WEP Warwick Evans Publishing

International Journal of Computer Science and Information Technology

ISSN: 3005-9682 (Print), ISSN: 3005-7140 (Online) | Volume 7, Number 3, Year 2025
DOI: https://doi.org/10.62051/ijcsit.v7n3.10
Journal homepage: https://wepub.org/index.php/IJCSIT/index

Ethical Challenges and Countermeasures in Computer Science Development under Artificial Intelligence

Tang Lei

Mongolian National University, Ulaanbaatar, Ulaanbaatar 16060, Mongolia
949342510@qq.com

ABSTRACT

As artificial intelligence technology deeply permeates the field of computer science, the discipline is undergoing a paradigm shift from instrumental rationality to value-based rationality. AI-driven algorithmic innovation, data utilisation, and autonomous decision-making systems have not only restructured the developmental logic of computer technology but also precipitated unprecedented ethical crises. This paper systematically analyses the ethical dilemmas confronting computer science and technology within the AI context across data governance, algorithmic application, technological power distribution, and global governance. Drawing upon representative case studies and international practices, it proposes pragmatic yet forward-looking countermeasures across four dimensions: legal regulation, technological reconfiguration, multi-stakeholder collaboration, and global governance. This framework provides theoretical underpinnings and practical pathways for advancing the sustainable development of computer science and technology within an ethical framework.

KEYWORDS

Artificial intelligence; Computer science and technology; Ethical challenges; Algorithmic governance; Data privacy; Global collaboration

1. INTRODUCTION

Since the concept of 'artificial intelligence' emerged in the 1950s, it has maintained a profound symbiotic relationship with computer science and technology. Entering the era of generative AI and autonomous agents, the developmental logic of computer science and technology has undergone a fundamental transformation: shifting from the traditional 'command-execution' model to a 'data-decision' paradigm, where algorithms increasingly replace human judgement and decision-making functions. Statistics indicate the global artificial intelligence industry has reached a scale of US$1.8 trillion. The computer technology innovations it drives are permeating critical domains such as medical diagnosis, judicial adjudication, and public safety. Large models like GPT-5 now achieve over 90% accuracy in specialised fields. While this technological leap unleashes immense productive forces, it simultaneously disrupts the equilibrium of traditional technological ethics. When autonomous driving systems face the dilemma of 'crashing into a barrier and sacrificing passengers' versus 'crashing onto the pavement and sacrificing pedestrians'; when medical AI refuses to provide diagnostic advice for specific groups due to biased training data; when recruitment algorithms implicitly discriminate against female job seekers – the ethical risks of computer science and technology have evolved from theoretical concerns into tangible challenges. The sharp contradiction between the pace of technological advancement and the lag in ethical regulation has made 'how to ensure computer technology upholds ethical boundaries in the AI era' a global challenge [1].

2. ETHICAL DIMENSIONS OF AI AND COMPUTER TECHNOLOGY CONVERGENCE

2.1. Paradigm Shift in Technological Ethics

Traditional computer technology ethics centred on 'tool safety,' focusing on explicit risks such as hardware failures and programme vulnerabilities. Ethical norms primarily revolved around technical operational standards and user responsibility. In contrast, AI-driven computer technology ethics exhibits a three-dimensional expansion: On the responsibility dimension, algorithmic autonomy extends accountability beyond the single user to encompass developers, trainers, deployers, and other stakeholders. In terms of impact, technological influence extends from the individual to societal structures, potentially triggering systemic discrimination and power imbalances; in terms of evaluation, ethical judgements have evolved from 'safety and usability' to 'fairness and justice', necessitating a dual consideration of efficiency and ethical values.

2.2. Defining Core Ethical Categories

Data Ethics: Focuses on the governance of AI training data throughout its lifecycle, addressing three core issues: informed consent for data collection, purpose limitation for data usage, and boundaries for data sharing. Its essence lies in resolving the conflict between 'data value utilisation' and 'protection of personality rights'.

