IT for Change's Input at PARIS21 Spring Meetings 2024

April 2024

IT for Change

PARIS21 Spring Meetings 2024

Building Trust in AI: What Role for Smart Regulation?

Anita Gurumurthy, IT for Change[1]

In discussions on digital tech and artificial intelligence (AI), trust tends to be technicalized – as a security issue or an issue for data protection – but trust is really the way a social contract shapes civic behavior based on the predictability of systems to serve the public interest. Trust in AI, therefore, is about trust in the 'purpose integrity' of the AI system.

I think that with AI, in many circles that may not be so techno-optimistic, the skepticism is about an innovation-first approach/a solutionism that has accompanied inscrutable AI models.

Outcome distortions in technology, in areas like hiring, criminal justice, welfare, and more are responsible for direct and indirect harm to individuals and communities. Evidence of the degradation caused by deepfakes and other misinformation to our global information environment and trust architectures also come to mind; not to mention the rise of financial speculation around AI innovation.

The AI ecosystem, despite its immense promises, is built not so much on public reasoning and public oversight, but private and opaque logics. So, when we think of de-risking, or making AI safe and trustworthy, we need to think not only about addressing the technical aspects endogenous to the AI system. It is about enabling a continuous recalibration of the inner workings of the system for alignment with the norms and principles adequate to a just and equal AI society.

So, what is AI? What is it really about? And why is its governance so vital?

AI is not merely a product, but rather a paradigm that co-constitutes contemporary social, economic, and political relations:

As a means of production, its economic value for countries and entities is deeply implicated in trade and intellectual property (IP)— and consequently, the competitiveness and fairness of market structures.

As the ordering principle of communicative rationality, AI is a determinant of the diversity and vitality of our public sphere and the arbiter of our communicative rights.

Within the emerging architectures of democracy, AI tools play a large part in the architecture of government— of civic spaces and infrastructures, including electoral infrastructure.

---

[1] Anita Gurumurthy, Executive Director, IT for Change, made the following input at PARIS21 Spring Meetings 2024 held in Washington D.C. from 3-4 April 2024. Know more about IT for Change's work here.

On the AI economy and AI's role as a factor of production, two important aspects need mention. One is that Free Trade Agreements (FTAs) are often used to deny developing countries the ability to have access to algorithms. This means that while data in various services' value chains flows out of these countries, AI capabilities are retained by transnational firms of rich countries. Secondly, and more insidiously, trade secrets are often invoked incorrectly and misused to lock up aggregate non-personal data (NPD)— often from public statistical sources.

Societal data like behavioral data that may not really qualify for trade secrets are also in the enclosures of big corporations. This means the innovation economy is starved of the raw material needed to build AI. And so, the absence of a governance regime for data and AI that addresses questions of the data commons and the enormous social and public value captured in private enclosures, is really the roadblock for development and for developing countries (like for instance, the Lagos State Government has stated that Uber breached a data-sharing agreement in the city, which required real-time access to trip and location data via backend API integration).

The recent UN resolution in the General Assembly (GA) talks about safe and trustworthy AI and underlines transparent, inclusive, and equitable use of AI— while respecting IP and privacy. But we do need to look closer at the contradictions in AI that arises from a highly skewed, anticompetitive regime of data underlying the AI and the way IP is used to kill innovation.

There is a global dynamic, here, of misrecognition rooted in under-representation; but also, maldistribution — rooted in economic injustice. We are staring at data inequity that inhibits the Majority World's right to innovate; the appropriation of the data commons for proprietary AI and a totalizing control of national public infrastructures, including under-sea cables, by corporations.

There are also questions for regulation about sectoral data— data of peasants and indigenous people in corporate AI value chains (John Deer e.g., also needs us to reflect on what happens to the documentation of the crop, the soil and the farm— and how the ownership of the knowledge should be pinned down. Similarly with the health data commons.

So, what would the meta principles for governance include?

1. The role of public reason and justification for the AI (what is the problem? Do we need the AI and how should data and AI solve the problem).

2. If the AI will generate fair economic and social value for the local context. Context here is not only a data field but a moral vector– we need to focus on regenerative AI, not hollow out local value through homogenizing cognitive tools like generative AI.

3. Protection and enhancement of people's rights and rights of nature— suggestions and considerations of ecological impact have been largely missing from the AI governance conversation, even as its massive carbon footprint looms over the world.

Global governance norms must protect and promote the data commons and the right to knowledge and innovation for all (UNESCO open science declaration).

They must provide for binding accountability and liability mechanisms to ensure that there is no evasion of responsibility for AI harms by state or private actors. The recent report of the UN AI advisory group highlights the shortcomings of self-regulation, and mentions the need for 'binding norms enforced by member states.'

At national levels, we need an independent regulator to tackle everything from risk assessment and mitigation to documentation, transparency and proactive disclosure, public domain registers, auditability, and the need for *post facto* adequacy (which includes building systems with the capacity to flag relevant information to verify the machine's inferences as well as obligations on public authorities to record justifications while using these systems).

Human rights need re-articulation in relation to AI— to cover social communications, food sovereignty, health, environment, gender equality, welfare delivery, work/employment etc., taking into account the prevailing governance regimes in this area.

The future of governance needs to be based on civic participation to uphold citizen interest. AI policy is not a technical issue and certainly its workings need not be the subject-matter of technical experts.

Civic AI literacy is vital and is often a forgotten frontier while we talk of public-private partnerships (PPPs); we should talk about public-community partnerships.

One final word on equitable AI futures— we spoke about debt, and we also spoke about capabilities to meet the AI age— how will the countries caught in debt make it? This is really the elephant in the room.

Yes, philanthropy has a role, and we are happy with efforts to support skilling, but we are talking about digital innovation ecosystems as the foundational infrastructures of tomorrow and that is squarely part of the sovereign priority of people. There has been a shocking neglect of public finance needed to build digital public goods— including compute power, in developing countries. The grand narrative of generative AI or AI led by large corporations, and the normalization of private finance for the profoundly vital domain of AI creates a legitimacy problem for alternative imaginations (like the one with open street mapping), with real consequence for resource optimized and context-specific public innovation. We are too caught up with the mythical power of big data models and the discursive power of the big companies.

You may have seen this in the news, the self-checkout system used by amazon in many stores in the US has been called off. "It relied on a host of cameras, sensors, and a vast array of sensitive equipment and 1,000 people staring at video feeds to do the job of one or two people sitting behind cash registers at each store."

This is a serious policy issue about jobs and we see a tail-wagging-the-dog situation.

The Global Digital Compact (GDC) is being discussed in the UN, and we need to make sure that official development assistance (ODA) is available for developing countries. In fact, like in the climate debate, we may need to introduce ideas like common but differentiated responsibilities. This won't go well with the powers that be, but for all the data extractivism we have seen, just like reparation of debt, we may need to invoke ideas that may seem radical, but are perfectly logical (like the vaccine discussion we had).

So, circling back, the public trust issue can only be solved through a public innovation ecosystem where private enterprise has a central role, but the ecosystem itself is governed by democratic norms and means, where AI can promote agency and autonomy — not stifle it.

Are we ready for this shift, is the question?