

# Bab 1

## Pendahuluan

### 1.1 Latar Belakang

Web semantik merupakan evolusi dan perluasan dari *www (World Wide Web)* yang ditampilkan tidak hanya dengan format bahasa manusia yang umum tetapi juga bisa dibaca dan digunakan oleh mesin. Web semantik juga merupakan perluasan web saat ini dimana informasi yang disampaikan memiliki definisi arti yang baik, memungkinkan komputer dan manusia bekerjasama lebih baik [1]. Perubahan dari teknologi web semantik yang sangat signifikan adalah perubahan dari “Web of documents” menjadi “Web of data”. Salah satu perkembangan teknologi yang digunakan pada web semantik adalah *ontology*.

*Ontology* merupakan suatu teori tentang makna dari suatu objek, properti dari suatu objek, serta relasi objek tersebut yang mungkin terjadi pada suatu domain pengetahuan. *Ontology* juga merupakan spesifikasi dari sebuah konseptual, dengan kata lain *ontology* adalah penjelasan dari sebuah konsep dan keterhubungannya dari sebuah ilmu tertentu.

*Ontology* mendeskripsikan data pada web dan keterhubungan antar data pada web. Dengan semakin berkembangnya teknologi web semantik, maka akan semakin banyak pula jumlah *ontology* yang tersedia di internet namun oleh karena itu maka ketersediaan *ontology* yang sangat banyak ini dapat menimbulkan heterogenitas. Heterogenitas *ontology* memunculkan permasalahan, misalnya terdapat dua *ontology* dengan nama yang berbeda, *ontology* tersebut memiliki struktur yang berbeda atau didefinisikan dengan cara yang berbeda padahal kedua *ontology* tersebut mengekspresikan pengetahuan yang sama namun dalam bahasa yang berbeda. Cara yang digunakan untuk mengurangi permasalahan heterogenitas *ontology* pada web semantik adalah *ontology matching*.

*Ontology matching* merupakan proses untuk membandingkan dua *ontology* dan menemukan keterhubungan diantara kedua *ontology* tersebut. Terdapat beberapa teknik yang dapat digunakan untuk menyelesaikan permasalahan heterogenitas dengan menggunakan *ontology matching* diantaranya adalah *terminological techniques (name-based technique)*, *structure-based techniques*, *instance-based techniques (extensional techniques)* dan *semantic-based techniques*. Dari beberapa teknik yang digunakan dalam *ontology matching* tersebut, salah satu teknik yang digunakan dalam menyelesaikan permasalahan heterogenitas pada tugas akhir ini adalah teknik *Semantic-based Ontology Matching*. Karakteristik dari *Semantic-based Ontology Matching* ini adalah model teori semantik digunakan untuk membenarkan hasilnya, metode yang seperti ini biasa disebut dengan metode deduktif. Metode deduktif ini tidak tampil secara baik untuk sebuah tugas induktif dasar seperti pencocokan *ontology*. Diperlukan fase preprocessing yang menyediakan ‘*anchors*’ yaitu entitas yang dideklarasikan, contohnya agardapat setara maka didasarkan pada entitas namanya atau input *user* untuk *instance*. Jadi *Semantic-based Ontology Matching* ini didasarkan pada penggunaan *resource* yang ada secara formal untuk menginisiasikan suatu keselarasan yang dapat dipertimbangkan lebih lanjut dengan metode deduktif [2].

Penggunaan metode *Semantic-Based Ontology Matching* ini diharapkan dapat memberikan suatu pemetaan kelas antar data *ontology* yang lebih luas, karena pemetaan kata berdasarkan hubungan semantiknya diyakini akan memberikan korespondensi yang lebih banyak sehingga dapat memberikan keluaran (*output*) *ontology alignment* baru yang mempunyai cakupan lebih luas.

## **1.2 Perumusan Masalah**

Dalam tugas akhir ini akan dirumuskan masalah sebagai berikut:

1. Bagaimana menerapkan metode *Semantic-based ontology matching* untuk menganalisis dan menentukan keterhubungan *ontology* pada web semantic?
2. Apakah metode *Semantic-based ontology matching* dapat menyelesaikan masalah heterogenitas pada web semantic?

