2025

Reimagining AI

For Environmental Justice and Creativity

COLLECTION OF ESSAYS

Edited by:
Jess Reia I MC Forelle I Yingchong Wang

2025

REIMAGINING AI

For Environmental Justice and Creativity

Editors:
Jess Reia, MC Forelle and Yingchong Wang

Contributors (alphabetical order by last name):
Ahmed Alrawi, Rafael Alvarado, Blair Attard-Frost, Lauren E. Bridges, Celia Calhoun, Coleen Carrigan, Jonathan Colmer, Yasmin Curzi, Ella Duus, Jonah Fogel, MC Forelle, Pedro Augusto P. Francisco, Sergio Guillen Grillo, Desiree Ho, Mehan Jayasuriya, Steven L. Johnson, Owen Kitzmann, Tamara Kneese, Jonathan Kropko, Kyrill Kunakhovich, Rachel Leach, Danila Longo, Siobhán Loughney, Maria Lungu, Christine Mahoney, Raheem Manning, Aaron Martin, Martina Massari, Shalini Misra, Andrew Mondschein, Peter Norton, Anne Pasek, Jess Reia, Andrea Roberts, Bryn Seabrook, Jessica Sewell, Anuti Shah, Mona Sloane, Andre Sobral, Will Straw, Maria Villanueva, Amanda Wyatt Visconti, Yingchong Wang, Keren Weitzberg, Megan Wiessner, Damien P. Williams, Caitlin Wylie.

Research team:
Anuti Shah, Jess Reia, MC Forelle, Rachel Leach, Sam Kane, Siobhán Loughney and Yingchong Wang

Revisions:
Anuti Shah, Rachel Leach and Siobhán Loughney

Design and layout:
Nela Díaz

Suggested citation:
Reia, J., Forelle, MC and Wang, Y. (2025). Reimagining AI for Environmental Justice and Creativity. Digital Technology for Democracy Lab, University of Virginia.

University of Virginia
Environmental Institute

University of Virginia
Digital Technology for Democracy Lab

University of Virginia
School of Data Science

CONTRIBUTORS

ALRAWI, Ahmed

ALVARADO, Rafael

ATTARD-FROST, Blair

BRIDGES, Lauren E.

CALHOUN, Celia

CARRIGAN, Coleen

COLMER, Jonathan

CURZI, Yasmin

DUUS, Ella

FOGEL, Jonah

FORELLE, MC

FRANCISCO, Pedro Augusto P.

GUILLEN GRILLO, Sergio

HO, Desiree

JAYASURIYA, Mehan

JOHNSON, Steven L.

KITZMANN, Owen

KNEESE, Tamara

KROPKO, Jonathan

KUNAKHOVICH, Kyrill

LEACH, Rachel

LONGO, Danila

LOUGHNEY, Siobhán

LUNGU, Maria

MAHONEY, Christine

MANNING, Raheem

MARTIN, Aaron

MASSARI, Martina

MISRA, Shalini

MONDSCHEIN, Andrew

NORTON, Peter

PASEK, Anne

REIA, Jess

ROBERTS, Andrea

SEABROOK, Bryn

SEWELL, Jessica

SHAH, Anuti

SLOANE, Mona

SOBRAL, Andre

STRAW, Will

VILLANUEVA, Maria

VISCONTI, Amanda Wyatt

WANG, Yingchong

WEITZBERG, Keren

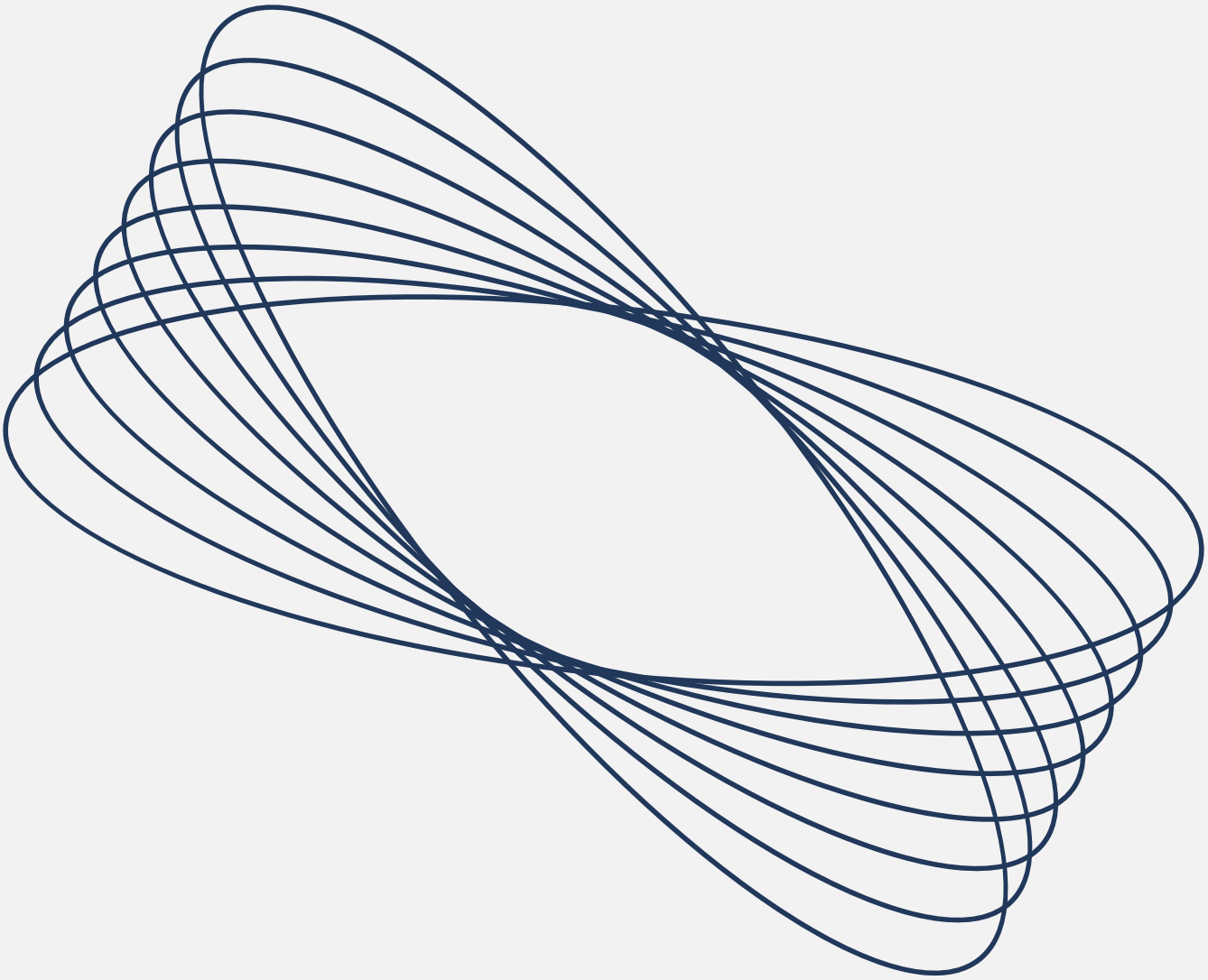
WIESSNER, Megan

WILLIAMS, Damien P.

WYLIE, Caitlin

TABLE OF CONTENTS

INTRODUCTION



CRITICAL CONVERGENCES: *Technology and Culture in a Burning Planet*

Jess Reia
MC Forelle
Yingchong Wang

Artificial intelligence (AI) is frequently presented as ubiquitous and inevitable, and indeed, today it has penetrated nearly every sector of global society, from health to education to finance, becoming the focus of many a national news story, international declaration, and intra-national political agendas. Despite its rising popularity, AI is not always visible. People everywhere constantly interact with AI-based systems making decisions for them in apps and services without being notified of the automated decision-making process.

The often-vague narrative about AI's potentialities and limitations contributes to the opacity of such systems whose social, environmental and cultural costs are still being measured and studied. Simultaneously, well-documented analyses of harms caused by overly techno-optimistic adoptions of AI guide us toward voices not always heard in the race for global AI leadership. At the heart of many of these conversations – both the ostentatious promises of techno-utopians and grounded and conscientious work of critical AI researchers – are questions about the current state, and possible future of, creativity and environmental sustainability in a world with, of, and through, AI. If AI, especially large-language models (LLMs), are contributing to a burning planet by using massive amounts of resources while also hindering creative industries, what comes next? If we could reimagine AI and its ecosystem, what would it look like?

To develop novel and multifaceted answers for these questions, we decided to explore it collectively and across disciplines and sectors. This collection of essays emerged from a workshop also titled "Reimagining AI

for Environmental Justice and Creativity," and is designed to become another useful resource for a broad audience of people with various backgrounds, skills, and interests. From educational and research contexts to policymaking and activism, we hope the ideas featured here will help us to reflect upon the challenges ahead when building, using and evaluating AI in different contexts.

The workshop took place at the University of Virginia (UVA)'s main campus in Charlottesville, United States, in October 2024. It brought together international speakers, the UVA community across schools and departments, and local government. It was designed as an opportunity to connect people, organizations, and resources in the networks we built over the last years of research and advocacy. In two days, participants engaged with panels open to the public and hands-on, invitation-only roundtables that explored overlapping issues that are not always visible in the public debate and scholarship, such as climate change, creative expression and the several roles that AI increasing embody in society.

AI and the environment

To understand the growing environmental impact of AI, we need to consider the infrastructure and materiality of big data and AI. AI systems, particularly large-scale models, require substantial computational power, leading to high energy consumption and increased greenhouse gas emissions. Additionally, the data centers housing AI servers consume vast amounts of water for cooling and generate electronic waste. The minerals needed for AI hardware further contribute to environmental degradation and geopolitical conflicts in the Global Majority, where so many natural resources are located – such as the lithium triangle (Chile, Argentina and Bolivia) and cobalt in the Democratic Republic of Congo. These impacts are globally connected and matter to everyone because they push us further into the climate crisis, resource depletion, and pollution that affect ecosystems and human and non-human health. Here we address some of these issues and proposed alternative ways to understand act to build environmental justice

AI and creative expression

One of the most profound intersections of technology and culture today is AI's relationship with creative expression. As generative AI products rapidly evolve to create increasingly refined text, images, and music, they pose fundamental questions about the essence of creativity. While these technologies democratize the creative tools, they also challenge traditional notions of authorship and original works. The utilization of AI has created new opportunities for creators; however, this creative frontier raises ethical issues with respect to copyright, compensation, and even cultural appropriation. Much current AI-generated work is being trained on datasets containing human-created work without explicit permission or attribution, raising legal and moral concerns about ownership. As these technologies become more common, there is an urgent need for new policy frameworks that both protect human creators while enabling innovation. The workshop and the collected essays contributed to this vital conversation, but the broader obligation falls on all stakeholders to engage in this critical conversation, to help

cultivate an ecosystem in which AI enhances human creativity rather than exploiting its value and diverse creativity is preserved rather than homogenized expressions.

The essays

We present these essays as a window into the many unresolved, at times conflicting, roles that artificial intelligence currently plays in the everyday work and imaginaries of experts and practitioners across disciplines. These pages are also a call for more critical, creative and nuanced explorations of technologies and how they shape the world around us. The collection is structured in four parts: Keeping AI within planetary boundaries, Regulating and governing AI for the public interest, Reframing understandings of AI, and Creating for an alternative AI future.

Keeping AI within planetary boundaries

This report begins with a collection of essays exploring the challenges and opportunities AI presents to environmental sustainability from the micro to the macro scale. Colmer leads off by pointing out that AI, like all emerging technologies, presents considerable potential to address environmental challenges; but, like all emerging technologies, requires a robust regulatory approach to balance the imperative for continued growth with the need to mitigate environmental costs. Pasek presents a more cautionary approach, reminding us that the costs of AI are already possible to measure, and those costs must be taken more seriously by regulators than the nebulous promises of revolutionary benefits that AI offers. Kneese and Bridges concur with these concerns, arguing that we must look beyond energy use and consider AI as both localized and supply-chain phenomena; Bridges further suggests the concept of "parasitic computation" as a more accurate framing of AI. Presenting another holistic lens, Wiessner puts forward a political ecology approach to AI that truly considers its impact not just on environmental sustainability, but further toward environmental justice. Loughney and Leach write from the student perspective, presenting concerns about the normalization of AI tools, even as it becomes clear that more energy and more data will not solve the problems of current AI systems. The final portion of this section turns toward research and governance, with Fogel and Jayasuriya identifying how grantmakers are supporting initiatives that harness AI to advance important climate research and community-driven interventions. Finally, our authors end by considering the imperative of Non-Western participation in global AI governance: Lungu considers the inverse disproportionality of Africa's contributions to AI development, in labor and resources, versus its presence in governance discussions; while Reia exhorts the importance of including Non-Western visions of the future that view degrowth and deceleration as not only possible, but necessary.

Regulating and governing AI for the public interest

We follow with a collection of essays that consider how AI systems might be governed, collaboratively developed, or resisted. It begins with essays that propose different governance strategies that provide alternative visions to current top-down approaches. Echoing many of our earlier writers, Attard-Frost, Shah and Alrawi. In each of their essays, illustrate how government has been slow to act in the face of AI, but how local, regional,

and decentralized initiatives show promise in countering this intransigence and putting power back in the hands of communities. Looking at the needs of specific communities, Misra and members of the Sloane Lab detail how emergency managers and university students, respectively, must be involved in AI system design and implementation for those systems to meet their needs. However, throughout these governance efforts, we must remain aware of the complications that AI systems are likely to introduce: Martin and Weitzberg consider how the increasing ubiquity of AI is inducing many governments to turn to the use of (privately owned) biometrics to authenticate human identity, and Seabrook discusses similar concerns about the long-term impacts of AI on higher education. Finally, our authors contemplate the possibility of refusal as a governance tactic, with Curzi outlining different approaches to the ethics of AI refusal, not to reject, but to encourage critical engagement with AI, and Mahoney suggesting that some forms of refusal may become new trends in production and corporate self-governance.

Reframing understandings of AI

The third section examines existing constructions of AI from multiple, disciplinary perspectives. Kunakhovich reminds us that our current fears about AI are a replay of how people once reacted to gramophones and film, and Carrigan reviews how the tech industry's "Bro Code" connects to a way our personal data is collected without consent. Wylie invites readers to question the common narratives around AI's labor-saving and problem-solving capacities, and Straw asks readers to think how AI reshapes cities and urban experiences. Francisco paints a vivid picture of the internet as a "Dark Forest" where users have to "hide" from AI-generated content, while Alvarado and Sobral push readers to rethink whether these systems deserve to be called "intelligent." The section ends with powerful reflections on what remains uniquely human. Johnson contends that the social trust that enabled innovation cannot be replaced by any amount of AI processing. Norton uses paint-by-numbers kits as a metaphor to discuss AI writing tools in educational settings, suggesting that the real value lies in the personal satisfaction of creative expressions. However, cognitive capacities, as Mondschein warns, may actually pay a price for surrendering our navigational reasoning to AI. Throughout these diverse perspectives, Forelle takes on a range of views and makes a strong argument challenging the idea that the development of AI is inevitable and that we as humans can shape technology to fit our values and needs.

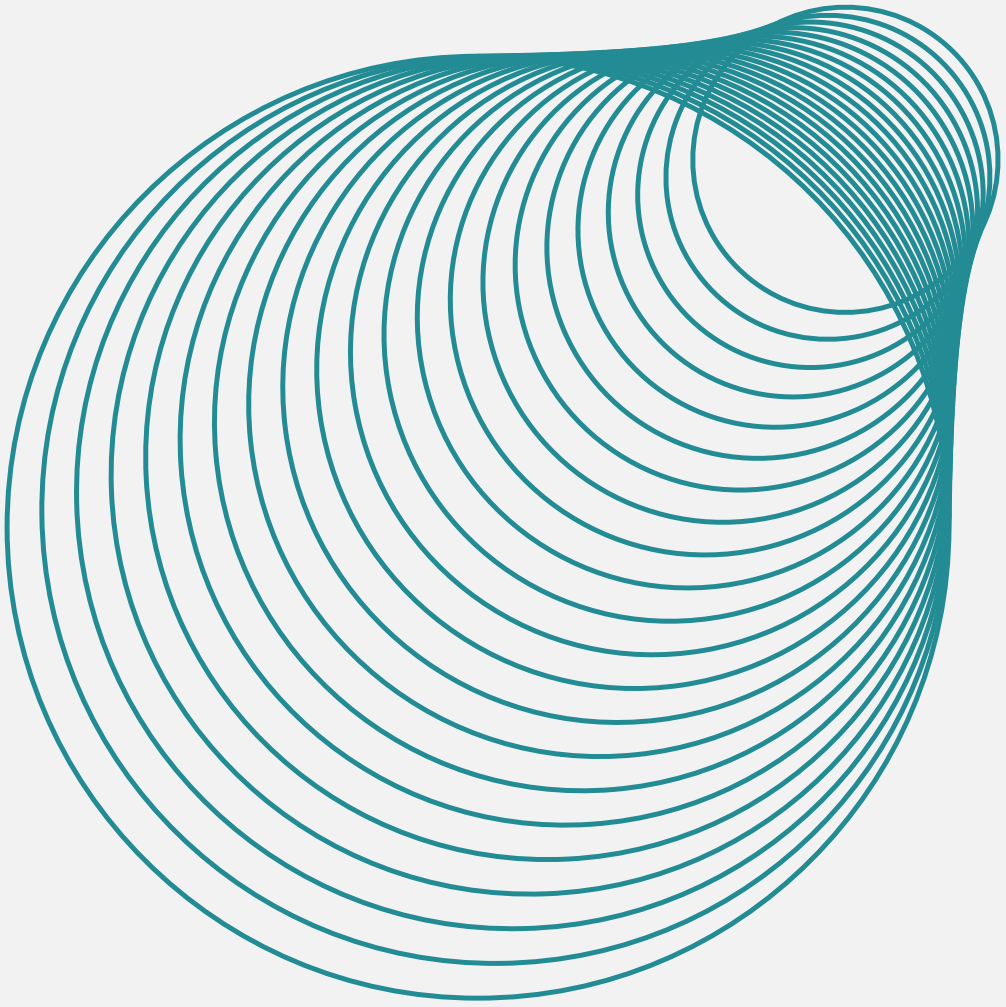
Creating for an alternative AI future

This section provides pathways toward more equitable, culturally sensitive, and environmentally responsible AI systems. Wang shows that established cultural policy models—from the market-oriented Facilitator to a more public-planned Architect approach—offer valuable frameworks to guide AI governance in creative domains. Visconti states that framing AI as part of the continuum of code rather than something completely new enables us to borrow decades of ethical design practices from digital humanities. Kropko and Manning both point out the democratic potential of AI—Kropko advocate free and open-source AI tools for community

and public interest, and Manning articulate how AI can promote creative equity if tools are employed with sufficient safeguards by sharing his experiences in planning Philadelphia's nightlife scenes. In this vision, Massari and Longo contend that AI should be seen as cultural infrastructure that will require settings favoring self-governance and democracy, proposing their project Civic Digital Twin as a model for integrating civic engagement in AI-driven urban planning. Williams draws our attention to the environmental costs of AI: how data center increase water consumption and impact resource-scarce communities. Both Grillo and Williams call for clarity in understanding what AI discourse does; Guillen Grillo suggests a "discursive cartography" to map out the different perspectives on AI and a "deliberative cartography" to help clarify the ways AI and forms of democratic governance work with each other, while Williams challenges us to question whether we should be pursuing AI at all and what it would take to build AI in truly sustainable ways.

As this collection illustrates, reimagining AI involves moving beyond the narrative of technology inevitability toward a more comprehensive understanding that spans environmental, creative, social, and ethical dimensions. The essays offer no single solution but rather a mosaic of perspectives that illuminate different pathways forward – challenging us to ask not just how we might regulate AI, but how we might fundamentally remodel its development, deployment, and governance. Across these four thematic sections, our contributors map the current landscape while signaling to more sustainable and imaginative horizons, reminding us that AI's future is still unwritten: it is a canvas on which we all have right and responsibility to draw. In bringing these voices into the conversation, we hope to inspire readers across sectors to engage critically and creatively with AI's evolving role in our shared world, promoting that technology serves humanity and the planet rather than the other way around.

Keeping AI within Planetary Boundaries



01.

DOES AI HAVE AN *Environment Problem?*

Jonathan Colmer

Department of Economics
University of Virginia

Advances in artificial intelligence (AI) have the potential to transform economic activity and redefine many aspects of our lives, including the way we work, learn, and play. However, as its influence grows, so do concerns about its environmental consequences and how they are distributed. AI is undeniably resource intensive, consuming large amounts of water and electricity. Yet, this is not unique to AI as nearly all forms of production and consumption impose environmental costs that are not fully accounted for in decision-making. The critical question is not whether AI uses resources, but how we can best align incentives to minimize these costs and whether the social benefits of AI outweigh the social costs (environmental and beyond).

Much of the public concern about the environmental costs associated with AI stems from eye-catching numbers about its total resource use. Newspaper articles highlighting the large amounts of water and electricity used to develop AI models, have portrayed AI as a "planet-eating" technology.[1] A more thoughtful evaluation, however, is required. While media narratives may raise awareness about environmental costs, they misdirect attention from more costly activities, obscure efficiency improvements, and largely ignore the potential benefits of AI activity.

It is important to understand the broader context. Investment in new data centers (which include all data processing and storage activities, not just AI) has surged in the last two years. Global data center electricity consumption, however, has remained relatively stable over time, accounting for ~1-1.5% of global electricity consumption.[2] The stability of aggregate electricity consumption is quite remarkable given the doubling of internet users since 2010 and a 25-fold expansion of internet traffic. This stability is due to rapid improvements in energy efficiency. The efficiency of AI-related computer chips has doubled every two to three years and modern AI-related chips use 99% less power to perform the same calculations as a model from 2008.[3] In addition, new cooling technologies have led to the development of data centers that consume zero water for cooling. Such advances have led Microsoft to reduce water intensity by over 80% since the early 2000s.[4]

These achievements highlight the importance of incentivizing further efficiency

improvements. If the social costs of energy and water use were reflected in market prices, incentives would be aligned, incentivizing further innovations and investments in efficiency increasing technologies.[5] Where environmental costs are incurred, it is also important to ensure that these burdens are not disproportionately borne by specific populations or regions, emphasizing the need for equitable and inclusive policy solutions.

AI also holds the potential to directly address environmental challenges by contributing to the development of technologies that address and mitigate environmental challenges like climate change and reduce the energy and resource intensity of economic activity more broadly.[6] These aren't just potential benefits — AI is already reducing the environmental costs of economic activity by optimizing the grid, increasing renewable energy deployment, improving supply chain efficiency, improving the monitoring and enforcement of conservation efforts, and advancing climate science. To fully realize these benefits, it is essential to ensure equitable access to AI-enabled solutions, allowing all communities and regions to share in the opportunities created by these advances.

Instead of singling out AI, we need evidence-based policies that align incentives to internalize environmental costs and ensure an equitable distribution of costs and benefits across all sectors and activities. This does not mean that the development of AI comes without problems. Serious concerns about algorithmic bias and discrimination, the concentration of market power, labor market disruptions, privacy violations, the misalignment of AI with human values or interests (which, ironically, may be good for the environment), accountability and transparency, and malicious use all demand careful regulatory and policy attention. To manage these risks, we need a systematic, comprehensive, and balanced approach. Addressing environmental challenges requires a similar framework — one that balances the imperative for continued growth in prosperity and well-being, particularly for the world's poorest populations, with the need to mitigate the environmental costs that such growth creates.

ENDNOTES

1 The Guardian. (2023, June 8). As the AI industry booms, what toll will it take on the environment? Retrieved from v The Guardian. (2024, May 30). The ugly truth behind ChatGPT: AI is guzzling resources at planet-eating rates. Retrieved from https://www.theguardian.com/commentisfree/article/2024/may/30/ugly-truth-ai-chatgpt-guzzling-resources-environment / Scientific American. (2023, December 7). AI's climate impact goes beyond its emissions. Retrieved from https://www.scientificamerican.com/article/ais-climate-impact-goes-beyond-its-emissions/ The Washington Post. (2024, September 18). A bottle of water per email: the hidden environmental costs of using AI chatbots. Retrieved from https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers/

2 International Energy Agency (2023). Data centres and data transmission networks. Retrieved from https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

3 International Energy Agency (2024). What the data centre and AI boom could mean for the energy sector. Retrieved from https://www.iea.org/commentaries/what-the-data-centre-and-ai-boom-could-mean-for-the-energy-sector.

4 Microsoft. (2024, July 25). Sustainable by design: Transforming datacenter water efficiency. Microsoft Cloud Blog. Retrieved from https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/07/25/sustainable-by-design-transforming-datacenter-water-efficiency/

5 Jaffe, A., Newell, G., and Stavins, R. (2003) "Technological Change and the Environment", Handbook of Environmental Economics, Volume 1, Edited by K. Mäller and J. Vincent.

6 Amodei, D. (2024) Machines of loving grace. Retrieved from https://darioamodei.com/machines-of-loving-grace

AI IS JUST A BUNCH *of Data Centers*

Anne Pasek

Department of Cultural Studies, Trent School of the Environment
Trent University

I am not an expert in AI. Unexpectedly, in conversations about how to react to this seemingly new technology, this is often a real asset. I am generally uninterested in the benchmark performance of new models and I am decidedly skeptical about utopian tales of the advanced AI to come. I am, however, an expert in ICT sustainability strategies and energy politics, long before the sector pivoted to AI. From that perspective, I see more continuities than disruptive breaks. The tech sector is still fundamentally in the business of data brokering, pushing the line of privacy and copyright law, and pursuing monopoly profits through platform lock-in. The current turn towards AI represents an intensification of these existing trends rather than a brand-new story.

This is especially the case when it comes to AI's environmental impacts. The ICT sector was already instigating a range of local land and water use conflicts with its often-clumsy data center expansion strategies, straining energy grids to shoulder new and substantial loads, and jumping between reporting methodologies to make all its accompanying carbon emissions seem to disappear. AI has made these trends worse. To look at only one company, Microsoft's emissions are 30% higher today than they were in 2020 because of its AI development priorities, and all the accompanying data center growth they entail. As a result, its goal to be carbon negative by 2030 is five times further away than it was a few short years ago.[1] Even the baroque system of compensatory carbon offsets and renewable energy credits on which the company had previously relied (with sometimes questionable credibility) to wash its hands clean cannot keep up with the current data center boom.

The AI-amplification of ICT's climate trajectory is simply unsustainable. Data centers currently make up 2.5% of energy demand in the USA; by some estimates, this could rise to up to 9% by the end of the decade.[2] The grid cannot accommodate that increase without real consequences to the reliability, price, and cleanliness of its energy supply. Ireland, which hosts a disproportionate share of Europe's data centers, is a warning sign: data centers are on track to take up a third of all demand on their national grid, threatening brown outs, rate hikes, and real barriers to progress on the country's national climate commitments.[3]

Communities in the US that are already host to dense data center clusters, such as Northern Virginia, are presently experiencing many of these problems (along with extensive noise and air pollution at the fence line).[4] The hyper concentration and continued expansion of these infrastructures is what AI hype cashes out to on the ground.

The strategies the sector has mustered in response to this crisis represent further continuities and cause for concern. Tech companies have generally relied on efficiency as a cure-all strategy for their environmental impacts, trusting that computational work will gradually get cleaner as chip designs improve over time.[5] The problem with this strategy is two-fold. Firstly, Moore's Law, which underwrote much of these efficiency gains, is coming to an end. It's not clear that there are many years of better chips ahead to bank on. Secondly, efficiency makes for cheaper goods, which in turn drives demand. In a context where CEOs describe the need for AI energy use as essentially infinite, it's naïve to assume that more environmentally friendly AI architectures or GPUs won't simply mean an overall increase in the kinds of AI products under development, with commensurately growing climate impacts. This approach misses the whole for the parts.

Policy makers should be very cautious about the limits of energy efficiency as an environmental strategy, just as they should maintain a disciplined skepticism about the novelty or exceptionalism of these technologies. Ultimately, it's all just data centers (and we perhaps have too many of them as it is). Grounding assessments in the churn of chips, concrete, and transmission lines that accrue around these structures will make for sounder, more accountable assessments of AI than any speculative account of the promises and perils of a wholly new technical frontier. That future is already here, and it's made of sprawling server racks in air-conditioned boxes.

ENDNOTES

1 Rathi, A & Bass, D. (2024). Microsoft's AI Push Imperils Climate Goal as Carbon Emissions Jump 30%. Bloomberg. Com. https://www.bloomberg.com/news/articles/2024-05-15/microsoft-s-ai-investment-imperils-climate-goal-as-emissions-jump-30

2 Aljbour, J. Wilson, T. & Patel, P. (2024) Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption. Palo Alto: Electric Power Research Institute. https://www.epri.com/research/products/000000003002028905.

3 International Energy Agency. (2024) Electricity 2024: Analysis and Forecast to 2026. IEA. https://www.iea.org/reports/electricity-2024. / Bresnihan, P., & Brodie, P. (2021). New extractive frontiers in Ireland and the moebius strip of wind/data. Environment and Planning E: Nature and Space, 4(4), 1645–1664. https://doi.org/10.1177/2514848620970121

4 Aljbour, J. Wilson, T. & Patel, P. (2024)

5 Pasek, A. (2023). From Atoms to Electrons: An Energy History and Future of Computing. In Digital Energetics, 15–42. Minneapolis: University Of Minnesota Press.

AI IMPACTS *Out of Frame*

Tamara Kneese

Climate, Technology, and Justice Program
Data & Society Research Institute

There is a massive amount of compute[1] required for AI, which requires GPUs to train and run high-energy workloads. In addition to chip design and fabrication, data centers and related energy infrastructures contribute to AI's environmental impacts. We have seen the failure to retire coal plants[2] and the revival of nuclear plants,[3] specifically, because of increasing energy demands from computing that outpaces renewable energy supplies, while data centers draw water[4] from drought-stricken areas. This is why Hugging Face researchers have called for the equivalent to an Energy Star rating system for assessing AI models,[5] because there are no clear, standardized metrics attached to models. But measuring energy and water consumption alone does not capture the full spectrum of impacts, which is why there is a need for more empirical, on-the-ground data from different regions around the world.

High-level frameworks for AI risk management and safety tend to ignore human rights impacts that are a matter of environmental justice. Frameworks substitute for engaging the diverse communities who are at the frontlines of both climate change and AI's harmful effects. Policy recommendations also tend to focus on technical evaluations and tweaks, which means that downstream repercussions and environmental considerations are left out of the equation. Looking at a model's potential for bias doesn't tell you much about how the model is connected to a global supply chain of labor exploitation and environmental degradation. Other impacts, including noise pollution and air pollution that leads to asthma, the loss of drinking water during droughts, the loss of agricultural land and displacement from homes, increased utility rates, and other social factors, are not captured by quantitative, technical measurements associated with AI infrastructures.

