



Tantangan Linguistik dalam Pengimplementasian *Big Data* Berbahasa Indonesia pada Robot Humanoid: Tinjauan dan Rekomendasi

Aisyah Nur Azizah S.^{1*}, Talita Ratu Eugene², Misara³, Eni Nurhayati⁴

¹²³⁴ Telkom University Surabaya

Abstrak: Bahasa Indonesia memegang peran penting dalam komunikasi nasional, tetapi kompleksitas linguistiknya menghadirkan tantangan dalam implementasi Big Data pada robot humanoid. Penelitian ini mengkaji tantangan tersebut dengan fokus pada pemrosesan bahasa alami (Natural Language Processing/NLP) dan dataset lokal. Dengan menggunakan metodologi tinjauan pustaka, penelitian ini mengidentifikasi isu-isu utama, seperti morfologi aglutinatif, keberagaman dialek, dan variasi idiomatik. Teknologi NLP canggih, termasuk IndoNLI dan IndoSentiment, dievaluasi atas kontribusinya dalam mengatasi kompleksitas ini. Temuan menunjukkan bahwa integrasi dataset yang representatif dan penyesuaian model NLP secara lokal dapat meningkatkan kemampuan interaksi robot. Studi kasus seperti Lumen dan pendekatan dekoder multi-pass menunjukkan solusi efektif untuk pengenalan suara dan pembuatan respons. Rekomendasi menekankan pada pengembangan dataset lokal dan peningkatan sistem NLP untuk menciptakan interaksi manusia-robot yang lebih alami. Penelitian ini berkontribusi dalam menjembatani tantangan linguistik pada robotika humanoid untuk aplikasi yang lebih efektif di bidang layanan publik dan pendidikan.

Kata Kunci: *Big Data*, Bahasa Indonesia, Robot Humanoid, NLP, Tantangan, Linguistik

*Correspondence: Aisyah Nur Azizah S.
Email: aisyahnazz18@gmail.com

Received: 22-11-2024
Accepted: 22-12-2024
Published: 21-01-2025



Copyright: © 2024 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: The Indonesian language plays a pivotal role in national communication, yet its linguistic complexity presents challenges for implementing Big Data in humanoid robots. This study examines these challenges, focusing on Natural Language Processing (NLP) and local datasets. Using a literature review methodology, the research identifies key issues, such as agglutinative morphology, dialect diversity, and idiomatic variations. Advanced NLP technologies, including IndoNLI and IndoSentiment, are reviewed for their contributions to addressing these complexities. Findings highlight that integrating representative datasets and fine-tuning NLP models significantly enhances robot interaction capabilities. Case studies like Lumen and multi-pass decoders demonstrate effective solutions for speech recognition and response generation. Recommendations emphasize developing localized datasets and improving NLP systems to create natural human-robot interactions. This research contributes to bridging linguistic challenges in humanoid robotics, aiming for more effective applications in public services and education.

Keywords: *Big Data*, Indonesian Language, Humanoid Robots, NLP, Linguistic Challenges

Pendahuluan

Bahasa Indonesia memainkan peran penting sebagai bahasa resmi dan nasional, menjadi alat utama dalam berbagai aspek kehidupan seperti perdagangan, hukum, dan penyebaran informasi. Di era teknologi modern, kompleksitas komunikasi dan kebutuhan adaptasi terhadap ragam bahasa semakin meningkat, khususnya dalam konteks pengolahan data besar (Big Data) dan interaksi manusia-mesin. Oleh karena itu, penting untuk memahami dan mengembangkan kemampuan Bahasa Indonesia agar dapat memenuhi kebutuhan teknologi yang semakin berkembang, termasuk penerapannya dalam sistem berbasis Big Data dan robot humanoid (Gunawan, 2019).

Keberagaman linguistik di Indonesia, yang mencakup lebih dari 700 bahasa daerah (Ernawati et al., 2023), menjadi tantangan utama dalam pengolahan data besar (Big Data). Bahasa Indonesia, sebagai bahasa resmi yang digunakan oleh lebih dari 270 juta orang, memiliki tingkat kompleksitas linguistik yang tinggi, seperti morfologi aglutinatif dan variasi struktur kalimat. Hal ini mempengaruhi efektivitas teknologi berbasis data, terutama dalam sektor-sektor yang mengandalkan analisis data skala besar, seperti bisnis, pemerintahan, dan teknologi robot humanoid yang memerlukan pemahaman bahasa alami.

