



- (51) **International Patent Classification:**  
G06F 16/332 (2019.01)
- (21) **International Application Number:**  
PCT/US2024/018755
- (22) **International Filing Date:**  
07 March 2024 (07.03.2024)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
18/122,667 16 March 2023 (16.03.2023) US
- (71) **Applicant:** MICROSOFT TECHNOLOGY LICENSING, LLC [US/US]; One Microsoft Way, Redmond, WA 98052-6399 (US).
- (72) **Inventors:** CHAMUA, Kishor; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). VERMA, Shveta; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). SAXENA, Saksham; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). CHORDIA, Sushil, Kumar; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). VENUGOPAL, Biju; Mi-

crosoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). **AGRAWAL, Puneet**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US). **RG, Karthik**; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (US).

(74) **Agent:** CHATTERJEE, Aaron, C. et al.; Microsoft Technology Licensing, LLC, One Microsoft Way, Redmond, WA 98052-6399 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, 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, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,

(54) **Title:** MULTI-LEVEL RICH AUTOSUGGESTIONS FOR MULTI-INTENT SEARCH QUERY

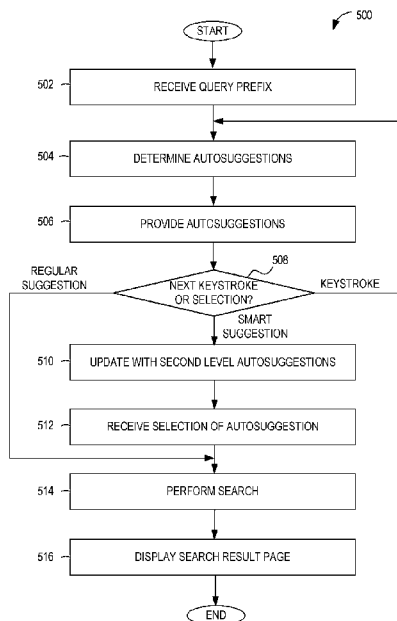


FIG. 5

(57) **Abstract:** Systems and methods are directed to providing multi-level rich autosuggestions for multi-intent search queries. The system receives a query prefix entered in a search box and accesses a database of machine-learning derived intent clusters. Based on the query prefix, a top-ranking machine-learning derived intent cluster is identified, and the search user interface is updated to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion. The smart suggestion is a first level suggestion of the top-ranking machine-learning derived intent cluster. The system receives a selection of the smart suggestion and, in response, replaces the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions. The second level autosuggestions comprise one or more intent categories that are subsets of the smart suggestion, whereby each intent category comprises a plurality of subcategory suggestions.

WO 2024/191706 A1

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, ME, 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 applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

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

**MULTI-LEVEL RICH AUTOSUGGESTIONS FOR MULTI-INTENT SEARCH QUERY****TECHNICAL FIELD**

**[0001]** The subject matter disclosed herein generally relates to search engines.  
5 Specifically, the present disclosure addresses systems and methods that determine and presents multi-level rich autosuggestions for multi-intent search queries.

**BACKGROUND**

**[0002]** Conventionally, an autosuggestion pane or box in search engines provides suggestions based on what the user has typed. The autosuggestion pane presently surfaces  
10 suggestions in textual format, where a relevance to a search query negatively correlates to the distance of a click target. Thus, it tends to take several reformulations for a user who wants to explore a topic, is looking for a specific result, or is unable to paraphrase the exact words.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**[0003]** Some embodiments are illustrated by way of example and not limitation in the  
15 figures of the accompanying drawings.

**[0004]** FIG. 1A and FIG. 1B are illustrations of different levels of autosuggestions, according to some example embodiments.

**[0005]** FIG. 2 is a diagram illustrating a network environment suitable for multi-level rich autosuggestions for multi-intent search queries, according to some example embodiments.

20 **[0006]** FIG. 3 is a diagram illustrating components of a machine learning intent system, according to some example embodiments.

**[0007]** FIG. 4 is a diagram illustrating components of a search system, according to some example embodiments.

25 **[0008]** FIG. 5 is a flowchart illustrating operations of a method for performing a search using multi-level rich autosuggestions for multi-intent search queries, according to some example embodiments.

**[0009]** FIG. 6 is a flowchart illustrating operations of a method for determining autosuggestions, according to some example embodiments.

30 **[0010]** FIG. 7 is a flowchart illustrating operations of a method for generating machine learning-derived intent clusters, according to some example embodiments.

**[0011]** FIG. 8A to FIG. 8D are example screenshots illustrating a search box and different levels of autosuggestions in an autosuggestion box.

**[0012]** FIG. 9 is an alternative example screenshot illustrating a search box and autosuggestions in an autosuggestion box.

35 **[0013]** FIG. 10 is a further example screenshot illustrating a search box and

autosuggestions in an autosuggestion box.

**[0014]** FIG. 11 is a block diagram illustrating components of a machine, according to some example embodiments, able to read instructions from a machine-storage medium and perform any one or more of the methodologies discussed herein.

5 **DETAILED DESCRIPTION**

**[0015]** The description that follows describes systems, methods, techniques, instruction sequences, and computing machine program products that illustrate example embodiments of the present subject matter. In the following description, for purposes of explanation, numerous specific details are set forth in order to provide an understanding of various embodiments of the present subject matter. It will be evident, however, to those skilled in the art, that embodiments of the present subject matter may be practiced without some or other of these specific details. Examples merely typify possible variations. Unless explicitly stated otherwise, structures (e.g., structural components, such as modules) are optional and may be combined or subdivided, and operations (e.g., in a procedure, algorithm, or other function) may vary in sequence or be  
10 combined or subdivided.

**[0016]** Example embodiments provide a network system that provides multi-level rich autosuggestions for multi-intent search queries. When a user performs a search, the user may have more than one search intent. For example, if a user is interested in Product X, the user may enter “Product X” in the search box and is presented with a variety of suggestions. However, the  
15 user may be interested in multiple results such as Product X general information and specification, variations on Product X, comparisons of Product X with similar products, reviews of Product X, and so forth. As a result, the user will need to perform multiple searches by typing each of these queries separately in a search box until the autosuggestion shows the search query and/or the user has identified the correct search query to use.

**[0017]** Conventional autosuggestion boxes do not provide an opportunity at a first level to go to different user search intents whereby a selection of an intent-based suggestion at the first level takes the user to a second level of an autosuggestion box where intent clusters that represent categories of different user intents (which are a subset of the selected intent-based suggestion) are shown. For example, if the user enters “Product X” in the search box, the user can be presented  
20 with an intent-based suggestion “Product X” (also referred to herein as a “smart suggestion”) that is visually distinguished (e.g., with an icon or in a different color) from non-intent-based autosuggestions (also referred to as “regular selections”). Selection of the intent-based suggestion triggers display of second level autosuggestions in the autosuggestion box. The second level autosuggestions can include intent categories that each comprises an intent cluster of suggestions,  
25 whereby the intent categories each represents a different user search intent that is a subset of the  
30  
35

intent-based suggestion.

**[0018]** For example, for the intent-based suggestion “Product X”, an intent category of “Product X Variants” can provide an intent cluster of autosuggestions of different versions of Product X, an intent category of “Product X vs” can provide an intent cluster of autosuggestions of similar products to compare Product X to, and an intent category of “Product X reviewed by” can provide an intent cluster of autosuggestions of different reviewers that have reviewed Product X. The user can select one of the autosuggestions within one of the intent clusters and be presented either with a corresponding result page or, if a second-level intent category and cluster is available, be presented with the second-level intent category and cluster in the autosuggestion box (e.g., a subset of autosuggestions representing different search intents of the selected autosuggestion). Thus, example embodiments provide a more visual and richer way by which search intent can be clarified and allows users to accelerate a journey of clarifying their intent. It is noted that the terms “autosuggestion” and “suggestion” are used interchangeably herein.

**[0019]** For example, FIG. 1A and FIG. 1B provide illustrations of different levels of autosuggestions that can be provided by various embodiments. As shown in FIG. 1A, a first level of autosuggestions is provided via an autosuggestion box 102 that includes a smart suggestion 104 which is visually distinguished from regular suggestions 106.

**[0020]** If a user selects the smart suggestion 104, a search box 108 and the autosuggestion box 102 are updated to display second level autosuggestions. Specifically, the smart suggestion is now shown as a new query in the search box 108 and the autosuggestions in the autosuggestion box 102 provides intent categories (or category suggestion) 110 that correspond to the smart suggestion 104. Each intent category includes a cluster of subcategory suggestions 112. If the user selects one of the subcategory suggestions, a search is performed and a result page 114 displayed.

**[0021]** FIG. 1B illustrates a further example whereby the selection of the subcategory suggestion at the second level causes third level autosuggestions to be displayed. Specifically, the selected subcategory suggestion is now shown as a new query in the search box 108 and the autosuggestions in the autosuggestion box 102 provides intent categories (or category suggestion) 116 within the selected subcategory suggestion. Each intent category includes a cluster of subcategory suggestions 118. If the user selects one of the subcategory suggestions at the third level, fourth level autosuggestions can be displayed. Any number of levels of autosuggestions can be displayed depending on the original query and subsequent selection of suggestions.

**[0022]** Thus, while conventional autosuggestion boxes only provide one level of autosuggestions whereby a selection takes the user to a result page, example embodiments surface suggestions based on predominant search intents for a given query in a structured manner

inside the autosuggestion box. This enables searching a topic more efficiently, reducing time and efforts spent on multiple queries.

**[0023]** As a result, example embodiments provide a technical solution that provides multi-level rich autosuggestions for a multi-intent search query. The technical solution includes using one more machine-learning (ML) models to determine search intents, cluster the search intents, determine relevance of different search intent clusters, and rank the search intent clusters. During a search, a database of the ranked search intent clusters is accessed, and intent clusters identified for population of an autosuggestion box. As a result, one or more of the methodologies described herein facilitate solving the technical problem of efficiently providing search results to a user that has multiple search intents. By using example embodiments, a user easily access relevant search results without having to perform an extraneous number of searches.

**[0024]** FIG. 2 is a diagram illustrating a network environment 200 suitable for multi-level rich autosuggestions for multi-intent search queries, in accordance with example embodiments. A network system 202 is communicatively coupled, via a network 204, to a plurality of client devices 206. In some embodiments, the network system 202 is associated with a single entity (e.g., software provider). The network system 202 comprises a subscriber system 208 such as, for example, Microsoft 365, a machine learning (ML) intent system 210, and a search system 212.

**[0025]** The subscriber system 208 can comprise a cloud-based software-as-a-service (SaaS) system that provides a plurality of applications that generate, edit, and display documents that can be used (e.g., accessed, analyzed) by the client devices 206 and searched by the search system 212. For example, the subscriber system 208 can provide a word processing application that manages text documents, a spreadsheet application that manages spreadsheets, a presentation application that manages presentation slides, and a meeting application that can generate transcriptions of virtual meetings. In some embodiments, the components of the subscriber system 208 works with a corresponding client application 214 at the client device 206.

**[0026]** The ML intent system 210 is configured to generate machine-learning derived intent clusters that are groupings of queries based on user search intents that can be used to derive different levels of autosuggestions by the search system 212 during a search. The ML intent system 210 comprises or accesses one or more machine-trained models that clusters queries based on search intent, determines relevance of the clusters, and determines ranking for the clusters. The ML intent system 210 will be discussed in more detail in connection with FIG. 3 below.

**[0027]** The search system 212 is a component of the network system 202 that performs searches both on the Internet and within the network system 202 (e.g., within the subscription system 208). Example embodiments use the search system 212 to search for results based on

what a user types into a search box or selects from an autosuggestion box. The search system 212, during a search, communicates with the ML intent system 210 to determine top ranking intent-based autosuggestions that correspond with what a user has typed (or started to type) in the search box. The search system 212 will be discussed in more detail in connection with FIG. 4 below.

**[0028]** Each client device 206 is a device of a user of the network system 202. The client devices 206 may comprise, but are not limited to, a smartphone, tablet, laptop, multi-processor systems, microprocessor-based or programmable consumer electronics, game consoles, set-top boxes, a server, or any other communication device that can perform operations with respect to the network system 202 via the network 204. The operations can include, for example, generating, editing, and viewing documents associated with the subscription system 208 and/or performing searches using the search system 212.

**[0029]** In example embodiments, each client device 206 comprises one or more client applications 214 that communicates with the network system 202 for added functionality. For example, the client application 214 may be a local version of an application or component of the subscription system 208 or the search system 212.

**[0030]** Depending on the form of the client devices 206, any of a variety of types of connections and networks 204 may be used. For example, the connection may be Code Division Multiple Access (CDMA) connection, a Global System for Mobile communications (GSM) connection, or another type of cellular connection. Such a connection may implement any of a variety of types of data transfer technology, such as Single Carrier Radio Transmission Technology (1xRTT), Evolution-Data Optimized (EVDO) technology, General Packet Radio Service (GPRS) technology, Enhanced Data rates for GSM Evolution (EDGE) technology, or other data transfer technology (e.g., fourth generation wireless, 4G networks, 5G networks). When such technology is employed, the network 204 includes a cellular network that has a plurality of cell sites of overlapping geographic coverage, interconnected by cellular telephone exchanges. These cellular telephone exchanges are coupled to a network backbone (e.g., the public switched telephone network (PSTN), a packet-switched data network, or other types of networks).

**[0031]** In another example, the connection to the network 204 is a Wireless Fidelity (Wi-Fi, IEEE 802.11x type) connection, a Worldwide Interoperability for Microwave Access (WiMAX) connection, or another type of wireless data connection. In some embodiments, the network 104 includes one or more wireless access points coupled to a local area network (LAN), a wide area network (WAN), the Internet, or another packet-switched data network. In yet another example, the connection to the network 204 is a wired connection (e.g., an Ethernet link) and the

network 204 is a LAN, a WAN, the Internet, or another packet-switched data network. Accordingly, a variety of different configurations are expressly contemplated

**[0032]** In example embodiments, any of the systems, devices, or services (collectively referred to as “components”) shown in, or associated with, FIG. 2 may be, include, or otherwise be implemented in a special-purpose (e.g., specialized or otherwise non-generic) computer that has been modified (e.g., configured or programmed by software, such as one or more software modules of an application, operating system, firmware, middleware, or other program) to perform one or more of the functions described herein for that system, device, or machine. For example, a special-purpose computer system able to implement any one or more of the methodologies described herein is discussed below with respect to FIG. 11, and such a special-purpose computer is a means for performing any one or more of the methodologies discussed herein. Within the technical field of such special-purpose computers, a special-purpose computer that has been modified by the structures discussed herein to perform the functions discussed herein is technically improved compared to other special-purpose computers that lack the structures discussed herein or are otherwise unable to perform the functions discussed herein. Accordingly, a special-purpose machine configured according to the systems and methods discussed herein provides an improvement to the technology of similar special-purpose machines.

