

ロングコンテキスト対応基盤モデルとAIエージェント構築に関する研究開発

実施者	株式会社ABEJA
概要	<ol style="list-style-type: none">1. ロングコンテキスト処理性能が高く、ローカル環境で動作可能な日本語LLMを開発し、プランニングおよびToolUseなどの能力を強化する。2. 1を用いた、実用的で汎用性の高いAIエージェントを開発する。3. 企業のミッションクリティカル業務への対応能力を検証するため、特定企業と連携し、提供されるデータを用いた実証を行う。

実施内容

- 1 ロングコンテキスト対応基盤モデルの開発
- 2 基盤モデルに基づく業務特化型AIエージェントの構築

開発される基盤モデル



企業のローカル環境上で
ミッションクリティカル業務に
適用できる
LLMおよびAIエージェントを開発

数十B以下の言語モデルに
追加学習で以下を強化

- ロングコンテキスト
 - タスクの複雑性に伴う多様なツールの利用
 - 複数のデータ参照や長い会話・思考履歴
- プランニング
 - 複雑なタスクを分解・順序立てて遂行
- ToolUse
 - 必要な外部関数や情報源を選定・実行
 - 繰り返し考える

社会実装の方法

- 実企業の業務特化エージェントとして使えるかの実証
- 自社ビジネスとしての展開
 - ABEJA Platformへの搭載
 - ABEJA Platformを通じて各顧客企業へ導入

- 成果物の公開



モデル



開発ノウハウ

※企業データを学習した
モデルは非公開

時空間MoEを用いた動画生成AI基盤モデルの研究開発と当該プラットフォームの満足度検証

実施者	株式会社AIdeaLab
概要	<ul style="list-style-type: none"> 本開発事業において2つの動画生成AI基盤モデルの開発 <ul style="list-style-type: none"> 時空間MoEによる小規模動画生成AI基盤モデルの開発 時空間MoEによる動画生成AI基盤モデルの開発 モデル公開を通じた、動画生成AI基盤モデル社会実装 <ul style="list-style-type: none"> 時空間MoEによる小規模動画生成AI基盤モデルやソースコードの公開と開発ノウハウの共有 動画生成AIプラットフォーム公開を通じた、動画生成AI基盤モデルの社会実装 <ul style="list-style-type: none"> 時空間MoEを含めた小規模動画生成AI基盤モデルのサービス提供やソリューション提供

実施内容

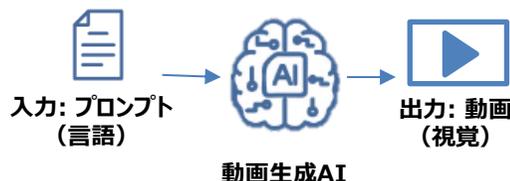
- 1 時空間MoEによる動画生成AI基盤モデルの開発
- 2 社会実装した動画生成AIプラットフォームの満足度検証
- 3 動画生成AI基盤モデルを開発するためのデータセットの構築
- 4 生成物の著作権に関する法務的リスクの調査

開発される基盤モデル

- 複数のWan 2.1 1.3Bのファインチューニングモデル
- 時空間MoEによる小規模動画生成AI基盤モデル
- 複数のWan 2.1 14Bのファインチューニングモデル
- 時空間MoEによる動画生成AI基盤モデル

社会実装の方法

- 自社ビジネスとして展開
 - 動画生成AIプラットフォームによるユーザーへの提供
 - 時空間MoEによる小規模動画生成AI基盤モデルのAPI提供
 - 基盤モデルによるソリューション提供



成果物の公開



ソースコード
(推論用など)



小規模
時空間MoE
動画生成
モデル



開発ノウハウ
(動画生成AIをMoE
するときの注意点)

一貫性のある日本語Full-Duplex-Speech マルチモーダルLLMの研究開発

実施者	AI inside株式会社
概要	<p>あらゆる人が直感的にAI Agentと協働できる、流暢で違和感のない音声対話と、現実の業務への対応力を備え、どこにでも導入できる小さなサイズの日本語特化マルチモーダルモデル</p> <p>開発したモデルは以下の二つの方向性の製品に搭載して社会実装する</p> <ul style="list-style-type: none">①当社主力事業DX Suiteとの連携による音声データ処理支援②新たに展開するAI Agentを中核に据えた対話支援

実施内容

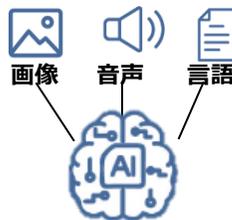
- 1 生成AI基盤モデルの開発
- 2 生成AI基盤モデルの開発に必要なデータセット収集
生成AI基盤モデルの
- 3 評価ベンチマーク (BuddyEval) の設計と公開
- 4 当該事業のノウハウと評価データセットの発信
- 5 研究開発成果の事業化



開発される基盤モデル

日本語音声対話とビジネス課題の解決に特化したモデル

- ・画像、音声に対応したマルチモーダル
- ・特定タスクで世界レベルの性能
- ・ローカルPC動作レベルの小ささ



社会実装の方法

自社サービスに搭載

- ・主力事業DX Suiteとの連携
- ・AI Agentを中核に据えた新サービス

成果物の公開



モデル
(API利用)



開発ノウハウ
(評価手法、一部データセット)

PLC制御におけるラダープログラム生成用大規模言語モデルの研究開発

実施者	Airion株式会社
概要	<p>工場制御設計にかかる工数の削減を目指し、ラダープログラムの生成に特化した30B規模のLLMを開発。協力会社から提供される実機で使われているラダープログラムデータを用いて既存のLLMを基盤に追加事前学習とファインチューニングを実施。</p> <p>社会実装に向けて設備メーカーを中心に実証実験を実施、具体的な成果を公開</p> <ul style="list-style-type: none"> • 実証実験では、実際の現場の制御設計担当者が本LLMを搭載したシステムを利用 • 実際にどの程度工数を削減できたかを測定し、成果として内容を公開

