

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
9 August 2001 (09.08.2001)

PCT

(10) International Publication Number
WO 01/56388 A2

(51) International Patent Classification⁷: **A04B**
(21) International Application Number: PCT/US01/03670
(22) International Filing Date: 5 February 2001 (05.02.2001)
(25) Filing Language: English
(26) Publication Language: English
(30) Priority Data:
09/498,910 4 February 2000 (04.02.2000) US
(71) Applicant: **GENERAL DYNAMICS INFORMATION SYSTEMS, INC.** [US/US]; 8800 Queen Avenue South, Bloomington, MN 55431 (US).

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

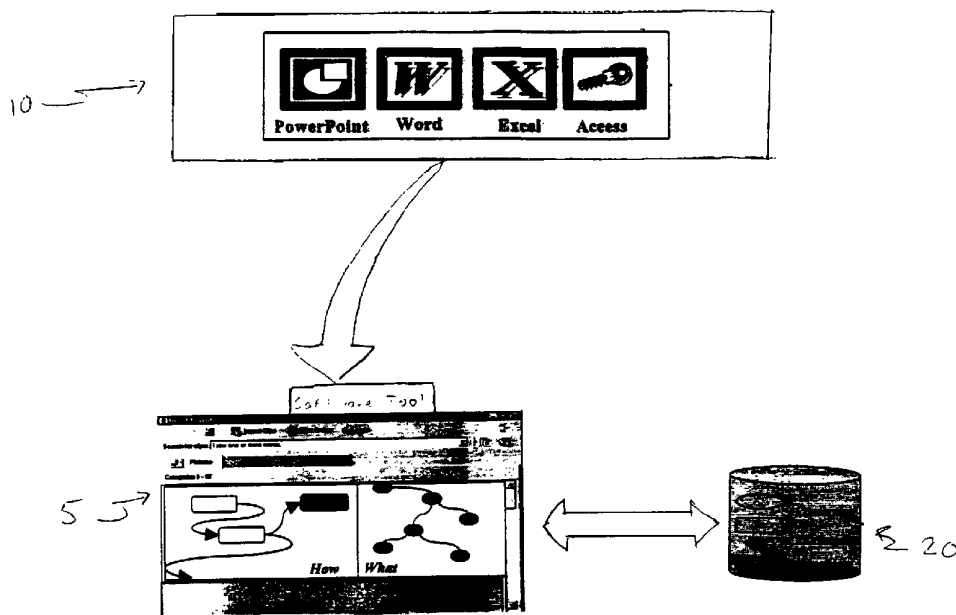
— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(72) Inventors: **MORK, Jay, Edward**; 4683 Parkridge Drive, Eagan, MN 55123 (US). **PHIPPS, Maria, Jeanne**; 8434 Spruce Court, Victoria, MN 55386 (US). **MURPHY, Michael, Emmett**; 3001 Foxpoint Road, Burnsville, MN 55337 (US).

(74) Agent: **MARINELLI, Joseph, F.**; Jenner & Block, One IBM Plaza, Chicago, IL 60611 (US).

(54) Title: APPARATUS AND METHOD FOR CREATING, MANAGING AND SHARING INFORMATION



(57) Abstract: A system allows a user to annotate documents by connecting any type of information in the document to concepts and/or processes making up a semantic ontology. The annotated information can be stored and accessed later to facilitate creation of new documents of a similar type. The system can use the information from the ontologies to automatically generate new documents by knowing what type of information to include and how to get the information. The system also allows users to create and edit ontologies. The system is user friendly in that it can be integrated with existing software applications with which users already are familiar.



WO 01/56388 A2

APPARATUS AND METHOD FOR
CREATING, MANAGING AND SHARING INFORMATION

Background of the Invention

1. Field of the Invention

The present invention relates to methods and devices for creating, managing, and sharing information. More particularly, the present invention relates to a software tool for creating, managing, and sharing information by applying semantic ontologies to bodies of information.

