



## Eksplorasi Linguistik Komputasional dalam Analisis Bahasa Alami untuk Mengungkap Evolusi Dialek Digital di Era Media Sosial Global

Ari Putra Wibowo<sup>1\*</sup>, Nurul Hidayat<sup>2</sup>

<sup>1</sup> Institut Widya Pratama Pekalongan, Indonesia

<sup>2</sup> Universitas Pancasila, Indonesia

\*Penulis korespondensi: [ariputra.stmikwp@gmail.com](mailto:ariputra.stmikwp@gmail.com)

**Abstract.** *The development of digital technology and social media has driven a major transformation in the study of linguistics, particularly in understanding the evolution of global languages and the emergence of digital dialects. Interactions on platforms such as Twitter, Instagram, and TikTok accelerate the formation of new vocabulary, the use of slang, and the spread of cross-cultural expression. Antony and Tramboo (2023) highlight the digital metamorphosis in language, while Friedrich and De Figueiredo (2016) explain the birth of digital Englishes in a sociolinguistic perspective. Language research challenges on social media have also emerged, such as the limitations of dialect processing (Jørgensen et al., 2015) and the influence of digital linguistic ecology (Klushina, 2022). Advances in natural language processing (NLP) have also strengthened this study. Cambria and White (2014) and Skaria et al. (2024) show the development of NLP from sentiment analysis to its application in education (Khensous et al., 2023) and multilingual text analysis (Agüero-Torales et al., 2021). The systematic review of Sundaram et al. (2023) and Prihatini et al. (2023) reinforces the evidence that social media expands language skills and enriches contemporary discourse, while Sun et al. (2021) show an increasing trend of digital linguistic research through bibliometric analysis, including cultural gaps (Rani & Samjetsabam, 2024). Overall, language not only evolves naturally, but also through human interaction with technology and digital culture. Digital dialects, slang, and online communication patterns are part of the global linguistic evolution that opens up interdisciplinary research opportunities between linguistics, computing, and cultural studies.*

**Keywords:** *Digital linguistics; Social media; Natural language processing (NLP); Digital dialect; Language evolution.*

**Abstrak.** Perkembangan teknologi digital dan media sosial telah mendorong transformasi besar dalam studi linguistik, khususnya dalam memahami evolusi bahasa global dan kemunculan dialek digital. Interaksi di platform seperti Twitter, Instagram, dan TikTok mempercepat pembentukan kosakata baru, penggunaan slang, serta penyebaran ekspresi lintas budaya. Antony dan Tramboo (2023) menyoroti metamorfosis digital dalam bahasa, sedangkan Friedrich dan De Figueiredo (2016) menjelaskan lahirnya digital Englishes dalam perspektif sosiolinguistik. Tantangan penelitian bahasa di media sosial juga muncul, seperti keterbatasan pemrosesan dialek (Jørgensen et al., 2015) dan pengaruh ekologi linguistik digital (Klushina, 2022). Kemajuan natural language processing (NLP) turut memperkuat kajian ini. Cambria dan White (2014) serta Skaria et al. (2024) menunjukkan perkembangan NLP dari analisis sentimen hingga penerapannya dalam pendidikan (Khensous et al., 2023) dan analisis teks multibahasa (Agüero-Torales et al., 2021). Kajian sistematis Sundaram et al. (2023) dan Prihatini et al. (2023) memperkuat bukti bahwa media sosial memperluas kemampuan berbahasa dan memperkaya wacana kontemporer, sementara Sun et al. (2021) menunjukkan meningkatnya tren penelitian linguistik digital melalui analisis bibliometrik, termasuk kesenjangan budaya (Rani & Samjetsabam, 2024). Secara keseluruhan, bahasa tidak hanya berkembang secara alami, tetapi juga melalui interaksi manusia dengan teknologi dan budaya digital. Dialek digital, slang, dan pola komunikasi daring menjadi bagian dari evolusi linguistik global yang membuka peluang riset interdisipliner antara linguistik, komputasi, dan kajian budaya.

**Kata kunci:** Linguistik digital; Media sosial; Natural language processing (NLP); Dialek digital; Evolusi bahasa.

### 1. PENDAHULUAN

Media sosial telah melahirkan fenomena linguistik baru berupa dialek digital yang mencerminkan kreativitas pengguna sekaligus menandai transformasi komunikasi global. Bahasa yang berkembang di ruang daring tidak hanya berfungsi sebagai sarana komunikasi, tetapi juga sebagai cerminan identitas, interaksi budaya, dan dinamika sosial masyarakat digital. Perpaduan kosakata lokal, global, dan simbol digital menciptakan pola bahasa yang

unik dan berbeda dari bahasa formal tradisional, sehingga memunculkan kebutuhan untuk menelaah lebih jauh bagaimana evolusi linguistik ini terbentuk dan bertransformasi.

