

アッペンジャパン概要

～世界No1アノテーションカンパニーが提供する
次世代のAI精度を向上するサービススキーム～

2024.2.22

アッペンジャパン株式会社





グローバルブランドから信頼されるAppen

ミッションクリティカルなお客様AIを支え続ける価値

Consumer tech



B2B tech



Hardware



Transport



Retail & ecommerce

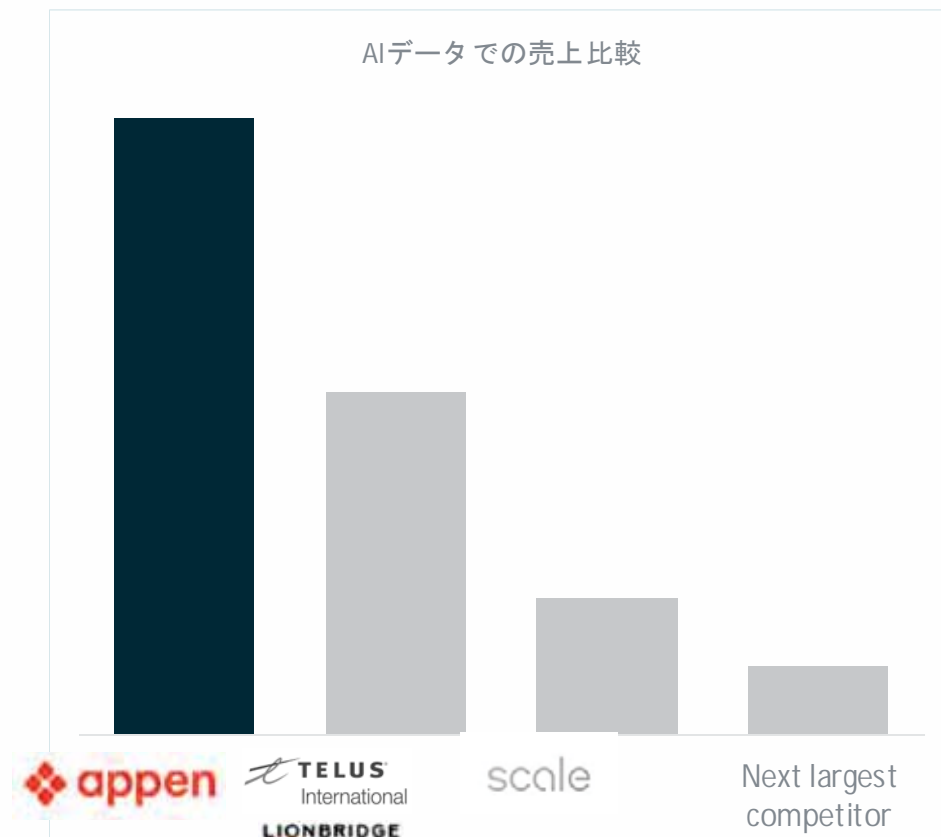


.. many more



グローバルトップの売上

トップのマーケットシェア



1. Lionbridge revenue source from Telus International IPO and other statements, Scale AI revenue estimated from public announcements related to funding rounds

This content is for internal use only

Performance

\$447M Record revenue in FY21

\$78.9M FY21 EBITDA Before FX Impact

40% Annualised revenue growth since 2016

32% Global Services growth 2H21 on 1H21

422% Revenue growth in China in FY21



グローバル規模の体制



米国

シアトル (US Headquarters)
ダラス (Secure Facility)
サンフランシスコ
デトロイト
ワシントンDC



100万人+

専門性を持つクラウドワーカー



70,000+

活動地域



グローバル

東京
シドニー (Corporate Headquarters)
エクセター (Secure Facility)
キャビテ (Secure Facility)
大連・無錫・北京 (Secure Facility)
上海
シンガポール



292+

言語カバレッジ



170

カ国



アッペンの最新サービス

お客様のAIプロジェクトを加速させる4つのモデルトレーニングサービス

アノテーション

画像・テキスト・ビデオ・点群・オーディオに最先端の技術でアノテーションを実施

例)

- 帳票OCR AI向けB-Box付与
- 不良検知AI向け欠陥ラベリング



プラットフォーム

自社の高機能プラットフォームにより、短期間に正解率の高いアノテーション結果を提供

仕様)

