

DAFTAR PUSTAKA

- [1] C. Wu, Y. Pan, H. Wu, and L. Ning, “Integrating Speech Recognition into Intelligent Information Systems: From Statistical Models to Deep Learning,” *Informatics*, vol. 12, no. 4, p. 107, Oct. 2025, doi: 10.3390/informatics12040107.
- [2] D. Reddy Yerramreddy, J. Marasani, P. S. Venkata Gowtham, G. Harshit, and Anjali, “Speech Recognition Paradigms: A Comparative Evaluation of SpeechBrain, Whisper and Wav2Vec2 Models,” in *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Pune, India: IEEE, Apr. 2024, pp. 1–6. doi: 10.1109/I2CT61223.2024.10544133.
- [3] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust speech recognition via large-scale weak supervision,” in *Proceedings of the 40th International Conference on Machine Learning*, in ICML’23. JMLR.org, 2023. <https://cdn.openai.com/papers/whisper.pdf>.
- [4] H. Kheddar, M. Hemis, and Y. Himeur, “Automatic speech recognition using advanced deep learning approaches: A survey,” *Inf. Fusion*, vol. 109, p. 102422, Sep. 2024, doi: 10.1016/j.inffus.2024.102422.
- [5] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, “wav2vec 2.0: a framework for self-supervised learning of speech representations,” in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, in NIPS ’20. Red Hook, NY, USA: Curran Associates Inc., 2020. <https://dl.acm.org/doi/abs/10.5555/3495724.3496768>.
- [6] F. Filippidou and L. Moussiades, “A Benchmarking of IBM, Google and Wit Automatic Speech Recognition Systems,” in *Artificial Intelligence Applications and Innovations*, vol. 583, I. Maglogiannis, L. Iliadis, and E. Pimenidis, Eds., in IFIP Advances in Information and Communication Technology, vol. 583. , Cham: Springer International Publishing, 2020, pp. 73–82. doi: 10.1007/978-3-030-49161-1_7.
- [7] M. Malik, M. K. Malik, K. Mehmood, and I. Makhdoom, “Automatic speech recognition: a survey,” *Multimed. Tools Appl.*, vol. 80, no. 6, pp. 9411–9457, Mar. 2021, doi: 10.1007/s11042-020-10073-7.
- [8] J. Li, “Recent Advances in End-to-End Automatic Speech Recognition,” *APSIPA Trans. Signal Inf. Process.*, vol. 11, no. 1, pp. 1–64, Apr. 2022, doi: 10.1561/116.000000050.
- [9] R. Prabhavalkar, T. Hori, T. N. Sainath, R. Schlüter, and S. Watanabe, “End-to-End Speech Recognition: A Survey,” *IEEEACM Trans. Audio Speech Lang. Process.*, vol. 32, pp. 325–351, 2024, doi: 10.1109/TASLP.2023.3328283.
- [10] A. Mohamed *et al.*, “Self-Supervised Speech Representation Learning: A Review,” *IEEE J. Sel. Top. Signal Process.*, vol. 16, no. 6, pp. 1179–1210, Oct. 2022, doi: 10.1109/JSTSP.2022.3207050.
- [11] S. Di Leo, L. De Cicco, and S. Mascolo, “Real-Time Speech-to-Text on Edge: A Prototype System for Ultra-Low Latency Communication with AI-Powered NLP,” *Information*, vol. 16, no. 8, p. 685, Aug. 2025, doi: 10.3390/info16080685.