実施内容

- ① データセット設計・構築
- ② LLMの学習・評価
- ③ ベンチマーク設計・構築・評価



開発される基盤モデル

以下のモデルを開発

- 工場設備の動作に関する仕様を言語で入力し、ラダープログラムを出力する30B程度のパラメータのLLM
- より正確には、ラダーと一対一対応する機械語を生成
- 以下を開発目標とする
 - 現場に導入することで制御設計工数の削減を達成
 - 欧州で主流のST言語による制御の自動化AIを上回る精度を達成



社会実装の方法

設備メーカーへの導入

- 本事業で開発したLLMを含むAIシステムを設備メーカー向けに有償で開発して導入する。
- 各メーカー向けのシステムから共通部分を切り出してパッケージ化し、より安価で幅広い拡大を狙う。

成果物の公開

- ベンチマーク用データ・評価指標
- 上記に基づく当社LLMの評価値
- 実証実験における結果
- ※ 製造情報等の機密を含むため、モデルそのものの一般公開は難しい



- ベンチマーク
- 当社モデルの評価値



- 実証実験結果

リモートセンシング用視覚言語モデルの研究開発

実施者	Degas株式会社
概要	<p>本事業において、以下の生成AI基盤モデルの開発、検証を行う。</p> <ul style="list-style-type: none">マルチスペクトル画像対応の地理空間基盤モデルと大規模言語モデル（LLM）の統合による、4Bパラメータの衛星画像対応視覚言語モデル（VLM） <p>社会実装に向けたビジネス展開に加え、ノウハウ等の成果物も公開・提供</p> <ul style="list-style-type: none">国際機関との実証実験を実施本開発事業で得られた成果物(モデル・ノウハウ・検証結果)の一部を公開

実施内容

- 1 事前追加学習データの収集
- 2 事前追加学習実施・検証



開発される基盤モデル

- 4Bパラメータの衛星画像対応視覚言語モデル（VLM）
- マルチスペクトル画像に対応
 - 一部タスクで世界最高レベルの性能



社会実装の方法

ビジネス展開

- 国際機関との実証実験を実施
- 基盤モデルを活用したアプリケーションやAIソリューションの各社への提供

成果物の公開

ソースコード
(ベンチマーク用など)

モデル
一部モデルを公開

開発ノウハウ
マルチモーダル化および検証結果

外科手術支援のための視覚・言語統合型AI基盤モデルの開発

実施者	Direava株式会社
概要	<p>【開発内容】 手術中の安全性および外科医の状況判断・意思決定を支援するための、術式・解剖構造・術者行動に関する知識を学習した日本語外科手術特化型LLMの開発</p> <p>【社会実装】 自社手術ナビゲーション製品への組み込みに向けて、外部医療機関の手術室にて実証</p>

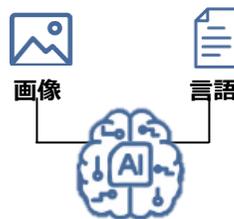
実施内容

- 1 外科専門言語コーパスの収集・術中画像と自然言語キャプションのデータセット構築
- 2 視覚・言語統合型AI基盤モデル開発
- 3 解剖部位アノテーション
- 4 医学的評価および外部検証



開発される基盤モデル

- 日本語外科手術に特化した7B規模のマルチモーダル基盤モデルを開発
- 独自で構築した高品質な外科手術データを学習
 - 臨床現場での実用に耐える性能



社会実装の方法

- 自社手術ナビゲーション製品への組み込み
- 医療AI規制に即した安全性・信頼性評価の実施



開発ノウハウ

一部の開発ノウハウを公表

大規模データを用いたリアルタイム音声基盤モデルの精度・機能向上による実用化加速

実施者	株式会社Kotoba Technologies Japan
概要	本事業では8B（80億パラメータ）の音声基盤モデルを、200万時間に及ぶ多様な音声データで事前学習し、下流タスクへ特化させる。対応言語は日英中韓を中心とした7カ国語、下流タスクは音声同時通訳、リアルタイムASR、音声合成とする。さらに、学習済みモデルを株式会社Kotoba Technologies Japanの実アプリケーション群で社会実装のPreview検証とフィードバック収集に活用し、その過程でモデルを蒸留して軽量化し、エッジ運用の実現可能性も示す。

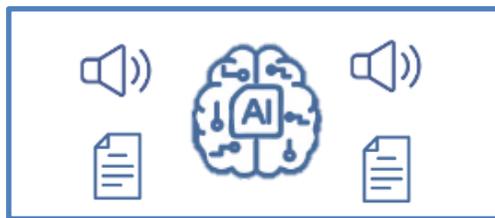
実施内容

- ① 大規模音声データセット構築
- ② 音声基盤モデル学習
- ③ 下流タスクへの特化
- ④ 社会実装と検証

開発される基盤モデル

音声データを、合成データを活用して事前学習・下流タスクのデータを合わせて7ヶ国語、200万時間まで増やして音声基盤モデルを学習する。

- 音声基盤モデルは、多言語化や超系列への対応ができるようになる。
- リアルタイム音声同時通訳、ASR、TTSなどの下流タスクで非常に高い性能を出す。



音声・テキストの入出力を行う下流タスクで高い性能を出せる音声基盤モデルを開発する。

社会実装の方法

開発したモデルをKotobaのiOSアプリ「同時通訳ーリアルタイム音声同時翻訳の新境地」（もしくは今後展開予定のDesktopアプリやAPI）などに展開して100,000分以上のユーザの使用を目指す。

開発したモデルを蒸留して小型のモデルを開発し、音声同時通訳やリアルタイムASRなどがノートPCやモバイルのエッジでも展開可能であることを実証する。

Factcheck RMとエージェントの開発研究

実施者	NABLAS株式会社
概要	<ul style="list-style-type: none"> ファクトチェックやハルシネーション対策に特化した強いモデルとファクトチェックエージェントの開発 <ul style="list-style-type: none"> 日本の高度なファクトチェック課題においてソースの真偽や信頼度の判定や分析を行う上で適切に分類・生成ができるモデル・エージェントを構築する ファクトチェック用の報酬モデルの構築や推論の際に並列サンプリングする推論時のスケールリングが可能なアプローチを構築することで広範囲に活用できる技術基盤を構築する 社会実装に向け、報道機関と共に本エージェントを実課題で検証しサービスの改善・運用に繋げていく <ul style="list-style-type: none"> 報道機関への事業展開と評価の基盤の構築（ベンチマークデータや評価方法の構築） 技術報告書やベンチマークデータの作成と公開し、業界全体で評価・改善が可能な環境を整備する