2. Description of the Related Art

Anyone who creates a document, such as a presentation, report, brief, paper, article or any other form of communication spends considerable time and energy in deciding what information to include in the document and how to obtain that information. For example, a business executive might create a business report that will be used to educate other business personnel. In creating the report, the executive must make several decisions. Among these decisions, executive must decide *what* information is available to include in the report, *what* information to include in the report, and *how* to obtain that information.

The process of deciding the *what* and *how* is the background effort that goes into creating the finished product – the report. While this decision-making process represents a large share of the effort required to make the report, the decision-making process is not apparent from the finished product. Rather, this decision-making process lies below the surface of the finished report. Although the report can be shared with others, only the executive possesses the knowledge of the *what* and *how*. Anyone wishing to create a similar report in the future will have to expend the same effort as the executive had in deciding the *what* and the *how* because the *what* and *how* remain the executive's unexpressed effort and knowledge.

Therefore, a method for capturing concepts, or the *what* and the *how*, is needed to prevent losing the knowledge behind reports. One approach to solving this problem is to develop semantic ontologies. A semantic ontology is a system in which concepts are described using a standard set of attributes, or a standard lexicon. A concept is described in terms of the semantic ontology by annotating the concept with attributes from the standard lexicon that are related by their meanings to the concept. The attributes not only describe the foreground information

-2-

associated with the concept, but also the background information including the knowledge of *what* information is to be included in the foreground information and *how* that information is obtained. Since the attributes describe the *what* and *how* of the concept, the background information is captured in the attributes and not lost.

5 For example, a specification sheet for an automobile might contain the following information: "TOP SPEED = 120 mph." Another person may be interested in creating a similar specification sheet for another automobile. If the person is not familiar with the concept of "speed" the person may not understand how to determine the vehicle's "top speed." What is missing is the additional information that explains *what* information is necessary to calculate top
10 speed and *how* to obtain that information. This missing information could be communicated through a semantic ontology. The semantic ontology in this example might include the attributes "distance," "time" and "measure." According to the ontology, speed could be expressed in terms of the attributes by saying, "speed is the measurement of the distance a vehicle travels in one hour." Therefore, the ontology communicates what information is necessary to calculate speed,
15 and how to obtain that information.

 Although semantic ontologies theoretically provide a means for communicating the *what* and *how* in documents, there have heretofore been several obstacles to their practical implementation. Ontologically annotating a document is a tedious, time consuming process. A document only can be annotated by the individual creating the document because only that
20 individual possess the knowledge of the *what* and *how*. To annotate a document, the individual first must determine whether the individual has access to any pre-existing suitable semantic ontology. Most individuals would not have access to any such ontology. Therefore, the individual must create a semantic ontology that is appropriate for the subject matter of the document. After creating the ontology, the individual must identify relevant data within the
25 document. Finally, the individual must annotate each datum with attributes from the ontology. These steps amount to a substantial effort on the part of the author. In most cases, individuals do not have the extra time required to annotate a document after it already has been created, much less create a suitable semantic ontology. In addition, the creator has little incentive to annotate the document because the annotation is primarily for the benefit of others – the creator already
30 knows the *what* and the *how*. Therefore, ontological processing is a tedious task with little reward for the creator of the document.

-3-

Another obstacle to implementing semantic ontologies is that there are no applications that allow authors to simultaneously create and ontologically annotate documents. Because the annotation is not performed in conjunction with the creation of the document, there is less incentive to annotate the document and a greater risk that the user will never annotate the document.

Therefore, it would be desirable to have a system to overcome the obstacles associated with implementing semantic ontologies to process information. It would be a significant advantage if such system automated the process of ontologically annotating documents.

