CALL FOR PAPERS
4th AAAI/ACM Conference on AI, Ethics, and Society
A single-track virtual conference

www.aies-conference.com
www.facebook.com/AIESConf
twitter.com/AIESConf

Submission deadline: 31 January 2021
Notification: 7 April 2021
Final version: April 30, 2021
Conference: May 19-20-21, 2021

Over the last few years, the world has awoken to the power that we have vested—often without thought or care—in the people and systems that collect, aggregate, analyse, and act on our data. At the same time, AI systems promise new ways to empower individuals and collectives to change society from the bottom up. International organisations, governments, universities, corporations, and philanthropists have recognised the urgent need to bring all of our intellectual tools to bear on charting a course through this uncertain new territory. Earlier iterations of this conference and others have seen the first fruits of these calls to action, as programs for research have been set out in many fields relevant to AI, Ethics, and Society.

The early days of shaking us awake are done: we now know, well, that we are increasingly reliant on AI systems that are radically changing the world around us, for better and worse. The next step is to chart a course forward, both by deepening our diagnosis of where we are now, and by developing new goals, models, and technical and regulatory systems to shape the future of AI and society toward how we collectively intend our societies to look.

To achieve these twin objectives—a richer understanding of where we are now, and technical and sociotechnical paths forward—we must draw on insights from across disciplines. AIES is convened each year by program co-chairs from Computer Science, Law and Policy, the Social Sciences, and Philosophy. Our goal is to encourage talented scholars in these and related fields to submit their best work related to the morality, law, and political economy of data and AI. Papers should be tailored for a multi-disciplinary audience without sacrificing excellence. In addition to the community of scholars who have participated in these discussions from the outset, we want to explicitly welcome disciplinary experts who are newer to this topic, and see ways to break new ground in their own fields by thinking about data and AI.

The following list of topics and examples is intended to be illustrative, not exhaustive.

- **Empirical research into the impacts of AI systems.**
    o Work bringing to light applications of AI with significant but insufficiently recognized impacts.
        · *E.g. detailing new and underexplored uses of AI in government, defence, healthcare, finance, political campaigning, marketing, digital platforms and other areas.*
    o Work advancing our theoretical understanding of how AI systems are changing societies.
        · *E.g. exploring how data and AI-driven policy-making leads to changes in how governments see citizens (and vice versa); how industry shapes social*

environments so that they are more susceptible to datafication; how AI systems can react to, produce and reproduce social inequality and prejudice, including racism and misogyny; the social consequences of automation; political economy of big tech.

- o Work investigating public or professional resistance to the deployment of AI systems.
- **Evaluative research into AI impacts.**
  - o Work deepening the moral diagnosis of existing and feasible AI systems.
    - · *E.g. theoretical accounts of why surveillance may be resisted or embraced; how it reshapes subjectivity and behavior; the kinds of manipulation it enables; accounts of the nature of discrimination as practiced by AI systems; existential risks posed by the development of AI systems.*
  - o Work evaluating existing and feasible AI systems against existing legal and regulatory regimes.
    - · *E.g. assessing the feasibility of 'black box' AI systems complying with existing administrative law; data protection implications of existing AI systems; impact of AI systems on antitrust issues.*
- **Evaluative research into the goals at which we should aim when redesigning AI systems.**
  - o Theoretical work aimed at addressing, understanding, or resolving evaluative uncertainty and disagreement about goals to aim at with AI systems.
    - · *E.g. determining how to think about discrimination in the age of AI; how to philosophically conceptualize alignment with human values.*
  - o Normative theory aiming to map out how AI systems could be used legitimately, and for social benefit.
    - · *E.g. re-examining the moral foundations of administrative law to devise standards for AI-assisted institutional decision-making.*
- **Technical research into the representation, acquisition, and use of ethical knowledge by AI systems.**
  - o How can ethical knowledge be represented as rules and constraints; as utility functions; as stories and scripts; as deep neural networks; etc?
    - · *E.g. ethical knowledge is learned by humans from limited amounts of experience and pedagogy; what does this mean for representation?*
  - o How should key concepts such as fairness and bias be formalized to allow properties of intelligent systems to be evaluated and guaranteed?
    - · *E.g. establishing "best practices" for training set curation to prevent or reduce transmission of existing societal bias to a learning system.*
- **Proposal and/or evaluation of technical methods for realising evaluative goals.**
  - o Work focusing on developing AI systems for specific application domains that advance valid evaluative goals.
    - · *E.g. 'Mechanism Design for Social Good' and related areas.*
  - o Work introducing mechanisms for procedural justice into AI systems as deployed in practice.
    - · *E.g. methods for making AI systems in practice better suited to democratic governance; design tools for introducing auditability trails into AI systems; explainable AI with a social purpose.*
- **Proposal and/or evaluation of sociotechnical methods for realising evaluative goals.**
  - o Work exploring the culture and practices of AI research and development to counteract structural injustice.
    - · *E.g. labor rights and employee activism in the tech sector; alternative socially-oriented methods for AI research and development such as data trusts and public benefit corporations; nature of collective mobilization in digitally distributed environments.*

