Abstract

The advance of technology has seen increase in applications that integrate new kinds of information, such as multimedia and scientific data, unstructured, semi-structured, structured or heterogeneous data being created and stored is exploding is collectively called “Dataspace”. Data being generated from various heterogeneous sources like, digital images, audio, video, online transactions, online social media, data from sensor nodes, click streams for different domains including, retail, medical, healthcare, energy, and day to day life utilities. Information Retrieval from heterogeneous information systems is required but challenging at the same as data is stored and represented in different data models in different information systems. Information integrated from heterogeneous data sources into single data source are faced upon by major challenge of information transformation were in different formats and constraints in data transformation are used in data integration for the purpose of integrating information systems, at the same is not cost effective.

Information retrieval from heterogeneous data sources remains a challenging issue, as the
number of data sources increases more intelligent retrieval techniques, focusing on information content and semantics, are required.

This paper describes the idea of Information retrieval system, Information integration which can be used in the Dataspace and heterogeneous data problems over the web.

References


Index Terms

Computer Science Information Sciences

Keywords

Information Retrieval; Dataspace; Information Retrieval Techniques; Heterogeneous Database;
Dataspace