View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by espace@Curtin

© 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Evidence/Discovery-Based Evolving Ontology (EDBEO)

Ehsan Nasiri Koozani, Omar Hussain, Tharam Dillon, Maja Hadzic
Digital Ecosystems and Business Intelligence Institute (DEBII)
ehsan.nasirikhoozani@postgrad.curtin.edu.au, O.Hussain@cbs.curtin.edu.au, Tharam.Dillon@cbs.curtin.edu.au, Maja.Hadzic@curtin.edu.au

Abstract

This paper presents a proposal for the development of an ontology evolution strategy which refines ontological relations in scientific ontologies.

In addition to experts' consensus, it is desirable to define ontological relations between any two concepts in a scientific ontology based on scientific evidence. To address this issue, we can relate ontological relations to different research results obtained from various studies. To implement this solution, our envisaged evidence/discovery-based methodology integrates a higher-level ontology (systematic review ontology) into a systematic review agent which employs a Fuzzy Inference System in order to automatically modify ontological relations of a domain ontology based on the evidence received from information resources. The evidence/discovery-based methodology will further use the domain ontology to discover novel connections between distinct literatures, thereby, enrich its conceptualization.

1. The significance of having evidence-based ontological relationships and the importance of undiscovered literature connections

Ontology is a branch of knowledge engineering the aim of which is to supply shared, formal, and explicit descriptions, definitions, classification, and organization for *concepts* and *their relationships* in a domain of knowledge [1]. Such combinations of concepts and their relationships constitute the foundation of theories in every domain.

Concepts in every domain represent the entities that experts in that domain deal with; and relationships link those entities together in a meaningful way. Although various domains encompass different types of concepts with a wide range of features and characteristics, the kinds of links (i.e., relationships) between those concepts have extensive commonalities across different domains. For example, the taxonomy (is-a) or meronomy (part-of) relationships are common in almost all ontologies. By and large, we can classify relationship types of ontologies into two categories: 1. *structural relationships*, and 2. *operational relationships*.

Structural relationships, which are of a rather static nature, incorporate those relationships such as "is-a" or "part-of" which represent constant and necessary structural links between the entities of a domain. Most ontologies incorporate this type of relationship between finite entities of a given domain. In contrast, operational relationships tap into the dynamic interactions between objects and the influence that different entities have on each other be it in the form of causation, change, regulation, or various types of mediation.

The structural type of relationship is more concerned with the classification or categorization of entities. Examples of such a relationship type can be found in chemistry where elements are classified according to certain categories. The operational relationship type, however, represents *scientific laws* by linking different entities in such a way that the discovered influences of one entity on another can be identified.

As the nature of entities and their corresponding concepts become more abstract and elusive, the complexity and ambiguity of their relationships increase. Therefore, it is easier to define a causal relationship between, for example, two physical entities such as *force* and *pressure* than between two psychological concepts such as *stress* and *emotional memory*.

The complexity of the establishment of an operational (and even structural) relationship between two such entities is due to the intuitive, ambiguous, and multi-dimensional nature of psychological concepts, as well as the existence of different definitions, worldviews, and methodological approaches [2] for their examination.

978-1-4244-9166-7/10/$26.00 ©2010 IEEE

Although it has been emphasized that ontology is a consensual knowledge framework [3], in the realm of science, however, it is desirable that ontology concepts, in addition to being consensual, also correspond to reality, and ontology relations represent real, physical relations in nature. According to this view, high-quality ontologies are illustrations of reality and incorporate universals that exist in the real world of space and time [4]. The incorporation of such concepts and their relationships in an ontological framework, however, needs preparation and inclusion of correct and evidence-based facts extracted from research literatures.

Our experience in the conceptualization of *Human Stress Ontology* (HSO) [5] brought to our attention the complexity and difficulty of defining relationships, particularly the operational ones, between various concepts in the human stress domain. Overall, different ontologies describe their relationships through the consensus of domain experts (e.g. in case of business-related ontologies) and/or by reference to the scientific facts (e.g. in case of biomedical ontologies). In the field of business, for example, experts can come to agreement that within the ontology X, "*manager*" and "*company*" are linked together via the relation "*manages*" (i.e. *manager manages company*). However, the defining of an operational relationship (e.g. *regulates*) in biomedical ontologies such as Gene Ontology (GO) [6] is supposed to be based on already existing scientific facts. For example, a statement such as "*cell cycle checkpoint regulates cell cycle*" in GO implies the constancy and necessity of the regulation of the cell cycle by the cell cycle checkpoint (i.e. *checkpoint necessarily regulates cell cycle*) [6]. Here, mere reference to consensus cannot bear much meaning as scientific laws have not been established through consensus.

