



令和5年度 我が国における生成AI基盤モデル 開発の加速化に向けた調査

最終報告書

2024年3月26日



Agenda

1. 背景と事業の全体像

2. 実施内容

- (1) 生成AI・基盤モデルに関する市場動向調査
- (2) 基盤モデルの開発に必要なリソースの特定・確保の方策
- (3) 開発成果の最大化の方策



1. 背景と事業の全体像



本業務の背景・目的 (仕様書を踏まえたBCG理解)

事業実施の基本方針：背景・経緯とゴール

本業務の背景



生成AIは産業活動・国民生活を大きく変える可能性を秘め、世界中で開発競争が激しさを増しており、その開発にタイムリーに関与しないことは、大きなリスクとなりうる

- 生成AIは技術革新のスピードが早く予見可能性が低い
- 最先端の技術情報にアクセスする機会を失う

上記を踏まえ、政府として取組の一層の加速が求められており、特に大規模基盤モデル(汎用モデル)はコアの技術基盤であり、開発力の底上げが必要である

- 可及的速やかに生成AIに関する汎用的な基盤モデルの研究力・開発力を国内に醸成することが重要
- 従来の政府による開発促進策では対応が難しく、市場原理を最大限尊重し、迅速、柔軟かつ集中的にプレイヤーの取組を加速する支援を行う必要
- あわせて、領域に特化した基盤モデルも大きな可能性が見込まれ、イノベーションの創出の観点でも重要 等

本事業の目的



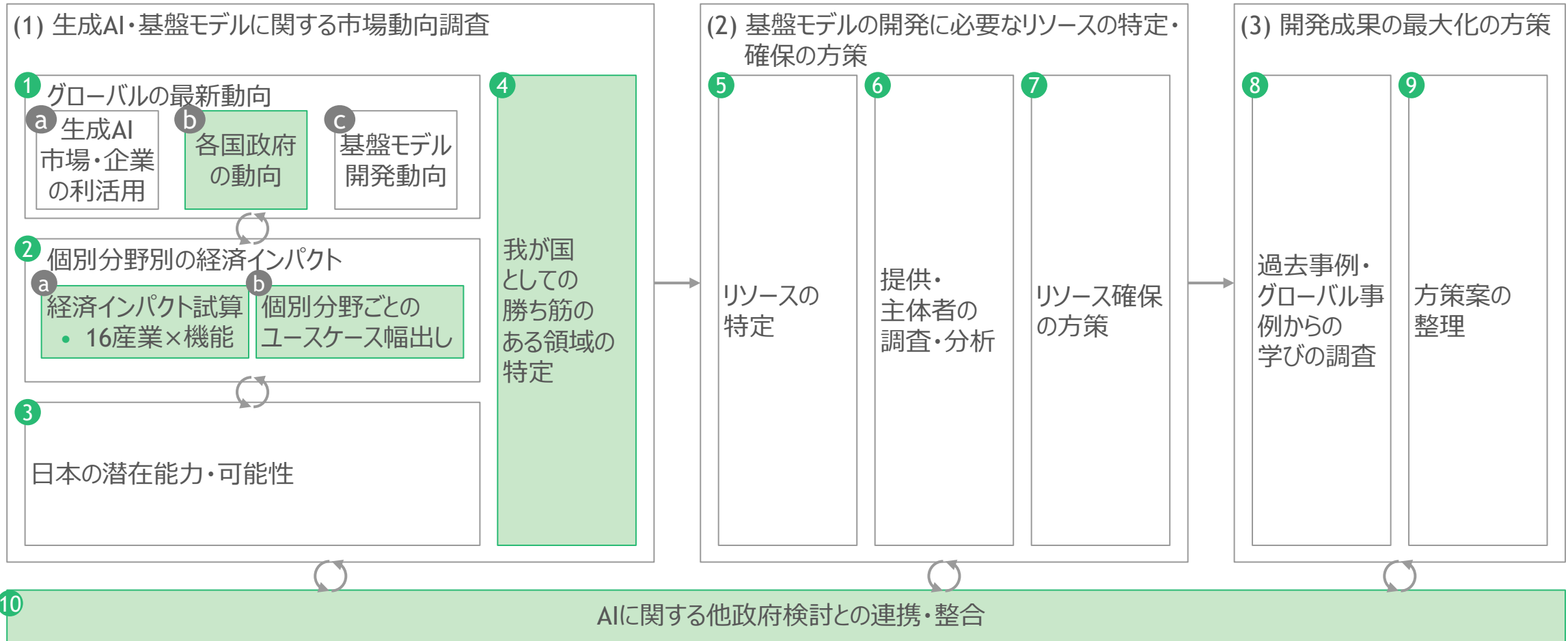
以下3つの調査・検討を通じ、国内における生成AI基盤モデル開発の加速化・効果の最大化を行う

