

## Education

---

### **M.Sc. Computer Science**, 2019

GPA 3.93

University of California Berkeley

Thesis: *"Inverse Reinforcement Learning for Dynamics"*

Advisor: Anca Dragan

### **B.A. Interdisciplinary Studies**, Technology and Inequality, 2018

Major GPA 3.97

### **B.A. Computer Science**, 2017

Major GPA 3.87

University of California Berkeley

Interdisciplinary Thesis: *"Serving Two Masters: An Expeditionary Attempt at Critical Technical Practice in Human Compatible Artificial Intelligence"*

Advisor: Shreeharsh Kelkar

High Honors, Phi Beta Kappa

## Research Interests

---

Technology and Inequality, AI and Justice, AI Policy, Human Compatible AI, Human Robot Interaction, Critical Technical Practices, Deep Reinforcement Learning, Value Sensitive Design, Science and Technology Studies, AI Pedagogy

## Research Positions

---

### **Interact Lab**, Berkeley, California

Graduate Researcher, May 2018 - May 2019

Developing algorithms for generating robust human models given goals and demonstrations.

### **Center For Human-Compatible AI**, Berkeley, California

Undergraduate Researcher, January 2017 - December 2017

- Worked on mechanism design guarantees and efficient algorithms for Negotiable Reinforcement Learning
- Explored applications of Rawlsian justice to AI/ML.

### **Berkeley School of Information**, Richmond, California

Data Analyst and Camp Counselor, June 2016 - August 2016

Served as a part-time camp counselor for Connected Camps and analyzed the chat and server logs for an ongoing research project.

### **Political Economy of the UC Research Group**, Berkeley, California

Web Design Research Assistant, January 2016 - May 2016

Designed a web tool that clarifies students' financial aid distributions.

## Grants and Affiliations

---

### **Center for Longterm Cybersecurity**, Berkeley, California

Research Grantee, January 2019 - Present

Conducting cross-disciplinary historical analysis of professionalization and pedagogy formation around issues such as security, safety, and public interest to inspire cybersecurity and AI/ML guidelines.

## **Algorithmic Fairness and Opacity Working Group**, Berkeley, California

*Affiliated Graduate Researcher*, August 2018 – Present

Working on proposals for *just* applications of AI/ML in the context of institutional decision-making.

## **Center For Human-Compatible AI**, Berkeley, California

*Affiliated Graduate Researcher*, May 2018 – Present

Building out interdisciplinary perspectives on Value Alignment and Legible Normativity.

## **Center For Science, Technology, Medicine, and Society**, Berkeley, California

*Undergraduate Liaison*, January 2016 – December 2017

Worked with CSTMS and a student group (SEE) to share STS perspectives with undergraduate engineering majors.

## **Presentations, Proceedings, and Papers**

---

**Andrus, M.** (2019, upcoming) *Serving Two Masters: An Expeditionary Attempt at Critical Technical Practice in Human Compatible Artificial Intelligence*. Paper presented at Society for Social Studies of Science (4S) Annual Meeting.

**Andrus, M.**, Gilbert, T.K., Boots, B. (2019) *Bias in Datasets and Fairness in Machine Learning*. Breakout session presented at Code For America Summit.

**Andrus, M.**, Gilbert, T.K. (2019) *Towards a Just Theory of Measurement: A Principled Social Measurement Assurance Program*. In Proceedings of the 2019 conference on AI Ethics and Society. Selected for oral presentation (15% of all submitted papers).

Hadfield-Menell, D., **Andrus, M.**, Hadfield, G. (2019) *Legible Normativity for AI Alignment: The Value of Silly Rules*. In Proceedings of the 2019 conference on AI Ethics and Society. Selected for oral presentation (15% of all submitted papers).

**Andrus, M.** (2019) *On Serving Two Masters: Directing Critical Technical Practice towards Human-Compatibility in AI*. Extended Abstract In Proceedings of the 2019 conference on AI Ethics and Society.

## **Outreach**

---

**GEESE** (Graduates for Engaged and Extended Scholarship in Engineering), January 2018 - Present  
Founding member, contributing to publications and events organization that are centered around developing interdisciplinary and socially motivated engineering pedagogies.

**BAIR AI4ALL Camp**, August 2017, August 2018

A summer camp for high-schoolers of underrepresented demographics in AI. Worked as a mentor for the 2017 camp, and helped organize and develop the curriculum for the 2018 camp.

**Connected Camps**, July-August 2016

Taught elementary- to middle-school aged children about programming and AI through Minecraft mods.

**Socially Engaged Engineers**, September 2015 - May 2017

Helped organize multiple events for the 2016 and 2017 Engineering Ethics Weeks and worked on the curriculum for our DeCal, a student led course for credit.

## Event Organization (selected)

---

### **Tech Skepticism and Worker Activism: Drawing Lessons for Research and Education**

Panel at UC Berkeley, 12/12/2019

A multi-panel event looking at recent crises in the tech industry (e.g. Google Walkout, Project Maven) to reflect on what graduate students can do now to prepare to be ethical researchers and educators.

### **Can Technology Solve Big Problems?**

Panel at UC Berkeley, 4/18/2016

An interdisciplinary panel for Berkeley's Engineering Ethics Week that looked critically at when technology and engineering are suited to solve problems and when other disciplines need to be engaged.

## Work Experience

---

### **Digital Democracy**, Oakland, California

*Software Engineering Intern*, June 2016 - August 2016

Helped develop a new front end for the Mapeo Desktop application.

### **Berkeley Student Cooperative**, Berkeley, California

*House Manager, Maintenance Manager, Network Manager*, August 2014 - December 2017

Worked in numerous manager positions at Davis House Cooperative, mediating conflicts, ensuring the safety and habitability of the building, and maintaining the network infrastructure.

### **Acha Carro**, Topanga, California

*Web Design Intern*, June 2014 - August 2014

Developed advertisements and began work on a mobile version of the website.

## Skills and Attributes

---

Fluent in Python (Tensorflow, Keras) and Java

Varying degrees of familiarity with too many disciplines.

Love for travel - recently made it to my 60<sup>th</sup> country!