Why Does a Responsible Climate Action AI Need the Arts and Humanities?

By Ramit Debnath

Image credit: Dall-E generated art in different style illustrating the impact of climate change on glaciers

Cambridge University's Ramit Debnath explores the potential of Responsible AI (RAI) in addressing global challenges, and explains why social sciences, philosophy, the arts and humanities have a critical role to play in shaping AI system design and presenting us with the best chance of securing a sustainable planetary future.

The AI revolution is driven by a global megatrend changing societies and economies via use of digital technologies, a process known as digitalisation. AI is already transforming job markets, business models, governance and societal welfare structures. This disruption has the potential to impact current actions and progress in addressing global challenges such as climate change, social inequality, clean water, health, human rights violations, migration, conflict and war.[1]

> By sharing stories, we can make complicated AI ideas more understandable to the public and start conversations about AI's role in climate justice.

Researchers and practitioners around the world are addressing these emerging AI challenges through the lens of responsible and safe AI. Where Responsible AI (RAI) refers to developing, deploying and using AI that is fair, accountable, transparent and ethical (FATE), and also to sustainability.

While the precise role of RAI in addressing current global challenges, such as climate change, remains unclear,[2] using AI to account for the ever-changing elements of climate change allows us to make more informed predictions

and ensuring their alignment with real-world conditions, but creating a true HITL system can be challenging. This is because, throughout the supply chain of these AI technologies, we continue to introduce biases through three key modes: biased datasets, biased programming and biased AI algorithm design, all of which fundamentally undermine the trustworthiness of these systems.[2] Applying these biases to complex decision-making tasks such as climate action, policing, judgement and healthcare, among others, can have devastating effects.

> **Human-in-the-loop (HITL) approaches are one of the main ways to make AI systems more reliable, fair, and simple to understand.**

about environmental changes, enabling us to implement mitigation and adaptation efforts sooner.[2] However, at present, addressing climate change using AI is difficult due to the vast number of variables linked with Earth's climate system data.

Humans in the loop

Human-in-the-loop (HITL) approaches are one of the main ways to make AI systems more reliable, fair, and simple to understand. HITL designs do this by incorporating human judgement into AI systems and allowing for a collaborative process. This is particularly important for climate-related AI applications, where the stakes are high and decisions can have far-reaching consequences for ecosystems, societies and the planet.

Experts argue that human intervention is essential for validating AI-driven climate models

The challenges of biased algorithmic design

The challenges of biased algorithmic design and programming usually have technical fixes, and the tech industry can deploy bias-correcting measures at scale. For instance, Deepmind researchers recently published a paper in Nature that illustrates the use of watermarks to enhance the transparency of texts produced by Large Language Models (LLMs).[3]

These watermarks assist in distinguishing between human-generated text and AI-generated text. The larger challenge, however, lies in the unbiasing of training datasets, whose bias is a result of deep-rooted digital inequalities and social injustices.

These inequalities can manifest in various forms, such as the inability to access an internet connection, digital illiteracy or a lack of access to affordable and accessible computing infrastructure, or exploitative labour practices of employing cheap data workers from the Global South.[4]

here, offering a broader ethical lens and deep understanding of human values, culture and societal contexts that can be invaluable in shaping RAI.

By analysing the social impacts of AI systems and uncovering potential biases within the AI supply chain, social scientists make significant

> **A distributive and procedural justice lens can help guide AI practitioners in developing responsible systems that align moral imperatives with practical viability.[2]**

A role for social sciences, philosophy, the arts and humanities

Climate change amplifies inequalities and injustices, disproportionately impacting vulnerable, resource-constrained groups. This makes it challenging to develop climate-focused AI that is unbiased, fair, and trustworthy in its decision-making. Social sciences, philosophy, the arts and the humanities play a critical role

contributions. They assist in comprehending the impact of AI on various demographic groups, its influence on the political economy, and its exploration of systemic issues such as environmental injustices or economic disparities, which technical models might otherwise overlook.[2] Social science researchers tease out the cultural dynamics and power structures that shape the spatial and temporal scope of AI implementation, ensuring the inclusion of diverse perspectives.

AI developers can gain a nuanced understanding of the communities affected by climate change by leveraging methodologies like ethnographic studies or stakeholder consultations – a perspective often overlooked in the current practice of training AI on the entire internet.

