

**PENGEMBANGAN AI *VTUBER* BERBASIS GPT-4 UNTUK INTERAKSI  
OBROLAN RINGAN PADA *LIVE STREAMING* YOUTUBE: STUDI  
KASUS CHANNEL KAIRA**

**(Skripsi)**

**Oleh**

**MUHAMMAD FADHILAH RAMADHANI**

**NPM 2157051006**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2025**

**PENGEMBANGAN AI *VTUBER* BERBASIS GPT-4 UNTUK INTERAKSI  
OBROLAN RINGAN PADA *LIVE STREAMING* YOUTUBE: STUDI  
KASUS CHANNEL KAIRA**

**Oleh**

**MUHAMMAD FADHILAH RAMADHANI  
NPM 2157051006**

**Skripsi**

**Sebagai Salah Satu Syarat untuk Mencapai Gelar  
SARJANA KOMPUTER**

**Pada**

**Jurusan Ilmu Komputer  
Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2025**

## ABSTRAK

### **PENGEMBANGAN AI *VTUBER* BERBASIS GPT-4 UNTUK INTERAKSI OBROLAN RINGAN PADA *LIVE STREAMING* YOUTUBE: STUDI KASUS CHANNEL KAIRA**

Oleh

**MUHAMMAD FADHILAH RAMADHANI**

Penelitian ini membahas pengembangan AI *VTuber* berbasis GPT-4 untuk menghasilkan interaksi obrolan ringan secara *real-time* pada *live streaming* YouTube, dengan studi kasus pada karakter “KAIRA” (Knowledge Artificial Intelligence Revolutionary Assistant). Latar belakang penelitian ini adalah keterbatasan *VTuber* konvensional yang dikendalikan manusia, seperti beban kerja tinggi, keterbatasan waktu, serta potensi kelelahan fisik dan mental. Sistem yang dikembangkan mengintegrasikan API GPT-4 untuk menghasilkan respons teks kontekstual, ElevenLabs untuk sintesis suara, dan VTube Studio untuk sinkronisasi animasi avatar. Pengujian dilakukan melalui dua tahap: evaluasi awal dengan 6 partisipan dan uji coba lapangan utama dengan 16 partisipan. Evaluasi subjektif menggunakan skala Likert dan evaluasi objektif menggunakan *cosine similarity* (IndoBERT dan IndoRoBERTa) untuk mengukur kesesuaian semantik antara respons AI dan ekspektasi pengguna. Hasil penelitian menunjukkan bahwa KAIRA mampu memberikan respons yang dinilai relevan dan natural, dengan nilai rata-rata penilaian Likert kategori “setuju” hingga “sangat setuju” dan skor *cosine similarity* yang konsisten di atas 0,6. Evaluasi penyaring topik game juga menunjukkan akurasi klasifikasi 50,33%, yang mengindikasikan perlunya peningkatan pada model deteksi topik. Penelitian ini menyimpulkan bahwa AI *VTuber* berbasis GPT-4 berpotensi menjadi alternatif interaksi virtual yang konsisten, efisien, dan menarik, meskipun masih terdapat ruang untuk optimasi dalam pengelolaan konteks dan akurasi penyaring topik.

Kata Kunci: AI *VTuber*, GPT-4, *Cosine Similarity*, YouTube *Live Streaming*, ElevenLabs

## **ABSTRACT**

### **DEVELOPMENT OF GPT-4-BASED AI VTUBER FOR CASUAL CONVERSATION INTERACTION IN YOUTUBE LIVE STREAMING: A CASE STUDY OF KAIRA CHANNEL**

**By**

**MUHAMMAD FADHILAH RAMADHANI**

This study presents the development of a GPT-4-based AI VTuber designed to deliver real-time casual chat interactions on YouTube Live, using the case study of “KAIRA” (Knowledge Artificial Intelligence Revolutionary Assistant). The research is motivated by the limitations of conventional human-controlled VTubers, including high workload, time constraints, and potential physical and mental fatigue. The system integrates the GPT-4 API to generate contextual text responses, ElevenLabs for speech synthesis, and VTube Studio for avatar animation synchronization. Testing was conducted in two stages: an initial evaluation with six participants and a main field trial with 16 participants. Subjective evaluation used a Likert scale, while objective evaluation employed cosine similarity (IndoBERT and IndoRoBERTa) to measure semantic alignment between AI responses and user expectations. Results indicate that KAIRA produced responses perceived as relevant and natural, with average Likert ratings in the “agree” to “strongly agree” range and cosine similarity scores consistently above 0.6. The game-topic filter evaluation yielded a classification accuracy of 50,33%, suggesting the need for improvements in topic detection. This study concludes that GPT-4-based AI VTubers have the potential to provide consistent, efficient, and engaging virtual interactions, although further optimization is needed for context management and topic filter accuracy.

**Keywords:** *AI VTuber, GPT-4, Cosine Similarity, YouTube Live Streaming, ElevenLabs*

Judul Skripsi

: **PENGEMBANGAN AI VTUBER  
BERBASIS GPT-4 UNTUK  
INTERAKSI OBROLAN RINGAN  
PADA *LIVE STREAMING*  
YOUTUBE: STUDI KASUS  
CHANNEL KAIRA**

Nama Mahasiswa

: **Muhammad Fadhilah Ramadhani**

Nomor Pokok Mahasiswa

: 2157051006

Program Studi

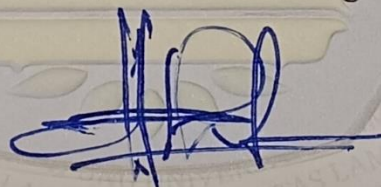
: Ilmu Komputer

Fakultas

: Matematika dan Ilmu Pengetahuan Alam

**MENYETUJUI**

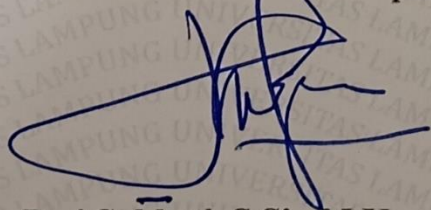
1. Komisi Pembimbing



**Dr. rer. nat. Akmal Junaidi, M.Sc.**  
NIP. 19710129 199702 1 001

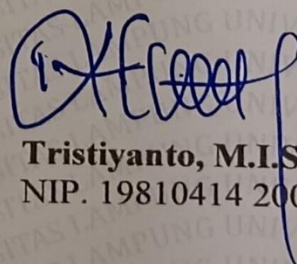
**MENGETAHUI**

Ketua Jurusan Ilmu Komputer



**Dwi Sakheti, S.Si., M.Kom.**  
NIP. 19680611 199802 1 001

Ketua Program Studi  
S1 Ilmu Komputer



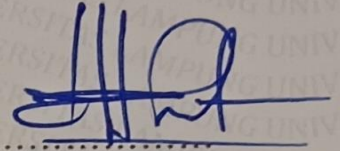
**Tristiyananto, M.I.S., Ph.D.**  
NIP. 19810414 200501 1 0001



## MENGESAHKAN

### 1. Tim Penguji

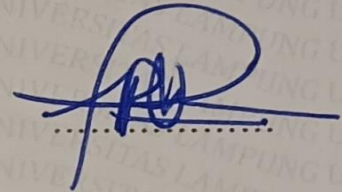
Ketua : **Dr. rer. nat. Akmal Junaidi, M.Sc.**



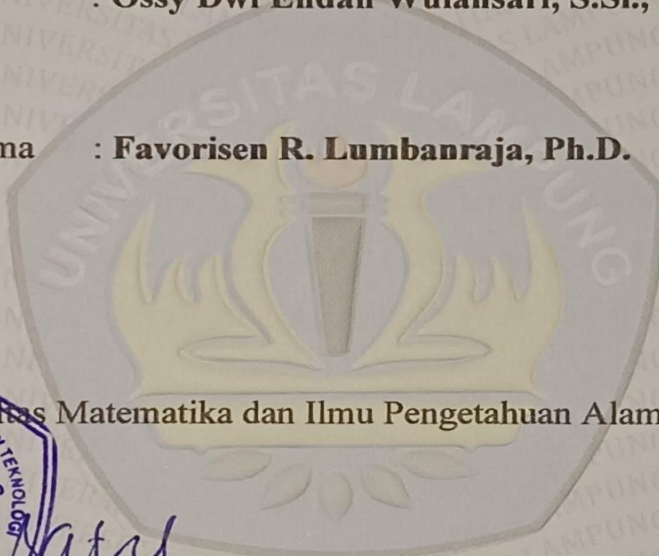
Sekretaris : **Ossy Dwi Endah Wulansari, S.Si., M.T.**



Penguji Utama : **Favorisen R. Lumbanraja, Ph.D.**



**Dr. Eng. Heri Satria, S.Si., M.Si.**  
NIP. 19711001 200501 1 002



Fakultas Matematika dan Ilmu Pengetahuan Alam

Tanggal Lulus Ujian Skripsi: **21 Juli 2025**

## PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Fadhilah Ramadhani

NPM : 2157051006

Dengan ini menyatakan bahwa skripsi saya yang berjudul “**Pengembangan AI Vtuber Berbasis GPT-4 Untuk Interaksi Obrolan Ringan Pada *Live Streaming Youtube: Studi Kasus Channel KAIRA***” adalah benar hasil karya tulis saya sendiri dan bukan karya orang lain. Seluruh tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya tulis ilmiah Universitas Lampung. Jika di kemudian hari terbukti bahwa skripsi saya adalah hasil penjiplakan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan peraturan yang berlaku.

Bandar Lampung, 14 Agustus 2025

Penulis,



Muhammad Fadhilah Ramadhani

NPM. 2157051006

## RIWAYAT HIDUP



Penulis dilahirkan di Gadingrejo pada 8 November 2003 sebagai anak pertama dari dua bersaudara dari pasangan Bapak Sugiyanto dan Ibu Eni Kartikasasi. Penulis menyelesaikan pendidikan Sekolah Dasar (SD) di SD N 8 Gadingrejo pada tahun 2015. Kemudian melanjutkan pendidikan Sekolah Menengah Pertama (SMP) di SMP N 1 Gadingrejo yang diselesaikan pada tahun 2018. Penulis kemudian melanjutkan lagi pendidikan Sekolah Menengah Atas (SMA) di SMA N 1 Gadingrejo yang diselesaikan pada tahun 2021. Penulis terdaftar sebagai mahasiswa S1 Ilmu Komputer, Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung pada tahun 2021 melalui jalur mandiri.

Selama menjadi mahasiswa, penulis melakukan beberapa kegiatan antara lain:

1. Menjadi Asisten Dosen Jurusan Ilmu Komputer untuk mata kuliah Pemrograman Berorientasi Objek (PBO) pada periode semester ganjil tahun ajaran 2023/2024.
2. Menjadi Asisten Dosen Jurusan Ilmu Komputer untuk mata kuliah *Internet of Things* (IoT) pada periode semester genap tahun ajaran 2023/2024.
3. Mengikuti program Gemastik XVI pada tahun 2023 pada divisi Piranti Cerdas dan Sistem Tanam.
4. Mengikuti Riset Independen dengan tema “Penerapan *Convolutional Neural Networks* untuk Pengenalan Wajah dalam Sistem Kehadiran Mahasiswa” yang bermitra dengan Jurusan Ilmu Komputer FMIPA, Unila pada tahun 2023.
5. Melaksanakan Kuliah Kerja Nyata (KKN) Periode II Universitas Lampung di Desa Karang Anyar pada tahun 2024.



## **PERSEMBAHAN**

Segala puji saya panjatkan kepada Allah Subhanahu wa Ta'ala segala limpahan rahmat dan karunia-Nya. Shalawat serta salam semoga senantiasa tercurah kepada Nabi Muhammad Shallallahu Alaihi wa sallam, sehingga penulis dapat menyelesaikan skripsi ini.

