KNOWLEDGE BASE CONSTRUCTION METHOD, DATA RETRIEVAL METHOD AND APPARATUS, AND CLOUD DEVICE

[0001] The present application claims priority to Chinese Patent Application No. 202310166877.7, filed with the China National Intellectual Property Administration on February 13, 2023 and entitled "KNOWLEDGE BASE CONSTRUCTION METHOD, DATA RETRIEVAL METHOD AND APPARATUS, AND CLOUD DEVICE", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

[0002] The present application relates to the field of computer technologies, and in particular, to a knowledge base construction method, a data retrieval method, an apparatus, and a cloud device.

BACKGROUND

[0003] A question-and-answer knowledge base is used in an intelligent customer service scenario. Based on a query text of a user, a corresponding answer text is retrieved from the question-and-answer knowledge base to provide a response to the user.

[0004] In related technologies, a segment of text in raw text data is used as an index, and the raw text is used as a value of this index, to construct the question-and-answer knowledge base. However, the question-and-answer knowledge base obtained based on this construction method is not applicable to retrieval for different modalities of answer data. This limitation restricts the application scope of the question-and-answer knowledge base in the intelligent customer service scenario.

SUMMARY

[0005] According to a plurality of aspects of the present application, a knowledge base construction method, a data retrieval method, an apparatus, and a cloud device are provided, to resolve a problem that a question-and-answer knowledge base constructed by using a related technology is not applicable to retrieval for different modalities of answer data.

[0006] According to a first aspect of embodiments of the present application, a knowledge base construction method is provided. The method includes:

acquiring a plurality of pieces of raw data, where modalities of the plurality of pieces of raw data are at least two of text, image, or video;

for raw data in the plurality of pieces of raw data, determining a key text of the raw data, determining a key of a key-value pair based on the key text, and determining a value of the key-value pair based on the raw data, where the key text is used for describing content of the raw data; and

constructing a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

[0007] According to a second aspect of embodiments of the present application, a data retrieval method is provided. The method includes:

receiving query data sent by a terminal device, where the query data is one of a text, an image, or a video;

retrieving, in a question-and-answer knowledge base, answer data corresponding to the query data, where a modality of the answer data is one of text, image, or video, and the question-and-answer knowledge base is constructed based on the knowledge base construction method according to any one of the first aspect; and

sending the answer data to the terminal device.

[0008] According to a third aspect of embodiments of the present application, a data retrieval method is provided. The data retrieval method is applied to a terminal device. The data retrieval method includes:

sending query data to a server; and

receiving answer data that is sent by the server and that is of the query data, where

the answer data is determined based on the data retrieval method according to the second aspect.

[0009] According to a fourth aspect of embodiments of the present application, a knowledge base construction apparatus is provided. The apparatus includes:

an acquiring module, configured to acquire a plurality of pieces of raw data, where modalities of the plurality of pieces of raw data are at least two of text, image, or video;

a determining module, configured to for raw data in the plurality of pieces of raw data, determine a key text of the raw data, determine a key of a key-value pair based on the key text, and determine a value of the key-value pair based on the raw data, where the key text is used for describing content of the raw data; and

a constructing module, configured to construct a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

[0010] According to a fifth aspect of embodiments of the present application, a cloud device is provided. The cloud device includes: a processor, a memory, and a computer program stored in the memory and capable of being run on the processor. When executing the computer program, the processor implements the knowledge base construction method according to the first aspect, the data retrieval method according to the second aspect and/or the data retrieval method according to the third aspect.

[0011] According to a sixth aspect of embodiments of the present application, a computer-readable storage medium is provided. The computer-readable storage medium stores computer-executable instructions. When being executed by a processor, the computer-executable instructions are used for performing the knowledge base construction method according to the first aspect, the data retrieval method according to the second aspect and/or the data retrieval method according to the third aspect.

[0012] According to a seventh aspect of embodiments of the present application, a computer program product is provided. The computer program product includes a computer program. The computer program is stored on a readable storage medium, at least one processor of an electronic device may read the computer program from the readable storage medium, and the at least one processor executes the computer program to enable the electronic device to perform the knowledge base construction method according to the first

aspect, the data retrieval method according to the second aspect and/or the data retrieval method according to the third aspect.

[0013] Embodiments of the present application are applied to an intelligent customer service scenario. The plurality of pieces of raw data are acquired, where the modalities of the plurality of pieces of raw data are at least two of text, image, or video. For the raw data in the plurality of pieces of raw data, the key text of the raw data is determined, the key of the key-value pair is determined based on the key text, and the value of the key-value pair is determined based on the raw data, where the key text is used for describing the content of the raw data. The question-and-answer knowledge base is constructed based on the key-value pairs corresponding to the plurality of pieces of raw data. In this way, a question-and-answer knowledge base that supports retrieval for different modalities of answers can be constructed.

