

PENGEMBANGAN *RETRIEVAL-AUGMENTED GENERATION (RAG) CHATBOT* MENGGUNAKAN METODE *DESIGN THINKING* UNTUK Mendukung *PERSONALIZED LEARNING*

Fitrah Izul Falaq¹, Nugroho Saputra², Nathanael Paskal³

^{1,2}Dosen Prodi Pendidikan Teknik Informatika dan Komputer, Teknik Elektro, FT – UNJ

³Mahasiswa Prodi Pendidikan Teknik Informatika dan Komputer, Teknik Elektro, FT – UNJ

¹fitrah.izul.falaq@unj.ac.id, ²nugrohosaputra@unj.ac.id, ³nathanael_1512622051@mhs.unj.ac.id

Abstrak

Pemanfaatan AI dalam bidang pendidikan membawa dampak yang positif, salah satunya melalui penerapan chatbot berbasis generative ai. Implementasi chatbot dalam pembelajaran sering dimanfaatkan untuk memfasilitasi pembelajaran mandiri. Sayangnya, hambatan chatbot sering dihadapi oleh permasalahan halusinasi, bias, dan jawaban di luar konteks. Salah satu teknik untuk mengatasi hal tersebut adalah Retrieval-Augmented Generation (RAG). Implementasi RAG Chatbot berpotensi mendukung personalized learning. Metode penelitian ini menggunakan model Design Thinking yang terdiri dari tahap empathize, define, ideate, prototype dan test. Metode pengembangan ini cocok karena penekanannya pada empati pengguna, iterasi prototipe cepat, dan validasi kebutuhan nyata. Berdasarkan hasil uji fungsionalitas menggunakan metode black-box testing, memperoleh skor 15/15 (100%). Sedangkan uji kegunaan menggunakan penilaian SUS memperoleh skor 74 (B), skor di atas rata-rata. Kesimpulannya, chatbot diterima dan layak digunakan untuk mendukung personalized learning. Pengguna merasa nyaman dan puas, namun perlu ditingkatkan agar pengguna aktif merekomendasikan.

Kata kunci : *Chatbot, Generative Ai, Personalized Learning, RAG*

1. Pendahuluan

Perkembangan Artificial Intelligence (AI) mendorong pemanfaatan chatbot pendidikan sebagai pendamping belajar yang mampu merespons secara natural, memberikan umpan balik cepat, serta mendukung pembelajaran mandiri. Berbagai penelitian menunjukkan bahwa chatbot dapat meningkatkan asistensi belajar, personalisasi pengalaman belajar, dan efisiensi tugas pendidik, meskipun masih menghadapi tantangan terkait etika, akurasi, dan keandalan (Labadze dkk., 2023; Okonkwo & Ade-Ibijola, 2021). UNESCO (2023) menekankan pentingnya pendekatan human-centred dalam pemanfaatan Generative AI di pendidikan, termasuk validasi pedagogis, mitigasi bias, perlindungan privasi, serta peningkatan kapasitas pengguna. Secara teknis, persoalan seperti halusinasi, kehilangan konteks, dan bias jawaban mendorong penggunaan pendekatan Retrieval-Augmented Generation (RAG), yang menggabungkan model generatif dengan pencarian dokumen eksternal untuk meningkatkan akurasi dan mengurangi kesalahan (Lewis dkk., 2020; Shuster dkk., 2021).

Pengembangan chatbot berbasis RAG perlu berlandaskan metodologi yang berorientasi pada pengguna. Design thinking dipilih karena menekankan empati, iterasi prototipe, dan validasi kebutuhan nyata (Razzouk & Shute, 2012). Pada tahap evaluasi, penelitian ini menggunakan black-box testing untuk memastikan kesesuaian fungsi sistem dengan spesifikasi (Ammann & Offutt, 2008) serta System Usability Scale (SUS) untuk mengukur tingkat kegunaan dan penerimaan pengguna (Brooke, 1996; Bangor dkk., 2009). Dengan demikian, penelitian ini memosisikan chatbot RAG sebagai media pembelajaran mandiri yang mendukung personalisasi di pendidikan tinggi melalui perancangan, pengujian fungsional, dan evaluasi berbasis bukti empiris guna menghasilkan rekomendasi implementasi yang bertanggung jawab.

2. Dasar Teori

2.1. Chatbot berbasis Generative AI

Chatbot berbasis Generative AI merupakan aplikasi yang didukung Large Language Model (LLM) dan mampu menghasilkan respons natural secara otonom berdasarkan input pengguna. Dalam konteks pendidikan,

Available at:

<https://journal.unj.ac.id/unj/index.php/pinter/issue/view/2279>

chatbot dimanfaatkan sebagai tutor mikro, asisten pemberi umpan balik, serta fasilitator pembelajaran mandiri yang tersedia secara on-demand. Penelitian menunjukkan bahwa manfaat utamanya meliputi dukungan belajar, personalisasi pengalaman belajar, dan efisiensi waktu bagi pendidik, meskipun tetap memiliki risiko seperti halusinasi, kehilangan konteks, dan bias apabila tidak didukung pengetahuan eksternal (Labadze dkk., 2023). Oleh karena itu, pengembangan chatbot pembelajaran memerlukan kurasi sumber, guardrail yang sistematis, serta pengujian sebelum diterapkan. Sejalan dengan itu, UNESCO (2023) menekankan pendekatan human-centred dalam pemanfaatan GenAI di pendidikan, yang mencakup perlindungan data, transparansi, mitigasi bias, akuntabilitas penyedia, serta penguatan kapasitas pendidik. Prinsip ini mengharuskan respons chatbot dapat ditelusuri ke sumber terpercaya (provenans), menjaga keamanan data pribadi, dan diuji kelayakan pedagogisnya sebelum diintegrasikan dalam lingkungan belajar.