Investment in AI is often justified by AI's potential to solve the climate crisis. But as we have seen with the growth of LLMs, companies like Microsoft[6] are blowing way past their emissions targets because of their new data centers, which themselves are carbon-intensive to build. This is why we can't just look at training and inference for calculating AI's energy footprint, but we must also look at mining, manufacturing, training, use, disposal –

the full life cycle. AI is being used to accelerate oil and gas extraction,[7] or other economic benefits, with very few environmental guardrails. Michigan just passed tax breaks[8] for a hyperscale data center that undermines the state's own climate goals, while customers might pay higher water and electricity bills.

Researchers also need to address the e-waste implications of rapid technological development. The push for generative AI innovation means that computer hardware is outdated more quickly, so devices must be refreshed more frequently. Hardware used in data centers and server farms contains valuable minerals that can be gleaned and reused, but it also contains hazardous materials including lead and mercury, meaning that the toxins that give electronics manufacturing workers cancer later harm the communities that are left to deal with e-waste. Refurbishment and circular design for hardware should be part of policy solutions that aim to mitigate the environmental impacts of AI.

For a short time, net zero goals were common for major companies and there have been pushes from employees[9] and even from shareholders[10] to adhere to climate pledges. But in the past two years, many companies have failed to meet their earlier promises.

The workers within tech companies who are focused on accountability, including corporate responsibility, RAI, and sustainability teams, have been cut or undermined during a period of mass layoffs while companies focus attention on generative AI models that depend on the exploitation of precarious data workers and the stolen labor of creatives while contributing to climate change.

Terms like AI, data center, and compute obscure disparities and forms of labor. Measuring and mitigating the climate impacts of AI across the supply chain and across its lifecycle requires attention to what such terms obscure. And it's all too common for tech companies to use the ambiguity of AI to justify their growing power, claiming that AI will help solve social problems like climate change, relying on speculative fantasy instead of tangible existing harms that are outside of frame. Policymakers should examine the complex relationships between tech and energy companies, local and state governments, and public utilities, capturing not only the technical measurement of greenhouse gas emissions and water consumption associated with AI infrastructures but also human rights impacts to public health and ecosystems.

ENDNOTES

1 Vipra, J. & Myers West, S. (2023, September 27). Computational Power and AI. AI Now Institute. https://ainowinstitute.org/publication/policy/compute-and-ai

2 Kishan, S. & Saul, J. (2024, January 25). AI Needs So Much Power That Old Coal Plants Are Sticking Around. Bloomberg.Com. https://www.bloomberg.com/news/articles/2024-01-25/ai-needs-so-much-power-that-old-coal-plants-are-sticking-around

3 Greshko, M. (2024). Nuclear power for AI: What it will take to reopen Three Mile Island safely. Nature, 634(8033), 272–273. https://doi.org/10.1038/d41586-024-03162-2

4 Hao, K. (2024, March 1). AI Is Taking Water From the Desert. The Atlantic. https://www.theatlantic.com/technology/archive/2024/03/ai-water-climate-microsoft/677602

5 Luccioni, S. (2024, May 9). Energy Scores for AI Models. Hugging Face. https://huggingface.co/blog/sasha/energy-star-ai-proposal

6 Calma, J. (2024, May 15). Microsoft's AI obsession is jeopardizing its climate ambitions. The Verge. https://www.theverge.com/2024/5/15/24157496/microsoft-ai-carbon-footprint-greenhouse-gas-emissions-grow-climate-pledge

7 Sharma, G. (2023, April 14). How Multibillion Dollar Investments In AI Are Driving Oil And Gas Sector Innovation. Forbes. https://www.forbes.com/sites/gauravsharma/2023/08/14/how-multibillion-dollar-investments-in-ai-are-driving-oil-and-gas-sector-innovation/

8 Ellison, G. (2024, November 15). Michigan Approves Tax Breaks for Hyperscale Data Centers. GovTech. https://www.govtech.com/policy/michigan-approves-tax-breaks-for-hyperscale-data-centers

9 Selyukh, A. (2019, September 19). Amazon Makes "Climate Pledge" As Workers Plan Walkout. NPR. https://www.npr.org/2019/09/19/762336929/amazon-makes-climate-pledge-as-workers-plan-walkout

10 Copley, M. (2023, April 9). Businesses face more and more pressure from investors to act on climate change. NPR. https://www.npr.org/2023/04/09/1168446621/businesses-face-more-and-more-pressure-from-investors-to-act-on-climate-change

FROM AI TO PC: *Reframing Artificial Intelligence as Parasitic Computation*

Lauren E. Bridges

Department of Media Studies
University of Virginia

Since ChatGPT's public launch in November 2022, pundits, journalists, and tech moguls have sounded alarms about the long-term risks of Artificial General Intelligence (AGI), particularly its potential to displace workers and threaten humanity. Yet this focus on hypothetical futures obscures the pressing issues posed by existing AI systems: the intensification of misinformation, civil rights abuses, environmental harm from fossil-fuel-reliant data centers, increased demand for rare materials, and a surge in global e-waste.[1]

While discussions around mitigating AI's problems—reducing biases, curbing hallucinations, and managing resource consumption—are growing, they often miss a critical point: the way we conceptualize and engage with AI might be the real problem. In this essay, I argue three things: (1) "Artificial Intelligence" is a fundamentally flawed concept; (2) we must focus on where AI operates to fully understand its social and environmental toll; and (3) "parasitic computation" better describes AI's processes and impact.

Artificial Intelligence is an Oxymoron

The term "Artificial Intelligence" has been misleading since its inception in 1956, when researchers proposed that learning and other features of intelligence could be replicated by machines.[2] Yet learning is not the same as understanding, as any student or teacher knows. The term "artificial" implies something constructed to imitate, while "intelligence" denotes the capacity to understand. Together, they amount to "the imitation of understanding."

This discrepancy is evident when AI falters. For instance, when former NBA player Brandon Hunter tragically passed in 2023, an AI-generated headline reportedly declared him "useless at

42."[3] While this was a glaring failure, AI's real utility lies in accuracy, not understanding—for example, screening for skin cancer or predicting food safety risks. However, this focus on AI's cognitive limits diverts attention from its broader dependencies and impacts.

Where Does AI Feed?

To fully understand AI's impacts, we must ask: where is AI? AI's environmental and social consequences vary dramatically depending on the location of its infrastructure. Data centers, the backbone of AI computation, are not evenly distributed—they cluster in regions where they can exploit local resources, often creating significant trade-offs.

For instance, in West London, a 2023 moratorium on new high-density housing was enacted because data centers had already secured future power allocations, prioritizing servers over people. In Northern Virginia, energy demands from data centers are straining the grid, leading to increased reliance on diesel generators, which exacerbate environmental injustices in nearby communities.[4]

These examples illustrate how the location of data centers shapes their impact. Concentrated industrial computation can drain local resources, exacerbate social inequalities, and strain environmental systems. As I argue in the next section, viewing AI as a parasitic process pushes us to consider not only how it operates but also where it imposes its burdens.

AI as Parasitic Computation

AI is powered by an extractive and resource-heavy process. It depends on massive datasets, often collected without consent; engineers and "ghost workers" who classify and train data; and significant natural resources for hardware, energy, and cooling systems. This process generates vast amounts of waste, from e-waste to carbon emissions.

Because of these dynamics, I find it useful to think of AI as a form of parasitic computation (PC). A parasite feeds on its host, depriving it of nutrients, and while some parasites (like certain fungi) can be beneficial, most cause harm. Similarly, AI "feeds" on resources—data, energy, labor—and leaves behind social and environmental consequences. Reframing AI as PC highlights its extractive and consumptive nature.

For example, claims like "Every AI prompt is equivalent to pouring a 16-ounce bottle of water on the ground" or "AI uses as much energy as a small country" underscore its resource intensity. While these estimations are eye-opening, they risk oversimplifying AI's impact. Such measurements often rely on incomplete industry data, leaving critics perpetually one step behind tech companies' claims of improved efficiency. Moreover, these generalizations overlook the critical question of location, underscoring the importance of linking parasitic computation to specific sites of impact.

Rethinking Our Relationship with AI

Addressing AI's challenges requires more than just mitigation strategies or efficiency improvements—it calls for a fundamental shift in how we think about and engage with these systems. By recognizing AI as parasitic computation, we can better understand its extractive nature and demand greater accountability for its social and environmental

costs. And by focusing on where AI operates, we can uncover the localized consequences often hidden behind sweeping narratives of progress.

Artificial Intelligence, as a term and a concept, obscures more than it reveals. Perhaps it's time to abandon the oxymoron altogether and start calling it what it is: a process that imitates understanding while feeding on the very resources—human, environmental, and infrastructural—that sustain it. Only then can we begin to grapple with the true cost of AI.

ENDNOTES

1 Hogan, M., & Lepage-Richer, T. (2024). Extractive AI. Centre for Media, Technology, and Democracy. https://static1.squarespace.com/static/5ea874746663b45e14a384a4/t/6664bcdf2cbd2d1d11ba6c03/1717877983870/CJT_Hogan+%26+Lepage-Richer.pdf

2 McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Magazine, 27(4), Article 4. https://doi.org/10.1609/aimag.v27i4.1904

3 Hart, J. (2023, September 15). MSN readers are furious after an apparently AI-generated headline announced "Brandon Hunter useless at 42" after the former NBA player suddenly died. Business Insider. https://www.businessinsider.com/apparently-ai-generated-obituary-headline-nba-player-brandon-hunter-useless-2023-9

4 Bridges, L. E. (2024). Competing digital capacities: Between state-led digital governance and local data center tradeoffs. Information, Communication & Society, 0(0), 1–18. https://doi.org/10.1080/1369118X.2024.2331765.

FROM "PROMISE AND PERIL" *to Political Ecologies of Automation*

Megan Wiessner

Digital Technology for Democracy Lab, Karsh Institute of Democracy
University of Virginia

Debates over the environmental significance of AI often assume that the environmental problem with AI lies on the production side—due to these technologies' expanding consumption of water, energy, and hardware—while the environmental promise of AI lies in its application as a source of environmental optimization and insight. This framing of AI's "promise and peril" for the environment is widespread in both academic research and popular media.[1] It is also misleading. This is partly because it creates a false equivalence between known harms and speculative gains; the costs intrinsic to all applications of machine learning are, after all, creating dramatic new pressures on supply chains and energy systems, while specialized ecological applications of these tools are nowhere near cancelling these out at a systemic level. But it's also misleading because it ignores the fact that even AI applications for sustainability are embedded in the politics of resources, energy, land, and labor. Optimizing vehicle routes, water usage, or grid storage is no guarantee of environmental justice.[2]

In my work researching technology in construction materials and the construction industry, I've seen the many ways in which AI tools are now deployed in the name of saving resources. Machine learning is now used in sawmills, for example, to optimize log cuts and materials usage. Concerns about climate change have inspired researchers, start-ups, and existing firms to explore how machine learning might be used to decrease the carbon intensity of the built environment: by generating novel concrete mixes that might sequester more carbon; by optimizing building layouts for thermal performance; or by analyzing supply chain data to identify more sustainable procurement options.[3] This all sounds promising, but in the contemporary political economy of the built environment, "sustainability"—which remains a marginal concern—is hard to disentangle from resource optimization, cost-cutting, and profit-seeking. Using less material also means spending less money. Low-carbon structures become assets within real estate portfolios, justifying emissions or exposure

to harmful entanglements elsewhere. Claims of sustainability allow real estate developers to generate rent premiums that contribute to gentrification.[4] Reducing supply chain risk or developing new processes in the name of efficiency become ways of reducing the influence of trade labor.[5] These reduced costs, the industry promises, will help it build more, and faster, for cheaper.

This last point raises another problem with the idea that AI will just optimize environmental harms away. AI systems are technologies of automation, and when automation succeeds, it enables new increases in energy consumption and material throughput, sometimes in unexpected ways. Here, it might be helpful to draw an analogy with technologies of a previous era: the steam-powered looms of the eighteenth and nineteenth centuries. These were more than time and energy-saving devices. With less dependence on the creative knowledge of weavers, the prices and quality of garments declined (the beginning of the fast fashion conundrum that AI, by accelerating design turnover, improving targeted advertising, and optimizing global shipping costs, is only continuing to fuel.)[6] The profits to be made with the new machines increased the demand for coal to operate more and more of them (inaugurating the same climate catastrophe that coal-powered data centers continue to fuel two centuries later).[7] The growing demand for cotton precipitated an expansion of the chattel slavery economy in the U.S. South, which in turn degraded soils and encouraged aggressive expansionist policy against Indigenous nations.[8]

AI-driven automation in manufacturing, energy prospecting, and privately developed military technology is likely to lead to doing more with more, not more with less. Market-driven investments in automation result not in the quicker and more judicious accomplishment of the same tasks by the same actors but in an increase in activity by more and more actors. This is as true in industrial sectors as it is with all the AI slop clogging your search results. Without clear boundaries, the prospect of optimization encourages more production, declining profit margins, and new cycles of investment in AI hardware operation, with all that the latter entails. Most applications of AI are not "about" the environment, but that doesn't mean they won't have environmental effects.

Instead of hoping new environmental insights will cancel this out, a more holistic assessment of AI's environmental implications would ask where AI is being taken up and assess what automation in those sectors is likely to do. Instead of leaning harder into the promise of optimization, the current moment could be one for thinking about material limits and what kind of political programs might be based around them. This rethinking of AI is not motivated by fear, ignorance, or a romantic attachment to what is unique about the human. Instead, it's a lucid acknowledgment of the real organizational power of these technologies, of their environmental politics, and of the importance of popular deliberation and control over their use.

ENDNOTES

1 Lohr, S. (2023). Will A.I. Ruin the Planet or Save the Planet?. The New York Times. https://www.nytimes.com/2024/08/26/climate/ai-planet-climate-change.html / Hiar, C. (2023) AI Might Drive down Carbon — If Its Energy Use Is Kept in Check. POLITICO. https://www.politico.com/news/2023/09/27/ai-carbon-dioxide-energy-ee-power-00118452 / Cowls, J. et al. (2023) The AI Gambit: Leveraging Artificial Intelligence to Combat Climate Change—Opportunities, Challenges, and Recommendations. AI & SOCIETY 38, no. 1 283–307. https://doi.org/10.1007/s00146-021-01294-x.

2 Anson, A. et al. (2022). Water Justice and Technology: The COVID-19 Crisis, Computational Resource Control, and Water Relief Policy. AI Now Institute. https://ainowinstitute.org/wp-content/uploads/2023/02/water-justice-technology.pdf

3 Li, Z. et al. (2022) Machine Learning in Concrete Science: Applications, Challenges, and Best Practices. Npj Computational Materials 8, no. 1 1–17. https://doi.org/10.1038/s41524-022-00810-x / Tien, W. et al. (2022) Machine Learning and Deep Learning Methods for Enhancing Building Energy Efficiency and Indoor Environmental Quality – A Review. Energy and AI. https://doi.org/10.1016/j.egyai.2022.100198.

4 Goodling, E. Green, J. & McClintock, N. (2015). Uneven Development of the Sustainable City: Shifting Capital in Portland, Oregon. Urban Geography 36, no. 4: 504–27, https://doi.org/10.1080/02723638.2015.1010791 / Dalton, B. & Fuerst, F. The 'Green Value' Proposition in Real Estate: A Meta-Analysis. Routledge Handbook of Sustainable Real Estate.

5 Anderson, L. (2024). Strikes and Labor Negotiations Highlight Need for Resilient Supply Chains. The Supply Chain Xchange. https://www.thescxchange.com/blog/supply-chain-resiliency-strikes.

6 See for instance Harreis, H. et al. (2023) Generative AI in Fashion. McKinsey & Company. https://www.mckinsey.com/industries/retail/our-insights/generative-ai-unlocking-the-future-of-fashion.

7 Malm, A. (2013). The Origins of Fossil Capital: From Water to Steam in the British Cotton Industry. Historical Materialism 21, no. 1: 15–68, https://doi.org/10.1163/1569206X-12341279 / Carruth, A. (2014) The Digital Cloud and the Micropolitics of Energy. Public Culture 26, no. 2 (73): 339–64. https://doi.org/10.1215/08992363-2392093.

8 Hornborg, A. (2006) Footprints in the Cotton Fields: The Industrial Revolution as Time–Space Appropriation and Environmental Load Displacement. Ecological Economics 59, no. 1: 74–81, https://doi.org/10.1016/j.ecolecon.2005.10.009 / Silkenat, D. (2022). Scars on the Land, 1st ed. Oxford University Press. https://doi.org/10.1093/oso/9780197564226.003.0002.

THE NEED FOR RESEARCH ON *AI-Driven Climate Solutions*

Jonah Fogel

University of Virginia

The escalating global climate crisis demands innovative, scalable, and equitable solutions. AI offers a new set of tools capable of producing insights otherwise unidentifiable using conventional computing techniques or too expensive in time and resources to be done conventionally. AI is being used now to deliver, among other things, innovative strategies for decarbonizing energy systems, conserving biodiversity, and promoting climate-resilient urban planning. In this frame, two critical research priorities have emerged: 1) how best to leverage AI to drive data-informed, equitable climate solutions, and 2) understanding and mitigating AI's environmental impacts. These priorities are critical for unlocking the potential of artificial intelligence while reducing pollution and minimizing potential harms to society.
AI tools hold the potential to provide actionable insights more quickly at lower costs, enabling more informed decision-making in climate action. However, many AI systems lack transparency, making them susceptible to biases that can exacerbate existing inequities. Without proper safeguards, these systems may disproportionately impact marginalized communities, hinder equitable climate interventions, and perpetuate structural disparities. This highlights an urgent need for research focused on building AI frameworks that embed fairness, transparency, and accessibility from the outset. Such research must explore methods for explainability, bias mitigation, and inclusive data collection to ensure AI contributes to just and sustainable climate solutions.

The exponential growth of data centers, driven in part by AI-specific workloads, is already creating unprecedented energy demands. For example, Virginia has experienced sharp increases in energy consumption linked to data centers, with demand projected to double within the next 10 years. A federal policy shifts away from renewable energy, and deregulation of emissions standards will slow

the adoption of sustainable practices and prolonging reliance on fossil fuels, making it harder to meet climate goals. Together, these issues threaten the state's ability to achieve renewable energy targets under policies like the Virginia Clean Economy Act. Research must explore sustainable approaches, such as developing energy-efficient algorithms, green computing, and integrating renewable energy sources into AI systems.

Advancing these research priorities will not only accelerate climate solutions but also position AI as a tool that aligns with the principles of justice and sustainability. To achieve these goals, interdisciplinary collaboration and forward-thinking policies will be crucial. By investing in these efforts now, we can ensure a future where AI serves as a force for good in addressing the climate crisis.

The Environmental Institute (EI) at UVA is deeply aligned with these research needs. The institute has invested in innovative projects which focus on using AI to address critical climate challenges and ensure just and sustainable outcomes for all communities.

Climate Justice Numerical Modeling: This project focuses on developing advanced numerical models that simulate the environmental and social impacts of climate policies, particularly those affecting vulnerable and marginalized communities. The models aim to provide policymakers with tools to ensure climate justice is a central component of decision-making processes.[1]

Blueprints AI: This project is dedicated to using artificial intelligence to generate equitable, just, and sustainable climate policies. These AI-driven frameworks will help guide policymakers at both the local and national levels in crafting climate strategies that are inclusive and responsive to all stakeholders, especially those in vulnerable communities.[2]

AI and Society: This project investigates the broader societal impacts of AI, particularly its ethical implications in areas like energy usage, workforce dynamics, and social equity. Researchers seek to ensure that AI technologies are implemented in ways that are fair and inclusive, addressing critical concerns around ethics in the context of climate solutions.[3]

AI for Localized Climate Policy: This project aims to leverage AI technologies to assist local governments in developing climate policies tailored to the specific needs of their communities. By incorporating local data and stakeholder input, the project helps create more effective and equitable climate policies, enhancing community resilience to climate change.[4]

These projects and emerging industry partnerships demonstrate EI's interdisciplinary strength, leveraging diverse academic and practical insights to tackle complex climate challenges, by funding projects that integrate expertise from data science, environmental science, statistics, media studies, history, and political science.

ENDNOTES

1 Pusede, S., Fields, K., Yu, S., Hadayia, J., Carlton, A.M., Cohen, R., & Gao, Z. (2021, October 13). A Novel Modeling Framework to Support Community Preferences in Air Quality Decision-making. University of Virginia Environmental Institute. https://environment.virginia.edu/our-work/climate-justice-numerical-modeling

2 Sloane, M., Nyelele, C., & Mamalakis, A. (2021, October 13). Blueprints for AI Equity and Transparency in Climate Science. University of Virginia Environmental Institute. https://environment.virginia.edu/our-work/blueprints-ai

3 Hicks, M. & Linstrum, E. (2022, April 4). Rethinking the Inevitability of AI: Historicizing Climate Impact and Discussing Best Practices for Energy and Water Consumption. University of Virginia Environmental Institute. https://environment.virginia.edu/ai-and-society

4 Mamalakis, A., Hartvigsen, T., Singh, S., & Papalexiou, S. (2025) Using Advanced AI for Simultaneous Bias-Correction and Down-scaling of Climate Model Projections. University of Virginia Environmental Institute. Retrieved from: https://environment.virginia.edu/our-work/ai-localized-climate-policy

ENVIRONMENTAL JUSTICE AND *AI Through an African Context*

Maria Lungu

Digital Technology for Democracy Lab, Karsh Institute of Democracy
University of Virginia

Artificial intelligence (AI) is projected to infuse up to $16 trillion into the global economy by 2030, prompting global conversations concerning ethical, social, and environmental implications.[1] Artificial intelligence is often viewed as a transformative technological advancement bolstering capitalistic structures and service delivery across sectors.[2] However, it is also viewed as a tool of power, exacerbating global inequities and exploiting natural resources.

More specifically, the environmental justice dimensions of AI remain underexplored, especially within the African context—a region integral to the AI supply chain through the extraction of critical raw materials.[3] This oversight continues to raise concerns given Africa's considerable role in the global AI supply chain, mainly through raw material extractions, but also through issues related to inequity and environmental harm.[4] As researchers, we question whether systems are fixable or should be completely reimagined. We also question what aspects of environmental justice concerning AI are overlooked, ignored, underplayed, etc. This essay highlights pointed topics in the African context to inform future research and policy considerations.

1. Unequal Distribution of Environmental Burdens

Artificial intelligence involves extracting and using critical raw materials like cobalt, graphite, platinum, tantalum, lithium, etc.[5] These raw materials are integral to the hardware supporting AI systems. Lithium and cobalt are essential for producing lithium-ion batteries, which are the primary energy source for many AI-enabled devices.[6] Graphite is a key anode material, enhancing battery efficiency and performance.[7] Tantalum is widely used in capacitors within

semiconductors, powering processors, and Graphics Processing Units (GPUs). Platinum creates high-sensitivity sensors and energy-efficient fuel cells.[8] Additionally, elements like neodymium and dysprosium enable the creation of high-strength magnets for robotics and autonomous systems.[9] Together, these materials underpin the energy storage, processing capabilities, and infrastructure for AI. However, this extraction occurs disproportionately in African nations like South Africa, Rwanda, Zambia, the Democratic Republic of Congo (DRC), and other central African countries.[10]

Artificial intelligence often benefits systems and institutions in wealthier nations despite efforts to extend development in the Global South. However, resource extraction and environmental degradation (deforestation, water contamination, soil erosion, etc.) are often concentrated in African countries.[11] Many African countries have resorted to source-banning unprocessed raw materials. For example, Zimbabwe recently instituted a lithium ban to prevent the industrial powers from capitalizing on material extraction without domestic value-addition.[12] Artificial intelligence conversations sparsely acknowledge these inequities or the environmental toll on resource-rich African regions. Addressing these inequities requires an inclusive dialogue with affected communities to consider what ethical sourcing, equitable resource-sharing mechanisms, and sustainable practices will look like (if possible).

2. Exploitation of local communities

One of the significant environmental justice concerns associated with AI supply chains involves the labor conditions for mineral extraction.[13] In some regions, the integration of automation and AI focuses on cutting labor costs, enhancing productivity, and saving resources such as fuel in developed institutions.[14] However, labor issues take on a more critical dimension in resource-rich African contexts, as mineral extraction often involves hazardous working conditions, exploitation, and inadequate compensation.[15]

For example, the labor conditions in many mining regions are sometimes marked by a lack of regard for worker safety or environmental protection.[16] For example, artisanal miners, often including children, are exposed to toxic chemicals without adequate safeguards.[17] Yet these injustices are frequently sidelined in global discussions about AI ethics, where the focus is on data privacy, algorithmic bias, and, ultimately, the environmental costs of material extraction. Addressing these injustices requires expanding the dialogue around AI ethics, labor rights, and the human cost of mineral extraction, ensuring AI advancements do not come at the expense of vulnerable populations.

3. Long-term environmental damage

In addition to labor conditions, the environmental justice rhetoric also focuses on the long-term environmental damage attributed to extraction. Unfortunately, the environmental consequences of mineral extraction extend beyond the immediate impact on African ecosystems. Once resources are extracted, sometimes regions are abandoned.[18] Communities are left with abandoned mines that become dangerous and toxic waste sites, and those same communities

suffer from polluted water and soil for generations.[19]

Often, the literature has considered how emissions from AI data centers in Western communities affect those regions. However, the long-term damages in Africa receive minimal attention in environmental justice narratives. Perhaps because it is too soon to understand what the damage looks like, or perhaps because there is an unwillingness to confront the extent of this damage.

4. Colonial legacies and resource governance

Finally, much can be said about mineral extraction to the detriment of African regions in the context of colonial legacies.[20] Currently, and unfortunately, colonial legacy underpins many of the current resource extraction dynamics in African countries.[21] Research has outlined evidence of multinational corporations extracting resources with minimal accountability, perpetuating a cycle of dependency and underdevelopment.[22] This is another context where environmental justice scholars must confront the neo-colonial structures that continue to disenfranchise African nations and communities, leaving them with degraded environments and structures.

Concluding: Toward inclusivity in environmental justice

One of the most important institutional changes to addressing these gaps includes increasing African representation in global governance conversations. These voices often need to be more adequately represented, and thus, their inclusion could foster better framing for environmental justice and AI ethics conversations. This is especially the case given how the resource-supplying African nations bear the environmental and social costs.[23]

Additionally, consider investing in sustainable extraction technologies and local economic development to mitigate environmental degradation and promote equity. Highlighting these issues in AI research and advocacy assists in a comprehensive understanding of the environmental justice implications of AI in the Global South. Such efforts are necessary for the global pursuit of AI to avoid perpetuating environmental injustices that disproportionately affect Africa, undermining the ethical aspirations of these transformative technologies.[24]

ENDNOTES

1 PwC analysis. (2017). Sizing the prize, PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution. https://www.pwc.com/gx/en/issues/artificial-intelligence/publications/artificial-intelligencestudy.html#:~:text=Total%20economic%20impact%20of%20AI%20in%20the%20period%20to%202030&text=AI%20could%20contribute%20up%20to,come%20from%20consumption%2Dside%20effects.

2 Lungu, M., & McCue, C. (2024). Evaluating Regulatory Structures for the Expansion of Artificial Intelligence. Journal of AI Law and Regulation, 1(3), 310-330.

3 Adams, R. (2022). AI in Africa: Key concerns and policy considerations for the future of the continent. Policy Brief. https://afripoli.org/ai-in-africa-key-concerns-and-policy-considerations-for-the-future-of-the-continent

4 Idu, M. A., Onodi, B. E., & Nwosu, U. (2022). Artificial Intelligence And Business Sustainability In Sub-Sahara Africa. JORMASSI Journal of Research in Management and Social Sciences, 8(2), 64-78.

5 Lakshmi, D., Nambiar, R., & Kumar, P. (2025). The role of artificial intelligence in materials science: a review. Advances in Sustainable Materials, 209-228.

6 *ibid.*

7 Lemma, A. (2025). Critical minerals, critical moment: Africa's role in the AI revolution. ODI. https://odi.org/en/insights/critical-minerals-critical-moment-africas-role-in-the-ai-revolution/

8 *ibid.*

9 Creamer, M. (2024). Artificial intelligence used to eliminate need for rare earths in permanent magnets. Mining Weekly. https://www.miningweekly.com/article/artificial-intelligence-used-to-eliminate-need-for-rare-earth-in-permanent-magnets-2024-06-11

10 Azaroual, F. (2024). Artificial Intelligence in Africa: Challenges and Opportunities. Research papers & Policy papers on Economic Trends and Policies, (2416).

11 *ibid.*

12 Fang, M. M. (2024). Climbing up the critical mineral value chains: The global south and green industrialization in an era of disruption. Vand. J. Transnat'l L., 57, 795.

13 Dauvergne, P. (2022). Is artificial intelligence greening global supply chains? Exposing the political economy of environmental costs. Review of International Political Economy, 29(3), 696-718.

14 Wang, Y., & Li, Y. (2023). Chinese economic growth and sustainable development: Role of artificial intelligence and natural resource management. Resources Policy, 85, 103996.

15 *ibid.*

16 Kapatamoyo, M. (2024). Artificial Intelligence in Natural Resources Management: Selected Case Studies from Africa. https://www.econstor.eu/bitstream/10419/302527/1/ITS-Seoul-2024-paper-128.pdf

17 Mienye, I. D., Sun, Y., & Ileberi, E. (2024). Artificial intelligence and sustainable development in Africa: A comprehensive review. Machine Learning with Applications, 100591.

18 Gikunda, K. (2024). Harnessing artificial intelligence for sustainable agricultural development in Africa: opportunities, challenges, and impact. arXiv preprint arXiv:2401.06171.

19 Singh, A., Kanaujia, A., Singh, V. K., & Vinuesa, R. (2024). Artificial intelligence for Sustainable Development Goals: Bibliometric patterns and concept evolution trajectories. Sustainable Development, 32(1), 724-754.

20 Hassan, Y. (2023). Governing algorithms from the South: a case study of AI development in Africa. AI & Society, 38(4), 1429-1442

21 Adams, R. (2021). Can artificial intelligence be decolonized?. Interdisciplinary Science Reviews, 46(1-2), 176-197.

22 *ibid.*

23 Bibri, S. E., Krogstie, J., Kaboli, A., & Alahi, A. (2024). Smarter eco-cities and their leading-edge artificial intelligence of things solutions for environmental sustainability: A comprehensive systematic review. Environmental Science and Ecotechnology, 19, 100330.

24 Lungu, M. (2024). Enhancing public service delivery in government procurement: A review exploring the role of artificial intelligence and automative structures. Handbook of Public Service Delivery, 188-204.

