American Journal of Innovations in Multidisciplinary Research

Volume 1, Issue 1 (May-June 2026) | ISSN: XXXX-XXXX

# Artificial Intelligence and Sustainable Development: A Multidisciplinary Framework for Inclusive Innovation

**Stuart Russell**

Professor
USA

**Abstract**

Artificial intelligence (AI) is increasingly embedded in the systems through which societies diagnose problems, allocate resources, deliver public services, design products, manage infrastructure, and generate knowledge. Its relevance to sustainable development lies not only in its capacity to process complex data or automate routine decisions, but also in its influence on who is visible in data, whose needs define innovation priorities, who has access to digital capability, and who bears the costs of technological change. AI can support climate modelling, precision agriculture, early-warning systems, health services, accessible education, energy management, and public-sector planning. At the same time, poorly governed AI can intensify exclusion, reproduce discrimination, concentrate economic value, increase surveillance, and add to environmental pressures through energy, water, and material use.

This paper develops a multidisciplinary framework for inclusive innovation at the intersection of AI and sustainable development. The framework connects technical performance with social justice, environmental responsibility, institutional capacity, and democratic accountability. It treats sustainable AI not as a collection of isolated applications but as a lifecycle question: a system should be assessed from problem selection and data collection through model development, procurement, deployment, monitoring, and retirement. The central argument is that an AI system can be considered developmentally valuable only when it addresses a legitimate public or community need, expands rather than narrows meaningful participation, and remains accountable for its social and ecological effects.

The paper uses an explanatory and integrative review of literature on AI, the Sustainable Development Goals (SDGs), responsible innovation, algorithmic fairness, digital inequality, climate action, education, health, agriculture, and governance. It distinguishes between AI as an enabling capability and AI as a socio-technical system shaped by data, compute infrastructure, institutions, incentives, regulations, and human judgement. The analysis considers both direct benefits, such as improved forecasting or resource optimization, and indirect risks, including biased outputs, loss of agency, unequal access to infrastructure, opaque decision-making, and rebound effects that can offset efficiency gains.

The paper concludes that inclusive innovation requires needs-led problem definition, participation by affected communities, accessible data and digital infrastructure, context-sensitive evaluation, human oversight, environmental proportionality, and mechanisms for redress.

**Keywords:** Artificial Intelligence, Sustainable Development, Inclusive Innovation, Sustainable Development Goals, Responsible AI, Digital Inclusion, AI Governance, Algorithmic Fairness, Climate Action, Public Interest Technology

## 1. Introduction

Artificial intelligence has moved from a specialized field of computing into a general-purpose capability that affects research, business, public administration, education, health, mobility, agriculture, media, and environmental management. Machine-learning systems can identify patterns in large data sets, generate predictions, support classification, optimize complex processes, and create text, images, software code, or scientific hypotheses. These capabilities are relevant to sustainable development because many development challenges involve interacting social, environmental, and economic conditions that are difficult to understand through a single discipline. Yet AI does not arrive in society as a neutral tool. It is designed within institutions, trained on historically produced data, deployed through uneven infrastructure, and governed by rules that can either protect or weaken the public interest.

The 2030 Agenda frames sustainable development as an integrated commitment to people, planet, prosperity, peace, and partnership. The SDGs recognize that poverty, health, education, gender equality, work, climate, biodiversity, institutions, and inequality are connected rather than separate policy domains. AI can help decision-makers interpret satellite imagery, forecast disease patterns, plan transport systems, reduce industrial waste, improve crop advice, expand assistive technologies, and target public services. However, the same systems can produce harms when they are deployed without representative data, accessible interfaces, environmental safeguards, or meaningful opportunities for affected groups to challenge decisions. The question is therefore not whether AI is good or bad for development in the abstract; it is whether particular applications are designed, governed, and evaluated in ways that improve real outcomes without shifting hidden costs onto vulnerable people or ecosystems.