- オンプレ/SaaS選択可能
- ユーザ課金ライセンス契約
- セットアップ、QAサービス付き



翻訳

専門のエキスパートをアサインし、確実な翻訳結果をAIやML向けに提供する事が可能

例)

- 自動車ダッシュボード表示翻訳(複数言語)トライアル



データ提供

スタジオ・車内・オフィスなど様々な場所で音声・テキスト・画像・動画・混合データなどを収集

例)

- 自動翻訳AI向け発話データ(多言語)
- 自動車ナンバープレート画像収集



あらゆるデータタイプに対応

画像、音声、動画、テキスト、点群データ、マルチモーダル、POIなど全てに対応

Image



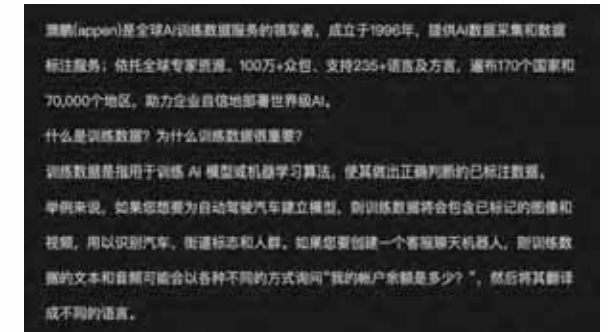
Audio



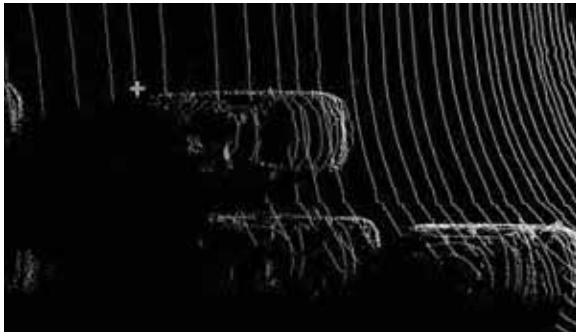
Video



Text



Lidar/3D



Multi-Modal



Synthetic Data



POI



アノテーションツールの提供

様々なデータへのキャプションニング、ラベリングを付与

- AppenのアノテーションツールにAIプリラベリング機能搭載
- データに独自のラベル、属性情報を付与可能
- 必要に応じてJSON形式等で出力



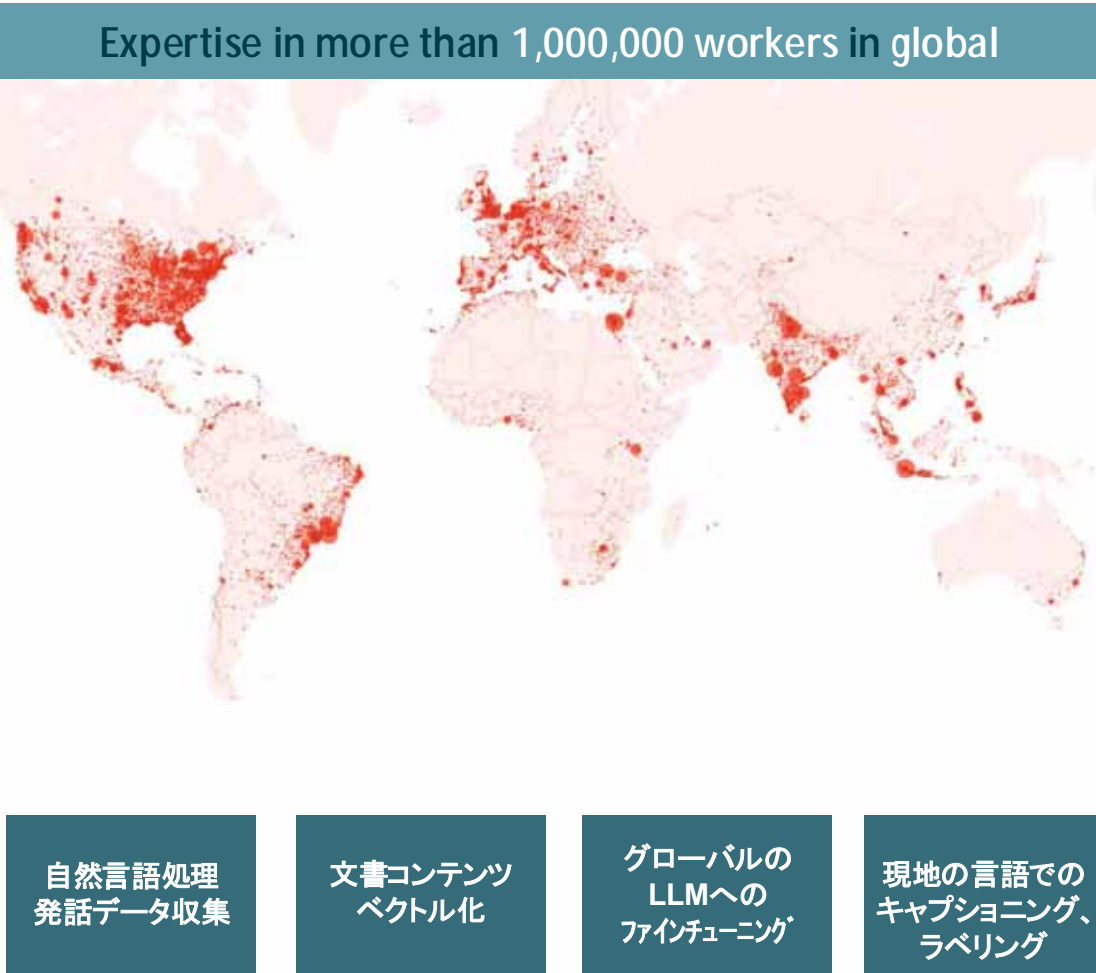
高度なユーザビリティプラットフォーム

- ツールをSaaS/オンプレミスで提供
- 画像、テキスト、音声、マルチモーダル等に対応
- 自動付与ロジック追加可能
- データタイプ毎のライセンス課金体系
- オプションサービスあり



