MULTIDISCIPLINARY RESEARCH INSTITUTE INDIA

International Journal of Advanced Scientific Research and Engineering Trends

ISSN: 2456-0774

Volume 10 Issue 1s (2026) | Pages: 152-158

Responsible AI — AI for Social Good and Sustainability

Sonal Jamdade[1,] Sudarshana Abbad[2], Pratima Kadam[3], Reshma Abhang[4], Snehal Ranesankpal[5], Aradhana Thorat[6]

[1,2,3,5]Dept. Of CSE, BVDUCOEP, India
[4,6]Dept. Of CSBS, BVDUCOEP, India

Peer Review Information

Type: *Article*
Received: *02 April 2026*
Revised: *28 April 2026*
Accepted: *05 June 2026*
Published: *23 June 2026*

Abstract

Artificial Intelligence (AI) has become a revolutionary technology that can cope with challenging issues of society and the environment. Nevertheless, the swift implementation of AI systems also triggered the ethical, social, and environmental issues, such as biases in the algorithm, a lack of transparency, uneven access, and a high level of energy usage. Responsible AI focuses on the development, design, and implementation of ethical, accountable, transparent, inclusive, and sustainable AI systems. In this paper, the author suggests a detailed model of Responsible AI to Social Good and Sustainability that incorporates the ethical governance, fairness, explainability, stakeholder involvement, and environment-awareness. Our review focuses on the existing literature covering AI use in social good, covering healthcare, environmental monitoring, smart cities, and community welfare and explains technical and socio-ethical issues. The framework brings in the multi-layered structure that includes data governance, model construction, deployment, impact assessment, and feedback process to guarantee quantifiable social and ecological results. Moreover, we discuss evaluation metrics, case studies, and strategies of implementation on how AI can be aligned with the United Nations Sustainable Development Goals (SDGs). Lastly, we list challenges, research gaps, and directions to go in order to develop Responsible AI as a useful working tool towards benefiting society, keeping it ecologically sustainable, and cordially advancing technology.

Keywords: Responsible AI; AI for Social Good; Sustainability; Ethical AI; Fairness; Explainability; Sustainable Development Goals.

How to Cite This Article

Jamdade, S., Abbad, S., Kadam, P., Abhang, R., Ranesankpal, S., & Thorat, A. (2026). Responsible AI — AI for social good and sustainability. *International Journal of Advanced Scientific Research and Engineering Trends*, 10(1s), 152–158.

Introduction

Introduction Artificial Intel The idea of ligence (AI) has not only been developed as but evolved as a rather computational technology, but has become a transformative force that affects virtually every aspect of human life. AI systems hold promise like never before in the fields of healthcare diagnostics and precision agriculture, as well as in smart transportation and climate modeling, to enhance the welfare of the population and environmental sustainability. Simultaneously, the swift growth of AI technologies has brought up important ethical, social, and environmentally critical issues. The concerns, including the bias of the algorithms, in transparency, and unfairness in decision-making, unfair access, and excessive energy consumption have highlighted the necessity of responsible design, development, and implementation of AI systems. Responsible AI is an approach to making AI systems ethical, accountable, transparent, inclusive, and sustainable, as well as according to the principles of technological progress to social values and environmental objectives. It does not only prioritize the technical correctness and efficiency of AI models but rather, the overall effect of AI on people, communities and the environment. Responsible AI seeks to ensure that the social benefits are as positive as possible and the possible harm is reduced by integrating ethical principles and sustainability goals into each step of the AI lifecycle, such as data collection and model training, deployment and monitoring. The idea of Responsible AI is especially applicable to AI for Social Good (AI4SG) that aims at applying AI to solving urgent global issues, including health inequity, climate change, poverty, disaster response, and equitable access to education. It has been demonstrated that although AI can create enormous social goodness, much of the current implementation lacks inclusiveness, lacks transparency, and resource wastage, which demonstrates the necessity of a single framework that can be systematically used to address these issues. This paper outlines a detailed model of Responsible AI to Social Good and Sustainability, which incorporates ethical guidelines, technical measures, and quantifiable indicators of impact. The framework aims to: Be fair and inclusive and AI systems must not reinforce social inequalities that already exist in society. Ensure that it is transparent and explainable so that stakeholders can be aware of how AI decisions are made. Integrate environmental sustainability, the reduction of the energy and resource footprint of AI systems. Support accountability and governance, so that organizations and policymakers can monitor, audit, and direct AI deployments. This paper discusses the current methods of using AI to support a social good and sustainability, research gaps, issues, and best practices through a thorough literature review. We also look at areas of application, such as healthcare, environmental surveillance, and smart cities, where Responsible AI can produce quantifiable social and ecological impacts. Lastly, the article addresses the evaluation measures, implementation plans, and the future research, in an attempt to give a feasible road-map of creating AI-based systems that are socially responsible and environmentally friendly

