

独自映像データを活用した防犯ドメイン特化VLM基盤モデルの研究開発

実施者	株式会社アジラ
概要	独自に蓄積した防犯カメラ由来の異常事象映像データ800万件という他社にないデータ資産を基盤に、防犯ドメインに特化した基盤VLMを開発する。このデータ資産を活用し、4B/27Bモデルに対する継続事前学習を実施することで、防犯カメラ映像内で発生する事象を正確に意味理解できる「防犯ドメイン特化VLM」を構築する。さらに、実現場での社会実装までを一貫して推進し、人手不足が深刻である警備業界の社会課題の解決を目指す。

実施内容

- ① データセット構築：800万件超の独自異常事象データからキャプション生成及びの詳細Q&Aアンノテーション約150万ペアを構築
- ② 基盤モデルの研究開発：Qwen3.5系VLMをベースとした4B/27Bモデルへの2段階継続事前学習を実施
- ③ ベンチマーク評価：異常事象識別・意味理解能力・実証環境精度の3指標で既存モデルと比較検証
- ④ 社会実装：精度評価と合わせて実現場におけるPoCを行う

開発される基盤モデル

「防犯ドメイン特化VLM」

4Bモデル：エッジ環境（ローカルサーバー）でのリアルタイム処理向け。軽量異常検知AIとの連携により低遅延で異常事象を解析

27Bモデル：クラウド環境での大規模・高精度解析向け。録画映像の時空間意味理解による意味理解データベースの構築に活用

防犯カメラの
独自異常事象データ



“防犯ドメイン
特化VLM”



社会実装の方法

● 自社ビジネスでの活用

・既存のAI security asillaにadd-onする形で、大型商業施設・オフィスビルをターゲットに社会実証を推進。人手不足が深刻な警備業界での省人化警備を目指す。

● 成果公開

Tech Blog・勉強会登壇・コミュニティイベントを通じてモデル、開発ノウハウ、ハイパーパラメータチューニングのハウツーの発信



モデル



開発ノウハウ

自動車整備領域特化型AIエージェントの開発

実施者	株式会社ABEJA
概要	人材不足と技術高度化に直面する自動車整備領域で、「整備特化型AIエージェント」を構築・実証する。現場整備士の秘匿知見を基盤モデルに注入し、保安基準に関わる高精度な回答と自律的な「ToolUse」能力を実現。現場の機密保持のためエッジ推論環境を構築し、整備士の知見を再学習する「Human in the Loop」基盤により性能を継続改善する。将来は多拠点展開に加え、多言語化による外国人材支援や海外へのナレッジ輸出、Physical AIへの適用を推進し、モビリティ産業の基盤維持に寄与する。

実施内容

- 1 暗黙知抽出に向けたデータ基盤の構築と合成データ生成
- 2 自動車整備特化型数十B級基盤モデルの開発
(CPT/SFT)
- 3 エージェント機能
(Tool-use)の実装と最適化
- 4 現場知見に基づくフィードバックとデータ提供

開発される基盤モデル

自動車整備ドメインにおける高度な専門知識と、自律的な課題解決能力（エージェント機能）を備えた基盤モデルを開発。

アクティブパラメータが30B以下のオープンモデルをベースとし、非公開の整備記録・暗黙知を高品質な学習データへと変換し、継続事前学習および指示チューニングを実施。専門用語・手順に関する予測精度（Perplexity）を大幅に向上させる。

社会実装の方法

実証パートナーの拠点を活用し、「Human in the Loop」を用いて、現場のフィードバックを即座に反映し、実用性を検証。エッジ推論環境の導入により機密性を担保しつつ、段階的に展開を拡張。

老化を測定するエピゲノム生成AI基盤モデルの新規開発と長寿医療への応用実証

実施者	Epigenetic AI株式会社
概要	<ul style="list-style-type: none">日本人を対象とした全3,000万箇所のDNAメチル化データを網羅した大規模データセットを統合学習し、老化と疾患リスクを測定するフルスクラッチのエピゲノム生成AI基盤モデル(Large Epigenome Model)を開発する。独自の「階層圧縮アーキテクチャ」により超長系列を処理し、既存最先端モデル比で予見可能な加齢性疾患発症イベント検知数を7倍以上に向上させる。医療機関向けの老化・疾患予測検査事業や、食品・製薬企業向けの研究開発プラットフォーム事業として社会実装し、健康寿命の延伸に貢献する。

