

harBest

AI開発を高速化する
データ作成プラットフォーム
「harBest」



agenda



会社紹介



harBestとは？



harBestの“クラウドパートナー”



作成できるデータ事例



品質を保つ、harBestの仕組み

会社紹介

株式会社APTOは、データ作成を中心としたAI作成支援プラットフォームを提供しています。

Mission

高品質なデータを大量に、
素早く、安く作成できること

Vision

ビジネスインパクトのあるAIを
誰でも簡単に実現できること

会社名

株式会社APTO[アプト]

事業内容

アノテーションデータ
プラットフォーム提供事業

代表者名

高品良

所在地

東京都渋谷区神南1-5-14
三船ビル4F

TEL

03-6416-0523

Mail

info@apto.co.jp

HP

<https://apto.co.jp>

harBestとは？

データ収集・作成・整理をカンタン Web オーダー



harBest スマホアプリ



一般ワーカー



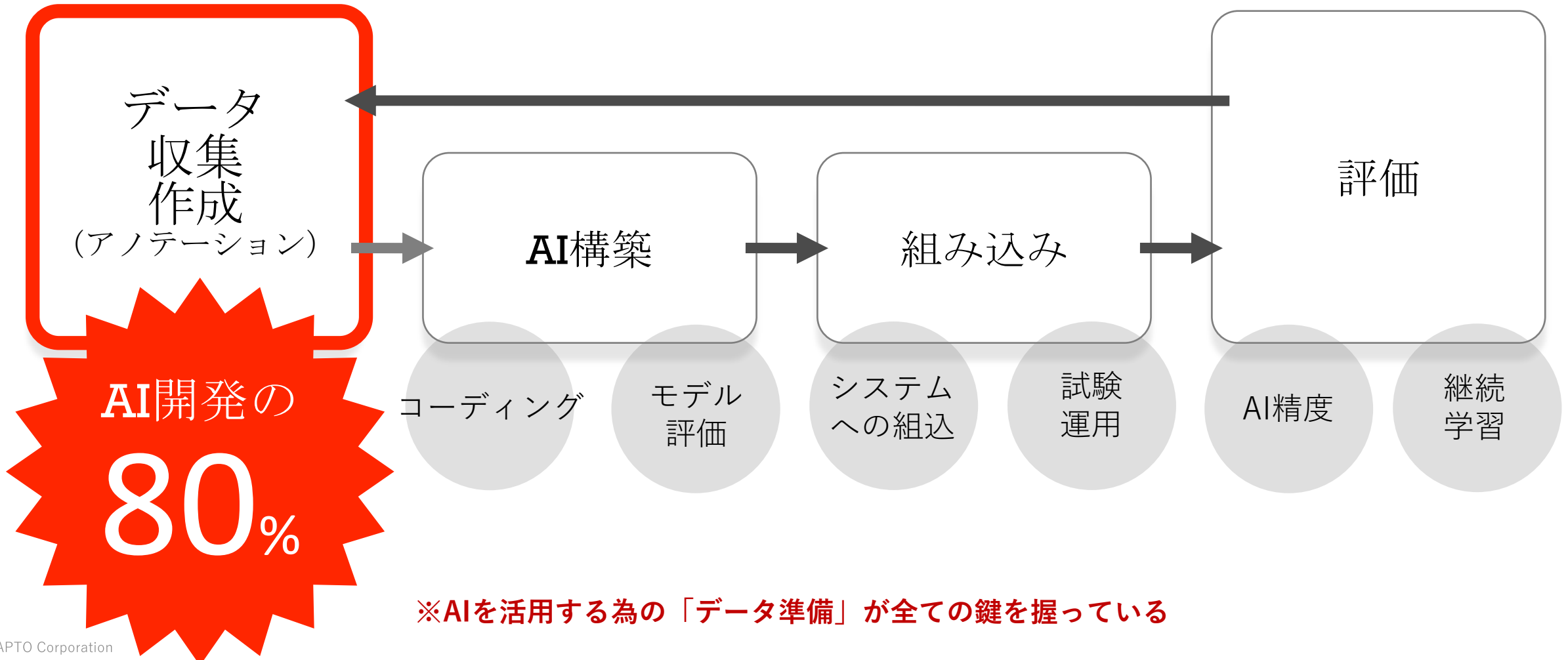
認定ワーカー



データ収集・作成・整理など出来る harBest スマホアプリを使って
面倒な作業をお客様に代わって実施します

例えば、AI開発のフローは以下になります

AI開発ではデータ収集・前処理に多くの労力が必要となることが多い



アノテーション（データの準備）って何？

あるデータに対して関連する情報（メタデータ）を注釈として付与することです

データに正解を付与し学習させることで、AIがデータを認識します。



What is COCO?



COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- ✓ Object segmentation
- ✓ Recognition in context
- ✓ Superpixel stuff segmentation
- ✓ 330K images (>200K labeled)
- ✓ 1.5 million object instances
- ✓ 80 object categories
- ✓ 91 stuff categories
- ✓ 5 captions per image
- ✓ 250,000 people with keypoints

Collaborators

Tsung-Yi Lin Google Brain
Genevieve Patterson MSR, Trash TV
Matteo R. Ronchi Caltech
Yin Cui Google
Michael Maire TTI-Chicago
Serge Belongie Cornell Tech
Lubomir Bourdev WaveOne, Inc.
Ross Girshick FAIR
James Hays Georgia Tech
Pietro Perona Caltech
Deva Ramanan CMU
Larry Zitnick FAIR
Piotr Dollár FAIR

Sponsors

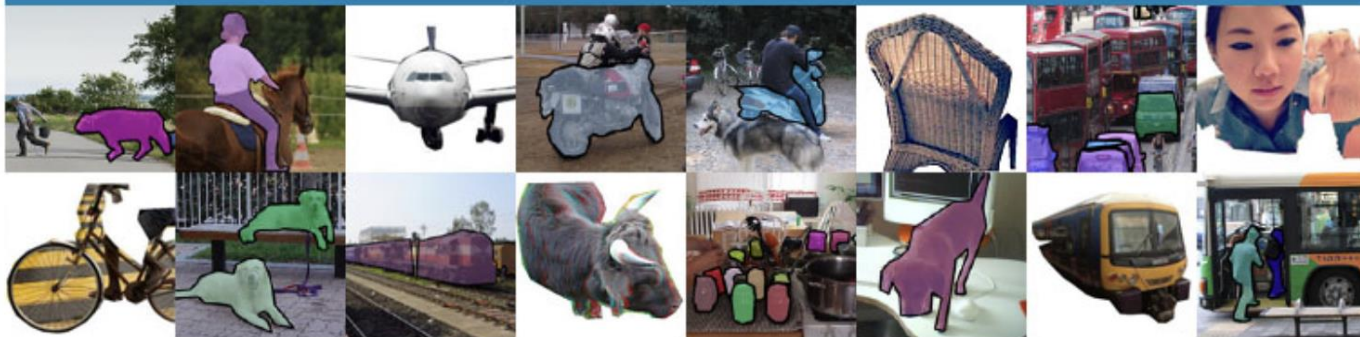


画像の塗り分けのデータ (Mscoco)

