UNIVERSITAS
GADJAH MADA
**Integrasi Large Language Models dalam Rancang Bangun Chatbot Berbasis Retrieval Augmented Generation
dan Cloud Computing untuk Efisiensi Akses Informasi pada Website Pendukung Akademik**
Yana Dayinta Nesthi, Ir. Yuris Mulya Saputra, S.T., M.Sc., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

DAFTAR PUSTAKA

[1] B. B. Irican, M. Sivri, V. Kokach, B. Kocaçınar, and F. P. Akbulut, "QBot: Domain-Specific Chatbots with Retrieval-Augmented Generation and Vector Embedding for Complex Documentation Queries," *2024 Innov. Intell. Syst. Appl. Conf. ASYU 2024*, 2024, doi: 10.1109/ASYU62119.2024.10757125.

[2] S. Pokhrel, B. K C, and P. B. Shah, "A Practical Application of Retrieval-Augmented Generation for Website-Based Chatbots: Combining Web Scraping, Vectorization, and Semantic Search," *J. Trends Comput. Sci. Smart Technol.*, vol. 6, no. 4, pp. 424–442, Jan. 2025, doi: 10.36548/jtcsst.2024.4.007.

[3] W. M. A. Dayata, "Carolinian Chatbot : End-to-End Question-Answer Pipeline on the University of San Carlos Policies , Procedures , and Guidelines Via Retrieval Augmented Generation," *2024 IEEE Int. Conf. Commun. Networks Satell.*, pp. 130–137, 2024, doi: 10.1109/COMNETSAT63286.2024.10862750.

[4] A. Šarčević, I. Tomičić, A. Merlin, and M. Horvat, "Enhancing Programming Education with Open-Source Generative AI Chatbots," *2024 47th ICT Electron. Conv. MIPRO 2024 - Proc.*, pp. 2051–2056, 2024, doi: 10.1109/MIPRO60963.2024.10569736.

[5] M. Maryamah, M. M. Irfani, E. B. Tri Raharjo, N. A. Rahmi, M. Ghani, and I. K. Raharjana, "Chatbots in Academia: A Retrieval-Augmented Generation Approach for Improved Efficient Information Access," *KST 2024 - 16th Int. Conf. Knowl. Smart Technol.*, pp. 259–264, 2024, doi: 10.1109/KST61284.2024.10499652.

[6] E. F. Trialopa, "Optimizing the Customer Service Performance with Using Chatbot by Utilizing Bot Accuracy, Journey Completion Rate, and Customer Satisfaction Score (CSAT) (Case Study: PT YAI)," UNIVERSITAS ISLAM INDONESIA YOGYAKARTA, 2022. [Online]. Available: https://dspace.uii.ac.id/handle/123456789/42045

[7] A. D. Vergaray, W. F. P. Robles, and J. A. S. Jiménez, "The Impact of

UNIVERSITAS
GADJAH MADA
**Integrasi Large Language Models dalam Rancang Bangun Chatbot Berbasis Retrieval Augmented Generation
dan Cloud Computing untuk Efisiensi Akses Informasi pada Website Pendukung Akademik**
Yana Dayinta Nesthi, Ir. Yuris Mulya Saputra, S.T., M.Sc., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

Chatbots on Customer Satisfaction: A Systematic Literature Review," *TEM J.*, vol. 12, no. 3, pp. 1407–1417, 2023, doi: 10.18421/TEM123-21.

[8] H. Tohir, N. Merlina, and M. Haris, "UTILIZING RETRIEVAL-AUGMENTED GENERATION IN LARGE LANGUAGE MODELS TO ENHANCE INDONESIAN LANGUAGE NLP," vol. 10, no. 2, pp. 352–360, 2024, doi: 10.33480/jitk.v10i2.5916.INTRODUCTION.

[9] A. K. Mohan, T. John, and P. S. Nikhil, "Contextual Conversational Intelligence : Leveraging Language Models For College Chatbot By Retrieval Augmented Generation," pp. 8–10, 2024, doi: 10.55041/IJSREM38284.

[10] I. Q. Huda and A. Prihanto, "Perbandingan Tahap Implementasi Dan Performa CHATBOT Telegram Dengan Platform Google DIALOGFLOW Dan Database POSTGRESQL," *J. Informatics Comput. Sci.*, vol. 5, no. 02, pp. 253–264, 2023, doi: 10.26740/jinacs.v5n02.p253-264.

[11] D. Baum, *Generative AI and LLMs for Dummies*. Snowflakes, 2024.

[12] R. Tayade *et al.*, "Comparison of Generative AI and Artificial Intelligence," vol. 8, no. 10, pp. 213–220, 2024.

