



- (51) International Patent Classification:
G06Q 50/32 (2012.01)
- (21) International Application Number:
PCT/US2014/012646
- (22) International Filing Date:
23 January 2014 (23.01.2014)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
13/801,148 13 March 2013 (13.03.2013) US
- (63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 13/801,148 (CON)
Filed on 13 March 2013 (13.03.2013)
- (71) Applicant: **GOOGLE INC.** [US/US]; 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US).
- (72) Inventors: **ALAKUIJALA, Jyrki, Antero**; Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US). **LYASHUK, Alexander**; Google Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043 (US).
- (74) Agents: **SHAW, Beth, Z.** et al.; Brake Hughes Bellermann LLP, c/o CPA Global, PO Box 52050, Minneapolis, MN 55402 (US).

(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).

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: PERSONALIZED SUMMARIES FOR CONTENT

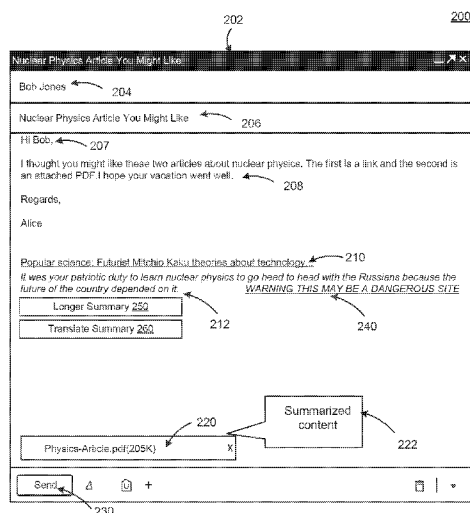


FIG. 2

(57) Abstract: A system detects a message from a sender to a recipient, the message including a reference to external content. The system accesses a user model comprising interest information about interests of the sender or the recipient. The system identifies interest content from the external content as relevant to an interest from the interest information, generates a summarized content from the external content and based on the interest content and containing only a subset of information in the external content, and modifies the message to include the summarized content in the message.

PERSONALIZED SUMMARIES FOR CONTENT

CROSS REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to, and is a continuation of, U.S. Nonprovisional Patent Application No. 13/801,148, filed on March 13, 2013, entitled “PERSONALIZED SUMMARIES FOR CONTENT”, the disclosure of which is incorporated by reference herein in its entirety.

BACKGROUND

[0002] This description relates generally to communications delivered by computer messages. Messages such as emails, social media posts, online chat discussions, or other electronic communication may include a reference to external or attached content. External content may include, for example, a PDF, a link to a webpage, a document, a video, any multimedia content, or an attachment.

SUMMARY

[0003] This disclosure includes a system and method that automatically and dynamically summarize an external content (e.g., a PDF, a webpage, a document, a video, an image, any multimedia content, an audio file, etc.) that is referenced in or by a message (e.g., attached or linked to), based on a context that includes special interests of a sender and a recipient of the message. The message may be an email message, a social media post, an interaction in an online chat discussion, or other electronic communication.

[0004] The system and method can include accessing a user model that includes interest information about a user, provided that each respective user consents to such access.

[0005] In one implementation, a method includes detecting a message from a sender to a recipient, the message including a reference to external content. The method includes accessing a user model including interest information about interests of the sender or the recipient. The method includes identifying interest content from the external content as relevant to an interest from the interest information, generating a summarized content from the external content, the summarized content being based on the interest content and containing only a subset of information in the external

content, and modifying the message to include the summarized content in the message. Other implementations of this aspect include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices.

[0006] These and other implementations can each include one or more of the following features. A system can identify common interests between the sender and the recipient using the user model, and generate the summarized content using the common interests. The system may model a relationship between the sender and the recipient as an entity-vector, and generate the summarized content using the entity-vector. A sender user model with interest information about interests of the sender and a recipient user model with interest information about interests of the recipient may be used to model the relationship. The system may enable the sender to edit the summarized content prior to delivering the message to the recipient. The system may determine a context of the message, and use the context to emphasize an entity in the summarized content. The context may include a greeting, a subject, or a text in the message. The system may determine that an entity in the summarized content relates to a topical news story from within a predetermined date range of a date of the message, and use the determination to boost the entity in the summarized content. The system may generate, based on the summarized content, a link directly to a portion of the external content, and provide the link in the message.

[0007] In one implementation, a method includes detecting a message from a sender to a recipient, the message including an attachment. The method includes accessing a user model including interest information about interests of the sender or the recipient. The method includes identifying interest content from the attachment as relevant to an interest from the interest information, generating a summarized content from the attachment, the summarized content being based on the interest content and containing only a subset of information in the attachment, and modifying the message to include the summarized content in the message. Other implementations of this aspect include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices.

[0008] In various implementations, the attachment or the external content may not be included in the body of the message, although a link (e.g., reference) to the

external content may be included in the body of the message.

[0009] One or more of the implementations of the subject matter described here may provide one or more of the following advantages. The summarized content reflects the interests of the recipient. The summarized content may also reflect common interests of the sender and the recipient.

[0010] The details of one or more implementations are set forth in the accompanying drawings and the description below. Other features will be apparent from the description and drawings, and from the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] FIG. 1 illustrates an example system.

[0012] FIG. 2 is an example user interface.

[0013] FIG. 3 is another example user interface.

[0014] FIG. 4 is a flow chart illustrating an example method of generating personalized summaries for content.

[0015] FIG. 5 is an example of a generic computer device that may be used with techniques described here.

[0016] FIG. 6 is an example of a generic computer device that may be used with techniques described here.

[0017] Like reference symbols in the various drawings indicate like elements.

DETAILED DESCRIPTION

[0018] FIG. 1 illustrates an example system 102. The system 102 can be implemented on one or more computers that communicate, for example, through a network. The system receives messages 104 with a reference to external content from a client device 106 and generates revised messages 108 with summarized content that includes extracts from the external content. Each message 104 is an electronic communication such as an e-mail, post, text message, etc. The message 104 can be, for example, in text form or in other forms, for example, in audio form or in image form.

[0019] The summarized content may be produced by an automatic process that reduces a document (e.g., the external content) using a computer program to create a summary that retains the most important points of the original document.

Technologies that can make a coherent summary may take into account variables such as length, writing style, and syntax. Personalization technology described here may enable dynamic insertion, customization, or suggestion of summarized content in any format that is relevant to an individual user, for example based on one or more of the user's behavior, implied preferences, and based on explicitly given information. The personalization may include relevant concepts which may be referred to here as "entities" (an "entity" may be a person, place, thing, or concept). Such entities may be extracted from documents using concept mining, which may involve aspects of artificial intelligence, statistics, data mining, and text mining.

[0020] The system 102 and the client device 106 can communicate through a network, for example, an intranet or the Internet. While the system 102 and the client device 106 are shown as two separate devices, in some implementations, the system 102 and the client device 106 can be the same device.

[0021] The system 102 includes an indexing engine 110 and a ranking engine 112. The indexing engine 110 maintains an index 114 for use by the system 102. The indexing engine 110 processes documents and updates index entries in the index 114, for example, using conventional or other indexing techniques.

[0022] In some implementations, the system 102 may maintain a user interest data store 120. In the instance where a user consents to the use of such data, the user interest data store 120 may include one or more entities, websites, text, search queries, demographic data, or other user interest information that may be used by system 102. In some implementations, the user interest data store 120 may classify search queries by user, by frequency, by location, and by date or time, where the user consents to the use of such data. In some implementations, the user interest data store 120 may be included as part of index 114.

