Paper No. 1

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

BMJ PUBLISHING GROUP LIMITED,

Petitioner,

v.

EVIDIUM, INC.,

Patent Owner.

Case No. IPR2027-00023

U.S. Patent No. 11,869,674

PETITION FOR *INTER PARTES* REVIEW

TABLE OF CONTENTS

I. INTRODUCTION 1
II. COMPLIANCE WITH IPR REQUIREMENTS 1
A. Standing 1
B. Payment of Fees 2
III. IDENTIFICATION OF CHALLENGED CLAIMS 2
IV. TECHNICAL BACKGROUND 3
A. Medical Ontologies and Terminologies 3
B. Healthcare Knowledge Graphs 4
C. Clinical Decision Support Systems 5
D. Natural Language Processing 5
V. THE '674 PATENT 6
A. AIA Status 6
B. Level of Skill in the Art 6
C. Overview of the '674 Patent 7
D. Prosecution History 9
VI. CLAIM CONSTRUCTION 10
VII. SUMMARY OF THE ASSERTED REFERENCES 10
A. Datla 11
B. Goldberg 11
C. Dang 12
D. Zhong 13
E. Van Assel 13
F. Kim 14
VIII. UNPATENTABILITY GROUNDS 14
A. Ground 1: Claims 1-3, 7-9, and 11-15 are Unpatentable Over Datla, Dang, and Goldberg 14
1. Motivation to Combine 15
2. Independent Claim 1 17

3. Claim 2 30
4. Claim 3 31
5. Claim 7 32
6. Claim 8 34
7. Claim 9 36
8. Claim 11 37
9. Independent Claim 12 41
10. Claim 13 46
11. Claim 14 47
12. Claim 15 47

B. Ground 2: Claims 4-5 are Unpatentable Over Datla, Dang, Goldberg, and Kim 48
1. Motivation to Combine 48
2. Claim 4 49
3. Claim 5 51

C. Ground 3: Claim 6 is Unpatentable Over Datla, Dang, Goldberg, and Van Assel 52

D. Ground 4: Claims 16-18 and 20-23 are Unpatentable Over Datla, Dang, and Van Assel 54
1. Motivation to Combine 54
2. Independent Claim 16 56
3. Claim 17 60
4. Claim 18 60
5. Claim 20 61
6. Claim 21 62
7. Claim 22 62
8. Claim 23 63

E. Ground 5: Claim 19 is Unpatentable Over Datla, Dang, Van Assel, and Kim 63

F. Ground 6: Claims 24-25 and 30 are Unpatentable Over Zhong

and Datla....64
1. Motivation to Combine....64
2. Independent Claim 24....65
3. Claim 25....74
4. Claim 30....75
G. Ground 7: Claim 26 is Unpatentable Over Zhong, Datla, and Dang....76
H. Ground 8: Claims 27 and 29 are Unpatentable Over Zhong, Datla, and Van Assel....77
1. Motivation to Combine....77
2. Claim 27....78
3. Claim 29....78
I. Ground 9: Claim 28 is Unpatentable Over Zhong, Datla, and Kim....79
IX. CONCLUSION....80

TABLE OF AUTHORITIES

Page(s)

Statutes

35 U.S.C. § 102 ..... *passim*

35 U.S.C. § 103 2

35 U.S.C. § 315 1

35 U.S.C. § 363 14

Regulations

37 C.F.R. § 42.10 vi

37 C.F.R. § 42.15 2

37 C.F.R. § 42.24 81

37 C.F.R. § 42.100 10

37 C.F.R. § 42.102 1

LISTING OF EXHIBITS

Exhibit No.	Description
1001	U.S. Patent No. 11,869,674 ("the '674 patent")
1002	File History of the '674 patent
1003	Declaration of Dr. Mark Alan Musen, M.D., Ph.D.
1004	Curriculum Vitae of Dr. Mark Alan Musen, M.D., Ph.D.
1005	U.S. Patent Publication No. 2019/0252074 ("Datla")
1006	U.S. Patent No. 11,636,949 ("Goldberg")
1007	U.S. Patent No. 12,248,826 ("Gnanasambandam")
1008	U.S. Patent Publication No. 2022/0208373 ("Dang")
1009	U.S. Patent No. 11,194,860 ("Zhong")
1010	Eryu Xia et al., "Mining Disease-Symptom Relation from Massive Biomedical Literature and Its Application in Severe Disease Diagnosis," IBM Research (2018) ("Xia")
1011	U.S. Patent Publication No. 2021/0233658 ("Van Assel")
1012	U.S. Patent Publication No. 2024/0221952 ("Kim")
1013	RESERVED
1014	RESERVED
1015	*Evidium, Inc. v. BMJ Publishing Group Limited*, Case No. 4:26-cv-05050-AMO (N.D. Cal.) (Dkt. 1) (May 28, 2026) (Complaint – Redacted Version)
1016	*Evidium, Inc. v. BMJ Publishing Group Limited*, Case No. 4:26-cv-05050-AMO (N.D. Cal.) (Dkt. 30) (Sept. 3, 2026) (Amended

Exhibit No.	Description
	Complaint – Redacted Version)
1017	*Evidium, Inc. v. BMJ Publishing Group Limited*, Case No. 4:26-cv-05050-AMO (N.D. Cal.) (Dkt. 30-15) (Amended Complaint Ex. O – Claim Chart for U.S. Patent No. 11,869,674)
1018	*Evidium, Inc. v. BMJ Publishing Group Limited*, Case No. 4:26-cv-05050-AMO (N.D. Cal.) (Dkt. 15) Proof of Service of Summons
1019	RESERVED
1020	U.S. Patent Publication No. 2022/0310260 ("Ullaskrishnan")
1021	U.S. Patent Publication No. 2019/0026437 ("Syeda-Mahmood")
1022	U.S. Patent No. 12,512,193 ("Marcjan")
1023	U.S. Patent Publication No. 2023/0377748 ("Yang")
1024	U.S. Patent No. 11,488,713 ("Velez")
1025	U.S. Patent Publication No. 2018/0108443 ("Li")
1026	U.S. Patent No. 12,014,135 ("Potts")
1027	U.S. Patent No. 7,260,480 ("Brown")
1028	Dan McCreary, Knowledge Graphs: The Third Era of Computing, Medium (Mar. 23, 2019) (https://dmccreary.medium.com/knowledge-graphs-the-third-era-of-computing-a8106f343450)
1029	"DocKG: A Knowledge Graph Framework for Health with Doctor-in-the-Loop," Ming Sheng et al. ("Sheng")
1030	"Learning a Health Knowledge Graph from Electronic Medical Records," Maya Rotmensch et al. (July 20, 2017) ("Rotmensch")

Exhibit No.	Description
1031	"Semantic Health Knowledge Graph: Semantic Integration of Heterogeneous Medical Knowledge and Services," Longxiang Shi et al. (Feb. 12, 2017) ("Shi")
1032	"Clinical Decision-Support Systems," Mark A. Musen et al. (2021) ("Musen Textbook")
1033	"Artificial Intelligence in Medical Diagnosis," Peter Szolovits et al. (1988) ("Szolovits")
1034	"*Internist-1*, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine," Randolph A. Miller et al. (1982) ("Miller")
1035	"INTERNIST-1 Properties: Representing Common Sense and Good Medical Practice in a Computerized Medical Knowledge Base," Fred E. Masarie et al., Computers and Biomedical Research (1985) ("Masarie")
1036	"Towards computer-assisted maintenance of medical knowledge bases," Dario A. Giuse, Artificial Intelligence in Medicine (1990) ("Giuse")
1037	"Automated Encoding of Clinical Documents Based on Natural Language Processing," Carol Friedman et al. (2004) ("Friedman")
1038	SNOMED CT Starter Guide (Version 2014-12-02) (Dec. 2014) (https://web.archive.org/web/20160718140111/http://doc.ihtsdo.org/download/doc_StarterGuide_Current-en-US_INT_20141202.pdf)
1039	About LOINC, LOINC from Regenstrief (https://web.archive.org/web/20220125211233/https://loinc.org/about)
1040	Mark Musen et al., "The National Center for Biomedical Ontology," Journal of the American Medical Informatics

Exhibit No.	Description
	Association (2012)
1041	Natalya Noy et al., "BioPortal: ontologies and integrated data resources at the click of a mouse," Nucleic Acids Research, 2009, Vol. 37 (May 2009)
1042	History of The BMJ (https://www.bmj.com/about-bmj/history-of-the-bmj)
1043	About – PubMed (https://web.archive.org/web/20200309080055/https://pubmed.ncbi.nlm.nih.gov/about/)
1044	BioPortal API Documentation (https://web.archive.org/web/20140725213752/http://data.bioontology.org/documentation)
1045-1049	RESERVED
1050	SNOMED CT and Context (https://docs.snomed.org/implementation-guides/context-representation-implementation-guide/4-snomed-ct-and-context)

MANDATORY NOTICES

1. Real Party in Interest

Petitioner BMJ Publishing Group Limited, which is wholly owned by the British Medical Association, is the real party-in-interest. No other party had access to or control over the Petition, and no other party funded or participated in preparation of this Petition.

2. Related Proceedings

The ’674 patent is related to the following U.S. Patents: 12,518,882 (“’882 patent”), 12,518,883 (“’883 patent”), 12,300,392, 12,525,362, 12,542,217, 12,640,274, and 12,725,716, and to pending U.S. Patent Application No. 19/722,113.

The related ’882 and ’883 patents are at issue in PGR2027-00001 and PGR2027-00002, respectively.

The ’674 patent and certain related patents are at issue in *Evidium, Inc. v. BMJ Publishing Group Limited*, Case No. 4:26-cv-05050-AMO (N.D. Cal.). The complaint (EX1015) was filed on May 28, 2026, and served on June 8, 2026 (EX1018).

3. Counsel

Petitioner identifies the following lead and back-up counsel:

Lead Counsel	Back-Up Counsel
Angelo J. Christopher	Seth D. Levy

Reg. No. 81,594 NIXON PEABODY LLP 70 West Madison St., Suite 5200 Chicago, IL 60602-4378 Phone: 312-977-4391 achristopher@nixonpeabody.com	Reg. No. 44,869 Jennifer Hayes Reg. No. 50,845 NIXON PEABODY LLP 300 South Grand Avenue, Suite 4100 Los Angeles, CA 90071-3151 Phone: 213-629-6000 slevy@nixonpeabody.com jenhayes@nixonpeabody.com

Pursuant to 37 C.F.R. § 42.10(b), a Power of Attorney has been filed herewith.

4. Service Information

Correspondence to lead and back-up counsel can be provided at the postal and electronic addresses above.

CLAIM LISTING

Claim 1	
[1pre]	A method of translating medical evidence into computational elements, said method comprising:
[1a]	inputting a portion of electronic text pertaining to a medical condition from a medical evidence source;
[1b]	processing said electronic text using natural language processing (NLP) to produce a plurality of diagnostic factors indicative of said medical condition;
[1c]	displaying said diagnostic factors on a computing device to a person;
[1d]	receiving input from said person via said computing device indicating a factor group that subsumes said diagnostic factors and a rank for said factor group; and
[1e]	storing said medical condition, said factor group and each of said diagnostic factors each as a separate entry in a database, said medical condition linked to said factor group and said factor group linked to each of said diagnostic factors, wherein an entry of each of said diagnostic factors in said database does not include information specific to said medical condition.
Claim 2	
2	A method as recited in claim 1 further comprising: displaying on said computing device a portion of said database which includes said medical condition entry, said factor group entry and each of said diagnostic factor entries and the links between them.
Claim 3	
3	A method as recited in claim 1 wherein said database is a graph database and said each separate entry is a node.
Claim 4	
4	A method as recited in claim 1 wherein said each diagnostic factor entry includes a unique identifier which links to an entry in a medical ontology representing said diagnostic factor.
Claim 5	
5	A method as recited in claim 1 wherein each said diagnostic factor

	entry includes at least one synonym of said diagnostic factor.
Claim 6	
6	A method as recited in claim 1 wherein each said diagnostic factor entry includes a temporality attribute, a negation attribute or an experiencer attribute.
Claim 7	
[7a]	A method as recited in claim 1 further comprising: receiving input from said person via said computing device indicating a demographics corresponding to one of said diagnostic factors, said demographics indicating a type of person to which said diagnostic factor occurs in the context of said medical condition; and
[7b]	storing said demographics in a demographics entry linked between said condition entry and said one of said diagnostic factors entry.
Claim 8	
[8a]	A method as recited in claim 1 further comprising: repeating said steps of inputting, processing, and displaying for a second medical condition different from said medical condition, at least one of said produced plurality of diagnostic factors for said second medical condition being the same as one of said diagnostic factors produced for said first medical condition;
[8b]	receiving input from said person via said computing device indicating a second factor group that subsumes said at least one diagnostic factor and a rank for said second factor group; and
[8c]	storing said second medical condition and said second factor group each as a separate entry in said database, said second medical condition linked to said second factor group and said second factor group linked to an entry in said database for said at least one diagnostic factor.
Claim 9	
9	A method as recited in claim 8 wherein traversing said database from said at least one diagnostic factor entry to said second medical condition produces a second rank for said at least one diagnostic factor in the context of said second medical condition which is different from a first rank produced by traversing from said at least one diagnostic factor entry to said medical condition.

Claim 11	
[11a]	A method as recited in claim 1 wherein said processing produces at least one laboratory test indicative of said medical condition;
[11b]	displaying said laboratory test on a computing device to said person;
[11c]	receiving input from said person via said computing device indicating an investigation group that subsumes said laboratory test and a rank for said investigation group; and
[11d]	storing said investigation group and said laboratory test each as a separate entry in said database, said medical condition linked to said investigation group and said investigation group linked to said laboratory test, wherein an entry of said laboratory test in said database does not include information specific to said medical condition.
Claim 12	
[12pre]	A method of determining a state of a patient, said method comprising:
[12a]	extracting, using natural language processing (NLP), from a medical evidence source a plurality of diagnostic factors indicative of a medical condition and storing each of said diagnostic factors and said medical condition as its own node in a database, said medical condition being linked to said diagnostic factors, wherein each of said diagnostic factor nodes do not include information specific to said medical condition;
[12b]	for at least one of said diagnostic factors, storing demographics required to indicate said medical condition on the basis of said at least one diagnostic factor in a demographics node located between said medical condition node and a node representing said at least one diagnostic factor;
[12c]	extracting from a medical evidence source a plurality of laboratory tests indicative of said medical condition and storing each of said laboratory tests as its own node in said database, said medical condition being linked to said laboratory tests, wherein each of said laboratory test nodes does not include information specific to said medical condition;

[12d]	for at least one of said laboratory tests, storing test results required to indicate said medical condition on the basis of said at least one laboratory test in a results node located between said medical condition node and a node representing said at least one laboratory test; and
[12e]	matching patient data from a patient medical chart to at least one of said diagnostic factor nodes, said demographics node, at least one of said laboratory test nodes, and said results node in order to display a result indicating a likelihood that said patient has said medical condition.
Claim 13	
13	A method as recited in claim 12 further comprising: determining a rank of said likelihood that said patient has said medical condition by reference to a node intermediate between said medical condition node and said diagnostic factor nodes that holds said rank of said diagnostic factors.
Claim 14	
14	A method as recited in claim 12 further comprising: determining a rank of said likelihood that said patient has said medical condition by reference to a node intermediate between said medical condition node and said laboratory test nodes that holds said rank of said laboratory tests.
Claim 15	
15	A method as recited in claim 12 further comprising: performing said extracting steps using natural language processing (NLP) that is initialized using existing diagnostic factor nodes and laboratory test nodes of said database.
Claim 16	
[16pre]	A method of determining a state of a patient, said method comprising:
[16a]	inputting a portion of a patient electronic health record (EHR) into natural language processing (NLP) to produce a plurality of diagnostic factors known to be present or absent in said patient;
[16b]	matching said diagnostic factors to entries in a database that represent said diagnostic factors, each of said diagnostic factor entries being linked to a medical condition entry and not including

	information specific to said medical condition to which said each diagnostic factor entry is linked;
[16c]	traversing a link from said each matched diagnostic factor entry to its linked medical condition and collecting a weight value indicative that said each matched diagnostic factor entry is caused by its linked medical condition;
[16d]	summing all collected weight values for each linked medical condition to produce a total weight value for each linked medical condition; and
[16e]	displaying on a computing device said linked medical conditions in order of weight value.
Claim 17	
17	A method as recited in claim 16 further comprising: displaying in association with each linked medical condition its corresponding total weight value.
Claim 18	
18	A method as recited in claim 16 further comprising: for each present diagnostic factor, traversing upward in a medical ontology link found in said matched diagnostic factor entry in order to infer other diagnostic factors present in said patient and matching these other diagnostic factors to diagnostic factor entries in said database.
Claim 19	
19	A method as recited in claim 16 wherein each said diagnostic factor entry includes at least one synonym of said diagnostic factor.
Claim 20	
20	A method as recited in claim 16 wherein each said diagnostic factor entry includes a temporality attribute, a negation attribute or an experiencer attribute.
Claim 21	
21	A method as recited in claim 16 further comprising: initializing said natural language processing with said entries and links of said database.
Claim 22	
22	A method as recited in claim 16 wherein a demographics entry of