It also would be desirable to have such a system that could be integrated with existing applications already used by authors to create documents. Such a system would capitalize on existing training, making it more likely that people would use the system successfully. Such a system would reduce or eliminate the loss of knowledge, increase the ability for people to share knowledge and reduce the time for people to create new documents of the same type as previously annotated documents. It would be another significant advantage if the system were developed to facilitate creating documents using a database of captured concepts.

Summary of the Invention

Semantic ontology provides a system in which a user organizes information into concepts. Those concepts are then described using a standard lexicon of attributes describing the concepts. The attributes describe the concepts in terms of the knowledge required to create the concept, including *what* information is necessary to express the concept and *how* that information is to be attained. In this manner, semantic ontology preserves the knowledge that is associated with a particular concept. The present invention provides a system for implementing semantic ontology in a way that avoids the obstacles previously associated with semantic ontology. The invention has several advantages over the prior art.

First, the present invention provides a software tool that automates the process of semantic ontology. Automating the process significantly decreases the time required to process a document according to a semantic ontology, which increases the likelihood that creators will be willing to annotate their documents.

-4-

Second, the present invention provides a software tool that is integrated into software applications that already are used to create documents. Therefore, document creators can simultaneously create their documents and process them according to the semantic ontology by cutting and pasting within a single application.

5 Third, the present invention provides a software tool that can assist document creators with creating new documents. The system can recommend to users what information to include in a document and can automatically attain and insert the information into the document. Alternatively, the system can describe how to attain the information. Document creators are more likely to contribute to the ontology by processing their documents according to the
10 ontology because the system provides a benefit in that it assists document creators in creating new documents.

These advantages are provided by an apparatus according to the present invention that has a memory for storing at least one ontology, a display for displaying the electronic document, means for selecting at least one datum within the document, means for annotating the
15 datum with at least one concept contained in the ontology, and a memory for storing the annotated datum. The stored ontology includes at least one concept or at least one process, and preferably includes one or more concepts and processes. Each concept and process can include a plurality of attributes.

An apparatus in accordance with the present invention also can create at least a
20 portion of an electronic document. The apparatus can include a memory for storing at least one ontology including at least one concept, means for selecting a concept from the stored ontology, means for copying the selected concept into the document, means for inputting information relating to the selected concept into the document, and a display for displaying the document. The memory preferably stores a plurality of ontologies. The apparatus can further include a
5 means for selecting an ontology from the plurality of ontologies. The ontology also can have at least one process for obtaining information and a means for selecting a process from the ontology. The apparatus can have a means for obtaining information as requested by a selected process. The stored concept can include a plurality of concept attributes and the stored processes can include a plurality of process attributes.

7 The present invention also involves a method for annotating an electronic document. The method includes the steps of providing an ontology having at least one concept, providing a body of information in electronic form, identifying at least one datum within the

-5-

body of information, matching the datum with at least one concept contained within the ontology, annotating the datum with the at least one concept, and storing in a memory the annotated datum. The system can store in an electronic form the datum, and the concept and the connection between them. The method also can include the steps of identifying at least one
5 attribute from the standard lexicon that describes the datum, annotating the datum with the attributes selected from the datum, storing in an electronic storage mechanism the datum annotated with at least one concept and at least one identified attribute. The ontology used has at least one concept and a standard lexicon of attributes to describe the concept. The system preferably uses multiple ontologies, each ontology having at least one concept and a standard
10 lexicon of attributes to describe the concept. The annotating step can include annotating the datum using a mark up language, and can further include the step of selecting the concept from a menu of terms electronically displayed, and storing the concept in the form of a mark up language. This method can be repeated for additional data in the document and storing the annotated data in the electronic storage mechanism.

15 The system also can create at least a portion of a document by selecting an ontology, selecting a concept from a menu of concepts belonging to the selected ontology, copying the selected concept to the document, and inputting information relating to the concept into the document. The selected concept can include attributes related to the concept. The user also can select a process from a menu of processes that can include attributes related to the
20 processes. The selected processes can be used to obtain information relating to the selected concept and automatically input the information relating to the concept into the document.