Literature Survey

It discusses the literature review on AI to benefit society, ethical AI systems, and AI sustainability that establish major themes, gaps, and trends.

1. AI for Social Good (AI4SG) An extensive index of AI to Social Good (AI4SG) lists more than 1,000 donated research papers and notes a tremendous rise in AI research in the context of societal issues in such domains as healthcare, transportation, environmental sustainability, and education. According to this literature, the application of AI can be found across predictive health models to the optimization of the public services, demonstrating the increasing pace of AI in the greater good. It, however, also recognizes such shared obstacles as a lack of data, privacy issues, and the necessity of long-term approaches to deployment. Luciano Floridi and others suggest seven key points of AI-system design that should really support social good and environmental safety by indicating that it should have clear ethical objectives, stakeholder engagement, and long-term assessment.

2. AI and Sustainable Development Goals (SDGs). AI has been directly aligning with the United Nations Sustainable Development Goals (SDGs), and it has been applied in several SDGs, such as health, education, climate action, and sustainable cities. Nevertheless, according to the analysis by experts, the deployment of AI is usually centred on well-resourced settings and that AI scaling to generate impactful SDG progress is still a concern because of data quality, talent shortages, and funding constraints. An extensive empirical study of academic literature explains that AI is now being used to apply to SDG-related studies, but a significant number of studies do not combine advanced AI methods with specific sustainability knowledge, which suggests the necessity of the interdisciplinary research framework.

3. Ethical Artificial Intelligence and Sustainability Frameworks. The field of AI ethics is a rich field of research that overlaps with sustainable development. The last systematic reviews of principles of responsible AI indicate that ethical frameworks focus on fairness, accountability, explainability, privacy, and inclusiveness as the basis of the sustainable AI implementation. A different study systematically maps the available AI ethics -based sustainability frameworks and toolkits and reports that mainstream standards and guidelines (of IEEE, OECD, GDPR etc.) discuss most common ethical concepts in detail but implementation is patchy.

4. Domain-Specific and Participatory Perspectives. Participatory AI methods are also highlighted by research, whereby the needs and contributions of community's influence what is considered as a social good, and to which the questions of equity, and contextual relevance, are answered. The involvement of stakeholders at the early stage improves the legitimacy and relevance of AI systems that are implemented in solving societal problems.

Responsible AI Framework

Corporate Social Responsibility AI Framework. To make Responsible AI (RAI) operational to serve social good and sustainability, we suggest a multi-layered framework, which incorporates the values of ethics, technical design, environment, and quantifiable social impact. The framework holds the AI systems ethical, accountable, transparent, inclusive, and sustainable during their lifecycle.

1. Framework Overview The framework has five interconnected layers: Data Governance Layer, Model Design Layer, Deployment and Integration Layer, Impact Evaluation Layer, Feedback and Governance Layer The framework will make AI systems responsible and have a positive social and environmental impact through all of its layers.

