

**PENGEMBANGAN CHATBOT “AKASHA” BERBASIS WEB DENGAN
PENDEKATAN C-RAG UNTUK INFORMASI AKADEMIK UNDIKSHA**



Oleh:

I GEDE SUDIARTIKA

NIM. 2115101080

**PROGRAM STUDI ILMU KOMPUTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA**

2025

**PENGEMBANGAN CHATBOT “AKASHA” BERBASIS WEB
DENGAN PENDEKATAN C-RAG UNTUK INFORMASI
AKADEMIK UNDIKSHA**

SKRIPSI

Diajukan kepada

Universitas Pendidikan Ganesha

Untuk Memenuhi Salah Satu Persyaratan

Dalam Menyelesaikan Program Sarjana

Program Studi Ilmu Komputer

Oleh

I GEDE SUDIARTIKA

NIM 2115101080

**PROGRAM STUDI ILMU KOMPUTER
JURUSAN TEKNIK INFORMATIKA
FAKULTAS TEKNIK DAN KEJURUAN
UNIVERSITAS PENDIDIKAN GANESHA
SINGARAJA**

2025

**DIAJUKAN UNTUK MELENGKAPI TUGAS-TUGAS
DAN MEMENUHI SYARAT-SYARAT UNTUK
MENCAPAI GELAR SARJANA**

Menyetujui,

Pembimbing I,

A stylized, handwritten signature in blue ink, featuring a large, sweeping loop that starts from the left, goes up and over, and then comes back down to the right.

Ir. Ketut Resika Arthana, S.T., M.Kom.

NIP. 198412012012121002

Pembimbing II,

A handwritten signature in blue ink, consisting of a series of fluid, connected loops and strokes.

Ir. Ketut Agus Seputra, S.ST, M.T.

NIP. 199008152019031018

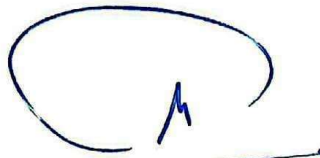
Skripsi oleh I Gede Sudiartika ini
telah dipertahankan di depan dewan penguji
pada tanggal 20 - 05 - 2025

Dewan Penguji,



I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 198910262019031004

(Ketua)



Ni Putu Novita Puspa Dewi, S.Kom., M.Cs
NIP. 199410032020122015

(Anggota)



I Ketut Resika Arthana, S.T., M.Kom.
NIP. 198412012012121002

(Anggota)



Ir. Ketut Agus Seputra, S.ST, M.T.
NIP. 199008152019031018

(Anggota)

Diterima oleh Panitia Ujian Fakultas Teknik dan Kejuruan
Universitas Pendidikan Ganesha
guna memenuhi syarat-syarat untuk mencapai gelar Sarjana Komputer

Pada:

Hari : Kamis
Tanggal : 11.0 JUL. 2020



Mengetahui

Ketua Ujian,

Made Windu Antara Kesiman, S.T., M.Sc., Ph.D.
NIP. 198211112008121001

Sekretaris Ujian,

I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs.
NIP. 198910262019031004

Mengesahkan,
Dekan Fakultas Teknik dan Kejuruan



Prof. Dr. Kadek Rihendra Dantes, S.T., M.T.
NIP. 197912012006041001

PERNYATAAN

Dengan ini saya menyatakan bahwa karya tulis yang berjudul “Pengembangan Chatbot “Akasha” Berbasis Web dengan Pendekatan C-RAG Untuk Informasi Akademik Undiksha” beserta seluruh isinya adalah benar-benar karya sendiri dan saya tidak melakukan penjiplakan dan pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung resiko/sanksi yang dijatuhkan kepada saya apabila dikemudian hari ditemukan adanya pelanggaran atas etika keilmuan dalam karya saya ini atau ada klaim terhadap keaslian karya saya ini.

Singaraja, 18 Februari 2025

Yang membuat pernyataan,



I Gede Sudiartika

NIM. 2115101080

PERNYATAAN PENGGUNAAN KECERDASAN BUATAN (AI)

Yang bertandatangan dibawah ini:

Nama : I Gede Sudiartika

NIM : 2115101080

Prodi : Ilmu Komputer

Fakultas : Fakultas Teknik dan Kejuruan (FTK)

Dengan ini menyatakan bahwa dalam penyusunan skripsi berjudul **“Pengembangan Chatbot “Akasha” Berbasis Web dengan Pendekatan C-RAG Untuk Informasi Akademik Undiksha”**, saya selaku penulis menggunakan kecerdasan buatan (AI) dalam bentuk:

