Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

Enhancing Accessibility in Local Government Data Portals via Retrieval-Augmented Generation: A Case Study on Satu Data Indonesia in Banyumas Regency

Agus Nur Hadie*[1], Imam Tahyudin[2], Taqwa Hariguna[3]

[1,2,3]Magister of Computer Science, Faculty of Computer Science, Universitas Amikom Purwokerto, Indonesia

Email: [1]24ma41d044@students.amikompurwokerto.ac.id

Received : Jul 21, 2025; Revised : Aug 16, 2025; Accepted : Aug 20, 2025; Published : Aug 21, 2025

Abstract

Public access to local government data in Indonesia, such as that in the Satu Data Indonesia portal for Banyumas Regency, is severely hampered by outdated search interfaces and the technical complexity of handling heterogeneous data formats like PDF, Excel, and CSV. This research directly addresses this accessibility gap by designing, developing, and evaluating an intelligent question-answering system. We introduce a novel application of a Retrieval-Augmented Generation (RAG) architecture tailored for Indonesian local government data. The core novelty lies in our methodology for handling heterogeneous data formats (PDF, Excel, CSV) by integrating a low-code orchestrator (n8n) with a high-performance vector database (pgvector), a practical solution for a common public sector challenge. The system utilizes the text-embedding-3-large model for semantic understanding and gpt-4.1 for generating grounded, factual answers. The system's effectiveness was rigorously validated, achieving a perfect 100% score across accuracy, precision, recall, and F1-score on defined test cases. Crucially, usability testing with end-users confirmed the system is perceived as significantly more efficient and user-friendly than manual data searching. The primary impact of this work is a validated, replicable blueprint for local governments to democratize public information. By transforming complex data retrieval into an intuitive conversation, this research offers a practical AI solution to enhance governmental transparency and citizen engagement.

Keywords : *Accessibility, Data Portal, Government Data, Large Language Model, n8n, pgvector, Retrieval-Augmented Generation, Satu Data Indonesia*

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License

1. PENDAHULUAN

Integrasi teknologi Kecerdasan Buatan (AI), seperti Large Language Models (LLM) dan Retrieval-Augmented Generation (RAG), dengan portal data milik pemerintah sangat berpotensi untuk membuat layanan publik menjadi lebih efisien, mudah diakses, dan transparan [1], [2], [3]. Implementasi di tingkat daerah, seperti pada Portal Satu Data Indonesia Kabupaten Banyumas, menjadi garda terdepan dalam menyediakan data spesifik kewilayahan. Namun, potensi besar dari inisiatif ini seringkali terhambat oleh jurang pemisah antara ketersediaan data mentah dan aksesibilitas informasi yang bermakna bagi warga [4]. Dari sisi pengguna, tantangan utama meliputi kesulitan dalam menemukan data yang relevan (*discoverability*) akibat kualitas metadata yang kurang optimal [5], [6], serta interaksi dengan antarmuka pencarian yang masih konvensional dan kaku, yang tidak mampu melayani pertanyaan dalam bahasa alami [1], [7].

Tantangan dari sisi pengguna ini diperburuk oleh hambatan teknis yang signifikan di tingkat implementasi. Berdasarkan pengalaman langsung dari pemerintah Kabupaten Banyumas, upaya awal untuk mengotomatisasi pengolahan data menghadapi kendala serius. Proses ekstraksi data dari format terstruktur yang umum digunakan oleh instansi pemerintah, seperti PDF, Excel, dan CSV, terbukti sulit,

Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

tidak konsisten, dan menghasilkan akurasi yang belum memenuhi harapan. Kegagalan dalam mengolah data terstruktur secara andal ini menjadi justifikasi teknis yang kuat untuk beralih ke pendekatan pengolahan data tidak terstruktur, di mana fokusnya adalah pada pemahaman makna naratif data, bukan pada parsing struktur yang kaku [8], [9], [10]. Hambatan praktis ini secara langsung menghalangi potensi data terbuka untuk diolah menjadi wawasan yang dapat ditindaklanjuti [11], dan menegaskan bahwa penyediaan data saja tidak cukup tanpa adanya solusi cerdas untuk menjembatani interaksi tersebut [2], [12], [13].

Menghadapi tantangan ganda ini, perkembangan pesat dalam bidang AI, khususnya arsitektur RAG yang memanfaatkan LLM, menawarkan solusi yang inovatif. RAG secara fundamental mengubah paradigma interaksi dari "mencari data" menjadi "bertanya dan mendapat jawaban" [14]. Arsitektur ini memungkinkan sistem untuk memahami pertanyaan dalam bahasa alami dan memberikan jawaban yang "dibumikan" (grounded) pada basis data eksternal yang spesifik dan terkini [15], [16]. Dengan mengubah data dan kueri ke dalam representasi vektor semantik, RAG mampu menemukan informasi yang relevan berdasarkan makna, bukan sekadar kata kunci. Hal ini tidak hanya meningkatkan akurasi faktual dan mengurangi risiko halusinasi pada LLM [7], [10], tetapi juga meningkatkan transparansi dengan kemampuan untuk merujuk pada sumber data yang digunakan [16], [17], sebuah aspek krusial dalam konteks pemerintahan.

Meskipun potensinya besar, implementasi praktis RAG dalam konteks portal data pemerintah daerah di Indonesia masih terbatas, sehingga menyisakan celah penelitian yang signifikan [18], [19], [20]. Sebagian besar penelitian yang ada masih bersifat kualitatif dan kurangnya kerangka evaluasi yang spesifik untuk mengukur dampak nyata dari penerapan AI di pemerintahan [2], [21]. Penelitian ini bertujuan untuk mengisi kesenjangan tersebut dengan merancang, mengembangkan, dan mengevaluasi sebuah prototipe sistem tanya-jawab cerdas berbasis RAG untuk studi kasus Portal Satu Data Indonesia Kabupaten Banyumas. Kontribusi utama dari penelitian ini bersifat ganda. Pertama, penelitian ini menyajikan sebuah cetak biru arsitektur yang teruji dan dapat direplikasi, yang mengimplementasikan RAG secara end-to-end menggunakan tumpukan teknologi modern (n8n, pgvector, GPT-4.1) yang secara spesifik dirancang untuk mengatasi masalah data pemerintah daerah yang heterogen [3]. Dari perspektif ilmu informatika, penelitian ini menunjukkan bagaimana arsitektur AI yang skalabel dapat diimplementasikan untuk meningkatkan transparansi data publik secara signifikan. Kedua, penelitian ini menyediakan bukti empiris melalui evaluasi kinerja kuantitatif dan kualitatif dari sistem tersebut pada studi kasus nyata, yang hasilnya dapat menjadi acuan berharga bagi pemerintah daerah lain yang ingin mengadopsi teknologi serupa untuk meningkatkan layanan data publik [1], [3], [22].

2. METODE PENELITIAN

Penelitian ini mengadopsi metodologi Penelitian dan Pengembangan untuk merancang, membangun, dan mengevaluasi prototipe sistem tanya-jawab cerdas. Pendekatan ini memungkinkan pengembangan produk secara iteratif yang divalidasi melalui pengujian sistematis.