IS GENAI *Worth the Costs?*

Siobhán Loughney

University of Virginia

Rachel Leach

University of Virginia

As research assistants studying the impact of the increasing integration of software and Artificial Intelligence into electric vehicles, we are highly interested in the expanding energy infrastructure required to support the proliferation of AI systems; this essay explores these themes and considers if the expansion of this technology is worth the costs.

GenAI and Energy:

To balance sustainability pledges with plans for rapid growth, large tech companies are turning to nuclear energy to power their endeavors. Take Microsoft, for example. Back in 2020, the company announced a commitment to being carbon-negative by 2030.[1] To back its ambitious AI plans, Microsoft has acquired a twenty-year power purchase agreement (PPA) for the reopening and operation of the Three Mile Island nuclear power plant in Dauphin County, PA, providing the company the power equivalent of 800,000 U.S. homes.[2] Although the plant operated until 2019, the 1979 partial meltdown of its second reactor still looms in the back of the public's mind. Furthermore, opinions are split on the use of nuclear energy to power data centers at the generation and use sites. In Pennsylvania, some locals fear a repeat of the infamous partial meltdown and raise concerns for health and safety, while others are hopeful for the economic boost the plant's jobs could bring or its contributions as a carbon-free energy source.[3] Down U.S. Highway 15, in Loudoun County, VA, residents have mixed opinions on the appeal, or lack thereof, of the influx of data centers in their towns.[4] The county now boasts the greatest concentration of data centers in the world.[5]

While residents of these areas and many others across the nation reckon with what the rapid expansion of AI and the related infrastructure means for their towns, from economic opportunity to climate impacts, industry has no plan to slow down. Global data center capacity demand is projected to increase by about 19 to 22 percent annually through 2030,[6] and tech companies are keeping pace. As tech giants push to revive nuclear energy to power their AI plans, the rest of us are left wondering how far is too far for these companies to go. The intensely extractive needs of AI, pulling on the energy grid, water, silicon, and many minerals,

seem to be taking priority while humans are left to watch their environmental safety take the backburner. AI's rapid expansion threatens "exacerbating droughts and desertification, disrupting ecosystems and fisheries, triggering conflict, and amplifying water inequalities by diverting water supplies towards technology hubs."[7] Bender et al. write, "Increasing the environmental and financial costs of these models doubly punishes marginalized communities that are least likely to benefit from the progress achieved by large LMs and most likely to be harmed by negative environmental consequences of its resource consumption."[8] Why must we divert all these resources to AI while people suffer the harms of the extraction? Is AI really a positive force in the world if it leaves ruins in its wake? At some point, we must question if the ever-mounting presence of AI is worth the massive drain on resources that it requires. In the U.S. alone, AI energy consumption has grown from 76 TWh in 2018 to 176 TWh in 2023, a shift from 1.9% to 4.4% of national energy consumption.[9] It is time to get more serious about the necessity of AI's expansion and curb the seemingly unbounded growth of its footprint on our planet. We must rethink the future we want to see and redefine what qualifies as an economic benefit– if communities are suffering through the effects of data center proliferation and the mass revival of nuclear energy while seeing little to no gains from the use of AI, we cannot allow industry's profits to define this era of exploitation as a positive force.

The limits of GenAI will not be solved through investing in more data and energy

AI models will give wrong answers regardless of how much data they are trained on and energy they are able to use. Researchers investigated the accuracy of various models including ChatGPT and Meta's LLaMA and as the models developed to include more data and parameters.[10] They found that as these models are scaled up through increased training data, they actually become less reliable, as instead of avoiding questions they do not know the answer to which they will answer inaccurately. Additionally, the researchers found that even with questions that humans could answer easily, they couldn't be sure a GenAI model could do the same.

Second, GenAI models are unable to accurately represent marginalized groups. As Ari Waldman, law professor at UC Irvine, puts it: "Algorithms increase the power of the past over the present and future."[11] In other words, because in the past, the government did not collect data on, and often did not even recognize, certain groups (queer and trans people, people of color), algorithms built based on "big data" are likely to be biased against, harmful to, or misrepresentative of, these groups.

In Critical Questions for Big Data, boyd & Crawford argue that knowledge produced through analysis of mass amounts of data is increasingly seen as "a higher form of intelligence" such that other forms of evidence that marginalized groups have historically relied on to advocate for and protect themselves is undervalued.[12] Take the courts for example; Waldman explains that in this

setting, "anecdotal accounts or ethnographic interviews" are seen as less persuasive than "hard data." This leads to the interviews, anecdotes, and personal stories that marginalized groups rely on to detail "police harassment and profiling" for example, to be undervalued.[13] As Jen Jack Gieseking points out, information on, by, and for marginalized groups will remain "small" so long as "big data" algorithms remain ascendent.[14] We should be keenly aware of the potential limits and drawbacks of algorithms built on data largely collected on, rather than for and with marginalized groups, as they proliferate.

Conclusions:

As fourth-year college students, we witnessed the entrance of large language models (LLMs) into the classroom in the midst of our undergraduate studies. The typical course syllabi sections on plagiarism and academic honesty quickly received artificial intelligence-specific addenda as universities grappled with decisions of managing AI use at the classroom and institutional levels. Some professors encourage students to use AI as a tool for brainstorming and drafting or a mechanism to prompt further research, like scrolling through Wikipedia before referencing a peer-reviewed journal.[15] Others strictly prohibited its use and issued strong warnings against generating written work with the aid of LLMs. Still, a greater, more concerning trend persists below the surface of AI in academics. While anyone could agree that relying solely on AI to generate academic work is dishonest, we are not paying close enough attention to how even its casual use for studying or assistance has eroded students' curiosity and confidence.

In several recent conversations with peers, as a question has arisen, a student has chimed in with "I'll ask ChatGPT," which seems to have eclipsed the phrase "google it." A recent report from Microsoft and Johns Hopkins suggests that this shift jeopardizes critical thinking and intellectual development. The researchers found that individuals who use GenAI to complete tasks produce a less "diverse set of outcomes" than those not using this technology. This can be explained, in part, by the fact that work guided by GenAI loses much of its grounding in personal experience and context. Additionally, this report finds that use of GenAI shifts users very understanding of what critical thinking entails. The researchers found that "knowledge workers" who use GenAI include "refining prompts" and "assessing AI generated content" in their definition of critical thinking, tasks that offload rather than develop skills such as analysis and synthesis which are typically understood as central to critical and reflective thinking.[16]

Through repeated exposure to even mild use of LLMs for academic purposes, students are becoming conditioned to turn to it as a first resort, and worse, are losing trust in themselves to seek out information or produce their own quality work. This phenomenon of self-doubt worsens the more we learn to trust blindly in LLMs.[17] Our parents once mocked our ease of access to seemingly endless information through search engines by reminding us that they grew up paging through physical library catalogs for school projects. Now, searching through online journals and news is falling by the wayside as asking LLMs for quick answers takes the driver's seat in information discovery. Even without the

question of the accuracy of the information provided by LLMs, students are certainly losing opportunities to happen upon information that spurs more questions or a more robust understanding of their original query. Without needing to comb through or even skim over articles and web pages in search of answers to our questions, we lose valuable time engaging with our own curiosities and incidental discovery of information.

As the use of AI expands and this technology is integrated into more and more products, from automobiles to writing assistants, we are left wondering if this technology is worth its cost. The steep energy cost and dubious accuracy of the technology coupled with its impact on the students around us makes us say no.

ENDNOTES

1 See: https://blogs.microsoft.com/blog/2020/01/16/microsoft-will-be-carbon-negative-by-2030/

2 Moss, S. (2024, September 20). Three Mile Island nuclear power plant to return as Microsoft signs 20-year, 835MW AI data center PPA. Data Center Dynamics. https://www.datacenterdynamics.com/en/news/three-mile-island-nuclear-power-plant-to-return-as-microsoft-signs-20-year-835mw-ai-data-center-ppa/

3 See: https://stateimpact.npr.org/pennsylvania/2024/09/25/fear-hope-among-mixed-reactions-as-three-mile-island-plans-to-restart/

4 See: https://loudouncountymagazine.com/about/f/loudoun-residents-weigh-in-on-data-centers

5 Mabeus-Brown, C. (2022, March 24). Report: Va. data center projects totaled $6.8B in 2021. Virginia Business. https://virginiabusiness.com/report-va-data-center-projects-totaled-6-8b-in-2021/

6 Srivathsan, B Sorel, M. Bhan, S. Sachdeva, P Batra, H. Sharma, R. Gupta, R. & Choudhary, S. (2024). AI data center growth: Meeting the demand. McKinsey & Company. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand

7 Hogan, M., & Richer, T. (2024). Extractive AI. Centre for Media, Technology and Democracy. https://www.mediatechdemocracy.com/climatetechhoganlepagericher

8 Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? . Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623. https://doi.org/10.1145/3442188.3445922

9 Shehabi, A., Smith, S., Sartor, D., Brown, R., Herrlin, M., Koomey, J., Masanet, E., Horner, N., Azevedo, I., & Lintner, W. (2016). United States Data Center Energy Usage Report (LBNL–1005775, 1372902; p. LBNL–1005775, 1372902). https://doi.org/10.2172/1372902

10 Zhou, L., Schellaert, W., Martínez-Plumed, F., Moros-Daval, Y., Ferri, C., & Hernández-Orallo, J. (2024). Larger and more instructable language models become less reliable. Nature, 634(8032), 61–68. https://doi.org/10.1038/s41586-024-07930-y

11 Waldman, A. E. (2024). Legibility Double Binds. Boston University Law Review, 104(4). https://www.bu.edu/bulawreview/files/2024/06/WALDMAN.pdf

12 boyd, d., & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, 15(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878.

13 Waldman, A.E. (2024).

14 Gieseking, J. J. (2018). Size Matters to Lesbians, Too: Queer Feminist Interventions into the Scale of Big Data. The Professional Geographer, 70(1), 150–156. https://doi.org/10.1080/00330124.2017.1326084

15 See: https://medschool.duke.edu/news/chatgpt-enters-classroom

16 Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025, April). The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers. Proceedings of the ACM CHI Conference on Human Factors in Computing Systems. https://www.microsoft.com/en-us/research/publication/the-impact-of-generative-ai-on-critical-thinking-self-reported-reductions-in-cognitive-effort-and-confidence-effects-from-a-survey-of-knowledge-workers/

17 See: https://pmc.ncbi.nlm.nih.gov/articles/PMC11082730/#:~:text=Blind%20trust%20in%20LLMs%20can, human%20ju dgment%20and%20LLM%20outputs.

REIMAGINING AI'S IMPACT ON THE ENVIRONMENT: *Finding Hope in Small-Scale Interventions*

Mehan Jayasuriya

Mozilla Foundation

Looking back on the conversations I had over the two days at the Karsh Institute during the "Reimagining AI" workshop, what stuck with me most were the questions that were raised about artificial intelligence, its impacts and its potential future. As a grantmaker, I held a different perspective from most of the others in the room: rather than engaging with these questions in the classroom, through publications or through policy, I fund projects that test out promising approaches to solving problems in the real world. In that context, there were two questions I asked on the panel about the environmental impacts of AI that seemed to resonate: How might we better measure the environmental and climate impacts of AI systems? And are there uses of AI which might help communities and activists address pressing environmental issues? This year, I am working with 10 projects around the world that I have funded to test approaches to answering those questions;[1] below I will give examples of two projects that exemplify our approach.

One thing we're already seeing in the field of measuring AI's considerable environmental impacts is that big tech companies will do whatever they can to obfuscate the true environmental costs of their products. One way they have historically done this is to build proprietary, internal metrics that paint a rosy picture of their impact; this also makes it very difficult to perform an "apples-to-apples" comparison between two different companies. At Mozilla, we are big believers in open-source tools providing a "neutral" alternative to proprietary tools and we hope to support such projects which might eventually become universal standards for

the industry. One such tool we are supporting this year is called CodeCarbon[2], a volunteer-run, open-source utility that estimates the amount of carbon dioxide (CO2) produced by the computing resources used to execute a piece of code. CodeCarbon can calculate the carbon impact of a piece of code based on the compute region in which it is run and can help developers understand how to reduce the environmental impact of their programs as they are being written. CodeCarbon will be used to measure the efficiency of projects during the Frugal AI Challenge[3] taking place during the AI Action Summit[4] in France in 2025.

On the other end of the spectrum, we know that around the world, communities are seeking to push back against environmentally harmful projects and extractive uses of their land, but sometimes lack the data and resources to make their case to regulators. In coastal Kenya, there was great concern in the community about a proposal to build a nuclear reactor in an area designated as a marine wildlife reserve. The project was moving full steam ahead (due in part to pressure from multinational corporations seeking a contract to build the reactor), but little time had been spent investigating the potential impact of the project on the wildlife who live in the nearby waters. The Center for Justice Governance & Environmental Action[5] (CJGEA) in Kenya proposed a project, led by researcher Benson Mbani[6], that would use computer vision to identify and count undersea species quickly and cheaply. Benson and his team collected hours of video footage shot by divers, which were then fed into the model, quickly producing scientifically accurate tallies of the many species that could be impacted by the proposed reactor. As a result of this project and CJGEA's ongoing activism, the local regulator in Kenya recently acknowledged that the proposed project site is an important environmental and cultural area and that a formal environmental impact study should be undertaken. CJEGA is acknowledged as a "consultant stakeholder" in the report that was published[7] and it is likely that their advocacy and tools will help shape the impact assessment to come.

While these are relatively small wins in the grand scheme of things—especially considering the vast sums of money being invested in the largest-scale AI products—we are heartened to see small projects like these that are challenging the status quo and having a real impact. In partnership with communities, academics, and activists, we believe that practitioners like CodeCarbon and Benson Mbani can help to nudge the future of AI in a direction that is less harmful to people and the environment.

ENDNOTES

1 Mozilla Foundation. (2024, February 5). Open-source AI for Environmental Justice: Introducing the 2024 Mozilla Technology Fund cohort. https://foundation.mozilla.org/en/blog/open-source-AI-for-environmental-justice/

2 CodeCarbon.io. (n.d.). Codecarbon.io. https://codecarbon.io/

3 Committee, F. A. C. O. (n.d.). Frugal AI Challenge. 2025 Frugal AI Challenge. https://frugalaichallenge.org/

4 elysee.fr. (n.d.). Artificial Intelligence Action Summit. https://www.elysee.fr/en/sommet-pour-l-action-sur-l-ia

5 Stanley, C. (n.d.). Center for justice. Center for Justice Governance & Environmental Action. https://www.centerforjgea.com/

6 Mbani, B. (n.d.). https://www.geomar.de/en/bmbani

7 NCEA - EN. (2025, February 3). Review of the sesa for the nuclear power programme-kenya. https://www.eia.nl/en/project-database/review-of-the-sesa-for-the-nuclear-power-programme-kenya/

AI MINIMIZATION AS TECHNOLOGICAL PROGRESS: *Lessons for a Global Governance Agenda*

Jess Reia

School of Data Science and Digital Technology for Democracy Lab
University of Virginia

In the book "The Right to be Cold," Sheila Watt-Cloutier[1] describes how the Inuit communities in the Arctic witness the climate crisis daily, while environmental justice issues remain an abstract concept for most people worldwide. Ailton Krenak, who writes from the Tropics in "Ideas to Postpone the End of the World,"[2] questions the legitimacy of institutions in charge of protecting heritage, knowledge and humanity – from universities to museums and intergovernmental organizations – in not letting "the planet be devoured by mining operations." Both activists define the relationship between their peoples and the land as intrinsically connected to their cosmovision. Indigenous communities have been experiencing threats to their world throughout centuries of colonization and, by using creativity as a survival strategy, offer possible paths forward.[3] These Non-Western cultures can offer us alternative ways to understand technology, centering environmental justice as something fundamentally part of our life in this planet.

Technology is a telling lens to examine how Western societies often imagine progress and innovation as replacements; this imagination suggests that improved devices and models will leave behind the obsolete, the unavailing and the past. Media scholars show us how this narrative is not entirely true, as technologies coexist, are repurposed and make comebacks. Specific domains, such as computer and data sciences, receive robust state and non-state investments to compete internationally in the progress race. Currently, data is a foundation

for other systems and imaginaries of efficiency (and innovation), having shapeshifted into an incredibly relevant asset over the past 150 years.[4] Big data became the cornerstone of automation and artificial intelligence (AI) models and systems. AI can be many things, including a computer system designed to perform tasks and decision-making on behalf (or instead) of humans. While AI-based automation might serve as a proper response in some cases, it is not an answer to all of our problems – or an idealized force that will take over Earth. Instead, the damage being done by these technologies is concrete, and it is already here, affecting real people and territories.

Western conceptions of progress usually prevent us from seeing degrowth, minimization and less technology as desirable perspectives for the future. When progress means plowing ahead at all costs, environmental concerns do not occupy the forefront of the technological and developmental agenda. The global competition for AI leadership is a great example of progress at all costs. Although AI has existed for decades, its presence in the public debate and everyday vocabulary has greatly increased in the last five years, especially after large language models (LLMs) and generative AI (genAI) were made available to non-specialist users worldwide. This proliferation amplifies existing problems (such as mis/disinformation and online gender-based violence) and caused new problems to emerge (like the easiness to create and share deep fakes during election campaigns). Investments in the AI industry reached levels never seen before and, suddenly, AI is being used in the platforms we interact with, in our personal communications, learning processes and government services. New actors and established big tech companies enter the profitable AI market, certifications are created and dangerous threats to democracy become visible.[5]

The way we talk about AI matters. From global summits and forums to roadmaps and policy briefs, there is a growing call for sustainable and public interest AI.[6] While AI is touted as a sustainable solution to climate change, presented in expos and trade shows around the world, large language models and data centers – such important components of the current AI agenda – are consuming massive amounts of energy, water and space. Greenwashing is not new, nor is it a specific problem of big tech companies. The "myth of sustainability" was "invented by corporations to justify their theft of our idea of nature. No company on this earth is sustainable, no matter what they say".[7]

We need to overcome the disgust caused by deceleration, regress, limitations and degrowth. For Yuk Hui, "the great acceleration that has taken place in recent decades has also led to various forms of destruction, cultural, environmental, social, and political."[8] Instead of only talking about impact or harm, maybe it is time to address AI's destruction, from extractivism and overconsumption to its role in warfare and politics. The turn of Big Tech to the far-right politics in the US impact communities and territories way beyond the country's borders. While burning fossil fuels, for example, AI can never be sustainable. We have been witnessing the symbiotic work of power, wealth and technosolutionism – the "desire to jump on technological solutions as a

quick and flawless way to solve complex real-world problems."[9]

Efforts emerging from civil society organizations such as the Virginia Data Center Reform Coalition, together with global demands to reduce AI's environmental harms across its entire supply chain and lifecycle,[10] offer guidance on potential next steps. What if, instead of plowing ahead on AI development and adoption, we asked more questions such as: Is AI the best solution to the problem in front of us? How much of the solution is also the problem? Or, being even more straightforward, do we need AI in a burning planet?

My call is to minimize AI use and acquisition. What if we reframe the way we see technological progress as minimization? These are five starting points for a global governance agenda:

Decreasing AI adoption: we need a global governance agenda that treats AI as a tool embedded in ethical, digital rights and environmental issues, concerned in strictly assessing the actual need for its deployment.

Fossil-free AI infrastructure: data centers, supercomputers and other infrastructure powering artificial intelligence require massive amounts of electricity and water to keep on functioning, sometimes with a short life cycle and still relying on fossil fuels. AI must be fossil-free, run on renewable energy and no longer a provider of services for oil and gas companies.

Destruction assessment: shift risk and impact assessments to an approach that evaluates its burdens and destruction. AI should be used when the benefits outweigh the destruction it causes.

Sovereignty-focused: mineral, technological, geopolitical and Indigenous sovereignty should be aspects considered in a process of AI minimization globally, in which nations and jurisdictions at the margins can have a voice, protect their territories and maintain agency.

Non-Western perspectives for the AI agenda: listen to Non-Western communities, land protectors, and other forms of collaborative knowledge production on how the AI agenda should advance.

AI alone won't save us. From the tropics to the arctic, there are other ways to reimagine AI and our relationship with technology.

ENDNOTES

1 Watt-Cloutier, S. (2019). Le droit au froid: Le combat d'une femme pour protéger sa culture, l'Arctique et la planète. Montréal: Écosociété.

2 Krenak, A. (2020). Ideas to postpone the end of the world. Toronto: Anansi International.

3 Danowski, D., & de Castro, E. V. (2016). The ends of the world. Cambridge, UK: Polity Press.

4 Leonelli, S. (2019). Data - from Objects to Assets. Nature, 574 (7778): 317–20. https://doi.org/10.1038/d41586-019-03062-w.

5 Two examples in different scales are private consultancy firms drafting municipal laws disguised as master plans in Brazil (Reia & Cruz, 2025) and the current take-over of entire governmental agencies in the US by big tech billionaires aligned with far-right agendas. Reia, J., & Cruz, L. (2025). Smart-Washing the City: A Study on the Privatization of Urban Digital Infrastructures in the Global South. International Journal of Communication, 19, 23. Retrieved from https://ijoc.org/index.php/ijoc/article/view/22203.

6 The 2025 AI Action Summit in Paris focused on building a "public interest AI platform" and discussed nuclear power as a "sustainable" response to the intensive use of resources required by large scale adoption of data-centric systems

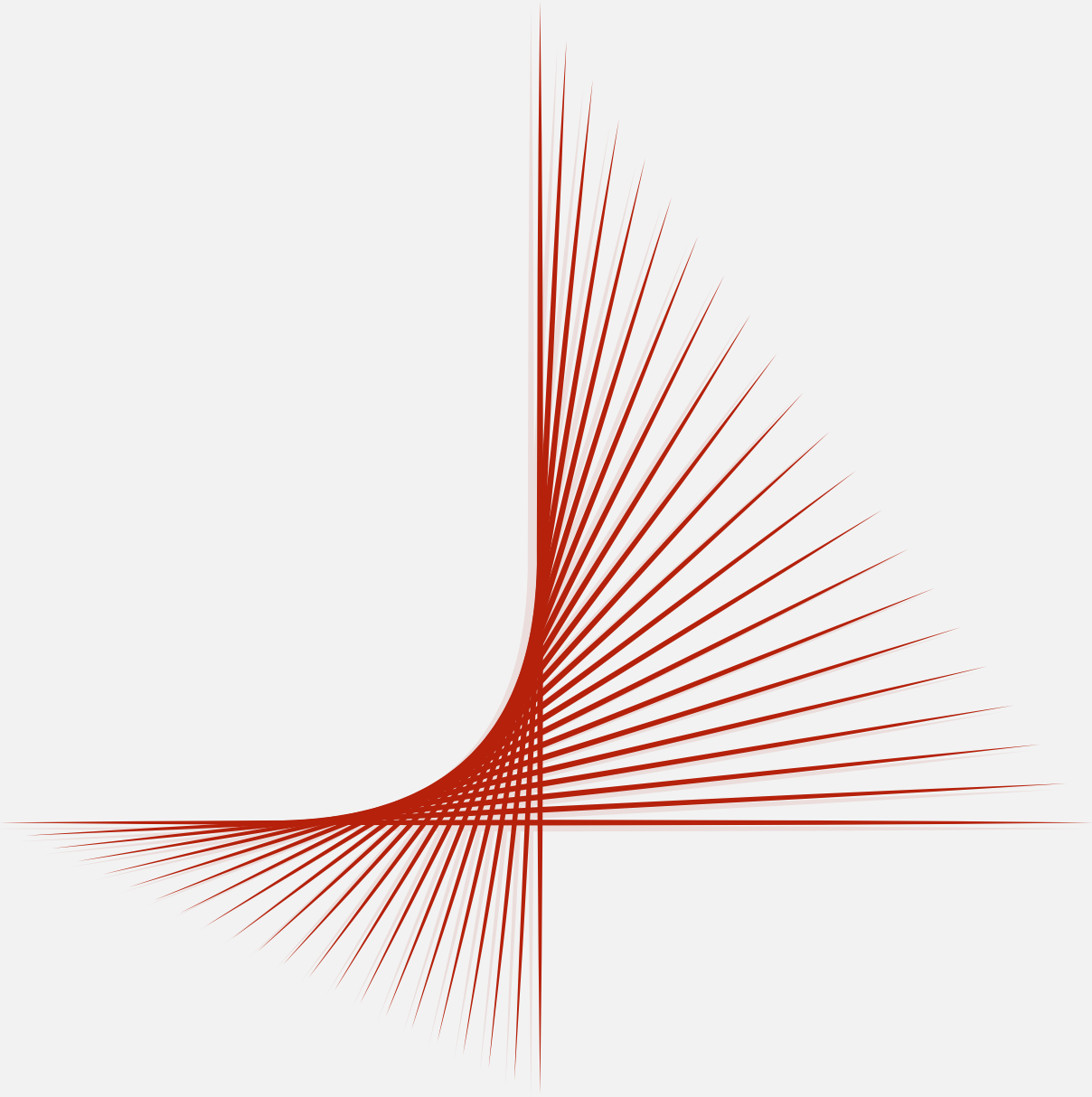
7 Krenak (2020, p. 15).

8 Hui, Y. (2016). The question concerning technology in China: An essay in cosmotechnics. Falmouth, UK: Urbanomic, p.7.

9 Read more at https://digitalrightswatch.org.au/2021/03/25/technosolutionism/.

10 See, for example the joint statement "Within Bounds: Limiting AI's environmental impact," signed for over 100 organizations in early 2025: https://greenscreen.network/en/blog/within-bounds-limiting-ai-environmental-impact/.

Regulating and Governing AI for the Public Interest



02.

AI REGULATION IN A *Post-Reality World*

Blair Attard-Frost

University of Toronto

Reality is being degenerated, divided into many incongruent realities, decoupled from ground truths. In their pursuit of profitable generative AI products, a small handful of big tech companies are eroding authenticity, trust, and equity across countless vectors of social and cultural life. They are eroding the quality and safety of information ecosystems. They are eroding the trustworthiness of collective memory and our capacity to collectively make sense of our world. Through extractive data scraping and model development practices – through rote theft – they are eroding the value of creative labor and the livelihoods of creative workers. They are eroding institutions. They are eroding trust and truth. They are eroding reality.

Government response to the slo-mo annihilation of reality has been meek. Around the world, new legislation and regulatory initiatives have emerged with an overwhelming emphasis on preventing tangible harms caused by AI systems.[1] Tangible harms are physical, psychological, economic, and environmental.[2] These harms are readily observable, quantifiable, contestable, and serviceable. Tangible harms can be serviced within neoliberal frameworks of justice that emphasize (1) individualistic accounts of harm and (2) processes of contestation and remediation that are rigidly formalized and procedural. Damages to bodily, material, cognitive, and emotional well-being are tangible harms caused by AI systems. Loss of finances, loss of resources, and denial of access to services and opportunities are tangible harms caused by AI systems. Damages to health and well-being incurred from soaring carbon emissions, degraded land, and depleted water supplies are tangible harms caused by AI systems.

Intangible harms – damages to social life and cultural production, to our capacities for shared sense-making and meaning-making, to the epistemic and ontological groundwork upon which we build all our institutions – are more pernicious. Intangible harms are collective harms that chip away at the stability and integrity of our shared values and worldviews. Intangible harms resist quantification, evade empirical observation, slide away from individualistic frameworks of justice that attempt to account for and remediate their damages. This does not mean that these harms do not exist, or that justice is not serviceable in response to

these intangible harms. However, we should not expect a just response to the AI-(de) generated erosion of reality to be found within large-scale governance systems. We should not expect the flurry of top-down legislative and regulatory initiatives that have recently been put forward by alliances of international, national, and industry leaders – in many cases, the very same leaders who accelerate this erosion under the auspices of "innovation" – to be effectively or justly enforced.

For those seeking to oppose the anti-reality machines and their makers, a viable path forward is to localize our regulatory thinking. AI regulation is not the preserve of a technocratic elite. Every day of our lives, we self-regulate the technologies used in our workplaces, in our professions, in our communities and our cities. We are regulatory experts within our own little spheres of our own post-reality worlds. We experience firsthand the stakes of our immediate material realities, and we can build the power needed to intervene in our own realities.

Through bottom-up regulatory action – through training and awareness-building and creating shared knowledge resources, through community and workplace guidelines for building and using AI, through collective bargaining and media engagement, open letters, petitions, and protests to raise awareness of harmful AI systems – we become AI regulators.[3] Recent regulatory initiatives in creative communities and workplaces provide a compelling vision of a path forward. The WGA and SAG-AFTRA labor strikes of 2023 offer a template for advancing self-determination in how AI is used in creative labor and for resisting harmful applications of generative AI in workplaces.[4] Online and off, communities of artists concerned about the intangible harms of generative AI have created open letters, knowledge resources, guidelines, and data masking tools, such as Glaze and Nightshade, to protect creative works against the industrial-scale theft being perpetrated by rogue AI developers.[5]

Against a backdrop of mounting authoritarianism and deregulatory sentiment in the halls of the U.S. government and abroad, we cannot rely solely on government regulators to protect us from tangible and intangible harms.[6] A more just and sustainable future for AI regulation can only be found by shifting our regulatory thinking toward small-scale, community-driven action. In a post-reality world – a world lacking reliable institutions – we must create our own regulatory institutions from the ground up.

ENDNOTES

1 Attard-Frost, B. & Lyons, K. (2024). AI governance systems: A multi-scale analysis framework, empirical findings, and future directions. AI and Ethics. https://link.springer.com/article/10.1007/s43681-024-00569-5, for an overview of AI governance practices and inventory of different types of harms addressed by Canada's AI governance systems.

2 Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N. . . . & Virk, G. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. AIES '23: AAAI/ACM Conference on AI, Ethics, and Society (723-741). https://dl.acm.org/doi/10.1145/3600211.3604673 / Attard-Frost, B. & Widder, D. G. (2024). The ethics of AI value chains. arXiv:2307.16787v3. https://arxiv.org/abs/2307.16787, for an overview of different types of harms caused by AI systems.

3 See Attard-Frost, B. (2023, December 16). AI countergovernance. Midnight Sun. https://www.midnightsunmag.ca/ai-countergovernance/, for an analysis of bottom-up AI regulation methods and governance practices

4 See WGA Contract 2023 (2023). Summary of the 2023 WGA MBA. https://www.wgacontract2023.org/the-campaign/summary-of-the-2023-wga-mba / SAG-AFTRA (2023). TV/theatrical: Regulating artificial intelligence. https://www.sagaftra.org/sites/default/files/sa_documents/AI%20TVTH.pdf, for a summary of AI regulations established by each union as a result of their 2023 strikes and collective bargaining

5 See Heikkilä, M. (2024, November 13). The AI lab waging a guerrilla war over exploitative AI. MIT Technology Review. https://www.technologyreview.com/2024/11/13/1106837/ai-data-posioning-nightshade-glaze-art-university-of-chicago-exploitation/, for a discussion of Glaze, Nightshade, and community organizing against generative AI developers by artists and researchers.

