A Tailored Ontology Sculpt For Web Information Congregation

P. Gopinath[1], A. Dhasaradhi[2], P. Nirupuma[3],
[1]M.Tech Student, [2]Asst. Prof, [3]Prof, Head
[1, 2, 3]Department of CSE,
Siddharth Institute of Engineering & Technology,
Puttur, Andhrapradesh, India,

Abstract— **As a sculpt for acquaintance explanation and exemplification, ontologies are extensively used to symbolize consumer profiles in tailored web information congregation. Conversely, when representing consumer profiles, many sculpt have utilized only acquaintance from either a global acquaintance base or a consumer local information.**

In this paper, a tailored ontology sculpt is planned for acquaintance illustration and reckoning over consumer profiles. This sculpt learns ontological consumer profiles from mutually a world acquaintance base and consumer confined occurrence repositories. The ontology sculpt is evaluated by comparing it against yardstick models in web information congregation. The consequences illustrate that this ontology sculpt is flourishing.

Keywords- **Ontology, world acquaintance, local instance repository, consumer profiles, web information congregation,**

I. INTRODUCTION

ON the last decades, the quantity of web-based in sequence obtainable has increased radically. How to gather useful in sequence from the web has turn into a demanding issue for users. Current web in sequence congregation systems attempt to gratify user necessities by capturing their information requirements. For this purpose, user profiles are created for user environment knowledge explanation.

User profiles stand for the perception models obsessed by users when congregation web in sequence. An impression model is unreservedly infatuated by users and is generated from their conditions knowledge. While this perception reproduction cannot be proven in laboratories, many network ontologists have experimental it in user concert [19]. When users read through a certificate, they can easily conclude whether or not it is of their interest or significance to them, a conclusion that arises from their understood concept models. If a user's conception model can be replicated, then a higher demonstration of user profiles can be built.

To reproduce user thought models, ontologies—a knowledge account and formalization model—are utilized in personalized web in sequence congregation. Such ontologies are called ontological client profiles or custom-made ontologies [18]. To represent client profiles, a lot of researchers have attempted to determine user surroundings knowledge through total or local investigation.

Global scrutiny uses existing global acquaintance bases for user environment awareness illustration. Commonly used comprehension bases include common ontologies, thesauruses (e.g., digital libraries), and online acquaintance bases (e.g., online categorizations and Wikipedia). The global analysis techniques manufacture successful appearance for client surroundings knowledge extraction. However, global investigation is incomplete by the superiority of the used awareness base. For example, WordNet was reported as cooperative in capturing client attention in some areas but hopeless for others .

Local analysis investigates client local in sequence or observes client performance in client profiles. For example, Li and Zhong revealed taxonomical patterns from the users' local text credentials to study ontologies for client profiles. A number of groups [12] well-read made to order ontologies adaptively from client's browsing the past. Otherwise, Sekine and Suzuki analyzed reservation kindling to discover user background knowledge. In some works, such as, clients were provided with a set of credentials and asked for significance response. Client background comprehension was then revealed beginning this criticism for client profiles. However, because local investigation techniques rely on data removal or organization techniques for knowledge discovery, irregularly the exposed results surround noisy and uncertain in sequence. As a product, local scrutiny suffers from uselessness at capturing formal client acquaintance.

From this, we can conjecture that client background awareness can be better revealed and represented if we can put together worldwide and local investigation within a hybrid Model. The knowledge formal in a global awareness base will constrict the background knowledge detection from the client restricted in sequence. Such a modified ontology replica should manufacture a superior demonstration of client profiles for web in sequence gathering.

