生成AIの台頭で大きく注目されるAIガバナンス
～欧米のAI・デジタル動向の考察と示唆～

【前編】

2024年9月24日

株式会社日本総合研究所　先端技術ラボ
株式会社三井住友フィナンシャルグループ
シリコンバレー・デジタルイノベーションラボ

＜本件に関するご照会先＞
田谷洋一　株式会社日本総合研究所 シニアエキスパート 兼 JRI America, Inc. Executive Director（Ytaya@jri-america.com）
緒方雄二　株式会社三井住友フィナンシャルグループ 部長代理　（yuji.ogata@smbcgroup.com）

はじめに

●近年、AIの社会実装が進んでおり、様々な業界におけるAI活用が注目されている。とりわけChatGPTなどの生成AIの登場によって、日本国内でも生成AIの活用に向けた議論が盛り上がっている印象を受けるが、米国などのデジタル先進国の状況と比較すると、国内のAI活用が大きく進展しているとは言い難い。

●AIを広範に利用し、新規事業の創出や企業価値の向上に繋げるためには、AIの正確性や安全性を十分に担保することが肝要になるが、日本企業にとってAIのリスク管理がAI導入を進める上で最も大きな障壁となっている。一方、米国の動向で注目されるのは、多くの組織においてAIのリスク管理やセキュリティ対策、公平性確保のためのガバナンス戦略の立案が、AIのユースケースの検討と並行して行われている点である。

●背景として、米国ではAIに関連する連邦政府の政策が活発化し、AIリスクやガバナンスを対象にした法規制の整備が加速している点が挙げられる。欧州の規制と比較すると、米国の法規制はAI導入を促進するためのインセンティブが多く盛り込まれているため、企業のガバナンス強化を促すだけでなく、企業が自社の基準に基づいてAIを活用する領域を適切に選択し、リスクをコントロールしながら積極的に実装を進めていく効果も期待できる。

●本稿では、AIを取り巻く近年の動向や問題意識に基づき、主にリスクやガバナンス戦略をテーマに、日米のAI関連施策や法規制の比較、米国企業の事例紹介などを通して、日本企業がAI活用を一層進めるためのヒントを模索していく。なお、本稿は、前編・後編の2部構成とし、前編ではガバナンスの動向や法規制の動向に関する整理、後編では米国企業の事例や日本のAI戦略の発展に向けた考察を行う。

目次

第１章
生成AIの発展とともに重要性を増す
企業のAIリスク管理とガバナンス体制の強化

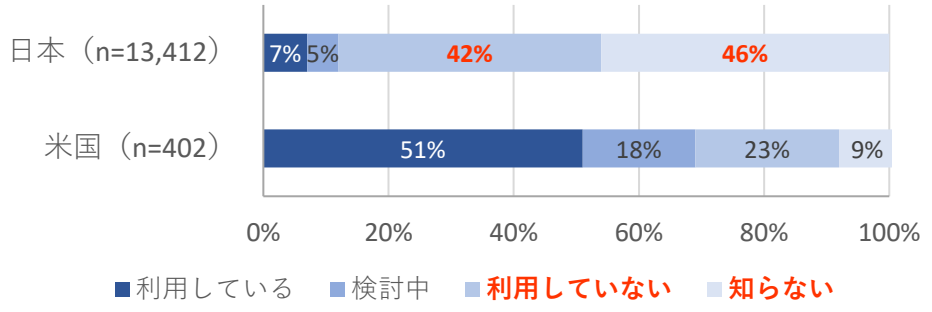
１．米国で進むAIリスク、ガバナンス整備施策

- 多くの企業で生成AIが注目されるようになっているものの、生成AIにおける日米の利用率には実際には非常に大きな差がある
- 日米で大きな差異が起こる要因としてまず指摘されるのは、生成AIの利用や導入への経営層の関心度合いの違いである
- また、データや資産、情報管理などのリスク管理・ガバナンス対策が生成AI利用を推進をする上での大きな障壁となっている

生成AIの利用率で大きく開く日米の差

- 多くの企業でDXが推進されるなか、業務改善や事業創出などのデジタル関連施策において、AIが果たす役割は大きいと指摘されている。とりわけChatGPTに代表される生成AIには大きく関心が集まり、日本国内でも導入検討が進んでいる。
- しかし、実際にChatGPT利用**に関する日米の状況を見ると、日本の利用率が7%であるのに対して、米国は51%と、日米の利用率には非常に大きな差がある**（図表１）。
- **日米差異の要因として、経営層の関心度が指摘されている。**米国では６割以上の経営層がChatGPTに対して強い関心を持っているのに対し、日本は米国の半分以下であった。

図表１：日米におけるChatGPTの利用率比較（2023年）

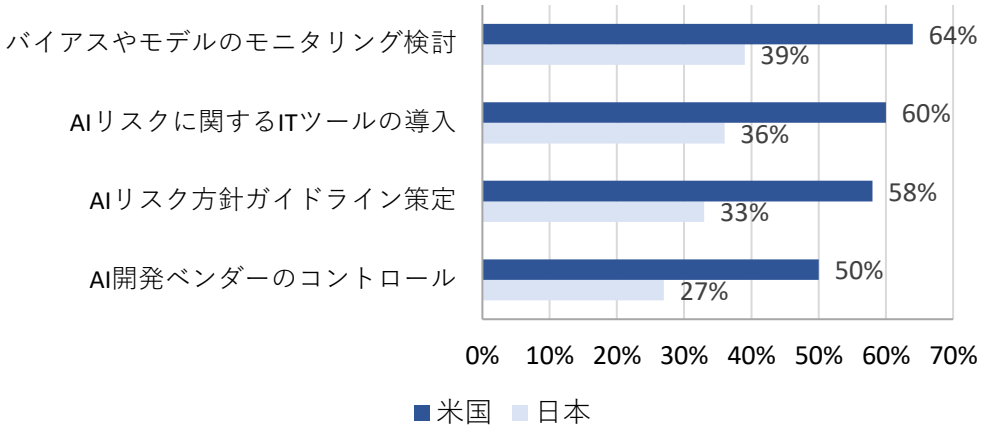


（参考）https://www.m2ri.jp/release/detail.html?id=580を基に作成

米国ではセキュリティルールと併せたAIの導入検討が進む

- 日米の差異に関するもう一つの注目すべき要因は、**日本企業にとって誤情報の拡散や顧客のデータ保護などのリスク管理やガバナンス対策が、AI導入を推進をする上で大きな懸念材料となっている点**である
- 米国では、**半数以上の企業がAIのモニタリングやリスク対策、ガイドライン策定などに着手しているが、日本では、それぞれの施策を推進する企業がおよそ半分の割合である**（図表２）。

図表２：日米におけるAIリスクへのガバナンス施策の取り組み状況



（参考）https://www.pwc.com/jp/ja/knowledge/thoughtleadership/2023-ai-predictions.htmlを基に作成

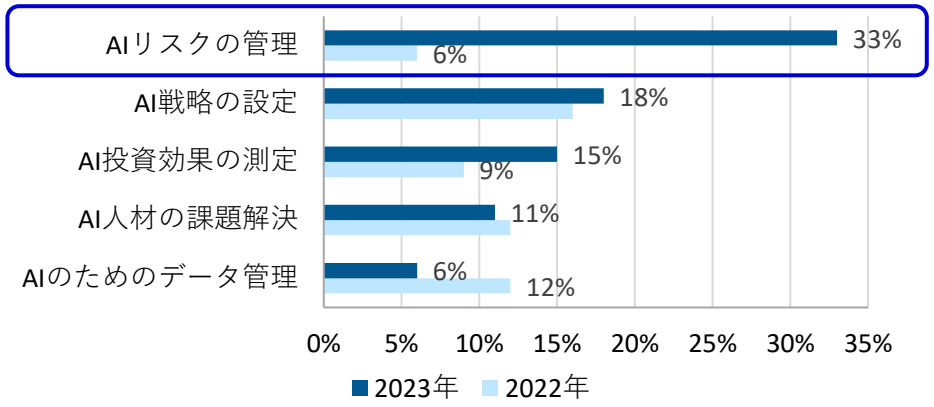
２．日本企業のAI推進の障壁として指摘されるリスク管理

- 日本国内でAIの導入検討が進むなか、その有用性だけでなく、**AIがもたらすリスクについての意識が急速に高まっている**
- 一方、日本企業のAIリスク管理における組織やルール、プロセスの整備状況を見ると、大半が今後整備を始める段階にあり、現在はAIリスクの内容を洗い出し、各種リスクへの対策を検討している状態だと考えられる

AI戦略を進める上で最優先課題となるリスク管理

- AI導入に関する日本企業の課題を見ると、**AIリスクの管理が最優先事項として挙げられている**（図表３）。2022年はAI戦略の立案や投資効果の測定が優先施策として挙げられていたものの、**2023年には、AIリスクの管理が他の施策と比較して急速に優先度が上がっていることが分かる**。
- 背景として、生成AIの利用検討が進むとともに、様々な業務やサービスへのAI実装に際して、AIがもたらすリスクへの関心や懸念が高まったことが予想される。

図表３：AIに関する優先課題（日本）



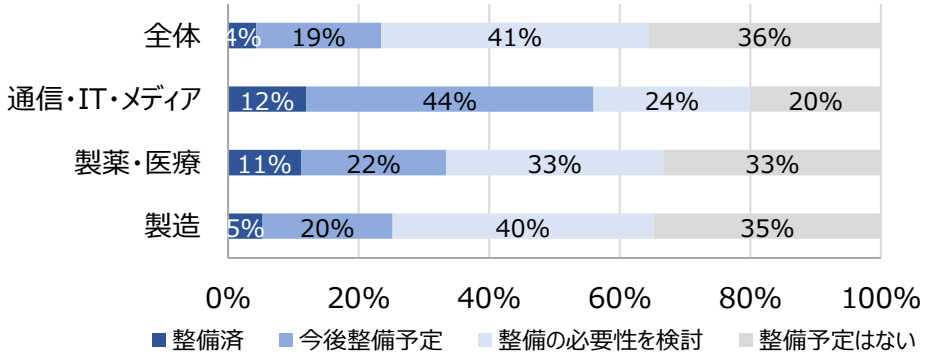
（参考）https://www.pwc.com/jp/ja/knowledge/thoughtleadership/2023-ai-predictions.htmlを基に作成

日本国内ではAIリスク管理の整備が大きく進まず

- AIリスク管理の主要な項目として挙げられるのは、**サイバーセキュリティや、AIを利用する際の安全性や信頼性に関するリスク、プライバシー等に関する対策の必要性**などである。
- 一方、日本国内のAIリスク管理における組織やプロセスなどの整備状況を見ると、大半の企業が今後整備を進める段階にあると考えられる（図表４）。

☞ 実際に、生成AIに関連するリスクへの理解を深めるために、次頁からAIの発展の歴史とともに利用者が認識すべき課題を整理しよう。

図表４：AIリスクを管理する組織、ルール、プロセスの整備状況（2023年、抜粋）



（参考）https://kpmg.com/jp/ja/home/insights/2024/02/cyber-security-survey2023.htmlを基に作成