3. Bagaimana nilai parameter *Similarity Threshold* mempengaruhi nilai performansi *precision*, *recall*, dan *f-measure* dari metode *Semantic-based ontology matching* dalam menentukan ketrhubungan *ontology* pada web semantik?

### **1.3 Tujuan**

Tujuan dari penelitian tugas akhir ini adalah:

1. Mengaplikasikan metode *Semantic-based Ontology Matching* untuk menangani masalah heterogenitas yang ada pada web semantik.
2. Menganalisis keterhubungan *ontology* pada web semantik berdasarkan parameter *Similarity Threshold* yang mempengaruhi nilai performansi *presicion*, *f-measure*, *recall* pada metode *Semantic-based ontology matching*.

### **1.4 Batasan Masalah**

Adapun batasan masalah dalam tugas akhir ini adalah sebagai berikut:

1. Data input yang digunakan adalah data yang berekstensi \*.owl dan \*.rdf.
2. Pengimplementasian program menggunakan Java.
3. Proses pencarian kemiripan *semantic (instance similarity measure)* menggunakan Lucene API.
4. Data yang digunakan dalam pencocokan *ontology* adalah data yang berbahasa Inggris.

### **1.5 Metodologi Penyelesaian**

Tahapan-tahapan yang akan ditempuh dalam pelaksanaan penelitian tugas akhir ini adalah sebagai berikut:

1. Studi literatur

Pencarian materi atau referensi yang dibutuhkan dan berhubungan dengan permasalahan seperti *Semantic web*, *ontology*, *Ontology Matching*, *Semantic-based Ontology Matching*, *Ontology Alignment*, *Precision*, *Recall*, dan *F-Measure*. Kemudian melakukan pendalaman materi atau referensi yang ditemukan untuk membantu pemecahan masalah pada penelitian ini.

## 2. Pengumpulan data

Pada tahapan ini terdapat proses mengumpulkan data yang akan digunakan sebagai data input kedalam sistem. Data input yang akan digunakan pada penelitian tugas akhir ini adalah data *ontology* yang berekstensi \*.owl dan \*.rdf

## 3. Perancangan sistem

Tahap penerapan metode *Semantic-based Ontology Matching* yang mencari keterhubungan dua buah *ontology* yang berbeda dengan cara mencocokkan kelas-kelas yang terdapat pada dua buah *ontology* tersebut berdasarkan *instance* yang dimilikinya. Inputan dari sistem berupa dua buah *ontology* yang berekstensi \*.owl atau \*.rdf. Kedua *ontology* tersebut kemudian akan dicari keterkaitannya dengan menggunakan metode *Semantic-based Ontology Matching* yang telah diimplementasikan kedalam sistem. Hasil *output* dari proses pencocokan ini berupa pemetaan keterkaitan kelas (*class mapping*) yang dibuat dalam format *ontology alignment* dengan ekstensi file \*.rdf dan menampilkan nilai performansi *recall*, *precision*, *f-measure*.

## 4. Implementasi sistem

Pada tahap ini sistem akan dibangun berdasarkan perancangan yang telah dibuat sebelumnya. Sistem akan dibangun dengan menggunakan bahasa pemrograman Java dan berbasis *desktop* dengan memanfaatkan *library* Jena API, Lucene API, dan *Alignment* API.