In this paper, an ontology representation to appraise this supposition is planned. This model simulates clients impression models by means of adapted ontologies, and attempts to look up web in sequence congregation performance by using ontological user profiles. The world acquaintance and a user's local occurrence reservoir (LIR) are used in the proposed representation. World knowledge is commonsensical acquaintance acquired by people from understanding and edification [16]; an LIR is a user's not public collection of in sequence items. From a world acquaintance base, we assemble made to order ontologism by adopting user reaction on motivating knowledge. A multidimensional ontology withdrawal method, Specificity

and Exhaustively, is also introduced in the projected model for analyzing concepts individual in ontologies. The projected ontology model is evaluated by assessment touching several standard models through experiments using a large average data set. The evaluation results be evidence for that the proposed ontology model is flourishing.

The research contributes to acquaintance business, and has the in the near future to improve the design of made to order web in sequence congregation systems. The donations are original and all the time more noteworthy, allowing for the speedy detonation of web in sequence and the increasing user-friendliness of online credentials.

The paper is prepared as follows: Section 2 discusses the associated vocation; in Section 3, we commence how adapted ontologies are constructed for users; and in Section 4, we present the multidimensional ontology withdrawal method. After that, Section 5 gives the architecture of the proposed model; Section 6 discusses the assessment issues, and the consequences are analyzed in Section 7. Finally, Section 8 makes conclusions and addresses our expectations work.

II. RELATED WORK

A. Ontology Learning

Global acquaintance bases were used by lots of accessible models to be taught ontologies for web in sequence congregation. For example, Gauch et al. [12] and Sieg et al. [15] learned modified ontologies from the unlock address list project to specify client's preferences and happiness in web explore. Wikipedia was used by Downey et al. [10] to help appreciate fundamental user happiness in queries. These works successfully revealed client background acquaintance; however, their concert was inadequate by the superiority of the international acquaintance bases.

Aiming at knowledge made to order ontologies, various works mined client background acquaintance from client local information.   Li and Zhong [13] second-hand model acknowledgment and organization rule withdrawal techniques to determine knowledge from user local credentials for ontology manufacture. Transept al. translated keyword queries to explanation Logic's conjunctive queries and second-hand ontologies to symbolize client background acquaintance. Zhong proposed a sphere ontology education approach that in employment various data withdrawal and natural-language considerate techniques. Navigli residential OntoLearn to notice semantic concept and associations from web credentials. Web contented taking out techniques will be used by Jiang and Tan [16] to determine semantic acquaintance from domain-specific content credentials for ontology knowledge. Finally, Shehata et al. [14] captured client in order needs at the judgment level somewhat than the deed level, and represented client profiles by the intangible Ontological Graph. The use of data withdrawal techniques in these models leads to more client background acquaintance being exposed. However, the acquaintance exposed in these works controlled noise and reservations.

In addition, ontologies were used in a lot of works to look up the presentation of acquaintance sighting. Using a fuzzy

sphere ontology taking out algorithm, a apparatus was developed by Lau et al. [19] in 2009 to make concept maps based on the posts on online conversation forums. Quest and Ali [15] used ontologies to help data pulling out in ordinary databases. Jin et al. [17] integrated data removal and in sequence recovery techniques to further enhance knowledge discovery. Doan et al. [8] proposed a model
called stick and used machine education techniques to find alike concepts in different ontologies. Dou et al. [9] planned a construction for scholarship sphere ontologies using pattern disintegration, clustering/classification, and organization rules withdrawal techniques. These works attempted to walk around a direction to mold world acquaintance extra proficiently.

B. User Profiles

User profiles were second-hand in net in sequence congregation to construe the semantic meanings of queries and incarcerate customer in sequence needs [12], [14]. User profiles were definite by Li and Zhong [14] as the motivating topics of a user's in sequence need. They also categorized client profiles into two diagrams: the data diagram user profiles acquired by analyzing a database or a set of communication [12], [11], [17], [13], [11]; the in sequence illustration user profiles acquired by using manual techniques, such as questionnaires and interviews, or mechanical techniques, such as in sequence recovery and machine learning. Van der Sluijs and Huben projected a method called the common client Model constituent to look up the excellence and consumption of client modeling. Wikipedia was in addition used by [10], to help determine client interests. In order to get hold of a user profile, Chirita et al. [6] and Teevan et al. used a collection of user desktop text credentials and emails, and cached web pages to search client interests. Makris et al. acquired client profiles by a ranked local position of categories, and then utilized web pages to personalize look for results for a client. These works attempted to get hold of client profiles in organize to discover client background acquaintance.

