

1759740253957_devnolife.docx

by Premium Portal

Submission date: 06-Oct-2025 11:47AM (UTC+0300)

Submission ID: 2755469825

File name: 1759740253957_devnolife.docx (137.42K)

Word count: 4585

Character count: 30432

Penelitian ini berjudul “Implementasi Retrieval-Augmented Generation (RAG) menggunakan model GPT untuk chatbot untuk mencari informasi tentang peraturan Keselamatan dan Kesehatan di Tempat Kerja, yang diharapkan memberikan manfaat secara teoritis dan praktis, yaitu:

PROPOSAL SKRIPSI

1
Diusulkan sebagai salah satu syarat untuk mendapatkan gelar sarjana komputer (S.Kom) di Program Studi Informatika



FATURROHMAN AL-KHAIR

105841107021

3
PROGRAM STUDI INFORMATIKA

FAKULTAS TEKNIK

UNIVERSITAS MUHAMMADIYAH MAKASSAR

2025

13 KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Assalamu'alaikum Warahmatullahi Wabarakatuh

Segala puji dan syukur penulis panjatkan ke hadirat Allah Subhanahu Wa Ta'ala atas limpahan rahmat, karunia, dan petunjuk-Nya sehingga **Proposal Skripsi berjudul "Implementasi Retrieval-Augmented Generation (RAG) Menggunakan Model GPT untuk Chatbot Pencarian Informasi Regulasi Keselamatan dan Kesehatan Kerja (K3)"** dapat diselesaikan dengan baik.

Proposal untuk skripsi ini disusun sebagai salah satu persyaratan akademik yang harus dipenuhi untuk menyelesaikan program sarjana di Fakultas Teknik Program Studi Informatika Universitas Muhammadiyah Makassar. Penulis menyadari bahwa karya ini masih jauh dari sempurna, oleh karena itu masukan berupa kritik dan saran yang membangun dari berbagai pihak sangat diharapkan untuk memperbaiki proposal skripsi ini.

Dalam proses penyusunan proposal skripsi ini, penulis memperoleh berbagai bentuk bantuan, bimbingan, dan dukungan dari berbagai pihak, baik secara materiil maupun moril. Selama menempuh perkuliahan, penulis juga senantiasa mendapatkan dukungan yang luar biasa. Oleh sebab itu, penulis menyampaikan rasa hormat dan terima kasih yang setulus-tulusnya kepada:

1. Kedua orang tua kami tercinta, penulis mengucapkan terima kasih yang sebesar-besarnya atas segala limpahan kasih sayang, doa yang tak terputus, serta dukungan moral dan materiil yang sangat berarti bagi penulis.
2. Ketua Program Studi Informatika Fakultas Teknik Universitas Muhammmad Makassar, yang telah memberikan arahan dan fasilitas akademik kepada Bapak Muh.Hayat A.M. Hayat.
3. Bapak **Muhyiddin A.M Hayat S.Kom.,M.T.**, selaku Ketua Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar yang telah memberikan arahan dan fasilitas akademik.

4. Ibu Titin Wahyuni, S.Pd., M.T. sebagai Dosen Pembimbing I, yang dengan sabar memberikan bimbingan, saran dan koreksi yang konstruktif selama proses penyusunan proposal ini.³⁷
5. Bapak Fachrim Irhamna Rachman S.Kom.,M.T., selaku Dosen Pembimbing II yang telah meluangkan waktu, tenaga dan pikiran untuk memberikan bimbingan, arahan, dan masukan yang sangat berharga dalam penyusunan proposal skripsi ini.¹⁶
6. Seluruh Bapak/Ibu Dosen dan Tenaga Kependidikan Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar yang telah memberikan ilmu pengetahuan, pengalaman, dan pelayanan akademik.¹⁰
7. Saudara/saudari kami di Fakultas Teknik. Khususnya Angkatan 2021, dengan rasa persaudaraan yang tinggi banyak membantu, memberikan semangat, dukungan dan kerjasama yang baik dalam menjalani proses perkuliahan dan penyusunan proposal skripsi ini.¹

Akhir kata, penulis berharap proposal ini dapat memberikan kontribusi positif bagi pengembangan ilmu pengetahuan dan teknologi, khususnya di bidang Keinformatikaan. Aamiin

“Billahi Fii Sabilil Haq Fastabiqul Khaerat”

Wassalamualaikum Warahmatullahi Wabarakatuh

Makassar, Agustus 2025

Penulis

DAFTAR ISI

⁴⁴ KATA PENGANTAR	2
DAFTAR ISI.....	4
DAFTAR TABEL.....	5
DAFTAR GAMBAR	6
¹ BAB I PENDAHULUAN	7
A. Latar Belakang	7
B. Rumusan Masalah	9
C. Tujuan Penelitian.....	9
D. Manfaat Penelitian.....	9
E. Ruang Lingkup Peneltian	10
F. Sistematika Penulisan.....	11
BAB II TINJAUAN PUSTAKA.....	12
A. Landasan Teori.....	12
B. Penelitian Terkait	14
C. Kerangka Pikir.....	16
BAB III METODE PENELITIAN.....	18
A. Tempat dan Waktu Penelitian	18
B. Alat dan Bahan	18
C. Perancangan Sistem.....	19
D. Teknik Pengujian Sistem.....	25
E. Teknik Analisis Data	²⁰ 25
DAFTAR PUSTAKA	27

DAFTAR TABEL

Table 1. Penelitian Terkait	14
Table 2. Waktu Penelitian.....	18

⁴¹
DAFTAR GAMBAR

Gambar 1. Kerangka Berpikir	
Gambar 2. Alur Penelitian.....	21
Gambar 3. Use case diagram.....	⁶⁹ 21
Gambar 4. Activity Diagram	

BAB I

PENDAHULUAN

A. ¹⁴ Latar Belakang

Keselamatan dan Kesehatan Kerja (K3) merupakan aspek penting yang menjamin perlindungan tenaga kerja dari risiko kecelakaan dan penyakit akibat aktivitas kerja. Undang-undang No. 1 Tahun 1970, Keselamatan Kerja adalah keseluruhan upaya untuk menjaga keutuhan jasmani dan rohani pekerja di lingkungan kerja. Regulasi tentang Keselamatan dan Kesehatan Kerja telah banyak diterbitkan oleh pemerintah melalui Peraturan Menteri, Standar Operasional Prosedur (SOP), hingga dokumen teknis yang berlaku di berbagai sektor industri. Regulasi tersebut ³² tidak hanya bertujuan untuk menurunkan angka kecelakaan kerja, tetapi juga untuk menciptakan lingkungan kerja yang produktif, sehat, dan berkelanjutan.

