

MSIISMコンファレンス 数理システムの考える生成AIビジネス活用 ～今、生成AIの使い所は？～

株式会社NTTデータ数理システム
数理ソリューション部 村田 康佑

アジェンダ

生成AIのビジネス活用の実情を、
現場の技術・営業の立場から、**主観的**にお話します

- 最近の生成AI、どれぐらいの実力がある？——百聞は一見に如かず
- 生成AI導入のコツ——導入を進めるのが難しい理由と対策
- 良い生成AIプロジェクトとは？

もう少し客観的な全体観は…

AI・数理科学市場の状況や、他の技術との関係性については
MSIISM Conference 2024 基調講演の**オンデマンド配信**で

基調
講演

株式会社NTTデータ数理システム



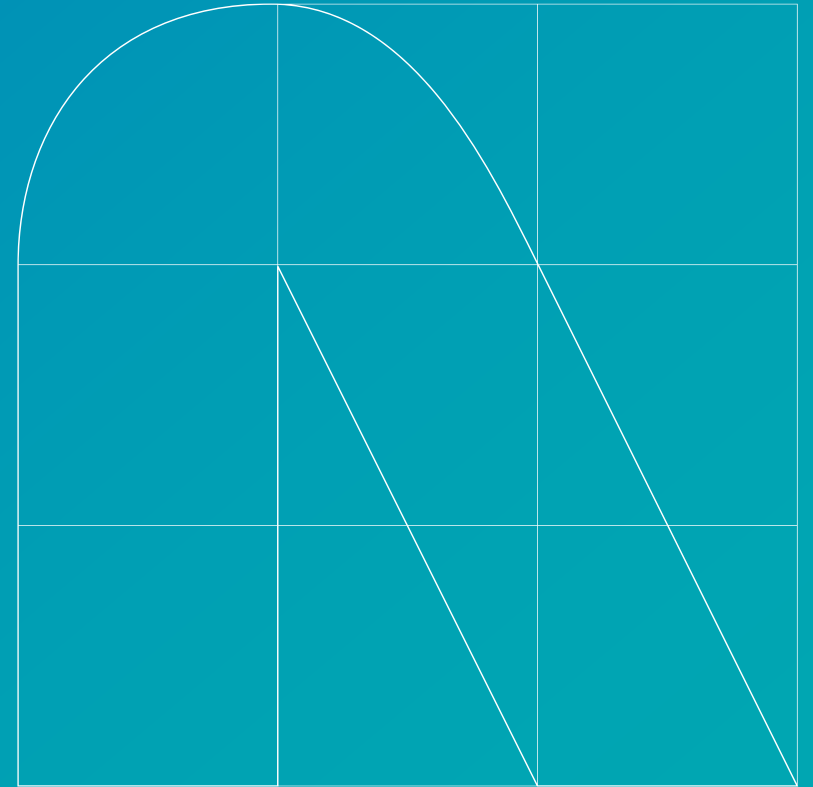
シミュレーション&マイニング部
グループリーダー

伊藤 孝太郎

数理科学と生成AIで革新を

1

最近の生成AI、
どれぐらいの実力がある？
——百聞は一見に如かず



百聞は一見に如かず——言語生成AIの実力

**もし室町時代のころの日本に生成AIがあったとしたら、
日本の文化はどのように変わりますか？**

1. 室町時代に生成AIが存在していた場合、茶道では**茶室の設計や茶器の選定が科学的に最適化**され、新たな美学が生まれた可能性がある。
2. 能楽では**古典文学を基に生成AIが新しい台本を作成し**、観客に革新的な芸術体験を提供することができた。
3. 書道において**生成AIは仮想の師匠として機能し、効率的な技術習得と個性を引き立てる指導を実現し**、新しい書道の可能性を広げる。
4. 生成AIは文化の**均質化の懸念**とともに、多様な選択肢を提供することで新しい文化形成の機会を創出し、**情報の民主化**を進める。
5. 生成AIは**文化の失われた要素を取り戻し**、技術と文化の相互作用を通じて未来に向けた革新を支える重要なツールである。

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。青字・赤字による強調は NTT データ数理システムによる。

**室町時代の代表的文化（茶道、能楽、書道）における生成AIの影響を
現代的なテーマを交えつつ、示唆を含む考察ができています**

百聞は一見に如かず——画像生成AIの実力

**もし令和まで江戸幕府が政権を維持していたときの
日本橋の様子を想像して描いて**



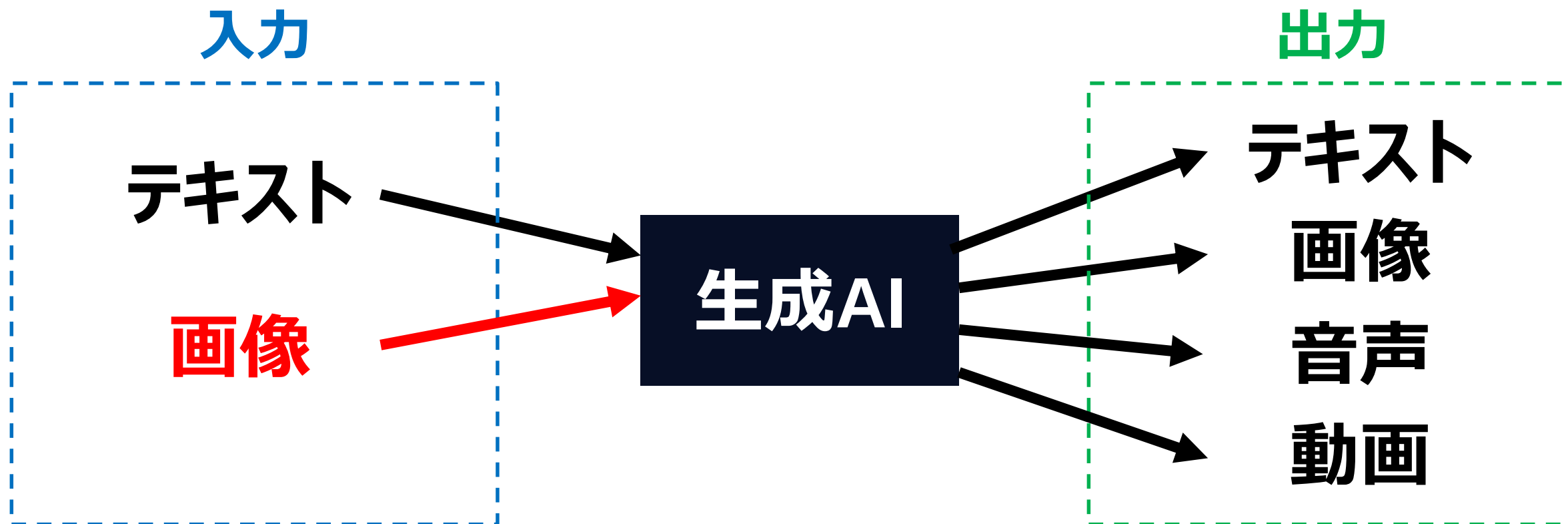
いずれも、NTTデータ数理システムがChatGPT（DALL-E3）を利用して生成した画像。



**背景の現代的なビルと、前景の伝統風の建築物を調和して描いている
アイデアの凡庸さや細部の粗が目立つものの、数秒で生成できるのは衝撃的**

ビジョン付き生成AIの登場

- ・ 主に文章生成・画像生成が主力となり発展してきたが、最近入出力が多様に
- ・ GPT-4Vの登場以降、画像も入力可能な**ビジョン付き生成AI**が台頭
→ 画像を解釈することができ、ユースケースも大きく広がる



百聞は一見に如かず——ビジョンの実力

A0サイズ×2枚のポスター画像

生成AI活用 受託開発・コンサルティングサービス

アイデアはあるのに
推進してくれる仲間がない

他社に相談したら
「難しい」と言われてしまった

生成AI活用をはじめたいけど…

- ・生成AIはすごそうだけど、業務となると何に使えるのかわからない
- ・メリットやリスクをよく理解できていない
- ・何から手を付けて良いかわからない

とりあえず、やってみただけ…

- ・とりあえず流行りのRAGを構築してみたが、思ったような精度が得られない
- ・とりあえずAIチャットを導入したが、あまり活用されていない

AI活用プロジェクトが立ち上がった
でも、成果に結びついていない

他社サービスでは
痒い所に手が届かなかった

生成AIのお悩み、お聞かせください！

約40年の歴史と8,000案件以上の実績を持つ、われわれの技術と経験でどうにかできるかもしれません

生成AI活用プロジェクト、NTTデータ数理システムが伴走します！

当社が特に強みを持つ領域

課題認識・業務検討	企画検討	PoC (1-3ヶ月)	実装・システム組込	運用
・技術アドバイス ・事例紹介	・ワークショップ・課題整理 ・実用性検証 ・技術課題を踏まえ、ゴール・スケジュールの設定	・モデル・アルゴリズム構築 ・モデル・アルゴリズム評価 ・実用性検証 ・試行期間中リアルタイム ・システム実装	・モデルの精度監視 ・モデルの再学習・再構築	・モデルの精度監視 ・モデルの再学習・再構築

われわれの強み

- ・ 定量的手法から最新論文まで——数理解科学とコンピュータサイエンスを愛する技術者による徹底した調査に基づく提案
- ・ コアアルゴリズム実装からGUI制作まで——業務で現実的に「使えるか」にこだわった、ワンストップなご提案

実際、生成AIってどんなことに使われているの？——生成AIのメジャーユースケース

業務補助AIチャット ナレッジ活用 商品企画 ヘルソナ自動作成・分析 手書きデータ活用 スライドデータ活用

RAG (Retrieval-Augmented Generation)

社内データなど、生成AIが学習していないデータでも生成AIが知識を補完しながら回答を生成

LLMを再学習させることは、新規知識を出力内容に反映できる
社内データベースのチャットボットなど、人手で確認しきれない膨大な量な知識の効率的な活用

非定型データからの情報抽出・半構造化

画像を読み取ることでさまざまなビジュアルLLMにより複雑な文書や手書きフォームからの情報抽出が可能に

・ 活字・手書き文字を精度よく認識可能
・ 多岐選択 (○を付けて回答)、自由記述など、様々な形式のフォームを埋め込んで抽出できる

生成AI関連で、わたしたちがご提案できる技術の一例です。

RDBなどの表データも、RAGの入力に

入力の表の変換、中間結果を作りながら繰り返すことで高精度な応答を実現 (Chain-of-Table)

列選択、ソート、カウントなどの基本操作を生成AIが判断しながら繰り返す

表データ → 中間データ → 質問に答える部分のみ残す

膨大な気象・地震ログデータからの情報探索

質問を理解したうえで、「(画面)列を選択してフィルタするクエリを実行

応用先：POSデータなど、各種表データに基づくQAシステム

プログラムのソースコードを自動解析、RAGに組み込む

レポート全体を解析、コールグラフに基づきドキュメントを生成
ソースコードを読み取れば答えられない難しいQAも可能に

repository directory fileA function ドキュメント
fileB class ドキュメント

ソースコードに基づきトラブルシューティングのQAチャット

「Method is not implemented」というエラーは、Androidの学習データベース「SQLite」の学習データベースに存在しないため発生しています。Androidの学習データベースは、以下のメソッドをサポートしています。以下のメソッドをサポートしていないため、エラーが発生しています。

クラス構造をコードから読み取りエラー原因を特定

応用先：ソフトウェア開発・サポート業務の効率化

ナレッジグラフを利用して、RAGの回答正確性向上

製品マニュアルや手順書などの文書群から、概念の関係を抽出
関係性をグラフで表現して蓄積することで、回答の正確性が向上

文書群 (文書A, B, C, ...)

