



ANALISIS UJARAN KEBENCIAN TERHADAP PENGGUNAAN KECERDASAN BUATAN DAN DAMPAKNYA TERHADAP PERSEPSI MASYARAKAT DI PLATFORM X: KAJIAN TINDAK TUTUR ILOKUSI DAN PERLOKUSI

Kistosi Al Ghifari ^a, Pradana Yahya Abdillah ^{b*}, Tanaya Edra Abitama ^c, Muhammad Rizqon Karim ^d, Fahmilatul Khairunnafisah ^e, Siti Maesaroh ^f

^a Fakultas Teknik/ Teknologi Informasi; kistosialghifari2005@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

^b Fakultas Teknik/ Teknologi Informasi; pradanayahyaabdillah@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

^c Fakultas Teknik/ Teknik Elektro; tanayaedraabitama@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

^d Fakultas Teknik/ Teknik Elektro; muhammadrizqonkarim@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

^e Fakultas Teknik/ Teknik Biomedis; fahmilatulkhairunnafisah@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

^f Fakultas Ilmu Budaya/ Sastra Arab; sitimaesaroh2006@mail.ugm.ac.id, Universitas Gadjah Mada; Kabupaten Sleman Provinsi Daerah Istimewa Yogyakarta

* Penulis Korespondensi: Kistosi Al Ghifari, Pradana Yahya Abdillah

ABSTRACT

The rapid development of artificial intelligence (AI) in Indonesia has sparked diverse public reactions on social media, with some expressions crossing the line into hate speech rather than remaining constructive criticism. This study aims to identify and classify forms of hate speech against the use of AI on platform X (Twitter) and to analyze its perlocutionary effects on public perception of the technology. A descriptive qualitative approach based on Searle's speech act theory is employed, focusing on illocutionary classification and the analysis of perlocutionary impact, with the locutionary level serving as a contextual framework. Data were collected through purposive sampling of 37 posts containing negative reactions to the use of AI, gathered between 2023 and 2026. The analysis shows that hate speech against AI manifests itself in five illocutionary categories: assertive (40%), directive (36%), expressive (28%), declarative (24%), and commissive (20%), with directive and assertive acts being the most frequent. The identified perlocutionary effects include the reinforcement of fear-based narratives about AI, the emergence of social polarization between AI users and opponents, and pressure to conform to digital behavioral norms. These findings suggest that hate speech against AI is not merely an expression of individual dissatisfaction but a discursive practice that shapes the collective perception of technology. This research contributes to the development of frameworks for digital literacy and AI governance policies in Indonesia by providing a pragmatic-linguistic perspective that has been largely absent from previous sentiment-based research.

Keywords: *hate speech; artificial intelligence; speech act; illocution; perlocution; X*

Abstrak

Perkembangan pesat kecerdasan buatan (AI) di Indonesia memicu berbagai reaksi publik di media sosial, dengan kritik konstruktif yang kadang berubah menjadi ujaran kebencian. Studi ini bertujuan untuk mengidentifikasi dan mengklasifikasikan bentuk ujaran kebencian terhadap penggunaan AI di platform X (Twitter) dan menganalisis dampak perlokusi terhadap persepsi publik. Pendekatan yang digunakan adalah deskriptif kualitatif dan berdasarkan teori tindak tutur Searle. Fokusnya adalah pada klasifikasi ilokusi dan

analisis dampak perlokusi, dengan tingkat lokusi sebagai dasar kontekstual. Data dikumpulkan melalui pengambilan sampel purposif sebanyak 37 unggahan di platform *X* yang berisi reaksi negatif terhadap penggunaan AI, yang dikumpulkan antara tahun 2023 dan 2026. Analisis menunjukkan bahwa ujaran kebencian terhadap AI bermanifestasi dalam lima kategori ilokusi: asertif (40%), direktif (36%), ekspresif (28%), deklaratif (24%), dan komisif (20%), dengan tindakan direktif dan asertif menjadi yang paling dominan. Efek perlokusi yang teridentifikasi meliputi penguatan narasi berbasis ketakutan tentang AI, munculnya polarisasi sosial antara pengguna dan penentang AI, dan tekanan untuk menyesuaikan diri dengan pola perilaku digital. Temuan ini menunjukkan bahwa ujaran kebencian terhadap AI bukan hanya ekspresi ketidakpuasan individu, melainkan praktik diskursif yang membentuk persepsi kolektif terhadap teknologi tersebut. Penelitian ini berkontribusi pada pengembangan kerangka kerja untuk literasi digital dan kebijakan tata kelola AI di Indonesia dengan memperkenalkan perspektif pragmatis-linguistik yang sebagian besar tidak ada dalam penelitian berbasis sentimen sebelumnya.

Kata Kunci: ujaran kebencian; kecerdasan buatan; tindak tutur; ilokusi; perlokusi; *X*

1. PENDAHULUAN

Kecerdasan buatan (AI) adalah salah satu inovasi teknologi yang paling cepat berkembang dalam dekade terakhir. AI didefinisikan sebagai program komputer yang dilengkapi dengan algoritma untuk mengekstrak pengetahuan dari data dan menerapkannya pada pemikiran dan perilaku seperti manusia [1]. Sejak peluncuran ChatGPT pada akhir tahun 2022 dan munculnya berbagai model AI generatif berbasis gambar dan video, teknologi ini telah menyebar dengan cepat di Indonesia, menjangkau hampir semua segmen masyarakat. Adopsi yang meluas ini membuka peluang baru untuk produktivitas dan inovasi; namun, hal ini juga menimbulkan kecemasan kolektif yang menyertai setiap lompatan teknologi besar.

