

DAFTAR PUSTAKA

- [1] PT Kereta Api Indonesia (Persero), “Laporan Tahunan dan Berkelanjutan,” 2023, [Online]. Available: www.kai.id
- [2] K. Perindustrian, “Indonesia Industry 4.0 Readiness Index,” *Kementrian Perindustrian Republik Indones.*, 2018.
- [3] V. M. Damanik, “Pengaruh Knowledge Sharing Terhadap Kinerja Karyawan Unit Commercial Passenger Marketing Pt Kereta Api Indonesia (Persero) Bandung,” 2020.
- [4] F. U. Ojika, W. O. Owobu, O. A. Abieba, O. J. Esan, B. C. Ubamadu, and A. I. Daraojimba, “AI-Enhanced Knowledge Management Systems: A Framework for Improving Enterprise Search and Workflow Automation through NLP and TensorFlow,” *Comput. Sci. IT Res. J.*, vol. 6, no. 3, pp. 201–230, Apr. 2025, doi: 10.51594/csitj.v6i3.1884.
- [5] K. Dong, I. Romanov, C. McLellan, and A. F. Esen, “Recent text-based research and applications in railways: A critical review and future trends,” *Eng. Appl. Artif. Intell.*, vol. 116, p. 105435, Nov. 2022, doi: 10.1016/j.engappai.2022.105435.
- [6] J. A. H. Alvaro and J. G. Barreda, “Advanced Engineering Informatics An advanced retrieval-augmented generation system for manufacturing quality control,” *Adv. Eng. Informatics*, vol. 64, no. March 2025, 2025, doi: 10.1016/j.aei.2024.103007.
- [7] Ototritas Jasa Keuangan, *Tata Kelola Kecerdasan Artifisial Perbankan Indonesia*. 2025.
- [8] KPMG, “The Future of AI Governance: The UAE Charter and Global Perspectives,” 2025.
- [9] X. Xu, H. Weytjens, D. Zhang, Q. Lu, I. Weber, and L. Zhu, “RAGOps: Operating and Managing Retrieval-Augmented Generation Pipelines,” Jun. 2025, [Online]. Available: <http://arxiv.org/abs/2506.03401>
- [10] A. Kimothi, *A Simple Guide to Retrieval Augmented Generation*. 2024.
- [11] R. Akkiraju *et al.*, “FACTS About Building Retrieval Augmented Generation-based Chatbots,” Jul. 2024, doi: <https://doi.org/10.48550/arXiv.2407.07858>.
- [12] A. Afzal, A. Kowsik, R. Fani, and F. Matthes, “Towards Optimizing and Evaluating a Retrieval Augmented QA Chatbot using LLMs with Human-in-the-Loop,” in *Proceedings of the Fifth Workshop on Data Science with Human-in-the-Loop (DaSH 2024)*, Association for Computational Linguistics, 2024, pp. 4–16. doi: 10.18653/v1/2024.dash-1.2.