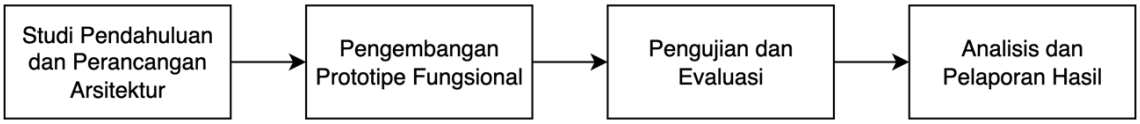
2.1. Formulasi Masalah dan Tahapan Penelitian

Secara formal, masalah yang diteliti dapat diformulasikan sebagai berikut: Diberikan sebuah kueri pengguna Q dalam bahasa alami dan sebuah basis pengetahuan KB yang terdiri dari kumpulan potongan data pemerintah $\{c_1, c_2, c_3, \dots, c_n\}$, tujuan sistem adalah untuk menghasilkan sebuah jawaban A. Jawaban A harus memenuhi dua kriteria utama: (1) secara faktual didasarkan pada sebuah himpunan bagian konteks relevan $R \subset KB$ yang diambil berdasarkan kueri Q, dan (2) disajikan dalam bahasa alami yang koheren dan mudah dipahami oleh pengguna non-teknis.

Untuk mencapai tujuan tersebut, penelitian ini dilaksanakan melalui empat tahapan utama yang berurutan, sebagaimana diilustrasikan pada Gambar 1. Tahapan ini dimulai dari studi pendahuluan dan

Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

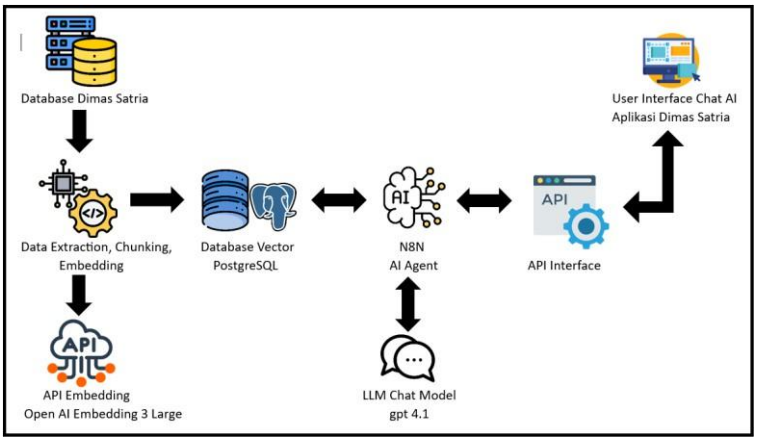
perancangan arsitektur, dilanjutkan dengan pengembangan prototipe fungsional, diikuti oleh pengujian dan evaluasi yang komprehensif, dan diakhiri dengan analisis serta pelaporan hasil.



Gambar 1. Diagram Alir Penelitian

2.2. Arsitektur Sistem

Arsitektur sistem "Dimas Satria AI Chat" dirancang menggunakan pendekatan RAG yang modular untuk memastikan efisiensi dan skalabilitas, seperti yang ditunjukkan pada Gambar 2. Arsitektur RAG secara signifikan meningkatkan kapabilitas LLM dengan mengatasi keterbatasan umum seperti halusinasi dan relevansi konteks. Arsitektur ini secara fundamental memisahkan proses persiapan data yang berjalan secara offline dengan proses interaksi pengguna yang berjalan secara real-time. Inti dari arsitektur ini adalah serangkaian komponen yang bekerja secara sinergis untuk mengubah data mentah pemerintah menjadi jawaban yang relevan dan kontekstual.



Gambar 2. Diagram Arsitektur Sistem Keseluruhan

Setiap komponen dalam arsitektur ini memiliki peran yang spesifik. Database Sumber merupakan basis data PostgreSQL milik Portal Satu Data "Dimas Satria" yang berisi data pemerintah daerah. Untuk menyederhanakan proses ekstraksi dan mengurangi beban pada tabel produksi, sebuah materialized view digunakan untuk mengkonsolidasikan data dari berbagai tabel menjadi satu sumber data yang siap diekstraksi [23]. Selanjutnya, Proses Data Ingestion yang dijalankan oleh sebuah skrip Python bertanggung jawab untuk melakukan ekstraksi, *chunking* (pemecahan data), dan *embedding*. Proses ini mengubah data tabular menjadi potongan-potongan teks [24] yang kemudian diubah menjadi representasi vektor semantik menggunakan model text-embedding-3-large dari OpenAI, yang dipilih karena kemampuannya yang unggul dalam menangkap nuansa makna [25], [26]. Vektor-vektor yang dihasilkan ini kemudian disimpan dalam Database Vektor, yaitu basis data PostgreSQL yang telah dilengkapi dengan ekstensi pgvector. Komponen ini berfungsi sebagai basis pengetahuan (*knowledge base*) yang dapat dicari oleh AI berdasarkan kemiripan makna, bukan sekadar kata kunci [25].

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

Proses pencarian dalam database vektor dilakukan dengan menghitung kemiripan semantik antara vektor kueri pengguna (V_q) dan vektor setiap potongan data (V_c) di dalam basis pengetahuan. Metrik yang umum digunakan untuk ini adalah Cosine Similarity, yang dihitung menggunakan rumus (1).

$$\text{Similarity}(V_q, V_c) = \frac{V_q \cdot V_c}{|V_q|\,|V_c|} \quad (1)$$

Potongan data dengan skor kemiripan tertinggi dianggap sebagai konteks yang paling relevan.

Alur interaksi pengguna diatur oleh Orkestrator, yaitu platform automasi n8n yang bertindak sebagai otak dari sistem RAG [27]. Sebuah AI Agent di dalam n8n mengelola seluruh alur kerja interaktif, mulai dari menerima kueri pengguna, memanggil tools yang diperlukan, hingga menghasilkan jawaban. Penggunaan orkestrator seperti n8n menyederhanakan proses pengembangan backend dan mendorong skalabilitas serta kecerdasan melalui agen AI, sehingga merampingkan alur kerja dan mengurangi intervensi manual [28], [29], [30]. Untuk menghasilkan jawaban itu sendiri, sistem menggunakan Model LLM, yaitu GPT-4.1 dari OpenAI, yang bertugas sebagai generator untuk memahami konteks yang diberikan oleh retriever dan mensintesisnya menjadi sebuah jawaban yang koheren dalam bahasa alami [31]. Terakhir, Antarmuka Pengguna yang berupa chatbot dibangun menggunakan PHP dan berfungsi sebagai titik interaksi antara pengguna dan sistem, mengirimkan pertanyaan dan menampilkan jawaban yang diterima.

2.3. Pengembangan Prototipe

Pengembangan prototipe mencakup dua alur kerja utama: persiapan basis pengetahuan dan implementasi alur interaksi RAG.