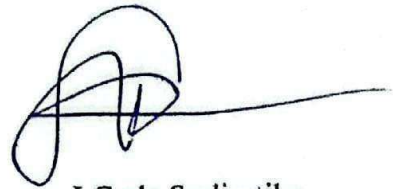
- **ChatGPT:** sebagai *brainstorming* alur kerja sistem RAG dan pengujian efektivitas *prompt*.
- **Scispace:** sebagai alat bantu pencarian artikel ilmiah terkait RAG
- **Claude:** sebagai alat bantu pengkodean pada *framework* LangChain dan LangGraph.
- **Windsurf:** ekstensi AI untuk Visual Studio Code agar mempercepat proses pengkodean dan *auto-complete*.

Penggunaan AI tersebut memenuhi prinsip-prinsip etika akademik dan **tidak menggantikan kontribusi intelektual saya sebagai penulis**. Semua hasil AI telah saya telaah, verifikasi, dan sesuaikan sesuai kebutuhan penelitian.

Demikian surat ini saya buat dengan sebenar-benarnya dan dapat digunakan sebagaimana mestinya.

Singaraja, 18 Pebruari 2025

Yang membuat pernyataan,

A handwritten signature in blue ink, consisting of a large, stylized 'A' shape with a horizontal line extending to the right.

I Gede Sudiartika

NIM. 2115101080



MOTTO

“Rencana tanpa laksana, hanya sekedar wacana”

PRAKATA

Puji syukur penyusun panjatkan ke hadapan Tuhan Yang Maha Esa, karena atas anugrah dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi ini yang berjudul **“PENGEMBANGAN CHATBOT “AKASHA” BERBASIS WEB DENGAN PENDEKATAN C-RAG UNTUK INFORMASI AKADEMIK UNDIKSHA”**. Skripsi ini ditujukan sebagai salah satu syarat untuk menyelesaikan Program Studi Strata Satu (S1) Ilmu Komputer Universitas Pendidikan Ganesha.

Dalam proses penyusunan skripsi ini tidak luput dari berbagai kekurangan, hambatan dan tantangan. Namun atas berkat dan tuntunan dari Ida Sang Hyang Widhi Wasa (Tuhan Yang Maha Esa) serta bantuan, saran, kritik dan kerjasama dari berbagai pihak, Proposal Penelitian ini dapat diselesaikan tepat waktu. Sebagai rasa hormat, penulis menyampaikan banyak terima kasih kepada:

1. Bapak Prof. Dr. I Wayan Lasmawan, M.Pd. selaku Rektor Universitas Pendidikan Ganesha.
2. Bapak Dr. Kadek Rihendra Dantes, S.T., M.T. selaku Dekan Fakultas Teknik dan Kejuruan.
3. Bapak I Ketut Resika Arthana, S.T., M.Kom., selaku dosen pembimbing I yang senantiasa membimbing dan memberikan saran kepada penulis.
4. Bapak Ir. Ketut Agus Seputra, S.T., M.T., selaku dosen pembimbing II yang turut serta memberikan saran dan masukan kepada penulis.
5. Bapak I Nyoman Saputra Wahyu Wijaya, S.Kom., M.Cs. selaku dosen penguji I yang turut serta memberikan arahan dan tuntunan kepada penulis.
6. Ibu Ni Putu Novita Puspa Dewi, S.Kom., M.Cs. selaku dosen penguji II yang turut serta memberikan arahan dan tuntunan kepada penulis.
7. Helpdesk UPA TIK Universitas Pendidikan Ganesha yang telah bersedia berkontribusi sebagai validator dalam penelitian ini.
8. Bapak I Made Tika dan Ibu Ni Wayan Sudiartini selaku orangtua penulis yang selalu mendukung dan memberikan semangat kepada penulis.
9. Ni Kadek Erna Yanti selaku adik dari penulis yang selalu memberi semangat kepada penulis.

10. Ni Komang Trisna Astiti selaku pacar dari penulis yang senantiasa memberikan semangat kepada penulis.

11. Rekan-rekan magang di UPA TIK Undiksha yang selalu mendukung dan memberikan semangat kepada penulis.

12. Serta semua pihak yang telah membantu memberikan saran dan dukungan kepada penulis yang tidak bisa disebutkan satu per satu.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna maka dari itu sangat diperlukan kritik dan saran dari berbagai pihak kepada penulis agar penelitian ini menjadi lebih baik lagi. Penulis berharap penelitian ini dapat berguna dan bermanfaat bagi kita semua. Akhir kata, penulis mengucapkan terimakasih.

