Policy Research Journal
ISSN (E): 3006-7030 ISSN (P) : 3006-7022 Volume 4, Issue 3, 2026

ARTIFICIAL INTELLIGENCE FOR ENVIRONMENTAL SUSTAINABILITY: CONCEPTUAL FOUNDATIONS, THEORETICAL PERSPECTIVES, AND A FRAMEWORK FOR BALANCED ASSESSMENT

Dr. Ifra Iftikhar[1], Prof. Dr. Sajjad Ahmad Paracha[2], Dr. Irem Sultana[*3]

[1]Associate Professor, Department of Media & Communication Studies, Lahore Garrison University
[2]Chairperson Department of Mass Communication, University of Southern Punjab, Multan
*[*3]Assistant Professor, Department of Mass Communication, Government College University Faisalabad*

[1]ifraiftikhar@lgu.edu.pk, [2]drsajadprachamcstudies@usp.edu.pk, [*3]iremsultana@gcuf.edu.pk

DOI: http://doi.org/10.5281/zenodo.21449295

Keywords
artificial intelligence, environmental sustainability, Green AI, ecological modernization, socio-technical systems, governance

Article History
Received: 31 January 2026
Accepted: 16 March 2026
Published: 30 March 2026

Copyright @Author

Corresponding Author: *
Dr. Irem Sultana

Abstract
Artificial intelligence (AI) is increasingly positioned as a critical instrument for addressing environmental degradation, climate change, and resource depletion, while simultaneously being implicated as a growing source of energy consumption, carbon emissions, and material extraction. This conceptual review critically examines AI's dual role as both an environmental solution and an environmental burden, integrating theoretical perspectives from environmental studies, science and technology studies, sociology, economics, and governance scholarship. Rather than cataloguing technical applications, the review foregrounds how foundational theories, including ecological modernization theory, socio-technical systems theory, the planetary boundaries framework, circular economy theory, the triple bottom line, diffusion of innovation theory, responsible innovation theory, and risk society theory, help explain why AI's environmental value is conditional rather than automatic. The review synthesizes recent empirical and conceptual literature on AI's applications across climate, energy, agriculture, biodiversity, water, waste, and urban systems, alongside literature documenting AI's energy, carbon, water, and material footprints. It concludes that AI functions as a capable but conditional sustainability tool whose net environmental value depends on governance quality, measurement rigor, and the extent to which efficiency gains are prevented from generating rebound effects. A conceptual framework linking AI technologies, environmental applications, governance mechanisms, and sustainable development outcomes is proposed to guide future interdisciplinary research.

1. Introduction

Humanity is confronted with a convergence of environmental problems that, separately, would exhaust existing governance capacity while collectively threatening the integrity of the Earth system. The accumulation of waste and pollution, freshwater stress, biodiversity loss, and climate change are now viewed as interrelated aspects of a single global issue rather than as distinct policy issues. In light of this, artificial intelligence has become one of the most important general-purpose technologies of the twenty-first century. Its applications span almost every aspect of environmental management, from the optimization of renewable energy grids to satellite-based deforestation monitoring.

The anticipation that AI can act as a decisive lever for environmental sustainability has been raised by the rapid spread of AI, especially since the introduction of deep learning and, more recently, large-scale foundation models. AI is often described by governments, international organisations, and tech companies as essential to reaching the Sustainable Development Goals (SDGs) and climate targets. Using a consensus-based expert elicitation approach, Vinuesa et al. (2020) projected that AI might facilitate 134 of the 169 SDG targets while also impeding 59 of them. This finding encapsulates the ambivalence present in the larger body of research. Therefore, it is best to think about AI as a technology whose environmental effects depend on how it is developed, used, and regulated rather than as a clear-cut answer.

A increasing volume of research that views AI's interaction to the environment as essentially two-sided has emerged as a result of this ambivalence. Van Wynsberghe (2021) formalised this concept by distinguishing between "sustainability of AI," which refers to the environmental cost of creating, training, and implementing AI systems, and "AI for sustainability," which refers to the application of AI to promote sustainable development. Since then, reviews have consistently verified that evaluating AI's environmental value requires taking into account the energy, water, and material intensity of the AI systems used to provide those applications in addition to the applications themselves (Bolón-Canedo et al., 2024; Verdecchia et al., 2023).

This review has three goals. First, by differentiating words that overlap, like Sustainable AI, Green AI, AI for Sustainability, and Responsible AI, it defines the conceptual landscape surrounding AI and environmental sustainability. Second, it looks at how fundamental theories from sociology, technology studies, environmental studies, and governance research shed light on the circumstances in which AI influences environmental outcomes in a favourable, unfavourable, or unclear way. Third, it offers a conceptual framework that connects AI technology, environmental results, and governance mechanisms by synthesising previous research on AI's environmental applications and costs.

This review is significant because it aims to shift the conversation from a straightforward list of AI applications to a more critical, theoretically based explanation of why AI's environmental impact is contingent. The empirical data on AI's overall environmental impact is actually conflicting, as the sections that follow demonstrate. While cross-country panel data indicates that AI development is generally linked to smaller ecological footprints and quicker energy transitions (Wang et al., 2024), the same literature also reveals rebound effects, measurement gaps, and governance flaws that can undermine or undo these benefits (Dauvergne, 2020). The main objective of this review is to comprehend this circumstance rather than to reach a straightforward conclusion.

2. Conceptual Foundations

The terminology used in the literature on AI and the environment are similar but not the same. Because each phrase suggests a distinct normative commitment and analytical focus, it is important to use them precisely.

