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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 November 2020 (12.11.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/226920 A1

(51) International Patent Classification:
G06F 40/35 (2020.01) *G06F 16/332* (2019.01)

(21) International Application Number:
PCT/US2020/030017

(22) International Filing Date:
27 April 2020 (27.04.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201910378798.6 08 May 2019 (08.05.2019) CN

(71) Applicant: **MICROSOFT TECHNOLOGY LICENSING, LLC** [US/US]; One Microsoft Way, Redmond, Washington 98052-6399 (US).

(72) Inventors: **FU, Hao**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **SONG, Ruihua**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US). **ZENG, Min**; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 98052-6399 (US).

(74) Agent: **SWAIN, Cassandra T.** et al.; MICROSOFT TECHNOLOGY LICENSING, LLC, One Microsoft Way, Redmond, Washington 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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: ELECTRONIC DEVICE AND METHOD FOR GENERATING CONTENT CONTAINING SPECIFIC WORD

(57) Abstract: The present disclosure provides a method for generating a content containing a specific word and an electronic device, the method includes: an electronic device pre-generating a first content while interacting with a user; for each specific word in a specific word stock: embedding the specific word into the first content to obtain a corresponding second content, and obtaining a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and outputting at least one of a plurality of the second contents according to the evaluation results, wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

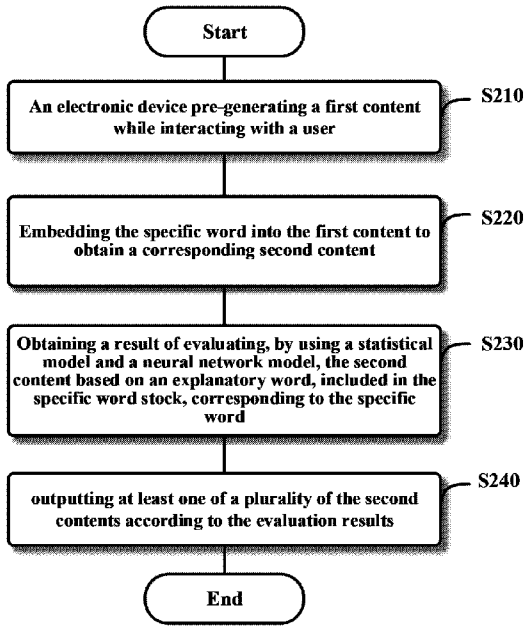


FIG 2

WO 2020/226920 A1

[Continued on next page]

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))*

ELECTRONIC DEVICE AND METHOD FOR GENERATING CONTENT CONTAINING SPECIFIC WORD

Technical Field

[0001] The present disclosure relates to article intelligence (AI) technology and more particularly to an electronic device and a method for generating a content containing a specific word.

Background Art

[0002] With the development of information technology and artificial intelligence, the chatbot are widely used in multi-modal human-computer interaction. For example, the chatbot may directly talk to the user, wherein after receiving information input by the user, the chatbot may analyze the information through a chatting neural network model , and generate a response sentence from the existing chat corpus according to the analyzed result.

[0003] Catchphrases with high timeliness often appear on popular social networks, and these catchphrases are popular with many users over a period of time. In addition, groups of specific fields may prefer to use certain hot words that have specific meanings in this field when communicating online or offline with each other. For example, many game players of a popular online game may like to use the name of a character in the game to express the meaning of the ability that the character represents. If the chatbot outputs a sentence containing such a specific word while chatting with the user, the user will feel intimate.

Summary

[0004] According to an aspect of the present disclosure, there is provided a method for generating a content containing a specific word, comprising: an electronic device pre-generating a first content while interacting with a user; for each specific word in a specific word stock: embedding the specific word into the first content to obtain a corresponding second content, and obtaining a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and outputting at least one of a plurality of the second contents according to the evaluation results, wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model

[0005] According to an aspect of the present disclosure, there is provided an electronic device, comprising: a processor; and a memory connected to the processor and storing

instructions that, when executed by the processor, cause the electronic device to: pre-generate a first content while interacting with a user; for each specific word in a specific word stock: embed the specific word into the first content to obtain a corresponding second content, and obtain a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and output at least one of a plurality of the second contents according to the evaluation results, wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

[0006] According to an embodiment of the present disclosure, a specific word stock is generated by the neural network model learning the specific word such as catchphrases on the social networks and the hot words in the specific fields regularly (for example, every day, every week, etc.) based on the specific corpus. When the electronic device interacts with the user, the neural network model and the specific word stock may be used to generate and output a content containing the specific words, such that the user feels intimate, thereby improving the user experience.

Brief Description of the Drawings

[0007] The above and other aspects and features of the present disclosure will be more clearly understood from the following detailed description of the appended claims.

[0008] Fig.1 schematically illustrates an application scenario according to an embodiment of the present disclosure;

[0009] Fig.2 is a flowchart that schematically illustrates a method for generating content containing a specific word according to an embodiment of the present disclosure;

[0010] Fig.3 schematically illustrates a user interface comparison before and after replacing a text span with a specific word according to an embodiment of the present disclosure;

[0011] Fig.4 is a block diagram schematically illustrating an electronic device according to an embodiment of the present disclosure.

Detailed Description

[0012] The detailed description made hereinafter with reference to the accompanying drawings will help to understand various embodiments of the present disclosure. The following detailed description includes various specific matters for understanding the present disclosure, but these matters should be considered as simple examples. Accordingly, those skilled in the art will recognize that various changes and modifications may be made to the various embodiments described herein without departing from the scope and sprit of

the disclosure. In addition, well-known functions or components are not described in detail for clarity and conciseness.