Algorithm Ethics: Concerns the fairness, transparency, and accountability of algorithmic decision-making, prioritising the governance of algorithmic discrimination, algorithmic hegemony, and algorithmic black-box issues. Its core objective is to prevent the abuse of algorithmic power.

Responsibility Ethics: Clarifying liability attribution when AI-driven computer systems cause harm. This necessitates establishing a 'developer-user-regulator' chain of responsibility to resolve the ethical dilemma of 'who is accountable'.

Global Ethics: Coordinating value differences and conflicting interests among nations in AI governance. This involves promoting an inclusive and equitable global governance framework to prevent widening technological divides and the emergence of governance hegemony.

3. KEY ETHICAL CHALLENGES IN COMPUTER SCIENCE DEVELOPMENT UNDER AI

3.1. Ethical Disorder in Data Governance: From Privacy Breaches to Data Colonialism

As the 'fuel' of computer technology in the AI era, data governance systems' ethical flaws have become the most prominent risk source, manifesting in three primary dimensions:

Covert and Massive Privacy Infringements: Traditional privacy protection relies on the principle of 'informed consent,' yet AI advancements have rendered this principle largely ineffective. On one hand, AI's multimodal perception capabilities can piece together comprehensive personal profiles from fragmented data, enabling the 'de-anonymisation' of even anonymised data through correlation analysis. On the other, 'blanket authorisation' has become an industry norm, compelling users to relinquish data rights when accessing free services. In 2024, an AI recommendation system on a social platform illegally collected users' geolocation data and chat records, exposing the privacy of millions and highlighting ethical vulnerabilities in data collection.

The Dilemma of Defining Data Ownership: Disputes over the ownership of AI training data have become a challenge in judicial practice. In the data infringement case between Company A and

Company B, where Company B unlawfully appropriated a voice dataset compiled by Company A over three years, the court ultimately recognised Company A's legitimate rights based on the 'substantial investment' principle. However, such rulings still lack explicit legislative underpinning. More complex is the commercialisation of public data: population health data collected by government departments being employed to train commercial AI models without providing reasonable compensation to data subjects, creating an ethical paradox of 'data exploitation'.

Inequality in global data governance: Imbalances in data resource distribution exacerbate the global digital divide. Currently, over 90% of high-quality training data is concentrated in a handful of developed nations, leaving developing countries facing a 'data deficit' predicament. Western tech giants exploit this imbalance by harvesting data from developing nations without sharing technological advancements, creating a phenomenon of 'data colonialism' — African medical data fuels the training of Western AI diagnostic models, while local healthcare institutions cannot afford to utilise these tools, further exacerbating global healthcare resource inequity [3].

3.2. Ethical Risks in Algorithm Application: From Discrimination Entrenchment to Power Monopoly

As the core engine of computer technology in the AI era, algorithms' ethical risks have permeated critical societal functions, exhibiting systemic characteristics:

Explicit algorithmic discrimination: Algorithms are not 'value-neutral' but reinforce and amplify existing societal biases. Recruitment AI automatically downgrades female candidates based on historical data suggesting 'women are more likely to take maternity leave'; A US judicial sentencing AI system assigned significantly higher recidivism risk scores to Black defendants than White defendants, despite comparable criminal circumstances. This cycle of 'data bias – algorithmic amplification – entrenched discrimination' transforms computer technology from a 'social tool' into a 'vehicle for discrimination,' severely eroding social fairness and justice.

The responsibility vacuum of algorithmic black boxes: The 'uninterpretability' of deep learning models complicates liability attribution. When medical AI-recommended treatments cause patient harm, developers evade responsibility by citing 'algorithmic autonomy,' while hospitals claim due diligence in review, leaving patients trapped in legal limbo. This accountability vacuum is even more pronounced in autonomous driving: in a 2023 accident where a collision occurred due to conflicting sensor data and algorithmic judgements, the vehicle manufacturer, software supplier, and user all deflected blame. Ultimately, the inability to pinpoint the specific logic behind the algorithmic decision rendered liability impossible to establish.