2. Framework Layers 1) Data Governance Layer Purpose: Secure input data of high quality, representative and ethical source. Critical Functions: Bias detection and mitigation, Privacy security (differential privacy, federated learning), Secure data storage and ethical data collection practices, Impact: Prevents the transmission of the past inequity of AI systems or strengthens discrimination. 2) Model Design Layer Mission: Build AI models that are responsible, interpretable, energy-saving, and sustainability-oriented. Important Uses Fairness-conscious algorithms (demographic parity, equal opportunity), Explainable AI (XAI) methods to make decisions interpretable Green AI strategies to reduce carbon footprint in training Impact: Fair AI models are technically sound, socially just, and sustainable. 3) Deployment and Integration Layer. Purpose: Adopt AI systems in a responsible way in the real world setting. Important Functions: Ongoing monitoring of ethical and performance indicators Risk management (identification of unintended harms or ethical drift) Interaction with the existing socio-technical systems to align with the stakeholders Impact: Provides assurances of safe and accountable and context-sensitive deployment. 4) Impact Evaluation Layer Purpose: Indicate AI results on social good and sustainability goals. Important Functions Social impact KPI definition (e.g., access improvement, equity index) Environmental metrics (energy consumption, CO_2 footprint) Fairness and transparency audits. Impact: Gives quantitative information about responsible AI performance and makes a decision. 5) Feedback & Governance Layer Purpose: make continuous improvement, stakeholder engagement, and policy compliance a reality. Significant Responsibilities: User, community, and regulatory feedback Loops User, community, and regulatory feedback loops Ethical, regulatory compliance checks (GDPR, AI Act, SDGs alignment) Model Updating Model updates over lessons learned Impact Ensures sustainable long-term AI systems, accountable and adaptive.

3. Key Advantages of the Framework Ethical Alignment: This will assure fairness, privacy, and human rights of AI. Sustainability Integration: Minimizes carbon footprint and is SDG-aligned. Stakeholder Participation: Involves the communities and regulators to achieve social legitimacy. Continuous Improvement: Feedback is continuous and it makes sure that it adapts to new challenges. Measurable Impact: Social and environmental KPIs can be assessed quantitatively. D. Ethical and Legal Compliance Problems. Unstable and changing laws: AI laws are constantly evolving and present organizations with the need to keep adapting. Interdisciplinary complexity: Ethical and legal issues frequently cut across technology, law, social sciences and environmental policy. Context-related factors: There are many differences between ethical risks in various fields and geographies; solutions that are universal do not work. Compliance measurement: Ethical compliance and societal impact is a complication that needs to be measured using standardized KPIs and audits.

Healthcare and Public Health

AI can reshape healthcare provision, increase access to care, and increase patient outcomes.

- Applications: Predictive diagnostics: AI models can be used to identify diseases at an early stage based on medical imaging and patient records, e.g. cancer, diabetes, or heart diseases. Individualized treatment planning: The algorithms suggest individual treatment according to the patient history, genomics, and clinical information.
- Resource optimization: AI forecasts the number of patients arriving at hospitals, which assists in streamlining the number of people, equipment, or medication to be used. Epidemic forecasting An AI model predicts the outbreak of a disease, allowing to respond in time and distribute resources.
- Responsible AI Impact: Minimizes healthcare inequality by providing access to predictive tools equally. Provides privacy assurance of sensitive patient information with federated learning and differential privacy. Conforms to SDG 3 (Good Health and Well-being).
- Sample: The case of the AI used by Google DeepMind to diagnose eye diseases in underserved territories proves to be very accurate without violating patients' privacy.

Environmental Sustainability

AI is important in environmental monitoring, climatic modelling, and management of natural resources, which results in environmental sustainability. Applications: Climate modeling and prediction: AI analyses satellite images and weather to forecast extreme weather conditions, Pollution and mitigation: Real-time air, water, and soil quality monitoring to guide policy changes, Resource optimization: AI optimizes energy consumption in industries, smart grids, and buildings, and lower its carbon footprint. Wildlife conservation IoT cameras and sensors can be used to track endangered species, identify poaching activity and guard habitats with AI.

