

ポルトガル編 Part 2： AI開発のため効率的なデータ基盤を 提供する DefinedCrowd

古田 哲也

DefinedCrowd Japan株式会社 代表執行役社長



JOI Venturesポルトガル編Part 2をお届けします。JOIとポルトガル投資貿易振興庁 (aicep) がタイアップして、日本でのビジネス拡大、日本企業との連携・協業を模索する、次世代を担う有望なベンチャー企業を紹介します。

今号は、AIの開発や導入に欠かせない機械学習用のトレーニングデータを、ヒューマン・イン・ザ・ループ (人間参加型) によって最適化する事業を展開する DefinedCrowdです。世界53カ国で10万人のタスクワーカーを「Neevo」というクラウドソーシングコミュニティとして抱え、音声、自然言語、コンピュータービジョンの3つの分野で高品質なトレーニングデータを提供しています。日本法人社長の古田氏にうかがいました。

これまでの歩みについて、お聞かせください

DefinedCrowd は、米ワシントン州シアトルに本社を構え、ポルトガルのリスボンとポルト、そして東京に拠点をもつグローバル企業です。CEOであるダニエル・ブラガは、2015年にシアトルで会社を立ち上げ、

R&Dセンターをリスボンとポルトに開設しました。

そして、スマートデータのためのプラットフォームを開発し、2016年には、シードラウンドの段階で、110万USドルに上る資金を投資家から集めることに成功しました。その投資家には、SONY、Amazon Alexaファンド、ポルトガル・ベンチャーズ (ポルトガル・フォメント銀行グループが出資する投資ファンド)、ビジーエンジェルズ (現Bynd Venture Capital) などが含まれています。

シリーズAラウンドでは、1180万USドルを、エヴォリューション・エクイティ・パートナーズ主導で集めるに至り、2018年、東京に事務所を開設しました。

その後、シリーズBラウンドでは、2020年に5050万USドルの資金を集め、これはアメリカで女性によって起業されたAI関連の企業において、最高額となりました。

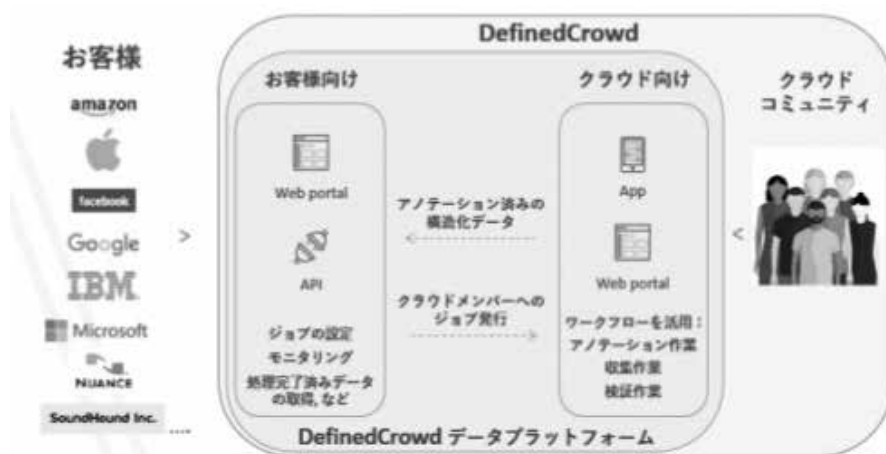
当社は、CBインサイツの「AIスタートアップ企業トップ100」や、WIREDの「最もホットなスタートアップ100社」、Forbesの「将来有望な米国のAI企業トップ50」に選ばれています。

世界には、AIトレーニングデータを収集する多くの企業があります。DefinedCrowdの強み、他の企業との違いは何でしょうか

私たちが提供する効率的なデータワークフローを用いることで、データサイエンティストは、トレーニングデータを収集、合成、充実、構造化することができます。企業の機械学習、AIイニシアティブ・トレーニングやモデリングの速度を上げるため、ヒューマン・イン・ザ・



ダニエル・ブラガCEOとDefinedCrowdの仲間たち



パートナーシップを締結しています。

データ活用が進むことで、どのようなビジネスチャンスが日本企業にもたらされると考えますか

日本企業には、事務手続のためのコストを減らし、音声認識の分野において顧客満足度を上げる可能性が十分にあります。コールセンターの応答、病

院の遠隔診療、カーナビの音声案内、さらにはテキスト・アプリで政府機関の文書管理、チャットボットで企業の情報共有といったことにまで広がります。

ヘルスケアの分野にもさらなる大きな可能性を見出しています。なぜならヘルスケアでは、異なる方言、アクセントを話す高齢者の言葉を理解することが必要とされているからです。

AIがハイレベルな出力を生み出すためにはデータの品質が重要であり、高品質なデータが新たなビジネスチャンスをもたらすでしょう。

将来に向けたビジョンをお聞かせください

DefinedCrowdでは、よりスマートな未来社会に向かって、人間と機械の自然な相互作用を創造するというビジョンをもっています。日本においては、音声認識、文字認識といった分野において、社会に大きく貢献したいと考えています。

人間とデジタル機器がボーダーレス化された社会の中で、誰もが必要としているものにアクセスできるようになると信じています。



ループ、全自動ツール、機械学習の能力を組み合わせることによって実現しています。私たちのヴァリュー・プロポジションは、品質、スピード、スケール（50以上の言語をカバー）であり、言語技術を駆使した自然な音声が強みです。

これまでに導入された企業や、提携事例などを教えてください

創業以来、トップクラスのAI企業やフォーチュン500にランキングされる企業に、音声、イメージ、テキストといったAIアプリケーションに必要なトレーニングデータを、プロジェクト固有の高品質データとして提供しています。クライアントには、Yahoo Japan、IBM、BMWやマスターカードといった企業があり、ヴァーチャルアシスタントやカスタマーサービスチャットボット、イメージ分析などのトレーニングに、当社のシステムを利用しています。

当社では、AIの出力能力を向上させるため、音声認識におけるワードエラー率を5%以下に抑えています。ワードエラーの減少により、エンドユーザーの使用感を著しく高め、結果として顧客満足度を上げているのです。

たとえば三井住友銀行では、当社が数時間で作成した数百件の「質問と回答」をAIに学習させ、きめ細かい学習データにより回答精度の向上に寄与することが確認できました。その結果、AIの学習データ作成作業は行員と遜色ない品質で対応可能と判断され、銀行業界で初めて当社へ業務委託いただきました。

また、ディープラーニングを活用して音から価値を創造する国立研究開発法人産業技術総合研究所発のベンチャー、Hmcomm株式会社とは、AIによる音声と感情のデータアノテーション領域における戦略的

お問い合わせ

JOI事業企画部

E-mail: bd@joi.or.jp、TEL: 03-5210-3311

DefinedCrowd Japan 株式会社
代表執行役社長 古田 哲也
<http://www.definedcrowd.com/>

ポルトガル投資貿易振興庁 (aicep)
東京オフィス 奥村 貴子
E-mail: takako.okumura@portugalglobal.pt
<http://www.portugalglobal.pt>