BRIEF DESCRIPTION OF DRAWINGS

[0014] The accompanying drawings described herein are used to provide a further understanding of the present application, and form part of the present application. Exemplary embodiments of the present application and descriptions thereof are used to explain the present application, and do not constitute any inappropriate limitation to the present application. In the accompanying drawings:

FIG. 1 is a diagram of an application scenario according to an exemplary embodiment of the present application;

FIG. 2 is a flowchart of steps of a knowledge base construction method according to an exemplary embodiment of the present application;

FIG. 3 is a schematic diagram of a question-and-answer knowledge base according to an exemplary embodiment of the present application;

FIG. 4 is a schematic diagram of another question-and-answer knowledge base according to an exemplary embodiment of the present application;

FIG. 5 is a flowchart of steps of a data retrieval method according to an exemplary embodiment of the present application;

FIG. 6 is a flowchart of steps of another data retrieval method according to an

exemplary embodiment of the present application;

FIG. 7 is a structural diagram of a knowledge base construction apparatus according to an exemplary embodiment of the present application;

FIG. 8 is a schematic structural diagram of a cloud device according to an exemplary embodiment of the present application.

DESCRIPTION OF EMBODIMENTS

[0015] To make the objectives, technical solutions, and advantages of the present application clearer, the technical solutions in the present application are clearly and completely described below with reference to specific embodiments of the present application and the accompanying drawings. Apparently, the described embodiments are merely some embodiments rather than all the embodiments of the present application. All other embodiments obtained by a person of ordinary skill in the art based on embodiments of the present application without paying creative efforts shall fall within the protection scope of the present application.

[0016] An intelligent customer service is an automatic question-and-answer product. In the field of e-commerce, in the face of rich product knowledge of merchants and massive questions that users may ask, if high-quality answers need to be provided for the users, there needs to be sufficient data in a question-and-answer knowledge base. In related technologies, an index (a key of a key-value pair) construction method is to directly construct a query index of a field. However, for this method being applied in a scenario of multi-source knowledge of the intelligent customer service, as the multi-source knowledge of the intelligent customer service is usually of different modalities, and each modality includes different types of data, with development of services, an amount of raw data also increases. Therefore, it is particularly important to design a question-and-answer knowledge base that can support retrieval for different modalities and different types of answer data while ensuring that a knowledge base construction method is scalable and can be applied to new raw data.

[0017] Based on the foregoing background, a knowledge base construction method provided in embodiments of the present application includes: acquiring a plurality of pieces of

raw data, where modalities of the plurality of pieces of raw data are at least two of text, image, or video; for raw data in the plurality of pieces of raw data, determining a key text of the raw data, determining a key of a key-value pair based on the key text, and determining a value of the key-value pair based on the raw data, where the key text is used for describing content of the raw data; and constructing a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data. In the present application, a question-and-answer knowledge base that supports retrieval for different modalities of answers can be constructed, and the construction method is scalable. When there is new raw data, the method may be used to obtain a key-value pair of the new raw data, and the key-value pair is incorporated into the question-and-answer knowledge base.

[0018] In an embodiment, the overall knowledge base construction method may be implemented by using a cloud computing system. In addition, a server involved in the knowledge base construction method may be a cloud server, to leverage cloud resources for running various neural network models. In addition to being applied to cloud, the knowledge base construction method may also be applied to a server-side device, for example, a conventional server or a server array. The specific implementation is not limited herein.

[0019] FIG. 1 is a diagram of an application scenario according to the present application. Raw data of different modalities, for example, text, image, or video, is included. Text data is from different types of texts, for example, a dialog, a comment, and a graph. The raw data is processed to construct a question-and-answer knowledge base. After the question-and-answer knowledge base is online for use, when a query text of a user is received, the question-and-answer knowledge base may be searched for corresponding answer data, and the answer data is returned to the user, to respond to the query text of the user.

[0020] In the following, the technical solution provided in each embodiment of the present application will be described in detail with reference to the accompanying drawings.

[0021] FIG. 2 is a flowchart of steps of a knowledge base construction method according to an exemplary embodiment of the present application. The method may be applied to a server. As shown in FIG. 2, the knowledge base construction method specifically includes the following steps.

[0022] S201: acquire a plurality of pieces of raw data.

[0023] Modalities of the plurality of pieces of raw data are at least two of text, image, or video.

