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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
10 July 2014 (10.07.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/106339 A1

(51) International Patent Classification:
G06F 17/20 (2006.01)

(21) International Application Number:
PCT/CN2013/070121

(22) International Filing Date:
6 January 2013 (06.01.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **EMPIRE TECHNOLOGY DEVELOPMENT LLC** [US/US]; 2711 Centerville Road, Suite 400, Wilmington, DE 19808 (US).

(72) Inventors; and
(71) Applicants *(for US only)*: **LI, Daqi** [CN/CN]; 239 Xiqi Road, Xi'an, Shaanxi 710003 (CN). **FANG, Jun** [CN/CN]; 127 Youyi Xilu, Xi'an, Shaanxi 710072 (CN).

(74) Agent: **CHINA SCIENCE PATENT & TRADEMARK AGENT LTD.**; 11/F., Bldg. D, International Finance and Economics Center, No.87, West 3rd Ring North Rd., Haidian District, Beijing 100089 (CN).

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

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

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

(54) Title: TEXT BILLING BASED ON SEMANTIC DATA REASONING

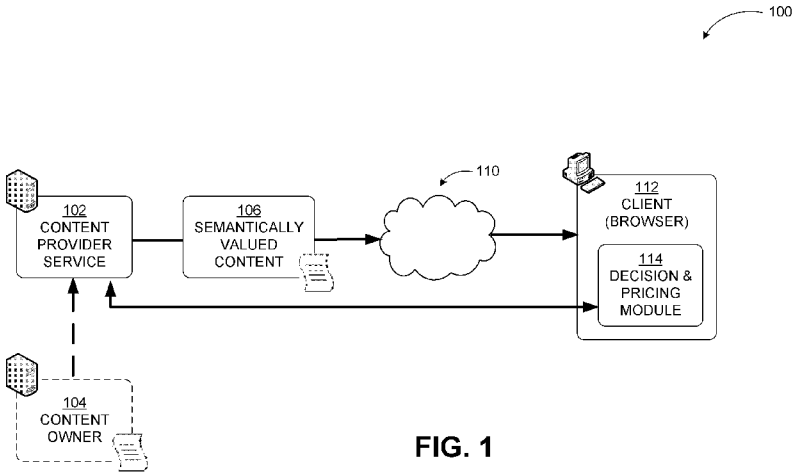

FIG. 1

(57) Abstract: Technologies are provided for text billing based on semantic data reasoning at a sentence level granularity. In some examples, price information may be embedded into sentences determined to be valuable in form of one or more semantic rules derived from ontology. When a client application selects a sentence (e.g., for copying) or downloads the textual content including the sentence, a value of the sentence may be determined by reasoning. The price may be retrieved from a pricing data store and the user charged accordingly.

WO 2014/106339 A1

TEXT BILLING BASED ON SEMANTIC DATA REASONING

BACKGROUND

[0001] Unless otherwise indicated herein, the materials described in this section are not prior art to the claims in this application and are not admitted to be prior art by inclusion in this section.

[0002] In the cloud computing environment, textual content is one of the many commodities that exchanges hands every day. Text owners publish textual content for users to browse and download portions of (or entire) content that is of interest to them. Similar to other forms of content, providing textual content is a service for which users are charged for. Determining how to price textual content may be a challenge for service providers. Conventional approaches typically create accounts for users and charge them based on predefined prices for entire content, based on amount of downloaded textual data, or even based on time a user spends browsing the textual content.

[0003] As mentioned above, one of the pricing strategies is to charge a user for the entire textual content such as a book or an article. The users may be interested in portions of the textual content, however. For example, one user may be interested in some sentences of an article, while another user may be interested for some other sentences, but neither user may be interested in the entire article. Even if the users were charged by the chapters of the article, the approach may not provide sufficient granularity. Furthermore, a user may download an entire textual content such as a book, from which multiple users may select and use distinct sentences resulting in lost value to the content provider.

SUMMARY

[0004] The present disclosure generally describes techniques for cloud-based text billing based on semantic data reasoning.

[0005] According to some examples, a method is described for providing text billing based on semantic data reasoning. The method may include pricing one or more rules in an ontology associated with textual content to be published; determining one or more valuable sentences within the textual content based on the priced rules; labeling the valuable sentences with price information; and publishing the textual content.

[0006] According to other examples, a content provider service executed on one or more servers is described for providing text billing based on semantic data reasoning. The content

provider service may include a textual content processing module. The textual content processing module may price one or more rules in an ontology associated with textual content to be published; determine one or more valuable sentences within the textual content based on the priced rules; label the valuable sentences with price information; and publish the textual content.

[0007] According to further examples, a content provider service for providing text billing based on semantic data reasoning is described. The content provider service may a textual content processing module executed on a server. The textual content processing module may price one or more rules in an ontology associated with textual content to be published; determine one or more valuable sentences within the textual content based on the priced rules; label the valuable sentences with price information; and publish the textual content. The content service provider may further include a decision and pricing module executed as part of a client application. The decision and pricing module may receive selection of one or more valuable sentences within the published textual content from a client application; retrieve one or more rules associated with the selected sentences; and determine a fee associated with the selected sentences employing reasoning based on the retrieved rules.

[0008] According to yet further examples, a computer readable medium may store instructions for providing text billing based on semantic data reasoning. The instructions may include actions similar to the method for providing text billing based on semantic data reasoning discussed above.

[0009] The foregoing summary is illustrative only and is not intended to be in any way limiting. In addition to the illustrative aspects, embodiments, and features described above, further aspects, embodiments, and features will become apparent by reference to the drawings and the following detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The foregoing and other features of this disclosure will become more fully apparent from the following description and appended claims, taken in conjunction with the accompanying drawings. Understanding that these drawings depict only several embodiments in accordance with the disclosure and are, therefore, not to be considered limiting of its scope, the disclosure will be described with additional specificity and detail through use of the accompanying drawings, in which:

FIG. 1 illustrates an example cloud-based text billing system based on semantic data reasoning, where at least a portion of pricing operations may be implemented at a client application;

FIG. 2 illustrates another example cloud-based text billing system based on semantic data reasoning, where textual content treatment and pricing operations may be implemented at a hosted service;

FIG. 3 illustrates an example framework for a text billing system based on semantic data reasoning;

FIG. 4 illustrates an example ontology segment instance for a text billing system based on semantic data reasoning;

FIG. 5 illustrates a general purpose computing device, which may be used for providing text billing based on semantic data reasoning;

FIG. 6 is a flow diagram illustrating an example method for providing text billing based on semantic data reasoning that may be performed by a computing device such as the computing device in FIG. 5; and

FIG. 7 illustrates a block diagram of an example computer program product, all arranged in accordance with at least some embodiments described herein.

DETAILED DESCRIPTION

[0011] In the following detailed description, reference is made to the accompanying drawings, which form a part hereof. In the drawings, similar symbols typically identify similar components, unless context dictates otherwise. The illustrative embodiments described in the detailed description, drawings, and claims are not meant to be limiting. Other embodiments may be utilized, and other changes may be made, without departing from the spirit or scope of the subject matter presented herein. It will be readily understood that the aspects of the present disclosure, as generally described herein, and illustrated in the Figures, can be arranged, substituted, combined, separated, and designed in a wide variety of different configurations, all of which are explicitly contemplated herein.

[0012] This disclosure is generally drawn, *inter alia*, to methods, apparatus, systems, devices, and/or computer program products related to text billing based on semantic data reasoning.

[0013] Briefly stated, technologies are generally described for text billing based on semantic data reasoning at a sentence level granularity. In some examples, price information may be embedded into sentences determined to be valuable in form of one or more semantic

rules derived from ontology. When a client application selects a sentence (*e.g.*, for copying) or downloads the textual content including the sentence, a value of the sentence may be determined by reasoning. The price may be retrieved from a pricing data store and the user charged accordingly.

[0014] FIG. 1 illustrates an example cloud-based text billing system based on semantic data reasoning, where at least a portion of pricing operations may be implemented at a client application, arranged in accordance with at least some embodiments described herein.

[0015] Users may browse and download portions of (or entire) textual content published online. Similar to other forms of content, providing textual content is a service for which users may be charged for. Semantic data is data organized according to a semantic data model that may be interpreted meaningfully without human intervention. The semantic data model (or ontology) is based on relationships between stored symbols and the real world. Thus, ontology is a formal, explicit specification of a shared conceptualization and renders shared vocabulary and taxonomy which may model a domain with the definition of objects and/or concepts and their properties and relations.

[0016] Semantic data is used in a wide variety of data management systems and applications, and may be organized based on binary models of objects, usually in groups of three parts: two objects and their relationship. For example, the data organization for representing a cup on a table may be in form of CUP TABLE. The objects (cup and table) may be interpreted with regard to their relationship (cup on table). The data may be organized linearly, indicating that since cup comes first in line, it is the object that acts. In other words, the position of the word may indicate that the cup is on the table and not that the table is on the cup. Databases designed around this concept may have greater applicability and may be more easily integrated into other databases. Semantic data systems are designed to represent the real world as accurately as possible within a data set. Data symbols may be organized linearly and hierarchically to assign certain meanings like the one described above. By representing the real world within data sets, semantic data may allow machines to interact with worldly information without human interpretation.

[0017] Description Logics (DL), which is a description language, may be used to provide a logical formalism for ontologies and semantic data by modeling concepts, roles, individuals and their relationships. More generally, DL may be used as a foundation for building knowledge bases from two main components, a terminological knowledge box (TBox) and an assertional knowledge box (ABox). The TBox refers to statements regarding the terminology itself describing the hierarchy of concepts and the relationships between

concepts, whereas the ABox refers to statements regarding the belonging of individuals to concepts and the relationships between individuals. In description logics, both TBox and ABox constitute a knowledge base. While the TBox tend to be more static in the sense that the definition of concepts is rather fixed, the ABox tend to be more dynamic in the sense that the population of individuals may change and be subject to extension through DL reasoning.

[0018] As shown in a diagram 100, a content provider service 102 may publish semantically valued content 106 over one or more networks 110, such as the Internet, to users. The content service provider 102 may generate the content or may receive it from one or more content owners 104. A user interested in a portion of (or entire) the published content may access the content through a client application 112 such as a browser. According to some examples, the client application 112 may include a decision and pricing module 114 for determining charges for selected (or downloaded) portions of the content by interacting with the content provider service 102. Price information may be embedded into sentences of the semantically valued content 106 determined to be valuable in form of one or more semantic rules derived from ontology. When a sentence within the semantically valued content 106 is selected (*e.g.*, for copying) or downloaded, a value of the sentence may be determined by reasoning at the content provider service 102 or by the decision and pricing module 114.

[0019] FIG. 2 illustrates another example cloud-based text billing system based on semantic data reasoning, where textual content treatment and pricing operations may be implemented at a hosted service, arranged in accordance with at least some embodiments described herein.

[0020] A core of textual data may be considered as information which maybe represented as semantic data. Thus, a value of the text may be loaded by semantic data. As mentioned above, semantic data may be mapped onto ontology (composed of rules and concepts). If the rules are labeled with a price, the text may be directly billed on semantic level. A system according to some example embodiments may also allow the price to be dynamically updated by the owner in real-time. The bases of the price calculation may be ontology and reasoning, so the price management may be viewed as ontology management working on semantic level. If a content (text) owner wishes to descend or ascend the price of the content, the price may be modified by adjusting the price label or tag on the ontology, which may be unique and assigned at a specific node of the network. Furthermore, copyright protection may also be supported. Because, the price is combined with the text, any attempts to illegally modify the text may destroy the unity of text and price.