6 See Edwards, B. (2024, November 6). Trump plans to dismantle Biden AI safeguards after victory. Ars Technica. https://arstechnica.com/ai/2024/11/trump-victory-signals-major-shakeup-for-us-ai-regulations/, for an analysis of measures that are likely to be taken to deregulate the tech industry in a second Trump term

DC'S AI REGULATION *as a Federal Framework*

Anuti Shah

University of Virginia

AI has many use cases, but are they all beneficial to the greater society? Many may seem helpful in the short run but can ultimately do more harm than benefit in the long run. This sentiment has already been seen; with the current surge of AI and Machine Learning algorithms, numerous companies have felt compelled to implement AI in their business without thinking about whether it is truly beneficial for their company and, in turn, the greater good of society. In our roundtable discussion about "Reimagining AI for Environmental Justice" with the UVA Karsh Institute of Democracy, a topic of discussion that resonated with me was how AI can be regulated. Since AI implementation is growing so rapidly, the government has not been able to develop comprehensive federal regulation to ensure that developing AI is not harmful to society in the long run. Despite this, state-level legislation has begun to successfully accomplish this goal; DC is a prime example of this. DC's AI values and regulation provide a crucial framework that can and should help frame federal AI regulation to help protect society from the countless potential harms of AI.

DC's AI legislation ensures that AI solutions are only adopted if they align with certain societal values. DC Mayor Muriel Bowser defined six core AI values in an executive order signed in February 2024. These values clearly benefit people, safety and equity, accountability, transparency, sustainability, privacy and cybersecurity. Any agency planning to deploy an AI tool first must "verify whether such employment is in alignment with the [above] AI values, assess impact to these AI values, consider what controls might be used to mitigate negative impacts, and document its review".[1] These key principles help ensure that deployed AI aligns with the government's goals and benefit society in the long run. DC did not stop there, though; the AI industry is everchanging, so the mayor also established a public AI Advisory Group and internal AI Taskforce. The advisory group holds public listening sessions to gain input on proper AI use and reviews specific AI tools and advises the mayor on its alignment with the AI values. The task force, led by Chief Technology Officer Stephen Miller, helps facilitate the work of the advisory group; both teams work together to advance value-driven AI in the district.[2]

One example of the value of this legislation is called DC Compass. This software allows

residents to ask questions in their preferred language, providing faster, well-cited answers with responsive maps, dashboards, statistics, natural language summaries, related datasets, and relevant DC government initiatives. DC Compass prioritizes equity by addressing barriers like data literacy gaps, language accessibility, and the overwhelming volume of datasets, ensuring all residents can easily access and understand government data. The tool enhances accountability and transparency through clear citation of sources, integration of feedback mechanisms, and detailed public documentation of its AI functionalities. Safety and privacy are safeguarded by using pre-screened, anonymized datasets and implementing robust security measures. Lastly, it promotes sustainability by controlling costs, limiting environmental impacts through the efficient use of AI technologies, and enhancing digital literacy within the community.[3]

The federal government has yet to pass comprehensive AI legislation, leaving a significant gap in regulating the development and deployment of AI technologies across the nation. This gap has only deepened in 2025, when the Trump administration revoked key Biden-era AI policies that focused on safety, transparency, and public accountability in AI systems. By dismantling these safeguards—many of which were designed to mitigate algorithmic bias, ensure ethical use, and promote interagency coordination—the federal government has signaled a return to a deregulation-first approach that prioritizes rapid innovation over long-term societal impact. This policy backslide has raised alarm among experts, further exposing vulnerable communities to the harms of unregulated AI deployment. In the absence of federal leadership, state and local efforts have become even more critical. DC's AI framework offers a valuable blueprint for how federal policy could approach this challenge by prioritizing societal values. By borrowing elements from DC's AI model, like having an AI taskforce and working group for each state and adopting the same values federally, federal policymakers could create a balanced approach that fosters innovation while protecting society from AI's potential harms; actions. A federal framework grounded in these principles would ensure that AI benefits all citizens equitably and responsibly, bridging the regulatory gap that currently exists.

ENDNOTES

1 Government of the District of Columbia. (2024). DC's AI Values and Strategic Plan. DC.gov techplan.dc.gov/page/dcs-ai-values-and-strategic-plan.

2 Edinger, J. & Pattison-Gordon, J. (2024) Values Drive the Work of AI Groups in Washington, D.C. GovTech. www.govtech.com/artificial-intelligence/values-drive-the-work-of-ai-groups-in-washington-d-c.

3 Office of the Chief Technology Officer. (2024). DC Compass AI Values Alignment Report. https://techplan.dc.gov/sites/default/files/dc/sites/itstrategicplan/publication/attachments/Final_Draft_DCCompass_Alignment.pdf

RETHINKING AI POWER: *Elevating Communities Through Decentralized Policy*

Ahmed Alrawi.

Digital Technology for Democracy Lab, Karsh Institute of Democracy
University of Virginia

In the rapidly evolving field of artificial intelligence (AI), establishing policies that fully address the concerns and needs of communities is becoming increasingly critical. The Decentralized AI Policy Model offers a transformative approach to AI governance, presenting a more inclusive alternative to the traditional centralized framework, which typically relies on a top-down approach where policies are set solely by state authorities or policymakers. Instead, the decentralized model emphasizes civic engagement and collaboration with local communities, non-profit organizations, and other essential stakeholders. The Decentralized AI Policy Model is not intended to undermine the authority of the state; rather, it seeks to enhance the quality and relevance of established policies by incorporating the lived realities, concerns, and expectations of those most affected by these technologies.

This decentralized model operates on the principle that effective AI policy requires a broad spectrum of voices, including those historically excluded from centralized policy discussions. The perspectives of local communities and organizations bring crucial insights into the ethical and societal impacts of AI, especially regarding privacy, safety, and fairness. For instance, many community members may express valid anxieties about the potential misuse of AI in surveillance applications, fearing a chilling effect that could deter freedom of expression and civic participation. By engaging with these groups, policymakers gain deeper understanding of the public's sentiments, enabling them to create policies that directly address these concerns, fostering greater transparency, accountability, and social trust in AI applications.

The decentralized model also recognizes the unique contributions of various stakeholders, each offering valuable expertise and insights into the policy-making process. Unlike traditional multistakeholder models, which often involve consultations but ultimately leave decision-making power concentrated in central authorities, the decentralized approach redistributes authority by embedding decision-making mechanisms directly within local communities, non-profits, and civic groups. This model prioritizes bottom-up governance, ensuring that those affected by AI policies have a direct role in shaping them rather than serving as mere advisors to centralized institutions. Non-profit organizations focused on social justice, privacy advocacy, and human rights can provide critical perspectives that ensure AI policies remain ethically grounded. Meanwhile, local community members can contribute insights based on their direct experiences and unique needs, whether in the realm of public safety, education, or healthcare. This level of engagement ensures that AI policies are designed with specific safeguards that respect individual privacy and uphold democratic values. For instance, when AI is implemented for local security or surveillance, community input can guide its use to be consensual and respectful, balancing the benefits of AI-driven insights with robust privacy protections.

In addition to addressing specific community needs, the Decentralized AI Policy Model provides a solution to several problems associated with overly centralized approaches. Centralized policies often lack the flexibility to adapt to rapidly changing technologies and the evolving concerns of the communities they impact. Because AI technologies advance at an unprecedented pace, regulatory frameworks that rely solely on centralized decision-making often struggle to keep up, leading to outdated or ineffective policies that fail to address emerging ethical dilemmas, privacy concerns, and algorithmic biases. Certainly, through prioritizing open dialogue and collaboration, decentralized policymaking promotes a more resilient governance framework adaptable to technological advancements and responsive to the concerns of the community. By incorporating diverse local perspectives, the decentralized model ensures that AI governance remains dynamic, iterative, and capable of evolving alongside technological innovations, rather than being reactive or rigid. Such an approach is mainly important as it guards against the potential misuse of AI technologies, mitigating the risks of surveillance overreach and safeguarding against any threats to civil liberties that may arise from AI's rapid deployment.

In a nutshell, the Decentralized AI Policy Model advances a path toward AI governance that is not only participatory, but deeply democratic. It reimagines AI policymaking as a collective endeavor where insights from the public, non-profit organizations, and various stakeholders shape policies that align with shared values and address actual community needs. Through this approach, AI policies are created in a way that respects diverse voices, minimizes the risks of surveillance, and emphasizes ethical standards, resulting in AI technologies that serve society more equitably and transparently. Rather than isolating policy decisions among a few decision-makers, this model offers a collaborative framework that empowers communities, enhances social trust, and ultimately leads to more just and effective AI governance.

STUDENTS IN THE DRIVER'S SEAT: *Establishing Collaborative Cultures of Technology Governance at Universities*

Celia Calhoun
Ella Duus
Desiree Ho
Owen Kitzmann
Mona Sloane

Sloane Lab
University of Virginia

In the age of AI, technologies that touch upon, direct, or deeply affect students' lives are everywhere. Tools such as career development platforms, content delivery systems, assessment tools, and technology infrastructure have transformed education and redefined the student experience. Unsurprisingly, students worry about the mass collection of their data and the unseen impacts of technologies on campus.[1] Web proctoring services and AI "detectors" can lead to false accusations of cheating.[2] Education technology ("EdTech") companies maintain individual data profiles that are frequently sold and exist forever[3] while course selection and major advising tools, possibly limit student autonomy.[4] Students fear that campus technologies compromise privacy and perpetuate bias.

Against that backdrop, it seems obvious that students should participate in technology and data governance issues at universities. However, respective efforts have been unsuccessful or remained overlooked. In 2018, the University of California, Los Angeles's Board on Privacy and Data Protection, which convened students, faculty, and administrators to review technology related policies, create standards, and resolve issues, was quietly replaced by smaller committees

without students.[5] This revocation of student power came at a time where student leaders were particularly concerned with how new technologies would negatively impact students, e.g., facial recognition.[6] UCLA is representative of a trend among American universities at large, wherein committees and boards dedicated to technology policy-making exist, but exclude student participation.[7]

Universities jockey to be seen as bastions of democracy, often emphasizing meaningful student leadership to attract applicants.[8] However, student leadership is not a major component in university decision making. With dramatically rising tuition costs, remaining competitive requires universities to prove that a college degree is worth the investment.[9] The number of university administrators bloats as universities offer more novel services while students' own change-making ability diminishes.[10] Students no longer have the power to pressure the administration for policy changes in important political and social issues like they did in the years following WWII .[11]

The recent development of AI policies at universities exemplifies the exclusion of students.[12] Student engagement on AI issues is often limited to designated areas (e.g., plagiarism or classroom uses) and not broader concerns like procurement and privacy. The University of Virginia's own AI policy creation in 2023 only involved student representatives in discussions around classroom-specific policies– not the broader impacts of AI on students or any other group.[13] A similar picture emerges at other US institutions: Yale, Boston University, University of Missouri, and many other higher education institutions convened task forces on AI in teaching with zero student representation.[14]

When it comes to the governance of their data, students have even less of a voice. There is limited legal recourse when their data rights are violated. The Family Education Rights and Protection Act (FERPA), designed to give students the right to access, amend, and dispute uses of their data, has become exceedingly difficult to enforce because student data is scattered among third-parties.[15] State-level student privacy laws vary widely, lagging behind the fast paced technology sector which oppose and lobby against expanding privacy laws.[16]

Due to the unstable footing students stand on regarding their data rights, establishing a culture of student involvement in technology governance is imperative. Even though they often lack involvement in internal decision-making at universities writ large,[17] students have clear ideas about how they want to participate in governance, especially when it comes to technology. For example, they repeatedly call for privacy, data, and technology boards where they can shape the strategic direction of their institution's technology use.[18] Students should occupy a meaningful seat at the table where procurement, vetting, certification, review, and accountability decisions take place.[19] Universities must diverge from their current path and allow students to have real impact in policy making and their technological and social futures.

ENDNOTES

1 Park, J., & Vance, A. (2021, October 25). Higher Education Voices: College Students' Attitudes Toward Data Privacy - Student Privacy Compass. Student Privacy Compass. https://studentprivacycompass.org/resource/higheredvoices2021/ / Jones, K. M. L., Asher, A., Goben, A., Perry, M. R., Salo, D., Briney, K. A., & Robertshaw, M. B. (2020). "We're being tracked at all times": Student perspectives of their privacy in relation to learning analytics in higher education. Journal of the Association for Information Science and Technology, 71(9), 1044–1059. https://doi.org/10.1002/asi.24358

2 Singer, N., & Krolik, A. (2021). Online cheating charges upend Dartmouth Medical School. The New York Times. https://www.nytimes.com/2021/05/09/technology/dartmouth-geisel-medical-cheating.html / Hill, K. (2022). Accused of Cheating by an Algorithm, and a Professor She Had Never Met. The New York Times. https://www.nytimes.com/2022/05/27/technology/college-students-cheating-software-honorlock.html / Brown, L. X. Z. (2020). How Automated Test Proctoring Software Discriminates Against Disabled Students - Center for Democracy and Technology. Center for Democracy and Technology. https://cdt.org/insights/how-automated-test-proctoring-software-discriminates-against-disabled-students/ / Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S. et al. Testing of detection tools for AI-generated text. Int J Educ Integr 19, 26 (2023). https://doi.org/10.1007/s40979-023-00146-z / Elkhatat, A.M., Elsaid, K. & Almeer, S. (2023). Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text. Int J Educ Integr 19, 17 https://doi.org/10.1007/s40979-023-00140-5 / Davalos, J., & Yin, L. (2024). AI Detectors Falsely Accuse Students of Cheating—With Big Consequences. Bloomberg. https://www.bloomberg.com/news/features/2024-10-18/do-ai-detectors-work-students-face-false-cheating-accusations / Ananny, M. (2022). Seeing Like an Algorithmic Error: What are Algorithmic Mistakes, Why Do They Matter, How Might They Be Public Problems? The Digital Public Sphere, A Healthy Digital Public Sphere. https://law.yale.edu/sites/default/files/area/center/isp/documents/ananny.pdf / Feathers, T. (2022). This Private Equity Firm Is Amassing Companies That Collect Data on America's Children – The Markup. The Markup. https://themarkup.org/machine-learning/2022/01/11/this-private-equity-firm-is-amassing-companies-that-collect-data-on-americas-children / EdTech Exposed. (n.d.)

3 Feathers, T. (2022). This Private Equity Firm Is Amassing Companies That Collect Data on America's Children – The Markup. The Markup. https://themarkup.org/machine-learning/2022/01/11/this-private-equity-firm-is-amassing-companies-that-collect-data-on-americas-children / EdTech Exposed. (n.d.). EdTech Law Center. https://edtech.law/wp-content/uploads/2024/06/edtech-exposed.pdf / Kelly, G., Graham, J., & Garton, S. (2023, June 27). 2023 State of Kids' Privacy. Common Sense Media. https://www.commonsensemedia.org/research/2023-state-of-kids-privacy

4 Richards, N. (2017). Intellectual Privacy: Rethinking Civil Liberties in the Digital Age. Oxford University Press, USA. / Johnson, J. A. (2017). Ethics and Justice in Learning Analytics. New Directions for Higher Education, 2017(179), 77–87. https://doi.org/10.1002/he.20245 / Jones, K.M.L., Goben, A., Perry, M.R., Regalado, M., Salo, D., Asher, A.D., ... Briney, K.A. (2023). Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios. portal: Libraries and the Academy 23(3), 485-515. https://dx.doi.org/10.1353/pla.2023.a901565. / Greenhalgh, S.P., DiGiacomo, D.K. and Barriage, S. (2023), "Platforms, perceptions, and privacy: ethical implications of student conflation of educational technologies", Information and Learning Sciences, Vol. 124 No. 9/10, pp. 247-265. https://doi.org/10.1108/ILS-03-2023-0030

5 Fredberg, S. (2020, February 13). Privacy and data protection board disbands, prompting concerns on policy revisions. Daily Bruin. https://dailybruin.com/2020/02/13/privacy-and-data-protection-board-disbands-prompting-concerns-on-policy-revisions

6 Burke, L. (2020). UCLA drops plan to use facial recognition security surveillance, but other colleges may be using technology. Inside Higher Ed | Higher Education News, Events and Jobs. https://www.insidehighered.com/news/2020/02/21/ucla-drops-plan-use-facial-recognition-security-surveillance-other-colleges-may-be

7 Caron, A. (2021). EDUCAUSE QuickPoll Results: Positioning Higher Education IT Governance as a Strategic Function. EDUCAUSE Review. https://er.educause.edu/articles/2024/2/educause-quickpoll-results-positioning-higher-education-it-governance-as-a-strategic-function / Neale, M., & Tryniecki, M. (2020, August 10). The Post-Pandemic Evolution of Student Data Privacy. EDUCAUSE Review. https://er.educause.edu/articles/2020/8/the-post-pandemic-evolution-of-student-data-privacy

8 Daniels, R.J., Shreve, G., & Spector, P. (2021). What Universities Owe Democracy. Baltimore: Johns Hopkins University Press. https://dx.doi.org/10.1353/book.97330. / Koppell, J. (2024, July 8). The Role of Universities in Shaping Democratic Values. The Chronicle of Higher Education. https://sponsored.chronicle.com/the-role-of-universities-in-shaping-democratic-values/index.html / Gallagher, T. (2021, Spring). The Democratic Imperative for Higher Education. Liberal Education. https://www.aacu.org/liberaleducation/articles/the-democratic-imperative-for-higher-education

9 Martin, E. A. (2024). College Rankings: What Type of Students Use Them and Who Benefits (Order No. 31631789). https://escholarship.org/uc/item/78f67056 / Jin, G. Z., & Whalley, A. (2007). The power of information: How do US news rankings affect the financial resources of public colleges? https://citeseerx.ist.psu.edu/documentrepid=rep1&type=pdf&doi=4241baef57794138a367aad6be6a1243d121ff2e / Kuh, G. D., & Pascarella, E. T. (2004). What Does Institutional Selectivity Tell Us About Educational Quality? Change: The Magazine of Higher Learning, 36(5), 52–59. https://doi.org/10.1080/00091380409604986

10 Jones, K. M. L., & Hinchliffe, L. J. (forthcoming). Preparing academic librarians to prioritize privacy in learning analytics projects: An evaluation of a professional development course. College &Research Libraries. https://osf.io/s5hwr / Bozeman, B., Youtie, J., & Jung, J. (2020). Robotic Bureaucracy and Administrative Burden: What Are the Effects of Universities' Computer Automated Research Grants Management Systems? Research Policy, 49(6), 103980. https://doi.org/10.1016/j.respol.2020.103980 / Schofer, E., Ramirez, F. O., & Meyer, J. W. (2020). The Societal Consequences of Higher Education. Sociology of Education. https://doi.org/10.1177/0038040720942912

11 Miles, J., Miller, M. T., & Nadler, D. P. (2020). Student Governance: Toward Effectiveness and the Ideal. Retrieved November 26, 2024, from College student journal website: https://www.semanticscholar.org/paper/Student-Governance%3A-Toward-Effectiveness-and-the-Miles-Miller/eeee5fafc2703e532df44cbb2de86c7e2d1c4a46 / Cintron, A. (2018). History. Retrieved from UVA Student Council website: https://www.uvastudco.com/history / Winant, G. (2024, April 24). Why Students Must Shout to Be Heard. The Chronicle of Higher Education. https://www.chronicle.com/article/why-students-must-shout-to-be-heard / Whittington, K. E. (2024). A Harvard Dean's Assault on Faculty Speech. The Chronicle of Higher Education. https://www.chronicle.com/article/a-harvard-deans-frontal-on-faculty-speech / Joselit, D. (2024). The Transactional University and Student Protest. October, 188, 175–177. https://doi.org/10.1162/octo_a_00523 / Cho, K. S. (2018). The Perception of Progress: Conceptualizing Institutional Response to Student Protests and Activism. In University of California, Los Angeles & Higher Education Research Institute, THOUGHT & ACTION [Journal-article]. https://www.researchgate.net/publication/336591731_The_Perception_of_Progress_Conceptualizing_Institutional_Response_to_Student_Protests_and_Activism / McKown, H. C. (1944). The Student Council. United States: McGraw-Hill. https://www.constitution.org/1-Activism/sg/mckown.htm / Haskins, C. H. (1923). The Rise of Universities. https://www.gutenberg.org/files/63574/63574-h/63574-h.html

12 Neale, M., & Tryniecki, M. (2020, August 10). The Post-Pandemic Evolution of Student Data Privacy. EDUCAUSE Review. https://er.educause.edu/articles/2020/8/the-post-pandemic-evolution-of-student-data-privacy

13 Heny, N., & Pennock, A. (2023). Generative AI has Arrived: What Does this Mean for Teaching and Learning at UVA? | Office of the Executive Vice President and Provost. Retrieved from Virginia.edu website: https://provost.virginia.edu/subsite/genai/task-force-report

14 Report of the Yale Task Force on Artificial Intelligence. (2024). AI At Yale. https://ai.yale.edu/report-of-the-yale-task-force-on-artificial-intelligence / Boston University AI Task Force | Office of the Provost. (2023). BU.edu. https://www.bu.edu/provost/about/committees/boston-university-ai-task-force/ / Mondelli, T., Schatz, E., & Trachtenberg, B. (2024). Task Force on Artificial Intelligence in the Learning Environment Report. In provost.missouri.edu. University of Missouri. https://provost.missouri.edu/wp-content/uploads/2024/08/TF-Report-AI-in-the-Learning-Environment-1.pdf

15 Center for Democracy & Technology. (2021). Commercial Companies and FERPA's School Official Exception: A Survey of Privacy Policies. In Center for Democracy & Technology. https://cdt.org/wp-content/uploads/2021/01/UPDATED-2021-01-21-Student-Privacy-2pg-Commercial-Edtech-Final.pdf / Jones, K.M.L., Goben, A., Perry, M.R., Regalado, M., Salo, D., Asher, A.D., ... Briney, K.A. (2023). Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios. portal: Libraries and the Academy 23(3), 485-515. https://dx.doi.org/10.1353/pla.2023.a901565. / Zeide, E. (2016). Student Privacy Principles for the Age of Big Data: Moving Beyond FERPA and FIPPs. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2821837 / Reisman, M. (2021, season-04). University Use of Big Data Surveillance and Student Privacy. Scholarship Repository. https://ir.law.fsu.edu/lr/vol48/iss2/6/

16 Student Privacy Compass. (n.d.). State Laws and Legislation - Student Privacy Compass. https://studentprivacycompass.org/state-laws-and-legislation/ / Fitzgerald, C., Williams, K., & Cross, R. J. (2024). The State of Privacy: How state "privacy" laws fail to protect privacy and what they can do better. Electronic Privacy Information Center. https://pirg.org/edfund/wp-content/uploads/2024/01/State-of-Privacy-Feb.-2024.pdf

17 Caron, A. (2021). EDUCAUSE QuickPoll Results: Positioning Higher Education IT Governance as a Strategic Function. EDUCAUSE Review. https://er.educause.edu/articles/2024/2/educause-quickpoll-results-positioning-higher-education-it-governance-as-a-strategic-function / Quaye, S. J., Linder, C., Stewart, T. J., & Satterwhite, E. M. (2021). A Review of Scholarship on College Student Activism from 2000 to 2020. Higher Education, 1–51. https://doi.org/10.1007/978-3-030-66959-1_5-1 / Patrick, J. (2022). Student Leadership and Student Government. Research in Educational Administration & Leadership, 7(1), 1–37. Retrieved from https://eric.ed.gov/?id=EJ1344509

18 Jones, K.M.L., Goben, A., Perry, M.R., Regalado, M., Salo, D., Asher, A.D., ... Briney, K.A. (2023). Transparency and Consent: Student Perspectives on Educational Data Analytics Scenarios. portal: Libraries and the Academy 23(3), 485-515. https://dx.doi.org/10.1353/pla.2023.a901565.

19 Sloane, M. (2023). Biden's AI executive order underlines need for student technology councils. Retrieved November 26, 2024, from Times Higher Education (THE) website: https://www.timeshighereducation.com/blog/bidens-ai-executive-order-underlines-need-student-technology-councils

SETTING UP HUMAN-AI *Teams in the Public Interest*

Shalini Misra

School of Public and International Affairs
Virginia Tech

How should AI systems be integrated into public sector settings for outcomes that are in the public interest? Public managers are told to adopt AI in their organizations, but are not always aware of whether AI is appropriate for a particular task or collaborative environment. Diffusion of AI systems in the public sector remains low, despite surging interest in adopting AI to improve public managerial decision-making. Prominent among the risks and challenges of AI adoption in the public sector is the need to uphold public sector values of transparency, democratic accountability, privacy, legitimacy, fairness, and equity. However, little attention is given to the cognitive and motivational factors that influence public managers to adopt AI.

We surveyed US-based emergency managers to understand their attitudes toward AI and their intentions to rely on AI in a set of decision-making scenarios relevant to crisis management. Emergency managers play an important role in society before, during, and after disasters. They work at all levels of the government, in non-profits, and the private sector.[1] While emergency managers had less positive attitudes toward AI and were less likely to rely on AI for decision-making, it wasn't because of wariness toward AI or lack of trust in AI. We found that public managers' humanistic and organizational needs are at least as important as technology design considerations for AI implementation in the public sector. We distill our findings into six insights for designing and implementing Human-AI teams in a way that aligns with public managers' cognitive capacities, responsibility to the public good, and organizational set up.

There's little trust in AI without transparency. If public managers are going to be asked to rely on AI for decision-making (sometimes overriding their intuition, experience, and expertise), they need to know and understand what factors the system used to determine the result. Managers need to be able to trace their decisions through a process that would satisfy their standards for rigor and transparency.

AI will place cognitive and administrative demands on public managers. AI is different from other types of technologies because of the need for intra and inter-organizational coordination, data infrastructure, organizational resources, expertise, operational capacity, and significant changes in organizational processes. Most AI systems are not designed for the public sector. The adoption of AI in the public sector organizations will need the establishment of the data infrastructure, training in the use of new AI systems, testing and evaluation protocols, and building in additional time and resources for decision-makers to verify AI. AI may unnecessarily replace current processes, tools, and technologies that work well without AI.

Inefficiencies and redundancies have value, especially when new technologies are adopted. Public managers are not comfortable integrating AI into their workflows without thorough vetting and evaluation. Any first run of any technology should be scrutinized and monitored with built-in redundancies. Even though checking the outputs of AI systems are likely to increase administrative burden, they are necessary to ascertain accuracy, consistency, and fairness of results.

Public managerial expertise and experience are undervalued in the discourse on Human-AI teaming. Discourses of human-AI collaboration often emphasize the potential value AI could bring to the table, such as speed, efficiency, pattern recognition, consistency, and accuracy for certain types of tasks. The skills, talents, and capacities humans bring to the table are given short shrift. The public managers we talked to emphasized the importance they placed on human input, their own extensive real-world experience, place-based knowledge and knowledge of their communities, and empathy in public managerial decision-making. Managers are more concerned about improving their own skills and those of their team members in AI environments, rather than concerns about narrow notions of efficiency or productivity.

Oversight and control over decisions are paramount. Among the organizational processes and work design conditions that managers said need to be place for public interest-centered AI integration are: (a) ground rules and shared understanding of how AI results should be interpreted; (b) systematic processes of experimentation and evaluation; and (c) organizational processes that enable managers to validate their analytical process, allow corrections, and review decision points.

Not all public managerial tasks are AI-appropriate. Managers distinguish between tasks that may be AI-appropriate under certain conditions and tasks that are inappropriate for AI. For example, some managers may be comfortable with AI assistance in crafting emergency preparedness messages, but not sending out the messages automatically and certainly not sending messages during an emergency. Many others noted the need for multilingual communication in the communities they serve as well as contextual knowledge about the community for emergency preparedness and crisis messaging.

Administrators and decision-makers who are thinking of implementing AI should rethink their program and policy design in light of these findings. In particular, they should

view adoption and implementation not just as a single decision but as a phased process that requires consultation at key points. Building in space, time, and resources for experimentation, evaluation, training, and collaborative deliberation routines is an important element of public interest-centered AI systems integration.

ENDNOTES

1 Misra, S., Katz, B., Roberts, P., Carney, M., & Valdivia, I. (2024). Toward a person-environment fit framework for artificial intelligence implementation in the public sector. Government Information Quarterly, 41(3), 101962. Available at: https://doi.org/10.1016/j.giq.2024.101962

ETHICS OF *AI Refusal*

Yasmin Curzi

Digital Technology for Democracy Lab, Karsh Institute of Democracy
University of Virginia

The conversation surrounding artificial intelligence (AI) is often constituted by eternal promises of innovation, efficiency, and progress. The lack of a reflective and critical approach to the technology's deployment and impact – especially when viewed through the lenses of data ethics, environmental justice, and digital sovereignty - nevertheless, remains overlooked.

In this essay, I propose an ethics of AI refusal to counteract the prevailing assumptions about AI's inevitability and its assumed benefits. Refusal, in this sense, is not simply about rejecting technology outright; rather, it is about questioning the frameworks within which AI is developed and operates, resisting the ways it consolidates power, and reframing what progress and creativity should mean in this context.

Hyping

Corporations are constantly evoking a narrative of "technological inevitability".[1] As they suggest, AI, much like the industrial revolutions before it, is an unstoppable force that will reshape the world regardless of human intervention. Hyping AI serves to justify its rapid – and unchecked –development, with the promise of greater efficiency, economic growth, and creativity.

Yet, this narrative overlooks critical considerations, such as the systems' actual efficiency and users' needs. From algorithmic biases in facial recognition[2] or nudity detection[3] to disastrous lapses in automated decision-making,[4] the failures of AI are many.[5] This is where the ethics of AI refusal could come into play. Instead of embracing the rhetoric of technological inevitability, refusal aims at challenging the very premise that AI must expand and continue to be applied to every instance of our lives. It insists that we ask whether this "progress" perpetuates deep inequalities with no greater benefit for society.

Expropriating

Another key issue in the ethics of AI refusal is the expropriation of human creativity – translated into the quantifiable word "data". AI, in its current form, is heavily reliant on vast datasets that include human-produced content – images, text, music, interactions

and more. These datasets, often scraped without explicit consent, raise concerns about intellectual property, authorship, and the devaluation of human labor. The AI models trained on this data can then produce "creative" outputs that mirror human styles and approaches, but without recognizing the cultural and intellectual labor that went into their creation and without offering compensation or recognition.