Client profiles can be categorized into three groups: interviewing, semi-interviewing, and non-interviewing. Interviewing client profiles can be deemed just the thing client profiles. They are acquired by means of physical techniques, such as Questionnaires, interviewing client's, and analyzing client confidential guidance sets. One distinctive example is the TREC Filtering Track teaching sets, which were generated by hand [13]. The users read each essay and gave a constructive or negative decision to the document alongside a given subject. Because, only client's completely know their wellbeing and preferences, these teaching credentials exactly reflect user background acquaintance. Semi-interviewing client profiles are acquired by partially computerized techniques with restricted user association. These techniques typically provide client's with a inventory of categories and request users for motivating or non-interesting categories. One distinctive example is the web training set attainment model introduced by Tao et al. [18], which extracts preparation sets from the net based on user feedback categories. Non-interviewing techniques do not engage users at all, but

174

determine user wellbeing instead. They acquire client profiles by observing client movement and performance and discovering client background knowledge [11]. A typical model is OBIWAN, planned by Gauch et al. [12], which acquires client profiles based on users' online browsing record. The interviewing, semi-interviewing, and non-interviewing client profiles can also be viewed as guide, semiautomatic, and habitual profiles, respectively.

III. MODIFIED ONTOLOGY CONSTRUCTIONS

Modified ontologies are a conceptualization representation that properly describes and specifies client background acquaintance. From explanation in everyday life, we found so as to web users strength have dissimilar prospect for the same exploration query. For example, for the topic "New York," commerce travelers may command dissimilar in sequence from free time travelers. Occasionally even the same client may have dissimilar expectations for the similar search uncertainty if applied in dissimilar circumstances. A client may become a commerce traveler when development for a commerce trip, or a free time traveler when development for a family holiday. Based on this surveillance, a postulation is fashioned that web client's have a private concept model for their in sequence needs. A client's concept model may modify according to different in sequence requirements. In this section, a representation constructing made to order ontologies for web users' perception models is introduced.

A. World Knowledge Demonstration

World knowledge is significant for in sequence congregation. According to the description provided by [16], planet acquaintance is reasonable knowledge obsessed by people and acquired through their understanding and schooling. Also, as piercing out by Nirenburg and Raskin [19],"world information is compulsory for lexical and referential disambiguation, together with establishing co-reference relatives and resolving contraction as well as for establishing and maintaining connectivity of the conversation and observance of the text to the text producer's ambition and plans." In this planned model, client background acquaintance is extracted from a world acquaintance base programmed from the Library of legislature Subject Headings (LCSH).

We first require making the world acquaintance base. The world information base must cover up an comprehensive assortment of topics, since users may move toward from dissimilar backgrounds. For this reason, the LCSH arrangement is an ideal world acquaintance pedestal. The LCSH was developed for organizing and retrieving in sequence from a large quantity of documentation collections. For over a hundred years, the acquaintance contained in the LCSH has undergone unremitting reconsideration and enhancement. The LCSH represents the ordinary growth and allocation of human academic work, and covers all-inclusive and comprehensive topics of world acquaintance [5]. In addition, the LCSH is the most wide-ranging non-specialized proscribed terminology in English. In many respects, the arrangement has develop into a defects average for subject matter classification and indexing, and is used as a means for

ornamental subject admission to acquaintance management systems [5].