[0013] Terms such as "first" and "second" used herein may be used to interpret various elements, and these elements are not limited by these terms. These terms may only be used to distinguish one element from another element. In addition, the singular forms used herein are intended to include the plural forms, unless the context clearly indicates otherwise.

[0014] The terms used in the detailed description is for the purpose of description, and is not intended to limit the present disclosure. All terms used herein have the same meaning as commonly understood by those skilled in the art unless they are otherwise defined.

[0015] Fig.1 schematically illustrates an application scenario according to an embodiment of the present disclosure.

[0016] A chatbot terminal application may be installed in the electronic device 101 in Fig.1. After the application is executed, the electronic device 101 may interact (e.g., chatting) with the user through the application. As shown in part A of Fig.1, the electronic device 101 may pre-generate the content "how to make money: what, how and where to sell?" (indicated by a dashed box) according to the context of chatting with the user.

[0017] By applying the method of the present disclosure (described in detail later), the catchphrase "Here comes the question" on the social networks may be extracted from a specific word stock and embed it into the pre-generated content. For example, you may replace "what" in the pre-generated content with the extracted "Here comes the question", so as to obtain the content "how to make money: here comes the question, how and where to sell?", the content may be output to the user interface of the electronic device 101 as shown in part B of Fig.1.

[0018] As used herein, "specific word" may refer to a term with high timeliness, such as a catchphrase on the social networks, a hot word in a specific area, etc., which are generally given meanings that are different from what are conventionally understood. For example, for a certain hot word that appears in a novel, game, or movie for a period of time, only those who understand the corresponding novel, game, or movie understand the meaning of the hot word.

[0019] Catchphrases may include many types of Internet slang, slogans, and snowclones. Internet slang refers to various slang obtained from the Internet, which may be intentional typo (for example, "teh" is a typo of "the") or an acronym (for example, "TL" or "DR" is the abbreviation of "Too long or "Did not read"). A slogan is a phrase that is often used online or offline, such as "You are the chosen one" which is often used as an independent

phrase that is different from the Internet slang. Snow clones are modern phrases that have been adapted from old sayings, famous epigrams, culture catchphrases, etc.

[0020] A method for generating a content containing a specific word according to an embodiment of the present disclosure will be described in detail below with reference to a flowchart shown in Fig. 2.

[0021] In step S210, the electronic device may pre-generate the first content when interacting with the user. As described above, a chatbot terminal application may be installed in the electronic device. After the application is executed, the electronic device may interact with the user through the application and may pre-generate the first content based on, for example, the context of the interaction with the user.

[0022] In step S220, for each specific word in the specific word stock, the specific word may be embedded into the first content to obtain the corresponding second content. The specific word stock may be generated and periodically updated based on a specific corpus by using a neural network model. For example, a specific corpus may include a corpus on the social networks, a corpus on a network communication platform in a specific field, a corpus on other web medias, and the like. Networked servers or electronic devices may use a neural network model to perform machine learning based on a specific corpus regularly (e.g., every day, every week, etc.), capture specific words such as catchphrases on social networks, hot words in specific fields, and generate or update the specific word stock.

[0023] According to an embodiment of the present disclosure, the specific word stock may include a plurality of specific word directories, and each of the specific word directories may include a specific word, an explanatory word corresponding to the specific word, an associated word of the specific word, and the like. The term "associated word" herein refers to the word adjacent to the corresponding specific word in the specific corpus, which may be captured from a specific corpus by using a neural network model. The term "explanatory word" herein is not necessarily a synonym or a parasynonym of the corresponding specific word, but may be any word or phrase that helps the neural network model to understand the corresponding specific word. For example, the explanatory word corresponding to the specific word may be determined or obtained by any one of the following manners by using a statistical model of the text or a neural network model.

[0024] Manner 1: for each specific word in the specific word stock, the statistical model of the text is used to determine, as the explanatory word corresponding to the specific word, a word that has a frequency of co-occurring with the specific word in the specific corpus being higher than a first threshold (which may be preset according to the needs and/or

arbitrarily based on the experience).

[0025] For example, for the catchphrase "waifu" in a specific word stock, it may find by using the statistical model of the text that the words "cute", "girl" and "date" have a high frequency co-occurring with "waifu" in the specific corpus, then the words "cute", "girl" and "date" are determined as the explanatory words of the catchphrase "waifu".

[0026] Manner 2: for each specific word in the specific word stock, the statistical model of the text is used to determine the explanatory word corresponding to the specific word based on the specific corpus and word-level local coherence.

[0027] The word-level local coherence refers to whether a word forms a feasible phrase with the immediately preceding word or the immediately following word. If a specific word shares some common viable adjacent words with a certain word, it means that they may be similar in syntax and semantics. Taking a sentence "have a date with my waifu" containing the specific word "waifu" in the specific corpus as an example, the "date with my..." and "have" in the sentence is a feasible adjacent word and phrase of the specific word "waifu". In order to obtain the explanatory word of the specific word "waifu", the statistical model of the text may be used to find words which are adjacent to the words and phrase "date with my..." and "have" in the specific corpus. For example, it is possible to find words such as "friends", "wife" and "husband". The found words may be determined as the explanatory words of "waifu".

