

# 1. Pendahuluan

## 1.1 Latar Belakang

Setiap orang dapat mencari informasi yang diinginkannya dengan menginputkan *keyword* atau *query* didalam *search engine*. *Search engine* akan mengembalikan *query result* berupa dokumen yang relevan dengan *query* yang diinputkan oleh *user*. Dari dokumen yang dihasilkan oleh *search engine*, *user* diharapkan puas terhadap dokumen yang dihasilkan. Kepuasan *user* terhadap dokumen yang dihasilkan inilah yang menjadi sebuah tantangan untuk memberikan suatu pengukuran terhadap dokumen yang dihasilkan dari inputan *user*. Sehingga diperlukan sebuah pengukuran dimana sebelum *user* menginputkan jenis *query* tertentu, *user* akan mengetahui jenis *query* seperti apa yang sesuai dengan *search engine* yang digunakan. Dalam hal ini yang menjadi *search engine* adalah Google dan Yahoo, dan jenis *query* yang digunakan berupa *common query language* yang *query*nya berupa simple dan boolean.

Permasalahan yang akan timbul adalah bagaimana cara mengukur kepuasan *user* terhadap dokumen yang dihasilkan. Kepuasan yang dimaksud disini adalah dokumen yang dihasilkan akan relevan dengan keinginan *user*. Namun karena keterbatasan untuk mengetahui algoritma dan cara bagaimana *search engine* menghasilkan *query result*, maka untuk menguji kehandalan dari *query result* tersebut dengan cara menganalisis efektivitas dari masing-masing *search engine*.

Efektivitas *search engine* diperlukan untuk mengukur sejauh mana hasil *query* yang disebut dengan dokumen relevan sesuai dengan yang diinginkan *user*. Efektivitas dapat di ukur dari beberapa parameter pengukuran diantaranya adalah *precision*. *Precision* adalah jumlah kelompok dokumen relevan dari jumlah dokumen yang ditemukan oleh *system*. *Precision* merupakan penilaian biner yang melihat efektivitas *search engine* dari hasil *query* relevan atau tidak relevan, ditemukan atau tidak ditemukan informasi yang diinginkan [7]. Selain itu *precision* mudah untuk diterapkan dalam mengukur efektivitas *search engine* [1].

Selain mengukur efektivitas dengan menggunakan *precision*, diperlukan sebuah penilaian subjektif yang dilakukan oleh *user* yang disebut dengan *satisfaction*. *Satisfaction* ini merupakan *questioner* yang akan menganalisis secara langsung interaksi antara *user* dengan *search engine*. Hal ini diperlukan untuk mengetahui sejauh mana penilaian *user* terhadap *query result* yang dihasilkan *search engine* secara subjektif.

## 1.2 Perumusan Masalah

Permasalahan yang dijadikan objek penelitian dan pengembangan tugas akhir ini adalah :

1. Bagaimana cara menentukan *search engine* yang akan dianalisis.
2. Bagaimana cara menentukan *Common Query Language* yang akan digunakan sebagai data testing untuk *search engine*.
3. Bagaimana cara menganalisis *precision search engine* berdasarkan *common query language*.

4. Bagaimana cara melakukan dan menganalisis *test satisfaction* pada *user*.

Dalam penelitian tugas akhir ini, objek penelitian dibatasi dengan ruang lingkup sebagai berikut :

1. *Search engine* yang dianalisis mengacu pada [www.hitwise.com](http://www.hitwise.com) yaitu ada dua *search engine*; Google dan Yahoo. *Search engine* tersebut memiliki *document collection* berupa *free text document*.
2. *Query* yang digunakan berdasarkan *Common Query Language* berupa; *simple* dan *boolean queries*.
3. *Test Environment* dibatasi pada 20 hasil temuan pertama dalam *search engine*[17]. *Document collection* sebanyak 189 dokumen yang didapat dari 7 *sample query*.
4. Pengujian dilakukan oleh *system* dan *user testing*, *user testing* akan melakukan *satisfaction* terhadap Google dan Yahoo berupa mengisi *questioner* yang berhubungan dengan *test relevansi search engine*.
5. *Document collection* Google dan Yahoo didapat pada tanggal 22 Februari 2009, pukul 11.00-13.30.
6. *Query* yang digunakan sebagai *inputan* adalah bahasa inggris.
7. Dokumen yang dicari didalam *search engine* memiliki filetype:htm.
8. Menggunakan *stopword* bahasa inggris.

### 1.3 Tujuan

Menganalisis kesesuaian *Common Query Language* untuk *search engine* Google dan Yahoo, berdasarkan hasil *precision* dari *system* yang dibuat dan *satisfaction* dari *user*, yang diperoleh dari setiap *query input* pada *search engine* yang *document collectionnya* berupa *free-text document*.

### 1.4 Metodologi penyelesaian masalah

Metodologi pembahasan yang digunakan dalam penelitian Tugas Akhir ini adalah:

1. Studi literatur  
Mengumpulkan literatur-literatur yang berhubungan dengan TA baik cetak maupun elektronik
2. Seleksi *search engine* dan *query*  
Seleksi *search engine* dilakukan dengan mengacu kepada rilis yang dibuat oleh [www.hitwise.com](http://www.hitwise.com) pada tanggal 26 April 2008. Ada dua *search engine* yang paling banyak digunakan yaitu; Google dan Yahoo. *Query inputan* mengacu pada *Common Query Language* berupa; *simple queries* dan *boolean*.
3. *Test environment* dan *test user*  
*Test environment* dilakukan pada 20 hasil temuan pertama dalam *search engine*. *Query result* yang dihasilkan akan diukur *precisionnya*. *User testing* akan melakukan *satisfaction* terhadap ke dua *search engine* yaitu; Google dan Yahoo berupa mengisi *questioner* yang berhubungan dengan *test relevansi search engine*.
4. Analisis *search engine*  
Analisis dilakukan dari hasil *system* dan *satisfaction* dari *user*. Dimana pengujian akan menilai *precision* dari setiap *Common Query Language* dari

masing-masing *search engine* dengan skenario penentuan *term* sebagai query input dan hasil *satisfaction* dari *user* berupa *questioner*.

5. Penyusunan laporan dan penarikan kesimpulan