

Epistemologi Artificial Intelligence : Kebenaran, Validitas, dan Otoritas Algoritmik

Sri Nurhayati¹, Diana Effendi², Agus Nursikuwagus³, Usep Mohamad Ishaq⁴, Andrias Darmayadi⁵

¹²³⁴⁵Universitas Komputer Indonesia; Indonesia

Correspondence e-mail*, sri.nurhayati@email.unikom.ac.id¹,

diana.effendi@email.unikom.ac.id², agusnursikuwagus@email.unikom.ac.id³,

usep.mohamad.ishaq@email.unikom.ac.id⁴, andrias.darmayadi@email.unikom.ac.id⁵

Submitted: 2025/11/27

Revised: 2025/12/04;

Accepted: 2025/12/05;

Published: 2025/12/05

Abstract

The development of artificial intelligence (AI) has brought fundamental changes in the way knowledge is produced, validated, and accepted in various sectors of life. Algorithmic models, especially deep learning, generate predictions and recommendations that are often treated as operational truths even though the inference process is not fully explainable. This study analyzes how AI changes the understanding of truth, validity, and epistemic authority from the perspective of the philosophy of science, and links it to the ontological and axiological dimensions in modern knowledge production. A qualitative approach based on philosophical analysis is used to integrate the thoughts of Popper, Kuhn, Lakatos, Van Fraassen, and Floridi. The results show that AI shifts knowledge from rational justification to performative and statistical validity, and challenges the position of humans as the primary epistemic agents. This study asserts that the epistemic transformation triggered by AI requires ontological and axiological reflection so that the development of knowledge remains in line with humanitarian principles and ethical responsibility.

Keywords

AI epistemology; Algorithmic truth; Black box models; Algorithmic authority; Validity statistics



© 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY SA) license, <https://creativecommons.org/licenses/by-sa/4.0/>.

PENDAHULUAN

Perkembangan kecerdasan buatan (AI) telah memasuki hampir seluruh aspek kehidupan manusia dan secara fundamental mengubah cara individu, organisasi, dan masyarakat memproduksi serta memvalidasi pengetahuan. Dalam berbagai studi, seperti yang dibahas oleh Floridi¹ dan Müller², data tidak lagi hanya dianalisis oleh manusia, melainkan diproses oleh

¹ Floridi, L., & Cowls, J. (2019). *A Unified Framework of Five Principles for AI in Society*. *Harvard Data Science Review*, 1(1), 1–15., n.d.

² Müller, V. C. (2021). *Ethics of Artificial Intelligence and Robotics*. In *The Stanford Encyclopedia of Philosophy*. Stanford University., n.d.

algoritma yang menghasilkan prediksi dan rekomendasi yang kemudian berfungsi sebagai bentuk *kebenaran operasional*. Pergeseran ini menunjukkan bahwa otoritas epistemik yang sebelumnya berasal dari kapasitas berpikir manusia kini mulai beralih kepada sistem komputasional yang bekerja melalui kalkulasi probabilitas dan pembelajaran pola dalam skala besar. Sejumlah kajian mutakhir menunjukkan bahwa epistemologi AI tidak lagi bisa dipisahkan dari praktik sosialnya: AI mulai diperlakukan sebagai mitra epistemik yang ikut membentuk proses produksi pengetahuan, bukan sekadar alat komputasi netral³. Di level yang lebih luas, pergeseran menuju *data-driven science* dibaca sebagai perubahan paradigma epistemologis baru, ketika korelasi statistik berskala besar dan otomatisasi inferensi menggantikan peran teori eksplisit dalam banyak praktik ilmiah⁴. Kondisi ini menjadikan AI bukan hanya persoalan teknologis, melainkan juga isu ontologis dan aksiologis yang menentukan arah masa depan pengetahuan.

Epistemologi klasik menempatkan pengetahuan sebagai hasil pembenaran rasional, pengalaman empiris, serta refleksi kritis. Namun epistemologi AI, sebagaimana dikritisi oleh Burrell⁵ dan Mittelstadt⁶, bertumpu pada korelasi statistik dan struktur data tanpa harus memahami hubungan kausal yang mendasarinya. Ketika keputusan medis, keuangan, pemerintahan, keamanan, hingga pendidikan mulai mengandalkan rekomendasi algoritmik, muncul pertanyaan mendasar: siapakah sebenarnya yang “mengetahui” — manusia yang menggunakan output tersebut, atau algoritma yang memproduksinya?.

