

ABSTRAK

Pada saat ini semakin banyaknya terjadi kasus transaksi barang ilegal dan data pribadi yang mengalami kebocoran. Transaksi ilegal dan penjualan data pribadi biasanya dilakukan pada *dark web* karena web tersebut memiliki enkripsi yang berlapis dan adanya sistem anonim pada saat mengaksesnya. Karena adanya jaringan yang berlapis-lapis tersebut membuat sulitnya untuk mengakses dan mengetahui konten pada *dark web*. Dengan menggunakan metode *crawling* dan juga TOR dapat untuk menembus dan menelusuri konten di *dark web*.

Metode perayapan di *dark web* dapat menggunakan *focused crawling* yang menggunakan pendekatan URL dengan melihat URL yang saling berhubungan dengan halaman URL utama pada topik yang sesuai dengan kata kunci yang dimasukkan. Pada penelitian-penelitian sebelumnya telah dilakukan *focused crawling* pada web biasa dengan mendapatkan hasil yang lebih maksimal dibandingkan mesin pencari umumnya.

Dari penelitian yang telah dilakukan sebelumnya dilakukan pengembangan dengan melakukan *crawling* pada *dark web* untuk memaksimalkan pencarian konten tidak terbatas di web permukaan saja. Hasil dari penelitian ini yaitu *focused crawler* ini dengan menggunakan metode *Regex Text* pada *keyword* ‘cocain’ didapatkan nilai akurasi yaitu 90,95% dengan waktu 15,82 detik dan pada *keyword* ‘hack’ didapatkan dengan akurasi 97.67% dengan waktu 16,03 detik. Selanjutnya pada proses *crawling* menggunakan *Regex Wildcard* pada *keyword* ‘cocain’ didapatkan nilai akurasi sistem sebesar 98,61% dengan waktu 16,3 detik dan pada *keyword* ‘hack’ dengan akurasi sebesar 99.67% dengan waktu 15,09 detik. Kemudian akurasi menggunakan *Regex Optional* yaitu pada *keyword* ‘cocain’ sebesar 98,61% dengan waktu 14,71 detik dan pada *keyword* ‘hack’ sebesar 99.67% dengan waktu 15.61 detik. Sistem *focused crawling* ini mampu melakukan fungsionalitasnya di *dark web*.

Kata Kunci: *Focused crawling, dark web, Tor, URL.*