[13] G. Rani, J. Singh, and A. Khanna, "Comparative Analysis of Generative AI Models," *2023 Int. Conf. Adv. Comput. Commun. Inf. Technol. ICAICCIT 2023*, pp. 760–765, 2023, doi: 10.1109/ICAICCIT60255.2023.10465941.

[14] B. Lin, "How a Decades-Old Technology and a Paper From Meta Created an AI Industry Standard ; Vector databases were an obscure technology until Pinecone and other startups turned it into a billion- dollar market amid the AI boom," pp. 1–5, [Online]. Available: https://www.proquest.com/blogs-podcasts-websites/how-decades-old-technology-paper-meta-created-ai/docview/3057375669/se-2?accountid=13771

[15] AWS Solution Library, "Guidance for Conversational Chatbots Using Retrieval Augmented Generation on AWS." Accessed: May 12, 2025. [Online]. Available: https://aws.amazon.com/solutions/guidance/conversational-chatbots-using-retrieval-augmented-generation-on-aws/?utm_source=chatgpt.com

UNIVERSITAS
GADJAH MADA
**Integrasi Large Language Models dalam Rancang Bangun Chatbot Berbasis Retrieval Augmented Generation
dan Cloud Computing untuk Efisiensi Akses Informasi pada Website Pendukung Akademik**
Yana Dayinta Nesthi, Ir. Yuris Mulya Saputra, S.T., M.Sc., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

[16] Google Cloud, "What is Retrieval-Augmented Generation (RAG)?" Accessed: May 12, 2025. [Online]. Available: https://cloud.google.com/use-cases/retrieval-augmented-generation?utm_source=chatgpt.com

[17] Google Cloud, "What are AI Hallucinations?" Accessed: May 12, 2025. [Online]. Available: https://cloud.google.com/discover/what-are-ai-hallucinations?hl=en

[18] AWS Solution Library, "How does Retrieval-Augmented Generation work?" Accessed: May 12, 2025. [Online]. Available: https://aws.amazon.com/what-is/retrieval-augmented-generation/?utm_source=chatgpt.com

[19] NVDIA, "Generative AI Chatbot Using RAG Technical Brief." Accessed: May 12, 2025. [Online]. Available: https://resources.nvidia.com/en-us-generative-ai-chatbot-workflow/knowledge-base-chatbot-technical-brief

[20] G. Sai Varshitha, C. Rupa, and D. Divya, "Real Time Blood Bank Communication Integrating Access Control Using XGBoost and Supabase," *2nd IEEE Int. Conf. Data Sci. Inf. Syst. ICDSIS 2024*, pp. 1–6, 2024, doi: 10.1109/ICDSIS61070.2024.10594708.

[21] n8n, "n8n Documentations." Accessed: Mar. 12, 2025. [Online]. Available: https://docs.n8n.io/?_gl=1*1q84ko2*_gcl_au*NTk0NzIzNDEzLjE3Mzg0OTE1NzA.*_ga*MTIyMjQ1MDk3OC4xNzQyNzA3MzIx*_ga_0SC4FF2FH9*MTc0MjgxMTQxMC4yLjEuMTc0MjgxMTQzNi4zNC4wLjA.

[22] G. Ower, "Not a programmer? You can mobilize data from biodiversity informatics APIs, too!," *Biodivers. Inf. Sci. Stand.*, vol. 6, pp. 6–9, 2022, doi: 10.3897/biss.6.93902.

[23] V. ALEXANDRIA, "USPTO ISSUES TRADEMARK : CHATWOOT," no. 6882688, pp. 1–2, 2022.

[24] K. Y. Chan, T. H. Yuen, and M. Co, "Using ChatGPT for medical education: the technical perspective," *BMC Med. Educ.*, vol. 25, no. 1, 2025, doi: 10.1186/s12909-025-06785-9.

[25] OpenAI, "OpenAI developer platform." Accessed: Apr. 21, 2025. [Online]. Available: https://platform.openai.com/docs/overview

UNIVERSITAS
GADJAH MADA
**Integrasi Large Language Models dalam Rancang Bangun Chatbot Berbasis Retrieval Augmented Generation
dan Cloud Computing untuk Efisiensi Akses Informasi pada Website Pendukung Akademik**
Yana Dayinta Nesthi, Ir. Yuris Mulya Saputra, S.T., M.Sc., Ph.D.
Universitas Gadjah Mada, 2025 | Diunduh dari http://etd.repository.ugm.ac.id/

[26] Schneider IT Management, "GPT-4o Mini in Azure: Performance Comparison - SCHNEIDER IT MANAGEMENT." [Online]. Available: https://www.schneider.im/microsoft-azure-ai-gpt-4o-mini-available/

[27] Artificial Analysis, "GPT-4o mini: Intelligence, Performance & Price Analysis." [Online]. Available: https://artificialanalysis.ai/models/gpt-4o-mini?utm_source=chatgpt.com