Smart Cities and Urban Planning

The AI facilitates effective, sustainable, and inclusive urbanization, improving the quality of life and resources. Improves sustainability and environmental performance in the city. Provides fair urban facility accessibility to people. Conforms to SDG 11 (Sustainable Cities and Communities). E.g. Barcelona employs AI-driven sensors and analytics to optimize the use of public lighting, waste management, and traffic flow and decrease energy use and emissions.

Humanitarian Aid and Disaster Management

AI can enhance both preparedness and response to disasters, which will save lives and decrease economic damage. Promotes fair disaster recovery with the needy in the hierarchy. Improves social strength and decreases environmental and economic consequences. Such as SDG 1 (No Poverty), SDG 11 and SDG 13. Case Study: Flood predictors in India are AI-powered to send alerts to local authorities in real-time and reduce the amount of human and economic resources wastage.

Education and Skill Development

AI will facilitate accessible, personalized, and inclusive education, which will assist in balancing social and digital disparities. Applications: Lessen inequality in education, offer equal access to learning resources, Upholds privacy, and ethical use of student data, Affirms SDG 4 (Quality Education). Example: Coursera and Khan Academy are platforms that use AI to suggest courses and recommend them based on the profile of the learners and their engagement.

Cross-Domain Integration

Multi-domain responsible AI is typically a blend of fields to enhance social good. For instance: Smart agriculture is a combination of AI, IoT, and weather modeling that enhances crop production and minimizes water and pesticide application in an effort to improve SDG 2 (Zero Hunger) and SDG 12 (Responsible Consumption). Healthcare-environment integration: AI forecasts the risk of diseases caused by pollution, providing environmental sustainability with the results of the health of the population.

Measurements and Accountability of Responsible AI

Responsible AI (RAI) has to evaluate and measure to make sure that AI systems are ethical, sustainable, and effective. Devoid of clear metrics, AI can generate technically sound outputs and not social good or environmental utility. This section provides the major dimensions, methodologies and quantitative/qualitative measures of assessing the RAI systems.

1. Dimensions of Evaluation: The responsible AI evaluation measures cut across four major dimensions: Social Impact Metrics, Environmental Sustainability Metrics, Ethical, and Fairness Metrics. Technical Performance Metrics, The dimension is a requirement to facilitate a holistic assessment of the AI that goes beyond conventional accuracy metrics.

2. Social Impact Metrics: AI systems that are responsible are meant to enhance the welfare of the society, equity and access to services. The important measures are: Equity and Inclusion Metrics. Demographic parity: Determines how well the outputs of AI are evenly spread among various population groups of people. Disparity ratio: Comparisons the results of advantaged and disadvantaged groups. Accessibility index: The index measures the accessibility of AI-enabled services based on the geographic or socioeconomic groups. Societal Benefit Metrics Coverage Percentage of target population positively impacted by AI interventions (e.g., number of patients benefited by healthcare AI). Adoption rate: Indicates the level of user interaction and AI solution adoption within communities. Outcome improvement: Measures the change in the social indicators (e.g. fewer hospital read missions, better educational outcomes). User and Community Feedback Scores or Stakeholder Satisfaction Surveys. Believes in, content, and sense of fairness. Example: To assess the effectiveness of an AI-based tele medicine platform, the proportion of rural patients served, waiting time reduction, and patient satisfaction can be used.

