(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 November 2015 (05.11.2015)

WIPO | PCT

(10) International Publication Number
WO 2015/166508 A1

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/IN2014/000290

(22) International Filing Date:
30 April 2014 (30.04.2014)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **HEWLETT-PACKARD DEVELOPMENT COMPANY, L.P.** [US/US]; 11445 Compaq Center Drive West, Houston, Texas 77070 (US).

(72) Inventor; and
(71) Applicant *(for US only)*: **PILLAI, Vinu** [IN/IN]; HP Avenue, 39/40, Electronics City, Hosur Road, Bangalore, Karnataka 560100 (IN).

(74) Agent: **ALANKI, N.V. Pradeep Kumar**; Global IP Services, 198F, 27th, Cross, 3rd Block, Jayanagar, Bangalore 560011, Karnataka (IN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to the identity of the inventor (Rule 4.17(i))*

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

(54) Title: CORRELATION BASED INSTRUMENTS DISCOVERY

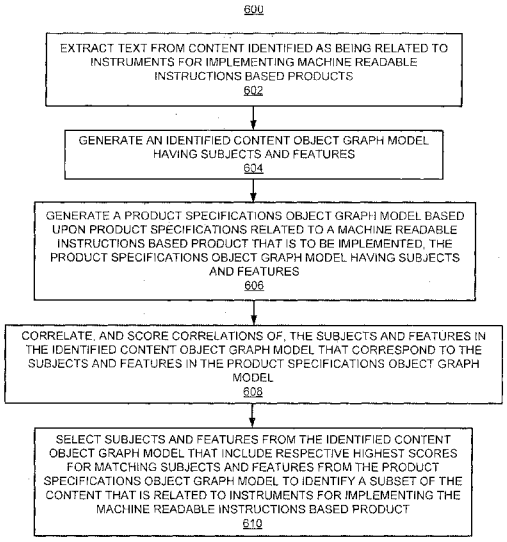


FIG. 6

(57) Abstract: According to an example, correlation based instruments discovery may include extracting text from content identified as being related to instruments for implementing machine readable instructions based products. An identified content object graph model having subjects and features may be generated. A product specifications object graph model having subjects and features may be generated based upon product specifications related to a machine readable instructions based product that is to be implemented. The subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model may be correlated and scored. Subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model may be selected to identify a subset of the content that is related to instruments for implementing the machine readable instructions based product.

WO 2015/166508 A1

CORRELATION BASED INSTRUMENTS DISCOVERY

BACKGROUND

[0001] In the architecture and design phase of a software (i.e., machine readable instructions) based application, typically a search is performed for a set of instruments that are best suited to deliver an end solution. Instruments typically include various machine readable instructions based applications, libraries, and/or components that are needed to implement the end solution. The search is typically based on a set of specifications outlined by stake holders associated with the end solution. Experts in a particular area associated with the end solution typically use their knowledge and recommendations from other experts to select a set of instruments that best meet the set of specifications and the end solution. For example, an expert typically analyzes the set of specifications and searches through a library of available instruments to select an instrument that best meets the set of specifications and the end solution.

BRIEF DESCRIPTION OF DRAWINGS

[0002] Features of the present disclosure are illustrated by way of example and not limited in the following figure(s), in which like numerals indicate like elements, in which:

[0003] Figure 1 illustrates an architecture of a correlation based instruments discovery apparatus, according to an example of the present disclosure;

[0004] Figure 2 illustrates further details of the correlation based instruments discovery apparatus of Figure 1, according to an example of the present disclosure;

[0005] Figure 3 illustrates a flowchart of correlation based instruments discovery, according to an example of the present disclosure;

[0006] Figure 4 illustrates a product specifications object graph model for a network monitoring tool example described herein, according to an example of the present disclosure;

[0007] Figure 5 illustrates word arrangement for correlation, according to an example of the present disclosure;

[0008] Figure 6illustrates a flowchart of a method for correlation based instruments discovery, according to an example of the present disclosure;

[0009] Figure 7illustrates further details of the flowchart of the method for correlation based instruments discovery, according to an example of the present disclosure; and

[0010] Figure 8illustrates a computer system, according to an example of the present disclosure.

DETAILED DESCRIPTION

[0011] For simplicity and illustrative purposes, the present disclosure is described by referring mainly to examples. In the following description, numerous specific details are set forth in order to provide a thorough understanding of the present disclosure. It will be readily apparent however, that the present disclosure may be practiced without limitation to these specific details. In other instances, some methods and structures have not been described in detail so as not to unnecessarily obscure the present disclosure.

[0012] Throughout the present disclosure, the terms "a" and "an" are intended to denote at least one of a particular element. As used herein, the term "includes" means includes but not limited to, the term "including" means including but not limited to. The term "based on" means based at least in part on.

[0013] In the architecture and design phase of a machine readable instructions based application, typically an expert analyzes a set of specifications outlined by stakeholders associated with an end solution, and searches through a library of available instruments to select an instrument that best meets the set of specifications and the end solution. Such an analysis is limited to the set of instruments that are available to the expert, the expert's experience, and/or the expert's bias towards selection of a particular instrument. Such an analysis thus omits consideration of numerous other instruments that may also provide a competitive offering, without an expert even being aware of the existence of such other instruments. An objective comparison of available instruments is often time consuming, for example, due to the dynamic nature of new instruments that come into existence. Stakeholders are also limited to reliance on expert opinion due to their lack of knowledge of particular instruments related to an end solution. Stakeholders are further limited to having to use an instrument that is recommended by an expert, without having a metric that indicates how well an instrument meets the original set of specifications on an ongoing basis.

[0014] According to examples, a correlation based instruments discovery apparatus and a method for correlation based instruments discovery are disclosed herein. The apparatus and method disclosed herein provide for the discovery of instruments based on facts from a given set of user specifications. The apparatus and method disclosed herein may provide a comparative report of the comparison on what instrument works best for a particular set of user specifications. The user specifications may be provided as product specifications related to a product that is to be implemented. The product may include any generic entity, such as, for example, a machine readable instructions based application, or a machine readable instructions and hardware based computer component that is used, for example, for a machine readable instructions based implementation. A product specifications satisfaction index may be used to measure how close an instrument selection matches the user product specifications. A cost based comparison may also be provided as an indication of a total cost accrued for a particular instrument. The product specifications satisfaction index and the cost may provide a balance between cost and the instrument capabilities that may be used to provide the best value for cost solution.

[0015] The apparatus and method disclosed herein may be used to implement other aspects related to instrument selection, such as, for example, a specification based recommendation engine (i.e., a mapping module as described herein) for sizing, and for costs associated with instruments for different industries. The apparatus and method disclosed herein may also be used to implement a knowledge helpdesk on instruments that are regularly updated, hardware sizing to determine what configuration best suits a set of product specifications, etc.

[0016] Generally, the apparatus and method disclosed herein may combine internal enterprise sources, external enterprise sources, and social media sources to build a comprehensive ecosystem of instruments which can be used by different applications. The internal enterprise sources may include, for example, instrument specifications (e.g., features related to instruments), instrument inventories (e.g.,

types of available instruments), and/or product assets (e.g., products that use multiple instruments). The external enterprise sources may include, for example, product specification websites (e.g., websites for features related to instruments), vendor portals (e.g., websites for different vendors for different instruments), product documentation (e.g., general information related to instruments), product forums (e.g., forums related to specific products that use instruments), framework specifications (e.g., specifications related to structure of instruments), and/or knowledge management websites (e.g., comprehensive knowledge websites related to instruments). The social media sources may include, for example, blogs (e.g., social blogs related to instruments), forums (e.g., social forums related to instruments), and/or knowledge management systems (e.g., comprehensive social websites related to instruments).

[0017] The apparatus and method disclosed herein may capture a standard set of specifications for an end solution (i.e., product specifications). A template may be used to capture the set of specifications for an end solution. The template may include a technology independent format, such as, for example, Extensible Markup Language (XML) for capturing specifications, and ensuring capture of all needed information. The template may be used to extract information that a specific instrument needs. The template may be modified as needed based on modifications to the instruments, and/or product specifications for the end solution. Additionally or alternatively, the template may be modified as needed based on results related to the end solution so that the results are optimized at each successive implementation of the apparatus and method disclosed herein.

[0018] The apparatus and method disclosed herein may facilitate application rationalization. Application rationalization is typically a tedious and cumbersome process which involves a deep study of a portfolio of instruments, and then deciding which instruments need to be retained or used, and which instruments need to be replaced. The apparatus and method disclosed herein may also serve as a recommendation engine for identifying the instruments which align with an

overall corporate strategy. The apparatus and method disclosed herein may further provide recommendations on the best strategy to perform rationalization.

[0019] The apparatus and method disclosed herein may provide a customized catalog. For example, the reach of instruments may be enhanced by providing a customized catalog for a given set of machine readable instructions applications.

