



LIONBRIDGE

AI データ サービス



ライオンブリッジは、質の高いデータ収集、注釈付け、そして人間参加型 (HITL: ヒューマンインザループ) の評価を通じて、ML、AI、生成 AI のモデルの開発と改良をサポートする AI データ サービスを提供しています。当社では AI システムがさまざまな言語、地域、実世界の場面で正確に動作するように、多様な言語を母国語とするコントリビューターで構成されるグローバル人材コミュニティを活用して、文化的に適切なデータセットを提供し、正確なモデル評価を大規模に実施しています。

◎ 高品質なデータから始まる AI のライフサイクル ◎

ライオンブリッジのデータ サービスは、収集から注釈付け、そして評価まで、AI ライフサイクルのあらゆる段階をサポートします。当社は倫理的に収集された高品質なデータ、専門家による注釈付け、人間参加型 (HITL: ヒューマンインザループ) の評価を提供することで、正確かつ安全で実用可能なモデルの生成を可能にしています。



新しいモデルを構築する場合も、既存のモデルを最適化する場合も、当社のサービスはモデルの精度、安全性、関連性を高めることを目的としているため、プロトタイプ作成から実稼働までのプロセスを確実に加速させることができます。

◎ ML/AI チームが直面する課題 ◎



◎ ライオンブリッジの AI データ サービス ◎

1

データの収集

- ▶ 音声、動画、テキストのデータセット
- ▶ コンテンツ キュレーション
- ▶ 翻訳
- ▶ テキスト読み上げと音声文字変換
- ▶ 分類法の開発
- ▶ インテント発話の作成



2

データの注釈付け

- ▶ 音声、動画、テキストのコンテンツ分類
- ▶ テキスト、画像、動画の注釈付け
- ▶ エモーショナル キャプチャ
- ▶ センチメント分析
- ▶ コンテンツの調整
- ▶ 文章要約
- ▶ 固有表現抽出とリンク作成



3

開発サポート

- ▶ 多言語プロンプト エンジニアリング
- ▶ プロンプト応答とリライト
- ▶ RAG の最適化とパターンのサポート
- ▶ モデルのレビューと評価
- ▶ レッド チーミングと敵対的トレーニング
- ▶ 現地市場に合わせた最適化と文化的な適応の向上



4

モデルの評価

- ▶ インテント作成とレビュー
- ▶ モデルの出力の検証とランク付け
- ▶ ジオロケーションに関する検証と関連性
- ▶ 多様性と包括性のテスト
- ▶ 出力のファクトチェックと関連性テスト
- ▶ 検索、製品、広告の関連性



5

ワークフォース管理

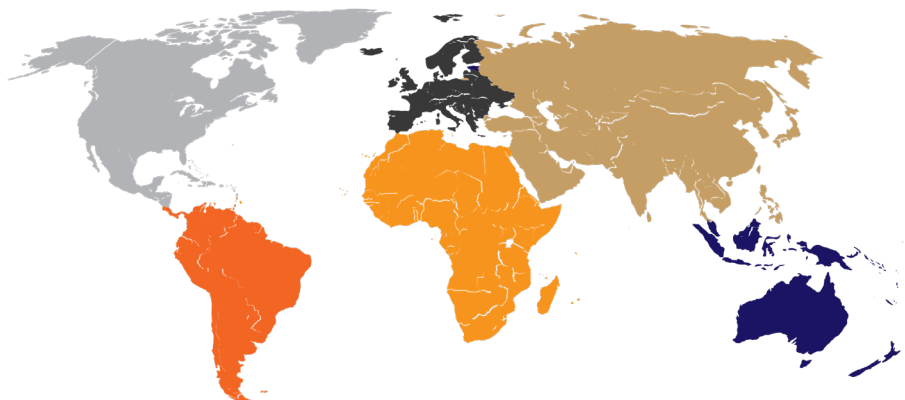
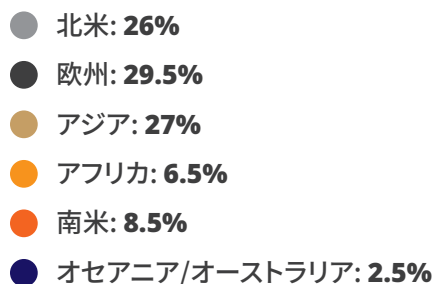
- ▶ AI 製品テスト
- ▶ セキュアな設備
- ▶ 計算言語学関連ソリューション
- ▶ データおよび ML エンジニアリング
- ▶ グローバルなコミュニティでの人材確保
- ▶ 専門知識



