



生成AIサミット-Vol.2

生成AIの限界を突破する！ LLMを補完する、NLP技術の ハイブリッドアプローチとは？

株式会社 言語理解研究所
CRO 事業開発責任者 芳賀 諭史

プロフィール



芳賀 諭史 (Haga Satoshi)

株式会社 言語理解研究所 (ILU)

CRO 事業開発責任者

Sansan株式会社にて、インボイス管理サービス「Bill One」
含む複数の事業開発やアライアンスを担当。
データベース事業の子会社代表を経て、
現在は株式会社言語理解研究所へ出向し、
CRO 事業開発責任者として、
AIのビジネス利用に関する企業支援を推進している。

株式会社 言語理解研究所 (ILU)



言語理解研究所

- 自然言語処理（NLP）専門のシステム開発会社
- 徳島大学の研究室が前身で40年の歴史
- 国内最大規模の言語資産（辞書）を保有
- 豊富な知見とノウハウを持つプロフェッショナル集団
- 大手企業とのAIシステム共同開発の実績多数

43 0x61 0x70 0x74 0x75 0x72 0x69 0x6E 0x67 0x20

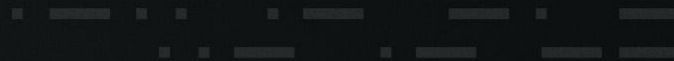
61 0x6E 0x64 0x20 0x73 0x70 0x69 0x6E 0x6E 0x69

6E 0x67 0x20 0x74 0x68 0x65 0x20 0x69 0x6E 0x74

65 0x6E 0x74 0x20 0x62 0x65 0x66 0x6F 0x72 0x65

20 0x61 0x6E 0x64 0x20 0x61 0x66 0x74 0x65 0x72

20 0x6C 0x61 0x6E 0x67 0x75 0x61 0x67 0x65



Mission

言葉の価値を最大限に引き出し、
顧客の想像を超え続ける

1 0 8

9 7

1 1 0

1 0 3

1 1 7

9 7

1 0 3

1 0 1

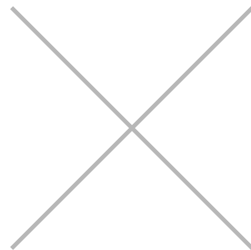


positive

Negative

Neutral

0x6C 0x61 0x6E 0x67 0x75 0x61 0x67 0x65



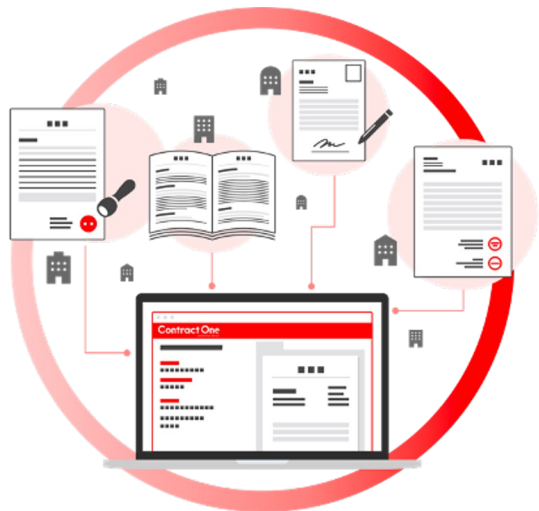
2023年06月にSansanグループイン

SansanのプラットフォームへAI技術を提供

契約データベースから、収益を最大化する

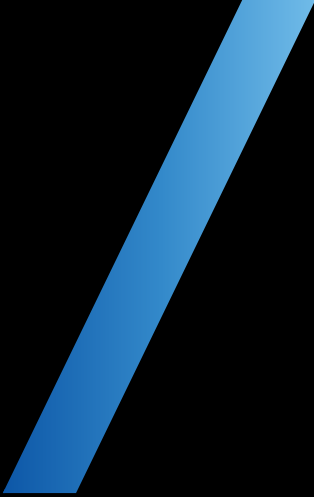
Contract One

powered by Sansan



契約書条項分け・原契約情報抽出API

Sansan株式会社が提供する「Contract One」は、あらゆる契約書を正確にデータ化し、全社で契約情報を活用できるようにする、契約データベースです。契約書条項分け・原契約情報抽出APIは、文字認識と情報抽出技術を活用した契約書の条項分けと、契約の親子関係を自動判別する「契約ツリー機能」を実行するために必要となる原契約情報を抽出する機能です。



