ABSTRACT

EKA MARLIANA. Information Retrieval for RSS News Document in Bahasa Indonesia. Supervised by FIRMAN ARDIANSYAH.

RSS (Really Simple Syndication) is a language derived from XML (Extensible Markup Language). The use of RSS as a syndication on Indonesian news sites has become widespread, as well as a syndicated news by news websites will continue to evolve in time, so it requires a search facility that can return information that explore the RSS data efficiently and effectively. Several studies have been conducted related to information retrieval, one of it was developed by Rahman (2006) which measured performance and compared the equality of returned XML document.

This research tries to implement the information retrieval using VSM (Vector Space Model) to build an RSS search facility and to analyze and compare the effects of the use of additional title weighting with normal weighting.

Test results show that the use of the normal weighting performs better than the use of weighting in the title. This is explained by the average precision value gotten from the test. At recall levels between 10% until 30% the average precision has the same value, at recall level 60% the average precision value of title weighting is higher than normal weighting, but between 40%, 50%, 70% until 100% the normal weighting precision is greater that of the title weighting.

Keywords: Information Retrieval, RSS, Vector Space Model.