* at ¶[0044]. Further, “each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol.” *Id.* at ¶[0024]. Thus, in my opinion, Taylor discloses the use of a database to select a proxy, *e.g.*, middlebox 40 according to the scraping strategy.

242. It would have been obvious to adapt a proxy rotator disclosed by Cousins’ reference to a Squid rotating proxy that is selected as part of a database as

disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (proxy rotation) to yield predictable results.

243. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

244. Thus, for the reasons discussed above, Cousins in view of Taylor teaches the limitations of claim 16.

Claim 19

245. **“A non-transitory computer readable medium for data extraction and gathering from a web server comprising instructions which, when executed by a computing within a service provider infrastructure, cause the computing device to perform: ...”** In my opinion, to the extent the preamble is limiting, and for at least the reasons discussed above in Grounds 1 and regarding claim 19, Cousins discloses this claim limitation.

246. **“... comparing parameters of a request received from a user device with a list of scraping strategies; ...”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 19, Cousins discloses this claim limitation.

247. **“... selecting, from the list of scraping strategies, a scraping strategy corresponding to the parameters of the request; ...”** In my opinion, for

at least the reasons discussed above in Grounds 1 and 2 regarding claim 19, Cousins discloses this claim limitation.

248. **“... identifying discrepancies between the request and the scraping strategy; ...”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 19, Cousins discloses this claim limitation.

249. **“... adjusting the parameters of the request according to the scraping strategy to form an adjusted request; ...”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 19, Cousins discloses this claim limitation.

250. **“... sending the adjusted request to a target through a proxy selected according to the scraping strategy.”** In my opinion, for at least the reasons discussed above in Grounds 1 and 2 regarding claim 19, Cousins discloses this claim limitation. However, if the Board does not agree that Cousins discloses “a proxy” then such a limitation would have been obvious under Cousins in view of Taylor.

251. As previously discussed, Cousins discloses that “in general, minions 124 are located topologically as close as possible to the target site’s server 708.” Ex. 1004 at 15:1-2. In addition, “[m]inions 124 may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as

Squid).” *Id.* at 16:28-29. Further, Squid, as known by a POSA, is a proxy rotator. Ex. 1009 at pp. 1-2.

252. Taylor discloses selecting a proxy based on the strategy of the type of communication detected, for example, “[b]ased upon the detected communication type identifier 32, the service provider device 12 can select an appropriate, associated middlebox 40 configured to proxy and monitor the application, workflow, or protocol traffic from the client device 15.” Ex. 1006 at ¶[0029].

253. In another example, Taylor discloses selecting a proxy based upon the type of application, workflow, or protocol being used, for example, “each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis.” *Id.* at ¶[0024].

254. It would have been obvious to adapt Cousins’ use of minions to include the use of minions as proxies, or the use of minions that use proxies as such an adaptation was specifically disclosed in Cousins. It would have also been obvious to adapt Cousins’ use of minions to include the functionality of a proxy as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (web scraping and proxies) to yield predictable results.

255. Thus, for the reasons discussed above, Cousins in view of Taylor teaches the limitations of claim 19.

256. I will now address the dependent claims.

Claim 7

257. Claim 7 depends from claim 1 and adds the limitation “**wherein the proxy is selected by a service provider infrastructure from a proxy database.**”

I incorporate my discussion of claim 1 in ground 3 discussed above. However, if the Board does not agree that Cousins, or Cousins in view of the common knowledge of a POSA, discloses a proxy “selected by a service provider infrastructure from a proxy database” then Taylor, in the same field, discloses such a limitation.

258. Taylor discloses “a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel” including a “virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox.” *Id.* at ¶[0007]. Taylor also discloses that the “service provider device is further configured to direct network communication to a proxy device which, in turn, exchanges communication with targeted server devices over a wide area network, such as the Internet.” *Id.* at ¶[0008]. Taylor further discloses a method that “includes receiving, by the service provider device,

a device access request from the client device ... [and] establishing, by the service provider device, a secure communication channel between the client device and a middlebox associated with the detected communication type identifier. *Id.* at ¶[0009].

259. Taylor discloses that the service provider device, *e.g.*, the service provider infrastructure, is configured to “detect the communication type identifier” associated with the request and based on a correspondence of the destination IP address and an entry in the local DNS database the service provider device “can select an appropriate middlebox 40” to proxy the network communication between a communication device and a server device. *Id.* at ¶[0044]. As “each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol” then each entry in the DNS database is associated with a particular middlebox proxy thereby, in my opinion, functioning as a proxy database.