### **Skripsi ini saya persembahkan kepada:**

Untuk kedua orang tua saya tercinta, yang setiap tetes doa dan kasih sayangnya menjadi cahaya penuntun di setiap langkah. Untuk seluruh kawan seperjuangan di Keluarga Besar Ilmu Komputer 2021 Universitas Lampung, yang bersama menempuh jalan penuh tawa, lelah, dan harapan hingga ke ujung perjalanan ini. Dan untuk diri saya sendiri, Muhammad Fadhilah Ramadhani, yang tetap berdiri teguh, menjaga api semangat agar tak pernah padam, hingga karya ini lahir sebagai bukti bahwa mimpi dapat dijangkau dengan ketekunan.

## MOTTO

*“... Janganlah kamu berputus asa dari rahmat Allah. Sesungguhnya tidak ada yang berputus asa dari rahmat Allah melainkan kaum yang kafir.”*  
(QS. Yusuf: 87)

*“Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan.”*  
(QS. Al-Insyirah: 5 – 6)

*“All stories eventually come to an end, but the memories and emotions they leave behind will continue to extend, no?”*  
(Elysia, Honkai Impact 3<sup>rd</sup>; HoYoVerse)

*“Aku tidak perlu menciptakan ulang dunia dari nol untuk menghasilkan sesuatu yang berarti. Cukup membuat sesuatu yang berguna, bisa dipertanggungjawabkan, dan membawa pengalaman baru — dan aku sudah melakukannya.”*

## SANWACANA

Puji syukur kehadirat Allah Subhanahu wa Ta'ala, yang telah memberikan rahmat dan hidayah-Nya kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini dengan tepat waktu. Shalawat serta salam penulis sanjungkan kepada Nabi Muhammad Shallallahu Alaihi wa sallam yang penulis harapkan syafaatnya di hari akhir kelak.

Skripsi yang berjudul **“Pengembangan AI Vtuber Berbasis GPT-4 Untuk Interaksi Obrolan Ringan Pada *Live Streaming* Youtube: Studi Kasus Channel KAIRA”** ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana pada Program Studi S1 Ilmu Komputer di Universitas Lampung.

Selama proses penulisan skripsi ini tidak terlepas dari dukungan banyak pihak yang telah membimbing, membantu dan memberi semangat kepada penulis, sehingga pada kesempatan ini penulis ingin menyampaikan ungkapan terima kasih kepada:

1. Kedua orangtua penulis, Bapak Sugiyanto dan Ibu Eni Kartikasasi yang selalu mendoakan, menyemangati, membiayai, serta mendukung baik secara moral maupun material.
2. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
3. Bapak Dwi Sakethi, S.Si., M.Kom. selaku Ketua Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
4. Ibu Yunda Heningtyas, M.Kom. selaku Sekeretaris Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Tristiyato, M.I.S., Ph.D. selaku Ketua Program Studi S1 Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

6. Bapak M. Iqbal Parabi, S.Si., M.T. selaku dosen Pembimbing Akademik selama menjadi mahasiswa di Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
7. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc, selaku dosen Pembimbing Utama yang telah membimbing sehingga penulis dapat menyelesaikan skripsi dengan baik.
8. Ibu Ossy Dwi Endah Wulansari, S.Si., M.T, selaku dosen Pembimbing Kedua yang sudah memberikan masukan, kritik, dan arahan dalam penyelesaian skripsi.
9. Bapak Favorisen Rosyking Lumbanraja, Ph.D. selaku dosen Pembahas yang sudah memberikan kritik dan saran.
10. Bapak dan Ibu Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah memberikan ilmu, pengetahuan serta pengalamannya selama penulis menjadi mahasiswa.
11. Ibu Ade Nora Maela selaku staf administrasi Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah banyak membantu proses administrasi dalam proses penyelesaian skripsi.
12. Seluruh staf Jurusan Ilmu Komputer FMIPA Universitas Lampung yang selalu membantu penulis mulai dari awal hingga akhir masa perkuliahan.
13. Teman-teman Ilmu Komputer Angkatan 2021 yang menjadi keluarga satu angkatan selama menjalankan masa studi di Jurusan Ilmu Komputer Universitas Lampung.
14. Untuk penulis sendiri, Muhammad Fadhilah Ramadhani yang sudah bertahan dan berjuang dalam semua proses menuntut ilmu, sehingga berhasil menyelesaikan pendidikan di Program Studi S1 Ilmu Komputer Universitas Lampung dengan tepat waktu.
15. Semua pihak yang turut membantu penulis dalam proses penyusunan penelitian skripsi ini.

Dalam proses penyusunan skripsi ini tentunya masih jauh dari kata sempurna dikarenakan terdapat kekurangan yang disebabkan oleh keterbatasan pengalaman dan pengetahuan penulis. Penulis sangat mengharapkan kritik dan saran dari para

pembaca sebagai pembelajaran untuk penulis. Semoga isi dari skripsi ini dapat bermanfaat bagi pihak yang membaca.

Bandar Lampung, 14 Agustus 2025

Muhammad Fadhilah Ramadhani  
NPM. 2157051006



## DAFTAR ISI

	Halaman
DAFTAR TABEL .....	vii
DAFTAR GAMBAR .....	viii
DAFTAR PSEUDOCODE .....	viii
I. PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah .....	4
1.4 Tujuan .....	5
1.5 Manfaat .....	5
II. TINJAUAN PUSTAKA .....	7
2.1 Penelitian Terdahulu .....	7
2.2 VTuber .....	9
2.3 ChatGPT .....	9
2.4 Text-to-Speech (TTS) .....	10
2.5 ElevenLabs .....	10
2.6 Use Case Diagram .....	11
2.7 Activity Diagram .....	12
2.8 Skala Likert .....	14
2.9 Cosine Similarity .....	15
2.10 IndoBERT .....	16
2.11 IndoRoBERTa .....	17
2.12 Research and Development (R&D) .....	17

<b>III.</b>	<b>METODE PENELITIAN .....</b>	<b>20</b>
3.1	Waktu dan Tempat .....	20
3.1.1	Waktu Penelitian.....	20
3.1.2	Tempat Penelitian .....	20
3.2	Alat dan Bahan Penelitian .....	20
3.2.1	Perangkat Keras .....	20
3.2.2	Perangkat Lunak .....	21
3.3	Metode Pengumpulan Data .....	21
3.3.1	Sampel dan Responden.....	21
3.3.2	Instrumen Kuesioner.....	22
3.3.3	Prosedur Pengumpulan Data.....	22
3.3.4	Teknik Analisis Data .....	22
3.4	Tahapan Penelitian .....	23
3.5	Perancangan Sistem.....	25
3.5.1	Use Case Diagram .....	25
3.5.2	Activity Diagram .....	26
3.5.1	Model Interaksi.....	28
3.6	Jadwal Penelitian.....	29
<b>IV.</b>	<b>HASIL DAN PEMBAHASAN .....</b>	<b>31</b>
4.1	Pengumpulan Informasi .....	31
4.2	Perencanaan.....	32
4.3	Implementasi .....	33
4.4	Evaluasi Awal .....	39
4.5	Revisi Prototipe .....	48
4.6	Evaluasi Lapangan Utama.....	49
4.7	Revisi Operasional .....	53
<b>V.</b>	<b>SIMPULAN DAN SARAN .....</b>	<b>57</b>
5.1	Simpulan.....	57
5.2	Saran.....	58
	<b>DAFTAR PUSTAKA .....</b>	<b>60</b>

## DAFTAR TABEL

Tabel	Halaman
1. Penelitian terdahulu.....	7
2. Komponen Use Case Diagram.....	12
3. Komponen Activity Diagram.....	14
4. Jadwal Penelitian.....	30
5. Contoh Tema, Pertanyaan, dan Ekspektasi Jawaban Partisipan .....	<b>Error! Bookmark not defined.</b>
6. Contoh Respon KAIRA .....	<b>Error! Bookmark not defined.</b>
7. Rata-rata dan Standar Deviasi Penilaian Tema Obrolan..	<b>Error! Bookmark not defined.</b>
8. Hasil Cosine Similarity Model IndoBERT .....	<b>Error! Bookmark not defined.</b>
9. Hasil Cosine Similarity Model IndoRoBERTa..	<b>Error! Bookmark not defined.</b>
10. Nilai Tertinggi dan Terendah Dari Masing-Masing Model ..	<b>Error! Bookmark not defined.</b>
11. Rata-rata dan Standar Deviasi Penilaian Uji Coba Publik....	<b>Error! Bookmark not defined.</b>
12. Klasifikasi dan Perilaku Respon KAIRA.....	<b>Error! Bookmark not defined.</b>

## DAFTAR GAMBAR

Gambar	Halaman
1. Alur Metodologi R&D Menurut Borg dan Gall.....	18
2. Tahapan Penelitian.....	25
3. <i>Use Case Diagram</i> AI VTuber KAIRA. ....	26
4. Activity Diagram Alur Chat.....	27
5. Activity Diagram Alur Input Gambar.....	28
6. Model Interaksi AI VTuber.....	29
7. Avatar Karakter yang Digunakan KAIRA. ....	<b>Error! Bookmark not defined.</b>
8. Diagram Rata-rata dan Standar Deviasi Penilaian Tema Obrolan.....	<b>Error! Bookmark not defined.</b>
9. Perbandingan Cosine Similarity IndoBERT dan IndoRoBERTa. ....	<b>Error! Bookmark not defined.</b>
10. Diagram Rata-rata dan Standar Deviasi Penilaian Hasil Penilaian Uji Coba Publik. ....	<b>Error! Bookmark not defined.</b>
11. Distribusi Waktu Respon KAIRA terhadap Setiap Interaksi	<b>Error! Bookmark not defined.</b>
12. Confusion Matrix Klasifikasi Tema.....	<b>Error! Bookmark not defined.</b>

## DAFTAR PSEUDOCODE

Pseudocode	Halaman
1. Fungsi Untuk Menyusun Prompt. ....	<b>Error! Bookmark not defined.</b>
2. Fungsi Untuk Mengirim Prompt ke GPT.....	<b>Error! Bookmark not defined.</b>
3. Fungsi FFMPEG Untuk Perbaiki Audio. ....	<b>Error! Bookmark not defined.</b>
4. Fungsi ElevenLabs Untuk Mengubah Respon GPT Menjadi Audio.....	<b>Error! Bookmark not defined.</b>
5. Potongan Kode Pemutar Audio.....	<b>Error! Bookmark not defined.</b>



## I. PENDAHULUAN

### 1.1 Latar Belakang

Perkembangan teknologi kecerdasan buatan atau *Artificial Intelligence* (AI) telah mengalami perkembangan pesat dan telah membuka peluang inovasi di berbagai bidang, termasuk industri hiburan digital. Salah satu fenomena penerapan teknologi AI yang menarik perhatian adalah kemunculan *Virtual YouTuber* (*VTuber*), karakter virtual yang melakukan *live streaming* untuk berinteraksi dengan penonton. *VTuber* sendiri pertama kali populer di Jepang dengan kemunculan *VTuber* pertama, Kizuna Ai, pada tahun 2016 dan popularitasnya semakin menyebar secara global (Puspitaningrum dan Prasetyo, 2019). *VTuber* umumnya dikendalikan oleh manusia menggunakan teknologi *motion capture*, tetapi ketergantungan pada *talent* manusia memiliki beberapa keterbatasan, seperti jadwal streaming yang tidak fleksibel, biaya produksi tinggi, dan risiko lainnya yang berkaitan dengan *talent*.

*VTuber* dengan *talent* manusia menghadapi beban kerja tinggi dalam produksi konten. Mereka harus mengintegrasikan banyak perangkat (misal *motion capture*, kamera, *software*) untuk mengendalikan avatar, sambil *multitasking* berinteraksi dengan penonton dan menjaga anonimitas. Penelitian terbaru menyebutkan bahwa sifat fragmentasi alat dan *multi-tasking* dalam *VTubing* menimbulkan beban kognitif dan fisik signifikan bagi *talent*, melebihi streamer biasa (Kim dkk., 2025). Hal ini dapat menyebabkan kelelahan dalam jangka panjang dan menuntut adanya penanganan khusus agar tidak mengganggu performa.

