

Hiroshi Nonaka

Phone: +1 (949) 685 7039
Email: hnonaka@soka.edu
Aliso Viejo, CA

Website: hironona.github.io/
GitHub: github.com/hironcode
LinkedIn: linkedin.com/in/hiroshi-nonaka/

INTERESTS

Representation Learning in Language Models, Vision Models, and Model-based Reinforcement Learning;
Language Grounding in World Models

EDUCATION

B.A., Soka University of America (SUA), Aliso Viejo, CA 08/2022 – 05/2026
Social and Behavioral Sciences with emphasis on Mathematics and Computer Science

- Topics: Machine Learning
- Relevant Coursework:
Intro to Computer Science | Algorithms and Data Structures | Linear Algebra | Intro to Probability and Statistics (in progress) | Calculus 2 | Multivariable Calculus | Real Analysis (in progress) | Regression Analysis | Differential Equations | Discrete Mathematics | Abstract Algebra | Cognitive Psychology (in progress) | Evolutionary Psychology | Principles of Economics | International Economics | Physics in Lagrangian Mechanics, Relativity, and Quantum Mechanics

Advanced Coursework in Machine Learning, The University of Tokyo
World Models | Deep Reinforcement Learning | Deep Generative Models | Global Consumer Intelligence/GCI (Statistical Machine Learning)

Extended Coursework, University of California, San Diego
Probability and Statistics for Machine Learning (Spring 2025)

Study Abroad, National Taiwan University 08/2024 – 12/2024
Chinese Language Study

EXPERIENCES AND PROJECTS

Research Assistant at Relational Cognition Lab, University of California, Irvine 09/2025 – Present
Supervisors: Dr. Matthew Shinkle & Dr. Anna Leshinskaya

- Apply mechanistic interpretability approaches to study moral reasoning in LLMs.

Presidential Research Assistant in Mathematics, Soka University of America 09/2023 – 09/2025
Supervisor: Dr. Katherine Perry

- Developed a program that extracts social networks from 1,200+ short stories generated by LLMs.
- Automated story generation pipeline with Gemini and OpenAI APIs.
- Integrated a RoBERTa-based sentiment classifier to predict relationship polarity.
- Our paper has been accepted to the NeurIPS 2025 WS on Evaluating the Evolving LLM Lifecycle: Benchmarks, Emergent Abilities, and Scaling and the EMNLP 2025 WS on Wordplay.

REU-CAAR Research Fellow, University of Maryland, College Park
Supervisor: Aviva Prins

06/2025 – 08/2025

- Model-free reinforcement learning in non-stationary Markov decision processes.
- Led the development of new *restart* frameworks, *adaptive* and *selective* restarts.
- Achieved near-optimal, 91% less dynamic regret than the baseline in an MDP designed to be challenging for exploration.
- Our paper has been accepted to the NeurIPS 2025 WS on Aligning Reinforcement Learning Experimentalists and Theorists (ARLET).

Research Fellow, Soka Institute for Global Solutions

10/2024 – 05/2025

- Selected as an undergraduate fellow at the Soka Institute for Global Solutions.
- Explored the application of the Free Energy Principle and Constructivism in international relations and formulated the concept of Variational Constructivism.
- Presented research at the SIGS symposium in Spring 2025.

World Models Research Project Member, The University of Tokyo

11/2024 – 01/2025

- Investigated spatiotemporal understanding of pre-trained world models using linear and non-linear probing.
- Managed the training pipeline of DreamerV3 (Hafner et al., 2024) and built a program collecting activations from the model.

Summer Research Fellow, Texas State University

06/2024 – 07/2024

Supervisor: Dr. Damian Valles

- Developed a new VLM architecture for emotion recognition in conversations based on LLaMA3-Instruct-8B.
- Improved the F1 score by 58% from the unimodal LLaMA3 on benchmarks.
- Our research paper was accepted by the IEEE UEMCON conference in 2024.

GENIAC Support Member, Ministry of Economy, Trade, and Industry of Japan

04/2024 – 06/2024

- Partook in GENIAC, an LLM development project initiated by METI of the Japanese government.
- Formulated weekly summaries of the data collection and curation teams' progress.
- Created fine-tuning data of 37,361 Japanese characters ($\approx 14,945$ words in English).

Inaugural Member of AI Economic and Social Theory Group, University of Tokyo 10/2023 – 04/2024

Instructor: Professor Keita Nishiyama

- Six-month-long discussion group on the intersection of AI, philosophy, linguistics, sociology, and physics.
- Delivered a final presentation on the impact of AI on liberal arts education.