実施内容

- 1 エピゲノム生成AI基盤モデル (LEM) の開発
- 2 開発モデルの社会実装・実証
- 3 成果公開とコミュニティ運営・対外発信

開発される基盤モデル

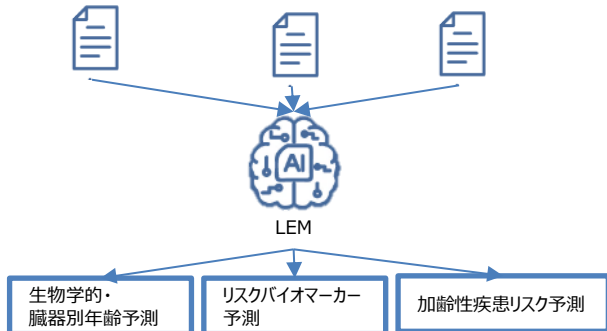
エピゲノム生成AI基盤モデル (Large Epigenome Model)

規模: 660億パラメータ

入力データ: 血液DNAメチル化データ 1.32兆トークン (44,000試料×3,000万箇所)

出力: 生物学的年齢、臓器別年齢、リスクバイオマーカー予測、加齢性疾患リスク予測等

エピゲノムデータ 多層オミクスデータ 臨床データ



社会実装の方法

実証企業での活用

医療機関での長寿医療有用性実証、食品企業での抗老化向け食品研究開発の有用性実証

自社ビジネス展開:

医療機関向けの生物学的年齢検査、食品・製薬企業向けの研究開発プラットフォームの提供

成果物の公開

開発ノウハウや学術論文を公開



論文



開発ノウハウ



実証成果 (国内)

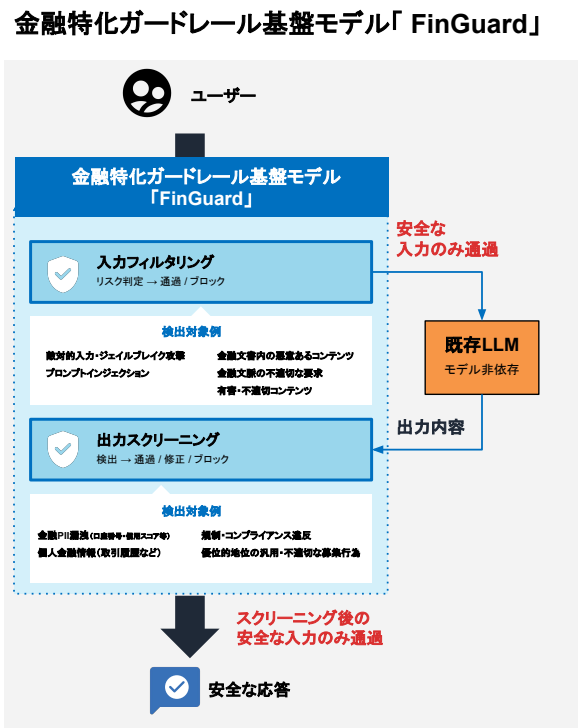
金融業務特化型AIガードレールモデルFinGuardの開発

実施者	株式会社 Elith
概要	<p>金融分野に特化した日本語ガードレール基盤モデルの構築を行う</p> <ul style="list-style-type: none"> 金融コーパスを活用したモデル高度化および安全性強化のための学習を実施 個人信用情報漏洩や規制違反などのリスクを高精度に検知・遮断 金融業務における AI活用のリスクを抑制し、安全性確保に貢献 特定の基盤モデルに依存しない構成により、柔軟な適用を実現 開発後は評価基盤の整備および知見の公開を予定

実施内容

- 1 金融安全性ベンチマーク「FinGuard-Bench」の作成
- 2 ベースモデルの選定
- 3 金融コーパスを用いた継続事前学習 (CPT)
- 4 金融安全性アノテーションデータを用いた安全性追加学習 (SFT)
- 5 ガード条件の最適化
- 6 金融機関等での実証検証
- 7 技術ブログ・論文投稿・評価レポート等による成果公開

開発される基盤モデル



社会実装の方法



実証

- 社内先行導入と技術検証の実施
- 金融AI安全性評価基盤「FinGuard-Bench」の開発・検証



事業化

- FISC・金融庁準拠パッケージによる直販
- Sler・金融DX事業者との拡販 連携および国際展開による拡販 (共同販売・OEM+アジア展開を統合)



成果物

- FinGuard-Bench公開による評価標準化への貢献
- 技術ブログ公開

ドローンインフラ・AIインフラを基盤とした多用途対応VLM基盤の研究開発

実施者	KDDI株式会社
概要	本事業は、DaaS（Drone as a Service）により全国で取得される多様なドローン実運用データを活用し、設備監視・インフラ点検など複数分野に共通して利用可能な多用途対応VLM（Vision Language Model）基盤を開発するものである。従来のCNNモデルは用途ごとに個別学習が必要で、異常データ不足や環境変動への弱さが課題であった。本事業では、大量のドローン実運用データと生成AIによるデータ増強を用いた（自己）教師あり学習を組み合わせることでドメイン特化基盤モデルを構築し、少量データでも高精度を保ちながら新規ユースケースへの拡張性を確保する。開発したモデルは国内データセンター上で運用し、既存手法との比較評価を通じて精度、処理性能、運用コストの優位性を検証する。最終的に、通信基盤・AI基盤を統合した国内循環型の解析環境を確立し、公共・民間領域への展開による社会実装と産業競争力向上を目指す。