3. Environmentally Friendly Sustainability Measures: The cost of the model development and deployment to the environment should be taken into account in responsible AI. Key metrics include: Energy Consumption Quantifies the amount of energy (kWH) consumed by the model training process, fine-tuning and inference. Allows the juxtaposition of green AI models to traditional models. Carbon Footprint Estimates CO 2 equivalent emissions of the AI computation. Buyers promote optimization of algorithms, cloud computing performance, and equipment selection. Resource Efficiency The metrics consist of memory footprint, processing time and hardware utilization. Trade-offs between environmental sustainability and model accuracy. Example, the Carbon-Aware AI framework created by Google Research is a data collection of energy energy used at each training run, and advocates scheduling computations to run at times when renewable energy is available.

4. Fairness and Ethics Metrics: Ethical metrics make AI systems transparent, responsible and impartial. Common metrics include: Bias and Fairness Statistical parity: The rate of prediction between groups. Equalized odds: This is used to guarantee equal false-positive and false-negative rates between groups. Counterfactual fairness: Tests whether manipulating sensitive characteristics (gender, race) has an outcome difference. Best Practices- explainability and Transparency Importance in features coverage: Measures the contribution of important features to the predictions. Local interpretability: Determine the extent to which XAI methods (SHAP, LIME) are able to clarify individual decisions.

Auditability: Measures the ease with which AI-based decisions are traceable, understandable, or verifiable. Accountability This is the presence of human-in-the-loop mechanisms. Proper records of decision making and assigning responsibilities. Example A loan approval system can use fairness measures, which assess if minority populations experience the same rates of approvals, and explainability measures, which measure whether a user can comprehend why a loan was granted or denied.

5. Built-in Evaluation Framework. All these metrics are multi-dimensionally aggregated into dashboards of a comprehensive assessment of Responsible AI: Dimension Examples Metrics Purpose Social Impact Coverage, adoption rate, equity index Make AI advantageous to communities. Environmental Energy use, CO 2 footprint, resource efficiency Minimize environmental impact. Ethical/Fairness Equalized odds, demographic parity, explainability Make responsible and understandable choices. Technical Reliability, scaleability, maintainability Accuracy, robustness, scalability Maintain reliability and efficiency. The evaluation can also take place through multi-stakeholder evaluations, such as policymakers, end-users, NGOs, and domain experts to make sure that AI interventions are in line with societal and environmental goals.

Difficulties and Constraints of Responsible AI

Although Responsible AI (RAI) has a tremendous potential to enhance a social good and sustainability, a number of technical, ethical, social, and operational issues are preventing its mass adoption. These challenges are essential to consider in order to develop viable, secure and efficient AI systems.

1. Technical Challenges Quality and Availability of Data. High quality, representative and diverse datasets are important in AI systems. Data in most aspects of social and environmental areas is limited, noisy, or biased and, therefore, can deliver inaccurate or unfair results. Example: AI in healthcare in low-resource areas might fail to have adequate patient data that results in non-population generalization models. Complexity and explainability of Algorithms. The current models of AI and Deep Learning in particular are usually opaque, which requires interpretation of decisions or debugging of failures. Absence of transparency may slow down trust, regulatory conformity and accountability in crucial applications. Energy and Computational Constraints. Training and deploying AI models especially in large scale social or environmental models consume large amounts of computation resources which contributes to growing carbon emissions and environmental impact. The lack of green or energy efficient infrastructure in certain areas can also improve the concern of sustainability. Generalization and Robustness. Artificial intelligence models can be effective in making predictions in a controlled setting but not in the real world because of the shift in distribution, dirty input, or malicious conditions. Scenario: AI used in flood prediction might not work during unprecedented climatic events because it is trained on the historical data.

2. Ethical Challenges Bias and Discrimination Training data may have historical or societal prejudices resulting in systemic discrimination of AI results particularly among marginalized persons. It is not an easy task to mitigate bias because there is a trade-off between being fair and accurate and putting into consideration the multi-stakeholder views. Accountability and Transparency. Most AI systems do not have trailings on decision-making, and it is sometimes challenging to hold anyone accountable in case of any harm. It is still difficult to provide meaningful human supervision without harming efficiency. Privacy and Consent The privacy rights may be violated depending on the collection of sensitive personal or community data to be used in AI applications. Privacy-sensitive methods, such as differentiated privacy or federated learning, can also cause technical complexity and decrease the model accuracy.