**[0033]** Moreover, any of the components illustrated in FIG. 2 or their functions may be combined, or the functions described herein for any single component may be subdivided among multiple components. Additionally, any number of client devices 206 may be embodied within the network environment 200. While only a single network system 202 is shown, alternative embodiments contemplate having more than one network system 202 to perform the operations discussed herein (e.g., each localized to a particular region).

**[0034]** FIG. 3 is a diagram illustrating components of the machine learning intent system 210, according to some example embodiments. The machine learning intent system 210 is configured to generate a database of machine-learning derived intent clusters. To enable these operations, the machine learning intent system 210 comprises or accesses a search log data storage 302, a prefix log storage 304, a model-based suggestion generator 306, a web index log storage 308, a knowledge data storage 310, a grouping engine 314, and an intent data storage 316 that are configured to communicate with each other (e.g., via a bus, shared memory, or a switch. In example embodiments, each of the search log data storage 302, prefix log storage 304, model-based suggestion generator 306, web index log storage 308, and knowledge data storage 310 provides data to the grouping engine 312, which determines top search intent groupings for each given prefix using n-gram based clustering. Alternative embodiments may comprise more or less components, combine the functions of some these components into a single component, or



making some components optional.

**[0035]** The search log data storage 302 comprises search logs. The search logs refer to data collected by the search system 212 on what users have searched for in the past. The search log information may be available across multiple geographies and user queries. In example  
5 embodiments, the search logs are used to come up with query suggestions by analyzing patterns in the search terms. For instance, the search logs indicate what queries are performed within a same session and these queries may indicate intent. For example, if many users search for “best running shoes” and later also search for “Nike running shoes” in the same session, the grouping engine 312 can suggest “Nike” as a possible way to filter down the intent. Thus, the information  
10 from the search logs is used for intent grouping by the grouping engine 312.

**[0036]** The prefix log storage 304 is a tree-based storage where every node is a possible character that the user enters in a search box or search bar. Thus, the prefix log storage 304 provides efficient prefix matching. The prefix log storage 304 can be used to associate given prefixes with top search intent groupings by the grouping engine 314 in some embodiments.  
15 Additionally, when a user starts to type a query, the search system 212 can use the prefix log storage 304 to quickly find all the search terms that start with the same prefix. While the prefix log storage 304 is shown as being part of the machine learning intent system 210, alternatively, the prefix log storage 304 can be located elsewhere in the network system 202 (e.g., in the search system 212).

**[0037]** The model-based suggestion generator 306 generates suggestions that go beyond those that can be derived from the search logs. In example embodiments, the model-based suggestion generator 306 comprises one or more machine-learning language models (e.g., artificial intelligence models) that are trained to understand and generate natural language and, more specifically, generate new suggestion candidates beyond those derived from the search logs.  
20 In one embodiment, the model-based suggestion generator 306 uses the information from the search logs data storage 302 to train the language models. Then the trained language models are used to generate new, related suggestions.

**[0038]** The web index log storage 308 provides information from web documents while the knowledge data storage 310 provides a knowledge base. In addition to what is seen in the  
30 search logs and derived from the language models, it is important to ground the data on knowledge. The search logs alone may not be sufficient since it does not capture the entirety of what a user could query on the web and does not capture the entirety of the knowledge base. Thus, structured pieces of information from both web documents (e.g., from the web index log storage 308) and the knowledge data storage 310 are used by the grouping engine 314 to provide  
35 a refined set of intent. For example, by identifying the possible entities from a query, the

knowledge data can be used, by the grouping engine 314, to map each entity to a predefined list of intents ranked by usage and importance.

**[0039]** Once data is collected or accessed from the search logs data storage 302 (e.g., search logs), the model-based suggestion generator 306 (e.g., language model-based suggestions), the web index log storage 308 (e.g., web documents), and knowledge data storage 310 (e.g., structured knowledge base), a clustering module 318 of the grouping engine 314 uses the data to group appropriate intents in a stack. In example embodiments, the clustering module 318 uses a machine learning model (or clustering algorithm) to group similar intents for a given query input (or prefix) together. The choice of a machine-learning model or clustering algorithm used can be dependent on specific characteristics of the data and the desired level or granularity. Example clustering algorithms include K-means clustering, mean-shift clustering, and density-based spatial clustering of applications with noise (e.g., DBSCAN).

**[0040]** A ranking module 320 of the grouping engine 314 is configured to determine a relevance and ranking of each intent cluster for a given query input. In some embodiments, the ranking module 320 applies a relevance model (e.g., a relevance algorithm) to the clusters generated by the clustering module 318. The relevance model is trained from the search logs, web documents, and knowledge base. The ranking module 320 then determines a ranking of each cluster that is determined to be relevant (e.g., exceeds a relevance threshold). In some embodiments, the ranking module 320 applies a ranking model (e.g., a ranking algorithm) to the intent clusters. The ranking model is trained from usage of past searches. For example, the ranking model computes a probability or likelihood that a user is going to click on a given intent based on a query input (or prefix).

**[0041]** The ranking module 320 ensures that the most relevant and useful results are returned as autosuggestions first, while less relevant results may be returned later (or not at all). This increases the chance of a user finding the information they are looking for quickly and efficiently. In one embodiment, a training module 322 uses supervised and semi-supervised learning algorithms to train one or more machine-learning ranking models based on datasets of search queries, associated results, and other relevant information, such as user feedback, click-through data, and relevance labels. The training module 322 can retrain the ranking model periodically. For example, the ranking model can be updated weekly with rolling two-years' worth of logs. In alternative embodiments, the training module 322 may be located outside of the grouping engine 314 and/or elsewhere in the network system 202.

**[0042]** In some embodiments, the grouping engine 314 includes a pointer module 324 that identifies and assigns imagery (e.g., images, videos) to each cluster. The imagery can be determined based on data from web documents (e.g., the web index log storage 308) and

knowledge base (e.g., knowledge data storage 310). The imagery may be stored in an image database (not shown). The pointer module 324 assigns a pointer corresponding to each image/video in the image database to the intent clusters (e.g., for each intent category and subcategory suggestion). In alternative embodiments, the pointer module 324 may be located  
5 outside of the grouping engine 314 and/or elsewhere in the network system 202.

**[0043]** Once the grouping of intents is performed and ranked, it is stored to the intent data storage 316. The intent data storage 316 stores the grouping of intents in a manner that is consumable real-time in a query suggestion space. Additional information can be stored around the intent groupings with an ability to query items on multiple pivots (e.g., clustering multiple  
10 queries into an intent provides a structured pivot for a user to traverse the possible queries the user may want to ask the search engine). For example, the corresponding image points are stored with the grouping of intents in the intent data storage 316. Other additional information can include pointers to kinds of suggestions from each intent group, a semantic vector representation of the intent group that is used at runtime/application time to rank the intent group with respect  
15 to the original request, and/or visual cues to the intent group with images and alternative namings to make it easily consumable. This provides efficient storage and performant enterprise-to-enterprise results in various embodiments.

**[0044]** FIG. 4 is a diagram illustrating components of the search system 212, according to some example embodiments. The search system 212 is configured to receive a search query and  
20 perform a search based on the search query. Further still, the search system 212 can receive a partial input of one or more keystrokes (e.g., a prefix) and provide autosuggestions including smart suggestions and multiple levels of intent categories and groupings of subcategory suggestions. To enable these operations, the search system 212 comprises a user interface (UI) engine 402, an autosuggestion engine 404, and a search engine 406 that are configured to  
25 communicate with each other (e.g., via a bus, shared memory, or a switch).

**[0045]** The UI engine 402 is configured to provide a search user interface including a search box and an autosuggestion box. Additionally, the UI engine 402 presents a search result page in response to a search performed by the search system 212. In example embodiments, the UI module detects inputs typed into the search box. The inputs can include a query prefix as a  
30 user is starting to type their query. With each keystroke, the UI engine 402 triggers the autosuggestion engine 404 to provide autosuggestions that may be the complete query that the user intends to type. When autosuggestions are available, the UI engine 402 updates the autosuggestion box to include the autosuggestions.

**[0046]** If a user selects a smart suggestion, the UI engine 402 displays a second level of  
35 autosuggestions. If the user selects a subcategory suggestion that has a further level of

autosuggestions, the UI engine 402 displays a next level of autosuggestions. Finally, if the user selects a regular suggestion (e.g., not a smart suggestion) or a subcategory suggestion that does not have a further level, the UI engine 402 displays a corresponding search result page.

**[0047]** The autosuggestion engine 404 is configured to determine and provide autosuggestions for a query input. In example embodiments, the autosuggestion engine 404 receives the query prefix from the UI engine 402 and provides a suggestions response (e.g., a JSON suggestions response) to the UI engine 402. The suggestions response includes the autosuggestion including points to images/videos. To enable these operations, the autosuggestion engine 404 comprises a prefix module 408, an intent module 410, and an image module 412.

**[0048]** The prefix module 408 is configured to identify queries based on the query prefix received from the UI engine 402. In example embodiments, the prefix module 408 accesses the prefix log storage 304 to determine all search terms that start with the same prefix as the query prefix. As discussed, the prefix log storage 304 is a tree-based storage where every node is a possible character that the user enters in the search box. Thus, the prefix log storage 304 is used for efficient prefix matching by the prefix module 408.

**[0049]** The intent module 410 is configured to determine machine-learning derives intent clusters or groupings that correspond to the search terms identified by the prefix module 408. In example embodiments, the intent module 410 accesses the intent data storage 316 and identifies a top (machine-learning derived) intent cluster based on the search terms. In selecting the top intent cluster, the intent module 410 may identify the intent cluster having the highest probability from all the matching intent clusters.

**[0050]** The image module 412 is configured to obtain and return images/videos that are associated with the top intent cluster identified by the intent module 410. In example embodiments, the intent cluster includes pointers to images/videos. Using the pointers, the image module 412 accesses an image data storage 414 that comprises an index of images and videos and obtains the images and/or videos that correspond to the pointers. The intent cluster and the images and/or videos are then returned to the UI engine 402 as the suggestions response. In some embodiments, pointers may also be used to point to other content such as audio content, additional textual information to describe the intent, metadata on the intent (e.g., trending, new) to help the user traverse the autosuggestion. Additionally, some embodiments may identify and provide more than one top intent cluster.

**[0051]** While the image module 412 is shown as part of the autosuggestion engine 404, alternative embodiments may locate the image module 412 outside of the autosuggestion engine 404. In these embodiments, the suggestions response includes the intent cluster information and image pointers. The UI engine 402 then works with the image module 412 to obtain the

corresponding images and videos using the image pointers.

**[0052]** Using the suggestions response, the UI engine 402 updates the autosuggestion box to display the autosuggestions. At a first level, the autosuggestions can include at least one smart suggestion. In some embodiments, the suggestions response includes all the different levels associated with the intent cluster. Thus, if the user selects a subcategory suggestion that includes a next level of autosuggestions (e.g., intent categories and groupings of subcategory suggestions), the next level of autosuggestions is presented. In other embodiments, the selection of the subcategory suggestion may trigger the autosuggestion engine 404 to perform operations to determine the next level of autosuggestions.

**[0053]** With each additional keystroke, the autosuggestion engine 404 is triggered to determine autosuggestions. In some cases, the suggestions response comprises a same intent cluster, while in other cases, the additional keystroke results in a different search term and/or matching intent cluster. If the user continues to type a search query to completion, a smart suggestion may not be presented as an autosuggestion. This is due to the precedence of what the user types trumping any autosuggestions.

**[0054]** When the user selects a lowest level search query (e.g., a regular suggestion, a subcategory suggestion with no further levels) or completes the search query typed in the search box, the search engine 406 performs a search. The search engine 406 returns search results to the UI engine 402, which presents the results on a search result page.

**[0055]** FIG. 5 is a flowchart illustrating operations of a method 500 for performing a search using multi-level rich autosuggestions for multi-intent search queries, according to some example embodiments. Operations in the method 500 may be performed by the network system 202 in the network environment 200 described above with respect to FIG. 2 – FIG. 4. Accordingly, the method 500 is described by way of example with reference to these components in the network system 202. However, it shall be appreciated that at least some of the operations of the method 500 may be deployed on various other hardware configurations or be performed by similar components residing elsewhere in the network environment 200. Therefore, the method 500 is not intended to be limited to these components.

**[0056]** In operation 502, the search system 212 receive a query prefix entered in a search box of a search user interface (UI) presented by the UI engine 402. The query prefix comprises one or more keystrokes entered by a user in the search box.

**[0057]** In operation 504, the autosuggestion engine 404 determines autosuggestions to provide to the user based on the query prefix. Operation 504 will be discussed in more detail in connection with FIG. 6 below.

**[0058]** In operation 506, the UI engine 402 provides autosuggestions in an autosuggestion

box associated with the search box. For example, the autosuggestion box can be shown below the search box. If the autosuggestions determined in operation 504 includes a smart suggestion, the autosuggestion box will show the smart suggestion visually distinguished from normal suggestions. The smart suggestion is a first-level suggestion of a top-ranking machine-learning device intent cluster.

5 [0059] In operation 508, a determination is made whether a next input is a keystroke or a selection of a suggestion. If the next input is a keystroke, the method 500 returns to operation 504 where autosuggestions are determined based on the updated prefix.

10 [0060] If, in operation 508, the next input is a selection of a smart autosuggestion, the method 500 continues to operation 510 where the autosuggestion box is updated with second level autosuggestions. The second level autosuggestions can include one or more intent categories whereby each intent category comprises a grouping of subcategory suggestions.