Menjalankan persona virtual yang berbeda dari identitas asli juga menimbulkan tantangan psikologis. Talent *VTuber* sering dituntut memerankan karakter fiksi yang mungkin jauh dari kepribadian aslinya, demi konsistensi persona dan hiburan penonton. Studi HCI melaporkan bahwa jika jarak antara kepribadian asli dengan persona *VTuber* terlalu besar, hal itu membebani mental talent dan dapat berujung pada *burnout* atau kelelahan emosional (Kim dkk., 2025).

Dari sisi keterbatasan fleksibilitas fisik, teknologi *VTuber* saat ini juga mengurangi kebebasan gerak talent dan menambah beban. Misalnya, kamera penangkap gerak mengharuskan *VTuber* tetap dalam area tertentu agar *tracking* avatar tidak hilang, dan perangkat *motion-tracker* harus dipasang ketat di tubuh demi akurasi (Kim dkk., 2025). Talent sering perlu melebih-lebihkan gerakan dan ekspresi agar avatar tampak hidup, yang tentu melelahkan secara fisik. Selain itu, penggunaan *headset* VR (HMD) yang berat dan alat murah berkualitas rendah dapat memicu ketidaknyamanan seperti pegal dan mual (*motion sickness*) selama *streaming* (Kim dkk., 2025).

Di sisi lain, kemajuan *Generative AI* seperti *Generative Pre-trained Transformer* (GPT) dari OpenAI menawarkan solusi potensial. GPT, khususnya untuk versi terbaru (GPT-4), mampu menghasilkan teks mirip manusia dengan pemahaman konteks yang baik (Radford dkk., 2023). GPT memiliki potensi menggunakan integrasi API (*Application Programming Interface*) pada pengembangan AI *VTuber* yang dapat menciptakan pengalaman interaksi yang lebih responsif dan realistis bagi penonton (Brown dkk., 2020).

Efisiensi dan skalabilitas menjadi alasan utama mempertimbangkan *VTuber* berbasis AI. Seorang manusia memiliki keterbatasan waktu, tenaga, dan fokus dalam melayani interaksi yang sangat tinggi volumenya. Riset menunjukkan bahwa manajemen interaksi real-time berjumlah besar (misal ribuan pesan *chat*) merupakan tantangan besar bagi streamer manusia (Amato dkk., 2024). Di sinilah AI bisa berperan, contoh kasusnya adalah

Neuro-sama, *VTuber* AI pertama yang menggabungkan *chatbot* dengan avatar virtual, di mana kepribadian dan ucapannya sepenuhnya dihasilkan AI berbasis LLM sehingga mampu berkomunikasi langsung dengan penonton di *live chat* (Amato dkk., 2024). *VTuber* AI semacam ini tidak mengalami lelah dan dapat aktif hampir 24/7, sehingga konten bisa lebih konsisten hadir dan *audiens* terlayani kapan saja.

AI juga dapat melengkapi peran *talent* manusia demi efisiensi kerja. Alih-alih menggantikan sepenuhnya, AI bisa digunakan untuk mengotomatiskan tugas-tugas teknis atau repetitif, sehingga *talent* manusia dapat fokus pada kreativitas. Sebagai contoh, Tang dkk. (2021) mengembangkan sistem AlterEcho yang secara otomatis memicu gerakan dan mimik avatar berdasarkan ekspresi wajah dan suara sang *VTuber* (Tang dkk., 2021). Dengan bantuan AI ini, *VTuber* dapat mempertahankan persona virtual secara konsisten (karakter avatar tetap bergerak/ekspresi tepat waktu) sambil mengurangi beban kognitif talent dalam mengontrol avatar secara manual (Tang dkk., 2021). Artinya, AI memungkinkan performa yang lebih efisien sehingga *talent* tidak perlu *multitasking* ekstrem untuk animasi, sehingga kelelahan berkurang dan interaksi dengan penonton tetap terjaga kualitasnya.

Selain mengurangi beban kerja *talent*, penggunaan AI juga membuka peluang baru untuk menciptakan bentuk interaksi yang lebih menarik dan responsif secara real-time. Salah satu aspek penting dalam dunia *VTubing* adalah keterlibatan penonton dalam *live chat*, yang jika ditangani secara efektif dapat meningkatkan kualitas pengalaman menonton secara keseluruhan. Sebuah eksperimen yang pernah dilakukan pada platform Twitch menunjukkan bahwa keterlibatan aktif penonton dalam *chat* dapat menambah nilai pada pengalaman menonton secara keseluruhan (Hamilton dkk., 2014). Penggunaan AI *VTuber* berbasis API GPT tidak hanya dapat meningkatkan kualitas interaksi, tetapi juga memberikan solusi bagi kreator konten dalam manajemen interaksi dengan penonton secara *real-time*. Dengan demikian, penelitian ini dapat memberikan kontribusi dalam

pengembangan teknologi *VTuber* serta memperkaya konten digital yang interaktif dan inovatif.

## 1.2 Rumusan Masalah

Dari latar belakang yang telah diuraikan, rumusan masalah dalam penelitian ini adalah:

- A. Bagaimana merancang sistem AI *VTuber* berbasis GPT yang terintegrasi dengan *platform* YouTube untuk menghasilkan interaksi obrolan ringan (*casual chat*) secara *real-time*, termasuk sinkronisasi respon teks, suara sintetis, dan animasi karakter?
- B. Sejauh mana AI *VTuber* berbasis GPT mampu meniru percakapan manusia dalam konteks obrolan santai (seperti hobi, kesukaan, atau kehidupan sehari-hari) berdasarkan penilaian pengguna terhadap kualitas interaksi (melalui skala Likert) dan kesamaan semantik (melalui pengukuran cosine similarity)?
- C. Apa saja hambatan teknis (seperti *latency*, manajemen konteks percakapan) dan etis (seperti risiko bias atau misinformasi) yang muncul dalam pengembangan AI *VTuber* ini, serta bagaimana strategi mengatasinya?

## 1.3 Batasan Masalah

Agar penelitian lebih terarah, beberapa batasan masalah berikut diterapkan:

- A. Teknologi:
  1. Hanya menggunakan API GPT dari OpenAI sebagai model utama untuk menghasilkan respon teks.
  2. Teknologi suara sintetis menggunakan ElevenLabs dan dengan menggunakan satu jenis suara karakter (Perempuan).
- B. *Platform*:
  1. Interaksi dilakukan eksklusif di YouTube *Live Chat*.
- C. Jenis Interaksi:
  1. Obrolan hanya pada topik ringan (*casual chat*) seperti kesukaan atau kehidupan sehari-hari.

D. Visualisasi Karakter:

1. Animasi karakter terbatas pada gerakan bibir (*lip-sync*) dan ekspresi wajah dasar (tersenyum). Tidak melibatkan *full-body motion* atau interaksi fisik.

E. Aspek non-Teknis:

1. Tidak membahas monetisasi konten, hak cipta karakter, atau dampak sosial jangka panjang dari AI VTuber.

## 1.4 Tujuan

Tujuan dari penelitian ini adalah:

- A. Mengembangkan AI *VTuber* berbasis GPT yang terintegrasi dengan *platform* YouTube untuk menghasilkan interaksi obrolan ringan secara *real-time*, termasuk sinkronisasi respon teks, suara sintetis, dan animasi karakter.
- B. Mengukur efektivitas AI *VTuber* berbasis GPT dalam meniru percakapan manusia pada konteks obrolan santai (seperti hobi, kesukaan, atau kehidupan sehari-hari) berdasarkan persepsi pengguna terhadap kualitas interaksi menggunakan skala Likert dan Tingkat kesamaan semantik antara ekspektasi pengguna dan respons AI menggunakan metode cosine similarity.
- C. Mengidentifikasi hambatan teknis dan etis dalam pengembangan sistem dan membuat strategi untuk mengatasinya.

## 1.5 Manfaat

Adapun manfaat dari penelitian ini adalah:

A. Bagi penonton:

1. Meningkatkan pengalaman interaksi selama *live streaming* melalui respon AI *VTuber* yang realistis dan relevan dengan topik obrolan santai (*casual chat*).
2. Menyediakan hiburan yang menarik dengan adanya karakter *VTuber* yang tidak menggunakan *talent* sebaliknya namun mampu berinteraksi kepada penonton.

B. Bagi kreator konten:



1. Membantu mengurangi ketergantungan pada *talent* manusia, serta menghemat biaya dan waktu produksi.

C. Bagi pengembang:

1. Menyediakan panduan teknis (seperti integrasi API GPT, sinkronisasi suara sintetis, dan sinkronisasi animasi) serta pertimbangan etis (manajemen bias, mitigasi misinformasi) untuk pengembangan AI *VTuber* berbasis GPT.

II. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

AI *VTuber* merupakan penerapan teknologi yang berada di ranah AI *assistant* atau *chatbot*. Penelitian terdahulu yang berkaitan dengan pengembangan dan pembuatan AI *assistant* atau *chatbot* dapat dilihat pada Tabel 1.

Tabel 1. Penelitian terdahulu

No	Penulis	Judul	Hasil	Evaluasi
1	(Rau dkk., 2024)	<i>Enhancing chatbot performance for imaging recommendations: Leveraging GPT-4 and context-awareness for trustworthy clinical guidance</i>	Chatbot accGPT-4 untuk rekomendasi pencitraan medis sesuai ACR <i>Appropriateness Criteria</i> menunjukkan akurasi 78% dan konsistensi 94%, efektif mengenali ketidakcocokan pencitraan diagnostik.	50 kasus klinis; uji akurasi dan konsistensi berdasarkan kategori “usually / may be appropriate”.

No	Penulis	Judul	Hasil	Evaluasi
2	(Van Poucke, 2024)	<i>ChatGPT, the perfect virtual teaching assistant? Ideological bias in learner-chatbot interactions</i>	Analisis linguistik menemukan bias ideologi dan ketidaksensitifan kultural dalam interaksi ChatGPT sebagai asisten virtual.	Analisis kualitatif korpus Reddit (15.909 kata) & ChatGPT (10.696 kata) menggunakan <i>Appraisal Theory</i> dan <i>Systemic Functional Linguistics</i> .
3	(Sathe dkk., 2024)	<i>How I GPT it: Development of Custom Artificial Intelligence (AI) Chatbots for Surgical Education</i>	Chatbot edukasi bedah berbasis GPT-4 mampu memberi jawaban medis akurat dan mendukung peningkatan interaktivitas pembelajaran.	Simulasi pembelajaran; evaluasi awal pada pertanyaan medis uji dan kemampuan chatbot menjawab dengan benar.
4	(Amato dkk., 2024)	<i>Can an AI-driven VTuber engage People? The KawAli Case Study</i>	VTuber KawAli berbasis LLM dinilai akurat, ekspresif, dan menyenangkan, meningkatkan keterlibatan dan kepercayaan pengguna.	32 partisipan; uji persepsi menggunakan penilaian akurasi, ekspresivitas, dan keterlibatan

No	Penulis	Judul	Hasil	Evaluasi
				selama live streaming.

## 2.2 VTuber

*VTuber* atau *Virtual YouTuber* adalah *streamer* yang menggunakan karakter visual yang dilengkapi dengan animasi dan teknologi *motion capture* secara *real-time* untuk berinteraksi dengan audiens melalui *platform* video, misalnya YouTube (Hermawan dkk., 2024). Istilah dari “*Virtual YouTuber*” sendiri diciptakan oleh *VTuber* Kizuna Ai pada akhir tahun 2016 yang dikenal secara luas di dunia sebagai *VTuber* pertama (Jamaluddin, 2021).

