BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Penelitian yang mengangkat masalah-masalah yang berkaitan dengan pemahaman suatu teks sudah banyak dilakukan dengan menggunakan pembangkit pertanyaan atau *question generation* sebagai solusinya [11] [12] [13] [14] [15]. Penelitian yang telah dilakukan sudah menggunakan teks berbahasa Inggris maupun bahasa Indonesia dengan metode-metode yang berbeda. Untuk penelitian menggunakan teks berbahasa Inggris, metode yang dilakukan adalah diantaranya menggunakan transformasi sintaksis dan metode statistik [11] [12]. Sedangkan penelitian yang menggunakan teks berbahasa Indonesia menggunakan metode analisis semantik dan berbasis *template* [13] [15]. Selain itu, penelitian pada teks berbahasa Indonesia dilakukan secara spesifik untuk kasus dalam bidang kesehatan [14].

Salah satu yang telah disebutkan pada penelitian teks berbahasa Indonesia adalah penelitian menggunakan metode berbasis *template* [15]. Penelitian ini menggunakan metode berbasis *template* dengan cara mencari kata khusus yang telah ditentukan untuk membangkitkan pertanyaan. Kata khusus menggambarkan kalimat jawaban untuk suatu kategori pertanyaan. Kategori pertanyaan yang dimaksud adalah definisi, alasan, dan metode. Dengan menggunakan cara ini, penelitian tersebut dapat menciptakan kalimat tanya *non-factoid* dengan kata tanya *apa yang dimaksud dengan*, *mengapa*, dan *bagaimana*.

Dari penelitian di atas, diketahui bahwa kalimat pertanyaan yang dapat dibangkitkan adalah kalimat *non-factoid*. Tetapi masih ada kata tanya yang belum teratasi seperti siapa, apa, mana, dan kapan. Kata-kata tanya tersebut dibutuhkan untuk mengatasi wacana atau teks *factoid*. Dengan menggunakan metode berbasis *template* seperti penelitian sebelumnya, penelitian ini akan membangkitkan kata-kata tanya tersebut dengan menggunakan analisis sintaksis. Analisis sintaksis bertujuan untuk mencari fungsi-fungsi sintaksis seperti subjek, predikat, objek, dan keterangan. Fungsi sintaksis dibutuhkan agar kata tanya siapa, apa, mana, dan

kapan dapat dibangkitkan karena kata tanya tersebut menanyakan fungsi sintaksis dalam suatu kalimat. Contohnya kata tanya siapa yang menanyakan subjek atau objek dalam suatu kalimat. Oleh karena itu, kata tanya siapa, apa, mana, dan kapan, serta mengapa dan bagaimana, dapat dibangkitkan dengan melakukan analisis sintaksis untuk mengetahui fungsi sintaksis dalam suatu kalimat. Hasil pertanyaan yang dibangkitkan perlu diuji dan diukur relevansinya dengan teks masukan untuk mengetahui akurasi dari metode yang digunakan.

1.2 Identifikasi Masalah

Identifikasi masalah berdasarkan yang telah dijelaskan pada latar belakang adalah dibutuhkannya pengukuran relevansi hasil pertanyaan terhadap teks masukan.

1.3 Maksud

Maksud dari penelitian ini yaitu membangun pembangkit pertanyaan otomatis.

1.4 Tujuan

Tujuan dari penelitian ini adalah mengukur relevansi hasil pembangkitan pertanyaan terhadap teks masukan.

1.5 Batasan Masalah

Batasan masalah dari aplikasi yang akan dibangun adalah sebagai berikut:

1. Masukan
   a. Masukan berupa file berisi teks bahasa Indonesia dalam bentuk bahasa baku.
   b. Ekstensi file masukan adalah .txt.
   c. Teks bahasa Indonesia yang menjadi masukan tidak memiliki simbol-simbol seperti simbol matematis, *emoticon*, dan sebagainya.
   d. Teks bahasa Indonesia yang menjadi masukan berupa teks soal cerita untuk pelajaran Bahasa Indonesia kelas 1 Sekolah Dasar.

2. Proses : Predikat yang dapat ditangani hanya berupa kata kerja.
3. Keluaran : Kata tanya yang dibangkitkan adalah siapa, apa, kapan, mana, mengapa, dan bagaimana.
4. Sumber daya : Basis data kata dasar dan jenis kata dalam Bahasa Indonesia berasal dari Kamus Online Kateglo.

