

生産現場における「生成 AI×オンプレ」の未来

株式会社エムニ



アジェンダ

1

会社概要・自己紹介

2

なぜ生成AI活用が進まないのか？

3

オンプレ開発の課題とコツ

4

生産現場における生成AIのキラーユースケース

5

まとめ

会社概要・自己紹介

会社概要

□ 株式会社エムニの概要

所在地	東京都千代田区東神田1丁目11-5 石田ビル東神田3階
設立年月日	2023年10月31日
代表者	下野 祐太
HP	https://emuniinc.jp

□ 株式会社エムニの特徴



京都大学発

&

M
MACC
松尾研發



技術顧問

□ 松尾先生による弊社のご紹介

エムニは、製造業における生成AIの活用に特化したスタートアップです。

代表の下野君は松尾研究所で3年間、製造業向けAI社会実装に深く携わってきた経験を持ち、その豊富な知識と実績が大きな強みとなっています。

また共同創業者の後藤君をはじめとして、有名ITベンチャーでのプロジェクト経験を持つメンバーが在籍しており、技術力と創造性に溢れたチームを形成しています。「AIで働く環境を幸せに、世界にワクワクを」というミッションのもと、エムニが製造業に革新をもたらし、未来の産業をリードすることを強く期待しています。



代表取締役 CEO

下野 祐太

SHIMONO YUTA

- ・京都大学大学院エネルギー科学研究科卒
- ・東京大学松尾研究所にて
プロジェクトマネージャーに従事
- ・PFN・DeNA・Recruit等の
複数企業のプロジェクトに取り組む
- ・松尾研起業クエスト1期生
- ・主に、AI開発・Biz面を担当

メディア実績



製造業×生成AI インタビュー

「製造業×生成AI インタビュー」の連載記事一覧です。



いいね 1 0 シェアする ポスト BI ブックマーク 0



製造業×生成AI インタビュー：

生産現場が注目する「生成AI×オンプレ」の未来 何が導入障壁になり得るか

現在、生産現場における生成AI活用では、オンプレミス環境下でのAIモデル運用に注目が集まっている。ただ、クラウド経由で生成AIサービスを利用する場合と異なり、オンプレミス環境ではさまざまな制約条件がある。これらを乗り越え、どのように実装を進めていくべきか。エムニの下野祐太氏に話を聞いた。

[池谷翼, MONOist] (2024年10月7日)



製造業×生成AI インタビュー：

産業界で使い倒せる日本語LLMを リコーが自前のモデル開発に取り組み始めたワケ

世界中で大規模言語モデル（LLM）の開発成果が次々に発表されている。国内でも日本語対応のLLM開発に取り組む企業が複数いるが、その1社がリコーだ。130億パラメーターの日本英両言語対応LLMを作った同社に、開発の狙いや今後の戦略を聞いた。

[池谷翼, MONOist] (2024年5月23日)



製造業×生成AI インタビュー：

“超高速”でアイデアを具体化 DNPが「生成AIラボ」で目指す共創活動

生成AIに関心を示し、自社サービスや業務への導入を検討する製造業は多い。だが、生成AIで何かできるのか、どういったサービスを作れるのかをイメージし、具体化していく仕組みが社内にあるだろうか。そのための仕掛けとして、東京都内に生成AIの共創施設をオープンしたDNPの和田剛氏と大竹宏之氏に話を聞いた。

[池谷翼, MONOist] (2024年1月11日)



製造業×生成AI インタビュー：

技術文書を100文字要約、アサヒビールがR&Dプロセスに生成AIを導入した2つの狙い

アサヒビールは2023年10月、R&D部門を主な対象として、マイクロソフトの「Azure OpenAI Service」を活用した社内情報検索システムのPOCを開始した。社内情報検索システム導入の理由や、R&Dプロセスにおける生成AI活用の期待感を同社担当者に聞いた。

[池谷翼, MONOist] (2023年11月8日)

製造業×生成AIのトレンドは“現場”にあり——松尾研究室のAIスタートアップが語る

2024年08月28日 19時09分 公開

[島田拓, ITmedia]

印刷

見る

Share

5

1

AIポータルメディアを運営するアイスマイラー（東京都渋谷区）は8月28日、AI製品に関する合同記者発表会を開催した。発表会では、東京大学でAIを研究している松尾・岩澤研究室発のAIスタートアップ・エムニ（東京都文京区）などが登壇。顧客である製造業界の生成AIトレンドを紹介した。

【訂正履歴：2024年8月29日午後8時 記事掲載当初「松尾研究所発」と記載していましたが、正しくは「松尾・岩澤研究室発」でした。お詫びして訂正します】



エムニ代表取締役の下野祐太さん

日本経済新聞

特集  AI PCとはなんだ!?

人類の可能性広げる AI PCで仕事が変わる!?



“製造業”が秘める AI活用 大ポテンシャル

松尾研時代から注目 製造業の世界
熟練工の暗黙知をAIで言語化
“爆速AI開発”ができるワケ

エムニ CEO
下野祐太

SPONSORED



メディア実績



A portrait of Yuta Shimono, Founder & CEO of Emune. He is a young man with dark hair and glasses, wearing a dark suit and white shirt. He is looking slightly to his right. To his right is a circular graphic containing his name and title, and below it is a circular logo for the Forbes 30 Under 30 list.

**YUTA
SHIMONO**

Founder & CEO



JAPAN 2025

弊社事業

オーダーメイドAI開発を土台としつつAI受託ノウハウを蓄積し、ソリューション及び製造業に特化したマルチプロダクトを展開

	戦略	研究	設計	調達	生技	保全	物流	保守
プロダクト	AI特許ロケット	-	-	-	AIインタビュー	AIインタビュー	-	AIチャットボット
ソリューション	AI特許検索	AIリサーチトレンド発掘	回路図面の情報抽出	-	報告書作成エージェント	-	-	-
AI受託ノウハウ	特許分類	物性値抽出	車両図解析	-	計画最適化	異音検知	-	クレーム自動対応
	侵害調査	転記自動化	図面解析	-	ヒヤリハット防止	異常検知	-	

土台となるオーダーメイドAI開発

取引先企業

製造業のエンタープライズ企業を中心とした取引実績

※ 取引先(一部)