	said database located between one of said matched diagnostic factor entries and its corresponding linked medical condition includes demographics of people who exhibit said matched diagnostic factor in the context of said linked medical condition, said method further comprising: only traversing said link from said one matched diagnostic factor entry to its linked medical condition when information of said patient from said EHR matches said demographics in said demographics entry.
Claim 23	
23	A method as recited in claim 16 wherein said displaying is for the purpose of patient diagnosis, for an administrative medical purpose, or for a medical preauthorization.
Claim 24	
[24pre]	A method of determining next-best actions for a medical condition, said method comprising:
[24a]	for each medical condition of a patient having a probability greater than a threshold of being present in said patient, performing the following
[24b]	identifying all known diagnostic factor entries in a database that each represent a diagnostic factor known to be present or absent in said patient, each known diagnostic factor entry being linked to a medical condition entry in said database representing said each medical condition but not including information specific to said each medical condition;
[24c]	identifying all unknown diagnostic factor entries in said database that each represent an unknown diagnostic factor not known to be present or absent in said patient, each unknown diagnostic factor entry being linked to said medical condition entry;
[24d]	identifying a rank of each unknown diagnostic factor entry that indicates likelihood that said corresponding unknown diagnostic factor indicates said medical condition, said rank being stored in a database entry link between said unknown diagnostic factor entry and said medical condition entry; and
[24e]	displaying on a computing device said unknown diagnostic factors corresponding to said unknown diagnostic factor entries in association with said each medical condition in order of said rank of

	said each unknown diagnostic factor.
Claim 25	
25	A method as recited in claim 24 further comprising: identifying said unknown diagnostic factor entries by traversing down through said database from said each medical condition entry and disregarding said known diagnostic factor entries.
Claim 26	
26	A method as recited in claim 24 wherein said rank of said each unknown diagnostic factor entry is located within a factor group entry located between said medical condition entry and said each unknown diagnostic factor entry.
Claim 27	
27	A method as recited in claim 24 further comprising: for each known diagnostic factor, traversing upward in a medical ontology link found in said identified known diagnostic factor entry in order to infer other known diagnostic factors present in said patient.
Claim 28	
28	A method as recited in claim 24 wherein each said known and unknown diagnostic factor entry includes at least one synonym of said known and unknown diagnostic factor.
Claim 29	
29	A method as recited in claim 24 wherein each said known and unknown diagnostic factor entry includes a temporality attribute, a negation attribute or an experiencer attribute.
Claim 30	
30	A method as recited in claim 24 wherein traversing said database from one of said unknown diagnostic factor entries to a first one of said medical conditions produces a first rank for said one of said unknown diagnostic factor entries in the context of said first medical condition which is different from a second rank produced by traversing from said one of said unknown diagnostic factor entries to a second one of said medical conditions.

I. INTRODUCTION

Petitioner BMJ Publishing Group Limited ("Petitioner" or "BMJ") requests *inter partes* review of claims 1-9 and 11-30 ("challenged claims") of U.S. Patent No. 11,869,674 ("'674 patent"). The '674 patent discloses methods of translating medical evidence into a database that stores diagnostic factors, laboratory tests, and medical conditions as separate linked entries, with condition-specific groupings, ranks, demographics, and test results kept outside the reusable factor and test entries, and methods of using that database to rank a patient's likely conditions and the unknown factors to ask about next. Each of these features was well known before its April 2022 priority date.

II. COMPLIANCE WITH IPR REQUIREMENTS

A. Standing

Petitioner certifies that the '674 patent is available for IPR and that Petitioner is not barred or estopped from requesting IPR of the challenged claims on the grounds below. Petitioner has not filed a civil action challenging the validity of any claim of the '674 patent.

This Petition is timely because it is filed within one year of service of Patent Owner's complaint alleging infringement of the '674 patent and more than nine months after the issue date. EX1001, code (45); EX1015; EX1018; 35 U.S.C. § 315(b); 37 C.F.R. § 42.102(a).

B. Payment of Fees

Pursuant to 37 C.F.R. § 42.15(b), the required fee is being submitted herewith. The Office is authorized to charge any deficiency, or credit overpayment, to Deposit Account No. 50-4181. Any additional fees due in connection with this Petition may be charged to the foregoing account.

III. IDENTIFICATION OF CHALLENGED CLAIMS

Petitioner identifies the following challenged claims and unpatentability grounds:

Ground	Claims	Basis
1	1-3, 7-9, 11-15	§103: Datla, Dang, and Goldberg
2	4-5	§103: Datla, Dang, Goldberg, and Kim
3	6	§103: Datla, Dang, Goldberg, and Van Assel
4	16-18, 20-23	§103: Datla, Dang, and Van Assel
5	19	§103: Datla, Dang, Van Assel, and Kim
6	24-25, 30	§103: Zhong and Datla
7	26	§103: Zhong, Datla, and Dang
8	27, 29	§103: Zhong, Datla, and Van Assel
9	28	§103: Zhong, Datla, and Kim

Petitioner's challenges are supported by the declaration and testimony of Mark A. Musen, M.D., Ph.D. *See* EX1003 ¶¶1-27.

IV. TECHNICAL BACKGROUND

The ’674 patent purports to transform unstructured medical evidence into computational elements stored in a database for determining a patient’s state or next-best actions. EX1003 ¶46. But every claimed technology—standardized medical ontologies, medical knowledge graphs, clinical decision support systems, and NLP-based clinical concept extraction—was well known before the priority date. *Id*.

A. Medical Ontologies and Terminologies

A medical ontology is a formal representation of medical concepts, such as conditions, findings, and investigations, and the relationships among them. EX1003 ¶47. Ontologies had long been used in clinical information systems, electronic health records (“EHRs”), and clinical decision support. *Id*.

The Systematized Nomenclature of Medicine—Clinical Terms (“SNOMED-CT”) is one of the most comprehensive and widely used clinical terminologies. EX1003 ¶¶48-49; EX1038. By 2022, SNOMED-CT contained hundreds of thousands of concepts in hierarchies connected by “is-a” and other relationships (e.g., Bacterial Pneumonia is-a Infective Pneumonia is-a Pneumonia), enabling ancestor-descendant traversal. EX1003 ¶¶50-53; EX1006, 1:54-56, 9:50-60, FIG. 6.

The Logical Observation Identifiers, Names, and Codes (“LOINC”), first developed in 1994, is a standard terminology for laboratory tests and clinical observations that, by 2022, included tens of thousands of codes. EX1003 ¶¶55-56;

EX1039. Many other terminologies (e.g., ICD, RxNorm, the Gene Ontology, and MeSH) were also widely used, and NCBO's BioPortal made hundreds of biomedical ontologies publicly available beginning in the mid-2000s. EX1003 ¶¶58-60; EX1040; EX1041.

The '674 patent concedes that SNOMED-CT and LOINC are pre-existing ontologies. EX1001, 4:6-17, 5:6-12, 6:62-7:12; EX1003 ¶¶54, 57.

B. Healthcare Knowledge Graphs

A knowledge graph is a structured representation of entities ("nodes") and the relationships among them ("edges") that enables efficient traversal, querying, and reasoning. EX1003 ¶61; EX1025 ¶[0003]. In healthcare, knowledge graphs represented conditions, findings (e.g., symptoms, signs, and test results), investigations (e.g., tests and procedures), medications, and their relationships, typically organizing nodes into condition, finding, and investigation categories. EX1003 ¶¶62-65; EX1006, 6:10-30; EX1011 ¶[0115]; EX1008, FIGS. 1A-1B.

The field had long addressed the problem of transforming narrative medical evidence into computational form. EX1003 ¶66; EX1034, 468-69; EX1035, 458-60. Practitioners built medical knowledge graphs from the medical literature (e.g., The BMJ), clinical guidelines, SNOMED-CT and LOINC, and EHRs. EX1003 ¶67. Graphs were constructed manually by experts, automatically using NLP and machine learning, or through human-in-the-loop hybrids. EX1003 ¶¶67-69;

EX1006, 6:5-30, 9:35-50; EX1028, 5; EX1029, 1-2, 11; EX1036, 21-22. By 2022, constructing and using medical knowledge graphs to integrate clinical evidence and support clinical reasoning and decision support was well known. EX1003 ¶70.

C. Clinical Decision Support Systems

Clinical decision support systems ("CDSS") are computer-based tools that provide clinicians and other healthcare stakeholders with person-specific, intelligently filtered, and timely information to enhance health-related decisions. EX1003 ¶71; EX1032, 803. CDSS have been used since the 1970s (e.g., MYCIN and INTERNIST-1). EX1003 ¶71; EX1033, 80, 83. INTERNIST-1 separately represented conditions and findings, associated findings with multiple conditions, limited certain conditions by demographics, ranked possible conditions, and suggested findings that would refine the differential. EX1003 ¶72; EX1034, 469-70; EX1035, 458-60, 462, 469-70. By the priority date, CDSS routinely used structured knowledge bases, including knowledge graphs linked to standard terminologies. EX1003 ¶73; EX1032, 811, 825-27.

D. Natural Language Processing

Natural language processing ("NLP") refers to computational techniques for extracting structured information from unstructured text and had long been used to extract clinical concepts and relationships from medical literature and clinical notes. EX1003 ¶74; EX1010, 1118-19; EX1037, 392-94. MetaMap, cTAKES, and

MedLEE extracted symptoms, diagnoses, medications, and laboratory tests from narrative text and mapped them to standard ontologies, and using NLP to populate knowledge graphs from medical literature was well known. EX1003 ¶75; EX1029, 2-5; EX1032, 811; EX1037, 392-98.

V. THE '674 PATENT

A. AIA Status

The '674 patent issued January 9, 2024, from Application No. 18/135,026, filed April 14, 2023, and claims priority to a provisional application filed April 15, 2022. EX1001, codes (21), (22), (45), (60).

B. Level of Skill in the Art

A person of ordinary skill in the art ("POSITA") would have had at least a master's degree in biomedical informatics, computer science, health informatics, or a closely related field and at least two years of experience working with medical knowledge representation systems, biomedical ontologies, knowledge graphs, or clinical decision support systems. EX1003 ¶¶28-32. A POSITA would also have had a good working knowledge of biomedical ontologies and terminologies, knowledge-graph construction and querying, clinical decision-support methodologies, graph-based reasoning, APIs for accessing structured medical knowledge, and integrating medical evidence from multiple sources. *Id*. Additional education can compensate for less professional experience, and vice versa. *Id*.

C. Overview of the ’674 Patent

The ’674 patent describes a “computational evidence platform” that “extracts clinical concepts from medical evidence sources and creates a database of elemental diagnostic factors and elemental investigations links to medical conditions.” EX1001, Abstract; EX1003 ¶77. The purported advance is converting narrative evidence into a “structured, computational representation system.” EX1001, 2:38-44. But the specification concedes that publishers “such as DynaMed, BMJ, UpToDate” already “assemble the published evidence through expert, human-centric activity,” and that current practice reduces narrative documents to other narrative documents. EX1001, 1:52-63; EX1003 ¶78.

The specification treats many claimed aspects as pre-existing. EX1003 ¶¶77-80. Evidence sources include the National Library of Medicine, JAMA Rational Clinical Evidence, “BMJ Best Practice (or other British Medical Journal databases),” SNOMED-CT, and LOINC, and the system “can extract clinical items from any textual data source.” EX1001, 6:62-7:12; EX1003 ¶¶79-80. The disclosed NLP processes are conventional (word and sentence segmenters, negation cues, and named entity recognition and linking). EX1001, 7:52-58; EX1003 ¶81. The database may be “[a]ny of a wide variety of” databases, including relational, object-store, or off-the-shelf JanusGraph graph databases. EX1001, 8:4-17; EX1003 ¶82.

Finding nodes are called "elemental diagnostic factors" ("EDFs"), which are signs and symptoms a physician can observe or a patient can describe; investigation nodes are "elemental diagnostic investigations" ("EDIs"), which are laboratory tests or specialist examinations. EX1001, 9:11-15, 9:52-67, 11:53-63; EX1003 ¶83. As shown in FIG. 4A, conditions link to "factor groups" containing EDFs and "investigation groups" containing EDIs, with demographics and result nodes in between. EX1001, 18:20-19:26, FIG. 4A; EX1003 ¶84.

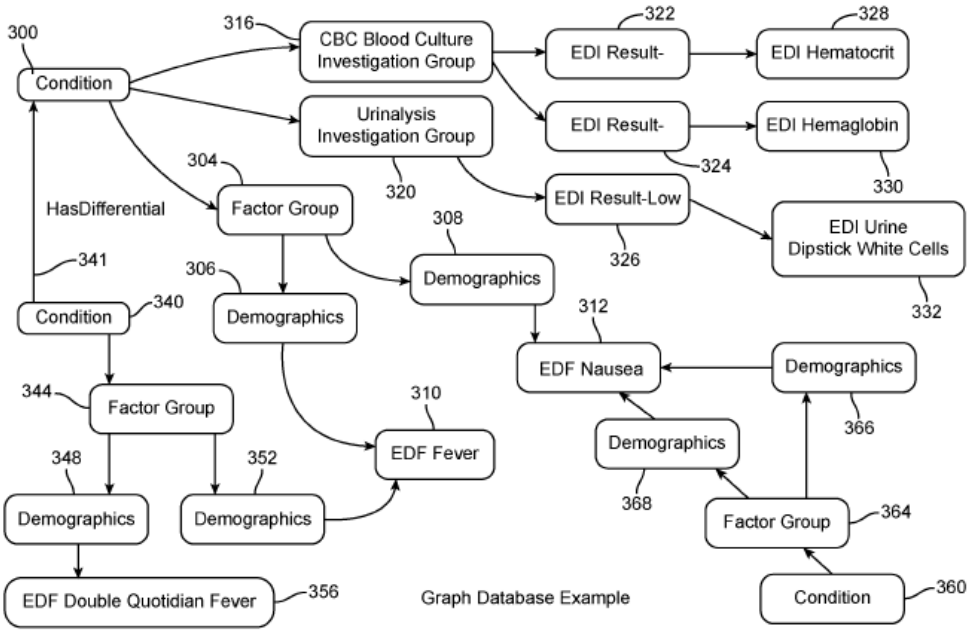


FIG. 4A

EX1001, FIG. 4A

According to the specification, EDFs "remain elemental, independent, and can be used computationally across conditions," while condition-specific ranks, demographics, and test results reside in the intermediate nodes. EX1001, 10:41-50,

22:5-10, 25:1-3. The platform then matches a patient's factors to the database to rank likely conditions and identifies unknown factors as next-best actions. EX1001, 22:33-35, 24:57-59, FIGS. 5-6.

D. Prosecution History

The application was filed April 14, 2023, with 30 claims and a Track One request. EX1002, 15, 87-93; EX1003 ¶87. The Examiner first rejected claims 1-15 and 24-30 as anticipated by U.S. Patent Publication No. 2022/0344020 ("VanNess"). EX1002, 184-99; EX1003 ¶88.

In response, for claim 1, the applicant argued that VanNess lacks electronic text from a "medical evidence source," NLP-produced diagnostic factors, displaying those factors, receiving input from a person indicating a factor group and rank, and factor entries without condition-specific information. EX1002, 202-18; EX1003 ¶¶90-91. For claim 12, the applicant argued VanNess lacks separate condition-independent factor and test nodes, intermediate demographics and results nodes, and matching chart data to those nodes. EX1002, 213-14; EX1003 ¶92. For claim 24, the applicant argued VanNess lacks iteration over probable conditions, separate known and unknown factor entries, and a rank stored between the unknown factor and condition entries. EX1002, 214-16; EX1003 ¶93.

The Examiner then maintained the VanNess rejection of claims 1-7 and 10, indicated that claims 8, 9, and 11 contained allowable subject matter, and allowed

claims 12-30. EX1002, 412-24; EX1003 ¶94. The applicant responded with "[m]inor" clarity amendments. EX1002, 441. The notice of allowance merely repeated the claim language. EX1002, 425-48, 469-78; EX1003 ¶¶95-96.

Each feature the applicant relied on during prosecution is disclosed in prior art the Examiner never applied. For example, Dang discloses intermediate feature nodes and a context node between disease and symptom nodes, and an observation-item node between disease and laboratory-test nodes; Goldberg discloses medical experts entering grouped, ranked clinical facts through an editor; and Zhong discloses ranking unknown symptoms using values stored on symptom-disease links. *Infra* §VIII; EX1003 ¶¶105, 156-170, 259, 279, 484-490.

VI. CLAIM CONSTRUCTION

Claims are given their plain and ordinary meaning as understood by a person of skill in the art. 37 C.F.R. § 42.100(b). Petitioner is unaware of any prior claim construction determination related to the '674 patent. Because the challenged claims are unpatentable under any interpretation consistent with their plain and ordinary meaning, the Board need not expressly construe any claim term. EX1003 ¶102.

VII. SUMMARY OF THE ASSERTED REFERENCES

The Examiner never applied any reference relied on herein in a rejection. EX1002, 184-99, 412-24. Goldberg was listed in an IDS filed with a post-allowance RCE, but the Examiner never applied it or combined it with any reference. EX1001,

code (56); EX1002, 491, 501-06; EX1003 ¶105. Datla, Dang, Zhong, Van Assel, and Kim were not cited or considered during prosecution. EX1001, code (56).