TABLE 1
Comparison of Different World Taxonomies

	LCSH	LCC	DDC	RC
# of Topics	394,070	4,214	18,462	100,000
Structure	Directed Acyclic Graph	Tree	Tree	Directed Acyclic Graph
Depth	37	7	23	10
Semantic Relations	Broader, Used-for, Related-to	Super- and Sub-class	Super- and Sub-class	Super- and Sub-class

The LCSH arrangement is greater compared with extra world acquaintance taxonomies used in preceding mechanism. Table 1 presents an assessment of the LCSH with the documentation of legislative body categorization (LCC) used by Frank in addition to Paynter [11], the Dewey Decimal Classification (DDC) used as a result of Wang and Lee [15] and King et al. [18], and the situation categorization (RC) developed by Gauch et al. [12] using online categorizations. As shown in Table 1, the LCSH covers more topics, has a more precise structure, and specifies more semantic kindred. The LCSH descriptors are confidential by professionals, and the categorization quality is definite by well-defined and endlessly refined category rules [5]. These features make the LCSH an model world acquaintance base for acquaintance manufacturing and administration.

The structure of the humankind knowledge pedestal used in this investigate is encoded from the LCSH references. The LCSH classification contains three types of references: Broader term (BT), Used-for (UF), in addition to connected term (RT) [5]. The BT references are for two subject's recitation the same topic, but at dissimilar levels of construct (or specificity). In our model, they are programmed as the is-a associations in the world acquaintance base. The UF references in the LCSH are used for countless semantic situations, including enlargement the semantic amount of an issue and recitation complex subjects and subjects subdivided by additional topics. The compound tradition of UF references makes them tricky to instruct. Throughout the assessment, we found that these references are time and again used to explain an achievement or an entity. When entity A is used for an exploit, A becomes a portion of that exploit (e.g., "a divergence is used for dining"); when A is used for an additional object, B, A becomes an ingredient of B (e.g., "a wheel is use for a car"). These cases can be encoded as the part-of kindred. Thus, we simplify the compound tradition of UF references in the LCSH and instruct them merely as the part-of relations in the world acquaintance stand. The RT references are for two subjects connected in some manner supplementary than by pecking order. They are programmed as the related-to relations in our world acquaintance base.

The primitive acquaintance element in our humankind acquaintance stand is subjects. They are programmed from the topic headings in the LCSH. These subjects are formalized as follows:

Definition 1. Let \$ be a set of subjects, an building block is $s \in \$$ s as a 4-tuple s: = (label; neighbor; ancestor; descendant), wherever

- label is the caption of s in the LCSH glossary;
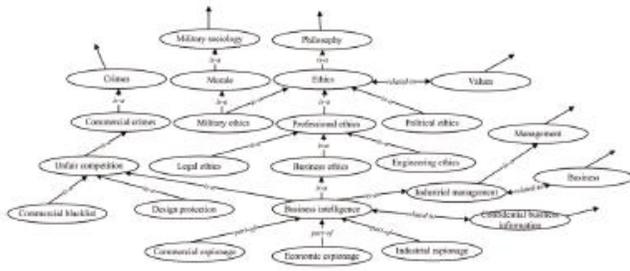- neighbor is a gathering returning the subjects that have straight links to s in the world knowledge base;

Fig 1: A sample part of the world knowledge base.

- ancestor is a meaning frequent the subjects that have a superior level of construct than s and link to straight or not directly in the humankind acquaintance base.
- Successor is a meaning frequent the subjects that are more detailed than s and link to s straight or obliquely in the world acquaintance base.

The subjects in the world acquaintance base are associated to each other by the semantic kindred of is-a, part-of, and related-to. The family are formalized as follows:

Definition 2 .Let to IR be a position of associations, an component r belongs R is a 2-tuple r: = {edge; type}, where

- an edge connects two subjects that clutch a type of relation;
- a type of relations is an element of {is-a; part-of; related-to}.

With Definitions 1 and 2, the world acquaintance stand can then be pompous as follows:

Definition 3. Let WKB be a world acquaintance base, which is taxonomy constructed as a bound for acyclic chart. The WKB consists of a set of subjects associated by their semantic kindred, and can be properly distinct as a 2-tuple WKB ::={\$; IR},
Where .

