

# **BAB I**

## **PENDAHULUAN**

### **1.1. Latar Belakang**

Internet merupakan media untuk publikasi. Internet menyimpan data berupa teks dan gambar yang dibuat dan disediakan untuk konsumsi publik ataupun pribadi. Terdapat sejumlah besar informasi yang disebarikan melalui internet. Pencarian konvensional yang berdasarkan string sering kali gagal menemukan halaman yang relevan dan menampilkan hasil yang tidak relevan dengan permintaan pengguna. Hal ini menunjukkan bahwa semua data tersedia di web, dan pengguna sulit menemukan apa yang dibutuhkan.

Telah banyak teknologi informasi yang diadopsi untuk mengatasi masalah ini, diantaranya teknik pencarian yang menganalisa struktur dari web untuk menentukan dokumen yang penting. Bahasa pemrograman berbasis web utamanya berfokus pada dokumen yang dapat dipahami oleh manusia dan juga pada struktur tampilan dokumen. Selain itu juga terdapat pencarian berbasis kata kunci yang menunjukkan bahwa hubungan sumber data dengan kata kunci kebanyakan berdasarkan frekuensi kemunculan kata kunci pada deskripsi teks (seperti *title*, *body*, *anchor*, *text*, dan lain sebagainya). Hal ini tidak dapat memastikan bahwa hasil yang diberikan menunjukkan hubungan makna antara kata kunci yang diinginkan pengguna (Li dkk., 2007).

Tim Burners Lee kemudian memunculkan ide semantik untuk menghubungkan jejaring web. Pencarian semantik berarti mencari informasi pada web dengan berdasarkan makna kata kunci. Yang mana makna kata kunci ini seringkali tidak ada pada kebanyakan pusat pencarian informasi, dan membutuhkan pemrograman yang kompleks. Semantik tidak hanya digunakan untuk menghubungkan dokumen satu sama lain tetapi juga mengenali dan memahami arti dari informasi tersebut (Hebeler dkk., 2009).

Teknologi semantik memanfaatkan ontologi untuk mengubah data yang tidak terstruktur menjadi terstruktur. Sehingga dapat meningkatkan nilai presisi hasil pencarian (Henry, 2014; Akbar, 2015). Selain itu pemanfaatan ontologi

dapat juga digunakan untuk memperluas skala hasil pencarian (Kara dkk., 2012). Dalam sistem pencarian, sekumpulan dokumen dan kata kunci yang telah diekstrak dan diperluas kemudian akan dihitung nilai kemiripannya. Dan algoritma pengukuran kemiripan yang digunakan adalah menggunakan model ruang vektor. Algoritma ini telah banyak diimplementasikan. Tidak hanya untuk mengukur kesamaan semantik antara dokumen dengan kueri, tetapi dapat juga diimplementasikan untuk mengukur kesamaan antara dua buah ontologi (*ontology alignment*) (Eidoon dkk., 2008).

Konsep pencarian semantik ini yang kemudian digunakan untuk melakukan penelitian. Penelitian semantik ini dibangun berbasis ontologi menggunakan algoritma model ruang vektor (*Vektor Space Model*). Data yang digunakan pada penelitian ini diambil dari web, terutama web yang memuat artikel berbahasa Indonesia terkait batik. Batik dipilih sebagai tema penelitian ini adalah karena batik merupakan warisan leluhur budaya bangsa Indonesia. Informasi batik banyak tersedia di internet. Tetapi pencarian berbasis kata kunci saat ini tidak mampu menampilkan hasil pencarian sesuai yang diinginkan pengguna. Hasil pencarian yang diberikan terkadang lebih condong pada sisi komersil batik. Dan mengabaikan sisi budaya ataupun sisi informatif lain yang dikandung oleh batik itu sendiri. Untuk itu dengan mengembangkan pencarian semantik berbasis ontologi menggunakan algoritma model ruang vektor diharapkan dapat lebih meningkatkan nilai presisi dan *recall* hasil pencarian.

## **1.2. Tujuan Penelitian**

Penelitian ini bertujuan untuk mengembangkan konsep pencarian semantik berbasis ontologi untuk memfokuskan hasil pencarian agar sesuai dengan struktur domain ontologi batik yang dibangun.

### **1.3. Manfaat Penelitian**

Beberapa manfaat yang ingin dicapai terkait penelitian ini antara lain:

1. Penelitian ini diharapkan dapat menambah referensi keilmuan terutama konsep pencarian semantik berbasis ontologi menggunakan algoritma model ruang vektor (*vector space model*) bagi academia.
2. Selain itu dapat memberikan informasi batik yang lebih luas dan mendalam bagi pengguna sehingga dapat lebih meningkatkan pemahaman pengguna mengenai batik.