Algorithmic Power Monopolisation: A handful of tech giants monopolise power by controlling core algorithms. Search engine algorithms dictate information distribution, e-commerce recommendation algorithms govern consumer choices, and social media algorithms shape public discourse. This 'algorithmic hegemony' not only stifles market competition but also moulds user perceptions through 'filter bubbles,' eroding the public's capacity for independent judgement. The EU's Digital Markets Act directly addresses these issues by designating companies like Google and Amazon as 'gatekeepers,' requiring them to disclose algorithmic recommendation logic. However, implementation effectiveness remains constrained by technical complexity.

3.3. Ethical Imbalances in Technological Advancement: From Tool Alienation to Civilisational Challenges

AI-driven computer technology development is precipitating deeper ethical crises, challenging core human civilisational values:

The risk of dissolving human agency: As AI systems replace human decision-making across increasing domains, humanity may fall into a 'technological dependency' trap. Intelligent writing

tools reduce content creation to ‘prompt debugging’; autonomous driving systems diminish human driving skills; medical AI weakens clinicians' clinical judgement. More gravely, the development of autonomous intelligent agents may lead to ‘decision outsourcing’—where corporations rely on AI for strategic decisions and governments on AI for public administration. Humans may gradually lose control over technology, falling into an alienated state of being ‘dominated by tools’.

The clash of creativity and labour ethics: AI-generated content is disrupting traditional creative ethics. In the case of Mr Li v Ms Liu concerning AI painting infringement, while the court recognised the copyright of human creators, the standard of ‘originality’ will be redefined when AI can independently produce novels, music, paintings, and other creative works. Simultaneously, AI's impact on the labour market is already evident: intelligent customer service replaces human agents, algorithmic editors supplant journalists, and industrial robots displace assembly-line workers. Without reasonable social safeguards, such transformation risks triggering mass unemployment and societal instability, contradicting the ethical premise that ‘technology should benefit humanity.’

The risk of technological runaway: Research into general artificial intelligence harbours ‘ethical blind spots.’ Current AI systems have demonstrated ‘emergent capabilities’ exceeding human expectations, with some large models autonomously optimising code and generating novel algorithms, their developmental trajectory increasingly slipping beyond human control. More alarmingly, autonomous weapons systems are breaching the ‘ban on killer robots,’ with AI-driven arms tested by certain nations capable of independently identifying and attacking targets without human intervention. Should such technological advancement spiral out of control, it would precipitate catastrophic ethical consequences.

3.4. Ethical Dilemmas in Global Governance: From Conflicting Standards to Weak Cooperation

The global nature of computer technology contrasts sharply with the regional scope of AI governance, presenting structural challenges to the global governance system:

Fragmented Governance Mechanisms: Current global AI governance exhibits a pattern of ‘multiple approaches advancing simultaneously’ yet ‘operating in isolation’. Organisations such as the OECD, UNESCO, and G7 have issued ethical principles, yet significant divergences exist in implementation pathways and regulatory priorities. The EU's Artificial Intelligence Act adopts a stringent risk-tiered regulatory model, while the US emphasises ‘industry self-regulation as the primary approach.’ China's Interim Measures for the Administration of Generative AI Services focuses on record-filing management and content review. Such divergent standards trap multinational corporations in a ‘compliance dilemma’ and create opportunities for ‘regulatory arbitrage.’

Inequality in discourse distribution: Global AI governance remains dominated by Europe and the US, whose ethical paradigms embed liberal value assumptions while neglecting the practical needs of developing nations. In AI standard-setting by international organisations, developing nations' participation falls below 30%. The Western-centric nature of language corpora leads to ‘cultural misinterpretations’ in non-Western contexts—one translation AI erroneously rendered an African tribal greeting as an insult, exposing the ethical risks of governance discourse imbalance.

