Kazuki Yano

✉ yano.kazuki@dc.tohoku.ac.jp   🌐 kazuki-ya.github.io   in linkedin.com/in/kazuki-yano
🎓 Google Scholar

Summary

Ph.D. student at Tohoku University specializing in LLM pre-training dynamics and training efficiency. First-author of publications at top-tier venues including **ICLR 2026**, **COLM 2025**, and **NAACL 2025**. Experienced in end-to-end large-scale model training on high-performance compute clusters.

Education

Tohoku University Sendai, Japan
Ph.D. in Information Science (expected) 2025 − 2028
Supervisor: Jun Suzuki

M.S. in Information Science 2023 − 2025
Supervisor: Jun Suzuki

Skills

- **Research areas:** LLM pre-training, parameter-efficient training, training dynamics, evaluation/benchmarking

- **Tools:** Python, PyTorch, distributed training, Linux, Git, Slurm, Docker/Singularity

- **Language:** Japanese (native), English (professional working proficiency)

Experience

SB Intuitions — Research Intern Nov 2024 − April 2026

- Conducted LLM development on a large compute cluster, spanning pre-training and post-training.

- Studied training efficiency and dynamics through controlled experiments (baselines/ablations) and analysis.

Preferred Networks Inc. — Research Intern Aug 2024 − Mar 2025

- Built synthetic-data pipelines for LLM training; evaluated continued pre-training (CPT) and supervised fine-tuning (SFT).

- Quantified effects of data/training choices with baselines/ablations and targeted diagnostics.

Rakuten Group, Inc. — Data Scientist (Part-time) Jul 2024 − May 2025

- Studied representation/activation steering for style control via inference-time interventions.

- Evaluated controllability across styles and analyzed internal representations.

LINE Corporation — Research Intern Aug 2023 − Sep 2023

- Developed an efficient adversarial-prompt optimization method for safety stress-testing aligned dialogue LMs.

Publications

- **Kazuki Yano**, Shun Kiyono, Sosuke Kobayashi, Sho Takase, and Jun Suzuki. *Pre-training LLM without Learning Rate Decay Enhances Supervised Fine-Tuning.* Proceedings of the Fourteenth International Conference on Learning Representations (ICLR) 2026. paper.

- **Kazuki Yano**, Jun Suzuki, and Shinji Watanabe. *Adapting Text LLMs to Speech via Multimodal Depth Up-Scaling.* Under review at the Annual Conference of the International Speech Communication Association (Interspeech) 2026.

- Ryo Fujii, Makoto Morishita, **Kazuki Yano**, and Jun Suzuki. *TimeMachine-bench: A Benchmark on Evaluating Model's Capability on Repository-level Migration Tasks.* Proceedings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL) 2026. paper.

- Keigo Shibata, **Kazuki Yano**, Ryosuke Takahashi, Wataru Ikeda, and Jun Suzuki. *Suppressing Final Layer Hidden State Jumps in Transformer Pretraining.* Findings of the 19th Conference of the European Chapter of the Association for Computational Linguistics (EACL) 2026. paper.

- **Kazuki Yano**, Sho Takase, Sosuke Kobayashi, Shun Kiyono, and Jun Suzuki. *Efficient Construction of Model Family through Progressive Training Using Model Expansion.* Proceedings of the Second Conference on Language Modeling (COLM) 2025. paper.

- Wataru Ikeda, **Kazuki Yano**, Ryosuke Takahashi, Jaesung Lee, Keigo Shibata, and Jun Suzuki. *Layer-wise Importance Analysis of Feed-Forward Networks in Transformer-based Language Models.* Proceedings of the Second Conference on Language Modeling (COLM) 2025. paper.

- **Kazuki Yano**, Takumi Ito, and Jun Suzuki. *STEP: Staged Parameter-Efficient Pre-training for Large Language Models.* Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL) 2025. paper.

- Keito Kudo, Hiroyuki Deguchi, Makoto Morishita, Ryo Fujii, Takumi Ito, Shintaro Ozaki, Koki Natsumi, Kai Sato, **Kazuki Yano**, Ryosuke Takahashi, Subaru Kimura, Tomomasa Hara, Yusuke Sakai, and Jun Suzuki. *Document-level Translation with LLM Reranking: Team-J at WMT 2024 General Translation Task.* Proceedings of the Ninth Conference on Machine Translation (WMT) 2024. paper.

- **Kazuki Yano**, Takumi Ito, and Jun Suzuki. *STEP: Staged Parameter-Efficient Pre-training for Large Language Models.* Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL) Student Research Workshop 2024.

Honors and Awards

- **Outstanding Paper Award** 2025
  The 31st Annual Meeting of the Association for Natural Language Processing (NLP 2025)
  Awarded to top 14 / 765 submissions

- **BOOST Program for next-generation AI areas (JST)** 2025
  BOOST: Broadening Opportunities for Outstanding young researchers and doctoral students in STrategic areas
  Research Fellow (subsidized by JST)

- **Encouragement Award** 2025
  The 31st Annual Meeting of the Association for Natural Language Processing (NLP 2025)
  Awarded to top 20 / 487 eligible submissions

- **Encouragement Award** 2024
  The 30th Annual Meeting of the Association for Natural Language Processing (NLP 2024)
  Awarded to top 18 / 427 eligible submissions

Academic Service

- Reviewer: ARR, TMLR, ICML, ICASSP.