[0028] Manner 3: for each specific word in the specific word stock, a topic model based on statistics is used to determine a topic of the specific word based on the specific corpus by using a topic model based on statistics as the explanatory word corresponding to the specific word.

[0029] Certain specific words belong to a specific topic, for example, the specific word "waifu" is related to topics such as anime or video games. Determining the topic for the specific word as the explanatory word of the specific word may help the neural network model to understand the specific word more accurately.

[0030] Manner 4: for each specific word in the specific word stock, the neural network model is used to predict a plurality of words, of which each may replace the specific word in a sentence containing the specific word in the specific corpus, and to select a word from the plurality of words having a prediction score higher than a second threshold (which may be preset according to the needs and/or arbitrarily based on the experience) as the explanatory word corresponding the specific word.

[0031] Taking the sentence "have a date with my waifu" containing the specific word

"waifu" as an example, in order to obtain the explanatory word of the specific word "waifu", the neural network model may be used to mask "waifu" in the sentence, such that the sentence becomes "have a date with my □□□", and a plurality of words appears in the position of the "□□□" based on the context. The word with a higher prediction score may be selected from a plurality of predicted words as the explanatory word of the specific word "waifu".

[0032] According to an embodiment of the present disclosure, any combination of the above manners 1 to 4 may also be used to determine or obtain the explanatory word corresponding to the specific word.

[0033] According to an embodiment of the present disclosure, the embedding of the specific word into the first content may include replacing a text span in the first content with the specific word. In Fig.3, it takes the specific word "waifu" and the text span "friend" as an example, and schematically shows the user interface comparison before and after replacing the text span with the specific word.

[0034] According to an embodiment of the present disclosure, no punctuation mark is included in the text span. The punctuation often plays an important role in the sentence structure. if the punctuation is changed, it may change the sentence structure and cause syntax errors. Therefore, after pre-generating the first content, only the text span that does not contain punctuation is replaced with the specific word, so it may reduce the replacements which may result in syntax errors .

[0035] According to an embodiment of the present disclosure, before replacing the text span in the first content with the specific word, whether the adjacent word of the text span in the first content includes at least one of the associated words of the specific word may be determined according to the specific word stock. If the determined result is yes, the replacement may be performed; otherwise, the replacement may be omitted.

[0036] In addition, in a specific corpus, certain specific words may often appear at the beginning or end of a sentence, or even act as an independent sentence. Such specific words may be identified by counting such usage in the specific corpus. It is conceivable to replace the text span at the beginning or end of the first content with such specific word.

[0037] According to an embodiment of the present disclosure, the electronic device 101 shown in Part A of Fig.1 pre-generates the first content "How to make money: what, how and where to sell?" is provided as an example, when a specific word stock contains the three specific words "Pinru's dress (品如的衣服)", "Khorium (氪金)" and "Here comes the

question", the electronic device 101 obtains a plurality of second contents such as "How to make money: Pinru's dress, how and where to sell?", "Khorium: what, how and where to sell?" and "How to make money: here comes the question, how and where to sell?" through the above step S220.

[0038] In step S230, for each specific word in the specific word stock, it may obtain a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word. For example, the evaluation is performed by the electronic device or a server connected thereto based on at least one of word-level co-occurrence, local coherence, topic, and contextual word embedding in the second content.

[0039] In step S240, at least one of a plurality of the second contents may be output according to the evaluation results. For example, according to the evaluation results of the above plurality of second contents obtained based on the first content "How to make money: what, how and where to sell?" shown in part A of Fig. 1, it is possible to output "How to make money: here comes the question, how and where to sell?" as shown in part B of Fig.1.

[0040] FIG. 4 schematically illustrates a block diagram of an electronic device 400 according to an embodiment of the present disclosure.

[0041] The electronic device 400 may include a memory 410 and a processor 420. The memory 410 may store program instructions for executing the method for generating a content containing a specific word described with reference to Fig. 2.

[0042] The memory 410 may be a volatile memory (e.g., a register, a cache, a random access memory (RAM), or the like), or a non-volatile memory (e.g., a read only memory (ROM), an electrically erasable programmable read only memory (EEPROM), a flash memory, or the like) or some combination thereof.

[0043] The processor 420 may be a physical or virtual processor, and may implement a method for generating a content containing a specific word according to an embodiment of the present disclosure by executing program instructions stored in the memory 410. The processor 420 may include a central processing unit (CPU), a microprocessor, a controller, a microcontroller, and the like.

[0044] In some embodiments, the electronic device 400 may be implemented as a variety of user terminals or service terminals. The service terminal may be a server, and a large electronic device provided by various service parties. The user terminal may be, for example, any type of mobile terminal, fixed terminal or portable terminal, such as a mobile phone, a multimedia computer, a multimedia tablet, an internet node, a communicator, a desktop

computer, a laptop computer, a notebook computer, a netbook computer, a tablet computer, a Personal Communication System (PCS) device, a personal navigation device, a Personal Digital Assistant (PDA), an audio/video player, a digital camera/camcorder, a positioning device, a television receiver, a radio broadcast receiver, an e-book device, a gaming device, or any combinations, including accessories and peripherals for these devices, or any combination thereof.

[0045] In some embodiments, the electronic device 410 may include a communication module 430, a power module 440, an audio module 450, a display 460, a bus 470, a camera module 480, or the like.

[0046] The communication module 430 may be configured to facilitate wired or wireless communication between the electronic device 400 and other devices. The electronic device 400 may access a wireless network based on a communication standard, such as WiFi, LTE, 5G and the like. In an exemplary embodiment, the communication module 430 may include a near field communication (NFC) module to perform a short-range communication.

