

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Perkembangan Internet saat ini terus berkembang dengan sangat pesat. Berbagai konten disediakan di Internet. Semakin banyaknya konten yang ada dalam Internet membuat layanan pencarian semakin dibutuhkan. Perkembangan mesin pencarian di Internet terus berkembang untuk memenuhi kebutuhan pengguna. Sampai saat ini telah banyak metode pencarian yang telah dikembangkan antara lain pencarian teks penuh, pencarian metadata, dan pencarian semantik

Pencarian teks penuh merupakan metode pencarian yang paling pertama dikembangkan. Pencarian dengan metode ini adalah melakukan pencarian dengan mencocokkan setiap kata yang ada dengan kata yang diberikan oleh pengguna[12]. Akan tetapi, metode ini masih memiliki kelemahan. Metode ini dapat menghasilkan hasil pencarian yang sangat banyak sehingga memungkinkan ketidakakuratan dalam pencarian. Metode ini juga tidak dapat menangani sinonim dari kata yang diberikan oleh pengguna.

Pencarian metadata merupakan metode pencarian yang lebih baik dibandingkan pencarian teks penuh karena pencarian metadata dapat menangani OR dan AND sehingga pengguna dapat lebih menspesifikan hasil pencarian yang diinginkan. Sebagai contoh, pengguna ingin mencari buku berdasarkan judul. Hasil pencarian berdasarkan judul buku dapat menghasilkan hasil pencarian yang banyak. Pengguna dapat menspesifikan buku dengan mencantumkan juga nama penerbit, nama pengarang dsb. Akan tetapi, metode ini masih memiliki kelemahan karena metode ini masih belum bisa menangani sinonim kata pencarian.

Pencarian semantik merupakan metode pencarian yang sedang dikembangkan pada *web 3.0*. metode ini dikembangkan untuk menangani kelemahan pada metode pencarian sebelumnya. Pencarian semantik dapat mengenali sinonim kata yang dibentuk berdasarkan graf. beberapa pemodelan data yang ada pada pencarian semantik adalah *weighted tree similarity*, *weighted directed acyclic graphs*, dan ontologi[12]. Pencarian semantik juga melakukan pencarian hingga isi konten dokumen. Sehingga memungkinkan hasil pencarian yang lebih relevan.

Institut Teknologi Telkom menyediakan fasilitas pencarian dokumen tugas akhir yang pernah dikerjakan oleh mahasiswa Institut Teknologi Telkom dalam bentuk *web*. Dalam *web* ini, pencarian masih menggunakan pencarian teks

penuh berdasarkan judul, nama mahasiswa, atau nim yang dimasukkan oleh pengguna. Kata kunci dari suatu judul tugas akhir tidak mencakup semua isi laporan dari tugas akhir tersebut sehingga pencarian dokumen tidak efektif. Pencarian semantik merupakan solusi untuk masalah pencarian pada kasus tersebut karena pencarian semantik melakukan pencarian tidak hanya berdasarkan judul tetapi juga berdasarkan konten.

Banyak metode yang dapat digunakan untuk mengimplementasikan pencarian semantik. Salah satunya adalah *weighted directed acyclic graph*. Operasi perhitungan pada *weighted directed acyclic graph* yaitu dengan membandingkan antara graf dokumen sebagai *dataset* dengan graf berdasarkan *query* yang diberikan pengguna. Hal terpenting dalam membangun pencarian semantic adalah ontology. Dengan penerapan ontologi juga dapat meningkatkan interoperabilitas dari sistem sehingga dapat diakses di berbagai *platform*.

Oleh karena itu, penulis sangat tertarik untuk melakukan pengimplementasian dan analysis terhadap penerapan pencarian semantik dengan *weighted directed acyclic graph* ini dalam pencarian dokumen tugas akhir mahasiswa Institut Teknologi Telkom. Diharapkan dengan pencarian semantik ini dapat membantu mahasiswa untuk menemukan informasi sesuai yang diharapkan.

