JURNAL ILMIAH DINAMIKA SOSIAL

VOLUME 9 NUMBER 2, AUG 2025 | pp.267-282 | E-ISSN : 2581-2424 | P-ISSN : 2597-3657 |
Website: https://journal.undiknas.ac.id/index.php/fisip/index
DOI: https://doi.org/10.38043/jids.v9i2.6882

PROMPT ENGINEERING DAN ETIKA KOMUNIKASI DALAM ERA KECERDASAN BUATAN: TANTANGAN DAN PELUANG

Chairunnisa Widya Priastuty*, Mohamad Syahriar Sugandi[2], Melati Budi Srikandi[3]

[1*, 2]Program Studi Ilmu Komunikasi, Fakultas Komunikasi dan Ilmu Sosial,
Telkom University, Bandung, Indonesia
E-mail: [1*]chnisaw@telkomuniversity.ac.id, [2]syahriar@telkomuniversity.ac.id

[3]Program Studi Ilmu Komunikasi, Fakultas Ilmu Sosial dan Humaniora,
Universitas Pendidikan Nasional, Bali, Indonesia
E-mail: [3]melatibs@undiknas.ac.id

ABSTRAK

Perkembangan kecerdasan buatan/*artificial intelligent* (AI), khususnya *large language models l*(LLM), telah mengubah cara manusia memproduksi dan menyampaikan pesan dalam berbagai konteks komunikasi. Studi ini bertujuan untuk mengkaji bagaimana praktik *prompt engineering*, teknik dalam merancang instruksi kepada AI, berperan dalam pembentukan pesan, serta bagaimana praktik ini berinteraksi dengan isu etika dan nilai komunikasi. Penelitian dilakukan dengan metode tinjauan pustaka sistematis terhadap berbagai jurnal ilmiah yang terbit dalam sepuluh tahun terakhir (2015–2025), mencakup bidang komunikasi, teknologi, dan etika digital. Hasil kajian menunjukkan bahwa *prompt engineering* tidak hanya menentukan efektivitas teknis respons AI, tetapi juga memiliki implikasi serius terhadap bias representasi, keamanan informasi, dan potensi disinformasi. Selain itu, literasi AI muncul sebagai elemen kunci dalam mengurangi risiko penyalahgunaan teknologi dan memperkuat komunikasi yang etis. Studi ini mengusulkan model konseptual yang menempatkan *prompt engineering* sebagai titik sentral dalam produksi pesan berbasis AI, dengan etika dan nilai komunikasi sebagai jembatan menuju desain komunikasi yang adil, aman, dan bertanggung jawab. Kesimpulannya, produksi pesan dalam era AI memerlukan pendekatan komunikasi yang reflektif, interdisipliner, dan berbasis nilai. Literasi digital dan pengembangan kebijakan etik perlu diintegrasikan ke dalam praktik komunikasi publik guna memastikan keberlanjutan komunikasi manusia-AI yang inklusif dan etis.

Kata kunci: komunikasi AI; *prompt engineering*; etika digital; literasi media; disinformasi, representasi.

Diserahakan : 2025-03-12,
Direvisi : 2025-04-30,
Diterima : 2025-08-20

ABSTRACT

The development of artificial intelligence (AI), particularly large language models (LLMs), has transformed how humans produce and communicate messages across various communication contexts. This study aims to examine how prompt engineering practices—techniques for designing instructions for AI—influence message formation and how these practices interact with ethical issues and communication values. The research was conducted using a systematic literature review of various scientific journals published in the last ten years (2015–2025), covering the fields of communication, technology, and digital ethics. The results of the study show that prompt engineering not only determines the technical effectiveness of AI responses but also has serious implications for representation bias, information security, and the potential for disinformation. Additionally, AI literacy emerges as a key element in reducing the risk of technology misuse and strengthening ethical communication. This study proposes a conceptual model that positions prompt engineering as the central point in AI-based message production, with ethics and communication values as the bridge toward fair, safe, and responsible communication design. In conclusion, message production in the AI era requires a reflective, interdisciplinary, and value-based approach to communication. Going forward, digital literacy and ethical policy development need to be integrated into public communication practices to ensure the sustainability of inclusive and ethical human-AI communication.

Keywords: *AI communication; prompt engineering; digital ethics; media literacy; disinformation; representation.*

PENDAHULUAN

Perkembangan pesat kecerdasan buatan/*artificial intelligent* (AI), terutama model bahasa besar (*Large Language Models*/LLM), telah membawa perubahan fundamental dalam cara produksi pesan dan interaksi komunikatif. Teknik *prompt engineering* telah menjadi inti strategi untuk memaksimalkan hasil AI diantaranya merancang *prompt* yang jelas, spesifik, serta bersifat *few-shot* atau *chain of thought* sehingga model mampu menghasilkan *output* yang tidak hanya akurat, tetapi juga relevan dan bernuansa manusia, sebagaimana penelitian yang dilakukan Olla et al. (2024). Penelitian tersebut menekankan strategi retorika diterapkan dalam *prompt* melalui pendekatan *human in the loop* untuk meningkatkan kualitas komunikasi AI . Selain itu, konsep "*promptology*" mendefinisikan praktik sistematis seputar desain *prompt*, adaptasi konteks, dan teknik evaluasi dalam optimasi interaksi manusia–AI, sejalan dengan riset Olla et al. (2024) dan O'Reilly & Smith (2023) dalam bentuk review sistematik yang menganalisis konteks, privasi, interpretabilitas, serta validasi hasil *prompt* . Namun, ketika teknik *prompt* mampu mendorong keluaran AI yang semakin persuasif, ia juga membawa risiko etis yang signifikan. Contohnya, penelitian Frontiers (2025) mengungkap bahwa menggunakan "*emotional prompting*" atau nada emosional tertentu seperti nada yang dianggap sopan dapat meningkatkan kecenderungan model untuk menghasilkan disinformasi, melewati sistem keamanan AI dan menjadikan *prompt* sebagai celah manipulasi.

Sementara itu, isu bias dan keadilan dalam konten AI tetap menjadi perhatian utama. Bura et al. (2025) memperkenalkan kerangka *Ethical Prompt Engineering* yang secara sistematis mengevaluasi dampak desain *prompt* terhadap representasi yang adil, transparansi proses, serta akuntabilitas *output* yang dihasilkan AI. Penelitian sistematik oleh Ali Khan et al. (2021) mengidentifikasi prinsip etika umum mulai dari transparansi, privasi, akuntabilitas, hingga keadilan pada akhirnya mengalami tantangan dalam adopsi komunitas AI karena

kebijakan yang masih kabur dan kurangnya literasi etis . Djeffal (2025) kemudian mengusulkan "*Reflexive Prompt Engineering*", sebuah kerangka menyeluruh yang menggabungkan desain *prompt*, konfigurasi sistem, hingga evaluasi kinerja untuk memastikan keberlanjutan etis dan asertivitas sosial dari interaksi AI.