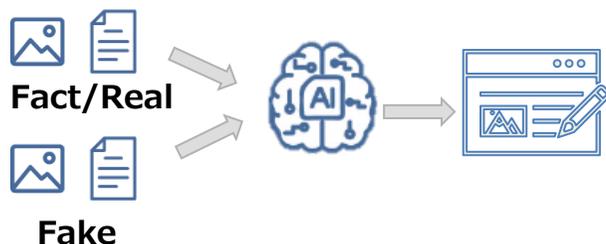
実施内容

- ① データセット構築
- ② リジექティブファインチューニング (RFT)
- ③ ルールベースの強化学習 (RL)
- ④ エージェント・フェイク検出ツールの構築
- ⑤ 報道実課題への実装・改良
- ⑥ ベンチマーク作成・公開

開発される基盤モデル

日本のファクトチェックやフェイク検出を高度に扱える報酬モデルやLLMの構築

- ファクトチェック向け報酬モデルFRMの構築
- 上記を元にしてGRPOアプローチによる強化学習を実施したLLMの構築
- 推論時に並列サンプリングし報酬モデルFRMで評価・統合する推論モデル



社会実装の方法

- ファクトチェックエージェントの実装とサービス構築
 - 報道の担当者が簡単に活用できるエージェントサービスの構築
 - 自社が保有する様々なフェイク検出・ファクトチェックモデルを活用できるエージェントやツールの構築
- ファクトチェックのデータセットや評価方法・指標の構築・公開



自律駆動R&DのためのAIエージェントアダプターの研究開発

実施者	株式会社NexaScience
概要	<p>本事業は、マルチAIエージェントシステムのための「AIエージェントアダプター基盤モデル」を開発し、AIエージェントによる高速かつ高精度な自律駆動R&Dを実現することを目的としています。</p> <p>🔑 核心技術 エージェントアダプター：エージェント間プロンプトチューニングを自律的に実施する基盤モデルです。エージェントごとに基盤モデルを微調整せずともマルチエージェントシステムを最適化できます。</p> <p>社会実装としては自社ビジネスとしての展開に加えて、論文調査→アイデア生成→実験→論文執筆を自律駆動できるOSS「AIRAS」に統合した状態でのモデルおよびソースコードの公開・提供も予定しています。</p>

実施内容

- ① エージェントアダプター基盤の学習方法の開発
- ② エージェントアダプター転移
- ③ 技術の開発
マルチエージェント調整
- ④ 機構の開発
- ⑤ AIRASとの統合
- ⑥ 評価と検証
データセット構築



開発される基盤モデル

AIエージェントアダプター基盤モデル

特徴

- 小規模パラメータ
- 基盤モデル非依存の設計
- ゼロショットタスク転移対応
- エージェント間通信最適化



社会実装の方法

自社ビジネスとしての展開

- AIエージェントアダプターを単独でSaaS提供
- OSS「AIRAS」の強化
- 研究開発者向け、知財担当者向け、投資家/事業開発者向けの自律AIプロダクトに活用
- 自律AIプロダクトの導入支援とカスタマイズに活用



要約タスク等を目標に出力形式への追従能力を高めたLLMの開発・データセット構築

実施者	Nishika株式会社
概要	<ul style="list-style-type: none"> • RAM8GB以下・GPU非搭載のPCでも動作する数BサイズのLLMを開発し、箇条書き／段落／テンプレート等、多様な出力形式に80%以上の精度で追従できる日本語モデルを実現 <ul style="list-style-type: none"> - 独自の日本語形式制御ベンチマークを整備し、自動評価パイプライン（形式遵守率／ROUGE-L／推論速度）実装 • 自社オンプレミス議事録ツール「SecureMemo」に組み込み、官公庁・民間企業でのトライアル評価を通じて実運用性を実証

実施内容

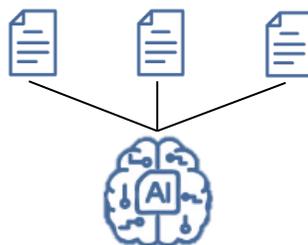
- 1 データセットの構築
 - 公開コーパス+自社データ, 形式制御プロンプトで合成データ作成
- 2 ベンチマークの作成
 - 文字起こし→要約の日本語ベンチマーク作成
- 3 追加学習および評価
 - 数BサイズのLLMベースにファインチューニング
- 4 実プロダクトへの組み込み
 - SecureMemo組み込み
 - 官公庁・民間企業にて実証



開発される基盤モデル

- 数Bパラメータ・量子化でRAM8GB以下のCPU推論に対応
- 箇条書き／段落／テンプレートなど多形式へ高い追従性能
- 形式遵守率≥80%、ROUGE-L既存8Bモデル比+5～10 pt
- LoRAのみで再学習可能な設計により、現場データで迅速にカスタマイズ

公開コーパス 自社データ 合成データ



数BサイズLLM

社会実装の方法

- SecureMemoに搭載、官公庁・医療機関等オンプレミス需要のある機関へ展開
- 契約下で官公庁・民間企業にモデル実証版を提供、実環境フィードバックを取得
- 形式制御データセット／ベンチマーク／合成データ作成ノウハウを一部公開