260. It would have been obvious to adapt a proxy rotator disclosed by Cousins’ reference to a Squid rotating proxy that is selected as part of a database as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (proxy rotation) to yield predictable results.

261. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

Claim 8

262. Claim 8 depends from claim 7 and adds the limitation “**wherein the proxy is selected according to the scraping strategy and includes geolocation settings, proxy type, proxy availability, or any combination thereof.**” I incorporate my discussion of claims 1 and 7 in ground 3 discussed above. Cousins discloses that policies may include geolocation settings, for example, “[p]olicies are maintained in databases sorted by different categories such as target site, time of day, location of target site, etc.” Ex. 1004 at 3:15-16. Furthermore, Cousins discloses that “certain types of requests may be transmitted to a specific minion 128” where “minion 124 may be restricted to a particular area (such as ‘stamps & coins’) so that the behavior of the minion 124 will emulate human behavior.” *Id.* at 16:18-21. Further, as discussed, Cousins discloses the minions “may also be implemented to be multithreaded and/or use multiple IP addresses and/or use proxies (such as Squid).” *Id.* at 16:28-29.

263. Taylor also discloses selecting proxies, *e.g.*, middleboxes, based on proxy type. For example, “the middlebox is configured to monitor and control the network traffic associated with a particular application, workflow, or protocol as executed by the client device.” Ex. 1006 at ¶[0017]. Further, Taylor discloses

selecting the proxy based on the type of application. For example, where “each middlebox 40 is configured to act as a proxy device for network traffic associated with a particular application, workflow, or protocol executed by the computer devices 18 within the residential network 14 on a per-flow or per-access basis” where “[f]or example, the first middlebox 40-1 can be configured to proxy and monitor network traffic pertaining to a particular type of video conferencing (e.g., SKYPE) while the second middlebox 40-2 can be configured to proxy and monitor network traffic pertaining to a particular type motion picture streaming (e.g., NETFLIX).” *Id.* at ¶0024].

264. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

Claim 13

265. Claim 13 depends from claim 1 and adds the limitation “**wherein the proxy is a proxy server integrated into a service provider infrastructure.**” I incorporate my discussion of claim 1 in ground 3 discussed above.

266. Taylor discloses “a service provider network system [that] includes an offsite or cloud-based host, such as a service provider device, which is configured to establish a secure communication channel” including a “virtual private network (VPN) connection, between a client device of a residential computer network and a cloud-based proxy device or middlebox.” *Id.* at ¶[0007].

267. It would have been obvious to adapt a proxy disclosed by Cousins' that is integrated into a service provider's infrastructure as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (proxies) according to known methods (proxies controlled by a service provider) to yield predictable results.

268. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

Claim 15

269. Claim 15 depends from claim 1 and adds the limitation “**wherein upon receiving the request, a service provider infrastructure performs the adjusting of the parameters of the request autonomously from the user device.**” I incorporate my discussion of claim 1 in ground 3 as discussed above.

270. Taylor discloses “a service provider device” as a part of the service provider infrastructure that performs an adjusting of parameter autonomously from the user device. For example, where a user device requests access to a specific server device but the server provider autonomously redirects the request. For example, “[f]or example, assume the case where a user of a first computer device 18-1 seeks to establish a video conferencing connection (e.g., SKYPE) with server device 22” and [i]n such a case, the computer device 18-1 generates a device access request 30 identifying the server device 22 and forwards the request within

the residential network 14 to the data communication device 16,” but “based upon the configuration of the data communication device 16 (e.g., as configured with the OpenFlow firmware), the data communication device 16 forwards the device access request 30 to the service provider device 12 via network communication channel 25 (e.g., over a WAN such as the Internet), rather than to the server device 22 identified by the device access request 30.” *Id.* at [0027].

271. It would have been obvious to adapt adjusting of the parameters of the request disclosed by Cousins’ that is autonomous as disclosed in Taylor. Such a combination would have been merely combining known prior art elements (information requests) according to known methods (forwarding network data) to yield predictable results.

272. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

Claim 17

273. Claim 17 depends from claim 16 and adds the limitation “**wherein the request optimizer can adjust the request received from the user device by modifying at least one of the parameters within the following components: a. URL address; b. headers; c. cookies; d. request body.**” In my opinion, for at least the reasons discussed above in claim 6 in Ground 1, claim 17 in Ground 2,

and claim 16 in Ground 3, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation.

274. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

Claim 18

275. Claim 18 depends from claim 16 and adds the limitation “**wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device.**” In my opinion, for at least the reasons discussed above in claim 18 in Ground 2, and claim 16 in Ground 3, Cousins, or Cousins in view of the common knowledge of a POSA, discloses this claim limitation.

