ABSTRACT

A medical information navigation engine ("MINE") is capable of inferring referral activity not reported into a referral workflow system by utilizing intent-based clustering of medical information. The intent based clustering reconciles received medical data, from a variety of sources, and then clusters the data by applying one or more clustering rules. After the referrals not otherwise reported are inferred, they may be utilized to generate metrics that can be utilized to enhance patient care, and reduce costs. Metrics may be generated for both in-network and out-of-network referrals in order to distinguish differences in reporting activity.
RULES FOR SIMILARITY

DATA

RECONCILIATION ENGINE (MAPPER)

506

526

Ex.

ADVIL

MOTRIN

IBUPROFEN

IBUPROFEN

FIG. 6
RULES FOR CLUSTERING

SIMILARITY MAPPER

PRESENTATION CLUSTERS

528

530

RULES FOR DYNAMICS

Ex. Of Rules for Clustering

\[
\begin{align*}
M_1(C) \\
M_3(C) \\
M_2(C) \\
M_2(C) \\
\ldots
\end{align*}
\]

Set of associated items + mapping based on category

FIG. 7
Optimizing Patient Retention

1. PATIENT RETENTION ENHANCEMENT
   A. Physician Communications
   B. Network Referral/Consultation
   C. Physician Socialization and Education
   D. Convergence Pathway for Specialty Identification and Credentialing with Clinical Integration

2. PATIENT RETENTION MANAGEMENT
   A. Data Collection
   B. Referral Workflows
   C. Executive Dashboard Metrics
   D. Referral Feedback Loops

3. PATIENT RETENTION ENABLERS
   A. Central Logic System Wide
   B. Professional Reciprocal Rates
   C. Hospital-Based Specialist Rates
   D. Integrated Telemedicine Programs
   E. Marketing

Dashboard Focus
Physician Referral Network

- MD3 Colonoscopy
- MD4 Pathology
- MD5 Orthopedics
- MD6 PT
- PCP1
- PCP2
- MD1 Cardiology
- MD2 Cardiology
- CPMC Echocardiogram
- Stanford Cath.

Optimize patient referrals for
- Patient experience (e.g. distance, access)
- Provider care outcomes and quality
- Cost of care

Fig. 8B
Distribution of Referral Services

In Network
- E&M 25%
- Surgery 21%
- Radiology 20%
- Hospital 13%
- Diagnostic Procedure 4%

Out of Network
- E&M 22%
- Surgery 26%
- Radiology 39%
- Hospital 0%
- Diagnostic Procedure 6%
- Injectable 6%
- DME 0%
- Home Care 0%
- PT/OT 1%

Fig. 9A
Example - Referral Inferred by the Model but Not Reported in workflow system

Evidence of Colonoscopy Referral
Patient is being referred by PCP for a colo in regards to BRBPR. Is colo ok to schedule and how soon? XXX, GI Surgery Scheduler

-- Telephone Encounter. 3/23/12

Evidence of the Colonoscopy Performed
COLONSCOPY REPORT Endoscopist: XXX, MD
Location: Surgecenter of Palo Alto. Referring physician: XXX, MD

-- Ambulatory Surgery Center. Progress Note. 5/1/12

Fig. 10A
Example - Referral Inferred by the Model but Not Reported

Evidence for Procedure Referral

Can I schedule this patient for an injection at sutter or in-office procedure? ----- Message ----- From: XXX. Sent: 5/10/2012 5:02 PM To: XXX, MD

-- Telephone Encounter. 5/10/12

Evidence for Referred Procedure

-- Informed Consent Note for spinal injection, 5/11/12

Fig. 10B
Example - Epic Referral Reporting

Incompletely Represents Referral Activity

- Encounter Note from Referring Physician, XXX, MD, Date: 4/26/12.
- YYY is a lovely 73 year old female with memory difficulties who returns for follow up.
  1) Further work-ups: a) MRI brain w/o gad ASAP to evaluate the right hemiplegia
  4) PT for the gait training and OT for the right upper extremity monoplegia

Fig. 10C

Epic Referral data for this patient: 4/26/12 to 5/31/12
- Referral for MRI, Brain, Internal Location.
  Date of Referral: 4/26/12, XXX, MD, Referring physician
Example - Referral Reporting Incompletely Represents Consulting Activity

**Pain Mgmt Clinic Consult Note. Date: 5/29/12**

REF PHYS: XXX., MD  
CHIEF COMPLAINT: chronic low back pain and left leg pain...  
PLAN: ... I offered the patient to perform a series of lumbar epidurals.

---

**Epic Referral data for this patient**

Referral for **Pain Mgmt Clinic**. Internal location.  
Date of Referral: 5/1/12. XXX, MD. Referring physician

---

Fig. 10D
Clinical Narrative is Key to Understanding Cost, Quality, and Patient Experience

Patient A: Telephone encounter note. 3/23/12

Patient is being referred by PCP for a colon in regards to BRBPR. Is colon ok to schedule and how soon? XXX, GI Surgery Scheduler

Patient A: Ambulatory Surgery Center. Progress Note. 5/1/12

COLONOSCOPY REPORT
Endoscopist: XXX, MD
Location: Surgecenter of Palo Alto. Referring physician: XXX, MD

Patient B: Encounter Note from Referring Physician, XXX, MD.
Date: 4/26/12

Plans:
1) Further work-ups: MRI brain w/o gad ASAP to evaluate the right hemiplegia
...
4) PT for the gait training and OT for the right upper extremity monoplegia

Patient B: Epic Referral data for this patient: 4/26/12 to 5/31/12

Referral for MRI, Brain, Internal location.
Date of Referral: 4/26/12. XXX, MD, Referring physician

Fig. 10E
ATTRIBUTE SELECTION (AUTOMATED AND/OR MANUAL)

DEFINE CRITERIA FOR CLUSTERING AND RELEVANT COMBINATIONS (AUTOMATED AND/OR MANUAL)

COMPUTE DISTANCE BETWEEN INCOMING DATA ATTRIBUTES AND CLUSTER. ARE ALL DISTANCES < MAXIMUM DISTANCE THRESHOLD?

