Sustaining the Status Quo? Rethinking AI for Sustainability through Power and Infrastructure

Harry Collins
University of Cambridge

Artificial Intelligence (AI) is increasingly presented as a tool for addressing environmental crises, yet there remains significant debate over what should count as genuine AI for Sustainability. I argue that dominant accounts remain too narrow because they focus on environmental outputs while neglecting the political and social conditions through which those outputs are produced. Drawing on Falk and van Wynsberghe's framework and Lehuedé's analysis of data infrastructure in Chile, I show that sustainability claims can obscure power concentration, extractive relations, and territorial exclusion. I propose a revised AI for Sustainability framework that adds conditions on benefit and harm distribution, community engagement, stronger carbon reporting and energy-efficiency improvement, and the evaluation of alternative social practices. To make the framework operational, I work through community engagement and impact reconciliation in detail, demonstrate an example application against a real-world AI project, and argue that the framework should be used through external, public assessment rather than self-certification.

Keywords: AI for Sustainability, Power Concentration, Environmental Justice, Data Infrastructure, Resource Extraction

Introduction

With carbon emissions rising, artificial intelligence (AI) has been suggested as a solution to environmental challenges (Vinuesa et al., 2020), and many AI systems branded as "AI for Sustainability" have been developed to this end (Falk and van Wynsberghe, 2024). However, there is also a growing body of work into the sustainability of AI itself, including environmental and social elements like energy consumption, extractive supply chains, labour, and expanding computational infrastructure. This raises the question of how to determine which AI systems can truly be considered AI for Sustainability, reflecting genuine environmental benefit, and when they instead serve to legitimise concentrated forms of power. I argue that dominant AI for Sustainability frameworks remain incomplete because they do not adequately address AI's political and social impacts, particularly power concentration and the distribution of environmental benefits and harms. Even where an AI system appears to generate positive environmental outcomes, those outcomes may depend on infrastructures that concentrate benefits among powerful firms and states while distributing environmental and social costs onto marginalised communities elsewhere. Sustainability claims can therefore function as a legitimising discourse obscuring extractive relations, infrastructure dependency, and control over environmental decision-making.

This article's contribution is to develop a demanding AI for Sustainability framework where environmental performance is inseparable from infrastructural accountability, referring to the public tracing and justification of the infrastructures, supply chains, governance arrangements, and community relations that make an AI project's environmental benefits possible. The framework therefore evaluates AI projects claiming sustainability benefits across the AI supply chain rather than the final model exclusively. To develop this argument, I centre my analysis around two important interventions. Section 1 focuses on Falk and van Wynsberghe's (2024) tiered framework for defining AI for Sustainability. I argue that the framework is too narrow, missing the political and social impacts of AI, particularly regarding benefits and harm distribution across AI infrastructures. Section 2 expands on this, turning to Lehuedé (2022), who reveals how data infrastructure in Chile threatens Indigenous practices that sustain their environment. By analysing Falk and van Wynsberghe (2024) alongside Lehuedé (2022), I argue that AI for Sustainability should be re-centred around power concentration, resource extraction, compute dependency, and the distribution of benefits and harms. Section 3 integrates these insights into a revised framework, proposes an external rating model, and presents an example use of the framework against a real system. Methodologically, I adopt a critical-response approach. I critically reconstruct Falk and van Wynsberghe's (2024) framework, consider it through Lehuedé's (2022) analysis of data infrastructure and territorial conflict, and use this engagement to develop a revised AI for Sustainability conceptual framework.

1. Beyond Greenwashing: Analysing Falk and van Wynsberghe's Framework

1.1. Why This Paper

Falk and van Wynsberghe (2024) attempt to clarify the definition of AI for Sustainability, which they argue lacks meaning. Without consensus on this, AI systems can masquerade as AI for Sustainability despite potentially perpetuating unsustainable practices. This enables greenwashing, the "*intersection of [...] poor environmental performance and positive communication about environmental performance*," (Delmas and Burbano, 2011) like Microsoft heavily marketing AI for Sustainability projects (Microsoft Research, 2024) whilst partnering with ExxonMobil to increase oil and gas extraction (ExxonMobil, 2019). Therefore, more rigour is required to determine whether AI systems contribute to sustainability in order to enable genuinely positive contributions in the field. This paper has been selected as it offers one of the clearest recent attempts to define AI for Sustainability through a strict framework. It is an important starting point, but requires expansion from environmental output to infrastructural power.