*VTuber* dibagi menjadi tiga jenis, yaitu bisnis, komunitas, dan individu. Salah satu perusahaan yang menjalankan bisnis *VTuber* adalah HololiveProduction yang dimiliki oleh perusahaan teknologi Jepang, COVER, yang pada tahun 2022, salah satu *talent* mereka yaitu Gawr Gura mendapat peringkat popularitas nomor satu dengan 4,3 juta *subscriber* dan 340 juta penayangan video (Lee, 2024).

## 2.3 ChatGPT

ChatGPT merupakan sebuah *language model* yang dikembangkan oleh OpenAI yang memiliki potensi revolusioner untuk menjembatani interaksi kita dengan teknologi (Aljanabi dkk., 2023). Dengan kemampuannya untuk menghasilkan teks seperti manusia dan menanggapi pertanyaan yang kompleks, ChatGPT telah memberikan dampak yang signifikan pada perkembangan teknologi terutama pada bidang integrasi dengan teknologi *computer vision* dan robotik (Aljanabi, 2023).

Saat ini, ChatGPT telah banyak diterapkan di berbagai bidang dan menunjukkan potensi yang menjanjikan. Hal ini telah menarik perhatian yang semakin besar dari kalangan industri dan akademisi (Hu, 2024). ChatGPT dapat membantu mendukung pembelajaran dengan berbagai cara, seperti menjelaskan konsep yang membingungkan di kelas, memberikan jawaban yang cepat dan akurat untuk pertanyaan, serta menyediakan umpan

balik yang dipersonalisasi dan saran perbaikan untuk meningkatkan pembelajaran (Al-Ghonmein dan Al-Moghrabi, 2024).

## 2.4 *Text-to-Speech* (TTS)

TTS adalah proses menghasilkan audio lisan berdasarkan teks yang dihasilkan secara otomatis menggunakan teknologi komputer (Yao dkk., 2024). Teknologi ini biasanya digunakan untuk membantu orang-orang disabilitas khususnya bagi mereka yang mengalami gangguan berbicara. Dalam penggunaan baru-baru ini, TTS banyak tertanam pada layanan robot dan humanoid untuk mempermudah berinteraksi dengan pemberi perintah (Ramli dkk., 2015).

Sistem TTS dikembangkan melalui kombinasi berbagai disiplin ilmu, seperti akustik, linguistik, pemrosesan sinyal, dan statistik. Tujuan utamanya adalah menghasilkan suara sintetis yang semirip mungkin dengan suara manusia asli dan mudah dipahami. Kecerdasan menunjukkan tingkat kejelasan suara yang dihasilkan, sedangkan kealamian menggambarkan seberapa nyaman suara tersebut untuk didengarkan (Oyucu, 2023).

## 2.5 ElevenLabs

ElevenLabs adalah *platform* AI yang menghasilkan suara manusia buatan menggunakan teknologi jaringan saraf dalam (*Deep Neural Network*). Alat ini dipakai untuk membuat narasi animasi dari teks skrip, memungkinkan revisi cepat dan suara yang terdengar alami. Meski berguna untuk efisiensi waktu, suara AI seperti ini masih kurang bisa menangkap emosi atau intonasi rumit seperti yang dilakukan pengisi suara manusia. Dengan memanfaatkan tools seperti ElevenLabs, kreator bisa lebih fokus mengembangkan visual, cerita, dan elemen penting lainnya dalam pembuatan konten (Chen, 2023).

Selain fitur-fitur tersebut, ElevenLabs juga mendukung teknologi suara dalam berbagai bahasa, termasuk Bahasa Indonesia, memungkinkan penggunaan satu jenis suara untuk beberapa bahasa sekaligus. Aplikasi ini menawarkan suara sintetis yang realistis dengan kualitas audio jernih, serta kemampuan kloning suara untuk meniru karakter vokal tertentu. Namun,

layanannya berbayar dengan batas 10.000 karakter pada versi gratis, dan masih memiliki keterbatasan dalam mengekspresikan variasi emosi atau intonasi secara mendalam. Meski demikian, teknologi ini tetap menjadi pilihan praktis untuk proyek yang memprioritaskan efisiensi, meski belum sepenuhnya menyamai fleksibilitas pengisi suara manusia (Dewatri dkk., 2023).


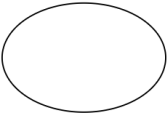


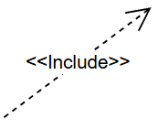
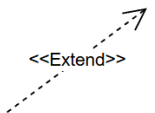
## 2.6 Use Case Diagram

*Use Case Diagram* merupakan salah satu diagram UML (*Unified Modeling Language*) yang digunakan pada tahap awal rekayasa perangkat lunak untuk menangkap kebutuhan fungsional suatu sistem (Alturas, 2023). Setiap *use case* merepresentasikan fitur atau fungsionalitas yang dibutuhkan *user*, dengan aktor sebagai entitas eksternal (manusia atau sistem lain) yang berinteraksi dengan sistem untuk mencapai suatu tujuan (Alturas, 2023). Diagram ini memberikan ringkasan visual tentang *use case*, para aktor, serta hubungan di antara mereka (Alturas, 2023). Notasi standar *Use Case Diagram* mencakup simbol aktor, *use case* (digambarkan dengan bentuk oval), dan asosiasi (garis) yang menghubungkan aktor dengan *use case* (Alturas, 2023).

Dalam pembuatannya, *Use Case Diagram* digambarkan dengan *boundary system* (kotak persegi yang mewakili batas sistem dan diberi nama sistem), simbol *use case* berupa oval yang diberi label sesuai fungsi atau layanan yang diberikan sistem, dan aktor digambarkan sebagai entitas luar yang berada di luar *boundary system* (umumnya dilambangkan figur stik) (Aquino dkk., 2020). Diagram ini juga menampilkan hubungan antar komponen. Garis asosiasi menghubungkan aktor ke *use case*, serta relasi khusus seperti “<<include>>” dan “<<extend>>” antar *use case* (Aquino dkk., 2020). Misalnya, relasi *extend* menunjukkan *use case* opsional yang memperluas *use case* utama, sedangkan *include* menunjukkan bahwa satu *use case* menyertakan fungsi *use case* lain sebagai bagian dari alur kerjanya (Aquino dkk., 2020). Ada pula generalisasi yang menunjukkan hierarki antar aktor atau *use case* dengan panah terbuka. Untuk lebih jelas mengenai

simbol-simbol yang ada dalam *Use Case Diagram* bisa dilihat pada Tabel 2.

**Tabel 2.** Komponen *Use Case Diagram*

No	Simbol	Nama	Keterangan
1	 Actor	Aktor	Aktor merepresentasikan pengguna atau entitas yang berinteraksi dengan sebuah sistem. Aktor bisa berupa manusia, perangkat eksternal, atau sistem lain yang terlibat dalam <i>use case</i> .
2		Use case	<i>Use case</i> menggambarkan fungsionalitas atau layanan yang disediakan oleh sistem kepada aktor. Biasanya berisi tindakan-tindakan atau skenario yang dilakukan oleh aktor dengan sistem.
3		Asosiasi	Asosiasi menggambarkan hubungan antara dengan <i>use case</i> .
4		Generelasisasi	Generalisasi menggambarkan hubungan hirarki antara aktor atau <i>use case</i> . Aktor atau <i>use case</i> yang lebih spesifik mewarisi atau memperluas fungsionalitas dari aktor atau <i>use case</i> yang lebih umum.
5		Include	<i>Include</i> menunjukkan bahwa suatu <i>use case</i> selalu menyertakan fungsionalitas <i>use case</i> lain sebagai bagian dari pelaksanaannya.
6		Extend	<i>Extend</i> menunjukkan bahwa suatu <i>use case</i> dapat memperpanjang fungsionalitas <i>use case</i> lain dalam kondisi tertentu.

## 2.7 Activity Diagram

*Activity Diagram* merupakan diagram yang menggambarkan alur kerja (*workflow*) atau rangkaian langkah-langkah operasional dalam sistem




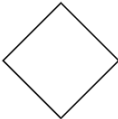

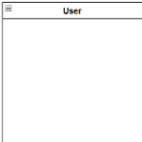
secara grafis. Secara definisi, *Activity Diagram* merepresentasikan proses bisnis atau operasional sistem secara langkah demi langkah, menunjukkan bagaimana aktivitas saling berurutan, termasuk cabang keputusan, iterasi, dan paralelisme dalam proses (Siewe dan Ngounou, 2025). Dengan begitu, *Activity Diagram* membantu *software engineer* memahami perilaku sistem atau skenario alur proses yang kompleks dengan lebih jelas (Siewe dan Ngounou, 2025).

*Activity Diagram* dianggap mampu merepresentasikan detail alur kontrol dan aksi di dalam sistem secara terstruktur. Keunggulan ini membuat *Activity Diagram* bermanfaat dalam memastikan bahwa model desain telah mencakup skenario keputusan dan proses paralel yang mungkin terjadi (Al-Fedaghi, 2021). Dengan kata lain, *Activity Diagram* memudahkan pengembang memverifikasi bahwa alur suatu kebutuhan atau *use case* tertentu telah digambarkan dengan benar dan lengkap dalam model, sebelum implementasi yang sebenarnya.

Dalam pembuatannya, terdapat simbol-simbol standar yang digunakan, di antaranya seperti *Start node* (bulatan hitam sebagai awal alur aktivitas), *End node* (bulatan hitam dengan lingkaran luar menandai akhir alur), *Activity/Action* (kotak atau bulatan lonjong yang merepresentasikan langkah proses), *Decision node* (simbol belah ketupat untuk percabangan kondisi), *Fork/Join node* (garis atau bar tebal horizontal/vertikal untuk memecah alur menjadi paralel dan menggabungkannya kembali), serta *Swimlanes* (kolom/baris pembagi area diagram untuk menunjukkan pemilik aktivitas) (Hendini, 2016). Untuk lebih jelas mengenai simbol-simbol yang ada dalam *Activity Diagram* bisa dilihat pada Tabel 3.



**Tabel 3.** Komponen *Activity Diagram*

No	Simbol	Nama	Keterangan
1		<i>Start</i>	Menunjukkan titik awal dari alur aktivitas dalam <i>Activity Diagram</i> .
2		<i>End</i>	Menunjukkan titik akhir dari alur aktivitas dalam <i>Activity Diagram</i> .
3		<i>Activity/Action</i>	Menyatakan langkah atau proses yang dilakukan dalam alur kerja sistem.
4		<i>Decision</i>	Menandakan percabangan kondisi untuk menentukan arah alur berdasarkan keputusan logis.
5		<i>Fork/Join</i>	<i>Fork</i> digunakan untuk memecah alur menjadi paralel; <i>Join</i> digunakan untuk menggabungkan kembali alur paralel.
6		<i>Swimlanes</i>	Digunakan untuk memisahkan tanggung jawab aktivitas berdasarkan aktor atau komponen sistem.

## 2.8 Skala Likert

Skala Likert merupakan salah satu instrumen psikometrik yang digunakan untuk mengukur persepsi, sikap, pendapat, atau preferensi seseorang terhadap suatu pernyataan yang diajukan dalam penelitian kuantitatif. Secara formal, skala Likert terdiri atas serangkaian pernyataan yang dinilai responden dengan memilih tingkat persetujuan atau ketidaksetujuan terhadap pernyataan tersebut (Joshi dkk., 2015). Jawaban tersebut

dikonversi menjadi data numerik, yang kemudian dianalisis secara statistik untuk menarik kesimpulan tentang sikap atau persepsi responden secara kolektif (Joshi dkk., 2015).

