



PENGEMBANGAN *CHATBOT* BERBASIS *RETRIEVAL AUGMENTED GENERATION* DENGAN STUDI KASUS DASAR POKOK AGRARIA

ANDIKA RISKY SURURI



PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Pengembangan Chatbot Berbasis *Retrieval Augmented Generation* dengan Studi Kasus Dasar Pokok Agraria” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, April 2025

Andika Risky Sururi
G6401211046

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

ANDIKA RISKY SURURI. Pengembangan *Chatbot* Berbasis *Retrieval Augmented Generation* dengan Studi Kasus Dasar Pokok Agraria. Dibimbing oleh IRMAN HERMADI.

Persoalan agraria di Indonesia menjadi salah satu persoalan yang perlu diperhatikan lebih mendalam sebab persoalan agraria erat kaitannya dengan persengketaan tanah. Berdasarkan Konsorsium Pembaruan Agraria (KPA), persoalan agraria kembali mengalami kenaikan sejak tahun 2021 hingga 2023. Banyak faktor yang menyebabkan persoalan agraria terjadi. Salah satu faktor penyebab tersebut adalah kurangnya pemahaman masyarakat terhadap peraturan perundang-undangan yang mengatur dasar pokok-pokok agraria. Penyuluhan kepada masyarakat dengan tujuan meningkatkan pemahaman kini menjadi solusi yang dijalankan dengan harapan peraturan perundang-undangan yang mengatur tentang dasar pokok agraria dapat diketahui dan dipahami oleh masyarakat. Namun, solusi ini menjadi kurang efektif berdasarkan sisi waktu dan biaya. Penyuluh perlu ke berbagai daerah dengan memakan waktu yang panjang untuk melakukan penyuluhan. Oleh karena itu, diperlukan pendekatan teknologi informasi yang menjembatani pengetahuan masyarakat terhadap peraturan perundang-undangan tentang agraria melalui *chatbot* yang mampu menjawab pertanyaan masyarakat seputar peraturan perundang-undangan dasar pokok agraria. *Chatbot* dibangun menggunakan *Large Language Model* dengan pendekatan *Retrieval Augmented Generation* (RAG) sehingga jawaban yang diberikan oleh *chatbot* bersumber dari peraturan perundang-undangan yang mengatur dasar pokok-pokok agraria. Melalui penelitian ini harapannya masyarakat mampu mendapatkan informasi terkait peraturan perundang-undangan tentang dasar pokok-pokok agraria dengan mudah dan mendorong peningkatan pemahaman masyarakat terhadap peraturan tersebut.

Kata Kunci: *Chatbot*, *Retrieval Augmented Generation*, Dasar Pokok Agraria.

ABSTRACT

ANDIKA RISKY SURURI. Chatbot Development Based on Retrieval Augmented Generation with Basic Agrarian Case Study. Supervised by IRMAN HERMADI.

Agrarian issues in Indonesia are one of the issues that need to be considered more deeply, because agrarian issues are closely related to land disputes. Based on the Agrarian Reform Consortium (KPA), agrarian issues have increased again from 2021 to 2023. Many factors cause agrarian issues to occur, one of the causes is the lack of public understanding of the laws and regulations governing the basic principles of agrarian law. Counseling the community with the aim of increasing understanding is now a solution that is implemented with the hope that the laws and regulations governing the basic principles of agrarian law can be known and understood by the community. However, this solution is less effective in terms of time and cost. Extension workers need to go to various regions and take a long time

to conduct counseling. Therefore, an information technology approach is needed that bridges the gap between laws and regulations on agrarian law and public understanding through chatbots that are able to answer public questions about basic agrarian laws and regulations. The chatbot is built using the Large Language Model with the Retrieval Augmented Generation (RAG) approach so that the answers given by the chatbot are sourced from laws and regulations governing the basic principles of agrarian law. Through this research, it is hoped that the public will be able to obtain information related to laws and regulations on the basic principles of agrarian law easily and encourage increased public understanding of these regulations.