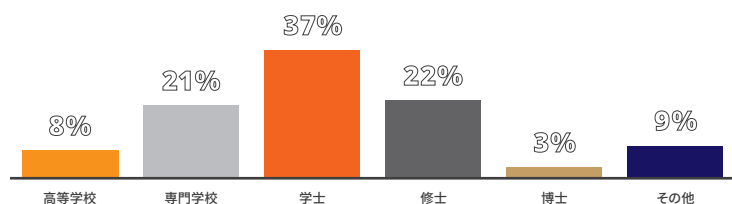


ライオンブリッジのグローバルな人材コミュニティには多様な言語や文化圏の人々が在籍しているため、あらゆるプロジェクトで幅広い視点と包括性が確保されます。当社ではこの多様性により、人間参加型 (HITL: ヒューマンインザループ) の評価の有効性を高め、AI モデルを実生活のニーズに適合させています。

高度に専門的で多様な専門知識と 関心を備えたグローバル コミュニティ



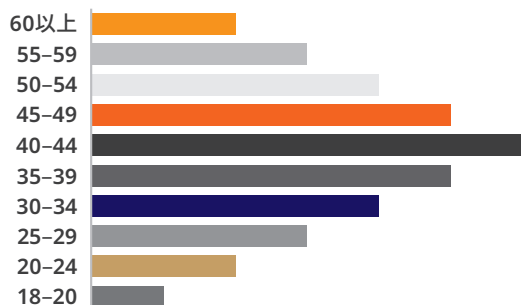
最終学歴



性別



年齢



60万名
のグローバル
コミュニティ

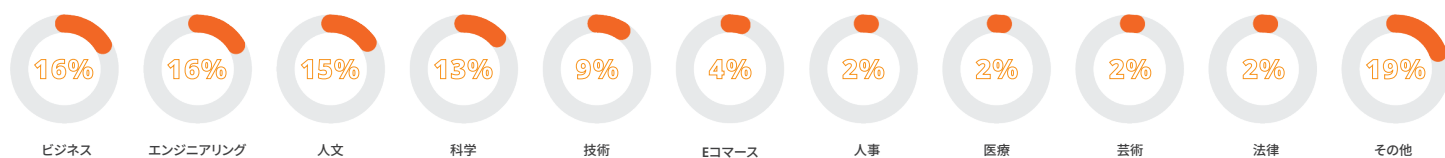


350以上
の対応言語



450以上
の地域

業種



このサービスが重要な理由



LIONBRIDGE AURORA AI STUDIO™ のご紹介

お客様のコンテンツ、アプリ、Web サイトなどのエンゲージメント強化にお役立ていただけます。オンデマンドでの豊富なプロジェクト ステータス分析データ、人員確保、タスク作成におけるカスタマイズ性などのメリットをお試しください。以下の機能に容易にアクセスできます。

- Web ベースのプロジェクト作成/管理ツール
- 管理されたエンドツーエンドの AI トレーニング ソリューション
- 50 万名の経験豊富なテスター、レビュー担当者、翻訳者からなるグローバル ネットワーク

ユースケース

データの収集



顔認識モデル

さまざまな照明条件、アングル、属性のもとで多様な顔画像を収集し、公平性と正確性を期すことができます。



ゴールデン セット

サンプル検索クエリ(テキスト)と予想される結果の画像、動画、文書などをペアにして、ゴールデンセットを作成できます。その結果を人間のレビュー担当者が検証することで、コンテンツの種類を問わず、正確なマルチモーダルのintent マッチングを行うことができます。

データの注釈付け



自動運転車両

画像やビデオフレームに、バウンディングボックスやポリゴンで歩行者、交通標識、車線標示などを示すラベルを付けることができます。



ヘルスケア AI

疾病検出モデル用の医療画像(X線、MRI)に注釈を付けたり、NLP タスクのために電子カルテ(EHR)にラベルを付けたりできます。



音声認識

音声クリップにトランスクリプション、話者 ID、感情タグなどのラベルを付けることができます。



レビュー用のセンチメント分析

レビューや投稿にセンチメント ラベル(肯定的、否定的、中立的)で注釈を付けることができます。

開発サポート



RAG (Retrieval-Augmented Generation、検索拡張生成) のセットアップ

外部のナレッジベース(企業 Wiki など)を統合するプロンプトを作成して、正確かつ最新の回答を得ることができます。



LLM によるカスタマー サポート

ブランドに合ったトーンで応答したり、特定のクエリをエスカレーションしたり、体系的な回答を提示したりするようにモデルを誘導するプロンプトを設計できます。

モデルの評価



機械翻訳 (MT)

人間参加型 (HITL: ヒューマンインザループ) の評価を行って、異なる言語やユースケースにおける複数の MT エンジンの出力を比較することができます。



偏見と公平性の監査

モデル (NLP、CV など) の評価を行うことで、さまざまな属性のグループ間でバランスの取れたパフォーマンスを確保できます。



チャットボット/LLM のパフォーマンス

評価担当者が、事前に定義されたガイドラインに照らして、チャットボットの応答の正確性、トーン、妥当性、安全性などを評価します。

詳しくはこちら

LIONBRIDGE.COM