The ethics of refusal, in this case, involves rejecting the idea that AI can or should replace human creativity. Instead, refusal calls for a reevaluation of what constitutes creativity in the age of AI. It encourages a more democratic and equitable approach to creative industries, one that safeguards the rights of creators and ensures that AI tools can be used to enhance, rather than replace, human expression. Furthermore, refusal in this domain asks how we can reimagine AI to work in collaboration with human creativity, rather than subsuming it.

Burning

The most relevant site for the ethics of AI refusal relies on its impact for environment. A recent International Energy Agency (IEA)[6] report highlights that data center electricity consumption is projected to double by 2026, driven largely by the rise of power-intensive workloads, including AI and cryptocurrency mining. Such data centers are often located in regions where energy production is heavily reliant on fossil fuels, which severely exacerbates environmental degradation. In addition, the growing demand for rare earth metals required to power AI hardware contributes to socioecological destruction, with mining in the Global South, an increasing carbon footprint,[7] and exploitative labor practices, including forced and child labor.[8]

Techno-chauvinists promote the rhetoric of "technological progress" to solve energy consumption, including tech giants turning to nuclear power plants acquisitions.[9] However, this movement disregards the significant environmental and social risks associated with nuclear energy, as well as its limited capacity to meet the insatiable power needs of the promised "AI-driven future".

Instead of hoping that nuclear energy can solve the energy demands of AI, we need to advocate for AI degrowth – a conscious effort to limit the unchecked expansion of AI at any costs.

Refusing

What I propose as an ethics of AI refusal is the idea of adopting critical engagements with these promises of AI – to continuously examine the assumptions that AI is inherently beneficial and inevitable. It is crucial to ask who benefits from its development, who bears the costs, and how its deployment aligns with the values of environmental justice, cultural sovereignty, and human dignity.

By refusing to accept AI as a panacea for societal problems, we can create space for alternative visions of progress – ones that prioritize collective well-being, ecological sustainability, and cultural creativity over profit and technological determinism. We must rethink the tradeoffs in AI development. Society wants development, but not at the cost of burning the planet in exchange for chatbots that suggest adding glue to our pizzas.[10]

ENDNOTES

1 For example, in 2021, Mark Zuckerberg announced that Meta (formerly Facebook) would invest $10 billion in developing the Metaverse, with plans to continue investing heavily. The goal was to create a fully immersive, AI-driven virtual world where users could interact and work in digital spaces. However, despite $36 billion spent by 2024, user engagement is still falling short of expectations.

2 Hao, K. (2019, January 21). AI is sending people to jail – and getting it wrong: Using historical data to train risk assessment tools could mean that machines are copying the mistakes of the past. MIT Technology Review. Retrieved from: https://www.technologyreview.com/2019/01/21/137783/algorithms-criminal-justice-ai/.

3 Mauro, G., & Schellmann, H. (2023, December 14). 'There is no standard': Investigation finds AI algorithms objectify women's bodies. The Guardian. Retrieved from: https://www.theguardian.com/technology/2023/feb/08/biased-ai-algorithms-racy-women-bodies.

4 Al-Sibai, N. (2024, December 5). Murdered insurance CEO had deployed an AI to automatically deny benefits for sick people. Futurism. Retrieved from: https://www.yahoo.com/news/murdered-insurance-ceo-had-deployed-175638581.html.

5 DiBenedetto, C. (2024, January 4). ChatGPT fails at diagnosing child medical cases. It's wrong 83 percent of the time. A new study suggests ChatGPT needs to go back to medical school. Available at: https://mashable.com/article/chatgpt-fails-to-diagnose-child-medical-cases.

6 Gooding, M. (2024, January 26). Global data center electricity use to double by 2026 - IEA report: AI and cryptocurrency workloads are driving up demand. DCD. https://www.datacenterdynamics.com/news/global-data-center-electricity-use-double-2026-iea-report/.

7 O'Donnell, J. (2024, December 13). AI's emissions are about to skyrocket even further. MIT Technology Review. https://www.technologyreview.com/2024/12/13/1108719/ais-emissions-are-about-to-skyrocket-even-further/

8 Rédaction Africanews. (2024, March 6). US tech giants absolved in cobalt child labour case. Africanews. https://www.africanews.com/2024/03/06/us-tech-giants-absolved-in-cobalt-child-labour-case/.

9 Mazhar, M. (2024, October 29). Microsoft, Google and Amazon turn to nuclear energy to fuel the AI boom: Companies investing in small modular reactors that proponents call a safer approach to nuclear energy. CBC Radio. https://www.cbc.ca/radio/.

10 Robison, K. (2024, May 23). Google promised a better search experience — now it's telling us to put glue on our pizza: Google's AI Overviews feature is delivering incorrect answers faster than ever. The Verge. https://www.theverge.com/2024/5/23/24162896/google-ai-overview-hallucinations-glue-in-pizza.

FROM CODE TO CURRICULUM:

Addressing the Hidden Costs of Generative AI in Higher Education

Bryn Seabrook

Department of Engineering and Society
University of Virginia

Generative AI's ability to produce realistic content raises questions about authenticity and the implications for trust in student work. Before examining the possibilities of generative AI use in the classroom, it is important to establish an understanding of ethical considerations of generative AI. One area that is underexplored about generative AI systems are those "hidden" costs that impact the environment in ways that everyday users might be too separated from to recognize the greater impact of complex sociotechnical systems.

Generative AI technologies present a difficult challenge in how to balance technological innovation with environmental responsibility. As consumers and passive bystanders of these products, users should be aware of their environmental impact. Generative AI demands high computational requirements, resulting in significant energy consumption, often sourced from non-renewable energy. The training process for generative AI can produce a considerable carbon footprint. Research indicates that training large models can emit as much carbon as several cars over their lifetimes. Beyond training, ongoing usage of AI tools contributes to carbon emissions through the continuous operation of data centers.

In addition to the energy demands of these systems, there is also the material demand for more sophisticated hardware. The consumer need for more powerful hardware to run AI applications can lead to increased electronic waste. Short product lifecycles may lead to quicker obsolescence of devices, exacerbating the electronic waste problem if proper

recycling practices are not followed. One important resource that is often overlooked in this discussion is water usage. Data centers require significant amounts of water for cooling purposes, which further exacerbates water scarcity issues.

Developing AI responsibly involves integrating sustainable practices in the design, training, and deployment phases, requiring a commitment from both developers and consumers. Do the benefits of AI tools outweigh their environmental costs? In the context of higher education, faculty and students may not be fully aware of the environmental implications of using generative AI tools. Colleges and universities have an opportunity to educate students about the environmental impacts of technology, which could help foster a culture of sustainability within academia. On a smaller scale, institutions can invest in energy-efficient hardware, utilize renewable energy sources, and promote practices that reduce overall energy consumption. One potential mitigation strategy is to implement carbon offset programs, but the success of this kind of strategy could fail if there are not enough institutions and corporations that promote such practices.

The Role of Faculty in Addressing AI Ethics

What can faculty do to address AI ethics in undergraduate classrooms? Creating intentional classroom space for ethical reflection of technology is not feasible in every kind of undergraduate course. The most important strategy that is easily implemented is for individual faculty to understand how generative AI fits in the context of their courses. Faculty do not have to entirely redesign their courses to account for generative AI tools. However, faculty are responsible for knowing how to ethically implement their individual course AI policy and/or the university policy.

For example, the policy that I adopted in my engineering ethics courses reads as follows:

Writer's AI Contract: AI Policy

One of the lessons that you will learn over the course of this semester is how to be intentional with the use of technology. To that end, one technology we have to consider is the use of AI and writing. The use of generative AI tools is generally permissible in this class, but you are required to indicate any use of such tools for work submitted. In addition to citing the specific tool used, please add an appendix that includes the prompts you used to get the desired results. Be aware that generative AI tools can produce incorrect or biased outputs. It is your responsibility to ensure that any content you use from AI tools is correct, unbiased, and not harmful.

There is a fair amount of writing in my undergraduate classes, both formal and informal writing styles. I wanted to be very clear that I would not prevent them from using tools that are readily available to them, but I also expect that they will not rely on these tools to replace their work and that they are responsible for the accuracy of their work. This kind of policy is not exactly replicable in a thermodynamics course, for example. The point here is that individual faculty do not simply create an AI policy as a blanket

statement of acceptance or rejection, but that faculty are clear about how generative AI tools can succeed and fail in the context of the learning objectives of their courses.[1]

ENDNOTES

1 O. Gambelin. (2024). Responsible AI: Implement an Ethical Approach in Your Organization. New York, New York: Kogan Page, 2024.

VERIFIED HUMAN?
Identity Inversions in Our New Machine Age

Aaron Martin

Department of Media Studies and School of Data Science
University of Virginia

Keren Weitzberg

School of Politics and International Relations, Institute for the Humanities and Social Sciences
Queen Mary University of London

In a world overrun by bots and AI agents, afflicted by automated disinformation, fraud and scams, and struggling to cope with an onslaught of machine-generated "slop," many worry about how we can ensure meaningful human exchange and prosperity in the future. Some have called for restrictions, a pause, or even a moratorium on AI, but for others, these technology-exacerbated problems necessitate a technology-enabled solution, namely biometrics: the automated measurement and recognition of our physical characteristics or behaviors. Specifically, it is argued that to ensure a trustworthy digital economy in which AI is ubiquitous, more and more of our interactions and transactions will soon necessitate strong forms of human identification and authentication based on biometrics. Probably one of the most extreme (and polarizing) advocates of such a view is Worldcoin (now rebranded as World), a project from the company Tools for Humanity (TfH). Co-founded by OpenAI's Sam Altman, TfH's World has biometrically registered millions of people across countries like Indonesia, Chile, and Kenya and is aggressively trying to expand its operations globally despite sustained regulatory pushback largely on privacy and data protection grounds.

Critically, through developments like these, we are witnessing an important shift in the purported objectives of the technologies of biometrics. Once intended as a technical means to assign "uniqueness" to people by distinguishing one person from another, biometrics are now being resignified as technologies for assuring humanness—distinguishing us from bots. TfH's World project, for example, anticipates a future in which humans are indistinguishable

from AI absent frequent biometric checks, and ordinary people have been rendered unemployed by computers, making it necessary to distribute a biometrically controlled universal basic income. How worried should we be about such a prospective future?

For sure, the signs of AI's degenerative effects on our societies are all around us: bots are everywhere and spreading—they are scraping the web, providing customer "service," polluting social media, and so on. Across a growing number of different sectors, automated agents[1] are being deployed to augment or in some cases replace the work of humans. Many fret about what this invasion means for online discourse and digital interactions, as well as for the quality and sustainability of our societies and politics.

In certain jurisdictions, regulators are beginning to intervene. Some are imposing transparency requirements on bot operators to make it clearer when we are interacting with a machine. However, this requires the goodwill of whomever is deploying bots. In other cases, regulators are forcing platforms to more proactively detect and remove "inauthentic" activity. The EU's Digital Services Act, for example, includes strong requirements to prevent intentional manipulation by bots. But these are only partial measures—the problem of AI and identity assurance is much bigger than a platform regulation issue. It cuts across a wide range of domains. For example, the humanitarian sector (where our research[2] is largely focused) is reflecting on the implications of AI on the potential risks of beneficiary "fraud". In a sector with notoriously weak identity management, the increasing digitization of aid, most notably humanitarian cash assistance, could be severely challenged by the misuse of AI to create false identities. Is more extensive use of biometrics the solution?

For well over a century, biometric technologies have been aimed at eliminating "fraud" through authentication, verification, and deduplication (i.e. finding people who are registered in a system or database using multiple different identities and deactivating duplicative data), using supposedly unique bodily characteristics, such as fingerprints or iris scans, to detect fraudsters. But the emergence of "synthetic identities" (which combine real and fake information to create a new identity that does not correspond to any real person) is pushing the boundaries of these technologies and imagining their use to new ends.

Companies like TfH's World claim to have the solution to these problems, i.e. a biometrically-enabled "proof of personhood."[3] Their sales pitch suggests that we can reclaim our digital sovereignty—and even our humanity—by relinquishing our data to them. In this radically libertarian and dystopian scenario in which a private company, not the state, is designated to provide a critical infrastructure, biometrics assume a new ontology (from individuating humans to distinguishing machines). Such a scenario invites the use of a surveillance technology that will no doubt encroach on different domains of social, economic, and political life. It also reduces humanness to a technical protocol for proving personhood—something that we ought to resist, no matter who is involved.

ENDNOTES

1 Crawford, K. (2025). AI agents will be manipulation engines. Wired. Retrieved 25 April 2025, from https://www.wired.com/story/ai-agents-personal-assistants-manipulation-engines/

2 Weitzberg, K., Cheesman, M., Martin, A., & Schoemaker, E. (2021). Between surveillance and recognition: Rethinking digital identity in aid. Big Data & Society, 8(1). https://doi.org/10.1177/20539517211006744

3 World (2024). Proof of personhood: What it is and why it's needed. Retrieved 25 April 2025, from https://world.org/blog/world/proof-of-personhood-what-it-is-why-its-needed

COULD ANTI-AI BE THE *Newest Pillar of ESG?*

Christine Mahoney

Batten School of Leadership & Public Policy
University of Virginia

"Man is infinitely perfectible," Rousseau said. It is not in the attainment of perfection but the purposeful striving toward perfection that we find fulfillment and a satisfying life. Humans have been advancing our techniques of artistic expression for millennia in an effort to better capture our mind's eye and the conjuring of our imaginations. The Dutch masters of the 17th century perfected realistic still lives of fruit, insects and dew as well as the golden light glinting off clouds at sunset, but their skill seemed to be made obsolete with the invention of the camera in the 1800s. New techniques in film developing made it possible not only to capture reality but to alter it in realistic ways, these trends only accelerated with the onset of motion pictures and video. Advances in computer-generated images left audiences wowed as they watched dinosaurs come to life and fly through the skies of Pandora in 3D in the early 2000s. We are now entering a new moment, one in which viewers are mesmerized by the most "perfect" magical AI-generated worlds of fairies and gnomes. A new medium has become available to artists that they have begun using in captivating, probing and amazing ways.

But just as all Media that has come before could be used for good (to delight, to question, to protest, to probe), or ill (to manipulate, dissemble, and deceive) this new media is and will increasingly be used by bad actors seeking to manipulate others' understanding of reality for their own gain.

There are three potential responses we might mount:

First, just as the Arts and Craft movement of the 1800s, in response to mass production of products during the Industrial Revolution, elevated and celebrated the human-made, we could likely see the rise in our valuation of *certifiably* human-made creations. The Arts & Crafts movement elevated and supported the training and thriving of woodworkers, stone masons, painters, weavers, gardeners, cooks, and more. We were already seeing the rise of a strong "makers-movement" in the 2000s, this will likely accelerate and expand as people come to value what is known to be real, human, man-made, and the inherent value in the time spent by the artist and artisan.

Second, we will have to ask ourselves: "at what cost?" Just as there was a realization of the environmental and social damage of the cold industrial production machine of the 1800s and the government regulation of that machine for the protection of society, we will need to see a new wave of regulation that attempts to control the worst outcomes of this new innovation. Just as there was a need for workers rights advocates, child protection advocates, and environmental protection advocates, we'll need to see the development of a new type of advocate – one that advocates for the protection of human-made creations.

However, government moves slowly, and the pace of the development of AI is so fast, there is also an opportunity around a third, private, self-regulation strategy. We will likely see the development of a sub-industry along the lines of socially and environmentally responsible businesses, but which incorporates the commitment to not use AI as part of their value proposition. Today's Google search takes 10 times amount of energy of last year's Google search. A junior staffer may be able to draft meeting minutes more quickly with AI, but is the production of thousands of pounds of CO2 worth the saving of 10 minutes? We've made significant advances in calculating the environmental impact of corporate and private sector behavior. It is not difficult to calculate the energy use and the related CO2 emissions and climate impacts of AI at the enterprise level. Just as a consumer might choose to buy shoes from an environmentally responsible company and doesn't use sweatshop labor, a new consumer of the mid-2000s may choose to purchase from a company that has made a commitment to not only not use slave labor, not dump effluent pollution into the water, but also not use AI unnecessarily. A corporate commitment to not use AI would not only be an environmental commitment, but also a societal cohesion commitment to the workforce. The measurement of corporate ESG (for Environmental, Social, and Governance) performance has come a long way over the past decade, with the Global Impact Investing Network now having mapped out over 700 metrics to capture corporate behavior. Eschewing AI might be the next commitment companies can make to signal to their customers they care about their environmental and social performance, and to give consumers alternatives to corporates that will throw out their next zero commitments for slick AI-generated social media posts.

Reframing Understandings of AI

03.

AI IN HISTORICAL *Perspective*

Kyrill Kunakhovich

Corcoran Department of History
University of Virginia

"For the last twenty years neither matter nor space nor time has been what it was from time immemorial. We must expect great innovations to transform the entire technique of the arts, thereby affecting artistic invention itself and perhaps even bringing about an amazing change in our very notion of art." So wrote the French poet Paul Valéry in 1928, nearly a century ago. What prompted his reflections was the spread of the gramophone, which allowed people to listen to music in the home – "at our own time, according to our own mood."[1] Seven years later, the German cultural critic Walter Benjamin quoted Valéry in a seminal essay on "The Work of Art in the Age of Mechanical Reproduction." He focused on another new technology, sound film: "a spectacle unimaginable anywhere at any time before this." With its rapid succession of images, Benjamin argued, film shortened our attention spans and made concentration impossible. Its viewers remained in a perpetual "state of distraction," mindlessly consuming "illusion-promoting spectacles and dubious speculations."[2]

The hopes, anxieties, and criticisms we voice about AI today are hardly new. To optimists, artificial intelligence promises "to transform at will an empty hour, an interminable evening, an endless Sunday, into an enchantment" – as the gramophone did for Valéry. To pessimists, it threatens "a tremendous shattering of tradition," as film did for Benjamin. What these two thinkers shared, despite their differing attitudes, was a strong sense that they were living in unprecedented times. New technologies seemed to be changing their whole world, for better or worse, and a hundred years later, we feel the same way.

Is this a case of history repeating – two eras of intense upheaval, the 1920s and the 2020s? The political theorist Marshall Berman suggested that it may be something more. A sense of living on the precipice, he wrote, is the ubiquitous experience of modernity: "To be modern is to find ourselves in an environment that promises us adventure, power, joy, growth, transformation of ourselves and the world – and, at the same time, that threatens to destroy everything we have, everything we know, everything we are." Since at least the eighteenth century, Berman argued, every generation has thought that its world was falling apart.[3] We feel this viscerally today,

but so did Valéry and Benjamin, and so, too, did Karl Marx a century before them when he remarked that "all that is solid melts into air."[4] Today's AI is certainly a new phenomenon with a new set of challenges and opportunities. But the experience of dealing with the new and of confronting technological disruption is itself not as new as we assume.

So, what can we learn from this experience? Benjamin's essay offers some ideas. It revolves around the concept of "aura" – that special quality of an authentic artwork. It is this aura that gets lost in reproduction, Benjamin argued. While an original work of art is venerated ritualistically – think of a statue on a pedestal in a museum – it makes no sense to venerate a copy. "The instant the criterion of authenticity ceases to be applicable to artistic production, the total function of art is reversed," Benjamin insisted. "Instead of being based on ritual it begins to be based on another practice – politics." Writing in 1935, amid the Nazi takeover of Germany, Benjamin was highly attuned to the political uses of art. That was one reason why he so distrusted film, whose "illusion-promoting spectacles" the Nazis used to full effect. But Benjamin realized that one could not fight Nazi propaganda with appeals to authority or tradition. In the age of mechanical reproduction, the aura was long gone, and politics became the only battleground.

What Benjamin's insight suggests is that the issue of AI today is first and foremost political. One topic we discussed at length was how to build a better AI, and we should certainly try to make this technology more inclusive, equitable, and transparent. But in the end, AI's impact will depend less on what it can do than on how we use it. It is tempting to view AI as so radically new that it requires a radically new mode of governance. To some enthusiasts, indeed, it is the governance that must conform to the technology and not the other way around. Yet, there is nothing unprecedented about a new technology with the potential to transform the world; as Berman reminds us, this is a constant feature of modernity. Our challenge, then, is not so different than it was a hundred years ago. What political guardrails can we put in place so that AI has more of Valéry's "enchantment" than Benjamin's "dubious speculations"?

ENDNOTES

1 Valéry, P. (1964). The Conquest of Ubiquity. (Ralph Mannheim, Trans.). Aesthetics. (Original work published 1928)

2 Benjamin, W. (1969). The Work of Art in the Age of Mechanical Reproduction. (Hannah Arendt, ed.; Harry Zohn, Trans.). Illuminations. (Original work published 1935)

3 Berman, M. (2010). All That Is Solid Melts into Air: The Experience of Modernity. (Hannah Arendt, ed.; Harry Zohn, Trans.). (Original work published 1982)

4 Marx, K. & Engels, F. (1888). The Manifesto of the Communist Party. (Samuel Moore, Trans.). Charles H. Kerr & Company. (Original work published 1888)

LIFE, LIBERTY AND THE PURSUIT OF CONVENIENCE— *on whose terms and at what costs?*

Coleen Carrigan

Department of Engineering and Society
University of Virginia

Computing products used by the world's majority are developed by one of the most segregated fields in the US labor force. Subfields in computer science (CS) that focus on theory/algorithms are even more dominated by men than CS more generally.[1] The cost of this problem can be seen in the field's lack of feminist leadership and tolerance for sexual and gender harassment,[2] which is significantly higher than in non-technical fields. The uneven distribution of opportunities, resources, and respect in CS education and its technical workforce denies some groups the skills required for leadership in the 21st century. Preferential treatment of men in this powerful sector undermines feminists' efforts to increase women's earning capacities, access to power, and our political and bodily sovereignty in broader culture.

The Bro Code and Its Influence

Gender harassment in CS (consent violations that consists of verbal, physical, and symbolic behaviors conveying hostility toward women and non-binary people) is core to what I call the "Bro Code." The Bro Code[3] works to preserve the technical workplace as a homosocial sanctuary for cisgender men through worksite norms that prize combativeness, long hours, racialized sexism, and a disdain for altruism (caring about humanity and the social good). Women of color in technical fields experience the greatest amount of mistreatment, including being paid less than their majority peers[4], blocked from advancing into leadership and often, harassed and stereotyped.[5] Not only are individuals harmed by the Bro Code, but, given the reach

of algorithmic machines and the surveillance tools that make them possible, the Bro Code is a problem of global proportions.

Along with laptops, cellphones, media platforms and chatbots, the Bro Code may be considered another significant output of computing. At a time in US society when women's freedom is under siege[6] and misogynistic rhetoric[7] terrifyingly high, predatory behaviors in high-tech sites of research and development may have implications for the many domains on which algorithms impinge. For example, I see a correlation between gender harassment in high-tech and the extraction of humans' lived experiences facilitated through digital devices. Using consent as a lens to compare gender violence in CS workplaces and human data extraction in social domains can make visible the asymmetrical relations of power in both the development and applications of artificial "intelligence."

Rapacious Data Collection

The labor needed to train machines to reason comes from us, we who consume computerize tools and platforms. This includes the capture and analyses of people's biological processes[8], behavior, communication, and social network patterns. Do digital users have a choice to opt out of being surveilled? Privacy experts[9] are concerned about the deluge of privacy agreements that users in the US encounter (which, on average, would take 76 working days a year to read).[10] This, combined with the highly-specialized language in which they are written, renders the Organization for Economic Cooperation and Development notice and consent statue coercive. Consideration of privacy includes not just data collection, but also data use and dissemination.[11] People feel a sense of powerlessness, forced to acquiesce[12] without knowing toward what ends their personal data will be used. These widespread data extraction practices[13] threaten privacy and security for everyone, but disproportionately harm vulnerable populations[14] like women of color.[15]

Bro Code bosses not only ask "people to trade privacy and security for convenience"[16] but they also harvest human labor, creativity and behavioral patterns. The benefits[17] of this activity are maldistributed.[18] Their latest sport is stalking highly prized data[19] like peer-reviewed books and research articles. Their data thirst is so strong, they are open to violating corporate policies and worse, the law.[20]

While others are rightfully framing the data extraction processes on which the computing industry depends as plagiarism[21] and theft,[22] framing privacy and freedom in feminist terms[23] can forge a bridge to broader social movements for democracy, racial justice and women's rights. I do not want to occlude the harms of sexual harassment by making crude parallels to digital bosses coercing us to submit to dodgy "terms of service." Being harassed at work is not the same lived experience as being spied on via digital platforms. However, in both settings, the computing industry crafts and deploys techniques of coercion to normalize consent violations and minimize alternative options. For example, retaliation[24] against those who report harassment in scientific workplaces is common and calls into question computer scientists' ability to refuse and resist gender violence when their livelihood and careers are at stake. Similarly notice and consent practices offer

no option for negotiations. Despite harmful outcomes and violations of legal precedent, Bro Code bosses give us two options: either submit to their terms or refuse digital products upon which the global economy and (often) our livelihoods depend.

Feminist Organizing to Protect Privacy

Bro Code bosses—leaders of powerful companies that produce computing commodities in highly segregated workplaces rife with gender and sexual harassment—are hawking a future dominated by machines that require unfettered access to our behaviors and our social and creative labors. We are often forced to make individual decisions about whether or not to relinquish our constitutional rights to privacy to subsist in the current digital economy—all in the name of convenience. This acclimatizes us to a political climate that imperils these rights, which has implications for US democracy.[25] Coercive data practices have enabled a massive apparatus of surveillance networks and communication platforms that powerful actors use to strip Americans of their bodily sovereignty.[26] The Bro Code has helped produce the conditions for regressive gender politics to spread and take hold in state institutions and public imaginaries. Feminists organizing around privacy and autonomy in the domain of healthcare in a post-Dobbs world can build productive alliances with activists challenging privacy and boundary violations by Big Tech corporations. This will leave Bro Code bosses open to greater scrutiny and debate, and reorient the struggle over technology, public welfare and justice in the US. We simply will not lay back and enjoy a future dominated by chatbots and autocratic Bro Code bosses.

ENDNOTES

1 Shoham, Y., Perrault, R., Brynjolfsson, E., Clark, J., Manyika, J., Niebles, J. C., Lyons, T., Etchemendy, J., Grosz, B., & Bauer, Z. (2018). The AI Index 2018 Annual Report. AI Index Steering Committee, Human-Centered AI Initiative, Stanford University.

2 Klein, F.K., Onovakpuri, U. & Scott, A. (2017). Tech Leavers Study: A first-of-its-kind analysis of why people volunarily left jobs in tech. Kapor Center for Social Impact. https://search.issuelab.org/resource/tech-leavers-study-a-first-of-its-kind-analysis-of-why-people-volunarily-left-jobs-in-tech.html

3 Carrigan, C. (2024). Cracking the Bro Code. The MIT Press. https://doi.org/10.7551/mitpress/14883.001.0001

4 West, S.M., Whittaker, M. and Crawford, K. (2019). Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. https://ainowinstitute.org/publication/discriminating-systems-gender-race-and-power-in-ai-2

5 National Academies of Sciences, E., and Medicine. (2022). Transforming Trajectories for Women of Color in Tech. The National Academies Press. https://doi.org/10.17226/26345

6 Tolentino, J. (2022, June 24). We're Not Going Back to the Time Before Roe v. Wade. We're Going Somewhere Worse. The New Yorker. https://www.newyorker.com/magazine/2022/07/04/we-are-not-going-back-to-the-time-before-roe-we-are-going-somewhere-worse

7 PBS NewsHour. (2024, December 10). Researchers report stunning surge of misogyny after the election [Video recording]. https://www.youtube.com/watch?v=dgxj_wfH-0w

8 Zuboff, S. (2019). Surveillance Capitalism and the Challenge of Collective Action. New Labor Forum, 28(1), 10–29. https://doi.org/10.1177/1095796018819461

9 Cate, F. & Mayer-Schönberger, V. (2013) International Data Privacy Law, Volume 3, Issue 2, May 2013, Pages 67–73, https://doi.org/10.1093/idpl/ipt005

10 Mcdonald, A. M., & Cranor, L. F. (2008). The Cost of Reading Privacy Policies. I/S: A Journal of Law and Policy for the Information Society, 4(3), 543–568.

11 Nissenbaum, H. (2004). Privacy as contextual integrity. Washington Law Review, 79(1), 119-158.

12 Andrejevic, M. (2014). The Big Data Divide. International Journal of Communication, 8, 1673–1689.

13 Crawford, K. (2022). Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press. https://yalebooks.yale.edu/book/9780300264630/atlas-of-ai/

14 Shephard, N. (2016, June 2). 5 reasons why surveillance is a feminist issue. LSE Engenderings. https://blogs.lse.ac.uk/gender/2016/06/02/5-reasons-why-surveillance-is-a-feminist-issue/

15 Noble, S.U. (2018) Algorithms of Oppression: How Search Engines Reinforce Racism. NYU Press. https://nyupress.org/9781479837243/algorithms-of-oppression/

16 Burgess, C. (2024, February 12). Is privacy being traded away in the name of innovation and security? CSO. https://www.csoonline.com/article/1306877/is-privacy-being-traded-away-in-the-name-of-innovation-and-security.html

17 Catsaros, O. (2023, June 1). Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds | Press | Bloomberg LP. Bloomberg Intelligence. https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

18 Taylor, C. (2019, December 2). Silicon Valley giants accused of avoiding over $100 billion in taxes over the last decade. CNBC. https://www.cnbc.com/2019/12/02/silicon-valley-giants-accused-of-avoiding-100-billion-in-taxes.html

19 Metz, C., Kang, C., Frenkel, S., Thompson, S. A., & Grant, N. (2024, April 6). How Tech Giants Cut Corners to Harvest Data for A.I. The New York Times. https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html

20 Grynbaum, M. M., & Mac, R. (2023, December 27). The Times Sues OpenAI and Microsoft Over A.I. Use of Copyrighted Work. The New York Times. https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

21 Wong, M. (2024, January 10). What If We Held ChatGPT to the Same Standard as Claudine Gay? The Atlantic. https://www.theatlantic.com/technology/archive/2024/01/chatgpt-plagiarism-copyright-harvard-claudine-gay/677086/

22 Mims, C. (2023, February 4). AI Tech Enables Industrial-Scale Intellectual-Property Theft, Say Critics. Wall Street Journal. https://www.wsj.com/articles/ai-chatgpt-dall-e-microsoft-rutkowski-github-artificial-intelligence-11675466857

23 Carrigan, C., Green, M. W., & Rahman-Davies, A. (2021). "The revolution will not be supervised": Consent and open secrets in data science. Big Data & Society, 8(2). https://doi.org/10.1177/20539517211035673

24 National Academies of Sciences, Engineering, and Medicine. (2018). Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine. Washington, DC: The National Academies Press. https://doi.org/10.17226/24994.