Konsep dasar skala Likert adalah menyediakan sejumlah pernyataan yang mewakili aspek-aspek tertentu dari sikap yang diukur, di mana responden menyatakan tingkat persetujuannya pada skala bertingkat. Skala yang paling umum digunakan adalah skala 5-poin, mulai dari 1 (Sangat Tidak Setuju) hingga 5 (Sangat Setuju), namun terdapat pula variasi skala 4 hingga 7 poin (Koo dan Yang, 2025). Skala Likert lima poin sering dipilih karena keseimbangannya antara kemudahan responden dalam menentukan pilihan serta sensitivitas data yang cukup baik (Koo dan Yang, 2025).

Skala Likert secara luas digunakan dalam berbagai bidang penelitian kuantitatif, mulai dari ilmu sosial hingga evaluasi sistem teknologi. Popularitas skala Likert didasarkan pada kemampuannya mengubah data kualitatif berupa persepsi atau sikap responden menjadi data numerik yang terstruktur (Schrum dkk., 2020). Secara umum, kemudahan desain, implementasi, serta analisis data yang didapatkan dari skala Likert menjadikannya pilihan populer dalam evaluasi. Dengan demikian, skala ini memberikan nilai tambah dalam proses penelitian yang membutuhkan pemahaman kuantitatif terhadap sikap atau persepsi responden (Schrum dkk., 2020).

## **2.9 Cosine Similarity**

*Cosine similarity* merupakan salah satu metode pengukuran kesamaan yang paling banyak digunakan dalam berbagai bidang seperti pengolahan bahasa alami (NLP), sistem rekomendasi, bahkan deteksi plagiarisme. Metode ini bekerja dengan menghitung nilai kosinus dari sudut antara dua vektor dalam ruang berdimensi- $n$ . Alih-alih memperhitungkan panjang vektor, *cosine similarity* hanya memperhatikan arah keduanya, sehingga sangat berguna ketika panjang vektor bukan merupakan faktor utama yang relevan (Hidayat, 2024).

Secara matematis, *cosine similarity* dihitung dengan membagi hasil perkalian titik (*dot product*) dua vektor dengan hasil kali dari panjang kedua vektor tersebut, dengan rumus sebagai berikut:

$$\text{sim}(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{||\mathbf{A}|| ||\mathbf{B}||}$$

dengan  $\mathbf{A} \cdot \mathbf{B}$  adalah perkalian titik, sedangkan  $||\mathbf{A}||$  dan  $||\mathbf{B}||$  adalah panjang vektor dari  $\mathbf{A}$  dan  $\mathbf{B}$ . *Cosine similarity* biasanya memiliki rentang nilai antara -1 sampai 1, dimana nilai 1 menunjukkan kesamaan arah vektor yang identik, nilai 0 menunjukkan bahwa arah vektor tidak memiliki kesamaan, dan nilai -1 menunjukkan bahwa arah vektor saling berlawanan.

## 2.10 IndoBERT

IndoBERT merupakan model bahasa pra-latih berbasis BERT yang secara khusus dikembangkan untuk bahasa Indonesia. Model ini hadir sebagai respons terhadap keterbatasan sumber daya NLP dalam bahasa Indonesia, seperti minimnya dataset beranotasi, sumber daya linguistik yang terbatas, dan kurangnya standar evaluasi yang menyeluruh. IndoBERT dilatih menggunakan skema *Masked Language Modeling* (MLM) dengan *framework* Huggingface dan *Next Sentence Prediction* (NSP), serta memanfaatkan korpus besar berbahasa Indonesia, termasuk Wikipedia Indonesia, berita daring dari Kompas, Tempo, dan Liputan6, serta korpus Web Indonesia (IdWaC) (Koto dkk., 2020).

Secara arsitektural, IndoBERT memiliki dua varian utama, yakni IndoBERT-*base* dan IndoBERT-*large*, yang masing-masing berbeda dari segi jumlah lapisan transformer dan parameter total. IndoBERT-*base* terdiri dari 12 *layer* transformer dan 110 juta parameter, sedangkan IndoBERT-*large* memiliki 24 *layer* dan 340 juta parameter. Kedua varian ini dirancang untuk dapat digunakan secara langsung atau melalui proses fine-tuning untuk berbagai tugas NLP (Geni dkk., 2023).

Keunggulan utama IndoBERT terletak pada kemampuannya menangkap makna semantik dalam bahasa Indonesia dengan akurasi yang lebih tinggi

dibandingkan model multilingual lain. Hal ini dimungkinkan karena IndoBERT hanya dilatih dengan data berbahasa Indonesia, sehingga distribusi kosakata dan struktur sintaksis yang ditangkap model lebih sesuai dengan karakteristik bahasa target (Yulianti dan Nissa, 2024).

### 2.11 IndoRoBERTa

IndoRoBERTa merupakan salah satu varian model bahasa berbasis transformer yang dikembangkan khusus untuk bahasa Indonesia. Model ini adalah adaptasi dari RoBERTa (*Robustly Optimized BERT Pretraining Approach*) yang dirancang dengan tujuan meningkatkan pemahaman konteks dan performa klasifikasi teks berbahasa Indonesia secara lebih optimal dibandingkan dengan model BERT multibahasa atau model-model sebelumnya seperti IndoBERT (Sagama dan Alamsyah, 2023).

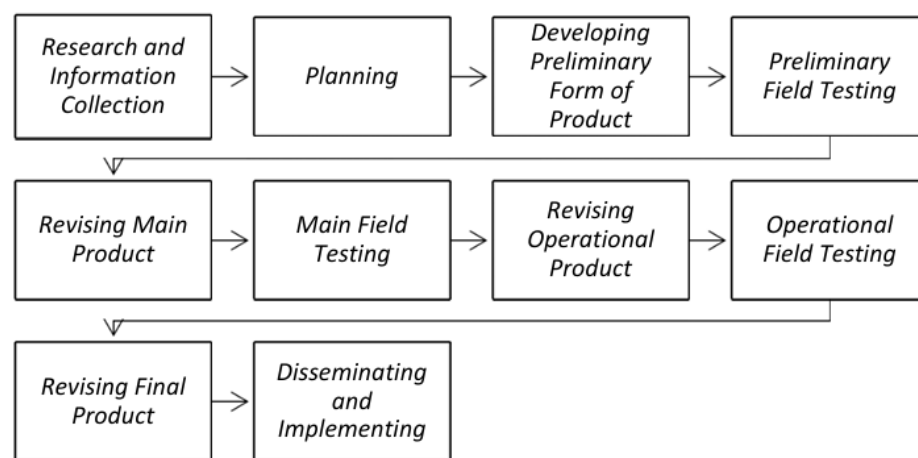
Secara teknis, IndoRoBERTa dilatih dengan dataset OSCAR (*Open Super-Large Crawled ALManaCH corpus*) (Faisal dkk., 2025), yaitu korpus data web skala besar berbahasa Indonesia. Penggunaan dataset ini bertujuan untuk menangkap variasi luas dalam struktur kalimat dan kosakata bahasa Indonesia, sehingga IndoRoBERTa dapat lebih baik dalam memahami berbagai variasi teks yang ada di Indonesia. Model ini memiliki sekitar 124 juta parameter, memungkinkan IndoRoBERTa secara efektif mempelajari pola-pola linguistik yang kompleks serta menghasilkan representasi semantik yang mendalam dan kontekstual dari teks berbahasa Indonesia.

### 2.12 Research and Development (R&D)

*Research and Development* (R&D) atau Penelitian dan Pengembangan merupakan rangkaian kegiatan yang bertujuan menciptakan produk baru yang inovatif atau menyempurnakan yang sudah ada. Proses R&D diawali dengan pengukuran kebutuhan, studi literatur, penelitian skala kecil, dan pertimbangan dari segi nilai. Tidak hanya berfokus pada penciptaan produk baru, R&D juga menguji keefektifan untuk memastikan produk berfungsi dengan baik di lingkungan yang ditargetkan (Judijanto dkk., 2024).

Pada 1960-an, R&D (di Eropa disebut RTD) menjadi model utama inovasi teknologi dan investasi bisnis, dengan tujuan meningkatkan pemahaman

manusia demi kemajuan masyarakat. Era 1980-an menandai perluasan konsep ini ke bidang pendidikan melalui adaptasi Borg & Gall, yang memanfaatkannya untuk pengembangan produk edukasi (seperti materi ajar) dan evaluasi program pembelajaran. Pada tahun 1992, Gay menambahkan bahwa R&D dalam pendidikan lebih berorientasi pada solusi praktis ketimbang sekadar uji teori. Karakteristik linear dan adaptifnya menjadikan R&D kerangka inovasi lintas disiplin yang bertahan hingga kini (Gustiani, 2019). Gambaran alur R&D menurut Borg dan Gall bisa dilihat pada Gambar 1.



**Gambar 1.** Alur Metodologi R&D Menurut Borg dan Gall (Gustiani, 2019).

Menurut Judijanto dkk. (2024) di dalam bukunya, tahapan dalam R&D meliputi beberapa langkah kritis dari penelitian awal hingga implementasi produk. Berikut adalah tahapannya:

1. Penelitian dan Pengumpulan Informasi: Melakukan studi literatur dan pengumpulan data untuk memahami masalah yang ada.
2. Perencanaan: Menyusun rencana atau desain awal berdasarkan informasi yang telah dikumpulkan.
3. Pengembangan Prototipe Awal: Membuat model awal dari produk.
4. Pengujian Awal: Melakukan uji coba lapangan terbatas untuk menilai efektivitas prototipe.

5. Revisi Produk: Mengadakan perbaikan berdasarkan *feedback* dari pengujian awal.
6. Uji Coba Lapangan Utama: Implementasi skala lebih luas untuk verifikasi lebih lanjut.
7. Revisi Operasional: Penyempurnaan akhir sebelum produk diluncurkan secara resmi.
8. Diseminasi dan Implementasi: Menyebarluaskan produk yang telah dikembangkan ke publik.

### III. METODE PENELITIAN

#### 3.1 Waktu dan Tempat

##### 3.1.1 Waktu Penelitian

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2024/2025, yang dimulai pada September 2024.

##### 3.1.2 Tempat Penelitian

Penelitian ini dilaksanakan di Jurusan Ilmu Komputer, Universitas Lampung, yang berlokasi di Jalan Prof. Dr. Ir. Sumantri Brojonegoro No. 1, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung, Lampung, 35141.

#### 3.2 Alat dan Bahan Penelitian

Penelitian ini melibatkan berbagai alat yang diperlukan untuk mendukung setiap tahap yang dilakukan sesuai dengan prosedur penelitian. Alat-alat tersebut dibedakan menjadi dua kategori, yaitu perangkat keras dan perangkat lunak.

##### 3.2.1 Perangkat Keras

Dalam pelaksanaan penelitian ini, penulis memanfaatkan laptop dengan spesifikasi yang memadai untuk mendukung proses pengembangan dan pengujian sistem. Perangkat keras yang digunakan terdiri dari komponen-komponen sebagai berikut:

- a) *Manufacturer* : Asus
- b) *Model* : ROG Strix
- c) *Processor* : AMD Ryzen 9 5900HX

- d) GPU : NVIDIA GeForce RTX 3070 Laptop GPU
- e) RAM : 32 GB
- f) *Storage* : SSD 970 EVO Plus 512 GB

### 3.2.2 Perangkat Lunak

Selama penelitian ini, penulis menggunakan berbagai perangkat lunak untuk mendukung proses pengembangan sistem. Setiap perangkat lunak dipilih berdasarkan fungsinya yang spesifik dan kontribusinya terhadap efisiensi pengembangan. Berikut ini adalah daftar perangkat lunak yang digunakan:

- a) Sistem Operasi Windows 10 Home
- b) Visual Studio Code
- c) Python
- d) Web Browser (Chrome)
- e) OpenAI *Platform*
- f) ElevenLabs
- g) VB-Audio Virtual Cable
- h) OBS Studio
- i) VTube Studio
- j) FFmpeg

## 3.3 Metode Pengumpulan Data

Penelitian ini memanfaatkan data primer melalui instrumen kuesioner terstruktur yang dirancang untuk mengevaluasi kapabilitas AI *VTuber* berbasis GPT, KAIRA, dalam obrolan ringan di *live streaming* YouTube.