- [13] Z. Li, Z. Wang, W. Wang, K. Hung, H. Xie, and F. Lee, “Retrieval-augmented generation for educational application: A systematic survey,” *Comput. Educ. Artif. Intell.*, vol. 8, no. May, p. 100417, 2025, doi: 10.1016/j.caeai.2025.100417.
- [14] M. Lindgren and L. Rådmann, “Comparative Analysis of Models for Document Question Answering using Retrieval-Augmented Generation,” Uppsala University, 2025.
- [15] S. Knollmeyer, O. Caymazer, and D. Grossmann, “Document GraphRAG: Knowledge Graph Enhanced Retrieval Augmented Generation for Document Question Answering Within the Manufacturing Domain,” *Electronics*, vol. 14, no. 11, p. 2102, May 2025, doi: 10.3390/electronics14112102.
- [16] Z. Wang, Z. Liu, W. Lu, and L. Jia, “Improving knowledge management in building engineering with hybrid retrieval-augmented generation framework,” *J. Build. Eng.*, vol. 103, p. 112189, Jun. 2025, doi: 10.1016/j.jobbe.2025.112189.
- [17] L.-C. Chen, M. S. Pardeshi, Y.-X. Liao, and K.-C. Pai, “Application of Retrieval-augmented Generation for Interactive Industrial Knowledge Management via a Large Language Model,” *Comput. Stand. Interfaces*, vol. 94, p. 103995, Aug. 2025, doi: 10.1016/j.csi.2025.103995.
- [18] L. Masanneck, S. G. Meuth, and M. Pawlitzki, “Evaluating base and retrieval augmented LLMs with document or online support for evidence based neurology,” *npj Digit. Med.*, vol. 8, no. 1, p. 137, Mar. 2025, doi: 10.1038/s41746-025-01536-y.
- [19] C. D. Yullinda, “Questions Answering System untuk Layanan Informasi Stunting Berbasis Open Source Language Model,” 2024.
- [20] G. J. B. Probst, S. Raub, and K. Romhardt, *Managing Knowledge: Building Blocks for Success*. 2000.
- [21] A. Thomas, “Digitally transforming the organization through knowledge management: a socio-technical system (STS) perspective,” *Eur. J. Innov. Manag.*, vol. 27, no. 9, pp. 437–460, 2024, doi: <https://doi.org/10.1108/EJIM-02-2024-0114>.
- [22] E. Tello-Leal, A. B. Rios-Alvarado, and A. Diaz-Manriquez, “A Semantic Knowledge Management System for Government Repositories,” in *2015 26th International Workshop on Database and Expert Systems Applications (DEXA)*, IEEE, Sep. 2015, pp. 168–172. doi: 10.1109/DEXA.2015.48.
- [23] A. Ait-Mlouk and L. Jiang, “KBot: A Knowledge Graph Based ChatBot for Natural Language Understanding Over Linked Data,” *IEEE Access*, vol. 8, pp. 149220–149230, 2020, doi: 10.1109/ACCESS.2020.3016142.
- [24] A. Gan *et al.*, “Retrieval Augmented Generation Evaluation in the Era of Large Language Models: A Comprehensive Survey,” 2025.

- [25] S. F. C. Haviana, M. A. Riyadi, and R. Kusumaningrum, "Evaluation of Chunking Strategies in RAG Application for Explicit Retrieval on Indonesian Language Scientific Papers," in *2025 12th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, 2025, pp. 59–65. doi: 10.1109/EECSI67060.2025.11290624.
- [26] D. Firdaus, I. Sumardi, and Y. Kulsum, "Integrating Retrieval-Augmented Generation with Large Language Model Mistral 7b for Indonesian Medical Herb," *J. Inform. Sunan Kalijaga*, vol. 9, no. 3, pp. 230–243, 2024, doi: <https://doi.org/10.14421/jiska.2024.9.3.230-243>.
- [27] S. Kahl *et al.*, "Evaluating the Impact of Advanced LLM Techniques on AI-Lecture Tutors for a Robotics Course," 2024, doi: 10.48550/arXiv.2408.04645.
- [28] B. Saha, U. Saha, and M. Malik, "QuIM-RAG: Advancing Retrieval-Augmented Generation With Inverted Question Matching for Enhanced QA Performance," *IEEE Access*, vol. 12, pp. 185401–185410, 2024, doi: 10.1109/ACCESS.2024.3513155.
- [29] S. Es, J. James, L. Espinosa Anke, and S. Schockaert, "RAGAs: Automated Evaluation of Retrieval Augmented Generation," in *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2024, pp. 150–158. doi: 10.18653/v1/2024.eacl-demo.16.
- [30] M. L. Husain, Y. Wibisono, and A. Anisyah, "Development of an Academic Services Chatbot Based on Retrieval-Augmented Generation (RAG)," *Brilliance*, vol. 5, no. 2, pp. 727–735, 2025, doi: <https://doi.org/10.47709/brilliance.v5i2.6719>.
- [31] G. \ Balakrishnan and A. Purwar, "Evaluating the Efficacy of Open-Source LLMs in Enterprise-Specific RAG Systems: A Comparative Study of Performance and Scalability".
- [32] "Evaluate your generative AI application locally with the Azure AI Evaluation SDK." [Online]. Available: <https://learn.microsoft.com/en-us/azure/ai-foundation/how-to/develop/evaluate-sdk>
- [33] M. B. Putra, M. R. Bisma Wiratara, S. V. Aliva, S. G. Rabiha, and G. F. Pakuning Desak, "Designing Knowledge Management System for Performance Efficiency Based Chatbot at PT.XYZ," in *2024 9th International Conference on Business and Industrial Research (ICBIR)*, IEEE, May 2024, pp. 1521–1526. doi: 10.1109/ICBIR61386.2024.10875920.
- [34] E. W. T. Ngai, M. C. M. Lee, M. Luo, P. S. L. Chan, and T. Liang, "An Intelligent Knowledge-based Chatbot for Customer Service," *Electron. Commer. Res. Appl.*, vol. 50, p. 101098, Nov. 2021, doi: 10.1016/j.elerap.2021.101098.