25 Jackson, L. (2021, May 21). Shoshana Zuboff Explains Why You Should Care About Privacy. The New York Times. https://www.nytimes.com/2021/05/21/technology/shoshana-zuboff-apple-google-privacy.html

26 Sorkin, A. R., Giang, V., Gandel, S., Hirsch, L., Livni, E., & Gross, J. (2022, June 30). Reconsidering Privacy Risks After Roe. The New York Times. https://www.nytimes.com/2022/06/30/business/dealbook/abortion-privacy-risks-data.html

AGAINST NEFARIOUS DISCOURSES OF AI *as Labor-Saving and Problem-Solving*

Caitlin D. Wylie

Department of Engineering and Society
University of Virginia

Discourse: *AI can replace human labor!* Automating away grunt work will give us all more time for creative, relational, expert work, or perhaps for free time to spend as we like… right? Research on prior technologies has found that "labor-saving" devices generally don't save labor; they change it. (A fellow participant raised this important point during our discussion at the workshop.) Ruth Schwartz Cowan, an influential historian of technology, made this insightful argument in a 1987 paper titled "Less Work for Mother?". She showed that mid-century American technologies marketed as reducing housework—e.g., the vacuum cleaner, the washing machine, the family car—actually created higher social expectations for cleanliness and the frequency of errand-running for middle-class women. These changing social norms, brought about by supposedly "labor-saving" technologies, thus laid more work on women.

So let's be extremely skeptical of these labor-saving claims, and instead ask, what kinds of labor might AI save? Whose labor is it? Without that labor, what might society look like? Let's work together to anticipate and recognize how AI use changes our social norms, so that we can push for the changes we want and resist the changes we don't want. Capitalism will always adapt to extract more labor from us; we should ask how we can oppose that pressure and use our time in ways that align with our values, e.g., rest, helping others, protecting the environment, etc.

AI can solve all problems! One of my undergraduate engineering students proposed a topic for their fourth-year capstone research that focuses on AI's environmental harms, specifically the high use of electricity and water by data centers. They argued that a solution

for this problem is using AI to optimize data centers' use of resources. Put another way, AI can solve a problem that AI creates. This idea is widespread among my engineering students and colleagues. Yet it overlooks so much, specifically by neglecting or misunderstanding the scale of possible optimization. I imagine that an AI model can reduce a data center's energy and water use by some small percentage, but that model also requires some amount of electricity and water to train and run itself. Is that operating cost worth it?

For me, the more nefarious side of this argument is the claim that if a data center has reduced its resource use by some percentage, then it is sustainable and responsible. Using AI to regulate its resource use also arguably creates more business for the data center. Resource optimization is certainly not a solution, and it can be used as a misleading form of "greenwashing", i.e., when a business claims sustainable practices in its marketing to improve its reputation without actually protecting the environment. A data center claiming that it uses AI resource optimization to use 12% less electricity and water than its competitors (or whatever the percentage might be) isn't doing much to protect the planet, in the grand scheme of things. They're still using inordinate amounts of resources, while trying to trick us into thinking that they're doing the best they can. I fear that the belief that AI can solve problems that it creates, in addition to any other kinds of problems, will help justify ever-expanding AI use and thus ever-growing resource demands. We must do better than this.

What can we do? One way to begin challenging these dangerous discourses is by improving understanding among the many groups affected by AI, such as the public (especially communities harmed by data centers), policymakers, tech companies, and tech experts. We have seen tech companies willfully mislead governments and the public, and we have seen governments struggle to understand tech products and systems. We need people to help translate each group's knowledge to help us all better understand how AI works, how regulation works, and how AI affects social and sociotechnical systems (e.g., the economy, the environment, public well-being). To bridge diverse worldviews and values, it would be powerful to have mediators who can facilitate communication among these groups. This expertise already exists in environmental justice, such as professional conflict mediators who help fossil fuel companies and frontline communities talk to each other about what they each need (e.g., the US Environmental Protection Agency's Conflict Prevention and Resolution Center https://www.epa.gov/eccr).

Understanding is the first step. Once we are on a more level playing field, then we can all join the game. This diverse collaboration among groups can help build socially beneficial values of consent, privacy, and public good directly into AI technologies and regulations. That work begins with getting our stories straight, tossing out misleading discourses, and then creating a shared narrative of what society with AI should be like.

BAKING *in Bias*

Jessica Ellen Sewell

Urban and Environmental Planning
University of Virginia

Imagine never having eaten a pizza, only having seen an image of one, and being asked to make one. You can see that it has something reddish on it, and cheese, but you don't know what the red stuff is, or what kind of cheese. You can use your skills to make something that looks right, but the chances are good that it won't taste quite right. The images and text created through AI algorithms often feel like this pizza. An AI-generated image may look like a real photo or painting, but an extra finger, a misplaced foot, or an odd blur lets us know that it is a simulacrum. This problem is built into predictive text or predictive images; they are never based on knowing the recipe, but on imitating what is created through the recipe. With more and better data and tweaks to the algorithms, the simulacrum can come closer to the real thing, but it will always be imitating the image, not following the recipe.

The quality of what is created through AI is rooted in the quality of the data it is based in. The maxim "garbage in, garbage out" summarizes the problem. While each AI program has its own corpus, they are generally trained on what is available online, including websites, digitized books and articles, and emails and other messages. This data often includes copyrighted images and texts, potentially causing significant harm to artists and writers whose work is imitated through the use of AI. In addition, this corpus is not neutral; it reflects and even magnifies prejudices. Even before AI programs became generally available, this prejudice was readily visible. For example, Wikipedia contributors creating entries about women at edit-a-thons found their entries deleted or challenged at a much higher rate than entries about men. Google image searches for professors and CEOs have returned largely white men, while searches for unprofessional hairstyles bring up Black women. AI programs trained on this biased data not only reproduce biases but magnify them.

The bias baked into AI through its corpus is not just a quality problem, it is an ethical problem. Using AI to sort through people's profiles to suggest who to interview for a job or who to treat as a potential criminal can cause direct harm, as can AI that has a harder time recognizing non-White individuals or understanding voices with certain accents or voice registers. As outputs from AI creep into

our everyday lives, in advertisements, blog posts, publications, and summaries that social media and software companies give us no way to turn off, they subtly influence us in similar ways that biased language and images created directly by humans do. They regularly present misinformation as fact and give us an image of the world that is whiter, straighter, and more male, skewing our sense of what is normal and what is possible.

AI programs' productions and corpus are also biased in ways we don't usually think about. In the design fields, particularly landscape architecture, there has recently been an interest in the ways that nonhuman species are affected by and participate in the world that humans shape and inhabit. Both AI programs' corpuses, and the purposes for which they have been designed, are biased towards humans. When we shape the world using AI tools, we shape a world that is intended to benefit humans, and that benefits some humans more than others. That world is less well suited to nonhuman species, and diminishes all experiences, human and nonhuman alike. By their nature, AI programs remix and rehash human-centric, biased bits of information, creating novelty through combination but never inventing anew, and never responding to the complex needs of a world that is not male, white, or even all human.

THE DIGITAL *and the Urban*

Will Straw

McGill University

This short essay is prompted by the conviction that we must think of the digital (and Artificial Intelligence in particular) in relation to the urban, and that the relationship between these two domains is the one which will most clearly shape our futures. The world is not fully digital, of course, nor is it fully urban, but both represent tendencies which seem unstoppable and definitive of contemporary life. It is by reflecting upon the relationship between the two that we may confront some of the most urgent questions facing us as citizens and societies.

Artificial intelligence, as the most recent and, perhaps, the most transformative expression of the digital, is not an urban phenomenon exclusively, of course. AI has found many of its most significant uses at the level of the nation state. Some of these uses may seem innocuous, like the chat bots which answer our questions about public benefits. Others seem more ominous – for example, the use of biometric recognition and other technologies of surveillance in the control of borders and control of populations. In our everyday lives, we perhaps encounter AI most frequently in the seemingly placeless operations of global capitalism, like the recommendations we receive from the algorithms of retail or streaming sites.

Still, as Federico Cugurullo et al suggest,[1] cities have become the privileged terrain for experimentation with AI. Procurement initiatives, AI-based experiments in "twinning," and AI-based enhancements to infrastructure all seem more appropriate and feasible at the level of the city than that of the nation state. At the municipal level, AI initiatives may be more easily scaled-up, to link with networks at the national or transnational levels, or scaled-down, to the level of the home or the individual.

The "smart city" initiatives of a decade or two ago laid the groundwork for a use of AI which promised the city-dweller a life characterized by heightened efficiency, democratic responsiveness, and the augmentation of sensory experience. And cities, it seems, much more than nation-states, have been caught up in a sense of themselves as places of innovation, in which the tools of AI might be marshalled to attract high-tech work forces and turn the "creativity," so valorized (if vaguely defined) in turn-of-the-millennium urban planning, into an asset whose value could be more easily measured.

In many respects, this "urbanization" of AI is mirrored in the contemporary urbanization of other phenomena. Many of the most influential tendencies within contemporary art, are about urban place-making or seek to intervene in the social relations of the city. Cultural policy is more and more the domain of municipal governments, as national states have vacated the terrain in the face of neoliberal defunding or perceived helplessness rooted in technological change. Even political struggle, the urbanist Federica Gatta[2] has argued, is more and more a struggle over space, in contexts which are mostly urban.

If the city is the most common terrain for the implementation of AI, we may only hope that the ways in which we see and judge it are themselves "urbanized." This is to hope that AI is judged in the terms used to evaluate other features of urban life. Rather than simply serving as the efficient "middle scale" for AI experiments, cities should infect this experimentation with the sorts of concerns and values which preside over the most progressive features of urban life.

Urban governance is confronted daily with the question of how we live together in various kinds of proximity - in proximity to other people, to other species, to circulating forms of knowledge and to environments of all kinds.

An urbanization of AI should compel us to hold AI up to the same sorts of judgement we apply to other features of urban life. The most important of these issues is the way in which, as a collectivity, we contend with social differences of all kinds. AI platforms and cities each generate spaces in which the differences of class, race, ethnicity, and sexuality (to name only a few) are structured and made meaningful. The struggle for a progressive governance of AI should, like the battles for urban democracy, be concerned with protocols for living together founded on the recognition of both social differences and collective rights.

ENDNOTES

1 Cugurullo, F., Caprotti, F., Cook, M., Karvonen, A., McGuirk, P., & Marvin, S., (2024). Conclusions: the present of Urban AI and the future of cities. (Cugurullo et al. eds.) Artificial Intelligence and the City: Urbanistic Perspectives on AI. https://doi.org/10.4324/9781003365877-26

2 Gatta, F. (2018). (Contre)Pouvoirs Urbains? Eléments pour une critique anthropologique de l'urbanisme participatif. Paris: Editions donner lieu.

FROM UTOPIAN DREAMS TO THE DARK FOREST: *GenAI and the Changing Landscape of the Web*

Pedro Augusto P. Francisco

Department of Engineering and Society
University of Virginia

The Internet user experience has changed dramatically over the past 40 years, especially regarding the Web. In the 1990s and early 2000s, surfing the internet – an expression that became outdated – was a novel experience for most people. There was a mix of curiosity, fascination, fear, distrust, and frustration. Despite the limitations of dial-up connections and the hesitation to meet and talk to strangers online, users were excited about new forms of communication and access to information, such as email exchanges and chat rooms, which later evolved into the first social networks.

This way of experiencing and interacting with the Internet was tied to a turn-of-the-century idealism, characterized by a belief in the transformative power of the Web to democratize access to information, promote connectivity, and empower individuals. It was the era of the Silicon Valley boom, with startups and enthusiasts envisioning a future where technology would break barriers and revolutionize communication. The Declaration of Independence of Cyberspace,[1] written in 1996, reflected this spirit of autonomy from traditional governments, driven by a utopian and technodeterministic vision of a more connected, free, and egalitarian world.

Today, our user experience reveals a very different Web. We live in the era of the Dark Forest of the Internet. Proposed by Yancey Strickler,[2] the Dark Forest argument is based on an idea originally presented by science fiction writer Liu Cixin in his book trilogy, The Three-Body Problem. As the argument goes, the perception of being the only intelligent species in the universe doesn't stem from actual solitude but rather from the fact that, like in a dark forest, all the other species

remain silent and hidden to avoid attracting the attention of predators.[3]

Similarly, faced with a profusion of bots, advertising, algorithms, clickbaits, and scams, internet users have stopped navigating open virtual environments, prioritizing interactions in closed spaces such as Discord groups, email newsletters, private messaging apps, or other environments accessible only by invitation. This is the only way to ensure interactions exclusively with other humans without the mediation and interference of unwanted agents. The gradual migration to these spaces leads to the current feeling that the open Internet, at least as we know it, has died.[4] Its real inhabitants are hidden in the protected corners of this dark forest.

With the rapid advent and popularization of Generative AI, especially Large Language Models (LLMs), the dark forest argument becomes even more relevant.[5] In a very short time, the Web has become flooded with content – texts, images, and videos – generated by AI. This content is posted on optimized pages where engagement with other bots further increases their visibility and reach. Suddenly, it has become difficult to identify what was made by humans and what was artificially generated. In doubt, one can only trust what is produced within the walled gardens of the Internet.

The most evident problem with this scenario is that the Internet's days of providing an environment for meetings and information exchange—which in turn stimulated spontaneity and creativity—are over. We now have a profusion of superficial texts and derivative content resulting from algorithms based on third-party intellectual property.

However, there is another more insidious problem. One proposed solution to encourage using open spaces on the Web would be creating means to attest to users' humanity. Thus, it would be possible to know beforehand who is a real user and who is just artificial intelligence pretending to be human. This solution is being offered by companies like Tools for Humanity, which developed a technology called World ID – an individual digital identity linked to each user's iris scan.[6] In other words, individuals provide sensitive personal data in exchange for their certificate of humanity. As if the privacy risks posed by this technology were not enough, the situation worsens because one of its founders is Sam Altman – CEO of OpenAI – the same company that developed ChatGPT. We then have the same actor offering both the antidote and the poison.

Thus, the dark forest analogy extends beyond user behavior to the actions of those who control the digital landscape. Much like racketeers who create problems only to sell solutions while trying to establish control over the neighborhood, some tech companies contribute to the chaos of the open Web while simultaneously offering the tools to navigate it. This dual role exacerbates the sense of distrust and forces users into walled gardens where they feel safer but are also more controlled and monitored.

Technological fixes[7] alone will not resolve the issues of trust and authenticity online—problems created by technological advancements themselves. Instead, users

should adopt a more proactive approach. If the current state of the web forces people to act under a survival instinct, there can be no trust in those who control the dark forest, primarily because they are the ones who cast shadows over those woods. If the web is now filled with corporate-created digital demons,[8] users must act as demonologists and exorcists, not cowering in fear but recognizing that demons exploit human vulnerabilities to present us with Faustian bargains. Relying on these deals and quick fixes can obscure the necessity for broader, systemic reforms that tackle the root causes of these problems.

ENDNOTES

1 Barlow, J. P. (2016) A Declaration of the Independence of Cyberspace. Electronic Frontier Foundation. https://www.eff.org/cyberspace-independence.

2 Strickler, Y. (2024). The Dark Forest Theory of the Internet. The Dark Forest Anthology of The Internet 19–25. https://www.ystrickler.com/the-dark-forest-theory-of-the-internet/

3 Liu, C. (2016). The Dark Forest. (Joel Martinsen, trans.). Fiction. London: Head of Zeus.

4 Nover, S. (2024). It's the End of the Internet as We Know It—and I Feel Fine. Fast Company. https://www.fastcompany.com/91246383/its-the-end-of-the-internet-as-we-know-it-and-i-feel-fine.

5 Appleton, M. (2024). The Expanding Dark Forest and Generative AI. The Dark Forest Anthology of the Internet, 129–39. The Dark Forest Collective. https://maggieappleton.com/ai-dark-forest/

6 Tong, A. (2024). Sam Altman's Rebranded Worldcoin Ramps up Iris-Scanning Crypto Project. Reuters. https://www.reuters.com/technology/artificial-intelligence/sam-altmans-rebranded-worldcoin-ramps-up-iris-scanning-crypto-project-2024-10-17/.

7 Rosner, L. (2004). Introduction. Technological Fix: Visions, Trials, and Solutions. Hagley Perspectives on Business and Culture. London: Routledge.

8 Doctorow, C. (2017). Cory Doctorow: Demon-Haunted World. Locus Online (blog). https://locusmag.com/2017/09/cory-doctorow-demon-haunted-world/.

AI AND *Transductive Mimesis*

Rafael Alvarado

School of Data Science
University of Virginia

It is a commonplace that ethical concerns surrounding artificial intelligence are framed in terms of bias. Although much good work has been done to expose the existence and effects of bias on socially deployed AI tools, overuse of the concept limits our thinking on the subject of AI ethics. The word derives from a concept in statistics used to designate all the things that can cause one to over or underestimate the true value of a population parameter. From here it has been generalized to define any kind of fallacy in human or machine reasoning. So now bias, in the forms of cognitive and algorithmic bias, has been used to name hundreds of ways that rational judgment can go awry, from the placebo effect to stereotyping—far beyond the ken of things like sample bias. It is not that there are not myriad ways that humans under and over-value certain forms of information. It is that the concept flattens our understanding of cognition into the rational and irrational. It demonizes reasonable judgments of value and taste and reduces all ethical questions to model estimations.

More important, an overextended idea of bias deflects our attention away from understanding how cognition works in the first place. Anthropologists, drawing from a rich ethnographic and archaeological record that spans the planet and thousands of years, have long advised against reducing human thinking to a concept of rationality rooted in a narrow ontology and have had comparative epistemology front of mind.

What is at stake here in relation to AI is how we conceptualize the relationship between the epistemic and the ethical, between how we—humans and machines—come to know the world and how we ought to act in it. In the context of AI, when exploring issues such as algorithmic bias, trust, and explainability, we ought to be framing these questions in terms of a rich set of ideas within which bias is only one. Among these ideas the concepts of representation and mimesis are particularly useful.

We live in an astoundingly interesting, if frightening, era in which the accumulated body of knowledge contained in the vast media corpus that is the internet has been condensed into Borgesian networks of model weights. Instead of thinking of these networks purely in terms of the biases they inherit and

amplify, we would do better to think of them as forms of transductive mimesis of human thought, copies of great patterns of collective representations, of Braudellian magnitude, waiting to be explored. Instead of seeking to remedy these representations through post hoc layers of counter-biases based on our current preferences, we might explore them to arrive at better understandings of human thinking and behavior.

We have the tools to pursue this work. A key to understanding these patterns is in the vector embedding spaces on which all LLMs are trained and the resulting layers of weights and connections that constitute the models. These exhibit complex geometries of semantically analogical and oppositional structures that have recently been opened up by methods such as dictionary learning and monosemanticity,[1] building on a longer tradition of computational semantics.[2] What is missing is a research framework grounded in the domain knowledge of human thinking from a comparative perspective. Particularly useful in this regard are symbolic anthropology and cognitive science, both of which contain rich frameworks that characterize human thinking in terms of geometries and topologies of meaning that align with computational models.[3]

How does this project connect to the ethics of AI? Philosophers and social theorists since Kant have long posited a connection between ontological categories and ethical judgments, pure and practical reason, worldview and ethos. By empirically describing the shared networks of meaning that shape human thinking, an entire field of experimental philosophy is opened up: we put ourselves in position to explore and test hypotheses about how overt human judgements are connected to latent perceptual models. As a result, we may become aware how own ontologies and how they influence our judgements. This is much more useful than hardcoding our own biases into AI models in the name of ethics.

ENDNOTES

1 Li, Y, Michaud, E., Baek, D., Engels, E., Sun, X., & Tegmark, M. (2024). The Geometry of Concepts: Sparse Autoencoder Feature Structure. arXiv. https://doi.org/10.48550/arXiv.2410.19750. / Ethayarajh, K. (2019). How Contextual Are Contextualized Word Representations? Comparing the Geometry of BERT, ELMo, and GPT-2 Embeddings. arXiv. https://doi.org/10.48550/arXiv.1909.00512. / Coenen, A., Reif, E., Yuan, A., Kim, B., Pearce, A., Viégas, F., & Wattenberg, M. (2019). Visualizing and Measuring the Geometry of BERT. arXiv. https://doi.org/10.48550/arXiv.1906.02715.

2 Widdows, D., and Kanerva, P. (2004). Geometry and Meaning. CSLI Publications, Center for the Study of Language and Information.

3 Lévi-Strauss, C. (1966). The Savage Mind. University of Chicago Press. / Gärdenfors, P. (2014). The Geometry of Meaning: Semantics Based on Conceptual Spaces. The MIT Press. https://doi.org/10.7551/mitpress/9629.001.0001.

ARTIFICIAL OR *Natural Intelligence?*

Andre Vinicius Leal Sobral

Digital Technology for Democracy Lab, Karsh Institute of Democracy
University of Virginia

Intelligence is one of the most puzzling assumptions regarding complex phenomena observed during our brief time on this planet. It is a Self-aggrandizing adjective chosen by dominant Western epistemology to congratulate its own cognitive accomplishments — an intelligence often denied to other species, yet now to machines under the label of "artificial." Are these systems merely fancy statistics, probabilistic machines, or stochastic parrots? Large language models dazzle the public with promises of productivity and social and cultural revolution, all while accelerating our march toward environmental collapse. While these robots are practical assistants for routine grammatical tasks and remarkably effective at generating repetitive text structures, they unsurprisingly lack depth and coherence when it comes to reflective thought.[1] Yet, these omnipresent genies have freely bestowed their gifts and worked tirelessly to fulfill countless desires of their users—not to win their favor, but to foster dependence.

The evident optimism surrounding tailored experiences with intelligent companions has been driven by both the race to consolidate a new market and a profit-driven agenda. The narrative surrounding innovative technology companies often centers on disruption—brave visionaries shaking the conservative foundations of the market. However, this facade conceals the collective effort required to produce and gather data to build such machines. These cultural works of text, sound, and visual arts are often carelessly appropriated, disregarding authorship and original intention, producing endless remixes of their contents. Billions of images, particularly those of faces, are captured under the promise of predictive applications that would ensure a safer world. Moreover, the idea that this artificial intelligence is sufficient to replace humans is unethical, coercing citizens into participating in a scientific experiment where the integrity of their lives and bodies is at stake.

Creativity, security, efficiency, convenience, and innovation are rendered weightless as machine intelligence is often tethered to the internet and the concept of cloud computing.

Faraway servers conceal the materiality of this electronic brain and depict agility in its responses, whether narrated in scarlet, disembodied voices, or formally prosaic text. The wizardry of computer technology distances users from the environmental costs of each mundane question, humorless pun, silly query, and frivolous command made for self-amusement. Nevertheless, artificiality leaks out of the genie's lamps into our digital spaces, producing informational junk that floods search queries, messaging apps, and social networks—much like oil spills do with natural environments, causing direct damage to users and transforming these spaces into dark reflections of their former purposes.

However, there is no inevitability, no technological determinism inherent to this scenario, so we must ask ourselves: Can a manufactured intelligence be based not only on collective works but also have communitarian aims? Can we envision synthetic intelligence devoted to protecting the environment? Is it possible to design intelligent helpers to support those committed to the fight for ecological preservation?

Beyond the origins and purposes of Artificial Intelligence, its application demands regulation and bold perspectives from lawmakers. Large language models must be explainable, with clear mechanisms for its users and the government. If these algorithms are capable of directly harming and influencing individuals in society, there is a necessity to restrict and supervise their use to prevent abuse. If these systems are built through theft and the aggressive appropriation of others' labor, and their data banks influence the distribution of liberty and justice, then these data banks must be open and publicly accessible. If their construction and use occur without consent, laws must be in place to ensure appropriate consequences for their misuse. If this technology poses a threat to the planet's future, strict limitations on its implementation and use are essential, including platform accountability laws, public oversight of digital infrastructure, regulations on carbon emissions and energy consumption, and the expansion and protection of digital human dignity and labor rights.

ENDNOTES

1 Zhang, C., Hofmann, F., Plößl, L. et al. (2024) Classification of reflective writing: A comparative analysis with shallow machine learning and pre-trained language models. Educ Inf Technol 29, 21593–21619. https://doi.org/10.1007/s10639-024-12720-0

LIMITS OF *Artificial Innovation*

Steven L. Johnson

McIntire School of Commerce
University of Virginia

For innovation and creativity, human judgement and shared experiences are irreplaceable.

Technology vendors are spending billions of dollars every year to push the frontier capabilities of large-language models. The ambition to build larger and larger generative AI (gen-AI) models is driven by great expectations for their dramatic problem-solving abilities. The hype for gen-AI includes hopes for addressing daunting problems, including the global climate crisis and cancer, by generating novel solutions.

In this essay, I define Artificial Innovation as the ability of gen-AI to produce exhaustive lists of potential responses to prompts through the recombination of possibilities already present in its training data. Central to applying this capability to creativity and innovation is the gen-AI can combine existing elements in novel ways that have never been seen before. Yet, as can be seen when gen-AI fabricates facts—often referred to as hallucinations—gen-AI lacks the expert knowledge to discern among all possibilities which potential solutions are grounded in reality. Further, because the diffusion of innovation is an inherently social process, Artificial Innovation, lacking human judgment and discernment, provides limited value in closing the large gap between ideas, implementation, and impact in a real-world context.

At first glance, the ability of gen-AI to quickly outperform humans in producing a greater variety of potential solutions sounds compelling. Yet, the history of innovation across diverse domains including science, arts, and technology suggests this form of Artificial Innovation is insufficient. There are nearly infinite ways to combine and recombine existing elements. While innovation is, by definition, characterized by novelty, novelty alone is insufficient for finding innovative solutions to important problems.

When humans react to novel situations, there is a fine balance between boredom, intrigue, and confusion. If something novel is too familiar, it leads to boredom. When something is too novel and lacks legibility, it leads to confusion and a desire to withdraw from or reject it. New creations are most compelling when they provide an optimal mix

of familiarity and novelty. In short, we are drawn in and intrigued by new things that are familiar enough to still be understandable and, thus, something we can implement or use.

Even when gen-AI can help identify plausible solutions and, with the help of experts, be used to identify promising solutions, there remains a gap between ideas and impact. For new ideas, solutions, or creative expressions to be accepted by others—to be valued and adopted—is a complex social process. Even when an invention is provably superior along many dimensions to existing solutions, adoption—not only for digital goods but also for physical goods and services—requires individuals to overcome the inertia of current use and an inherent skepticism to change.

Proposed solutions are most likely to be valued in a community of practice when they incorporate ideas that are novel to that community and are proposed by individuals who are socially active within the community.[1] Being immersed in the context where a solution will be evaluated and, potentially, adopted provides important information about both what solutions are most likely to be easily adopted as well as knowledge of how to present those solutions in readily understandable ways. These are human capabilities that build upon interpersonal relationships and contextualized knowledge, capabilities far beyond what is possible with Artificial Innovation.

In summary, the ability of gen-AI to perform Artificial Innovation—to surface exhaustive lists of potentially innovative solutions to challenging problems—is inadequate for navigating the complex social processes required to progress from ideas to impacts. Human experts are in the best position to identify from all possible AI outputs what will be viewed as creative or innovative by the intended audience. Core community members have the social capital and local knowledge that leads to the acceptance of new ideas. Lacking these abilities, expecting AI to solve thorny societal problems remains more hype than reality.

ENDNOTES

1 Safadi, H., Johnson, S. L., & Faraj, S. (2021). "Who contributes knowledge? Core-periphery tension in online innovation communities." Organization Science, 32(3), 752-775.

IF YOU LISTEN,
You Can Hear What Chatbots Are Trying to Tell Us about Schooling

Peter Norton

Department of Engineering and Society
University of Virginia

Dan Robbins invented paint-by-numbers in 1950. Others had tried, but Robbins' kits were the first to sell. In 1951, they were a regional novelty; by 1953, they were a national fad.

The ads promised customers that with a paint-by-numbers kit, they'd paint a work of art – one "with your signature." Some highbrows cautioned that paint-by-numbers cannot make anyone an artist and that the results may be decorative but were not art. Robbins, however, was a believer. He said his kits applied a process he called "the Craft Master system," by which "absolutely anybody can paint a technically perfect picture in oils."

This difference of perspective never matured into a major controversy. No art student we know of was tempted to pass off her paint-by-numbers landscape as her original work. No art school convened meetings to develop policies to restrict or ban the Craft Master system. No entrepreneur came up with a paint-by-numbers detection kit.

One reason, of course, is that a paint-by-numbers kit could only promise its customers a picture that thousands of other customers would also paint. This offered the sham artist no safe path to a successful deception. But if we set aside this constraint, we can see that the temptation to engage in such a fraud was negligible anyway.

Even at the hands of a skillful painter, paint-by-numbers kits yield impersonal work that may be attractive but that express little or nothing personal from the painter. From viewers, they will evoke little response, apart perhaps from admiration for the painter's precision brushwork. While a few people might welcome a means by which to achieve an unearned

reputation as an artist by deception, paint-by-numbers will not serve the purpose.

The painter may find pleasure in painting by numbers – a pleasure that might also be found from coloring the pages of an elaborate coloring book or from assembling a jigsaw puzzle of a Rembrandt – but there will be little lasting attachment to the finished work. The painting may, in some sense, be expressive, but it will not be the painter's own expression. For the painter, the act of painting may be rewarding, but the picture will not be an enduring source of true, personal satisfaction.

For reasons such as these, paint-by-numbers did not confront the art world with an existential threat or cause a panic demanding the convening of expert panels.

Many of us used training wheels to help us learn to ride a bicycle. They kept us from falling over, but a point soon came when they were no longer of use. We could have kept using them, of course – but the pleasure of riding a bike lay in exercising the skill of cycling. Once this begins, training wheels are a nuisance. Cycling clubs don't have to have anxious meeting to develop policies about training wheels. Cyclists don't want them.

In schooling, however – and especially in higher education – large language models are causing such anxieties about students' writing. Administrators invite teachers to panel discussions and draft policies. In meetings, one person will warn others that teachers can expect most students to use AI chatbots to write much of their written assignments for them. Another will say teachers have a duty to police their students' work for evidence of such shortcuts. This reminder will lead a third person to explain that there is no reliable way to prove that a student has passed off chatbot-written text as their own – an observation that returns the cycle of comments to its starting point. The rounds are repeated until the meeting is adjourned.

Maybe we don't have to ride this merry-go-round. There is a pleasure in writing, and in finding ways to write well, that is not unlike the pleasure of painting or bike riding. This satisfaction, as an experience most people know firsthand, requires no explanation. There is pleasure, psychologists tell us, in "self-efficacy" – that feeling we get when we find we can do something that demands some skill from us and can do it in our own way. This is the kind of effect that makes painting by numbers a mere novelty, and training wheels irrelevant. The pleasure is in the painting. The pleasure is in the cycling.