成果物の公開



データセット

一部公開



開発ノウハウ

合成データ作成

openBIMにおけるBIM情報要件の生成基盤モデルの研究開発

実施者	ONESTRUCTION株式会社
概要	<ul style="list-style-type: none"> 建設DXを進めるため、建設物の3次元データ（BIM）の標準化されたデータ運用を社会浸透させていく必要がある。弊社では国際標準であるopenBIM®の普及を目指している。 現状、専門知識がないとデータ交換のための情報要件（IDS）を作るのに多大な時間を要してしまう。本事業で開発する基盤モデルが専門家の知識を学習し、様々な要件からIDSを自動生成する。 これにより大手企業だけでなく中小企業も含め、幅広い建設会社がIDSを作成することを容易にし、国際標準に沿ったBIMデータの運用により建設業界全体の生産性を向上させる。

実施内容

- ① 学習・評価データセット構築
- ② 異なるサイズのモデル開発
- ③ スケーラブルなインフラ構築
- ④ 評価データセット、ノウハウの公開



開発される基盤モデル

入力：自然言語プロジェクト要件

- （例：建築物の種類、必要な設備、求められる性能など）
- 国際標準（IFC）および関連ドキュメント（ISO規格、仕様書など）
 - 建設業界専門用語・知識ベース
 - 過去のIDSサンプルデータ

出力：XML形式で記述されたIDS

- 建設プロジェクトに必要なデータ項目とその精度を厳密に定義
- IFCスキーマ（ISO 16739）との整合性を保ったデータ定義

要件指示



ナレッジ



技術書



社会実装の方法

自社ビジネスとしての展開

- 自社サービスであるopenBIM®ソリューション『OpenAEC®』への組み込み
- APIプラットフォームの構築及びSI

成果物の公開

- 評価データの公開
- テックブログ、勉強会でのノウハウ公開
- 一部モデルの公開（パラメーターの少ないもの）



モデル
（一部公開）



開発ノウハウ

自律稼働デバイスに向けた高精度軽量VLMの開発

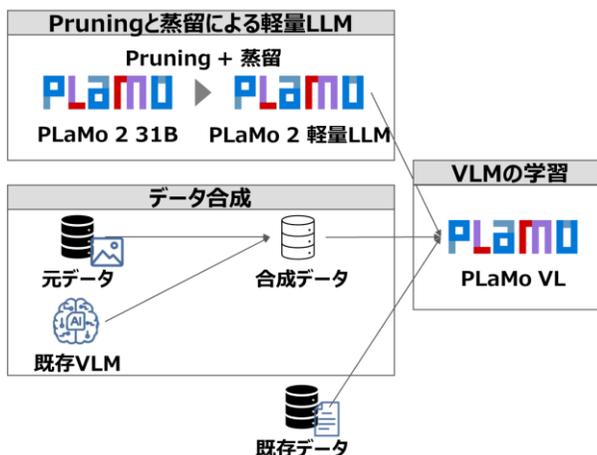
実施者	株式会社Preferred Networks
概要	<p>生成AIの物理世界への応用を見据え、視覚情報と言語情報を統合したVLMが注目されている。本開発事業においては、ローカルやエッジでVLMを動作させることを想定した、高精度かつ軽量で動作可能なVLMの構築を目指し、以下を実施する。</p> <ul style="list-style-type: none"> PLaMo 2を用いた高品質データの生成と、合成データ基盤の画像への拡張 生成した高品質データを用いたモデルの学習とpruningや蒸留を施した高精度軽量VLMの開発 ドローンや監視カメラへの応用を見据えたユーザ企業との社会実装に向けた実証

実施内容

- ① データセットの構築
- ② 基盤モデル開発
- ③ モデルの外部提供の実証
- ④ ノウハウの共有および公開

開発される基盤モデル

Pruningと蒸留を施した軽量PLaMo 2に対し、既存のVLM等を用いて生成した合成データ等で学習することで、高精度かつ軽量のVLMを学習する



社会実装の方法

- インターネット通信が制限される環境を想定し、複数のユーザ企業と実際の業務で使用する想定のデータを活用した実証を行う
- ドローンに搭載したカメラで撮影した空撮画像に対して、開発したVLMで異常等を検知する検証を行う
- 監視カメラで撮影した工場内の動画に対して、開発したVLMで作業員の行動等を分析する検証を行う
- ベースモデルとなるVLMについては公開し、テックブログ等を通じて開発ノウハウも継続して公開する



モデル



開発ノウハウ

視覚接地した文書特化型 視覚言語基盤モデルの構築

実施者	Sansan株式会社
概要	<p>視覚接地した文書特化型視覚言語基盤モデル「Cello」の構築を行う</p> <ul style="list-style-type: none"> 既存の基盤モデル「Viola」に判断根拠となる位置情報を出力する機能を追加 <ul style="list-style-type: none"> 矩形を介したエンジン間連携により、細かい文字の補正が可能 「Viola」を補正するために残存する工程を撤廃可能になる 位置を表現する特殊トークン導入により、VQAの枠組みで実現可能 <p>開発後は当社データ化による社会貢献と論文等において知見の公開を予定</p>

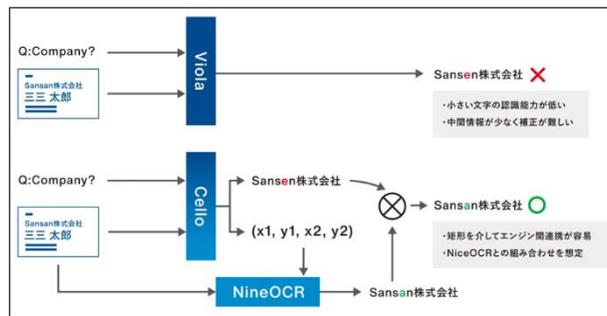
実施内容

- ① 小規模ファインチューニングのデータセット構築
- ② 文書画像を扱う視覚言語基盤モデルの拡張（小規模ファインチューニング）
- ③ 事前学習データセット構築
- ④ 位置情報付き事前学習
- ⑤ ファインチューニング
- ⑥ 基盤モデルの最適化（事前学習強化）
- ⑦ ファインチューニング（最適化後）
- ⑧ 論文執筆