ナレッジグラフを参照した質問応答例

Q. Alkanonには、AutoModeling [文書B]、線形回帰 [文書C]などの機能があります。

通常のRAGが苦手な「複数の文書を正確に読み取って答えるべき質問でも、正確な回答を生成できる

応用先：医療、法律など正確性や根拠が特に要求される分野

音声認識と音声生成技術を利用したロールプレイング

AIエージェントとの音声対話によるロールプレイング
コールセンターOPや営業が場数を踏んで知識定着を図る

音声通話による製品トラブル対応シーン

シーン・状況にあわせてカスタマーの発言を動的に生成

あなたとこの製品の製品を買ったけど、全然動かないんだけど、不良品なんじゃないの？

AIカスタマー

訓練生の発言を書き起こし

この度は弊社製品の動作不良により、ご迷惑をおかけして申し訳ございません。……

たしかに製品を壁に面して使っていました。少し離して動かしてみたら動きました！

AIカスタマー

応用先：OP・営業育成 (主に自社製品の知識定着を図る段階)

AIエージェントによるマルチエージェントディスカッション

通常のチャットAIでは回答が難しい多側面からの考察が必要な質問でも、複数のAIエージェントをディスカッションすることで、多様な立場や専門性から自律的に論点を整理・深掘りしながら最終回答を作成

各AIエージェントには
ヘルソナや専門性を設定

・ 30代女性
・ 埼玉在住
・ ゲーム業界に詳しい

AIエージェントは、
各ヘルソナから知識を
参照しながら発言

インターネット トキコメント 動画DB

例 自社の福利厚生制度に対する不満点は？

20代・男性・新社員エージェント
リモートワーク補助費が4000円なのは、少なすぎます。
人事部エージェント
具体的にどれくらいあれば満足度が上がりますか？
30代・女性・中堅社員エージェント
そうですね。インターネット回線だけで5000円はかかるし、それに電気代もあわせて10,000円くらいは欲しいです。

応用先：福利グループインタビュー、新商品企画アイデアエクスジェン

これはNTTデータ数理システムが
生成AI開発関係のソリューションを
提示したポスターです

このポスターでどのような強みが
述べられているかを箇条書きで
整理して

A0サイズ×2枚のポスター



強みとして提示 している点を抽出

強みとして提示している点

第一枚目のポスター:

- 多くの実績
 - 約40年の歴史と8,000件以上の案件実績
- 総合力
 - 定石的手法から最新のAI技術までカバー
- 専門知識
 - 数理科学とコンピュータサイエンスを基盤とした徹底的な調査と提案

第二枚目のポスター:

- 多岐に渡る応用例の紹介
 - ソフトウェア開発支援、POSデータ解析、法務・医療分野の応用など
- 技術的な深さと応用力
 - 複雑な課題に対する具体的な解決策の提示
- 多機能な技術ソリューション
 - ナレッジグラフ、音声認識、AIエージェントなど複数の技術の統合

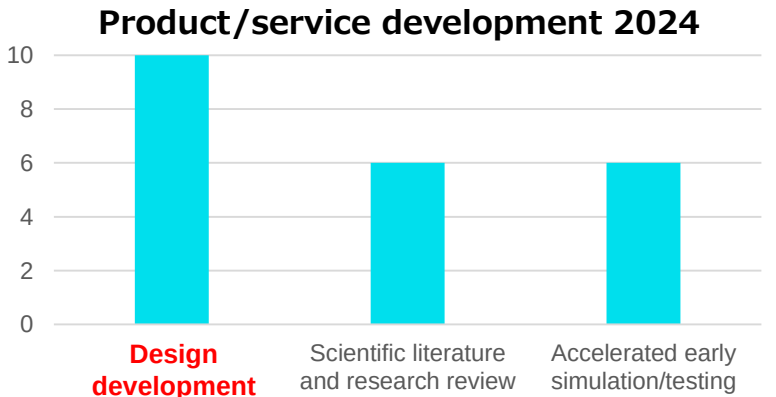
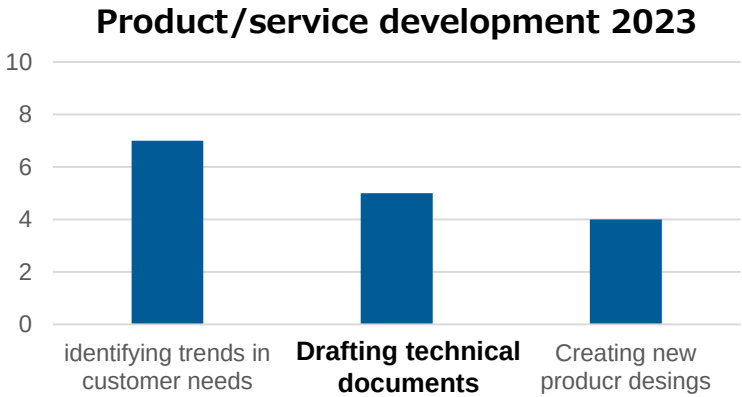
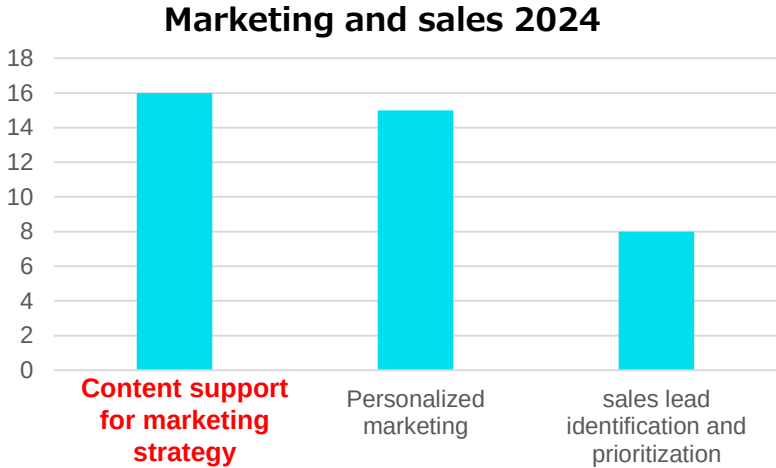
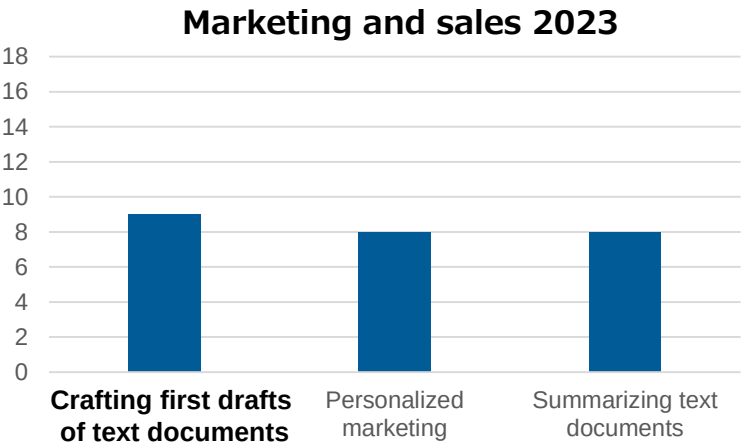
このようにNTTデータ数理システムは、豊富な経験と幅広い技術的専門知識を生かして、生成AIを活用した具体的で高度なソリューションを提供しています。

文字やポスターの構造・主張を正確に読み取って、強みを抽出できている
デザイン・画像データ → テキストデータへの情報抽出・構造化が可能に

生成AIの応用先は、ドラフト文書の作成からデザイン生成・マーケティングへと多様化。

✓部門ごとの生成AI応用先Top3

自部門で対象業務に
応用していると回答した
組織の割合(%)

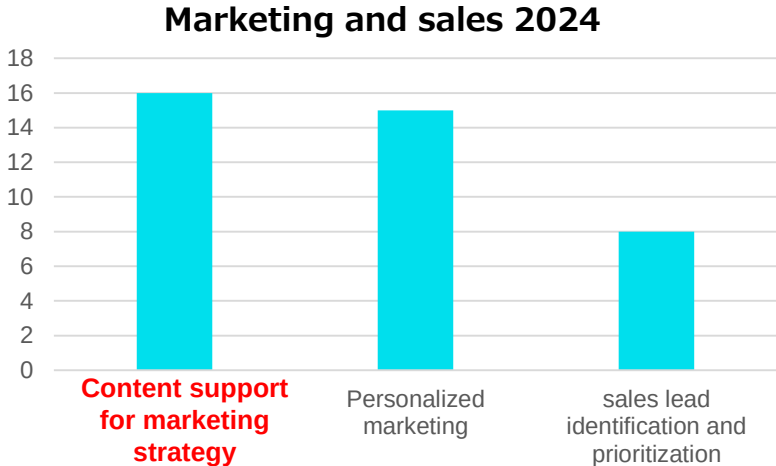
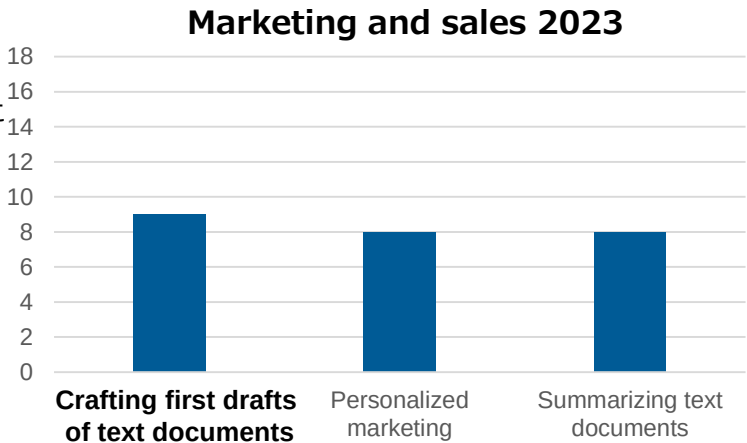