- (1) 生成AI、特に基盤モデル分野に関する市場動向調査
 - 汎用基盤モデルの競争状況および日本における開発プレイヤーの余地を検討することで、我が国の開発力の底上げに向けた勝ち筋のある基盤モデル開発のアプローチ・領域を特定する
- (2) 基盤モデルの開発に必要なリソースの特定、及び確保の方策に係る調査・分析
 - 生成AIのコア技術である基盤モデルの開発に必要なリソースについての調査・分析を実施し、確保の方策について示唆を得る
- (3) 開発支援の成果を最大化する方策に係る調査・分析
 - 開発支援の成果を最大化する方策についての調査・分析を実施し、その方策を明らかにする

弊社の国内・グローバル、技術エキスパート、類似事例の知見を活用し、仕様書記載項目を効率的・効果的に実施するとともに、BCG独自の創意工夫も含めて実施

事業実施の基本方針：全体アプローチ

□ :仕様書記載内容 ■ :追加提案



予備調査事業: 検討スケジュール

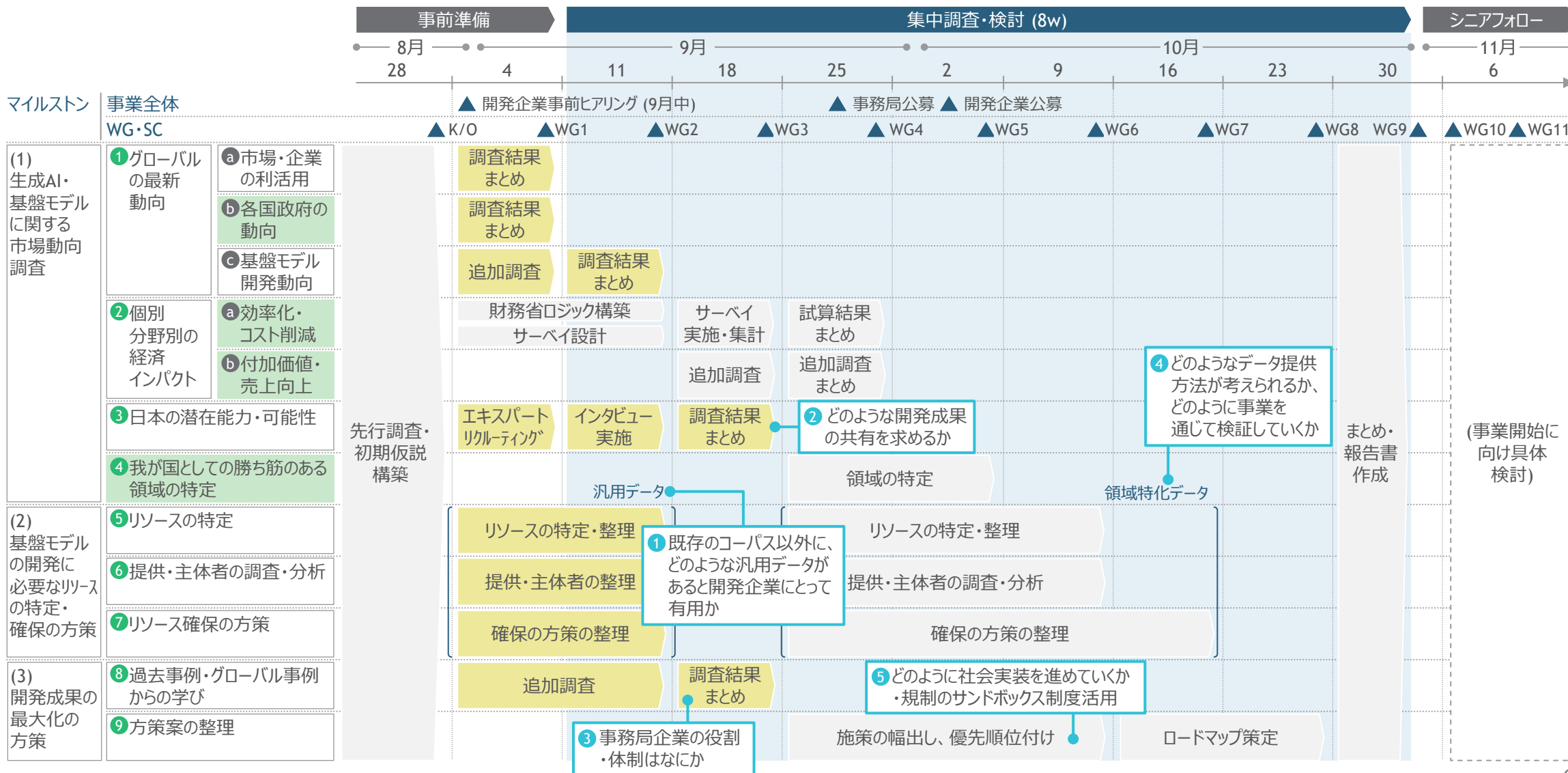
□ : 仕様書記載内容

背景と事業の全体像

■ : 追加提案

■ : 1サイクル目に向けた検討

□ : これまでの打合せで頂いた論点



2. 实施内容

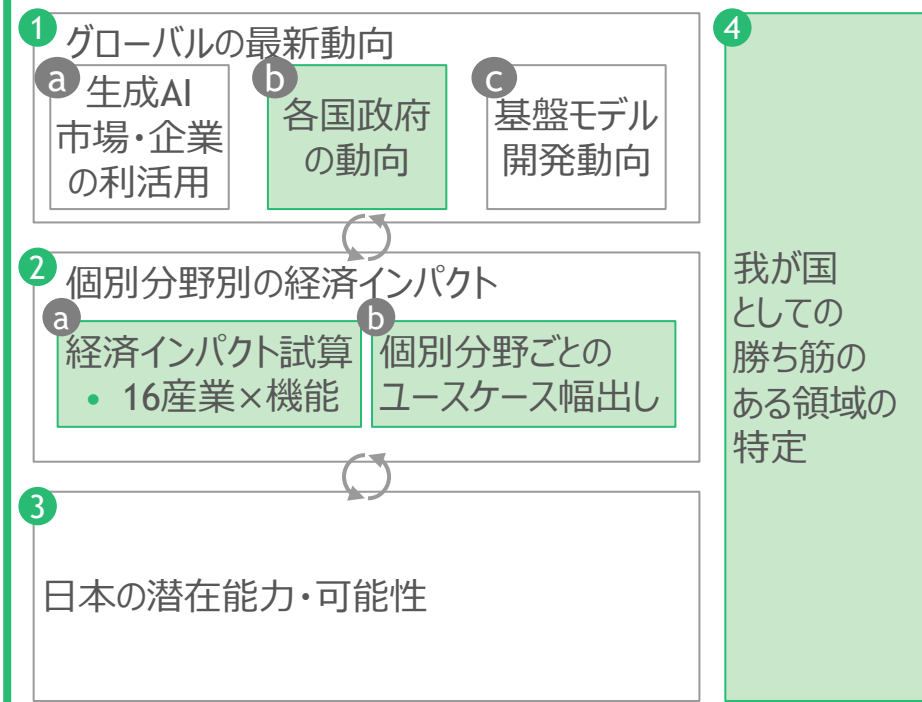


(1) 生成AI・基盤モデルに関する市場動向調査

事業実施の基本方針 : (1) 生成AI・基盤モデルに関する市場動向調査

□ :仕様書記載内容 ■ :追加提案

(1) 生成AI・基盤モデルに関する市場動向調査



(2) 基盤モデルの開発に必要なリソースの特定・確保の方策



(3) 開発成果の最大化の方策



10 AIに関する他政府検討との連携・整合

我が国としての基盤モデル開発の勝ち筋を、グローバルの動向/経済インパクト/日本の潜在能力・可能性の調査・分析を通じて特定

(1)生成AI・基盤モデルに関する市場動向調査: 検討アプローチ

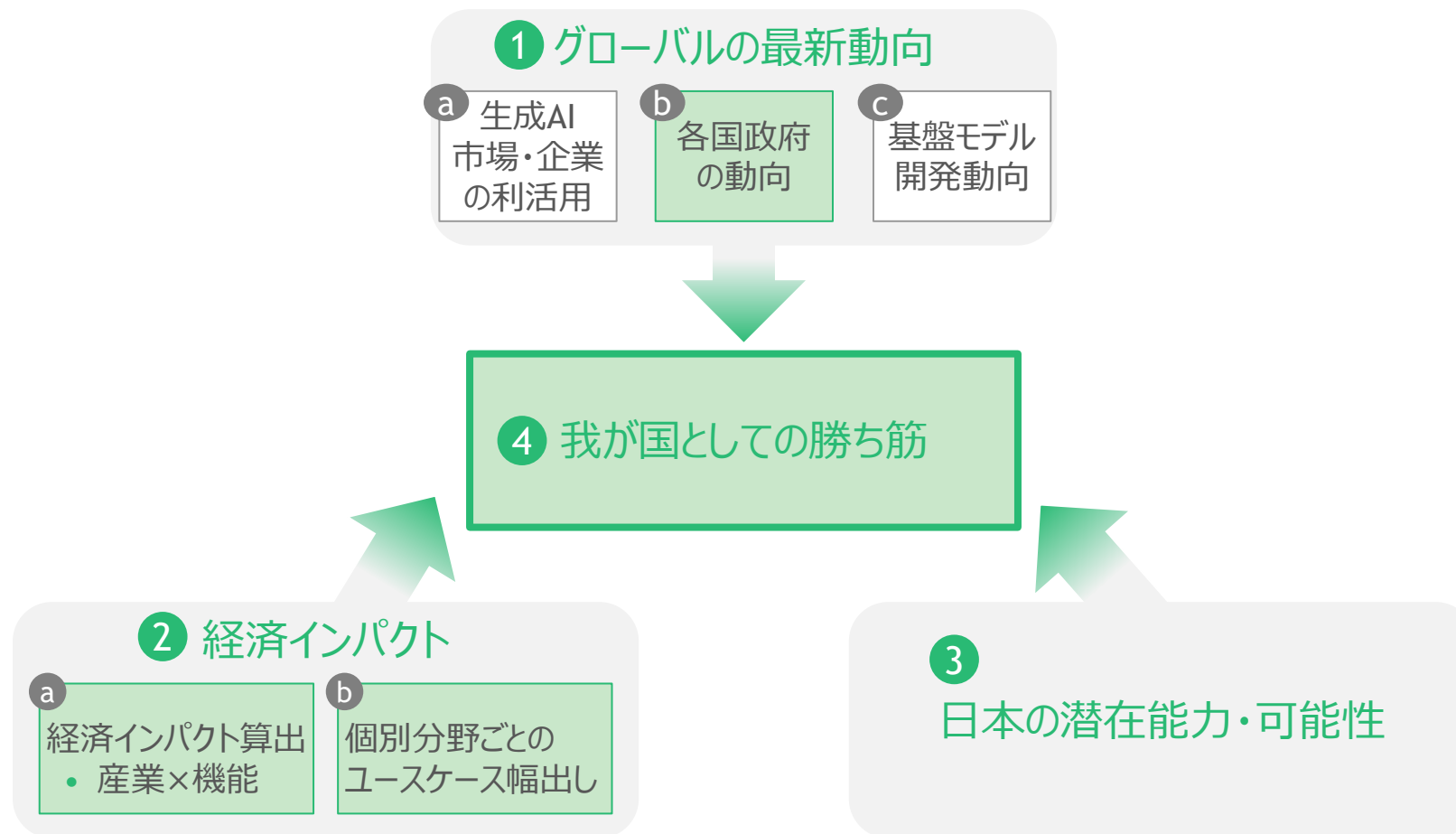
仕様書の記載内容

以下の項目について、生成AI、特に基盤モデル分野等における市場動向、企業動向、開発動向等について、調査・分析を行う

