APTO

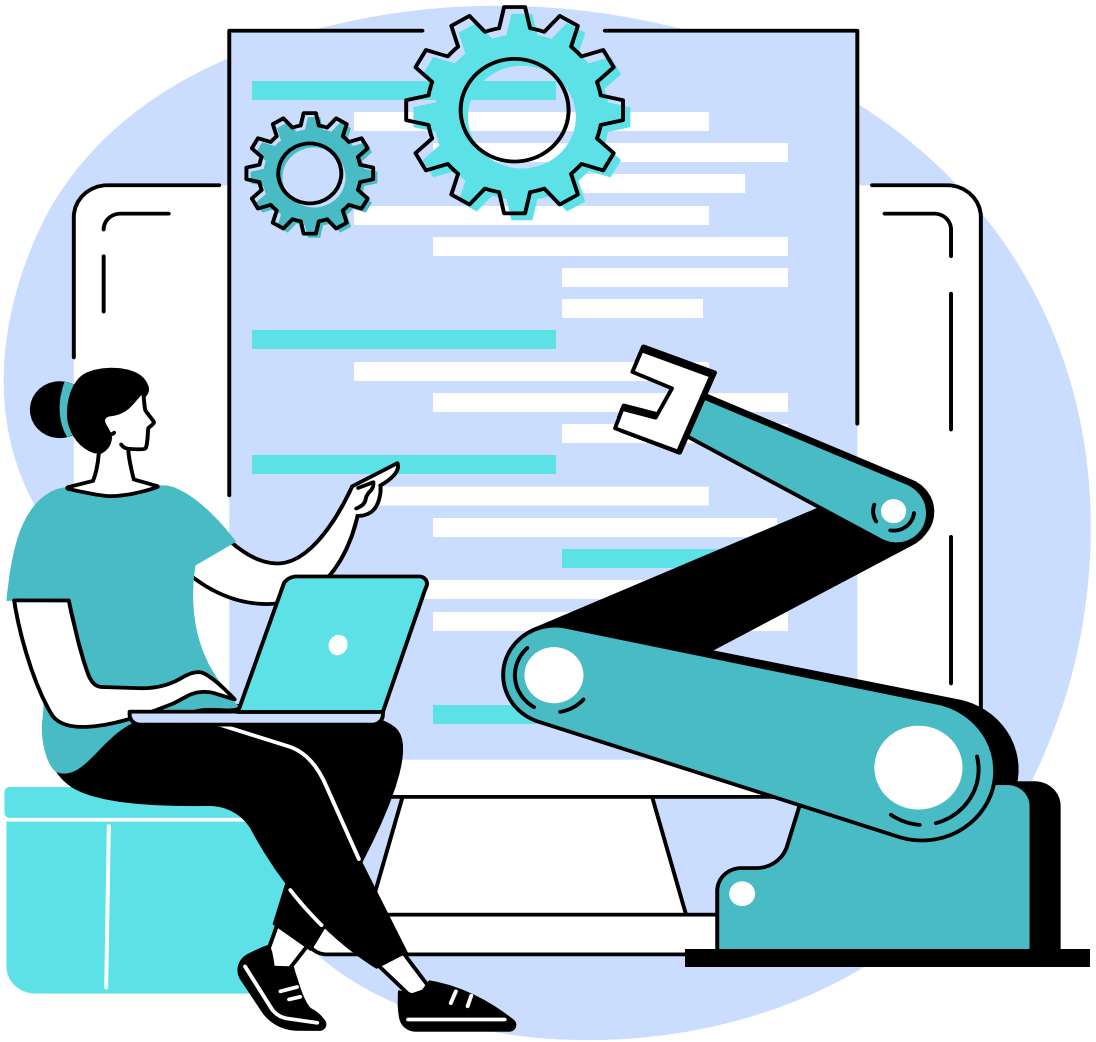
APTO お役立ち資料

3

3分で分かる

これからのAI開発に不可欠

Human-in-the-Loop
(ヒューマンインザループ)



目次

AI開発においてその精度を向上させるにはHITLアプローチが不可欠です。
本資料では以下の順序で HITL（Human-in-the-Loop）について解説させていただきます。

1. AI開発について …… 3~4
2. Human-in-the-Loopとは …… 5~7
3. Human-in-the-Loop 事例 …… 8
4. 弊社でご支援可能な“HITL”アプローチ …… 9
5. 弊社提供サービス“harBest”について …… 10