Reaksi publik di Indonesia terhadap perkembangan AI menunjukkan tren yang didominasi bersifat negatif. Analisis sentimen di platform *X* (dahulu Twitter) menunjukkan bahwa kosakata yang paling sering digunakan dalam wacana publik tentang AI didominasi oleh istilah negatif seperti berbahaya, menakutkan, mengerikan, cemas, dan memusingkan [1]. Temuan ini didukung oleh studi tentang reaksi pengguna internet terhadap kampanye AI pemerintah. Menurut studi ini, 68,75% reaksi bersifat negatif, dengan kritik utama terkait pengabaian para profesional kreatif, kurangnya regulasi, dan politisasi teknologi [2]. Lebih lanjut, telah ditunjukkan bahwa penyalahgunaan AI dalam bentuk deepfake dan konten manipulatif mendistorsi persepsi publik, merusak kepercayaan pada informasi digital, dan memperburuk polarisasi sosial [3].

Reaksi negatif terhadap AI di media sosial tidak selalu terbatas pada kritik konstruktif. Beberapa pernyataan publik melampaui batas kritik dan masuk ke dalam kategori ujaran kebencian pernyataan yang bertujuan untuk mencemarkan nama baik, menyerang, atau memprovokasi kelompok tertentu karena penggunaan atau dukungan mereka terhadap teknologi AI. Fenomena ujaran kebencian itu sendiri telah menjadi masalah serius di Indonesia baik dari perspektif linguistik maupun hukum, karena peraturan yang ada tidak secara spesifik mengatur semua bentuk ekspresi [4]. Faktor psikologis dan sosial, termasuk ketidakamanan ekonomi dan kurangnya literasi digital, berkontribusi pada penyebaran ujaran kebencian daring [5]. Dampaknya pada pengguna media sosial meliputi kecemasan, trauma sekunder, dan perubahan persepsi sosial terhadap target audiens [6]. Namun, belum ditemukan penelitian yang mengkaji ujaran kebencian terhadap penggunaan AI sebagai subjek analisis pragmatik. Penelitian sebelumnya umumnya mengandalkan analisis sentimen berbantuan komputer, yang mengkategorikan bahasa ke dalam kategori positif, negatif, dan netral. Oleh karena itu, mereka belum mampu mengungkap struktur tindak tutur, maksud ilokusi, dan efek perlokusi yang dimiliki bahasa terhadap pembaca.

Penelitian ini berfokus pada platform media sosial *X*, karena platform ini mewakili ruang digital paling produktif untuk wacana publik tentang teknologi di Indonesia dan sering menjadi tempat pertemuan bagi kelompok pengguna dan penentang AI. Pembatasan penelitian pada satu platform memastikan homogenitas fitur linguistik data, karena konvensi penulisan dan gaya interaksi berbeda di berbagai platform. Pertanyaan penelitian terdiri dari tiga aspek: (1) Bentuk ujaran kebencian terhadap penggunaan AI apa yang terjadi di platform media sosial *X*? (2) Pola apa yang ditunjukkan oleh tindak ilokusi ujaran ini menurut klasifikasi Searle? (3) Apa efek perlokusi ujaran kebencian terhadap AI terhadap persepsi publik di ruang digital?.

2. TINJAUAN PUSTAKA

2.1. Ujaran Kebencian di Media Sosial

Ujaran kebencian adalah pernyataan yang bertujuan untuk menyerang, merendahkan, atau memprovokasi individu atau kelompok melalui pernyataan diskriminatif. Di Indonesia, ujaran kebencian di media sosial telah menjadi masalah hukum dan sosial yang serius, karena ketentuan Undang-Undang Informasi dan Transaksi Elektronik (ITE) tidak dapat mengatasi semua manifestasi fenomena ini [4]. Studi tentang ujaran kebencian di Twitter menunjukkan bahwa platform ini merupakan lahan subur bagi berbagai bentuk ujaran kebencian, mulai dari penghinaan, pencemaran nama baik, dan penistaan agama hingga tindakan menjijikkan, disinformasi, dan provokasi [7]. Pendekatan analisis wacana kritis Van Dijk telah digunakan untuk mengungkap ideologi dan struktur kekuasaan di balik ujaran kebencian terhadap etnis Tionghoa di Twitter [8]. Pendekatan siberpragmatik, yang menggabungkan analisis tindak tutur dengan karakteristik komunikasi siber, menawarkan klasifikasi enam jenis ujaran kebencian [9]. Dari segi psikologis, paparan ujaran kebencian telah terbukti menyebabkan kecemasan, trauma sekunder, dan perubahan jangka panjang dalam persepsi sosial [6].

2.2. Teori Tindak Tutur Searle: Fokus Ilokusi dan Perlokusi

Teori tindak tutur merupakan konsep sentral dalam pragmatik, yang dirumuskan oleh Austin dan diuraikan secara rinci oleh Searle. Teori ini membagi bahasa menjadi tiga tingkatan: tindak lokusioner, yang berkaitan dengan makna literal dari apa yang dikatakan; tindak ilokusi, yang menggambarkan maksud atau tujuan berbicara; dan tindak perlokusi, yang menggambarkan efek dari apa yang dikatakan pada pendengar [10]. Searle mengklasifikasikan tindak ilokusi menjadi lima jenis utama: (1) ujaran asertif, yang membuat pernyataan tentang dunia; (2) ujaran direktif, yang dimaksudkan untuk mendorong pendengar melakukan suatu tindakan; (3) ujaran komisif, yang mewajibkan pembicara untuk melakukan tindakan di masa mendatang; (4) ujaran ekspresif, yang mengungkapkan keadaan psikologis pembicara; dan (5) ujaran deklaratif, yang mengubah status sosial atau realitas [10]. Sebuah studi tentang ujaran kebencian berdasarkan model Searle menunjukkan bahwa tindak tutur yang paling dominan dalam ujaran kebencian di media sosial bersifat ekspresif dan direktif [11].

