

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Saat ini, tingkat kepopuleran penggunaan *website* sebagai platform utama sangat meningkat, hal ini dikarenakan sifat *website* yang sangat fleksibel dan mudah di gunakan di berbagai platform secara instan. Berbeda dengan aplikasi *smartphone* ataupun *software* komputer/laptop yang pada saat penggunaan awal harus dilakukan proses instalasi, *website* dapat langsung digunakan tanpa perlu membutuhkan proses ini dan hanya perlu mengetikkan alamat situs, situs web langsung dapat digunakan.

Namun, seiring peningkatan kepopulerannya ini, *website* bisa dapat disalahgunakan dengan mudah. Salah satu contoh penyalahgunaan *website* yang perlu diperhatikan adalah adanya oknum-oknum yang memberikan informasi rahasia perusahaan, pemerintah serta informasi sensitif lainnya secara publik. Perilaku ini dapat berdampak buruk bagi perusahaan terkait, pemerintah atau bahkan secara individu.

Salah satu solusi yang dapat digunakan untuk menghadapi masalah seperti ini adalah penggunaan web *crawler*. Menurut Wikipedia, penjelasan web *crawler* adalah “Web *Crawler* atau sering juga di sebut *spiders* adalah sebuah *tool* untuk mengindeks dan mengunduh konten dari internet, lalu di simpan ke dalam *database* mesin pencari ”. Sudah ada banyak metode web *Crawling* yang beredar saat ini, terutama untuk penggunaannya di situs web konvensional. Namun penyebaran data sensitif ke publik tidak hanya dilakukan di situs web konvensional saja. Hal ini juga terjadi di web gelap atau lebih sering dikenal dengan *dark web*. Sayangnya belum banyak metode *web crawling* yang di sediakan untuk penggunaan di *dark web* mengingat cara akses *dark web* yang cukup sulit dan membutuhkan metode yang berbeda dengan *dark web* konvensional.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka rumusan masalah pada penelitian ini adalah sebagai berikut :

- a. Dapatkah metode *in-depth website crawl* ini diimplementasikan pada *darkweb*?
- b. Seberapa besar persentase akurasi dan performa penggunaan metode *in-depth website crawl* dalam mengumpulkan URL dari *dark web*?

1.3 Tujuan dan Manfaat

Penelitian ini bertujuan untuk:

- a. Di dapatkannya sebuah program *crawling* yang dapat beroperasi di *dark web*.
- b. Di dapatkan program *crawling darkweb* dengan nilai akurasi lebih besar (>) dari 80% dan nilai performa *crawling* untuk 1 URL lebih kecil (<) dari 5 detik.

1.4 Batasan Masalah

Adapun Batasan masalah pada tugas akhir ini yaitu :

- a. Halaman situs yang dijadikan objek pada penelitian ini adalah halaman situs yang termasuk ke dalam kategori *dark web*.
- b. *Crawler* yang buat hanya memiliki tujuan untuk mengumpulkan *database* berupa tautan dan tidak akan melakukan klasifikasi.
- c. Pada tugas akhir ini tidak akan dibahas mengenai performansi jaringan.

1.5 Metode Penelitian

Metode yang digunakan pada pengerjaan Tugas Akhir ini yaitu:

1. Studi Literatur

Studi literatur dilakukan untuk mendapatkan informasi mendalam mengenai permasalahan yang ditemui.

2. Bimbingan dan diskusi

Bimbingan dan diskusi dilakukan dengan penjelasan progres dengan pembimbing dan melalui bimbingan mingguan dan *chat* WhatsApp.

3. Penyusunan Buku Tugas Akhir

Penyusunan buku Tugas Akhir ini merupakan dokumentasi dari apa saja yang telah dilakukan selama pengerjaan Tugas Akhir. Buku ini ditujukan untuk memudahkan orang lain membaca dan mengembangkannya.

1.6 Ringkasan Sistematika Penulisan

Ringkasan sistematika laporan yaitu sebagai berikut:

Bab I Pendahuluan

Pada bab ini dikemukakan latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, metode penelitian, dan rincian sistematika penulisan;

Bab II Tinjauan Pustaka

Pada bab ini berisi tentang teori - teori yang digunakan dalam pengerjaan Tugas Akhir;

Bab III Perancangan Sistem

Bab ini menjelaskan mengenai perancangan sistem yang dibuat untuk memenuhi tujuan Tugas Akhir;

Bab IV Hasil dan Analisis

Pada bab ini berisi tentang hasil pengujian dan analisis pengujian;

Bab V Simpulan dan Saran

Pada bab ini berisi tentang simpulan pengerjaan Tugas Akhir serta saran untuk penelitian selanjutnya.