Singaraja, 19 September 2024

(I Gede Sudiartika)



DAFTAR ISI

PRAKATA.....	iv
DAFTAR ISI	ix
DAFTAR TABEL.....	xi
DAFTAR GAMBAR	xii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan.....	5
1.4 Batasan Masalah Penelitian.....	5
1.5 Manfaat Penelitian.....	6
BAB II KAJIAN PUSTAKA DAN LANDASAN TEORI	8
2.1 Kajian Pustaka	8
2.2 Landasan Teori	10
2.2.1 Universitas Pendidikan Ganesha	10
2.2.2 Sumber Informasi Akademik Undiksha.....	11
2.2.3 Chatbot.....	12
2.2.4 <i>Large Language Model</i> (LLM).....	13
2.2.5 <i>Retrieval-Augmented Generation</i> (RAG)	14
2.2.6 Tools.....	23
2.2.7 <i>Retrieval Augmented Generation Assessment</i> (RAGAS)	28
2.2.8 <i>User Experience Questionnaire</i> (UEQ)	31
BAB III METODOLOGI PENELITIAN.....	34
3.1 Kerangka Penelitian	34
3.2 Identifikasi Permasalahan.....	34
3.2.1 Identifikasi Sumber Informasi Akademik Terdahulu.....	35
3.2.2 Identifikasi Permasalahan Pengguna	36
3.3 Tahap Pembuatan Model CRAG	40
3.3.1 Penyiapan Data	40
3.3.2 Perancangan	41
3.3.3 Pengkodean.....	46

3.4 Evaluasi	48
3.4.1 <i>Retrieval Evaluations</i>	49
3.4.2 <i>Generation Evaluations</i>	50
3.4.3 <i>User Experience Questionnaire (UEQ)</i>	51
BAB IV HASIL DAN PEMBAHASAN.....	54
4.1 Antarmuka Chatbot.....	54
4.2 Proses <i>Indexing</i>	60
4.2.1 Pengumpulan Dataset PDF	60
4.2.2 Pengumpulan <i>Scraping Link</i>	61
4.2.3 Ekstraksi dan Membuat Objek Document	63
4.2.4 <i>Chunking</i>	66
4.2.5 <i>Embedding</i>	67
4.3 Pendekatan <i>Corrective RAG</i>	69
4.3.1 <i>Agent State</i>	69
4.3.2 <i>Query Expansion Agent</i>	71
4.3.3 <i>Retrieval Agent</i>	74
4.3.4 <i>Corrective Agent</i>	75
4.3.5 <i>Generator Agent</i>	78
4.4 Alur <i>Graph</i>	82
4.5 Pembandingan <i>Naïve-RAG</i>	83
4.6 Evaluasi RAGAS.....	84
4.6.1 <i>Context Precision</i>	85
4.6.2 <i>Context Recall</i>	87
4.6.3 <i>Faithfulness</i>	89
4.6.4 <i>Response Relevancy</i>	90
4.7 Perbandingan <i>Corrective RAG</i> dengan <i>Naïve RAG</i>	92
4.8 <i>User Experience Questionnaire</i>	94
BAB V PENUTUP.....	100
5.1 Kesimpulan.....	100
5.2 Saran	102
DAFTAR PUSTAKA	104
LAMPIRAN.....	108
RIWAYAT HIDUP	124

DAFTAR TABEL

Tabel 2.1 Metode <i>Chunking</i> dan penerapannya	17
Tabel 3.1 Pertanyaan untuk survei kebutuhan chatbot.....	38
Tabel 3.2 Sumber data Chatbot Akademik Undiksha	41
Tabel 3.3 Kuisisioner Pengalaman Pengguna.....	51
Tabel 4.1 Detail key pada <i>State Global</i>	70
Tabel 4.2. Penjelasan aturan pada <i>Corrective Prompt</i>	76
Tabel 4.3. Penjelasan aturan pada <i>Generator Prompt</i>	79
Tabel 4.4 Skor <i>Context Precision</i>	86
Tabel 4.5 Skor <i>Context Recall</i>	88
Tabel 4.6 Skor <i>Faithfulness</i>	89
Tabel 4.7 Hasil Cosine Similiarity dari Pertanyaan asli dan pertanyaan buatan...	91
Tabel 4.8 Perbandingan <i>Corrective RAG</i> dengan <i>Naive RAG</i>	93
Tabel 4.9 Pertanyaan UEQ.....	95
Tabel 4.10 Hasil Survei UEQ.....	96

