Abstract

Finding meaningful information among the billions of information resources on the Web is a difficult task due to growing popularity of the Internet. The future of World Wide Web (WWW) is the Semantic Web, where ontologies are used to assign (agreed) meaning to the content of the Web. On the Semantic Web, data will inevitably be linked to many different ontologies, and information processing across ontologies is not possible without knowing the semantic mappings between them. As the resources on the Semantic Web are annotated using these ontologies, new search techniques are required to find specific information. For this, architecture has been proposed for ontology based semantic web crawler. This architecture can exploit the semantic metadata to efficiently discover and extract information from the Semantic Web. In this paper Semantic matching between content of downloaded web page and ontology is used to guide the crawler towards relevant information.

References

- M. Ehrig and A. Maedche, "Ontology-focused crawling of web documents,"
A Novel Architecture of Ontology-based Semantic Web Crawler

- Debajyoti, Arup Biswas, Sukanta &quot;A New Approach to Design Domain Specific Ontology Based Web Crawler&quot;, 10th International Conference on Information Technology – 2007 IEEE.
- Nigel Shadbolt, Wendey hall, Tim Berners-Lee &quot;the semantic web revisited&quot; IEEE intelligent system(2006).
- L. Page, S. Brin, R. Motwani, T. Winograd. &quot;The PageRank Citation Ranking: Bringing Order to the Web&quot;; Stanford Digital Library Technologies Project.

Index Terms

Computer Science
Keywords
Ontology  Semantic Web  Crawler