Brief Description of the Drawings

FIG. 1 is a schematic diagram of a system in accordance with the present invention;

FIG. 2 is a schematic diagram of a screen display from a system using the present invention;

FIG. 3 is a schematic diagram of a second screen display for a system using the present invention; and

FIG. 4 is a schematic diagram of a third screen display for a system using the present invention.

Detailed Description of the Preferred Embodiment

Referring to the figures, an apparatus for processing information according to the present invention is shown in FIG. 1. The system includes a software tool 5 and an ontology server 20. Software tool 5 is an application that can be launched as an applet within existing software applications 10. Existing software applications 10 can include applications for creating documents. For example, the system can be used in conjunction with Microsoft™ Word™, Excel™, Access™ and PowerPoint™ software.

Software tool 5 can be utilized to serve either of two objectives. First, software tool 5 can be launched within an existing document created in one of existing software applications 10 to process existing data contained in a document. Each datum can be, for example, any length of text, numerical information, a picture, a chart or any other type of recordable intelligence. Second, software tool 5 can be launched within a blank or existing document created in one of the existing software applications 10 to generate new information to be inputted into the blank or existing document. As shown in FIG. 2, software tool 5 can be launched, e.g., by clicking on an icon 14 on a tool bar of software application 10 displayed to the user, e.g., on a CRT screen display. Clicking on the icon can allow the user to “pull down” a menu 8.

Menu 8 presents a series of options to the user. “Search Ontologies” option 6 permits a user of the system to search existing ontologies. Existing ontologies preferably are stored on ontology server 20. When option 6 is selected, existing ontologies are recalled from ontology server 20. Recalled existing ontologies are displayed to the user.

-7-

Each existing ontology recalled from ontology server 20 includes a listing of concepts 25, and preferably processes 30, with which the ontology is associated. Each concept 25 or process 30 included in an existing ontology includes a list of attributes 35 that describe each concept 25 or process 30. Attributes 35 that describe concept 25 describe what type of information is relevant to concept 25, and why that particular information is relevant to concept 25. Attributes 35 that describe processes 30 describe where information that is relevant to process 30 is to be found and how such information can be attained.

Software tool 5 may be launched to process existing data in a document. Existing data may be selected in a document using a keyboard, mouse or similar pointing device.

“Search Ontologies” option 6 is used to annotate existing data within the document, preferably using an existing ontology and its associated concepts 25 and processes 30. After an existing ontology is found that matches information already contained in the document, then that existing ontology may be selected using a computer keyboard, mouse or similar pointing device. After the existing ontology is selected, a listing of concepts 25 and processes 30 that are related to the existing ontology are displayed on graphical window 18, as shown in FIG. 3. Concepts 25 and processes 30 can be found that match data in the existing document. A concept 25 or process 30 can be selected by clicking on the concept or process in graphical window 18, or using a mouse or similar computer pointing device to drag concept 25 or process 30 onto the matching data to be processed within the existing document. This annotates the data with the concept or process information. When the user is finished annotating all the desired data in the document, the user can save the document to an electronic memory. The saving step saves each annotated datum, i.e., the datum, the related concept or process and the fact of the connection between the datum and the concept or process.

“Edit An Ontology” option 7 allows a user to edit an existing ontology if desired by the user because, e.g., an existing ontology cannot be found to match precisely the information already contained in the document. “Edit An Ontology” option 7 allows the user to add or delete concepts 25 or processes 30 to an existing ontology. “Edit An Ontology” option 7 also allows the user to add or delete attributes 35 to concepts 25 or processes 30. The edited ontology is then stored in ontology server 20, and becomes part of the existing ontologies.

“Create New Ontology” option 9 allows a user to create an entirely new ontology. A new ontology is created by defining the concepts and processes, and the attributes of the concepts and

processes for the ontology. The new ontology is then stored in ontology server 20, and becomes an existing ontology.