[0020] The apparatus and method disclosed herein may provide an indication of whether the needs of a stakeholder are met with respect to an end solution. For example, projects that employ a variety of instruments typically do not have a clear picture of how effectively the instruments have been designed, and the degree to which they have met the specifications. The apparatus and method disclosed herein may ensure that stakeholders' needs are met, and the right instruments used in information technology (IT) are used to provide a solution to the stakeholders. A stakeholder may also be able to determine the best combination of solutions that fits a budget and matches all of the stakeholders' needs.

[0021] The apparatus and method disclosed herein may be used to provide support for a company's portfolio of instruments with fact based recommendations to thus increase the probability of a customer adopting a certain product as it is backed by the customer specifications. The apparatus and method disclosed herein may provide for feature discovery, which may be used to perform a more economical instrument rationalization within a project. Based on a wider choice of available instruments, a project may also be sized based on a budget. The apparatus and method disclosed herein may also facilitate licensing of products based on the identification of appropriate instruments for a given project. With respect to knowledge and helpdesk systems, data that is needed to implement such systems may be enhanced within an enterprise with more updated content. For example, the apparatus and method disclosed herein may provide an efficient way of obtaining the best and latest content crowd sourced and coupled with facts. The apparatus and method disclosed herein may also facilitate implementation of a services ecosystem based on the instruments that are discovered for a particular

project. Further, the apparatus and method disclosed herein may be implemented in a variety of domains, such as, for example, healthcare, education, tourism, travel, eCommerce, etc.

[0022] Figure 1 illustrates an architecture of a correlation based instruments discovery apparatus (hereinafter also referred to as "apparatus 100"), according to an example of the present disclosure. Referring to Figure 1, the apparatus 100 is depicted as including an extraction module 102 to receive content from a variety of sources that include internal enterprise sources 104, external enterprise sources 106, and social media sources 108 to build a comprehensive ecosystem of instruments, which may be used by different applications. The internal enterprise sources 104 may include, for example, instrument specifications, instrument inventories, and/or product assets. The external enterprise sources 106 may include, for example, product specification websites, vendor portals, product documentation, product forums, framework specifications, and/or knowledge management websites. The social media sources 108 may include, for example, blogs, forums, and/or knowledge management systems. As described herein with reference to Figures 1 and 2, the extraction module 102 may receive product specifications 110 related to a product 112, and extract data from the internal enterprise sources 104, the external enterprise sources 106, the social media sources 108, and the product specifications 110. The extraction module 102 may include a crawler 114, a data source management module 116, and connectors 118, the operation of each of which is described below, to extract the data from the internal enterprise sources 104, the external enterprise sources 106, the social media sources 108, and the product specifications 110.

[0023] An object graph model generation module 120 is to extract the subjectivity being described in the text of the data from the internal enterprise sources 104, the external enterprise sources 106, the social media sources 108, and the product specifications 110. Further, the object graph model generation module 120 may extract the features and the sentiment if any associated with

those features using techniques such as subjectivity extraction and classification, implicit and explicit feature detection, and word sense disambiguation. The object graph model generation module 120 may include a feature extraction module 122, a subjectivity classification module 124, a natural language processing module 126, and a word sense disambiguator module 128, the operation of each of which is described below, to generate object graph models130of subjects that are identified by the object graph model generation module 120. Generally, the object graph model generation module 120 may perform subjectivity, and implicit and explicit feature extraction on the extracted text from the identified content of the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108, to generate an identified content object graph model. Further, the object graph model generation module 120 may perform subjectivity, and implicit and explicit feature extraction on the text from the product specifications110 to generate a product specifications object graph model.

[0024] A mapping module 132 is to identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model. The mapping module 132 may correlate, and score the correlation of, the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model. Further, the mapping module 132 may select subjects from the identified content object graph model that include respective highest scores for matching subjects from the product specifications to identify instruments for implementing the machine readable instructions based product. The mapping module 132 may include a category extraction module 134, a correlation module 136, a scoring module 138, and a comparison module 140, the operation of each of which is described below, to generate an instrument recommendation 142based on a subset with the highest match. The instrument recommendation 142may be further optimized using learning capabilities to fine tune the output for more accurate matches. Generally, the mapping module 132

may compare the data patterns extracted from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108, with the data patterns available from the product specifications 110. The mapping module 132 may perform text analytics and word/text ranking to determining the word categories, subjectivity with implicit and explicit features detection, tone of the sentences, and assign a score to each matching entity. A correlation model may be generated to arrive at an index, and to determine the degree of match between the entities. A comparison may be performed with all entities and their associated features to determine the overall score, and to arrive at the final index. Based on the comparison, the greatest matching subset may be presented as the instrument recommendation 142.

[0025] The modules and other elements of the apparatus 100 may be machine readable instructions stored on a non-transitory computer readable medium. In this regard, the apparatus 100 may include or be a non-transitory computer readable medium. In addition, or alternatively, the modules and other elements of the apparatus 100 may be hardware or a combination of machine readable instructions and hardware.

[0026] Referring to Figures 1 and 2, the crawler 114of the extraction module 102 may extract text from content of the product specifications 110, and from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108. The internal enterprise sources 104, the external enterprise sources 106, the social media sources 108 may be configurable, and updated with newer content. The crawler 114 may operate in conjunction with a scheduler to implement automatic (e.g., without human intervention) updates related to the crawling operation.

[0027] The data source management module 116of the extraction module 102 may manage various aspects of the extraction. For example, the data source management module 116 manages the uniform resource locators (URLs) that are to be accessed for the internal enterprise sources 104, the external enterprise

sources 106, and the social media sources 108. According to another example, the data source management module 116 manages how many times the URLs are to be accessed, and/or sequence of the URLs that are to be accessed.

[0028] The connectors 118 of the extraction module 102 may provide the connection to a variety of sources that include internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108. Referring to Figures 1 and 2, the connectors 118 may be used to connect to the internal enterprise sources 104 via a corporate intranet. Further, the connectors 118 may be used to connect to the external enterprise sources 106 and the social media sources 108 via the Internet. Generally, the connectors 118 may provide the connection to communicate with a specific type of data source.

[0029] The feature extraction module 122 of the object graph model generation module 120 may extract features and the sentiment, if any, associated with the features. The features and sentiment may be extracted based on implicit and explicit feature detection, and word sense disambiguation determined by the word sense disambiguator module 128. The implicit and explicit feature detection may identify what specific feature is being described within a sentence and for a particular subject. For example, a sentence may discuss the feature of a screen for a mobile phone. The implicit feature detection may be based, for example, on an evaluation of content adjacent the content that is being evaluated.

[0030] The subjectivity classification module 124 of the object graph model generation module 120 may extract the subjectivity being described in the text. The subjectivity may be extracted based on subjectivity extraction and classification. For example, the subjectivity classification module 124 extracts the subjectivity being described in a sentence (e.g., what the sentence is describing, what is the subject, etc.). For example, a sentence may discuss the feature of a screen for a subject of a mobile phone.

[0031] The natural language processing module 126of the object graph model generation module 120 may be used for language processing for content of the internal enterprise sources 104, the external enterprise sources 106, the social media sources 108, and the product specifications 110. The natural language processing module 126 may derive meaning from the content of the internal enterprise sources 104, the external enterprise sources 106, the social media sources 108, and the product specifications 110. The feature extraction module 122, the subjectivity classification module 124, and the word sense disambiguator module 128 may all be part of the natural language processing module 126.

[0032] The object graph model generation module 120 may generate the object graph models 130for entities(e.g., subjects) that are extracted by the extraction module 102. The object graph model 130 may be generated for each entity that is discovered in the text content. The object graph model 130 may be built by linking each subject and the features that are related to that subject and other appropriate subjects. Similarly, subjects may be linked to other appropriate subjects. In this manner the object graph model 130 may include multiple relationships between subjects and features.

[0033] The information determined by the object graph model generation module 120 may be stored in a graph database 202. The graph database 202 may be organized to store the object graph models 130 that include multiple relationships between subjects and features. In this manner, the object graph models 130 may be navigated to ascertain subjects and associated subjects and/or features. For example, a particular server may be created as a subject and include a plurality of features. The features may be specified, for example, in a product specification website for one of the external enterprise sources 106. The subject of the particular server may be further linked to subjects of other servers and related features to provide the object graph model 130 for servers.

[0034] The category extraction module 134 of the mapping module 112 may extract the appropriate category of entities from the graph database 202 for analyzing the product specifications 110 against the appropriate content from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108. For example, as described herein, the category extraction module 134 may identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model.

[0035] The correlation module 136 of the mapping module 112 may correlate the data extracted by the extraction module 102 to the object graph model 130 generated by the object graph model generation module 120to determine the nearest set of matching subjects and features. The correlation module 136 may create a correlation model to arrive at the index, and to determine the degree of match between the entities. For example, as described herein, the correlation module 136 may correlate the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model.

