Katarzyna (Kasia) Kobalczyk

knk25@cam.ac.uk | kasia-kobalczyk.github.io | github.com/kasia-kobalczyk
Trinity College, CB2 1TQ Cambridge, UK | +44 7934 433342

PhD student in the ML and AI lab of Prof. Mihaela van der Schaar at the University of Cambridge.
Research interests: large language models, probabilistic machine learning, meta-learning, sample-efficient decision-making,
.

EDUCATION

UNIVERSITY OF CAMBRIDGE PhD in Applied Mathematics and Theoretical Physics
2023 - 2027 (Expected) | Cambridge, UK
Research in the ML and AI lab of Prof. Mihaela van der Schaar.

UNIVERSITY OF CAMBRIDGE MASt in Mathematical Statistics
2022 - 2023 | Cambridge, UK | Grade: Honours with Merit
Part III essay titled: *Meta-learning for multimodal task distributions* , supervised by Prof. Sergio Bacallado.

UNIVERSITY OF WARWICK BSc in Mathematics and Statistics
2019 - 2022 | Coventry, UK | Grade: First Class Honours
Ranked 1st in the Statistics student cohort; 92.5%, 89.3%, 89.4% yearly grade averages.

WORK EXPERIENCE

META Machine Learning Research Scientist Intern
Summer 2024 | Menlo Park, CA, United States
- Internship with the Adaptive Experimentation Team with a focus on Bayesian Optimisation.

CITADEL Quantitative Research Intern
Summer 2023 | London, UK
- Developing and back-testing quantitative trading strategies for index rebalance.
- Portfolio optimization with conic programming.

G-RESEARCH Summer Research Intern
Summer 2022 | London, UK
- Change-point analysis and anomaly detection for high-dimensional time-series data.

SCHRODERS Data Insights Unit Intern
Summer 2021 | London, UK
- Application of Bayesian item-response theory to survey data analysis.
- Explainability and error sensitivity research in mean-variance portfolio optimisation.

SHELL PLC Data Science Intern
Summer & Winter 2020 | Krakow, Poland
- Clustering analysis and decision tree learning for customer segmentation.
- Modelling and analysis of traffic data with spatio-temporal networks
- Machine learning for fraud detection.

PUBLICATIONS

Conference & Journal papers:
1. **Kobalczyk**, K., & van der Schaar, M. (2025a). *Preference Learning for AI Alignment: a Causal Perspective* . Forty-Second International Conference on Machine Learning

2. Pouplin*, T., **Kobalczyk** *, K., & van der Schaar, M. (2025). *LLMs for Generalizable Language-Conditioned Policy Learning under Minimal Data Requirements* . Forty-Second International Conference on Machine Learning. https://arxiv.org/abs/2412.06877 **[Spotlight]**
3. **Kobalczyk**, K., & van der Schaar, M. (2025b). *Towards Automated Knowledge Integration From Human-Interpretable Representations* . The Thirteenth International Conference on Learning Representations. https://openreview.net/forum?id=NTHMw8S1Ow **[Spotlight]**
4. **Kobalczyk** *, K., Astorga*, N., & van der Schaar, M. (2025). *Active Task Disambiguation with LLMs* . The Thirteenth International Conference on Learning Representations. https://openreview.net/forum?id=JAMxRSXLFz **[Spotlight]**
5. Walley, G., Shenvi, A., Strong, P., & **Kobalczyk**, K. (2023). *cegpy: Modelling with chain event graphs in Python* . Knowledge-Based Systems, 274, 110615. https://doi.org/10.1016/j.knosys.2023.110615

Workshops & Pre-prints:

1. **Kobalczyk** *, K., Fanconi*, C., Sun, H., & van der Schaar, M. (2025). *Few-shot Steerable Alignment: Adapting Rewards and LLM Policies with Neural Processes* . 2nd Workshop on Models of Human Feedback for AI Alignment (MoFA) at ICML 2025. https://arxiv.org/pdf/2412.13998 **[Oral]**
2. Piskorz*, J., **Kobalczyk** *, K., & van der Schaar, M. (2025). *Eliciting Numerical Predictive Distributions of LLMs Without Auto-Regression* . Workshop on Foundation Models for Structured Data (FMSD) at ICML 2025
3. Laguna*, S., **Kobalczyk** *, K., Vogt, J. E., & van der Schaar, M. (2025). *Interpretable Reward Modeling with Active Concept Bottlenecks* . Workshop on Programmatic Representations for Agent Learning (FMSD) at ICML 2025
4. **Kobalczyk**, K., & van der Schaar, M. (2024). *Informed Meta-Learning* . SPIGM and MFHAIA workshops at ICML 2024. https://arxiv.org/abs/2402.16105v4
5. Zhu, M., **Kobalczyk**, K., Petrovic, A., Nikolic, M., van der Schaar, M., Delibasic, B., & Lio, P. (2023). *Tabular Few-Shot Generalization Across Heterogeneous Feature Spaces* . https://arxiv.org/abs/2311.10051

* Equal contribution.

AWARDS

Warwick Statistics Prize | 2023
Awarded for the best overall performance in the Mathematics and Statistics degree

Institute of Mathematics and its Applications (IMA) Prize | 2023
Awarded for an outstanding performance in mathematics-oriented subjects

Outstanding Academic Excellence Prize | 2022
Awarded for the best 2nd year examination results in the Department of Statistics

SKILLS

PROGRAMMING LANGUAGES
Proficient: Python
Intermediate: R
Basic: Stan, SQL, Bash, Matlab

FRAMEWORKS & SOFTWARE
PyTorch, Hugging Face, Scikit-learn,
Weights & Biases, Hydra,
Git, Conda, Docker, JAX

SPOKEN LANGUAGES
Native fluency: English, Polish
Intermediate: German