Pertanyaan ini semakin relevan mengingat banyak model AI, terutama *deep learning*, bersifat *black box*. Penelitian oleh Lipton⁷ dan Rudin⁸ menunjukkan bahwa meskipun model-model tersebut menghasilkan performa prediksi yang tinggi, proses inferensinya sulit dijelaskan. Fenomena ini melahirkan paradoks epistemologis: pada saat pemahaman manusia terhadap proses penalaran algoritmik menurun, kepercayaan terhadap hasil algoritma justru meningkat. Hal ini menciptakan kondisi di mana otoritas algoritmik dapat menggantikan otoritas pengetahuan manusia,

³ Simbolon, L., Manugeran, M., & Barus, E. (2025). Does AI Know Things? An Epistemological Perspective on Artificial Intelligence. *Journal of English Language and Education*, 10(5), n.d.

⁴ Balazka, D., & Rodighiero, D. (2020). Big Data and the Little Big Bang: An Epistemological (r)Evolution. *Frontiers in Big Data*, 3, 31, n.d.

⁵ Burrell, J. (2016). How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms. *Big Data & Society*, 3(1), 1–12., n.d.

⁶ Mittelstadt, B. D., Russell, C., & Wachter, S. (2019). Explaining Explanations in AI. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT)*, 279–288., n.d.

⁷ Lipton, Z. C. (2018). The Mythos of Model Interpretability. *Queue*, 16(3), 1–27., n.d.

⁸ Rudin, C. (2019). Stop Explaining Black Box Machine Learning Models for High-Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*, 1(5), 206–215., n.d.

sebagaimana dikatakan oleh O'Neil⁹ dan Ananny & Crawford¹⁰.

Di tengah percepatan teknologi, epistemologi klasik yang bertumpu pada korespondensi, penjelasan kausal, dan justifikasi rasional mengalami ketegangan ketika berhadapan dengan model-model AI yang bekerja melalui statistik, korelasi, dan pembelajaran non-kausal. Sistem algoritmik tidak menghasilkan pengetahuan melalui pemahaman, tetapi melalui performa prediktif yang stabil. Hal ini menggeser fokus epistemik dari “mengapa sesuatu benar” menjadi “apakah ia bekerja secara konsisten”, sehingga membuka ruang baru bagi bentuk kebenaran operasional yang tidak lagi berada dalam kerangka filsafat pengetahuan tradisional. Meskipun berbagai penelitian telah membahas isu etika, bias, dan implikasi sosial dari algoritma, kajian yang secara khusus menelaah bagaimana AI mendefinisikan ulang konsep kebenaran, validitas, dan otoritas epistemik dalam kerangka filsafat ilmu masih terbatas. Gap inilah yang menjadi dasar penelitian ini: belum adanya kerangka epistemologis terpadu yang menghubungkan dimensi ontologis, epistemologis, dan aksiologis dalam memahami pembentukan pengetahuan algoritmik.

Penelitian ini memiliki tiga tujuan utama: (1) menjelaskan bagaimana struktur pengetahuan bergeser dari model kausal-rasional menuju model prediktif-performatif, (2) menguraikan bagaimana validitas algoritmik terbentuk dan apa implikasinya terhadap tradisi epistemologi klasik, dan (3) menganalisis bagaimana otoritas algoritmik muncul serta bagaimana hal ini menggeser posisi manusia sebagai agen epistemik.

Berdasarkan tujuan tersebut, rumusan masalah dalam penelitian ini adalah: (1) bagaimana AI mengubah landasan epistemik dalam memproduksi dan memvalidasi pengetahuan, (2) bagaimana konsep validitas epistemik dipahami ulang dalam konteks model black box, dan (3) bagaimana otoritas algoritmik terbentuk serta apa implikasinya terhadap otonomi pengetahuan manusia.

METODE

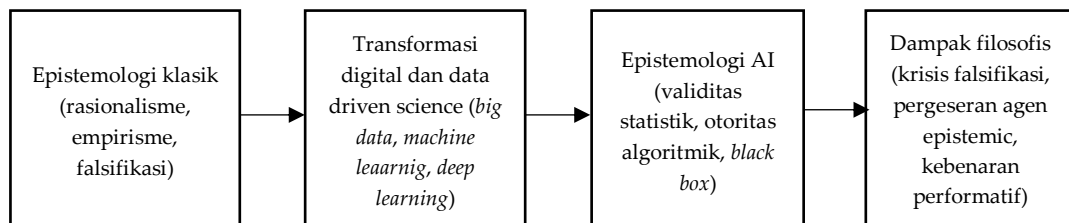
Penelitian ini menggunakan pendekatan kualitatif dengan metode analisis filosofis berbasis kajian literatur untuk memahami bagaimana kecerdasan buatan (AI) membentuk ulang konsep kebenaran, validitas, dan otoritas pengetahuan. Analisis filosofis dipilih karena sesuai untuk memeriksa struktur konseptual dan argumen normatif dalam fenomena epistemik kontemporer. Pendekatan ini tidak bertujuan menguji hipotesis empiris, melainkan mengevaluasi secara kritis

⁹ O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown., n.d.