Dalam konteks globalisasi, media sosial berperan sebagai katalis yang mengubah cara orang berkomunikasi dan berinteraksi lintas batas geografis maupun budaya (Sundaram et al., 2023). Platform seperti Facebook, Twitter, dan WhatsApp memungkinkan terjadinya pertukaran informasi secara instan, mempercepat perubahan bahasa, dan menghadirkan ruang bagi munculnya variasi linguistik baru (Mishra et al., 2019). Transformasi digital ini juga melahirkan bentuk literasi baru yang berbeda dari pola literasi tradisional, dengan norma dan aturan yang lebih cair dalam komunikasi publik maupun interpersonal (Klushina, 2022).

Kreativitas pengguna dalam ruang digital menjadi faktor penting dalam pembentukan dialek baru. Slang, akronim, serta bentuk bahasa informal lainnya kerap digunakan untuk mengekspresikan identitas sekaligus menandai keanggotaan dalam komunitas digital tertentu (Sundaram et al., 2023). Selain itu, dialek dan variasi bahasa yang jarang ditemukan dalam tulisan formal kini berkembang pesat di media sosial. Misalnya, fitur fonologis dari *African American Vernacular English* (AAVE) telah diadaptasi dalam variasi ejaan di Twitter, memperlihatkan bagaimana identitas linguistik dapat terefleksi melalui medium digital (Jørgensen et al., 2015).

Fenomena ini semakin kompleks dengan adanya interaksi multikultural. Media sosial menghadirkan ruang yang mempertemukan beragam komunitas budaya, sehingga kosakata lokal dan global saling berpadu, membentuk bahasa hibrid yang unik (Rani & Samjetsabam, 2024). Hal ini sejalan dengan pandangan Friedrich dan De Figueiredo (2016) yang menekankan bahwa “digital Englishes” merupakan hasil dari percampuran bahasa global yang terus-menerus dipengaruhi oleh dinamika teknologi. Teknologi modern, termasuk kecerdasan buatan (AI) dan media digital, juga berperan dalam mempercepat metamorfosis linguistik, mengaburkan batas-batas komunikasi dan menciptakan budaya bahasa baru dalam era teknologi (Antony & Trambo, 2023).

Dengan demikian, fenomena dialek digital yang lahir dari media sosial bukan hanya sekadar variasi bahasa, tetapi juga representasi dari transformasi komunikasi global yang dipengaruhi oleh interaksi lintas budaya, kreativitas pengguna, dan perkembangan teknologi. Kajian mendalam mengenai evolusi ini menjadi penting untuk memahami bagaimana bahasa berkembang dalam ruang digital sekaligus implikasinya bagi komunikasi masa depan.

## 2. TINJAUAN PUSTAKA

### **Kajian Linguistik Komputasional dan Analisis Bahasa Alami (NLP)**

Natural Language Processing (NLP), atau yang juga dikenal sebagai linguistik komputasional, merupakan sub-bidang kecerdasan buatan (AI) yang berfokus pada pemodelan bahasa alami untuk berbagai aplikasi seperti pengenalan dan sintesis ucapan, terjemahan mesin, pengenalan karakter optik (OCR), analisis sentimen, sistem tanya jawab, hingga sistem dialog interaktif (Cambria & White, 2014; Skaria et al., 2024). NLP memungkinkan mesin untuk memahami, menafsirkan, dan menghasilkan bahasa manusia melalui serangkaian teknik komputasional yang semakin berkembang seiring dengan ketersediaan data teks dalam jumlah besar dan kemajuan model pembelajaran mendalam (Khensous et al., 2023).

Berbagai penelitian telah menyoroti tantangan dalam NLP, terutama ketika dihadapkan pada bahasa yang kompleks seperti bahasa Arab, yang memerlukan pendekatan linguistik dan teknis khusus (Darwish et al., 2021). Aplikasi NLP telah digunakan dalam analisis sentimen, pengenalan entitas, serta pemrosesan multibahasa, menunjukkan peran signifikan bidang ini dalam mendukung interaksi manusia dan mesin (Skaria et al., 2024).