1.6 Metodologi Penelitian

Metodologi penelitian yang digunakan adalah metodologi eksperimen. Eksperimen adalah suatu penelitian yang dengan sengaja peneliti melakukan manipulasi terhadap satu atau lebih variabel dengan suatu cara tertentu sehingga berpengaruh pada satu atau lebih variabel lain yang diukur [27]. Langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut.

1. Melakukan kajian secara induktif yang berkait erat dengan permasalahan yang hendak dipecahkan. Penalaran induktif merupakan prosedur yang berpangkal dari peristiwa khusus sebagai hasil pengamatan empirik dan berakhir pada suatu kesimpulan atau pengetahuan baru yang bersifat umum.
2. Mengidentifikasi dan mendefinisikan masalah.
3. Melakukan studi literatur dan beberapa sumber yang relevan. Studi literatur adalah teknik pengumpulan data dengan mengadakan studi penelaahan terhadap buku-buku, literatur-literatur, catatan-catatan, dan laporan-laporan yang ada hubungannya dengan masalah yang dipecahkan [25]. Studi literatur dilakukan untuk mencari apakah ada literatur yang sudah mengatasi masalah yang ditemukan atau pendekatan solusi yang dapat digunakan.
4. Analisis masalah dan solusi. Menganalisa masalah yang teridentifikasi dan melakukan analisa solusi yang dapat dilakukan setelah mencari pendekatan yang lebih baik.
5. Melaksanakan penelitian. Setelah masalah dan solusi dianalisa, penelitian dilakukan dengan implementasi dalam membangun perangkat lunak.

Pembangunan perangkat lunak menggunakan model *prototype*. Berikut adalah proses-proses yang terjadi pada model tersebut.

a. Pengumpulan Kebutuhan

   Mengidentifikasikan semua kebutuhan, yaitu aturan-aturan tata bahasa Indonesia, literatur, dan jurnal yang berhubungan dengan penelitian yang sedang dilakukan.

b. Desain *Prototype*

   Menganalisa, merancang, dan mendesain sistem yang akan dibuat seperti menganalisa basis data, analisa solusi dari masalah, merancang tampilan antar muka beserta jaringan semantiknya, dan *flowchart* setiap proses yang dilakukan dalam sistem.

c. Pembangunan *Prototype*

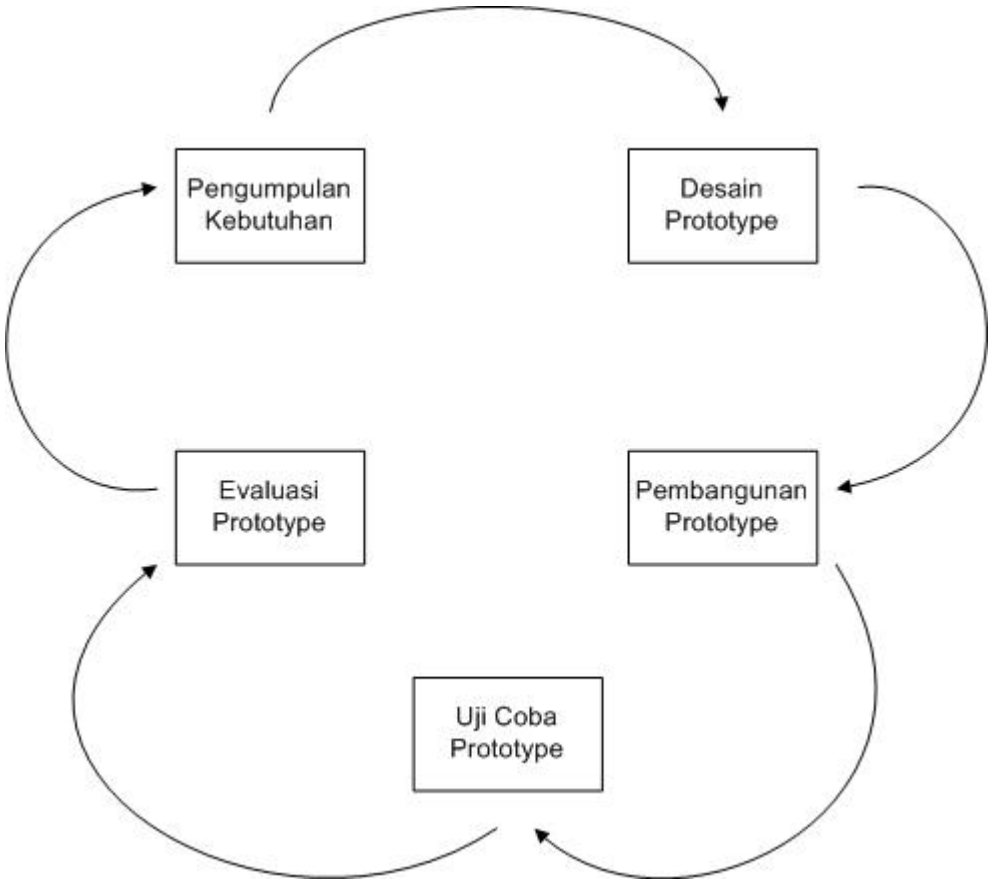
   Hasil pada tahap sebelumnya diimplementasikan pada sistem dengan menggunakan bahasa pemrograman yang digunakan, yaitu Groovy dengan menggunakan *framework* Grails.

