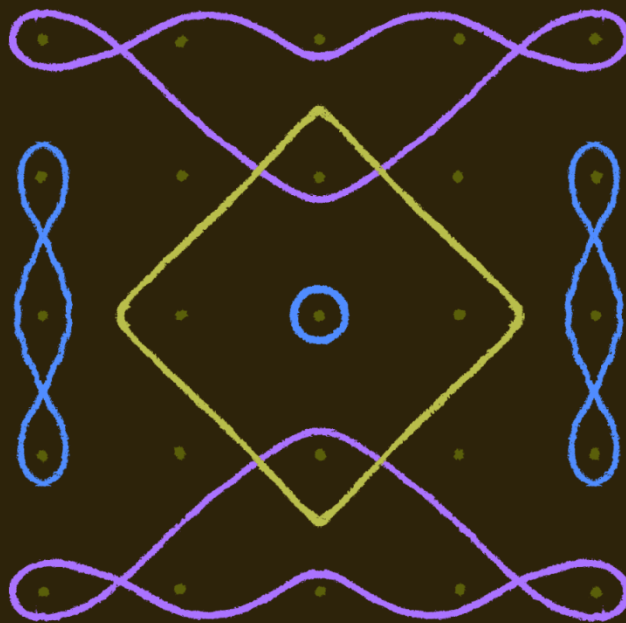


Reframing Impact:
AI Summit 2026

AI for Good

Abeba Birhane

February 2026



AINOW



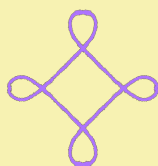
aapti institute



THE MAYBE

This piece is part of [Reframing Impact](#), a collaboration between AI Now Institute, Aapti Institute, and The Maybe. In this series we bring together a wide network of advocates, builders, and thinkers from around the world to draw attention to the limitations of the current discourse around AI, and to forge the conversations we want to have.

In the run-up to the 2026 India AI Impact Summit, each piece addresses a field-defining topic in AI and governance. Composed of interview excerpts, the pieces are organized around a frame (analysis and critique of dominant narratives) and a reframe (provocations toward alternative, people-centered futures).



Abeba Birhane founded and leads the AI Accountability Lab (AIAL) at Trinity College Dublin, where she is also an assistant professor of AI at the School of Computer Science and Statistics. Her research focuses on AI accountability. She formerly served on the United Nations Secretary-General's AI Advisory Body and currently serves on the AI Advisory Council in Ireland.

In this conversation, Birhane critiques the “AI for good” framing embraced by industry and international organizations for its technocratic treatment of complex political and social issues. Such framing, she argues, puts a shiny veneer on bad data, extractive and exploitative practices, and weak evidence. Instead of pouring resources into Big Tech-driven models, she urges us to support smaller, community-based efforts that actually deliver—doing work that serves communities without making grandiose claims. She proposes that governments move away from uncritical adoption of AI via “AI for social good” initiatives and instead demand sound, empirical evidence for claims.

Following is a lightly edited transcript of the conversation.

FRAME: By taking a technosolutionist approach to complex, layered issues, the current “AI for good” conversation misses the point and obscures the “bad” within the AI industry.

Building AI for good without addressing sociopolitical issues and data challenges is like “building a palace with rotting wood.”

Oftentimes the idea [behind AI for social good] is to extend AI tools to solve complex socioeconomic, political questions. Fundamentally, these are not questions that can be solved by AI or any other technology. The UN’s Sustainable Development Goals (SDGs)—things like eliminating hunger, removing gender violence, or expanding access to education—are inherently issues that require political will, issues that require restructuring existing systems, issues that require political negotiation. So AI or other technological tools simply do not solve these problems.

When you come down another level, current AI systems—from large language models to simple tools that are used in hiring or in government—are inherently also built with data sets and with ideologies that encode and exacerbate inequality, societal norms, stereotypes, etc. Even if you have good intentions, trying to use those tools to solve complex problems is naive at best.