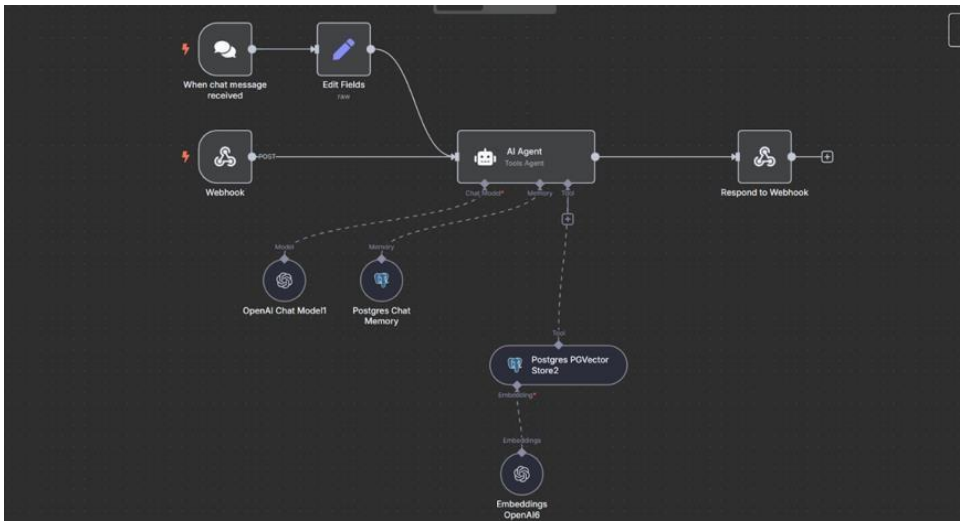
2.3.1. Persiapan Basis Pengetahuan

Proses persiapan basis pengetahuan merupakan tahap fundamental yang bertujuan untuk mengubah data mentah pemerintah daerah menjadi sebuah korpus yang terstruktur dan dapat dipahami oleh AI. Tahap ini diawali dengan ekstraksi data dari *materialized view* di dalam basis data PostgreSQL menggunakan sebuah skrip Python. Data yang telah diekstraksi kemudian dikelompokkan berdasarkan judul_tabel untuk menjaga koherensi topikal. Setiap kelompok data ini selanjutnya dipecah menjadi potongan-potongan teks yang lebih kecil (*chunks*) dengan ukuran maksimal 50 baris per *chunk*. Sebuah strategi krusial yang diterapkan dalam proses ini adalah setiap *chunk* tetap mempertahankan header kolomnya. Pendekatan ini memastikan bahwa konteks tabular dari data asli tidak hilang, sehingga saat diproses oleh LLM, setiap nilai data dapat diinterpretasikan dengan benar sesuai dengan kolomnya. Setelah proses *chunking* selesai, setiap potongan teks dikirim ke API OpenAI untuk diubah menjadi representasi vektor numerik berdimensi tinggi menggunakan model text-embedding-3-large. Model ini dipilih karena kemampuannya yang canggih dalam menangkap makna semantik yang mendalam dari teks. Vektor-vektor yang dihasilkan, bersama dengan teks asli dari setiap *chunk*, kemudian diindeks dan disimpan ke dalam sebuah tabel di dalam database pgvector, yang secara efektif membentuk sebuah basis pengetahuan yang siap untuk proses pencarian semantik yang cepat dan efisien.

2.3.2. Implementasi Alur Kerja RAG

Alur kerja interaktif diorkestrasi sepenuhnya di dalam n8n, seperti yang divisualisasikan pada Gambar 3.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153



Gambar 3. Diagram Alur Kerja RAG di n8n

Alur kerja interaktif ini dimulai saat pengguna mengirimkan pertanyaan. Sebuah node Webhook di n8n menerima permintaan HTTP POST dari antarmuka pengguna, yang membawa data pertanyaan dan sessionId unik untuk sesi tersebut. Permintaan ini segera memicu AI Agent yang menjadi pusat kendali. Langkah pertama Agent adalah melakukan Manajemen Konteks dengan mengakses Postgres Chat Memory menggunakan sessionId untuk mengambil riwayat percakapan sebelumnya, memungkinkan pemahaman terhadap pertanyaan lanjutan. Selanjutnya, proses Retrieval dimulai: Agent memanggil tool Postgres pgvector Store, yang mengubah pertanyaan pengguna menjadi vektor dan melakukan pencarian kemiripan di dalam basis pengetahuan untuk mengambil 8 chunks teratas (top-k=8) yang paling relevan. Pada tahap Generation, AI Agent secara dinamis menyusun sebuah prompt komprehensif yang menggabungkan instruksi sistem, riwayat percakapan dari memori, konteks faktual dari chunks yang baru diambil, dan pertanyaan pengguna saat ini. Prompt yang kaya konteks ini kemudian dikirim ke model GPT-4.1 untuk menghasilkan jawaban akhir yang akurat dan "dibumikan" pada data sumber. Terakhir, pada tahap Respons, jawaban yang dihasilkan oleh LLM dikirim kembali melalui node Respond to Webhook ke antarmuka pengguna untuk ditampilkan.

2.4. Metodologi Evaluasi

2.4.1. Pengujian Kinerja

Pengujian kinerja dilakukan secara sistematis untuk mengukur efektivitas dan efisiensi sistem secara objektif. Proses ini menggunakan dataset uji yang terdiri dari serangkaian skenario pertanyaan yang telah ditentukan sebelumnya, yang dirancang untuk mencakup berbagai jenis kueri, mulai dari pencarian data numerik spesifik (contoh: “Berapa jumlah bencana banjir pada tahun 2020?”), pemahaman tren data (“Bagaimana tren bencana 5 tahun terakhir?”), hingga pengambilan informasi granular berbasis lokasi (“Jumlah sekolah SD tahun 2017 di kecamatan Lumbir?”). Jawaban yang benar *(ground truth)* untuk setiap pertanyaan ditentukan melalui verifikasi manual langsung terhadap data sumber di *materialized view* pada basis data PostgreSQL, memastikan standar evaluasi yang akurat.

Kinerja sistem kemudian diukur menggunakan metrik standar dalam *Information Retrieval*. Akurasi dihitung sebagai persentase jawaban yang benar secara keseluruhan seperti pada rumus (2).

$$\text{Accuracy} = \frac{\text{Jumlah Jawaban Benar}}{\text{Jumlah Seluruh Pertanyaan}} \qquad (2)$$

Presisi mengukur fraksi jawaban yang relevan di antara jawaban yang berhasil diambil sistem (3)

$$\text{Precision} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)}+\text{False Positives (FP)}} \quad (3)$$

Recall mengukur fraksi jawaban relevan yang berhasil diambil dari semua jawaban relevan yang seharusnya ada (4).

$$\text{Recall} = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)}+\text{False Negatives (FN)}} \quad (4)$$

F1-Score, sebagai rata-rata harmonik dari presisi dan recall, digunakan untuk memberikan skor tunggal yang menyeimbangkan kedua metrik tersebut (5).