- [12] Z. Kozhirkbayev, “Kazakh Speech Recognition: Wav2vec2.0 vs. Whisper,” *J. Adv. Inf. Technol.*, vol. 14, no. 6, pp. 1382–1389, 2023, doi: 10.12720/jait.14.6.1382-1389.
- [13] J. C. Vásquez-Correa and A. Álvarez Muniain, “Novel Speech Recognition Systems Applied to Forensics within Child Exploitation: Wav2vec2.0 vs. Whisper,” *Sensors*, vol. 23, no. 4, p. 1843, Feb. 2023, doi: 10.3390/s23041843.
- [14] C. Özkurt, “Investigation of the Effectiveness of Audio Processing and Filtering Strategies in Noisy Environments on Speech Recognition Performance,” Feb. 22, 2024, *In Review*. doi: 10.21203/rs.3.rs-3973856/v1.
- [15] P. Eickhoff, M. Möller, T. P. Rosin, J. Twiefel, and S. Wermter, “Bring the Noise: Introducing Noise Robustness to Pretrained Automatic Speech Recognition,” in *Artificial Neural Networks and Machine Learning – ICANN 2023*, vol. 14260, L. Iliadis, A. Papaleonidas, P. Angelov, and C. Jayne, Eds., in Lecture Notes in Computer Science, vol. 14260, Cham: Springer Nature Switzerland, 2023, pp. 376–388. doi: 10.1007/978-3-031-44195-0_31.
- [16] A. Pervaiz *et al.*, “Incorporating Noise Robustness in Speech Command Recognition by Noise Augmentation of Training Data,” *Sensors*, vol. 20, no. 8, p. 2326, Apr. 2020, doi: 10.3390/s20082326.
- [17] H. Kheddar, Y. Himeur, S. Al-Maadeed, A. Amira, and F. Bensaali, “Deep transfer learning for automatic speech recognition: Towards better generalization,” *Knowl.-Based Syst.*, vol. 277, p. 110851, Oct. 2023, doi: 10.1016/j.knosys.2023.110851.
- [18] P. Netinant, R. Fooprateepsiri, A. Rukhiran, and M. Rukhiran, “An AIoT-Based Framework for Automated English-Speaking Assessment: Architecture, Benchmarking, and Reliability Analysis of Open-Source ASR,” *Informatics*, vol. 13, no. 2, p. 19, Jan. 2026, doi: 10.3390/informatics13020019.
- [19] E. Kumalija and Y. Nakamoto, “Performance evaluation of automatic speech recognition systems on integrated noise-network distorted speech,” *Front. Signal Process.*, vol. 2, p. 999457, Sep. 2022, doi: 10.3389/frsip.2022.999457.
- [20] M. S. Islam Ridoy, S. Akter, and Md. A. Rahman, “Adaptability of ASR Models on Low-Resource Language: A Comparative Study of Whisper and Wav2Vec-BERT on Bangla,” in *2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM)*, Gazipur, Bangladesh: IEEE, Jun. 2025, pp. 1–6. doi: 10.1109/NCIM65934.2025.11159987.
- [21] D. Moser, N. Stanic, and M. Sariyar, “Benchmarking speech-to-text robustness in noisy emergency medical dialogues: an evaluation of models under realistic acoustic conditions,” *JAMIA Open*, vol. 8, no. 6, p. ooaf147, Nov. 2025, doi: 10.1093/jamiaopen/ooaf147.
- [22] A. Trabelsi, L. Wery, S. Warichet, and E. Helbert, “Is Noise Reduction Improving Open-Source ASR Transcription Engines Quality?,” in *Proceedings of the 16th International Conference on Agents and Artificial*

- Intelligence*, Rome, Italy: SCITEPRESS - Science and Technology Publications, 2024, pp. 1221–1228. doi: 10.5220/0012457100003636.
- [23] D. Ferdiansyah and C. S. K. Aditya, “Implementasi Automatic Speech Recognition Bacaan Al-Qur’an Menggunakan Metode Wav2Vec 2.0 dan OpenAI-Whisper,” *J. Tek. Elektro Dan Komput. TRIAC*, vol. 11, no. 1, pp. 11–16, Jun. 2024, doi: 10.21107/triac.v11i1.24332.
- [24] M. Fuckner, S. Horsman, P. Wiggers, and I. Janssen, “Uncovering Bias in ASR Systems: Evaluating Wav2vec2 and Whisper for Dutch speakers,” in *2023 International Conference on Speech Technology and Human-Computer Dialogue (SpeD)*, Bucharest, Romania: IEEE, Oct. 2023, pp. 146–151. doi: 10.1109/SpeD59241.2023.10314895.
- [25] S. Alharbi *et al.*, “Automatic Speech Recognition: Systematic Literature Review,” *IEEE Access*, vol. 9, pp. 131858–131876, 2021, doi: 10.1109/ACCESS.2021.3112535.
- [26] A. Mehrish, N. Majumder, R. Bharadwaj, R. Mihalcea, and S. Poria, “A review of deep learning techniques for speech processing,” *Inf. Fusion*, vol. 99, p. 101869, Nov. 2023, doi: 10.1016/j.inffus.2023.101869.
- [27] N. Babić and J. Galić, “An Analysis of Speech Emotion Recognition Based on Hybrid DNN-HMM Framework,” in *2023 31st Telecommunications Forum (TELFOR)*, Belgrade, Serbia: IEEE, Nov. 2023, pp. 1–4. doi: 10.1109/TELFOR59449.2023.10372735.
- [28] H. Tomar, D. Deshwal, and N. Trivedi, “Multilingual audio classification: a scalable approach using Wav2Vec2 and qdrant,” *Int. J. Inf. Technol.*, Dec. 2025, doi: 10.1007/s41870-025-02982-3.
- [29] V. J. Reddi *et al.*, “MLPerf Inference Benchmark,” in *2020 ACM/IEEE 47th Annual International Symposium on Computer Architecture (ISCA)*, Valencia, Spain: IEEE, May 2020, pp. 446–459. doi: 10.1109/ISCA45697.2020.00045.
- [30] C.-T. Lu and L.-Y. Wang, “Implementation of an Automatic Meeting Minute Generation System Using YAMNet with Speaker Identification and Keyword Prompts,” *Appl. Sci.*, vol. 14, no. 13, p. 5718, Jun. 2024, doi: 10.3390/app14135718.
- [31] D. Sanapala, K. Choudhary, and S. Shetty, “Multi-Person Speech Curator for Minutes of Meetings along with Meeting Summarization and Language Translation,” in *Proceedings of the 2023 7th International Conference on Advances in Artificial Intelligence*, Istanbul Turkiye: ACM, Oct. 2023, pp. 6–12. doi: 10.1145/3633598.3633600.
- [32] K. Wei, B. Li, H. Lv, Q. Lu, N. Jiang, and L. Xie, “Conversational Speech Recognition by Learning Audio-Textual Cross-Modal Contextual Representation,” *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 32, pp. 2432–2444, 2024, doi: 10.1109/TASLP.2024.3389630.
- [33] S. U. Meshram, “Evolution of Modern Web Services – REST API with its Architecture and Design”.
- [34] G. Blinowski, A. Ojdowska, and A. Przybylek, “Monolithic vs. Microservice Architecture: A Performance and Scalability Evaluation,” *IEEE Access*, vol. 10, pp. 20357–20374, 2022, doi: 10.1109/ACCESS.2022.3152803.

