

DAFTAR PUSTAKA

- Ahmad Assegaf, I., Taufik Syastra, M., Kurniawan, R., Soleh Fajari, M., Novarini, R., Karim Harahap, A., Maulid, E., Irfayanti, Y., & Kurnia Wijayanti, E. (2024). Pengembangan Chatbot Konsultasi Kesehatan Mental Kesehatan Mental Berbasis Open Ai Model Gpt-3.5 Turbo Menggunakan Media Whatsapp. *Jurnal Informatika Teknologi Dan Sains*. <https://doi.org/https://doi.org/10.51401/jinteks.v6i4.4749>
- Ariantini, D. A. R., Lumenta, A. S. M., & Jacobus, A. (2016). Pengukuran Kemiripan Dokumen Teks Bahasa Indonesia Menggunakan Metode Cosine Similarity. *Jurnal Teknik Informatika*, 9(1). <https://doi.org/10.35793/jti.9.1.2016.13752>
- Bai, J., Bai, S., Chu, Y., Cui, Z., Dang, K., Deng, X., Fan, Y., Ge, W., Han, Y., Huang, F., Hui, B., Ji, L., Li, M., Lin, J., Lin, R., Liu, D., Liu, G., Lu, C., Lu, K., ... Zhu, T. (2023). *Qwen Technical Report*. <https://doi.org/https://doi.org/10.48550/arXiv.2309.16609>
- Bansal, A. (2023). Optimizing RAG with Hybrid Search and Contextual Chunking. *Journal of Engineering and Applied Sciences Technology*, 1–5. [https://doi.org/10.47363/JEAST/2023\(5\)E114](https://doi.org/10.47363/JEAST/2023(5)E114)
- Bhat, S. R., Rudat, M., Spiekermann, J., & Flores-Herr, N. (2025). *Rethinking Chunk Size For Long-Document Retrieval: A Multi-Dataset Analysis*. <https://doi.org/https://doi.org/10.48550/arXiv.2505.21700>
- 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, 2020-December*.
- Cao, H. (2024). *Recent advances in universal text embeddings: A Comprehensive Review of Top-Performing Methods on the MTEB Benchmark*.
- Chase, H. (2022). *LangChain*. <https://github.com/langchain-ai/langchain>
- Chen, B., Zhang, Z., Langrené, N., & Zhu, S. (2024). *Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review*.
- ChromaDB. (2025). *Concepts - Chroma Cookbook*. <https://cookbook.chromadb.dev/core/concepts/>
- Cormack, G. V., Clarke, C. L. A., & Buettcher, S. (2009). Reciprocal rank fusion outperforms condorcet and individual rank learning methods. *Proceedings - 32nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009*. <https://doi.org/10.1145/1571941.1572114>

- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1*, 4171–4186. <https://arxiv.org/abs/1810.04805v2>
- Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2023). RAGAS: Automated Evaluation of Retrieval Augmented Generation. *EACL 2024 - 18th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of System Demonstrations*, 150–158. <https://arxiv.org/abs/2309.15217v1>
- Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2025). *Ragas: Automated Evaluation of Retrieval Augmented Generation*. <https://doi.org/https://doi.org/10.48550/arXiv.2309.15217>
- Fatonah, F. R., Maylawati, D. S., & Nurlatifah, E. (2024). Chatbot Edukasi Pra-Nikah berbasis Telegram Menggunakan Bidirectional Encoder Representations From Transformers (BERT). *Jurnal Algoritma*, 21(2), 29–40. <https://doi.org/10.33364/algoritma/v.21-2.1657>
- Firdaus, V. A. H., Saputra, P. Y., & Suprianto, D. (2020). Intelligence chatbot for Indonesian law on electronic information and transaction. *IOP Conference Series: Materials Science and Engineering*, 830(2). <https://doi.org/10.1088/1757-899X/830/2/022089>
- Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2023a). *Retrieval-Augmented Generation for Large Language Models: A Survey*.
- Green, B. F., Wolf, A. K., Chomsky, C., & Laughery, K. (1961). Baseball: An automatic question-answerer. *Proceedings of the Western Joint Computer Conference: Extending Man's Intellect, IRE-AIEE-ACM 1961*, 219–224. <https://doi.org/10.1145/1460690.1460714>
- Guntoro, Loneli Costaner, & Lisnawita, L. (2020a). Aplikasi Chatbot untuk Layanan Informasi dan Akademik Kampus Berbasis Artificial Intelligence Markup Language (AIML). *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, 11(2), 291–300. <https://doi.org/10.31849/digitalzone.v11i2.5049>
- Gustavian Rahardjo, H. (2024). *Prompt Engineering GPT-3.5 Turbo untuk Ekstraksi Data Curriculum Vitae (CV) dalam Proses Pemilihan Calon Pekerja (Studi Kasus PT. NexWave)*.
- Hadi, M. U., Tashi, Q. Al, Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Hassan, S. Z., Shoman, M., Wu, J., Mirjalili, S., & Shah, M. (2025). Large Language Models: A Comprehensive Survey of its Applications, Challenges,