But in schooling, writing's pleasures are distrusted to the point that teachers routinely assign papers with mandatory minimum length limits, in an effort to force the writing out of students. Students feel compelled to fill pages – and their writing often shows it. Perhaps some such pressures may be practical necessities, but to work well they may only supplement, not substitute for, the pleasures of writing.

If teachers fear that AI chatbots will write students' papers for them, they devise policies to deter recourse to chatbots, find means of detecting chatbot prose, and devise assignments that make chatbots less useful. But they can also take this problem as a symptom demanding their attention.

Writing, like painting and bike riding, can be an inherently attractive and rewarding mode of expression – a mode so rich in its own satisfactions that delegating it to a robot makes no more sense than fastening training wheels to your bike, or following a paint-by-numbers kit instead of expressing one's own creativity with a brush.

If educators fear chatbots, the problem may lie as much in the norms of institutional education as in the chatbots' powers.

AI AND EVERYDAY TRAVEL: *Relinquished Intelligences*

Andrew Mondschein

Department of Urban and Environmental Planning
University of Virginia

Taking the promises of artificial intelligence (AI) at face value, AI claims it will solve society's problems more effectively than current, presumably human, intelligences can. In an urban context, this promise manifests as a host of public-private initiatives, often lumped into the already worn catchall of "smart cities." Smart cities applications encourage optimization, seeking to enable urban functions that are more efficient, safe, sustainable, and resilient. While these may be laudable goals, scholars and activists alike have made penetrating critiques of urban optimization. They have emphasized threats to equity, democracy, and the likelihood of entrenching systematic biases and power relationships in AI urbanism, even as our human-centered approaches have tended to come up short. These critiques call for instituting strong ethical and political control over smart cities. While we may ultimately exert social control over AI urbanism, these approaches still substitute one mode of intelligence for another, diminishing human intelligence in urban settings.

Urban transportation puts the tradeoffs between AI and human intelligence into high contrast. AI and machine learning are already widely deployed in the transportation sector. Driverless vehicles are a prime example of this, but even without automation, algorithmic routing services such as Waze and Apple Maps have already pushed many drivers and other travelers into obligatory symbiosis with AI. To save time, people relinquish parts of their spatial cognitive faculties to navigation services. It's not just routing: search algorithms suggest shops, health care, restaurants, and other activities, optimizing for crowdsourced service quality and other factors including advertising revenues. As a result, many of us are now living in an experiment where the choices of where and how to travel in cities are now made or mediated externally, using algorithmic intelligence instead of our own innate cognitive processes.

Are there consequences to this tradeoff, beyond an ironic nostalgia for paper maps and getting stuck in traffic? Urban cognition – the set of choices we make to live our lives in cities – draws upon neurobiological structures and processes. Navigation and other tasks that extend our lives into the surrounding environment are seated primarily within the brain's hippocampus. The hippocampus links spatial knowledge with other types of episodic memory and serves as a key site for facilitating personal goal attainment. The neurobiological foundations of urban cognition do not sanctify it but simply underscore that urban cognition is evolutionary, developmental, and part of how we define ourselves as human. To leave behind these distinctive modes of human learning is to relinquish the intelligence that comes with them. This is a choice we may be making without much conscious consideration, while multiple findings show that cognition has a "use it or lose it quality." To not exercise our spatial cognition is not just a matter of capability but also potentially opens us up to long-term consequences, with evidence of increased susceptibility to dementia.

Beyond direct health consequences, algorithmic technologies change how we engage with our surrounding environments. Because they solve problems that previously were ours to address, we become more passive travelers and dwellers, forgoing active engagement with environments. Consequently, the environment's ability to serve as a communicative partner is reduced, with signs, architecture, and even natural features losing their salience. Even if human intelligence in transportation systems results in suboptimal outcomes – crowding, inefficient behaviors, missed opportunities – these intelligences are resilient and diverse. Despite the pitfalls and frustrations, relying on our own cognition to move through the city is resilient to potential disruptions that could shut down navigation and automation systems. Further, the distinctiveness of each person's urban intelligence is itself valuable, providing diverse perspectives and strategies for dealing with shared urban terrains.

Transportation is just one aspect of urban intelligence, and individual behavior sits at one end of a spectrum of engagement with cities that spans to community action and governance. Residential location selection, social tie formation, youth development, urban planning, and other social functions take place across urban environments, building human intelligence through action and learning. AI, whether deployed to ease individual burdens or address broader societal objectives, is likely to function at the expense of human intelligence-building practices. Is this a problem? Can we find other venues, whether virtual realms or other aspects of our lives, to substitute for the cognitively active facets of urban life? Perhaps, but this intelligence will likely require intentionality around experiential learning and equitably providing opportunities to develop empowering, actionable intelligence that urban life once provided. In the meantime, we need to think clearly about AI in cities, and how it may impact our own urban intelligence.[1 2 3]

ENDNOTES

1 Batty, M. (2013) Big Data, Smart Cities and City Planning. Dialogues in Human Geography 3, no. 3: 274–79.

2 Mondschein, A. & Moga, S.T. (2018) New Directions in Cognitive-Environmental Research. Journal of the American Planning Association 84, no. 3–4: 263–75.

3 Schiller, D., Eichenbaum, H., Buffalo, E. A., Davachi, L., Foster, D. J., Leutgeb, S. & Ranganath, C. Memory and Space: Towards an Understanding of the Cognitive Map. The Journal of Neuroscience 35, no. 41 (2015): 13904–11.

IMAGINE
Saying "No"

MC Forelle

Department of Engineering and Society
University of Virginia

The "Reimagining AI" workshop was a fascinating event, mostly due to the incredible breadth of disciplines, experiences, and priorities that the participants brought to the room. The space was full of hope, of ideas for how artificial intelligence might be shaped, guided, harnessed for the greater good. However, the moment that has stayed with me most since the event was, when given the prompt "If you could do one thing to address the problems with AI right now, what would you do," Anne Pasek (assistant professor of Media, Culture, and the Environment at Trent University in Ontario, Canada) replied: "I would stop all construction of new data centers for a year." It was the most ambitious answer anyone gave and, frankly, the most compelling. Why should we not set our sights that high?

In my field, science and technology studies, the apparent inevitability of "advanced" technologies like artificial intelligence is understood as technological determinism, "the idea that technology develops as the sole result of an internal dynamic, and then, unmediated by any other influence, molds society to fit its patterns."[1] This pattern of thought sublimates human thought, action, and meaning to the supposedly natural advancement of technology – people are powerless to resist the forward march of progress, it declares. Today, we see technological deterministic thinking in basically all prevailing discourses about artificial intelligence, even those that seek to mitigate its harms: from the assertions that AI was bound to be developed someday, so it might as well be developed "ethically"; to the claims that, now that it's here, you can't put that toothpaste back in the tube so we might as well figure out how to use it for good.

I reject this premise on its face. There is no reason why theory, scholarship, or advocacy should always already be yielding to the technological determinism of AI's evangelists. If they can start from the premise that AI is inevitable, then we should start from the premise that it is not. If they demand that we demonstrate that AI does real harm before it can be regulated, then we should demand that they demonstrate it produces no harm before it can be distributed. If they want to build new data centers to support the compute requirements of their new AI systems, then we should demand

they detail precisely how this can be done without causing damage before we let them.

It was refreshing to share space with other scholars who agreed that we can be ambitious in our rejection of technological determinism. Together we brainstormed how we might make the case for this ambition, how to identify the roadblocks in front of it, and how to collectivize and organize power behind it. We discussed efforts already underway in academia,[2] in advocacy,[3] and (until recently) in government.[4]

It felt important, and still does, to remember that there is always the possibility of resistance and refusal. When she received her National Book Foundation Medal in 2014, renowned science fiction author Ursula Le Guin famously chided the audience about the power that profit had over the book industry, telling them, "We live in capitalism, its power seems inescapable – but then, so did the divine right of kings. Any human power can be resisted and changed by human beings."[5] I hear, as well, that we live in technological determinism, so the power of technology seems inescapable. But any human-made technology can – and, now more than ever, must – be resisted and changed by human beings.

ENDNOTES

1 Winner, L. (1980). Do Artifacts Have Politics? Daedalus 109, no. 1: 122.

2 Edwards, D., Cooper, Z. G. T., & Hogan, M., (2024). The Making of Critical Data Center Studies. Convergence. https://doi.org/10.1177/13548565231224157 / Kneese, T. & Young, M. (2024). Carbon Emissions in the Tailpipe of Generative AI. Harvard Data Science Review, no. 5 Special Issue https://doi.org/10.1162/99608f92.fbdf6128.

3 Data & Society. (2024). Climate, Technology, and Justice. Data & Society. https://datasociety.net/research/climate-technology-and-justice/.

4 Strickland, E. (2023). What You Need to Know About Biden's Sweeping AI Order - IEEE Spectrum. IEEE Spectrum. https://spectrum.ieee.org/biden-ai-executive-order.

5 Le Guin, U. K. (2014) National Book Foundation Medal: Ursula's Acceptance Speech, Speech in Acceptance of the National Book Foundation Medal for Distinguished Contribution to American Letters https://www.ursulakleguin.com/nbf-medal.

Creating for an Alternative AI Future

04.

CULTURAL POLICY MODELS AND *Potential Applications on AI Governance*

Yingchong Wang

School of Data Science
University of Virginia

The emergence of artificial intelligence in creative and cultural domains presents unique regulatory challenges that extend beyond traditional technology governance frameworks. Cultural policies reflect fundamental values about governmental operations and intended objectives,[1] making these models particularly relevant for understanding AI's role in creativity and cultural development. Cultural policy frameworks are especially appropriate for analyzing AI governance because they address similar core tensions: balancing artistic excellence with democratic access, preservation with innovation, and national identity with global exchange.[2] This theoretical lens of cultural policy models helps identify critical preconditions for designing responsible AI governance in cultural domains.

The Facilitator Model

This model, exemplified by the United States' approach to cultural policy, suggests a market-oriented approach to regulation through tax incentives and indirect support. According to Mulcahy (2006), this involves governmental strategies that promote culture through leveraging private patronage through tax policies. Applied to AI governance, this approach would emphasize creating favorable conditions for responsible AI development through market mechanisms while maintaining light regulatory oversight.

The Patron Model

This model, characterized by the arm's length principle through arts councils, provides insights for establishing independent oversight bodies in creative domains. Hillman-Chartrand and McCaughey (1989) demonstrate how this approach effectively maintains professional standards while preserving independence from both governmental and commercial interests.[3] In the context of AI governance,

this model suggests establishing autonomous institutions that can evaluate and guide AI's cultural impact while ensuring public benefit without direct political intervention.

The Architect Model

The social-democratic approach to cultural policy, prevalent in Nordic countries, integrates cultural development within comprehensive social welfare objectives. Mulcahy (2006) illustrates how these systems prioritize accessibility, sustainability, and representativeness in cultural development. Applied to AI governance, this model suggests embedding AI development within broader social policies to ensure these technologies serve community needs and promote cultural democracy while maintaining high standards of excellence.

The Engineer Model

The Engineer model, historically characterized by direct state ownership and control of cultural production means, demonstrates potential pitfalls in over-regulation of creative domains. As both Mulcahy (2006) and Hillman-Chartrand and McCaughey (1989) observe, this approach often resulted in constrained artistic freedom and the subordination of cultural expression to political objectives. This model's limitations are particularly instructive for AI governance, warning against excessive governmental control that could stifle innovation and restrict creative freedom. Instead, its historical challenges suggest the importance of developing frameworks that protect creative autonomy while establishing necessary baseline standards for responsible AI development.

Implications

1. Understanding Cultural Policy Landscapes

For effective AI governance in creative domains, policymakers and scholars must first thoroughly understand their existing cultural policy frameworks. Different nations and regions operate under distinct models - from the market-oriented Facilitator approach to the more socially integrated Architect model. This understanding is crucial for developing AI governance mechanisms that can be effectively integrated into existing institutional structures while addressing new challenges posed by AI technologies. Rather than creating entirely new frameworks, policy development could build upon established cultural policy foundations while adapting to emerging technological realities.

2. Leveraging AI as a Catalyst for Policy Evolution

The emergence of AI in creative and cultural domains presents an opportunity to modernize cultural policy approaches. The complex interplay between AI, creativity, and cultural development requires a more nuanced regulatory framework that balances multiple objectives: market sensitivity, social responsibility, necessary oversight, and democratic access. This balance is particularly crucial as AI technologies challenge traditional notions of creativity, authorship, and cultural production. Cultural policy models can evolve to address these new dynamics while maintaining their fundamental role in protecting and promoting cultural diversity.

3. Facilitating Cross-sectoral Dialogue

The integration of AI into cultural and creative sectors necessitates enhanced dialogue between traditionally separated policy domains. Technology policy experts, cultural administrators, artists, and AI developers must collaborate to develop governance frameworks that are both technologically informed and culturally sensitive. This cross-sectoral approach can help identify potential synergies and conflicts between technological innovation and cultural preservation, leading to more effective and balanced policy solutions.

The examination of cultural policy models offers valuable insights for developing AI governance in creative domains. These established models demonstrate how different societies have balanced competing priorities in cultural development. Understanding these policy approaches becomes essential for developing frameworks that effectively integrate technological innovation while preserving cultural values. This historical perspective, combined with a contemporary understanding of AI's potential and challenges, provides a foundation for thoughtful policy development that can serve both technological advancement and creativity development in the digital age.

ENDNOTES

1 Mulcahy, K. V. (2006). Cultural policy: Definitions and theoretical approaches. The journal of arts management, law, and society, 35(4), 319-330.

2 Matarasso, F., & Landry, C. (1999). Balancing act: twenty-one strategic dilemmas in cultural policy (Vol. 4). Council of Europe.

3 Hillman-Chartrand, H., & McCaughey, C. (1989). The Arm's Length Principle and the Arts: an International Perspective –Past, present and future. In M. Cummings Jr. & J. M. D. Schuster (Eds.), Who's to Pay For The Arts? The International Search for Models of Arts Support (Vol. 3, pp. 43–80)

REIMAGINING AI, *Black Heritage Tourism, and Preservation*

Andrea R. Roberts

Center for Cultural Landscapes, School of Architecture
University of Virginia

Artificial Intelligence (AI) has become a new arbiter of power, specifically computing power within the sphere of virtual and participatory public heritage curation. AI has become a compass and a proxy for legitimacy because for-profit technology firms can devise algorithms that determine what we see when we wander and explore in places without visible built heritage. When contemplating heritage sites in the United States, our current preservation frameworks and design lexicon situate the work of educating the public as being about sites, buildings, or landmarks but not always landscapes. So much of the technology applied to participatory heritage, museum, and house museum websites not only recreates sense of place through 3D or virtual realities but also gives access to tools that can curate a tour. In this way, these sites take from available public knowledge possibilities and are still driven by the priorities and potentials of the present Western aesthetics and preservation regulatory frameworks. One of the most significant conversations witnessed during the Reimagining AI workshop was a discussion of AI tools that curate tours or experiences of heritage landscapes and sites. AI that could make new meanings out of available data can open portals of awareness and possibility as well. Consider the work of design professional and educator Curry Hacket, who has created countless AI generated graphics of Black life offering the possibilities of what might have been or may be in another parallel existence, leveraging symbols of Black heritage from around the world.

The work of the UVA Center for Cultural Landscapes is one concerned with places, specifically largely invisible sites and landscapes for which no public record or map exists. One would argue that the creation of crowdsourcing maps, like The Texas Freedom Colonies Project Atlas and Study and the working Arc of Enslaved Communities Map, might bring into the public domain more accurate or inclusive information about place

meaning and heritage. However, might the context and meaning be lost as the data forms are disaggregated and new relationalities are formed that may reveal more about place vulnerabilities but concurrently erase some of the meanings of these places hidden in the complex archival materials, documents, and responses to survey prompts?

In other instances, AI offers heritage scholars and preservationists an opportunity to create dialogues between archives. Spatially, platforms like Esri's Arc GIS integrate artificial intelligence (AI) into its tools to help users analyze spatial data, make predictions, and perform other tasks. The capabilities are present, but there is little discussion about risks that surface in the world of heritage because it is disconnected from privacy and power issues emerging when considering who may have access or control over publicly available place data. While much is available in the public domain, what are the connections AI makes that lead to the over-exposure of African American rural land, for example, to unscrupulous vultures who seek opportunities to attain land with unstable titles? There is a direct correlation, for example, between incidences of heir property and the historic Black settlements whose history and locations remain largely hidden because of the existence of data outside the public domain. While AI can help quickly locate settlements by using AI to calculate concentrations of population, land acquisition, and cultural anchor institutions, AI cannot presuppose what the origin stories, circulation patterns, and hidden meanings are embedded in private memories. The tension here is between how we determine access and control over data that present us with opportunities to piece together hidden heritage while also examining the risk of those data elements getting into the wrong hands. This is partly related to how much historic preservation and heritage conservation fail to consider development and environmental risk until it is too late. Thus, reimagining AI means reimagining the barriers between fields of study, professions, and constructions of things we take for granted, such as what makes a place a place and what are the hidden notions of place in the Black imagination that might also be unleashed by AI.

WHEN I CLOSE MY EYES, *I Envision AI as a Collaborator*

Maria Villanueva

University of Virginia

Artificial intelligence has become an inevitable companion in digital creation, transforming how we make images, audio, and video. But as these technologies evolve, we must carefully consider their role in our creative process as artists and creators. AI doesn't create from nothing. It learns by consuming human history, drawing patterns from our collective creativity. When we generate an image or analyze data, we're not witnessing pure machine invention, we are witnessing a sophisticated reflection of our accumulated knowledge. We are at a crossroads where we move from imagining into fully recognizing and taking a path where we claim our seat at the table and become part of the decision-making process as large contributors to this knowledge. We are facing a challenge and an opportunity much like the birth of social media,[1] and how it needs to be collaborative in order to exist and persist after its novelty wears out.

For marketing and graphic design, AI offers relief from tedious tasks. Background removal, image tracing, and complex editing now happen with unprecedented speed. Large Language Models like ChatGPT enable us refine grammar and structure, but these tools should not be replacements for creating the work or replacing an original idea. The real challenge isn't technological capability but understanding our relationship with the tools. These tools are aids, not substitutes for original thought and creative invention. Rather than replacing human creativity, AI can serve as a strategic tool that streamlines the technical process, freeing artists to focus on the nuanced, emotional core of the work. This shift, however, raises important questions about value. Under capitalism, the worth of art is often tied to the labor invested in its creation. The more time and effort that goes into making something, the higher its perceived value. As AI reduces the labor involved, it challenges this conventional mindset. If an image or a piece of music can be produced quickly and effortlessly, does that diminish its worth? Or does it redefine what we consider valuable in art?

Not all AI is equally accessible. Sophisticated technologies remain hidden, available primarily to those with significant resources. Successful artists like Refik Anadol showcase AI's potential with large budgets and a large collaborative human team, but they represent a narrow slice of the creative experience. We must consider that not every artist has access to advanced technologies or funding. The same struggles remain for artists having to navigate complex proposal processes and access to new tools with limited resources. Our fundamental challenge is to critically understand these tools while democratizing their availability.

Collective engagement is key in shaping the path for collaborative AI. We need artists actively contributing to policy discussions to ensure transparency in training practices, fair use of data, and ethical standards. This is not about resisting technology, but making sure it enhances rather than undermines human creativity. We must create paths that resist easy algorithmic reproduction-developing practices that draw deeply from personal narrative, experimental techniques, and complex human intuition. Our vision should be a relationship with AI as a collaborative partner rather than a competitor. We are not passive recipients of technological change; we are active participants in its unfolding and design. Researchers have questioned "How can we systematically design machines as teammates in a human centric way?"[2]

The future is not predetermined. It will be shaped by those who engage critically, ask difficult questions, those who insist on being heard. In this digital age we, as creators must claim our place in shaping the future of AI collaboration. To do so, we must ensure that AI remains a tool that amplifies human creativity rather than a force that diminishes it. This means advocating for respect and recognition of the work and data that form its foundation-our own art, writing, and shared human history. AI would not exist without these contributions, and our value should be reflected not just in acknowledgment but in fair treatment and ethical use of our work. The worth of art goes beyond the labor it involves; it is about the emotion, vision, and human experience it carries. The stories we tell and the art we create possess depth and complexity that algorithms cannot replicate. To truly thrive in this new age, we must make sure that our creativity is not seen as a mere dataset to be mined but as a testament to the human spirit.

ENDNOTES

1 Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons, 53(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

2 Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. Information and Management, 57(2). https://doi.org/10.1016/j.im.2019.103174

REFRAMING AI *with the Digital Humanities*

Amanda Wyatt Visconti

Scholars' Lab, University of Virginia
Association for Computers & the Humanities
Digital Humanities Slack

I direct a digital humanities (DH) center[1] that's focused on ethical, creative experimentation at the intersections of humanities, culture, and tech since 2006. A common definition of DH encompasses both using digital methods (such as coding and mapping) to explore humanities research questions (such as concerns of history, culture, and art); and asking humanities-fueled questions about technology (such as ethical design review of tools like specific instances of AI). I always add a third core feature of DH: a set of socially just values and community practices around labor, credit, design, collaboration, inclusion, and scholarly communication, inseparable from best-practice DH.

I write this piece as someone with expertise in applicable DH subareas-research programming, digital scholarly design, and the ethical review of digital tools and interfaces- but not as someone with particular experience related to ML, LLMs, or other "AI" knowledges. A field of new and rapidly evolving tools means true expertise in the capabilities and design of AI is rare; often we are either talking about secondhand experiences of these tools (e.g. "Microsoft Co-Pilot let me xyz") or about AI as a shorthand for desired computing capabilities, unfounded on familiarity with current research papers or understanding of codebases. (A value-neutral claim: science fiction authors without technical skillsets have helped us imagine and later create).

Convergence on the term "data science" has both inspired new kinds of work, and elided contributions of the significantly overlapping field of library and information studies. Similarly, "AI" as the shorthand for the last few years' significant steps forward in ML (and LLMs in particular) obscures the work of the digital humanities and related critical digital research and design fields such as Science and Technology Studies (STS). When we use the term "AI", it's tempting to frame our

conversations as around a wholly new thing, focusing on longer-term technical aspirations uninhibited by practical considerations of direct audience needs, community impacts, resources. While that's not necessarily a bad way to fuel technological creativity, it's too often the only way popular conversations around AI proceed. In one research blog post[2] exploring the moral and emotional dimensions of technological design, L.M. Sacasas lists 41 questions we can ask when designing technologies, from "What sort of person will the use of this technology make of me?" to "Can I be held responsible for the actions which this technology empowers? Would I feel better if I couldn't?" We don't need to reinvent digital design ethics for AI-we've already got the approaches we need (though those can always be improved).

I frame "AI" as just being *code*: a milestone advancement in code, sure, yet still part of the long history of programming and its packagings (codebase, repo, library, plugin…). Thinking of AI as part of this continuity of codework, instead of as its own unique thing, makes it easier to remember we already have years of experience designing and analyzing the ethics and societal impacts of this work-so much so that I've started assuming people who say "LLM" or "ML" rather than "AI" when starting conversations are more likely to be conversant with the specifics of current AI tech, as well as its ethical implications. The terms we use for our work and scholarly conversations are strategic: matching the language of current funding opportunities, job ads. We've seen similar technologically-vague popularizing on terms with past convergences of tech interest too, including massive open online courses (MOOCs), "big data", and the move from "humanities computing" to the more mainstreamed "digital humanities".

Digital humanities centers like our Scholars' Lab offer decades of careful, critical work evaluating existing tools, contributing to open-source libraries, and coding and designing technology in-house-all founded on humanities skills related to history, ethics, narrative, and more strengths necessary to generative critique and design of beneficial tech. Some of the more interesting LLM-fueled DH work I've seen in the past couple years has involved an AI first- or second-pass at a task, followed by verification by humans-for situations where the verification step is neither more onerous nor more error-prone than a human-only workflow. For example:

» the Marshall Project had humans pull out interesting text from policies banning books in state prisons, used AI to generate useful summaries of these, then had humans check those summaries for accuracy[3]

» Scholars Ryan Cordell and Sarah Bull[4] tested Chat GPT's utility in classifying genres of historical newspaper and literary text from dirty OCR and without training data, and in OCR cleanup, with promising results

» My Scholars' Lab colleague Shane Lin has been exploring AI applications for OCRing text not well-supported by current tools, such as writing in right-to-left scripts

» Archaeologists restoring the HMS Victory[5] applied an AI-based algorithm to match very high-resolution, high-detailed images

stored in different locations to areas of a 3D model of the ship

One of DH's strengths has been its focus on shared methods and tools across disciplines, regardless of differences in content and disciplinary priorities, with practitioners regularly attending interdisciplinary conferences (especially unusual within the humanities) and discussing overlapping applications of tools across research fields. DH experts also prioritize non-content-agnostic conversations, prompted by the frequency with which we borrow and build on tools created for non-academic uses. For example, past Scholars' Lab DH Fellow Ethan Reed found utility in adapting a sentiment analysis tool from outside his field to exploring the emotions in Black Arts Poetry works, but also spent a significant portion of his research writing[6] critiquing the biased results based on the different language of sentiment in the tool's Rotten Tomatoes training dataset. (ML training sets are an easy locus for black boxing biases, context, and creator and laborer credit-similar to known issues with text digitization work, as explored by Aliza Elkin's gorgeous *Hand Job* zine series[7] capturing Google Books scans that accidentally caught the often non-white, female or non-gender-conforming hands of the hidden people doing the digitizing.)

We already know where to focus to produce more beneficial, less harmful, creative digital tools: social justice. At the *Reimagining AI* roundtable, my table's consensus was that issues of power and bias are key not just to reducing ML harms, but to imagining and harnessing positive potential. Key areas of concern included climate terrorism (e.g. reducing the energy costs of data centers), racism (e.g. disproportionate negative impacts on BIPoC compounding existing economic, labor, and police violence threats), human rights (e.g. provision of a universal basic income easing concerns about areas ML may beneficially offset human labor), and intertwined ableist and computing access issues (e.g. AI search-result "slop" is terrible for screen readers, low-bandwidth internet browsing). In our existing scholarly fields and advocacy goals, where are current gaps in terms of abilities, resources, scale, efficiencies, audiences, ethics, and impacts? After identifying those major needs, we're better positioned to explore how LLMs might do good or ill.

ENDNOTES

1 The Scholars' Lab at University of Virginia; see https://scholarslab.org/

2 Sacasas, L. M. (2024). Do Artifacts Have Ethics? L.M. Sacasas. https://thefrailestthing.com/2014/11/29/do-artifacts-have-ethics/

3 Calderón, A. R. (2023). Decoding Bureaucracy. Medium. https://generative-ai-newsroom.com/decoding-bureaucracy-5b0c1411171

4 See: https://bsky.app/profile/sarahebull.bsky.social/post/3kg2vhh45tm2y

5 National Museum of the Royal Navy. (2024). Archaeology meets AI in ground-breaking collaboration between The

National Museum of the Royal Navy and the University of Southampton. https://www.nmrn.org.uk/news/archaeology-meets-ai-ground-breaking-collaboration-between-national-museum-royal-navy-and

6 See https://scholarslab.lib.virginia.edu/people/ethan-reed for related research blog essays.

7 Zines are free to download at http://alizaelk.in/digitize

LET'S MAKE USE OF FREE AND OPEN SOURCE AI: *Civic and Public Interest Technology*

Jonathan Kropko

School of Data Science
University of Virginia

The modern discourse surrounding AI emphasizes the money and resources that are needed to develop and use it. But it can be easy to forget that there are many AI models and tools that are either free and open source, or free to use for smaller applications. While tech workers are deservedly well paid, there are also many tech-skilled individuals who seek to volunteer their time and effort for good causes. Let's use this to do good things in the world.

Tech-based volunteerism exists all over the world to bolster civic society and benefit underserved communities. Nascent movements such as civic tech, public interest tech, and tech for good are supported by NGOs and nonprofits such as Code for America, the New America Foundation, and the Alliance of Civic Technologists. These groups working at local and national levels have developed projects such as GetCalFresh,[1] an easy and accessible guide that helps people in California apply for the SNAP program, Court Bots[2] that remind people of their court dates and help them navigate to the right locations to help avoid bench warrants, fines, and jail time, and projects to help wildlife hospitals understand where most vehicle-animal collisions happen to advocate for land bridges across highways.

Organizations that work to benefit vulnerable communities or assist with environmental conservation are rarely well-funded enough to afford IT departments, data science teams, on premises computing clusters, and other tools that enable large enterprises to take advantage of AI, data, and tech. But the kinds of volunteer tech projects described above use free and open source software along with low- or no-cost hosting and deployment strategies to provide the benefit of tech to the organizations that otherwise would not be able to use it. Many of the volunteers do so out of

altruism and a sense of community obligation, but there are many self-interested reasons why skilled technicians volunteer: these projects offer a chance to network effectively and to learn new coding languages and methods, and they make for impressive pieces in a professional portfolio.

All of the ingredients are readily available for civic tech to become even more prominent component of modern tech as it interacts with society. But the missing piece is organizational, and that's where academia can lead. By adopting a scientific approach to civic tech via peer review and dissemination of knowledge, we can have a profound impact by replicating successful projects in other locations: every state and town could use a Court Bot, for example. As members of the academy, we should work to change the standards for promotion and tenure to include volunteer work that uses advanced STEM skill, and we should create journals and conferences to encourage, adjudicate, and popularize this work. If we do, then we can foster a great democratizing force within tech and AI.

ENDNOTES

1 See: https://www.getcalfresh.org/

2 See: https://github.com/CivicTechAtlanta/georgia-courtbot

REIMAGINING AI AS CULTURAL INFRASTRUCTURE: *Bridging Heritage, Urban Life, and Digital Responsibility*

Martina Massari

Department of Architecture
University of Bologna

Danila Longo

Department of Architecture
University of Bologna

Artificial Intelligence (AI) permeates academic, policy and economic debate. The United States is leading the critical discussion on its multilayered risks and perspectives, and yet, AI's cultural dimension remains underexplored. In contrast, Europe - and Italy in particular – has taken a slower, more cautious approach, advocating for AI that prioritize societal well-being over purely economic objectives. Europe has long emphasized the integration of digitalization with cultural policy, producing guidelines to ensure technology aligns with values of trust, creativity, adaptability, and social responsibility.[1] The European Union's AI Act exemplifies this vision, aiming for a human-centric and ethical use of AI.[2] However, the Act rarely addresses culture explicitly - a clear oversight given AI's reliance on cultural inputs such as texts, images, and values that shape its design and application. This omission expresses a gap in AI policy: the failure to consider cultural preservation, diversity, and local traditions, which often resist easy incorporation into predominantly techno-optimistic framework.