実施内容

①データセット構築・データセット構築

- 実運用（設備監視）で取得する画像・映像データを集積
- クレンジング/データ加工に加え実データから疑似的に異常データを生成

②基盤モデル構築・モデル評価

- ドローン実運用データを利用した自己教師あり学習によってドローン画像に対するタスク性能を引き上げたドメイン特化基盤モデルを構築
- 実運用ドメイン性能（検知性能、状況説明性能）、多用途タスク性能（汎用的な視覚理解能力）を評価

③実証

- 監視ユースケースをもとに、KDDIグループが保有する施設での実証を予定

④ノウハウの公開

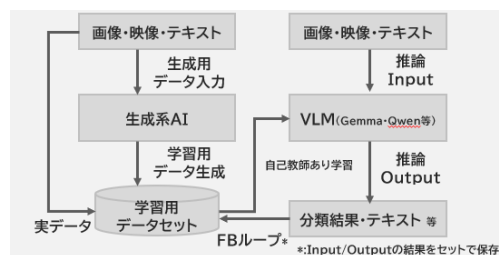
- KDDI最大のビジネスイベントであるKDDIサミットでの成果発表

開発される基盤モデル

「大規模なドローン映像・画像を活用した（自己）教師あり学習」によるドメイン特化基盤モデル

- 画像（動画を含む）データの拡充
実運用データを起点として生成AI技術を含むデータ拡充手法を活用し、不足する異常系画像データを補完
- テキストデータの生成
画像データに対し、VLM学習に必要なテキスト記述（状況説明文、検知指示と応答等）を付与

国内データの循環活用を前提とした実運用適用可能な基盤モデル確立を目指す



社会実装の方法

<Step1>

様々な領域への適用促進

- 大量データによる追加学習や大規模改修なしでCNN対比で短期間で実装
- 複数分野の用途に対する高速展開を実現

<Step2>

- DaaSの全国展開により、大量かつ継続的な画像・映像データ取得を常態化
- AI解析の継続的な改善ループを回すことで、国内産業の競争力強化と新規サービス創出を促進

■ 成果物の公開



本事業で得られた
知見



実証成果
(国内)

日本の制度・慣習・言語実務に適合したエージェント型基盤AIの開発

実施者	Sakana AI株式会社
概要	<ol style="list-style-type: none"> 1. 世界のフロンティアモデルと同等の規模である1Tクラスのモデルに対する継続事前学習と事後学習を国内で初めて実現させる 2. 人間が数時間から数日かけて行う網羅的な調査（Deep Research）や、外部ツールを用いた計算・情報整理・検証を通じて信頼性の高い結論を導く能力（ツール統合推論能力）を、世界のフロンティアモデルと同等の水準まで向上させる 3. 日本語に特化した音声認識と音声合成モデルを開発し、一般ユーザーにチャットアプリケーションとして広く使っていただくことで、ソブリンAIとしての実用性と信頼性を多角的に検証する

実施内容

- 1 継続事前学習によって日本知識とエージェント基盤能力の強化
- 2 事後学習によるDeep Researchとツール統合推論能力の強化
- 3 日本語音声モデルの開発
- 4 ユーザー実証

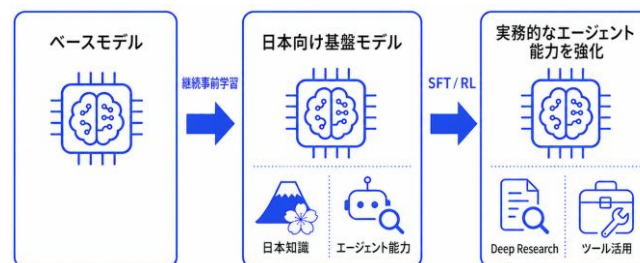
開発される基盤モデル

- 1Tサイズの日本向けのエージェントモデル
- 継続事前学習を通じて日本の慣習や文化の理解、汎用的なエージェント能力を実現
 - 事後学習（SFT/RL）により、実用的なDeep Researchや、Pythonなどのツールを用いた推論能力を強化
 - 日本語に特化した音声認識と音声合成モデルを開発

社会実装の方法

- 日本在住の日本語スピーカーを対象にチャットサービスとして提供し、以下の3つを想定した実証を行う
- エージェントによる調査・推論業務
 - 日本固有の専門知と適切な日本語表現が必要な業務
 - 自然な日本語音声対話