必要枚数

約20万枚

Dataset examples



精度を高めようとするほど
学習させるデータ量は増えていく

データの課題

データ活用をした企画をしたいけど…
こんなお悩みございませんか？

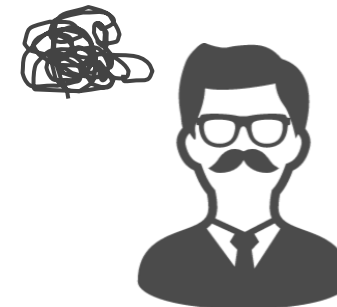
肝心なデータがない…



作業を
したいが人員が
足りない…



自社にデータあるけど
とても整理
しきれない…



課題

データ準備に時間と労力とお金がかかるからこそ、日本のAI開発は海外に後れをとっている

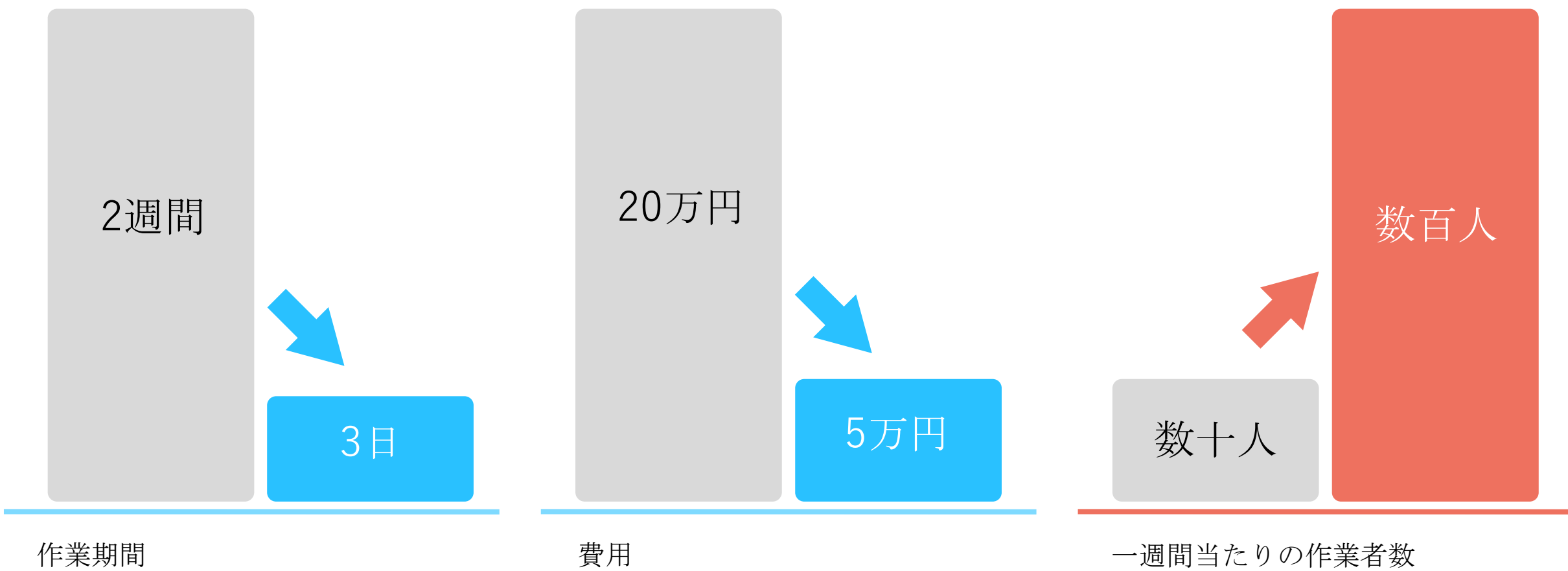
harBestで解決します

harBestでデータに関する様々なお悩みを解決します



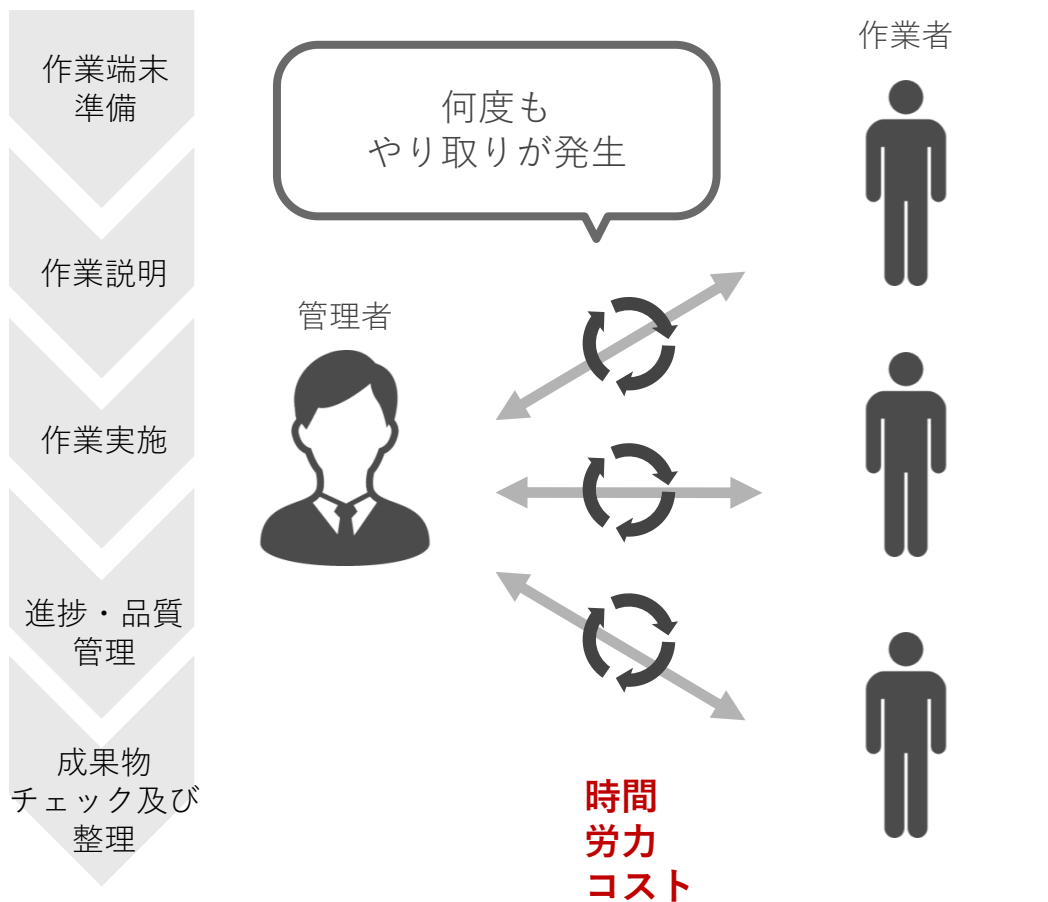
実績

画像分類作業15万件をharBestで実施した結果



管理プロセスを一元管理しコストを最適化

今までの作業イメージ

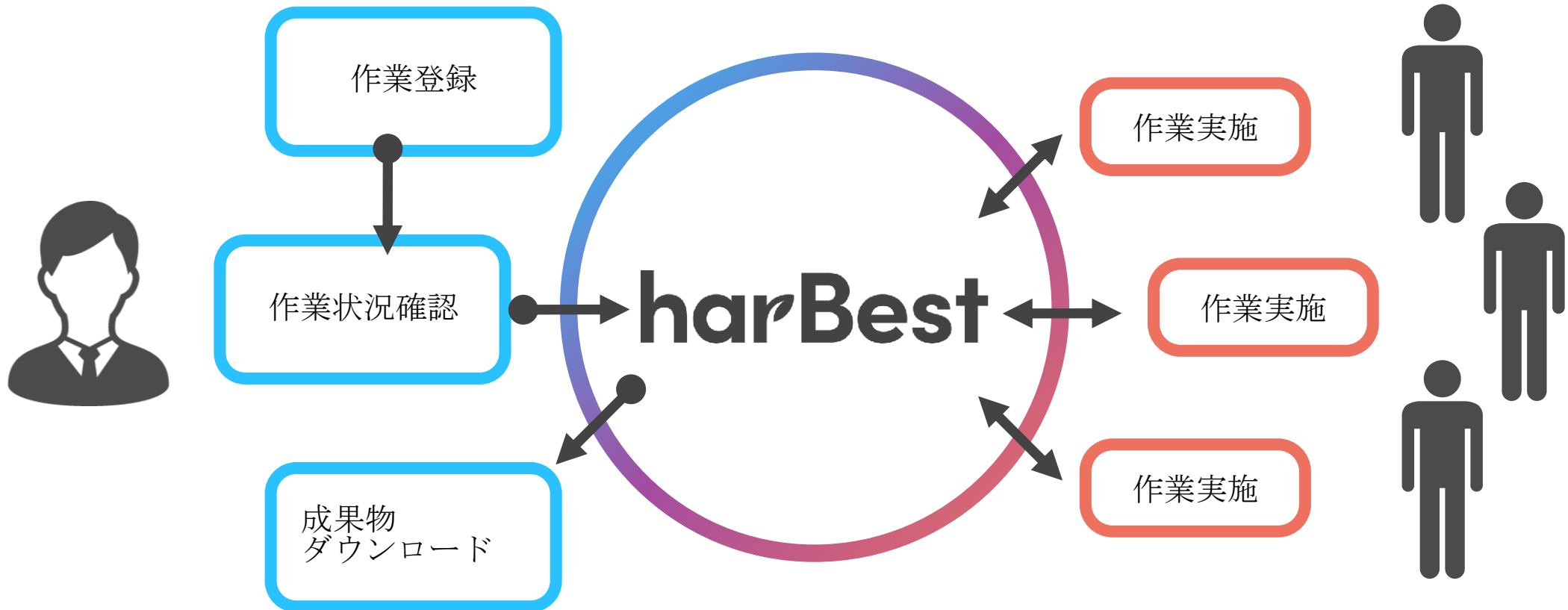


