

# Roadmap for an MIT Initiative on Computing, Data, and Racial Justice

Spring 2021

As advances in computing and data sciences impact more and more facets of people’s daily lives, new technologies risk exacerbating existing inequalities. Datasets of past events or outcomes often reflect historical patterns of systematic prejudice and injustice—both the visible and the obfuscated—which in turn can reinscribe biases within the algorithms that are trained upon them today. More generally, technologies that are developed by people with particular horizons— informed by their own backgrounds and experiences—often fail to recognize, let alone address, the priorities of other groups and communities.

Given MIT’s leadership in computing and the data sciences, and the significant stakes for broader society, MIT has a special responsibility and opportunity to foster actively antiracist computing practices. We propose an Institute-wide Initiative on Computing, Data, and Racial Justice. Like similar Institute initiatives, the new effort can focus the attention of members of the MIT community on high-priority topics (“top-down” signaling) while supporting widespread improvisation, creativity, and collaboration among individual scholars, students, and research groups (“bottom-up” participation).

The new Initiative on Computing, Data, and Racial Justice should support and accelerate efforts already underway in various departments, labs, and centers, and should complement—and be informed by—MIT’s [Strategic Action Plan on Diversity, Equity, and Inclusion](#). The Initiative can coordinate and expand upon existing efforts in research, teaching, public policy, and engagements with civic groups and private-sector partners beyond campus. The Initiative can further foster new, robustly multidisciplinary efforts with emphases upon accountability and inclusivity in design, implementation, and evaluation of new technologies, and on creating equitable benefit for members of marginalized groups, drawing on the experience and expertise of members from all five Schools at MIT.

## New Efforts

### 1. Sharing Resources

A first step for the new Initiative will be to highlight efforts and align pipelines on computing, data, and antiracism that are already underway across the Institute. We envision an easy-to-navigate “dashboard”-style website, modeled on the “Climate Science and Climate Risk” site ([climateprimer.mit.edu](http://climateprimer.mit.edu)) produced by Prof. Kerry Emanuel and colleagues in MIT’s Department of Earth, Atmospheric, and Planetary Sciences.

Such a site would be useful for researchers and students on campus, enabling members of the MIT community who are already immersed in such efforts to connect with each other, to amplify ongoing efforts, and highlight opportunities for others who are eager to get involved. A well-designed portal or dashboard site would also help connect MIT efforts with colleagues at other academic institutions, private-sector companies, and civic groups.

By featuring links to recent research papers, collaborations, and events, as well as to syllabi, the new open-access [MIT Case Studies Series on Social and Ethical Responsibilities of Computing](#), and other teaching materials, the site would make original materials broadly accessible for specialists and nonspecialists alike.

## **2. Learning from the Past**

In many fields of engineering today, specific mechanisms have been developed to investigate episodes that cause harm, so as to learn from past mistakes and improve designs for the future. Within the United States, for example, the [National Transportation Safety Board](#) (NTSB) conducts detailed “after incident” investigations following accidents involving the civil transportation sectors. Teams of specialists work to identify causes and prepare formal reports with recommendations. The investigative framework is designed to feed back into daily practice and teaching, for the benefit of individual companies, regulators, researchers, and broader civil society.

Although some large, private tech companies presently conduct their own internal investigations, there does not yet exist a shared framework for systematically investigating—and learning from—unintended consequences and potential harms caused by specific computing practices and/or datasets. Such a framework should be grounded in insights and research methodologies drawn from across scientific and engineering fields as well as humanities and social sciences, drawing on the full range of MIT’s strengths.

We propose creating a framework for robust “after incident” investigations in computing and the data sciences, the results from which would be shared openly. In the near term, such a framework could help students and researchers at institutions like MIT learn from unanticipated events or inadvertent harms that can arise from well-intentioned research projects. Building such a framework for review and recommendations within academic research settings would be an important start, offering learning opportunities for students and rising computing professionals, even if such reforms might take longer to implement among private companies.

## **3. Reducing Threats when Conducting Research**

For several decades, academic research involving human subjects conducted within the United States has been governed by requirements for review by Institutional Review Boards (IRBs), such as MIT’s [Committee on the Use of Humans as Experimental Subjects](#) (COUHES). Historically, the IRB infrastructure developed in response to harms—both to individual subjects and to vulnerable populations—from research in the biomedical and social sciences. Most concerns centered around potential harms arising from the act of collecting data for a given study, such as the (now-infamous) [Tuskegee Syphilis Study](#) of hundreds of Black men whose illnesses went untreated for decades.

