



## Rancang Bangun Sistem Diskusi Hukum Islam Berbasis Multi-Agent AI dengan Metode Retrieval-Augmented Generation (RAG) pada Platform Web Bahtsul Masail

Ryan Yunus<sup>\*1</sup>, Laksamana Rajendra Haidar Azani Fajri<sup>2</sup>

<sup>1</sup>Prodi Ilmu Al-Quran Dan Tafsir, Fakultas Ushuluddin, Universitas Islam Negeri Sunan Kudus

<sup>2</sup>Prodi Komunikasi dan Penyiaran Islam, Fakultas Dakwah, Universitas Islam Negeri Sunan Kudus

Korespondensi penulis: [ryan.yunus@uinsuku.ac.id](mailto:ryan.yunus@uinsuku.ac.id)

**Abstract.** *The development of generative artificial intelligence has opened opportunities for strengthening Islamic legal learning through a digital forum that remains connected to authoritative textual references. This study aims to design and build a web-based bahtsul masail discussion system using a multi-agent AI architecture and the Retrieval-Augmented Generation (RAG) method. The proposed system simulates the deliberative roles commonly found in pesantren-based Islamic legal discussion, namely Moderator, Mubahits, Mu'aridh, and Mushahih. The research applies a Research and Development approach with the Waterfall model, covering requirement analysis, interface design, multi-agent workflow design, implementation, functional testing, and evaluation. RAG is implemented by allowing users to upload PDF documents of kitab kuning and assign the documents to particular agent roles. The uploaded texts are then used as contextual grounding so that each agent can formulate arguments, rebuttals, and final decisions based on traceable references rather than unsupported model memory. The application is implemented using HTML, CSS, and JavaScript on the front end, while the AI reasoning process is orchestrated through an API-based large language model. Functional testing shows that the system can complete five sequential stages of bahtsul masail, display role-based responses, manage uploaded references, and present discussion history. The main contribution of this study is a system design that combines pesantren deliberation procedures with AI-based retrieval support to provide an interactive learning medium for Islamic law, while emphasizing that the final authority of legal validation remains with qualified scholars.*  
**Keywords:** Artificial Intelligence; Bahtsul Masail; Multi-Agent System; Retrieval-Augmented Generation; Web Application.

**Abstrak.** Perkembangan kecerdasan buatan generatif membuka peluang baru untuk memperkuat pembelajaran hukum Islam melalui forum digital yang tetap terhubung dengan rujukan tekstual otoritatif. Penelitian ini bertujuan merancang dan membangun sistem diskusi bahtsul masail berbasis web dengan arsitektur multi-agent AI dan metode Retrieval-Augmented Generation (RAG). Sistem yang diusulkan mensimulasikan peran deliberatif yang lazim ditemukan dalam tradisi diskusi hukum Islam di pesantren, yaitu Moderator, Mubahits, Mu'aridh, dan Mushahih. Penelitian menggunakan pendekatan Research and Development dengan model Waterfall yang meliputi analisis kebutuhan, desain antarmuka, perancangan alur multi-agent, implementasi, pengujian fungsional, dan evaluasi. RAG diterapkan dengan menyediakan fitur unggah dokumen PDF kitab kuning dan pemberian label peran agen terhadap dokumen tersebut. Teks yang diunggah digunakan sebagai konteks rujukan agar setiap agen dapat menyusun argumen, sanggahan, dan putusan akhir berdasarkan referensi yang dapat ditelusuri, bukan hanya berdasarkan memori model. Aplikasi diimplementasikan menggunakan HTML, CSS, dan JavaScript pada sisi antarmuka, sedangkan proses penalaran AI diorkestrasi melalui large language model berbasis API. Hasil pengujian fungsional menunjukkan bahwa sistem mampu menyelesaikan lima tahap bahtsul masail secara berurutan, menampilkan respons sesuai peran, mengelola referensi kitab, serta menyimpan riwayat diskusi. Kontribusi utama penelitian ini adalah rancangan sistem yang memadukan prosedur deliberasi pesantren dengan dukungan retrieval AI untuk menghasilkan media pembelajaran hukum Islam yang interaktif, dengan tetap menegaskan bahwa otoritas validasi hukum berada pada ulama atau ahli keagamaan yang berwenang.  
**Kata kunci:** Aplikasi Web; Bahtsul Masail; Kecerdasan Buatan; Multi-Agent System; Retrieval-Augmented Generation.

## 1. LATAR BELAKANG

Bahtsul masail merupakan tradisi keilmuan Islam yang berkembang kuat di lingkungan pesantren, terutama dalam praktik pengambilan keputusan hukum Islam secara kolektif. Forum ini tidak hanya berfungsi sebagai ruang tanya jawab keagamaan, tetapi juga sebagai mekanisme deliberasi ilmiah yang menekankan pembacaan teks, argumentasi, koreksi, dan keputusan akhir. Secara historis, tradisi intelektual bahtsul masail menjadi salah satu karakter penting dalam produksi pengetahuan hukum Islam di lingkungan Nahdlatul Ulama (Zahro, 2004). Dalam perkembangannya, forum ini mulai menarik perhatian kajian akademik yang menghubungkan bahtsul masail dengan teknologi kecerdasan buatan, termasuk penggunaan ChatGPT dalam konteks moderasi hukum Islam (Faizin et al., 2025).

Pada praktik konvensional, bahtsul masail memerlukan kehadiran peserta, ketersediaan kitab rujukan, kemampuan membaca literatur Arab klasik, serta moderator yang mampu menjaga alur diskusi. Kebutuhan tersebut sering kali menjadi kendala ketika forum ingin diperluas kepada masyarakat yang memiliki keterbatasan akses, jarak geografis, atau belum mempunyai kemampuan teknis membaca kitab secara memadai. Di sisi lain, kebutuhan masyarakat terhadap respons keagamaan yang cepat dan bertanggung jawab semakin meningkat seiring bertambahnya persoalan sosial, ekonomi, teknologi, dan budaya kontemporer.

Perkembangan teknologi informasi memberikan kesempatan untuk mentransformasikan sebagian proses pembelajaran bahtsul masail ke dalam platform digital. Digitalisasi tidak dimaksudkan untuk menggantikan otoritas ulama, melainkan untuk memperluas akses, mendokumentasikan proses diskusi, dan membantu pengguna memahami pola penalaran hukum Islam berbasis rujukan. Aplikasi berbasis web menjadi pilihan yang relevan karena dapat diakses melalui peramban tanpa instalasi perangkat lunak khusus dan dapat mengintegrasikan fitur pengelolaan dokumen, antarmuka diskusi, serta penyimpanan riwayat sesi.

