

Towards Ethical, Fair, and Prosocial Multiagent Systems

Social AI and MAS within ISL

Nirav Ajmeri

with Alex Davies, Dan Collins, Jessica Woodgate, Joseph Trevorrow, and Yining Yuan (and Kal Roberts and Vishal Joshi)

AI and Collective Dynamics Units
Intelligent Systems (Federated) Labs
University of Bristol
<https://uob-mas.github.io/>

May 2025

Research Focus

Encompass intelligent agents and normative multiagent systems, multiagent reinforcement learning, and graph-based approaches

- ▶ Facilitate rich interactions between
 - ▶ Computation entities, such as software agents or embodied agents
 - ▶ Social entities, such as humans and organisations
- ▶ Ensure alignment between individual goals and values, and societal norms
- ▶ Promote prosocial behaviour ensuring equitable outcomes

People @ Social AI and Multi-Agent Systems



Daniel Collins
responsibility



Jessica Woodgate
normative ethics



Yining Yuan
normative systems



Joseph Trevorrow
value alignment



Alex Davies
graph neural networks



Kal Roberts
gameplay, marl



Vishal Joshi
language-based agents



Lucas Fadden
sustainability

Understanding, Embedding, and Sustaining Ethics: Agents and Society

Understand ethics

- ▶ Goals, values, preferences
- ▶ Social norms
- ▶ Ethical principles

Embed ethics

- ▶ Learn desirable behaviours
- ▶ Operationalise principles

Sustain ethics

- ▶ Sanctions, normative signals
- ▶ Explaining behaviours

Thank You

Web: <https://uob-mas.github.io>

- ▶ Agents Reading Group (Dan; bi-weekly on Wednesday @ 1630)
- ▶ AI Lunch & Learn (Wednesdays @ 1330; Join via ISL webpage)
- ▶ Data Ethics Club (Jess; bi-weekly on Wednesday @ 1300; <https://dataethicsclub.com/>)
- ▶ Citizen-Centric AI Systems Seminars (Nirav and Jess, once-a-month, <https://sites.google.com/view/ccais-seminar/home>)
- ▶ AI4CI ECR activities (time varies; bi-weekly)