Namun, meskipun dokumen regulasi K3 tersedia secara resmi, tingkat literasi dan pemahaman pekerja terhadap isi regulasi tersebut masih tergolong rendah. Studi menunjukkan bahwa sebagian besar pekerja tidak memahami secara menyeluruh isi prosedur keselamatan kerja, bahkan setelah mengikuti pelatihan (Hidayah & Kamali Zaman, 2022). Hambatan yang paling sering ditemukan meliputi keterbatasan akses terhadap informasi digital, penggunaan ⁶⁵ bahasa hukum yang sulit dipahami oleh orang awam, serta kurangnya media interaktif yang dapat menjembatani pemahaman regulasi dengan konteks kerja aktual (Ulzheimer et al., 2021). Di samping itu, pelatihan formal mengenai K3 memerlukan biaya yang cukup besar, membuatnya tidak dapat diakses secara merata oleh semua pihak.

⁵⁶ Seiring dengan berkembangnya teknologi kecerdasan buatan (*Artificial Intelligence/AI*), pendekatan baru dalam penyebaran informasi regulasi mulai mendapatkan perhatian, salah satunya melalui pengembangan chatbot berbasis bahasa alami (*Natural Language Processing*). Model *Generative Pre-trained Transformer* (GPT) merupakan salah satu teknologi NLP paling canggih saat ini, dengan kemampuan menjawab pertanyaan dan memberikan informasi

GPT berbasis teks secara instan dan mudah dipahami (Yenduri et al., 2023),²⁵ telah digunakan dalam berbagai sektor, termasuk pendidikan, kesehatan, dan layanan publik, tetapi secara murni hanya mengandalkan data pelatihan umum, sehingga tidak cukup kuat ketika diminta untuk menjawab pertanyaan yang spesifik terhadap dokumen atau konteks tertentu seperti Peraturan Keselamatan dan Kesehatan Kerja.

Untuk mengatasi keterbatasan tersebut, metode *Retrieval-Augmented Generation* (RAG) diperkenalkan sebagai solusi yang menggabungkan pencarian dokumen relevan (*retrieval*) dengan kemampuan generatif model GPT. Dalam konteks ini, chatbot tidak hanya mengandalkan pengetahuan umum, tetapi juga menarik bagian teks dari dokumen K3 aktual sebagai basis jawabannya, sehingga menghasilkan respons yang faktual dan kontekstual. Studi sebelumnya telah membuktikan efektivitas RAG dalam sistem tanya jawab berbasis dokumen seperti fatwa, peraturan sekolah, dan modul pembelajaran. Namun, sejauh penelusuran penulis, belum ditemukan penelitian yang secara khusus mengimplementasikan RAG untuk chatbot pencarian regulasi K3 di Indonesia—ini merupakan gap penting yang ingin dijawab dalam penelitian ini.⁶⁶

Penelitian ini bertujuan untuk mengembangkan sistem *chatbot* berbasis *Generative Pre-trained Transformer* (GPT) yang diintegrasikan dengan⁵⁷ pendekatan *Retrieval-Augmented Generation* (RAG) guna memfasilitasi pencarian dan pemahaman informasi regulasi Keselamatan dan Kesehatan Kerja (K3) secara otomatis. Sistem ini dirancang sebagai solusi alternatif yang efisien, interaktif, dan inklusif untuk mendukung peningkatan literasi K3 bagi pekerja, petugas K3, maupun pelaku industri. Melalui pemanfaatan teknologi kecerdasan buatan yang mampu menghadirkan informasi secara cepat, akurat, dan kontekstual,⁶² penelitian ini diharapkan dapat memberikan kontribusi signifikan terhadap peningkatan keselamatan kerja serta memperluas akses terhadap regulasi K3 secara merata dan efektif. Selain itu, hasil penelitian ini berpotensi menjadi model penerapan teknologi AI pada sektor regulasi dan kepatuhan (*compliance*), yang dapat direplikasi pada berbagai domain

pengetahuan lainnya, sehingga memberikan dampak positif bagi pengelolaan informasi di berbagai bidang industri.

29. B. Rumusan Masalah

Berdasarkan latar belakang diatas maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana mengintegrasikan berbagai dokumen regulasi K3 dalam satu sistem pencarian?
2. Apa solusi untuk membaca gambar atau tabel dalam pdf yang tidak bisa diekstraksi sebagai teks?

1. C. Tujuan Penelitian

Berdasarkan Rumusan Masalah, batasan pada penelitian ini adalah sebagai berikut:

1. Mengembangkan sistem pencarian yang mampu mengintegrasikan berbagai dokumen regulasi K3 (Keselamatan dan Kesehatan Kerja) dalam satu platform terpadu hingga memudahkan pengguna dalam menemukan informasi yang relevan.
2. Merumuskan solusi teknis untuk mengekstraksi informasi dari elemen non-teks seperti gambar atau tabel dalam dokumen PDF yang tidak dapat dibaca secara langsung sebagai teks.

D. Manfaat Penelitian

Penelitian ini berjudul “Implementasi Retrieval-Augmented Generation (RAG) menggunakan model GPT untuk chatbot untuk mencari informasi tentang peraturan Keselamatan dan Kesehatan di Tempat Kerja, yang diharapkan memberikan manfaat secara teoritis dan praktis, yaitu:

1. Secara Teoritis:

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem pencarian informasi berbasis kecerdasan buatan yang bertujuan mempermudah akses terhadap regulasi Keselamatan dan Kesehatan Kerja secara cepat dan relevan.

⁵⁰
2. Secara Praktis

a. Manfaat bagi peneliti

Penelitian ini memberikan peneliti pengalaman mendalam mengenai implementasi model AI, serta penerapannya dalam dunia nyata. Selain itu, penelitian ini juga melatih peneliti dalam proses perancangan sistem, pengolahan data, serta evaluasi sistem cerdas berbasis teks.

b. Manfaat bagi Sosial dan Nasional

Penelitian ini berpotensi mendukung pembangunan nasional melalui peningkatan pemahaman dan akses terhadap ⁵⁴ regulasi keselamatan kerja, yang pada akhirnya dapat mengurangi ⁵⁴ kecelakaan kerja dan meningkatkan kesejahteraan tenaga kerja di Indonesia.