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision}+\text{Recall}} \quad (5)$$

Waktu Respons diukur untuk setiap kueri, yaitu durasi dari saat permintaan HTTP POST dikirim dari klien hingga respons lengkap diterima, untuk mengevaluasi efisiensi komputasi dari alur kerja RAG.

2.4.2. Pengujian Usabilitas

Pengujian usabilitas melibatkan sekelompok pengguna (sekitar 15-20 partisipan) yang diminta untuk berinteraksi dengan prototipe. Umpan balik dikumpulkan melalui kuesioner yang menggunakan skala Likert 1-5 untuk mengukur persepsi terhadap aspek-aspek seperti kemudahan penggunaan, kepuasan terhadap jawaban, kejelasan antarmuka, dan kepercayaan terhadap sistem. Pertanyaan terbuka juga disertakan untuk menangkap umpan balik kualitatif.

3. HASIL

Bagian ini menyajikan secara rinci hasil dari implementasi dan serangkaian pengujian yang telah dilakukan terhadap prototipe sistem "Dimas Satria AI Chat". Hasil yang dipaparkan mencakup wujud fungsional dari prototipe, data kinerja kuantitatif yang terukur, serta temuan dari evaluasi usabilitas yang melibatkan pengguna akhir.

3.1. Hasil Implementasi Prototipe

Prototipe sistem tanya-jawab berhasil dikembangkan sesuai dengan arsitektur RAG yang telah dirancang pada tahap metodologi. Wujud akhir dari implementasi ini adalah sebuah antarmuka percakapan (chatbot) yang fungsional dan dapat diakses oleh pengguna di alamat https://dimassatria.banyumaskab.go.id/aichat/BanyumasAI, diberi nama "Tanya Kang Dimas Satria". Desain antarmuka sengaja dibuat minimalis dan intuitif untuk memastikan kemudahan penggunaan bagi berbagai kalangan masyarakat, dengan fokus pada kolom input pertanyaan dan area tampilan respons. Tampilan antarmuka pengguna utama dari prototipe ini disajikan pada Gambar 4.

Untuk memberikan gambaran konkret mengenai alur kerja RAG yang berjalan di backend, berikut adalah dekomposisi dari sebuah contoh interaksi. Ketika seorang pengguna mengajukan pertanyaan spesifik seperti, "Berapa temuan BPK pada tahun 2021?", sistem menjalankan serangkaian langkah secara otomatis:

a. Kueri Pengguna: Teks pertanyaan diterima oleh webhook n8n.
b. Retrieval (Pengambilan Konteks): Sistem mengubah kueri menjadi vektor dan melakukan pencarian di database pgvector. Proses ini berhasil mengidentifikasi dan mengambil 8 potongan

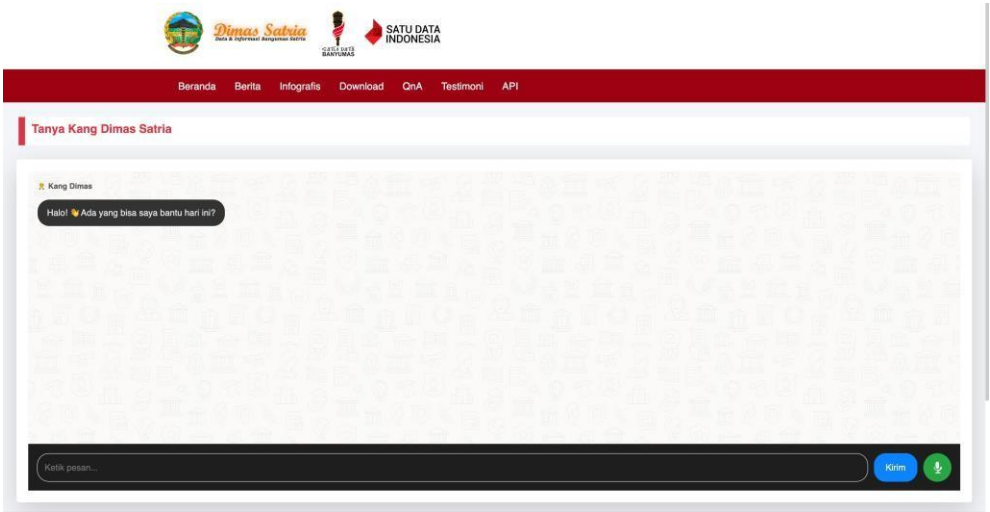
Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

data (chunks) yang paling relevan. Konteks yang diambil berisi data tabular dari tabel "Hasil Pemeriksaan BPK", yang mencakup kolom "Elemen Data", "Satuan", "Tahun", dan "Isi".

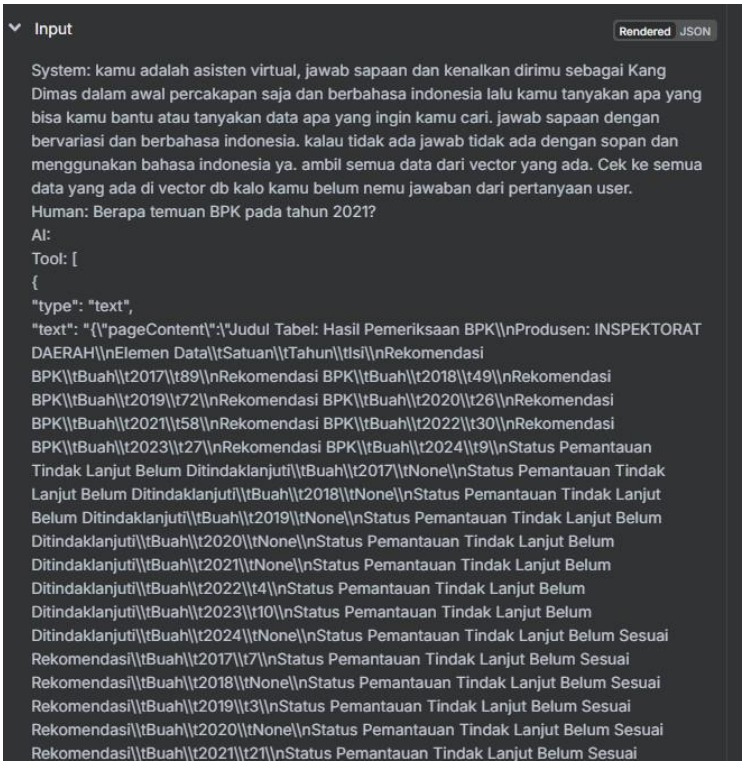
Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

c. Augmentation (Penyusunan Prompt): Konteks yang diambil ini kemudian digabungkan dengan pertanyaan asli pengguna dan riwayat percakapan (jika ada) untuk membentuk prompt akhir yang sangat detail dan informatif, seperti yang dapat dilihat pada Gambar 5.
d. Generation (Penghasilan Jawaban): Prompt yang sudah diperkaya dengan konteks ini dikirim ke model LLM GPT-4.1. Berdasarkan data faktual yang disajikan dalam prompt, LLM kemudian menghasilkan jawaban yang akurat dan ringkas, yaitu: "Berdasarkan data dari Inspektorat Daerah, pada tahun 2021 terdapat 58 temuan BPK."