Of greater concern is the defining of operational relationships between two stress-related concepts since, compared to Genetics, there are less certain or established facts or scientific laws in domains such as human stress. For example, a statement such as "*stress response reduces Gonadotropin secretion*" when represented in the HSO cannot be considered as a certain or necessary relationship.

Another problem regarding the expression of ontological relationships relates to the alteration or evolution of such relationships as a result of the changes occurring in a domain of interest. Such alterations in business-related ontologies might arise from changes in the goals and needs of stakeholders or the occurrence of conflicts in their perspectives. Different ontology refinement and evolution strategies, which are mainly based on consensus between ontology users, have been designed to resolve such conflicting situations in business fields [7]. Such methods are more concerned with the changes which might occur in the definition of various concepts and entities across different enterprises. For example, [8] in their attempt to address the issue of ontology integration, re-verification, and maintenance have introduced a formal articulation algorithm the aim of which is to discover various links between different evolving ontologies. In a previous attempt to resolve the context-induced inconsistencies and contradictions which might take place in a knowledge base, CYC [9] proposes the notion of microtheories. The multitude of common sense assertions (microtheories) in CYC allow the knowledge base or system builder to make other assertions about every context-dependent assumption, resolving their possible contradictory implications in different contexts. It is suggested, however, that this assertion-making strategy be implemented using argumentation reasoning.

Despite such efforts to solve the inconsistency and context-dependency issues of ontologies, to our knowledge, there is no ontology evolution methodology for addressing the ongoing change of relationships between concepts in scientific ontologies such as biomedical ontologies in a systematic evidence-based manner. In order to design acceptable, reliable, and effective ontologies for scientific domains, we need to define evidence-based (not merely consensual) relationships between concepts with the capability to change and evolve in response to new incoming research results and contributions.

The issue of undiscovered connections in the scientific literature [10] is another important issue which, we propose, can be integrated with the topic of ontology relationships. Undiscovered connections are those implicit and plausible linkages which might exist between disjoint fragments of various research works. Such linkages are being studied systematically by methods developed mainly in the field of Literature Based Discovery (LBD) [10]. Integration of LBD into the ontology evolution process, particularly in scientific ontologies, can enhance the number of ontology concepts, discover novel relationships between ontology concepts, and improve the comprehensiveness of the ontology framework. In response to these needs, we envisage the establishment of an ontology evolution methodology, *Evidence/Discovery-Based Evolving Ontology* (EDBEO), on the basis of which researchers can define evidence-based relationships as well as discover new relationships between different concepts in an ontology. To deal with such issues, we incorporate methods of systematic review as effective strategies to find provable evidences for ontology relationships, and

LBD techniques to discover new relationships and connections between ontology concepts.

2. Systematic Review

The importance of evidence-based theories in different areas of science, particularly medicine, has been highlighted in methodological studies of systematic review and meta-analysis. A systematic review can be defined as a succinct summary of the best existing evidence for a certain research issue such as a clinical question. It utilizes explicit methods to identify, combine, and examine high quality evidences from relevant studies in order to determine "*the whole truth*" [11] and augment, in a scientific way, the validity of assertions made in a domain. As science is a collaborative and cumulative process, dispersed parts of this process can be combined through the systematic review [12]. Furthermore, a systematic review considers contradictory results of isolated studies of a specific topic as positive and negative particulars of a probabilistic distribution of results, rather than contradictions [13].

In a simple synthesis systematic review, the researcher uses an *election mode* or *voting method* to investigate each study individually and counts the results or votes with respect to the target question. For example, the researcher may learn that among 50 elected studies, 30 studies demonstrate positive results, 15 display negative results and the rest of 5 point to no significant results [12].

