

# 1. PENDAHULUAN

## 1.1 Latar Belakang

Pada sebuah situs web terdapat arsitektur yang digunakan dalam membangun *website* tersebut. Arsitektur informasi merujuk ke disain (rancangan) yang berisi gambaran atau deskripsi terinci tentang struktur sebuah sistem informasi. Arsitektur informasi yaitu aspek untuk mengatur dan menyederhanakan informasi, merancang, mengintegrasikan dan menggabungkan informasi, menciptakan cara untuk menemukan, memahami, maupun saling bertukar dalam mengelola informasi [1].

Website Fakultas Universitas Telkom dibuat sebagai media dalam memberikan informasi dan memperlancar setiap civitas akademika yang ada di masing-masing fakultas. Dalam *website* tersebut tentu dibuat kontennya berdasarkan kebutuhan pengguna masing-masing fakultas. Dalam membuat konten tersebut pasti terdapat referensi/standar yang digunakan sebagai patokan. Dalam hal ini Universitas Telkom juga mempunyai standarisasi yang telah dibuat. Namun standarisasi yang dimiliki oleh Universitas Telkom baru dikeluarkan Januari 2016 dan jauh sebelumnya masing-masing fakultas telah memiliki website, Sehingga hasil yang diperoleh tidak secara maksimal maka hanya akan membantu masing-masing fakultas nantinya dalam melakukan pengupdetan terhadap struktur label website fakultas sesuai dengan standarisasi. Website Fakultas yang digunakan dalam penelitian ini yaitu Fakultas Industri dan Fakultas Informatika. Konten *website* yang dimaksudkan adalah struktur label website fakultas. Struktur informasi mendefinisikan cara utama dimana pengguna dapat menavigasi [2]. Berbicara mengenai struktur informasi berarti dapat dikaitkan dengan taksonomi informasi. Taksonomi dapat diartikan sebagai pengelompokan suatu hal berdasarkan hierarki (tingkatan) tertentu. Di mana taksonomi yang lebih tinggi bersifat lebih umum dan taksonomi yang lebih rendah bersifat lebih spesifik [3]. Dilakukan perhitungan kemiripan masing-masing struktur label website terhadap standarisasi. Untuk mengetahui pengimplementasian masing-masing website fakultas walaupun standarisasi dikeluarkan jauh sebelumnya.

Pencarian kemiripan antar fakultas dengan standarisasi dapat dilihat dari 3 sisi yaitu *label similarity*, *structural similarity*, dan *behavioural similarity*. Pada tugas akhir ini dilakukan dari sisi *structural similarity* dengan mengibaratkan setiap struktur label *website* masing-masing fakultas dan standarisasi konten *website* kedalam *graph*, sehingga diperoleh 3 *graph* yang berbeda. Pencarian kemiripan dilakukan dengan membandingkan masing-masing *graph* fakultas dengan *graph* standarisasi dengan menggunakan algoritma *exhaustive with pruning*.

Algoritma *exhaustive with pruning* melakukan pencarian semua kemungkinan pemetaan dan melakukan pemangkasan untuk menjaga hanya pemetaan dengan kesamaan tertinggi. Pemangkasan dilakukan jika jumlah kemungkinan pemetaan melebihi nilai *pruneat* maka dilakukan pengambilan kemungkinan pemetaan sebanyak nilai *pruneto*. Perhitungan Pemetaan antar label *node graph* yang dibandingkan menggunakan *syntactic similarity*. Kemudian dari hasil pemetaan dilakukan perhitungan *graph edit distance* yaitu, menghitung *node insertions/deletions*, menghitung *edge insertions/deletions*, menghitung *node substitutions*. Hasil dari perhitungan tersebut digunakan dalam menentukan kemiripan setiap fakultas dengan standarisasi dengan menggunakan perhitungan *graph edit similarity*.

## 1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, maka dapat dirumuskan sebagai berikut:

1. Bagaimana hasil dan analisis yang diperoleh dari perhitungan kemiripan struktur label pada masing-masing *website* fakultas terhadap Standarisasi Konten *Website* Telkom University menggunakan algoritma *exhaustive with pruning*?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah diatas, tujuan yang ingin dicapai pada tugas akhir ini adalah:

1. Mendapatkan hasil dan melakukan analisis hasil yang diperoleh dari perhitungan kemiripan struktur label pada masing-masing *website* fakultas terhadap Standarisasi Konten *Website* Telkom University menggunakan algoritma *exhaustive with pruning*.

## 1.4 Batasan Masalah

Dalam Tugas Akhir ini ada beberapa batasan masalah, yaitu:

1. Adapun struktur label pada *website* fakultas yang akan dilakukan analisis yaitu *Website* Fakultas Informatika dan Fakultas Industry yang ada di Universitas Telkom dengan Standarisasi Konten *Website* Fakultas Telkom University yang dikeluarkan oleh Direktorat Sistem Informasi Telkom University.
2. Pengambilan data struktur label pada *website* fakultas dilakukan dengan cara di crawling menggunakan scrapy pada tanggal 29 desember 2016.
3. Hasil dari scrapy diolah dalam Excel secara manual.
4. Pemetaan dilakukan dengan memperhatikan nilai *similarity* dari label/teks pada node *graph* yang dibandingkan tanpa mempedulikan arti/makna dari label/string tersebut.
5. Penelitian yang dilakukan hanya sampai pada perhitungan dan analisis kemiripan tanpa membandingkan dengan menggunakan algoritma lain.

## 1.5 Metodologi Penyelesaian Masalah

Metodologi dalam proses penyelesaian masalah ini terdiri dari beberapa tahapan, yaitu:

### 1. Studi Literatur

Pencarian materi-materi dan referensi yang berkaitan tentang arsitektur informasi, pengambilan struktur *website*, dan perhitungan tingkat kemiripan.

### 2. Pengambilan Struktur Label

Pada tahapan ini akan dilakukan pengambilan struktur label pada masing-masing *website* fakultas dengan di crawling menggunakan scrapy.

### 3. Perhitungan Kemiripan

Melakukan pencarian kemiripannya menggunakan *string edit distance*, *string edit similarity*, *graph edit distance* dan *graph edit similarity*.

### 4. Pengujian dan Analisis Hasil Pengujian

Tahap ini akan dilakukan pengujian terhadap program yang telah dibuat dalam menghitung tingkat kemiripan konten web.

### 5. Laporan Tugas Akhir

Pembuatan laporan tugas akhir berdasarkan hasil pengujian dan analisis yang dilakukan dengan mencari tingkat kemiripan struktur label pada masing-masing *website* fakultas dengan standarisasi konten web fakultas.