Jake Fawkes

jake.fawkes@st-hughs.ox.ac.uk | +44 7923 101627 | Jakefawkes | Oxford, UK

Education

Sep 2025 – Sep 2020 | **PhD in Statistical Machine Learning** | Department of Statistics, University of Oxford
- Research at the intersection of **machine learning** and **causal inference** with three main strands , **machine learning** to estimate and test for **causal effects**,**causality** for **algorithmic fairness**, and understanding the **mathematical structure** of **causal inference**.
- Supervised by Prof. Robin Evans and Prof. Dino Sejdionvic.

Jul 2020 – Sep 2016 | **MMath in Mathematics** | Jesus College, University of Oxford
- Focused on pure mathematics courses in Algebra, Functional Analysis and Combinatorics for the Bachelors section of the course before switching to a mix of statistics machine learning and pure mathematics in the masters year.
- Achieved first class honours all years, with an average year mark of 75.4 and final year mark of 77.4.

Skills

Software: Python, Git, LaTeX | ML Package: PyTorch, Pandas, Scikit-learn, Numpy, Stan | Research Interests: Causal machine learning, Algorithmic Fairness, AI ethics, Robustness, Domain Generalisation, Multi-Enviroment Learning, Category Theory for ML.

Publications and Working Papers

Jan 2025 | **The Hardness of Validating Observational Studies with Experimental Data**
Published at the International Conference on Artificial Intelligence and Statistics (AISTATS) 2025
Jake Fawkes, Michael O'Riordan, Athanasios Vlontzos, Ciarán M Gilligan-Lee

Jan 2025 | **Is Merging Worth It? Securely Evaluating the Information Gain for Causal Dataset Acquisition** code | pdf
Published at the International Conference on Artificial Intelligence and Statistics (AISTATS) 2025
Jake Fawkes*, Lucile Ter-Minassian*, Desi Ivanova, Uri Shalit, Chris Holmes

Oct 2024 | **The Fragility of Fairness: Causal Sensitivity Analysis for Fair Machine Learning** code |pdf
Published at Neural Information Processing Systems (Neurips) 2024. Dataset & Benchmarks track.
Jake Fawkes*, Nic Fishman*, Mel Andrews, Zachary C. Lipton

July 2024 | **Doubly Robust Kernel Statistics for Testing Distributional Treatment Effects**
Published in Transactions on Machine Learning Research (TMLR) 2024 code |pdf
Jake Fawkes, Robert Hu, Robin J. Evans, Dino Sejdinovic

May 2024 | **The Role of Learning Algorithms in Collective Action**
Published at the International Conference of Machine Learning (ICML), 2024 pdf
Omri Ben-Dov*, **Jake Fawkes***, Samira Samadi, Amartya Sanyal

June 2023 | **Returning The Favour: When Regression Benefits From Probabilistic Causal Knowledge** code |pdf
Oral at the International Conference of Machine Learning (ICML), 2023.
Shahine Bouabid*, **Jake Fawkes***, Dino Sejdinovic

Feb 2022 | **Selection, Ignorability and Challenges with Causal Fairness** pdf
Oral at Conference on Causal Learning and Reasoning (CLeaR), 2022.
Jake Fawkes, Robin J. Evans, Dino Sejdinovic.

June 2023 | **Predicting Disease Onset from Electronic Health Records for Population Health Management: a Scalable and Explainable Deep Learning Approach** pdf
Published in Frontiers in Artificial Intelligence edition on Transforming Population Health with P4 Healthcare.
Robert Grout, Rishab Gupta, Ruby Bryant, Mawada A. Elmahgoub, Yijie Li, Khushbakht Irfanullah, Rahul F. Patel, **Jake Fawkes**, and Catherine Inness

June 2023 | **Results on Counterfactual Invariance** pdf
Poster at ICML workshop on Spurious Correlations, Invariance and Stability, 2023.
Jake Fawkes, Robin J. Evans.