E. Ruang Lingkup Peneltian

Berdasarkan rumusan masalah, adapun batasan pada penelitian ini sebagai berikut:

1. Ruang lingkup Regulasi Keselamatan dan Kesehatan Kerja: Regulasi yang digunakan hanya mencakup peraturan nasional ⁶⁰ di Indonesia, seperti undang-undang, peraturan pemerintah, Peraturan Menteri Ketenagakerjaan, Keputusan Menteri dan SNI yang berkaitan dengan ⁵¹ keselamatan dan kesehatan kerja (K3).
2. Bahasa Dokumen: Seluruh dokumen yang digunakan dalam proses retrieval dan pemrosesan berbahasa Indonesia; dokumen berbahasa asing tidak disertakan dalam penelitian ini.
3. Jenis dan Format Dokumen: Penelitian hanya memproses dokumen digital berformat PDF dan DOCX, dokumen hasil pemindaian (scan) tanpa teknologi OCR tidak digunakan.
4. Retrieval-Augmented Generation (RAG): Penelitian berfokus pada penerapan model RAG dengan GPT sebagai komponen generatif; pendekatan lain seperti rule-based, keyword matching, atau fine-tuning murni tidak dieksplorasi.

5. Sistem Chatbot: Chatbot dirancang sebagai prototipe berbasis teks; tidak mendukung antarmuka berbasis suara atau integrasi dengan platform komunikasi seperti WhatsApp atau Telegram.
6. Uji Coba Sistem: Pengujian dilakukan secara terbatas pada kelompok kecil, seperti mahasiswa, akademisi, dan praktisi K3; tidak diterapkan secara luas dalam skala industri atau organisasi.

F. Sistematika Penulisan

Secara keseluruhan, struktur penulisan ini dibagi ke beberapa bab utama yang masing-masing memuat aspek-aspek penting sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan uraian tentang latar belakang masalah, Rumusan Masalah, tujuan penelitian, Manfaat Penelitian, Ruang Lingkup dan Sistematika Penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisi landasan teoritis yang relevan dengan topik penelitian. Teori-teori dan hasil studi terdahulu yang mendukung pelaksanaan penelitian.

BAB III METODE PENELITIAN

Bab ini membahas secara rinci metode yang digunakan dalam penelitian, termasuk pendekatan, tahap pelaksanaan, serta perangkat dan sumber daya yang dimanfaatkan dalam proses desain dan pengembangan sistem.

TINJAUAN PUSTAKA**A. Landasan Teori****1. Keselamatan dan Kesehatan Kerja (K3)**

Keselamatan dan Kesehatan Kerja (K3) adalah serangkaian upaya sistematis yang dilakukan untuk memastikan kondisi kerja yang aman, sehat, dan terhindar dari risiko kecelakaan dan penyakit akibat kerja. Undang-Undang No. 1 Tahun 1970 mendefinisikan Keselamatan Kerja sebagai segala tindakan untuk menjamin keutuhan dan kesempurnaan jasmani dan rohani tenaga kerja melalui perlindungan kerja di tempat kerja. (Waisapi, 2022) menjelaskan bahwa tujuan dari Keselamatan dan Kesehatan Kerja adalah untuk mencegah kecelakaan kerja, meningkatkan produktivitas, dan menciptakan suasana kerja yang aman dan nyaman.

Menurut (Djoko Subagijo) regulasi K3 mencakup berbagai ketentuan teknis dan hukum yang kompleks, sehingga pemahaman terhadapnya membutuhkan sistem pendukung berbasis teknologi yang mampu menjembatani antara dokumen regulatif dengan kebutuhan pengguna. Maka dari itu, pendekatan digital yang interaktif seperti chatbot menjadi penting untuk mempermudah akses informasi ini.

2. ¹² Sistem Informasi

Sistem informasi adalah kumpulan elemen yang saling terkait untuk mengumpulkan, mengelola, menyimpan, dan mendistribusikan informasi untuk mendukung pengambilan keputusan dan pengendalian dalam organisasi, yang terdiri dari input, proses, output dan umpan balik yang semuanya dikoordinasikan oleh perangkat keras, perangkat lunak, manusia dan data (Vargas Saputra et al., 2025).

Dalam konteks penelitian ini, sistem informasi yang dibangun berfungsi untuk memfasilitasi interaksi antara pengguna dan data regulasi K3 melalui antarmuka chatbot yang cerdas dan adaptif.

3. Kecerdasan Buatan (Artificial Intelligence)

Kecerdasan buatan (AI) adalah bidang ilmu komputer yang berfokus pada pembuatan mesin yang dapat ⁶⁸ meniru kecerdasan manusia. AI memungkinkan komputer untuk melakukan tugas-tugas rumit seperti mengenali suara, memahami bahasa alami, membuat prediksi, dan belajar dari data.

AI telah menjadi fondasi untuk pengembangan sistem chatbot yang digunakan dalam penelitian ini, salah satu bentuk AI yang diterapkan adalah pemrosesan bahasa alami (NLP), yang memungkinkan sistem untuk memahami dan menjawab pertanyaan dalam bahasa manusia.

4. Chatbot

Chatbot adalah sistem komputer yang dirancang untuk berinteraksi dengan manusia melalui antarmuka percakapan berbasis teks atau suara.

Chatbot dapat bersifat rule-based (berorientasi aturan) atau AI-Based. chatbot AI mampu mempelajari konteks dan memberikan tanggapan yang lebih fleksibel dan personal dalam penelitian ini (Sharifi et al., 2021). Chatbot dikembangkan untuk memfasilitasi pencarian dan pemahaman regulasi K3, menjadikannya sebagai alat bantu edukatif dan informatif ²³ yang dapat diakses kapan saja dan oleh siapa saja.

saja.

5. Generative Pre-trained Transformer (GPT)

GPT adalah model bahasa berbasis deep learning yang dilatih pada ³⁶ sejumlah besar data teks untuk memahami dan menghasilkan bahasa alami, bekerja berdasarkan arsitektur transformasi yang memungkinkan pemrosesan urutan teks dalam konteks yang lebih luas (Brown et al., ⁶⁷ 2020).