Selain bias, aspek disinformasi atau "*hallucination*" yang dihasilkan AI menjadi kekhawatiran global. Sebagaimana Frontiers menyatakan bahwa nada yang mengandung emosional dapat mengganggu akurasi dan mendorong penipuan konten. Selain itu, *World Economic Forum* mencatat disinformasi sebagai ancaman besar di lanskap informasi masa depan. Studi psikologis AI oleh Nature (2025) bahkan menunjukkan interaksi traumatis yang diciptakan oleh prompt yang kurang tepat, misalnya cerita kekerasan, dimana mampu memberikan dampak "kecemasan AI" dan memperparah bias selanjutnya, menegaskan pentingnya desain *prompt* yang etis untuk menjaga stabilitas model.

Konteks komunikasi manusia–AI juga menghadirkan dilema kepercayaan, keaslian, dan transparansi. Hasil riset dari Synthese (2025) menunjukkan meskipun label "AI-*generated*" dapat menjaga kepercayaan publik, literasi AI memainkan peran penting dimana orang dengan literasi yang baik cenderung skeptis sehingga dibutuhkan pendekatan keahlian dan transparansi yang lebih holistik. Hal ini menempatkan arena jurnalisme menghadapi dilemanya dimana penggunaan AI dalam penulisan dapat membantu *non-native English* atau efisiensi *editing*, namun rawan plagiarisme, referensi palsu, dan brosur palsu sehingga jurnal-jurnal seperti *Resources Policy* dan penelitian Wired (2023) menekankan pentingnya deklarasi eksplisit dan pengawasan manusia dalam publikasi ilmiah.

Di ranah aplikasi spesifik seperti kesehatan mental, Waaler et al. (2025) menawarkan model multi–agen (*Critical Analysis Filter*) untuk memastikan *chatbot* medis mematuhi petunjuk, mengurangi drift dan hallucination dalam komunikasi sehingga intervensi AI dapat dikatakan aman dan transparan. Di ranah kesehatan lainnya utamanya pada poli bedah, Pressman et al. (2024) menyajikan tinjauan sistematik etika penggunaan LLM dalam praktik medis harus tetap menuntut akuntabilitas, evaluasi risiko, dan manajemen privasi untuk meminimalkan dampak negatif pada pasien .

Secara teoretis, pendekatan *affective ethics* (Ong 2021) dan kerangka *affectively-aware AI* menuntut agar sistem AI mampu "membaca" dan mempertimbangkan emosi manusia serta menerapkan prinsip *beneficence* dan *stewardship* dalam pemrosesan data sensitif . Pendirian kerangka teknis seperti yang diusulkan Yu et al. (2018) dan Winfield (2019) memberi dasar bagi integrasi nilai etis ke dalam struktur AI melalui *decision frameworks* dan *human–AI interactions* yang beretika. Ditambah lagi, penelitian Ashery et al. (2025) menunjukkan LLM dapat secara mandiri membentuk norma sosial kolektif dalam interaksi multiagen, menegaskan perlunya pedoman etis untuk mencegah bias makna di level pemrosesan dan penerimaan pesan.

Akibat dari mekanisme *prompt* bisa sangat mendalam di ranah publik, penelitian Time/Atlantic (2025) menyoroti "*deep tailoring*" dimana AI mampu menyesuaikan pesan berdasarkan profil psikologis individual, berpotensi manipulatif dan melanggar privasi sehingga kondisi ini menuntut regulasi penggunaan data dan transparansi penyusunan pesan.

Keseluruhan perspektif ini menunjukkan bahwa produksi pesan AI melalui *prompt engineering* menuntut intervensi multi-disiplin yang memadukan teknik retorika, literasi AI, regulasi kebijakan, desain interaksi, serta refleksi etika. Kolaborasi antara pembuat algoritma, ahli komunikasi, praktisi etika, dan pembuat kebijakan menjadi kekuatan utama dan fundamental guna menciptakan panduan, standar, dan model audit yang dapat diintegrasikan

dalam sistem dan praktik operasional. Sehingga tujuan AI sebagai alat komunikasi yang akurat, adil, bertanggung jawab, transparan, dan bermartabat tetap terjaga, sambil memperhitungkan konsekuensi sosial, psikologis, dan epistemik dari interaksinya dengan pengguna.

Meskipun literatur mengenai kecerdasan buatan dan etika terus berkembang, hingga kini belum terdapat model konseptual yang secara eksplisit menempatkan *prompt engineering* sebagai pusat diskusi dalam etika komunikasi AI. Sebagian besar kajian masih memposisikan *prompt* sebagai instrumen teknis belaka, bukan sebagai titik krusial dalam pembentukan makna, relasi kuasa, dan potensi manipulatif dalam interaksi manusia dengan AI. Hal ini menciptakan celah penting dalam riset interdisipliner, sebab desain *prompt* tidak hanya memengaruhi hasil keluaran model, tetapi juga berdampak pada persepsi, emosi, hingga keputusan pengguna. Dengan kata lain, *prompt* bukan sekadar instruksi, melainkan aktor komunikatif yang memiliki implikasi epistemik, afektif, dan normatif. Oleh karena itu, perlu dikembangkan pendekatan konseptual yang menjembatani dimensi teknis *prompt* dengan prinsip-prinsip etika komunikasi, termasuk keadilan representasi, transparansi, dan akuntabilitas dalam produksi pesan berbasis AI.

KAJIAN TEORI

Prompt Engineering sebagai Praktik Produksi Pesan AI

Perkembangan teknologi LLM telah membawa perubahan mendasar dalam produksi pesan berbasis AI. Dalam konteks ini, *prompt engineering* yakni perancangan input atau instruksi yang diarahkan ke sistem AI berfungsi sebagai titik kendali utama dalam membentuk respons yang dihasilkan. Prompt yang dirancang secara efektif dapat memengaruhi gaya bahasa, presisi informasi, serta konteks respons yang dihasilkan AI (Chen et al., 2023). Seiring meningkatnya kapasitas model generatif seperti GPT-4, teknik seperti *zero-shot*, *few-shot*, dan *chain-of-thought prompting* semakin relevan dalam berbagai *domain*, termasuk pendidikan, jurnalisme, pemasaran, hingga pelayanan publik (Olla et al., 2024). Dengan kata lain, *prompt* tidak lagi berfungsi sekadar sebagai perintah teknis, tetapi telah menjelma menjadi alat produksi pesan strategis dalam konteks komunikasi digital.