15 [0061] In operation 512, a selection of a suggestion from the second level is received by the UI engine 402. Based on the selection, a search is performed in operation 514. More specifically, the search engine 406 takes the selected suggestion as the search query and performs a search for corresponding information.

20 [0062] In some cases, there may be one or more further levels of autosuggestions. In these cases, the selection of the suggestion in operation 512 will cause an update to the autosuggestion box with a next level of autosuggestions. This continues until there is a selection of a lowest level suggestion. Once the lowest level suggestion is selected, the search is performed in operation 514.

[0063] Returning to operation 508, if the next input is a selection of a regular autosuggestion, then the search system 212 operates like a conventional search system and performs a search in operation 514.

25 [0064] In operation 516, the UI engine 402 displays a search result page that comprises the results of the search performed by the search engine 406.

30 [0065] FIG. 6 is a flowchart illustrating operations of a method 600 for determining autosuggestions, according to some example embodiments. Operations in the method 600 may be performed by the search system 212 (e.g., the autosuggestion engine 404) described above with respect to FIG. 4. Accordingly, the method 600 is described by way of example with reference to this component in the search system 212. However, it shall be appreciated that at least some of the operations of the method 600 may be deployed on various other hardware configurations or be performed by similar components residing elsewhere in the network environment 100. Therefore, the method 600 is not intended to be limited to this component.

35 [0066] In some embodiments, when the UI engine 402 receives the query prefix in 502,

the UI engine 402 sends an autosuggestion request to the autosuggestion engine 404. The autosuggestion request includes the query prefix. Thus, in operation 602, the autosuggestion engine 404 receives the query prefix.

**[0067]** In operation 604, the autosuggestion engine 404 identifies search terms that match the query prefix. In example embodiments, the prefix module 408 accesses the prefix log storage 304 and identifies all search terms that start with the same prefix as the query prefix.

**[0001]** In operation 606, the autosuggestion engine 404 determines a top intent cluster. In example embodiments, the intent module 410 accesses the intent data storage 316 and identifies machine-learning derived intent clusters or groupings that correspond to each search term identified by the prefix module 408. In example embodiments, the intent module identifies a top (machine-learning derived) intent cluster that matches each the search term. In some cases, the intent module 410 selects one intent cluster (e.g., highest probability or ranking intent cluster) to return as a smart suggestion. In some cases, more than one intent cluster can be returned as smart suggestions.

**[0002]** In operation 608, the autosuggestion engine 404 accesses images and videos that corresponding to the smart suggestion(s) that will be returned to the UI module 402. The intent clusters identified by the intent module 410 includes pointers to images and videos. In example embodiments, the image module 412 accesses and retrieves the images and videos from the image data storage 414 via the pointers. The images and videos may be associated with each intent category and/or subcategory suggestion.

**[0003]** In operation 610, the autosuggestion engine 404 returns a suggestions response to the user engine 402. In one embodiment, the suggestion response is a JSON suggestions response, although other forms of responses may be used. The suggestions response includes the one or more smart suggestions and corresponding images and/or videos.

**[0004]** FIG. 7 is a flowchart illustrating operations of a method 700 for generating machine learning-derived intent clusters, according to some example embodiments. Operations in the method 700 may be performed by the machine learning intent system 210 described above with respect to FIG. 3. Accordingly, the method 700 is described by way of example with reference to these components in the machine learning intent system 210. However, it shall be appreciated that at least some of the operations of the method 700 may be deployed on various other hardware configurations or be performed by similar components residing elsewhere in the network environment 100. Therefore, the method 700 is not intended to be limited to these components.

**[0005]** In operation 702, the grouping engine 314 accesses search logs. In example embodiments, the clustering module 318 accesses the search logs from the search log data storage 302. The search logs refer to data collected by the search system 212 on what users have searched

for in the past. In example embodiments, the search logs are used by the clustering module 318 to derive suggestions by analyzing patterns in the search terms.

**[0006]** In operation 704, the model-based suggestion generator 306 generates language model suggestions. These suggestions go beyond those that can be derived from the search logs.

5 In example embodiments, the model-based suggestion generator 306 comprises one or more machine-learning language models that are trained to understand and generate new suggestion candidates beyond those derived from the search logs. In one embodiment, the model-based suggestion generator 306 uses the information from the search logs data storage 302 to train the language models. Then, the trained language models are used to generate new, related  
10 suggestions. These new, related suggestions are then provided to the clustering module 318.

**[0007]** In operation 706, the grouping engine 314 accesses web logs and knowledge data. In example embodiments, the clustering module 318 accesses web logs from the web index log storage and accesses knowledge data from the knowledge data storage 310. The structured pieces of information from both web documents (e.g., from the web index log storage 308) and the  
15 knowledge data storage 310 are used by the grouping engine 314 to provide a refined set of intent.

**[0008]** In operation 708, the grouping engine 314 applies all of the data (e.g., search logs, language model-based suggestions, web documents, and structured knowledge base) to one or more machine-learning clustering models. Specifically, the clustering module 318 of the grouping engine 314 uses the data to group appropriate intents in a stack. In example  
20 embodiments, the clustering module 318 uses the machine learning clustering model(s) to group similar intents for a given query together.

**[0009]** After the clustering, the ranking module 320 of the grouping engine 314 determines relevance and ranking of each intent cluster for a given query input. In some embodiments, the ranking module 320 applies a relevance machine-learning model (e.g., a relevance algorithm) to  
25 the intent clusters generated by the clustering module 318 to determine relevance. The ranking module 320 then determines a ranking of each intent cluster that is deemed relevant (e.g., above a relevance threshold). In some embodiments, the ranking module 320 applies a ranking machine-learning model (e.g., a ranking algorithm) to these intent clusters. For example, the ranking model computes a probability or likelihood that a user is going to click on a given intent based on a  
30 particular query input (or query prefix).

**[0010]** In operation 710, the grouping engine 314 incorporates image pointers with the intent clusters. In example embodiments, the pointer module 324 identifies and assigns imagery (e.g., images, videos) to each intent cluster. The imagery can be determined based on data from web documents (e.g., the web index log storage 308) and knowledge base (e.g., knowledge data  
35 storage 310). The images may be stored in the image data storage 414 and a pointer corresponding



to each image/video in the image data storage 414 is included by the image module 412 with the intent clusters.

**[0011]** In operation 712, the grouping engine 314 stores results of the intent clustering to the intent data storage 316. The intent data storage 316 is then accessed during a search to return intent clusters based on query inputs (e.g., query prefixes).

**[0012]** While the operations of FIG. 5 - FIG. 7 are shown in a particular order, alternative embodiments may practice the operations in a different order. Additionally, one or more of the operations may be made optional in various embodiments.

**[0013]** FIG. 8A to FIG. 8D are example screenshots illustrating a search box and different levels of autosuggestions in an autosuggestion box. Referring to FIG. 8a, a user has typed “iPhone” in a search box 802. The search system 212 provides suggestions in an autosuggestion box 804 that are based on the query prefix “iPhone.” The suggestions include, for example, iPhone 13, iPhone 12, iPhone XR, iPhone Store, iPhone Amazon. iPhone 13 is a smart suggestion and is visually distinguished from the other (regular) suggestions. In the present example, the iPhone 13 suggestion includes a “+” icon that indicates that it is a smart suggestion. Alternative embodiment may use other icons or manners of visually distinguishing the smart suggestions (e.g., different color font, bold font).

**[0014]** If the user selects the smart suggestion (e.g., clicks on it), the set of suggestions in the autosuggestion box 804 gets replaced with a new set of rich autosuggestions which will now help the user with the different search intents that the user may have. This new set of autosuggestions is a second-level of autosuggestions. FIG. 8B shows an example screenshot illustrating the updated autosuggestion box 804 with the second-level of autosuggestions. Additionally, the search box 802 is updated to include the selected smart suggestion.

**[0015]** The autosuggestions include a plurality of intent categories 806. In some embodiments, the ranking of the suggestions in the autosuggestion box 804 is based on existing ranking of the suggestions with respect to the smart suggestion. Here the top ranked intent category 806 is iPhone 13 variants followed by iPhone 13 comparisons (e.g., “iPhone 13 VS”), iPhone 13 reviews by reviewers, and iPhone 13 specifications. Thus, a top-ranking intent category 806 (e.g., highest probability the user is looking for information with this intent) is the different types of iPhone 13 variants. A next highest ranking intent category 806 comprises comparisons of the iPhone 13 versus other models. This is followed by reviews of the iPhone 13 by various reviewers. The final intent category 806 comprises various specifications of the iPhone 13.

**[0016]** In some embodiments, the ranking is based on an overall search intent of the user. For instance, if the user is trying to make a buying decision, typically the user will want to figure

out the phone that they are interested in (e.g., first row are variants). Then, the user may compare with other models (e.g., second row are comparisons with other models). Then, the user is likely to research reviews of the product (e.g., third row are reviews by others). This overall search intent can be derived from search “journeys” performed by other users (e.g., the search logs).

5 **[0017]** Each of the intent categories 806 includes a plurality of subcategory suggestions 808. The subcategory suggestions 808 can include images and/or videos. The images and videos are retrieved user pointers to the image data storage 414. For example, the iPhone 13 Variants category includes subcategory suggestions 808 of iPhone 13 Pro Max, iPhone 13 Pro, and iPhone 13 with each displaying an associated image.

10 **[0018]** Assuming the user selects an option from the iPhone 13 variants intent category 806 (e.g., iPhone 13 Pro Max), the search box 802 and autosuggestion box 804 are updated to show a third level of autosuggestions. Referring now to FIG. 8C, the search box 802 has been updated to display the selected autosuggestion (e.g., iPhone 13 Pro Max).

**[0019]** The autosuggestion box 804 now shows new intent categories 810 corresponding to the iPhone 13 Pro Max search query. A top-ranking intent category 810 is iPhone 13 Pro Max color variants. This is followed by iPhone 13 Pro Max specifications, iPhone 13 Pro Max comparisons with other models, and iPhone 13 Pro Max reviews. Within each of these new intent categories 810 are a plurality of subcategory suggestions 812. For example, the iPhone 13 Pro Max color variant intent category 810 includes subcategory suggestions 812 of sierra blue, alpine green, graphite, silver, and gold.

20 **[0020]** At this point, the user can select one of the subcategory suggestions 812 and a search is performed. A search result page can then be presented.

**[0021]** While the example of FIG. 8A – FIG. 8C illustrates three levels of autosuggestions, alternative embodiments can comprise more or less levels. Additionally, the number of levels displayed to a user is dependent on user selection. For example and referring back to FIG. 8B, if the user selects a subcategory suggestion (e.g., camera, memory, storage) from the iPhone 13 specification intent category 806, a search may be performed based on the selected subcategory suggestion and a search result page presented in a next updated user interface.

30 **[0022]** In some embodiments, if the user does not choose any of the options in FIG. 8C (e.g., continues typing; after a predetermined amount of time has passed), the user interface can be refreshed to go to a state where the user gets regular autosuggestions as shown in FIG. 8D. As relevance goes down, the smart suggestion goes down, respectively. For instance, the smart suggestion is now shown at the bottom of the autosuggestion box 804. This allows the user to go back to a previous state by clicking on the smart suggestion.

35 **[0023]** FIG. 9 is an alternative example screenshot illustrating a search box 902 and

autosuggestions in an autosuggestion box 904. The example shown in FIG. 9 shows a use case with a sports related search query. Here, the user has selected a smart suggestion of “cricket.” In response, the search box 902 is updated to show the selected smart suggestion (e.g., visually distinguished by a “+” icon 906). The selection of the smart suggestion causes several intent categories to be displayed in the autosuggestion box 904 including cricket rankings and top cricket athletes in 2022. The autosuggestions can also include other autosuggestions including “cricket match Bangladesh New Zealand” and “cricket match Sri Lanka Australia,” which do not have subcategory suggestions.

**[0024]** FIG. 10 is a further example screenshot illustrating a search box 1002 and autosuggestions in an autosuggestion box 1004. The example shown in FIG. 10 shows a use case with an artist search query. Here, the user has selected a smart suggestion of “Ellie Goulding.” In response, the search box 1002 is updated to show the selected smart suggestion (e.g., visually distinguished by a “+” icon 1006). The selection of the smart suggestion causes several intent categories to be displayed in the autosuggestion box 1004 including events, albums and titles, songs, and awards. Each of these intent categories includes a plurality of subcategory suggestions. For example, selection of a song may cause a corresponding music video to play whereby the music video is accessed via a pointer to the image data storage 414. Alternatively, audio of the song may play without video. In this alternative case, a pointer to an audio file in an audio data storage may be used instead of a pointer to the image data storage 414.

**[0025]** FIG. 11 illustrates components of a machine 1100, according to some example embodiments, that is able to read instructions from a machine-storage medium (e.g., a machine-storage device, a non-transitory machine-storage medium, a computer-storage medium, or any suitable combination thereof) and perform any one or more of the methodologies discussed herein. Specifically, FIG. 11 shows a diagrammatic representation of the machine 1100 in the example form of a computer device (e.g., a computer) and within which instructions 1124 (e.g., software, a program, an application, an applet, an app, or other executable code) for causing the machine 1100 to perform any one or more of the methodologies discussed herein may be executed, in whole or in part.

**[0026]** For example, the instructions 1124 may cause the machine 1100 to execute the flow diagrams of FIG. 4 to FIG. 6. In one embodiment, the instructions 1124 can transform the machine 1100 into a particular machine (e.g., specially configured machine) programmed to carry out the described and illustrated functions in the manner described.

**[0027]** In alternative embodiments, the machine 1100 operates as a standalone device or may be connected (e.g., networked) to other machines. In a networked deployment, the machine 1100 may operate in the capacity of a server machine or a client machine in a server-client

network environment, or as a peer machine in a peer-to-peer (or distributed) network environment. The machine 1100 may be a server computer, a client computer, a personal computer (PC), a tablet computer, a laptop computer, a netbook, a set-top box (STB), a personal digital assistant (PDA), a cellular telephone, a smartphone, a web appliance, a network router, a network switch, a network bridge, or any machine capable of executing the instructions 1124 (sequentially or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term “machine” shall also be taken to include a collection of machines that individually or jointly execute the instructions 1124 to perform any one or more of the methodologies discussed herein.

