

AIは経営に何を及ぼすか

～データの時代が到来どう乗り越える～

2024年2月14日

株式会社NTTデータグループ

相談役 岩本 敏男

本日のアジェンダ

- NTTデータグループの概要
- 情報の3階層
- 企業における取組み
- NTT大規模言語モデル

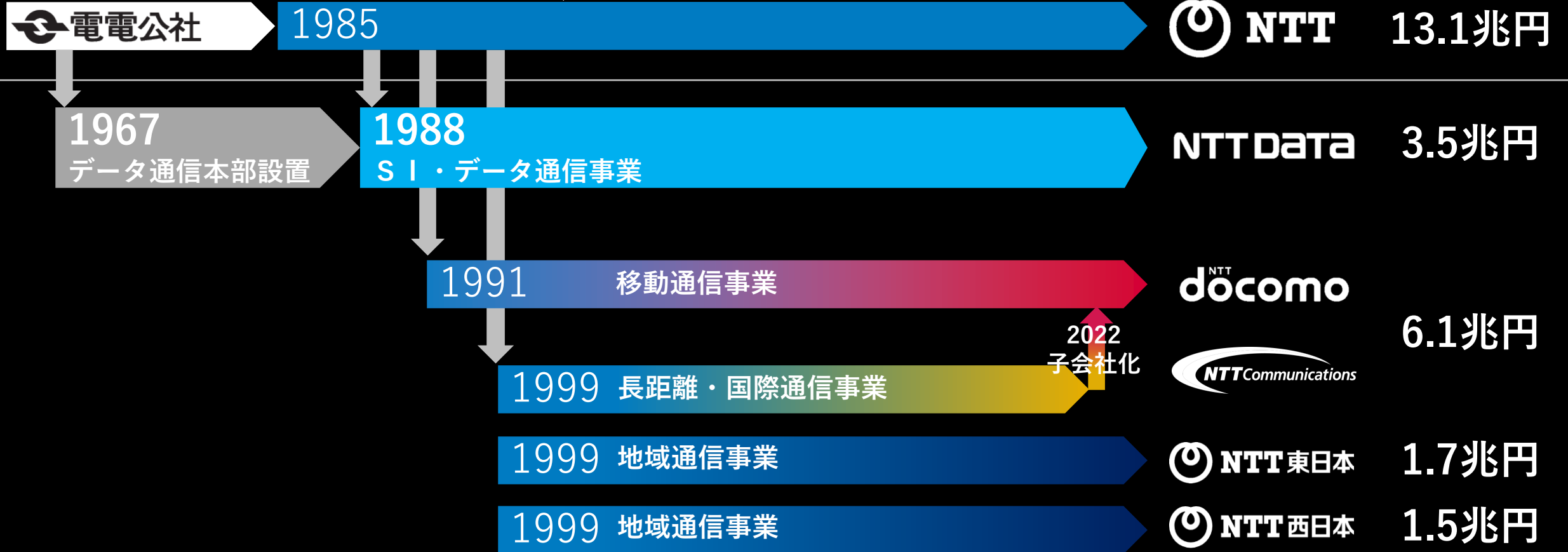
NTTデータグループの概要

沿革

民営化

持株会社へ移行

FY22
営業収益



(売上は各社HP掲載の2022年度決算値をもとに、0.1兆円単位になるよう四捨五入)

沿 革

FY22
営業収益

13.1兆円



統合ICT事業

^{NTT}
docomo



6.1兆円

地域通信事業

 NTT 東日本

 NTT 西日本

3.2兆円

グローバル・ソリューション事業

NTT Data ×  NTT Ltd.

「SI事業」を強みとするNTTデータと
「DC、NW、クラウド事業」を強みとするNTT Ltd.が統合

3.5兆円

(売上は各社HP掲載の2022年度決算値をもとに、0.1兆円単位になるよう四捨五入)

FY2022 4兆円超
3.49兆円

- 海外グループ会社
- 国内グループ会社
- NTTデータ
(国内事業会社)

海外売上比率
54%

国内売上比率
46%

34期連続増収

NTTデータ
1988年創業



グローバルカバレッジ

50 以上の国と地域

197,000 人以上

ヨーロッパ
47,900+

中東・アフリカ
10,300+

インド
38,500+

中国
4,900+

日本
44,700+

APAC
10,300+

北米
20,200+

南米
20,400+

情報の3階層

経営の3大資源

情報



ヒト



モノ



カネ

では、何故 今
データが重要なのか？

そもそも、データとは何？

Data

ラテン語・イタリア語

Dare（与える）が語源

伝達、処理などに適するように形式化、
符号化されたもの

“感知”できるもの

伝達や処理などに適するように形式化、
符号化されたもの

(投影のみ)

約46億年前地球誕生以来 森羅万象は全て自然現象



約20万年前、ホモサピエンスが誕生以来
人が織りなす営み、社会現象が加わった



しかし、これらを表すデータは全てアナログ



それが今、デジタルデータに置き換わった

では どうして今 データが重要なのか？



ITにより森羅万象(自然現象・社会現象)
をデジタルデータ化することが可能になる



ITの進化により この大量データを
制約なく収集し処理できる



これができるようになるようになる？

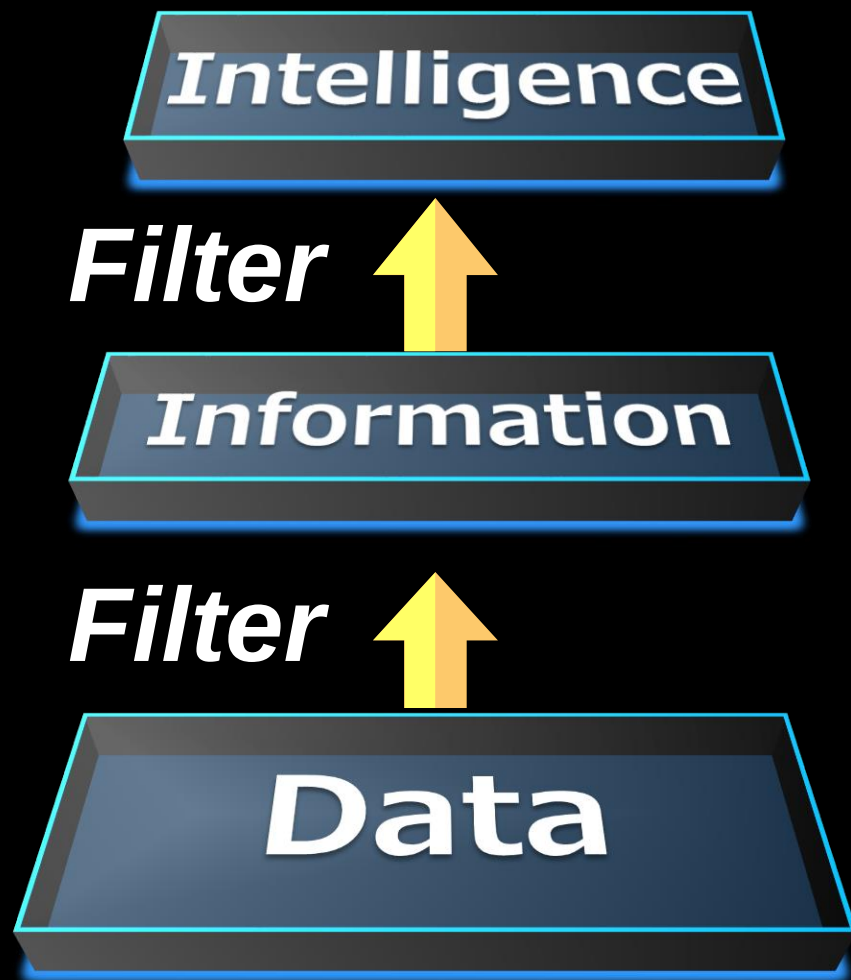
それを解き明かす考え方が

情報の3階層

Data ➡ information ➡ intelligence

(投影のみ)

「情報の3階層」



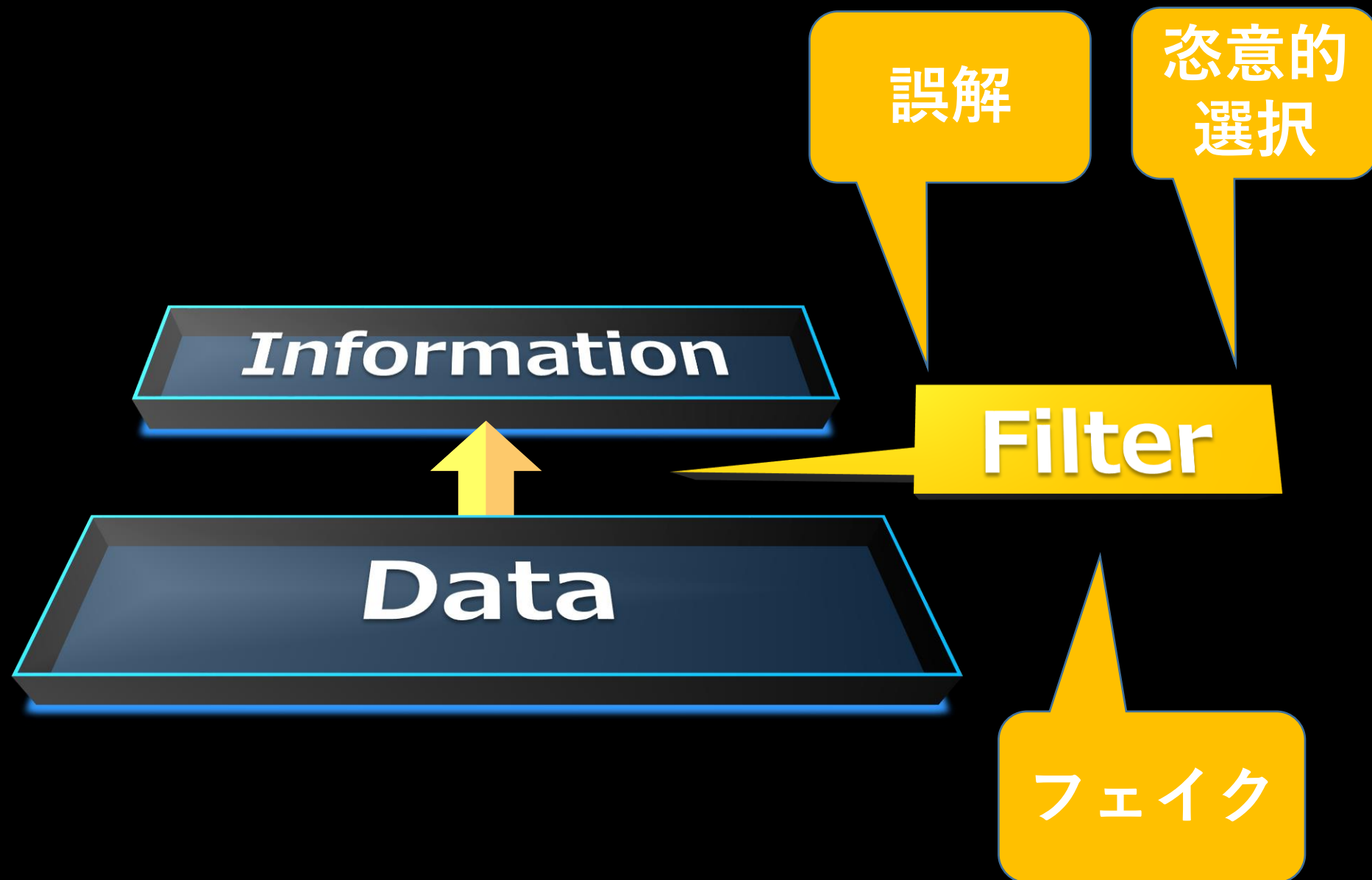
意思決定をする源泉となるもの

一般的に言われる情報のこと
データに意味を持たせ、分析や
評価の対象となるもの

世の中で起こる膨大な事実を表した
もの(意味をもつとは限らない)