If software tool 5 is launched to generate new information to include in a document, the user can use "Search Ontologies" option 6 to search for existing ontologies that would match the subject of the document. When an existing ontology is found to match the subject of the document to be created, selecting that existing ontology will allow the user to view concepts 25 and processes 30 to include in the document. "Document Recommend" option 11 allows the user to create a document by selecting an existing ontology and one or more concepts 25 or processes 30. Software tool 5 would then compare the selected concepts 25 and processes 30 with the concepts 25 and processes 30 in other documents created by other users in that ontology context and recommend other concepts 25 and processes 30 to accompany the already selected concepts 25 and processes 30 in the document being created.

After the user has selected an existing ontology and related concepts 25 and processes 20 to include in the document, software tool 5 can use the attributes 35 associated with any included processes 30 to obtain data to be included in the document according to concepts 25. Attributes 35 associated with processes 20 provide instructions to the user's computer to search specific databases for the required data. The user's computer could then import that data into the document. For example, if the concept 25 was an automobile's top speed, a process 35 could instruct the user's computer to access a database that contains the vehicle specifications including the vehicle's top speed. The user's computer would then import the data into the document where top speed is addressed.

After the data in the document is annotated with selected concepts 25 and processes 30, the user can use software tool 5 to create relationships between concepts 25 and processes 30. Relationships are created by selecting two related concepts 25, two related processes 30 or a related concept 25 and a related process 30 by clicking on each of them in a graphical window. Software tool 5 then launches relationship display window 50 shown in FIG. 4. Relationship display window 50 lists operators 55 that describe possible relationships between the selected concepts 25 or processes 30. For example, in FIG. 4, the user has selected two concepts between which to create a relationship. The user can select an operator 55 that describes the appropriate relationship between the two selected concepts. After the relationship is selected, software tool 5 translates the selected concepts 25 and processes 30 and

-9-

corresponding operator 55 into a markup language, such as DAML. A markup language is capable of being processed and stored in binary format in a computer memory.

Therefore, the present system overcomes many of the obstacles that previously had existed and prevented practical implementation of semantic ontologies to process
5 information. The system allows for the automated annotation of documents using electronically stored ontologies. The system also allows users to create and edit ontologies. The system can make recommendations to an author regarding the types of information that should be included in a given type of document. Finally, the system can automatically generate documents based on the stored ontologies, by knowing the *what* and the *how* that go into making that type of
10 document. The system is easy to use because it can be integrated with existing software applications with which users already are familiar.

Whereas the present invention has been described with respect to specific embodiments thereof, it will be understood that various changes and modifications will be suggested to one skilled in the art and it is intended that the invention encompass such changes
15 and modifications as fall within the scope of the appended claims.

-10-

What is claimed is:

1. An apparatus for annotating an electronic document, comprising:
a storage mechanism for storing at least one ontology, said ontology including at least one concept;
a display for displaying the electronic document;
5 means for selecting at least one datum within the document;
means for annotating the datum with at least one concept contained in the ontology; and
a memory for storing the annotated datum.
2. The apparatus of claim 1 wherein said stored ontology further includes at least one process for obtaining information.
3. The apparatus of claim 2 wherein said stored concept includes a plurality of concept attributes and said stored process includes a plurality of process attributes.
4. An apparatus for creating at least a portion of a document, comprising:
a memory for storing at least one ontology, said ontology including at least one concept;
5 means for selecting a concept from said stored ontology;
means for copying said selected concept into the document;
means for inputting information relating to said selected concept into the document; and
a display for displaying the document.
5. The apparatus of claim 4 wherein said memory stores a plurality of ontologies and further comprising means for selecting an ontology from said plurality of ontologies.
6. The apparatus of claim 4 wherein said ontology further comprises at least one process for obtaining information and further comprising means for selecting a process from said ontology.

-11-

7. The apparatus of claim 6 further comprising means for obtaining information as requested by a selected process.

8. The apparatus of claim 6 wherein said stored concept includes a plurality of concept attributes and said stored process includes a plurality of process attributes.

