Jacek Karwowski

jacek.karwowski@cs.ox.ac.uk ⋄ (+48) 731 77 22 64

EDUCATION

Ph.D. in Computer Science	*2022 - now*
University of Oxford	
Specialization: Probability theory and category theory	
M.Sc. in Pure Mathematics	*2018 - 2021*
University of Warsaw	
Thesis: Formal semantics of a reversible language	
B.Sc. in Computer Science	*2018 - 2021*
University of Warsaw	
Thesis: Low-latency Apache Parquet library (blogpost)	
B.Sc. in Mathematics	*2015 - 2018*
University of Warsaw	
Thesis: Products in positive opetopic sets	

PUBLICATIONS

V. Choudhury, J. Karwowski, and A. Sabry, "Symmetries in reversible programming: from symmetric rig groupoids to reversible programming languages," *Proceedings of the ACM on Programming Languages*, vol. 6, pp. 1–32, Jan. 2022

N. Ackerman, C. E. Freer, Y. Kaddar, J. Karwowski, S. Moss, D. Roy, S. Staton, and H. Yang, "Probabilistic Programming Interfaces for Random Graphs: Markov Categories, Graphons, and Nominal Sets," *Proceedings of the ACM on Programming Languages*, vol. 8, pp. 61:1819–61:1849, Jan. 2024

R. Douglas, J. Karwowski, C. Bae, A. Draguns, and V. Krakovna, "Limitations of Agents Simulated by Predictive Models," Feb. 2024. arXiv:2402.05829 [cs]

J. Karwowski, O. Hayman, X. Bai, K. Kiendlhofer, C. Griffin, and J. Skalse, "Goodhart's Law in Reinforcement Learning," Oct. 2023. arXiv:2310.09144 [cs]

ACADEMIC

Facilitator - Introduction to AI Alignment, Oxford AI Safety Society	Hilary Term 2023
Class Tutor - Lambda Calculus and Types, University of Oxford	Hilary Term 2023
Reviewer - 35th International Conference on Computer Aided Verification	2023
PhD Scholarship - Improving the Long-Term Future, Open Phillantropy	2022 - 2025
Scientific Program Officer - PL in ML Conference	2017
Scientific Lead - Student ML research group, University of Warsaw	2015 - 2016

WORK EXPERIENCE

Research Scholar - Stanford Existential Risks Initiative 2023

As a part of SERI ML Alignment Theory Scholars program, I worked with Victoria Krakovna at DeepMind and others on understanding how can powerseeking arise in predictive models (a paper is currently under review).

Researcher - Oxford AI Safety Labs 2022 - 2023

AI Safety Labs is a research program intended to introduce students to the technical research on AI safety. I worked with Joar Skalske at the Future of Humanity Institute and others on understanding Goodhart's law in reinforcement learning (a paper is currently under review).

Quantitative Developer - Xantium 2021 - 2022

Xantium is the algorithmic trading division of Tudor Corporation. I worked on volatility trading, where I built models, implemented execution and managed datasets. Unfortunately, details of my work there are classified under an NDA.

SWE Intern - G-Research 2020

· G-Research is a quantitative research company, developing algorithms for predicting movements in financial markets. I worked on programmers tools for the simulations platform. Unfortunately, details of my work there are classified under an NDA.

Research Intern - Indiana University 2019

· As a part of Global Talent Attraction Program, I conducted research on understanding the semantics of reversible programming languages, in the context of Homotopy Type Theory (paper).

SWE Intern - Microsoft 2019

· I worked on Office365 back-end, adding new features and optimizing the existing ones, and deploying code to production. Details of my work there are classified under an NDA.

Research Intern - Golem Factory GmbH 2017 - 2018

· Golem is a blockchain startup, creating a platform for sharing computational resources using Ethereum network. I developed a probabilistic verification algorithm for programs executed in an untrusted environment, and implemented it for the use-case of hyper-parameter optimization for neural networks (whitepaper, blogpost, another blogpost).

Data Scientist - Applica.ai 2016 - 2017

· Applica is a startup working on ML tools for NLP. I was responsible for creating new models: from reading papers and designing experiments to prototyping models in Python and implementing them on the production cluster.

Student Intern - University College London 2014 - 2015

· Over the academic year, I have been working on a platform for testing high-frequency trading algorithms. Then, during the summer, I worked on a open-source medical data platform.