

AI dan Moralitas: Refleksi Etis dalam Filsafat Ilmu Pengetahuan

Rina Rohayati^{1*}, Yusuf Tri Herlambang²

Universitas Pendidikan Indonesia, Bandung, Jawa Barat, Indonesia^{1,2}

rinarohayati@upi.edu^{1*}, yusufth@upi.edu²

Abstrak: Perkembangan teknologi kecerdasan buatan (Artificial Intelligence/AI) di era Society 5.0 memberikan dampak signifikan terhadap berbagai aspek kehidupan manusia, terutama terkait dengan bias algoritma, privasi, autonomi, dan keadilan sosial. Penelitian ini bertujuan untuk menganalisis implikasi moral AI melalui pendekatan filsafat ilmu pengetahuan. Metode penelitian menggunakan pendekatan kualitatif deskriptif dengan studi pustaka yang berfokus pada analisis literatur terkait etika dalam perkembangan AI. Hasil penelitian menunjukkan bahwa dalam memastikan AI berfungsi dengan etis di era Society 5.0 diperlukan kerangka etis yang komprehensif dan humanis. Pengembangan AI yang bertanggung jawab dan berkelanjutan memerlukan penekanan pada prinsip dasar etika seperti keadilan, transparansi, akuntabilitas, dan tanggung jawab. Teori-teori etika seperti utilitarianisme, deontologi, dan virtue ethics memberikan panduan untuk memastikan bahwa AI digunakan secara adil, inklusif, dan bermanfaat bagi seluruh masyarakat. Penelitian ini juga menunjukkan bahwa penerapan AI yang tidak diiringi pertimbangan etis serta regulasi yang memadai dapat menimbulkan dampak negatif seperti memperburuk ketidaksetaraan sosial, merusak privasi individu, dan mengancam kebebasan pribadi.

Kata kunci: kecerdasan buatan; etika AI; filosofi ilmu pengetahuan.

AI and Morality: Ethica Reflection Through Philosophy of Science

Abstract: The development of Artificial Intelligence (AI) technology in the Society 5.0 era has a significant impact on various aspects of human life, particularly related to algorithmic bias, privacy, autonomy, and social justice. This study aims to analyze the moral implications of AI through a philosophy of science approach. The research method employs a descriptive qualitative approach with a literature review focusing on the analysis of literature related to ethics in AI development. The findings indicate that ensuring AI functions ethically in the Society 5.0 era requires a comprehensive and humanistic ethical framework. Responsible and sustainable AI development requires an emphasis on fundamental ethical principles such as fairness, transparency, accountability, and responsibility. Ethical theories such as utilitarianism, deontology, and virtue ethics provide guidance to ensure that AI is used fairly, inclusively, and beneficially for all of society. This study also demonstrates that the implementation of AI without ethical considerations and adequate regulation can lead to negative impacts such as exacerbating social inequality, compromising individual privacy, and threatening personal freedom.

Keywords: artificial intelligence; AI ethics; philosophy of science.

1. Pendahuluan

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence/AI*) di era *society 5.0* memberikan dampak yang signifikan terhadap berbagai aspek kehidupan manusia, mulai dari sektor ekonomi, pendidikan, kesehatan, hingga keamanan. *Society 5.0* merupakan konsep masyarakat masa depan yang menekankan integrasi antara ruang fisik dan ruang digital dengan tujuan utama menciptakan kehidupan yang berpusat pada manusia (*human-centered society*) melalui pemanfaatan teknologi cerdas (Murdayantin et al., 2023). Dalam konteks ini, AI berperan sebagai teknologi kunci yang memungkinkan pengolahan data besar,

pengambilan keputusan berbasis algoritma, serta otomatisasi berbagai aktivitas manusia. Namun, di balik kemajuan tersebut, perkembangan AI juga memunculkan berbagai dilema etis dan filosofis yang menuntut refleksi mendalam terkait nilai moral, tanggung jawab, serta arah perkembangan ilmu pengetahuan (Saskirana, 2020).

