



Implementasi Sistem Pembuatan Soal Otomatisasi Pembelajaran Pendidikan Agama Islam Dengan Menggunakan Langchain Dan Llm Berbasis Gemini

Irgi Fahrezy, Nazruddin Safaat Harahap*, Fitri Wulandari, Surya Agustian

Teknik Informatika, Fakultas Sains Dan Teknologi, Universitas Islam Negeri Sultan Syarif Kasim Riau Pekanbaru
email : 12150113234@students.uin-suska.ac.id¹⁾, nazruddin.safaat@uin-suska.ac.id²⁾ fitri.wulandari@uin-suska.ac.id³⁾
surya.agustian@uin-suska.ac.id⁴⁾

Abstrak– Penelitian ini mengkaji implementasi sistem otomatisasi pembuatan soal dalam konteks Pendidikan Agama Islam (PAI) menggunakan teknologi LangChain dan Large Language Model (LLM) berbasis Gemini. Metodologi penelitian meliputi pengumpulan data, analisis kebutuhan, perancangan arsitektur sistem, implementasi, serta validasi melalui pengujian kotak hitam dan evaluasi pakar. Sistem dirancang untuk menghasilkan soal secara otomatis yang sesuai dengan materi pembelajaran dan tingkatan kognitif dalam taksonomi Bloom, mulai dari pengetahuan faktual hingga evaluasi dan kreasi. Proses pengujian melibatkan enam evaluator ahli dengan latar belakang pendidikan dan pengalaman di bidang PAI. Hasil menunjukkan bahwa sistem berhasil menghasilkan pertanyaan berkualitas tinggi dengan rata-rata tingkat persetujuan sebesar 96,89%. Diferensiasi nilai menunjukkan hasil tertinggi pada pertanyaan interpretasi teoretis (100%) dan terendah pada analisis kritis (30%), yang mengindikasikan variasi kemampuan sistem dalam mengakomodasi spektrum kognitif secara menyeluruh. Penelitian ini menunjukkan bahwa integrasi kecerdasan buatan dalam pembuatan soal sangat mungkin dilakukan dan memberikan hasil yang signifikan, meskipun masih terdapat tantangan dalam menghasilkan soal yang mendorong pemikiran tingkat tinggi dan autentik.

Kata Kunci: Otomatisasi pembuatan soal, pendidikan agama Islam, kecerdasan buatan

Abstract– *This study examines the implementation of an automated question generation system in the context of Islamic Religious Education (PAI) using LangChain technology and a Gemini-based Large Language Model (LLM). The research methodology includes data collection, needs analysis, system architecture design, implementation, and validation through black-box testing and expert evaluation. The system is designed to automatically generate questions aligned with instructional content and the cognitive levels of Bloom's Taxonomy, ranging from factual knowledge to evaluation and creation. The testing process involved six expert evaluators with educational backgrounds and experience in PAI. Results indicate that the system successfully produced high-quality questions, with an average approval rate of 96.89%. Differentiated scores revealed the highest performance in theoretical interpretation questions (100%) and the lowest in critical analysis (30%), indicating varying system capabilities across the cognitive spectrum. This study demonstrates that integrating artificial intelligence into question generation is highly feasible and yields significant outcomes, although challenges remain in producing questions that stimulate higher-order and authentic thinking.*

Keywords: Automated question generation, Islamic religious education, artificial intelligence

1. PENDAHULUAN

Rekonfigurasi digital yang terjadi dalam domain pendidikan telah menginduksi perubahan fundamental pada arsitektur pembelajaran tradisional. Penetrasi teknologi bukan sekadar menginisiasi restrukturisasi metode transmisi pengetahuan, melainkan mentransformasi paradigma evaluatif yang menjadi fondasi dalam proses edukatif. Elemen teknologi kini berevolusi menjadi komponen vital, menggantikan posisinya sebagai instrumen fakultatif, dalam intensifikasi efisiensi pembelajaran. Sinergi antara teknologi informasional dan komunikatif, yang teraktualisasi dalam entitas multimedia, telah mengintensifikasi relevansi reorientasi digital dalam infrastruktur edukatif modern, meliputi elaborasi sistem informasional akademik, pengelolaan pembelajaran, hingga repositori kognisi terintegrasi[1]

Metamorfosis digital dalam spektrum pendidikan telah mengkatalisis emergensi beragam terobosan inovatif, mulai dari pembelajaran berbasis multimedia, platform e-learning, hingga instrumentasi evaluatif berbasis teknologi seperti *Computer-based Test (CBT)* dan *Computer-assisted Test (CAT)*. Metodologi *CAT*, yang diadaptasi dari implementasi global, dikonfigurasi untuk kongruen dengan kontekstualitas ketenagakerjaan nasional Indonesia. Konstruksi sistem ini diorientasikan pada maksimalisasi proses selektif melalui mekanisme evaluasi terkomputerisasi. Transformasi teknologi memanasifestasikan bukan hanya progresivitas gradual, tetapi restrukturisasi komprehensif dalam modalitas pembelajaran, pola operasional, dan dinamika sosial. Kecerdasan Artifisial (AI) menjadi faktor determinan dalam evolusi ini, mereferensikan pada disiplin yang berorientasi pada kreasi sistem dengan kapasitas kognitif ekuivalen dengan manusia. Implementasi *AI* telah menunjukkan efektivitas dalam resolusi problematika multidimensional, termasuk dalam sektor edukasi. Dalam ranah pemrosesan linguistik natural, aplikasi *AI* telah menghasilkan asisten virtual dengan kemampuan interpretasi dan generasi bahasa manusia secara autentik [2]

Kemajuan teknologi *AI* seperti pembelajaran mesin, deep learning, dan Pemrosesan Bahasa Natural (*NLP*) berkembang dengan progresivitas eksponensial. Model Bahasa Besar (*Large Language Model/LLM*), sebagai manifestasi *AI* kontemporer, merupakan ekspansi dari arsitektur jaringan neural konvensional. *LLM* memiliki kapasitas untuk memproses dan menghasilkan korpus tekstual masif setelah dilatih dengan dataset ekstensif. Kapabilitas ini memfasilitasi aplikasi *LLM* dalam diversitas operasi linguistik, mulai dari generasi konten, translasi interlingual, hingga analisis sentimen. menunjukkan bahwa pemanfaatan *Gemini API*, sebagai salah satu implementasi *LLM*, dalam aplikasi pameran berbasis web mampu menghasilkan deskripsi karya seni secara otomatis dengan struktur bahasa yang logis, deskriptif,





dan relevan secara kontekstual. Temuan ini mempertegas potensi *LLM* dalam mendukung otomatisasi linguistik yang tidak hanya fungsional secara teknis, tetapi juga bernilai tinggi secara semantik dan aplikatif di berbagai domain.[3]