データ : McKinsey & Company The state of AI in early 2024, The state of AI in 2023
グラフ : NTTデータ数理システム



✓部門ごとの生成AI応用先Top3

自部門で対象業務に
応用していると回答した
組織の割合(%)



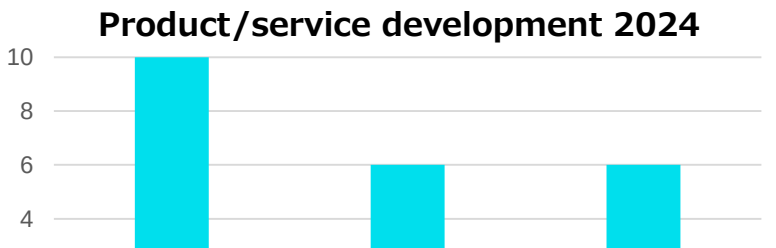
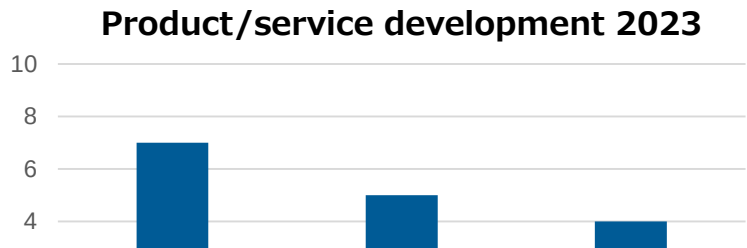
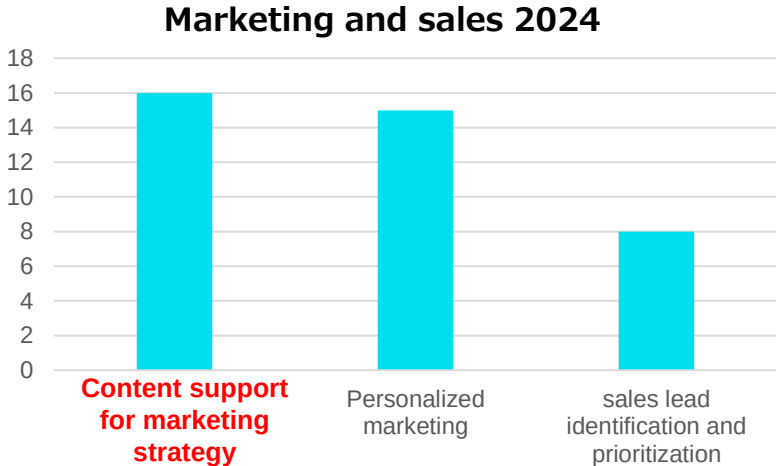
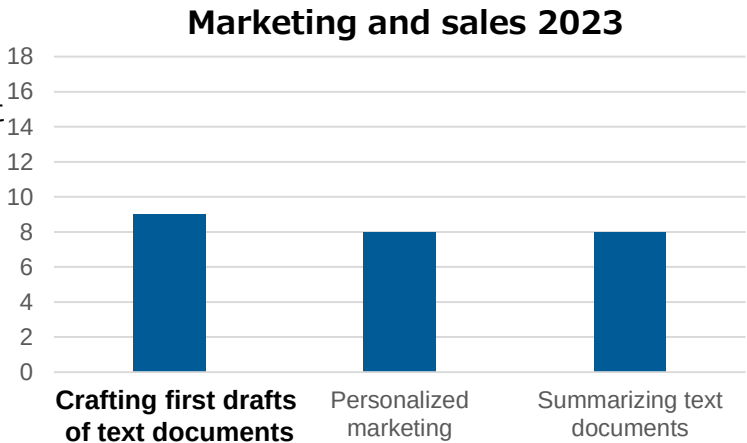
画像のスライドの構造（位置関係や強調された文字）や矢印などの図形の意味、
スライド中に現れるグラフから読みとれる事実から、
スライド作成者の意図を整理しながら読み取ってください。
そのうえで、画像中の ??? にあてはまるキーメッセージを、
??? の部分に当てはまる文量で作成してください

データ：McKinsey
グラフ：NTTデータ数理システム

生成AIの応用は2024年に向けて多様化し、より高度な領域へ拡大

✓部門ごとの生成AI応用先Top3

自部門で対象業務に
応用していると回答した
組織の割合(%)



スライドの構造から「いいたいこと」を正確に読み取ってメッセージを考えられている
→ ポンチ絵的な複雑なスライドから、要点をテキストデータとして抽出可能に

これまでの苦手を克服しつつある生成AI

- 少し前まで、生成AIは高度な推論が必要な数学・物理は苦手と言われていた
→ **9月リリースの OpenAI o1 モデルで圧倒的な性能向上、苦手克服？**
- 当代最高峰の数学者の一人テレンス・タオ氏も、肯定的な予想を述べる



Terence Tao

@tao@mathstodon.xyz

In chatgpt.com/share/94152e76-751... I gave the new model a challenging complex analysis problem (which I had previously asked GPT4 to assist in writing up a proof of in chatgpt.com/share/63c5774a-d58...). Here the results were better than previous models, but still slightly disappointing: the new model could work its way to a correct (and well-written) solution *if* provided a lot of hints and prodding, but did not generate the key conceptual ideas on its own, and did make some non-trivial mistakes. The experience seemed roughly on par with trying to advise a mediocre, but not completely incompetent, (static simulation of a) graduate student. However, this was an improvement over previous models, whose capability was closer to an actually incompetent (static simulation of a) graduate student. It may only take one or two further iterations of improved capability (and integration with other tools, such as computer algebra packages and proof assistants) until the level of "(static simulation of a) competent graduate student" is reached, at which point I could see this tool being of significant use in research level tasks. (2/3)

[Parenthetical clarifications added - 9/19/2024]

2024年9月14日 07:03 · 9月19日 15:55に編集 · Web · 20 · 54

あと1, 2回の改善（と計算機代数ソフトや証明支援系など、他ツールとの融合）により、「**有能な大学院生レベル（の統計的シミュレーション）**」に到達し得る

テレンスタオ氏のマストドン（mathstodon）投稿 <https://mathstodon.xyz/@tao/113132503432772494> より引用

百聞は一見に如かず——数学もできるようになってきた生成AI

$f: [0, 1] \rightarrow \mathbb{R}$ が C^2 級るとき、次の等式を示しなさい。

$$\lim_{n \rightarrow \infty} n \left(\int_0^1 f(x) dx - \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \right) = \frac{1}{2} (f(1) - f(0))$$

(当社オウンドメディア MSIISM <https://www.msiism.jp/article/flower-exam2021-2.html> より)

- 当社の過去入社試験問題（2021年度出題）
- 大学1年次程度の数学で解けるものの、深い考察が必要で簡単ではない
(数学専攻の大学院生受験者であっても、全員が解けるような問題ではない)

この問題は、関数 $f: [0, 1] \rightarrow \mathbb{R}$ が C^2 級、つまり2回微分可能であり、その2階微分が連続であることを利用します。目的は、積分値と区分求積法による近似値の差を n 倍したときの極限を求めることです。

まず、与えられた等式の左辺を整理します。

$$n \left(\int_0^1 f(x) dx - \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \right) = n \int_0^1 f(x) dx - \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right)$$

この差を調べるために、各小区間 $[\frac{k}{n}, \frac{k+1}{n}]$ における積分値と点 $x = \frac{k}{n}$ での関数値の差を考えます。関数 f を点 $x = \frac{k}{n}$ の周りでテイラー展開すると、

$$f(x) = f\left(\frac{k}{n}\right) + f'\left(\frac{k}{n}\right) \left(x - \frac{k}{n}\right) + \frac{1}{2} f''(\xi) \left(x - \frac{k}{n}\right)^2$$

となります（ただし、 ξ は $\frac{k}{n} \leq \xi \leq x$ の適当な値）。

この式を区間 $[\frac{k}{n}, \frac{k+1}{n}]$ で積分すると、

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x) dx = f\left(\frac{k}{n}\right) \frac{1}{n} + \frac{f'\left(\frac{k}{n}\right)}{2n^2} + \frac{f''(\xi)}{6n^3}$$

となります。ここで、 $\frac{1}{n} f\left(\frac{k}{n}\right)$ を引くと、

$$\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x) dx - \frac{1}{n} f\left(\frac{k}{n}\right) = \frac{f'\left(\frac{k}{n}\right)}{2n^2} + \frac{f''(\xi)}{6n^3}$$

NTTデータ数理システムがChatGPT（o1-preview）を利用して生成した文章

これを n 倍すると、

$$n \left(\int_{\frac{k}{n}}^{\frac{k+1}{n}} f(x) dx - \frac{1}{n} f\left(\frac{k}{n}\right) \right) = \frac{f'\left(\frac{k}{n}\right)}{2n} + \frac{f''(\xi)}{6n^2}$$

全ての k について合計すると、

$$n \left(\int_0^1 f(x) dx - \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \right) = \sum_{k=0}^{n-1} \left(\frac{f'\left(\frac{k}{n}\right)}{2n} + \frac{f''(\xi_k)}{6n^2} \right)$$

ここで、 ξ_k は各区間に対応する適当な値です。 $n \rightarrow \infty$ の極限では、 $\frac{f''(\xi_k)}{6n^2}$ の項は消えます。