Each asserted reference is analogous art because it is in the '674 patent's field of endeavor—automating access to medical data and evidence for clinical diagnosis. EX1001, 1:15-19; EX1005 ¶[0001]; EX1006, 1:17-20; EX1008 ¶[0001]; EX1009, 1:6-10; EX1011 ¶¶[0001], [0115]; EX1012 ¶¶[0015]-[0016], [0038]-[0039].

A. Datla

Datla published August 15, 2019, from an application filed April 15, 2019, with an international filing date of October 24, 2017. EX1005, codes (22), (43), (86). Datla is prior art at least under 35 U.S.C. § 102(a)(1).

Datla discloses an NLP engine that extracts patient symptoms and demographic parameters from natural-language input and queries a knowledge graph built from a corpus of medical information—with "symptoms as root nodes connected by edges to the diseases and medical conditions associated with those symptoms"—to identify "a ranked list of one or more medical conditions, diagnoses, treatments, and/or tests for the patient." EX1005, claim 1, ¶¶[0006]-[0008], [0028], [0030]-[0036], [0050]-[0057]; EX1003 ¶103.

B. Goldberg

Goldberg issued April 25, 2023, from an application filed August 12, 2019. EX1006, codes (22), (45). Goldberg is prior art at least under 35 U.S.C. § 102(a)(2).

Goldberg discloses a hybrid knowledge graph diagnostic system that integrates clinical knowledge from books, articles, and publications with ontology-based graphs (including SNOMED and LOINC) and selects graph pathways to generate differential diagnoses. EX1006, Abstract, 5:51-62, 6:10-24; EX1003 ¶104. Its knowledge base includes node sets for diseases, associated conditions, symptoms, complications, laboratory tests, and treatments, which medical experts populate through a template-based editor capturing categorized clinical rules with probability values and source citations. EX1006, 6:20-30, 9:41-49, 16:8-24, FIGS. 1, 4, 8A-8B; EX1003 ¶104.

C. Dang

Dang published June 30, 2022, from an application filed December 31, 2020. EX1008, codes (22), (43). Dang is prior art at least under 35 U.S.C. § 102(a)(2).

Dang discloses a graph-based clinical diagnosis system whose FIG. 1A meta-model includes Disease, Symptom, Sign, Context, "Lab Test & Check," and "Observation Item" node types, and whose FIG. 1B graph includes factual (symptom, context, and test), intermediate feature, and disease nodes, with conditional probabilities P(f|d) on the feature-disease connections. EX1008 ¶¶[0027]-[0031], FIGS. 1A-1B; EX1003 ¶106. Dang detects factual nodes from patient records, infers evidential, feature, and disease nodes, and outputs a sorted list of possible diseases. EX1008 ¶¶[0047]-[0050], [0056]; EX1003 ¶106.

D. Zhong

Zhong issued December 7, 2021, from an application filed July 11, 2016. EX1009, codes (22), (45), (65). Zhong is prior art at least under 35 U.S.C. § 102(a)(1).

Zhong discloses systems for "question generation to obtain more related medical information based on observed symptoms from a patient." EX1009, Abstract. Zhong collects observed symptoms, identifies likely diseases from a knowledge graph whose edges carry symptom-disease probability values, scores candidate (unobserved) symptoms, and presents the highest-ranked as questions, repeating "until a disease is diagnosed or the patient reports no more observed symptoms." EX1009, 3:19-4:31, claim 1, FIGS. 3-6; EX1003 ¶107.

E. Van Assel

Van Assel published July 29, 2021, from an application filed January 23, 2020. EX1011, codes (22), (43). Van Assel is prior art at least under 35 U.S.C. § 102(a)(1).

Van Assel discloses a computer-implemented diagnosis system whose diagnostic engine uses a knowledge graph, or "medical ontology, linking together medical diseases, symptoms, risk factors etc.," in which each concept has a label and a unique identifier and is related to parent concepts by is-a links. EX1011 ¶¶[0113]-[0118], [0124]; EX1003 ¶108. Van Assel records whether each finding is present,

absent, or unknown and when it was reported. EX1011 ¶¶[0112], [0245], [0263]; EX1003 ¶109.

F. Kim

Kim published July 4, 2024, from a national-phase application filed January 17, 2024, with an international filing date of July 21, 2021. EX1012, codes (22), (43), (86). Kim is prior art at least under 35 U.S.C. § 102(a)(2). *See* 35 U.S.C. § 363.

Kim discloses extracting medical concepts into a graph database, or knowledge graph, whose entries are annotated with unique ontology identifiers (e.g., "SNOMED CT code SCTID: 93880001") and include synonyms. EX1012, Abstract, ¶¶[0038], [0050]-[0054]; EX1003 ¶¶110-111.

VIII. UNPATENTABILITY GROUNDS

A. Ground 1: Claims 1-3, 7-9, and 11-15 are Unpatentable Over Datla, Dang, and Goldberg

Claims 1-3, 7-9, and 11-15 would have been obvious over Datla, Dang, and Goldberg. EX1003 ¶¶113-328. In the combination, Datla supplies the medical-evidence corpus, NLP processing, condition-independent graph nodes, and diagnostic outputs; Dang supplies an intermediate feature-node tier and context, test, and observation-item node types; and Goldberg supplies human editorial review of grouped, ranked clinical facts and condition-specific test results. *Id*. ¶¶114-122.

1. Motivation to Combine

A POSITA would have combined Datla, Dang, and Goldberg because each addresses medical knowledge graph construction for clinical decision support. Datla builds a knowledge graph from medical evidence sources for automated diagnosis; Dang builds a symptom-disease knowledge graph from medical guidelines for machine-guided diagnosis; and Goldberg builds a hybrid knowledge graph integrating clinical knowledge with ontology-based knowledge. EX1005 ¶¶[0001], [0028]; EX1008 ¶¶[0003], [0025]; EX1006, 5:51-62; EX1003 ¶¶114-122.

Datla describes its graph as "a tree-like structure with a plurality of nodes connected by one or more edges," so a POSITA implementing Datla would have looked to other knowledge-graph descriptions, such as Dang, for architectural detail. EX1005 ¶[0031]; EX1003 ¶116. Dang discloses a three-tier architecture of factual nodes (symptom, test, and context nodes), intermediate feature nodes, and disease nodes, and a graph that places a Context node between Disease and Symptom nodes and an Observation Item node between Disease and Lab Test & Check nodes. EX1008 ¶¶[0027]-[0029], [0047], FIGS. 1A-1B; EX1003 ¶117. The feature tier reflects the inferential chain from raw observations to diseases. EX1008 ¶¶[0031]-[0032], [0040], [0047]-[0052]; EX1003 ¶¶118-119.

Datla's graph is built automatically, but Goldberg explains that "due to the complexity of such medical content, human experts may be required to understand

the text and use pre-defined templates to process the data and define the clinical relations," and it provides a template-based editor through which experts enter and review disease-feature associations with probability values. EX1006, 6:20-30, 9:41-49, FIGS. 8A-8B; EX1003 ¶¶120-121; *see also* EX1029, 3-4 (confirming that automatic construction methods "cannot be directly applied to the medical domain" and that "an automatic method that involves an appropriate amount of expert effort is required"). A POSITA incorporating Dang's feature-node architecture into Datla therefore would have been motivated to employ a human expert review and confirm (or create) the feature-node groupings and their ranks. EX1003 ¶120. Dang expressly invites this, stating that its graph "may be generated according to any known method in the art." EX1008 ¶[0027]; EX1003 ¶160.

A POSITA would have reasonably expected success because the combination uses each reference for its known purpose: Datla's NLP-based extraction and knowledge graph, Dang's graph architecture and node types, and Goldberg's condition-specific clinical facts and editorial review. EX1003 ¶122. The result—a medical knowledge graph with expert-confirmed factor groupings and ranks built from NLP-extracted concepts—is predictable, and the combination produces no unexpected result. *Id*.

2. Independent Claim 1

Each limitation of claim 1 is disclosed or rendered obvious by the combination. EX1003 ¶¶123-175.

a. Element [1pre]

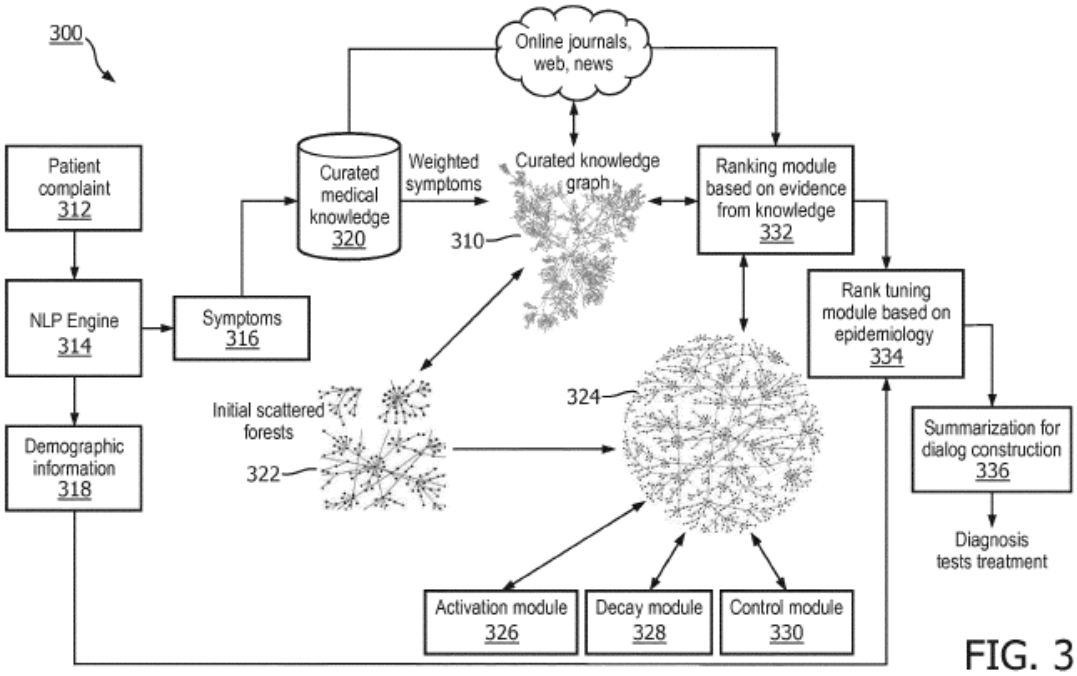
To the extent the preamble is limiting, Datla discloses a method of translating medical evidence into computational elements. EX1003 ¶¶124-128. Datla translates narrative text from medical journals, Wikipedia clinical-medicine pages, and PubMed into the nodes and edges of a knowledge graph that can be algorithmically queried and traversed. EX1005 ¶¶[0030]-[0031], [0034]-[0035], [0047]-[0048]; EX1003 ¶¶125-126. Dang does the same with medical guidelines, and such translation was well known, as Friedman's MedLEE-based coding of clinical narrative shows. EX1008 ¶[0025]; EX1037, 392; EX1003 ¶¶127-128.

b. Element [1a]

Datla inputs a portion of electronic text pertaining to a medical condition from a medical evidence source to build its knowledge graph. EX1003 ¶¶129-132. Datla's "corpus of medical information can be any source of information, including but not limited to medical journals, newspapers, online sources such as Wikipedia, and other sources," and Datla retrieves biomedical articles "that can answer generic clinical questions." EX1005 ¶¶[0030], [0033]; EX1003 ¶130. Datla analyzes Wiki pages under the "clinical medicine" category, each of which "is related to other medical

conditions," as well as PubMed and PubMed Central articles. EX1005 ¶¶[0030], [0035], [0047]-[0048]; EX1003 ¶130.

Datla FIG. 3 shows medical evidence sources including "Online journals, web, news" used to create weighted patient symptoms and a knowledge graph. EX1005 ¶¶[0025], [0065]-[0067], FIG. 3; EX1003 ¶131.



EX1005, FIG. 3

Datla FIG. 5 likewise shows the "Corpus of medical information" (i.e., medical evidence sources) input into the system. EX1005 ¶[0027], FIG. 5; EX1003 ¶132.

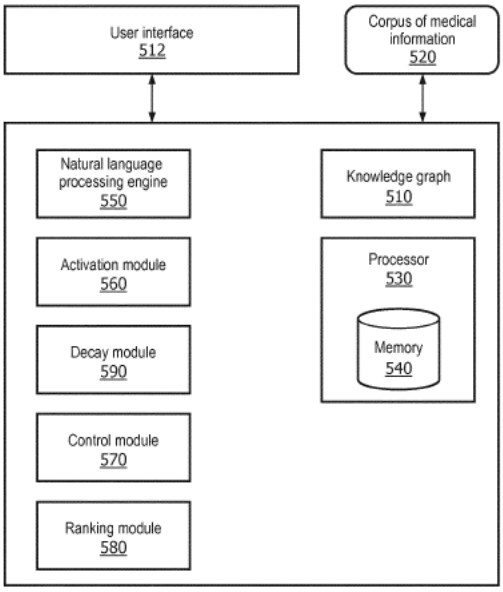


FIG. 5

EX1005, FIG. 5

c. Element [1b]

Datla discloses or renders obvious processing the electronic text using natural language processing (NLP) to produce a plurality of diagnostic factors indicative of said medical condition. EX1003 ¶¶133-143. Datla's knowledge graph includes various node types including nodes for symptoms (diagnostic factors), medical conditions, diagnoses, tests, procedures, medications, or other clinical concepts. EX1005 ¶¶[0030]-[0031]; EX1003 ¶134. Datla discloses processing text from the corpus of medical information to create these nodes. EX1005 ¶¶[0030]-[0031], [0035]; EX1003 ¶138. For example, Datla extracts the "classic and common signs and symptoms" of each Wiki page's diagnosis to build a graph with "symptoms as root nodes connected by edges to the diseases and medical conditions associated

with those symptoms." EX1005 ¶[0040]; EX1003 ¶138. These symptoms are diagnostic factors indicative of a medical condition. EX1003 ¶¶134, 138-139.

Datla also discloses the use of NLP. For example, Datla's "natural language processing engine extracts at least one patient symptom," may "extract keywords related to symptoms, such as lab results, procedures, and/or demographic information," and applies bag-of-words, vector-embedding, and TF-IDF techniques. EX1005 ¶¶[0046], [0052], [0065], [0081]; EX1003 ¶¶135-137. Datla further explains that its graph-building approach includes: "(i) topical keyword analysis in which the most clinically relevant keywords from the given topic descriptions, summaries, and clinical notes are identified" and "(ii) diagnostic inference using reasoning based on the topical keywords to generate the diagnoses, tests, and treatments using the underlying clinical contexts represented within either a key-value memory network or a knowledge graph, both powered by an external clinical knowledge source." EX1005 ¶[0034]; EX1003 ¶¶138-141. A POSITA would have understood, or at least found obvious, that Datla's extraction of keywords regarding symptoms, signs, diagnoses, tests, and treatments (EX1005 ¶¶[0028], [0034], [0052]) would be accomplished via NLP. EX1003 ¶140.

To the extent Patent Owner contends[1] that Datla does not use NLP to produce diagnostic factors for its knowledge graph, it would have been obvious to employ Datla's NLP engine for this purpose. EX1003 ¶142. Using the NLP engine to extract the same types of information from the corpus would have been a natural and obvious extension of Datla's NLP engine. *Id*. This further would have been obvious in view of Dang, which expressly teaches using NLP to extract a set of diseases and symptoms from medical information. EX1008 ¶[0025]; EX1003 ¶¶141-142. Indeed, using NLP for this purpose was well known in the art. *Supra* §§IV.B, IV.D. A POSITA would have had a reasonable expectation of success because natural language processing is a common method for automatically extracting text. EX1003 ¶142.

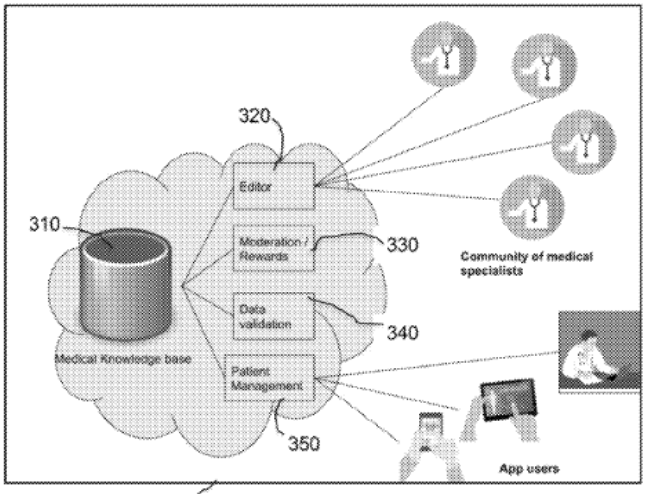
d. Element [1c]

Datla displays diagnostic factors on a computing device to a person when it provides diagnostic outputs, including the symptom-based reasoning behind them, to clinicians through "any user interface that allows information to be conveyed, such as a speaker or screen." EX1005 ¶¶[0031], [0062]; EX1003 ¶¶144-145.

[1] Patent Owner cannot credibly dispute that Datla uses NLP as claimed without contradicting its allegation that creating databases from journal content "inherently involves the use of NLP." EX1017, 9; EX1003 ¶143.