- ① グローバルの最新動向 (米欧中含む3カ国以上)
- ② 生成AIの個別分野別の経済インパクト (情報通信業、製造業、金融業を含む3種類以上)
- ③ 日本の生成AI分野における潜在能力・可能性等に関するヒアリングなどの調査、調査結果に対する分析

検討アプローチ

□:仕様書記載内容 ■:追加提案



開発領域として意義のある粒度で調査・分析を実施

(1)生成AI・基盤モデルに関する市場動向調査:調査・分析における粒度

考え方

生成AI基盤モデルの開発にあたっては、必要なリソース(データ)が特定可能な粒度で分析する必要

よって、産業x機能の粒度での個別分野の調査・分析を行った上で、有望な個別分野についてはユースケースレベルまで調査・分析を実施

- 産業**
 産業分類大分類を基本とし、必要となるデータが大きく異なる産業は中分類レベルで再分類
- 機能**
 企業活動基本調査での事業組織分類を基本とし、必要となるデータが大きく異なる機能は再分類

調査・分析における粒度

		機能								合計値
		R&D	SW開発 コンテンツ制作	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務横断 (資料作成等)	
産業	自動車・鉄道・船舶・航空									
	製薬・素材・化学									
	その他製造業									
	卸売・小売									
	医療									
	介護									
	建設									
	サービス									
	運輸・郵送									
	情報通信									
	メディア・エンタメ									
	金融・保険									
	プロフェッショナルサービス									
	教育									
	不動産									
	エネルギー									
	農林水産・鉱石砂利									
	合計値									

"経済インパクト" とともに "グローバルの市場動向" と "日本の潜在能力・可能性" を鑑み 基盤モデルの開発余地が大きい領域を特定

Illustrative

1 グローバルの最新動向