harBestによる作業イメージ



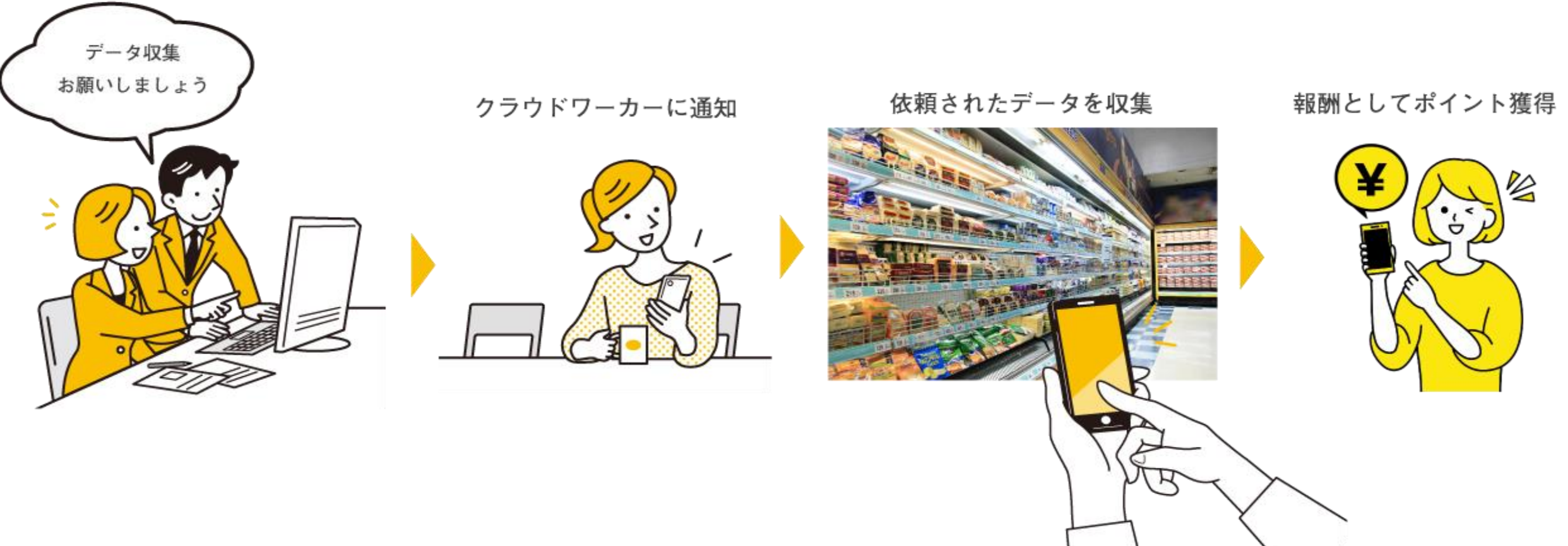
harBestはデータ作成の発注プラットフォーム

中間のBPO業者を介さず、直接パートナーに作業依頼ができるから早い・安い



発注と受託のフローイメージ

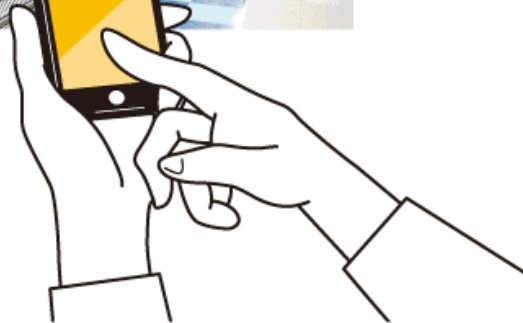
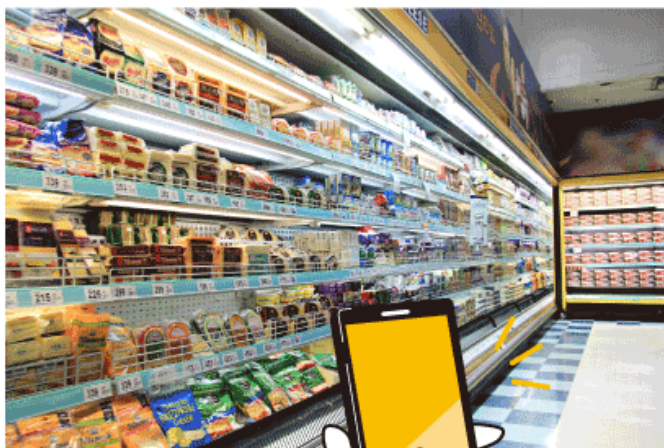
発注するとすぐにクラウドワーカーに通知され作業を開始します