To the extent the element requires displaying diagnostic factors to a person during editorial review,[2] Goldberg supplies that teaching. EX1003 ¶146. Goldberg's editor 320 is "an interface in the form of a mobile app or web interface for data entry by community members," through which medical experts use "a template-based editor" to translate "complex medical concepts and relations" into machine-readable form. EX1006, 6:25-30, 7:57-59, FIG. 3; EX1003 ¶¶146-147.

FIG. 3



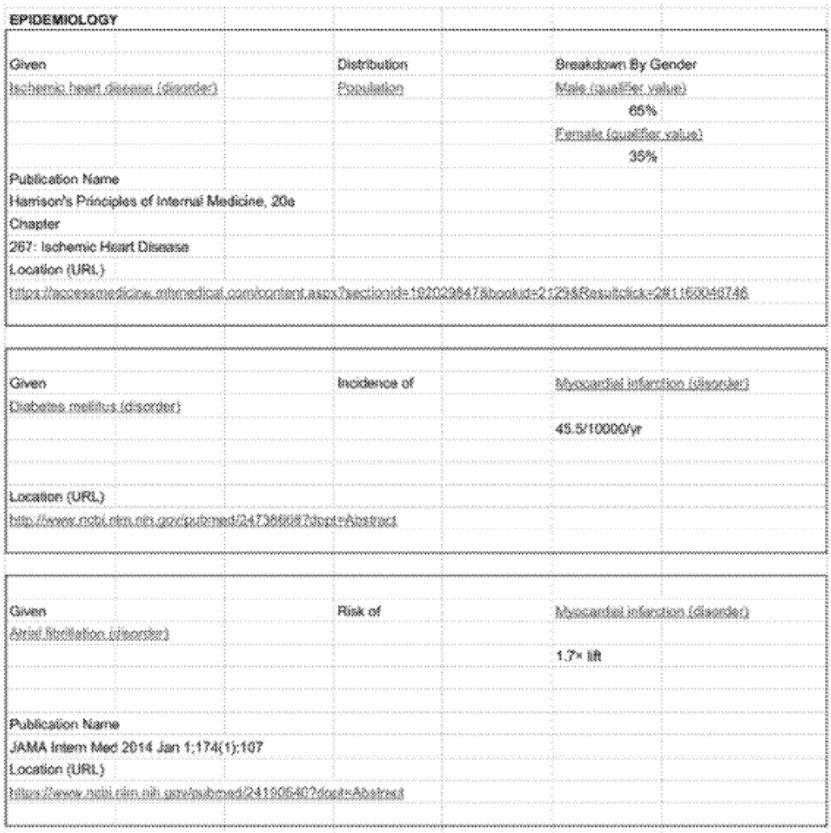
EX1006, FIG. 3

FIGS. 8A and 8B show the editor displaying clinical rules for myocardial infarction "as they appear in medical literature," organized by category with probability values and source citations. EX1006, 9:61-10:3, FIGS. 8A-8B; EX1003

[2] Patent Owner alleges that this element is met when the information is "displayed as part of the knowledge graph or in any user interface." EX1017, 9-10.

¶¶148, 151. FIG. 8A shows an EPIDEMIOLOGY panel with gender distribution, incidence, and risk data, each in a "Given [condition]" card. EX1006, FIG. 8A; EX1003 ¶149.

FIG. 8A



EX1006, FIG. 8A

FIG. 8B shows PHYSICAL SIGNS/SYMPTOMS (e.g., dyspnea, 70-80%), IMAGING (e.g., ST segment depression, 40-50%), and COMPLICATIONS entries. EX1006, FIG. 8B; EX1003 ¶150.

Figure 8B



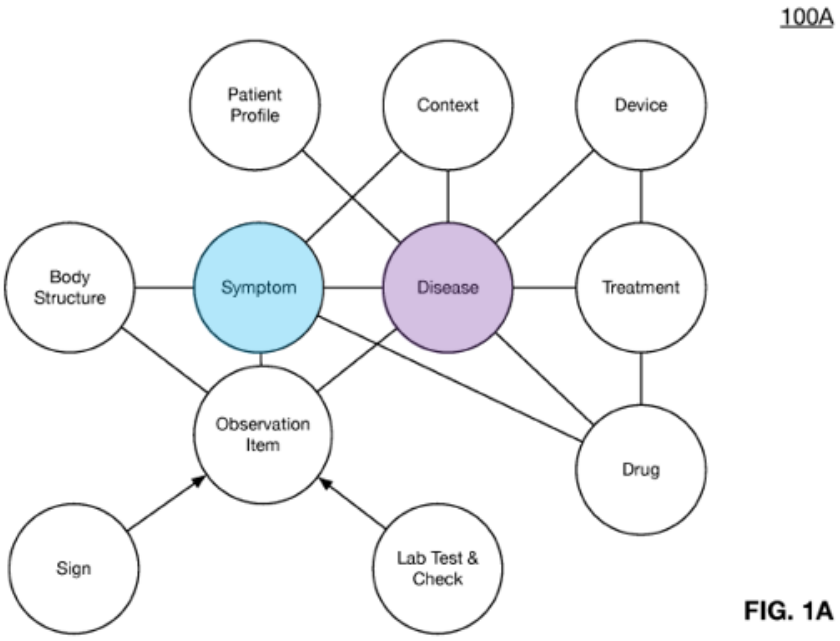
EX1006, FIG. 8B

Displaying extracted diagnostic factors to an expert during knowledge-base construction was well known and routine in the art and would have been obvious. EX1037, FIG. 5; EX1003 ¶152. In the combination, the diagnostic factors that Datla's NLP extracts are displayed to a human expert on Goldberg's editor so the

expert can review and confirm the feature-node groupings and ranks, with a reasonable expectation of success. EX1003 ¶153; *infra* §VIII.A.2.e.

e. Element [1d]

Element [1d] requires a factor group that subsumes the diagnostic factors and has a rank, and human input indicating that group and rank. EX1003 ¶¶154-155. Datla supplies the graph built from medical evidence, Dang supplies intermediate factor-group nodes, and Goldberg supplies the human input. *Id*. ¶¶156-170. Dang's FIG. 1A illustrates distinct node types connected by directional edges. EX1008 ¶[0027], FIG. 1A.



EX1008, FIG. 1A (annotated)

Dang's FIG. 1B illustrates a three-tier graph in which feature nodes $\{f_1 \ldots f_k\}$ are positioned between factual nodes and disease nodes $\{d_1 \ldots d_n\}$. EX1008 ¶¶[0028]-

[0029], FIG. 1B; EX1003 ¶156. For example, feature node f_1 receives connections from symptom nodes s_1 and s_2 and context node c_1 and connects to disease node d_1 through conditional probability $P(f_1|d_1)$. EX1008 ¶[0030], FIG. 1B; EX1003 ¶¶157-158.

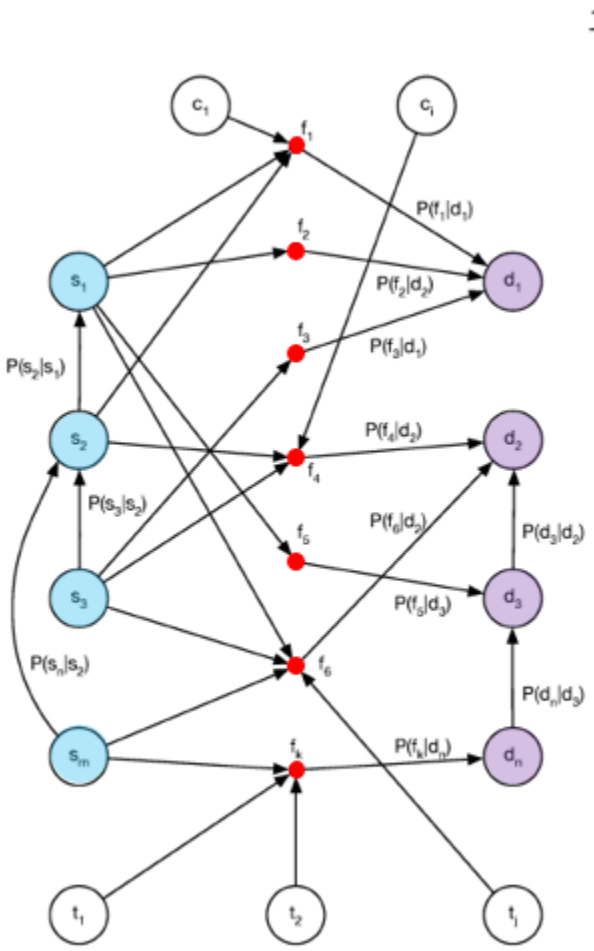


FIG. 1B

EX1008, FIG. 1B (annotated)

Dang's feature nodes are factor groups because each subsumes multiple diagnostic-factor nodes, consistent with the factor groups of the '674 patent's FIG. 4A. EX1001, FIG. 4A; EX1003 ¶¶157-158. Each feature node also has an associated

rank: the disease-feature connection carries P(f|d), the likelihood of the feature when the disease is present, and Dang computes an aggregated "Confidence of feature = Prod(Confidence of evidence*weight)." EX1008 ¶¶[0031], [0049]. These values quantify the strength of the association between the grouped factors and the disease. EX1003 ¶159.

Goldberg supplies the claimed human input. Goldberg's experts use the template-based editor and a "crowdsourcing and community-based editing module" to enter disease-feature associations with probability values (ranks) and source citations, and FIGS. 8A-8B show factors organized by category with Probability and Strength of Source values. EX1006, 5:42-50, 6:20-30, 10:28-32, FIGS. 8A-8B; EX1003 ¶¶160-162; *see also* EX1029, 6-7 (confirming the benefits of involving subject matter experts in knowledge graph creation). It would have been obvious to a POSITA to have a human subject matter expert use Goldberg's editor to indicate the factor groups (Dang's feature nodes) and their ranks (probability values, as both Dang and Goldberg teach) as this is a known technique applied to a known architecture in a known system, with the predictable result of a more accurate graph. EX1003 ¶¶163-164.

Patent Owner's infringement allegations confirm that the combination meets this element. Patent Owner contends that a "rank for said factor group" is met by diagnostic factors "grouped and ranked as 'Key,' 'Other,' 'Common,' and

'Uncommon'" through "rigorous editorial methodology," and by "'Strong' risk factors" versus "'Other' risk factors." EX1017, 12; EX1003 ¶¶165-166. Under that reading, a rank is a categorical priority designation that editorial personnel apply to a group. EX1003 ¶166. Goldberg's Type of Rule/Feature categories (e.g., SYMPTOM/COMPLAINT, RISK FACTORS, LABORATORY TESTS), with expert-assigned Probability and Strength of Source values, function the same way, as do Dang's feature-node probability and confidence values. EX1006, 5:63-6:9, 6:20-30, FIGS. 8A-8B; EX1008 ¶¶[0031], [0049]; EX1003 ¶¶167-169. An expert already reviewing and creating factor groups in Goldberg's editor would have assigned or confirmed a categorical designation for each group as a predictable extension of Goldberg's workflow. EX1006, 9:41-49; EX1003 ¶170.

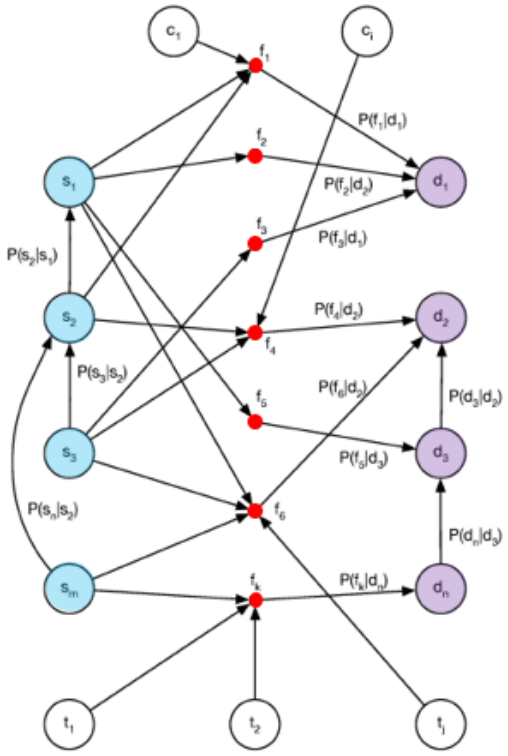
f. Element [1e]

Datla stores medical conditions, symptoms, tests, and treatments as separate interconnected nodes, and its symptom nodes are condition-independent: "a symptom of a fever will be connected by edges to hundreds of other nodes, as a fever is a symptom of many patient scenarios." EX1005 ¶¶[0030]-[0032]; EX1003 ¶173.[3]

[3] Patent Owner alleges in its infringement contentions that a graph containing separate symptom and condition nodes, like Datla and Dang, satisfies this claim element. EX1017, 17-20, 62-63, 74.

Condition-specific relationship information resides on the weighted edge, not in the symptom node. EX1005 ¶[0032]; EX1003 ¶¶173, 252. Datla's system includes a "database," which a POSITA would have understood to be a graph database. EX1005 ¶[0078], FIG. 5; EX1003 ¶251.

Dang's three-tier architecture explicitly discloses the recited linkage: disease nodes (medical conditions) link to feature nodes (factor groups), which link to symptom nodes (diagnostic factors). EX1008 ¶¶[0028]-[0030], FIG. 1B; EX1003 ¶174.



EX1008, FIG. 1B (annotated)

Dang's symptom nodes are also condition-independent, because the same symptom node can connect to feature nodes for different diseases and the condition-specific probability resides on the feature-disease connection. EX1008 ¶¶[0030]-[0031], FIG. 1B; EX1003 ¶174.[4]

Representing each concept once and placing relationship-specific information on the relationship was a well-established knowledge-engineering principle and would have been obvious to a POSITA. EX1003 ¶174. The combination thus stores the condition, the factor group, and each diagnostic factor as separate linked entries, with each factor entry free of condition-specific information. *Id*. ¶175.

3. Claim 2

Datla, Dang, and Goldberg render obvious displaying a portion of the database including the medical condition entry, factor group entry, each diagnostic factor entry, and the links between them. EX1003 ¶¶176-181.

Goldberg's editor displays clinical data organized by condition (e.g., Myocardial Infarction), Type of Rule/Feature categories (factor groups), and individual factors with their Probability, Strength of Source, and Link to Source—

[4] Patent Owner alleges that a diagnostic factor node holds no condition-specific information when another node is located between the diagnostic factor and condition nodes. EX1017, 61, 94-95, 126.

the condition entry, factor group entries, diagnostic factor entries, and the links between them. EX1006, 9:61-10:3, FIGS. 1, 4, 8A-8B; EX1003 ¶178. Dang's FIG. 1B likewise visualizes disease, feature, and symptom nodes with their connecting edges. EX1008 ¶¶[0028]-[0031], [0034], FIG. 1B; EX1003 ¶179.

In the combination, displaying these entries and links is necessary to Goldberg's editorial workflow, because the expert cannot review or edit factor-group assignments without seeing them. EX1003 ¶180. INTERNIST-1's knowledge-base editing tools displayed the same types of elements decades earlier. *Id*. ¶181.

4. Claim 3

Datla, Dang, and Goldberg disclose graph databases in which each separate entry is a node. EX1003 ¶¶182-187.

Datla's knowledge graph, stored in a graph database, is "a tree-like structure with a plurality of nodes connected by one or more edges," in which each condition, symptom, test, and other clinical concept is a node. EX1005 ¶[0031]; EX1003 ¶185.

Dang likewise discloses this; for example, FIG. 1B illustrates a graph with nodes connected by edges carrying conditional probabilities. EX1008 ¶¶[0028]-[0031], FIG. 1B; EX1003 ¶185.

Goldberg also stores its medical knowledge in a graph database and describes a "hybrid knowledge graph" of interlinked nodes and edges. EX1006, 5:51-62, 6:10-34; EX1003 ¶186.

In the combination, the condition, factor group, and each diagnostic factor are nodes in a graph database. EX1003 ¶187.

