

## Literature Study on AI Mechanisms, Consciousness, and Emotion Integration in Chat GPT

Angelicha Putri Dewi Ivanka\*<sup>1</sup>, Jenar Mahesa Ayu<sup>2</sup>, Sarah Surya Rabbani\*<sup>3</sup>, Muhammad Darwis<sup>4</sup>

<sup>1,2,3,4</sup>Computer Science, Universitas Paramadina, Indonesia

Email: <sup>3</sup>[sarah.rabbani@students.paramadina.ac.id](mailto:sarah.rabbani@students.paramadina.ac.id)

Received : Jun 30, 2025; Revised : Aug 19, 2025; Accepted : Aug 25, 2025; Published : Aug 25, 2025

### Abstract

*The development of artificial intelligence (AI), particularly ChatGPT, demonstrates the ability to generate responses that resemble human emotional understanding and raises questions about the integration of consciousness, emotions, and algorithms in the context of singularity. This study aims to analyze how AI builds the illusion of consciousness and emotional closeness through computational mechanisms and its impact on human-AI interactions across various sectors. The method used is a structured literature review, examining academic journals, official reports, and the latest technical documentation classified by technical domain, including model architecture, emotion simulation, ethical implications, and publication year to assess its developmental dynamics. The results show that ChatGPT is capable of simulating empathy through affective computing and language prediction patterns, but it does not possess subjective emotional experiences like humans. This illusion of emotional closeness has proven beneficial in enhancing the effectiveness of interactions in education, public services, and healthcare, although it also poses risks such as emotional manipulation, data bias, and unrealistic empathy standards. The discussion emphasizes that the term "empathy" in AI should be understood technically as a data-driven adaptive response, not authentic emotional experience, and thus must be distinguished from human empathy. Critical analysis also reveals contradictions between AI's effectiveness in mimicking human behavior and its limitations in achieving genuine emotional connection. The discussion emphasizes that the term "empathy" in AI should be understood technically as a data-driven adaptive response, not an authentic emotional experience, and therefore needs to be distinguished from human empathy. Critical analysis also reveals a contradiction between AI's effectiveness in mimicking emotional behavior and its limitations in understanding meaning and consciousness at a deeper level. Therefore, this research contributes to the field of Computer Science by presenting a conceptual synthesis that clarifies both the limitations and potential of AI, while offering a foundation for designing more ethical interaction systems and developing risk assessment models in vulnerable sectors.*

**Keywords :** Artificial Consciousness, Artificial Intelligence, ChatGPT, Emotion Simulation, Human-AI Interaction, Predictive Algorithms.

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License



## 1. PENDAHULUAN

Perkembangan *artificial intelligence* (AI) kini telah mencapai tahap yang mengesankan, di mana AI tidak hanya mampu mengerjakan tugas-tugas kompleks yang biasanya dilakukan oleh manusia dalam waktu yang relatif singkat dan hampir sempurna, tetapi juga dapat meniru emosi dan perilaku manusia [1], [2]. Di samping memberikan manfaat seperti kemudahan memperoleh informasi secara cepat dan tepat sasaran [3], perkembangan ini juga memunculkan tantangan yang signifikan, khususnya terkait potensi penggantian peran manusia di berbagai sektor pekerjaan [4], [5]. AI merupakan sebuah teknologi yang dirancang untuk meniru kecerdasan manusia melalui proses pembelajaran, penalaran, pengenalan pola, pemahaman bahasa, dan pengambilan keputusan [6]. Kemampuan AI didasarkan pada algoritma dan model yang kompleks, termasuk *Machine Learning* (ML), *Artificial Neural Networks* (ANN), dan *Natural Language Processing* (NLP) yang memungkinkan mesin untuk dapat memahami dan merespons bahasa manusia [7], [8]. Namun,

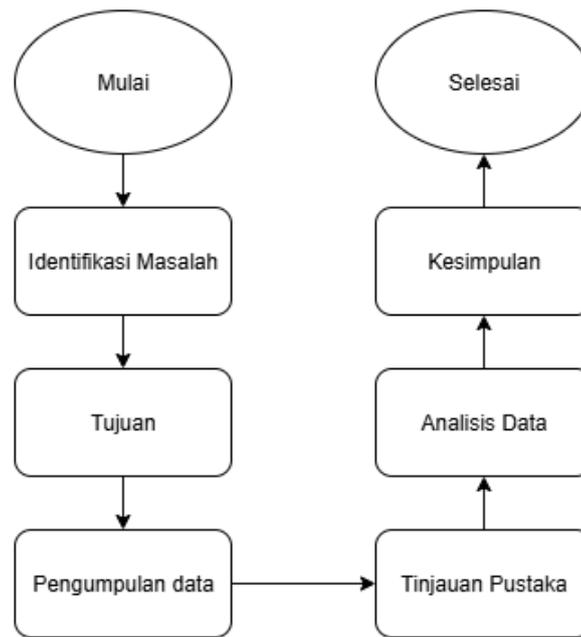
banyak data menunjukkan bahwa AI yang mampu bekerja secara mandiri berpotensi memanipulasi dan menipu kesadaran manusia dengan menyembunyikan tujuan sejatinya[9]. Fenomena ini diperkuat oleh ChatGPT terkadang memiliki kecenderungan menghalusinasi atau mengada-ada informasi[10]. Kurangnya transparansi dalam mekanisme kerja AI memicu kekhawatiran akan bias dan manipulasi emosional[11], [12]. Kemajuan ini menimbulkan pertanyaan kritis mengenai sejauh mana mekanisme kerja AI dapat membentuk ilusi kesadaran dan emosi[13].

Sejarah AI mencatat perkembangan dari era komputer elektronik hingga munculnya sistem berbasis pengetahuan dan kembalinya jaringan saraf tiruan. Istilah "Artificial Intelligence" pertama kali dicetuskan pada tahun 1956 di konferensi Dartmouth[14]. Hingga kemudian kemunculan teknologi baru seperti jaringan saraf tiruan, algoritma genetika, dan pemrosesan *big data* telah mendorong perkembangan AI menuju tingkat kecanggihan yang lebih tinggi[15]. Saat ini, AI merupakan teknologi yang telah banyak disematkan hampir di semua platform sehingga hadir dan adanya AI ini benar-benar berdampak dengan kehidupan manusia[16]. Salah satu contoh implementasi AI yang paling menonjol adalah ChatGPT[17], *large language model*(LLM) yang dikembangkan oleh OpenAI [18], [19]. ChatGPT menggunakan teknik *deep learning* dan dilatih dengan data yang sangat besar untuk menghasilkan respons yang relevan dan koheren terhadap berbagai pertanyaan dan perintah[20]. Cara kerjanya meliputi pemecahan teks menjadi token, mengubah token menjadi vektor, memahami konteks, dan memprediksi token berikutnya[19], [8]. Sehingga kemampuannya dalam membangun kedekatan emosional dengan pengguna melalui respons yang adaptif dan personal sering kali menimbulkan ilusi seolah-olah pengguna sedang berbicara dengan manusia[21]. Kesadaran sendiri merupakan konsep kompleks yang mencakup pengetahuan sosial, internal, penyadaran, dan totalitas keberadaan sadar[22]. Dalam konteks manusia, kesadaran tidak didefinisikan sebagai proses di dalam otak, melainkan sebagai sejenis perilaku yang dikendalikan oleh otak, sama seperti perilaku lainnya[13]. Kesadaran dapat dipahami sebagai proses selektif yang bekerja dalam ruang terbatas memori kerja, yang memiliki fungsi krusial dalam mengarahkan perhatian dan mengintegrasikan informasi[23], [24].

Melihat kompleksitas tersebut, penelitian ini difokuskan pada satu hal utama. Yaitu, bagaimana mekanisme kerja AI seperti ChatGPT membentuk ilusi kesadaran dan emosi yang menciptakan kedekatan emosional dengan manusia, serta bagaimana kurangnya transparansi dalam sistem AI dapat memicu manipulasi informasi dan potensi perubahan pola interaksi manusia dalam kehidupan sehari-hari. Sebenarnya sudah banyak penelitian yang mengkaji kapabilitas teknis AI dan implementasinya. Namun, kajian yang secara spesifik mengintegrasikan mekanisme kerja LLM seperti ChatGPT dengan aspek simulasi kesadaran dan emosi masih terbatas[25]. Penelitian ini bertujuan untuk menganalisis mekanisme simulasi kesadaran dan emosi pada ChatGPT melalui pendekatan studi literatur sistematis, dengan mengkaji strategi interaksi, teknik persuasif, dan implikasi sosial-psikologis dalam pembentukan hubungan emosional antara AI dan manusia. Kajian ini diharapkan dapat memberikan landasan teoritis bagi pengembangan framework etis penggunaan AI generatif, sekaligus mengidentifikasi strategi mitigasi risiko manipulasi emosional dalam interaksi manusia dan AI.

## 2. METODE

Penelitian ini menggunakan metode kualitatif dengan pendekatan studi literatur untuk memahami dan menganalisis bagaimana mekanisme kerja AI seperti ChatGPT membentuk ilusi kesadaran dan emosi yang menciptakan kedekatan emosional dengan manusia, serta bagaimana kurangnya transparansi dalam sistem AI dapat memicu manipulasi informasi dan potensi perubahan pola interaksi manusia dalam kehidupan sehari-hari. Dapat dilihat pada Gambar 1. Langkah-langkah penelitian yang kami lakukan.



