Abstract

DEEP Web contains very large and valuable information than the surface Web. However, making use of such consolidated information requires substantial efforts since the pages are generated for visualization not for data exchange. Thus, Extracting structured data from deep Web pages is a challenging problem due to the underlying intricate structures of such pages. So, extracting information from searchable Websites has been a key step for Web information
integration. We discuss some of the underlying problems and issues central to extending information retrieval systems.

Reference

- http://www.tcp.ca/Jan96/BusandMark.html
  [formerlyhttp://www.tcp.ca/Jan96/BusandMark.html]
- The 1999 NEC study report on average Web document

Index Terms

Computer Science  Ubiquitous Computing

Key words

Deep Web  Surface web

Web-mining