

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Perkembangan teknologi memang bukanlah menjadi suatu hal baru di era modern seperti sekarang ini, terutama pada teknologi informasi dan komunikasi berkembang begitu pesat. Informasi digital merupakan menjadi sebuah sarana yang wajib untuk mendapatkan informasi yang cepat, efektif dan akurat. Media cetak seperti buku merupakan salah satu sumber informasi yang sangat penting, biasanya untuk pendataan setiap buku mengenai informasi dan jenis-jenis buku dilakukan di suatu dokumen. Namun dari setiap waktu ke waktu dokumen akan bertambah, yang mengakibatkan pendataan informasi buku bisa terjadi banyak kesalahan dan kurang efektif. Maka dari permasalahan tersebut perlu dibangun sebuah sistem agar dalam pendataan informasi dapat dilakukan secara efektif oleh sistem yaitu dengan cara mengekstraksi informasi tersebut dengan mudah, cepat dan benar.

Ekstraksi informasi merupakan suatu bidang ilmu untuk pengolahan Bahasa alami, dengan cara mengubah teks tidak terstruktur menjadi informasi dalam bentuk terstruktur. Data terstruktur merupakan data yang telah tersusun dan tertata sehingga mudah untuk melakukan pencarian informasi data, sebaliknya data tidak terstruktur merupakan data yang masih acak atau belum terorganisir. Ekstraksi Informasi bertujuan untuk mengekstraksi sekumpulan data teks untuk mendapatkan “fakta-fakta berkaitan dengan kejadian (events), entitas atau keterhubungan (relationship)” dalam bentuk informasi terstruktur sebagai masukan basis data [1]. Menurut pendekatan [2] untuk membangun sistem ekstraksi informasi dibagi menjadi dua yaitu *knowledge engineering* dan *automatic training*, pendekatan *knowledge engineering* atau *rule based* menggunakan komponen berupa grammar rules yang ditulis pakar. Sedangkan *automatic training* pembentuk rules yang dilakukan secara otomatis dengan mempelajari dari data latihan yang ada. Pada penelitian sebelumnya yang telah dilakukan Ekstraksi nama lokasi dari tweets informasi lalu lintas menggunakan metode rule base menghasilkan keakuratan

94.5% , 95.10% dan 94.8% penelitian dilakukan oleh Yuda munarko berdasarkan data Surabaya traffic service, Ditlantas Jatim dan Radio Surabaya [3].

Berdasarkan uraian tersebut dibangun sebuah sistem ekstraksi informasi dokumen teks pada buku dengan menggunakan metode rule base agar dalam pendataan informasi yang berkaitan dengan buku dapat menghasilkan data akurat, cepat mudah dan tidak memakan waktu yang lama untuk mengolahnya. Dengan adanya sistem tersebut diharapkan dapat menunjang dan memudahkan pendataan buku agar cepat, akurat dan efektif.

### **1.2 Perumusan Masalah**

Berdasarkan latar belakang yang sudah dikemukakan, dapat diambil rumusan masalah yaitu bagaimana mendapatkan informasi sebuah buku yang tidak terstruktur menjadi informasi terstruktur.

### **1.3 Maksud dan Tujuan**

Penelitian ini bermaksud membuat sistem sistem ekstraksi informasi pada buku. Sedangkan tujuan yang ingin dicapai dari penelitian ini adalah agar memperoleh hasil Ekstraksi informasi buku text yang benar.

### **1.4 Batasan Masalah**

Agar pengerjaan tugas akhir ini lebih terarah, maka diperlukan batasan masalah. Sehingga pembahasan dibatasi pada ruang lingkup:

1. Dokumen yang digunakan sebagai masukan berupa buku ekstensi .txt hasil konversi OCR dari hasil scan JPG.
2. Buku masukan hanya buku yang menggunakan berbahasa indonesia.
3. Keluaran yang dihasilkan dari sistem adalah judul, nama penulis, tahun terbit, dan ISBN.
4. Menggunakan metode Rule Base.

