

ABSTRAK

Metode *Web Scraping* memiliki banyak metode beragam. Dalam melakukan *Web Scraping* yang menjadi fokus utama adalah mengidentifikasi *website* tersebut apakah merupakan *website* statis atau dinamis dan juga struktur data dari *website* tersebut. Dengan banyaknya metode *Web Scraping*, ini membuat pengguna cukup kesulitan untuk memilih metode yang sesuai dengan kebutuhan, terutama di zaman sekarang di mana kebutuhan akan informasi sangat diperlukan dengan cepat dan tepat.

Tugas akhir ini bertujuan untuk membandingkan beberapa metode dalam melakukan *scraping* data dari beberapa karakteristik *web* yang berbeda dan melakukan perbandingan di beberapa parameter untuk mencari metode yang sesuai dan tepat dalam menangani tiap karakteristik *website* tersebut.

Hasil penelitian ini mengambil data pada Simple Web sebagai sampel untuk *website* statis dan Kompas dan Tokopedia sebagai sampel untuk *website* dinamis dengan pengujian sebanyak 20 kali pada setiap parameternya. Pada penelitian kali ini menggunakan 3 metode yaitu *Regular Expression*, *Parsing DOM* dan *XPATH*. Pada penelitian ini dengan menggunakan metode *Regular Expression* dan mendapatkan rata-rata penggunaan *CPU* sebesar 3,57%, *Memory* sebesar 5,69 KB dan waktu yang dibutuhkan 1,94s. Dengan metode *Parsing DOM* rata-rata penggunaan *CPU* sebesar 2,63%, *Memory* sebesar 388,04 KB dan waktu yang dibutuhkan 2,76s. Kemudian dengan menggunakan metode *XPATH* rata-rata penggunaan *CPU* sebesar 2,9%, *Memory* sebesar 1,41 KB dan waktu yang dibutuhkan 2,32 s.

Kata kunci: *Web Scraping, Parsing DOM, XPath, Regular Expression, Dynamic Web Content*