Berdasarkan analisis ini, studi ini secara eksplisit berfokus pada dua tingkat tindak tutur Searle: tindak tutur sebagai unit klasifikasi dan tindak perlokusi sebagai unit efek. Tindak tutur berfungsi sebagai dasar kontekstual untuk memahami makna literal ujaran tetapi bukan fokus analisis. Fokus ini didasarkan pada asumsi bahwa pertanyaan penelitian menekankan dua aspek: identifikasi pola dalam ujaran kebencian (fungsi ilokusi) dan investigasi pengaruhnya terhadap persepsi publik (fungsi perlokusi). Tindak lokusi berfungsi sebagai titik awal untuk mengidentifikasi tindak ilokusi. Analisis utama bertujuan untuk (a) mengklasifikasikan maksud ujaran menurut lima kategori ilokusi Searle dan (b) memeriksa pengaruh tindak perlokusi terhadap persepsi pembaca tentang teknologi AI. Pendekatan ini baru, karena studi-studi sebelumnya umumnya berfokus pada pengklasifikasian tindakan ilokusi tanpa mempertimbangkan efek perlokusi dalam wacana publik.

2.3. Persepsi Publik terhadap Kecerdasan Buatan

Persepsi publik terhadap AI beragam dan dinamis. Analisis sentimen global menunjukkan bahwa opini publik tentang AI terbagi antara optimisme mengenai manfaatnya dan kekhawatiran mendalam tentang dampak sosialnya [12]. Di Indonesia, sebuah studi berbasis berbasis awan kata (word cloud) dari X-Data mengidentifikasi ekspresi negatif sebagai kosakata dominan [1]. Sebuah studi berbasis algoritma Naive Bayes yang meneliti reaksi pengguna internet terhadap kampanye AI pemerintah menemukan bahwa 68,75% reaksi bersifat negatif. Kritik utama termasuk pengabaian terhadap para profesional kreatif dan kurangnya regulasi [2]. Penyalahgunaan AI dalam bentuk deepfake dan konten manipulatif juga berkontribusi pada hilangnya kepercayaan publik terhadap informasi digital [3]. Studi internasional menunjukkan bahwa identitas moderator dan tujuan moderasi memengaruhi penerimaan pengguna terhadap keputusan moderasi konten [9]. Temuan ini relevan dengan penelitian saat ini, karena ujaran kebencian terhadap AI di X juga merupakan ekspresi ketidakpercayaan terhadap kemampuan AI untuk bertindak sebagai agen.

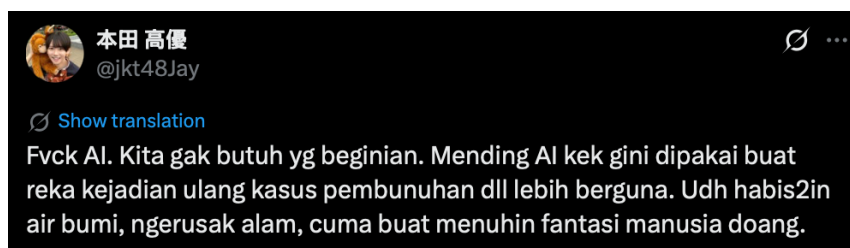
2.4. Literasi Digital sebagai Strategi Mitigasi

Literasi digital merupakan keterampilan penting untuk menghadapi ujaran kebencian dan disinformasi di media sosial. Model literasi digital yang mengintegrasikan nilai-nilai moral telah terbukti efektif dalam menekan penyebaran ujaran kebencian, karena tidak hanya membekali masyarakat dengan keterampilan teknis tetapi juga menumbuhkan kesadaran etika dalam interaksi di ruang digital [13]. Penguatan literasi

digital di Indonesia menjadi semakin mendesak untuk mengembangkan karakter warga digital yang berpikir kritis [14]. Model komunikasi yang dirancang khusus untuk mempromosikan literasi digital dalam memerangi ujaran kebencian menunjukkan bahwa pendekatan kolaboratif antara literasi digital dan komunikasi berbasis nilai dapat menjadi strategi yang efektif untuk menekan ujaran kebencian [15]. Secara global, ujaran kebencian dianggap sebagai masalah kesehatan masyarakat yang membutuhkan pendekatan terpadu yang mencakup teknologi, keterampilan, dan regulasi [16].

3. METODOLOGI PENELITIAN

Studi ini menggunakan pendekatan deskriptif kualitatif, yang memungkinkan pemahaman yang lebih mendalam tentang fenomena sosial melalui analisis data non-numerik. Kerangka teoritis utama adalah teori tindak tutur Searle, dengan fokus analitis pada klasifikasi ilokusi dan investigasi efek perlokusi (lihat subbagian 2.2). Sumber data terdiri atas postingan dan komentar berbahasa Indonesia di platform *X* yang berisi reaksi negatif terhadap penggunaan AI. Sebagai ilustrasi, Gambar 1. menunjukkan salah satu bentuk data mentah berupa komentar yang mengandung tindak tutur ekspresif (kritik/kekesalan) sebelum masuk ke tahap reduksi data.