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Large Language Model (LLM)	13
Gambar 2.2 Proses Chunking.....	16
Gambar 2.3 <i>Embedding Model</i> pada RAG	19
Gambar 2.4 Proses Penyimpanan dan pencarian dalam <i>database vector</i>	20
Gambar 2.5 ChatGPT	22
Gambar 2.6 Logo Bahasa Pemrograman Python	23
Gambar 2.7 Proses <i>Scraping</i>	25
Gambar 3.1 Kerangka Penelitian Chatbot Akademik Undiksha	34
Gambar 3.2 Website Pedoman Studi Undiksha	35
Gambar 3.3 Kalender Akademik Undiksha	36
Gambar 3.4 Jumlah responden terhadap survei peneliti	39
Gambar 3.5 Grafik informasi akademik yang sering diakses mahasiswa.....	39
Gambar 3.6 Statistik tingkat kesulitan mahasiswa.....	40
Gambar 3.7 Alur Proses CRAG	43
Gambar 3.8 Arsitektur Multi-Agent pada Chatbot Akademik Undiksha.....	44
Gambar 3.9 Contoh penerapan LangGraph	46
Gambar 3.10 Pengkodean Node dan Edge pada framework LangGraph	47
Gambar 4.1 Antarmuka chatbot	54
Gambar 4.2 Menu Dashboard	55
Gambar 4.3 Tampilan Manajemen Dataset PDF	55
Gambar 4.4 Tampilan Manajemen Dataset Link	56
Gambar 4.5 Tampilan Menu Basis Data Vektor.....	56
Gambar 4.6 Hasil dari proses <i>chunking</i>	57
Gambar 4.7 Tampilan Menu <i>History Chat Process</i>	58
Gambar 4.8 Proses <i>embedding query</i> (pertanyaan).....	58
Gambar 4.9 Proses pengambilan <i>chunk (retrieval)</i>	59
Gambar 4.10 Proses koreksi (<i>corrective</i>).....	59
Gambar 4.11 Proses membuat jawaban akhir (<i>generate</i>)	60
Gambar 4.12. Sumber dataset PDF	61

Gambar 4.13 Kode program <i>scraping</i>	62
Gambar 4. 14 Kode untuk membersihkan karakter enter berlebih	62
Gambar 4.15 Kode program untuk menghapus karakter non-ASCII	62
Gambar 4.16 Kode program memuat metadata PDF dan Link.....	63
Gambar 4.17 Kode program.....	64
Gambar 4.18 Hasil dari ekstraksi file PDF di folder “dataset”	64
Gambar 4.19 Menambahkan metadata tahun dan sumber data.....	65
Gambar 4.20 Membuat objek Document pada dataset link	65
Gambar 4.21 Menggabungkan dataset PDF dan dataset link	66
Gambar 4.22 Kode program untuk melakukan <i>chunking</i>	66
Gambar 4. 23 Hasil <i>Chunking</i>	67
Gambar 4.24 <i>Embedding</i> dan membuat basis data vektor	68
Gambar 4.25 Basis data vektor yang dihasilkan	69
Gambar 4.26 <i>State Global</i> pada chatbot “AKASHA”	70
Gambar 4.27 Kamus singkatan dan istilah akademik Undiksha.....	72
Gambar 4.28 Membuat pola regex untuk mencari akronim	72
Gambar 4.29 Fungsi untuk memperluas pertanyaan.....	73
Gambar 4.30 Kelas <i>Query Expansion Agent</i>	73
Gambar 4.31 Proses <i>Query Expansion Agent</i>	74
Gambar 4.32 Proses <i>Retrieval Agent</i>	74
Gambar 4.33 <i>Corrective Agent Prompt</i>	75
Gambar 4.34 Proses <i>Corrective Agent</i>	78
Gambar 4.35 Hasil <i>Corrective Agent</i>	78
Gambar 4.36 <i>Generator Agent Prompt</i>	79
Gambar 4.37 Proses <i>Generator Prompt</i>	82
Gambar 4.38 Hasil <i>Generator Agent</i>	82
Gambar 4.39 Membuat alur <i>graph</i> menggunakan LangGraph	82
Gambar 4.40 Memandu Helpdesk UPA untuk validasi RAGAS	85
Gambar 4.41 Proses validasi <i>Context Precision</i>	86
Gambar 4.42 Daftar Semua <i>Chunk</i> Pada Basis Data Vektor.....	87
Gambar 4.43 Proses Validasi <i>Faithfulness</i>	89
Gambar 4.44 Diagram UEQ.....	99