276. Thus, for the reasons discussed above, Cousins in view of Taylor teaches this limitation.

**GROUND 4: OBVIOUSNESS OF CLAIMS 2, 3, 20, and 21 BY
COUSINS (Ex. 1004) IN VIEW OF RFC 7231 (Ex. 1005)**

The Combination of Cousins and RFC 7231

277. In my opinion, it would have been obvious to combine Cousins with RFC 7231.

278. Cousins is titled “Minimal Impact Crawler” and discloses the use of minions to retrieve crawl requests and crawl the target web sites. *See* Ex. 1004 at Abstract, code (57). Cousins further discloses that “[c]rawlers gather this information by using the Internet in a process known as crawling” and that “[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” *Id.* at 1:17-20. RFC 7231 is the standards track document to define the Hypertext Transfer Protocol (HTTP/1.1) semantics and content and thus, in my opinion, are highly relevant to the protocol features disclosed in Cousins.

279. Thus, in my opinion, it would have been obvious to combine Cousins and RFC 7231.

280. In my opinion, Cousins is directed to accessing information from target web sites using standard web protocols, including TCP/IP and HTTP over the Internet. A POSA would recognize the use of RFC 7231 standards in the use of Internet communications. RFC 7231 specifies standards track protocols for the

Internet community as it describes the usage of metadata, request methods, header fields, responses, and security considerations. Ex. 1005 at 3-5. Thus, it would have been obvious to adapt Cousins to adhere to the RFC 7231 standards that entails network communications.

281. Cousins and RFC 7231 provide a teaching, suggestion, or motivation to combine Cousins with RFC 7231 standards for consistency and interoperability with other network devices and applications.

282. Such a combination represents combining prior art elements (network components including clients, proxies, and target servers) according to known methods (Internet standards) to yield predictable results (compatibility and interoperability). In addition, the combination is merely a simple modification of one known element (network communications) for another (standards-based communication). Additionally, the combination is the use of a known technique (network communications) to control similar devices (standards) in the same way.

283. In my opinion, the limitations of claims 2, 3, 20, and 21 would have been obvious to a POSA based on the disclosures in Cousins and RFC 7231.

284. In my opinion, Cousins in view of RFC 7231 discloses each of the limitations of claims 2, 3, 20, and 21 of the '770 Patent as discussed below. .

Claims 2 and 20

285. Claim 2 depends from claim 1. Claim 20 depends from claim 19.

Claim 2 and claim 20 add the limitation **“further comprising: checking and analyzing the metadata of the reply received from the target.”** I incorporate my discussion of claim 1 and claim 19 in grounds 1, 2, and 3 discussed above. In my opinion, Cousins in view of RFC 7231 discloses this claim limitation. Cousins discloses that “computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20.

286. RFC 7231 discloses “[f]or the purposes of HTTP, a ‘representation’ is information that is intended to reflect a past, current, or desired state of a given resource, in a format that can be readily communicated via the protocol, and that consists of a set of representation metadata and a potentially unbounded stream of representation data.” Ex. 1005 at 7. RFC 7231 further explains that the metadata is analyzed to determine how to interpret data, for example, “[r]epresentation header fields provide metadata about the representation” and “[w]hen a message includes a payload body, the representation header fields describe how to interpret the representation data enclosed in the payload body.” *Id.* at 8.

287. It would have been obvious to adapt Cousins’ use of a crawler using standard web protocols (TCP/IP, HTTP, etc.) to include the analyzing of metadata

as discussed in RFC 7231. Such a combination represents combining prior art elements (standard web protocols and metadata) according to known methods (the use of HTTP protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites).

288. Thus, for the reasons discussed above, Cousins in view of RFC 7231 teaches the limitations of claims 2 and 20.

Claims 3 and 21

289. Claim 3 depends from claim 2. Claim 21 depends from claim 20. Claim 3 and claim 21 add the limitation **“further comprising: forwarding the reply to the user device pursuant to the analysis of the metadata.”** I incorporate my discussion of claims 2 and 20 discussed above. In my opinion, Cousins in view of RFC 7231 discloses this claim limitation. Cousins discloses that “computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20.

290. RFC 7231 discloses analysis of metadata that includes where “HTTP semantics include the intentions defined by each request method (Section 4), extensions to those semantics that might be described in request header fields ... and the meaning of other control data and resource metadata that might be given in response header fields.” Ex. 1005 at 6. RFC 7231 describes a tunnel mode as bi-

directional where “[t]he CONNECT method requests that the recipient establish a tunnel to the destination origin server identified by the request-target and, if successful, thereafter restrict its behavior to blind forwarding of packets, in both directions, until the tunnel is closed.” *Id.* at 30.