DENSO
Crafting the Core

LIXIL

TOKYO GAS

DAICEL

MEIDEN
Quality connecting the next

TAIYO YUDEN

不二製

株式会社 三井ハイテック

 空気をかえよう
Easte

 より、そう、ちから。
東北電力

RESONAC

RYOBI
リョービ株式会社

 **dynabook**



四国電力グループ
四国計測

 **大東建託**

 **Narita Airport**


東京都

 **NTT東日本**

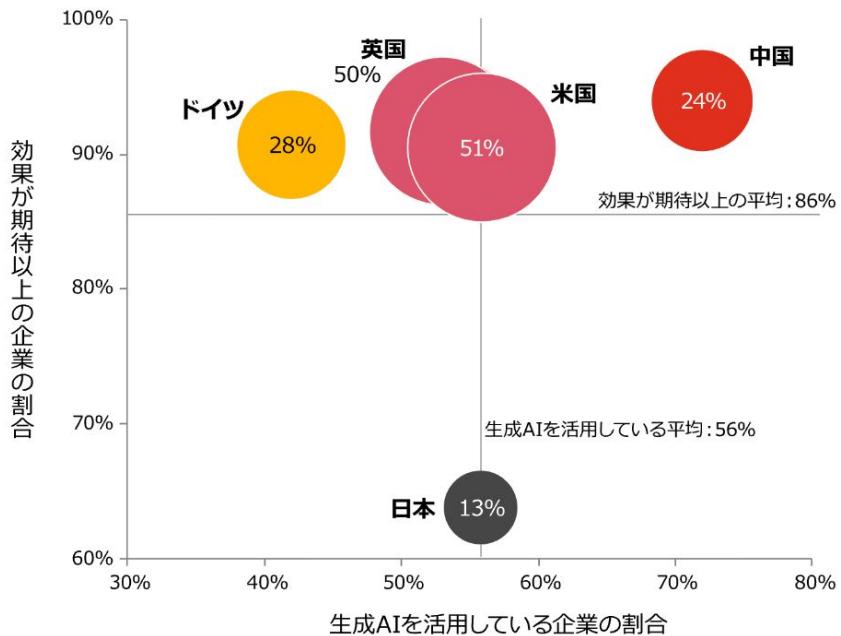
 **STNet**

 **Sound One**

なぜ生成AI活用が進まないのか？

日本の現状 | 他国との比較

他国と比較して生成AI自体利用はしているものの、期待値を大きく下回る



- 米国
 - 中国よりも生成AIを活用している企業の割合が低いものの、生成AIの効果が期待以上の割合が同程度に高い
 - 「期待を上回る」割合は中国の2倍以上
- 中国
 - 効果が期待以上の企業の割合、生成AIを活用している企業の割合が最も高く、生成AI導入を積極的に推進
 - 「期待を上回る」割合は米・英・独に見劣り
- ドイツ
 - 他国と比べて生成AIを活用している企業の割合は低いものの、生成AIの効果が期待以上の企業の割合が米・英・中に見劣りしない水準
 - 「期待を上回る」割合は中国より高く、米・英に次ぐ水準
- 日本
 - 生成AIを活用している企業の割合は平均的な水準にあるものの、生成AIの効果が期待以上の企業の割合は、他国と比べて低い
 - 「期待を上回る」割合は、米・英の1/4、独・中の半分程度
- 効果が期待以上の企業の割合：
生成AIを「既に活用している」を選択した企業の内、生成AIの効果が「期待を大きく上回っている」「期待通りの効果があった」と回答した割合
- 生成AIを活用している企業の割合：
生成AIの推進度合いとして「社外向けの生成AI活用サービスを提供している」「社内業務等で生成AIを活用している」と回答した割合
- 円の大きさ：
生成AIを「既に活用している」を選択した企業の内、生成AIの効果が「期待を大きく上回っている」と回答した割合

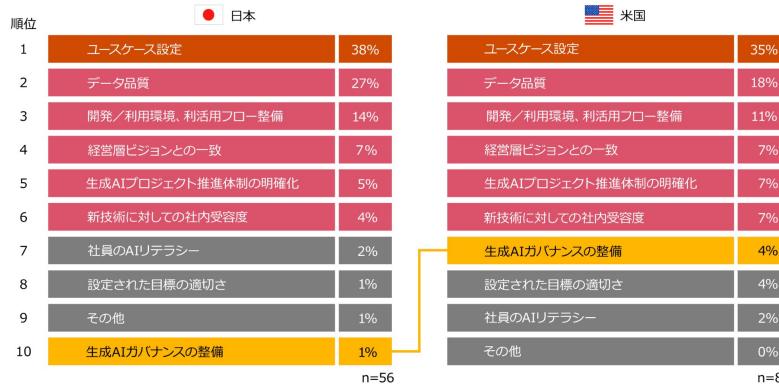
なぜ期待値を超えることができないのか？

ユースケースの設定が最重要にも関わらず、現場ベースでの取り組みができていない

□ 成果を出せた要因

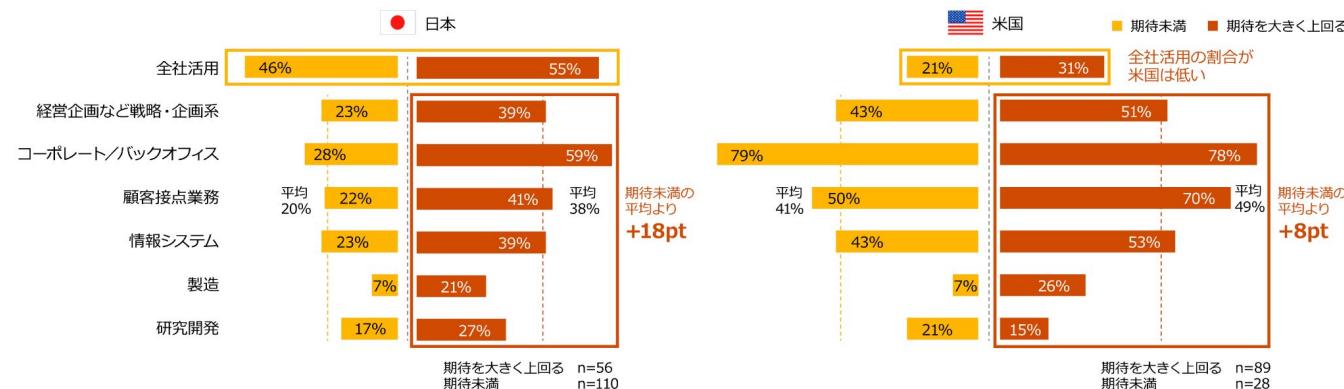
日本・米国共に
最も重要なのは
ユースケース設定
その後データ品質や
環境・利活用フローの
整理が続く

<https://pwc.com/jp/ja/knowledge/thoughtleadership/generative-ai-survey2024-us-comparison.html>



□ 生成AIの使い所

日本では**全社活用**が多いが、
米国では**経営や
コーポレートといった
特定業務に特化した
具体的なユースケース推進が
先行している**

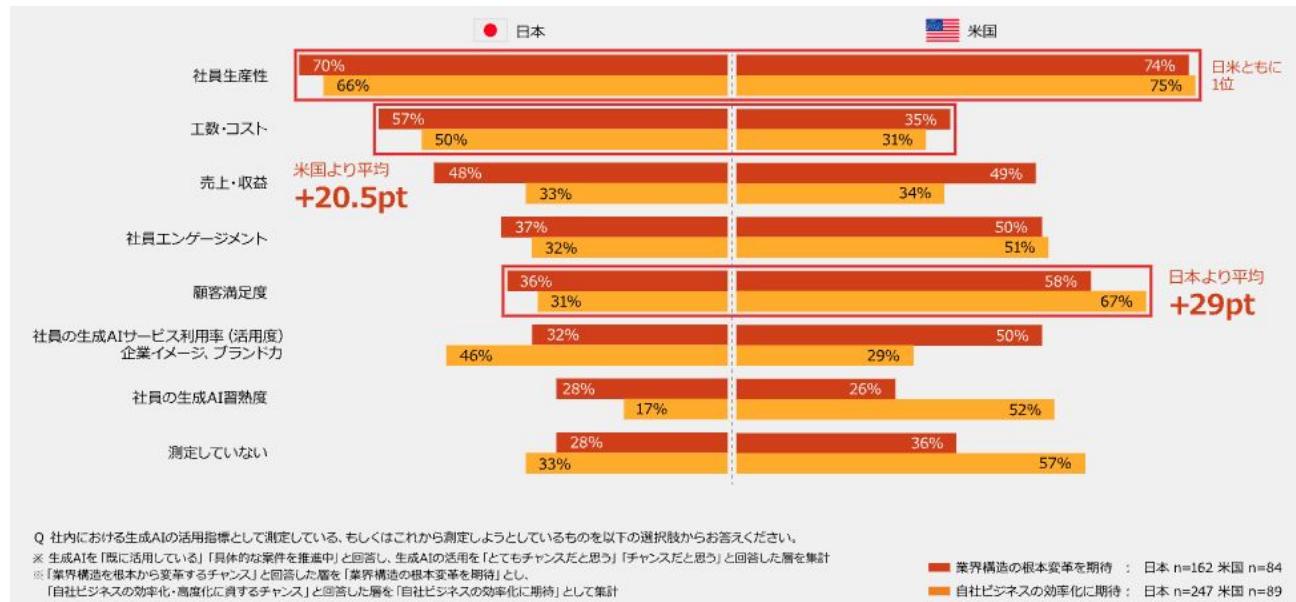


考察 | なぜ期待値を超えることができていないのか？

具体的なユースケースが定まっていないことから業務効率化に終止してしまい、米国では重要視されている顧客価値向上にアプローチできていないのでは

□ 生成AI活用の指標

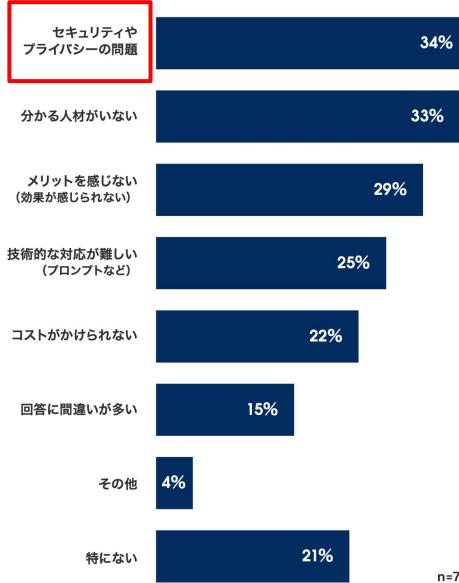
日本では工数・コストがより重要視されている傾向、一方米国では顧客満足度をより重要視する傾向が分かる