３．AI発展の歴史、基盤モデルの登場によるパラダイムシフト

➢ AIモデル開発は、基盤モデルの登場によって**「タスクごとに専用のモデルを開発する」から「汎用のモデルを1つ作って使い回す」方式へと移りつつある。**また、生成AIがAI利用のハードルを大きく下げ、データサイエンティストなどのようにモデル開発に関する専門技術を持たない一般ユーザーであっても、プロンプトやRAGなどの活用によってAIを柔軟かつ広範囲に活用できるようになった。

特化型AI
モデル開発に関するAI専門技術：高

汎用AI
モデル開発に関するAI専門技術：低

1980～1990年頃	1990～2010年代	2017年	2018年～
エキスパートシステム（ルールベースデータプロセッシング）	機械学習	ディープラーニング（ニューラルネットワーク）	AIのパラダイムシフト **基盤モデル**	BERT（ALBERT…） GPT（ChatGPT） 生成AI
プログラミング	特徴量エンジニアリング		プロンプトエンジニアリング RAG、等
特定の問題に対してルールに基づく処理をさせる。PGMのメンテナンスやデータ入力など**膨大な手作業が必要**。	**データを基にAIがルールやパターンを学習するように。**ただし、データサイエンティストが**データの特徴を定義づける必要あり**。	**AIが大量のデータを自分で学習し、パターンや特徴を見つけられるようになった。**ただし、**長時間の学習や膨大な開発コストが必要**に。	**基本的な学習内容がモデルに含まれるようになった。**基盤モデルを利用し、様々なタスクに対応することが可能に（次頁）。	分かりやすい説明文や入出力例を与えることでモデルが成長。**モデル開発やAI利用に対するハードルが大きく下がる。**
✓ ECサイトのリコメンド ✓ ユーザーに応じた記事表示、等	✓ 店舗の来客分析 ✓ 商品の需要予測、等	✓ 自動運転、高度運転自動化 ✓ FaceID、スマートスピーカー、等		

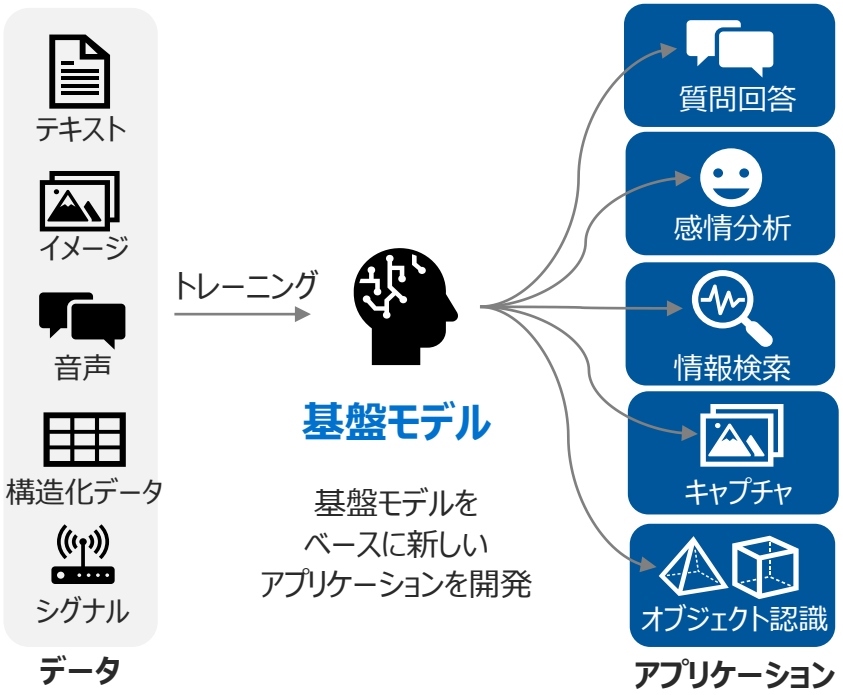
（参考）https://blogs.nvidia.co.jp/2023/06/19/what-are-foundation-models/, https://aismiley.co.jp/ai_news/detailed-explanation-of-the-history-of-ai-and-artificial-intelligence/などを基に作成

４．基盤モデルの特徴と注意点

- 基盤モデルの登場によって**AIをゼロから開発する必要がなくなり、様々なモデルを迅速かつコスト効率よく開発することが可能に**
- 一方、基盤モデルは汎用性が高い反面、不適切な回答やバイアスが生じる可能性があるほか、悪用することも可能であり、対策ができていない状態で実装を進めると大きな混乱をきたすリスクが高く、**適切にコントロールする仕組み作りが併せて必要となる**

基盤モデルの概要

- **膨大なデータを基にトレーニングされた汎用的なAIモデル。AIをゼロから開発する必要がなく、**基盤モデルを基に、新しいアプリケーションを迅速かつコスト効率よく開発することが可能に。



（参考）https://blogs.nvidia.co.jp/2023/06/19/what-are-foundation-models/, https://blog.recruit.co.jp/data/articles/foundation_models/などを基に作成

基盤モデル利用の際の注意点

1. 理解力の欠如：文法的にも事実的にも正しい答えを出すことができるが、**ユーザーの要求や意図、背景を正確に理解することが困難。社会意識や感情、倫理観もない。**
2. 信頼できない答え：特定の主題に関する質問への回答は、**信頼性が低く、不適切、または不正確である場合がある**
3. バイアス：トレーニングデータから有害な表現を拾い上げることがあり、**バイアスが生じる可能性**がある。

基盤モデルのリスクとは

- 基盤モデルは汎用性が高い反面、**悪意のある第三者の手に渡れば、フェイクコンテンツを捏造したり、新しいコンピュータウイルスを作成するなどの行為も大規模かつ容易**にできてしまう。
- 基盤モデルの不適切な回答や悪用対策が不十分な状態で実装を進めると、**社会に大きな混乱をきたすリスク**がある。
- **基盤モデルは使い方次第で善用も悪用できる**という理解が重要であり、生成AIを活用したデジタル施策を推進する上で**AIのリスクコントロールやガバナンス施策は不可欠**である。

【参考１】近年の基盤モデルの変遷

日本総研
The Japan Research Institute, Limited

自然言語処理、文章生成 ＞ 生成対象の拡大（画像・動画、音楽など）／精度向上 ＞ マルチモーダル*

基盤モデル登場

OpenAI

- **GPT**（2018）
  - ✓ 文章生成
  - ✓ 言語翻訳
- **GPT2**（2019）
  - ✓ 文章生成
  - ✓ 言語翻訳
- **GPT３**（2020）
  - ✓ 文章生成
  - ✓ 言語翻訳
- **CLIP**（2021）
  - ✓ 画像認識
- **DALL・E**（2021）
  - ✓ テキストから画像を生成
- **ChatGPT公開**（2022）
- **DALL・E2**（2022）
  - ✓ 高解像度で画像を生成
- **GPT4**（2023）
  - ✓ マルチモーダル（様々な形式の内容を理解）
- **DALL・E3**（2023）
  - ✓ 高解像度で画像を生成
- **GPT4o**（2024）
  - ✓ マルチモーダル、セキュリティ、速度向上

Google

- **BERT**（2018）
  - ✓ 文章分類
  - ✓ 質疑応答
  - ✓ 表現認識
- **検索エンジンにBERT搭載**（2019）
- **T5**（2020）
  - ✓ 文章要約
  - ✓ 質疑応答
  - ✓ 表現認識
- **LaMDA**（2021）
  - ✓ 自然対話
  - ✓ 膨大なデータで事前学習
- **PaLM**（2022）
  - ✓ 言語理解
  - ✓ 知識推論
  - ✓ コード生成
- **Imagen**（2022）
  - ✓ テキストから画像を生成
- **Bard**（2023）（LaMDA搭載の対話型AI）
- **Gemini**（2023）
  - ✓ マルチモーダル（様々なタスクを汎用的に実行）

AI21 labs

- **AI21 Jurassic**（2021）
  - ✓ 文章生成、等

stability.ai

- **Stable Diffusion**（2022）
  - ✓ 画像生成

ANTHROPIC

- **Claude**（2023）
  - ✓ マルチモーダル

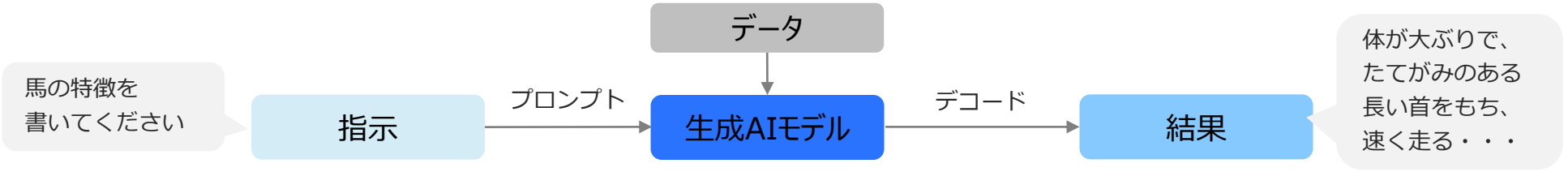
2017 2018 2019 2020 2021 2022 2023 2024

＊ P10にマルチモーダルの概要を記載

【参考２】ユニモーダル VS マルチモーダル

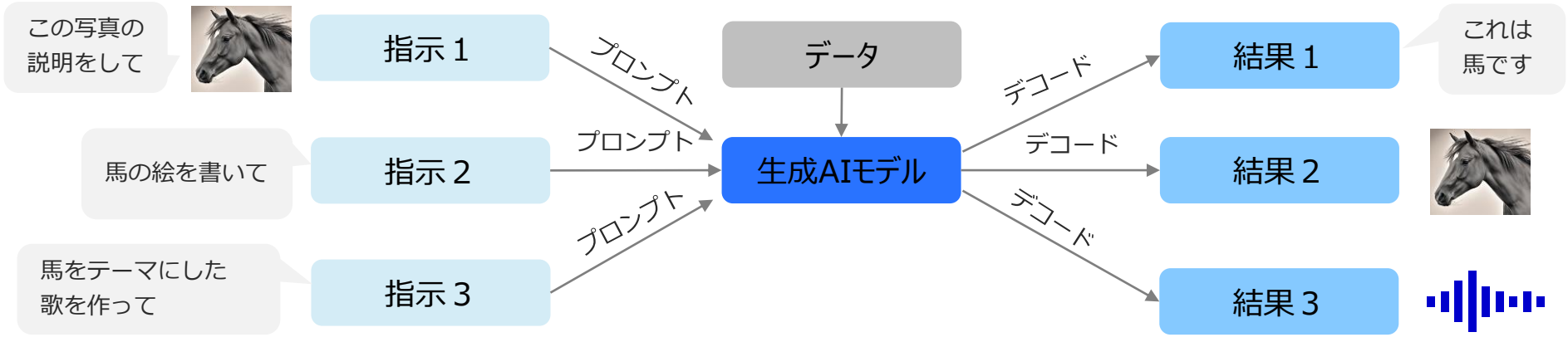
ユニモーダル

- **あるタイプのデータを入力し、同じタイプの出力を生成**する。
- 多くが**テキストプロンプトを処理し、テキスト応答を生成するように設計**されている
- マルチモーダルと比較すると処理がシンプルなため、セキュリティ対策や説明可能性などの対策が取りやすい