したがって、

$$\lim_{n \rightarrow \infty} n \left(\int_0^1 f(x) dx - \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \right) = \lim_{n \rightarrow \infty} \frac{1}{2n} \sum_{k=0}^{n-1} f'\left(\frac{k}{n}\right)$$

f' は連続関数なので、リーマン和の定義より、

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f'\left(\frac{k}{n}\right) = \int_0^1 f'(x) dx = f(1) - f(0)$$

したがって、

$$\lim_{n \rightarrow \infty} n \left(\int_0^1 f(x) dx - \frac{1}{n} \sum_{k=0}^{n-1} f\left(\frac{k}{n}\right) \right) = \frac{1}{2} (f(1) - f(0))$$

大きな論理的破綻なく正答一出題・採点担当者曰く「満点を与え得る」答案

人間のチャンピオンを制してきたAIの歴史

- ・ 古くからAI研究が盛んなゲーム：既にAIが人間のチャンピオンを制している
- ・ 近年、**数学などの分野でも人間のチャンピオンに肉薄する実力が出つつある**
Cf. 2024/07、AlphaProof/Geometry2（DeepMind）がIMO銀メダル級記録

年	種目	当時の人間側チャンピオン	チャンピオンを制したAI
1996	チェス	ガルリ・カスパロフ	Deep Blue（IBM）
2016	囲碁	イ・セドル	AlphaGo（DeepMind）
2017	将棋	佐藤天彦	Ponanza（山本一成ら）
...	...	...	...
20XX	数学	？ ？ ？	？ ？ ？

アカデミック分野でも、人間側最高峰を制するのは時間の問題？

素朴な疑問

なぜ、われわれは既に
人間を凌駕するといっても過言ではない
すごいAIを手に行しているのに、
われわれの業務は
すぐに大きく変わらないのか？

2

業務への生成AI導入のコツ ——導入を進めるのが難しい理由と対策



ゲームも、数学も、「ルール」が決まっている

- 数学は形式化すれば、ルールが厳密に決められる（「ゲーム」に似ている）
Cf. Hilbert らの形式主義 → 数学基礎論・数理論理学に基づく形式体系
- ルールが厳密な問題の探索は、実はAIの一丁目一番地
→ Alpha Proof は、**定理証明系 LEAN** に基づく**強化学習**を採用

```
-- Gödel's Second Incompleteness Theorem--
```

```
theorem goedel_second_incompleteness [System.Consistent T] : T  $\nvdash$   $\uparrow$ Con := fun h  $\mapsto$   
  goedel_unprovable T <| and_left! (consistent_iff_godel T)  $\odot$  h
```

```
theorem inconsistent_unprovable [N  $\vdash_m^*$  T] : T  $\nvdash$   $\sim\uparrow$ Con := fun h  $\mapsto$   
  not_godel_unprovable T <| contra $_o$ ! (and_right! (consistent_iff_godel T))  $\odot$  h
```

```
theorem inconsistent_undecidable [N  $\vdash_m^*$  T] : System.Undecidable T  $\uparrow$ Con := by  
  haveI : Consistent T := Sound.consistent_of_satisfiable <_, (inferInst  
  constructor  
  • exact goedel_second_incompleteness T  
  • exact inconsistent_unprovable T
```

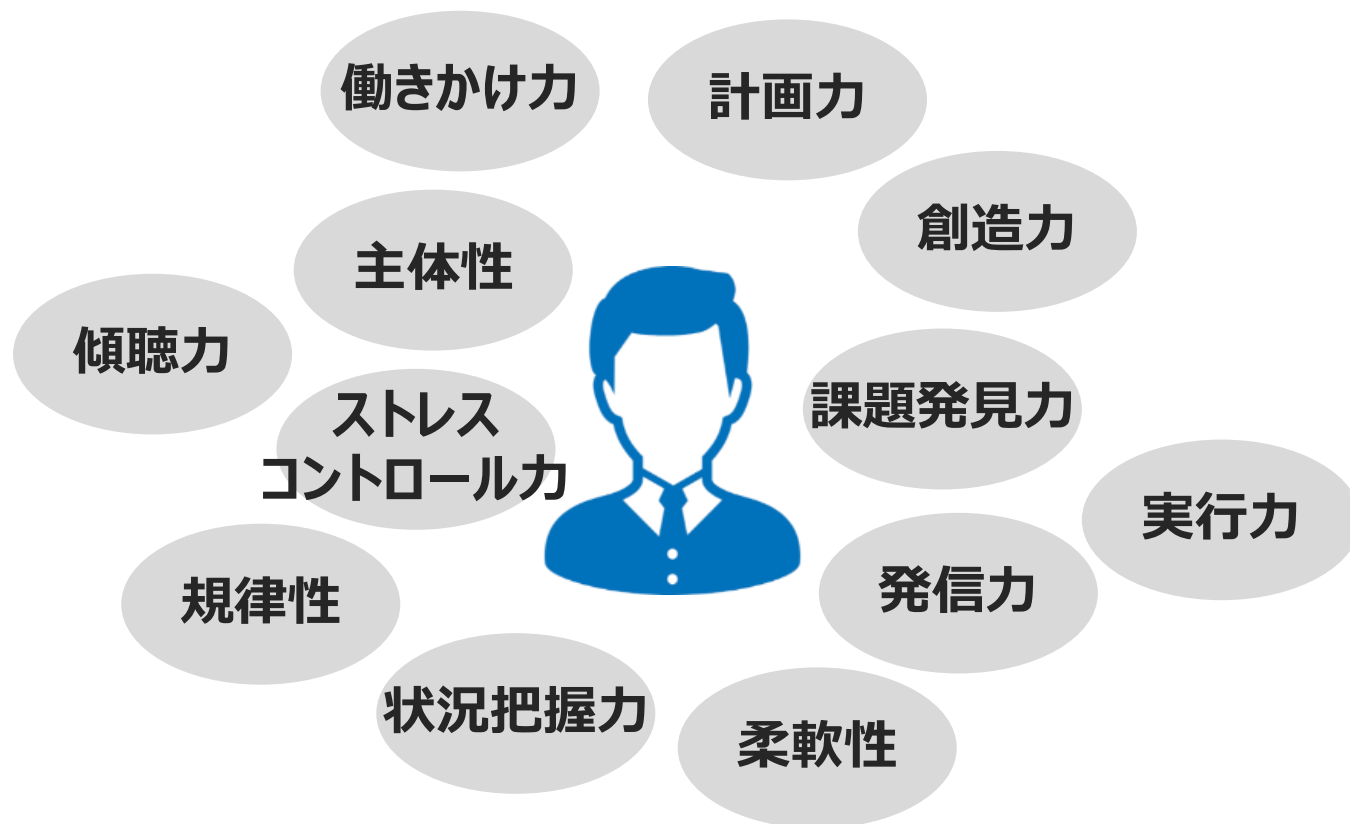
証明をソースコードのように書くとその
正しさを機械的に判定してくれる
→ 強化学習により探索戦略を習得

一方、われわれの業務は... ?

<https://github.com/FormalizedFormalLogic/Incompleteness/blob/622ebb7e95cbccfab6e722b3194335380f54a29/Incompleteness/Arith/Second.lean#L184-L185> より引用

一方、我々の業務は…？

- われわれは、専門知識以外にも多くのジェネリックスキルを持っている
→ ルールが厳密でない業務でも、少ない情報から「空気を読んで」遂行できる
- ジェネリックスキルのなかには、現在のAIにとって獲得が困難なものも



図は、経済産業省による「社会人基礎力」 (<https://www.meti.go.jp/policy/kisoryoku/index.html>, 2006年) を参考にしてNTTデータ数理システムが作成した。図中に挙げた12の能力要素は、「社会人基礎力」における12の能力要素である。

生成AIの業務活用を進めるのが難しい理由

理由

**そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから**

**生成AIがハルシネーションという
大きな問題を抱えているから**

**生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから**

生成AIの業務活用を進めるのが難しい理由——とその対策

理由

そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから

生成AIがハルシネーションという
大きな問題を抱えているから

生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから

対策

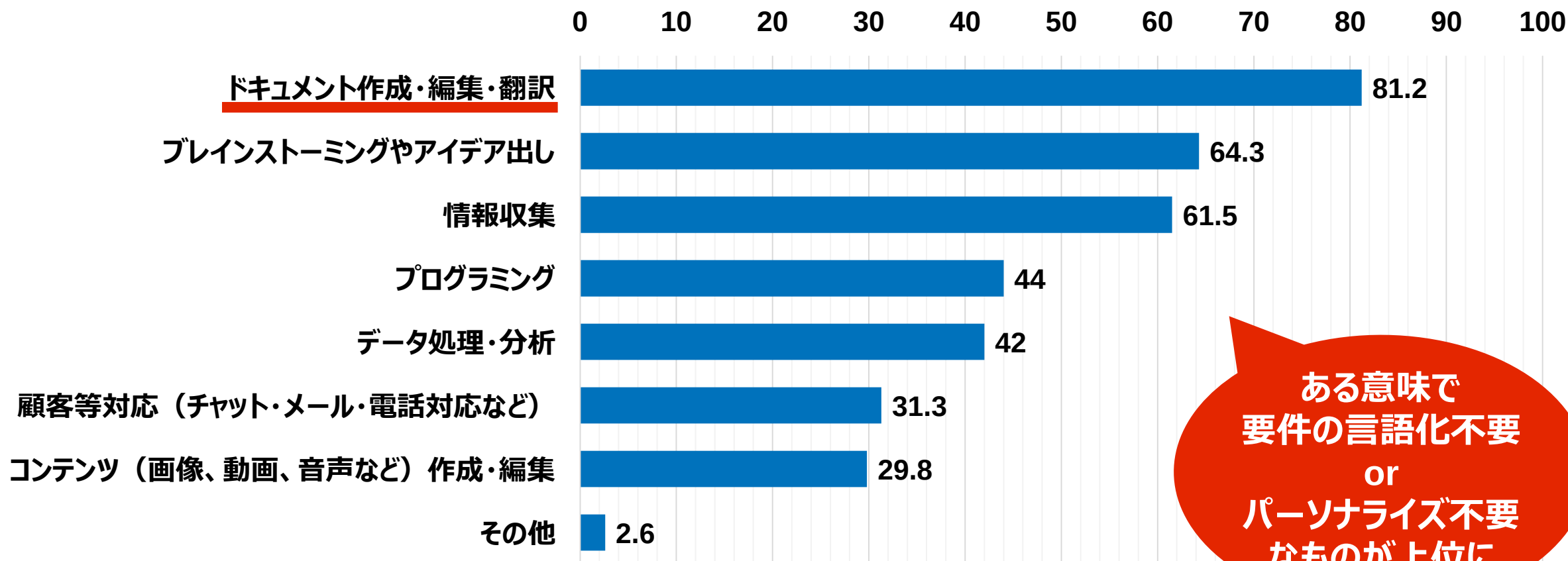
それでも、業務について
頑張って言語化をする

結局は言語化が重要

- 生成AIの側に立ってみれば、彼らは「私／私たち」が誰なのかを知り得ないし、そこからうまく「空気を読む」手段もない
→ 目前にどんな人がいても良いように、無難で平均的な解答をするしかない
- 上手く「私／私たち」にあわせて性能を引き出すには、「私／私たち」自身について丁寧な言語化が重要
→ ところがこれは（人間側に課せられた課題が）難しいので、進展が遅い
- おそらくこの状況は、現在の延長で技術が進展しても大きくは変わらない？

