

Employment

- 2023–Pres **Research Scientist**, *Google*, London, UK.
Google Research; Cerebra Group.
- 2021–2023 **Bioethicist**, *Google*, Palo Alto, CA, US.
Google Health; Quality, Regulatory, Safety, and Trust Group.
- 2020–2021 **Postdoctoral Research Fellow**, *Stanford University*, Palo Alto, CA, US.
McCoy Family Center for Ethics in Society and Institute for Human-Centered AI.
- 2019–2020 **Research Assistant**, *University of Cambridge*, Cambridge, UK.
Leverhulme Centre for the Future of Intelligence.

Education

- 2017–2020 **PhD in Philosophy**, *University of Bristol*.
No Corrections. Supervisors: Richard Pettigrew and Brad Hooker.
- 2016–2017 **MA in Philosophy**, *University of Bristol*.
Distinction in all Components. Ranked 1st out of a class of 24.
- 2013–2016 **BSc in Philosophy, Logic and Scientific Method**, *London School of Economics*.
First Class Honours. Ranked joint-1st out of a class of 51.

Research

- AOS Ethics of Artificial Intelligence, Machine Learning, and Data Science
- AOC Moral Philosophy, Decision Theory, Philosophy of Science

Papers

- 2022 **Grote, T. and Keeling, G.**, "Enabling Fairness in Healthcare Through Machine Learning" in *Ethics and Information Technology* 24(3), pp.1-13.
- Keeling, G.**, "Automated Vehicles and the Ethics of Classification" in *Autonomous Vehicle Ethics: The Trolley Problem and Beyond*, R. Jenkins et al. (eds), OUP.
- Keeling, G. and Nyrup, R.**, "Explainable Machine Learning, Patient Autonomy, and Clinical Reasoning" in *Oxford Handbook of Digital Ethics*, C. Véliz (ed), OUP.
- Grote, T. and Keeling, G.**, "On Algorithmic Fairness in Medical Practice" in *Cambridge Quarterly of Healthcare Ethics* 31(1), pp.83-94.
- Keeling, G. and Burr, C.**, "Digital Manipulation and Mental Integrity" in *The Philosophy of Online Manipulation*, M. Klenk and F. Jongepier (eds), Routledge.
- Keeling, G., and Paterson, N.**, "Proper Functions: Etiology Without Typehood" in *Biology & Philosophy* 37(19), pp.1-17.

- 2021 **Bommasani, R., Hudson, D., ... , Keeling, G., ... , and Liang, P.**, "On the Opportunities and Risks of Foundation Models", arxiv.org/abs/2108.07258.
Keeling, G., "A Dilemma for Reasons Additivity" in *Economics & Philosophy* , pp.1-23.
- 2020 **Keeling, G.**, "Why Trolley Problems Matter for the Ethics of Automated Vehicles" in *Science and Engineering Ethics*, 26, pp.293-307.
- 2019 **Keeling, G., Evans, K., Thornton, S.M., Mecacci, G., Santoni de Sio, F.**, "Four Perspectives on What Matters for the Ethics of Automated Vehicles" in *Road Vehicle Automation 6*, Lecture Notes in Mobility, Springer, pp.49-60.
- 2018 **Keeling, G.**, "Legal Necessity, Pareto Efficiency and Justified Killing in Autonomous Vehicle Collisions" in *Ethical Theory and Moral Practice*, 21, pp.413-427.
Keeling, G., "Against Leben's Rawlsian Collision Algorithm for Autonomous Vehicles" in *Philosophy and Theory of Artificial Intelligence*, Müller, V. (ed) Springer, pp.259-272.

Talks & Panels

- 2022 **Moral Imagination: A Method for Culture Change in Technology Development**, with Amanda McCroskery, Renee Shelby, Sandra Blascovich, and Ben Zevenbergen.
 NeurIPS Workshop on Cultures in AI/AI in Culture
Panel Discussion: Fairness in Machine learning, with Sam Corbett-Davies.
 Neukom Institute for Computational Science, Dartmouth College (Invited)
Algorithmic Bias, Generalist Models, and Clinical Medicine.
 Workshop on the Ethics of Influence, University of Oxford (Invited)
Lossy Operational Values: The Case of Fair Machine Learning.
 Technology and Human Values Conference, Google (Invited)
Ethical Considerations: Synthetic Characters and Geriatric Loneliness.
 Responsibility and Autonomy in AI, University of Oxford (Invited)
Autonomous Vehicles: Safely Navigating Low-Probability High-Impact Events.
 Philosophy and AI Seminar, Facebook/Meta AI (Invited)
- 2021 **Legal Necessity and Justified Killing in Autonomous Vehicle Collisions.**
 Center for Human-Compatible AI, UC Berkeley (Invited)
Enabling Fairness in Healthcare through Machine Learning.
 Stereotyping and Medical AI Colloquium, King's College London (Invited)
Digital Manipulation and Mental Integrity, with Christopher Burr.
 Philosophy of Online Manipulation Book Workshop
A Dilemma for Reasons Additivity.
 Political Theory Workshop, Stanford University (Invited)
- 2020 **On Algorithmic Fairness in Medical Practice**, with Thomas Grote.
 Workshop on Algorithmic Fairness, University of Copenhagen
Decision Support Systems and Clinical Reasoning, with Rune Nystrup.
 Leverhulme Centre for the Future of Intelligence, University of Cambridge (Invited)
 Philosophy of Medical AI, University of Tübingen (Invited)

