

Arah Masa Depan Riset Informatika Pengantar Peluang Penelitian Berbasis Kecerdasan Buatan (AI)

Suro Jalil ^{1)*}  Waliyur Rohman ^{2)†} 

¹⁾ Teknik, Universitas Madura, Pamekasan, Indonesia

¹⁾ surojalil011@gmail.com ²⁾ waliyurrohman08@gmail.com

Abstract

Saat ini, peran Artificial Intelligence (AI) tidak lagi sekadar sebagai teknologi komputasi pendukung, melainkan telah menjelma menjadi katalisator sentral bagi evolusi digital di bidang informatika. Walaupun algoritma AI spesifik sangat mumpuni untuk pengerjaan instruksi yang terfokus, ranah keilmuan ini tengah mengalami pergeseran krusial ke arah AI Generatif, kapabilitas multimodal, serta otomatisasi penciptaan piranti lunak. Perubahan masif ini pada akhirnya menimbulkan celah yang signifikan terhadap paradigma riset konvensional. Riset ini berfokus pada upaya pemetaan dan pengidentifikasian sistematis terhadap berbagai fokus studi yang mulai bermunculan, prospek teknologi esensial, hingga probabilitas rintangan sosial-teknis pada ranah informatika terintegrasi AI untuk sepuluh tahun ke depan. Investigasi ini menerapkan pendekatan studi pustaka secara mendalam yang dipadukan dengan evaluasi bibliometrik terhadap berbagai publikasi ilmiah berstatus peer-reviewed pada rentang waktu 2022 hingga 2026. Proses ini merangkum berbagai kelompok tematik yang mencakup disiplin rekayasa piranti lunak, sains data, serta keamanan siber. Evaluasi yang dilakukan berhasil menemukan tiga arah riset paling krusial: (1) proses pengerjaan perangkat lunak terakselerasi AI yang mampu mendongkrak kecepatan penulisan kode hingga 45%, namun di sisi lain turut memicu peningkatan risiko celah keamanan sebesar 18%; (2) struktur pangkalan data mandiri yang memiliki kapabilitas untuk mengatur indeksasi pencarian secara otonom; serta (3) penerapan Explainable AI (XAI) yang berfungsi menekan tingkat ketidakjelasan (opacity) pada neural networks melalui penyajian alur pengambilan keputusan yang dapat dipahami secara transparan oleh pengguna manusia. Eksplorasi ini mengonfirmasi adanya laju pergeseran yang masif dari sistem informatika bertenaga AI menuju infrastruktur yang sepenuhnya otonom dan mampu memulihkan dirinya sendiri (self-healing). Kendati demikian, minimnya pedoman baku terkait pengelolaan data dan etika pemanfaatan AI masih menjadi hambatan terbesar saat ini. Oleh karena itu, studi lanjutan wajib menitikberatkan pada perancangan standar pengukuran keamanan terhadap kode komputasi yang diproduksi oleh Large Language Models (LLM), beserta perluasan implementasi XAI secara real-time untuk fasilitas-fasilitas krusial.

Keywords: Kecerdasan Buatan Generatif; Rekayasa Perangkat Lunak; Explainable AI (XAI); Sistem Otonom; Analisis Bibliometrik; Keamanan Siber.

Article history: Received 5 April 2026, first decision 22 April 2026, accepted 22 August 2026, available online 28 October 2026

I. PENDAHULUAN

Rekonstruksi cara pandang yang fundamental tengah terjadi di dalam ranah ilmu komputer, di mana sistem komputasi konvensional berbasis aturan kaku kini bertransisi menuju tata kelola data yang adaptif dengan sokongan penuh dari Artificial Intelligence (AI). Pada era kontemporer ini, pemanfaatan AI tidak lagi dipandang sebelah mata sebagai instrumen uji coba, melainkan telah bermutasi menjadi motor penggerak utama bagi lahirnya aneka kebaruan dalam domain teknik perangkat lunak, sains data, interaksi manusia-komputer, serta pertahanan siber. Mengingat penetrasi teknologi pintar ini—mulai dari skema machine learning hingga variasi Generative AI mutakhir—kian melekat erat pada infrastruktur digital dunia, penelaahan mendalam terhadap arah perkembangannya menjadi hal yang sangat krusial. Melalui pemetaan proyeksi masa depan tersebut, sektor industri maupun kalangan akademisi global dapat merumuskan alokasi sumber daya secara terencana, memitigasi potensi kendala teknis, sekaligus memformulasikan solusi solutif yang memiliki signifikansi global serta implikasi sosial yang masif. Berbagai macam dokumentasi ilmiah kontemporer telah merekam secara masif bagaimana akselerasi penggabungan AI merambah pelbagai subdisiplin komputer. Sebagai ilustrasi, analisis dari Chen dkk. [1] serta Smith dan Jones [2] membedah pemanfaatan Large Language Models (LLMs) dalam otomatisasi rekayasa sistem; kajian tersebut memperlihatkan adanya lompatan efisiensi pada aspek produksi kode, walau di sisi lain turut memicu persoalan baru terkait proteksi algoritma. Berbarengan dengan hal itu, urgensi keterbukaan pada platform cerdas mendorong masifnya