Artificial Intelligence (AI) refers to computing systems capable of executing activities that traditionally require human cognition, including perception, prediction, categorisation, and decision-making, most commonly through machine learning models trained on massive datasets. Environmental Sustainability signifies the potential of human systems to meet existing requirements without damaging the ecological integrity, resource base, or adaptive capacity available to future generations. The literature has created a number of intermediate constructs to describe the junction between these two ideas.

The largest of these concepts is "sustainable AI," which Van Wynsberghe (2021) defines as a movement to refocus the entire AI lifecycle toward social justice and ecological integrity, from idea generation to governance. Rather than handling the two aspects of the AI-environment relationship independently, it purposefully includes both. Bolón-Canedo et al. (2024) break down Green AI into "green-in-AI," which is the design of energy-efficient algorithms and hardware, and "green-by-AI," which is the application of AI to support environmentally beneficial outcomes in other sectors. Green AI is a more specific and technical concept that focuses on lowering the computational and energy costs of AI systems themselves. AI for Sustainability (or AI for Good/AI for Earth in policy discourse) refers to the application-oriented use of AI to advance environmental or developmental goals, largely without reference to AI's own footprint; this is the framing implicit in most sectoral reviews of AI in energy, agriculture, or conservation. As demonstrated by studies like Rolnick et al. (2022) on machine learning applications throughout the climate value chain, climate AI is a domain-specific subset focused on mitigation and adaptation to climate change. Environmental Intelligence is a wider, more applied phrase describing AI-enabled sensing, monitoring, and decision-support systems for ecological management. Responsible AI is the phrase most frequently used in policy and industry contexts (e.g., OECD, UNESCO); it encompasses fairness, transparency, accountability, and safety in addition to environmental issues. It is more expansive than sustainable AI. Digital Sustainability, lastly, situates AI within the wider environmental footprint of digital infrastructure, including networks, data centers, and end-user devices, treating AI as one component of a larger sociotechnical system whose sustainability must be analysed holistically.

Table 1. Conceptual Landscape of AI and Environmental Sustainability

Concept	Primary Focus	Key Reference
Sustainable AI	Whole AI lifecycle; both AI-for and sustainability-of AI	Van Wynsberghe (2021)
Green AI	Energy/computational efficiency of AI itself (green-in) and AI-enabled environmental solutions (green-by)	Bolón-Canedo et al. (2024); Verdecchia et al. (2023)
AI for Sustainability	Application of AI to environmental/development goals	Nishant et al. (2020); Vinuesa et al. (2020)
Climate AI	AI applications specific to climate mitigation/adaptation	Rolnick et al. (2022)
Environmental Intelligence	AI-enabled sensing, monitoring, decision support	Nishant et al. (2020)
Responsible AI	Fairness, transparency, accountability, safety, and sustainability	Policy discourse (OECD, UNESCO)
Digital Sustainability	Footprint of the wider digital/AI infrastructure stack	Dauvergne (2020)

Table 1 distinguishes overlapping concepts frequently conflated in the literature.

These ideas are not so much opposing definitions as they are alternative perspectives on the same fundamental conflict: while AI can be utilised to further environmental objectives, the AI systems that do so have their own material and energy costs. The theoretical discussion that follows is motivated by this tension.

3. Foundational Theories

An explanation of AI and the environment that is solely application-based runs the risk of becoming descriptive rather than explanatory. In order to illustrate why AI's environmental contribution is conditional rather than automatic, this part draws on eight theoretical traditions that were chosen for their combined coverage of the technological, social, economic, and governance aspects of the AI-environment interaction. Every theory is analysed for its history, underlying presumptions, applicability to artificial intelligence and environmental sustainability, advantages, disadvantages, and potential for theoretical growth.

3.1 Ecological Modernization Theory

Mol and Spaargaren (2000) developed ecological modernisation theory, which had its roots in European environmental sociology in the 1980s. They contended that industrial modernisation and environmental protection are not mutually exclusive but rather that institutional and technological innovation can "decouple" economic growth from environmental degradation. Applied to AI, ecological modernisation theory defines AI-enabled efficiency gains, in energy grids, industrial processes, or logistics, as instances of reflexive modernisation in which technology corrects the ecological adverse effects of prior industrial expansion. Its power rests in explaining why governments and companies so readily adopt an optimistic, techno-solutionist narrative around AI: the notion legitimises continuous technological and economic development while promising environmental improvement. Its main drawback, which is well-documented in the AI literature, is that it undervalues rebound effects and fails to sufficiently account for situations in which efficiency gains are absorbed by higher consumption rather than being translated into absolute decreases in resource use (Dauvergne, 2020). Future theoretical developments may incorporate degrowth or post-growth perspectives with ecological modernisation to better explain situations in which decoupling does not occur at the necessary size or speed.

3.2 Socio-Technical Systems Theory and Science and Technology Studies (STS)

Socio-technical systems theory, developed by scholars including Bijker, Hughes, and Pinch (1987) within the broader field of science and technology studies, holds that technologies cannot be understood in isolation from the social, institutional, and infrastructural arrangements in which they are embedded. When applied to AI, this viewpoint maintains that an algorithm's environmental impact is never solely defined by its technical design; rather, it develops from the interaction of data infrastructures, energy grids, organisational practices, and user behaviour. This framing is especially useful for explaining why technically similar AI systems can produce different environmental outcomes depending on where and how they are deployed, such as an AI-optimized irrigation system embedded in a well-managed water utility versus the same system operating without adequate institutional oversight. One of the strengths of the theory is that it does not treat technology as a neutral, self-contained variable. On the other hand, the theory has a tendency toward descriptive complexity, which can make it difficult to make predictions that are generalisable. A promising addition is a closer connection with actor-network theory (Latour, 2005), which adds non-human elements like algorithms, sensors, datasets, and causal agency to these systems. This gives us a fuller picture of how AI infrastructures create environmental outcomes along with human actors.