[0023] A user interest engine 130 may analyze message 104 to determine if one or more terms from the message are associated with a recipient's interest, using user interest data store 120, for example, and in the cases where the users consent to the use of such data. The user interest engine 130 may use the user interest data store 120 to identify interests in the message using techniques described herein, for example as discussed in more detail below with respect to FIGs. 2-4. In some implementations, user interest engine 130 may analyze a message to locate one or more terms from the message 104 in common with an interest of a recipient of the

message 104. A summary engine 140 may use the results of that analysis to generate a summarized content for a referenced external content (e.g., an attachment or a link) referenced by or associated with the message.

[0024] The ranking engine 112 (e.g., the summary engine 140) may use the index 114 to identify information, such as portions of documents or quick-links to parts of the document related to the message 104, for example, using conventional or other information retrieval techniques. The ranking engine 112 calculates scores for the entities in the documents, for example, using one or more ranking signals. Each signal provides information about the document itself or the relationship between the document and the message. One example signal is a measure of the overall quality of the document. Another example signal is a measure of the number of times the terms of the message occur in the document or in a portion of the document. Other signals can also be used. The ranking engine 112 ranks the responsive documents or portions of documents using the scores.

[0025] Although user interest engine 130 and summary engine 140 are depicted as part of ranking engine 112, in various implementations, user interest engine 130 or summary engine 140 may be included as part of indexing engine 110, or as a separate, single engine within system 102.

[0026] As an example, a sender of a message 104 might include the subject of the message 104 as “beach vacations” and attached external content (e.g., a document about Florida) to the message 104, without any detailed explanation of the document. The recipient may not want to review or scroll through the entire document to find relevant information, especially if the document is long. The summary engine 140 may therefore provide a summarized content for the document that refers to a “Florida beach vacation package for \$1000” to be included as part of the message 104 (e.g., as message 108). The summarized content may be an extract of text from the document. Other examples of summaries generated by the system 102 are described in more detail below with respect to FIGs. 2-6.

[0027] FIG. 2 is an example user interface. As shown in FIG. 2, a system such as system 102 may present a user interface 200 including a message 202. In this example, the message 202 is an e-mail message. In various implementations, however, the message could be a social media post, a text message, or other electronic communication. The message 202 includes a recipient name 204, a title 206, a

greeting 207, text 208, a link 210, and an attachment 220. In the example shown in FIG. 2, a sender is drafting an e-mail to a recipient 204 “Bob Jones” with the title 206 “Nuclear Physics Article You Might Like”. The greeting 207 “Hi Bob,” and text 208 of the message may be entered and edited by the sender in any e-mail application, web-based email program, etc. The sender includes a link 210 to external content (a website entitled “Popular science Futurist Mitchio Kaku theories about technology...”) and an attachment 220 “Physics-Article.pdf”.

[0028] In the instance where a user consents to the use of such data, the system 102 generates personalized summaries of each of the external contents—i.e., of the website from the link 210 and of the content of attachment 220. The personalized summaries include summarized content 212 and summarized content 222. The system may gather entities from the context of the message 202, and use the entities to emphasize content in the summarized content 212 or the summarized content 222. For example, in the message 202 the title 206 includes “Nuclear Physics”. The system may use the entity “nuclear physics”, which was gathered from the context, to emphasize content related to nuclear physics in the summarized content 212.

[0029] In some implementations, in the instance where a user consents to the use of such data, the system 102 generates the personalized summaries automatically as the sender types the message 202. In various implementations, the system creates the summarized content 212 and summarized content 222 after a sender requests such summary generation. Alternatively or additionally, in cases where users consent, the system generates the summaries after the sender hits the “send” button 230. In such an implementation, the sender may not view the summarized content 212 and summarized content 222 before sending the message 202, but the recipient would see the summaries in the e-mail message 202 upon receiving the message 202. In various implementations, the sender may view the message 202 with the summarized content 212 and summarized content 222 in a sent message folder of an e-mail application after the e-mail is sent.

[0030] As shown in message 202, summarized content can take various forms. For example summarized content 222 is shown in a pop-up style window which may appear if a user hovers a cursor over attachment 220, or may appear as text within message 202. Summarized content 212 appears as text below link 210.

[0031] In some implementations, the sender may review and customize the

summarized content (e.g., change or add to) or remove the summarized content 212 and summarized content 222 before sending the message 202 (e.g., before hitting the send button 230.) For example, a sender may wish to customize the text even more specifically for the recipient by adding to the text or by removing some of the summarized content. In some implementations, a user may select a graphical user interface element 250 to opt for a longer summarized content 212 than a default length, which may be set for example as one sentence. In some implementations, a user may opt to have the system automatically convert the summarized content 212 to another language, for example by selecting a graphical user interface element 260.

[0032] In some implementations, the system may determine a scoring for some or all of the external content, and use the scoring in generating the summarized content 222. The scoring may be based on, for example, a political score, a technicality, or a spam score. For example, the system may determine that link 210 is related to a webpage whose domain is associated with some spam, and will assign a spam score to the link 210. As a result, the system may include a spam score 240 or other warning in association with the summarized content 212 or elsewhere in the message 202. As another example, the system may determine that a link is related with an acclaimed scientific publication, and may reflect that in association with the summarized content 212.

[0033] In some implementations, the system may adapt the summarized content to a vocabulary or style of the sender, with the consent of the users. In some implementations, the system may generate the language of the summarized content based on an age or skill level of the sender or recipient, with the consent of the users. For example, the summarized content 212 may be personalized differently based on native language or reading level.

[0034] In some implementations, a topicality score may change the summarized content. For example, the system may use entities that have a high topicality in news sources to increase those entities in value in summary generation. As an example, the system may determine that an entity in the external content (e.g., in link 210 or attachment 220) is related to a topical news story. For example, the system may determine that “nuclear physics” was an entity that appeared in many news stories (e.g., that day, or within the past week of message 202 being created), and may increase the entity “nuclear physics” when it generates summarized content.

The system may boost terms (e.g., score terms higher when selecting what terms to include in a summary) if those terms are associated with the entity “nuclear physics” when the system generates summarized content 212 and summarized content 222. Alternatively or additionally, the system may emphasize terms (e.g., provide in bold, colored, underlined, or italicized formats) if those terms are associated with the entity “nuclear physics” when the system provides the summarized content 212 and the summarized content 222. In some implementations, the system may attempt to present summarized content in the same order as the content is presented in the external content.

[0035] In some implementations, in instances where the users have consented to the use of such data, the system may access user interest models of interests of the sender and recipient from various sources, such as e-mail, social network groups, search history, education, age, gender, or other sources. The system may model interests of users using, for example, an entity-vector. Such models may be generated offline, for example, once a day, or once a week, etc. The system may determine common interests of multiple users, for example based on user interest models of both a sender and a recipient of message 202. Information in an individual user model is protected such that other users cannot determine its contents.