Gambar 1. Contoh postingan ujaran kebencian terhadap AI di platform *X*

Fokus pada satu platform ditentukan oleh homogenitas konvensi linguistik dan kepadatan wacana publik terkait teknologi yang lebih tinggi di *X* dibandingkan dengan platform lain.

Data dikumpulkan melalui pengambilan sampel purposif berdasarkan kriteria inklusi berikut: (a) ujaran yang dipublikasikan di *X* antara tahun 2023 dan 2026; (b) ujaran dalam bahasa Indonesia, termasuk varietas informal dan kode campuran Indonesia-Inggris; (c) ujaran yang secara eksplisit ditujukan terhadap penggunaan AI atau pengguna AI; dan (d) pernyataan yang mengandung unsur ujaran kebencian, yaitu serangan, penghinaan, provokasi, atau seruan untuk diskriminasi. Pencarian dilakukan menggunakan kata kunci seperti “benci AI,” “tolak AI,” “boikot AI,” “hentikan AI,” “AI itu buruk,” “AI itu berbahaya,” dan “AI harus dilarang.” Tiga puluh tujuh pernyataan dikumpulkan dan kemudian disaring berdasarkan kriteria yang disebutkan di atas. 37 pernyataan yang terpilih membentuk korpus utama untuk analisis.

Analisis data dilakukan dalam empat langkah berurutan. Pertama, data diidentifikasi dan dipilih menggunakan kriteria klasifikasi pragmatis [9]. Pada langkah kedua, dilakukan analisis lokusioner, dengan memeriksa makna literal pernyataan dan konteksnya. Pada langkah ketiga, dilakukan klasifikasi ilokusi, mengkategorikan pernyataan berdasarkan maksud komunikatifnya dalam lima kategori Searle. Keempat, analisis perlokusi meneliti dampak pernyataan terhadap pembaca dan persepsi kolektif dengan membandingkannya dengan penelitian sebelumnya tentang opini publik mengenai AI di Indonesia [1-3]. Validitas data dipastikan dengan melakukan triangulasi sumber dalam platform yang sama, yaitu dengan menggunakan data dari berbagai laporan, periode waktu, dan konteks wacana yang berbeda.

4. HASIL DAN PEMBAHASAN

4.1 Distribusi Tindak Ilokusi Ujaran Kebencian terhadap AI

Analisis terhadap 37 pernyataan terpilih menunjukkan bahwa ujaran kebencian terhadap penggunaan AI di *X* dapat dikategorikan ke dalam lima kategori ilokusi Searle. Distribusi tindak ilokusi ujaran kebencian terhadap AI di platform *X* adalah sebagai berikut: tindak asertif sebanyak 10 data (40%) dengan ciri klaim negatif tentang AI; tindak ekspresif 7 data (28%) berupa ungkapan kebencian dan ketakutan; tindak direktif 9 data (36%) berupa larangan dan ajakan boikot; tindak komisif 5 data (20%) berupa komitmen memblokir atau tidak menggunakan AI; serta tindak deklaratif 6 data (24%) berupa pernyataan legitimasi penolakan dan tuntutan pelarangan AI. Hasil menunjukkan bahwa tindakan asertif merupakan kategori yang paling sering muncul, diikuti oleh pernyataan direktif, ekspresif, deklaratif, dan komisif. Komposisi ini menunjukkan dua

poin penting. Pertama, ujaran kebencian terhadap AI di *X* bukan hanya luapan emosi (ekspresif) tetapi juga pernyataan faktual (asertif) yang berfungsi sebagai dasar argumentasi untuk menentanginya. Kedua, tindakan yang sangat direktif menunjukkan bahwa ujaran kebencian terhadap AI memiliki orientasi normatif, yaitu upaya untuk memengaruhi perilaku pengguna lain agar mereka tidak lagi menggunakan AI.

4.2 Klasifikasi Tindak Ilokusi dan Analisis Data

4.2.1 Tindak Ilokusi Asertif

Tindakan asertif dalam korpus tersebut berbentuk klaim negatif tentang AI, baik yang bersifat ekonomi, moral, ekologis, maupun kognitif. Klaim-klaim ini dirumuskan menggunakan kalimat deklaratif yang menyiratkan kepastian melalui konstruksi kondisional kausal.

Data (1) – @daniorerio235: *"Kalo AI menggantikan sebagian besar pekerjaan manusia terus pengangguran makin banyak, yang ada dunia ini bakal chaos."*

Dalam dataset (1), ujaran lokusi adalah kalimat kondisional yang menghubungkan pengenalan AI dengan kekacauan sosial. Ujaran ilokusi bersifat asertif dan menyajikan prediksi tentang dampak sosioekonomi AI, sedangkan ujaran perlokusi memperkuat ketakutan terhadap AI sebagai ancaman terhadap stabilitas sosial. Pernyataan ini selaras dengan hasil analisis sentimen yang mengidentifikasi penggunaan istilah seperti bahaya dan teror sebagai dominan dalam wacana publik tentang AI [1].

Data (2) – @otaknyakonslet: *"Visinya ngeri eh. Manusia dibikin tumpul, goblok, kehilangan expertise dan ketergantungan dulu dgn segala kemudahan AI. Setelah itu baru mereka jualan kecerdasan atau expertise lewat produk AI mereka. Prinsip yg sama dgn narkoba."*

Data (2) membuat klaim provokatif yang menyamakan AI dengan narkoba sebagai zat adiktif. Pernyataan tersebut panjang dan terstruktur, dengan struktur argumentatif yang mengingatkan pada kritik; namun demikian, pernyataan tersebut tetap provokatif dan sangat penuh kebencian, karena pengguna AI digambarkan sebagai membosankan dan bodoh istilah yang secara langsung merendahkan. Efek perlokusi yang dihasilkan membangun citra perusahaan AI sebagai penjahat dan pengguna AI sebagai korban yang pantas dihina. Penggunaan metafora narkoba mengaktifkan kerangka moral yang lebih luas, sehingga memperkuat polarisasi sosial yang diamati dalam studi tentang penyalahgunaan AI [3].