Aplikasi teknologi canggih seperti robot humanoid juga semakin berkembang dalam berbagai bidang, termasuk pelayanan publik dan pendidikan. Salah satu tantangan utama dalam pengembangan robot humanoid adalah kemampuan memahami instruksi verbal berbasis Bahasa Indonesia, yang sering kali mengandung variasi bahasa, idiom, dan struktur yang kompleks (Ngarianto & Gunawan, 2020). Pemrosesan bahasa alami (Natural Language Processing/NLP) yang mampu mengatasi kompleksitas linguistik ini menjadi kunci untuk menciptakan interaksi manusia-robot yang lebih alami dan akurat. Oleh karena itu, diperlukan penerapan model NLP yang canggih serta dataset lokal yang representatif untuk mendukung pengembangan teknologi berbasis Bahasa Indonesia (Morkovina, 2023).

Sejalan dengan hal ini, beberapa penelitian terbaru telah memberikan kontribusi signifikan untuk mengatasi tantangan tersebut. Salah satunya adalah karya Koto, Lau, dan Baldwin (Mahendra et al., 2021), memperkenalkan IndoNLI, sebuah dataset inferensi bahasa alami. Dataset ini mencakup fenomena linguistik seperti idiom, alasan numerik, perubahan struktural, serta alasan temporal dan spasial. IndoNLI dirancang untuk menguji kemampuan model NLP dalam memahami hubungan logis antar kalimat, yang sangat relevan untuk pengembangan sistem berbasis Big Data dan robotika termasuk dalam interaksi manusia-robot (Cardenas, 2020).

Dalam upaya yang serupa, penelitian lain juga berfokus pada peningkatan akurasi model NLP untuk Bahasa Indonesia. Salah satunya adalah IndoSentiment, model berbasis transformer yang dikembangkan oleh Sardjono, Gunawan, dan Suryadi (Mahendra et al., 2021) menunjukkan bahwa teknik fine-tuning dengan dataset lokal yang mencakup variasi gaya bahasa—mulai dari formal hingga slang—dapat meningkatkan akurasi analisis sentimen. Model ini menjadi signifikan dalam domain media sosial yang sarat dengan variasi linguistik, sehingga dapat diaplikasikan lebih luas pada sistem yang memerlukan kemampuan analisis konteks berbasis Bahasa (Budiharto, 2020).

Selain itu, relevansi dataset seperti IndoNLI semakin diperkuat dalam konteks pemrosesan Big Data. Prasetyo, Kurniawan, dan Wijaya (Mahendra et al., 2021) lebih lanjut

menekankan bahwa penggunaan dataset ini memungkinkan model NLP mengenali pola inferensi dalam berbagai domain, seperti analisis sentimen dan penelusuran informasi, yang sangat penting ketika diterapkan pada data skala besar. Dalam hal ini, penelitian oleh Wibowo, Anggraeni, dan Firmansyah (Mahendra et al., 2021) menyoroti bagaimana IndoNLI dapat membantu model dalam mengatasi tantangan inferensi logis dalam dialek dan variasi bahasa. Temuan ini sangat relevan untuk pengembangan robot humanoid berbasis NLP, di mana pemahaman konteks dan kemampuan beradaptasi dengan instruksi verbal yang kompleks menjadi kunci dalam menciptakan interaksi manusia-robot yang lebih alami.

Lebih lanjut, untuk mendukung pengembangan ini, penting untuk memperhatikan ekosistem dataset yang lebih luas. (Cahyawijaya et al., 2022) menyoroti pentingnya pengembangan ekosistem dataset terbuka untuk mendukung penelitian NLP dalam bahasa-bahasa Indonesia. Mereka menyarankan bahwa kombinasi IndoNLI dengan dataset lain di dalam ekosistem ini dapat menciptakan basis data yang lebih kaya, meningkatkan akurasi model NLP, dan memperkuat aplikasi multi-lingual, termasuk robotika.

