

ABSTRAK

Penulisan karya tulis ini bertujuan untuk membantu para pelajar SMA untuk mendapatkan informasi mengenai Perguruan Tinggi dengan lebih mudah dengan dibuatnya sistem pencarian. Adapun yang menjadi latar belakang penulisan ini karena masih banyaknya pelajar SMA yang masih kesulitan dalam mencari informasi mengenai penerimaan mahasiswa baru. Walaupun pihak Universitas sudah memberikan informasi tentang PMB (penerimaan mahasiswa baru) di *website* mereka, namun tidak sedikit juga para pelajar yang tidak tahu *link website* tiap Universitas yang ada. Oleh karena diperlukan sistem pencarian untuk membantu para pelajar tersebut.

Dalam mendapatkan data mengenai PMB digunakan *web crawler* yang menjadi fokus utama dalam penelitian ini. Kemudian setelah data berhasil didapatkan maka proses selanjutnya adalah pembuatan sistem pemerolehan informasi menggunakan model pemerolehan *boolean*. Dalam penelitian ini, *web crawler* yang digunakan adalah *apache nutch* yang memiliki *setting depth* dimana *depth* ini merupakan iterasi pengambilan data tiap *website*. Kemudian *depth* tersebut akan dibandingkan berdasarkan data PMB dari suatu Universitas yang telah di *crawling*. Dan hasil yang didapat yaitu, pertama sistem pemerolehan informasi ini berhasil mencari data mengenai PMB dari Universitas yang ada di Yogyakarta. Yang kedua yaitu, dibutuhkan *setting depth 2* dan dengan *keyword "pmb"*, untuk berhasil mendapatkan data mengenai penerimaan mahasiswa baru dari 19 Universitas yang ada di Yogyakarta.

Kata kunci : sistem pemerolehan informasi, model pemerolehan boolean, *web crawler*, *apache nutch*, *depth*, penerimaan mahasiswa baru.

ABSTRACT

Writing this paper aims to help high school students to get information about the University more easily with the creation of a search system. As for the background of this writing because there are still many high school students who are still difficulty in finding information about the admission of new students. Although the university has provided information about the PMB (new admissions) on their website, but not a few also the students who do not know any university website links. Because the search system is needed to help the students.

In obtaining data about PMB used web crawler which became the main focus in this research. Then after the data was obtained then the next process is making information acquisition system using boolean acquisition model. In this research, the web crawler used is apache nutch which has depth setting where this depth is iteration of data retrieval each website. Then the depth will be compared based on PMB data from a University that has been in crawling. And the results obtained is, first system pemrolehan this information managed to find data about PMB from the University in Yogyakarta. The second is, it takes depth setting 2 and with keyword "pmb", to successfully get data about the admission of new students from 19 universities in Yogyakarta.

Keyword: information retrieval, boolean model, web crawler, apache nutch, depth, new students admission.