[0036] In some implementations, the system may use a function of the entity-vectors from two user interest models to determine common interests of two people, e.g., a function that depends on multiplying the two entity-vectors. Such multiplication may be done for example when message 202 is created or drafted, or when the message 202 is sent to a recipient. In some implementations, for example with group messages, the system may normalize an entity-vector for a group of users. It will be understood that there may, in some cases, be multiple senders and recipients of messages, for example in a chat discussion or social media forum. The system may gather entities from the entity-vectors and the system may use those entities to emphasize content in the summarized content of a message (such as the summarized content 212 shown in FIG. 2). In such a way, the system may generate summarized content that is of personal interest to the recipient.

[0037] It will be understood that the message 202 is an example only, and in various implementations, the message 202 may include other elements, such as a sender, other recipients, an image, other attachments or links, a date, a time, other

metadata, etc.

[0038] FIG. 3 is an example user interface. As shown in FIG. 3, a system such as system 102 may present a user interface 300 including a message 302. In this example, the message 302 is an e-mail message.

[0039] The message 302 includes a recipient name 304, a title 306, a greeting 307, text 308, and a link 310. In the example shown in FIG. 3, a sender is drafting an e-mail to a recipient 304 “Jennifer” with the title 306 “Malaria Article”. The greeting 307 “Dear Jennifer,” and text 308 of the message may be entered and edited by the sender in any e-mail application, web-based email program, etc. The sender includes a link 310 to external content, which in this example is a website with the title “The National Department of Health of Southern Africa”. The recipient may be unaware of what the link 310 references or what the sender wants the recipient to look at in the link, and the recipient may have to access the external content and then scroll through irrelevant content before finding relevant information.

[0040] In the instance where the users consent to the use of such data, the system may include one or more personalized summaries of the content, with quick-links (e.g., a modified version of a hyperlink) that bring the reader directly to a relevant place in the content. The system may gather entities from the context of the message 302 (e.g., title 306, greeting 307, text 308, or other information associated with the message 302) and emphasize the gathered entities in the generation of summarized content. In various implementations, for example as shown in FIG. 3, the system may generate multiple summaries and multiple quick links. For example, as shown in FIG. 3, the system may gather the entity “malaria” and use the entity to generate summarized content 312 “Eliminating Malaria May be Possible through effective vaccination yet to be developed” with quick link 314 “The National Department of Health of Southern Africa#Vaccines”. As another example, the system may gather the entity “medications” and use the entity to generate a summarized content 320 “Several medications are available to prevent malaria in travelers to malaria-endemic countries” and an associated quick link 322 “The National Department of Health of Southern Africa#Medications.” In some implementations, the sender may edit the summaries and the quick link 314 or quick link 322, or add other summaries or quick links to the message 302.

[0041] FIG. 4 a flow chart illustrating an example method of generating

personalized summaries for content. For convenience, the steps of the flow chart are described with reference to a system that performs the steps. The system can be, for example, system 102 described above with reference to FIG.1, component 112 described above with reference to FIG. 1, or other systems.

[0042] In the example of FIG. 4, the system detects a message from a sender to a recipient, the message including a reference to external content (402). For example, the message may be one of message 202 shown in FIG. 2 or message 302 shown in FIG. 3. The reference to external content may be a link 210 (which refers to an web page) or an attachment 220 as shown in FIG. 2, for example. The system accesses a user interest model comprising information about interests of the sender or the recipient (404). In instances where a user consents to the use of such data, the interests may be stored in the user interest data store 120 as shown in FIG. 1, in the index 114, or in one or more other data stores accessible by the system. Interests may include, for example, entities, such as geographic locations associated with the user, terms from prior search queries of the user, demographic data, social media data (such as recent purchases or self-supplied interests of close friends), and other information, in cases where a user consents to the use of such data.

[0043] The system (e.g., user interest engine 130 shown in FIG. 1) identifies interest content from the external content as relevant to an interest from the interest information (406). The interest content may include entities from an entity-vector associated with the user, for example, as described above with respect to FIG. 2. The system generates a summarized content from the external content (408). The summarized content may be based on the interest content and containing only a subset of information in the external content. The system modifies the message to include the summarized content in the message (410) for example as shown in FIG. 2 and FIG. 3.

[0044] In some implementations, entities or interests that are identified in the external content are analyzed (e.g., using user interest engine 130) to determine whether the entities or interests relate to user-identified interests. The system may identify entities using, for example, a data graph such as a knowledge graph or other technology. In some implementations, the system may utilize a map or a graph to identify entities on a webpage that may interest a particular user. For example, any entity name in a header can be identified using a knowledge graph or similar data. A header of a webpage may partially or fully match an entity name and may match

multiple entities. In various implementations, headers can be ranked by how well they match the entities, and how many of the entities are matched that are related to a user's interests.

[0045] In some implementations, the system scores potential summarized content (e.g., in steps 408 by weighting the amount of overlap via either the message's context (e.g., text, subject, greeting) or the sender or recipient's interests. In some implementations, the system scores potential summarized content based on topicality or other factors.

[0046] FIG. 5 shows an example of a generic computer device 500, which may be system 102 of FIG. 1, and which may be used with the techniques described here. Computing device 500 is intended to represent various example forms of computing devices, such as laptops, desktops, workstations, personal digital assistants, cellular telephones, smart phones, tablets, servers, and other computing devices, including wearable devices. The components shown here, their connections and relationships, and their functions, are meant to be examples only, and are not meant to limit implementations described or claimed in this document.

[0047] Computing device 500 includes a processor 502, a memory 504, a storage device 506, and expansion ports 510 connected via an interface 508. In some implementations, computing device 500 may include transceiver 546, communication interface 544, and a GPS (Global Positioning System) receiver module 548, among other components, connected via interface 508. Device 500 may communicate wirelessly through communication interface 544, which may include digital signal processing circuitry where necessary. Each of the components 502, 504, 506, 508, 510, 540, 544, 546, and 548 may be mounted on a common motherboard or in other manners as appropriate.

[0048] The processor 502 can process instructions for execution within the computing device 500, including instructions stored in the memory 504 or on the storage device 506 to display graphical information for a GUI on an external input/output device, such as display 516. Display 516 may be a monitor or a flat touchscreen display. In some implementations, multiple processors or multiple buses may be used, as appropriate, along with multiple memories and types of memory. Also, multiple computing devices 500 may be connected, with each device providing

portions of the necessary operations (e.g., as a server bank, a group of blade servers, or a multi-processor system).

[0049] The memory 504 stores information within the computing device 500. In one implementation, the memory 504 is a volatile memory unit or units. In another implementation, the memory 504 is a non-volatile memory unit or units. The memory 504 may also be another form of computer-readable medium, such as a magnetic or optical disk. In some implementations, the memory 504 may include expansion memory provided through an expansion interface.

[0050] The storage device 506 is capable of providing mass storage for the computing device 500. In one implementation, the storage device 506 may be or contain a computer-readable medium, such as a hard disk device, an optical disk device, or a tape device, a flash memory or other similar solid state memory device, or an array of devices, including devices in a storage area network or other configurations. A computer program product can be tangibly embodied in such a computer-readable medium. The computer program product may also contain instructions that, when executed, perform one or more methods, such as those described above. The computer- or machine-readable medium is a storage device such as the memory 504, the storage device 506, or memory on processor 502.

[0051] The interface 508 may be a high speed controller that manages bandwidth-intensive operations for the computing device 500 or a low speed controller that manages lower bandwidth-intensive operations, or a combination of such controllers. An external interface 540 may be provided so as to enable near area communication of device 500 with other devices. In some implementations, controller 508 may be coupled to storage device 506 and expansion port 514. The expansion port, which may include various communication ports (e.g., USB, Bluetooth, Ethernet, wireless Ethernet) may be coupled to one or more input/output devices, such as a keyboard, a pointing device, a scanner, or a networking device such as a switch or router, e.g., through a network adapter.