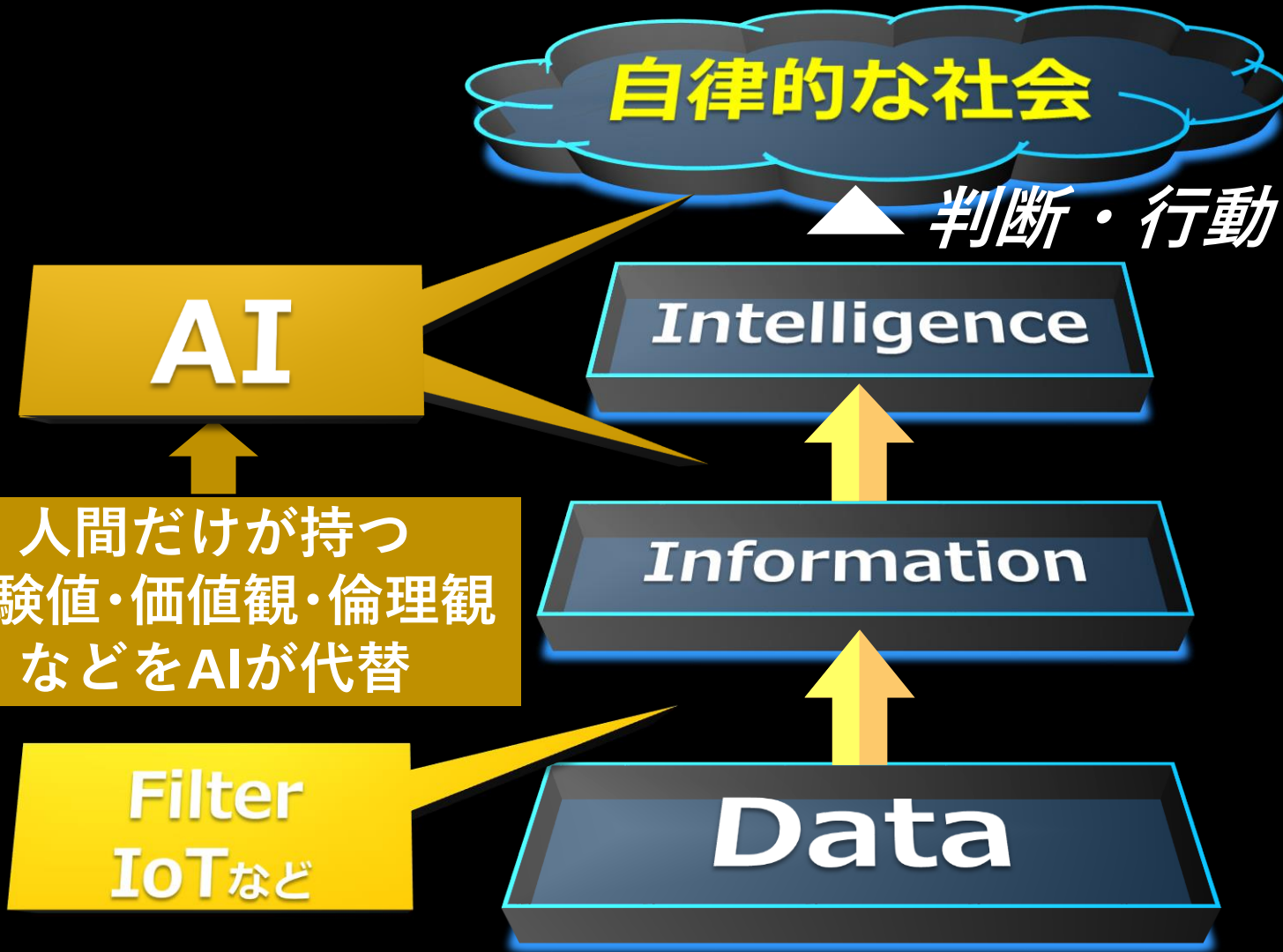
「情報の3階層」の フィルターには注意が必要

「情報の3階層」



AIが判断・行動まで行う

人が判断しなくても
自律的に動き出す
社会へ



A I のガバナンス

どのように考えるべきか？

Eurasia Group が予想する 4 つの AI リスク

(1) 偽情報 (Disinformation)

(2) 拡散 (Proliferation)

(3) 大量解雇 (Mass Displacement)

(4) 人間の代替 (Deeper/Complex Interactions)

科学技術の進化を
止める必要はない

ELSIを意識した
マネジメントが必要

積極的な A I 活用に向け
人類共通のルールが必要

ELSI

ゲノム解析プロジェクトの中から
約30年前に生まれた言葉

新たに開発された技術を、社会で実用化
する上で生じる「技術以外の課題」のこと

「E」 Ethical (倫理的)

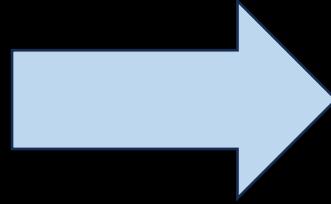
「L」 Legal (法的)

「S」 Social (社会的)

「I」 Issues (問題)

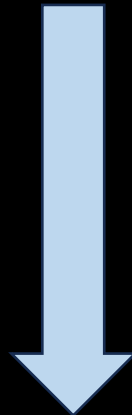
ELSI重要性の背景

最新テクノロジー



社会への実装

目的と結果のミスマッチ



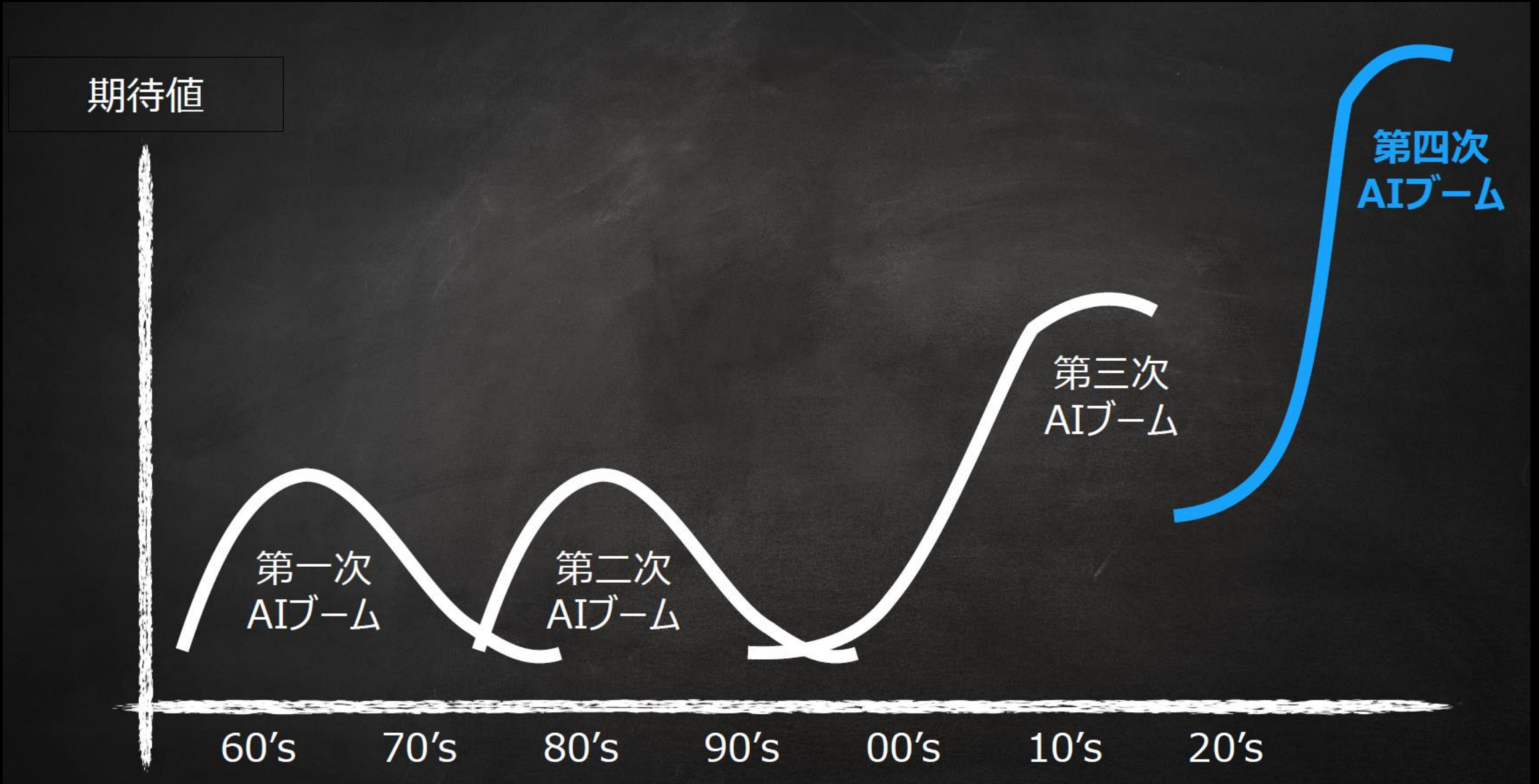
安全性の確保
責任の所在

社会に受け入れられるか
非差別や公平性の確保

発生可能性のある課題を予想して対応

企業における取り組み

勢いを増すAI



生成AIの爆発的な浸透

2021年1月
論文+デモ
CLIP
DALL-E

2022年4月
強化版デモ
DALL-E2

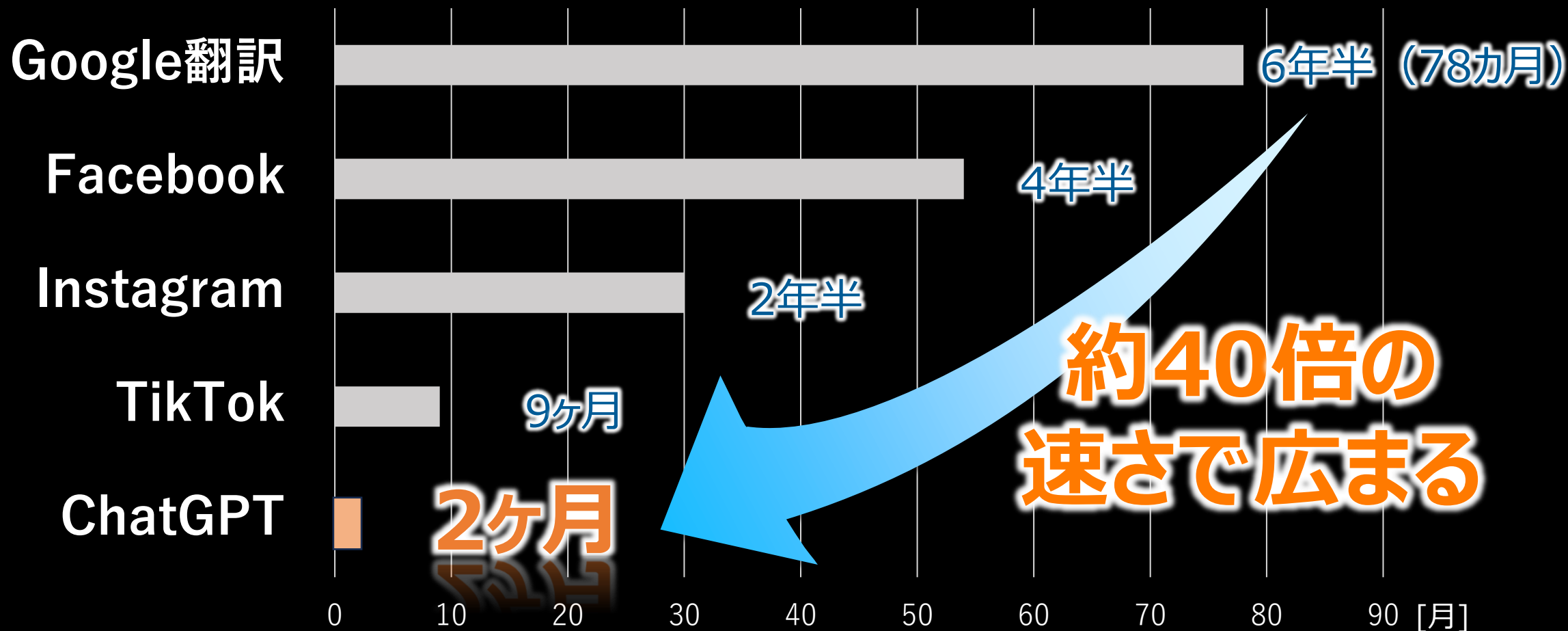
2022年8月
オープンソース
Stable
Diffusion

2022年10月
一大ブーム
DAU1,000万
派生続々登場

2022年10月
IT大手が続々
同等機能の
製品搭載を発表

生成AIの爆発的な浸透

月間のアクティブユーザー数（MAU）が1億人達成するまでにかった期間



これまでのAIと生成AIの違い

生成AIには従来のAIにはなかった「使い勝手の良さ」がある。

	これまでのAI (統計モデル、機械学習モデル etc.)	生成AI
学習	都度、専用データで学習が必要	事前に巨大なデータを学習済み
インプット	画像、数値、文字列などのデータ	「これをして欲しい」という 言葉での指示
アウトプット	推論結果 (予測、分類、スコア、検出、識別 etc.)	様々なコンテンツ (テキスト、画像、音声、映像)

AIを動かす知識と学習が必要

AIと対話しながら業務や創作活動が可能

企業におけるAIへの取り組みの方向性

企業は「積極的なAI活用の推進」と「AIガバナンスの徹底」の両輪で取り組み、ビジネスを拡大する必要性が高まっている



人の能力に迫るAI技術の登場

活用範囲の拡大
新たな価値創出が期待される

AI活用につまわるリスク

各国法規制の制裁金リスク
AIの不適切な利用による信用の低下

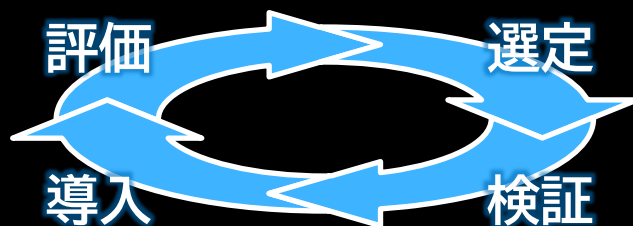


AIを活用したサービス

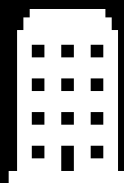
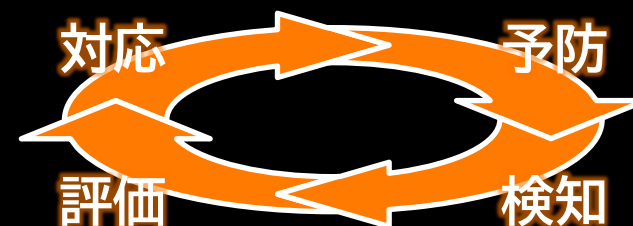
推進と統制の両輪



