

## 1. Pendahuluan

Pada bagian ini berisi empat sub-bagian, yaitu Latar Belakang, Topik dan Batasannya, Tujuan, dan Organisasi Tulisan. Berikut penjelasannya.

### Latar Belakang

Setiap perguruan tinggi memiliki media untuk menyajikan informasi terkait perguruan tinggi tersebut. Informasi adalah penerangan, pemberitahuan, kabar, dan berita terkait suatu hal [8]. Salah satu media untuk menyajikan informasi tersebut adalah situs web. Situs web adalah program komputer yang menjalankan peladen yang menyediakan akses kepada beberapa laman [8].

Salah satu elemen penting di dalam situs web adalah *Information Architecture* dari website itu sendiri. *Information Architecture* berfokus dalam perencanaan struktur informasi yang akan di sajikan kepada pengguna. *Information Architecture* berfokus dalam pengorganisasian, pengurutan, dan pengelompokan informasi untuk memaksimalkan *User Experience* [10]. Salah satu tahap penting dalam merancang *Information Architecture* adalah perancangan *headings label* yang tepat. *Headings label* adalah label yang mendeskripsikan informasi yang tertera di bawahnya, contohnya adalah judul dalam suatu berita. *Headings label* yang kurang tepat pada suatu informasi dapat memberikan *User Experience* yang buruk terhadap pengguna. *Labeling* adalah bentuk representasi, kita menggunakan label untuk mempresentasikan bagian besar dari sebuah informasi [10]. *Labeling* yang buruk dapat membuat pengguna merasa tidak mendapatkan informasi yang mereka cari pada website yang digunakan, sehingga kepercayaan pengguna terhadap informasi yang terdapat pada website menurun. Untuk meminimalisir masalah tersebut, kita harus mencoba sebaik mungkin untuk membuat *label* yang memiliki pemahaman bahasa yang sama dengan pengguna. Salah satu contoh ketidakcocokan *headings label* dengan konten dibawahnya pada situs web Telkom University dapat dilihat pada tautan berikut <https://telkomuniversity.ac.id/telkom-university-berbagi-dengan-500-anak-yatim/>, pada link tersebut terdapat ketidakcocokan antara *headings label* dan konten dimana *headings label* tidak terlalu mendeskripsikan konten dibawahnya, hal ini dibuktikan atas survey terhadap 5 responden, dimana 3 dari 5 responden menjawab bahwa judul berita pada artikel tersebut kurang mendeskripsikan konten dibawahnya. Oleh sebab itu dalam merancang sistem pelabelan yang baik, sumber pelabelan yang tepat adalah dari pengguna itu sendiri. Jika dapat diciptakan sistem pelabelan yang bersumber *directly from user* maka itulah sistem pelabelan terbaik yang ada [5]. Penerapan sistem rekomendasi *headings label* yang bersifat *directly from user* seharusnya akan meningkatkan ketepatan *headings label* pada situs perguruan tinggi Telkom University.

Salah satu cara untuk mendapatkan informasi dari pengguna adalah melalui *Crowdsourcing*. *Crowdsourcing* adalah salah satu cara untuk mengumpulkan informasi pada kalangan tertentu, sehingga memungkinkan pengumpulan informasi dengan cepat dan dalam skala yang besar. Dalam penerapannya *Crowdsourcing* memerlukan *crowdworker*. *Crowdworker* adalah pengguna produk yang sekiranya memiliki kualifikasi untuk dapat dimanfaatkan untuk mengerjakan suatu masalah, *Crowdworker* berperan sebagai pelaksana tugas seperti melakukan masukan, pendapat atau penilaian terhadap tugas yang di berikan [2]. Pada *Crowdsourcing* pemilihan *crowdworker* yang tepat sangat menentukan kualitas pengerjaan tugas yang di hasilkan. Oleh sebab itu diperlukan kualifikasi tertentu untuk memilih *crowdworker* [4]. Apabila tidak memungkinkan untuk dilakukan seleksi *crowdworker*, maka dapat pula dilakukan seleksi input *crowdworker* dimana sistem harus memiliki mekanisme penyeleksian input dari *crowdworker* [4]. Salah satu mekanisme seleksi input *crowdworker* adalah menerapkan sistem yang dapat menerapkan sistem *Majority Decision* [14]. Sistem dengan *majority decision* adalah sistem yang dapat menyimpulkan keputusan dari mayoritas, yang pada kasus ini adalah menyimpulkan saran saran dari *crowdworker* akan judul berita mana yang cocok. Salah satu metode untuk mendapatkan jawaban mayoritas *crowdworker* adalah *tf-idf* [14]. Dalam penerapannya *Crowdsourcing* juga memerlukan *Natural Language Processing* untuk mengolah inputan yang di dapatkan dari *crowdworker*. *Natural Language Processing* adalah beragam teknik komputasi untuk menganalisa dan merepresentasikan bahasa alami agar dapat di proses komputer [9]. Penggunaan *Natural Language Processing* adalah untuk mengubah inputan *crowdworker* yang bersifat teks tidak terstruktur menjadi teks terstruktur sehingga mempermudah pengolahan data input oleh komputer. *Text Mining* sangat bergantung terhadap teknologi *Natural Language Processing* dalam penerapannya [6]. *Text Mining* adalah proses ekstraksi informasi dari teks yang tidak terstruktur. Teks disini adalah bahasa manusia yang belum terstruktur seperti pesan. Sebelum penambangan teks dilakukan maka kita perlu menerapkan praproses kedalam program penambangan teks. Hal ini dikarenakan representasi teks dalam bentuk alaminya sangat sulit untuk dimengerti oleh komputer. Oleh sebab itu perlu dirancang representasi teks agar lebih mudah diproses dan dimengerti oleh program [13].

### Perumusan masalah

Adapun permasalahan yang ada pada penelitian ini adalah bagaimana menerapkan *Crowdsourcing* dalam rekomendasi *headings label* informasi situs web Telkom University. Untuk menyelesaikan masalah tersebut seleksi terhadap input *crowdworker* pada penelitian ini juga harus diperhatikan agar input terbaik yang akan direkomendasikan.

### **Batasan Masalah**

Adapun batasan masalah pada penelitian ini adalah sebagai berikut:

- Sistem akan diujikan pada situs web berita Telkom University (<https://telkomuniversity.ac.id/all-news/>)
- Artikel yang diuji adalah 10 artikel pada rentang tanggal 2 oktober 2018 – 19 oktober 2018
- Rekomendasi pada tiap artikel pengujian adalah 11 rekomendasi

### **Tujuan**

Adapun tujuan dari penelitian ini yaitu menerapkan *Crowdsourcing* untuk rekomendasi *Headings Label* informasi situs web Telkom University. Analisis akan input yang dihasilkan *Crowdworker* juga diperlukan agar dapat disimpulkan seleksi input yang tepat dalam rekomendasi *headings label* informasi. Penerapan algoritma *Text Mining* pada inputan dari *crowdworker* dan menganalisa kualitas *headings label* yang di hasilkan dari *Crowdsourcing*.

### **Organisasi Tulisan**

Bagian selanjutnya dari paper ini akan membahas studi terkait akan penelitian ini yaitu pada bagian II. Selanjutnya perancangan dari sistem yang akan dibuat akan dijabarkan pada bagian III. Hasil dan evaluasi dari perancangan sistem pada bagian III akan dijabarkan pada bagian IV. Kesimpulan dari penelitian ini akan dijelaskan pada bagian V dari paper ini.