NTT DATA

金融高度化セミナー「生成AI 基礎と実装 」

金融機関における生成AIの活用とその課題

2024年4月19日

NTTデータ
金融イノベーション本部
ビジネスデザイン室
イノベーションリーダーシップ統括部
統括部長　山本　英生

自己紹介

株式会社NTTデータ
金融イノベーション本部ビジネスデザイン室
イノベーションリーダーシップ統括部長

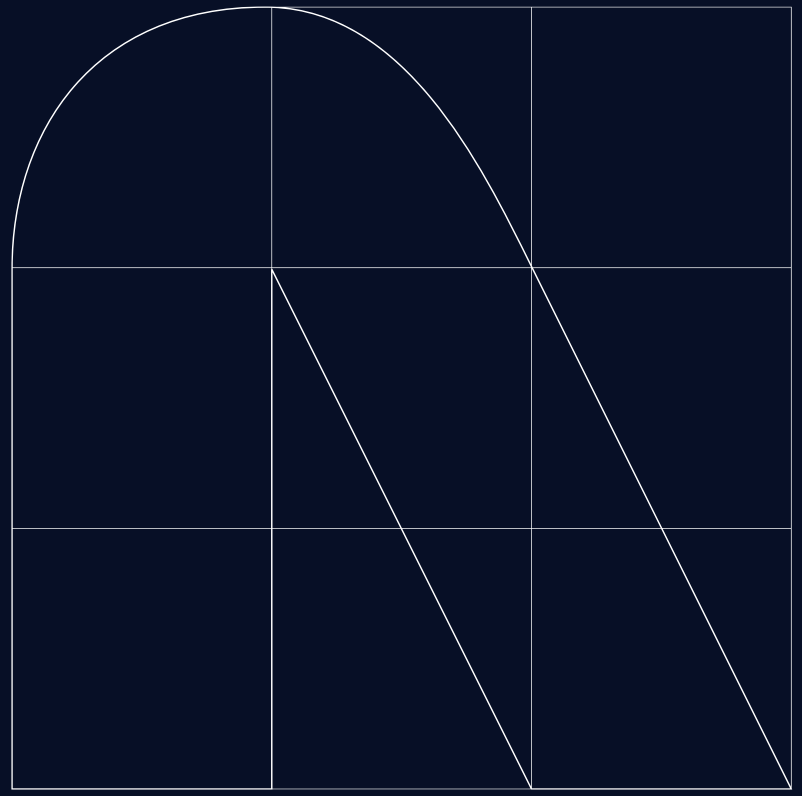
山本 英生

1996年 NTTデータ入社。
システム開発を経験後、金融機関のITグランドデザインなど多くのコンサルティング案件に従事。
現在は金融分野でのITトレンドの情報発信や、ITグランドデザイン・先進技術領域（AI、RPA、クラウド、量子コンピュータ、Web3など）のコンサルティングなど幅広く担当。
業界誌（金融財政事情等）等への寄稿・講演実績多数。
書籍「Web3と自律分散型社会が描く銀行の未来」（一般社団法人金融財政事情研究会）を執筆。

目次

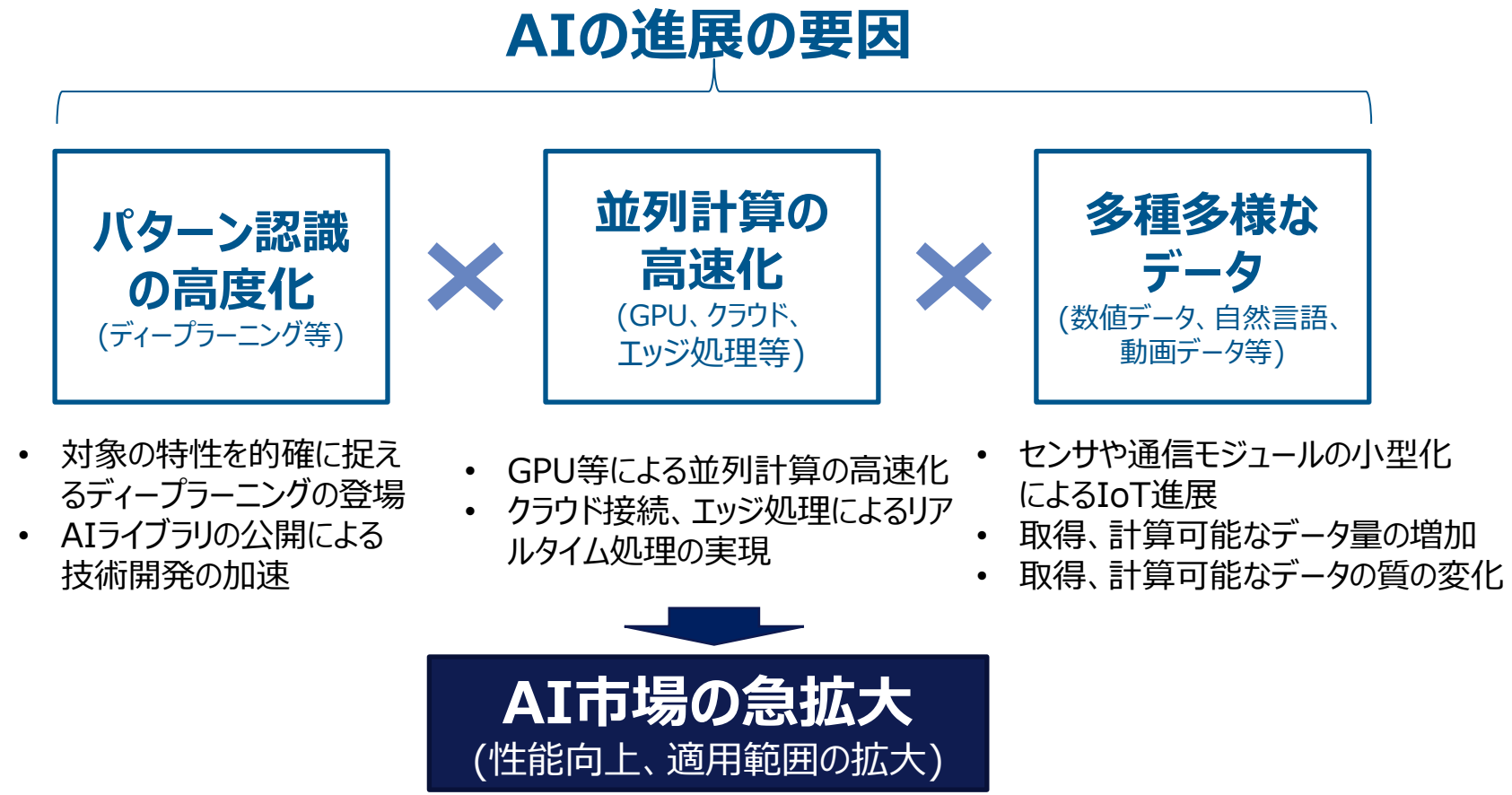
1

金融でのAIの活用



2010年代のAIの進展の要因

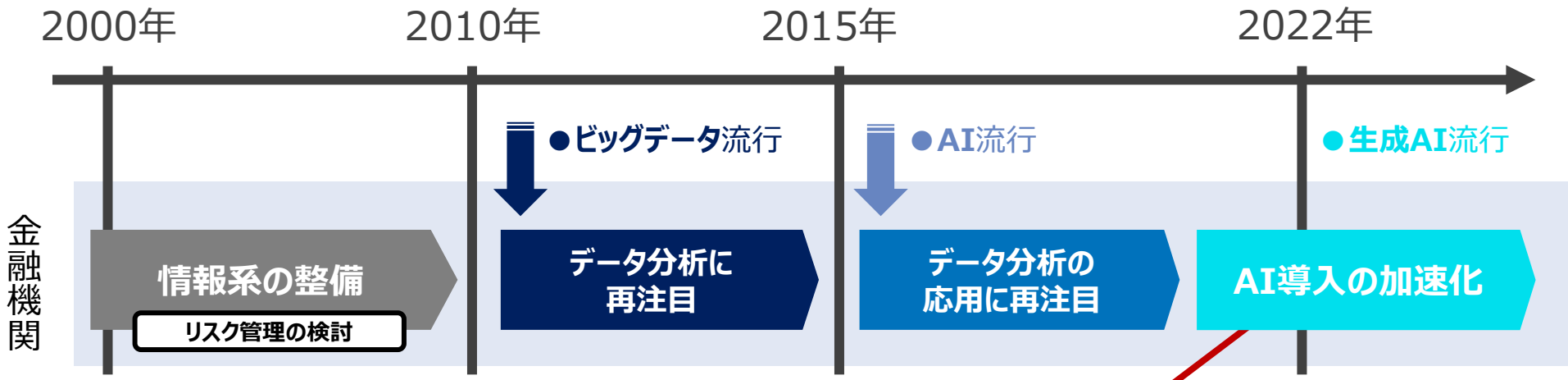
- AIの発展には、人工知能技術の発展・計算環境の進化に加え、取得、計算可能なデータ量の増加、質の変化が大きく関与



技術的なブレイクスルーを機に金融をはじめ多くの分野でAIが利用されている

金融機関におけるデジタル技術の変遷

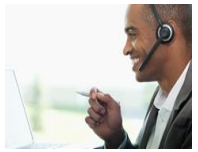
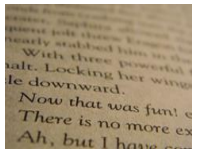
- 金融機関では情報系の整備やビッグデータの活用の検討を経て2015年頃からAI活用が活況
- 近年は生成AIの流行により、産業分野全体でAI導入が加速化している状況
- 既存のAI、生成AIを対象に、AIのリスクを適切にコントロールする「ガバナンス」への関心



- 全産業の中でも**金融・保険のAI導入率は高い** ※総務省「日本企業のAI・IoTの導入状況」等から
- 銀行業界では低金利の継続や異業種の参入等の事業環境の変化に対応すべく、**メガバンク・大手地銀を中心**に業務効率化・生産性向上を目的とした**積極的なAI導入**が進んでおり、AIの知見を蓄積している
- **導入されたAIのリスクを適切にコントロールするための「ガバナンス」への関心も高まりつつある**

金融分野におけるAIで利用されるデータ

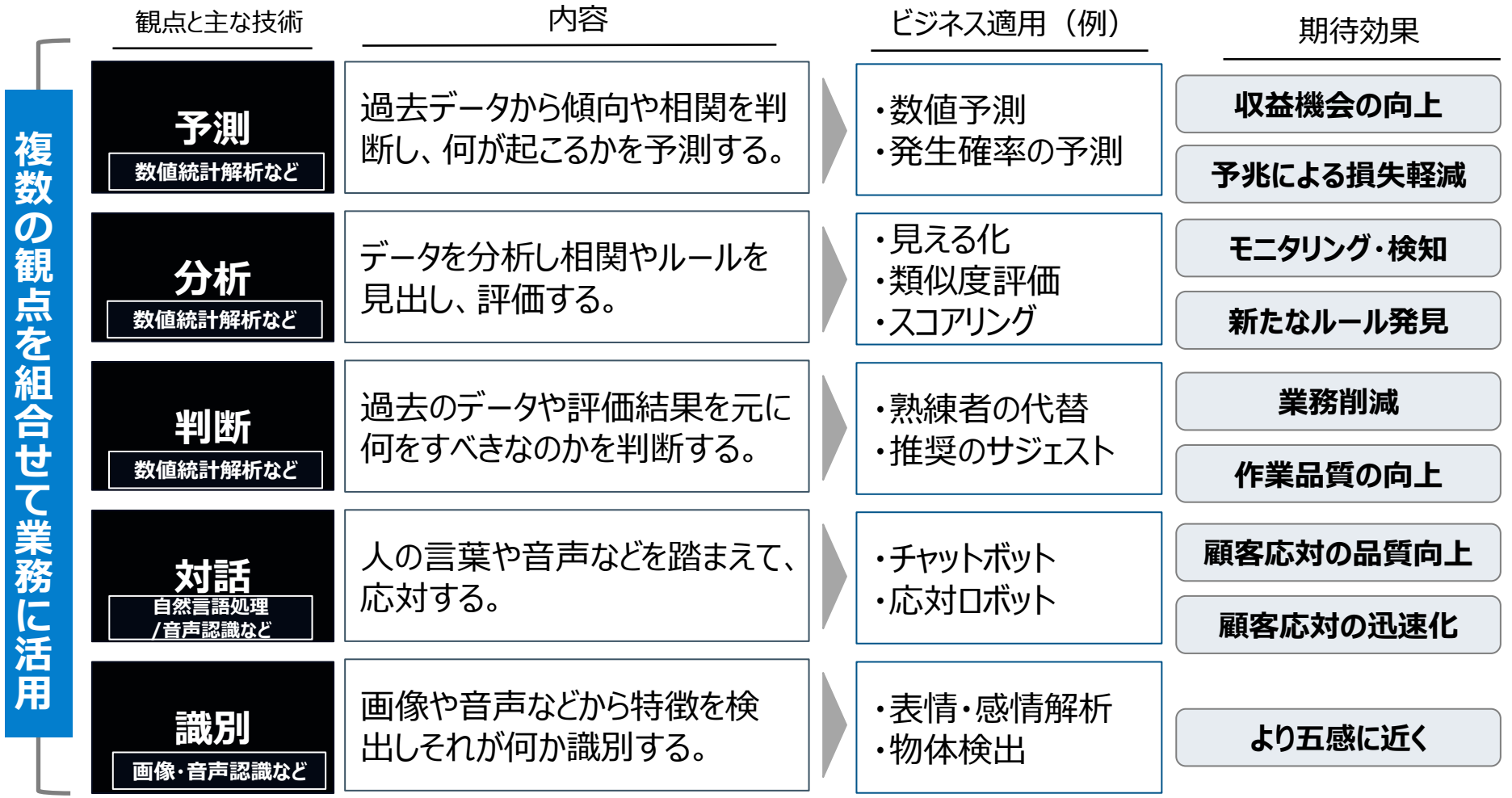
- AIは様々な角度からビジネスバリューが期待されており、複数の金融機関にて多種多様なデータを利用したAIを導入