A multidisciplinary approach is essential because technical accuracy alone cannot establish whether an AI system is socially desirable. A model can have high predictive performance while using data collected without consent, making decisions that affected individuals cannot understand, or redirecting scarce resources away from communities with limited digital visibility. Likewise, a system that appears efficient at the organizational level may create additional energy demand, require costly cloud infrastructure, displace frontline workers, or reinforce language barriers. Sustainable AI therefore requires insights from computer science, environmental studies, economics, public policy, law, ethics, sociology, education, public health, design, and local knowledge systems. Each discipline helps reveal a different part of the socio-technical environment in which AI operates.

Inclusive innovation provides a useful organizing principle. It asks who participates in innovation, who benefits from it, who is excluded, and how benefits and risks are distributed. In this paper, inclusive innovation means more than making a digital service available to a larger number of users. It means giving communities, workers, learners, patients, farmers, public servants, and civil-society groups a meaningful role in problem definition, design, testing, evaluation, and governance. It also means recognizing that access is shaped by affordability, language, disability, gender, location, connectivity, legal status, literacy, trust, and the ability to obtain remedy when a system causes harm.

## 2. Literature Review

The literature on AI and sustainable development identifies substantial potential alongside substantial uncertainty. Studies linking AI to the SDGs emphasize applications in health, education, energy, agriculture, cities, climate action, and environmental monitoring. These studies also stress that AI can create negative interactions by increasing inequality, weakening privacy, amplifying misinformation, or encouraging high-consumption patterns. This duality is important because it moves the discussion beyond technological optimism. AI should be assessed through the conditions under which it is used, the institutions that control it, and the outcomes it produces for different groups rather than through a generic assumption that digital innovation is inherently sustainable.

A second body of literature focuses on ethical and human-rights dimensions. Recurring principles include human dignity, non-discrimination, fairness, privacy, transparency, accountability, safety, and meaningful human oversight. Research on algorithmic bias demonstrates that performance can vary across demographic groups when datasets, labels, measurement practices, or deployment contexts are uneven. Scholars also note that transparency is not achieved merely by publishing a technical description. People affected by an AI-supported decision need information that is understandable, timely, relevant to the decision, and connected to a process for questioning or appealing the outcome.

A third theme concerns the political economy of AI. Developmental benefits depend on access to connectivity, data resources, local research capacity, affordable compute, relevant language technologies, and skilled institutions. When these conditions are concentrated in a small number of countries or firms, AI can deepen dependency rather than build local capability. This concern is particularly important for low-resource languages, rural areas, smaller enterprises, public-service organizations, and community groups that may be expected to use systems designed elsewhere without the ability to adapt, audit, or govern them. Inclusive innovation therefore requires investment not only in applications but also in public digital infrastructure, skills, standards, and local research ecosystems.

A fourth theme concerns environmental responsibility. AI can improve forecasting, coordinate energy systems, optimize logistics, support climate science, and identify environmental change. At the same time, AI models rely on physical systems: data centers, semiconductor supply chains, networks, cooling systems, and electricity. Environmental assessment must therefore consider the full lifecycle of an application, including energy use, water requirements, equipment replacement, e-waste, and potential rebound effects. The relevant question is not whether a model is computationally sophisticated; it is whether the expected public benefit is proportionate to the resources consumed and whether lower-impact alternatives have been considered.

**Table 1: Summary of Key Literature on Artificial Intelligence and Sustainable Development**

| S. No. | Author(s) / Year | AI and Sustainability Focus | Study Context | Key Contribution |
|---|---|---|---|---|
| 1 | United Nations (2015) | Sustainable Development Goals | 2030 Agenda | Established an integrated framework that connects social, economic, environmental, and institutional development priorities. |