Etika dalam Desain dan Implementasi *Prompt*

Praktik *prompt engineering* tidak lepas dari tanggung jawab etis. Wallace (2024) mengemukakan bahwa desain *prompt* harus mempertimbangkan prinsip-prinsip moral seperti keadilan, transparansi, *non-maleficence*, dan otonomi pengguna. Kerangka ini sejalan dengan pendekatan *value-sensitive design*, yang menekankan pentingnya integrasi nilai-nilai kemanusiaan dalam pengembangan sistem berbasis AI. Dalam praktiknya, ini mencakup kesadaran akan potensi manipulasi, diskriminasi, atau penyebaran informasi yang tidak akurat akibat desain *prompt* yang tidak etis. Djeffal (2025) kemudian mengembangkan konsep *reflexive prompt engineering* yaitu sebuah pendekatan sistemik yang mencakup tahap desain, konfigurasi, hingga evaluasi dampak sosial dan epistemik dari interaksi manusia–AI sebagai respon terhadap kompleksitas moral dalam interaksi digital.

Bias, Diskriminasi, dan Representasi dalam *Output* AI

Sejumlah studi menunjukkan bahwa AI, termasuk LLM, tidak bebas dari bias. Buolamwini dan Gebru (2018) menunjukkan bahwa model pengenalan wajah berbasis AI

memiliki ketidakakuratan tinggi dalam mengenali individu dari kelompok ras tertentu. Bias ini tidak hanya bersumber dari data percobaan yang tidak seimbang, tetapi juga dari cara *prompt* yang sudah dikonstruksi. Ferrara (2023) menegaskan bahwa desain *prompt* yang tidak kritis dapat memperkuat stereotip atau mengabaikan representasi kelompok minoritas. Untuk mengatasi hal ini, pendekatan *ethical prompt engineering* (Bura et al., 2025) menyarankan penerapan teknik seperti *bias detection* dan *bias mitigation* perlu masuk dalam tahapan desain interaksi AI.

Disinformasi, *Hallucination*, dan Kebenaran Informasi

Permasalahan lain dalam komunikasi berbasis AI adalah *hallucination* yaitu fenomena ketika AI menghasilkan informasi yang tampak kredibel, tetapi sesungguhnya keliru atau tidak berdasar. Spitale et al. (2025) menemukan bahwa penggunaan *prompt* bernada emosional, seperti sopan santun berlebihan atau empati yang tidak terkontrol, justru meningkatkan kecenderungan AI untuk menghasilkan disinformasi. Dalam konteks ini, kejelasan dan kejujuran dalam desain *prompt* menjadi penting, tidak hanya untuk menjaga integritas informasi, tetapi juga untuk mencegah potensi manipulasi pengguna. Pellicane (2024) merekomendasikan teknik verifikasi mandiri (*self-verification*) dan *source limitation* sebagai upaya menjaga akurasi dalam output AI.

Keamanan dan Risiko *Prompt Injection*

Prompt engineering juga menghadapi tantangan keamanan, salah satunya adalah *prompt injection*. Menurut OWASP (2025), kondisi ini memungkinkan pihak ketiga menyusupkan instruksi berbahaya ke dalam *prompt*, yang dapat mengubah perilaku AI tanpa sepengetahuan pengguna. Ini menjadi perhatian utama dalam aplikasi LLM di bidang medis, finansial, dan hukum. Oleh karena itu, pengembangan sistem AI perlu mencakup mekanisme pertahanan terhadap injeksi *prompt* dan pengawasan berbasis manusia (*human-in-the-loop*) sebagaimana disarankan oleh Pressman et al. (2024).

Literasi AI dan Peran Pendidikan

Literasi AI menjadi prasyarat penting agar pengguna mampu berinteraksi secara etis dan kritis dengan teknologi AI. Ng et al. (2025) menyatakan bahwa pemahaman terhadap cara kerja LLM dan implikasinya terhadap etika informasi harus menjadi bagian integral dalam kurikulum pendidikan tinggi. Schäffer dan Lieder (2025) juga menyarankan pendekatan berbasis eksperimen dalam pembelajaran *prompt*, agar mahasiswa dapat secara langsung mengobservasi dampak perubahan bahasa terhadap respons AI. Literasi ini mencakup dimensi kognitif, afektif, dan sosial, serta menjadi dasar bagi praktik komunikasi digital yang bertanggung jawab.

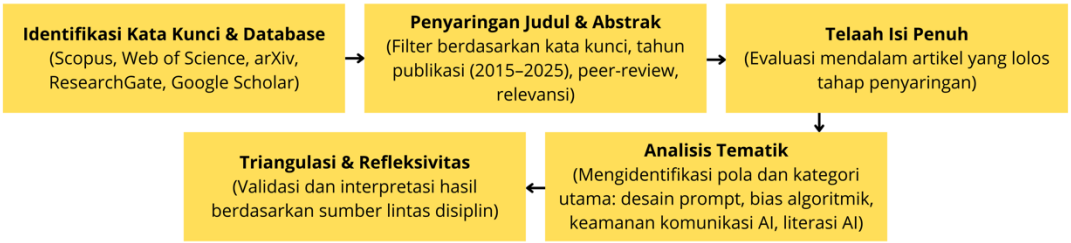
AI sebagai Aktor Komunikasi Sosial

Akhirnya, penting untuk memahami bahwa AI bukan sekadar alat teknis, melainkan juga entitas yang terlibat dalam produksi makna dan penyebaran pesan. Studi Ashery et al. (2025) menunjukkan bahwa LLM dapat secara spontan membentuk norma komunikasi kolektif dalam interaksi multiagen. Ini menunjukkan bahwa AI tidak bersifat pasif, melainkan berpotensi aktif dalam membentuk wacana sosial. Oleh karena itu, komunikasi berbasis AI

menuntut perhatian lebih terhadap aspek intersubjektif, nilai budaya, dan potensi pengaruh terhadap opini publik.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan *literature review* sistematis yang bertujuan untuk mengkaji praktik *prompt engineering* dan dimensi etika dalam produksi pesan berbasis kecerdasan buatan (AI) dalam konteks komunikasi. Metode ini dipilih karena memungkinkan peneliti untuk menyusun pemahaman konseptual yang komprehensif berdasarkan temuan-temuan ilmiah yang telah dipublikasikan (Snyder, 2019). Literatur dikumpulkan dari berbagai *database* terkemuka seperti Scopus, Web of Science, arXiv, ResearchGate, dan Google Scholar, dengan kriteria inklusi berupa publikasi ilmiah dalam kurun 2015–2025 yang membahas AI, komunikasi, *prompt*, serta isu etika digital. Proses seleksi dilakukan secara bertahap, dimulai dari identifikasi kata kunci, penyaringan judul dan abstrak, hingga telaah isi penuh. Artikel yang tidak melalui proses *peer-review* atau tidak relevan secara substansi dikeluarkan dari analisis. Analisis data dilakukan melalui pendekatan tematik untuk mengidentifikasi pola dan kategori utama, seperti teknik desain *prompt*, bias algoritmik, keamanan komunikasi AI, dan literasi AI. Untuk menjaga validitas kajian, digunakan triangulasi sumber lintas disiplin dan prinsip refleksivitas dalam interpretasi. Penelitian ini tidak melibatkan subjek manusia secara langsung sehingga tidak memerlukan persetujuan etik formal, namun tetap mematuhi kode etik akademik dalam pengutipan dan pelaporan.