[0021] The example configuration, shown in a diagram 200, is similar to that of the diagram 100 in some ways. Semantically valued content 230 may be published to users by a content service provider 202 over one or more networks 210, such as the Internet. The raw content may be provided to the content provider service 202 by one or more content owners 204. The content provider service 202 may process the content, embed pricing information labeling sentences deemed valuable, and publish the semantically valued content 230. When a user shows interest on portions of the published content and copies, downloads, or otherwise consumes the content (text), the content provider service 202 may determine values through reasoning, add pricing for total selected portions of the content and inform a client application 222 of the user employing modules such as a decision and pricing module 224, a valuation module 226, and similar ones. Such modules may be distinct modules or applications within the content provider service 202 or integrated modules or applications performing multiple tasks.

[0022] FIG. 3 illustrates an example framework for a text billing system based on semantic data reasoning, arranged in accordance with at least some embodiments described herein.

[0023] In a system according to some example embodiments, rules among the ontology may be assigned respective prices by the content owner. After some pre-treatment, the valuable content portions (*e.g.*, sentences) may be determined and labeled. In an example configuration, where the pricing may be performed on the client side, valuable sentences may be determined once the client application (*e.g.*, browser) downloads the content. Subsequently, the rules which valuable sentences include may be retrieved. Having been reasoned, the price of rule for each sentence may be attained. The remaining rules may be deduced by analogy. The total fee may be calculated and the browser charged for.

[0024] As shown in a diagram 300, an example scenario may begin with a text owner 304 providing price of a rule (344) to ontology 350. Textual data 342 may be subjected to pre-treatment (346) (*e.g.*, creation of semantic model) such that it is suitable for billing. Valuable sentences may be determined (348) with priced rules from the ontology 350 by processing the pre-treated text (346). Next, the determined valuable sentences may be labeled (or tagged) (352) associating them with their assigned price(s) and published over a network 310.

[0025] On the user side, a client application 332 may receive user selection of one or more valuable sentences (354), upon which the client application 332 may retrieve the rules for the selected valuable sentences (356). Reasoning may be performed on the semantic data

and the price for the selected sentences may be totaled up (358) so the user can be charged on a sentence basis for the consumed data. If the text owner 304 decides to change the price for one or more sentences, those prices may be adjusted by modifying the rule price label at the ontology 350.

[0026] FIG. 4 illustrates an example ontology segment instance for a text billing system based on semantic data reasoning, arranged in accordance with at least some embodiments described herein.

[0027] As depicted in a diagram 400, an ontology may have a hierarchical structure of concepts 468 spreading from a root node 462 and coupled through rules 464. In a system according to embodiments, each rule 464 may be labeled with a price 466. For computation purposes, it may be assumed that some example text waiting for billing is T, which may be transformer to semantic data form. O may represent the ontology with respect to T. $R^O = \{r_1, \ldots r_k, \ldots\}$ may be a rule set of O, where r_k is an arbitrary rule and k a positive integer. Herein, the valuable content in T may be included by O. O may be built and managed by a variety of techniques. In some examples, common or domain ontologies from the semantic web may be customized to fit the needs to a particular textual content provider and employed to manage the price rules. In other examples, the ontologies may be refined for particular textual content using commonly available ontology compilation tools.

[0028] On the text owner side, the text owner may be enabled to label every rule of the ontology a price, which may be used to evaluate the charge of the browser. The text owner may issue an order to every valuable sentence rule and the specific prices may be computed automatically according to:

$$[1] \qquad P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}},$$

where P^{r_k} is the price of rule r_k, *total* is the total price of the whole text, |R| is the total number of valuable rules included in the ontology O, and seq^{r_k} is the sequence of r_k. The prices computed according to [1] may be automatically labeled on the ontology, O, with the help of an ontology management tool. The text owner may also be enabled to adjust the price this way. Embodiments are not limited to the pricing computation according to [1], and other pricing formulas may also be employed using the principles described herein. Furthermore, rules among R which are not ordered (by the text owner) may be interpreted as the price of an associated sentence being equal to zero.

[0029] In pre-treatment of text, T may be divided into single sentences first with the help of diverse logical list separators in T. Thus, the T may be transformed to a sentence set $S = \{s_1, \ldots, s_i, \ldots\}$, where s_i is an arbitrary sentence and i is a positive integer. At a second sub-

step, firstly, a sentence s_i may be selected from S, and the concept in the selected sentence, s_i, may be determined. Next, the determined concept may be mapped onto O. If a direct link exists between the concept and the ontology structure, then every rule may be stored into a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to s_i. Subsequently, s_i may be removed from S, if $S \neq \varnothing$ and the second sub-step may be repeated. Otherwise, the process may end. This way, every sentence may be partitioned and their respective rules ascertained.

[0030] As mentioned previously, not all sentences among T may be valuable for pricing. So, a system according to some example embodiments may determine valuable sentences through a two stage process employing direct matching and relatedness computation. Direct matching may determine a sentence which conformably includes a valuable rule as discussed above. Direct matching may be performed by selecting a sentence s_i from S, extracting the rule set R^{s_i} associated with s_i, and designating s_i as a free sentence, if there does not exist any r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$. Otherwise, s_i may be designated as a valuable sentence and S reduced by $S = S - s_i$. The direct matching process may be performed iteratively and end if all of sentences of S have been tested.

[0031] After direct matching, some sentences related to valuable rules in S may still exist. This maybe induced by the reason of semantic presentation differences. Thus, a relatedness computation may be carried out in succession after direct matching. Sentences sufficiently related to the valuable rules (*e.g.*, above a predetermined threshold) may also be labeled as valuable sentences.

[0032] Relatedness computation may begin with selection of a sentence s_i from S, followed by extraction of a rule set R^{s_i} associated with s_i. If the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, then s_i may be designated a valuable sentence and S reduced by $S = S - s_i$. The relatedness may be computed according to equation [2], which is based on the assumption that words which appear in the entire published content may have some semantic relatedness.

$$[2] \qquad \mathrm{rel}(r_j, r_k) = \frac{\mathrm{hits}(r_j + r_k)}{\min(\mathrm{hits}(r_j), \mathrm{hits}(r_k))},$$

where $\mathrm{rel}(r_j,r_k)$ is relatedness between r_j and r_k, $\mathrm{hits}(r_j+r_k)$ is a number of pages returned by searching word1 and word2, $\min(\mathrm{hits}(r_j),\mathrm{hits}(r_k))$ is the minimum number of pages returned by searching word1 and word2 separately. Semantic relatedness computed by using [2] may have a value between 0 and 1. The relatedness value may be closer to 1, if the two words are closely related to each other.

[0033] As long as word1 and word2 are related to each other, the returned pages from a separate search and a jointly search may have some association. By using the minimum number of pages instead of a maximum number or an average number in equation [2] following phenomenon may be avoided. When hits of one word (*e.g.*, word1) are many while hits of the other (word2) are few, if many appearances of word2 are accompanied by word1, a false positive result may be deduced that the words are strongly related.

[0034] Once the semantically valued text is published, a client application may receive user selection of desired text and determine valuable sentences in the desired text according to the label or tag associated with the sentences. The valuable sentences may be collected into a sentence set $S' = \{s_1,\ldots,s_m,\ldots\}$, where s_m is an arbitrary valuable sentence and m is a positive integer. The rules associated with the valuable sentences may be ascertained, referred to as R^{s_m} denoting the rules with respect to s_m, $R^{s_m} = \{r_1,\ldots,r_n,\ldots\}$.

[0035] Once the rules for the valuable sentences are ascertained, the price of s_m may be computed by extracting a rule r_n from R^{s_m} and reasoning r_n employing a reasoner. A semantic reasoner, reasoning engine, rules engine, or a reasoner, is an application or module capable of inferring logical consequences from a set of asserted facts or axioms. Using rules that are commonly specified by means of an ontology language and often a description language, reasoners may use first-order predicate logic to perform reasoning. Some reasoners may also employ probabilistic processing. The reasoner may return a "consistent" result or an "inconsistent" result. Inconsistent result may be regarded as the rule indicating the sentence is free, while consistent result may indicate that the sentence is associated with a fee, depicted in f^{r_n}.

[0036] There may exist two cases where r_n is valuable, namely r_n matches a valuable rule of O or is sufficiently related to a valuable rule as discussed above. f^{r_n} may be computed differently in these two cases. If r_n matches a valuable rule in O, f^{r_n} may be equal to P^{r_n} as determined in [1]. If r_n is sufficiently related to a valuable rule, f^{r_n} may be computed according to:

$$[3] \qquad f^{r_n} = \mu * \max_{r_x \in R^O, \mathrm{rel}(r_n, r_x) \geq \eta} [\mathrm{rel}(r_n, r_x) * P^{r_x}],$$

where μ is an adjustment coefficient between 1 and 0.

[0037] Next, the r_n may be removed from R^{s_m}, if $R^{s_m} \neq \varnothing$ and the computation using [3] may be repeated, else the total fee of s_m may be computed using:

$$[4] \qquad F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n},$$

where F^{s_m} is the total fee for s_m.

[0038] FIG. 5 illustrates a general purpose computing device, which may be used for providing text billing based on semantic data reasoning, arranged in accordance with at least some embodiments described herein.

[0039] For example, the computing device 500 may be used as part of a textual content service for providing text billing based on semantic data reasoning as described herein. In an example basic configuration 502, the computing device 500 may include one or more processors 504 and a system memory 506. A memory bus 508 may be used for communicating between the processor 504 and the system memory 506. The basic configuration 502 is illustrated in FIG. 5 by those components within the inner dashed line.

[0040] Depending on the desired configuration, the processor 504 may be of any type, including but not limited to a microprocessor (μP), a microcontroller (μC), a digital signal processor (DSP), or any combination thereof. The processor 504 may include one more levels of caching, such as a level cache memory 512, a processor core 514, and registers 516. The example processor core 514 may include an arithmetic logic unit (ALU), a floating point unit (FPU), a digital signal processing core (DSP Core), or any combination thereof. An example memory controller 518 may also be used with the processor 504, or in some implementations the memory controller 518 may be an internal part of the processor 504.

[0041] Depending on the desired configuration, the system memory 506 may be of any type including but not limited to volatile memory (such as RAM), non-volatile memory (such as ROM, flash memory, *etc.*) or any combination thereof. The system memory 506 may include an operating system 520, a content billing service 522, and program data 524. The content billing service 522 may include a labeling module 525 and a rule module 526 for embedding sentences with value information and providing rules for pricing based on semantic data reasoning as described herein. The program data 524 may include, among other data, rule data 528 or the like, as described herein.