| S. No. | Author(s) / Year | AI and Sustainability Focus | Study Context | Key Contribution |
|---|---|---|---|---|
| 2 | Vinuesa et al. (2020) | AI and SDG interactions | Multi-SDG evidence synthesis | Identified enabling as well as inhibiting pathways through which AI can affect progress across the SDGs. |
| 3 | UNESCO (2021) | Ethics of AI | Global normative guidance | Connects AI governance with human rights, inclusion, environmental protection, and accountability. |
| 4 | Jobin, Ienca & Vayena (2019) | AI ethics principles | Global guideline review | Mapped recurring principles including transparency, justice, responsibility, privacy, and non-maleficence. |
| 5 | Floridi et al. (2018) | Beneficial AI | AI4People framework | Proposed a framework around beneficence, non-maleficence, autonomy, justice, and explicability. |
| 6 | Buolamwini & Gebru (2018) | Algorithmic bias | Gender classification systems | Demonstrated performance disparities and the need for representative data and disaggregated evaluation. |
| 7 | Klerkx, Jakku & Labarthe (2019) | Digital agriculture | Agrifood systems | Explored how data-driven technologies reshape agricultural knowledge, practices, and governance. |
| 8 | Rolnick et al. (2022) | Machine learning for climate | Cross-sector applications | Outlined application opportunities in electricity, transport, buildings, agriculture, and earth systems. |
| 9 | Kaack et al. (2022) | Climate mitigation and AI | Governance and emissions | Examined how AI can support mitigation while creating risks linked to energy use and policy choices. |
| 10 | Suresh & Guttag (2021) | Sources of algorithmic harm | Machine-learning lifecycle | Explained how harms can emerge from data, modelling, deployment, feedback, and institutional context. |
| 11 | NIST (2023) | AI risk management | Governance and operations | Provided a practical framework for managing trustworthy AI risks across organizational processes. |
| 12 | UNESCO (2023) | Generative AI in education | Education and research | Highlighted human-centred safeguards, capacity building, inclusion, and protections for learners and educators. |

## 3. Object of Study

The object of this study is the relationship between artificial intelligence and sustainable development, with specific attention to the conditions under which AI can support inclusive innovation. The study considers AI as a socio-technical system that includes data, algorithms, compute infrastructure, interfaces, users, institutions, financing, regulation, and environmental resources. It therefore examines AI not only as a tool for automation or prediction, but also as a system of choices about which problems are prioritized, what kinds of knowledge are recognized, and how benefits and burdens are distributed.

The study examines AI applications across multiple development domains, including health, education, agriculture, climate action, energy, urban services, public administration, accessibility, and small-enterprise support. It considers how AI may create value through early warning, resource optimization, decision support, personalized services, scientific discovery, and improved access to information. It also examines risks related to bias, exclusion, surveillance, opacity, misinformation, job quality, unequal infrastructure, environmental footprint, and weak accountability. The purpose is not to treat all AI applications as comparable. The risks and benefits of a clinical decision-support system, an agricultural advisory tool, a public-benefit eligibility model, and a generative AI learning assistant differ substantially in evidence requirements, affected rights, data sensitivity, and acceptable levels of uncertainty.

**Key Points of Object of Study**

- To examine AI as a socio-technical system rather than only as a computational tool or model.
- To assess how AI can contribute to social, economic, environmental, and institutional dimensions of sustainable development.
- To distinguish between technical performance, real-world effectiveness, public value, and equitable distribution of benefits.
- To identify participation barriers linked to income, language, gender, disability, geography, connectivity, and digital capability.
- To analyse how data quality, representation, consent, privacy, and local context affect the legitimacy of AI-supported decisions.
- To evaluate risks related to algorithmic bias, opacity, misinformation, surveillance, and loss of human agency.
- To examine energy, water, hardware, and e-waste considerations across the lifecycle of AI systems.
- To study the role of public institutions, research organizations, private firms, communities, and civil society in AI governance.
- To compare needs-led innovation with technology-first deployment and identify conditions for responsible scaling.
- To develop a multidisciplinary framework for inclusive, accountable, and environmentally proportionate AI innovation.