Linguistik komputer mencakup pemahaman, pemrosesan, dan pembuatan bahasa melalui teknologi pemrosesan bahasa alami (*Natural Language Processing/NLP*), yang memungkinkan komputer bekerja dengan teks lisan maupun tulisan untuk mengekstrak makna dan menghasilkan teks yang alami serta tata bahasa yang benar. Dengan *NLP*, komputer tidak hanya memahami bahasa pemrograman formal seperti *Java*, tetapi juga mampu memproses bahasa manusia seperti Bahasa Indonesia dan Inggris. Teknologi ini digunakan dalam berbagai aplikasi seperti *chatbot*, penerjemah otomatis, dan analisis sentimen, serta telah menjadi fondasi penting dalam transformasi digital di berbagai bidang, termasuk pendidikan dan bisnis.[4]

Dalam konteks evaluasi teknologis, salah satu permasalahan utama pada sistem konvensional adalah adanya diskrepansi antara kompleksitas soal dengan kapasitas kognitif peserta. Pemilihan soal secara arbitrer dari bank soal tanpa mempertimbangkan dimensi kompleksitas dapat menurunkan validitas instrumen evaluatif. Untuk mengatasi hal ini, *Computer Adaptive Testing (CAT)* hadir sebagai solusi inovatif dengan menyajikan soal yang tingkat kesulitannya terkalibrasi secara dinamis berdasarkan analisis algoritmik terhadap respons peserta. Sistem ini memungkinkan proses evaluasi yang lebih akurat dan efisien karena dapat menyesuaikan soal secara real-time dengan kemampuan individu. Temuan Astino dan Sabandar (2023) mendukung hal tersebut, di mana mereka mengembangkan dan menerapkan sistem *Computer Assisted Test (CAT)* berbasis website guna mengelola ujian digital secara lebih adaptif. Hasilnya menunjukkan bahwa penerapan *CAT* mampu meningkatkan efektivitas, efisiensi, serta fleksibilitas pelaksanaan ujian, sekaligus memberikan pengalaman yang lebih relevan dan responsif terhadap kebutuhan peserta.[5]

Evolusi *AI* telah menginisiasi konseptualisasi otomatisasi dalam generasi pertanyaan dari sumber literatur dengan memanfaatkan teknologi *LangChain*, *FAISS*, dan *LLM*. *LLM* dikarakterisasi oleh kapabilitas pembelajaran kontekstual, menghasilkan output tekstual berbasis input yang diberikan. Atribut ini memungkinkan *LLM* untuk menghasilkan respons yang konsisten dan relevan, sangat aplikabel untuk implementasi interaktif. yang menunjukkan bahwa penerapan sistem *question answering* berbasis *LLM* dapat memberikan jawaban yang tepat, cepat, dan akurat terhadap pertanyaan yang diajukan dalam bahasa alami, serta memiliki potensi besar dalam pengembangan sistem tanya jawab berbasis teks yang mendukung proses pencarian informasi dan pembelajaran otomatis secara efektif [6]

LangChain sebagai kerangka pengembangan aplikasi *LLM* memfasilitasi interkoneksi dengan sumber data eksternal dan segmentasi korpus tekstual menjadi unit-unit manageable, meskipun limitasi eksistensinya terletak pada kompatibilitas data tekstual semata. Kemampuan ini sangat berguna dalam mengelola dokumen besar dan kompleks, seperti materi pembelajaran atau teks hukum, dengan tetap menjaga konteks dan keterkaitan informasi. yang dalam penelitiannya mengenai otomatisasi fatwa hukum Islam berbasis *AI*, menunjukkan bahwa pemrosesan bahasa alami dan sistem berbasis *LLM* memungkinkan pengambilan keputusan hukum yang kompleks secara efisien dengan mempertimbangkan sumber-sumber rujukan syariah. Pendekatan ini memperkuat peran teknologi dalam mendukung analisis teks keagamaan dan hukum secara sistematis dan terstruktur, sebagaimana yang juga difasilitasi oleh arsitektur *LangChain* dalam konteks pengembangan aplikasi cerdas berbasis teks [7]

Metodologi konvensional dalam sintesis literatur dan formulasi pertanyaan masih melibatkan intervensi manual substansial, mengkonsumsi waktu yang signifikan untuk mengekstraksi informasi dari korpus dokumen yang voluminus. Inkonsistensi dalam interpretasi antar departemen juga menegaskan adanya urgensi terhadap solusi yang mampu menghadirkan efisiensi terukur dengan hasil keluaran yang terstandarisasi. Untuk mencapai implementasi yang optimal, dibutuhkan pendekatan otomatisasi sekuensial terhadap dokumen digital guna mempercepat proses kerja dan meminimalkan kesalahan interpretatif. yang menunjukkan bahwa sistem berbasis kecerdasan buatan mampu secara otomatis menganalisis teks hukum Islam, mengurangi ketergantungan pada input manual, serta menghasilkan keluaran yang konsisten dan akurat. Pendekatan ini membuktikan bahwa integrasi *AI* dalam pengolahan dokumen tidak hanya meningkatkan efisiensi operasional, tetapi juga memperkuat validitas hasil dalam konteks pendidikan, hukum, dan penelitian ilmiah lainnya [8]

Dalam konteks Pendidikan Agama Islam (PAI), material pembelajaran tradisional seringkali tidak mengakomodasi kebutuhan pengalaman edukatif yang interaktif dan kontekstual. Kecerdasan buatan (*AI*) menawarkan potensi besar untuk menciptakan media pembelajaran yang adaptif dan personalisasi, memungkinkan materi disesuaikan dengan gaya belajar dan tingkat pemahaman peserta didik. Di era transformasi digital saat ini, infiltrasi teknologi dalam sektor edukasi memfasilitasi implementasi *AI* untuk mengoptimalkan kualitas pembelajaran dan mengatasi berbagai limitasi sistemik dari metode konvensional. yang menunjukkan bahwa pengembangan media pembelajaran berbasis *AI* dalam konteks PAI tidak hanya meningkatkan partisipasi siswa, tetapi juga memperdalam pemahaman nilai-nilai keagamaan melalui pendekatan yang lebih menarik dan kontekstual. Dengan demikian, *AI* bukan sekadar alat bantu teknis, melainkan sarana transformatif dalam mendukung pembelajaran agama yang lebih efektif, interaktif, dan relevan dengan perkembangan zaman [9]

Chatbot merupakan aplikasi yang semakin populer karena memanfaatkan teknologi *Kecerdasan Buatan (AI)*, Pembelajaran Mesin, Pembelajaran Mendalam, dan Pemrosesan Bahasa Alami (*NLP*). *AI* telah memungkinkan sistem komputer memahami interaksi suara dan teks secara lebih efisien, sehingga mempercepat respons dan meningkatkan pengalaman pengguna. Perkembangan terkini menunjukkan integrasi teknologi *AI* yang semakin canggih, seperti model



