

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

BRIGHT DATA LTD.,

Petitioner

v.

OXYLABS, UAB,

Patent Owner

Case IPR2023-01441

Patent No 7,010,526

DECLARATION OF KEVIN C. ALMEROTH, Ph.D.

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	PROFESSIONAL BACKGROUND.....	1
III.	COMPENSATION	13
IV.	LEGAL CONSIDERATION	13
A.	Anticipation.....	14
B.	Obviousness	Error! Bookmark not defined.
C.	Claim Interpretation	14
V.	OVERVIEW OF TASK & BASIS FOR OPINIONS	15
VI.	LEVEL OF SKILL IN THE RELEVANT ART	16
VII.	OVERVIEW OF THE PRIOR ART.....	17
A.	U.S. Patent Pub. No. 2003/0163454 (“Jacobsen”).....	17
B.	U.S. Patent No. 7,082,427 (“Seibel”)	19
VIII.	OVERVIEW OF THE ’526 PATENT.....	23
A.	Independent Claims 1 and 34.....	26
B.	Prosecution History.....	28
IX.	CLAIM CONSTRUCTION.....	28
A.	Creating at least a second data miner in response to the customer request for information	28
X.	JACOBSEN DOES NOT DISCLOSE EACH CLAIM FEATURE	31
A.	[1e], [34d]: Jacobsen Does Not Create a Data Miner In Response to a Customer Request for Information	31
1.	Creating ... in response to the customer request for information	31
2.	Second Data Miner.....	34
XI.	SEIBEL DOES NOT DISCLOSE EACH CLAIM FEATURE	37
A.	[1b]-[1c], [34b]: Petitioner’s Mapping of the Data Store is Inconsistent with Seibel’s Disclosure.....	37
B.	[1e], [34d]: Seibel Does Not Create a Data Miner In Response to a Customer Request for Information.....	41
XII.	CONCLUSION.....	45

I. INTRODUCTION

1. My name is Dr. Kevin C. Almeroth, Ph.D. I have been retained by Oxylabs, UAB as an independent expert consultant in this inter partes review (“IPR”) proceeding before the United States Patent and Trademark Office (“PTO”).

II. PROFESSIONAL BACKGROUND

2. All of my opinions stated in this Declaration are based on my own personal knowledge and professional judgment. In forming my opinions, I have relied on my knowledge and experience in designing, developing, researching, and teaching the technology referenced in this Declaration.

3. I am over 18 years of age and, if I am called upon to do so, I would be competent to testify as to the matters set forth herein. I understand that a copy of my current curriculum vitae, which details my education and professional and academic experience, is being submitted as EX2002. The following provides a brief overview of some of my experience that is relevant to the matters set forth in this Declaration.

4. I am currently a Professor Emeritus in the Department of Computer Science at the University of California, Santa Barbara (UCSB). While active at UCSB, I held faculty appointments and was a founding member of the Computer Engineering (CE) Program, Media Arts and Technology (MAT) Program, and the Technology Management Program (TMP). I also served as the Associate Director

of the Center for Information Technology and Society (CITS) from 1999 to 2012. I have been a faculty member at UCSB since July 1997.

5. I hold three degrees from the Georgia Institute of Technology: (1) a Bachelor of Science degree in Information and Computer Science (with minors in Economics, Technical Communication, and American Literature) earned in June 1992; (2) a Master of Science degree in Computer Science (with specialization in Networking and Systems) earned in June 1994; and (3) a Doctor of Philosophy (Ph.D.) degree in Computer Science (Dissertation Title: Networking and System Support for the Efficient, Scalable Delivery of Services in Interactive Multimedia System, minor in Telecommunications Public Policy) earned in June 1997. During my education, I have taken a wide variety of courses as demonstrated by my minors. My undergraduate degree also included a number of courses more typical of a degree in electrical engineering including digital logic, signal processing, and telecommunications theory.

6. One of the major concentrations of my research over the past 30+ years has been the delivery of multimedia content and data between computing devices, including various network architectures. In my research, I have studied large-scale content delivery systems, and the use of servers located in a variety of geographic locations to provide scalable delivery to hundreds or thousands of users simultaneously. I have also studied smaller-scale content delivery systems in which

content is exchanged between individual computers and portable devices. My work has emphasized the exchange of content more efficiently across computer networks, including the scalable delivery of content to many users, mobile computing, satellite networking, delivering content to mobile devices, and network support for data delivery in wireless networks.

7. In 1992, the initial focus of my research was on the provision of interactive functions (e.g., VCR-style functions like pause, rewind, and fast-forward) for near video-on-demand systems in cable systems; in particular, how to aggregate requests for movies at a cable head-end and then how to satisfy a multitude of requests using one audio/video stream broadcast to multiple receivers simultaneously. This research has continually evolved and resulted in the development of techniques to scalably deliver on-demand content, including audio, video, web documents, and other types of data, through the Internet and over other types of networks, including over cable systems, broadband telephone lines, and satellite links.

8. An important component of my research has been investigating the challenges of communicating multimedia content, including video, between computers and across networks including the Internet. Although the early Internet was used mostly for text-based, non-real time applications, the interest in sharing multimedia content, such as video, quickly developed. Multimedia-based

applications ranged from downloading content to a device to streaming multimedia content to be instantly used. One of the challenges was that multimedia content is typically larger than text-only content, but there are also opportunities to use different delivery techniques since multimedia content is more resilient to errors. I have worked on a variety of research problems and used a number of systems that were developed to deliver multimedia content to users. One content-delivery method I have researched is the one-to-many communication facility called “multicast,” first deployed as the Multicast Backbone, a virtual overlay network supporting one-to-many communication. Multicast is one technique that can be used on the Internet to provide streaming media support for complex applications like video-on-demand, distance learning, distributed collaboration, distributed games, and large-scale wireless communication. The delivery of media through multicast often involves using Internet infrastructure, devices and protocols, including protocols for routing and TCP/IP.

9. Starting in 1997, I worked on a project to integrate the streaming media capabilities of the Internet together with the interactivity of the web. I developed a project called the Interactive Multimedia Jukebox (IMJ). Users would visit a web page and select content to view. The content would then be scheduled on one of a number of channels, including delivery to students in Georgia Tech dorms delivered

via the campus cable plant. The content of each channel was delivered using multicast communication.

10. In the IMJ, the number of channels varied depending on the capabilities of the server including the available bandwidth of its connection to the Internet. If one of the channels was idle, the requesting user would be able to watch their selection immediately. If all channels were streaming previously selected content, the user's selection would be queued on the channel with the shortest wait time. In the meantime, the user would see what content was currently playing on other channels, and because of the use of multicast, would be able to join one of the existing channels and watch the content at the point it was currently being transmitted.

11. The IMJ service combined the interactivity of the web with the streaming capabilities of the Internet to create a jukebox-like service. It supported true Video-on-Demand when capacity allowed but scaled to any number of users based on queuing requested programs. As part of the project, we obtained permission from Turner Broadcasting to transmit cartoons and other short-subject content. We also connected the IMJ into the Georgia Tech campus cable television network so that students in their dorms could use the web to request content and then view that content on one of the campus's public access channels.

12. More recently, I have also studied issues concerning how users choose content, especially when considering the price of that content. My research has examined how dynamic content pricing can be used to control system load. By raising prices when systems start to become overloaded (i.e., when all available resources are fully utilized) and reducing prices when system capacity is readily available, users' capacity to pay as well as their willingness can be used as factors in stabilizing the response time of a system. This capability is particularly useful in systems where content is downloaded or streamed on-demand to users.