[0042] The computing device 500 may have additional features or functionality, and additional interfaces to facilitate communications between the basic configuration 502 and any desired devices and interfaces. For example, a bus/interface controller 530 may be used to facilitate communications between the basic configuration 502 and one or more data storage devices 532 via a storage interface bus 534. The data storage devices 532 may be one or more removable storage devices 536, one or more non-removable storage devices 538, or a combination thereof. Examples of the removable storage and the non-removable storage devices include magnetic disk devices such as flexible disk drives and hard-disk drives (HDD), optical disk drives such as compact disk (CD) drives or digital versatile disk (DVD)

drives, solid state drives (SSD), and tape drives to name a few. Example computer storage media may include volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information, such as computer readable instructions, data structures, program modules, or other data.

[0043] The system memory 506, the removable storage devices 536 and the non-removable storage devices 538 are examples of computer storage media. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD), solid state drives, or other optical storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which may be used to store the desired information and which may be accessed by the computing device 500. Any such computer storage media may be part of the computing device 500.

[0044] The computing device 500 may also include an interface bus 540 for facilitating communication from various interface devices (*e.g.*, one or more output devices 542, one or more peripheral interfaces 544, and one or more communication devices 566) to the basic configuration 502 via the bus/interface controller 530. Some of the example output devices 542 include a graphics processing unit 548 and an audio processing unit 550, which may be configured to communicate to various external devices such as a display or speakers via one or more A/V ports 552. One or more example peripheral interfaces 544 may include a serial interface controller 554 or a parallel interface controller 556, which may be configured to communicate with external devices such as input devices (*e.g.*, keyboard, mouse, pen, voice input device, touch input device, *etc.*) or other peripheral devices (*e.g.*, printer, scanner, *etc.*) via one or more I/O ports 558. An example communication device 566 includes a network controller 560, which may be arranged to facilitate communications with one or more other computing devices 562 over a network communication link via one or more communication ports 564. The one or more other computing devices 562 may include servers at a datacenter, customer equipment, and comparable devices.

[0045] The network communication link may be one example of a communication media. Communication media may typically be embodied by computer readable instructions, data structures, program modules, or other data in a modulated data signal, such as a carrier wave or other transport mechanism, and may include any information delivery media. A "modulated data signal" may be a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media may include wired media such as a wired network or

direct-wired connection, and wireless media such as acoustic, radio frequency (RF), microwave, infrared (IR) and other wireless media. The term computer readable media as used herein may include both storage media and communication media.

[0046] The computing device 500 may be implemented as a part of a general purpose or specialized server, mainframe, or similar computer that includes any of the above functions. The computing device 500 may also be implemented as a personal computer including both laptop computer and non-laptop computer configurations.

[0047] Example embodiments may also include methods for text billing based on semantic data reasoning. These methods can be implemented in any number of ways, including the structures described herein. One such way may be by machine operations, of devices of the type described in the present disclosure. Another optional way may be for one or more of the individual operations of the methods to be performed in conjunction with one or more human operators performing some of the operations while other operations may be performed by machines. These human operators need not be collocated with each other, but each can be with a machine that performs a portion of the program. In other examples, the human interaction can be automated such as by pre-selected criteria that may be machine automated.

[0048] FIG. 6 is a flow diagram illustrating an example method for providing text billing based on semantic data reasoning that may be performed by a computing device such as the computing device in FIG. 5, arranged in accordance with at least some embodiments described herein.

[0049] Example methods may include one or more operations, functions or actions as illustrated by one or more of blocks 622, 624, 626, 628, 630, and/or 632, and may in some embodiments be performed by a computing device such as the computing device 500 in FIG. 5. The operations described in the blocks 622-632 may also be stored as computer-executable instructions in a computer-readable medium such as a computer-readable medium 620 of a computing device 610.

[0050] An example process for text billing based on semantic data reasoning may begin at block 622, "PRICE RULE ON ONTOLOGY", where one or more text owners such as the text owner 304 may provide a pricing directive, which may be automatically processed to associate a rule in the ontology with a price.

[0051] Block 622 may be followed by block 624, "DETERMINE VALUABLE SENTENCES AMONG TEXT", where a content provider service such as the content

provider service 202 may determine the valuable sentences through direct matching and/or relatedness computation.

[0052] Block 624 may be followed by block 626, "LABEL VALUABLE SENTENCES", where one or more sentences may be tagged with price information based on the associated rule in the ontology by the content provider service 202. In some embodiments, pricing information embedded textual content may be published for users to consume in a cloud-based environment. Blocks 628 through 632 may be performed upon publishing of the textual content.

[0053] At block 628, "RECEIVE SELECTION OF VALUABLE SENTENCE", the client application 332 or the content provider service 202 may receive a user selection of a valuable sentence.

[0054] Block 628 may be followed by block 630, "RETRIEVE RULE(S) FOR VALUABLE SENTENCE", where one or more rules for the user selected sentence may be retrieved from the ontology 350 for reasoning and fee calculation by the client application 332 or the content provider service 202.

[0055] Block 630 may be followed by block 632, "PERFORM REASONING AND ADD PRICING", where the client application 332 or the content provider service 202 may use a reasoner to determine the price of selected valuable sentences, and then sum the prices up for charging to the user.

[0056] FIG. 7 illustrates a block diagram of an example computer program product arranged in accordance with at least some embodiments described herein.

[0057] In some examples, as shown in FIG. 7, the computer program product 700 may include a signal bearing medium 702 that may also include one or more machine readable instructions 704 that, when executed by, for example, a processor, may provide the functionality described herein. Thus, for example, referring to the processor 504 in FIG. 5, the content billing service 522 may undertake one or more of the tasks shown in FIG. 7 in response to the instructions 704 conveyed to the processor 504 by the medium 702 to perform actions associated with text billing based on semantic data reasoning as described herein. Some of those instructions may include, for example, pricing rule on ontology, determining valuable sentences among text, labeling valuable sentences, receiving selection of valuable sentence, retrieving rule(s) for valuable sentence, and performing reasoning and adding pricing, according to some embodiments described herein.

[0058] In some implementations, the signal bearing medium 702 depicted in FIG. 7 may encompass a computer-readable medium 706, such as, but not limited to, a hard disk drive, a

solid state drive, a Compact Disc (CD), a Digital Versatile Disk (DVD), a digital tape, memory, *etc.* In some implementations, the signal bearing medium 702 may encompass a recordable medium 708, such as, but not limited to, memory, read/write (R/W) CDs, R/W DVDs, *etc.* In some implementations, the signal bearing medium 702 may encompass a communications medium 710, such as, but not limited to, a digital and/or an analog communication medium (*e.g.*, a fiber optic cable, a waveguide, a wired communications link, a wireless communication link, *etc.*). Thus, for example, the program product 700 may be conveyed to one or more modules of the processor 704 by an RF signal bearing medium, where the signal bearing medium 702 is conveyed by the wireless communications medium 710 (*e.g.*, a wireless communications medium conforming with the IEEE 802.11 standard).

[0059] According to some examples, a method is described for providing text billing based on semantic data reasoning. The method may include pricing one or more rules in an ontology associated with textual content to be published; determining one or more valuable sentences within the textual content based on the priced rules; labeling the valuable sentences with price information; and publishing the textual content.

[0060] According to other examples, the method may further include receiving selection of one or more valuable sentences within the published textual content; retrieving one or more rules associated with the selected sentences; and determining a fee associated with the selected sentences employing reasoning based on the retrieved rules. Pricing the one or more rules in the ontology may include receiving an order from a textual content owner for valuable sentence rules; and determining a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, $|R|$ is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k.

[0061] According to further examples, the method may also include if a rule among R is not ordered, designating a sentence associated with the unordered rule as a free sentence without a price to be charged. The method may yet include dividing the textual content into single sentences using logical separators to obtain a transformed sentence set; for each sentence in the transformed sentence set, determining a concept; mapping the determined concept onto the ontology; and if a direct link exists between the concept and the ontology, storing a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

[0062] According to yet other examples, determining the one or more valuable sentences may include employing a two stage process comprising direct matching and

relatedness computation. The direct matching may include selecting a sentence s_i from a transformed sentence set S; extracting a rule set R^{s_i} associated with s_i; designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as $S = S - s_i$; and iteratively performing direct matching through S. The method may further include upon completion of the direct matching, selecting a sentence s_i from S; extracting the rule set R^{s_i} associated with s_i; and if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designating the sentence s_i a valuable sentence and reducing S as $S = S - s_i$.

[0063] According to yet further examples, the method may include determining the relatedness between r_j and r_k, employing: $\mathrm{rel}(r_j, r_k) = \frac{\mathrm{hits}(r_j+r_k)}{\min(\mathrm{hits}(r_j),\mathrm{hits}(r_k))}$, where $\mathrm{rel}(r_j,r_k)$ is the relatedness between r_j and r_k, $\mathrm{hits}(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $\min(\mathrm{hits}(r_j),\mathrm{hits}(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately. The method may also include determining the fee associated with the selected sentences by employing a semantic reasoning engine, where the semantic reasoning engine returns one of a "consistent" result and an "inconsistent" result. The inconsistent result may be interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result may be interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

[0064] According to some examples, the fee may be computed according to: $f^{r_n} = \mu * \max_{r_x \in R^O, \mathrm{rel}(r_n,r_x) \geq \eta} [\mathrm{rel}(r_n, r_x) * P^{r_x}]$, where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness. The method may further include removing r_n from rule set R^{s_m} for the selected sentences until $R^{s_m} = \varnothing$; and determining a total fee F^{s_m} of selected sentences s_m as: $F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n}$. The semantic reasoning engine may employ first order predicate logic or probabilistic processing. The selection of one or more sentences may be determined based on one of copying or downloading of the one or more sentences. The method may also include enabling a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology or employing a description language to provide a logical formalism for the ontology.

[0065] According to other examples, a content provider service executed on one or more servers is described for providing text billing based on semantic data reasoning. The content provider service may include a textual content processing module. The textual content

processing module may price one or more rules in an ontology associated with textual content to be published; determine one or more valuable sentences within the textual content based on the priced rules; label the valuable sentences with price information; and publish the textual content.

[0066] According to further examples, the content provider service may include a decision and pricing module, which may receive selection of one or more valuable sentences within the published textual content from a client application; retrieve one or more rules associated with the selected sentences; and determine a fee associated with the selected sentences employing reasoning based on the retrieved rules. The textual content processing module may further receive an order from a textual content owner for valuable sentence rules; and determine a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, $|R|$ is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k.

[0067] According to yet other examples, the textual content processing module may if a rule among R is not ordered, designate a sentence associated with the unordered rule as a free sentence without a price to be charged. The textual content processing module may also divide the textual content into single sentences using logical separators to obtain a transformed sentence set; for each sentence in the transformed sentence set, determine a concept; map the determined concept onto the ontology; and if a direct link exists between the concept and the ontology, store a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

[0068] According to yet further examples, the textual content processing module may employ a two stage process comprising direct matching and relatedness computation. The textual content processing module may perform direct matching by: selecting a sentence s_i from a transformed sentence set S; extracting a rule set R^{s_i} associated with s_i; designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as $S = S - s_i$; and iteratively performing direct matching through S. The textual content processing module may also upon completion of the direct matching, select a sentence s_i from S; extract the rule set R^{s_i} associated with s_i; and if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designate the sentence s_i a valuable sentence and reduce S as $S = S - s_i$.