- 2019 **Proper Functions: Etiology Without Typehood**, *with Niall Paterson*.
 Philosophy of Biology and Cognitive Sciences Workshop, U. of the Basque Country
 British Society for the Philosophy of Science Annual Conference, University of Durham
Treatment Recommender Systems and Informed Consent, *with Rune Nyrup*.
 Issues in Explainable AI, Saarland University (Invited)
Object Classification and the Reasonable Belief Standard.
 Automated Vehicles Symposium, Orlando FL (Invited)
 Research Institute for Ethics and Law, Swansea University (Invited)
 2nd Centre for Ethics, Philosophy and Public Affairs (CEPPA) Graduate Conference
 on Moral and Political Philosophy, University of St Andrews
 iCog5: Approaches to Higher Cognitive Function, University of Reading
- 2018 **Uncertainty About Persons in Autonomous Vehicle Collisions**.
 Automated Vehicles Symposium, San Francisco, CA (Invited)
Panel Discussion: Ethical and Social Implications of Autonomous Vehicles, *with
 Nicholas G. Evans, Noah Goodall, Ryan Jenkins, and Katherine Evans*.
 Automated Vehicles Symposium, San Francisco, CA (Invited)
Building Machines that Learn and Think About Morality, *with Christopher Burr*.
 Philosophy After AI Symposium, Annual Convention of the Society for the Study of
 Artificial Intelligence and the Simulation of Behaviour, University of Liverpool
- 2017 **Against Leben's Rawlsian Algorithm for Autonomous Vehicles**.
 Philosophy and Theory of Artificial Intelligence, University of Leeds
 Artificial Ethics Symposium, University of Southampton
Reliability Weighted Belief Revision in Peer Disagreements, *with Nemo D'Qrill*.
 Graduate Conference on Epistemology of Disagreement, University of Tartu
 Ninth European Congress of Analytic Philosophy, LMU Munich
Blame in Autonomous Vehicle Collisions.
 10th Symposium on Computing and Philosophy, Annual Convention of the Society for
 the Study of Artificial Intelligence and the Simulation of Behaviour, University of Bath
- 2016 **Neuroscience, Criminal Liability and the Voluntary Act Requirement**.
 Sheffield Undergraduate Philosophy Conference, University of Sheffield
The Evolution of Moral Responsibility.
 Bristol Undergraduate Rationality Research Group, University of Bristol
- 2015 **Evolutionary Biology, Moral Intuitions and the Is-Ought Problem**.
 3rd Annual LSE-Bayreuth Undergraduate Conference, LSE
 Reading Undergraduate Philosophy Conference, University of Reading

Service

- 2023–Pres **Programme Committee: Philosophy After AI Symposium**.
 Society for the Study of Artificial Intelligence and the Simulation of Behaviour

- 2022-Pres **Programme Committee: Fairness Accountability and Transparency (FAccT).**
Association for Computing Machinery (ACM).
- 2021-2022 **Programme Committee: Explanatory AI: Between Ethics and Epistemology.**
Issues in XAI Series: Delft University of Technology.
- 2021-2021 **Programme Committee: Bias and Discrimination in Algorithmic Decisions.**
Volkswagen Foundation BIAS Research Initiative, Leibniz University, Hannover.
- 2019–2020 **Programme Committee: Philosophy After AI Symposium.**
Society for the Study of Artificial Intelligence and the Simulation of Behaviour
- 2018–2019 **President.**
British Postgraduate Philosophy Association
Executive Committee.
British Philosophical Association