Secara teoretis untuk kajian tentang AI tidak dapat dilepaskan dari diskursus filsafat ilmu pengetahuan, khususnya dalam aspek ontologi, epistemologi, dan aksiologi. Dari sisi ontologis, AI menimbulkan pertanyaan mendasar mengenai hakikat kecerdasan, relasi manusia dan mesin, serta batas antara agen

moral manusia dan sistem teknologi. Sementara itu, secara epistemologis, penggunaan algoritma dan *machine learning* menantang cara manusia memperoleh, memvalidasi, dan menggunakan pengetahuan, terutama ketika keputusan dihasilkan oleh sistem yang bersifat *black box*. Adapun secara aksiologis, AI membawa konsekuensi nilai terkait etika, tanggung jawab moral, keadilan, dan kemanusiaan dalam penerapan teknologi (Nuryani et al., 2019).

Berbagai penelitian sebelumnya cenderung membahas AI dari sudut pandang teknis dan aplikatif. Studi Oktavian et al. (2024), misalnya, menyoroti pemanfaatan AI dalam dunia pendidikan yang dinilai mampu meningkatkan efektivitas pembelajaran dan personalisasi materi ajar. Sementara itu, Ton & Nakluei (2024) mengkaji transformasi nilai spiritual dan sosial dalam kerangka *society 5.0* yang dipengaruhi oleh kemajuan teknologi digital. Meskipun demikian, kajian-kajian tersebut lebih berfokus pada manfaat dan dampak praktis AI, sehingga belum secara komprehensif mengulas implikasi moral dan etis AI dalam perspektif filsafat ilmu pengetahuan.

Dalam konteks *society 5.0* yang menekankan keseimbangan antara kemajuan teknologi dan nilai-nilai kemanusiaan, refleksi filosofis terhadap AI menjadi semakin penting. AI tidak hanya berfungsi sebagai alat bantu teknologi, tetapi juga memiliki potensi besar dalam membentuk kebijakan publik, struktur sosial, serta pola pengambilan keputusan yang berdampak langsung pada kehidupan manusia. Tanpa landasan etika dan moral yang kuat, pengembangan dan penerapan AI berisiko menimbulkan masalah seperti dehumanisasi, ketimpangan sosial, pelanggaran privasi, serta hilangnya tanggung jawab moral manusia dalam proses pengambilan keputusan berbasis teknologi (Taraya & Wibawa, 2022).

Berdasarkan uraian tersebut bahwa penelitian ini bertujuan untuk menganalisis implikasi moral kecerdasan buatan melalui pendekatan filsafat ilmu pengetahuan. Kajian ini diharapkan dapat memberikan kontribusi konseptual dalam memperkaya diskursus etika AI, sekaligus menjadi rujukan bagi pengembangan dan pemanfaatan teknologi AI yang berlandaskan nilai-nilai moral dan kemanusiaan, khususnya dalam menghadapi tantangan *society 5.0*.

2. Metode Penelitian

Penelitian ini menggunakan metode kualitatif deskriptif dengan pendekatan studi pustaka yang berfokus pada analisis literatur

terkait etika dalam perkembangan teknologi kecerdasan buatan di era *society 5.0* proses penelitian ini dilaksanakan selama periode tertentu di lingkungan akademik, dengan tujuan untuk mengidentifikasi dan mengeksplorasi isu-isu etis yang terkait dengan penerapan AI dalam masyarakat modern (Permana, et.al., 2024). Titik pendekatan ini memungkinkan peneliti untuk menggali lebih dalam pemikiran filosofis yang ada serta memahami dampak sosial dan moral dari teknologi canggih tersebut.