REJECTION FROM PRESENTATION CLUSTER

INCLUDE IN PRESENTATION CLUSTER

STOP

FIG. 11
SYSTEMS AND METHODS FOR PATIENT RETENTION IN NETWORK THROUGH REFERRAL ANALYTICS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This non-provisional application claims the benefit of provisional application No. 61/754,527 filed on Jan. 18, 2013, entitled “Systems and Methods for Patient Retention in Network Through Referral Analytics”, including Appendix A, which application and appendix are incorporated herein in their entirety by this reference.

[0002] This application is also a continuation-in-part application which claims the benefit of application Ser. No. 13/223,228 filed on Aug. 31, 2011, entitled “Medical Information Navigation Engine (MINE) System”, which application claims priority to U.S. Provisional Application No. 61/379,228 filed on Sep. 1, 2010, of the same title, both applications are incorporated herein in their entirety by this reference.

BACKGROUND

[0003] The present invention relates generally to a medical information engine, and particularly to management and consolidation of medical information which enables patient retention in a medical network by leveraging analysis of referral practices.

[0004] Despite rapid growth of innovation in other fields in recent decades, the world of medical information, including patient medical records, billing, referrals, and a host of other information, has enjoyed little to no useful consolidation, reliability, or ease-of-access, leaving medical professionals, hospitals, clinics, and even insurance companies with many issues, such as unreliability of medical information, uncertainty of diagnosis, lack of standard, and a slew of other related problems.

[0005] One of the challenges facing those in the medical or related areas is the number of sources of information, the great amount of information from each source, and consolidation of such information in a manner that renders it meaningful and useful to those in the field in addition to patients. Obviously, this has contributed to increased medical costs and is perhaps largely attributed to the field suffering from an organized solution to better aid the medical professionals, to better aid those requiring more reliable patient history and those requiring more control and access over such information.

[0006] Currently, when a patient sees various medical professionals over the years, there is no method for universally tracking recommendations, thoughts, prescriptions, referrals, or diagnosis. This hinders the job of insurance companies in making certain requisite determinations, physicians making decisions that directly impact the health of the patient, and hospitals and other medical institutions that similarly rely but do not have the benefit of the requisite information, not to mention the patient.

[0007] For consolidated provider networks it may be possible to centralize a patient’s records. However, these records are typically not available outside of the network, and moreover, often the records generated outside of network are often isolated. Patient referrals, thus can cause increased costs for an insurer, and may result in fragmented patient information. In order to reduce operating costs, there is significant pressure to optimize patient retention within a medical network. In order to effectively implement mechanisms for this retention, accurate referral data is required.

[0008] It is therefore apparent that an urgent need exists for a medical information navigation engine (“MINE”) capable of managing medical information in a manner that is beneficial, reliable, portable, flexible, and efficiently usable to those in the medical field, including patients. A functionality of such a system would be the ability to track patient referrals, and generate models which assist insurers and care providers to track referral practices. Such a MINE should also be capable of reconciling and intent-based clustering of patient data by applying at least one clustering rule to the reconciled medication information, and presenting the clustered reconciled medical information to a user.

SUMMARY

[0009] To achieve the foregoing and in accordance with the present invention, systems and methods for managing medical information are provided. In particular, systems and methods for referral analysis using intent-based clustering of medical information is provided.

[0010] In one embodiment, a medical information navigation engine (“MINE”) is capable of analyzing referral data by inferring referral activity not reported into a referral workflow system utilizing intent-based clustering of medical information. The intent based clustering reconciles received medical data, from a variety of sources, and then clusters the data by applying one or more clustering rules. These rules may include similarity rules, in some embodiments, where distance between a patient attribute and the reconciled is computed. If the distance is below a threshold then the referral can be modeled for.

[0011] After the referrals not otherwise reported are inferred, they may be utilized to generate metrics that can be utilized to enhance patient care, and reduce costs. The metrics can be generated by adding the inferred referral activity to reported referral activity to generate a combined referral number, and dividing the reported referral activity by the combined referral number. This calculation may be performed for both in network and out of network referrals in order to distinguish differences in reporting activity. The clustering can also be utilized to identify the type of referral (i.e. radiology) in order to analyze trends, and develop patient retention practices.

[0012] Note that the various features of the present invention described above may be practiced alone or in combination. These and other features of the present invention will be described in more detail below in the detailed description of the invention and in conjunction with the following figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] In order that the present invention may be more clearly ascertained, some embodiments will now be described, by way of example, with reference to the accompanying drawings, in which:

[0014] FIG. 1 shows a medical system 100, in accordance with an embodiment of the invention;

[0015] FIG. 2 shows further details of the MINE 112 of FIG. 1, in accordance with an embodiment of the invention;

[0016] FIG. 3 shows an exemplary embodiment implementing the system 100 using various devices;
[0017] FIG. 4 shows further details of the system 100, in accordance with an embodiment of the invention;
[0018] FIG. 5 shows further details of the engine 502 and the block 504 of FIG. 4;
[0019] FIGS. 6 and 7 each show examples of applying the rules 526, and 528 and 530, to the data 506 to yield certain beneficial results;
[0020] FIGS. 8A-8C show exemplary diagram for implementing patient retention in a medical network;
[0021] FIGS. 9A and 9B show example diagrams for analytics generated by the MINE 112 for patient referrals;
[0022] FIGS. 10A-10E show example evidence analyzed by the MINE 112 to generate referral analytics; and
[0023] FIG. 11 shows a flow chart of the steps performed by the block 504 of FIG. 5 in applying the rules therein, in accordance with an exemplary method of applying intent-based clustering and display to the data 506 of FIG. 5.