PUBLICATIONS

- **Hiroshi Nonaka***, Simon Ambrozak*, Sofia R. Miskala-Dinc†, Amedeo Ercole†, and Aviva Prins, “Efficient Restarts in Non-Stationary Model-Free Reinforcement Learning,” *Under review at Workshop on Aligning Reinforcement Learning Experimentalists and Theorists at 39th Conference on Neural Information Processing Systems (NeurIPS)*, San Diego, CA, USA, 2025.
- **Hiroshi Nonaka** and K. E. Perry, “Evaluating LLM Story Generation through Large-scale Network Analysis on Social Structures,” *Under review at Workshop on Evaluating the Evolving LLM Lifecycle at 39th Conference on Neural Information Processing Systems (NeurIPS)*, San Diego, CA, USA, 2025.

- **H. Nonaka** and D. Valles, “Fully Auto-Regressive Multi-modal Large Language Model for Contextual Emotion Recognition,” *2024 IEEE 15th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, Yorktown Heights, NY, USA, 2024, pp. 0291-0299, doi: 10.1109/UEMCON62879.2024.10754673.

POSTER AND PRESENTATIONS

- S. Ambrozak, **H. Nonaka**, A. Ercole, S. Miskala-Dinc, and A. Prins, “Be Less Optimistic: Tuned optimism & efficient restarts in model-free RL,” *Final Presentation for REU-CAAR*, College Park, MD, USA, 2025.* (Presentation)
- **H. Nonaka**, S. Ambrozak, S. Miskala-Dinc, A. Ercole, and A. Prins, “Learning to Let Go: Efficient Restarts and Calibrated Optimism in Model-Free Reinforcement Learning,” *Summer Undergraduate Research Conference*, College Park, MD, USA, 2025. *(Poster Presentation)
- **H. Nonaka**, A. Imadate, and T. Ogawa, “Investigating the Understanding of Spatio-temporal Information in World Models,” *Final Project Presentation Session*, Bunkyo City, Tokyo, Japan, 2025.* (Poster Presentation)

HONORS & AWARDS

- Soka Ikeda Scholarship 08/2025
The most prestigious award given to one student in the class for academic excellence and leadership, which covers the entire cost of attendance.
- Award of Excellence in Research 08/2025
Award given to only one student at Soka University of America based on faculty nominations.
- Toshizo Watanabe Scholarship 05/2025
Grant of \$7,500 awarded to selected applicants.
- Soka Institute for Global Solutions Student Fellowship 10/2024
- Soka Merit Award 08/2024
Scholarship awarded to the top 3 students in the class who demonstrate academic excellence.
- Award of Excellence in Academics 08/2024
Award given to only one student at Soka University of America based on faculty nominations.
- Pacific Basin Research Center Summer Grant 05/2024
Grant awarded to selected students for their participation in research programs and activities.
- Best Idea Award 05/2024
Award of \$1,000 given to the top selected teams at a business pitch competition, The Fowler Global Social Innovation Challenge 2024.
- Soka Merit Award 08/2023
- Outstanding Student Award 07/2023
Selected as one of the top 1 % of students in Global Consumer Intelligence at the University of Tokyo.

TEACHING EXPERIENCES

- **Introduction to Data Science.** Soka University of America. Fall 2025. Tutor under Dr. Hassan Attarchi.
- **Introduction to Computer Science.** Soka University of America. Spring 2024 & Spring 2025. Tutor under Dr. Katherine Perry.

ACTIVITIES

- Club President**, Interdisciplinary AI 02/2025 – Present
- Club promoting the interdisciplinary understanding of artificial intelligence.
 - Organized an on-campus discussion event in March 2025.
- Organizer**, Nest of Mathematicians 04/2024 – Present
- [Webpage](#) for students to share STEM content.
 - Created featured lists of free online courses in machine learning and summer research programs to enhance student access to those fields.
- Club Secretary**, Soka Student Movement for Culture of Peace 09/2022 – 04/2024
- Club promoting “Culture of Peace,” which was proposed by the United Nations, on-campus.
 - Created the design format of the [club website](#).
 - Co-hosted events on social media and culture of peace and on liberal arts education and culture of peace.

CERTIFICATES

- Undergraduate (NSF) Basic Course. CITI Program. 06/2024
- Greedy Algorithms, Minimum Spanning Trees, and Dynamic Programming. Stanford. 05/2023
- Graph Search, Shortest Paths, and Data Structures. Stanford. 04/2023
- Divide and Conquer, Sorting and Searching, and Randomized Algorithms. Stanford. 04/2023

SKILLS

- **Computer Languages:** Python, JavaScript, HTML/CSS, SQL
- **Relevant Libraries:** PyTorch, Transformers, Scikit-learn, NumPy, Pandas, SciPy, Matplotlib, Seaborn, TensorFlow, SpaCy, NetworkX
- **Tools:** Linux, Git, GitHub, Hugging Face, Slurm, Gemini API, OpenAI API

LANGUAGES

- English: Professional (TOEFL iBT 112: Reading 29, Listening 27, Speaking 29, Writing 27)
- Japanese: Native
- Mandarin Chinese: Intermediate