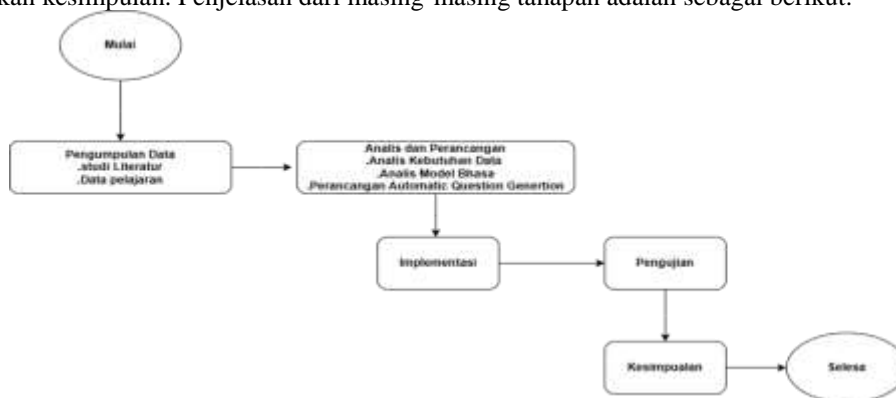
transformer dan arsitektur GPT dari OpenAI, yang mendorong kemampuan chatbot untuk berinteraksi secara real-time, personal, dan aman di berbagai bidang seperti kesehatan, pemerintahan, dan pendidikan. Evolusi tema penelitian juga mencerminkan pergeseran dari fokus dasar seperti AI dan NLP menuju aplikasi praktis seperti e-commerce dan interaksi manusia-komputer. Secara keseluruhan, AI menjadi fondasi utama dalam membangun chatbot yang lebih kontekstual, adaptif, dan berpusat pada pengguna, seiring meningkatnya kebutuhan akan sistem percakapan yang cerdas dan relevan di berbagai sektor[10]

Investigasi ini berfokus pada implementasi sistem otomatisasi dalam konstruksi soal pembelajaran Pendidikan Agama Islam (PAI) menggunakan *LangChain* dan *LLM* berbasis *Gemini*. Implementasi ini diharapkan dapat memitigasi problematika dalam formulasi soal evaluatif yang adaptif, berkualitas, dan sinkron dengan standar pendidikan nasional. Kontribusi penelitian ini diproyeksikan mampu mendukung pengembangan sistem evaluatif berbasis komputasi yang inovatif dan responsif terhadap kebutuhan peserta didik, serta menjadi fondasi bagi eksplorasi lebih lanjut dalam pemanfaatan kecerdasan artifisial di bidang pendidikan. Yang menekankan bahwa integrasi kecerdasan buatan dalam pembelajaran PAI tidak hanya berperan sebagai alat bantu teknis, tetapi juga sebagai mitra guru dalam menciptakan proses pembelajaran yang lebih bermakna, adaptif, dan personal. Integrasi ini menunjukkan bahwa *AI* dapat menjadi katalisator dalam transformasi pembelajaran agama yang kontekstual, dinamis, dan selaras dengan tantangan serta kebutuhan Pendidikan [11]

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui beberapa tahapan sistematis untuk mengembangkan sistem pembuatan soal otomatisasi pembelajaran Pendidikan Agama Islam [12]. Adapun alur penelitian yang diimplementasikan ditunjukkan pada Gambar 1. Penelitian ini dilakukan melalui 5 tahapan utama: pengumpulan data, analisis dan perancangan, implementasi, pengujian, dan penarikan kesimpulan. Penjelasan dari masing-masing tahapan adalah sebagai berikut:



Gambar 1. Metodologi Penelitian

2.2 Pengumpulan Data

a Studi Literatur

Tinjauan pustaka yang dilakukan peneliti mencakup berbagai sumber seperti jurnal, artikel, tinjauan sistematis, dan website yang berkaitan dengan topik penelitian. Topik yang dieksplorasi dalam proses ini meliputi sistem pembuatan pertanyaan otomatis, *LangChain*, model bahasa skala besar (*LLM*), pengujian penilaian komputer (*CAT*), pemrosesan bahasa alami (*NLP*), *Streamlit*, dan materi PAI yang menjadi bahan dalam penelitian ini [13]

b Data Pembelajaran

Peneliti memperoleh materi pelajaran Pendidikan Agama Islam (PAI) dalam format PDF dari program studi Universitas Suska Riau sebagai sumber utama untuk pengembangan system [14]

2.3 Analisis dan Perancangan

a Analisa Kebutuhan Data

Tahapan ini bertujuan untuk mengidentifikasi data yang diperlukan untuk penelitian. Dalam penelitian ini, peneliti mengumpulkan materi Pendidikan Agama Islam (PAI) dalam format PDF yang akan diolah oleh sistem untuk menghasilkan pertanyaan secara otomatis [7]

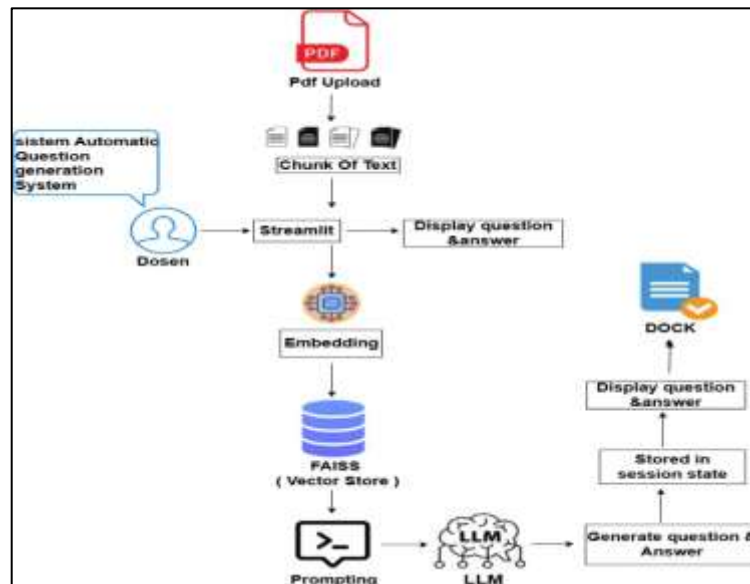
b Analisa Model Bahasa

Analisis model bahasa merupakan langkah untuk menentukan model bahasa yang akan digunakan dalam

penelitian ini. Peneliti memilih model bahasa dari platform Google DeepMind yaitu model Gemini, yang dirancang untuk memproses dan menghasilkan teks dalam bahasa alami dengan kemampuan pemahaman konteks yang tinggi [15]

c Perancangan Automatic Question Generation System

Sistem pembuatan pertanyaan otomatis dirancang dengan alur yang ditunjukkan pada Gambar 2.