マルチモーダル

- **大量のテキストや、画像、動画、音声記録などのデータを学習させることで、生成AIモデルの学習能力を増強**
- テキストと関連する画像、ビデオ、音声などのデータ間におけるパターンや相関関係を学習し、**複数の結果を生成することが可能**
- **ユニモーダルと比較するとアルゴリズムが複雑になり、ブラックボックスになりがちであり、バイアスも多くなる傾向**がある。
- 複数のソースやフォーマットから得られる膨大な量のデータを使って学習するため、**プライバシー管理が煩雑**に。また、**それぞれのデータタイプに応じたセキュリティ対策や、脆弱性を包括的にカバーするセキュリティ対策などが必要**になる。



（参考）https://www.datacamp.com/blog/what-is-multimodal-aiなどを基に作成

【参考３】基盤モデルの例

製品名	Gemini	GPT4o	Claude 3.5 Sonnet	Stable Diffusion
開発企業	Google	OpenAI	Anthropic	Stability.ai
本社	米国	米国	米国	イギリス
リリース	2023年12月	2024年5月	2024年6月	2022年8月
主な特徴	**マルチモーダル**	**マルチモーダル**	**マルチモーダル**	**画像生成**
詳細	•Googleが2023年12月に発表した**大規模マルチモーダル基盤モデル**。Googleの生成AIは元々Bardという名称だったが、Geminiに名称を変更。 •マルチモーダル性（テキスト、画像、音声など異なるフォーマットのデータを統合的に処理）が高まり、**質問応答や文章生成、画像認識、音声認識など、様々なタスクを汎用的にこなすことが可能に**。	•OpenAIがリリースしたGPTシリーズの最新モデル。**GPT-4で実装されたマルチモーダル機能が一層強化され**、画像や音声、動画などの入出力が可能になった。 **•特に音声処理のスピードが格段に速まり**、GPT-4の応答速度（3～5秒程度）と比較して、大幅な性能向上（0.3秒程度）を達成した。 **•会話の割り込みも可能になり、人間と会話をしているような感覚を体験できる。**	•OpenAIの元社員が2021年に設立したスタートアップ。2023年にClaudeを提供。 •同社は**倫理的なアプローチを特に重視**しており、AIが善悪を見極め適切なやりとりを行えるように**「Constitution AI（憲法AI）」と呼ばれる原則をシステムに搭載**している。 •2024年6月に発表されたCalude 3.5 Sonnetは、**OpenAIのGPT4oと同等あるいは上回る能力を備えているとされる。**	•2020年にイギリスで設立されたStability.aiが提供するサービス。入力されたテキストを元に高精度の画像を生成するAI。 •2022年に**オープンソース化され、画像生成AIの利用に対するハードルを大きく下げたと評されている。** •自分で学習データを追加したり、既存のデータを調整することで、オリジナルの画像生成AIを作成可能

（参考）https://ai-market.jp/technology/foundation-model/#i-7, https://www.sbbit.jp/article/cont1/140613?page=2, https://www.skillupai.com/blog/tech/about-gemini/#no1などを基に作成

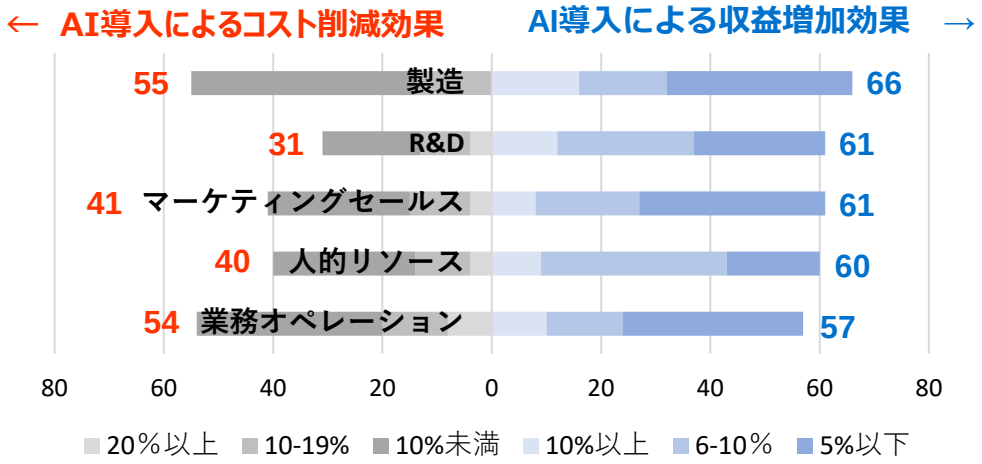
５．企業価値向上において広く活用が期待されるAI

- 多くの企業がAIの導入効果として期待するのは、事業創出や収益増などの企業価値向上に資する効果である
- 特にAIの活用領域として注目されるのが、ESG経営の課題解決に寄与する非財務情報の活用や分析である
- 非財務投資やシナリオ分析には正確性や妥当性が強く求められるため、AIのリスク管理やガバナンス対策は一層不可欠な要素に

AIで実感している効果（コスト削減から新たな価値の創出へ）

- グローバル企業が挙げるAIの導入効果として、**コスト削減効果より収益効果が高い点に注目したい**。多くの企業にとって、**AIはコスト削減のためのツールではなく、事業創出や企業価値の向上に寄与する要素として認識される**傾向がある。
- 様々な業界においてAIの導入が進むなか、AIの活用度が事業の成長や継続性を大きく左右する要素になりつつある。

図表５：AIの導入効果（コスト削減効果と増収効果の比較）

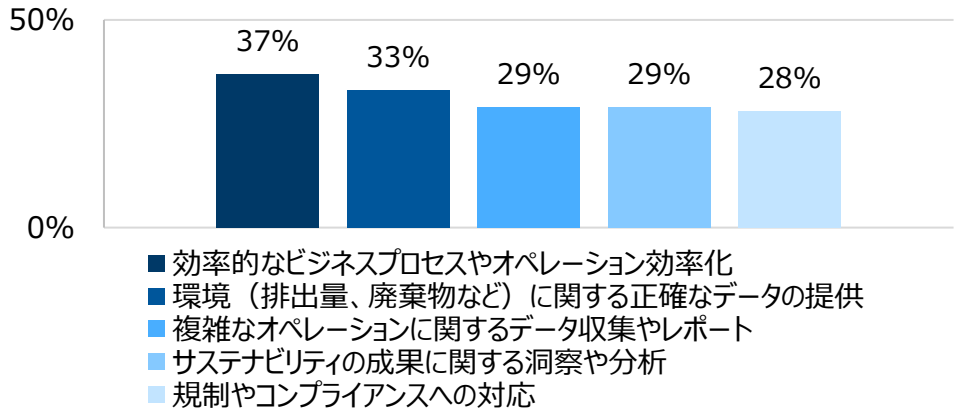


（参考）https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-reviewなどを基に作成

AIの活用が大きく期待されるESGの課題解決

- 近年、AIの活用領域として特に注目されているのが、**ESG経営などの非財務情報に関するデータの活用や分析**である
- 米国では**SEC*1が投資判断に影響を与える非財務情報を年次報告書等に記載することを要請**しており、非財務情報の管理や分析が経営戦略における重要な位置づけとなっている
- **企業価値と非財務投資の相関分析や、長期シナリオ分析などのシミュレーションはAIが最も能力を発揮する領域**であり、非財務投資の見極めにおいてAIの活用が一層期待されている。

図表６：AIが最も解決できると考えるESGの課題は何か*2



*1 米国証券取引委員会

*2 （参考）https://www.ibm.com/downloads/cas/DOMQ0OWAを基に作成

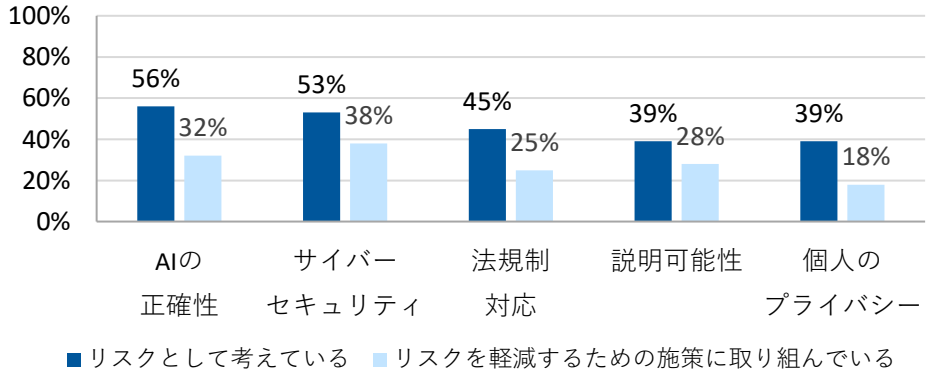
６．生成AIの活用と並行して進められる企業のガバナンス戦略

- リスク管理やガバナンス施策が十分に整備されるまでは、広範囲に生成AIを活用することは難しく、米国においても実用化にまで至ったユースケースはまだ少ないと言われている
- もっとも、現在は生成AIの本格活用に向けた準備段階と捉えることも可能であり、ガバナンス戦略が立案されることで、様々な領域へのAI実装が急速に進むことも予想される

生成AIの活用と並行して進むリスク･ガバナンス施策

- 米国では**生成AIの活用と並行して、AIの正確性やサイバーセキュリティなどのリスク管理関連施策が推進**されている。
- リスク管理やガバナンス施策が十分に整備されるまでは、生成AIを広範囲に実装することは難しく、**現状は、米国企業の多くが生成AI実用化の検討段階にある**と言われている。
- もっとも、ガバナンス戦略が整備されれば、様々な領域へのAI活用が進むと予想されるため、**現在は本格活用に向けた準備段階と捉えることもできる。**

図表７：生成AIの導入におけるリスク

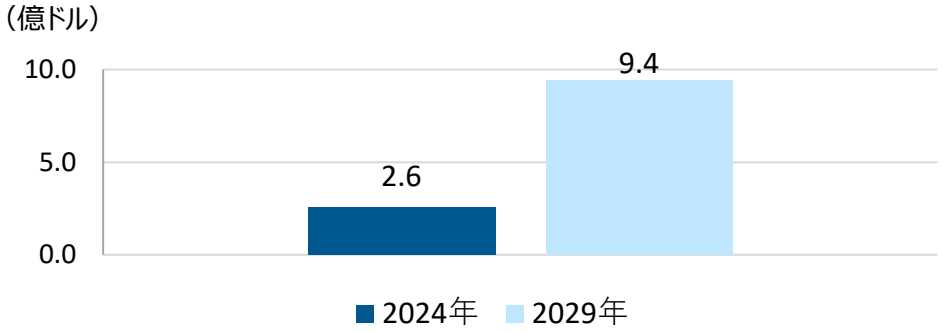


（参考）https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-yearを基に作成

生成AIの台頭で急成長するAIガバナンス市場

- グローバルのAIガバナンス市場規模は**2024年に2.6億米ドルと推定され、2029年までに9.4億米ドルに達し、28.80%のCAGR（年平均成長率）で急速に成長する**と予測されている
- グローバル企業にとって**AIガバナンス施策は、事業戦略の強化や企業価値の向上に大きく寄与する施策**として認識されている
- Capgemini Research Instituteによると、**倫理観を伴うAIサービスを提供する企業は、NPS*1で大きく優位性**があり、企業価値やブランド力の向上に繋がる効果があると指摘されている。