### 3.3.1 Sampel dan Responden

Pemilihan sampel dilakukan secara purposive sampling dengan melibatkan beberapa partisipan yang memiliki pengalaman dengan *VTuber*, *streamer*, atau interaksi dengan *chatbot*. Kriteria ini dipilih agar ekspektasi mereka terhadap gaya bahasa dan konten obrolan ringan KAIRA bisa mencerminkan audiens target riset.



### 3.3.2 Instrumen Kuesioner

Pada pengujian awal, penulis akan menyiapkan daftar tema obrolan ringan seperti hobi dan minat, aktivitas sehari-hari, candaan ringan, dan rekomendasi santai, sedangkan partisipan akan merumuskan pertanyaan dan ekspektasi jawaban dari pertanyaannya. Partisipan akan diminta untuk menuliskan ekspektasi jawaban secara utuh yang akan digunakan untuk menghitung kemiripan semantik. Partisipan juga akan mengisi penilaian dengan skala Likert 1 – 5 dari jawaban yang diberikan KAIRA dengan seberapa dekatnya dengan ekspektasi mereka. Pada pengujian kedua, dilakukan pengujian dengan skala yang lebih besar tanpa terbatas dengan tema. Setelah pengujian selesai akan diberikan form penilaian terhadap pengalaman para partisipan saat menggunakan KAIRA.

### 3.3.3 Prosedur Pengumpulan Data

Pada tahap pengujian awal, beberapa partisipan akan dilibatkan untuk bergabung dalam *streaming* KAIRA yang bersifat privat. Partisipan mengajukan pertanyaan yang telah mereka susun berdasarkan tema yang dipilih oleh penulis. Setelah KAIRA merespon partisipan mencocokkan dengan ekspektasi jawaban mereka, nilai kecocokan tersebut akan diberikan kepada penulis melalui form penilaian menggunakan skala Likert 1 – 5. Pada tahap pengujian kedua akan melibatkan kontribusi dengan komunitas yang lebih luas untuk bergabung dalam *streaming* KAIRA. Pada tahap ini tidak dibatasi tema dalam obrolan, hanya saja masih di ranah obrolan santai. Setelah *streaming* berakhir, para partisipan akan mengisi form penilaian pengalaman selama mengikuti *streaming* KAIRA.

### 3.3.4 Teknik Analisis Data

Analisis data kuantitatif dilakukan dengan beberapa pendekatan. Pertama, distribusi skor Likert dihitung untuk menilai sejauh mana KAIRA memenuhi ekspektasi interaksi obrolan ringan berdasarkan persepsi partisipan. Kedua, analisis semantik diterapkan dengan menghitung skor cosine similarity antara ekspektasi partisipan dan

jawaban KAIRA, sehingga diperoleh ukuran objektif kesesuaian semantik. Selain itu, sistem penyaring topik dievaluasi menggunakan dataset berlabel manual untuk mengukur kinerja klasifikasi percakapan bertema game dan non-game. Hasil klasifikasi dianalisis melalui perhitungan akurasi dan divisualisasikan dalam bentuk confusion matrix untuk mengidentifikasi distribusi kesalahan klasifikasi.

### 3.4 Tahapan Penelitian

Penelitian ini memiliki delapan tahapan yang mengacu pada model R&D yang dijelaskan sebagai berikut:

#### A. Pengumpulan Informasi

Tahapan ini bertujuan untuk memahami kebutuhan dasar dalam membangun sistem AI *VTuber*. Informasi diperoleh melalui observasi pola interaksi *VTuber* dalam sesi *freetalk* di YouTube, yang kemudian dijadikan dasar untuk menentukan fitur minimum yang harus dimiliki oleh sistem.

#### B. Perencanaan

Berdasarkan informasi yang dikumpulkan, disusunlah perencanaan sistem AI *VTuber* dengan pendekatan modular. Perencanaan meliputi pemilihan teknologi pendukung (seperti model bahasa, TTS, dan platform avatar), penentuan alur interaksi, serta penyusunan parameter evaluasi untuk mengukur performa sistem dalam menanggapi obrolan santai.

#### C. Pengembangan Prototipe Awal

Pada tahap ini, model awal AI *VTuber* berbasis GPT mulai diimplementasikan sesuai desain, membangun alur percakapan, dan menyiapkan integrasi dengan perangkat lunak pendukung lainnya.

#### D. Pengujian Awal

Pada tahap ini, prototipe AI *VTuber* diuji oleh beberapa partisipan dalam sesi *live stream* terbatas. Setiap partisipan menyusun pertanyaan

dan menuliskan ekspektasi jawaban sesuai tema yang telah disiapkan. Partisipan kemudian mengajukan pertanyaan tersebut langsung pada *live chat* YouTube. Setelah sesi berakhir, partisipan akan memberikan skor Likert 1 – 5 untuk kemiripan jawaban AI dengan ekspektasi mereka sesuai tema (1 = sangat tidak sesuai, 5 = sangat sesuai). Untuk mendukung evaluasi, analisis kesamaan semantik dilakukan dengan menghitung *cosine similarity* antara teks ekspektasi dan jawaban AI menggunakan model IndoBERT untuk melakukan *sentence embedding* atau mengubah teks menjadi representasi vektor. Selain menggunakan IndoBERT, juga akan menggunakan IndoRoBERTa untuk melihat perbandingan pengukuran model-model tersebut. Skor semantik ini melengkapi skor Likert dan membantu mengukur sejauh mana KAIRA meniru nuansa obrolan ringan alami.

#### E. Revisi Prototipe

Berdasarkan data pengujian awal dan umpan balik yang diberikan akan dilakukan perbaikan pada prototipe, menyempurnakan alur respon, dan sinkronisasi perangkat lunak jika diperlukan.

#### F. Uji Coba Lapangan Utama

Versi prototipe revisi akan diujikan pada audiens publik dalam sesi *live streaming* terbuka tanpa pembatasan tema tetapi masih dalam konteks obrolan ringan. Semua partisipan yang terlibat akan diminta untuk mengisi form pengalaman penggunaan KAIRA setelah sesi *live streaming* berakhir.

#### G. Revisi Operasional

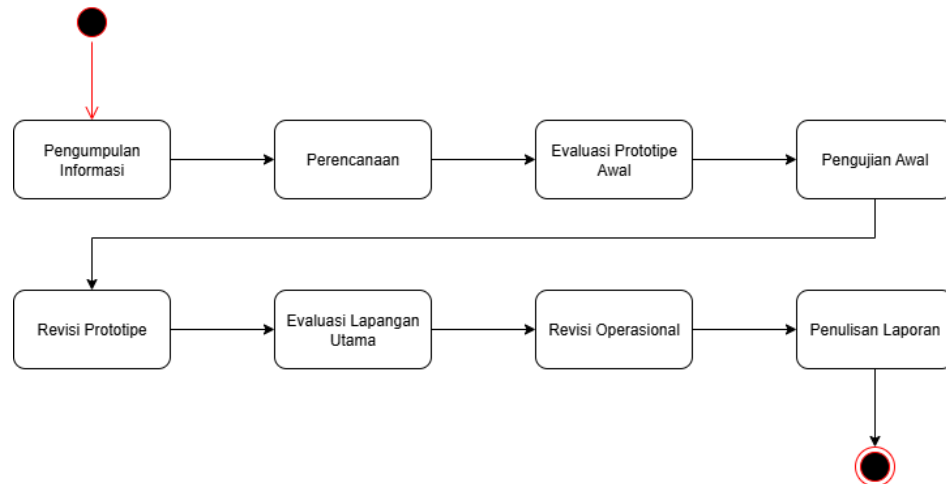
Tahap revisi operasional dilakukan untuk menyempurnakan prototipe KAIRA agar siap menuju produk final. Pada tahap ini dilakukan penyempurnaan performa sistem, peningkatan kualitas interaksi, serta evaluasi teknis lanjutan untuk memastikan prototipe berfungsi secara optimal. Kegiatan meliputi optimasi operasional berdasarkan hasil uji coba lapangan, pengujian ulang kinerja sistem pada kondisi nyata, dan penilaian akurasi mekanisme pendukung seperti penyaring topik. Hasil

dari tahap ini menjadi dasar untuk memutuskan kesiapan sistem dalam implementasi final.

#### H. Penulisan Laporan

Tahap terakhir meliputi pendokumentasian lengkap serta rekomendasi implementasi untuk komunitas *streaming* YouTube.

Alur dari tahapan penelitian ini bisa dilihat pada Gambar 2.



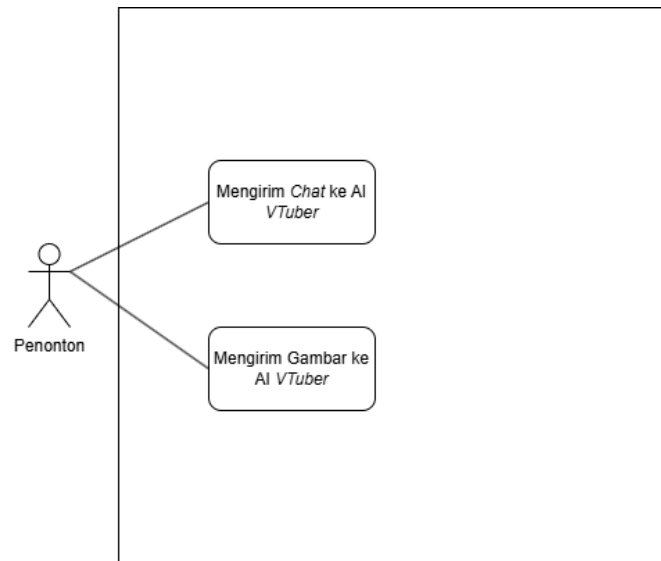
**Gambar 2.** Tahapan Penelitian.

### 3.5 Perancangan Sistem

Perancangan sistem bertujuan untuk menggambarkan alur interaksi antara pengguna dan sistem, serta proses-proses yang terjadi di dalam sistem secara menyeluruh sebelum diimplementasikan. Perancangan ini meliputi penyusunan *use case diagram* untuk memodelkan kebutuhan fungsional sistem dan *activity diagram* untuk memvisualisasikan alur aktivitas yang terjadi dalam sistem.

#### 3.5.1 Use Case Diagram

Diagram ini memodelkan bagaimana pengguna (penonton) dapat berinteraksi dengan sistem, serta fitur-fitur yang diakses. *Use case diagram* dari KAIRA dapat dilihat pada Gambar 3.