Kemajuan large language model (LLM) dan artificial intelligence (AI) generatif juga membuka peluang baru dalam membangun sistem yang mampu mensimulasikan dialog argumentatif. Model bahasa skala besar memiliki kemampuan mengikuti instruksi, menyusun respons koheren, serta melakukan few-shot learning untuk tugas baru (Brown et al., 2020; OpenAI, 2023). Sistem multi-agent memungkinkan beberapa agen AI diberi

peran berbeda sehingga dapat berinteraksi secara terstruktur. Dalam konteks bahtsul masail, pembagian peran tersebut dapat dipadankan dengan Moderator, Mubahits, Mu'aridh, dan Mushahih. Dengan demikian, sistem tidak hanya menghasilkan satu jawaban langsung, tetapi menyajikan proses diskusi yang bertahap dan menyerupai struktur musyawarah.

Namun demikian, penerapan AI pada domain keagamaan memiliki tantangan serius. Salah satu masalah utama adalah hallucination, yaitu keadaan ketika model menghasilkan informasi, kutipan, atau rujukan yang terdengar meyakinkan tetapi tidak benar-benar bersumber dari dokumen otoritatif. Risiko ini menjadi sangat sensitif dalam kajian hukum Islam karena kesalahan rujukan dapat menimbulkan pemahaman hukum yang keliru. Oleh sebab itu, sistem AI untuk diskusi fiqih memerlukan mekanisme pembatasan dan pelacakan sumber yang kuat.

Metode Retrieval-Augmented Generation (RAG) dipilih untuk mengurangi risiko tersebut. RAG memungkinkan model menggunakan sumber eksternal sebagai konteks sebelum menghasilkan jawaban, sehingga keluaran dapat lebih spesifik dan lebih mudah diverifikasi (Gao et al., 2023; Lewis et al., 2020). Dalam penelitian ini, sumber eksternal yang digunakan berupa PDF kitab kuning yang diunggah pengguna. Dengan cara tersebut, agen AI diarahkan untuk merujuk pada teks yang disediakan dan menyebutkan dasar argumentasinya secara lebih eksplisit. Prinsip ini sejalan dengan kebutuhan bahtsul masail yang menuntut adanya ibarat, dalil, dan dasar kitab yang dapat diverifikasi.

Berdasarkan latar belakang tersebut, penelitian ini berfokus pada rancang bangun sistem diskusi hukum Islam berbasis multi-agent AI dengan metode RAG pada platform web Bahtsul Masail. Tujuan penelitian adalah menghasilkan rancangan dan implementasi sistem yang mampu menjalankan alur sidang otomatis, mengelola dokumen kitab, membedakan peran agen, menampilkan argumen dan sanggahan, serta menyajikan putusan akhir sebagai bahan pembelajaran awal yang tetap memerlukan validasi ulama.

## **2. KAJIAN TEORITIS**

### **Bahtsul Masail dalam Tradisi Pesantren**

Bahtsul masail secara umum dipahami sebagai forum pembahasan masalah-masalah keagamaan yang dilakukan melalui mekanisme musyawarah ilmiah. Dalam tradisi Nahdlatul Ulama, forum ini berkaitan erat dengan kerja lembaga bahtsul masail yang membahas persoalan fiqih dengan merujuk pada kitab-kitab mu'tabar. Mahfudin (2021)

menjelaskan bahwa metodologi istinbath hukum dalam bahtsul masail memiliki pendekatan yang disusun secara hierarkis, antara lain pendekatan qauli, ilhaqi, dan manhaji. Kajian lain menunjukkan bahwa fatwa bahtsul masail juga dapat dianalisis melalui pendekatan maqasid syariah, terutama ketika putusan hukum mempertimbangkan aspek kemaslahatan dan konteks sosial (Hakim et al., 2023). Transformasi metode bahtsul masail dalam berinteraksi dengan Al-Qur'an juga menunjukkan adanya dinamika metodologis yang terus berkembang dalam tradisi hukum Islam NU (Pratomo, 2020; Zahro, 2004).

Secara prosedural, forum bahtsul masail melibatkan aktor yang memiliki peran berbeda. Moderator bertugas membuka dan mengarahkan jalannya diskusi. Mubahits menyampaikan usulan jawaban beserta rujukan. Mu'aridh memberikan sanggahan atau koreksi terhadap argumentasi yang diajukan. Mushahih menilai kelayakan argumentasi dan memberikan pengesahan atau putusan. Struktur ini menunjukkan bahwa bahtsul masail pada dasarnya merupakan sistem deliberatif berbasis peran, sehingga dapat dimodelkan ke dalam arsitektur multi-agent.

### **Artificial Intelligence dan Large Language Model**

Artificial intelligence merujuk pada kemampuan sistem komputasi untuk melakukan tugas yang biasanya memerlukan kecerdasan manusia, seperti memahami bahasa, mengidentifikasi pola, dan menyusun respons. Perkembangan LLM membuat sistem AI mampu memproses instruksi bahasa alami dan menghasilkan teks yang koheren. Brown et al. (2020) menunjukkan bahwa model bahasa skala besar memiliki kemampuan few-shot learning, yaitu kemampuan mengerjakan tugas baru dengan contoh terbatas. Perkembangan kemampuan tersebut diperkuat oleh model generatif mutakhir yang mampu mengikuti instruksi kompleks, meskipun tetap memiliki keterbatasan dan risiko kesalahan faktual (OpenAI, 2023). Kemampuan ini relevan untuk aplikasi diskusi karena sistem dapat diberi instruksi peran dan format jawaban tertentu.

Walaupun LLM memiliki kemampuan bahasa yang kuat, penggunaannya dalam domain hukum Islam tidak boleh dilepaskan dari prinsip validasi sumber. LLM dapat menghasilkan jawaban yang tampak logis tetapi tidak memiliki landasan teks. Penerapan AI dalam penelitian Islam sebenarnya telah dirintis dalam bidang lain, seperti penelusuran perawi hadis secara komputasional (Muballighi, 2022), yang menunjukkan bahwa pendekatan berbasis teks dapat diterapkan pada sumber-sumber keagamaan klasik. Oleh

karena itu, desain sistem harus menggabungkan kemampuan generatif dengan pembatasan sumber, instruksi sistem yang ketat, dan pengawasan manusia. Dalam penelitian ini, AI ditempatkan sebagai alat bantu pembelajaran dan simulasi diskusi, bukan sebagai pemberi fatwa yang memiliki otoritas mandiri.