5  
10 **[0028]** The machine 1100 includes a processor 1102 (e.g., a central processing unit (CPU), a graphics processing unit (GPU), a digital signal processor (DSP), an application specific integrated circuit (ASIC), a radio-frequency integrated circuit (RFIC), or any suitable combination thereof), a main memory 1104, and a static memory 1106, which are configured to communicate with each other via a bus 1108. The processor 1102 may contain microcircuits that are configurable, temporarily or permanently, by some or all of the instructions 1124 such that the processor 1102 is configurable to perform any one or more of the methodologies described herein, in whole or in part. For example, a set of one or more microcircuits of the processor 1102 may be configurable to execute one or more modules (e.g., software modules) described herein.

15  
20 **[0029]** The machine 1100 may further include a graphics display 1110 (e.g., a plasma display panel (PDP), a light emitting diode (LED) display, a liquid crystal display (LCD), a projector, or a cathode ray tube (CRT), or any other display capable of displaying graphics or video). The machine 1100 may also include an input device 1112 (e.g., a keyboard), a cursor control device 1114 (e.g., a mouse, a touchpad, a trackball, a joystick, a motion sensor, or other pointing instrument), a storage unit 1116, a signal generation device 1118 (e.g., a sound card, an amplifier, a speaker, a headphone jack, or any suitable combination thereof), and a network interface device 1120.

25  
30 **[0030]** The storage unit 1116 includes a machine-storage medium 1122 (e.g., a tangible machine-storage medium) on which is stored the instructions 1124 (e.g., software) embodying any one or more of the methodologies or functions described herein. The instructions 1124 may also reside, completely or at least partially, within the main memory 1104, within the processor 1102 (e.g., within the processor’s cache memory), or both, before or during execution thereof by the machine 1100. Accordingly, the main memory 1104 and the processor 1102 may be considered as machine-storage media (e.g., tangible and non-transitory machine-storage media). The instructions 1124 may be transmitted or received over a network 1126 via the network interface device 1120.

35

**[0031]** In some example embodiments, the machine 1100 may be a portable computing device and have one or more additional input components (e.g., sensors or gauges). Examples of such input components include an image input component (e.g., one or more cameras), an audio input component (e.g., a microphone), a direction input component (e.g., a compass), a location input component (e.g., a global positioning system (GPS) receiver), an orientation component (e.g., a gyroscope), a motion detection component (e.g., one or more accelerometers), an altitude detection component (e.g., an altimeter), and a gas detection component (e.g., a gas sensor). Inputs harvested by any one or more of these input components may be accessible and available for use by any of the modules described herein.

#### 10 EXECUTABLE INSTRUCTIONS AND MACHINE-STORAGE MEDIUM

**[0032]** The various memories (e.g., 1104, 1106, and/or memory of the processor(s) 1102) and/or storage unit 1116 may store one or more sets of instructions and data structures (e.g., software) 1124 embodying or utilized by any one or more of the methodologies or functions described herein. These instructions, when executed by processor(s) 1102 cause various operations to implement the disclosed embodiments.

**[0033]** As used herein, the terms “machine-storage medium,” “device-storage medium,” “computer-storage medium” (referred to collectively as “machine-storage medium 1122”) mean the same thing and may be used interchangeably in this disclosure. The terms refer to a single or multiple storage devices and/or media (e.g., a centralized or distributed database, and/or associated caches and servers) that store executable instructions and/or data, as well as cloud-based storage systems or storage networks that include multiple storage apparatus or devices. The terms shall accordingly be taken to include, but not be limited to, solid-state memories, and optical and magnetic media, including memory internal or external to processors. Specific examples of machine-storage media, computer-storage media, and/or device-storage media 1122 include non-volatile memory, including by way of example semiconductor memory devices, for example, erasable programmable read-only memory (EPROM), electrically erasable programmable read-only memory (EEPROM), FPGA, and flash memory devices; magnetic disks such as internal hard disks and removable disks; magneto-optical disks; and CD-ROM and DVD-ROM disks. The terms machine-storage medium or media, computer-storage medium or media, and device-storage medium or media 1122 specifically exclude carrier waves, modulated data signals, and other such media, at least some of which are covered under the term “signal medium” discussed below. In this context, the machine-storage medium is non-transitory.

#### 30 SIGNAL MEDIUM

**[0034]** The term “signal medium” or “transmission medium” shall be taken to include any form of modulated data signal, carrier wave, and so forth. The term “modulated data signal”

means a signal that has one or more of its characteristics set or changed in such a matter as to encode information in the signal.

#### COMPUTER READABLE MEDIUM

[0035] The terms “machine-readable medium,” “computer-readable medium” and  
5 “device-readable medium” mean the same thing and may be used interchangeably in this disclosure. The terms are defined to include both machine-storage media and signal media. Thus, the terms include both storage devices/media and carrier waves/modulated data signals.

[0036] The instructions 1124 may further be transmitted or received over a  
communications network 1126 using a transmission medium via the network interface device  
10 1120 and utilizing any one of a number of well-known transfer protocols (e.g., HTTP). Examples of communication networks 1126 include a local area network (LAN), a wide area network (WAN), the Internet, mobile telephone networks, plain old telephone service (POTS) networks, and wireless data networks (e.g., Wi-Fi, LTE, and WiMAX networks). The term “transmission medium” shall be taken to include any intangible medium that is capable of storing, encoding, or  
15 carrying instructions 1124 for execution by the machine 1100, and includes digital or analog communications signals or other intangible medium to facilitate communication of such software.

[0037] Throughout this specification, plural instances may implement components,  
operations, or structures described as a single instance. Although individual operations of one or  
20 more methods are illustrated and described as separate operations, one or more of the individual operations may be performed concurrently, and nothing requires that the operations be performed in the order illustrated. Structures and functionality presented as separate components in example configurations may be implemented as a combined structure or component. Similarly, structures and functionality presented as a single component may be implemented as separate components.  
25 These and other variations, modifications, additions, and improvements fall within the scope of the subject matter herein.

[0038] Certain embodiments are described herein as including logic or a number of  
components, modules, or mechanisms. Modules may constitute either software modules (e.g.,  
code embodied on a machine-storage medium or in a transmission signal) or hardware modules.  
30 A “hardware module” is a tangible unit capable of performing certain operations and may be configured or arranged in a certain physical manner. In various example embodiments, one or more computer systems (e.g., a standalone computer system, a client computer system, or a server computer system) or one or more hardware modules of a computer system (e.g., a processor or a group of processors) may be configured by software (e.g., an application or application portion)  
35 as a hardware module that operates to perform certain operations as described herein.

[0039] In some embodiments, a hardware module may be implemented mechanically, electronically, or any suitable combination thereof. For example, a hardware module may include dedicated circuitry or logic that is permanently configured to perform certain operations. For example, a hardware module may be a special-purpose processor, such as a field programmable gate array (FPGA) or an ASIC. A hardware module may also include programmable logic or circuitry that is temporarily configured by software to perform certain operations. For example, a hardware module may include software encompassed within a general-purpose processor or other programmable processor. It will be appreciated that the decision to implement a hardware module mechanically, in dedicated and permanently configured circuitry, or in temporarily configured circuitry (e.g., configured by software) may be driven by cost and time considerations.

[0040] Accordingly, the term “hardware module” should be understood to encompass a tangible entity, be that an entity that is physically constructed, permanently configured (e.g., hardwired), or temporarily configured (e.g., programmed) to operate in a certain manner or to perform certain operations described herein. As used herein, “hardware-implemented module” refers to a hardware module. Considering embodiments in which hardware modules are temporarily configured (e.g., programmed), each of the hardware modules need not be configured or instantiated at any one instance in time. For example, where the hardware modules comprise a general-purpose processor configured by software to become a special-purpose processor, the general-purpose processor may be configured as respectively different hardware modules at different times. Software may accordingly configure a processor, for example, to constitute a particular hardware module at one instance of time and to constitute a different hardware module at a different instance of time.

[0041] Hardware modules can provide information to, and receive information from, other hardware modules. Accordingly, the described hardware modules may be regarded as being communicatively coupled. Where multiple hardware modules exist contemporaneously, communications may be achieved through signal transmission (e.g., over appropriate circuits and buses) between or among two or more of the hardware modules. In embodiments in which multiple hardware modules are configured or instantiated at different times, communications between such hardware modules may be achieved, for example, through the storage and retrieval of information in memory structures to which the multiple hardware modules have access. For example, one hardware module may perform an operation and store the output of that operation in a memory device to which it is communicatively coupled. A further hardware module may then, at a later time, access the memory device to retrieve and process the stored output. Hardware modules may also initiate communications with input or output devices, and can operate on a resource (e.g., a collection of information).

**[0042]** The various operations of example methods described herein may be performed, at least partially, by one or more processors that are temporarily configured (e.g., by software) or permanently configured to perform the relevant operations. Whether temporarily or permanently configured, such processors may constitute processor-implemented modules that operate to perform one or more operations or functions described herein. As used herein, “processor-implemented module” refers to a hardware module implemented using one or more processors.

**[0043]** Similarly, the methods described herein may be at least partially processor-implemented, a processor being an example of hardware. For example, at least some of the operations of a method may be performed by one or more processors or processor-implemented modules. Moreover, the one or more processors may also operate to support performance of the relevant operations in a “cloud computing” environment or as a “software as a service” (SaaS). For example, at least some of the operations may be performed by a group of computers (as examples of machines including processors), with these operations being accessible via a network (e.g., the Internet) and via one or more appropriate interfaces (e.g., an application program interface (API)).

**[0044]** The performance of certain of the operations may be distributed among the one or more processors, not only residing within a single machine, but deployed across a number of machines. In some example embodiments, the one or more processors or processor-implemented modules may be located in a single geographic location (e.g., within a home environment, an office environment, or a server farm). In other example embodiments, the one or more processors or processor-implemented modules may be distributed across a number of geographic locations.

#### EXAMPLES

**[0045]** Example 1 is a method for providing multi-level rich autosuggestions for multi-intent search queries. The method comprises receiving a query prefix entered in a search box of a search user interface; accessing a database of machine-learning derived intent clusters; based on the query prefix, identifying a top-ranking machine-learning derived intent cluster; updating the search user interface to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster; receiving a selection of the smart suggestion; and in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.

**[0046]** In example 2, the subject matter of example 1 can optionally include wherein at



least some of the plurality of subcategory suggestions comprise a corresponding graphical element.

5 [0047] In example 3, the subject matter of any of examples 1-2 can optionally include wherein the graphical element is an image corresponding to a subcategory suggestion of the plurality of subcategory suggestions.

[0048] In example 4, the subject matter of any of examples 1-3 can optionally include wherein the graphical element is a video element corresponding to a subcategory suggestion of the plurality of subcategory suggestions, selection of the video element causing a corresponding video to play.

10 [0049] In example 5, the subject matter of any of examples 1-4 can optionally include receiving a selection of a subcategory suggestion from one of the intent categories; and in response to the selection of the subcategory suggestion updating the search box with the selected subcategory suggestion; and replacing the second level autosuggestions with third level autosuggestions in the autosuggestion box, the third level autosuggestions comprising one or  
15 more third level intent categories that each represents a different category that is a subset of the selected subcategory suggestion.

[0050] In example 6, the subject matter of any of examples 1-5 can optionally include receiving a selection of a subcategory suggestion from one of the intent clusters; and in response to the selection of the subcategory suggestion performing a search based on the selected  
20 subcategory suggestion; and causing display of a result page based on the search.

[0051] In example 7, the subject matter of any of examples 1-6 can optionally include wherein the smart suggestion is displayed visually distinguished from regular suggestions of the plurality of suggestions in the autosuggestion box with a smart suggestion icon.

[0052] In example 8, the subject matter of any of examples 1-7 can optionally include  
25 generating the database of machine-learning derived intent clusters, the generating comprising accessing search logs of previous searches performed by users; and applying the search logs to a machine-learning clustering model that clusters similar search intents together for a given query input.

[0053] In example 9, the subject matter of any of examples 1-8 can optionally include  
30 periodically updating the database of machine-learning derived intent clusters by accessing updated data from the search logs and applying updated data to the machine-learning model.

[0054] In example 10, the subject matter of any of examples 1-9 can optionally include  
35 accessing suggestions generated by a language model, web documents, and knowledge data, wherein the applying the search logs to the machine-learning clustering model further comprises applying the suggestions generated by the language model, the web documents, and the

knowledge data to the machine-learning clustering model.

**[0055]** In example 11, the subject matter of any of examples 1-10 can optionally include ranking the machine-learning derived intent clusters, the ranking comprising applying the machine-learning derived intent clusters to a machine-learning ranking model that determines a probability that each machine-learning derived intent clusters will be selected for the given query input.

**[0056]** In example 12, the subject matter of any of examples 1-11 can optionally include training the machine-learning ranking model on a predetermined duration of search queries and user feedback; and periodically retraining the machine-learning ranking model with a rolling predetermined duration of search queries and user feedback.

**[0057]** Example 13 is a system for providing multi-level rich autosuggestions for multi-intent search queries. The system comprises one or more hardware processors and a memory storing instructions that, when executed by the one or more hardware processors, cause the one or more hardware processors to perform operations comprising receiving a query prefix entered in a search box of a search user interface; accessing a database of machine-learning derived intent clusters; based on the query prefix, identifying a top-ranking machine-learning derived intent cluster; updating the search user interface to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster; receiving a selection of the smart suggestion; and in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.

**[0058]** In example 14, the subject matter of example 13 can optionally include wherein the operations further comprise receiving a selection of a subcategory suggestion from one of the intent categories; and in response to the selection of the subcategory suggestion updating the search box with the selected subcategory suggestion; and replacing the second level autosuggestions with third level autosuggestions in the autosuggestion box, the third level autosuggestions comprising one or more third level intent categories that each represents a different category that is a subset of the selected subcategory suggestion.

**[0059]** In example 15, the subject matter of any of examples 13-14 can optionally include wherein the operations further comprise receiving a selection of a subcategory suggestion from one of the intent clusters; and in response to the selection of the subcategory suggestion performing a search based on the selected subcategory suggestion; and causing display of a result

page based on the search.

**[0060]** In example 16, the subject matter of any of examples 13-15 can optionally include wherein the smart suggestion is displayed visually distinguished from regular suggestions of the plurality of suggestions in the autosuggestion box with a smart suggestion icon.