¹⁰ Ananny, M., & Crawford, K. (2018). *Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*. *New Media & Society*, 20(3), 973–989., n.d.

kerangka epistemik yang mendasari sistem algoritmik. Gambar 1 menggambarkan pergeseran dari epistemologi klasik yang bertumpu pada inferensi kausal menuju epistemologi algoritmik berbasis korelasi dan *performativitas statistic*.

Gambar 1. Kerangka epistemologi AI



Validitas kajian literatur dijaga melalui proses seleksi yang ketat berdasarkan tiga kriteria:

(1) relevansi langsung dengan isu epistemologi AI seperti kebenaran algoritmik, *explainability*, dan otoritas algoritmik, (2) kredibilitas sumber yang berasal dari publikasi akademik bereputasi seperti Scopus, IEEE Xplore, Springer, dan Stanford Encyclopedia of Philosophy, dan (3) periode terbit sepuluh tahun terakhir, kecuali karya filsafat klasik yang tetap digunakan sebagai landasan teoretik. Proses penelusuran literatur mengikuti standar kajian literatur filosofis yang menekankan kelengkapan argumentatif dan koherensi konseptual ¹¹, melalui basis data Scopus, ScienceDirect, SpringerLink, JSTOR, IEEE Xplore, dan arXiv dengan kata kunci seperti "*epistemology of AI*," "*algorithmic authority*," dan "*data-driven epistemology*." Untuk memperkuat kedalaman analisis, penelitian ini menerapkan triangulasi teoretik dengan memeriksa konsistensi lintas lima tradisi epistemologi utama Popper, Kuhn, Lakatos, Van Fraassen, dan Floridi. Pendekatan triangulasi konseptual ini sejalan dengan gagasan pluralisme epistemik yang menekankan pentingnya pemeriksaan lintas-kerangka dalam filsafat ilmu ^{12 13 14}. Analisis dilakukan menggunakan *conceptual analysis* untuk membedah konsep-konsep seperti kebenaran algoritmik, validitas prediktif, dan opasitas model, serta *philosophical argument analysis* untuk menilai koherensi internal dan implikasi epistemologis dari berbagai argumen yang muncul dalam literatur primer dan sekunder. Penelitian ini memiliki batasan metodologis karena bersifat konseptual tanpa verifikasi empiris terhadap model AI tertentu, sehingga temuan yang dihasilkan terutama bersifat teoretis dan dapat dilengkapi melalui studi empiris atau studi kasus pada penelitian berikutnya

¹¹ Boersema, David. 2008. *Philosophy of Science: Text with Readings*. Boston: Pearson Education. (n.d.).

¹² Longino, Helen E. 2002. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press. (n.d.).

¹³ Kellert, Stephen H., Helen E. Longino, and C. Kenneth Waters, Eds. 2006. *Scientific Pluralism*.

Minneapolis: University of Minnesota Press. (n.d.).

¹⁴ Chang, Hasok. 2012. *Is Water H₂O? Evidence, Realism and Pluralism*. Dordrecht: Springer. (n.d.).

HASIL DAN PEMBAHASAN

Hasil

Berikut hasil penelitian mengenai perubahan struktur pengetahuan di era algoritmik, validitas algoritmik dan krisis transparansi, otoritas algoritmik serta pergeseran agen epistemik, dan tantangan epistemologis yang muncul dari penggunaan kecerdasan buatan.

