

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi saat ini sudah berkembang dengan sangat pesat. Pemanfaatannya sudah mencakup di berbagai bidang kehidupan. Teknologi informasi saat ini sudah banyak memberikan kemudahan. Salah satunya dalam penggunaan mesin pencarian atau *Search Engine*. Pemanfaatan mesin pencarian dalam menyaring informasi sangat dibutuhkan saat ini, mengingat banyaknya informasi yang telah beredar di internet. Sebagaimana mekanisme mesin pencarian ini akan menyaring informasi yang dibutuhkan oleh pengguna berdasarkan kata kunci atau *keyword* yang digunakan, tentu ini akan sangat bermanfaat untuk mendukung pengguna dalam mendapatkan informasi yang valid dan akurat di internet. Oleh karena itu mesin pencarian merupakan salah satu alat bantu yang sangat dibutuhkan oleh pemerintahan untuk memudahkan warganya dalam mengakses dan menyaring informasi yang bersumber langsung dari mereka di internet.

Perkembangan teknologi informasi dan komunikasi ini menjadikan pelaksanaan pemerintah pada era yang dinamakan *e-government*, dimana diperlukan adanya keterbukaan informasi antara pemerintah dengan masyarakat untuk mewujudkan adanya keterbukaan publik. *E-government* atau *electronic government* merupakan penggunaan teknologi informasi oleh pemerintah untuk memberikan informasi dan pelayanan publik (Lenak dkk, 2021). Informasi yang disediakan oleh pemerintah ini salah satunya adalah dalam media website, sehingga pemerintahan Sumatera Barat mempunyai banyak website yang aktif dalam memberikan informasi yang berguna untuk masyarakatnya. Akan tetapi informasi ini tidak bisa

tersampaikan dengan baik karena banyaknya sumber informasi yang terpisah berdasarkan masing-masing website dinas yang mengeluarkan informasi tersebut.

Dalam mesin penelusuran dibutuhkan pengumpulan beberapa sumber website yang akan dijadikan acuan dalam mencari informasi yang dicari oleh pengguna melalui keyword yang dimasukkan kedalam kolom pencarian. Teknik yang digunakan dalam mengumpulkan konten-konten website yang dibutuhkan ini adalah dengan membuat sebuah *Web Crawler*. *Crawler* adalah program yang menelusuri World Wide Web dengan cara yang metodis, otomatis dan teratur. Istilah lain untuk web *crawler* adalah ant, automatic indexer, bots, web spiders atau web robots (Halim dkk, 2017). Web *crawler* dapat dijadikan sebuah alat untuk mengambil konten-konten yang ditampilkan oleh halaman website, kemudian konten-konten tersebut dikelompokkan dalam satu atau lebih atribut (Ilmawan, L. B, 2018).

Untuk melakukan pencarian konten yang berhubungan dengan kata kunci yang dicari maka digunakan metode *regular expression* atau *Regex*. *Regex* merupakan sebuah konstruksi bahasa yang mencocokkan teks berdasarkan pola-pola tertentu, terutama untuk kasus yang kompleks (Astuti dkk, 2018). Melakukan pencarian dengan metode *Regex* akan lebih menghemat waktu dan memori komputer. *Regular search expression* (REGEX) adalah metode atau algoritma pencarian string yang memiliki konsep dasar pencarian bukan hanya kepada string secara spesifik, namun melakukan pencarian berdasarkan pola dari string tersebut (Nagpure, 2019). Sedangkan Metode pencarian lainnya yang adalah *Binary Search*, yaitu sebuah metode atau algoritma fungsi pencarian yang dirancang untuk berjalan dengan konsep memanfaatkan array dari biner yang sudah tersusun secara urut (Syahputra & Sinurat, 2016). *Binary search* merupakan salah satu algoritma yang cukup efisien, sehingga membuat algoritma tersebut menjadi algoritma pencarian yang populer dan sering digunakan. Namun metode *Binary search* ini mempunyai kekurangan yaitu data harus diurutkan atau disorting terlebih dahulu. Hal ini menjadikan Algoritma Binary Search tidak baik untuk melakukan pencarian terhadap data berantai.