3. Literature Based Discovery

Literature Based Discovery (LBD) is the process of searching for significant, complementary, implicit, and hidden relationships among information contained in disjoint published literatures in order to discover new knowledge [10]. LBD was introduced as a strategy to deal with the fragmentation and overspecialization of a huge array of information embedded in a broad range of information resources. This fragmentation has caused scientists to communicate within the restricted space of their own fragment without having much communication with other related domain communities [10]. As a result, researchers may fail to stay aware of important results published in other related fields which may have significant implications for their own work. As current literature searching strategies through the use of regular search engines are not capable of identifying such important links between discoveries of various research works, LBD was suggested as a practical strategy to discover implicit, novel, and unpublished connections between concepts of dispersed literatures [14]. A typical example of a LBD process is as follows:

The results of one experiment published in a paper indicate that A affects B. Another research report which may be published in a different journal states that B affects C. As the reader can deduce, a logical inference extracted from these two statements can assert that A affects C; however, the two mentioned research works may be unaware of this connection until we manage to link their results together and discover such a connection. The discovered connection can later be tested to see whether in reality there exists such a relation between A and C.

One basic problem in the process of LBD is that, in actual fact, there might exist an infinite number of such possible relations between numerous concepts across various literatures. Given that only a finite number of such connections can be plausible, novel, and testable, several LBD methods and tools have been created to limit the number of these intermediate concepts (say B concepts) and draw relevant and reasonable hypotheses [14].

For example, [15] utilizes metadata section of the Medline database, i.e. Medical Subject Headings (MeSH) [16] (instead of extraction from title, abstract, or text bodies) to pull out and rank important concepts in the process of finding intermediate concepts.

4. A methodology for evidence/discovery-based evolving ontology (EDBEO)

The proposed EDBEO methodology is an ontology evolution strategy which aims to address two relationship issues in a scientific ontology: 1) the evidentiality of existing relationships between represented concepts in an ontology; and 2) the discovery of new relationships and concepts for the enrichment of that ontology. To address these issues, EDBEO methodology will incorporate two major components: 1) an *Automated Systematic Review Agent* (ASRA); and 2) an ontology-based literature discovery framework.

4.1. Automated Systematic Review Agent

The first stage in every systematic review is to explicitly and precisely formulate the core research question and define its elements. It has been emphasised that the research question in a systematic review is a multidimensional *conceptual structure* to which other phases of the systematic review procedure must conform [12]. In order to formulate a consistent and consensual conceptual structure and, consequently, gain germane and comparable results, [12] suggest that having a shared formalized vocabulary of relevant concepts in the form of a *scientific research ontology* can be effective and helpful. With respect to the formulation of the core research question in a systematic review process, the scientific research ontology can serve to provide terminological homogeneity for concepts used by different investigators and, in turn, make the retrieved information consistent with the consequent results. Moreover, it can improve information extraction tools by facilitating the identification, extraction, and association of related terms in a scientific text [12].

In contrast to [12] which used ontology as a facilitating tool for the conduction of systematic review, our proposed *Automated Systematic Review Agent* (ASRA) implements systematic review as a basis for the refinement of ontology relationships.

ASRA uses two distinct ontologies in order to establish evidence-based relationships between ontology concepts. The first ontology is the *Domain Ontology* (DO) for which concepts we aim to establish evolving evidential relationships. The second ontology is a higher-level ontology, called the *Systematic Review Ontology* (SRO), which can be similar to the abovementioned scientific research ontology designed by [12].

Based on the DO concepts, knowledge from scientific research ontologies (and/or libraries), and a survey of article result and conclusion sections, the researcher first builds the SRO. This higher-level ontology encompasses those terms (e.g. verbs) appearing in scientific literatures which indicate the relationship between two specific concepts with respect to the research results. A survey of scientific article conclusions or abstracts reveals that most of them contain statements about the relation between two concepts for which the research has obtained some results. For example, the conclusion section of study *X* states that "*stress response reduces Gonadotropin secretion*". The SRO repository contains a rich terminology of specific nouns and phrases such as "*stress response"*, and "*Gonadotropin secretion"* as well as specific verbs such as "*reduce"*. The specific nouns, in fact, represent those professional concepts between which a scientific theory (contained in relevant articles) aims to prove a connection. The SRO also incorporates various synonyms and different spellings of these key concepts in its repository. For example, for the verb *"reduce"* there will come all its synonyms e.g. *"decrease", "diminish",* or *"lessen"* which are usually used by researchers to state the result of their studies. Establishment and specification of such synonyms can be based on linguistic guidelines and critical thinking. Using annotation tools, the ASRA uses the SRO to identify, retrieve, and import such key terminologies to a separate database. Such specialized terminologies in the SRO can also be used for the selection of relevant articles as not every article contains recognizable statements about the proved or disproved connections between two concepts in the form of a scientific theory.