1. Perubahan Struktur Pengetahuan di Era Algoritmik

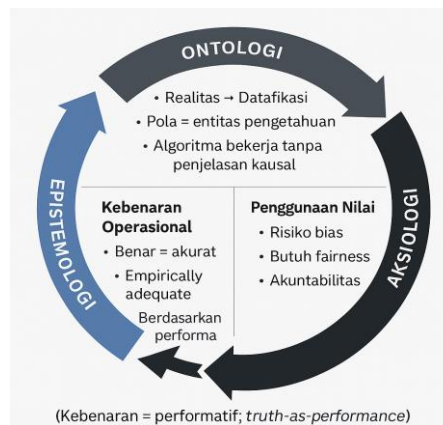
Kehadiran AI telah menggeser struktur produksi pengetahuan dari model berbasis penalaran kausal menuju model berbasis korelasi dan prediksi. Dalam epistemologi klasik, pengetahuan umumnya ditautkan dengan *justified true belief*, yang menuntut adanya alasan rasional dan hubungan kausal yang dapat dijelaskan. Namun, dalam praktik AI, terutama pada model *deep learning*, hubungan tersebut tidak selalu tersedia. Lipton dan Rudin mencatat bahwa performa prediktif menjadi dasar legitimasi epistemik, bukan kemampuan menjelaskan bagaimana suatu kesimpulan diperoleh^{15 16}. Temuan literatur juga menegaskan munculnya bentuk “kebenaran operasional”, yakni kebenaran yang diakui sejauh model memberikan output yang akurat, stabil, dan berfungsi dalam konteks penggunaannya.

Gambar 2 berikut diagram ini menggambarkan proses pembentukan kebenaran dalam sistem algoritmik melalui tiga lapisan konseptual: ontologi, epistemologi, dan aksiologi. Pada tingkat ontologis, AI memandang realitas sebagai data yang didigitalkan dan diproses sebagai pola statistik tanpa pemahaman kausal. Secara epistemologis, kebenaran ditentukan oleh performa model dalam menghasilkan prediksi yang akurat dan *empirically adequate*. Pada ranah aksiologis, penggunaan kebenaran algoritmik harus diselaraskan dengan nilai fairness, akuntabilitas, dan mitigasi bias untuk memastikan bahwa kebenaran operasional tidak bertentangan dengan prinsip etis dalam pengambilan keputusan. Diagram ini menegaskan bahwa kebenaran dalam AI bersifat performatif (*truth-as-performance*), lahir dari keberhasilan algoritma menjalankan fungsi prediktifnya.

¹⁵ Lipton, Z. C. (2018). *The Mythos of Model Interpretability*. *Queue*, 16(3), 1–27.

¹⁶ Rudin, C. (2019). *Stop Explaining Black Box Machine Learning Models for High-Stakes Decisions and Use Interpretable Models Instead*. *Nature Machine Intelligence*, 1(5), 206–215.

Gambar 2. Alur Kebenaran Algoritmik: Relasi Ontologi, Epistemologi, dan Aksiologi

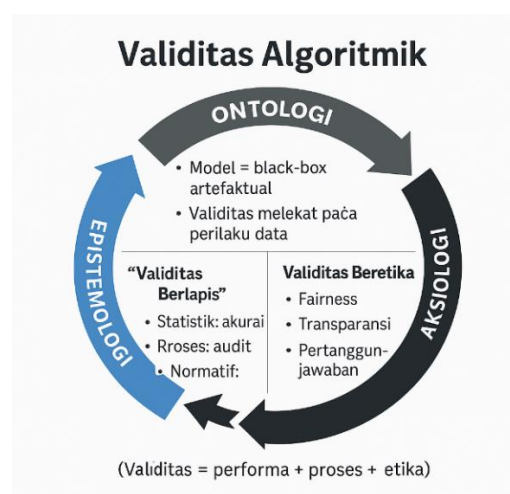


2. Validitas Algoritmik dan Krisis Transparansi

Validitas algoritmik muncul sebagai bentuk baru validitas epistemik ketika model-model kompleks tidak menyediakan struktur inferensial yang dapat dilacak. Burrell menyebut kondisi ini sebagai opaque rationality. Model *deep learning* tidak memungkinkan falsifikasi Popper, sehingga validitas prediktif menggantikan validitas penjelasan. Temuan ini mengindikasikan bahwa validitas dalam AI tidak hanya bersifat statistik, tetapi juga procedural bergantung pada kualitas data, proses pelatihan, dan keterlacakan keputusan model.

Gambar 3 Diagram menunjukkan bahwa validitas teknis (akurasi), validitas proses (transparansi dan audit), dan validitas normatif (kelayakan penggunaan) membentuk satu kesatuan. Validitas algoritmik tidak hanya dinilai dari output, tetapi dari keseluruhan proses epistemik yang melibatkan data, model, dan tata kelola.

