

DAFTAR PUSTAKA

- [1] M. M. Henry, N. A. Heryanto, M. Isnan, K. S. Nugroho, and B. Pardamean, "Automatic Multiple Choice Question Generation: A Systematic Literature Review," *Proceeding 2024 9th Int. Conf. Inf. Technol. Digit. Appl. ICITDA 2024*, pp. 1–7, 2024, doi: 10.1109/ICITDA64560.2024.10810102.
- [2] H. F. Yusuf, I. Hidayah, and T. B. Adji, "Syntagmatic Distractor Generation for Vocabulary Multiple Choice Questions," *Proc. - Int. Conf. Educ. Technol. ICET*, pp. 94–99, 2023, doi: 10.1109/ICET59790.2023.10453362.
- [3] W. González Hernández, "The Teaching-Learning Process or the Teaching Process and the Learning Process," *Cult. Psychol.*, vol. 29, no. 1, pp. 96–115, 2023, doi: 10.1177/1354067X221097610.
- [4] J. Segone Dirane, "Teaching and Learning Using Technology," in *The Future of Television and Video Industry*, IntechOpen, 2024. doi: 10.5772/intechopen.114246.
- [5] Dr. Lohans Kumar Kalyani, "The Role of Technology in Education: Enhancing Learning Outcomes and 21st Century Skills," *Int. J. Sci. Res. Mod. Sci. Technol.*, vol. 3, no. 4, pp. 05–10, Apr. 2024, doi: 10.59828/ijrmst.v3i4.199.
- [6] N. V. Thakre, "LEARNING SYSTEM APPLICATION," *Gurukul Int. Multidiscip. Res. J.*, Jun. 2024, doi: 10.69758/GIMRJ2406I8V12P120.
- [7] H. Setiawan, I. Hidayah, and S. S. Kusumawardani, "Automatic Item Generation with Reading Passages: A Systematic Literature Review," *Proc. - Int. Conf. Educ. Technol. ICET*, vol. 2022–Octob, pp. 250–255, 2022, doi: 10.1109/ICET56879.2022.9990738.
- [8] J. Xie, N. Peng, Y. Cai, T. Wang, and Q. Huang, "Diverse Distractor Generation for Constructing High-Quality Multiple Choice Questions," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 30, pp. 280–291, 2022, doi: 10.1109/TASLP.2021.3138706.
- [9] Z. Qiu, X. Wu, and W. Fan, "Automatic Distractor Generation for Multiple Choice Questions in Standard Tests," in *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, Stroudsburg, PA, USA: International Committee on Computational Linguistics, 2020, pp. 2096–2106. doi: 10.18653/v1/2020.coling-main.189.
- [10] S. Jiang and J. Lee, "Distractor generation for chinese fill-in-the-blank items," *EMNLP 2017 - 12th Work. Innov. Use NLP Build. Educ. Appl. BEA 2017 - Proc. Work.*, pp. 143–148, 2017, doi: 10.18653/v1/w17-5015.
- [11] M. J. Gierl and H. Lai, "The Role of Item Models in Automatic Item Generation," *Int. J. Test.*, vol. 12, no. 3, pp. 273–298, Jul. 2012, doi: 10.1080/15305058.2011.635830.
- [12] Y. Hoshino, "Relationship between types of distractor and difficulty of multiple-choice Vocabulary tests in sentential context," *Lang. Test. Asia*, vol. 3, no. 1, pp. 1–14, 2013, doi: 10.1186/2229-0443-3-16.
- [13] S. Moore, E. Costello, H. A. Nguyen, and J. Stamper, "An Automatic Question Usability Evaluation Toolkit," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 14830 LNAI, pp. 31–46, 2024, doi: 10.1007/978-3-031-64299-9_3.
- [14] F. R. Ardiyanto, "The Influence of Students' Vocabulary Mastery Toward Their TOEFL Reading Comprehension Achievement," *Edudikara J. Pendidik. dan Pembelajaran*, vol. 9, no. 1, pp. 37–44, 2024, doi: 10.32585/edudikara.v9i1.357.