13. As a parallel research theme, starting in 1997, I began researching issues related to wireless devices and sensors. In particular, I was interested in showing how to provide greater communication capability to "lightweight devices," i.e., small form-factor, resource-constrained (e.g., CPU, memory, networking, and power) devices. Starting in 1998, I published several papers on my work to develop a flexible, lightweight, battery-aware network protocol stack. The lightweight protocols we envisioned were similar in nature to protocols like Bluetooth, Universal Plug and Play (UPnP) and Digital Living Network Alliance (DLNA).

14. From this initial work, I have made wireless networking—including ad hoc, mesh networks and wireless devices—one of the major themes of my research. My work in wireless network spans the protocol stack from applications through to the encoding and exchange of data at the data link and physical layers.

15. At the application layer, even before the large-scale “app stores” were available, my research looked at building, installing, and using apps for a variety of purposes, from network monitoring to support for traditional computer-based applications (e.g., content retrieval) to new applications enabled by ubiquitous, mobile devices. For example, my research has looked at developing applications for virally exchanging and tracking “coupons” through “opportunistic contact” (i.e., communication with other devices coming into communication range with a user). In many of the courses I have taught there is a project component. Through these projects I have supervised numerous efforts to develop new “apps” for download and use across a variety of mobile platforms.

16. Toward the middle of the protocol stack, my research also looked to build wireless infrastructure support to enable communication among a set of mobile devices unaided by any other kind of network infrastructure. These kinds of networks are useful either in challenged network environments (e.g., when a natural disaster has destroyed existing infrastructure) or when suitable support for network communication never existed. The deployment of such networks (or even the use of traditional network support) are critical to support services like disaster relief, catastrophic event coordination, and emergency services deployment.

17. Yet another theme is monitoring wireless networks, in particular different variants of IEEE 802.11 compliant networks, to (1) understand the

operation of the various protocols used in real-world deployments, (2) use these measurements to characterize use of the networks and identify protocol limitations and weaknesses, and (3) propose and evaluate solutions to these problems. I have successfully used monitoring techniques to study wireless data link layer protocol operation and to improve performance by enhancing the operation of such protocols. For wireless protocols, this research includes functions like network acquisition and channel bonding. Protecting networks, including their operation and content, has been an underlying theme of my research almost since the beginning of my research career. Starting in 2000, I have been involved in several projects that specifically address security, network protection, and firewalls. After significant background work, a team on which I was a member successfully submitted a \$4.3M grant proposal to the Army Research Office (ARO) at the Department of Defense to propose and develop a high-speed intrusion detection system. Key aspects of the system included associating streams of packets and analyzing them for viruses and other malware. Once the grant was awarded, we spent several years developing and meeting the milestones of the project. A number of my students worked on related projects and published papers on topics ranging from intrusion detection to developing advanced techniques to be incorporated into firewalls. I have also used firewalls, including their associated malware detection features, in developing

techniques for the classroom to ensure that students are not distracted by online content.

18. Recent work ties some of the various threads of my past research together. I have investigated content delivery in online social networks and proposed reputation management systems in large-scale social networks and marketplaces. On the content delivery side, I have looked at issues of caching and cache placement, especially when content being shared, and the cache has geographical relevance. We were able to show that effective caching strategies can greatly improve performance and reduce deployment costs. Our work on reputation systems showed that reputations have economic value, and as such, creates a motivation to manipulate reputations. In response, we developed a variety of solutions to protect the integrity of reputations in online social networks. The techniques we developed for content delivery and reputation management were particularly relevant in peer-to-peer communication and recommendations for downloadable “apps.”

19. As an important component of my research program, I have been involved in the development of academic research into available technology in the marketplace. One aspect of this work is my involvement in the Internet Engineering Task Force (IETF). The IETF is a large and open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. I have been

involved in various IETF groups including many content delivery-related working groups like the Audio Video Transport (AVT) group, the MBone Deployment (MBONED) group, Source Specific Multicast (SSM) group, the Inter-Domain Multicast Routing (IDMR) group, the Reliable Multicast Transport (RMT) group, the Protocol Independent Multicast (PIM) group, etc. I have also served as a member of the Multicast Directorate (MADDOGS), which oversaw the standardization of all things related to multicast in the IETF. Finally, I was the Chair of the Internet2 Multicast Working Group for seven years.

20. My involvement in the research community extends to leadership positions for several academic journals and conferences. I am the co-chair of the Steering Committee for the ACM Network and System Support for Digital Audio and Video (NOSSDAV) workshop and on the Steering Committees for the International Conference on Network Protocols (ICNP), ACM Sigcomm Workshop on Challenged Networks (CHANTS), and IEEE Global Internet (GI) Symposium. I have served or am serving on the Editorial Boards of IEEE/ACM Transactions on Networking, IEEE Transactions on Mobile Computing, IEEE Network, ACM Computers in Entertainment, AACE Journal of Interactive Learning Research (JILR), and ACM Computer Communications Review. I have co-chaired a number of conferences and workshops, including the IEEE International Conference on Network Protocols (ICNP), IEEE Conference on Sensor, Mesh and Ad Hoc

Communications and Networks (SECON), International Conference on Communication Systems and Networks (COMSNETS), IFIP/IEEE International Conference on Management of Multimedia Networks and Services (MMNS), the International Workshop On Wireless Network Measurement (WiNMee), ACM Sigcomm Workshop on Challenged Networks (CHANTS), the Network Group Communication (NGC) workshop, and the Global Internet Symposium, and I have served on the program committees for numerous conferences.

21. Furthermore, in the courses I have taught at UCSB, a significant portion of my curriculum covered aspects of the Internet and network communication, including the physical and data link layers of the Open System Interconnect (OSI) protocol stack, and standardized protocols for communicating across a variety of physical media such as cable systems, telephone lines, wireless, and high-speed Local Area Networks (LANs). The courses I have taught also cover most major topics in Internet communication, including data communication, multimedia encoding, and mobile application design. My research and courses have covered a range of physical infrastructures for delivering content over networks, including cable, Integrated Services Digital Network (ISDN), Ethernet, Asynchronous Transfer Mode (ATM), fiber, and Digital Subscriber Line (DSL). For a complete list of courses I have taught, see my curriculum vitae (CV) (EX2002).

22. In addition, I co-founded a technology company called Santa Barbara Labs that was working under a sub-contract from the U.S. Air Force to develop very accurate emulation systems for the military's next generation internetwork. Santa Barbara Labs' focus was in developing an emulation platform to test the performance characteristics of the network architecture in the variety of environments in which it was expected to operate, and, in particular, for network services including IPv6, multicast, Quality of Service (QoS), satellite-based communication, and security. Applications for this emulation program included communication of a variety of multimedia-based services, including video conferencing and video-on-demand.

23. In addition to having co-founded a technology company myself, I have worked for, consulted with, and collaborated with companies for nearly 30 years. These companies range from well-established companies to start-ups and include IBM, Hitachi Telecom, Turner Broadcasting System (TBS), Bell South, Digital Fountain, RealNetworks, Intel Research, Cisco Systems, and Lockheed Martin.

24. Through my graduate education, leadership with CITS, involvement in TMP, role in the development of the Internet2 infrastructure, and consulting with ISPs, I have gained a strong understanding in the role of the Internet in our society and the challenges of deploying large-scale production networking infrastructure. CITS, since its inception, has looked at the role of the Internet in society, including

how the evolution of technology have created communication opportunities and challenges, including, for example through disruptive technologies like P2P. TMP looks to focus on non-purely technical issues, including, for example, state-of-the-art business methods, strategies for successful technology commercialization, new venture creation, and best practices for fostering innovation. Through my industry collaborations and Internet2 work, I have developed significant experience in the challenges of deploying, monitoring, managing, and scaling communication infrastructure to support evolving Internet services like streaming media, conferencing, content exchange, social networking, and e-commerce.