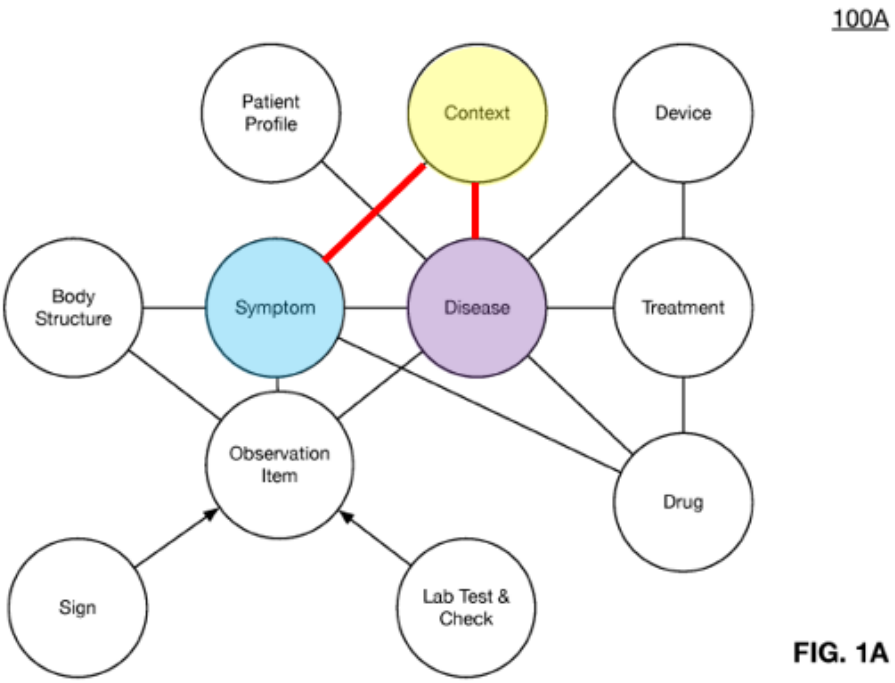
5. Claim 7

a. Element [7a]

Goldberg discloses receiving input from a person via a computing device indicating demographics corresponding to diagnostic factors. EX1003 ¶¶188-189. Goldberg's editor templates are "designed to support different types of clinical facts e.g. etiological, demographic, causal relations between disorders, findings and complications, risk factors, etc." EX1006, 13:16-22; EX1003 ¶190. FIG. 8A's EPIDEMIOLOGY section shows expert-entered gender distribution (65% male, 35% female) for myocardial infarction, and FIG. 4 shows a RISK FACTORS entry for Celiac disease, "Periods of life (qualifier value):ADULTHOOD," specifying the type of person to whom the factor applies for that condition. EX1006, FIGS. 4, 8A; EX1003 ¶¶191-192. That is the '674 patent's "age and sex restriction." EX1001, 10:41-44; EX1003 ¶192. Because Goldberg's editor is already in the combination for Element [1d], receiving demographic input through it requires no further modification. EX1003 ¶193.

b. Element [7b]

Dang supplies the structure for storing demographics in an entry linked between the condition and diagnostic factor entries. EX1003 ¶195. Dang's FIG. 1A places a Context node, connected by one edge to the Disease node and by another to the Symptom node, on the path between them, just as the '674 patent describes demographics node 306 as "in between the EDF node and the condition" even though factor group 304 also lies on that path. EX1008 ¶[0027], FIG. 1A; EX1001, 25:15-21, FIG. 4A; EX1003 ¶196.



EX1008, FIG. 1A (annotated)

It would have been obvious to store condition-specific demographics in Dang's context node, which is structurally linked between the condition (Disease) and the

diagnostic factor (Symptom). EX1003 ¶196; *see also infra* §VIII.A.9.c; EX1003 ¶¶260-262; EX1008 ¶¶[0027], [0036]. Zhong confirms that demographics warranted their own node in a diagnostic graph. EX1009, 3:20-25; EX1003 ¶198. In the combination, demographics entered through Goldberg's editor are stored in Dang's context node linked between the condition and factor entries. EX1003 ¶197.

6. Claim 8

a. Element [8a]

Datla's graph is built from a corpus covering hundreds of conditions, and the same diagnostic factor links to many of them: fever "will be connected by edges to hundreds of other nodes." EX1005 ¶¶[0030]-[0032], [0036], [0039]; EX1003 ¶¶200-201. Datla's graph-building necessarily repeats the inputting, processing, and displaying steps for multiple conditions, and Datla resolves variant terminology to canonical entries using the Wikipedia API, so a factor appearing for two conditions maps to the same node. EX1005 ¶[0040]; EX1003 ¶201. The '674 patent's FIG. 4A shows the same shared-factor arrangement, with fever 310 under both condition 300 and condition 340. EX1001, 18:20-35, FIG. 4A; EX1003 ¶202.

b. Element [8b]

This element repeats Element [1d] for the second medical condition. Goldberg's editor builds the knowledge base condition by condition, as FIGS. 4 (Celiac disease) and 8A (myocardial infarction) show, so repeating the editorial step

for a second condition is the combined system's normal operation. EX1006, 6:35-52, FIGS. 4, 8A; EX1003 ¶¶204-206.

c. Element [8c]

Because each diagnostic factor in Datla is a single, condition-independent node linked to many conditions through separate edges, the second condition's factor group links to the same factor entry already stored for the first condition. EX1005 ¶[0031]; EX1003 ¶¶208-209. Dang likewise connects the same symptom node to feature nodes for different diseases. EX1008 ¶[0030], FIG. 1B; EX1003 ¶174.

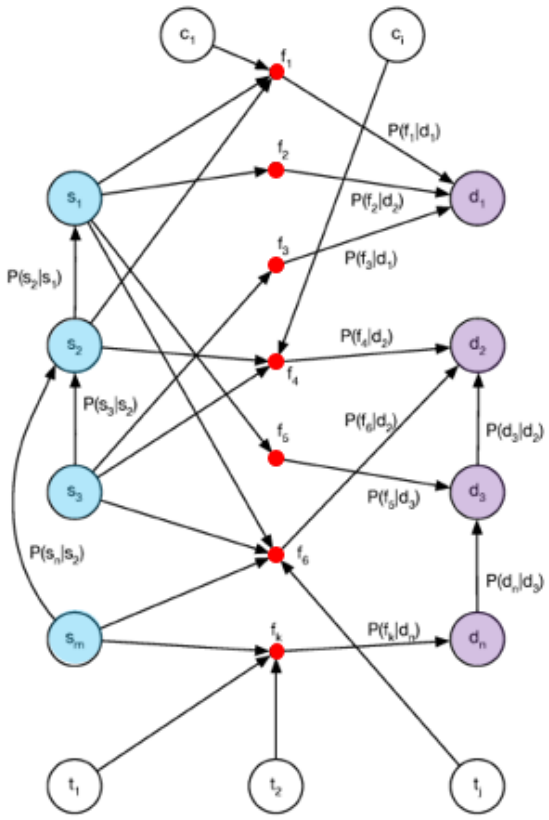


FIG. 1B

EX1008, FIG. 1B (annotated)

Goldberg's FIG. 5 shows the same diagnostic factors across multiple diseases with different likelihood values, confirming reuse of factor entries, and the '674 patent describes the same architecture: "EDFs remain elemental, independent, and can be used computationally across conditions." EX1006, 9:13-29, FIG. 5; EX1001, 10:49-50; EX1003 ¶¶210-211.

7. Claim 9

Claim 9 requires that traversing from a shared factor to the second condition produces a different rank than traversing to the first. EX1003 ¶¶212-213. Datla

weights each edge by "the frequency with which the two nodes are associated within a source or a corpus," and "[t]he weight may be variable and based on the strength of the relationship." EX1005 ¶[0032]; EX1003 ¶214. Because fever co-occurs with different conditions at different frequencies, its edge weights to those conditions differ, and traversing to each produces a different rank. EX1005 ¶¶[0031]-[0032]; EX1003 ¶¶215-216.

Goldberg independently assigns condition-specific probability values to the same factor, and in the combination, those ranks reside in each condition's factor group rather than in the shared factor entry, just as the '674 patent describes: "the noise value for fever 310 will be the same for condition 300 as well as condition 340, although the rank of factor groups 304 and 344 will be different." EX1006, 6:35-53, 9:13-29, FIG. 5; EX1001, 25:1-3; EX1003 ¶¶217-219.

8. Claim 11

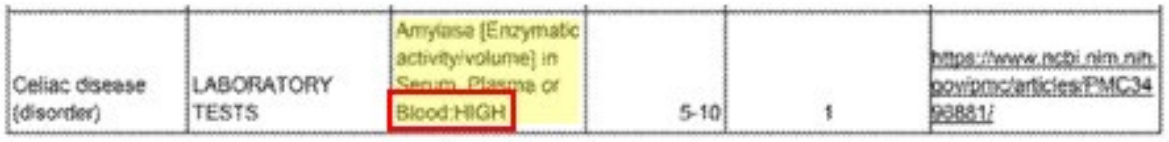
a. Element [11a]

Datla's graph includes "tests" as a node category, and Datla uses the extracted keywords "to generate the diagnoses, tests, and treatments using the underlying clinical contexts," thus producing laboratory tests indicative of the condition. EX1005 ¶¶[0031], [0034]; EX1003 ¶¶221-222.

b. Element [11b]

Goldberg's editor displays all clinical fact types to medical experts; its knowledge base includes a LABORATORY TESTS category and FIG. 1 shows

entries such as "Amylase [Enzymatic activity/volume] in Serum, Plasma or Blood:HIGH" for Celiac disease. EX1006, 13:16-21, FIGS. 1, 4, 8B; EX1003 ¶¶224-225.[5]



Celiac disease (disorder)	LABORATORY TESTS	Amylase [Enzymatic activity/volume] in Serum, Plasma or Blood:HIGH	5-10	1	https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3498881/

EX1006, FIG. 1 (excerpted and annotated)

c. Element [11c]

Element [11c] requires a rank for the investigation group. The '674 patent describes ranking investigation groups by the order in which they should be performed, with the final ranking determined by human input. EX1001, 13:38-43, 13:43-53; EX1003 ¶227. An investigation group's ranking is also the ranking of each of its subsumed laboratory tests. EX1001, 26:61-27:6; EX1003 ¶228. A POSITA would therefore have understood the claimed "rank for said investigation group" to be a priority or ordering assigned to a group of laboratory tests for a given condition. EX1003 ¶229. This rank tells the clinician which test or tests to perform first and how strongly the group's tests bear on the condition. *Id.*

[5] Patent Owner alleges that this element is satisfied merely by "[p]roviding documentation that is available to clinicians." EX1017, 54.

Goldberg Discloses the Investigation Group and Expert Input. Goldberg's LABORATORY TESTS category is an investigation group: just as the SYMPTOM/COMPLAINT and RISK FACTORS categories subsume diagnostic factors, it subsumes a condition's laboratory tests, such as Amylase (5-10%), Ceruloplasmin (2-5%), and Endomysium Ab (90-95%) for Celiac disease. EX1006, FIGS. 1, 4; EX1003 ¶231. The expert enters or confirms these Probability values, which rank the tests' diagnostic significance, through Goldberg's editor. EX1006, FIGS. 1, 4; EX1003 ¶231. Where an investigation group holds one test, or the expert assigns a group's tests the same value, this probability value is also the rank "applicable to each of the EDIs to which it is linked," as the '674 patent describes. EX1001, 21:43-45; EX1003 ¶232.

Dang Supplies a Rank Held at the Group Level. Dang's knowledge graph includes "a set of factual nodes (symptom nodes, test nodes and context nodes), a set of feature nodes, and a set of disease nodes," where test nodes connect upward to feature nodes and each feature node connects to a disease node through a conditional probability. EX1008 ¶¶[0028]-[0029], [0031], FIG. 1B. Dang computes a value at the feature-node level. EX1008 ¶[0049] ("Confidence of feature = Prod(Confidence of evidence*weight)"); EX1003 ¶233.

It would have been obvious for the expert to assign each investigation group a priority ordering in the combined system. EX1003 ¶234. Goldberg uses its

knowledge base to suggest a "next best step" for patient evaluation. EX1006, 7:25-36; EX1003 ¶234. A system recommending which test to perform next would benefit from knowing which group of tests clinicians perform first for that condition, and an expert already using Goldberg's editor to group and value laboratory tests by condition would naturally record that ordering at the group level, as Dang's feature-node rank allows. EX1003 ¶234. Applying a known clinical convention (tiered testing) to a known editorial workflow (Goldberg) and graph structure (Dang) would predictably yield an investigation group with an editorially assigned rank. *Id.*

A POSITA would have been motivated to prioritize tests within an investigation group for various reasons, such as ranking less expensive tests over more expensive ones, less invasive tests over more invasive ones, and locally available tests over those requiring transfer to another facility. EX1003 ¶235.

d. Element [11d]

In the combination, the investigation group is a feature node in Dang's three-tier architecture, adopting Goldberg's LABORATORY TESTS grouping. EX1003 ¶237. Dang's architecture includes sets of test, feature, and disease nodes. EX1008 ¶[0029], FIG. 1B. The combined system therefore stores the investigation group and laboratory test as separate entries, with the condition linked to the investigation group and the investigation group linked to the laboratory test, as claimed. EX1003 ¶¶237-238.

Datla's test nodes, like its symptom nodes, link to multiple conditions through weighted edges and therefore include no condition-specific information. EX1005 ¶¶[0031]-[0032]; EX1003 ¶239. Goldberg stores each disease-feature association with distinct Source Disorder, Type of Rule/Feature, and Rule Target fields, and the Type field is separately queryable across diseases. EX1006, 8:47-62, FIGS. 1, 4; EX1003 ¶240. Applying the claim 1 architecture to the investigation side yields the claimed condition-investigation group-laboratory test chain, with condition-specific information in the investigation group rather than the test entry. EX1003 ¶242.

9. Independent Claim 12

Each limitation of claim 12 is disclosed or rendered obvious by the combination for the reasons given above and below; claims 12-15 do not depend on Goldberg's human-input rationale. *Supra* §VIII.A.1; EX1003 ¶¶243-244.

a. Element [12pre]

To the extent the preamble is limiting, Datla discloses a method of determining a state of a patient. EX1003 ¶¶245-247. Datla "accepts natural language input from a medical professional about a patient's scenario" and identifies "a ranked list of one or more medical conditions," which is how the '674 patent describes determining patient state. EX1005 ¶¶[0001], [0007], [0028]; EX1001, 22:33-35, 24:57-59; EX1003 ¶246. Dang likewise determines "a possible disease list." EX1008 ¶[0056]; EX1003 ¶247.

b. Element [12a]

Element [12a] combines Elements [1a]-[1b] and [1e], without the factor group. *Supra* §§VIII.A.2.b-c, VIII.A.2.f. Datla extracts diagnostic factors from medical journals, Wikipedia, and PubMed; to the extent Datla does not expressly label graph construction as NLP, Dang does, and Datla already includes an NLP engine and updates its graph with "new journal articles." EX1005 ¶¶[0030], [0032], [0034]-[0036], [0040], [0052]; EX1008 ¶¶[0025], [0027]; EX1003 ¶¶249-250. Datla stores each symptom and condition as its own linked node, with condition-specific weights on edges rather than in the symptom node, as does Dang. EX1005 ¶¶[0030]-[0032], [0078]; EX1008 ¶¶[0027]-[0031]; EX1003 ¶¶251-254.

c. Element [12b]

The '674 patent describes its demographics node as holding condition-specific "age and sex restriction[s]" that are "unique to the condition and the EDF, but separate from the EDF." EX1001, 10:41-49, 19:27-34; EX1003 ¶256. Datla requires stored demographic knowledge: it extracts patient demographics and lowers a disease's rank "if a disease is not common for a demographic," which is possible only if the system stores the demographics in which each disease is or is not common. EX1005 ¶¶[0051]-[0052], [0061], [0076]; EX1003 ¶257. Storing such knowledge to limit diagnostic inference was long established: INTERNIST-1's

"TABOOs" identified diseases that "should not be considered" in certain patient populations, such as endometriosis in males. EX1035, 462; EX1003 ¶258.

Datla does not specify how that knowledge is represented, but Dang supplies a context node located between the disease and symptom nodes, and a POSITA would have stored Datla's demographic knowledge there for the reasons given for claim 7. EX1008 ¶[0027], FIG. 1A; *supra* §VIII.A.5; EX1003 ¶¶259-263. Because a demographic qualifier's diagnostic significance depends on the disease, a POSITA would have stored it on the path between the specific factor and the specific disease rather than in either node, as Zhong's demographic nodes and demographic-conflict filtering confirm. EX1009, 3:20-25, 4:18-20, 6:4-5; EX1003 ¶¶264-270.

d. Element [12c]

Datla's graph includes test nodes extracted from its corpus, and Dang's graph includes "a lab test and check node" and test nodes $\{t_1 \ldots t_j\}$ extracted from medical guidelines using NLP. EX1005 ¶¶[0031], [0034]; EX1008 ¶¶[0025], [0027]-[0029], FIGS. 1A-1B; EX1003 ¶¶271-272. For the reasons given for Element [12a], each test would be stored as its own condition-independent node, consistent with LOINC's identification of tests independent of any disease and with Goldberg's LOINC-coded test entries, whose condition-specific probability is held in the rule rather than the test identity. EX1006, 6:17-20, FIGS. 1, 4; EX1003 ¶¶273-274.

e. Element [12d]

Datla extracts "lab results" from patient input but does not expressly store, in its graph, the results required to indicate a condition. EX1005 ¶¶[0031], [0052]; EX1003 ¶276. Goldberg does: its LABORATORY TESTS rule for Celiac disease, "Amylase [Enzymatic activity/volume] in Serum, Plasma or Blood:HIGH," has two colon-separated components—the test identity and the result qualifier "HIGH"—and Goldberg's other test rules, such as "Ceruloplasmin [Mass/volume] in Serum or Plasma:LOW," follow the same pattern, as FIG. 4 shows. EX1006, 6:65-7:2, FIGS. 1, 4; EX1003 ¶277.

