

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Internet merupakan sumber informasi yang sangat besar, baik informasi itu berupa teks, media atau data dalam bentuk lain. Setiap halaman *web* menampilkan data dalam bentuk yang berbeda-beda. Data-data yang ditampilkan sangat berguna baik secara fungsi, analisis, maupun untuk keperluan bisnis. Sayangnya sebagian besar *website* tidak membuka data-data ini secara bebas. Kebanyakan *website* tidak menyediakan fasilitas untuk menyimpan data tersebut ke penyimpanan lokal. Dari permasalahan inilah maka muncul rekayasa perangkat lunak untuk mengambil data-data tersebut.

Kay Russel mengungkapkan "Web scraping is the process of gathering and organizing unstructured information pages and data on the World Wide Web". [1]. Pernyataan tersebut menjelaskan bahwa web scraping digunakan untuk mendapatkan informasi dari halaman dan data yang tidak terstruktur dari *website*. Web scraping juga bisa diterapkan untuk mempermudah mendapatkan informasi berupa berkas, media secara otomatis tanpa harus mengunjungi *website* yang dituju, penggunaan seperti ini akan memperingan pekerjaan dari segi waktu dan tenaga.

Karya ilmiah ini akan menerapkan teknik *web scraping* untuk mendapatkan buku elektronik gratis dari suatu *website* setiap hari secara otomatis tanpa seseorang harus mengunjungi *website* tersebut.

Sebuah *website* yang bernama [www.packtpub.com](http://www.packtpub.com) mempunyai layanan untuk membeli buku-buku elektronik tentang ilmu teknologi informasi. Selain

menjual buku-buku elektronik berbayar Packtpub juga menyediakan buku elektronik gratis yang hanya berlaku untuk satu hari saja. Fitur ini disediakan Packtpub agar user tertarik untuk mendaftar ke *website* tersebut dan menjadi pelanggan Packtpub.

Buku-buku elektronik gratis hanya bisa didapatkan oleh member yang telah terdaftar di Packtpub. Member akan mendapatkan tautan untuk dapat mengunduh buku elektronik tertentu setelah melakukan klaim atas buku elektronik tersebut. Maka untuk mendapatkan semua buku elektronik setiap hari member harus mengklaim tautan tersebut setiap hari.

Berdasarkan permasalahan itulah penulis menyusun skripsi ini yang berjudul "APLIKASI 'BOOKSCRAP' UNTUK MENGUNDUH BUKU ELEKTRONIK GRATIS DENGAN MENERAPKAN METODE WEB SCRAPING", untuk membantu member mendapatkan buku elektronik gratis secara otomatis dari Packpub.

### **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah diuraikan maka dapat dirumuskan masalahnya "Bagaimana membuat aplikasi "bookscrap" untuk mengunduh buku elektronik gratis dengan menerapkan metode web scraping?"

### **1.3 Batasan Masalah**

Dalam penyajian karya ilmiah ini diperlukan batasan masalah agar pembahasan lebih fokus, terarah dan tidak melebar, maka penulis memberikan batasan-batasan pembahasan masalah sebagai berikut :

1. Aplikasi ini dikembangkan hanya sampai pada tahap *prototype*.

2. Aplikasi ini hanya diterapkan pada website [www.packtpub.com](http://www.packtpub.com).
3. Aplikasi ini akan mengunduh buku elektronik menggunakan identitas pengguna *website* Packtpub.
4. Pada karya ini tidak membahas aspek hukum secara mendalam mengenai e-book yang diunduh
5. Aplikasi ini akan dibagi menjadi beberapa bagian yaitu *scraper*, API dan halaman muka pengguna.
6. Aplikasi ini tidak menggunakan konsep kecerdasan buatan untuk melakukan *scraping*.
7. Aplikasi ini berjalan pada server Linux Ubuntu Server 14.04.
8. Aplikasi ini dibuat menggunakan bahasa pemrograman Node.JS, PHP dengan *framework* CodeIgniter 3.0 dan Smarty 3.0, serta MySQL sebagai RDBMS-nya.

#### **1.4 Maksud dan Tujuan Penelitian**

##### **1.4.1 Maksud**

Maksud diadakannya penelitian ini adalah sebagai syaran kelulusan dan mendapatkan gelar kesarjanaan komputer pada jurusan Teknik Informatika STMIK AMIKOM Yogyakarta.

##### **1.4.2 Tujuan Penelitian**

Tujuan yang ingin dicapai dari penelitian ini adalah :

1. Membangun aplikasi Bookscrap untuk mengunduh buku elektronik gratis dengan menerapkan metode web scraping.

2. Membantu mempermudah member Packtpub dalam mendapatkan buku-buku elektronik gratis.

## **1.5 Metode Penelitian**

### **1.5.1 Metode Pengumpulan Data**

Metode penelitian yang dilakukan untuk mendapatkan informasi-informasi tentang objek permasalahan ini adalah :

1. Metode Pengamatan (*Observation*)

Metode pengumpulan data dengan cara mengamati langsung pada permasalahan yang akan diteliti untuk memperoleh informasi.