25. I am a Member of the Association of Computing Machinery (ACM) and a Fellow of the Institute of Electrical and Electronics Engineers (IEEE).

III. COMPENSATION

26. My work on this matter is being billed at my customary rate of \$750 an hour. Also, I am being reimbursed for reasonable expenses I incur in relation to my services. I have no pecuniary interest in the outcome of this proceeding. I understand I will be paid regardless of the outcome of any proceeding in which my work is used.

IV. LEGAL CONSIDERATION

27. I am not an attorney. My understanding of the law is based on information provided by counsel for Oxylabs, UAB.

A. Anticipation

28. I understand that a reference anticipates a claim if it discloses each and every element recited in the claim, arranged as in the claim, so as to enable a person of ordinary skill in the art (POSITA) to make and use the claimed invention without the need for undue experimentation in light of the general knowledge available in the art. I understand that a U.S. Patent document is presumed to have sufficient description to include sufficient detail for a person of ordinary skill in the art to make and use the subject matter that document describes.

29. I also understand that a claim limitation may be present in an alleged prior art reference either expressly or inherently, but inherency may be established only if such limitation is necessarily present in the reference. I understand inherency cannot be established by probabilities or possibilities, and the mere fact that a certain thing may result from a given set of circumstances is not sufficient to establish inherency.

B. Claim Interpretation

30. I understand that a claim term is interpreted according to its ordinary and customary meaning as a POSITA would have understood the term in light of the surrounding claim language, other claims, the specification, and the patent's prosecution history, which are referred to as intrinsic evidence. I also understand that prior art references cited in the patent's prosecution history are considered

intrinsic evidence. I further understand that evidence outside the patent and its prosecution history (e.g., dictionaries and technical articles), may inform the context in which a POSITA would have understood the claims of a patent. I understand this ordinary and customary meaning applies absent unique circumstances, such as where a patent clearly expresses an intent to set forth a special meaning for a term.

V. OVERVIEW OF TASK & BASIS FOR OPINIONS

31. I have been asked to review the '526 Patent. I have been asked to provide opinions related to certain issues from the perspective of a person of ordinary skill, having knowledge of the relevant art, as of May 8, 2002, and—except where otherwise noted—the opinions stated in this declaration are from that perspective. The qualifications and abilities of such a person are described in paragraph 35 below.

32. More particularly, I have been informed that this proceeding is currently at a preliminary stage during which the Patent Trial and Appeal Board (the Board) will decide whether the Petition establishes a reasonable likelihood of at least one claim of the '526 Patent being invalid. I also understand that the Board may review the Petition under a stricter “compelling merits” standard that requires the Petition to make a stronger showing of invalidity; however, my analysis is not performed using this higher standard. At this preliminary stage, I have been asked to focus my analysis on the Petition’s proposed invalidity theories for the '526 Patent’s independent claims 1 and 34.

33. My opinions are based on my education, training, and experience as well as items that I reviewed to prepare my opinions, including the Petition, the documents listed in the Exhibit List included with the Petition, and any other items that I reference in my below analysis.

34. I understand that other issues may arise that require further explanation, and I will provide that explanation if appropriate. As a result, I respectfully reserve the right to update and supplement this declaration and the information and opinions provided herein.

VI. LEVEL OF SKILL IN THE RELEVANT ART

35. I understand that a patent must be analyzed from the perspective of a person of ordinary skill in the art (“POSITA”) as of the time of invention of the patent, which is typically considered to be the patent’s earliest filing date. While I do not necessarily agree with every aspect of Dr. Williams’ explanation of the level of ordinary skill in the art, I do not currently believe the aspects of that explanation I would change or clarify would impact the analysis set forth in this declaration. For purposes of this declaration, my opinions are therefore provided from the perspective of the POSITA described in paragraph 30 of Dr. Williams’ declaration (EX1008) as of May 8, 2002, which is produced below:

either: (1) a Bachelor Degree or higher in the field of electrical engineering, computer engineering, or computer science and at least

one year of related practical experience; or (2) at least four years of practical experience in data management.

Such a POSITA would have been familiar with the concepts and types of computing networks and processes described below in paragraphs.

VII. OVERVIEW OF THE PRIOR ART

36. Dr. Williams primarily relies on two cited references, each directed to different problems than those faced by the inventors of the '526 patent.

A. U.S. Patent Pub. No. 2003/0163454 (“Jacobsen”)

37. U.S. Patent Pub. No. 2003/0163454 is titled “Subject specific search engine” and lists as its first inventor Brian Jacobsen. I will refer to this reference as “Jacobsen.”

38. Jacobsen discloses a subject specific search engine (SSSE) that performs filtering of websites before they are saved in a database so that only relevant documents will be stored. EX1003, Abstract, ¶[0010]. For example, the system can be configured to collect and index only the pages deemed relevant for a specific subject, (e.g., law or medicine) similar to a law or medical library. *Id.*, ¶¶[0011]-[0020]. An example of this is shown in FIG. 3B, which depicts a filter 13 that screens information 12 such that only information 14 that passes through the filter will be indexed 15 and stored in database 9.

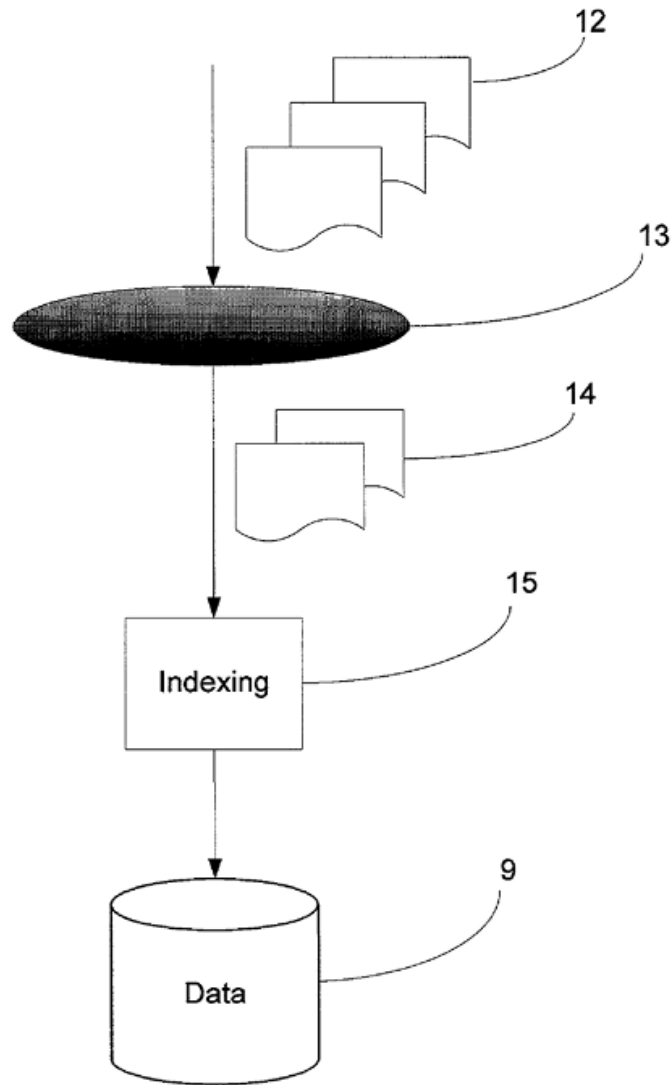


Figure 3B

Id., ¶[0047].