- [15] P. Pascasarjana, F. Teknik, and U. G. Mada, "PARADIGMATIC PADA ENGLISH VOCABULARY TEST," 2023.
- [16] C. Y. Yeung, J. Lee, and B. Tsou, "Difficulty-aware Distractor Generation for Gap-Fill Items," *Proc. 17th Annu. Work. Australas. Lang. Technol. Assoc.*, pp. 159–164, 2019, [Online]. Available: <https://pypi.org/project/pytorch-pretrained-bert/> <https://www.aclweb.org/anthology/U19-1021>
- [17] Y. Susanti, R. Iida, and T. Tokunaga, "Automatic generation of english Vocabulary tests," *CSEDU 2015 - 7th Int. Conf. Comput. Support. Educ. Proc.*, vol. 1, pp. 77–87, 2015, doi: 10.5220/0005437200770087.
- [18] Y. Susanti, T. Tokunaga, H. Nishikawa, and H. Obari, "Automatic distractor generation for multiple-choice English Vocabulary questions," *Res. Pract. Technol. Enhanc. Learn.*, vol. 13, no. 1, p. 15, Dec. 2018, doi: 10.1186/s41039-018-0082-z.
- [19] S. Murugan and B. Sadhu Ramakrishnan, "Automatic Morpheme-based Distractors Generation for Fill-in-the-Blank Questions using Listwise Learning-To-Rank Method for Agglutinative Language," *Eng. Sci. Technol. an Int. J.*, vol. 26, p. 100993, Feb. 2022, doi: 10.1016/j.jestch.2021.04.012.
- [20] S. Panda, F. P. Gomez, M. Flor, and A. Rozovskaya, "Automatic Generation of Distractors for Fill-in-the-Blank Exercises with Round-Trip Neural Machine Translation," in *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2022, pp. 391–401. doi: 10.18653/v1/2022.acl-srw.31.
- [21] T. Leslie and M. J. Gierl, "Using Automatic Item Generation to Create Multiple-Choice Questions for Pharmacy Assessment," *Am. J. Pharm. Educ.*, vol. 87, no. 10, p. 100081, Oct. 2023, doi: 10.1016/j.ajpe.2023.100081.
- [22] F. Falcão, D. M. Pereira, N. Gonçalves, A. De Champlain, P. Costa, and J. M. Pêgo, "A suggestive approach for assessing item quality, usability and validity of Automatic Item Generation," *Adv. Heal. Sci. Educ.*, vol. 28, no. 5, pp. 1441–1465, Dec. 2023, doi: 10.1007/s10459-023-10225-y.
- [23] M. J. Gierl, J. Shin, and T. Firoozi, "Automatic item generation," in *International Encyclopedia of Education (Fourth Edition)*, Elsevier, 2023, pp. 193–200. doi: 10.1016/B978-0-12-818630-5.10026-0.
- [24] S. E. Prasetyo, T. B. Adji, and I. Hidayah, "Automated Item Generation: Model and Development Technique," in *2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE)*, IEEE, Sep. 2020, pp. 64–69. doi: 10.1109/ICITACEE50144.2020.9239243.
- [25] I. Conference, R. Advances, and N. L. Processing, "Student Research Workshop associated with Recent Advances in Natural Language Processing," *Language (Baltim.)*, vol. 2021, no. Ranlp, pp. 7–9, 2011.
- [26] A. Killawala, I. Khokhlov, and L. Reznik, "Computational Intelligence Framework for Automatic Quiz Question Generation," in *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, IEEE, Jul. 2018, pp. 1–8. doi: 10.1109/FUZZ-IEEE.2018.8491624.
- [27] Y. A. Bachiri and H. Mouncif, "Applicable strategy to choose and deploy a MOOC platform with multilingual AQG feature," *Proc. - 2020 21st Int. Arab Conf. Inf. Technol. ACIT 2020*, 2020, doi: 10.1109/ACIT50332.2020.9300051.