Weakened Collaborative Mechanisms: Effective mechanisms for multi-stakeholder collaborative governance remain unestablished. Roles within AI governance remain ambiguous for governments, enterprises, academia, and civil society organisations. Businesses frequently refuse algorithm disclosure under ‘technical confidentiality’ grounds, while academic ethical research struggles to translate into policy practice. Civil society lacks institutionalised channels for regulatory participation. Initiatives like the Global Partnership on Artificial Intelligence have attempted to foster multi-stakeholder engagement, yet their lengthy processes and broad, abstract agendas have rendered them largely ceremonial, failing to establish genuine collaborative governance networks.

4. ANALYSIS OF THE CAUSES OF ETHICAL CHALLENGES

4.1. Suppression of Value Logic by Technical Logic

AI-driven computer technology development adheres to a 'priority of efficiency' technical logic, creating inherent tension with the 'fairness and justice' value logic advocated by ethics. In R&D, enterprises prioritise quantifiable metrics like model accuracy and computational speed, often marginalising ethical considerations as 'technical obstacles.' Internal documents from one AI firm reveal that despite awareness of gender bias risks in its recruitment algorithm, it proceeded with deployment to enhance hiring efficiency—demonstrating technology's overwhelming dominance over ethical values. This 'technological supremacism' disregards the social attributes of computing technology, causing technological development to deviate from humanistic concerns.

4.2. Lagging and Ambiguous Legal Regulations

Legal frameworks struggle to keep pace with technological iteration, creating regulatory gaps. In data governance, China's Personal Information Protection Law establishes fundamental principles but lacks clear provisions on the reasonable boundaries for AI training data usage or criteria for data ownership attribution. In algorithm governance, no dedicated legislation regulates algorithmic fairness and transparency; in liability determination, core issues such as the attribution principles for 'algorithmic responsibility' and burden of proof allocation remain unresolved. This legal lag renders ethical disputes difficult to resolve through institutionalised channels, forcing reliance on case-by-case adjudication and resulting in unstable governance outcomes.

4.3. Mismatch Between Technological Complexity and Regulatory Capabilities

The complexity of AI technology far exceeds the capabilities of traditional regulatory approaches. The 'black-box nature' of deep learning models renders effective scrutiny by regulators challenging; the dynamic evolutionary capacity of AI systems renders 'one-off approvals' ineffective; and cross-border technology flows render single-nation regulation insufficient. Although China has established an algorithm registration system, insufficient technical expertise among regulators and limited intelligence in regulatory tools render the registration largely pro forma, failing to identify potential ethical risks. This capability gap leaves regulators ill-equipped to address complex AI systems.

4.4. Global Governance: Diverging Interests and Value Conflicts

Global AI governance faces the dual challenge of diverging interests and conflicting values. On the interests front, developed nations seek to maintain technological monopolies, reluctant to share core technologies or governance influence; developing countries urgently demand technological inclusivity, opposing 'data colonialism'. Regarding values, Western nations prioritise 'individual rights', placing data privacy protection foremost; developing countries emphasise 'development rights', seeking equilibrium between technological innovation and ethical regulation. This divergence hinders the formation of global governance consensus, trapping international cooperation in a state of 'discussion without resolution.'

5. ETHICAL GOVERNANCE STRATEGIES FOR COMPUTER SCIENCE AND TECHNOLOGY IN THE CONTEXT OF ARTIFICIAL INTELLIGENCE

5.1. Establishing a Rule-of-Law-Based Ethical Regulatory Framework

Refine specialised legislation and standards systems: Expedite the formulation of the Artificial Intelligence Ethics Review Act, establishing statutory benchmarks for algorithmic transparency, fairness, and accountability. Implement mandatory access reviews for AI systems in high-risk domains such as healthcare and the judiciary. Establish a tiered and categorised ethical standards framework: Prohibit certain applications (e.g., autonomous weapon systems), restrict others (e.g., commercial use of emotion recognition technology), and encourage specific uses (e.g., climate prediction models) to define clear boundaries for technological advancement. Develop algorithmic fairness assessment guidelines referencing ISO/IEC standards, specifying quantitative metrics for data bias and corrective methodologies.

