Derck Prinzhorn

derckprinzhorn@gmail.com | derckprinzhorn.com | GitHub | Scholar | LinkedIn

Education

University of Amsterdam Sep 2023 – present
MSc Artificial Intelligence *Amsterdam, Netherlands*
Thesis (in progress): AI Control, supervised by Maksym Andriushchenko

University of Amsterdam Sep 2020 – Jun 2023
BSc Artificial Intelligence *Amsterdam, Netherlands*
Thesis: Benchmarking Conformal Prediction Methods for Time Series Regression

Work Experience

Prinzhorn Solutions, *Founder* Apr 2025 – present
Working with companies to help them better understand and manage risks associated with adopting AI systems.

Wisr, *Co-Founder* Sep 2024 – Oct 2025
Worked on a startup helping teachers save time with AI grading.

Dutch Police, *AI Architect* Apr 2023 – Apr 2025
Defined reference architectures for AI, MLOps, and AI security.

University of Amsterdam, *Teaching Assistant* Sep 2022 – Sep 2023
Machine Learning, Reinforcement Learning, Bayesian Statistics, Data Structures & Algorithms.

LeerLevels, *Software Engineer* Oct 2021 – Jan 2023
Developed grading algorithms, search engines, and recommendation systems.

Research Experience

Max Planck Institute for Intelligent Systems & ELLIS Institute Tübingen, *Research Intern* Jan 2026 – present
Research on AI control under supervision of Maksym Andriushchenko.

Aithos Research Lab, *Research Engineer* Apr 2025 – Jan 2026
Worked on evals for AI value systems and moral competence.

University of Amsterdam, *Research Intern* Jan 2024 – Feb 2025
Conformal prediction for time series; 3D diffusion models for radiotherapy dose prediction; physics benchmarking in video generation models.

Supervised Program for Alignment Research (SPAR), *Research Intern* Jul 2024 – Oct 2024
AI control and safety research with Aryan Bhatt (Redwood Research); studied monitor improvement methods.

Publications

C. Zhang, D. Cherniavskii, A. Tragoudaras, A. Vozikis, T. Nijdam, **D. Prinzhorn**, et al. Morpheus: Benchmarking Physical Reasoning of Video Generative Models with Real Physical Experiments. *arXiv preprint arXiv:2504.02918*, 2025. (Under review)

T. Nijdam, **D. Prinzhorn**, J. de Heus, T. Brouwer. HIVE: A Hyperbolic Interactive Visualization Explorer for Representation Learning. *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 655–660, 2025.

D.W.E. Prinzhorn, T. Nijdam, P.A. van der Linden, A. Timans. Conformal Time Series Decomposition with Component-wise Exchangeability. *13th Symposium on Conformal and Probabilistic Prediction with Applications (COPA 2024)*; *Proceedings of Machine Learning Research (PMLR)*, 2024.

G. de Jong, M.J. Meijer, **D.W.E. Prinzhorn**, H. Ruiter. Reproducibility Study of FairAC. *Transactions on Machine Learning Research (TMLR)*, 2024. Poster at NeurIPS 2024.

Awards

AmsterdamAI Thesis Award — Bachelor thesis on conformal prediction for time series (2023)

Volunteering

Stichting Hoormij | Board Advisor & Quality Council	Jan 2021 – present
Forward Incubator | Startup Mentor	Dec 2024 – May 2025
AI Safety Amsterdam (AISA) | Member	Sep 2023 – May 2025
Dutch Nao Team | ML Engineer & Vice Chair	Sep 2022 – Mar 2024
Programme Committee AI, UvA | Student Member	Sep 2021 – Apr 2023
Stichting Studiezalen | Mentor	Feb 2020 – Oct 2021