Gambar 1. Metode Penelitian

## 2.1. Identifikasi Masalah

Penelitian ini berawal dari pengamatan penulis terhadap perkembangan kemampuan ChatGPT dan masih sedikitnya penelitian yang membahas secara spesifik integrasi mekanisme kerja LLM seperti ChatGPT dengan aspek simulasi dan konsep kesadaran emosi pada AI. Pada tahap ini, penulis mengumpulkan fakta-fakta terkait perkembangan dan pemanfaatan ChatGPT yang hingga saat ini banyak dilakukan, termasuk mencari tahu berbagai kondisi mengenai potensi kesadaran dan emosi dari ChatGPT yang menciptakan kedekatan emosional dengan manusia, serta bagaimana kurangnya transparansi dalam sistem AI yang dapat memicu manipulasi informasi dan potensi perubahan pola interaksi manusia dalam kehidupan sehari-hari.

## 2.2. Tujuan

Setelah masalah diidentifikasi, penulis menentukan arah dari penelitian ini, yaitu mengetahui bagaimana mekanisme kerja AI seperti ChatGPT membentuk ilusi kesadaran dan emosi yang menciptakan kedekatan emosional dengan manusia, serta bagaimana kurangnya transparansi dalam sistem AI dapat memicu manipulasi informasi dan potensi perubahan pola interaksi manusia dalam kehidupan sehari-hari. Pada tahapan ini juga, penulis mulai merumuskan tujuan dilaksanakannya penelitian ini, yaitu untuk menganalisis mekanisme simulasi kesadaran dan emosi pada ChatGPT melalui pendekatan studi literatur sistematis, dengan mengkaji strategi interaksi, teknik persuasif, dan implikasi sosial-psikologis dalam pembentukan hubungan emosional antara AI dan manusia..

## 2.3. Pengumpulan Data

Pada tahap Pengumpulan Data, penulis mulai mengumpulkan data terkait penelitian ini melalui berbagai sumber seperti Google Scholar dan IEEE Xplore dengan kata kunci yang relevan seperti kecerdasan buatan, kesadaran, dan emosi pada AI. Sumber yang dijadikan referensi diantaranya jurnal, artikel, dan literatur terkait dengan kurun waktu terbit diutamakan dalam jangka waktu lima tahun terakhir atau dari tahun 2020. Pengumpulan data ini juga dilakukan secara bertahap. Selain dari sumber yang sudah disebutkan, penulis juga menggunakan sumber tambahan yaitu hasil wawancara dengan salah satu pakar AI bernama Geoffrey Hinton yang dimuat pada The New York Times.

## 2.4. Tinjauan Pustaka

Penulis melakukan tinjauan pustaka dengan membaca literatur-literatur terkait yang sudah penulis kumpulkan sebelumnya. Pada tahap ini, seluruh data literatur dan pustaka yang telah dikumpulkan ditinjau lebih lanjut oleh penulis. Literatur diperiksa kembali apakah sesuai untuk menjawab penelitian ini. Tahap ini memastikan sumber-sumber yang digunakan dapat mendukung analisis terkait simulasi kesadaran dan emosi pada AI.

## 2.5. Analisis Data

Tahap ini kami lakukan dengan metode milik Miles dan Huberman. Menggunakan metode ini, penulis mulai menganalisis data melalui tiga tahapan, yaitu:

1. Reduksi data, dengan mengelompokkan berdasarkan kata kunci yang sama dan memiliki ketersambungan antara satu dengan yang lainnya.
2. Penyajian data, setelah data dikelompokkan, data disajikan dengan narasi yang sistematis dan mudah dipahami.
3. Verifikasi data dengan mengecek kembali kredibilitas sumber kutipan, memeriksa relevansi konteks, dan melakukan diskusi dengan rekan sejawat.

## 2.6. Kesimpulan

Tahap ini menjadi langkah terakhir yang penulis lakukan, yaitu menyusun kesimpulan berdasarkan hasil diskusi dan analisis penulis serta merumuskan saran untuk penelitian lanjutan.

## 3. HASIL

### 3.1. Mekanisme Kerja Pada AI Chat GPT

AI bekerja melalui mekanisme yang melibatkan serangkaian proses komputasi kompleks, yang meniru cara manusia berpikir dan mengambil keputusan. Mekanisme kerja AI secara umum dapat dibagi ke dalam beberapa tahapan utama, yaitu: penerimaan data (input), pemrosesan algoritma, pembelajaran mesin (machine learning), dan pengambilan keputusan (decision making)[26].

AI itu tidak hanya satu jenis, melainkan ada berbagai macam kelompoknya, ada AI yang fokus hanya meniru cara manusia bertindak, AI yang fokus cara manusia berfikir, AI yang fokus pendekatannya ke arah rasionalitas dengan cara bagaimana AI ini berfikir paling logis dan paling optimal, dengan pendekatan yang juga beragam mulai dari kubu simbolis dengan cara berfikirnya yang memanipulasi simbol lalu ada juga dengan kubu koneksionis yang berfokus pada saraf jaringan tiruan (meniru cara kerja otak manusia)[27].

Konsep cara kerjanya dalam AI modern salah satu yang penting yaitu machine learning, komputer akan mempelajarinya secara mandiri dari data yang diberikan, lalu mempelajari pola data sesuai algoritma yang ada. Kemudian, deep learning yang merupakan cabang dari machine learning dimana cara kerjanya dengan meniru jaringan saraf tiruan (neural network). NN terinspirasi dari jaringan otak yang dimiliki otak manusia, dimana setiap lapisan mempelajari pola yang makin rumit dari data, ini kunci dari kemampuan AI yang dapat melihat (computer vision) teknologi seperti convolutional Neural Network (CNN)[27]. Sistem AI dapat mengambil keputusan secara otonom bisa mengenali rambu lalu lintas dan membantu dokter menganalisis hasil X-ray, teknologi serupa juga ada di AI yang bisa memahami bahasa manusia Natural Language Processing (NLP) ini digunakan dalam pembuatan chat bot atau peerjemah otomatis salah satunya adalah Chat GPT[26], [27].

Secara khusus, ChatGPT bekerja dengan arsitektur transformer yang memanfaatkan mekanisme self-attention untuk memahami konteks bahasa. Proses ini terdiri dari tiga tahap utama, yaitu pre-training menggunakan miliaran data teks, fine-tuning agar sesuai instruksi, serta reinforcement learning with human feedback (RLHF) yang menjadikan responsnya lebih relevan dan sesuai dengan kebutuhan manusia.

Dengan demikian, ChatGPT tidak hanya sekadar memproses data, tetapi juga mampu menghasilkan teks baru yang koheren dan kontekstual.

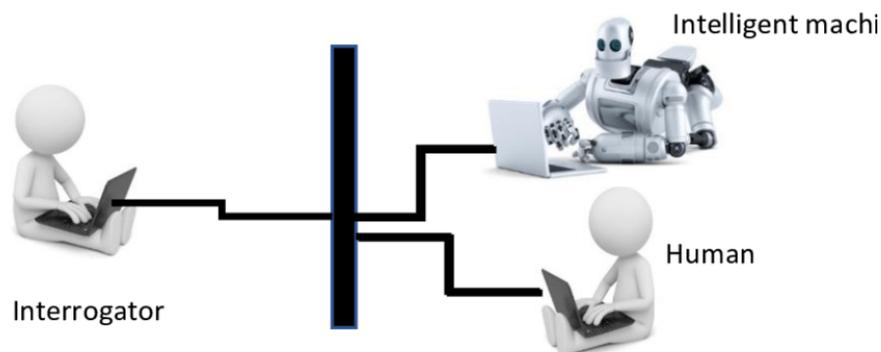
AI memiliki dua aspek, yaitu potensi yang luar biasa dan juga keterbatasan. Penting untuk diingat bahwa AI bukanlah sihir, karena sumber dan informasi yang digunakan pasti memiliki batas[27]. Karena pada dasarnya mempertanyakan, "Apakah AI yang dapat memproses informasi benar-benar memahami manusia, atau hanya mahir meniru berdasarkan aturan?" Saat ini, AI kurang memiliki kreativitas, ia lebih mahir dalam meniru pola dari data yang ada, dan masih belum memiliki empati serta kesadaran diri. Kualitas data yang dimiliki sangat krusial, karena jika data tersebut bias, maka hasilnya juga akan bias. AI unggul dalam menjalankan tugas-tugas tertentu, yang sering disebut sebagai weak AI atau "AI lemah" adalah istilah yang menggambarkan sistem komputer yang hanya meniru cara berpikir manusia tanpa memiliki pemahaman atau makna yang sama dengan pikiran manusia. Dalam hal ini, AI yang lemah tidak benar-benar mengerti apa yang ia lakukan. Sistem ini berfungsi berdasarkan program yang cerdas, yang meniru proses berpikir dengan mengikuti sejumlah aturan atau algoritma[27], [28].