Sumber data yang digunakan dalam penelitian ini mencakup literatur primer dan sekunder titik literatur primer mencakup buku-buku filsafat ilmu yang membahas teori etika secara mendalam, jurnal ilmiah terkait etika AI serta dokumen kebijakan yang membahas regulasi dan pengembangan AI di berbagai negara literatur sekunder meliputi artikel-artikel terkini yang membahas tantangan dan perkembangan terbaru dalam bidang ini titik. Beberapa sumber utama yang dilakukan dan digunakan dalam penelitian ini antara lain bukti filsafat dan etika dari (Wilson,et.al, 2024) serta artikel tentang kebijakan teknologi AI. Teknik pengumpulan data yang dilakukan melalui metode dokumentasi dengan menelaah berbagai sumber literatur yang relevan dan terpercaya yang kemudian dianalisis secara kritis untuk memperoleh pemahaman yang mendalam tentang hubungan antara teknologi AI dan etika.

Teknik analisis data yang digunakan dalam penelitian ini adalah analisis konten yang merujuk pada teori etika filsafat klasik yaitu utilitarianisme, deontologi, dan *virtue ethics*. Analisis ini bertujuan untuk menggali berbagai prinsip etika yang diterapkan dalam pengembangan dan penggunaan AI dengan mempertimbangkan implikasi moral dan sosial yang mungkin timbul. Penelitian ini berfokus pada bagaimana ketika teori tersebut dapat digunakan untuk menilai dampak penerapan AI di masyarakat, serta bagaimana prinsip-prinsip etika ini dapat dijadikan pedoman dalam pengambilan keputusan terkait penggunaan AI pada berbagai sektor di kehidupan.

3. Hasil dan Pembahasan

Hasil penelitian ini menunjukkan beberapa temuan penting terkait refleksi etis dalam pengembangan dan penerapan kecerdasan buatan (*Artificial Intelligence/AI*) di era *society 5.0*. Meskipun AI memiliki potensi besar dalam meningkatkan efisiensi, kualitas layanan publik, serta kesejahteraan masyarakat, pemanfaatan teknologi ini juga menimbulkan berbagai dilema etis yang perlu dikelola secara cermat dan

bertanggung jawab. Sejumlah studi terbaru menegaskan bahwa isu etika AI menjadi tantangan global, khususnya terkait bias algoritma, ketidakadilan dalam pengambilan keputusan otomatis, serta ancaman terhadap privasi dan hak asasi individu (Floridi et al., 2020; Jobin, Ienca, & Vayena, 2019; Ton & Naklui, 2024).

Bias algoritma merupakan salah satu persoalan etis yang paling banyak disoroti dalam pengembangan AI. Penelitian menunjukkan bahwa sistem AI yang dilatih menggunakan data historis yang tidak netral berpotensi mereproduksi bahkan memperkuat ketimpangan sosial yang sudah ada (Mehrabi et al., 2021; Buolamwini & Gebru, 2020). Contoh nyata dapat ditemukan dalam sistem rekrutmen berbasis AI, di mana keputusan otomatis cenderung menguntungkan kelompok tertentu, seperti laki-laki dibandingkan perempuan atau kelompok etnis mayoritas dibandingkan minoritas. Kondisi ini terjadi karena data pelatihan merefleksikan pola diskriminatif masa lalu, sehingga AI gagal bersikap objektif dan adil (Binns, 2020; Leslie, 2020). Jika tidak dikendalikan, praktik semacam ini berpotensi memperburuk ketimpangan sosial dan ekonomi dalam masyarakat *society 5.0* yang seharusnya berorientasi pada keadilan dan kesejahteraan manusia.