In the next stage, the ASRA implements the retrieved theories (statements such as *X influences Y*) of article conclusions to refine the ontology relationships of the domain ontology (DO). ASRA implements the *voting method* applied in a simple synthesis systematic review. The ASRA repository, which combines all theoretical statements indicating negative or positive connections between two concepts, then undergoes a statistical analysis to calculate the percentage of statements which have reported negative or positive connections between two given concepts.

The link between the evidence gained from the SRO and the change of relation in the DO can be established in an automated way such that the more evidence (higher percentage of positive correlations) is received from the SRO, the more a relevant ontology relation in the DO will change toward a positive indication. For example, the more supporting evidence is received about the existence of a positive relation (e.g. *increase*) between two concepts (e.g. stress response and Gonadotropin secretion), the stronger (more proving or certain) the ontology relationship between those concepts becomes in the DO. In this way, for example, the existing "*is likely to increase*" relationship between two DO concepts (e.g. *stress response is likely to increase Gonadotropin secretion*) can be turned to "*is highly likely to increase*" relationship (i.e. *stress response is highly likely to increase Gonadotropin secretion*) when a significant percentage of proving evidence is received from the SRO.

An effective way to modulate such ontology relationships is to consider them as fuzzy variables which can take various linguistic values. For example, "*increases*" can take different values of "*is likely to*", "*is highly likely to*", "*is less likely to*", "*does not*", etc according to the percentage of supporting evidence being stored and accumulated in the SRO. In this manner, the ASRA might keep giving weights to the relationship "*is less likely to increase*" (e.g. "*is likely to increase*",) until it gets to the point where, with sufficient confidence, it can be asserted that a "*is highly likely to increase*" relationship exists between the two concepts. In contrast, the incoming disproving research results can continuously decrease the value of "*is less likely to increase*" until it gets to the "*does not increase*" point.

To achieve this goal in an automatic and consistent manner, we implement *Fuzzy Inference System* [17] based on *Fuzzy Logic* [18]. Fuzzy logic enables one to reason with natural language variables which lack finite and distinguishing boundaries in the experience of everyday life. It provides an effective method for dealing with infinite values such as "*very*" or "*almost*" to which computation the conventional two-valued (0 and 1) Boolean logic cannot be applied.

Fuzzy inference systems are composed of different stages of fuzzification of input data, application of fuzzy rules, and defuzzification of output results. In the fuzzification stage, crisp values are transformed into grades of membership. Fuzzy rules are conditional statements which appear in the form of "*IF x is A: THEN y is B*". The first part of the fuzzy rule (i.e. *IF x is A*) is called the *antecedent*. Respectively the second part of the fuzzy rule (i.e. *THEN y is B*) is denoted as the *consequent*. At the defuzzification stage, the fuzzy outcomes of each variable will be turned into a single number which is quantifiable for decision making purposes. We explain this process using our previous example (*stress response increases Gonadotropin secretion*).

The antecedent (input) part of the fuzzy rules for each ontology statement is composed of two parts:

1. *Type of Research Results* (TRR) of the SRO which might be in the form of verifying, disproving, or neutral statements such as "*stress response increased Gonadotropin secretion*", "*stress response decreased Gonadotropin secretion*" or "*stress response had no significant association with Gonadotropin secretion*". The system reduces all such statements to three general categories of *positive conclusion* (*+ve*), *neutral conclusion* (*NEU*), and *negative conclusion* (*–ve*). A positive conclusion represents the likelihood of an affirmative relationship being present; whereas, a negative conclusion represents the likelihood of a disproving relationship being present between the two concepts. Neutral conclusion is used to represent the likelihood of no significant relationship existing between the two concepts;

2. The second part of the input gives an account of the level of severity of the TRR (Type of Research Results) which shows the *level of proof* (LOP) obtained from the considered articles. The LOP is represented by a trapezoidal function that has three fuzzy sets: Low (L), Medium (M) and High (H). The membership function of each fuzzy set is defined as:

$\mu_{\text{Low}}(\text{LOP}) = 1$ (if $0 < x < 20$); $\frac{30-x}{10}$ (if $21 < x < 30$); 0 (if $31 < x < 100$)