[0052] The computing device 500 may be implemented in a number of different forms, as shown in the figure. For example, it may be implemented as a standard server 530, or multiple times in a group of such servers. It may also be implemented as part of a rack server system. In addition, it may be implemented in a personal computer such as a laptop computer 522, or smart phone 536. An entire

system may be made up of multiple computing devices 500 communicating with each other. Other configurations are possible.

[0053] FIG. 6 shows an example of a generic computer device 600, which may be system 102 of FIG. 1, and which may be used with the techniques described here. Computing device 600 is intended to represent various example forms of large-scale data processing devices, such as servers, blade servers, datacenters, mainframes, and other large-scale computing devices. Computing device 600 may be a distributed system having multiple processors, possibly including network attached storage nodes that are interconnected by one or more communication networks. The components shown here, their connections and relationships, and their functions, are meant to be exemplary only, and are not meant to limit implementations described or claimed in this document.

[0054] Distributed computing system 600 may include any number of computing devices 680. Computing devices 680 may include a server or rack servers, mainframes, etc. communicating over a local or wide-area network, dedicated optical links, modems, bridges, routers, switches, wired or wireless networks, etc.

[0055] In some implementations, each computing device may include multiple racks. For example, computing device 680a includes multiple racks 658a – 658n. Each rack may include one or more processors, such as processors 652a-552n and 662a-562n. The processors may include data processors, network attached storage devices, and other computer controlled devices. In some implementations, one processor may operate as a master processor and control the scheduling and data distribution tasks. Processors may be interconnected through one or more rack switches 658, and one or more racks may be connected through switch 678. Switch 678 may handle communications between multiple connected computing devices 600.

[0056] Each rack may include memory, such as memory 654 and memory 664, and storage, such as 656 and 666. Storage 656 and 666 may provide mass storage and may include volatile or non-volatile storage, such as network-attached disks, floppy disks, hard disks, optical disks, tapes, flash memory or other similar solid state memory devices, or an array of devices, including devices in a storage area network or other configurations. Storage 656 or 666 may be shared between multiple processors, multiple racks, or multiple computing devices and may include a computer-readable medium storing instructions executable by one or more of the

processors. Memory 654 and 664 may include, e.g., volatile memory unit or units, a non-volatile memory unit or units, or other forms of computer-readable media, such as a magnetic or optical disks, flash memory, cache, Random Access Memory (RAM), Read Only Memory (ROM), and combinations thereof. Memory, such as memory 654 may also be shared between processors 652a-552n. Data structures, such as an index, may be stored, for example, across storage 656 and memory 654. Computing device 600 may include other components not shown, such as controllers, buses, input/output devices, communications modules, etc.

[0057] An entire system, such as system 102, may be made up of multiple computing devices 600 communicating with each other. For example, device 680a may communicate with devices 680b, 680c, and 680d, and these may collectively be known as system 102. As another example, system 102 of FIG. 1 may include one or more computing devices 600 as indexing engine 110, a separate computing device 600 as system 102, and one or more computing devices 600 as index 114.

Furthermore, some of the computing devices may be located geographically close to each other, and others may be located geographically distant. The layout of system 600 is an example only and the system may take on other layouts or configurations.

[0058] Various implementations can include implementation in one or more computer programs that are executable or interpretable on a programmable system including at least one programmable processor, which may be special or general purpose, coupled to receive data and instructions from, and to transmit data and instructions to, a storage system, at least one input device, and at least one output device.

[0059] These computer programs (also known as programs, software, software applications or code) include machine instructions for a programmable processor, and can be implemented in a high-level procedural or object-oriented programming language, or in assembly/machine language. As used herein, the terms “machine-readable medium” “computer-readable medium” refers to any non-transitory computer program product, apparatus or device (e.g., magnetic discs, optical disks, memory (including Read Access Memory), Programmable Logic Devices (PLDs)) used to provide machine instructions or data to a programmable processor.

[0060] The systems and techniques described here can be implemented in a computing system that includes a back end component, for example, as a data server,

or that includes a middleware component such as an application server), or that includes a front end component such as a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the systems and techniques described here, or any combination of such back end, middleware, or front end components. The components of the system can be interconnected by any form or medium of digital data communication such as a communication network). Examples of communication networks include a local area network ("LAN"), a wide area network ("WAN"), and the Internet.

[0061] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other.

[0062] A number of implementations have been described. Nevertheless, various modifications may be made without departing from the spirit and scope of the disclosure. In addition, the logic flows depicted in the figures do not require the particular order shown, or sequential order, to achieve desirable results. In addition, other steps may be provided, or steps may be eliminated, from the described flows, and other components may be added to, or removed from, the described systems. Accordingly, other implementations are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A method comprising:
 - detecting, by a microprocessor of a computing device, a message from a sender to a recipient, the message including a reference to external content;
 - accessing a user model comprising interest information about interests of the sender or the recipient;
 - identifying interest content from the external content as relevant to an interest from the interest information;
 - generating a summarized content from the external content and based on the interest content and containing only a subset of information in the external content; and
 - modifying the message to include the summarized content in the message.
2. The method of claim 1, wherein the identified interest content comprises interests identified as common interests between the sender and the recipient using the user model of the sender and the recipient.
3. The method of any of the preceding claims, wherein a relationship between the sender and the recipient is modeled as an entity-vector, and wherein the summarized content is generated using the entity-vector.
4. The method of any of the preceding claims, further comprising:
 - using a sender user model comprising interest information about interests of the sender and a recipient user model comprising interest information about interests of the recipient to model the relationship.
5. The method of any of the preceding claims, wherein the identifying interest content comprises:
 - enabling the sender to edit the summarized content prior to delivering the message to the recipient.

6. The method any of the preceding claims, further comprising:
 - determining a context of the message; and
 - using the context to emphasize an entity in the summarized content.
7. The method of claim 6, wherein the context includes a greeting, a subject, or a text in the message.
8. The method of any of the preceding claims further comprising:
 - determining that an entity in the summarized content relates to a topical news story from within a predetermined date range of a date of the message; and
 - using the determination to boost the entity in the summarized content.
9. The method of any of the preceding claims, further comprising:
 - generating, based on the summarized content, a link directly to a portion of the external content; and
 - providing the link in the message.
10. A system comprising:
 - at least one processor; and
 - a memory that stores instructions that, when executed by the at least one processor, cause the system to perform operations of:
 - detecting a message from a sender to a recipient, the message including a reference to external content;
 - accessing a user model comprising interest information about interests of the sender or the recipient;
 - identifying interest content from the external content as relevant to an interest from the interest information;
 - generating a summary of the external content, the summary being based on the interest content and containing only a subset of information included in the external content; and
 - modifying the message to include the summary in the message.

11. The system of claim 10, wherein the identified interest content comprises interests identified as common interests between the sender and the recipient using the user model of the sender and the recipient.

12. The system of claim 10 or claim 11, the instructions causing the system to further perform the operations of:

- determining a scoring of the summary; and
- presenting the scoring with the summary.

