VLDB 2010
The 36th International Conference on Very Large Data Bases
13 to 17 September 2010

Singapore

Reviews For Paper

Track	Demonstrations
Paper ID	900
Title	Facetedpedia: Enabling Faceted Search for Wikipedia

Masked Reviewer ID: Assigned_Reviewer_1
Review:

Question	
Detailed Comments	The Facetedpedia system described in this paper supports faceted search for the popular Wikipedia. The main interesting and new aspect of the system is its capability to dynamically discover query-dependent faceted interfaces. In contrast, existing similar systems are not query-dependent. The paper outlines the key algorithms and techniques used in the implementation of the system, and it also provides fairly detailed information about the implementation. A separate paper describing this work has been accepted by the WWW conference. The proposed demo scenario is very detailed and should interest the visitors of the demo. Facetedpedia should make an interesting demo and it has some nice features. Some of the weaknesses of this paper include: (1) the underpinning research does not seem to be fundamentally new; (2) only about 30% of the articles in Wikipedia are utilized in the system; and (3) only one demo scenario will be demonstrated.
Overall Rating	Neutral

Masked Reviewer ID: Assigned_Reviewer_2
Review:

Question	
Detailed Comments	This paper proposes a faceted search system Facetedpedia for Wikipedia that dynamically discovers a query-dependent faceted interface for a set of Wikipedia articles. Faceted Wikipedia search for complex queries [ref. 1] and geospatial semantics [ref. 2] has been studied, and a web site is available for public access (http://www.dbpedia.org). As Facetedpedia dynamically categorizes the search results based on ad hoc queries, it may significantly slow down the response time. In the example, why does the user need to select "what kind of entities are you looking for"? What are the options for the user? The system needs the user to input the domain of the search. I am wondering that since the system has already pre-computed the results of the categorization and obtains the keywords input by the user. Is it possible that the system can figure out the user's targeted domain or give some possible suggestions based on the available knowledge, rather than asking the user to manually select the domains. The system needs to obtain the information from the external searching engine, which is google in the example. This could be a system performance bottleneck, because the other processes will wait for the result from the external searching engine. There are ambiguities in the system, as shown in Figure 3. the system may return 'universal studio' and '{universal studio}' as different items. Similar ambiguities are found for 20th_Century_Fox and Metro-Goldwyn-Mayer. It would be better, if the system can eliminate the ambiguities in the searching result. Otherwise, the user may be confused about which path to go next. [ref. 1] Rasmus Hahn, Christian Bizer, Christopher Sahnwaldt, Christian Herta, Scott Robinson, Michaela Bˆ¢rgle, Holger Dˆ¢wiger, Ulrich Scheel: Faceted Wikipedia Search. 13th International Conference on?Business Information Systems (BIS 2010), Berlin, Germany, May?2010 [ref. 2] Christian Becker, Christian Bizer: Exploring the Geospatial Semantic Web with DBpedia Mobile. Journal of Web Semantics: Science, Services and?Agents on?the World Wide Web, Vol. 7, Pages 278iV286, 2009.
Overall Rating	Neutral

Masked Reviewer ID: Assigned_Reviewer_3
Review:

Question	
Detailed Comments	This proposal proposes demonstrating the facet generation techniques for Facetedpedia (proposed in [4]). The technique seems interesting and useful for wikipedia search and browsing. 1. In the proposed technique, why do we consider the categorization of the attribute articles, but not directly the target articles? 2. What can the demo show beyond the paper [4]?
Overall Rating	Accept