Selain untuk melakukan pencarian, dalam penelitian ini penggunaan *Regex* juga ditujukan untuk membersihkan konten yang telah di kumpulkan oleh *crawler* dari tag-tag html ataupun simbol-simbol yang tidak dibutuhkan. Informasi yang telah di ekstrak akan dicocokkan sesuai pattern yang telah ditentukan menggunakan *Regex* (Dewi dkk, 2019).

Dari latar belakang diatas dapat disimpulkan bahwa Pemerintahan Propinsi Sumatera Barat membutuhkan sebuah mesin penelusuran yang akan mengumpulkan semua konten informasi dari semua website pemerintah propinsi Sumatera Barat dan memudahkan masyarakat dalam mencari informasi yang mereka butuhkan dengan mudah. Untuk itu penulis melakukan penelitian sebagai Tugas Akhir (Tesis) dengan judul “Global Search Engine dalam Mengindeks Website Pemerintahan Propinsi Sumatera Barat Menggunakan Metode Regex.”

1.2 Perumusan Masalah

Berdasarkan latar belakang diatas dapat ditemukan beberapa permasalahan seperti berikut :

1. Bagaimana merancang sebuah sistem yang dapat mengumpulkan semua konten informasi dari berbagai website pemerintahan propinsi Sumatera Barat dan mengindeksnya di dalam sebuah database.
2. Bagaimana menerapkan metode regex dalam mempercepat dan mempermudah dalam melakukan pencarian konten yang berhubungan dengan kata kunci yang dimasukan secara akurat.

1.3 Batasan Masalah

Agar penelitian ini sesuai dan lebih terarah dalam proses pengumpulan data maupun informasi yang dibutuhkan, maka penulis menetapkan batasan-batasan masalah yang akan diteliti, yaitu:

1. Penelitian ini menggunakan metode menggunakan metode *Regex (Regular Expression)* dalam melakukan pencarian.
2. Data yang digunakan untuk penelitian ini memakai data semua website yang dikelola Dinas Komunikasi dan Informatika Provinsi Sumatera Barat.

1.4 Tujuan Penelitian

Adapun beberapa tujuan penelitian yang hendak dicapai diantaranya sebagai berikut:

1. Merancang mesin pencarian yang akan membantu pemerintah dan masyarakat provinsi Sumatera Barat dalam memberi dan menerima informasi yang berguna.
2. Menerapkan metode *Regex* atau *Regular Expression* dalam mengumpulkan konten-konten dari semua website pemerintahan provinsi Sumatera Barat dan dalam melakukan metode pencarian ke dalam konten website yang telah dikumpulkan.

1.5 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian Implementasi *Metode Regular Expression* (*Regex*) ini adalah:

1. Sebagai sarana bagi penulis untuk menerapkan dan mengembangkan wawasan dibidang ilmu pengetahuan komputer. Kemudian mendalami metode metode *Regular Expression* (*Regex*), sehingga penulis mampu untuk mengembangkan dan memanfaatkan dalam pembuatan mesin pencarian.
2. Penelitian ini diharapkan membantu Dinas Komunikasi dan Informatika provinsi Sumatera Barat dalam mengindeks semua website yang di kelola.

1.6 Sistematika Penulisan

Sistematika yang digunakan dalam penyusunan thesis ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisi latar belakang, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penelitian.

BAB II. LANDASAN TEORI

Pada bab ini dijelaskan tentang landasan teori yang menjabarkan tentang teori-teori yang berkaitan dengan metode dan sesuai dengan judul penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini membahas tentang analisa dan penggunaan secara matematis metode *Regex*.

BAB IV ANALISA DAN PERANCANGAN

Bab ini membahas tentang analisa permasalahan, analisa kebutuhan sistem, menerapkan metode *Regex*.

BAB V IMPLEMENTASI SISTEM

Bab ini membahas hasil implementasi Mesin pencarian menggunakan metode *Regex*.

BAB VI KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dan hasil penelitian sistem pencarian menggunakan metode *Regex* pada website pemerintahan provinsi Sumatera Barat dan memberikan saran untuk penelitian berikutnya.