DETAILED DESCRIPTION

[0024] The present invention will now be described in detail with reference to several embodiments thereof as illustrated in the accompanying drawings. In the following description, numerous specific details are set forth in order to provide a thorough understanding of embodiments of the present invention. It will be apparent, however, to one skilled in the art, that embodiments may be practiced without some or all of these specific details. In other instances, well known process steps and/or structures have not been described in detail in order to not unnecessarily obscure the present invention. The features and advantages of embodiments may be better understood with reference to the drawings and discussions that follow.

[0025] 1. System Architecture

[0026] Referring now to FIG. 1, medical system 100 is shown, in accordance with some embodiments. The system 100 is shown to include medical source 114, a medical information navigation engine (MINE) 112, and medical information consumers (also referred to herein as “output” or “medical output”) 117. The medical source 114 is shown to include one or more electronic health records (EHR) 118 and 120, health information exchange (HIE) 122, and a picture archiving and communication system (PACS) 124. The MINE 112 is shown to include interface 113, a back-end medical processor 116, and a front-end medical processor 115.

[0027] “Medical information”, as used herein, refers to any health-related information, including but not limited to patient medical records, patient entered information, care team entered information, healthcare device generated information, referral information, and billing information.

[0028] The source 114 generally provides various medical information to the MINE 112. For example, the EHRs 118 and 120 each may provide information such as medical records and billing, the HIE 122 may provide information such as medical records, and the PACS 124 may provide information such as diagnostic imaging and reports.

[0029] The medical information consumers 117, which may be made of a host of entities or individuals, such as patients, clinics, medical institutions, health organization, and any other medical-related party, use information that is provided by the processor 115 of MINE 112 and that can, by way of example, consist of patients, medical systems, medical organization administrators, medical researchers, and/or EHR users. For example, user-customized processed medical information is provided by the processor 115 to a number of users within the medical information consumers 117. In this case, the processor 115 generates user-customized processed medical information to a plurality of users, with at least a portion of the user-custome processed medical information being provided to each of the users based on the relevancy of the portion being provided of each user’s specific function or role and each user’s associated security privileges.

[0030] The processor 116, in some embodiments, indexes, maps, and consolidates medical information, received from the interface 113, and tags this information, and determines to reconcile the tagged information. In some methods and embodiments, information that is extracted from images is tagged to enhance recall of search queries. Indexing, at least in part, processes document and converts them into formats that allows for quick searching across a large collection of documents.

[0031] The information in the MINE 112 is encrypted and secure to ensure privacy of sensitive medical information.

[0032] It is understood that the sources 114 of FIG. 1 includes merely some examples of the sources that communicate with the MINE 112 and that other sources, known to those in the field, are contemplated. Similarly, the output 117 may be used by those or entities not discussed herein but that are contemplated and within the scope and spirit of the invention.

[0033] The interface 113 serves to receive information that is in various forms, such as but not limited to text, html, CCD, CCR, HL7 and any other type or formatted information. The interface 113 then provides to the processors 115 and 116 information, as needed.

[0034] The processor 116 receives some of the medical information that the interface 113 processes and performs certain tasks to process it, such as indexing, semantic meta-tagging, and reconciliation. Indexing takes processed documents and converts them into formats that make it easy to quickly search across a large collection of documents. Semantic meta-tagging embeds information into the medical information that is relevant thereunto and that can later be used to search for certain information for the purpose of reconciliation and search, among many others.

[0035] One aspect of consolidation, reconciliation and deduplication, generally refers to removing of redundant patient medical records, such as, multiple records for the same individual appearing as though the records are for different individuals or multiple data elements that are recorded similarly but slightly differetly in the different sources. In this case, the processor 116 recognizes that the records belong to a single individual or are the same data and just recorded differently and automatically consolidates them. The patient or user of the system 100 may also manually perform reconciliation. The processor 116 advantageously determines whether or not reconciliation is performed.

[0036] The processor 116 outputs the indexed, tagged and reconciled information to the processor 115. The foregoing tasks are a generalization and further details of each are provided below.

[0037] The processor 115 performs certain tasks on the information provided by the interface 113 and the processor 116, which include query, search, presentation, and quality checking. The output of the processor 115 is the output of the MINE 112, or output 117.
The MINE 112, through the processor 115, in some embodiments and methods, invites members of a medical care team to join it thereby allowing distributed user-organized care teams.

Querying, as performed by the processor 115, is the ability to receive, as input, a free text query, from a user, (i.e., a query without any restrictions on the structure)—and converting the free text query into commands to a medical search engine, such as Medical Lexical Search Engine and the MATRIX (Medical Application Terminology Relationship IndexX) Concept Search Engine, using a sophisticated query processing engine optimized to work with medical queries. The results of the search engine are sent to the presentation display planner—which decides the most relevant presentation given the user’s organization and role (e.g. the provider, search query program, a healthcare administrator, a study administrator, and the patient). The presentation discussed below, receives such information. In some embodiments and methods, the medical information or user information is processed to suggest relevant queries.