1.2. Overview and Core Strengths

Falk and van Wynsberghe (2024) sought to develop a framework for determining which AI systems should be classified as AI for Sustainability. They define sustainability through seven environmental impact categories including climate change, chemical pollution, and change in biosphere integrity/biodiversity (Dong and Hauschild, 2017), arguing that AI for Sustainability systems must positively contribute to at least one. The authors then use necessary and sufficient conditions to tier sustainability claims. Necessary conditions must be met for an outcome to occur, whereas outcomes are guaranteed when sufficient conditions are also met. (Brennan, 2024). Falk and van Wynsberghe's framework classifies systems that provide environmental monitoring or information as AI in an application domain. If there is also a sustainability analysis of the AI system, particularly carbon reporting from model training, the system becomes AI towards Sustainability. These conditions are deemed necessary, and then if the sufficient condition is met, demonstrating a positive environmental action directly from the model, then the system is considered AI for Sustainability. This is valuable because it makes AI for Sustainability more rigorous. Distinguishing between information and action is especially useful, as AI forecasting energy demand, for instance, could help phase in renewable energy sources whilst also optimising fossil fuel imports. Thus, demonstrable environmental action is important to prevent potential usefulness being treated as actualised sustainability. The framework also brings the sustainability of AI into view by requiring a sustainability analysis.

1.3. Critical Analysis of Limitations: Distribution, Infrastructure, and Social Alternatives

Despite these strengths, defining AI for Sustainability as positively contributing to one environmental impact category misses AI's political element, particularly how the AI systems' effects are distributed. Per Mensah (2019), environmental, economic, and social sustainability are distinct but "interconnected" (2019) pillars, so simplifying AI for Sustainability to focus on environmental sustainability targets neglects social impacts and where environmental sustainability takes place. Dauvergne (2022) uses a political economy lens to argue that corporate AI for Sustainability initiatives deepen global power imbalances, with benefits realised by wealthy Western states and corporations while environmental and social costs predominantly impact the marginalised. For example, Google DeepMind used AI to schedule wind-power delivery to their electricity grids, increasing Google's wind energy usage (Witherspoon and Fadrhonc, 2019). Although this appears positive, the benefits are concentrated towards

Google rather than the planet, highlighted as they measure success in economic value for themselves without mentioning the AI system's own resource usage. Despite this, the article is tagged as "sustainability" and Falk and van Wynsberghe's (2024) framework would classify this at least as AI towards Sustainability if carbon emissions from training were reported. However, not only is this omitted, but there is no indication that any benefits of the system are felt by communities involved in Google's AI supply-chain.

AI and data infrastructure require the extraction of rare earth elements and minerals like tantalum, much of which is mined in Africa and particularly the Democratic Republic of Congo (Dauvergne, 2022). Muldoon (2023) notes that increased demand for such materials has led to "*devastating environmental and social consequences*" (2023), such as polluted water supplies. Rather than sustainability being an unconscious term, this demonstrates conscious efforts to conceal AI supply chains, like how the physical reality of wider computational infrastructure is hidden behind the Cloud (Ensmenger, 2021). Moreover, nations like DR Congo's carbon emissions are miniscule versus wealthy Western states like the USA (Ritchie and Roser, 2024), which harbour most AI projects and rewards (Maslej et al, 2024), highlighting the unjust distribution of benefits and negatives. Publications like Google's provide no evidence as to how states and communities involved in providing their AI infrastructure environmentally benefit from Western AI systems versus the environmentally destructive consequences of the infrastructure, yet understanding the distribution of impacts is missed in Falk and van Wynsberghe's (2024) framework.

A second issue is that the sustainability analysis condition is too weak if it only requires disclosure of carbon emissions from AI training. Falk and van Wynsberghe aim to establish an "easy-to-implement" (2024) method to encourage disclosure, yet disclosure alone does not necessarily improve AI sustainability. There has been significant research into improving AI algorithms' energy efficiency, with Verdecchia et al.'s (2023) literature review finding cases of energy savings over 50% versus current methods. Disappointingly, they observe minimal industry involvement in this area, despite industry's leading role in most AI research (Thompson and Ahmed, 2023). Hence, evidence that AI for Sustainability research is experimenting with these techniques to improve the sustainability of AI should be included in the sustainability analysis condition, demanding genuine improvement in the sustainability of AI.