各国政府のAI戦略及び関連する取組みを要素別に調査・整理の上、注力領域を特定
各国政府の動向 [初期調査]

国	注力領域
EU	AIの倫理的な使用、AIの透明性、AIのセキュリティ
米国	AIの安全保障、AIのセキュリティ、AIの倫理的な使用
中国	AIの安全保障、AIのセキュリティ、AIの倫理的な使用
韓国	AIの安全保障、AIのセキュリティ、AIの倫理的な使用

2 経済インパクト

AI導入による効率化・コスト削減インパクトを17産業×8機能で分解し算出
効率化・コスト削減インパクトを17産業×8機能で分解し算出

産業	効率化・コスト削減インパクト (億円)
自動車・鉄道	605
製造業	945
卸売・小売	331
医療	15
建設	869
サービス	24
金融・保険	21
教育	7
不動産	314
エネルギー	11
その他	13
合計	4,895

3 日本の潜在能力・可能性

日本の潜在能力・可能性においては、産業的国際競争力・データをもつ製造業・医療や、
日本独自の文化を有するカスタマーサービス・アニメーションが有望

③ 日本の潜在能力・可能性

エキスパートの育成

- 製造業**
 - 日本が産業的に強い製造業・モノづくりにおいて、最先端技術を開発する能力を有している。特に、高度な技術力を持つ製造業は、AIの導入による効率化・コスト削減に大きく貢献する可能性がある。
- 医療**
 - 高齢化の進展や医療費の増加に伴って、医療現場でAIを活用する需要が高まっている。AIが診断や治療の最適化に貢献することで、医療の質を向上させる可能性がある。
- アニメーション**
 - アニメーション制作は、AIの導入による効率化・コスト削減の効果が期待されている。AIがキャラクターの生成や背景の作成に貢献することで、制作コストを削減し、制作スピードを向上させる可能性がある。
- カスタマーサービス**
 - AIチャットボットや音声アシスタントの導入により、顧客サービスの効率化・コスト削減が期待されている。AIが顧客の問い合わせに対応することで、顧客満足度を向上させる可能性がある。
- その他**
 - AIの導入による効率化・コスト削減の効果が期待されている。AIが業務の自動化やデータ分析に貢献することで、生産性を向上させる可能性がある。

