



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



# **PENGEMBANGAN SISTEM CERDAS UNTUK MANAJEMEN, ANALISIS, DAN OTOMATISASI DOKUMEN BERBASIS WEB**

**SUBJUDUL:**

**Analisis Teks Dokumen Administratif untuk Peringkasan  
Menggunakan NER-NLP**

**SKRIPSI**

**Dibuat untuk Melengkapi Syarat-Syarat yang Diperlukan untuk  
Memperoleh Diploma Empat Politeknik**

**Muhammad Nathan Sunarto**

**2207412024**

**TI-CCIT-8A**

**PROGRAM STUDI TEKNIK INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER  
POLITEKNIK NEGERI JAKARTA**

**2026**



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta



**PENGEMBANGAN SISTEM CERDAS UNTUK  
MANAJEMEN, ANALISIS, DAN OTOMATISASI  
DOKUMEN BERBASIS WEB**

**SUBJUDUL:**

**Analisis Teks Dokumen Administratif untuk Peringkasan  
Menggunakan NER-NLP**

**SKRIPSI**

**Muhammad Nathan Sunarto**

**2207412024**

**PROGRAM STUDI TEKNIK INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA DAN KOMPUTER  
POLITEKNIK NEGERI JAKARTA**

**2026**



## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

### SURAT PERNYATAAN BEBAS PLAGIARISME

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Nathan Sunarto  
NIM : 2207412024  
Jurusan/Program Studi : Teknik Informatika dan Komputer / Teknik Informatika  
Judul Skripsi : PENGEMBANGAN SISTEM CERDAS UNTUK MANAJEMEN, ANALISIS, DAN OTOMATISASI DOKUMEN BERBASIS WEB  
Sub Judul : Analisis Teks Dokumen Administratif untuk Ringkasan Menggunakan NER-NLP

Menyatakan dengan sebenarnya bahwa skripsi ini benar-benar merupakan hasil karya saya sendiri, bebas dari peniruan terhadap karya dari orang lain. Kutipan pendapat dan tulisan orang lain ditunjuk sesuai dengan cara-cara penulisan karya ilmiah yang berlaku.

Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa dalam skripsi ini terkandung ciri-ciri plagiat dan bentuk-bentuk peniruan lain yang dianggap melanggar peraturan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Depok, 08 Juni 2026

Yang membuat pernyataan,



Muhammad Nathan Sunarto

NIM 2207412024



**© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta**

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**SURAT PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK  
KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Politeknik Negeri Jakarta, saya bertanda tangan di bawah ini:

Nama : Muhammad Nathan Sunarto

NIM : 2207412024

Jurusan/Program Studi : T. Informatika dan Komputer/Teknik Informatika

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Politeknik Negeri Jakarta Hak Bebas Royalti Non-Eksklusif atas karya ilmiah saya yang berjudul:

**Analisis Teks Dokumen Administratif untuk Peringkasan**

**Menggunakan NER-NLP.**

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non Eksklusif ini, Politeknik Negeri Jakarta berhak menyimpan, mengalihmediakan/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Depok, 30 Juni 2026

Yang membuat pernyataan



Muhammad Nathan Sunarto

NIM.2207412024



© Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

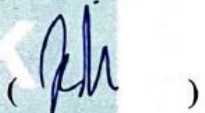
LEMBAR PENGESAHAN

Skripsi diajukan oleh:

Nama : Muhammad Nathan Sunarto  
NIM : 2207412024  
Program Studi : Teknik Informatika  
Judul Skripsi : Pengembangan Sistem Cerdas Untuk Manajemen, Analisis, Dan Otomatisasi Dokumen Berbasis Web  
Judul Skripsi : Analisis Dokumen Administratif Untuk Peringkasan Menggunakan NER-NLP

Pada hari ini Selasa Tanggal 23, bulan Juni tahun 2026, telah diadakan Sidang Skripsi untuk saudara Muhammad Nathan Sunarto dan dinyatakan **LULUS**

Disahkan oleh

Pembimbing I : Risna Sari, S.Kom., M.T.I.. (  )  
Penguji I : Dr. Dewi Yanti Liliana, S.Kom., M.Kom. (  )  
Penguji II : Rizki Elisa Nalawati, S.T., M.T. (  )  
Penguji III : Ariawan Andi Suhandana, S.Kom., M.T.I. (  )

Mengetahui:

Jurusan Teknik Informatika dan Komputer  
Ketua



Dr. Anita Hidayati, S.Kom., M.Kom.  
NIP. 19790832003122003



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## **Analisis Teks Dokumen Administratif untuk Identifikasi Bagian Teks Variabel Dinamis**