Strengthen judicial adjudication and redress mechanisms: consolidate case law experience in AI-related disputes, clarifying copyright attribution for AI-generated works, standards for data infringement determinations, and rules for allocating the burden of proof in algorithmic discrimination cases. Establish specialised AI ethics tribunals incorporating computer experts and ethicists into case adjudication to address the challenge of establishing technical facts. Establish an AI Ethics Compensation Fund to provide swift redress for victims of algorithmic discrimination and data misuse, mitigating the high costs of individual rights enforcement.

Strengthen full lifecycle oversight: Construct an end-to-end regulatory framework encompassing 'pre-deployment review – in-process monitoring – post-incident accountability'. Pre-launch: Establish an ethical impact assessment system for AI systems, requiring enterprises to submit ethical evaluation reports prior to deployment. During operation: Utilise regulatory sandboxes for dynamic monitoring of high-risk AI systems. Pilot experiences in Beijing, Shenzhen and other locations demonstrate that regulatory sandboxes can effectively detect latent risks such as algorithmic discrimination. Post-incident: Implement strict accountability mechanisms, imposing penalties including fines and licence revocation on enterprises violating ethical norms, while pursuing legal liability of responsible individuals.

5.2. Promoting Ethical Reconstruction at the Technological Level

Developing ethically embedded technical architectures: Embedding ethical principles into the design phase of computer systems to create 'ethics-first' technical frameworks. Mandatorily integrating 'life-priority' modules into autonomous driving systems to ensure human life takes precedence in emergencies; incorporating 'fairness verification' protocols into medical AI to automatically detect and rectify discriminatory decisions targeting specific demographics. Drawing on the '3×3 Decision Verification Matrix', establish technical verification standards for algorithmic ethics, conducting ethical audits across three dimensions: objective setting, data input, and decision output.

Develop explainable AI and transparent algorithmic technologies: Overcome bottlenecks in explainable AI, mandating that AI systems in high-risk domains possess 'decision traceability' capabilities. Medical AI must not only provide treatment recommendations but also explain their rationale (e.g., symptom matching, drug allergy risks); judicial AI sentencing recommendations must include algorithmic reasoning explanations. Establish a tiered algorithmic transparency system: disclose foundational algorithmic logic to the public while providing full algorithmic parameters to regulators, balancing 'technical transparency' with 'trade secret protection'.

Develop a technical framework for data ethics governance: advance privacy-preserving computation technologies such as federated learning and homomorphic encryption to achieve 'data usability

without visibility,' unlocking data value while safeguarding security. Establish data provenance systems with end-to-end labelling of training datasets, clarifying informed consent status and ownership to prevent misuse at source. Promote data intellectual property registration systems to provide legal protection for datasets requiring substantial corporate investment, thereby encouraging lawful data transactions.

5.3. Cultivating a Diverse and Collaborative Governance Ecosystem

Establish multi-stakeholder governance mechanisms: Construct a collaborative governance framework characterised by 'government leadership, corporate responsibility, think tank support, and societal oversight'. Government shall formulate regulations and oversee enforcement; enterprises shall fulfil primary responsibility by establishing internal AI ethics committees and integrating ethical review into R&D processes; think tanks shall provide professional support while universities and research institutions conduct ethical theory research and technology assessments; civil society organisations shall exercise oversight through dedicated AI ethics reporting platforms to promptly identify violations. Ethics committees must include philosophers and sociologists constituting no less than 40% of membership to ensure substantive non-technical perspectives.

Strengthen corporate ethical responsibility: Promote the establishment of AI ethical compliance systems within enterprises, incorporating ethical performance into corporate assessment metrics. Require large technology companies to publish annual AI ethics reports, disclosing information such as data usage, algorithmic fairness assessment results, and ethical risk mitigation measures. Encourage corporate involvement in ethical standard-setting. The participation of enterprises such as Huawei and Alibaba in international rule-making demonstrates that deep corporate engagement enhances the feasibility of standards. Offer incentives like tax breaks and project prioritisation to companies practising ethical responsibility, creating a positive incentive for 'ethical compliance yielding benefits'.

