

BAB I

PENDAHULUAN

1.1.Latar belakang

Perkembangan teknologi dan informasi semakin hari kian berkembang dengan pesat, khususnya perkembangan internet. Internet kini menjadi konsumsi harian bagi kita. Internet semakin diandalkan dalam berbagai hal seperti, komunikasi, media informasi, bisnis, dan masih banyak hal yang dilakukan dengan menggunakan internet.

Informasi yang tersebar di internet terdiri dari berbagai dokumen baik berupa tulisan maupun lisan. Dokumen tersebut dapat digunakan sebagai bahan penelitian *analitic linguistic*. Penelitian *analitic linguistic* merupakan penelitian yang merujuk pada bahasa sebagai bahan penelitian. Sekumpulan dokumen baik berupa tulisan maupun lisan yang disimpan dan diproses di dalam komputer untuk tujuan penyelidikan linguistik disebut dengan korpus. (Renouf, A. 1988).

Seiring dengan peningkatan jumlah informasi, kebutuhan perangkat pencarian informasi yang efektif dan efisien sangat diperlukan, sehingga pencarian akan informasi didapat dengan lebih cepat dan akurat. Untuk menunjang kebutuhan yang semakin banyak tersebut, dibutuhkan *analitic linguistic* yang banyak pula. Namun sayangnya di Indonesia masih sedikit penelitian terkait *analitic linguistic*.

Penelitian mengenai *analitic linguistic* memerlukan waktu yang sangat lama, dengan tahapan yang sangat banyak diantaranya pengumpulan kata, pelabelan. Oleh karena itu penulis ingin menyediakan alat *tagging* yang dapat digunakan sebagai pelabelan terhadap korpus Bahasa Indonesia.

Pembuatan alat *tagging* korpus Bahasa Indonesia ini menjadi sangat penting. Hal ini menjadi sangat penting karena dengan semakin banyaknya penelitian mengenai *analitic linguistic* maka akan menambah database korpus yang telah diberi label. Yang pada akhirnya dapat memperbaiki sistem alat penerjemah pada Bahasa Indonesia.

Penelitian *alanitic linguistic* pada penulisan ini bertujuan untuk menyediakan data agar computer dapat mengerti akan struktur bahasa Indonesia. Pada prosesnya terdapat beberapa tahapan yakni, *Crawling, Tokenisasi, Parsing* dan *Tagging*.

Pada penelitian ini penulis memfokuskan pada *tagging* korpus. Penulis memilih fokus pada proses *tagging* karena di Indonesia masih sangat jarang penelitian mengenai *tagging*.

Korpus Bahasa Indonesia dapat menjadi sebuah *golden* korpus ketika proses *tagging* yang dilakukan sudah banyak dan dapat berkembang secara otomatis dengan didukung *data mining*.

Penulis mengharapkan dengan penelitian ini dapat mempermudah dan mempercepat bagi peneliti untuk membuat sebuah *data mining* dan akan menghasilkan sebuah *golden* korpus dalam Bahasa Indonesia.

1.2.Perumusan Masalah

Berdasarkan latar belakang yang penulis sampaikan dapat dirumuskan permasalahan sebagai berikut:

- Bagaimana membangun alat *tagging* korpus Bahasa Indonesia?

1.3.Tujuan dan Manfaat Penelitian

Tujuan yang penulis harapkan dari penelitian ini yaitu:

- Terbentuknya alat *tagging* untuk korpus Bahasa Indonesia

Manfaat yang di dapat dari penelitian ini antara lain:

- Terbentuknya *data mining* korpus Bahasa Indonesia
- Terbentuknya *golden* korpus Bahasa Indonesia

1.4.Batasan Masalah

Dalam sebuah penelitian diperlukan adanya batasan, sehingga penelitian yang dilakukan lebih terarah dan fokus pada tujuan, maka dari itu berikut ini beberapa batasan, diantaranya:

- Aplikasi yang dibangun berbasis web

- Aplikasi yang di bangun hanya fokus sebagai alat input (*tagging*) yang bersifat manual.
- Aplikasi ini menggunakan database MySQL dan NOSQL.
- *Tagset* yang digunakan dalam aplikasi ini menggunakan tagset yang disusun oleh Dinakaramani.

1.5.Sistematika Penulisan

Penyusunan tugas akhir ini mengikuti sistematika sebagai berikut:

Bab I : Pendahuluan

Bab ini berisikan latar belakang dari penulisan proposal tugas akhir, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, dan sistematika dari penulisan proposal tugas akhir ini

Bab II : Landasan Teori

Bab ini berisikan pembahasan teori berupa materi yang dapat penulis gunakan sebagai pendukung penelitian yang penulis lakukan. Pada bab ini pula berisikan penelitian terkait dan posisi penelitian.

Bab III : Metodologi Penelitian

Bab ini berisikan tahapan yang akan dilakukan dalam melakukan penelitian diantaranya tahapan pengembangan system serta rancangan penelitian. Rancangan penelitian menjelaskan mengenai cara menyelesaikan permasalahan, jenis penelitian yang dilakukan, rancangan analisis dan pengumpulan data, tempat lokasi serta waktu dilaksanakannya penelitian.

Bab IV : Analisis dan Perancangan Sistem

Bab ini berisikan hasil analisis yang dilakukan oleh penulis untuk menunjang kebutuhan system aplikasi kemudian dilanjutkan dengan perancangan system aplikasi yang akan dibangun pada penelitian ini.

Bab V : Implementasi dan Testing

Bab ini berisikan implementasi yang dilakukan oleh penulis untuk membangun kebutuhan sistem aplikasi kemudian dilanjutkan

dengan pengujian sistem aplikasi yang akan dibangun pada penelitian ini.

Bab VI : Kesimpulan dan Saran

Bab ini berisikan kesimpulan dari penelitian yang dilakukan oleh penulis dan kemudian dilanjutkan saran untuk penelitian selanjutnya.



STT - NF