Abstrak

Label pada *website* digunakan untuk merepresentasi informasi yang ada pada *website* tersebut. Agar tidak terjadi kekeliruan penyampaian informasi oleh pemilik *website* kepada *user*, maka *developer website* perlu untuk mendesain label yang baik sehingga label tersebut dapat merepresentasikan informasi yang ingin disampaikan pemiliki *website* kepada *user* [1].

Salah satu cara dalam mendesain sistem *labeling* pada *website* adalah dengan membandingkan dan mempelajari sistem *labeling* yang digunakan pada *website*-*website* kompetitor [1]. Dalam Tugas Akhir ini akan dilakukan rekomendasi sistem labeling pada *website* universitas dengan cara melakukan perbandingan label antar *website* universitas.

Metode yang digunakan adalah dengan melakukan ekstraksi data menggunakan Scrapy untuk mendapatkan label yang digunakan pada *website* universitas dengan peringkat 10 besar di *Web*ometrics Indonesia dan *website* Telkom University untuk kemudian membandingkan label dengan cara menghitung kemiripan label menggunakan metode *Levenshtein Distance*. Dilakukan penyederhanan jumlah label dengan menerapkan *rule*. Hasil perbandingan label tersebut kemudian dianalisis dengan melakukan perbandingan dengan indeks pada dokumen terkait aturan pendidikan tinggi dan melakukan *expert review*.

Hasil ekstraksi data label pada 11 *website* universitas menghasilkan 854 label setelah dilakukan *preprocessing*. Perbandingan label dapat menunjukkan persentase kemunculan label pada setiap *website* universitas maupun keseluruhan *website* universitas, jumlah label yang di hasilkan setelah proses perbandingan adalah 706 label. Analisis terhadap hasil dengan melakukan perbandingan dengan indeks dokumen aturan pendidikan tinggi menunjukkan penggunaan dokumen aturan pendidikan tinggi dapat membantu untuk mempertimbangkan label apa saja yang perlu untuk di tampilkan pada *website* universitas di Indonesia. Analisis terhadap hasil dengan melakukan *expert review* menunjukkan 25 dari 706 label tidak digunakan untuk label yang di tampilkan pada *website* universitas.

Kesimpulan akhir dari penelitian ini adalah hasil perbandingan label dapat digunakan sebagai rekomendasikan standar sistem *labeling* yang dapat digunakan untuk membangun atau memperbaiki *website* universitas di Indonesia dengan cara data-data tersebut di pilah-pilah kembali sesuai dengan kebutuhan masing-masing dari pembuat *website* atau pembuat standar *website* di masing-masing universitas.

Kata Kunci : arsitektur informasi *web*, sistem *labeling*, perbandingan label, ekstraksi informasi, *levenshtein distance*