AI開発について

昨今、様々な業界でAIを活用する企業が増えております。
ただし業務で活用できるAIを開発するためには、AIの精度を高めるアプローチが必要です。

精度の高いAIを開発するためにはどんなアプローチが必要なのか？
本資料ではHITLアプローチによるAIの精度向上について解説させていただきます。

ピーマンの自動収穫機 AGRIST様
https://harbest.io/case/34/

AI議事録 オルツ様
https://harbest.io/case/162/

AI開発について

AIを開発するにあたり要素は多岐に渡ります。

重要な要素としては、データ・モデル・計算リソース等が挙げられます。

その中でも特にAndrew氏のMLOpsに関する研究によって

Data-Centric AI（データの質に焦点を当てたAI開発手法）が注目されているように

「教師データ（AIが学習をするためのデータ）の質」が非常に重要です。

では「どうやって高品質なデータを収集し、ラベリングすればいいのか？」

次ページから説明いたします。

AIシステム ＝ Model ＋ Data

Data-Centric → データに比重を置いたアプローチ方法

└データを変更しAIの精度を改善

Good Data(質の高いデータ)が必要

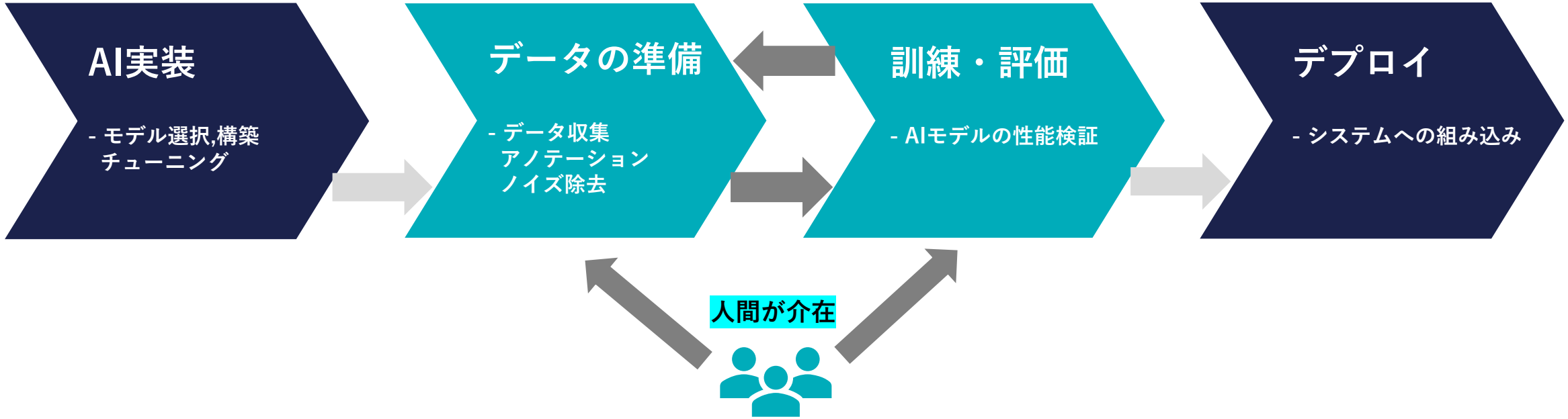