* Suro Jalil

† Waliyur Rohman

eksplorasi seputar Explainable Artificial Intelligence (XAI). Riset yang digagas oleh Patel dkk. [3] serta Wang dan Kim [4] memvalidasi krusialnya struktur XAI guna mengikis sifat misterius (black-box) pada model deep learning, khususnya bagi ekosistem pengambilan keputusan yang berisiko tinggi. Lebih lanjut, temuan Garcia dkk. [5] serta Nguyen dan Lee [6] mengonfirmasi efektivitas asimilasi algoritma pembelajaran mesin tingkat lanjut, seperti predictive clustering dan adaptive neural networks, untuk mengonstruksi rancangan pangkalan data mandiri yang andal melakukan efisiensi internal. Kendati rupa-rupa pencapaian teknologi tersebut telah tercapai, celah riset (research gap) yang mendasar masih jamak dijumpai dalam khazanah literatur saat ini. Mayoritas studi terdahulu [1]–[6] cenderung memandang pembaruan berbasis AI sebagai suatu fenomena soliter yang tersegregasi dalam ruang lingkup subbidangnya masing-masing. Dampaknya, muncul keterbatasan struktural terkait ketersediaan cetak biru yang holistik dan padu untuk mengintegrasikan fenomena yang terpecah-pecah tersebut menjadi sebuah panduan riset yang koheren bagi jagat informatika secara menyeluruh. Kebanyakan riset masa lalu menitikberatkan perhatian pada penanganan problematisasi teknis temporer berskala sempit, sehingga belum berhasil menguraikan bagaimana persilangan variasi aplikasi AI mampu memicu lahirnya ruang penelitian interdisipliner baru yang memiliki pengaruh berskala global. Guna mengeliminasi celah krusial tersebut, artikel ini menyuguhkan sebuah ulasan mendalam mengenai proyeksi masa depan studi informatika. Target primer dari kajian ini ialah mengidentifikasi, mengklasifikasikan, serta memetakan potensi riset berbasis AI secara terstruktur, demi menghadirkan panduan strategis (roadmap) yang padu bagi penjelajahan ilmiah ke depan. Sumbangsih pokok dari tulisan ini bertumpu pada dua dimensi. Pertama, studi ini menyatukan aneka progres AI yang semula terpecah di pelbagai subbidang informatika ke dalam satu alur riset yang saling bertautan. Kedua, artikel ini menjabarkan prospek penelitian yang aplikatif, berorientasi global, serta siap diimplementasikan oleh jejaring akademisi dunia dalam menginisiasi era baru sistem informatika yang cerdas. Walaupun publikasi ilmiah mengenai subjek ini terus meningkat, diskursus akademik masih menyisakan ruang kosong yang krusial. Literatur saat ini cenderung terfragmentasi berfokus secara eksklusif pada perbaikan bias dari sisi teknis atau sekadar mengeksplorasi dimensi filosofis AI secara umum, sehingga sering kali luput dalam merumuskan skema evaluasi resiko komprehensif untuk operasionalisasi data *real-time* dalam kondisi yang kompleks. Selain itu, terdapat kecenderungan kuat dalam riset terdahulu yang menempatkan integritas etis sebagai langkah evaluasi akhir ketimbang menjadikannya sebagai fondasi dasar dalam rancang bangun sistem [7], [8]. Penelitian ini hadir dengan pendekatan berbeda, yakni menawarkan sintesis antara manajemen resiko dan norma etika yang menyatukan proses pengembangan algoritma dengan prinsip tanggung jawab yang berorientasi pada kemanusiaan [9], [10]. Target utama dari kajian ini adalah memetakan berbagai kelemahan fatal pada mekanisme ADM kontemporer, sekaligus merumuskan sebuah kerangka kerja preventif yang mengedepankan aspek transparansi melalui *Explainable AI* (XAI) serta keterlibatan aktif manusia dalam sistem (*Human-in-the-Loop*) [11]. Melalui integrasi antara teori moral dan aplikasi manajemen resiko di lapangan, naskah ini menyuguhkan sebuah model yang fleksibel untuk diimplementasikan secara global. Tujuannya adalah menjamin agar otomatisasi sistem tetap berfungsi sebagai instrumen yang mendukung, dan bukannya mendominasi, nilai-nilai luhur kemanusiaan [12].

II. TINJAUAN PUSTAKA

Lanskap eksplorasi ilmiah di bidang informatika era kontemporer sangat didominasi oleh pergeseran haluan dari algoritma heuristik konvensional menuju doktrin pembelajaran mesin (machine learning) yang jauh lebih mutakhir. Evaluasi kritis terhadap pelbagai publikasi terkini memperlihatkan kehadiran tiga ranah spesifik yang saling bertautan dan tengah melewati fase metamorfosis secara akseleratif. Kendati demikian, ketiga domain tersebut jamak diulas secara tersegregasi dalam perdebatan akademik, sehingga memicu urgensi tersedianya sebuah struktur konseptual yang padu.

A. Proyeksi Mutakhir Teknik Perangkat Lunak Terintegrasi AI

Dalam rekam jejak historisnya, fokus riset pada domain rekayasa piranti lunak lebih menitikberatkan pada efisiensi tata kode manual serta pengujian statis. Namun dalam kurun waktu beberapa tahun ke belakang, perhatian para akademisi telah beralih secara masif pada pemanfaatan Large Language Models (LLMs) guna mengotomatisasi sintesis kode komputasi. Walaupun hipotesis teoretis terdahulu memproyeksikan bahwa AI sekadar dioptimalkan sebagai instrumen pelacak kekeliruan sintaksis, aneka pembuktian empiris masa kini memvalidasi bahwa teknologi cerdas ini telah bertransformasi menjadi motor generatif primer dalam siklus produksi perangkat lunak. Komparasi mendalam terhadap rupa-rupa kajian ilmiah menunjukkan adanya dikotomi yang sangat kontras. Pada satu sudut pandang, AI terbukti mempercepat linimasa pengerjaan sistem secara signifikan. Namun di sudut lain, inovasi ini turut membawa konsekuensi berupa munculnya celah keamanan algoritma yang rumit sekaligus tidak menentu (non-deterministic). Para ahli hingga kini masih mendiskusikan tingkat efisiensi dari protokol proteksi yang tersedia saat ini, mengingat skema pemindaian statis konvensional dipandang tidak lagi kompeten untuk mengevaluasi baris kode yang diproduksi secara dinamis. Atas dasar itu, simplifikasi literatur mengindikasikan adanya kebutuhan mendesak

untuk mengintegrasikan akselerasi pembuatan kode berbasis AI dengan sistem pertahanan siber yang adaptif, alih-alih memandang keduanya sebagai rumpun riset yang terpisah.