39. Jacobsen further discloses that the system may include a site ranking feature that will rank the most relevant websites in the database according to a user query. EX1003, ¶[0067]-[0069]. This feature is accomplished by two steps: a word scoring step and a word ranking step. *Id.* For the word scoring, each webpage in the database is parsed and assigned a word score, which counts the number of times

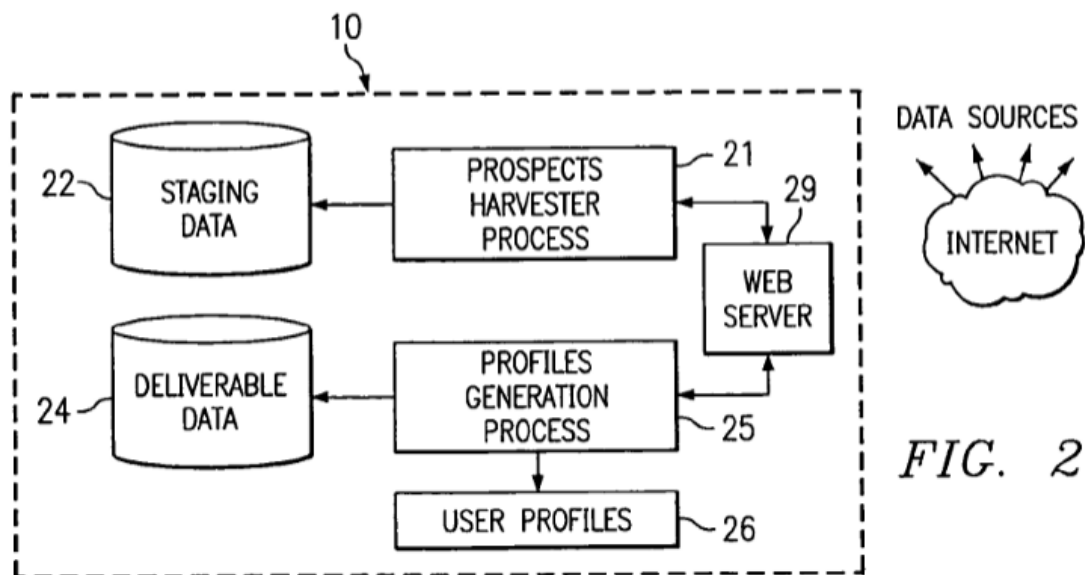
that specific words appear on the web page. *Id.* For the word ranking, a user query will be parsed to identify non-trivial words and then word scores for the corresponding non-trivial words will be summed to identify the webpages in which the words occur most often. *Id.* The webpages may be sorted based on this ranking so the page with the most matching words appears at the top of the results.

B. U.S. Patent No. 7,082,427 (“Seibel”)

40. U.S. Patent No. 7,082,427 is titled “Text indexing system to index, query the archive database document by keyword data representing the content of the documents and by contact data associated with the participant who generated the document” and lists as its first inventor John C. Seibel. I will refer to this reference as “Seibel.”

41. Seibel discloses “[a] text indexing system for collecting business intelligence about a client, as well as for identifying prospective customers of the client” EX1004, Abstract. “The general purpose of system 10 is to analyze a client's marketing and sales cycle in order to reveal inefficiencies and opportunities, then to relate those discoveries to net revenue estimates.” *Id.*, 2:60-64.

42. FIG. 2 of Seibel (reproduced below “illustrates the various functional elements of lead generator system 10,” which are “partitioned between two distinct processes.” *Id.*, 3:59-62. As shown, these two distinct processes include a prospects harvester process 21 and a profiles generation process 25. *See id.*



Prospects Harvester Process

43. The “prospects harvester process 21 uses a combination of external data sources, client internal data sources and user parameter extraction interfaces, in conjunction with a search, recognition and retrieval system, to harvest contact information from the web and return it to a staging database 22.” EX1004, 3:63-7:1. This process 21 collects information “from both inside the client’s organization and outside the organization.” *Id.* 4:1-3.

44. FIG. 3 of Seibel “illustrates ... a first embodiment of the prospects harvester” that performs the process 21 (EX1004, 2:27-28) and includes a crawler process 31 that is operable with five types of data sources: newsgroups, discussion groups, mailing lists, internal documents, and websites. *See* EX1004, 4:66-5:57. Thus, the prospects harvester process 21 may utilize a crawler process 31 (depicted

in FIG. 3) to collect information from various data sources and return it to a staging data base 22. EX1004, 3:63-4:3.

Profiles Generation Process

45. Seibel's "profiles generation process 25 analyzes the user's *in-house files* and records related to the user's existing customers to identify and group those customers into profile categories based on the customer's buying patterns and purchasing volumes." EX1004, 4:23-27. Unlike prospects harvester process 21, the profiles generation process 25 only analyzes internal information, rather than information from the internet (e.g., websites). For example, Seibel states that "prospects harvester process 21 uses a combination of external data sources [and] client internal data" (EX1004, 3:63-64), whereas Seibel's description of profiles generation process 25 is limited to analyzing "a company's *internal* client records" along with "other *internal* files" to create a profile (EX1004, 4:31-37). This emphasis on the use of internal documents is repeated throughout Seibel's disclosure. *see e.g.*, EX1004, 3:15-17 (system 10 "[e]xamines *internal* sales and marketing materials" to "generate[] profiles of the client[]")

46. A POSITA would understand such internal documents to be distinct from webpages found on the internet. *see* EX1004, 5:36-57 (distinguishing "internal documents" as being different than "web pages from internet web sites"); 9:12-17 ("Internal documents contain business information about sales, marketing, and

human resources. External sources consist primarily of the public domain in the Internet.”). Using the internal information, “[p]rofiles generation process 25 can be used to create a user (client) profiles database 26, which stores profiles of the client and its customers.” EX1004, 4:47-52; *see also id.*, 4:31-37 (process 25 mines “ a company’s *internal* client records” along with “other *internal* files” to create a profile).

47. “FIGS. 4 and 5 illustrate a database server system [41].” EX1004, 2:29-39. And this “system 41 could be used to implement the profiles generation process 25” depicted in FIG. 2. EX1004, 5:60-64. As depicted, system 41 includes input data 42 that may include various client and customer data from different *internal* data sources that deliver sales information to data mart 41b. *See* EX1004, 5:65-6:35. As explained above, the input data 42 for system 41 does not include webpages or other external documents. EX1004, 4:31-37, *see also* EX1004, 5:65-66 (describing input data as customer data), 5:48-50 (describing customer data as internal data). System 41 also includes a Predefined Query Repository Component 41c that “stor[es] predefined queries” (EX1004, 6:44-45) and a Business Intelligence Charts Repository Component 41d that “stores the most frequently visited charts” (EX1004, 6:58-62). Sever 41e can either pull charts directly from Business Intelligence Charts Repository Component 41d or can run a “predefined query” from the Predefined Query Repository Component 41c. EX1004, 6:62-7:1

VIII. OVERVIEW OF THE '526 PATENT

48. The '526 Patent discloses a data mining system 10 that can be configured for responding to customer requests for data, enterprise data analysis, competitive intelligence, trending, discovery, web portal services, clustering and taxonomy creation, or the like. EX1001, Abstract, 3:7-19. An example of the overall architecture of the system can be seen in the illustrative component diagram of FIG. 1 (below). *See Id.*

