AI and Global Governance

Ethics and Governance of Artificial Intelligence Initiative

Across the globe, a new wave of AI-based technologies and applications are being rolled out in schools, homes, and hospitals, with digital leaders across the high tech, telecom, and financial services sectors among the early adopters. AI promises enormous benefits for the social good and can improve human well-being, safety, and productivity. But it also poses significant risks for workers, developers, firms, and governments alike. As a global society, we are only beginning to understand the ethical, legal, and regulatory challenges associated with AI, as well as to develop appropriate governance models and responses.

Many of the new technologies and applications challenge existing institutions, incumbent industries, and familiar power dynamics. Some of these effects resemble patterns of disruption at the precipice of the Internet revolution. As in the early days of the Internet, policymakers at the national and international level are struggling with the question how to best deal with these tectonic shifts and how to regulate new technology to embrace its promise, but also to address a broad variety of concerns, ranging from the impact of automation on employment to privacy, safety, and justice. These past experiences with global Internet governance might be helpful today as policymakers deal with the latest shockwaves caused by AI. For instance, models of multi-layered and multi-stakeholder Internet governance might also inspire and influence AI governance. At the same time, these modes have worked better for some types of challenges and not others; the difficulty in solving challenges related to encryption and cybersecurity, as well as harmful speech and misinformation online, suggest limitations and a need for imagination.

Perhaps even more so than with the Internet, policymakers - often navigating in the dark - face hard problems and difficult choices when addressing AI systems. Today, it remains unclear what the actual impact of AI-based technologies on the digital economy and society at large will be. It is even uncertain whether the right metrics and instruments are readily available to measure AI's societal impact. Moreover, the discussion about the normative implications and tradeoffs of AI is at an early stage, with many of the challenging ethical questions (related to fairness, equality, and inclusion, among other topics) surfacing slowly when compared with the rapid proliferation of AI in everyday life. And even where a value consensus can be reached in a timely manner, it is often unclear what the best policy instruments are to address the different challenges within and across the broad range of contexts in which AI systems are deployed - at the national and, even more so, at the global level.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

23 Everett St., 2nd Floor
Cambridge, MA 02138

cyber.harvard.edu
hello@cyber.harvard.edu
@BKCHarvard
617.495.7547

Not all of these challenges are radically new, of course, but they cannot be addressed by simply reapplying conventional approaches. Our interactions with national and international policymakers have revealed three cross-sectional sets of questions that can be prioritized and seem worthy to be addressed in order to bolster the use of AI for the public good: (1) How can the information asymmetries between technologists and policymakers be reduced, given the complex "black box" nature of AI technology? (2) What are all the instruments available in the governance toolbox - borrowing also from previous global governance challenges - that can be applied to AI issues? (3) How can academia contribute to the formation of robust discussion forum to inform the debate about global AI governance challenges and solutions?

The AI and Global Governance track levels information asymmetries between the public and private sectors in order to give policymakers the tools they need to effectively navigate a complex, highly technical, and global policy space.

In undertaking this work, the **Berkman Klein Center** will draw upon its close relationship to the Global Network of Internet & Society Research Centers and its ties to international organizations as diverse as the World Economic Forum, the OECD, and the ITU.

By working closely with various stakeholders, we seek to address some of the following challenges:

› **Public-Private Interfaces**: The development and use of AI are significantly shaped by decisions-makers in the private sector. A handful of leading companies that operate at a global level make important design choices that have lasting implications for societies that increasingly rely on AI-based technologies. From a public interest perspective, policy-makers are called upon to create robust interfaces between the private and public sector. Approaches include traditional modes of regulation in areas like data protection, as the right to explanation in the EU suggests. Other approaches might be based on incentives or alternative modes in which governments can exercise their power, for instance through procurement processes. We will help policy-makers to to identify the tools that will be most effective, by drawing on our experiences studying policy successes and failures across technologies and issues, ranging from cloud computing to cybersecurity.

› **Legal Interoperability**: As policymakers across jurisdictions begin to develop sector-specific regulations and laws aimed at governing AI-based applications at the local, national, regional, and international levels, there is a strong possibility that we could end up with a fragmented set of interventions that are unlikely to work together across boundaries and borders. We aim to support the development of bottom-up design principles that could enable increased legal interoperability across these interventions.

› **Data Inclusivity**: It is well understood that biased or incomplete training data can make AI tools biased themselves. A facial recognition AI only trained on white male faces may fail when deployed in the real world with a diversity of users. Policymakers are in a position to help address this problem, as they have access to massive, inclusive data sets. Yet, by the same token, there are privacy risks that arise in giving private sector actors access to this breadth of information.

Pillars of Impact

In building solutions that address these challenges, our institutions are making a series of investments, most significantly in:

1. **Bridging Information Asymmetries:** The speed of advancement and our limited understanding of technological and social implications of AI challenge our conceptions of autonomy, agency, and accountability. Simultaneously, we see an expanding divide between the small group of AI experts and the large population affected by these technologies. Caught in the middle are policymakers who lack the time and expertise to learn all the nuances of every new technological advancement, and they answer to a public who often have access to even less information. Academic institutions are one of the few places that can bridge the divide between not just AI experts and policy experts, but also between the social sciences, humanities, and hard sciences. As we've learned from Internet policymaking, this is far harder than simply putting people in a room. The Berkman Klein Center and MIT Media Lab play three roles that can help this knowledge exchange: (1) we understand how to straddle disciplines because we are interdisciplinary knowledge creators, drawing on a diverse community of faculty, fellows, and students; (2) we are experienced at translating research into practice; and (3) we are experienced conveners, facilitating conversations across disciplines and between the public and private in order to advance the public interest.

2. **Expanding the Governance Toolbox:** Policymakers wield a far more diverse toolbox than just the traditional instruments of regulation and legislation. In studying several different digital technologies, from the Internet itself, to cloud computing, to digital currencies, to broadband infrastructure deployment, we have charted the range of tools available to policymakers. These include multistakeholder approaches, norms and standard setting processes, procurement powers, among others. Successful policymakers often deploy these varied tools and levers of influence in strategic ways. To policymakers less familiar with them, their applications and their nuances can seem daunting. Activating our diverse networks, the Berkman Klein Center and MIT Media Lab will create a playbook that will guide public-sector decision-makers through these tools and their applications.

3. **Building Stakeholder Relationships:** Many of the policy questions emerging from AI have global implications - and typically no single country can solve the problems on its own. It is increasingly important for policymakers to have robust global relationships in order to develop effective solutions. We will leverage an international network of academic institutions through the Global Network of Internet and Society Centers, as well as their deep relationships with global policymakers in order to directly support public-sector decision-makers in responding to global AI challenges.

About the Ethics and Governance of Artificial Intelligence Initiative

The rapidly growing capabilities and increasing presence of AI-based systems in our lives raise pressing questions about the impact, governance, ethics, and accountability of these technologies around the world. How can we narrow the knowledge gap between AI "experts" and the variety of people who use, interact with, and are impacted by these technologies? How do we harness the potential of AI systems while ensuring that they do not exacerbate existing inequalities and biases, or even create new ones? At the Berkman Klein Center, a wide range of research projects – including the one outlined above – community members, programs, and perspectives seek to address the big questions related to the ethics and governance of AI under the Ethics and Governance of AI Initiative, launched in 2017.