3.3 Planetary Boundaries Framework

According to the Planetary Boundaries Framework created by Rockström et al. (2009), mankind must operate within nine essential Earth-system boundaries, including climate change, freshwater use, biodiversity loss, and land-system change. These constraints must be respected to keep Earth stable and habitable.

The environmental impact of artificial intelligence can be assessed using this methodology. Instead of assessing if AI just enhances efficiency, it tackles a more fundamental question: Does AI contribute to keeping human activities within Earth's ecological limits? Thus, reducing energy use or emissions in one area does not make an AI system environmentally sustainable. Its overall influence should help maintain global human pressure within safe environmental limits.

Most AI research for environmental sustainability quantify success by energy efficiency or emissions per unit of production. Although these gains are important, they do not show if AI is reducing environmental strain sufficiently to stay within Earth's ecological limits. AI can become more efficient as resource use and emissions rise. This indicates a major literature limitation.

Its scientific underpinning and capacity to shift focus from tiny, incremental gains to long-term environmental sustainability make the Planetary Boundaries Framework powerful. It assesses whether technology advancement improves Earth system health and resilience, not just optimisation. However, implementing this approach to AI is difficult. AI is employed in various industries, making it hard to quantify its impact on environmental issues including carbon emissions, energy use, water use, and resource extraction. AI's effects are scattered across many applications, hence there is no universally acknowledged way for tying AI to planetary limits.

Recently published Green AI research suggests adding a digital infrastructure indicator to the Planetary Boundaries Framework to solve this constraint. This metric would evaluate AI systems' environmental footprint, encompassing computational energy, hardware materials, and digital infrastructure resource demands (Bolón-Canedo et al., 2024). An overall assessment of whether AI contributes to environmental sustainability within Earth's ecological boundaries would be possible.

3.4 Circular Economy Theory

Circular economy theory, synthesised by Ghisellini, Cialani, and Ulgiati (2016), advocates replacing the linear take-make-dispose model of production with closed-loop systems that minimise resource extraction and waste through reuse, remanufacturing, and recycling. The concept of circular economy intersects with artificial intelligence in two different ways: first, AI has the potential to enable circularity in other industries, such as through predictive maintenance, waste-sorting robotics, and materials-flow optimisation; second, AI hardware itself is a candidate for circular design due to its reliance on rare earth elements and its contribution to electronic waste. The theory's strength is its systemic, lifetime orientation, which opposes the inclination to assess AI applications separately from their downstream disposal and upstream material origins. Its limitation is that circular economy models were developed primarily for physical manufacturing systems and require adaptation to capture the specific dynamics of compute-intensive digital technologies, where the dominant environmental costs arise from energy consumption during use rather than from material flows alone. A helpful theoretical expansion makes a clear connection between hardware lifespan assessment for AI accelerators and data-center infrastructure and the concepts of the circular economy.

3.5 Triple Bottom Line Theory

According to Elkington's (1997) triple bottom line theory, an organization's performance should be assessed in terms of economic profit, social justice, and environmental preservation (sometimes known as the "three Ps": profit, people, and planet). The triple bottom line, when applied to AI, offers a helpful remedy for purely techno-economic assessments of AI systems, which often claim cost savings or efficiency benefits without matching social or environmental accounting. It also contributes to the explanation of corporate incentives surrounding the deployment of AI: companies that use AI for sustainability purposes typically do so when economic and reputational benefits coincide with environmental claims. This dynamic fosters greenwashing when environmental performance is claimed but not independently confirmed. One of the strengths of the theory is that it provides a structure for multi-criteria evaluation that is both intuitive and widely accepted. However, the primary limitation of the theory is that it does not provide a mechanism for resolving trade-offs in situations where the three dimensions are in conflict. For example, when an artificial intelligence system reduces environmental impact but eliminates labour or concentrates economic gains. Integrating the triple bottom line with stakeholder theory provides a partial solution by identifying which interests should be prioritised in such trade-offs.

3.6 Diffusion of Innovation Theory

Rogers (2003) developed the diffusion of innovation theory, which emphasises elements including relative benefit, compatibility, complexity, trialability, and observability to explain how and why new technologies spread throughout a community over time. Applied to environmental AI, the theory helps explain uneven global adoption: AI-enabled precision agriculture or smart-grid technologies diffuse rapidly where infrastructure, capital, and skilled labour are available, but adoption lags in lower-income regions and smaller enterprises, reproducing existing digital divides. This has obvious sustainability consequences, since the countries and sectors most exposed to climate risk are usually those with the least potential to embrace AI-based mitigation and adaptation tools. The theory's ability to forecast adoption trends and pay attention to institutional and social factors that go beyond technical performance are its strongest points. The theory's strength is its predictive description of adoption dynamics, as well as its attention to social and institutional aspects other than sheer technical performance. Its main shortcoming is that it was designed to explain adoption rather than environmental consequences, and it offers little about what happens after diffusion, such as rebound effects or unequal distribution of environmental benefits and disadvantages. Combining diffusion of innovation theory and environmental justice theory provides one approach to closing this gap.

3.7 Responsible Innovation Theory

According to Stilgoe, Owen, and Macnaghten's (2013) responsible innovation theory, innovation processes should be anticipatory, reflexive, inclusive, and responsive, incorporating ethical and societal considerations throughout technology development rather than treating them as an afterthought. Applied to AI and the environment, the theory underlies proposals for lifecycle environmental assessment, stakeholder consultation, and adaptive governance to be integrated into AI development from the outset rather than patched once systems are implemented at scale. Its strength is its process orientation, delivering practical suggestions on how organisations could operationalise sustainability commitments throughout design rather than merely following deployment. One of its limitations is that responsible innovation frameworks have traditionally been applied more to biotechnology and nanotechnology than to artificial intelligence. Furthermore, it is still a challenge to translate the principles of these frameworks to the fast-paced, compute-intensive, and frequently opaque development cycles of contemporary AI, particularly foundation models. Specifically for foundation-model development, where lifecycle decisions are made by a small number of firms with outsized aggregate environmental influence, an important extension would be to formalise environmental anticipatory governance. This would be a significant example of an extension.