- o   Work proposing and evaluating methods for responsible and inclusive innovation with active involvement from those affected by new technologies.
  - ▪   *E.g. methods for participatory design and responsible innovation practices.*
- **Proposal and/or evaluation of legal and regulatory approaches for realising evaluative goals.**
  - o   Work exploring the relative merits of using legal instruments such as antitrust, consumer protection, and data protection to regulate the impacts of AI.
    - ▪   *E.g. comparative analysis of data protection regimes; arguments for or against explicit regulation of automated decision-making; ongoing prospects for transnational regulation.*
  - o   Work exploring the role of public law in constraining public use of AI and related technologies.
    - ▪   *E.g. investigation of how administrative law needs to be revised to accommodate AI (or vice versa).*

Submitted papers should address these or related topics in ways that make a substantive contribution to knowledge in one or more fields. A paper should clearly establish its research contribution, its relevance, and its relation to prior research.

Submitted papers must be 6-10 pages (including all figures and tables) in AAAI two-column format, plus unlimited pages for (non-discursive) references. This typically corresponds to no more than 8,000 words for the main content. For the AAAI format, see the templates provided at
https://www.aaai.org/Publications/Templates/AuthorKit21.zip
        [The AAAI formatting templates are intended for final camera-ready copy of accepted papers. The AAAI copyright block is hard-coded into the AAAI paper templates to retain proper spacing, and cannot be removed. It is not considered binding until a paper is accepted and a signed copyright form is submitted by the author. At the initial submission for review stage, it is not necessary to submit source files as supplementary material.]

Optionally, authors can upload supplementary materials (e.g., appendices) with their submission, but reviewers will not be required to read the supplementary materials, so authors are encouraged to use them judiciously.

Authors should note that changes to the author list after the submission deadline are not allowed. At least one author of each accepted paper is required to register for, attend, and present the work at the conference.

All submissions must be submitted through the EasyChair link on the conference website.

Review will be double-blind, so authors should remove identifying information from their papers. However, to assist selecting reviewers, authors should report the paper's primary disciplines on the first page.

IMPORTANT NOTICE: All submitted papers must meet the above criteria. However, to accommodate the publishing traditions of different fields, authors of accepted papers can provide a one-page abstract of the paper for the conference proceedings, along with a URL pointing to the full paper. Authors should guarantee the link to be reliable for at least two years. This option is available to accommodate subsequent publication in journals that would not consider results that have been published in preliminary form in a conference proceedings. Such papers must be submitted electronically and formatted just like papers submitted for full-text publication.

Papers submitted to AIES-2021 may not be published or accepted for publication at an archival conference or journal prior to submission to AIES.

The proceedings of the conference will be published in the AAAI and ACM Digital Libraries.

Recognizing that a multiplicity of perspectives leads to stronger science, the conference organizers actively welcome and encourage people with differing identities, expertise, backgrounds, beliefs, or experiences to participate.

IMPORTANT DATES
Submission Deadline: January 31, 2021
Notification: April 7, 2021
Final version: April 30, 2021
Conference: May 19-20-21, 2021

Conference program co-chairs:
Marion Fourcade (UC Berkeley)
Benjamin Kuipers (Michigan)
Deirdre Mulligan (UC Berkeley)
Seth Lazar (Australian National University)