製造業における生成AI活用のハードル

『セキュリティ及びプライバシーの問題』が現状の生成AI活用及び将来的に生成AIの活用を広げる上での障壁となっている

Q4：Q1で「活用の予定はない」と答えた方への質問です。なぜ生成AIを活用しないのか、理由として当てはまるものを選択してください（複数回答可）

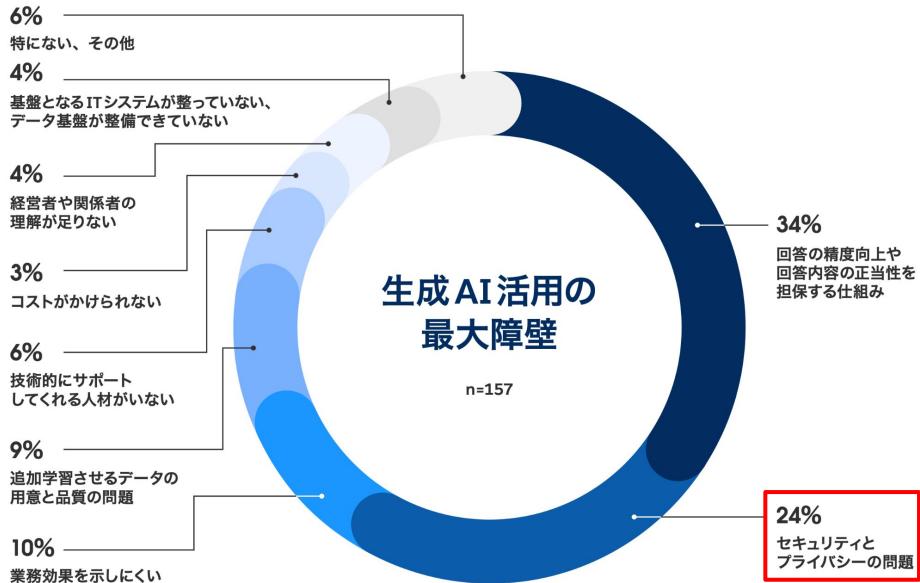


- 最も多かった回答は「セキュリティやプライバシーの問題」で34%、次点で多かった回答は「分かる人材がない」で33%だった

- セキュリティや人材の問題がAI活用の障壁であることがうかがえる。これらの回答は、AI導入に向けた教育とセキュリティ対策の重要性を示唆している

- また、メリットを感じないと答えた割合も高く、AIの有効性を広めるための情報提供が求められている

Q6：生成AIの活用を広げる上で一番の障壁だと感じていることを選択してください。



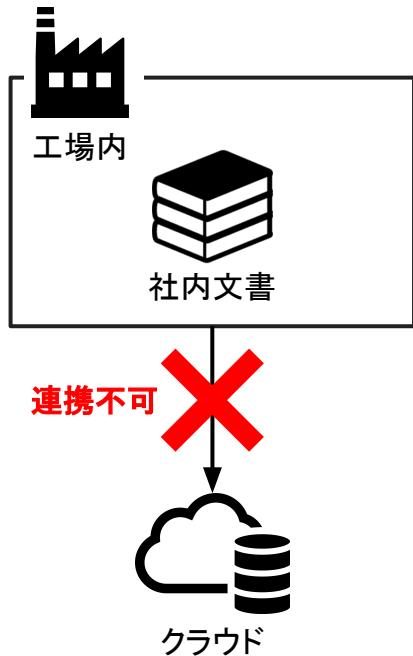
オンプレ開発の課題とコツ

生産現場におけるオンプレ活用のポテンシャル

クラウドを利用することなく工場内の閉じた環境で生成AIを利用することが可能に

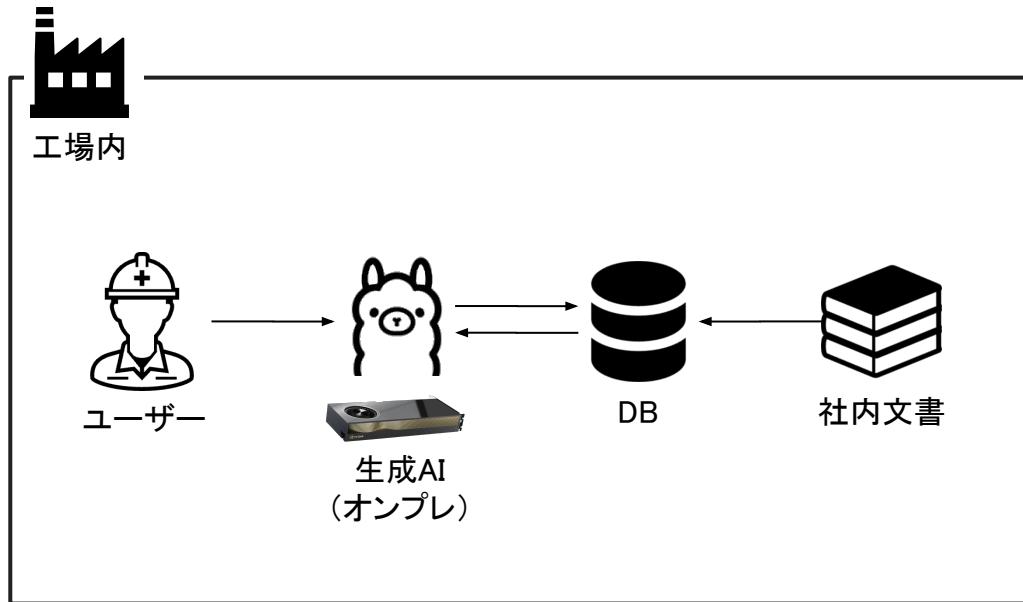
課題

工場内ではクラウドを利用できない



実現できること

オンプレ環境で開発することで、工場内完結で生成AIが利用可能



生産現場における生成AI活用の課題

高いセキュリティと引き換えに、精度改善の難しさと運用の難しさが発生する



課題1：LLMの種類が限定的

解決策：オープンLLM + PEFT/QLoRA + RAG



課題2：前処理の複雑さ

解決策：AI-OCR + 図面認識 + 用語正規化 + RAG



課題3：大規模利用の難しさ

解決策：ユースケース選定 + 量子化/vLLM最適化 + ハイブリッド



課題4：運用保守のコスト

解決策：MLOps/LLMOps + 監視
(Prometheus/Grafana/DCGM) + AIOps

生産現場における生成AI活用の課題

高いセキュリティと引き換えに、精度改善の難しさと運用の難しさが発生する



課題1：LLMの種類が限定的

解決策：オープンLLM + PEFT/QLoRA + RAG



課題2：前処理の複雑さ

解決策：AI-OCR + 図面認識 + 用語正規化 + RAG



課題3：大規模利用の難しさ

解決策：ユースケース選定 + 量子化/vLLM最適化 + ハイブリッド



課題4：運用保守のコスト

解決策：MLOps/LLMOps + 監視
(Prometheus/Grafana/DCGM) + AIOps

Closed LLMとOpen LLMの性能の違い

OpenLLMとClosedLLMで性能としては肉薄しているものの、安定性や深い思考力・画像に対する処理などまだ性能差は残っているのが現状

Closed LLM

OpenAIのgpt、AnthropicのClaude、GoogleのGeminiなど、企業が開発し、内部のトレーニングデータや詳細なモデルアーキテクチャが公開されていないLLM。高度な推論能力と幅広い知識を持つ