分類		内容	金融分野での利用データ例	データ種別
数値統計分析		● 与えられたデータを分析し、**パターン**を導き出す。 ● 過去のデータを活用し、未来を**予測**する。	・口座残高 ・格付情報 ・財務諸表	構造化
画像認識		● **画像**、**動画**から**オブジェクト**や特徴を**認識・検出**する ● 大量の画像を活用し、物体やシーンを認識する。	・表情 ・建物写真 ・衛星写真	非構造化
音声認識		● 人間の会話などの**音声**を認識し、**テキスト化**や特徴を抽出する ● 音声の抑揚などから発話者の**感情**を読み取る。	・電話音声 ・打合せ音声 ・機械の動作音	非構造化
自然言語処理		● 人の**言語**や文脈を**解釈**したり、生成する。	・メール ・日報 ・会議録	非構造化
最適化		● ルール・制約の範囲内で可能な限り良い解を見つける。	・配送計画 ・シフトスケジュール ・資産ポートフォリオ	構造化/非構造化

構造化データ：あらかじめ定義され、ある定められた構造となるよう成形されたデータ。リレーショナルDBがよい例。
非構造データ：音声や画像、自然言語など使用時まで処理されないデータ。

AIに期待するビジネスバリュー

- AIは様々な角度からビジネスバリューが期待され、AIの導入が進み、企業の業務のあり方が大きく変化

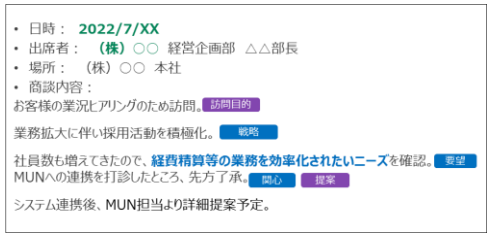


ドメイン特化型自然言語処理

- 2018年にGoogleが発表した自然言語処理技術BERTに、業務ドメインの特化を施すソリューションをNTTデータが提供
- ドメイン特化BERTによる高性能な言語モデルに加え、それらを業務に適用する技術ノウハウ・アプリケーションを具備

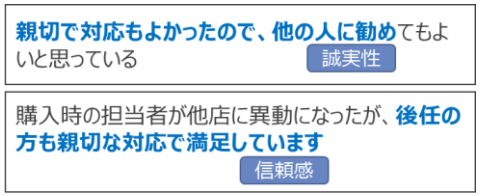
業務アプリケーション　業務に適したアプリケーション（※下記はアプリケーションの一部です）

日報解析



コンプライアンスチェック

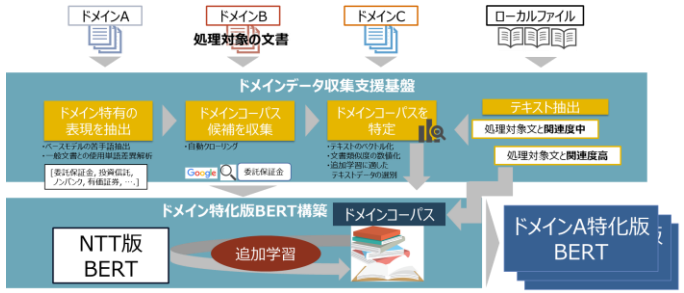
アンケート分析



ドメイン特化BERT　言語解析を支える技術

ドメイン特化BERTフレームワーク

日本語モデルの中で最大サイズのコーパスをもとに作成したNTT版BERTに加え、ドメインコーパスの追加学習が可能なフレームワーク



ドメイン特化モデル

Web上から収集した金融ドキュメントを追加学習したモデル。
外務員資格試験を解かせるとドメイン特化前より50点近く得点アップ。

一種外務員試験の問題例

文書	日経225先物には、制限値幅は定められていない。
回答	×（問題文は正しくない）

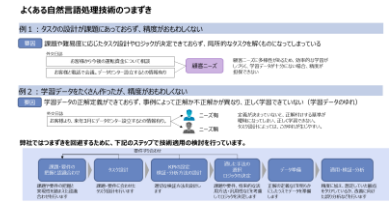
利用モデル	得点
NICT版BERTモデル	263点
ドメイン特化前	281点
ドメイン特化BERT-FW 自動構築モデル	328点

技術ノウハウ

技術検証・適用の知見

事例

検証観点

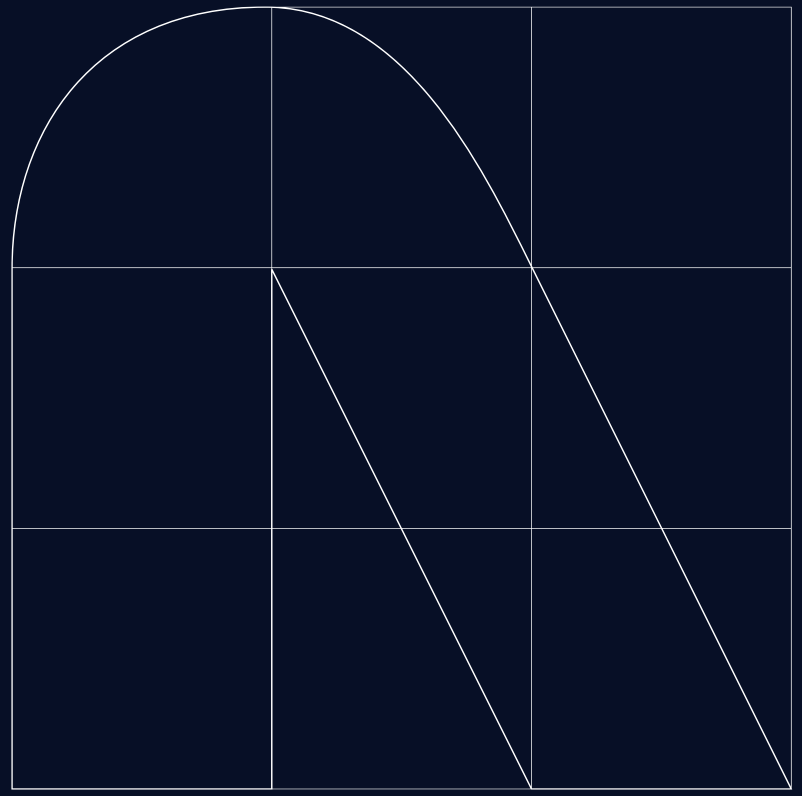


BERT技術

佐藤他(2022) BERT入門　プロ集団に学ぶ新世代の自然言語処理（AI/Data Science実務選書）

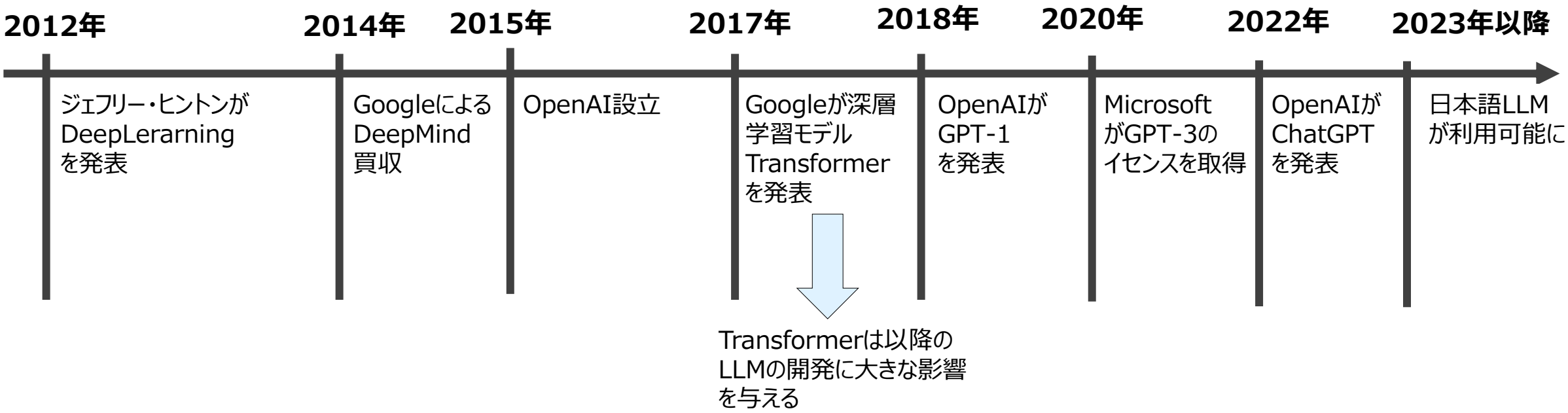
2

生成AIの活用の期待



2022年からの生成AIブームまでの流れ

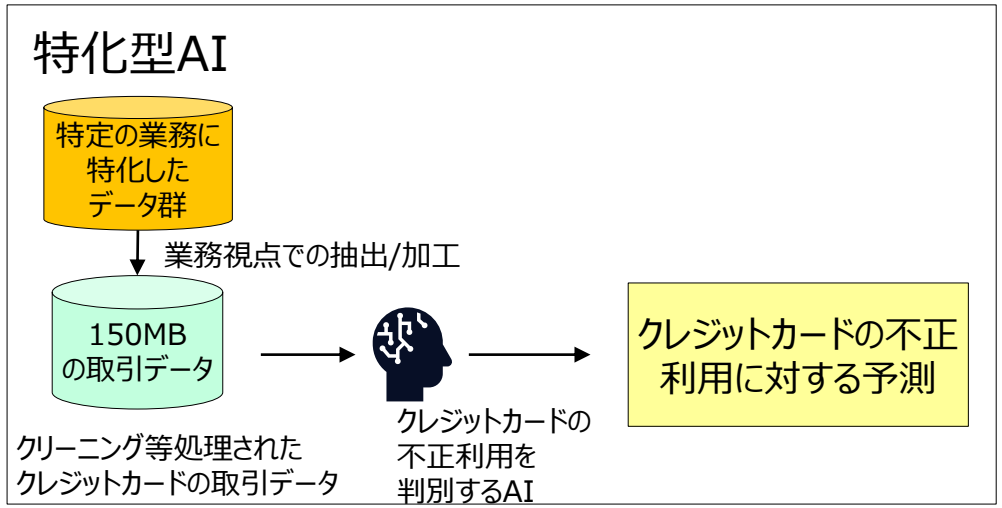
- OpenAIのサム・アルトマンが、2014年のGoogleによるDeepMindの買収にAI技術の不安や危機感を察知
- アルトマンは、汎用AI実現のため大規模言語モデル(LLM)であるGPTを開発
- MicrosoftはOpenAIと提携を模索、OpenAIのGPT-3のライセンスを2020年に取得
- OpenAIは2022年11月に、ChatGPTを発表
- 2023年に入り、日本語LLMが利用可能に