d. Uji Coba *Prototype*

   Prototype yang sudah dibangun akan diujicobakan dengan data uji berupa data teks berbahasa Indonesia.

e. Evaluasi *Prototype*

   User sebagai pengguna dari *prototype* akan melakukan evaluasi hasil pertanyaan yang berhasil dibangkitkan dengan memilih pertanyaan mana saja yang relevan dengan teks yang menjadi masukan sebelumnya. Jika hasil kurang memuaskan, maka langkah pertama akan kembali dilakukan.



Gambar 1.1 Model Pengembangan Perangkat Lunak *Prototype*

6. Analisis hasil penelitian. Hasil akhir dari penelitian ini adalah nilai akurasi pada relevansi pertanyaan-pertanyaan yang dibangkitkan dengan teks yang menjadi masukan. Pengukuran dilakukan dengan menghitung pertanyaan-pertanyaan yang relevan dibagi dengan total pertanyaan yang dibangkitkan.

1.7 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini disusun untuk memberikan gambaran umum tentang penelitian pembangkitan pertanyaan yang dilakukan. Sistematika penulisan laporan tugas akhir ini adalah sebagai berikut.

BAB 1 PENDAHULUAN

Menguraikan latar belakang masalah tentang penelitian pembangkitan pertanyaan yang pernah dilakukan dalam bahasa Indonesia. Mengidentifikasi masalah yang terdapat pada salah satu penelitian pembangkitan pertanyaan, kemudian menentukan tujuan penelitian terkait dengan permasalahan yang berhasil teridentifikasi. Menguraikan batasan-batasan masalah agar penelitian tidak terlalu melebar. Menentukan metode penelitian yang dipakai dan menjabarkan sistematika penulisan pada penelitian ini.

BAB 2 LANDASAN TEORI

Berfungsi sebagai kamus teori dalam penelitian ini yang menjelaskan teori-teori atau aturan-aturan yang dipakai pada penelitian yang menjadi acuan pada proses analisis dan perancangan. Teori seperti tata bahasa Indonesia, *pre-processing* yang digunakan, pembentukan pertanyaan, *flowchart*, pohon keputusan, UML, dan bahasa pemrograman.

BAB 3 ANALISIS DAN PERANCANGAN

Penjelasan analisis dan perancangan yang dilakukan pada penelitian ini, seperti analisis masalah, analisis solusi, analisis basis data, analisis kebutuhan fungsional dan non fungsional, analisis pengujian, serta perancangan sistem yang mencakup *Unified Modelling* Language atau UML, perancangan basis data, perancangan prosedural, perancangan struktur menu, perancangan tampilan antar muka, dan jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas tentang pengujian akurasi metode template berdasarkan analisis sintaksis yang dibangun.

BAB 5 KESIMPULAN DAN SARAN

Berisi kesimpulan dari penelitian pembangkitan pertanyaan teks berbahasa Indonesia menggunakan *template* berdasarkan analisis sintaksis dan berisi saran untuk penelitian selanjutnya dalam bidang yang sama agar melengkapi kekurangan dari penelitian yang telah dilakukan.