Search, as performed by the processor 115, is built around the concept of Zero-Click Relevance—on the ability to get to all the relevant information in an actor in the healthcare system requires by typing in just a single query. The search engine, within the processor 115, performing the search comprises an indexing and searching, as will become apparent shortly. Optionally, search results may be securely embedded into third party programs. In some embodiments, searching involves determining presenting (also referred to herein as “providing”) access to specific relevant data based on a search query, the patient, and the user’s specific function and/or role and security privileges. A user may be within the output 117 and security privileges are either determined by the MINE 112 or the provider or both. The information that is uploaded to the MINE 112 by users, such as in output 114 (in some embodiments) is searched by the processor 115. The uploaded information may include information such as but not limited to status posts, records, and images. Such user-uploaded information is routed automatically to the output 117, as needed.

Some aspects of the search are now discussed relevant to an example. Assuming, by way of example, that Dr. Smith, an internal medicine physician, sees a new patient, Joan Sample, who presents with a complaint of chest pain. Joan has brought several continuity-of-care documents (CCDs) and a 600-page pdf file representing of her medical chart. She has seen a cardiologist who uses NextGen’s electronic medical record (EMR) and a gastroenterologist who uses EMR’s EMR and she has recently visited a local emergency room. Dr. Smith uses the search of the various methods and embodiments of the invention to efficiently assemble the relevant information he needs. Dr. Smith selects Joan Sample as the patient and enters the clinical context “chest pain” in the search bar of a screen presented by the MINE 112 (examples of such screens are shown in subsequent figures herein). He is presented with relevant lab results, such as CKMB, troponin, and amylose; relevant diagnostic results, such as prior electrocardiograms (EKGs) and the most recent chest computed tomography (CT) scan; and all progress notes and consult reports in which concepts relevant to chest pain, like “GERD” and “cardiac stress test”, are mentioned. Two distinct types of searches are combined, in accordance with a method and embodiment of the invention, to retrieve information medically relevant to Joan’s complaint: 1) Lexical search, where text in the patient record is searched for occurrences of the search term, its variants and synonyms; and 2) Medical concept search, where data that is medically related to the search term is retrieved. Medical concept search finds relevant structured data with standardized codes, such as lab results, and text results, such as progress notes, which include terms medically related to the search term. In Joan’s case, a search for “chest pain” returns a CKMB lab result and a reference to the most recent chest CT scan. Accordingly and advantageously, the Lexical and Medical concept search solves Dr. Smith’s information overload problem by returning information in a chart most relevant to determining the etiology of Joan’s chest pain complaint. Further, in some embodiments, the presentation, discussed shortly, presents a united view of Joan’s history by reconciling and de-duplicating data from multiple sources that may be coded and described differently. Redundant data is automatically reconciled even if it is described differently by different sources.

Presentation, as performed by the processor 115, is displaying health information to the requesting user in a way that reduces the number of clicks and maximizes the amount of meaningful information delivered based on the interpreting the intent of the user query.

Quality checking, as performed by the processor 115, is checking the quality of medical information provided by various sources, i.e. source 114, by the patients, structured data, and unstructured data, in a wiki-like manner setting whereby the users can help maintain and improve the quality of information displayed. The foregoing tasks, performed by the processor 115, are further described in detail below. Additionally, the users or patients may make comments regarding medical information, in a wiki-like manner.

In summary, the MINE 112 transacts medical information including the interface 113 receiving medical information from a number of medical sources (such as within the source 114) for processing, identifying, mapping, and consolidating by the medical processor 116, providing access to specific relevant data, based on a user’s security privileges, within the identified, mapped, and consolidated medical information, based on user-specific functions or roles, performed by the processor 115, and generating user-customized processed medical information to a number of users, such as within the output 117, with at least a portion of the user-customized processed medical information being provided to each of the users based on its relevancy to each user’s specific function or role and each user’s associated security privileges.

FIG. 2 shows further details of some embodiments of the system 100, particularly the MINE 112 thereof. That is, the processor 116 is shown to include an indexing and metal tagging module 234, which includes an indexing module and a meta tagging module (both of which are not shown in FIG. 2 in the interest of clarity), which may be a module, as shown in FIG. 2 or two physically separate modules. The processor 116 is further shown to include a reconciliation and de-duplication module 236, which also can be broken out into two modules, a reconciliation module and a de-duplication module, and a code and semantic mapping module 238, which also may be a single module or multiple modules. The modules 234, 236, and 238 communicate with one another.
to the servers 232, including but not limited to PCs or servers. The display 340 is used to construct presentation and display information to users, such as the patient’s records, billing information, and other types of medical information. The display 340, in some embodiments, also performs processing of some of the functions of the processor 115.

[0047] The foregoing modules may be software programs, executed by a computer or computing engine of suitable sorts, or may be implemented in hardware.

[0048] FIG. 3 shows an exemplary embodiment implementing the system 100 using various devices. That is, the medical system 330 is analogous to the system 100 and is shown to include the sources 114 coupled to communicate, securely, through the secure communication link 342, to the interface 113. The link 342 may be any suitable communication channel allowing information, of various formats and types, to be transferred to the interface 113 in a secure and encrypted fashion. Exemplary communication channels of which the link 342 is made include the Internet, VPN connections over the Internet, private dedicated digital lines such as T1, T3, E1, and SONET, and other fiber optic formats.

[0049] The interface 113, in some embodiments, is a software program that executes on one or more servers 232, which can be a server of any kind of suitable computing engine, such as personal computer (PC). The servers 232 receive secure information through the link 342 from the sources 114. The processor 116, in some embodiments, includes the module 236 and one or more servers 234, which may be any suitable computing engine, similar to the servers 232, including but not limited to PCs or servers.

[0050] The module 236 and servers 234 perform the tasks discussed above relative to the processor 116 and the display 340 and servers 238 perform the tasks discussed above relative to the processor 115 though these processors may and often perform additional tasks related to medical information, some examples of which are presented and discussed below and the rest of which are contemplated and achieve the various advantages, results and functions presented herein.

