
Abstrak

Seiring dengan meningkatnya penggunaan internet, maka penggunaan situs web sebagai sarana penyaji informasi juga ikut meningkat. Salah satu penyaji informasi tersebut adalah perguruan tinggi. Banyak perguruan tinggi yang menyajikan informasi terkait perguruan tinggi melalui situs web. Dalam pelaksanaannya banyak situs web perguruan tinggi yang memiliki *headings label* kurang tepat dengan informasi yang di sajikan. *Headings label* adalah label yang mendeskripsikan informasi yang tertera di bawahnya, contohnya adalah judul dalam suatu berita. Untuk mengatasi masalah ini dapat di gunakan metode pelabelan yang bersifat *directly from user* yang berarti pelabelan bersumber langsung dari pengguna situs web tersebut. Untuk mendapatkan *headings label* yang bersumber dari pengguna dapat digunakan metode *Crowdsourcing*. *Crowdsourcing* adalah metode pemberian tugas terhadap *crowdworker* untuk mendapatkan informasi dan hasil dengan waktu yang relatif cepat. *Crowdworker* adalah kelompok tertentu yang di pilih berdasarkan beberapa kriteria, yang dianggap cocok untuk mengerjakan tugas yang ingin di berikan. Metode *Crowdsourcing* yang di padukan dengan *Natural Language Processing* dan *Text Mining* dapat di manfaatkan sebagai sumber pelabelan yang bersifat *directly from user*. Dari sistem yang dibuat 7 dari 10 judul hasil rekomendasi *crowdworker* berhasil mengalahkan judul asli pada situs web Telkom University. Sistem juga dapat menampilkan token – token pada tiap artikel yang memiliki relevansi tinggi terhadap artikel

Kata kunci : crowdsourcing, natural language processing, headings label, text mining, crowdworker, information architecture