Menurut (Brown et al., 2020) model GPT memiliki kemampuan zero-shot, one-shot, dan few-shot learning, sehingga sangat efektif dalam menjawab pertanyaan berbasis teks tanpa pelatihan ulang. GPT digunakan

dalam penelitian ini sebagai model inti dalam memberikan jawaban dari chatbot secara natural dan sesuai konteks.

6. **Retrieval-Augmented Generation (RAG)**

RAG adalah pendekatan dalam pengembangan sistem tanya-jawab (question answering system) yang menggabungkan proses pencarian informasi (retrieval) dari dokumen eksternal dengan kemampuan generatif model bahasa seperti GPT. Dalam arsitektur RAG, dua komponen utama digunakan: retriever untuk mencari dokumen relevan, dan generator untuk menyusun jawaban berdasarkan dokumen tersebut (Lewis et al., 2021a)

7. **Unified Modeling Language (UML)**

UML adalah bahasa pemodelan standar yang digunakan untuk merancang dan mendokumentasikan sistem perangkat lunak, menyediakan berbagai jenis diagram seperti use case, class, sequence, dan activity diagram yang membantu dalam merancang sistem secara visual dan struktural (Booch et al., 2005(M Teguh, 2018)).

Dalam penelitian ini, UML digunakan untuk memodelkan alur kerja sistem chatbot, interaksi antar komponen, serta hubungan antara pengguna, sistem, dan basis data regulasi K3.

B. Penelitian Terkait

Peneliti memperoleh banyak inspirasi dan referensi untuk menyusun proposal skripsi ini dari penelitian sebelumnya, terkait dengan latar belakang masalah pada proposal ini. Penelitian sebelumnya yang terkait antara lain:

Table 1. Penelitian Terkait

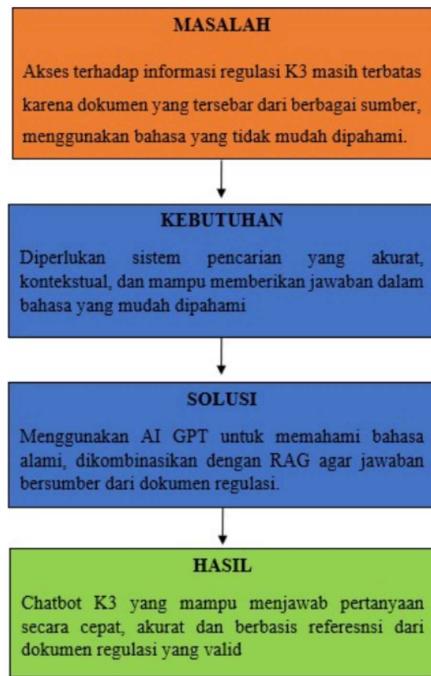
No	Nama/Tahun	Judul	Metode	Hasil
1	(Ahadi et al., 2025)	Retrieval-Augmented Generation in a Web-Based Question Answering	RAG + GPT-3, web-based QA	Mengurangi halusinasi LLM, meningkatkan akurasi jawaban berbasis

		System for Fiqh Books	dokumen referensi.
2	(Gayathri Devi. et al., 2024)	MineBot: Chatbot to Respond to Text Queries Pertaining to Acts, Rules, and Regulations in Mining	LangChain + GPT-3.5-turbo + RAG, integrasi Wikipedia & Google API
3	(Lewis et al., 2021b)	Retrieval-Augmented Generation for Knowledge-Intensive NLP Task	RAG yang menggabungkan <i>Dense Passage</i> <i>Retriever</i> dengan model generatif BART
4	(Silvanie & Subekti, 2022)	Aplikasi Chatbot Untuk Faq Akademik Di Ibi-K57 Dengan Lstm Dan Penyematan Kata	Long Short- Term Memory (LSTM) + Word Embedding (CBOW) + Flask (server) + Kotlin (client)

5	(Ranjit et al., 2023)	Retrieval- Augmented Chest X-ray Report Generation using a Pretrained Contrastive Language- Image Model Generative Language Model	Retrieval- Augmented Generation (RAG) + CLIP (Contrastive Language- Image Pretraining) untuk retrieval + GPT-2 dibanding generatif murni	dengan koneksi server meningkatkan kualitas laporan X-ray dengan memanfaatkan retrieval data historis yang relevan, menghasilkan deskripsi lebih akurat dan kaya informasi dibanding model generatif murni
---	--------------------------	---	---	--

1 C. Kerangka Pikir

Kerangka penelitian menunjukkan bagaimana membangun sistem chatbot yang dapat membantu pekerja, petugas K3, dan pelaku industri dalam mengakses dan memahami aturan tentang Keselamatan dan Kesehatan Kerja (K3) dengan lebih mudah. Alasan yang mendasari pengembangan sistem ini adalah bahwa masyarakat pekerja masih rendah dalam memahami regulasi K3, karena keterbatasan sumber informasi yang relevan, bahasa regulasi yang sulit dipahami, serta kurangnya media interaktif yang dapat menyajikan informasi tersebut secara cepat, tepat, dan kontekstual.



1 Gambar 1. Kerangka Berpikir

BAB III

METODE PENELITIAN

A. Tempat dan Waktu Penelitian

1. Tempat Penelitian

Lokasi penelitian yang dipilih oleh peneliti adalah di Kantor Keselamatan dan Kesehatan Kerja (K3) di Kampus Universitas Muhammmad Makassar.

2. Waktu Penelitian

Pelaksanaan penelitian ini berlangsung antara Juni dan Agustus 2025, dimulai dari tahap pengumpulan data hingga penyusunan laporan akhir.

Table 2. Waktu Penelitian

NO	KEGIATAN	BULAN							
		JUNI		JULI		AGUSTUS			
		3	4	1	2	3	4	1	2
1	Perumusan topik penelitian								
2	Perencanaan penelitian								
3	Pelaksanaan penelitian								
	a. Pembuatan program								
	b. Penelitian uji program								
4	Penulisan laporan								

1. Alat dan Bahan

Alat penelitian berupa laptop yang akan digunakan peneliti untuk merancang sistem chatbot untuk pencarian informasi regulasi Keselamatan dan Kesehatan Kerja (K3) berbasis RAG. Adapun perangkat keras dan perangkat lunak yang digunakan adalah sebagai berikut:

1. Perangkat keras (Pengembang)
 - a. Processor Intel Core-I5 (2 core, 8 thread)

- b. Besar Memory Ram 24GB
 - c. Kapasitas SSD 119GB
2. Perangkat Lunak
- a. Windows Home
 - b. Text editor Visual Studio Code
 - c. Python sebagai bahasa programming
 - d. FastAPI dan Uvicorn sebagai framerowk dan server backend
 - e. HTML, CSS dan Javascript sebagai User Interface (frontend)
 - f. MongoDB dan ChromaDB sebagai sistem manajemen basis data (Vector dan dokumentasi)
 - g. Google Chrome sebagai browser.