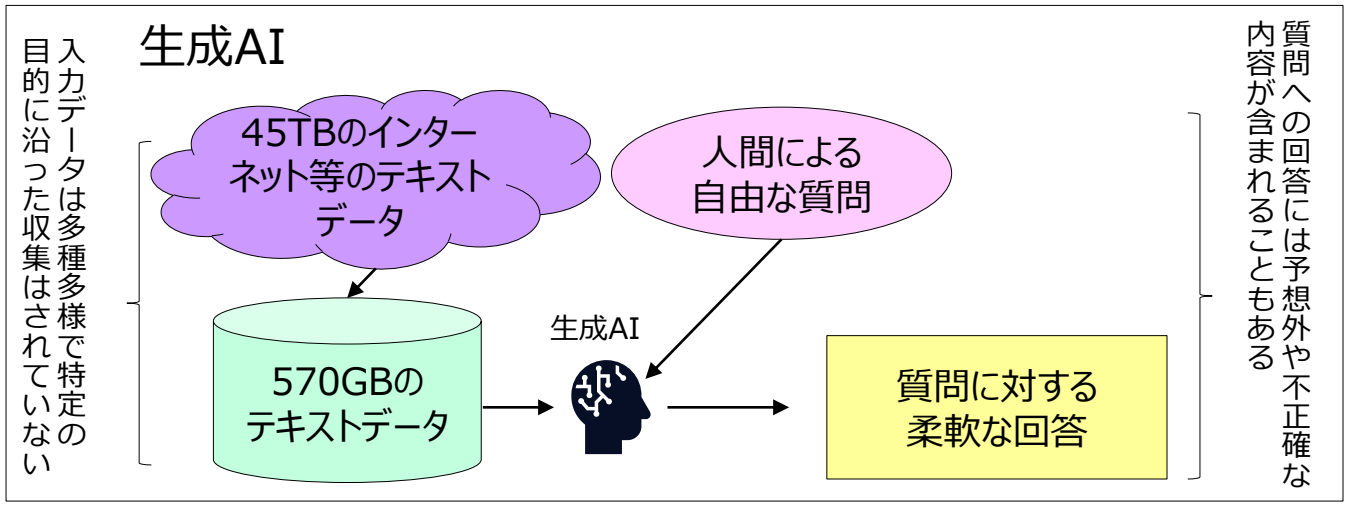


これまでのAI(特化型AI)と生成AIの違い

- **特化型AIは、AIが解決する課題にあったデータ群を抽出、加工し、適切なデータを学習**
- クレジットカードの不正利用の検出に特化したAIは、150MBのデータ量で学習し一定の精度で予測する能力を持つ
  (https://www.kaggle.com/datasets/mlg-ulb/creditcardfraudの内容から)

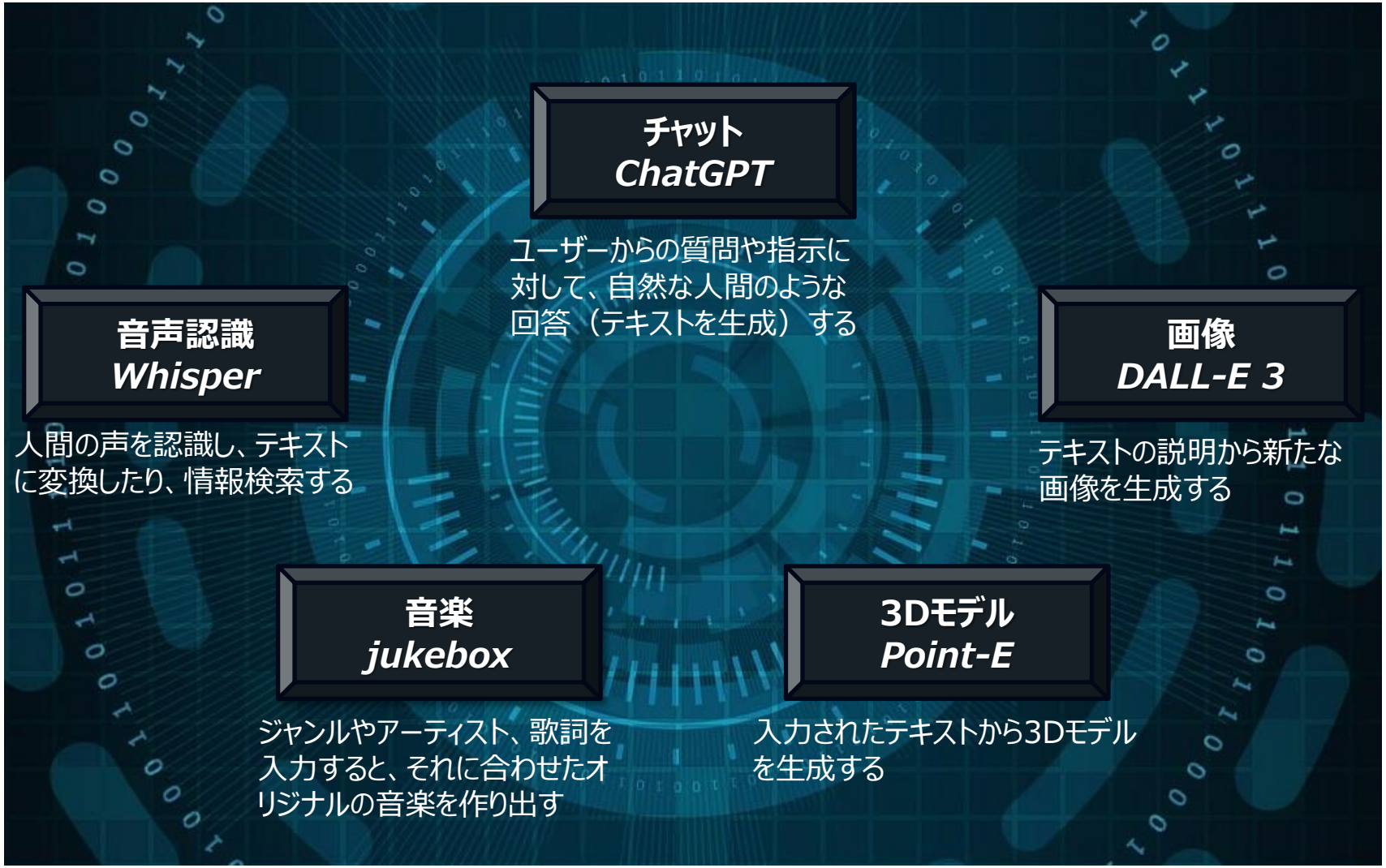
- **生成AIは、Web上の自然言語等の大量のデータを学習　(良質なテキストデータを選定する場合もある)**
  - 生成AIの中でも特に注目されているChatGPTは、大規模言語モデル(LLM)をベースに開発
  - GPT-3はWebなどの45TBのデータを前処理した570GBのテキストデータを使い学習
  - **入力データは多岐に渡り単一の目的に特化して収集されていないが、人間の自由な質問に答える能力を持つ**





OpenAIの生成AI（Generative AI）

- 最も活発なOpenAIの生成AI(Generative AI)領域は、多様なデータフォーマットをカバーしている

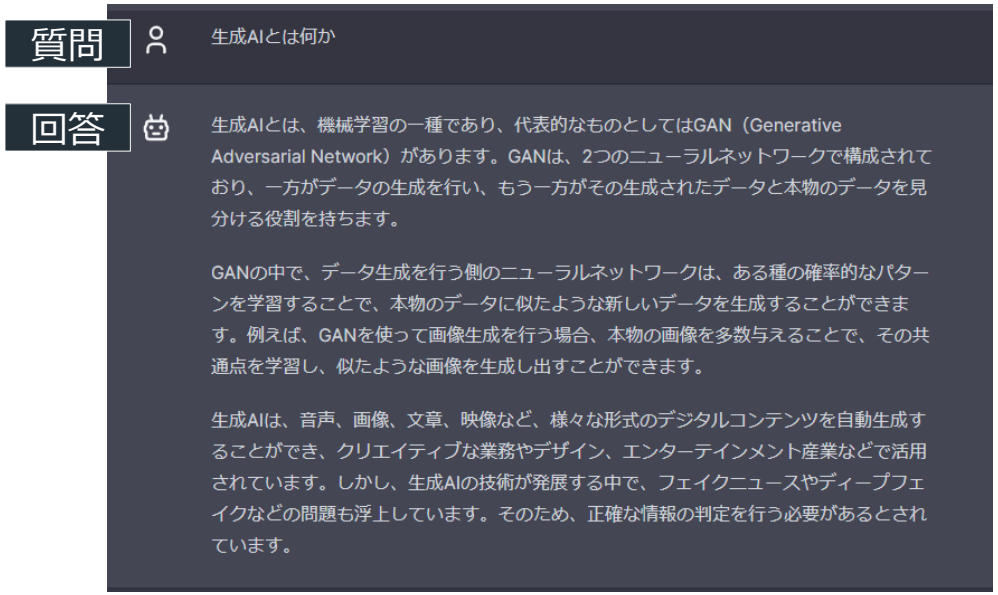


OpenAIの生成AI（Generative AI）：ChatGPT

- 生成AIの中でも特に注目されているChatGPTは、OpenAIが2022年11月に公開したチャットボットOpenAIのGPT-3.5ファミリーの言語モデルを基に構築されており、教師あり学習と強化学習による転移学習で実装
- 多くの人が使いやすい簡易なUIでチャット形式で利用でき、かつチャットの会話履歴を記憶し回答可能