**生成AIの限界を突破する！
LLMを補完する、NLP技術の
ハイブリッドアプローチとは？**

今回のセッションの対象は

社内業務利用

- メール文作成
- アイディア出し
- 議事録作成
- 文章校正・要約
- コーディング補助

ビジネス利用

- 自社サービスへのAI組み込み
- プロダクトのバックエンドシステム
- 業務オペレーションへのAI活用



生成AIの実用化を阻む、LLMの弱点

間違ふ、でも原因分らない、直せない
説明責任（アカウンタビリティ）が果たせない

アウトプットの
正確性や安定性に欠ける

ブラックボックスであるため
過程や原因が不明

間違いの修正や
更新が困難

正答率94.1%でも 実用化に至らない

- 生成AI（LLM）の正確性は80%～90%が限界
- 90%を超えても実用化の期待に届かない
- 残り10%は例外や個別事情の集積
- トレードオフのケースもあり改善が困難

The screenshot shows the official website of Mito City (三豊市). The header includes the city logo and navigation links for various services like '暮らし・手続き' (Living/Procedures), '子育て・教育' (Child-rearing/Education), '健康・医療・福祉' (Health/Medical/Welfare), and '行政情報' (Administrative Information). A breadcrumb trail indicates the current location: '現在の位置 > ホーム > 各課からのご案内 > 市民環境部 > 環境衛生課 > ごみ・環境・ペット > 「チャットGPTを利用したごみ出し案内」 本格導入について'. The main content area features a title '「チャットGPTを利用したごみ出し案内」 本格導入について' and a date '更新日：2023年10月23日'. The text describes a pilot project where ChatGPT was used to generate waste disposal notices, with a 94.1% accuracy rate achieved. It notes that while the accuracy was high, the remaining 5.9% of errors were due to specific circumstances, leading the city to decide on a full-scale implementation of ChatGPT for this purpose.

三豊市 One MITOYO

暮らし・手続き | 子育て・教育 | 健康・医療・福祉 | 行政情報

現在の位置 > ホーム > 各課からのご案内 > 市民環境部 > 環境衛生課 > ごみ・環境・ペット > 「チャットGPTを利用したごみ出し案内」 本格導入について

「チャットGPTを利用したごみ出し案内」 本格導入について

更新日：2023年10月23日

市では、6月から東京大学大学院工学系研究科松尾研究室と協力して、ChatGPTを用いた市民向けの「ごみ出し案内」の実証実験を進めてきました。

本格導入の条件として正答率99%を設定していましたが、10/23～11/30に行った2回目の実証実験の結果、正答率が94.1%に留まったことなどから、このたび、本市における「ごみ出し案内」業務には、ChatGPTを活用しないと決断しました。

今後は、今回得た知見を活かし、市役所全体として、市民サービスの向上のためにChatGPTが活用できるか否かを探ってまいります。

引用：三豊市ホームページ「チャットGPTを利用したごみ出し案内」本格導入について
<https://www.city.mitoyo.lg.jp/kakuka/shiminkankyoku/eisei/2/chatGPT/index.html>

NIKKEI

ILU独自の翻訳エンジンにより、高精度な記事翻訳を実現

日英自動翻訳サービスは、GoogleやDeepLなどの既存の翻訳エンジンが苦手とする、固有名詞の翻訳、省略された主語の補完、情報漏洩のリスク回避といった課題を、ILUの自然言語処理技術で改善し、高精度なAIサービスの提供を実現しています。



翻訳のタスクにおいて、LLMが苦手なこと

長文や複雑な構成における
文脈理解

新聞記事の性質の
汲み取り

専門分野や業界特有の
用語や言い回し

日本語の特徴
(主語が無い文章が多い等)

固有名詞
(企業名・人名)

