

DAFTAR PUSTAKA

- [1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is All You Need,” in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [2] P. Rasmussen, P. Paliychuk, T. Beauvais, J. Ryan, and D. Chalef, “Zep: A Temporal Knowledge Graph Architecture for Agent Memory,” in *arXiv preprint arXiv:2501.13956*, 2025.
- [3] A. Rogers, O. Kovaleva, and A. Rumshisky, “A Primer on Neural Network Models for Natural Language Processing,” *Journal of Artificial Intelligence Research*, vol. 61, pp. 65–95, 2020.
- [4] L.-A. Kaffee, R. Biswas, C. M. Keet, E. K. Vakaj, and G. de Melo, “Multilingual Knowledge Graphs and Low-Resource Languages: A Review,” *Transactions on Graph Data and Knowledge*, vol. 1, no. 1, pp. 10:1–10:19, 2023.
- [5] Z. Zhang, Y. Liu, X. Wang, and Z. Chen, “A Comprehensive Survey of Memory-Augmented Neural Networks for Natural Language Processing,” in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2023, pp. 3542–3565.
- [6] S. Borgeaud, A. Mensch, J. Hoffmann, T. Cai, E. Rutherford, K. Millican, G. van den Driessche, J.-B. Lespiau, B. Damoc, A. Clark *et al.*, “Improving Language Models by Retrieving from Trillions of Tokens,” in *International Conference on Machine Learning*. PMLR, 2022, pp. 2206–2240.
- [7] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, “Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks,” in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474.
- [8] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, M. Wang, and H. Wang, “Retrieval-Augmented Generation for Large Language Models: A Survey,” 2024.
- [9] V. Karpukhin, B. Oguz, S. Min, P. Lewis, L. Wu, S. Edunov, D. Chen, and W.-t. Yih, “Dense Passage Retrieval for Open-Domain Question Answering,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020, pp. 6769–6781.
- [10] A. Hogan, E. Blomqvist, M. Cochez, C. d’Amato, G. d. Melo, C. Gutierrez, S. Kirrane, J. E. L. Gayo, R. Navigli, S. Neumaier *et al.*, “Knowledge Graphs,” *ACM Computing Surveys*, vol. 54, no. 4, pp. 1–37, 2021.
- [11] B. Cai, Y. Xiang, L. Gao, H. Zhang, Y. Li, and J. Li, “Temporal Knowledge Graph Completion: A Survey,” in *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 2023, pp. 6545–6553.
- [12] R. Chen, “Retrieval-Augmented Generation with Knowledge Graphs: A Survey,” 2025.

- [13] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin, W. X. Zhao, Z. Wei, and J.-R. Wen, "A Survey on Large Language Model based Autonomous Agents," 2023.
- [14] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. Narasimhan, and Y. Cao, "ReAct: Synergizing Reasoning and Acting in Language Models," in *International Conference on Learning Representations*, 2023.
- [15] T. Schick, J. Dwivedi-Yu, R. Dessì, R. Raileanu, M. Lomeli, L. Zettlemoyer, N. Cancedda, and T. Scialom, "Toolformer: Language Models Can Teach Themselves to Use Tools," in *Advances in Neural Information Processing Systems*, vol. 36, 2023.
- [16] Y. Qin, S. Liang, Y. Ye, K. Zhu, L. Yan, Y. Lu, Y. Lin, X. Cong, X. Tang, B. Qian *et al.*, "Tool Learning with Foundation Models," 2023.
- [17] X. Zhao *et al.*, "AGENTiGraph: A Multi-Agent Knowledge Graph Framework for Interactive, Domain-Specific LLM Chatbots," 2025.
- [18] R. Zhao, S. Conia, E. Peng, M. Li, and S. Potdar, "AgREE: Agentic Reasoning for Knowledge Graph Completion on Emerging Entities," 2025.
- [19] J. Sun, C. Xu, L. Tang, S. Wang, C. Lin, Y. Gong, H.-Y. Guo, and N. Duan, "Think-on-Graph: Deep and Responsible Reasoning of Large Language Model on Knowledge Graph," in *The Twelfth International Conference on Learning Representations*, 2023.
- [20] J. Jiang, K. Zhou, Z. Dong, K. Ye, W. X. Zhao, and J.-R. Wen, "StructGPT: A General Framework for Large Language Model to Reason over Structured Data," in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, 2023, pp. 9237–9251.
- [21] B. Wilie, K. Vincentio, G. I. Winata, S. Cahyawijaya, X. Li, Z. Y. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar, and A. Purwarianti, "IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding," in *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*. Association for Computational Linguistics, 2020, pp. 843–857.
- [22] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, "IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics, 2020, pp. 757–770.
- [23] A. F. Aji, G. I. Winata, S. Cahyawijaya, F. Koto, I. G. W. Mahendra, J. V. D. Koto, A. Romadhony, K. Kurniawan, D. Moeljadi, E. M. Ponti, A. Purwarianti, and S. Ruder, "One Country, 700+ Languages: NLP Challenges for Underrepresented Languages of Indonesia," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, 2022, pp. 7226–7249.

- [24] M. D. Al Kautsar, A. Candra, M. A. Al Hakim, M. S. Kahfi, F. Koto, A. F. Aji, P. Limkonchotiwat, E. Chuangsuwanich, and G. I. Winata, "Seadialogues: A multilingual culturally grounded multi-turn dialogue dataset on southeast asian languages," *arXiv preprint arXiv:2508.07069*, 2025.
- [25] Gemini Team, "Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities," *arXiv preprint arXiv:2507.06261*, 2025.
- [26] Gemma Team, "Gemma 3 technical report," *arXiv preprint arXiv:2503.19786*, 2025.
- [27] Google DeepMind, "Embeddinggemma: Powerful and lightweight text representations," *arXiv preprint arXiv:2509.20354*, 2025.
- [28] N. Francis, A. Green, P. Guagliardo, L. Libkin, T. Lindaaker, V. Marsault, S. Plantikow, M. Rydberg, P. Selmer, and A. Taylor, "Cypher: An Evolving Query Language for Property Graphs," in *Proceedings of the 2018 International Conference on Management of Data*, 2018, pp. 1433–1445.
- [29] J. Wei, X. Wang, D. Schuurmans, M. Bosma, E. Chi, Q. Le, and D. Zhou, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," in *Advances in Neural Information Processing Systems*, vol. 35, 2022, pp. 24 824–24 837.
- [30] A. Maharana, D.-H. Lee, S. Tulyakov, M. Bansal, F. Barbieri, and Y. Fang, "Evaluating Very Long-Term Conversational Memory of LLM Agents," *arXiv preprint arXiv:2402.17753*, 2024.
- [31] A. Hadiifar *et al.*, "Lexical diversity for evaluation of ai-generated content," *arXiv preprint arXiv:2311.05252*, 2023.
- [32] P. Laban *et al.*, "Causal news corpus: Extracting causal relations from news," in *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC)*, 2022.
- [33] Z. Gekhman *et al.*, "Evaluating the long-term factuality of large language models," *arXiv preprint arXiv:2403.18802*, 2024.
- [34] M. Fernández-Pichel, J. C. Pichel, and D. E. Losada, "Evaluating search engines and large language models for answering health questions," *npj Digital Medicine*, vol. 8, no. 1, Mar. 2025. [Online]. Available: <http://dx.doi.org/10.1038/s41746-025-01546-w>
- [35] J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative agents: Interactive simulacra of human behavior," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, 2023, pp. 1–22.