Julian Skirzyński

Curriculum Vitae — March 2025

jskirzynski@ucsd.edu
www.jskirzynski.com

Education

University of California, San Diego 2022 – Present
Ph.D. Candidate in Computer Science & Engineering
Thesis: Empirical Toolbox for Evaluating and Redesigning Human-AI Collaboration
Advisor: Berk Ustun

McGill University 2017 – 2020
M.S. in Computer Science
Thesis: Language-Conditional Imitation Learning
Advisor: David Meger

University of Warsaw 2012 – 2018
M.S. in Cognitive Science
B.S. in Mathematics, Cognitive Science
Advisors: Andrzej Skowron (Mathematics); Piotr Wasilewski (Cognitive Science)

Academic Positions

Max Planck Institute for Intelligent Systems, Tübingen, Germany 2019 – 2023
Researcher
Led a project on interpretable RL policies and interventions that use them to aid human planning.
Advisor: Falk Lieder

Research Interests

Areas: Machine Learning, Cognitive Science, Human-Computer Interaction
Topics: Decision-Making, Interpretability, Explainability, Reinforcement Learning, Experimental Design, Human-AI Collaboration
Applications: Social Sciences, Medicine, Consumer Finance, Criminal Justice

Awards & Honors

Pierre Arbour Foundation Scholarship 2018 – 2019
McGill University Graduate Excellence Award 2018
McGill - University of Warsaw Exchange Scholarship 2015
University of Warsaw Academic Excellence Scholarship 2014 – 2017

Preprints

[1] On the Value of Interpretability in Human Decision-Making
Julian Skirzyński, Elena Glassman, Berk Ustun
In Submission, 2025

[2] Discrimination Exposed? On the Reliability of Explanations for Discrimination Detection
Julian Skirzyński, Davind Danks, Berk Ustun
In Submission, 2025

Papers

*equal contribution

Google Scholar

[3] Automatic Discovery and Description of Human Planning Strategies
Julian Skirzyński, Yash Raj Jain, Falk Lieder
Behavior Research Methods, 2023

[4] Boosting Human Decision-making with AI-Generated Decision Aids
Frederic Becker*, **Julian Skirzyński***, Bas van Opheusden, Falk Lieder
Computational Brain & Behavior, 2022

[5] Automatic Discovery of Interpretable Planning Strategies
Julian Skirzyński, Frederic Becker, Falk Lieder
Machine Learning, 2021

[6] Object [Re] Cognition with Similarity
Łukasz Sosnowski, **Julian Skirzyński**
International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, 2018

[7] A Framework for Analysis of Granular Neural Networks
Julian Skirzyński
International Joint Conference on Rough Sets, 2017

Refereed Workshop Papers

[8] On Interpretability and Overreliance
Julian Skirzyński, Elena Glassman, Berk Ustun
Interpretable AI: Past, Present and Future, *NeurIPS Workshop*, 2024

[9] Language-Conditional Imitation Learning
Julian Skirzyński, Bobak Baghi, David Meger
Visually Grounded Interaction and Language, *NAACL Workshop*, 2021

Teaching

DSC291 – Interpretability & Explainability in Machine Learning 2023
Teaching Assistant
Co-designed curriculum and held weekly office hours for serving 20+ PhD/MS students
Delivered guest lectures on ML interpretability methods and cognitive biases in AI-assisted decision-making
Completed teaching development workshop on graduate-level instruction

Software

GitHub
Strategy Extraction from RL Policies – Algorithm to extract interpretable decision trees from RL policies
Human Planning Strategy Analysis – Framework for identifying strategies used in human planning tasks

Selected Industry Positions

Educational Entertainment One, Warsaw, Poland 2021 – 2024
Lead Technical Architect
Designed algorithms (AI, NLP) and supported the production process for a story-driven mobile game for learning English.

Academic Service

Journal Reviewing

Machine Learning 2022

Conference Program Committee

NeurIPS – Annual Conference on Neural Information Processing Systems 2023 – 2025
ICML – International Conference on Machine Learning 2025
ICLR – International Conference on Learning Representations 2024
FAccT – ACM Conference on Fairness, Accountability and Transparency 2022 – 2025
ICML Workshop RL4RealLife – International Conference on Machine Learning 2021
IPMU – Information Processing and Management of Uncertainty in Knowledge-Based Systems 2018

Personal

Language Skills : English, Polish, German (Conversational)
Software Skills : Python, R, C++, Flask, AWS, PyTorch, CPLEX, JavaScript, Jira
Interests : Soccer, Groundhopping, Traveling, Fantasy Literature, Record Collecting
Other : Peer tutoring, Co-author of "Triozy polskie", a textbook for learning Polish by foreigners