

世界トップクラスのトレーニングデータでAI開発をリードするアッペン 「NexTech Week 第7回 AI・人工知能 EXPO【春】」にて 合成データとリアルワールドデータの併用について紹介！

世界トップクラスのトレーニングデータでAI開発をリードするアッペン（Appen）の日本法人であるアッペンジャパン株式会社（本社：東京都千代田区、代表：吉崎哲郎）は、合成データ・プラットフォームとリアルワールドの併用に関する最新のアップデート情報を、2023年5月10日（水）から12日（金）開催の「NexTech Week 第7回 AI・人工知能 EXPO【春】」にて紹介いたします。



当社は、AIライフサイクルデータを提供するリーディングカンパニーとして、AIに学習させたいデータにタグ付けを行う「アノテーション」によりAIモデルの精度を高めることができる「アッペンデータアノテーションプラットフォーム」を提供しています。AIデータにとって重要なのはデータのボリュームと正確性です。しかし、実際に十分なデータの量を確保することはむずかしく、確保できたとしても、データの正確性に欠けることが課題としてありました。こうした課題を解決するの現実世界から取得したデータではなく、実験室やコンピューターで生成したデータである「合成データ」の活用です。当社は2022年3月に、AI視覚システム用の合成訓練データを作成する世界有数のプラットフォームを開発するマインドテック・グローバルと提携し、370万ドルの資金調達を達成しました。また保有するデータ精度を高めるための「アッペンデータアノテーションプラットフォーム」により、合成データのみでは、担保することが難しい正確性を持ったデータを生成することができます。

アッペンジャパンは、2023年5月10日（水）から12日（金）開催の「NexTech Week 第7回 AI・人工知能 EXPO【春】」小間番号10-12にて、アッペンとマインドテック・グローバルの両社より合成データとリアルデータの活用事例や組み合わせ方、それらの特徴や使い分けについてご紹介させていただきます。著作権への配慮や危険が伴う環境など、開発期間やデータ準備期間なども合わせて適切な使い分けをご紹介いたします。

ぜひこの機会にブースにお立ち寄りください。

【合成データとリアルワールドデータの併用が必要な理由】

合成データとリアルワールドデータを併用することで、AIモデルの学習データセットにドメイン知識を取り込むことができます。ドメイン知識とは、AIモデルが動作するコンテキストに関連する私たちがすでに知っている事柄のことです。例えば、画像認識では、画像の回転などのオーグメント（合成データの種類）を使用し、AIモデルに有用なドメイン知識、つまり、分類には回転は関係ない（猫の写真を回転させても猫である）ことを教えます。同様に、コンピュータの動作に関するマシンデータのパターンが正常か悪意があるかを判断する際には、コンピュータが使用されている状況について詳しく知る必要があります。

また、AIのデータセットは不完全です。ロングテールのシナリオでモデルをトレーニングするための十分なデータがないことが多いです。AIの性能は、こうした頻度の低いシナリオでモデルがどれだけうまく振る舞うかで決まることが多くなっています。合成データとリアルワールドデータの併用はこれを解決することができます。

【合成データプラットフォーム「マインドテックカメレオンプラットフォーム」について】

当社はパートナー企業であるAI視覚システム用の合成訓練データを作成する世界有数のプラットフォームを開発するマインドテック・グローバルと提携し、カメレオン・プラットフォームを合成データのプラットフォームとして活用しています。マインドテックのカメレオン・プラットフォームは、AIデータ生成のあらゆる側面を正確にコントロールすることができ、必要とされる膨大な多様なデータセットを迅速かつ効率的に作成することが可能です。このプラットフォームは、使いやすさと柔軟性のためにUIとコマンドラインの統合を特徴としています。ア트워크からスマートアセットへの変換、データの拡張とキュレーションまで、エンドツーエンドのプラットフォームであり、AIを活用して現実世界の統計や特性を確実に保持することができます。実世界のデータは不足しており、ラベル付けに時間がかかり、絶えず変わるデータプライバシー規制を厳格に順守する必要があります。「ノーコード」の「セルフサービス」プラットフォームにより、AI開発者とデータ科学者は、コンピューターで生成されたシーンやシナリオを構築することで、高品質で正確にラベル付けされ、プライバシー規制に準拠した合成データの作成を迅速に自動化できます。

アッペンジャパン合成データに関するお問い合わせはこちらまで：<https://appen.co.jp/contact/>

【アッペン（Appen）について】

アッペン（<https://appen.co.jp/>）は、世界で最も革新的な人工知能システムを構築し、継続的なAIの学習や改善に欠かせない画像、文章、発話、音声、映像、その他のデータを収集、ラベリングするサービスを提供しています。アッペンは、235以上の言語に対応する専門知識、100万人以上の熟練したクラウドワーカーによって構成される全世界170カ国に及ぶグローバルネットワーク、業界最先端のAI支援型データアノテーション・プラットフォームによって、世界中のテクノロジー、自動車、金融、小売、そして医療メーカーのリーダー企業、および各国の政府に対して、高品質で、セキュリティレベルが高く、スピードが伴ったサービスを提供します。1996年に設立されたアッペンは、現在、オーストラリア本社をはじめ、アメリカ、イギリス、フィリピン、中国、韓国、日本のオフィスを拠点に、多くの顧客を支援しています。

【会社概要】

会社名 : アッペンジャパン株式会社
住所 : 東京都千代田区丸の内1-5-1 新丸の内ビルディング 9F
代表者名 : 吉崎哲郎
ホームページ : <https://appen.co.jp/>