開発される基盤モデル

視覚接地した文書特化型視覚言語基盤モデル「Cello」

- 位置を表現する特殊トークン導入により、VQAの枠組みで実現可能
- 出力結果の補正が容易となり、既存の基盤モデルに比べデータ化工程数の削減が可能



既存の基盤モデル「Viola」と「Cello」の比較図

社会実装の方法

自社データ化システムへの組込による社会提供

- 迅速にデータ化されるプロダクトの提供
- 当社プロダクト導入企業への付加価値提供

成果物の公開

- 構築後に事前学習の効率化方法に関して論文公開



ソースコード
論文に関するコード・
データ処理に関する
疑似コードのみ



開発ノウハウ

長尺コンテキストを理解する大規模映像基盤モデルの開発

実施者	SDio株式会社
概要	<p>長尺映像データの深い文脈理解を実現する、大規模映像基盤モデル(LVM)を開発する。</p> <ul style="list-style-type: none"> ● 数時間にも及ぶ映像コンテキストの長尺処理能力を実証 ● 階層的メモリ、MoE等の先端技術を組み合わせ、日本語の映像に対する高い理解度とコスト効率を両立させる ● 社会実装では、現在提供中のAI映像解析プラットフォームの精度や文脈理解を高めると同時に、ファインチューニングによる業界特化ソリューションの基盤を構築し組み合わせ展開する ● 開発で得られた研究知見、ベンチマーク、一部ツールは公開し、国内のAI開発コミュニティへ貢献する

実施内容

- 1 長尺コンテキスト用のデータセット構築
- 2 モデルの開発・評価
- 3 ファインチューニングパイプラインのプロトタイプを実装
- 4 ベンチマーク公開
- 5 コミュニティへの参画・開発知見を共有

開発される基盤モデル

長尺映像における文脈、日本語を深く理解する、7B-13Bの大規模映像基盤モデル

- 階層的メモリ機構により、数時間級の映像から因果関係や予兆を推論
- MoEと独自アーキテクチャにより、高いコスト効率を達成する
- オープンソースやTVデータで学習し、日本語の映像・音声・テキストにおいて最高水準を達成する
- 業界特化アダプタにより、特定産業への迅速な対応が可能に



社会実装の方法

API提供と業界特化ソリューションで事業を展開

- 既存AI事業のLVM化
- メディア・小売業界の顧客へ、より高付加価値な分析を提供し、国内事業基盤を強化
- グローバル市場への展開
- 業界特化アダプタを迅速かつ低コストで特化ソリューションを開発・提供

成果物の公開



AIによる分子の発明を実現させるタンパク質-リガンド分子相互作用基盤モデルの開発

実施者	SyntheticGestalt株式会社
概要	<p>タンパク質とリガンド分子間の相互作用を深層学習によって理解し、特定のタンパク質ターゲットに対して最適な結合特性と合成容易性を併せ持つ新規化合物を生成するための基盤モデル「Gestalt Interaction」と、それに分子の生成機能付与した「Synthetic Ligand」の開発を目的とする。</p> <p>本事業は、創薬分野を主要応用領域としつつも、材料科学、農業化学、環境技術、食品科学など幅広い産業領域における分子設計イノベーションを可能にする国産基盤モデルの確立を目指す。「Gestalt Interaction」は当社が開発済みの低分子構造理解に特化した基盤モデル（SG4D10B）を拡張し、化合物の生成機能を追加し、タンパク質構造情報を条件として組み込んだより実地的な統合基盤モデルとなる。</p>

実施内容

①トレーニングデータの開発

ドッキングシミュレーションによって複合体データを生成する

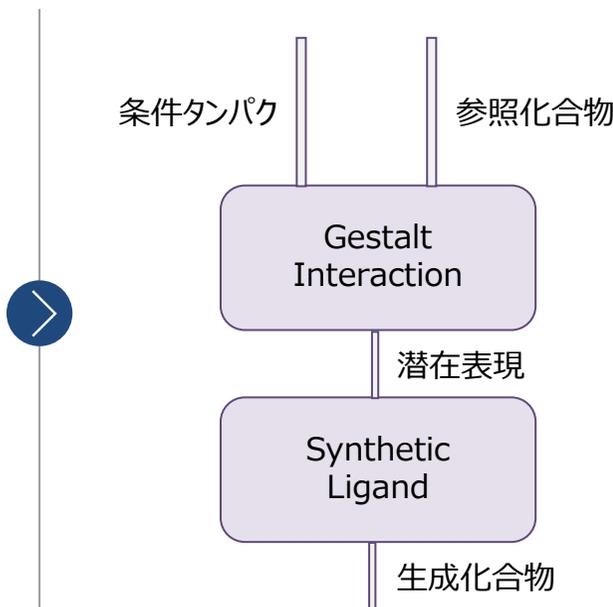
②Gestalt Interaction の開発

モデル・アーキテクチャ：化合物、相互作用、タンパク質の3点を入力とする。相互作用は Heterogeneous Graph として表現したもの、Equivariant Graph Neural Network によって Encoder と Decoder を構成する。

③Synthetic Ligand の開発

強化学習を用いて生成分子が結合親和性を最適化

開発される基盤モデル



社会実装の方法



モデル

有償モデルを Google Cloud の Marketplace にて流通。ユーザーに合わせてWeb App 化なども行い提供する。



開発
ノウハウ

開発手法に関しては、特許を申請した後に、Top tier のジャーナルへ投稿する予定。

完全自動運転に向けた車載可能なフィジカル基盤モデルの開発

実施者	Turing株式会社
概要	<ul style="list-style-type: none"> 映像・センサ・制御信号を統合的に理解し、実世界の状況変化や移動体の動きを予測できるフィジカル基盤モデルを開発 マルチモーダル生成AIを活用し、言語・視覚・行動を単一系列で出力する新たな推論アーキテクチャを構築する 将来的な完全自動運転の社会実装を見据え、基盤技術の一部は他領域でも再利用可能な共通基盤として公開予定

