

2025年10月7日



AI時代の“データ整備基盤”とは？ Alteryx Oneが実現する AI Data Clearing Houseの全貌

アルテリックス・ジャパン合同会社

ソリューションエンジニアリング リードセールスエンジニア

新郷 美紀

alteryx

急速に進化・普及する生成AIソリューション

AI進化を支える3つの要因

- 膨大な学習データ

近年、AIモデルに使われるデータポイント数は飛躍的に増加し、精度向上に寄与

- 計算能力(コンピュート環境)の飛躍的向上

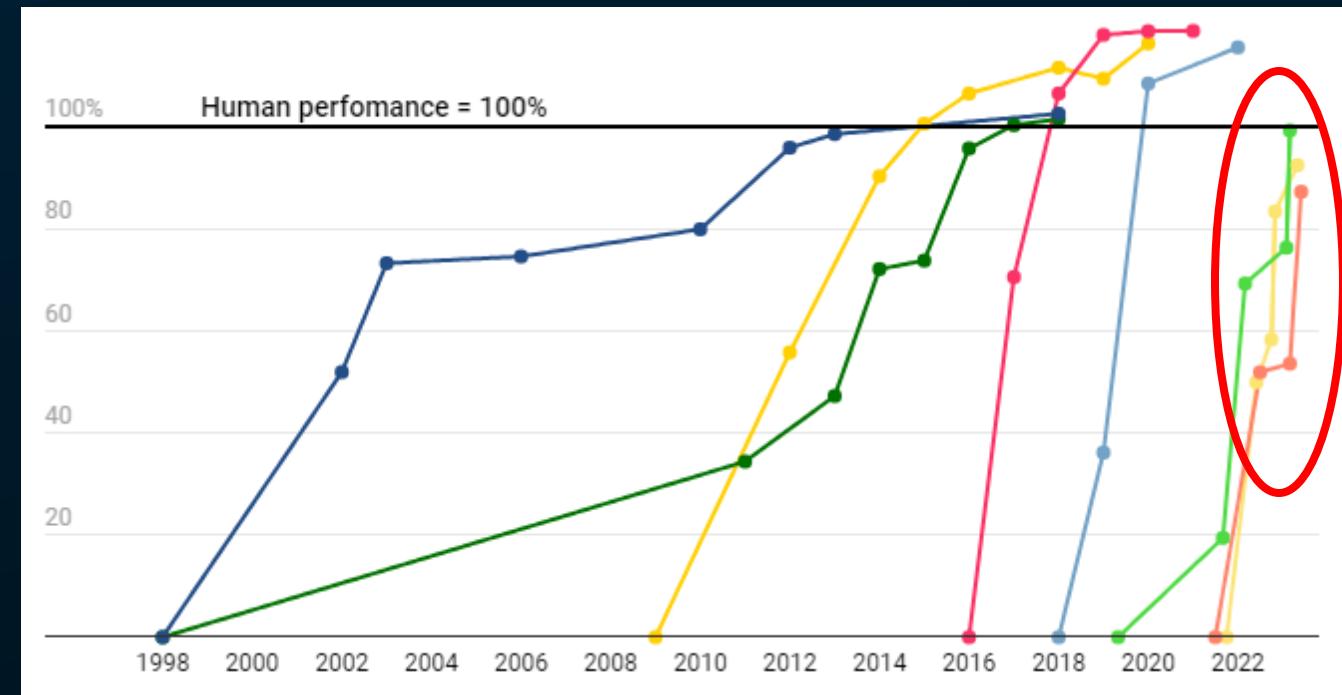
過去5年間で、AIトレーニングに使われる計算リソースは指数関数的に増加

- アルゴリズムの進化

より効率的に大規模データを扱える手法が登場し、学習スピードと精度が飛躍的に向上

AIはすでに多くのタスクで人間を上回っており、さらに新たなタスクで人間を超えるスピードも加速

- 手書き文字認識
 - 音声認識
 - 画像認識
 - 読解力
- 言語理解
 - 常識的な文章補完
 - 小学生レベルの数学
 - コード生成



最先端AIのベンチマーク性能（人間との比較） – 2023年8月時点
AIはすでに多くのタスクで人間の能力を上回っており、新たな分野で人間を超えるスピードも加速

Douwe Kielaほか:「Dynabench: Rethinking Benchmarking in NLP」
<https://aclanthology.org/2021.nacl-main.324.pdf> を元にAlteryxが作成