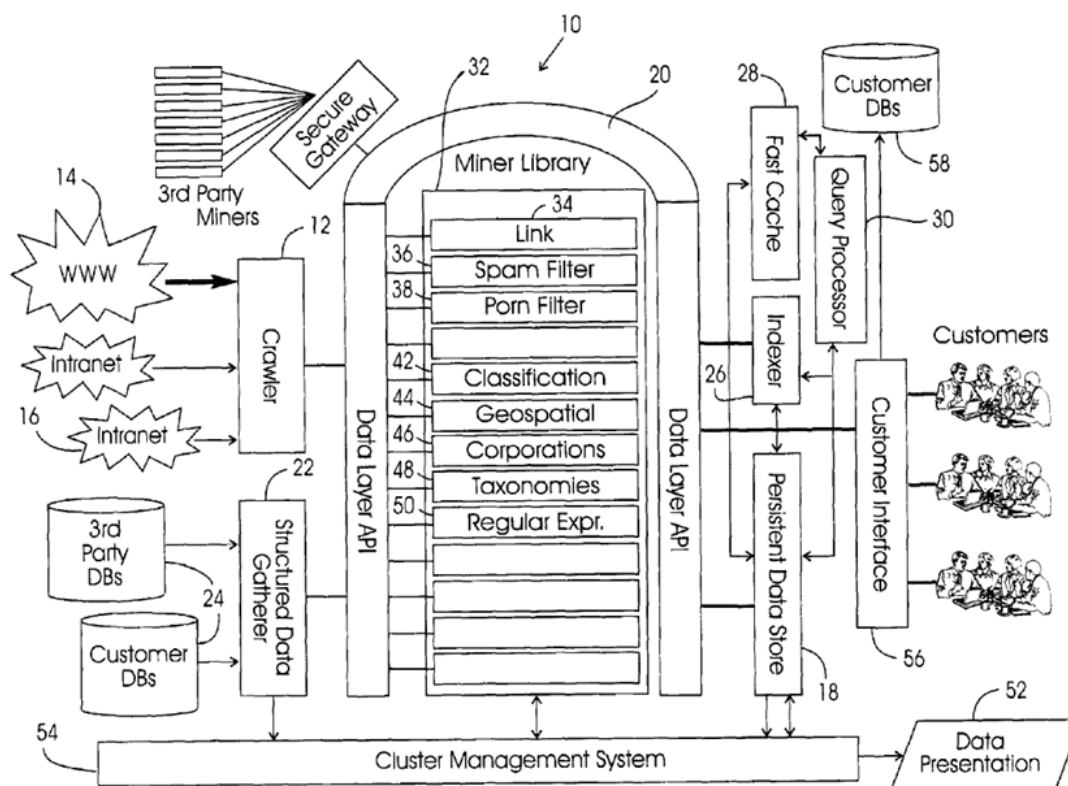


FIG 1

49. As depicted, the system 10 can include a web crawler 12 for accessing the internet or intranet to record and store web pages in a data store 18. EX1001, 3:25-38. The webpages stored in data store 18 are parsed (i.e., tokenized) to identify

the words that occur on the web pages and indexed by indexer 26. The indexer 26 may be “a conventional inverted file text indexer” that “indexes web documents and provides a keyword search application programming interface (API) to the documents.” *Id.*, 4:49-62. Thus, the indexing and tokenization of the webpages allow a user to search the data store 18 via a conventional keyword search. *See id.*, 4:54-62.

50. Additionally, a data mining layer of the system includes a plurality of data miners, which may be stored in miner library 32, that can access the webpages stored in data store 18. EX1001, 5:47-50. The data miner is “an analysis engine that generates an output, and specifically an output that can include one or more keys representing characteristics of an entity, using a set of rules.” *Id.*, 5:63-66. Examples of some of these miners include:

a link miner 34 which returns links to/from a page, a spam filter 36 for identifying “spam” in the data store 18, a porn filter 38 for identifying pornographic pages in the data store 18, a classification miner 42 that classifies pages based on the occurrence of patterns of terms in the pages, a geospatial miner 44 which identifies any geographic information on a Web page, a corporations miner 46, a taxonomies miner 48 that returns pages having a predefined taxonomy category, a regular expression (regex) miner 50 that provides a stream of pages containing a defined regex, and so on.

Id., 5:52-63.

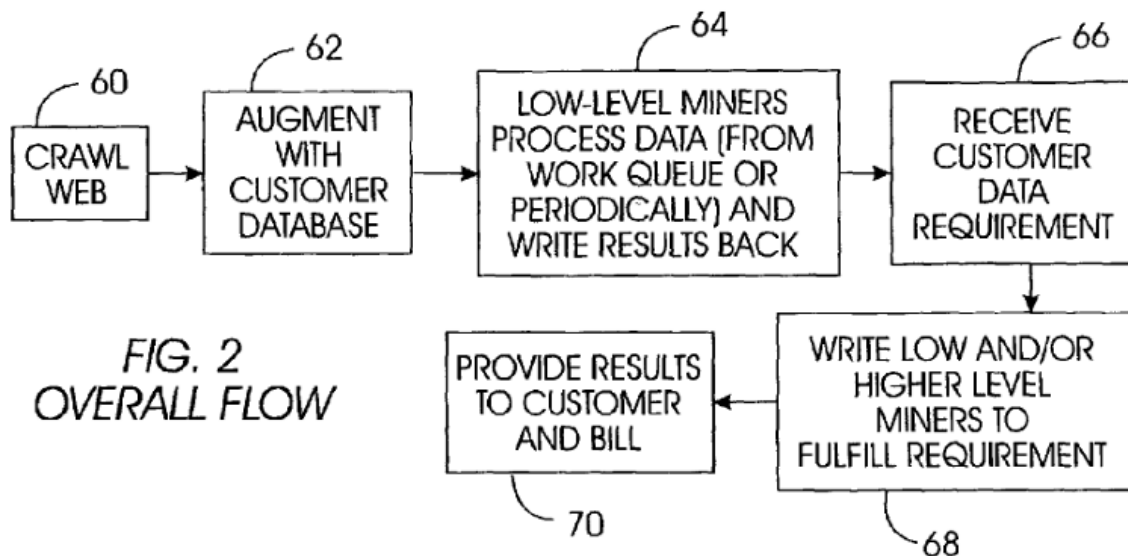
51. The miners can access the webpages in the datastore to extract new data, relations, or information from the webpages. *See* EX1001, 6:25-40. Some of

the miners can be classified “low level” miners that analyze the tokenized, indexed webpages and write back the results of their processing to the data store 18 for other miners and end users to access. *See* EX1001, 7:28-33, 8:18-27. Some “high level” miners can process data that includes the outputs of other miners. *Id.*, 8:32-39 (high level miner may “require entities for processing that have been tagged with keys output by lower level miners”).

52. The described miners can be pre-existing (e.g., created during implementation of the mining system or during a previous operations) or the system can create new miners during operation, such as in response to a customer request. EX1001, 8:32-39. As an example, a high level miner can be created in response to the query “what do our commercial customers in the Pacific Northwest think of our competitor's health care products in terms of brand name strength and value?” EX1001, 8:40-44. This created miner “might specify that it wishes to receive only pages from the Pacific Northwest, as indicated by a geospatial key appended to entities by the geospatial miner, and only if the competitor's name is featured in the entity, as indicated by a key appended to the entity by a proper name miner.” *Id.*, 8:44-54. Results of the miner’s data processing can then be sent to a customer. *Id.*, 8:54-56.