291. RFC 7231 further explains forwarding the response by establishing the tunnel based on the metadata header analysis, for example, where “[t]he recipient proxy can establish a tunnel either by directly connecting to the request-target or, if configured to use another proxy, by forwarding the CONNECT request to the next inbound proxy” where a successful “response indicates that the sender (and all inbound proxies) will switch to tunnel mode immediately after the blank line that concludes the successful response’s header section.” *Id.* at 31.

292. It would have been obvious to combine Cousins’ use of a crawler to forward retrieved data from a target web site to a user device based on an analysis of the metadata with the use of standard web protocols that include the analyzing of metadata as discussed in RFC 7231. Such a combination represents combining prior art elements (standard web protocols and metadata) according to known methods (the use of HTTP protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites and sending the data back to the user device).

293. Thus, for the reasons discussed above, Cousins in view of RFC 7231 teaches the limitations of claims 3 and 21.

GROUND 5: OBVIOUSNESS OF CLAIM 18 BASED ON COUSINS

(Ex. 1004) IN VIEW OF RFC 7540 (Ex. 1008)

The Combination of Cousins and RFC 7540

294. In my opinion, it would have been obvious to combine Cousins with RFC 7540.

295. Cousins is titled “Minimal Impact Crawler” and discloses the use of minions to retrieve crawl requests and crawl the target web sites. *See* Ex. 1004 at Abstract, code (57). Cousins further discloses that “[c]rawlers gather this information by using the Internet in a process known as crawling” and that “[e]ssentially, computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20. RFC 7540 is the standards track document to define the Hypertext Transfer Protocol (HTTP/2) semantics and thus, in my opinion, are highly relevant to the protocol features disclosed in Cousins.

296. Thus, in my opinion, it would have been obvious to combine Cousins and RFC 7540.

297. In my opinion, Cousins is directed to accessing information from target web sites using standard web protocols, including TCP/IP and HTTP over the Internet. A POSA would recognize the use of RFC 7540 standards in the use of Internet communications. RFC 7540 specifies standards track protocols for the

Internet community as it describes the usage of protocol, frames, streams and multiplexing, message exchanges, connection management, and security considerations. Ex. 1008 at 2-4. Thus, it would have been obvious to adapt Cousins to adhere to the RFC 7540 standards that entails network communications.

298. Cousins and RFC 7540 provide a teaching, suggestion, or motivation to combine Cousins with RFC 7540 standards for consistency and interoperability with other network devices and applications.

299. Such a combination represents combining prior art elements (network components including clients, proxies, and target servers) according to known methods (Internet standards) to yield predictable results (compatibility and interoperability). In addition, the combination is merely a simple modification of one known element (network communications) for another (standards-based communication). Additionally, the combination is the use of a known technique (network communications) to control similar devices (standards) in the same way.

300. Thus, for the reasons discussed above, Cousins in view of RFC 7540 teaches the limitations of claim 18.

Claim 18

301. Claim 18 depends from claim 16 and adds the limitation **“wherein the request optimizer utilizes the same proxy for an unlimited number of requests received from the same user device.”** I incorporate my discussion of claim 16 in grounds 1, 2, and 3 discussed above. In my opinion, Cousins in view of RFC 7540 discloses this claim limitation. Cousins discloses that “computers of the crawler take on the role of browsing users and access the target web sites using the standard web protocols (TCP/IP, HTTP, etc.) over the internet.” Ex. 1004 at 1:17-20.

302. RFC 7540 discloses connection management that includes persistent connections where “HTTP/2 connections are persistent” and that “[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or HTTP/2 connections are persistent” and “[f]or best performance, it is expected that clients will not close connections until it is determined that no further communication with a server is necessary (for example, when a user navigates away from a particular web page) or until the server closes the connection.” Ex. 1008 at 65.

303. It would have been obvious to adapt Cousins’ use of a crawler using standard web protocols (TCP/IP, HTTP, etc.) to include the use of a persistent

connect to the same proxy as discussed in RFC 7540. Such a combination represents combining prior art elements (standard web protocols and connection methods) according to known methods (the use of HTTP protocols in servicing requests for data from target web sites) to yield predictable results (retrieving data from the target web sites).

304. Thus, for the reasons discussed above, Cousins in view of RFC 7540 teaches the limitations of claim 18.

.....

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 11 May, 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
- Cellular and PCS Standards
- Cellular Telephone Architecture
- Digital Signal Processing

- Telecommunications Technology
- VoIP Technology
- Computer Networking
- 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.