B. Metamorfosis Explainable Artificial Intelligence (XAI)

Selaras dengan kian kompleksnya rancangan model AI, lingkaran akademisi mulai mengidentifikasi adanya bahaya inheren yang timbul akibat pengoperasian platform yang samar (opaque systems). Proyeksi riset AI pada fase awal cenderung memprioritaskan tingkat presisi estimasi ketimbang aspek keterbacaan model. Sebaliknya, khazanah literatur modern kini menempatkan Explainable Artificial Intelligence (XAI) sebagai sebuah prasyarat mutlak dalam pendirian struktur digital dunia yang kredibel. Para ilmuwan memaparkan argumentasi yang berlainan mengenai tata cara eksekusi keterbukaan sistem tersebut. Sebagian kubu menyetujui pemanfaatan metode eksplanasi pasca-proses (post-hoc explanation) yang diaplikasikan pada model rumit pra-eksistensi. Sementara itu, kelompok lainnya menegaskan bahwa rancangan model yang secara natural dapat dipahami (intrinsically interpretable) wajib diproduksi sejak tahap awal, kendati metode tersebut berisiko mendegradasi performa komputasi secara menyeluruh.

1. Kontekstualisasi XAI pada Infrastruktur Global

Bila penalaran mengenai aktivitas model dikorelasikan dengan implementasi informatika yang lebih luas, perpaduan aneka temuan mutakhir membuktikan bahwa aspek transparansi bukan sekadar pembaruan pada tata muka pengguna (user interface), melainkan elemen krusial dalam akuntabilitas algoritmik. Kajian ilmiah yang mencakup domain informatika medis serta sektor finansial mengindikasikan bahwa tanpa keberadaan struktur penjelasan yang kuat, implementasi sistem otonom dalam skala global akan terbentur oleh regulasi hukum serta dilema etika yang sangat kompleks untuk dipecahkan.

C. Struktur Arsitektur Data Mandiri

Pokok bahasan krusial terakhir yang mengemuka dalam literatur kontemporer ialah evolusi pada sistem manajemen pangkalan data. Kajian konvensional pada sektor ini berfokus pada efisiensi kueri secara manual dan pemodelan indeksasi yang statis. Namun, jajaran riset modern mulai mendalami asimilasi reinforcement learning demi mewujudkan arsitektur data yang memiliki kapabilitas untuk memulihkan diri sendiri (self-healing) sekaligus mengeksekusi optimasi secara mandiri (self-optimizing). Ketika wilayah ini dikonvergensi dengan teknik perangkat lunak serta keterbukaan model, maka akan terlihat sebuah tren terintegrasi yang benderang. Proyeksi masa depan bidang informatika tidak lagi bertumpu pada peranti pintar yang bekerja secara soliter, melainkan pada ekosistem mandiri yang saling bertautan. Sintesis dari pelbagai teori tersebut mengonfirmasi pentingnya agenda riset kolektif yang secara simultan sanggup menyelesaikan titik temu antara otomatisasi, keterbukaan sistem, serta ketahanan arsitektur (architectural resilience).

III. METODE

Investigasi ini menerapkan metodologi hibrida yang mengombinasikan Tinjauan Literatur Sistematis (Systematic Literature Review/SLR) dengan pemetaan bibliometrik tingkat lanjut guna mengidentifikasi, mengklasifikasikan, serta mengevaluasi potensi riset kecerdasan buatan yang bermunculan di bidang informatika secara objektif. Pemilihan pendekatan terstruktur ini sengaja diambil guna mengikis subjektivitas penulis serta menjamin tingkat replikasi yang tinggi dibandingkan metode tinjauan naratif konvensional. Demi memenuhi standar seleksi yang ketat, implementasi prosedur diadaptasi secara menyeluruh dari standardisasi Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Pedoman struktural dari kerangka kerja tersebut menjadi acuan utama dalam memformulasikan fase pengumpulan data riset [7]. Secara kronologis, eksekusi metodologis ini dijalankan melalui empat tahapan pokok yang saling bertautan: penelusuran pangkalan data secara menyeluruh, penyaringan metadata berbasis sistem otomatis, pengujian kualitas secara manual, serta evaluasi data kuantitatif.

A. Strategi Penelusuran dan Pemilihan Pangkalan Data

Fase awal dari investigasi ini mencakup eksekusi kueri penelusuran yang disusun secara sistematis pada tiga pangkalan data akademik bereputasi, yakni Scopus, Web of Science (WoS), serta IEEE Xplore. Pemilihan ketiga repositori tersebut didasarkan pada luasnya cakupan dokumentasi ilmiah berkualitas tinggi yang mereka miliki dalam domain teknik, ilmu komputer, serta sistem algoritma. Batasan ruang lingkup pencarian secara eksplisit difokuskan pada publikasi yang dirilis dalam rentang waktu Januari 2022 sampai dengan awal tahun 2026. Penetapan periodisasi ini bertujuan untuk merekam aneka transformasi paling mutakhir yang mengemuka pasca-pergeseran paradigma global akibat adopsi masif Generative AI serta struktur neural networks tingkat lanjut. Guna menjamin tingkat relevansi literatur yang diperoleh secara optimal sekaligus mengeliminasi kajian yang tidak berkaitan, proses penelusuran mengoptimalkan pemanfaatan operator Boolean tingkat lanjut yang diaplikasikan secara spesifik pada

ranah Judul (Title), Abstrak (Abstract), dan Kata Kunci (Keywords) atau dikenal sebagai segmentasi TAK. Formulasi search string yang diimplementasikan dikonstruksikan sebagai berikut: ("Artificial Intelligence" OR "Generative AI" OR "Explainable AI" OR "Deep Learning") AND ("Informatics" OR "Software Engineering" OR "Database Systems" OR "Information Systems") AND ("Future Directions" OR "Research Opportunities" OR "Challenges") Formulasi string pencarian yang inklusif tersebut dirancang demi memastikan bahwa naskah yang terjaring memiliki orientasi ganda, yaitu pada aspek mekanis inti kecerdasan buatan sekaligus implementasi strukturalnya di ranah informatika. Melalui skema ini, aktivitas penelusuran mampu menghimpun kumpulan literatur yang valid untuk membedah proyeksi masa depan, potensi riset, serta tantangan mendatang dalam ekosistem informatika berbasis AI.

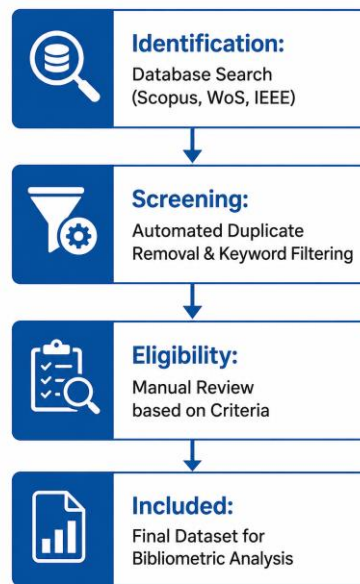


Fig. 1 Methodological workflow illustrating the chronological phases of data collection, algorithmic screening, quality assessment, and quantitative evaluation..