Limitations, and Future Prospects. *Authorea Preprints*.
<https://doi.org/10.36227/TECHRXIV.23589741.V8>

Harahap, A. M. (2020). Pengantar Hukum Ketenagakerjaan. In *Literasi Nusantara*.
<https://doi.org/10.2307/j.ctt1zqrn98.22>

Huang, Y., & Huang, J. (2024). *A Survey on Retrieval-Augmented Text Generation for Large Language Models*. <http://arxiv.org/abs/2404.10981>

Junaidi, Tumuhury, H., Saida, H. S. F., Abas, M., & Setiawan, R. (2023). Dasar Hukum Ketenagakerjaan. In *Pt Mafy Media Literasi Indonesia* (Vol. 44, Issue 8).

Kean Gao, A. (2023). Vec2Vec: A Compact Neural Network Approach for Transforming Text Embeddings with High Fidelity. *ArXiv Preprint*.

Kusumaningrum, R., Hanifah, A. F., Khadijah, K., Endah, S. N., & Sasongko, P. S. (2023a). Long Short-Term Memory for Non-Factoid Answer Selection in Indonesian Question Answering System for Health Information. *International Journal of Advanced Computer Science and Applications*, 14(2).
<https://doi.org/10.14569/IJACSA.2023.0140246>

Kusumaningrum, R., Hanifah, A. F., Khadijah, K., Endah, S. N., & Sasongko, P. S. (2023b). Long Short-Term Memory for Non-Factoid Answer Selection in Indonesian Question Answering System for Health Information. *International Journal of Advanced Computer Science and Applications*, 14(2).
<https://doi.org/10.14569/IJACSA.2023.0140246>

Lakatos, R., Pollner, P., Hajdu, A., & Joo, T. (2024). *Investigating the performance of Retrieval-Augmented Generation and fine-tuning for the development of AI-driven knowledge-based systems*. 1–18.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 2020-December*.

Li, J., Chen, J., Ren, R., Cheng, X., Zhao, W. X., Nie, J.-Y., & Wen, J.-R. (2024). *The Dawn After the Dark: An Empirical Study on Factuality Hallucination in Large Language Models* (Vol. 1). <https://doi.org/https://doi.org/10.48550/arXiv.2401.11817>

Lubis, A. T. U. BR., Harahap, N. S., Agustian, S., Irsyad, M., & Afrianty, I. (2024a). Question Answering System pada Chatbot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 955–964.
<https://doi.org/10.57152/malcom.v4i3.1378>

