

Abstract

Labels on the website used to represent the information contained on the website. To avoid confusion delivery of information by the owner of the website to the user, the website developers need to design a good label so that the label can represent information to be conveyed to the user the owner of the website [1].

One way in designing a website labeling system is to compare and study the labeling system used on the websites of competitors [1]. In this final project will be carried out on the labeling system on the university website by doing a comparison between the university website label.

The method used is to perform data extraction using scrapy to get the labels used on the website with the top 10 universities in the Webometrics Indonesia and Telkom University's website and then compare the label by calculating the similarity labels using methods Levenshtein Distance. Simplification is done by applying the rule number label. Label comparison results are then analyzed by comparison with the index on higher education rules related documents and perform expert review.

Results of the label on the data extraction 11 854 university websites generate labels after preprocessing. Comparison label may indicate the percentage of appearance of labels on each university's website as well as the entire university website, the number of labels that produced after the comparison process are 706 labels. Analysis of the results by comparison with an index higher education rules document shows the use of document rules of higher education may help to consider any labels that need to be displayed on the website of the university in Indonesia. Analysis of the results by the expert review showed 25 of 706 labels are not used for the label displayed on the university website.

The final conclusion of this study was the comparison of labels can be used as a recommended standard labeling system that can be used to build or repair the university's website in Indonesia by these data sorted into in accordance with the needs of each of the manufacturer's website or standards makers website in each of the universities.

Keywords: *Web Information Architecture, Systems Labeling, Label Comparison, Information Extraction, Levenshtein Distance*