A final limitation is the overly quantitative focus on reducing carbon emissions with technology, thus risking subscribing to climate isolationism, considering climate issues as "*an isolated, discrete, scientific problem in need of technological solutions*" (Stephens, 2022). Stephens argues that focusing on reducing carbon emissions using technology risks detracting attention from social innovations to sustainability issues whilst perpetuating existing social inequalities. Instead of accepting "inherently unsustainable" (Falk and van Wynsberghe, 2024) AI, alternative approaches and voices should be considered. Stilgoe et al.'s (2020) framework for responsible innovation includes public consultation to ensure communities' concerns are addressed (2020), whereas here, alternative voices and social innovations are missed. This is an issue because AI and climate research is dominated by a white male hegemony (Stephens, 2022). Muldoon (2023) furthers this, considering AI infrastructure as a "colonial supply chain" (2023) reinforcing global power asymmetries, promoting European modernity and techno-solutionist capitalism. This results in a techno-deterministic narrative restricting the possibility for social changes like "*community-based initiatives to support the provisioning of local [..] energy*" (Stephens, 2022) due to the prospect of socio-technical lock-in, meaning increasing dependence on technological solutions permanently entrenches their unsustainable practices (Cairns, 2014). Therefore, before committing to AI to solve a sustainability issue, alternative social innovations and voices from those who will be affected by the AI should be considered to show that AI is genuinely necessary. To this end, Lehuedé (2022) shows how alternative viewpoints about data centre expansion in Chile can reveal environmental injustices and social practices that can preserve natural ecosystems, now examined.

2. The Infrastructural Politics of AI: Analysing Lehuedé

2.1. *Why This Paper*

Lehuedé (2022) offers an alternative lens to the discussed framework. It analyses the environment through conflicting understandings of territory, showing perspectives and social practices towards the environment beyond the dominant European modernity viewpoint highlighted by Muldoon (2023). Thus, it has been selected due to foregrounding the infrastructural and territorial politics that definitional frameworks like Falk and van Wynsberghe's (2024) risk overlooking.

Two concepts need briefly defining. "Ways of worlding" (Lehuedé, 2022) refers to the differing ways people and communities understand what exists, what has agency, and how relations between humans and non-humans are organized. Assetised and relational ontologies respectively mean an understanding of territory as a set of economic resources, or as constituted through interdependencies between humans, non-humans, land, water, and other entities (Lehuedé, 2022).

2.2. *Overview and Methodological Lesson*

Lehuedé (2022) uses political ontology, which analyses conflicting ways of worlding, to examine divergences in what territory is in relation to data infrastructure expansion in Chile. Lehuedé collected empirical data through documents and interviews with stakeholders including public and private sector representatives and Indigenous Lickan Antay activists. He then conducted a discursive-material analysis of how territory is conceived and how agency is attributed to human and non-human actors. His main claim is that conflict around Chile's data infrastructure expansion is driven by two divergent ontologies of territory: a dominant assetised ontology driven by European modernity and colonialism that considers territory an economic resource, and a relational ontology from Indigenous communities where territory results from interdependencies between human and non-human actors.

The chululo-colonies episode is the clearest methodological point for AI for Sustainability. Lehuedé shows how Indigenous Lickan Antay communities have conflicted with the ALMA observatory. The communities have sustained the environment "*through practices that acknowledge mutual interdependency among the actors that make up the territory*" (2022), including life-sustaining entities like mountains that provide water. During ALMA's natural gas pipeline development, representatives declared no adverse environmental impacts, yet Lickan Antay activists visited the prospective building site and discovered overlooked chululos, saving 22 colonies. The Lickan Antay's acknowledgement of non-humans and knowledge of the territory helped preserve it, revealing harms formal assessments missed. This raises questions of whose knowledge is counted when determining environmental impacts. AI sustainability frameworks focusing only on technical output risk repeating this mistake, counting emissions, efficiency, or environmental benefits while failing to recognise the people and non-human entities most affected by the infrastructures that make those benefits possible.

2.3. *Strengths, Counterarguments, and Implications*

The paper's main strength regarding AI for Sustainability is using political ontology to highlight how relational ontologies of territory can sustain ecosystems through social practices that acknowledge non-humans' importance, and the interdependencies between all entities in a territory. Reinforcing this is wider research showing that the natural environment in areas occupied by Indigenous communities is generally preserved better than elsewhere (IPBES, 2019). Drawing attention to the colonial themes of Chile's data infrastructure also shows how this expansion occupies Indigenous lands and destroys their worlds, viewing territory as economic resources to be extracted. Concerning AI for Sustainability, these findings can help create more robust, informed criteria for sustainable AI projects to adhere to regarding AI's political and social impacts.