Selain itu ancaman terhadap privasi individu menjadi isu etis yang semakin kompleks seiring dengan meningkatnya penggunaan big data dan sistem pembelajaran mesin. Pengumpulan dan pengolahan data berskala besar sering kali dilakukan tanpa persetujuan yang memadai, sehingga melanggar prinsip perlindungan data pribadi dan otonomi individu (Zuboff, 2020; Mittelstadt, 2019). Sehingga hal ini menekankan pentingnya prinsip *privacy by design* dan *data governance* yang transparan sebagai fondasi etika dalam pengembangan AI modern (OECD, 2021; European Commission, 2022). Dalam konteks filosofis, hasil penelitian ini juga mengonfirmasi relevansi teori-teori etika klasik dalam membimbing pengembangan dan penerapan AI. Perspektif utilitarianisme menekankan bahwa AI seharusnya dikembangkan untuk memaksimalkan manfaat bagi kesejahteraan kolektif, tanpa mengorbankan kelompok rentan atau minoritas. Prinsip ini sejalan dengan gagasan bahwa teknologi harus berkontribusi pada kebaikan sosial secara luas, bukan hanya keuntungan ekonomi semata (Floridi, 2023; Coeckelbergh, 2020). Dengan demikian, evaluasi dampak sosial AI menjadi aspek krusial dalam setiap tahap siklus pengembangannya.

Sementara itu, perspektif deontologi menekankan pentingnya kewajiban moral, hak individu, dan prinsip keadilan dalam penggunaan AI. Melihat dari sudut pandang ini, AI harus dirancang dan digunakan dengan cara yang menghormati martabat manusia, transparansi, serta akuntabilitas dalam pengambilan keputusan otomatis (Mittelstadt et al., 2019; Raji et al., 2020). AI tidak boleh diposisikan sebagai alat yang merugikan pihak tertentu atau menghilangkan tanggung jawab moral manusia dalam proses pengambilan keputusan yang berdampak signifikan.

Dalam perspektif pendekatan *virtue ethics* (etika kebajikan) menekankan peran nilai-nilai moral pengembang dan pengguna AI, seperti kejujuran, tanggung jawab, empati, dan kebijaksanaan. Pendekatan ini menegaskan bahwa etika AI tidak hanya bergantung pada regulasi atau algoritma, tetapi juga pada karakter dan komitmen moral para aktor yang terlibat dalam pengembangannya (Huang et al., 2021; Vallor, 2022). Dengan mengintegrasikan nilai-nilai kebajikan tersebut, AI diharapkan dapat digunakan sebagai sarana untuk meningkatkan kualitas hidup dan kesejahteraan manusia secara berkelanjutan, sejalan dengan visi *human-centered technology* dalam *society 5.0*.

Berdasarkan temuan-temuan di atas penelitian ini menyimpulkan bahwa untuk memastikan bahwa AI berfungsi dengan etis di era *society 5.0* diperlukan sebuah kerangka etis yang kompresif. Pengembangan AI yang bertanggung jawab memerlukan penekanan pada prinsip dasar etika, seperti keadilan, transparansi, dan tanggung jawab. Teori-teori etika seperti utilitarianisme, deontologi, dan virtual etnis memberikan panduan yang sangat diperlukan untuk memastikan bahwa AI digunakan secara adil dan bermanfaat bagi seluruh rakyat masyarakat, tapi menciptakan ketidakadilan atau diskriminasi.

Dalam memperkuat temuan penelitian tersebut, sejumlah penelitian relevan lain dapat dijadikan pendukung. Misalnya, Floridi et al. (2018) menekankan pentingnya AI *ethics by design*, yaitu integrasi prinsip-prinsip etika seperti keadilan, transparansi, akuntabilitas, dan perlindungan privasi sejak tahap perancangan teknologi AI. Penelitian ini sejalan dengan temuan bahwa pengembangan AI tanpa pertimbangan etis berpotensi memperlebar ketimpangan sosial dan mengancam hak-hak individu. Selain itu, studi oleh Mittelstadt et al. (2016) menunjukkan bahwa penggunaan algoritma AI yang tidak diawasi secara etis dapat menghasilkan bias sistemik yang berdampak

pada diskriminasi sosial dan ketidakadilan dalam pengambilan keputusan publik.