### **ABSTRAK**

*Dokumen administratif merupakan salah satu sumber informasi penting yang digunakan dalam berbagai aktivitas organisasi dan institusi. Namun, proses identifikasi informasi penting pada dokumen administratif masih banyak dilakukan secara manual sehingga membutuhkan waktu yang cukup lama dan berpotensi menimbulkan kesalahan. Penelitian ini bertujuan mengembangkan sistem analisis dokumen administratif yang mampu melakukan ekstraksi informasi penting dan menghasilkan ringkasan dokumen secara otomatis. Sistem yang dikembangkan memanfaatkan PaddleOCR untuk mengekstraksi teks dari dokumen PDF digital, PDF hasil pemindaian, dan DOCX, serta model IndoBERT yang telah melalui proses fine-tuning untuk melakukan Named Entity Recognition (NER) terhadap entitas administratif seperti nomor surat, jenis dokumen, tanggal, pengirim, penerima, perihal, lokasi, waktu, dan isi dokumen. Selain itu, sistem mengimplementasikan LayoutLMv3 untuk mendukung ekstraksi informasi pada dokumen yang mengandung tabel dan menghasilkan ringkasan dokumen berdasarkan hasil identifikasi entitas. Model NER dilatih melalui 13 skenario eksperimen dengan berbagai strategi optimasi, dan konfigurasi terbaik diperoleh pada Run ID #13 dengan nilai F1-score sebesar 88,56%. Sistem kemudian diimplementasikan dalam bentuk REST API menggunakan FastAPI dan di-deploy pada AWS EC2. Hasil pengujian menunjukkan bahwa sistem mampu mengidentifikasi informasi penting secara otomatis dan menghasilkan ringkasan dokumen yang relevan, sehingga dapat membantu meningkatkan efisiensi pengelolaan dan analisis dokumen administratif.*

**POLITEKNIK  
NEGERI  
JAKARTA**



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## KATA PENGANTAR

Puji dan syukur penulis panjatkan kehadiran Allah SWT atas segala rahmat, hidayah, dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Analisis Dokumen Administratif untuk Identifikasi Teks Variabel Dinamis Menggunakan OCR dan Named Entity Recognition Berbasis IndoBERT" dengan baik dan tepat waktu.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Terapan (S.Tr.Kom.) pada Program Studi Teknik Informatika, Jurusan Teknik Informatika dan Komputer, Politeknik Negeri Jakarta. Penelitian ini berfokus pada pengembangan sistem yang mampu melakukan ekstraksi teks dari dokumen administratif menggunakan PaddleOCR, mengidentifikasi entitas penting seperti nomor surat, jenis dokumen, tanggal, pengirim, penerima, perihal, dan lokasi menggunakan model IndoBERT yang telah melalui proses fine-tuning, serta menghasilkan ringkasan dokumen secara otomatis melalui layanan REST API yang diimplementasikan pada lingkungan AWS EC2. Sistem ini diharapkan dapat membantu meningkatkan efisiensi proses analisis dokumen administratif serta mengurangi waktu yang dibutuhkan pengguna dalam memahami informasi penting yang terkandung dalam dokumen.

Dalam proses penyusunan skripsi ini, penulis menyadari bahwa keberhasilannya tidak terlepas dari dukungan, bimbingan, dan bantuan berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dr. Anita Hidayati, S.Kom., M.Kom., Ketua Jurusan Teknik Informatika dan Komputer Politeknik Negeri Jakarta, atas dukungan dan arahan yang diberikan kepada seluruh mahasiswa.
2. Ibu Euis Oktavianti, S.Si., M.T.I., Kepala Program Studi Teknik Informatika Politeknik Negeri Jakarta, atas bimbingan dan fasilitas akademik yang menunjang proses penelitian.



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

3. Ibu Risna Sari sebagai Dosen Pembimbing, yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penelitian dan penulisan skripsi ini.
4. Seluruh Dosen dan Staf Akademik Jurusan Teknik Informatika dan Komputer Politeknik Negeri Jakarta, atas ilmu dan pengalaman yang telah diberikan kepada penulis selama masa perkuliahan.
5. Kedua orang tua dan seluruh keluarga penulis, yang senantiasa memberikan doa, semangat, dukungan moral, dan materi yang tidak ternilai selama penulis menempuh pendidikan.
6. Seluruh rekan mahasiswa seperjuangan angkatan 2022, atas kebersamaan, diskusi, dan semangat yang selalu menginspirasi penulis selama masa studi.
7. Seluruh jajaran team FC Barcelona yang telah menghibur, memberikan semangat serta dukungan dengan cara memenangkan liga.
8. Semua pihak yang telah membantu dan mendukung yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, penulis mengharapkan kritik dan saran yang bersifat membangun demi penyempurnaan skripsi ini di masa mendatang. Besar harapan penulis agar skripsi ini dapat memberikan manfaat bagi pembaca, khususnya bagi pengembangan ilmu pengetahuan di bidang Teknik Informatika dan Rekayasa Perangkat Lunak.