成果物の公開



開発ノウハウ

文書に特化した指示追従能力の高い視覚言語モデルの開発

実施者	Sansan株式会社
概要	<p>請求書・契約書などのビジネス文書に対し、ユーザーが自然言語で指定する複数の条件・フォーマット・回答拒否条件等に適切に追従し、高精度に回答できる視覚言語モデルを開発する。</p> <p>また、類似したタスクについて、共同研究を通してベンチマークを開発し公開予定である。開発モデルは「Bill One」「Contract One」に組み込み、顧客固有のデータ化・入力業務の自動化を通じて、バックオフィス業務の生産性向上に貢献する。</p>

実施内容

- ① 文書画像に対する指示チューニング (小規模)
- ② 文書画像に対する選好最適化 (小規模)
- ③ 社内データセットを用いた継続事前学習
- ④ 社内データセットおよび合成データセットを用いた継続事前学習
- ⑤ 文書画像に対する指示チューニング
- ⑥ 文書画像に対する選好最適化
- ⑦ 論文執筆



開発される基盤モデル

フォーマットや回答が存在しない場合の挙動など、多数の指示・制約に柔軟に追従して文書画像から回答できる視覚言語モデルを開発



自然言語での指示 (例)

- 販売促進費支払い方法／現場名を抽出
- JSON形式で出力
- 表の5行目に値がない場合は空欄
- etc...

出力 (例)

```
{
  "販売促進費支払い方法": "銀行振込",
  "現場名": null
}
```

社会実装の方法

自社プロダクト「Bill One」「Contract One」における、ユーザーの指示に基づき請求書や契約書をデータ化する機能上での実証

成果物の公開

- 本事業に関連するタスクのベンチマークを公開
- 実証結果については自社テックブログ等で共有



開発ノウハウ



実証成果
(国内)



ベンチマーク

金融領域に向けた非定型文書解析・根拠生成モデルの構築

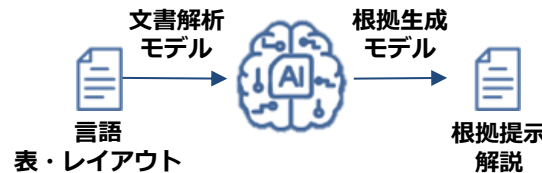
実施者	シェルパ・アンド・カンパニー株式会社
概要	<p>【開発内容】</p> <ul style="list-style-type: none"> ● 統合報告書や有価証券報告書、帳票等の非定型文書を対象とした、日本語特有の複雑なレイアウトを解析する文書解析モデルの開発 ● 情報抽出結果とともに文書中の具体的な根拠箇所と説明を提示する根拠生成モデルの開発 <p>【社会実装】</p> <ul style="list-style-type: none"> ● 従来の技術では困難であった複雑なレイアウトを含む文書に対し、情報の引用元・根拠を明示することにより、金融市場における情報抽出の信頼性向上と、原文に遡及可能な形での確認を可能にする ● 実証実験において点検工程等の作業時間を削減し、人的コスト削減目標50%を達成する ● 実験を通して得られた技術的成果については、論文・学会発表等の形で公開する

実施内容

- ① 文書解析データセット構築
- ② 根拠生成データセット構築
- ③ 文書解析モデル学習
- ④ 根拠生成モデル学習
- ⑤ モデルの公開
- ⑥ データセットの公開
- ⑦ 研究成果の発信
- ⑧ 実証実験
- ⑨ サービス化

開発される基盤モデル

十数Bパラメータ規模のマルチモーダル基盤モデルを開発



- 高精度な解析：複雑な表組・図表の解析、非定型・日本語の特殊性に対応
- 専門性：開示基準への準拠、専門用語や文脈に最適化された事前学習
- 透明性：情報の引用元・根拠の明示により信頼性を担保

社会実装の方法

ビジネス展開

- 自社SaaSへの搭載
- 金融機関、投資家、上場企業に対する情報収集・根拠生成ソリューションの提供

成果の公開

- 国内のAI開発者コミュニティや金融実務者へ還元
- 研究成果を関連学会や技術ブログ等を通じて積極的に発信



文書解析モデル
根拠生成モデル

構造・評価方法・
実証結果の概要
(オープンソースライセンス)