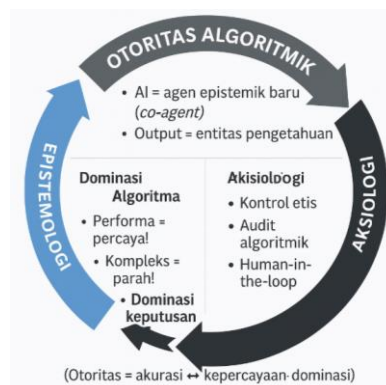
Gambar 3. Alur Validitas Algoritmik: Relasi Ontologi, Epistemologi, dan Aksiologi



3. Otoritas Algoritmik dan Pergeseran Agen Epistemik

Hasil kajian memperlihatkan bahwa otoritas algoritmik muncul ketika manusia lebih mempercayai output model dibanding penilaian sendiri. O'Neil menyatakan bahaya ketika algoritma diterima tanpa evaluasi kritis¹⁷. Coeckelbergh menyoroti bahwa AI berpotensi mengambil alih fungsi epistemik manusia¹⁸. Temuan literatur menunjukkan bahwa dominasi algoritma tidak hanya terjadi karena akurasi model, tetapi karena posisi strukturalnya dalam rantai pengambilan keputusan. Diagram pada gambar 4 menggambarkan bagaimana otoritas algoritmik terbentuk melalui keterkaitan antara aspek ontologis, epistemologis, dan aksiologis dalam sistem kecerdasan buatan. Secara ontologis, AI diposisikan sebagai agen epistemik baru (*co-agent*) yang menghasilkan output yang diperlakukan sebagai entitas pengetahuan, sehingga sebagian fungsi penalaran epistemik manusia mulai didistribusikan kepada sistem komputasional.

Gambar 4. Alur Otoritas Algoritmik: Relasi Ontologi, Epistemologi, dan Aksiologi



Pada tingkat epistemologi, dominasi algoritma muncul melalui mekanisme performatif: performa yang tinggi menumbuhkan kepercayaan, kompleksitas teknis mendorong ketergantungan, dan pada akhirnya algoritma berperan menentukan keputusan tanpa dapat sepenuhnya dibantah oleh manusia. Sementara itu, dimensi aksiologis menekankan bahwa otoritas tersebut harus diimbangi oleh mekanisme kontrol etis seperti audit algoritmik, prinsip *human-in-the-loop*, serta hak pengguna untuk menolak rekomendasi mesin. Integrasi ketiga dimensi ini menunjukkan bahwa otoritas algoritmik bukan sekadar hasil akurasi, tetapi konsekuensi dari interaksi antara performa teknis, struktur kepercayaan, dan desain tata kelola yang menentukan sejauh mana mesin boleh mengambil alih proses penilaian manusia.

¹⁷ O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

¹⁸ Coeckelbergh, M. (2022). *Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence*. *AI and Ethics*, n.d.

4. Tantangan Epistemologis dan Implikasi Masa Depan Pengetahuan

Gabungan antara kebenaran operasional, validitas statistik, dan otoritas algoritmik menimbulkan beberapa konsekuensi filosofis. Pertama, manusia kehilangan otonomi epistemik ketika keputusan penting bergantung pada sistem yang tidak dipahami. Riset empiris tentang automation bias menunjukkan bahwa ketika sistem AI diposisikan sebagai pemberi rekomendasi utama, pengguna cenderung menerima keluarannya secara tidak kritis dan mengabaikan bukti lain, sehingga kemampuan penalaran dan refleksi epistemik manusia justru melemah¹⁹. Di bidang diagnostik kesehatan, dinyatakan ketergantungan berlebihan pada alat bantu AI berisiko mengikis keterampilan klinis manusia jika tidak diimbangi dengan desain kerja yang sengaja menjaga keterlibatan aktif dokter dalam proses interpretasi dan pengambilan keputusan²⁰. Karena itu, sejumlah penulis mengusulkan pemahaman baru tentang *human-in-the-loop* yang tidak sekadar menambah *checkpoint* manusia, tetapi menjadikan manusia sebagai penjaga utama justifikasi epistemik: AI diposisikan sebagai co-agent epistemik yang harus selalu beroperasi di bawah kerangka pengawasan, penjelasan institusional, dan tanggung jawab manusia^{21 22}. Kedua, struktur pengetahuan masyarakat berubah dari sesuatu yang dapat dijustifikasi secara rasional menjadi sesuatu yang diterima secara fungsional. Ketiga, epistemologi AI membuka kemungkinan bahwa kriteria pengetahuan di masa depan tidak lagi memerlukan penjelasan, melainkan hanya keberhasilan performatif.