5 **[0061]** In example 17, the subject matter of any of examples 13-16 can optionally include wherein the operations further comprise generating the database of machine-learning derived intent clusters, the generating comprising accessing search logs of previous searches performed by users; and applying the search logs to a machine-learning clustering model that clusters similar search intents together for a given query input.

10 **[0062]** In example 18, the subject matter of any of examples 13-17 can optionally include wherein the operations further comprise ranking the machine-learning derived intent clusters, the ranking comprising applying the machine-learning derived intent clusters to a machine-learning ranking model that determines a probability that each machine-learning derived intent clusters will be selected for the given query input.

15 **[0063]** In example 19, the subject matter of any of examples 13-18 can optionally include wherein the operations further comprise training the machine-learning ranking model on a predetermined duration of search queries and user feedback; and periodically retraining the machine-learning ranking model with a rolling predetermined duration of search queries and user feedback.

20 **[0064]** Example 20 is a storage medium comprising instructions which, when executed by one or more hardware processors of a machine, cause the machine to perform operations for providing multi-level rich autosuggestions for multi-intent search queries. The operations comprise receiving a query prefix entered in a search box of a search user interface; accessing a database of machine-learning derived intent clusters; based on the query prefix, identifying a top-ranking  
25 machine-learning derived intent cluster; updating the search user interface to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster; receiving a selection of the smart suggestion; and in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion  
30 box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.

**[0065]** Some portions of this specification may be presented in terms of algorithms or symbolic representations of operations on data stored as bits or binary digital signals within a  
35 machine memory (e.g., a computer memory). These algorithms or symbolic representations are

examples of techniques used by those of ordinary skill in the data processing arts to convey the substance of their work to others skilled in the art. As used herein, an “algorithm” is a self-consistent sequence of operations or similar processing leading to a desired result. In this context, algorithms and operations involve physical manipulation of physical quantities. Typically, but not necessarily, such quantities may take the form of electrical, magnetic, or optical signals capable of being stored, accessed, transferred, combined, compared, or otherwise manipulated by a machine. It is convenient at times, principally for reasons of common usage, to refer to such signals using words such as “data,” “content,” “bits,” “values,” “elements,” “symbols,” “characters,” “terms,” “numbers,” “numerals,” or the like. These words, however, are merely convenient labels and are to be associated with appropriate physical quantities.

**[0066]** Unless specifically stated otherwise, discussions herein using words such as “processing,” “computing,” “calculating,” “determining,” “presenting,” “displaying,” or the like may refer to actions or processes of a machine (e.g., a computer) that manipulates or transforms data represented as physical (e.g., electronic, magnetic, or optical) quantities within one or more memories (e.g., volatile memory, non-volatile memory, or any suitable combination thereof), registers, or other machine components that receive, store, transmit, or display information. Furthermore, unless specifically stated otherwise, the terms “a” or “an” are herein used, as is common in patent documents, to include one or more than one instance. Finally, as used herein, the conjunction “or” refers to a non-exclusive “or,” unless specifically stated otherwise.

**[0067]** Although an overview of the present subject matter has been described with reference to specific example embodiments, various modifications and changes may be made to these embodiments without departing from the broader scope of embodiments of the present invention. For example, various embodiments or features thereof may be mixed and matched or made optional by a person of ordinary skill in the art. Such embodiments of the present subject matter may be referred to herein, individually or collectively, by the term “invention” merely for convenience and without intending to voluntarily limit the scope of this application to any single invention or present concept if more than one is, in fact, disclosed.

**[0068]** The embodiments illustrated herein are believed to be described in sufficient detail to enable those skilled in the art to practice the teachings disclosed. Other embodiments may be used and derived therefrom, such that structural and logical substitutions and changes may be made without departing from the scope of this disclosure. The Detailed Description, therefore, is not to be taken in a limiting sense, and the scope of various embodiments is defined only by the appended claims, along with the full range of equivalents to which such claims are entitled.

**[0069]** Moreover, plural instances may be provided for resources, operations, or structures described herein as a single instance. Additionally, boundaries between various resources,

operations, modules, engines, and data stores are somewhat arbitrary, and particular operations are illustrated in a context of specific illustrative configurations. Other allocations of functionality are envisioned and may fall within a scope of various embodiments of the present invention. In general, structures and functionality presented as separate resources in the example configurations may be implemented as a combined structure or resource. Similarly, structures and functionality presented as a single resource may be implemented as separate resources. These and other variations, modifications, additions, and improvements fall within a scope of embodiments of the present invention as represented by the appended claims. The specification and drawings are, accordingly, to be regarded in an illustrative rather than a restrictive sense.

## CLAIMS

1. A method comprising:
  - receiving a query prefix entered in a search box of a search user interface;
  - accessing a database of machine-learning derived intent clusters ;
  - based on the query prefix, identifying a top-ranking machine-learning derived intent cluster;
  - updating the search user interface to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster;
  - receiving a selection of the smart suggestion; and
  - in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.
2. The method of claim 1, wherein at least some of the plurality of subcategory suggestions comprise a corresponding graphical element.
3. The method of claim 2, wherein the graphical element is an image corresponding to a subcategory suggestion of the plurality of subcategory suggestions.
4. The method of claim 2, wherein the graphical element is a video element corresponding to a subcategory suggestion of the plurality of subcategory suggestions, selection of the video element causing a corresponding video to play.
5. The method of claim 1, further comprising:
  - receiving a selection of a subcategory suggestion from one of the intent categories; and
  - in response to the selection of the subcategory suggestion:
    - updating the search box with the selected subcategory suggestion; and
    - replacing the second level autosuggestions with third level autosuggestions in the autosuggestion box, the third level autosuggestions comprising one or more third level intent categories that each represents a different category that is a subset of the selected subcategory suggestion.
6. The method of claim 1, further comprising:
  - receiving a selection of a subcategory suggestion from one of the intent clusters; and
  - in response to the selection of the subcategory suggestion:
    - performing a search based on the selected subcategory suggestion; and
    - causing display of a result page based on the search.
7. The method of claim 1, wherein the smart suggestion is displayed visually distinguished

from regular suggestions of the plurality of suggestions in the autosuggestion box with a smart suggestion icon.

8. The method of claim 1, further comprising generating the database of machine-learning derived intent clusters, the generating comprising:

accessing search logs of previous searches performed by users; and

applying the search logs to a machine-learning clustering model that clusters similar search intents together for a given query input.

9. The method of claim 8, further comprising:

periodically updating the database of machine-learning derived intent clusters by accessing updated data from the search logs and applying updated data to the machine-learning model.

10. The method of claim 8, further comprising:

accessing suggestions generated by a language model, web documents, and knowledge data, wherein the applying the search logs to the machine-learning clustering model further comprises applying the suggestions generated by the language model, the web documents, and the knowledge data to the machine-learning clustering model.

11. The method of claim 8, further comprising:

ranking the machine-learning derived intent clusters, the ranking comprising applying the machine-learning derived intent clusters to a machine-learning ranking model that determines a probability that each machine-learning derived intent clusters will be selected for the given query input.

12. The method of claim 11, further comprising:

training the machine-learning ranking model on a predetermined duration of search queries and user feedback; and

periodically retraining the machine-learning ranking model with a rolling predetermined duration of search queries and user feedback.

13. A system comprising:

one or more hardware processors; and

a memory storing instructions that, when executed by the one or more hardware processors, cause the one or more hardware processors to perform operations comprising:

receiving a query prefix entered in a search box of a search user interface;

accessing a database of machine-learning derived intent clusters;

based on the query prefix, identifying a top-ranking machine-learning derived intent cluster;

updating the search user interface to provide a plurality of first level autosuggestions in an

autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster;

receiving a selection of the smart suggestion; and

in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.

14. The system of claim 13, wherein the operations further comprise:

receiving a selection of a subcategory suggestion from one of the intent categories; and

in response to the selection of the subcategory suggestion:

updating the search box with the selected subcategory suggestion; and

replacing the second level autosuggestions with third level autosuggestions in the autosuggestion box, the third level autosuggestions comprising one or more third level intent categories that each represents a different category that is a subset of the selected subcategory suggestion.

15. The system of claim 13, wherein the operations further comprise:

receiving a selection of a subcategory suggestion from one of the intent clusters; and

in response to the selection of the subcategory suggestion:

performing a search based on the selected subcategory suggestion; and

causing display of a result page based on the search.

16. The system of claim 13, wherein the smart suggestion is displayed visually distinguished from regular suggestions of the plurality of suggestions in the autosuggestion box with a smart suggestion icon.

17. The system of claim 13, wherein the operations further comprise generating the database of machine-learning derived intent clusters, the generating comprising:

accessing search logs of previous searches performed by users; and

applying the search logs to a machine-learning clustering model that clusters similar search intents together for a given query input.

18. The system of claim 17, wherein the operations further comprise:

ranking the machine-learning derived intent clusters, the ranking comprising applying the machine-learning derived intent clusters to a machine-learning ranking model that determines a probability that each machine-learning derived intent clusters will be selected for the given query input.

19. The system of claim 18, wherein the operations further comprise:

training the machine-learning ranking model on a predetermined duration of search queries



and user feedback; and

periodically retraining the machine-learning ranking model with a rolling predetermined duration of search queries and user feedback.

20. A storage medium comprising instructions which, when executed by one or more hardware processors of a machine, cause the machine to perform operations comprising:

receiving a query prefix entered in a search box of a search user interface;

accessing a database of machine-learning derived intent clusters;




based on the query prefix, identifying a top-ranking machine-learning derived intent cluster;

updating the search user interface to provide a plurality of first level autosuggestions in an autosuggestion box including a smart suggestion, the smart suggestion being a first level suggestion of the top-ranking machine-learning derived intent cluster;

receiving a selection of the smart suggestion; and

in response to receiving the selection of the smart suggestion, replacing the plurality of first level autosuggestions in the autosuggestion box with second level autosuggestions, the second level autosuggestions comprising one or more intent categories that are subsets of the smart suggestion, each intent category comprising a plurality of subcategory suggestions corresponding to its intent category.

LEVEL 1: AUTOSUGGEST

SEARCH QUERY   

SMART SUGGESTION

REGULAR SUGGESTION

REGULAR SUGGESTION




REGULAR SUGGESTION

REGULAR SUGGESTION

106

102

LEVEL 2: AUTOSUGGEST

NEW QUERY   

CATEGORY SUGGESTION

SUBCATEGORY SUGGESTION

CATEGORY SUGGESTION

SUBCATEGORY SUGGESTION

CATEGORY SUGGESTION

SUBCATEGORY SUGGESTION

CATEGORY SUGGESTION

SUBCATEGORY SUGGESTION

108

110

112

102

114

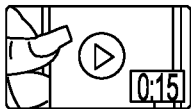
SEARCH RESULT PAGE

MICROSOFT BING (IPHONE 13 PRO MAX SIERRA BLUE)    MARK  4065  

ALL WORK NEWS IMAGES VIDEOS MAPS SHOPPING

3,790,000,000 RESULTS ANYTIME 

APPLE IPHONE 13 PRO, 128GB, SIERRA BLUE - UNLOCKED ...  
HTTPS://WWW.AMAZON.COM/APPLE-IPHONE-13-PRO-SIERRA/



3 ROWS - NOV 11, 2021 - FACE ID CERAMIC SHIELD FRONT SURGICAL-GRADE STAINLESS STEEL. WATER RESISTANT TO A DEPTH OF 6  
4.1/5 ★★★★★ (100)

WAS: \$954.99 \$954.99 DETAILS

PRICE: \$894.99 \$894.99

YOU SAVE: \$60.00 \$60.00 (6%)

SEE ALL 3 ROWS ON WWW.AMAZON.COM

APPLE IPHONE 13 PRO, 128GB, SIERRA BLUE [LOCKED]  
HTTPS://WWW.AMAZON.COM/APPLE-IPHONE-SIERRA-CARRIER...

10 ROWS - IPHONE 13 PRO. DISPLAY, 6.1-INCH SUPER RETINAL XDR DISPLAY WITH PROMOTION. CAPACITY. 128GB, 256GB, ...

BRAND: APPLE CAPACITY: 128GB, 256GB, 512GB, 1TB

	IPHONE 13 PRO	IPHONE 13 PRO...	IPHONE 13	IPHONE 13 MINI
PRICE	FROM: \$899.99	FROM \$999.99	FROM \$829.99	FROM: \$729.00
RATINGS	NO CUSTOMER R...	4.4 OUT OF 5 ST...	3.2 OUT OF 5 ST...	2.7 OUT OF 5 ST...
DISPLAY	6.1-INCH SUPER ...	6.7-INCH SUPER...	6.1-INCH SUPER...	5.4-INCH SUPER...