**Table 2: Object of Study Framework**

| S. No. | Study Component | Description | Purpose |
|---|---|---|---|
| 1 | AI System | Data, model, interface, compute resources, users, and decision processes. | Define the technical and organizational unit of analysis. |
| 2 | Sustainable Development Outcome | Expected social, economic, environmental, or institutional benefit. | Connect technical activity with measurable public-interest goals. |
| 3 | Inclusion and Participation | Access, affordability, language, disability, representation, consent, and community voice. | Assess who can shape, use, and benefit from innovation. |
| 4 | Data and Infrastructure | Data quality, interoperability, connectivity, digital public infrastructure, and local capability. | Understand the practical conditions needed for reliable deployment. |
| 5 | Governance and Rights | Transparency, privacy, fairness, safety, oversight, redress, and procurement standards. | Protect people from unjustified risks and enable accountability. |
| 6 | Environmental Proportionality | Energy, water, material use, lifecycle impacts, and lower-impact alternatives. | Ensure ecological costs are visible and justified by expected benefit. |
| 7 | Institutional Capacity | Skills, maintenance, budgets, evaluation, and adaptation capacity. | Assess whether an intervention can be sustained after launch. |
| 8 | Outcome Measurement | Equity, effectiveness, resilience, user experience, and unintended consequences. | Evaluate public value over time rather than relying on model accuracy alone. |

## 4. Research Methodology

This paper uses an explanatory and integrative review design. It does not collect new survey responses, run a new machine-learning experiment, or evaluate a single deployed product. Instead, it synthesizes scholarship, policy frameworks, technical guidance, and institutional reports concerning AI, sustainable development, responsible innovation, inclusion, and environmental governance. This approach is appropriate because the research question concerns relationships that extend across sectors, populations, institutions, and technologies. A single domain-specific evaluation would not be sufficient to capture the social and ecological conditions that determine whether AI contributes to development outcomes.

The review is structured around six analytical domains: development needs and public value; participation and inclusion; data and infrastructure; model and deployment risks; environmental proportionality; and governance and accountability. Sources are assessed according to the development problem addressed, the AI capability involved, the population affected, the evidence of benefit, the distribution of risk, the governance mechanism proposed, and the environmental implications identified. The analysis differentiates between proof of technical feasibility and proof of real-world public value. For example, an accurate prediction system may still be unsuitable when frontline institutions cannot act on the prediction, when data subjects lack meaningful consent, or when the recommended action is inaccessible to the people most affected.

### 4.1 Research Design

- Study Type: Explanatory and literature-based research.
- Approach: Integrative review and conceptual synthesis.
- Unit of Analysis: AI-enabled systems, projects, and governance arrangements connected to sustainable-development objectives.
- Primary Domains: Public value, inclusion, data governance, technical reliability, environmental impact, and accountability.
- Institutional Context: Governments, public-service providers, research institutions, firms, civil-society organizations, and communities.
- Outcome Variables: Accessibility, equity, effectiveness, resilience, environmental proportionality, transparency, and remedy.
- Ethical Position: No personal, confidential, or human-subject data were collected for this study.

### 4.2 Methodological Steps

- Step 1: Define sustainable development outcomes as social, economic, environmental, and institutional goals that require evidence beyond model accuracy.
- Step 2: Identify development contexts in which AI may be useful, unnecessary, inappropriate, or secondary to non-technical solutions.
- Step 3: Review evidence on inclusion, access, and participation, including language, disability, connectivity, affordability, and local knowledge.
- Step 4: Examine data quality, representation, consent, privacy, and the risk of bias in AI-supported decisions.
- Step 5: Assess model design, evaluation, human oversight, transparency, and routes for contesting harmful or inaccurate outcomes.
- Step 6: Analyse the environmental footprint of compute, energy use, water use, hardware, and potential rebound effects.
- Step 7: Compare governance mechanisms such as risk assessments, procurement rules, audit practices, impact evaluations, and independent oversight.
- Step 8: Develop policy and design recommendations that align innovation with equity, public value, ecological responsibility, and institutional capacity.