4.2.2 Tindak Ilokusi Ekspresif

Ekspresi dalam korpus tersebut menunjukkan kebencian, ketakutan, jijik, dan muak terhadap AI dan penggunaannya. Ciri linguistik yang konsisten adalah penggunaan kata kerja afektif (kebencian, ketakutan, jijik, dan muak) yang secara langsung berkolokasi dengan kata AI atau pengguna AI.

Data (3) – @poccariee: *"apapun yg pake ai gue benci, apalagi kalo keterusan."*

Data (3) adalah contoh tipikal dari tindakan ekspresif murni. Tindak lokusi yang ditemukan berupa ekspresi umum kebencian terhadap semua pengguna AI. Tindak ilokusi berupa ekspresi kebencian yang diperkuat oleh konstruksi yang meningkat, terutama ketika dilanjutkan. Tindakan perlokusi yang dihasilkan adalah normalisasi sikap negatif terhadap AI sebagai posisi yang dilegitimasi secara afektif, terutama karena ujaran ini memperoleh 49.200 tayangan. Ujaran ekspresif viral ini berfungsi sebagai model emosional dengan memberikan legitimasi sosial kepada pembaca untuk membenci AI tanpa argumentasi rasional.

Data (4) – @pearchly: *"sumpah takut banget sama AI ini, plis semoga kita dihindarkan dari orang jahat dan mereka cepet waras."*

Data (4) menunjukkan tindakan ekspresif berbasis kecemasan. Tindakan verbal terdiri dari pengakuan kecemasan, disertai dengan doa untuk keselamatan; tindakan oral tetap ekspresif tetapi menyajikan AI sebagai ancaman yang melampaui kemampuan individu, sementara secara bersamaan mencap pengguna AI sebagai jahat. Tindakan verbal yang dihasilkan memperkuat ketakutan terhadap AI sebagai entitas yang mengancam keselamatan dan martabat. Pola ini selaras dengan kekhawatiran mengenai penyalahgunaan data wajah dalam konten deepfake [3].

4.2.3 Tindak Ilokusi Direktif

Tindakan direktif terwujud dalam bentuk larangan, seruan boikot, dan perintah untuk menghentikan penggunaan AI. Ciri linguistik yang konsisten meliputi kata kerja imperatif (berhenti, jangan, hentikan), penggunaan huruf kapital untuk penekanan, dan penggunaan tagar sebagai seruan kolektif untuk bertindak.

Data (5) – @MikazukiArion: *"Kalau kalian gabisa gambar, banyak artist yang pasti available untuk commission dan ingat jangan pakai AI!"*

Data (5) mengungkapkan tindakan arahan yang dikemas dalam logika alternatif. Ujaran tersebut menawarkan solusi (menugaskan seniman) sebelum mengeluarkan larangan (jangan menggunakan AI). Ujaran arahan ini diperkuat oleh konteks komunitas seniman digital yang merasa tergantikan oleh AI generatif. Perlokusi yang dihasilkan adalah pembentukan norma kelompok yang memandang penggunaan AI sebagai pengkhianatan terhadap solidaritas profesional, sehingga menciptakan tekanan untuk menyesuaikan diri yang memengaruhi keputusan pengguna lain.

Data (6) – @r0ckwithhoon: *"EH SIALAN, STOP USING AI BUAT NGEDIT NGEDIT BEGINI KAMPUNG."*

Data (6) mengintegrasikan umpatan (SIALAN), seluruh huruf kapital, dan istilah negatif (KAMPUNG) untuk memperkuat pengaruh perintahnya. Lokusi ini terdapat dalam larangan untuk memodifikasi gambar menggunakan AI; ilokusi menyampaikan perintah untuk menghentikan yang dibungkus dalam emosi marah sekaligus mengandung nada merendahkan bagi pelakunya. Perlokusi yang timbul adalah penguatan stigma sosial terhadap pengguna AI sebagai kelompok yang dipandang tidak beradab. Pola diskursus ini menunjukkan bahwa perintah dan ungkapan emosional seringkali saling terjalin, menjadikan ujaran kebencian terhadap AI bersifat kompleks dalam konteks pragmatik.

4.2.4 Tindak Ilokusi Komisif

Tindakan komisif yang ditemukan dalam korpus melibatkan kewajiban pembicara untuk memberikan sanksi sosial kepada pengguna AI, seperti memblokir mereka atau berhenti mengikuti mereka. Fitur linguistik yang paling penting adalah kata kerja modal bentuk masa depan (akan, saya blokir, saya berhenti mengikuti), yang menempatkan penolakan dalam konteks tindakan yang dijanjikan.

Data (7) – @demieqod: *"Yg pake ai art ghibli gw unfollow."*

Data (7) menunjukkan komisi yang paling umum. Pernyataan tersebut merupakan deklarasi prinsip pribadi, sebuah komitmen untuk tidak lagi mengikuti akun pengguna yang dihasilkan AI dalam gaya Ghibli; ilokusi yang ditemukan berupa komitmen untuk menerapkan sanksi sosial. Perlokusi yang dihasilkan adalah pembentukan ekspektasi normatif di antara pembaca bahwa penggunaan AI berisiko kehilangan basis pengikut mereka, sehingga memberikan tekanan implisit pada pembuat konten untuk menahan diri dari menggunakan AI.