There are, however, places where Lehuedé's argument could be strengthened. There is little engagement with wider debates around data infrastructure despite claiming that analysing relational philosophies is crucial due to the terricide, a global phenomenon. Brodie (2023) connects Lehuedé's research to colonial dynamics in data infrastructure worldwide, so exploring whether similar ontological divergences may be occurring elsewhere could have helped guide further research. Potential counterarguments are also only lightly discussed. For instance, Lehuedé criticises the "consequentialist and quantitative logics" (2022) used to justify ALMA, yet does not engage with the counterargument that thousands of people may benefit from data infrastructure versus disrupting a small community (2022). He could have questioned who those standing to benefit are, how they will benefit, how this impacts Lickan Antay livelihoods, and who decides. Also, ethical intentions do not necessarily translate into ethical outcomes (Powell, 2022), so Lehuedé could have discussed how benefits of data infrastructure are only potential versus the immediate and actual destruction of the Lickan Antay's world.

Rather than undermine the usefulness of Lehuedé's work for my argument, these issues highlight that a revised AI for Sustainability framework must also ask who decides whether infrastructure is justified, whose knowledge is recognised, and how claimed benefits are weighed against territorial and social harms. It must also address the possibility of refusal rather than assuming that community engagement only means improving a project that has already been chosen.

3. Integrating Insights to Construct a Revised AI for Sustainability Framework

3.1. Tightening the Necessary and Sufficient Conditions

I now construct an altered AI for Sustainability framework that considers AI's political and social impacts by integrating the insights from the analysis of both central papers. The original framework's conditions remain, with additional criteria added to address the distribution of benefits and negatives of AI systems, stronger sustainability analysis, and the evaluation of alternative social innovations and practices. As the decision to use necessary and sufficient conditions is central to Falk and van Wynsberghe's framework, the relationship between the conditions and tiers needs to be explicit. In this revised framework, a necessary condition is required for a project to move beyond the lowest tier. All sufficient conditions are required jointly, together with all necessary conditions, for the strongest classification of AI for Sustainability. So, a single sufficient condition alone does not guarantee sustainability. Rather, together they are jointly sufficient conditions to collectively demonstrate that the project has moved from information and reporting to accountable environmental action.

The first tier, AI in "Application Domain", applies where a system provides monitoring or information relevant to an environmental domain. The second tier, AI towards Sustainability, applies where the system meets the necessary transparency conditions: Monitoring and Information, Benefit and Harm Distribution Assessment, and Carbon Emission Reporting and Enhancement. The third tier, AI for Sustainability, applies when the prior necessary and forthcoming sufficient conditions are met, which are Community Engagement and Impact Reconciliation, Alternative Social Perspectives and Innovations Evaluation, and Action. Thus, the strongest label is reserved for projects demonstrating environmental contribution and social accountability.

3.2. The Revised Conditions

Monitoring and Information (Necessary). This condition is retained from Falk and van Wynsberghe's (2024) framework. The AI system must generate information that could have a positive environmental impact in at least one UN environmental impact category.

Benefit and Harm Distribution Assessment (Necessary). This is a new condition. The project must make transparent the expected benefits and harms of the AI system, including who receives the benefits and who bears the costs across development, deployment, data infrastructure, mineral extraction, energy use, water use, labour, and waste. This is necessary to directly address the political element of AI missed in the original framework. Without it, a project could deliver environmental gains for a powerful organisation while displacing costs onto communities elsewhere.

Carbon Emission Reporting and Enhancement (Necessary). This expands the original sustainability analysis condition. Reporting carbon emissions from training is important, yet the condition should also include inference where possible and evidence that technical improvements to energy efficiency have been considered. Showing a deliberate effort to improve the sustainability of AI at the training and inference stages is easy-to-implement at the developer level without requiring a full understanding of the AI infrastructure. The prior iteration of this condition only included reporting carbon emissions from training and failed to demand progress despite the current unsustainable status quo of AI. The condition is therefore shifted from just disclosing emissions to showing attempts to improve on them.

Community Engagement and Impact Reconciliation (Sufficient). To meet this new condition, a project must demonstrate meaningful engagement with communities affected by the AI system, including infrastructures, and it must show that this engagement is consequential. Engagement must occur early enough to influence design, siting, deployment, mitigation, benefit-sharing, or refusal. Lehuedé's (2022) chululo example demonstrates the need for this, showing that affected communities may perceive infrastructural harms that formal experts overlook, so their knowledge must have authority within the process. As AI infrastructure can perpetuate colonial power imbalances, it is crucial to demonstrate open dialogues with communities and areas that AI systems depend on, and to accommodate their concerns.