OpenAI

gpt

ANTHROPIC

Claude



Google

Gemini

Open LLM

OpenAIのgpt-oss、MetaのLLaM4、AlibabaのQwen32B、など、オープンソースで開発され、研究者や企業が自由に利用・カスタマイズできるLLM



OpenAI

gpt-oss



Meta AI



Llama



Qwen 32B

参考)最新技術 | gpt-oss

gpt-o4-miniと同等精度のモデルがオープンソースとして公開
クローズドな環境におけるLLM活用のゲームチェンジャーになり得る

□ 精度の高さ

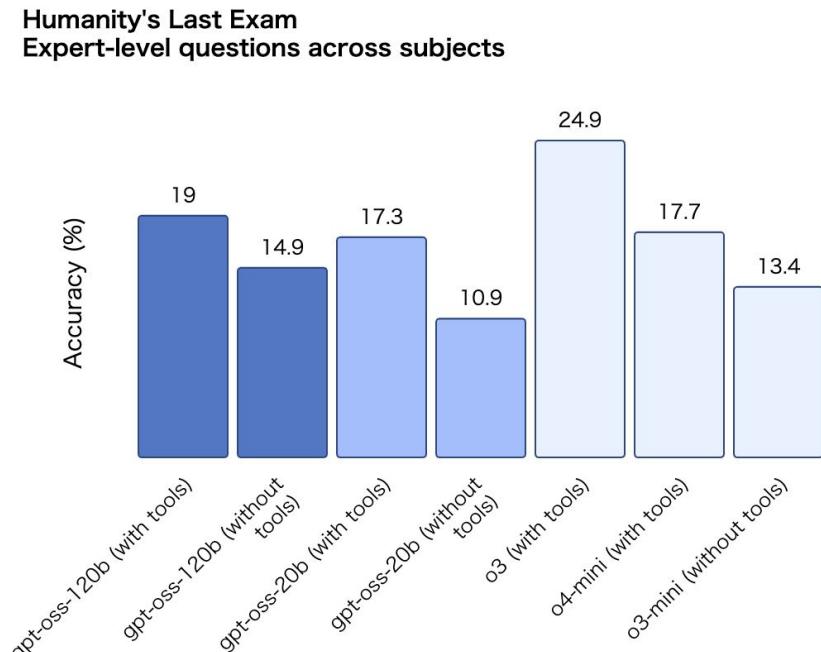
- 右図のように高い性能を示している
- またFine-tuning等の特化学習が現実的に可能

□ モデルの安定性・安心感

- これまでよく利用していた LLMと基本同じ
※ Copilotも中身は同じ
- 日本語も安定して出力でき、
思想の偏りも中華系 LLMより少ない

□ モデルのコンパクトさ

- gpt-oss 20Bであれば
数十万円程度の消費者向け PCで動作可能
- 120Bも大規模なシステムは不要

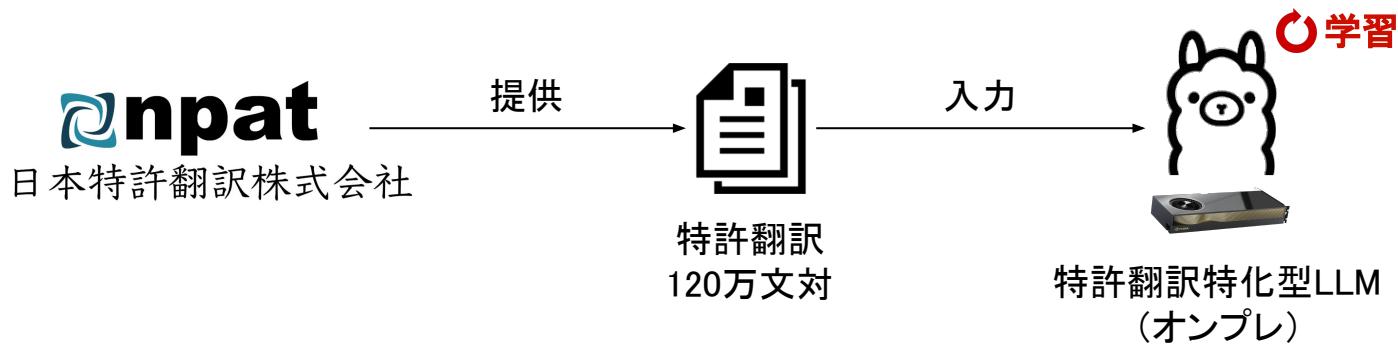


事例 | 独自翻訳LLMのオンプレ開発

特許翻訳に特化したLLMをオンプレ上で独自開発し調査を効率化

特許翻訳特化型LLMを用いることで、外国公報を1つ数十円で調査可能に

事前学習

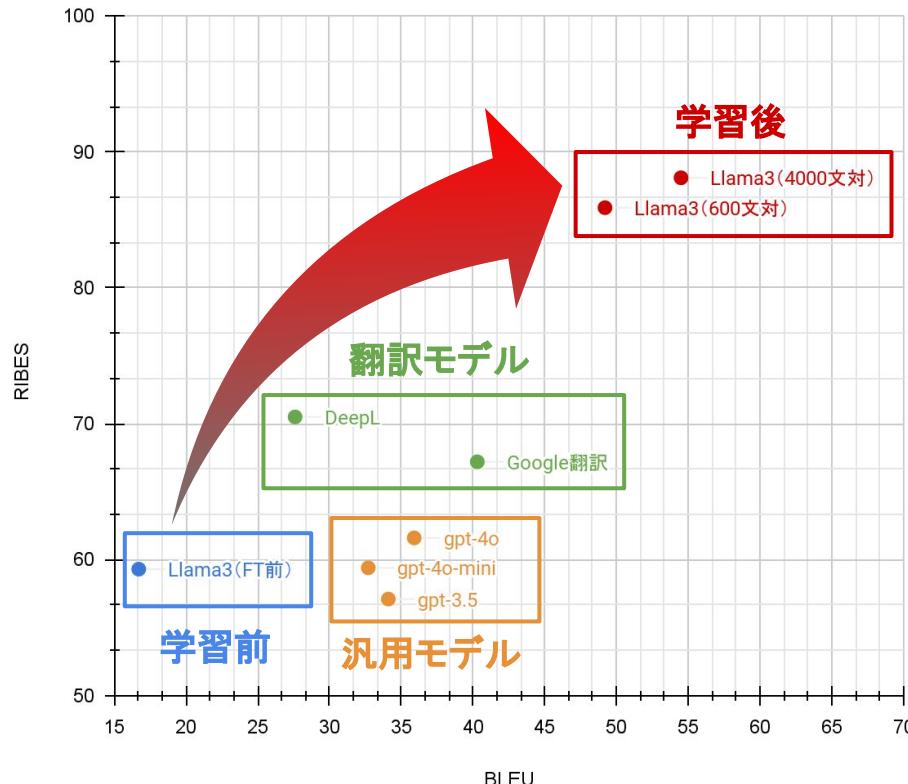


実運用



事例 | 独自翻訳LLMのオンプレ開発

ファインチューニングを実施することで、特許翻訳性能が大幅に向上



事例 | GPT-4oを超えるまで

問題設定の適切さに加えて、多くの課題を乗り換えることで達成

□ GPT-4oを超える特許翻訳性能を獲得できた要因

- データの種類と品質
外国語特許 ⇄ 日本語特許という対応関係が明瞭かつ、プロの手による高品質な翻訳結果
- 概念を捉えるというタスクとの相性
Fine-tuningで知識を取得するのは難しいが、翻訳の様な言葉の概念を理解することは得意
また今回は特定領域だけに特化させれば十分であるため、そこも相性が良かった

□ 独自LLMを開発する上でのハマリ所

- 些細なミスで正常に動かない
例)余分なインデックスが含まれる、接頭辞が抜けていた等
- 評価データの難易度が高すぎるため、そのままでは改善の打ち手が通用しなかった
タスクを分解して、ルールベースで補える部分は別で後処理を加えた