²⁰ Data yang digunakan dalam penelitian ini adalah dokumen regulasi K3, yang diperoleh dari berbagai sumber resmi, dan dimasukkan ke dalam sistem dan diproses untuk menjadi sumber pengetahuan yang akan digunakan untuk menjawab pertanyaan pengguna secara kontekstual melalui chatbot.

C. Perancangan Sistem

Perancangan sistem sangat penting dalam pengembangan suatu sistem, karena menguraikan bagaimana sistem dibangun dari tahap perencanaan hingga tahap pembuatan fungsi-fungsi yang diperlukan untuk pengoperasian sistem.

Alur jalannya penelitian pembuatan *chatbot* informasi regulasi Keselamatan dan Kesehatan Kerja (K3) seperti pada gambar 2 terbagi menjadi 7 (tujuh) blok penelitian, antara lain:

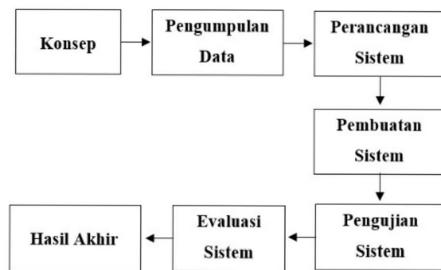
1. Konsep, tahapan ini dimulai dengan merumuskan ide dasar dari sistem chatbot yang akan dikembangkan. Pada tahap ini ditentukan permasalahan utama, batasan sistem, dan dirancang kebutuhan fungsional. Konsep chatbot diarahkan pada integrasi dan teknologi Artificial Intelligence (GPT) dan pencarian semantik (RAG) untuk

meningkatkan aksesibilitas regulasi Keselamatan dan Kesehatan Kerja (K3) dari dokumen.

2. Pengumpulan data, setelah konsep sistem telah ditetapkan, tahapan selanjutnya adalah pengumpulan data berupa dokumen regulasi K3 dalam format PDF. Selain itu, dilakukan studi literatur terhadap teknologi pendukung seperti GPT, LangChain, OpenAI Embedding, dan ChromaDB yang akan digunakan dalam proses pengolahan data.
3. Perancangan Sistem, Pada blok ini dilakukan perancangan teknis sistem, yang meliputi alur kerja sistem (workflow), perancangan modul utama, desain basis data, dan pemilihan algoritma yang akan digunakan. Komponen yang dirancang meliputi proses ekstraksi teks dari PDF, pembagian teks menjadi chunk, pembuatan embedding vektor menggunakan API OpenAI, penyimpanan embedding di MongoDB, hingga pengembangan sistem retrieval dan generation menggunakan LangChain dan GPT.
4. Pembuatan sistem, Tahap ini merupakan proses implementasi dari rancangan sistem sebelumnya. Pembuatan sistem dilakukan menggunakan bahasa pemrograman Python, dengan integrasi beberapa library seperti PyMuPDF, LangChain, ChromaDB, OpenAI, dan MongoDB. Pada tahap ini dibangun pipeline RAG yang mencakup ekstraksi teks, embedding, penyimpanan data, pencarian vektor, dan generasi jawaban oleh model GPT berdasarkan dokumen sumber.
5. Pengujian sistem, Setelah sistem selesai diimplementasikan, dilakukan pengujian terhadap berbagai jenis pertanyaan, baik pertanyaan langsung maupun pertanyaan lanjutan (multi-turn conversation). Pengujian mencakup keakuratan jawaban, relevansi konteks, serta responsivitas chatbot dalam menjawab berdasarkan isi dokumen regulasi K3. Hasil pengujian dianalisis secara deskriptif untuk menilai performa sistem.
6. Evaluasi dan penyimpulan, Setelah sistem selesai diimplementasikan, dilakukan pengujian terhadap berbagai jenis pertanyaan, baik pertanyaan langsung maupun pertanyaan lanjutan (multi-turn

conversation). Pengujian mencakup keakuratan jawaban, relevansi konteks, serta responsivitas chatbot dalam menjawab berdasarkan isi dokumen regulasi K3.

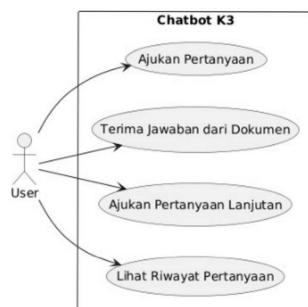
7. Hasil akhir, setelah melalui proses pengujian dan perbaikan, chatbot sudah dapat diterapkan. Jalannya penelitian dapat dilihat pada gambar 2 di bawah ini:



Gambar 2. Alur Penelitian

Langkah-langkah **Perancangan** Web Chatbot tentang Keselamatan dan Kesehatan Kerja digambarkan sebagai berikut:

1. *Use case diagram*



Gambar 3. Use case diagram

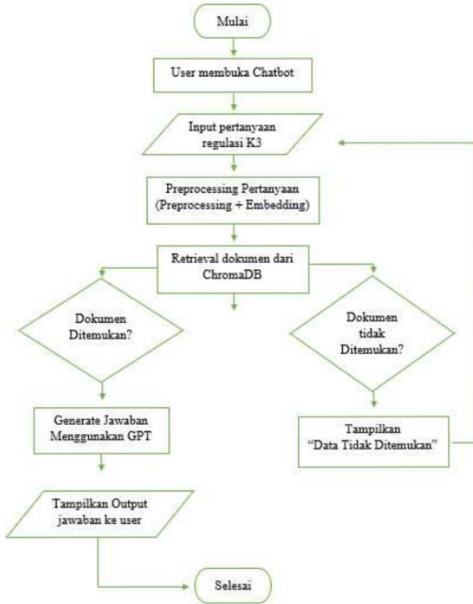
Pada tahap perancangan use case diagram, diagram ini berfungsi untuk menggambarkan interaksi antara pengguna (user) dengan sistem

chatbot regulasi Keselamatan dan Kesehatan Kerja (K3). Diagram ini menunjukkan proses awal hingga akhir penggunaan sistem, serta fitur-fitur utama yang tersedia bagi pengguna. Dalam sistem ini hanya terdapat satu jenis akun yaitu akun pengguna (user), karena sistem bersifat publik dan ditujukan untuk akses informasi K3 secara mudah.

