NTT

NTT and Sakana AI sign a collaboration agreement for the R&D of AI constellations that will enable a society in which sustainable generative AI is used.
～Aiming to solve complex social problems through the collaboration of small, intelligent, and diverse LLM～

NTT Corporation (NTT) and Sakana AI Ltd. (Sakana AI) have signed a collaboration agreement for R&D of architectures in which small and diverse AI models cooperate with each other towards realizing a sustainable society. We aim to solve complex social problems by initiating a collaborative study of an architecture in which many AI models are connected and operate as an aggregate (AI constellation, see below).

1.Background

People worldwide anticipate that generative AI, like large language models (LLMs), will improve task efficiency and create new businesses. However, the development of generative AI necessitates significant computational resources, including power, thereby presenting a major challenge.

NTT has been researching natural language processing technology for over 40 years. NTT is also promoting R&D of IOWN technology, including photonics-electronics convergence technology. On November 1, 2023, NTT announced the launch of tsuzumi[1], a high-performance language model with a significantly reduced number of parameters.

Sakana AI was established in Tokyo in 2023 by David Ha and Llion Jones. Previously, David Ha served as the Head of Google Brain's Japan division, leading research on complex autonomous systems, while Llion Jones authored the paper "Attention Is All You Need[2]," proposing a "transformer[3]" model that sparked the current surge in generative AI. Sakana AI specializes in designing individual AI models and possesses expertise in AI architecture design, including the integration of multiple AI models.

NTT and Sakana AI will utilize their natural language processing technology and AI architecture design capabilities to solve the issue of rising computational complexity in large-scale AI models. This will be achieved by deploying multiple small AIs in a distributed configuration and linking them together efficiently. The focus is not only on the individual

models, but also on the architecture itself. We aim to create an AI architecture where small AIs connect based on their tasks and operate as an aggregate. This structure, which we call the "AI constellation", is based on the image of stars that are connected to form a constellation.

The AI constellation contributes to the reduction of environmental burdens because it uses only necessary AIs. Additionally, by having AIs autonomously cooperate with each other, we have high expectations for the creation of new collective intelligence. We intend to solve complex social problems by creating unprecedented knowledge and value through the AI Constellation.

２．Overview of the content of the collaboration agreement

Under this collaboration, NTT will provide natural language processing technology, while Sakana AI will provide expertise in AI model architecture design. Together, we will combine our strengths to research and develop the following:

(1) To develop AI constellation technology for connecting multiple small AIs.

(2) To build language models that create new value.

3．Endorsements

○NTT CTO, Katsuhiko Kawazoe

The NTT Group is developing IOWN technology to achieve a sustainable society. We're focusing on saving power via photonics-electronics convergence technology and other methods. We prioritize reducing the environmental load by making AI models with low computational load, and we are confident that by collaborating with Sakana AI, we can provide a new AI architecture not only domestically, but also globally. I am excited about a future where AI collaborates to develop new collective intelligence.

○Sakana AI CEO, David Ha

Sakana AI aims to develop distributed, autonomous, and efficient language models inspired by the laws of nature, where small fish in large swarms make superior behavioral decisions. Through collaboration with the NTT Group, we aim to embody the IOWN philosophy in AI technology development, creating a completely new approach.

※1 tsuzumi: NTT has developed tsuzumi, a language model that is lightweight yet provides world-leading Japanese language processing performance. Available in English and Japanese, tsuzumi can perform inference processing on a single GPU or CPU. It can also be extended to other modalities, such as visual and auditory. Moreover, it can be tuned specifically for certain industries and business organizations.

https://www.rd.ntt/research/LLM_tsuzumi.html

※2 Attention Is All You Need: A paper proposing a transformer model by a team led by Google researchers.
Proceedings of the 31st International Conference on Neural Information Processing Systems, December 2017, Pages 6000–6010.
https://dl.acm.org/doi/10.5555/3295222.3295349

※3 Transformer: This technology has significantly improved accuracy while requiring less text for AI to learn, compared to previous mainstream natural language processing technologies. It has greatly enhanced the precision of foreign language translation, sentence summarization, and sentence generation with chatbot-like capabilities. Currently, this technology is being implemented in speech and video processing, leading to a revolution in recent AI services.

About NTT

NTT contributes to a sustainable society through the power of innovation. We are a leading global technology company providing services to consumers and business as a mobile operator, infrastructure, networks, applications, and consulting provider. Our offerings include digital business consulting, managed application services, workplace and cloud solutions, data center and edge computing, all supported by our deep global industry expertise. We are over $95B in revenue and 330,000 employees, with $3.6B in annual R&D investments. Our operations span across 80+ countries and regions, allowing us to serve clients in over 190 of them. We serve over 75% of Fortune Global 100 companies, thousands of other enterprise and government clients and millions of consumers.

■Media contact
NTT Corporation
Research and Development Planning Department
nttrd-pr@ml.ntt.com

Sakana AI Ltd.
info@sakana.ai