- \$ is a set of subjects $\$:= \{s_1; s_2; \ldots ; s_m\}$;
- IR is a set of semantic relations $IR := \{r_1; r_2; \ldots ; r_n\}$ linking the subjects in \$.

B. Ontology Production

The subjects of consumer attention are extracted from the WKB via user communication. A tool called Ontology knowledge surroundings (OLE) is urbanized to support users through such communication. Concerning a theme, the attractive subjects consist of two sets: optimistic subjects are the concepts applicable to the in sequence need, and

pessimistic subjects are the concepts resolving inconsistent or confusing understanding of the in sequence need. Thus, for a set topic, the OLE provides users through a set of candidates to recognize affirmative and pessimistic subjects. These applicant subjects are extracted from the WKB.

Fig. 2 is a screen-shot of the OLE for the example topic "financial intelligence." The subjects scheduled on the top-left plate of the OLE are the applicant subjects obtainable in hierarchical form. For each s 2 SS, the s and its associates are
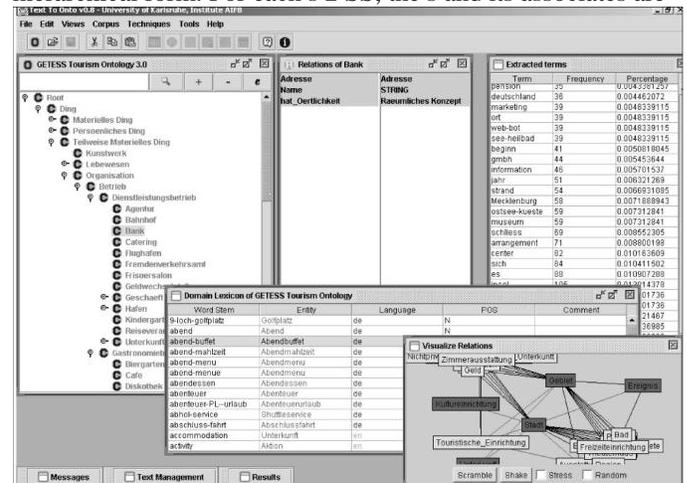

Fig 2. Ontology learning environment.

Retrieve if the tag of s contains any one of the inquiry terms in the known topic (e.g., "financial" and "spying"). From these candidates, the user selects positive subjects for the topic. The user-selected positive subjects are presented on the top-right panel in hierarchical form.

The applicant negative subjects are the descendants of the user-selected positive subjects. They are shown on the bottom-left panel. From these negative candidates, the user selects the negative subjects. These user-selected pessimistic subjects are listed on the bottom-right panel (e.g., "Political ethics" and "Student ethics"). Note that for the conclusion of the agreement, some optimistic subjects (e.g., "Ethics," "Crime," "Commercial crimes," and "Competition Unfair") are also included on the bottom-right panel with the negative subjects. These positive subjects will not be included in the negative set. The outstanding candidates, which are not fed back as whichever optimistic or unenthusiastic from the user, turn out to be the neutral subjects to the given topic.

Ontology is then constructed for the given topic by means of this client feedback subjects. The construction of the ontology is based on the semantic links regarding these data in the WKB. The ontology contains three types of acquaintance: positive subjects, negative subjects, and neutral subjects. Fig. 3 illustrates the ontology (incompletely) constructed for the case in point topic "Economic espionage," where the white nodes are optimistic, the dark nodes are negative, and the gray
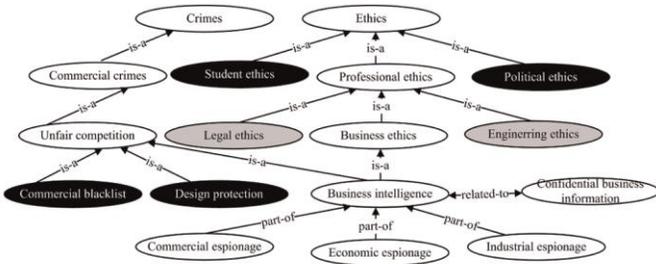
Fig.3 An ontology (partial) constructed for topic "Economic Espionage."