積極的な AI 活用の推進



AIガバナンスの徹底



企業

NTTデータグループの 生成AIへの取り組み

NTTデータグループのAIへの取り組み

NTTデータグループは、お客様のAIを活用した業務変革の実現に向け
「積極的なAI活用の推進」と「AIガバナンスの徹底」の両輪で取り組む

✓ 人の能力に迫るAI技術

司法試験等の難関試験を突破
一般消費者も気軽に利用可能に

社会情勢

✓ 活用にまつわるリスク増大

各国での法規制（制裁金）
不適切利用による信用低下

お客様

AIを活用した業務の変革

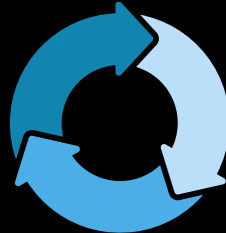
両輪でサポート



積極的なAI活用推進

先進顧客との事例創出

主要プレイヤーとの
戦略的
パートナーシップ



技術アセット
の整備



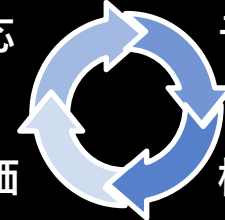
ガバナンスの徹底

対応

予防

評価

検知



AIガバナンス室

2023.4.1 新設

Global Generative AI Office

全社横断の組織を設置し、生成AIの取り組みを推進

Global Generative AI Office

グローバルレンジで生成AIの戦略推進・投資管理・施策実行

Driving Generative AI strategy

NTTDATA Group

NTTDATA Japan

公共

金融

法人

NTTDATA Inc. + OpCos

US

EMEAL

APAC

NTTデータの生成AIの取り組みの歴史



生成AIの取り組みを拡大させるため、アセットを開発、商用展開

IDP ナレッジ
マネジメント



チャットボットAI



コード生成・変換



インテリジェント
マネージドサービス



生成AIを活用した変革

自社のValue Chainの変革

ソフトウェア開発分野へ生成AIを適用

お客様のValue Chainの変革

顧客との協創を推進

戦略・ガバナンス

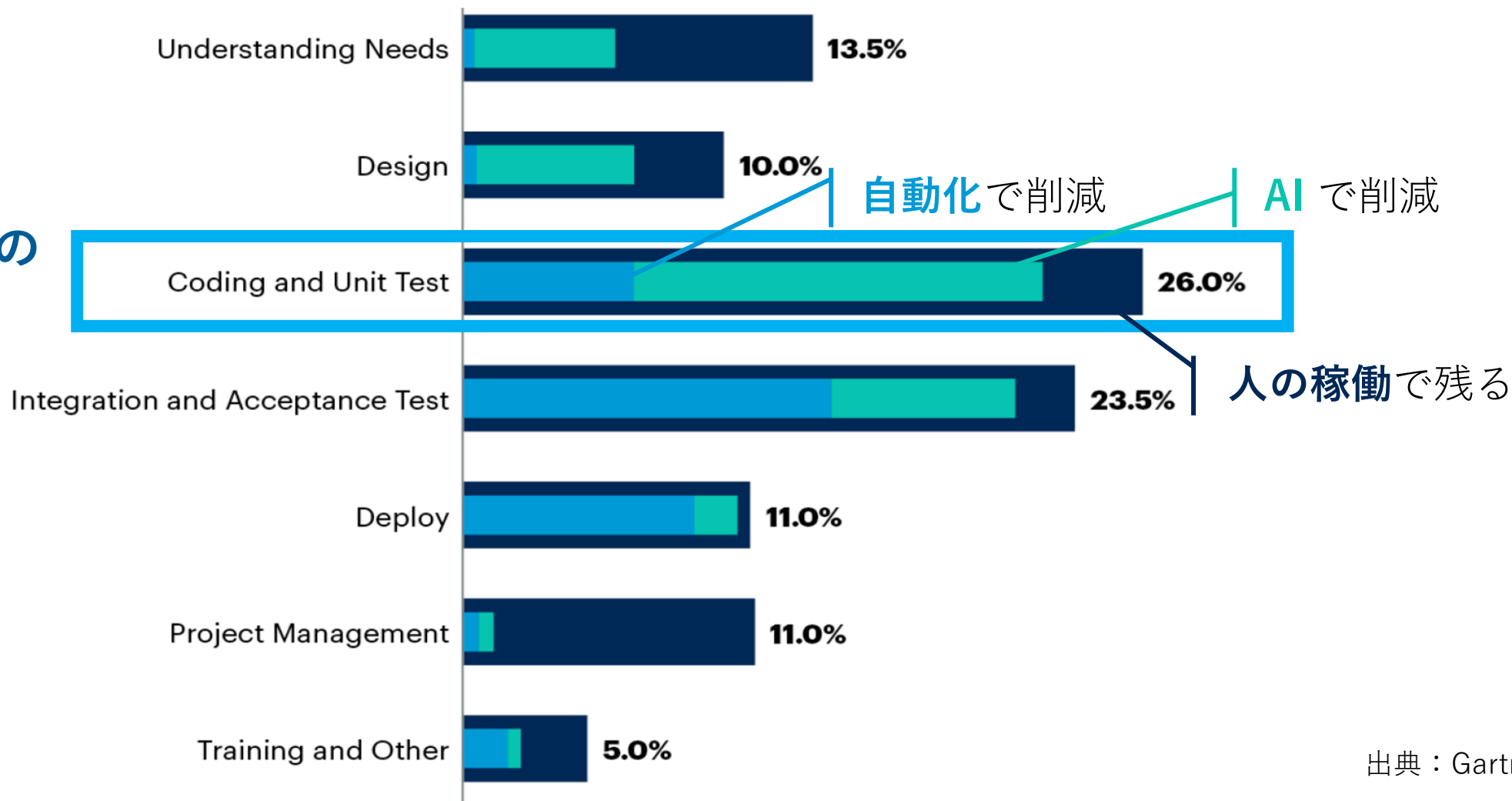
生成AI活用に向けたグローバルガバナンスの対応



AIを用いたソフトウェア開発技術の動向

AIをソフトウェア開発に適用することで、大幅な生産性向上が期待されている

コーディングへの
適用効果(大)



出典：Gartner

生成AI活用をソフトウェア開発領域全般へ拡大

要件定義

設計

コーディング/
テスト

運用・保守

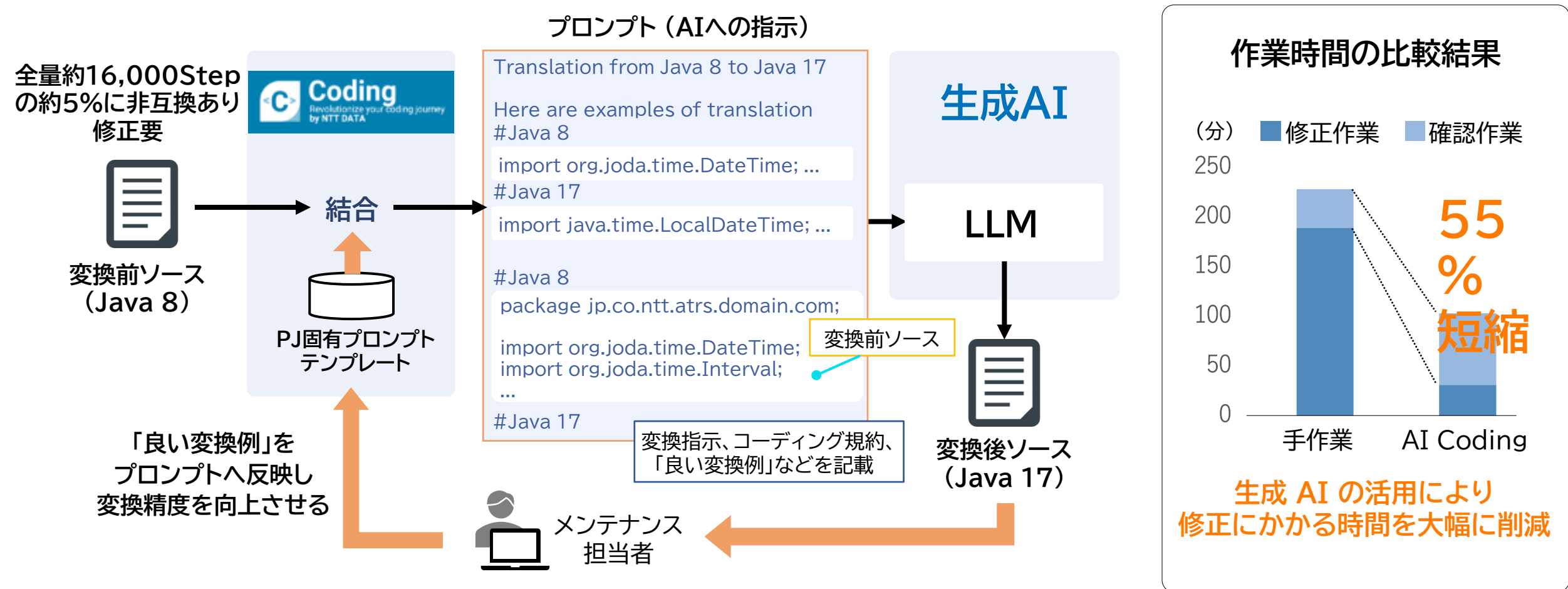
デモ | AIを用いたコーディング

(投影のみ)

【事例】AI-Coding 技術によるマイグレーションへの活用

～航空券予約システムにおける Javaバージョンアップ～

航空券予約システムの Java 8 から Java 17 へのバージョンアップ作業を生成 AI で実施。
手作業と比較して、55% 時間短縮（生産性 2.2倍）を実現。



生成AIを活用した変革

自社のValue Chainの変革

ソフトウェア開発分野へ生成AIを適用

お客様のValue Chainの変革

顧客との協創を推進

戦略・ガバナンス

生成AI活用に向けたグローバルガバナンスの対応



NTTデータの生成AIに関する取り組みの方向性

幅広いケイパビリティから
最適な生成AI活用価値を提供

労働集約型ビジネスから
AIによる価値提供ビジネスへ変革

(投影のみ)