3.8 Environmental Governance Theory and Risk Society Theory

The environmental governance theory investigates the institutions, rules, and actors which encompass the state, the market, and civil society that shape collective responses to environmental problems. On the other hand, the risk society theory, which was developed by Beck (1992), contends that late-modern societies are increasingly organised around the management of manufactured risks that escape traditional forms of accountability. These viewpoints, when taken together, are extremely pertinent to artificial intelligence, which poses environmental hazards that are innovative, dispersed, and not effectively represented by existing regulatory categories. These risks include energy demand spikes, opaque supply chains, and algorithmic judgement errors in environmental management. Specifically, Nishant et al. (2020) claim that the most significant long-term environmental value of artificial intelligence may not lie in the direct optimisation of technical processes, but rather in its ability to boost environmental governance itself. This is accomplished through enhanced monitoring, enforcement, and policy formulation abilities. A limitation of this theoretical pairing is that governance theories frequently lag behind the pace of technological change, which results in regulatory frameworks, such as early AI legislation, that struggle to keep up with rapidly evolving foundation models and their environmental implications. Therefore, the strength of this theoretical pairing is that it focuses on institutional capacity and accountability rather than technology performance alone. The creation of anticipatory and adaptive governance models that are specifically tailored for rapidly evolving general-purpose technologies is an extension that has the potential to be fruitful.

Table 2. Comparative Summary of Foundational Theories

Theory	Core Assumption	Key Strength	Key Limitation
Ecological Modernization	Technology can decouple growth from degradation	Explains techno-optimist policy narratives	Underplays rebound effects
Socio-Technical Systems / STS	Technology is embedded in social-institutional context	Avoids technological determinism	Limits generalizable prediction
Planetary Boundaries	Earth-system processes have absolute safe limits	Provides absolute, science-based benchmark	Hard to attribute to a diffuse general-purpose technology
Circular Economy	Closed-loop systems minimize extraction and waste	Lifecycle, systemic orientation	Built for material flows, not compute/energy use
Triple Bottom Line	Performance spans economic, social, environmental axes	Multi-criteria evaluation structure	No mechanism to resolve trade-offs
Diffusion of Innovation	Adoption follows predictable social/technical factors	Explains uneven global AI adoption	Silent on post-adoption environmental consequences
Responsible Innovation	Ethics should be embedded throughout innovation	Actionable, process-oriented guidance	Poorly adapted to fast-moving foundation models
Environmental Governance / Risk Society	Manufactured risks require new accountability forms	Centers institutional capacity and accountability	Regulation lags technological pace

IEER Institute for Excellence in Education & Research

Table 2 compares the eight theoretical traditions applied to AI and environmental sustainability in this review.

When taken as a whole, these theories provide support for a single proposition: the contribution that artificial intelligence makes to the environment is not a characteristic that is inherent to the technology itself, but rather an emergent result of the social, institutional, and governance systems that it is embedded within. This premise serves as the foundation for the rest of the literature review, which begins with an examination of the environmental applications and costs of artificial intelligence (AI) and then returns, in Section 6, to the circumstances in which the benefits of AI are expected to outweigh its drawbacks.

4. AI as a Tool for Environmental Sustainability: Application Domains

The empirical and review literature on AI's applications in the environment is centred on four recurrent functional roles rather than a single list of sectors. These roles are as follows: forecasting and prediction, system optimisation, remote sensing and monitoring, and decision support (Nishant et al., 2020; Rolnick et al., 2022). It appears that the environmental value of artificial intelligence is not derived from sector-specific originality but rather from its generic capacity to extract patterns from massive, complex, and rapidly changing environmental data streams. This is because these functions are repeated across a variety of domains, including climate, energy, agriculture, biodiversity, water, waste, and urban areas.

4.1 Climate Change Mitigation and Energy Systems

Applications in the energy sector are the most developed area of environmental AI. In order to facilitate a more profound integration of low-carbon electricity sources, machine learning is utilised to forecast the generation of renewable energy, to balance the intermittent supply of solar and wind energy against the demand for it, and to optimise the operation of smart grids (Rolnick et al., 2022). Cross-country panel evidence is among the most robust quantitative support in this body of research. Wang, Li, and Li (2024), who analysed data from 67 countries between 1993 and 2019, discovered that higher levels of artificial intelligence development were associated with lower ecological footprints and carbon emissions, as well as with faster energy transitions, with the largest estimated effect on energy transition itself. Nevertheless, this effect was not uniform; it was more pronounced in contexts with greater trade openness and higher existing AI development, and it varied with industrial structure. This suggests that the climate benefits of artificial intelligence are moderated by pre-existing economic and institutional conditions rather than solely arising from the adoption of AI.

4.2 Agriculture, Biodiversity, and Forest Management

In the field of agriculture, artificial intelligence (AI) provides support for precision farming by means of computer-vision-based crop and pest monitoring, yield prediction, and optimised irrigation and fertiliser application. This allows for significant reductions in input use while simultaneously maintaining or improving yields. When it comes to the conservation of biodiversity, AI, when paired with remote sensing and acoustic or camera-trap sensor networks, makes it possible to identify species, detect poaching, and track land-use change and deforestation at scales that are not accessible through manual survey approaches. Applications for forest management extend this skill to include the prediction of wildfire danger and the detection of unlawful logging. This is a reflection of the ecological and geographic diversity of the underlying challenges, as described by reviews across different domains, which define biodiversity and conservation applications as being methodologically varied in comparison to the more standardised energy-sector applications (Nishant et al., 2020).