9. A method for annotating an electronic document, comprising the steps of:

- (a) providing an ontology having at least one concept;
- (b) providing a body of information in electronic form;
- (c) identifying at least one datum within the body of information;
- 5 (d) matching the datum with at least one concept contained within the ontology;
- (e) annotating the datum with the at least one concept; and
- (f) storing in a memory the annotated datum.

10. The method of claim 9 wherein said annotated datum storing step comprises storing in an electronic form the datum, the at least one concept and the connection between them.

11. The method of claim 10 further comprising the steps of:

identifying at least one attribute from the standard lexicon that describes the datum;

annotating the datum with the attributes selected for the datum; and

5 storing in an electronic storage mechanism the datum, annotated with the at least one concept and the at least one identified attribute.

12. The method of claim 9 wherein said ontology providing step comprises providing an ontology having at least one concept and a standard lexicon of attributes to describe the at least one concept;

-12-

13. The method of claim 9 wherein said ontology providing step further comprises the step of providing a plurality of ontologies, each ontology having at least one concept and a standard lexicon of attributes to describe the at least one concept.

14. The method of claim 9 wherein said annotating step comprises annotating the datum using a mark up language.

15. The method of claim 14 wherein said annotating step further comprises selecting the at least one concept from a menu of terms electronically displayed to a user.

16. The method of claim 14 wherein said annotated datum storing step further comprises storing the at least one concept in the form of mark up language.

17. The method of claim 9 further comprising repeating steps (a)-(e) for additional data within the body of information and storing in the memory the annotated data.

18. A method for creating at least a portion of a document, comprising the steps of:

selecting an ontology;

selecting a concept from a menu of concepts belonging to the selected ontology;

5 copying the selected concept to the document; and

inputting information relating to the concept into the document.

19. The method of claim 18 wherein the selected concept includes attributes related to the concept.

20. The method of claim 18 further comprising the step of selecting a process from a menu of processes.

21. The method of claim 20 wherein the selected process includes attributes related to the process.

-13-

22. The method of claim 20 further comprising the step of using the selected process to obtain information relating to the selected concept and automatically inputting the information relating to the concept into the document.