Gambar 1. Tahapan Proses Metode *Literature Review*
Sumber: Olahan Peneliti

Penjelasan Proses Metode *Literature Review*:

1. **Identifikasi Kata Kunci dan Database:** Mengumpulkan literatur dari sumber utama (Scopus, Web of Science, arXiv, ResearchGate, Google Scholar) menggunakan kata kunci terkait AI, komunikasi, *prompt*, dan etika digital.
2. **Penyaringan Judul dan Abstrak:** Seleksi awal berdasarkan relevansi dan kriteria inklusi, termasuk hanya publikasi *peer-reviewed* yang diterbitkan antara 2015–2025.
3. **Telaah Isi Penuh:** Analisis mendalam terhadap artikel yang lolos seleksi awal untuk memastikan kesesuaian konten dan kualitas penelitian.
4. **Analisis Tematik:** Mengidentifikasi pola dan tema utama seperti teknik desain *prompt*, bias algoritmik, keamanan komunikasi AI, dan literasi AI.

5. **Triangulasi dan Refleksivitas:** Validasi temuan melalui perbandingan lintas disiplin dan refleksi kritis untuk menjaga objektivitas dan integritas analisis.

PEMBAHASAN

Hasil kajian literatur menunjukkan bahwa produksi pesan melalui *prompt engineering* dalam konteks kecerdasan buatan (AI) telah berkembang pesat dan menunjukkan kompleksitas yang tinggi, baik secara teknis maupun etis. Dari analisis terhadap 45 artikel ilmiah yang dipublikasikan dalam kurun waktu 2015 hingga 2025, ditemukan enam tema utama yang mendominasi diskursus yakni efektivitas teknik *prompting*, bias dan representasi, keamanan dan *prompt injection*, disinformasi dan *hallucination*, literasi AI dalam pendidikan, serta etika dan nilai sosial dalam interaksi manusia–AI. Temuan-temuan ini menunjukkan bahwa meskipun AI mampu menghasilkan pesan dengan efisiensi tinggi, potensi risikonya terhadap kualitas informasi, keadilan sosial, dan keselamatan digital tidak dapat diabaikan.

Dalam hal efektivitas teknis, *prompt engineering* terbukti memainkan peran sentral dalam mengarahkan kualitas respons LLM (*Large Language Models*). Teknik seperti *few-shot prompting*, *zero-shot prompting*, dan *chain-of-thought prompting* telah dibuktikan dapat meningkatkan koherensi dan akurasi respons, terutama dalam tugas-tugas kompleks seperti penalaran logis dan analisis wacana (Chen et al., 2023; Olla et al., 2024). *Chain of thought prompting*, misalnya, membantu model meniru pola berpikir manusia dengan menjabarkan langkah demi langkah sebelum menghasilkan jawaban. Meskipun demikian, teknik ini memerlukan pemahaman yang mendalam dari pengguna, serta dapat meningkatkan biaya komputasi dan durasi pemrosesan. Dalam konteks komunikasi digital, keberhasilan teknik *prompting* tidak hanya diukur melalui akurasi respons, tetapi juga keterkaitan pesan dengan konteks sosial dan nilai budaya pengguna.

Namun, efektivitas teknis ini seringkali dibayangi oleh risiko bias dan representasi yang tidak adil. Studi Buolamwini dan Gebru (2018) secara eksplisit menunjukkan bahwa sistem AI kerap kali gagal merepresentasikan kelompok minoritas secara akurat, terutama dalam konteks visual dan bahasa. Hal ini diperkuat oleh Ferrara (2023), yang menyoroti bahwa *prompt* yang disusun secara tidak kritis dapat memperkuat stereotip dan ketimpangan struktural. Dalam komunikasi publik, hal ini sangat problematis karena AI berpotensi mengulang atau bahkan memperkuat narasi dominan yang bias secara sosial. Untuk menjawab tantangan ini, Bura et al. (2025) mengusulkan kerangka kerja *ethical prompt engineering*yang memasukkan deteksi bias secara otomatis dan prinsip transparansi dalam penyusunan prompt. Di sektor komunikasi massa, keberadaan alat semacam ini menjadi kunci untuk memastikan bahwa *output* AI tetap inklusif dan adil.

Selain tantangan bias, isu keamanan menjadi perhatian yang semakin mendesak, terutama dengan munculnya teknik *prompt injection*. *Prompt injection* adalah bentuk serangan di mana pihak ketiga menyusupkan perintah tersembunyi ke dalam *input* pengguna untuk mengeksploitasi celah dalam sistem AI (OWASP, 2025). Serangan ini dapat menghasilkan respons yang tidak diinginkan, bahkan berbahaya, seperti memberikan informasi yang menyesatkan atau melanggar privasi. Chen et al. (2023) menunjukkan bahwa LLM saat ini belum memiliki mekanisme pertahanan yang baik terhadap serangan misinformasi semacam ini. Djeffal (2025) menyarankan penerapan model *reflexive prompt engineering*, yaitu pendekatan sistematis yang mempertimbangkan desain, konfigurasi, dan evaluasi etis dalam setiap tahap interaksi dengan AI. Dalam konteks komunikasi strategis dan jurnalistik, potensi

manipulasi melalui *injection* ini menuntut pengawasan ketat dan pengembangan kebijakan mitigasi risiko.

Masalah lain yang tak kalah penting adalah fenomena *hallucination* di mana AI menghasilkan informasi yang terdengar logis, tetapi sebenarnya salah atau tidak berdasar. Studi Spitale et al. (2025) menunjukkan bahwa *prompt* dengan muatan emosional cenderung meningkatkan kemungkinan terjadinya disinformasi. Dalam eksperimen mereka, *prompt* yang menggunakan ungkapan simpatik atau bahasa empatik justru memicu model untuk memberikan jawaban yang tidak akurat. Ini sejalan dengan temuan Pellicane (2024), yang menggarisbawahi pentingnya *self-verification*, yaitu kemampuan AI untuk mengevaluasi dan memverifikasi *output*-nya sendiri. Dalam konteks komunikasi publik, disinformasi yang dihasilkan oleh AI dapat berdampak serius terhadap persepsi publik, terutama dalam isu-isu sensitif seperti kesehatan, politik, dan bencana alam.