## 5. Data Analysis

The data analysis in this paper is a qualitative synthesis of the reviewed literature. The evidence is interpreted by tracing how a development problem becomes an AI proposal, how the proposal is shaped by data and infrastructure, how it is deployed through institutions, and how it affects different groups and ecological systems. Since the study does not use a new dataset, it does not claim a universal numerical estimate of AI impact on the SDGs. Rather, it identifies recurring patterns that help explain why apparently promising AI applications can produce very different outcomes across settings. The central finding is that inclusive innovation is cumulative: public value depends on the alignment of problem definition, participation, data governance, implementation capacity, environmental responsibility, and accountability. Failure in any one domain can weaken the value of a technically capable system.

**5.1 AI Capability and Sustainable Development Alignment**

AI can create value when it is aligned with a clearly specified development problem and a feasible pathway to action. In agriculture, for example, an AI-supported advisory system may identify pest risk or irrigation needs, but its usefulness depends on whether farmers can access the advice in a relevant language, trust the source, obtain recommended inputs, and adapt recommendations to local soil, weather, and market conditions. In health, a triage model may identify patients at elevated risk, but it does not improve outcomes unless a clinic has staff, medicines, referral pathways, and a safe process for communicating results. Similar reasoning applies to education, climate adaptation, public-benefit delivery, and urban services. AI should be evaluated as one component in a wider service system rather than as an independent intervention.

The literature suggests that technology-first deployment often creates a mismatch between model capability and institutional need. Organizations may procure an AI product because it appears innovative, while the underlying problem is inadequate data collection, staffing, governance, or service design. A needs-led approach begins by asking what decision must be improved, who makes that decision, what evidence is already available, what non-AI options exist, and what harms could follow from an error. In some cases, a simple dashboard, better records, community outreach, or rule-based workflow may be more reliable, affordable, and understandable than a complex model. The appropriate solution is therefore the least complex option that can meet the public objective safely and effectively.

Alignment also requires a theory of change. An AI system should state how its output will lead to a real benefit, what actors must act on the output, which assumptions must hold, and what indicators will show progress. Without this chain, organizations can mistake technical novelty for impact. A developmentally aligned AI project should include baseline conditions, expected beneficiaries, groups that may face new burdens, mechanisms for feedback, and criteria for suspension or redesign when the system does not deliver its intended value.

**Key Observation:**

- A technically accurate model does not guarantee a useful or equitable development outcome.
- Needs-led problem definition reduces the risk of deploying AI where simpler, lower-cost, or community-based solutions are more appropriate.
- A clear theory of change should connect model outputs, institutional action, beneficiary experience, and measurable public value.

**5.2 Inclusion, Access, and the Distribution of Benefits**

Inclusive innovation requires attention to the unequal conditions under which people encounter AI. Connectivity, device ownership, digital literacy, language, disability access, gender norms, time availability, legal identity, and trust in institutions influence whether individuals can use an AI-enabled service. A system delivered through a smartphone application may appear widely available while excluding people who share devices, cannot afford data, rely on low-bandwidth networks, use a low-resource language, or need an accessible interface. These barriers do not merely reduce uptake; they can distort the data used to evaluate the system, making it appear successful for connected users while masking exclusion among those it fails to reach.

Volume 1, Issue 1 (May-June 2026) | ISSN: XXXX-XXXX

Participation must extend beyond consultation. Communities and frontline workers often possess knowledge about local risks, seasonal patterns, cultural norms, service bottlenecks, and safety concerns that are not visible in administrative data. Their involvement can improve problem definition, data interpretation, interface design, testing, and evaluation. Participatory approaches also make it more likely that a system will be trusted and used responsibly. However, participation should be structured and resourced; it cannot depend on unpaid labour or symbolic meetings. Stakeholders should receive clear information, reasonable time, accessible formats, and a meaningful ability to influence decisions.

**Key Observation:**

- Access should be assessed through affordability, accessibility, language, connectivity, trust, and the ability to obtain support, not only through application availability.
- Participatory design improves relevance when community knowledge and frontline experience are treated as evidence rather than as an afterthought.
- Average performance can conceal uneven burdens; evaluation should examine outcomes for groups that may be underrepresented or disproportionately harmed.