[0024] In an embodiment of the present application, it may be understood that a database includes a plurality of pieces of raw data. The plurality of pieces of raw data are from different data sources, and data of the data sources is of a text modality, an image modality, or a video modality. A data source of the text modality includes different types of data sources, for example, a user comment data source, a person-to-person dialog data source, a product attribute data source, a product knowledge graph data source, a product detail page data source, and a product home page data source. The user comment data source includes: a comment text made by a user on quality of a product, an express speed of the product, a service of a merchant, or the like. The person-to-person dialog data source includes: a dialog text between a user and a merchant manual customer service. The product attribute data source includes: text description for a product attribute (for example, category, size, color). The product knowledge graph data source includes: a knowledge graph in which a node represents a product, and an association between nodes represents an association between products. The product detail page data source includes images containing text content. The text content is used for describing a product, may be recognized by using an OCR (optical character recognition, optical character recognition) technology, and is used as raw data. The product home page data source includes images containing text content. The text content is also used for describing a product, may also be recognized by using the OCR technology, and is used as the raw data.

[0025] In an embodiment of the present application, the image may be, for example, an image including a product object, or may be an image including a text, for example, a product detail page or a product home page. The video is, for example, a product introduction video, a product use manner video, or a product installation/mounting manner video.

[0026] In addition, for a piece of raw data, for example, a comment text is a piece of raw data, a person-to-person dialog is a piece of raw data, a description about a product attribute is a piece of raw data, a knowledge graph is a piece of raw data, an entire text recognized in an image is a piece of raw data, an image is a piece of raw data, and a video is a piece of raw data.

[0027] A source of the raw data is not limited in embodiments of the present application.

[0028] S202: for raw data in the plurality of pieces of raw data, determine a key text of the raw data, determine a key of a key-value pair based on the key text, and determine a value of the key-value pair based on the raw data.

[0029] In an embodiment of the present application, at least one key text can be correspondingly determined for one piece of raw data. The key text is used for describing content of the raw data. The determination of the key text of the raw data includes: converting the raw data into a natural language text, and performing key text extraction on the natural language text to obtain the key text. The key text includes at least one of a content topic, a core opinion, a keyword, or a key entity of the raw data.

[0030] If the raw data is a text, the conversion of the raw data into a natural language text includes: processing, by using a prompt (a natural language model) template, the raw data to obtain the natural language text. The prompt template is pre-trained, and can convert the raw data into the natural language text. If the raw data is an image or a video (including a plurality of image frames), the image or the video may be input into a pre-trained multi-modality recognition model for processing, to obtain a natural language text describing the corresponding image or video.

[0031] In an embodiment of the present application, the performing key text extraction on the natural language text to obtain the key text includes: inputting the natural language text into a unified model to perform the key text extraction to obtain the key text.

[0032] In an embodiment of the present application, the unified model includes a plurality of pre-trained sub-models that facilitate the unification of a plurality of NLP (natural language) tasks, for example, a matching sub-model, a classification sub-model, or a sorting sub-model, or the like. The matching sub-model is used for implementing a matching task, the classification sub-model is used for implementing a classification task, and the sorting sub-model is used for implementing a sorting task. The natural language text is input into the unified model and processed by using the sub-models, to obtain at least one key text.

[0033] For example, core content, such as the content topic, the core opinion, the keyword, and the key entity of the raw data, can be extracted by using the unified model. For the content topic, for example, the raw data is about a product or quality. For the core opinion, for

example, the raw data is positive data or negative data. The key word, for example, is a key word in the raw data that describes an attribute of the content topic. The key entity, for example, is a location, time, a product name, or a product category.

[0034] In an optional embodiment, the key text may be used as the key of the key-value pair, and the natural language text or the raw data may be used as the value of the key-value pair, to obtain the key-value pair, so as to construct a question-and-answer knowledge base. One piece of raw data corresponds to a natural language text, a natural language text may correspond to one key, and the key is a key text or a combination of a plurality of key texts in the key texts of the raw data.

[0035] For example, refer to FIG. 3. The raw data is X, the natural text corresponding to the raw data is N, and the key text corresponding to the raw data includes a topic A, a core opinion B, a keyword C, a keyword D, a key entity E, and a key entity F. The key-value pairs included in the constructed question-and-answer knowledge base are shown in FIG. 3.

[0036] Further, the determination of the value of the key-value pair based on the raw data includes: segmenting the raw data based on the key text to obtain a data segment corresponding to the key text, where a modality of the data segment is the same as the modality of the raw data; and determining the value of the key-value pair based on the data segment.

[0037] In addition, if the raw data is a text, the segmentation of the raw data based on the key text to obtain a data segment corresponding to the key text includes: using a machine reading comprehension (MRC, Machine Reading Comprehension) technology to extract, from the raw data, a data segment describing the key text.

[0038] Specifically, the key text and the corresponding raw data are input into an understanding model based on the machine reading comprehension technology, and then the data segment describing the key text is output. The understanding model can remove redundant noise data from the raw data, to improve the overall quality of the data segment. In an embodiment of the present application, a key text has a corresponding data segment in the raw data.