Conference Organization

- 2021 **Issues in XAI-2: Explaining and Understanding in Medicine.**
Leverhulme Centre for the Future of Intelligence, University of Cambridge
- 2019 **Joint Biology and Philosophy Workshop.**
University of Bristol; Keynotes: Samir Oksaha, Mark Beaumont and Tom Williams
- 2018 **British Postgraduate Philosophy Association Annual Conference.**
University of Birmingham; Keynotes: Alison Jaggar and Veronique Munoz-Darde
- 2016 **British Undergraduate Philosophy Society Spring Conference.**
University of Cambridge; Keynote: Richard Holton
- 2015 **British Undergraduate Philosophy Society Autumn Conference.**
London School of Economics; Keynote: Tim Williamson
British Undergraduate Philosophy Society Spring Conference.
University of Southampton; Keynotes: Mike Otsuka and Helen Beebee
- 2014 **British Undergraduate Philosophy Society Autumn Conference.**
Birkbeck, University of London; Keynotes: M.M. McCabe and Onora O'Neill

Teaching

Stanford University

- 2021 **Ethics, Law and Politics of Artificial Intelligence.**
BS/MS Primary Instructor. Selected Topics: Intelligence explosion hypothesis and AGI safety, local model-agnostic explanations and counterfactual explanations, political microtargeting, fairness impossibility theorems, and ethics of autonomous systems.

University of Bristol

- 2020 **Political Philosophy.**
2nd Year BA Seminar Tutor. Selected Topics: Political obligation, distributive justice, boundaries of the state, democracy, hate speech, feminist political theory.

- 2019 **What is Democracy?.**
2nd and 3rd Year BA Guest Seminar Tutor. Topic: Arrow's impossibility theorem.
- 2018-2019 **Research Skills for MSc Migration and Mobility Studies.**
MSc Seminar Tutor (2018) and Guest Lecturer on Statistics (2019). Selected Topics: Interviews, ethnography, statistical methods, textual analysis, project planning.
- 2017-2018 **Introduction to Formal Logic.**
1st Year BA, BSc and MSci Seminar Tutor. Selected Topics: Syntax and semantics for propositional and first-order logic, natural deduction, soundness and completeness.
- [Guest Lectures](#)
- 2023 **Fair Machine Learning: Theory and Practice.**
BA/BS Lecture, School of Public and International Affairs, Princeton University.
- 2022 **Fair Machine Learning.**
Executive AI Principles Ethics Fellowship, Google.
- 2021 **Black Boxes, Explainability, and the GDPR.**
JD/LLM Lecture, Faculty of Law, Chinese University of Hong Kong.
- 2020 **Ethics of Medical Decision Support Systems.**
MA Lecture, Department of Philosophy and Religious Studies, Utrecht University.

[Awards](#)

- 2017 **Arts and Humanities Research Council PhD Studentship, ca. £55,000.**
Awarded through the South, West and Wales Doctoral Training Partnership.
- 2016 **Andrea Mannu Prize, London School of Economics.**
Joint best performance across all undergrad philosophy programmes (joint-1st/51).

[Wider Engagement](#)

- 2023 **Lecture: Ethics and Language Models.**
Penningtons Manches Cooper/Ethical Reading Public Event
- 2022 **Careers Talk: Practical Ethics in Silicon Valley.**
Uehiro Centre for Practical Ethics, University of Oxford
- Teaching AI Ethics: Lessons from Stanford and Google.**
National Symposium on Socially Responsible STEM, City, University of London
- 2021 **Careers Talk: Practical Ethics in Tech.**
Let's Phi: Beyond Academic Philosophy
- Lecture: AI Ethics and the Long Term Future.**
creAltivity Workshop for High School Students
- Philosophy Careers Outside Academia.**
MA Seminar, Department of Philosophy, University of Bristol.
- Lecture: Trolley Problems and Automated Vehicles.**
Ethical Reading, Reading, UK

- 2020 **Lecture: Medical Reasoning in the Age of Artificial Intelligence.**
Milton Keynes Artificial Intelligence Expert Forum
- 2019 **Panel Discussion: Artificial Intelligence in Healthcare**, with Kourosh Saeb-Parsy, Slawomir Nasuto, and Weizi Vicky Li.
Austin Vita/Ethical Reading Public Event on AI in Healthcare
- 2018 **Lecture: The Ethics of Automated Vehicles.**
St Edward's School, Oxford

Public Policy

- 2020-21 I have consulted for the UK Law Commission Automated Vehicle Review, alongside the UK Government's Centre for Connected and Automated Vehicles and the UK Government's Centre for Data Ethics and Innovation.

Peer Review

Referee for *Nature*, *Journal of Applied Philosophy*, *Utilitas*, *Journal of Ethics*, *Synthese*, *Journal of Medical Ethics*, *Ethics and Information Technology*, *Science and Engineering Ethics*, *Philosophy and Technology*, *Res Publica*, *AI and Society*, *IEEE Transactions on Technology and Society*, *IEEE Intelligent Transportation Systems Transactions*, *IEEE Access*, *Virtual Reality*, and *International Journal of Technoethics*. I have also refereed for ACM FAccT, NeurIPS and AAAI/ACM Conference on AI, Ethics and Society, alongside the MIT Press.