

# Pengembangan Aplikasi Percakapan Suara Secara Real-Time untuk Simulasi Wawancara Kerja Berbasis Kecerdasan Buatan

Muhamad Jikril Aryanda<sup>1</sup>, Erick Andhika

<sup>1</sup>Program Studi Teknik Komputer, Politeknik Sukabumi

Jl. Babakan Sirna No.25, Benteng, Kec. Warudoyong, Kota Sukabumi, Jawa Barat 43132

jikrilar2903@gmail.com

---

## Abstrak

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) telah menghadirkan peluang baru dalam membangun aplikasi interaktif yang dapat mendukung berbagai kebutuhan manusia. Salah satu penerapannya adalah pada simulasi wawancara kerja, di mana calon pelamar dapat berlatih berkomunikasi secara langsung dengan sistem yang mampu merespon suara pengguna secara *real-time*. Tugas akhir ini bertujuan untuk merancang dan mengembangkan aplikasi percakapan suara berbasis *web* yang memanfaatkan teknologi *speech recognition*, *natural language processing (NLP)*, dan *text-to-speech (TTS)* untuk menghadirkan pengalaman wawancara yang lebih realistis. Implementasi sistem dilakukan dengan mengintegrasikan teknologi *WebRTC* sebagai media komunikasi suara *real-time* serta model *AI* untuk memahami pertanyaan dan memberikan respon secara cepat dan kontekstual. Hasil pengujian menunjukkan bahwa aplikasi mampu memberikan respon dengan tingkat latensi rendah dan akurasi pengenalan suara yang cukup baik dalam skenario wawancara kerja. Dengan adanya aplikasi ini, pengguna diharapkan dapat meningkatkan kemampuan komunikasi, mengasah keterampilan menjawab pertanyaan, serta memperoleh pengalaman simulasi wawancara yang lebih efektif dan efisien sebelum menghadapi wawancara kerja sesungguhnya.

**Kata kunci:** Aplikasi *Web*, Kecerdasan Buatan, Percakapan Suara *Real-Time*, Simulasi Wawancara Kerja, *WebRTC*, *NLP*, *TTS*

## Abstract

*The rapid advancement of Artificial Intelligence (AI) has created new opportunities to develop interactive applications that support various human needs. One of its applications is in job interview simulations, where applicants can practice direct communication with a system capable of responding to user speech in real time. This study aims to design and develop a web-based voice conversation application that integrates speech recognition, natural language processing (NLP), and text-to-speech (TTS) technologies to provide a more realistic interview experience. The system is implemented by utilizing WebRTC as a medium for real-time voice communication and AI models to interpret questions and deliver contextual responses quickly. The testing results show that the application can provide responses with low latency and satisfactory speech recognition accuracy in job interview scenarios. This application is expected to help users improve their communication skills, practice answering interview questions, and gain a more effective and efficient interview simulation experience before facing actual job interviews.*

**Kata kunci:** *Web Application*, *Artificial Intelligence*, *Real-Time Voice Conversation*, *Job Interview Simulation*, *WebRTC*, *NLP*, *TTS*

---

## I. PENDAHULUAN

Kemajuan teknologi merupakan hal yang tak terhindarkan dalam perkembangan manusia. Asisten virtual atau kecerdasan buatan berkembang pesat, dengan berbagai perusahaan teknologi besar berlomba menghadirkan layanan terbaik, seperti Google Assistant, Siri milik Apple, dan Alexa dari Amazon. Salah satu jenis asisten virtual yang banyak dibahas adalah voice assistant, yaitu asisten digital yang memanfaatkan pengenalan suara dan pemrosesan bahasa alami (*NLP*) untuk menjalankan perintah pengguna[1].

Salah satu contoh pengembangan *voicebot* adalah aplikasi desktop bernama *Chatterbot*. Aplikasi ini memanfaatkan *Natural Language Processing (NLP)* sebagai inti sistem untuk menentukan respons yang diberikan bot. Dengan menjadikan *NLP* sebagai “otak” atau *brainfile*, proses pengembangan menjadi lebih mudah karena respons-respons bot dapat dikelola secara terstruktur.