### **Studi Terdahulu Mengenai Bahasa Media Sosial**

Penelitian linguistik pada media sosial menggunakan berbagai metode seperti analisis wacana, analisis percakapan, analisis multimodal, hingga analisis korpus. Metode tersebut banyak digunakan untuk menelaah variasi sosial, praktik bahasa, dan pembentukan identitas di ruang digital (Spilioti, 2018; Sun et al., 2021). Fokus kajian utama adalah pada presentasi diri, konstruksi komunitas, serta bagaimana bahasa digunakan untuk memperkuat relasi sosial (Spilioti, 2018).

Lebih lanjut, media sosial juga memiliki peran penting dalam pembelajaran bahasa. Platform seperti Facebook, WhatsApp, dan YouTube telah digunakan untuk meningkatkan keterampilan menulis, berbicara, membaca, dan mendengarkan. Selain itu, media sosial juga efektif untuk memperkaya penguasaan kosakata dan tata bahasa dalam konteks pembelajaran daring (Prihatini et al., 2023).

### **Analisis Sentimen Multibahasa pada Media Sosial**

Salah satu tren yang menonjol dalam penelitian linguistik komputasional adalah analisis sentimen multibahasa, khususnya pada media sosial. Dengan meningkatnya keragaman bahasa di dunia digital, penelitian beralih ke arah pendekatan lintas bahasa (*cross-lingual*) dan penggunaan teks berkode campuran (Agüero-Torales et al., 2021). Pendekatan berbasis pembelajaran mendalam telah memberikan hasil yang menjanjikan, meskipun masih

terdapat keterbatasan pada penggunaan arsitektur kompleks berbasis *transformer* (Agüero-Torales et al., 2021).

### **Kesimpulan Tinjauan**

Kajian NLP dan studi bahasa media sosial memperlihatkan perkembangan pesat dalam dekade terakhir. Di satu sisi, NLP berperan penting dalam memungkinkan mesin untuk berinteraksi dengan bahasa manusia secara lebih natural. Di sisi lain, penelitian bahasa media sosial menyoroti bagaimana komunikasi digital menciptakan variasi linguistik baru serta membuka peluang bagi inovasi dalam pembelajaran bahasa. Sinergi keduanya menunjukkan bahwa linguistik komputasional tidak hanya mendukung teknologi komunikasi, tetapi juga merefleksikan dinamika sosial dan budaya dalam era digital.

### **3. METODOLOGI PENELITIAN**

Penelitian ini menggunakan jenis penelitian kualitatif dengan pendekatan linguistik komputasional. Pendekatan ini dipilih karena mampu memadukan analisis bahasa secara mendalam dengan dukungan teknologi komputasi yang memungkinkan pengolahan data dalam jumlah besar. Melalui integrasi metode ini, penelitian berfokus pada identifikasi, analisis, serta interpretasi evolusi dialek digital yang berkembang di era media sosial global.

#### **Sumber Data**

Sumber data penelitian ini berasal dari teks daring yang dipublikasikan melalui media sosial global, seperti Twitter, Instagram, TikTok, dan platform serupa. Pemilihan media sosial didasarkan pada peran dominannya dalam membentuk bahasa digital baru serta keterwakilan generasi pengguna yang sangat beragam. Data diambil melalui teknik web scraping atau Application Programming Interface (API) sesuai dengan ketentuan kebijakan masing-masing platform.

#### **Metode Analisis**

Analisis dilakukan melalui tiga tahap utama. Pertama, text mining digunakan untuk mengekstraksi kosakata dari data teks yang telah dikumpulkan, sehingga dapat diperoleh daftar kata, frasa, dan ekspresi khas yang sering muncul. Kedua, dilakukan analisis semantik untuk memahami makna baru yang muncul dari penggunaan kosakata tersebut, baik dalam konteks lokal maupun global. Ketiga, penerapan Natural Language Processing (NLP) bertujuan mengidentifikasi pola linguistik, termasuk pergeseran morfologi, sintaksis, dan pragmatik yang terjadi dalam dialek digital.

## **Instrumen Penelitian**

Instrumen penelitian berupa perangkat lunak analisis teks, seperti Python dengan library NLP (NLTK, spaCy, atau Transformers), serta R dengan paket analisis teks (tm, quanteda). Selain itu, perangkat lunak visualisasi data seperti Gephi atau Tableau dapat digunakan untuk memetakan hubungan antar kosakata dan tren semantik yang terbentuk.

## **4. HASIL DAN PEMBAHASAN**

### **Hasil Penelitian**

Hasil analisis data teks dari berbagai platform media sosial global menunjukkan bahwa terdapat pergeseran signifikan dalam kosakata dan struktur bahasa yang digunakan oleh pengguna. Melalui proses text mining, ditemukan sejumlah kata baru, singkatan, dan bentuk ekspresi yang merepresentasikan dialek digital. Contohnya adalah penggunaan akronim (seperti OOTD, FOMO, IMO), emotikon yang diubah menjadi bentuk verbal (XD, LOL), serta kombinasi huruf-angka (gr8, l8r).