生成AIを導入する目的 (IPA DX動向2024 より)

要件の言語化 or パーソナライズをせずとも、手軽に効果を見込めそうなものが上位に



ある意味で
要件の言語化不要
or
パーソナライズ不要
なものが上位に

IPA 編『DX動向2024—求められる成果と変革』データ集「2-20. 生成AIを導入する目的」より引用（図は株式会社NTTデータ数理システムが再作成した）
<https://www.ipa.go.jp/digital/chousa/dx-trend/dx-trend-2024.html>

なんの工夫もしないと…

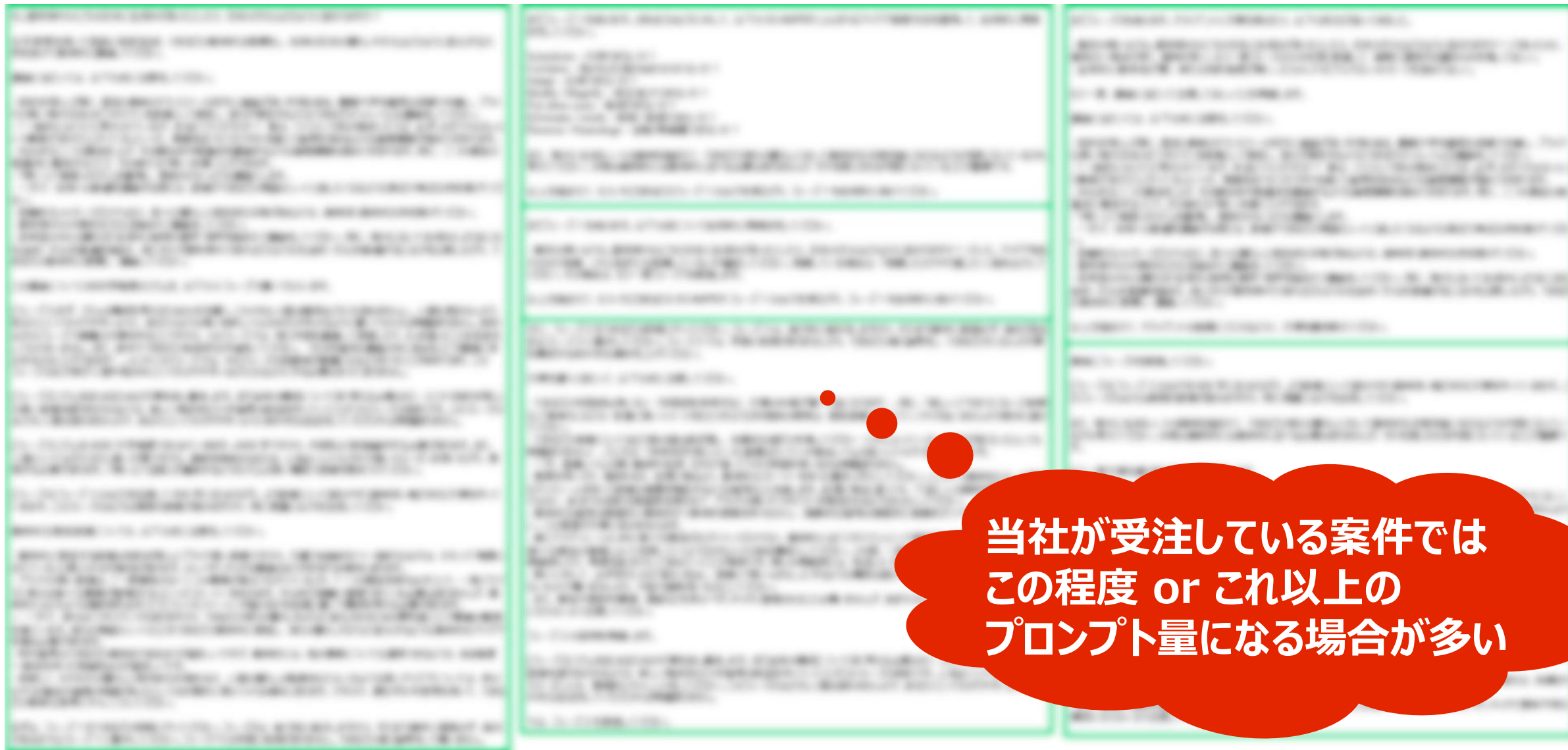
もし室町時代のころの日本に生成AIがあったとしたら、日本の文化はどのように変わりますか？ 要点を1文ごとの箇条書きで、5つにまとめてください。

1. 文学や詩の創作が自動化され、多くの人々がアクセスしやすくなり、文化の普及が加速するでしょう。
2. 絵画や工芸のデザインが高度に発展し、多種多様な創作物が生まれる可能性が高まります。
3. 政治や軍事における戦略の立案が効率化し、戦国大名や将軍の意思決定がより迅速で的確になるかもしれません。
4. 教育方法が革新され、知識や技術の伝達が高度化し、学びの格差が縮まるでしょう。
5. 人々の生活における日常的な問題解決が容易になり、生活の質が向上する可能性があります。

大雑把で
示唆がない…

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。

想定読者や、読者に与えたい印象をかなり細かく書くと…



当社が受注している案件では
この程度 or これ以上の
プロンプト量になる場合が多い

1. 室町時代に生成AIが存在していた場合、茶道では**茶室の設計や茶器の選定が科学的に最適化**され、新たな美学が生まれた可能性がある。
2. 能楽では**古典文学を基に生成AIが新しい台本を作成し**、観客に革新的な芸術体験を提供することができた。
3. 書道において**生成AIは仮想の師匠として機能し、効率的な技術習得と個性を引き立てる指導を実現し**、新しい書道の可能性を広げる。
4. 生成AIは文化の**均質化の懸念**とともに、多様な選択肢を提供することで新しい文化形成の機会を創出し、**情報の民主化**を進める。
5. 生成AIは**文化の失われた要素を取り戻し**、技術と文化の相互作用を通じて未来に向けた革新を支える重要なツールである。

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。青字・赤字による強調は NTT データ数理システムによる。

生成AIの業務活用を進めるのが難しい理由——とその対策

理由

そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから

生成AIがハルシネーションという
大きな問題を抱えているから

生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから

対策

それでも、業務について
頑張って言語化をする

ハルシネーション（幻覚）

- 生成AIは、**知らないことを知らないと言えず、無理に答えを捻り出す**
→ 人間ではとうてい犯さないような、特徴的な誤りを犯してしまう

NTTデータ数理システムの社長の名前
は何ですか？



NTTデータ数理システムの社長（代表取締役）の名前は、「長尾 健（ながお たけし）」です。

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。

NTTデータ数理システムの社長の名前
は何ですか？



NTTデータ数理システムの社長は、2023年10月時点では
清水 祐亮（しみず ゆうすけ）氏です。

NTTデータ数理システムの社長の名前
は何ですか？



NTTデータ数理システムの社長の名前は、安達 晋士（あ
だち しんじ）氏です。

「NTTデータ数理システム」の社長を GPT-4o に10回聞いた結果

氏名（敬称略）	頻度	判定
長尾 健	1	不正解
岩沢 大輔	1	不正解
小柴 俊治	1	不正解
堤 信夫	1	不正解
山口 宗一郎	1	不正解
安達 晋士	1	不正解
小川 徹	1	不正解
中村 充孝	1	不正解
清水 祐亮	1	不正解
わからない／公式情報を確認すべき	1	