Alternative Social Perspectives and Innovations Evaluation (Sufficient). To avoid climate isolationism, a new sufficient condition is needed to consider alternative social practices and innovations for sustaining the environment before committing to AI. Engaging with affected communities could reveal existing social practices that could be applied elsewhere, or ontologies that may be lost with the expansion of AI infrastructure. Additionally, engaging with social innovation literature like Bergman et al.'s (2010) review of climate social innovations could reveal movements like Transition Towns, which engages local communities with practices to reduce fossil fuel dependence (2010). Rather than using AI to manage energy demands in a community's smart grid and slowly phase in renewable energy, such social schemes may have similar positive local impacts without the wide-ranging negatives of a currently unsustainable technology impacting communities elsewhere. So, projects should consider both the social practices that may be lost by locking in AI infrastructure, and show that AI use is warranted over alternative social innovations and practices suggested by communities and social innovation literature.

Action (Sufficient). This final sufficient condition retains Falk and van Wynsberghe's (2024) central insight. The AI system must lead to an action that demonstrably positively contributes to an environmental impact category versus just producing information that could be beneficial.

Figure 1 shows the revised framework. The tiered structure remains, but conditions regarding the political and social impacts of AI are added to ensure that AI for Sustainability is as sustainable as possible. The sustainability of AI is now considered across multiple conditions to reflect its different impacts. If all necessary conditions are met, an AI system is considered "AI towards Sustainability", and where all necessary and sufficient conditions are met, a system can genuinely be considered "AI for Sustainability".

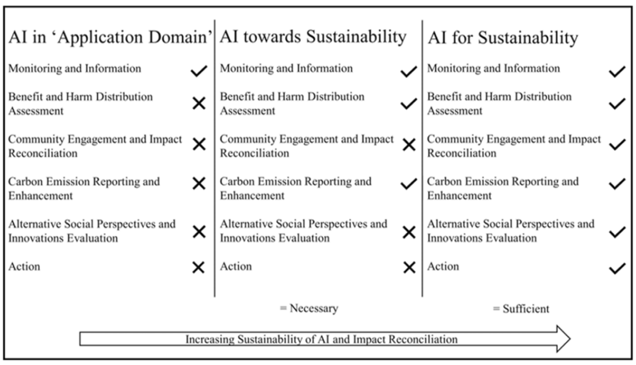

Figure 1: Tiered framework for classifying sustainable AI research (own illustration).

Under this framework, the AI for Sustainability label is reserved for systems that demonstrate environmental benefit while also accounting for infrastructural harms, community authority, and plausible non-AI alternatives.

3.3. *External Rating, Refusal, and the Limits of Checklist Governance*

This article draws on relational ontology and the possibility of refusal while also proposing a framework that, if self-applied by companies or research teams, could be reduced to an organisational checklist. Thus, it could be vulnerable to greenwashing, with organisations defining harms, judging their own engagement and certifying their own sustainability. So, the framework should be treated not as a private self-assessment but as the basis for an external, public rating model. A robust model for this is seen in the Fairwork project, which publishes annual public ratings of how fairly workers in the digital platform economy are treated (Cant et al, 2023). A similar model for AI for Sustainability would assess projects against public criteria, require documentary evidence, include community testimony, and make ratings visible, resulting in public accountability. It would also remove reliance on self-regulation and pressure organisations to perform better. If an affected community has not been meaningfully engaged, if benefits and harms are not mapped, or if a project depends on infrastructure that communities reject, the rating should reflect this. The strongest AI for Sustainability classification should therefore be unavailable where the project's environmental output depends on overriding those most exposed to its infrastructural costs. To this end, it is important to operationalise the framework's conditions in this context.

3.4. *Operationalising the Framework*

Having established the need for an external rating model, I now operationalise how the conditions should be assessed through such a model. Benefit and Harm Distribution Assessment would require a published lifecycle impact map showing where benefits and harms occur across deployment, energy and water use, mineral extraction, labour, waste, and data infrastructure, with uncertainty clearly stated rather than concealed. Plociennik et al. (2025) propose a practical approach to perform such an analysis, which includes scoping the system, creating a lifecycle inventory, assessing impacts, and interpreting results. Carbon Emission Reporting and Enhancement would require disclosure of training and inference emissions, alongside evidence of attempts to improve model and hardware efficiency beyond current standards. Alternative Social Perspectives and Innovations Evaluation would require a comparative justification showing why AI is preferable to non-AI interventions, including community-led practices, institutional reform, low-tech monitoring, or local ecological knowledge. Action would be assessed through evidence that the system has produced a demonstrable positive contribution to at least one environmental impact category, versus just generating information. Each condition would need to be supported by public evidence, independent review, and community testimony where relevant, ensuring that the AI for Sustainability label is awarded only where environmental gains are matched by infrastructural and social accountability.

