

AIの進化とABEJAの歩み

ディープラーニングという技術革新が発表された 2012年に、ABEJAを設立
AIの進化とともに成長

ABEJAの歩み



AIの進化

2012	画像認識の競技会でトロント大学のグループがディープラーニングを用いて圧勝	2015	AIの画像認識精度が人間を上回る	2019	Googleが自然言語処理の革新的技術BERTを検索エンジンに導入	2022	OpenAIがChatGPTを公開 GPT-3.5、GPT-4が発表
2014	オックスフォード大学のマイケル Aオズボーン准教授らが、現在人間が担っている仕事の 47%は20年後になくなると発表 Amazonが世界初のスマートスピーカーを発売	2016	囲碁AIのAlphaGoが世界トップ棋士に勝利する	2020	スマートスピーカーの年間出荷台数が 1億台を突破	2024	OpenAIがSoraを発表
		2017	日本ディープラーニング協会設立、岡田代表取締役 CEOが理事に就任	2022	言語の理解度を図るベンチマークテストSuperGLUEのスコアでAIが人間を上回る	2024	OpenAIがo1 Previewを発表

NEDO公募のLLM開発事業(第二期)

経済産業省「GENIAC」の元、
NEDO公募の「競争力ある生成 AI基盤
モデルの開発」(第二期)に採択

- 取組内容:特化型モデル開発のためのモデルの小型化の研究開発
- 実施期間: 2024年10月～2025年4月
- 助成対象:主に計算リソースに係るコスト

GENIAC:日本の生成AIの開発力強化を目的としたプロジェクト

NEDO : 国立研究開発法人
新エネルギー・産業技術総合開発機構

- 持続可能な社会の実現に必要な研究開発の推進を通じて、イノベーションを創出する、国立研究開発法人

LLMの社会実装における課題

精度とコストの
トレードオフ



- LLMはサイズが大きいほど大規模な計算リソースが必要となり、コストがあがる
- サイズの大きい汎用的なモデルにおいても、用途に合わせたデータを用いて追加学習等のカスタマイズが必要となる
- 多くのケースでは、実用に向け、精度とコストのトレードオフが生じ、実運用可能な性能とコストを両立させることが求められる

第一期

日本語 LLM の構築と周辺技術 (RAG、Agent 機能) の研究開発

実施期間 2024年2月～8月

- 特定領域やタスクに応じた特化型の小型モデルの方が、コストや利便性等の面で有用なケースが少なくない
- 進化の早い LLM と社会実装に必要な周辺領域の研究を継続的に実施

第二期

特化型モデル開発のためのモデルの小型化の研究開発

実施期間 2024年10月～2025月4月

事業の状況

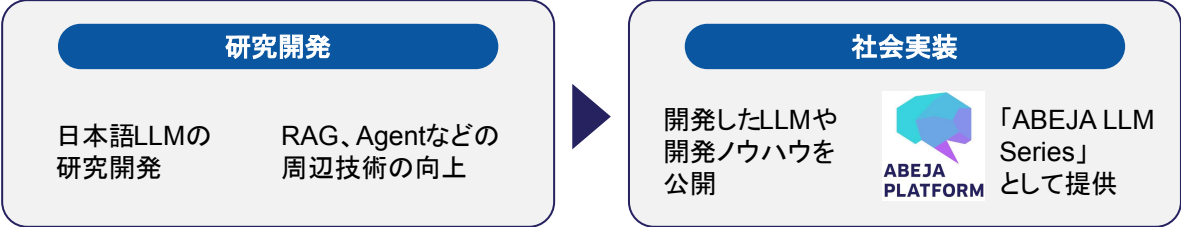
NEDO公募の「ポスト 5G情報通信システム基盤強化研究開発事業／ポスト 5G情報通信システムの開発」に、LLM 開発事業案が採択

- 取組内容: 日本語 LLM と周辺技術 (RAG、Agent 機能) の研究開発
- 実施期間: 2024 年 2 月～8 月
- 助成対象: 主に LLM 構築に必要な計算リソースに係るコスト

NEDO: 国立研究開発法人
新エネルギー・産業技術総合開発機構

- 持続可能な社会の実現に必要な研究開発の推進を通じて、イノベーションを創出する、国立研究開発法人

実施内容



LLM の課題 社会実装への制約要因

情報の質

- 知識のカットオフ
「既存」のデータに基づくため最新情報に未対応
- ハルシネーション
事実に基づかない誤情報を生成する可能性がある

利用コスト

- LLM の利用時に大規模な計算リソースを消費する
- 投資対効果を勘案すると、利用に制約が生じる

課題解決

RAG と Agent 機能の向上

精度及び計算コストパフォーマンスの向上、利用シーンの拡がりにより、LLM の社会実装を強力に推進



RAG

- LLM と外部データを結びつけ、外部データの知見を組み込んだ回答を生成できる技術
- 外部データを入れ替えるだけで、関連する高精度な回答を行うことが可能



Agent 機能

- 入力情報を基に LLM が自律的にアクションを計画・実行し、外部データを用いた回答が作成可能
- Agent 機能の向上を行うことで、LLM の適用範囲がより一層拡大