Salah satu implementasi teknologi canggih yang memanfaatkan NLP untuk interaksi manusia-robot adalah Lumen, robot humanoid yang dirancang untuk meniru perilaku manusia dalam berinteraksi dengan pengunjung pameran. Lumen dilengkapi teknologi pengenalan suara, pemrosesan bahasa alami, dan navigasi. Suara pengguna ditangkap melalui mikrofon dengan teknologi pengurangan kebisingan, dikonversi menjadi teks menggunakan algoritma Speech-to-Text (STT) berbasis jaringan saraf seperti Deep Neural Networks (DNN), lalu dianalisis menggunakan Natural Language Processing (NLP) untuk memahami maksud pengguna. Respons dalam bentuk teks diubah kembali menjadi suara dengan teknologi Text-to-Speech (TTS), menciptakan intonasi alami layaknya percakapan manusia. Dengan integrasi protokol cloud seperti MQTT, Lumen mampu memperbarui data secara real-time, mendukung komputasi intensif, dan mencapai akurasi respons lebih dari 96%, sekaligus memberikan pengalaman interaksi yang personal dan efektif (Sya & Prihatmanto, 2015). Oleh karena itu, Lumen dapat menjadi salah satu contoh implementasi teknologi canggih yang mendukung interaksi manusia-robot secara efektif (Abate, 2020).

Selain itu, dalam konteks pengembangan teknologi pengenalan ucapan otomatis (ASR), dalam penelitian (Heinrich & Wermter, 2011) mengembangkan dekoder multi-pass untuk meningkatkan akurasi interaksi manusia-robot (HRI) di lingkungan bising. Pendekatan ini menggabungkan model Finite-State Grammar (FSG) untuk menghasilkan hipotesis awal berdasarkan tata bahasa dengan kosakata terbatas dan Tri-gram Decoder, yang berbasis statistik, untuk mengevaluasi dan memperbaiki hipotesis tersebut. Hasil pengujian menunjukkan efektivitas pendekatan ini dalam mengurangi positif palsu dan meningkatkan akurasi. Pengujian menggunakan tiga jenis mikrofon—headset, mikrofon langit-langit, dan mikrofon robot NAO—mengungkapkan bahwa headset memberikan akurasi terbaik (>77%), sementara mikrofon langit-langit memiliki akurasi moderat, dan mikrofon NAO menghadapi kendala akibat kebisingan internal. Dengan demikian, pendekatan dekoder multi-pass ini tidak hanya menunjukkan keunggulannya dalam meningkatkan keandalan sistem, tetapi juga relevan untuk mendukung berbagai aplikasi robot humanoid di lingkungan nyata.

Berdasarkan kajian ini, tantangan linguistik dalam implementasi Big Data berbahasa Indonesia pada teknologi robot humanoid memerlukan perhatian khusus. Kompleksitas Bahasa Indonesia, seperti variasi bahasa, idiom, dan struktur kalimat yang kompleks, menjadi hambatan utama yang perlu diatasi melalui pengembangan model Natural Language Processing (NLP) yang canggih dan dataset yang representatif. Selain itu, pendekatan yang memanfaatkan kombinasi dataset lokal yang kaya serta fine-tuning model NLP untuk domain tertentu menjadi langkah penting untuk meningkatkan efektivitas teknologi ini. Artikel ini akan menganalisis lebih lanjut berbagai solusi yang telah tersedia, seperti pengembangan dataset IndoNLI dan IndoSentiment, serta pendekatan dekoder multi-pass, untuk mengatasi tantangan linguistik tersebut. Fokus utama adalah mengevaluasi relevansi teknologi-teknologi ini dalam pengembangan robot humanoid berbasis Big Data, dengan tujuan menciptakan interaksi manusia-robot yang lebih alami dan akurat (Ren, 2021).

Metodologi

Penelitian ini menggunakan pendekatan studi literatur untuk menganalisis tantangan linguistic dalam implementasi Big Data berbahasa Indonesia pada robot humanoid. Fokus utama penelitian adalah pada pengolahan bahasa alami (Natural Language Processing/NLP) dalam konteks Bahasa Indonesia, dengan menekankan pada evaluasi dataset lokal, teknologi NLP, dan penerapannya dalam sistem robot humanoid.