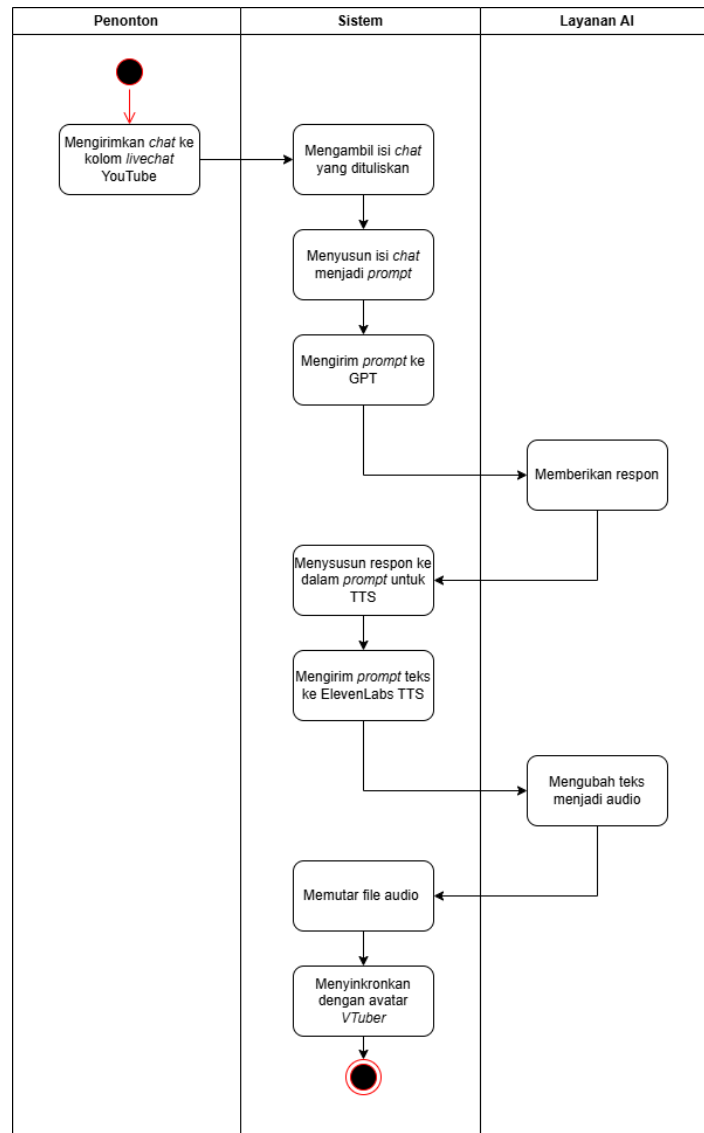
**Gambar 3.** *Use Case Diagram AI VTuber KAIRA.*

### 3.5.2 Activity Diagram

Diagram ini digunakan untuk memvisualisasikan alur proses atau aktivitas yang terjadi di dalam sistem KAIRA untuk membantu memahami langkah-langkah proses yang berjalan di dalam sistem.

#### A. Diagram Alur *Input Chat*

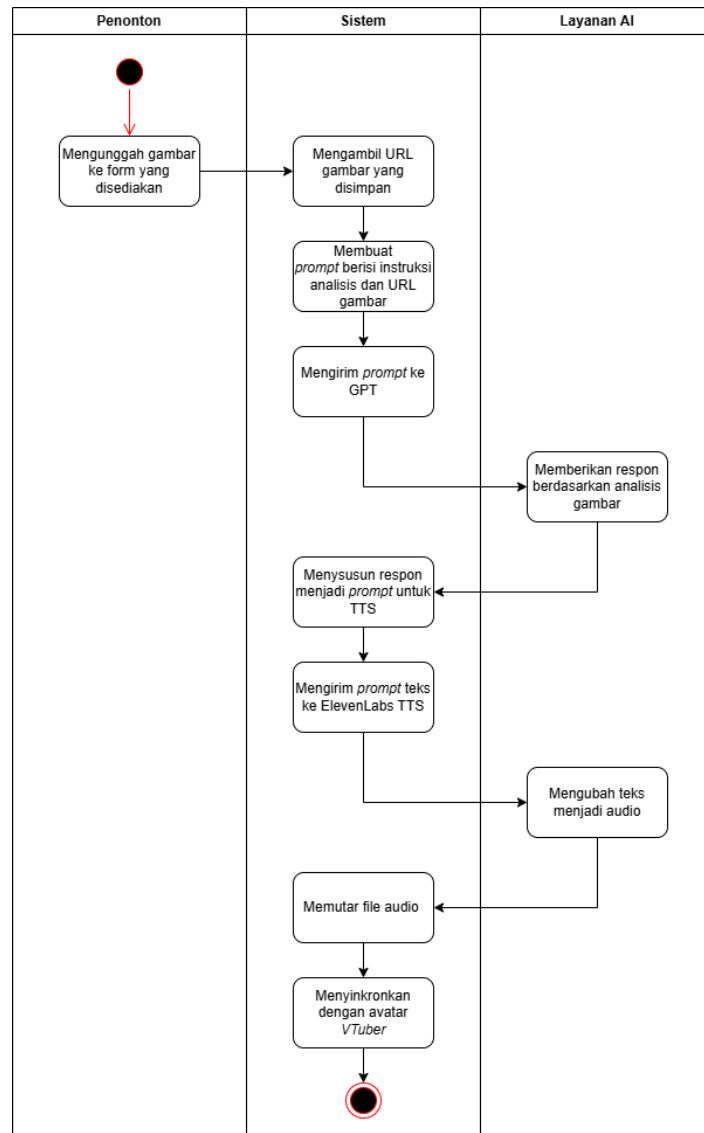
Diagram ini menggambarkan alur kerja mulai dari penonton mengirim pesan di *live chat* hingga sistem menghasilkan respon suara yang disinkronkan dengan avatar *VTuber*. *Activity diagram* alur *chat* bisa dilihat pada Gambar 4.



**Gambar 4.** *Activity Diagram Alur Chat.*

#### B. Diagram Alur *Input* Gambar

Diagram ini menggambarkan alur kerja mulai dari penonton mengunggah gambar pada form yang disediakan hingga sistem menghasilkan respon suara yang disinkronkan dengan avatar *VTuber*. *Activity diagram* alur *chat* bisa dilihat pada Gambar 5.

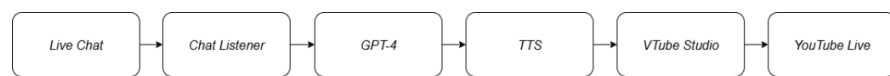


**Gambar 5.** Activity Diagram Alur Input Gambar.

### 3.5.1 Model Interaksi

Model interaksi menggambarkan alur komunikasi antara pengguna, sistem AI *VTuber*, dan *platform live streaming*. Alur interaksi dimulai ketika pengguna mengirim pesan melalui YouTube *Live Chat*, yang kemudian ditangkap oleh komponen *Chat Listener*. Pesan tersebut diteruskan ke GPT-4 untuk diproses menjadi respons teks yang kontekstual dan relevan dengan percakapan. Respons teks selanjutnya dikirim ke layanan *text-to-speech* (TTS) ElevenLabs, yang mengubah teks menjadi audio.

Audio yang dihasilkan akan diputar melalui VTube Studio, yang menampilkan avatar AI *VTuber* seolah berbicara secara *real-time*. Hasil akhir berupa kombinasi audio dan animasi avatar kemudian ditampilkan melalui siaran YouTube Live, sehingga pengguna dapat mendengar dan melihat respons AI *VTuber* secara langsung. Model interaksi ini memberikan gambaran alur komunikasi dari *input* pengguna hingga *output* audiovisual, yang menjadi dasar pengembangan dan evaluasi prototipe AI *VTuber*. Visualisasi model interaksi sistem ditunjukkan pada Gambar 6.



**Gambar 6.** Model Interaksi AI *VTuber*.

### 3.6 Jadwal Penelitian

Jadwal pelaksanaan setiap tahapan penelitian disajikan dalam bentuk tabel linimasa (*Gantt chart*). Tujuannya agar memudahkan untuk melihat urutan kegiatan, durasi, dan periode waktu pelaksanaan penelitian secara keseluruhan. Periode penelitian dimulai pada September 2024. Tabel linimasa bisa dilihat pada Tabel 4.





## V. SIMPULAN DAN SARAN

### 5.1 Simpulan

Berdasarkan hasil pengembangan, pengujian, dan analisis terhadap sistem AI VTuber bernama KAIRA, dapat diambil kesimpulan sebagai berikut:

- A. Sistem AI *VTuber* KAIRA berbasis GPT-4 berhasil dikembangkan dan diintegrasikan dengan avatar virtual secara *real-time*, memanfaatkan *platform* YouTube *Live Chat*, layanan *text-to-speech*, dan layanan GPT. Sistem ini mampu menjalankan sesi *live streaming* tanpa keterlibatan manusia secara langsung.
- B. Performa KAIRA sebagai AI *VTuber* menunjukkan hasil yang cukup baik dalam menciptakan percakapan yang menyenangkan dan relevan dengan konteks obrolan santai. Penilaian partisipan dari dua tahap uji coba menunjukkan bahwa KAIRA mampu menjawab pertanyaan dengan cukup baik, memahami maksud percakapan, serta menjaga kenyamanan gaya bahasa. Namun, masih ditemukan beberapa kekurangan, terutama dalam hal kecepatan respon dan kedalaman ekspresi karakter yang membuat KAIRA belum sepenuhnya terasa sebagai sosok virtual yang “hidup”. Hal ini menunjukkan bahwa meskipun teknologi AI dapat menghadirkan interaksi yang menyerupai manusia, aspek emosional dan spontanitas masih menjadi tantangan yang perlu dikembangkan lebih lanjut.

- C. Evaluasi tambahan pada tahap revisi operasional menunjukkan bahwa akurasi sistem penyaring topik untuk membedakan percakapan bertema game dan non-game masih rendah, yaitu sebesar 50,33%. Hal ini menunjukkan adanya keterbatasan korpus dan kesalahan klasifikasi, meskipun sistem tetap mampu memberikan respons relevan berkat desain routing yang tidak sepenuhnya membatasi percakapan.
- D. Pengembangan KAIRA sebagai AI VTuber menghadirkan tantangan teknis dan etis. Dari sisi teknis, hambatan utama meliputi keterlambatan respon akibat proses berlapis, pemahaman konteks yang kadang terlalu literal terhadap candaan atau bahasa gaul, minimnya ekspresi avatar, dan pengaturan teknis yang membuat sebagian pesan tampak diabaikan. Dari sisi etis, potensi misinformasi tetap ada meski identitas KAIRA sebagai AI disampaikan secara jelas. Kepribadian netral dapat terasa kurang autentik, dan interaksi jangka panjang berisiko menimbulkan keterikatan emosional pengguna, yang menimbulkan pertanyaan etis tentang hubungan manusia dengan karakter virtual.

## 5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan pertimbangan untuk pengembangan sistem dan penelitian lanjutan.

Pertama, peningkatan kecepatan respon sistem menjadi salah satu prioritas penting. Disarankan dilakukan perbaikan pada waktu tanggap, baik dalam pengelolaan permintaan API maupun manajemen antrian pesan. Tujuannya adalah agar interaksi terasa lebih *real-time* dan responsif, serta mengurangi jeda yang dirasakan pengguna.

Kedua, diperlukan pengembangan ekspresi karakter virtual untuk memperkuat kesan “hidup” dari KAIRA di mata penonton. Sistem dapat ditingkatkan pada aspek visual seperti sinkronisasi bibir (*lip-sync*), ekspresi wajah, serta gerakan tubuh yang relevan terhadap konteks obrolan. VTuber Studio sebagai *platform* penyaji avatar virtual mendukung fitur pendeteksian ekspresi berbasis webcam atau perangkat pelacak (*tracker*),

serta memungkinkan penggunaan *hotkey* dan *plugin* eksternal untuk mengatur gerakan. Pengembang dapat memanfaatkan skrip pemicu animasi berdasarkan jenis emosi atau kata kunci dalam output GPT, seperti animasi senyum dan lambaian tangan saat percakapan ceria, atau ekspresi murung untuk nada sedih. Implementasi ini dapat dilakukan melalui integrasi VTube Studio API atau *middleware* seperti VTS WebSocket Server yang dikontrol menggunakan Python atau Node.js. Namun perlu diperhatikan bahwa tidak semua model avatar memiliki parameter ekspresi atau gerak tubuh lengkap, terutama pada model dengan tampilan setengah badan.

Ketiga, eksplorasi penerapan sistem di *platform live streaming* lain seperti Twitch atau TikTok Live bisa menjadi arah penelitian selanjutnya. Hal ini memungkinkan pengujian terhadap efektivitas sistem dan respon audiens dalam konteks *platform* dengan karakteristik pengguna yang berbeda.

Keempat, pengembangan interaksi multibahasa dan kemampuan merespon emosi secara kontekstual juga sangat potensial. Dukungan terhadap berbagai bahasa serta kemampuan mengenali dan merespons ekspresi emosional dalam obrolan dapat meningkatkan kualitas komunikasi dan memperluas jangkauan sistem secara global.