- [28] B. Das, M. Majumder, S. Phadikar, and A. A. Sekh, "Automatic generation of fill-in-the-blank question with corpus-based distractors for e-assessment to enhance learning," *Comput. Appl. Eng. Educ.*, vol. 27, no. 6, pp. 1485–1495,

- [29] M. Liu, V. Rus, and L. Liu, "Automatic Chinese Multiple Choice Question Generation Using Mixed Similarity Strategy," *IEEE Trans. Learn. Technol.*, vol. 11, no. 2, pp. 193–202, Apr. 2018, doi: 10.1109/TLT.2017.2679009.
- [30] Y. Susanti, H. Nishikawa, T. Tokunaga, and O. Hiroyuki, "Item Difficulty Analysis of English Vocabulary Questions," in *Proceedings of the 8th International Conference on Computer Supported Education*, SCITEPRESS - Science and Technology Publications, 2016, pp. 267–274. doi: 10.5220/0005775502670274.
- [31] M. Wisnu Prabowo, I. Budi, and H. Budi Santoso, "Developing Question Generation System for Bahasa Indonesia Using Indonesian Standard Language Regulation," in *2021 10th International Conference on Software and Computer Applications*, New York, NY, USA: ACM, Feb. 2021, pp. 258–261. doi: 10.1145/3457784.3457824.
- [32] M. Moron, J. Scocoza, L. Chiruzzo, and A. Rosa, "A Tool for Automatic Question Generation for Teaching English to Beginner Students," in *2021 40th International Conference of the Chilean Computer Science Society (SCCC)*, IEEE, Nov. 2021, pp. 1–5. doi: 10.1109/SCCC54552.2021.9650423.
- [33] P. R. Rao, T. N. Jhavar, Y. A. Kachave, and V. Hirlekar, "Generating QA from Rule-based Algorithms," in *2022 International Conference on Electronics and Renewable Systems (ICEARS)*, IEEE, Mar. 2022, pp. 1697–1703. doi: 10.1109/ICEARS53579.2022.9751723.
- [34] P. Shuai, L. Li, S. Liu, and J. Shen, "QDG: A unified model for automatic question-distractor pairs generation," *Appl. Intell.*, vol. 53, no. 7, pp. 8275–8285, Apr. 2023, doi: 10.1007/s10489-022-03894-6.
- [35] C. Zhang, Y. Sun, H. Chen, and J. Wang, "Generating Adequate Distractors for Multiple-Choice Questions," in *Proceedings of the 12th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, SCITEPRESS - Science and Technology Publications, 2020, pp. 310–315. doi: 10.5220/0010148303100315.
- [36] L. Wang, D. Sheng, and H.-T. Zheng, "Question Generation via Multi-stage Answers Editing Network," in *2021 International Joint Conference on Neural Networks (IJCNN)*, IEEE, Jul. 2021, pp. 1–8. doi: 10.1109/IJCNN52387.2021.9534027.
- [37] B. Leite, H. L. Cardoso, L. P. Reis, and C. Soares, "Factual Question Generation for the Portuguese Language," in *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, Aug. 2020, pp. 1–7. doi: 10.1109/INISTA49547.2020.9194631.
- [38] A. Kilgarriff and D. Tugwell, "Word sketch: Extraction and display of significant collocations for lexicography," *Proc. ACL Work. Collocation Comput. Extr. Anal. Exploit.*, pp. 32–38, 2001, [Online]. Available: <http://www.kilgarriff.co.uk/Publications/2001-KilgTugwell-ACLcollos-Sketches.pdf>
- [39] C.-H. Lee, T.-Y. Chen, L.-P. Chen, P.-C. Yang, and R. T.-H. Tsai, "Automatic Question Generation from Children's Stories for Companion Chatbot," in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, IEEE, Jul. 2018, pp. 491–494. doi: 10.1109/IRI.2018.00078.