4 基盤モデル開発の有望領域

(1) 生成AI・基盤モデルに関する市場動向調査
我が国としての有望領域の特定

8つの領域をカバーすることで効率化・コスト削減による経済インパクトの9割に当たる15兆円の創出が見込める
② 領域に特化した効率的な基盤モデル開発: ④ 効率化・コスト削減における有望領域

効率化・コスト削減による経済インパクト初期試算結果

産業	単位:億円	機能								合計値
		R&D	SW開発 ソフトウェア制作	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務構築 (資料作成等)	
自動車・鉄道	605	4	51	1,234	195	307	721	1,019	4,135	
製造業	945	32	454	3,066	1,404	561	1,781	2,695	10,938	
卸売・小売	331	128	4,422	341	12,160	2,584	615	2,660	16,444	
医療	15	220	277	86	660	11,911	615	2,660	27,587	
建設	869	79	366	2,639	1,133	535	1,428	2,319	9,367	
サービス	24	339	426	132	1,016	16,560	948	4,096	23,540	
運輸・郵便	12	166	209	65	498	8,112	464	2,007	11,532	
テレコム/ソフトウェア	125	5,745	84	12	296	1,140	995	1,378	9,775	
メディア/エンタメ	12	57	363	17	1,021	789	186	568	3,012	
金融・保険	21	131	392	44	1,083	5,884	1,275	1,548	10,378	
アパレル/ファッション	432	786	149	123	542	5,080	948	1,402	9,461	
教育	7	94	118	36	281	4,580	262	1,133	6,510	
不動産	314	29	133	955	410	194	517	839	3,389	
エネルギー	11	1	18	298	102	73	205	235	944	
農林水産・ 鉱石燃料	13	0	14	174	49	43	198	158	650	
合計値	4,895	8,049	8,131	12,766	22,759	66,425	15,291	31,913	170,229	

有望領域 (初期仮説)

- オペレーション/サービス/CS
- 業務構築 (資料作成等)
- 製造業×R&D-
セールス/マーケティング
- バックオフィス
- 卸売・小売×セールス/
マーケティング
- 医療×オペレーション/
サービス/CS
- SW開発/コンテンツ制作
- 金融・保険×セールス/
マーケティング×オペレーション/
サービス/CS

効率化・コスト削減

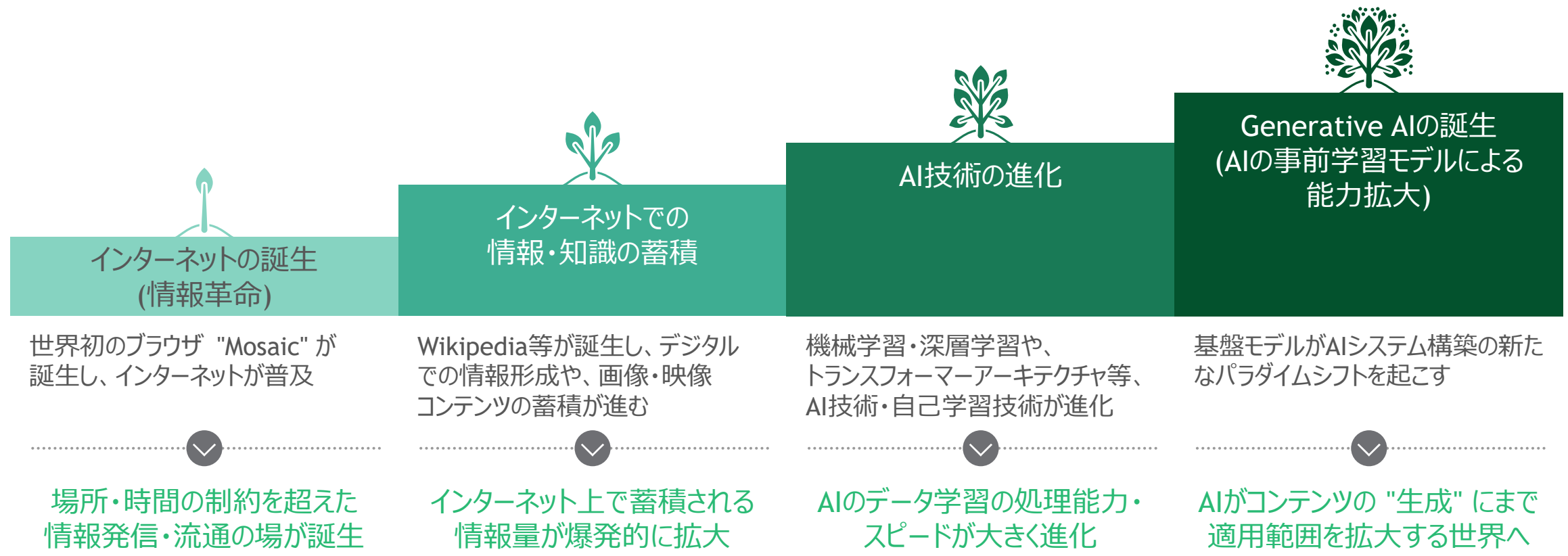
5,000億円以上
1,000億~5,000億円未満
500億~1,000億円未満
500億円未満

Copyright © 2022 by Boston Consulting Group. All rights reserved.