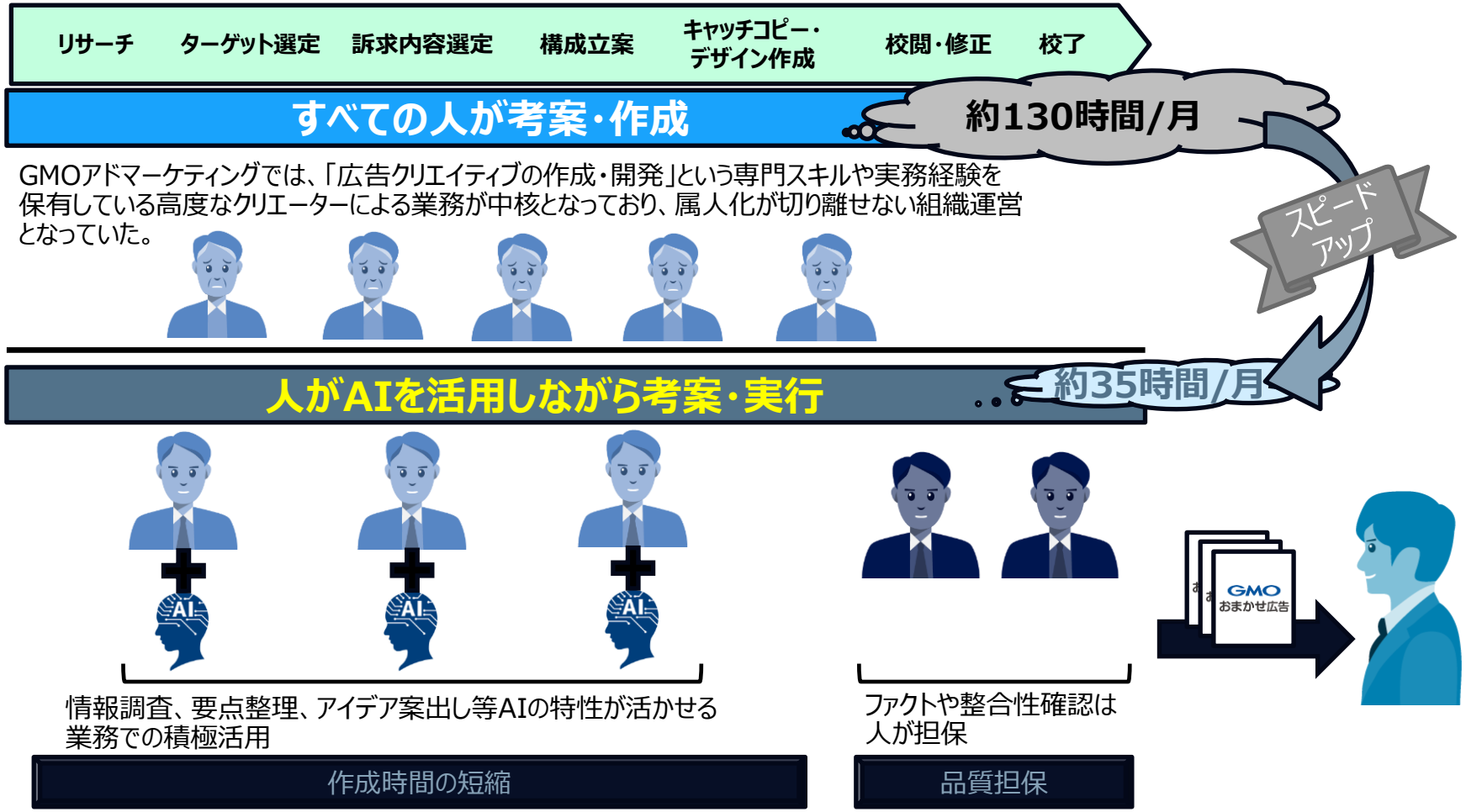
*ChatGPT*はなぜ流行ったのか？

- 高度な自然言語処理能力
- 多岐にわたる応用範囲
- ユーザビリティの向上
- 進化と改善への取り組み



【事例】GMOアドマーケティング | 「ChatGPT」活用の背景と作業効率化

- GMOアドマーケティングは、インターネット広告事業において「ChatGPT」を活用し、デジタル広告のクリエイティブ作成時間を従来の1/3以下に削減

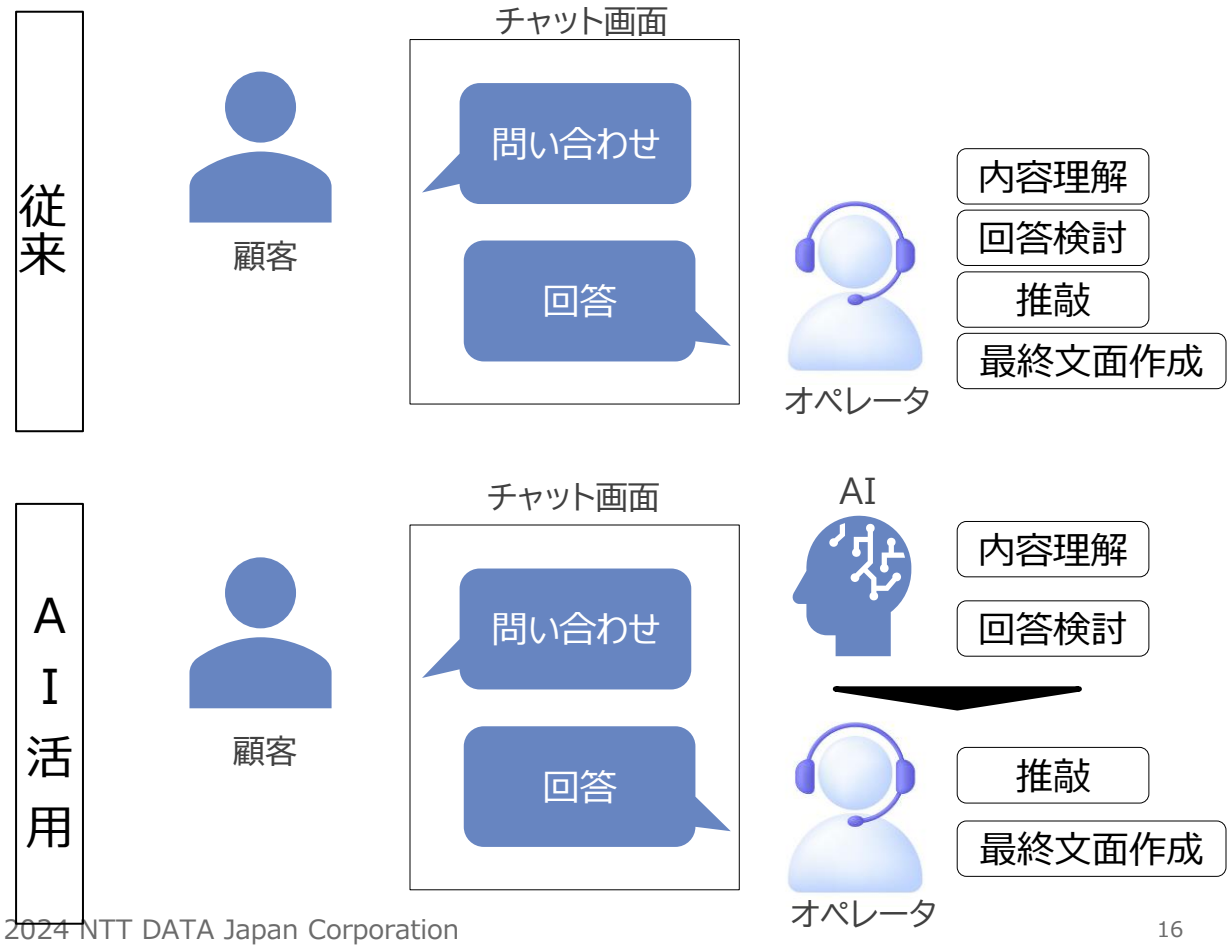


出典情報をもとに弊社にて作成
出典：GMOアドマーケティングニュースリリース
https://www.gmo.jp/news/article/8372/

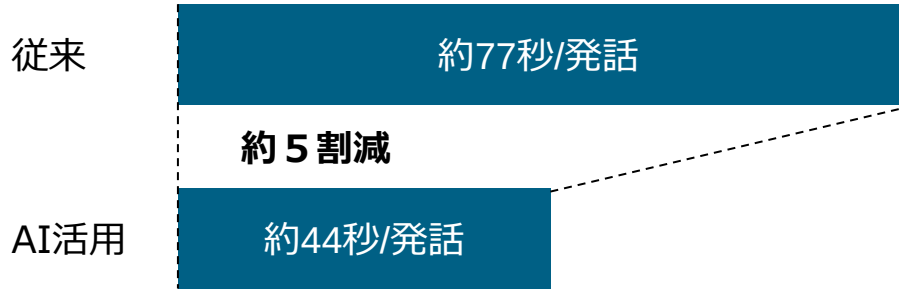
【事例】 東京海上日動の「ELYZA Brain」活用

- 高い精度と品質が求められる「お客様応対業務」で言語生成AI活用の実証実験を実施
- お客様への応対文面の作成業務において約50%の省力化に成功

生成AIの活用



発話時間の削減効果



オペレータからの評価

- 入力がスムーズになった
- 返信内容を均一化できる

出典：日経新聞（顧客応対、生成AIで5割短縮　イライザと東京海上日動）

生成AIをビジネスに活用したユースケース

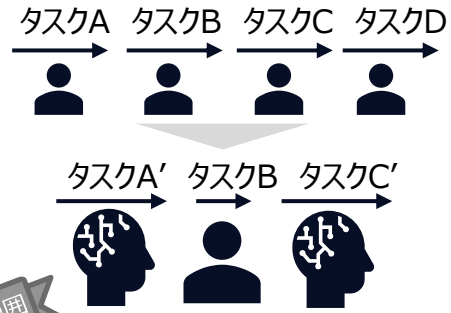
- 生成AIでテキスト生成・プログラミング・文書の要約・機械翻訳など、多岐にわたるタスクを遂行
- これらのタスク遂行能力は、ビジネスの様々な領域で活用可能

金融サービス・業務支援

既存業務の効率化

生成AIを既存タスクの補助・代替として使い、個人の**生産性の"量"**の部分が向上

単に特定のタスクを生成系AIに置換えるだけでなく、業務フロー全体を生成系AI前提で見直したほうがよい場合も



活用範囲

【例：口座開設申込】
新規口座の開設などの記入支援、プロセスの自動化　など

アイデア出し・補完

生成AIをアイデア出しや補完に使うことで、壁打ちのなどの効率性・品質を向上し、**生産性の"質"**の部分を向上

例えば調査の第一段階を行うことや評価観点のMECE向上のための補完

活用範囲

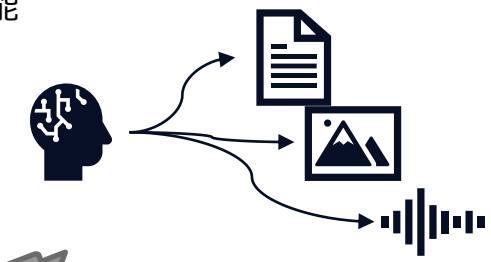
【例：リスクマネジメント】
マーケットレポート等の分析、経済・政治的リスクの予測

【例：文章作成・校正】
XX文書の草稿作成、文書校正

非構造データによるビジネス価値向上

テキスト・画像・音声のような従来では作成が困難だった**非構造データを生成**を活用した新規・高付加価値のビジネスの考案

ユーザやアクセスごとに毎回違うデータを生成するコンテンツのバリエーションや、非構造データのも含め情報提供のパーソナライズも可能



活用範囲

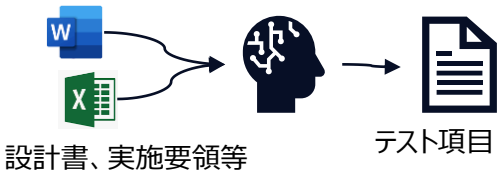
【例：顧客の資産管理】
パーソナライズされた投資推奨

【例：ファイナンシャルプランニング】
予算編成、退職計画など、パーソナライズされた財務計画

システム開発支援

開発プロセス適用

要件定義、設計、製造、テスト工程のEnd-to-End で**生産性向上**を目的に、生成AI を活用したシステム開発プロセス標準化を推進。



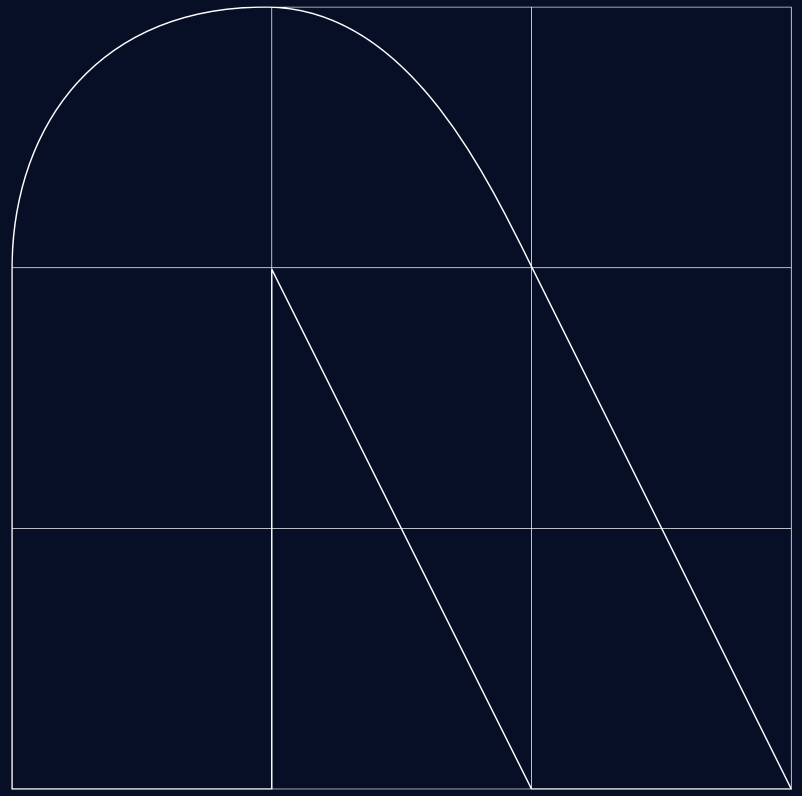
活用範囲

【例：テスト項目の生成支援】
設計書、実施要領・テスト観点、過去テスト項目などをインプットに、網羅的なテスト項目を生成支援

出典情報をもとに当社にて作成
出典：「金融機関はChatGPTにどう対処すべきか」（日本総合研究所）
https://www.jri.co.jp/MediaLibrary/file/column/opinion/pdf/14206.pdf

