

報道関係者各位

2024年9月12日
FastLabel 株式会社

FastLabel、「VLM 開発用データ 支援サービス」を提供開始
学習データの収集、データセットの販売、アノテーション、RAG データ作成サービスなどを通じて
VLM の開発で発生するリソース不足やデータの質の担保といった課題を解消！

Data-centric AI 開発に必要なデータ収集・生成からアノテーション、モデル開発、MLOps 構築までの全工程を支援する FastLabel 株式会社（本社：東京都渋谷区、代表取締役：鈴木健史、以下「当社」）は、大規模視覚言語モデル（以下、VLM）開発用の高品質で権利クリアな日本語データセット販売サービスをはじめとした、「VLM 開発用データ 支援サービス」を 2024 年 9 月 12 日より提供開始したことをお知らせします。

VLM（大規模視覚言語モデル）開発用データ 支援サービスを提供開始



■「VLM 開発用データ 支援サービス」提供の背景

米 OpenAI 社の「GPT-4」、米 Meta 社の「Llama2」を始めとする LLM が公開されたことを受け、コンテンツを生成することが可能な生成 AI（Generative AI）技術が急速に発展しています。併せて、「GPT-4o」や「Gemini」に代表される VLM（大規模視覚言語モデル）も登場し、活用の可能性が期待されています。一方で、VLM の学習には過去の技術をさらに上回る質・量の画像データおよびテキストデータ収集、作成が必要とされますが、社内に既に存在するデータだけで十分な質・量を確保できない場合は、外部から大量のデータを権利クリアな状態で調達することが必要です。そこで、当社は VLM の学習データとして利用される画像・動画・文書データのデータセット販売、同データセットに対するアノテーション代行サービス、マルチモーダル RAG データ作成サービスを「VLM 開発用データ 支援サービス」として提供し、企業のデータ収集・作成において発生するリソース不足やデータの質の担保といった課題を解決します。

「VLM 開発用データ 支援サービス」紹介ページはこちら

<https://fastlabel.ai/service/vlm>

■当社が提供する「VLM 開発用データ 支援サービス」の特徴

<データセット販売の特徴>

- 100 万点超の多様な学習データ
当社パートナーのストックフォトサービスが所有する 100 万点超の画像・動画・文書データを VLM の学習データとしてご利用いただけます。
- 権利クリア
販売するデータセットは全て当社でゼロから作成しており、権利リスクがなく、安心してご利用頂けます。
- 高品質
当社ではデータの正確性や多様性に関する細かな品質管理基準を設定しており、データ作成工程、データ作成後の品質検査工程で基準への適合を担保しております。

<VLM 向けアノテーション代行サービスの特徴>

- 柔軟な構造のテキストアノテーションを付与することが可能
 - キャプション
 - 画像の特定の領域(矩形で指示)を根拠とした QA
 - 多肢選択式の QA
 - OCR
 - その他要件に応じて対応可能
- セキュアなデータ管理
お客様のクラウド環境に保管された機密性の高いデータは、当社環境のストレージに保存することなくアノテーションが可能です。
- 当社独自のプラットフォームを活用したスムーズな検品
プラットフォーム機能により、効率的なレビューと作業員間コミュニケーションが実現します。必要に応じてお客様側でランダムサンプリング機能を利用した抜取での受入れ検査が可能です。

<マルチモーダル RAG データ作成サービスの元となるデータ>

- お客様が社内で保有している PowerPoint、PDF、Excel 等の図表を含むドキュメント
- お客様が事業活動の中で蓄積される画像・動画データ
- お客様が開発する VLM の適用領域における信頼性の高い外部ドキュメント(IR 情報等)

■今後の展望

当社は今後、国内外の大手精製 AI ベンダーをはじめとした多くのお客様へのデータ提供の経験を元にデータの品質に関する客観的な基準の整備、品質定量化手法の開発を進めて参ります。当社は引き続き、データセットおよび AI プラットフォームのご提供、データ作成代行やコンサルティング等のサービスを通じて AI 開発プロセスを効率化し、AI 開発に取り組む企業、研究機関のインフラとなることで、日本社会における AI 革命を下支えしていきたいと考えております。

■ データセットの購入およびサンプル画像の確認方法

ご購入またはお見積り、サンプル画像の確認や新規撮り下ろしをご希望のお客様は、以下のメールアドレス宛、もしくはフォームよりお問い合わせください。今回販売を開始するデータセットにご希望の要件のデータがない場合は、ご要件に応じた新規撮り下ろし対応も可能です。

お問い合わせ：pr@fastlabel.ai

お問い合わせフォーム URL：<https://fastlabel.ai/service/vlm>

■ 当社のデータセット事業について

当社は、データセットの販売のみならず、以下に対応したデータ収集代行サービスを展開しています。

- ・画像（人物撮り下ろし、帳票データ収集 他）
- ・動画（スポーツ動画撮影、走行車両撮影 他）
- ・音声（スクリプト発話、ノイズ収録 他）
- ・テキスト（会話コーパス 他）* VLM 開発用途含む

【データ収集代行サービスの取り組み事例】

- ・ソニーネットワークコミュニケーションズ株式会社：<https://fastlabel.ai/casestudy/snc>
- ・AWL 株式会社：<https://fastlabel.ai/casestudy/awl>

お問い合わせ：pr@fastlabel.ai

お問い合わせフォーム URL：<https://fastlabel.ai/contact-pr>

FastLabel 株式会社について

当社は、“AI インフラを創造し、日本を再び「世界レベル」へ”をパーパスに掲げ、2020 年 1 月の創業以来、教師データ作成代行やアノテーションツールの提供を主力事業として展開してきました。

現在では、AI 開発に必要な学習データ素材の収集、アノテーション、モデル開発、MLOps 構築までの全工程をサポートする事業を展開しております。

社名：FastLabel 株式会社

代表者：代表取締役 CEO 鈴木健史

事業内容：AI 開発を包括的に支援する「AIPaaS」(※) の提供（AI データプラットフォーム、アノテーションサービス）

設立：2020 年 1 月 23 日

本社所在地：〒151-0051 東京都渋谷区千駄ヶ谷五丁目 27 番 5 号

URL：<https://fastlabel.ai/>

※) AIPaaS…AI Process as a Service、AI 開発プロセスを高速化・高度化させるサービス、プロダクトの総称