[0069] The textual content processing module may determine the relatedness between r_j and r_k employing: $rel(r_j, r_k) = \frac{hits(r_j+r_k)}{min(hits(r_j),hits(r_k))}$, where $rel(r_j,r_k)$ is the relatedness between r_j and r_k, $hits(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $min(hits(r_j),hits(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately. The decision and pricing module may determine the fee associated with the selected sentences by employing a semantic reasoning engine, and where the semantic reasoning engine returns one of a "consistent" result and an "inconsistent" result. The inconsistent result may be interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result may be interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

[0070] According to some examples, the fee may be computed according to: $f^{r_n} = \mu * \max_{r_x \in R^O, rel(r_n,r_x) \geq \eta} [rel(r_n, r_x) * P^{r_x}]$, where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness. The decision and pricing module may remove r_n from rule set R^{s_m} for the selected sentences until $R^{s_m} = \varnothing$; and determine a total fee F^{s_m} of selected sentences s_m as: $F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n}$. The semantic reasoning engine may employ first order predicate logic or probabilistic processing. The selection of one or more sentences may be determined based on one of copying or downloading of the one or more sentences. The textual content processing module may further enable a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology. The textual content processing module may also employ a description language to provide a logical formalism for the ontology.

[0071] According to further examples, a content provider service for providing text billing based on semantic data reasoning is described. The content provider service may a textual content processing module executed on a server. The textual content processing module may price one or more rules in an ontology associated with textual content to be published; determine one or more valuable sentences within the textual content based on the priced rules; label the valuable sentences with price information; and publish the textual content. The content service provider may further include a decision and pricing module executed as part of a client application. The decision and pricing module may receive selection of one or more valuable sentences within the published textual content from a client

application; retrieve one or more rules associated with the selected sentences; and determine a fee associated with the selected sentences employing reasoning based on the retrieved rules.

[0072] According to some examples, the textual content processing module may receive an order from a textual content owner for valuable sentence rules; and determine a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, $|R|$ is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k. The textual content processing module may further divide the textual content into single sentences using logical separators to obtain a transformed sentence set; for each sentence in the transformed sentence set, determine a concept; map the determined concept onto the ontology; and if a direct link exists between the concept and the ontology, store a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

[0073] According to other examples, the textual content processing module may iteratively perform direct matching through a transformed sentence set S by: selecting a sentence s_i from S; extracting a rule set R^{s_i} associated with s_i; and designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as $S = S - s_i$. The textual content processing module may also upon completion of the direct matching, select a sentence s_i from S; extract the rule set R^{s_i} associated with s_i; if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designate the sentence s_i a valuable sentence and reduce S as $S = S - s_i$.

[0074] According to yet other examples, the textual content processing module may determine the relatedness between r_j and r_k employing: $rel(r_j, r_k) = \frac{hits(r_j+r_k)}{min(hits(r_j),hits(r_k))}$, where $rel(r_j,r_k)$ is the relatedness between r_j and r_k, $hits(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $min(hits(r_j),hits(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately. The decision and pricing module may determine the fee associated with the selected sentences by employing a semantic reasoning engine, and where the semantic reasoning engine may return one of a "consistent" result and an "inconsistent" result with the inconsistent result being interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result being interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

[0075] According to further examples, the fee may be computed according to: $f^{r_n} = \mu * \max_{r_x \in R^O, rel(r_n, r_x) \geq \eta} [rel(r_n, r_x) * P^{r_x}]$, where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness. The decision and pricing module may remove r_n from rule set R^{s_m} for the selected sentences until $R^{s_m} = \emptyset$; and determine a total fee F^{s_m} of selected sentences s_m as: $F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n}$. The semantic reasoning engine may employ first order predicate logic or probabilistic processing. The selection of one or more sentences may be determined based on one of copying or downloading of the one or more sentences. The textual content processing module may enable a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology.

[0076] There is little distinction left between hardware and software implementations of aspects of systems; the use of hardware or software is generally (but not always, in that in certain contexts the choice between hardware and software may become significant) a design choice representing cost vs. efficiency tradeoffs. There are various vehicles by which processes and/or systems and/or other technologies described herein may be effected (*e.g.*, hardware, software, and/or firmware), and that the preferred vehicle will vary with the context in which the processes and/or systems and/or other technologies are deployed. For example, if an implementer determines that speed and accuracy are paramount, the implementer may opt for a mainly hardware and/or firmware vehicle; if flexibility is paramount, the implementer may opt for a mainly software implementation; or, yet again alternatively, the implementer may opt for some combination of hardware, software, and/or firmware.

[0077] The foregoing detailed description has set forth various embodiments of the devices and/or processes via the use of block diagrams, flowcharts, and/or examples. Insofar as such block diagrams, flowcharts, and/or examples contain one or more functions and/or operations, it will be understood by those within the art that each function and/or operation within such block diagrams, flowcharts, or examples may be implemented, individually and/or collectively, by a wide range of hardware, software, firmware, or virtually any combination thereof. In one embodiment, several portions of the subject matter described herein may be implemented via Application Specific Integrated Circuits (ASICs), Field Programmable Gate Arrays (FPGAs), digital signal processors (DSPs), or other integrated formats. However, those skilled in the art will recognize that some aspects of the embodiments disclosed herein, in whole or in part, may be equivalently implemented in

integrated circuits, as one or more computer programs running on one or more computers (*e.g.*, as one or more programs running on one or more computer systems), as one or more programs running on one or more processors (*e.g.*, as one or more programs running on one or more microprocessors), as firmware, or as virtually any combination thereof, and that designing the circuitry and/or writing the code for the software and or firmware would be well within the skill of one of skill in the art in light of this disclosure.

[0078] The present disclosure is not to be limited in terms of the particular embodiments described in this application, which are intended as illustrations of various aspects. Many modifications and variations can be made without departing from its spirit and scope, as will be apparent to those skilled in the art. Functionally equivalent methods and apparatuses within the scope of the disclosure, in addition to those enumerated herein, will be apparent to those skilled in the art from the foregoing descriptions. Such modifications and variations are intended to fall within the scope of the appended claims. The present disclosure is to be limited only by the terms of the appended claims, along with the full scope of equivalents to which such claims are entitled. It is to be understood that this disclosure is not limited to particular methods, reagents, compounds compositions or biological systems, which can, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting.

[0079] In addition, those skilled in the art will appreciate that the mechanisms of the subject matter described herein are capable of being distributed as a program product in a variety of forms, and that an illustrative embodiment of the subject matter described herein applies regardless of the particular type of signal bearing medium used to actually carry out the distribution. Examples of a signal bearing medium include, but are not limited to, the following: a recordable type medium such as a floppy disk, a hard disk drive, a Compact Disc (CD), a Digital Versatile Disk (DVD), a digital tape, a computer memory, a solid state drive, *etc.*; and a transmission type medium such as a digital and/or an analog communication medium (*e.g.*, a fiber optic cable, a waveguide, a wired communications link, a wireless communication link, *etc.*).

[0080] Those skilled in the art will recognize that it is common within the art to describe devices and/or processes in the fashion set forth herein, and thereafter use engineering practices to integrate such described devices and/or processes into data processing systems. That is, at least a portion of the devices and/or processes described herein may be integrated into a data processing system via a reasonable amount of experimentation. Those having skill in the art will recognize that a typical data processing system generally includes one or more

ot a system unit housing, a video display device, a memory such as volatile and non-volatile memory, processors such as microprocessors and digital signal processors, computational entities such as operating systems, drivers, graphical user interfaces, and applications programs, one or more interaction devices, such as a touch pad or screen, and/or control systems including feedback loops and control motors (*e.g.*, feedback for sensing position and/or velocity of gantry systems; control motors for moving and/or adjusting components and/or quantities).

[0081] A typical data processing system may be implemented utilizing any suitable commercially available components, such as those typically found in data computing/communication and/or network computing/communication systems. The herein described subject matter sometimes illustrates different components contained within, or connected with, different other components. It is to be understood that such depicted architectures are merely exemplary, and that in fact many other architectures may be implemented which achieve the same functionality. In a conceptual sense, any arrangement of components to achieve the same functionality is effectively "associated" such that the desired functionality is achieved. Hence, any two components herein combined to achieve a particular functionality may be seen as "associated with" each other such that the desired functionality is achieved, irrespective of architectures or intermediate components. Likewise, any two components so associated may also be viewed as being "operably connected", or "operably coupled", to each other to achieve the desired functionality, and any two components capable of being so associated may also be viewed as being "operably couplable", to each other to achieve the desired functionality. Specific examples of operably couplable include but are not limited to physically connectable and/or physically interacting components and/or wirelessly interactable and/or wirelessly interacting components and/or logically interacting and/or logically interactable components.

[0082] With respect to the use of substantially any plural and/or singular terms herein, those having skill in the art can translate from the plural to the singular and/or from the singular to the plural as is appropriate to the context and/or application. The various singular/plural permutations may be expressly set forth herein for sake of clarity.

[0083] It will be understood by those within the art that, in general, terms used herein, and especially in the appended claims (*e.g.*, bodies of the appended claims) are generally intended as "open" terms (*e.g.*, the term "including" should be interpreted as "including but not limited to," the term "having" should be interpreted as "having at least," the term "includes" should be interpreted as "includes but is not limited to," *etc.*). It will be further

understood by those within the art that if a specific number of an introduced claim recitation is intended, such an intent will be explicitly recited in the claim, and in the absence of such recitation no such intent is present. For example, as an aid to understanding, the following appended claims may contain usage of the introductory phrases "at least one" and "one or more" to introduce claim recitations. However, the use of such phrases should not be construed to imply that the introduction of a claim recitation by the indefinite articles "a" or "an" limits any particular claim containing such introduced claim recitation to embodiments containing only one such recitation, even when the same claim includes the introductory phrases "one or more" or "at least one" and indefinite articles such as "a" or "an" (*e.g.*, "a" and/or "an" should be interpreted to mean "at least one" or "one or more"); the same holds true for the use of definite articles used to introduce claim recitations. In addition, even if a specific number of an introduced claim recitation is explicitly recited, those skilled in the art will recognize that such recitation should be interpreted to mean at least the recited number (*e.g.*, the bare recitation of "two recitations," without other modifiers, means at least two recitations, or two or more recitations).

[0084] Furthermore, in those instances where a convention analogous to "at least one of A, B, and C, *etc.*" is used, in general such a construction is intended in the sense one having skill in the art would understand the convention (*e.g.*, " a system having at least one of A, B, and C" would include but not be limited to systems that have A alone, B alone, C alone, A and B together, A and C together, B and C together, and/or A, B, and C together, *etc.*). It will be further understood by those within the art that virtually any disjunctive word and/or phrase presenting two or more alternative terms, whether in the description, claims, or drawings, should be understood to contemplate the possibilities of including one of the terms, either of the terms, or both terms. For example, the phrase "A or B" will be understood to include the possibilities of "A" or "B" or "A and B."

[0085] In addition, where features or aspects of the disclosure are described in terms of Markush groups, those skilled in the art will recognize that the disclosure is also thereby described in terms of any individual member or subgroup of members of the Markush group.