4.3 Water Resource Management, Waste, and Circular Economy

The applications of artificial intelligence in water management include the optimisation of treatment processes, the identification of leaks in distribution networks, and the predictive modelling of water quality and demand. Optimising the recovery of resources and increasing the lifespans of equipment through predictive maintenance are two ways that predictive models assist the objectives of the circular economy. Computer-vision-based sorting systems in trash management increase recycling accuracy. Carbon accounting has also evolved as a separate application domain, with artificial intelligence being used to automate the estimation and verification of greenhouse gas emissions across corporate supply chains. This is a function that is becoming increasingly tied to environmental, social, and governance (ESG) reporting standards.

4.4 Smart Cities, Transportation, and Environmental Monitoring

Applications of artificial intelligence in urban settings include traffic optimisation, building energy management, and air quality monitoring. These applications are frequently included into larger "smart city" platforms that are formed with the intention of reducing the environmental imprint of urban systems. Examples of applications that contribute to sustainable transportation include the optimisation of routes for logistics fleets and the development of control systems for electric and autonomous cars with the goal of lowering emissions. The ability of artificial intelligence to handle satellite and sensor data at scale is essential to environmental monitoring in a broader sense. This capability enables AI to support applications ranging from monitoring air quality to predicting natural disasters, such as flood, landslip, and severe weather forecasting and early-warning systems. The increasing frequency of catastrophic weather occurrences is reflected in the fact that studies have shown that artificial intelligence-based catastrophe prediction and climate-resilient urban planning are among the sub-domains of environmental AI research that are expanding at the quickest rate.

There is a common criterion that can be found in the literature across all four areas, and that is that technological feasibility does not guarantee that there will be a net environmental benefit. On the other hand, effectiveness is contingent upon the quality of the data, the interpretability of the model, the availability of the infrastructure, and the degree to which operational performance measures are tied to genuine sustainability outcomes rather than being considered as proxies for them. This distinction will be discussed in further detail in the following section.

5. Environmental Costs and Risks of AI

Artificial intelligence is not necessarily beneficial to the environment, and a significant body of critical literature has evolved specifically to examine the environmental costs associated with it. These expenses are incurred at each and every stage of the artificial intelligence lifecycle, beginning with the fabrication of hardware and continuing through the training of models, deployment, and in the end, disposal.

5.1 Energy Consumption and Carbon Emissions

The process of training large machine learning models requires a significant amount of energy. In an early study that has been widely cited, Strubell, Ganesh, and McCallum (2019) demonstrated that training certain large natural language processing models could generate carbon emissions that are comparable to the lifetime emissions of several automobiles. This study has drawn sustained attention to the carbon cost of model training. Bolón-Canedo et al. (2024) note that selecting more efficient architectures can reduce energy use by large margins without comparable loss of accuracy, highlighting the fact that artificial intelligence's energy intensity is a design choice rather than an inevitability. Subsequent research has demonstrated that the choices made regarding model architecture can have a significant impact on this footprint. Particularly in light of the fact that foundation models are being implemented to massive user bases, inference, which is the recurring use of trained models at scale, is being widely recognised as a source of aggregate energy consumption that is equivalent to or even larger than training.

5.2 Water Consumption, Hardware, and Material Extraction

For the purpose of cooling, data centers that support artificial intelligence training and inference require a significant amount of water. This expense is typically externalised to local water systems and is unevenly distributed geographically, frequently harming regions that are already experiencing water stress. Artificial intelligence gear, which includes specialised accelerators, is dependent on rare earth elements and other key minerals. The extraction of these minerals comes with considerable ecological and societal consequences, which is reminiscent of criticisms that have been levelled against the material basis of digital technology in general for a long time (Crawford, 2021). This cost is rarely reflected into assessments of artificial intelligence's environmental performance, which tend to focus narrowly on operational energy use rather than complete lifecycle impact. The rapid obsolescence cycle of AI-specific hardware also leads to increased amounts of electronic waste, which is another factor that contributes to the development of electronic waste.

5.3 Rebound Effects and Environmental Inequality

Perhaps the most conceptually significant risk that has been reported in this body of research is the rebound effect. This phenomenon describes a situation in which the efficiency gains that result from the use of AI are cancelled out or possibly even reversed by increased consumption. According to Dauvergne (2020), artificial intelligence-driven advertising and recommendation systems actively boost consumption. Furthermore, it is frequently the case that efficiency advantages gained in one segment of a supply chain are reinvested into expanding production rather than being translated into absolute savings in resource use. It is important to avoid making the mistake of comparing advances in efficiency at the micro level with environmental benefits at the macro level. This dynamic is reminiscent of the classic Jevons paradoxs. When a technology becomes more efficient, people often use it more, which can increase total resource consumption instead of reducing it. The problem of environmental inequality is connected to this issue. The advantages of AI-enabled sustainability tools and the environmental costs of AI infrastructure are not evenly distributed. Data centers, mineral extraction, and the disposal of electronic waste have a disproportionate impact on communities and regions with lower incomes. On the other hand, the efficiency benefits of AI are primarily accrued to users and businesses with higher incomes.