3

金融での生成AIの活用



国内金融機関における生成AI（ChatGPT）の導入状況①

- 金融業界でもChatGPTの積極的な利用を行うニュースが顕著 【2023年4月～6月】

No.	企業名	取り組み状況
1	MUFG	三菱UFJ銀行が「Azure OpenAI Service対応チーム」を組成。文書検索や企画書の作成補助、顧客対応支援といった用途での可能性を検証。同年度内にグループ全体で共通化できる仕組みの構築を目指す。
2	SMBCグループ	日本総研、NECと共同で、Azure OpenAI Serviceを活用したAIアシスタントツール(SMBC-GPT)の実証実験を開始。同年秋までに三井住友銀行の全行員に展開予定。
3	みずほ銀行	Azure OpenAI Serviceの活用検討を開始。稟議書や契約書の作成支援、事業手続きやシステムに関する社員からの照会対応、自然言語による金融関連のデータ収集およびコード生成などへの活用想定。
		自社システムの開発や保守に生成AIを活用する実証実験を富士通と始めた。米OpenAIが提供する「ChatGPT」を使用し、設計書の記載ミスや漏れを自動で検出し、システム開発の品質を向上させる。
4	大和証券	「ChatGPT」を導入し全社員約9,000人を対象に利用を開始。Azure OpenAI Serviceを利用し、セキュアな環境にて、全業務に利用可。
5	東京海上日動	Azure OpenAI Serviceをもとに保険代理店からの照会などに対応する対話型AIの開発を開始。同年6月に一部支店で試験活用を始める。
6	MS&AD	NECおよびアクセンチュアと連携し、米OpenAIが提供する「ChatGPT」を活用した社内の業務効率化の取り組みに加え、事故対応サービスにおける社員とAIの新たな協業モデルの構築を発表。
7	損保ジャパン	DX推進部開発推進グループでプログラミングの支援に「ChatGPT」を活用。

出典情報をもとに当社にて作成
出典：日本経済新聞

国内金融機関における生成AI（ChatGPT）の導入状況②

- 各業界で生成AIの実証実験/業務適用が拡大 【2023年7月～9月】

No.	企業名	取り組み状況
1	ゆうちょ銀行	Azure OpenAI Serviceを用いた社内チャットボットの実証実験を開始。ChatGPTをカスタマイズし、情報ソースの明確化や文章構成のチューニングを行った。
2	みずほ信託銀行	エクサウィザーズの「exaBase IRアシスタント powered by ChatGPT」で業務提携。決算説明会や株主総会、記者会見などの想定問答集を自動作成で業務効率向上を支援する。
3	大分銀行	生成AIを用いた業務生産性向上の実証実験を開始。Azure OpenAI ServiceとPaLM2を活用し、融資稟議書の作成支援に活用する。
4	ほくほくFG	ほくほくフィナンシャルグループ(北陸銀行・北海道銀行)は富士通のFujitsu Kozuchi (code name) - Fujitsu AI Platformを用いた業務効率化の実証実験を行う。
5	宮崎銀行	生成AIの業務適用に関する実証実験を開始する。融資関連の書類作成などをChatGPTにて支援し、行員によって個人差のあった部分を解消する。
6	住友生命	Azure OpenAI Serviceを利用した社内チャットサービスをグループ内1万人向けに展開。文書・資料作成など日常業務の効率化だけでなく、新規事業の創造に対しても活用していく見通し。
7	オリコ	Azure OpenAI Serviceの実証実験を開始した。業務効率化のみならず、サービス品質や提供スピードの向上など、業務への活用可能性を様々な角度から検証する。
8	SBI HD	「SBI生成AI室」を設立し、グループ内でのノウハウ共有やAI利活用戦略の検討を推進する。また、グループ内のAzure OpenAI Serviceの導入を実施した。

出典情報をもとに当社にて作成
出典：各社リリースニュース

金融機関における生成AI（ChatGPT）の導入状況

継続して導入が広がりつつ、PoCの結果を公表する企業も徐々に増加。【2023年10月～12月】

No.	企業名	取り組み状況
1	MUFG	事務や営業といった銀行業務への生成AI導入で、労働時間削減効果が**月22万時間以上**に相当すると試算。
2	肥後銀行	生成AI（**ChatGPT**）の業務活用を進めるとともに、行員のＩＴリテラシー向上に資する実証実験を開始。情報収集（企業調査、市場調査等）、企画書等の素案作成、アイデアの創出などで活用
3	福井銀行	**Azure OpenAI Service**上に生成 AI「ChatGPT」を組み込んだ生成 AI ツール「Asis-AI」の共同開発を株式会社ギブリーと開始。 企画業務における論点整理、融資稟議資料・提案資料の作成・校正・要約などで活用
4	横浜銀行	従業員専用の情報分析プラットフォーム「行内ChatGPT」を導入。Azure上で一般的な「ChatGPT」の機能に加え、行内の各種規程やマニュアルなど情報照会（**RAG**）に対応。
5	遠州信用金庫	**チャットGPT**を活用した業務効率化を開始。各種データを集計するプログラム構築に貢献するなど、本部から営業店への還元スピードが早くなり、営業の方針転換や顧客満足度向上につなげる
6	明治安田生命	**Azure OpenAI Service**上に資料作成や社内照会等の日々の業務をサポートする当社専用のシステム「ＡＩアシスタント」を導入。
7	セゾンカード	**Azure OpenAI Service**上にAI アシスタントサービス「SAISON ASSIST」の提供を開始、12 月より社内情報回答チャットボット「アシストくん」のテスト運用を開始。
8	東京海上日動	ELYZAと共同で保険契約者からの問い合わせを受けるコールセンターに生成AIを導入し、**作業時間を約50%短縮**する効果を確認した。

出典情報をもとに当社にて作成
出典：各社リリースニュース、日本経済新聞、ニッキン

金融機関における生成AI（ChatGPT）の導入状況

PoCの結果を公表する企業や、生成AI向けの専用保険開発といったサービスにも派生。【2024年1月～3月】

No.	企業名	取り組み状況
1	みずほ銀行	システム運用における品質向上と効率化を目指す実証実験を、2023年8月から3ヵ月間、IBMと共同で実施。イベント検知における**エラーメッセージの監視と対応において98%の精度**
2	紀陽銀行	**Azure OpenAI Service**上にChatGPTを導入し、本部業務への活用可否・ユースケースを検討。
3	常陽銀行 足利銀行	**Azure OpenAI Service**上にChatGPTを導入し、情報収集や資料構成案の作成等にかかる業務の効率化を進め、営業活動や企画業務等へ充当する時間を捻出するなど生産性の向上に繋げる。
4	佐賀銀行	Azure OpenAI Service上にChatGPTを導入し、システムベンダーの力を借りずに**若手行員プロジェクトチームにて内製化**
5	静岡銀行	生成 AI を活用し、銀行特有の慣習や専門知識を理解したうえで、IT 企画業務に関する相談やアドバイスに対応する **CIO を想定した対話パートナー**の実現（擬人化）に向けた実証実験を開始
6	三菱HCキャピタル	**Azure OpenAI Service**を活用したセキュアな生成AIを営業事務をはじめとするさまざまな業務へ適用し、その生産性の向上、成長の加速を図る
7	あいおいニッセイ同和損保	「**生成 AI 専用保険**」を株式会社Archaicと共同開発。生成 AI の利用により、知的財産権の侵害や情報漏洩等が発生した際に、企業が負担する様々な費用を補償する。 ガバナンス体制の構築支援や事故発生後のコンサルティングサービスをパッケージで提供。

出典情報をもとに当社にて作成
出典：各社リリースニュース、ニッキン

金融機関が直面するChatGPT活用における6つの課題

- 2023年3月時点ではChatGPTの活用に慎重な姿勢をみせた企業も少なくなかったが、4月に入ると、大手金融機関が軒並みChatGPT導入
- 大手損保でも導入が相次ぐ一方で、ChatGPT活用の課題も具体化

正確性

例）金融商品の説明が誤っていたり、情報が古かったりする場合がある

不確実性

例）金融商品に関する顧客からの問い合わせで想定外の返答をする恐れがある

自社の最新データベースやWebサイト/法令と連携されるなど、不正確・想定外の情報を返さないように制御する仕組みを設ける

説明責任/透明性

例）ChatGPTが出した保険料の見積もり根拠が不明

返答の根拠となるリンクや情報ソースの明示する仕組みを設ける

データ保護

例）機密情報の漏洩や、FISC安全対策基準で求められるデータ保管への適合リスク

閉域網での利用や暗号化といった漏洩対策、使用するリージョンを検討

サイバーセキュリティー

例）AIが人間になりすまして口座情報やクレジットカード番号を詐取

AIの悪用を想定した新たな検知の仕組みの導入なだが必要に

リテラシー

例）金融関連の質問をしても月並みの回答しか返ってこない

「プロの銀行員としての回答して」と入力するなど、ChatGPTの特性を知った活用方法ついて学ぶ

出典情報をもとに当社にて作成
出典：日経XTECH「ChatGPT活用における6つの課題」

金融機関における生成AIの対顧客利用: 克服すべき課題

- **生成AIの対顧客利用には大きな可能性がある**一方で、金融機関特有の課題をクリアする必要がある
- データ、セキュリティ、誤り、規制の4つの観点から、克服すべき課題を整理

データ品質の確保

- 生成AIの学習に用いるデータの品質が結果に大きく影響
- **金融データの正確性、最新性、代表性の担保**が重要
- データのバイアスによる不適切な結果の回避

セキュリティとプライバシーの確保

- **顧客の機密情報や個人情報保護のプライオリティの意思統一**
- 対プロンプトインジェクションなど生成AIシステムへのセキュリティ対策の徹底
- 顧客のプライバシー保護や業務上の秘密と説明責任のバランス

誤った結果や判断のリスク対策

- **生成AIによる誤った金融アドバイスや意思決定の可能性の配慮とその対策**
- 人間による監督と生成AIの結果をオーバーライドする対策
- **AIの限界と人間の関与の明確化 (Human in the loop)**

規制対応と内部統制

- 金融規制のほか、法的規制への適合性の確認と対応
- **企業内のAIガバナンスの枠組み構築と内部統制の強化**
- 生成AIの利用に対する説明可能性と公平性の担保

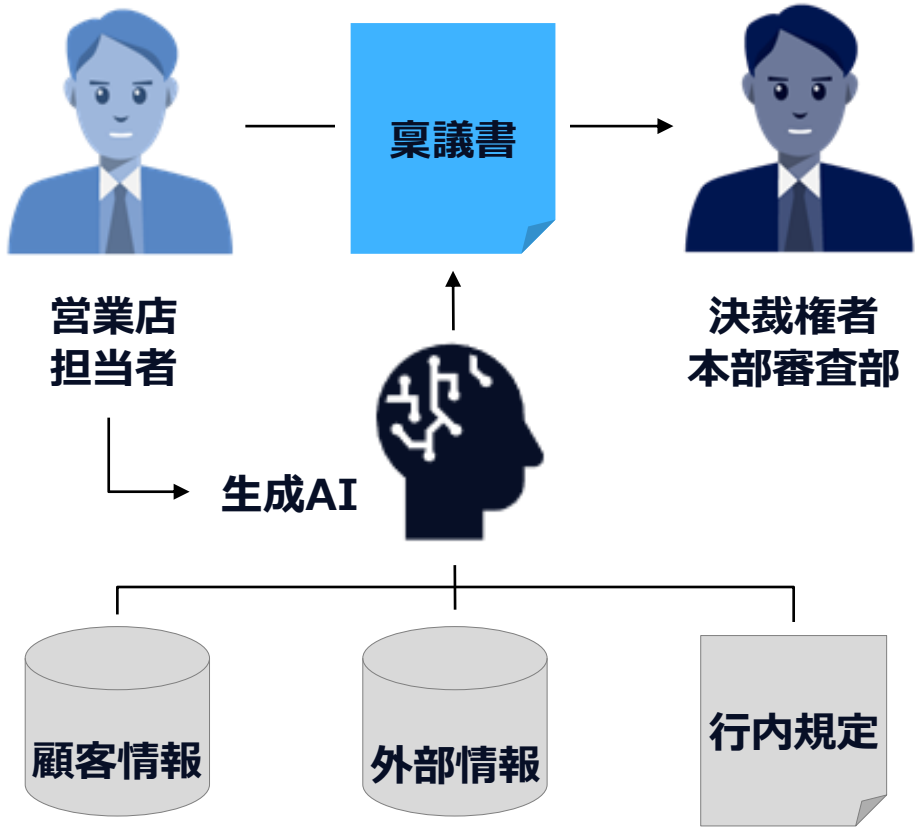
金融における生成AI活用のユースケース　①融資稟議作成

- 金融機関が融資審査を行う際に必要となる稟議書の作成を、生成AIが支援

概要

- **金融機関が融資審査を行う際に必要となる稟議書の作成を支援**
- **顧客の企業情報や過去の事例などをインプットとして稟議作成負荷を軽減**
- **生産性の向上に加え、属人化したノウハウの標準化等の効果も期待される**
- **大手行に限らず、地方地銀でも活用を模索する動きが活発化している**