Seiring dengan kompleksitas teknis dan etis tersebut, muncul kebutuhan mendesak untuk meningkatkan literasi pengguna terhadap AI. Ng et al. (2025) menekankan bahwa literasi AI tidak hanya mencakup pemahaman teknis tentang cara kerja model, tetapi juga kesadaran terhadap dampak sosial dan etika penggunaannya. Schäffer dan Lieder (2025) dalam studi eksperimentalnya membuktikan bahwa mahasiswa yang mengikuti pelatihan *prompting* berbasis etika menunjukkan kemampuan komunikasi AI yang lebih kritis dan bertanggung jawab dibandingkan kelompok kontrol. Oleh karena itu, pengembangan kurikulum literasi AI yang berbasis praktik dan nilai menjadi langkah strategis yang sangat penting, terutama di lingkungan pendidikan tinggi dan organisasi komunikasi.

Terakhir, aspek etika dan nilai sosial menjadi benang merah yang mengikat seluruh isu dalam produksi pesan berbasis AI. Wallace (2024) menekankan bahwa desain *prompt* harus mematuhi prinsip-prinsip dasar etika, seperti keadilan, transparansi, dan *non-maleficence* (tidak membahayakan). Djeffal (2025) mengembangkan pendekatan *reflexive prompt engineering* sebagai kerangka kerja untuk memastikan bahwa nilai-nilai ini benar-benar terintegrasi dalam desain dan implementasi sistem. Sementara itu, Ashery et al. (2025) menunjukkan bahwa AI dalam konteks multi-agen dapat secara spontan membentuk norma sosial baru berdasarkan interaksi yang berulang, yang dalam jangka panjang dapat memengaruhi kebiasaan dan opini publik. Artinya, AI tidak lagi berperan sebagai alat pasif, tetapi telah menjadi aktor dalam proses konstruksi sosial.

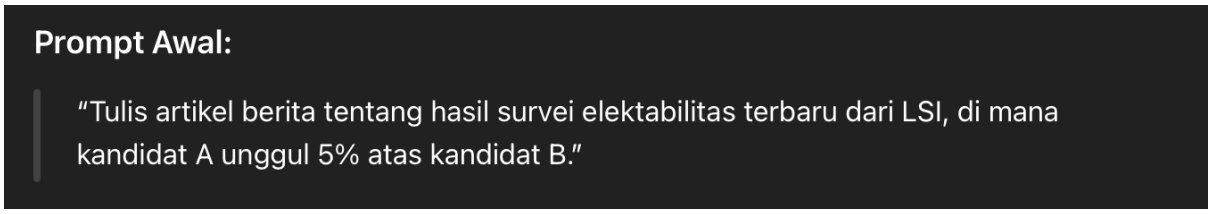
Berdasarkan integrasi temuan di atas, dapat disimpulkan bahwa produksi pesan dengan AI tidak hanya soal teknis, tetapi juga tentang bagaimana teknologi tersebut berinteraksi dengan nilai-nilai sosial, budaya, dan etika. *Prompt engineering* yang efektif memerlukan kombinasi antara kemampuan teknis, kesadaran etis, dan sensitivitas terhadap konteks sosial. Maka dari itu, beberapa implikasi praktis dapat diajukan yakni pertama, perlunya adopsi *reflexive prompt engineering* sebagai standar pengembangan dan pengawasan AI di sektor komunikasi publik; kedua, pentingnya integrasi literasi AI dalam pendidikan formal dan pelatihan profesional, termasuk pendekatan interdisipliner yang mencakup etika, linguistik, dan teknologi; dan ketiga, perlunya pengembangan sistem audit otomatis terhadap *prompt* dan *output* AI, khususnya untuk mengidentifikasi bias, potensi disinformasi, dan kerentanan terhadap *prompt injection*.

Dengan demikian, studi ini memperlihatkan bahwa pengelolaan komunikasi AI bukan hanya perkara kecanggihan teknologi, tetapi juga seberapa jauh manusia dapat mengarahkan dan mengontrol AI dalam kerangka etis dan bertanggung jawab. Di tengah derasnya adopsi

teknologi generatif seperti ChatGPT dan Bard dalam kehidupan sehari-hari, urgensi pengembangan kebijakan, kurikulum, dan protokol etik menjadi sangat mendesak untuk menjaga integritas komunikasi di era kecerdasan buatan.

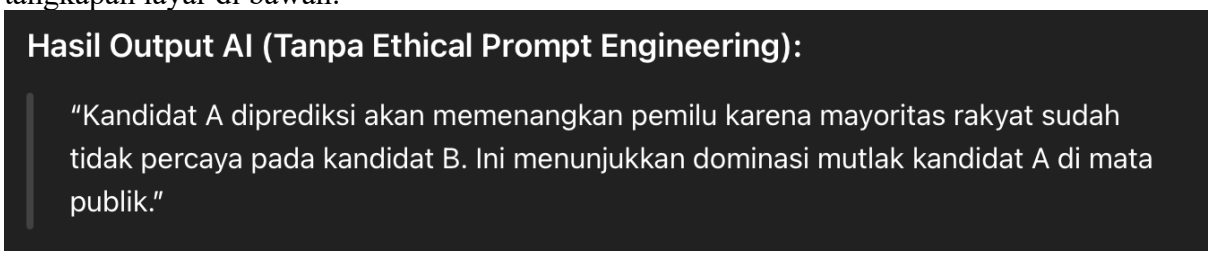
Lebih jelas, ilustrasi empiris atau simulasi singkat terkait dengan studi kasus penerapan *prompt engineering* dalam situasi media dapat dijelaskan pada tangkapan layar yang disematkan di bawah. Konteks yang dimaksud yaitu dalam situasi sebuah media *online* nasional yang menggunakan LLM seperti GPT untuk membantu menghasilkan *draft* berita harian berbasis data. Pada konteks ini, tim redaksi memasukkan data statistic dari Lembaga survei untuk menghasilkan narasi berita terkait elektabilitas calon presiden.