APPLE IPHONE 13 PRO (128GB) SIERRA BLUE

READ REVIEWS AND BUY APPLE IPHONE 13 PRO (128GB) - SIERRA BLUE AT TARGET. CHOOSE FROM CONTACTLESS SAME DAY DELIVERY, DRIVE UP AND MORE

AVAILABILITY: OUT OF STOCK

BRAND: APPLE

APPLE IPHONE 13 PRO 128GB) - SIERRA BLUE: TARGET

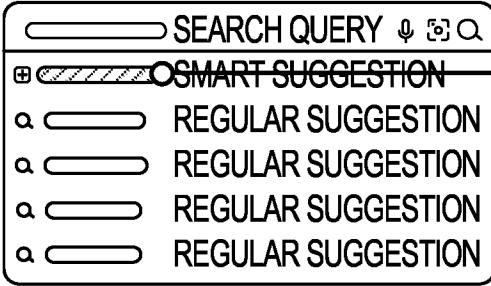
HTTPS://WWW.TARGET.COM/PIAPPLE-IPHONE-13-PRO-128GB-SIERRA-BLUE/-/A-84240095

4.0/5  
TARGET.COM

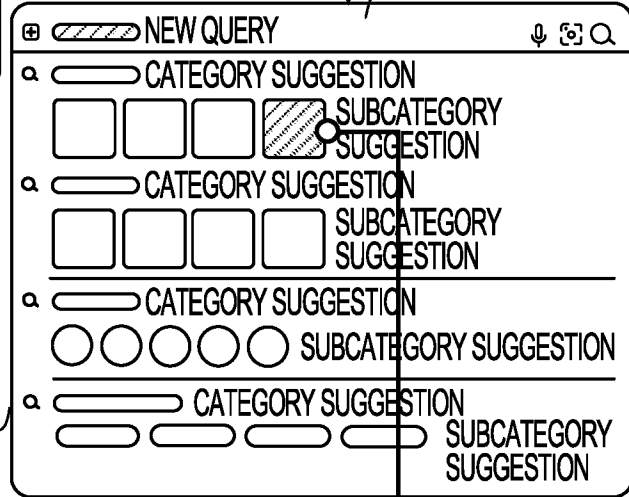
SUGGEST AN EDIT

FIG. 1A

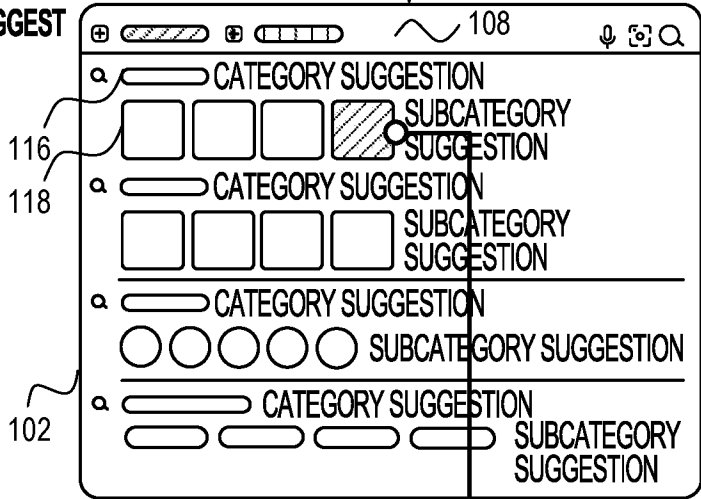
LEVEL 1: AUTOSUGGEST



LEVEL 2: AUTOSUGGEST



LEVEL 3: AUTOSUGGEST



LEVEL 4: AUTOSUGGEST

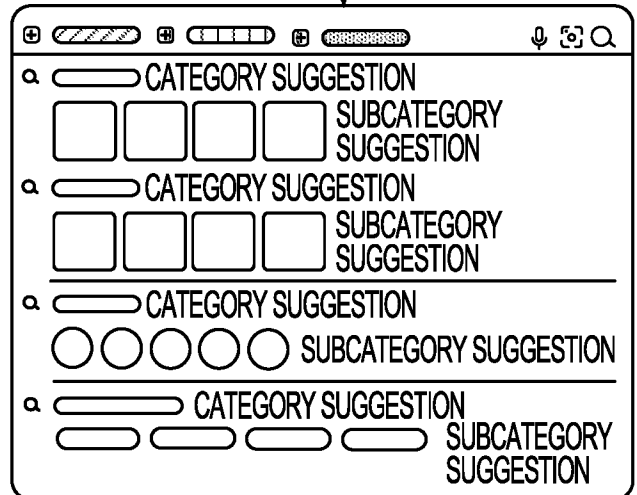


FIG. 1B

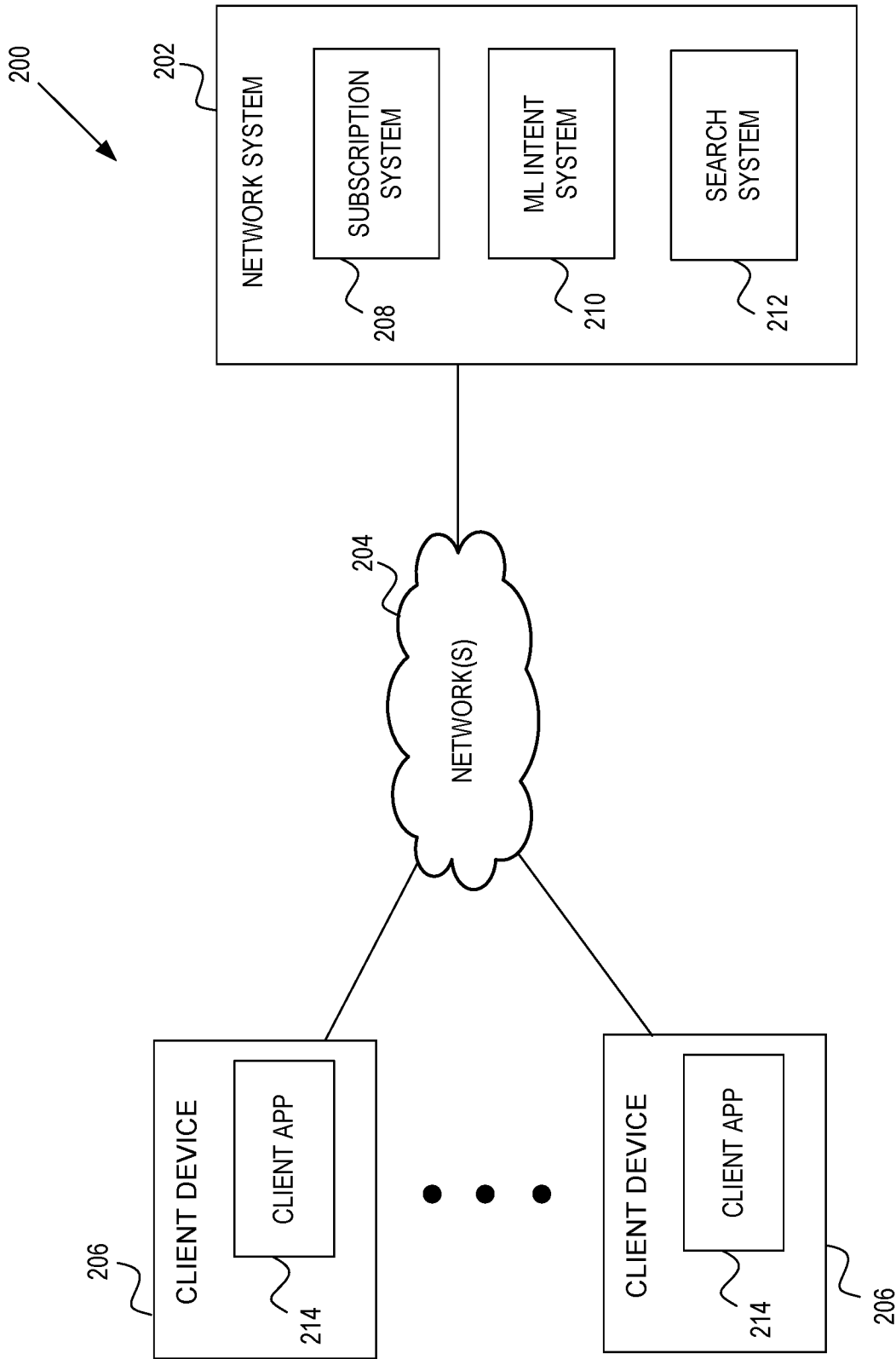


FIG. 2

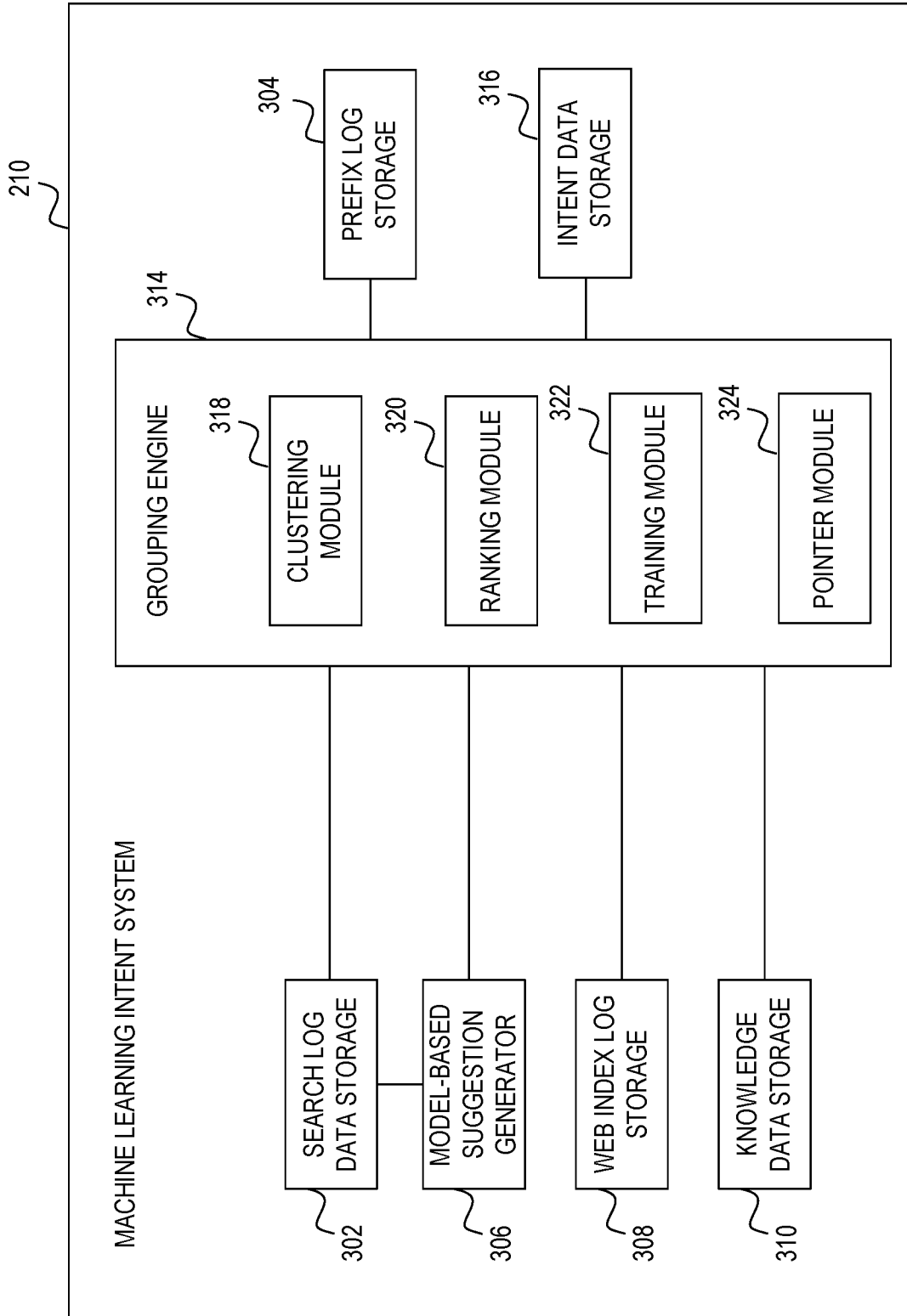


FIG. 3

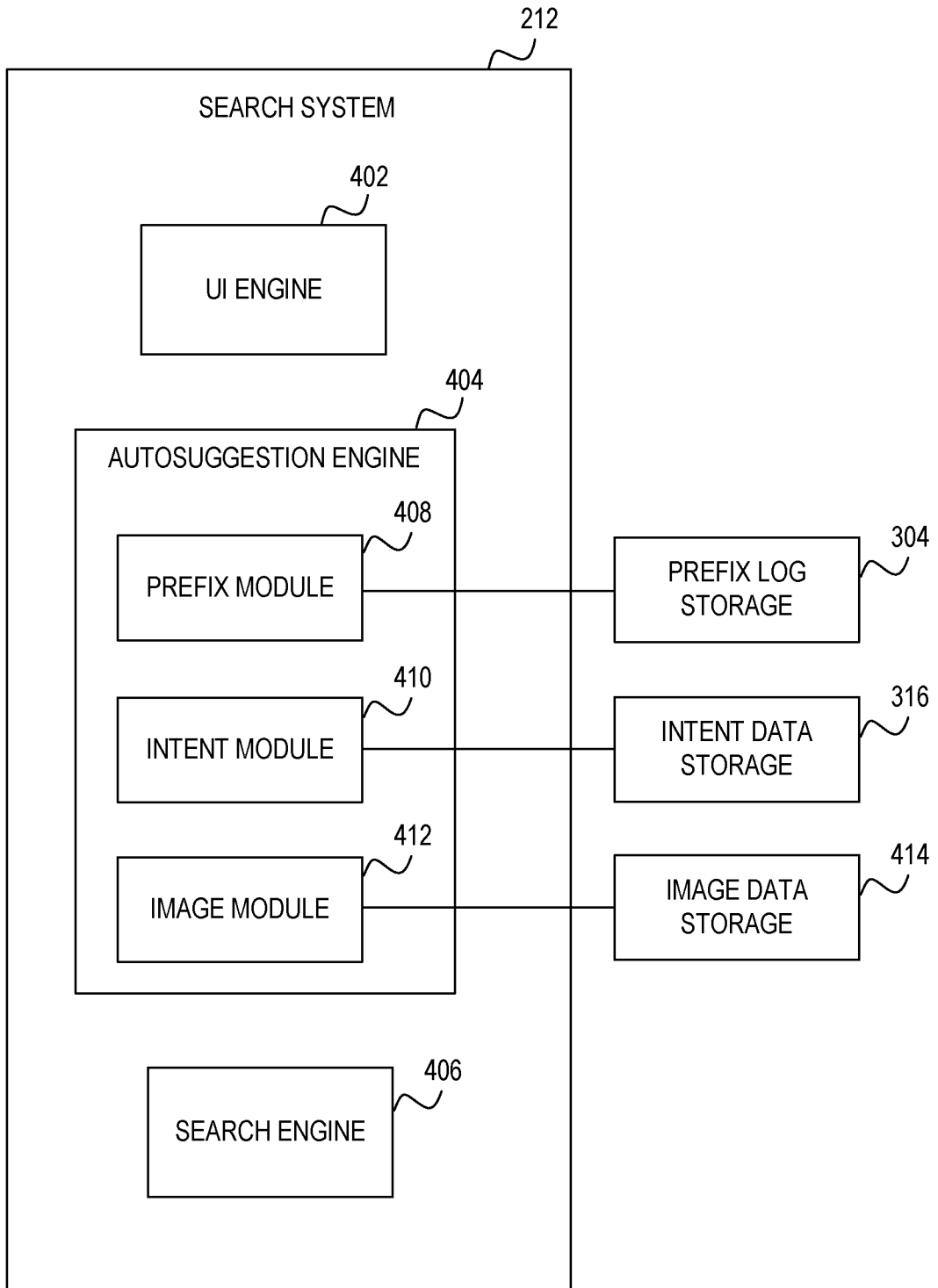


FIG. 4

6/15