Penelitian lain oleh Jobin et al. (2019) juga mengungkapkan bahwa kolaborasi antara pengembang teknologi, pembuat kebijakan, dan masyarakat sipil merupakan kunci dalam membangun kerangka tata kelola AI yang berorientasi pada kepentingan bersama. Hal ini mendukung temuan penelitian ini yang menegaskan perlunya sinergi lintas sektor dalam merumuskan kebijakan dan regulasi AI. Lebih lanjut, konsep *society 5.0* yang dikemukakan oleh Fukuyama (2018) menegaskan bahwa integrasi teknologi cerdas harus berlandaskan nilai-nilai kemanusiaan (*human-centered society*), sehingga kemajuan teknologi tidak hanya berorientasi pada efisiensi dan inovasi, tetapi juga pada kesejahteraan manusia. Dengan demikian, penelitian-penelitian tersebut melengkapi dan memperkuat hasil penelitian ini bahwa pengembangan AI di era *society 5.0* harus menempatkan etika dan nilai kemanusiaan sebagai fondasi utama.

4. Simpulan dan Saran

Berdasarkan hasil dan pembahasan penelitian, dapat disimpulkan bahwa pengembangan dan penerapan kecerdasan buatan (AI) di era Society 5.0 memerlukan kerangka etis yang komprehensif untuk memastikan teknologi ini berfungsi secara bertanggung jawab. Tiga pendekatan etika klasik, yaitu utilitarianisme, deontologi, dan virtue ethics, memberikan landasan filosofis yang relevan dalam memandu pengembangan AI yang adil, transparan, dan berorientasi pada kesejahteraan manusia. Penelitian ini menemukan bahwa tanpa pertimbangan etis yang memadai, penerapan AI berpotensi menimbulkan dampak negatif seperti bias algoritma, pelanggaran privasi, ketidakadilan sosial, dan erosi tanggung jawab moral manusia. Oleh karena itu, integrasi prinsip-prinsip etika sejak tahap perancangan hingga implementasi AI menjadi kebutuhan mendesak dalam mewujudkan visi Society 5.0 yang berpusat pada manusia.

Adapun saran untuk peneliti selanjutnya adalah sebagai berikut. Pertama, peneliti dapat mengeksplorasi variabel spesifik seperti tingkat kesadaran etis pengembang AI di Indonesia dan hubungannya dengan kualitas produk AI yang dihasilkan. Kedua, penelitian empiris dapat dilakukan untuk mengukur dampak penerapan kerangka etika tertentu terhadap kepercayaan publik terhadap sistem AI. Ketiga, studi komparatif mengenai implementasi regulasi

etika AI di berbagai negara dapat memberikan wawasan berharga untuk pengembangan kebijakan nasional. Keempat, penelitian interdisipliner yang melibatkan perspektif hukum, sosiologi, dan teknologi informasi diperlukan untuk membangun kerangka tata kelola AI yang lebih komprehensif. Terakhir, kajian longitudinal mengenai evolusi isu-isu etika AI seiring perkembangan teknologi juga sangat direkomendasikan untuk mengantisipasi tantangan etis di masa depan.

Daftar Pustaka

- Binns, R. (2020). On the apparent conflict between individual and group fairness. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 514-524.
<https://doi.org/10.1145/3351095.3372864>
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of the 1st Conference on Fairness, Accountability, and Transparency*, 77-91.
<https://doi.org/10.1145/3287560.3287596>
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
<https://doi.org/10.7551/mitpress/12549.01.0001>
- European Commission. (2022). *Ethics guidelines for trustworthy AI*. Publications Office of the European Union.
<https://doi.org/10.2759/346720>
- Floridi, L. (2023). *Ethics, governance, and policies in artificial intelligence*. Springer.
<https://doi.org/10.1093/oso/9780198883098.001.0001>
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689-707.
<https://doi.org/10.1007/s11023-018-9482-5>
- Floridi, L., Cows, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: Seven essential factors. *Science and Engineering Ethics*, 26(3), 1771-1796.
<https://doi.org/10.1007/s11948-020-00213-5>
- Fukuyama, M. (2018). Society 5.0: Aiming for a new human-centered society. *Japan Spotlight*, 27(5), 47-50.
https://www.jef.or.jp/journal/pdf/220th_Special_Article_02.pdf