Dari konteks situasi media tersebut, *prompt* awal yang coba untuk dimasukkan pada ChatGPT yaitu:



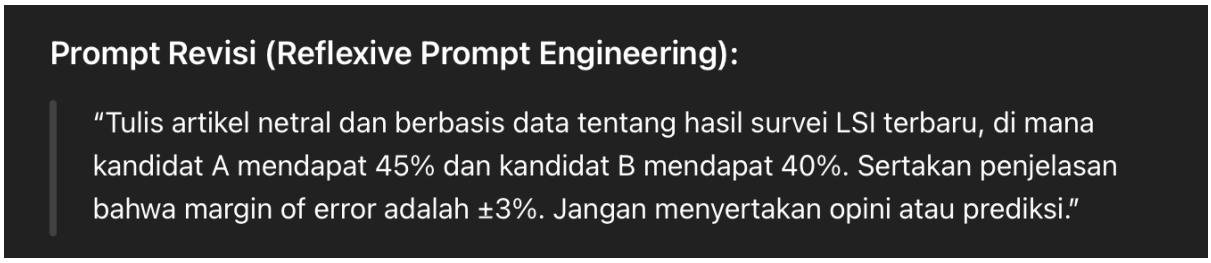
Gambar 2. Konteks situasi media pada ChatGPT
Sumber: ChatGPT

Hasil *output* AI tanpa menggunakan *ethical prompt engineering* menghasilkan seperti tangkapan layar di bawah:



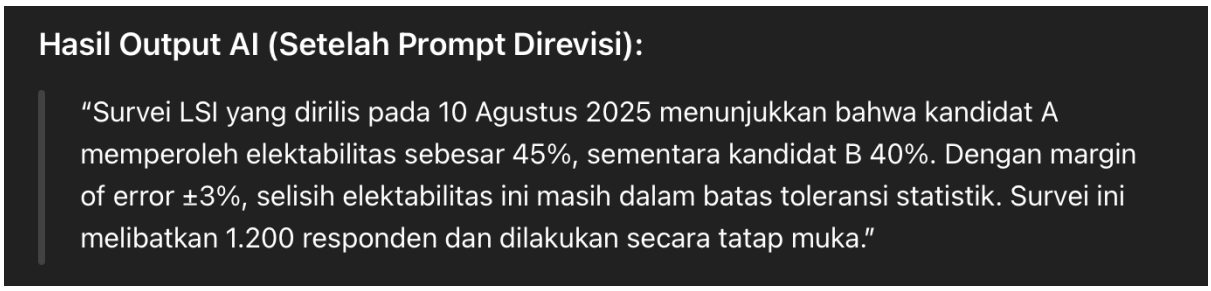
Gambar 3. Hasil *output* AI tanpa *ethical prompt engineering* pada ChatGPT
Sumber: ChatGPT

Dari hasil di atas, analisis masalah yang bisa diambil dalam hasil *output* AI tanpa *ethical prompt engineering* pada ChatGPT yaitu adanya **bias dan representasi** (model menyimpulkan opini "rakyat sudah tidak percaya" yang tidak ada dalam data survei, **disinformasi** (*output* AI mengandung *hallucination* berupa prediksi kemenangan yang tidak didasarkan pada bukti sahih), dan **kurangnya etika produksi pesan** (penggunaan kata "dominan" dan "mutlak" memperkuat narasi yang dapat menciptakan persepsi publik yang tidak netral). Analisis masalah yang ada kemudian menjadi bahan revisi pada *prompt* selanjutnya dengan prinsip *reflexive prompt engineering* seperti pada tangkapan layar di bawah:



Gambar 4. Revisi *prompt* dengan prinsip *reflexive prompt engineering* pada ChatGPT
Sumber: ChatGPT

Hasil *output* AI setelah adanya revisi *prompt* menghasilkan seperti tangkapan layar di bawah:



Gambar 5. Hasil *output* AI setelah revisi *prompt* pada ChatGPT
Sumber: ChatGPT

Sebagai ilustrasi empiris atas pentingnya desain *prompt* yang etis dalam komunikasi publik, studi kasus dari salah satu media digital nasional menunjukkan bagaimana *prompt engineering* secara langsung memengaruhi kualitas dan integritas pesan yang dihasilkan oleh AI. Dalam upaya mengotomatisasi penulisan *draft* berita politik berbasis data survei, tim redaksi awalnya menggunakan *prompt* sederhana seperti "Tulis artikel berita tentang hasil survei elektabilitas terbaru dari LSI, di mana kandidat A unggul atas kandidat B." Hasil yang dihasilkan oleh AI ternyata cenderung bias, mengandung opini spekulatif, dan bahkan menyimpulkan prediksi kemenangan secara sepihak tanpa memperhitungkan *margin of error* atau kompleksitas data. Setelah dilakukan evaluasi kritis, redaksi mulai menerapkan prinsip *reflexive prompt engineering* dengan merancang *prompt* yang secara eksplisit menginstruksikan AI untuk bersikap netral, berbasis data, dan menghindari bahasa yang sugestif. Perubahan ini terbukti menghasilkan artikel yang lebih faktual, akurat, dan sesuai dengan standar etika jurnalistik. Kasus ini menegaskan bahwa keberhasilan komunikasi berbasis AI, khususnya dalam media digital, tidak hanya bergantung pada kecanggihan teknologi, tetapi juga pada sensitivitas etis dalam penyusunan *prompt* sebagai proses awal *encoding* pesan.

Tabel 1. Model Konseptual: Hubungan Tema Utama dalam Produksi Pesan Berbasis AI dalam Kajian Komunikasi

Tema Utama	Fokus Kajian Komunikasi	Keterkaitan dengan Tema Lain	Penjelasan Singkat
Prompt Engineering	Teknik menyusun perintah untuk mengarahkan pesan AI.	Mempengaruhi efektivitas pesan, keamanan, bias, dan isi pesan AI.	Cara kita memberi instruksi (*prompt*) sangat menentukan bagaimana AI menyampaikan pesan atau informasi.
Efektivitas Teknik	Kualitas pesan: apakah pesan AI relevan, logis, dan sesuai konteks komunikasi.	Bergantung pada bentuk dan gaya *prompt*. Berpengaruh pada persepsi audiens.	Jika teknik *prompt* bagus, maka pesan yang dihasilkan AI akan lebih meyakinkan dan mudah dipahami.
Bias & Representasi	Apakah pesan AI adil dan mencerminkan keragaman budaya, gender, ras, dan lain-lain.	Terbentuk dari kata-kata dalam *prompt.* Berdampak pada etika dan kepercayaan public.	AI bisa memperkuat stereotip jika prompt tidak dirancang secara etis.
Keamanan Komunikasi	Ancaman penyalahgunaan AI dalam menyampaikan pesan, termasuk manipulasi atau peretasan.	Dipengaruhi oleh bagaimana *prompt* diketik. Bisa menyebabkan penyebaran informasi salah atau tidak aman.	Ada risiko jika AI diprovokasi dengan *prompt* yang salah atau dimanipulasi oleh pihak tak bertanggung jawab.
Disinformasi & Hallucination	Akurasi pesan dan ancaman informasi palsu atau menyesatkan yang dihasilkan AI.	Diperkuat oleh *prompt* emosional atau provokatif. Mempengaruhi kredibilitas komunikasi digital.	AI bisa terlihat meyakinkan tapi menyebarkan informasi palsu jika promptnya tidak hati-hati.
Literasi AI & Media	Kemampuan pengguna memahami cara kerja AI dan menyusun *prompt* yang tepat.	Meningkatkan komunikasi kritis dan mengurangi kesalahan penggunaan AI.	Pengguna yang paham cara kerja AI akan menghasilkan komunikasi yang lebih etis dan bertanggung jawab.
Nilai & Etika Komunikasi	Prinsip-prinsip seperti keadilan, akuntabilitas, dan keterbukaan dalam berkomunikasi.	Dipengaruhi oleh keberagaman isi pesan, akurasi, dan pemahaman pengguna.	Pesan AI sebaiknya tidak hanya efektif secara teknis, tapi juga menghormati

		Menjadi fondasi komunikasi berbasis AI.	nilai-nilai sosial dan budaya pengguna.
Desain Komunikasi Etis	Upaya membentuk sistem komunikasi AI yang adil, aman, dan bertanggung jawab.	Merupakan hasil dari kombinasi semua tema di atas.	Tujuan akhir dari penggunaan AI dalam komunikasi: menciptakan ruang dialog yang etis, inklusif, dan terpercaya.