[0039] For example, if the raw data is “parameters of this projector are as follows: the body weight is 1 kg, the zoom ratio is fixed focus, and the light source type is LED light

source", obtained key texts include: projector, weight, zoom ratio, fixed focus, light source type, and LED light source. A data segment corresponding to "projector" may be the entire raw data, a data segment corresponding to "weight" may be "the body weight is 1 kg", a data segment corresponding to "zoom ratio" and/or "fixed focus" is "the zoom ratio is fixed focus", a data segment corresponding to "light source type" and/or "LED light source" is "the light source type is LED light source".

[0040] In an embodiment of the present application, if the raw data is an image, the entire image may be used as the value of the key-value pair, or the raw data may be segmented, and a local image obtained through segmentation is a data segment used as the value of the key-value pair. The segmentation may be implemented by using image recognition or image processing technologies. For example, an object included in an image is a set of tableware, and a key text is "spoon", a local image containing a spoon in the image may be used as a data segment of the "spoon".

[0041] Further, if the raw data is a video, the entire video may be used as the value of the key-value pair, or a part of image frames in the video may be used as the value of the key-value pair, or a local image in an image in the video may be used as the value of the key-value pair. In embodiments of the present application, a specific segmentation method for the image or the video is not limited.

[0042] Further, the determination of the key of the key-value pair based on the key text includes: expanding the key text to obtain a first query text, and determining the first query text as the key of the key-value pair.

[0043] In an embodiment of the present application, one or more key texts may be expanded, to obtain the first query text. For example, for the raw data X, the corresponding key texts include the topic A, the core opinion B, the keyword C, the keyword D, the key entity E, and the key entity F. The topic A may be expanded, to obtain a corresponding first query text. Alternatively, the topic A and the core opinion B are expanded, to obtain a corresponding first query text. A plurality of first query texts may be correspondingly obtained for one piece of raw data. The first query text is determined as the key of the key-value pair, in other words, an index.

[0044] In the present application, the key text may be expanded by inputting the key text

into a pre-trained expansion model, so that the key text is expanded into a natural language query text. For example, if the topic A is a projector, and the keyword C is a weight, a first query text after expansion may be “what is the weight of the projector”. It may be understood that the first query text after expansion includes the key text.

[0045] Further, after the expanding the key text to obtain a query text, the method further includes: encoding the first query text to obtain an encoding vector, and determining the encoding vector as the key of the key-value pair.

[0046] In an embodiment of the present application, the first query text is encoded by using an encoder of a pre-trained BERT language model. The first query text is input into the encoder for encoding, and the encoding vector of the first query text is output.

[0047] It may be understood that, using the encoding vector as the key (that is, the index) of the key-value pair can effectively compensate for a poor effect of recalling semantic content by using a query text index.

[0048] S203: construct the question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

[0049] Further, the method further includes: performing data mining on the raw data to obtain target data; generating a second query text of the target data based on the target data; and determining the target data as the value of the key-value pair, and determining the second query text as the key of the key-value pair, to construct the question-and-answer knowledge base.

[0050] In an embodiment of the present application, data obtained through mining the raw data forms a key-value pair, and the key-value pair is added to the constructed question-and-answer knowledge base, to improve a data volume of the question-and-answer knowledge base. For example, user comment data is mined to obtain target data which is comment data, and a plurality of pieces of comment data form a comment knowledge base. Product attribute data is mined to obtain target data which is product data, and a plurality of pieces of product data form a product knowledge base. Other data is mined to obtain target data which is general data, and a plurality of pieces of general data form a general knowledge base. For the target data in the comment knowledge base, the product knowledge base, and the general knowledge base, a key of the target data is determined, and the target data is used

as a value, so that a key-value pair can be obtained and is added to the question-and-answer knowledge base constructed in the foregoing manner.

[0051] For example, FIG. 4 is a schematic diagram of a question-and-answer knowledge base obtained according to the present application. A key of the question-and-answer knowledge base may be a text or an encoding vector, and a value corresponding to the key may be a text, an image, or a video. For example, in FIG. 4, an encoding vector b is an encoding vector of a query text a, and the encoding vector b and the query text a both correspond to a same value which is a text g. An encoding vector d is an encoding vector of a query text c, and the encoding vector d and the query text c both correspond to a same value which is an image h. An encoding vector f is an encoding vector of a query text e, and the encoding vector f and the query text e both correspond to a same value which is a video k.

[0052] The question-and-answer knowledge base construction method provided in the present application can resolve a problem of uniform retrieval of different modalities of answer data, has strong scalability, and can effectively address the inflexibility of question-and-answer knowledge base construction methods in related technologies. In the present application, a billion-level question-and-answer knowledge base can be quickly constructed, thereby improving retrieval coverage for online query data.

[0053] FIG. 5 is a flowchart of steps of a data retrieval method according to an exemplary embodiment of the present application. The data retrieval method is applied to a server and specifically includes the following steps.

[0054] S501: receive query data sent by a terminal device.