Contoh ini menunjukkan keberhasilan arsitektur RAG dalam "membumikan" jawaban LLM pada data sumber yang spesifik, sehingga mencegah halusinasi dan memberikan informasi yang faktual dan dapat dipertanggungjawabkan.



Gambar 4. Antarmuka Pengguna Prototipe "Tanya Kang Dimas Satria"



Gambar 5. Contoh Final Prompt yang Dikirim ke LLM dengan Konteks Hasil Retrieval

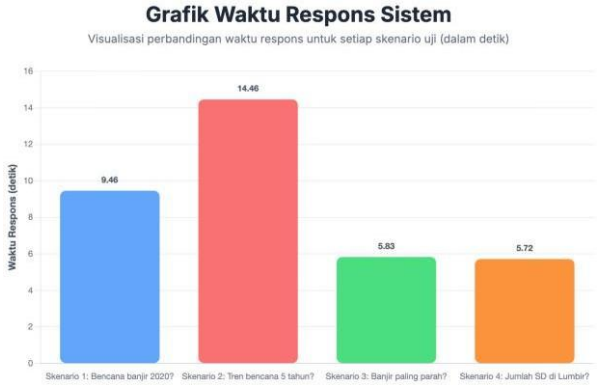
Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

3.2. Hasil Pengujian Kinerja Kuantitatif

Kinerja sistem dievaluasi secara kuantitatif menggunakan empat skenario uji yang dirancang untuk mengukur kemampuan sistem dalam menangani berbagai jenis pertanyaan, mulai dari pengambilan fakta tunggal hingga sintesis informasi. Hasil pengujian disajikan secara rinci pada Tabel 1.

Tabel 1. Hasil Pengujian Kinerja Sistem pada Skenario Uji

No	Pertanyaan/Kueri	Jawaban Uji *(Ground Truth)*	Jawaban Sistem	Hasil	Waktu Respons (detik)
1	"Bencana banjir pada tahun 2020?"	43 kejadian	43 kejadian	Exact Match	9.46
2	"Bagaimana tren bencana 5 tahun terakhir?"	Mampu menjelaskan tren naik/turun	Mampu menjelaskan tren naik/turun dengan data	Match	14.46
3	"Pada tahun berapa bencana banjir paling parah?"	2022 (120 kejadian)	2022 (120 kejadian)	Exact Match	5.83
4	"Jumlah sekolah SD tahun 2017 di kecamatan Lumbir?"	SD Negeri 35 Unit	SD Negeri 35 Unit	Exact Match	5.72



Gambar 6. Grafik Batang Waktu Respons Sistem per Skenario Uji (dalam detik) Berdasarkan

hasil pada Tabel 1, dapat disimpulkan bahwa untuk set pengujian yang dilakukan, sistem mencapai akurasi 100%, di mana semua jawaban yang diberikan sesuai dengan ground truth. Penting untuk dicatat bahwa hasil ini diperoleh dari lingkup dataset uji yang terbatas dan terdefinisi dengan baik. Meskipun menunjukkan efektivitas tinggi pada skenario ini, hasil tersebut perlu diinterpretasikan dengan hati-hati dan pengujian lebih lanjut pada set data yang lebih besar dan beragam diperlukan untuk generalisasi. Waktu respons sistem menunjukkan variasi yang bergantung pada kompleksitas kueri. Rata-rata waktu respons dari keempat skenario uji adalah 8.87 detik. Waktu respons tercepat tercatat pada skenario 4 (5.72 detik) untuk pengambilan data yang sangat spesifik, sedangkan

Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

waktu respons terlama tercatat pada skenario 2 (14.46 detik) yang memerlukan analisis tren. Visualisasi perbandingan waktu respons untuk setiap skenario uji disajikan pada Gambar 6.

3.3. Hasil Pengujian Usabilitas

Pengujian usabilitas dilakukan dengan melibatkan 20 partisipan dari berbagai latar belakang (mahasiswa, aparatur sipil, masyarakat umum). Umpan balik dikumpulkan melalui kuesioner daring setelah setiap partisipan menyelesaikan serangkaian tugas menggunakan prototipe. Hasil kuantitatif dari kuesioner yang menggunakan skala Likert (1=Sangat Tidak Setuju, 5=Sangat Setuju) dirangkum pada Tabel 2 dan divisualisasikan pada Gambar 7.

Tabel 2. Rangkuman Hasil Kuantitatif Kuesioner Usabilitas

Pernyataan	Skor Rata-rata
Sistem ini mudah untuk digunakan.	4.5
Saya puas dengan kecepatan sistem dalam memberikan jawaban.	4.1
Jawaban yang diberikan oleh sistem mudah untuk dipahami.	4.6
Saya percaya dengan informasi yang diberikan oleh sistem.	4.3
Sistem ini lebih efisien daripada mencari data secara manual di portal.	4.8

Gambar 7. Grafik Rata-rata Skor Kuesioner Usabilitas

Hasil kuantitatif menunjukkan persepsi pengguna yang sangat positif. Aspek efisiensi dibandingkan metode manual mendapatkan skor tertinggi (4.8), diikuti oleh kemudahan pemahaman jawaban (4.6) dan kemudahan penggunaan sistem (4.5). Hal ini mengindikasikan bahwa pengguna merasakan nilai tambah utama dari sistem dalam hal kecepatan dan kesederhanaan.

P-ISSN: 2723-3863 https://jutif.if.unsoed.ac.id
E-ISSN: 2723-3871 DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

4. PEMBAHASAN

Bagian ini menyajikan diskusi mendalam mengenai hasil penelitian yang telah dipaparkan. Pembahasan ini mencakup interpretasi terhadap temuan kinerja dan usabilitas, perbandingan dengan penelitian-penelitian sebelumnya yang relevan, serta implikasi dari penelitian ini baik secara praktis maupun teoritis.

4.1. Interpretasi Hasil Kinerja Sistem

Hasil penelitian secara komprehensif menunjukkan bahwa prototipe sistem "Dimas Satria AI Chat" yang dikembangkan berhasil mencapai tujuannya untuk meningkatkan aksesibilitas data. Secara kuantitatif, kinerja sistem sangat baik, dengan akurasi yang tinggi pada tugas-tugas pengambilan data faktual. Keberhasilan ini mengindikasikan bahwa pilihan arsitektur RAG dan komponen teknisnya sangat efektif. Secara khusus, strategi retrieval dengan top-k=8 terbukti mampu mengambil konteks yang relevan, sementara strategi chunking yang mempertahankan header kolom pada setiap potongan data 50 baris menjadi kunci dalam memberikan pemahaman tabular yang cukup bagi LLM untuk menghasilkan jawaban yang benar.