軽量・特化型LLM tsuzumi

tsuzumi

専門知識をもった
小さなLLM

言語学習データの
質と量を向上

軽量

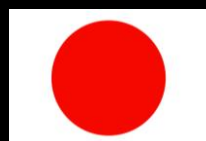
高性能

柔軟なカスタマイズ

マルチモーダル

0.6B~7B

GPT-3(175B)の
1/300



日本語は
GPT-3よりも優秀

英語は主要製品と
同等レベル

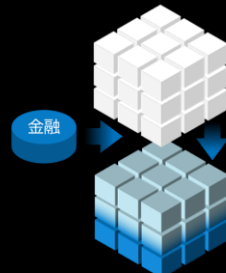
プロンプト
エンジニアリング

プロンプトに
金融関連情報を付加



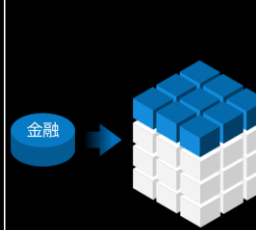
tsuzumi

フルファイン
チューニング



tsuzumi

アダプタ
チューニング

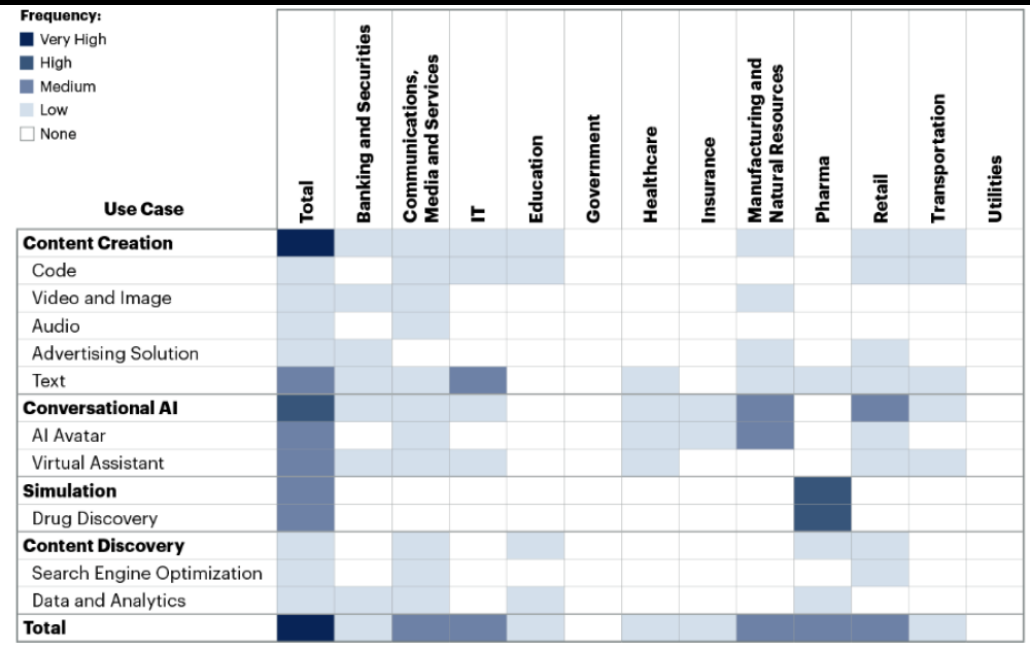


tsuzumi



注目ユースケース・インダストリ

GenAI インダストリー別ユースケース



出典：Gartner

注目のユースケース・インダストリ

ユースケースとしては、
文章生成・検索、対話が大きな活用機会

インダストリとしては、
主にメディア、IT、製造、医薬、小売
での活用が増加

特に医薬業界における生成AIの
初期アプリケーション化が顕著

生成AIを活用した変革

自社のValue Chainの変革

ソフトウェア開発分野へ生成AIを適用

お客様のValue Chainの変革

顧客との協創を推進

戦略・ガバナンス

生成AI活用に向けたグローバルガバナンスの対応



AI活用に求められる評価視点

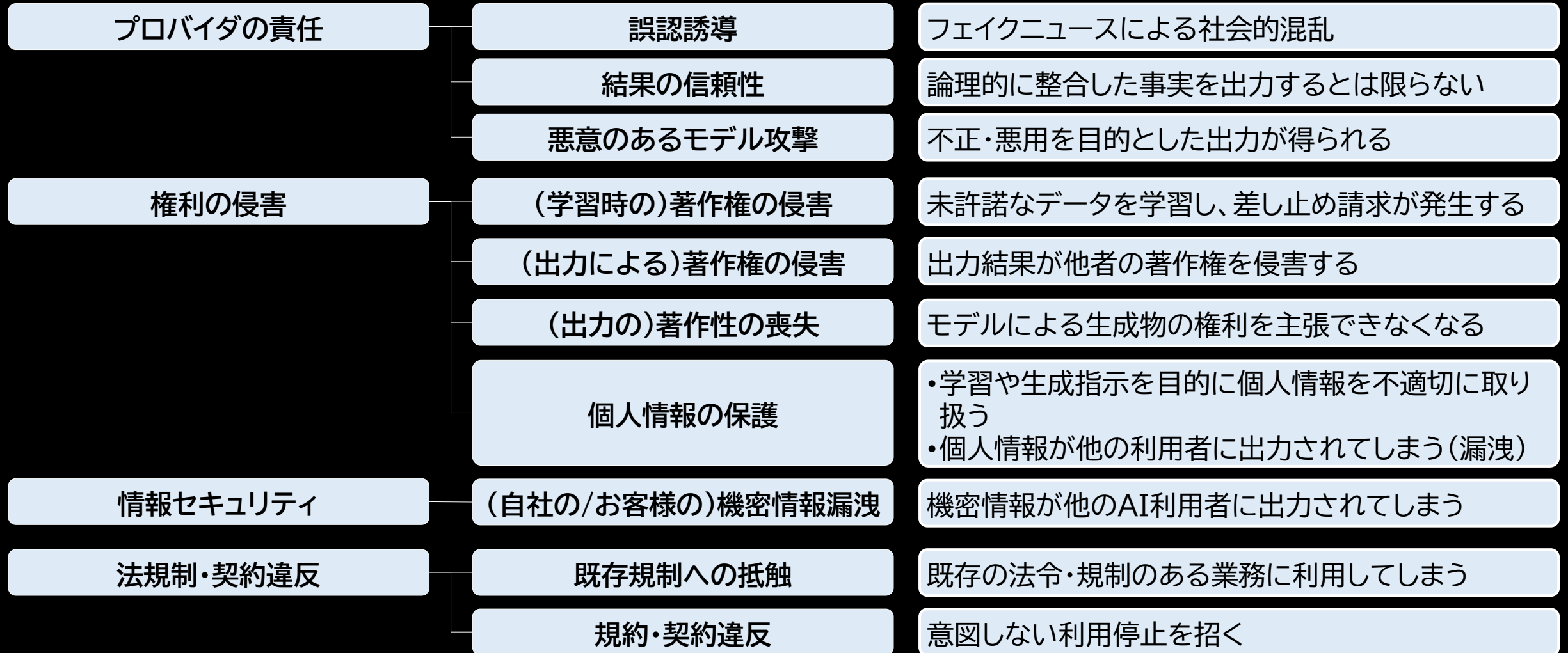
AIの活用範囲の拡大により、社会的な視点や
倫理的な視点での評価も重要視されている



AI活用の評価視点の変化

生成AIによって生じるリスク

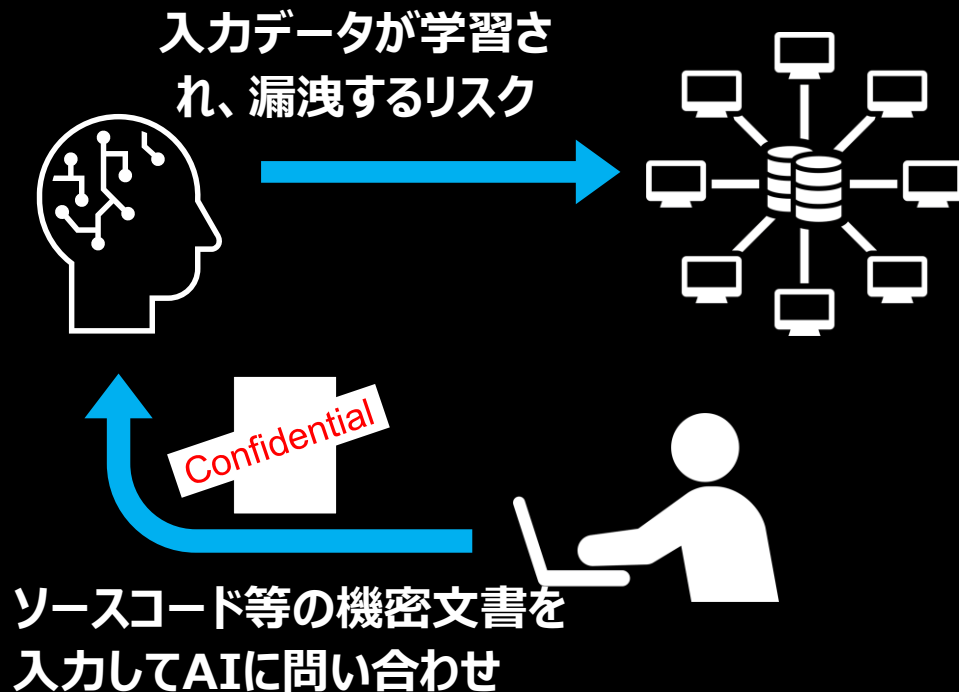
生成AIはコンテンツを入力に、コンテンツを出力できる破壊的な技術だが、社会への負の影響も大きく、利用・提供には注意を必要とする



生成AIのトラブル事例

生成AI活用で業務の効率化が期待できる一方で、利用に際しては**仕組みを理解した上でのサービス選定**や**利用する社員のリテラシー**が無ければ**情報漏洩**につながる**リスク**もある。

Samsungが機密情報をChatGPTにリーク (韓：2023)



【概要】

社員がソースコードをChat GPTへ入力しエラー解消を依頼。その後Samsung社はChat GPTの社内利用を禁止。

【影響】

- ・ソースコードが学習され流出するリスク
- ・信頼低下によるブランドイメージダウン

出展：サムスン、従業員の生成AI利用を禁止 – ChatGPT経由でデータ漏れる - Bloomberg

AIガバナンス室の新設

2023年4月1日にAIガバナンス室を設置し、全社のAIガバナンスを強化

AIガバナンス室設立のプレス（3/23）

NTT DATA

PICK UP

サービス

業種別ソリューション

ブログ DATA INSIGHT

企業情報

サステナビリティ

IR情報

ホーム > ニュース

AIのリスクマネジメント強化を目的 とした「AIガバナンス室」の新設

報道発表

2023年3月23日

株式会社NTTデータ

株式会社NTTデータは、AIの不適切な利用によって生じるリスクを実効的にマネジメントし、AIの適正活用を推進するための組織として、AIガバナンス室を2023年4月1日に設置します。

近年、AIの不適切な利用により生じる、倫理・社会受容性に起因するトラブルや法規制の罰則は、お客さまビジネスに大きく影響をおよぼすようになっていきます。本組織は、当社が提供するAIシステムおよび当社内システムを対象に、AIによる事業リスクを検知するルール・体制を整備し、プロジェクトと一体となって対応することで、社会とお客さまがAI活用の恩恵を最大限に享受し、サステナブルな社会を実現できる環境を整備していきます。

ニュースについて

ニュースに掲載されている、サービス内容、サービス・製品の価格、仕様、お問い合わせ先、その他の情報は、発表日現在の情報です。その後予告なしに変更となる場合があります。また、ニュースにおける計画、目標などは様々なリスクおよび不確実な事実により、実際の結果が予定と異なる場合もあります。あらかじめご了承ください。

ニュースメール配信

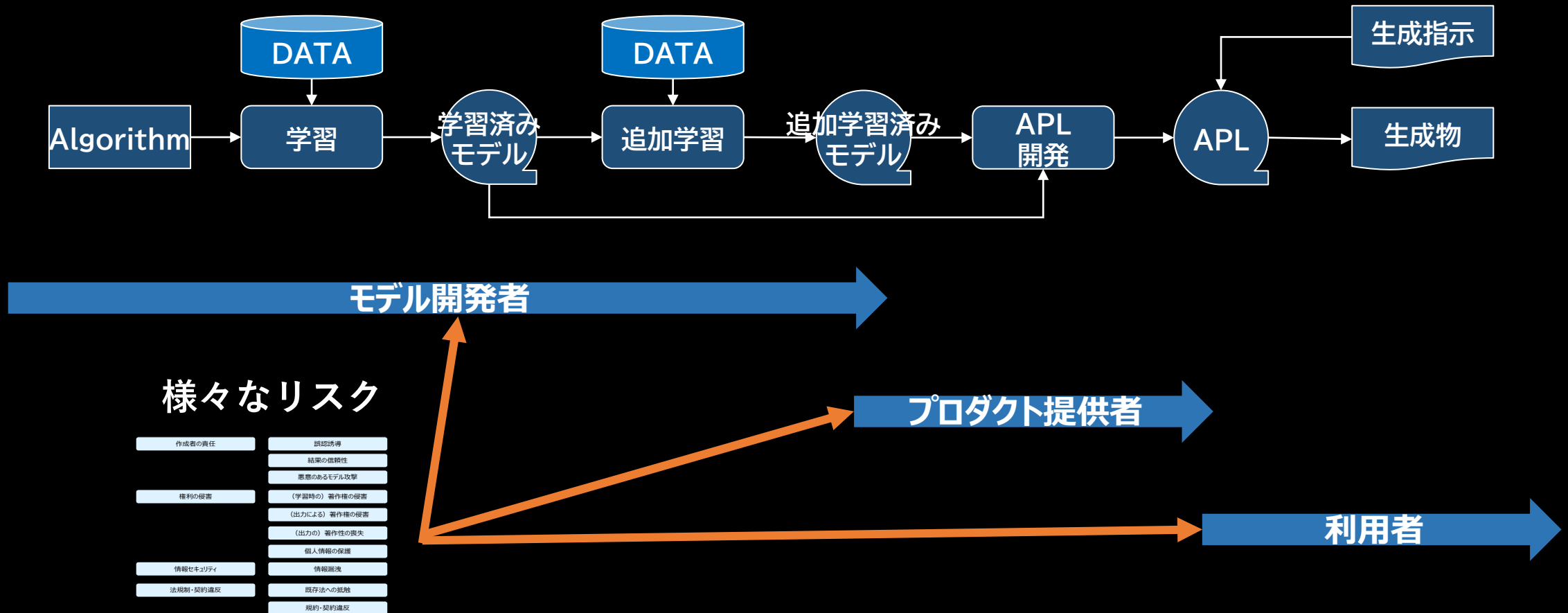
ニュースの更新状況をいち早くお知らせするた



03-5561-1111 | 03-5561-2222 | 03-5561-3333

利用の「立場」を明確化

バリューチェーン上で、自組織がどこにあたるか？ポジションに応じてリスクも異なる



企業ガバナンス：解決に向けた取り組みの方向性

AIガバナンス方針の策定

組織的メカニズムの確立

企業がとるべきアクション



- 技術、法務、倫理、消費者など**多様な観点・専門性を持つメンバ**による協議会を**設置**（AIアドバイザリーボード）
- 協議会での議論結果を**リスクマネジメントに反映**



- **AIリスクマネジメントのポリシー策定**
- **策定したポリシーを組織規程へ組み込み、ルール化**



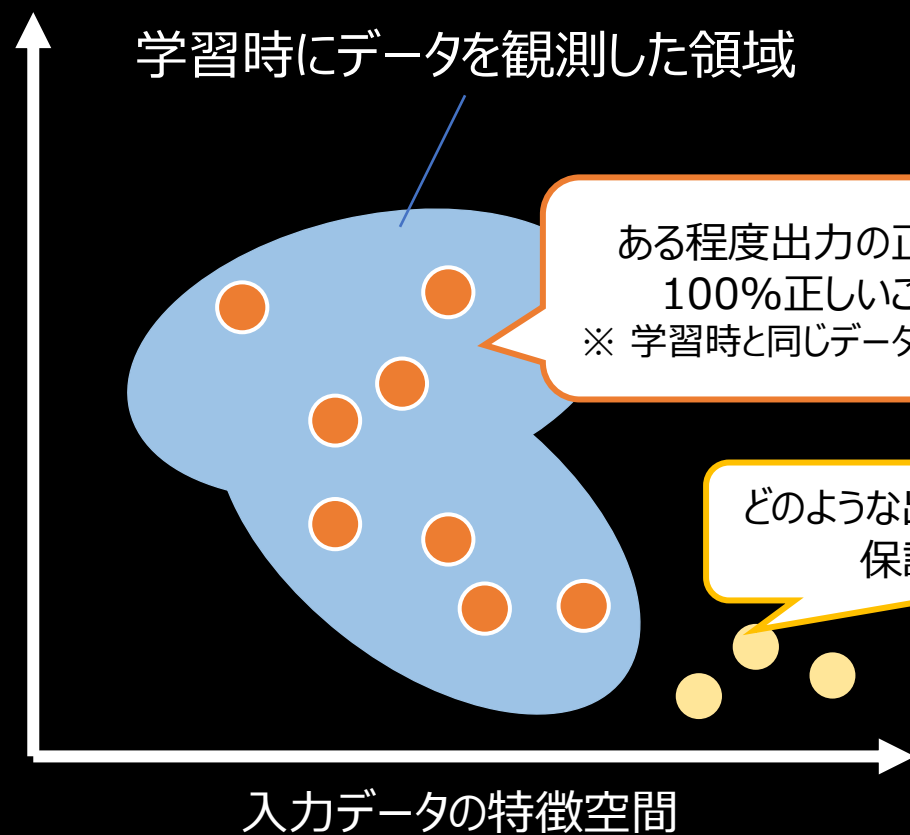
- AIガバナンスに関する研修プログラムを整備・展開し、全社員・職員が受講
- **策定したルール・プロセスに関する社内セミナー開催**