### 5.3 Environmental Proportionality and Resource Responsibility

AI can contribute to environmental sustainability by improving weather forecasting, environmental monitoring, energy-system management, logistics, agricultural planning, building efficiency, and scientific discovery. These opportunities are important, particularly where complex data can help institutions anticipate risk or reduce waste. However, environmental benefit must be demonstrated rather than assumed. A system that improves operational efficiency may still increase total consumption when lower costs or faster processes encourage additional use. This rebound effect means that evaluation should consider system-wide consequences instead of relying only on direct efficiency gains.

The environmental footprint of AI includes electricity for training and inference, data-center cooling, water use, hardware manufacturing, supply-chain impacts, and electronic waste. The footprint varies across model size, task, location, energy source, hardware life cycle, and frequency of use. A responsible approach therefore considers whether a smaller model, a more efficient workflow, a shared public infrastructure, a local data-processing solution, or a non-AI alternative could achieve the same objective. Environmental proportionality does not require rejecting computationally demanding systems in every case; it requires making resource use visible and weighing it against credible, measurable benefits.

**Key Observation:**

- Environmental benefit should be evaluated at the system level, including possible rebound effects and infrastructure impacts.
- Resource-intensive AI should be justified by a credible public benefit and assessed against lower-impact alternatives.
- Procurement, reporting, and lifecycle review can make energy, water, hardware, and e-waste considerations operational rather than rhetorical.

### 5.4 Governance, Capacity, and Co-Innovation

AI governance for sustainable development must combine rules, institutional capacity, and practical implementation. Principles such as fairness and transparency are necessary but insufficient when organizations lack skilled staff, data stewardship, security controls, procurement standards, evaluation methods, or the authority to pause a harmful system. Governance should therefore be translated into operational practices: impact assessments before deployment, documentation of data and model limitations, independent review for high-stakes use, user testing, audit trails, incident reporting, clear roles for human oversight, and accessible grievance mechanisms.

Capacity building is especially important for public institutions, schools, clinics, local governments, small enterprises, and community organizations. These actors may face pressure to adopt AI while having limited bargaining power with vendors or limited access to technical expertise. Sustainable innovation requires the ability to ask informed questions, review claims, negotiate data-sharing terms, understand model limitations, and maintain systems after pilots end. Investments in open standards, public-interest research, digital public infrastructure, local-language resources, and interdisciplinary education can reduce dependency and broaden the distribution of capability.

**Key Observation:**

- Ethical principles require operational governance practices, skilled institutions, and enforceable routes for accountability.
- Capacity building enables organizations and communities to evaluate, adapt, and govern AI rather than becoming passive recipients of external systems.
- Co-innovation improves local relevance and legitimacy by involving technical, domain, and community expertise throughout the lifecycle.

## 6. Discussion

The literature indicates that AI should be approached as a development capability with conditions, not as a universal solution. Its value depends on whether it expands the ability of people and institutions to make better decisions, access essential services, protect ecosystems, and respond to complex risks. This is why technical progress must be accompanied by public-interest governance. The most meaningful question is not whether an organization has adopted AI, but whether its use is justified by a real need, supported by evidence, accessible to affected groups, and accountable for outcomes that matter in everyday life.

The proposed Inclusive AI for Sustainable Development Framework brings together six dimensions. First, needs-led problem definition asks whether AI is appropriate and what public purpose it serves. Second, equitable participation and accessibility ensure that affected people can influence design and use systems meaningfully. Third, data and infrastructure readiness address the quality, representativeness, privacy, interoperability, and local capacity required for reliable use. Fourth, responsible lifecycle management embeds testing, documentation, human oversight, monitoring, and redress across the system. Fifth, environmental proportionality makes resource use and ecological trade-offs visible. Sixth, public accountability requires transparent governance, procurement discipline, evaluation, and enforceable remedy.