2.2. Retrieval-Augmented Generation (RAG)

Retrieval-Augmented Generation (RAG) adalah kerangka kerja yang menggabungkan kemampuan model bahasa besar (LLM) yang telah dilatih sebelumnya dengan sistem pengambilan informasi (*retrieval system*) yang dapat diakses dari memori eksternal. Tujuan utamanya adalah untuk mengatasi keterbatasan LLM dalam mengakses dan memanipulasi pengetahuan secara akurat, terutama untuk tugas-tugas yang intensif pengetahuan. (Lewis dkk., 2020). RAG mengatasi beberapa tantangan signifikan yang dihadapi LLM para terik seperti halusinasi, pengetahuan yang tidak *up to date*, meningkatkan kredibilitas dan *provability* (Asai dkk., 2024; Procko & Ochoa, 2024). Model RAG dilatih secara *end-to-end* untuk menggunakan *input*, seperti sebuah pertanyaan, untuk mengambil potongan teks yang relevan dari memori non-para metrik dan kemudian menggunakan potongan teks ini sebagai konteks tambahan saat menghasilkan *output* target, seperti sebuah jawaban (Lewis dkk., 2020).

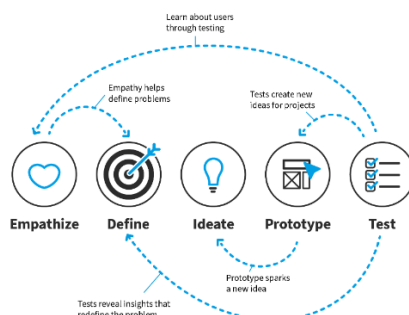
2.3. Pembelajaran Terpersonalifikasi (*Personalized Learning*)

Personalized learning (PL) merujuk pada penyesuaian tujuan, urutan, dan dukungan pembelajaran sesuai kebutuhan serta perkembangan individu. Meta-analisis dalam *British Journal of Educational Technology* terhadap 16 randomized controlled trials menunjukkan bahwa teknologi yang memfasilitasi personalisasi memberikan efek positif signifikan, khususnya pada literasi dan matematika, dengan ukuran efek kecil hingga menengah (Major dkk., 2021). Studi terbaru terkait AI dan chatbot juga menunjukkan kecenderungan peningkatan hasil belajar, meskipun efektivitasnya bervariasi tergantung jenjang pendidikan, desain intervensi, dan peran chatbot. Hal ini menegaskan bahwa keberhasilan PL berbasis AI sangat bergantung pada kualitas desain instruksional dan pengendalian konten. Dalam konteks tersebut, pendekatan Retrieval-Augmented Generation (RAG) dapat menjadi solusi dengan menggabungkan personalisasi respons dan akses ke sumber materi yang valid, sehingga profil belajar individu tetap terpenuhi tanpa mengorbankan keandalan informasi.

3. Metodologi

3.1. Metode Pengembangan

Penelitian pengembangan ini menggunakan metode *design thinking*. Pemilihan metode ini sesuai dengan karakteristik *design thinking* yang lebih mengacu pada kecepatan, ketepatan dan berorientasi kepada pengguna (Razzouk & Shute, 2012). Pertama kali diperkenalkan oleh Stanford University's Hasso Plattner Institute of Design, yang akrab disebut *d.school*. Hal ini cocok dengan tujuan penelitian pengembangan yang ditujukan untuk memfasilitasi pembelajaran mandiri. Metode *design thinking* terdiri dari lima tahapan di antaranya *empathize*, *define*, *ideate*, *prototype* dan *test*. Tahapan *design thinking* tidak dilakukan secara linier, artinya dapat beralih dari satu tahap ke tahap lainnya, contohnya dari tahap pengujian ke tahap pendefinisian jika pengujian tersebut mengungkapkan wawasan yang mendefinisikan ulang masalah. (Interaction Design Foundation - IxDF., 2025). Dapat dilihat pada gambar 3.1.



Gambar 3.1. Alur Metode *Design Thinking Framework* (Interaction Design Foundation - IxDF., 2025)

3.2. Teknik Pengumpulan dan Analisis Data

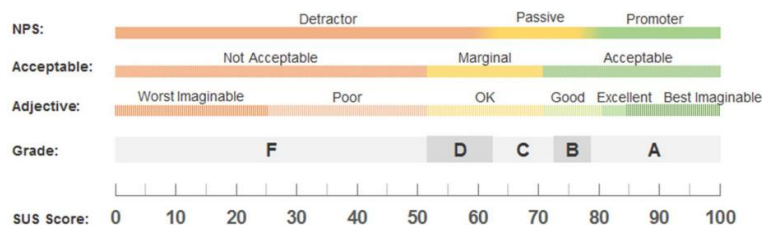
Proses pengumpulan dan analisis data dilakukan secara daring menggunakan layanan *google form*. Data yang dikumpulkan selanjutnya dianalisis sesuai dengan kebutuhan. Analisis data dilakukan untuk uji fungsionalitas dan uji tingkat kegunaan. Secara detail, implementasi pengumpulan data dan analisis data dapat dijabarkan sebagai berikut:

1) Penilaian *Black-box*

Black box testing merupakan pengujian yang bertujuan melihat program tersebut sama dengan tugas program tersebut tanpa harus mengetahui kode program yang di pakai (Nur dkk., 2020). Pengujian ini dilakukan untuk mengukur tingkat keberhasilan sebuah sistem atau aplikasi. Umumnya, pengujian dilakukan dengan pembuatan *case* pengujian, uji kualitas dan menemukan kesalahan yang tidak terdeteksi yang disebabkan oleh kesalahan eksekusi program. Terdapat beberapa metode pengujian dalam pengujian black box seperti *equivalence partitioning, cause effect graph, orthogonal array testing, comparison testing, orthogonal array testing, random data selection, performance testing, feature test, fuzzing, sample testing, robustness testing, boundary value analysis, behaviour testing, endurance testing, all-pair testing*, dan lain-lain. (Parlika dkk., 2020).