LABORATORY TESTS	Amylase [Enzymatic activity/volume] in Serum, Plasma or Blood:HIGH	5-10	
COMPLICATION	Anemia of chronic disorder (disorder)		5-10
SIGNS	Ataxia (finding)	1-2	
ASSOCIATED CONDITION	Autoimmune hepatitis (disorder)	1-2	5-10
LABORATORY TESTS	Ceruloplasmin [Mass/volume] in Serum or Plasma:LOW	2-5	

EX1006, FIG. 4 (excerpted and annotated)

Goldberg's qualifier is the content the '674 patent places in a results node: "the result of the test (described using terms such as 'high' or 'low', using ranges, using binary values, using ordinal types, etc.)." EX1001, 22:5-10; EX1003 ¶278. Dang's FIG. 1A positions an Observation Item node between the Disease node and the Lab Test & Check node, a natural location for a result indicating a disease, particularly because symptoms, signs, and context have their own node types. EX1008 ¶[0027], FIG. 1A; EX1003 ¶¶279-281.

It would have been obvious to store Goldberg's condition-specific results in a results node between the condition and test nodes, such as a node like Dang's Observation Item node. EX1003 ¶271. Datla, Dang, and Goldberg address the same problem; Datla already extracts "lab results" as diagnostic inputs; Goldberg shows a knowledge graph storing the result associated with a disease; and storing a condition-specific result in the test node would break that node's reuse across conditions—the same principle that keeps condition-specific information out of factor nodes. EX1005 ¶[0052]; EX1006, FIG. 1; EX1003 ¶¶283-287. Independently of Dang, implementing Goldberg's colon-delimited separation as its own node is a routine graph-design choice. EX1003 ¶288. Storing a value such as "HIGH" in a node and comparing it with patient data was routine, so a POSITA would have reasonably expected success. EX1006, 6:17-20; EX1003 ¶289.

f. Element [12e]

Datla matches patient data from a patient medical chart to diagnostic factor nodes. Datla uses MIMIC discharge notes, which contain "comprehensive clinical scenarios," and its NLP engine extracts the patient's symptoms, which the system uses to query the graph, activating matching symptom nodes. EX1005 ¶¶[0039], [0052], [0055], [0057]; EX1003 ¶¶292-293. Datla's engine also extracts "lab results," which correspond to its test nodes, and Datla displays ranked conditions, adjusted by demographics, through a user interface—a result indicating a likelihood

that the patient has the condition. EX1005 ¶¶[0031], [0052], [0061]-[0062]; EX1003 ¶¶294-296.

In the combination, matching patient data to the demographics and results nodes would have been obvious. EX1003 ¶297. Datla already extracts the patient's demographics and lab results, and the combination provides corresponding nodes against which to match them. EX1005 ¶¶[0052], [0061]; EX1003 ¶298. Dang's diagnosis process matches patient data to all factual node types—"symptom nodes, test nodes, context nodes"—using semantic matching, and Goldberg uses demographics and lab results entered for a patient to generate a differential. EX1008 ¶[0047], FIG. 5; EX1006, 8:14-16; EX1003 ¶299. Adding those nodes without matching patient data to them would serve no purpose, and the matching yields the same ranked output Datla already produces. EX1005 ¶¶[0055], [0057], [0061]-[0062]; EX1003 ¶¶300-301.

10. Claim 13

Datla stores rank information as weighted edges between conditions and diagnostic factors and uses those weights to determine likelihood through activation and decay. EX1005 ¶¶[0006], [0032], [0057]; EX1003 ¶¶302-306. Dang's intermediate feature nodes, between disease and factual nodes, carry P(f|d) and an aggregated feature confidence, and Dang determines disease likelihood from those values: "Confidence of disease = Sum(Confidence of evidence*TF-IDF)." EX1008

¶¶[0028]-[0031], [0049]-[0050]; EX1003 ¶¶307-311. A POSITA would have incorporated Dang's feature tier into Datla because it organizes the same rank information more structurally, aggregates evidence from multiple factual nodes before disease-level inference in a manner mirroring clinical reasoning, and uses each element for its known purpose with predictable results. EX1003 ¶¶312-315.

11. Claim 14

Claim 14 applies claim 13's intermediate-node architecture to laboratory tests. EX1003 ¶¶316-317. Dang's factual tier expressly includes test nodes $\{t_1 \ldots t_j\}$, and its feature nodes aggregate evidence from all factual nodes, so the same feature-level confidence values hold the rank of the tests and determine disease likelihood. EX1008 ¶¶[0028]-[0031], [0049]-[0050]; EX1003 ¶¶318-320. Because Datla's graph includes test nodes, incorporating Dang's feature tier provides the claimed intermediate node for tests without additional motivation, and Goldberg's ranked LABORATORY TESTS category provides further support. EX1005 ¶¶[0031], [0034]; EX1006, FIGS. 1, 4; EX1003 ¶¶321-323.

12. Claim 15

Datla's NLP is initialized using existing database entries. EX1003 ¶¶325-326. Datla's KV-MemNN model is trained on Wikipedia clinical-medicine pages organized by diagnosis and their associated clinical concepts, and Datla resolves extracted diagnosis terms against existing entries: "[s]ince certain diagnosis terms

from MIMIC-II do not exactly match the Wikipedia page titles, some embodiments use the Wikipedia API to search for the most appropriate Wiki page by using each diagnosis term as the search keyword." EX1005 ¶¶[0035], [0039]-[0041]; EX1003 ¶327. In the combined graph, a POSITA would have initialized the NLP with the existing diagnostic factor and laboratory test nodes to recognize and resolve newly extracted concepts. EX1003 ¶328.

B. Ground 2: Claims 4-5 are Unpatentable Over Datla, Dang, Goldberg, and Kim

Claims 4-5 would have been obvious over Datla, Dang, Goldberg, and Kim. EX1003 ¶¶329-352. In the combination, Kim supplies ontology identifiers and synonyms for recognizing semantically equivalent concepts across sources. EX1012 ¶¶[0041], [0050]-[0055], [0060].

1. Motivation to Combine

A POSITA would have added Kim to the Ground 1 combination for three reasons. EX1003 ¶¶330-335. First, Kim, like Datla and Goldberg, concerns medical knowledge graphs supporting clinical decision-making, and a system ingesting concepts from heterogeneous sources benefits from recognizing when sources use different words for the same concept—the problem Kim solves by annotating expressions that mean "semantically the same" with "the same identifiers of an ontology." EX1012 ¶¶[0038]-[0039], [0050]; EX1003 ¶331.

Second, Datla and Goldberg already point toward standard ontologies. Datla states that its system "may extract term frequency-inverse document frequency (TFIDF) weighted topical keywords from the given descriptions, summaries or notes and map them to categories represented in one or more ontologies, including but not limited to the following controlled clinical ontologies: SNOMED CT for diagnoses, LOINC for tests, and/or RxNorm for treatments," and Goldberg derives graphs from SNOMED and LOINC. EX1005 ¶[0054]; EX1006, 6:10-24; EX1003 ¶332. Kim supplies the conventional practice of storing the ontology code and synonyms in the graph entry. EX1012 ¶¶[0051], [0054]-[0055]; EX1003 ¶¶332, 335.

Third, annotating graph entries with ontology codes and synonyms was known, and Kim identifies off-the-shelf tools, such as MedCAT, for doing so. EX1012 ¶[0053]; EX1003 ¶333. Kim's benefits apply to any multi-source medical graph, not just its structured-report setting. EX1012 ¶¶[0050]-[0051], [0054]; EX1003 ¶334.

2. Claim 4

Kim discloses concepts "represented as nodes" that "may be identified by a unique ontology code," explaining that "[u]sing unique node ids/codes, e.g. ontology codes, allows the unambiguous identification of every concept instance," and that graph entries "may themselves be annotated with identifiers from several different ontologies," such as "the SNOMED CT code SCTID: 93880001." EX1012

¶¶[0041], [0051]-[0052], [0060]; EX1003 ¶¶337-338. A SNOMED CT identifier designates, and thereby links to, the ontology entry representing the concept, and Kim uses the identifier to reach "related concepts." EX1012 ¶[0052]; EX1003 ¶¶338-339. Given Kim's teachings, it would have been obvious to include the corresponding unique identifier in each diagnostic factor node. EX1012 ¶¶[0041], [0051]; EX1003 ¶340.

A POSITA would have done so for three reasons. First, the identifier resolves terminology variation: Datla acknowledges that "certain diagnosis terms from MIMIC-II do not exactly match the Wikipedia page titles," and Kim teaches that ontology identifiers let differently worded expressions be recognized as the same concept. EX1005 ¶[0040]; EX1012 ¶[0050]; EX1003 ¶342. Second, a shared code lets extracted patient concepts be matched to nodes by code rather than by imperfect text matching. EX1005 ¶¶[0040], [0055]; EX1003 ¶343. Third, the identifier connects the node to the ontology's structure, such as the SNOMED is-a relations Goldberg illustrates. EX1006, 1:65-2:1, 9:50-55, FIG. 6; EX1012 ¶[0052]; EX1003 ¶344. A POSITA would have reasonably expected success because Goldberg shows clinical graphs combined with ontologies, MedCAT automates identifier assignment, and an identifier is a node property that leaves Datla's edges and activation unchanged. EX1006, 5:47-62; EX1012 ¶[0053]; EX1005 ¶¶[0032], [0036]; EX1003 ¶345.

3. Claim 5

Patent Owner alleges that SNOMED CT-coded entries incorporate "SNOMED CT concept entries," which "inherently include synonym descriptions as part of their standard data model." EX1017, 27-28. Goldberg's hybrid knowledge graph includes "a first knowledge graph [] derived from a healthcare Ontology which includes a large vocabulary of medical concepts. (e.g., SNOMED)," and its clinical graph "links medical concepts that are described in the Ontological graphs." EX1006, 6:10-30. Under Patent Owner's reading, the combination's diagnostic factor entries, drawn from Goldberg's SNOMED-derived concepts, would include synonyms. *Supra* §VIII.B.1.

Kim discloses that "the graph database includes synonyms, translations, concepts and relationships," and ties synonyms to individual entries: "in order to add translations or synonyms for terms of existing ontology concepts to one or more entries of the graph database, the respective concept might have to be identified using existing terms describing it." EX1012 ¶¶[0054]-[0055], [0060]; EX1003 ¶347.

It would have been obvious to include in each diagnostic factor node at least one synonym (e.g., "pyrexia" for fever) for three reasons. EX1003 ¶¶348-349. First, as Kim teaches, synonyms prevent duplicate nodes as Datla's graph is updated with "new journal articles." EX1005 ¶[0032]; EX1012 ¶[0054]; EX1003 ¶349. Second, synonyms improve matching of clinical text to the correct node, which Datla

recognizes is imperfect. EX1005 ¶¶[0040], [0052], [0055]; EX1012 ¶¶[0060], [0073]; EX1003 ¶350. Third, synonyms complement claim 4's identifier, because standard ontologies provide multiple terms for each concept. EX1012 ¶[0055]; EX1003 ¶351. Storing text strings as node properties is routine, so success would have been expected. EX1012 ¶[0054]; EX1003 ¶352.

C. Ground 3: Claim 6 is Unpatentable Over Datla, Dang, Goldberg, and Van Assel

Claim 6 would have been obvious over Datla, Dang, Goldberg, and Van Assel, which supplies a stored temporality attribute. EX1003 ¶¶353-365. Claim 6 is disjunctive, so a temporality attribute suffices. The '674 patent describes temporality as "whether or not the patient is experiencing the factor now," "a binary value indicating whether the factor was experienced in the past or now." EX1001, 19:36-40, 21:14-24; EX1003 ¶364.

Datla may "filter out negated clinical concepts in order to give more weight to positive clinical manifestations," but it does not expressly state that it stores an attribute in each factor entry. EX1005 ¶[0054]; EX1003 ¶360. Van Assel does: each entry may include "date information indicating the day on which the entry was reported"; Van Assel's validity module compares the time since an entry was reported with "stored information relating to each concept, indicating the duration of time for which information relating to that concept is valid"; and Van Assel stores permanence in the knowledge base as the concept "Is Permanent." EX1011

¶¶[0264], [0270], [0279]; EX1003 ¶¶361-363. These attributes indicate whether a factor is current or historical—the function of the claimed temporality attribute. EX1003 ¶¶364-365.

Patent Owner's infringement allegations confirm that Van Assel's attributes meet claim 6. Patent Owner contends that time-qualified metrics, such as "Spirometry performed < 6 months ago" and "Flu vaccination <12 months," are "temporality attributes, as claimed," and that "'No' values" indicating whether a concept appears in a patient record "demonstrate negation attributes." EX1017, 28-29. Van Assel's stored reporting date and validity duration are the same kind of time qualifiers, and its present/absent finding values are the same kind of "No" values. EX1011 ¶¶[0245], [0264], [0270].

A POSITA would have incorporated Van Assel's temporality attribute into Datla's diagnostic factor entries for the reasons given for Ground 1 and because both references perform automated diagnosis using structured knowledge bases and probabilistic models. *Supra* §VIII.A.1; EX1011 ¶¶[0001]-[0004]; EX1005 ¶¶[0001], [0028]; EX1003 ¶¶354-356. Van Assel teaches that temporal validity affects accuracy: a symptom reported months ago may no longer be relevant, while a chronic condition's permanence must be tracked. EX1011 ¶¶[0264], [0270], [0279]; EX1003 ¶357. Datla processes clinical narratives containing temporal information but stores no temporality attribute; adding one would let the system distinguish

current from historical factors, using each element for its known purpose with predictable results. EX1005 ¶[0039]; EX1003 ¶¶357-358.

D. Ground 4: Claims 16-18 and 20-23 are Unpatentable Over Datla, Dang, and Van Assel

Claims 16-18 and 20-23 would have been obvious over Datla, Dang, and Van Assel. EX1003 ¶¶366-443. Datla and Dang disclose or render obvious each element; Van Assel provides additional, express disclosure to the extent Patent Owner contends otherwise. *Id*. ¶¶366, 379.

1. Motivation to Combine

A POSITA would have combined Datla and Dang for the reasons given for Ground 1 and because both generate, from medical literature, a graph linking findings to diseases and use it with patient data to rank likely diagnoses. *Supra* §VIII.A.1; EX1005 ¶¶[0001], [0006]-[0007]; EX1008 ¶¶[0002], [0025]; EX1003 ¶¶367-368. Datla requires demographic knowledge but does not specify its representation, and Dang supplies context and patient-profile nodes; Datla weights edges by co-association frequency, while Dang teaches disease-conditional probabilities that it states improve accuracy; and both include laboratory tests as graph concepts. EX1005 ¶¶[0031]-[0032], [0037], [0061]; EX1008 ¶¶[0027]-[0031], [0036], [0040]; EX1003 ¶¶369-371. Both use conventional graph structures, NLP, and scoring arithmetic, so success would have been expected. EX1003 ¶372.

A POSITA also would have incorporated Van Assel's ontology-based knowledge graph into the Datla/Dang system. Van Assel, like Datla and Dang, is directed to knowledge-graph-based medical diagnosis from patient findings. EX1011 ¶¶[0113], [0115]; EX1003 ¶¶373-374. First, Van Assel teaches a medical ontology knowledge graph in which each concept has a label and a unique identifier (an IRI) and is related to parent and child concepts by is–a relationships. EX1011 ¶¶[0115], [0117]–[0118], [0123]–[0124]; EX1003 ¶374. Second, Van Assel extracts both present and absent findings via NLP and maps them to knowledge-base concepts. EX1011 ¶¶[0112], [0125], [0245]; EX1003 ¶375. Third, Van Assel uses its ontology hierarchy to identify all ancestors (hypernyms) and descendants (hyponyms) of a node, and extends concept lists with neighboring concepts from the knowledge base. EX1011 ¶¶[0300], [0325], [0349]–[0350]; EX1003 ¶375.

A POSITA would have recognized that linking each diagnostic factor entry to an ontology concept (as in Van Assel), and using the ontology's is–a hierarchy to relate specific patient findings to more general ones, would improve the Datla/Dang system's ability to match patient findings recorded at different levels of specificity to the symptom nodes in the diagnostic graph. EX1003 ¶376. This is the use of a known technique to improve a similar system in the same way, with predictable results and a reasonable expectation of success, because Van Assel's ontology

operates on the same kinds of concepts (symptoms, diseases, risk factors) as Datla and Dang. *Id.*

A POSITA also would have been motivated to use Van Assel's present/absent coding in the combined Datla/Dang system. Datla may "filter out negated clinical concepts in order to give more weight to positive clinical manifestations," and Van Assel teaches that retaining both present and absent findings improves diagnosis. EX1005 ¶[0054]; EX1011 ¶[0219]; EX1003 ¶378. Recording a negated concept as absent, rather than discarding it, is a known technique to improve a similar system in the same way, and a POSITA would have reasonably expected success because Datla's NLP already identifies the negation. EX1003 ¶378.

2. Independent Claim 16

a. Element [16pre]

Element [16pre] is identical to Element [12pre] and is met for the same reasons. *Supra* §VIII.A.9.a; EX1003 ¶¶380-381.

b. Element [16a]

Datla inputs portions of patient records into its NLP. Datla extracts weighted keywords "from the given descriptions, summaries or notes," including "actual admission notes derived from MIMIC-III," which are part of the patient's EHR, and Dang begins from the patient's "electronic medical record (EMR)." EX1005 ¶¶[0048], [0050], [0054], [0057]; EX1008 ¶[0036]; EX1003 ¶¶383-384. Datla's NLP engine extracts present symptoms and may "filter out negated clinical

concepts," distinguishing present from absent findings. EX1005 ¶¶[0052], [0054]; EX1003 ¶¶385-386.