[0055] The query data is one of a text, an image, or a video. If the query data is an image or a video, the image or the video may be converted into a query text, and the query text may describe content of the image or the video.

[0056] S502: retrieve, in a question-and-answer knowledge base, answer data corresponding to the query data.

[0057] A modality of the answer data is one of text, image, or video, and the question-and-answer knowledge base is constructed based on the foregoing knowledge base construction method.

[0058] Further, the retrieval of the answer data corresponding to the query data in the

question-and-answer knowledge base includes: determining a value of a key-value pair of which a key is determined based on the query data as the answer data; and/or,

encoding based on the query data to obtain a query encoding vector; and determining, in the question-and-answer knowledge base, a target encoding vector whose similarity with the encoding vector is greater than a threshold; and determining, in the question-and-answer knowledge base, a value of a key-value pair of which a key is the target encoding vector as the answer data.

[0059] Specifically, the determination of a value of a key-value pair of which a key is determined based on the query data as the answer data includes: if the query data sent by the terminal device is text, retrieving a corresponding value in the question-and-answer knowledge base as the answer data by using the query data as the key; if retrieval of the corresponding value is failed, encoding the query data to obtain an encoding vector, and retrieving the corresponding value in the question-and-answer knowledge base as the answer data by using the encoding vector as the key.

[0060] Further, if the query data is an image or a video, extraction may be performed on the image or the video, to obtain a corresponding query text, and then the query text is used as the key or an encoding vector of the query text is used as the key, to retrieve a corresponding value as the answer data.

[0061] S503: send the answer data to the terminal device.

[0062] In an embodiment of the present application, the server may retrieve the answer data of the query data based on the knowledge base, to provide a high-quality answer for the user.

[0063] FIG. 6 is a flowchart of steps of another data retrieval method according to an exemplary embodiment of the present application. The data retrieval method is applied to a terminal device and specifically includes the following steps.

[0064] S601: send query data to a server.

[0065] The query data is one of a text, an image, or a video.

[0066] S602: receive answer data that is sent by the server and that is of the query data.

[0067] The answer data is determined based on the foregoing data retrieval method.

[0068] For a specific implementation process of this embodiment, reference may be made

to the foregoing embodiments. Details are not described herein again.

[0069] In an embodiment of the present application, refer to FIG. 7. In addition to a knowledge base construction method, a knowledge base construction apparatus 70 is further provided. The knowledge base construction apparatus 70 includes:

an acquiring module 71, configured to acquire a plurality of pieces of raw data, where modalities of the plurality of pieces of raw data are at least two of text, image, or video;

a determining module 72, configured to, for raw data in the plurality of pieces of raw data, determine a key text of the raw data, determine a key of a key-value pair based on the key text, and determine a value of the key-value pair based on the raw data, where the key text is used for describing content of the raw data; and

a constructing module 73, configured to construct a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

[0070] In an optional embodiment, the determining module 72 is specifically configured to: convert the raw data into a natural language text; and perform key text extraction on the natural language text to obtain the key text, where the key text includes at least one of a content topic, a core opinion, a keyword, or a key entity of the raw data.

[0071] In an optional embodiment, when performing the key text extraction on the natural language text to obtain the key text, the determining module 72 is specifically configured to: input the natural language text into a unified model to perform the key text extraction to obtain the key text.

[0072] In an optional embodiment, the determining module 72 is specifically configured to: segment the raw data based on the key text to obtain a data segment corresponding to the key text, where a modality of the data segment is the same as the modality of the raw data; and determine the value of the key-value pair based on the data segment.

[0073] In an optional embodiment, if the raw data is a text, when segmenting the raw data based on the key text to obtain the data segment corresponding to the key text, the determining module 72 is specifically configured to: use a machine reading comprehension technology to extract, from the raw data, a data segment describing the key text.

[0074] In an optional embodiment, the determining module 72 is specifically configured to: expand the key text to obtain a first query text; and determine the first query text as the

key of the key-value pair.

[0075] In an optional embodiment, the determining module 72 is further configured to: encode the first query text to obtain an encoding vector; and determine the encoding vector as the key of the key-value pair.

[0076] In an optional embodiment, the determining module 72 is further configured to: perform data mining on the raw data to obtain target data; generate a second query text of the target data based on the target data; and determine the target data as the value of the key-value pair, and determine the second query text as the key of the key-value pair, to construct the question-and-answer knowledge base.

[0077] In an embodiment of the present application, a data retrieval apparatus (not shown) is further provided. The data retrieval apparatus includes:

a receiving module, configured to receive query data sent by a terminal device, where the query data is one of a text, an image, or a video;

a retrieving module, configured to retrieve, in a question-and-answer knowledge base, answer data corresponding to the query data, where a modality of the answer data is one of text, image, or video, and the question-and-answer knowledge base is constructed based on the foregoing knowledge base construction method; and

a sending module, configured to the answer data to the terminal device.

