

Yoshihiro Izawa

izawa-yoshihiro146@g.ecc.u-tokyo.ac.jp | github.com/Omusubi0123 | www.linkedin.com/in/yoshihiro-izawa-104262343/

Saitama, Japan

Portfolio Website: <https://www.omusubiy.com/>

RESEARCH INTERESTS

Large language models and foundation models, especially mechanistic interpretability and the understanding of internal representations, behaviors, and reasoning mechanisms of learned systems toward more transparent and reliable AI.

EDUCATION

The University of Tokyo, Tokyo, Japan April 2026 – March 2028 (Expected)

Master of Information Science and Technology, Information Communication Engineering

The University of Tokyo, Tokyo, Japan April 2023 – March 2026

Bachelor of Engineering, Information Communication Engineering

Grad Thesis: “[Steering at the Source: Style Modulation Heads for Robust Persona Control](#)”, Dean’s Award (Research)

PUBLICATION

Y. Izawa*, Gouki Minegishi, Koshi Eguchi, Sosuke Hosokawa, Kenjiro Taura,
“[Steering at the Source: Style Modulation Heads for Robust Persona Control](#)”, **ICML 2026**, ICLR 2026 WS

Y. Izawa*, Gouki Minegishi, Yoko Yamakata,
“[Anosognosia in LLMs: Probing Self-Awareness of Quantized Computational Substrate](#)”, Preprint (2026)

WORK HISTORY

LY Corporation, Tokyo, Japan August 2026 – September 2026

Machine Learning Engineer Intern (Visual Understanding Team)

- Additional Training of MLLMs in Japanese

The University of Tokyo Hospital, Tokyo, Japan September 2025 – September 2026

Technical Support (Division of Cardiovascular Medicine, AI Development Team)

- Translated medical requirements from physicians into technical specs and developed codebase.
- Implemented information extraction pipelines from medical documents using vLLM.
- Trained Masked Autoencoder (MAE) model for various medical downstream tasks.

Akari Inc, Tokyo, Japan August 2023 - December 2025

Student Intern (Algorithm Engineer)

- Trained domain-specific Embedding and Reranker models. ([technical blog](#))
 - Generated high-quality synthetic training datasets using local Large Language Models (LLMs).
 - Fine-tuned BERT-based models with parallel and distributed computing using DeepSpeed.
- Developed various algorithms leveraging LLM, including Search Agents and document retrieval systems.

PROJECTS

Generative AI Ideathon of SoftBank November 2025

Result: 11th out of 44,572 ideas

- Team Leader: Orchestrated the overall project, from core idea conceptualization to the presentation.
- Conceptualized a high-value solution that led to a successful patent filing by the organizers (SoftBank Cop.).

UGIP Data Hackathon 2025 (collaboration between The University of Tokyo and Softbank): March 2025

Result: 3rd prize

- Developed [TherapysToy](#), an AI-powered plush toy (Raspberry Pi) to maintain mental health of children.
- Team Leader: Specified mental health problems, polished the business model and presentation slide. Implemented backend and AI model services.

Personal DevOps Infrastructure October 2024 - Present

- Managing a dedicated Linux server at home for personal project hosting (backend and database servers).
- Leveraged Docker to containerize independent services, streamlining deployment and resource management.