Beberapa penelitian sebelumnya telah dilakukan terkait dengan pengembangan sistem percakapan suara berbasis kecerdasan buatan, namun dengan fokus dan tujuan yang berbeda-beda. Penelitian oleh Ethiraj V, David A dan lainnya. (2025) menekankan bahwa pipeline interaksi suara yang terdiri atas *Automatic Speech Recognition (ASR)*, *Large Language Model (LLM)*, dan *Text-to-Speech (TTS)* masih memerlukan optimasi lebih lanjut agar benar-benar dapat mendukung komunikasi real-time. Selain itu, sebagian besar sistem yang ada masih bergantung pada layanan cloud, sehingga menimbulkan kendala berupa biaya tinggi, keterbatasan bandwidth, serta isu privasi pengguna[2]. Permasalahan lain juga muncul dalam penerapan nyata, seperti penurunan akurasi akibat kebisingan lingkungan, variasi aksen, maupun perbedaan kecepatan berbicara. Sementara itu, penelitian oleh Di Leo S, De Cicco L, Mascolo S (2025) menekankan pentingnya pemrosesan suara secara real-time di *edge computing* untuk menekan *delay* sekaligus menjaga privasi pengguna. Hasil penelitian tersebut menunjukkan bahwa implementasi sistem STT berbasis *edge* mampu mencapai latensi sub-detik, meskipun masih terdapat keterbatasan dalam hal skalabilitas dan integrasi penuh dengan pipeline percakapan[3].

Sehingga, diperlukan sebuah sistem yang tidak hanya mampu melakukan pengenalan dan pemrosesan suara, tetapi juga memberikan pengalaman percakapan yang responsif, natural, serta mendekati kondisi wawancara kerja yang sebenarnya.

Pengembangan aplikasi percakapan suara secara real-time berbasis kecerdasan buatan menjadi langkah penting untuk menjawab berbagai tantangan tersebut, khususnya dalam meningkatkan efektivitas latihan wawancara dan kesiapan pengguna menghadapi proses rekrutmen. Dengan memanfaatkan integrasi teknologi seperti *ASR*, *NLP*, *TTS*, dan *WebRTC*, sistem ini diharapkan mampu menghadirkan interaksi suara dua arah yang cepat, adaptif, dan kontekstual. Selain itu, aplikasi ini diharapkan dapat menjadi media pembelajaran modern yang membantu pengguna melatih kemampuan komunikasi, mengurangi kecemasan, serta meningkatkan kepercayaan diri sebelum menghadapi wawancara kerja yang sesungguhnya.

## II. TINJAUAN PUSTAKA

### A. Kecerdasan Buatan

Kecerdasan buatan (*AI*) adalah “payung istilah” yang digunakan untuk menyebut simulasi yang dilakukan oleh mesin-mesin atau alat, yang terhubung dengan samudera data, yang menyerupai kecerdasan manusia. Tidak diragukan lagi, *AI* sudah memberi dampak positif dalam banyak aspek kehidupan manusia yaitu ekonomi, pendidikan, pemerintahan, hingga pertahanan dan keamanan. Namun, *AI* bagaikan dua sisi mata uang yang juga memberikan dampak negatif. Adanya dampak multidimensi yang ditimbulkan oleh *AI* membawa pada suatu pertanyaan tentang cara mengimbangi kemajuan *AI* agar tetap terarah pada koridor yang diinginkan.

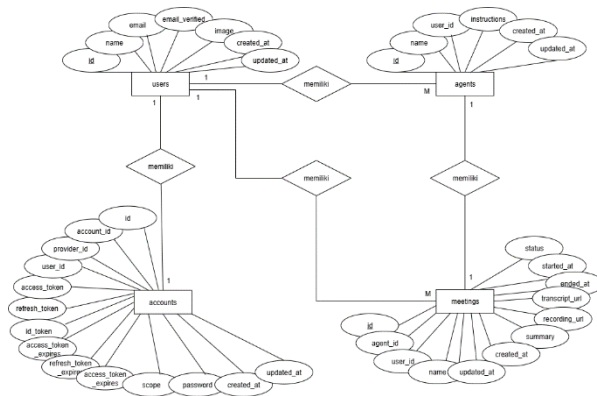
### B. *Natural Language Processing*

*Natural Language Processing (NLP)* merupakan salah satu cabang ilmu Artificial Intelligence (*AI*) yang berfokus pada pengolahan bahasa alami. Bahasa alami sendiri adalah bahasa yang secara umum digunakan oleh manusia dalam berkomunikasi satu sama lain. Bahasa yang diterima oleh komputer harus diproses dan dipahami terlebih dahulu agar maksud dari pengguna bisa dipahami dengan baik oleh computer.