The framework also clarifies why inclusion and sustainability cannot be separated. A system that reduces emissions but excludes rural communities, uses inaccessible interfaces, or shifts costs onto low-income households is not fully sustainable. Likewise, a system that reaches many users but relies on excessive resource consumption, weak privacy protections, or unjustified surveillance does not represent inclusive progress.

## 7. Conclusion

This paper examined artificial intelligence and sustainable development through a multidisciplinary framework for inclusive innovation. The analysis shows that AI can support meaningful progress in areas such as climate resilience, health, education, agriculture, energy, urban services, accessibility, and public administration. Yet its benefits are neither automatic nor evenly distributed. Data gaps, infrastructure inequality, algorithmic bias, weak governance, inaccessible design, environmental costs, and limited institutional capacity can transform a promising application into a source of exclusion or harm.

The central conclusion is that sustainable AI must be evaluated through its full lifecycle and its real-world consequences. Developmentally valuable AI begins with a legitimate need, includes affected communities in appropriate ways, uses data responsibly, operates with human oversight, makes risks understandable, provides routes for redress, and demonstrates that its environmental costs are proportionate to its benefits. These conditions do not prevent innovation; they create the trust, relevance, and accountability required for innovation to endure.

Future research should examine long-term evidence on AI-enabled public services, local-language technologies, climate-related applications, educational tools, worker impacts, and community-governed data systems.

## 8. References

1. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610-623. https://doi.org/10.1145/3442188.3445922
2. Benjamin, R. (2019). Race after technology: Abolitionist tools for the New Jim Code. Polity Press.
3. Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. Patterns, 2(2), 100205. https://doi.org/10.1016/j.patter.2021.100205
4. Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Proceedings of Machine Learning Research, 81, 1-15.
5. Dignum, V. (2019). Responsible artificial intelligence: How to develop and use AI in a responsible way. Springer.
6. Floridi, L., Cowls, J., Beltrametti, M., et al. (2018). AI4People - An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28, 689-707. https://doi.org/10.1007/s11023-018-9482-5
7. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1, 389-399. https://doi.org/10.1038/s42256-019-0088-2
8. Kaack, L. H., Donti, P. L., Strubell, E., et al. (2022). Aligning artificial intelligence with climate change mitigation. Nature Climate Change, 12, 518-527. https://doi.org/10.1038/s41558-022-01377-7
9. Klerkx, L., Jakku, E., & Labarthe, P. (2019). A review of social science on digital agriculture, smart farming and agriculture 4.0: New contributions and a future research agenda. NJAS: Wageningen Journal of Life Sciences, 90-91, 100315. https://doi.org/10.1016/j.njas.2019.100315
10. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. https://doi.org/10.6028/NIST.AI.100-1

11. Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449.
12. Raji, I. D., Smart, A., White, R. N., et al. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 33-44. https://doi.org/10.1145/3351095.3372873
13. Rolnick, D., Donti, P. L., Kaack, L. H., et al. (2022). Tackling climate change with machine learning. ACM Computing Surveys, 55(2), Article 42. https://doi.org/10.1145/3485128
14. Stanford Institute for Human-Centered Artificial Intelligence. (2025). AI Index Report 2025. Stanford University.
15. Suresh, H., & Guttag, J. V. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. Proceedings of the 2021 ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization, Article 17, 1-9.
16. United Nations. (2015). Transforming our world: The 2030 Agenda for Sustainable Development. A/RES/70/1.
17. United Nations Secretary-General's High-Level Advisory Body on Artificial Intelligence. (2024). Governing AI for humanity: Final report. United Nations.
18. United Nations Educational, Scientific and Cultural Organization. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.
19. United Nations Educational, Scientific and Cultural Organization. (2023). Guidance for generative AI in education and research. UNESCO.
20. Vinuesa, R., Azizpour, H., Leite, I., et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11, 233. https://doi.org/10.1038/s41467-019-14108-y
21. World Bank. (2021). World Development Report 2021: Data for better lives. World Bank.