$\mu_{\text{Medium}}(\text{LOP}) = 0$ (if $0 < x < 20$); $\frac{x-20}{10}$ (if $21 < x < 30$); 1 (if $30 < x < 55$); $\frac{70-x}{15}$ (if $56 < x < 70$); 0 (if $71 < x < 100$)

$\mu_{\text{High}}(\text{LOP}) = 0$ (if $0 < x < 55$); $\frac{x-55}{15}$ (if $56 < x < 70$); 1 (if $70 < x < 100$)

The output (consequent) of the fuzzy rules is one of the four linguistic values of: *does not*, *is less likely to*, *is likely to*, and *is highly likely to*. Depending on the values of the TRR and LOP of the input, the ontology relation variable (e.g. *increases*) takes one of these output linguistic values.

The fuzzy rules are in the following form:

IF *TRR* = *+ve* and *LOP*= *L* THEN Output = *is likely to*
IF *TRR* = *+ve* and *LOP*= *M* THEN Output = *is likely to*
IF *TRR*= *+ve* and *LOP*= *H* THEN Output = *is highly likely to*
IF *TRR* = *NEU* and *LOP*= *L* THEN Output = *is less likely to*
If *TRR* = *NEU* and *LOP*= *M* THEN Output = *is less likely to*
IF *TRR* = *NEU* and *LOP*= *H* THEN Output = *is less likely to*
IF *TRR* = *-ve* and *LOP*= *L* THEN Output = *is likely to*
IF *TRR* = *-ve* and *LOP*= *M* THEN Output = *is less likely to*
IF *TRR* = *-ve* and *LOP*= *H* THEN Output = *does not*

At the defuzzification stage, the output of fuzzy rules will be achieved by the process of centroid defuzzification. In the following example, we demonstrate how the fuzzy system works to modulate the ontology relation of *increases* between two concepts of *stress response* and *Gonadotropin secretion* based on receiving evidence from literature results.

In the first stage, the agent specifies each type of research results (TRR) in specified article conclusions or abstract sections and calculate their percentage (LOP) of appearing in a given set. The result of such calculation is used to fuzzify the antecedent inputs. Then, depending on the obtained value of each input, the output value is determined according to the fuzzy rules. For example, if the system identifies that the percentage of verifying or positive TRR ranges between 55-90% of all the SRO statements (including verifying, disproving, and neutral), then the system output will be *is highly likely to*. In other words, if at least 55% of the TRR belongs to the Positive conclusion (+ve) category, then the *is likely to increase* relation between the two concepts will be changed to *is highly likely to increase*, i.e. "*Stress Response is highly likely to increase Gonadotropin secretion*".

The verifying, disproving, or neutral statements of the TRRs may appear in different linguistic forms. Therefore, the system should employ logical operators of OR, AND (maximum, minimum) in order to reduce different statements of the same type to a predefined category of the TRR. For example, we may have the following calculation for the *positive conclusion* (+ve) category of the TRR:

IF *"stress response* ***escalated*** *Gonadotropin secretion"* OR *"stress response* ***rises*** *Gonadotropin secretion"* OR *"stress response* ***has positive association with*** *Gonadotropin secretion"* OR *"stress response* ***predicted*** *Gonadotropin secretion"* THEN

"Stress Response is highly likely to increase Gonadotropin secretion".

In the above example, the LOP (level of proof) of 55-90% is the Union (∪) of the LOP of these verifying statements: *escalated*, *rises*, *has positive association with*, and *predicted*.

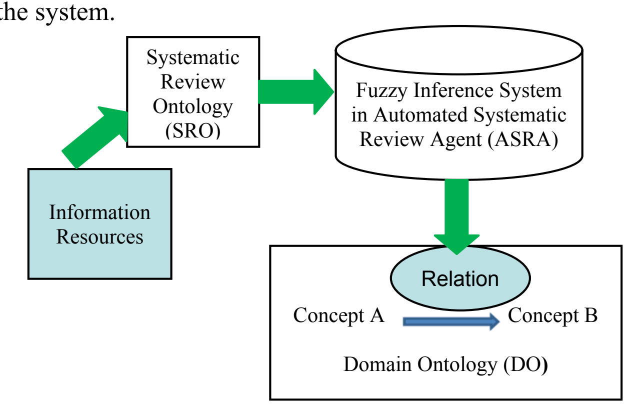
Notice that the statements after THEN are the ontology statements in the DO which are aimed to be modified according to the evidence they receive.