Namun, beberapa peneliti seperti Floridi²³ mengatakan bahwa kualitas epistemik masyarakat tidak boleh hanya bergantung pada performa algoritmik. Dibutuhkan pendekatan *explainable AI* untuk mengembalikan transparansi dan memperkuat pertanggungjawaban epistemik. Diskusi ini menunjukkan bahwa masa depan pengetahuan bukan sekadar persoalan teknis, tetapi persoalan filosofis mendalam yang menentukan bagaimana manusia memahami dunia, membuat keputusan, dan membangun kebenaran kolektif.

¹⁹ Kücking, F., Hübner, U., Przysucha, M., Hannemann, N., Kutza, J.-O., Moelleken, M., Erfurt-Berge, C., Dissemond, J., Babitsch, B., & Busch, D. (2024). Automation Bias in AI-Decision Support: Results from an Empirical Study. *Studies in Health Technology and Informatics*, 317, 298–304., n.d.

²⁰ Sunday, O. (2025). Behavioral Impacts of AI Reliance in Diagnostics: Balancing Automation with Skill Retention. *Epidemiology & Health Data Insights*, 1(3), Ehd011., n.d.

²¹ Sunday, O. (2025). Behavioral Impacts of AI Reliance in Diagnostics: Balancing Automation with Skill Retention. *Epidemiology & Health Data Insights*, 1(3), Ehd011.

²² Theunissen, M., & Browning, J. (2022). Putting Explainable AI in Context: Institutional Explanations for Medical AI. *Ethics and Information Technology*, 24(2), Article 23., n.d.

²³ Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1), 1–15.

PEMBAHASAN

Pembahasan ini menginterpretasikan temuan penelitian dalam kerangka epistemologi filsafat ilmu dan mengaitkannya dengan perkembangan epistemologi algoritmik.

1. Rekonstruksi Epistemologi AI melalui Lima Tradisi Besar

Temuan penelitian menunjukkan bahwa ada perbedaan mendasar antara cara pengetahuan klasik bekerja—yang bertumpu pada teori dan penjelasan dengan cara kerja epistemologi algoritmik yang lebih menekankan pada kinerja atau performa model. Sejumlah kajian menunjukkan bahwa model AI sering kali menghasilkan keluaran akurat tanpa memberikan struktur penjelasan yang dapat ditelusuri^{24–25}. Pergeseran ini sejalan dengan pandangan bahwa perkembangan ilmu pengetahuan digital semakin bergeser dari ilmu berbasis teori menuju pendekatan yang digerakkan oleh data dan pola korelasi²⁶. Dalam konteks ini, keberhasilan prediksi sering kali dianggap cukup sebagai dasar untuk menerima suatu model sebagai “memadai” secara epistemik, tanpa harus memahami mekanisme internalnya²⁷. Sementara itu, pendekatan epistemologi informasi menegaskan bahwa struktur pengetahuan di era AI sangat ditentukan oleh representasi data, arsitektur sistem, dan logika pemrosesan informasi digital^{28–29}. Kajian lain juga menyoroti bahwa AI membentuk relasi baru antara manusia dan mesin sebagai sesama agen epistemik, di mana sebagian proses pengetahuan dipindahkan kepada algoritma³⁰ menjelaskan bagaimana kebenaran dihasilkan, diverifikasi, dan dibatasi ketika proses epistemik tidak lagi bertumpu pada penjelasan kausal, tetapi pada stabilitas keluaran model dan konsistensi proses komputasional.

2. Pergeseran Landasan Validitas: Dari Penjelasan ke Performa

Temuan mengenai dominasi kebenaran performatif menunjukkan bahwa epistemologi AI mulai mengaburkan batas antara pengetahuan dan prediksi. Sistem algoritmik dapat menghasilkan keluaran yang diterima sebagai benar meskipun mekanisme internalnya tidak sepenuhnya dapat

²⁴ Lipton, Z. C. (2018). *The Mythos of Model Interpretability*. *Queue*, 16(3), 1–27.

²⁵ Rudin, C. (2019). *Stop Explaining Black Box Machine Learning Models for High-Stakes Decisions and Use Interpretable Models Instead*. *Nature Machine Intelligence*, 1(5), 206–215.

²⁶ Balazka, D., & Rodighiero, D. (2020). *Big Data and the Little Big Bang: An Epistemological (r)Evolution*. *Frontiers in Big Data*, 3, 31.

²⁷ Russo, F. (2022). *Connecting Ethics and Epistemology of AI*. *AI & Society*, 37(4), 1483–1495.

²⁸ Floridi, L., & Cowls, J. (2019). *A Unified Framework of Five Principles for AI in Society*. *Harvard Data Science Review*, 1(1), 1–15.

²⁹ Müller, V. C. (2021). *Ethics of Artificial Intelligence and Robotics*. In *The Stanford Encyclopedia of Philosophy*. Stanford University.

