Abstract

Web is loaded with information and is getting overloaded with data each passing day. There needs to be efficient development in the area of information retrieval so the required data can be fetched accurately and efficiently. In this paper the two promising areas Natural Language Processing and Web technologies which can be combined together to enable the enterprise to combine the unstructured and structured data in ways that was not handled efficiently by traditional tools. The better understanding of web information can be done by integrating NLP and Web portal. The paper also explores various techniques used both areas. Also the NLP frameworks which can be used for the future work in this area.

References


10. Suma Adindla and Udo Kruschwitz, "Combining the Best of Two Worlds: NLP and IR for Intranet Search”, in 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology.


**Index Terms**

Computer Science

Information Systems

**Keywords**

NLP, information Retrieval, semantic Assistance, entity extraction, visual page segmentation, semi-markov conditional random fields, hierarchical conditional random field and Web Portal.