



ヒトデータを AI・研究・ビジネスにつなげる ためのアノテーション事業

顔・身体・音声・生体データを扱う研究・開発のために、
学習データ設計から伴走いたします

データに「意味」を与え、活用可能な状態にする

アノテーションとは、画像・音声数値などのデータに対して

「これは何の状態か」「どのような意味を持つか」というラベル（タグ）を付ける作業です。

生のデータは、そのままではAIにとって意味を持たない単なる数値や記号の羅列に過ぎません。



画像データの場合

「異常がある／ない」「状態A／状態B」など



音声・波形データの場合

「通常／注意が必要」「発話区間／無音」など

アノテーションは、データを「使える状態」に整えるための重要な準備工程です。

「ヒトデータ」 のアノテーションとは

顔画像・音声・生体信号などの「ヒトデータ」においては、例えば次のような形で「人が見て判断できる状態」をラベルとして整理していきます。

顔画像の例)

- 汗が出ている状態
- 顔が紅潮している状態

音声・生体信号の例)

- 緊張している／落ち着いている状態
- 負荷が高い／安定している状態 など

その結果、データは単なる記録ではなく、状態を判定・予測する材料となり、**研究・AI・事業の意思決定に使える「生きたデータ」**になります。



「ヒトデータ」活用の難しさ

ヒトデータを用いた実際の研究・開発では、以下のような状況がよく起こっています。



生データはあるが
ビジネスにどう生かせばよいかわからない



学習データを
どのように作ればよいか判断できない



ヒトデータのアノテーションは
手間がかかるため
研究と並行して内製するのが難しい



顔画像や生体データなどの
個人情報の扱いに不安がある

ヒトデータを学習データとしてどう設計し、その判断や運用を誰が引き受けるのか
といった課題が、ヒトデータの活用が進まない大きな原因となっています。

オルチェのアノテーション

ヒトデータのアノテーションは、決まったやり方やテンプレートを当てはめれば成立する工程ではありません。扱うデータの性質や研究・開発の目的など様々な要素によって、求められる設計が大きく変わってくるからです。

創業以来、ヒト理解を専門に研究・開発を担ってきたオルチェなら、

ヒトデータ特有の曖昧さや個体差を前提に、お客様ごとの目的に合わせて「どのような意味をラベルとして与えるか」という形と判断基準を整理・設計し、それに基づくアノテーション作業を一体で行うことができます。さらに、その後の開発まで一貫して関わっていくことも可能です。個人情報にも配慮した安全な運用を行っています。



オルチェのアノテーションの特長

オルチェは、次の5つを特長としたアノテーション設計で、お客様の目的実現に伴走します。

01



人に特化

曖昧さ・個体差の理解

02



顧客最適

伴走型の設計

03



開発

まで担える

実装・運用視点

04



個人情報・ 生体データ への配慮

安全な取り扱い体制

05



ヒトデータ生成技 術による 学習データ拡張

合成データ・品質向上

01 人に特化

顔画像、身体、音声、生体信号などのヒトデータは、**正解が一つに定まらないことが多い**ため、曖昧さや個体差を前提に、観測項目・判断基準・粒度を設計する必要があります。

オルチェは、ヒトデータを専門領域として扱い、この前提に立ってアノテーション設計に向き合っています。



02 顧客最適（伴走型の設計）

アノテーション設計は「付けられる項目」から決めるものではなく、**最終的に何を実現したいのか**、データがどのように使われるのかから逆算して設計します。

オルチェは、観測項目・判断基準・粒度を、案件ごとにお客様と一緒に整理することを重視しています。

 **実現したいゴール**
データ活用の最終目的

↓ 逆算して設計

 **最適な項目・粒度**
顧客と共に要件定義

03 開発まで担える

アノテーション単体に閉じず、解析・AI活用・PoC・システム開発までを一貫して担うことも可能です。

オルチェなら、後段で成立しない設計（学習できない粒度、検証できない判断基準など）を事前に避けることができます。



04 個人情報・生体データへの配慮

顔画像や生体データは取り扱い制約が強く、運用・管理を前提に設計する必要があります。

オルチェは、個人情報・生体データを扱う体制と経験を前提に、設計・アノテーション・活用まで対応しています。



05 ヒトデータ生成による学習データ拡張

データ拡張とは、ただデータ量を増やすためのものではなく、以下のような目的で行います。

- ✓ 実データでは得られにくい **希少ケース・境界ケース** を補完する
- ✓ 設計した判断基準が学習・検証に耐えるかを確認する
- ✓ モデルの挙動を事前に検証する

