

# 1. Pendahuluan

## 1.1. Latar Belakang

*Website* semenjak penemuannya telah mengalami banyak perkembangan fungsi, salah satunya adalah sebagai sarana penyebaran informasi [2], namun organisasi informasi yang buruk menyebabkan informasi tersebut sulit ditemukan oleh *user* [3]. Oleh karena itu dalam melakukan pengembangan sebuah *website*, diperlukan pengorganisasian informasi untuk membuat informasi menjadi mudah di temukan oleh *user* [1]. Untuk membuat pengorganisasian informasi yang baik maka perlu diterapkan konsep-konsep Arsitektur Informasi pada *Web* dalam melakukan pengembangan sebuah *website*.

Arsitektur Informasi *Web* merupakan struktur rancangan (desain) *web* agar dapat tersusun suatu informasi yang tepat (terorganisasi) dan mudah ditemukan isinya [4]. Arsitektur Informasi pada *Web* memiliki beberapa komponen yaitu Sistem Organisasi, Sistem Navigasi, Sistem Pencarian, dan Sistem *Labeling* [1]. Salah satu komponen penting dalam Arsitektur Informasi pada *Web* adalah Sistem *Labeling*.

Label pada *website* digunakan untuk merepresentasi informasi yang ada pada *website* tersebut [1]. Sistem *Labeling* yang baik, diperlukan agar tidak terjadi kekeliruan penyampaian informasi oleh pemilik *website* kepada *user* serta memudahkan *user* dalam menemukan informasi [1]. Terdapat beberapa cara dalam mendesain Sistem *Labeling*, Salah satunya adalah dengan membandingkan dan mempelajari sistem *labeling* yang digunakan pada *website-website* yang sejenis dengan *website* yang akan dibuatkan sistem *labeling* nya [1].

Dalam Tugas Akhir ini akan dilakukan pembentukan rekomendasi sistem *labeling* dengan melakukan perbandingan *website*. *Website* yang akan digunakan sebagai perbandingan label adalah *website* universitas. Rekomendasi penggunaan sistem *labeling* pada *website* universitas diperlukan karena universitas sebagai sebuah institusi pendidikan saat ini menggunakan *website* sebagai salah satu media yang digunakan baik oleh intern universitas (mahasiswa, staff, maupun dosen) maupun publik untuk mendapatkan informasi [3]. Justifikasi pemilihan *website* yang digunakan sebagai *sample* perbandingan adalah universitas dengan peringkat 10 besar di *Webometrics* Indonesia ditambah dengan *website* Telkom Univeristy. Label yang akan dibandingkan pada *website* Universitas adalah label pada sistem navigasi, karena setiap *website*

universitas yang dijadikan *sample* memiliki sistem navigasi. Dengan melakukan perbandingan antar *website*, kita dapat menemukan pola label apa saja yang biasa digunakan pada sistem navigasi *website* universitas, sehingga hasil perbandingan dapat digunakan untuk membantu dalam pembuatan maupun perbaikan terhadap *website* universitas.

Dalam Tugas Akhir ini akan dilakukan ekstraksi data menggunakan Scrapy untuk mendapatkan label yang digunakan pada 11 *website* universitas di Indonesia. Scrapy adalah *framework* aplikasi untuk menelusuri dan mengekstrak data dari halaman *web*, Scrapy melakukan *scraping*, yaitu proses mengekstraksi data dari suatu halaman *web* dengan mendefinisikan *template* untuk setiap halaman yang hendak diekstrak [5]. Berdasarkan pada referensi [5], teknik ini dapat digunakan karena tingkat akurasi hasil ekstraksi tinggi, karena pendefinisian *template* yang seakurat mungkin.

Kemudian, melakukan perbandingan label dengan cara menghitung kemiripan label menggunakan metode *Levenshtein Distance*. Metode *Levenshtein Distance* digunakan karena metode ini dapat menghitung kemiripan antara 2 buah *string* berdasarkan pada huruf yang terkandung dalam *string* tersebut. Metode ini mendukung Tugas Akhir ini, karena label yang memiliki *string* berbeda walaupun secara makna memiliki makna yang sama akan dianggap sebagai label yang berbeda, karena akan dilihat seberapa persen penggunaan setiap label pada keseluruhan *website* universitas.

Untuk memperkuat hasil rekomendasi label, hasil perbandingan label antar *website* universitas akan dibandingkan dengan indeks yang diperoleh pada dokumen Standar Nasional Pendidikan Tinggi yang dikeluarkan oleh Dikti (SN Dikti) [6] dan dokumen Standar dan Prosedur Akreditasi Institusi Pendidikan Tinggi (AIPT) [7] serta dengan menggunakan *expert review* terhadap daftar label. Penggunaan [6] dan [7] sebagai penguat rekomendasi dapat dilakukan karena dokumen-dokumen tersebut menjelaskan aturan mengenai pendidikan tinggi yang ada di Indonesia salah satunya aturan mengenai universitas yang ada di Indonesia. *Expert review* dilakukan dengan meminta *review* kepada orang yang ahli di bidang *website* universitas dan pendidikan tinggi, hal ini perlu dilakukan karena ada kemungkinan tidak keseluruhan informasi mengenai pendidikan tinggi bisa didapatkan melalui [6] dan [7] sehingga untuk menutupi kelemahan tersebut diperlukan *expert review*.

*Output* hasil perbandingan label dan *expert review* akan berupa perankingan label-label. Hasil perankingan kemudian dijadikan sebagai rekomendasi sistem pelabelan.