5.4 Greenwashing, Governance Failure, and the Digital Divide

Greenwashing is a practice in which environmental claims concerning artificial intelligence systems are stated without proper verification. The absence of standardised metrics that can be independently verified for the environmental footprint of artificial intelligence creates settings that are conducive to greenwashing. Nishant et al. (2020) identify overreliance on historical training data, uncertain behavioural responses to AI-based interventions, cybersecurity vulnerabilities, and, most importantly, difficulty in measuring the actual effects of AI interventions as recurring methodological barriers across the field. This risk is compounded by weak governance, which is a cause for concern. Finally, the digital divide, which refers to unequal access to the infrastructure, capital, and expertise required to deploy environmental artificial intelligence, means that the regions and communities that are most susceptible to the effects of climate change are frequently the least able to adopt AI-based adaptation and mitigation tools. This reinforces rather than reduces the environmental and economic disparities that already exist.

6. Balancing Benefits and Costs

Sections 4 and 5 present an empirical picture that is neither consistently positive nor consistently negative. On the one hand, country-level data correlates AI development with reduced ecological footprints, fewer carbon emissions, and speedier energy transitions (Wang et al., 2024), and sector-specific assessments record large, well-documented advances in monitoring, forecasting, and optimisation. On the other side, the same body of work details increased energy and water demand from AI infrastructure, material extraction costs, rebound effects, and governance failures that can counteract or reverse these gains (Dauvergne, 2020; Strubell et al., 2019).

Economically, AI's environmental applications generate productivity advantages and cost savings that create significant market incentives for adoption, but these same motivations might encourage superficial or inflated environmental claims where verification is inadequate. Surveilling the environment and making the best use of resources with AI can help with disaster preparedness and service delivery, especially in places with limited data. However, some groups may be able to get these benefits more easily than others. From the point of view of governance, artificial intelligence has the potential to boost environmental institutions by enhancing the evidence foundation for policy and enforcement (Nishant et al., 2020). However, the rate of development of AI frequently exceeds the capacity of regulatory frameworks to manage the environmental threats that it poses. The trade-off between AI's potential for the environment and its resource footprint brings up questions of fairness and intergenerational justice that are similar to and, in some ways, more heated than long-running debates in environmental ethics about the pros and cons of technological progress.

The theoretical discussion in Section 3 suggests that these trade-offs are not resolved by technology choice alone. Ecological modernisation theory correctly identifies AI's potential for decoupling economic activity from environmental harm, but only under specific conditions: when efficiency gains are not absorbed by rebound consumption, when clean electricity is available to power AI infrastructure, and when governance mechanisms translate technical potential into verified outcomes. Socio-technical systems theory maintains this conditionality by stressing that the same algorithm might create varied outcomes depending on its institutional setting. The planetary boundaries framework adds a further, more demanding requirement: that AI's net contribution be assessed not only in relative, efficiency-based terms but against absolute ecological thresholds. Taken together, the literature and theory reviewed here support a conditional rather than categorical conclusion: AI contributes positively to environmental sustainability where high-quality data, explicit sustainability metrics, clean energy infrastructure, and robust governance are jointly present, and its contribution is uncertain or negative where these conditions are absent.

7. Emerging Trends

The link between artificial intelligence and the environment is being reshaped by a number of technologies. According to Rolnick et al. (2022), foundation models and generative artificial intelligence have significantly increased the scale of computation that is required for both training and inference. This has led to an intensification of concerns regarding aggregate energy demand. This is occurring at the same time that these same models are increasingly being applied to climate modelling, materials discovery, and Earth observation tasks. In parallel, the Green Artificial Intelligence movement has developed from a specialised technical concern into a well-established research field. Verdecchia, Sallou, and Cruz (2023) conducted a comprehensive review of ninety-eight studies and discovered that the majority of Green AI research is still narrowly focused on the model-training stage. Furthermore, the involvement of industry remains limited. This suggests that the field, despite its growth, has not yet reached the level of practical

Several new AI technologies are helping reduce the environmental impact of AI itself. There are two noteworthy cases. Instead of transferring data to big, centralised data centers for processing, these technologies perform AI computations directly on low-power devices such as sensors, smartphones, or Internet of Things (IoT) devices. This strategy eliminates the need to transport huge amounts of data over networks, lowering energy consumption, reducing latency, and minimising the stress on energy-intensive data centers.

A similar approach is federated learning, which allows AI models to be trained across multiple local devices without transferring raw data to a central server. By dispersing computation over multiple devices, federated learning can cut communication costs, increase data privacy, and lessen the energy demands of centralised computing.

The Growing digital trend is the utilisation of the use of digital twins. A digital twin is a virtual representation of a real-world object, system, or environment that can be used to simulate, monitor, and optimise its performance. In environmental applications, digital twins are being employed for climate modelling, urban planning, energy management, and infrastructure design. A more advanced variant, known as climate digital twins, builds highly comprehensive virtual simulations of the Earth's climate system, enabling researchers and policymakers to test different climatic scenarios and evaluate the possible effects of environmental policies before enacting them.

AI is also altering carbon markets. Machine learning algorithms are enhancing the accuracy of measuring, reporting, and verifying greenhouse gas emissions, making carbon accounting more dependable and transparent. AI can also automate parts of the carbon credit certification and trading process, helping improve the efficiency and credibility of carbon markets. At the same time, AI-powered tools are becoming an integral part of Environmental, Social, and Governance (ESG) reporting by enabling organizations to collect, analyze, and disclose sustainability data more effectively.

Lastly, Earth observation is one of the most well-known uses of AI in environmental sustainability. Earth Observation AI can identify and track changes in forests, oceans, agricultural land, glaciers, biodiversity, and urban areas with previously unheard-of speed and accuracy by fusing satellite pictures with machine learning. This technology allows governments and researchers to track environmental change over large geographic areas and over long periods, making it one of the most reliable and widely adopted applications of AI for environmental monitoring.

