Abstract

The increase number of manuscripts and their diversity add the difficulty of searching and arranging for relevant manuscripts. The quality of search results provided by search engines has not been maximized in response to user requests because it does not involve semantic elements in the search process. It is necessary to build an information retrieval system for manuscript that makes it easier for researchers finding the title of the manuscript accordance with the topic of their research.

Information retrieval system for manuscript is built using semantic web. Manuscript data used in this research are Indonesian manuscript. Stages build system include data crawler process, build ontologies, NLP process, SPARQL query representation process and indexing process.

Information retrieval system for Indonesian manuscript can display the title and link of manuscript based on the search sentence entered. Tests are conducted on 3 types of search sentences with recall and precision methods. The recall value indicates that the owned
manuscripts are returned 93.3% by information retrieval system. The precision value indicates that the results are returned 100% relevant by information retrieval system.

References


Index Terms

Computer Science
Information Sciences

Keywords
Manuscript, crawler, semantic web, ontology, NLP, SPARQL query