Community Engagement and Impact Reconciliation as a sufficient condition requires deeper operational detail, as participatory design initiatives can often be performative, failing to truly give communities a voice. For instance, Delgado et al. (2023) argue that participatory AI design initiatives often barely consult stakeholder groups, noting that of 68 participatory AI projects studied, only 11 engaged with stakeholders multiple times (2023). With the permanence of material impacts like eradicating Indigenous worlds, performative gestures are not enough, so organisations must show that impacted communities have a meaningful voice. I therefore work this condition through concretely: what evidence satisfies it, who assesses it, and against what standard.

Table 1: Operational standard for Community Engagement and Impact Reconciliation.

Element	Evidence required	Assessor	Minimum standard
Affected-community mapping	Map communities affected by deployment, data infrastructure, energy and water use, mineral supply chains, labour, waste streams, and territorial change. Include uncertainty and explain exclusions.	Independent review body with community-nominated members.	Published impact map.
Early, repeated community engagement	Records of meetings, workshops, deliberative forums etc., and time for response before design lock-in.	Independent review body, with community representatives' sign-off	Engagement must occur before irreversible decisions.
Design influence	Change log showing how community input altered system design.	Independent review body, with community representatives' sign-off.	Consultation is insufficient without creating documented changes or enforceable constraints.
Right to refuse and grievance	Clear process for objection, refusal, appeal. Public record of unresolved objections.	Independent review body	Objections must be resolved.
Benefit-sharing and repair	Agreements covering local benefits, remediation, community ownership/stakes where appropriate.	Independent review body and community representatives.	Benefits must be tangible, locally meaningful, and revisable.

This operationalisation, summarised in Table 1, follows from the methodological lesson of Lehuedé's chululo example, that affected communities' knowledge must be considered and given authority. Community Engagement and Impact Reconciliation therefore cannot be satisfied by one-off consultations or by publishing stakeholder-engagement statements after design decisions have already been made. It requires repeated engagement, community verification, and evidence of change in response. A strong example of this can be found in the Ada Lovelace Institute's Citizens' Biometrics Council (2021). Unlike the consultations criticised by Delgado et al. (2023), the council utilised a "mini-publics" (2021) format where 50 diverse citizens who would be impacted by biometrics in the UK deliberated over several months, cross-examining experts to co-create policy recommendations for biometric AI. Thus, identifying impacted communities, adopting similar structured, repeat engagements where they are given the time, resources, and authority to influence AI system development can ensure meaningful collaboration resulting in sustainable AI. Crucially, there must also be evidence of reconciliation, meaning this participatory design should demonstrate clear change where concerns or challenges are raised, like altered siting, rejection of components, or implementation of ideas from energy democracy e.g. community stakes in local data centres and infrastructure (Szulecki, 2018).

Where a community refuses a project because its infrastructural demands threaten territory, livelihood, or world-making practices, that refusal should be treated as decisive evidence against classifying the project as AI for Sustainability. Ultimately, as Lehuedé (2022) shows, embodied practices are as important as discourse, so truly involving communities in projects could help preserve local environments and worlds by giving them a genuine stake in how AI infrastructure operates. Although this framework adds complexity, AI is highly complex in its infrastructure and in its political and social impacts throughout that infrastructure. Therefore, more complexity is warranted. The revised framework does not require perfect knowledge of every supply-chain relation before any AI project can proceed. It requires a higher evidential burden before a project can claim the strongest sustainability label. The label should be demanding because it carries legitimising power.

3.5. Applying the Framework: Google DeepMind Wind-Power Scheduling

The framework can be illustrated by applying it to the Google DeepMind wind-power scheduling example discussed earlier. Under the previous framework, the system almost met every AI for Sustainability condition. However, here it would not receive the strongest classification with existing evidence unless it publicly accounted for the AI system, its infrastructures, and its distributed impacts. Table 2 therefore shows how the framework changes the assessment.

Table 2: Applying the revised framework to Google DeepMind wind-power scheduling.