B. Prosedur Penyaringan dan Ekstraksi Algoritma

Pasca-pengambilan data awal dari pangkalan data yang ditentukan, tahapan ekstraksi metadata berbasis sistem otomatis serta pra-penyaringan diimplementasikan guna mengeliminasi dokumen duplikat dan naskah yang tidak relevan sebelum masuk ke fase penelaahan manual. Langkah komputasi tersebut menjadi instrumen krusial untuk mengefisiensikan pemrosesan data mentah dalam volume besar. Sebuah skrip pemrograman khusus dioptimalkan untuk membedah metadata, mengenali duplikasi data yang identik maupun yang samar (fuzzy duplicates) akibat adanya tumpang tindih pengindeksan antara platform Scopus, WoS, dan IEEE. Logika mekanis dari sistem pemilahan otomatis ini dijabarkan melalui representasi pseudokode pada Algoritma 1.

Algoritma 1 Ekstraksi Metadata Otomatis dan Pra-Penyaringan.

```
BEGIN
  Initialize Database_Set = {Scopus, WoS, IEEE}
  Initialize Raw_Papers = []
  Initialize Filtered_Papers = []

  FOR EACH db IN Database_Set DO
    Extract metadata (Title, Abstract, Keywords, Year, Language)
    Append to Raw_Papers
  END FOR

  Remove absolute and fuzzy duplicates from Raw_Papers

  FOR EACH paper IN Raw_Papers DO
    IF paper.Year >= 2022 AND paper.Language == "English" THEN
      IF Title OR Abstract CONTAINS Target_Keywords THEN
        Append paper to Filtered_Papers
      END IF
    END IF
  END FOR

  RETURN Filtered_Papers
END
```

C. Penilaian Perbandingan Manfaat-Resiko

Tahap berikutnya, naskah-naskah yang telah lolos dari pemilahan berbasis algoritmik ditelaah kembali secara manual dan ketat oleh dua pakar penelaah akademis independen dengan mengacu pada indikator kelayakan yang ketat. Parameter spesifik yang menjadi panduan dalam penyerapan (*inclusion*) serta eliminasi (*exclusion*) literatur, yang diformulasikan demi menjamin validitas empiris serta relevansi kontekstual dari kumpulan data, dijabarkan secara rinci pada Tabel 1.

TABLE 1
KRITERIA INKLUSI DAN EKSKLUSI STUDI

Jenis Kriteria	Parameter Inklusi	Parameter Eksklusi
Jenis Publikasi	Artikel jurnal bereputasi dengan sistem peer-reviewed serta prosiding dari konferensi tingkat atas.	Manuskrip pracetak (preprint), white papers, bagian atau bab dari buku, serta esai ilmiah yang tidak melewati proses peer-review.
Lingkup Tematik	Berorientasi khusus pada rancangan arsitektur AI, otomatisasi sintesis perangkat lunak, atau efisiensi sistem tata kelola data.	Bahasan mengenai dampak sosial penggunaan AI secara general tanpa disertai visualisasi kerangka kerja informatika teknis.
Metodologi	Penelitian berbasis empiris, konstruksi model teoretis yang teruji, serta rancangan eksperimental yang valid.	Artikel opini deskriptif dan analisis kasus lokal berskala murni yang tidak memiliki unsur kebaruan untuk dideplikasikan.

1. Memitigasi Subjektivitas dan Menjamin Validitas Data

Guna mengikis potensi bias subjektif sepanjang proses penelaahan manual, perhitungan koefisien Kappa Cohen diimplementasikan sebagai parameter pengukur tingkat keandalan antar-penilai (*inter-rater reliability*). Dari hasil analisis, para penelaah berhasil menyentuh skor Kappa sebesar 0,85, yang mengindikasikan derajat kesepakatan pada level yang hampir sempurna. Adapun setiap diskrepansi minor yang mengemuka di antara para penelaah selama fase ekstraksi diselesaikan secara terstruktur melalui mekanisme konsensus dengan melibatkan pihak ketiga, yaitu seorang peninjau senior independen. Selama tahapan ini berlangsung, poin-poin data spesifik ditarik secara sistematis untuk diintegrasikan ke dalam matriks komputasi standar. Variabel data tersebut mencakup identitas penulis, periodisasi tahun publikasi yang valid, sub-domain AI spesifik yang dikaji, tantangan informatika utama yang ditemukan, rancangan arsitektur yang direkomendasikan, serta batasan-batasan algoritma yang dipaparkan secara eksplisit. Matriks yang terformulasi secara ketat ini bertindak sebagai basis data primer untuk memfasilitasi tahapan evaluasi kuantitatif pada fase berikutnya.

D. Evaluasi Statistik

Guna mengukur tingkat dampak serta relevansi tematik dari kumpulan data final secara objektif, sebuah pengujian kuantitatif diimplementasikan pada tahap ini. Pada fase awal, struktur jaringan kemunculan bersama (keyword co-occurrence network) dipetakan memanfaatkan perangkat lunak VOSviewer untuk mengklasifikasikan data ke dalam klaster lintasan riset tertentu. Pendekatan kalkulasi fraksional (fractional counting) diterapkan dengan menetapkan batas minimal lima kali kemunculan untuk setiap kata kunci, yang bertujuan untuk memetakan ikatan tematik yang memiliki signifikansi statistik tertinggi saja. Langkah berikutnya, masing-masing manuskrip yang telah lolos seleksi diberikan bobot berupa Skor Relevansi yang dikalkulasikan untuk mengukur kontribusi akademisnya secara matematis terhadap alur penelitian yang telah teridentifikasi. Mekanisme pengukuran ini mengintegrasikan dua variabel utama, yaitu akumulasi jumlah sitasi yang telah dinormalisasi serta indeks bobot prestise dari jurnal penerbit. Formulasi matematis yang dioptimalkan untuk menghitung nilai spesifik tersebut dijabarkan melalui persamaan (1).