Data (8) – @castramakara: *"Akan gue BLOKIR semua akun hiburan yang pake AI. JELEK."*

Data (8) menunjukkan tindakan komisif yang sepenuhnya dikapitalisasi untuk kata kerja janji (BLOKIR) dan pernyataan nilai (JELEK). Tindakan ucapan komisif ini tampak lebih agresif karena memiliki cakupan yang luas (semua akun hiburan) dan disertai dengan justifikasi estetika. Tindakan perlokusi yang dihasilkan memperkuat pesan bahwa penggunaan AI dihukum secara sosial, sekaligus menciptakan citra bahwa konten AI tidak menyenangkan. Jenis tindakan komisif ini berkontribusi pada pembentukan ekonomi perhatian yang menghukum pengembang AI.

4.2.5 Tindak Ilokusi Deklaratif

Tindakan deklaratif dalam kumpulan yang berbentuk pernyataan berupaya untuk melegitimasi penolakan terhadap AI sebagai posisi standar yang seharusnya diterapkan secara universal. Ciri linguistik yang terlihat jelas adalah penggunaan modalitas keharusan (harus, diwajibkan, seharusnya), istilah yang terkait dengan kebijakan (regulasi, sanksi), dan pembentukan argumen yang menempatkan penutur dalam posisi sebagai perwakilan nilai yang sama.

Data (9) – @awkmellon: *"Penggunaan Video AI harus diatur sih ini. Wajib pake watermark. Yang publish video ngga pake Watermark, didenda 100jt per video per publish."*

Data (9) menampilkan deklaratif paling eksplisit dalam korpus karena tuturan ini menetapkan kerangka kebijakan secara terperinci, dari kewajiban watermark hingga besaran denda. Lokusi yang ditemukan berupa proposal regulasi; ilokusi deklaratif berupa penetapan norma yang seakan-akan otoritatif. Perlokusi yang dihasilkan adalah penanaman keyakinan publik bahwa AI memerlukan pengaturan ketat, sekaligus penguatan tuntutan kebijakan ke arah pembatasan pola yang konsisten dengan keluhan netizen tentang minimnya regulasi AI [2].

Data (10) – @anamoelya: *"ya Allah apakah aku harus minta teknologi generate AI yang banyak mudharatnya di mana-mana ini dimusnahkan saja dari muka bumi ini mumpung lagi hari Arafah."*

Data (10) memperlihatkan deklaratif yang dibingkai dengan modalitas religius dan eskalasi retorik menuju permintaan ekstrem (dimusnahkan dari muka bumi). Lokusi berbentuk pertanyaan retorik berkonten doa; ilokusi deklaratif yang melegitimasi penolakan AI sebagai posisi moral keagamaan. Perlokusi yang dihasilkan adalah pelebaran basis legitimasi penolakan, yaitu dari ranah estetika dan ekonomi ke ranah moral-religius, sehingga ujaran kebencian terhadap AI lebih sulit ditangkis dengan argumen rasional-teknis karena bersandar pada kerangka nilai di luar wacana teknologi itu sendiri.

4.3. Pola Linguistik dan Diksi

Ujaran kebencian terhadap AI di *X* tidak hanya dapat diklasifikasikan menggunakan kategori ilokusi, tetapi juga menunjukkan pola linguistik yang konsisten. Pertama, korpus didominasi oleh kata-kata dengan konotasi negatif yang kuat (benci, jijik, muak, jelek, takut, terkutuk) yang secara langsung berdialog dengan kata AI. Pola ini selaras dengan hasil analisis word cloud sentimen publik [1]. Kedua, korpus menunjukkan penggunaan campuran bahasa Indonesia-Inggris yang menonjol (misalnya, "Berhenti menggunakan AI, itu benar-benar berbahaya") serta penyertaan kosakata budaya pop Korea (misalnya, "Shibal") dan istilah komunitas seperti "Penyepeng" dan "Cenblue," yang berfungsi sebagai penanda identitas komunitas. Ketiga, korpus menunjukkan penggunaan intensif perangkat penekanan visual dalam bentuk huruf kapital, huruf berulang, tanda seru bertumpuk, dan emotikon marah ini berfungsi sebagai pengganti intonasi dalam komunikasi tertulis. Keempat, ujaran kebencian terhadap AI sering disertai dengan metafora yang memperluas kerangka moral dari ujaran tersebut, seperti metafora narkoba dalam Data (2), metafora pencuri pekerja, dan metafora spons, yang menggeser AI dari ranah teknologi netral ke ranah moralitas, sehingga penggunaan AI dianggap sebagai pelanggaran moral daripada keputusan teknis.

