

PENERAPAN *NAMED ENTITY RECOGNITION* (NER) UNTUK EKSTRAKSI OTOMATIS ENTITAS PADA TEKS BERITA ONLINE PEMERINTAH DAERAH

Nama Mahasiswa : Rama Otari
Nim : 6304221463
Dosen Pembimbing : Elvi Rahmi, M.Kom.

ABSTRAK

Berita pemerintah daerah memiliki peran penting sebagai sumber informasi publik, namun umumnya masih disajikan dalam bentuk teks tidak terstruktur sehingga menyulitkan proses pencarian dan analisis informasi. Penelitian ini bertujuan menerapkan metode *Named Entity Recognition* (NER) untuk mengekstraksi entitas penting sekaligus membangun sistem berbasis web yang mengintegrasikan proses *scraping* berita, prapemrosesan teks, ekstraksi entitas, dan penyajian hasil secara terstruktur. Metode yang digunakan memanfaatkan model *spaCy pretrained* berbasis *multilingual* dan *IndoBERT pretrained* untuk mengenali entitas PER, ORG, LOC, dan DATE pada berita Detik.com. Sistem dikembangkan menggunakan *framework Flask* dan dievaluasi dengan membandingkan hasil ekstraksi terhadap data *ground truth* menggunakan metrik *precision*, *recall*, dan *F1-score*. Hasil pengujian menunjukkan nilai *precision* sebesar 67%, *recall* sebesar 76,2%, dan *F1-score* sebesar 71,2%. Kinerja terbaik diperoleh pada entitas PER dan DATE, sementara entitas LOC dan ORG masih menunjukkan hasil yang kurang optimal. Secara keseluruhan, sistem ini dapat dimanfaatkan sebagai alat bantu awal dalam analisis berita pemerintah daerah secara lebih efisien dan sistematis.

Kata kunci: *Named Entity Recognition, spaCy pretrained, IndoBERT pretrained, Berita Pemerintah Daerah, Ekstraksi Entitas*

**IMPLEMENTATION OF NAMED ENTITY RECOGNITION (NER) FOR
AUTOMATIC EXTRACTION OF ENTITIES IN REGIONAL
GOVERNMENT ONLINE NEWS TEXTS**

Name	: Rama Otari
Student ID	6304221463
Supervisor	: Elvi Rahmi, M.Kom.

ABSTRACT

Local government news is an important source of public information, but it is generally presented in unstructured text, making it difficult to search and analyze. This study aims to apply Named Entity Recognition (NER) to extract key entities and develop a web-based system that integrates news scraping, text preprocessing, entity extraction, and structured result presentation. The method utilizes a multilingual pretrained spaCy model and an IndoBERT pretrained model to recognize entities such as PER (person), ORG (organization), LOC (location), and DATE in news from Detik.com. The system was developed using the Flask framework and evaluated by comparing extraction results with ground truth data using precision, recall, and F1-score metrics. The evaluation results show a precision of 67%, recall of 76.2%, and an F1-score of 71.2%. The best performance was achieved for PER and DATE entities, while LOC and ORG entities showed less optimal results. Overall, this system can be used as an initial tool to support the analysis of local government news more efficiently and systematically.

Keywords: *Named Entity Recognition, spaCy pretrained, IndoBERT pretrained, Local Government News, Entity Extraction*