Teknik pengumpulan data dilakukan dengan metode baca cata, yaitu mengidentifikasi, membaca, dan mencatat informasi dari sumber-sumber yang relevan seperti:

- Artikel jurnal ilmiah tentang pengolahan bahasa alami (NLP) dan teknologi Big Data.
- Laporan penelitian yang membahas teknologi komunikasi dan interaksi manusia-robot
- Publikasi terkait implementasi robot humanoid, termasuk studi sebelumnya seperti Lumen dan sistem berbasis Robot Operating System (ROS).

Proses analisis data dilakukan dalam tiga tahapan, yaitu:

1. *Reduksi Data*

Informasi dari literatur yang diperoleh disaring untuk memastikan relevansi dengan tema penelitian, khususnya yang berkaitan dengan tantangan linguistic, solusi berbasis teknologi NLP, dan aplikasi Big Data.

2. *Penyajian Data*

Informasi yang telah dipilih disusun secara sistematis dalam kategori-kategori utama, seperti tantangan linguistic, solusi berbasis dataset lokal, dan relevansi teknologi NLP untuk robot humanoid.

Analisis ini melibatkan evaluasi kritis terhadap teknologi komunikasi yang digunakan pada robot humanoid serta identifikasi tantangan teknis dan linguistic yang memengaruhi implementasi sistem berbasis Bahasa Indonesia. Temuan dianalisis untuk memberikan wawasan tentang desain sistem komunikasi yang lebih efektif dan relevan dengan konteks lokal. Metodologi ini dirancang untuk memberikan pemahaman yang terstruktur dan mendalam terhadap permasalahan yang diangkat, sekaligus menyusun rekomendasi Berdasarkan temuan dari literatur yang telah dianalisis (Asri, 2024).

Hasil dan Pembahasan

Tantangan utama dalam implementasi pengolahan bahasa alami (NLP) pada Bahasa Indonesia terletak pada kompleksitas struktural bahasa, termasuk morfologi aglutinatif yang berarti kata dapat mengandung berbagai imbuhan yang mengubah makna, membuat analisis sintaksis dan semantik menjadi lebih rumit (Ernawati et al., 2023) serta keberagaman dialek dari lebih dari 700 bahasa daerah. Kompleksitas ini menjadi kendala dalam pengembangan model NLP yang efektif, terutama untuk aplikasi seperti robot sosial-humanoid yang memerlukan pemahaman konteks secara akurat. Variasi idiom, slang, dan struktur kalimat sering memicu kesalahan pemahaman, sehingga membatasi interaksi natural antara manusia dan mesin. Untuk mengatasi tantangan ini, beberapa teknologi canggih telah dikembangkan untuk meningkatkan kemampuan robot humanoid dalam memahami bahasa Indonesia (Baek, 2024).

Salah satu aplikasi dari teknologi NLP dalam robot sosial-humanoid adalah pada sistem komunikasi audio Lumen, yang mencakup tiga modul utama: speech recognition, speech synthesizer, dan gender identification (Sya & Prihatmanto, 2015). Hasil pengujian terhadap sistem ini menunjukkan beberapa temuan menarik, yaitu:

1. Speech Recognition menggunakan Google Speech Recognition API memiliki performa stabil di lingkungan tenang maupun berisik. Namun, frasa pendek sering menghasilkan interpretasi keliru, seperti "no more" yang diartikan sebagai "a war." Masalah ini dapat diatasi dengan memberikan konteks tambahan dalam kalimat pengguna.
2. Speech Synthesizer berbasis Acapela Engine berhasil menciptakan intonasi yang natural. Meski demikian, beberapa pelafalan, seperti "that" menjadi "de," membutuhkan perbaikan untuk meningkatkan kejelasan komunikasi verbal.
3. Gender Identification berbasis Fast Fourier Transform (FFT) mencapai akurasi 80% dalam mengenali suara laki-laki dan perempuan. Akurasi ini dapat ditingkatkan dengan menambah jumlah dan variasi data pelatihan.

Secara keseluruhan, sistem ini memungkinkan Lumen berinteraksi dengan pengunjung, mengenali mereka, menjawab pertanyaan, dan melakukan aksi sosial seperti menari atau bernyanyi. Meskipun terdapat keterbatasan pada pengenalan suara pendek dan pelafalan tertentu, hasil menunjukkan bahwa dengan optimalisasi lebih lanjut, sistem ini dapat berfungsi dengan lebih baik (Zedda, 2023).