13. The system of claim 10, claim 11, or claim 12, the instructions causing the system to further perform the operations of:

- determining a context of the message; and
- using the context to emphasize an entity in the summary.

14. The system of claim 10, claim 11, claim 12, or claim 13, the instructions causing the system to further perform the operations of:

- enabling the sender to edit the summary prior to delivering the message to the recipient.

15. The system of claim 10, claim 11, claim 12, claim 13, or claim 14, the instructions causing the system to further perform the operations of:

- providing an option for the sender to translate the summary into another language.

16. The system of claim 10, claim 11, claim 12, claim 14, or claim 15, the instructions causing the system to further perform the operations of:

- generating, based on the summary, a link directly to a portion of the external content; and
- providing the link in the message.

17. A non-transitory computer readable medium including instructions that when executed cause a system to:
- detect a message from a sender to a recipient, the message including a reference to external content;
 - access a user model comprising interest information about interests of the sender or the recipient;
 - identify interest content from the external content as relevant to an interest from the interest information;
 - generate a summarized content from the external content, the summarized content containing only a subset of information in the external content based on the interest content; and
 - modify the message to include the summarized content in the message.
18. The non-transitory computer readable medium of claim 17, wherein a relationship between the sender and the recipient is modeled as an entity-vector, and wherein the summarized content is generated using the entity-vector.
19. The non-transitory computer readable medium of claim 17 or claim 18, the instructions further including instructions that when executed cause the system to:
- enabling the sender to edit the summarized content prior to delivering the message to the recipient.
20. The non-transitory computer readable medium of claim 17, claim 18, or claim 19, the instructions further including instructions that when executed cause the system to:
- generate, based on the summarized content, a link directly to a portion of the external content; and
 - provide the link in the message.