おまけ)検索についても、ただRAGを組むだけでは特許の言葉の使い回しに対応することが難しかった

生産現場における生成AI活用の課題

マネージドな機能が利用できないため前処理を自前で作り込む必要性が生じる



課題1：LLMの種類が限定的

解決策：オープンLLM + PEFT/QLoRA + RAG



課題2：前処理の複雑さ

解決策：AI-OCR + 図面認識 + 用語正規化 + RAG



課題3：大規模利用の難しさ

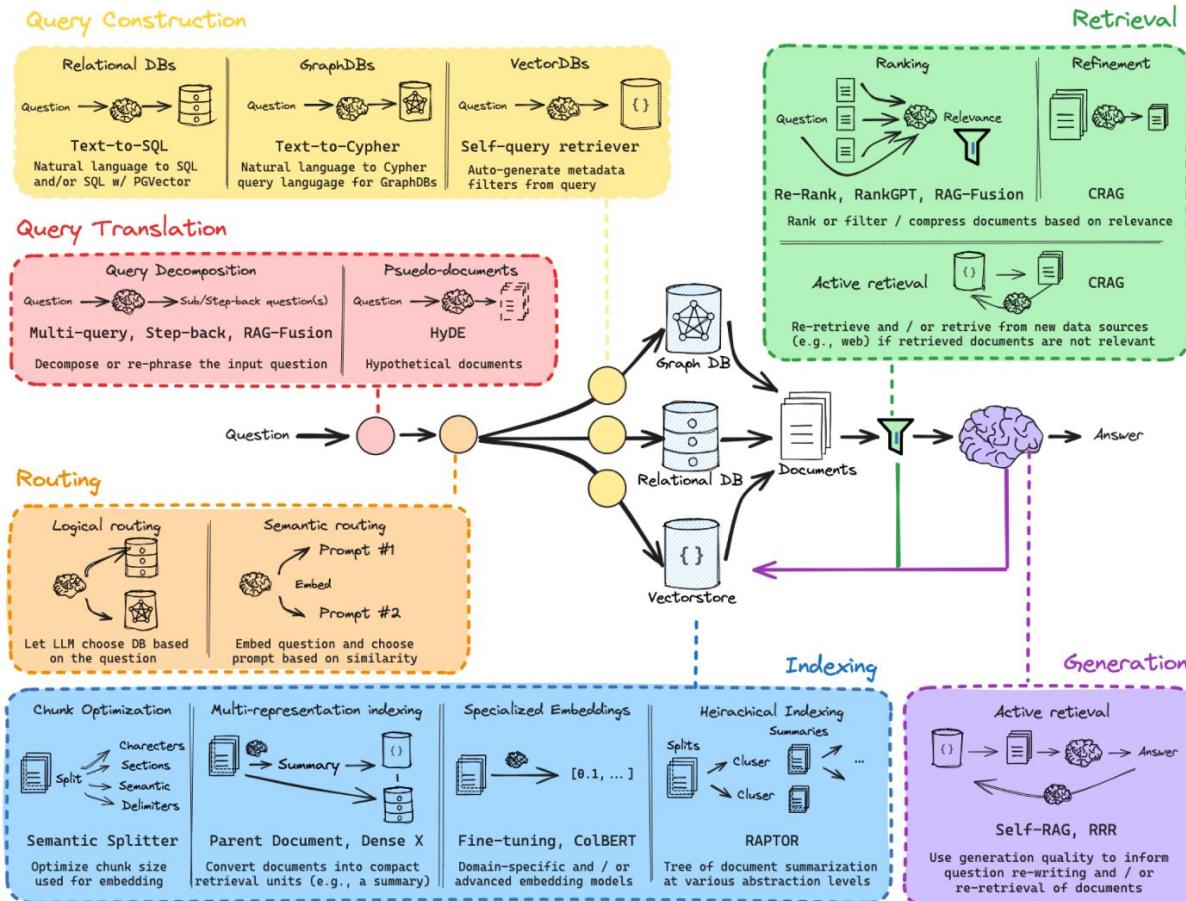
解決策：ユースケース選定 + 量子化/vLLM最適化 + ハイブリッド



課題4：運用保守のコスト

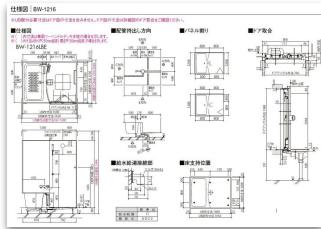
解決策：MLOps/LLMOps + 監視
(Prometheus/Grafana/DCGM) + AIOps

参考 | チャットボット開発で考慮するべきこと

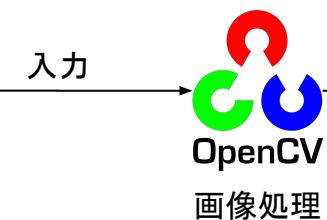


事例 | 図面からの情報抽出

AI-OCRと画像処理技術を活用して、図表・図面から必要な情報を抽出して検索や回答時に自由度高く参照・表示させることが可能に



図表



出力

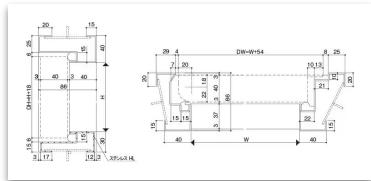


AI-OCR

出力



一覧表



図面



記号検出
アルゴリズム

生産現場における生成AI活用の課題

運用シーンを想定して、適切なユースケースで活用することが重要になる



課題1：LLMの種類が限定的

解決策：オープンLLM + PEFT/QLoRA + RAG



課題2：前処理の複雑さ

解決策：AI-OCR + 図面認識 + 用語正規化 + RAG



課題3：大規模利用の難しさ

解決策：ユースケース選定 + 量子化/vLLM最適化 + ハイブリッド



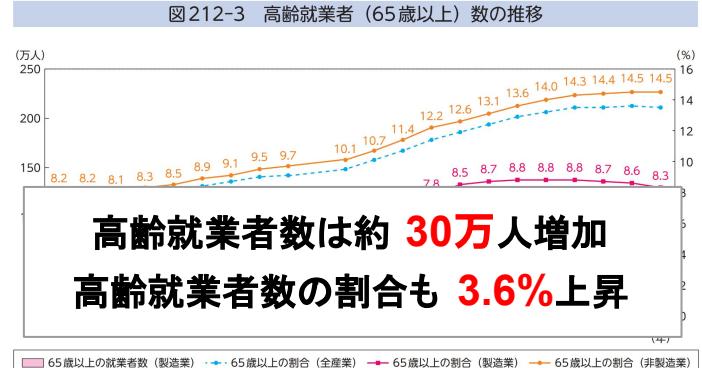
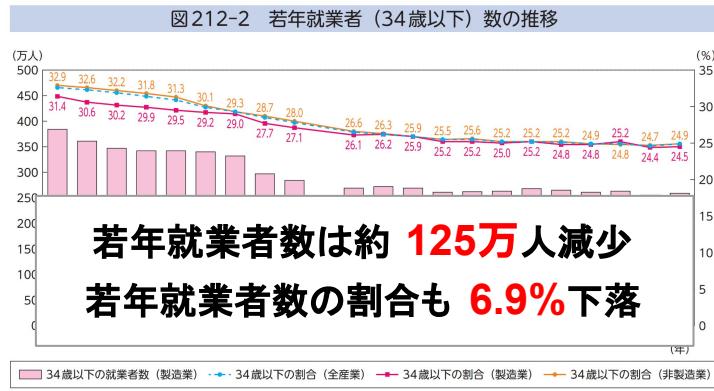
課題4：運用保守のコスト