何ができるの？

データ収集から加工までの流れを一元的に行えます。もちろん別々に利用することも可能です

データ収集

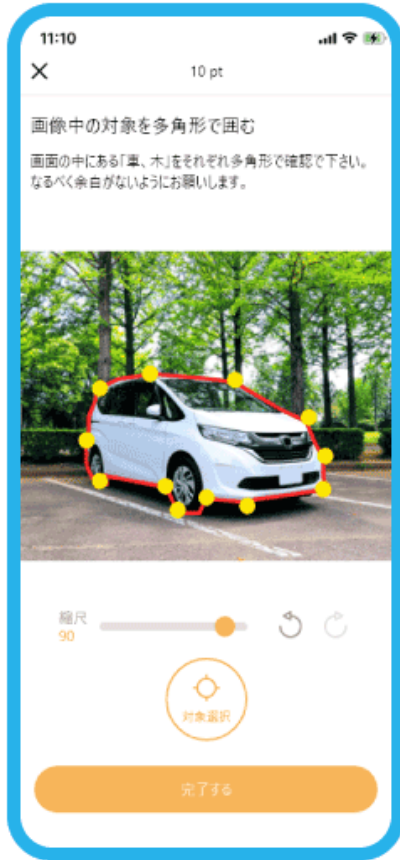


データ分析・加工



データ加工のイメージ

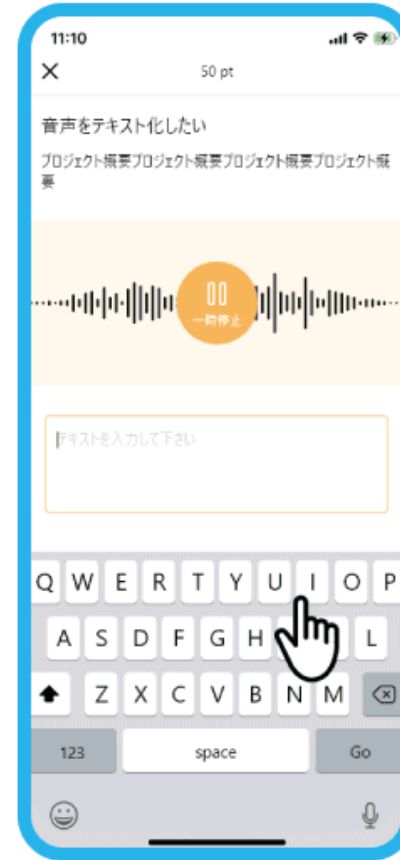
画像・文章・音声などのデータ作成を依頼出来ます



矩形・多角形・塗り潰し



文章分類・文脈ラベル



音声文字起こし

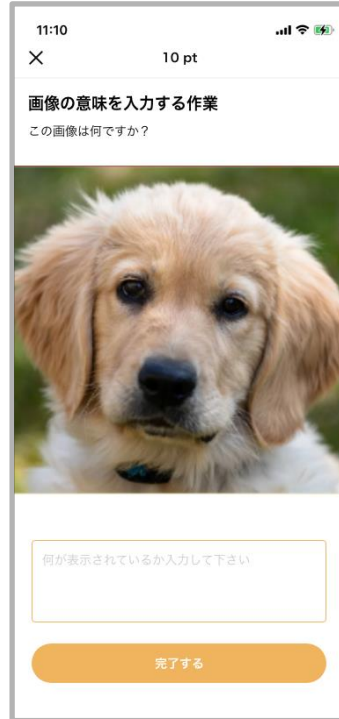
※一例

画像データを収集・作成・整理

選択肢



テキスト入力



塗り潰し



矩形



多角形



アップロード

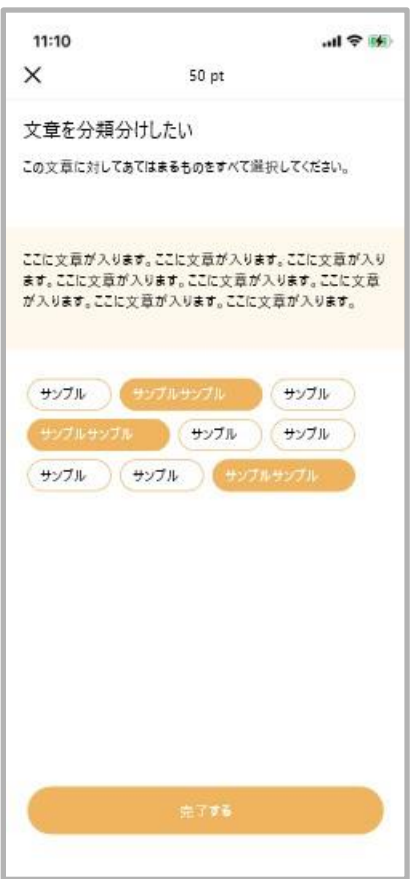


文章／テキストデータを作成・整理

1 : 1



1 : N



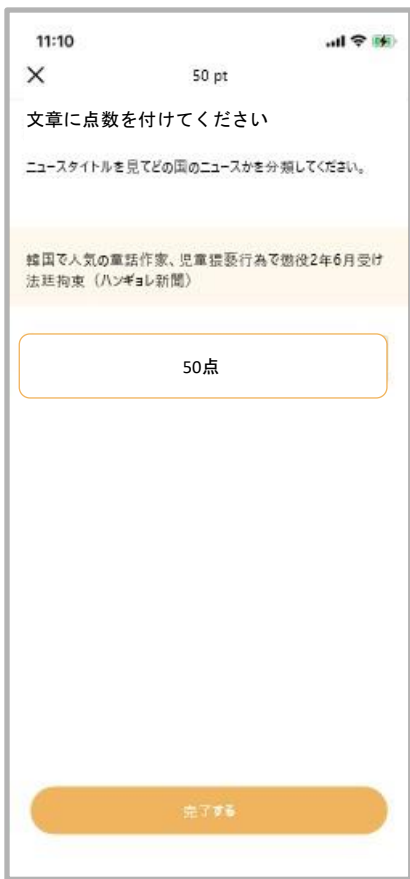