[0078] In an optional embodiment, the retrieving module is specifically configured to: determine a value of a key-value pair of which a key is determined based on the query data as the answer data; and/or,

encode based on the query data to obtain a query encoding vector; determine, in the question-and-answer knowledge base, a target encoding vector whose similarity with the encoding vector is greater than a threshold; and determine, in the question-and-answer knowledge base, a value of a key-value pair of which a key is the target encoding vector as the answer data.

[0079] In an embodiment of the present application, another data retrieval apparatus (not shown) is further provided. The data retrieval apparatus is applied to a terminal device. The data retrieval apparatus includes:

a sending module, configured to send query data to a server; and

a receiving module, configured to receive answer data that is sent by the server and that is of the query data, where the answer data is determined based on the foregoing data retrieval method.

[0080] In an embodiment of the present application, the knowledge base construction apparatus is provided and applied to an intelligent customer service scenario. The plurality of pieces of raw data are acquired, where the modalities of the plurality of pieces of raw data are at least two of text, image, or video. For the raw data in the plurality of pieces of raw data, the key text of the raw data is determined, the key of the key-value pair is determined based on the key text, and the value of the key-value pair is determined based on the raw data, where the key text is used for describing the content of the raw data. The question-and-answer knowledge base is constructed based on the key-value pairs corresponding to the plurality of pieces of raw data. In this way, a question-and-answer knowledge base that supports retrieval for different modalities of answers can be constructed.

[0081] In addition, some procedures described in the foregoing embodiments and accompanying drawings include a plurality of operations that appear in a specific sequence. However, it should be clearly understood that the operations may not be performed in the sequence in which the operations appear in this specification or be performed in parallel. The sequence numbers are merely used for distinguishing different operations, and do not indicate any execution sequence. In addition, the procedures may include more or fewer operations, and the operations may be performed in sequence or in parallel. It should be noted that, descriptions such as “first” and “second” in this specification are used to distinguish different messages, devices, modules, and the like, do not indicate a sequence, and do not constitute a limitation to different types of “first” and “second”.

[0082] FIG. 8 is a schematic structural diagram of a cloud device 80 according to an embodiment of the present application. The cloud device 80 is configured to run the foregoing knowledge base construction method or image processing method. As shown in FIG. 8, the cloud device includes: a memory 84 and a processor 85.

[0083] The memory 84 is configured to store a computer program, and may be configured to store various other information to support operations on the cloud device. The memory 84 may be an object storage service (Object Storage Service, OSS).

[0084] The memory 84 may be implemented by any type of volatile or non-volatile storage device or a combination thereof, such as a static random access memory (SRAM), an electrically erasable programmable read-only memory (EEPROM), an erasable programmable read-only memory (EPROM), a programmable read-only memory (PROM), a read-only memory (ROM), a magnetic memory, a flash memory, a magnetic disk, or an optical disc.

[0085] The processor 85, coupled with the memory 84, is configured to execute the computer program in the memory 84, to: acquire a plurality of pieces of raw data, where modalities of the plurality of pieces of raw data are at least two of text, image, or video; for raw data in the plurality of pieces of raw data, determine a key text of the raw data, determine a key of a key-value pair based on the key text, and determine a value of the key-value pair based on the raw data, where the key text is used for describing content of the raw data; and construct a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

[0086] Further optionally, when determining the key text of the raw data, the processor 85 is specifically configured to: convert the raw data into a natural language text; and perform key text extraction on the natural language text to obtain the key text, where the key text includes at least one of a content topic, a core opinion, a keyword, or a key entity of the raw data.

[0087] Further optionally, when performing the key text extraction on the natural language text to obtain the key text, the processor 85 is specifically configured to: input the natural language text into a unified model to perform the key text extraction to obtain the key text.

[0088] Further optionally, when determining the value of the key-value pair based on the raw data, the processor 85 is specifically configured to: segment the raw data based on the key text to obtain a data segment corresponding to the key text, where a modality of the data segment is the same as the modality of the raw data; and determine the value of the key-value pair based on the data segment.

[0089] Further optionally, if the raw data is a text, when segmenting the raw data based on the key text to obtain the data segment corresponding to the key text, the processor 85 is specifically configured to: use a machine reading comprehension technology to extract, from

the raw data, a data segment describing the key text.

[0090] Further optionally, after expanding the key text to obtain a query text, the processor 85 is further configured to: encode the first query text to obtain an encoding vector; and determine the encoding vector as the key of the key-value pair.

[0091] Further optionally, the processor 85 is further configured to: perform data mining on the raw data to obtain target data; generate a second query text of the target data based on the target data; and determine the target data as the value of the key-value pair, and determine the second query text as the key of the key-value pair, to construct the question-and-answer knowledge base.