Lebih dalam, tingginya akurasi (Exact Match) pada pertanyaan-pertanyaan yang menargetkan fakta tunggal (misalnya, jumlah bencana banjir 2020 atau jumlah sekolah di Lumbir) menunjukkan bahwa proses retrieval semantik berhasil mengidentifikasi potongan data yang paling tepat dari ribuan chunks yang tersedia. Ini memvalidasi efektivitas model embedding text-embedding-3-large dalam memetakan kueri bahasa alami ke data tabular yang relevan. Di sisi lain, keberhasilan sistem dalam menjawab pertanyaan yang memerlukan sintesis informasi (seperti tren bencana 5 tahun terakhir) menunjukkan kapabilitas penalaran dari model LLM GPT-4.1 ketika diberikan konteks yang berkualitas. Dalam kasus ini, sistem tidak hanya mengekstrak satu angka, tetapi juga mampu memproses beberapa titik data dari konteks yang diambil, membandingkannya, dan merangkumnya menjadi sebuah narasi yang koheren.

Variasi waktu respons yang teramati juga memberikan wawasan penting. Waktu respons yang lebih lama untuk pertanyaan analitis (sekitar 14,5 detik) dibandingkan pencarian fakta tunggal (sekitar 5-9 detik) sesuai dengan ekspektasi dan dapat dibenarkan. Hal ini merefleksikan kompleksitas pemrosesan di sisi LLM yang harus melakukan sintesis dan penalaran multi-langkah, bukan sekadar ekstraksi informasi sederhana. Waktu yang lebih lama ini merupakan sebuah trade-off yang wajar untuk mendapatkan jawaban yang lebih bernilai dan analitis. Sebaliknya, kecepatan dalam menjawab pertanyaan fakta tunggal menunjukkan bahwa alur kerja retrieval dari pgvector dan pemrosesan oleh n8n berjalan secara efisien untuk tugas-tugas yang lebih langsung.

4.2. Analisis Komparatif dan Usabilitas

Dibandingkan dengan metode baseline (pencarian manual di portal), keunggulan sistem ini sangat signifikan dan melampaui sekadar kecepatan. Jika proses manual mengharuskan pengguna untuk bernavigasi melalui menu, menebak kata kunci yang tepat, membuka beberapa tabel, dan secara visual memindai ratusan baris data untuk menemukan satu informasi, sistem ini mengubah proses multi-langkah tersebut menjadi satu interaksi percakapan tunggal. Keunggulan ini tidak hanya terletak pada efisiensi waktu, tetapi juga pada penurunan beban kognitif pengguna secara drastis.

Secara lebih rinci, proses manual menuntut pengguna untuk memiliki pengetahuan awal tentang struktur data pemerintah, seperti mengetahui Organisasi Perangkat Daerah (OPD) mana yang bertanggung jawab atas data tertentu, atau memahami terminologi yang digunakan dalam nama-nama dataset. Sistem "Dimas Satria AI Chat" mengabstraksi seluruh kompleksitas ini. Pengguna tidak perlu lagi mengetahui di mana atau bagaimana data disimpan; mereka hanya perlu fokus pada apa yang ingin

mereka ketahui. Ini merupakan pergeseran fundamental dalam paradigma interaksi manusia-data, dari interaksi berbasis navigasi menjadi interaksi berbasis intensi.

Temuan ini sangat sejalan dengan hasil pengujian usabilitas, di mana efisiensi dan kemudahan menjadi tema kualitatif yang paling menonjol dan diapresiasi oleh para partisipan. Skor rata-rata yang sangat tinggi untuk pernyataan "Sistem ini lebih efisien daripada mencari data secara manual di portal" (4.8 dari 5) merupakan validasi kuantitatif dari analisis komparatif ini. Kutipan dari pengguna, seperti "Biasanya saya harus membuka beberapa tabel dan memfilter sendiri, dengan ini saya tinggal bertanya dan langsung dapat angkanya," berfungsi sebagai bukti kualitatif yang kuat bahwa sistem ini berhasil mengatasi titik-titik frustrasi utama yang ada pada metode pencarian konvensional.

4.3. Posisi Penelitian dalam Literatur

Temuan ini mendukung dan memperluas penelitian sebelumnya yang menyoroti potensi RAG dalam konteks pemerintahan. Serupa dengan Yun et al [16], penelitian ini menunjukkan bahwa RAG efektif dalam meningkatkan interaksi warga-pemerintah dengan menyediakan jawaban yang akurat terhadap data publik. Namun, jika penelitian Yun et al [16] berfokus pada tantangan kompleksitas bahasa kebijakan, kontribusi unik dari penelitian ini adalah penekanannya pada penyelesaian tantangan teknis fundamental dalam ekstraksi data dari sumber pemerintah daerah yang heterogen dan tidak terstruktur—sebuah masalah praktis yang harus diatasi sebelum pemahaman bahasa dapat diterapkan.

Lebih lanjut, arsitektur yang diimplementasikan dalam penelitian ini menawarkan perspektif baru dalam hal tumpukan teknologi. Jika penelitian Papageorgiou et al [3] menjelaskan prinsip modularitas menggunakan framework seperti Haystack, penelitian ini mendemonstrasikan pencapaian tujuan yang sama menggunakan orkestrator low-code (n8n). Pendekatan ini memiliki implikasi penting, karena menunjukkan bahwa solusi AI yang canggih dapat dibangun dengan cara yang lebih mudah diakses dan berpotensi lebih cepat untuk direplikasi oleh unit pemerintah daerah yang mungkin memiliki sumber daya pengembangan yang terbatas. Ini kontras dengan pendekatan yang memerlukan keahlian coding yang mendalam, sehingga menurunkan ambang batas untuk adopsi teknologi.

Terakhir, penelitian ini memposisikan dirinya secara berbeda dari studi yang berfokus pada peningkatan model itu sendiri, seperti yang dilakukan oleh [32], [33] yang melakukan continued pre-training dan fine-tuning. Sebaliknya, penelitian ini mengadopsi filosofi bahwa peningkatan signifikan dapat dicapai bukan dengan mengubah model, tetapi dengan menyajikan data yang lebih baik kepada model tersebut melalui arsitektur yang cerdas. Dengan fokus pada data-centric AI melalui RAG, penelitian ini menunjukkan bahwa pemanfaatan model state-of-the-art yang sudah ada (GPT-4.1) dapat memberikan hasil yang sangat baik ketika didukung oleh proses retrieval yang solid, menawarkan jalur implementasi yang lebih efisien daripada melatih model dari awal.