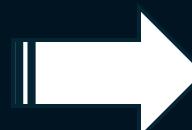
生成AIの期待と現実

なぜ生成AIの企業利用が拡大しているか？ (ユーザ視点)

- ① 技術的ナレッジに精通してなくても直感的に利用可能
- ② 多様かつ複雑な問い合わせに即応
- ③ 業務直結タスク（資料作成・プログラミング）に対応
- ④ 業務効率化・コスト削減・スピード向上
- ⑤ DX推進、競争力確保の観点から企業内利用が促進

商用利用が限定的な理由（課題）

- ① 回答結果の精度が保証できない
- ② 情報の鮮度・信頼性が不明
- ③ 企業特有のナレッジ反映が不十分
- ④ ガバナンス・監査対応の不足



結果としてPoC留まりが多発

Alteryxのソリューションアプローチ



生成AI時代に Alteryxがお客様にご提供する価値

生成AIの利用に関する客先課題とAlteryxによる解決案

お客様が抱える課題

課題	詳細
なぜこの結論に至ったかが、わからない？	<ul style="list-style-type: none">ロジックがブラックボックス化している (複雑なロジック実装は容易だが、思考プロセスが不明。妥当性を検証する術がない)
毎回同じ答えが得られるか不明？	<ul style="list-style-type: none">処理結果の再現性が担保されない (自動化プロセスで結論の正しさを簡単にチェックできない)
業務プロセスへの統合は？	<ul style="list-style-type: none">複雑なプロセス全体の再現性確保が難しい (生成AIを通常の処理ロジックに組み込む際の可視化が困難)

Alteryxによる解決策

課題	解決策
なぜこの結論に至ったかが、わからない？	<ul style="list-style-type: none">複数モデル比較（Azure OpenAI, Google Gemini等）でアウトプットを検証従来ロジックとの比較をワークフローで可視化し優位性を評価
毎回同じ答えが得られるか、不明？	<ul style="list-style-type: none">サンプルデータによる自動検証で再現性を担保データ変化を可視化し予期せぬ変化を発見
業務プロセスへの統合は？	<ul style="list-style-type: none">生成AIを通常の処理と同じインターフェースで統合完全なデータリネージで処理フローを可視化し正当性を担保

より具体的な課題と主な対処方法

課題	一般的な対処方法	Alteryxの特色
回答ロジックのブラックボックス化 (Reasoningの不透明さ)	<ul style="list-style-type: none"> ・チェーン・オブ・ソートやステップごとの検証を行う ・解答根拠を可視化する仕組みを導入 	<ul style="list-style-type: none"> ・データ処理過程の可視化：ワークフローで変換・加工・分析をすべて記録 ・Reasoning検証：回答の根拠となるデータをワークフローで裏取り
ハルシネーション（誤生成）対策	<ul style="list-style-type: none"> ・前処理／後処理で出力検証 ・複数モデルを比較（クロスチェック） ・確かめ算（数値・ロジック検証）を組み込み 	<ul style="list-style-type: none"> ・複数モデル比較：異なるLLMの回答を並列処理 ・前後比較ワークフロー：元データと生成結果を突合して整合性チェック
出典不明・データ正確性	<ul style="list-style-type: none"> ・RAG (Retrieval Augmented Generation) で参照データを明示 ・出典表示義務のある仕組みを導入 ・データガバナンスルールを適用 	<ul style="list-style-type: none"> ・出所トレーサビリティ：入力データ、加工プロセス、出力までを完全ログ化 ・ガバナンス基盤：監査ログ・アクセス権管理・データ利用履歴を一元管理
モデルごとの精度差と最適モデルの選定	ベンチマークテスト（MMLU, TruthfulQA等） 業務データでのA/Bテスト	<ul style="list-style-type: none"> ・評価指標を定量化：精度・速度・コストを比較し、最適モデルを選定 ・モデル比較：複数LLMの出力を並列で検証