- Lubis, A. T. U. BR., Harahap, N. S., Agustian, S., Irsyad, M., & Afrianty, I. (2024b). Question Answering System pada Chatbot Telegram Menggunakan Large Language Models (LLM) dan Langchain (Studi Kasus UU Kesehatan). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 955–964. <https://doi.org/10.57152/malcom.v4i3.1378>
- Makruf, S., Sunan Gunung Djati Bandung, U., Yudha Pratama UIN Sunan Gunung Djati Bandung, B., Nur Muslimah UIN Sunan Gunung Djati Bandung, A., Ilham Pratama, M., & Shaleh, C. (2025). Teori Hukum dan Filsafat Hukum: Membangun Responsivitas terhadap Perubahan Ekonomi dan Sosial Corresponding Author. *Jurnal Ilmu Hukum & Ekonomi Syariah*, 10, 94–112. <https://doi.org/10.32505/muamalat.v10i1.9614>
- Mardiyati, S. (2023). Jaminan Hak Pekerja Yang Terkena Pemutusan Hubungan Kerja Dalam Perjanjian Kerja Waktu Tertentu. *Disiplin : Majalah Civitas Akademika Sekolah Tinggi Ilmu Hukum Sumpah Pemuda*, 29(1), 33–42. <https://doi.org/10.46839/DISIPLIN.V29I1.99>
- Matindas, C. L. (2018). Analisis Hukum Ketenagakerjaan Tentang Perlindungan Buruh/Pekerja Berdasarkan Undang-Undang Nomor 13 Tahun 2003. *Lex Privatum*, VII(3), 149–157.
- Matsumoto, N., Choi, H., Moran, J., Hernandez, M. E., Venkatesan, M., Li, X., Chang, J. H., Wang, P., & Moore, J. H. (2025). ESCARGOT: an AI agent leveraging large language models, dynamic graph of thoughts, and biomedical knowledge graphs for enhanced reasoning. *Bioinformatics*, 41(2), 2–5. <https://doi.org/10.1093/bioinformatics/btaf031>
- Mavroudis, V. (2024). *LangChain*. <https://doi.org/10.20944/preprints202411.0566.v1>
- Maybury, M. T. (2003). Toward a question answering roadmap. *New Directions in Question Answering*, 2003, 8–11.
- Millen Alhafidz, B., Rachmawati, E., & Eko Yunanto, P. (2025). *Sistem Question Answering pada Data Kesehatan Menggunakan Model pre-trained BERT*.
- Minaee, S., Mikolov, T., Nikzad, N., Chenaghlu, M., Socher, R., Amatriain, X., & Gao, J. (2024). *Large Language Models: A Survey*. <https://doi.org/http://arxiv.org/abs/2402.06196>
- Mucci, T. (2025). *What is Question Answering?* | IBM. <https://www.ibm.com/think/topics/question-answering>
- OpenAI. (2024). *New embedding models and API updates* . <https://openai.com/index/new-embedding-models-and-api-updates/>

- Petukhova, A., Matos-Carvalho, J. P., & Fachada, N. (2025). Text clustering with large language model embeddings. *International Journal of Cognitive Computing in Engineering*, 6, 100–108. <https://doi.org/10.1016/j.ijcce.2024.11.004>
- Prasetya, M. R. A., & Priyatno, A. M. (2022). Dice Similarity and TF-IDF for New Student Admissions Chatbot. *RIGGS: Journal of Artificial Intelligence and Digital Business*, 1(1). <https://doi.org/10.31004/riggs.v1i1.5>
- Radford Alec, Wu Jeffrey, Child Rewon, Luan David, Amodei Dario, & Sutskever Ilya. (2019). Language Models are Unsupervised Multitask Learners | Enhanced Reader. *OpenAI Blog*, 1(8).
- Rafifah, J., Angelina, R., Jumhana, E., & Fatima, N. A. (2025). Implementasi Dan Tantangan Perlindungan Hukum Terhadap Hak-Hak Pekerja Dalam Hubungan Kerja Di Indonesia. *Konsensus: Jurnal Ilmu Pertahanan, Hukum Dan Ilmu Komunikasi Volume 2 Nomor 1 Tahun 2025*, 2, 147–157. <https://doi.org/10.62383/konsensus.v2i1.626>
- Rakhmawati, N. A., Satriawan, I., Gunawan, Y., & Noviyani, R. F. (2023a). Development Of Indonesian Constitution Question Answering System. *UPB Scientific Bulletin, Series C: Electrical Engineering and Computer Science*, 85(4).
- Ramadhan, T. I., Supriatman, A., & Kurniawan, T. R. (2024). Passage Retrieval untuk Question Answering Bahasa Indonesia Menggunakan BERT dan FAISS. *Jurnal Algoritma*, 21(2), 156–163. <https://doi.org/10.33364/algoritma/v.21-2.2100>
- Ramos, J. (2003). *Using TF-IDF to Determine Word Relevance in Document Queries*.
- Robertson, S. (2004). Understanding inverse document frequency: On theoretical arguments for IDF. *Journal of Documentation*, 60(5), 503–520. <https://doi.org/10.1108/00220410410560582>
- Rosita, K., & Waluyo, W. (2023a). Hukum Ketenagakerjaan sebagai Instrumen Pelindung Bagi Tenaga Kerja dan Pengusaha dalam Penanganan Masalah Hubungan Kerja. *Jurnal Hukum Dan Pembangunan Ekonomi*, 11(1), 41. <https://doi.org/10.20961/hpe.v11i1.68677>
- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). *A Systematic Survey of Prompt Engineering in Large Language Models: Techniques and Applications*. <https://arxiv.org/abs/2402.07927v2>
- Samudra, G., Turmudi Zy, A., & Ermantor. (2025). Implementasi Retrieval Augmented Generation (RAG) Dalam. *JSAI: Journal Scientific and Applied Informatics*, 8(1). <https://doi.org/10.36085>