A. Independent Claims 1 and 34

53. Independent claim 1 of the '526 Patent is directed to “a method for data mining.” EX1001, 11:63. Figure 2 of the '526 Patent, below, illustrates a partial flowchart diagram of an illustrative method of data mining for some of the steps covered by independent claim 1.



See EX1001, 8:13-39.

54. As illustrated, and with reference to FIG. 1, a crawler 12 may crawl the web to add webpages to a data store 18. *Id.* Once within the data store, at least some data miners may process the stored webpages to include the results of the data miners. EX1001, 8:18-31. A customer may submit a request for accessing the stored webpages, such as via customer interface 56. *Id.*, 7:39-42, 8:32-33. In response to the request, high level miners can be created and can process the stored webpages and results from data miners that are associated with the webpages. *Id.*, 8:33-39.

The stored webpages will be modified to include the output of the created data miner and results will be provided to the customer. *Id.*, 8:33-56.

55. Claim 1 recites **[1a]** a method for data mining, comprising:
- [1b]** crawling the Internet to return Web pages to a data store;
 - [1c]** accessing the data store with at least a first data miner to process data in the store and output at least one result;
 - [1d]** receiving a customer request for information;
 - [1e]** creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output;
 - [1f]** at least partially based on the output of the second data miner, responding to the customer request for information.

EX1001, 11:63-12:7.

56. Claim 34 is directed to an information service that includes limitations similar to the Limitations [1a]-[1f] of claim 1 and recites [34a] An information service, comprising:

- [34b]** means for accessing, with at least a first data miner, a data store containing information representing at least one Web page to process data in the store and output at least one result;
- [34c]** means for receiving a customer request for information;
- [34d]** means for creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output; and

[1e] means for responding to the customer request for information at least partially based on the output of the second data miner.

EX1001, 13:42-54.

B. Prosecution History

57. During prosecution of the '526 Patent, claims 1 and 34 were rejected as being anticipated by Chipman (EX1010 – U.S. Patent Pub. No. 2002/0091923). *See* EX1002, 89. In response to the rejection, Applicant amended both claims 1 and 34 to recite replace the term “establishing” with “creating” for Limitations [1e] and [34d]. EX1002, 74, 79 (“~~establishing~~ creating at least a second data miner ...”). Applicant also provided remarks that “no new miners are created” in Chipman. EX1002, 83. Instead, Chipman discloses that “the request is passed to pre-existing search engines” and “picking a pre-existing search engine” fails to disclose “creating” a miner in response to the customer request for information. EX1002, 83.

IX. CLAIM CONSTRUCTION

58. I understand claims are to be given their ordinary and customary meaning to a POSITA in view of the claims, specification, and prosecution history.

A. Creating at least a second data miner in response to the customer request for information ...

59. Claims 1 and 34 recite “...creating at least a second data miner in response to the customer request for information...”. *See* §VIII.A.

60. The Petition and Dr. Williams contend that the prosecution history of the '526 Patent would change how a POSITA would interpret the above recited limitation. Pet., 13-14; EX1008, ¶49¹. Specifically, Dr. Williams recites that based on the prosecution history (*See* §VIII.B), “a POSA would understand the limitations of claims 1 and 34 to require creating a new data miner to search the results of a first data miner to provide customized results to the user based on the customized parameters of the user search request.” EX1008, ¶49. Dr. Williams also cites to Merriam-Webster’s definition of creating to support his interpretation. EX1008, ¶49 (citing EX1007, 2) (defining “creating” as to make or bring into existence something new.”). When comparing Dr. Williams’ proposed construction to the original claim language, it is apparent that Dr. Williams omits the requirement that the second data miner is created “in response to the customer request for information.” *See* EX1008, ¶49 (creating at least a second new data miner ~~in response to the customer request for information~~ to search the results of a first data miner to provide customized results to the user based on the customized parameters of the user search request)

¹ The Petition and Dr. Williams’ declaration (EX1003) are virtually identical. Although I primarily address Dr. Williams’ declaration herein, these remarks would similarly apply to the Petition.

(emphasis added by underlining new text and striking-through deleted text found in claim 1). Dr. Williams' construction is mistaken.

61. Dr. Williams' proposed construction is contrary to how a POSITA would have interpreted the claim in light of the specification and prosecution history. To illustrate, as described with reference to FIG. 2 of the '526 Patent, "customer requests for information can be received" at block 66 and then at block 68, additional low level miners can be written in response, or high level miners, if required and not yet written, can be created." EX1001, 8:32-35. Referring to the prosecution history, "creating" a data miner was distinguished from the process of "picking a pre-existing search engine." Neither of these disclosures would lead a POSITA to interpret the claimed feature by ignoring the requirement that the second data miner is created "in response to the customer request for information." Additionally, the dictionary definition of "creating" fails to show how a POSITA would arrive at Dr. Williams' proposed construction as it is only relevant to the term "creating" and not the remaining limitations in the claim. *See* EX1007, 2.

62. Dr. Williams provides no explanation or connection between the prosecution history (or dictionary definition) and his proposed construction. There is simply no reasoning or evidence that support Dr. Williams' proposed changes to the claim language and, thus, a POSITA would not have interpreted the claim as Dr. Williams proposes.

X. JACOBSEN DOES NOT DISCLOSE EACH CLAIM FEATURE

63. Dr. Williams’ analysis of claim 1 and claim 34 is identical. *Compare* EX1008, ¶¶70-83 *with* EX1008, ¶¶90-101. Therefore, while I only explicitly address the deficiencies of Dr. Williams’ analysis of claim 1, these deficiencies must also apply to claim 34.

A. [1e], [34d]: Jacobsen Does Not Create a Data Miner In Response to a Customer Request for Information

64. Both claims 1 and 34 require creating a second data miner in response to the customer request for information. *See supra*, §VIII.A. Both claims also require a second data miner accessing at least the result of the first data miner to generate an output. *See supra*, §VIII.A. Jacobsen does not expressly disclose the term “data miner.” Instead, Dr. Williams identifies Jacobsen’s “site ranking feature” as allegedly corresponding to the second data miner. EX1008, ¶79, 97 (“second data miner, e.g., site ranking feature”).

1. *Creating ... in response to the customer request for information*

65. As explained above in Section IX.A, Limitations [1e] and [34d] require “creating at least a second data miner *in response to the customer request for information*” and **not** creating a data miner to provide customized results to the user based on parameters of a search request. Dr. Williams fails to address how Jacobsen creates this site ranking feature *in response to the customer request for information*. EX1008, ¶¶79-81. Dr. Williams’ only discussion of “creating” a data miner is found

in a single sentence: “Jacobsen discloses creating a second data miner, e.g., site ranking feature, in response to the customer request for information.” EX1008, ¶79. However, Dr. Williams does not cite to any portion of Jacobsen to support this assertion. EX1008, ¶79. Further, the portions of Jacobsen (EX1003, ¶¶[0026], [0069]) cited by Dr. Williams do not disclose creating the site ranking feature. Accordingly, there is nothing in Dr. Williams’ declaration that would lead a POSITA to understand that Jacobsen discloses creating the site ranking feature in response to the customer request for information. Therefore, Jacobsen does not teach or disclose every feature of claims 1 and 34.