[0036] The scoring module 138 of the mapping module 112 may use the correlation performed by the correlation module 136 to determine the nearest set of matching objects and features by scoring the subjects and features. The scoring module 138 may include text analytics and word/text ranking capabilities to determine the word categories, subjectivity with implicit and explicit features detection, tone of the sentences, and assign a score to each matching entity. For example, as described herein, the scoring module 138 may score the correlation of the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model. An example of a scoring pattern used by the scoring module 138 may match a predetermined number of features for a given subject to the predetermined number of features for another given subject.

[0037] The comparison module 140 of the mapping module 112 may compare the data patterns extracted from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108, with the data patterns available from the product specifications 110. Alternatively or additionally, the comparison module 140 may compare the data patterns extracted from the web with the data patterns available from the product specifications 110. The comparison module 140 may compare all entities and their associated features to determine the overall score, and to arrive at a final index. The greatest matching subset may be presented as the instrument recommendation 142. For example, as described herein, the comparison module 140 may select subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify instruments for implementing the machine readable instructions based product (i.e., the product 112).

[0038] As shown in Figure 2, at 200, the information such as the product specifications 110 and the instrument recommendation 142 may be respectively received from and provided to a user via a variety of channels. For example, the channels include mobile phones, tablets, etc.

[0039] As shown in Figure 2, the apparatus 100 may include various other components, such as, for example, a catalog administration module 204, a notification module 206, a reports module 208, a search module 210, and a data access layer 212. The catalog administration module 204 may be used to define terms (e.g., any subjects, features) that are used in the graph database 202. The notification module 206 may be used to trigger alerts related to the content of the graph database 202. For example, based on the identification of a certain subject, feature, or score, an alert may be triggered. The reports module 208 may generate reports related, for example, to a cost based comparison as an indication of a total cost accrued for a particular instrument. The reports may include a product specifications satisfaction index, and the cost may provide a balance between cost

and the instrument capabilities that may be used to provide the best value for cost solution. The search module 210 may be used to enter a search term, and may search through the graph database 202 to generate results related to the search term. The data access layer 212 may be used to access the object graph model data stored in the graph database 202.

[0040] Figure 3 illustrates a flowchart 300 of correlation based instruments discovery, according to an example of the present disclosure.

[0041] Referring to Figure 3, at block 302, the data source management module 116 of the extraction module 102 may organize a list of URLs to be crawled as categories for the purpose of scheduling. For example, the data source management module 116 may identify content related to instruments for implementing machine readable instructions based products from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108.

[0042] At block 304, the data source management module 116 of the extraction module 102 may pick a group of URLs associated with a particular category to be crawled.

[0043] At block 306,for each URL that is picked at block 304, a determination is made as to whether a URL of the URLs identified at block 304 has already been visited.

[0044] In response to a determination that the URL of the URLs identified at block 304 has already been visited, at block 308,a determination is made as to whether content associated with the URL has changed.

[0045] In response to a determination that content associated with the URL has not changed, at block 310, further processing related to the URL of the URLs identified at block 304 is stopped.

[0046] In response to a determination that content associated with the URL has changed, or in response to a determination that the URL of the URLs identified at block 304 has not been visited, at block 312, the extraction module 102 and the object graph model generation module 120 may process text from the URL. For example, according to a network monitoring tool example described herein, the URLs subject to the extraction include URL1 and URL2. At block 314, the text processed by the extraction module 102 and the object graph model generation module 120 may be stored as the object graph models 130 in the graph database 202.

[0047] At block 316,theextraction module 102 may extract text content from the URL of the URLs identified at block 304.Instead of separate extractors, a common strategy may be followed, which is to convert the URL content into plain text after removing the Hyper Text Markup Language (HTML) tags from HTML pages, and then performing context and subjectivity extraction.

[0048] At block 318, the object graph model generation module 120 may perform subjectivity extraction, along with extraction of implicit and explicit features from the text. For example, the feature extraction module 122, the subjectivity classification module 124, and the word sense disambiguator module 128 may perform subjectivity extraction, along with extraction of implicit and explicit features from the text. For the network monitoring tool example described herein, URL1 may yield "Solar Winds Network Configuration Manager" as a subject with features, attributes, and/or properties as "manage complex network environments", "automated network discovery", "importing discovered devices", etc. URL1 may yield many other subject and associated features, attributes, and/or properties.

[0049] At block 320, the object graph model generation module 120 may generate the object graph models 130 of subjects and features that are extracted by the object graph model generation module 120. Figure 4 illustrates an object graph model for the network monitoring tool example described herein. The object graph model may represent entities and associated features, attributes, and/or

properties. The object graph model may be visualized as a node with multiple branches. Referring to Figure 4, the nodes (i.e., entities) may be represented as larger circles at 400, and the associated features, attributes, and/or properties that are determined from specific content may be represented as smaller circles at 402. For example, nodes E1 and E2 may represent entities, and A1-A7 may represent the associated features, attributes, and/or properties. A node (e.g., E1) may be related to another node (e.g., E2) through an implicit or an explicit relationship based on the content. Further, an attribute may itself become a node if it has other discovered attributes in the content.

[0050] For the network monitoring tool example described herein, a product specifications object graph model maybe designated as G1, and identified content object graph models for URL1 and URL2 may be respectively designated as G2 and G3.

[0051] At block 322, the text for the product specifications 110 related to the product 112 may be received. For example, for the network monitoring tool example described herein, the text for the product specifications 110 may indicate "we need a network monitoring tool for our large network. We need to monitor routers, switches, firewalls, different circuits, VPN connections and servers to name a few. The number of devices is expected to be more than 500 and hence we need a robust tool for capturing data from a large number of devices. We would like something that has all of the features but does not have a large footprint and it should not interrupt normal activity, such as causing the core switches and CPU to spike. We would also like to have a vendor with great tech support who will assist with the installation and teach the engineers of the proper usage and capabilities of the tool."

[0052] At block 324, the extraction module 102 and the object graph model generation module 120 may process the text from the product specifications 110. Specifically, at block 316, the extraction module 102 may extract text content from the product specifications 110. At block 318, the object graph model generation

module 120 may perform subjectivity extraction, along with extraction of implicit and explicit features from the text. For example, the feature extraction module 122, the subjectivity classification module 124, and the word sense disambiguator module 128 may perform subjectivity extraction, along with extraction of implicit and explicit features from the text. At 320,a product specifications object graph model may be generated for the product specifications 110.

[0053] For the network monitoring tool example described herein, the product specifications 110 refer to network monitoring tool and network as two entities which may be treated as subjects. Referring to Figure 4, assuming the object graph model of Figure 4 represents the product specifications object graph model G1, these entities may be symbolically visualized as E1 and E2 respectively (i.e., the primary subjects). The network monitoring tool (E1's) attributes are "footprint" as "low", "interrupt normal activity" as "minimal", "vendor tech support" as "great", and "installation" as "assist/assistance". The network (E2's) attributes are "size" as "large", "monitor" as "routers, switches, firewalls, different circuits, VPN connections, and servers", and "number of devices" as "more than 500". E1 and E2 may be related to each other in the first sentence which indicates that E1 is used to monitor E2. The attributes "routers", "firewalls", "VPN Connections", and "servers" may themselves operate as subjects or nodes, may include further attributes based on the content from other parts of the text, and may be related in a similar manner.

[0054] At block 326, with respect to the product specifications 110, the mapping module 132 may list each subject and its associated features.

[0055] At block 328, the mapping module 132 may perform a search for each subject and its associated features in the graph database 202. For example, the category extraction module 134 may identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model.

[0056] At block 330,the correlation module 136 may perform a correlation between the entities of the product specifications 110 and entities retrieved from the graph database 202. For example, the correlation module 136 may correlate the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model. For the network monitoring tool example described herein, the correlation module 136 may perform a correlation between G1-G2, and G1-G3 to extract the highest optimum results. Each node of the object graph models G1, G2, and G3 may be iterated, and the features, attributes, and/or properties may be examined. Various scenarios may exist in the correlation process. For example, a simple case may be where G2's or G3's attribute values are similar to that of G1. In such cases, identification may involve a text comparison. However, the attribute values of G2 or G3 may be synonyms or contextually similar terms to the attribute values of G1. In this regard, a combination of word sense disambiguation techniques may be coupled with using a dictionary of words or capabilities of a the natural language processing module 126. For example, a lexical database such as WORDNET may be used. In this regard, nouns, verbs, adjectives, and adverbs may be grouped into sets of cognitive synonyms (i.e., synsets), each expressing a distinct concept. Synsets may be interlinked by conceptual-semantic and lexical relations. The resulting network of meaningfully related words and concepts may be navigated using application programming interfaces (APIs). Nouns, verbs, adjectives, and adverbs may be distinguished based on their use of different grammatical rules. Every synset may include a group of synonymous, words, or collocations (a collocation is a sequence of words that form a specific meaning, such as "car pool"). Different senses of a word may be in different synsets. Every attribute value in a source (e.g., G1) may be analyzed againsteach attribute value in targets (e.g., G2, G3, ... etc.) within the WORDNET hierarchy.