Despite the simplicity of the voting approach in a systematic review, it has been suggested that it can hardly be a reliable method, as further examination of various studies may reveal that there have existed diverse degrees of contrast or variations in the study subjects (e.g. various age or sex groups), different adopted research methods, or in general, various contexts for different studies. This makes it difficult to compare results of selected studies through a simple voting method [19].

One of the most common methods used in a systematic review is *meta-analysis*. As a quantitative research analysis, meta-analysis uses a host of statistical methods to combine the empirical results of various studies, which have used a multitude of data sets and methods, in order to provide more insight and stronger explanatory power for individual studies, evaluate conflicting scientific evidences of various results, and clarify controversies [19]. In a meta-analysis procedure, each original individual study is considered to be part of a larger study. Data from each single study is amalgamated to produce one single and final result, thereby summarizing the evidence as a whole and producing valid generalizations [12].

The incorporation of an automated meta-analysis agent into the EDBEO methodology is not practical as the meta-analysis process requires access to all collected data, implemented methods, and different aspects of the context in which each relevant study has been carried out. Retrieval of such complicated information requires too much time and manual work. However, it is possible to implement such an ambition provided all researchers begin to store their obtained data and other information relevant to their studies in a specific database where various data and information about research results can be stored, encoded, annotated, and subject to knowledge engineering technologies. In this way, we will be able to have a special data repository to which a *meta-analysis agent* can have access so that it can encode, classify, and analyse the results of various studies, and thereby, produce statistically meaningful conclusions as to the evidentiality of a certain ontology relationship between two concepts. Such conclusions can then be applied to the ontology relationships of a domain ontology as described in the previous section.

Results of a multitude of studies may point to the existence of correlation between different concepts. However, there is not much consensus, common definitions, or indices for the meaning and context of those concepts. The proposed *meta-analysis agent* can encode the results of various studies according to the predefined definitions and contexts of each study. The system can ask the researcher to ensure his concepts match the definition of predefined concepts of the domain ontology. However, if an introduced concept is completely or relatively new to the ontology, the agent will add that concept with its relevant context to the ontology. In this way, other researchers will be able to learn about newly-arrived concepts and novel phenomena in their domain of interest, start their own investigation of those concepts, and add their results to the system.



4.2. Ontology-based literature discovery framework

The second component of the EDBEO is a literature-based discovery framework which uses the concepts in the domain ontology in order to discover new connections between domain concepts in the literature. Similar to [15] method of utilizing MeSh terms for the extraction and ranking of interesting concepts, our approach uses the domain ontology to facilitate the process of finding intermediate concepts in the literature. The discovered concepts and connections, will then be added to the domain ontology to enrich its conceptual specification of the domain as well as promote its comprehensiveness and incorporate new discoveries. The new added discoveries, which are represented as new *concept-relationship-concept* combinations, will be further linked to the ASRA framework to undergo the relationship refinement process.

In a related work, [20] applied ontology matching techniques to perform the process of *concept discovery* for the *enrichment* of the domain ontology.

5. Conclusion

Based on the assumption that ontological relations such as "*causes*" in scientific domains are grounded in scientific evidence, and with regard to the importance of discovering new connections and knowledge in a domain literature, we proposed the design of an ontology evolution methodology (EDBEO) by means of which evidence-based ontologies, which can also be a foundation for literature-based discovery, can be established. The EDBEO methodology has the potential to enhance the accuracy and evidentiality of ontology relations as well as facilitate literature reviews and literature-based discoveries.

Conventional search methods require that researchers undertake extensive Web searches, paper reviews, and statistical analysis to learn whether a theory which has been addressed by a broad range of studies is evidence-based and, if so, to what degree. Moreover, they need to keep track of the emerging study results to keep up to date with any supporting evidence or refutation of a scientific hypothesis in their domain of interest. EDBEO can be regarded as a facilitating tool with which investigators can keep pace with the latest research results about various degrees of proof or refutation of a scientific theory in the form of explicit facts represented in a domain ontology. Furthermore, by retrieval, analysis, and representation of the results of the latest research works, EDBEO has the potential to produce a state-of-the-art report to researchers, enabling them to write more comprehensive and accurate review papers.