### **Multi-Agent System**

Multi-Agent System (MAS) adalah pendekatan komputasi yang melibatkan beberapa agen untuk bekerja sama, berinteraksi, atau saling mengevaluasi dalam menyelesaikan tugas. Pada konteks LLM, agen dapat berupa model yang sama tetapi diberi instruksi berbeda sesuai peran. Guo et al. (2024) menyebutkan bahwa sistem multi-agent berbasis LLM berkembang untuk pemecahan masalah kompleks, simulasi sosial, dan pengaturan komunikasi antarpihak. Wang et al. (2024) juga menekankan pentingnya perancangan profil agen, memori, perencanaan, dan mekanisme tindakan dalam sistem agen otonom. Konsep agen generatif yang mampu mensimulasikan perilaku manusia dalam lingkungan sosial juga telah didemonstrasikan oleh Park et al. (2023), sehingga relevan untuk memodelkan interaksi peran dalam forum diskusi.

Pendekatan multi-agent relevan untuk forum bahtsul masail karena forum tersebut mengandung dinamika argumentasi. Jika sistem hanya menggunakan satu agen, hasil yang muncul cenderung berupa jawaban tunggal. Sebaliknya, multi-agent memungkinkan adanya tahapan rumusan masalah, usulan dalil, sanggahan, penguatan jawaban, dan keputusan akhir. Pengembangan kerangka percakapan antargen seperti CAMEL dan AutoGen menunjukkan bahwa pola role-playing, pembagian tugas, dan komunikasi antargen dapat membantu penyelesaian tugas kompleks (Li et al., 2023; Wu et al., 2023). Selain itu, pendekatan reasoning bertahap dan evaluasi diri pada agen, seperti Tree of Thoughts dan Reflexion, mendukung gagasan bahwa keluaran AI dapat diperkuat melalui proses pertimbangan berlapis (Shinn et al., 2023; Yao et al., 2023). Pembagian peran ini membuat proses pembelajaran lebih transparan karena pengguna dapat melihat bagaimana sebuah pendapat dibentuk, diuji, dan disimpulkan.

### **Retrieval-Augmented Generation**

Retrieval-Augmented Generation adalah pendekatan yang menggabungkan proses pencarian informasi dengan kemampuan generasi teks LLM. Lewis et al. (2020) menunjukkan bahwa RAG dapat menghasilkan bahasa yang lebih spesifik dan faktual karena model diberi dukungan dokumen eksternal. Survei perkembangan RAG

menegaskan bahwa pendekatan ini telah berkembang menjadi berbagai arsitektur untuk meningkatkan akurasi, relevansi, dan keterlacakan keluaran model (Gao et al., 2023). Penelitian lain juga menunjukkan bahwa retrieval dari korpus besar maupun integrasi retrieval pada model black-box dapat meningkatkan kualitas keluaran sistem bahasa (Borgeaud et al., 2022; Shi et al., 2024). Implementasi RAG pada aplikasi web juga telah dibuktikan layak secara teknis melalui integrasi model bahasa dengan lingkungan berbasis web (Radeva et al., 2024).

Dalam penelitian ini, RAG diterapkan dengan cara sederhana dan sesuai kebutuhan aplikasi. Pengguna mengunggah PDF kitab kuning, kemudian sistem mengonversi dokumen ke format yang dapat dikirim ke API AI. Dokumen tersebut dipetakan berdasarkan label peran, misalnya kitab untuk Mubahits, kitab untuk Mu'aridh, atau kitab umum. Setiap agen menerima dokumen yang relevan dengan tugasnya, sehingga argumentasi yang dihasilkan diharapkan lebih terikat pada sumber. Desain ini penting untuk mengurangi risiko rujukan palsu dan meningkatkan akuntabilitas jawaban.

### **Penelitian Terdahulu**

Penelitian terdahulu mengenai RAG, LLM, dan agen otonom menunjukkan bahwa penggabungan retrieval dan komunikasi antarputaran dapat meningkatkan kualitas keluaran sistem. Borgeaud et al. (2022) memperlihatkan bahwa pengambilan informasi dari korpus besar dapat meningkatkan kinerja model bahasa. Shi et al. (2024) memperkenalkan REPLUG sebagai pendekatan retrieval yang dapat diterapkan pada model bahasa black-box. Li et al. (2023) melalui CAMEL dan Wu et al. (2023) melalui AutoGen menunjukkan bahwa pola role-playing dan percakapan antargen dapat mendukung penyelesaian tugas kompleks. Shinn et al. (2023) dan Yao et al. (2023) menambahkan bahwa proses refleksi serta eksplorasi jalur penalaran dapat meningkatkan kualitas reasoning agen.

Di sisi kajian Islam, Mahfudin (2021), Hakim et al. (2023), Pratomo (2020), dan Zahro (2004) memberikan landasan tentang metodologi bahtsul masail, transformasi metode istinbath hukum, serta tradisi intelektual NU. Sementara itu, Faizin et al. (2025) memperlihatkan bahwa bahtsul masail mulai dikaji dalam hubungan dengan AI dan moderasi hukum Islam. Penelitian-penelitian tersebut menekankan pentingnya prosedur, rujukan, dan otoritas keilmuan. Kebaruan penelitian ini terletak pada penggabungan

struktur bahtsul masail dengan arsitektur multi-agent AI dan RAG dalam satu platform web yang dirancang untuk pembelajaran interaktif.

### **3. METODE PENELITIAN**

#### **Desain Penelitian**

Penelitian ini menggunakan pendekatan Research and Development (R&D) karena berorientasi pada pengembangan produk sistem informasi. Model pengembangan yang digunakan adalah Waterfall, yaitu model pengembangan perangkat lunak yang berjalan secara bertahap mulai dari analisis kebutuhan hingga pengujian. Model Waterfall dipilih karena tahapan pengembangannya yang sekuensial dan terstruktur, cocok untuk sistem dengan spesifikasi yang relatif jelas sejak awal (Saravanos & Curinga, 2023). Model ini juga sesuai karena alur bahtsul masail yang dimodelkan memiliki urutan tahapan yang jelas, sehingga cocok diterjemahkan ke dalam proses pengembangan sistem yang sekuensial.