実施内容

- 1 画像圧縮トークナイザの学習**
 - 走行映像を情報量1KB以下の可変長トークンに変換
- 2 実走行データによるフィジカル基盤モデルの学習**
 - 実走行環境における多様な運転シーンを収集し、視覚・言語・行動が統合された大規模学習を実施
- 3 自動運転システムへの統合**
 - 学習済みモデルを軽量化・量子化し、車載GPU上でのリアルタイム推論に対応する構成で実証



開発される基盤モデル

- 100億パラメータ規模の視覚／言語／行動を統合するフィジカル基盤モデルを開発
- 視覚-言語-行動の自己回帰系列生成に対応し、現実世界の認知・判断・計画を一体的に実行
 - 走行映像・センサ・制御信号に基づき、自動運転タスクと連動した学習を実施
 - 複数の評価指標・ベンチマークにて検証を行い、国際的に最高水準の性能を目指す



社会実装の方法

- 自社ビジネスとしての展開
- 自動車メーカーや大手サプライヤーとの共同開発・技術実証 (PoC) を通じて、カメラ+AI 方式の自動運転システムをADAS製品へ応用し、販売する

成果物の公開



建築現場の施工管理を自動化するAI基盤モデルの開発

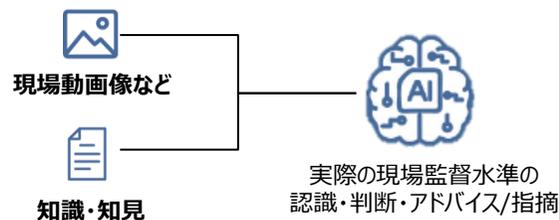
実施者	Zen Intelligence株式会社
概要	<ul style="list-style-type: none"> 建築現場の施工管理の自動化を見据えた、現場特化型基盤モデルを開発 <ul style="list-style-type: none"> 建築現場の空間・物体・意味データを時系列で統合したマルチモーダルデータセットを構築する 実際の現場監督程度の認識・判断・アドバイス/指摘ができる Large Vision-Language Model (LVLM) を開発する 社会実装に向けビジネス展開に加えて、本事業で開発したモデルのアーキテクチャや得られたノウハウ・工夫点などを公開予定

実施内容

- 1 建築現場に特化したマルチモーダルデータセットの構築
- 2 Large Vision-Language Model(LVLM) の開発
- 3 モデルの導入実証
- 4 モデルのアーキテクチャ・ノウハウの公開（一部）
- 5 社会実装に向けたビジネス展開

開発される基盤モデル

- 建築現場の三次元空間データを中核にマルチモーダルデータセットを構築
- 現場監督の品質・工程・安全項目に関する知識・知見
 - 部材・建材等の情報や設置箇所
- 現場特化型の基盤モデルを開発する
- 最大7BのLVLM
 - 実際の現場監督程度の認識・判断・アドバイス/指摘を目指す



社会実装の方法

- 自社ビジネスとしての展開
- zenshotへ搭載し、各企業へ導入予定（カスタマイズ・コンサルティング・展開）
- 成果物の公開
- モデルのアーキテクチャやデータフォーマット、課題設定やモデル構築上の工夫点など



ソースコード
(一部)



開発ノウハウ

低分子化合物の生物活性を世界最高精度で予測する、創薬生成AI基盤モデルの研究開発

実施者	アリヴェクス株式会社
概要	<p>開発内容と目標</p> <ul style="list-style-type: none"> 低分子化合物と創薬標的の結合強度を高速・高精度・絶対予測する独自技術ModBindをオラクルとし、化合物の大規模生成と絞り込みを組み合わせることで、能動学習により精度高く効率的にAIの追加学習を実施し、世界最高精度の低分子創薬AIを開発する <p>社会実装に向けた取り組み</p> <ul style="list-style-type: none"> 本事業で開発したAIモデル・データは、本事業終了後に学会発表・論文等で公開を予定している 本事業終了後もさらなる検討を継続し、モデルの販売・提携・生成された化合物のライセンスなどを行う

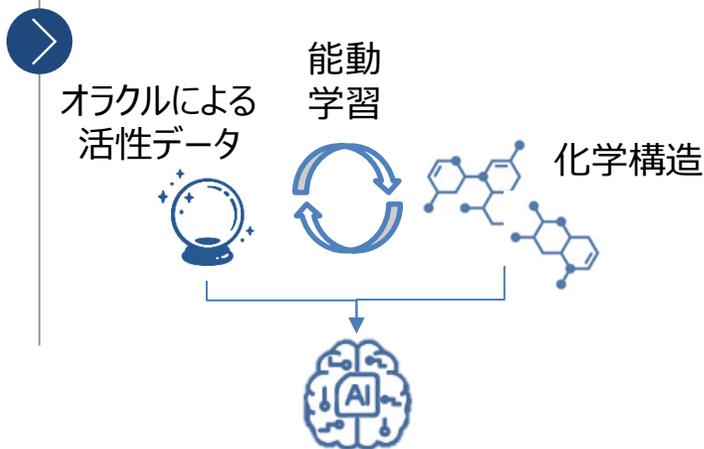
実施内容

- 1 ModBindオラクル、および化合物大規模生成と絞り込みの、能動学習プロセスへの統合
- 2 能動学習を利用したモデルの追加学習
- 3 ベンチマークデータでのモデル精度検証
- 4 論文等での成果の公開(本事業終了後)
 - 構築したモデル
 - オラクルの生成データ