1.2 Rumusan Masalah

Rumusan masalah yang akan dilakukan penelitian sebagai berikut:

1. Bagaimana mengimplementasikan *semantic search* dengan *weighted directed acyclic graph* pada kasus pencarian dokumen tugas akhir di Institut Teknologi Telkom?
2. Bagaimana perancangan *ontology* pada kasus pencarian dokumen tugas akhir di Institut Teknologi Telkom?
3. Bagaimana relevansi hasil yang dihasilkan sistem terhadap hasil yang dianggap relevan?
4. Bagaimana kinerja algoritma *weighted directed acyclic graph* dengan melakukan pengujian secara algoritma dan relevansi *output system*?

1.3 Tujuan

Tujuan dari penelitian ini antara lain:

1. Mengimplementasikan *semantic search* dengan *weighted directed acyclic graph* pada kasus pencarian dokumen tugas akhir di Institut Teknologi Telkom.
2. Merancang ontologi pada kasus pencarian dokumen tugas akhir di Institut Teknologi Telkom.
3. Melakukan analisis terhadap hasil yang dihasilkan sistem terhadap hasil yang dianggap relevan.
4. Melakukan pengukuran terhadap algoritma *weighted directed acyclic graph* dengan pengujian algoritma dan relevansi *output system*.

1.4 Lingkup Penelitian

Dengan begitu luasnya teknologi *web semantic*, penulis membatasi penelitian ini dengan rincian sebagai berikut:

1. *Dataset* dalam penelitian tugas akhir ini adalah metadata dokumen tugas akhir fakultas informatika Institut Teknologi Telkom dengan Mahasiswa angkatan 2008 dan 2009.
2. Bahasa tektual yang digunakan adalah *ontology*, XML, RDF, dan OWL.
3. Algoritma yang digunakan dalam penelitian ini adalah *weighted directed acyclic graphs*.
4. Tools yang digunakan untuk pembuatan *ontology* adalah Protégé 3.4.8.
5. Analisis utama dalam penelitian ini adalah performansi sistem dalam hal akurasi.

1.5 Metodologi Penelitian

Metodologi yang digunakan dalam penelitian ini antara lain:

1. Studi Literatur
Mempelajari literatur-literatur yang berkaitan dengan *semantic web* dikarenakan *semantic web* merupakan teknologi yang masih baru dan masih dalam tahap pengembangan. Studi literatur disini adalah mempelajari literatur yang berkaitan dengan *ontology*, *weighted directed acyclic graph*, OWL, dan *information retrieval*.
2. Observasi dan pengumpulan data
Mempelajari tugas akhir yang ada di Institut Teknologi Telkom sebagai bahan untuk melakukan perancangan *ontology*. Pada tahap ini juga dilakukan identifikasi kelemahan-kelemahan yang ada pada sistem yang telah ada.

3. Implementasi sistem

Tahap ini yaitu melakukan implementasi perancangan sistem yang telah dirancang sebelumnya dengan memasukkan *dataset* yang telah dikumpulkan dalam tahap sebelumnya ke dalam sistem. Tahap ini juga mengimplementasikan *ontology* dengan menggunakan Protégé 3.4.8 agar dapat digunakan oleh sistem dan juga menerapkan *weighted directed acyclic graph* untuk melakukan penghitungan kemiripan antar dokumen.

4. Pengujian sistem

Sistem akan dilakukan pengujian hasil pencarian *semantic search* berdasarkan dengan *precision*, *recall*, dan *F-Measure* serta melakukan analisis berdasarkan hasil pengujian sistem. Pengujian sistem dilakukan terhadap implementasi sistem yang telah dibuat. Pada tahap ini akan dilakukan penilaian apakah sistem yang dibuat dapat menghasilkan hasil yang lebih baik daripada sistem lama yang telah ada.

5. Pembuatan laporan

Tahap ini melakukan pendokumentasian atas semua kegiatan penelitian yang telah dilakukan sekaligus melaporkan hasil yang telah didapatkan.