1/6

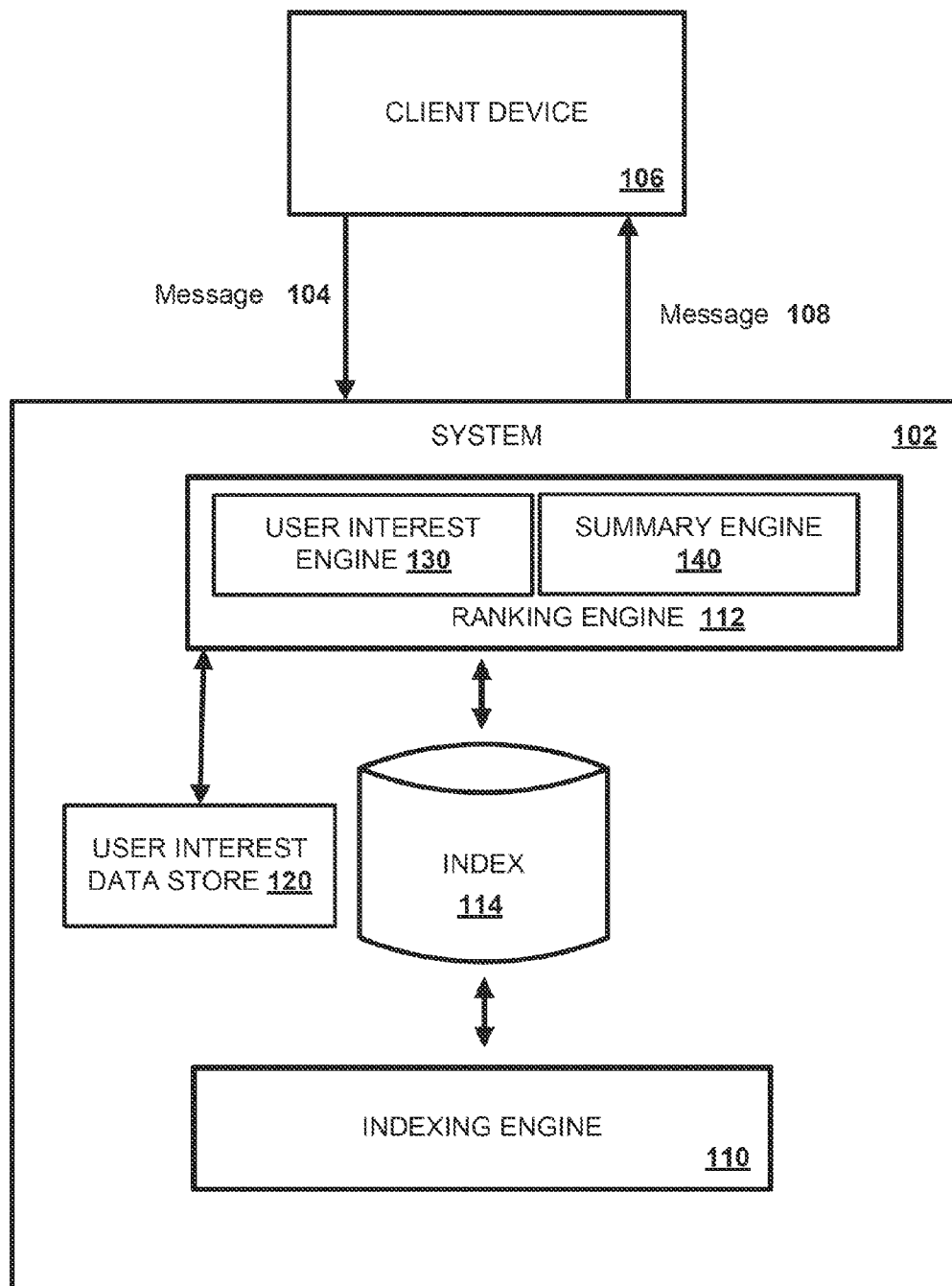


FIG. 1

2/6

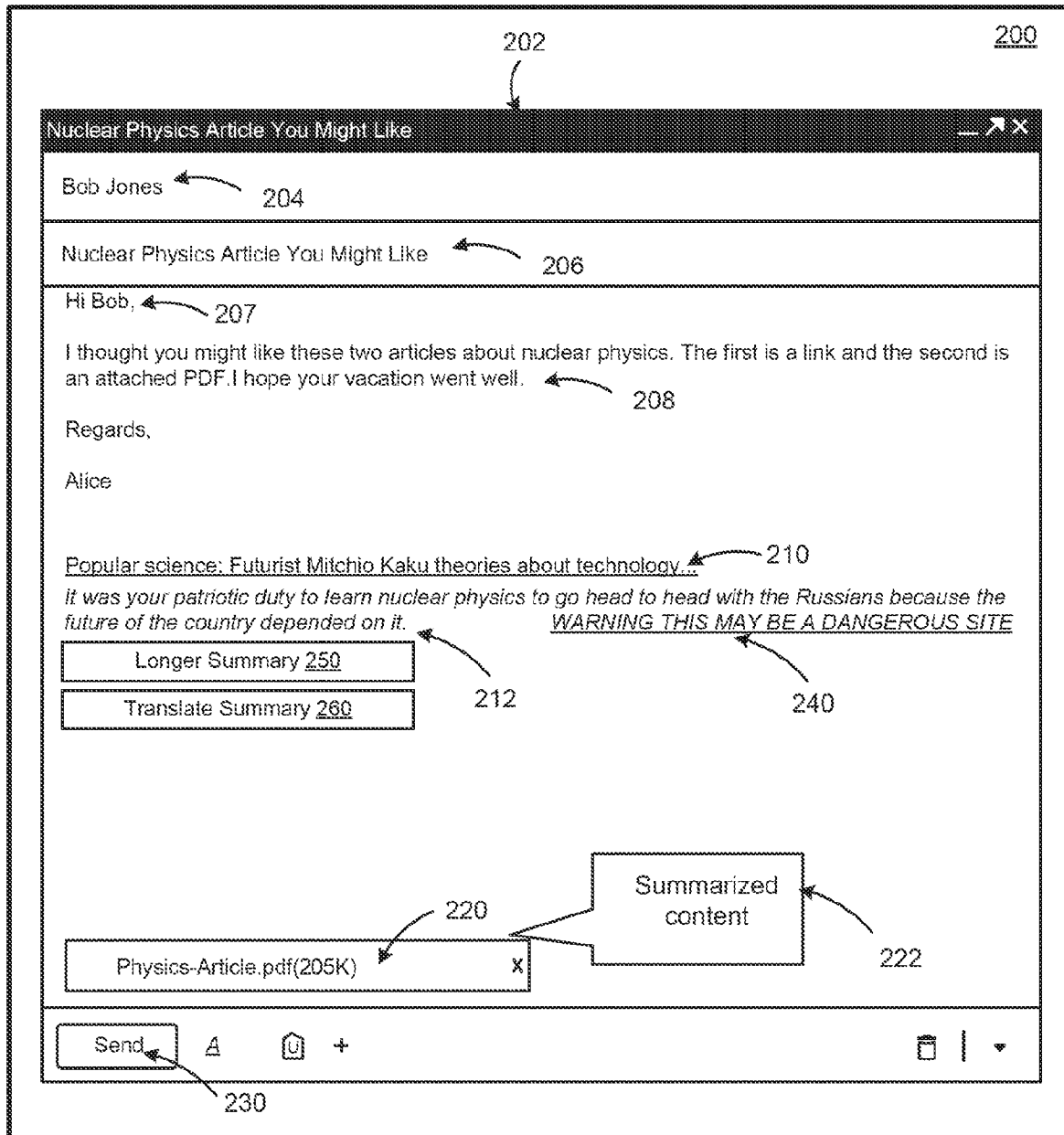


FIG. 2

3/6

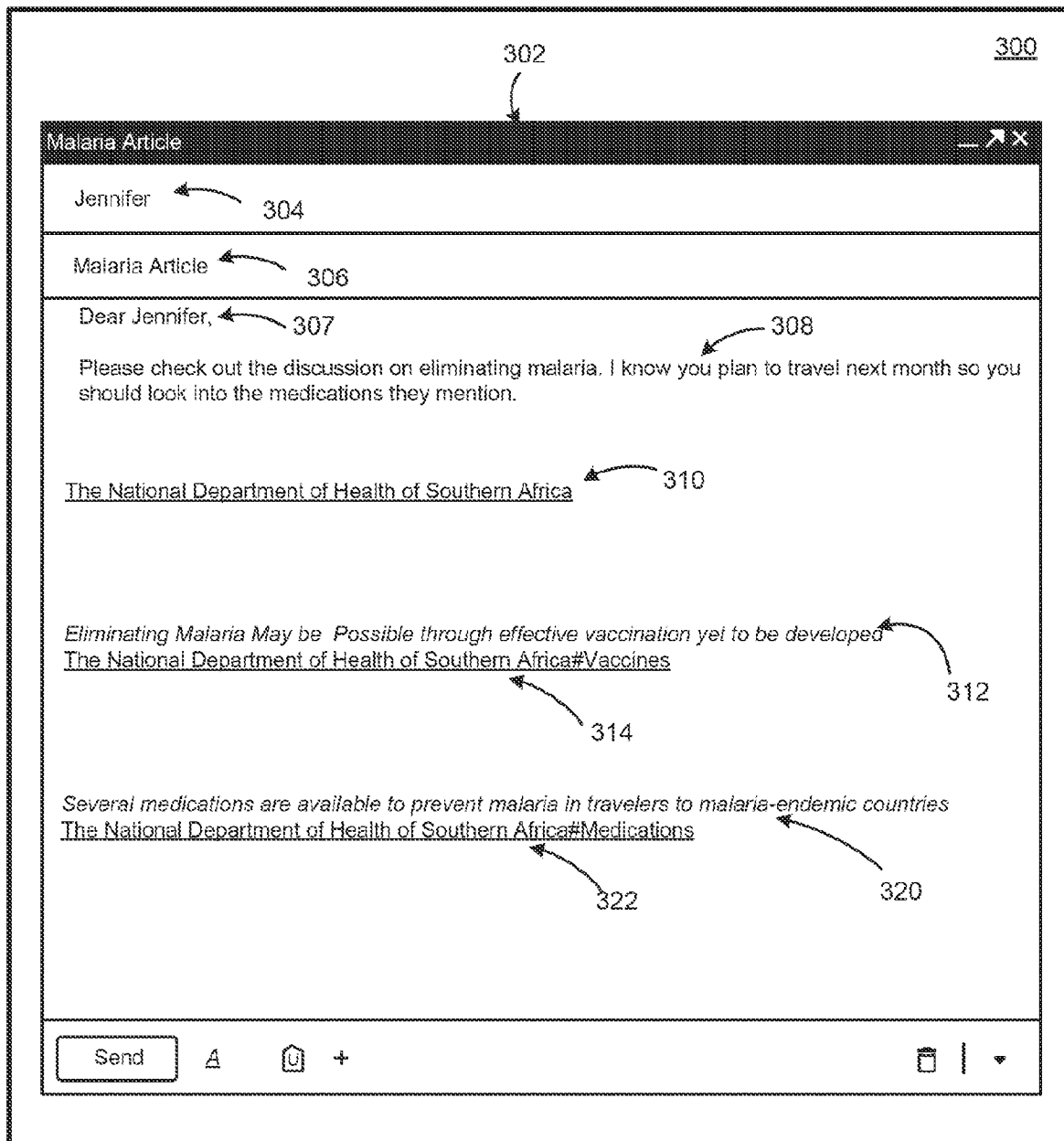


FIG. 3

4/6

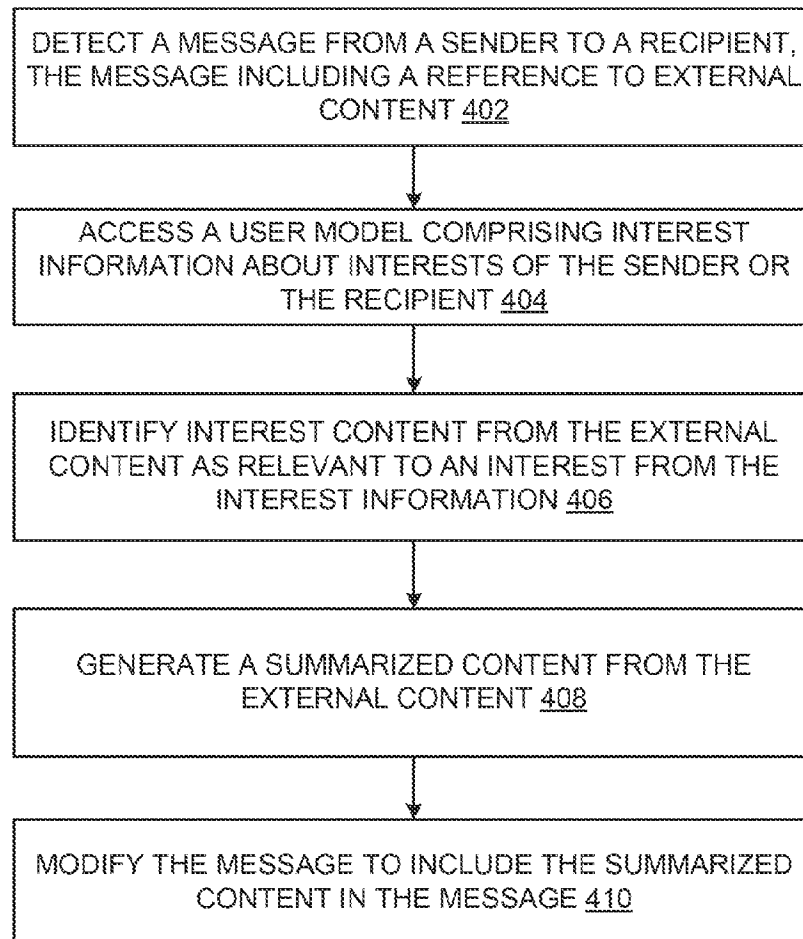


FIG. 4

5/6

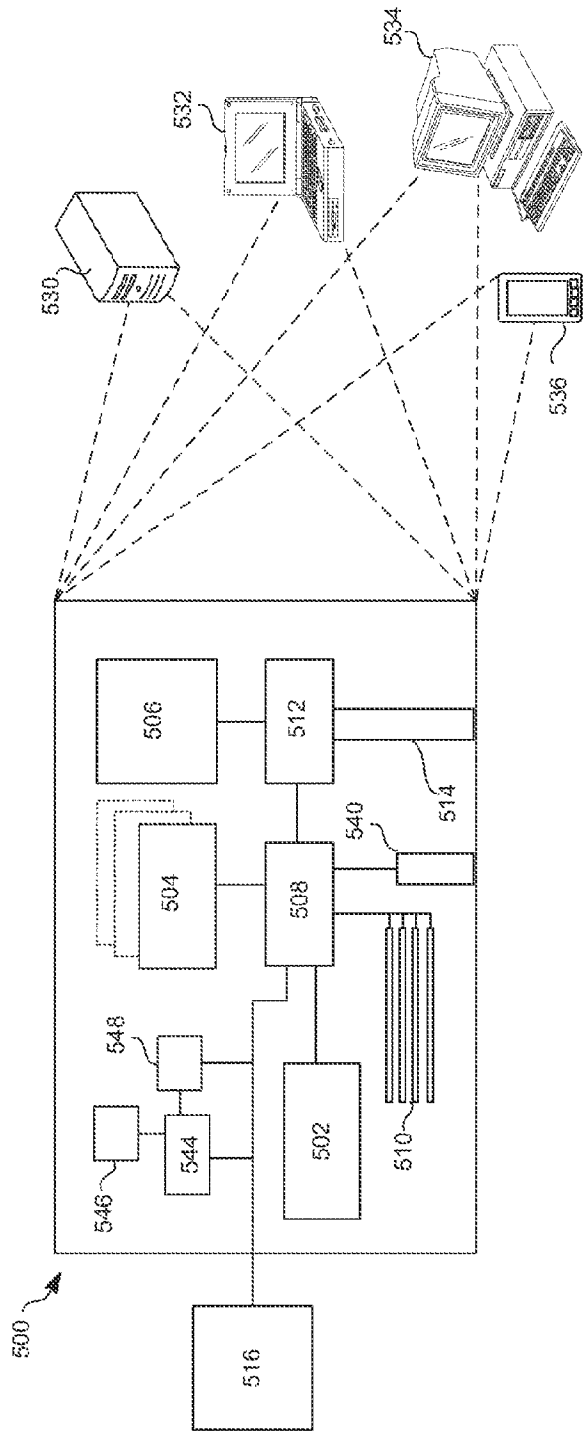


FIG. 5

6/6

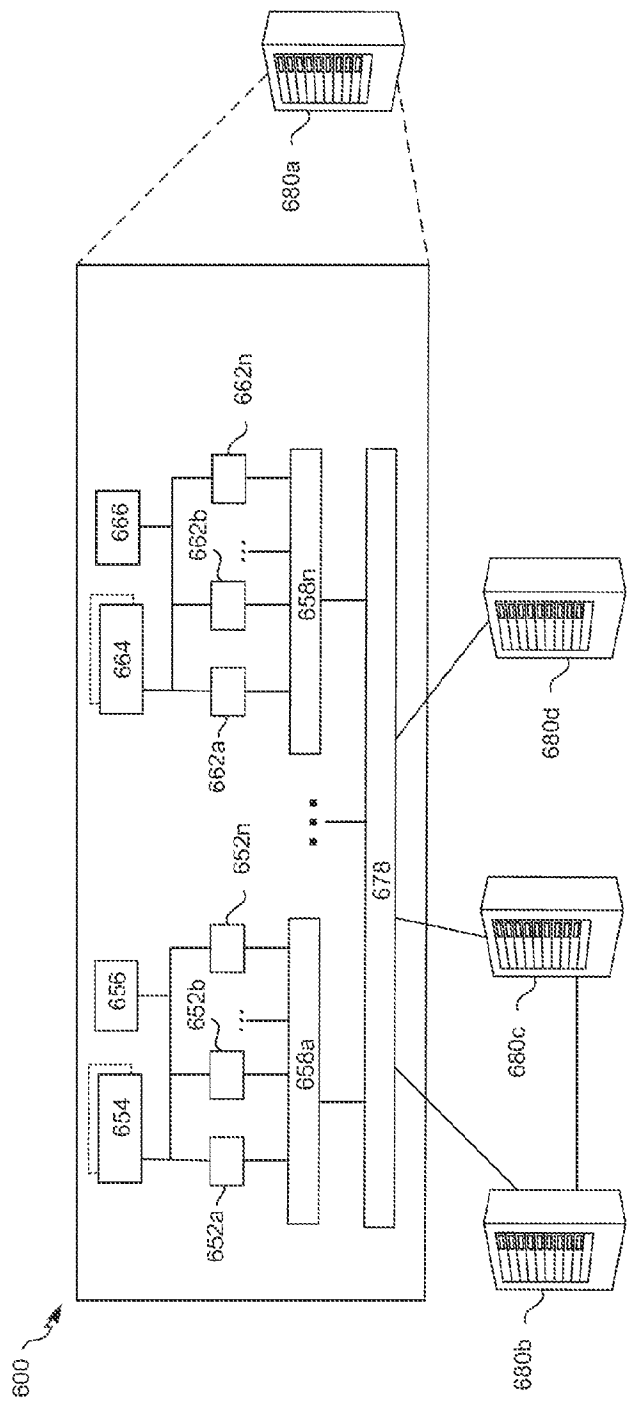


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2014/012646**A. CLASSIFICATION OF SUBJECT MATTER****G06Q 50/32(2012.01)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)

G06Q 50/32; G06F 15/16; G06N 3/00

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

Korean utility models and applications for utility models

Japanese utility models and applications for utility models

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

eKOMPASS(KIPO internal) & Keywords: message, reference, user, model, external, content, summary