[0086] As will be understood by one skilled in the art, for any and all purposes, such as in terms of providing a written description, all ranges disclosed herein also encompass any and all possible subranges and combinations of subranges thereof. Any listed range can be easily recognized as sufficiently describing and enabling the same range being broken down into at least equal halves, thirds, quarters, fifths, tenths, *etc.* As a non-limiting example, each range discussed herein can be readily broken down into a lower third, middle third and upper

third, *etc.* As will also be understood by one skilled in the art all language such as "up to," "at least," "greater than," "less than," and the like include the number recited and refer to ranges which can be subsequently broken down into subranges as discussed above. Finally, as will be understood by one skilled in the art, a range includes each individual member. Thus, for example, a group having 1-3 cells refers to groups having 1, 2, or 3 cells. Similarly, a group having 1-5 cells refers to groups having 1, 2, 3, 4, or 5 cells, and so forth.

[0087] While various aspects and embodiments have been disclosed herein, other aspects and embodiments will be apparent to those skilled in the art. The various aspects and embodiments disclosed herein are for purposes of illustration and are not intended to be limiting, with the true scope and spirit being indicated by the following claims.

CLAIMS

WHAT IS CLAIMED IS:

1. A method for providing text billing based on semantic data reasoning, the method comprising:

pricing one or more rules in an ontology associated with textual content to be published;

determining one or more valuable sentences within the textual content based on the priced rules;

labeling the valuable sentences with price information; and

publishing the textual content.

2. The method of claim 1, further comprising:

receiving selection of one or more valuable sentences within the published textual content;

retrieving one or more rules associated with the selected sentences; and

determining a fee associated with the selected sentences employing reasoning based on the retrieved rules.

3. The method of claim 2, wherein pricing the one or more rules in the ontology comprises:

receiving an order from a textual content owner for valuable sentence rules; and

determining a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, $|R|$ is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k.

4. The method of claim 3, further comprising:

if a rule among R is not ordered, designating a sentence associated with the unordered rule as a free sentence without a price to be charged.

5. The method of claim 2, further comprising:

dividing the textual content into single sentences using logical separators to obtain a transformed sentence set;

for each sentence in the transformed sentence set, determining a concept;

mapping the determined concept onto the ontology; and

if a direct link exists between the concept and the ontology, storing a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

6. The method of claim 2, wherein determining the one or more valuable sentences comprises:

employing a two stage process comprising direct matching and relatedness computation.

7. The method of claim 6, wherein the direct matching comprises:

selecting a sentence s_i from a transformed sentence set S;

extracting a rule set R^{s_i} associated with s_i;

designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as $S = S - s_i$; and

iteratively performing direct matching through S.

8. The method of claim 7, further comprising:

upon completion of the direct matching, selecting a sentence s_i from S;

extracting the rule set R^{s_i} associated with s_i; and

if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designating the sentence s_i a valuable sentence and reducing S as $S = S - s_i$.

9. The method of claim 8, further comprising determining the relatedness between r_j and r_k, employing: $\mathrm{rel}(r_j, r_k) = \frac{\mathrm{hits}(r_j+r_k)}{\min(\mathrm{hits}(r_j),\mathrm{hits}(r_k))}$, where $\mathrm{rel}(r_j,r_k)$ is the relatedness between r_j and r_k, $\mathrm{hits}(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $\min(\mathrm{hits}(r_j),\mathrm{hits}(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately.

10. The method of claim 9, further comprising determining the fee associated with the selected sentences by employing a semantic reasoning engine, wherein the semantic reasoning engine returns one of a “consistent” result and an “inconsistent” result.

11. The method of claim 10, wherein the inconsistent result is interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result is interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

12. The method of claim 11, wherein the fee is computed according to:

$$f^{r_n} = \mu * \max_{r_x \in R^O, rel(r_n, r_x) \geq \eta} [rel(r_n, r_x) * P^{r_x}],$$

where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness.

13. The method of claim 11, further comprising:

removing r_n from rule set R^{s_m} for the selected sentences until $R^{s_m} = \varnothing$; and

determining a total fee F^{s_m} of selected sentences s_m as:

$$F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n}.$$

14. The method of claim 11, wherein the semantic reasoning engine employs one of first order predicate logic and probabilistic processing.

15. The method of claim 2, wherein the selection of one or more sentences is determined based on one of copying or downloading of the one or more sentences.

16. The method of claim 2, further comprising enabling a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology.

17. The method of claim 2, further comprising employing a description language to provide a logical formalism for the ontology.

18. A content provider service executed on one or more servers for providing text billing based on semantic data reasoning, the content provider service comprising:

a textual content processing module configured to:

price one or more rules in an ontology associated with textual content to be published;

determine one or more valuable sentences within the textual content based on the priced rules;

label the valuable sentences with price information; and

publish the textual content.

19. The content provider service of claim 18, further comprising a decision and pricing module configured to:

receive selection of one or more valuable sentences within the published textual content from a client application;

retrieve one or more rules associated with the selected sentences; and

determine a fee associated with the selected sentences employing reasoning based on the retrieved rules.

20. The content provider service of claim 19, wherein the textual content processing module is further configured to:

receive an order from a textual content owner for valuable sentence rules; and

determine a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, |R| is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k.

21. The content provider service of claim 20, wherein the textual content processing module is further configured to:

if a rule among R is not ordered, designate a sentence associated with the unordered rule as a free sentence without a price to be charged.

22. The content provider service of claim 19, wherein the textual content processing module is further configured to:

divide the textual content into single sentences using logical separators to obtain a transformed sentence set;

for each sentence in the transformed sentence set, determine a concept;

map the determined concept onto the ontology; and

if a direct link exists between the concept and the ontology, store a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

23. The content provider service of claim 19, wherein the textual content processing module is further configured to:

employ a two stage process comprising direct matching and relatedness computation.

24. The content provider service of claim 23, wherein the textual content processing module is further configured to perform direct matching by:

selecting a sentence s_i from a transformed sentence set S;

extracting a rule set R^{s_i} associated with s_i;

designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as $S = S - s_i$; and

iteratively performing direct matching through S.

25. The content provider service of claim 24, wherein the textual content processing module is further configured to:

upon completion of the direct matching, select a sentence s_i from S;

extract the rule set R^{s_i} associated with s_i; and

if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designate the sentence s_i a valuable sentence and reduce S as $S = S - s_i$.

26. The content provider service of claim 25, wherein the textual content processing module is further configured to determine the relatedness between r_j and r_k employing:

$rel(r_j, r_k) = \frac{hits(r_j+r_k)}{min(hits(r_j),hits(r_k))}$, where $rel(r_j,r_k)$ is the relatedness between r_j and r_k, $hits(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $min(hits(r_j),hits(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately.

27. The content provider service of claim 26, wherein the decision and pricing module is further configured to determine the fee associated with the selected sentences by employing a semantic reasoning engine, and wherein the semantic reasoning engine returns one of a "consistent" result and an "inconsistent" result.

28. The content provider service of claim 27, wherein the inconsistent result is interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result is interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

29. The content provider service of claim 27, wherein the fee is computed according to:

$$f^{r_n} = \mu * \max_{r_x \in R^O, rel(r_n, r_x) \geq \eta} [rel(r_n, r_x) * P^{r_x}],$$

where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness.

30. The content provider service of claim 29, wherein the decision and pricing module is further configured to:

remove r_n from rule set R^{sm} for the selected sentences until $R^{sm} = \varnothing$; and

determine a total fee F^{sm} of selected sentences s_m as:

$$F^{sm} = \sum_{r_n \in R^{sm}} f^{r_n}.$$

31. The content provider service of claim 27, wherein the semantic reasoning engine employs one of first order predicate logic and probabilistic processing.

32. The content provider service of claim 19, wherein the selection of one or more sentences is determined based on one of copying or downloading of the one or more sentences.

33. The content provider service of claim 19, wherein the textual content processing module is further configured to enable a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology.

34. The content provider service of claim 19, wherein the textual content processing module is further configured to employ a description language to provide a logical formalism for the ontology.

35. A content provider service for providing text billing based on semantic data reasoning, the content provider service comprising:

a textual content processing module executed on a server, the textual content processing module configured to:

price one or more rules in an ontology associated with textual content to be published;

determine one or more valuable sentences within the textual content based on the priced rules;

label the valuable sentences with price information; and

publish the textual content; and

a decision and pricing module executed as part of a client application, the decision and pricing module configured to:

receive selection of one or more valuable sentences within the published textual content from a client application;

retrieve one or more rules associated with the selected sentences; and

determine a fee associated with the selected sentences employing reasoning based on the retrieved rules.

36. The content provider service of claim 35, wherein the textual content processing module is further configured to:

receive an order from a textual content owner for valuable sentence rules; and

determine a price for each rule based on $P^{r_k} = \frac{total*(|R|+1)}{2|R|*\,seq^{r_k}}$, where P^{r_k} is the price of rule r_k, *total* is total price of the whole text, $|R|$ is total number of valuable rules included in the ontology O, and seq^{r_k} is a sequence of r_k.

37. The content provider service of claim 35, wherein the textual content processing module is further configured to:

divide the textual content into single sentences using logical separators to obtain a transformed sentence set;

for each sentence in the transformed sentence set, determine a concept;

map the determined concept onto the ontology; and

if a direct link exists between the concept and the ontology, store a rule associated with the sentence in a rule set $R^{s_i} = \{r_1,\ldots,r_j,\ldots\}$, where R^{s_i} denotes the rule set with respect to a sentence s_i within the textual content, and removing the sentence s_i from S.

38. The content provider service of claim 35, wherein the textual content processing module is further configured to:

iteratively perform direct matching through a transformed sentence set S by:

selecting a sentence s_i from S;

extracting a rule set R^{s_i} associated with s_i; and

designating s_i as a free sentence, if there does not exist a rule r_j that satisfies $P^{r_j} \neq 0$, where $r_j \in R^{s_i}$, else designating s_i as a valuable sentence and reducing S as S = S - s_i.

39. The content provider service of claim 38, wherein the textual content processing module is further configured to:

upon completion of the direct matching, select a sentence s_i from S;

extract the rule set R^{s_i} associated with s_i; and

if the relatedness between r_j and r_k, where $r_j \in R^{s_i}$, $r_k \in R^O$, and $P^{r_k} = 0$, exceeds a predetermined threshold η, designate the sentence s_i a valuable sentence and reduce S as S = S - s_i.

40. The content provider service of claim 39, wherein the textual content processing module is further configured to determine the relatedness between r_j and r_k employing:

$rel(r_j, r_k) = \frac{hits(r_j+r_k)}{min(hits(r_j),hits(r_k))}$, where $rel(r_j,r_k)$ is the relatedness between r_j and r_k, $hits(r_j+r_k)$ is a number of pages returned by searching a first word and a second word, and $min(hits(r_j),hits(r_k))$ is a minimum number of pages returned by searching the first word and the second word separately.

41. The content provider service of claim 40, wherein the decision and pricing module is further configured to determine the fee associated with the selected sentences by employing a semantic reasoning engine, and wherein the semantic reasoning engine returns one of a "consistent" result and an "inconsistent" result with the inconsistent result being interpreted as a retrieved rule indicating a sentence associated with the retrieved rule being free of charge and the consistent result being interpreted as the retrieved rule indicating the sentence associated with the retrieved rule being associated with a fee.