Enhance public AI ethical literacy: Integrate AI ethics education into the national education system, introducing AI ethics courses in primary and secondary schools to cultivate young people's awareness of technological ethics; Conduct public outreach activities on AI ethics through case studies and scenario simulations to help citizens understand the harms of algorithmic bias and data misuse alongside preventive measures. Establish public participation channels in AI governance by soliciting open feedback during policy formulation, ensuring governance rules reflect public concerns and avoiding 'technocratic bias.'

5.4. Participating in the Construction of a Global Ethical Governance System

Promote international alignment of governance rules: Building upon the Global AI Governance Initiative, advance the establishment of mutual recognition mechanisms for AI ethics within multilateral frameworks such as the G20. Actively participate in standard-setting by international organisations including ISO and IEEE, championing issues like algorithmic transparency and linguistic diversity on the global agenda, and embedding the principle of 'people-centred development with equal emphasis on security' into international regulations. Strengthen rule coordination with developing nations to reflect the interests of the Global South and prevent Eurocentric monopolisation of regulatory frameworks.

Undertake global ethical capacity-building cooperation: Through mechanisms like the Digital Silk Road, assist developing countries in building AI ethical governance capabilities. Establish AI ethics training centres to provide technical training and policy advice; Share open-source ethical assessment tools to reduce governance costs for developing nations; Jointly pursue data governance cooperation to help developing countries establish data protection systems and guard against 'data colonialism.' Such collaboration will bridge the global digital divide while enhancing China's voice in global governance.

Establish a global risk prevention mechanism: Create an international AI safety early-warning network to monitor technological developments in high-risk areas such as autonomous weapons research and general artificial intelligence, enabling shared risk intelligence. Establish a transnational AI ethics violation sanction mechanism to impose joint sanctions on enterprises breaching global ethical consensus, curbing 'regulatory arbitrage' practices. Convene a global AI ethics summit to foster international consensus on core issues such as autonomous weapons prohibition and data privacy protection, constructing a new global governance paradigm of 'shared risk prevention and shared benefits'.

6. CONCLUSION

The advancement of computer science and technology driven by artificial intelligence, while catalysing a productivity revolution, has simultaneously precipitated four major ethical challenges: disorder in data governance, ethical lapses in algorithmic design, imbalances in technological development, and failures in global governance. These challenges fundamentally stem from the contradiction between the pace of technological advancement and the lag in ethical norms, the conflict between the instrumental rationality of technology and the value rationality of society, and the imbalance between the global distribution of technological resources and the allocation of governance discourse power. Addressing these challenges necessitates establishing a four-dimensional response system: 'legal regulation – technological reconstruction – multi-stakeholder collaboration – global governance.' This involves establishing ethical baselines through legal means, embedding ethical principles through technological innovation, forming governance synergy through multi-stakeholder participation, and resolving governance dilemmas through global cooperation.

REFERENCES

[1] Chen, L. Q. (2025). Ethical Challenges and Response Strategies in Artificial Intelligence Development [J]. Jiangnan Forum, 62(2), 72-77.

[2] Du, Yanyong. (2024). Ethical Review of Artificial Intelligence: Current Status, Challenges, and Solutions [J]. Journal of Donghua University (Social Sciences Edition), 24(2), 32-39.

[3] Cao Jianfeng. (2023). Towards Trustworthy AI: Governance Challenges and Responses for ChatGPT-like Generative Artificial Intelligence [J]. Journal of Shanghai University of Political Science and Law (Rule of Law Forum), 38(4), 28-42.

[4] Zhang Ming. (2024). Legal Regulation of Discrimination in Artificial Intelligence Algorithms [J]. Legal Studies, 46 (3): 124-140.

[5] Li Hua (2025). Technical Pathways and Ethical Values of Explainable AI [J]. Journal of Computers, 48 (2): 312-328.