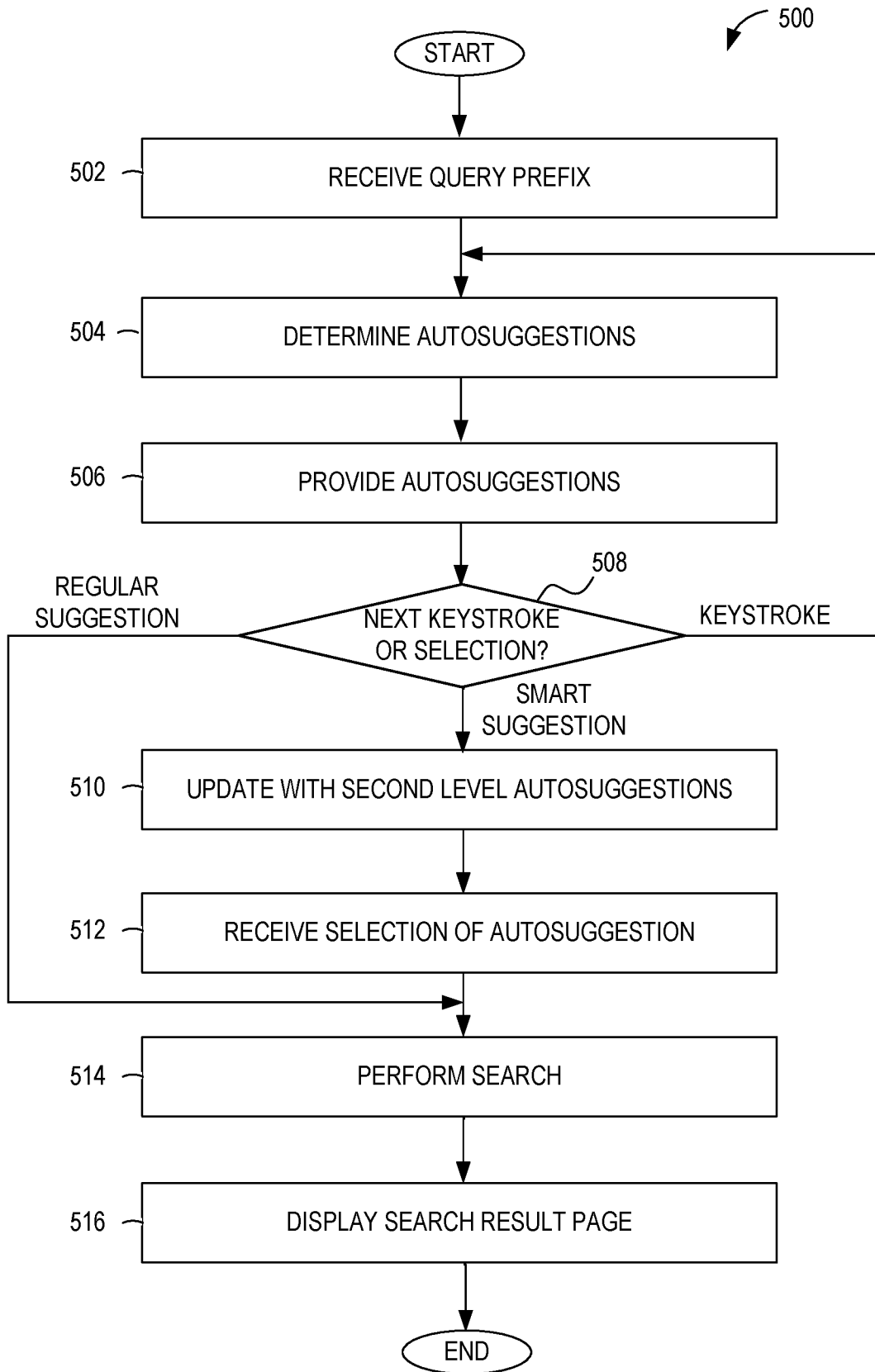


FIG. 5

7/15

504

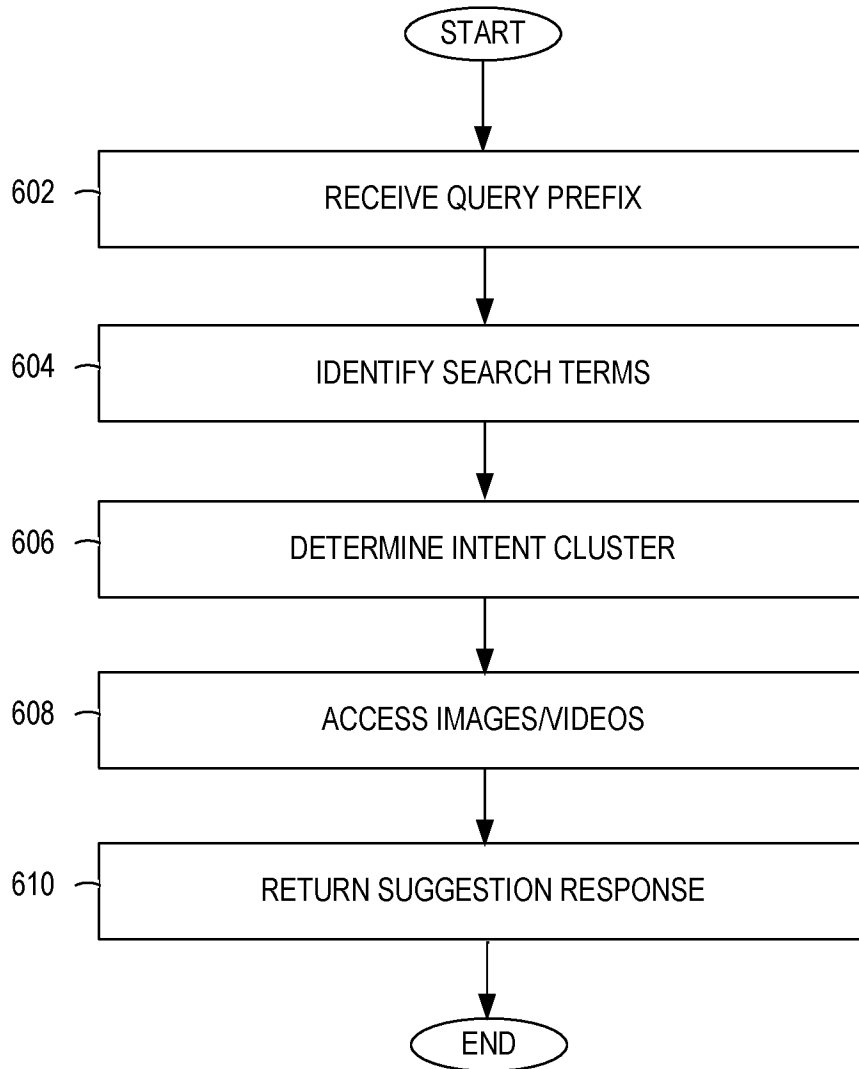


FIG. 6



8/15

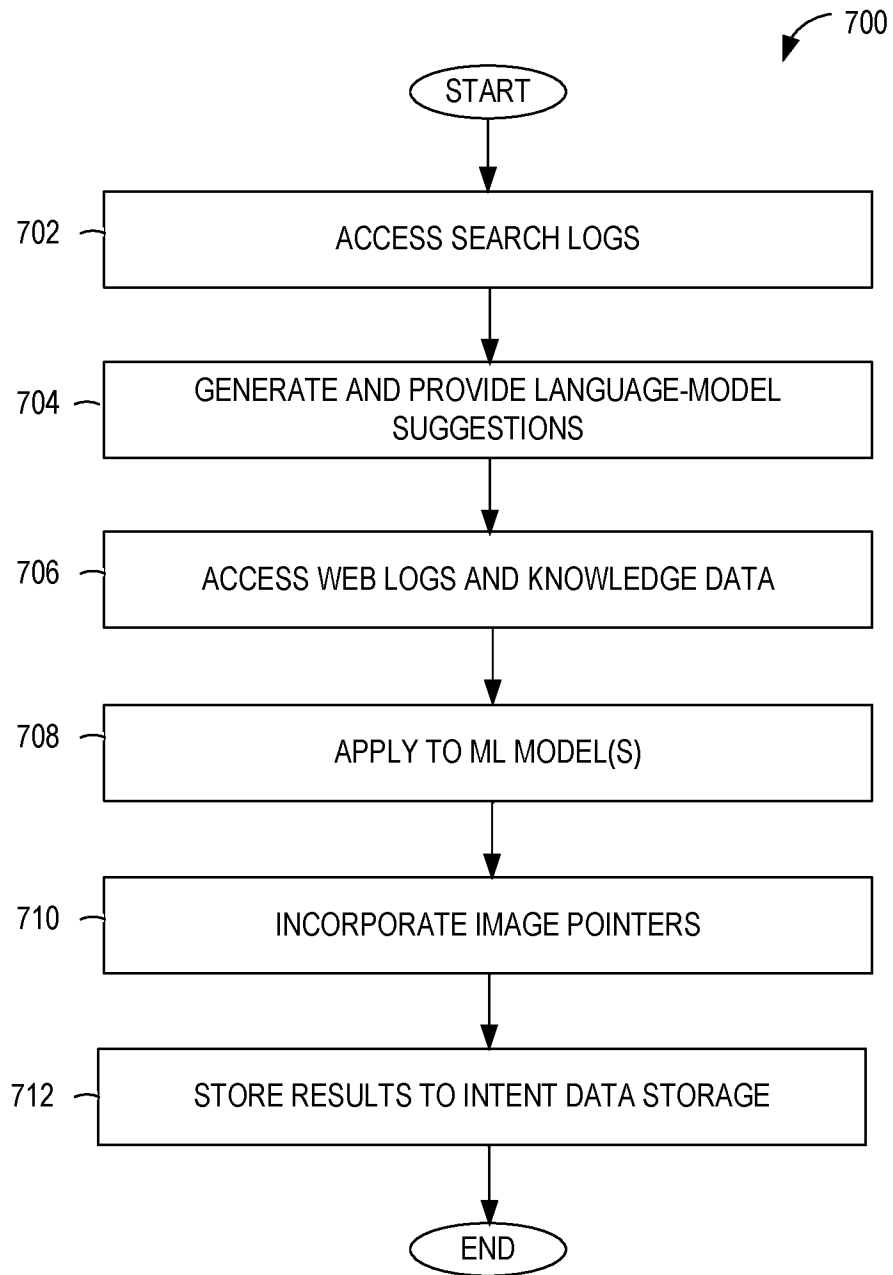


FIG. 7

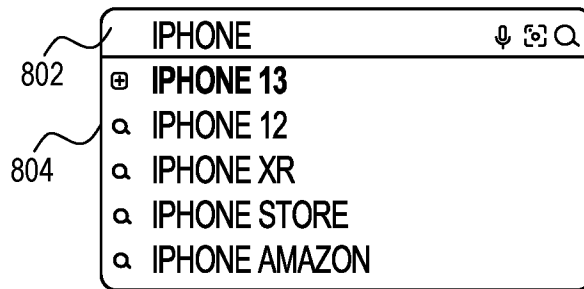


FIG. 8A



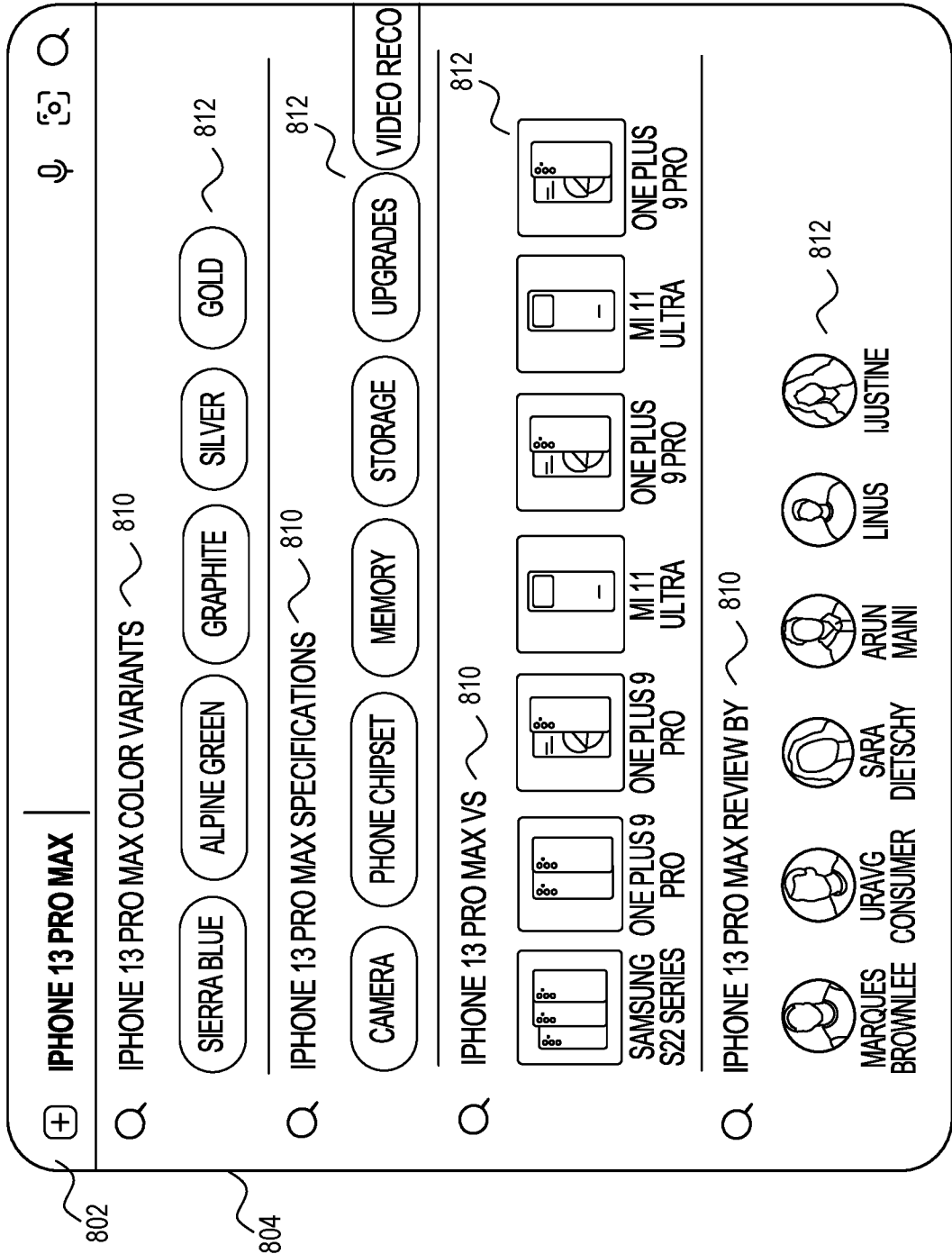


FIG. 8C

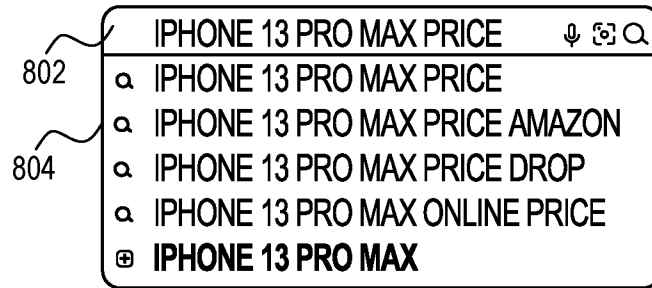


FIG. 8D

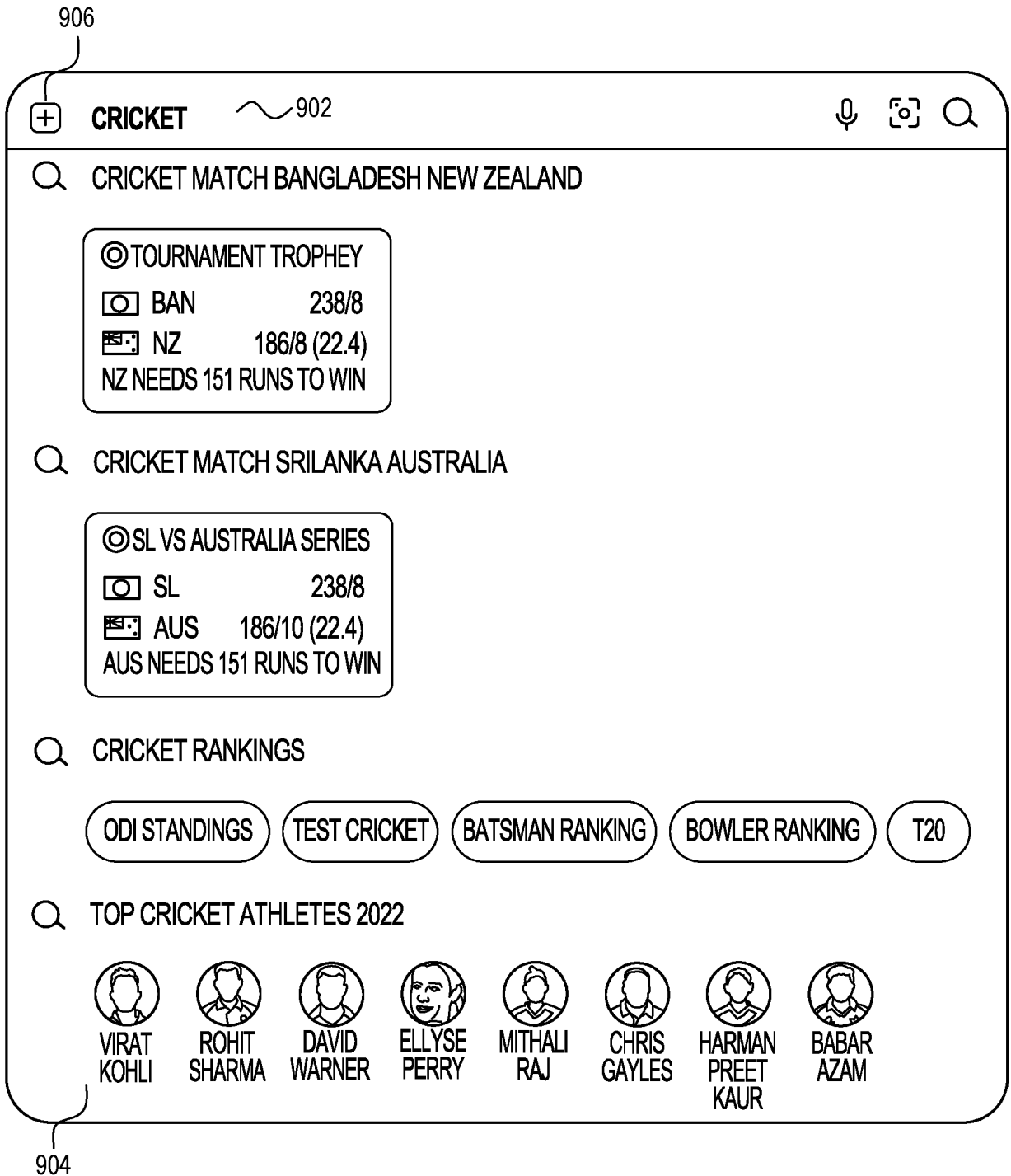


FIG. 9

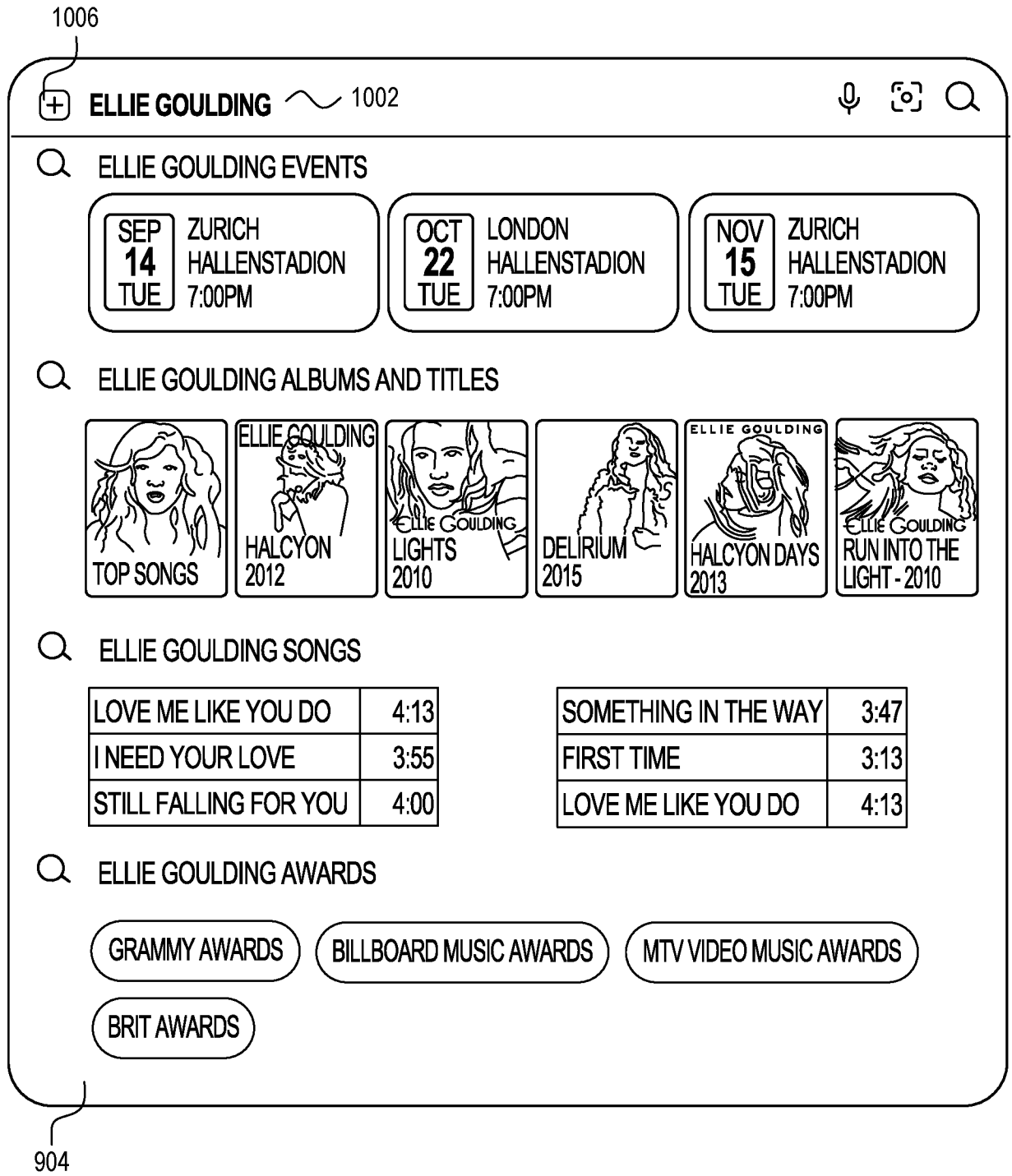


FIG. 10

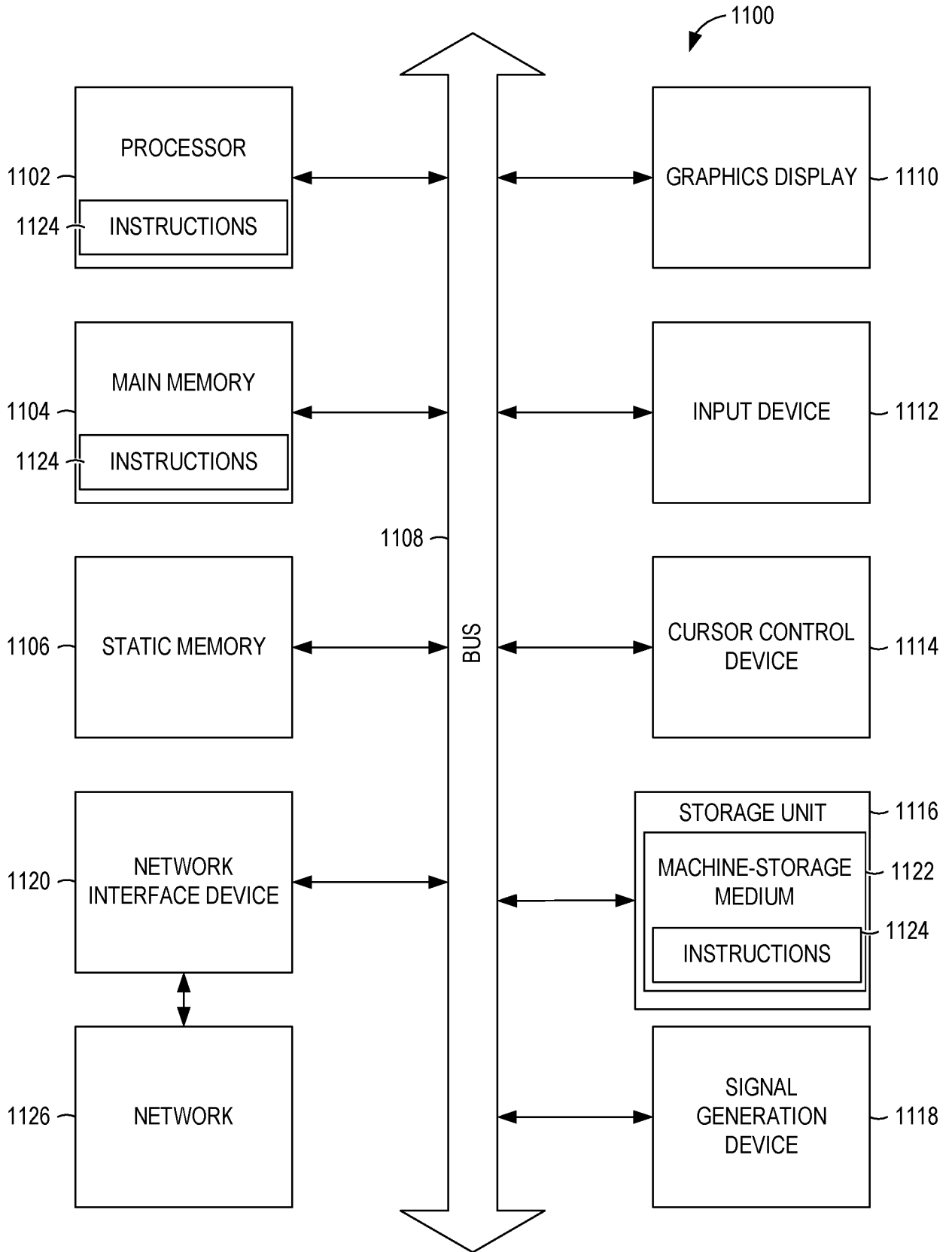


FIG. 11



# INTERNATIONAL SEARCH REPORT

International application No  
PCT/US2024/018755

**A. CLASSIFICATION OF SUBJECT MATTER**  
INV. G06F16/332  
ADD.

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 G06N**

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)

**EPO-Internal, WPI Data**

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
<b>X</b>	<p><b>US 9 576 074 B2 (MICROSOFT CORP [US]; MICROSOFT TECHNOLOGY LICENSING LLC [US]) 21 February 2017 (2017-02-21) column 16; figure 5 column 8; figure 2 column 17 column 15; figure 4</b></p> <p style="text-align: center;">-----</p>	<b>1 - 20</b>

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  
  
**4 June 2024**

Date of mailing of the international search report  
  
**21/06/2024**

Name and mailing address of the ISA/  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer  
  
**Krawaritis, Achilles**

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2024/018755

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 9576074	B2 21-02-2017	CN 105453082 A	30-03-2016
		EP 3011479 A1	27-04-2016
		US 2014379744 A1	25-12-2014
		US 2017090595 A1	30-03-2017
		WO 2014204691 A1	24-12-2014
-----			