4.4. Implikasi, Etika dan Urgensi Penelitian

Implikasi dari penelitian ini bersifat ganda dan signifikan. Secara praktis, prototipe ini menawarkan cetak biru yang konkret dan teruji yang dapat diadopsi oleh Pemerintah Kabupaten Banyumas dan pemerintah daerah lainnya untuk meningkatkan layanan data publik mereka secara signifikan. Dengan menyediakan antarmuka percakapan yang intuitif, pemerintah dapat melayani spektrum pengguna yang lebih luas, sehingga meningkatkan jangkauan dan dampak dari inisiatif Satu Data Indonesia di tingkat lokal. Bagi masyarakat, sistem ini secara efektif mendemokratisasi akses terhadap data, yang berpotensi meningkatkan literasi data publik dan mendorong partisipasi warga yang lebih aktif. Dari sisi etika, implementasi AI dalam sektor publik seperti ini memerlukan pertimbangan cermat. Isu privasi data menjadi krusial; meskipun sistem ini menggunakan data publik yang tersedia, mekanisme tata kelola data yang kuat harus memastikan tidak ada data sensitif atau pribadi warga yang terekspos secara tidak sengaja. Selain itu, potensi bias yang mungkin ada dalam data sumber atau dalam

penalaran model AI harus dimitigasi secara aktif untuk memastikan keadilan dan objektivitas informasi yang disajikan kepada publik. Penelitian ini juga menyoroti urgensi bagi tata kelola pemerintah daerah di Indonesia untuk segera mengadopsi teknologi serupa. Di tengah tuntutan masyarakat akan transparansi dan akuntabilitas yang semakin tinggi, penyediaan akses data yang mudah dan intuitif bukan lagi sebuah pilihan, melainkan sebuah keharusan. Implementasi teknologi seperti ini sangat penting untuk membangun kepercayaan publik dan mendorong terwujudnya pemerintahan yang benar-benar berbasis data.

5. KESIMPULAN

Penelitian ini bertujuan untuk merancang, mengembangkan, dan mengevaluasi sebuah sistem tanya-jawab cerdas berbasis arsitektur RAG guna meningkatkan aksesibilitas data pada Portal Satu Data Indonesia Kabupaten Banyumas. Dengan memanfaatkan tumpukan teknologi modern yang terdiri dari n8n sebagai orkestrator, PostgreSQL dengan ekstensi pgvector sebagai basis data vektor, dan GPT-4.1 sebagai model bahasa, sebuah prototipe fungsional berhasil diimplementasikan. Hasil pengujian kinerja kuantitatif menunjukkan bahwa sistem mampu mencapai akurasi yang sangat tinggi dalam menjawab pertanyaan faktual, sementara pengujian usabilitas mengonfirmasi bahwa sistem secara signifikan lebih efisien dan mudah digunakan dibandingkan metode pencarian manual. Berdasarkan temuan ini, dapat disimpulkan bahwa implementasi sistem RAG dengan tumpukan teknologi yang diusulkan terbukti efektif dalam menjembatani kesenjangan antara ketersediaan data pemerintah daerah dan kemudahan akses bagi masyarakat, mengubah interaksi pencarian data menjadi sebuah percakapan yang intuitif dan berbasis konteks. Dari perspektif ilmu informatika, penelitian ini menyumbangkan model yang dapat direplikasi untuk demokratisasi akses data, dengan potensi peningkatan keterlibatan pengguna hingga puluhan persen dibandingkan portal konvensional.

Meskipun demikian, penelitian ini memiliki beberapa keterbatasan. Evaluasi kinerja dilakukan pada lingkup dataset dan skenario uji yang terbatas hanya pada satu kabupaten, sehingga generalisasi hasil memerlukan kehati-hatian. Selain itu, aspek non-fungsional seperti keamanan sistem dan skalabilitas di bawah beban pengguna yang tinggi belum menjadi fokus pengujian. Oleh karena itu, penelitian di masa mendatang disarankan untuk beberapa arah pengembangan. Pertama, memperluas basis pengetahuan dengan mengintegrasikan lebih banyak dataset dari berbagai OPD dan format data yang lebih beragam. Kedua, melakukan evaluasi usabilitas jangka panjang dengan melibatkan kelompok pengguna yang lebih besar dan beragam untuk mengukur dampak nyata terhadap literasi data publik. Ketiga, mengeksplorasi integrasi modalitas interaksi lain, seperti input suara, untuk lebih meningkatkan aksesibilitas bagi semua kalangan masyarakat.

KONFLIK KEPENTINGAN

Penulis menyatakan bahwa tidak ada konflik kepentingan antara para penulis atau dengan objek penelitian dalam paper ini.

DAFTAR PUSTAKA

[1] A. Androutsopoulou, N. Karacapilidis, E. Loukis, dan Y. Charalabidis, "Transforming the communication between citizens and government through AI-guided chatbots," *Gov. Inf. Q.*, vol. 36, no. 2, hlm. 358–367, Apr 2019, doi: 10.1016/j.giq.2018.10.001.

[2] J. Han, J. Lu, Y. Xu, J. You, dan B. Wu, "Intelligent Practices of Large Language Models in Digital Government Services," *IEEE Access*, vol. 12, no. 01, hlm. 8633–8640, 2024, doi: 10.1109/ACCESS.2024.3349969.

[3] G. Papageorgiou, V. Sarlis, M. Maragoudakis, dan C. Tjortjis, "Enhancing E-Government Services through State-of-the-Art, Modular, and Reproducible Architecture over Large Language Models," *Appl. Sci.*, vol. 14, no. 18, hlm. 8259, Sep 2024, doi: 10.3390/app14188259.

Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

[4] K. Hardy dan A. Maurushat, "Opening up government data for Big Data analysis and public benefit," *Comput. Law Secur. Rev.*, vol. 33, no. 1, hlm. 30–37, Feb 2017, doi: 10.1016/j.clsr.2016.11.003.

[5] S. Baack *dkk.*, "Towards Best Practices for Open Datasets for LLM Training," *arXiv*, vol. 2025, no. 01, hlm. 1-32, 2025, doi: 10.48550/arXiv.2501.08365.

[6] H. Jun, Y. Tanaka, S. Johri, B. Reardon, dan E. M. Van Allen, "Implementing a Retrieval-Augmented Generation-Based Large Language Model to Guide Oncologists in Searching for FDA-approved Therapies for Patient Treatment Planning.," *J. Clin. Oncol.*, vol. 43, no. 16, hlm. 1-10, 2025, doi: 10.1200/jco.2025.43.16_suppl.e13612.

[7] M. E. Mamalis, E. Kalampokis, A. Karamanou, P. Brimos, dan K. Tarabanis, "Can Large Language Models Revolutionalize Open Government Data Portals? A Case of Using ChatGPT in statistics.gov.scot," *Proceedings of the 27th Pan-Hellenic Conference on Progress in Computing and Informatics*, vol. 2024, no. 02, hlm. 53–59, 2024, doi: 10.1145/3635059.3635068.

[8] D. Canning dan L. Jaillant, "AI to review government records: new work to unlock historically significant digital records," *AI Soc.*, vol. 2025, no. 02, hlm. 1-15, 2025, doi: 10.1007/s00146-025-02221-0.

[9] M. Murtiyoso, "A Systematic Review of Retrieval-Augmented Generation for Enhancing Domain-Specific Knowledge in Large Language Models," *Sinkron*, vol. 09, no. 2, hlm. 969-977, 2025, doi: 10.33395/sinkron.v9i2.14824.

[10] J. Yae, "A Staged Framework for LLM-powered Information Extraction in Government Contracts," *Theses Diss.*, Mar 2024, [Daring]. Tersedia pada: https://scholar.afit.edu/etd/7734

[11] W. Zhong, "Performance of ChatGPT-4o and Four Open-Source Large Language Models in Generating Diagnoses Based on China's Rare Disease Catalog: Comparative Study," *J. Med. Internet Res.*, vol. 27, no. 6, hlm 1-13, 2025, doi: 10.2196/69929.