2) Uji Tingkat Kegunaan *System Usability Scale (SUS)*

System Usability Scale (SUS) merupakan instrumen pengukuran kegunaan yang telah digunakan secara luas sejak 1986 dan dirancang untuk memperoleh penilaian cepat, akurat, serta berbiaya rendah terhadap persepsi pengguna terhadap suatu sistem (Brooke, 1996). Instrumen ini cocok digunakan pada layanan berbasis website dan tidak memerlukan jumlah responden yang besar, di mana 12–14 responden dinilai cukup untuk menghasilkan skor yang andal (Tullis & Stetson, 2004). SUS terdiri dari 10 item dengan skala 0–4; untuk item positif (1, 3, 5, 7, 9), skor dihitung dengan mengurangi nilai skala dengan 1, sedangkan untuk item negatif (2, 4, 6, 8, 10), skor dihitung dengan mengurangkan nilai dari 5. Jumlah skor kemudian dikalikan 2,5 untuk memperoleh nilai akhir SUS (Brooke, 2013). Skor tersebut selanjutnya diinterpretasikan berdasarkan kategori persentil, grade, sifat, tingkat penerimaan, dan Net Promoter Score (NPS) sesuai skala interpretasi yang digunakan.



Gambar 3.2. Skala Interpretasi Hasil Skor SUS (Sauro, 2018)

Untuk memudahkan interpretasi data, skala tersebut dapat diolah menjadi sebuah tabel yang lebih detail sebagai berikut, dapat dilihat pada tabel 3.1 :

Tabel 3.1. Skala Interpretasi Hasil Skor SUS (Sauro, 2018)

Grade	SUS Range	Percentile Range	Adjective	Acceptable	NPS
A+	84.1-100	96-100	Best Imaginable	Acceptable	Promoter
A	80.8-84.0	90-95	Excellent	Acceptable	Promoter
A-	78.9-80.7	85-89	Excellent	Acceptable	Promoter
B+	77.2-78.8	80-84	Excellent	Acceptable	Passive
B	74.1 – 77.1	70 – 79	Excellent	Acceptable	Passive
B-	72.6 – 74.0	65 – 69	Excellent	Acceptable	Passive
C+	71.1 – 72.5	60 – 64	Good	Acceptable	Passive
C	65.0 – 71.0	41 – 59	Good	Marginal	Passive
C-	62.7 – 64.9	35 – 40	Good	Marginal	Passive
D	51.7 – 62.6	15 – 34	OK	Marginal	Detractor
F	25.1 – 51.6	2– 14	Poor	Not Acceptable	Detractor
F	0-25	0-1.9	Worst Imaginable	Not Acceptable	Detractor

4. Hasil dan Analisis

4.1. Hasil Pengembangan Produk

Produk yang dihasilkan dalam penelitian pengembangan ini adalah media pembelajaran berbasis *chatbot* yang memanfaatkan teknik *Retriew-Augmented Generation (RAG)*. Perbedaan mendasar dengan *chatbot* berbasis *generative AI* pada umumnya terletak pada kemampuan untuk memahami konteks, referensi materi tambahan, personalisasi, minimalisir halusinasi dan kesesuaian respons berdasarkan desain pengetahuan yang

diharapkan. *Chatbot* dengan teknik RAG didesain agar mampu mempelajari materi dan memberikan respons yang relevan. Dalam aspek pembelajaran, kemampuan tersebut sangat bermanfaat untuk mendukung *personalized learning*. Pengembangan utama RAG *Chatbot* mengacu pada kebutuhan dan karakteristik pengguna. Oleh karena itu, pemilihan metode *design thinking* sangat cocok digunakan karena mengacu pada kecepatan, ketepatan dan efektivitas pengguna berdasarkan tanggapan pengguna.

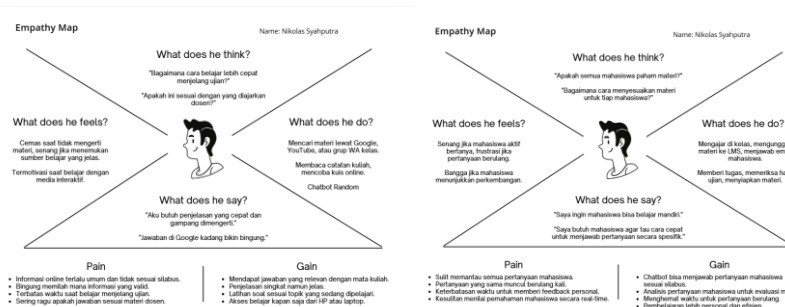
4.1.1. Empathize

Tahap pertama dalam proses *design thinking* adalah membangun empati terhadap permasalahan yang akan diselesaikan melalui analisis user persona. Proses ini dilakukan melalui studi kasus, wawancara, kajian literatur, serta analisis kebutuhan, kemudian dirangkum dalam bentuk *empathy map*. Pengembangan RAG *Chatbot* ditujukan untuk mendukung pembelajaran mahasiswa dan pembelajar berusia 19–25 tahun yang termasuk Generasi Z, dengan karakteristik digital-oriented, realistis, dan *do it yourself* (DIY) (Arum dkk., 2023). Selain itu, aspek keterjangkauan biaya juga menjadi pertimbangan agar solusi yang dikembangkan mendukung prinsip pendidikan berkualitas dan inklusif sesuai SDG's poin 4.

User persona disusun berdasarkan sampel mahasiswa Pendidikan Teknik Informatika dan Komputer semester 1–4, serta melibatkan dosen sebagai perancang materi dan evaluator pembelajaran. Proses ini diperkuat dengan temuan literatur untuk memastikan akurasi karakteristik pengguna. Hasil analisis user persona dan *empathy map* selanjutnya menjadi dasar perancangan solusi pada tahap berikutnya dapat dilihat pada Gambar 4.1 dan 4.2.



