Timothius T / 215116485

Tropical web crawl adalah salah satu cara web crawl yang menjelajahi web dengan berfokus pada suatu topik. Topical web crawl berguna untuk mengambil informasi dengan cara web crawl dan menyaringnya sebelum link dibuka untuk mempersingkat waktu dan mendapat informasi yang sesuai dengan kebutuhan.

Topical web crawl juga dapat menjadi referensi yang biasanya dilihat pada google ads ataupun ads engine yang lain. Seperti saat kita mencari kota Surabaya maka ads juga dapat menampilkan yang ada pada Surabaya. Lalu pada news google juga bisa menampilkan banyak web yang memiliki news sama.

Salah satu cara yang dapat dilakukan dengan mencari anchor text dari link yang dituju tersebut. Contohnya bila mencari informasi tentang crawling maka anchor text dari link tersebut kemungkinan mengandung kata “crawling” seperti anchor text ‘temukan cara crawling disini’ lalu kemudian akan mengarah ke suatu link yang tidak mengandung kata crawling tetapi terdapat informasi mengenai crawling.

Ada juga cara dengan membuka link itu sendiri dan melihat isi web content yang ada pada halaman tersebut lalu dicocokkan dengan topik yang sedang dicari. Topik tersebut tidak hanya mencakup informasi saja tetapi juga bisa seperti “PageRank yang tinggi”

Terkadang anchor text tidak mengandung kata yang diinginkan. Di saat itulah link context dapat digunakan. Link context dapat memberikan patokan yang lebih efisien dari anchor text. Cara meng-ekstrak link context yang paling mudah adalah mengambil tulisan atau paragraf di sekitarnya.

Terdapat pula banyak cara yang lebih efisien dengan menggunakan resource eksternal. Sebagai contoh menggunakan web directory, web text index, backlink, atau berbagai jenis sumber web lainnya.

Cara-cara tersebut tidak menutup kemungkinan jika information retrieval dapat dilakukan tanpa membuka link tersebut dan hanya mengandalkan koneksi pada website tersebut.