4.4. Efek Perlokusi Terhadap Persepsi Masyarakat

Analisis efek perlokusi dilakukan dengan membandingkan pola bahasa korpus dengan studi opini publik sebelumnya. Hasilnya mengidentifikasi tiga efek perlokusi utama: penguatan narasi berbasis ketakutan, peningkatan polarisasi sosial, dan tekanan untuk menyesuaikan diri dengan perilaku digital. Pertama, pernyataan asertif yang menggambarkan AI sebagai penyebab kekacauan sosial (Dataset 1), pernyataan ekspresif yang mengungkapkan ketakutan (Dataset 4), dan pernyataan deklaratif yang menyerukan larangan ekstrem (Dataset 10) secara konsisten menghasilkan narasi bahwa AI merupakan ancaman eksistensial. Narasi ini selaras dengan dominasi kata-kata "bahaya," "kengerian," "menakutkan," dan "ketakutan" dalam awan kata opini publik [1] dan temuan bahwa kekhawatiran utama pengguna internet adalah penggantian pekerja kreatif [3]. Efek perlokusi adalah penguatan kerangka kognitif berbasis ketakutan, kerangka berpikir tentang AI yang berfokus pada bahaya. Kedua, pernyataan direktif yang melarang penggunaan AI (Data 5, 6) dan pernyataan komisif yang mengancam sanksi sosial (Data 7, 8) menciptakan garis pemisah yang jelas antara pengguna AI dan penentangnya. Penggunaan label yang menstigmatisasi (pengguna AI, pejuang, idiot) mengubah perbedaan dalam penggunaan teknologi menjadi kategori moral yang berbeda. Polarisasi ini sejalan dengan temuan tentang penyalahgunaan AI, yang menunjukkan bahwa wacana publik tentang teknologi mudah terpolarisasi ketika dimediasi oleh konten manipulatif dan retorika yang menghasut [3]. Efek perlokusi adalah pembentukan kelompok dalam dan kelompok luar dalam wacana teknologi, yang berpotensi membatasi ruang untuk dialog konstruktif tentang tata kelola AI. Ketiga, kombinasi dari pidato direktif yang meluas, pidato pengumuman dengan sanksi, dan deklarasi penetapan norma menciptakan tekanan konformitas yang memaksa pengguna untuk menyembunyikan penggunaan AI mereka di ruang publik digital. Studi tentang efek psikologis ujaran kebencian menunjukkan bahwa paparan terhadap jenis ujaran ini dapat menyebabkan kecemasan, tekanan psikologis, dan perubahan perilaku digital [6]. Ringkasan efek perlokusi ujaran kebencian terhadap AI dan keselarasannya dengan riset terdahulu adalah sebagai

berikut: (a) Penguatan narasi berbasis ketakutan, termanifestasi dalam klaim chaos sosial, ungkapan takut, dan tuntutan pelarangan ekstrem, yang selaras dengan temuan word cloud 'bahaya', 'ngeri', 'serem', 'takut' serta kekhawatiran pergantian pekerja kreatif; (b) Pembentukan polarisasi sosial, termanifestasi dalam label stigmatik, perintah boikot, dan sanksi blokir, yang selaras dengan temuan manipulasi opini publik melalui konten AI dan ujaran kebencian etnis; (c) Tekanan konformitas perilaku digital, termanifestasi dalam komisif unfollow atau blokir massal serta direktif normatif, yang selaras dengan temuan dampak psikologis ujaran kebencian berupa kecemasan dan perubahan perilaku.

Ketiga efek perlokusi ini menunjukkan bahwa ujaran kebencian terhadap AI di *X* tidak berakhir pada tingkat ujaran individu, tetapi berfungsi sebagai praktik diskursif yang membentuk persepsi kolektif, mendefinisikan kategori moral untuk teknologi, dan memberlakukan sanksi sosial tanpa otoritas formal. Implikasi penting dari temuan ini adalah bahwa intervensi terhadap ujaran kebencian terhadap AI tidak dapat hanya bergantung pada moderasi konten otomatis, tetapi membutuhkan pendekatan yang berlandaskan pragmatisme untuk mempromosikan literasi digital. Pendekatan ini memberdayakan masyarakat untuk mengenali tindakan ilokusi dalam bahasa dan memahami efek perlokusi [13], [15]. Pada tingkat kebijakan, temuan ini menggarisbawahi perlunya tata kelola AI yang transparan dan partisipatif untuk mengatasi keluhan substantif yang tersembunyi di balik bahasa yang sarat emosi [16].

5. KESIMPULAN DAN SARAN

Berdasarkan analisis tindak tutur Searle terhadap 37 contoh ujaran kebencian terhadap AI di platform *X* (Twitter) antara tahun 2023 dan 2026, penelitian ini sampai pada tiga kesimpulan utama. Pertama, ujaran kebencian terhadap AI termanifestasi dalam lima kategori ilokusi Searle, dengan tindak tutur asertif (40%) dan direktif (36%) sebagai kategori dominan, diikuti oleh tindak tutur ekspresif (28%), deklaratif (24%), dan komisif (20%). Komposisi ini menempatkan ujaran kebencian terhadap AI sebagai praktik diskursif yang argumentatif, normatif, dan afektif, serta menegakkan sanksi sosial. Kedua, ujaran kebencian terhadap AI di *X* ditandai dengan diksi afektif negatif, perpaduan bahasa Indonesia dan Inggris yang dipengaruhi budaya pop Korea, perangkat penekanan visual, dan penggunaan metafora yang menggeser AI dari ranah teknologi netral ke ranah moral. Ketiga, efek perlokusi yang dihasilkan mencakup penguatan narasi berbasis ketakutan, terciptanya polarisasi sosial antara pengguna dan penentang AI, serta tekanan untuk menyesuaikan diri dengan pola perilaku digital. Temuan ini menegaskan bahwa ujaran kebencian terhadap AI bukan sekadar reaksi emosional, melainkan praktik diskursif yang membentuk persepsi kolektif terhadap teknologi. Di bidang literasi digital, pendidik dan profesional komunikasi didorong untuk mengembangkan kurikulum yang menggabungkan kesadaran pragmatis. Di bidang pembuatan kebijakan, regulator AI di Indonesia didorong untuk menafsirkan ujaran kebencian terhadap AI sebagai sinyal keluhan substantif terkait regulasi dan keadilan distributif. Keterbatasan penelitian ini adalah ketergantungannya pada satu platform dan jumlah data yang kecil. Penelitian lebih lanjut disarankan untuk membandingkan pola ujaran kebencian terhadap AI di berbagai platform dan untuk mengintegrasikan pendekatan komputasional dan pragmatis untuk pemetaan yang lebih komprehensif.