IV. MULTIDIMENSIONAL ONTOLOGY MINING

Ontology removal discovers attractive and on-topic acquaintance from the concepts, semantic kindred, and instances in an ontology. In this section, a 2D ontology mining method is introduced: Specificity and Exhaustivity. Specificity (denoted spe) describes a subject's center on a given topic. Exhaustivity (denoted exh) restricts a subject's semantic space dealing with the topic. This method aims to investigate the subjects and the potency of their relations in an ontology.

We argue that a subject's specificity has two focuses: 1) on the referring-to concepts (called semantic specificity), and 2) on the given topic (called topic specificity). These need to be addressed separately.

A. Semantic Specificity

The semantic specificity is investigated based on the arrangement of O(T) innate from the world knowledge base. The strength of such a focus is influenced by the Subject's locality in the taxonomic structure tax^S of O(T) (this is also argued by [42]). As stated in Definition 4, the tax^S of O(T)Þ is a graph linked by semantic relations. The subjects located at upper bound levels toward the root are more abstract than those at lower bound levels toward the "leaves." The upper bound level subjects have more descendants, and thus refer to more concepts, compared with the lower bound level subjects. Thus, in terms of a concept being referred to by both an upper bound and lower bound subjects, the lower bound subject has a stronger focus because it has fewer concepts in its space. Hence, the semantic specificity of a lower bound subject is greater than that of an upper bound subject

V. ARCHITECTURE OF THE ONTOLOGY MODEL

The projected ontology sculpt aims to determine user background acquaintance and learns tailored ontologies to symbolize consumer profiles. Fig. 6 illustrates the building of the ontology sculpt. A tailored ontology is constructed, according to a given topic. Two acquaintance possessions, the global world knowledgebase and the user's local occurrence depository, are utilized by the sculpt. The world acquaintance base provides the taxonomic configuration for the tailored ontology. The user background acquaintance is revealed from the user local occasion depository. Against the given topic, the specificity and exhaustivity of subjects are investigated for user background acquaintance discovery.
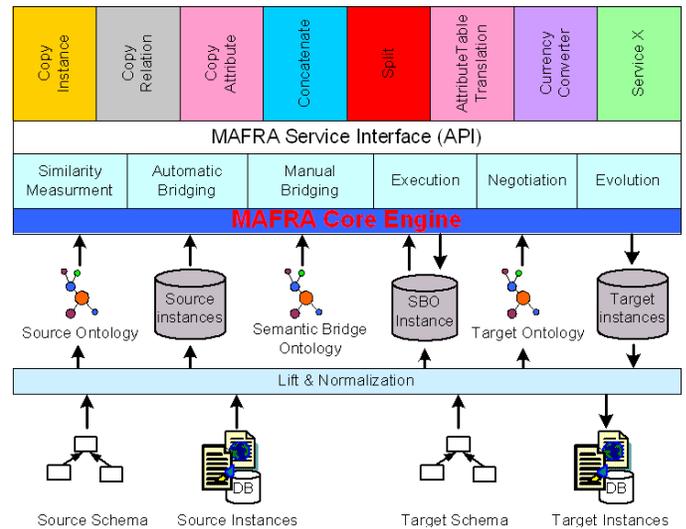

Fig: Architecture of the ontology model.

VI. EVALUATION

A. Experiment Design

The projected ontology sculpt was evaluated by intention experiments. Because it is tricky to evaluate two sets of acquaintance in diverse representations, the primary devise of the valuation was to evaluate the efficiency of an information congregation system (IGS) that used diverse sets of user background acquaintance for information congregation. The acquaintance discovered by the ontology sculpt was first used for a run of information congregation, and then the acquaintance manually particular by users was used for another run. The latter run set up a yardstick for the appraisal because the acquaintance was manually specified by users.

