



ABSTRACT

PERSONALIZATION ON NEWS SEARCH ENGINE USING GOOGLE+ USER PROFILE DATA AND ELASTICSEARCH

Geiya Wihanadieptha Poetra

13/350347/PA/15637

Personalization has proven to be a useful method in helping users find relevant items and documents. More and more digital media implement some form of personalization to help users find related news. Even though there are a lot of work in the field of personalization, search engines are only personalized to a limited degree. Social media are also a widely used platform that has become an identity in the virtual world.

In this research we investigate on the possibility of building a near real-time personalized search engine with Google+ user profile data as the source of personalization using Elasticsearch. The author gather news articles data from *detik.com* and *okezone.com*. Search engine are using *Elasticsearch*, multiple sub fields are created for different analyzer, one is *Indonesian* analyzer which *stem* and remove *stopwords* from the news articles. Other sub field are using *search analyzer* that apply *synonym graph filter* that allow us to use multiple words synonym at search time.

The system then tested by 10 different users to measure the *precision*. From this research are concluded that creating a personalized search engine using Elasticsearch with Google+ user profile data are possible and could be a great starting user profile data for further personalization. The performance of the personalized search engine has measured and have score with *precision* are 0,79.

Keywords: Personalization, Search Engine, Elasticsearch, Google+, Profiling



INTISARI

PERSONALISASI PADA MESIN PENCARIAN BERITA MENGUNAKAN DATA PROFIL PENGGUNA GOOGLE+ DAN ELASTICSEARCH

Geiya Wihanadieptha Poetra

13/350347/PA/15637

Personalisasi sudah terbukti menjadi metode yang berguna dalam membantu pengguna dalam menemukan barang dan dokumen yang relevan. Semakin banyak media digital mengimplementasikan beberapa bentuk personalisasi untuk membantu pengguna menemukan berita yang berhubungan. Walaupun banyak pekerjaan dalam bidang personalisasi, mesin pencarian hanya dapat di personalisasi hingga tingkatan tertentu.

Media sosial juga sudah menjadi pelantar yang digunakan secara luas yang sudah menjadi identitas di dalam dunia virtual. Dalam penelitian ini kita menyelidiki atas kemungkinan untuk membangun personalisasi mesin pencarian yang mendekati waktu nyata dengan data profil pengguna Google+ sebagai sumber personalisasi menggunakan Elasticsearch. Penulis mengumpulkan data artikel berita dari *detik.com* dan *okezone.com*. Mesin pencarian menggunakan *Elasticsearch*, beberapa sub bidang dibuat untuk analyzer yang berbeda, salah satunya adalah *Indonesian analyzer* yang *stem* dan menghapus *stopwords* dari artikel berita. Sub bidang lainnya menggunakan *search analyzer* yang menerapkan *synonym graph filter* yang membuat kita dapat menggunakan synonym beberapa kata pada *search time*.

Sistem lalu di uji oleh 10 pengguna berbeda untuk mengukur *presisi*. Dalam penelitian ini dapat disimpulkan bahwa membuat personalisasi mesin pencarian menggunakan Elasticsearch dengan data profil pengguna Google+ adalah memungkinkan dan dapat menjadi data awal profil pengguna yang baik untuk personalisasi lebih lanjut. Kinerja dari personalisasi mesin pencarian sudah terukur dan mendapatkan nilai *presisi* adalah 0,79.

Kata kunci: Personalization, Search Engine, Elasticsearch, Google+, Profiling