Jakarta, 9 Juni 2026

Muhammad Nathan Sunarto

NIM 2207412024



## DAFTAR ISI

|                                                                                |    |
|--------------------------------------------------------------------------------|----|
| BAB I PENDAHULUAN .....                                                        | 1  |
| 1.1. Latar Belakang Masalah .....                                              | 1  |
| 1.2. Perumusan Masalah .....                                                   | 3  |
| 1.3. Batasan Masalah .....                                                     | 4  |
| 1.4. Tujuan Penelitian .....                                                   | 5  |
| 1.5. Manfaat Penelitian .....                                                  | 5  |
| 1.6. Sistematika Penulisan .....                                               | 6  |
| BAB II TINJAUAN PUSTAKA .....                                                  | 7  |
| 2.1. Karakteristik Dokumen Administratif JTik sebagai Sumber Data .....        | 7  |
| 2.2. Ekstraksi Teks dan Tata Letak dari PDF dengan PyMuPDF dan PaddleOCR ..... | 9  |
| 2.3. Analisis Tata Letak Dokumen dan Deteksi Bagian Struktural .....           | 12 |
| 2.4. Named Entity Recognition (NER) untuk Entitas Administratif .....          | 12 |
| 2.5. Perbandingan dengan Penelitian Terdahulu .....                            | 24 |
| BAB III METODOLOGI PENELITIAN .....                                            | 27 |
| 3.1. Rancangan Penelitian .....                                                | 27 |
| 3.2. Tahapan Penelitian .....                                                  | 27 |
| 3.3. Objek Penelitian .....                                                    | 30 |
| 3.4. Sumber Data .....                                                         | 32 |
| 3.5. Teknik Pengumpulan Data .....                                             | 32 |
| 3.6. Teknik Analisis Data .....                                                | 33 |
| BAB IV HASIL DAN PEMBAHASAN .....                                              | 34 |
| 4.1 Analisis Kebutuhan .....                                                   | 34 |



## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

|                                |     |
|--------------------------------|-----|
| 4.2 Perancangan Sistem .....   | 40  |
| 4.3 Permodelan Algoritma ..... | 49  |
| 4.4 Pengujian .....            | 86  |
| BAB V PENUTUP .....            | 111 |
| 5.1 Kesimpulan .....           | 111 |
| 5.2 Saran .....                | 111 |
| DAFTAR PUSTAKA .....           | 112 |
| LAMPIRAN .....                 | 116 |





**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**DAFTAR TABEL**

|            |                                                   |     |
|------------|---------------------------------------------------|-----|
| Tabel 4.1  | Tabel Kebutuhan Functional .....                  | 37  |
| Tabel 4.2  | Tabel Kebutuhan non functional .....              | 39  |
| Tabel 4.3  | Matriks Prioritas Kebutuhan .....                 | 40  |
| Tabel 4.4  | Tabel Daftar Endpoint API .....                   | 49  |
| Tabel 4.5  | Tabel Hasil Pengujian Functional Blackbox .....   | 88  |
| Tabel 4.6  | Hasil Evaluasi Model .....                        | 92  |
| Tabel 4.7  | Tabel Confusion Matriks Per Label .....           | 93  |
| Tabel 4.8  | Tabel Skenario dan Hasil UAT .....                | 100 |
| Tabel 4.9  | Tabel Tanggapan Terbuka Peserta UAT .....         | 103 |
| Tabel 4.10 | Tabel Evaluasi Pencapaian Tujuan Penelitian ..... | 108 |

**POLITEKNIK  
NEGERI  
JAKARTA**



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**DAFTAR GAMBAR**

|                                                   |    |
|---------------------------------------------------|----|
| Gambar 3.1 Diagram AI engineering Lifecycle ..... | 28 |
| Gambar 4.1 Diagram Arsitektur System .....        | 29 |
| Gambar 4.2 Diagram Pipeline System .....          | 44 |
| Gambar 4.3 Diagram Arsitektur API .....           | 47 |
| Gambar 4.4 Diagram Endpoint API .....             | 48 |
| Gambar 4.5 Hasil Confusion Matriks .....          | 96 |





**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## **BAB I**

### **PENDAHULUAN**

#### **1.1. Latar Belakang Masalah**

Dokumen administratif merupakan salah satu media utama yang digunakan oleh institusi pendidikan, pemerintahan, maupun organisasi dalam menjalankan proses administrasi dan komunikasi formal. Berbagai jenis dokumen seperti surat undangan, surat permohonan, surat keterangan, nota dinas, pengumuman, serta dokumen akademik lainnya mengandung informasi penting yang digunakan sebagai dasar pengambilan keputusan dan pelaksanaan kegiatan organisasi. Informasi tersebut umumnya berupa nomor surat, tanggal dokumen, pengirim, penerima, perihal, lokasi kegiatan, dan jenis dokumen. Keakuratan dalam mengidentifikasi informasi tersebut sangat penting karena berkaitan langsung dengan validitas dokumen dan efektivitas proses administrasi.