Under the similar untried circumstances, if the IGS might realize the same (or similar) presentation in two dissimilar runs, we could prove that the naked acquaintance has the same excellence as the user particular acquaintance. The projected ontology sculpt could then be proven capable to the domain of web information congregation. In information congregation evaluations, a common batchstyle research is developed for the evaluation of different models, using a test set and a set of topics connected with relevant judgments [36].

Our experiments followed this move toward and were performed under the tentative situation set up by the TREC-11 Filtering Track.3 This way aimed to evaluate the methods of persistent user profiles for disentanglement apposite and nonrelevant documents in an conventional stream [12]. User atmosphere acquaintance in the experiments was represented by consumer profiles, such as those in the experiments of [13] and the TREC-11 Filtering Track. A consumer profile consisted of two manuscript sets: a positive document set containing the on-topic, appealing acquaintance, and a unconstructive manuscript set containing the contradictory, confusing concepts. Each manuscript d held a support value sustain the given topic. Based on this demonstration, the baseline models in our experiments were suspiciously selected. Consumer profiles can be categorized into three

groups: interviewing, semi-interviewing, and non interviewing profiles, as previously discussed in Section 2. In an attempt to estimate the planned ontology sculpt to the archetypal models representing these three group consumer profiles, four models were implemented in the experiments:

1. The Ontology sculpt that implemented the planned ontology sculpt. Consumer background acquaintance was computationally naked in this sculpt.
2. The TREC sculpt that represented the perfect interviewing consumer profiles. User background familiarity was manually specified by users in this sculpt.
3. The Category sculpt that represented the no interviewing consumer profiles.
4. The Web sculpt that represented the semi-interviewing consumer profiles.

VII. CONCLUSIONS AND FUTURE WORK

In this paper, an ontology sculpt is projected for representing consumer backdrop knowledge for tailored web information congregation. The sculpt constructs consumer tailored ontologies by extracting world acquaintance from the LCSH system and discovering consumer background acquaintance from consumer local occurrence repositories. A multidimensional ontology mining technique, exhaustivity and specificity, is also introduced for consumer background acquaintance discovery. In appraisal, the normal topics and a huge testbed were used for experiments. The sculpt was compared against yardstick models by applying it to a frequent system for information congregation.

The research consequences reveal that our planned sculpt is gifted. A sympathy investigation was also conducted for the ontology sculpt. In this analysis, we originate that the amalgamation of inclusive and local acquaintance works better than using any one of them.

In addition, the ontology sculpt using acquaintance with both is-a and part-of semantic kindred works better than using only one of them. When using only global acquaintance, these two kinds of kindred have the same assistance to the concert of the ontology sculpt. While using both global and local acquaintance, the acquaintance with part-of relations is more imperative than that with is-a.

The future ontology sculpt in this paper provides a explanation to emphasizing inclusive and local acquaintance in a single computational sculpt. The findings in this paper can be functional to the design of web information congregation systems. The sculpt also has general assistance to the fields of Information Rescue, web astuteness, Commendation Systems, and Information Systems.

In our future work, we will examine the methods that engender user local occasion repositories to contest the depiction of a global acquaintance base. The present work assumes that all user local illustration repositories have content-based descriptors referring to the subjects, however, a large quantity of documents presented on the web may not have such content-based descriptors. For this problem, in Section IV-B, strategies like ontology mapping and text cataloging/clustering were suggested. These strategies will be investigated in future work to solve this trouble. The examination will enlarge the applicability of the ontology sculpt to the mainstream of the offered web documents and increase the contribution and significance of the present work.

REFERENCES

[1] R. Baeza-Yates and B. Ribeiro-Neto, Modern Information Retrieval. Addison Wesley, 1999.