[0047] The power module 440 provides power to various components of the electronic device 400. The power module 440 may include a power management system, one or more batteries, and a wireless or wired charger.

[0048] The audio module 450 is configured to output and/or input audio signals. For example, the audio module 450 includes a microphone (MIC) .

[0049] The display 460 may include a liquid crystal display (LCD) and a touch panel, wherein the touch panel may comprise one or more touch sensors to sense gestures on the touch panel.

[0050] The memory 410, the processor 420, the communication module 430, the power module 440, the audio module 450, the display 460, and the camera module 480 may be connected with the bus 470. The bus 470 may provide an interface between the processor 420 and the remaining components of electronic device 400. In addition, the bus 470 may also provide an interface for each component of the electronic device 400 to access the memory 410 and an interface for the various components to mutually access each other.

[0051] The functions described above herein may be performed, at least in part, by one or more hardware logic devices. For example, available hardware logic devices include Field Programmable Gate Arrays (FPGA), Application Specific Integrated Circuits (ASICs), Application Specific Standard Products (ASSP), System on Chips (SOC), Complex Programmable Logic Devices (CPLDs), and the like.

Exemplary clause

[0052] According to an aspect of the present disclosure, there is provided a method for generating content containing specific word, comprising: an electronic device pre-generating a first content while interacting with a user; for each specific word in a specific word stock: embedding the specific word into the first content to obtain a corresponding second content, and obtaining a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and outputting at least one of a plurality of the second contents according to the evaluation results, wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

[0053] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine, as the explanatory word corresponding to the specific word, a word with a frequency, higher than a first threshold, of co-occurring with the specific word in the specific corpus.

[0054] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine the explanatory word corresponding to the specific word based on the specific corpus and word-level local coherence.

[0055] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine a topic of the specific word based on the specific corpus as the explanatory word corresponding to the specific word.

[0056] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the neural network model is used to predict a plurality of words each of which can replace the specific word in a sentence containing the specific word in the specific corpus, and to select from the plurality of words a word with a prediction score higher than a second threshold as the explanatory word corresponding the specific word.

[0057] According to an embodiment of the present disclosure, the embedding of the specific word into the first content comprises: replacing a text span in the first content with the specific word.

[0058] According to an embodiment of the present disclosure, no punctuation mark is included in the text span.

[0059] According to an embodiment of the present disclosure, wherein, for each specific word in the specific word stock, a word immediately adjacent to the specific word in the

specific corpus is stored in the specific word stock as an associated word with the specific word, using the neural network model; and wherein the replacing of the text span in the first content with the specific word comprises: determining whether one or more words immediately adjacent to the text span in the first content include at least one of one or more associated words with the specific word, based on the specific word stock; and if the determination result is affirmative, performing the replacement.

[0060] According to an embodiment of the present disclosure, the evaluation is based on at least one of word co-occurrence, local coherence, topic, and contextual word embedding in the second content.

[0061] According to an aspect of the present disclosure, there is provided an electronic device, comprising: a processor; and a memory connected to the processor and storing instructions that, when executed by the processor, cause the electronic device to: pre-generate a first content while interacting with a user; for each specific word in a specific word stock: embed the specific word into the first content to obtain a corresponding second content, and obtain a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and output at least one of a plurality of the second contents according to the evaluation results, wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

[0062] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine, as the explanatory word corresponding to the specific word, a word with a frequency, higher than a first threshold, of co-occurring with the specific word in the specific corpus.

[0063] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine the explanatory word corresponding to the specific word based on the specific corpus and word-level local coherence.

[0064] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the statistical model is used to determine a topic of the specific word based on the specific corpus as the explanatory word corresponding to the specific word.

[0065] According to an embodiment of the present disclosure, for each specific word in the specific word stock, the neural network model is used to predict a plurality of words each of which can replace the specific word in a sentence containing the specific word in the specific corpus, and to select from the plurality of words a word with a prediction score

higher than a second threshold as the explanatory word corresponding the specific word.

[0066] According to an embodiment of the present disclosure, the embedding of the specific word into the first content comprises: replacing a text span in the first content with the specific word.

[0067] According to an embodiment of the present disclosure, no punctuation mark is included in the text span.

[0068] According to an embodiment of the present disclosure, wherein, for each specific word in the specific word stock, a word immediately adjacent to the specific word in the specific corpus is stored in the specific word stock as an associated word with the specific word, using the neural network model; and wherein the replacing of the text span in the first content with the specific word comprises: determining whether one or more words immediately adjacent to the text span in the first content include at least one of one or more associated words with the specific word, based on the specific word stock; and if the determination result is affirmative, performing the replacement.

[0069] According to an embodiment of the present disclosure, the evaluation is based on at least one of word co-occurrence, local coherence, topic, and contextual word embedding in the second content.

[0070] Program codes for implementing the methods of the present disclosure may be written in any combination of one or more programming languages. The program codes may be provided to a general-purpose computer, a special purpose computer, or a processor or controller of other programmable data processing apparatus such that the program codes, when executed by the processor or controller, causes the functions/operations specified in the flowcharts and/or block diagrams to be implemented. The program codes may execute entirely on the machine, partly on the machine, execute as part of the stand-alone software package, and execute partly on the remote machine or execute entirely on the remote machine or server.