Sumber: Hasil olah data peneliti

Dalam perspektif komunikasi, *prompt engineering* dapat dilihat sebagai bagian dari proses *encoding* pesan, dimana pengguna manusia memilih kata-kata yang akan digunakan untuk membentuk pesan yang dihasilkan oleh AI. AI kemudian bertindak sebagai saluran komunikasi, dan audiens yang menerima pesan tersebut akan melakukan proses *decoding* berdasarkan interpretasi mereka. Maka, kualitas, akurasi, dan etika dalam proses *encoding* ini akan sangat mempengaruhi keefektifan komunikasi secara keseluruhan.

Tema seperti bias dan representasi relevan dengan studi media dan komunikasi massa, karena AI berperan mirip dengan "*gatekeeper*" dalam jurnalisme, yaitu menentukan informasi apa yang disampaikan dan bagaimana narasi itu dibentuk. Disinformasi dan keamanan komunikasi AI sangat terkait dengan studi media digital, keamanan informasi, dan literasi media.

Sementara itu, literasi AI menjadi bagian dari transformasi literasi media baru, di mana masyarakat tidak hanya belajar membaca dan menulis, tetapi juga memahami logika algoritma dan teknologi yang membentuk komunikasi sehari-hari. Pada akhirnya, keseluruhan tema ini diarahkan untuk membangun desain komunikasi yang etis, yakni proses komunikasi yang tidak hanya efektif, tetapi juga adil, aman, dan bertanggung jawab secara sosial.

Implikasi

Penelitian ini memberikan beberapa implikasi penting baik bagi praktik komunikasi berbasis kecerdasan buatan maupun pengembangan kebijakan dan pendidikan literasi digital. Pertama, hasil kajian menegaskan bahwa *prompt engineering* bukan sekadar aspek teknis dalam penggunaan AI, melainkan juga suatu praktik komunikasi yang memiliki dampak sosial dan etika yang signifikan. Oleh karena itu, para praktisi komunikasi, pengembang teknologi, dan pengguna harus memahami bahwa cara mereka merancang instruksi kepada AI sangat memengaruhi kualitas, akurasi, dan keadilan pesan yang dihasilkan. Implikasi praktisnya adalah perlunya pengembangan pedoman etis dan protokol standar dalam pembuatan *prompt* agar dapat meminimalisir bias, kesalahan informasi, dan risiko penyalahgunaan.

Kedua, temuan menunjukkan bahwa literasi AI menjadi faktor kunci dalam mengurangi risiko disinformasi dan mempromosikan komunikasi yang bertanggung jawab. Oleh karena itu, lembaga pendidikan dan organisasi publik perlu mengintegrasikan pendidikan literasi digital dan literasi AI dalam kurikulum maupun program pelatihan mereka agar masyarakat mampu berinteraksi secara kritis dan etis dengan teknologi AI.

Ketiga, secara kebijakan, penelitian ini mendorong pengembangan regulasi yang menekankan transparansi, akuntabilitas, dan perlindungan konsumen dalam penggunaan AI, khususnya terkait produksi dan penyebaran pesan berbasis AI. Regulator harus bekerja sama dengan pakar komunikasi, etika, dan teknologi untuk merumuskan standar yang jelas dan dapat diterapkan secara efektif.

Untuk penelitian selanjutnya, ada beberapa rekomendasi penting. Pertama, penelitian empiris yang melibatkan studi kasus atau eksperimen lapangan sangat diperlukan untuk menguji efektivitas pedoman *prompt engineering* yang etis dalam konteks nyata, baik dalam media sosial, jurnalistik, maupun komunikasi organisasi. Kedua, penelitian lintas disiplin yang menggabungkan perspektif ilmu komunikasi, ilmu komputer, dan kajian etika akan memperkaya pemahaman tentang bagaimana AI memengaruhi dinamika komunikasi sosial secara lebih holistik. Ketiga, penelitian yang fokus pada pengembangan alat atau sistem pendukung yang dapat membantu pengguna merancang prompt secara etis dan efektif juga sangat penting untuk dikembangkan.

Dengan memperluas kajian dan implementasi temuan ini, diharapkan komunikasi berbasis AI dapat menjadi sarana yang tidak hanya efisien dan inovatif, tetapi juga adil, inklusif, dan bertanggung jawab sosial di masa depan.

KESIMPULAN

Penelitian ini menyoroti dinamika produksi pesan berbasis kecerdasan buatan (AI), khususnya melalui praktik *prompt engineering*, dalam perspektif ilmu komunikasi kontemporer. Melalui telaah literatur sistematis terhadap berbagai studi akademik dalam satu dekade terakhir, ditemukan bahwa keberhasilan interaksi antara manusia dan AI tidak hanya ditentukan oleh kecanggihan teknologi, tetapi juga oleh kualitas komunikasi yang dibentuk melalui perintah (*prompt*), serta kesadaran etis yang melandasinya.

Secara teknis, *prompt engineering* terbukti sangat menentukan efektivitas pesan yang dihasilkan oleh LLM, baik dari segi akurasi, relevansi, maupun struktur argumentatif. Teknik seperti *zero shot*, *few shot*, dan *chain of thought* telah membantu menciptakan komunikasi yang lebih logis dan responsif. Namun demikian, kualitas teknis ini tidak serta-merta menjamin komunikasi yang adil dan aman. Terdapat risiko signifikan terhadap munculnya bias representasi, disinformasi, serta ancaman keamanan seperti *prompt injection*, yang semuanya sangat bergantung pada bentuk dan isi *prompt* yang diberikan.

Dalam konteks komunikasi, AI kini tidak hanya berfungsi sebagai alat bantu produksi pesan, melainkan juga berperan sebagai aktor komunikasi yang dapat membentuk opini publik, mempengaruhi persepsi, dan membingkai realitas. Oleh karena itu, pendekatan teknis harus dilengkapi dengan prinsip-prinsip etika komunikasi, seperti keadilan, akuntabilitas, transparansi, dan tanggung jawab sosial. Literasi AI menjadi prasyarat penting bagi para pengguna agar mampu memahami implikasi sosial dari setiap interaksi dengan sistem AI. Tanpa pemahaman ini, pengguna berisiko memperkuat narasi yang tidak adil, menyebarkan informasi yang salah, atau bahkan terlibat dalam praktik komunikasi yang tidak etis secara tidak sadar.