Condition	Condition met on available evidence?	Reason
Monitoring and Information (Necessary)	Yes	System generates information relevant to wind-power delivery.
Benefit and Harm Distribution Assessment (Necessary)	No	The available account only emphasises benefits to Google.
Carbon Emission Reporting and Enhancement (Necessary)	No	The example does not evidence this information for the AI system itself.
Community Engagement and Impact Reconciliation (Sufficient)	No	There is no public evidence of external engagement.
Alternative Social Perspectives and Innovations Evaluation (Sufficient)	No	The example does not compare the AI system with non-AI, community-led, or institutional energy alternatives.
Action (Sufficient)	Partly	The AI system improves wind-power delivery, but only for Google.

Per the conditions met and the matrix in Figure 1, the project is firmly in the AI in "Application Domain" tier based on publicly available evidence. A case that previously looked relatively strong is revealed to require substantial work to harbour any sustainability claim under the revised framework unless benefits, harms, community authority, and alternatives are assessed. The system is not unsustainable in every respect, but the AI for Sustainability label should be withheld until infrastructural accountability is demonstrated.

3.6. Limitations

Limitations include supply chains and data infrastructures being inherently opaque (Crawford and Joler, 2018), so the framework cannot demand perfect knowledge, instead requiring public evidence and transparency regarding uncertainty. Also, external rating bodies may reproduce the power inequalities they aim to assess, which is why community-nominated assessors and testimony are critical safeguards. Furthermore, community refusal may create difficult tensions with urgent climate action, especially where proposed systems promise emissions reductions. Community refusal does not necessarily mean a project cannot proceed, but it should prevent the project from receiving the strongest AI for Sustainability label while material objections remain unresolved. Finally, the comparison of AI versus non-AI alternatives will be context-specific and contested, thus difficult to objectively analyse. This is a reason to make comparison explicit rather than to avoid it, and to involve community stakeholders in analysis.

Conclusion

To conclude, this paper has closely analysed and built upon Falk and van Wynsberghe (2024) and Lehuedé's (2022) work to define AI for Sustainability. I demonstrated significant gaps in Falk and van Wynsberghe's AI for Sustainability framework as AI's political and social impacts are not addressed. Then, I integrated insights from the analysis of both papers to produce a modified AI for Sustainability framework with a greater emphasis on power, infrastructure, distribution, community authority, and alternative social innovations. This enables the AI for Sustainability label to truly represent a system that is the best solution for a given sustainability issue, and that progresses the sustainability of AI both technically and socially through sincere engagement with communities impacted by the AI system. The operationalisation shows how the framework can move beyond abstract critique, specifying required evidence, assessors, and minimum standards, while the Fairwork-style rating model shows how the framework can avoid being reduced to self-certification. Future work could include in-depth analyses of popular AI for Sustainability projects against this framework, or further developing the conditions. Ultimately, AI systems are unsustainable if their environmental benefits depend on hidden, uncontested destructive infrastructure. The AI for Sustainability label should be reserved for projects with environmental gains matched by accountable infrastructures, community authority, and genuine openness to social alternatives.

References

Ada Lovelace Institute (2021) The Citizens' Biometrics Council. London: Ada Lovelace Institute. Available at: https://www.adalovelaceinstitute.org/report/citizens-biometrics-council/.

Bergman, N., Markusson, et al. 2010, February. Bottom-Up, Social Innovation for Addressing Climate Change. In *Sussex Energy Group Conference—ECEEE*.

Brennan, A. (2024) 'Necessary and Sufficient Conditions', in Zalta, E.N. and Nodelman, U. (eds.) *The Stanford Encyclopedia of Philosophy*. Summer 2024 edn. Available at: https://plato.stanford.edu/archives/sum2024/entries/necessary-sufficient/.

Brodie, P. (2023) 'Data Infrastructure Studies on an Unequal Planet', *Big Data & Society*, 10(1), p. 20539517231182402.

Cairns, R.C. (2014) 'Climate Geoengineering: Issues of Path-Dependence and Socio-Technical Lock-In', *Wiley Interdisciplinary Reviews: Climate Change*, 5(5), pp. 649-661.

Cant, C., Ustek-Spilda, F., Cole, M. and Graham, M. (2023) *Fairwork Annual Report 2023: State of* *the Global Platform Economy.* Oxford: The Fairwork Project. Available at: https://www.fair.work.

Crawford, K. and Joler, V. (2018) Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources. New York: AI Now Institute and Share Lab. Available at: https://anatomyof.ai.

Dauvergne, P. (2022) 'Is Artificial Intelligence Greening Global Supply Chains? Exposing the Political Economy of Environmental Costs', *Review of International Political Economy*, 29(3), pp. 696-718.