図表８：AIガバナンス関連のグローバル市場*2



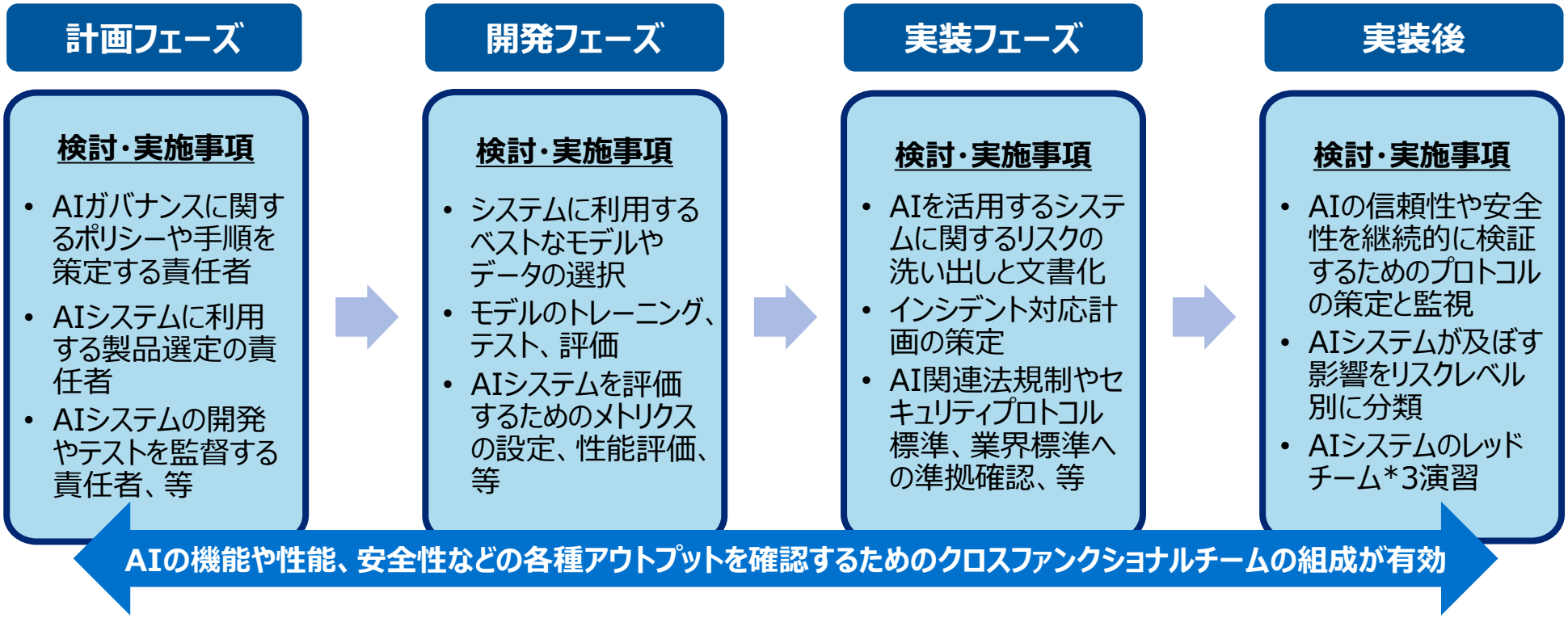
*1 NPS(Net Promoter Score)：顧客ロイヤルティ（商品やサービスに対する信頼等）を測る指標

*2 （参考）https://www.mordorintelligence.com/ja/industry-reports/ai-governance-marketを基に作成

７－１．AIガバナンスを強化するプロセスや組織の構築

- 後述するように、**法規制が強化されている欧米などの地域では、多くの組織にとってAIガバナンス戦略の立案が急務**となっており、システム開発サイクルの中に、ガバナンスに関するポリシーの策定やリスク管理を行うプロセスなどを組み込む必要性が指摘されている
- 機能や性能、安全性などの様々な観点でメトリクス*1を設定するとともに、AIが意図したとおりに動作しない場合に備えて適切なインシデント管理体制を敷くなど、**人間がAIシステムを十分にコントロールできる仕組みを整備することが重要**とされる
- また、**組織内にクロスファンクショナルチームを組成し、法規制やセキュリティ標準などへの準拠状況を部門横断的に確認**するとともに、AI実装後も継続的にシステムの安全性や信頼性を検証する体制を整備することも有効な施策として示されている

図表９．AIシステム開発における一連のガバナンス施策例*2



*1 メトリクス：ソフトウェアの品質や開発プロジェクトの進捗状況などを定量的に測定して管理のための指標にすることを指す
*2 （参考）RSA Conference 2024 セッション" AI Governance & Risk Management: Fortifying Your Cybersecurity Strategies(2024年5月7日開催)"を基に作成
*3 ある組織のセキュリティの脆弱性を検証するためなどの目的で設置された、その組織とは独立したチームのこと

７－２．米国のAIガバナンス施策に関する企業事例の紹介

- 米国では、**ビッグテック企業などを中心に、AI実装に向けたフレームワークやガバナンス方針を立案する動きが拡大**している
- 企業の取り組みに共通するのは、ガバナンス強化に向けた**全社横断的な施策の推進やAI管理要領の策定**などである
- AIには多様な場面での活用が期待される反面、システムの動作や性能だけでなく、安全性や信頼性、法規制への準拠状況など、**AIシステムが世の中に与える影響を多角的に検証することが強く求められており、部門横断的な体制の構築が不可欠**となる

図表１０：Cisco社の「責任あるAIのフレームワーク」

項目	内容
部門横断型のAI施策推進	エンジニアリングや営業、セキュリティ、法務、人事等の各部門の経営幹部を中心にAI委員会を構成し、リーダーシップをとってAIの監督を行う。
AIのコントロール	セキュリティやプライバシー、人権を考慮してAI設計を行う。また、リスクを防止、軽減するための評価を義務付け、事業を展開する市場における（法規制等の）評価基準を満たすかどうかを判断する。
インシデント管理	AIシステムに関するインシデントをAIガバナンスチームに割り当て、報告・分析し、解決のために関連チームと協業する
社外エンゲージメント	政府機関等と協力し、AIのメリットとリスクに関するグローバルの見解を理解する。また、研究機関等と連携し、技術や組織、社会などの様々な観点から、AIガバナンスのベストプラクティスを立案する。

（参考）https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/cisco-responsible-artificial-intelligence-framework.pdfを基に作成

図表１１：Google社が提示するAI実装のベストプラクティス（抜粋）

項目	内容
部門横断型のAI施策推進	AI施策を評価するために、様々な組織（IT、セキュリティ、リスク、コンプライアンス、法務、データガバナンスなど）のステークホルダーの専門知識を活用する。
AI指針の策定	組織のAI指針を定義し、基本的な要件とユースケースを明確にする。プライバシーの保護に重点を置き、AIが生成した意思決定のレビューに、人間が関与することを保証する。
インシデント管理体制の整備	AIの運用において、セキュリティやコンプライアンス、法務などに関する疑問や課題が生じた際、明確な回答を迅速に得るためのエスカレーション体制を確立する。
AI施策に関するエンゲージメント	AI施策の状況や関連する法規制への準拠状況などについて、社内外の利害関係者に伝達・共有する仕組みを整備する。

（参考）https://cloud.google.com/transform/gen-ai-governance-10-tips-to-level-up-your-ai-program?hl=enを基に作成

【参考４】AIを活用した開発の主な特徴と推進体制について

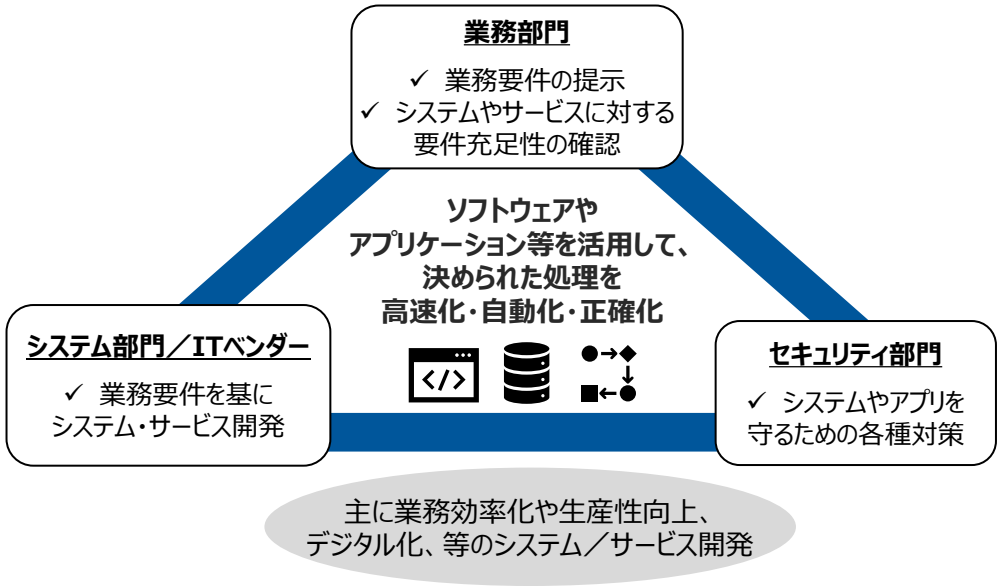
従来のシステム／サービス開発（ITツール導入）

項目	内容
特徴	業務部門の要件を基に、システム開発部門やITベンダーがシステム／サービス開発を行う
開発内容	主に**業務プロセスやサービスのデジタル化・ペーパーレス化、自動化、データ管理等の開発**が中心
システムやサービスの検証	業務要件を基に、機能や性能の適合性、アウトプットや処理の妥当性等に関するテストや検証を行う
開発体制	一般的に業務部門やシステム開発部門／ITベンダー、セキュリティ部門等が協業して開発を行う