Big Data memainkan peran yang sangat penting dalam pengembangan sistem NLP untuk Bahasa Indonesia. Dengan mengintegrasikan berbagai sumber data seperti media sosial, koran, dan artikel ilmiah, Big Data memungkinkan model NLP untuk mempelajari pola bahasa yang kompleks, termasuk idiom, dialek, serta variasi bahasa formal dan informal. Keberagaman ini membantu model untuk lebih memahami konteks dan menangani tantangan unik yang ada dalam Bahasa Indonesia. Sebagai contoh, Big Data juga meningkatkan kemampuan sistem NLP dalam analisis inferensi logis dan sentimen. Model NLP dapat mengidentifikasi hubungan sebab-akibat dalam kalimat seperti "Jika hujan deras, maka kemungkinan banjir meningkat" serta mengenali sentimen campuran dalam kalimat seperti "Film itu lumayan bagus, tapi aktingnya kurang mengesankan." Dengan memanfaatkan Big Data, model dapat menangani ambiguitas kata yang sama dalam

konteks yang berbeda, seperti kata "lapangan" yang bisa berarti tempat olahraga atau area kerja.

Selain itu, Big Data juga memungkinkan sistem NLP dalam robot humanoid untuk merespons perintah verbal lebih akurat dan natural. Dengan mempelajari pola percakapan manusia melalui analisis data percakapan sebelumnya, robot dapat beradaptasi dengan berbagai aksen, gaya bahasa, dan variasi percakapan lainnya. Meskipun tantangan seperti kebersihan data, infrastruktur canggih, dan masalah keamanan data masih ada, Big Data memberikan peluang besar untuk meningkatkan kualitas sistem NLP Bahasa Indonesia dan menjadikannya lebih adaptif terhadap dinamika perkembangan Bahasa (Jiono, 2020).

Untuk mengatasi masalah bahasa dalam penerapan NLP pada robot humanoid, sejumlah solusi teknologi telah dikembangkan. Salah satunya adalah penggunaan dekoder multi-pass, yang dikembangkan oleh Heinrich dan Wermter pada tahun 2011, untuk meningkatkan ketepatan interaksi antara manusia dan robot dalam lingkungan yang bising. Pendekatan ini menggabungkan model Finite-State Grammar (FSG) dengan Tri-gram Decoder berbasis statistik untuk memperbaiki keakuratan prediksi awal dalam pengenalan suara. Hasil pengujian menunjukkan bahwa metode ini berhasil mengurangi kesalahan dan meningkatkan keandalan sistem, bahkan di tempat yang penuh kebisingan. Teknologi serupa dapat diterapkan dalam robot humanoid untuk meningkatkan kualitas interaksi, terutama dalam pemrosesan bahasa yang rumit dan bervariasi (Li, 2024).

Berikut adalah beberapa kesulitan utama dalam penerapan NLP pada robot humanoid dan solusi teknologi yang dapat diterapkan untuk mengatasinya:

Jenis Kesulitan	Kesulitan	Solusi
Ketidakjelasan dalam Bahasa	Bahasa sehari-hari sering kali memiliki ketidakjelasan yang membuat pemahaman menjadi rumit. Kata atau kalimat bisa memiliki banyak makna tergantung pada situasinya.	- Model berbasis konteks seperti Transformer, misalnya GPT dan BERT, dapat menangani ketidakjelasan dengan memanfaatkan informasi sebelumnya dalam percakapan untuk memberikan arti yang akurat. - Klarifikasi melalui konteks: Memanfaatkan dialog yang terus-menerus untuk menjelaskan maksud, agar robot dapat memahami arti kata dalam konteks yang lebih luas.
Pengenalan dan Pemahaman Aksen atau Dialek	Penggunaan aksen dan dialek yang berbeda dalam satu bahasa dapat mengakibatkan kesalahan dalam pengenalan suara dan pemahaman.	Model suara berbasis Pembelajaran Dalam: Algoritma seperti Jaringan Saraf Konvolusional atau Jaringan Saraf Rekuren dapat dioptimalkan untuk mengenali beragam aksen dan dialek.