66. In addition to Dr. Williams’ deficient analysis, a POSITA would have understood that Jacobsen’s site ranking feature does not teach creating at least a second data miner in response to the customer request for information. Jacobsen’s site ranking feature (the alleged second data miner) is not created in response to a keyword search query (the alleged customer request). A POSITA would have understood that Jacobsen’s site ranking feature is a part of the ranking capability of conventional search engines. *Compare* EX1003, ¶[0069] (describing the site ranking feature as determining which pages are most relevant by counting the number of times a keyword appears) *with* EX1001, 1:19-31 (describing relevancy rankings of conventional search engines as determining which pages are most relevant by keyword frequency of occurrence). In both cases, a POSITA would have

understood that ranking feature allows pages to be sorted by the number of times (i.e., frequency) a keyword from the query occurs within each page. *Id.*

67. Jacobsen does not disclose how its site ranking feature is implemented within its subject specific search engine. *See* EX1003, ¶[0069]; EX2003, 4-5 (“How systems carry out their search and matching functions differs according to which theoretical model of information retrieval underlies the system's design philosophy.”). Further, although not alleged by Dr. Williams, the site ranking feature could have been implemented in several different ways and a POSITA would **not** find Jacobsen’s silence sufficient for an inherent disclosure. EX2003, 4-5. Thus, Jacobsen does not disclose, expressly or inherently, creating a new scoring algorithm in response to a keyword search query.

68. To illustrate this point, one possible method of implementing this site ranking feature is via a scoring algorithm. EX2003, 5. Yet, a POSITA would understand that this scoring algorithm would not be created in response to the user query and would have been created during implementation of Jacobsen’s search engine. *See id.* For example, if a user entered several keyword search queries at various times, all queries would be passed to the same scoring algorithm and Jacobsen would not create a new scoring algorithm for each query. *See* EX2004, 32 (Figure 1.3 disclosing a software architecture in which the ranking/scoring algorithm would process multiple user search queries). In other words, all queries are passed

to the *existing* scoring algorithm. See EX2004, 31-32 (disclosing software architecture—including a ranking algorithm—that is required to implement the retrieval process). Thus, a POSITA would not understand the use of a ranking feature—such as the site ranking feature disclosed in Jacobsen—as teaching a new scoring algorithm in response to a keyword search query.

2. *Second Data Miner*

69. As discussed above, Dr. Williams identifies Jacobsen’s “site ranking feature” as allegedly corresponding to the second data miner. EX1008, ¶¶79, 97. Yet, Dr. Williams provides no explanation of how this “site ranking feature” constitutes a “data miner.” Further, as explained below, a POSITA would have understood Jacobsen’s “site ranking feature” to be different from the claimed “second data miner.”

70. As explained in the background section of the ’526 Patent, keyword searching and ranking were known functions of existing search engines. EX1001, 1:19-31. The ’526 Patent describes that existing search engines “access Web pages and ... generat[e] an inverted index of words that occur on the Web pages.” *Id.*, 1:19-23. The “[q]ueries are responded to by accessing the index using the requested key words as entering arguments, and then returning from the index the URLs that satisfy the queries. The page identifications that are returned are usually ranked by relevance using, e.g., link information or key word frequency of occurrence.” *Id.*,

1:25-31. Dr. Williams acknowledges this disclosure of a search engine's conventional ranking feature as allegedly being "Applicant Admitted Prior Art." EX1008, ¶¶52-53.

71. The '526 Patent describes its invention against this backdrop of the keyword searching and ranking features of conventional search engines. For example, the '526 Patent describes that its system 10 also includes keyword search functionality that it attributes to indexer 26—which is described as being distinct from the data miners (e.g., 34-50) and the data miner library 32. *See* EX1001, 4:49-5:9 (describing indexer 26 and miners working together). For example, in a preferred embodiment, the "indexer 26 is a generalization of a conventional inverted file text indexer" that "indexes web documents and provides a keyword search application programming interface (API) to the documents." EX1001, 4:49-62. Thus, a POSITA would understand that the miners described in the '526 Patent do more than simply provide a keyword search and ranking feature, as this functionality is attributed to a different component (e.g., indexer 26). Further, the '526 Patent describes how using the disclosed data miner would provide a different result than using a conventional keyword search when responding to the same query. *Compare* EX1001, 1:46-59 (describing analysis of a question using a conventional keyword search in the **background** of the '526 Patent) *with* 8:40-56 (describing analysis of the same question using data miners).

72. To further illustrate, the terms “miner” and “data mining element” are described as an “analysis engine that generates an output, and specifically an output that can include one or more keys representing characteristics of an entity, using a set of rules.” EX1001, 5:63-66. Dr. Williams fails to explain why a POSITA would understand Jacobsen’s site ranking feature to read on the claimed data miner—despite the ’526 Patent distinguishing its data miners from the ranking features of conventional search engines. *See* EX1001, 1:25-31. Thus, the Petition cannot establish that a POSITA would understand Jacobsen’s site ranking feature to teach the claimed second data miner.

73. Further, a POSITA would understand the claimed miners to be distinct from the sorting/ranking algorithms used in the keyword search functionality of conventional search engines. For example, the Modern Information Retrieval Textbook, published in 1999, provides a definition for “data mining” as “data mining: extraction of new data, relations, or partial information from any type of data.” EX2004, 73. A POSITA would not understand sorting pages based on the number of times a word occurs as extracting new data, relations, or partial information. Thus, a POSITA would understand that the keyword search and ranking functionality (i.e., data retrieval and sorting) is not the same as the claimed data miners.

XI. SEIBEL DOES NOT DISCLOSE EACH CLAIM FEATURE

74. Dr. Williams’ analysis of claim 1 and claim 34 is largely identical. *Compare* EX1008, ¶¶165-190 *with* EX1008, ¶¶209-228. Therefore, while I explicitly address the deficiencies of Dr. Williams’ analysis of claim 1, these deficiencies must also apply to claim 34.

A. [1b]-[1c], [34b]: Petitioner’s Mapping of the Data Store is Inconsistent with Seibel’s Disclosure

75. Limitation [1b]-[1c] recite “crawling the Internet to return Web pages to a data store” and accessing the data store with a first data miner. *See supra*, §VIII.A. Limitation [34b] recites “means for accessing, with at least a first data miner, a data store containing information representing at least one Web page.” *Id.* For claim 1, Dr. Williams contends that Seibel teaches these limitations because “Seibel discloses crawling the Internet using processes 21, 31 to return Web pages to a data store, e.g., database server system 41.” EX1008, ¶175; *see also* EX1008, ¶210 (“The processes of the lead generation system 10 include crawling the Internet using a process 21 or crawler process 31 and then storing the results in a data store, e.g., database server system 41.”). This is incorrect.

76. There is no disclosure in Seibel that prospects harvester process 21 returns Web pages to database server system 41. *See generally* EX1004; *see also* §VII.B. Further, Dr. Williams provides no citation to support this assertion and, instead, relies on disparate citations from two distinct processes of Seibel without

providing any explanation as to why these processes are connected. *See* EX1008, ¶¶171-175; *see also* §VII.B (Seibel describes process 21 and system 41 as distinct aspects of “two distinct processes”). As acknowledged by Petitioner, the prospects harvester process 21 “harvest[s] contact information from the web and ***return[s] it to a staging data base 22.***” Pet., 53 (quoting EX1004, 3:63-4:1). Seibel never discusses this staging data base 22 outside of this quote. *See generally* EX1004. Because there is no disclosure in Seibel that equates staging data base 22 and system 41, a POSITA would not understand prospects harvester process 21 to return anything—let alone web pages—to the database server system 41.