① グローバルの最新動向

Generative AIの台頭により、AIが "生成" 領域に進化し、拡張されることで、大きな社会的変化を引き起こす

生成AI登場による社会的変化



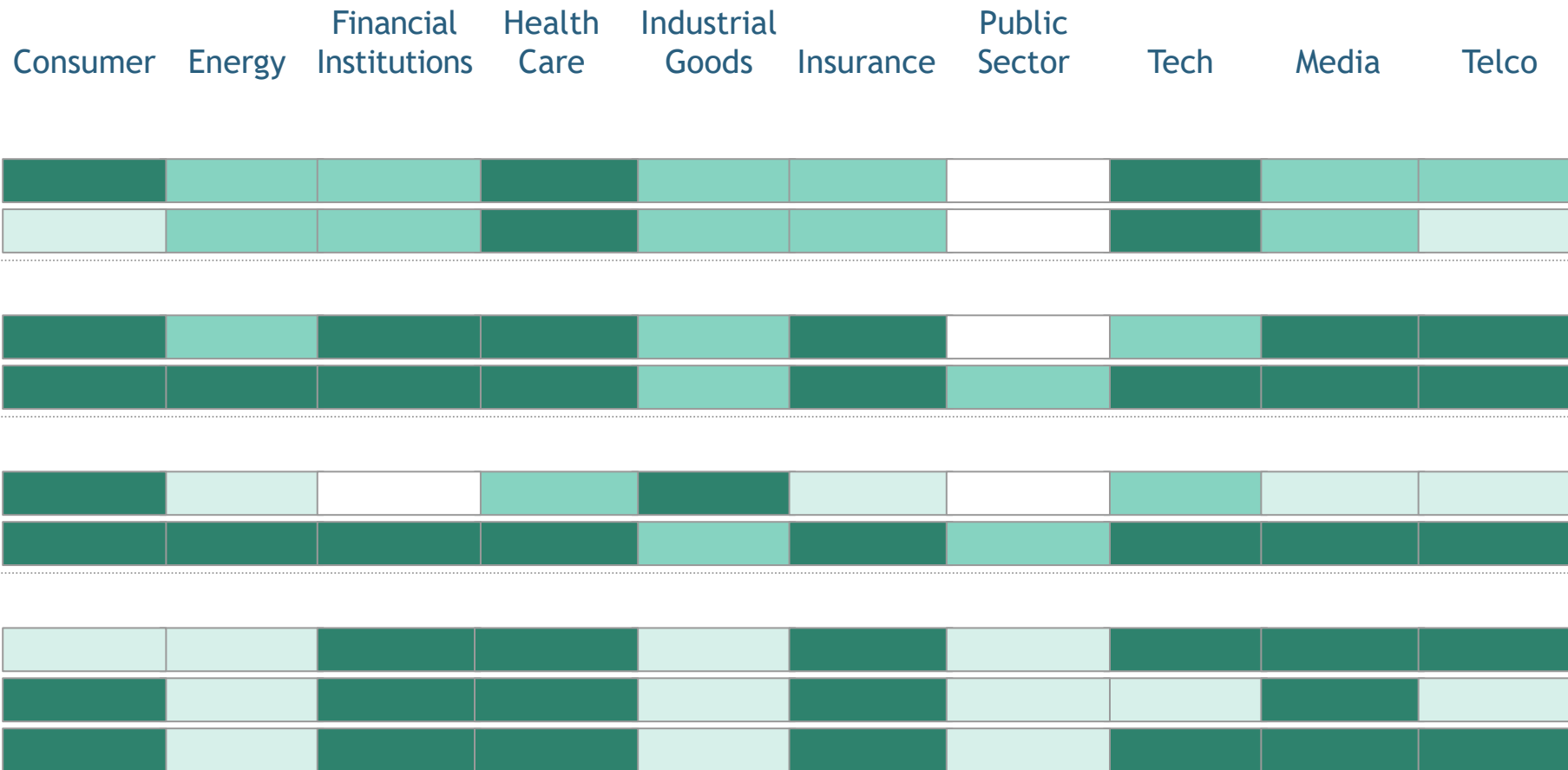
BCGのグローバル調査においても、全産業・全機能に対してインパクトが見込まれているが、インパクトの大きさは領域によって異なる

潜在的な経済効果



小 大

産業



機能

基盤モデル開発の有望領域を見極める上では、4つの要素を考慮する必要があり、各国政府の動向が重要となる



規制

生成 AIに関する規制/法律¹

- データ保護とガバナンス
- データとコンテンツ作成の責任
- コンプライアンスと報告義務



基盤モデルの性質

どの基盤モデルがグローバルで勝利できるか

- 大規模基盤モデル
- 領域固有の基盤モデル (地域、セクター、アプリケーション、...)