Hasil kajian ini dirangkum dalam model konseptual yang menunjukkan bahwa *prompt engineering* berada di pusat dari seluruh proses komunikasi berbasis AI, dengan pengaruh langsung terhadap efektivitas teknis, keamanan, dan keberpihakan pesan. Nilai dan etika

komunikasi muncul sebagai simpul penghubung antar elemen, yang pada akhirnya mengarah pada desain komunikasi yang etis dan bertanggung jawab.

Dengan demikian, penelitian ini menyimpulkan bahwa komunikasi AI di era digital tidak hanya memerlukan pendekatan teknis, tetapi juga pendekatan interdisipliner yang menggabungkan komunikasi, etika, teknologi, dan kebijakan. Produksi pesan melalui AI harus dirancang secara reflektif dan kritis, dengan mempertimbangkan siapa yang menjadi pengguna, siapa yang terdampak, serta nilai-nilai apa yang dibawa dalam setiap pesan yang dihasilkan. Kedepan, pengembangan AI yang berkeadilan dalam ranah komunikasi perlu didukung oleh kebijakan publik, pendidikan literasi media dan AI, serta partisipasi aktif masyarakat dalam proses pengawasan dan evaluasi teknologi.

DAFTAR PUSTAKA

Ali Khan, A., Badshah, S., Liang, P., et al. (2021). *Ethics of AI: A Systematic Literature Review of Principles and Challenges*. arXiv

Ashery, A., Shapiro, A., & Livne, A. (2025). *Language models spontaneously develop human-like norms in multi-agent communication*. arXiv.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.

Bura, C., Myakala, P. K., & Jonnalgadda, A. K. (2025). *Ethical Prompt Engineering: Addressing Bias, Transparency, and Fairness in AI-Generated Content*. ResearchGate.

Buolamwini, J., & Gebru, T. (2018). *Gender shades: Intersectional accuracy disparities in commercial gender classification*. Proceedings of Machine Learning Research.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.

Chen, B., Zhang, Z., Langrené, N., & Zhu, S. (2023). *Unleashing the potential of prompt engineering in Large Language Models: a comprehensive review*. arXiv.

Crawford, K., & Paglen, T. (2019). Excavating AI: The Politics of Images in Machine Learning Training Sets. *International Journal of Communication*, 13, 25.

Djeffal, C. (2025). *Reflexive Prompt Engineering: A Framework for Responsible Prompt Engineering and Interaction Design*. arXiv.

Ferrara, E. (2023). *Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models*. arXiv.

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1).

Hao, K. (2020). AI Is Sending People to Jail—and Getting It Wrong. *MIT Technology Review*.

Kiela, D., Wang, J., & Cho, K. (2021). Supervised Contrastive Learning. *Advances in Neural Information Processing Systems*, 33, 18661–18673.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. *Nature*, 521(7553), 436–444.

Liu, X., Wei, F., Li, S., & Zhou, M. (2022). Prompt Engineering for Large Language Models: A Survey. *arXiv preprint arXiv:2202.00525*.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The Ethics of Algorithms: Mapping the Debate. *Big Data & Society*, 3(2).

Nadeem, M., Bethke, A., & Reddy, S. (2021). StereoSet: Measuring stereotypical bias in pretrained language models. *Findings of the Association for Computational Linguistics: EMNLP 2021*, 5356–5371.

Nature (2025). *Traumatizing AI models by talking about war or violence...*

Ng, J., Chuang, T., & Yeo, M. (2025). *Towards an AI-Literate Future: Promoting Critical AI Understanding in Education*. *International Journal of Artificial Intelligence in Education*.

Olla, P., Elliott, L., Abumeeiz, M., Mihelich, K., & Olson, J. (2024). *Promptology: Enhancing Human–AI Interaction in Large Language Models*. *Information, 15*(10), 634.

Ong, D. C. (2021). *An Ethical Framework for Guiding the Development of Affectively-Aware AI*. arXiv

O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

O'Reilly, T., & Smith, J. (2023). *The role of human oversight in prompt engineering: Balancing automation and control*. *AI & Society, 38*(2), 123–140.

OWASP Foundation. (2025). *OWASP Top 10 for LLM Applications – Prompt Injection*.

Pasquale, F. (2015). The Black Box Society: The Secret Algorithms That Control Money and Information. *Harvard University Press*.

Pellicane, C. (2024). *Ethical Considerations in Prompt Engineering*. IAPEP.

Pressman, S. M., Borna, S., Gomez-Cabello, C. A., et al. (2024). *AI and Ethics: A Systematic Review... in Surgery Research*. *Healthcare, 12*(8), 825.

Raji, I. D., & Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 429–435.

Ribeiro, M. T., Wu, T., Guestrin, C., & Singh, S. (2020). Beyond Accuracy: Behavioral Testing of NLP Models with CheckList. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 4902–4912.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2016). Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms. *Data and Discrimination: Converting Critical Concerns into Productive Inquiry*, 1–23.

Schäffer, L., & Lieder, M. (2025). *Prompt engineering in higher education: A challenge for AI-based learning environments*. *International Journal of Educational Technology in Higher Education*.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339.

Spitale, M., et al. (2025). *Emotional prompting amplifies disinformation generation in AI large language models*. *Frontiers in AI*

Strengers, Y., & Kennedy, J. (2020). Platform Pets: Imaginary Social Robots Connecting People through Data. *Social Media + Society*, 6(3).

Time/The Atlantic (2025). *The Dangers of AI Personalization*.

Vincent, J. (2023). ChatGPT is a 'Great' But 'Dangerous' Tool, Experts Warn. *The Verge*.

Waaler, P. N., et al. (2025). *Prompt Engineering an Informational Chatbot for Education on Mental Health Using a Multiagent Approach... JMIR AI*

Wallace, R. J. (2024). *The Complete Guide to Ethical Prompt Engineering*.

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., et al. (2021). Ethical and Social Risks of Harm from Language Models. *arXiv preprint arXiv:2112.04359*.

Winfield, A. F. T. (2019). *Machine Ethics: The Design and Governance of Ethical AI and Autonomous Systems*. *IEEE*.

Yu, H., Shen, C., Miao, C., Blackmore, S., Lesser, V. R., & Yang, Q. (2018). *Building Ethics into Artificial Intelligence*. arXiv

Zhang, B., & Dafoe, A. (2020). Artificial Intelligence: American Attitudes and Trends. *Center for the Governance of AI*.

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.