77. Further, even considering Dr. Williams’ faulty assumption that prospects harvester process 21 and server system 41 are somehow connected, there is nothing in the record that would lead a POSITA to believe that prospects harvester process 21 is returning web pages to this system. Dr. Williams’ theory rests on the misguided assumption that the prospects harvester process 21 and the profiles generation process 25 are a part of the same process. *See* EX1008, ¶¶170-175. But this notion is explicitly rebutted by Seibel, which “partition[]s” these “two distinct processes.” EX1004, 3:59-62 (stating the “functions of system 10 are partitioned between two distinct processes” and immediately describing the prospects harvester process 21 and the profiles generation process 25). Yet, Dr. Williams’ analysis of Limitation [1b] fails to connect these two distinct processes disclosed in Seibel.

78. For example, after discussing the prospects harvester process 21 (FIG. 2) and data sources for the prospects harvester process 21, 23 (FIG. 3) in paragraphs 171-173 of his Declaration, Dr. Williams abruptly turns to database server system 41 (FIG. 4) in paragraphs 174 without making any connection between these two processes. Seibel is clear that database server system 41 which is “used to implement the profiles generation process 25” and that this profiles generation process 25 is distinct from the prospects harvester process 21. EX1004, 3:59-62, 5:60-64.

79. Further, to the extent Dr. Williams intends to assert that acquisition component 41a (*see* EX1008, ¶174) crawls the internet to return webpages to system 41, Seibel does not disclose this. As explained in Section VII.B, profiles generation process 25 (implemented by system 41) only analyzes internal documents. EX1004, 4:31-37 (process 25 analyzes “ a company’s *internal* client records” along with “other *internal* files” to create a profile). A POSITA would understand such internal documents to be distinct from webpages found on the internet. EX1004, 5:36-57 (distinguishing internal documents from web pages from internet web sites); 9:12-17 (same).

80. This is confirmed by Seibel’s FIG. 4, which illustrates data acquisition component 41a receiving input data 42, which “can be the client's sales data, customer-contact data, customer purchase data and account data etc.” EX1004 5:65-66. This “client” and “customer” data is described as internal data. EX1004, 4:31-

37 (process “mine[s] a company’s internal client records”), 5:48-50 (“Other internal Sources are customer information, sales records, accounting records”). To further illustrate, none of the “[v]arious data sources” described with respect to input data 42 are webpages from the internet. *See* EX1004, 5:67-6:10. The fact that data acquisition component 41a may “extract[] data from different data sources” does not mean that one of these data sources would be webpages from Internet web sites. In fact, a POSITA would understand that Seibel expressly describes the data sources 42 as being distinct from webpages.

81. Thus, Dr. William’s assertion that “Seibel discloses crawling the Internet using processes 21, 31 to return Web pages ... database server system 41” (EX1008, ¶175) is contrary to the teachings of Seibel and the understanding of a POSITA. Although, Dr. Williams alters this statement in his analysis of Limitation [34b], this does not change his analysis. *See* EX1008, ¶210 (“The processes of the lead generation system 10 include crawling the Internet using a process 21 or crawler process 31 and then storing the results in a data store, e.g., database server system 41.”). Dr. Williams still maps to prospects harvester process 21 and staging data base 22 as corresponding to the data store containing information representing at least one web page (EX1008, ¶¶211-212) and then jumps to profiles generation process 25 as corresponding to the first data miner that accesses this data store

(EX1008, ¶¶213-214). For the same reasons discussed above, this is inconsistent with Seibel's teachings.

B. [1e], [34d]: Seibel Does Not Create a Data Miner In Response to a Customer Request for Information

82. The claims require creating a second data miner in response to the customer request for information. *See supra*, §VIII.A. Petitioner identifies Seibel's "server 41e" as allegedly corresponding to the second data miner. Pet., 57-58.

83. As explained above in Section IX.A, Limitations [1e] and [34d] require "creating at least a second data miner *in response to the customer request for information*" and not creating a data miner to provide customized results to the user based on parameters of a search request. Dr. Williams fails to address how Seibel creates this server 41e *in response to the customer request for information*. EX1008, ¶182-186. Dr. Williams' only discussion of "creating" a data miner is found in a single conclusory sentence: "Seibel discloses creating at least a second data miner in response to the customer request for information." EX1008, ¶182, 219. Dr. Williams does not cite to any portion of Jacobsen to support this assertion, but adds an example in the following sentence which adds that "Seibel discloses a second data miner, e.g., server 41e runs a query using different parameters (e.g., a quarterly revenue query where the parameters passed are the specific region, the specific year, and the specific quarter) to extract data from data mart 41b to feed the data to the requested chart." *Id.*, (citing EX1004 at 6:44-7:2). However, Dr.

Williams fails to explain how a POSITA would understand this cited portion of Seibel to disclose creating a data miner.

84. To start, server 41e is a web server “such as Microsoft IIS or Apache server or any other commercially available web server.” EX1004, 7:25-15. A POSITA would understand such a server to be a computing device that is part of (or operates with) the architecture of the lead generator system. For example, the web server may be an application such as an Apache server, installed on computing hardware and, thus, the web server is created when it is installed or implemented. *See* EX2005, 1, 8. Consequently, a POSITA would understand this web server 41e is not created in response to the customer request for information.

85. To the extent Dr. Williams intends to assert that server 41e is what is creating a second data miner in response to the customer request for information, this opinion is unclear and unexplained. *See* EX1008, ¶183 (describing Server 41e as what “extracts (e.g., mines) data”), 185 (describing “server 41e” as one of multiple “data miners.”). The only other aspect of Seibel that is included in Dr. Williams’ analysis is that server 41e runs a query. EX1008, ¶182 (citing EX1004, 6:44-7:2). This citation describes system 41 may include a “***Predefined*** Query Repository Component 41c [that] is the central storage for ***predefined*** queries.” EX1004, 6:44-45. But a POSITA would not understand running a ***predefined*** query to disclose creating a data miner.

86. These predefined queries can be used by server 41e to obtain charts in the event the desired chart is not stored in the chart repository 41d, as described in Seibel:

When an end user requests a chart, system 41 first queries the chart repository 41d to see if there is an existing chart. If there is a preexisting chart, server 41e pulls that chart directly from the repository. If there is no preexisting chart, server 41e runs the corresponding predefined query from the query repository 41c in order to extract data from data mart 41b and Subsequently feed the data to the requested chart.

EX1004, 6:62-7:2. Thus, when server 41e runs a query—as alleged by Petitioner—the “server 41e *runs the corresponding predefined query* from the query repository 41c ...” See EX1004, 6:62-7:2.

87. A POSITA would not understand running this predefined query as creating a second data miner in light of the specification and prosecution history. As explained in Paragraphs 49 and 52-53 of Dr. Williams declaration, Applicant overcame a rejection based on the Chipman reference (EX1010) during prosecution by arguing that Chipman did not disclose “creating” a data miner. See also EX1002, 83. Specifically, Applicant argued that “picking a pre-existing search engine” or passing a request to “pre-existing search engines” failed to disclose the “creating” limitation. EX1002, 83. Here, Seibel similarly discloses storing predefined queries and then running (i.e., picking) the “corresponding predefined query from the query

repository 41c.” EX1004, 6:62-7:2. This “predefined” query would have been created in advance of (i.e., before) server 41e runs the query. EX2006, 1 (“predefined” means being “defined *in advance*”). Thus, similar to the arguments made during prosecution, a POSITA would not have understood picking a predefined query to disclose “creating a second data miner in response to the customer request for information,” as claimed.

XII. CONCLUSION

88. All statements made herein of my own knowledge are true, all statements made herein on information and belief are believed to be true, and these statements were made with the knowledge that willful false statement and the like are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001.

Respectfully submitted,

Dated: January 5, 2024



Kevin C. Almeroth, Ph.D.