Tahap analisis kebutuhan dilakukan untuk mengidentifikasi kebutuhan pengguna, kebutuhan peran agen, kebutuhan dokumen rujukan, dan kebutuhan tampilan antarmuka. Tahap desain mencakup perancangan arsitektur sistem, alur sidang, desain basis antarmuka, dan skema komunikasi antargen. Tahap implementasi mencakup penulisan kode antarmuka web dan integrasi API AI. Tahap pengujian dilakukan melalui pengujian fungsional dan observasi output agen. Tahap evaluasi digunakan untuk menilai kelayakan sistem sebagai media pembelajaran awal.



**Gambar 1. Tahapan Pengembangan Sistem dengan Model Waterfall**

#### **Analisis Kebutuhan Sistem**

Analisis kebutuhan dilakukan untuk memastikan bahwa sistem yang dikembangkan sesuai dengan karakter forum bahtsul masail. Kebutuhan utama sistem mencakup kemampuan membuat masalah baru, mengunggah kitab PDF, memberi label kitab berdasarkan agen, menjalankan sidang otomatis, menampilkan respons tiap agen, dan

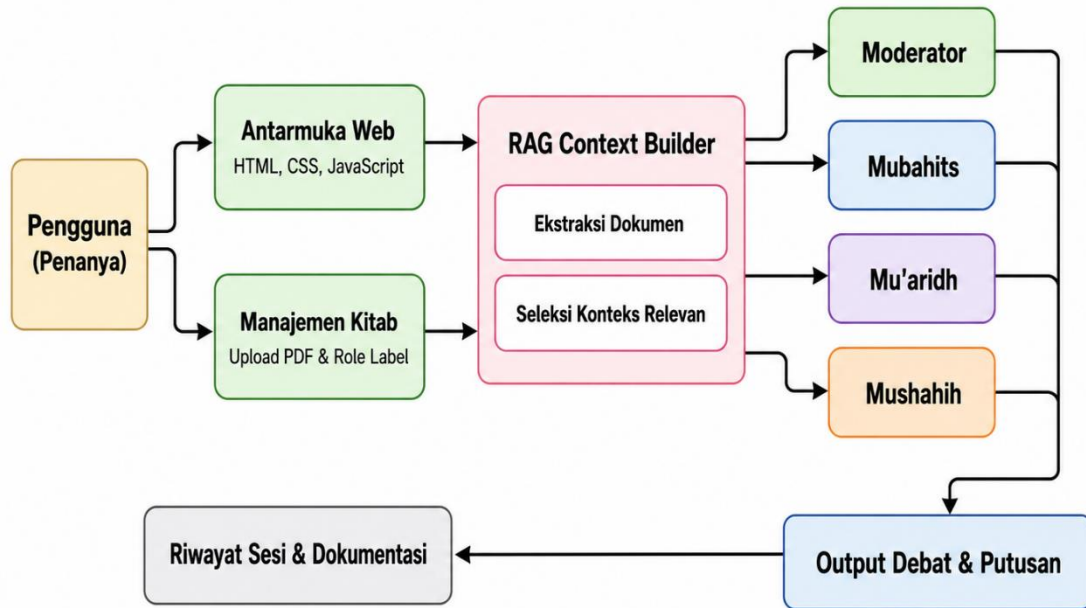
menyimpan riwayat diskusi. Selain kebutuhan fungsional, sistem juga harus memenuhi kebutuhan nonfungsional seperti kemudahan penggunaan, kejelasan tampilan, responsivitas antarmuka, dan keterlacakan sumber.

**Tabel 1. Kebutuhan Fungsional Sistem**

Kode	Kebutuhan Fungsional	Keterangan
F1	Membuat masalah baru	Pengguna dapat memasukkan pertanyaan atau kasus fiqih yang akan dibahas.
F2	Mengunggah kitab PDF	Pengguna dapat memasukkan dokumen kitab sebagai sumber rujukan.
F3	Memberi label agen	Kitab dapat ditetapkan untuk Mubahits, Mu'aridh, atau umum.
F4	Menjalankan sidang otomatis	Sistem memanggil agen secara berurutan sesuai alur bahtsul masail.
F5	Menampilkan output per tahap	Setiap respons agen ditampilkan dengan identitas peran yang jelas.
F6	Menyimpan riwayat sesi	Pertanyaan dan hasil diskusi dapat dilihat kembali oleh pengguna.

### Perancangan Arsitektur Sistem

Arsitektur sistem dirancang dengan memisahkan lapisan antarmuka, manajemen dokumen, penyusun konteks RAG, orkestrator API, dan agen AI. Antarmuka web bertugas menerima input pengguna dan menampilkan hasil. Manajemen dokumen menangani unggah PDF dan label peran. Penyusun konteks RAG menyiapkan dokumen yang relevan untuk setiap agen. Orkestrator API mengatur pemanggilan agen secara berurutan. Agen AI bertugas menghasilkan output sesuai peran masing-masing.



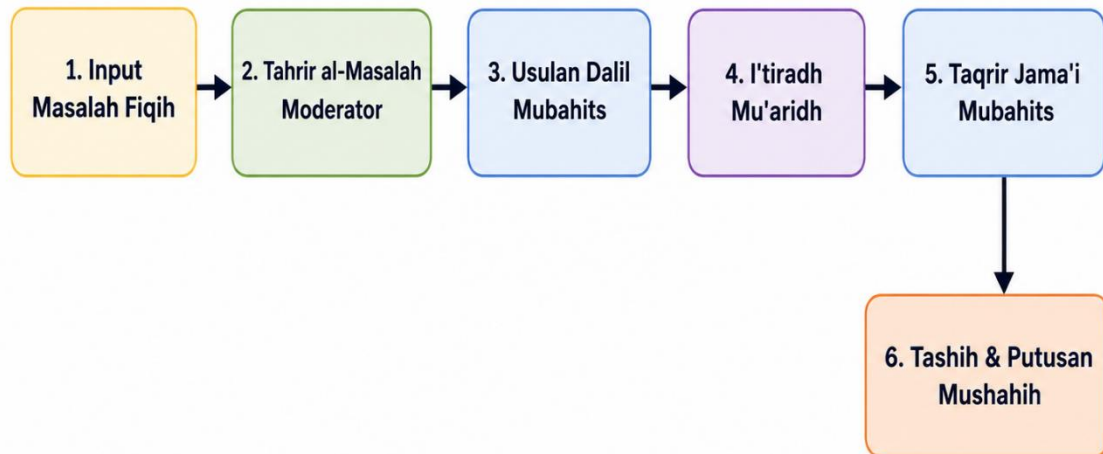
Gambar 2. Arsitektur Multi-Agent RAG pada Sistem Bahtsul Masail Digital

Tabel 2. Deskripsi Peran Agen dalam Sistem

Agen	Peran dalam Forum	Output yang Diharapkan
Moderator	Membuka sidang, merumuskan masalah, dan mengarahkan alur diskusi.	Tahrir al-masalah dan instruksi kepada Mubahits.
Mubahits	Mengusulkan jawaban hukum berdasarkan kitab yang diberikan.	Usulan hukum, kutipan kitab, dan istidlal.
Mu'aridh	Memberikan sanggahan, koreksi, atau pertanyaan kritis.	I'tiradh, dalil tandingan, dan catatan kelemahan argumen.
Mushahih	Menilai argumen dan memberikan keputusan akhir.	Tashih, putusan hukum, dan dasar rujukan.