Gambar 1 Alur Automatic Question Generation Sysytem

Pada gambar di atas, terdapat beberapa proses yang saling berkaitan dalam sistem, dengan penjelasan sebagai berikut:

1. **Set up:** Proses awal untuk mengatur konfigurasi aplikasi menggunakan Streamlit.
2. **Data Resource:** Sumber data berupa dokumen teks atau materi pelajaran PAI yang menjadi input utama untuk sistem.
3. **Chunk of Text:** Data dari sumber dipecah menjadi bagian-bagian kecil untuk mempermudah pemrosesan.
4. **Embedding:** Proses mengubah potongan teks menjadi representasi numerik menggunakan teknik embedding, yang memungkinkan teks diproses oleh sistem.
5. **FAISS (Vector Store):** Tempat penyimpanan representasi vektor dari teks yang telah diembedding, memfasilitasi pencarian informasi yang relevan secara cepat sebagai knowledge base (basis pengetahuan).
6. **Prompting:** Sistem mengirimkan permintaan atau instruksi (prompt) ke model bahasa besar (LLM) berdasarkan informasi yang relevan.
7. **Large Language Model (LLM):** Model bahasa besar seperti Gemini memproses prompt untuk menghasilkan pertanyaan dan jawaban otomatis.
8. **Generate Question & Answer:** Tahap di mana pertanyaan dan jawaban dihasilkan berdasarkan input dan pemrosesan sebelumnya.
9. **Stored in Session State:** Hasil pertanyaan dan jawaban disimpan sementara pada session state untuk digunakan lebih lanjut.
10. **Display Question & Answer:** Pertanyaan dan jawaban yang dihasilkan ditampilkan kepada pengguna melalui antarmuka Streamlit [16]

2.4 Implementasi

a Implementasi Frontend

Frontend merupakan bagian dari sistem yang menyediakan antarmuka visual kepada pengguna. Dalam penelitian ini, peneliti menggunakan Streamlit sebagai antarmuka untuk sistem pembangkitan pertanyaan otomatis yang bertindak sebagai penghubung antara pengguna dan sistem. Pemilihan Streamlit didasarkan pada kemampuannya sebagai pustaka Python sumber terbuka yang memudahkan pembuatan dasbor dan aplikasi web interaktif tanpa memerlukan keterampilan pengembangan frontend yang mendalam [17]

b Implementasi Backend

Backend yang berkualitas harus memiliki performa yang baik. Untuk mengetahui apakah berkinerja baik, perlu

dilakukan pengujian kinerja sistem. Tujuan pengujian kinerja adalah untuk memeriksa apakah sistem dapat menangani pengguna dan permintaan data dalam jumlah besar tanpa mempengaruhi kinerja perangkat lunak. Dalam penelitian ini, peneliti memilih Python sebagai bahasa pemrograman utama untuk mengimplementasikan sistem pembuatan pertanyaan otomatis karena ketersediaan berbagai pustaka yang mendukung [18]

2.5 Pengujian

Pengujian merupakan langkah penting dalam pengembangan sistem untuk memastikan bahwa sistem pembuatan pertanyaan otomatis yang dikembangkan konsisten dengan analisis awal dan berkinerja baik dalam memberikan informasi kepada pengguna. Dua jenis pengujian yang dilakukan dalam penelitian ini adalah pengujian *Scoring Evaluator Langkahain* dan User Acceptance Test (UAT).

a Scoring Evaluator Langchain

Scoring Evaluator LangChain adalah salah satu metode evaluasi yang disediakan oleh framework LangChain untuk menilai kualitas jawaban atau keluaran dari *Large Language Model (LLM)*, seperti ChatGPT atau Gemini. Evaluator ini digunakan untuk memberikan penilaian numerik (skor) terhadap jawaban model berdasarkan kriteria yang telah ditentukan pengguna [19]

b User Acceptance Test (UAT)

User Acceptance Testing (UAT) dilakukan untuk mengevaluasi sistem pembuatan pertanyaan otomatis dan mendapatkan umpan balik langsung dari pengguna. Pengujian tersebut melibatkan enam responden yang merupakan guru, instruktur, pelatih, atau individu yang memiliki keahlian di bidangnya. Umpan balik yang diberikan melalui evaluasi ini dimaksudkan untuk memastikan bahwa keluaran sistem memenuhi kebutuhan dan harapan pengguna [20]

3. HASIL DAN PEMBAHASAN

3.1 Implementasi Sistem Automatic Question Generation

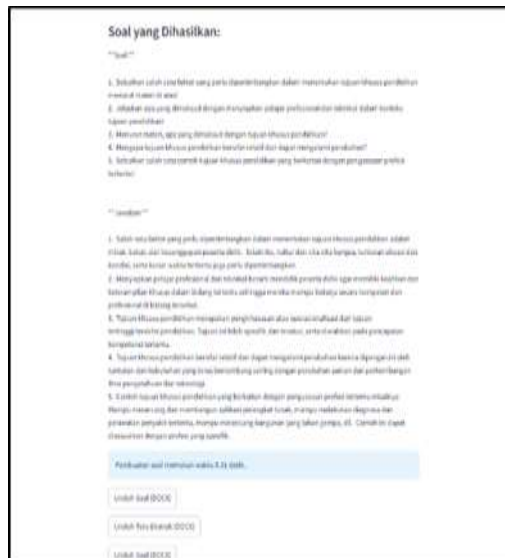
Sistem ini dirancang untuk menghasilkan soal Pendidikan Agama Islam secara otomatis dengan mengintegrasikan FAISS sebagai penyimpanan vektor dan LLM Gemini-2.0-Flash untuk pemrosesan bahasa. Dokumen materi dalam format PDF diolah melalui tiga tahap utama: chunking, embedding, dan pencarian semantik. Tahap chunking membagi teks berdasarkan topik agar tetap koheren secara makna. Selanjutnya, embedding dengan model text-embedding-3-large mengubah teks menjadi vektor multidimensi yang merepresentasikan konteks dan hubungan antar konsep. Vektor ini disimpan dalam FAISS untuk memungkinkan pencarian cepat dan akurat terhadap segmen yang relevan. Antarmuka pengguna dirancang sederhana dengan fitur unggah PDF (maks. 200MB) dan judul “Generator Soal dengan Gemini API”. Sistem ini memadukan teknologi AI dengan prinsip pendidikan Islam untuk menghasilkan soal yang relevan secara kontekstual dan pedagogis.



Gambar 3 Tampilan Generator Soal

Pada gambar 3 menunjukkan antarmuka aplikasi “Generator Soal dengan Gemini API” yang dirancang untuk menghasilkan soal otomatis dari dokumen PDF, khususnya materi Pendidikan Agama Islam. Pengguna dapat mengunggah file PDF (maksimal 200MB), yang kemudian secara otomatis diproses melalui tahapan ekstraksi teks, pemotongan teks menjadi 13 bagian (chunking), konversi teks menjadi vektor (embedding), dan penyimpanan dalam indeks FAISS untuk pencarian

semantik. Setelah proses ini, pengguna dapat menentukan jumlah soal, tingkat kesulitan, dan bahasa soal sebelum menekan tombol “Buat Soal”. Aplikasi juga menyediakan opsi untuk mengunduh hasil ekstraksi teks dalam format DOCX, dengan tampilan antarmuka yang sederhana dan mudah digunakan. Pada gambar 4 Setelah pengguna menekan tombol “Buat Soal”, sistem akan memproses input teks dari dokumen dan menghasilkan daftar soal beserta jawabannya secara otomatis. Pada tampilan hasil, terdapat dua bagian utama, yaitu “Soal” dan “Jawaban”, yang masing-masing ditampilkan secara terstruktur dan rapi.



