Abstract

The semantic web is a synergetic movement led by International standards body, the WWW Consortium (W3C). It aims at converting the current web dominated by unstructured and semi structured documents into a "web of data". Here two techniques of semantic web crawling are reviewed, one is ontology based and other is based on Lexical database. For this, architecture has been proposed which is a combination of above two techniques. The future of WWW is semantic web where Ontology and Lexical database are used for effective and fast searching by the web crawler. It is used for Information retrieval and question answering system. Ontology is a formal designation of shared approach it is basically approach of entities and their attributes.

References

Ontological Paradigm for Focused Crawling based on Lexical Analysis

Available: iteseer. ist. psu. edu/aggarwal01intelligent. html.
- Methods in "WordNet and their Application to Information Retrieval on the Web" 7th ACM.
- J. Lu, L. Ma, L. Zhang, C. Wang, Y. Pan, and Y. Yu, SOR: a practical system for Ontology storage, reasoning and search.

Index Terms

Computer Science
Information Systems
Keywords
Crawler  Semantic web  Ontology  Lexical database