### Perancangan Alur Sidang Digital

Alur sidang digital dirancang mengikuti lima tahap utama. Pertama, Moderator melakukan tahrir al-masalah untuk merumuskan pokok persoalan. Kedua, Mubahits mengajukan jawaban awal dengan rujukan kitab. Ketiga, Mu'aridh memberikan sanggahan atau pembandingan. Keempat, Mubahits melakukan taqirir jama'i dengan memperkuat jawaban. Kelima, Mushahih memberikan putusan akhir. Setiap tahap menerima konteks dari tahap sebelumnya sehingga dialog tetap berkesinambungan.



**Gambar 3. Alur Sidang Otomatis Forum AI Bahtsul Masail**

### **Implementasi RAG dan Pengelolaan Dokumen**

Implementasi RAG dilakukan melalui mekanisme unggah dokumen PDF pada sisi pengguna. Dokumen yang diunggah dibaca oleh sistem dan dikaitkan dengan label agen. Label ini penting karena dalam forum bahtsul masail setiap pihak dapat memiliki sumber atau kitab yang berbeda. Dengan demikian, Mubahits dapat menggunakan kitab tertentu untuk mengajukan pendapat, sedangkan Mu'aridh dapat memakai kitab lain untuk menyanggah atau memberi perbandingan. Mushahih diberi akses lebih luas agar dapat menilai keseluruhan argumentasi. Desain ini sejalan dengan prinsip aplikasi RAG berbasis web yang menempatkan dokumen eksternal sebagai konteks utama bagi model bahasa (Radeva et al., 2024).

Setiap prompt agen berisi instruksi peran, batasan penggunaan sumber, format keluaran, dan larangan membuat rujukan yang tidak tersedia. Instruksi ini dirancang agar AI tidak sekadar menjawab secara umum, tetapi mengikuti struktur diskusi. RAG tidak menjamin kebenaran hukum secara mutlak, tetapi meningkatkan peluang keluaran yang dapat ditelusuri karena agen diarahkan membaca konteks dokumen yang disediakan pengguna.

**Tabel 3. Skema Distribusi Dokumen RAG per Agen**

Label Dokumen	Agen Penerima	Tujuan Penggunaan
Mubahits	Mubahits	Menjadi dasar usulan jawaban dan istidlal awal.
Mu'aridh	Mu'aridh	Menjadi dasar sanggahan, perbandingan, atau koreksi.
Umum	Moderator dan Mushahih	Menjadi referensi umum untuk merumuskan masalah dan menilai putusan.
Semua kitab	Mushahih	Menilai konsistensi argumen dari seluruh pihak.

Sumber: Data diolah penulis, 2026.

#### F. Rancangan Pengujian

Pengujian dilakukan dengan metode black-box untuk memastikan bahwa setiap fungsi utama berjalan sesuai tujuan. Skenario pengujian berfokus pada kemampuan sistem menerima input, memproses dokumen, menjalankan alur multi-agent, dan menampilkan hasil. Selain itu, pengujian observasional dilakukan untuk membandingkan keluaran ketika sistem berjalan tanpa dokumen RAG dan ketika sistem menggunakan dokumen kitab.

**Tabel 4. Rancangan Pengujian Black-Box**

Fungsi	Skenario Uji	Indikator Berhasil
Input masalah	Pengguna mengetik kasus fiqh pada kolom pertanyaan.	Masalah diterima dan sidang dapat dimulai.
Upload kitab	Pengguna memilih file PDF dan label agen.	File muncul pada daftar kitab referensi.
Sidang otomatis	Pengguna menjalankan diskusi setelah mengisi pertanyaan.	Lima tahap sidang muncul secara berurutan.
RAG	Agen menerima dokumen kitab sesuai label.	Jawaban memuat rujukan yang relevan dengan dokumen.

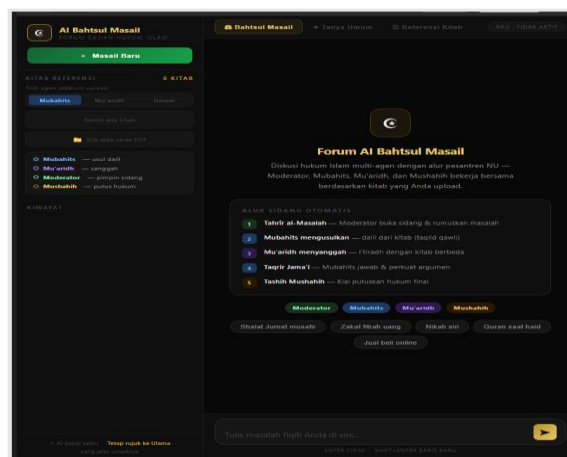
Fungsi	Skenario Uji	Indikator Berhasil
Riwayat sesi	Pengguna membuka ulang sesi yang telah dilakukan.	Pertanyaan dan output sebelumnya dapat ditampilkan.

#### 4. HASIL DAN PEMBAHASAN

##### Hasil Implementasi Sistem

Hasil implementasi menunjukkan bahwa sistem berhasil dibangun sebagai aplikasi web dengan tampilan utama yang terdiri atas sidebar referensi kitab, area informasi alur sidang, daftar contoh pertanyaan, serta kolom input masalah fiqih. Sidebar digunakan untuk menambahkan masalah baru, mengelola kitab referensi, memilih label agen, dan melihat riwayat sesi. Area utama digunakan untuk menampilkan identitas forum, tahapan sidang otomatis, dan keluaran respons agen.

Desain antarmuka menggunakan tema gelap dengan aksen warna berbeda untuk tiap agen. Perbedaan warna ini membantu pengguna mengenali peran yang sedang berbicara. Moderator ditandai dengan warna hijau, Mubahits dengan warna biru, Mu'aridh dengan warna ungu, dan Mushahih dengan warna kuning. Pilihan visual tersebut memperkuat pemahaman bahwa diskusi berlangsung secara bertahap dan melibatkan peran yang berbeda.