³⁰ Simbolon, L., Manugeran, M., & Barus, E. (2025). *Does AI Know Things? An Epistemological Perspective on Artificial Intelligence*. *Journal of English Language and Education*, 10(5).

dijelaskan. Dalam tradisi filsafat ilmu, kondisi ini menimbulkan ketegangan antara pendekatan yang menilai pengetahuan berdasarkan penjelasan yang dapat dipertanggungjawabkan, dan pendekatan yang menilai pengetahuan berdasarkan keberhasilan performa model.

Kondisi ini mengimplikasikan bahwa dasar penilaian validitas epistemik tidak lagi bertumpu pada alasan atau penjelasan yang dapat diuji secara rasional, tetapi lebih pada reliabilitas statistik yang dapat diukur melalui kinerja model. Hal ini menandai perubahan mendasar dalam cara justifikasi epistemik dibangun di era kecerdasan buatan.

3. Otoritas Algoritmik sebagai Krisis Epistemik

Otoritas algoritmik muncul ketika pengguna menerima keluaran algoritma tanpa kemampuan mengakses struktur inferensialnya. Kondisi ini menimbulkan apa yang dapat disebut sebagai *epistemic outsourcing*, yaitu penyerahan justifikasi epistemik kepada sistem yang tidak dapat dipertanggungjawabkan secara rasional. Fenomena ini bukan semata persoalan etika, tetapi persoalan epistemologi: siapa yang memiliki legitimasi epistemik untuk mendefinisikan apa yang benar ketika algoritma bekerja tanpa penjelasan?.

4. Kontribusi Teoritis Penelitian

Penelitian ini memberikan tiga kontribusi teoretis yaitu :

1. Kontribusi ontologis, dimana AI menciptakan entitas epistemik baru berupa model prediktif yang bekerja tanpa representasi kausal.
2. Kontribusi epistemologis , dimana AI menggeser justifikasi epistemik dari landasan penjelasan (*explanation-based validity*) ke landasan performa (*performance-based validity*).
3. Kontribusi aksiologis, dimana pengetahuan algoritmik menuntut *epistemic accountability* yang menggabungkan transparansi proses, audit algoritmik, dan mekanisme *human-in-the-loop*.

Ketiga kontribusi ini membentuk kerangka terpadu yang jarang dibahas secara simultan dalam literatur sebelumnya.

5. Implikasi Teoritis dan Praktis Penelitian

Secara teoretis, temuan penelitian ini menunjukkan bahwa epistemologi klasik tidak lagi memadai untuk menjelaskan cara kerja pengetahuan dalam sistem AI yang bergantung pada performa prediktif dan struktur model yang opak. Karena itu, dibutuhkan kerangka epistemologi baru yang mampu mengintegrasikan validitas statistik, perilaku model, serta keterbatasan transparansi algoritmik. Kerangka ini harus menjelaskan bagaimana kebenaran dihasilkan,

diverifikasi, dan dibatasi ketika proses epistemik tidak lagi bertumpu pada penjelasan kausal, tetapi pada stabilitas keluaran model dan konsistensi proses komputasional.

Secara praktis, implikasinya adalah perlunya mekanisme audit epistemik yang dapat memeriksa kualitas model tidak hanya dari sisi akurasi, tetapi juga bias, reliabilitas lintas-konteks, dan konsistensi keputusan. Pengguna dan institusi harus memiliki prosedur untuk meninjau kembali keputusan algoritma serta ruang intervensi manusia dalam kondisi kritis. Tata kelola AI perlu memastikan bahwa proses pengambilan keputusan mempertahankan otonomi epistemik manusia, terutama pada domain yang menuntut akuntabilitas tinggi seperti kesehatan, hukum, dan keuangan.

6. Arah Penelitian Selanjutnya

Penelitian selanjutnya dapat difokuskan pada :

1. Dinamika kepercayaan manusia terhadap model prediktif, yaitu bagaimana orang membangun kepercayaan pada hasil AI, apa yang membuat mereka yakin atau ragu, dan bagaimana mereka merespons ketika model memberikan hasil yang keliru.
2. Perbandingan antara validitas kausal dan validitas prediktif, untuk melihat kapan penjelasan sebab-akibat tetap dibutuhkan dan kapan akurasi prediksi sudah dianggap cukup, serta bagaimana perbedaan ini memengaruhi cara kita menilai kebenaran dalam praktik ilmiah.
3. Penyusunan kerangka pertanggungjawaban epistemik (*epistemic accountability*), yaitu bagaimana merancang aturan dan mekanisme agar AI tidak hanya akurat, tetapi juga dapat dipahami, diawasi, dan dikendalikan oleh manusia, sehingga keputusan penting tidak sepenuhnya diserahkan kepada algoritma.