42. The content provider service of claim 41, wherein the fee is computed according to:

$$f^{r_n} = \mu * \max_{r_x \in R^O, rel(r_n, r_x) \geq \eta} [rel(r_n, r_x) * P^{r_x}],$$

where μ is an adjustment coefficient between 1 and 0, P^{r_k} is a price for rule r_x, R^O is a set of rules associated with the ontology, and η is a predetermined threshold for relatedness.

43. The content provider service of claim 42, wherein the decision and pricing module is further configured to:

remove r_n from rule set R^{s_m} for the selected sentences until $R^{s_m} = \varnothing$; and

determine a total fee F^{s_m} of selected sentences s_m as:

$$F^{s_m} = \sum_{r_n \in R^{s_m}} f^{r_n}.$$

44. The content provider service of claim 43, wherein the semantic reasoning engine employs one of first order predicate logic and probabilistic processing.

45. The content provider service of claim 35, wherein the selection of one or more sentences is determined based on one of copying or downloading of the one or more sentences.

46. The content provider service of claim 35, wherein the textual content processing module is further configured to enable a textual content owner to modify pricing for one or more sentences by adjusting rule pricing in the ontology.

47. A computer readable storage medium with instructions stored thereon, which when executed on one or more computing devices execute a method for providing text billing based on semantic data reasoning, wherein the method includes action of claims 1 through 17.

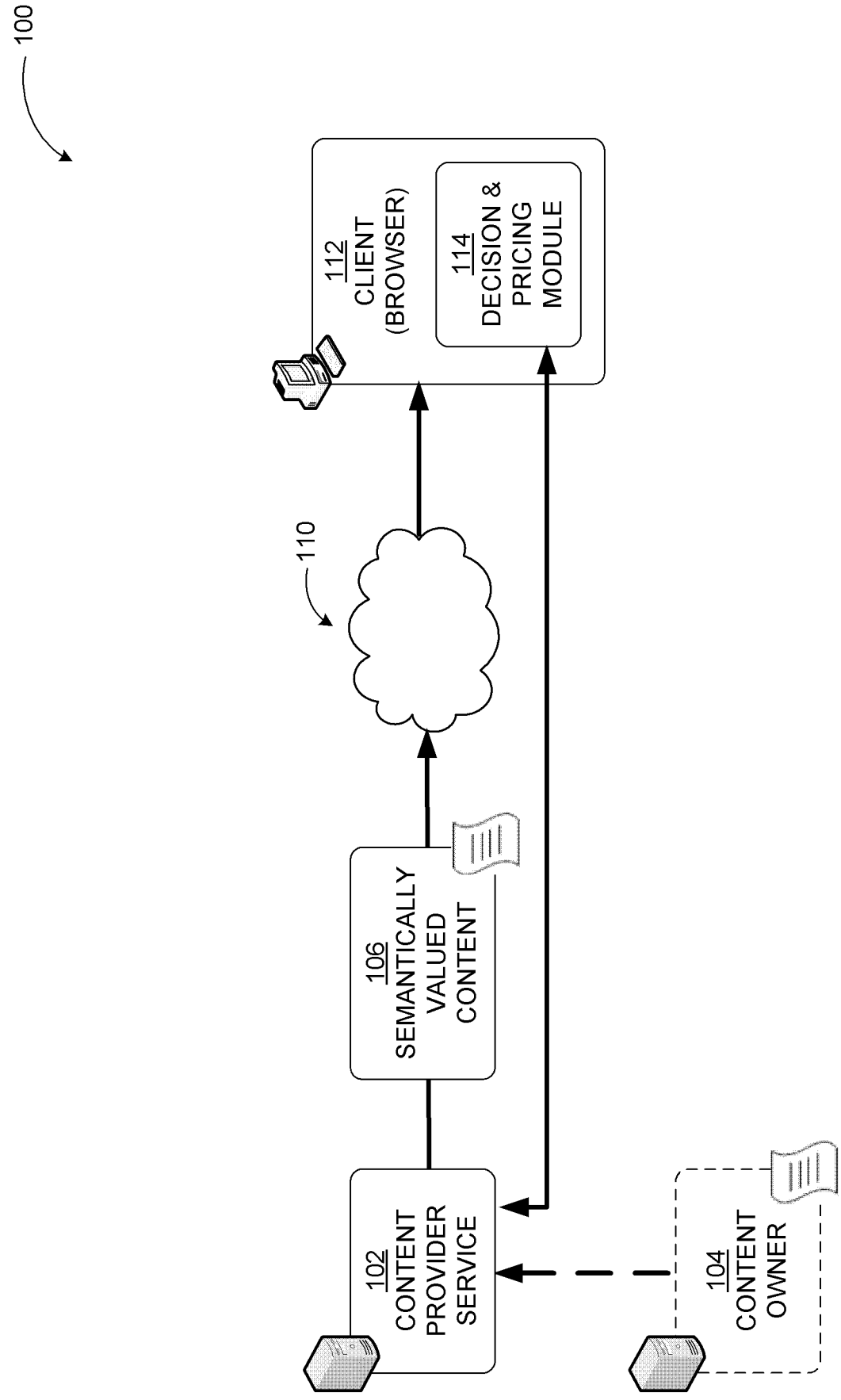
100
112
CLIENT
(BROWSER)
114
DECISION &
PRICING
MODULE
110
106
SEMANTICALLY
VALUED
CONTENT
102
CONTENT
PROVIDER
SERVICE
104
CONTENT
OWNER

FIG. 1

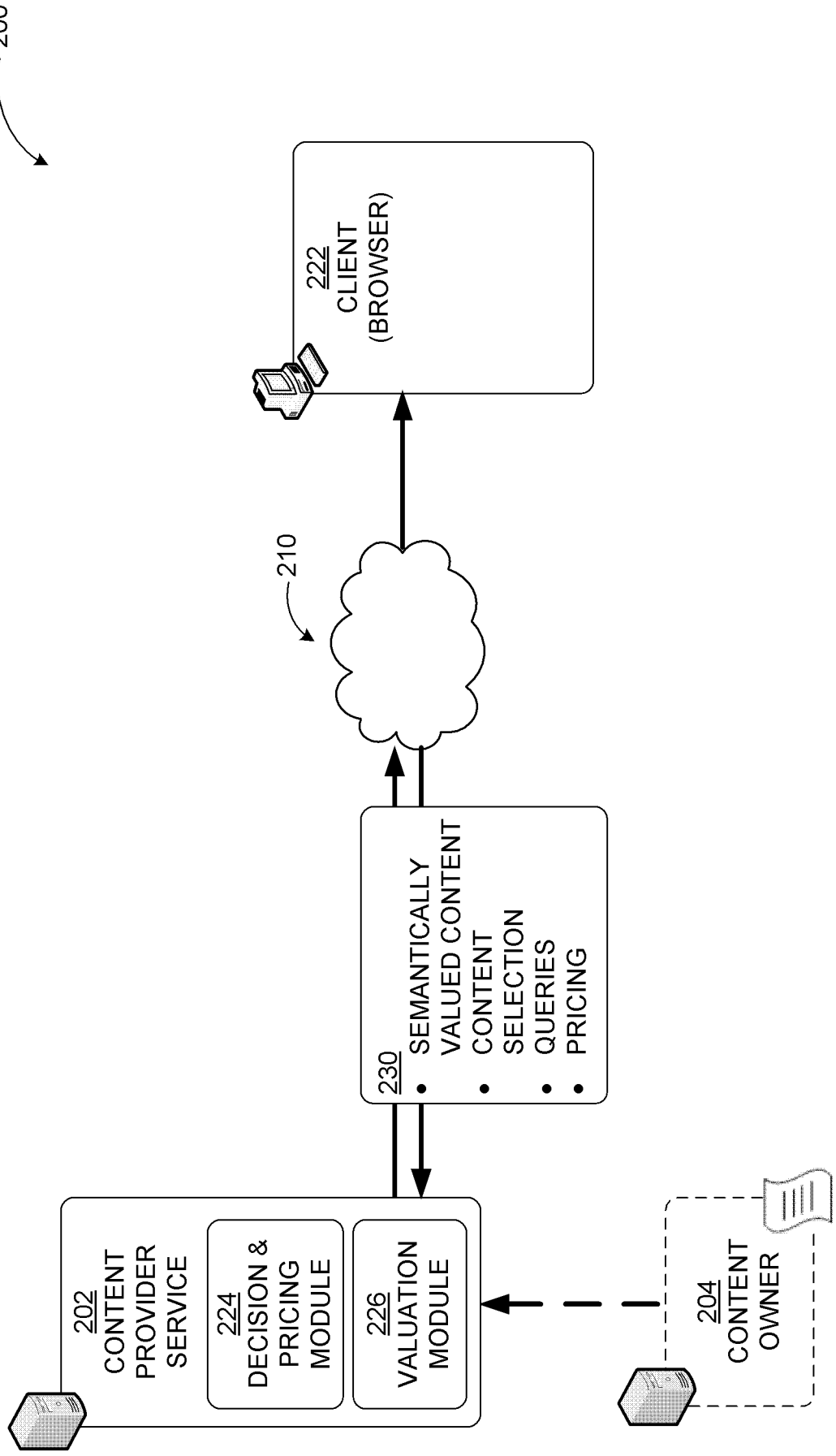
200
222
CLIENT
(BROWSER)
210
230
SEMANTICALLY
VALUED CONTENT
CONTENT
SELECTION
QUERIES
PRICING
202
CONTENT
PROVIDER
SERVICE
224
DECISION &
PRICING
MODULE
226
VALUATION
MODULE
204
CONTENT
OWNER

