

モデル組み合わせによる日本語大規模基盤モデル開発と半自動データ作成フレームの構築

実施者	株式会社ELYZA
概要	日本語処理能力の高い高性能汎用基盤モデルを開発 • Mixture of Experts（MoE）、日本特有のデータの学習、フィードバック学習などを重点的に実施することで、推論効率の向上、日本特有の知識への回答精度の向上、ハルシネーションの抑制を目指す。 • 上記を通じ、開発した汎用基盤モデルの実用化および領域特化モデルの開発に繋げる。 社会実装に向け、自社プロダクトへの組み込み・API提供を進め、また一部モデル・ノウハウ等の成果物も公開・提供

実施内容

1 事前学習

2 指示学習

3 フィードバック学習

4 評価データセットの整備

5 コミュニティへの貢献

6 成果物の公開

開発される基盤モデル

グローバルモデルに匹敵する、日本語処理能力の高い汎用基盤モデルを開発

- MoE： 推論効率を向上
- 日本特有のデータの学習・フィードバック学習： 日本特有の知識への回答精度の向上・ハルシネーションの抑制

上記モデルの開発を踏まえ、領域特化モデルの開発に繋げる

汎用モデルの実用化
(本事業)

特化モデル開発
(本事業以降)

社会実装の方法

自社ビジネスとしての展開

- 自社既存プロダクトへの組み込み・API提供
- パートナー企業と共同で拡販

成果物の公開



ソースコード
評価などに関
して極力公開



モデル
事前学習済み



開発ノウハウ
事前学習部分
など

End-to-End音声基盤モデルの開発

実施者	株式会社Kotoba Technologies Japan
概要	End-to-Endで音声を入出力できる基盤モデルを構築、開発のインフラを整える - 大規模音声データセットを用いて、最大規模のEnd-to-End音声基盤モデルを学習 - 音声基盤モデルの評価パイプラインを整備 モデル学習のノウハウ / 学習済みモデルの一部を公開する、社会実装に向けた国内でのビジネス展開を行う - モデル学習のノウハウをテックブログ、コードベースなどの形で公開 - 学習済み小規模モデルも安全性を考慮した上で公開 - 開発した音声基盤モデルのAPIとしての公開、商用パッケージング

実施内容

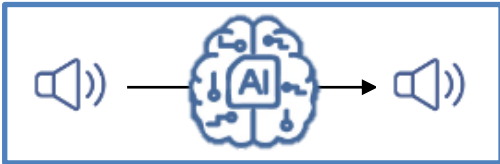
- ① データセットの構築
- ② 音声基盤モデル開発
- ③ 学習モデル評価
- ④ 知見の公開
- ⑤ コミュニティへの貢献
- ⑥ API化、商用パッケージング



開発される基盤モデル

音声AIの汎用性の強化に向けて、End-to-Endの音声基盤モデルの開発をスケールアップ：

- ・ **学習データ**：日英の大規模音声データセットを数十万時間活用
- ・ **モデル**：最大70億パラメータの音声モデルとしては最大規模のモデルを開発



音声の入出力をEnd-to-End基盤モデルで行う。

社会実装の方法

自社ビジネスとしての展開

- ・ **API提供**：開発した音声基盤モデルを企業・デベロッパーに向けて商用提供
- ・ **商用パッケージング**：音声基盤モデルをビジネスに向けてパッケージングして販売

成果物の公開



ソースコード
追加学習・推論用のコードなど



モデル
小規模モデルの重みを学術機関に向けて公開、企業にも条件付きで公開



開発ノウハウ
テックブログなどを通じてノウハウを公開(前処理、評価、分散並列学習など)

論理推論を可能とする大規模言語モデルの研究開発

実施者	富士通株式会社
概要	法務・金融・医療など規制規則への準拠が必要な業界でも生成AIを利用可能にするため、知識処理技術ナレッジグラフの生成と推論に特化した、以下2つの大規模言語モデル（LLM）を開発 ・ナレッジグラフ生成LLM）自然言語テキストからナレッジグラフを生成する大規模言語モデル ・ナレッジグラフ推論LLM）ナレッジグラフに基づいて質問回答を論理推論する大規模言語モデル 社会実装に向け、自社ビジネスに関連させた展開を図り、モデル・開発ノウハウ等の成果物も公開・提供 ・本開発事業で得られた成果物の一部を、Hugging Face/GitHub等にてOSS公開

実施内容

① データセットの構築

事前学習用に、自然言語テキスト⇔ナレッジグラフの対訳コーパスを整備
ナレッジグラフの生成タスクと推論タスクの指示学習データを作成

② アーキ・ハイパラの選定

③ LLMの開発と精度向上

上記学習データを用いて7B～46.7B級(MoE構成含む)のLLMからベースとなる1つを選定し、事前学習・指示学習を実施

④ 成果物の公開

開発される基盤モデル

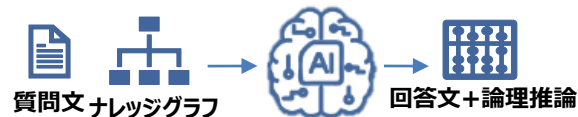
・ナレッジグラフ生成LLM

関係抽出ベンチマーク（英語：ReDocRED, 日本語：JacRED）で世界最高性能+10ポイントを達成



・ナレッジグラフ推論LLM

マルチホップQAベンチマーク（英語：HotpotQA, 日本語：JEMHopQA）と法的判断予測ベンチマークで世界最高性能+10ポイントを達成



社会実装の方法

自社ビジネスとして展開

- ・**Fujitsu Kozuchi**) 自社AIプラットフォームへの搭載により、企業・一般向けに広く試用環境を提供
- ・**Fujitsu Uvance**) 各種ビジネスオフリングに組み込み、グローバルにソリューション展開

成果物の公開



ソースコード
モデル利用のためのコード



モデル



開発ノウハウ