- [35] I. Shabani, E. Mëziu, B. Berisha, and T. Biba, “Design of Modern Distributed Systems based on Microservices Architecture,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 2, 2021, doi: 10.14569/IJACSA.2021.0120220.
- [36] K. Arya Bayu Wirya Kesuma, I. N. Y. A. Wijaya, and I. G. J. E. Putra, “Implementasi Next.Js, Typescript, Dan Tailwind Css Untuk Pengembangan Aplikasi *Frontend* Sistem Inventory Perusahaan Apar (Studi Kasus: CV Indoka Surya Jaya),” *Jikom J. Inform. Dan Komput.*, vol. 14, no. 2, pp. 95–108, Oct. 2024, doi: 10.55794/jikom.v14i2.195.
- [37] L. Fitriani and M. T. Alfarisi, “Rancang Bangun Aplikasi Monitoring Kebutuhan Masyarakat Untuk Relawan TIK Berbasis Web dan Mobile,” *J. Algoritma*, vol. 22, no. 2, Nov. 2025, doi: 10.33364/algoritma/v.22-2.2574.
- [38] T. Wijaya, “TEKNIK CODING HYBRID ANTARA TYPESCRIPT DAN JAVASCRIPT DALAM *FRAMEWORK* ANGULAR,” *Inf. Syst. J.*, vol. 8, no. 01, pp. 21–28, Jun. 2025, doi: 10.24076/infosjournal.2025v8i01.2048.
- [39] A. Muhamad Saepuloh and S. Ginting, “PERANCANGAN SISTEM INFORMASI MANAJEMEN PROYEK DENGAN MENGGUNAKAN SOFTWARE NEST.JS BERBASIS WEB DI PT. MITRA PAJAKKU,” *INFOKOM Inform. Komput.*, vol. 10, no. 1, pp. 1–9, Nov. 2022, doi: 10.56689/infokom.v10i1.818.
- [40] A. Fu’adi and A. Prianggono, “Analisa dan Perancangan Sistem Informasi Akademik Akademi Komunitas Negeri Pacitan Menggunakan Diagram UML dan EER,” *J. Ilm. Teknol. Inf. Asia*, vol. 16, no. 1, pp. 45–54, Mar. 2022, doi: 10.32815/jitika.v16i1.650.
- [41] I. R. Mukhlis and R. Santoso, “Perancangan Basis Data Perpustakaan Universitas Menggunakan MySQL dengan Physical Data Model dan Entity Relationship Diagram,” *J. Technol. Inform. JoTI*, vol. 4, no. 2, pp. 81–87, Apr. 2023, doi: 10.37802/joti.v4i2.330.
- [42] O. D. Alao, E. A. Priscilla, R. C. Amanze, S. O. Kuyoro, and A. O. Adebayo, “User-Centered/User Experience Uc/Ux Design Thinking Approach for Designing a University Information Management System,” *Ingénierie Systèmes Inf.*, vol. 27, no. 4, pp. 577–590, Aug. 2022, doi: 10.18280/isi.270407.
- [43] H. Wang, Y. Zhang, and X. Fu, “Innovative Applications of Black Box Testing: A New Strategy for Teaching Assessment in Higher Education,” *IEEE Access*, vol. 13, pp. 90514–90526, 2025, doi: 10.1109/ACCESS.2025.3568825.
- [44] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Comput. Sci.*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.