### C. Percakapan Suara *Real-Time*

Percakapan suara *real-time* adalah proses komunikasi suara yang terjadi secara langsung (sinkron) antara dua pihak atau lebih tanpa jeda waktu yang signifikan. Dalam konteks teknologi informasi, percakapan suara real-time mengacu pada transmisi sinyal audio melalui jaringan komputer atau internet sehingga pengguna dapat berbicara dan mendengar

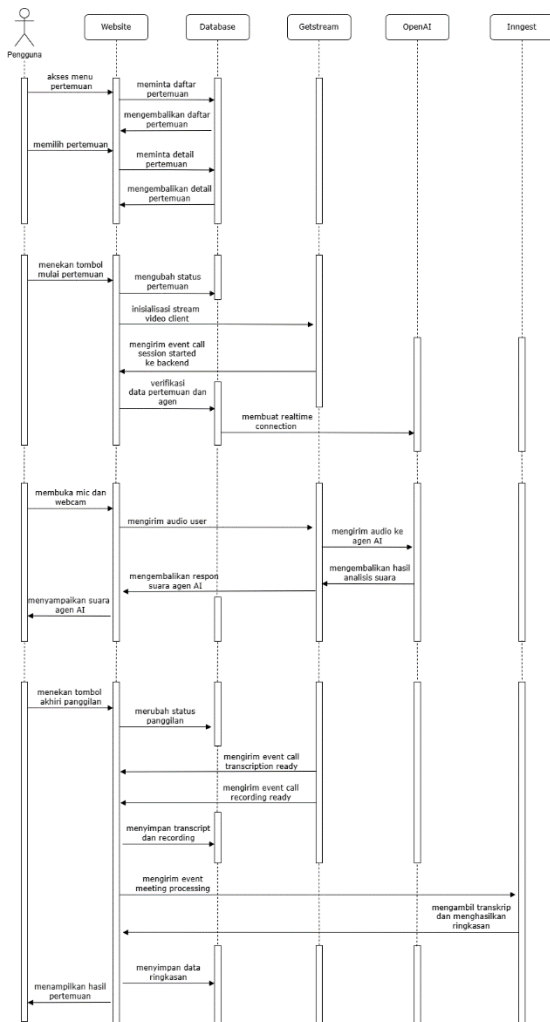




Gambar 3. Entity Relationship Diagram

#### 4. Sequence Diagram

Sequence diagram digunakan untuk menggambarkan alur interaksi antar komponen dalam sistem berdasarkan urutan waktu. Diagram ini menunjukkan bagaimana proses berlangsung dan bagaimana objek saling bertukar pesan ketika pengguna memulai pertemuan, berinteraksi dengan agen AI, hingga sistem menghasilkan ringkasan percakapan. Sequence Diagram sistem *Meet AI* dapat dilihat pada Gambar 4.



Gambar 4. Sequence Diagram

#### E. Implementasi Sistem

Sistem diimplementasikan menggunakan *framework* Next.js dengan bahasa pemrograman JavaScript dan TypeScript, serta PostgreSQL sebagai basis data utama. Proses autentikasi memanfaatkan layanan Better Auth, sementara interaksi suara dan pemrosesan bahasa alami menggunakan OpenAI Realtime API. Selain itu, sistem juga didukung Inngest sebagai layanan *background job* untuk memproses pembuatan ringkasan pertemuan. Seluruh komponen *frontend* dibangun dengan Tailwind CSS agar antarmuka lebih responsif dan konsisten.

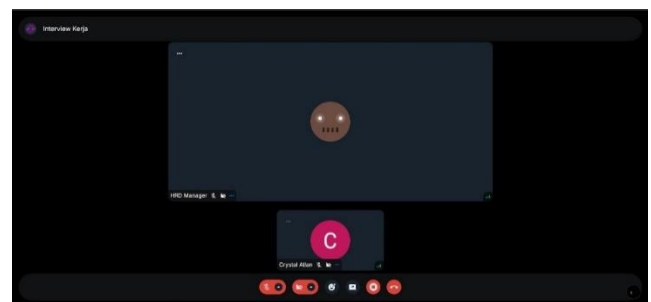
#### F. Pengujian Sistem

Pengujian sistem dilakukan menggunakan metode Black Box Testing dengan tujuan memvalidasi fungsi-fungsi utama pada aplikasi *Meet AI*, khususnya proses *voice interaction*, pemberian respons real-time oleh agen AI, serta kemampuan sistem dalam menghasilkan ringkasan percakapan secara otomatis. Setiap fitur diuji berdasarkan kesesuaian output yang dihasilkan dibandingkan dengan output yang diharapkan.

### IV. HASIL DAN PEMBAHASAN

#### G. Hasil Perancangan

Antarmuka sistem *Meet AI* dirancang menyerupai platform pertemuan modern yang menyediakan fitur percakapan suara interaktif dengan agen AI. Pengguna dapat memulai sesi pertemuan, berinteraksi melalui suara, serta menerima respons secara real-time. Setiap percakapan diproses oleh model AI untuk menghasilkan tanggapan kontekstual dan ringkasan otomatis. Tampilan utama dari agen AI dalam aplikasi ini dapat dilihat pada gambar 5.