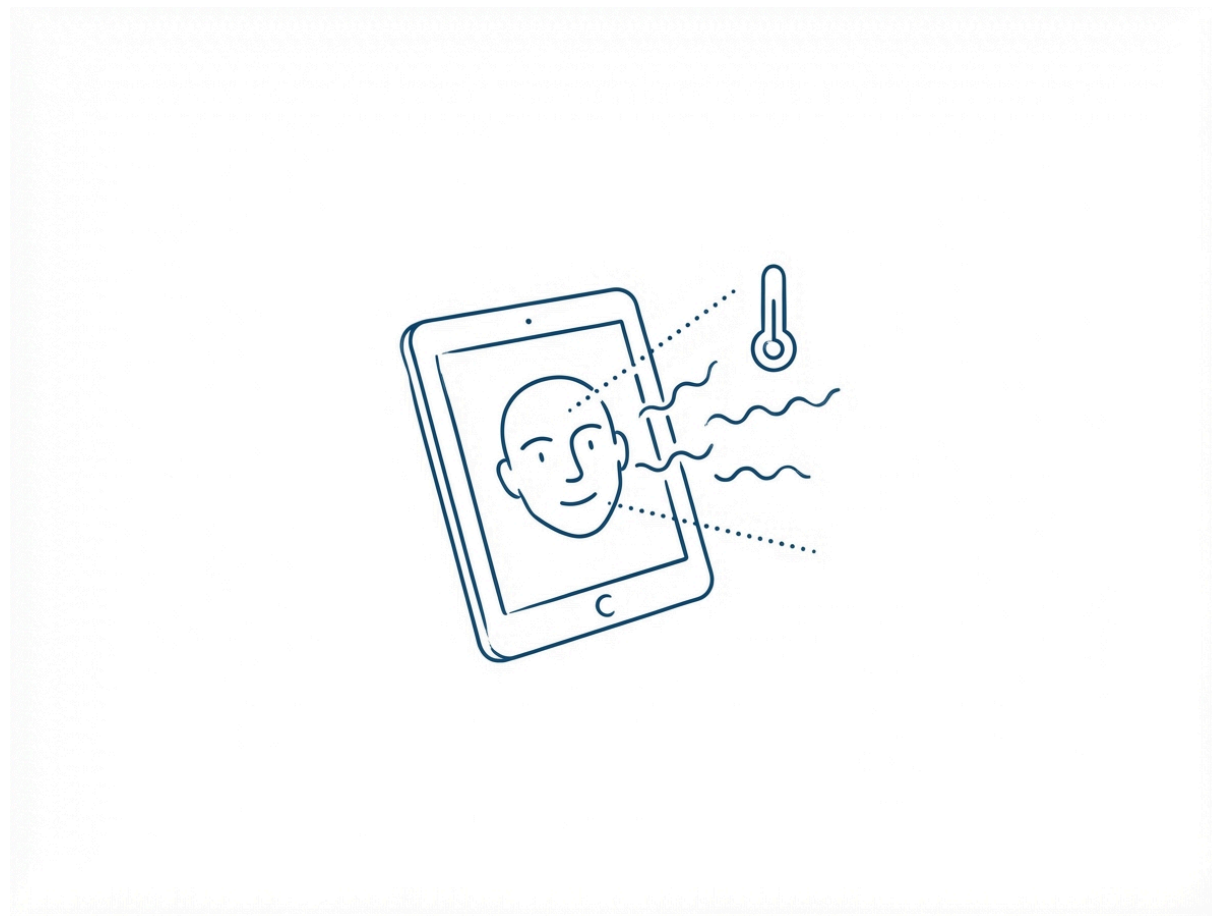
オルチェは、独自のヒトデータ（合成データ）生成技術を用いて、アノテーション設計と整合した形で、効果的に学習データを拡張することができます。



事例紹介 | 概要と目的

作業現場などにおいて非接触で取得可能な顔画像を用いて、**熱中症リスクを推定するAI開発**が進められていました。

そこで求められる学習データの作成において、**オルチェ**がどのようなアノテーション設計を行い、**企画実現につなげたのか**を、事例としてご紹介します。



Point1：何を観測するか (適切な分解)

熱中症リスクとは、深部体温上昇という生理学的に定義された状態を指します。

しかし、作業現場などにおいて深部体温を直接測ることはできないため、顔画像からどのような変化を兆候として扱うかを設計する必要がありました。

オルチェは、「深部体温上昇が顔にどのように現れるか」という観点から、右の3点を観測項目として定義しました。



Point2：判断基準の設計

(迷いを設計で扱う)

実画像では、分解した項目であっても判断に迷うケースが多く存在します。

- 紅潮と日焼けの区別
- 汗と照明反射の見分け
- 疲労感と個人差の境界線

などにおいて、基準が曖昧なまま主観で判断を進めると、学習データとして成立しなくなります。

そこでオルチェは、正解を無理に決めさせるのではなく、**判断基準と行動ルールそのものを設計し**、アナテーターの判断行動のブレを防ぎました。

設計された3つの行動ルール



断定できない場合の選択肢を用意する



複数の状態が同時に成り立つことを許容する



見える情報のみで判断する

Point3：中間生成物として扱い、学習・検証につなげる

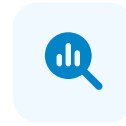
境界ケース・迷うケースほど、頻度は低い一方で、AI開発においてはとても重要なデータとなります。