[0057] Figure 5 illustrates word arrangement for correlation, according to an example of the present disclosure. Referring to Figure 5, if "8" is an attribute value from G1, and "5" is an attribute value from G2, the shortest distance between these two words (5 → 1 → 3 → 8) may be treated as a measure of how close these match. Similarly if "8" is an attribute value from G1,and "12" is an attribute value from G3, the shortest distance between these two (12 → 10 → 5 → 1 → 3 → 8) may be treated as a measure of how close these attributes match. In the first case, the distance is 3 (n – 1, where n is 4 and represents the number of nodes in the path), and in the second case the distance is5 (n – 1, where n is 6). Thus the shortest value in these two correlations (G1→ G2 and G1 → G3) yields3 as the highest optimum result, and thus G2 may be chosen as a potential search result. Further analysis may be performed for other attributes.

[0058] The scoring module 138 may score the correlation between the entities of the product specifications 110 and entities retrieved from the graph database 202. For example, the scoring module 138 may score the correlation of the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model.

[0059] At block 332, the comparison module 140 may select the entities with the highest match and output the entities as the instrument recommendation 142. For example, the comparison module 140 may select the top five (or otherwise user-configurable) entities with the highest match, and output the entities as the instrument recommendation 142. For example, the comparison module 140 may select subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify instruments for implementing the product 112.

[0060] Figures 6and 7 respectively illustrate flowcharts of methods 600 and 700 for correlation based instruments discovery, corresponding to the example of the correlation based instruments discovery apparatus 100 whose construction is

described in detail above. The methods 600 and 700 may be implemented on the correlation based instruments discovery apparatus 100 with reference to Figures 1-3by way of example and not limitation. The methods 600 and 700 may be practiced in other apparatus.

[0061] Referring to Figure 6, for the method 600, at block 602, the method may include extracting text from content identified as being related to instruments for implementing machine readable instructions based products. For example, referring to Figures 1 and 2, the extraction module 102 may extract text from content identified as being related to instruments for implementing machine readable instructions based products.

[0062] At block 604, the method may include generating an identified content object graph model having subjects and features. For example, referring to Figures 1 and 2, the object graph model generation module 120 may generate an identified content object graph model having subjects and features.

[0063] At block 606, the method may include generating a product specifications object graph model based upon product specifications related to a machine readable instructions based product that is to be implemented. For example, referring to Figures 1 and 2, the object graph model generation module 120 may generate a product specifications object graph model based upon the product specifications 110 related to a machine readable instructions based product that is to be implemented. The product specifications object graph model may include subjects and features.

[0064] At block 608, the method may include correlating, and scoring correlations of, the subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model. For example, referring to Figures 1 and 2, the mapping module 132 may correlate, and score correlations of, the subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model.

[0065] At block 610, the method may include selecting subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content that is related to instruments for implementing the machine readable instructions based product. For example, referring to Figures 1 and 2, the comparison module 140 may select subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content (e.g., as the instrument recommendation 142) that is related to instruments for implementing the machine readable instructions based product.

[0066] According to an example, extracting text from content identified as being related to instruments for implementing machine readable instructions based products may further include using a crawler to extract the text from content identified as being related to instruments for implementing machine readable instructions based products.

[0067] According to an example, extracting text from content identified as being related to instruments for implementing machine readable instructions based products may further include extracting text from content identified as being related to instruments for implementing machine readable instructions based products from internal enterprise sources, external enterprise sources, and social media sources.

[0068] According to an example, the method 600 may further include performing subjectivity and feature extraction on the extracted text from the identified content to generate the identified content object graph model having subjects and features. According to an example, performing subjectivity and feature extraction on the extracted text from the identified content to generate the identified content object graph model having subjects and features may further include performing implicit and explicit feature extraction on the extracted text from the

identified content to generate the identified content object graph model having subjects and features.

[0069] According to an example, the method 600 may further include performing subjectivity and feature extraction on text from the product specifications to generate the product specifications object graph model having subjects and features. According to an example, performing subjectivity and feature extraction on text from the product specifications to generate the product specifications object graph model having subjects and features may further include performing implicit and explicit feature extraction on the text from the product specifications to generate the product specifications object graph model having subjects and features.

[0070] According to an example, scoring correlations of the subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model may further include determining a degree of match between the subjects and features in the identified content object graph model and the subjects and features in the product specifications object graph model.

[0071] Referring to Figure 7, for the method 700, at block 702, the method may include identifying content related to instruments for implementing computer products from internal enterprise sources, external enterprise sources, and social media sources. For example, referring to Figures 1 and 2, the data source management module 116 may identify content related to instruments for implementing computer products from the internal enterprise sources 104, the external enterprise sources 106, and the social media sources 108.

[0072] At block 704, the method may include extracting text from the identified content. For example, referring to Figures 1 and 2, the crawler 114 may extract text from the identified content.

[0073] At block 706, the method may include performing subjectivity, and implicit and explicit feature extraction on the extracted text from the identified content to generate an identified content object graph model. For example, referring to Figures 1 and 2, the object graph model generation module 120 may perform subjectivity, and implicit and explicit feature extraction on the extracted text from the identified content to generate an identified content object graph model. The identified content object graph model may link subjects and features extracted from the text from the identified content.

[0074] At block 708, the method may include receiving product specifications related to a computer product that is to be implemented. For example, referring to Figures 1 and 2, the extraction module 102 may receive product specifications 110 related to a computer product (e.g., the product 112) that is to be implemented.

[0075] At block 710, the method may include extracting text from the product specifications. For example, referring to Figures 1 and 2, the crawler 114 may extract text from the product specifications 110.

[0076] At block 712, the method may include performing subjectivity, and implicit and explicit feature extraction on the text from the product specifications to generate a product specifications object graph model. For example, referring to Figures 1 and 2, object graph model generation module 120 may perform subjectivity, and implicit and explicit feature extraction on the text from the product specifications110 to generate a product specifications object graph model. The product specifications object graph model may link subjects and features extracted from the text from the product specifications.

[0077] At block 714, the method may include identifying subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model. For example, referring to Figures 1 and 2, the category extraction module 134 may identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model.

[0078] At block 716, the method may include correlating, and scoring the correlation of, the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model. For example, referring to Figures 1 and 2, the correlation module 136 and the scoring module 138 may respectively correlate, and score the correlation of, the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model. The scoring may include determining a degree of match between the identified subjects and features in the identified content object graph model and the subjects and features in the product specifications object graph model.

[0079] At block 718, the method may include selecting subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content that is related to instruments for implementing the computer product. For example, referring to Figures 1 and 2, the comparison module 140 may select subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content (e.g., as the instrument recommendation 142) that is related to instruments for implementing the computer product.

[0080] Figure 8 shows a computer system 800 that may be used with the examples described herein. The computer system 800 may represent a generic platform that includes components that may be in a server or another computer system. The computer system 800 may be used as a platform for the apparatus 100. The computer system 800 may execute, by a processor (e.g., a single or multiple processors) or other hardware processing circuit, the methods, functions and other processes described herein. These methods, functions and other processes may be embodied as machine readable instructions stored on a computer readable medium, which may be non-transitory, such as hardware

storage devices (e.g., RAM (random access memory), ROM (read only memory), EPROM (erasable, programmable ROM), EEPROM (electrically erasable, programmable ROM), hard drives, and flash memory).

[0081] The computer system 800 may include a processor 802 that may implement or execute machine readable instructions performing some or all of the methods, functions and other processes described herein. Commands and data from the processor 802 may be communicated over a communication bus 804. The computer system may also include a main memory 806, such as a random access memory (RAM), where the machine readable instructions and data for the processor 802 may reside during runtime, and a secondary data storage 808, which may be non-volatile and stores machine readable instructions and data. The memory and data storage are examples of computer readable mediums. The memory 806 may include a correlation based instruments discovery module 820 including machine readable instructions residing in the memory 806 during runtime and executed by the processor 802. The correlation based instruments discovery module 820 may include the modules of the apparatus 100 shown in Figure 1.

[0082] The computer system 800 may include an I/O device 810, such as a keyboard, a mouse, a display, etc. The computer system may include a network interface 812 for connecting to a network. Other known electronic components may be added or substituted in the computer system.