”Data-Centric AI について詳しくはこちら“

https://harbest.io/documents/738

Human-in-the-Loop とは

データの質を向上させ、精度の高いAIを開発するためには
“HITL (Human-in-the-Loop)”のアプローチが有効です。

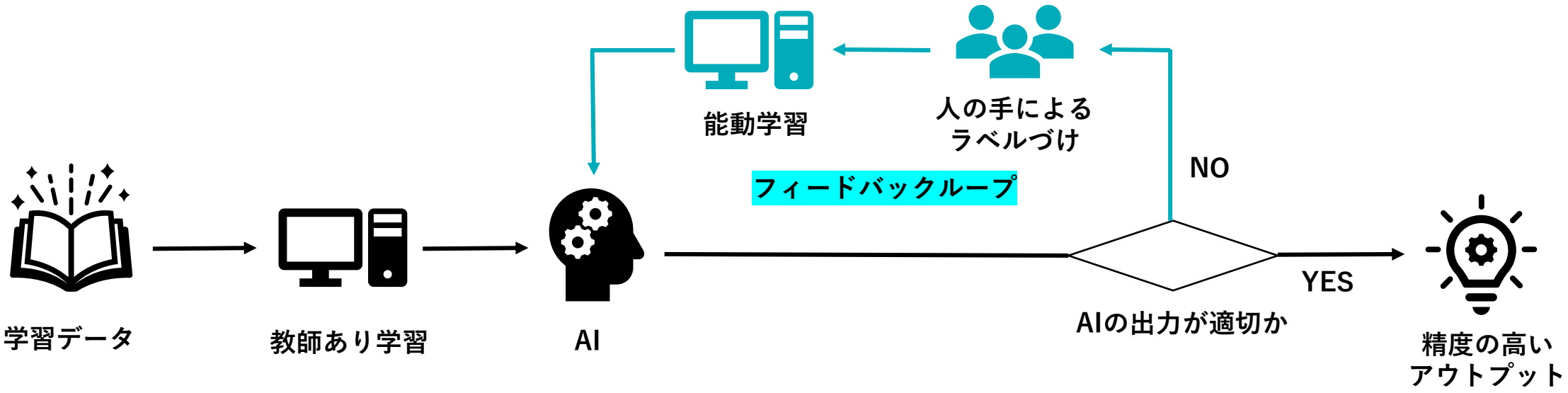
AI開発における HITL (Human-in-the-Loop) とは、
"データの準備 ～ モデルの訓練や評価” という一連の機械学習プロセスの中に、
人間との相互作用(フィードバック)が含まれることを指します。



Human-in-the-Loop とは

HITLの具体的な手法としては以下のような流れとなります。

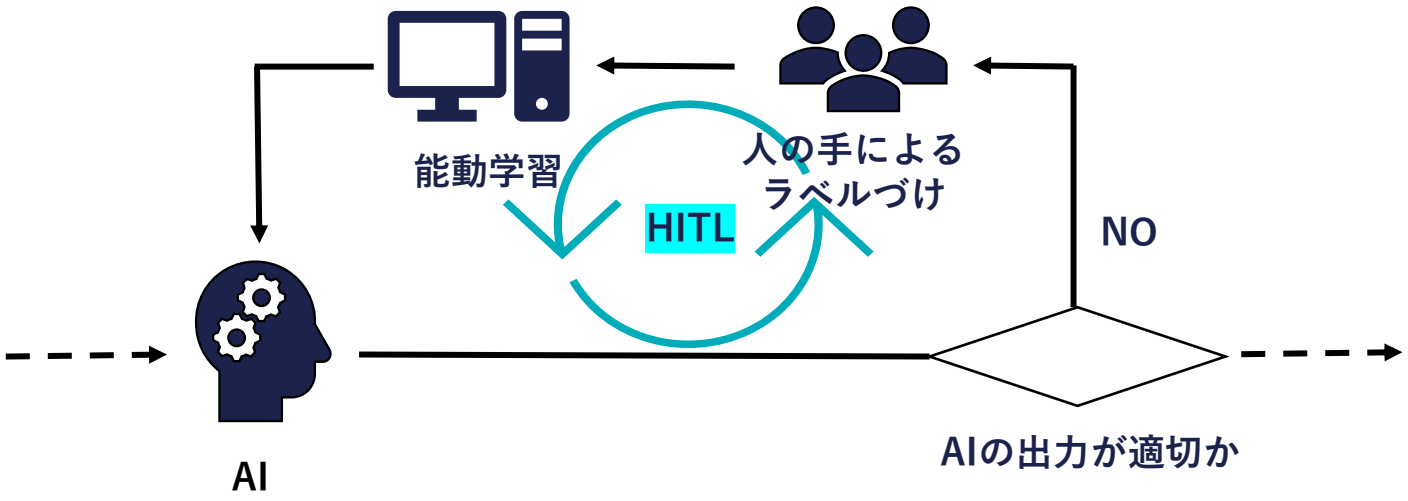
①大量の学習データを用い、教師あり学習で予測モデルを生成する
②学習と性能検証を繰り返し、十分な性能が得られれば本番環境へ移行する
③未知のデータに対する出力結果を観測し、問題なければそのまま成果とする
④AIからの出力に誤りがあれば、人間によるオペレーションで対処する
⑤AIが誤ったデータに対して人間がラベルづけの作業を行い、AIに再学習させる（能動学習）