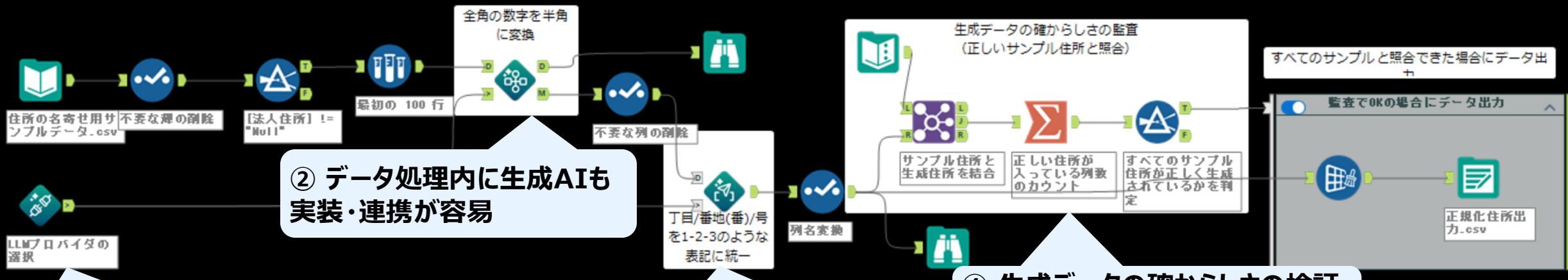


生成AIを安全・確実にエンタープライズ利用するためには
クリアリングハウス基盤が必須！

生成AIをAlteryxで利用する価値

日本の住所の名寄せ

- ・住所の数字に全角文字が含まれるものを半角数字に統一
- ・丁目、番地（番）、号などを排除し、1-2-3などの表記に統一
- ・名寄せ後の正しい住所サンプルと生成AIの出力結果を比較し精度



- ① 利用する生成AIプラットフォーム、モデルの確認・変更が容易
- ② データ加工処理の高度化、精度向上に生成AI機能の利用も容易
- ③ 生成AI機能の自動化、再現性の確認の仕組みの実装も容易
- ④ データガバナンス機能による生成AIの活用（監査条件の高度化・柔軟な対応）が促進可能

生成AIを利用した日本の住所の名寄せ結果（参考）

Alteryx Designer x64 - *日本の住所の名寄せ_v3.ymd

ツール、ヘルプ、およびリソースを検索…

★お気に入り ★推奨 ◀ 入出力 ●準備 ■結合 ●バース ■変換 ■インデータベース ■レポート ●ドキュメンテーション ●空間 ●機械学習 ●テキストマイニング ■コンピュータービジョン ◁ インターフェース ●データ調査 ■予測 ■AB テスト

エクスプローラー コメント ツールコンテナ

データクリンズPro (32) - 設定

Null データを削除

Null 行

Null 列

クリーニングする列を選択 検索

列	タイプ
元の住所	V_WString
変換後の住所	V_WString
*Unknown	Unknown

文字*を削除

先頭と末尾の空白

タブ、改行、重複した空白

すべての空白

HTMLタグ

不可視文字

文字

次を除く
例: ABC

数字

次を除く
例: 123

句読点と特殊文字

次を除く
例: .!?