[0083] What has been described and illustrated herein is an example along with some of its variations. The terms, descriptions and figures used herein are set forth by way of illustration only and are not meant as limitations. Many variations are possible within the spirit and scope of the subject matter, which is intended to be defined by the following claims -- and their equivalents -- in which all terms are meant in their broadest reasonable sense unless otherwise indicated.

What is claimed is:

1. A method for correlation based instruments discovery, the method comprising:

extracting text from content identified as being related to instruments for implementing machine readable instructions based products;

generating, by a processor, an identified content object graph model having subjects and features;

generating a product specifications object graph model based upon product specifications related to a machine readable instructions based product that is to be implemented, the product specifications object graph model having subjects and features;

correlating, and scoring correlations of, the subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model; and

selecting subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content that is related to instruments for implementing the machine readable instructions based product.

2. The method of claim 1,wherein extracting text from content identified as being related to instruments for implementing machine readable instructions based products further comprises:

using a crawler to extract the text from content identified as being related to instruments for implementing machine readable instructions based products.

3. The method of claim 1, wherein extracting text from content identified as being related to instruments for implementing machine readable instructions based products further comprises:

extracting text from content identified as being related to instruments for implementing machine readable instructions based products from internal enterprise sources, external enterprise sources, and social media sources.

4. The method of claim 3, wherein the internal enterprise sources include sources related to at least one of instrument specifications, instrument inventories, and product assets.

5. The method of claim 3, wherein the external enterprise sources include sources related to at least one of product specification websites, vendor portals, product documentation, product forums, framework specifications, and knowledge management websites.

6. The method of claim 3, wherein the social media sources include sources related to at least one of blogs, forums, and knowledge management systems.

7. The method of claim 1, further comprising:

performing subjectivity and feature extraction on the extracted text from the identified content to generate the identified content object graph model having subjects and features.

8. The method of claim 7, wherein performing subjectivity and feature extraction on the extracted text from the identified content to generate the identified content object graph model having subjects and features further comprises:

performing implicit and explicit feature extraction on the extracted text from the identified content to generate the identified content object graph model having subjects and features.

9. The method of claim 1, further comprising:

performing subjectivity and feature extraction on text from the product specifications to generate the product specifications object graph model having subjects and features.

10. The method of claim 9, wherein performing subjectivity and feature extraction on text from the product specifications to generate the product specifications object graph model having subjects and features further comprises:

performing implicit and explicit feature extraction on the text from the product specifications to generate the product specifications object graph model having subjects and features.

11. The method of claim 1,wherein scoring correlations of the subjects and features in the identified content object graph model that correspond to the subjects and features in the product specifications object graph model further comprises:

determining a degree of match between the subjects and features in the identified content object graph model and the subjects and features in the product specifications object graph model.

12. The method of claim 1, wherein the identified content object graph model links the subjects and features from the extracted text from the identified content.

13. A correlation based instruments discovery apparatus comprising:

a processor; and

a memory storing machine readable instructions that when executed by the processor cause the processor to:

identify content related to instruments for implementing computer products from internal enterprise sources, external enterprise sources, and social media sources;

extract text from the identified content;

perform subjectivity, and implicit and explicit feature extraction on the extracted text from the identified content to generate an identified content object graph model;

receive product specifications related to a computer product that is to be implemented;

extract text from the product specifications;

perform subjectivity, and implicit and explicit feature extraction on the text from the product specifications to generate a product specifications object graph model;

identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model;

correlate, and score the correlation of, the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model; and

select subjects and features from the identified content object graph model that include respective predetermined scores for matching subjects and features from the product specifications object graph model to identify a subset of the content that is related to instruments for implementing the computer product.

14. The correlation based instruments discovery apparatus according to claim 13, wherein the predetermined scores represent the highest scores.

15. A non-transitory computer readable medium having stored thereon machine readable instructions to provide correlation based instruments discovery, the machine readable instructions, when executed, cause a processor to:

identify content related to instruments for implementing computer products from internal enterprise sources, external enterprise sources, and social media sources;

extract text from the identified content;

perform, by a processor, subjectivity, and implicit and explicit feature extraction on the extracted text from the identified content to generate an identified content object graph model;

receive product specifications related to a computer product that is to be implemented;

extract text from the product specifications;

perform subjectivity, and implicit and explicit feature extraction on the text from the product specifications to generate a product specifications object graph model;

identify subjects and features in the identified content object graph model that correspond to subjects and features in the product specifications object graph model;

correlate, and score the correlation of, the identified subjects and features in the identified content object graph model to the subjects and features in the product specifications object graph model, wherein the scoring includes determining a degree of match between the identified subjects and features in the identified content object graph model and the subjects and features in the product specifications object graph model; and

select subjects and features from the identified content object graph model that include respective highest scores for matching subjects and features from the product specifications object graph model to identify a subset of the content that is related to instruments for implementing the computer product.

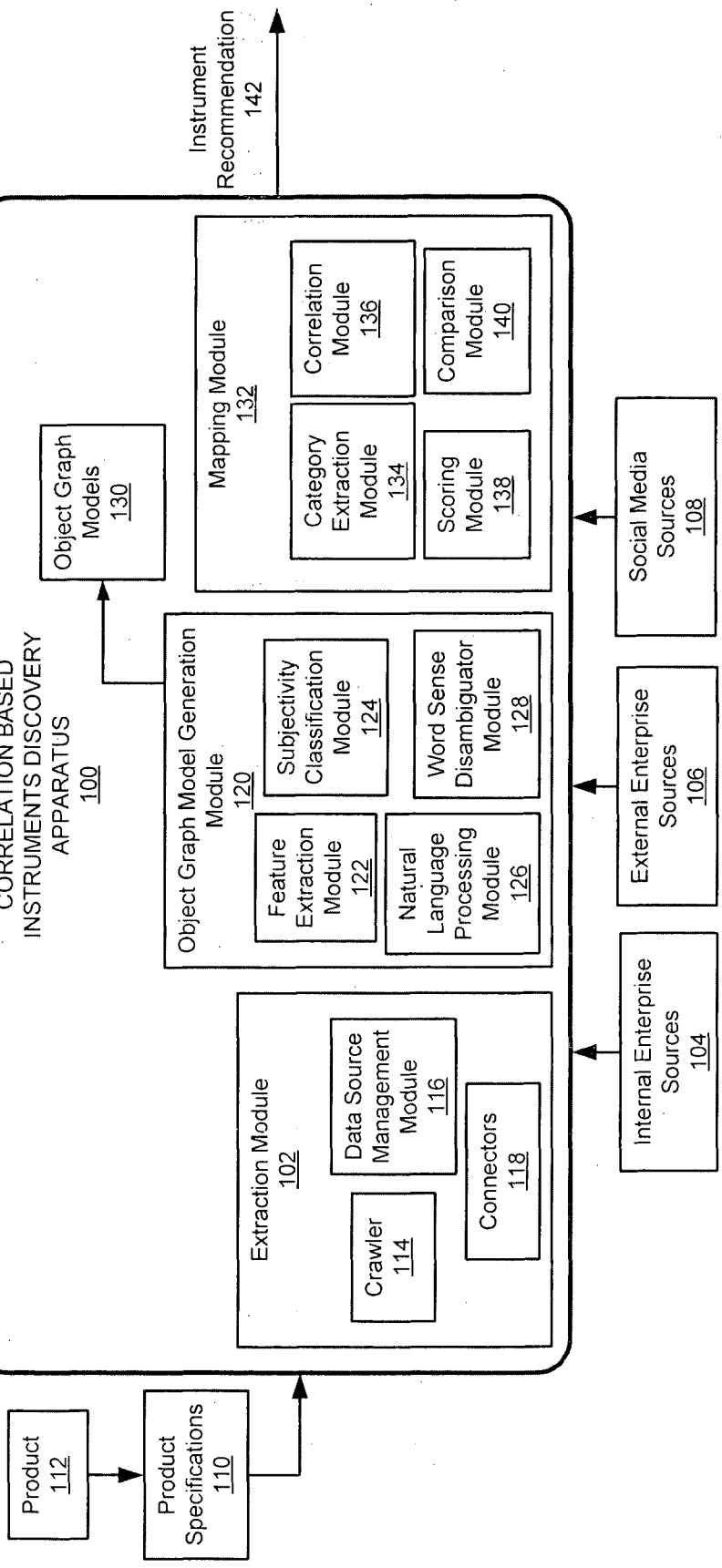

Instrument Recommendation 142
CORRELATION BASED INSTRUMENTS DISCOVERY APPARATUS 100
Object Graph Models 130
Mapping Module 132
Correlation Module 136
Comparison Module 140
Category Extraction Module 134
Scoring Module 138
Object Graph Model Generation Module 120
Subjectivity Classification Module 124
Word Sense Disambiguator Module 128
Feature Extraction Module 122
Natural Language Processing Module 126
Extraction Module 102
Data Source Management Module 116
Crawler 114
Connectors 118
Social Media Sources 108
External Enterprise Sources 106
Internal Enterprise Sources 104
Product 112
Product Specifications 110