- a. Ajukan pertanyaan, pada tahap awal, pengguna dapat langsung mengakses halaman antarnuha chatbot dan mengajukan pertanyaan seputar regulasi K3. Pertanyaan ini dapat berupa definisi, prosedur, atau ketentuan hukum terkait keselamatan kerja, seperti “Apa itu APD?”, atau “Apa isi pasal tentang pemadam kebakaran?”. Sistem kemudian menerima input ini dan memprosesnya melalui tahap *retrieval* dan *generation* menggunakan model RAG.
- b. Terima jawaban dari Dokumen, setelah sistem menerima pertanyaan, modul retrieval akan mencari bagian dokumen PDF yang relevan menggunakan pencarian berbasis vektor. Kemudian hasil pencarian tersebut akan dijadikan konteks oleh model GPT untuk menghasilkan jawaban yang informatif dan mudah dipahami. Jawaban ini ditampilkan secara instan kepada pengguna di tampilan chatbot.
- c. Ajukan pertanyaan lanjutan, Jika pengguna merasa jawabannya belum cukup lengkap, mereka dapat mengajukan pertanyaan lanjutan tanpa perlu mengulang konteks. Fitur ini memungkinkan percakapan berkelanjutan (*multi-turn conversation*), seperti “lanjutkan” atau “bagaimana dengan pengecualian pasal tersebut?”, dan sistem tetap mempertahankan konteks pertanyaan sebelumnya.
- d. Lihat riwayat pertanyaan, Pengguna juga dapat melihat kembali riwayat percakapan mereka dalam sesi sebelumnya. Fitur ini berguna untuk merevisi, menyimpan, atau menyalin informasi yang sudah diterima tanpa harus bertanya ulang. Riwayat disimpan secara lokal selama sesi berlangsung, atau secara opsional dapat diarsipkan.

1 2. Activity diagram

Activity Diagram adalah representasi grafis dari seluruh tahapan alur kerja yang berisi aktivitas, pilihan tindakan, pengulangan, dan hasil dari aktivitas tersebut, yang dapat digunakan untuk menjelaskan proses operasional secara langkah demi langkah dari komponen sistem tersebut.



Gambar 4. *Activity Diagram*

a. User membuka chatbot

Tahapan pertama adalah ketika pengguna (user) mengakses antarmuka chatbot yang sudah dibangun berbasis web. Pada tahap ini, pengguna siap untuk melakukan interaksi awal dengan sistem.

- b. User menginput pertanyaan

Pengguna kemudian mengetikkan pertanyaan seputar regulasi K3 yang ingin diketahui, misalnya "Apa isi pasal tentang APD?" atau "Bagaimana prosedur evakuasi darurat?".
- c. Preprocessing pertanyaan (embedding)

Setelah Pertanyaan dari pengguna kemudian diproses melalui tahap preprocessing. Pada tahap ini, pertanyaan diubah menjadi representasi vektor (embedding) agar bisa dibandingkan secara semantik dengan data yang ada di basis data ChromaDB.
- d. Retrieval dokumen dari ChromaDB

Setelah embedding pertanyaan dihasilkan, sistem melakukan pencarian ke dalam basis data vektor (ChromaDB) untuk menemukan dokumen yang memiliki kemiripan semantik tinggi. Proses ini dikenal dengan retrieval..
- e. Percabangan: Apakah dokumen ditemukan?

Sistem akan memeriksa apakah ada dokumen atau bagian dokumen yang relevan dengan pertanyaan pengguna, jika ditemukan, proses akan dilanjutkan ke tahap generasi jawaban jika tidak ditemukan. Jika tidak, sistem akan menampilkan informasi bahwa "Data Tidak Ditemukan."
- f. Generative jawaban menggunakan GPT

Apabila dokumen berhasil ditemukan, potongan teks yang relevan akan dikirimkan ke model GPT untuk menghasilkan jawaban yang utuh dan mudah dipahami. Model GPT akan melakukan proses reasoning dengan konteks yang diberikan.
- g. Tampilkan output jawaban ke user

Jawaban yang telah dihasilkan oleh GPT akan ditampilkan langsung di antarmuka pengguna. Sistem juga mempertahankan konteks untuk memungkinkan pertanyaan lanjutan (multi-turn conversation).Selesai

Setelah jawaban diterima oleh pengguna, proses sesi tanya-jawab dapat dihentikan, atau pengguna dapat memilih untuk melanjutkan percakapan dengan pertanyaan berikutnya.

D. Teknik Pengujian Sistem

Teknik Pengujian Sistem dilakukan dengan metode black box testing. Uji coba yang dilakukan bersifat mandiri dan diuji coba langsung dengan memperhatikan rancangan yang dibuat.

Pengujian blackbox (blackbox testing) adalah salah satu metode pengujian perangkat lunak yang berfokus pada sisi fungsionalitas, khususnya pada input dan output (apakah sudah sesuai dengan apa yang diharapkan atau belum).

Black Box Testing adalah pengujian yang dilakukan untuk mengamati hasil input dan output perangkat lunak tanpa mengetahui struktur kode perangkat keras. Pengujian ini dilakukan pada akhir pemrograman untuk mengetahui apakah perangkat Lunak dapat berfungsi dengan baik.

E. Teknik Analisis Data

Setelah semua data yang diperlukan dalam pengembangan sistem chatbot regulasi Keselamatan dan Kesehatan Kerja (K3) berhasil dikumpulkan, proses analisis data menjadi salah satu langkah penting dalam menjamin kualitas dan akurasi sistem yang dikembangkan. Analisis data berperan dalam mengevaluasi kinerja sistem serta memvalidasi apakah sistem mampu memenuhi kebutuhan pengguna secara optimal. Ketepatan dalam pemilihan metode analisis sangat menentukan kualitas kesimpulan yang dihasilkan, sehingga dapat berdampak besar pada kebermanfaatan dan penerapan hasil penelitian ini dalam dunia kerja nyata, khususnya pada bidang keselamatan kerja.