FIGURE 1

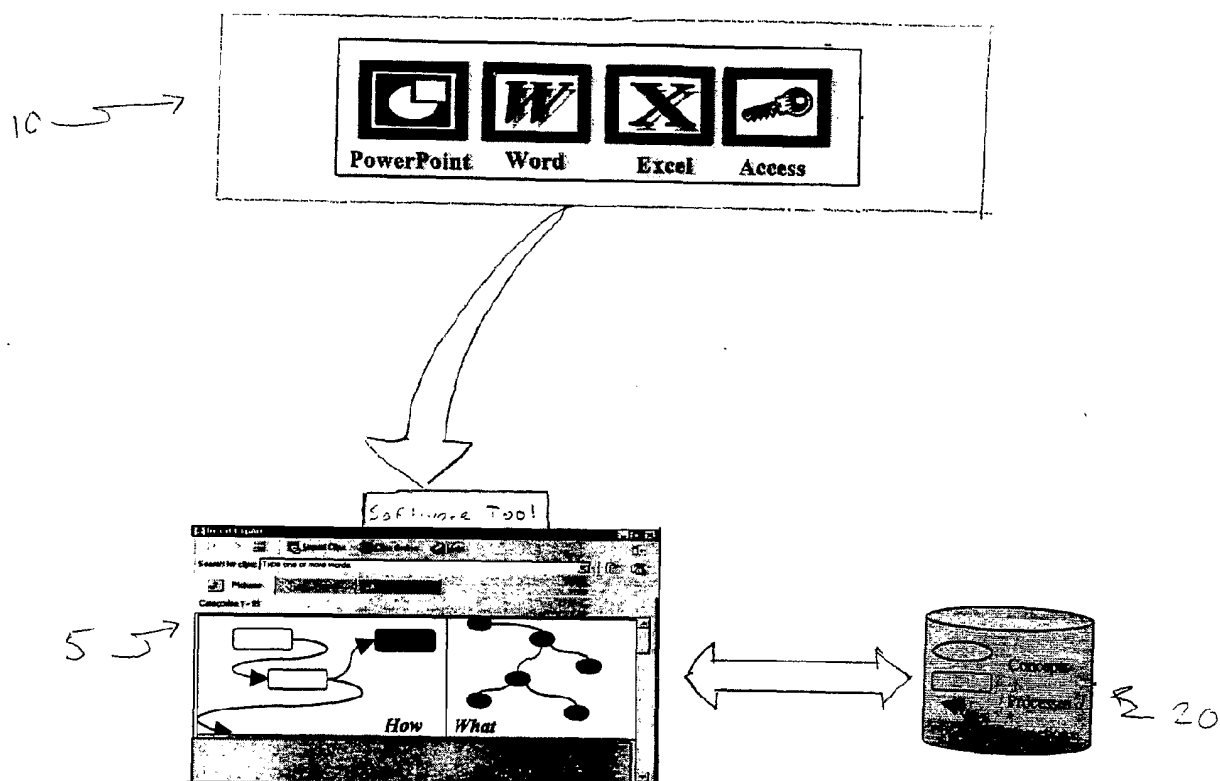


FIGURE 2

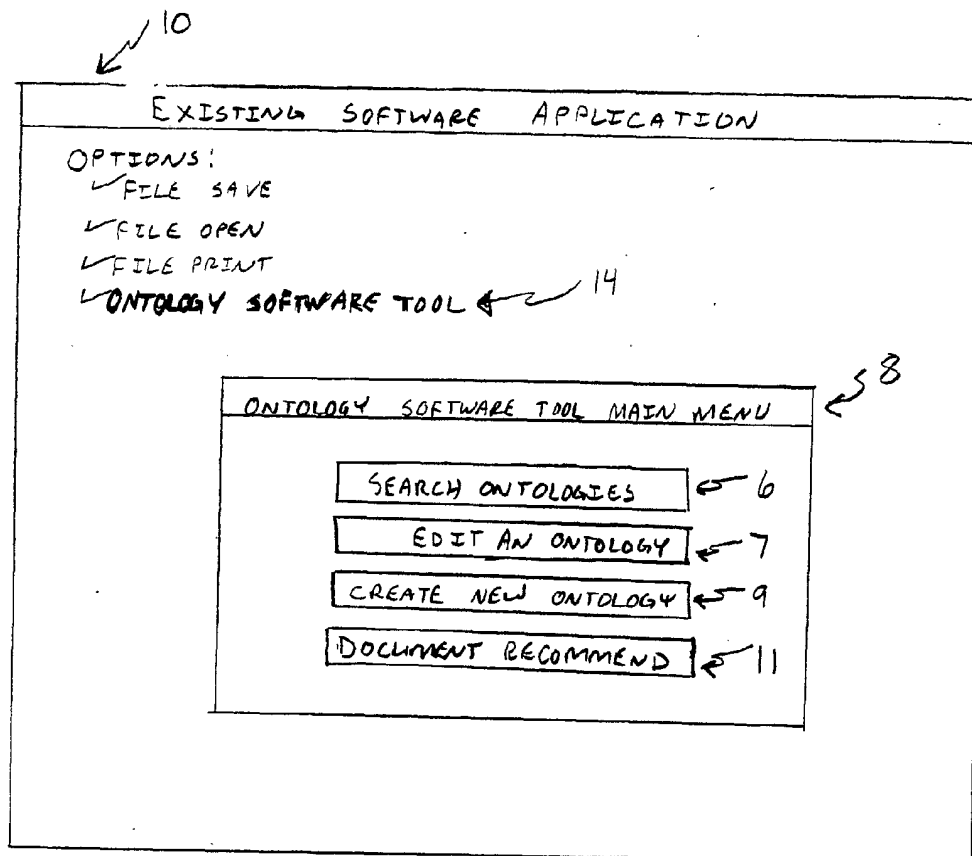