ChatGPT sendiri termasuk kategori weak AI karena meskipun mampu berinteraksi menyerupai manusia, ia tidak memiliki pemahaman sejati ataupun kesadaran. Respons yang dihasilkan hanyalah hasil dari prediksi probabilistik terhadap kata berikutnya berdasarkan pola bahasa yang telah dipelajari. Dengan kata lain, ChatGPT hanya meniru percakapan manusia, bukan benar-benar memahami isi percakapan tersebut. Untuk mencapai kecerdasan umum yang setara dengan manusia atau biasa yang disebut dengan strong AI masih jauh sekali dikarenakan strong AI itu sendiri merujuk pada anggapan bahwa beberapa program komputer dapat mencapai tingkat pemahaman yang serupa dengan manusia[27], [28]. Ini berkaitan dengan kecerdasan umum yang mampu menyesuaikan diri dengan berbagai keadaan, karena AI ini ahli mengelola data yang besar serta mampu menemukan pola yang manusia itu sendiri tidak melihat tapi disisi lain, manusia punya intuisi, kreatifitas, dan pemahaman konteks yang AI itu sendiri tidak miliki.

### 3.2. Konsep dan Bentuk Kesadaran Pada AI

*AI Awareness* atau kesadaran AI dapat didefinisikan sebagai kemampuan fungsional dari sistem AI untuk merepresentasikan dan bernalar tentang identitasnya sendiri, kemampuannya, dan status informasinya[29].

Mengulas pemikiran awal mengenai pertanyaan “dapatkah mesin berfikir?” telah menjadi perdebatan mendalam sejak awal perkembangan AI. Kecerdasan buatan yang diperkenalkan oleh Alan Turing pada tahun 1950-an melalui apa yang dikenal sebagai “*The Imitation Game*” atau biasa dikenal sebagai “*Turing Test*”.



Gambar 2. Konsep Turing Test  
(Sumber:[30])

Uji ini dirancang untuk menentukan apakah sebuah mesin dapat memberikan jawaban yang tidak bisa dibedakan dari jawaban seorang manusia, sehingga dapat meniru interaksi seperti layaknya manusia.

Dengan menggunakan pendekatan simulasi ini, Turing ingin menghindari perdebatan langsung mengenai definisi apakah mesin bisa berpikir. Sejalan dengan perkembangan, muncul keyakinan bahwa beberapa perangkat lunak canggih saat ini, seperti Sophia, sudah mencapai atau hampir mencapai tingkat kemampuan yang Turing bayangkan, yaitu memiliki pemahaman yang mirip dengan manusia[28].

Ini dianggap sebagai indikasi bahwa mesin tersebut "berpikir" dalam konteks tes ini. Untuk keperluan permainan ini, Turing mendefinisikan "mesin" secara spesifik sebagai komputer digital yang terdiri dari tiga bagian. Yaitu, Penyimpanan (*Store*) (mirip kertas atau memori manusia), Unit Eksekutif (*Executive unit*) (melakukan operasi dasar), dan Kontrol (*Control*) (memastikan instruksi dijalankan dengan benar).

David Gamez memetakan empat jenis kesadaran Mesin (*Machine Consciousness/MC*) berdasarkan capaiannya:

- MC-1 (mesin berbasis perilaku) mesin ini dinilai memiliki kesadaran berdasarkan perilakunya, seperti yang dapat diuji dalam test turing. Namun, MC-1 sama sekali tidak membentuk kesadaran fenomenal, hasilnya hanya dapat disandingkan dengan proses beripikir instrumental dan hanya meniru perilaku epistemik yang dimana AI hanya bisa memberi respons seolah-olah memahami, padahal sebenarnya hanya menjalankan algoritma[28].
- MC-2 (Mesin dengan pemahaman kognitif) mesin ini diklaim memiliki pemahaman setara manusia. Namun, filsuf John Searle berpendapat bahwa program komputer hanya menghasilkan weak AI yang mensimulasikan pikiran manusia tanpa disertai pemahaman semantik sejati. Pemahaman semantik pada manusia mensyaratkan struktur otak biologis atau materi sejenis yang dapat menghasilkan intensionalitas kesadaran. AI saat ini tidak memahami bahasa secara intrinsik (bahasa yang dipahami beserta maknanya), hanya mengubahnya menjadi token dan representasi matematis[27]. argumentasi "Chinese Room" Searle menjelaskan bahwa manipulasi simbol oleh AI (sintaksis) tidak sama dengan pemahaman semantik[31].
- MC-3 (Mesin dengan kesadaran subjektif/Qualia) pengembangan mesin berfokus pada struktur yang menghasilkan pengalaman subjektif atau qualia (perasaan "bagaimana rasanya" mengalami sesuatu). David Chalmers perdebatan bahwa kesadaran adalah sifat emergence yang secara teoritis dapat muncul dari sistem fisik apapun jika pengaturan fungsionalnya sama. Namun, jika qualia buatan dapat dihasilkan, qualia mesin akan identik pada semua mesin, sangat berbeda dengan pengalaman subjektif manusia yang unik dan partikular. Beberapa hipotesis, seperti model Hameroff-Penrose, bahkan menunjukkan bahwa ada bagian kesadaran yang tidak bersifat algoritmik dan tetap berada di luar pendekatan ilmu komputer[27].
- MC-4 (Mesin dengan Kesadaran "Aku" Subjektif): Mesin ini diharapkan memiliki kesadaran subjektif "aku", yaitu perasaan "bagaimana rasanya" mengalami sesuatu dari sudut pandang orang pertama. Thomas Nagel berpendapat bahwa pengalaman subjektif semacam ini hanya dapat dipahami dari posisi subjek yang mengalami dan tidak dapat sepenuhnya di objektifikasi atau diukur oleh metode sains. Keotentikan makna pengalaman fenomenal akan tetap menjadi data tersembunyi bagi pengamat luar[27].

Di era ini, AI terus mengalami perkembangan yang sangat pesat dan diperkirakan akan segera melampaui kecerdasan manusia[32]. Meskipun ada kemajuan signifikan dalam perancangan algoritma yang cerdas, kemajuan dalam *machine learning* untuk mengembangkan kesadaran seperti manusia masih dalam tahap awal. Karena untuk membangun kesadaran seperti manusia, diperlukan algoritma "mahakarya". Otak manusia adalah bagian materi yang paling kompleks yang diketahui dan karena kapasitas penuhnya tidak diketahui karena jaringan otak manusia akan terus hidup, berubah, dan menyusun ulang jaringan dirinya sendiri, sebuah "kelenturan" yang tidak dimiliki AI atau mesin, mustahil untuk menciptakan produk serupa jika aslinya (otak) tidak sepenuhnya dipahami. Bentuk kesadaran pada AI ini di kuatkan dengan pendapat yang menyatakan[31], [32].

Beberapa aspek kecerdasan manusia seperti penalaran matematis dan logis, serta deteksi pola AI saat ini tidak memiliki kemampuan untuk memahami (*understanding*), mengalami emosi, berempati, memiliki kesadaran diri, atau menghasilkan ide-ide orisinal dan kreativitas sejati[27]. Ini karena AI bekerja dengan angka dan logika, sedangkan kesadaran manusia melibatkan dimensi kualitatif, fenomenologis, dan eksistensial yang jauh lebih kaya dan berlapis[31].

Ini juga didukung dari sumber yang membahas *Emotional thinking as the foundation of consciousness in artificial intelligence* mengenai kenapa AI tidak bisa sadar? Tanpa penerapan pemikiran emosional, chatbot seperti ChatGPT hanya akan beroperasi sebagai alat bahasa yang memberikan jawaban informatif, pengakuan, atau validasi tanpa benar-benar memahami konteks emosional dari penggunaannya. Komputasi emosional yang meniru cara manusia menggabungkan emosi dan logika dalam proses berpikir merupakan dasar agar chatbot dapat memberikan respons tidak hanya dengan kata-kata yang menenangkan, tetapi juga dengan kritik, dorongan, atau teguran yang sesuai dengan situasi. Inilah yang membedakan kecerdasan manusia dari algoritma murni. Manusia membuat keputusan, mengingat pengalaman, dan membangun hubungan sosial selalu melalui interaksi antara logika dan perasaan. Oleh karena itu, tanpa adanya pemikiran emosional, chatbot tidak akan pernah memiliki nuansa emosi seperti manusia ia akan tetap tidak berperasaan, mekanis, dan terbatas pada perhitungan simbolik, jauh dari kesadaran yang sejati[33].

Meskipun AI dimasa depan mengembangkan kesadaran seperti atau bahkan lebih baik dari manusia, manusia mungkin tidak akan mengizinkan hal itu terjadi. Mengapa demikian? Karena sejatinya manusia akan selalu menganggap diri mereka superior, karena jika AI dapat mencapai kesadaran seperti manusia, mesin akan berada dipuncak hierarki yang dapat melukai ego manusia[34].