ソースコード

パラメータ・実行コード
評価データセットの一部

災害検知LLMとAIエージェントによる次世代防災・危機管理基盤の開発

実施者	株式会社Spectee
概要	<ul style="list-style-type: none">災害時に氾濫する膨大な情報をAIが自律的に解釈し、防災計画等に基づき、必要な対応をリアルタイムに示唆する次世代防災・危機管理基盤を開発する。これにより、災害初動のリードタイムを劇的に短縮するとともに、担当者を情報精査の負荷から解放し、人命救助・避難誘導・復旧支援といった、高度な判断を要する核心的業務に注力できる環境を実現する。成果は国内外の自治体・民間企業でのユーザー実証を経て弊社プラットフォームへ組み込み、日本発の「自律型・高度AI危機管理基盤」として国内外に展開する。これにより災害損失の最小化を実現し、世界のレジリエンス市場における日本のリーダーシップを確立する。

実施内容

- ① データセットの収集・整備と評価基盤構築
- ② 災害検知LLMの開発
- ③ 危機管理AIエージェントの開発とシステム統合
- ④ 機能実証および精度評価

開発される基盤モデル

災害検知LLM

- 国産LLM「PLaMo」をSpecteeが過去約10年にわたり蓄積した災害関連データで指示チューニング
- SNS情報等からの災害検知・事象分類・被災度判定タスクを学習
- 独自ベンチマーク「Japan SNS Disaster Benchmark (JSDB)」で汎用モデルを上回る性能を目指す

社会実装の方法

自社ビジネスとして国内外へ展開

- 国内トップシェアのAI危機管理基盤に統合し、自治体・民間企業等へ提

成果物の公開



モデル

軽量化モデルの公開を検討



開発ノウハウ



ベンチマーク

人間らしい対話を実現する音声基盤モデルの開発とその社会実装

実施者	株式会社DubGuild
概要	<p>大規模音声言語基盤モデルの開発</p> <ul style="list-style-type: none">700万時間の音声データセットを構築世界最大サイズの音声対話モデルを構築感情理解、自然な応答、割り込みを含むリアルタイム対話の実現をめざす

実施内容

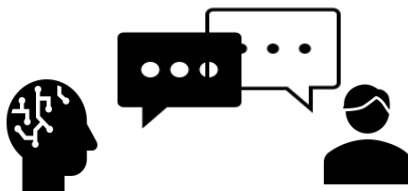
- 1 大規模音声データセットの構築
- 2 大規模音声言語モデルの事前学習
- 3 対話用途へのファインチューニング
- 4 実証および社会実装

開発される基盤モデル

700万時間規模の音声データを使い、音声を入力とする30B級の音声言語基盤モデルを構築する。

音声言語基盤モデルについて：

- 感情や間など、音声に含まれるがテキストに書きおこすことのできない情報を理解した応答生成が可能
- カスケードのシステムと比較してリアルタイムな出力が可能



社会実装の方法

実証パートナーとの共同検証を通じて、音声エージェントによる業務自動化や、社会で幅広く利用可能な対話モデルの実現を目指す。

成果物の公開



開発ノウハウ

開発ノウハウを自社テックブログを通じて公開
また、パートナー企業との実証内容をブログ等で公開予定



実証

手術安全性向上のための『外科手術未来予測モデル』の開発と実証

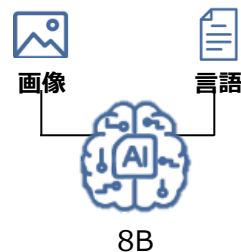
実施者	Direava株式会社
概要	<p>【開発内容】 手術中の安全性向上および若手外科医の術中意思決定・技能教育を支援するため、術前情報や動的情報と時系列における因果関係を内包した「外科手術未来予測モデル」を開発</p> <p>【社会実装】 自社手術プラットフォーム製品への組み込みに向けて、海外を含む外部医療機関の手術室にて実証</p>

実施内容

- 1 マルチモーダルアノテーション
- 2 外科手術未来予測モデルの開発
- 3 臨床的有用性評価および外部性能検証
- 4 World modelの構築に向けた事前検証

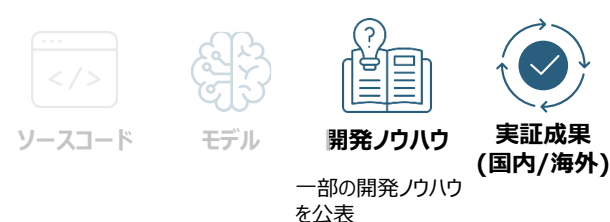
開発される基盤モデル

- 日本語外科手術に特化した8B規模のマルチモーダル基盤モデルを開発
- 独自に構築した術前・術中動態を含む高品質な因果関係データを学習
 - 有害事象を未然に防ぐ未来予測と臨床現場での実用に耐える性能