Seiring dengan meningkatnya transformasi digital di berbagai institusi, jumlah dokumen administratif yang dihasilkan dan diproses setiap hari terus bertambah. Pemerintah Indonesia melalui kebijakan Sistem Informasi Kearsipan Nasional (SKN) 2025-2029 juga mendorong pengintegrasian sistem informasi kearsipan dengan Sistem Pemerintahan Berbasis Elektronik (SPBE), yang semakin mempercepat volume digitalisasi dokumen administrasi. Meskipun sebagian besar dokumen telah tersedia dalam format digital, proses identifikasi informasi penting masih banyak dilakukan secara manual. Staf administrasi sering kali harus membaca keseluruhan isi dokumen untuk menemukan informasi yang diperlukan sebelum melakukan pengarsipan, pelaporan, atau penyusunan tindak lanjut. Proses tersebut membutuhkan waktu yang relatif lama, terutama ketika jumlah dokumen yang harus diproses sangat banyak. Selain itu, proses manual juga berpotensi menimbulkan kesalahan identifikasi informasi akibat kelelahan maupun perbedaan interpretasi pengguna terhadap isi dokumen. Penelitian oleh (A. Yani 2024) menunjukkan bahwa sistem otomatisasi berbasis web mampu meningkatkan efisiensi pekerjaan administrasi secara signifikan dibandingkan dengan proses manual.



### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian , penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Permasalahan lain yang sering ditemukan adalah keberagaman format dokumen administratif. Meskipun memiliki struktur yang relatif serupa, setiap jenis dokumen dapat memiliki tata letak, gaya penulisan, serta susunan informasi yang berbeda. Variasi tersebut menyebabkan proses ekstraksi informasi secara otomatis menjadi lebih kompleks karena sistem harus mampu mengenali pola informasi yang muncul pada berbagai bentuk dokumen. Pada dokumen hasil pemindaian (scanned document), tantangan menjadi lebih besar karena informasi yang tersedia masih berbentuk citra dan perlu dikonversi terlebih dahulu menjadi teks yang dapat diproses oleh komputer.

Perkembangan teknologi Artificial Intelligence (AI), khususnya pada bidang Natural Language Processing (NLP) dan Computer Vision, membuka peluang untuk mengatasi permasalahan tersebut. sistem manajemen dokumen adaptif yang dapat belajar dari struktur dokumen baru secara terus-menerus untuk meningkatkan efisiensi pengelolaan dokumen dalam jangka panjang. Salah satu teknologi yang banyak digunakan adalah Optical Character Recognition (OCR), yaitu teknologi yang mampu mengubah teks pada gambar atau dokumen hasil pemindaian menjadi teks digital yang dapat diproses lebih lanjut. Setelah teks berhasil diekstraksi, proses identifikasi informasi penting dapat dilakukan menggunakan teknik Named Entity Recognition (NER), yaitu salah satu tugas dalam NLP yang bertujuan untuk mengenali dan mengklasifikasikan entitas tertentu ke dalam kategori yang telah ditentukan, seperti nama orang, organisasi, lokasi, tanggal, maupun informasi administratif lainnya. Selain itu, teknologi text summarization dapat dimanfaatkan untuk menghasilkan ringkasan dokumen secara otomatis sehingga pengguna dapat memahami isi dokumen dengan lebih cepat dan efisien.

Penelitian mengenai ekstraksi informasi dokumen menggunakan pendekatan OCR dan NER telah menunjukkan hasil yang menjanjikan dalam membantu proses otomatisasi dokumen. Sebuah penelitian pada tahun 2025 berhasil mengimplementasikan sistem OCR dan NLP menggunakan PyTesseract dan model fine-tuned IndoBERT untuk pengelolaan dokumen administratif, yang mencapai akurasi klasifikasi 93,6% dan F1-Score NER sebesar 0,989. Namun, sebagian besar penelitian masih berfokus pada identifikasi entitas umum dan

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

belum secara khusus menangani kebutuhan identifikasi teks variabel dinamis pada dokumen administratif berbahasa Indonesia. Padahal, informasi seperti nomor surat, perihal, pengirim, penerima, dan jenis dokumen merupakan komponen penting yang sering digunakan dalam proses administrasi dan pengarsipan dokumen. Penelitian oleh Ramdhani (2025) menunjukkan bahwa model deep learning untuk NER berbahasa Indonesia belum sepenuhnya efektif dalam menangani kekhususan struktur dokumen administratif yang semi-terstruktur, sehingga diperlukan pendekatan hybrid yang menggabungkan deep learning dengan rule-based untuk mencapai akurasi yang lebih tinggi. Pendekatan hybrid yang dikembangkan oleh Ramdhani (2025) berhasil mencapai skor F1 rata-rata 98% pada sepuluh jenis dokumen kepegawaian.