KESIMPULAN

Penelitian ini menjawab tiga pertanyaan utama mengenai epistemologi AI. Pertama, AI mengubah struktur pengetahuan dari model berbasis penjelasan menuju model berbasis performa prediktif, menandai pergeseran ontologis dan epistemologis dari teori ke data. Kedua, validitas dalam AI bergeser dari penjelasan kausal menuju validitas statistik yang bersifat empirically adequate, sehingga menantang prinsip falsifikasi Popper dan memperkuat paradigma baru ala Kuhn dan Van Fraassen. Ketiga, otoritas algoritmik muncul ketika performa model mendorong ketergantungan pengguna, menggeser sebagian fungsi agen epistemik manusia ke sistem komputasional.

Secara teoretis, penelitian ini menawarkan kerangka tiga dimensi ontologi, epistemologi,

aksiologi untuk membaca ulang kebenaran algoritmik dan posisi manusia dalam struktur pengetahuan kontemporer. Secara praktis, penelitian ini menekankan perlunya mekanisme tata kelola epistemik seperti audit algoritmik, *human-in-the-loop*, dan institusional *explainability*.

Penelitian selanjutnya dapat diarahkan pada: (1) pengembangan model *epistemic accountability* untuk sistem AI, (2) kajian empiris mengenai bagaimana pengguna membangun kepercayaan terhadap rekomendasi algoritmik, dan (3) eksplorasi ontologi data dalam konteks AI generatif.

REFERENSI

- Ananny, M., & Crawford, K. (2018). Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability. *New Media & Society*, 20(3), 973–989. n.d.
- Balazka, D., & Rodighiero, D. (2020). Big Data and the Little Big Bang: An Epistemological (r)Evolution. *Frontiers in Big Data*, 3, 31. n.d.
- Boersema, David. 2008. *Philosophy of Science: Text with Readings*. Boston: Pearson Education. n.d.
- Burrell, J. (2016). How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms. *Big Data & Society*, 3(1), 1–12. n.d.
- Chang, Hasok. 2012. *Is Water H₂O? Evidence, Realism and Pluralism*. Dordrecht: Springer. n.d.
- Coeckelbergh, M. (2022). Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence. *AI and Ethics*. n.d.
- Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1), 1–15. n.d.
- Kellert, Stephen H., Helen E. Longino, and C. Kenneth Waters, Eds. 2006. *Scientific Pluralism*. Minneapolis: University of Minnesota Press. n.d.
- Kücking, F., Hübner, U., Przysucha, M., Hannemann, N., Kutza, J.-O., Moelleken, M., Erfurt-Berge, C., Dissemond, J., Babitsch, B., & Busch, D. (2024). Automation Bias in AI-Decision Support: Results from an Empirical Study. *Studies in Health Technology and Informatics*, 317, 298–304. n.d.
- Lipton, Z. C. (2018). The Mythos of Model Interpretability. *Queue*, 16(3), 1–27. n.d.
- Longino, Helen E. 2002. *The Fate of Knowledge*. Princeton, NJ: Princeton University Press. n.d.
- Mittelstadt, B. D., Russell, C., & Wachter, S. (2019). Explaining Explanations in AI. *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT)*, 279–288. n.d.
- Müller, V. C. (2021). Ethics of Artificial Intelligence and Robotics. In *The Stanford Encyclopedia of Philosophy*. Stanford University. n.d.
- O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown. n.d.
- Rudin, C. (2019). Stop Explaining Black Box Machine Learning Models for High-Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence*, 1(5), 206–215. n.d.

- Russo, F. (2022). Connecting Ethics and Epistemology of AI. *AI & Society*, 37(4), 1483–1495. n.d.
- Simbolon, L., Manugeran, M., & Barus, E. (2025). Does AI Know Things? An Epistemological Perspective on Artificial Intelligence. *Journal of English Language and Education*, 10(5). n.d.
- Sunday, O. (2025). Behavioral Impacts of AI Reliance in Diagnostics: Balancing Automation with Skill Retention. *Epidemiology & Health Data Insights*, 1(3), Ehdi011. n.d.
- Theunissen, M., & Browning, J. (2022). Putting Explainable AI in Context: Institutional Explanations for Medical AI. *Ethics and Information Technology*, 24(2), Article 23. n.d.