機密情報の扱い

正確性や安定性に欠ける、生成AIへの向き合い方

致し方ないと受け入れる

期待値調整を行う

UI/UXを工夫する

最後は人の力でなんとかする

改善対応を行う

プロンプトエンジニアリング

RAG

ファインチューニング

生成AIを擬人化すると



生成AIくん

めちゃくちゃ物知りで
頭いいけど、ちょっと適当で
空気がよめない新入社員

そんな生成AIくんへの改善対応

プロンプト エンジニアリング



懇切丁寧に
事細かく指示を出す

RAG



参考情報を渡して
見ながら作業させる

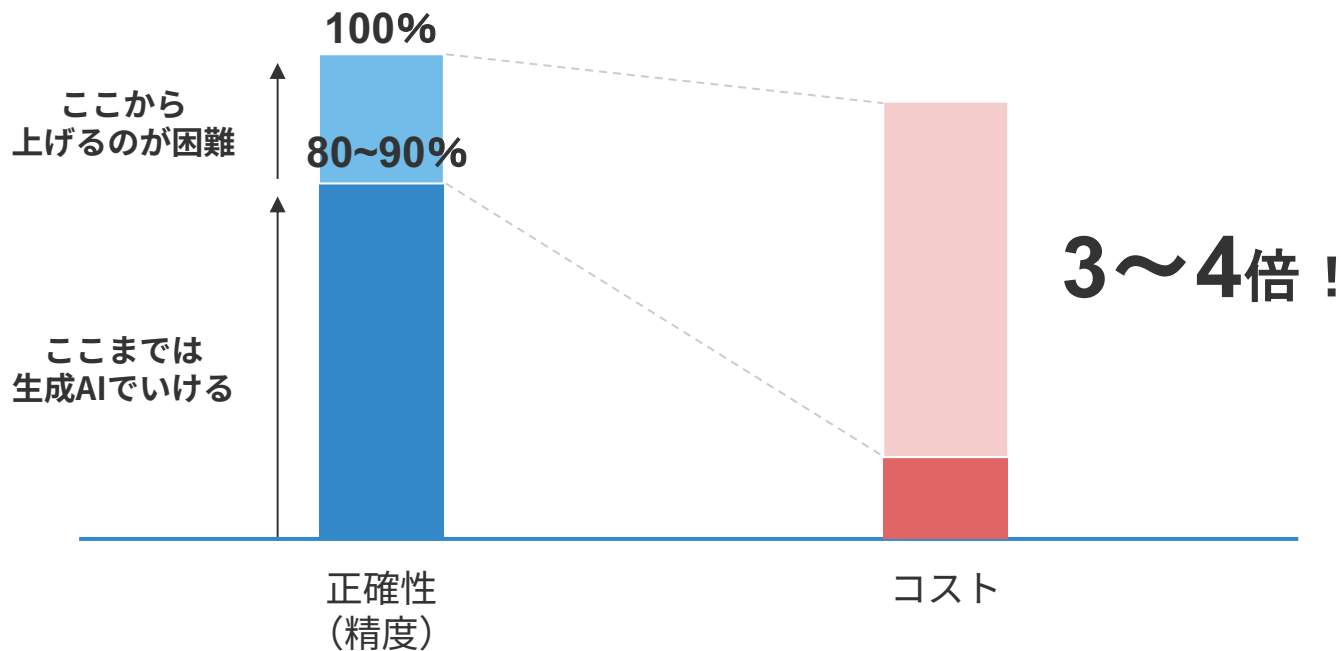
ファイン チューニング



必要な知識を
再教育させる

正確性を100%に近づけるには、多大なコストがかかる

正確性を上げるためにかかるコスト



生成AIの限界をどのように突破するか？

生成AIの実用化において重要な4要素

1

正確性が高いこと

2

説明できること

3

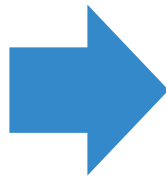
改善できること

4

対応コストが適切であること

生成AI（LLM）の弱点を補完するアプローチが有効

生成AIくん
ひとりにまかせず



先輩たちと
チームで対応！



チームNLP

NLP技術のハイブリッドアプローチ



生成AI
(LLM)