We suggest researchers should be aware of any single study conducted in the world regarding the creation of any proving or disproving evidence for a given theory. The EDBEO methodology aims to do this job in an automated way, compensating for the inability of researchers to be cognisant of all existing studies on a certain topic. It also has the potential to represent and predict the emergence and discovery of new implicit connections between scientific theories in any domain of interest.

Another important implication of this proposal is that, to our knowledge, it is the first proposed methodology which aims to integrate methods of systematic review and literature-based discovery as plausible strategies for the evolution of scientific ontologies.

6. References

[1] R. Studer, V. R. Benjamins, and D. Fensel, "Knowledge engineering: principles and methods," *Data & Knowledge Engineering,* vol. 25, pp. 161-197, 1998.

[2] H. T. Hunt, "Why psychology is/is not traditional science: The self-referential bases of psychological research and theory," *Review of General Psychology,* vol. 9, pp. 358-374, 2005.

[3] P. Compton and R. Jansen, "A philosophical basis for knowledge acquisition," *Knowledge Acquisition,* vol. 2, pp. 241-257, 1990.

[4] B. Smith, "Beyond concepts: Ontology as reality representation," in *International Conference on Formal Ontology and Information Systems*, Turin, November 2004.

[5] E. N. Khoozani and M. Hadzic, "Human stress ontology: Multiple applications and implications of an ontology framework in the mental health domain," in *3rd International Conference on Health Informatics*, Valencia, Spain, 2010, pp. 228-234.

[6] "Gene Ontology Consortium. The Gene Ontology (GO) project in 2006," *Nucleic Acids Research,* vol. 34, pp. D322–326, 2006.

[7] P. De Leenheer and T. Mens, "Ontology evolution: State of the art and future directions," in *Ontology management, semantic web, semantic web services, and business applications* M. Hepp, P. De Leenheer, A. De Moor, and Y. Sure, Eds., 2008, pp. 131-176.

[8] M. Prasenjit and G. Wiederhold, "An ontology-composition algebra," in *Handbook on ontologies*, S. Staab and R. Studer, Eds. Berlin Heidelberg: Springer, 2004, pp. 93-113.

[9] D. B. Lenat, "Cyc: A large scale investment in knowledge infrastructure," *Communications of the ACM,* vol. 38, 1995.

[10] D. R. Swanson, N. R. Smalheiser, and A. Bookstein, "Information discovery from complementary literatures: Categorizing viruses as potential weapons," *Journal of the American Society for Information Science and Technology,* vol. 52, pp. 797–812, 2001.

[11] C. D. Mulrow, D. J. Cook, and F. Davidoff, "Systematic reviews: critical links in the great chain of evidence," *Annals of Internal Medicine,* vol. 126, p. 389, 1997.

[12] J. C. de Almeida Biolchini, P. G. Mian, A. C. C. Natali, T. U. Conte, and G. H. Travassos, "Scientific research ontology to support systematic review in software engineering," *Advanced Engineering Informatics,* vol. 21, pp. 133-151, 2007.

[13] T. C. Taveggia, "Resolving research controversy through empirical cumulation: Toward reliable sociological knowledge," Sociological methods & research, vol. 2, p. 395, 1974.

[14] M. C. Ganiz, W. M. Pottenger, and C. D. Janneck, "Recent advances in literature based discovery," Lehih University 2005.

[15] P. Srinivasan, "Text mining generating hypotheses from MEDLINE," Journal of the American Society for Information Science and Technology, vol. 55, pp. 396-413, 2004.

[16] H. J. Lowe and G. O. Barnett, "Understanding and using the medical subject headings (MeSH) vocabulary to perform literature searches," Journal of the American Medical Association, vol. 271, pp. 1103-1108, 1994.

[17] K. M. Passino and S. Yurkovich, Fuzzy control. Menlo Park, CA: Addison Wesley Longman, 1998.

[18] L. A. Zadeh, "Fuzzy logic and approximate reasoning," Synthese, vol. 30, pp. 407-428, 1975.

[19] T. D. Stanley, "Wheat from chaff: Meta-analysis as quantitative literature review," The Journal of Economic Perspectives, vol. 15, p. 131, 2001.

[20] S. Castano, A. Ferrara, and G. N. Hess, "Discovery-driven ontology evolution," in 3rd Workshop on Semantic Web Applications and Perspectives (SWAP-06), Pisa, Italy, 2006.