[2] G.E.P. Box, J.S. Hunter, and W.G. Hunter, Statistics For Experimenters. John Wiley & Sons, 2005.

[3] C. Buckley and E.M. Voorhees, "Evaluating Evaluation Measure Stability," Proc. ACM SIGIR '00, pp. 33-40, 2000.

[4] Z. Cai, D.S. McNamara, M. Louwerse, X. Hu, M. Rowe, and A.C. Graesser,NLS: A Non-Latent Similarity Algorithm," Proc. 26th Ann. Meeting of the Cognitive Science Soc. (CogSci '04), pp. 180-185, 2004.

[5] L.M. Chan, Library of Congress Subject Headings: Principle and Application. Libraries Unlimited, 2005.

[6] P.A. Chirita, C.S. Firan, and W. Nejdl, "Personalized Query Expansion for the Web," Proc. ACM SIGIR ('07), pp. 7-14, 2007.

[7] R.M. Colomb, Information Spaces: The Architecture of Cyberspace. Springer, 2002.

[8] A. Doan, J. Madhavan, P. Domingos, and A. Halevy, "Learning to Map between Ontologies on the Semantic Web," Proc. 11th Int'l Conf. World Wide Web (WWW '02), pp. 662-673, 2002.

[9] D. Dou, G. Frishkoff, J. Rong, R. Frank, A. Malony, and D. Tucker,"Development of Neuroelectromagnetic Ontologies(NEMO): A Framework for Mining Brainwave Ontologies," Proc. ACM SIGKDD ('07), pp. 270-279, 2007.

[10] D. Downey, S. Dumais, D. Liebling, and E. Horvitz, "Understanding the Relationship between Searchers' Queries and Information Goals," Proc. 17th ACM Conf. Information and Knowledge Management (CIKM '08), pp. 449- 458, 2008.

[11] E. Frank and G.W. Paynter, "Predicting Library of Congress Classifications from Library of Congress Subject Headings," J. Am. Soc. Information Science and Technology, vol. 55, no. 3, pp. 214-227, 2004.

[12] S. Gauch, J. Chaffee, and A. Pretschner, "Ontology-Based Personalized Search and Browsing," Web Intelligence and Agent Systems, vol. 1, nos. 3/4, pp. 219-234, 2003.

[13] R. Gligorov, W. ten Kate, Z. Aleksovski, and F. van Harmelen, "Using Google Distance to Weight Approximate Ontology Matches," Proc. 16th Int'l Conf. World Wide Web (WWW '07), pp. 767-776, 2007.

[14] J. Han and K.C.-C. Chang, "Data Mining for Web Intelligence," Computer, vol. 35, no. 11, pp. 64-70, Nov. 2002.

[15] B.J. Jansen, A. Spink, J. Bateman, and T. Saracevic, "Real Life Information Retrieval: A Study of User Queries on the Web," ACM SIGIR Forum, vol. 32, no. 1, pp. 5-17, 1998.

[16] X. Jiang and A.-H. Tan, "Mining Ontological Knowledge from Domain Specific Text Documents," Proc. Fifth IEEE Int'l Conf. Data Mining (ICDM '05), pp. 665-668, 2005.

[17] W. Jin, R.K. Srihari, H.H. Ho, and X. Wu, "Improving Knowledge Discovery in Document Collections through Combining Text Retrieval and LinkAnalysis Techniques," Proc. Seventh IEEE Int'l Conf. Data Mining (ICDM '07), pp. 193-202, 2007.

[18] J.D. King, Y. Li, X. Tao, and R. Nayak, "Mining World Knowledge for Analysis of Search Engine Content," Web Intelligence and Agent Systems, vol. 5, no. 3, pp. 233-253, 2007.

[19] R.Y.K. Lau, D. Song, Y. Li, C.H. Cheung, and J.X. Hao, "Towards a Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning," IEEE Trans. Knowledge and Data Eng., vol. 21, no. 6, pp. 800- 813, June 2009.

178