To the extent Patent Owner contends that Datla does not disclose Element [16a], Van Assel teaches producing a plurality of diagnostic factors known to be present or absent in said patient. Van Assel's diagnostic engine uses an NLP module to extract medical concepts from physician notes, store them in the patient's clinical history, and code them into concept mapping to the knowledge base. EX1011 ¶[0239]; EX1003 ¶387. The clinical notes indicate whether the concept was present or absent. EX1011 ¶¶[0137], [0145]; EX1003 ¶387. Van Assel then assigns each extracted concept a value indicating "whether the concept is validly present (1), validly absent (-1) or this is unknown (0)." EX1011 ¶[0245]; EX1003 ¶387. Van Assel likewise extracts from patient input "concepts which are used in the model and are identified by the user as absent." EX1011 ¶¶[0128], [0130]; EX1003 ¶387.

A POSITA would have recorded Datla's negated concepts as absent, as Van Assel does, rather than discarding them, because "[h]aving values for more of the nodes enables the diseases to be computed with greater certainty," and would have expected success because Datla already may "filter out negated clinical concepts." EX1005 ¶[0054]; EX1011 ¶[0219]; EX1003 ¶378.

c. Element [16b]

Datla matches extracted patient symptoms to symptom nodes when it "queries the knowledge [graph] using the extracted one or more patient symptoms," and those nodes are condition-independent, as explained for Element [1e]. EX1005 ¶¶[0031], [0055]; *supra* §VIII.A.2.f; EX1003 ¶¶389-390. Van Assel confirms that a common knowledge-graph architecture stores each concept as only a label and identifier, with relational information in separate triples. EX1011 ¶¶[0117], [0123]-[0124]; EX1003 ¶391. Van Assel's knowledge base applies this architecture at scale: 31.7 million triples over 8.3 million concepts, none holding condition-specific information. *Id*.

d. Element [16c]

Datla traverses links from matched symptom nodes to linked conditions through "one-hop expansion of the symptom nodes," with activation weights that propagate along each link. EX1005 ¶¶[0055], [0057]; EX1003 ¶¶393-395. Because Datla's edges exist because the symptom is a manifestation of the linked condition, the propagated activation and the pre-computed edge weights are weight values indicating that the factor is caused by the condition. EX1005 ¶¶[0031]-[0032], [0057]; EX1003 ¶¶396-397. Dang's inference likewise traverses from matched factual nodes through feature nodes to disease nodes, collecting P(f|d) and confidence values, and its TF-IDF weights are tied to individual feature-disease pairs. EX1008 ¶¶[0031], [0037], [0049]-[0050], [0053]; EX1003 ¶¶398-401. It

would have been obvious to store such a disease-specific weight on each Datla symptom-condition edge and collect it during one-hop expansion. EX1003 ¶401.

Van Assel provides additional evidence and teaches that diagnosis considers causal and probabilistic relationships between risk-factors, diseases, and symptoms, where such relationships describe the strength of the connections between the concepts. EX1011 ¶¶[0003], [0208]-[0209], FIG. 3(a); EX1003 ¶402.

e. Element [16d]

Datla's control module "controls the accumulation of activations at each node," so activation from multiple symptom nodes accumulates (sums) at each condition node. EX1005 ¶¶[0006], [0057]-[0058], [0071], [0084]; EX1003 ¶¶403-405. Dang expressly sums: "Confidence of disease = Sum(Confidence of evidence*TF-IDF)," adding a weighted contribution from each matched finding. EX1008 ¶¶[0048], [0050]; EX1003 ¶¶406-407.

f. Element [16e]

Datla provides the conditions, ranked by accumulated activation weight, to the clinician "via any user interface that allows information to be conveyed, such as a speaker or screen," and Dang sorts its possible disease list and outputs it, including "via a graphical user interface." EX1005 ¶¶[0011], [0062]; EX1008 ¶[0056]; EX1003 ¶¶408-410.

3. Claim 17

Datla identifies "a ranked list" of conditions and seeks to recommend diagnoses "in addition to the rationale for each option," which communicates the basis for the ranking, including the underlying weight. EX1005 ¶¶[0003], [0007]; EX1003 ¶¶411-414. A ranked list displaying conditions in order of weight necessarily associates each condition with information reflecting its total weight, and Dang likewise displays a possible disease list sorted by probability. EX1008 ¶[0056]; EX1003 ¶¶414-415.

4. Claim 18

In Van Assel, each concept is stored as an entry with a label and an identifier that "uniquely identifies" the concept and "may be a link to an internet based resource," e.g., a "Frontal headache" concept with an "iri" field. EX1011 ¶¶[0117]-[0118], [0123]; EX1003 ¶¶416-417. Because Van Assel expresses each is-a relationship in terms of IRIs (e.g., "<Breast cancer, isA, Cancer>"), the IRI in an entry is the medical ontology link from which its ancestors are located—the same structure the '674 patent describes, using the factor "as an index into the medical taxonomy." EX1011 ¶[0124]; EX1001, 20:28-30, 23:55-57; EX1003 ¶418. Van Assel's graph is a directed acyclic graph defined by "subClassOf" edges, and it retrieves "[a]ll of the ancestors" of each node, which is traversing upward. EX1011 ¶¶[0332], [0334], [0349]-[0350]; EX1003 ¶419.

Because a child concept is a kind of its parent, a patient with a present child finding (e.g., frontal headache) necessarily has the parent finding (e.g., headache), so applying Van Assel's ancestor retrieval to each present finding infers other present findings. EX1003 ¶420. Van Assel itself replaces patient EHR entries with their IRIs and extends them with ancestor concepts, and its distinction between present (1) and absent (-1) findings lets the inference be limited to present factors. EX1011 ¶¶[0245], [0299]-[0300]; EX1003 ¶¶420-421; EX1010, 1119. In the combination, the inferred findings are matched to symptom nodes as in Element [16b], so they contribute to scoring. EX1008 ¶¶[0047]-[0050]; EX1011 ¶[0300]; EX1003 ¶422.

5. Claim 20

Claim 20 recites the same limitation as claim 6, and that analysis applies equally. *Supra* §VIII.C. Van Assel's stored reporting date, validity duration, and permanence attributes satisfy the disjunctive claim. EX1011 ¶¶[0264], [0270], [0279]; EX1003 ¶¶423-424.

A POSITA would also have been motivated to incorporate Van Assel's temporality attribute into the diagnostic factor entries of the Datla/Dang system for the reasons given for Ground 3. *Supra* §VIII.C. For example, temporal context improves diagnostic accuracy by disregarding stale symptoms. EX1011 ¶¶[0264], [0270], [0279]; EX1003 ¶377.

6. Claim 21

Claim 21 is substantively the same as the NLP-initialization limitation of claim 15, and the same analysis applies. *Supra* §VIII.A.12. Datla's model is trained using Wikipedia clinical-medicine pages organized by diagnosis, so the NLP is initialized with existing entries and the links among them, and Datla resolves MIMIC-II diagnosis terms against existing pages using the Wikipedia API. EX1005 ¶¶[0039]-[0041]; EX1003 ¶¶425-429.

7. Claim 22

Claim 22 requires a demographics entry between the factor and the condition, and traversal gated on the patient matching those demographics. EX1003 ¶431. Dang's Context node supplies the first requirement for the reasons given for Elements [7b] and [12b], and Dang's FIG. 1B feeds context node c1 into feature node f1 alongside symptom nodes s1 and s2, so context determines whether and how strongly a factor supports a disease. *Supra* §§VIII.A.5.b, VIII.A.9.c; EX1008 ¶¶[0027]-[0030], FIGS. 1A-1B; EX1003 ¶¶432-433. For the second, Datla lowers a condition's rank on a demographic mismatch and also excludes candidates by demographics, filtering articles by "location, gender" and "demographic information." EX1005 ¶¶[0037], [0047], [0061]; EX1003 ¶434.

Applying the demographic check as a gate on traversal would have been obvious for three reasons. EX1003 ¶¶435-436. First, lowering a rank and excluding

a candidate are two known, predictable options, and Datla uses both. *Id*. ¶435. Second, with demographics stored on a node between the factor and the condition, the natural place to check them is when traversal reaches that node. *Id*. Third, Datla's control module already limits propagation so that "there is no runaway activation among the nodes," and stopping propagation across a link whose demographics the patient does not meet keeps implausible conditions out of the ranking, consistent with Datla's goal "to improve the accuracy of the recommendations relative to the patient scenario of interest." EX1005 ¶¶[0058], [0061]; EX1003 ¶435. Zhong confirms the approach, disregarding candidate symptoms whose demographics conflict with the patient's. EX1009, 4:18-20, 6:4-5; EX1003 ¶¶435-436.

8. Claim 23

Claim 23 is disjunctive, and Datla displays ranked conditions for patient diagnosis, its core purpose: "automated methods and systems to provide a clinical diagnosis of a patient's symptoms." EX1005 ¶[0001]; EX1003 ¶¶437-439.

E. Ground 5: Claim 19 is Unpatentable Over Datla, Dang, Van Assel, and Kim

Claim 19 would have been obvious over Datla, Dang, Van Assel, and Kim. EX1003 ¶¶440-443. A POSITA would have combined Datla, Dang, and Van Assel for the reasons given for Ground 4 and added Kim for the Ground 2 reasons, which do not depend on Goldberg: Datla ingests concepts from multiple sources, updates its graph "with new journal articles," and recognizes that clinical terms do not always

match reference terms. *Supra* §§VIII.B.1, VIII.D.1; EX1005 ¶¶[0032], [0040]; EX1003 ¶442. Neither Dang nor Van Assel alters how Datla's factor nodes are populated or matched. Claim 19 recites the same limitation as claim 5, and that analysis applies equally for claim 19. *Supra* §VIII.B.3; EX1012 ¶¶[0054]-[0055], [0060]; EX1003 ¶443.

F. Ground 6: Claims 24-25 and 30 are Unpatentable Over Zhong and Datla

Claims 24-25 and 30 would have been obvious over Zhong and Datla. EX1003 ¶¶444-513. In the combination, Zhong supplies the iterative next-best-action pipeline and its symptom-disease knowledge graph, and Datla supplies journal-based graph construction and condition-independent factor nodes. *Id*. ¶¶445-450.

1. Motivation to Combine

A POSITA would have combined Zhong and Datla because both use a medical knowledge graph to rank candidate diseases from a patient's findings and to guide further information gathering. EX1009, Abstract; EX1005 ¶¶[0001], [0006]; EX1003 ¶¶445-446. Zhong does not detail how its graph is built but teaches that it should be "in line with current medical findings," and Datla teaches constructing and continuously updating such a graph from a corpus that includes medical journals. EX1009, 3:25-29; EX1005 ¶¶[0030]-[0032]; EX1003 ¶447. Both link a single symptom node to multiple disease nodes, and Datla's shared, condition-

independent nodes with weights on the edges avoid duplication and improve scalability, a benefit a POSITA would have applied to Zhong's graph, which already stores symptom-disease values on edges. EX1009, 4:19-25, claim 1, FIG. 2; EX1005 ¶¶[0031]-[0032]; EX1003 ¶¶448, 473-476. Datla's NLP would also allow free-text input to be matched to Zhong's symptom nodes. EX1005 ¶¶[0052], [0055]; EX1003 ¶449. The combination applies known techniques to a similar system with predictable results, and neither reference teaches away. EX1003 ¶450.

2. Independent Claim 24

a. Element [24pre]

To the extent the preamble is limiting, Zhong discloses a method of determining next-best actions for a medical condition. Zhong is directed to "question generation to obtain more related medical information based on observed symptoms from a patient"—determining, for a patient with incomplete symptom information, which additional symptoms to ask about next to reach an accurate diagnosis. EX1009, Abstract, 1:21-25; EX1003 ¶¶452-453.

As shown in FIG. 3, Zhong collects observed symptoms (step 305), generates possible diseases (step 310), identifies candidate symptoms associated with those diseases (step 315), scores each candidate (step 330), and generates questions about the selected candidates (step 340). EX1009, 3:39-4:31, FIG. 3; EX1003 ¶454.

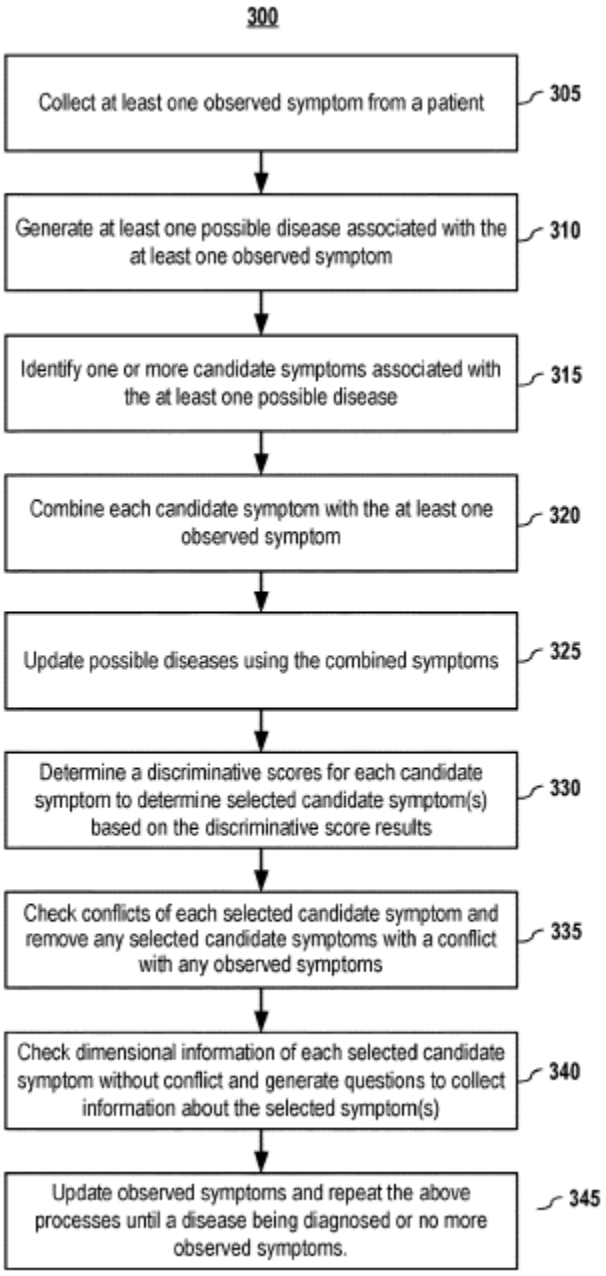


FIG. 3

EX1009, FIG. 3

The generated questions asking about the presence or absence of previously unknown diagnostic factors are the next-best actions. EX1003 ¶454. The process repeats: "the observed symptoms are updated after the information is collected from

the generated questions and the above steps may be repeated until a disease is diagnosed or the patient reports no more observed symptoms." EX1009, 4:27-31 (step 345); EX1003 ¶455. FIGS. 4-6 elaborate on steps within FIG. 3: FIG. 4 expands candidate generation and probability-based selection (steps 310-315); FIG. 5 expands discriminative scoring (step 330); and FIG. 6 expands conflict filtering, question generation, and updating (steps 335-345). EX1009, FIGS. 3-6; EX1003 ¶456.

b. Element [24a]

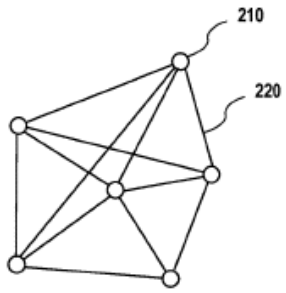
Element [24a] defines the conditions for which Elements [24b]-[24e] are performed: only those having a probability above a threshold. EX1003 ¶¶457-459. Zhong queries its knowledge graph with the observed symptoms and carries forward only the above-threshold diseases into candidate identification and scoring. EX1009, 3:65-4:2, 4:38-45; EX1003 ¶460. Zhong's claim 1 recites this structure: "forming a set of likely diseases comprising diseases from the plurality of diseases that have a probability above a threshold and a set of unlikely diseases that do not have a probability above the threshold." EX1009, claim 1; EX1003 ¶461. Zhong thus performs candidate identification, scoring, and output (Elements [24b]-[24e]) for the claimed conditions. EX1003 ¶462.

c. Element [24b]

Element [24b] requires (1) entries representing factors known to be present or absent, (2) each linked to a condition entry, and (3) each lacking condition-specific information. EX1003 ¶465.

Known factors. Zhong collects "a set of one or more observed symptoms (SO) . . . from the patient" (step 305), and, as the patient answers generated questions, updates the observed set, so it tracks factors known to be present or absent. EX1009, 3:41-44, 4:27-31, 6:20-28, FIG. 3 (step 345), FIG. 6 (step 625); EX1003 ¶¶466-467.

Entries linked to conditions. Zhong's knowledge graph is "a graph encoding domain knowledge, where a node (e.g., node 210) in the graph is an entity, such as a symptom, a dimension, a demographic, and a disease, and an edge (e.g., edge 220) in the graph represents a relationship between two entities." EX1009, 3:19-25, FIG. 2.