Human-in-the-Loop とは

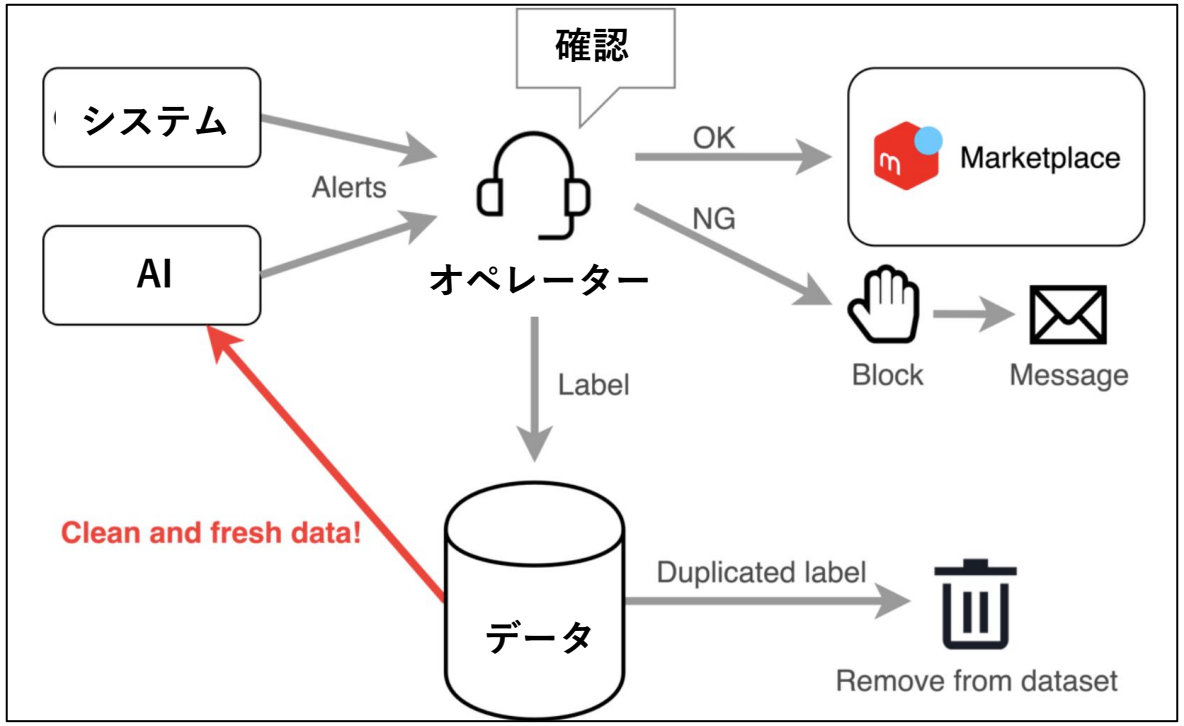
HITLで重要なのはAIから出力された結果を人間がチェックし、
そこから新たな学習データを得るという一連の流れです。

AIが間違えてしまったデータや未知のデータに対して、
人間がラベルづけ（アノテーション）を行い集中的に再学習を行う、
すなわち能動学習をループで行うことがHITLのアプローチとなります。



Human-in-the-Loop 事例

以下は国内最大級のフリマサイト ”メルカリ様“ のHITLの事例です。
違反出品の検知システムにおいて、システムによって検知された商品は
オペレーターが全て確認をし、必要に応じてラベルづけを行うことで
継続的にAIの精度を向上を実現されております。