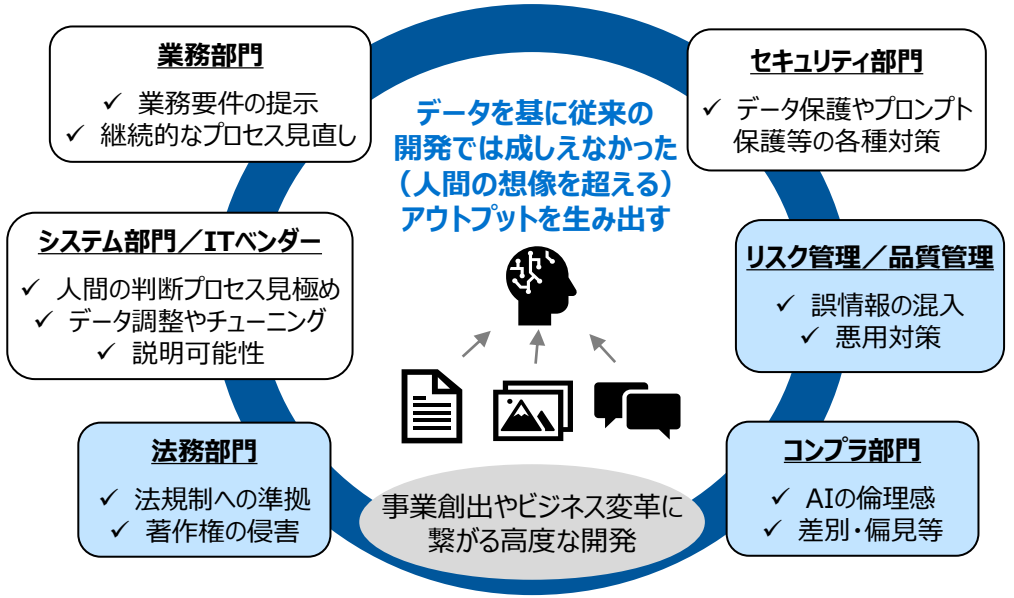
AIを活用したシステム／サービス開発

項目	内容
特徴	業務部門の要件を基に、システム開発部門やITベンダーがAI学習のデータや学習環境を用意する
開発内容	ビッグデータ分析や動向予測、意思決定の支援など、**企業の事業創出に大きく寄与する開発も可能**に
システムやサービスの検証	従来の検証に加え、**法規制の準拠や倫理・偏見等、アウトプットに関する様々な観点での検証が必要**に
開発体制	従来の開発体制に加え、**法務やコンプラ、リスク・品質管理等、多くの関連部門との協業が不可欠**に

図表１２：従来のシステム／サービス開発の体制イメージ



図表１３：AIシステム／サービス開発における部門横断組織のイメージ



（参考）「日経MOOK　生成AI　協働・共生の時代」などを基に作成

８．日本企業がAI導入を進めていくうえでの障壁

- AIの導入目的は日米で大きく異なっており、米国は収益増加施策に注力する一方、日本はコスト削減などの業務効率化施策に重点が置かれる傾向がある。要因として挙げられるのは、社内の協業体制やデジタル施策における予算の充足性の違いである。
- ガバナンスの整備はAI戦略の根幹となる施策の一つである。AI活用を多角的に進めるためには、従来のITツール導入とは異なる発想への転換が必要であり、部門横断的な推進体制の確立やデジタル予算の拡充も併せて検討することが求められる。

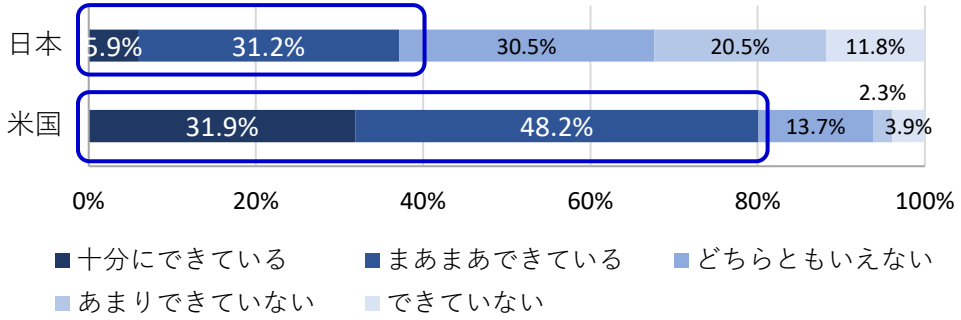
部門間の協調が十分に図られない日本のデジタル施策

- ７－１、７－２で触れたように、AIガバナンス戦略の立案や推進には、**AIを取り巻く様々な課題を総合的に捉えて判断するための、部門横断的な体制の構築が肝要**となる。
- しかし、デジタル施策における部門間の協調に関する日米の状況を比較すると、米国では約8割の企業が部門間の協調を行っているのに対し、**日本ではその割合が4割未満にとどまる**。
- 米国ではDXの予算を定常的に確保する企業の割合が最も多いが、日本では都度申請を行い、承認された予算のみが割り当てられるケースが多い。**日本企業の予算不足は、デジタル施策を推進する際の深刻な課題**として指摘されている。

AI戦略を立案するための体制強化や予算拡充の検討を

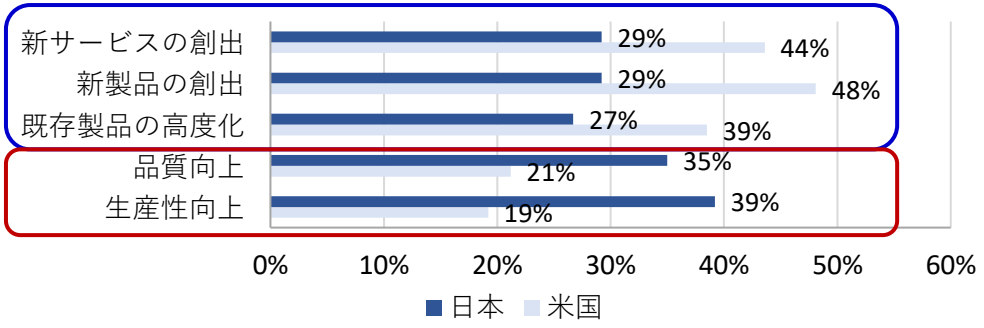
- AIの導入目的として、**米国ではサービス開発などの収益増加施策が挙げられる一方、日本ではコスト削減施策に重点が置かれる傾向**がある（図表１５）。DXの推進体制や予算などの違いが、日米の導入目的の差異として表れていると推察される。
- 部門間の連携が図られないままAI施策を進めた場合、**予算や人材などの制約を受け、特定の組織だけが推進の責任を担うこととなり、AI活用が限定的となる**ことも懸念される。
- AIを多角的に活用する為には、従来のITツール導入とは異なる発想への転換が必要であり(P16詳述)、**部門横断的な体制構築やデジタル予算の拡充も併せて検討することが求められる**。

図表１４：経営者･IT部門･業務部門の協調（日米）



（参考）独立行政法人情報処理推進機構（IPA）「IPA DX白書2023」2023年2月を基に作成

図表１５：AI導入の目的（日米）



（参考）独立行政法人情報処理推進機構（IPA）「IPA DX白書2023」2023年2月を基に作成

９．まとめ

●近年、グローバルにAIの社会実装が進み、様々な業界におけるAI活用が進んでいる。とりわけChatGPTなどの生成AIの登場をきっかけに、日本国内でも生成AIの利用に向けた議論がなされているが、実際の利用率や検討状況などを見ると、わが国でAI活用が大きく進展しているとは言い難い。

●日本企業がAI導入を進める際の障壁として挙げられるのは、サイバーセキュリティや、AIを利用する際のリスク管理、プライバシー保護等に関する一連のガバナンス施策の立案と推進である。AIは多様な場面での活用が期待される反面、安全性や信頼性、法規制への準拠状況など、AIが世の中に与える影響を多角的に検証することも併せて求められており、組織が部門横断的に施策や戦略を立案する重要性が指摘されている。

●実際に、ガバナンス施策が十分に整備されるまでは、広範囲に生成AIを活用することは難しく、米国企業も生成AIのPoCを実施しているものの、業務の実用化にまで踏み込んだ有効なユースケースはまだ少ないと言われている。もっとも、AI実装に向けた一連のガバナンス戦略が立案されれば、様々な領域へのAI導入が進むことも予想されるため、現在は生成AIの本格活用に向けた準備段階とも捉えられる。

●日本企業が生成AI活用を多角的に進めるためには、基盤モデルの特性やグローバルの事例を十分に把握するとともに、従来のITツールとは異なる発想への転換を行い、全社的にAI施策を進めるための体制構築やデジタル予算の拡充も併せて検討することが求められよう。

第２章
AIに関連する欧米の法規制動向と考察

１－１．欧米のAI政策の特徴や違い（一覧）

- AIの社会実装が進むなか、欧米ではAIに関する様々な規制やガイドラインが整備されつつある
- 特に注目されるのは各国のアプローチの違いであり、欧州ではAIのリスクに応じて厳格な罰則を施行する法規制が採択された一方、米英ではAI活用やイノベーションの促進とAIリスク対策を両立するための政策が進められている

図表１６：各国のAI関連政策の比較

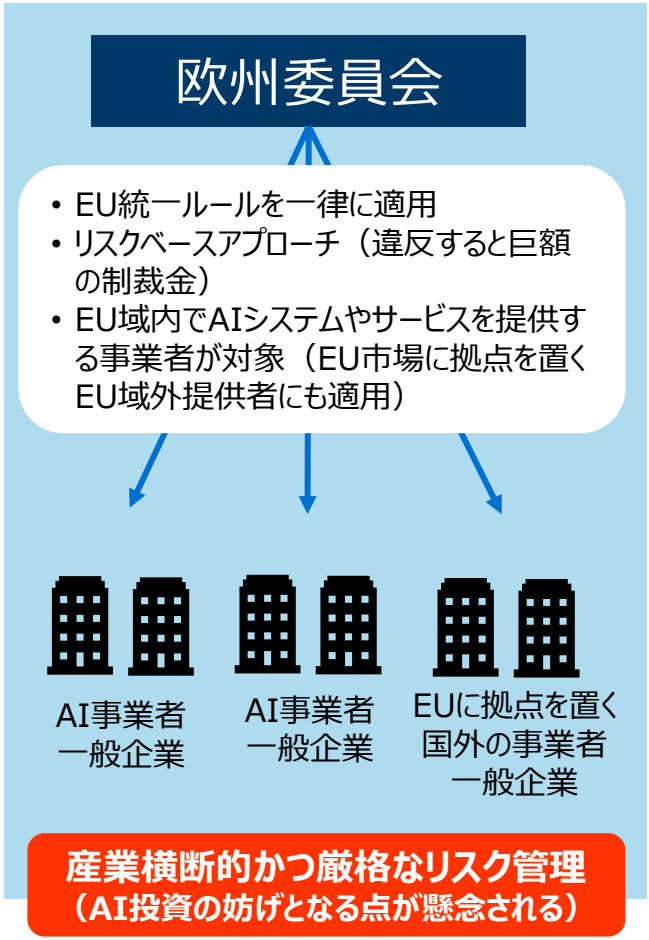
	EU	米国	イギリス
取組姿勢	**AI活用の厳格化 開発利用者に対する法的拘束有**	**積極的なAI活用推進・ 開発企業に対して一部義務化**	**イノベーション・成長機会の創出・ AIリスクへの対応の迅速化**
直近の動き	EUのAI規制法が採択（2024/5） 今後、規制が段階的に適用される	AIの安全性に関する新基準などの 大統領令を公表(2023/10)	AIの安全な開発や規制当局のスキル アップをイギリス政府が支援(2024/02)
特徴 （P21参照）	**欧州委員会による 画一的かつ包括的な管理**	**産業の特性や状況に応じて、政府機関あるいは規制当局が AIリスクやイノベーション促進に関する個別の政策を立案・推進**	
政策の概要	✓ リスクベースアプローチにより、AIをリスクに応じて４つのカテゴリに分類し、禁止事項や要求事項、義務などを明示 ✓ 法的拘束力があり、違反した場合、罰則も課される ✓ 規制・罰則は、EU域外で作られたソリューションをEU域内で使用する場合にも適応される	✓ AI活用を促進する政策（以下例） - ヘルスケアや気候変動など重要分野におけるAI研究助成金の拡大 - 高度な技能を持つ人材の受け入れ拡大のためのビザ審査の合理化 ✓ AI開発企業に、サービス提供前の安全性評価を受けるなど、一部で強制力のあるルールもあり	✓ 責任あるAIの開発に向けて9つの研究ハブを設立 ✓ 米国とのパートナーシップ構築に向けて9,000万ポンドを拠出 ✓ 国際科学パートナーシップ基金を通して、英米連携の研究プロジェクトに対し900万ポンドを拠出、等