FIG. 1

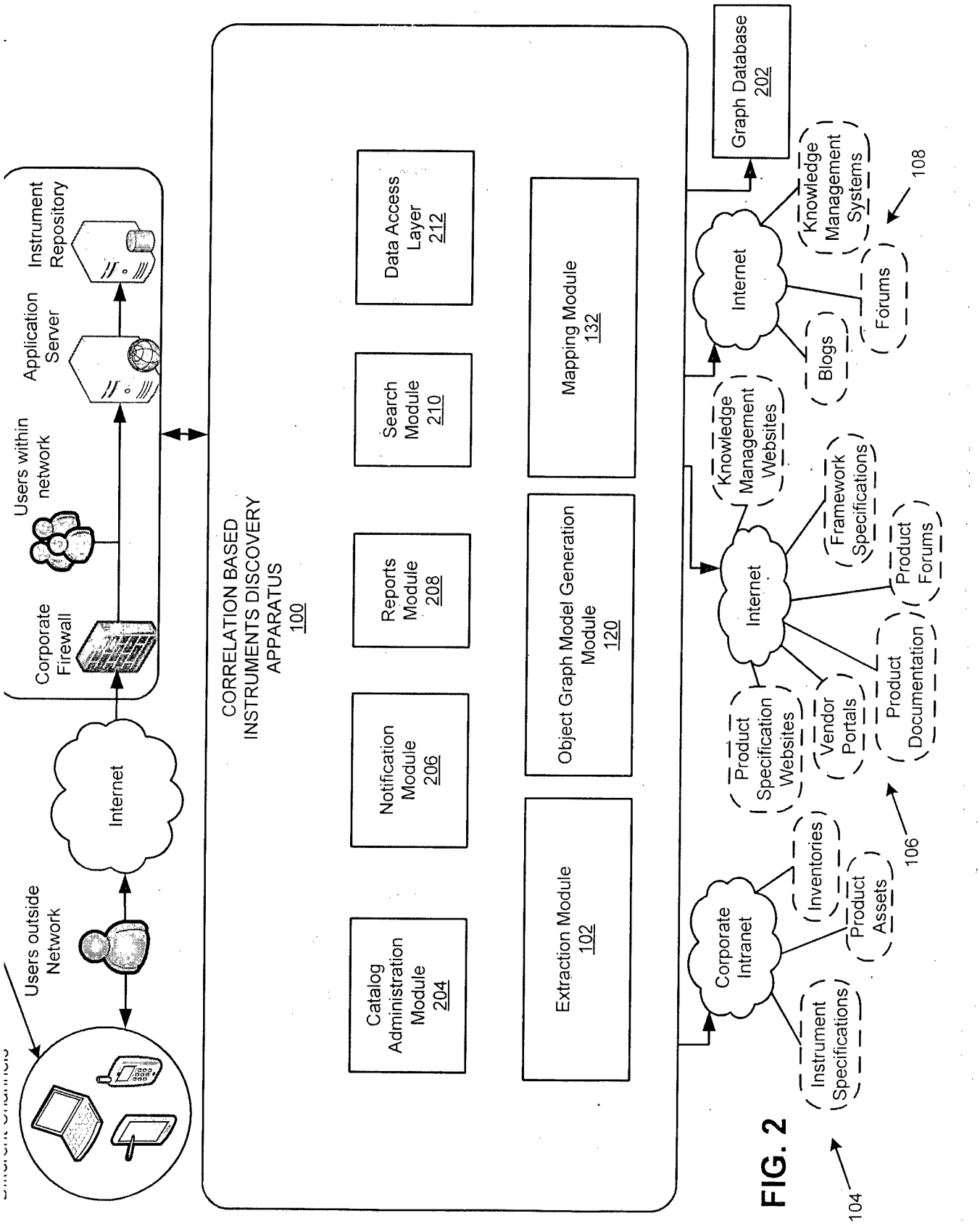

Users outside Network
Internet
Corporate Firewall
Users within network
Application Server
Instrument Repository
CORRELATION BASED INSTRUMENTS DISCOVERY APPARATUS 100
Catalog Administration Module 204
Notification Module 206
Reports Module 208
Search Module 210
Data Access Layer 212
Extraction Module 102
Object Graph Model Generation Module 120
Mapping Module 132
Graph Database 202
Corporate Intranet
Inventories
Product Assets
Instrument Specifications
Internet
Product Specification Websites
Vendor Portals
Product Documentation
Framework Specifications
Product Forums
Knowledge Management Websites
Internet
Blogs
Forums
Knowledge Management Systems
104
106
108

FIG. 2

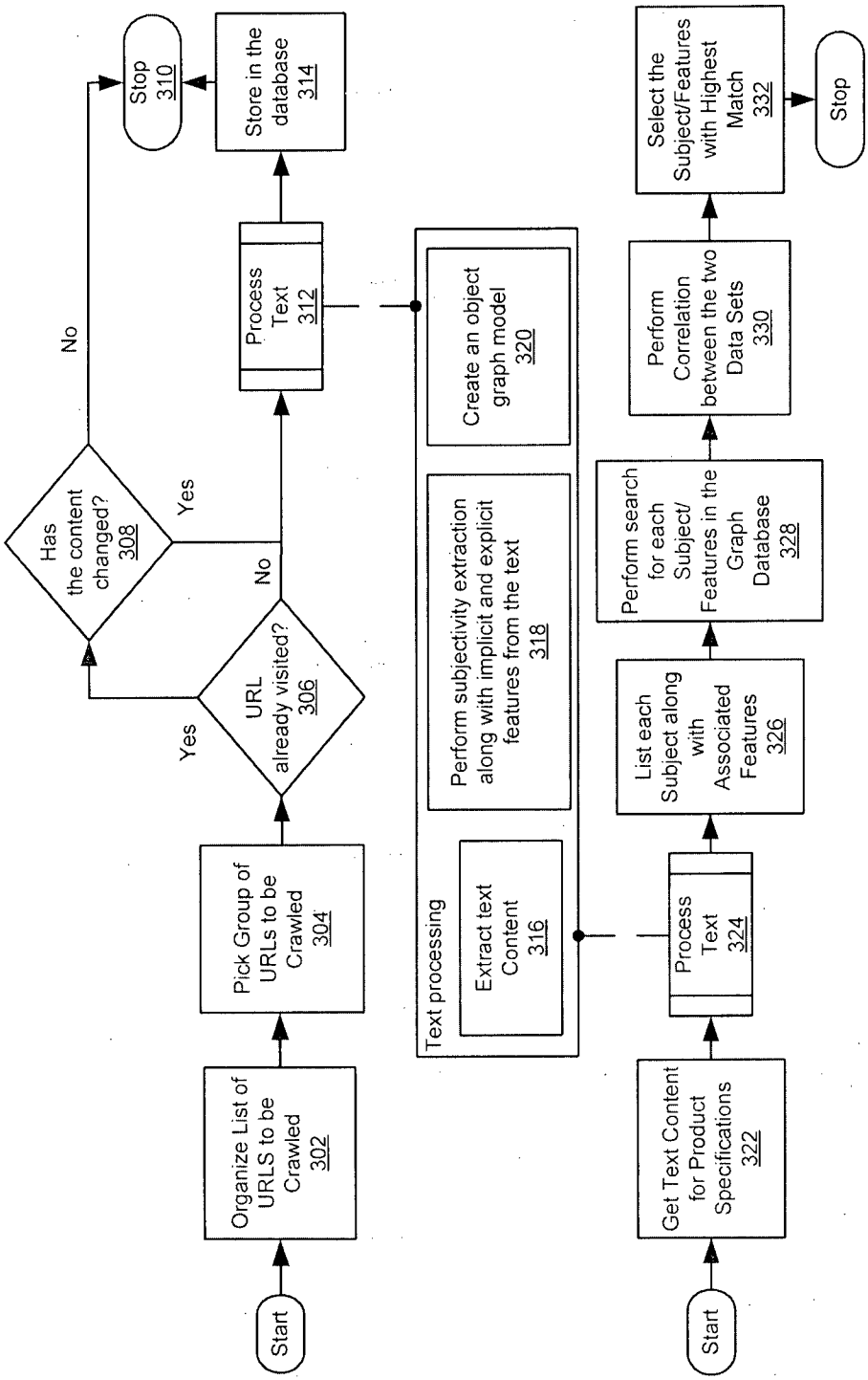

Start
Organize List of URLS to be Crawled 302
Pick Group of URLs to be Crawled 304
URL already visited? 306
Yes
No
Has the content changed? 308
Yes
No
Stop 310
Process Text 312
Store in the database 314
Text processing
Extract text Content 316
Perform subjectivity extraction along with implicit and explicit features from the text 318
Create an object graph model 320
Start
Get Text Content for Product Specifications 322
Process Text 324
List each Subject along with Associated Features 326
Perform search for each Subject/ Features in the Graph Database 328
Perform Correlation between the two Data Sets 330
Select the Subject/Features with Highest Match 332
Stop