Research in various areas of computing and the data sciences also involves the use of data from human subjects, but such efforts have rarely been subject to systematic review within the existing IRB structure. For example, research projects on facial recognition algorithms have typically used large collections of photographs from publicly accessible repositories, rather than

collecting photographs directly from individual participants. Such scenarios include no provision for “informed consent” from the people whose data are incorporated into the studies, nor any systematic means to identify potential harms that might arise from the studies.

Several [recent controversies](#) highlight potential negative impacts—both for individuals and for certain populations—from research involving the analysis of large volumes of publicly accessible data. Such potential impacts include unintentional violations of privacy expectations (from the collation of separate, large datasets) and differential impacts from policing or surveillance on members of marginalized communities. These negative impacts typically arise from the uses to which data from human subjects are put, even in those cases for which data collection posed no direct threat to individuals.

Given the ubiquity of human-sourced data in contemporary computing and data fields, and the distinct threats and vulnerabilities that they pose (often disproportionately to communities of color), we propose the creation of a new review mechanism specifically designed to minimize potential harms to human subjects arising from research projects in computing and the data sciences. Building on recent experience with the [Legal, Ethical, and Equity committee](#) for campus planning, such a review and oversight structure must incorporate the input and experience from faculty, students, and staff drawn from across all five Schools of the Institute.

Several academic and professional societies are presently considering the creation of such review structures, including research fields focused on computer vision and natural-language processing. Computer scientists increasingly recognize that IRBs do not have the right tools to address issues of dataset safety and other potential biases and harms associated with machine-learning systems, nor do other organizations exist to which researchers may turn for help. MIT should pioneer these efforts, ensuring that such vital review and oversight functions can be performed rigorously and well.

#### **4. Accountable research practices**

Scholars in several fields of study, including environmental sciences, urban planning, and health sciences, place increasing emphasis upon the development and implementation of accountable and inclusive research practices, in which community residents and organizations, academic and public institutions collaborate in the design, implementation, and evaluation of new data science methods and technologies.

It is imperative to develop such [participatory approaches](#) to machine learning as well; early efforts are underway to understand how to implement such practices and to mitigate the potential [harms](#) of such approaches. MIT has the potential to lead in the innovation of accountable research practices in AI and data science, building on the scholarship and experience of MIT faculty who already lead in related efforts that span MIT’s five Schools and College. We propose a set of new activities to catalyze research, innovation, and experimentation in this space through workshops to share efforts and build new collaborations, together with longitudinal working groups to catalyze and seed longer-terms collaboration and proposals on such research.

## 5. Training our students in antiracist computing practices

Efforts to foster actively antiracist computing practices must extend to our undergraduate and graduate teaching efforts. As a start, we propose to build on efforts already underway in the Social and Ethical Responsibilities of Computing (SERC) to develop original [Case Studies](#) and Active Learning Projects (ALP) on social, ethical, and policy aspects of computing, including a “focal area” emphasizing computing and racial justice. All such materials will be made freely available (at MIT and beyond) on various open-access platforms, so that faculty can embed the new materials into existing computing courses; the original materials will be made broadly accessible to the public and for other educators to incorporate in their classes as well. The Case Studies and ALP efforts provide a direct vehicle for bringing to our students wider awareness of the patterns of systematic prejudice and injustice that reinscribe existing biases as we develop new technologies.

These proposals are only a start. As part of a new MIT Initiative on Computing, Data, and Racial Justice, they would help the MIT community assess where we are now and take concrete steps to ameliorate particular harms or vulnerabilities. They would establish a multidisciplinary framework within which members of our community could foster new collaborations—across campus and beyond—to pursue research, teaching, and broader engagements, harnessing computing and the data sciences to proactively work against systemic racism.

Prepared by members of the SERC Action Group on Computing, Data, and Antiracism

Co-chairs:

Julie Shah (Aero/Astro, CSAIL): Associate Dean for Social and Ethical Responsibilities of Computing

David Kaiser (STS, Physics): Associate Dean for Social and Ethical Responsibilities of Computing

Members:

Sandy Alexandre (Literature)

Bill Freeman (EECS)

Chakanetsa Mavhunga (STS)

Amy Moran-Thomas (Anthropology)

Aude Oliva (EECS, Quest)

Ronitt Rubinfeld (EECS)

Justin Steil (DUSP)

Collin Stultz (EECS, IMES)

Jessika Trancik (IDSS)

Katlyn Turner (Media Lab)

Ariel White (Political Science)

Karen Zheng (ORC, Sloan)