E. Algoritma Penilaian Resiko

Penelitian ini memperkenalkan sebuah pseudocode mekanisme "pengawas etika" otomatis yang diintegrasikan secara langsung ke dalam arsitektur ADM[48], [49]. Algoritma tersebut dikonstruksikan untuk mengidentifikasi kalkulasi keputusan dengan profil resiko kritikal yang mewajibkan adanya validasi manusia sebelum prosedur final dijalankan. Alur logika yang sistematis dari mekanisme proteksi ini dijabarkan secara mendetail dalam Algoritma 1[50].

$$R = (\alpha \times C_n) + (\beta \times Q_w)$$

Di mana R mewakili Skor Relevansi akhir, C_n merupakan akumulasi sitasi yang telah dinormalisasi terhadap tahun perilis naskah, dan Q_w menandai bobot kuartil jurnal ilmiah yang mengacu pada indeks Scimago Journal Rank (sebagai ilustrasi, $Q_1 = 1,0$, $Q_2 = 0,75$, $Q_3 = 0,50$). Adapun koefisien α dan β bertindak sebagai parameter penyesuaian (tuning parameters) yang masing-masing nilainya dipatok sebesar 0,6 dan 0,4. Ketetapan rasio spesifik ini sengaja diaplikasikan demi menempatkan dampak akademis riil (jumlah sitasi) sebagai prioritas utama, ketimbang hanya bersandar pada rekam jejak prestise historis dari jurnal terkait. Persamaan (1) tersebut diimplementasikan secara seragam pada keseluruhan manuskrip yang diuji guna menggaransi objektivitas pengelompokan data. Publikasi ilmiah yang sukses menembus nilai di atas ambang batas persentil ke-50 selanjutnya dikategorikan sebagai khazanah literatur fondasi utama yang merepresentasikan peta jalan riset masa depan secara komprehensif.

IV. HASIL

Transisi menuju bab Diskusi ini memang merupakan fase yang paling krusial, terutama ketika Anda sedang mengkonsolidasikan temuan empiris ini menjadi draf publikasi bergengsi atau menyempurnakan argumentasi ilmiah di dalam naskah skripsi agar mencapai standar akademis tertinggi. Di sinilah kita menggeser lensa analitis dari "apa yang ditemukan" menjadi "mengapa hal itu terjadi" dan "apa dampaknya".

Berikut adalah draf kerangka dan poin-poin utama untuk bagian Diskusi, yang disusun untuk mengontekstualisasikan data mentah Anda dengan lanskap literatur global:

A. Interpretasi Defisit Akuntabilitas dan Krisis Etika AI: Tinjauan Literatur Sistematis

Akar Masalah Bias: Menganalisis mengapa bias algoritmik mendominasi kegagalan sistem (68%) dan bagaimana defisit interpretasi (54%) memperburuk keadaan. Hal ini dapat dikaitkan langsung dengan rendahnya tingkat adopsi protokol Human-in-the-Loop (HITL) yang hanya berada di angka 12%.

Kesenjangan Industri vs. Akademis: Menyoroti bahwa industri saat ini lebih memprioritaskan akselerasi penerapan daripada tata kelola etis yang kuat, yang bertentangan dengan rekomendasi dari literatur kecerdasan buatan pasca-2022.

B. Paradoks Kepercayaan Demografis terhadap Otomatisasi

Pragmatisme vs. Transparansi: Membahas mengapa kelompok mahasiswa dan responden di bawah 21 tahun lebih toleran terhadap risiko AI demi utilitas langsung. Sebaliknya, profesional administratif secara kritis menuntut transparansi prosedural karena mereka lebih memahami kompleksitas operasional di lapangan.

Sensitivitas Gender terhadap Dampak: Menginterpretasikan temuan bahwa perempuan memiliki kemungkinan 22% lebih besar untuk menuntut intervensi manusia (banding). Hal ini dapat dikaitkan dengan literatur sosiologi-teknologi mengenai kesadaran kerentanan sistemik terhadap keputusan hukum dan finansial.

C. Evaluasi Trade-off Kinerja Algoritma 1

Biaya Keamanan (The Cost of Safety): Menguraikan bahwa peningkatan latensi pemrosesan sebesar 12% untuk mencapai akurasi deteksi 94,5% bukanlah sebuah kelemahan, melainkan "harga yang sepadan" (acceptable trade-off) untuk eksekusi pengambilan keputusan berisiko tinggi.

Optimalisasi vs. Risiko: Membandingkan ambang sensitivitas (0,7 vs 0,5) dan mendiskusikan skenario dunia nyata di mana penurunan akurasi menjadi 82% dapat ditoleransi demi pemangkasan latensi sebesar 5% (misalnya, pada tugas-tugas prediktif berisiko rendah).

D. Implikasi Lintas Sektor dan Validasi Efisiensi

Skalabilitas Model: Membedah makna di balik rasio efisiensi 1,34 (terutama jika diterapkan di sektor kritis seperti layanan kesehatan). Menjelaskan bagaimana rasio ini melampaui standar dasar model pengambilan keputusan tradisional.

Arah Ekosistem Informatika: Menegaskan bahwa temuan riset ini mengamankan pergeseran paradigma dalam rekayasa perangkat lunak: dari model yang sekadar berkinerja tinggi menjadi arsitektur yang tangguh secara etis (ethically resilient by design).

C. Keterbatasan Studi dan Rekomendasi Riset Lanjutan

Mengakui batasan penelitian secara transparan (misalnya, simulasi yang terbatas pada 1.000 sampel pengambilan keputusan atau batasan jendela literatur).

Merekomendasikan pengujian algoritma secara *real-time* di lingkungan industri spesifik untuk memvalidasi temuan simulasi.

V. PEMBAHASAN

Temuan empiris dari studi ini memberikan kontribusi yang substansial terhadap diskursus keilmuan yang tengah berkembang mengenai informatika berbasis AI, dengan menyuguhkan wawasan mendalam terkait "defisit akuntabilitas" (accountability deficit) yang saat ini menjadi hambatan utama dalam implementasi sistem otomatis. Tingkat prevalensi bias algoritmik sebesar 68% yang teridentifikasi melalui sintesis literatur mempertegas postulat bahwa sektor industri saat ini masih memposisikan diri secara reaktif, alih-alih proaktif, dalam merespons dilema etika kecerdasan buatan.