FIGURE 3

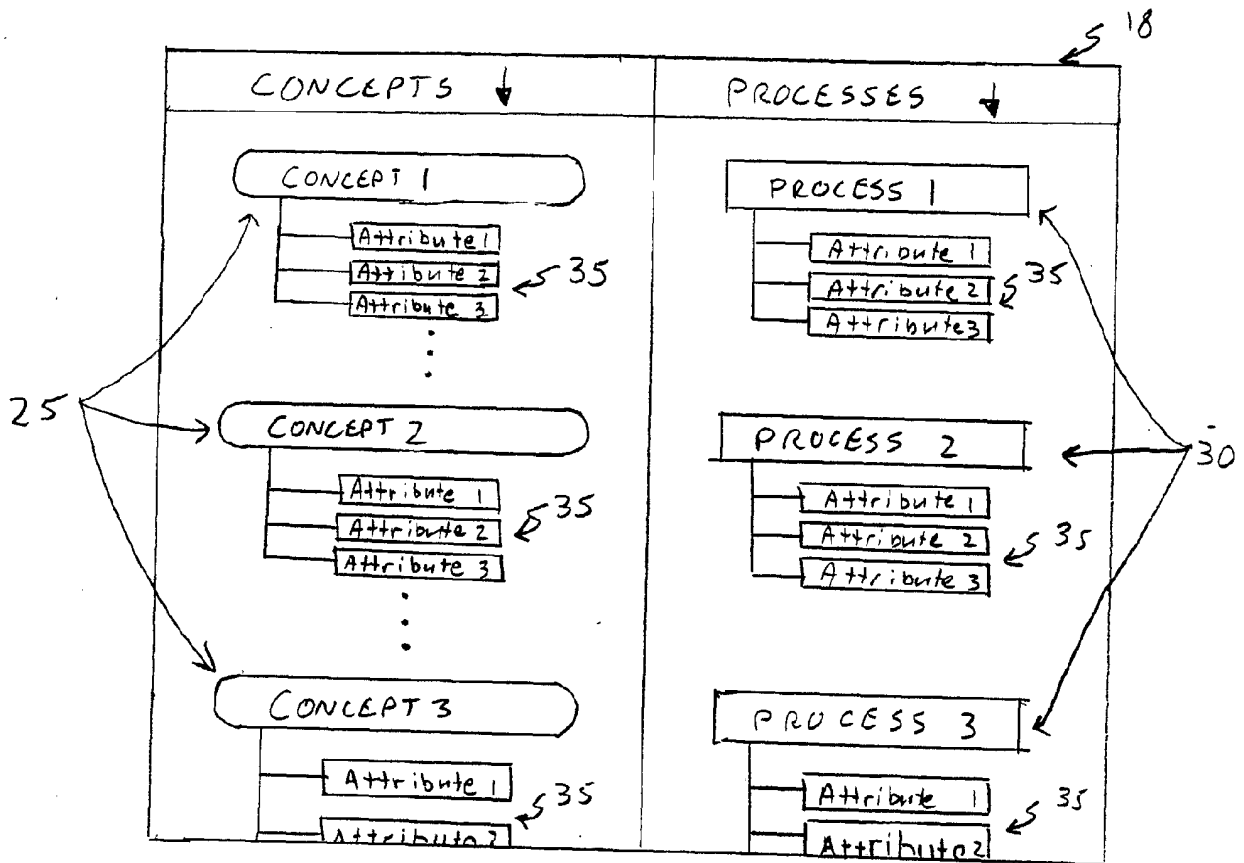


FIGURE 4