

## ABSTRAK

Fitriyono, Aji 2016. *Pencarian Produk Dengan Teknik Web Scraping Berbasis Website*. Tugas Akhir. Program Studi Teknik Informatika (S1). STIKI Malang. Pembimbing: Koko Wahyu P, S.Kom, M.T.I, Co. Pembimbing : Sugeng Widodo, M.Kom.

Kata Kunci: Internet, *Search Engine*, *Web Scraping*, HTML DOM, Datatable.

Saat ini dunia teknologi informasi telah berkembang pesat khususnya pada pemanfaatan internet. Dengan adanya internet akan mempermudah dan mempercepat proses pencarian dan penyebaran informasi. Hal tersebut dimanfaatkan sebagai penyebaran informasi produk pada situs jual beli.

Perlu adanya sebuah *search engine* yang dikhususkan hanya mencari informasi produk saja. Proses pengambilan informasi produk dari *website* disebut juga *web scraping*. Pengambilan ditujukan pada situs jual beli yang memiliki fitur pencarian dan hasil pencarian terbuka tanpa di acak. Dengan memanfaatkan teknik *regular expression (regex)* sebagai alat untuk menentukan pola yang mengawali dan mengakhiri suatu konten utama pada halaman *situs*, serta menggunakan HTML DOM (Document Object Model) sebagai alat pengambilan *tag* atau *elemen* HTML tertentu saja. Kedua teknik tersebut sangat membantu dalam menemukan konten yang berisi informasi produk. Konten berisi pola-pola tag HTML yang mengapit informasi produk. Setiap tag akan di pilih *id* dan *class* yang unik untuk di tentukan blok dan baris konten informasi produk. Sehingga percobaan terhadap semua kemungkinan konten tidak perlu dilakukan. Selanjutnya hasil *scraping* ditampung kedalam bentuk *informasitable*. Penggunaan *Jquery Datatable* di pilih karena memiliki fitur lengkap dalam menginformasikan data berupa *table*.

*Search Engine* Pencari Produk diharapkan dapat membantu masyarakat atau pengguna internet untuk mencari produk dengan mudah, cepat dan efisien tanpa harus mengunjungi satu per satu situs. Diharapkan dapat menginformasikan sebagian besar dari informasi/kategori produk seperti nama, gambar, harga, penjual, kondisi, keterangan, dan situs yang di ambil datanya. Ketiga situs tersebut adalah situs *jualo.com*, *bukalapak.com*, dan *bhineka.com*.