

**NTT と Sakana AI、サステナブルな生成 AI 社会に向けた
AI コンステレーション研究で連携
～小さく賢い多様な LLM の集合により複雑な社会課題の解決をめざす～**

日本電信電話株式会社（本社：東京都千代田区、代表取締役社長：島田明、以下 NTT）と Sakana AI 株式会社（本社：東京都港区、CEO：David Ha、以下 Sakana AI）は、サステナブル社会の実現に向けて、小型で多様な AI 同士が協調するアーキテクチャの研究開発に関して協力することを定める連携協定を締結しました。多数の AI モデルがつながり、集合体として動作するアーキテクチャ（AI コンステレーション、後述）の共同検討を開始し、複雑な社会課題の解決をめざします。

1. 背景

現在、大規模言語モデル（LLM）をはじめとした生成 AI が、様々な業務を効率化し、新たな事業創出をもたらすことが期待されています。一方、その開発にあたっては電力を含む巨大な計算機資源が必要であり大きな課題となっています。

NTT は、40 年以上の自然言語処理技術研究を行ってきたことに加え、光電融合技術をはじめとした IOWN 先端技術の研究開発を推進しています。また、2023 年 11 月 1 日には、高度な性能を維持しつつ、パラメタ数を大幅に削減した言語モデル「tsuzumi^{*1}」の発表を行いました。

Sakana AI は、Google Brain の日本部門統括として、複雑系、自律システムの研究を指揮してきた David Ha 氏と、現在の生成 AI の爆発的な普及につながった「トランスフォーマー^{*2}」モデルの提案論文である「Attention Is All You Need^{*3}」の著者の 1 人である Llion Jones 氏が、2023 年に東京で立ち上げた AI 企業です。Sakana AI は、単一の AI モデルの設計スキルだけでなく、複数の AI を組み合わせるエージェント機能の設計など、AI アーキテクチャ設計スキルも有しています。

NTT と Sakana AI は、これまで培った自然言語処理技術や AI アーキテクチャの設計スキルに基づき、現状の大規模 AI モデルの計算量が増大していることの課題に対処していきます。多数の小型 AI を分散的に配置し、それら AI 同士を効率的に連携させることで、単一の AI モデルの省電力化を推進することに加えて、アーキテクチャ 構造自体からの省電力にも取り組んでいきます。わたしたちは、タスクに応じて必要な小型 AI がつながり、集合体として動作するアーキテクチャを、星がつながり星座になっていくイメージをもって、



「AI コンステレーション」と名付けました。

AI コンステレーションでは、環境負荷の低減だけでなく、AI 同士を自律的に協調させることで、新たな集合知が生まれる期待もあります。AI コンステレーションにより、今までにない、知識・価値を生み出すことで複雑な社会課題の解決をめざします。

2. 連携協定の概要

本連携では、NTT が有する自然言語処理技術と Sakana AI の AI モデル・アーキテクチャの設計スキルを組み合わせ、下記についての共同検討を開始します。

- (1) 複数の小型 AI を連携させる AI コンステレーション技術の創出
- (2) 新しい価値を創出する言語モデルの構築

3. エンドースメント

○NTT 代表取締役副社長 (CTO) 川添 雄彦

NTT グループは IOWN 技術を中心として、サステナブル社会の実現に取り組んでいます。光電融合技術などで省電力に取り組むだけでなく、計算負荷が低い AI モデルを作ることで、環境負荷を押し下げていくことは重要です。今回、Sakana AI 社と連携することで、国内のみならずグローバルに新しい AI アーキテクチャを提供していけると考えています。さらには、AI 同士が連携することで、新たな集合知が生まれていく未来にとっても期待しています。

○Sakana AI CEO, David Ha

Sakana AI は、小さな個々の魚が大きな群れを形成する中で優れた行動判断を行っていくような自然界の法則に着想を得て、分散的、自律的、効率的な新しい言語モデルの開発をめざしています。NTT グループとの連携を通じ、IOWN の理念を AI の分野でもより具体化し、こうした全く新しい技術の開発を進めてまいります。

※1 tsuzumi : NTT は、軽量でありながら世界トップレベルの日本語処理性能を持つ大規模言語モデル「tsuzumi」を開発しました。「tsuzumi」は英語と日本語に対応し、1GPU や CPU での推論動作を実現します。更に、視覚や聴覚といったモーダルに対応し、特定の業界や企業組織に特化したチューニングが可能です。

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

※2 トランスフォーマー：それまでに主流だった自然言語処理技術と比較して、AI に学習させる文章量が少ない一方で、精度向上を実現した技術。外国語の翻訳や文章の要約、チャットボットのような文章生成の精度を飛躍的に向上させた。音声や映像処理に活用する技術も提案されており、昨今の AI サービスに革新をもたらした技術。



※3 Attention Is All You Need : Google を中心とした研究チームから、トランスフォーマー※2 モデルが提案された論文。

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>

■本件に関する報道機関からのお問い合わせ先

日本電信電話株式会社

研究企画部門

nttrd-pr@ml.ntt.com

Sakana AI 株式会社

info@sakana.ai