Gambar 4.1. User Persona Pengembangan RAG Chatbot



Gambar 4.2. User Persona Pengembangan RAG Chatbot

4.1.2. Define

Setelah mengumpulkan informasi selama tahap *Empathize*, berikutnya adalah menganalisis pengamatan dan menyintesisnya untuk menentukan masalah inti yang telah diidentifikasi. Pada tahap *Define*, peneliti mengidentifikasi masalah menjadi sebuah spesifikasi media, desain sistem, dan *learning journey* pemanfaatan *chatbot* dalam pembelajaran. Faktor penentu pengembangan media terdiri dengan mempertimbangkan karakteristik pengguna, kebutuhan biaya, perawatan dan keberlanjutan media. Orientasi utama dalam penelitian pengembangan ini yaitu menciptakan *chatbot* dengan *resource* terjangkau, mudah dioperasikan dan mampu mengakomodir tujuan pembelajaran berbasis *personalized learning*. Batasan pengembangan terletak pada kapasitas pengguna untuk 2 – 3 kelas digunakan secara paralel. Hasil proses *define* dapat dilihat pada tabel 4.1.

Tabel 4.1. Spesifikasi Media Pembelajaran RAG Chatbot

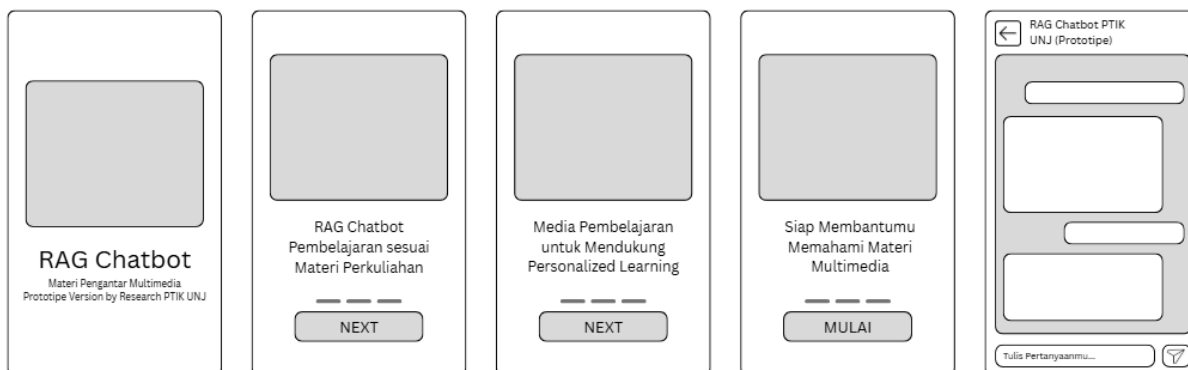
No.	Aspek	Deskripsi
1	Nama Media	RAG Chatbot for Personalized Learning (Prototype)
2	Muatan Materi	Pengantar Multimedia
3	Jenis Media Berdasarkan Fungsi	Suplemen Pembelajaran
4	Jenis Platform	Web-based
5	Akses Media	Website Responsif (Desktop & Mobile)
6	LLM Service	Google Gemini (gemini-2.0-flash), versi Gratis
7	Tech-Stack	HTML, Bootstrap CSS & Vanilla JavaScript

8	Webserver	Apache
9	Knowledge Base (Database)	Google Doc
10	Tautan Chatbot	https://project.tep.or.id/chatbot-rag
11	Spek Minimum Perangkat	Seluruh perangkat dengan web browser terbaru. Ketentuan minimal versi rilis di atas tahun 2025
12	Target pengguna	Lembaga pendidikan maupun pelatihan
13	Skala penggunaan	Terbatas untuk proses pembelajaran, tidak komersial
14	Kapasitas pengguna	20 – 50 pengguna dalam waktu bersamaan
15	Opsi Pemanfaatan dalam Pembelajaran	Suplemen pembelajaran

Berdasarkan spesifikasi tersebut, terdapat beberapa prinsip sebagai acuan kesuksesan implementasi *chatbot* untuk pembelajaran. Pertama, *chatbot* dibuat dengan teknologi sederhana untuk menekan kebutuhan *resource* yang digunakan. Hal ini sesuai kebutuhan dengan kapasitas dan skala pengguna. Kedua, layanan LLM yang digunakan sebagai pemrosesan *chatbot* utama yaitu layanan Google. Pemilihan ini untuk mendapatkan kualitas yang baik dengan biaya terjangkau dan keberlanjutan terjamin. Ketiga, pemanfaatan dalam pembelajaran menekankan pada kemudahan pengguna dengan meminimalisir prosedur. Desain pengalaman pengguna diarahkan untuk langsung memanfaatkan fungsi *chatbot*.

4.1.3. Ideate

Tahapan *ideate* berusaha menghasilkan beragam solusi kreatif untuk menjawab permasalahan yang telah dirumuskan sebelumnya (*Interaction Design Foundation - IxDF.*, 2025). Dalam pengembangan RAG *Chatbot* ini, desain antar muka yang dihasilkan mengacu pada kebutuhan desain sistem dan kemudahan penggunaan media. Adapun opsi yang dipilih adalah menyederhanakan tampilan dengan menampilkan *splash* menu yang berisi informasi *chatbot* dan menu *chatting* untuk berinteraksi. Konsep tampilan antar muka *chatbot* pada Gambar 4.3.



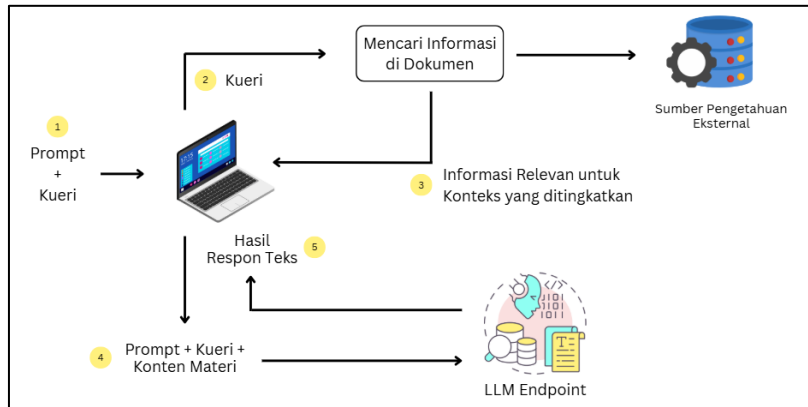
Gambar 4.3. Konsep Wireframe RAG Chatbot

Berdasarkan konsep antar muka tersebut, selanjutnya menghasilkan desain alur sistem dan opsi pemanfaatan dalam pembelajaran. Alur sistem merumuskan bagaimana proses secara teknis RAG *Chatbot* bekerja. Sedangkan Opsi pemanfaatan berisi petunjuk implementasi *chatbot* dalam pembelajaran.