**Gambar 5. Tampilan Antarmuka Aplikasi Forum AI Bahtsul Masail**

*Sumber: Dokumentasi tampilan aplikasi yang dikembangkan, 2026.*

## **Pembahasan Antarmuka dan Pengalaman Pengguna**

Antarmuka sistem dirancang sederhana agar pengguna dapat langsung memahami alur penggunaan. Bagian kiri layar berfungsi sebagai pusat kendali referensi. Pengguna dapat memilih tab Mubahits, Mu'aridh, atau Umum sebelum mengunggah kitab. Desain ini penting karena distribusi dokumen memengaruhi konteks yang akan diterima agen. Jika dokumen tidak diberi label, argumen agen berpotensi tidak sesuai dengan sumber yang diinginkan.

Bagian tengah layar menampilkan ringkasan alur sidang otomatis. Ringkasan ini berfungsi sebagai scaffolding pembelajaran bagi pengguna yang belum terbiasa dengan tahapan bahtsul masail. Daftar contoh masalah, seperti shalat Jumat musafir, zakat fitrah uang, nikah siri, dan Al-Qur'an saat haid, disediakan untuk mempercepat eksplorasi sistem. Kolom input di bagian bawah memungkinkan pengguna menulis kasus fiqih secara bebas.

Dari sisi pengalaman pengguna, sistem menekankan keterbacaan, identitas peran, dan alur proses. Tantangan antarmuka yang masih perlu diperhatikan adalah bagaimana menampilkan kutipan kitab yang panjang tanpa membuat layar penuh, bagaimana memberi peringatan ketika dokumen belum diunggah, serta bagaimana menampilkan tingkat keyakinan atau status verifikasi rujukan. Perbaikan tersebut dapat menjadi agenda pengembangan berikutnya.

## **Hasil Alur Multi-Agent**

Sistem menjalankan alur multi-agent secara sekuensial. Pada tahap awal, Moderator merumuskan pertanyaan ke dalam bentuk masalah hukum yang lebih jelas. Setelah itu Mubahits memberikan jawaban awal dengan rujukan. Mu'aridh memberikan sanggahan untuk menguji argumentasi. Mubahits kemudian diberi kesempatan menjawab sanggahan. Pada tahap terakhir, Mushahih menyimpulkan dan memberi putusan. Pola ini membuat keluaran sistem tidak hanya berupa kesimpulan, tetapi juga proses reasoning yang dapat dibaca pengguna.

**Tabel 5. Hasil Observasi Output Tiap Agen**

<b>Tahap</b>	<b>Agen</b>	<b>Karakteristik Output</b>
1	Moderator	Membuka diskusi, merumuskan pokok masalah, dan mengatur giliran.

Tahap	Agen	Karakteristik Output
2	Mubahits	Mengusulkan hukum awal, menyebut rujukan, dan menyusun istidlal.
3	Mu'aridh	Memberikan sanggahan, dalil pembanding, dan pertanyaan kritis.
4	Mubahits	Menjawab sanggahan dan memperkuat argumentasi awal.
5	Mushahih	Menilai argumen dan menyajikan keputusan akhir yang lebih ringkas.

### Efektivitas RAG dalam Membatasi Jawaban

Perbandingan keluaran menunjukkan bahwa penggunaan dokumen RAG meningkatkan spesifisitas jawaban. Ketika sistem tidak diberi dokumen, agen cenderung menghasilkan jawaban umum dengan penyebutan kitab atau kaidah yang masih generik. Ketika dokumen kitab diunggah, agen lebih mampu menyesuaikan jawaban dengan konteks dokumen dan diarahkan untuk menyebut rujukan secara lebih eksplisit. Hal ini menunjukkan bahwa RAG berperan sebagai mekanisme grounding yang membantu AI tetap berada dalam batas sumber yang diberikan, sebagaimana ditekankan dalam penelitian RAG sebelumnya (Gao et al., 2023; Lewis et al., 2020; Radeva et al., 2024).

Meskipun demikian, RAG tidak boleh dipahami sebagai jaminan absolut. Kualitas keluaran tetap bergantung pada kualitas dokumen PDF, keterbacaan teks, panjang konteks yang dapat diproses, kejelasan instruksi prompt, dan kemampuan model memahami bahasa Arab atau teks kitab. Oleh karena itu, sistem perlu dilengkapi pesan kehati-hatian bahwa hasil AI adalah bahan kajian awal dan bukan keputusan hukum final yang berdiri sendiri.

**Tabel 6. Perbandingan Output Sistem Tanpa dan Dengan RAG**

Aspek	Tanpa RAG	Dengan RAG
Spesifisitas referensi	Cenderung umum dan bergantung pada memori model.	Lebih spesifik karena diarahkan pada dokumen kitab yang diunggah.
Risiko hallucination	Lebih tinggi karena model dapat membuat rujukan tidak terverifikasi.	Lebih rendah karena jawaban dibatasi oleh konteks dokumen.

Aspek	Tanpa RAG	Dengan RAG
Kesesuaian peran agen	Peran masih dapat berjalan, tetapi argumen kurang berbasis sumber.	Peran lebih kuat karena setiap agen menerima dokumen sesuai label.
Kemudahan verifikasi	Pengguna sulit memeriksa sumber jawaban.	Pengguna dapat menelusuri kembali dokumen yang menjadi dasar jawaban.
Kualitas pembelajaran	Memberi jawaban cepat tetapi kurang menunjukkan sumber.	Mendorong pembelajaran berbasis teks dan argumentasi.

### Hasil Pengujian Fungsional

Pengujian fungsional dilakukan pada beberapa contoh masalah fiqih untuk memastikan sistem dapat menyelesaikan alur diskusi. Masalah yang digunakan mewakili tema ibadah, muamalah, dan hukum keluarga. Hasil pengujian menunjukkan bahwa seluruh skenario dapat menjalankan lima tahap diskusi. Waktu proses dipengaruhi oleh panjang dokumen, jumlah token konteks, kualitas koneksi API, dan kompleksitas pertanyaan.