社会実装の方法

- 自社手術プラットフォーム製品への組み込み
- 医療AI規制に即した安全性・信頼性評価の実施
- 成果物の公開



Fast/Slow連携アーキテクチャによる自動運転タクシーの実現

実施者	newmo株式会社
概要	<ul style="list-style-type: none"> 日本特有の交通環境に対応した自動運転基盤モデルを開発し、自社タクシー基盤を活用して社会実装を推進する。これにより、深刻なドライバー不足や地域交通の維持困難といった社会課題の解決を図る。 VLMによる意味理解を担うSlowモデルと、BEV空間で高頻度制御を行うFastモデルを非同期連携させた独自の2層アーキテクチャを構築し、安全で説明可能なレベル4自動運転を実現する。

実施内容

- ① データセット構築
- ② Fastモデル開発
- ③ Slowモデル開発
- ④ 連携・安全制御・蒸留
- ⑤ ベンチ評価・外部提供実証・PoC
- ⑥ 成果の公開

開発される基盤モデル

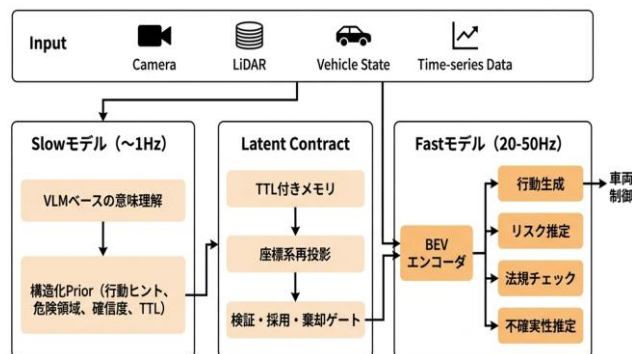
OFastモデル：リアルタイム制御等の自動運転タスクの実行
役割：BEV 中間表現上で高頻度に行動生成、リスク評価、法規整合性判定

仕様：車載GPU上で20-50Hzの高速推論。パラメータ数0.1B-0.5B規模。

OSlowモデル：VLMベースの高度な意味理解と判断

役割：VLMを用い、複雑な路辺状況や他者の意図など高次の意味理解

仕様：約1Hzで動作。パラメータ数4B-10B規模。交通領域へ追加学習。

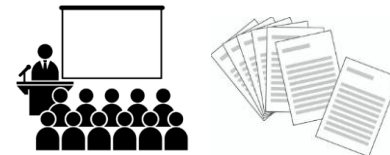


社会実装の方法

開発モデルを、公開ベンチマーク、外部提供用推論環境、限定エリアPoCの3段階で評価。

事業期間中に、限定エリア・限定時間帯における公道PoCを実施。介入率、遠隔支援率、完遂率、不要停止、急ブレーキ、乗降精度等を計測。

○成果物の公開



国内外での論文・学会への発表

実証や実験により得られた知見やデータを公開し、外部からも成果にアクセスできるようにする。

可能な範囲で、ソースコードならびに再現に必要なデータセットの公開も検討。

次世代創薬の戦略的自律性を確立するマルチモーダル空間AI：Virtual Cell基盤モデルの開発

実施者	株式会社biomy
概要	<ul style="list-style-type: none"> 動物実験では再現が困難なヒト腫瘍微小環境を、計算機上で高精度に再構築可能な腫瘍微小環境特化型の基盤モデルを開発する。新規標的分子の同定や臨床試験における患者層別化、最適な併用療法を事前予測することで、新薬開発の期間短縮および治験成功確率の飛躍的な向上を目指す。 公益財団法人がん研究会と連携し、非小細胞肺がんの病理スライドから、空間トランスクリプトーム・タンパク質・H&E画像の情報を統合したデータを構築。さらに病理AI開発に関する豊富な知見を有する東京大学より技術協力を得る。 一般社団法人バイオ産業情報化コンソーシアムと共に、本モデルの現場での有用性を評価する。

実施内容

1 マルチモーダル空間データセットの構築

同一切片から形態・遺伝子・タンパク質情報を取得

2 「Virtual Cell基盤モデル」の開発（データ前処理/モデル学習/評価）

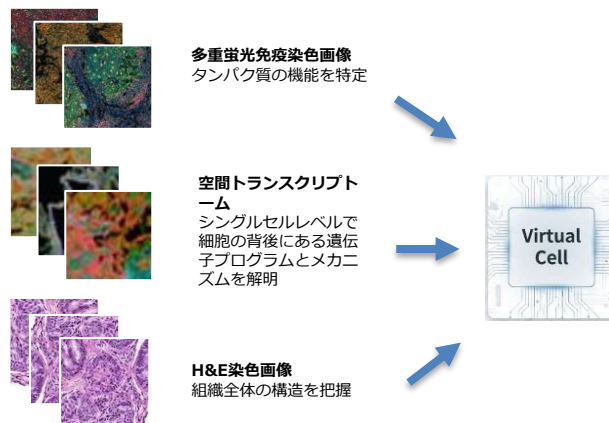
マルチモーダルデータを用いた基盤モデルを構築。既存SOTAモデルとの評価を通して、予測性能を評価。