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pengembangan sistem analisis dokumen administratif berbasis OCR dan Named Entity Recognition menggunakan model transformer IndoBERT. Sistem yang dikembangkan mampu menerima dokumen administratif dalam format PDF maupun dokumen hasil pemindaian, melakukan ekstraksi teks menggunakan OCR, mengidentifikasi teks variabel dinamis melalui model NER, serta menghasilkan ringkasan dokumen secara otomatis. Dengan adanya sistem ini, diharapkan proses identifikasi informasi penting pada dokumen administratif dapat dilakukan secara lebih cepat, akurat, dan efisien sehingga mampu membantu pekerjaan administrasi serta mendukung pengelolaan dokumen digital di lingkungan institusi.

## 1.2. Perumusan Masalah

1. Bagaimana merancang pipeline yang mampu mengekstrak teks dan informasi tata letak dari dokumen PDF administratif menggunakan PyMuPDF, serta mengidentifikasi bagian-bagian struktural dokumen (header, body, tabel, tanda tangan) berdasarkan fitur spasial dan tekstual?

2. Bagaimana membangun model Named Entity Recognition (NER) yang dapat mengenali entitas-entitas kunci dalam dokumen administratif, yaitu NOMOR\_SURAT, JENIS\_DOKUMEN, TANGGAL, PENGIRIM, PENERIMA, PERIHAL, dan LOKASI?

(Penelitian terkini menunjukkan bahwa model transformer seperti IndoBERT



#### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

dapat di-fine-tuning untuk mengenali entitas pada teks berbahasa Indonesia dengan performa yang baik [Muchtar dkk., 2025; Ramdhani, 2025].)

3. Bagaimana menghasilkan ringkasan otomatis dari isi dokumen (bagian body) yang informatif dan akurat?
4. Bagaimana mengevaluasi kinerja sistem dalam mengekstraksi entitas dan menghasilkan ringkasan dibandingkan dengan anotasi manual?

### 1.3. Batasan Masalah

- Jenis dan format dokumen: Dokumen administratif dalam format PDF yang berasal dari lingkungan JTik, meliputi Surat Tugas, Nota Dinas, Surat Undangan, dan Surat Keputusan. Diasumsikan dokumen lahir digital (bukan hasil pindaian) sehingga teks dapat diekstrak langsung menggunakan PyMuPDF.
- Bahasa dokumen: Bahasa Indonesia.
- Entitas yang diekstrak: Terbatas pada header Header (teks kop instansi), Body (teks isi utama), Table (area tabel), Signature (area tanda tangan),
- Pendekatan NER: Menggunakan pustaka IndoBERT dengan model bahasa Indonesia yang telah dilatih ulang (fine-tuning) atau menggunakan pendekatan berbasis aturan untuk entitas tertentu. Penelitian oleh Muchtar dkk. (2025) menunjukkan bahwa pendekatan fine-tuning IndoBERT dapat menghasilkan F1-score yang stabil untuk pengenalan entitas seperti PER, ORG, LOC, DATE, dan CRD pada teks berbahasa Indonesia.
- Metode peringkasan: Menggunakan metode ekstraktif sederhana (Template-based/Rulebased) karena keterbatasan data latih untuk model abstraktif.
- Output sistem: Berupa file JSON yang berisi entitas yang ditemukan beserta posisinya (bounding box) dan ringkasan teks.

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang menggunakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**1.4. Tujuan Penelitian**

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. Membangun pipeline otomatis untuk ekstraksi teks dan tata letak dari dokumen PDF administratif menggunakan PyMuPDF.
2. Mengembangkan modul identifikasi bagian struktural dokumen (header, body, tabel, tanda tangan) berdasarkan informasi spasial.
3. Melatih dan mengevaluasi model NER untuk mengenali entitas-entitas administratif pada teks dokumen.
4. Mengimplementasikan modul peringkasan teks untuk menghasilkan ringkasan isi dokumen.
5. Mengukur akurasi ekstraksi entitas dan kualitas ringkasan yang dihasilkan.