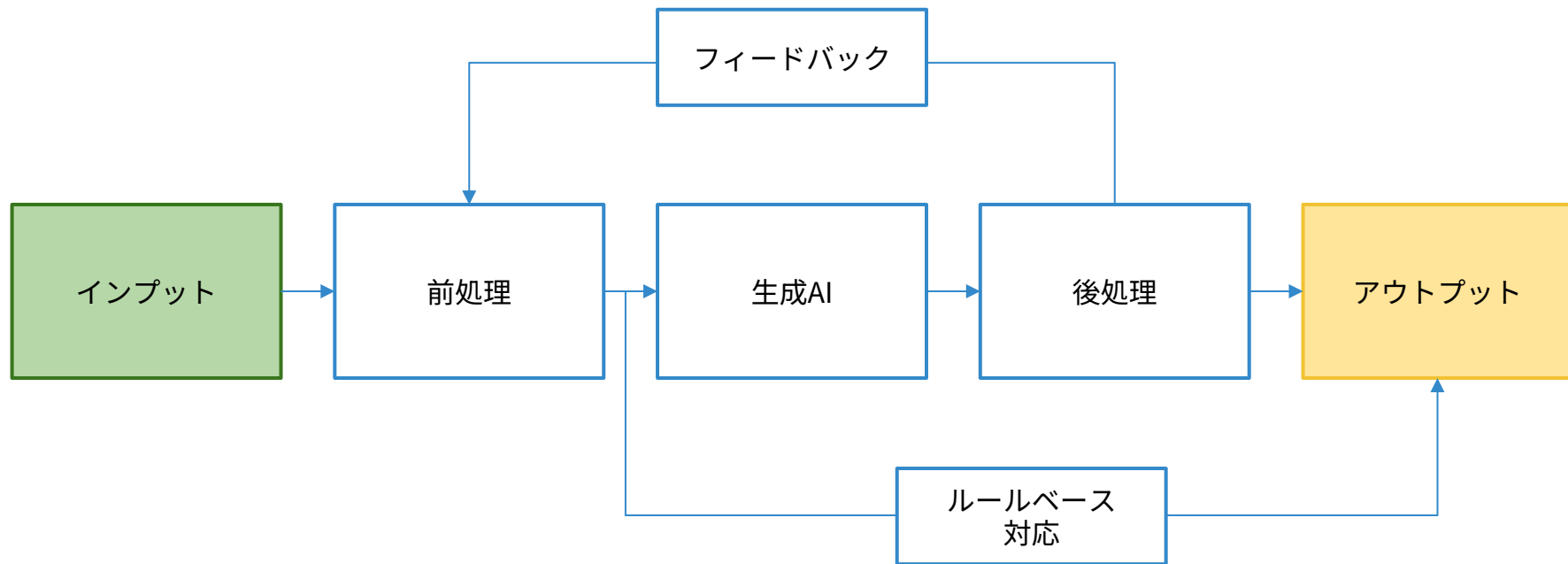
The diagram illustrates a hybrid NLP approach. It features a blue circle on the left containing the text '生成AI (LLM)' and a blue square on the right containing the text 'ルールベースAI (知識モデル)'. These two shapes are connected by a large white 'X' symbol, representing the combination of the two technologies.

ルールベースAI
(知識モデル)