Gambar 5. Hasil perancangan percakapan agen AI

#### H. Hasil Pengujian

Pengujian Black Box pada fitur Agen AI dilakukan untuk memastikan setiap fungsi dapat menerima input

pengguna, memprosesnya, dan menghasilkan respons sesuai yang diharapkan. Berdasarkan serangkaian pengujian, seluruh fungsi inti seperti pemanggilan agen, pemrosesan percakapan, dan pemberian respons berjalan dengan baik. Tingkat keberhasilan respons mencapai sekitar 94%, di mana ketidaksesuaian umumnya terjadi pada variasi pertanyaan atau frasa yang terlalu umum sehingga tidak terdeteksi oleh model. Tabel pengujian Black Box untuk menu Agen AI dapat dilihat pada Tabel berikutnya.

**Tabel 1. Blackbox Testing Menu Agen AI**

No	Input (User)	Expected Output	Hasil Pengujian
1	Klik tombol Buat Agen Baru	Menampilkan dialog tambah agen berisi input <i>Nama</i> dan <i>Instruksi</i>	Sesuai harapan
2	Isi semua field lalu klik tombol Buat	Data agen tersimpan dan muncul pada daftar agen	Sesuai harapan
3	Klik tombol Batal pada dialog tambah agen	Menutup dialog tanpa menyimpan data	Sesuai harapan
4	Klik salah satu agen pada daftar	Menampilkan halaman detail agen sesuai data	Sesuai harapan
5	Klik tombol ubah pada detail agen	Menampilkan dialog ubah agen dengan data sebelumnya	Sesuai harapan

Selanjutnya merupakan tabel pengujian *Blackbox* pada sistem terkait percakapan suara dengan agen AI.

**Tabel 2. Blackbox Testing Percakapan Suara Agen AI**

No	Input (User)	Expected Output	Hasil Pengujian
1	Klik tombol "Mulai Pertemuan" dari halaman Detail Pertemuan.	Navigasi ke halaman "Lobby Pertemuan" yang menampilkan pratinjau dan kontrol <i>microphone</i> serta <i>speaker</i> .	Sesuai harapan
2	Klik tombol "Gabung Panggilan".	Navigasi ke halaman "Ruang Pertemuan" dengan koneksi <i>microphone</i> aktif.	Sesuai harapan
3	Berbicara setelah berhasil masuk Ruang Pertemuan.	Input suara diterima; ikon <i>microphone</i> di avatar pengguna menampilkan indikator suara; agen AI merespons.	Sesuai harapan
4	Klik ikon End Call (tombol merah) di bilah kontrol.	Panggilan berakhir, dan sistem menavigasi ke halaman Detail Pertemuan (Ringkasan).	Sesuai harapan

5	Uji interaksi verbal dengan agen AI (berikan pertanyaan terkait topik pertemuan)	Agen AI merespons ucapan Pengguna dengan respons suara yang relevan dan kontekstual.	Sesuai harapan
---	--	--	----------------

[3] S. Di Leo, L. De Cicco, and S. Mascolo, "Real-Time Speech-to-Text on Edge: A Prototype System for Ultra-Low Latency Communication with AI-Powered NLP," *Information (Switzerland)*, vol. 16, no. 8, Aug. 2025, doi: 10.3390/info16080685.

## V. KESIMPULAN

Penelitian ini menghasilkan aplikasi *Meet AI*, yaitu platform percakapan suara berbasis teknologi *Speech-to-Text*, *Natural Language Processing*, dan integrasi *OpenAI Realtime API*. Sistem mampu menyediakan layanan simulasi pertemuan interaktif, menghasilkan ringkasan otomatis, serta menyimpan riwayat percakapan dalam basis data untuk kebutuhan evaluasi pengguna. Hasil pengujian menunjukkan bahwa fitur utama, seperti agen AI, pengelolaan pertemuan, dan proses pembuatan ringkasan, berjalan dengan baik dan memberikan pengalaman penggunaan yang stabil.

Ke depan, sistem dapat dikembangkan lebih lanjut dengan peningkatan kemampuan analisis percakapan, optimasi model bahasa agar lebih adaptif terhadap konteks pengguna, serta integrasi dengan platform konferensi atau aplikasi pesan instan untuk memperluas cakupan penggunaan.

## REFERENSI

[1] K. Kusmayadi, D. Kurniadi, and R. Setiawan, "Aplikasi Voice Assistant Pada Smartwach Menggunakan Open Artificial Intelligence," *Jurnal Algoritma*, vol. 21, no. 2, pp. 309–320, Nov. 2024, doi: 10.33364/algoritma/v.21-2.1499.

[2] V. Ethiraj, A. David, S. Menon, and D. Vijay, "Toward Low-Latency End-to-End Voice Agents for Telecommunications Using Streaming ASR, Quantized LLMs, and Real-Time TTS," Aug. 2025, [Online]. Available: <http://arxiv.org/abs/2508.04721>