オルチェでは、アノテーション結果を中間生成物として保持することにより、こうしたケースも適宜拡張・検証を行いながら、実用性の高い学習データにつなげました。



中間生成物として保持

アノテーション結果を正解ラベルではなく「判断の状態」としてそのまま記録



分析・特定

判断が割れた領域・境界ケースを特定し
足りないデータ分布を可視化



拡張・補強（合成データ）

境界ケースを重点的に生成・拡張し
学習データを補強する

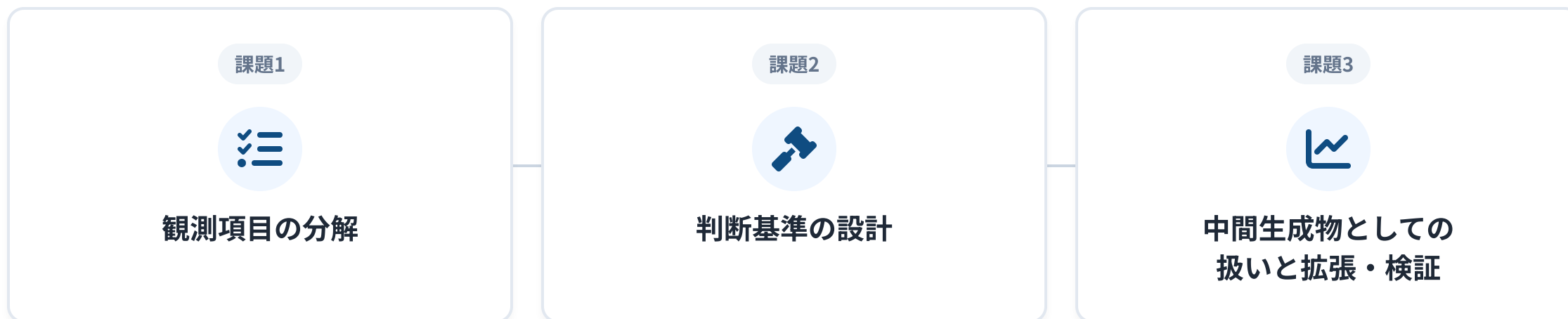


検証（テスト）

微妙なケースに対するモデル挙動を確認し
実用性を担保する

このCASEが示していること

このケースでは、以下の3つの要素を**一体の設計工程**として扱っている。

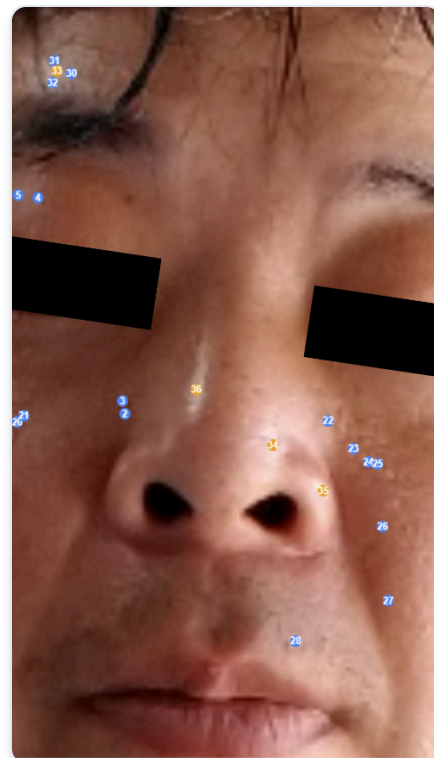


ヒトデータの**アノテーション**を「**設計工程**」として引き受けることで、**初めて実用的な学習データが成立します。**

アノテーションの具体例

実際のアノテーション作業において、汗粒や汗のテカリにラベル付けを行っている例です。

個人差、照明条件、眼鏡や帽子などの装飾品がある場合でも、定義された基準に従って精密に付与します。



i 青・黄色のドットと数字は、実際に付与されたアノテーションポイント（ランドマーク）を示しています。

まずは、状況整理から

アノテーション設計に悩んだ段階で、相談できる先として――

👉 状況と目的の整理から伴走

オルチェのアノテーション事業は、作業を請け負う前に、状況と目的を整理するところから関わります。

- ✓ まだ仕様が固まっていない
- ✓ 何をどこまでアノテーションすべきか分からない
- ✓ 学習や検証につながるか不安がある

といった段階での相談を前提としています。

☞ 具体的な相談例

例えば、次のような相談に対応しています。

- ❓ この企画で、アノテーションはどこが重要か
- ❓ 観測項目・判断基準・粒度は妥当か
- ❓ アノテーションを内製すべきか、外注すべきか
- ❓ 個人情報・生体データの扱いは問題ないか
- ❓ 将来的な学習・検証・拡張まで見据えた設計になっているか

※アノテーション単体の相談も、設計のみの相談も可能です。

まずは、現在の状況をお聞かせください。企画段階・検討段階からでも問題ありません。

アノテーション設計に悩んだ段階で、相談できる先であること
それが、オルチェのアノテーション事業です。



MAIL CONTACT

info@oruche.co.jp