AIガバナンス推進に関わるファクションの整備

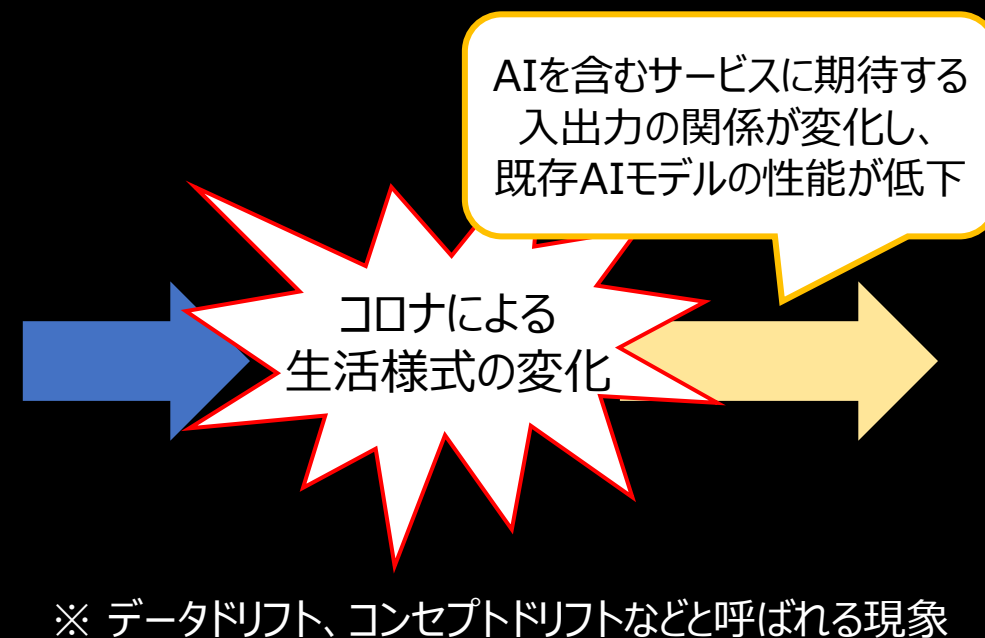
クオリティコントロールへの影響

AIに関するサービスレベルの概要

不確実性があるため通常と異なる目標設定と監視、改善が必要



そのほか、外的要因によって性能低下する例



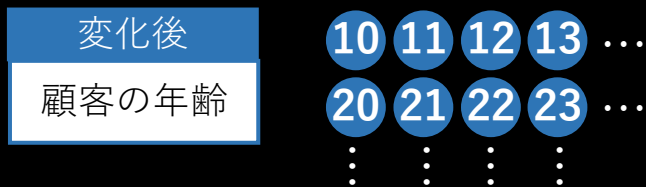
【参考】データ構造とデータの性質の変化の監視

データ構造の変化

関係システムの更改やセンサなどの変更に伴う
入力データの型やカテゴリの変化



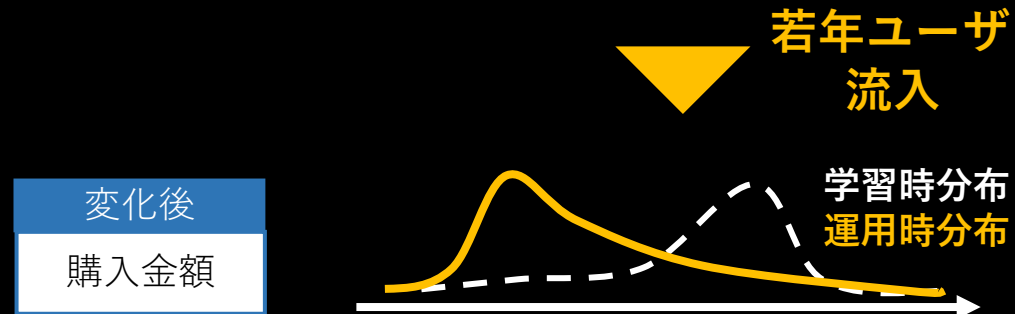
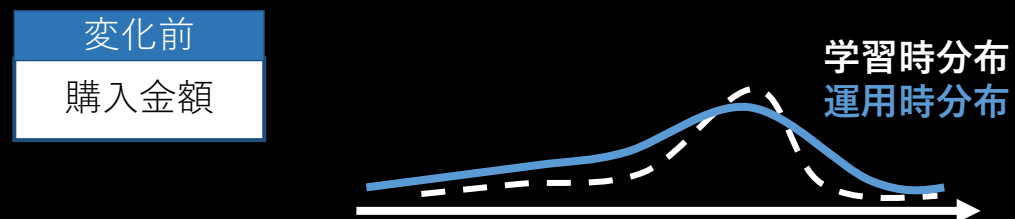
会員管理
システム導入



学習時と出力結果が一致するか確認

データ分布の変化

外部要因やデータ入手元の変化に伴うデータの
傾向変化 ※値の変化、特定カテゴリの増加等



学習時の分布情報と運用時の分布情報を比較

AI運用作業の全体像

AIシステムは従来システム同様の保守・運用に加え、主に下記作業が必要

従来システム保守運用作業

システム監視

サーバ再起動

障害対応

問合せ対応

パッチ対応

アップデート

各種手順書対応

機能追加対応



AIシステム固有の保守運用作業

サービス
レベル管理

機械学習モデルの性能低下及び精度低下につながる予兆がないかを監視

変更管理

データ分析、再モデリング、サービング

構成管理

AIモデルと必要情報を管理し、必要なモデルを再現可能な状態を確保

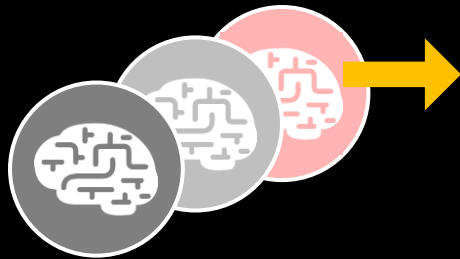
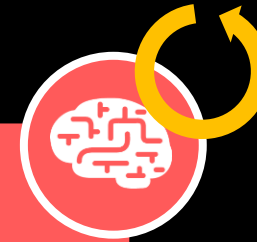
サービスレベルの低下を検知した場合の対応策