3. Legal and Regulatory issues. Fragmented Legal Landscape Regulation of AI is taking varied forms in different countries, posing compliance related issues to multinational AI implementations. They can be the GDPR in Europe, CCPA in the United States, and the nascent AI Acts in Asia. Liability and Accountability Problems It is not certain how AI decisions will be legally responsible, particularly when the outcomes are damaging and unforeseen. Autonomous systems, e.g., medical AI, create a question of accountability between developers, deployers, or end-users. Intellectual Property Intellectual property may infringe on ownership, making it difficult to protect the AI-created products and innovate.

4. Social and Operational Problems. Trust and Adoption Mistrust, lack of understanding, or job displacement are some of the reasons why communities are resistant to the use of AI. Development of trust involves transparency, stakeholder involvement and explicit benefits and risk communication. Digital Divide Inequality in digital infrastructure and AI literacy may increase disparities, restricting the access of Responsible AI to underserved groups. Fitting into Existing Systems Responsible deployment of AI can be costly and time-intensive and may need changes to the legacy systems, staff training, and interoperability.

5. Environment and Green and Sustainability Issues. High Energy Consumption Massive electricity is consumed by large-scale AI models, which are a source of carbon emissions. Model design and green data centers related to energy efficiency are not universally available at present. Trade-offs Between Accuracy and Sustainability-By minimizing energy use, predictive accuracy may be minimized by simplifying models, and this may lead to a trade-off between performance and environmental responsibility.

6. Weaknesses of present Responsible AI Methods. Lack of Standardization No universal system and criterion of determining Responsible AI. Indicators of fairness, sustainability and social impact differ substantially in application with cross-comparison challenging. Scalability Challenges A large number of RAI strategies are pilot-level or prototypical and it is difficult to scale them to national or international implementation. Minor Multi-Stakeholder Engagement. The design of ethical AIs does not tend to include the end-users, marginalized populations, or local authorities, which makes society less relevant and equitable. Problem in Impact Measurement. Measuring social good and sustainability is naturally complicated, and proxy metrics might not be able to address long-term consequencesing areas.

Future Research

Directions Although Responsible AI (RAI) has gone a long way into advancing social good, ethical compliance and sustainability, the gaps and opportunities in research are still critical. To do the job in the future, there should be a technical strength, moral responsibility, environmental sustainability and social acceptance.

1. Advanced Ethical Artificial Intelligence Research. Fairness Metrics that is context-aware. The existing measures of fairness do not usually reflect the intricate social settings. The current study ought to be extended in the future by the creation of dynamic measures of fairness that must take into account historical, socio-economic, and cultural backgrounds. Case in point: Modifying the indicators of loan approval equity to reflect the economic differences across regions. Elucidability in Complex Models. Deep learning models are becoming applied in high-stakes applications, but are not transparent. The future research ought to be on scalable, human-interpretable explanations of large AI models, as this will make its decisions accessible and challengeable to its stakeholders. Ethical Decision-Making Frameworks It is a frontier to create AI systems that are capable of ethical reasoning and value-dependent decisions. Future studies need to consider hybrid approaches to AI models based on symbolic reasoning and machine learning to conform AI judgments to human values.

2. Improving Data Protection and Control. Privacy-Preserving Learning Federated and Privacy-Preserving Learning AI can use sensitive data to train its models without privacy loss. The future studies should consider the strong, scalable federated learning systems that can balance the performance of models, fairness, and privacy. Procedures of Data collection and data curation, Bias-resistant Data collection and curation techniques must be developed and enhanced in order to allow automated recognition and correction of dataset biases. Subsequent directions of work ought to be synthetic data generation, augmenting datasets, and sampling based on fairness. Tracing Data Provenance Research on traceable and auditable data pipelines will increase accountability enabling regulators and communities to confirm AI decision-making processes.