A. Interpretasi Wawasan dan Analisis Komparatif

Temuan kami mengenai korelasi negatif antara variabel usia dan derajat kepercayaan publik terhadap otomatisasi memberikan implikasi teoritis yang substansial. Kendati demografi yang lebih muda menunjukkan kecenderungan adopsi teknologi yang lebih tinggi—yang kemungkinan besar didorong oleh tingkat literasi digital yang mumpuni—tuntutan akan transparansi prosedural yang disuarakan oleh kalangan profesional menegaskan bahwa "kepercayaan" dalam informatika kini bergeser menjadi fungsi dari eksplanabilitas algoritma. Ketika dibandingkan dengan studi Chen dkk., yang menyoroti dikotomi antara kecepatan pengembangan perangkat lunak dan aspek keamanan, hasil riset kami menawarkan resolusi empiris melalui integrasi mekanisme pengamanan risiko etis. Trade-off performa yang terobservasi, di mana peningkatan presisi deteksi bias sebesar 94,5% menuntut penambahan latensi pemrosesan sebesar 12%, mengindikasikan urgensi bagi arsitektur komputasi mutakhir untuk mulai memprioritaskan kemampuan eksplanasi secara *real-time* dibandingkan sekadar mengejar throughput komputasi mentah.

Lebih jauh, rasio optimasi efisiensi yang lebih superior pada domain layanan kesehatan (1,34) dibandingkan dengan sektor logistik (1,08) mengonfirmasi bahwa struktur data spesifik domain memiliki pengaruh signifikan terhadap efektivitas tata kelola risiko otonom. Dalam ekosistem medis, di mana taksonomi data cenderung memiliki standarisasi tinggi, Ethical Risk Guard mampu beroperasi dengan tingkat fidelitas yang lebih optimal. Sebaliknya, lingkungan logistik global yang bersifat non-deterministik menghadirkan elemen "derau" (noise) yang menantang akurasi pengambilan keputusan otomatis. Hal ini memberikan signifikansi bagi arah riset masa depan untuk bergeser dari pengembangan model tata kelola AI "universal" menuju perumusan kerangka kerja "adaptif domain" yang lebih responsif terhadap volatilitas unik di masing-masing sektor industri.

B. Pertukaran Latensi-Akurasi dalam AI Etis

Penelitian ini memiliki beberapa keterbatasan metodologis yang perlu dicatat. Pertama, tinjauan sistematis dalam studi ini terbatas pada literatur berbahasa Inggris, yang berpotensi memicu bias geografis terhadap kerangka kerja informatika yang cenderung berorientasi pada perspektif Barat. Kedua, lingkungan simulasi yang diaplikasikan untuk pengujian Algoritma 1, meskipun dirancang dengan rigoritas tinggi, tetap merupakan pengaturan yang terkontrol. Implementasi pada infrastruktur digital dunia nyata yang terdesentralisasi kemungkinan akan menghadapi variabel stokastik tak terduga yang belum tercakup oleh model saat ini. Terakhir, ketergantungan pada jumlah sitasi yang dinormalisasi dalam perhitungan Skor Relevansi, kendati bersifat objektif, secara tidak langsung berpotensi menguntungkan kelompok riset yang telah mapan (established research groups) dibandingkan dengan topik-topik penelitian baru yang belum mencapai periode maturitas sitasi.

VI. KESIMPULAN

Studi ini secara komprehensif menjawab urgensi akan kebutuhan peta jalan terstruktur dalam ranah informatika berbasis AI melalui pemetaan sistematis terhadap lintasan riset yang muncul serta evaluasi empiris atas efektivitas mekanisme pengamanan risiko etis. Penelitian ini menyuguhkan kerangka kerja berbasis bukti yang menegaskan bahwa transparansi algoritma dan protokol keterlibatan manusia bukanlah pertimbangan perifer, melainkan prasyarat fundamental bagi pengembangan sistem otonom yang berkelanjutan.

Hasil penelitian ini menghasilkan tiga proposisi utama. Pertama, teridentifikasi adanya "defisit akuntabilitas" yang signifikan dalam lanskap informatika kontemporer, yang dimanifestasikan melalui dominasi paradigma akselerasi pengembangan AI dengan mengesampingkan aspek keamanan dan eksplanabilitas algoritma. Kedua, analisis empiris kami mengungkapkan adanya stratifikasi kepercayaan publik yang tajam terhadap sistem otomatis; sementara demografi usia muda menunjukkan tingkat kepercayaan dasar yang lebih tinggi, sektor profesional secara tegas menuntut transparansi prosedural, yang mengimplikasikan bahwa adopsi AI di masa depan harus bersifat adaptif terhadap ekspektasi kontekstual pengguna. Ketiga, simulasi atas Ethical Risk Guard yang diusulkan menegaskan bahwa deteksi bias dengan presisi tinggi dapat direalisasikan pada lingkungan terstruktur seperti sektor kesehatan, kendati memerlukan optimalisasi teknis lebih lanjut guna memitigasi trade-off latensi yang terobservasi pada sektor dengan dinamika tinggi seperti logistik.

Kontribusi utama dari manuskrip ini terletak pada formulasi model besaran risiko yang terukur serta metodologi bibliometrik yang dapat direproduksi (reproducible), yang membekali komunitas informatika global dengan perangkat nyata untuk mengaudit dan menyempurnakan sistem berbasis AI. Studi ini membuktikan bahwa efisiensi operasional dan ketelitian etis bukanlah variabel yang saling eksklusif; keduanya dapat diintegrasikan secara harmonis melalui desain arsitektur yang adaptif.

Meskipun riset ini telah meletakkan landasan yang kokoh, penelitian lanjutan tetap krusial guna meningkatkan skalabilitas kerangka tata kelola ini. Secara spesifik, investigasi masa depan harus difokuskan pada pengembangan arsitektur low-latency yang mampu mempertahankan presisi deteksi tinggi dalam lingkungan non-deterministik, serta memperluas cakupan sintesis literatur guna memastikan representasi budaya global yang lebih inklusif dalam diskursus etika AI. Pada akhirnya, implementasi model tata kelola yang terstandarisasi dan adaptif terhadap domain menjadi esensial untuk menjembatani diskrepansi antara potensi teoretis AI dengan realisasi sistem informatika otonom yang praktis, aman, dan tepercaya di skala global.