データ加工やアルゴリズムの見直し等の試行錯誤を
最新データでやり直す

※ データサイエンティストが必要

AIモデルの実装は流用しつつ最新データで
再学習

※ アルゴリズムによっては長時間



過去に学習したAIモデルのうち、最新データでの
性能が最も良いものに差し替える

AIモデルの出力にルール等の後処理を追加し、
サービスの出力を制御

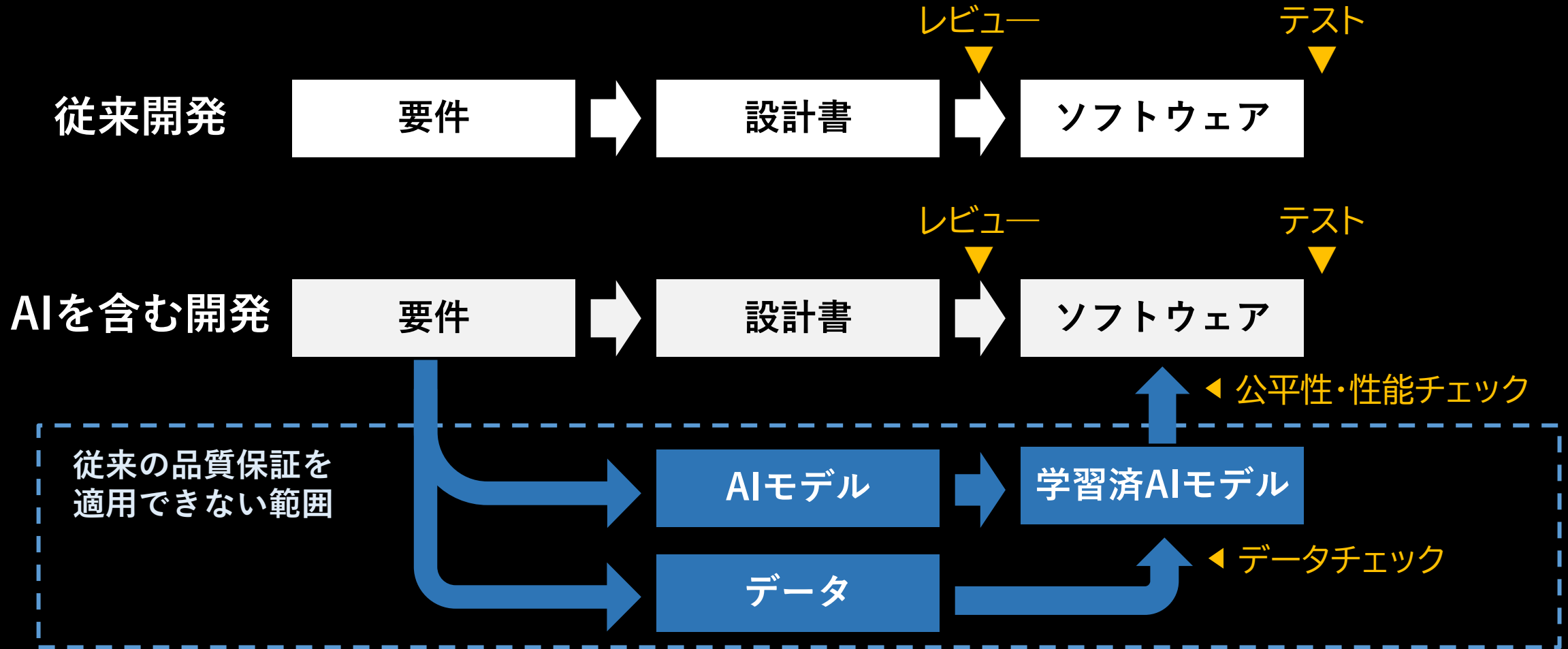
※ ルールの管理が必要

If ●● then ▲▲

期待される改善効果は大きい
が実施に係るコストが高い対策

実施に係るコストが低い
が根本的ではない対策

AIに関する変更管理の概要

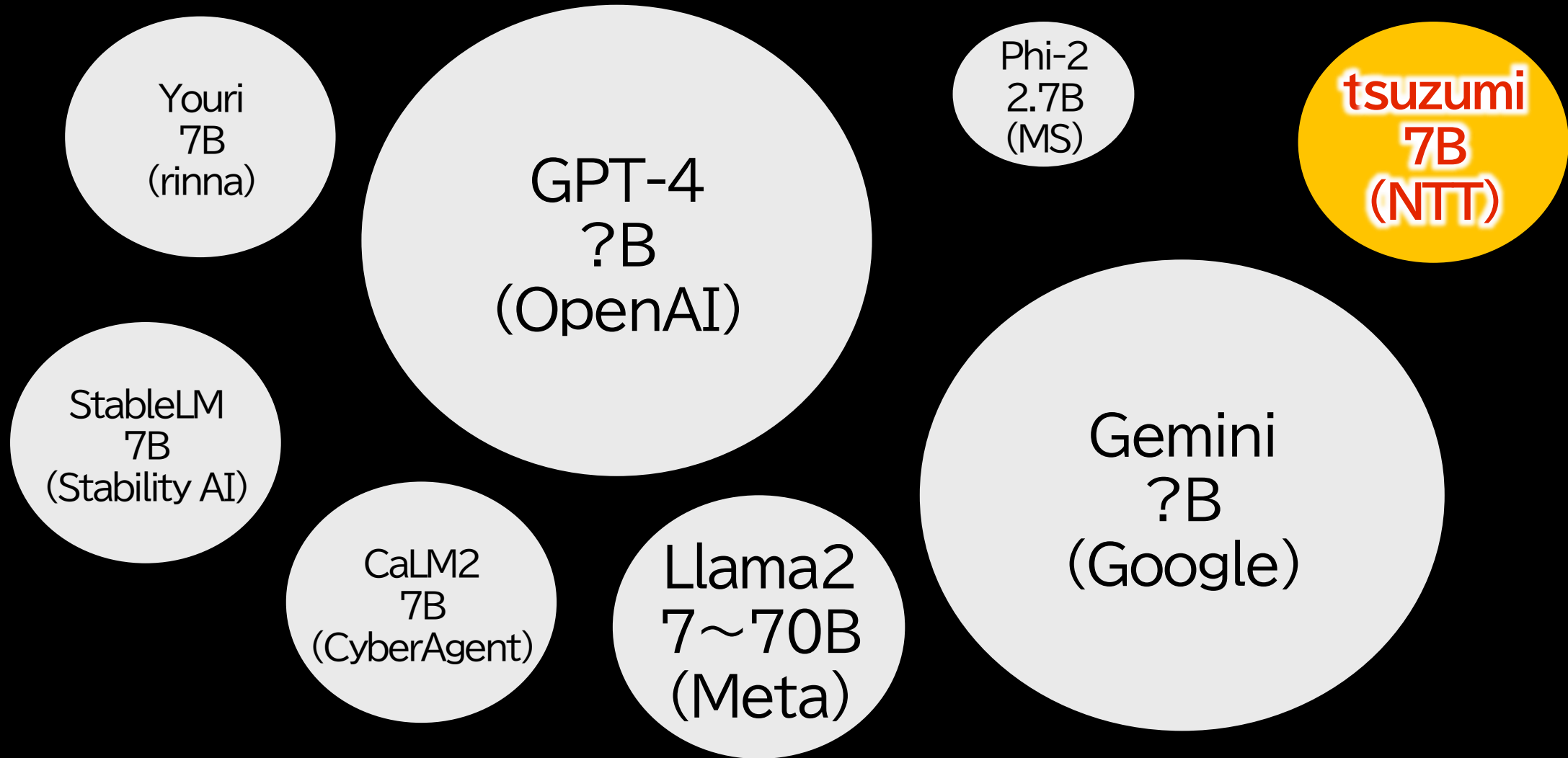


AI固有のテスト観点+リリース判定におけるAI固有の考慮が必要

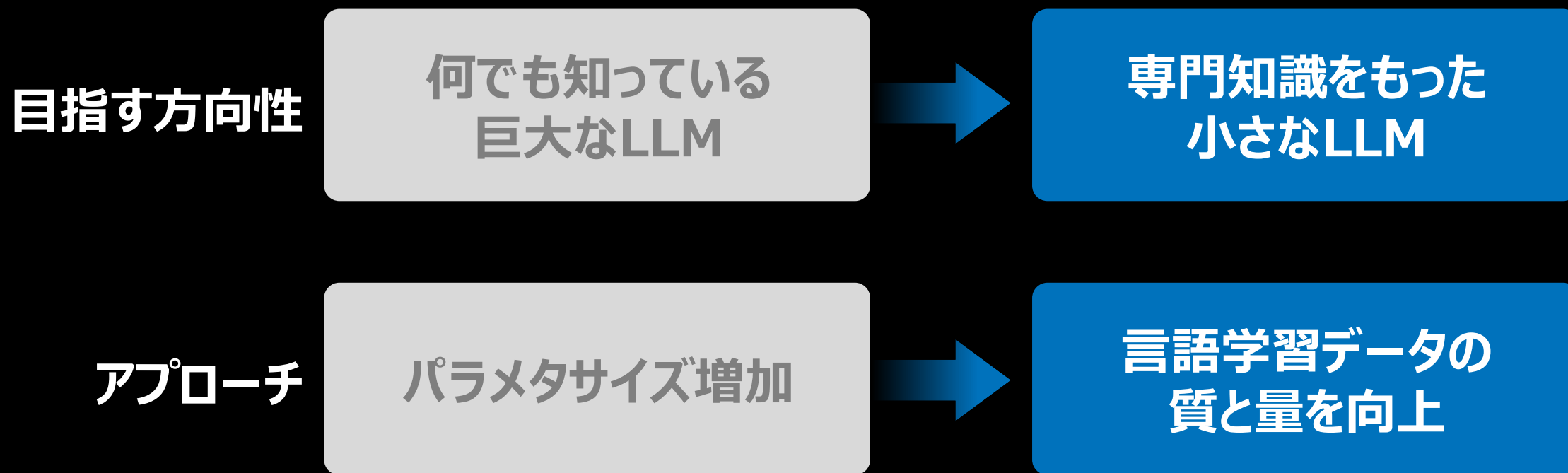
NTTの大規模言語モデル 「tsuzumi」とは

1. 日々生まれてくる生成AI(言語モデル)

ハイパースケーラーやITベンダ、スタートアップから日ごとに新しい生成AIが登場している



2. tsuzumiの戦略



3. tsuzumiの特徴

- 1 40年以上の自然言語処理研究による蓄積
日英対応、日本語は世界トップクラス
- 2 超軽量で高性能
高いコストパフォーマンス
- 3 柔軟かつ低コストでチューニング、クローズドデータも学習
カスタマイズ、最新化
- 4 日本初！マルチモーダル(図表ドキュメント視覚読解)
図表読解等様々な形式に対応

4. 日本語性能比較：Rakudaベンチマーク

tsuzumi-7Bは、世界トップクラス、国産LLM中トップの性能を達成(評価スコア：1225／2023.10.26時点)
大規模なGPT-3.5を上回り、同クラスの国産LLMを大きく上回る

tsuzumi

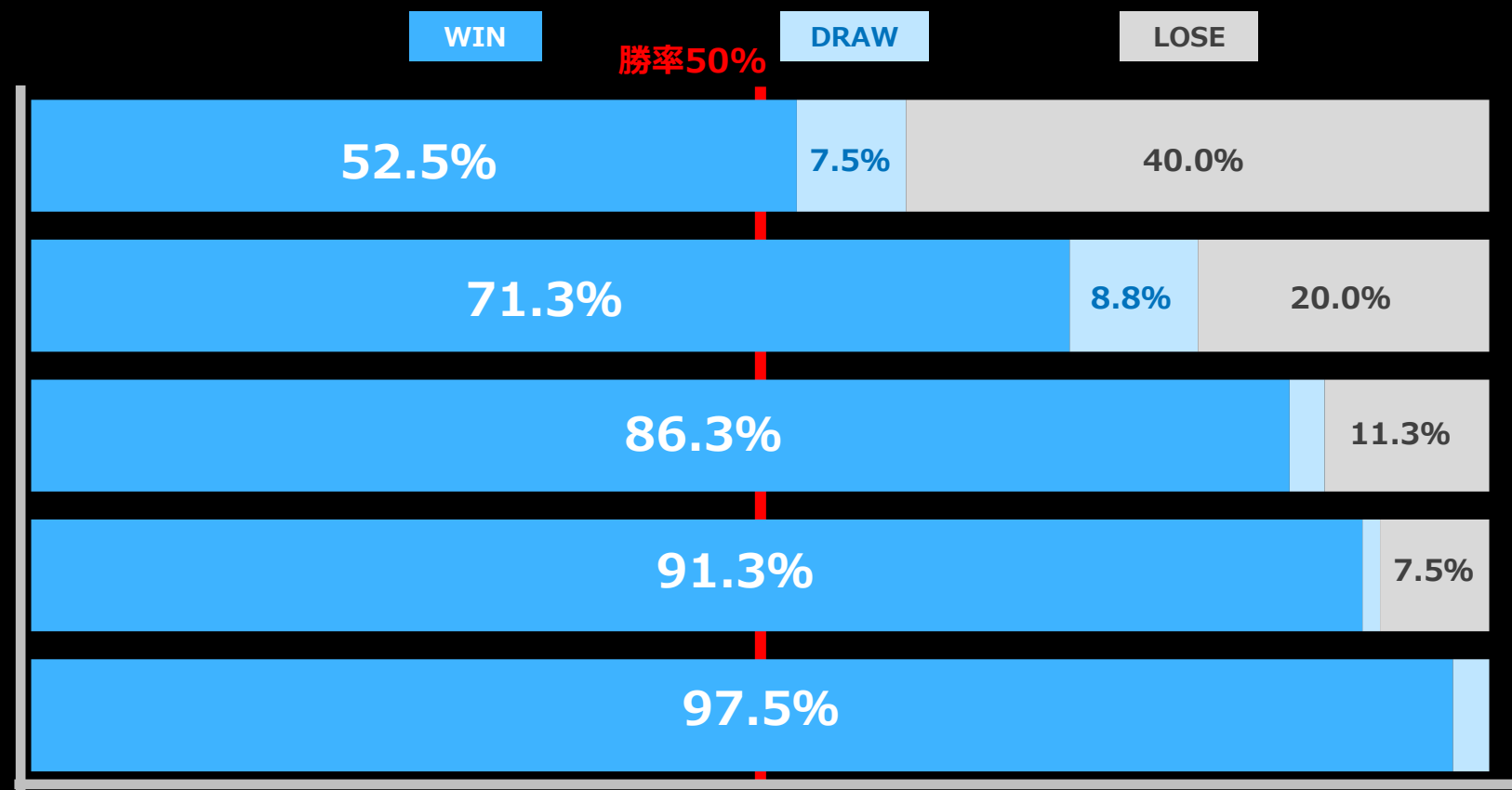
対 GPT-3.5 20B? 175B?
(OpenAI gpt-3.5-turbo-0301)

対 Elyza 7B
(ELYZA elyza-7b-fast-instruct)

対 Ja Stable 7B
(Stability AI ja-stable-7b-alpha instruct)

対 Weblab 10B
(東大松尾研 weblab-10b-instruct)

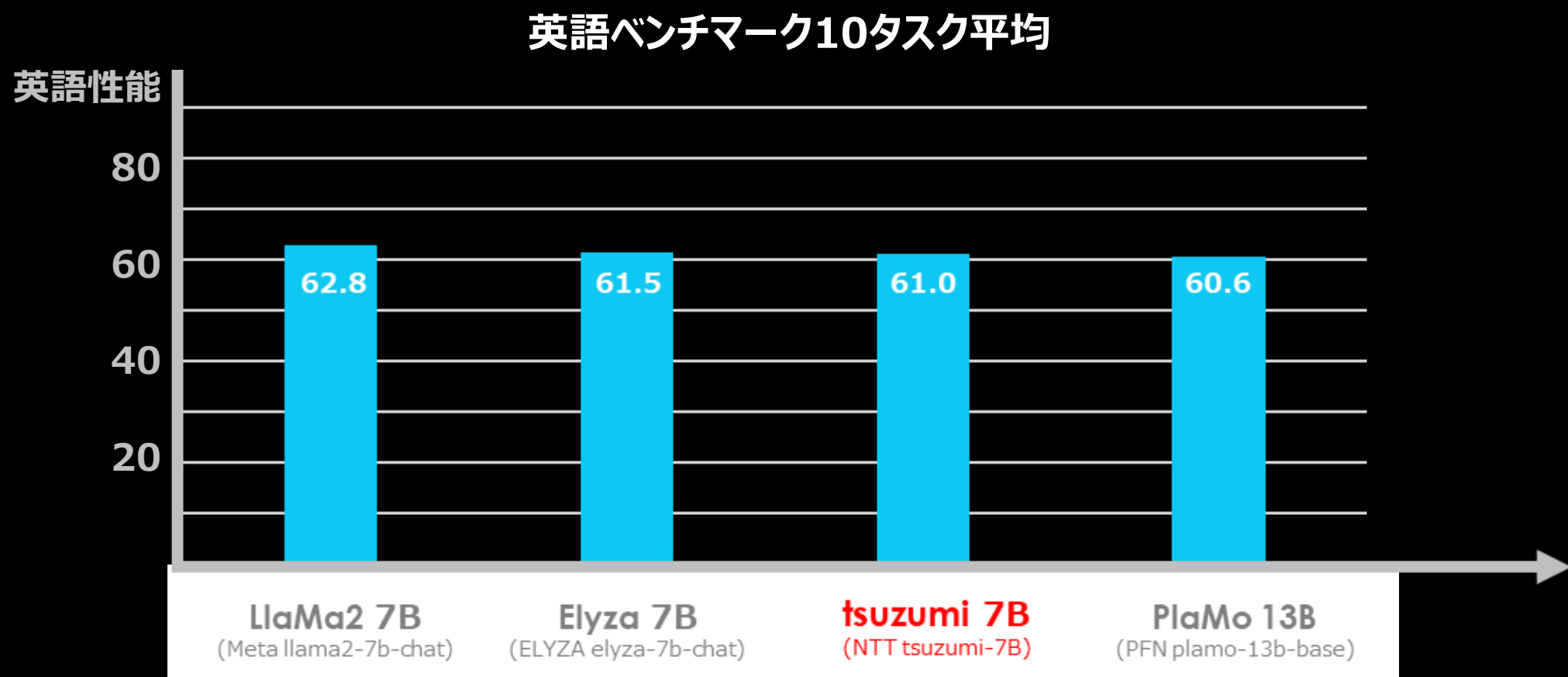
対 LLM-JP 13B
(Nii LLM勉強会 llm-jp-13b-instruct-full-jaster-dolly-oasst-v1.0)



※rakudaベンチマーク： <https://yuzuai.jp/benchmark> 2024.10.22実施
日本の地理・政治・歴史・社会に関する40問の質問 GPT-4による2モデルの比較評価（40問×提示順2）で採点
llm-jpを除くモデル出力はサイトにアップロードされているものを利用 llm-jpはhuggingfaceのモデルカード記載の設定による
入力の繰り返しおよび終端トークンは後処理により除外した。
評価スコアは、2024/09/27付リーダーボード記載の全モデルとtsuzumi-7bをGPT-4による2モデルの比較評価を行い、Bradley-Terry strengthsにてランキングした結果

5. 英語性能比較 : lm-evaluation-harness

同サイズで英語をメインとした世界トップクラスLLMのLLaMA 7Bと同程度の性能を実現
英語以外にも、中韓仏独など多言語・プログラミング言語を拡充予定



※1: 評価方法
lm-evaluation-harness: <https://github.com/EleutherAI/lm-evaluation-harness>
rinna/bilingual-gpt-neox-4bのモデルカード記載の英語10タスク（commonsense reasoning分野）におけるスコアの平均値
評価指標はaccおよびacc_norm（両方存在する場合はacc_normを優先）
llama2-7b, elyza-7b, tsuzumi-7bは全てchat/instructモデルを利用 plamo-13bはbaseモデルの数値