貿易制限

国家間での財・サービスの移転制限

- 計算能力に影響する米中の半導体輸出制限
- 米中間の投資の制限



地域の文化・言語の特性

文化的／言語的特異性や、既存のエコシステムプレイヤーへの依存度合い



1. およびデータ保護、AI等の分野に関連する。
出所: BCG分析

各国政府のAI戦略及び関連する取組みを要素別に調査・整理の上、注力領域を特定

各国政府の動向

								
AIの位置づけ	産業領域でのAI活用は民間企業が主導し、政府は準公共領域を支援	米中への対抗から域内間でのデータ活用促進・厳しい規制・ルール導入	将来的なドイツの競争力保持を目指した取組み	AI人材確保、オープンデータポリシー、倫理枠組みに焦点を当てた取組み	科学・AI大国としての地位を保持するため中長期的な取組み	都市国家としての行政・市民サービス向上に活用	公共部門のデータを民間でのAI開発に提供する等積極的支援	AIでの覇権を目指して研究開発、人材育成、AI倫理、関連法規制の整備を推進
ルール・規制	民間開発によるリスク抑制に取組み ・「AI権利章典のための青写真」	域内でのルール整備を推進 ・ AI規制法案・製造責任規定等	グローバル標準の策定をリード ・ ISO/IEC42005 (AI影響評価等)	倫理的規制の整備 ・ カナダとパートナーシップ締結	AI利用を規定する法律はない ・ 現行法でカバー	企業の利活用におけるリスクおよび対応方針を策定	AI産業の信頼性強化のため民間主導で生成AI技術に対応したAI倫理の標準ルールを策定	生成AIにおけるガバナンス強化 ・ ガイドライン等のルール導入
AI活用/サービス	輸送、ヘルスケア、農業	ヘルスケア、交通、農業	ヘルスケア、介護、気候環境、航空宇宙、モビリティ、農業	健康、交通、環境、防衛、安全保障	ヘルスケア、社会福祉、国防	ヘルスケア、行政、物流、教育、国境警備	交通・物流、ヘルスケア、製造、公共サービス	農業、防衛、デジタルエコミー、教育、雇用、ヘルスケア、起業、社会福祉、交通
データ	民間での推進に委ね、公共分野のみ政府が介入 ・ eHealth Exchange	企業間・地域間のデータ連携を推進 ・ Gaia-X (産業間でのデータ連携基盤) ・ Catena-X (自動車産業でのデータ連携基盤)			—	各他プレイヤーの機能を活用	各他プレイヤーの機能を活用	AI基盤の国家標準システム確立に向け、主要民間企業を指定企業として設定 ・ 百度、アリババ等
基盤モデル	—	EU域内での汎用基盤モデル: OpenGPT-X に資金提供 ・ 米中への対抗を図り、域内のデータ連携基盤であるGaia-Xと互換 ・ ドイツ語、英語、フランス語、スペイン語、イタリア語に対応 ・ ドイツを中心に、約1,400万ユーロの資金提供	—	—	自国内企業によるAI基盤モデル開発支援 ・ 安全で信頼性の高い基盤モデルの開発・構築支援を目的に1億ポンドを拠出 ・ 政府内にタスクフォース立ち上げ	—	民間での開発支援のため、韓国語データを整備	主要民間企業を指定企業として設定し、データ・リソースを集約 ・ AI基盤の国家標準システム確立を目指す ・ 百度、アリババ等

国内背景および地政学的観点も踏まえ、各国のAIにおける政府の介入・取組みが異なる

政府の介入度合い

A 企業主導型

国の例



主な取組み

民間での積極的な開発・活用を背景に、産業領域は民間での推進に委ね、公共分野のみ政府が介入

- 民間での推進によるリスク抑制は政府としても取組み

B 利活用促進型



国家規模によるデータ制約を背景に、AI基盤・モデルの開発は他プレイヤーに委ね、それらの自国内の利活用を促進

- シンガポール：企業の利活用におけるリスクおよび対応方針を策定
- 韓国：生成AI技術に対応したAI倫理の標準ルールを策定

C 米中対抗型



米中への対抗から域内間でのAI活用促進およびAI基盤開発を進めるとともに、厳しい規制・ルールを導入

- EU：域内でのデータ連携基盤を整備するとともに、それと互換するAI基盤モデルの開発を支援
- ドイツ：ISO等のグローバル標準の策定をリード
- フランス：倫理的規制の整備を進め、カナダとパートナーシップ締結
- イギリス：科学・AI大国としての地位の保持を目的に、自国内企業によるAI基盤モデル開発支援