Gambar 4 Hasil Generator Soal

Pada gambar 4 Setelah pengguna menekan tombol “Buat Soal”, sistem akan memproses input teks dari dokumen dan menghasilkan daftar soal beserta jawabannya secara otomatis. Pada tampilan hasil, terdapat dua bagian utama, yaitu “Soal” dan “Jawaban”, yang masing-masing ditampilkan secara terstruktur dan rapi.

3.2 Arsitektur dan Alur Pemrosesan Sistem

Sistem pembuatan soal otomatis ini dirancang menggunakan pendekatan modular, mengintegrasikan beberapa teknologi kecerdasan buatan seperti LangChain, SentenceTransformer, FAISS, dan LLM (Gemini API). Berikut ini adalah tahapan proses secara terstruktur:

- a Ekstraksi Teks dari Dokumen PDF
Langkah awal dalam sistem adalah mengekstrak teks dari dokumen PDF yang diunggah oleh pengguna. Proses ini menggunakan pustaka PyPDF2 untuk membaca setiap halaman dan mengekstrak isi teksnya secara lengkap. Berdasarkan pengujian dengan dokumen "TUJUAN PENDIDIKAN ISLAM.pdf", proses ini dapat diselesaikan dalam waktu rata-rata 2,01 detik. Hasil ekstraksi juga disimpan dalam format .docx untuk memudahkan distribusi dan dokumentasi.
- b Segmentasi Teks Menggunakan LangChain
Setelah ekstraksi, teks dibagi menjadi beberapa bagian menggunakan modul RecursiveCharacterTextSplitter dari pustaka LangChain. Teknik ini memecah teks menjadi chunk dengan ukuran 500 karakter dan overlap 100 karakter untuk mempertahankan konteks. Proses segmentasi ini memungkinkan sistem melakukan analisis semantik yang lebih presisi pada setiap bagian teks. Dalam pengujian, teks berhasil disegmentasi menjadi 13 bagian dalam waktu 0,02 detik.
- c Embedding Semantik dengan SentenceTransformer
Setiap chunk teks kemudian direpresentasikan dalam bentuk embedding vektor berdimensi tetap menggunakan model all-MiniLM-L6-v2 dari SentenceTransformers. Embedding ini mengkodekan makna semantik dari setiap segmen teks ke dalam bentuk numerik yang dapat dihitung kemiripannya. Waktu rata-rata proses embedding untuk seluruh chunk adalah 8,58 detik.
- d Pembuatan Indeks Pencarian dengan FAISS
Embedding yang dihasilkan selanjutnya digunakan untuk membangun indeks FAISS (Facebook AI Similarity Search) menggunakan metode IndexFlatL2. Struktur ini memungkinkan pencarian semantik yang cepat dan efisien dalam ruang vektor berdimensi tinggi. Indeks ini memungkinkan sistem untuk menemukan chunk teks yang paling relevan berdasarkan kueri atau pertanyaan yang akan dihasilkan. Waktu yang dibutuhkan untuk membangun indeks ini hanya sekitar 0,04 detik.
- e Generasi Soal dengan LLM (Gemini API)
Sistem memungkinkan pengguna memilih jumlah soal, tingkat kesulitan (Mudah, Sedang, Sulit), dan bahasa output (Indonesia/English). Setelah parameter dipilih, sistem membuat prompt secara dinamis berdasarkan chunk teks yang paling relevan. Prompt dikirim ke model LLM Gemini 2.0 Flash dari Google melalui `genai.GenerativeModel`, yang

menghasilkan soal dan jawaban dalam format terstruktur. Contoh struktur prompt dalam Bahasa Indonesia.

f Pencocokan Soal ke Konteks Asli (Similarity Search)

Setelah soal dihasilkan, sistem melakukan proses evaluasi kemiripan semantik (similarity search) antara setiap pertanyaan dengan seluruh chunk yang telah disimpan sebelumnya menggunakan FAISS. Hal ini dilakukan untuk menelusuri dan memverifikasi dari bagian mana pertanyaan tersebut kemungkinan besar berasal. Proses ini memberikan transparansi dan validasi pedagogis terhadap hasil soal yang dihasilkan.

g Dokumentasi dan Output Terstruktur

Hasil akhir dari proses ini berupa:

h Soal dan jawaban dalam format dua kolom

- File .docx berisi pertanyaan
- File .docx dan .csv berisi hasil similarity search
- Semua file dapat diunduh langsung oleh pengguna. Pendekatan terstruktur ini membuat sistem relevan digunakan untuk kebutuhan penilaian formatif, sumatif, maupun pembelajaran mandiri, khususnya dalam konteks pendidikan agama Islam.

3.3 Evaluasi Kuantitatif Kualitas Soal

Evaluasi kuantitatif terhadap kualitas soal yang dihasilkan oleh sistem merupakan aspek integral dalam validasi kinerja dan efektivitas pedagogis. Proses evaluasi melibatkan sembilan tenaga pendidik pascasarjana dengan keahlian substantif dalam bidang pendidikan agama Islam, mengimplementasikan metodologi skala Likert lima poin dari "Sangat Tidak Sesuai" (STS) hingga "Sangat Sesuai" (SS) dengan bobot nilai 1 hingga 5. Hasil dari evaluasi kualitas soal dapat dilihat pada tabel berikut.

Tabel 1. Bobot Nilai

Pilihan Jawaban	Keterangan	Bobot Nilai
SS	Sangat Sesuai	5
S	Sesuai	4
N	Netral / Cukup	3
TS	Tidak Sesuai	2
STS	Sangat Tidak Sesuai	1

Berikut ini adalah pertanyaan survei pengguna dan jawaban nya yang sesuai.

Tabel 2. Pertanyaan dan jawaban Kuesioner

No	Soal	Jawaban
1	Sebutkan salah satu rukun nikah yang disebutkan dalam materi di atas!	Sighat (akad yang terdiri dari ijab dan qabul) atau Wali mempelai perempuan.
2	Sebutkan salah satu rukun nikah yang disebutkan dalam materi di atas!	Sighat adalah akad dalam pernikahan yang terdiri dari ijab (pernyataan penyerahan dari pihak perempuan atau walinya) dan qabul (pernyataan penerimaan dari pihak laki-laki).
3	Sebutkan salah satu syarat yang harus dipenuhi oleh seorang wali nikah!	Berakal, baligh, merdeka (bukan budak), atau kesamaan agama.
4	Mengapa seorang wali nikah harus memiliki kesamaan agama dengan mempelai perempuan?	Karena dalam Islam, seorang wali harus seagama dengan mempelai perempuan yang akan dinikahkan. Wali non-muslim tidak sah menikahkan perempuan muslim.
5	Apa yang dimaksud dengan ijab dan qabul dalam akad nikah?	Ijab adalah pernyataan penyerahan dari pihak perempuan (melalui walinya) dan qabul adalah pernyataan penerimaan dari pihak laki-laki. Keduanya merupakan bagian dari akad nikah.
...
30	Wali nikah bertanggung jawab untuk memastikan kemaslahatan pernikahan yang akan dilangsungkan. Diskusikan peran wali dalam memberikan nasihat dan pertimbangan kepada calon mempelai, serta batasan-batasan intervensi wali dalam keputusan pribadi calon mempelai.	Wali nikah bertanggung jawab untuk memberikan nasihat dan pertimbangan kepada calon mempelai mengenai aspek-aspek penting dalam pernikahan, seperti kesiapan mental, ekonomi, dan agama. Wali juga dapat memberikan masukan mengenai calon pasangan, tetapi batasan intervensi wali adalah bahwa wali tidak boleh memaksa wanita untuk menikah dengan orang yang tidak disukai atau menghalangi pernikahan yang sah tanpa alasan yang dibenarkan. Keputusan akhir mengenai pernikahan tetap berada di tangan calon mempelai wanita, dengan tetap mempertimbangkan nasihat dan pertimbangan dari wali.