FIG. 2

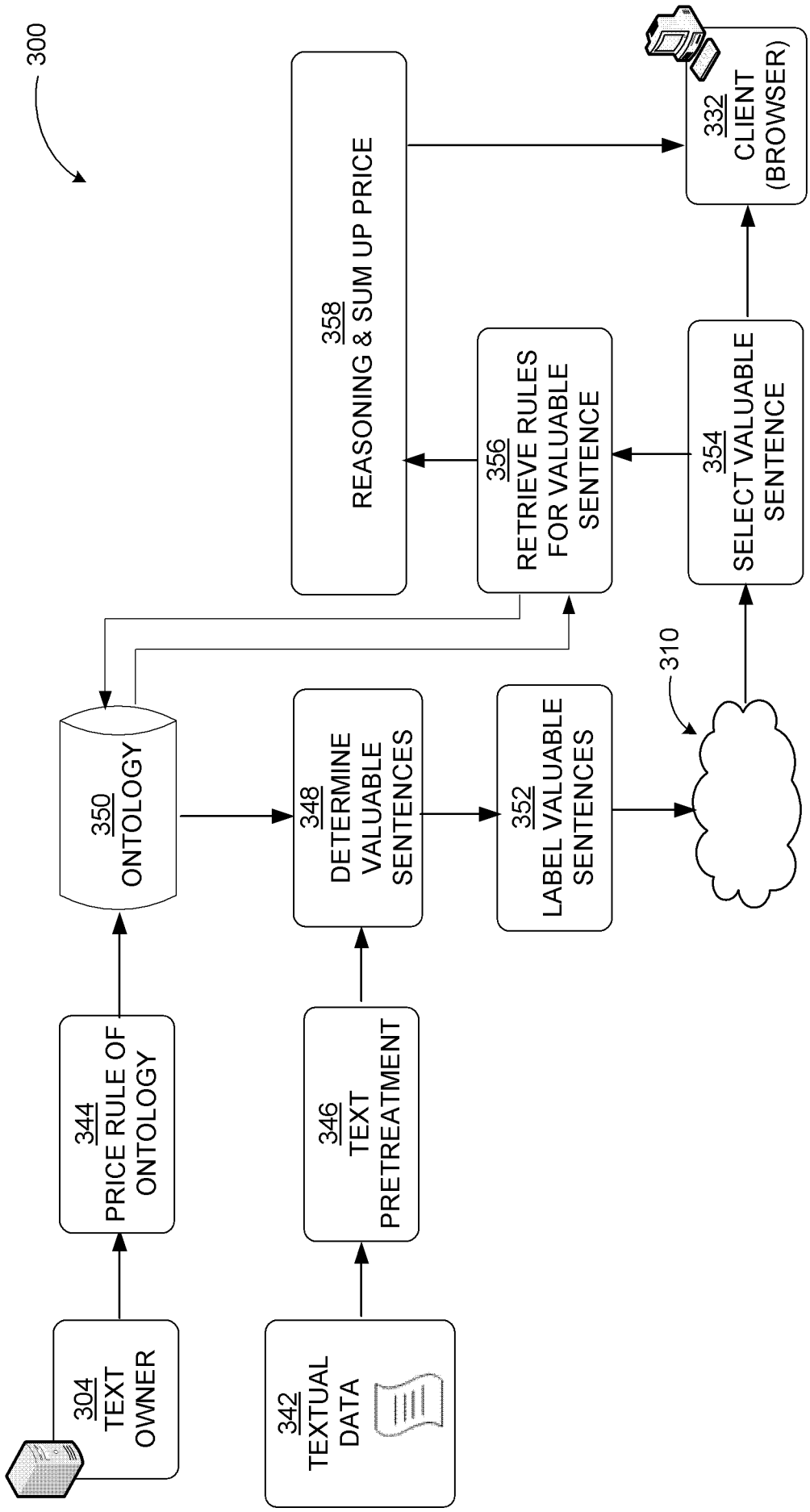
300
304 TEXT OWNER
342 TEXTUAL DATA
344 PRICE RULE OF ONTOLOGY
346 TEXT PRETREATMENT
350 ONTOLOGY
348 DETERMINE VALUABLE SENTENCES
352 LABEL VALUABLE SENTENCES
310
354 SELECT VALUABLE SENTENCE
356 RETRIEVE RULES FOR VALUABLE SENTENCE
358 REASONING & SUM UP PRICE
332 CLIENT (BROWSER)

FIG. 3

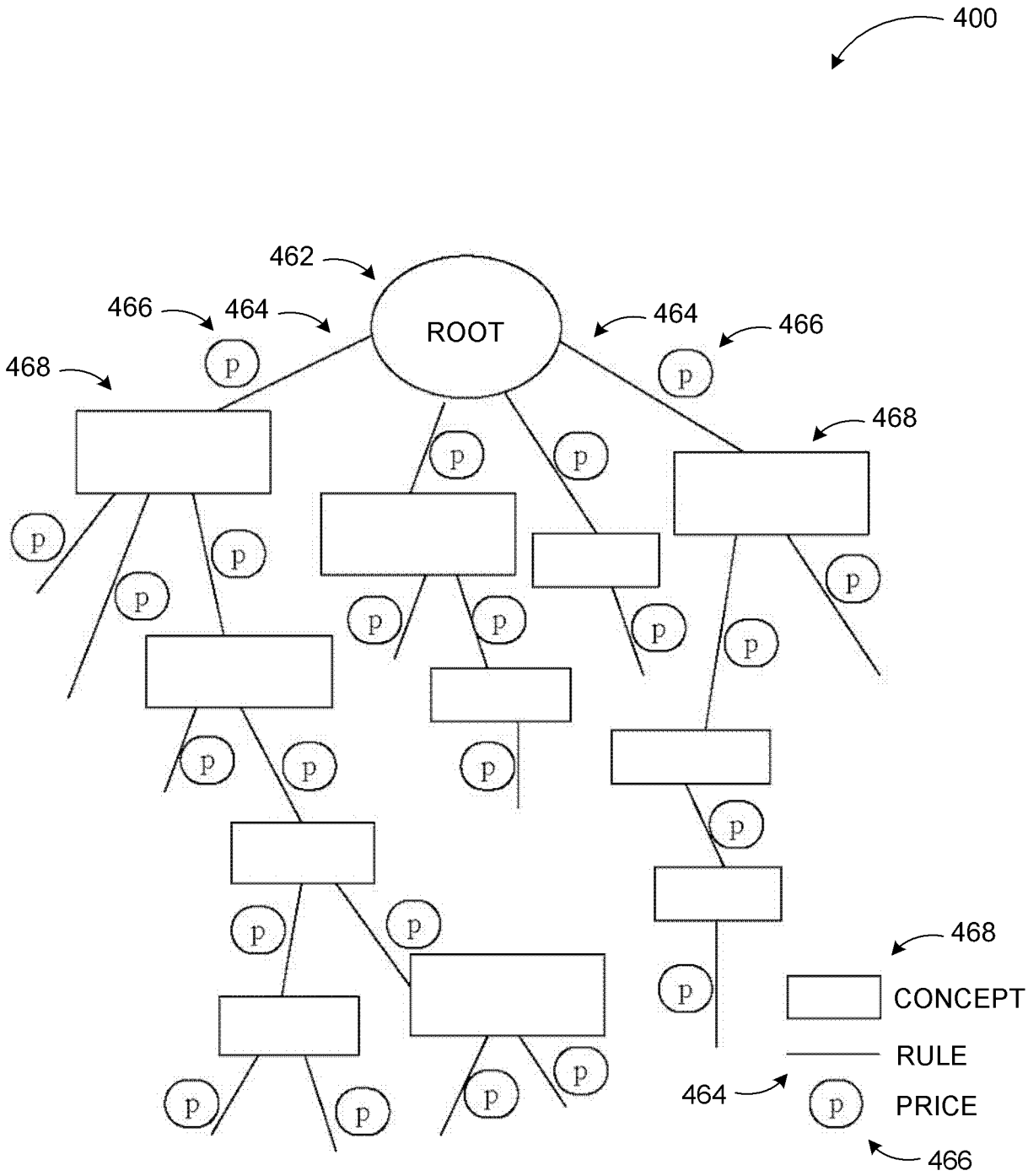
400
462
ROOT
466
464
464
466
468
468
468
CONCEPT
RULE
464
PRICE
466

FIG. 4

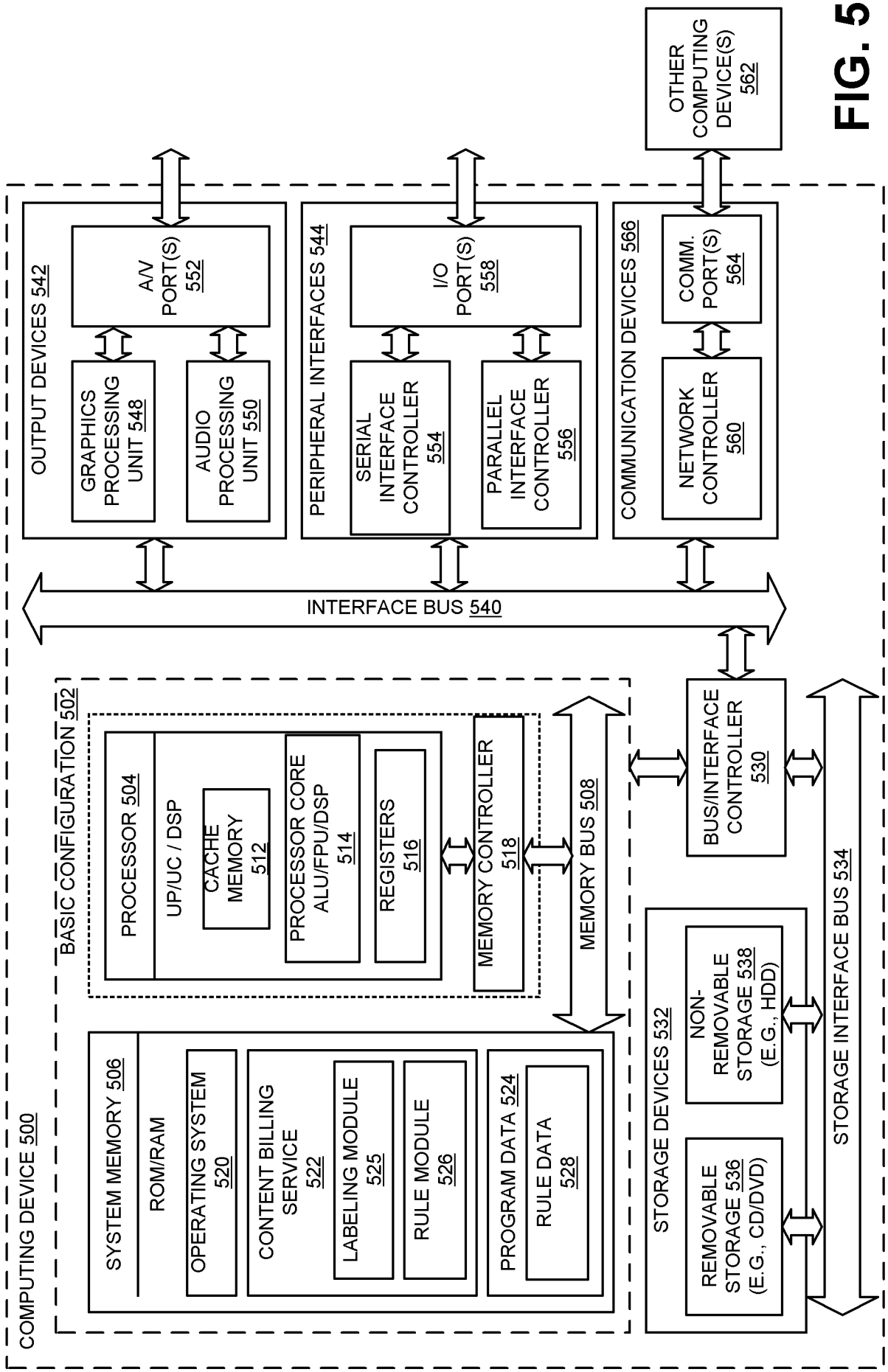
COMPUTING DEVICE 500
BASIC CONFIGURATION 502
PROCESSOR 504
UP/UC / DSP
CACHE MEMORY 512
PROCESSOR CORE ALU/FPU/DSP 514
REGISTERS 516
MEMORY CONTROLLER 518
SYSTEM MEMORY 506
ROM/RAM
OPERATING SYSTEM 520
CONTENT BILLING SERVICE 522
LABELING MODULE 525
RULE MODULE 526
PROGRAM DATA 524
RULE DATA 528
MEMORY BUS 508
BUS/INTERFACE CONTROLLER 530
STORAGE DEVICES 532
REMOVABLE STORAGE 536 (E.G., CD/DVD)
NON-REMOVABLE STORAGE 538 (E.G., HDD)
STORAGE INTERFACE BUS 534
INTERFACE BUS 540
OUTPUT DEVICES 542
GRAPHICS PROCESSING UNIT 548
AUDIO PROCESSING UNIT 550
A/V PORT(S) 552
PERIPHERAL INTERFACES 544
SERIAL INTERFACE CONTROLLER 554
PARALLEL INTERFACE CONTROLLER 556
I/O PORT(S) 558
COMMUNICATION DEVICES 566
NETWORK CONTROLLER 560
COMM. PORT(S) 564
OTHER COMPUTING DEVICE(S) 562