3. Sustainability Research in the Environment. Green AI and energy efficient models. To sustain AI models in the long term, it is important that the effects on the environment are reduced. The studies are to be done on model compression and quantization and adaptive computation methods to minimize the use of energy. Lifecycle Assessment of AI Systems Complete systems to assess the overall environmental impact of AI, both in terms of training and deployment, are required. The application of carbon footprint monitoring tools and energy optimization algorithms to the AI pipelines can be incorporated into future research. Artificial Intelligence in Climate Change Resilience and Adaptation. Creating AI, capable of forecasting, countering, and adapting to environmental issues, including severe weather conditions, the lack of water, and deforestation. The ecological sustainability can be improved through cross-disciplinary studies where AI, climate science, and social sciences have been integrated.

4. Social and Human Research. Inclusive AI Design Future studies need to engage the marginalized and underrepresented groups in the process of designing, testing, and controlling AI. Social relevancy, trust and adoption of AI solutions can be enhanced using participatory approaches. Human-in-the-Loop Systems The study of human judgment-based decision systems based on the application of AI technology to high-stakes problems. The automation and human control so that accountability and ethics are taken care of. AI Literacy and Trust Building Increasing the public awareness about AI by educating, being transparent, and using interactive tools. The effective communication methods that reduce fear, promote adoption, and build trust in the society should be studied in future.

5. Regulatory, Legal and Governance Research. Responsible AI Metrics Research should be standardized to standardize metrics of fairness, social impact, sustainability and accountability. This allows inter-domain benchmarking, auditing and regulatory compliance. Comparative Analysis of International AI Regulations Comparative research on AI regulations across the world (GDPR, EU AI Act, AI ethics guidelines) can be used to inform best practices on legal compliance. Setting up adaptive and flexible governance models that adapt to technological and social changes. Liability and Accountability Models The legal frameworks of AI liability, especially of autonomous and high-risk systems, are not adequately developed. In future research efforts, it is possible to investigate mechanisms to distribute responsibility among developers, deployers and end-users.

6. Multidisciplinary and Multidomestic Research. Integration Across Sectors It can be said that AI applications frequently entail cross-domain integration, e.g., connecting the healthcare, climate, and urban planning AI systems. Future studies must be on AI frameworks that are interoperable enabling the movement of data and insights that are safe and responsible across sectors. Longitudinal Effective-Impact

Studies. It is important to evaluate the long-term social and environmental impact of AI to be able to understand systemic effects. Longitudinal data collection, causal inference, and impact assessment frameworks could be included in the future work. Ethical-Conscious AI in New technology. The principles of responsible AI must be extended to the new technology, such as AI-based cars, robots, IoT, and blockchain. The possibility of AI being ethical, transparent, and sustainable in the rapidly changing fields should be studied.