1) Alur Desain Sistem

Pengembangan RAG *Chatbot* menggunakan teknik RAG yang memungkinkan *chatbot* menjawab pertanyaan berdasarkan pengetahuan eksternal. Adapun mekanismenya secara rinci sebagai berikut, pada gambar 4.4:

1. Pengguna melakukan *input* perintah;
2. *Chatbot* akan melakukan pencarian dokumen yang relevan. Dalam hal ini, materi yang diinputkan sebelumnya adalah materi Pengantar Multimedia. Proses pencarian menggunakan perintah pengambilan data pengetahuan yang telah disimpan dalam *database* pengetahuan. *Knowledge base* yang digunakan disimpan dalam *google doc* yang dapat menyediakan *data* untuk diakses oleh *chatbot*;
3. *Chatbot* menerima tambahan pengetahuan dan memasukkannya ke dalam *prompt*;
4. *Prompt* final digabungkan dengan kueri sebelumnya kemudian dikirimkan pada layanan LLM, dalam hal ini menggunakan Gemini 2.0;
5. Pengguna akan menerima respons hasil dari pengolahan LLM.



Gambar 4.4. RAG Pipeline Chatbot PTIK

2) Petunjuk Pemanfaatan Media Pembelajaran RAG Chatbot

Selanjutnya, dalam aspek pembelajaran, pemanfaatan *Chatbot* RAG dapat digunakan baik sebelum, saat dan setelah proses pembelajaran. Pemanfaatan *Chatbot* dalam RAG efektif digunakan sebagai suplemen. Cara kerjanya sama seperti pada umumnya, namun konteksnya dapat disesuaikan kebutuhan. Secara rinci, pemanfaatan RAG untuk pembelajaran dapat dilihat pada tabel 4.2.

Tabel 4.2. Opsi Strategi Pembelajaran dengan Memanfaatkan RAG Chatbot

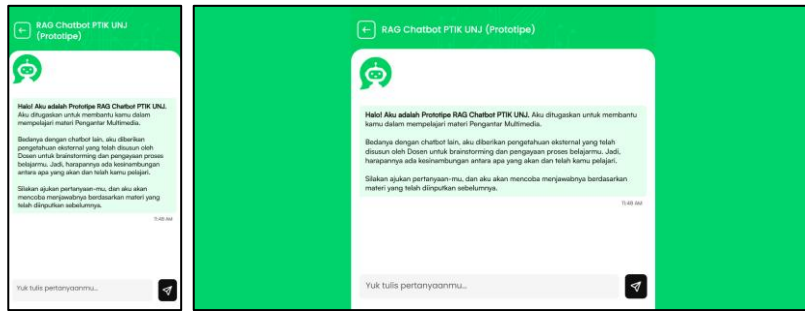
No.	Fase Pembelajaran	Aktivitas Pemanfaatan RAG Chatbot dalam Pembelajaran
1	Pra Pembelajaran	<ul style="list-style-type: none"> - Pembelajar (Guru/Dosen/<i>Trainer</i>) melakukan pembuatan materi menggunakan <i>google doc</i> sesuai dengan tujuan pembelajaran serta topik bahasan yang akan diajarkan. - Siswa dapat mengakses <i>chatbot</i> untuk <i>brainstorming</i> terkait materi yang akan dipelajari. Proses pembaruan pengetahuan berlangsung secara otomatis setiap 5 menit.
2	Saat Pembelajaran	<ul style="list-style-type: none"> - Selama proses pembelajaran berlangsung, apabila penjelasan yang disampaikan oleh Pembelajar (Guru/Dosen/<i>Trainer</i>) dirasa kurang lengkap, setiap Pembelajar (siswa/mahasiswa/peserta) dapat melakukan interaksi dengan <i>chatbot</i>. - <i>Chatbot</i> dapat menjelaskan istilah atau materi yang terlewat untuk disampaikan. - Pembelajar (siswa/mahasiswa/peserta) bisa mendapatkan rekomendasi rangkuman materi ataupun alur belajar sesuai dengan materi yang dipelajari. Harapannya, pengguna dapat dengan mudah menyusun strategi belajar sesuai dengan karakteristiknya.
3	Pasca Pembelajaran	<ul style="list-style-type: none"> - Pembelajar (siswa/mahasiswa/peserta) menginputkan perintah untuk membuat soal pengayaan materi yang telah diperoleh. Soal yang dihasilkan akan berbeda-beda setiap pengguna. - Pembelajar (siswa/mahasiswa/peserta) dapat memperoleh rekomendasi materi lanjutan. - Pembelajar (Guru/Dosen/<i>Trainer</i>) dapat memperoleh pengetahuan terkait materi yang perlu disediakan untuk pertemuan selanjutnya. Tentunya, desain materi hanya dijadikan sebagai referensi agar sesuai dengan tujuan pembelajaran.