As scholars in urban technology, it is our responsibility to delve deeper into AI's role in cultural heritage and urban environments. We must explore and understand the social responsibilities associated with AI-driven infrastructures and test frameworks that support cultural diversity, encourage public participation, and uphold local governance.

AI in Urban Spaces: beyond a Passive Tool

AI offers numerous benefits for the valorization of cultural heritage. It enhances access to cultural resources, supports research, creates "memory insurance policies" for artifacts at risk of destruction, and adapts cultural content to evolving social contexts, fostering cultural continuity. However, as AI increasingly shapes cultural narratives, critical questions about accountability arise. Who bears responsibility when algorithmic decisions exacerbate social inequalities? Who is accountable for potential cultural polarization? Addressing these issues demands a clear framework for accountability in AI applications, particularly in urban environments where AI is evolving from a passive tool to an active agent of change.

Embedded AI systems collect data and inform public sphere dynamics, transforming cities into experimental grounds for AI-human interaction[3] without clearly defined boundaries for action. Urban AI plays a vastly more influential role, directly intervening in shaping social dynamics, cultural expressions, and even political engagement. As an urban agent, AI's influence over urban space can alter local identities and public sphere outcomes. AI-driven digital platforms are not neutral entities; they reflect the ideologies and power structures governing their design and deployment. Acting as political actors, these platforms influence information flows, public opinion, and even social movements, sometimes surpassing traditional governance in their societal impact. As a consequence, such systems risk exacerbating socio-political divides, privileging certain voices while marginalizing others. Data-driven urban regeneration can overlook local structural vulnerabilities, imposing one-size-fits-all models that fail to consider unique socio-political contexts. This evidence calls for frameworks that foster self-governance and draw from deliberative democracy encouraging local control over cultural narratives.

Visual tools, such as interactive and locally contextualized AI-generated images, can aid this effort by more effectively conveying urban and cultural nuances than abstract concepts. However, standard AI-generated imagery often risks homogenizing cultural representation, depicting similar architectural and commercial patterns that may not reflect local diversity. Here, an art-science approach can provide greater nuance. By involving artists in AI development, more inclusive visual representations of heritage can emerge - images that resonate with diverse cultural audiences. Artistic collaborations can introduce symbolic choices that balance simplicity and detail, avoiding the distractions of excessive digital realism. This approach can make AI-generated content both familiar and universal, enriching public understanding of local heritage while preserving its distinctiveness.

Civic Digital Twin: a High-Stakes Project for Local Governance

To counteract these risks while leveraging an art-science collaboration, a possible model is the "Civic Digital Twin". As demonstrated in Bologna (Italy) it is a project that combines local stakeholder collaboration to integrate citizen engagement into AI-driven urban planning. This project emphasizes two dimensions: focusing on the sociotechnical processes of reproducing cities rather than just technical models and recognizing the uniqueness of local urban cultural heritage instead of treating cities as abstract entities. This project considers AI as a "boundary object," operating across multiple contexts and serving as a site for experimentation and cross-disciplinary collaboration. By exploring alternative possibilities in both research and practice, the project shifts AI applications from passive observation to active societal participation.

The Civic Digital Twin sees AI as a form of relational infrastructure that can either exploit or bridge cultural divides, influencing who is included or excluded in public life. This digital platform reimagines urban landscapes, framing AI as an orientation tool-a "navigation map" for public life-that can critically evaluate its own role in promoting equitable access and civic representation. Cultivating this future literacy will transform uncertainty into a resource, enabling communities to co-create inclusive and resilient urban ecosystems.

By positioning AI as an essential and inclusive public infrastructure, society can engage in conversations that not only support technological advancement but also prioritize cultural diversity, civic engagement, and social responsibility.

ENDNOTES

1 Floridi, L. et al. (2018). "AI4People—An Ethical Framework for a Good AI Society." Minds and Machines.

2 European Commission (2021). Artificial intelligence act. European Parliament: European Parliamentary Research Service. Available at: https://www.iisf.ie/files/UserFiles/cybersecurity-legislation-ireland/EU-AI-Act.pdf

3 Floridi, L. et al. (2018). AI4People—An Ethical Framework for a Good AI Society. Minds and Machines.

REIMAGINING AI: *Equity, Creativity, and Governance in the Nighttime Economy*

Raheem Manning

Nighttime Economy & Business Development
City of Philadelphia

In October 2024, I had the privilege of participating in the "Reimagining AI for Environmental Justice and Creativity" panel at the University of Virginia. This gathering of multidisciplinary scholars, policymakers, activists, and practitioners delved into the intricate dynamics of power and trust in technology and governance, especially concerning artificial intelligence (AI). Reflecting on the discussions, I am compelled to explore the intersection of AI, environmental justice, and creativity, emphasizing the imperative of inclusive and equitable technological development. My perspective on these topics is shaped by my role as the Director of Nighttime Economy and Business Development for the City of Philadelphia, where I work at the nexus of policy, economic development, and community engagement.

AI and Environmental Justice: A Policy Perspective

AI has the potential to revolutionize environmental monitoring, policymaking, and advocacy. However, its deployment must be approached with caution to prevent exacerbating existing inequalities. Marginalized communities often bear the brunt of environmental degradation and may be further disadvantaged by AI systems that do not consider their unique contexts. For instance, AI-driven environmental policies that rely solely on data without community engagement can overlook localized knowledge and needs, leading to ineffective or even harmful outcomes.

During the panel, one of my fellow panelists eloquently discussed the risks of AI reinforcing historical patterns of environmental racism, highlighting how predictive models used in urban planning often replicate exclusionary policies. Building on this, I emphasized the necessity of community-led AI initiatives,

drawing from my experience in local government. In my work, I have seen firsthand how policies that exclude direct input from affected communities fail to generate trust or effective solutions. Philadelphia's nighttime economy, for example, thrives on dynamic interactions between businesses, policymakers, and residents-an approach that should inform AI-driven policy frameworks for environmental justice.

Creativity, AI, and the Nighttime Economy

AI's role in creative fields is a double-edged sword. On one hand, AI can augment human creativity by providing new tools for expression and innovation. On the other hand, it raises concerns about the commodification and potential devaluation of human artistic endeavors. The panel highlighted the need to balance technological advancement with the preservation of human creativity's intrinsic value.

As someone deeply engaged in the cultural and economic development of Philadelphia's nightlife and creative industries, I brought up the importance of AI in democratizing creative opportunities. One of my fellow panelists explored the ways AI is used in music production and visual arts. Building on their insights, I underscored how AI could serve as a tool for creative equity-helping independent artists and nightlife entrepreneurs gain access to resources traditionally dominated by larger institutions. However, I also stressed that AI must be implemented with safeguards to ensure it does not displace artists or erode the authenticity of cultural expression, particularly for communities whose voices have historically been marginalized.

Building Trust in AI through Governance and Transparency

A recurring theme in our discussion was the critical importance of trust in the deployment of AI technologies. For AI to be a tool for good, it must be transparent, accountable, and aligned with the values of the communities it serves. This involves not only technical robustness but also ethical considerations, such as data privacy, consent, and the mitigation of biases.

From my vantage point as a policy advocate and government official, I emphasized that trust is not just about technical transparency-it is about governance. AI governance should mirror effective public governance: inclusive, community-driven, and responsive to public needs. I shared how Philadelphia is working to implement public safety and economic development programs that require a balance between technological efficiency and public trust. The Liberty Bell Safe Certification Program, for example, is designed to support nightlife businesses in creating safer spaces through training and certification. If AI tools are integrated into these initiatives-whether for crowd management, security, or business analytics-it must be done in ways that reinforce trust, not erode it.

Conclusion: A Vision for Equitable AI

Reflecting on the insights from the panel, it is evident that the intersection of AI, environmental justice, and creativity presents both challenges and opportunities. To harness AI's potential for positive impact, we must adopt a holistic approach that integrates technical innovation with ethical considerations and community engagement.

This means drawing from best practices in policymaking-ensuring that AI governance is as participatory and inclusive as the best examples of civic engagement.

The panel discussion reaffirmed my belief that AI must not be a tool wielded only by technologists and corporations, but one that is co-created with policymakers, artists, and community members. My work in Philadelphia's nighttime economy has shown me that innovation thrives when it is deeply rooted in community engagement. AI must be no different. By fostering interdisciplinary collaboration and a commitment to equity, we can reimagine AI as a force that not only advances technology but also amplifies justice, creativity, and cultural resilience.

REIMAGINING "AI'S" ENVIRONMENTAL *and Sociotechnical Materialities*

Damien P. Williams

UNC Charlotte

There are numerous assumptions bundled into the current thinking around what "artificial intelligence" does and is, and around whether we should even be using it and, if so, how. Those pushing "AI" adoption tend to presuppose it necessarily will be good for something- that it will be useful and solve some problem- without ever defining exactly what that problem might be. Often, we see that there are these pushes towards paradigms of efficiency and ease of work and "rote" tasks being taken off our hands without anyone ever asking the fundamental follow-up question of "...okay but does it actually do any of that?" Relatedly, it's often assumed that "artificial intelligence" will become or will make other things "better" in some nebulous way if only we just keep pushing, just keep building, just keep moving towards the next model of it. If we keep doing that, then eventually, we're assured, "in just ten years," "AI" will turn into the version of itself that will solve all our problems. But this notion that in ten years, "AI" will be embedded in everything and will be inescapable and perfect is something we've been hearing for the past 50 years. This recurrent technosocial paradigm of "AI Summer" and "AI Winter" exists for a reason; these hype-cycles pushing towards automation, neural nets, big data, or algorithms over and over again represent externalities which must be addressed in a deeper way through questions like, "What are the values of the people who push 'AI's' 'inevitability,' and what are their actual goals?" Because, while people might think they mean the same things when they say "AI," or are indicating the same kinds of needs to be met, in truth, we're very often talking past each other. Without a clear understanding of what it is we each and all actually think of as the "good" of "AI" technology- without confronting that question in a very direct and intentional way- different groups will just keep pushing in different directions, and whoever has the

predominant access to and control over the levers of power wins the right to define the problems that "AI" seeks to address. But in many cases, those are problems they and their vision of "AI" helped to create.

Current estimates hold that water consumption increased ~34% in areas where Microsoft and Google placed datacenters for search and "AI," and that every email's worth of text you have an LLM "AI" write consumes a pint of water. Put another way, imagine if every time you composed 150 of your own words, you had to just take out a 16 oz water bottle, fill it up, and dump it in the trash. We're not just talking about water for cooling servers, either. In thermal power plants, you need water to turn into steam to run turbines, and then to cool the systems which do that, as well. So the more energy needed, the more water used in production and cooling. And while many highlight that some systems only use this water once and then release it, even that is a process and a period of capturing that water, both removing the water from use, and potentially trapping and killing organisms living in it. Additionally, the water returned after the "once through" process has a significantly higher temperature than when it started. It should be said that the numbers in this discussion are estimates based on known figures for chip performance, electricity production, and whatever data's been wrenched from "AI" corporations. They're estimated because these companies do not release their *actual* resource consumption numbers.

Further, the data centers that support "AI" are oftentimes built in communities that are already resource scarce, and pulling water from or putting emissions into these communities ensures that "AI's" harms are necessarily disproportionately enacted on the people who can least afford to bear them. Rather than rulemakers just paying lip-service to people's grievances, logging them in a repository somewhere, and making whatever rules they intended to make to begin with, both the creation and regulation of "AI" must be directed by those whom it's most likely to harm. But while marginalized communities absolutely must have meaningful input when it comes to technologies which will be wielded against them, there also has to be a centralized response in the form of some standard-setting body. And, recursively, that standard-setting body will have to be meaningfully responsive to the needs of those most likely to be harmed if said regulations and standards go wrong.

And so, we have to ask our questions: Who is most harmed by current uses of "AI"? What does the energy footprint of a data center *actually* look like? How much water and fossil fuel does it take to run "AI's" servers and their computations? What are their carbon and waste heat emissions? Because the more we dig down on this, the more we truly confront the next questions: Should we be doing "AI" differently? What would it take to build "AI" in a different way? What would it take to power "AI" in a truly renewable way? And what and whom do we even want "AI" to be for? If it helps, you can try to think of it as a game:

First major "AI" firm to use only renewable energy sources, an open source and radical consent model for the collection and use of training data, and a community partnership regulatory process which centers and heeds the needs of the most marginalized, wins.

EXPLORING DISCURSIVE AND DELIBERATIVE *Cartographies for AI and Environmental Justice*

Sergio Guillen Grillo

Department of Engineering and Society
University of Virginia

Policy makers, corporations, and citizens are scrambling globally to make sense of and respond to the expanding impact of artificial intelligence (AI) in countless domains of commercial and social activity. The voracity for natural resource and service infrastructure resulting from current and projected levels of expansion in data center facilities could roll back a significant share of recent achievements in energy system decarbonization and water conservation,[1] and there are calls for more systematic assessment regarding the potential compensating impacts of accelerating energy transition innovations.[2] The vertiginous expansion of AI applications across a vast swath of commercial and public interest domains, as diverse as talent recruitment, justice administration, medical diagnosis, transportation, publishing, entertainment, defense, and political campaigning, suggests that its pace of adoption has far surpassed the social and institutional capabilities to collectively assess their impacts and achieve basic agreements about how to bound, track and regulate them.[3]

This opportunity to be a guest commentator for the Karsh Institute's Reimagining AI panel on AI and Environmental Justice has highlighted vital questions about responsible innovation, justice and environmental reflexivity in the Anthropocene.[4] As an educator in tech ethics and environmental peacebuilding, the realization that struck me most prominently when preparing to join a discussion entwining these concepts was

the discursive haziness that predominates in much of the public discussion about AI in contrast to the much greater discursive clarity that accompanies the notion of Environmental Justice. The applications and envisioned outcomes of AI are increasingly being touted (and sometimes decried) in an expanding number of fields, portraying it as a tool with seemingly unlimited potential but glaringly unspecified contradictions. This point was made extremely eloquently by Damien Williams during the panel's question period:

(...) while we might think we're talking about the same things and the same kinds of needs to be met we're very often talking past each other and one group of people means one thing by what it is they're trying to achieve, and we mean something very totally different. And without a clear understanding of what it is they actually think of as good and as the good of this technology, without being made to confront that question in a very real direct and intentional way, we're just going to keep pushing in different directions (...) [Time stamp 47:00]

In comparison, discussing Environmental Justice refers to a radical, normative project associated with a strongly networked, yet multifaceted, global movement.[5] As a plural network, the Environmental Justice Movement continues to evolve, and at times sees internal contestation of some of its prevailing perspectives, but it can nonetheless be characterized as having a clear collective intent: to confront "the uses of state, social, corporate, and colonial power of [sic] vested in systemic practices around the world that marginalize, disenfranchise, and systemically impose environmental violence by corrupting the systems of relations between humans and the more-than-human world."[6]

Contemplating this contrast, it is troubling that such a consequential collection of tools and agendas continue to both traverse and transform our public sphere while cloaked in such discursive ambiguity. This is not to say that everyone who interacts with AI systems does so ambiguously and without a specific perspective about its social, economic or political implications. Certainly, those actively involved in the development and deployment of AI applications, as well as those engaging critically with it, have substantial insights into the range and social distribution of its benefits and setbacks. However, as highlighted in Damien William's previously cited comment, the notion of AI seems to navigate the general public discourse with very little specificity about what it is meant to accomplish, for whom it will accomplish that, and whom it should be accountable to.

I believe an important element of the work required to reimagine AI through the lenses of environmental justice and creativity involves categorizing the diverse perspectives and agendas relative to AI so they can be recognized and addressed more clearly in public discourse. I refer to this task as *the discursive cartography of AI.* The work of mapping AI discourses could be tackled at two different levels. A broader approach would involve mapping the full range of prevailing discourses about AI. This approach would place AI at the center of an extensive discursive mapping effort. However, focusing through the specific lens of environmental justice, an alternate, more manageable approach would be to

explore where different perspectives about AI are in the environmental discourses landscape. For either of these approaches, the discourse mapping work pioneered by John Dryzek in his 1997 book The Politics of the Earth, now in its fourth edition (2022) can serve as a powerful model.[7] Dryzek categorizes ten main branches, and their associated offshoots, in a tree of environmental discourses that represent the prevailing shared frameworks for apprehending environmental issues and coordinating action in response to them. The four central elements he uses to analyze and classify discourses are: (a) what basic entities are recognized or constructed, (b) what assumptions about natural relationships are embodied, (c) who are identified as agents and what are seen as their motives, and (d) what are the prevailing metaphors and rhetorical devices. The high-level discourses in Dryzek's classification span a broad historical range and include "limits to growth", "administrative rationality", "sustainability", "green radicalism", and "ecological modernization". Specific concepts, policies and sociotechnical artifacts can be examined through the lens of these diverse discourses. Environmental justice, for example, can be understood as a more specific sub-discourse of green radicalism, contesting the omission from the mainstream environmental narrative and movements of marginalized peoples and their disproportionate harm from environmental pollution and degradation. However, environmental justice as a concept has now gained ground in other prominent discourses, such as administrative rationality and even ecological modernization. AI, as a sociotechnical artifact can also be examined through the lens of diverse environmental discourses such as ecological modernization, administrative rationality, democratic pragmatism, and as has been done through this panel, environmental radicalism. This work can be promising as a first step in advancing the discursive cartography of AI.

As a complement to the effort of a discursive cartography, I also propose a parallel effort of deliberative cartography. By deliberative cartography, I mean the mapping of the political landscape and pathways through which AI and democratic governance are mutually impacting one another. Because of this reciprocal effect this mapping task involves at least two different elements. The first element entails determining the sites and processes in the deliberative system[8] through which shared understanding and consequential decisions about AI, oriented towards the public good, need to be achieved. It involves the central questions of "where" and "how" we need to come together to explore the potential benefits and impacts of AI, and work to generate legitimate, responsible, and consequential agreements about its development, deployment, and regulation. The second element relates to assessing both AI's possible contributions to overcoming certain limitations of public deliberation, and its capacity to erode the quality of public discourse. Each of these questions is currently the subject of intense but frequently fragmented discussion, with some focusing on the AI applications' role enhancing deliberation, others looking at the erosion of public discourse, and yet others exploring the roles of deliberative forums for discussing and proposing AI regulation. A thread that connects these three questions can greatly assist democratic renewal scholars

to seek a more integrative engagement with the questions of AI's roles and impacts.

AI is increasingly impacting global economic and ecological systems as well as citizen's everyday lives. This dazing pace of expansion surpasses our current ability to collectively examine its public impacts and make decisions about its regulation. Contemplating the notions of AI and environmental justice side by side reveals the high level of discursive ambiguity that has prevailed in public discussions about AI. A discursive cartography of AI can help make the range of discourses about AI more visible in the public sphere and support more productive communication and public decision-making. A deliberative cartography of AI can help focus our efforts to advance the most constructive and just sites and approaches for inclusive claim-making and accountability in the governance of AI.

ENDNOTES

1 Truby, J. (2024) AI is already killing us. Journal of Internet Law 27(7), SSRN: https://ssrn.com/abstract=4908637

2 Luers, A., Koomey, J., Masanet, E., Gaffney, O., Creutzig, F., Lavista Ferres, J., & Horvitz, E. (2024). Will AI accelerate or delay the race to net-zero emissions? Nature, 628(8009), 718-720. https://doi.org/10.1038/d41586-024-01137-x

3 Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': Regulatory competition for artificial intelligence. Law, Innovation and Technology, 13(1), 57–84. https://doi.org/10.1080/17579961.2021.1898300

4 Dryzek, J.S. & Pickering, J. (2018). The politics of the Anthropocene. Oxford University Press.

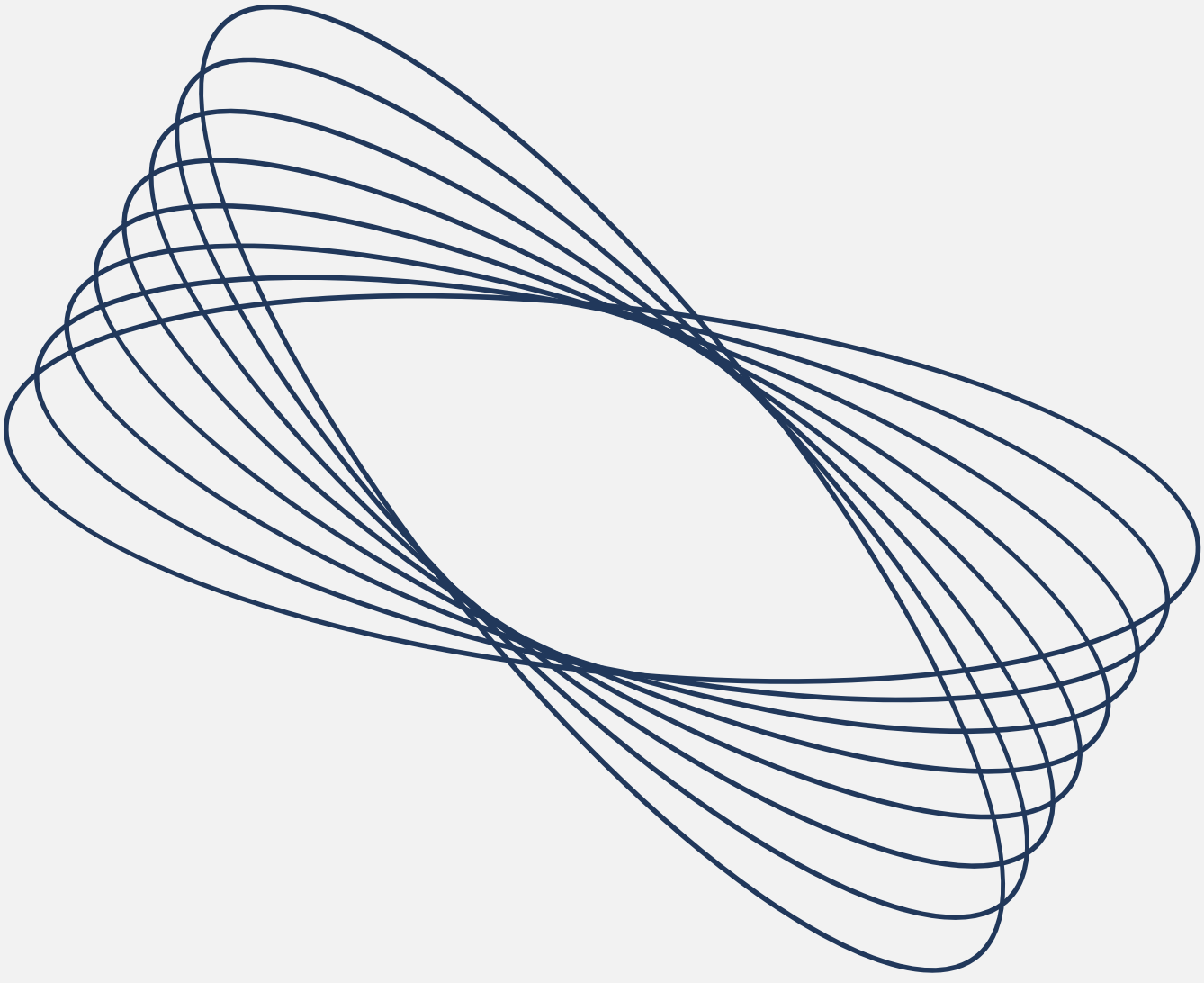
5 Dryzek, J.S. (2022) The Politics of the Earth (4th edition). Oxford University Press.

6 Figueroa, R. M. (2023). Environmental justice. In Hale, B., Light, A., & Lawhon, L. A. (Eds.), The Routledge companion to environmental ethics (pp. 767-782). Routledge.

7 Dryzek, J.S. (2022).

8 Bächtiger, A., Dryzek, J. S., Mansbridge, J., & Warren, M. E. (2018). Deliberative democracy: An introduction. In A. Bächtiger, J. S. Dryzek, J. Mansbridge, & M. E. Warren (Eds.), The Oxford handbook of deliberative democracy (pp. 1-31). Oxford University Press.

EDITORS AND ORGANIZERS



Jess *Reia*

Jess Reia is an Assistant Professor of Data Science and a Faculty co-lead at the Digital Technology for Democracy Lab at the University of Virginia. They are also a Visiting Scholar at Fudan University in Shanghai and a Non-Resident Fellow at the Center for Democracy & Technology in Washington, D.C. Reia works primarily on topics of data justice, human rights and technology policy transnationally. A policymaker by training, they have collaborated with governments for over a decade and conducted research that has been published in four languages. Reia is also a public scholar whose writing and interviews were featured in numerous outlets, including Estadão, Le Devoir and BBC. Before joining UVA, they were appointed Mellon Postdoctoral Researcher at McGill University and worked at the Center for Technology & Society at FGV Law School in Rio de Janeiro.

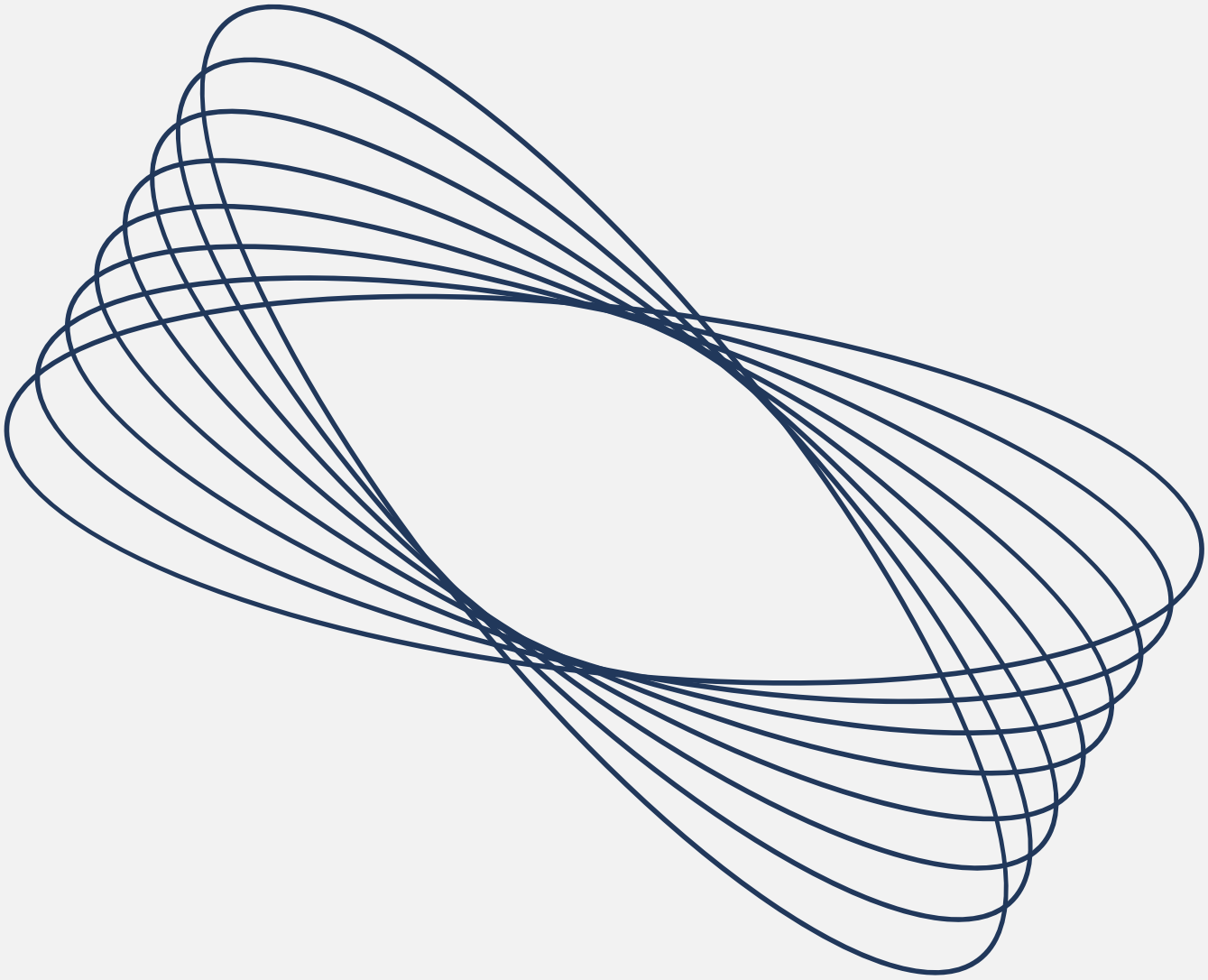
MC *Forelle*

MC Forelle is an assistant professor in Engineering and Society at the University of Virginia School of Engineering and Applied Science. Their work examines the intersection of law, technology, and culture, focusing on materiality, sustainability, and practices of resistance and change. Their current book project studies the paradigm shift in repair presented by software-enabled, autonomous, and connected automotive technologies – particularly with the global rise of internal combustion engine bans – and the failure of industry or government in accounting for the long horizon of automotive devices and infrastructures. They are also exploring new approaches to studies of the supply chains involved in semiconductor manufacturing, use, maintenance, and disposal. Prior to joining UVA, MC was Cornell Presidential Postdoctoral Research Fellow, and received their PhD in Communication from the University of Southern California Annenberg School for Communication and Journalism.

Yingchong *Wang*

Yingchong Wang is an Andrew Mellon Postdoctoral Researcher in the School of Data Science at the University of Virginia. Her work explores creative tourism, responsible AI implementation in cultural organizations, and data-informed policy for creative planning. With academic experience spanning multiple countries and institutions like the Palace Museum in Beijing and Lincoln Center, Yingchong built her expertise through a Ph.D. from Ohio State University's Department of Arts Administration, Education, and Policy, a master's in Arts Management from Carnegie Mellon University, and a B.A. in English from Tsinghua University. Beyond her academic pursuits, she is an accomplished cellist who won first prize at the 2019 Osaka International Music Competition in the China region. Yingchong is currently leading a research project on the ethical implementation of large-scale models in museums at UVA School of Data Science.

Acknowledgements



The collection of essays and the workshop were made possible by the UVA Digital Technology for Democracy Lab at the Karsh Institute of Democracy, the School of Data Science and the Department of Engineering and Society, with the generous support of the UVA Environmental Institute.

We would like to say thank you to Jane Kulow, Laurent Dubois, Amy Tackitt, Vivian Feggans, Julie Lasseter and Jonah Fogel for their help from the moment we started brainstorming this project to the publication of these essays.

Endless gratitude to all the participants and speakers who shared their time, energy and expertise with us during the two-day workshop and on these pages. This work would not have been possible without Anuti Shah, Rachel Leach, Sam Kane and Siobhán Loughney, who were research assistants and collaborators in this project.

A special thanks to the designer Nela Díaz, who made our ideas come to life in such a beautiful way.

And thank you to the Karsh Institute of Democracy for hosting the workshop.