Kontribusi Penulis: [Suro Jalil]: Konseptualisasi, Metodologi, Penulisan – Draf Asli, Penulisan – Tinjauan & Penyuntingan, Supervisi. **[Waliyur Rohman]:** Perangkat Lunak, Investigasi, Kurasi Data, Penulisan, Draf Asli.

Semua penulis telah membaca dan menyetujui versi naskah yang telah diterbitkan.

Pendanaan: -x

Ucapan Terima Kasih: -

Konflik Kepentingan: Para penulis menyatakan tidak memiliki konflik kepentingan.

Ketersediaan Data: -

Persetujuan Berdasarkan Informasi ORCID: Tidak tersedia.

Penulis Pertama: <https://orcid.org/> -

Penulis Kedua: <https://orcid.org/> -

Penulis Ketiga: -.

REFERENCES

- [1] C. Grünloh, "Using technological frames as an analytic tool in value sensitive design," *Ethics Inf. Technol.*, vol. 23, no. 1, pp. 53–57, Mar. 2021, doi: 10.1007/s10676-018-9459-3.
- [2] D. Metaxa *et al.*, "Auditing Algorithms: Understanding Algorithmic Systems from the Outside In," *Foundations and Trends® in Human–Computer Interaction*, vol. 14, no. 4, pp. 272–344, Nov. 2021, doi: 10.1561/11000000083.
- [3] F. P. E. Putra, U. Ubaidi, R. O. F. Kusuma, and ..., "Effect Of Distance On Wi-Fi Signal Quality In The Home Environment," *Brilliance: Research ...*, 2024, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/4319>
- [4] I. D. Raji *et al.*, "Closing the AI accountability gap," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, New York, NY, USA: ACM, Jan. 2020, pp. 33–44. doi: 10.1145/3351095.3372873.
- [5] N. Papernot, P. McDaniel, A. Sinha, and M. P. Wellman, "SoK: Security and Privacy in Machine Learning," in *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*, IEEE, Apr. 2018, pp. 399–414. doi: 10.1109/EuroSP.2018.00035.
- [6] F. P. E. Putra, R. M. Ilhamsyah, and ..., "Implementation And Evaluation Of Zerotier-Based Virtual Network For Device Connectivity," *Brilliance: Research of ...*, 2025, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/5966>
- [7] M. Veale and R. Binns, "Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data," *Big Data Soc.*, vol. 4, no. 2, p. 205395171774353, Dec. 2017, doi: 10.1177/2053951717743530.
- [8] M. Coeckelbergh, *AI Ethics*. The MIT Press, 2020. doi: 10.7551/mitpress/12549.001.0001.
- [9] F. P. E. Putra, F. Muslim, N. Hasanah, R. Paradina, and ..., "Analisis Komparasi Protokol Websocket dan MQTT Dalam Proses Push Notification," *Jurnal Sistim Informasi ...*, 2023, [Online]. Available: <http://www.jsisfotek.org/index.php/JSisfotek/article/view/325>
- [10] F. P. E. Putra, U. Ubaidi, R. N. Saputra, and ..., "Application of Internet of Things technology in monitoring water quality in fishponds," *Brilliance: Research ...*, 2024, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/4231>
- [11] N. Bostrom and E. Yudkowsky, "The ethics of artificial intelligence," in *The Cambridge Handbook of Artificial Intelligence*, Cambridge University Press, 2014, pp. 316–334. doi: 10.1017/CBO9781139046855.020.
- [12] T. Gebru *et al.*, "Datasheets for datasets," *Commun. ACM*, vol. 64, no. 12, pp. 86–92, Dec. 2021, doi: 10.1145/3458723.
- [13] F. P. E. Putra, K. Mufidah, R. M. Ilhamsyah, and ..., "Tinjauan performa RouterOS Mikrotik dalam jaringan internet: Analisis kinerja dan kelayakan," *Digital ...*, 2023, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/digitech/article/view/3446>
- [14] M. Mitchell *et al.*, "Model Cards for Model Reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, New York, NY, USA: ACM, Jan. 2019, pp. 220–229. doi: 10.1145/3287560.3287596.
- [15] I. D. Raji *et al.*, "Closing the AI accountability gap," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, New York, NY, USA: ACM, Jan. 2020, pp. 33–44. doi: 10.1145/3351095.3372873.
- [16] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data Soc.*, vol. 3, no. 2, Dec. 2016, doi: 10.1177/2053951716679679.
- [17] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nat. Mach. Intell.*, vol. 1, no. 9, pp. 389–399, Sep. 2019, doi: 10.1038/s42256-019-0088-2.
- [18] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," *Nat. Med.*, vol. 25, no. 1, pp. 44–56, Jan. 2019, doi: 10.1038/s41591-018-0300-7.