文字列の列を置換

日本の住所の名寄せ

・住所の数字に全角文字が含まれるものを作角数字に統一
・丁目、番地（番）、号などを排除し、1-2-3などの表記に統一
・名寄せ後の正しい住所サンプルと生成AIの出力結果を比較し精度

結果 - データクリンズPro (32) - 出力

2/2 フィールド | 20 のレコードが表示されました

レコード	元の住所	変換後の住所
1	京都府京都市右京区西院月双町5	京都府京都市右京区西院月双町5
2	広島県東広島市西条吉行東一丁目4番14号	広島県東広島市西条吉行東1-4-14
3	愛知県名古屋市中村区名駅2丁目29番16号	愛知県名古屋市中村区名駅2-29-16
4	愛知県瀬戸市暁町3-143 暁西部地区工業団地内	愛知県瀬戸市暁町3-143 暁西部地区工業団地内
5	東京都世田谷区代田2-31-8	東京都世田谷区代田2-31-8
6	東京都中央区日本橋富沢町10-16 MYARK日本…	東京都中央区日本橋富沢町10-16 MYARK日本…
7	東京都中央区明石町8-1	東京都中央区明石町8-1
8	東京都八王子市越野24-1	東京都八王子市越野24-1
9	東京都千代田区内神田2丁目2-1錦倉河岸ビル	東京都千代田区内神田2-2-1 錦倉河岸ビル
10	東京都墨田区錦糸1丁目2番1号アルカセントラ…	東京都墨田区錦糸1-2-1 アルカセントラル5階
11	東京都文京区本郷1-28-24	東京都文京区本郷1-28-24
12	栃木県小山市城東2-32-17	栃木県小山市城東2-32-17
13	神奈川県横浜市神奈川区入江二丁目18番地	神奈川県横浜市神奈川区入江2-18
14	茨城県つくばみらい市中志筑2481-1	茨城県つくばみらい市中志筑2481-1
15	茨城県坂東市馬立1204-3	茨城県坂東市馬立1204-3
16	茨城県日立市大沼町3-19-3	茨城県日立市大沼町3-19-3

名寄せ前と名寄せ後のデータ

Alteryxの生成AI関連機能の具体例

Alteryxの生成AI接続機能（接続の柔軟性・安全性・効率化）

項目	概要
AIモデルの柔軟活用	複数モデルを簡単切替、要件に応じた最適選択
再現性と自動化	一度の設定で再利用・自動化、安定した運用を実現
効率化	接続設定をワークフロー内で再利用、作業削減
高精度な分析	クリーンデータ処理 + 生成AIで精度向上
セキュリティとガバナンス	接続認証・利用履歴を統制し、監査対応を強化

一度設定すればドロップダウンで複数の生成AIプロバイダ・モデルから容易に選択可能

Alteryxの生成AIによるデータ処理（処理の簡便化・高度化・再現性）

サンプル	従来手法	Alteryx + 生成AI
文書情報抽出 (契約書・請求書など)	正規表現・ルール設定を文書タイプごとに作成・修正	「請求日と金額を抽出」と指示するだけ、フォーマットが変わっても対応可能
データ列名の統一・整理	手動で列のマッピングルールを作成	「列名を統一」と指示するだけ
顧客レビュー要約・感情分析	キーワード集計・ネガポジ分析を手作業	レビューを要約し、感情をスコアリング

生成AIの活用事例：営業・マーケティング向けカスタマ360のチャットボット



まとめ

Alteryxは「柔軟性 × 高度化 × 自動化 × ガバナンス」で、生成AIにおいても安全かつ戦略的に活用できる基盤を提供

1. 柔軟な生成AIプラットフォーム選択

利用する生成AIの **モデル確認・変更が容易** で、ニーズに応じた最適な選択が可能

2. 高度かつ高精度なデータ加工

データ処理に生成AIを組み込み、**分析の精度向上と効率化を実現**

3. 自動化と再現性の確保

生成AI機能をワークフローに組み込み、**自動化・再現性ある分析環境** を実装可能

4. ガバナンスと監査対応

データガバナンス機能により、**生成AI活用の透明性・信頼性を担保**

アンケートにご協力お願いいたします



アンケートのご回答画面を弊社ブースにてご提示いただけますと
粗品を進呈いたします。

URL: <https://forms.office.com/r/mQjRNpijSp>

プライバシーポリシー :

<https://www.alteryx.com/ja/privacy-policy>

Thank You

alteryx