- [40] M. Setyowati, A. K. Latifa, E. Pratiwi, and S. Mabagits, "Student's Vocabulary Mastery on TOEFL Test: Does It Correlate with Reading Comprehension?," vol. 434, no. Iconelt 2019, pp. 245–249, 2020, doi: 10.2991/assehr.k.200427.048.

- [41] D. D. Qian, "Investigating the Relationship Between *Vocabulary* Knowledge and Academic Reading Performance: An Assessment Perspective," *Lang. Learn.*, vol. 52, no. 3, pp. 513–536, Sep. 2002, doi: 10.1111/1467-9922.00193.
- [42] W. Tampubolon, F. Sinamo, and E. Saragih, "Vocabulary Assesment Strategy," *Proj. (Professional J. English Educ.*, vol. 5, no. 2, p. 403, 2022, doi: 10.22460/project.v5i2.p403-413.
- [43] A. Ghio, "Achile : un dispositif de décodage acoustico-phonétique et d'identification lexicale indépendant du locuteur à partir de modules mixtes Alain Ghio HAL Id : tel-01663493," no. November 1997, 2018.
- [44] D. Trajanov *et al.*, "Review of Natural Language Processing in Pharmacology," *Pharmacol. Rev.*, vol. 75, no. 4, pp. 714–738, 2023, doi: 10.1124/pharmrev.122.000715.
- [45] D. Khurana, A. Koli, K. Khatte, and S. Singh, "Natural language processing: state of the art, current trends and challenges," *Multimed. Tools Appl.*, vol. 82, no. 3, pp. 3713–3744, Jan. 2023, doi: 10.1007/s11042-022-13428-4.
- [46] R. Patil, S. Boit, V. Gudivada, and J. Nandigam, "A Survey of Text Representation and Embedding Techniques in NLP," *IEEE Access*, vol. 11, pp. 36120–36146, 2023, doi: 10.1109/ACCESS.2023.3266377.
- [47] S. Pais, J. Cordeiro, and M. L. Jamil, "NLP-based platform as a service: a brief review," *J. Big Data*, vol. 9, no. 1, p. 54, Dec. 2022, doi: 10.1186/s40537-022-00603-5.
- [48] B. Das, M. Majumder, S. Phadikar, and A. A. Sekh, "Multiple-choice question generation with auto-generated *distractors* for computer-assisted educational assessment," *Multimed. Tools Appl.*, vol. 80, no. 21–23, pp. 31907–31925, Sep. 2021, doi: 10.1007/s11042-021-11222-2.
- [49] P. Baldwin *et al.*, "A Natural-Language-Processing-Based Procedure for Generating *Distractors* for Multiple-Choice Questions," *Eval. Health Prof.*, vol. 45, no. 4, pp. 327–340, Dec. 2022, doi: 10.1177/01632787211046981.
- [50] K. K. Maurya and M. S. Desarkar, "Learning to Distract: A Hierarchical Multi-Decoder Network for Automated Generation of Long *Distractors* for Multiple-Choice Questions for Reading Comprehension," in *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, New York, NY, USA: ACM, Oct. 2020, pp. 1115–1124. doi: 10.1145/3340531.3411997.
- [51] X. Zhou, S. Luo, and Y. Wu, "Co-Attention Hierarchical Network: Generating Coherent Long *Distractors* for Reading Comprehension," *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 05, pp. 9725–9732, Apr. 2020, doi: 10.1609/aaai.v34i05.6522.
- [52] S. Fyffe, P. Lee, and S. Kaplan, "'Transforming' Personality Scale Development: Illustrating the Potential of State-of-the-Art Natural Language Processing," *Organ. Res. Methods*, vol. 27, no. 2, pp. 265–300, Apr. 2024, doi: 10.1177/10944281231155771.