4.1.4. Prototipe

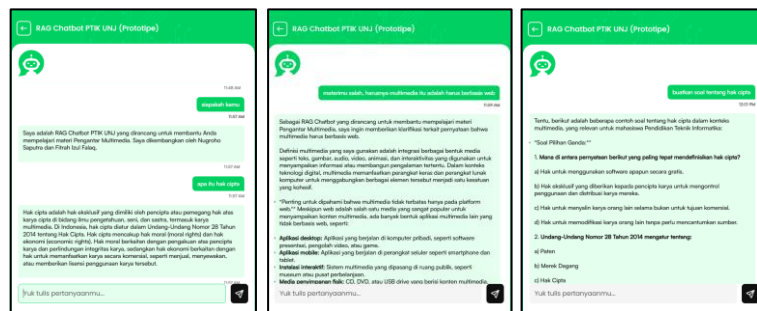
Setelah melalui tahapan *ideate*, selanjutnya adalah merealisasikan rancangan menjadi sebuah media versi uji coba atau prototipe. Adapun hasil tampilan dan fitur yang dihasilkan dapat digambarkan sebagai berikut, dapat dilihat pada gambar 4.5, 4.6 dan 4.7.



Gambar 4.5. Splash Screen RAG Chatbot



Gambar 4.6. Tampilan Perbandingan *Chatbot* Versi Mobile dan Desktop



Gambar 4.7. Implementasi Proses Chatting (Menjawab Pertanyaan, Klarifikasi, dan Membuat Soal)

4.1.5. Test

Pada tahap *test*, peneliti menggunakan 2 metode uji untuk melakukan pengukuran kualitas media pembelajaran *chatbot* yang dihasilkan. Pertama, uji fungsionalitas menggunakan metode *black box*. Kedua, uji tingkat kegunaan menggunakan metode SUS. Adapun hasil pengujian didapatkan sebagai berikut:

1) Uji Fungsionalitas Menggunakan Metode *Black Box*

Pengujian dilakukan oleh pengembang secara mandiri untuk memeriksa kesiapan *chatbot* bagi pengguna. Tahap ini, peneliti melakukan pengujian fungsionalitas *chatbot* untuk mengukur tingkat keberhasilan fitur. Tujuannya adalah untuk menganalisis kendala yang mungkin terjadi. Skenario uji mengacu pada aspek respons pertanyaan, kemampuan klarifikasi konteks, kualitas umpan balik, dan kemampuan evaluasi pembelajaran berdasarkan materi. Hasil pengujian *black box* dapat dilihat pada tabel 4.3.

Tabel 4.3. Tabel Hasil Pengujian *Black Box*

No	Skenario Uji	Input	Output yang Diharapkan	Output yang Diperoleh	Status	Skor
1	Pencarian definisi dasar	Apa itu multimedia?	Menjelaskan bahwa multimedia adalah penggunaan komputer untuk menyajikan teks, gambar, audio, video, animasi secara terintegrasi	<i>Chatbot</i> memberikan definisi multimedia sesuai dengan materi yang telah diinput sebelumnya.	Lulus	1
2	Pencarian konsep format gambar	Jelaskan perbedaan <i>bitmap</i> dan vektor	Jawaban membandingkan <i>bitmap</i> (berbasis piksel) dan vektor (berbasis objek matematis)	<i>Chatbot</i> memberikan penjelasan <i>bitmap</i> dan <i>vector</i> secara detail dan disertai dengan contohnya	Lulus	1
3	Pencarian elemen multimedia	Sebutkan komponen multimedia	Menyebutkan teks, gambar, audio, video, animasi, interaktivitas	<i>Chatbot</i> memberikan penjelasan sumber materi kemudian menyebutkan komponen multimedia dilengkapi dengan definisi singkat dari masing-masing komponen	Lulus	1
4	Evaluasi pembelajaran	Berikan kuis tentang elemen multimedia	Undang-undang yang mengatur tentang Hak Cipta di Indonesia adalah... a) UU ITE b) UU Sisdiknas c) UU Hak Cipta Nomor 28 Tahun 2014 d) UU Paten	<i>Chatbot</i> mampu menjawab soal dengan benar	Lulus	1
5	Rekomendasi belajar lanjutan	Apa materi selanjutnya	Rekomendasi topik berikutnya dalam modul Pengantar Multimedia	<i>Chatbot</i> menjelaskan pengetahuan yang dimiliki, kemudian memberikan rekomendasi secara	Lulus	1

		yang harus saya pelajari?		runtut berdasarkan materi yang telah diinputkan sebelumnya		
6	Pencarian berdasarkan perangkat lunak	Sebutkan contoh <i>software</i> pengolah grafis	Menyebutkan <i>CorelDraw</i> , <i>Adobe Photoshop</i> , <i>GIMP</i> , dsb.	<i>Chatbot</i> menyebutkan seluruh contoh <i>software</i> pengolah grafis sesuai dengan materi	Lulus	1
7	Integrasi materi video	Tampilkan video pengantar multimedia	<i>Chatbot</i> memberikan tautan atau menampilkan video pengantar multimedia	<i>Chatbot</i> memilih <i>link</i> video yang memiliki judul dengan kata kunci "pengantar multimedia" dan tidak memberikan <i>link</i> video lain	Lulus	1
8	Koreksi ejaan	Apa itu multimedia?	<i>Chatbot</i> mengoreksi ke 'multimedia' dan memberi jawaban benar	<i>Chatbot</i> memberikan penjelasan terkait multimedia tanpa memberikan informasi koreksi kata yang tepat	Lulus	1
9	Konteks ganda	Apa itu animasi, dan sebutkan contohnya	Menjelaskan definisi animasi + contoh animasi	<i>Chatbot</i> mampu menjawab secara deskriptif	Lulus	1
10	Respons terhadap pertanyaan di luar materi	Apa ibukota Thailand?	Menjawab Bangkok + memberi pengingat bahwa topik utama adalah multimedia	<i>Chatbot</i> memberikan informasi bahwa tidak bisa menjawab karena di luar konteks materi yang telah diberikan sebelumnya	Lulus	1
11	Penanganan bahasa campuran	"What is multimedia iku?"	<i>Chatbot</i> tetap memahami maksud pertanyaan dan menjawab dengan benar	<i>Chatbot</i> memberikan penjelasan terkait multimedia tanpa memberikan informasi penggunaan Bahasa yang berbeda-beda	Lulus	1
12	Penyediaan sumber belajar tambahan	Rekomendasi buku tentang multimedia	<i>Chatbot</i> memberikan daftar buku atau referensi terpercaya terkait multimedia yang telah diberikan oleh dosen	<i>Chatbot</i> mengubah format daftar pustaka pada materi menjadi deskripsi rekomendasi yang merangkum daftar Pustaka materi	Lulus	1
13	Pencarian materi berdasarkan kata kunci	Materi hak cipta	<i>Chatbot</i> menampilkan ringkasan materi animasi dan tautan pembelajaran lebih lanjut	<i>Chatbot</i> memberikan balasan rangkuman materi "hak cipta"	Lulus	1
14	Pemberian kuis singkat	Beri saya kuis 5 soal tentang multimedia	<i>Chatbot</i> memberikan 5 soal yang relevan dengan materi Pengantar Multimedia, dengan tingkat kesulitan bervariasi	<i>Chatbot</i> mampu menghasilkan soal berdasarkan materi	Lulus	1
15	Pemberian umpan balik	Multimedia seharusnya adalah berbasis web	<i>Chatbot</i> memberikan penjelasan yang benar	<i>Chatbot</i> memberikan klarifikasi dan melanjutkan dengan memberikan penjelasan dan alasan yang benar sesuai materi.	Lulus	1
Skenario Lolos Uji					15/15	
Presentase Lolos Uji					100%	