解決策：MLOps/LLMOps + 監視
(Prometheus/Grafana/DCGM) + AIOps

生産現場における 生成AIのキラーユースケース

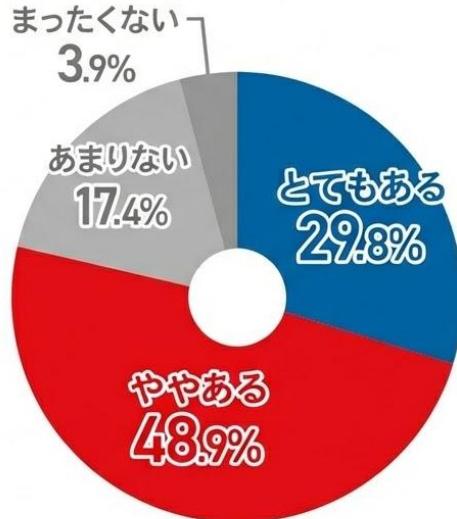
社会背景 | 製造業の現状

人員不足・高齢化に伴い、ベテラン等の業務の属人性を低減する必要性が高まっている



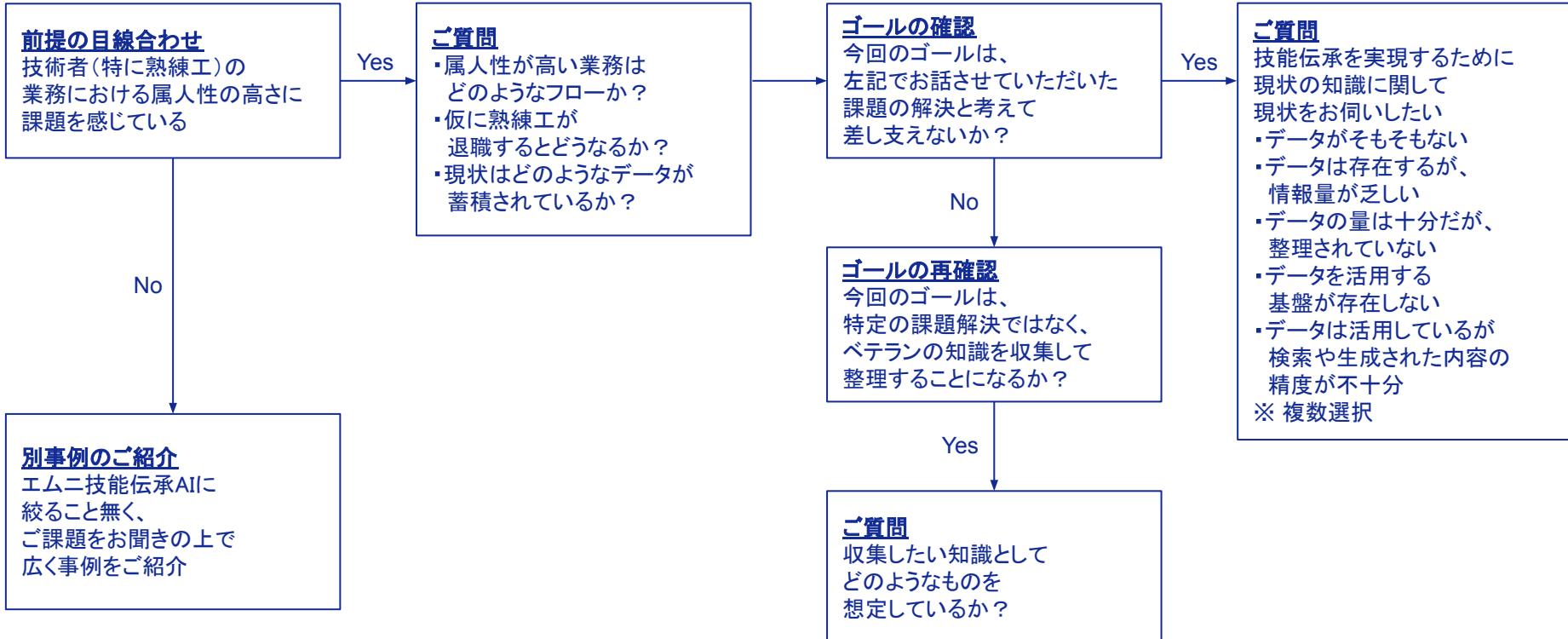
現在、属人化していると感じる業務はありますか？

【製造業の管理職】



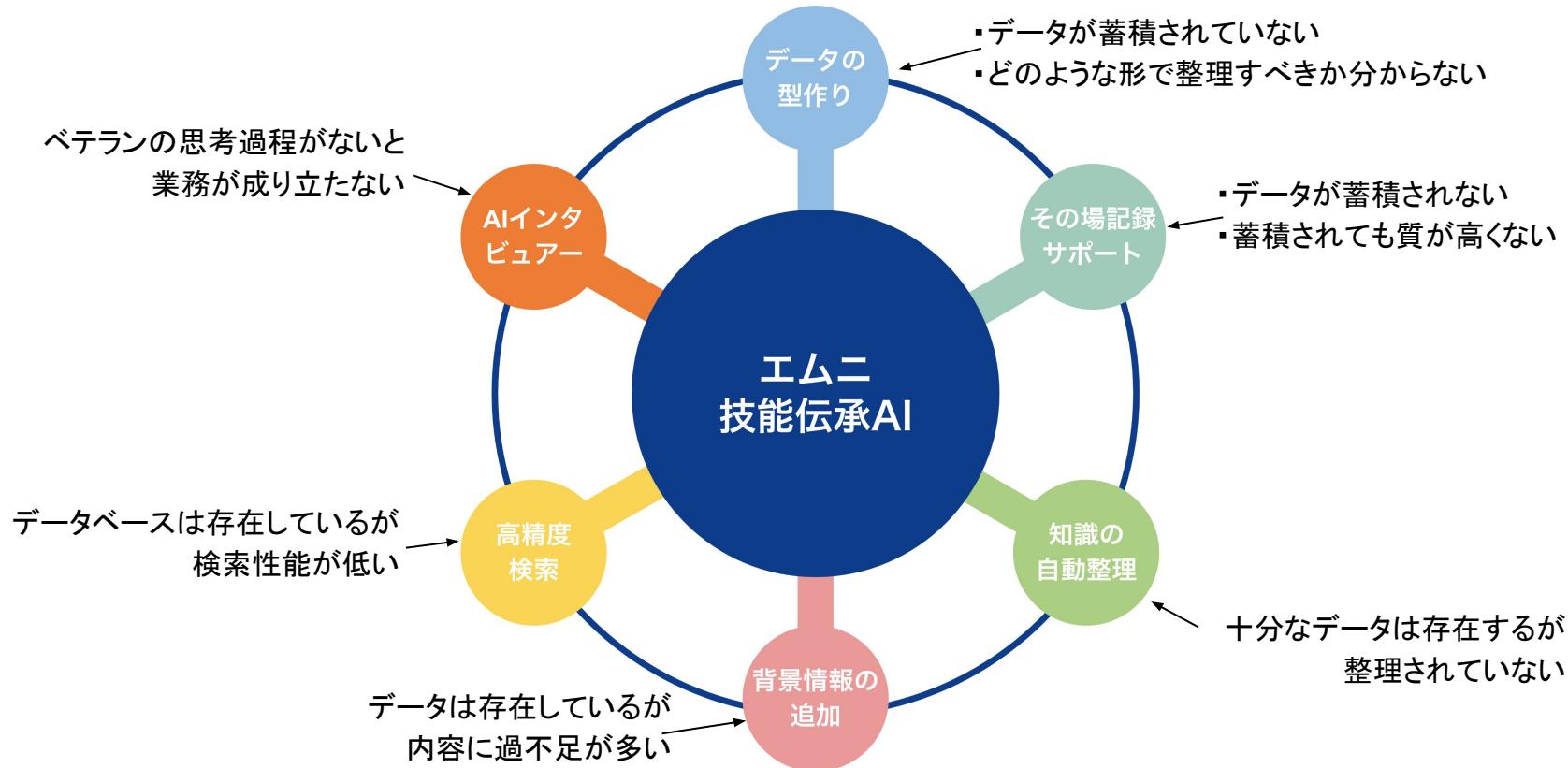
背景 | 課題感に関するお伺い

技能伝承を実現するために、まずは現状の業務内容やご課題感についてお伺いしたい



アプローチ概要

生成AIを活用した5つのモジュールによって、脱属人化を実現する



アプローチ詳細1 | データの型作り

データ全体の理想的な形式を定義して既存データの整理をする

□ データの理想的な状態とは

本来構造化されているべきものが記述されている状態

- 例)トラブル事例

トラブル報告書が記録されているだけではなく、
 対象の設備情報・部品の情報・製品の情報まで
 含めて保持している状態が理想
 ⇒ 設備の部品などの情報まで含めて
 データを保持するように設計