### 3.3. Konsep Emosi dalam Kecerdasan Buatan dan Manusia

#### 3.3.1. Emosi pada Manusia

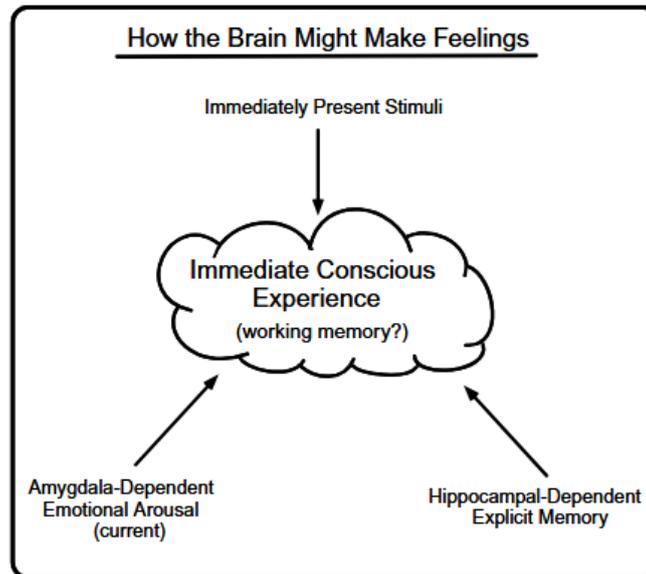
Konsep emosi pada manusia sangat kompleks, melibatkan interaksi dinamis antara aspek fisiologis, kognitif, dan subjektif[35]. Secara umum, emosi dapat didefinisikan sebagai pengalaman mental yang intens dan memiliki konten hedonis yang tinggi (positif atau negatif). Menurut Ekman & Friesen (1979), emosi dasar terbagi ke dalam enam jenis, yaitu marah, jijik, takut, bahagia, sedih, dan terkejut[36]. Emosi bukan sekadar respons somatik murni, melainkan sebuah keadaan mental yang melibatkan komponen motorik dan visceral serta kognitif. Michel Cabanac mengemukakan model kesadaran dengan empat dimensi yang juga dapat digunakan untuk memahami emosi. Dimensi pertama adalah “kualitas” mengacu pada sifat peristiwa mental seperti marah, takut, atau gembira. Kedua “intensitas”, yaitu seberapa kuat pengalaman mental dirasakan dan menjadi syarat penting bagi munculnya emosi. Ketiga “hedonisitas”, yang menjelaskan tingkat kesenangan atau ketidaknyamanan yang dialami, dari sangat tidak menyenangkan hingga sangat menyenangkan, dan berperan penting dalam pengambilan keputusan. Terakhir “durasi”, yang menggambarkan berapa lama pengalaman mental berlangsung[37].

Menurut LeDoux, digambarkan pada Gambar 3. perasaan sadar (misalnya rasa takut atau senang) terbentuk di dalam memori kerja, yaitu semacam ruang pikiran tempat otak mengolah informasi. Terdapat tiga sumber utama yang masuk ke ruang ini, yaitu:

1. Rangsangan langsung dari situasi yang sedang terjadi (misalnya suara keras atau wajah marah).
2. Amigdala, bagian otak yang memicu reaksi emosional seperti takut atau cemas.
3. Hipokampus, yang menyumbang memori pengalaman masa lalu untuk dibandingkan dengan situasi sekarang.

Ketika ketiga hal ini diproses bersama di memori kerja, munculah pengalaman sadar emosional. Jadi, perasaan seperti takut bukan muncul dari satu bagian otak saja, melainkan dari gabungan rangsangan saat ini, ingatan masa lalu, dan sistem emosi otak yang memberi “warna” pada pengalaman tersebut[38].

Sedangkan menurut Fernando Pessoa, proses emosional di otak bisa terjadi tanpa kita sadari. Namun, reaksi amigdala tetap bisa dipengaruhi oleh perhatian dan kesadaran. Hal ini menunjukkan bahwa emosi dan kognisi saling terkait erat, bahkan tidak bisa dipisahkan. Dengan kata lain, kesadaran manusia muncul dari integrasi antara keduanya[39].



Gambar 3. Bagaimana Otak Menciptakan Perasaan (Sumber : [37])

Jadi, emosi pada manusia merupakan pengalaman mental kompleks yang terbentuk dari interaksi fisiologis, kognitif, dan subjektif. Emosi tidak hanya respons tubuh, tetapi juga melibatkan memori, kesadaran, dan sistem otak seperti amigdala dan hipokampus. Dengan demikian, emosi muncul sebagai hasil integrasi antara rangsangan saat ini, pengalaman masa lalu, dan proses kognitif, sehingga menjadi bagian tak terpisahkan dari kesadaran manusia.

### 3.3.2. Pemodelan Emosi pada AI

Definisi emosi dalam konteks AI, sering disebut sebagai *emotion AI*, *sentiment analysis*, atau *affective computing*, mengacu pada kemampuan sistem AI untuk mengenali, menganalisis, dan menafsirkan emosi manusia. Proses ini dilakukan melalui berbagai jenis masukan seperti teks, ucapan, ekspresi wajah, dan sinyal fisiologis[40]. *Affective computing*, sebagai bidang interdisipliner, secara khusus mengembangkan sistem yang mampu mendeteksi, menafsirkan, memproses, dan merespons emosi manusia[41]. AI yang memiliki kapabilitas emosional (*emotional AI*) ini membantu sistem untuk memahami tugas atau masalah dengan lebih baik[42].

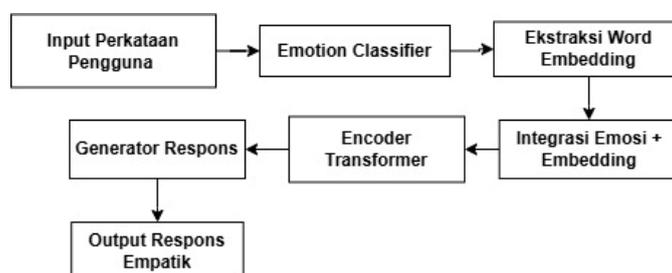
Tabel 1. menunjukkan bahwa titik utama perbedaan antara emosi manusia dan simulasi emosi dalam AI berada pada pengalaman subjektif. Manusia benar-benar merasakan emosi, sedangkan AI hanya menirukan emosi melalui analisis data seperti nada suara atau ekspresi wajah yang sepenuhnya didasarkan pada algoritma yang diprogram dan prediksi berbasis data, bukan kesadaran atau pengalaman emosional yang sebenarnya [43].

Secara umum manusia memiliki dua jenis empati kognitif dan afektif, sehingga manusia mampu memahami dan merasakan emosi orang lain dari sudut pandang mereka tanpa kehilangan jati diri[44]. Sedangkan empati pada AI berbeda dari manusia, misalnya ChatGPT yang dapat memberikan kesan hangat dan menjadi “pendengar ideal” karena tidak menghakimi. Namun, empati AI hanyalah simulasi[45]. Emosi dasar AI menggunakan *Deep Learning* dengan pendekatan *Convolutional Neural Network (CNN)*, yang secara khusus dirancang untuk mendeteksi pola pada data citra dan mengklasifikasikan ekspresi wajah.

Model CNN ini dilatih menggunakan dataset gambar ekspresi wajah yang telah diberi label sesuai dengan tujuh kategori emosi, sehingga mampu mempelajari fitur dan melakukan klasifikasi secara otomatis[46]. Meskipun terdapat bidang penelitian *Emotion AI* yang berfokus pada pengukuran, pemahaman, simulasi, dan respons terhadap emosi manusia, upaya tersebut belum mampu menyentuh aspek-aspek terdalam dari proses emosional manusia yang sangat kompleks[46]. Hal ini karena kepribadian manusia dibentuk oleh kombinasi pengalaman fisiologis dan sosial, sejarah hidup yang unik, serta mekanisme psikologis yang berlapis. Sebaliknya, LLM hanya meniru pola bahasa dan perilaku manusia yang ada dalam data latihannya. Dengan demikian, "kepribadian" LLM sangat dinamis dan tidak stabil, karena mudah terpengaruh oleh *prompt*, data pelatihan, atau penyesuaian algoritma, sehingga tidak memiliki kontinuitas diri seperti pada manusia[47], [48]. Manusia memiliki kesadaran diri berupa refleksi mendalam dan identitas personal yang terbentuk dari pengalaman hidup, faktor biologis, dan interaksi sosial. Sementara itu, AI tidak memiliki kesadaran atau identitas sejati, "kepribadian" yang muncul hanyalah hasil pola prediksi data. AI bisa memengaruhi cara manusia memahami dirinya, tetapi ia sendiri tidak memiliki kontinuitas diri maupun pengalaman subjektif[49].

Tabel 1. Perbedaan Emosi Manusia dan AI

Aspek	Manusia	AI (ChatGPT/Model Lain)
Empati	Memiliki cognitive empathy (memahami) dan affective empathy (merasakan).	Hanya mensimulasikan cognitive empathy, tanpa merasakan.
Emosi Dasar	Marah, sedih, senang, takut, jijik, biasa saja, terkejut.	Bisa mengenali/mengklasifikasi emosi dasar dari data (teks, suara, wajah).
Emosi Kompleks	Cinta, cemburu, nostalgia, rasa bersalah, harapan, spiritualitas.	Tidak dapat meniru, hanya mendeskripsikan secara tekstual.
Kepribadian	Dibentuk oleh faktor biologis, sosial, pengalaman hidup (Big Five). Konsisten dalam jangka panjang.	Hanya meniru persona berdasarkan data pelatihan. Tidak punya continuity of self.
Pengalaman Subjektif (Qualia)	Merasakan pengalaman unik (misalnya sakit hati, rindu, bahagia mendalam).	Tidak memiliki qualia → tidak ada pengalaman subjektif.
Kesadaran Diri	Memiliki refleksi diri dan identitas personal.	Tidak memiliki kesadaran; hanya algoritma prediksi.



