

学習によって識字率がアップする AI サービス 手書き文字をデータ化する「Tegaki」提供開始

手書き文字の認識率 99.22%*の研究を元に
データ入力作業の効率化とコスト削減を実現

株式会社 Cogent Labs（本社：東京都渋谷区、代表取締役：飯沼純、エリック ホワイトウェイ、以下「コージェントラボ」）は、2017年8月より、手書き文字を高精度でデータ化する「Tegaki（<https://tegaki.ai/>）」の一般販売を開始いたします。

「Tegaki」は、既存技術では自動認識による読み取りが難しくキーボード入力が必要だった手書き文字を、スピーディーに高精度でデータ化します。独自開発の AI 技術でデータを処理・学習することによって、読み取り精度は、継続的に向上します。

コージェントラボは「Tegaki」を通じて、データ入力にかかる業務の効率化とコスト削減、労働生産性の向上を、業界を問わず実現します。

*「Tegaki」の認識率について、すべての手書き帳票で初回から 99.22%の認識率を保証するものではありません。99.22%の認識率については末尾をご覧ください。

Tegaki 利用イメージ(フォーム設定画面)



■サービス概要

「Tegaki」は、製造、販売、金融、医療、サービス、教育など、あらゆる業界で使われている、手書き帳票に書かれた文字を高精度に認識できるサービスとして開発しました。手書き文字と活字を同時に認識することが可能です。

読み取りたい帳票はWEBブラウザ上で指定します。データの認識と読み取り結果はAPIで提供しており、現在ご利用中のシステムと接続して手軽に導入できます。

「Tegaki」4つの特徴

- ✓学習性：読み取り精度の継続的な向上：ディープラーニングを活用したアルゴリズムから構成され、データを処理しながら学習することで、読み取り精度が継続的に向上。
- ✓拡張性：他言語への対応：トレーニングデータを準備する事で各種他言語の手書き文字読み取りへの適用可能。業界用語への対応：言語モデルを備えており、業界特有の言語モデルと連携すれば読み取り精度を高めることが可能。
- ✓利便性：プロセスの簡素化：活字・手書き文字双方に加え、チェックボックス・丸囲い文字などを一つのサービスで読み取れるため、フォーム毎OCRとの使い分けや選別をする手間が不要。
- ✓経済性：コスト削減：シングルライン10文字（ひらがな5文字、カタカナ5文字）の場合、人による入力17.6円～のところ、tegakiでは1円以下と、およそ94%のコスト削減を実現※1)

コージェンラボでは、これまでに十数社とPoC（Proof of Concept：コンセプト検証）を実施、さまざまな検証を重ねて今回のサービス開始に至りました。すでに検証を経て、複数社に導入が決まり、他にも多くの企業が導入検討を進めています。「Tegaki」は、銀行口座・クレジットカードなど金融機関や電気・ガス・水道など公共料金の申込書、お客様の声やアンケートの用紙、病院の問診票や診断書、テスト答案など、幅広いシーンでの活用を想定しています。

※1 日本データエントリー協会 平成28年(2016年)度 データエントリー料金資料より (<https://www.jdea.gr.jp/index.html>)

■サービス価格

いくつかのプランをご用意しています。スタンダードタイプは月額20万円(20万円分のデータ化費用含む)から。1フィールドはボックスキャラクター0.2円、シングルライン0.8円となります。帳票1枚当たりのデータ化費用は、その帳票のフィールド数とフィールドのタイプによって異なります。

エンタープライズ版では、セキュリティおよび個人情報保護などへの対応について、個別に専任担当がついて導入まで相談/サポートさせていただきます。エンタープライズ版の価格は個別にお見積りとなります。

フィールドのタイプについて

品番	品名	単価	数	金額
0123456	千原印刷 A4 サイズ	2088	10	20,880

品番: 6 x ボックスキャラクター
 品名: 1 x シングルライン
 単価: 1 x シングルライン
 数: 1 x シングルライン
 金額: 1 x シングルライン

■参考：増え続けているドキュメント処理の需要

マイナンバー関連業務、ストレスチェック制度の導入、電力小売／ガス小売自由化などの制度改革に伴いドキュメントは増え続けており、国内ドキュメントアウトソーシングサービス市場の2015年～2020年の年間平均成長率は7.1%、2020年の市場規模は4,023億4,500万円と見込まれています。(※2)