[0071] According to an embodiment, the machine-readable medium may be a tangible medium that may contain or store a program for use by or in connection with an instruction execution system, apparatus, or device. The machine readable medium may be a machine readable signal medium or a machine readable storage medium. A machine-readable medium may include, but is not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, or device, or any suitable combination thereof. More specific examples of machine-readable storage media may include electrical connections based on one or more wires, a portable computer disk, a hard disk, a random

access memory (RAM), a read only memory (ROM), an erasable programmable read only memory (EPROM), a flash memory, fiber optics, a CD-ROM, optical storage devices, magnetic storage devices, or any suitable combination thereof.

[0072] In addition, although the operations are depicted in a particular order, it may be appreciated that such operations are performed in the particular order shown or in the order, or that all illustrated operations should be performed to achieve the desired result. Multitasking and parallel processing may be advantageous in certain circumstances. Likewise, although several specific implementation details are included in the above discussion, these should not be construed as limitation of the scope of the disclosure. Certain features described in the context of separate implementations may also be implemented in combination in a single implementation. Conversely, various features described in the context of a single implementation may be implemented in a plurality of implementations, either individually or in any suitable sub-combination.

[0073] Although the subject matter has been described in language specific to structural features and/or methodological actions, it may be appreciated that the subject matter defined in the appended claims is not necessarily limited to the specific features or actions described. Rather, the specific features and actions described above are merely exemplary forms of implementing the claims.

[0074] Those skilled in the art will appreciate that steps, measures, and technical solutions in the operations, methods, and processes discussed in the embodiments of the present invention may be substituted, changed, combined, or deleted. In addition, other steps, measures, and technical solutions in the operations, methods, and processes discussed in the present invention may be substituted, changed, rearranged, combined or deleted. In addition, the prior art having other steps, measures, and technical solutions in the operations, methods, and processes discussed in the present disclosure may be substituted, changed, rearranged, combined, and deleted.

[0075] While the example embodiments of the present disclosure have been shown and described, it will be understood that various modifications and changes may be made without departing from the spirit and scope of the disclosure. Therefore, the disclosure is intended to cover all modifications and variations within the scope of the claims.

Claims

1. A method for generating a content containing a specific word, comprising:

an electronic device pre-generating a first content while interacting with a user;

for each specific word in a specific word stock:

embedding the specific word into the first content to obtain a corresponding second content, and

obtaining a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and

outputting at least one of a plurality of the second contents according to the evaluation results,

wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

2. The method according to claim 1, wherein, for each specific word in the specific word stock, the statistical model is used to determine, as the explanatory word corresponding to the specific word, a word with a frequency, higher than a first threshold, of co-occurring with the specific word in the specific corpus.

3. The method according to claim 1, wherein, for each specific word in the specific word stock, the statistical model is used to determine the explanatory word corresponding to the specific word based on the specific corpus and word-level local coherence.

4. The method according to claim 1, wherein, for each specific word in the specific word stock, the statistical model is used to determine a topic of the specific word based on the specific corpus as the explanatory word corresponding to the specific word.

5. The method according to claim 1, wherein, for each specific word in the specific word stock, the neural network model is used to predict a plurality of words each of which can replace the specific word in a sentence containing the specific word in the specific corpus, and to select from the plurality of words a word with a prediction score higher than a second threshold as the explanatory word corresponding the specific word.

6. The method according to claim 1, wherein the embedding of the specific word into the first content comprises: replacing a text span in the first content with the specific word.

7. The method according to claim 6, wherein no punctuation mark is included in the text span.

8. The method according to claim 6,

wherein, for each specific word in the specific word stock, a word immediately adjacent to

the specific word in the specific corpus is stored in the specific word stock as an associated word with the specific word, using the neural network model; and

wherein the replacing of the text span in the first content with the specific word comprises:

determining whether one or more words immediately adjacent to the text span in the first content include at least one of one or more associated words with the specific word, based on the specific word stock; and

if the determination result is affirmative, performing the replacement.

9. The method according to claim 1, wherein the evaluation is based on at least one of word co-occurrence, local coherence, topic, and contextual word embedding in the second content.

10. An electronic device, comprising:

a processor; and

a memory connected to the processor and storing instructions that, when executed by the processor, cause the electronic device to:

pre-generate a first content while interacting with a user;

for each specific word in a specific word stock:

embed the specific word into the first content to obtain a corresponding second content, and

obtain a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word; and

output at least one of a plurality of the second contents according to the evaluation results,

wherein the specific word stock is generated and periodically updated based on a specific corpus using the neural network model.

11. The electronic device according to claim 10, wherein, for each specific word in the specific word stock, the statistical model is used to determine, as the explanatory word corresponding to the specific word, a word with a frequency, higher than a first threshold, of co-occurring with the specific word in the specific corpus.

12. The electronic device according to claim 10, wherein, for each specific word in the specific word stock, the statistical model is used to determine the explanatory word corresponding to the specific word based on the specific corpus and word-level local coherence.

13. The electronic device according to claim 10, wherein, for each specific word in the specific word stock, the statistical model is used to determine a topic of the specific word

based on the specific corpus as the explanatory word corresponding to the specific word.

14. The electronic device according to claim 10, wherein, for each specific word in the specific word stock, the neural network model is used to predict a plurality of words each of which can replace the specific word in a sentence containing the specific word in the specific corpus, and to select from the plurality of words a word with a prediction score higher than a second threshold as the explanatory word corresponding the specific word.