## 1.2. Rumusan Masalah

Adapun rumusan masalah dalam Tugas Akhir ini sebagai berikut :

1. Bagaimana cara membandingkan label *website* universitas ?
2. Bagaimana cara melakukan pengolahan data hasil perbandingan label menjadi rekomendasi label untuk *website* universitas ?

## 1.3. Tujuan

Berdasarkan pada rumusan masalah, maka adapun tujuan dalam Tugas Akhir ini sebagai berikut :

1. Melakukan perancangan program untuk membandingkan label *website* universitas dengan metode *Levenshtein Distance*.
2. Melakukan pengolahan data dengan meranking label berdasarkan hasil perbandingan label untuk mendapatkan rekomendasi label.

## 1.4. Batasan Masalah

Batasan masalah pada Tugas Akhir ini sebagai berikut :

1. Pembuatan indeks pada [7] dan [6] dilakukan dengan cara melakukan *preprocessing* data teks terhadap dokumen, yaitu proses pengubahan bentuk data yang belum terstruktur menjadi data yang terstruktur sesuai dengan kebutuhan.
2. *Preprocessing* terhadap dokumen dibatasi sampai proses *stemming*, yaitu proses pengambilan kata dasar dari kata kerja yang berimbuhan untuk menghasilkan indeks.
3. Indeks yang di hasilkan melalui *preprocessing* dokumen hanya akan menghasilkan indeks dengan 1 buah kata saja, sehingga label yang dapat dibandingkan dengan indeks dokumen hanya label yang terdiri dari 1 buah kata.
4. *Website* yang digunakan sebagai *sample* terbatas pada *website* yang berbahasa Indonesia.
5. Label pada *website* yang dibandingkan terbatas pada label sebagai sistem navigasi khususnya navigasi global. Navigasi global pada umumnya terdiri dari *top menu*, *main menu* dan *footer menu* pada sebuah *website*.
6. *Preprocessing* terhadap data label hasil ekstraksi data dari *web* dilakukan dengan semi manual.

7. Pengembangan *tools* untuk *preprocessing* dokumen dan perbandingan label menggunakan Java.
8. Rekomendasi label hanya terbatas pada rekomendasi terhadap label navigasi global berupa *string* yang dapat digunakan pada *website* universitas, tidak membahas bagaimana sistem organisasi dari label navigasi atau bagaimana susunan penempatan label navigasi pada *website*.

### 1.5. Metodologi Penyelesaian Masalah

Metodologi penyelesaian masalah dalam penelitian ini meliputi :

1. Studi Literatur

Melakukan studi literatur terhadap materi arsitektur informasi pada *web* khususnya tentang sistem *labeling*, studi literatur terkait *preprocessing* dokumen, melakukan studi literatur terkait ekstraksi informasi pada *web* serta penggunaan metode *Levenshtein Distance* untuk perbandingan *string*.

2. Pendefinisian Label

Melakukan pendefinisian terhadap label dari *website* universitas yang akan dibandingkan.

3. Perancangan Program

Melakukan perancangan program untuk ekstraksi label pada *web* menggunakan Scrapy. Melakukan perancangan program untuk *preprocessing* dokumen. Melakukan perancangan program untuk membandingkan label menggunakan metode Levenshtein.

4. Pengumpulan dan Pengolahan Data

Melakukan pengumpulan data label dari setiap *website* universitas dengan melakukan ekstraksi informasi menggunakan Scrapy. Melakukan pengumpulan daftar kata pada dokumen sebagai indeks pada dokumen. Melakukan pengolahan data label dengan menggunakan metode *Levenshtein Distance*. Hasil perbandingan antar label kemudian diurutkan berdasarkan nilai *similarity*. Daftar label yang sudah diurutkan di validasi untuk mendapatkan rekomendasi label.

5. Pelaporan

Melakukan dokumentasi mulai dari tahap studi literatur, pendefinisian label, perancangan program, pengumpulan data, serta hasil penelitian dalam bentuk buku Tugas Akhir.

## 1.6. Sistematika Penulisan

Tugas akhir dengan judul “**Analisis Perbandingan Label pada Website Universitas di Indonesia sebagai Rekomendasi Sistem Pelabelan Web pada Website Universitas di Indonesia**” ini disusun dengan sistematika penulisan sebagai berikut :

1. Pendahuluan

Bab 1, menjelaskan latar belakang masalah yang diangkat, perumusan masalah, tujuan, batasan masalah dan metodologi penyelesaian masalah serta sistematika penulisan dari tugas akhir ini.

2. Tinjauan Pustaka

Bab 2 ini, menjelaskan teori-teori yang mendukung dan terkait dalam tugas akhir ini.

3. Perancangan Sistem

Bab 3 ini, memuat gambaran umum sistem, langkah-langkah pengumpulan data, langkah-langkah dalam melakukan *preprocessing* terhadap data label dan dokumen, langkah-langkah pemrosesan terhadap data label dan data dokumen, dan langkah-langkah dalam melakukan pemilihan label sebagai rekomendasi.

4. Pengujian dan Analisis

Bab 4, memuat hasil analisis terhadap data-data label yang dijadikan sebagai rekomendasi sistem pelabelan *website* universitas.

5. Kesimpulan dan Saran

Bab 5 ini, menjelaskan kesimpulan hasil dari penelitian tugas akhir yang dilakukan dan saran yang membangun untuk kedepannya.