開発される基盤モデル

世界最高精度で低分子化合物の生物活性を予測する創薬生成AI基盤モデル

- 幅広く化合物構造をカバーする、複数の創薬標的モデル



社会実装の方法

モデルのライセンスおよび、モデルを用いた実際の創薬と成果のライセンス

- 論文等での成果の公開(本事業終了後)
- 構築したモデル
 - オラクルの生成データ



モデル



オラクルデータ

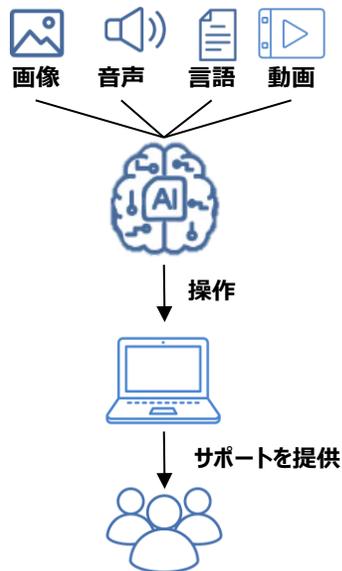
日本のカスタマーサポートのためのオムニモーダルエージェントモデルの開発

実施者	カラクリ株式会社
概要	本事業では、日本のカスタマーサポートに特化したオムニモーダルエージェントモデルを開発する。既存の日本語LLMに加え、画像・音声・動画を統合的に理解できる能力を付与し、非エンジニアの現場担当者でも自然言語や視覚的指示でAIに業務委任が可能となることを目指す。音声・動画データを新たに整備し、Trainiumを用いた大規模学習と強化学習により、GPT-4o等フロンティアモデルと同等以上の性能の実現を目指す。モデルはオープンウェイトで、ノウハウ、ベンチマーク、コードと共に公開し、ベンチマークおよび現場実証により性能を検証する。

実施内容

- 1 データセット構築
- 2 モデルの開発
- 3 ベンチマークの構築
- 4 モデルの公開
- 5 ノウハウ・コード・ベンチマークの公開

開発される基盤モデル



社会実装の方法

- ・既存のサービスに組み込みユーザーへ提供
- ・モデルを公開し一般ユーザーが使いやすく
- ・ベンチマークを公開し、カスタマーサポート領域での生成AI開発を加速させる
- ・ソースコードや開発ノウハウを公開し、多くの生成AIモデル開発者が効率的に開発できるようにする



モデル



ベンチマーク



ソースコード



開発ノウハウ

暗黙知の抽出を目指した業界特化のドキュメント読解基盤モデルの開発

実施者	ストックマーク株式会社
概要	<ul style="list-style-type: none"> 本事業では、主に製造業を対象とした情報密度や専門性の高い複雑なドキュメント読解を行うための業界特化のマルチモーダル（視覚・言語）基盤モデルの開発を行う。 社会実装に向け、国内でのビジネス展開に加え、一部モデル・ノウハウ等の成果物も公開・提供 <ul style="list-style-type: none"> - PaaS化し他社へ販売、情報収集・検索SaaSの既存自社サービスへの組み込み - 開発したマルチモーダル基盤モデルの公開、開発ノウハウの提供

実施内容

- ① 一般のドキュメント読解を主な対象とした30B規模の視覚言語基盤モデルの開発
- ② 製造業のドキュメント読解に特化した30B規模の視覚言語基盤モデルの開発
 - 製造業のドキュメント収集
 - 合成データ作成
 - モデル学習



開発される基盤モデル

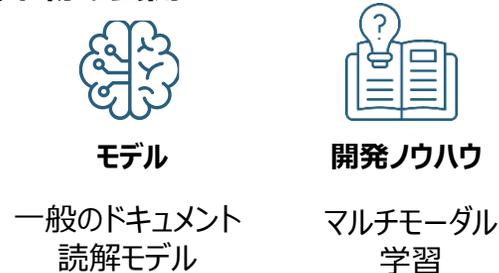
- 業界特化モデルのベースとなる一般のドキュメント読解を主な対象とした30B規模の視覚言語基盤モデル、及び製造業のドキュメント読解に特化した視覚言語基盤モデルの開発を行う。
- 製造業を含むビジネスドキュメントのベンチマークにおいて世界最高性能を目指す。



社会実装の方法

- 自社ビジネスでの展開
 - PaaS化によって他社へ販売
 - 自社既存サービスへ組み込み・販売
 - AI型の情報収集/検索SaaSサービス
 - 大手企業300社に導入済み

成果物の公開



中規模（10B-40B）の業界・タスク特化型LLMの研究開発

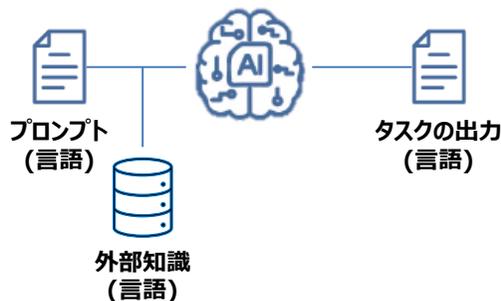
実施者	株式会社野村総合研究所
概要	<ul style="list-style-type: none"> 下記を特徴とする中規模（10B-40B）の業界・タスク特化型LLMの構築手法を確立する <ol style="list-style-type: none"> 様々な業界・タスクに適用可能 ベースモデル非依存で新しいモデルも自由に選択可能 高品質な専門知識テキストコーパス構築と継続事前学習 タスク特化の合成教師データ生成とファインチューニング 特化型LLMの量子化による推論の省資源化と精度維持の両立 金融業界（証券、保険）の実務を具体的な題材として、構築と実証を行う

実施内容

- ① ベースモデルの選定（10B-40BのオープンLLMを複数）
- ② 専門知識テキストコーパスの構築
- ③ 継続事前学習
- ④ タスク特化合成教師データ生成
- ⑤ タスク特化ファインチューニング
- ⑥ 実務タスクによる実証

開発される基盤モデル

- 複数の10B-40B規模オープンLLMをベースとした業界知識適応モデル
 - 今回は金融業界を題材に構築
- 上記モデルを各タスクごとに合成教師データでファインチューニングしたタスク特化モデル
 - 今回は証券、保険の複数のタスクを題材に構築・検証



社会実装の方法

- 自社ビジネスとしての展開
 - 個社向けに特化型モデル構築・導入サービスを展開
 - AIEージェントから呼び出すモデルとしても活用しながら、顧客業務プロセスへ組み込む
 - 様々な業界向けに構築・展開していく（まずは金融から）