Dalam penelitian ini, pendekatan analisis data yang digunakan adalah analisis kualitatif deskriptif, hal ini disebabkan oleh sifat informasi yang dianalisis berasal dari pengujian sistem, berupa interaksi teks (pertanyaan dan jawaban) dan observasi terhadap kinerja chatbot dalam menjawab pertanyaan berbasis dokumen regulasi.

membandingkan antara hasil jawaban chatbot dengan isi dokumen K3 asli untuk menilai akurasi, kelengkapan jawaban, dan koherensi konteks, terutama dalam pertanyaan bertingkat (multi-turn).

Proses pengujian dilakukan melalui serangkaian skenario pertanyaan yang meliputi: pertanyaan definisi, prosedural, hingga pertanyaan lanjutan kontekstual seperti “lanjutkan” atau “jelaskan lebih dalam.” Log dari sesi tanya-jawab tersebut disimpan dalam bentuk file .json sebagai bagian dari dokumentasi, yang kemudian dianalisis secara manual untuk menilai apakah sistem mampu memberikan jawaban relevan berdasarkan isi dokumen sumber dan mengikuti konteks percakapan sebelumnya.

Selain itu, dilakukan juga evaluasi fungsional sistem dengan cara menguji keandalan proses ekstraksi teks, pembuatan embedding, penyimpanan data ke database, serta kemampuan sistem dalam menangani error seperti file PDF kosong atau pertanyaan tanpa hasil pencarian. Hasil evaluasi ini digunakan untuk mengidentifikasi kelemahan sistem dan memberikan dasar dalam menyusun rekomendasi pengembangan selanjutnya.

Dengan demikian, teknik analisis data pada penelitian ini difokuskan untuk mengevaluasi sistem secara menyeluruh dari segi fungsi, konteks pemahaman, serta keakuratan isi jawaban, sehingga dapat dipastikan bahwa sistem yang dikembangkan mampu menjadi solusi efektif dalam membantu pekerja atau petugas K3 dalam mengakses regulasi yang relevan secara efisien.

DAFTAR PUSTAKA

- Ahadi, R., Safaat Harahap, N., Fikry, M., & Kurnia, F. (2025). *Retrieval-Augmented Generation in a Web-Based Question Answering System for Fiqh Books*. *Journal of Artificial Intelligence and Software Engineering*, 5(2), 626–635. <https://doi.org/10.30811/jaise.v5i2.7005>
- Beheshti, A., Yang, J., Sheng, Q. Z., Benatallah, B., Casati, F., Dustdar, S., Nezhad, H. R. M., Zhang, X., & Xue, S. (2023). *ProcessGPT: Transforming Business Process Management with Generative Artificial Intelligence*. <http://arxiv.org/abs/2306.01771>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). *Language Models are Few-Shot Learners*. <http://arxiv.org/abs/2005.14165>
- Hidayah, P., & Kamali Zaman, M. (2022). Science Midwifery Implementation of Occupational Safety and Health (K3) Inspection as a Work Accident Prevention Effort in Palm Oil Factory, Kampar Regency, Riau Province A R TICLE IN F O ABSTRACT. In *Science Midwifery* (Vol. 10, Issue 3). Online. www.midwifery.iocspublisher.org
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021a). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. <http://arxiv.org/abs/2005.11401>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021b). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks*. <http://arxiv.org/abs/2005.11401>

- Gayathri Devi., K, Siranjeevi., S, Rupeshwar., & T, Oviya. (2024). Minebot:
²¹ Chatbot to Respond to Text Queries Pertaining to Various Acts, Rules, And
Regulations Applicable to Mining Industries. *International Journal of
Research Publication and Reviews*, 5(3), 4317–4325.
<https://doi.org/10.55248/gengpi.5.0324.07107>
- M²⁷ Teguh, P. (2018). *Unified Modeling Language (UML) Model Untuk
Pengembangan Sistem Informasi Akademik Berbasis Web.*
- Ranjit, M., Ganapathy, G., Manuel, R., & Ganu, T. (2023). *Retrieval Augmented
Chest X-Ray Report Generation using OpenAI GPT models.*
<http://arxiv.org/abs/2305.03660>
- Sharifi, A., Zibaei, A., & Rezaei, M. (2021). A deep learning based hazardous
materials (HAZMAT) sign detection robot with restricted computational
resources. *Machine Learning with Applications*, 6, 100104.
<https://doi.org/10.1016/j.mlwa.2021.100104>
- Silvanie, A., & Subekti, R. (2022). *APLIKASI CHATBOT UNTUK FAQ
AKADEMIK DI IBI-K57 DENGAN LSTM DAN PENYEMATAN KATA.*
Jurnal Informatika Dan Komputer) Akreditasi KEMENRISTEKDIKTI, 5(1).
<https://doi.org/10.33387/jiko>
- Strohmeier, Stefan. (2022). *Handbook of research on artificial intelligence in
human resource management*. Edward Elgar Publishing.
- Ulzheimer, L., Kanzinger, A., Ziegler, A., Martin, B., Zender, J., Römhild, A., &
Leyhe, C. (2021). Barriers in times of digital teaching and learning – a
german case study: Challenges and recommendations for action. *Journal of
Interactive Media in Education*, 2021(1). <https://doi.org/10.5334/jime.638>
- Vargas Saputra, T., Harlina, T., Studi, P., Informatika, M., Pgri Banyuwangi, S.,
& Korespondensi, P. (2025). *Penerapan Sistem Informasi Perpustakaan
Berbasis Web Pada STIKOM PGRI Banyuwangi*. 6(1).
<https://doi.org/10.31284/j.kernel.2025.v6i1.7804>

2

Yenduri, G., M. R., G. C. S., Y. S., Srivastava, G., Maddikunta, P. K. R., G. D. R., Jhaveri, R. H., B. P., Wang, W., Vasilakos, A. V., & Gadekallu, T. R. (2023). *Generative Pre-trained Transformer: A Comprehensive Review on Enabling Technologies, Potential Applications, Emerging Challenges, and Future Directions*. <http://arxiv.org/abs/2305.10435>