- [35] Y. Gao *et al.*, “Retrieval-Augmented Generation for Large Language Models: A Survey,” Mar. 2024, [Online]. Available: <http://arxiv.org/abs/2312.10997>
- [36] H. Yu, A. Gan, K. Zhang, S. Tong, Q. Liu, and Z. Liu, “Evaluation of Retrieval-Augmented Generation: A Survey,” Jul. 2024, doi: 10.1007/978-981-96-1024-2_8.
- [37] S. König, “Automated Evaluation and Parameter Optimisation for Retrieval Augmented Generation Systems,” Technische Universität Wien, 2025.
- [38] G. B and A. Purwar, “Evaluating the Efficacy of Open-Source LLMs in Enterprise-Specific RAG Systems: A Comparative Study of Performance and Scalability,” Jun. 2024, doi: <https://doi.org/10.48550/arXiv.2402.11281>.
- [39] M. F. Amin, “Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial,” *J. Eng. Res.*, vol. 6, no. 5, 2022.
- [40] Y.-H. Lee, Y.-H. Chen, and C.-Y. Guo, “A New Machine Learning Approach to Optimize Correlated Biomarkers,” *Commun. Stat. - Theory Methods*, vol. 54, no. 23, pp. 7515–7526, 2025, doi: 10.1080/03610926.2025.2477289.
- [41] J. J. Berman, “Understanding Your Data,” in *Data Simplification*, Elsevier, 2016, pp. 135–187. doi: 10.1016/B978-0-12-803781-2.00004-7.
- [42] B. Imeah, E. Penz, M. Rana, and C. Trask, “Economic analysis of new workplace technology including productivity and injury: The case of needle-less injection in swine,” *PLoS One*, vol. 15, no. 6, p. e0233599, Jun. 2020, doi: 10.1371/journal.pone.0233599.
- [43] M. A. Zailani *et al.*, “Drone versus ambulance for blood products transportation: an economic evaluation study,” *BMC Health Serv. Res.*, vol. 21, no. 1, p. 1308, Dec. 2021, doi: 10.1186/s12913-021-07321-3.
- [44] S. K. Ahmed *et al.*, “Using thematic analysis in qualitative research,” *J. Med. Surgery, Public Heal.*, vol. 6, p. 100198, Aug. 2025, doi: 10.1016/j.glmedi.2025.100198.
- [45] S. K. Ahmed *et al.*, “Using Thematic Analysis in Qualitative Research,” *J. Med. Surgery, Public Heal.*, vol. 6, p. 100198, Aug. 2025, doi: 10.1016/j.glmedi.2025.100198.
- [46] M. Vaziri, L. Mandel, C. Spiess, and M. Hirzel, “PDL: A Declarative Prompt Programming Language,” Oct. 2024, doi: <http://dx.doi.org/10.48550/arXiv.2410.19135>.
- [47] C. Pathade, “Red Teaming the Mind of the Machine: A Systematic Evaluation of Prompt Injection and Jailbreak Vulnerabilities in LLMs,” May 2025, [Online]. Available: <http://arxiv.org/abs/2505.04806>
- [48] “Faithfulness.” [Online]. Available:

<https://docs.ragas.io/en/v0.1.21/concepts/metrics/faithfulness.html>

- [49] “Answer Relevance.” [Online]. Available: https://docs.ragas.io/en/v0.1.21/concepts/metrics/answer_relevance.html
- [50] “Context Utilization.” [Online]. Available: https://docs.ragas.io/en/v0.1.21/concepts/metrics/context_utilization.html
- [51] “Faithfulness.” [Online]. Available: <https://www.confident-ai.com/docs/metrics/single-turn/faithfulness-metric>
- [52] “Answer Relevancy.” [Online]. Available: <https://www.confident-ai.com/docs/metrics/single-turn/answer-relevancy-metric>
- [53] “Contextual Relevancy.” [Online]. Available: <https://www.confident-ai.com/docs/metrics/single-turn/contextual-relevancy-metric>
- [54] “GroundednessEvaluator Class.” [Online]. Available: <https://learn.microsoft.com/en-us/python/api/azure-ai-evaluation/azure.ai.evaluation.groundednessevaluator?view=azure-python>
- [55] “RelevanceEvaluator Class.” [Online]. Available: <https://learn.microsoft.com/en-us/python/api/azure-ai-evaluation/azure.ai.evaluation.relevanceevaluator?view=azure-python>
- [56] “RetrievalEvaluator Class.” [Online]. Available: <https://learn.microsoft.com/en-us/python/api/azure-ai-evaluation/azure.ai.evaluation.retrievalevaluator?view=azure-python>
- [57] S. Barnett, S. Kurniawan, S. Thudumu, Z. Brannelly, and M. Abdelrazek, “Seven Failure Points When Engineering a Retrieval Augmented Generation System,” Jan. 2024, [Online]. Available: <http://arxiv.org/abs/2401.05856>
- [58] M. S. Omonov and Y. Ahn, “Towards Smart Public Administration: A TOE-Based Empirical Study of AI Chatbot Adoption in a Transitioning Government Context,” *Adm. Sci.*, 2025, doi: <https://doi.org/10.3390/admsci15080324>.
- [59] Z. Iba and A. Wardhana, *Metode Penelitian*, 1st ed. Purbalingga: Eureka Media Aksara, 2023.
- [60] M. S. B. Yusof, “ABC of Content Validation and Content Validity Index Calculation,” *Educ. Med. J.*, vol. 11, no. 2, pp. 49–54, 2019, doi: <https://doi.org/10.21315/eimj2019.11.2.6>.
- [61] S. C. Izah, L. Sylva, and M. Hait, “Cronbach’s Alpha: A Cornerstone in Ensuring Reliability and Validity in Environmental Health Assessment,” *ES Energy Environ.*, vol. 23, no. 1057, pp. 1–14, 2024, doi: <https://dx.doi.org/10.30919/esee1057>.

- [62] S. König, “Automated Evaluation of Parameter Optimisation for Retrieval Augmented Generation Systems,” 2025. doi: <https://doi.org/10.34726/hss.2025.122443>.
- [63] Q. H. Nguyen *et al.*, “Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil,” *Math. Probl. Eng.*, vol. 2021, pp. 1–15, Feb. 2021, doi: 10.1155/2021/4832864.
- [64] F. S. Nahm, “Receiver operating characteristic curve: overview and practical use for clinicians,” *Korean J. Anesthesiol.*, vol. 75, no. 1, pp. 25–36, Feb. 2022, doi: 10.4097/kja.21209.
- [65] B. Sarmah, B. Hall, R. Rao, S. Patel, S. Pasquali, and D. Mehta, “HybridRAG: Integrating Knowledge Graphs and Vector Retrieval Augmented Generation for Efficient Information Extraction,” Aug. 2024, [Online]. Available: <http://arxiv.org/abs/2408.04948>
- [66] Y. Du *et al.*, “Context Length Alone Hurts LLM Performance Despite Perfect Retrieval,” Oct. 2025, [Online]. Available: <http://arxiv.org/abs/2510.05381>
- [67] B. Chen, Z. Zhang, N. Langrené, and S. Zhu, “Unleashing the potential of prompt engineering for large language models,” *Patterns*, vol. 6, no. 6, p. 101260, Jun. 2025, doi: 10.1016/j.patter.2025.101260.