[0092] In an optional embodiment, the processor 85, coupled with the memory 84, is configured to execute the computer program in the memory 84, to: receive query data sent by a terminal device, where the query data is one of a text, an image, or a video; retrieve, in a question-and-answer knowledge base, answer data corresponding to the query data, where a modality of the answer data is one of text, image, or video, and the question-and-answer knowledge base is constructed based on any of the foregoing knowledge base construction methods; and send the answer data to the terminal device.

[0093] Further optionally, when retrieving, in the question-and-answer knowledge base, the answer data corresponding to the query data, the processor 85 is specifically configured to: determine a value of a key-value pair of which a key is determined based on the query data as the answer data; and/or,

encode based on the query data to obtain a query encoding vector; determine, in the question-and-answer knowledge base, a target encoding vector whose similarity with the encoding vector is greater than a threshold.

[0094] In an optional embodiment, the processor 85, coupled with the memory 84, is configured to execute the computer program in the memory 84, to: send query data to a server, and receive answer data that is sent by the server and that is of the query data, where the answer data is determined based on the foregoing data retrieval method.

[0095] Further, as shown in FIG. 8, the cloud device further includes: a firewall 81, a load balancer 82, a communication component 86, a power supply component 83, and the like. Only some components are schematically provided in FIG. 8, which does not mean that the

cloud device includes only the components shown in FIG. 8.

[0096] The cloud device provided in this embodiment of the present application can obtain a compressed visual network model. The compressed visual network model occupies a relatively small memory and has relatively fast calculation efficiency without affecting recognition accuracy.

[0097] Correspondingly, in an embodiment of the present application, a computer-readable storage medium storing a computer program is further provided. When the computer program/instructions is/are executed by a processor, the processor is enabled to implement the steps in the methods shown above.

[0098] Correspondingly, in an embodiment of the present application, a computer program product including a computer program/instructions is further provided. When the computer program/instructions is/are executed by a processor, the processor is enabled to implement the steps in the methods shown above.

[0099] The communication component in FIG. 8 is configured to facilitate communication in a wired or wireless manner between a device in which the communication component is located and another device. The device in which the communication component is located may access a communication standard-based wireless network, for example, Wi-Fi, 2G, 3G, 4G/LTE, 5G, or another mobile communication network, or a combination thereof. In an exemplary embodiment, the communication component receives a broadcast signal or a broadcast-related text from an external broadcast management system through a broadcast channel. In an exemplary embodiment, the communication component further includes a near field communication (NFC) module, to promote short-range communication. For example, the NFC module may be implemented based on a radio frequency identification (RFID) technology, an infrared information association (IrDA) technology, an ultra-wideband (UWB) technology, a Bluetooth (BT) technology, and another technology.

[0100] The power supply component in FIG. 8 provides power to various components of a device in which the power supply component is located. The power supply component may include a power supply management system, one or more power supplies, and another component associated with generating, managing, and allocating power for the device in which the power supply component is located.

[0101] In several embodiments provided in the present application, it should be understood that the disclosed system and method may be implemented in another manner. For example, the system embodiment described above is merely an example. For example, division into units is merely logical function division and during actual implementation, there may be another division manner. For example, a plurality of units or components may be combined or integrated into another system, or some features may be ignored or not performed. In addition, the displayed or discussed mutual couplings or direct couplings or communication connections may be implemented by using some interfaces. The indirect couplings or communication connections between systems or units may be implemented in electronic, mechanical, or another form.

[0102] The units described as separate components may or may not be physically separated, and components displayed as units may or may not be physical units, in other words, may be located in one position, or may be distributed on a plurality of network units. Some or all of the units may be selected according to actual needs to achieve the objectives of the solutions of the embodiments.

[0103] In addition, functional units in embodiments of the present application may be integrated into one processing unit, or each of the units may be physically separated, or two or more units may be integrated into one unit. The foregoing integrated unit can be implemented in the form of hardware or in the form of hardware plus a software functional unit.

[0104] The foregoing integrated unit implemented in the form of a software functional unit may be stored in a computer-readable storage medium. The foregoing software functional unit is stored in a storage medium, including a plurality of instructions for enabling a computer device (which may be a personal computer, a server, a network device, or the like) or a processor to perform some steps of the method in each embodiment of the present application. The foregoing storage medium includes any medium that can store program code, for example, a USB flash drive, a removable hard disk, a read-only memory (read-only memory, ROM), a random access memory (random access memory, RAM), a magnetic disk, or an optical disc.

[0105] A person skilled in the art can clearly understand that, for convenience and ease of description, the division of the foregoing functional modules is only used as an example. In

actual applications, the foregoing functions can be assigned to different functional modules as needed. In other words, an internal structure of the system can be divided into different functional modules to complete all or part of the functions described above. For the specific operating process of the system described above, reference may be made to the corresponding process in the aforementioned method embodiments. Details are not described herein again.