イメージ

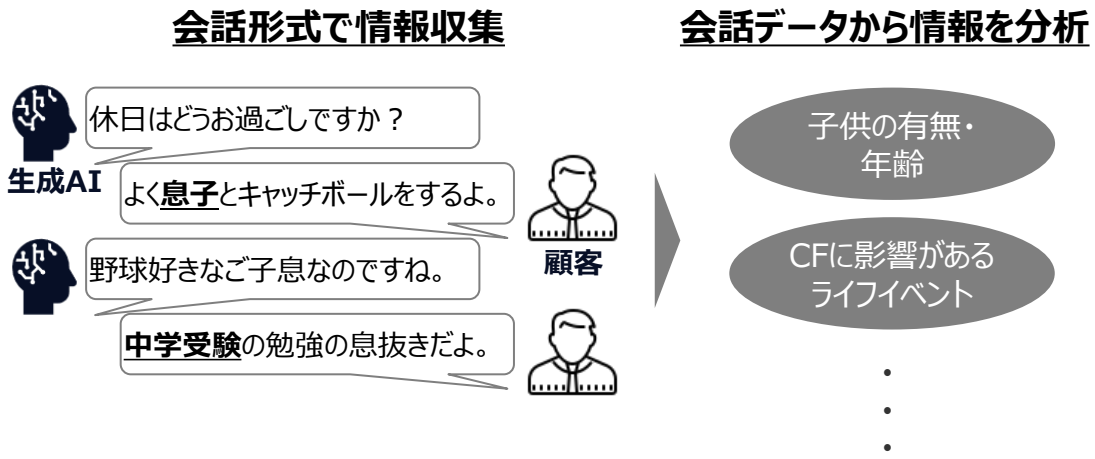


金融における生成AI活用のユースケース　②ファイナンシャルプランナーの支援

- FPの業務における「顧客のニーズ把握」と「提案内容の最適化」の領域に生成AIを活用
- FPの生産性は向上し、顧客に対してもより適切なアドバイスを提供できるようになる

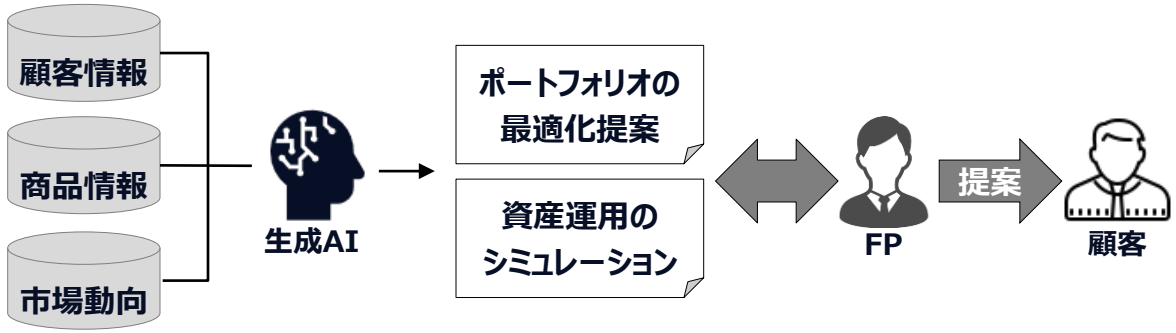
資産運用ニーズに関する情報を顕在化

- **会話形式で顧客と質疑応答を行い、個人の行動履歴、家族構成、ライフスタイル等から資産運用の目標やリスク許容度などを聞き出す**



提案作成の支援

- **入力値をもとにポートフォリオ最適化や資産運用シミュレーションを実施し、FPのアドバイス作成を補佐**
- **FPは生成AIからの情報と自身の専門的な知識を組み合わせ、顧客のニーズに合わせた提案を行う**



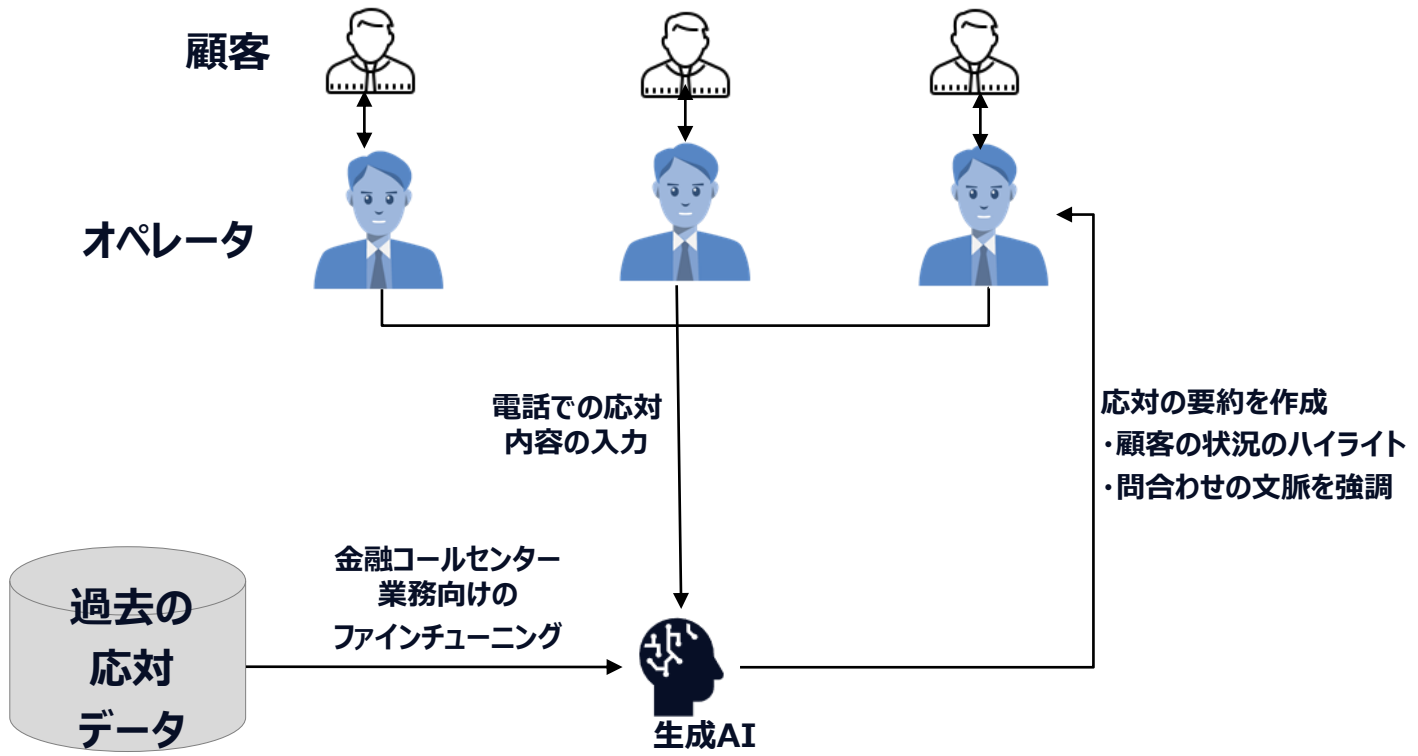
金融における生成AI活用のユースケース　③コールセンター入電業務の要約サポート

- ヘルプデスクのオペレーターの入電応対内容の要約を生成AIがサポート
- 生成AIのサポートにより新人のオペレーターでも熟練オペレーターと同レベルの要約作成が効率的に可能

概要

- **金融機関のコールセンター用に言語モデルをファインチューニング**
- **代表的なアフターコールワークである顧客とオペレーターとの応対内容の要約を生成AIがサポート**
- **金融業務での、顧客の問題理解や、会話で問題解決に至るまでの情報を生成AIに組み込み、オペレーターの対応内容の要約を効率化**
- **新たに作られた要約は、新人オペレーターの教育や、オペレーターが参照する文書としても活用可能**

イメージ



出典情報をもとに当社にて作成
出典：https://dl.acm.org/doi/abs/10.1145/3604237.3626838
Fine-Tuning Pretrained Language Models to Enhance Dialogue Summarization in Customer Service Centers

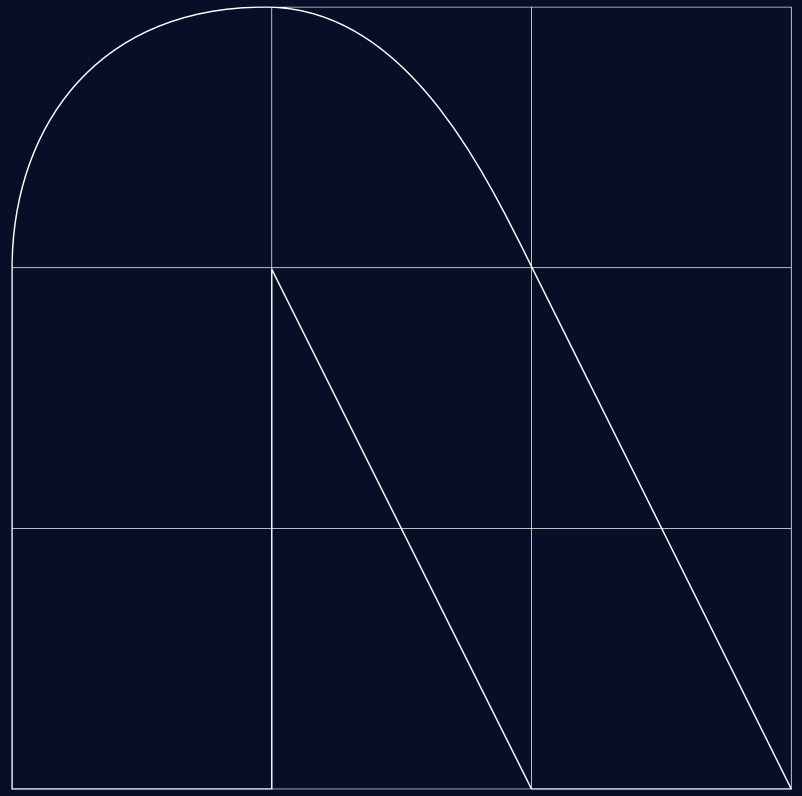
生成AI利用上の注意点

- 生成AIは自分の課題についての考えを整理するのに有効に作用
- 注意すべき点もいくつかあり、生成AIツールを利用するうえでの初歩的な注意点を紹介

回答をそのまま利用	著作権に関するリスク	個人情報や機密の漏洩
• **生成AIツールの出した回答を自らの業務の成果物としてそのまま利用することは自分自身の学びや成長ならない**	• **生成AIツールの学習データに他者の著作物が含まれている可能性がある** • **それらの使用が図らずも著作権侵害や剽窃につながるおそれがある**	• **生成AIツールに個人情報や機密情報を入力すると、情報が意図せず流出・漏洩してしまう可能性がある**
生成AIを利用して成果物を作っても、第三者(上司、同僚、取引先)のレビューが入る 第三者に内容に質問されても、「生成AIで作った」という理由で、意見の根拠の説明を逃れることはできない 仮に思考を整理するために生成AIを利用したとしても、書かれていることの確からしさや根拠を出典などを理解して説明することが必要	成果物の作成は、他者の文献を引用し自分の考えを主張することが重要 生成AIは人類が書いてきた大量の文章から、指示に沿ったパターンを見出し、合致しそうなコンテンツを生成 指示の内容によって、どんなデータが出てきたか、その諸元を明らかにしておくことは必要	生成AIに指示をするときに「あなた自身のこと」をいろいろと書いていないか？ 生成AIはこれらの指示文を学習して新たな自らの回答に使う 指示するときの設定で指示文の内容を学習させないなど注意 あなたの指示文で所属する組織の機密が生成AIの学習に利用され、責任を問われる場合もある
・**RAG(Retrieval-Augmented Generation)などを用い**て生成AI外の知識を優先的に回答する方法も	・生成AIを利用した画像等のアップロードや販売は、既存の著作物との「類似性」や「依拠性」の評価が必要	・個人情報や機密の漏洩が看過できない場合には**オンプレミスでの生成AI利用も**考慮