企業に業務改善としてのドキュメント処理効率化ニーズが元々ある中、このようにドキュメントが増えることによって、インソーシング、アウトソーシングともデータ入力業務は増えており、ユーージェントラボが2016年6月にAIによる手書き認識についての研究を発表して以来、多くの問い合わせを頂いております。

※2 IDCより (<http://www.idc-japan.co.jp/Press/Current/20161026Apr.html>)

■認識率 99.22%について

ユーージェントラボの開発したエンジンを使用して下記の文章を認識した場合の認識率は99.22%となります。

元の手書き (5人の筆跡)

認識結果

<p>青銅の魔人 江戸川乱歩 童車ゆき</p> <p>冬の夜、月のさえた晩、銀座通りに近い橋のたもとと交番に、ひとりの警官が夜の見はりについていました。一時をとくにすぎた真夜中です。ひるまは電車やバスや自動車が、縦横にはせちがう大通りも、まるでいなかの原っぱのようにさびしいのです。月の光に、四本の電車のレールがキラキラ光っているばかり、動くものは、何もありません。東京中の人々が死にたえてしまったようなさびしさです。警官は、交番の赤い電灯の下に、じっと立って、注意がかくあたりを見まわしていました。濃い口ひげの下から、息をするたびに、白い煙のようなものが立ちのぼっています。寒さに息がおおるのです。「オヤ、へんなやつだなあ。よっぽらいかな。」警官が思わずひとりごとをつぶやきました。キラキラ光った電車のレールのまんま中を、ひとりの男が歩いてくるのです。青い色の背広に、青い色のソフトをかぶった大男です。この寒いのに外套も着ていません。その男の歩きかたが、じつにへんなのです。お巡りさんが、よっぽらいかと思っただけ、むりはありません。しかし、よく見ると、よっぽらいともちがいます。右ひだりにヨロヨロするのではなく、なんだか両足とも義足でもはめてあるような歩きかたなのです。人間の足で歩くのではなく、機械でできた足で歩いているような感じなのです。顔は帽子のかげになって、よく見えませんが、なんだかドス黒い顔で、それが少しもわき見をしないで、夢遊病者のように正面をむいたまま、ガックリガックリ歩いているのです。</p>	<p>青銅の魔人江戸川乱歩の音冬の夜、月のさえた晩、銀座通りに近い橋のたもとと交番に、ひとりの警官が夜の見はりについていました。一時をとくにすぎた真夜中です。ひるまは電車やバスや自動車が、縦横にはせちがう大通りも、まるでいなかの原っぱのようにさびしいのです。月の光に、四本の電車のレールがキラキラ光っているばかり、動くものは、何もありません。東京中の人々が死にたえてしまったようなさびしさです。警官は、交番の赤い電灯の下に、じっと立って、注意がかくあたりを見まわしていました。濃い口ひげの下から、息をするたびに、白い煙のようなものが立ちのぼっています。寒さに息がおおるのです。「オヤ、へんなやつだなあ。よっぽらいかな。」警官が思わずひとりごとをつぶやきました。キラキラ光った電車のレールのまんま中を、ひとりの男が歩いてくるのです。青い色の背広に、青い色のソフトをかぶった大男です。この寒いのに外套も着ていません。その男の歩きかたが、じつにへんなのです。お巡りさんが、よっぽらいかと思っただけ、むりはありません。しかし、よく見ると、よっぽらいともちがいます。右ひだりにヨロヨロするのではなく、なんだか両足とも義足でもはめてあるような歩きかたなのです。人間の足で歩くのではなく、機械でできた足で歩いているような感じなのです。顔は帽子のかげになって、よく見えませんが、なんだかドス黒い顔で、それが少しもわき見をしないで、夢遊病者のように正面をむいたまま、ガックリガックリ歩いているのです。</p>
---	---

■Cogent Labs について

コージェントラボは、人工知能を活用して人々の働き方や生活の未来を形作ることに取り組み、実生活の課題を解決する、使いやすく直感的なソリューションを提供します。時系列データを活用した予測、情報抽出、自然言語と音声処理、強化学習など、多分野にまたがる専門性を活用することで、最新の研究内容を実ビジネスに応用します。

会社名： 株式会社 Cogent Labs

代 表： 飯沼 純 / エリック ホワイトウェイ

所在地： 東京都渋谷区代官山町 20-23 TENOHA LAB

U R L : <https://cogent.co.jp/>

事 業：最先端の人工知能の研究・開発と関連ソリューションサービスの提供