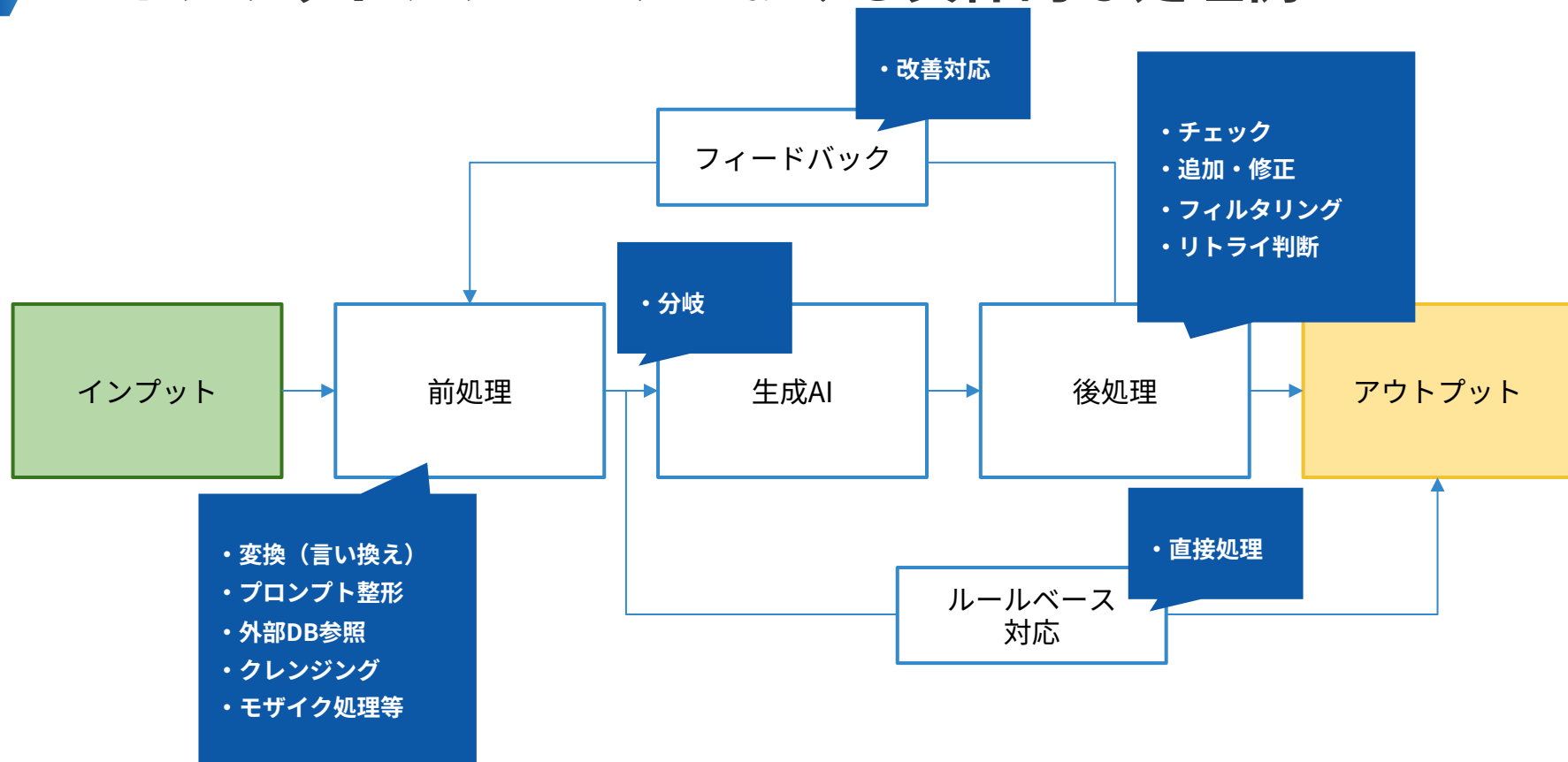
生成AIとルールベースAIの比較

生成AI（LLM）	ルールベースAI（知識モデル）
広範囲のドメインやタスクに汎用的に対応可能	特定のドメインやタスクに最適化しやすい
透明性・解釈性が低い（ブラックボックス）	透明性・解釈性が高い（明確なルール）
柔軟だがやや安定性に欠ける	硬直だが安定的
一般に高精度だが、タスクによっては精度低下	一定の条件下であれば高精度
追加学習のコストが高い	更新やチューニングが比較的容易