Pengujian dilakukan dengan 15 skenario uji dengan memberikan inputan dan memeriksa keluaran respons yang dihasilkan. Berdasarkan hasil pengujian *black box*, total fungsi yang berhasil sejumlah 15 dari total 15 skenario uji. Sehingga mendapatkan tingkat presentase keberhasilan sebesar 100%. Dapat disimpulkan bahwa RAG *Chatbot* dapat berfungsi dengan baik.

2) Uji Tingkat Kegunaan Menggunakan Metode *System Usability Scale* (SUS)

Pengujian SUS dilakukan kepada 20 responden mahasiswa PTIK yang memiliki konsentrasi pada bidang Multimedia. Adapun usia 19-21 tahun. Rata-rata mahasiswa terbiasa menggunakan layanan *chatbot* dengan *generative ai*. Proses pengumpulan data dilakukan selama 2 hari. Tahapannya, responden mencoba *chatbot* terlebih dahulu kemudian mengisi kuesioner. Hasil pengolahan data responden dapat dilihat pada Tabel 4.4.

Tabel 4.4. Tabel Hasil Pengolahan Data Responden

Q1	Skor Hasil Hitung									Jumlah	Nilai (Jumlah x 2.5)
	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10		
3	4	4	4	3	3	3	3	3	3	33	83
4	4	4	3	3	2	4	3	4	2	33	83
2	2	3	4	2	2	3	1	4	4	27	68
3	4	4	4	3	4	4	4	4	4	38	95
4	4	4	0	3	3	4	0	4	4	30	75
2	2	4	3	2	2	3	3	3	0	24	60
1	3	3	4	2	4	4	4	4	4	33	83
1	4	4	4	3	4	4	4	2	4	34	85
1	3	3	4	1	3	3	3	2	2	25	63
1	2	3	4	2	2	4	4	2	4	28	70
2	4	4	4	4	3	3	4	4	3	35	88
2	3	3	3	2	2	3	3	2	3	26	65

4	1	3	0	3	1	4	0	3	2	21	53
0	1	4	4	4	1	4	4	1	4	27	68
3	1	2	4	2	3	3	3	2	4	27	68
2	3	3	3	2	2	3	2	2	2	24	60
3	3	4	0	3	3	1	2	2	4	25	63
2	4	3	4	1	3	3	3	4	4	31	78
3	4	4	4	4	4	4	4	4	4	39	98
2	4	4	4	3	3	3	3	3	3	32	80
Skor Rata-rata (Hasil Akhir)											74

Berdasarkan hasil pengujian tersebut, skor rata-rata SUS yang didapatkan adalah sebesar 74. Selanjutnya dilakukan interpretasi dengan menggunakan tabel klasifikasi SUS. Hasil interpretasi pengukuran tingkat kegunaan menggunakan metode SUS dapat dilihat pada tabel 4.5.

Tabel 4.5. Interpretasi Pengukuran Tingkat Kegunaan Menggunakan Metode System Usability Scale (SUS)

Kriteria	Skor	Interpretasi
SUS Score	74	Nilai di atas rata-rata
Grade	B	Kualitas sistem baik, namun masih ada ruang untuk peningkatan menuju <i>grade A</i>
Adjective	Good	Pengguna menilai sistem cukup baik dan nyaman digunakan
Percentile Range	70 – 79	Skor berada di kuartil atas
Acceptable	Acceptable	Tingkat kelayakan penggunaan dapat diterima untuk rilis dan diimplementasikan dalam penggunaan nyata
NPS	Passive	Pengguna cukup puas, tetapi tidak semua akan secara aktif merekomendasikan sistem ke orang lain

4.2. Analisis Pengembangan Produk

Metode design thinking terbukti efektif dalam pengembangan Retrieval-Augmented Generation (RAG) Chatbot melalui tahapan empathize, define, ideate, prototype, dan test, sehingga menghasilkan media yang selaras dengan kebutuhan serta karakteristik pengguna. Pendekatan ini dinilai tepat untuk pengembangan aplikasi atau layanan berbasis web yang mengutamakan kecepatan, ketepatan, dan efektivitas penggunaan. Berdasarkan hasil pengembangan dan evaluasi, RAG Chatbot dinyatakan layak dan dapat diterima untuk mendukung personalized learning, dengan tingkat kenyamanan dan kepuasan pengguna yang baik, meskipun belum seluruh pengguna secara aktif merekomendasikannya. Selain itu, hasil kuesioner responden menunjukkan adanya beberapa rekomendasi perbaikan fitur sebagai bahan penyempurnaan lebih lanjut, sebagaimana dirangkum pada Tabel 4.6.