15. The electronic device according to claim 10, wherein the embedding of the specific word into the first content comprises replacing a text span in the first content with the specific word;

wherein, for each specific word in the specific word stock, a word immediately adjacent to the specific word in the specific corpus is stored in the specific word stock as an associated word with the specific word, using the neural network model; and

wherein the replacing of the text span in the first content with the specific word comprises:

determining whether one or more words immediately adjacent to the text span in the first content include at least one of one or more associated words with the specific word, based on the specific word stock; and

if the determination result is affirmative, performing the replacement.

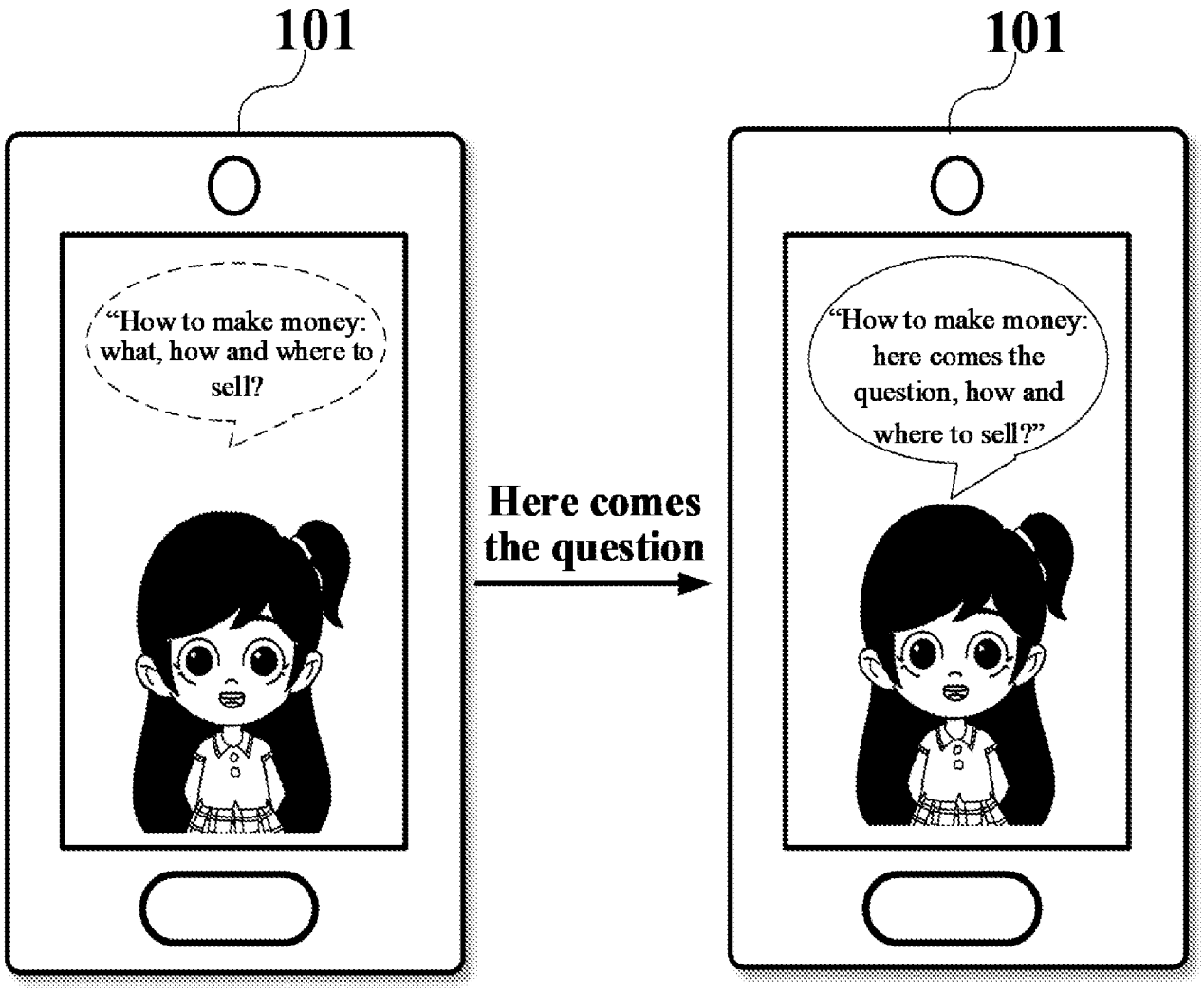

101
101
"How to make money: what, how and where to sell?
"How to make money: here comes the question, how and where to sell?"
Here comes the question
A
B