4

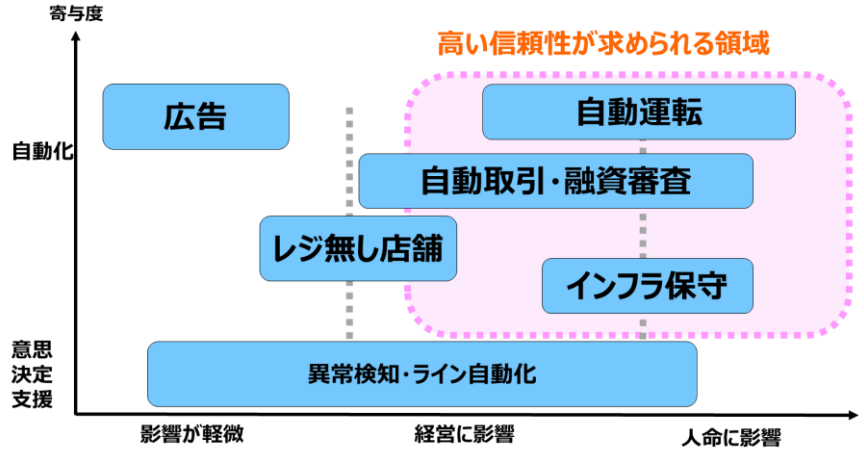
生成AI時代のAIガバナンスと課題



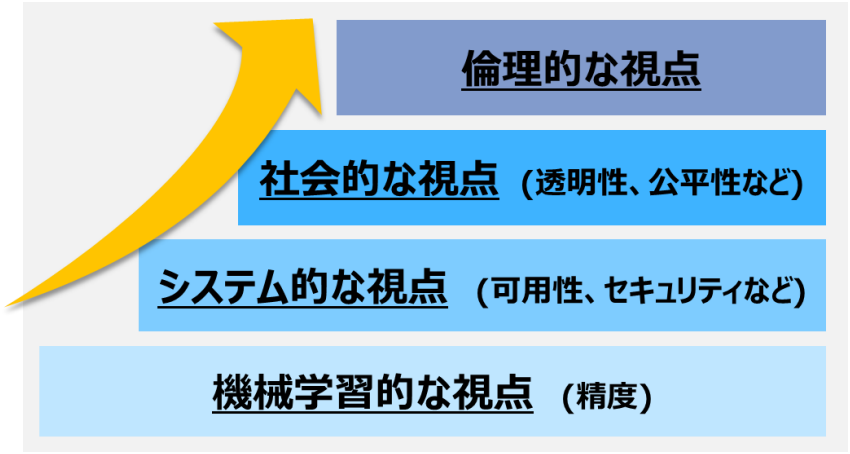
AIガバナンスが求められつつある背景

- AI・データ活用が拡大する一方、活用に際しての「社会的・倫理的な視点」での評価が重要

AI・データ活用の活用範囲の拡大



AI・データ活用の評価視点の変化



社会的・倫理的に問題になったAI事例

国内：取り組み中断やサービス停止に発展した事例

問題点	発生事象	概要	影響
偽情報拡散	水害画像のフェイク拡散 (日:2022)	2022年9月下旬の台風15号により水害が発生した静岡県に関して、フェイク画像がTwitter上で拡散された問題。不自然な個所からフェイクを疑われ、画像生成AIで作成したフェイク画像と認めて謝罪。	• 内閣官房長官が注意喚起
プライバシー侵害	監視カメラによる出所者の顔認識(日:2021)	顔認識カメラを使って、刑務所からの出所者と仮出所者の一部を駅構内で検知する防犯対策を実施していることが大手新聞社により報じられた。	• 取り組み停止
他者創作物の侵害の恐れ	AIイラスト生成技術の公開停止 (日:2022)	日本某社が公開したイラスト生成AIが1日で公開停止になった問題。他人の画像を勝手に読み込ませることが可能な点に懸念の声が上がった。	• サービス停止

国外：訴訟問題に発展した事例

問題点	発生事象	概要	影響
人種差別	顔認識AIで誤認逮捕 (米:2020)	監視カメラに映っていた人物と別人にもかかわらず、顔認識システムでの判定結果で、黒人男性が万引き犯に間違われ誤認逮捕。	• 利用停止 • 米連邦議会では政府機関による顔認識の使用を禁じる法案が提出
ソーシャルスコアリング	育児給付詐欺検出システムによる誤った告発 (蘭:2021)	オランダ政府運用の育児給付詐欺を検出システムにより、26,000の無実の家族が社会給付詐欺で誤って告発。被害家族の多くは、有色種や低所得者、移民の履歴を持った人であった。	• 2021年に内閣総辞職
	債務回収システムによる債務通知の誤送付 (豪:2016)	社会福祉省が導入したシステムで、債務が無い生活保護者や先住民族なども含む40万人余りの市民に誤った債務通知を送った。	• 被害者40万人余りによる集団訴訟に発展し、最終的に被害者と18億ドルで和解

AIガバナンスを取り巻く国内外の動向①

- 国際的にAIガバナンスの在り方が議論され始め、国内関連省庁でもルールメイクに向けた動きが活発化

欧州のAI規制法案（2024年3月13日可決）

欧州

EUの立法機関である欧州議会が世界初となるAIの包括的な規制法案を可決した。EU加盟国が5月に正式に承認し、2025年の早期に発効し、2026年から適用される見通し

https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

経済産業省(AI全般のルールメイク、ガイドライン策定を所掌)

経産省

2021年、AI原則の実践の在り方に関する検討会で、「**我が国のAIガバナンスの在り方**」を策定。AIガバナンス要素をレイヤで区分して整理

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_1.pdf

広島AIプロセス(2023年5月～)

G7

- 生成AIの活用や開発、規制に関する国際的なルール作りの推進を目的とする、G7広島サミットの首脳宣言に盛り込まれた新たな枠組み
- 9月7日閣僚級会合で全AI関係者向け国際的指針の年内策定を合意

【指針を構成する主な項目】

① 高度AIシステムの適切な安全対策及び導入前の社会的リスクの考慮
② 高度AIシステム導入後の脆弱性の特定と低減に向けた努力
③ モデルの能力、限界、適切・不適切な利用領域の公表
④ **プライバシーポリシー及びAIガバナンスポリシー等のリスク管理計画及び低減手法の開発及び開示**

https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/kakuryoukyuu.pdf

https://www8.cao.go.jp/cstp/ai/ai_senryaku/4kai/susumekata.pdf

金融庁（ルール・ガイドラインを金融実務に応用する調整を所掌）

金融庁

2021年、「モデル・リスク管理の原則」（MRM原則）を公表し**金融機関に対するモデル*・リスクに関する期待目線**を8つの原則ベースで提示

*定量的な手法で、理論や仮定に基づきインプットを処理し、アウトプット（推定値、予測値、スコア、分類等）を出力するもの

AIガバナンスを取り巻く国内外の動向②

- 海外の大手金融機関では、原理原則的なAIガバナンス指針から、現場業務への実装に注力

JPMorgan Chase

モデル開発時からガバナンスに注力、モデルのコンプライアンスの継続監視や、責任分担も明確

運用指針：AI モデルの開発、テスト、導入、**継続的な監視のための適切なポリシー、手順、管理を確立**

体制：コンプライアンスの所有権と説明責任を確立、監視委員会による確立された責任分担

出典情報をもとに当社にて編集
https://www.technologyreview.com/2023/07/18/1075972/good-governance-essential-for-enterprises-deploying-ai/

Bank of America

現場でのガバナンス体制の構築と最高リスク責任者との密な連携の実現

体制：ガバナンスの体制が組織ごとに配置され、最高リスク責任者にレポートする仕掛け

社会受容：AIの規律を期待する**規制当局に先んじるようにするガバナンスを実施**

出典情報をもとに当社にて編集
https://www.fastcompany.com/90465134/the-top-qualification-for-an-ai-governance-officer-courage

Wells Fargo

米国政府のガイドラインを行内で実装し、回復性、モデル開発とレビューの分離などを推進

運用指針：2022年の**ホワイトハウスのAI権利章典青写真の実装を行内で実施**

利用・運用：特定のモデルでその不利な決定につながった属性とシグナルを関連付けるツールを構築、**Stanford大学の人間中心人工知能研究グループと協業**

出典情報をもとに当社にて編集
https://www.fastcompany.com/90465134/the-top-qualification-for-an-ai-governance-officer-courage

HSBC

上位層がAIの導入を決定するほか、人材育成プログラムを実施しAI活用の底上げを図る

体制：AI の実験やの使用を考えている社員は、その**AIのアイデアを上級リーダーで構成される倫理委員会**に持ち込む

利用・運用：AIの公平性、透明性、説明可能性などの倫理的な観点を考慮した取り組みを実施

出典情報をもとに当社にて編集
https://ffnews.com/magazine/exclusive-truth-accuracy-and-ai-ash-booth-hsbc-in-the-fintech-magazine/

AIガバナンスを取り巻く国内外の動向③

海外の大手金融機関は、AIガバナンスに対して人・金・時間といったリソースを相応に投入していると考えられる

観点	海外大手金融機関のAIガバナンスの特徴
運用指針	ガイドラインを制定し、現場レベルの規定に落とし込んでいる
セキュリティ	経営上位層のコミットメントが明確
体制	行内の責任分担が明確
	開発とレビューで体制を分けるなど相互牽制が効くようにしている
利用・運用	データやモデルの偏りを十分な量でテストしている
	AIモデルのパフォーマンスとコンプライアンス異常のリアルタイム監視をしている
社会受容	AIガバナンスの取り組みを銀行の重要な役職がアピールしている

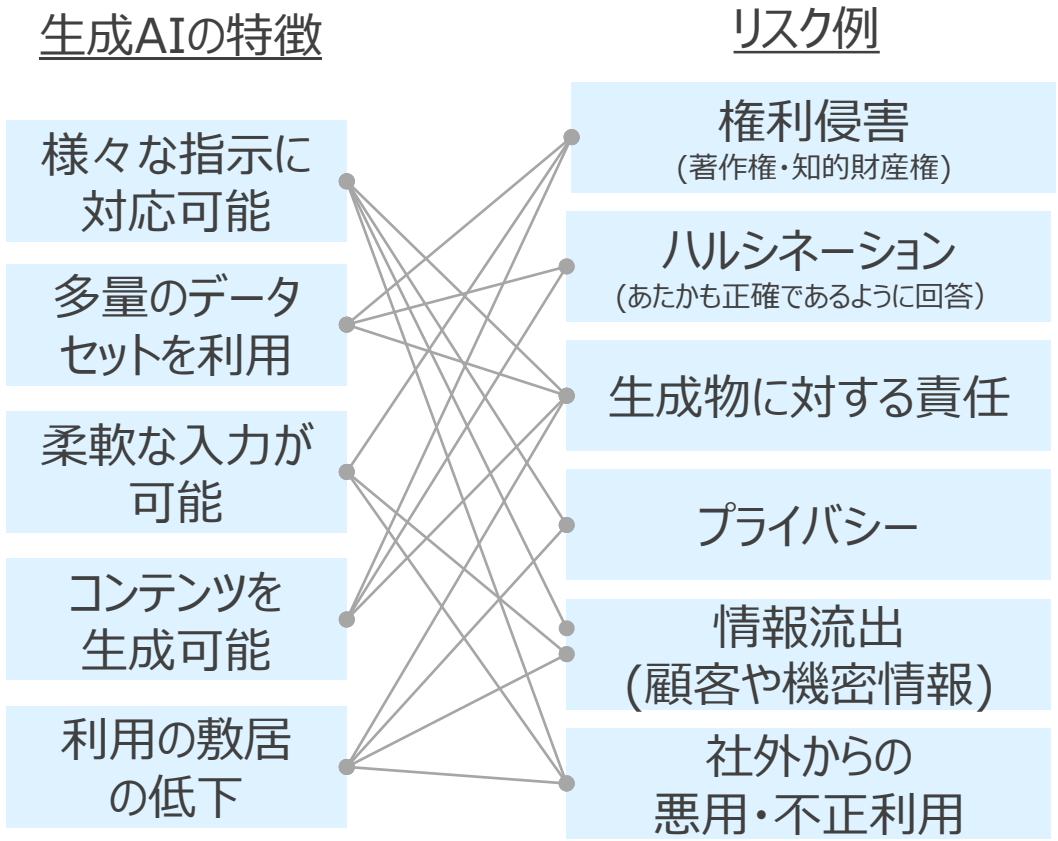
従来AI／生成AIの差異とリスク

- 生成AIの登場によってリスクの新規発生／深刻化が生じており、従前のリスクコントロールでは不十分

従来AI (生成AI 登場前のAI)	データの分類などを学習し予測を行い、結果を出力 **“特定業務等の定められた行為の自動化”**
生成AI	無数の多種多様なデータをもとにパターンや関係を学習 **“指示に従いコンテンツ/成果物を生成”**

