

1. Pendahuluan

Latar Belakang

Di era perkembangan teknologi saat ini, dengan kemajuan dan meningkatnya kebutuhan informasi dalam masyarakat, terutama dengan pesatnya perkembangan teknologi internet. Internet telah menjadi sumber informasi yang penting bagi masyarakat. Informasi tersebut baik berupa informasi teknologi, sosial, pendidikan, lowongan kerja dan sebagainya. Selain itu, dengan adanya teknologi seperti smartphone, komputer dan tablet, pengguna dapat menggunakan layanan internet dengan nyaman dan dapat digunakan dimana saja.

Mengenai informasi lowongan kerja saat ini tidak hanya didapatkan pada koran, brosur atau mendatangi perusahaan, tetapi informasi dapat ditemukan melalui internet. Dalam internet terdapat banyak web pihak ketiga yang menyediakan sebuah informasi lowongan kerja dari berbagai perusahaan. Hal ini menjadi salah satu cara paling umum untuk dilakukan para pencari kerja secara online. *Web* pihak ketiga sangat membantu memudahkan para pencari kerja untuk mendapatkan informasi lowongan kerja. Namun beberapa web pihak ketiga terkadang memiliki informasi yang berbeda karena perusahaan pemilik lowongan tidak memberikan informasi lowongan kerja perusahaannya kepada semua *web* pihak ketiga, sehingga membuat para pencari kerja mengunjungi satu persatu *web* untuk mendapatkan informasi yang diinginkan, hal ini tentu membutuhkan waktu yang lama.

Teknik *Web Scraping* adalah suatu proses dalam pengambilan sebuah dokumen semi-terstruktur dari internet, umumnya berupa halaman-halaman *web* yang ditulis dalam bahasa markup HTML atau XHTML, dan menganalisis dokumen tersebut untuk mengambil data tertentu dari halaman tersebut yang akan digunakan untuk kepentingan lain [1].

Beberapa penelitian terkait mengenai implementasi teknik *Web Scraping* telah dilakukan sebelumnya antara lain: implementasi *Web Scraping* untuk mengambil data artikel berita pada tiga *web*, metode yang digunakan dalam penelitian ini adalah *Regular Expression* untuk mengenali elemen yang dicari pada *web* berita [2], implementasi *Web Scraping* dengan metode HTML DOM untuk mendownload data publikasi artikel ilmiah dari Google Scholar berdasarkan *id* atau *class* yang terdapat pada *source code web* Google Scholar [3], implementasi *Web Scraping* untuk pengambilan data pada situs *Marketplace* [4]. Terdapat beberapa metode yang dapat digunakan dalam implementasi *Web Scraping* seperti *Regular Expression*, HTML DOM, dan XPath [5]. Pada penelitian ini, penulis menggunakan *Web Scraping* untuk mengumpulkan data mengenai informasi lowongan kerja antara lain nama pekerjaan, nama perusahaan, logo perusahaan, lokasi, persyaratan, tanggal posting, dan alamat URL dari web pihak ketiga yaitu Kalibr, Jobstreet, dan Glints. Metode yang digunakan dalam *Web Scraping* pada penelitian ini adalah *Regular Expression* yang digunakan untuk mengenali elemen class tersebut. Data yang telah berhasil didapatkan disimpan dalam *database*, kemudian ditampilkan pada sistem informasi berbasis *web* yang dibuat.

Topik dan Batasannya

Topik yang dibahas dalam tugas akhir ini adalah bagaimana implementasi *Web Scraping* dengan metode *Regular Expression* untuk mendapatkan data mengenai informasi lowongan kerja dari masing-masing *web* pihak ketiga lowongan kerja dan bagaimana pengembangan sistem informasi berbasis *web* untuk menampilkan data yang didapatkan dari *Web Scraping*. Batasan masalah dalam penelitian ini yaitu proses *Web Scraping* dilakukan pada *web* pihak ketiga lowongan kerja antara lain Kalibr, Jobstreet, dan Glints. Pengambilan data berupa nama pekerjaan, nama perusahaan, logo perusahaan, lokasi, persyaratan, tanggal posting, dan alamat URL lowongan. Proses *Web Scraping* dilakukan dalam tiga hari sekali.

Tujuan

Tujuan tugas akhir ini adalah untuk mengetahui implementasi teknik *Web Scraping* dengan metode *Regular Expression* untuk mendapatkan data mengenai informasi lowongan kerja dari masing-masing *web* pihak ketiga lowongan kerja dan mengembangkan sistem informasi berbasis *web* untuk menampilkan data yang didapatkan dari *Web Scraping*.

Organisasi Tulisan

Organisasi tulisan setelah pendahuluan pada penelitian ini yaitu studi terkait mengenai web lowongan kerja, web scraping, Python, *Regular Expression* dan Django Framework. Pada bagian sistem yang dibangun menjelaskan metodologi pada penelitian ini.