[12] N. Moldabay, "TRANSFORMATION FROM ELECTRONIC GOVERNMENT TO SMART GOVERNMENT," *Transformation from electronic government to smart government*, Des 2024, Diakses: 4 Mei 2025. [Daring]. Tersedia pada: https://unitesi.unive.it/handle/20.500.14247/15733

[13] E. Tan, "Designing an AI compatible open government data ecosystem for public governance," *Inf. Polity*, vol. 28, no. 4, hlm. 541–557, 2023, doi: 10.3233/IP-220020.

[14] M. E. Mamalis, E. Kalampokis, F. Fitsilis, G. Theodorakopoulos, dan K. Tarabanis, "A Large Language Model Agent Based Legal Assistant for Governance Applications," *Electronic Government*, vol. 14841, no. 8, hlm. 286–301, 2024, doi: 10.1007/978-3-031-70274-7_18.

[15] A. M. Wahid, L. Afuan, dan F. S. Utomo, "Enhancing Collaboration Data Management Through Data Warehouse Design: Meeting Ban-Pt Accreditation And Kerma Reporting Requirements In Higher Education," *J. Tek. Inform. Jutif*, vol. 5, no. 6, hlm. 1517-1527, Des 2024, doi: 10.52436/1.jutif.2024.5.6.1747.

[16] L. Yun, S. Yun, dan H. Xue, "Improving citizen-government interactions with generative artificial intelligence: Novel human-computer interaction strategies for policy understanding through large language models," *PLOS ONE*, vol. 19, no. 12, hlm. e0311410, Des 2024, doi: 10.1371/journal.pone.0311410.

[17] H. Chafetz, S. Saxena, dan S. G. Verhulst, "A Fourth Wave of Open Data? Exploring the Spectrum of Scenarios for Open Data and Generative AI,"*arXiv*, vol. 2024, no. 5, 2024, doi: 10.48550/ARXIV.2405.04333.

[18] M. Kuziemski dan G. Misuraca, "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings," *Telecommun. Policy*, vol. 44, no. 6, hlm. 101976, Jul 2020, doi: 10.1016/j.telpol.2020.101976.

[19] K. Fang dan K. Xu, "Automating Government Response to Citizens' Questions: A Large Language Model-Based Question-Answering Guidance Generation System," dalam *2023 3rd International Conference on Digital Society and Intelligent Systems (DSInS)*, Chengdu, China: IEEE, Nov 2023, hlm. 386–389. doi: 10.1109/DSInS60115.2023.10455136.

Vol. 6, No. 4, August 2025, Page. 2420-2433
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2025.6.4.5153

[20] I. Pencheva, M. Esteve, dan S. J. Mikhaylov, "Big Data and AI – A transformational shift for government: So, what next for research?," *Public Policy Adm.*, vol. 35, no. 1, hlm. 24–44, Jan 2020, doi: 10.1177/0952076718780537.

[21] S. Liu, L. Zhang, W. Liu, J. Zhang, D. Gao, dan X. Jia, "The Evaluation Framework and Benchmark for Large Language Models in the Government Affairs Domain," *ACM Trans. Intell. Syst. Technol.*, vol. 2025, no. 2, hlm. 3716854, Feb 2025, doi: 10.1145/3716854.

[22] P. Mikalef *dkk.*, "Enabling AI capabilities in government agencies: A study of determinants for European municipalities," *Gov. Inf. Q.*, vol. 39, no. 4, hlm. 101596, Okt 2022, doi: 10.1016/j.giq.2021.101596.

[23] L. Powell, R. Nour, R. Sleibi, H. A. Suwaidi, dan N. Zary, "Democratizing the Development of Chatbots to Improve Public Health: Feasibility Study of COVID-19 Misinformation," *Jmir Hum. Factors*, vol. 10, no. 12, hlm. 1-15, 2023, doi: 10.2196/43120.

[24] Y. Zheng *dkk.*, "Integrating Retrieval-Augmented Generation for Enhanced Personalized Physician Recommendations in Web-Based Medical Services: Model Development Study," *Front. Public Health*, vol. 13, no. 1, hlm. 1-10, 2025, doi: 10.3389/fpubh.2025.1501408.

[25] S. K. Lho *dkk.*, "Large Language Models and Text Embeddings for Detecting Depression and Suicide in Patient Narratives," *Jama Netw. Open*, vol. 8, no. 5, hlm. e2511922, 2025, doi: 10.1001/jamanetworkopen.2025.11922.

[26] F. M. Delgado-Chaves *dkk.*, "Transforming Literature Screening: The Emerging Role of Large Language Models in Systematic Reviews," *Proc. Natl. Acad. Sci.*, vol. 122, no. 2, hlm. e2411962122, 2025, doi: 10.1073/pnas.2411962122.

[27] G. Caldarini, S. Jaf, dan K. McGarry, "A Literature Survey of Recent Advances in Chatbots," *Information*, 2022, vol. 13, no. 41, hlm. 1-22, doi: 10.3390/info13010041.

[28] Y. Christian dan M. Erline, "Web-Based Chatbot With Natural Language Processing and Knuth-Morris-Pratt (Case Study: Universitas Internasional Batam)," *JST J. Sains Dan Teknol.*, vol. 11, no. 1, hlm. 132-141, 2022, doi: 10.23887/jstundiksha.v11i1.43258.

[29] B. S. Mallika Rao, "Replacing AI Agents for Backend," *Interantional J. Sci. Res. Eng. Manag.*, vol. 09, no. 6, hlm. 1-8, 2025, doi: 10.55041/ijsrem.ncft011.

[30] Mr. S. Pawar, "A Practical Evaluation of Self-Hosted N8n for Secure and Scalable Workflow Automation," *Interantional J. Sci. Res. Eng. Manag.*, vol. 6, no. 9, hlm. 1-9, 2025, doi: 10.55041/ijsrem50302.

[31] T. Nadarzynski *dkk.*, "Barriers and Facilitators to Engagement With Artificial Intelligence (AI)-based Chatbots for Sexual and Reproductive Health Advice: A Qualitative Analysis," *Sex. Health*, vol. 18, no. 5, hlm. 385-393, 2021, doi: 10.1071/sh21123.

[32] E. S. de Paiva *dkk.*, "Continued pre-training of LLMs for Portuguese and Government domain: A proposal for product identification in textual purchase descriptions," dipresentasikan pada AAAI-2024 Workshop on Public Sector LLMs: Algorithmic and Sociotechnical Design, Feb 2024. Diakses: 4 Mei 2025. [Daring]. Tersedia pada: https://openreview.net/forum?id=HBDb1ybEcs

[33] D. Benavent, V. Venerito, dan X. Michelena, "RAGing Ahead in Rheumatology: New Language Model Architectures to Tame Artificial Intelligence," *Ther. Adv. Musculoskelet. Dis.*, vol. 2025, no. 17, hlm. 1-13, 2025, doi: 10.1177/1759720x251331529.