FIG. 3

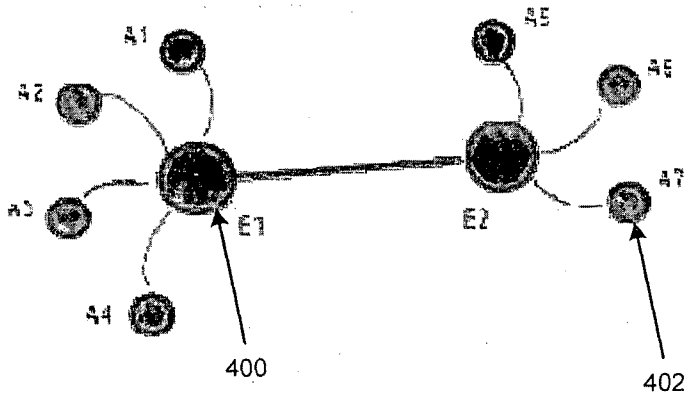
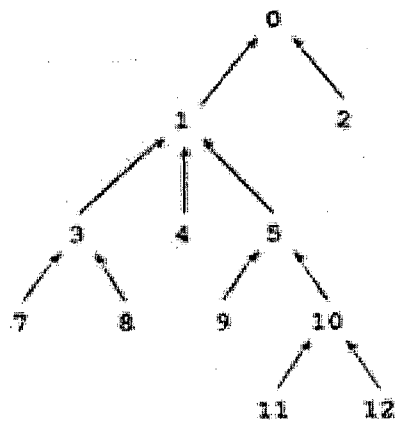


FIG. 4



The shortest path between 5 and 8 is
5 - 1 - 3 - 8

FIG. 5

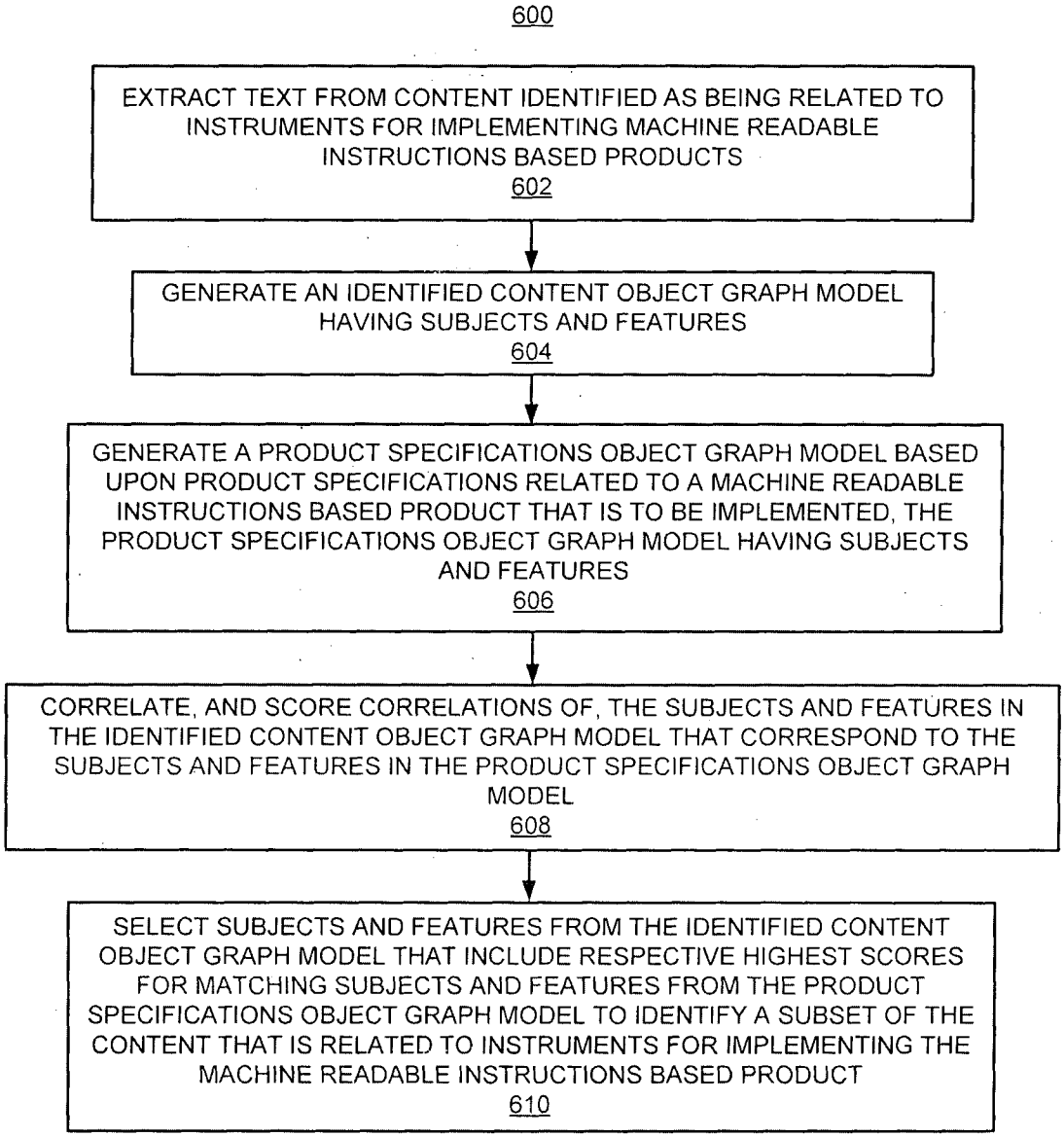

600
EXTRACT TEXT FROM CONTENT IDENTIFIED AS BEING RELATED TO INSTRUMENTS FOR IMPLEMENTING MACHINE READABLE INSTRUCTIONS BASED PRODUCTS
602
GENERATE AN IDENTIFIED CONTENT OBJECT GRAPH MODEL HAVING SUBJECTS AND FEATURES
604
GENERATE A PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL BASED UPON PRODUCT SPECIFICATIONS RELATED TO A MACHINE READABLE INSTRUCTIONS BASED PRODUCT THAT IS TO BE IMPLEMENTED, THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL HAVING SUBJECTS AND FEATURES
606
CORRELATE, AND SCORE CORRELATIONS OF, THE SUBJECTS AND FEATURES IN THE IDENTIFIED CONTENT OBJECT GRAPH MODEL THAT CORRESPOND TO THE SUBJECTS AND FEATURES IN THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL
608
SELECT SUBJECTS AND FEATURES FROM THE IDENTIFIED CONTENT OBJECT GRAPH MODEL THAT INCLUDE RESPECTIVE HIGHEST SCORES FOR MATCHING SUBJECTS AND FEATURES FROM THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL TO IDENTIFY A SUBSET OF THE CONTENT THAT IS RELATED TO INSTRUMENTS FOR IMPLEMENTING THE MACHINE READABLE INSTRUCTIONS BASED PRODUCT
610