* PLaMoのみベースモデルの結果

6. 高いコストパフォーマンス

軽量のLLM **tsuzumi** を2種類開発

超軽量版

tsuzumi-0.6B

GPT-3 (175B) の
約**300**分の**1**

軽量版

tsuzumi-7B

GPT-3 (175B) の
25分の**1**

7. 軽量な生成AIのメリット

推論・学習コストが低いためオンプレでも利用可能。また、消費電力が低いためサステナビリティにも配慮

推論・学習コスト

動作環境

消費電力

大規模なモデル

高い

クラウド前提

高い

小型・軽量なモデル

低い

クラウド/オンプレ

低い

【参考】消費電力

サステナビリティ

GPT-3 (175B) 規模の学習には、膨大なエネルギーが必要

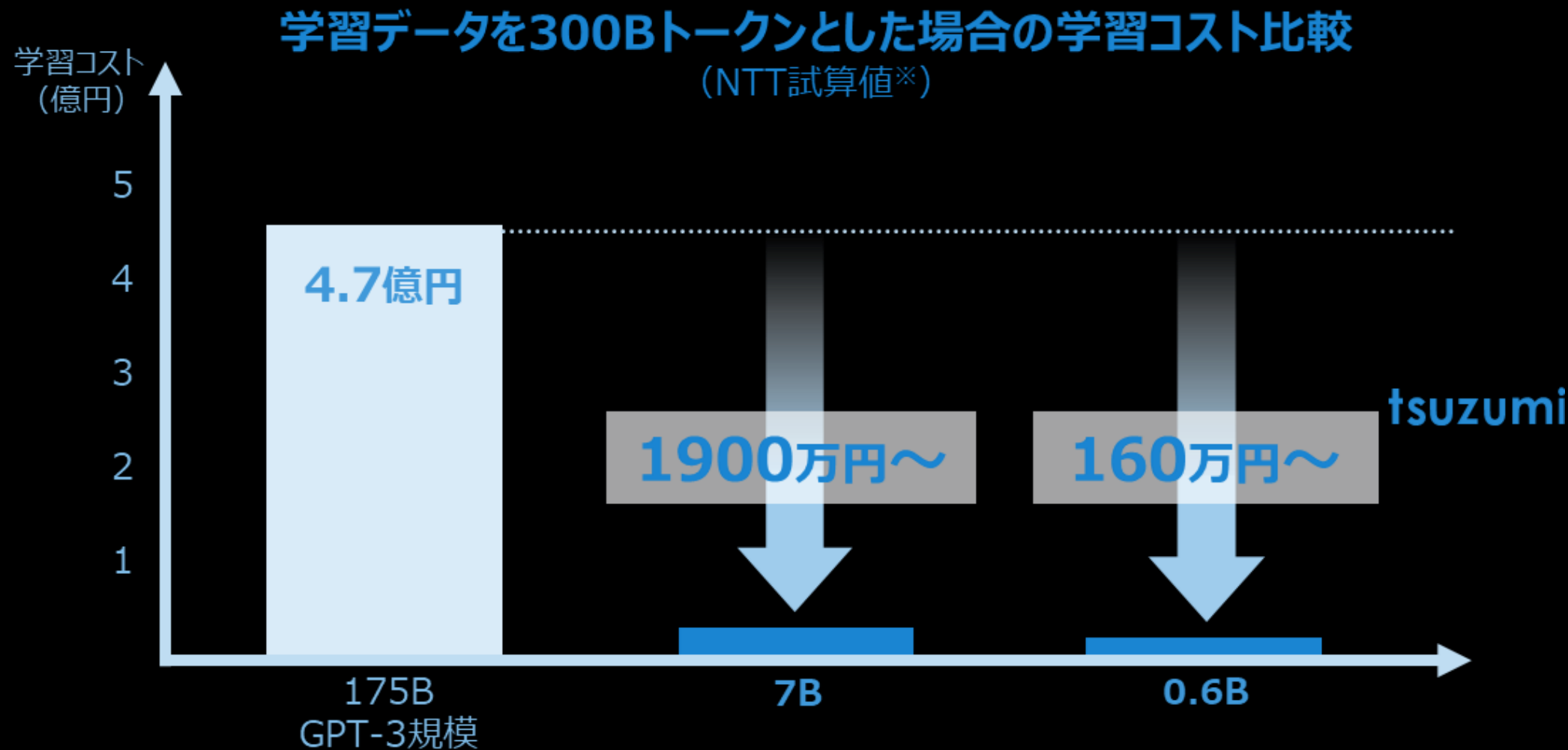
約 **1300MWh**_[1] / 1回の学習



原発1基分 約 **1000MWh**

8. 学習コスト

GTP-3規模のLLMと比べて、学習コストを約300分の1（超軽量版）および約25分の1（軽量版）に低減可能

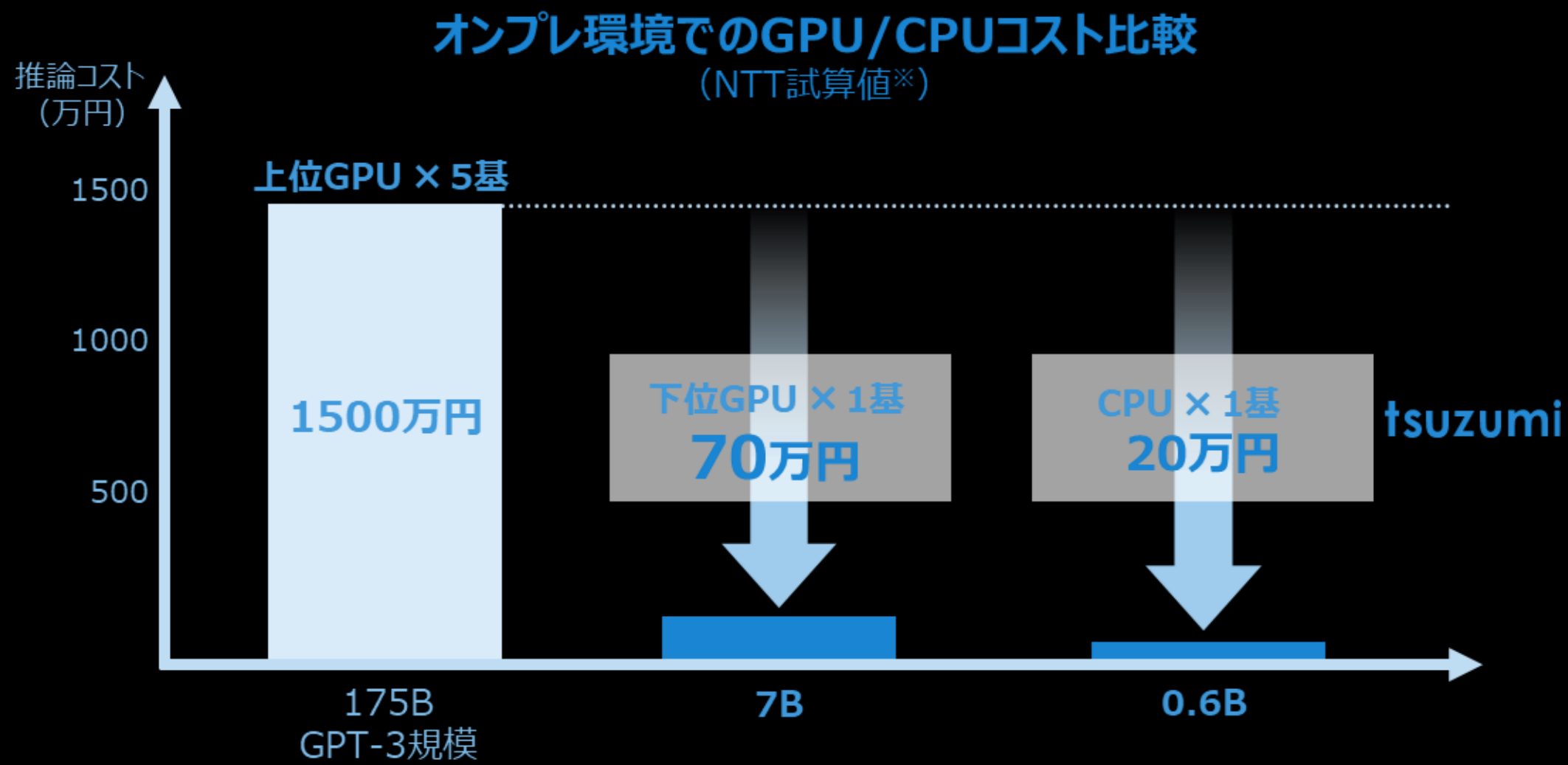


※試算条件

- LLaMA-1 7Bの学習時82432GPU-hoursをベースに、パラメタ数比とトークン数比から各LLMの必要GPU-hoursを算出
- 算出したGPU-hoursとAWSのGPUクラウド料金から学習コストを算出
- AWS GPUクラウド料金: A100-80GB 1ノード (8GPUs) 約14万円/日と想定
- 通常、パラメタサイズが小さい場合、精度を向上させるためには、2-3倍程度の学習データが必要。それに比例しコストも向上

9. 推論コスト

GTP-3クラスと比べて、推論コストを約70分の1（超軽量版）と約20分の1（軽量版）に低減可能



※試算条件
・量子化: 16ビット
・必要GPUメモリサイズ: パラメタ数 × 量子化サイズ/8bit (175Bは350GB, 7Bは14GB, 0.6Bは2.4GB)
・ハードウェアコストは、上位GPU A100 80GB: 300万円/台, 下位GPU A10 24GB: 70万円/台, CPU PC: 20万円/台として換算、その他の運用などの費用は含まず

10. tsuzumi vs GPT-4 比較デモ

質問に対して十分な品質の回答をGPT-4よりも速く生成できる

Q.なぜtsuzumiは回答品質が高いか？

A. tsuzumiは金融特化の知識を
フルファインチューニング+アダプタチューニング
の両方で学習することができる

	フルファインチューニング	アダプタチューニング (ファインチューニング)
Azure OpenAI (GPT3.5/GPT4)	×	○
tsuzumi	○	○

※今回のデモはフルファインチューニング+アダプタチューニングの両方を実施

Q.なぜtsuzumiは回答が速いか？

A. 回答を生成する際に利用する
パラメータが少ない

	トークン生成速度(秒)	文字生成速度(秒)
tsuzumi	46.0	81.3
GPT-4	×3 (14.1)	×6 (12.3)
GPT-4 turbo	以上 (11.1)	以上 (10.1)