EX1009, FIG. 2

Each observed symptom maps to a symptom node, which is the known diagnostic factor entry. EX1003 ¶467. Zhong's claim 1 confirms that "an edge that

links a symptom node to a disease node has an associated symptom-disease pair value," and Zhong generates possible diseases "by querying the knowledge graph" from the observed symptoms—traversing those links. EX1009, 3:45-49, 8:58-67, claim 1; EX1003 ¶¶468-469.

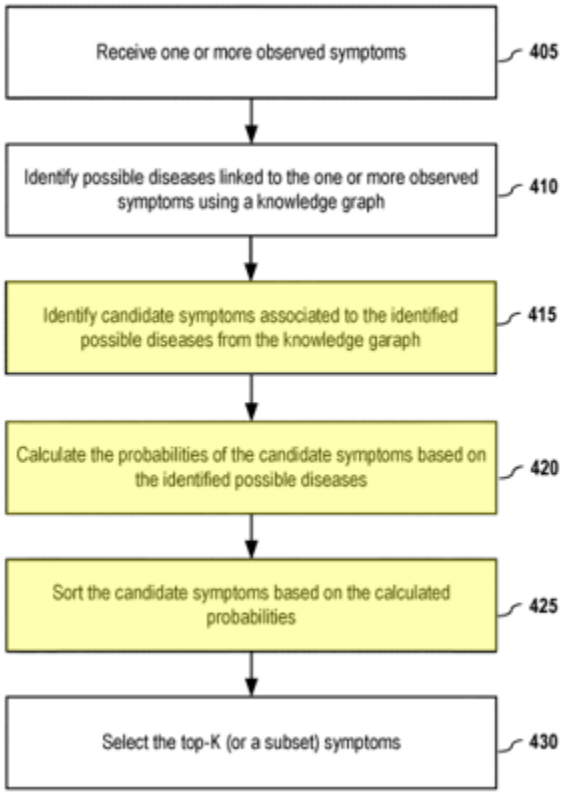
No condition-specific information. Zhong's graph has two node types connected by edges and places condition-specific information on the edge—a value "that represents a probability between the disease represented by the disease node given an observation of the symptom associated with the symptom node." EX1009, 3:19-25, 8:59-67, claim 1. A symptom node is described only as an entity representing "a symptom," so one node (e.g., fever) links to multiple diseases through separate edges, each carrying its own probability. EX1003 ¶470.

To the extent Zhong is not explicit, Datla supplies this limitation: Datla's fever node "will be connected by edges to hundreds of other nodes," with the condition-specific weight on each edge. EX1005 ¶¶[0031]-[0032]; EX1003 ¶¶471-472; *supra* §VIII.A.2.f. A POSITA would have adopted Datla's condition-independent nodes in Zhong's graph so the same symptom node is reused across conditions, avoiding duplication and ensuring consistency as the graph is updated, with predictable results. *Supra* §VIII.F.1; EX1003 ¶¶473-476.

d. Element [24c]

At step 315, Zhong identifies "a set of candidate symptoms (SC) associated with the set of possible diseases . . . using the knowledge graph," and explains that "a candidate symptom may be an unobserved symptom or a symptom not reported to the automatic diagnosis system"—a factor not known to be present or absent. EX1009, 3:50-54; EX1003 ¶¶477-479.

FIG. 4 details this step: at step 415, "M (an integer number larger than 1) candidate symptoms (SC) associated with the likely diseases are identified using the knowledge graph," and "each possible disease may associate to one or more candidate symptoms." EX1009, 4:45-51, FIG. 4; EX1003 ¶480.



EX1009, FIG. 4 (annotated)

Zhong identifies all such entries because it extracts the full set of M candidates linked to the likely diseases before any narrowing; only later steps (steps 420-430 and 505-525) calculate, sort, and select the top-K and top-N. EX1009, 4:45-6:2, FIGS. 4-5; EX1003 ¶481. Each candidate is linked to the condition entry because Zhong reaches it by traversing the symptom-disease edges from the likely diseases. EX1009, 3:19-25, 4:45-48, claim 1; EX1003 ¶¶482-483.

e. Element [24d]

Element [24d] requires (1) a rank indicating the likelihood that the unknown factor indicates the condition, and (2) that rank stored in the link between the entries. EX1003 ¶¶484-485.

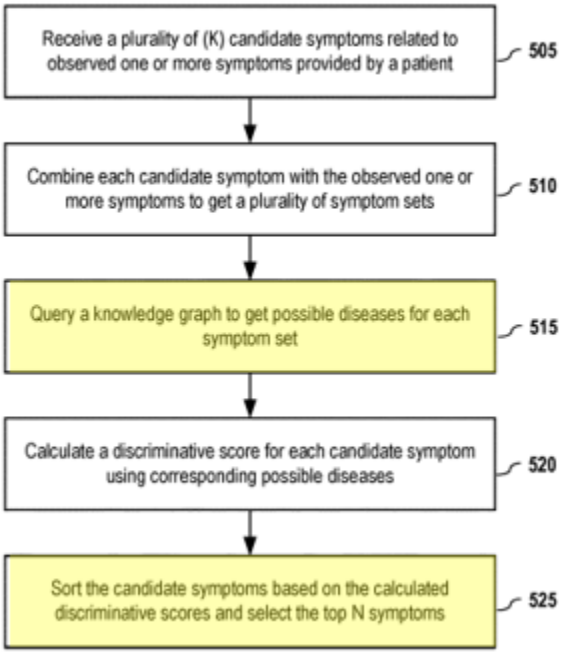
Zhong's edges carry a symptom-disease pair value "that represents a probability between the disease represented by the disease node given an observation of the symptom associated with the symptom node." EX1009, 8:60-67, claim 1. A higher value means a stronger association with the disease, so the value ranks how likely the factor is to indicate the condition. EX1003 ¶486. Zhong uses these values as inputs at step 420, where "the probabilities of the identified candidate symptoms are calculated based on the likely disease" as max P(SC|SO) (Equation 1), and at step 425 the candidates are "sorted based on the calculated probabilities." EX1009, 4:52-66, FIG. 4; EX1003 ¶487. Because "an edge . . . represents a relationship between two entities," the edge holding that value is the "database entry link between" the symptom node and the disease node. EX1009, 3:19-25, claim 1, FIG. 2 (edge 220); EX1003 ¶¶488-490.

f. Element [24e]

Display. Zhong outputs questions about the selected candidates through output interface 130 and claims "outputting one or more indicia for determining absence or presence of a candidate symptom from the selected subset of candidate

symptoms for the patient." EX1009, 2:64-67, 4:20-24, 6:13-14, 9:37-39, claim 1, FIGS. 1, 3 (step 340), 6 (step 620); EX1003 ¶493.

In association with each condition. Candidates are identified as those "associated with the likely diseases" (step 415); at step 515 Zhong queries the graph to "get likely diseases and unlikely diseases" for each candidate; and each candidate's discriminative score measures how well it separates the likely from the unlikely diseases (Equation 2). EX1009, 4:45-48, 5:13-15, 5:34-49, FIG. 5. Each question thus is presented because of its relevance to, and ability to discriminate among, the likely conditions. EX1003 ¶494.



EX1009, FIG. 5 (annotated)

In order of rank. Zhong sorts candidates by probability (step 425) and selects the "top-K candidate symptoms," then sorts them by discriminative score—derived from the edge values—and selects the "top-N" (step 525). EX1009, 4:65-66, 5:55-58, 5:67-6:2; EX1003 ¶¶495-496. The questions output at step 340 (FIG. 6, step 620) correspond to these sorted candidates, and displaying them in rank order is the purpose of Zhong's sort-select-output pipeline, which prioritizes the most diagnostically valuable unknown factors. EX1003 ¶¶497-498.

3. Claim 25

Claim 25 requires identifying the unknown factor entries by (1) traversing down through the database from each condition entry and (2) disregarding the known factor entries. EX1003 ¶¶499-500.

Traversing down. Zhong identifies candidate symptoms by starting from the above-threshold disease nodes and following the edges to the linked symptom nodes: at step 415, "M (an integer number larger than 1) candidate symptoms (SC) associated with the likely diseases are identified using the knowledge graph," and "each possible disease may associate to one or more candidate symptoms." EX1009, 4:45-51, FIG. 4 (step 415); EX1003 ¶501. Because disease nodes and symptom nodes sit at different tiers connected by symptom-disease edges, querying from each disease to its linked symptoms is the claimed downward traversal. EX1009, 3:19-29, claim 1, FIG. 2; *supra* §VIII.F.2.c; EX1003 ¶502.

Disregarding known entries. Zhong defines candidates as the complement of the known set: "a candidate symptom may be an unobserved symptom or a symptom not reported to the automatic diagnosis system," so only symptom nodes outside the observed set are identified. EX1009, 3:52-54; EX1003 ¶503. Zhong further enforces that separation: "the selected candidate symptoms are checked for conflicts with observed symptoms and any selected candidate symptom with a conflict is removed from further consideration." EX1009, 4:15-18, FIG. 3 (step 335); EX1003 ¶504. Thus, Zhong identifies the unknown factor entries by traversing down from each condition entry while disregarding the known factor entries. EX1003 ¶505.

4. Claim 30

Claim 30 requires that the same unknown factor have different ranks depending on which condition it is traversed to. EX1003 ¶¶506-507.

Zhong stores a value on each symptom-disease edge "that represents a probability between the disease represented by the disease node given an observation of the symptom associated with the symptom node." EX1009, 8:59-67, claim 1; EX1003 ¶508. Because each edge carries a value specific to its symptom-disease pair, the same symptom node linked to two disease nodes carries different values on the two edges; a symptom like fever has a different probability for one disease than for another, and identical values would be a coincidence that is unlikely in a real-world medical knowledge graph. EX1003 ¶509. Traversing from a candidate

symptom node to a first disease node thus produces the rank stored on that edge, and traversing from the same node to a second disease node produces a different rank stored on a different edge. *Supra* §VIII.F.2.e; EX1003 ¶510.

Datla independently confirms this architecture. Datla's edges "are also weighted based on a relationship between the nodes," such as "the frequency with which the two nodes are associated within a source or a corpus," and "[t]he weight may be variable and based on the strength of the relationship." EX1005 ¶[0032]; EX1003 ¶511. Because different conditions co-occur with the same factor at different frequencies, and Datla's fever node "will be connected by edges to hundreds of other nodes," each with its own weight, the weight from a factor to one condition differs from the weight to another. EX1005 ¶¶[0031]-[0032]; EX1003 ¶¶512-513.

G. Ground 7: Claim 26 is Unpatentable Over Zhong, Datla, and Dang

Claim 26 would have been obvious over Zhong, Datla, and Dang. EX1003 ¶¶514-520. A POSITA would have combined Zhong and Datla for the reasons given for Ground 6 and added Dang's feature-node architecture because both Zhong and Dang use knowledge graphs for diagnostic inference and question generation, Dang's intermediate tier groups symptom-level evidence into clinically meaningful clusters absent from Zhong's flat graph, and Dang states that its graph "may be generated according to any known method in the art." *Supra* §VIII.F.1; EX1009,

Abstract; EX1008 ¶¶[0003], [0025], [0027]-[0031]; EX1003 ¶¶515-516. Each element is used for its known purpose with predictable results. EX1003 ¶516.

Zhong and Datla store the rank directly on the symptom-disease edge, without an intermediate node. EX1009, 8:59-67, claim 1; EX1005 ¶[0032]; EX1003 ¶518. Dang supplies the factor group: its feature nodes are intermediate nodes between the factual and disease nodes that subsume symptom nodes and carry a rank—P(f|d) on the disease-feature connection and an aggregated feature confidence. EX1008 ¶¶[0028]-[0031], [0049], FIG. 1B; EX1003 ¶¶519-520. For the reasons given for claim 13, it would have been obvious to hold each unknown factor's rank at the feature node, between the condition and factor entries. *Supra* §VIII.A.10; EX1003 ¶520.

H. Ground 8: Claims 27 and 29 are Unpatentable Over Zhong, Datla, and Van Assel

Claims 27 and 29 would have been obvious over Zhong, Datla, and Van Assel. EX1003 ¶¶522-537.

1. Motivation to Combine

A POSITA would have combined Zhong and Datla for the reasons given for Ground 6 and added Van Assel for the reasons given for Grounds 3 and 4, which apply equally to the Zhong-Datla system. *Supra* §§VIII.C, VIII.D.1, VIII.F.1; EX1003 ¶¶522-523. Van Assel performs knowledge-graph-based diagnosis using an ontology whose concepts carry identifiers and is-a relationships; linking factor

entries to ontology concepts would improve matching of findings recorded at different specificity levels; and temporal context improves accuracy because factors can become stale. EX1011 ¶¶[0113], [0115], [0117]-[0118], [0123]-[0124], [0264], [0270], [0279]; EX1003 ¶¶524-527.

2. Claim 27

Claim 27 recites the same upward ontology traversal as claim 18, applied to known factors, and the claim 18 analysis applies. *Supra* §VIII.D.4; EX1003 ¶¶528-529. Van Assel stores IRI-identified concepts related by is-a links and retrieves "[a]ll of the ancestors" of each node. EX1011 ¶¶[0118], [0124], [0334], [0349]-[0350]; EX1003 ¶530. Because a patient with a known child finding has the parent finding, applying that retrieval to each known finding infers other known findings, and Van Assel's present/absent distinction limits the inference to present factors, the logically valid direction. EX1011 ¶[0245]; EX1003 ¶531.

3. Claim 29

Claim 29 recites the same limitation as claims 6 and 20, applied to known and unknown entries. *Supra* §§VIII.C, VIII.D.5; EX1003 ¶533. Zhong stores no temporality, negation, or experiencer attributes in its symptom entries, and Datla states only that it may "filter out negated clinical concepts," but Van Assel's stored reporting date, validity duration, and permanence attributes satisfy the disjunctive claim. EX1011 ¶¶[0264], [0270], [0279]; EX1005 ¶[0054]; EX1003 ¶¶533-537.

Because each Zhong symptom node is a factor entry whether the symptom is known or unknown for a given patient, adding the attribute to each node adds it to both. EX1009, 3:19-25, 3:52-54; EX1003 ¶533.

I. Ground 9: Claim 28 is Unpatentable Over Zhong, Datla, and Kim

Claim 28 would have been obvious over Zhong, Datla, and Kim. EX1003 ¶¶538-541. A POSITA would have combined Zhong and Datla for the reasons given for Ground 6 and added Kim for the reasons given for Ground 2. *Supra* §§VIII.B.1, VIII.F.1; EX1003 ¶540. Because Zhong's graph is built and continually updated from multiple sources using Datla's techniques, and Zhong matches patient-reported symptoms to its symptom nodes, the combination faces the terminology-variation problem Kim addresses. EX1009, 3:25-29, 3:41-49; EX1005 ¶[0032]; EX1012 ¶¶[0050]-[0051], [0054]; EX1003 ¶540. Kim's synonyms are node properties that leave Zhong's edge values and scoring unchanged. EX1003 ¶540.

Claim 28 recites the same synonym limitation as claims 5 and 19, applied to known and unknown entries; because each Zhong symptom node is a factor entry whether the symptom is observed or a candidate, adding synonyms to each node adds them to both. *Supra* §§VIII.B.3, VIII.E; EX1009, 3:19-25, 3:52-54; EX1012 ¶¶[0054]-[0055], [0060]; EX1003 ¶541.

IX. CONCLUSION

For at least the foregoing reasons, claims 1-9 and 11-30 are unpatentable, and Petitioner requests that the Board institute *inter partes* review and cancel those claims.

Respectfully submitted,

Dated: October 23, 2026

By: /s/ *Angelo J. Christopher*

Angelo J. Christopher
Reg. No. 81,594
NIXON PEABODY LLP
70 West Madison St., Suite 5200
Chicago, IL 60602-4378
Tel. 312-977-4391
Fax 844-450-2568
Lead Counsel for Petitioner

CERTIFICATION OF WORD COUNT

Pursuant to 37 C.F.R. § 42.24(d), the undersigned hereby certifies that the foregoing Petition for *Inter Partes* Review is produced using a 14-point Times New Roman font and contains approximately 13,917 words, which is less than the 14,000 total words permitted by the U.S. Patent Trial and Appeal Board rules. Counsel relies on the word count of the computer program used to prepare the Petition on October 23, 2026.

Dated: October 23, 2026 By: */s/ Angelo J. Christopher*
Lead Counsel for Petitioner

CERTIFICATE OF SERVICE

The undersigned hereby certifies that a copy of the foregoing Petition for *Inter Partes* Review, together with all exhibits filed therewith, was served on October 23, 2026, by Priority Mail Express at the correspondence address of record for the '674 patent in the USPTO Patent Center:

58766 - Beyer Law Group LLP (Evidium, Inc.)
P.O. Box 51887
Palo Alto, CA 94303-1887
United States

A courtesy copy was also sent via electronic mail to Patent Owner's litigation counsel at the following addresses:

Michael J. Sacksteder
MSacksteder@Fenwick.com

David L. Hayes
DHayes@Fenwick.com

Jessica L. Benzler
jbenzler@fenwick.com

Payam Ahmadi
PAhmadi@fenwick.com

Fenwick & West LLP

By: */s/ Angelo J. Christopher*
Lead Counsel for Petitioner