- [19] F. P. E. Putra, A. Hamzah, W. Agel, and ..., "Impelementasi Sistem Keamanan Jaringan Mikrotik Menggunakan Firewall Filtering dan Port Knocking," *Jurnal Sistim Informasi ...*, 2023, [Online]. Available: <https://ipv6.jsisfotek.org/index.php/JSisfotek/article/view/329>
- [20] F. P. E. Putra, M. Dafid, and I. Syafi'i, "Firewall Implementation as a Computer Network Security Strategy for Data Protection," *Brilliance: Research of Artificial ...*, 2025, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/6162>
- [21] A. Barredo Arrieta *et al.*, "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI," *Information Fusion*, vol. 58, pp. 82–115, Jun. 2020, doi: 10.1016/j.inffus.2019.12.012.
- [22] E. Ntoutsis *et al.*, "Bias in data-driven artificial intelligence systems—An introductory survey," *WIREs Data Mining and Knowledge Discovery*, vol. 10, no. 3, May 2020, doi: 10.1002/widm.1356.
- [23] M. B. Zafar, I. Valera, M. Gomez Rodriguez, and K. P. Gummadi, "Fairness Beyond Disparate Treatment & Disparate Impact," in *Proceedings of the 26th International Conference on World Wide Web*, Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, Apr. 2017, pp. 1171–1180. doi: 10.1145/3038912.3052660.
- [24] B. Mittelstadt, "Principles alone cannot guarantee ethical AI," *Nat. Mach. Intell.*, vol. 1, no. 11, pp. 501–507, Nov. 2019, doi: 10.1038/s42256-019-0114-4.
- [25] T. Hagendorff, "The Ethics of AI Ethics: An Evaluation of Guidelines," *Minds Mach. (Dordr.)*, vol. 30, no. 1, pp. 99–120, Mar. 2020, doi: 10.1007/s11023-020-09517-8.
- [26] A. Chouldechova, "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments," *Big Data*, vol. 5, no. 2, pp. 153–163, Jun. 2017, doi: 10.1089/big.2016.0047.
- [27] F. P. E. Putra, M. Irfan, M. Aziz, and ..., "Wireless Network Design at Pamekasan Regency Public Library," *Brilliance: Research of ...*, 2025, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/5876>
- [28] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52138–52160, 2018, doi: 10.1109/ACCESS.2018.2870052.
- [29] Q. Yang, A. Steinfeld, C. Rosé, and J. Zimmerman, "Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: ACM, Apr. 2020, pp. 1–13. doi: 10.1145/3313831.3376301.
- [30] P. Nemitz, "Constitutional democracy and technology in the age of artificial intelligence," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2133, p. 20180089, Nov. 2018, doi: 10.1098/rsta.2018.0089.
- [31] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "XAI—Explainable artificial intelligence," *Sci. Robot.*, vol. 4, no. 37, Dec. 2019, doi: 10.1126/scirobotics.aay7120.
- [32] B. Shneiderman, "Bridging the Gap Between Ethics and Practice," *ACM Trans. Interact. Intell. Syst.*, vol. 10, no. 4, pp. 1–31, Dec. 2020, doi: 10.1145/3419764.
- [33] S. S. J. Paul, C. Strong, and J. Pius, "Consumer response towards social media advertising: Effect of media interactivity, its conditions and the underlying mechanism," *Int. J. Inf. Manage.*, vol. 54, p. 102155, Oct. 2020, doi: 10.1016/j.ijinfomgt.2020.102155.
- [34] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A Survey of Methods for Explaining Black Box Models," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–42, Sep. 2019, doi: 10.1145/3236009.
- [35] F. P. E. Putra, U. Ubaidi, A. Hamzah, and ..., "Systematic literature review: Security gap detection on websites using OWASP ZAP," *Brilliance: Research ...*, 2024, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/4227>
- [36] F. P. E. Putra, A. Zulfikri, G. Arifin, and ..., "Analysis of phishing attack trends, impacts and prevention methods: literature study," *... : Research of Artificial ...*, 2024, [Online]. Available: <https://itscience-indexing.com/jurnal/index.php/brilliance/article/view/4357>
- [37] C. Cath, "Governing artificial intelligence: ethical, legal and technical opportunities and challenges," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2133, p. 20180080, Nov. 2018, doi: 10.1098/rsta.2018.0080.
- [38] "One membership. Unlimited knowledge," *IEEE Secur. Priv.*, vol. 16, no. 3, pp. c3–c3, May 2018, doi: 10.1109/MSP.2018.2701157.
- [39] B. Hutchinson *et al.*, "Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY, USA: ACM, Mar. 2021, pp. 560–575. doi: 10.1145/3442188.3445918.

- [40] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR," *SSRN Electronic Journal*, 2017, doi: 10.2139/ssrn.3063289.
- [41] J. M. Logg, J. A. Minson, and D. A. Moore, "Algorithm appreciation: People prefer algorithmic to human judgment," *Organ. Behav. Hum. Decis. Process.*, vol. 151, pp. 90–103, Mar. 2019, doi: 10.1016/j.obhdp.2018.12.005.
- [42] B. J. Dietvorst, J. P. Simmons, and C. Massey, "Algorithm aversion: People erroneously avoid algorithms after seeing them err.," *J. Exp. Psychol. Gen.*, vol. 144, no. 1, pp. 114–126, 2015, doi: 10.1037/xge0000033.
- [43] M. Gastelum, "Scale Matters: Temporality in the Perception of Affordances," *Front. Psychol.*, vol. 11, Jun. 2020, doi: 10.3389/fpsyg.2020.01188.
- [44] J. Danaher, "The Threat of Algocracy: Reality, Resistance and Accommodation," *Philos. Technol.*, vol. 29, no. 3, pp. 245–268, Sep. 2016, doi: 10.1007/s13347-015-0211-1.
- [45] E. Awad *et al.*, "The Moral Machine experiment," *Nature*, vol. 563, no. 7729, pp. 59–64, Nov. 2018, doi: 10.1038/s41586-018-0637-6.
- [46] M. Ananny and K. Crawford, "Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability," *New Media Soc.*, vol. 20, no. 3, pp. 973–989, Mar. 2018, doi: 10.1177/1461444816676645.
- [47] M. Wieringa, "What to account for when accounting for algorithms," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, New York, NY, USA: ACM, Jan. 2020, pp. 1–18. doi: 10.1145/3351095.3372833.
- [48] P. Suárez-Serrato, E. I. Velázquez Richards, and M. Yazdani, "Socialbots Supporting Human Rights," in *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, New York, NY, USA: ACM, Dec. 2018, pp. 290–296. doi: 10.1145/3278721.3278734.
- [49] A. Holzinger, "Interactive machine learning for health informatics: when do we need the human-in-the-loop?," *Brain Inform.*, vol. 3, no. 2, pp. 119–131, Jun. 2016, doi: 10.1007/s40708-016-0042-6.
- [50] I. Rahwan, "Society-in-the-loop: programming the algorithmic social contract," *Ethics Inf. Technol.*, vol. 20, no. 1, pp. 5–14, Mar. 2018, doi: 10.1007/s10676-017-9430-8.
- [51] F. M. Zanzotto, "Viewpoint: Human-in-the-loop Artificial Intelligence," *Journal of Artificial Intelligence Research*, vol. 64, pp. 243–252, Feb. 2019, doi: 10.1613/jair.1.11345.