

governments should make companies parties to international agreements. Including private companies in high diplomacy may veer toward unprecedented, but excluding those who have so much control would doom any governance structure that excludes them before it even starts.

- **Impermeable:** For AI governance to work, it must be impermeable; given AI's ability to easily proliferate, just one defection from the regime could allow a dangerous model to escape. Therefore, any compliance mechanisms should be watertight, with easy entry to compel participation and costly exit to deter noncompliance.
- **Targeted:** Given AI's general-purpose nature and the complexities involved in governing it, a single governance regime is insufficient to address the various sources of AI risk. In practice, determining which tools are appropriate to target which risks will require developing a live, working taxonomy of discrete potential AI impacts. AI governance must therefore be targeted, risk-based, and modular rather than one-size-fits-all.

Governing AI will be among the international community's most difficult challenges in the coming decades. As important as the imperative to regulate AI is the imperative to regulate it correctly. Current debates on AI policy too often tend toward a false debate between progress and doom (or geopolitical and economic advantages versus risk mitigation). And rather than think creatively, solutions too often resemble paradigms for yesterday's problems. This will not work in the age of AI.

Good policymaking will be vital, but getting there rests on good institutions. To build these institutions, the international community will need to agree on a conceptual framework for how to think about AI. We offer these principles as a start. **F&D**

IAN BREMMER is president and founder of Eurasia Group and GZERO Media. **MUSTAFA SULEYMAN** is CEO and cofounder of Inflection AI.

Fostering More Inclusive Democracy with AI



Hélène Landemore

AI can enhance democratic institutions by ensuring citizens' voices are truly heard

People worry that artificial intelligence is, or will soon be, undermining democracy. They fear AI will take away jobs, destabilize the economy, and widen the divide between the rich and the poor. This could further concentrate power in the hands of a few tech companies and weaken government structures designed to regulate them. Some also fear that tech giants and government may increasingly delegate human decision-making to machines, eventually replacing democracy with “algocracy,” rule not by the people but by algorithm.

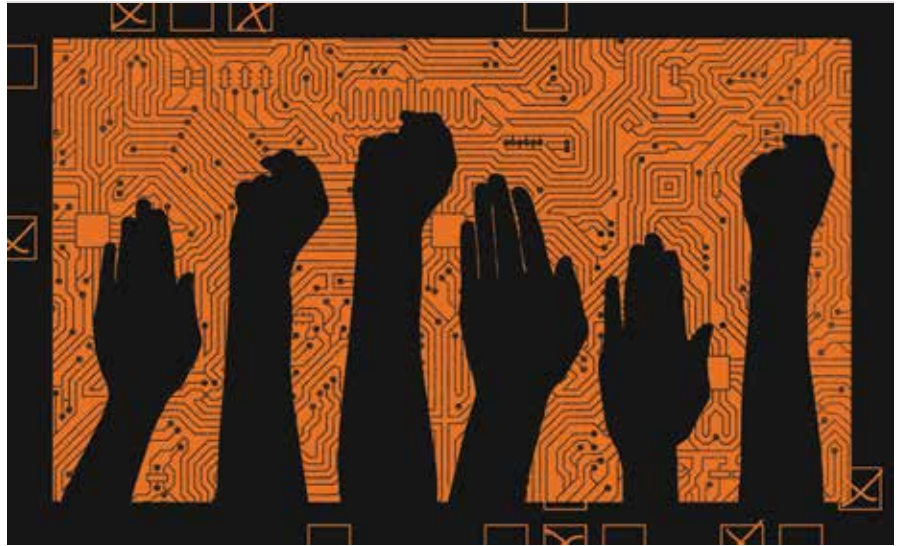
This dystopian vision misses our current capacity to shape AI development. We, as human societies, have the political ability (at least for now) and the responsibility to address the harm AI could inflict on us. We also have the technological opportunity to harness AI to enhance our democracy in a way that strengthens our collective ability to govern—rather than simply regulate—AI.

Like other ethical and political challenges, such as gene editing, AI governance requires not just more expert intervention and regulation but more citizen voice and input—for example, on how to navigate the distributive impact of AI on the economy. Like other global concerns, such as climate change, AI governance requires this democratic voice to be heard at the level of international institutions. Luckily, AI has the potential to usher in a more inclusive, participatory, and deliberative form of democracy, including at the global scale.

Participatory experiments

For 40 years many governments have engaged in experiments aiming to include ordinary citizens in policymaking and lawmaking in richer ways than through voting alone. These experiments have mostly been local and small-scale, much like the citizens' assemblies and juries that have proliferated on climate and other issues. A 2020

“AI regulation is likely to be better enforced and more effective in AI-empowered democracies.”



Organisation for Economic Co-operation and Development report found close to 600 such cases in which a random sample of citizens engages deeply with an issue and formulates informed policy recommendations (and in one case even proposals).

But some of these political experiments have also aimed for mass participation, as in the participatory constitutional processes organized in Brazil, Kenya, Nicaragua, South Africa, and Uganda in the 1980s and 1990s, and more recently in Chile, Egypt, and Iceland, which have used mass consultations and crowdsourcing to reach out to ordinary people. Not every attempt has been successful, of course, but all are part of a significant trend.