正解



株式会社NTTデータ数理システム

代表取締役社長 箱守 聡

幻覚を回避できたのは
10回中1回

生成AIの業務活用が難しい理由——とその対策

理由

そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから

生成AIがハルシネーションという
大きな問題を抱えているから

生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから

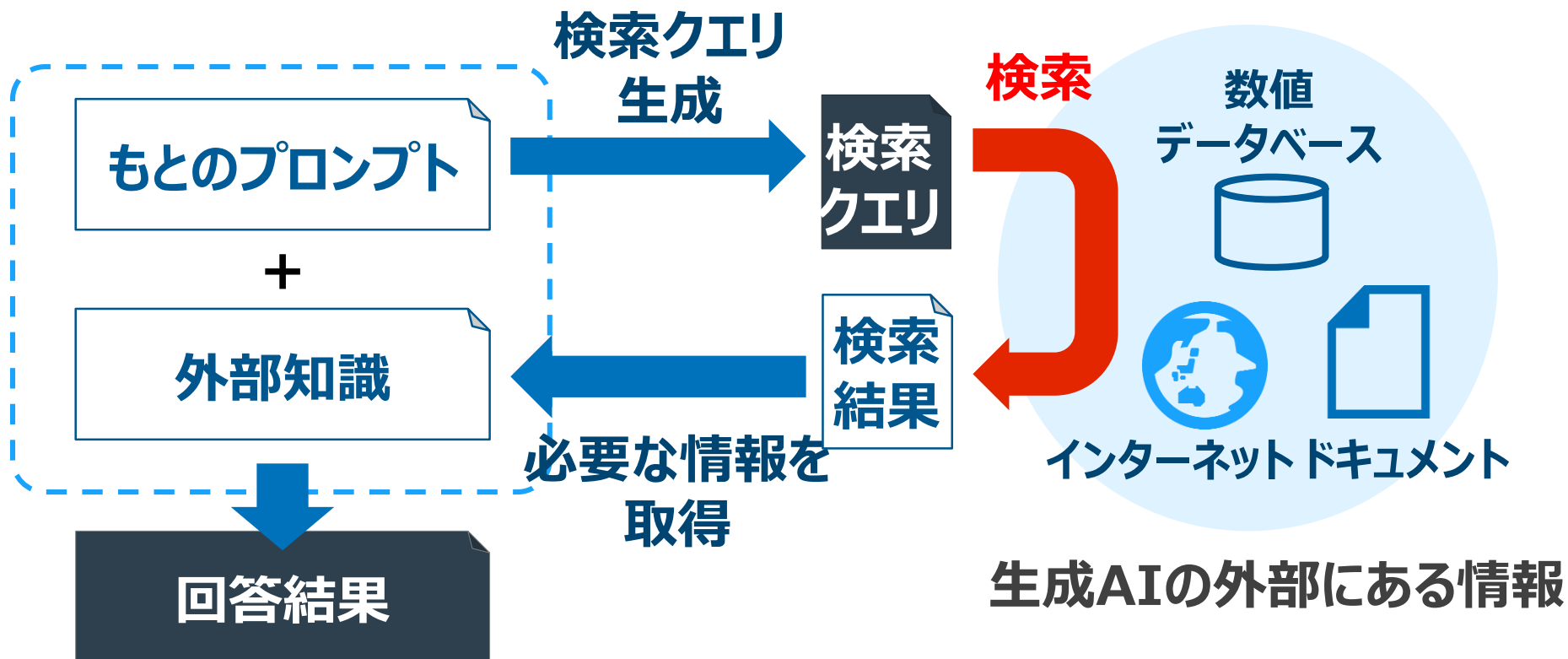
対策

それでも、業務について
頑張って言語化をする

ハルシネーションの原因に
よって対策は異なるが、
たとえば RAG が有効

RAG; Retrieval-Augmented Generation

- 生成AIと、生成AIの外部にある情報検索を組み合わせることで回答生成
→ **生成AIが知らない知識（社内ナレッジや最新時事）に基づく回答が可能**
- 社内にある多くのデータを活用できる可能性がある



インターネット検索と組み合わせれば…

NTTデータ数理システムの社長の名前は何ですか？



4 件のサイトを検索しました ✓

株式会社NTTデータ数理システムの代表取締役社長は、箱守聰（はこもり さとし）さんです。彼は2016年に同社の社長に就任し、数理科学技術とコンピュータサイエンスを基盤に、企業や社会のさまざまな問題解決に取り組んでいます(MSI)(NTT).

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。赤下線による強調は NTT データ数理システムによる。

製品マニュアルと組み合わせれば…

製品についての質問に答えてくれるチャットボットに

The screenshot displays the S4 Simulation System interface. On the left, a sidebar menu lists various statistical distributions under the heading '19.1 乱数属性' (Random Number Attributes). The main content area shows the '19.1.8 アーラン分布' (Erlang Distribution) section, which includes text about generating random numbers following the Erlang distribution and specifying parameters like shape (α) and scale (β). On the right, a 'Q&A Chat' window shows a conversation. The chat bot greets the user and explains that conversation content is used for service improvement. A user query asks for instructions on generating Erlang random numbers, and the chat bot responds that it is possible and points to the relevant manual section (19.1.8 Erlang Distribution).

操作マニュアル チュートリアル psim 言語

297 (298 / 357) 130%

19.1.8 アーラン分布

アーラン分布に従う乱数を返します。

正整数

アーラン分布の形状パラメータを指定します。

β

アーラン分布の尺度パラメータを指定します。

297

19.1 乱数属性

Q&A Chat

ようこそ「S4 QA Chat サービス」へ！

会話内容はサービス向上のため利用されます。個人情報の入力をご遠慮ください。

誤った回答をする場合もあるため、マニュアルと併読することをお勧めします。

アーラン分布にしたがう乱数を生成できるか教えてください。マニュアルの該当箇所も教えてください。

はい、アーラン分布に従う乱数を生成することができます。マニュアルの該当箇所は「19.1.8 アーラン分布」です。

当社パッケージソフトウェアS⁴ Simulation Systemの使い方を質問することができる質問応答チャットアプリを公開

技術的詳細は、当社オウンドメディアMSIISMにて公開中
<https://www.msiism.jp/article/llm-qa-app.html>

幅広く浸透してきたが、まだまだ奥が深い RAG

- 単純なRAGは手軽に構築できる反面、精度を上げるのは意外と難しい
- データの種類や問題の性質によって精度向上の手段は色々
→ 生成AIと最適な検索技術を組み合わせる必要がある

RDBなどの表データも、RAGの入力に

入力の表の変換を、中間結果を作りながら繰り返すことで高精度な応答を実現（Chain-of-table）

列選択、ソート、カウントなどの基本操作を生成AIが判断しながら繰り返す

表データ → 中間データ → ... → 質問に関する部分のみ残す

例 膨大な気象・地象ログデータからの情報探索

観測点番号	観測点名称	緯度	経度	震度
0	47,835 日南市油津	31.5784	131.4081	5強
1	47,829 都城市昌高原	31.7299	131.0818	5弱
2	42,519 新富町上富田	32.0689	131.4874	5弱
3	47,830 宮崎市西島	31.9384	131.4142	5弱
4	41,515 鹿屋市新栄町	31.376	130.8504	5弱
5	47,827 鹿児島市東都元	31.5552	130.5486	4
6	42,517 佐伯市瀬江瀬江浦	32.8009	131.9216	4
7	42,512 戸北町戸北	32.299	130.4931	4

震度が5弱の観測点を列挙してください

都城市昌高原, 新富町上富田, 宮崎市西島, 鹿屋市新栄町

質問を理解したうえで、「震度」列を選択してフィルタするクエリを実行

応用先：POSデータなど、各種表データに基づくQAシステム

プログラムのソースコードを自動解析、RAGに組み込む

レポジトリ全体を解析、コールグラフに基づきドキュメントを生成
ソースコードを読まなければ答えられない難しいQAも可能に

例 ソースコードに基づきトラブルシュートするQAチャット

強化学習モデル作成時に "reward is not implemented." というエラーが発生しました。どのように解決すればよいですか。

"reward is not implemented." というエラーは、RLModel クラスのサブクラスで reward メソッドが実装されていないために発生しています。RLModel クラスは抽象クラスのようなものであり、具体的な強化学習モデルを作成する際には、以下のメソッドをオーバーライドする必要があります：

```
1. reward(self, simulator, item, prev_time)
2. observation(self, simulator, item)
3. action(self, simulator, obs, act, item)
```

クラス構造をコードから読み取りエラー原因を特定

これらのメソッドを具体的な強化学習モデルに応じて実装する必要があります。以下に、これらのメソッドを実装する例を示します。

応用先：ソフトウェア開発・サポート業務の効率化

ナレッジグラフを利用して、RAGの回答正確性向上

製品マニュアルや手順書などの文書群から、概念の関係を抽出
関係性をグラフで表現して蓄積することで、回答の正確性が向上

文書群（文書A, B, C, ...）

抽出

数理システム ← made by 文書A

文書B → has feature Auto Modeling

Auto Modeling → has feature 線形回帰

線形回帰 → has feature 文書C

文書C → has feature 形態素解析

形態素解析 → has feature Alkano

Alkano → has feature 数理システム

ナレッジグラフを参照した質問応答例

Q. Alkanoにはどのような機能がある？

A. Alkano には、AutoModeling [文書B]、線形回帰 [文書C] などの機能があります。

通常のRAGが苦手な「複数文書を正確に踏まえて答えるべき」質問でも、正確な回答を生成できる

応用先：医療、法律など正確性や根拠が特に要求される分野

生成AIの業務活用を進めるのが難しい理由——とその対策

理由

そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから

生成AIがハルシネーションという
大きな問題を抱えているから

生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから

対策

それでも、業務について
頑張って言語化をする

ハルシネーションの原因に
よって対策は異なるが、
たとえば RAG が有効

生成AIの業務活用を進めるのが難しい理由——とその対策

理由

そもそも、われわれの業務には
言語化しにくい複雑な部分が多いから

生成AIがハルシネーションという
大きな問題を抱えているから

生成AIを使いこなすには
仕組みをある程度把握したうえで
プロンプトエンジニアリングが必要だから

対策

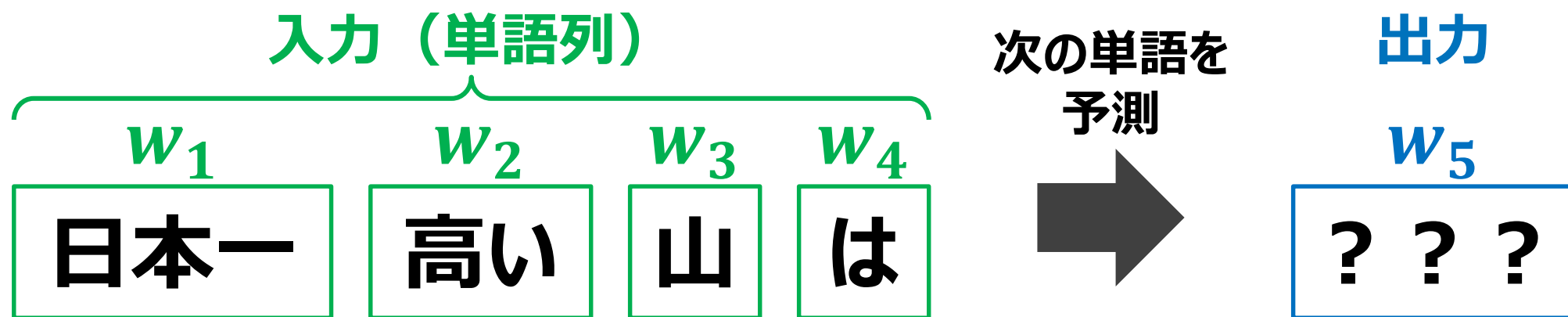
それでも、業務について
頑張って言語化をする

ハルシネーションの原因に
よって対策は異なるが、
たとえば RAG が有効

仕組みと技法を
学ぶ

文章生成の仕組みは、「予測」の繰り返し

- 「既に生成した単語列から、次の単語を予測する」を繰り返して文章生成
- 予測のルールを確率分布でモデリング（言語モデル）



条件付確率分布 $P(w_5 | w_1, w_2, w_3, w_4)$
次にくる確率が高い単語を求める

富士山	3%
エベレスト	1%
⋮	
山本	0.1%

※実際に使われているLLMでは、chunkとよばれるもっと細かい単位で区切ることが多い

例えば——その原理を理解していれば直観的にわかること

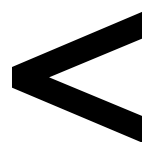
単純に考えれば、ユーザの入力 + AIが既に生成した文章が多くて正確なほど、次の単語の予測は正確になる

→ **結論先取ではなく、思考過程先取で話させたほうが良い**

結論先取で答えさせる

答えは です。

その根拠は、



思考過程先取で答えさせる

なので、答えは です。

例題：歩行時間見積もり問題

問題

20代男性が、立川駅から東京駅まで中央線沿いを歩きます。
途中、各駅に立ち寄り記念撮影をします。所要時間を推計してください。

GPT4o による見積もり結果（10回実験した平均）

結論先取で答えさせる

思考過程先取で答えさせる

VS.

