

ABSTRAK

Sulitnya mengelola dokumen administratif di Kapel St. Yohanes Rasul yang dilakukan secara manual sehingga rentan terhadap *human error* seperti kesalahan pencatatan hingga kehilangan dokumen. Penelitian ini bertujuan mengatasi masalah tersebut dengan mengolah data arsip milik kapel yang berformat PDF dan citra (*image*), untuk diklasifikasikan ke dalam tiga kelas (surat perizinan, undangan, dan peminjaman) serta mengekstrak informasi penting di dalamnya. Sistem terintegrasi ini menerapkan teknologi *Optical Character Recognition* (OCR) menggunakan PyTesseract dan *Natural Language Processing* (NLP) berbasis model *fine-tuned* IndoBERT. Hasil pengujian model menunjukkan performa yang sangat tinggi, di mana sistem klasifikasi mencapai akurasi 93,6% pada data uji baru, sedangkan model *Named Entity Recognition* (NER) berhasil meraih F1-Score sebesar 0.989. Pengujian penerimaan pengguna juga menunjukkan bahwa sistem ini sangat efektif, berhasil meningkatkan efisiensi proses pengelolaan dokumen hingga 69,6% berdasarkan analisis waktu dan mendapatkan skor System Usability Scale (SUS) 77.5 yang menandakan *usability* yang baik. Dapat disimpulkan, sistem ini terbukti fungsional, efisien, dan dapat diterima dengan baik oleh pengguna untuk meningkatkan akurasi dan efisiensi pengelolaan dokumen administratif.

Kata Kunci: Pengelolaan dokumen, *Optical Character Recognition*, *Natural Language Processing*