Some governments have also rolled out broad multi-format consultation campaigns. The 2019 Great National Debate launched by French President Emmanuel Macron in response to the yellow vest movement, with some 1.5 million participants, is one example. Another is the EU-wide Conference on the Future of Europe, which invited citizens from EU member countries to weigh in on reforms to EU policies and institutions, prompting 5 million people to visit the website and 700,000 to engage in debate.

Despite some online elements, these have been mostly low-tech, analog processes, involving no AI whatsoever.

Politicians, overwhelmed by the raw and multifaceted data or unsure of its meaning, have as a result easily ignored the citizens' input. People were allowed to speak but were not always heard. And the level of deliberation, even for those involved, was often superficial.

Enhanced deliberation

We now have the chance to scale and improve such deliberative processes exponentially so that citizens' voices, in all their richness and diversity, can make a difference. Taiwan Province of China exemplifies this transition.

Following the 2014 Sunflower Revolution there, which brought tech-savvy politicians to power, an online open-source platform called pol.is was introduced. This platform allows people to express elaborate opinions about any topic, from Uber regulation to COVID policies, and vote on the opinions submitted by others. It also uses these votes to map the opinion landscape, helping contributors understand which proposals would garner consensus while clearly identifying minority and dissenting opinions and even groups of lobbyists with an obvious party line. This helps people understand each other better and reduces polarization. Politicians then use the resulting information to shape public policy responses that take into account all viewpoints.

Over the past few months pol.is has evolved to integrate machine learning with some of its functions to render the experience of the platform more deliberative. Contributors to the platform can now engage with a large language model, or LLM (a type of AI), that speaks on behalf of different opinion clusters and helps individuals figure out the position of their allies, opponents, and everyone in between. This makes the experience on the platform more truly deliberative and further helps depolarization. Today, this tool is frequently used to consult with residents, engaging 12 million people, or nearly half the population.

Corporations, which face their own governance challenges, also see the potential of large-scale AI-augmented consultations. After launching its more classically technocratic Oversight Board, staffed with lawyers and experts to make decisions on content, Meta (formerly Facebook) began experimenting in 2022 with Meta Community Forums—where randomly selected groups of users from several countries could deliberate on climate content regulation. An even more ambitious effort, in December 2022, involved 6,000 users from 32 countries in 19 languages to discuss cyberbullying in the metaverse over several days. Deliberations in the Meta experiment were facilitated on a proprietary Stanford

F&D

Online editions in six languages

Léala en
español



Disponible
en français !



Читайте
по-русски!

اقرأ باللغة
العربية !

请阅读中文版 !



日本語はこちら !

www.imf.org/fandd

University platform by (still basic) AI, which assigned speaking times, helped the group decide on topics, and advised on when to put them aside.

For now there is no evidence that AI facilitators do a better job than humans, but that may soon change. And when it does, the AI facilitators will have the distinct advantage of being much cheaper, which matters if we are ever to scale deep deliberative processes among humans (rather than between humans and LLM impersonators, as in the Taiwanese experience) from 6,000 to millions of people.

Translation, summarization, analysis

The applications of AI in deliberative democracy are still in the exploratory phase. Instantaneous translation among multilingual groups is the next frontier, as is summarization of collective deliberations. According to recent research, AI is 50 percent more accurate than human beings when it comes to summarization (as evaluated by trained undergraduates comparing AI summaries and human coders' summaries of deliberation transcripts). Some amount of human judgment will, however, likely be necessary for many of these tasks. In such cases AI can still serve as a useful aid to human analysts, facilitators, and translators.

More ways that AI can enhance democracy are on the horizon. OpenAI, the company that launched ChatGPT, recently introduced a grant program called Democratic inputs to AI. The grants subsidized the 10 most promising teams in the world working on algorithms that serve human deliberation (full disclosure: I am on the board of academic advisors who helped formulate the grant call and select the winners). These tools can hopefully soon be deployed to serve, among other goals, global deliberation on AI governance, in line with the vision of OpenAI CEO Sam Altman.

Addressing risks

Deploying AI in democracy has its risks—like data bias, privacy concerns, potential for surveillance, and legal

challenges—in almost every field. It also raises the problem of the digital divide and the potential exclusion of illiterate and techno-skeptical groups. Many of these problems will need to be addressed politically, economically, legally, and socially first and foremost, rather than through technology alone. But technology can help here too.

For example, privacy and surveillance concerns may be remediated by something such as zero-knowledge protocols (also called zero-knowledge proofs, or ZKP), which aim to verify or “prove” identity without collecting data on participants (for example, through text messaging authentication or through blockchain). ZKP can be used both for online voting and in deliberative contexts—for example, to share sensitive information or play the role of whistleblower. Meanwhile, generative AI can make previously scarce knowledge and tutoring resources available to everyone who needs them. As a custom-tailored interlocutor for citizens, it can explain technical policy issues in people's particular cognitive style (including through images) and convert their oral input into written input as needed.

Despite its limitations and risks, AI has the potential to bring about a better, more inclusive version of democracy, one that would in turn equip governments with the legitimacy and knowledge to oversee AI development. AI regulation is likely to be better enforced and more effective in AI-empowered democracies.

Still, there is a risk that democracy itself could be a casualty of the AI revolution. Urgent investment is needed in AI tools that safely augment the participatory and deliberative potential of our governments. **F&D**

HÉLÈNE LANDEMORE is a professor of political science at Yale University. She is also a fellow at the Ethics in AI Institute at the University of Oxford and an advisor to the Democratic inputs to AI program at OpenAI.