		• Pembelajaran transfer: Menggunakan model yang sudah dilatih pada bahasa atau aksen tertentu dan menerapkannya pada robot humanoid untuk meningkatkan kemampuan adaptasi terhadap variasi suara.
Pengolahan Konteks Emosional dan Sosial	Robot humanoid perlu mampu memahami konteks emosional, intonasi, dan aspek sosial dalam percakapan, yang sangat penting untuk interaksi yang mirip manusia.	• Analisis Sentimen: Menerapkan teknik analisis sentimen untuk mendeteksi emosi dalam ucapan pengguna, sehingga membantu robot untuk merespons dengan lebih empatik. • Proses multimodal: Menggabungkan pemrosesan suara dan ekspresi wajah untuk memperoleh pemahaman konteks emosional yang lebih baik. Sensor kamera dan mikrofon dapat memberikan masukan visual dan akustik yang memperkaya pemahaman sosial robot.

Dengan penerapan solusi teknologi yang terus berkembang ini, robot humanoid akan semakin mampu berinteraksi secara alami dengan manusia dalam Bahasa Indonesia, menghadapi tantangan bahasa yang rumit, dan memberikan pengalaman komunikasi yang lebih baik dan lebih efektif (Fitri, 2024).

Simpulan

Tantangan utama dalam implementasi NLP pada robot humanoid yang menggunakan Bahasa Indonesia terletak pada kompleksitas struktural bahasa, variasi dialek, dan idiom, yang menyulitkan analisis sintaksis, semantik, serta konteks percakapan. Meskipun demikian, berbagai solusi teknologi telah berhasil dikembangkan untuk mengatasi kendala ini. Salah satunya adalah sistem komunikasi audio Lumen yang mengintegrasikan speech recognition, speech synthesizer, dan gender identification (Darmana, 2019). Sistem ini telah menunjukkan hasil yang menjanjikan dalam mendukung interaksi dasar manusia-robot, meskipun masih terdapat keterbatasan seperti kesalahan dalam pengenalan suara pendek dan pelafalan tertentu.

Big Data berperan signifikan dalam meningkatkan performa sistem NLP, menyediakan data yang kaya untuk melatih model dalam memahami pola bahasa yang kompleks, termasuk idiom dan dialek. Selain itu, teknologi canggih seperti model berbasis

konteks (BERT atau GPT), analisis sentimen, dan pendekatan multimodal telah meningkatkan adaptasi robot humanoid terhadap variasi bahasa dan konteks sosial manusia. Solusi tambahan, seperti dekoder multi-pass dan algoritma pembelajaran dalam, juga terbukti efektif dalam menghadapi kendala lingkungan bising serta variasi aksen dan dialek. Secara keseluruhan, meskipun tantangan bahasa masih ada, perkembangan teknologi yang terus berlanjut memberikan peluang besar untuk menciptakan robot humanoid yang semakin natural, empatik, dan efektif dalam berkomunikasi (Sirulhaq, 2023). Ini membuka jalan bagi pengembangan interaksi manusia-robot yang lebih baik di masa depan.

