Interdisciplinary Ethics of AI Workshops for Future Policymakers:

Pilot Project at Georgetown University, 2019-21

Interdisciplinary Ethics of AI Workshops for Future Policymakers: Pilot Project at Georgetown University

1. Overview

In this pilot project pursued by Georgetown University's Ethics Lab and the Center for Security and Emerging Technology (CSET), we delivered a set of three workshops on AI ethics for future policymakers assembled from three major fellowship programs in 2020-21.

- **Workshop 1 (March 2020) – AI & Bias/Discrimination**. How can we make AI systems more just? This workshop scrutinized the predictive capabilities of AI to investigate how AI systems can be biased, discriminatory, or otherwise unfair. Our case study delved into the increasingly popular use of AI at various stages of the hiring process. Our goal was to help participants better understand how applicants' rights to fair and equitable treatment can be adversely affected by problematic, overhyped AI tools.

- **Workshop 2 (November 2020) – AI & National Security**. What role should AI play in advancing national security? This workshop explored uses of AI in transnational security contexts in order to identify ethical considerations relevant to risk assessments in policymaking. Our case study centered on a fictionalized US military mission overseas involving the option to deploy an AI-based weapon with an autonomous mode. Participants grappled with ethical risks and uncertainties raised by the potential use of such a weapon.

- **Workshop 3 (May 2021) – AI & Surveillance**. How can we protect human rights in the age of AI-driven surveillance technologies? This workshop explored AI-driven surveillance technologies deployed by governments and corporations to better understand the ethical values and tradeoffs in various use contexts. While our case study focused on facial recognition in domestic policing, we also considered use of surveillance by non-democratic regimes and U.S. foreign policy goals of counteracting authoritarian surveillance tools.

Based on a holistic assessment of the workshop series—including feedback from participant surveys after each workshop—we believe that the model that we piloted can be effective, viable, and merits further refinement. The model also involves a number of

challenges that need to be approached deliberately and thoughtfully by institutions wishing to deliver rigorous co-designed AI ethics workshops for future policymakers. We discuss these points in more detail below.

Opportunities to Replicate and Scale

Piloting interdisciplinary workshops on AI ethics has been worthwhile and rewarding. Based on our experience, doing so successfully requires dedication, flexibility, and patience on the part of institutional partners, as well as deep substantive expertise in ethics and technical and policy fields. Successful workshops also require an ability to apply abstract

ethical concepts to granular real-world scenarios in ways that are engaging for participants with varied backgrounds and starting points.

In our view, effective workshop replication involves a number of ingredients, including:

- **Expertise in relevant fields.** Philosophers trained in ethics—and technology ethics more specifically—should ideally be paired with experts in the technical dimensions of AI and policymaking. Being able to draw on expertise in all of these fields is crucial to develop nuanced case studies, plan workshop content, and help participants acquire new insights.

- **Expertise in translational ethics methodology.** The aim of these workshops was to introduce theoretical yet agile ethical frameworks to help future leaders navigate real-world decisions related to building, deploying, and governing emerging technologies. These learning goals require methodologies that help scaffold "translational ethics"—ethical content and skills at stake in highly granular, highly constrained, and often strategic-level decision-making. Ethics Lab pursues this goal with methods inspired by design thinking; other methodologies are possible as well.

- **Expertise in workshop facilitation.** Leading workshops in ways that are compelling, turn participants' contributions into useful takeaways, and balance didactic goals with discussion requires a background in multi-disciplinary workshop facilitation. These skills are especially important when experts from different fields play a role in facilitating a workshop.

- **Cooperative, open-minded attitude.** For a project team with members from various disciplines to function well, a willingness to better understand each other's starting points and assumptions, to adjust flexibly, and adopt an attitude of experimentation about goals and tactics, are all important for a productive and enjoyable working partnership.

- **Dedication of time & effort.** Designing and delivering successful workshops requires considerable time for research, brainstorming, prototyping, and iterative back-and-forth between sub-teams on the project. Sometimes, backing up a few steps is helpful to make progress. Ensuring sufficient lead time to conceptualize, draft, and re-draft workshop plans is especially important, leaving enough time for fresh perspectives and tweaks.

While the workshop model that we piloted is replicable in principle, keeping the above points in mind, it may admittedly be difficult to scale. Achieving many of the goals outlined above will likely require Principal Investigators (and project teams) who already have many of the relevant skills and experiences noted above and are motivated to collaborate in sustained ways.

The PIT-UN Network could explore a variety of avenues to support organizations interested in hosting inter-disciplinary AI ethics workshops. For example, members could be given opportunities to learn more about facilitation methodologies, including tips and best practices, as well as opportunities to practice those skills and/or see them in action in demos or workshops.

6. Summary & Contact Information

Our pilot project generated a promising model with potential to build a strong cross-disciplinary cohort, particularly when in-person gatherings can resume. Although this project is completed, we are considering pursuing other funding opportunities to rerun the workshop series.

We would be glad to answer questions and provide input for organizations thinking about delivering workshops of the kind that we have piloted. We can be reached at ethicslab@georgetown.edu.

7. Workshop Materials

Workshop 1: AI & Justice
- [Key Takeaways](#)
- [Run of Show](#)
- Collateral
  - o [Talkback Prompts](#)
  - o [Interest Bearers Worksheets](#)
  - o [Group Questions](#)
  - o [Questions & Criteria Worksheet](#)
  - o [Markup Brochure (for 11x17)](#)
  - o [Markup Brochure (36x54)](#)
- Images
  - o [Group Markup](#)
  - o [Group Discussion](#)
  - o [Group Worksheets](#)

Workshop 2: Ethics of Risk and AI in National Security
- [Key Takeaways](#)
- [Run of Show](#)
- Collateral
  - o Briefing Materials
    - ▪ [Covering Note + Scenario](#)
    - ▪ [Further Details](#)
    - ▪ [Video — Scenario](#)
    - ▪ [Video — Background Concepts 1](#)
    - ▪ [Video — Background Concepts 2](#)
  - o Miro Activity Boards
    - ▪ [Pre-Workshop Talkbacks](#)
    - ▪ [Workshop Boards](#)
    - ▪ [PDF of Group Worksheets](#)
- Images
  - o [Screenshot of Miro action](#)
  - o [Screenshot of a group worksheet](#)

Workshop 3: Surveillance Technologies & Human Rights

- [Key Takeaways](#)
- [Run of Show](#)
- Collateral
    - [Jamboard](#)
    - [PDF of Jamboard Worksheets](#)
- Images
    - [Screenshot of warm-up exercise (1)](#)
    - [Screenshot of warm-up exercise (2)](#)
    - [Screenshot of small group case study work (1)](#)
    - [Screenshot of small group case study work (2)](#)
    - [Screenshot of small group case study work (3)](#)