FIG. 6

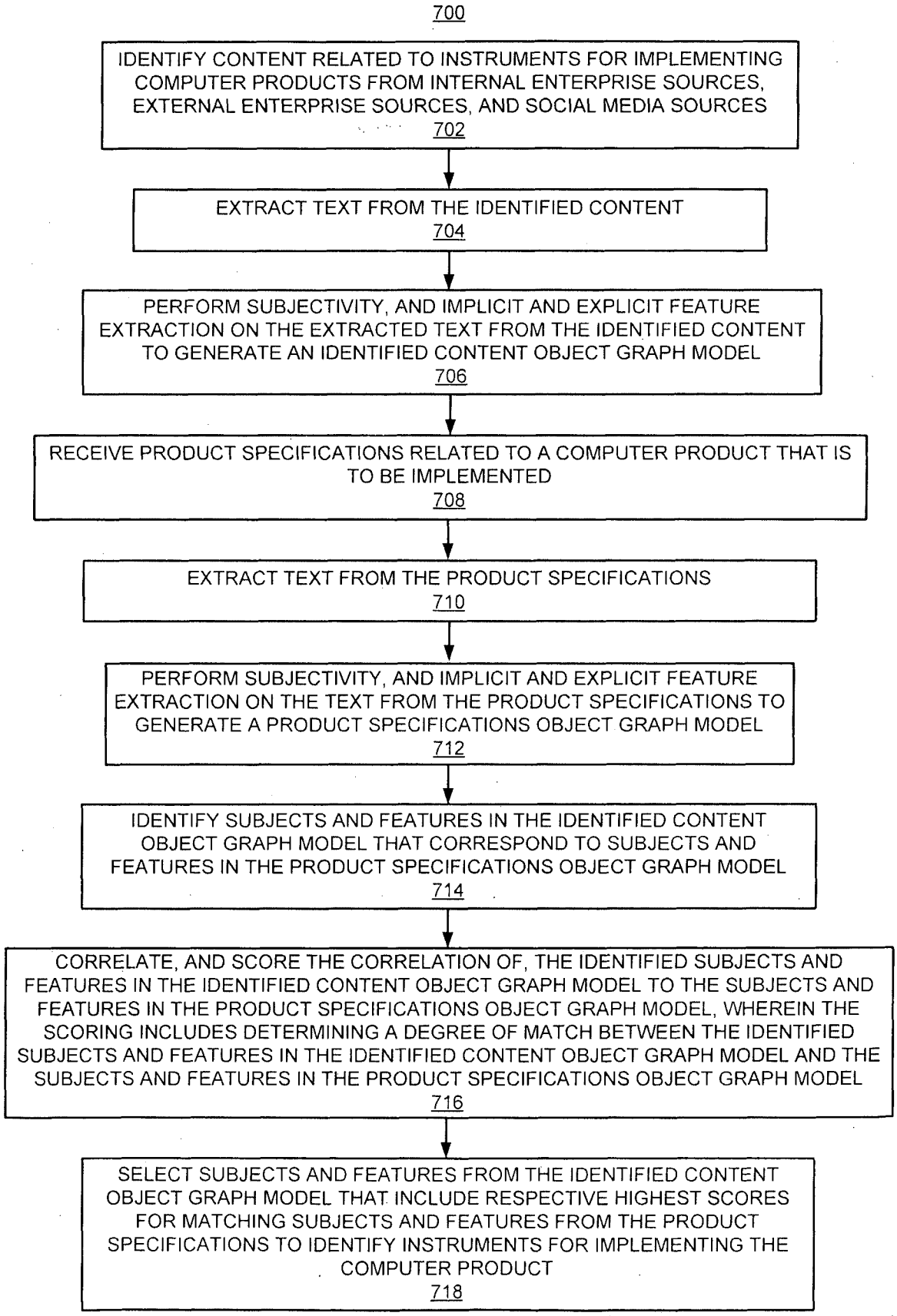

700
IDENTIFY CONTENT RELATED TO INSTRUMENTS FOR IMPLEMENTING COMPUTER PRODUCTS FROM INTERNAL ENTERPRISE SOURCES, EXTERNAL ENTERPRISE SOURCES, AND SOCIAL MEDIA SOURCES
702
EXTRACT TEXT FROM THE IDENTIFIED CONTENT
704
PERFORM SUBJECTIVITY, AND IMPLICIT AND EXPLICIT FEATURE EXTRACTION ON THE EXTRACTED TEXT FROM THE IDENTIFIED CONTENT TO GENERATE AN IDENTIFIED CONTENT OBJECT GRAPH MODEL
706
RECEIVE PRODUCT SPECIFICATIONS RELATED TO A COMPUTER PRODUCT THAT IS TO BE IMPLEMENTED
708
EXTRACT TEXT FROM THE PRODUCT SPECIFICATIONS
710
PERFORM SUBJECTIVITY, AND IMPLICIT AND EXPLICIT FEATURE EXTRACTION ON THE TEXT FROM THE PRODUCT SPECIFICATIONS TO GENERATE A PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL
712
IDENTIFY SUBJECTS AND FEATURES IN THE IDENTIFIED CONTENT OBJECT GRAPH MODEL THAT CORRESPOND TO SUBJECTS AND FEATURES IN THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL
714
CORRELATE, AND SCORE THE CORRELATION OF, THE IDENTIFIED SUBJECTS AND FEATURES IN THE IDENTIFIED CONTENT OBJECT GRAPH MODEL TO THE SUBJECTS AND FEATURES IN THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL, WHEREIN THE SCORING INCLUDES DETERMINING A DEGREE OF MATCH BETWEEN THE IDENTIFIED SUBJECTS AND FEATURES IN THE IDENTIFIED CONTENT OBJECT GRAPH MODEL AND THE SUBJECTS AND FEATURES IN THE PRODUCT SPECIFICATIONS OBJECT GRAPH MODEL
716
SELECT SUBJECTS AND FEATURES FROM THE IDENTIFIED CONTENT OBJECT GRAPH MODEL THAT INCLUDE RESPECTIVE HIGHEST SCORES FOR MATCHING SUBJECTS AND FEATURES FROM THE PRODUCT SPECIFICATIONS TO IDENTIFY INSTRUMENTS FOR IMPLEMENTING THE COMPUTER PRODUCT
718

FIG. 7

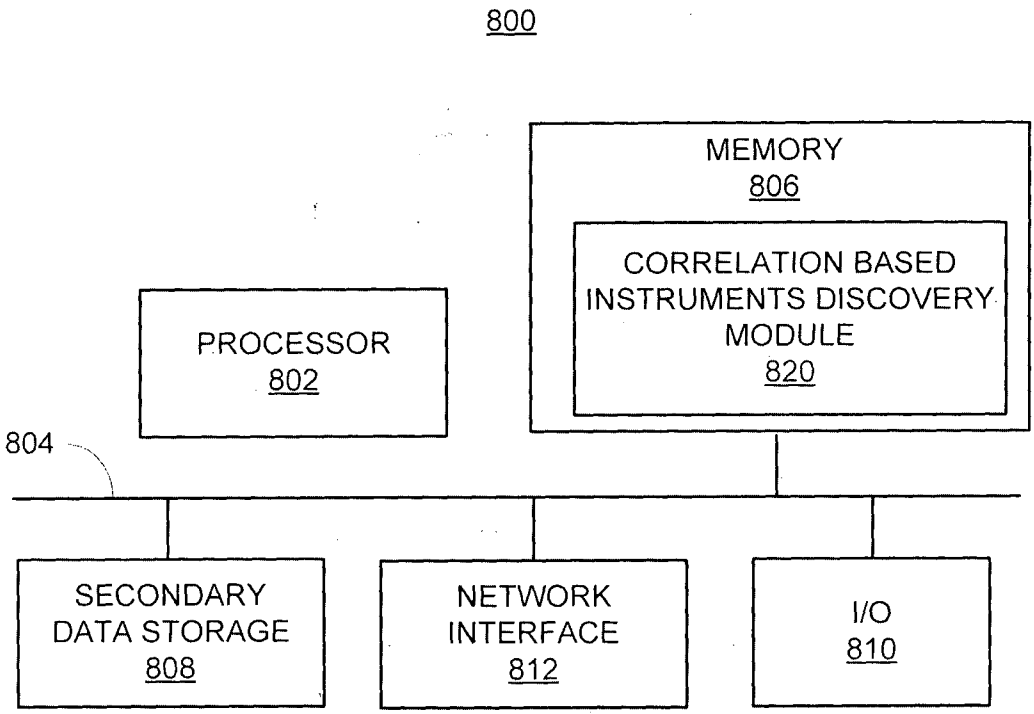

800
MEMORY
806
CORRELATION BASED INSTRUMENTS DISCOVERY MODULE
820
PROCESSOR
802
804
SECONDARY DATA STORAGE
808
NETWORK INTERFACE
812
I/O
810

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IN2014/000290

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI: correlation, search+, instructions, template?, match+, stakeholder, instrument, register+, model, extract+, text, librar+, graph, discovery, crawler, content, website, score, subject, feature?, enterprise, source?, specification?, forum, blog

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2009210419 A1 (CHITNIS UPENDRA ET AL.) 20 August 2009 (2009-08-20) claims 1-17, description, paragraphs [0029]-[0043], figures 1-5	1-15
A	US 2008189402 A1 (INTERNATIONAL BUSINESS MACHINES CORP.) 07 August 2008 (2008-08-07) the whole document	1-15
A	US 7912843 B2 (YAHOO! INC.) 22 March 2011 (2011-03-22) the whole document	1-15
A	US 2009327264 A1 (MICROSOFT CORP.) 31 December 2009 (2009-12-31) the whole document	1-15

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
18 January 2015	**28 January 2015**

Name and mailing address of the ISA/CN	Authorized officer
STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA(ISA/CN) **6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088 China**	**LI,Min**
Facsimile No. **(86-10)62019451**	Telephone No. **(86-10)62413700**

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/IN2014/000290

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2009210419 A1	20 August 2009	Non e	
US 2008189402 A1	07 August 2008	Non e	
US 7912843 B2	22 March 2011	US 8145649 B2	27 March 2012
US 2009327264 A1	31 December 2009	US 2012030200 A1	02 February 2012