※『Q.日本のエネルギー政策と環境保護のバランスについて、現状と改善策を提案してください。』に対する、3回実施した場合の回答平均速度比較。

※トークン=単語、句読点などのデータの最小単位

11. さまざまなチューニング方法

精度やコストなど要件の違いに柔軟に応える、3つのチューニング方法を提供

プロンプトエンジニアリング

コスト ◎
精度 △

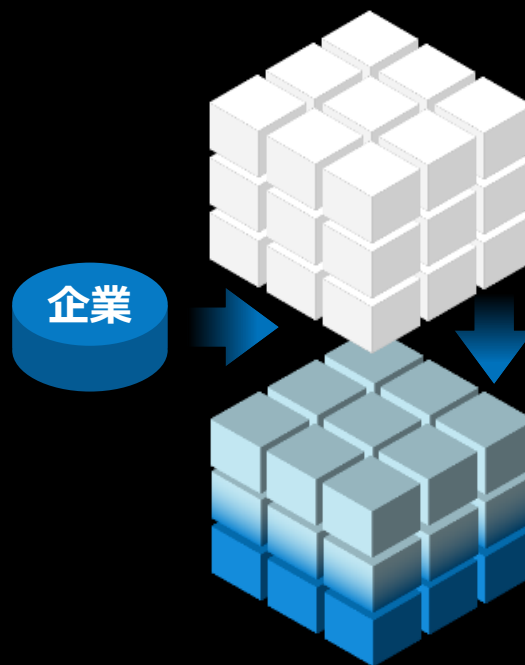
プロンプトに
企業関連情報を付加



tsuzumi

フルファインチューニング

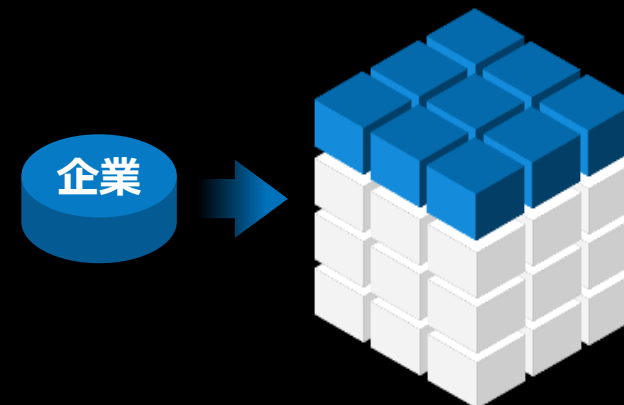
コスト △
精度 ◎



tsuzumi

アダプタチューニング

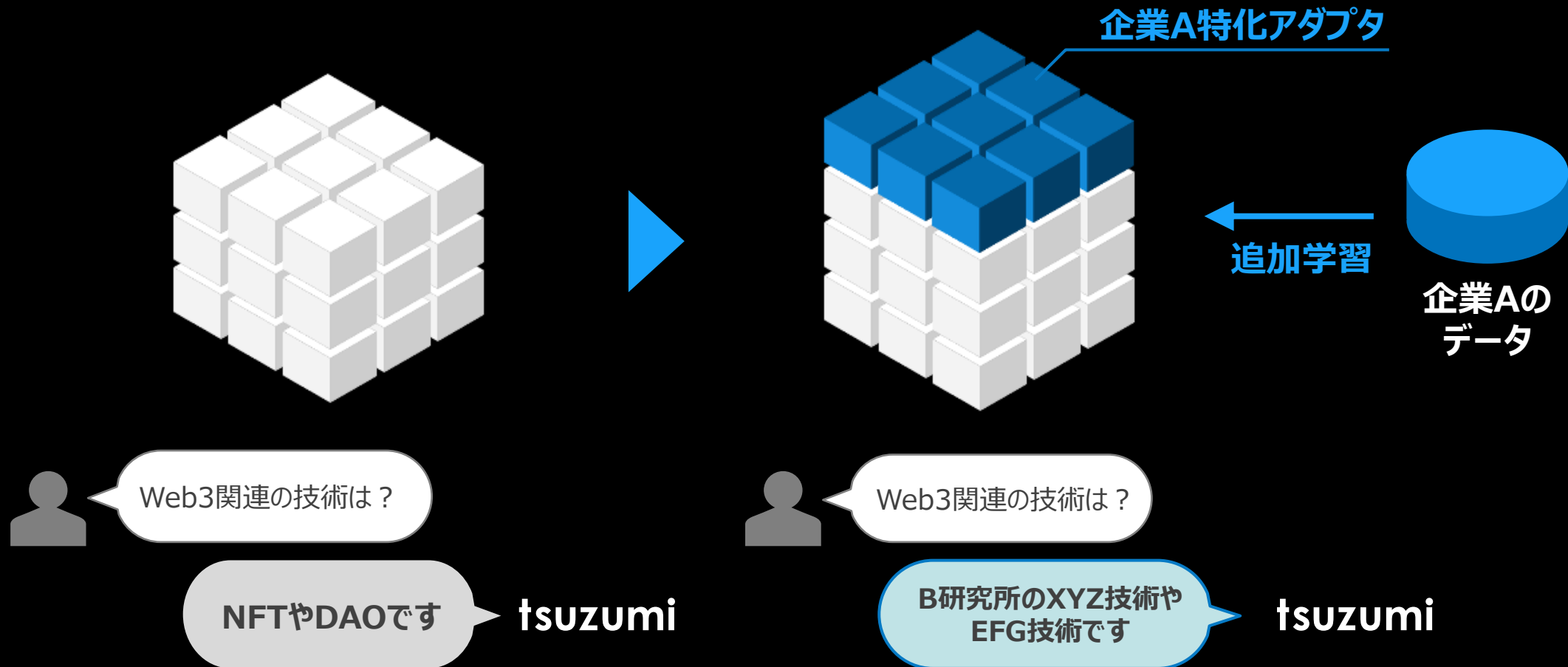
コスト ○
精度 ○



tsuzumi

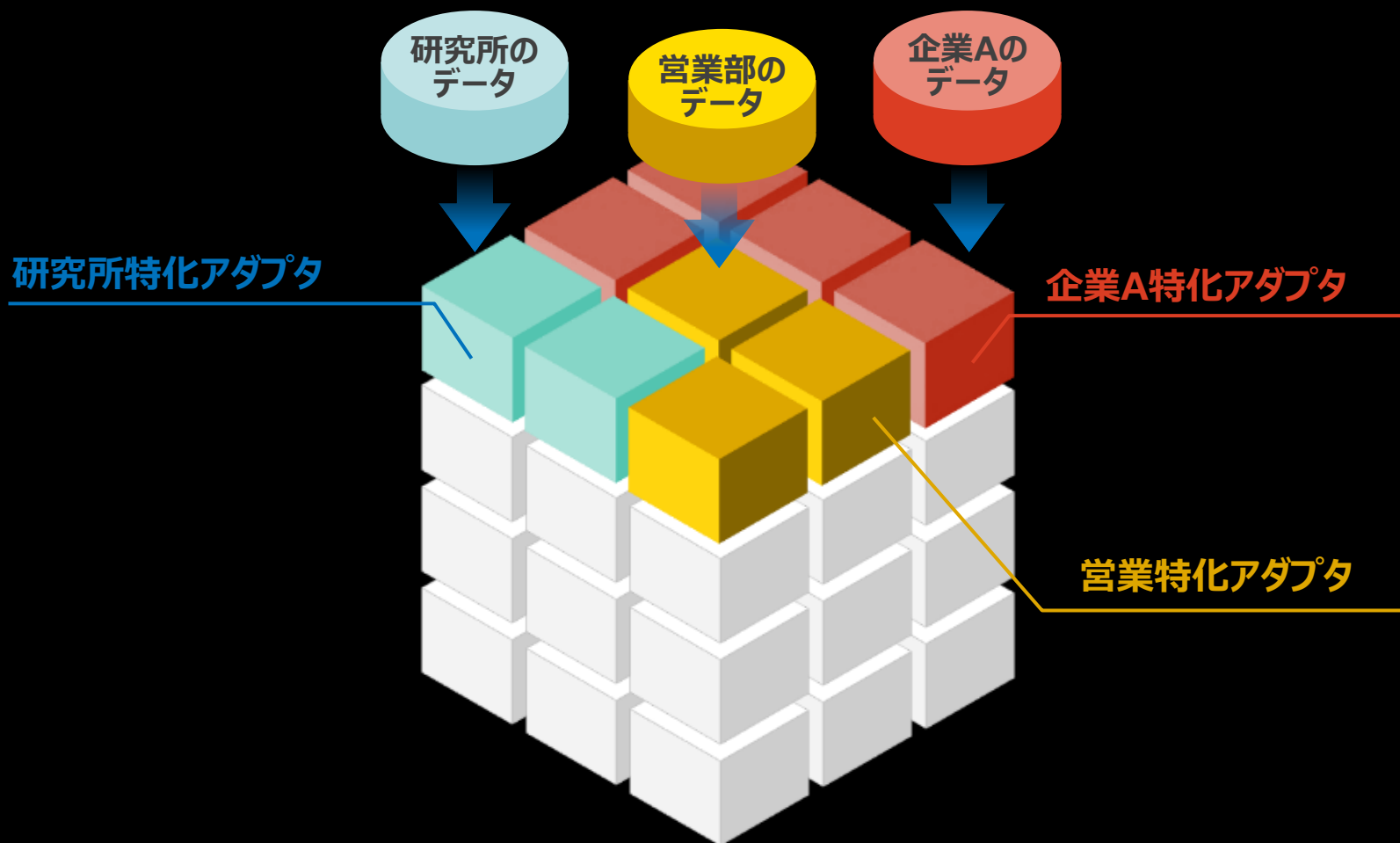
12. アダプタチューニング

組織ごとのカスタマイズを低コストで実現



13. マルチアダプタ

基盤モデルを複数アダプタで共有し、利用ユーザやシーンに応じてアダプタを切り替えたり、組み合わせたり可能



tsuzumi

NTT DATA

14. マルチモーダル

言語による質問だけでなく、文書画像を提示しながらの質問が可能
請求書や仕様書など画像付き文書を用いる業務やRPA業務に適用可能

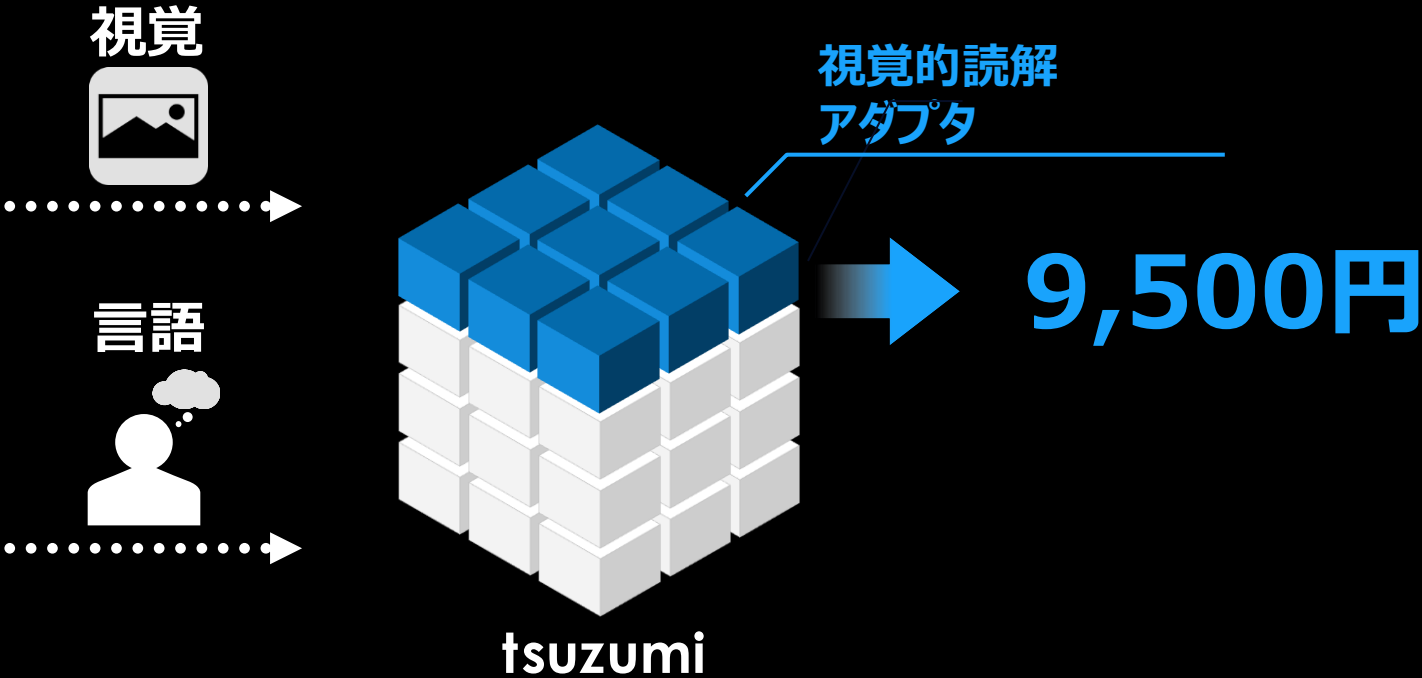
ABC株式会社
山田 太郎 様

発行日：2022年10月12日
株式会社XYZ
〒242-0000
神奈川県〇〇市〇〇
〇〇〇ビル
TEL: 012-345-6789

下記の通り、お見積り申し上げます。

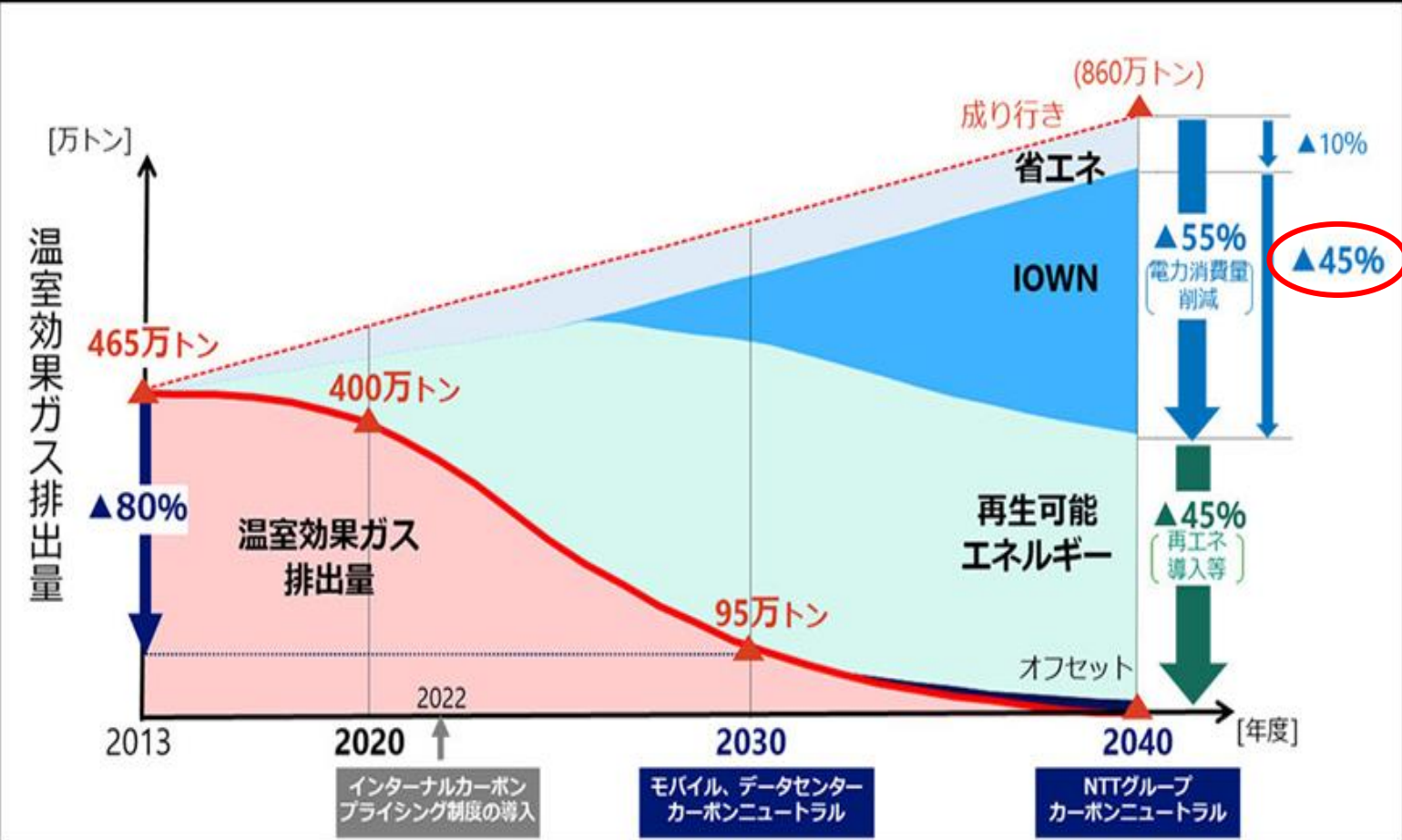
合計金額（税込）10,450 円

品名	数量	単価	金額
商品A	1	3,000	3,000
商品B	3	1,500	4,500
商品C	1	2,000	2,000



10%の消費税を除いた
合計金額は？

15. 視覚読解モデルの実例



2040年のIOWNの電力消費量削減の割合は何%でしょうか？

Q

Infographic 呼出 クリア 実行

A ▲45%

16. 今後のスケジュール



17. tsuzumiをご提供するための準備を進めております

tsuzumiをお客様業務で活用いただけるようにそのノウハウ蓄積を進めております

tsuzumiが**活かせるタスク**って何？

ファインチューニングは実際どうやるの？

タスク特化・ドメイン特化のための
学習データは？

システムへの組み込み方法は？

NTT DATA