

Grammar Based Annotation Content Mapping in the Social Page in the Context of Short Text

T. Sravanthi¹, G. Sowmya²

¹PG Scholar, Dept of CSE, Marri Laxman Reddy Institute of Technology, Dt: R.R District, AP, India

²Assistant Professor, Dept of CSE, Marri Laxman Reddy Institute of Technology, Dt: R.R District, AP, India

ABSTRACT

Technology and its social value have significantly changes the global village to the extent of information in single hand taking forward to the best level of path of information technology. These days If we look forward to the Industrial of Domain of Information Technology has changed a lot to social value media of software development either a service based or a product based organization, there would be good and healthy competition leads to today's magnificent world to be better look and having the best of best to next level of research. If we consider the Wikipedia and text content may or may not be accurate, but looking to nest level it's a free and open source no revenue getting generated and lastly informatics one. But now question came into mind that is it developed to best of technology which is changing environment, do think so, Hence we considered the accuracy and fast ontology based annotation to make the significant value to the nest level by putting the best flow architecture in the annotation of text or any content but limited to some symbols.

Index Terms:- Annotation, Grammar, Ontology, Comparison

1. INTRODUCTION

Annotation stands for the process of describing images, and retrieval stands for the process of finding images. The two major approaches to image retrieval are content based image retrieval that analyzes the actual image data, and metadata-based approach that retrieves images based on human-annotated metadata. Also relevance feedback has been used in image retrieval complementing text-based systems. In this paper the retrieval is done by using the annotated metadata, and not the content-based analysis or relevance feedback. The research problem is: "How should the metadata be created and what kind of system could interpret the metadata to make it easy to find images for an average user.

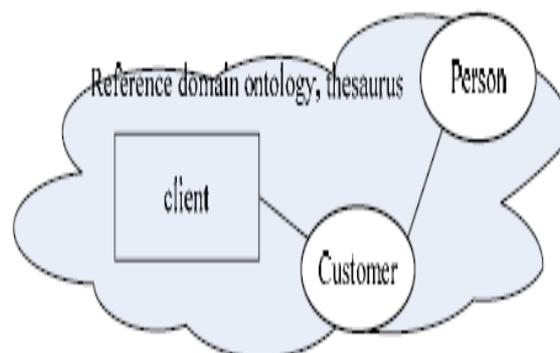


Fig.1.1 Overview of Annotation Based Ontological Search

The word ontology comes from the Greek words onto (being) and logos (study): Ontology is the study of being. On some occasions ontology is treated as a synonym for metaphysics, domain, or context. The process starts from acquirement of domain knowledge, possibly with the aid of domain experts. When new information is received the ontologist can build a prototype, possibly consulting the programmers. The prototype is then evaluated by the client to decide how to refine it. If client evaluation is not needed the ontologist can proceed by gathering more information that is added into the ontology. The ontology can be refined also according to the client's feedback without acquiring domain knowledge from anywhere else.

2. RELATED WORK

The usual case is that the annotator has a number of fields where the required values can be set. Some fields take text as value and some take values like integer, boolean, date, etc. Some values might be more or less predefined. There might be a menu with a number of color options associated to a field, and properties like length could be constrained into the metric system. Many systems require giving proper values to certain fields before accepting an annotation. Then again all fields do not have to be given values because these might not be essential or there possibly does not exist values for all the fields with all the annotated items.

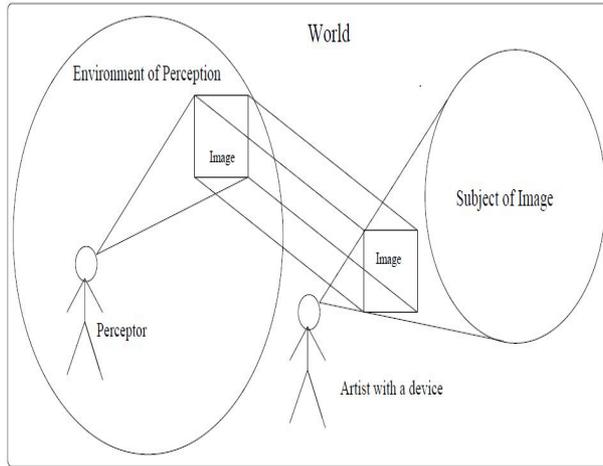


Fig.2.1 Relational Model for world of perceptual Fact

Ontology-based design is the only way to construct structure-based systems, and in general structure-based systems can be called ontology-based systems. The methods explained here are on very general level and ontologies have to be understood to thoroughly understand the structure-based paradigm. From a human point of view, the individual particulars are those things that a human would approximately consider as a unique as semblance of physical and abstract things that is so unique that there cannot be two absolutely identical of the same kind, or if there are; these too can be enumerated and called individuals. The particulars are often concrete, like individual humans, stars, and planets, but also all the categories can be called individuals. Every category and ontology is unique: if there are two ontologies that are not identical, then they are individual, and if two ontologies are identical, then they are the same individual ontology. And every category is unique, because there is no other category in it's specific place in some ontology, of which all are unique.

3. METHODOLOGY

The metadata that describes images can be roughly divided in two parts. One part concerns the concepts that give information about the creator of the image, tools used in the process of creating the image, art style of the image and the artist, price, and other explicit properties of the image. The other part describes what is actually in the image, the implicit properties that can be understood by perception the image itself. These two parts cannot be cleanly separated, and both have to be taken into account when analyzing an image. In the retrieval sense, photos that are annotated with plain text behave similarly to plain text documents because both contain text, which can be exploited by conventional text-based retrieval techniques.