8. Governance, Policy, and Regulation

The literature consistently identifies governance as the key factor that distinguishes AI deployments that actually benefit the environment from those that don't (Nishant et al., 2020). A number of recent advances in institutional and regulatory frameworks are directly relevant. Although a risk-based regulatory framework for AI systems is established under the European Union's AI Act, environmental consequences are only indirectly addressed through openness and documentation requirements rather than as a separate risk category. This gap has been the subject of ongoing criticism. Both the OECD AI Principles and the UNESCO Recommendation on the Ethics of Artificial Intelligence articulate environmental sustainability as a component of responsible AI governance. However, neither of these instruments is legally binding, which limits the extent to which they can be enforced. The International Organization for Standardisation (ISO) has established standards that are pertinent to artificial intelligence management systems and environmental management. These standards offer organisations the opportunity to adopt voluntary technical benchmarks; nevertheless, adoption is still inconsistent and mainly unconfirmed by independent auditors.

There is a need for environmental auditing of artificial intelligence systems, standardised sustainability metrics that are reported alongside model performance benchmarks, and a closer integration between AI governance and existing climate policy instruments, such as carbon pricing and disclosure regimes. This is in addition to the formal regulation that is currently existing. According to the ideas of responsible innovation and environmental governance that were described in Section 3, it is suggested that such procedures are most effective when they are incorporated early on in the development lifecycle of artificial intelligence, rather than being applied retrospectively after systems have already been deployed at scale. The development of foundation models is concentrated among a few number of companies; therefore, governance frameworks that particularly address this concentrated capacity, rather than presuming a dispersed population of independent developers, represent a policy goal that has not been thoroughly examined but is becoming increasingly essential.

9. Research Gaps

There are a number of gaps that repeat. Conceptually, the proliferation of overlapping words, Sustainable AI, Green AI, AI for Sustainability, Responsible AI, has not yet settled on a stable, shared lexicon, complicating cross-study comparison. According to Verdecchia et al. (2023), a large portion of the literature is methodologically review-based, conceptual, or case-study-driven rather than experimental, and studies often measure operational or technical performance rather than verified environmental outcomes, undermining causal claims about the true environmental benefit of AI. Theoretically, existing frameworks, including many that were discussed in Section 3, were built prior to the introduction of large-scale foundation models. These frameworks need to be modified in order to take into consideration the concentrated, compute-intensive, and fast growing character of contemporary artificial intelligence development. The primarily indirect consideration of environmental risk in important regulatory instruments like the EU AI Act and the voluntary, uneven adoption of pertinent international standards are clear examples of policy gaps. Despite the fact that the dynamics of innovation diffusion, which were discussed in Section 3, suggest that these contexts face unique adoption barriers, the empirical literature is dominated by studies from high-income countries and China. As a result, the relationship between artificial intelligence and the environment in low- and middle-income countries, which are frequently the most climate-vulnerable, is relatively underexamined. Lastly, there are still interdisciplinary gaps between environmental and social science research, which is better suited to assess systemic, long-term, and distributive outcomes, and technical AI research, which tends to optimise narrow performance metrics; closing this gap is perhaps the most important task facing the field.

10. Proposed Conceptual Framework

This section proposes a conceptual framework that links artificial intelligence technologies to sustainable development outcomes through five interconnected components. These components are as follows: (1) AI technologies and infrastructure; (2) environmental applications; (3) governance mechanisms and ethical safeguards; (4) stakeholder interactions; and (5) net environmental and developmental outcomes. This section is a synthesis of the theoretical and empirical material that was reviewed earlier.

At the foundation of the framework, artificial intelligence technologies and infrastructure, which include model architectures, hardware, and data systems, generate environmental applications (the functions of forecasting, optimisation, monitoring, and decision-support that are catalogued in Section 4) as well as direct environmental costs (the burdens of energy, water, material, and waste that are catalogued in Section 5). These two output streams are not independent; as ecological modernisation and socio-technical systems theory jointly imply, the same underlying artificial intelligence infrastructure can generate predominantly positive or predominantly negative net effects depending on the governance mechanisms and ethical safeguards that mediate its deployment. These safeguards include environmental auditing, lifecycle assessment, regulatory oversight, and responsible-innovation practices that are embedded during design rather than added after implementation.

A dynamic that is explained by stakeholder and environmental justice considerations is that governance mechanisms do not operate solely on technology; rather, they operate through stakeholder interactions among developers, deploying organisations, regulators, affected communities, and civil society. The relative power and access of these stakeholders shapes whose interests are reflected in how artificial intelligence systems are designed and evaluated. The final component of the framework, which is the net environmental and developmental outcomes, is purposefully not treated as a fixed endpoint but rather as a feedback variable. Verified outcomes (or the absence of them) inform subsequent

governance adjustments and stakeholder demands, thereby creating an iterative cycle rather than a linear pipeline from technology to impact.
There are two assertions made by this framework that represent a departure from more traditional accounts that are focused on applications. First, it sees environmental applications and environmental costs as concurrent, co-occurring outputs of the same technical basis, rather than treating them as different categories that need to be weighed independently. Second, it views governance and stakeholder interaction not as external limits on an otherwise technological process, but rather as inherent determinants of whether the net effect of artificial intelligence is good or bad. This is in line with the conditional conclusion that was reached in Section 6. In the future, empirical research could be conducted to operationalise this framework by determining whether the presence of particular governance mechanisms, such as mandatory environmental disclosure or independent lifecycle auditing, is associated with measurable improvements in net environmental outcomes across comparable AI deployments.