At an even higher level, if you look at major companies and corporations like Microsoft and Google, which have been trading their AI for social good initiatives on the one hand, but on the other hand exacerbating inequality, powering genocide, and powering war—they are the very corporations that are exacerbating environmental destruction. So the entire effort becomes a bit of an oxymoron.

[AI for good] is a way to paint a positive image of AI technologies, especially in light of a lot of the backlash—like the resist or refuse AI grassroots movement that’s emerging. So “AI for good” allows companies to say “Look, we’re doing something good! Everything about AI is not bad. And you can’t criticize us.” I think to the naive listener, those ideas might be convincing because it’s easy to believe that AI is like magic—that it can do anything and everything.

You have to ask: “What are we trying to solve? What are the systems we are using?” It’s only when you start asking those kinds of questions, you start to realize that a lot of the claims around “AI for good” start to crumble and don’t stand up to scrutiny.

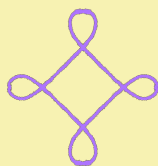
There may be surface improvements, but there are significant risks, too.

We’re going to have to wait and see [what happens]. Because a lot of the changes that the deployment of AI or other technologies bring in are very nuanced. There might be some surface-level improvements. For example, you might see some people having access to the internet or increased access to various services.

However, there is also underlying destruction and division that these systems are creating. My suspicion is it will take a while before it dawns on us to what extent the AI systems are really altering the social fabric, encoding existing norms and stereotypes in a way that makes the rich richer and more powerful.

It’s sometimes really scary the way you see some African governments jumping on the AI bandwagon and buying into this rhetoric that AI is going to “leapfrog” the continent into prosperity—with very little thought into the impact on people’s freedom of movement, freedom of speech, and on broader knowledge ecosystems. It’s a gradual regression.

I think the current framing is doing a lot of damage in terms of people’s consciousness, and people’s knowledge and understanding of what AI is.



REFRAME: Birhane argues that the path to actual “good” involves building small, community-based models and demanding robust evidence for social benefit claims.

Small, community-driven organizations are actually building “for good.”

I know of a lot of small startups or community-driven initiatives in small organizations that are doing an amazing job in using technology, or building technology from the ground up, that is allowing people to learn and helping people interact. I also know of several initiatives that are using technology to advance scientific knowledge, public communication, etc.—and you never hear those initiatives being framed as, or labeling themselves as “AI for social good.” Their efforts come from the sheer need to care for their communities, yet they don’t get the credit, they don’t get the accolades, and they are not often framed as AI for social good.

On the other hand, you see these large organizations in public claiming to do “AI for social good,” but in fact opposing—or supporting initiatives that regress—social progress.

I would be in favor of abandoning the term entirely and actually supporting small communities and small initiatives that are doing excellent work, without claiming “to do good.”

Demanding evidence of claims of benefits (rather than just documenting risk) is one way to pierce through the hype.

A lot of the claims are made by Big Tech corporations or AI vendors that have vested interest in ensuring that there is massive AI uptake and that these products are integrated across society. The problem is that policies are being made, massive investments are being made, and AI systems are being adopted and integrated into all kinds of spaces—just based on potentials and promises.

Unfortunately, at these summits, there has been very little discussion around whether there is empirical evidence for the positive claims that we are aiming for or we are hoping that AI will bring. Is there any empirical evidence? And, to what extent is the empirical evidence sound?

My expectation, my intuition is that there will be a lot of “Global South, Global Majority” narrative—but those at the margins of society will be put at further harm, and will be further disadvantaged by these [AI for good] narratives and by the uncritical adoption of AI systems. You would hope that governments would demand a little bit more scrutiny and more actual empirical evidence to support these claims, rather than kind of going with the “vibe.”

What I would love is for these governments to actually make decisions—or demand things—based on what is best for the people, especially for the people at the margins of society, and what is good for the environment.