[0051] The processor 115, in some embodiments, includes display and visualization 340 executing on one or more servers 238, which may be any suitable computing engine, similar to the servers 232, including but not limited to PCs or servers. The display 340 is used to construct presentation and display information to users, such as the patient’s records, billing information, and other types of medical information. The display 340, in some embodiments, also performs processing of some of the functions of the processor 115.

[0052] As shown in FIG. 3, the servers 232 are coupled to the module 236 and the servers 234, and to the display 340 and the servers 238 and the module 236 and servers 234 are coupled to the display 340 and the servers 238.

[0053] In some embodiments, the interface 113, servers 232, module 236, servers 234, display 340, and servers 238 are remotely located relative to the sources 114 and in some embodiments, remotely located relative to one another. Further, they are considered a part of the Internet cloud where, performing their tasks in a manner known as “cloud-computing”. However, other manner of achieving the functions and advantages of the invention, including various other of implementation, not shown in FIG. 3 or other figures herein and/or not discussed are contemplated.

[0054] FIG. 4 shows further details of the module 236 of FIG. 2, in accordance with an embodiment of the invention. The module 236 is shown to include a reconciliation engine (also referred to hereinafter as the “mapper”) 502 responsive to data 506, which is, at least in part, within the source 114, and is shown to provide reconciled information that is provided to the intent-based presentation block 504.

[0055] The engine 502 advantageously learns, through history, ontology, user-input, the type of user, and a host of other factors, similarities between various information from the data 506, defines characteristics thereof, models this information conceptually, pre-selects and sorts information before providing it the block 504 for presentation in the format of a display, or other known types of presentations. Such processing entails the use of various sets of rules, at various stages, as will be evident shortly relative to subsequent figures and discussions.

[0056] Presentation by the block 504 is intent-based, that is, the user of the module 236 along with history, and other factors are used to determine the information to be presented. With time, as the engine 502’s knowledge of medical information, such as drugs, type of users, diagnosis, the relationship between various diagnosis/diseases relative to each other and relative to various medications, and other information, increases, the information presented by 504 becomes increasingly intent-based.

[0057] The engine 502 is shown to include a conceptual model block 508, which conceptually models the data 506, such as to determine similarities, an example of which is provided and discussed in subsequent figures.

[0058] FIG. 5 shows further details of the engine 502 and the block 504 of FIG. 4. The engine 502 is shown to include a reconciler block 510 that receives data 506 and a similarity mapper 512, which generally performs the tasks of the block 508 in FIG. 1. The block 504 is shown to include a presentation cluster block 514, which is shown to receive information from the mapper 512, and a data cluster 516.

[0059] A set of similarity rules 526, which identify similarities of various types of information, and define characteristics thereof, is shown being utilized by the reconciler 510. The rules 526 are applied to the data 506 to identify similar concepts, which unlike prior art techniques, is not to look for matches and rather to correlate information based on concepts. Through feedback from users 518, this becomes a learned process with improved and more sophisticated conceptual similarity detection. The similarity mapper 512 maps the reconciled information, generated by the reconciler 510.

[0060] Another set of rules, namely, a set of clustering rules 528, is provided to the presentation cluster block 514 for determining which information, if any, to cluster or group. The block 514 also receives as input, user intent query 540, from a user, and applies the rules 528 to the latter. The rules 528 are used by the block 514 to group information received from the mapper 512, based on the user intent query 540, and in the process additional apply a set of dynamics (time) rules 530 there to. The rules 530 serve to identify what is to be looked at to find what information has been changed over time. In this respect, feedback from the user, through 542, is utilized. Similarly, the rules 528 utilize feedback from the user. Additionally, feedback from the user is utilized, at 534, to accumulate concept-based information and definitions in a Wiki-style fashion.

[0061] The presentation cluster block 514 generates output data clusters 516. The cluster 516 information may be displayed and/or presented in other manners, such as with an Application Programming Interface (API), and it further may
receive user feedback and use the same to further refine rules for clustering and similarity mappings.

[0062] The rules 526, 528, and 530 are independent of one another in some embodiments of the invention. In other embodiments, information flows between. Advantageously, these rules, partly because they are applied at different stages in the processing of the data 506, allow for a learned and conceptualized process as opposed to a hard decision. For example, in current techniques, where only one set of rules are utilized early on in the processing of the data, a hard decision is made with no flexibility to alter this decision thereby increasing the risk of mis-categorization and/or identification of relevant information. In contrast, thereto, the different sets of rules of the embodiment of FIG. 5, breakdown categories, such as similarity, display, and history, allows configuration of various aspects thereof.

[0063] By way of example, in prior art techniques, where the data is regarding electronic devices and a cell phone is to be identified, where the single set of rules, made early on in the process, is based on the lack of a keyboard, and a central processing unit, the device may be erroneously identified as an electronic tablet, with no recourse. Whereas, the embodiment of FIG. 5 allows for progressive learning of various attributes of the device by, for example, using the above exemplary rules as the rules 526 but based on the rules 530 and 528, introducing attributes, such as size of the device, that allow for a more accurate identification of the device. And further, due to the user feedback and query, allow for dynamically altering the rules.

[0064] Use of various rules, such as rules 526, 528, and 530, at various stages of processing, allows flexibility in applying the rules to achieve greater accuracy of clustering. In medical applications in particular, information is oftentimes duplicated for various reasons, such as lack of standardization of names of medications, shorthand identification of information, and a slew of other reasons. In this regard, flexibility of applying rules is vital. While three sets of rules are shown in the figures and discussed herein relative to various embodiments, it is understand that a different number of rules may be employed.