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 7178099 B2 (DAVID MEYER, et al.) 13 February 2007 See claims 1-2, 4, 12, 15 and figures 1, 7A.	1-3, 10-12, 17-19
Y	US 2011-0289161 A1 (CLAIBORNE R. RANKIN, JR. et al.) 24 November 2011 See paragraphs [0005], [0072], [0134], claims 1-2, 5 and figure 29.	1-3, 10-12, 17-19
Y	US 2004-0143636 A1 (ERIC J HORVITZ et al.) 22 July 2004 See paragraph [0171], claims 1, 25 and figures 1-3.	3, 18-19
A		1-2, 10-12, 17
A	US 2008-0281927 A1 (LUCRETIA H. VANDERWENDE et al.) 13 November 2008 See abstract, paragraphs [0044]-[0045], claims 1-2 and figures 1-4.	1-3, 10-12, 17-19
A	KR 10-2013-0007244 A (NHN CORPORATION) 18 January 2013 See abstract, claims 1-2 and figures 1-3.	1-3, 10-12, 17-19

☐ 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

12 August 2014 (12.08.2014)

Date of mailing of the international search report

12 August 2014 (12.08.2014)

Name and mailing address of the ISA/KR

International Application Division
Korean Intellectual Property Office
189 Cheongsa-ro, Seo-gu, Daejeon Metropolitan City, 302-701,
Republic of Korea

Facsimile No. +82-42-472-7140

Authorized officer

OH, Eung Gie

Telephone No. +82-42-481-8744



INTERNATIONAL SEARCH REPORTInternational application No.
PCT/US2014/012646**Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)**

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.: 7
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
Claims 7 is unclear, because it refers to multiple dependent claim 6 which do not comply with PCT Rule 6.4(a).
3. ☒ Claims Nos.: 4-6, 8-9, 13-16 and 20
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of any additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/012646

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 7178099 B2	13/02/2007	US 2002-0143871 A1	03/10/2002
US 2011-0289161 A1	24/11/2011	US 2011-0288962 A1	24/11/2011
		US 2011-0289009 A1	24/11/2011
		US 2011-0289010 A1	24/11/2011
		US 2011-0289106 A1	24/11/2011
		WO 2011-146631 A2	24/11/2011
		WO 2011-146631 A3	19/01/2012
US 2004-0143636 A1	22/07/2004	AU 2001-247555 A8	24/09/2001
		AU 2001-249261 A8	24/09/2001
		AU 2001-47555 A1	24/09/2001
		AU 2001-49261 A1	24/09/2001
		AU 2005-202775 A1	19/01/2006
		AU 2005-202775 B2	10/06/2010
		CA 2511102 A1	30/12/2005
		CA 2511102 C	18/06/2013
		CN 100442268 C	10/12/2008
		CN 100505703 C	24/06/2009
		CN 100594701 C	17/03/2010
		CN 101044470 A	26/09/2007
		CN 1221912 C	05/10/2005
		CN 1304985 C	14/03/2007
		CN 1429364 A	09/07/2003
		CN 1429372 A	09/07/2003
		CN 1434406 A	06/08/2003
		CN 1551011 A	01/12/2004
		CN 1577359 A	09/02/2005
		CN 1577359 B	02/05/2012
		CN 1577359 C0	09/02/2005
		CN 1591398 A	09/03/2005
		CN 1591398 B	11/01/2012
		CN 1591398 C0	09/03/2005
		CN 1591399 A	09/03/2005
		CN 1591399 B	09/03/2011
		CN 1591399 C0	09/03/2005
		CN 1716921 A	04/01/2006
		CN 1716921 C0	04/01/2006
		CN 1734440 A	15/02/2006
		CN 1734440 B	28/04/2010
		CN 1734440 C	28/04/2010
		CN 1734440 C0	15/02/2006
		EP 1264238 A2	11/12/2002
		EP 1287444 A2	05/03/2003
		EP 1287444 B1	18/07/2007
		EP 1326189 A2	09/07/2003
		EP 1326189 A3	17/08/2005
		EP 1452988 A1	01/09/2004
		EP 1494409 A2	05/01/2005

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/012646

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		EP 1494409 A3	27/04/2005
		EP 1494409 B1	12/11/2008
		EP 1613009 A1	04/01/2006
		EP 1639428 A2	29/03/2006
		EP 1852789 A1	07/11/2007
		JP 04471671 B2	02/06/2010
		JP 04612341 B2	12/01/2011
		JP 04724659 B2	13/07/2011
		JP 04896342 B2	14/03/2012
		JP 05000746 B2	15/08/2012
		JP 05000747 B2	15/08/2012
		JP 05192624 B2	08/05/2013
		JP 05243679 B2	24/07/2013
		JP 2003-248647 A	05/09/2003
		JP 2003-527713 A	16/09/2003
		JP 2004-266815 A	24/09/2004
		JP 2004-506961 A	04/03/2004
		JP 2005-115914 A	28/04/2005
		JP 2006-018833 A	19/01/2006
		JP 2007-536600 A	13/12/2007
		JP 2010-231805 A	14/10/2010
		JP 2010-231806 A	14/10/2010
		JP 2012-074061 A	12/04/2012
		JP 2012-084160 A	26/04/2012
		KR 10-0869624 B1	21/11/2008
		KR 10-0938518 B1	25/01/2010
		KR 10-0984446 B1	29/09/2010
		KR 10-1042083 B1	16/06/2011
		KR 10-1064769 B1	14/09/2011
		KR 10-1083468 B1	16/11/2011
		KR 10-1143167 B1	10/07/2012
		KR 10-1159349 B1	25/06/2012
		KR 10-1201183 B1	13/11/2012
		KR 10-2008-0056320 A	20/06/2008
		TW 237769 A	11/08/2005
		TW 237769 B	11/08/2005
		US 06513026 B1	28/01/2003
		US 06601012 B1	29/07/2003
		US 07243130 B2	10/07/2007
		US 07249159 B1	24/07/2007
		US 07417650 B1	26/08/2008
		US 07444383 B2	28/10/2008
		US 07457879 B2	25/11/2008
		US 07565403 B2	21/07/2009
		US 07634528 B2	15/12/2009
		US 07743340 B2	22/06/2010
		US 07844666 B2	30/11/2010
		US 07975015 B2	05/07/2011
		US 08019834 B2	13/09/2011
		US 08024415 B2	20/09/2011

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/US2014/012646

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
		US 08086672 B2	27/12/2011
		US 08701027 B2	15/04/2014
		US 2002-0054117 A1	09/05/2002
		US 2002-0087649 A1	04/07/2002
		US 2003-0046421 A1	06/03/2003
		US 2004-0030753 A1	12/02/2004
		US 2004-0039786 A1	26/02/2004
		US 2004-0098462 A1	20/05/2004
		US 2004-0128359 A1	01/07/2004
		US 2004-0199663 A1	07/10/2004
		US 2004-0254998 A1	16/12/2004
		US 2007-0214228 A1	13/09/2007
		US 2007-0288932 A1	13/12/2007
		US 2009-0299934 A1	03/12/2009
		WO 2001-069387 A2	20/09/2001
		WO 2001-069387 A3	10/10/2002
		WO 2001-069432 A2	20/09/2001
		WO 2001-069432 A3	27/12/2002
		WO 2005-006130 A2	20/01/2005
		WO 2005-006130 A3	17/08/2006
US 2008-0281927 A1	13/11/2008	None	
KR 10-2013-0007244 A	18/01/2013	None	