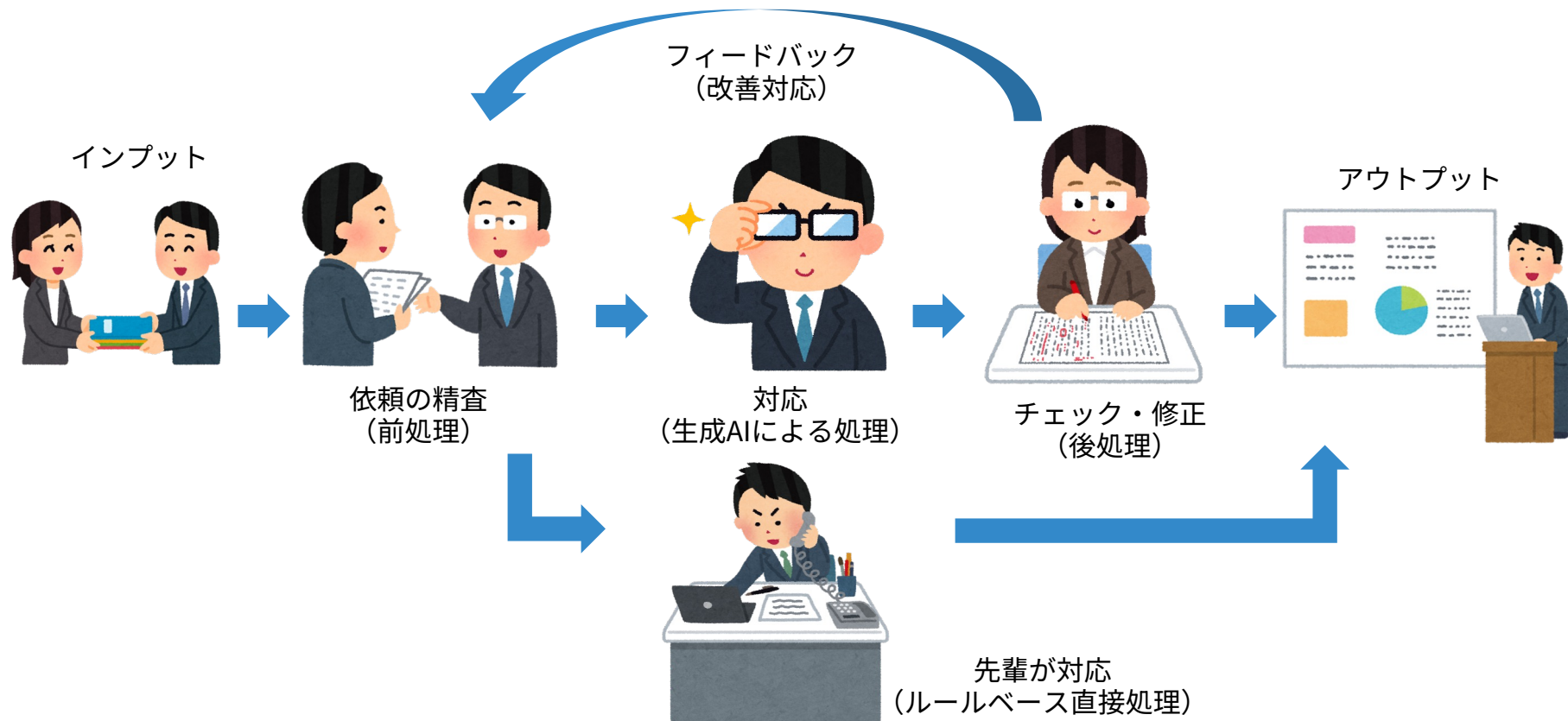
ハイブリッドアプローチのプロセス構成例



ハイブリッドアプローチにおける具体的な処理例



また例えると

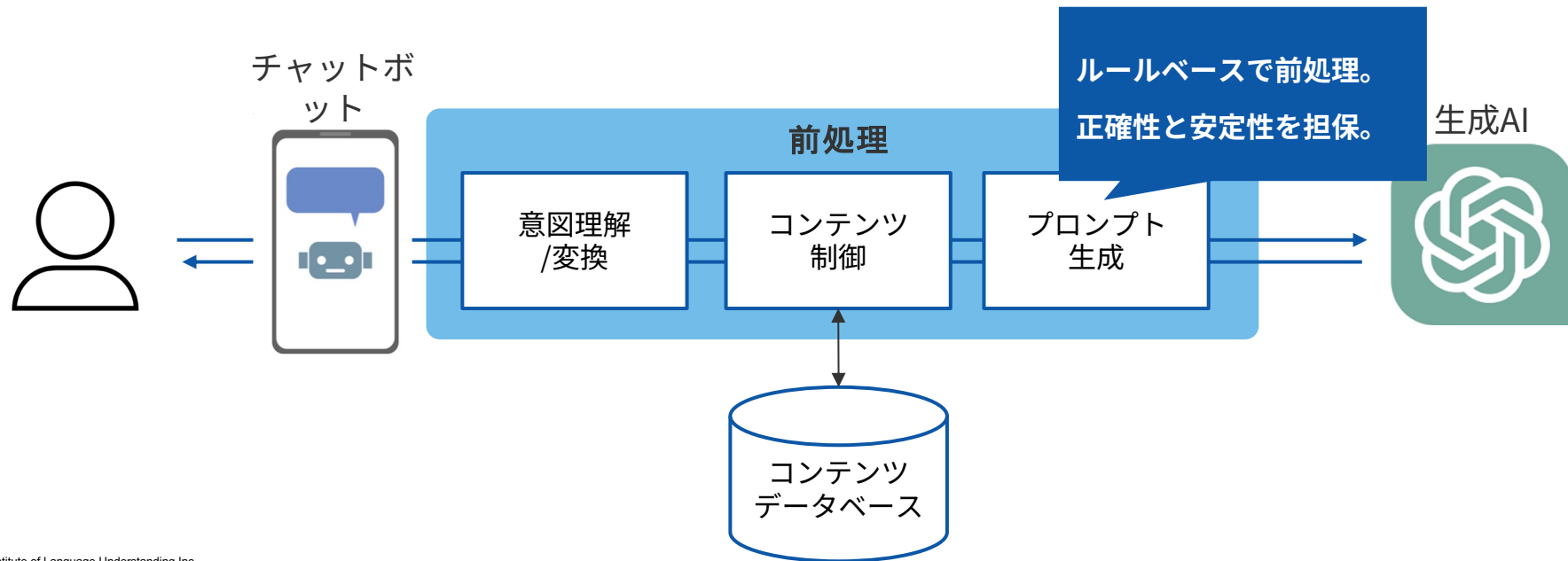


ハイブリッドアプローチのメリット

- ルールベースAIを適切に組み合わせることで、正確性や安定性が向上
- 実用化において、LLMの弱点となる説明と改善が可能になる
- すべて生成AIで処理するよりも、大抵のケースで高速かつ低コスト
- 疎結合であるため、生成AIの新しいモデルが出ても切り替えが可能
- 生成AI・LLMの進化に追従でき、常に最適な構成を維持できる

事例：サービス紹介チャットボット

- ユーザからの問い合わせに対し、自社サービスを適切に紹介するチャットボットサービス。
- 扱うサービスの特性上、正確性と安定性を重視。
- サービスポリシーに照らして、リードしたい方向性やNG対応が必須設定。



まとめ

- 実用化のためには、**正確であること**、**説明できること**、**改善できること**が大事
- **生成AIとルールベースAIを組み合わせること**が最適解
- 重要なのは、どの部分を生成AIにまかせて、
どの部分をルールベースにするかの**判断・見極め**



言語理解研究所

言語理解研究所（ILU）の提供サービス

01

コンサルティング
/PoCサービス

02

カスタマイズ型
AIエンジン受託開発

03

ソリューション提供
ILUアノテーション

言語理解研究所（ILU）の「クラフトAI」

高精度かつ改善可能な「実用化向け」のカスタマイズ型AIシステム

- AI・NLP技術のプロフェッショナル集団が構築
- 生成AIとルールベースを組み合わせる、ハイブリッドアプローチ
- NLP技術と国内最大規模の辞書を用いて、人手でルールを構築
- 言葉や単語、フレーズの概念化技術でルール構築のコストを大幅に圧縮
- 各企業の業務に即した、個別カスタマイズと最適化を実施

positive
Negative
Neutral

国内最大規模の言語資産（辞書）



終わりに

- 生成AIはビジネスに大きな変革をもたらす技術であることは確実
- 生成AI導入は所詮ツールで手段
- しかし、こと生成AIに関しては「手段」が「目的」化しても良い
- 大事なのは、乗り遅れず向き合い続けること



言語理解研究所