[0065] 11. Medical Information Management

[0066] For a better understanding of the flexibility the rules of FIGS. 5-7 offer, an example is now presented. Suppose the data 506 carries medical information for which a particular condition, e.g. diabetes, is to be detected. Rule 526 allows for a similarity between lab results and “diabetes” to be identified but that is nearly where the application of rule 526 ends until further information is known and extracted later in the processing of the data 506. Namely, when rule 528 is applied to the outcome identified by Rule 526, the lab results are crawled or inspected for “diabetes” or another identifier for “diabetes”. Additionally, the presence of various relevant labs is detected and the association between the presence of the labs and the problem of diabetes and perhaps, hemoglobin A1c (a measure of average blood glucose concentration over the past 30 to 120 days, used in the diagnosis and treatment of diabetes) is made. Next, the rule 530 is applied to the outcome of the application of rule 528 where patient data is used or a correlation between a problem and a treatment for a large percent of the patient population is made. Specifically, the percentage of patients with diabetes is detected. The parameter of time allows for the latter detection, otherwise, for example, at the application of rule 526 or even rule 528, a large patient base could not have been correlated.

[0067] The user input at 540 and the user feedback at 518 all help in the clustering of data. At the application of rule 526, a determination is made as to how things are similar until a user asks about the similarity after which a more educated application of rules is performed. Thus, no decision is made until the end or at the output of the block 514, in real-time.

[0068] During the application of rule 526, the system is informed of key parameters but not how to put the data together. Then, during the application of the rule 528, the system is informed of how to put the data together (cluster) by aligning the data in a particular manner. Application of rule 530 determines how things change over time, or not, but not what the correlation or similarity actually is, which is generally done through the rule 528. Part of the reason for splitting the rules is to delay decision-making as long as possible in an effort to cleverly use more information, such as that provided by the user, for an increasingly accurate finding.

[0069] The outcome of the data cluster 516 can be transmitted to another processor, system, user, or any other entity.

[0070] Another example of the manner rules are employed, outside of the medical community, is for example, in determining parameters of hair where rule 526 is used to look for length of hair and rule 528 uses the outcome of the length of hair to further determine alopecia as compared with normal hair growth. Rule 530 may then be used to determine a percentage of a demographic that has experienced baldness. Further examples of the application of these rules is shown and discussed relative to FIGS. 6-11.

[0071] As with the blocks of the MINI 112 of FIG. 1, it is understood that the blocks shown in FIG. 5, such as block 510 and 514, and 516 may be independently a machine and/or processor or a part of a machine and/or processor. They may alternatively, be carried out in software programs.

[0072] FIGS. 6 and 7 each show examples of applying the rules 526, and 528 and 530, to the data 506 to yield certain beneficial results. In FIG. 6, rule 526 is applied to the data 506 to identify the medication named “Advil” as an “ibuprofen”. Similarly, “Motrin” is identified as Ibuprofen, therefore, allowing more flexibility to a patient and a medical professional in deciding to use these drugs. Rules for similarity specific what characteristics need to be looked at to determine similarity of an object to another object.

[0073] Using the same example, rules 528 and 530 may be applied to the outcome of the rule 526 to the data 506 to determine other information based on the intent of the user. For example, the dosage of Ibuprofen, from all sources, even those with other ingredients, may be determined by applying rule 528, after applying rule 526 such that the outcome of rule 526 detects Ibuprofen type of medications and rule 528 narrows the detection to those with a threshold dosage.

[0074] FIG. 7 shows an application of the rules 528 and 530 where a set of associated terms, m1(c), has been identified and another set of associated items, m2(c), has been identified. For example m1(c) is one medication, m2(c) is another medication and “c” are particular characteristics of each medication such as, but not limited to, brand name, generic name, dosage, prescription instructions (sig), prescription date, and start date. Rules for dynamics, rule 526, is the time base characteristics. Rules for clustering, rules 528, would be probability of matches of other characteristics. For example, for a given medication such as oral contraceptive pills (OCPs), the rules of dosage might be ignored such that different prescriptions with different doses would be considered the same for clustering purposes. In this example, the use of
oral contraceptive medications at all dosages is contraindi-
cated for women with a genetic predisposition or other risk
factors associated with thrombotic events (e.g., venous
thromboembolism). Another medication where dosage might
be very important to outcomes would not be clustered
together if the dosage were different. For example, Warfarin,
a medication commonly used to prevent blood clotting, has
a very narrow therapeutic window and its metabolism widely
varies among individuals. The dosage of Warfarin is highly
correlated to outcomes of interest. Physicians routinely pre-
scribe different dosages of Warfarin to treat or prevent throm-
botic events or predisposing conditions such as pulmonary
embolism or atrial fibrillation.

[0075] For further clarification, in the example of Advil and
Ibuprofen, if the intent is to investigate whether the current
dosage of Ibuprofen is too high, all sources of Ibuprofen (even
those with other ingredients) are better to be identified, in
which case the set m1(c) may represent all sources. The date
may also be an indicator, such as the last day or week this
medication was prescribed or consumed and may accordingly
be a part of the set, m2(c). In contrast, if contributors to (or
indicators of) a chronic condition is the intent, a longer
history (months, years) and the chronic condition itself would
be “related” despite low “similarity”. Thus, through the
flexibility of the application of various rules, such as rules 526,
528, and 530, there can be different ways of displaying infor-
mation, from the data clusters 516, about a concept under
different intents.

[0076] By way of further explanation, the rules 528 are
used to determine what is considered inside the cluster and the
rules 530 is how things in the cluster change over time.
Though, in other embodiments, other types of rules are antici-
pated and other numbers of rules are anticipated. Using the
Advil/Ibuprofen example above, the rules 528 are used to
determine whether other medicines belong in the display
cluster, for instance if they contain Ibuprofen but they also
contain other things (such as sleep aid, Contrem) that may or
can not “belong” in the display cluster. The embodiment of
FIG. 5 advantageously learns whether they “belong” or not.
Through the rule 530, the Ibuprofen cluster might emphasize
recent events (post week) in the ranking. Other clusters may
interact differently with time.