FIG 1

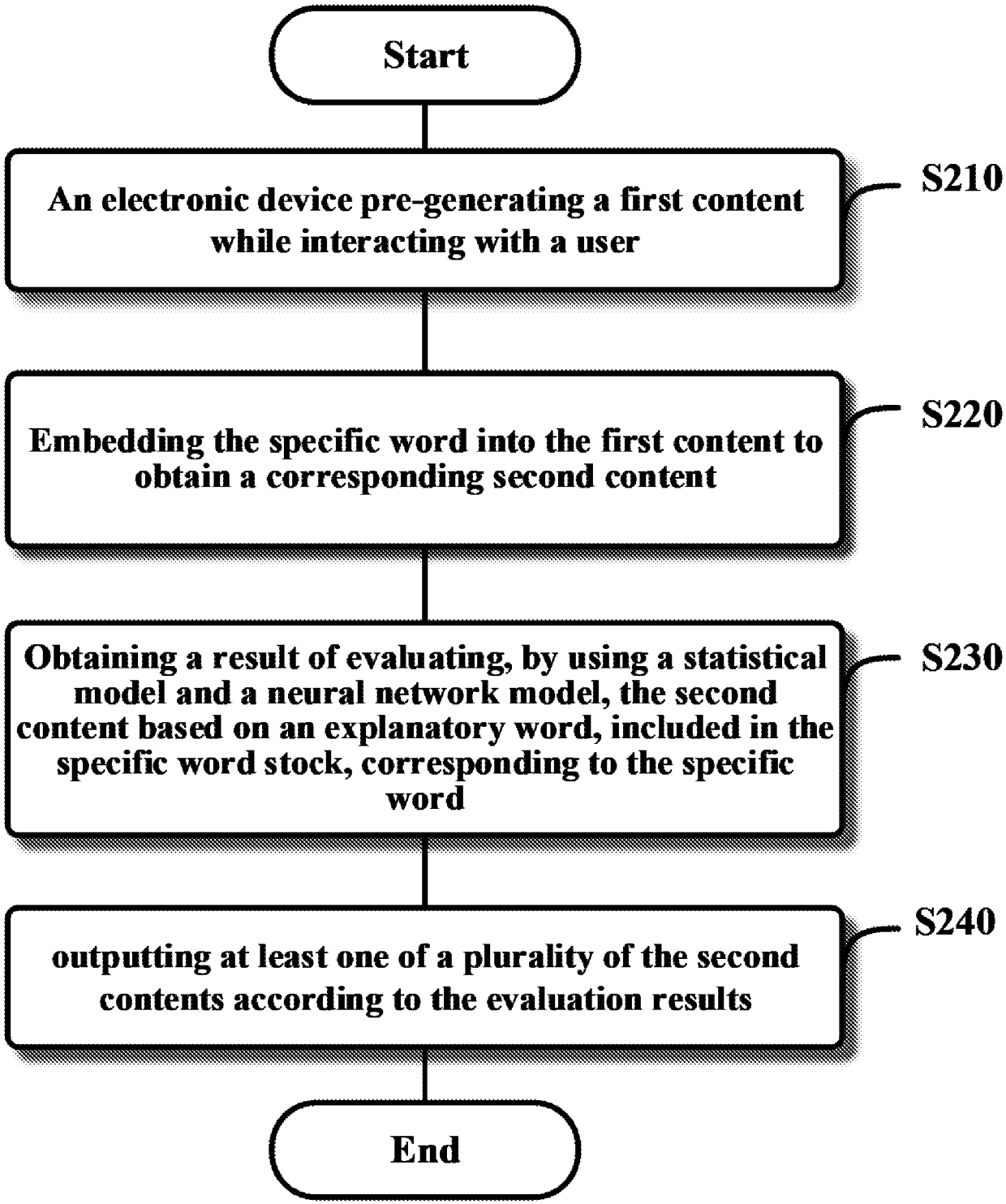

Start
An electronic device pre-generating a first content while interacting with a user
S210
Embedding the specific word into the first content to obtain a corresponding second content
S220
Obtaining a result of evaluating, by using a statistical model and a neural network model, the second content based on an explanatory word, included in the specific word stock, corresponding to the specific word
S230
outputting at least one of a plurality of the second contents according to the evaluation results
S240
End

FIG 2

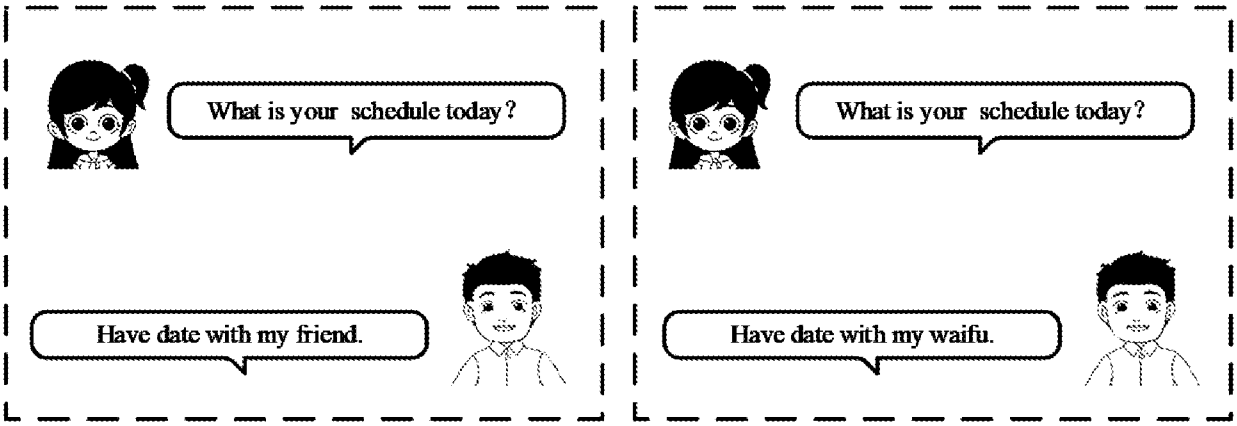

What is your schedule today?
Have date with my friend.
What is your schedule today?
Have date with my waifu.