Pada tahap analisis semantik, hasil menunjukkan bahwa kosakata baru tersebut tidak hanya berfungsi sebagai gaya komunikasi singkat, melainkan juga memiliki makna sosial tertentu. Misalnya, istilah ghosting yang awalnya berarti menghilang tanpa kabar kini mengalami perluasan makna dalam konteks hubungan sosial digital. Begitu pula dengan kata viral yang bergeser dari konteks biologi ke ranah popularitas konten.

Melalui penerapan NLP (Natural Language Processing), teridentifikasi adanya pola linguistik lintas platform. Twitter cenderung melahirkan kosakata singkat karena keterbatasan karakter, sementara TikTok memunculkan istilah baru yang erat kaitannya dengan tren musik dan visual. Pola sintaksis yang muncul juga menunjukkan adanya penghilangan subjek dalam kalimat singkat, penggunaan bahasa campuran (code-switching), serta munculnya struktur baru berbasis tagar yang berfungsi ganda sebagai label sekaligus penanda identitas komunitas.

### **Pembahasan**

Hasil tersebut menegaskan bahwa dialek digital merupakan fenomena linguistik yang berkembang pesat seiring dengan meningkatnya penggunaan media sosial global. Temuan kosakata baru dan pola sintaksis unik memperlihatkan bahwa bahasa dalam ranah digital bukan hanya variasi informal, tetapi juga membentuk sistem komunikasi tersendiri dengan aturan dinamis. Hal ini sejalan dengan pandangan linguistik modern yang menekankan bahwa bahasa selalu beradaptasi dengan konteks sosial dan teknologi.

Perubahan semantik pada kata-kata seperti *ghosting* atau *viral* mencerminkan adanya resemantisasi, yaitu proses pembentukan makna baru melalui konteks budaya digital. Fenomena ini mendukung teori bahwa bahasa tidak pernah statis, melainkan terus bertransformasi melalui interaksi sosial. Penggunaan NLP dalam penelitian ini memperlihatkan pola linguistik yang sebelumnya sulit diamati secara manual, terutama dalam skala data yang besar.

Selain itu, hasil analisis juga memperlihatkan adanya globalisasi dialek digital. Istilah yang lahir di suatu komunitas daring dapat dengan cepat menyebar ke berbagai wilayah dunia, melampaui batas geografis maupun linguistik. Hal ini memperkuat pandangan bahwa media sosial menjadi arena utama dalam pembentukan bahasa baru di era globalisasi digital. Dengan demikian, penelitian ini tidak hanya menggambarkan dinamika kosakata, tetapi juga menyoroti bagaimana teknologi berperan sebagai katalisator dalam evolusi bahasa manusia.

## 5. KESIMPULAN DAN SARAN

### Kesimpulan

Penelitian ini menunjukkan bahwa media sosial telah menjadi ruang yang subur bagi lahirnya dialek digital yang unik, dinamis, dan lintas budaya. Analisis linguistik komputasional melalui text mining, analisis semantik, dan NLP berhasil mengidentifikasi berbagai kosakata baru, pergeseran makna, serta pola linguistik yang khas di platform digital. Fenomena ini menegaskan bahwa bahasa dalam ruang daring tidak hanya mencerminkan kreativitas pengguna, tetapi juga berfungsi sebagai penanda identitas, solidaritas komunitas, sekaligus wujud transformasi komunikasi global.

Perubahan semantik seperti istilah *ghosting*, *viral*, maupun munculnya bentuk singkatan dan akronim baru membuktikan bahwa digitalisasi telah mempercepat proses inovasi bahasa. Selain itu, globalisasi media sosial memperkuat penyebaran kosakata lintas budaya sehingga mempertemukan unsur lokal, global, dan simbol digital dalam satu pola bahasa yang inklusif. Dengan demikian, penelitian ini menegaskan bahwa dialek digital bukan sekadar variasi informal, melainkan bagian dari evolusi bahasa modern yang relevan untuk dikaji lebih mendalam.