[0077] FIG. 11 shows a flow chart of the steps performed by
the block 514 of FIG. 5 in applying the rules therein, in
accordance with an exemplary method of applying intent-
based clustering and display to the data 506 of FIG. 5.

[0078] In FIG. 11, at step 1151, an automatic or manual,
or a combination, of attribute selection is performed by applying
rules 526 and 528, in accordance with a method and embed-
diment of the invention. Accordingly, attributes to be included
in clustering are selected. For example, for a presentation of
medications, several attributes might be included. To present
results to a user, another machine, processor or the like, for a
medication history intent, medication brand name, generic
name, coding system, drug code, and ingredient might be
selected. This may be done by a user, manually, or automati-
cally by the block 514, in an exemplary embodiment of the
invention.

[0079] Next, at step 1153, the criteria for clustering relevant
combinations is defined, manually, automatically or using a
combination thereof. The matching criteria for each of these
attributes are defined as a maximum distance along with an
appropriate definition of distance. Examples of “maximum
distance” are “exact match”, “close match”, “loose
match” or distance-\(x\) where \(x\) is an appropriate maximum
distance. Examples of distance measures are numerical dif-
ference, any edit distance, semantic distance, abstract dis-
tance along an ontology or graph, etc. Rules for relevant
combinations of the selected attributes are also defined during
this step. In some methods and embodiments, attributes can
be combined to create a composite threshold that the data 506
can be measured against. With reference to the medication
history intent presented hereinabove, all medications in the
patient history that are close matches on brand name or
generic name might be included, along with exact semantic
matches on a particular drug ingredient or exact numerical
matches on a group of related drug codes.

[0080] Next, at 1155, the distance between the data 506’s
attributes and cluster is computed. For example, it is deter-
mined whether all distances are less than the maximum dis-
tance threshold and if so, the cluster is updated to include such
data in the presentation, at step 1159, otherwise, step 1157,
such data is rejected (not included) in the cluster.

[0081] III. Patient Retention in Network

[0082] Now that the mechanisms of the MINE 112 have
been described in detail, attention should be turned to FIGS.
8A-8C, which disclose means for patient retention through
referral analytics and management. The MINE 112 provides
the ability to track patient referrals in a manner which may
guide practitioners to refer patients in network, and can pro-
vide tools for targeting physician education and training on
referral practices.

[0083] Medical foundations refer substantial patient care
outside the Medical Network where equal or superior care is
available within the network. These referrals are often unco-
ordinated and not uniformly reported using existing stacking
systems. For most medical providers, it is not clear where
these patients are referred or for what reason(s) and the
related costs. To ensure affordable, quality care and positive
patient experience better referral analytics are desired.

[0084] FIG. 8A illustrates that the optimization of patient
retention in network relies upon three core areas: 1) patient
retention enhancement, 2) patient retention management, and
3) enablers of patient retention. The first category of retention
optimization relies upon medical network procedures and
education in order to increase patient retention. This is well
known, but in order to effectively implement these actions,
retention analytics, including referral analytics, are desirable.
For example, a procedure for placing quality standards for
referrals outside of network may not be effective if the prac-
tice area rarely engages in referral practices. In a similar
manner, patient retention enablers, such as marketing, central
logic systems, and set payment rates enable the retention of
patients through affordability and quality of care.

[0085] As previously noted, in order to drive enablers to
enhance patient retention, patient retention management sys-
tems are employed. These systems collect data, generate
referral analytics, and otherwise compile and analyze relevant
data. The MINE 112 is a particularly powerful tool
able of performing many of these activities.

[0086] Moving to FIG. 8B, a physician referral network is
illustrated, for context. The goal of any patient referral opti-
mization is to ensure the best patient experience, outcomes
and cost. This illustration provides a representative example
of what a typical physician referral network may look like,
and relative costs associated with the various providers.

[0087] FIG. 8C provides an example diagram for a com-
plete referral lifecycle. In this example, a patient has an initial
encounter which yields a visit note, a referral and a claim. The referral may be for a specialist or an internal referred encounter. A procedure note or consult note are generated from this interaction. A return encounter is then made, with the generation of another visit note. These records of the encounter may be utilized by the MINE 112 to determine referral activity in ways currently not possible without manual review of each note.

The MINE 112 is able to analyze all referral data to establish an analytical dashboard that provides benchmarking, outcome measures, and enables knowledge driven care. Existing referral workflow systems (i.e. Epic, etc.) only have the ability to report, track and manage referrals if entered. Currently, most referral data is undefined, untrackable and unmanageable, thus making referral workflow systems inefficient. The MINE 112 is capable of picking out the referral data in a manageable form for consumption by the referral workflow systems. With more robust referral data, these referral workflow systems will be able to assist in the development of strategies that optimize access, continuity of care, resource management, affordability, quality, and patient/physician satisfaction.

The disclosed system may then utilize the referral data to generate one or more metrics of compliance by physicians to preferred referral reporting. An example of such a metric is defined by Formula A:

\[
\text{Formula A} = \frac{\text{Number of in network referrals reported}}{\text{Total number of in network referrals}}
\]

Further, a second metric can be computed which reveals opportunities for improvement of out of network reporting. An example of such a metric is defined by Formula B:

\[
\text{Formula B} = \frac{\text{Number of out of network referrals reported}}{\text{Total number of out of network referrals}}
\]

The denominator values for these metrics necessitate the usage of the MINE 112 to scour patient records in order to capture all incidences where a referral is given. In order to ensure complete capture of referrals, the MINE 112 data was compared to auditing data. The results of these comparisons are provided in Table 1.

### Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Referral entries in workflow system</th>
<th>Inferred referrals not already captured</th>
<th>Total</th>
<th>Formula B results</th>
</tr>
</thead>
<tbody>
<tr>
<td>MINE Analysis</td>
<td>256</td>
<td>173</td>
<td>429</td>
<td>60%</td>
</tr>
<tr>
<td>Internal Audit</td>
<td>218</td>
<td>131</td>
<td>349</td>
<td>62%</td>
</tr>
</tbody>
</table>

As can be seen, given similar samples, the results of the automated MINE 112 modeling of referrals is statistically similar to the results of a costly and time consuming audit. Also striking, is that the referral reporting practices differ significantly when it is an in network, HMO style; practice, as compared to an out of network, PPO style arrangement. Table 2 provides the referral reporting rates for these two scenarios:

### Table 2

<table>
<thead>
<tr>
<th>Method</th>
<th>Referral entries in workflow system</th>
<th>Inferred referrals not already captured</th>
<th>Total</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>In network</td>
<td>29,456</td>
<td>9,542</td>
<td>38,998</td>
<td>76%</td>
</tr>
<tr>
<td>Out of network</td>
<td>13,365</td>
<td>9,479</td>
<td>22,844</td>
<td>59%</td>
</tr>
</tbody>
</table>

The MINE 112 is also able to utilize its analytical capabilities to pinpoint the distribution of referral services. FIG. 9A provides breakdowns of referral types. Interestingly, the referral distributions have significant differences between in network and out of network referrals. For example, radiology and surgery are referred more frequently when out of network, and hospital and physical therapy is more often referred in network.

FIG. 9B provides a graph illustrating where referrals were located when not in the referral workflow system. The overwhelming majority of these referrals were sent out of network, which on a whole is more costly, and does not deliver a higher level of care. For this very reason, the capture of un-reported referral data is critical to the control over health care costs, and the optimization of patient retention through better outcomes at lower costs.

Lastly, while the architecture of the MINE 112, and the analytics performed to capture referral data that is otherwise not reported has been disclosed in considerable detail, specific examples of referral modeling are helpful for further clarification. FIGS. 10A-10E provide such examples.

For instance, in FIGS. 10A and 10B, physician notes can be analyzed for a specific procedure referral, even when the referral has not been reported into a referral workflow system. In contrast, FIGS. 10C and 10D provide an instance where referral activity is being reported in the workflow system (MRI referral) but the physician’s notes provide a more complete story (referral is also for occupational therapy).

FIG. 10E provides an example of which data is deemed relevant for some embodiment of the MINE 112 as part of the analysis. This includes cost data, quality data, and patient experience data.

While this invention has been described in terms of several embodiments, there are alterations, modifications, permutations, and substitute equivalents, which fall within the scope of this invention. Hence, it should also be noted that there are many alternative ways of implementing the methods and apparatuses of the present invention. It is therefore intended that the following appended claims be interpreted as including all such alterations, modifications, permutations, and substitute equivalents as fall within the true spirit and scope of the present invention.

What is claimed is:

1. In a medical information navigation engine ("MINE"), a method for referral analysis, the method comprising:
   - inferring referral activity not reported into a referral workflow system utilizing intent-based clustering of medical information;
generating reporting metrics from the inferred referral activity; and
displaying the metrics on a dashboard.

2. The method of claim 1 wherein the intent-based clustering of medical information further comprises:
   receiving medical information from a plurality of medical sources;
   reconciling the medical information; and
   clustering the reconciled medical information, wherein the clustering includes applying at least one clustering rule to the reconciled medical information.

3. The method of claim 2 wherein the reconciling further comprises applying at least one similarity rule.

4. The method of claim 3 wherein the at least one similarity rule includes comparing patient data attributes.

5. The method of claim 4 further comprising computing a distance between the patient data attributes and clustered reconciled medical information.

6. The method of claim 5 wherein if the distance is less than a threshold, then the clustered reconciled medical information is included as a referral.

7. The method of claim 1 wherein the metric includes adding the inferred referral activity to reported referral activity to generate a combined referral number, and dividing the reported referral activity by the combined referral number.

8. The method of claim 7 wherein the metric is performed for in network referrals and out of network referrals.

9. A medical information navigation engine ("MINE") comprising:
   an inference engine configured to infer referral activity not reported into a referral workflow system utilizing intent-based clustering of medical information;
   an analyzer configured to generate reporting metrics from the inferred referral activity; and
   a display configured to display the metrics on a dashboard.

10. The MINE of claim 9 wherein the inference engine further comprises:
    a medical information interface configured to receive medical information from a plurality of medical sources;
    a reconciliation engine configured to reconcile the medical information; and
    an intent-based presentation engine configured to cluster the reconciled medical information, wherein the clustering includes applying at least one clustering rule to the reconciled medical information.

11. The MINE of claim 10 wherein the reconciling further comprises applying at least one similarity rule.

12. The MINE of claim 11 wherein the at least one similarity rule includes comparing patient data attributes.

13. The MINE of claim 12 wherein the intent-based presentation engine is further configured to compute a distance between the patient data attributes and clustered reconciled medical information.

14. The MINE of claim 13 wherein if the distance is less than a threshold, then the clustered reconciled medical information is included in a presentation cluster prepared for the user.

15. The MINE of claim 13 wherein the analyzer is further configured to generate the metrics by adding the inferred referral activity to reported referral activity to generate a combined referral number, and dividing the reported referral activity by the combined referral number.

16. The MINE of claim 11 wherein the analyzer is further configured to generate the metrics for in network referrals and out of network referrals.

* * * * *