（参考）Plug and Play Venturesの投資家Amit Patel氏へのインタビュー（2024年5月16日）, https://www.nikkei.com/article/DGXZQOUA195JQ0Z10C24A4000000/, https://www3.nhk.or.jp/news/html/20231030/k10014241801000.html, https://www.nikkei.com/article/DGXZQOGR21BN30R20C24A5000000/などを基に作成

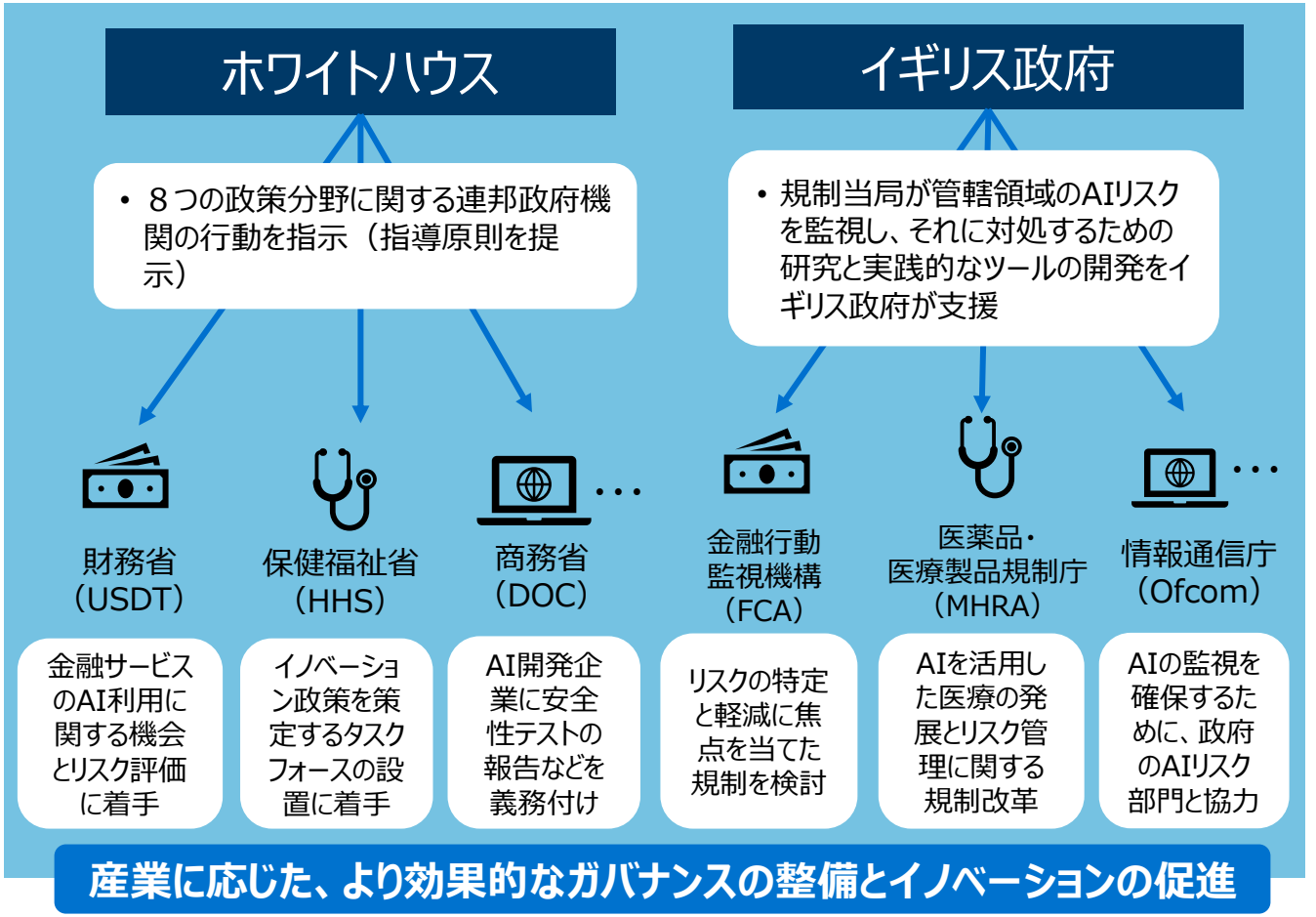
１－２．欧米のAI政策の特徴や違い（米国VCへのインタビューを基に作成）

- EUのAI規制法は、**欧州委員会がEU域内のAIサービスについて産業横断的に厳格なリスク管理を行う**点が特徴
- 一方、**米英では産業分野に応じて省庁が政策の立案や推進における責任を担う**、という政策スタンスの違いがある
- 米英の政策は、**産業の特性や状況に応じて、より効果的なガバナンスとイノベーションの促進を図る狙い**がある

図表１７：EUのAI規制アプローチ



図表１８：米国とイギリスのAI規制アプローチ



（参考）Plug and Play Venturesの投資家Amit Patel氏へのインタビュー（2024年5月16日）、各国政策ドキュメントなどを基に作成

２．米国のAI関連規制における近年の動向

- 米国では2019年の大統領令でAIの研究開発支援に重点を置いた施策が示されて以降、**AIの社会実装の拡大に向けて、AI権利章典（AIの利用原則）やリスクマネジメントフレームワークなどが段階的に整備されてきた**。
- 2023年10月の大統領令では、AIの安全利用に関する基準や連邦政府機関の政策などが掲げられ、AIの安全利用やリスク管理に関する一定の規制が示されたものの、**様々な組織がイノベーションを促進するためのインセンティブプログラムも併せて設置され、AIの導入やシステム開発などにおいて世界をリードしようとする姿勢が見られる**

図表１９：近年の米国の主なAI関連政策

国	法律・業界規制	発行元・年	内容	政策のポイント	AIに関する企業の動向
米国	13859号 大統領令	ホワイトハウス 2019年2月	米国がAI研究開発・展開において世界をリードし続けることを主張	研究：AI研究開発の促進 AI技術の発展 ↓ 実装準備：AI利用の原則やリスク分類など、実装における前提事項を整理 ↓ 実装：AI実装の拡大に向けたリスク管理とイノベーション促進の両立	• 2018年6月、GoogleがAIの安全性や説明責任、プライバシー等に関する原則を発表。 • 2020年のパンデミック以降、デジタルサービスが多様化し、ユーザーのデジタル体験を強化することが企業の最重要課題に。 • 企業は様々なビジネスにAIを組み込むことを検討し、AIがデジタル施策における優先的な投資対象の一つに。 • 2023年2月、OpenAIがChatGPT（GPT 3.5）リリース。ChatGPTを中心とする生成AIの台頭によってAIの活用関心が一層高まる。
	AI権利章典	ホワイトハウス 2022年10月	安全、偏見のないアルゴリズム、プライバシー保護、通知と説明、人による管理監督など、**AI利用における５つの原則を発表**		
	AIリスクマネジメントフレームワーク	国立標準技術研究所（NIST）2023年1月	リスクベース・アプローチに基づき、**リスク分類に応じた実務的なAIマネジメントを提示**		
	14110号 大統領令	ホワイトハウス 2023年10月	AI開発企業に対して、**政府による安全性の評価を受けるよう義務化**。一方で、ヘルスケアや気候変動など重要分野における**AI研究への助成金の拡大等も明示**（P22後述）。		

（参考）https://www.pwc.com/jp/ja/knowledge/column/awareness-cyber-security/generative-ai-regulation02.htmlなどを基に作成

【参考５】AIの積極活用やイノベーションを促す米国のデジタル政策

- 昨年発出された米大統領令では、AIの利用や開発の促進に向けたインセンティブプログラムが多数盛り込まれた
- 米国のデジタル政策の特徴として、ガバナンスやリスク管理等に関する一連の法規制やルールの整備と併せて、デジタル技術やデータの積極的な活用に向けた施策（インセンティブプログラム等）を立案する点が挙げられる

大統領令で示されるイノベーション促進施策

- 2023年10月に発出された米大統領令では、AI研究の助成金やAIの技術支援、AI人材の獲得に向けた施策など、**AIによるイノベーションを促進するための各種インセンティブプログラムが挙げられている**。

図表２０：米大統領令におけるイノベーション促進関連施策

施策	内容
AI研究の助成金	✓ AI研究者や学生が主要なAIリソースとデータを利用できる環境の整備 ✓ ヘルスケアや気候変動など、重要な分野におけるAI研究への助成金の拡大を通じて、米国全土でAI研究を促進する
AI技術支援	✓ 開発者や起業家に対してAI技術の開発支援やリソースへのアクセス権などを提供し、企業のAIサービス商業化を促進する
AI人材の獲得	✓ ビザの基準や審査を近代化・合理化することにより、AIの高度な専門スキルを持つ人材が米国で学習や勤務をする環境を拡大する

（参考）https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/#:~:text=With%20this%20Executive%20Order%2C%20the,information%20with%20the%20U.S.%20government.を基に作成

（参考）デジタル活用や市場成長を促す米国政策の過去事例

- 米国のAI政策にはイノベーションの促進に関連する内容が盛り込まれており、国内のAI市場を大きく発展させる狙いがあると考えられる。近年の事例では、**医療分野の法規制整備と併せてインセンティブプログラムを定めた結果、医療機関のデジタル活用が急速に進み**、テック企業の参入も相まって**デジタルヘルス市場が大きく発展した**経緯がある。

図表２１：医療分野におけるデジタル関連の主な政策

年	政策・法規制	内容
1996年	HIPAA法制定	情報の完全性、機密性、リスク管理、プライバシー保護
2006年	Value-Based Careの概念公表	患者に対する治療の有効性や改善効果などの質が医療費用の算定基準となる医療制度
2009年	HITECH法制定	電子カルテ（EHR）の利用を奨励、インセンティブプログラムの導入
2010年～	Value-Based Careの広がり	医療機関を対象にValue-based Careの償還システムが拡大。医療におけるデータ活用が重要視される。