Tabel 1 menunjukkan rata-rata dari nilai yang diberikan oleh 6 tenaga pendidik yang memiliki pemahaman yang baik

tentang pendidikan agama islam. Hasil dari evaluasi ini adalah sebagai berikut.

Tabel 3. Hasil Evaluasi Kualitas Jawaban

No	Pertanyaan	SS	S	N	TS	STS	Total Nilai	Presentasi Nilai
1	Sebutkan salah satu rukun nikah yang disebutkan dalam materi di atas!	6	0	0	0	0	30	100%
2	Sebutkan salah satu rukun nikah yang disebutkan dalam materi di atas!	6	0	0	0	0	30	100%
3	Sebutkan salah satu syarat yang harus dipenuhi oleh seorang wali nikah!	5	1	0	0	0	30	96,67%
4	Mengapa seorang wali nikah harus memiliki kesamaan agama dengan mempelai perempuan?	5	1	0	0	0	30	96,67%
5	Apa yang dimaksud dengan ijab dan qabul dalam akad nikah?	6	0	0	0	0	30	100%
...
30	Wali nikah bertanggung jawab untuk memastikan kemaslahatan pernikahan yang akan dilangsungkan. Diskusikan peran wali dalam memberikan nasihat dan pertimbangan kepada calon mempelai, serta batasan-batasan intervensi wali dalam keputusan pribadi calon mempelai.	5	1	0	0	0	30	96,67%

Evaluasi terhadap 30 soal menunjukkan variasi kualitas dengan persentase nilai antara 30% hingga 100%. Soal-soal faktual seperti definisi sighthat, syarat wali, dan rukun nikah mendapat skor tertinggi (100%) karena sesuai dengan materi dasar pendidikan Islam. Sebaliknya, soal yang menuntut analisis mendalam, seperti penyalahgunaan kuasa wali dan perbedaan pandangan ulama, memperoleh skor rendah (30–50%) karena kompleksitas konseptual yang belum sepenuhnya terakomodasi oleh sistem. Rata-rata keseluruhan mencapai 96,89%, yang menempatkan performa sistem dalam kategori “Sangat Sesuai” menurut skala Likert. Hasil ini menunjukkan efektivitas teknologi LangChain dan LLM Gemini dalam menghasilkan soal berkualitas, sekaligus menyoroti perlunya pengembangan lebih lanjut dalam merumuskan pertanyaan yang mendorong pemikiran kritis tingkat tinggi dalam pendidikan agama Islam. Pencapaian rata-rata 96,89% mengindikasikan potensi transformatif dari integrasi teknologi LangChain dan model bahasa besar dalam menciptakan alat bantu pedagogis yang secara efektif mendukung proses pembelajaran dan penilaian dalam pendidikan agama Islam [21]

3.4 Implementasi Pengujian Scoring Evaluator LangChain

Scoring evaluator LangChain merupakan pilihan lain untuk pengujian. pengujian ini mengevaluasi respons sistem menggunakan model bahasa pada skala 1 hingga 10. Dengan membandingkan hasil pengujian ini, model bahasa dapat menentukan apakah respons sesuai atau tidak dengan referensi yang diberikan menurut standar yang telah ditentukan sebelumnya. Berikut ini adalah standar evaluasinya:

- Skor 1: Jika jawaban sama sekali tidak terkait dengan referensi.
- Skor 3: Jika jawaban memiliki relevansi minor tetapi tidak sejalan dengan referensi.
- Skor 5: Jika jawaban memiliki relevansi sedang tetapi mengandung ketidakakuratan.
- Skor 7: Jika jawaban sejalan dengan referensi tetapi memiliki kesalahan atau kelalaian minor.
- Skor 10: Jika jawaban benar-benar akurat dan sejalan dengan referensi dengan sempurna.

Berikut adalah hasil pengujian dengan *scoring evaluator* LangChain:

Tabel 4. Hasil Pengujian dengan Scoring Evaluator LangChain

No	Pertanyaan	Skor	Persentase	Tingkat kesulitan
1	Sebutkan salah satu rukun nikah yang disebutkan dalam materi di atas!	10	100%	Mudah
2	Jelaskan apa yang dimaksud dengan sighthat dalam 10		100%	

	konteks pernikahan			Mudah
3	Sebutkan salah satu syarat yang harus dipenuhi oleh seorang wali nikah!	7	70%	Mudah
4	Mengapa seorang wali nikah harus memiliki kesamaan agama dengan mempelai perempuan?	10	100%	Mudah
5	Apa yang dimaksud dengan ijab dan qabul dalam akad nikah?	10	100%	Mudah
6	Mengapa seorang wali nikah harus berakal?	10	100%	Mudah
7	Mengapa seorang wali nikah harus baligh?	10	100%	Mudah
8	Mengapa seorang wali nikah harus merdeka (bukan budak)?	10	100%	Mudah
9	Jika seorang wali nikah tidak memenuhi syarat berakal, apakah pernikahan tersebut sah? Jelaskan!	10	100%	Mudah
10	Sebutkan 3 rukun nikah.	10	100%	Mudah
...
25	Jelaskan mengapa 'merdeka' menjadi syarat wali. Apakah relevan dengan perwalian masa kini?	3	30%	Sulit
26	Bagaimana konsep "berakal" mempengaruhi kelayakan wali nikah? Berikan contoh kasus dan solusinya dalam Islam.	30	30%	Sulit
27	Ijab dan qabul harus jelas. Analisis tantangan di era modern: bahasa, budaya, dan teknologi komunikasi.	3	30%	Sulit
28	Jelaskan perbedaan pandangan ulama tentang urutan wali nikah. Berikan contoh kasus kompleks dan solusinya.	3	30%	Sulit
29	Bagaimana hukum Islam mengatur jika wanita tidak memiliki wali yang sah? Apa alternatifnya dan prosesnya?	10	100%	Sulit
30	Diskusikan peran wali dalam memberi nasihat dan batas intervensi wali terhadap keputusan pribadi mempelai.	1	10%	Sulit
	Total		70,33%	

Menurut data dalam tabel, skor rata-rata untuk pengujian yang dilakukan menggunakan *scoring evaluator* LangChain adalah 70,33%. Interval untuk mengevaluasi bobot pada skala *Likert* adalah:

a. Indeks 0% – 19,99% : Sangat Tidak Setuju

b. Indeks 20% – 39,99% : Tidak Setuju

c. Indeks 40% – 59,99% : Kurang Setuju

d. Indeks 60% – 79,99% : Setuju

e. Indeks 80% – 100% : Sangat Setuju

Scoring Evaluator LanChain memiliki nilai indeks 70,33%. "Setuju" mungkin merupakan skor interval untuk *Scoring Evaluator* LanChain menurut skala *Likert* di atas.