FIG 3

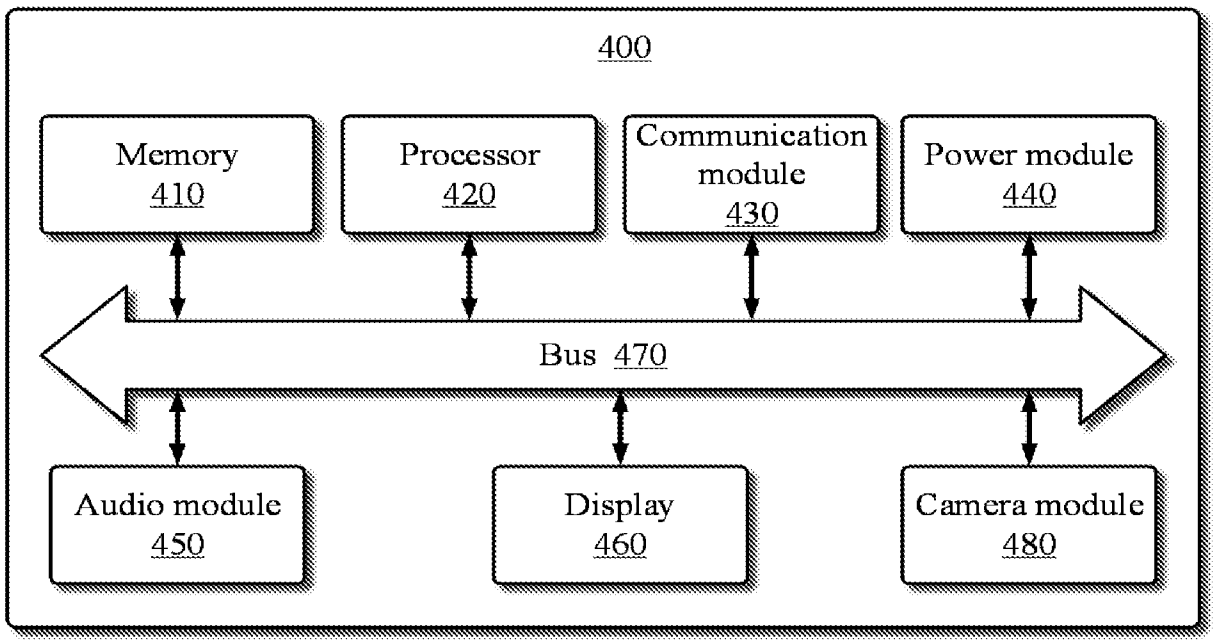

400
Memory
410
Processor
420
Communication
module
430
Power module
440
Bus 470
Audio module
450
Display
460
Camera module
480

FIG 4

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/030017

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06F40/35 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

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2018/365212 A1 (BANERJEE SIDDHARTHA [US] ET AL) 20 December 2018 (2018-12-20) abstract paragraph [0002] - paragraph [0007] paragraph [0034] - paragraph [0042] paragraph [0055] - paragraph [0056] paragraph [0077] - paragraph [0080] paragraph [0082] - paragraph [0089] paragraph [0092] - paragraph [0106] paragraph [0109] - paragraph [0134] ----- -/--	1-15

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
10 July 2020	20/08/2020

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

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2020/030017

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DI WANG ET AL: "Steering Output Style and Topic in Neural Response Generation", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 10 September 2017 (2017-09-10), XP080819802, abstract page 1, left-hand column, line 6 - page 1, right-hand column, line 17 page 2, left-hand column, line 12 - page 3, left-hand column, line 30 page 4, left-hand column, line 5 - page 4, right-hand column, line 13 page 4, right-hand column, line 48 - page 6, left-hand column, line 16	1-15
A	PRASOMPHAN SATHIT: "Improvement of Chatbot in Trading System for SMEs by Using Deep Neural Network", 2019 IEEE 4TH INTERNATIONAL CONFERENCE ON CLOUD COMPUTING AND BIG DATA ANALYSIS (ICCCBDA), IEEE, 12 April 2019 (2019-04-12), pages 517-522, XP033552779, DOI: 10.1109/ICCCBDA.2019.8725745 ISBN: 978-1-7281-1409-5 [retrieved on 2019-05-28] abstract page 1, left-hand column, line 1 - page 2, right-hand column, line 17 page 3, left-hand column, line 36 - page 3, right-hand column, line 52 page 4, right-hand column, line 45 - page 5, left-hand column, line 4	1-15
A	US 2014/337814 A1 (KALNS EDGAR T [US] ET AL) 13 November 2014 (2014-11-13) abstract paragraph [0015] - paragraph [0023] paragraph [0026] - paragraph [0030] paragraph [0052] - paragraph [0053] paragraph [0065] - paragraph [0068]	1-15
A	US 2018/276293 A1 (ROZSA ENIKO I [CA] ET AL) 27 September 2018 (2018-09-27) abstract paragraph [0031] - paragraph [0035] paragraph [0041] - paragraph [0045] paragraph [0056]	1-15

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2020/030017

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018365212 A1	20-12-2018	US 2018365212 A1	20-12-2018
		US 2019121850 A1	25-04-2019
US 2014337814 A1	13-11-2014	NONE	
US 2018276293 A1	27-09-2018	CA 2741212 A1	27-11-2012
		CN 103548023 A	29-01-2014
		DE 112012001750 T5	06-02-2014
		TW 201314476 A	01-04-2013
		US 2012303356 A1	29-11-2012
		US 2013290366 A1	31-10-2013
		US 2018276293 A1	27-09-2018
		WO 2012162822 A1	06-12-2012