100万人を超えるクラウドワーカーの協業

170カ国、292言語に対応可能なエキスパートがお客様LLMをサポート



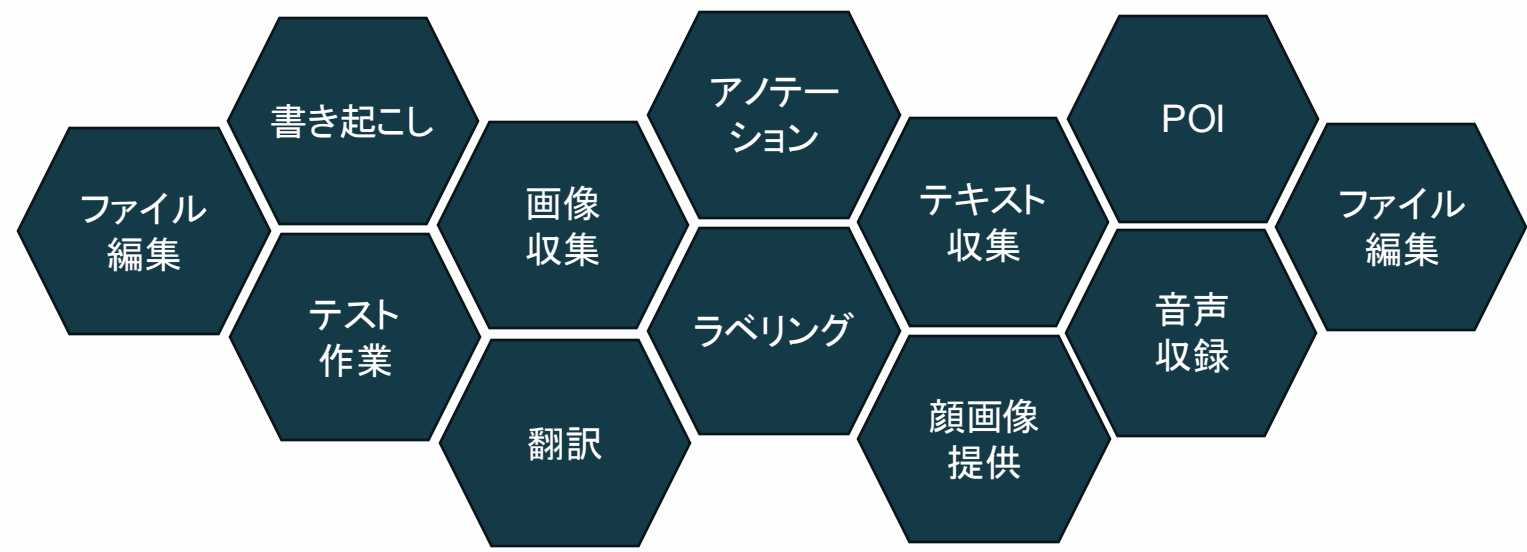
取り扱い可能言語（一部）

Achinese	Bengali	Chinese Traditional	Fijian
Afrikaans	Betawi	Chinese	Filipino
Algerian Arabic	Bhojpuri	Corsican	Finnish
Algerian Saharan Arabic	Bihari	Creek	French
Amharic	Bislama	Croatian	Galician
Arabic-10+ varieties	Bodo	Cypriot Arabic	Gan Chinese
Aragonese	Borana-Arsi-Guji Oromo	Czech	Ganda
Arbëreshë Albanian	Bosnian	Danish	Georgian
Armenian	Breton	Dari	German
Arvanitika Albanian	Brunei	Dhivehi	Gheg Albanian
Assamese	Bulgarian	Dotyali	Goan Konkani
Asturian	Burmese	Dutch	Greek
Awadhi	Campidanese Sardinian	Dzongkha	Gujarati