- Sasongko Wibowo, J., Februariyanti, H., & Listiyono, H. (2024). *Model Penjawab Pertanyaan Otomatis Berdasarkan Peringkat Relevansi Kalimat Menggunakan Model BERT* (Vol. 5, Issue 3). <https://doi.org/https://doi.org/10.30645/kesatria.v5i3.427>
- Septiani, D., & Isabela, I. (2022). *SINTESIA: Jurnal Sistem dan Teknologi Informasi Indonesia Analisis Term Frequency Inverse Document Frequency (Tf-Idf) Dalam Temu Kembali Informasi Pada Dokumen Teks*.
- Simmons, R. F. (1973). Semantic Networks: Their Computation and Use for Understanding English Sentences. *Computer Models of Thought and Language*, 63–113.
- Soepomo, I. (1999). Pengantar Hukum Perburuhan. In *Pengantar Hukum Perburuhan*.
- Șorecău, M., & Șorecău, E. (2023). An Alternative Application to ChatGPT that Uses Reliable Sources to Enhance the Learning Process. *International Conference Knowledge-Based Organization*, 29(3). <https://doi.org/10.2478/kbo-2023-0084>
- Sree Mala, C., Gezici, G., & Giannotti, F. (2025). *Hybrid Retrieval for Hallucination Mitigation in Large Language Models: A Comparative Analysis*. <https://doi.org/https://doi.org/10.48550/arXiv.2504.05324>
- Syah, M. I., Harahap, N. S., Novriyanto, & Sanjaya, S. (2024). *Penerapan Retrieval Augmented Generation Menggunakan Langchain dalam* (Vol. 6, Issue 2).
- Tribber, Y., Kusnadi, & Asfi, M. (2024). *Implementasi Retrieval Augmented Generation untuk Layanan Informasi Kampus dengan Chatbot Berbasis AI*. <https://doi.org/https://seminar.iaii.or.id/index.php/SISFOTEK/article/view/560>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems, 2017-December*.
- Voorhees, E. M. (1999). The TREC-8 Question Answering Track report. In *In: Proceedings of TREC-8* (Vol. 7, Issue 4, pp. 361–378). <https://doi.org/10.1017/S1351324901002789>
- Voorhees, E. M. (2003). Overview of the TREC 2003 Question Answering Track. *It: Proceedings of TREC'03*, 99–116.
- Wahidur, R. S. M., Kim, S., Choi, H., Bhatti, D. S., & Lee, H. N. (2025). Legal Query RAG. *IEEE Access*, 13, 36978–36994. <https://doi.org/10.1109/ACCESS.2025.3542125>
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E. H., Le, Q. V., & Zhou, D. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems*, 35.

- Wilensky, R., Chin, D. N., Luria, M., Martin, J., Mayfield, J., & Wu, D. (2000). The Berkeley UNIX consultant project. *Artificial Intelligence Review*, 14(1–2), 43–88. <https://doi.org/10.1023/a:1006500224529>
- Yang, A., Li, A., Yang, B., Zhang, B., Hui, B., Zheng, B., Yu, B., Gao, C., Huang, C., Lv, C., Zheng, C., Liu, D., Zhou, F., Huang, F., Hu, F., Ge, H., Wei, H., Lin, H., Tang, J., ... Qiu, Z. (2025). *Qwen3 Technical Report*. <https://doi.org/https://doi.org/10.48550/arXiv.2505.09388>
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). *A Survey of Large Language Models*. <https://doi.org/https://doi.org/10.48550/arXiv.2303.18223>