□ どうやって理想を定義するのか？

作成したいアプリケーションと

貴社のご要望に合わせて弊社の専門家が定義

トラブル報告書

Name	Type	Description
id	<code>uuid</code>	トラブルを一意に示すID
content	<code>str</code>	トラブルの内容
part	<code>str</code>	トラブル箇所
action	<code>str</code>	トラブルに対応するための行動
machine	<code>str</code>	対象の設備
reasoning	<code>str</code>	どうやってそのような対策にたどり着くのか

作業指示書

Name	Type	Description
id	<code>uuid</code>	作業指示書を一意に示すID
instructions	<code>List[str]</code>	step-by-stepの指示
machine	<code>str</code>	対象の設備
possible_failure	<code>str</code>	その作業時に発生しうるトラブル・失敗

アプローチ詳細2 | その場記録サポート



項目ごとに入力することで自然に構造化、またリアルタイムでAIがレビューを実施

コメント付きセクションエディタ（サンプル）

● 選択範囲にコメント ✎ コメントを全削除

テキストをドラッグで選択 → 「選択範囲にコメント」をクリック

セクション1：現象 section-1

I

セクション2：発生したトラブル事象 section-2

セクション3：検知方法 section-3

コメント 0件

まだコメントがありません。
左側のテキストを選択して「選択範囲にコメント」を押してください。

アプローチ詳細3 | 知識の自動整理

作成したデータの型を埋めて知識を整理する

□ 知識の整理方法

- 1で紹介した方法で知識のスキーマを決定する
- 現状蓄積されている既存のデータを入力として、上記で定義したスキーマに基づいてLLMでデータの型を埋める
- 埋めた知識の内容を人間がチェックを行う



既存文書

トラブル報告書		
Name	Type	Description
id	uuid	トラブルを一意に示すID
content	str	トラブルの内容
part	str	トラブル箇所
action	str	トラブルに対応するための行動
machine	str	対象の設備
reasoning	str	どうやってそのような対策にたどり着くのか



Name	内容
id	66515923-6484-4cf3-bef0-50cea9182aff
content	製品にキズがついてしまう
part	工程xxxx / 装置yyy
action	パラメタを調整して抑止した
reasoning	?

アプローチ詳細4 | 背景情報の追加



LLMの推論によって既存資料に記載されていない背景情報を追加する

□ 知識の整理方法

- ③で作成した整理済みの知識について、埋まっていない背景情報をLLMで推論
- 正誤を確認できる熟練工によって、LLMの出力内容を確認して修正を行う

Name	内容
id	66515923-6484-4cf3-bef0-50cea9182aff
content	製品にキズがついてしまう
part	工程xxxx / 装置yyy
action	パラメタを調整して抑止した
reasoning	?



LLM

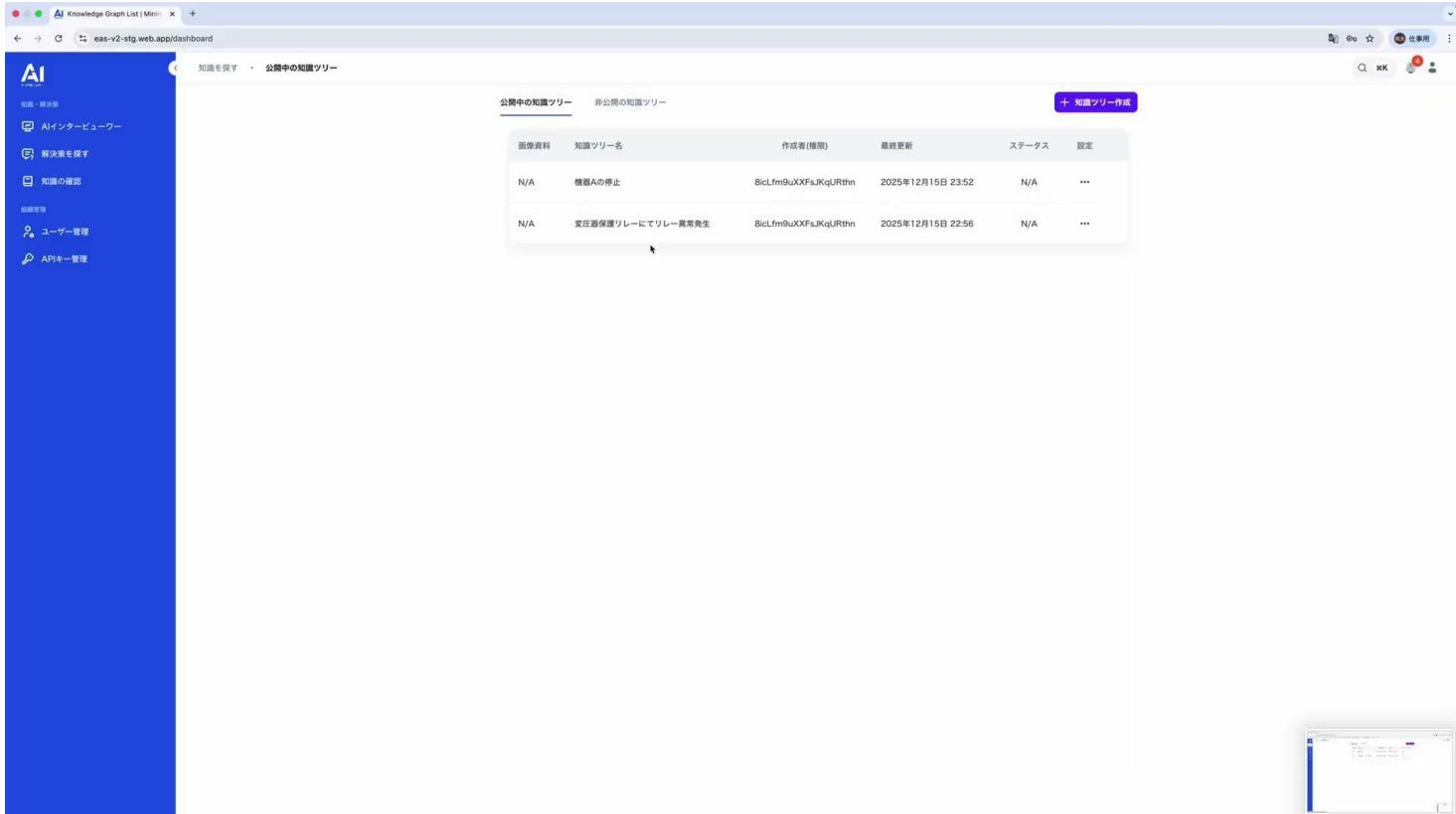
Name	内容
id	66515923-6484-4cf3-bef0-50cea9182aff
content	製品にキズがついてしまう
part	工程xxxx / 装置yyy
action	パラメタを調整して抑止した
reasoning	位置ずれで干渉し擦過痕が発生



熟練工

アプローチ詳細6 | AIインタビュアー

LLMによる深堀り質問を通して、熟練工の考え方や判断基準を抽出



The screenshot shows a web-based dashboard titled "AI Knowledge Graph List". The left sidebar contains navigation links: "知識・検索", "AIインタビュワー", "新決策を探す", "知識の確認", "組織管理", "ユーザー管理", and "APIキー管理". The main content area has a header "知識を探索 公開中の知識ツリー". Below it is a table with two rows of data:

公開資料	知識ツリーネーム	作成者(権限)	最終更新	ステータス	設定
N/A	機器Aの停止	BicLfm9uXXFsJKqURthn	2025年12月15日 23:52	N/A	...
N/A	変圧器保護リレーにてリレー異常発生	BicLfm9uXXFsJKqURthn	2025年12月15日 22:56	N/A	...