Bacanese Malay	Catalan	Eastern Canadian Inuktitut	Gulf Arabic
Bagheli	Cebuano	Eastern Mari	Haitian Creole
Balinese	Central Bikol	Eastern Oromo	Hakka Chinese
Bambara	Central Khmer	Eastern Yiddish	Halh Mongolian
Bashkir	Central Kurdish	Egyptian Arabic	Haryanvi
Basque	Central Malay	English-15+ varieties	Hassaniyya
Bavarian	Central Pashto	Esperanto	Hausa
Belarusian	Chhattisgarhi	Ewe	Hawaiian
Bemba	Chinese Simplified	Faroese	Hebrew



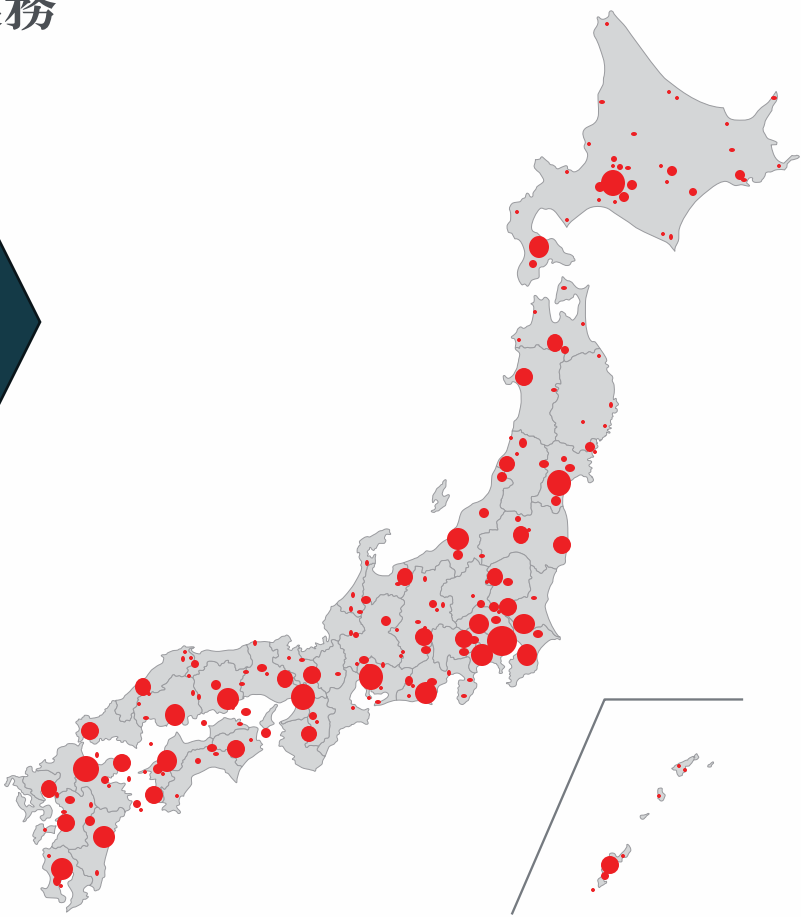
日本でのクラウドワーカービジネス

日本国内6万人を超えるクラウドワーカーによる
シンプルタスクの一括オペレーション及び情報収集業務



クラウドワーカー依頼例（調査系）

東京都内5万件の飲食店情報収集	道路看板の文字書き起こし
ガソリンスタンド100箇所動画収集	全国農作物収穫量調査のチェック
自動車左折動画の撮像	積雪・霧での走行画像収集