FIG. 5

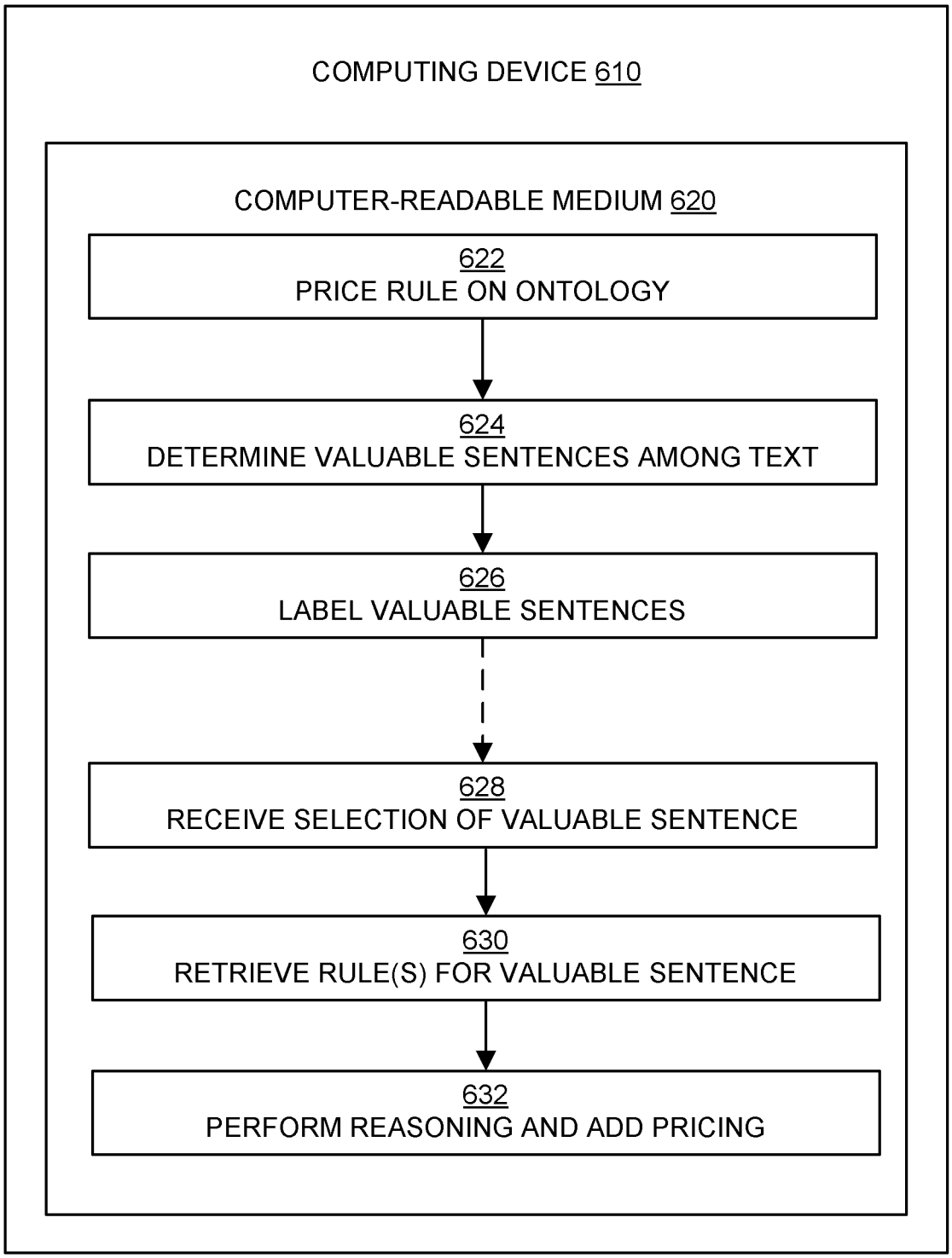
COMPUTING DEVICE 610
COMPUTER-READABLE MEDIUM 620
622
PRICE RULE ON ONTOLOGY
624
DETERMINE VALUABLE SENTENCES AMONG TEXT
626
LABEL VALUABLE SENTENCES
628
RECEIVE SELECTION OF VALUABLE SENTENCE
630
RETRIEVE RULE(S) FOR VALUABLE SENTENCE
632
PERFORM REASONING AND ADD PRICING

FIG. 6

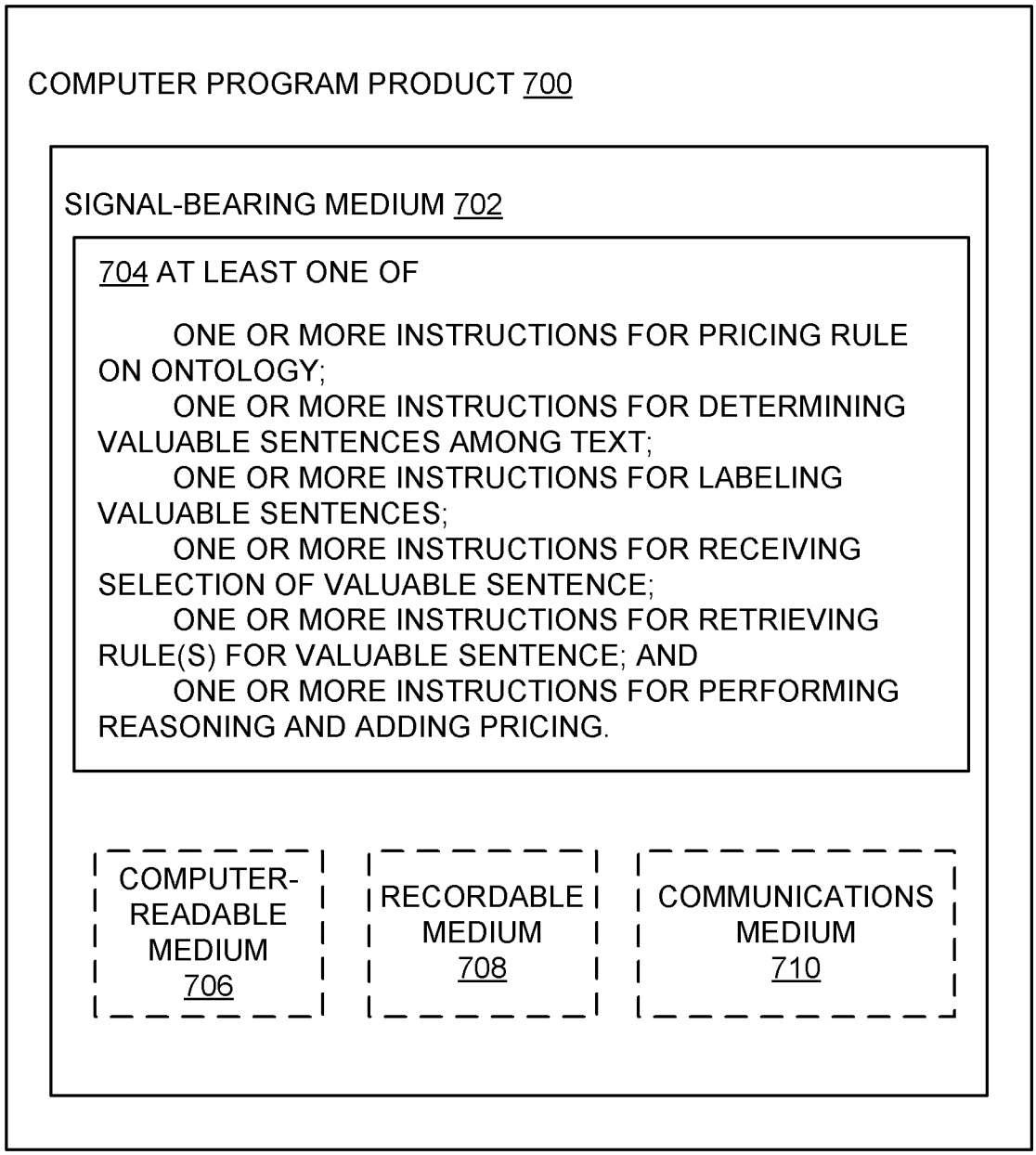
COMPUTER PROGRAM PRODUCT 700
SIGNAL-BEARING MEDIUM 702
704 AT LEAST ONE OF
ONE OR MORE INSTRUCTIONS FOR PRICING RULE ON ONTOLOGY;
ONE OR MORE INSTRUCTIONS FOR DETERMINING VALUABLE SENTENCES AMONG TEXT;
ONE OR MORE INSTRUCTIONS FOR LABELING VALUABLE SENTENCES;
ONE OR MORE INSTRUCTIONS FOR RECEIVING SELECTION OF VALUABLE SENTENCE;
ONE OR MORE INSTRUCTIONS FOR RETRIEVING RULE(S) FOR VALUABLE SENTENCE; AND
ONE OR MORE INSTRUCTIONS FOR PERFORMING REASONING AND ADDING PRICING.
COMPUTER-READABLE MEDIUM 706
RECORDABLE MEDIUM 708
COMMUNICATIONS MEDIUM 710

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/070121

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/20 (2006.01) i

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F 15/, G06F 17/, G06N 5/, G06Q 10/, G06Q 30/, G06Q 50/

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)

CNABS,DWPI: sentence?, content+, rule?, text?, semantic, ontology, word?, pric+, min+, associat+, reason+, bill+, principle?, string?, passage?, valu+, threshold+, relat+, conclu+, inferenc+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US2007055569A1 (CONTEXTWEB) 08 March 2007 (08.03.2007) the abstract, the description paragraphs [0032] to [0046] and [0083], the figures 1 and 14	1-2,6,15-19,23,32-35,45-47
A		3-5,7-14,20-22,24-31,36-44
A	US7693704B2 (Lexis-Nexis GROUP, a division of Reed Elsevier Inc.) 06 April 2010 (06.04.2010) the whole document	1-47
A	CN101620596A (University of Northeastern) 06 January 2010 (06.01.2010) the whole document	1-47

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

* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 September 2013 (30.09.2013)	**10 Oct. 2013 (10.10.2013)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer GUO,Yongqiang Telephone No. (86-10)62412084

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/070121

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US2007055569A1	08.03.2007	WO2007022137A2	22.02.2007
		WO2007022137A3	21.12.2007
		EP1913544A2	23.04.2008
		AU2006279694A1	22.02.2007
		IN200800128P2	05.12.2008
		JP2009505246A	05.02.2009
		CN101253526A	27.08.2008
		CN101253526B	19.10.2011
US7693704B2	06.04.2010	US2005144169A1	30.06.2005
CN101620596A	06.01.2010	CN101620596B	15.02.2012