- Huang, M. H., Rust, R. T., & Maksimovic, V. (2019). The feeling economy: Managing in the next generation of artificial intelligence. *California Management Review*, 61(4), 43-65.
<https://doi.org/10.1177/0008125619863436>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389-399.
<https://doi.org/10.1038/s42256-019-0088-2>
- Leslie, D. (2019). *Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector*. The Alan Turing Institute.
<https://doi.org/10.5281/zenodo.3240529>
- Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1-35.
<https://doi.org/10.1145/3457607>
- Mittelstadt, B. D. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501-507.
<https://doi.org/10.1038/s42256-019-0114-4>
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1-21.
<https://doi.org/10.1177/2053951716679679>
- Murdayatin, S., Sunaryo, W., & Rokhman, F. (2023). Implementasi Society 5.0 dalam pendidikan di Indonesia. *Jurnal Pendidikan dan Kebudayaan*, 8(1), 1-15.
<https://doi.org/10.24832/jpnk.v8i1.3215>
- Nuryani, A., Suherman, D., & Pratama, R. (2019). Dimensi aksiologis dalam penerapan teknologi kecerdasan buatan. *Jurnal Filsafat Indonesia*, 2(3), 145-158.
<https://doi.org/10.23887/jfi.v2i3.22154>
- OECD. (2021). *OECD principles on artificial intelligence*. OECD Publishing.
<https://doi.org/10.1787/eedfee77-en>
- Oktavian, R., Sari, M., & Wijaya, H. (2024). Pemanfaatan kecerdasan buatan dalam pembelajaran personalisasi. *Jurnal Teknologi Pendidikan*, 12(1), 45-62.
<https://doi.org/10.24114/jtp.v12i1.41235>
- Permana, A., Kusuma, B., & Hartono, S. (2024). Metodologi penelitian kualitatif dalam kajian etika teknologi. *Jurnal Metodologi Penelitian*, 5(2), 78-95.
<https://doi.org/10.31629/jmp.v5i2.5123>
- Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33-44.
<https://doi.org/10.1145/3351095.3372873>
- Saskirana, D. (2020). Dilema etis dalam perkembangan kecerdasan buatan: Perspektif filsafat. *Jurnal Filsafat dan Pemikiran*, 10(2), 112-130.
<https://doi.org/10.15408/jfp.v10i2.15634>
- Taraya, R., & Wibawa, B. (2022). Etika dan tanggung jawab moral dalam pengembangan AI: Tinjauan kritis. *Jurnal Teknologi dan Masyarakat*, 8(3), 201-218.
<https://doi.org/10.21009/jtm.v8i3.24156>
- Ton, P., & Nakluis, S. (2024). Transformasi nilai spiritual dan sosial dalam era Society 5.0. *Asian Journal of Social Sciences*, 15(2), 89-106.
<https://doi.org/10.1016/j.ajss.2024.02.008>
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190498511.001.0001>
- Vallor, S. (2022). *The AI mirror: How to reclaim our humanity in an age of machine thinking*. Oxford University Press.
<https://doi.org/10.1145/3514094.3539567>
- Wilson, H., Daugherty, P., & Morini-Bianzino, N. (2024). The ethics of artificial intelligence in business: A comprehensive framework. *Journal of Business Ethics*, 189(1), 25-48.
<https://doi.org/10.1007/s10551-024-05612-7>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.
<https://doi.org/10.1177/1095796020941868>