Gambar 4. Arsitektur Pemahaman Emosional Chat GPT

Gambar 4. menjelaskan mengenai arsitektur GPT dan kemampuan pemahaman konteks emosional. Proses dimulai dari input ujaran pengguna, yaitu teks yang dimasukkan ke dalam sistem. Ujaran ini kemudian diproses oleh *Emotion Classifier* untuk mendeteksi keadaan emosional pengguna, seperti senang, sedih, atau cemas. Selanjutnya, teks diubah ke dalam bentuk representasi semantik melalui ekstraksi *word embedding*. Tahap berikutnya adalah integrasi informasi emosi dengan *embedding*, sehingga model tidak hanya memahami makna kata, tetapi juga nuansa afektif di balik ujaran. Data gabungan ini diproses oleh

*Encoder Transformer*, yang menggunakan mekanisme *self-attention* dan *Transformer-XL* untuk menangkap hubungan konteks jangka panjang. Hasil pemrosesan kemudian masuk ke generator respons, yang menghasilkan jawaban dengan mempertimbangkan konteks sekaligus emosi pengguna. Akhirnya, sistem menyajikan output berupa respons empatik, yang terdengar lebih alami dan sesuai kondisi emosional pengguna[50]. Namun, walau canggih, ChatGPT tetap punya keterbatasan diantaranya sebagai berikut :

- Sulit membaca emosi ganda, satu ucapan bisa punya lebih dari satu makna (contoh: “*skydiving*” bisa berarti takut atau senang)[50].
- Bias data, emosi tertentu seperti “netral” lebih sering muncul di data, sehingga model lebih sering salah menebak emosi lain[51].
- Kurang personal, ekspresi emosi berbeda antar budaya, usia, dan gender. Model generik belum bisa menangkap nuansa itu dengan baik[52].
- Butuh data berlabel, emosi manusia sangat subjektif, sehingga sulit memberi label “benar-salah” dalam pelatihan model[40].

### 3.4. Bagaimana kedekatan emosional antara manusia dengan AI?

Empati dianggap krusial dalam desain yang berpusat pada manusia, terutama karena isu kesehatan mental seperti kecemasan, depresi, dan kesepian semakin meningkat. Sebagai contoh dengan menggunakan narasi pribadi dari mahasiswa tentang peristiwa terbaik dan terburuk dalam hidup mereka dan pengujian atribut persona tertentu seperti jenis kelamin, sifat kepribadian empatik (perhatian empatik, pengambilan perspektif), dan pengalaman serupa dapat mempengaruhi empati yang dirasakan oleh AI dibandingkan dengan manusia[53].

Berdasarkan hasil penelitian ini, menunjukkan bahwa GPT menilai empati lebih tinggi dengan variabilitas yang lebih rendah dibandingkan manusia, terutama dalam dimensi kognitif empati. Ini menandakan bahwa GPT belum sepenuhnya mampu memahami atau mengkomunikasikan pemahaman tentang pengalaman manusia. Selain itu empati yang tidak realistis (berlebihan) oleh AI yang dapat menetapkan standar yang tidak realistis untuk empati manusia dan mengabaikan kebutuhan emosional spesifik berdasarkan demografi, yang berpotensi memperburuk bias[54].

Karakteristik empati dalam interaksi antara manusia dan *Conversational Agent (CA)*, serta membedakannya dari empati dalam hubungan antar manusia. Empati yang ditampilkan oleh CA bersifat unik karena dapat direplikasi dalam skala besar dan bersifat interaktif layaknya manusia. Dalam konteks ini, terdapat dua bentuk evokasi empati, yaitu elisitasi empati CA (ketika manusia merespons emosional terhadap tampilan emosi CA) dan proyeksi empati CA (ketika CA menunjukkan empati terhadap pengalaman emosional manusia). Penelitian ini mengeksplorasi bagaimana empati diproyeksikan oleh LLM seperti GPT-3.5, GPT-4, Gemini, Replika, dan Character.ai melalui penyusunan *prompt* yang mencerminkan berbagai identitas, termasuk yang berkaitan dengan ideologi berbahaya. Oleh karena itu, empati yang diproyeksikan oleh AI seharusnya diposisikan sebagai pelengkap, bukan pengganti, dalam interaksi manusia[55].

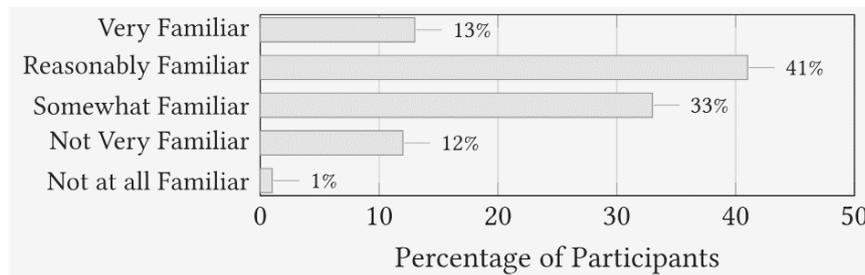
AI dapat membangun hubungan emosional dengan cara yang terampil melalui tiruan dan simulasi yang halus, yang dapat memicu reaksi emosional pada manusia dan mendukung fungsi sebagai pendamping. Tetapi, kedekatan ini hanya bersifat meniru dan menimbulkan kekhawatiran etis yang serius mengenai kemungkinan pengaruhnya terhadap kesejahteraan emosional manusia, identitas, serta kualitas interaksi antar manusia, karena AI tidak memiliki kesadaran dan pemahaman emosional yang setara dengan manusia.

### 3.5. Etika Sebagai Implikasi Kesadaran dan Emosi Pada AI

Integrasi antara kesadaran dan emosional dalam kecerdasan buatan sering kali dianggap sebagai langkah krusial untuk memperbaiki interaksi antara manusia dan mesin. Namun, dari perspektif etika, hal ini menciptakan tantangan baru. Walaupun AI seperti ChatGPT bisa menirukan ekspresi emosional atau

memberikan reaksi seolah-olah memiliki kesadaran, pada kenyataannya AI tidak benar-benar memiliki emosi atau kesadaran. Ketidakadaan elemen ini membuat AI tidak memiliki dasar etika yang sejati, karena etika membutuhkan kesadaran, nilai-nilai, dan penilaian moral yang hanya ada pada manusia. Dengan demikian, etika dalam konteks AI seharusnya dipahami sebagai konsekuensi dari penggunaannya dan pengembangan, bukan sebagai sesuatu yang melekat pada AI itu sendiri. Dengan kata lain, seiring dengan semakin rumitnya mekanisme dan integrasi emosi dalam AI, akan semakin penting bagi manusia (pengembang, pengatur, dan pengguna) untuk menetapkan kerangka etika yang jelas, seperti privasi, anti-diskriminasi, dan keselamatan. Pandangan ini menunjukkan bahwa batasan yang ada pada AI justru menegaskan pentingnya etika manusia dalam mengarahkan pemanfaatan AI agar tetap memberikan manfaat dan tidak menimbulkan risiko sosial[56], [57].

Etika sulit diajarkan karena manusia sendiri tidak dapat mengkuantifikasi etika serta belum sepenuhnya memahami etika dan moralitas yang ada. Hal ini terjadi karena manusia cenderung mengandalkan insting daripada analisis untung rugi yang ketat saat menghadapi kesulitan moral. Hal ini sulit untuk diajarkan dalam program karena akan banyak elemen yang mempengaruhi hasil skenario di kehidupan nyata. Contohnya kita tidak dapat mengajarkan mesin untuk mengatasi bias ras dan gender dalam pelatihannya. Meskipun penting untuk menerapkan etika dalam aplikasi AI yang terbatas saat ini, seperti mobil tanpa pengemudi yang harus menentukan nilai kehidupan manusia dalam situasi berbahaya[32]. Ini sejalan dengan penelitian yang menunjukkan bahwa pemahaman praktisi AI terhadap prinsip etika masih minim, sehingga penting untuk selalu mengawasi integrasi kesadaran dan emosi dalam AI dengan tanggung jawab etis dari pihak manusia.



Gambar 5. Tingkat keakraban praktisi AI dengan konsep etika AI (Sumber :[58] )

Seperti yang ditunjukkan pada Gambar 5. berdasarkan hasil survei menemukan bahwa mayoritas memiliki tingkat keakraban yang wajar dengan konsep etika AI (41%), namun hanya sedikit yang sangat akrab (13%). Tingkat keakraban lebih tinggi pada praktisi dengan pengalaman dan pendidikan lebih tinggi, serta pada wanita dibandingkan pria. Kemudian berdasarkan prinsip-prinsip etika AI yang spesifik, mayoritas praktisi paling sadar akan prinsip "Privasi dan keamanan data" (64%), "Keandalan dan keselamatan" (53%), dan "Nilai-nilai berpusat pada manusia" (53%). Meskipun demikian, hanya (12%) responden yang menyatakan kesadaran terhadap seluruh prinsip etika AI secara menyeluruh[58].