Kelima, peningkatan akurasi sistem penyaring topik perlu menjadi salah satu fokus utama. Perluasan korpus data dengan mengambil lebih banyak subreddit yang berkaitan dengan tema yang digunakan, penyesuaian threshold, dan penerapan metode *filtering* hibrida yang dapat meningkatkan ketepatan klasifikasi tanpa mengorbankan kelancaran interaksi.

## DAFTAR PUSTAKA

- Al-Fedaghi, S. (2021). Validation: Conceptual versus Activity Diagram Approaches. (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, 12(6), 287–297. [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org) [Diakses: Jun. 16, 2025]
- Al-Ghonmein, A. M., & Al-Moghrabi, K. G. (2024). The potential of ChatGPT technology in education: advantages, obstacles and future growth. *IAES International Journal of Artificial Intelligence*, 13(2), 1206–1213. <https://doi.org/10.11591/ijai.v13.i2.pp1206-1213>
- Aljanabi, M. (2023). ChatGPT: Future Directions and Open possibilities. *Mesopotamian Journal of Cybersecurity*, 2023, 16–17. <https://doi.org/10.58496/MJCS/2023/003>
- Aljanabi, M., Ghazi, M., Ali, A. H., & Abed, S. A. (2023). ChatGpt: Open Possibilities. *Iraqi Journal for Computer Science and Mathematics*, 4(1), 62–64. <https://doi.org/10.52866/20ijcsm.2023.01.01.0018>
- Alturas, B. (2023). Connection between UML use case diagrams and UML class diagrams: a matrix proposal. *International Journal of Computer Applications in Technology*, 72(3), 161–168. <https://doi.org/10.1504/IJCAT.2023.133294>
- Amato, N., De Carolis, B., De Gioia, F., Venezia, M. N., Palestra, G., & Loglisci, C. (2024). Can an AI-driven VTuber engage People? The KawAi Case Study. In *Joint Proceedings of the ACM IUI Workshops 2024*. ACM IUI Workshops, Greenville, South Carolina, USA, 1–10.
- Aquino, E., de Saqui-Sannes, P., & Vingerhoeds, R. A. (2020). A Methodological Assistant for Use Case Diagrams. In *Proceedings of the 8th International Conference on Model-Driven Engineering and Software Development*. MODELSWARD 2020, Valletta, Malta, 227–236. <https://doi.org/10.5220/0008938002270236>

- 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., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D. (2020). *Language Models are Few-Shot Learners*. <https://doi.org/10.48550/arXiv.2005.14165>
- Chen, H.-W. (2023). *Endoscopic Endonasal Skull Base Surgery For Pituitary Lesions: An AI-Assisted Creative Workflow To Develop An Animated Educational Resource For Patients And Physicians*.
- Dewatri, R. A. F., Aqthar, A. Z. Al, Pradana, H., Anugerah, B., & Nurcahyo, W. H. (2023). Potential Tools to Support Learning: OpenAI and Elevenlabs Integration. *ODELIA: Southeast Asia Journal on Open Distance Learning*, 01(02), 59–69.
- Faisal, M. R., Fitriani, K. E., Mazdadi, M. I., Indriani, F., Nugrahadi, D. T., & Prastya, S. E. (2025). Enhancing Natural Disaster Monitoring: A Deep Learning Approach to Social Media Analysis Using Indonesian BERT Variants. *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, 7(1), 77–89. <https://doi.org/10.35882/ijecemi.v7i1.38>
- Geni, L., Yulianti, E., & Sensuse, D. I. (2023). Sentiment Analysis of Tweets Before the 2024 Elections in Indonesia Using Bert Language Models. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 9(3), 746–757. <https://doi.org/10.26555/jiteki.v9i3.26490>
- Google LLC. (2025). *Stream Latency*. [https://support.google.com/youtube/answer/7444635?hl=id&ref\\_topic=9257892&sjid=11086992659188788035-NC](https://support.google.com/youtube/answer/7444635?hl=id&ref_topic=9257892&sjid=11086992659188788035-NC) [Diakses: Jun. 17, 2025]
- Gustiani, S. (2019). Research And Development (R&D) Method As A Model Design In Educational Research And Its Alternatives. *HOLISTICS JOURNAL*, 11(2), 12–22. <https://www.researchgate.net/publication/365669215> [Diakses: Mei. 2, 2025]
- Hamilton, W. A., Garretson, O., & Kerne, A. (2014). Streaming on twitch: Fostering participatory communities of play within live mixed media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1315–1324. <https://doi.org/10.1145/2556288.2557048>
- Hendini, A. (2016). Pemodelan UML Sistem Informasi Monitoring Penjualan Dan Stok Barang (Studi Kasus: Distro Zhezha Pontianak). *JURNAL KHATULISTIWA INFORMATIKA*, 4(2), 107–116.

- Hermawan, H., Subarkah, P., Utomo, A. T., Ilham, F., & Saputra, D. I. S. (2024). VTuber Personas In Digital Wayang: A Review Of Innovative Cultural Promotion For Indonesian Heritage. *Jurnal Pilar Nusa Mandiri*, 20(2), 165–175. <https://doi.org/10.33480/pilar.v20i2.5921>
- Hidayat, A. F. (2024). *Evaluasi Keandalan Cosine Similarity dalam Mendeteksi Plagiarisme Kode Program*. IF2123 Aljabar Linier dan Geometri, Institut Teknologi Bandung.
- Hu, J. (2024). Core Technology, Typical Applications, And Future Development Prospects of ChatGPT. *Highlights in Science, Engineering and Technology CSIC*, 85, 347–353.
- Jamaluddin, A. (2021). The Effectiveness Of Vtuber As A Medium Of Information. In *World Islamic Social Science Congress*. Universiti Islam Sultan Sharif Ali, Brunei Darussalam. <https://www.researchgate.net/publication/370779992> [Diakses: Sep. 30, 2025]
- Joshi, A., Kale, S., Chandel, S., & Pal, D. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7(4), 396–403. <https://doi.org/10.9734/bjast/2015/14975>
- Judijanto, L., Muhammadijah, M., Utami, R. N., Suhirman, L., Laka, L., Boari, Y., Lembang, S. T., Wattimena, F. Y., Astriawati, N., Laksono, R. D., MARS, & Yunus, M. (2024). *Metodologi Research And Development (Teori dan Penerapan Metodologi RnD)* (Sepriano, Efitra, Nurzatul Dihniah, & Yayan Agusdi, Eds.). PT. Sonpedia Publishing Indonesia, Jakarta. <https://www.researchgate.net/publication/381290945> [Diakses: Mei. 2, 2025]
- Kim, D., Lee, S., Jun, Y., Shin, Y., & Lee, J. (2025). VTuber’s Atelier: The Design Space, Challenges, and Opportunities for VTubing. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*, Yokohama, Japan. ACM, New York, NY, USA, 1–23. <https://doi.org/10.1145/3706598.3714107>
- Koo, M., & Yang, S.-W. (2025). Likert-Type Scale. *Encyclopedia MDPI*, 5(18), 1–11. <https://doi.org/10.3390/encyclopedia5010018>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). *IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP*. <http://arxiv.org/abs/2011.00677>
- Lee, C.-M. (2024). The Key Factors Affecting Audience’s Support for VTubers. *Journal of Management and Information*, 29(113), 45–78. <https://www.researchgate.net/publication/384159907>
- Nielsen, J. (2023, August 2). *The Need for Speed in AI*. UX Tigers Fearless Usability. <https://www.uxtigers.com/post/ai-response-time> [Diakses: Jun. 17, 2025]

- Oyucu, S. (2023). A Novel End-to-End Turkish Text-to-Speech (TTS) System via Deep Learning. *Electronics (Switzerland)*, 12(8), 1–19. <https://doi.org/10.3390/electronics12081900>
- Puspitaningrum, D. R., & Prasetyo, A. (2019). Fenomena “Virtual Youtuber” Kizuna Ai di Kalangan Penggemar Budaya Populer Jepang di Indonesia. *Mediator: Jurnal Komunikasi*, 12(2), 128–140. <https://doi.org/10.29313/mediator.v12i2.4758>
- Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2023). *Language Models are Unsupervised Multitask Learners*. <https://api.semanticscholar.org/CorpusID:160025533> [Diakses: Sep. 6, 2025]
- Ramli, I., Jamil, N., Seman, N., & Ardi, N. (2015). An Improved Syllabification for a Better Malay Language Text-to-Speech Synthesis (TTS). *Procedia Computer Science*, 76, 417–424. <https://doi.org/10.1016/j.procs.2015.12.280>
- Rau, A., Bamberg, F., Fink, A., Hien Tran, P., Reisert, M., & Russe, M. F. (2024). Enhancing chatbot performance for imaging recommendations: Leveraging GPT-4 and context-awareness for trustworthy clinical guidance. *European Journal of Radiology*, 181, 1–5. <https://doi.org/10.1016/j.ejrad.2024.111756>
- Sagama, Y., & Alamsyah, A. (2023). Multi-Label Classification of Indonesian Online Toxicity using BERT and RoBERTa. In *2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. Bali, Indonesia, 143–149. <https://doi.org/10.1109/IAICT59002.2023.10205892>
- Sathe, T. S., Roshal, J., Naaseh, A., L’Huillier, J. C., Navarro, S. M., & Silvestri, C. (2024). How I GPT It: Development of Custom Artificial Intelligence (AI) Chatbots for Surgical Education. *Journal of Surgical Education*, 81(6), 772–775. <https://doi.org/10.1016/j.jsurg.2024.03.004>
- Schrump, M. L., Johnson, M., Ghuy, M., & Gombolay, M. C. (2020). Four Years in Review: Statistical Practices of Likert Scales in Human-Robot Interaction Studies. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*. Association for Computing Machinery, New York, NY, USA, 43–52. <https://doi.org/10.1145/3371382.3380739>
- Shi, Y., & Deng, B. (2024). Finding the sweet spot: Exploring the optimal communication delay for AI feedback tools. *Information Processing & Management*, 61(2). <https://doi.org/10.1016/j.ipm.2023.103572>
- Siewe, F., & Ngounou, G. M. (2025). On the Execution and Runtime Verification of UML Activity Diagrams. *Software MDPI*, 4(1), 4. <https://doi.org/10.3390/software4010004>
- Tang, M. T., Zhu, V. L., & Popescu, V. (2021). Alterecho: Loose avatar-streamer coupling for expressive VTubing. In *2021 IEEE International Symposium on*



*Mixed and Augmented Reality (ISMAR)*. Bari, Italy, 128–137.  
<https://doi.org/10.1109/ISMAR52148.2021.00027>

- Van Poucke, M. (2024). ChatGPT, the perfect virtual teaching assistant? Ideological bias in learner-chatbot interactions. *Computers and Composition*, 73, 102871. <https://doi.org/10.1016/j.compcom.2024.102871>
- Wang, Y. L., & Lo, C. W. (2025). The effects of response time on older and young adults' interaction experience with Chatbot. *BMC Psychology*, 13(1), 1–12. <https://doi.org/10.1186/s40359-025-02459-9>
- Yao, Y., Liang, T., Feng, R., Shi, K., Yu, J., Wang, W., & Li, J. (2024). SR-TTS: a rhyme-based end-to-end speech synthesis system. *Frontiers in Neurorobotics*, 18, 1322312. <https://doi.org/10.3389/fnbot.2024.1322312>
- Yulianti, E., & Nissa, N. K. (2024). ABSA of Indonesian customer reviews using IndoBERT: single-sentence and sentence-pair classification approaches. *Bulletin of Electrical Engineering and Informatics*, 13(5), 3579–3589. <https://doi.org/10.11591/eei.v13i5.8032>
- Zhou, K., Ethayarajh, K., Card, D., & Jurafsky, D. (2022). *Problems with Cosine as a Measure of Embedding Similarity for High Frequency Words*. <http://arxiv.org/abs/2205.05092>