**1.5. Manfaat Penelitian**

Manfaat penelitian ini dibedakan menjadi dua aspek:

**1.5.1 Manfaat Teoritis**

1. Memberikan kontribusi pada pengembangan sistem pemahaman dokumen berbahasa Indonesia, khususnya untuk domain administratif.
2. Menjadi referensi bagi peneliti lain dalam mengintegrasikan OCR, NER, dan peringkasan teks pada dokumen semi-terstruktur.

**1.5.2 Manfaat Praktis**

1. Memudahkan staf administrasi JTIK dalam mencari dan memanfaatkan informasi dari arsip dokumen.
2. Mengurangi waktu dan potensi kesalahan dalam identifikasi informasi penting pada dokumen administratif.
3. Menjadi dasar pengembangan sistem manajemen dokumen yang lebih cerdas di lingkungan institusi.



**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## 1.6. Sistematika Penulisan

Laporan penelitian ini disusun dalam lima bab yang saling berkaitan untuk memberikan gambaran yang utuh dan runtut mengenai seluruh proses pengerjaan. Bagian awal adalah Bab I Pendahuluan, yang berfungsi sebagai pengantar untuk menjelaskan alasan mengapa penelitian ini dilakukan, masalah apa yang ingin diselesaikan, serta tujuan dan batasan yang ditetapkan.

Pembahasan kemudian berlanjut ke Bab II Tinjauan Pustaka, yang akan mengupas teori-teori pendukung dan menelaah penelitian terdahulu yang relevan sebagai landasan berpikir. Setelah bekal teori cukup, Bab III Metode Penelitian akan menjelaskan langkah-langkah teknis yang diambil, mulai dari perancangan sistem hingga strategi pengujian yang akan dilakukan.

Hasil nyata dari perancangan tersebut akan disajikan dalam Bab IV Hasil dan Pembahasan, yang menampilkan prototipe aplikasi yang telah jadi serta analisis kinerjanya. Akhirnya, laporan ditutup dengan Bab V Penutup, yang berisi kesimpulan menyeluruh dari penelitian serta saran-saran untuk pengembangan lebih lanjut di masa depan.

**POLITEKNIK  
NEGERI  
JAKARTA**

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

## **BAB V PENUTUP**

### **5.1 Kesimpulan**

Penelitian ini berhasil mengembangkan sistem analisis dokumen administratif berbasis OCR (PaddleOCR), NER (IndoBERT fine-tuning), dan ringkasan otomatis. Model NER mencapai performa terbaik dengan F1-score 95,15% (Run ID #13). Sistem mampu mengekstrak teks dari PDF digital, PDF scan, dan DOCX, mengidentifikasi entitas administratif (nomor surat, tanggal, pengirim, penerima, perihal, lokasi), serta menghasilkan ringkasan dokumen dalam format JSON. Hasil pengujian fungsional dan User Acceptance Testing (UAT) menunjukkan bahwa sistem berfungsi sesuai kebutuhan dan diterima baik oleh pengguna, meskipun masih terdapat keterbatasan pada dokumen dengan struktur kompleks (TOR, lampiran, transkrip). Seluruh tujuan penelitian tercapai.

### **5.2 Saran**

Untuk pengembangan lebih lanjut, disarankan agar jumlah dan variasi dataset pelatihan diperbanyak, terutama untuk entitas LOKASI dan PERIHAL yang masih memiliki tingkat akurasi lebih rendah, sehingga model dapat belajar lebih banyak pola penulisan dokumen administratif dari berbagai instansi. Penelitian selanjutnya juga dapat mengoptimalkan pra-pemrosesan citra pada dokumen hasil pemindaian berkualitas rendah, misalnya dengan teknik denoising dan deskewing, guna meningkatkan kualitas ekstraksi teks oleh PaddleOCR. Selain itu, metode ringkasan dapat ditingkatkan dari ekstraktif menjadi abstraktif berbasis transformer agar ringkasan yang dihasilkan lebih alami dan informatif. Terakhir, sistem ini berpotensi dikembangkan menjadi platform yang lebih komprehensif dengan menambahkan fitur klasifikasi dokumen, pencarian semantik, serta integrasi dengan sistem administrasi institusi, kemudian diuji coba pada lingkungan operasional nyata dengan jumlah pengguna yang lebih besar.

## DAFTAR PUSTAKA

Ramdhani, T.W. (2025) Pendekatan Hybrid Rule Based dan Deep Learning untuk Named Entity Recognition pada Dokumen Kepegawaian Pemerintah di Indonesia, Disertasi, Fakultas Ilmu Komputer Universitas Indonesia. Tersedia di: <https://www.lib.ui.ac.id>.