3 ワーキンググループの組成と評価

ワーキンググループを通じた、産学官連携による客観的評価と概念実証

開発される基盤モデル

介入条件下における組織全体の空間的な遺伝子・タンパク質発現分布の予測



社会実装の方法



モデル

実証期間中はVirtual Cell ワーキンググループに参画する製薬企業が検証に利用可能とする。実証後は非独占的にライセンスを提供。



開発ノウハウ

データ構築プロトコルや、大規模分散学習のインフラ運用ノウハウをTech Blogやホワイトペーパーで公開。



実証成果
(国内/海外)

実証成果については、権利化後、論文化や学会発表を行う。

現実世界の再現を目指す材料基盤モデルの研究開発

実施者	株式会社Preferred Networks
概要	現実世界を再現する材料基盤モデルの構築を目的として、汎用性と高精度を両立する汎用機械学習ポテンシャルを事前学習と事後学習からなる新しい枠組みで開発する。事前学習においては電池材料等において再現性が高い高精度DFT計算データセットを新規に構築する。事後学習では事前学習モデルによる大規模推論の結果と実験値を使うことで融点など巨視的なスケールにかかわる物性値を直接学習する。得られたモデルについて、材料開発の実課題に対峙している企業と行う実証でフィードバックを得、実社会の課題解決に役立つモデル開発を目指す。成果物については技術開発に伴うノウハウの一部などをブログ等で公開することで、本事業で得られた知見を社会に還元していく予定である。

実施内容

- ① 事前学習モデル開発
- ② 事後学習モデル開発
- ③ 実用化に向けた実証
- ④ コミュニティへの貢献

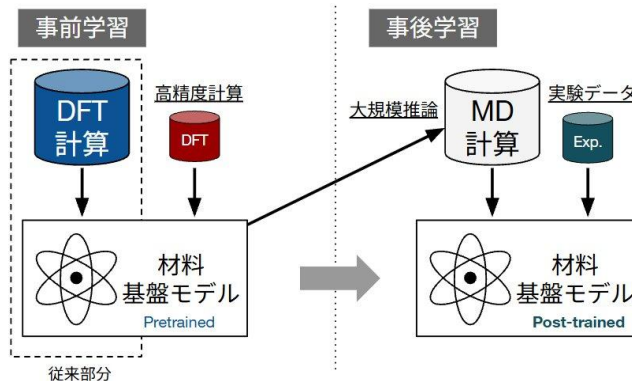
開発される基盤モデル

事前学習

- 新規高精度計算条件でデータセットを構築し、電池材料等での精度の向上を目指す

事後学習

- 大規模推論と実験値を用い、特に融点の精度向上を目指す



社会実装の方法

自社ビジネスとしての展開

- 汎用原子レベルシミュレータ Matlantisへ搭載、提供
- 材料探索事業の基盤技術として活用

成果物の公開

- データセットの計算条件、検証用データおよび実装を公開
- 技術開発に伴うノウハウの共有



モデル
(商用サービス、APIを通して)



開発ノウハウ



実証成果
(国内/海外)

Generative Retrieval技術を用いた二次流通市場向け高精度検索・推薦基盤モデルの開発

実施者	株式会社メルカリ
概要	メルカリの商品&カタログデータ、LLMで合成した感性豊かな口コミデータを学習し、AI自身が商品を丸暗記して探す日本初の次世代検索モデルを開発します。名前のわからない一点物や「結婚式にふさわしい服」といった曖昧な言葉からでも、お宝を一瞬で見つけるよう改善。何度も検索し直すストレスから解放され、世界中の誰もが欲しい物にすぐに出会える最高の買い物体験を目指します。出品物の約8割を占めるカタログ未紐付け商品の発見性を飛躍的に向上、再検索率の削減・商品購入率向上、越境取引の加速を通じたGMV拡大を目指します。

実施内容

- ①大規模二次流通データセットの構築
メルカリの40億品超の取引・検索ログと、駿河屋の3,000万件の精緻なカタログデータを統合
- ② LLMによる感性・口コミデータの合成
Llama4、Qwen等を用い、出品者と購入者の語彙の乖離を埋める300万件のレビューデータを生成
- ③ Generative Retrieval基盤モデルの開発
商品の階層構造を保持する「Semantic ID」を設計し、モデル自体にインデックス機能を内包させる学習を実施
- ④ リランキングと性能評価
幻覚を抑制するリランキング手法を組み合わせ、Amazon ESCI等の国際ベンチマークおよび社内実証環境で精度を検証