系列（固有表現）



翻訳・要約



文章点数評価



音声データを作成・整理

音声収録



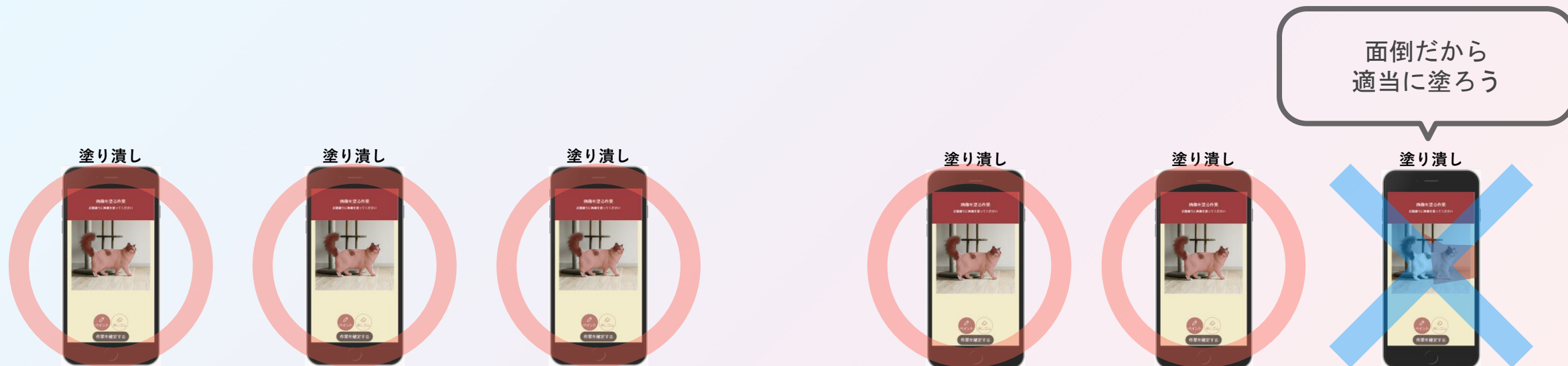
音声書き起こし



サービス紹介動画

<https://youtu.be/MkxERuHqwx4>

データチェックは必ず奇数人数で行い自動評価します



信頼度が高いデータと評価

信頼度が低いデータと評価

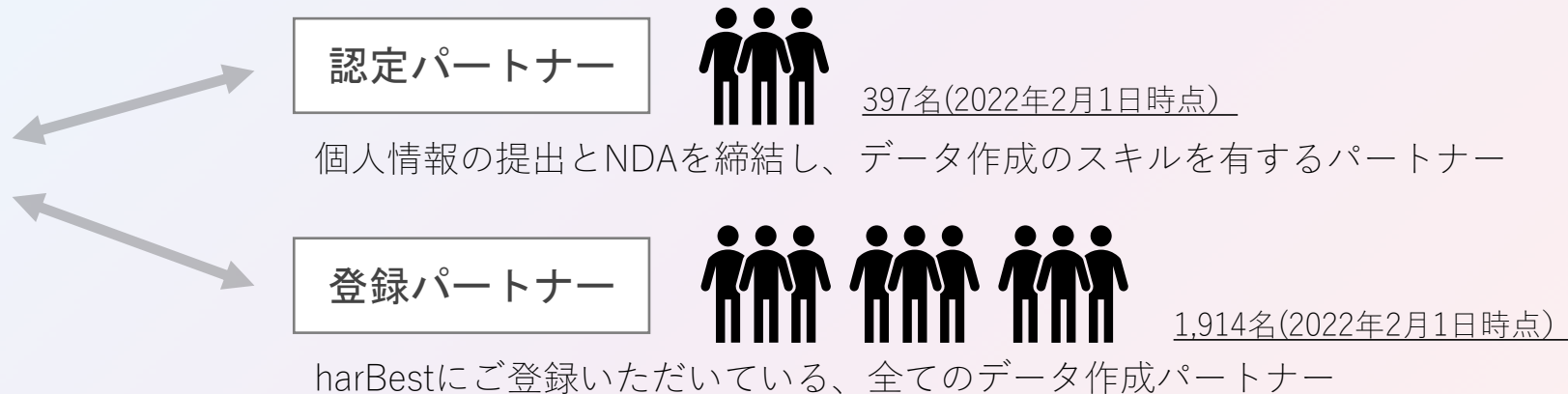
▶ データの信頼性を向上

※単一者への回答などで、チェック不要の場合は除きます

品質を上げる為の認定パートナー制度

パートナーのスキルに応じて作業案件を変え、品質を担保

harBest
作業者を評価



		登録パートナー	認定パートナー
作業難易度「低」	公開可	○	○
	公開不可	×	○
作業難易度「中」	公開可	○	○
	公開不可	×	○
作業難易度「高」	公開可	○	○
	公開不可	×	○

パートナーへの報酬スコアとトレーニング

① 作業品質をスコア化し、各パートナー事にランク付けしています

ランク					
レーティング	1~20	21~40	41~60	61~80	81~100
獲得pt倍率	0.5	0.8	1.0	1.5	2.0

品質スコアが高いほど、レーティングが高まりパートナーさんが受け取る報酬が高くなる仕組みになっています。

作業品質高い = 高レーティング = 高報酬
 作業品質低い = 低レーティング = 低報酬

② 認定パートナーには、定期的なデータ作成のトレーニングを実施



対象物を四角形で囲んでください。



対象物を多角形で囲んでください。

クライアントに

早く・安く・大量に

そして

正確に

AI開発・活用を加速させ
Creativeな事業推進を実現

パートナーに

年齢性別国籍問わず

好きな時間に

好きな場所で

すきま時間の仕事を換金し
Creativeな時間創造を実現