クラウドワーカー分布イメージ

高度なセキュリティとコンプライアンス

4つの施策により徹底したデータ管理を遂行



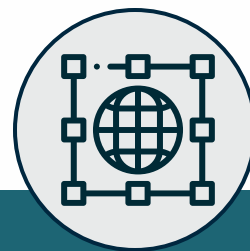
セキュアな体制(認証)

- GDPR
- SOC 2 Type II
- ISO27001
- HIPAA



セキュアなワーカー

- 豊富な経験を保持
- 事前研修
- スクリーニング
- ペナルティ制度



セキュアな拠点

- セキュアなサイト
- 認証の継続取得
- セキュアマネージドサービス
- 厳選された管理者のみによるアクセス



セキュアな環境

- 堅牢データセンターを利用
- ダウンロード不可
- 再ログインの制限
- アサインされたリソースのみログイン



コンピュータビジョン事例1

-電子機器メーカー様-

AI-OCR帳票認識の性能評価向けに2000枚以上の帳票提供とアノテーション実施

会社概要：

企業名：某企業

【業界】 家電、エレクトロニクス機器

【業務】 製造、販売、ソフトウェア開発

【従業員】 40,000人以上

【売上】 2兆5,000億円以上

【拠点】 グローバル各所

ビジネス状況(課題)：

- 納品書、請求書等の日本語帳票の生成、チェックAI-OCRを構築
- 大量の評価用学習データの収集とB-Boxなどのアノテーション作業が短期間で必要
- 将来的な拡張を踏まえ、新たなデータ提供ベンダーを検討

サービス提供内容：

- 請求書、領収書、納品書、見積書を合計2,200枚を顧客情報や印影画像をダミー名称に修正して提供
- 帳票内のタイトル、文言、表にバウンディングボックスをセットし、顧客指定のラベル情報を付与。JSONファイルに変換して提供
- 学習を止めないよう、完成分を分割して都度納品

アッペンを選んだ理由：

- グローバルで教師データを提供している実績あり
- データ提供とアノテーションの両方をワンストップでデリバリー
- 大量データをタイムリーに提供できるリソース体制
- 高品質で精度の高いアノテーション作業の実施

実績(ビジネス効果)：

 **2,200枚**  **8週間**  **95%**

提供期間 合格率



コンピュータビジョン事例2

-エンジン、コンプレッサー製造企業様-

欠陥部品判定AIのモデル構築に向けた線キズ、異物付着等の特殊画像データを提供

会社概要：

企業名：某企業

【業界】 繊維機械、自動車部品メーカー

【業務】 エンジン、コンプレッサ設計、製造

【従業員】 70,000人以上

【売上】 2兆5000億円以上

【拠点】 グローバル各所

ビジネス状況(課題)：

- 製造工場でのDXの一環として、不良品検知のAI化を推進
- 9つの製造ラインにおいて、異常データのサンプル画像の収集が必要
- お客様の製造部品に近いプロダクトの線キズ、異物付着、切り子、亀裂、鋳巣などの画像が必要

サービス提供内容：

- エンジン、コンプレッサー部品の欠陥画像をリアル画像で収集。正常データと異常データを1セットとし、60セット180枚を提供
- 欠陥箇所に対して、バウンディングボックスをセットし欠陥クラスを付与。最終的にPascal VOC形式で提供
- お客様指定のデータ格納ファイル構成、ファイル命名規則に従って、堅牢なセキュリティ管理下のデータセンター経由で提供

アッペンを選んだ理由：

- 特殊な欠陥データの収集が可能
- 画角、背景マスキング、解像度などの細かな要件に対応
- プロトタイプ向けの小規模データ収集が可能
- Pascal VoCへの変換要件に経験豊富な技術者が対応