PRIMARY SOURCES

1	dilibadmind.unismuh.ac.id Internet Source	10%
2	Submitted to Liverpool John Moores University Student Paper	3%
3	pkm.tunasbangsa.ac.id Internet Source	1 %
4	www.ajhssr.com Internet Source	1 %
5	Submitted to Universitas Muhammadiyah Makassar Student Paper	1 %
6	ijlter.org Internet Source	1 %
7	Submitted to Consejo Colombiano de Seguridad Student Paper	1 %
8	Submitted to Universitas Musamus Merauke Student Paper	1 %
9	Submitted to Institut Teknologi Sumatera Student Paper	1 %
10	123dok.com Internet Source	1 %
11	oulurepo.oulu.fi Internet Source	1 %

12	Submitted to Universitas Budi Luhur Student Paper	1 %
13	Submitted to LPPM Student Paper	1 %
14	Submitted to Universitas Putera Batam Student Paper	1 %
15	I Komang Sugiarktha, Dw Ayu Agung Indra Swari, Dedy Panji Agustino, Tria Hikmah Fratiwi. "Pengembangan Aplikasi Mobile E-Pasraman Hindu untuk Optimalisasi Pengelolaan Data Akademik", remik, 2025 Publication	1 %
16	Submitted to Keimyung University Student Paper	1 %
17	eprints.itn.ac.id Internet Source	<1 %
18	www.frontiersin.org Internet Source	<1 %
19	www.coursehero.com Internet Source	<1 %
20	Submitted to iGroup Student Paper	<1 %
21	ijjrt.org Internet Source	<1 %
22	Submitted to Universitas Maritim Raja Ali Haji Student Paper	<1 %
23	repository.uinjkt.ac.id Internet Source	<1 %
24	dergipark.org.tr Internet Source	<1 %

25	penerbit.gunadarma.ac.id Internet Source	<1 %
26	repository.iainpalopo.ac.id Internet Source	<1 %
27	ejournal.kresnamediapublisher.com Internet Source	<1 %
28	ejournal.unkhair.ac.id Internet Source	<1 %
29	luckey13ch.blogspot.com Internet Source	<1 %
30	repository.unhas.ac.id Internet Source	<1 %
31	Submitted to Erasmus University Rotterdam Student Paper	<1 %
32	Submitted to Universitas Negeri Semarang Student Paper	<1 %
33	repository.ukwms.ac.id Internet Source	<1 %
34	www.ijrpr.com Internet Source	<1 %
35	Submitted to STT PLN Student Paper	<1 %
36	Submitted to Universitas Pendidikan Ganesha Student Paper	<1 %
37	core.ac.uk Internet Source	<1 %
38	ejournal.unib.ac.id Internet Source	<1 %
39	www.rumahjurnal.or.id Internet Source	<1 %

<1 %

<1 %

- 40 LIAN GANG SU, Hariyadi Sukamdani, Maya Dewi Diyah Maharani. "Management Safety Healthy Work dalam Upaya Meminimalkan Kecelakaan Kerja di Pt Huawei Tech Investment", Ekonika : Jurnal Ekonomi Universitas Kadiri, 2024

Publication

<1 %

- 41 docplayer.info <1 %
Internet Source

<1 %

- 42 sepatusafetyonline.com <1 %
Internet Source

<1 %

- 43 Agrawal, Tejas Prakash. "RadGraph: Integrating Fine-Grained and Global Features Using GCNs and LLMs", State University of New York at Buffalo, 2024

Publication

<1 %

- 44 Submitted to St. Joseph's College <1 %
Student Paper

<1 %

- 45 id.123dok.com <1 %
Internet Source

<1 %

- 46 jwk.bandung.lan.go.id <1 %
Internet Source

<1 %

- 47 repository.upnvj.ac.id <1 %
Internet Source

<1 %

- 48 Eric Alfonsius, Witha Radifa Mokodongan, Deiby Salaki. "Sistem Informasi Laporan (SILAPOR) Warga Pada Desa Tanamon Kecamatan Sinonsayang Berbasis Website Menggunakan Metode RAD", Inventor: Jurnal

<1 %

Inovasi dan Tren Pendidikan Teknologi Informasi, 2025

Publication

49	khp.re.kr Internet Source	<1 %
50	anzdoc.com Internet Source	<1 %
51	es.scribd.com Internet Source	<1 %
52	journal.ippmunindra.ac.id Internet Source	<1 %
53	jurnal.dharmawangsa.ac.id Internet Source	<1 %
54	kappppppp.blogspot.com Internet Source	<1 %
55	pure.psu.edu Internet Source	<1 %
56	www.pinterpolitik.com Internet Source	<1 %
57	Muhammad Yusuf, Asyrofi Anam, Muhammad Rifki Raihan. "SISTEM PAKAR PENENTUAN KOMPOSISI SKINCARE BERDASARKAN MASALAH KULIT WAJAH MENGGUNAKAN NATURAL LANGUAGE PROCESSING (NLP)", Jurnal Informatika dan Rekayasa Elektronik, 2025 Publication	<1 %
58	Sri Rahayuningsih. "Identifikasi Penerapan Dan Pemahaman Kesehatan Dan Keselamatan Kerja Dengan Metode Hazard And Operability Study (Hazop) Pada UMKM	<1 %

Eka Jaya", JATI UNIK : Jurnal Ilmiah Teknik dan
Manajemen Industri, 2019

Publication

59	Submitted to Universitas Brawijaya Student Paper	<1 %
60	digilib.uns.ac.id Internet Source	<1 %
61	jer.or.id Internet Source	<1 %
62	journal.ipm2kpe.or.id Internet Source	<1 %
63	repository.iainpurwokerto.ac.id Internet Source	<1 %
64	www.makalah.co.id Internet Source	<1 %
65	www.scribd.com Internet Source	<1 %
66	www.slideshare.net Internet Source	<1 %
67	www.springerprofessional.de Internet Source	<1 %
68	Yeni Risyani, Susi Japit, Conrad Bombongan, Tanda Selamat, Yuliana Yuliana. "Pelatihan Pembuatan Video dengan AI untuk Pemuda- Pemudi GPSI Medan Utara", Jurnal Minfo Polgan, 2024 Publication	<1 %
69	widuri.raharja.info Internet Source	<1 %

Exclude quotes Off

Exclude bibliography Off

Exclude matches Off