## **1.5 Metodologi Penelitian**

Metodologi penelitian yang dipakai adalah metode deskriptif. Metodologi deskriptif merupakan yang mempelajari masalah-masalah yang ada serta tata cara kerja yang berlaku. Pada metode ini menghimpun data, menyusunnya secara sistematis, faktual dan cermat. Metode penelitian ini memiliki dua tahapan yaitu pengumpulan data dan pembangunan perangkat lunak

### **1. Identifikasi Masalah**

Tahap pertama dalam penelitian adalah perumusan masalah. Peneliti merumuskan masalah ataupun kendala pada proses ekstraksi informasi pada buku.

### **2. Pengumpulan Data**

Metode pengumpulan data yang digunakan dalam kasus ini adalah sebagai berikut:

#### **2.1 Studi Literatur**

Studi literatur adalah teknik pengumpulan data dengan cara mengumpulkan literatur, jurnal, browsing internet dan bacaan-bacaan yang ada kaitannya dengan topik yang diambil.

#### **2.2 Pengumpulan Data**

Pada tahap ini diambil pengumpulan data buku dengan beberapa buku. Data sampel ini digunakan sebagai data training saat melakukan ekstraksi informasi pada buku.

### **3. Analisis dan Perancangan Sistem**

Tahap analisis dan perancangan sistem, peneliti melakukan analisis dan perancangan sistem yang akan dibangun pada aplikasi berdasarkan dari permasalahan dan data yang telah diperoleh, sehingga tercapai suatu tujuan dari peneliti ini. Pada tahapan ini terdapat dua bagian, yaitu :

### 3.1 Analisis Sistem

Proses melakukan analisis pada sistem yang akan dibangun, berikut adalah cakupan dalam tahapnya :

1) Identifikasi Buku

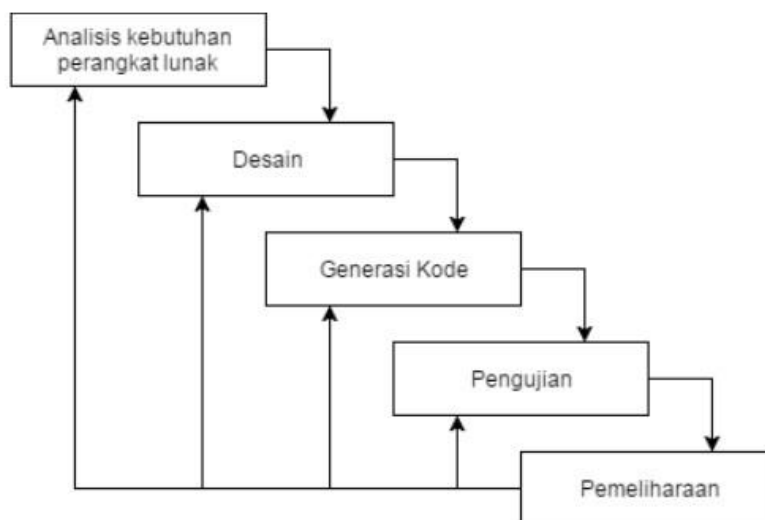
Pada tahapan analisis identifikasi buku adalah mengganti data buku yang digunakan pada penelitian dan mengamati metode yang sesuai untuk penambihan ekstraksi informasi pada buku.

2) Pembentukan aturan

Pada tahap ini mengambil kata kunci dan aturan yang telah di analisis pada data buku yang akan digunakan untuk ekstraksi informasi.

### 3.2 Pembangunan Sistem

Metode analisis data dalam pembuatan perangkat lunak menggunakan metode waterfall. Menurut Pressman metode *waterfall* adalah model klasik yang bersifat sistematis, berurutan dalam membangun software. Terdiri dari 5 tahapan yang saling terkait dan mempengaruhi yang meliputi beberapa proses berikut :



**Gambar1. 1 Waterfall Model**

Berikut ini akan dijelaskan mengenai tahap-tahap pembangunan sistem dengan menggunakan metode *waterfall*.