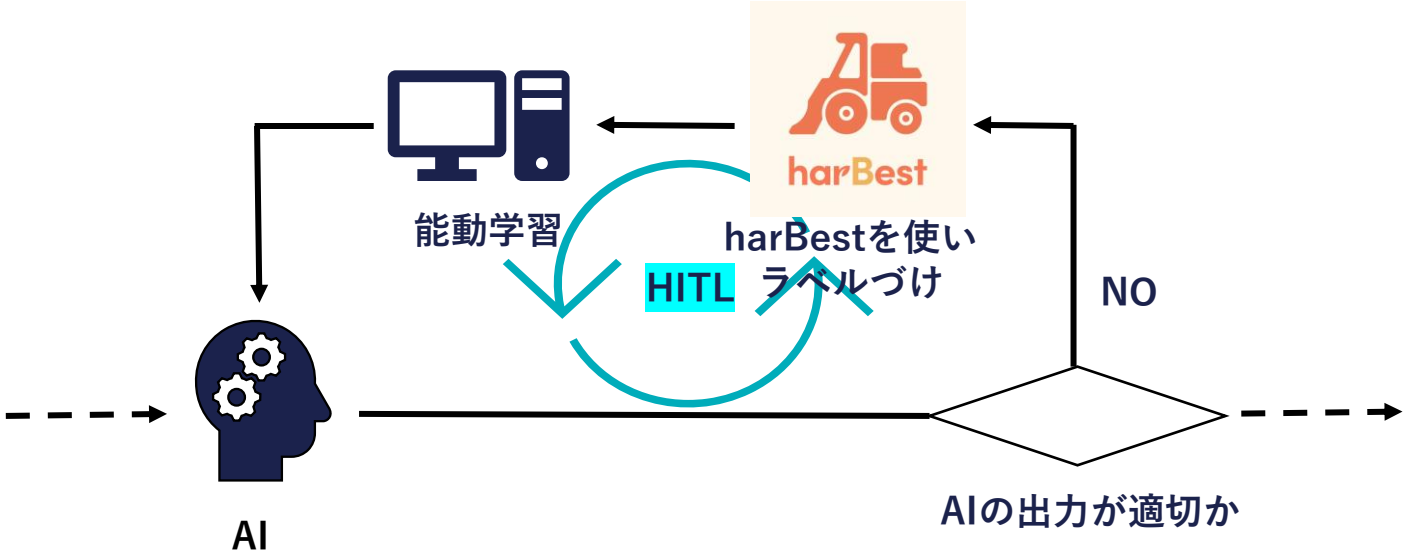


参照: https://engineering.mercari.com/blog/entry/2020-03-25-180000/

弊社でご支援可能な”HITL”アプローチ

弊社提供サービス“harBest”は、
プラットフォームを通して日本中のクラウドワーカーに
データ収集やアノテーション作業を簡単に依頼することが可能です。

AI開発の工程において、出力が適正か否かをAIエンジニアが判断し、
実際にラベルづけを行う際はharBestを用いることで
最速で質の高いラベルづけを実現できます。



弊社提供サービス“harBest”について

弊社提供サービス “harBest” は
データセントリックなAI開発をサポートするプラットフォームです。

高品質な教師データの作成を
早く・安く・大量に

早さ
harBestを通じてクラウドワーカーが
圧倒的な早さで大量のデータ作成を実現

質
独自の品質管理体制で高品質なデータを提供
データ作成の二度手間も防止

種類
画像・動画・音声・テキスト 等
様々なデータに形式に対応

株式会社オルツ様の事例

IT・インターネット

https://harbest.io/case/162

日本発、大規模言語処理モデルOriginal LLM「LHTM-2」の開発に挑む。

事例概要

自然言語の処理において、精度の高いデータ提供を ”harBest” を通して実現。

「パーソナルAI」を掲げているだけに、対話コミュニケーションと自然言語処理系のご依頼が多いが

「パラメーターをいかに増やすか」や「変な日本語とかそのあたりも含めて直してください」等の

"わがままを聞いてもらえるデータ屋" としてオルツ様のLLM開発をご支援。

会社概要

AI開発の8割を占めるデータ収集・アノテーションに
特化したサービスを提供しております

社名	株式会社APTO
代表	高品 良
事業内容	① アノテーションツール提供 ② アノテーション受託 / AI開発受託 ③ データセット販売 ④ ノーコード / ローコードプラットフォーム提供
受賞歴	Tech Chrunch japan 2021、Tribus、Gstartup
設立	2020年1月20日
拠点	本社：東京都渋谷区神南1-5-14 三船ビル403
Webサイト	https://harbest.io/

Best Venture 100
ベンチャー通信が「これから成長が期待される100社」をご紹介

2024年度ベストベンチャー100に選出

無料相談受付中

GANに関するお悩みをお持ちの方は
下記からお気軽にお問い合わせください。

webフォームから

click

無料お問い合わせ
AI開発相談はこちら

メールでのお問い合わせ

info@apto.co.jp

会社名・氏名・メールアドレス・電話番号を
ご記入の上、お問い合わせください

※当社のホームページからでも資料請求・お問い合わせが可能です。
https://harbest.io/