Inclusive by design

Philosophers add another crucial dimension by questioning the ethical foundations of AI in climate action. They probe fundamental

Image credit: https://www.campaignasia.com/article/ai-paints-two-futures-climate-action-or-not-for-wwf-exhibition/484138

questions like, 'What is fairness in an AI system?' They also explore fundamental questions, such as 'What constitutes fairness in an AI system?' and 'Which values should an AI system prioritise?'

Philosophy encourages AI developers to reflect on the implications of the technology, not only on a societal level but also in terms of individual rights and freedoms. For instance, when designing an AI for climate-related resource distribution, it's crucial to question its fairness and align its decision-making process with human-centred values. A distributive and procedural justice lens can help guide AI practitioners in developing responsible systems that align moral imperatives with practical viability.[2] AI systems become inclusive by design.

for a wider range of people to understand. By sharing stories, we can make complicated AI ideas more understandable to the public and start conversations about AI's role in climate justice.

Similarly, digital art serves as an effective medium for storytelling and educating about the shortcomings inherent in AI. Artists are leveraging these technologies to address issues such as bias resulting from misrepresentations of culture and social dynamics. For example, in 2024, the UN Headquarters used AI art to compile millions of photos of coral reefs, many of which are under threat from rising ocean temperatures brought on by climate change.[5] Similarly, in 2024, the World Wide Fund for Nature curated an exhibition that showcases a series of 20 AI-generated paintings. These paintings depict two futures: one in which

> **Including insights from the arts and humanities fosters a more holistic, people-centred approach to AI development while aligning with planetary health goals.**

The power of stories

The humanities (which include history, literature and culture studies) have a lot to say about the beneficial and adverse effects that technology has had on society over time. By looking into the early days of the AI development model, these historical stories can reveal patterns of inequality and warn us not to make the same mistakes again. Literature and stories, on the other hand, help connect scientific experts with regular people, making the effects of AI easier

society addresses climate change, and another in which it does not, illustrating the perilous trajectory we are on and the pressing need for urgent action to restore nature.[6] Including insights from the arts and humanities fosters a more holistic, people-centred approach to AI development while aligning with planetary health goals. It encourages empathy, cultural sensitivity and awareness of the real-world consequences of technology.

Author bio

Ramit Debnath is a university assistant professor and an academic director at the University of Cambridge. He leads the Cambridge Collective Intelligence & Design Group and co-directs the climaTRACES lab at CRASSH. Ramit works at the intersection of computational social sciences, responsible AI design and climate action, especially interested in how individual behaviour rolls into the social dynamics of collective decision-making and how can the emergent AI help? He has visiting faculty roles at Caltech and the Indian Statistical Institute, and serves on the steering committee for the Cambridge's Centre for Human-Inspired AI and the Centre for Data-driven Discovery. Ramit has a background in electrical engineering and computational social sciences, and an MPhil and PhD from Cambridge as a Gates Scholar.

Cite as: Ramit Debnath (2024). 'Why Does a Responsible Climate Action AI need the Arts and Humanities?' *The New Real Magazine*, Edition Two. pp 76-80. https://doi.org/10.2218/newreal.10934

References:

1. Creutzig, F., et al. (2022) 'Digitalization and the Anthropocene,' Annu. Rev. Environ. Resour. 47, 479-509.

2. Debnath, R., et al. (2023) 'Harnessing Human and Machine Intelligence for Planetary-level Climate Action,' npj Clim. Action 2, 20.

3. Dathathri, S., et al. (2024) 'Scalable Watermarking for Identifying Large Language Model Outputs,' Nature 634, 818-823.

4. Perrigo, B. (2023) 'Exclusive: OpenAI Used Kenyan Workers on Less Than $2 Per Hour to Make ChatGPT Less Toxic,' TIME.

5. Lennon, C. (2024) 'AI-powered Art puts 'Digital Environmentalism' on Display at UN Headquarters,' UN News.

6. Watson, I. (2024) 'AI Paints Two Futures - Climate Action or Not - for WWF Exhibition. Campaign.'

Links:

1. Cambridge Collective Intelligence & Design
2. climaTRACES lab