	従来AI	生成AI
特徴	事前に目的に沿うデータを学習させて、回答を得る	条件を与えるだけで、**様々な指示に対応可能**
学習	具体的なデータセットで特徴や傾向を学習	**無数の多種多様なデータをもとにして学習**
入力	定められた形式のデータセットを用意して投入	**様々なデータ形式や指示のインプットが可能**
出力	学習済みのデータのパターンをもとに適切な答えを提示	入力された指示に**従って新しいコンテンツを生成する**
利用	定まったデータを用意し、モデルに読み込ませ回答を得る	プログラミング等ではなく**「質問文」形式で利用することが可能**

生成AIでは広範囲に多様な対応ができ、
活用の幅が期待されているがリスクも広がる



従来のAIリスクに加えて、
新たなリスクやリスクの深刻化が発生するのではないか

金融機関における生成AIのガバナンスリスク

- AIの利活用にあたっては、銀行業務特有の問題発生を予見して利用する現場へとルールの浸透が必要

組織・仕組みにおける対応例

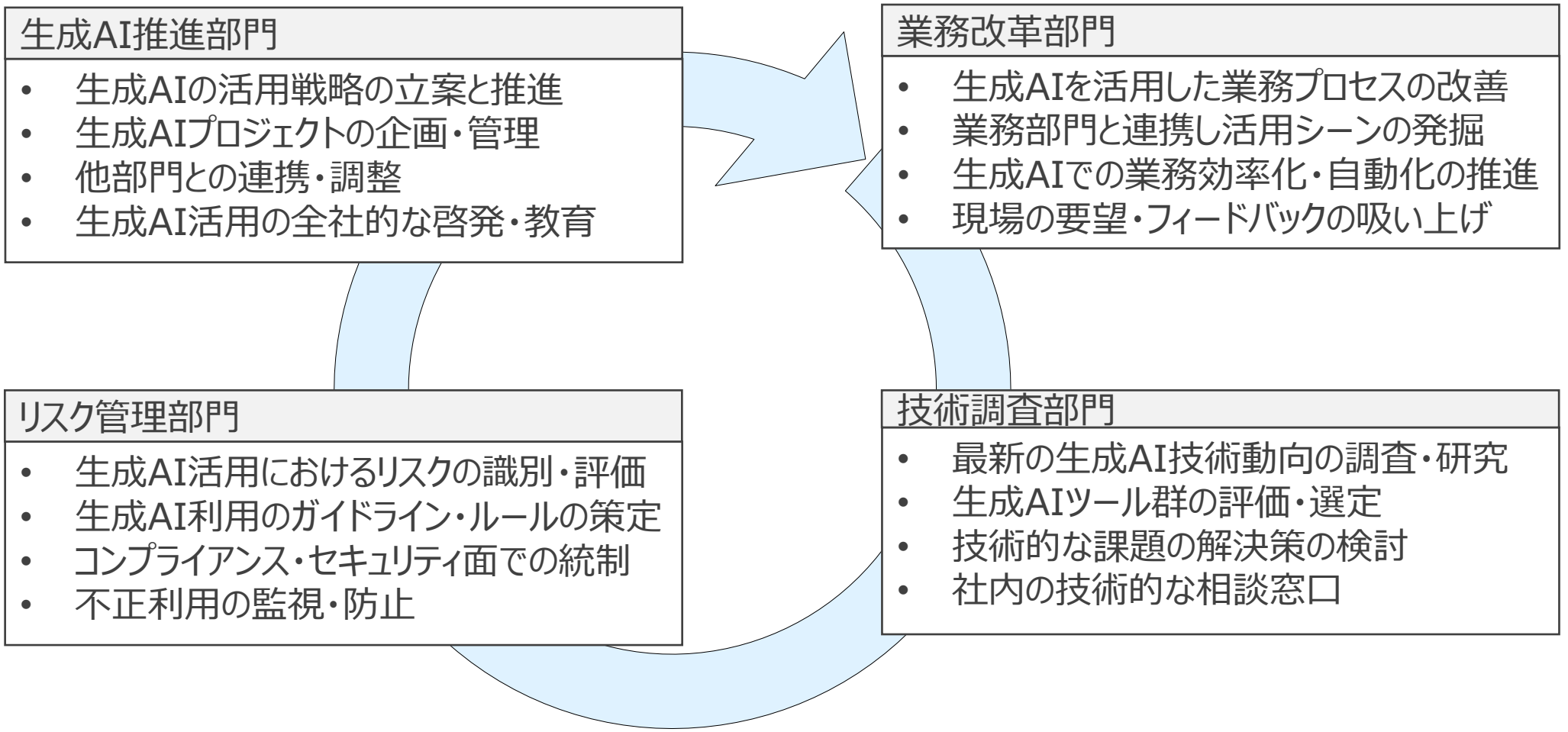
運用指針	● ガイドラインとしてAI利用の“憲章”は表明されているが具体的なAIの管理手続きが明確でない ● 現場でのAI利用に関する統制の効かせ方
セキュリティ	● クラウド、オンプレミスともAIシステムのセキュリティ担保の確認 ● データインジェクション、プロンプトインジェクションなど外部攻撃への対策。
体制	● AI推進とAI統制の役割の明確化(業務部門の推進と、IT部門等の統制の役割、コミュニケーション) ● 業務部門で利用されるAIのニーズ収集と全社的な教育体制が必要となる
利用・運用	● 社員/顧客に対して誤った回答をしないかガードレールの整備が必要 ● 著作権・個人情報保護・機密情報への配慮や浸透 ● 品質確認のめ継続的なモニタリングをすることが必要
社会受容	● 各地域の政府等の最新の規制に対応することが必要 ● 不測の事態におけるベンダーの責任分解 ● AIが代替した業務をしていた人の配慮

業務利用におけるリスク例

インプット	個人情報等の流出	● 顧客・取引先データのAI利用における閲覧が制限されるべき情報が他者に提示される可能性がある
インプット	機密情報等の流出	● 要約等の利用における投入データの守秘義務や機密保持への抵触の懸念
モデル	個人情報や顧客情報の再利用	● 重要な取引先のデータ等が生成AIに取り込まれ社外に流出してしまう。
モデル	過去の一時点の情報	● 過去の時点の情報での回答となり、最新でない情報の可能性がある。
アウトプット	知的財産や著作権の侵害	● 生成AIで作られたデータが著作物を模倣したり、知的財産に抵触する可能性
アウトプット	不正確な回答(ハルネーション)	● 回答に誤りがある可能性や銀行の価値観や業務との整合性を鑑みた際の回答の正しさ

金融機関の生成AI活用を支える組織体制(例)

- 生成AI推進部門が全体の司令塔となり、戦略の立案と推進を主導
- 組織の規模や特性に合わせ、リスク管理部門、業務改革部門、技術調査部門が役割を調整し安定的な活用を実践



金融機関における生成AIの活用の課題

- 金融機関の持つ情報は、企業業績や個人の行動傾向など機密性が高いものが多数
- データの保護とセキュリティを最優先に考慮しつつ、生成AIの精度と信頼性を考慮し活用形態を決定
- **自社の機密データをクラウド環境に入力できるか否かが判断のポイント**

クラウド環境での生成AIの活用

メリット	●初期投資コストが低く、スモールスタートが可能 ●スケーラビリティが高く、柔軟な拡張が可能 ●最新の生成AI技術を利用可能
デメリット	●データ保護とセキュリティ面での懸念が残る ●クラウド、生成AI双方のベンダーロックインのリスク ●ネットワーク遅延や通信コストの発生

オンプレミス環境での生成AIの活用

メリット	●データ保護とセキュリティ面での管理が容易 ●自社のニーズに合わせたカスタマイズが可能 ●ネットワーク遅延や通信コストの低減
デメリット	●初期投資コストが高く、専門人材の確保が必要 ●スケーラビリティに制約があり、拡張に時間とコストがかかる ●最新のAI技術の導入に遅れが生じる可能性

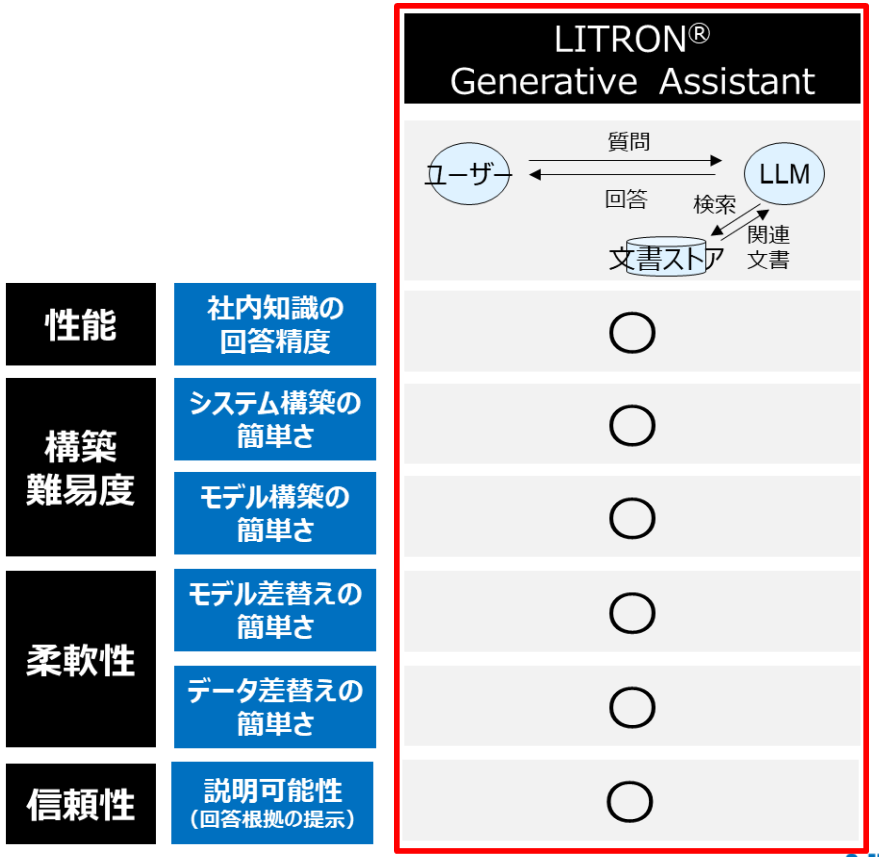
生成AIを安全に利用可能にするLITRON® Generative Assistant

- 生成AIを利用する場合に、利用者の個人情報や企業の機密情報の流出、不正確な回答などが懸念される
- LITRON® Generative Assistantにより、手元にある資料を保護しながら生成AIを利用可能

業務利用におけるリスク例

インプット	個人情報等の流出	● 顧客・取引先データのAI利用における閲覧が制限されるべき情報が他者に提示される可能性がある
インプット	機密情報等の流出	● 要約等の利用における投入データの守秘義務や機密保持への抵触の懸念
モデル	個人情報や顧客情報の再利用	● 重要な取引先のデータ等が生成AIに取り込まれ社外に流出してしまう
モデル	過去の一時点の情報	● 過去の時点の情報での回答となり、最新でない情報の可能性がある
アウトプット	知的財産や著作権の侵害	● 生成AIで作られたデータが著作物を模倣したり、知的財産に抵触する可能性
アウトプット	不正確な回答(ハルネーション)	● 回答に誤りがある可能性や銀行の価値観や業務との整合性を鑑みた際の回答の正しさ

LITRON® Generative Assistantでは、知能(LLM)と知識(社内文書ストア)を分離することで、生成AIによる柔軟な回答と、社内文書に基づく信頼性のある回答が可能




NTT DATA