- [53] Microsoft, "Concepts of ai and machine learning in azure: Language models," [learn.microsoft.com](https://learn.microsoft.com/en-us/azure/aks/concepts-ai-ml-language-models). [Online]. Available: <https://learn.microsoft.com/en-us/azure/aks/concepts-ai-ml-language-models>
- [54] "Fundamental Text Processing," in *Language Intelligence*, Wiley, 2025, pp. 31–61. doi: 10.1002/9781394297290.ch3.
- [55] K. Vayadande, A. M. Kulkarni, G. B. Yadav, R. Kumar, and A. R. Sawant, "A Comprehensive Analysis of Various Tokenization Techniques and Sequence-to-Sequence Model in Natural Language Processing," in *How Machine Learning is Innovating Today's World*, Wiley, 2024, pp. 1–11. doi: 10.1002/9781394214167.ch1.

- [56] P. Vadlapati, "Tokenization Beyond NLP: Potential Applications in Data Analytics, Cybersecurity, and Beyond," *INTERANTIONAL J. Sci. Res. Eng. Manag.*, vol. 08, no. 12, pp. 1–7, Dec. 2024, doi: 10.55041/IJSREM9532.
- [57] J. Singh and V. Gupta, "Text Stemming," *ACM Comput. Surv.*, vol. 49, no. 3, pp. 1–46, Sep. 2017, doi: 10.1145/2975608.
- [58] J.-M. Torres-Moreno, "Beyond Stemming and Lemmatization: Ultra-stemming to Improve Automatic Text Summarization," 2012, [Online]. Available: <http://arxiv.org/abs/1209.3126>
- [59] X. Y. Tan and C. W. Tan, "From Complexity to Clarity: Tatabahasa-Centric Lemmatization in Malay Texts," in *2024 3rd International Conference on Digital Transformation and Applications (ICDXA)*, IEEE, Jan. 2024, pp. 1–5. doi: 10.1109/ICDXA61007.2024.10470673.
- [60] A. K. Ojo, "Text Preprocessing Techniques," in *Text Mining Techniques with Applications, Edition 1*, BP International, 2024, pp. 76–87. doi: 10.9734/bpi/mono/978-81-972870-5-3/CH7.
- [61] S. Park, J.-H. Song, and Y. Kim, "A Neural Language Model for Multi-Dimensional Textual Data based on CNN-LSTM Network," in *2018 19th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*, IEEE, Jun. 2018, pp. 212–217. doi: 10.1109/SNPD.2018.8441130.
- [62] A. Vaswani *et al.*, "Attention Is All You Need," Aug. 2023, [Online]. Available: <http://arxiv.org/abs/1706.03762>
- [63] I. M. W. P. Winarta, S. A. I. Alfarozi, and I. Hidayah, "Atomic Evaluation using Large Language Model for Automated Essay Exam Scoring," in *2024 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT)*, IEEE, Nov. 2024, pp. 21–27. doi: 10.1109/COMNETSAT63286.2024.10862459.
- [64] W. X. Zhao *et al.*, "A Survey of Large Language Models," Mar. 2025, [Online]. Available: <http://arxiv.org/abs/2303.18223>
- [65] A. Grattafiori *et al.*, "The Llama 3 Herd of Models," Nov. 2024, [Online]. Available: <http://arxiv.org/abs/2407.21783>
- [66] Z. Lu *et al.*, "Small Language Models: Survey, Measurements, and Insights," Feb. 2025, [Online]. Available: <http://arxiv.org/abs/2409.15790>
- [67] F. Wang *et al.*, "A Comprehensive Survey of Small Language Models in the Era of Large Language Models: Techniques, Enhancements, Applications, Collaboration with LLMs, and Trustworthiness," Dec. 2024, [Online]. Available: <http://arxiv.org/abs/2411.03350>
- [68] B. Bergner, A. Skliar, A. Royer, T. Blankevoort, Y. Asano, and B. E. Bejnordi, "Input prompt," 2023.