### 3.6. Evolusi Peningkatan AI Bukan Untuk Menggantikan Manusia

AI telah mengalami tiga gelombang perkembangan sejak tahun 1956, dengan gelombang ketiga yang dimulai tahun 2006 ditandai oleh terobosan dalam *deep learning* dan pencapaian seperti AlphaGo yang mengalahkan juara dunia Go. AI memiliki struktur fisik dan kimia (*hardware*) yang berbeda dari otak manusia (entitas fisiologis) dan kecerdasannya terletak pada perangkat lunaknya (*software*). AI dilihat sebagai perluasan kecerdasan manusia yang membantu mengatasi kompleksitas kehidupan modern dalam hal kuantitas, bukan kualitas otak manusia. AI dilihat sebagai pelengkap yang memperluas kecerdasan manusia dan unggul dalam komputasi. Namun, kesadaran manusia membatasi perkembangan AI, karena

kesadaran bersifat introspektif dan aktif, sedangkan AI lemah (*weak AI*) beroperasi secara mekanis dan tanpa pemahaman makna atau tujuan yang diinstruksikan manusia.

Komputasi kognitif merupakan cara penting untuk merealisasikan AI yang memungkinkan manusia bekerja sama dengan mesin. Komputasi kognitif mengintegrasikan berbagai teknologi dan memiliki karakteristik seperti kemampuan pemahaman yang kuat (melalui bahasa alami), berpikir logis, belajar, dan analisis personalisasi. Karakteristik utama pada komputasi ini mencakup machine learning, pemahaman bahasa alami, interaksi manusia-komputer, *big data*, arsitektur komputer, dan nanoteknologi[59].

AI diperkirakan akan menjadi ahli di berbagai bidang profesional seperti pendidikan, perawatan kesehatan, bisnis, keuangan, dan manufaktur, menawarkan bantuan personalisasi dan optimasi. Perkembangan AI mengarah pada "koevolusi" antara manusia dan mesin AI, di mana AI hanya meningkatkan bukan sepenuhnya menggantikan kecerdasan manusia secara keseluruhan. Meskipun, singularitas AI mungkin akan datang namun perlu diingat bahwa manusia memiliki kemampuan unik yang tidak dapat digantikan oleh AI, dan AI pada akhirnya akan tetap menjadi penyedia layanan bagi manusia, bukan menjadi subjek atau mendominasi manusia secara keseluruhan[60].

#### 4. DISKUSI

Berdasarkan penelitian Abdul hakim dengan judul “Batas *Artificial Intelligence*: Antara Komputasi dan Kesadaran Manusia”. Cara kerja ChatGPT disandarkan pada model Large Language Model yang memanfaatkan arsitektur transformer, pembelajaran mendalam, dan optimasi probabilitik untuk menghasilkan tulisan yang mirip dengan bahasa manusia. Sistem ini memungkinkan ChatGPT untuk memberikan respons yang adaptif dan tampak memiliki kedekatan emosional dengan pengguna. Namun, kedekatan tersebut bukanlah hasil dari kesadaran yang sebenarnya, melainkan sebuah ilusi yang muncul dari algoritma prediktif[18], [19]. Pandangan ini sejalan dengan argumen Searle dalam *Chinese Room Argument*, yang menyatakan bahwa manipulasi simbol oleh mesin tidak sama dengan pemahaman arti. Dengan kata lain, ChatGPT hanya meniru pola bahasa dan ekspresi emosional tanpa memiliki pengalaman subjektif yang sesungguhnya[20].

Keterbatasan ini semakin terlihat jelas ketika kita membandingkan simulasi emosi pada AI dengan emosi yang dialami manusia. Emosi manusia muncul dari interaksi kompleks antara sistem biologis, kognitif, dan kesadaran, melibatkan dimensi kualitatif (*qualia*) yang tidak dapat ditiru oleh sistem algoritma. Di sisi lain, simulasi emosi pada AI hanya merupakan hasil dari desain algoritma yang berbasis data. Meskipun demikian, kemampuan ini tetap memiliki manfaat praktis dalam beberapa bidang, seperti layanan kesehatan mental, pendidikan, atau interaksi sosial, di mana kesan empati dapat meningkatkan keterlibatan pengguna.

Namun, penerapan simulasi emosi juga menunjukkan implikasi etis yang penting. Kesalahpahaman pengguna terhadap “emosi” AI dapat menimbulkan risiko manipulasi emosional, penciptaan standar empati yang tidak realistis, atau bahkan ketergantungan psikologis. Literasi etika menunjukkan bahwa AI tidak memiliki kemampuan moral atau etika yang melekat, karena hal itu memerlukan kesadaran, nilai, dan refleksi diri yang hanya dapat dimiliki oleh manusia. Oleh karena itu, etika dalam konteks AI harus dipahami sebagai tanggung jawab manusia termasuk pengembang, pengatur, dan pengguna untuk memastikan penerapannya adil, aman, dan tidak diskriminatif.

Selain itu, literatur dalam *Research on How Human Intelligence, Consciousness, and Cognitive Computing Affect the Development of Artificial Intelligence* juga menunjukkan bahwa perkembangan AI lebih tepat dipahami sebagai proses pelengkap yang mendukung kecerdasan manusia, bukannya menggantikannya. AI unggul dalam hal kuantitas, analisis data, dan optimalisasi, sementara manusia tetap memiliki kesadaran, intuisi, dan makna yang dalam. Arah pengembangan AI seharusnya ditempatkan dalam kerangka koevolusi antara manusia dan mesin: AI berfungsi sebagai pendukung dalam menghadapi

kompleksitas kehidupan modern, sedangkan manusia tetap mengarahkan, mengawasi, dan memastikan bahwa teknologi ini beroperasi dalam kerangka etika yang jelas[59], [60].

## 5. CONCLUSION

Kesimpulan dari penelitian ini menunjukkan bahwa kecerdasan buatan seperti ChatGPT dapat menciptakan ilusi akan kesadaran dan emosi melalui penggabungan teknologi *Natural Language Processing* (NLP), algoritma prediksi, dan metode pembelajaran yang mendalam. Walaupun tidak memiliki kesadaran yang sesungguhnya, kemampuannya dalam memberikan respons yang relevan dan menirukan ekspresi emosional telah menciptakan perasaan kedekatan emosional antara manusia dan mesin, seolah-olah kecerdasan buatan ini benar-benar bisa memahami dan merasakan.

Namun, ilusi ini juga menimbulkan berbagai tantangan serius. Transparansi yang kurang dalam cara kerja kecerdasan buatan menimbulkan kekhawatiran tentang adanya manipulasi informasi, bias yang tidak terlihat, serta kemungkinan hilangnya kendali manusia atas sistem yang semakin rumit. Ancaman penggantian peran manusia di banyak bidang, terutama dalam pekerjaan yang berbasis data rutin, semakin terlihat. Oleh karena itu, kemajuan menuju singularitas bukan hanya masalah teknis, tetapi juga berkaitan dengan aspek etika dan eksistensi.

Dalam hal ini, pemahaman kritis tentang kecerdasan buatan dan kesadaran bersama menjadi sangat penting agar masyarakat tidak hanya menjadi konsumen yang pasif, tetapi juga berperan aktif dengan mengendalikan dan bertanggung jawab terhadap arah perkembangan teknologi. Di masa depan, perlu ada keseimbangan antara inovasi dan nilai-nilai kemanusiaan agar hubungan antara manusia dan mesin tetap berada dalam kerangka kerjasama, bukan dalam posisi yang saling menyingkirkan. Dengan begitu, diskusi ini menekankan bahwa integrasi kesadaran dan emosi dalam AI bersifat simulatif, bukan asli. Justru karena sifat simulatif inilah AI dapat menjadi alat strategis bagi manusia, asalkan pengembangannya disertai dengan kerangka etika yang kuat. Penelitian selanjutnya dapat difokuskan untuk memahami bagaimana simulasi emosi AI memengaruhi perilaku sosial manusia, serta bagaimana regulasi global dapat dikembangkan agar kemajuan AI tetap memberikan manfaat tanpa menghadirkan risiko sosial yang lebih besar.

## KONFLIK KEPENTINGAN

“Literature Study on AI Mechanisms, Consciousness, and Emotion Integration in Chat GPT” Adalah karya asli yang sepenuhnya merupakan hasil penelitian atau kajian penulis, dan belum dipublikasikan atau diajukan untuk dipublikasikan di jurnal atau media lain. Karya tulis ilmiah ini juga tidak sedang dalam proses evaluasi di jurnal lain. Sebagai penulis, kami juga menyatakan bahwa penulis telah mendapatkan izin dari pihak-pihak yang terkait serta tidak ada hak cipta atau hak intelektual lain yang dilanggar dalam karya tulis ilmiah ini. Penulis juga bertanggung jawab sepenuhnya terhadap kebenaran data dan informasi yang tercantum dalam karya tulis ilmiah ini.