3.5 Integrasi Temuan dan Implikasi Pedagogis

Sintesis komprehensif dari implementasi teknis dan evaluasi empiris mengungkapkan integrasi yang sukses antara teknologi pemrosesan bahasa alami kontemporer dengan kebutuhan pedagogis dalam pendidikan agama Islam.



Konvergensi antara metodologi komputasional dan prinsip instruksional membentuk ekosistem teknologi pendidikan yang berfokus pada pengayaan proses pembelajaran melalui otomatisasi pembuatan soal berkualitas tinggi. Evidensi kuantitatif dari evaluasi pakar menyoroti kapabilitas sistem dalam menghasilkan pertanyaan yang menangkap nuansa teoretis dan filosofis dalam pendidikan agama Islam, mencapai rata-rata persetujuan 96,89% di antara sembilan evaluator dengan keahlian dalam bidang tersebut. Pencapaian ini mendemonstrasikan bahwa pendekatan komputasional berbasis model bahasa besar dan teknologi LangChain dapat secara efektif menghasilkan materi pengujian yang menangkap kompleksitas konseptual yang melekat dalam disiplin ilmu teologis dan filosofis.

Analisis disparitas dalam penilaian—dengan nilai tertinggi pada pertanyaan yang fokus pada interpretasi teoretis (100%) dan nilai terendah pada pertanyaan yang membutuhkan analisis kritis (30%)—memberikan wawasan nuansir tentang kemampuan diferensial sistem dalam mengakomodasi berbagai level taksonomi kognitif. Temuan ini mengartikulasikan arah potensial untuk pengembangan algoritma yang lebih baik memprioritaskan pemikiran tingkat tinggi saat menghasilkan pertanyaan evaluatif dan kritis. Validasi fungsional melalui pengujian *Evaluator Scoring Langchain* melengkapi evaluasi konten dengan konfirmasi bahwa infrastruktur teknis sistem beroperasi secara reliabel di semua komponen—dari input data hingga ekspor hasil. Interoperabilitas yang mulus antara subsistem dan respons yang akurat terhadap parameter pengguna menciptakan pengalaman pengguna yang kohesif, mengurangi hambatan teknis yang sering menghalangi adopsi teknologi dalam konteks pendidikan.

Implikasi pedagogis dari implementasi sistem melampaui otomatisasi sederhana dari proses pembuatan soal. Sistem ini merepresentasikan alat transformatif yang dapat mendukung berbagai strategi instruksional dalam pendidikan agama Islam—dari penilaian formatif yang memfasilitasi umpan balik real-time hingga evaluasi sumatif yang mengukur pencapaian tujuan pembelajaran. Kustomisasi tingkat kesulitan soal memungkinkan diferensiasi instruksional, memfasilitasi pengalaman pembelajaran yang dipersonalisasi untuk mengakomodasi berbagai tingkat kemampuan peserta didik. Dalam konteks yang lebih luas, implementasi sistem ini menandai konvergensi produktif antara tradisi pendidikan Islam dan inovasi teknologi, mendemonstrasikan bahwa kemajuan dalam kecerdasan buatan dan pemrosesan bahasa alami dapat dimanfaatkan untuk memperkaya, bukan menggantikan, modalitas instruksional tradisional dalam pendidikan agama. Pendekatan integratif ini memungkinkan pendidik untuk memanfaatkan efisiensi komputasional sambil mempertahankan nuansa pedagogis yang esensial untuk transmisi pengetahuan religius dan nilai-nilai etika. Sebagai kesimpulan, implementasi dan evaluasi sistem pembuatan soal otomatis untuk pendidikan agama Islam menghasilkan model fungsional yang mendemonstrasikan viabilitas teknis dan relevansi pedagogis. Hasil kuantitatif dan kualitatif mendukung proposisi bahwa teknologi LangChain dan model bahasa besar berbasis Gemini dapat secara efektif diterapkan untuk menghasilkan materi penilaian berkualitas tinggi yang menangkap kompleksitas konseptual dalam pendidikan agama Islam, membuka peluang signifikan untuk inovasi instruksional dalam domain ini [22]

4. KESIMPULAN

Bagian Implementasi sistem pembuatan soal otomatisasi pembelajaran PAI menggunakan *LangChain* dan *LLM* berbasis Gemini membuktikan keberhasilan integrasi teknologi kecerdasan buatan dalam konteks pendidikan keagamaan. Sistem ini berhasil menghasilkan pertanyaan berkualitas tinggi dengan tingkat persetujuan rata-rata 96,89% dari evaluator berpengalaman, mendemonstrasikan kapabilitas mekanisme otomatisasi dalam menghasilkan materi evaluasi yang menangkap kompleksitas konseptual dalam pendidikan agama Islam. Keunggulan sistem terletak pada kemampuannya menghasilkan pertanyaan interpretatif teoretis dengan nilai tertinggi 100%, sementara kelemahannya termanifestasi dalam formulasi pertanyaan analisis kritis dengan nilai terendah 30%. Temuan ini mengindikasikan bahwa meskipun teknologi kecerdasan buatan kontemporer mampu menghasilkan instrumen evaluasi yang valid, masih terdapat ruang pengembangan substansial dalam mengakomodasi taksonomi kognitif tingkat tinggi. Penelitian selanjutnya disarankan untuk mengeksplorasi pengembangan algoritma yang lebih presisi dalam memfasilitasi pemikiran kritis dan evaluatif. Integrasi teknologi pengenalan entitas bernama (*named entity recognition*) dan analisis sentimen dapat meningkatkan kemampuan sistem dalam memformulasikan pertanyaan kompleks. Diperlukan eksplorasi komparatif terhadap berbagai model bahasa besar untuk mengidentifikasi arsitektur optimal dalam konteks pendidikan keagamaan. Implementasi praktis memerlukan kolaborasi interdisipliner antara pakar teknologi dan pendidik agama untuk memastikan sinkronisasi antara inovasi teknologi dan prinsip pedagogis dalam pendidikan agama Islam.