Present | **A Unifying Category for Causal Interventions with An Intuitive Axiomatisation**
Working Paper
Jake Fawkes, Robin Evans, Rob Cornish

Experience

Sep 2025 – Mar 2025 | **Research Science Intern** | **Valence Labs**
Project: Detecting double gene knockouts in genetic perturbation data
- Project applying causal representation learning methodology to single cell data. Working under the supervision of Dr Jason Hartford and Dr Kristina Ulicna.

Aug 2024 – June 2024 | **Research Science Intern** | Causal Inference Lab, **Spotify**
Project: Combining Observational And Experimental Data to Improve Causal Effect Estimation
- Worked to how **A/B tests** can be combined with **large observational datasets** to improve estimation of causal effects. Written up in conference length submission to **AiStats**.
- Follow up on improved **Neural Net Architectures** for learning from experimental data (**Pytorch**).
- Supervised by Dr. Ciarán Gilligan-Lee.

Dec 2023 – Oct 2023 | **Visiting Researcher** | Rational Intelligence Lab, CISPA – Helmholtz Center for Information Security
Project: Data Curation for Model Robustness and Out of Distribution Generalisation
- Research project aiming to understand how **subsampling** and **data curation** methods can be applied to improve **robustness** properties of machine learning models and enable **out of distribution generalisation**.
- Project inspire by recent work demonstrating CLIP models are more **robust** within image classification tasks primarily due to the quality of training data.
- Self proposed project for which a Helmholtz visiting researcher grant was awarded.
- Supervised by Dr. Krikamol Muandet.

Apr 2023 – Oct 2022 | **Responsible AI Research Intern** | **Accenture** and **Alan Turing Institute**, UK
- Focused on the application of Responsible AI methodology within healthcare settings, including **auditing model fairness**, producing **model card** templates, and advising on **explainability**.
- First project involved working in a team designing **NLP** solutions to processing **electronic health records** for **disease prediction**. Worked within a team of data scientists as the sole Responsible AI consultant, advising on **machine learning fairness**, **model interpretability** and **transparency**.
- Second project involved working within one of Accenture's wider Responsible AI teams to design a **Python package** for a large **pharmaceutical company** to **audit** their machine learning models. Worked in a team of software engineers to select, design, and code auditing functions.
- Internship through the **Alan Turing Institute**, supervised at Accenture's by **Ali Shah**.

Awards

Feb 2023 | **Helmholtz Visiting Researcher Grant**
Sep 2020 | **ESPRC studentship for DPhil in Statistics and Machine Learning**
2017-2020 | **Lawrence Scholarship for Exam Performance from Jesus College**

Service and Outreach

Reviewer
- Reviewed for UAI (2022 - 2025), FAccT (2025), TMLR, and AFCP (2022).

Ongoing – Sep 2023 | **Statistical Consultant** | Cure Parkinsons
- Assisted in the presentation and analysis of data from a Cure Parkinson's study into the unique challanges faced by menopausal women with Parkinson's.

Sep 2022 – Jun 2022 | **Mfano Africa Mentor** | University of Oxford
Project: Variational Autoencoders for Fraud Detection
- Provided mentorship for a masters level student from the Ivory Coast for a 3 month research project.
- Met weekly to discuss the students ideas, relevant papers and code, and to guide research directions.
- Project was assessed via a research presentation and project write-up which the student passed.

Invited Talks

Jul 2023 | **Returning The Favour: When Regression Benefits From Probabilistic Causal Knowledge** | International Conference of Machine Learning

Sep 2022 | **Returning The Favour: When Regression Benefits From Probabilistic Causal Knowledge** | CISPA – Helmholtz Center for Information Security

Sep 2022 | **Problems due to Selection and an Over-commitment to Ignorability in Causal Fairness** | Pacific Causal Inference Conference 2022

Apr 2022 | **Selection, Ignorability and Challenges with Causal Fairness** | CLeaR