Qi Sun

lfsm.martin@gmail.com • Google Scholar Link

Education

Institute of Science Tokyo — **Ph.D. in Computer Science**
Advisor: Rio Yokota — *2024–Present*

Tokyo Institute of Technology — **M.S in Computer Science**
Advisor: Rio Yokota — *2022–2024*

Dalian University of Technology — **B.E in Computer Science**
2018–2022

Experience

Sakana AI — **Research Scientist (Founding Team)**
Built a leading Japan-based start up, RL for foundation models. — *2023.9–Present*

- **Sakana Fugu: Multi-Agent Orchestration System (Product)**
  Turned my research into a product offering an API service that dynamically coordinates frontier LLMs (GPT, Gemini, Claude) to achieve state-of-the-art performance, on tasks like GPQAD, LCBv6, and SWEPro. I led model training and infrastructure, overseeing end-to-end system design including distributed training, scalable serving with cloud infrastructure on Google Cloud Platform.
  URL: https://sakana.ai/fugu-beta/
- **Evolutionary Context Search for Skill Acquisition**
  We propose Evolutionary Context Search (ECS), a method that searches for optimal context to inject new knowledge into LLMs. ECS outperforms various RAG baselines, shows that task-driven context search is more effective than similarity-based approaches, and opens a new paradigm for skill acquisition.
  URL: https://arxiv.org/abs/2602.16113
- **Trinity: An Evolved LLM Coordinator (ICLR 2026)**
  We introduce a lightweight coordinator that orchestrates multiple diverse LLMs via a tri-role protocol, trained with an reinforcement learning. Trinity achieves state-of-the-art on LiveCodeBench and consistently outperforms all individual models it coordinates.
  URL: https://arxiv.org/abs/2512.04695
- **Transformers^2: Self-adaptive LLMs (ICLR 2025)**
  We propose Singular Value Fine-tuning (SVF), an fine-tuning technique with few parameters and graceful performance. Additionally, we introduce a novel self-adaptation framework enabling models to dynamically adjust their internal representations in response to diverse prompts, receving 1K+ star in github.
  URL: https://github.com/SakanaAI/self-adaptive-llms
- **Transformer Layers as Painters (AAAI 2025)**
  We conduct a comprehensive investigation into the structural and functional properties of pretrained transformer layers. Our analysis reveals remarkable insights about the robust performance characteristics of Large Language Models under various intervention strategies.
  URL: https://arxiv.org/abs/2407.09298
- **An Evolved Universal Transformer Memory (ICLR 2025)**
  We build a small neural network on top of pretrained LLM to intelligently identify and drop non-essential tokens, substantially enhancing long sequence processing capabilities while maintaining contextual coherence.
  URL: https://x.com/SakanaAILabs/status/1866286131685498920
- **Evolutionary Optimization of Model Merging Recipes (NMI)**
  We present a novel evolutionary approach to merging different foundation models, effectively combining their complementary capabilities. We also demonstrate emergent abilities in the resulting models.
  URL: https://x.com/SakanaAILabs/status/1770613032198279663

Tokyo Institute of Technology **Research Assistant**
Large-scale foundation model training, data generation, and evaluation. *2022.9–Present*

- Develop vision language model framework and set up training on 3k GPUs.
- Collect and filter Japanese language training corpus.

Large-scale Artificial Intelligence Open Network(LAION) **Contributor**
Large-scale strategy data generation. *2022.9–Present*

- Create 800B tokens strategy game dataset leveraging super-computer Fugaku.

Intel AI Lab **Intern**
Acceleration of alphaFold2 inference on CPU. *2021.9–2022.4*

- Achieve 23x faster inference speed up on intel xeon processor.
- Bring codebase from JAX/Haiku into Pytorch with Jit for better CPU performance.
- Use Intel Neural Compressor for low-precision acceleration.

Peer-Reviewed Publication

Xu, J., Q. Sun*., Schwendeman, P., Nielsen, S., Cetin, E., & Tang, Y.*
Trinity: An Evolved LLM Coordinator.
In Proceedings of the 14th International Conference on Learning Representations (ICLR 2026).

Nielsen, S., Cetin, E*., Schwendeman, P*., Q. Sun., Xu, J., & Tang, Y.*
Learning to Orchestrate Agents in Natural Language with the Conductor.
In Proceedings of the 14th International Conference on Learning Representations (ICLR 2026).

Q. Sun, Pickett, M*., Nain, A. K., & Jones, L. Transformer Layers as Painters.*
In Proceedings of the 39th AAAI Conference on Artificial Intelligence (AAAI 2025).

Q. Sun., Cetin, E*., & Tang, Y*. Transformer2: Self-adaptive LLMs.*
In Proceedings of the 13th International Conference on Learning Representations (ICLR 2025).

Cetin, E., Q. Sun, Zhao, T., & Tang, Y. An Evolved Universal Transformer Memory.
In Proceedings of the 13th International Conference on Learning Representations (ICLR 2025).

Akiba, T., Shing, M., Tang, Y., Q. Sun, & Ha, D. Evolutionary Optimization of Model Merging Recipes.
Nat Mach Intell 7, 195–204 (2025). https://doi.org/10.1038/s42256-024-00975-8.

Nakamura, T., Mishra, M., Tedeschi, S., et al.
Aurora-M: The First Open Source Multilingual Language Model Red-teamed according to the U.S. Executive Order.
In Proceedings of the 2025 International Conference on Computational Linguistics (COLING 2025).

Honors & Awards

Science Tokyo BOOST Scholarship, 2024 - 2027, 10,800,000 JPY

Monbukagakusho Honors Scholarship, 2022 - 2023, 300,000 JPY

China National Scholarship, 2018 - 2019, 8,000 CNY

Merit Student of Dalian University of Technology