2. Metode Kepustakaan (*Library*)

Pengumpulan data dengan membaca buku-buku, literatur-literatur serta berbagai kode sumber dari komunitas di internet yang berhubungan dengan permasalahan sebagai referensi dan bahan pembandingan.

### **1.5.2 Metode Analisis**

Pada tahap ini dilakukan analisis sebagai berikut.

1. Analisis Proses Bisnis Pada Sistem Manual.
2. Analisis Kebutuhan Sistem
  - a. Analisis Kebutuhan Fungsional
  - b. Analisis Kebutuhan Non Fungsional
3. Analisis Kelayakan Sistem
  - a. Analisis Kelayakan Teknologi
  - b. Analisis Kelayakan Hukum

### 1.5.3 Metode Perancangan

Menggambarkan spesifikasi sistem baru yang akan dibuat, pada tahapan ini ada beberapa dokumen yang akan dibuat, meliputi:

#### 1.5.1.1. Pemodelan Proses (*Process Modeling*)

Pemodelan proses adalah cara formal untuk menggambarkan bagaimana bisnis beroperasi, mengilustrasikan aktivitas-aktivitas yang dilakukan dan bagaimana data berpindah di antara aktivitas-aktivitas itu. Penulis menggunakan UML (*Unified Modeling Language*) untuk merepresentasikan pemodelan proses pada sistem yang dirancang.

#### 1.5.1.2. Pemodelan Data (*Data Modeling*)

Data model adalah cara formal untuk menggambarkan data yang digunakan dan diciptakan dalam suatu sistem bisnis. Penulis menggunakan ERD (Entity Relationship Diagram) untuk merepresentasikan pemodelan data sistem yang dirancang.

#### 1.5.1.3. Desain Antar Muka (*Interface Design*)

Antarmuka pengguna merupakan tampilan dimana pengguna berinteraksi dengan sistem.

### 1.5.4 Metode Pengembangan

Pada tahap pengembangan, metode yang digunakan untuk mengembangkan aplikasi ini adalah metode *Prototyping*. *Prototyping* adalah proses iteratif dalam pengembangan sistem di mana kebutuhan diubah ke dalam sistem yang sudah bisa bekerja dengan baik (*working system*) yang secara terus menerus diperbaiki melalui

kerjasama antara pengguna dan tim pengembang (*developer*). Implementasi metode *Prototyping* ini akan dibagi menjadi 4 fase, yaitu:

1. Tim pengembang berdiskusi dengan pengguna untuk menentukan tujuan umum dan kebutuhan dari aplikasi yang akan dibuat.
2. Tim pengembang kemudian membangun *prototype* yang sesuai dengan hasil diskusi pada fase pertama. Ketika sebuah *prototype* telah selesai, pengguna mencoba menggunakan *prototype* yang telah dibangun itu dan menyampaikan pada tim pengembang tentang apa saja yang telah sesuai dengan hasil fase I dan apa saja yang belum sesuai.
3. Tim pengembang kemudian memberikan *feedback* untuk memperbaiki *prototype*.
4. Kemudian *prototype* versi terbaru kembali ke pengguna dan begitu seterusnya sampai memenuhi kebutuhan dan tujuan umum dari aplikasi.

#### **1.0 Sistematika Penulisan**

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

##### **1.6.1 BAB I PENDAHULUAN**

Berisi latar belakang masalah, rumusan masalah, batasan masalah, maksud dan tujuan penelitian, metode penelitian dan sistematika penulisan.

##### **1.6.2 BAB II LANDASAN TEORI**

Berisi uraian beberapa landasan teori yang akan digunakan dalam menyelesaikan skripsi.

#### 1.6.3 BAB III ANALISIS DAN PERANCANGAN

Berisi tentang gambaran umum aplikasi, identifikasi masalah, solusi masalah, analisis kebutuhan sistem, analisis kelayakan sistem dan perancangan sistem.

#### 1.6.4 BAB IV IMPLEMENTASI DAN PEMBAHASAN

Berisi tentang setiap tahapan dalam pembuatan sistem dimulai dari perancangan sistem sampai dengan implementasi program.

#### 1.6.5 BAB V PENUTUP

Berisi hasil penelitian yang memuat kesimpulan dan saran dari hasil pembahasan Bab IV.

#### DAFTAR PUSTAKA

Berisi dari sumber-sumber pustaka yang diambil dari buku, majalah, narasumber maupun data internet.