成果物の公開



モデル

タスク特化モデルは
非公開



開発ノウハウ

論文執筆、
学会発表

SIP-jmed-LLMをベースとした医療用用途特化LLMの開発

実施者	株式会社プレジジョン がん研究センター 東京大学 九州大学
概要	<ul style="list-style-type: none"> 本プロジェクトは、国立情報学研究所 大規模言語モデル研究開発センターが中心となって開発するソブリンLLMであるllm-jpシリーズを医学コーパス800億文字で追加学習した「SIP-jmed-llm-jp (13B, 8×13B, 172B)」を核に①診療録自動校正、②個人情報匿名化等の5タスクに特化したマルチモーダル基盤モデルを6か月で構築することを目指す。 700施設で2030年度までに2万人の医師利用・年間12億円のSaaS収益化を図る。 医師の文章作成時間を50%削減し、患者待機時間を30%以上短縮させる。結果として、LLM、LMMの開発の基盤となり、日本の優れた医療を持続可能にする

実施内容

- 全体：医療特化LLMの設計・パッケージ構成案策定
- ① 医療従事者・患者の会話をカルテ記事に変える医療LLMの開発
 - ② RAG・匿名化の知識構造と評価整備
 - ③ 放射線所見構造化モデルの開発
 - ④ DPC分類支援モデルの構築

開発される基盤モデル

SIP-jmed-llm-jp 8 × 13Bを中心に医療ドメインのライセンス上透明性の高い医療テキスト生成をし、蒸留向けデータを学習し、13BサイズのLLMの開発を実現する。

他、医学知識によるRAGの作成など日本の医療現場でLLMの導入の壁となっている課題を解決するための医療LLMを開発する。

蒸留



8x13B



13Bサイズ等

社会実装の方法

▼ 社会実装

- ・プレジジョン社内商品への組み込み

▼ LLMモデルの公開

- ・作成モデルの公開

成果物の公開



ソースコード
モデルとのサ
ンプルコード
のみ



モデル
2モデルの
公開



開発ノウハウ

長期記憶メカニズムと対話型学習を融合した最先端の生成AI基盤モデルの研究開発

実施者	楽天グループ株式会社
概要	Attention機構を持つ生成AIモデルが抱える課題に対して具体的な解決策を推進する □ パーソナライズ機能：ユーザーの嗜好を反映したパーソナライズされた応答を可能にする対話型学習能力の実現 □ メモリー機能：ロングコンテキスト処理能力の飛躍的な向上による長期記憶能力の具備 □ アクション機能：AIエージェントによる検索、レコメンデーション、広告の最適化と購買体験の向上 □ コスト効率性：既存の生成AIに比較して、運用コスト（例：推論時間）の大幅な削減

実施内容

開発される基盤モデル

社会実装の方法

1 [モデルの開発]

約4~5兆トークンの厳選したオープンテキストデータを学習データとし、自己教師あり学習を実施

2 [ベンチマーク比較]

既存の最先端オープンソースモデル単体との性能と比較し、同等かそれ以上のベンチマークスコアを達成

3 [AIエージェント戦略]

当該基盤モデルを適用し、顧客ニーズに合わせた柔軟な対応を可能にするAIエージェントのプロトタイプを開発

4 [モデルの公表計画]

モデルのアーキテクチャや評価手法を公開、学習済みモデルはオープンウェイトで公開。ただし、学習データや学習コードは非公開

「Factorization Memory」を 中核技術としたロングコンテキストを 効率的に処理できるLLM

長期間・高精細なユーザ行動履歴を
コンテキストとしてLLMに与えることが可能



高度にパーソナライズされたAIエージェントの構築
に寄与

- ・ 検索・レコメンデーション・広告の最適化
- ・ AIをベースにした購買体験の向上

1. 当社サービス/外部企業への実装

国内最大規模のユーザーベースを持つ当社サービスに本成果を適用することで、いち早く、多くの方に、最先端のAIエージェント体験を届ける。また、将来的には外部企業への提供も実施

2. オープンウェイト公開

開発した基盤モデルのオープンウェイト公開を通して、公共性を確保。開発者コミュニティとの連携や、論文等による学術界へのフィードバックを通して、迅速な技術革新を実現

3. 継続改善体制構築

コミュニティからのフィードバックや、モデルのアップデートを通じた継続的な改善体制を構築し、汎用性と持続性を担保

ドキュメント群の高度な解析を通じて知の活用を支援するマルチモーダルLLM

実施者	株式会社リコー
概要	<p>日本のモノづくりの知の継承の危機を解決することをめざし「企業の知の結晶である様々なドキュメント群の高度な解析を通じて知の活用を支援するマルチモーダルLLM(LMM)」を開発</p> <ul style="list-style-type: none"> 複雑な図表を含む文書画像から情報を抽出、解析して判断する高いマルチモーダルリーズニング性能 プライベートLMMを想定した低コスト運用にも対応(小型化/高速推論など) <p>社会実装とともに、モデル・開発ノウハウ・ベンチマークを公開することで今後の基盤モデル開発に貢献</p>

実施内容

- ① 複雑な図表を含む文書画像から情報を抽出、解析して判断する高いマルチモーダルリーズニング性能の獲得
- ② 運用コストを大きく削減するモデルの高効率化
- ③ 開発コストを大きく削減する効率的な開発プロセスの確立
- ④ 学会発表・モデル公開・ベンチマーク方法公開
- ⑤ 社外向けの自社技術（LLMを含むAI関連）の紹介
- ⑥ GENIACコミュニティでの情報共有

開発される基盤モデル

ドキュメント群の読み取りに加えて解析も可能なマルチモーダルLLM

- 高度な解析対応
- 低コスト運用対応
- 高いカスタマイズ性



プライベートLMM対応
/カスタム対応

社会実装の方法

自社ビジネスとして展開

- マルチモーダルモデルのAPI提供
- マルチモーダルモデルのライセンス提供
- マルチモーダルモデルに付帯するビジネス（カスタマイズ・コンサルティング・展開）

成果物の公開



ソースコード
一部のみ



モデル
顧客データを含む
モデルは除く



開発ノウハウ