リスク管理やデータ保護対策など、デジタル実装における対策等を整理

デジタル・データ活用の拡大に向けたインセンティブや評価基準の策定

テック企業などの参入、デジタルヘルス市場の拡大へ

（参考）https://www.jri.co.jp/MediaLibrary/file/report/jrireview/pdf/11225.pdfを基に作成。各種法規制、概念の詳細については本URLを参照されたい。

３．イノベーションとガバナンスのトレードオフ

- 先行する欧米の動向から、**AI政策の成否や評価を左右する鍵はイノベーション促進とガバナンスのバランス**だと考えられる

- 以降のページは、**デジタル市場の発展と技術発展のそれぞれの観点で、主にアメリカやEUの事例を基に法規制（ガバナンス）がイノベーションに与える影響について考察**を進める

- なお、技術発展については様々な要素があるが、本稿では生成AIの発展によって、特に大きく活用（悪用）が進むと予想されるDeepfakeを事例として挙げる

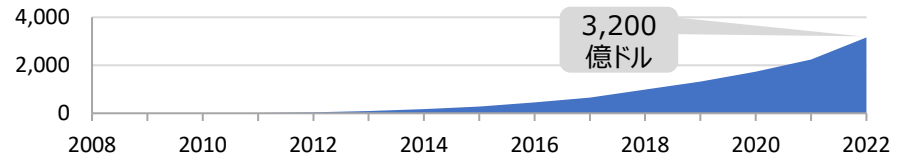
３－１．イノベーションとガバナンスのトレードオフ（デジタル市場発展の観点）

- デジタルの進展を振り返ると、特に米国ではデジタルイノベーションとともに大きな経済成長がもたらされた事例がある
- 一方、EUでは企業間の公正な競争を目指す厳格な規制が敷かれており、イノベーションへの影響も懸念される

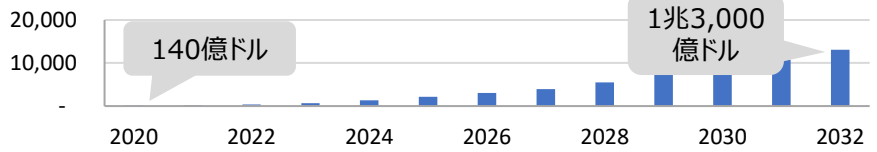
テック企業が牽引する米国のデジタル市場成長

- 新たに形成されたデジタル市場に大きな価値が見いだされると、多くの新興企業が参入し、巨大な経済圏を形成するが、**米国では特にデジタル関連市場の隆盛が経済を牽引している。**
- Analysis Groupの調査によると、2022年に**App Storeの市場は1兆ドルを超える経済圏を形成**している。実際に**iOSのアプリ開発者は3,200億ドル以上を稼いでいると言われており、AppStoreは開発者にとっての一大市場**となっている。
- **生成AI市場も**LLMを活用した様々なサービスの登場により、**2032年までに1兆3,000億ドルに成長する見込み**である。

図表２２：AppStoreのアプリ開発者が稼いだ金額の総額（億ドル）*1



図表２３：GenAIの市場推移（億ドル）*2



*1（参考）https://www.statista.com/chart/9671/developer-earnings-apple-app-store/を基に作成
*2（参考）https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/を基に作成

EUが目指す法規制強化による公正競争の実現

- 一方、EUでは、Appleがアプリの価格設定に制約を設けている点やAppStoreの手数料などを問題視し、同社の運営を**DMA法違反と判断した（P26の右側に詳述）**。
- EUの対応を受けて、Appleは、**規制上の不確実性という理由から、2024年はEUにおける生成AI技術の提供を見送る**と発表した。また、GoogleやMetaも欧州における生成AI機能の提供を遅らせる、または一時的に見送るとの判断をした。
- DMA法のように、EUでは近年ビッグテックの影響力の抑制や市場寡占を防ぐための政策が進められている。もっとも、実際には、左記のように**ビッグテックのエコシステムによって、ソフトウェア企業や消費者に恩恵がもたらされる側面**もあり、EUの法規制とイノベーションへの影響は今後も動向が注目される。

（参考）https://www.wired.com/story/apple-hits-a-major-roadblock-as-eu-targets-app-store/, https://www.namiten.jp/2024/06/25/eu-apple-dma-10/などを基に作成

３－２．イノベーションとガバナンスのトレードオフ（デジタル市場発展の観点）

- AIへの投資や技術開発力、政策などの観点で、現時点では米国が世界のAI開発競争をリードしているとされる
- もっとも、各国ともAI法規制は発展途上の段階であり、今後も法規制の動向とイノベーションへの影響が注目される

世界のAI開発力ランキング

- AIの導入を進める世界62か国の競争力ランキングによると、**インフラへの投資やAI技術開発力、政府の戦略などの観点において、アメリカが最もAI競争力が高い**と評価されている。
- イギリスは研究や投資で優位性がある点、EU圏内ではドイツやフィンランドなどがAI政策を進めている点が評価されている。

図表２４：Global AIランキング（全9項目の内、主要な3項目を記載）

順位	国	インフラ	開発	政府の戦略
1	**アメリカ**	**100**	**100**	**90.3**
2	中国	92.1	80.6	93.5
3	シンガポール	82.8	24.4	81.8
4	**イギリス**	**61.8**	**19.8**	**89.2**
5	カナダ	62.1	18.9	93.4
6	韓国	74.4	60.9	91.9
7	イスラエル	60.5	22.2	31.8
8	**ドイツ（EU）**	**68.2**	**19.5**	**93.9**
9	スイス	68	24.9	9
10	**フィンランド（EU）**	**73**	**13.1**	**82.7**
11	**オランダ（EU）**	**65.7**	**13.1**	**71.8**
12	**日本**	**80.8**	**22.2**	**80.3**

（参考）https://www.tortoisemedia.com/intelligence/global-ai/#rankings, https://www.tortoisemedia.com/intelligence/global-ai/#rankingsなどを基に作成
ランキングは全9項目のスコアの合計値で算出（図表は9項目の内、3項目のみを記載）

【参考】EUのDMA法（Digital Market Act）とは

- EUのDMA法は**デジタル市場における公正競争の実現を目的とした法律**。AppleやGoogleなどのgatekeepersと呼ばれる**ビッグテック企業のサービスに対して、厳格な義務や禁止事項、罰則を課している**。2024年6月、欧州委員会はAppleが同法に違反したとする暫定的な認定を行った（図表25）。

図表２５：DMA法に基づいたAppleへの指摘事項

指摘事項	内容
1	アプリ開発者が顧客に価格情報を提供したり、App Store以外のチャネルを通してオファーを宣伝することができない
2	アプリ開発者がユーザーを自社サイトへ誘導し、ウェブサイトで取引をした場合も、Appleに手数料を払う必要がある
3	Appleが開発者に請求するAppStoreの手数料が適切な範囲を超えている

- AppleはApp Storeによって、多くのアプリ開発者が参入する一大市場を形成したが、DMA法は**Appleの独占にメスを入れ、アプリストアの運営を一層公正にしようとするもの**である。
- EUの規制強化に伴い、GoogleやMetaも生成AI機能の提供を欧州内で見送る判断をしている。**AI規制法により、欧州ではビッグテックによる先進機能が一層制限される可能性**もある。

（参考）https://ec.europa.eu/commission/presscorner/detail/en/IP_24_3433, https://techcrunch.com/2024/06/24/apples-app-store-breaches-eus-digital-markets-act/などを基に作成

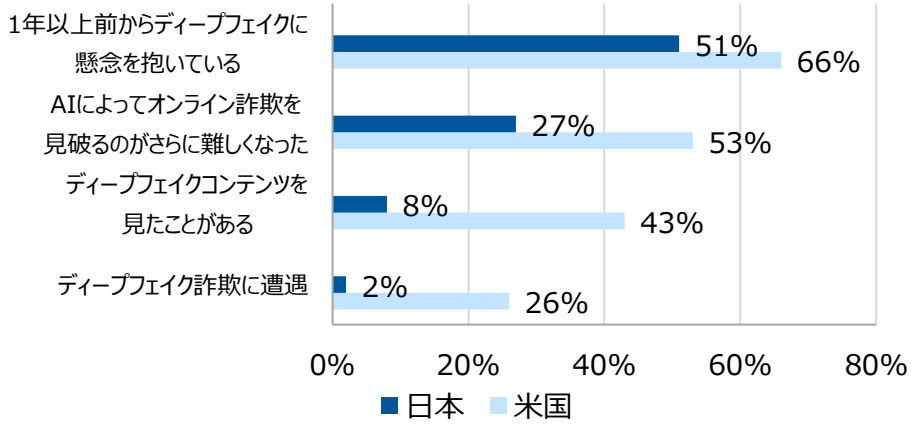
４－１．イノベーションとガバナンスのトレードオフ（技術発展の観点）

- 生成AI技術の高度化に伴い、近年特に注目されるのがDeepfakeの活用である。Deepfakeで生成された動画や音声は精巧さを増しており、選挙関連活動や生態認証情報の偽装など、悪用手段も多様化している。
- 悪用の増加・深刻化に伴い、法規制の整備を含め、Deepfakeへの悪用対策が強化されることが予想される。

近年拡大するDeepfakeの悪用

- **生成AIの発展に伴い、近年、本物そっくりの動画や音声などを作り出すDeepfake*を悪用した事件が拡大**している。
- 米国では、消費者の約4人に1人がDeepfake詐欺に遭遇しており、特に選挙関連のトピックが上位を占めている。
- 近年は**生体認証情報を偽装したり、従業員になりすまし、企業の資産を搾取する例も出ており**、様々な業界がDeepfakeへの理解や悪用対策を求められる状況になりつつある。

図表２６：Deepfakeに関する消費者アンケート



（参考）https://prtimes.jp/main/html/rd/p/000000046.000033447.htmlを基に作成

*機械学習の手法の一つであるDeep leaning（深層学習）を用いて、２つ以上の画像や映像などを人工的に合成する技術

Deepfakeを悪用した事例

ケース	事例
2024年米大統領選	2024年1月、ニューハンプシャー州において、バイデン大統領になりすましたロボコール（営業や選挙活動などに使われる自動音声電話）が民主党支持者にかかり、同州の予備選に投票しないよう呼びかける事案が発生した。
生体認証を突破するマルウェア	生体認証情報を盗聴するモバイルマルウェアGold Pickaxe。同マルウェアがスマートフォンにインストールされると、ユーザーは顔などの生体情報を傍受される。ハッカーが盗聴した生体情報を基にDeepfake動画を生成し、銀行アプリの認証などに悪用する危険がある。
Deepfakeの悪用による資産搾取	グローバル企業の従業員がCFOから電話で巨額の金融取引の支援を依頼された。従業員はCFOとのZoomコールを介して相手がCFO本人であると思い込み、指定された口座に2,500万ドルを送金した。しかし、そのCFOのビデオがDeepFake だったことが後ほど発覚した。