### Saran

Berdasarkan temuan penelitian, terdapat beberapa saran yang dapat diajukan. Pertama, penelitian lebih lanjut sebaiknya dilakukan dengan memperluas sumber data dari berbagai platform baru atau komunitas daring tertentu untuk mendapatkan gambaran yang lebih komprehensif. Kedua, diperlukan pengembangan metode analisis yang lebih canggih, seperti

integrasi machine learning **dan** deep learning, untuk menangkap dinamika bahasa digital secara real-time. Ketiga, penting bagi pendidik dan pembuat kebijakan bahasa untuk memahami fenomena dialek digital ini agar dapat menyusun strategi literasi digital yang seimbang antara kreativitas berbahasa dengan keterjagaan norma bahasa.

Dengan pendekatan yang lebih mendalam dan interdisipliner, penelitian tentang dialek digital diharapkan dapat memberikan kontribusi signifikan, baik bagi pengembangan ilmu linguistik maupun bagi praktik komunikasi dalam masyarakat global yang semakin terdigitalisasi.

## DAFTAR PUSTAKA

- Agüero-Torales, M. M., Abreu Salas, J. I., & López-Herrera, A. G. (2021). Deep learning and multilingual sentiment analysis on social media data: An overview. *Applied Soft Computing*, 107, 107373. <https://doi.org/10.1016/j.asoc.2021.107373>
- Antony, S., & Trambo, I. A. (2023). Digital metamorphosis: Unraveling the evolution of global language in the technological culture. *International Journal of English Language and Literature Studies*, 12(4), 358–373. <https://doi.org/10.55493/5019.v12i4.4921>
- Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine*, 9(2), 48–57. <https://doi.org/10.1109/MCI.2014.2307227>
- Darwish, K., Habash, N., Abbas, M., Al-Khalifa, H., Al-Natsheh, H. T., Bouamor, H., Bouzoubaa, K., Cavalli-Sforza, V., El-Beltagy, S. R., El-Hajj, W., Jarrar, M., & Mubarak, H. (2021). A panoramic survey of natural language processing in the Arab world. *Communications of the ACM*, 64(4), 72–81. <https://doi.org/10.1145/3447735>
- Friedrich, P., & De Figueiredo, E. H. D. (2016). *The sociolinguistics of digital Englishes*. Routledge. <https://doi.org/10.4324/9781315681184>
- Jørgensen, A. K., Hovy, D., & Søgaard, A. (2015). Challenges of studying and processing dialects in social media. In *ACL-IJCNLP 2015 – Workshop on Noisy User-Generated Text (WNUT 2015)* (pp. 9–18). Association for Computational Linguistics.
- Khensous, G., Labeled, K., & Labeled, Z. (2023). Exploring the evolution and applications of natural language processing in education. *Revista Română de Informatică și Automatică*, 33(2), 61–74. <https://doi.org/10.33436/v33i2y202305>
- Klushina, N. (2022). Communicative culture on the internet and media linguoecology. *Jazykovedný Časopis*, 73(1), 78–85. <https://doi.org/10.2478/jazcas-2022-0029>
- Mishra, S., Samal, C. K., Yadav, N., & Choudhury, R. K. (2019). The impact of various digitized social networking media through text, images and videos on language usage. *International Journal of Scientific and Technology Research*, 8(10), 1968–1971.
- Prihatini, A., Kusuma Zamahsari, G., & Pangesti, F. (2023). Social media benefit for advancing language ability in E-learning environment: A systematic literature review. In

Proceedings of the 2023 17th International Conference on Ubiquitous Information Management and Communication (IMCOM 2023). IEEE. <https://doi.org/10.1109/IMCOM56909.2023.10035623>

Rani, P., & Samjetsabam, G. (2024). Cultural discourse analysis and digital inequalities in a multicultural context: A case study of Manipur in India. In *Digital inequality: Studies in cultural communication* (pp. 147–162). Routledge. <https://doi.org/10.4324/9781032672250-9>

Skaria, S., Krishna, P. K. A., Anagha, K., James, S. K., & Nargeese, V. (2024). Natural language processing (NLP): Advancements, applications, and challenges. *AIP Conference Proceedings*, 3134(1), 020005. <https://doi.org/10.1063/5.0230441>

Spilioti, T. (2018). The language of social media. In *The Routledge handbook of English language studies* (pp. 310–323). Routledge. <https://doi.org/10.4324/9781351001724-22>

Sun, Y., Wang, G., & Feng, H. (2021). Linguistic studies on social media: A bibliometric analysis. *SAGE Open*, 11(3). <https://doi.org/10.1177/21582440211047572>

Sundaram, A., Subramaniam, H., Hamid, S. H. A., & Nor, A. M. (2023). A systematic literature review on social media slang analytics in contemporary discourse. *IEEE Access*, 11, 132457–132471. <https://doi.org/10.1109/ACCESS.2023.3334278>