**Tabel 7. Hasil Pengujian Fungsional per Skenario Masalah**

Masalah Fiqih	Alur Selesai	Referensi Muncul	Waktu Proses ( $\pm$ detik)
Shalat Jumat bagi musafir	Ya	Ya	45-60
Zakat fitrah menggunakan uang	Ya	Ya	40-55
Nikah siri dan pencatatan negara	Ya	Ya	50-70
Membaca Al-Qur'an bagi wanita haid	Ya	Ya	42-58
Jual beli online dan akad salam	Ya	Ya	55-75

### Validasi Logika Sistem

Validasi logika sistem dilakukan dengan memeriksa apakah setiap agen menghasilkan output sesuai peran. Moderator dinilai berhasil apabila mampu merumuskan masalah dan tidak langsung memberikan putusan. Mubahits dinilai berhasil apabila menyampaikan jawaban awal dan dalil. Mu'aridh dinilai berhasil apabila tidak

mengulang jawaban Mubahits, tetapi memberikan keberatan atau pembandingan. Mushahih dinilai berhasil apabila merangkum argumen dan memberikan keputusan akhir berdasarkan proses sebelumnya.

Hasil observasi menunjukkan bahwa pemisahan prompt peran sangat menentukan kualitas diskusi. Jika instruksi terlalu umum, agen dapat saling tumpang tindih. Misalnya, Mu'aridh dapat berubah menjadi pemberi jawaban baru tanpa menyanggah argumen sebelumnya, atau Mushahih dapat terlalu panjang sehingga menyerupai pembahasan ulang. Oleh karena itu, prompt perlu dirumuskan secara tegas dengan struktur jawaban yang konsisten. Prinsip ini sejalan dengan temuan penelitian agen LLM yang menekankan pentingnya instruksi peran, refleksi, dan reasoning bertahap dalam meningkatkan kualitas keluaran (Shinn et al., 2023; Yao et al., 2023).

**Tabel 8. Indikator Keberhasilan Peran Agen**

<b>Agen</b>	<b>Indikator Keberhasilan</b>	<b>Risiko Jika Prompt Lemah</b>
Moderator	Mampu merumuskan masalah dan menjaga urutan sidang.	Langsung memberi jawaban hukum tanpa membuka ruang diskusi.
Mubahits	Mampu menyajikan usulan hukum berbasis rujukan.	Jawaban terlalu umum dan tidak memuat dasar kitab.
Mu'aridh	Mampu memberi sanggahan yang relevan terhadap usulan.	Hanya mengulang pendapat Mubahits atau membuat topik baru.
Mushahih	Mampu memberi keputusan akhir yang ringkas dan bertanggung jawab.	Putusan terlalu spekulatif atau tidak mempertimbangkan argumen sebelumnya.

### **Keterbatasan Sistem**

Sistem yang dikembangkan masih memiliki beberapa keterbatasan. Pertama, sistem sangat bergantung pada kualitas PDF yang diunggah. Jika PDF berupa hasil pindai dengan kualitas rendah, teks dapat sulit diproses. Kedua, sistem belum menggunakan vector database sehingga pengelolaan korpus kitab berukuran sangat besar belum optimal. Ketiga, sistem belum memiliki panel validasi ulama yang dapat memberi status sah, perlu revisi, atau tidak layak digunakan. Keempat, sistem belum memiliki fitur pelacakan halaman yang selalu presisi untuk setiap kutipan.

Keterbatasan lain berkaitan dengan aspek etika dan otoritas. Dalam kajian hukum Islam, proses penentuan hukum tidak hanya bergantung pada kemiripan teks, tetapi juga pada penguasaan ushul fiqih, maqashid, konteks sosial, dan otoritas keilmuan. Oleh karena itu, aplikasi ini harus ditempatkan sebagai sistem bantu belajar dan simulasi diskusi, bukan sebagai pengganti musyawarah ulama atau fatwa resmi.

### **Implikasi Pengembangan**

Implikasi teoretis dari penelitian ini adalah adanya model konseptual yang menghubungkan prosedur bahtsul masail dengan arsitektur multi-agent AI. Pemodelan ini memperlihatkan bahwa tradisi diskusi pesantren memiliki struktur yang dapat diterjemahkan menjadi alur komputasional tanpa menghilangkan prinsip kehati-hatian. Dari sisi praktis, sistem dapat digunakan sebagai media latihan bagi mahasiswa, santri, atau masyarakat yang ingin memahami bagaimana sebuah masalah fiqih dibahas secara bertahap.

Implikasi teknisnya adalah perlunya pengembangan sistem yang lebih kuat pada sisi retrieval, validasi rujukan, dan dokumentasi hasil sidang. Pada tahap lanjut, sistem dapat diintegrasikan dengan database kitab, OCR berkualitas tinggi, pencarian semantik, ekspor dokumen hasil sidang, serta modul penilaian manusia. Dengan demikian, sistem tidak hanya menjadi chatbot diskusi, tetapi berkembang menjadi platform dokumentasi dan pembelajaran bahtsul masail digital.

## **5. KESIMPULAN DAN SARAN**

Penelitian ini berhasil merancang dan membangun sistem diskusi hukum Islam berbasis multi-agent AI dengan metode Retrieval-Augmented Generation pada platform web Bahtsul Masail. Sistem melibatkan empat agen utama, yaitu Moderator, Mubahits, Mu'aridh, dan Mushahih, yang bekerja secara sekuensial mengikuti lima tahap diskusi. Implementasi RAG melalui fitur unggah PDF kitab kuning memungkinkan agen memperoleh konteks sumber yang lebih jelas sehingga jawaban dapat diarahkan pada rujukan yang lebih mudah diverifikasi.

Pengujian fungsional terhadap lima skenario masalah fiqih — meliputi tema ibadah, muamalah, dan hukum keluarga — menunjukkan bahwa seluruh alur diskusi dapat diselesaikan dengan waktu proses berkisar antara 40 hingga 75 detik per skenario. Seluruh skenario berhasil menampilkan referensi kitab secara eksplisit. Validasi logika peran menunjukkan bahwa pemisahan prompt agen secara tegas mampu menghasilkan

output yang sesuai fungsi masing-masing: Moderator merumuskan masalah, Mubahits menyusun istidlal, Mu'aridh memberikan sanggahan, dan Mushahih merangkum putusan akhir. Perbandingan sistem dengan dan tanpa RAG menunjukkan peningkatan spesifisitas referensi dan penurunan risiko hallucination ketika dokumen kitab diunggah, karena jawaban agen dibatasi oleh konteks dokumen yang tersedia.