例題：歩行時間見積もり問題

問題

20代男性が、立川駅から東京駅まで中央線沿いを歩きます。
途中、各駅に立ち寄り記念撮影をします。所要時間を推計してください。

GPT4o による見積もり結果（10回実験した平均）

結論先取で答えさせる

平均

4時間54分

（標準偏差：73.23分）

VS.

思考過程先取で答えさせる

平均

10時間04分

（標準偏差：87.48分）

※講演者が実際に同条件で歩いた実績値は（食事など除き）**12時間程度**

結論先取の場合の回答例（5時間と回答）



推計所要時間は**5時間**です。

見積もりの根拠

1. **距離:** 立川駅から東京駅までの距離は約40kmです。
2. **歩行速度:** 一般的な徒歩の平均速度は時速約5kmです。この速度で歩くと、約8時間かかりますが、写真撮影のための時間を考慮する必要があります。
3. **撮影時間:** 各駅での写真撮影や休憩のために、駅ごとに約5分程度を要すると仮定します。中央線の主要駅は8駅程度あり、その合計で約40分。
4. **総合的な時間:** 歩行に要する時間（約8時間）から撮影時間を引いて、最終的な所要時間を5時間と見積もりました。

結論先取の場合の回答例——結論を出すまで



推計所要時間は

**「5時間」という結論を出すまでに
できる推論がほとんどない**

まず無思考で「5時間」と宣言



推計所要時間は 5時間 です。

見積もりの根拠

距離・歩行速度の事実認識はおおむね正しく、
妥当な値（8時間）は出せているが...

1. **距離:** 立川駅から東京駅までの距離は約40kmです。
2. **歩行速度:** 一般的な徒歩の平均速度は時速約5kmです。この速度で歩くと、約8時間かかりますが、写真撮影のための時間を考慮する必要があります。
3. **撮影時間:** 各駅での写真撮影や休憩のために、駅ごとに約5分程度を要すると仮定します。中央線の主要駅は8駅程度あり、その合計で約40分。
4. **総合的な時間:** 歩行に要する時間（約8時間）から撮影時間を引いて、最終的な所要時間を5時間と見積もりました。

無思考で考えた「5時間」と
つじつまを合わせようとして破綻している

思考過程先取の場合（10時間と回答）



立川駅から東京駅まで徒歩で移動する際の推定時間を算出するためには、以下の要素を考慮します。

1. 距離

- 立川駅から東京駅までの距離: 約40km。具体的には、中央線を利用した場合の最短距離はおおよそこの範囲になります。

2. 歩行速度

- 平均的な歩行速度: 一般的な成人の歩行速度は、時速約4～5kmです。ここでは4.5km/hを基準にします。

3. 時間計算

- 必要な時間の計算: 距離（40km）を歩行速度（4.5km/h）で割ります。

$$\text{必要な時間} = \frac{\text{距離}}{\text{歩行速度}} = \frac{40 \text{ km}}{4.5 \text{ km/h}} \approx 8.89 \text{ 時間}$$

4. 各駅での写真撮影

- 写真撮影による時間ロス: 各駅での写真撮影には、移動時間に加え、1駅あたり平均2～3分程度かかると想定します。中央線の各駅は立川から東京までに約20駅あるため、20駅での撮影にかかる時間は以下のよう
に計算できます。

$$\text{写真撮影時間} = 20 \text{ 駅} \times 3 \text{ 分/駅} = 60 \text{ 分} = 1 \text{ 時間}$$