- [69] L. Chen and G. Varoquaux, "What is the Role of Small Models in the LLM Era: A Survey," 2024, [Online]. Available: <http://arxiv.org/abs/2409.06857>
- [70] T. B. Brown *et al.*, "Language Models are Few-Shot Learners," Jul. 2020, [Online]. Available: <http://arxiv.org/abs/2005.14165>
- [71] OpenAI *et al.*, "GPT-4 Technical Report," Mar. 2024, [Online]. Available: <http://arxiv.org/abs/2303.08774>
- [72] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing," Jul. 2021, [Online]. Available: <http://arxiv.org/abs/2107.13586>
- [73] T. Wu *et al.*, "PromptChainer: Chaining Large Language Model Prompts through Visual Programming," Mar. 2022, [Online]. Available:

- [74] M. McTear and M. Ashurkina, "Advanced Prompt Engineering," in *Transforming Conversational AI*, Berkeley, CA: Apress, 2024, pp. 115–144. doi: 10.1007/979-8-8688-0110-5_6.
- [75] T. Z. Zhao, E. Wallace, S. Feng, D. Klein, and S. Singh, "Calibrate Before Use: Improving Few-Shot Performance of Language Models," Jun. 2021, [Online]. Available: <http://arxiv.org/abs/2102.09690>
- [76] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large Language Models are Zero-Shot Reasoners," Jan. 2023, [Online]. Available: <http://arxiv.org/abs/2205.11916>
- [77] A. Scius-Bertrand, M. Jungo, L. Vögtlin, J.-M. Spat, and A. Fischer, "Zero-Shot Prompting and Few-Shot Fine-Tuning: Revisiting Document Image Classification Using Large Language Models," in *Pattern Recognition*, A. Antonacopoulos, S. Chaudhuri, R. Chellappa, C.-L. Liu, S. Bhattacharya, and U. Pal, Eds., Cham: Springer Nature Switzerland, 2025, pp. 152–166.
- [78] J. Wei *et al.*, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," *Adv. Neural Inf. Process. Syst.*, vol. 35, 2022.
- [79] P. Lewis *et al.*, "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Adv. Neural Inf. Process. Syst.*, vol. 2020-Decem, no. NeurIPS, 2020.
- [80] L. Ouyang *et al.*, "Training language models to follow instructions with human feedback," Mar. 2022, [Online]. Available: <http://arxiv.org/abs/2203.02155>
- [81] V. Zhelezniak, A. Savkov, A. Shen, and N. Y. Hammerla, "Correlation Coefficients and Semantic Textual Similarity," May 2019, [Online]. Available: <http://arxiv.org/abs/1905.07790>
- [82] Alexander S. Gillis, "Definition proof of concept (POC)," 2023. [Online]. Available: <https://www.techtarget.com/searchcio/definition/proof-of-concept-POC>
- [83] B. Rogers, "Peterson's Toefl Practice Tests," Colorado: Petersons, 2001.
- [84] Pyle, "Cliffs TOEFL Prepratation Guide," 5th ed, Ed., Nebraska: Cliff Notes, 1995.
- [85] B. Rogers, *The Complete Guide To The TOEFL® Test PBT Edition Answer Key and Audio Scripts*, 1st ed. Massachusetts: Heinle ELT, 2011.
- [86] P. J. Sharpen, "Baron's Practice Exercises for TOEFL," 6th ed., New York: Barron's Educational Series, 2009.
- [87] D. Phillips, "Longman Complete Course for the TOEFL Test Preparation for the Computer and Paper Test," 2nd ed., Pearson College Div, 2001, p. 682.
- [88] S. Abdaljalil, H. Kurban, P. Sharma, E. Serpedin, and R. Atat, "SINdex: Semantic INconsistency Index for Hallucination Detection in LLMs," 2025, [Online]. Available: <http://arxiv.org/abs/2503.05980>
- [89] M. R. LYNN, "Determination and Quantification Of Content Validity," *Nurs. Res.*, vol. 35, no. 6, p. 382-386, Nov. 1986, doi: 10.1097/00006199-198611000-00017.