Conclusion

The increase in the use of Artificial Intelligence (AI) opens new possibilities of solving complicated issues in society and environment. Yet, in their absence, AI systems may perpetuate inequality, support bias, erode privacy, and amount to environmental degradation. This makes it clear that Responsible AI (RAI) is now more than ever needed; a process that incorporates ethical, social, legal, and environmental values in the AI life-cycle so that the growth of technologies does not come at the cost of the societal and ecological well-being. This paper has provided the general overview of what Responsible AI is regarding social good and sustainability, its principles, frameworks, applications, evaluation metrics, challenges, and future research directions. We suggested the multi-layered Responsible AI model whereby: Data Governance Layer: Entails the provision of quality data, reduction of bias, privacy, and ethical source. Model Design Layer: Fair, explainable, energy-efficient models that are in line with the sustainability objectives. Deployment & Integration Layer: Responsible deployment and implementation of AI with monitoring, risk management and alignment with stakeholders. Impact Evaluation Layer: The means of assessing both social and environmental impacts using both quantitative and qualitative measures. Feedback & Governance Layer: Ensuring stakeholder involvement, the enhancement of the process, and adherence to ethical and legal requirements. Within this framework, AI systems can be designed, implemented, and evaluated to support United Nations Sustainable Development Goals (SDGs) across healthcare, education, environmental management, disaster response, and smart cities. Responsible AI enables predictive diagnostics, patient-centered care, and equitable healthcare access. It supports climate modeling, pollution monitoring, and resource optimization, while enhancing traffic management, energy efficiency, urban planning, and community security in smart cities. The application of AI to disaster management increases early warning, resource distribution and real-time loss estimation. Within the education sector, AI offers personalized education, skill evaluation, and on-demand education opportunities to all. The Responsible AI needs to be evaluated using multi-dimensional measures that consider social impact, environmental sustainability, ethical compliance and technical performance. Such measures make sure that AI systems do not only provide accuracy and efficiency but also have real societal and environmental value. In spite of its potential, Responsible AI has some serious challenges and limitations, among which there are technical (data quality, model transparency, energy consumption), ethical (bias, privacy, accountability), legal (regulations, liability, intellectual property), and operational (trust, digital divide, integration with legacy systems) ones. The solution to these challenges requires interdisciplinary studies, involvement of stakeholders and strong governance systems. In future, the research directions will be: Creating contextualized fairness metrics, explainable and ethics conscious AI systems. Improving privacy-preserving and bias-resistant data algorithms and safe federated learning systems. Minimizing the environmental impact of AI by using green AI, energy-saving computer algorithms, and lifecycle analyses. Making AI human-centered and inclusive in order to guarantee equal benefits. Homogenisation of evaluation measures, systems of governance and regulation compliance across sectors and regions. Exploration of cross-domain integration and long-term societal effect, making AI deal with the highly complicated, interconnected global problems. To sum up, Responsible AI provides an innovative solution to use artificial intelligence in social benefit and sustainability. Through a systematic integration of ethical standards, stakeholder involvement, environmental awareness, and quantifiable appraisals in the design and implementation of AI, one can tailor the benefits of society to the minimal number of harms, such that AI can become an important part of a fairer, more sustainable, and stronger society.

References

1. Z. R. Shi, C. Wang, and F. Fang, Artificial Intelligence for Social Good: A Survey, arXiv preprint, 2020.
2. L. Floridi et al., “How to Design AI for Social Good: Seven Essential Factors,” Sci. Eng. Ethics, 26, pp. 1771–1796, Apr. 2020.
3. Sustainable AI: AI for sustainability and the sustainability of AI, AI & Ethics, vol. 1, pp. 213–218, Feb. 2021.
4. Z. Chen et al., “A Survey on AI Sustainability: Emerging Trends on Learning Algorithms and Research Challenges,” arXiv, 2022.
5. L. Cheng, K. R. Varshney, and H. Liu, “Socially Responsible AI Algorithms: Issues, Purposes, and Challenges,” arXiv, 2021.
6. M. Tahaei et al., “A Systematic Literature Review of Human Centered, Ethical, and Responsible AI,” arXiv, 2023.
7. S. E. Bibri et al., “Environmentally sustainable smart cities and their converging AI, IoT, and Big Data technologies,” Energy Informatics, 2023.
8. F. Siddiqui et al., “Leveraging AI for Social Impact in Environmental Sustainability,” Int. J. Sci. Res. Sci. Eng. Technol., vol. 12, no. 4, pp. 253–266, Jul. 2025.
9. AI ethics and sustainability: Accelerating paradigm shifts toward sustainable development, J. Strategy & Innovation, 2025.
10. Towards AI ethics led sustainability frameworks and toolkits: Review and research agenda, J. Sustainable Finance and Accounting, 2024.