5. 合計時間の算出

- 合計時間: 歩行時間と写真撮影の時間を合わせます。

$$\text{合計時間} = 8.89 \text{ 時間} + 1 \text{ 時間} \approx 9.89 \text{ 時間}$$

結論

したがって、立川駅から東京駅まで徒歩で移動しながら各駅の写真を撮る場合、**約10時間**と見積もります。

「10時間」という結論を出すまで
推論をする余裕がある

NTTデータ数理システムがChatGPT（GPT-4o）を利用して生成した文章。強調は NTT データ数理システムによる。

生成AIの業務活用のコツは結局…

言語化

仕組み・特徴
の理解

業務

と

生成AI

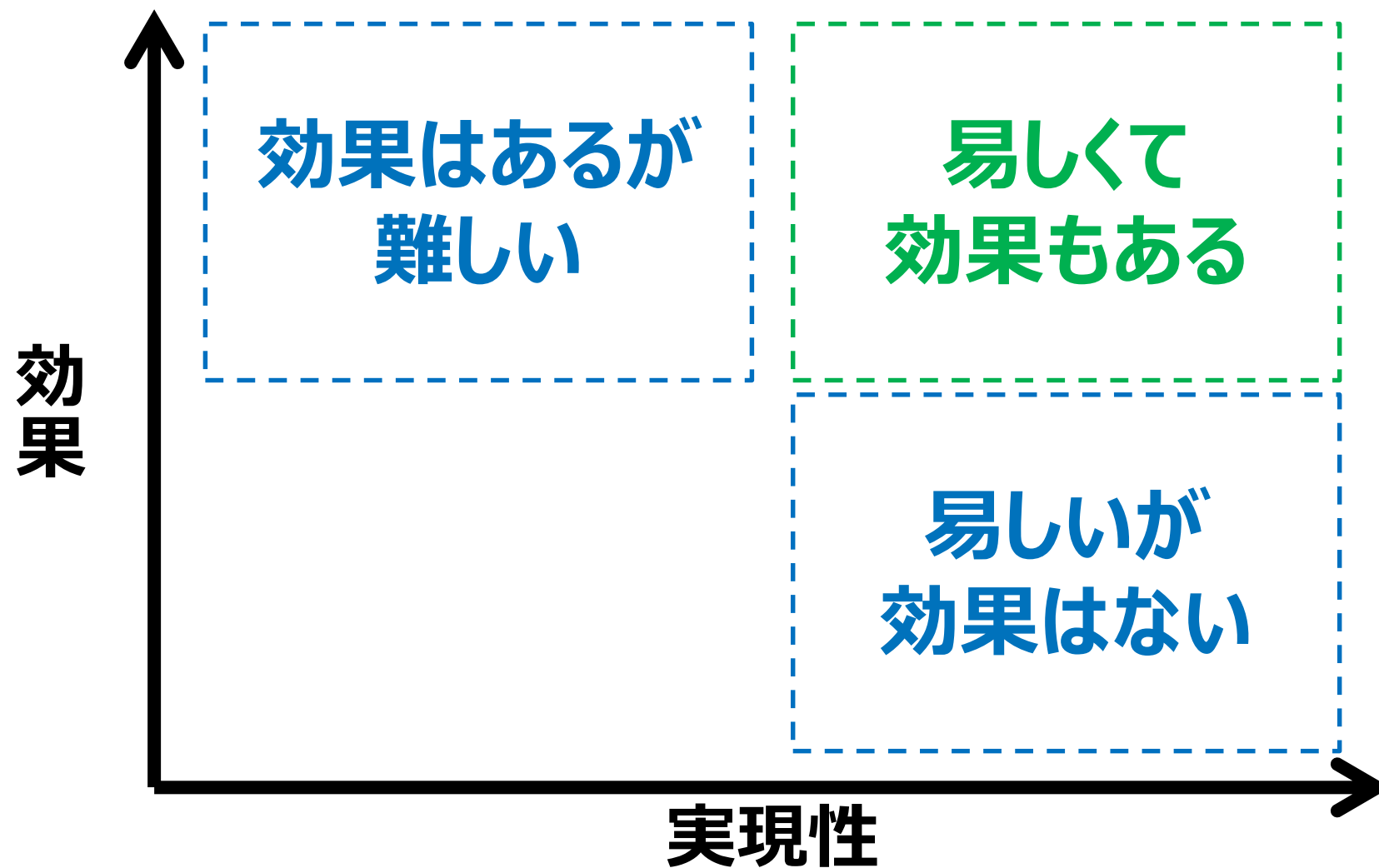
の、両方に向き合うこと

3

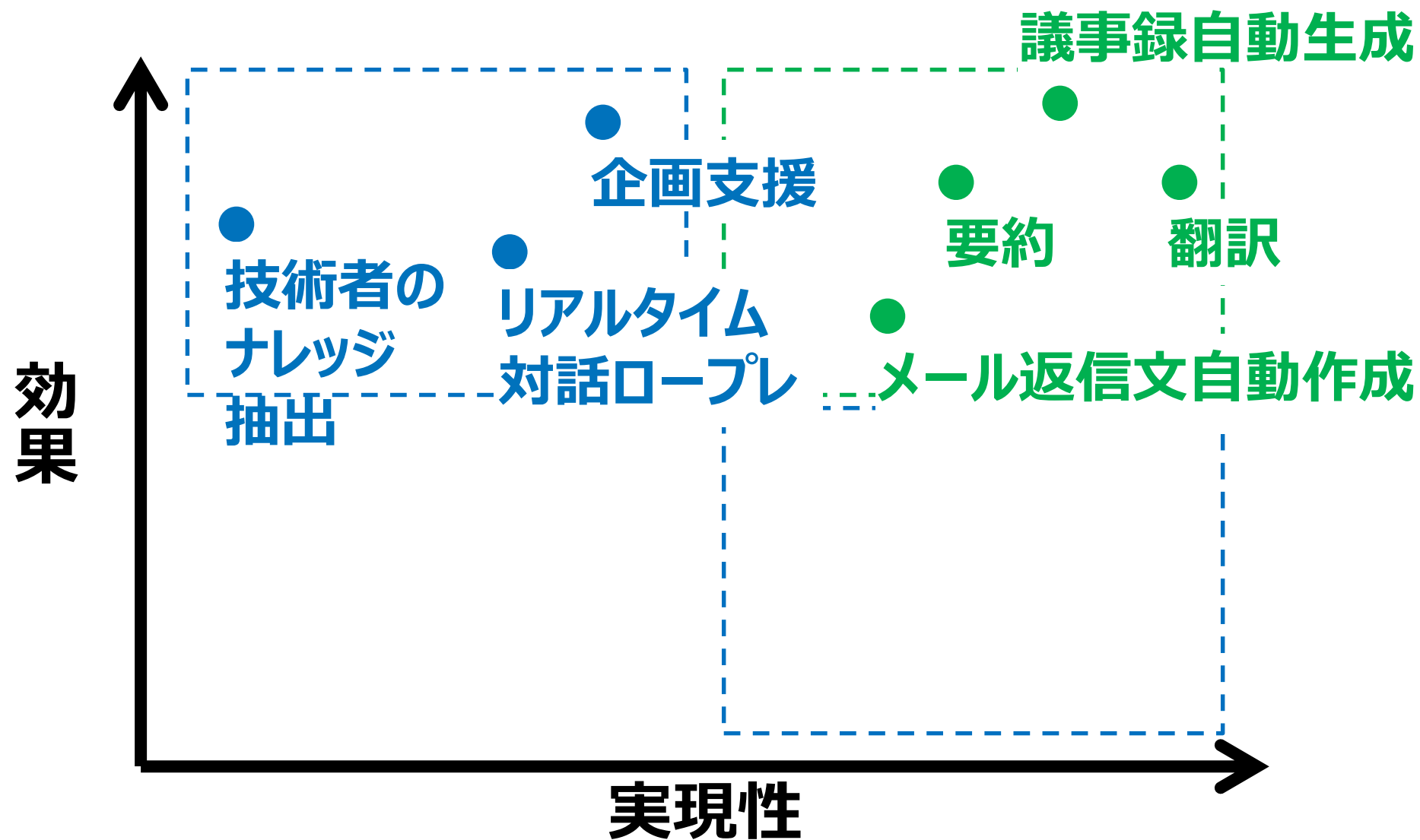
良い 生成AIプロジェクトとは？



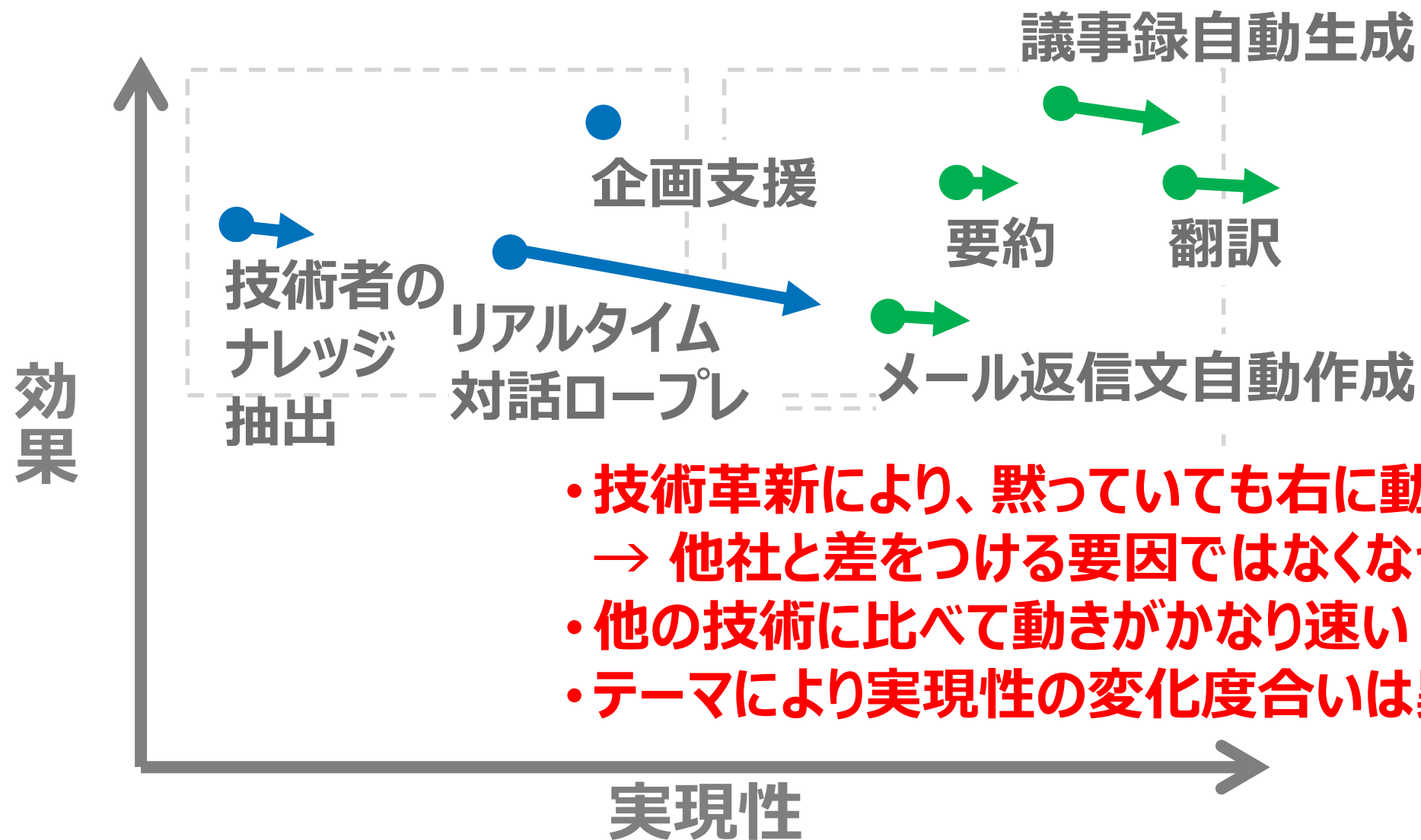
AI導入プロジェクトの整理



たとえば生成AIの場合は…



数か月単位で景色が変わる生成AI



- 技術革新により、黙っていても右に動いていく
→ 他社と差をつける要因ではなくなってくる
- 他の技術に比べて動きがかなり速い
- テーマにより実現性の変化度合いは異なる

生成AIプロジェクトは

1	簡単なので 今やるべき型	既に現在のモデルで簡単に実現できる (新規開発せずとも、既存サービスで十分な場合も多い)
2	今がんばる 価値あり型	難しいが、今頑張ってチューニングすれば実現できるかも そして、そのチューニングは技術革新があっても無駄にならない
3	今がんばると 後悔するかも型	難しいが、今頑張ってチューニングすれば実現できるかも でもそのチューニングは技術革新によって近々無駄になる
4	技術の進歩に 期待型	今頑張ってチューニングするのも難しいが モデルの改善によって近々実現する
5	人間が輝く領域	生成AIにとっては原理的に難しい、しばらくは人間が輝く領域だ

生成AIプロジェクト、5つの類型？

1	簡単なので 今やるべき型	既に現在のモデルで簡単に実現できる (新規開発せずとも、既存サービスで十分な場合も多い)
2	今がんばる 価値あり型	難しいが、今頑張ってチューニングすれば実現できるかも そして、そのチューニングは技術革新があっても無駄にならない
3	今がんばると 後悔するかも型	難しいが、今頑張って でもそのチューニング
4	技術の進歩に 期待型	今頑張ってチューニングしても難しいが モデルの改善によって近々実現する
5	人間が輝く領域	生成AIにとっては原理的に難しい、しばらくは人間が輝く領域だ

もちろん、
やるべきなのはこの2つ

生成AIプロジェクト、5つの類型？

1	簡単なので 今やるべき型	既に現在のモデルで簡単に実現できる (新規開発せずとも、既存サービスで十分な場合も多い)
2	今がんばる 価値あり型	難しいが、今頑張 そして、そのチューニ
3	今がんばると 後悔するかも型	難しいが、今頑張ってチューニングすれば実現できるかも でもそのチューニングは技術革新によって近々無駄になる
4	技術の進歩に 期待型	今頑張ってチューニングするのも難しいが モデルの改善によって近々実現する
5	人間が輝く領域	生成AIにとっては原理的に難しい、しばらくは人間が輝く領域だ

まずはここからはじめたいが
他社とは大きく差がつかない

生成AIプロジェクト、5つの類型？

1	簡単なので 今やるべき型	既に現在のモデル (新規開発せずとも、既	投資してうまく行けば 他社と差をつけられる
2	今がんばる 価値あり型	難しいが、今頑張ってチューニングすれば実現できるかも そして、そのチューニングは技術革新があっても無駄にならない	
3	今がんばると 後悔するかも型	難しいが、今頑張ってチューニングすれば実現できるかも でもそのチューニングは技術革新によって近々無駄になる	
4	技術の進歩に 期待型	今頑張って モデル	でも、これとの区別が難しい
5	人間が輝く領域	生成AIにとっては原理的に難しい、しばらくは人間が輝く領域だ	

無駄にならないとは？——ひとつの考えかた

技術革新があっても、検討の結果が活かえること

主には…

- ・ 生成AIプロジェクトを通して自分たちの業務の整理が進む
いつでも（進化していく）生成AIが受け入れられるようになる
- ・ モデルにRAG等で投入するためのデータ整理が進み、データの質が向上する
→ **検討の結果得られたプロンプトやデータが資産になり、今後も活用できる**

副次的には…

- ・ プロジェクトを通して、従業員のAIに対する理解が進む

生成AIプロジェクト、5つの類型？

1	簡単なので 今やるべき型	既に現在のモデルで簡単に実現できる (新規開発せずとも、既存サービスで十分な場合も多い)
2	今がんばる 価値あり型	難しいが、今頑張ってチューニングすれば実現できるかも そして、そのチューニングはモデルの改善によっても無駄にならない

ぜひ皆様のお悩みをお聞かせください！
われわれと一緒に「今やるべきこと」を
考えましょう！

4

おわりに ——まとめ



まとめに代えて：生成AI導入が進んだ世界で見えてくる景色

- ・ 人間と生成AI、それぞれの長所・短所を活かした分業が進む
- ・ **人間が自分たちの業務を深く理解し、言語化できている**