11. Conclusion

The purpose of this analysis was to analyse artificial intelligence as a tool for environmental sustainability via a perspective that was purposefully speculative. The review argued that the environmental usefulness of AI is conditional rather than intrinsic. Artificial intelligence demonstrates a large and increasingly well-documented capability in forecasting, optimisation, monitoring, and decision support across a wide range of domains, including climate, energy, agriculture, biodiversity, water, waste, and infrastructure systems. However, the costs of artificial intelligence in terms of energy, water, materials, and governance are not insignificant, nor are they automatically outweighed by the environmental applications of AI. Rebound effects, measurement gaps, and unequal access can erode or reverse the benefits that were predicted (Dauvergne, 2020; Verdecchia et al., 2023).
There is a common understanding that can be found among the eight different theoretical traditions that are discussed in Section 3. These traditions include ecological modernisation, socio-technical systems, planetary boundaries, circular economy, triple bottom line, diffusion of innovation, responsible innovation, environmental governance and risk society theory. The contribution that artificial intelligence makes to the environment is not a result of the technology itself, but rather of the interaction between technology and the institutional, economic, and social systems that surround it. This realisation has direct repercussions for the four communities that are actively involved in relation to this topic. It recommends that researchers should go beyond technical performance measurements and toward verifiable, standardised environmental outcome measures. Additionally, they should move toward theoretical frameworks that are expressly fitted to the scope and concentration of contemporary foundation-model development. Researchers should take this into consideration. This highlights the importance for policymakers to incorporate environmental risk into regulatory frameworks for artificial intelligence in a more explicit manner, rather than in an indirect manner, and to close the gap between voluntary technological standards and accountability mechanisms that are enforceable. Because of the established risk of greenwashing in this sector, it suggests that genuine environmental claims for artificial intelligence systems require independent verification rather than self-reported efficiency measurements. This is a significant implication for the industry. It argues that the distribution of artificial intelligence's environmental costs and benefits across geographies, income groups, and generations needs at least as much examination as its overall technical performance. This means that it is important for society to take this into consideration.
Therefore, artificial intelligence should be understood as both an opportunity and a challenge for environmental sustainability. It is a technology that has the potential to significantly enhance humanity's capacity to monitor, forecast, and manage environmental systems. However, the ultimate contribution of this technology will be

determined less by its computational sophistication and more by the quality of the governance, measurement, and institutional arrangements within which it is embedded.

REFERENCES

Beck, U., & Ritter, M. (1994). Risk Society: Towards a New Modernity.

Bijker, W., Hughes, T. P., & Pinch, T. (1989). The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology. https://doi.org/10.2307/3105993Bolón-Canedo, V., Morán-Fernández, L., Cancela, B., & Alonso-Betanzos, A. (2024). A review of green artificial intelligence: Towards a more sustainable future. Neurocomputing, 599, 128096. https://doi.org/10.1016/j.neucom.2024.128096

Crawford, K. (2021). Atlas of AI: Power, Politics and the Planetary Costs of Artificial Intelligence. Perspectives on Science and Christian Faith. https://doi.org/10.12987/9780300252392

Dauvergne, P. (2020). Is artificial intelligence greening global supply chains? Exposing the political economy of environmental costs. Review of International Political Economy, 29, 696 - 718. https://doi.org/10.1080/09692290.2020.1814381

Elkington, J. (2006). Governance for Sustainability. Corporate Governance: An International Review, 14, 522-529. https://doi.org/10.1111/j.1467-8683.2006.00527.x

Ghisellini, P., Cialani, C., & Ulgiati, S. (2016). A review on circular economy: The expected transition to a balanced interplay of environmental and economic systems. Journal of Cleaner Production, 114, 11–32. https://doi.org/10.1016/j.jclepro.2015.09.007

Latour, B. (2005). Reassembling the social: An introduction to actor-network-theory. Oxford University Press.

Mol, A. P. J., & Spaargaren, G. (2000). Ecological modernisation theory in debate: A review. Environmental Politics, 9(1), 17–49. https://doi.org/10.1080/09644010008414511

Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. Int. J. Inf. Manag., 53, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F. S., III, Lambin, E. F., Lenton, T. M., Scheffer, M., Folke, C., Schellnhuber, H. J., Nykvist, B., de Wit, C. A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P. K., Costanza, R., Svedin, U., … Foley, J. A. (2009). A safe operating space for humanity. Nature, 461(7263), 472–475. https://doi.org/10.1038/461472a

Rogers, E. M. (2003). Diffusion of innovations (5th ed.). Free Press.

Rolnick, D., Donti, P., Kaack, L., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A., Maharaj, T., Sherwin, E. D., Mukkavilli, S., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., . . . Bengio, Y. (2019). Tackling Climate Change with Machine Learning. ACM Computing Surveys (CSUR), 55, 1 - 96. https://doi.org/10.1145/3485128

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation*. The Ethics of Nanotechnology, Geoengineering and Clean Energy. https://doi.org/10.1016/j.respol.2013.05.008

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and Policy Considerations for Deep Learning in NLP. ArXiv, abs/1906.02243. https://doi.org/10.18653/v1/p19-1355Van Wynsberghe, A. (2021). Sustainable AI: AI for sustainability and the sustainability of AI. AI and Ethics, 1, 213 - 218. https://doi.org/10.1007/s43681-021-00043-6

Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of Green AI. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13. https://doi.org/10.1002/widm.1507

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature Communications, 11, Article 233. https://doi.org/10.1038/s41467-019-14108-y

Wang, Q., Li, Y., & Li, R. (2024). Ecological footprints, carbon emissions, and energy transitions: the impact of artificial intelligence (AI). Humanities and Social Sciences Communications, 11. https://doi.org/10.1057/s41599-024-03520-5

IEER
Institute for Excellence in Education & Research