

---

## AI and Inclusion

---

AI-based technologies and applications offer tremendous opportunities to build a better world. Advanced systems can be used to enhance educational instruction, improve the efficiency of transportation infrastructure, fight epidemics, support the elderly, and combat corruption, to name just a few examples. However, perhaps more so than ever before, these same AI-based technologies can also deepen existing divides, gaps, and inequalities, and even create new ones if not developed and deployed in thoughtful ways with appropriate safeguards and support mechanisms in place.

**In a world challenged by growing domestic and international inequalities, it is essential that we respond quickly to the ways in which AI may exacerbate existing imbalances and generate new disparities.** Recent evidence suggests that traditionally marginalized populations – in terms of ethnicity, race, gender, gender and sexual identity, religion, national origin, location, age, skill and educational level and/or socioeconomic status – have limited and uneven access to AI-based technologies and their benefits. This amplifies existing structural inequalities, especially for those at the intersections of such groups.

Classification and categorization are not only at the heart of the notion of social inclusion, but also at the core of AI-powered technologies, which revolutionize the ways in which data and people are systematized, often with unforeseen consequences. As we become more dependent upon autonomous systems, **we must reconceptualize social inclusion itself in order to fully understand the challenges and opportunities that AI systems may present.** Consider the algorithms used in hiring processes that attempt to classify applicants as prospectively successful (or unsuccessful) hires based on historical patterns, yet may instead identify and reinforce the biases of their designers or available training data. AI's ability to categorize, cluster, and pattern match creates new pathways for discrimination and exclusion that call for new methods of analysis. These applications and systems, and their tangible consequences on people's lives, cannot be simply extricated from the larger context of social inclusion; rather, they must be continually reevaluated in dialogue across geographies, disciplines, and sectors, alongside social inclusion and its traditional precepts.

The deep knowledge gap separating most people affected by AI systems and their designers creates an information asymmetry that complicates and often threatens progress toward a more inclusive future. The youth population is a particularly vulnerable group, as very little has been done to empower young people to critically engage the discourse surrounding the next generation of technologies that have a marked potential to shape their lives for better or for worse. Taking into account the massive impact AI is predicted to have on the future of work and employment, it is critical that youth are well-equipped to participate in the debates about the ethics and governance of AI.

Progress towards a more inclusive society in which AI technology is predicted to play a prominent role will require interdisciplinary dialogue and collaboration that bring together a variety of stakeholders, including governments, companies, international organizations, and civil society groups. It also demands a concerted effort to include individuals who may be the most flagrantly affected by the next generation of technologies and yet the least empowered to engage and be heard in conversations related to AI development and applications.

**The AI and Inclusion track will foster the design and deployment of AI to benefit all members of society, including traditionally underserved communities.** It will advance these objectives through research, learning, education, and engagement across local and global communities to close existing digital divides and participation gaps.

The **Berkman Klein Center** and **MIT Media Lab** are engaging with a diverse group of stakeholders to inform, support, and strengthen efforts aimed at ensuring that AI systems are used in ways that advance diversity and inclusiveness. Drawing upon our breadth of partnerships and our experience translating research efforts into outreach and educational offerings for policymakers and technology leaders, we seek to address the following challenges:

- **Reconceptualizing Inclusion:** We will reexamine the notion of social inclusion and how it should be applied to AI. How do different individuals across a variety of backgrounds, geographies, and disciplines approach and understand the key challenges and opportunities at the intersection of AI and inclusion? How can such diverse viewpoints and perspectives inform a new understanding of inclusion, as AI applications may unleash new dynamics that differentiate between “in” and “out”?
- **Empowering Groups:** We will focus on empowering underrepresented groups to critically engage with the discourse around AI. How do different groups, communities, and geographies use and relate to AI, and how does this affect our understanding of the promise and limitations of AI systems? How can we empower underrepresented groups — including youth — to think more deeply, critically, and creatively about AI’s current and future impacts on their lives?
- **Working with Stakeholders:** Partnerships with a diverse group of stakeholders will inform research initiatives and potential policy interventions. How can we improve collaborations across sectors to increase information flows between underrepresented communities, AI experts, inclusion experts, and the public at large?

## Pillars of Impact

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

1. Co-hosting the **Global Symposium on AI & Inclusion (November 8-10, 2017)**, which aimed to identify and address an array of issues involving AI and Inclusion from an interdisciplinary perspective. This was the first of several events in the works that will bring together interdisciplinary, global stakeholders to establish interfaces among individuals and groups with a shared interest in creating a more inclusive society supported by AI.
2. Co-developing and curating **educational resources – “learning experiences”** – with youth and underrepresented populations. A learning experience is a semi-structured activity one completes to gain knowledge about a specific topic (e.g., the intersection between AI and wellbeing; AI and news). Each is organized around a measurable goal, incorporates multimedia materials, and encourages the production of an artifact (e.g., a blog post, video, infographic). The resources are designed to empower underrepresented groups — such as the youth population — to think more deeply about AI systems.
3. Conducting short, action-oriented **research sprints** among experts about specific topics (e.g., AI & Youth), with a focus on dimensions of inclusion. Insights gained from this work will be shared with diverse stakeholders as part of our translational research and educational mission.

---

For more information, please contact Sandra Cortesi at [scortesi@cyber.harvard.edu](mailto:scortesi@cyber.harvard.edu) or visit <https://cyber.harvard.edu/research/ai>.