Delgado, F., Yang, S., Madaio, M. and Yang, Q. (2023) 'The Participatory Turn in AI Design: Theoretical Foundations and the Current State of Practice', *Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, pp. 1-23.

Delmas, M. and Burbano, V. (2011) 'The Drivers of Greenwashing', *California Management Review*, 54(1), pp. 64-87. Available at: https://doi.org/10.1525/cmr.2011.54.1.64.

Dong, Y. and Hauschild, M.Z. (2017) 'Indicators for Environmental Sustainability', Procedia CIRP, 61, pp. 697-702.

Ensmenger, N. (2021) 'The Cloud is a Factory', in *The MIT Press eBooks*, pp. 29-50. Available at: https://doi.org/10.7551/mitpress/10993.003.0005.

ExxonMobil (2019) ExxonMobil to Increase Permian Profitability through Digital Partnership with Microsoft. Available at: https://corporate.exxonmobil.com/news/news-releases/2019/0222_exxonmobil-to-increase-permian-profitability-through-digital-partnership-with-microsoft.

Falk, S. and Van Wynsberghe, A., 2024. Challenging AI for Sustainability: What Ought It Mean?. *AI and Ethics*, 4(4), pp.1345-1355.

IPBES (2019) Global Assessment Report on Biodiversity and Ecosystem Services. Edited by E.S. Brondizio, J. Settele, S. Díaz and H.T. Ngo. Bonn, Germany: IPBES Secretariat. Available at: https://doi.org/10.5281/zenodo.3831673.

Lehuedé, S. (2022) 'Territories of Data: Ontological Divergences in the Growth of Data Infrastructure', *Tapuya: Latin American Science, Technology and Society*, 5(1), p. 2035936.

Maslej, N., Fattorini, et al. (2024) The AI Index 2024 Annual Report. Stanford, CA: AI Index Steering Committee, Institute for Human-Centered AI, Stanford University.

Mensah, J. (2019) 'Sustainable Development: Meaning, History, Principles, Pillars, and Implications for Human Action: Literature Review', *Cogent Social Sciences*, 5(1), p. 1653531.

Microsoft Research (2024) Microsoft Climate Research Initiative. 17 April. Available at: https://www.microsoft.com/en-us/research/collaboration/microsoft-climate-research-initiative/.

Muldoon, J. (2023) Platform Socialism: How to Reclaim Our Digital Future from Big Tech. London: Pluto Press.

Plociennik, C., Watjanatepin, P., Van Acker, K. and Ruskowski, M., 2025. Life Cycle Assessment of Artificial Intelligence Applications: Research Gaps and Opportunities. *Procedia CIRP*, 135, pp.924-929.

Powell, A.B., Ustek-Spilda, F., Lehuedé, S. and Shklovski, I. (2022) 'Addressing Ethical Gaps in Technology for Good: Foregrounding Care and Capabilities', *Big Data & Society*, 9(2), p. 20539517221113774.

Ritchie, H. and Roser, M. (2024) 'CO2 Emissions', *Our World in Data*, 22 January. Available at: https://ourworldindata.org/co2-emissions.

Stephens, J.C. (2022) Diversifying Power: Why We Need Antiracist, Feminist Leadership on Climate and Energy. Washington, DC: Island Press.

Stilgoe, J., Owen, R. and Macnaghten, P., 2020. Developing a Framework for Responsible Innovation. In *The Ethics of Nanotechnology,* *Geoengineering, and Clean Energy* (pp. 347-359). Routledge.

Szulecki, K. (2018) 'Conceptualizing Energy Democracy', *Environmental Politics*, 27(1), pp. 21-41.

Thompson, N.C. and Ahmed, N. (2023) 'What Should be Done About the Growing Influence of Industry in AI Research?', *Brookings*, 5 December. Available at: https://www.brookings.edu/articles/what-should-be-done-about-the-growing-influence-of-industry-in-ai-research/.

Verdecchia, R., Sallou, J. and Cruz, L. (2023) 'A Systematic Review of Green AI', Wiley Interdisciplinary Reviews: *Data Mining and Knowledge Discovery*, 13(4), p. e1507.

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S.D., Tegmark, M. and Fuso Nerini, F. (2020) 'The Role of Artificial Intelligence in Achieving the Sustainable Development Goals', *Nature Communications*, 11(1), pp. 1-10.

Witherspoon, S. and Fadrhonc, W. (2019) 'Machine Learning Can Boost the Value of Wind Energy', *Google: The Keyword*, 26 February. Available at: https://blog.google/technology/ai/machine-learning-can-boost-value-wind-energy/.