

# 実践3：データにタグ付けする「アノテーション」

土井 伸洋

表1 主なアノテーション・ツール (Google アカウントなどの連携あり)

画像を扱えるもののうち、開始までの手間が少ない物を優先して掲載

名前	開発元	タイプ	費用	アカウント連携注2	URL	備考
Annofab	国内	クラウド	無料枠あり注1	○	https://annofab.com/	今回の実験で利用する。管理機能が強力で画像のほかLiDARデータにも対応している。自動車関連での実績が豊富
FastLabel					https://fastlabel.ai/	国内での実績が豊富で多様なアノテーション方式に対応している。また入力データ自体の購入サービスもある
ScaleAI	海外(米)				https://scale.com/	アノテーション・ツール最大手。コンペや研究用の大型データ・セット作成実績が多数ある
LabelBox					https://labelbox.com/	多様なアノテーション方式、多種の入出力データ形式に対応している
Hasty	海外(独)				https://hasty.ai/	各種アノテーションにおけるAIによる強力な入力アシスト機能がある
VoTT	OSS	スタンドアロン	無料	不要	https://github.com/microsoft/VoTT	マイクロソフトが開発/公開。2022年11月で更新が停止している
Label Studio					https://labelstud.io/	クラウド・ストレージを利用したデータ同期機能がある。有料のエンタープライズ版がある
CVAT					https://github.com/opencv/cvat	Intel社が開発/公開。既存のアノテーション・フォーマットへのエクスポート機能が充実している
Labelme					https://github.com/wkentaro/labelme	MITが開発したLabelMeにインスパイアされ作成されたOSS。Pythonで構築されている

※：2023年1月、筆者調べ、公開情報にもとづいて作成

注1：無料で始めることができる。一定以上のストレージ容量、先進機能、サポートなどを利用する場合は追加費用がある

注2：Google アカウントなどとの連携で即登録&開始できる

前章でアノテーションの仕様を決めたので、本章で実際にアノテーションの作業に入ります。ここでのアノテーション対象は画像とし、アノテーション・ツールにはクラウド型のAnnofabを使用します。アノテーション作業の手順はもちろんのこと、作業におけるポイントについても紹介します。

## アノテーション・ツールを選定するときのポイント

### ● 無料で利用できるツールが複数展開されている

アノテーションを行うツールを、アノテーション・ツールと呼びます。機械学習におけるデータの重要性が高まる中で、国内外のベンダから多数の有料、無料ツールが展開されているほか、オープンソース製品の開発も盛んに行われています。

アノテーション・ツールは、対象である画像や動画(点群データ、音声、テキストの場合もある)を読み込み、矩形/ポリゴン/塗りつぶしなどの作図ツールを使ってアノテーション作業を行うものです。アノテーション結果は、txtやJSON、既存のアノテーション形式で得られるものが多いです。

表1に主なアノテーション・ツールをまとめました。画像を扱えるもののうち、特に使用開始までの手間が少ないものを掲載しました。登録が不要、またはGoogleアカウント連携のみで利用可能ですので、試す際に参考にしてください。今回はこのうちの1つ、Annofabを実験で使います注1。

注1：筆者の属する会社で開発したのですが、汎用的に使ってもらえるので紹介しています。