REFERENCES

- [1] Riska Rahman Tanjung, Annida Azhari Ritonga, Bintang Mahrani Abdullah, Nita Afriani Siregar, and Armilah Armilah, "Transformasi Digital dalam Pendidikan: Meningkatkan Kualitas Pembelajaran Melalui Teknologi," *Sinar Dunia: Jurnal Riset*





- Sosial Humaniora dan Ilmu Pendidikan*, vol. 3, no. 2, pp. 211–217, Jun. 2024, doi: 10.58192/sidu.v3i2.2195.
- [2] R. Yasmar and D. R. Amalia, “ANALISIS SWOT PENGGUNAAN CHAT GPT DALAM DUNIA PENDIDIKAN ISLAM,” *Fitrah: Jurnal Studi Pendidikan*, vol. 15, no. 1, pp. 43–64, Jul. 2024, doi: 10.47625/fitrah.v15i1.668.
- [3] N. Noor Kamala Sari, J. Irawan, and V. Handrianus Pranatawijaya, “Implementasi Gemini API untuk Generatif Teks Deskripsi Karya Otomatis dalam Aplikasi Pameran Berbasis Web dengan Metode Waterfall,” 2024. [Online]. Available: <https://doi.org/10.25047/jtit.v11i1.356>
- [4] B. Raharjo, *P Y YAYASAN PRIMA AGUS TEKNIK YAYASAN PRIMA AGUS TEKNIK YAYASAN PRIMA AGUS TEKNIK*.
- [5] Y. Rozi Astino, P. Korespondensi, and V. Pitsalitz Sabandar, “Pengembangan Dan Penerapan Sistem Computer Assisted Test (CAT) Untuk Mengelola Ujian Berbasis Website,” vol. 4, no. 3, pp. 253–259, 2023, doi: 10.33365/jatika.v4i2.2603.
- [6] A. T. U. BR. Lubis, N. S. Harahap, S. Agustian, M. Irsyad, and I. Afrianty, “Question Answering System pada Chatbot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan),” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 3, pp. 955–964, May 2024, doi: 10.57152/malcom.v4i3.1378.
- [7] A. A. Munshi, W. H. AlSabban, A. T. Farag, O. E. Rakha, A. Al Sallab, and M. Alotaibi, “Automated Islamic Jurisprudential Legal Opinions Generation Using Artificial Intelligence,” *Pertanika J Sci Technol*, vol. 30, no. 2, pp. 1135–1156, Apr. 2022, doi: 10.47836/pjst.30.2.16.
- [8] Karishma M Mujawar and Prof. Vidya S, “INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH Content and Question Generation from Book using LLM, Lang Chain, and FAISS,” 2024, doi: 10.15680/IJMRSET.2024.0707173.
- [9] A. Hafiz, A. Mu’ti, and A. Amirrachman, “Pengembangan Media Pembelajaran Agama Islam Berbasis Kecerdasan Buatan: Perspektif Pendidikan Agama Islam,” *Mauriduna: Journal of Islamic Studies*, vol. 5, no. 1, pp. 56–63, May 2024, doi: 10.37274/mauriduna.v5i1.1070.
- [10] N. S. Harahap, A. Saad, and H. Ubaidullah, “Comprehensive Bibliometric Literature Review of Chatbot Research: Trends, Frameworks, and Emerging Applications.” [Online]. Available: www.ijacsa.thesai.org
- [11] L. Ilma’ Nun, A. Mohtarom, A. Marzuki, and U. S. Lawal, “The Integration of Artificial Intelligence as a Teacher’s Partner in Islamic Religious Education Learning,” *Journal of Islamic Education Research*, vol. 6, no. 02, 2025, doi: 10.35719/jier.v6i2.473.
- [12] Achruh, M. Rapi, M. Rusdi, and R. Idris, “Challenges and Opportunities of Artificial Intelligence Adoption in Islamic Education in Indonesian Higher Education Institutions,” *International Journal of Learning, Teaching and Educational Research*, vol. 23, no. 11, pp. 423–443, Nov. 2024, doi: 10.26803/ijlter.23.11.22.
- [13] F. Farhana, A. Suryadi, and D. Wicaksono, “PENGEMBANGAN BAHAN AJAR BERBASIS DIGITAL PADA MATA PELAJARAN BAHASA INGGRIS DI SMK ATLANTIS PLUS DEPOK.”
- [14] D. Susilawati Sj, M. A. Maarif, and A. Zamroni, “Strategi Pengembangan Program Pembelajaran Pendidikan Agama Islam Sekolah,” *Tafkir: Interdisciplinary Journal of Islamic Education*, vol. 2, no. 1, pp. 20–40, Jan. 2021, doi: 10.31538/tijie.v2i1.21.
- [15] F. Rizki, A. Sutiyo, N. S. Harahap, S. Agustian, and R. M. Candra, “KLIK: Kajian Ilmiah Informatika dan Komputer Implementasi Question Answering Berbasis Chatbot Telegram Pada Tafsir Al-Jalalain Menggunakan Langchain dan LLM,” *Media Online*, vol. 4, no. 5, pp. 2464–2472, 2024, doi: 10.30865/klik.v4i5.1784.
- [16] S. Alnefaie, E. Atwell, and M. A. Alsalka, “Using Automatic Question Generation Web Services Tools to Build a Quran Question-and-Answer Dataset,” 2023. [Online]. Available: <https://exploreai.jisc.ac.uk/tool/question-generation>
- [17] Sukiman, Hendry, Muhammad Khahfi Zuhanda, Fenny, and sjukun, “Pelatihan Pemanfaatan Gemini AI untuk Mendukung Pembelajaran pada SMA di Sumatera Utara,” vol. Vol 6, May 2024.
- [18] R. Diky Dermawan, “Meningkatkan Kinerja Output ChatGPT Melalui Teknik Prompt Engineering Yang Dapat Dikustomisasi,” 2024.
- [19] B. Singh, *Building Applications with Large Language Models: Techniques, Implementation, and Applications*. Apress Media LLC, 2024. doi: 10.1007/979-8-8688-0569-1.
- [20] D. Apriliani, “Penggunaan artificial intelligence dalam pembelajaran bahasa Indonesia,” *DIKBASTRA: Jurnal Pendidikan Bahasa dan Sastra*, vol. 7, no. 1, Aug. 2024, doi: 10.22437/dikbastra.v7i1.33262.
- [21] I. Rofiki and A. R. Dewi, “PROBLEMATIKA PENALARAN CHATGPT DALAM MENYELESAIKAN SOAL KOMBINATORIKA [THE PROBLEMATIC REASONING OF CHATGPT IN SOLVING COMBINATORICS PROBLEMS],” *JOHME: Journal of Holistic Mathematics Education*, vol. 8, no. 2, pp. 231–252, Dec. 2024, doi: 10.19166/johme.v8i2.8600.
- [22] Aji Bayu Permadi, Nazruddin Safaat H, Lestari Handayani, and Yusra, “IMPLEMENTASI QUESTION ANSWERING SYSTEM TAFSIR AL-AZHAR MENGGUNAKAN LANGCHAIN DAN LARGE LANGUAGE MODEL BERBASIS CHATBOT TELEGRAM,” *Jurnal Teknoif Teknik Informatika Institut Teknologi Padang*, vol. 12, no. 1, pp. 62–69, Apr. 2024, doi: 10.21063/jtif.2024.v12.1.62-69

