

日本のカスタマーサポートのための高品質AIエージェントモデルの開発・事業成果概要

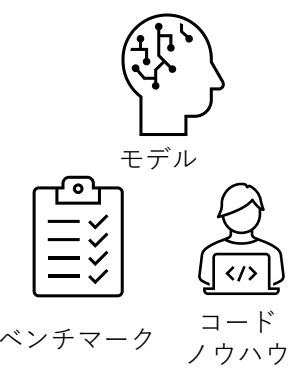
実施期間 2024年度～2025年度

実施者

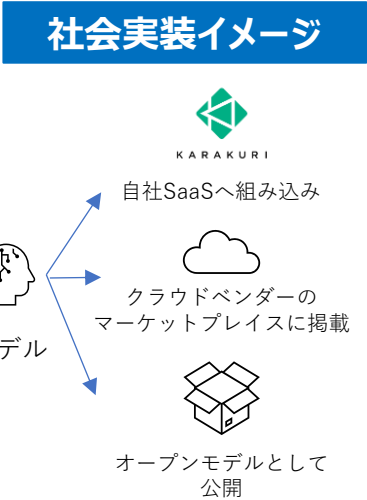
カラクリ株式会社

事業概要

日本に特化した高品質なカスタマーサポート向けAIエージェントを実現するためのモデルの開発を行う。



本プロジェクトは日本に特化した高品質なカスタマーサポート向けAIエージェントを実現するためのモデルの開発を目指している。データの作成、ベンチマークの作成、モデルの開発をそれぞれ行う。データについてはすでに作成済みのカスタマーサポート用データセットをさらに拡充し、高品質なカスタマーサポートAIエージェントを作るための学習データを整える。加えて、カスタマーサポートの性能を測るベンチマークを作成し、モデルの構築、および評価を行う。



日本のカスタマーサポートの効率化の実現のため、日本に特化したAIエージェントモデルを構築し、以下のような形で社会に広く提供を行う。

1. 自社SaaSへの組み込み
当社はSaaSとしてすでにカスタマーサポート向けのサービス群を提供しており、その中に組み込んで提供する。
2. クラウドベンダーのマーケットプレイスへ掲載
Bedrock Marketplaceなどに掲載することで各社の環境で開発者がすぐに使える形で提供する。
3. オープンモデルとして公開
Huggingfaceなどにオープンモデルとして公開する。

事業成果

カスタマーサポートのためのComputer Using Agent（CUA）モデルのコア技術となる日本語Vision Language Modelを開発した。同パラメータ数以下のオープンモデルの中では、各種VLMベンチマークで最高水準のスコアを記録し、クローズドなフロンティアモデルと比較しても遜色のない結果を示している。

本モデルはQwen2.5 VL 32Bをベースとし、AWS Trainiumを用いて学習された。QwenのVLベースモデルをTrainiumで学習したのは世界初であり、その際のノウハウとコードは今後公開予定である。

また、日本のカスタマーサポート向けAIエージェントに求められる能力を測定するベンチマークも構築しており、こちらも順次公開していく方針である。

本モデルはすでにユーザー企業において実証実験が開始されており、その成果についても追って公表する予定である。

	JDocQA (Acc)	JMMMU	Heron
Qwen/Qwen2.5-VL-32B-Instruct	25.2	48.8	74.8
gpt-4o-2024-11-20	22	57.5	93.7
meta-llama/llama-3.2-11B-Vision-Instruct	15.4	36.4	38.1
今回開発モデル 1（32B, SFT）	26.3	55.2	81
今回開発モデル 2（32B, GRPO）	27.8	53.8	84.7