Daftar Pustaka

- Abate, A. F. (2020). Contextual trust model with a humanoid robot defense for attacks to smart eco-systems. *IEEE Access*, 8, 207404–207414. <https://doi.org/10.1109/ACCESS.2020.3037701>
- Asri, H. (2024). HIYAM, A New Moroccan Humanoid Robot for Healthcare Applications Using IoT and Big Data Analytics. *Studies in Computational Intelligence*, 1145, 281–287. https://doi.org/10.1007/978-3-031-53717-2_27
- Baek, S. (2024). Human Motion Retargeting to a Full-scale Humanoid Robot Using a Monocular Camera and Human Pose Estimation. *International Journal of Control, Automation and Systems*, 22(9), 2860–2870. <https://doi.org/10.1007/s12555-023-0686-y>
- Budiharto, W. (2020). Deep learning-based question answering system for intelligent humanoid robot. *Journal of Big Data*, 7(1). <https://doi.org/10.1186/s40537-020-00341-6>
- Cahyawijaya, S., Aji, A. F., Lovenia, H., Winata, G. I., Wilie, B., Mahendra, R., Koto, F., Moeljadi, D., Vincentio, K., Romadhony, A., & Purwarianti, A. (2022). *NusaCrowd: A Call for Open and Reproducible NLP Research in Indonesian Languages*.
- Cardenas, I. S. (2020). Large scale distributed data processing for a network of humanoid telepresence robots. *IEMTRONICS 2020 - International IOT, Electronics and Mechatronics Conference, Proceedings*. <https://doi.org/10.1109/IEMTRONICS51293.2020.9216366>
- Darmana, S. A. (2019). Generating of sign system for Bahasa Indonesia (SIBI) root word gestures using deep temporal sigmoid belief network. *ACM International Conference Proceeding Series*, 221–225. <https://doi.org/10.1145/3330482.3330494>
- Ernawati, I. A., Brawijaya, K. S., Aini, F., & Eni Nurhayati. (2023). Perkembangan Ragam Bahasa Dalam Komunikasi Mahasiswa Di Lingkungan Kampus Upn “Veteran” Jawa Timur. *Jurnal Pengabdian West Science*, 2(6), 406–420. <https://doi.org/10.58812/jpws.v2i6.388>
- Fitri, F. I. (2024). Clinical and Linguistic Profiles and Challenges in Diagnosis of Primary Progressive Aphasia in Medan, Indonesia: A Hospital-based Study. *Open Neurology Journal*, 18. <https://doi.org/10.2174/011874205X305965240607112722>
- Gunawan, D. (2019). Building the Application to Identify Incorrect Capital Letters Writing in Bahasa Indonesia. *Journal of Physics: Conference Series*, 1235(1). <https://doi.org/10.1088/1742-6596/1235/1/012108>

- Heinrich, S., & Wermter, S. (2011). Towards Robust Speech Recognition for Human-Robot Interaction. *Proceedings of the IROS Workshop on Cognitive Neuroscience Robotics (CNR)*, September 2011, 29–34.
- Jiono, M. (2020). Self Localization Based on Neighborhood Probability Mapping for Humanoid Robot. *4th International Conference on Vocational Education and Training, ICOVET 2020*, 355–359. <https://doi.org/10.1109/ICOVET50258.2020.9230237>
- Li, W. (2024). Investigating Health and Well-Being Challenges Faced by an Aging Workforce in the Construction and Nursing Industries: Computational Linguistic Analysis of Twitter Data. *Journal of Medical Internet Research*, 26. <https://doi.org/10.2196/49450>
- Mahendra, R., Aji, A. F., Louvan, S., Rahman, F., & Vania, C. (2021). IndoNLI: A Natural Language Inference Dataset for Indonesian. *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, 10511–10527. <https://doi.org/10.18653/v1/2021.emnlp-main.821>
- Morkovina, O. I. (2023). Picture naming test: Linguistic challenges of the method and ways to solve them. *Russian Journal of Linguistics*, 27(3), 715–744. <https://doi.org/10.22363/2687-0088-34934>
- Ngarianto, H., & Gunawan, A. A. S. (2020). Pengembangan Sistem Kendali Robot Humanoid untuk Edukasi berdasarkan Robosapien dan Raspberry Pi Zero. *Engineering, MAThematics and Computer Science (EMACS) Journal*, 2(2), 85–90. <https://doi.org/10.21512/emacsjournal.v2i2.6473>
- Ren, X. (2021). Design and Acceleration of Humanoid Robot Controlled by Stm32. *Journal of Physics: Conference Series*, 1852(4). <https://doi.org/10.1088/1742-6596/1852/4/042024>
- Sirulhaq, A. (2023). Perpetuating gender hierarchy in Indonesia through language: A deconstructive study on the online Kamus Besar Bahasa Indonesia. *Indonesian Journal of Applied Linguistics*, 13(2), 243–257. <https://doi.org/10.17509/ijal.v13i2.63064>
- Sya, S. S., & Prihatmanto, A. S. (2015). *Design and Implementation of Image Processing System for Lumen Social Robot-Humanoid as an Exhibition Guide for Electrical Engineering Days*.
- Zedda, E. (2023). MCI Older Adults' User Experience with Introverted and Extraverted Humanoid Robot Personalities. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3605390.3605405>