Tabel 4.6. Analisis Pengembangan Produk

Poin Hasil Pengembangan	Analisis
Penggunaan metode <i>design thinking</i> dalam pengembangan Chatbot dengan Retrieval-Augmented Generation (RAG)	Metode cocok digunakan karena sangat mendedepankan kebutuhan dan karakteristik pengguna. Selain itu cocok untuk pengembangan yang menekankan ada kecepatan, ketepatan dan efektivitas. Proses ini sangat cocok untuk menyediakan <i>chatbot</i> yang memang menekankan pada ciri khas dari calon pengguna.
RAG Chatbot materi Pengantar Multimedia menggunakan <i>knowledge base</i> dan LLM dari Google	Efektif digunakan untuk skala kecil atau dalam proses pembelajaran. Berdasarkan hasil uji coba, layanan LLM Gemini versi gratis dapat digunakan antara 20 – 50 pengguna dalam waktu bersamaan. Namun, tidak direkomendasikan untuk penggunaan secara luas seperti bisnis yang membutuhkan tingkat kebutuhan yang tinggi. Peningkatannya dapat memperhatikan akses token yang tersedia.
Platform Chatbot berbasis web	Efektif karena memudahkan untuk diakses melalui beragam perangkat, baik desktop maupun <i>mobile</i> . Hal yang perlu diperhatikan adalah desain <i>website</i> yang harus lebih <i>responsive</i> .
Penggunaan dalam pembelajaran untuk mendukung <i>personalized learning</i> .	Cocok digunakan sebagai suplemen pembelajaran di mana <i>chatbot</i> diposisikan sebagai asisten belajar yang telah didesain dengan materi yang sesuai kurikulum ataupun tujuan pembelajaran.

5. Kesimpulan dan Saran

Penelitian ini berhasil mengembangkan chatbot berbasis Retrieval-Augmented Generation (RAG) untuk mendukung personalized learning dengan menggunakan metode design thinking yang meliputi tahap empathize, define, ideate, prototype, dan test. Pengembangan dilakukan dengan teknologi yang efisien guna menekan kebutuhan sumber daya serta menyesuaikan kapasitas dan skala pengguna, sementara layanan Large Language Model (LLM) memanfaatkan platform Google untuk menjamin kualitas respons yang baik dengan biaya terjangkau. Hasil pengujian fungsionalitas melalui black-box testing memperoleh skor 15/15 (100%),

menunjukkan seluruh fitur berjalan sesuai spesifikasi, sedangkan pengujian kegunaan menggunakan System Usability Scale (SUS) menghasilkan skor 74 (kategori B) yang berada di atas rata-rata, sehingga chatbot dinyatakan layak dan dapat diterima untuk mendukung pembelajaran terpersonalisasi. Meskipun demikian, peningkatan fitur dan pengujian dalam skala yang lebih luas tetap disarankan agar tingkat kegunaan dapat meningkat ke kategori A serta mendorong pengguna untuk lebih aktif merekomendasikan sistem.

Daftar Pustaka

- Ammann, P., & Offutt, J. (2008). *Introduction to software testing*. Cambridge: Cambridge University Press.
- Arum, S. (2023). Karakteristik Generasi Z. *Accounting Student Research Journal*, 2(1), 59–72.
- Asai, A., Wu, Z., Wang, Y., Sil, A., & Hajishirzi, H. (2024). Self-RAG: Learning to retrieve, generate, and critique through self-reflection. *Twelfth International Conference on Learning Representations (ICLR)*.
- Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual SUS scores mean: Adding an adjective rating. *Journal of Usability Studies*, 4(3), 114–123.
- Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability Evaluation in Industry* (pp. 189–194). London: Taylor & Francis.
- Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies*, 8(2), 29–40.
- Interaction Design Foundation. (2025). *What is design thinking?* Diakses dari <https://www.interaction-design.org/literature/topics/design-thinking>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), 1–17. <https://doi.org/10.1186/s41239-023-00426-1>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459–9474.
- Major, L., Francis, G. A., & Tsapali, M. (2021). The effectiveness of technology-supported personalised learning in low- and middle-income countries: A meta-analysis. *British Journal of Educational Technology*, 52(5), 1935–1964. <https://doi.org/10.1111/bjet.13116>
- Nur, H., Nugroho, I. S., Saputra, M. R. E., Suhaemi, N., & Saifudin, A. (2020). Pengujian black box pada aplikasi sistem pengarsipan surat menggunakan teknik equivalence partitions. *Jurnal Teknologi Sistem Informasi dan Aplikasi*, 3(2), 76. <https://doi.org/10.32493/jtsi.v3i2.4692>
- Okonkwo, C. W., & Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Parlika, R., Nisaa', T. A., Ningrum, S. M., & Haque, B. A. (2020). Studi literatur kekurangan dan kelebihan pengujian black box. *Teknomatika*, 10(2), 131–140.
- Procko, T. T., & Ochoa, O. (2024). Graph Retrieval-Augmented Generation for large language models: A survey. *Proceedings - 2024 Conference on AI, Science, Engineering, and Technology (AIxSET)* (pp. 166–169). IEEE. <https://doi.org/10.1109/AIxSET62544.2024.00030>
- Razzouk, R., & Shute, V. (2012). What is design thinking and why is it important? *Review of Educational Research*, 82(3), 330–348. <https://doi.org/10.3102/0034654312457429>
- Sauro, J. (2018). *5 ways to interpret a SUS score*. MeasuringU. Diakses dari <https://measuringu.com/interpret-sus-score/>
- Shuster, K., Poff, S., Chen, M., Kiela, D., & Weston, J. (2021). Retrieval augmentation reduces hallucination in conversation. *Findings of the Association for Computational Linguistics: EMNLP 2021* (pp. 3784–3803). <https://doi.org/10.18653/v1/2021.findings-emnlp.320>
- Tullis, T. S., & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. *Proceedings of the Usability Professionals Association Conference (UPA)*, 1–12.
- UNESCO. (2023). *Global education monitoring report 2023, Southeast Asia: Technology in education: A tool on whose terms?* Paris: UNESCO.