PoC実施及び準備状況

トヨタ自動車を始めとしてPoC実施中8社、実施準備中5社以上

PoC実施中企業	対象業務	取組詳細
トヨタ自動車株式会社	生産技術	設備監視における異常判定
輸送機器メーカー	生産技術	生産工程における異常対応の迅速化
株式会社明電舎	保全	機器の故障発生時の保全対応迅速化
インフラ系メーカー	生産技術	異常判定における熟練工のノウハウ抽出
株式会社ダイセル	生産技術	生産安定化のためのノウハウ抽出
輸送機器メーカー	設計	ベテランが持つ設計時のノウハウ抽出
化学メーカー	生産技術	異常発生時の対応迅速化、及びFMEA作成
輸送機器メーカー	保全	トラブル発生時の負荷軽減・復旧迅速化



経済産業省主催の
懸賞金活用型プログラムに
以下2社と共同エントリー中

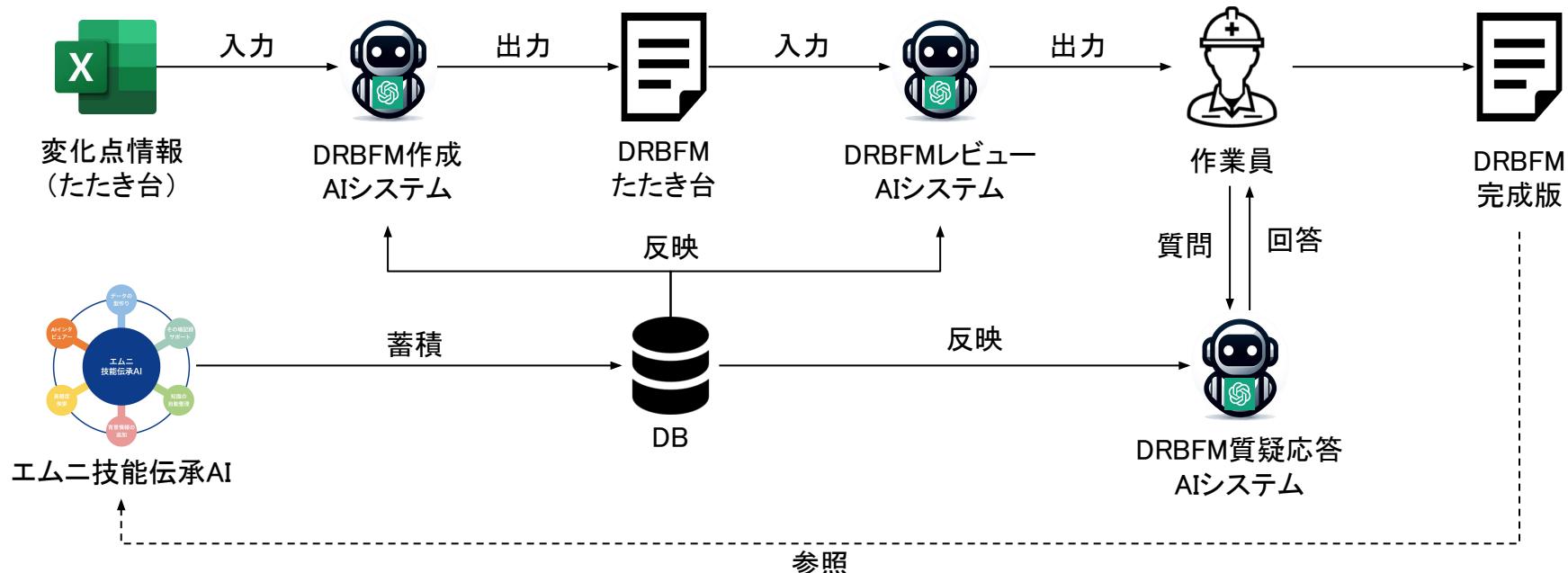
領域01：製造業

 DAICEL  エムニ 領域 01	 MEIDEN <i>Quality connecting the next</i>  エムニ 領域 01
<small>大阪府大阪市</small> <small>株式会社ダイセル</small> <small>株式会社エムニ</small>	
<small>東京都品川区</small> <small>株式会社明電舎</small> <small>株式会社エムニ</small>	
株式会社ダイセル	株式会社明電舎
株式会社エムニ	株式会社エムニ

参考)導入効果 | リスク観点の見逃し削減

「変化点からDRBFMの叩き台を作成→AIの自動レビュー→チャットボット経由で知識を参照した効率的な確認」により、リスク観点をスケモレなく抽出することが可能になる

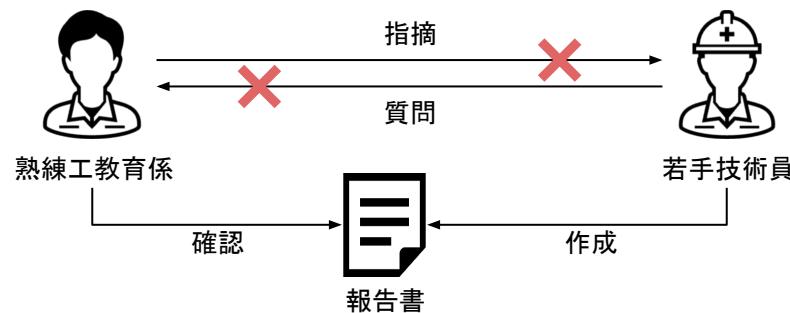
□ 目指す姿



参考)テーマ例 | 現場の若返り実現

経験等の熟練工の知識をAIインタビューで取り出すことによって、
若手はいつでも質問して熟練工の経験に基づいた回答を参照することを可能にする

- 現状…熟練工の考えはなかなか伝わらず、若手は中々質問することができない



- 目指す姿…若手はいつでも質問することができ、熟練工の経験に紐づいた回答を得ることができる



※ 熟練工が回答しやすいようにクイズ形式に変換を行う

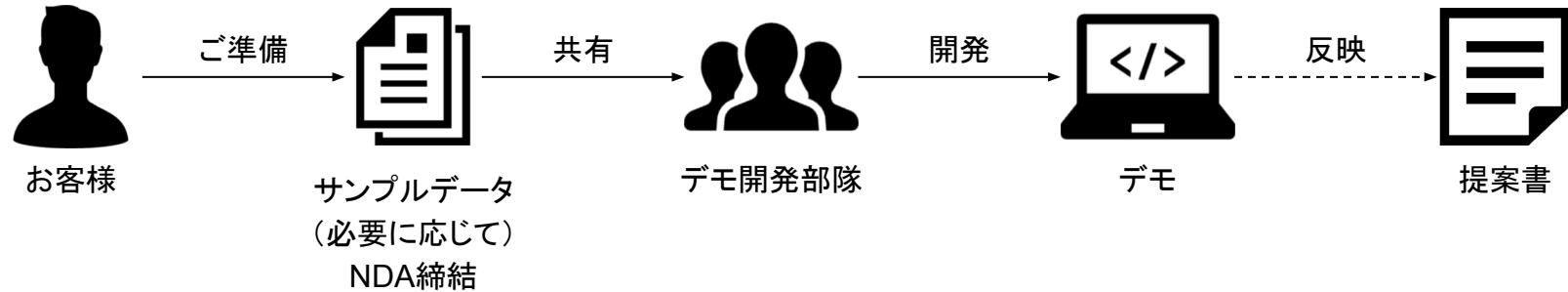
終わりに

無料 爆速デモ開発



無料でデモを爆速開発し、AI活用を検討するまでの不確実性を低減

□ デモの流れ



□ 取り組み例

紙芝居ではなく、実際の動作イメージや精度が分かるようなデモシステムを開発可能

- **1日で高速デモ開発**
論文サンプルを頂いた翌日に、調査効率化 AIのデモシステムをご説明
- **1週間で複雑な要件を整理・プロトタイプまで開発**
100ページに及ぶサンプル資料を確認し、要件整理から実際に動くシステムまで構築

Thank you .