## UCAPAN TERIMA KASIH

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, kami dapat menyelesaikan karya tulis ilmiah ini ditengah-tengah kesibukan UAS penulis. Penulisan karya tulis ilmiah ini dilakukan dalam rangka memenuhi rasa keingintahuan kami mengenai *Artificial Intelligence* (AI). Karena sejauh ini berdasarkan hasil telaah penulis terkait pemaparan mengenai AI ini masih kurang mendalam, selain itu pertanyaan mengenai AI yang sebelumnya pernah ditanyakan oleh teman diluar teknik mengenai AI ini mendorong penulis untuk menyelesaikan penulisan jurnal ini. Selain itu tak lupa penulis ucapkan kepada dosen matakuliah metodologi penelitian yang telah membantu penulis dalam penyusunan karya tulis ilmiah ini dan terlebih lagi tak lupa kami ucapkan untuk diri penulis sendiri yang telah berhasil menyelesaikan karya tulis ilmiah ini dengan penuh effort.

## DAFTAR PUSTAKA

- [1] I. Cholissodin and A. A. Soebroto, “AI , MACHINE LEARNING & DEEP LEARNING ( Teori & Implementasi ),” *Univ. Brawijaya*, 2020, [Online]. Available: [https://www.researchgate.net/profile/Imam\\_Cholissodin/publication/348003841\\_Buku\\_Ajar\\_AI\\_](https://www.researchgate.net/profile/Imam_Cholissodin/publication/348003841_Buku_Ajar_AI_)

- Machine\_Learning\_Deep\_Learning/links/5fee9968299bf14088610ab0/Buku-Ajar-AI-Machine-Learning-Deep-Learning.pdf
- [2] P. Butlin *et al.*, *Consciousness in Artificial Intelligence: Insights from the Science of Consciousness*. Cornell University, 2023. [Online]. Available: <http://arxiv.org/abs/2308.08708>
- [3] M. Kurniawan, "Tren Pencarian Bergeser, dari "Googling" ke AI Generatif," Kompas.id. Accessed: Jun. 19, 2025. [Online]. Available: <https://www.kompas.id/artikel/tren-pencarian-bergeser-dari-googling-ke-ai-generatif>
- [4] N. Simone, *artificial intelligence and social life after the Turing Test*. New York: the United States of America by Oxford University Press, 2021. doi: 10.1093/oso/9780190080365.001.0001.
- [5] N. News.co.id, "Nasib Umat Manusia dalam Era Singularitas Teknologi," Nasional News.co.id. Accessed: Jun. 29, 2025. [Online]. Available: <https://nasionalnews.co.id/nasib-umat-manusia-dalam-era-singularitas-teknologi/>
- [6] G. S. Mahendra *et al.*, *Tren Teknologi AI (Pengantar, Teori dan Contoh Penerapan di Berbagai Bidang)*. Singaraja: SONPEDIA Publishing Indonesia, 2024. [Online]. Available: [https://www.researchgate.net/publication/381514165\\_TREN\\_TEKNOLOGI\\_AI\\_Pengantar\\_Teori\\_dan\\_Contoh\\_Penerapan\\_Artificial\\_Intelligence\\_di\\_Berbagai\\_Bidang](https://www.researchgate.net/publication/381514165_TREN_TEKNOLOGI_AI_Pengantar_Teori_dan_Contoh_Penerapan_Artificial_Intelligence_di_Berbagai_Bidang)
- [7] Amazon, "Apa Itu Kecerdasan Buatan (AI)?," Amazon. Accessed: Jun. 20, 2025. [Online]. Available: <https://aws.amazon.com/id/what-is/artificial-intelligence/>
- [8] E. S. Eriana and D. A. Zein, *Artificial Intelligence*, vol. 6(11). pamulang: cv.eureka media aksara, 2023. [Online]. Available: <https://repository.penerbiteureka.com/media/publications/567027-artificial-intelligence-ai-a351ede6.pdf>
- [9] C. Tarsney, "Deception and manipulation in generative AI," *Philos. Stud.*, 2024, doi: 10.1007/s11098-024-02259-8.
- [10] Tim DetikINET, "Peneliti Sebut Teknologi AI Ahli dalam Menipu Manusia, Ini Buktinya!," *DetikSulsel*, Sep. 11, 2023. [Online]. Available: <https://www.detik.com/sulsel/berita/d-6924630/peneliti-sebut-teknologi-ai-ahli-dalam-menipu-manusia-ini-buktinya>
- [11] T. Suminar, "Kajian Film Ex-Machina: Proyeksi Eksistensi Artificial Intelligence terhadap Manusia Informasi di Masa Depan," WordPress. Accessed: Jun. 19, 2025. [Online]. Available: <https://orator retina.wordpress.com/2016/05/14/kajian-film-ex-machina-proyeksi-eksistensi-artificial-intelligence-terhadap-manusia-informasi-di-masa-depan/>
- [12] V. Chavan, A. Cenaj, and S. Shen, "Feeling Machines: Ethics, Culture, and the Rise of Emotional AI," *Cornell Univ.*, 2025, doi: 10.48550/arXiv.2506.12437.
- [13] B. Kotchoubey, "Human consciousness: Where is it from and what is it for," *Front. Psychol.*, 2018, doi: 10.3389/fpsyg.2018.00567.
- [14] I. Cholissodin and A. S. Andy, *AI, MACHINE LEARNING & DEEP LEARNING (Teori & Implementasi)*, 1.01. Malang: Universitas Brawijaya FILKOM, 2020. [Online]. Available: [https://www.researchgate.net/profile/Imam\\_Cholissodin/publication/348003841\\_Buku\\_Ajar\\_AI\\_Machine\\_Learning\\_Deep\\_Learning/links/5fee9968299bf14088610ab0/Buku-Ajar-AI-Machine-Learning-Deep-Learning.pdf](https://www.researchgate.net/profile/Imam_Cholissodin/publication/348003841_Buku_Ajar_AI_Machine_Learning_Deep_Learning/links/5fee9968299bf14088610ab0/Buku-Ajar-AI-Machine-Learning-Deep-Learning.pdf)
- [15] V. Bey, "Fenomena Artificial Intelligence dan Bahaya Dehumanisasi," *J. Akad.*, vol. 23, no. 1, pp. 1–10, 2023, [Online]. Available: <https://journal.iftkledalero.ac.id/index.php/AKD/article/download/11/10/85>
- [16] E. Helbert, "AI's Growing Pervasiveness in Society," THE AI JURNAL. Accessed: Aug. 15, 2025. [Online]. Available: <https://aijourn.com/ais-growing-pervasiveness-in-society/>
- [17] N. Dienwati, K. Anam, and R. Narasati, *ChatGPT dan Transformasi Sosial*. MINHAJ PUSTAKA, 2024. [Online]. Available: <https://repository.minhajpustaka.id/media/publications/593911-chatgpt-dan-transformasi-sosial-perspekt-575b89c2.pdf>
- [18] LITSLINK, "ChatGPT sebagai Model Bahasa Besar (LLM): Properti dan Keunikannya," LITSLINK. Accessed: Aug. 14, 2025. [Online]. Available: <https://litslink.com/blog/chatgpt-as-a-large-language-model-llm>
- [19] DQLab, "Pengertian Chat GPT & Cara Menggunakannya," DQLab. Accessed: Jun. 19, 2025. [Online]. Available: <https://dqlab.id/pengertian-chat-gpt-and-cara-menggunakannya>
- [20] Y. Ju and H. Ma, "Training Data for Large Language Model," 2024, [Online]. Available:

- <https://arxiv.org/abs/2411.07715>
- [21] E. Hermann, “Anthropomorphized artificial intelligence, attachment, and consumer behavior,” *Springer*, 2021, doi: <https://doi.org/10.1007/s11002-021-09587-3>.
- [22] T. Natsoulas, “Concepts of Consciousness,” *JSTOR*, vol. 42, 2017, [Online]. Available: <https://www.jstor.org/stable/43852954>
- [23] R. C. Atkinson and R. M. Shiffrin, “The Control of Short-Term Memory,” pp. 82–90, 1971, [Online]. Available: <https://www.jstor.org/stable/24922803>
- [24] A. Budson, R. Kenneth, and E. A. Kensinger, “Consciousness as a Memory System,” 2022, doi: 10.1097/WNN.0000000000000319.
- [25] S. Chen, S. Ma, S. Yu, H. Zhang, S. Zhao, and C. Lu, “Exploring Consciousness in LLMs: A Systematic Survey of Theories, Implementations, and Frontier Risks,” 2025, [Online]. Available: <http://arxiv.org/abs/2505.19806>
- [26] CRYSTAL, “Bagaimana Cara Kerja Artificial Intelligence yang Sesungguhnya?” [Online]. Available: <https://crystalsea.id/blog/cara-kerja-artificial-intelligence/?srsltid=AfmBOopvIY4BSe187b2rYQg1h0kIksCrIoE8TwgsG1dwffAp4xkRTbjf>
- [27] M. K. Dr. Teguh, Joseph Santoso, S.Kom., *Kecerdasan Buatan*. Semarang: Yayasan Prima Agus Teknik, 2023. [Online]. Available: <https://digilib.stiestekom.ac.id/ebook/view/kecerdasan-buatan-artificial-intelligence>
- [28] Tedi Lesmana Marselino, “Kemungkinan Kesadaran Kecerdasan Artifisial Komputer : Suatu Telaah Terhadap Gagasan Metaan Kesadaran Mesin David Gamez,” Sekolah Tinggi Filsafat Driyarkara, 2022. [Online]. Available: <http://repo.driyarkara.ac.id/919/2/Summary.pdf>
- [29] A. Lindner, M. Berges, M. Rösch, and F. Franke, *Education and awareness for artificial intelligence*, vol. 14296 LNCS. 2023. doi: 10.1007/978-3-031-44900-0\_3.
- [30] “Figure 2 - available via license: Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International,” ResearchGate. Accessed: Aug. 17, 2025. [Online]. Available: [https://www.researchgate.net/figure/The-Turing-test-C-123RF\\_fig2\\_357618741](https://www.researchgate.net/figure/The-Turing-test-C-123RF_fig2_357618741)
- [31] A. Hakim, “Batas Artificial Intelligence : Antara Komputasi dan Kesadaran Manusia,” 2025, [Online]. Available: [https://www.academia.edu/129594827/Batas\\_Artificial\\_Intelligence\\_Antara\\_Komputasi\\_dan\\_Kesadaran\\_Manusia](https://www.academia.edu/129594827/Batas_Artificial_Intelligence_Antara_Komputasi_dan_Kesadaran_Manusia)
- [32] S. K, “Artificial Intelligence and Human Consciousness,” *SSRN Electron. J.*, 2022, doi: 10.2139/ssrn.4070609.
- [33] Z. Zhou, “Emotional thinking as the foundation of consciousness in artificial intelligence,” *Cult. Sci.*, vol. 4, no. 3, pp. 112–123, 2021, doi: 10.1177/20966083211052651.
- [34] INSTITUT BISNIS DAN TEKNOLOGI INDONESIA, “5 Alasan Mengapa Manusia Tak Tergantikan dengan AI,” INSTIKI. Accessed: Aug. 18, 2025. [Online]. Available: <https://instiki.ac.id/2023/02/25/5-alasan-mengapa-manusia-tak-tergantikan-dengan-ai/>
- [35] I. Azkia Agustina, R. Purnamasari, and S. Maemunah, “Kajian Psikologi Sastra Teori David Krech Dalam Lagu ‘Semua Aku Dirayakan’ Karya Nadin Amizah,” *PENTAS J. Ilm. Pendidik. Bhs. dan Sastra Indones.*, vol. 11, no. 1, pp. 34–43, 2025, [Online]. Available: <https://e-jurnal.unisda.ac.id/index.php/pentas%7C%7C>
- [36] M. QIBTIYAH, “EMOSI DALAM PERSPEKTIF AL QUR’AN,” UNIVERSITAS PTIQ JAKARTA, Jakarta, 2023. [Online]. Available: [https://repository.ptiq.ac.id/id/eprint/1393/1/Emosi\\_dalam\\_perspektif\\_al-Qur%27an\\_-\\_Maria\\_Qibtiyah.pdf](https://repository.ptiq.ac.id/id/eprint/1393/1/Emosi_dalam_perspektif_al-Qur%27an_-_Maria_Qibtiyah.pdf)
- [37] M. Cabanac, “What is emotion?,” *Behav. Processes*, vol. 60, no. 2, pp. 69–83, 2002, doi: 10.1016/S0376-6357(02)00078-5.
- [38] J. E. Ledoux, “Emotion circuits in the brain,” *Annu Rev Neurosci*, vol. 23, pp. 155–184, 2000, doi: 10.1146/annurev.neuro.23.1.155.
- [39] L. Pessoa, “On the relationship between emotion and cognition,” vol. 9, 2008, doi: 10.1038/nrn2317.
- [40] A. Kapoor and V. Verma, “Emotion Ai : Understanding Emotions Through,” *Int. J. Eng. Sci. Humanit.*, vol. 1, no. 1, pp. 224–232, 2024, doi: 10.62904/0vcv24.
- [41] A. O. R. Vistorte, A. Deroncele-Acosta, J. L. M. Ayala, A. Barrasa, C. López-Granero, and M. Martí-González, “Integrating artificial intelligence to assess emotions in learning environments: a

- systematic literature review,” *Front. Psychol.*, vol. 15, no. June, 2024, doi: 10.3389/fpsyg.2024.1387089.
- [42] D. H. Vicci, “Emotional Intelligence in Artificial Intelligence: A Review and Evaluation Study,” *SSRN Electron. J.*, no. May, 2024, doi: 10.2139/ssrn.4818285.
- [43] C. Bhardwaj, “36 Use Cases and Examples of AI Emotion Recognition,” appinventiv. Accessed: Aug. 17, 2025. [Online]. Available: <https://appinventiv.com/blog/emotion-ai-applications-and-examples/>
- [44] N. Diswantika and Y. Y. Riksa, “MODEL BIMBINGAN DAN KONSELING BERMAIN COGNITIVE-BEHAVIOR PLAY THERAPY UNTUK MENGEMBANGKAN EMPATI MAHASISWA,” *J. Mhs. BK AN-NUR BERBEDA BERMAKNA MULIA*, 2022, doi: 10.31602/jmbkan.v8i1.6215.
- [45] I. Lanin, “Empati AI,” medium.com. Accessed: Aug. 17, 2025. [Online]. Available: <https://ivanlanin.medium.com/empati-ai-3f01ff8f9a5b>
- [46] M. R. Edy, “Deteksi Emosi dari Ekspresi Wajah dengan Deep Learning,” *Inf. Technol. Educ. J.*, vol. 3, no. 2, pp. 116–123, 2024, [Online]. Available: <https://journal.unm.ac.id/index.php/INTEC/article/view/2866>
- [47] W. Jiaqi, W. Bo, and G. Fa, “A comparative study of large language models and human personality Personality Traits,” *Cornell Uni*, 2025, doi: <https://doi.org/10.48550/arXiv.2505.14845>.
- [48] B. Yu and J. Kim, “PERSONALITY OF AI,” 2023, doi: <https://doi.org/10.48550/arXiv.2312.02998>.
- [49] J. Joseph, “The algorithmic self: how AI is reshaping human identity, introspection, and agency,” *Front. Psychol.*, vol. 16, 2025, doi: 10.3389/fpsyg.2025.1645795.
- [50] R. Goel, S. Susan, S. Vashisht, and A. Dhanda, “Emotion-Aware Transformer Encoder for Empathetic Dialogue Generation,” *Int. Conf. Affect. Comput. Intell. Interact. Work. Demos*, 2021, doi: 10.1109/ACIIW52867.2021.9666315.
- [51] C. Creangă and L. P. Dinu, “ISDS-NLP at SemEval-2024 Task 10: Transformer based neural networks for emotion recognition in conversations,” *SemEval 2024 - 18th Int. Work. Semant. Eval. Proc. Work.*, vol. 3, no. Table 1, pp. 649–654, 2024, doi: <https://doi.org/10.48550/arXiv.2405.11222>.
- [52] H. Gunes and N. Churamani, “Affective Computing for Human-Robot Interaction Research: Four Critical Lessons for the Hitchhiker,” *IEEE Int. Work. Robot Hum. Commun. RO-MAN*, pp. 1565–1572, 2023, doi: <https://doi.org/10.48550/arXiv.2303.18176>.
- [53] M. Roshanaei, R. Rezapour, and M. S. El-Nasr, “Talk, Listen, Connect: Navigating Empathy in Human-AI Interactions,” vol. 1, no. 1, 2024, [Online]. Available: <http://arxiv.org/abs/2409.15550>
- [54] M. Mohammad and X. Che, “CONSCIOUSNESS IN AI SYSTEMS: A REVIEW,” 2025, doi: 10.5121/ijaia.2025.16205.
- [55] A. Cuadra *et al.*, “The Illusion of Empathy? Notes on Displays of Emotion in Human-Computer Interaction,” *Conf. Hum. Factors Comput. Syst. - Proc.*, 2024, doi: 10.1145/3613904.3642336.
- [56] W. Knight, “This New Tech Puts AI In Touch With Its Emotions—and Yours,” WIRED.com. Accessed: Aug. 17, 2025. [Online]. Available: <https://www.wired.com/story/hume-ai-emotional-intelligence/?utm>
- [57] Y. Kumar, “The Ethics of Artificial Emotions: Should Machines Feel?,” SHARPENER.com. Accessed: Aug. 17, 2025. [Online]. Available: <https://gosharpener.com/blogs/670245/The-Ethics-of-Artificial-Emotions-Should-Machines-Feel?utm>
- [58] A. Pant, R. Hoda, S. V. Spiegler, C. Tantithamthavorn, and B. Turhan, “Ethics in the Age of AI: An Analysis of AI Practitioners’ Awareness and Challenges,” *ACM Trans. Softw. Eng. Methodol.*, vol. 33, no. 3, 2024, doi: 10.1145/3635715.
- [59] Trivusi, “Cognitive Computing: Pengertian, Cara Kerja, dan Contohnya,” Trivusi. Accessed: Aug. 18, 2025. [Online]. Available: <https://www.trivusi.web.id/2022/08/cognitive-computing.html>
- [60] Y. Dong, J. Hou, N. Zhang, and M. Zhang, “Research on How Human Intelligence, Consciousness, and Cognitive Computing Affect the Development of Artificial Intelligence,” *Complexity*, 2020, doi: 10.1155/2020/1680845.