実績(ビジネス効果)：



5種類
欠陥データ

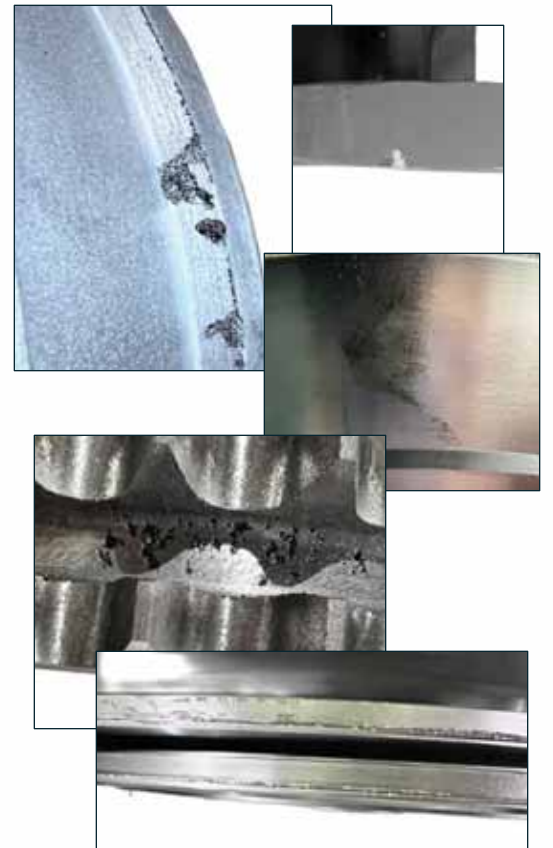


Pascal
VoC対応



提供期間

6週間



コンピュータビジョン事例3

-家電、エレクトロニクス企業様-

製造、ロジスティクス、建設などの作業進捗AIの実現に向けた空間認識データを生成

会社概要：

企業名：某企業

【業界】 家電製造、エレクトロニクスメーカー

【業務】 家庭機器、産業機器の製造、販売

【従業員】 240,000人以上

【売上】 7兆3000億円以上

【拠点】 グローバル各所

ビジネス状況(課題)：

- モノづくり、運搬といった作業の進捗を正確に把握する管理ソリューションのニーズが顕在化
- AIを使った作業工程管理と進捗状況を把握する自動化技術が未成熟
- 研究を通じて海外の大学と共同で論文を作成する必要があるが、成果を出すためのアノテーション作業が膨大

サービス提供内容：

- 機器製造、物販などの作業を撮影した擬似空間ビデオデータ200本以上に対するアノテーション作業
- 作業者と机、マザーボード、ドライバー、はんだごてなどの物体の位置関係と関連性を複数のクラスから選択して付与
- 位置関係をB-Boxによる座標出力し、関係性を事前定義したクラスでJson形式で出力

アッペンを選んだ理由：

- 大量データを短期間で対応することが可能
- 細かなアノテーション要件を迅速、かつ正確に把握
- お客様の予算にミートした価格での契約が可能
- 今後のデータ収集のサービスにも対応可能なCapability

実績(ビジネス効果)：



5,246
ファイル出力



合格率

94%



Json
対応



テキストアノテーション事例 1

-大手広告代理店D社-

各業界に特化したSNS分析・マーケティング活用ソリューションのためのコーパス(辞書)作成作業

会社概要：

企業名：	某企業
【業界】	広告、出版。マーケティング
【業務】	広告制作、デジタルマーケティング、コンサルティング、事業開発
【拠点】	日本国内メイン

ビジネス状況(課題)：

- ネット上の企業/製品に対する評価をボジネガ分析するニーズは非常に高いものの、自然言語のため、定量把握には毎回手間がかかり大変な分析作業が求められていた。
- 特にSNS投稿では略語なども多く、文法上も正しい文章とは限らないため、正しい評価ができていたとは言えなかった。
- SNS投稿をカバーする業界別の汎用辞書を作成し、AIによる自然言語処理と組み合わせることで精度の高いマーケティング分析が可能となる。

サービス提供内容：

- SNSからある特定の業界・企業・製品について記載されたコメントを収集。1 カテゴリに対して1万件のコメントを収集。
- 収集したコメントに対して、ボジネガ分析を行う。ポジティブの場合は想起・期待・レビュー・レコメンドといった形でさらに細分化。
- 既存の辞書に含まれないSNS特有の単語・表現を人手により抽出し、カテゴリ分け、正しい意味の付与、マーケティング観点からのラベル付けを行う。

アッペンを選んだ理由：

- 大量のクラウドリソースが利用可能なため、SNS用語に詳しい作業者を事前テストにより応募者の中からスクリーニング。より高い精度で分析作業が可能。
- 予算をあらかじめ設定しておき、プロジェクト中に予算の範囲内であれば仕様変更に対応していくという、柔軟なプロジェクト進行が可能。
- 日本語だけでなく中長期的にグローバルな辞書作成にも問題なく対応できる。

実績(ビジネス効果)：



1万件のコメント分析



提供期間

4週間



THANK YOU