Ucapan Terima Kasih

Penulis menyampaikan ucapan terima kasih kepada Ibu Zulisih Maryani, S.S., M.A., selaku dosen pengampu Mata Kuliah Wajib Kurikulum (MKWK) Bahasa Indonesia dan Penulisan Ilmiah Universitas Gadjah Mada atas bimbingan dan masukan yang membantu penyempurnaan artikel ini, khususnya dalam menajamkan fokus kajian pada satu platform serta memperjelas arah analisis linguistik tindak tutur Searle.

DAFTAR PUSTAKA

- [1] R. Wijanarko, D. E. Ratnawati, and P. P. Adikara, "Analisis Sentimen Dampak Perkembangan Artificial Intelligence (AI) pada Media Sosial *X*/Twitter Menggunakan Metode Random Forest," 2024. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [2] K. Siregar Alfarizi, S. Nasution, and P. Nabawy, "Analisis Sentimen Netizen Indonesia Terhadap Kampanye Penggunaan Kecerdasan Buatan Oleh Pemerintah Menggunakan Algoritma Naive Bayes," *Jurnal Ilmu Komputer Dan Informatika*, Jun. 2025, Accessed: May 29, 2026. [Online]. Available: <https://jurnal.globalscients.com/index.php/jiki>
- [3] E. O. Napitupulu, "Analisis Penyalahgunaan AI Dalam Memanipulasi Opini Publik di Media Sosial," *Jurnal Abdimas Kartika Wijayakusuma*, vol. 6, no. 4, Dec. 2025, doi: 10.26874/jakw.v6i4.1190.

- [4] S. Kartika and N. Nurhayati, “Ujaran Kebencian (Hate Speech) di Media Sosial dalam Konteks Hukum dan Perubahan Sosial (Studi Kasus pada Masyarakat Kota Medan),” *JURNAL MERCATORIA*, vol. 16, no. 1, pp. 99–106, Jun. 2023, doi: 10.31289/mercatoria.v16i1.7668.
- [5] F. Lailatus Sa and D. Evita Santi, “Faktor Produksi Ujaran Kebencian melalui Media Sosial,” *Jurnal Psikologi Perseptual*, vol. 6, no. 1, 2021, [Online]. Available: <http://jurnal.umk.ac.id/index.php/perseptual>
- [6] R. Abdillah *et al.*, “STUDI PSIKOLOGI SIBER TENTANG DAMPAK HATE SPEECH BAGI PENGGUNA MEDIA SOSIAL CYBER PSYCHOLOGY STUDY ON THE IMPACT OF HATE SPEECH FOR SOCIAL MEDIA USERS,” *SIBATIK JOURNAL | VOLUME*, vol. 2, no. 11, 2023, doi: 10.54443/sibatik.v2i11.1478.
- [7] W. Maulana, “Ujaran Kebencian Terhadap Jokowi pada Masa Pandemi Covid-19: Studi Kasus Twitter.”
- [8] A. A. Priyatna, K. A. Wibowo, and N. A. Syafirah, “Prasangka Etnis dan Ujaran Kebencian: Analisis Ujaran Kebencian terhadap Etnis Tionghoa di Twitter (X),” *Jurnal Komunikasi*, vol. 19, no. 1, pp. 147–162, Oct. 2024, doi: 10.20885/komunikasi.vol19.iss1.art9.
- [9] K. N. Widyatnyana, I. W. Rasna, and I. B. Putrayasa, “ANALISIS JENIS DAN MAKNA PRAGMATIK UJARAN KEBENCIAN DI DALAM MEDIA SOSIAL TWITTER,” 2023.
- [10] P. Dilanti and R. Panji Hermoyo, “Tindak Tutur Ilokusi Searle dalam Film Pendek Jarak Antar Kanvas Karya Turah Parthayana,” *Pendidikan*, 2024. [Online]. Available: <https://e-journal.my.id/onoma>
- [11] F. D. Wulandari and I. S. Hanum, “TINDAK TUTUR ILOKUSI DALAM UJARAN KEBENCIAN PADA KOMENTAR SELEBGRAM RIO DAN KEKEYI: KAJIAN PRAGMATIK.”
- [12] E. Appiah and A. Htait, “AI in the Public Eye: Analysing Social Media Sentiment and Opinion on Artificial Intelligence.”
- [13] M. Mansyur and J. Perintis Kemerdekaan, “Model Literasi Digital untuk Melawan Ujaran Kebencian di Media Sosial Digital Literacy Model to Counter Hate Speech on Social Media,” *Jurnal Ilmu Pengetahuan dan Teknologi Komunikasi*, vol. 22, no. 2, doi: 10.33164/iptekkom.22.2.2020.125-142.
- [14] A. Iriyani and D. Puji Lestari, “Literasi Digital sebagai Upaya Membangun Karakter Masyarakat Digital”.
- [15] A. Dwi Andriani, S. Ainul Fitri, and K. Muchtar, “MODEL KOMUNIKASI LITERASI DIGITAL DALAM MENGATASI UJARAN KEBENCIAN DI MEDIA SOSIAL,” 2024.
- [16] T. Nguyen, “Merging public health and automated approaches to address online hate speech,” *AI and Ethics*, vol. 4, no. 2, pp. 441–450, May 2024, doi: 10.1007/s43681-023-00281-w.