Keywords: Chatbot, Retrieval Augmented Generation, Agrarian Principles.

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PENGEMBANGAN *CHATBOT* BERBASIS *RETRIEVAL AUGMENTED GENERATION* DENGAN STUDI KASUS DASAR POKOK AGRARIA

ANDIKA RISKY SURURI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:

1. Hafidlotul Fatimah Ahmad S.Kom., M.Kom
2. Rina Trisminingsih S.Kom., M.T., Ph.D

Judul Skripsi : Pengembangan *Chatbot* Berbasis *Retrieval Augmented Generation* dengan Studi Kasus Dasar Pokok Agraria

Nama : Andika Risky Sururi
NIM : G6401211046

Disetujui oleh

Pembimbing :
Irman Hermadi, S.Kom, MS, Ph.D



Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:
Dr. Sony Hartono Wijaya S.Kom, M.Kom
NIP.198108092008121002



Tanggal Ujian:
20 Mei 2025

Tanggal Lulus:
(tanggal penandatanganan oleh Dekan
Fakultas/Sekolah ...)



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2024 sampai bulan Maret 2024 ini ialah pengembangan *chatbot* berbasis RAG, dengan judul "Pengembangan *Chatbot* Berbasis *Retrieval Augmented Generation* dengan Studi Kasus Dasar Pokok Agraria".

Terima kasih penulis ucapkan kepada para pembimbing, Irman Hermadi, S.Kom, MS, Ph.D yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada pembimbing akademik, moderator seminar, dan penguji luar komisi pembimbing. Ungkapan terima kasih juga disampaikan kepada ayah, ibu, serta seluruh keluarga yang telah memberikan dukungan, doa, dan kasih sayangnya.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, April 2025

Andika Risky Sururi

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	2
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	3
2.1 <i>Chatbot</i>	3
2.2 <i>Retrieval Augmented Generation</i>	4
2.3 Implementasi RAG	4
2.4 <i>Elasticsearch</i>	6
2.5 <i>Retriaval Augmented Generation Assesment (RAGAS)</i>	6
III METODE	8
3.1 Data Penelitian	8
3.2 Tahapan Penelitian	8
3.3 Pengambilan Data	9
3.4 Ekstraksi Teks	9
3.5 <i>Parsing Teks</i>	9
3.6 <i>Pembangunan Knowledge Base</i>	10
3.7 Integrasi <i>Large Language Model</i>	10
3.8 Penerapan <i>Prompting</i> Pada LLM	11
3.9 Evaluasi	11
IV HASIL DAN PEMBAHASAN	13
4.1 <i>Knowledge Base</i>	13
4.2 Konfigurasi <i>Prompting</i>	16
4.3 Hasil <i>Retrieval</i>	17
4.4 <i>History</i> dan <i>Interface Chatbot</i>	18
4.5 Hasil Evaluasi	20
V SIMPULAN DAN SARAN	25
5.1 Simpulan	25
5.2 Saran	25
DAFTAR PUSTAKA	26
RIWAYAT HIDUP	28

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR TABEL

1	Penelitian RAG terdahulu	5
2	Deksripsi jenis pertanyaan	12
3	Ayat dengan panjang lebih dari 85 kata	14
4	Contoh jawaban dari setiap jenis pertanyaan	21
5	Hasil evaluasi RAG	24

DAFTAR GAMBAR

1	Arsitektur <i>transformer</i>	3
2	Proses <i>naive retrieval augmented generation</i>	5
3	<i>Text searching</i> pada <i>elasticsearch</i>	6
4	Tahapan penelitian	8
5	Tahapan memuat data ke dalam <i>knowledge base</i>	10
6	Skema evaluasi	11
7	Histogram sebaran jumlah kata pada ayat	13
8	Data pada <i>knowledge base</i>	16
9	Penerapan <i>prompting</i>	17
10	Hasil <i>retrieved context</i>	18
11	<i>History</i> percakapan	18
12	<i>Activity diagram chatbot</i>	19
13	Tampilan <i>chatbot</i>	19
14	Diagram jenis pertanyaan	20