## 5. Pengujian

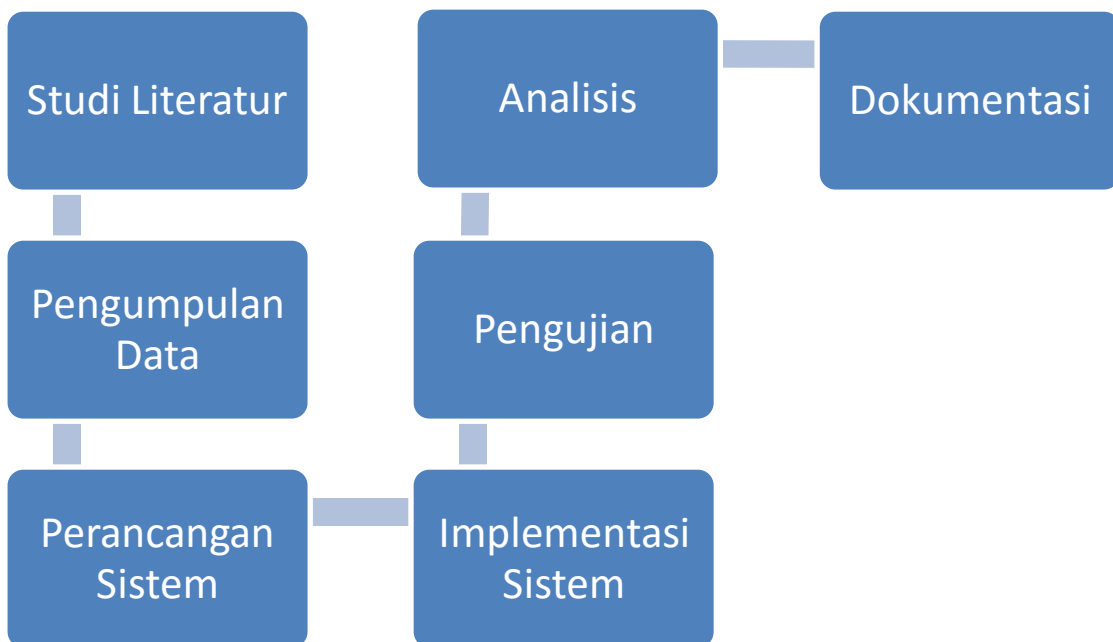
Pada tahap ini dilakukan pengujian terhadap sistem yang dibangun dengan inputan *dataset* yang telah dibuat sebelumnya untuk mengukur pengaruh dari parameter *Similarity Threshold* terhadap metode *Semantic-based Ontology Matching*.

## 6. Analisis

Melakukan analisis terhadap pengaruh parameter *Similarity Threshold* terhadap nilai performansi *recall*, *precision*, dan *f-measure* dari metode *Semantic-based Ontology Matching*.

## 7. Dokumentasi

Dokumentasi dilakukan untuk mencatat setiap hal penting dalam pelaksanaan penelitian tugas akhir ini.



Gambar 1.2: Metode penyelesaian masalah

## 1.6 Sistematika Penulisan

Secara garis besar skripsi ini terdiri dari 5 (lima) bab dengan beberapa sub bab. Agar mendapat arah dan gambaran yang jelas mengenai hal yang akan ditulis, berikut ini adalah sistematika penulisan yang penulis buat secara lengkap:

### BAB 1: Pendahuluan

Pada bab ini penulis menjelaskan latar belakang, perumusan masalah, tujuan, batasan masalah, metodologi penyelesaian, serta sistematika penulisan dari tugas akhir ini.

### BAB 2: Tinjauan Pustaka

Pada bab ini akan diuraikan konsep tentang pengenalan tentang web semantik, komponennya, dan *ontology*, serta metode apa saja yang dapat digunakan dalam *ontology matching*

### BAB 3: Perancangan Sistem

Dalam bab ini, penulis mencoba untuk menguraikan tentang bagaimana gambaran umum sistem yang akan dibangun, dengan menggambarkan setiap proses yang ada didalam sistem. Serta menjelaskan setiap kebutuhan dari sistem yang akan dibangun ini.

### BAB 4: Pengujian dan Analisis

Dalam bab ini penulis membahas tentang tujuan pengujian, prosedur-prosedur yang dilakukan untuk pengujian, menjelaskan skenario pengujian, serta menuliskan analisis yang dibuat berdasarkan pengujian yang dilakukan.

### BAB 5: Kesimpulan dan Saran

Pada bagian bab ini merupakan bab terakhir yang berisikan kesimpulan dari pembahasan yang diuraikan diatas serta saran-saran yang dianggap perlu dalam usaha menuju perbaikan dan kesempurnaan.