[0106] After considering the specification and practicing the invention disclosed herein, a person skilled in the art may easily conceive of other implementations of the present application. The present application is intended to cover any variations, uses, or adaptive changes of the present application. These variations, uses, or adaptive changes follow the general principles of the present application and include common general knowledge or common technical means in the art, which are not disclosed in the present application. The specification and embodiments are only regarded as exemplary. The true scope and spirit of the present application are indicated by the following claims.

[0107] It should be understood that the present application is not limited to the precise structures described above and shown in the accompanying drawings, and various modifications and changes can be made without departing from the scope of the present application. The scope of the present application is limited only by the attached claims.

CLAIMS

1. A knowledge base construction method, comprising:

acquiring a plurality of pieces of raw data, wherein modalities of the plurality of pieces of raw data are at least two of text, image, or video;

for raw data in the plurality of pieces of raw data, determining a key text of the raw data, determining a key of a key-value pair based on the key text, and determining a value of the key-value pair based on the raw data, wherein the key text is used for describing content of the raw data; and

constructing a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

2. The knowledge base construction method according to claim 1, wherein the determining a key text of the raw data comprises:

converting the raw data into a natural language text; and

performing key text extraction on the natural language text to obtain the key text, wherein the key text comprises at least one of a content topic, a core opinion, a keyword, or a key entity of the raw data.

3. The knowledge base construction method according to claim 2, wherein the performing key text extraction on the natural language text to obtain the key text comprises:

inputting the natural language text into a unified model to perform the key text extraction to obtain the key text.

4. The knowledge base construction method according to any one of claims 1 to 3, wherein the determining a value of the key-value pair based on the raw data comprises:

segmenting the raw data based on the key text to obtain a data segment corresponding to the key text, wherein a modality of the data segment is the same as the modality of the raw data; and

determining the value of the key-value pair based on the data segment.

5. The knowledge base construction method according to claim 4, wherein if the raw data is a text, the segmenting the raw data based on the key text to obtain a data segment corresponding to the key text comprises:

using a machine reading comprehension technology to extract, from the raw data, a data segment describing the key text.

6. The knowledge base construction method according to any one of claims 1 to 3,

wherein the determining a key of a key-value pair based on the key text comprises:

expanding the key text to obtain a first query text; and

determining the first query text as the key of the key-value pair.

7. The knowledge base construction method according to claim 6, wherein after the expanding the key text to obtain a query text, the method further comprises:

encoding the first query text to obtain an encoding vector; and

determining the encoding vector as the key of the key-value pair.

8. The knowledge base construction method according to any one of claims 1 to 3, wherein the method further comprises:

performing data mining on the raw data to obtain target data;

generating a second query text of the target data based on the target data; and

determining the target data as the value of the key-value pair, and determining the second query text as the key of the key-value pair, to construct the question-and-answer knowledge base.

9. A data retrieval method, comprising:

receiving query data sent by a terminal device, wherein the query data is one of a text, an image, or a video;

retrieving, in a question-and-answer knowledge base, answer data corresponding to the query data, wherein a modality of the answer data is one of text, image, or video, and the question-and-answer knowledge base is constructed based on the knowledge base construction method according to any one of claims 1 to 8; and

sending the answer data to the terminal device.

10. The data retrieval method according to claim 9, wherein the retrieving, in a question-and-answer knowledge base, answer data corresponding to the query data comprises:

determining a value of a key-value pair of which a key is determined based on the query data as the answer data;

and/or

encoding based on the query data to obtain a query encoding vector;

determining, in the question-and-answer knowledge base, a target encoding vector whose similarity with the encoding vector is greater than a threshold; and

determining, in the question-and-answer knowledge base, a value of a key-value pair of which a key is the target encoding vector as the answer data.

11. A data retrieval method, applied to a terminal device, wherein the data retrieval method comprises:

sending query data to a server; and

receiving answer data that is sent by the server and that is of the query data, wherein the answer data is determined based on the data retrieval method according to claim 9 or 10.

12. A knowledge base construction apparatus, comprising:

an acquiring module, configured to acquire a plurality of pieces of raw data, wherein modalities of the plurality of pieces of raw data are at least two of text, image, or video;

a determining module, configured to, for raw data in the plurality of pieces of raw data, determine a key text of the raw data, determine a key of a key-value pair based on the key text, and determine a value of the key-value pair based on the raw data, wherein the key text is used for describing content of the raw data; and

a constructing module, configured to construct a question-and-answer knowledge base based on key-value pairs corresponding to the plurality of pieces of raw data.

13. A cloud device, comprising: a processor, a memory, and a computer program stored on the memory and capable of being run on the processor, wherein when executing the computer program, the processor implements the knowledge base construction method according to any one of claims 1 to 8 and/or the data retrieval method according to any one of claims 9 to 11.