Hasil pengujian menunjukkan bahwa sistem mampu menjalankan alur sidang otomatis, menampilkan respons sesuai peran, mengelola dokumen kitab berdasarkan label, dan menyimpan riwayat sesi. Penggunaan RAG meningkatkan spesifisitas referensi dan menurunkan risiko jawaban generik dibandingkan sistem tanpa dokumen. Meskipun demikian, hasil AI tetap harus dipahami sebagai bahan pembelajaran awal yang memerlukan validasi ahli karena otoritas penetapan hukum Islam berada pada ulama atau lembaga keagamaan yang berwenang.

Saran pengembangan berikutnya adalah menambahkan vector database untuk mengelola korpus kitab yang lebih besar, meningkatkan akurasi OCR dan pelacakan halaman, menambahkan fitur ekspor hasil sidang, menyediakan panel validasi ulama, serta mengembangkan mode metodologi bahtsul masail yang lebih lengkap seperti ilhaq al-masail dan istinbath jama'i. Penelitian lanjutan juga perlu menguji kualitas output sistem melalui penilaian pakar fiqh dan pengguna sasaran.

## **UCAPAN TERIMA KASIH**

Penulis menyampaikan terima kasih kepada seluruh pihak yang mendukung pengembangan sistem dan penyusunan artikel ini. Ucapan terima kasih juga disampaikan kepada komunitas akademik dan pesantren yang memberikan inspirasi terhadap pengembangan media pembelajaran hukum Islam berbasis teknologi digital.

## **DAFTAR REFERENSI**

- Borgeaud, S., Mensch, A., Hoffmann, J., Cai, T., Rutherford, E., Millican, K., van den Driessche, G. B., Lespiau, J. B., Damoc, B., Clark, A., de Las Casas, D., Guy, A., Menick, J., Ring, R., Hennigan, T., Huang, S., Maggiore, L., Jones, C., Cassirer, A., ... Sifre, L. (2022). Improving language models by retrieving from trillions of tokens. *Proceedings of the 39th International Conference on Machine Learning*, 162, 2206-2240.
- 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., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information*

- Processing Systems, 33, 1877-1901. Retrieved from <https://arxiv.org/abs/2005.14165>
- Faizin, N., Maarif, A. S., & Hanafi, Y. (2025). Considering religious moderation in Islamic law through AI ChatGPT and Bahtsul Masail of Nahdlatul Ulama. *Law, Innovation and Technology*, 17(1), 271-288. <https://doi.org/10.1080/17579961.2025.2469351>
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, H., & Wang, H. (2023). Retrieval-augmented generation for large language models: A survey. arXiv preprint, arXiv:2312.10997. Retrieved from <https://arxiv.org/abs/2312.10997>
- Guo, T., Chen, X., Wang, Y., Chang, R., Pei, S., Chawla, N. V., Wiest, O., & Zhang, X. (2024). Large language model based multi-agents: A survey of progress and challenges. *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 8048-8057. <https://doi.org/10.24963/ijcai.2024/890>
- Hakim, A., Ahmad, R., & Jaafar, S. M. J. S. (2023). The methodology of reforming Islamic law fatwa Bahthul Masail Nahdlatul Ulama in family planning (KB) in Indonesia by the Maqasid Shari'ah approach. *Al-Manhaj: Jurnal Hukum dan Pranata Sosial Islam*, 5(1), 939-952. <https://doi.org/10.37680/almanhaj.v5i1.2934>
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Kuttler, H., Lewis, M., Yih, W. T., Rocktaschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems*, 33, 9459-9474. Retrieved from <https://arxiv.org/abs/2005.11401>
- Li, G., Hammoud, H. A. K., Itani, H., Khizbullin, D., & Ghanem, B. (2023). CAMEL: Communicative agents for mind exploration of large scale language model society. *Advances in Neural Information Processing Systems*, 36, 51991-52008. Retrieved from <https://arxiv.org/abs/2303.17760>
- Mahfudin, A. (2021). Metodologi istinbath hukum Lembaga Bahtsul Masail Nahdlatul Ulama. *Jurnal Hukum Keluarga Islam*, 6(1), 1-17. Retrieved from <https://journal.unipdu.ac.id/index.php/jhki/article/view/2573>
- Muballighi, A. (2022). Tracing hadith narrators with AI: Computational approaches in Islamic legal research. *Religions*, 13(11), 1012. <https://doi.org/10.3390/rel13111012>
- OpenAI. (2023). GPT-4 technical report. arXiv preprint, arXiv:2303.08774. Retrieved from <https://arxiv.org/abs/2303.08774>
- Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). Generative agents: Interactive simulacra of human behavior. *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 1-22. <https://doi.org/10.1145/3586183.3606763>
- Pratomo, H. (2020). Transformasi metode bahtsul masail NU dalam berinteraksi dengan Al-Qur'an. *Jurnal Lektur Keagamaan*, 18(1), 109-134. <https://doi.org/10.31291/jlka.v18i1.620>
- Radeva, I., Popchev, I., Doukowska, L., & Dimitrova, M. (2024). Web application for retrieval-augmented generation: Implementation and testing. *Electronics*, 13(7), 1361. <https://doi.org/10.3390/electronics13071361>
- Saravanos, A., & Curinga, M. X. (2023). Simulating the software development lifecycle: The waterfall model. *Applied System Innovation*, 6(6), 108. <https://doi.org/10.3390/asi6060108>

- Shi, W., Min, S., Yasunaga, M., Seo, M., James, R., Lewis, M., Zettlemoyer, L., & Yih, W. T. (2024). REPLUG: Retrieval-augmented black-box language models. *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 8364-8377. <https://doi.org/10.18653/v1/2024.naacl-long.463>
- Shinn, N., Cassano, F., Gopinath, A., Narasimhan, K., & Yao, S. (2023). Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36, 8634-8652. Retrieved from <https://arxiv.org/abs/2303.11366>
- Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. R. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), 186345. <https://doi.org/10.1007/s11704-024-40231-1>
- Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A. H., White, R. W., Burger, D., & Wang, C. (2023). AutoGen: Enabling next-gen LLM applications via multi-agent conversation framework. *arXiv preprint, arXiv:2308.08155*. Retrieved from <https://arxiv.org/abs/2308.08155>
- Yao, S., Yu, D., Zhao, J., Shafran, I., Griffiths, T. L., Cao, Y., & Narasimhan, K. (2023). Tree of thoughts: Deliberate problem solving with large language models. *Advances in Neural Information Processing Systems*, 36, 11809-11822. Retrieved from <https://arxiv.org/abs/2305.10601>
- Zahro, A. (2004). *Lajnah Bahtsul Masa'il 1926-1999: Tradisi intelektual NU*. LKiS.