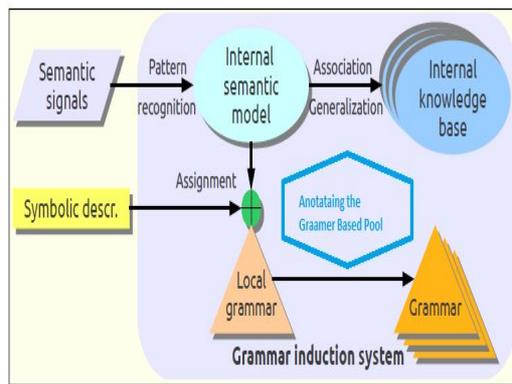


Fig.3.1. Architecture Flow of the Text and Grammar Based Content Mapping

The generic text-based information retrieval is carried on so that first a user types a query that consists of 1 to n keywords into a query field of a search interface. The search engine compares the keywords with a set of documents gathered from a database and gives them priority values. For example, if the keyword is book, document A contains two instances of book and document B contains only one instance, then A gets a higher priority. The documents are presented to the user, highest priority first. When the size and amount of the documents grows, the classical problems of text-based information retrieval start rising. Irrelevant documents are retrieved and the user has to use time filtering the information again, usually by browsing through the search results. The fitness of generic text-based retrieval is depicted. When the recall gets higher the precision gets lower, and when precision gets higher. The retrievers usually do not have to give values to all the fields of the search interface to find the wanted images and can be totally unaware of the system beforehand, when in contrast the annotators are usually trained for the job. When the annotator sets a certain value to a field like integer 1, the retrievers can possibly query a certain range of values such as integers between [0 10], all dates between years 1500-2000, and character strings just as with the text-based paradigm. Again, the use of an agreed vocabulary helps to find the wanted images, and in contrast, using incorrect vocabulary might make the search very hard. Fields with pre-defined value specifications are very helpful when the vocabulary used in the annotation process is unknown. It would be very time-consuming and unnecessary to always give values to all the possible fields if there are many of these. One simple way of executing a field-based search is to first give values to only a few fields and start the search. If the precision is too low the retriever can set more constraints by giving values for a few more fields or giving more accurate values for some fields.

4. CONCLUSION AND FUTURE WORK

Technology and its implementation have its own build in feature as of changes common in the industry of Information technology. All ontologies describe and relate entities in different ways, but still the boundaries of the description power of any ontology remains the same. No matter what kind of formal language is used in constructing ontology or creating queries. Looking forward for the enhancement which in term will be done by one or more is the Security mechanism involved in the open data source content of accuracy.

REFERENCE

- [1] Altavista search engine: <http://www.altavista.com>
- [2] Ilkka Autio, Tapio Elomaa: Flexible view recognition for indoor navigation based on Gabor filters and support vector machines. *Pattern Recognition* vol. 36, nro. 12, p.2769-2779, 2003.
- [3] S. Cucerzan. Large-scale named entity disambiguation based on Wikipedia data. *Proc. of Empirical Methods in NLP*, 2007.
- [4] P. Ferragina and A. Gulli. A personalized search engine based on web-snippet hierarchical clustering. In *Proc. WWW*, 801{810, 2005.
- [5] P. Ferragina and U. Scaiella. TAGME: On-the-y annotation of short text fragments (by Wikipedia entities). In *Proc. ACM CIKM*, 2010 (poster).
- [6] E. Gabrilovich and S. Markovitch. Feature generation for text categorization using world knowledge. In *Proc. IJCAI*, 1048{1053, 2005.
- [7] E. Gabrilovich and S. Markovitch. Wikipedia-based semantic interpretation for natural language processing. *J. Artif. Int. Res.*, 34(1):443{498, 2009.
- [8] J. Hu, L. Fang, Y. Cao, H.-J. Zeng, H. Li, Q. Yang, and Z. Chen. Enhancing text clustering by leveraging Wikipedia semantics. In *Proc. ACM SIGIR*, 179{186, 2008.
- [9] X. Hu, N. Sun, C. Zhang, and T.-S. Chua. Exploiting internal and external semantics for the clustering of short texts using world knowledge. In *Proc. ACM CIKM*, 919{928, 2009.
- [10] S. Kulkarni, A. Singh, G. Ramakrishnan, and S. Chakrabarti. Collective annotation of Wikipedia entities in web text. In *Proc. ACM KDD*, 457{466, 2009.
- [11] O. Medelyan, D. Milne, C. Legg, and I. H. Witten. Mining meaning from Wikipedia. *Int. J. Hum.-Comput. Stud.*, 67(9):716{754, 2009.
- [12] R. Mihalcea and A. Csomai. Wikify!: linking documents to encyclopedic knowledge. In *Proc. ACM CIKM*, 233{242, 2007.
- [13] D. Milne and I. H. Witten. An effective, low-cost measure of semantic relatedness obtained from Wikipedia links. In *Proc. AAAI Workshop on Wikipedia and Artificial Intelligence*, 2008.
- [14] D. Milne and I. H. Witten. Learning to link with Wikipedia. In *Proc. ACM CIKM*, 509{518, 2008.
- [15] M. Sahami and T. Heilman. A web-based kernel function for measuring the similarity of short text snippets. In *Proc. WWW*, 377{386, 2006.
- [16] F. M. Suchanek, G. Kasneci, and G. Weikum. YAGO: A core of semantic knowledge. In *Proc. WWW*, 697{706,