1. Analisis kebutuhan perangkat lunak

Langkah ini merupakan analisis terhadap kebutuhan *software*, dan tahap dalam pengumpulan kebutuhan penelitian secara lengkap tentang ekstraksi informasi pada buku. Data yang diambil adalah data buku dalam file ekstensi .txt dan mengumpulkan data-data tambahan baik yang ada dalam jurnal, artikel, maupun dari internet.

2. Desain

Setelah kebutuhan data dan pemroses telah dikumpulkan, maka tahap selanjutnya adalah perancangan arsitektur sistem secara keseluruhan. Untuk kebutuhan perangkat lunak dalam pengolahan data diperlukan *software* yang diantaranya penyunting kode sumber yang berjalan pada sistem

3. Generasi Kode

Setelah perancangan sistem secara keseluruhan, maka tahap selanjutnya adalah mengimplementasikan ke dalam bahasa pemrograman. Tahap dimana aturan *rule-base* digunakan dalam menyelesaikan permasalahan dalam ekstraksi informasi pada buku.

4. Pengujian

Tahapan ini digunakan untuk mengintegrasikan program dan pengujian tujuan, apabila program yang dibuat berjalan sesuai dengan harapan, dan sesuai dengan pendekatan aturan maka ekstraksi informasi buku siap digunakan oleh pengguna sistem.

5. Pemeliharaan

Tahapan ini merupakan akhir dari pembuatan sebuah *software* atau sistem. Setelah melakukan analisis, desain dan pengkodean maka sistem yang sudah jadi akan digunakan oleh *user*. Kemudian *software* yang telah dibuat harus dilakukan pemeliharaan secara berkala.

### **3.3 Pengujian Sistem**

Tahapan selanjutnya adalah pengujian sistem yang telah dibangun. Pengujian ini dilakukan dengan menggunakan pengujian black box dan pengujian aturan hasil ekstraksi informasi.

### **3.4 Penarikan Kesimpulan**

Tahapan penarikan kesimpulan, peneliti memberikan kesimpulan terhadap sistem yang telah dibangun berdasarkan dari hasil pengujian yang telah dilakukan.

## **1.6 Sistematikan Penulisan**

Sistematika penulisan ini disusun untuk memberikan gambaran umum tentang penulisan tugas akhir yang akan dilakukan. Sistematika penulisan tugas akhir ini adalah sebagai berikut.

### **BAB 1 PENDAHULUAN**

Pada bab ini membahas uraian mengenai latar belakang masalah yang diambil, rumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian dan sistematika penulisan.

### **BAB 2 LANDASAN TEORI**

Pada bab ini berisi bahasan berupa bahan-bahan kajian, konsep dasar, dan teori dari para ahli yang berkaitan dengan penelitian. Meninjau permasalahan dan hal-hal yang berguna dari penelitian-penelitian dan sintesis serupa yang pernah dikerjakan sebelumnya dan menggunakan sebagai acuan pemecahan masalah pada penelitian.

### **BAB 3 ANALISIS DAN PERANCANGAN SISTEM**

Pada bab ini berisi tentang analisa kebutuhan sistem dengan penjabaran mengenai perancangan aplikasi serta perancangan sistem yang akan dibangun.

## **BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM**

Bab ini membahas mengenai solusi beserta implementasi dari masalah yang telah dianalisis. Dan menentukan bagaimana sistem dirancang, dibangun, dan diuji yang sesuai dengan hasil penelitian.

## **BAB 5 KESIMPULAN DAN SARAN**

Bab ini akan membahas mengenai penarikan kesimpulan dari penelitian yang sudah dikerjakan dan menjawab hal-hal yang melatarbelakangi masalah pada bab 1, dan berisi pula saran yang mampu menjadi masukan dan pengembangan selanjutnya dari penulis penelitian maupun masukan dari pengguna sistem.