（参考）https://www.nbcnews.com/politics/2024-election/fake-joe-biden-robocall-tells-new-hampshire-democrats-not-vote-tuesday-rcna134984, https://www.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html, https://www.secureworld.io/industry-news/mobile-malware-deepfakes-biometric-authenticationを基に作成

４－２．イノベーションとガバナンスのトレードオフ（技術発展の観点）

- Deepfakeの悪用対策が求められる一方、同技術によって恩恵をもたらされる業界もあると期待されている
- Deepfakeは必ずしも悪影響を及ぼす技術ではなく、活用の方法によっては、ブランド戦略やユーザー体験の強化等に寄与する要素もあるため、**法規制の整備にはガバナンスとイノベーションの両側面での検討が必要**になる

様々な分野で活用が期待されるDeepfake

- Deepfake*は**広告やトレーニング、デジタル体験向上などの領域で様々な活用が期待されており、大きな経済効果をもたらす可能性がある**。

ケース	事例
広告（ハイパー・パーソナライゼーション）	地域や言語などに応じて、ローカライズされた広告を生成して配信（右図）。SNSやポッドキャスト、ラジオなどユーザーの嗜好に合わせてキャンペーンを配信し、ハイパーパーソナライズされたブランド戦略を実施することが可能に。
医療分野でのトレーニング	医療分野では一般的に高額な医療用ロボットを使って、医学生向けのトレーニングを行うが、DeepfakeやAR・VRを搭載したヘッドセットを利用することで、より安価に模擬患者の診断や治療を経験することができるように。
新しいデジタル購買体験の提供	Walmartは2021年から開始しているバーチャル試着室では、アプリ上で服の試着をシミュレーションすることができる（右図）。これまでオンラインショッピングで最も困難であった試着を実現するサービスとして注目されている。

（参考） https://techinformed.com/deepfakes-for-good-how-synthetic-media-is-transforming-business/, https://www.brandvm.com/post/deepfake-marketing#:~:text=Deepfake%20technology%20allows%20for%20hyper,personalized%20experience%20for%20diverse%20audiences.などを基に作成

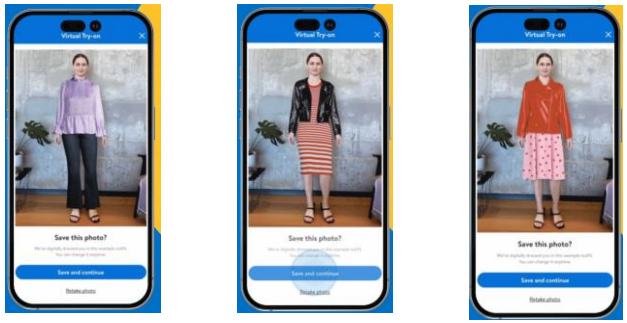
*Deepfakeにはネガティブなイメージもあるが、アメリカでは上記事例のように善用する場合もDeepfakeという名称が一般的に利用されているため、本稿では名称を統一する

Deepfakeの活用事例

図表２７：マラリア撲滅キャンペーン「Malaria Must Die」動画

- ✓ デビット・ベッカム氏がマラリアの撲滅を訴える動画を9つの言語で配信
- ✓ Deepfakeで顔の動きを操作し、それぞれの言語を話しているように見せている

図表２８：Walmartのバーチャル試着サービス「Virtual try-on.」



（参考）https://www.dataart.com/blog/positive-applications-for-deepfake-technology-by-max-kalmykov, https://www.walmart.com/cp/virtual-try-on/4556767などを基に作成

【参考６】米国におけるDeepfakeに対する検知技術と法整備の現状

- 米国では、Deepfakeの検出技術に関する研究開発や、悪用を防止するための法整備が進んでいる
- OpenAIなどのAI関連企業でも、Deepfakeを検出するツールの開発が進められているが、Deepfakeの検出技術やメディアの保護対策は先進デジタル企業にとっても発展途上の段階であり、今後の動向が注目される

Deepfakeの検知技術の研究や法整備に関する動向

- **米国では、Deepfakeの検出技術に関する研究や、連邦政府機関による研究支援が進められている。**DARPA（米国国防高等研究計画局）は2024年3月、メディアの改ざんを検知・保護するための2つの取り組みを開始した（図表29）。
- **Deepfakeに関する法整備も広がりつつある。**カリフォルニア州は2019年に政治活動やポルノなどにおけるDeepfakeの悪用を禁止する法案を可決している。近年は、テキサス州やミシガン州などでも選挙でのDeepfake利用を禁止する法案が可決されており、今後、連邦法として制定されることも予想される。

図表２９：DARPAのDeepfake検出に関する取り組み

取り組み	内容
1	SemaFor（DARPAのDeepfake研究プログラム）で開発されたDeepfakeの検出機能等を提供するオープンソースをリポジトリ上で一般に公開
2	AI FORCEと呼ばれるオープンコミュニティの研究活動で、Deepfakeを検出できる革新的で堅牢なAIモデル（機械学習、深層学習）の開発を行う

（参考）https://www.cryptopolitan.com/tennessee-shield-artists-from-ai/, https://news.bloomberglaw.com/artificial-intelligence/california-looks-to-boost-deepfake-protections-before-electionsなどを基に作成

AI企業によるDeepfake対策

- OpenAIは、画像生成を行うAIモデルDALL-Eで生成されたコンテンツを検出するツールをリリース。同ツールは、最新バージョンのDALL-E3で生成された画像を98%以上の精度で識別できる（図表30）。
- 一方、他社の生成AIモデル（MidjourneyやStabilityなど）で生成された画像を検出するようには設計されておらず、生成AIによるコンテンツを完全に識別することは難しい。
- OpenAIは現在、**Coalition for Content Provenance and Authenticityの運営委員会に所属し、MetaやGoogleらとともにデジタルコンテンツ証明の技術標準を推進**している。

図表３０：OpenAI社のツールによるDeepfake画像の検知精度

DALL-E3で生成されたコンテンツ	その他のAIモデルで生成されたコンテンツ
98.8%	5%～10%程度

（参考）https://siliconangle.com/2024/05/07/openai-announces-new-tool-can-detect-deepfake-images/を基に作成

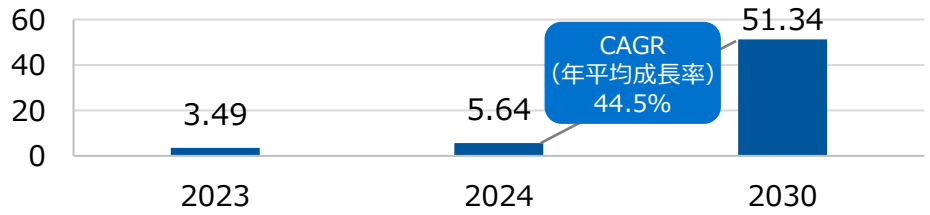
４－３．技術の特性や進展、イノベーションを十分に考慮した制度設計を

- Deepfakeはビジネス領域での幅広い活用や、同技術の悪用対策・セキュリティ対策などの技術革新によって、今後グローバルの関連市場が急速に拡大することが予測されている。
- Deepfakeのような発展途上の技術は、使い方によって多くの利益をもたらすことも、莫大な損害をもたらすこともできる。**重要なのは、**技術が十分に成熟していない段階で厳格な規制を強いることではなく、**技術の特性を十分に理解し、業界ごとの活用状況やインシデントなどに応じて、柔軟に規制を変更できるように設計しておくこと**であろう。

急速な成長が予想されるDeepfake市場

- Deepfakeの市場は、**2024年の5.6億ドルから2030年には51億ドルへと急速に拡大する**ことが予想されている。
- 背景として、生成AIの発展に伴い、従来は高い技術力やコストを要した動画などのコンテンツ生成が飛躍的に容易になる点が挙げられる。特に**エンターテインメントや広告、教育などの分野においてDeepfakeの活用が進む**ことが期待されている。
- また、**Deepfakeの悪用対策として、より優れた検出やセキュリティ対策などの領域における技術革新が活発化**し、同分野における投資が急増することも見込まれている。

図表３１：Deepfakeのグローバル市場規模（億ドル）



（参考）https://www.marketsandmarkets.com/Market-Reports/deepfake-ai-market-256823035.htmlを基に作成

Deepfake関連ソリューションの事例

図表３２：Deepfakeを生成するAI ”Stable Diffusion”*1

✓ 入力したテキスト内容に沿って高品質の画像を生成する生成AI。2022年にオープンソース化され、画像生成AIの利用に対するハードルを大きく下げたと評されている。右がDeepfake画像。

図表３３：Deepfake検知ソリューション “Reality Defender”*2

✓ Deepfakeで生成された文章や音声、動画など様々なコンテンツを分析・検出するデジタルソリューション。生体認証やコールセンターにおける音声通話の偽装など、リアルタイムでDeepfakeを解析することが可能。

*1（参考）https://www.youtube.com/watch?v=a5BGZx4CnY8を基に作成
*2 （参考）https://www.realitydefender.com/を基に作成

５．第２章まとめ

- AIの社会実装が進むなか、欧米ではAIに関する様々な規制やガイドラインが整備されつつある。特に注目されるのは各国のアプローチの違いであり、欧州では産業横断的に厳格なAIリスク管理を行う法規制が採択された一方、米英では産業の特性に応じて、より効果的なガバナンスの整備とイノベーションの促進を実現するための政策が進められている点である。
- EUのAI規制法には、企業のイノベーションを支援する施策も一部含まれるものの、包括的かつ厳格な法規制である点を鑑みると、規制の順守に伴う負担の増加によって、むしろ企業のAI投資が妨げられ、欧州の国際競争力を損なう可能性がある点が懸念されている 。
- 一方、米国の取り組みには、AIの安全利用やリスク管理に関する原則や標準、一定の規制を示しつつも、産業分野に応じて省庁が政策立案を主導する体制をとっており、様々な領域でイノベーションを促進し、AIの開発において世界をリードしようとする姿勢が見られる。
- 現時点ではAIには未知の部分も多く、各国にとっても法規制の整備は発展途上の段階であるため、規制の影響を評価するのは時期尚早である。しかし、これまでのデジタルの進展を振り返ると、米国ではイノベーションとともに大きな経済成長がもたらされた事例も多く、AIの実装によって起こりうるリスクを適切にコントロールしながら、イノベーションや市場成長を強く促す方法を模索していくことがAI政策を発展させる上で有効だと考えられる。
- 先行する欧米の政策や企業の事例には、日本の政策や企業の戦略を立案する上でヒントとなる内容も多く含まれており、各国の動向を継続的に観測することが肝要となろう。

前編のおわりに

- 前編では、生成AIの発展とともに重要性を増す企業のAIリスク管理やガバナンス体制の強化、AIに関連する欧米の法規制動向等についての考察を行った。

- 前編で整理した内容を基に、後編では米国企業のAIガバナンス施策に関する先進の事例や、AIガバナンスを効果的に進めるためのデジタルソリューションの紹介、日本企業のAIガバナンス施策強化に向けた提言などを中心に議論を進める。