Tandiono, S.M. (2025) Evaluasi Optimalisasi Metode Ekstraksi Teks Berbasis Teknologi OCR NER dan LLM, MBKM Report, Universitas Multimedia Nusantara. Tersedia di: <https://kc.umn.ac.id/id/eprint/36386/> (Diakses: 9 Juni 2026).

Muchtar, N.U., Arifianto, D. dan Umlasari, R. (2025) 'Analisis Kinerja Transformer Untuk Named Entity Recognition (NER) Menggunakan IndoBERT Pada Transkrip Video Politik Berbahasa Indonesia', *Discovery : Jurnal Ilmu Pengetahuan*, 10(2), pp. 180-199. doi: 10.33752/jd.v10i2.10133.

Hasanah, U. dkk. (2025) PENERAPAN OPTICAL CHARACTER RECOGNITION (OCR) DAN NATURAL LANGUAGE PROCESSING (NLP) UNTUK PENGELOLAAN INFORMASI DOKUMEN (STUDI KASUS PADA KAPEL ST YOHANES RASUL PAROKI ST YUSUP KARANGPILANG), Buku karya ilmiah, Universitas Telkom. Tersedia di: <https://openlibrary.telkomuniversity.ac.id>.

Cui, C. dkk. (2025) PaddleOCR-VL: Boosting Multilingual Document Parsing via a 0.9B Ultra-Compact Vision-Language Model, *arXiv preprint*, Tersedia di: <https://arxiv.org/abs/2510.14528> (Diakses: 9 Juni 2026).

Hidayat, T., Nugroho, A. dan Wibowo, B. (2025) 'Extractive Indonesian Automated Text Summarization with IndoBERT and One-Dimensional Convolutional Neural Network', dalam *Lecture Notes in Computer Science*, Springer. doi: 10.1007/978-3-031-89045-5\_8.

Liu, Z., Shi, K. dan Chen, N.F. (2025) 'Transformer-based document-level discourse processing: Exploiting prior language knowledge and hierarchical parsing', *Computer Speech & Language*, 94, p. 101809.

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta

2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

A. Yani and A. S. Alimsyah, "Pengembangan Teknologi Internet dalam Administrasi Persuratan: Membangun Sistem Surat Otomatis," Jurnal Citra Pendidikan, vol. 5, no. 4, pp. 99-109, 2025.

Zain, M.A., Nugroho, A. dan Lestari, D. (2025) 'Data Augmentation using Large Language Models for Indonesian Named Entity Recognition', Journal of ICT Research and Applications, 19(1), pp. 45-58.

Al Faraby, S., & Romadhony, A. (2022). Pengaruh Distribusi Panjang Data Teks pada Klasifikasi: Sebuah Studi Awal. Jurnal Media Informatika Budidarma, 6(3), 1501. <https://doi.org/10.30865/mib.v6i3.4259>

Binmakhashen, G. M., & Mahmoud, S. A. (2020). Document layout analysis: a comprehensive survey. ACM Computing Surveys (CSUR), 52(6), 1-36.



**POLITEKNIK  
NEGERI  
JAKARTA**

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**DAFTAR RIWAYAT HIDUP**

Lahir di Manado, 13-12-2004

Penulis merupakan mahasiswa Program Studi Teknik Informatika, Jurusan Teknik Informatika dan Komputer, Politeknik Negeri Jakarta. Penulis menempuh pendidikan dasar di SDS ANGKASA 1 kemudian melanjutkan pendidikan di SMPN 50 JAKARTA dan SMAS ANGKASA 1. Pada tahun 2022, penulis melanjutkan pendidikan Diploma IV/Sarjana Terapan pada Program Studi Teknik Informatika, Politeknik Negeri Jakarta.

**POLITEKNIK  
NEGERI  
JAKARTA**

## LAMPIRAN

### Lampiran 1 Transkrip Wawancara

|                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|-------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Judul Penelitian  | PENGEMBANGAN SISTEM CERDAS UNTUK MANAJEMEN, ANALISIS, DAN OTOMATISASI DOKUMEN BERBASIS WEB                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Nama Peneliti     | Aurio Hendrianoko Rajaa<br>Muhammad Nathan                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
| Hari/Tanggal      | Selasa/Mar 03 2026                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
| Narasumber        | Mba Faras                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| Daftar Pertanyaan | <p>Bagaimana alur SOP pada penulisan surat?</p> <p>Apakah sudah menggunakan sistem otomatisasi seperti mail merge atau masih manual?</p> <p>Bagaimana cara penyimpanan dan pengelolaan dokumen saat ini?</p> <p>Apa saja jenis-jenis dokumen yang sering dibuat?</p> <p>Bagaimana perbedaan format antar jenis dokumen tersebut?</p> <p>Bagaimana proses pembuatan dokumen khusus seperti surat peringatan (SP)?</p> <p>Berapa jumlah dokumen yang diproses dalam satu periode (misalnya per tahun)?</p> <p>Bagaimana proses pembuatan dokumen magang?</p> <p>Bagaimana cara distribusi dokumen kepada pihak terkait?</p> <p>Apakah masih memerlukan hard copy atau sudah full digital?</p> <p>Seberapa sering template dokumen diperbarui?</p> <p>Apa kendala utama dalam proses pengelolaan dokumen saat ini?</p> |

#### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, /penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Hasil  
Transkrip

Peneliti:

Mba, saya mau tanya terkait alur kerja pembuatan dokumen di sini. Secara umum, workflow-nya itu seperti apa dari awal sampai dokumen jadi?

Narasumber:

Kalau secara umum sih, biasanya kita mulai dari kebutuhan dulu ya. Misalnya ada permintaan surat, nanti kita lihat jenisnya apa, terus kita pakai template yang sudah ada. Kadang kalau banyak datanya, kita pakai mail merge buat generate banyak dokumen sekaligus.

Peneliti:

Berarti sudah ada automation sedikit ya, seperti mail merge?

Narasumber:

Iya, tapi belum semuanya. Masih banyak juga yang manual, apalagi kalau datanya beda-beda atau inputnya tidak terstruktur.

Peneliti:

Kalau penyimpanan dokumennya sendiri gimana?

Narasumber:

Disimpan di folder, jadi ada central repository gitu. Tapi ya masih folder biasa, belum ada sistem khusus.

Peneliti:

Untuk kolaborasi atau kerja bareng tim, apakah sudah ada sistemnya?

Narasumber:

Belum terlalu, biasanya masih sharing file aja.

Peneliti:

Apa saja jenis-jenis dokumen yang sering dibuat di sini?

Narasumber:

Banyak sih. Yang sering itu seperti surat tugas, surat keterangan, surat keputusan (SK), sama notulensi.

Peneliti:

Kalau untuk yang eksternal?

Narasumber:

Ada juga undangan, surat pengadaan, perbaikan kelas, sama surat jawaban. Kadang juga surat pengantar, misalnya untuk MoU atau

**Hak Cipta :**

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

|                                                                                                                                                |  |
|------------------------------------------------------------------------------------------------------------------------------------------------|--|
| keperluan lain.                                                                                                                                |  |
| Peneliti:<br>Untuk formatnya apakah berbeda-beda?                                                                                              |  |
| Narasumber:<br>Iya, tapi sebagian besar sudah ada templatnya. Cuma tetap perlu disesuaikan lagi.                                               |  |
| Peneliti:<br>Kalau untuk dokumen seperti surat peringatan (SP), bagaimana formatnya?                                                           |  |
| Narasumber:<br>SP itu biasanya ada SP1, SP2, SP3, formatnya hampir sama, cuma beda di tingkatannya.                                            |  |
| Peneliti:<br>Dalam satu tahun kira-kira berapa banyak dokumen yang dibuat?                                                                     |  |
| Narasumber:<br>Kurang lebih sekitar 1500 dokumen per tahun.                                                                                    |  |
| Peneliti:<br>Untuk kasus tertentu seperti magang, itu prosesnya bagaimana?                                                                     |  |
| Narasumber:<br>Kalau magang, biasanya mahasiswa isi Google Form dulu. Setelah itu, datanya kita ambil, terus kita buat suratnya secara manual. |  |
| Peneliti:<br>Berarti belum otomatis ya dari form ke dokumen?                                                                                   |  |
| Narasumber:<br>Belum, masih input satu-satu.                                                                                                   |  |
| Peneliti:<br>Untuk distribusi dokumen, biasanya dikirim bagaimana?                                                                             |  |
| Narasumber:<br>Sekarang kebanyakan lewat email. Jadi sudah tidak perlu hard copy untuk beberapa jenis dokumen.                                 |  |
| Peneliti:<br>Kalau untuk template sendiri, apakah sering berubah?                                                                              |  |



## © Hak Cipta milik Jurusan TIK Politeknik Negeri Jakarta

### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengumukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Narasumber:

Iya, biasanya ada update. Misalnya sekarang pakai template 2025–2026, nanti bisa berubah lagi.



### Hak Cipta :

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
  - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penulisan laporan, penulisan kritik atau tinjauan suatu masalah.
  - b. Pengutipan tidak merugikan kepentingan yang wajar Politeknik Negeri Jakarta
2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin dari Jurusan TIK Politeknik Negeri Jakarta

Bukti Dokumentasi



### Lampiran 2 Dokumentasi Pelaksanaan *User Acceptance Testing*