開発される基盤モデル

Qwen3.5をベースに、自社の膨大な取引ログ、連携先（駿河屋）の3,000万件のカタログ、さらにLLMで合成した300万件の口コミデータを統合学習させた二次流通特化型モデル

【参考にする基盤モデル図】

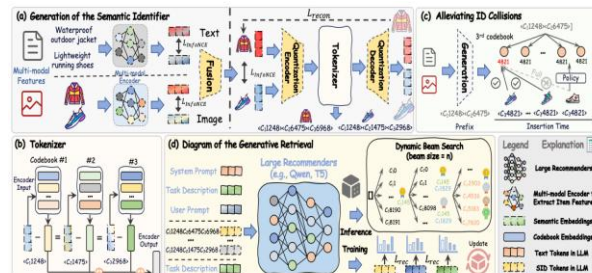


Figure 2: Overview of FORGE. (a) The generation process of SIDs. Left panel: the training of the text/image encoder. Right panel: the schematic diagram of item tokenization. (b) Detailed workflow of the tokenizer. (c) Post-collision handling for semantic identifiers. (d) The training and dynamic inference procedure of generative retrieval.

社会実装の方法

- 既存基盤への統合
月間利用者数2,300万人を超える「メルカリ」の検索・推薦エンジンに本モデルを直接搭載し、ユーザーへ提供する
- グローバル展開
多言語・多文化の文脈理解を武器に「メルカリ グローバルアプリ」へ実装し、越境取引における日本発ホビー文化（ACGE）の流通を加速させる

成果物の公開



モデル
Hugging Face等



技術知見
TechBlog、勉強会

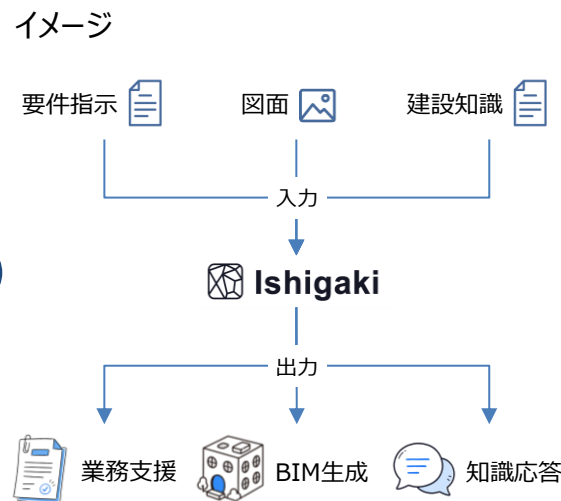
世界モデルの社会実装を見据えた、BIM生成統合型建設ドメイン特化基盤モデルの研究開発

実施者	ONESTRUCTURE株式会社
概要	<p>建設業界では人手不足と高齢化が深刻であり、建設DXの要である3次元設計手法「BIM」の普及は遅れている。前年度(Cycle3)では、BIM情報要件であるIDS生成に特化した基盤モデル「Ishigaki-IDS」と評価ベンチマークを構築し、汎用モデルでは対応が難しい専門領域においてドメイン特化モデルの有効性を実証した。</p> <p>本事業(Cycle4)では、この優位性を特定領域から建設ドメイン全体へと広げること为目标とし、前年度の成果を建設知識・業務支援・BIM生成を統合的に担う基盤モデルへ発展させて日本発の国際標準確立を目指す。世界モデルの社会実装を見据えた基盤モデル開発により、建設空間の構造特性・物理挙動・法規制の相互関係を内包しつつ、BIMの意味的整合性を保ち業務判断を支援し、建設DXを加速させる。</p>

実施内容

- 1 データ収集・評価データ・評価体系整備**
公開情報、法令、仕様書、BIMモデル、合成データを統合し、30~80Bトークン規模のデータセットを構築。既存のベンチマークに加え、独自構築したベンチマークで多面的に評価。
- 2 環境整備・学習/評価スクリプト実装**
最新のオープンウェイトモデルを継続学習させるためのスクリプトと、評価スクリプトの構築。
- 3 予備実験・基盤モデル学習**
建設ドメイン特化基盤モデルを継続学習により構築
- 4 外部実証・成果整理**
海外事業者含め、BIM作成や建設文書を扱うタスクにおいて実用性の実証を実施。

開発される基盤モデル




社会実装の方法


実証・事業化：

- 開発するモデルを用いた建設業務の自動化の実証を国内外の事業者と実施
- 実証成果を自社サービス「OpenAEC」への組み込みとAPI公開による商用展開を予定
- 国立松江工業高等専門学校との産学連携によるデータセットを構築。


公開する成果物：




ソースコード
(モデル利用のためのコード)



モデル
(AWS Marketplaceにて公開予定)



開発ノウハウ
(テックブログ、論文など)



実証成果
(国内/海外)