D 中央集権型



党/政府の強力な統制により、AI基盤の国家標準システム確立に向け、政府が主要民間企業を指定企業として設定

- データ・開発リソースを集約し、開発スピード・成果を最大化



米国のAI戦略概観



前提となる社会情勢/背景

自由な社会・経済を最優先し、政府の関与 (規制) は最小限とするのが基本的な考え方

近年は、2000年代の世界経済を牽引した米国巨大IT企業に対する国内外からの強い反発への対応と健全な市場維持のため、規制と自由のバランスを模索



AI戦略の位置付け/特徴

AI分野における世界的リーダーとしての地位を強化する手段としてAI戦略を位置づけ

産業領域でのAI活用は民間企業が主導し、政府は準公共領域を中心に支援



取組状況

[規制・ルール]
健全な市場の維持・促進のための最小限の規制

[AI基盤]
主な取組はなし



EUのAI戦略概観



前提となる社会情勢/背景

歴史的に、政府は個人の人権やプライバシーを保護する責務を負うという考え方

加えて、EUを単一市場として捉え、域内市場を外部から守る意識/米中への対抗意識も強い



AI戦略の位置付け/特徴

厳しい規制・ルールの導入による米中による占有への対抗

域内の経済成長/イノベーション活性化のため、**国・産業を超えたAI・データ基盤**を構築する動き



取組状況

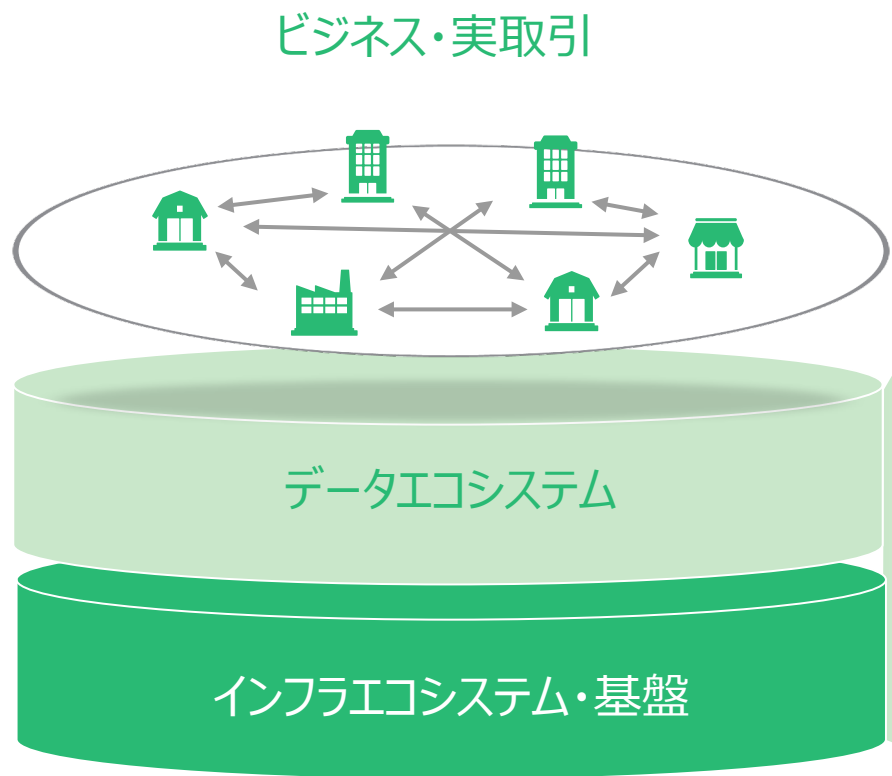
[規制・ルール]
AI Actを発表し、リスクカテゴリに応じた規則を制定

[AI基盤]
Gaia-Xとの互換性を持つEU域内でのLLM:OpenGPT-Xの開発



従来から系列がないEUでは、企業間取引を円滑化するために、データや規格の標準化が進められてきた

EUの動向: 全体像



International Data Space

- 「データエコシステム構築」のため、データ交換の標準アーキテクチャー、ルール・ガイドライン、データ交換のための技術要素 (コネクター) 等を定義・提供

GAIA-X

- 「インフラエコシステム構築」のため、関連するルール・ポリシー、通信・相互接続サービス等を定義・提供
- データ交換においては、IDS定義に準拠し、IDSコネクターを活用

データ・インフラエコシステムを活用した産業別データスペース (Catena-X, アグリGaia 等)



EUでは多くの標準化が進められ、複数取組みを跨る形で、より包括的な企業間連携を実現しようとしている

EUの動向: EUにおける標準化の偏移



個別領域・機能における標準化

- 安全規格、技術企画、EDI等の特定業務領域でのデータ・通信の標準化

安全・技術企画
(CEN/CENELEC/ETSI)

EDI標準 (TDI/EDIFACT)



産業・テーマ横断での標準化

- 業界横断もしくは、バリューチェーン・テーマ横断での標準化、ユースケース確立を推進

Industry 4.0 (製造)

Catena-X (自動車産業)



複数産業・テーマをつないだインフラまで含む標準化

- 複数のエコシステムやデータスペースを跨ったデータ流通のための標準化
- 更に、データインフラ領域まで踏み込み標準化を推進

新規検討中のユースケース

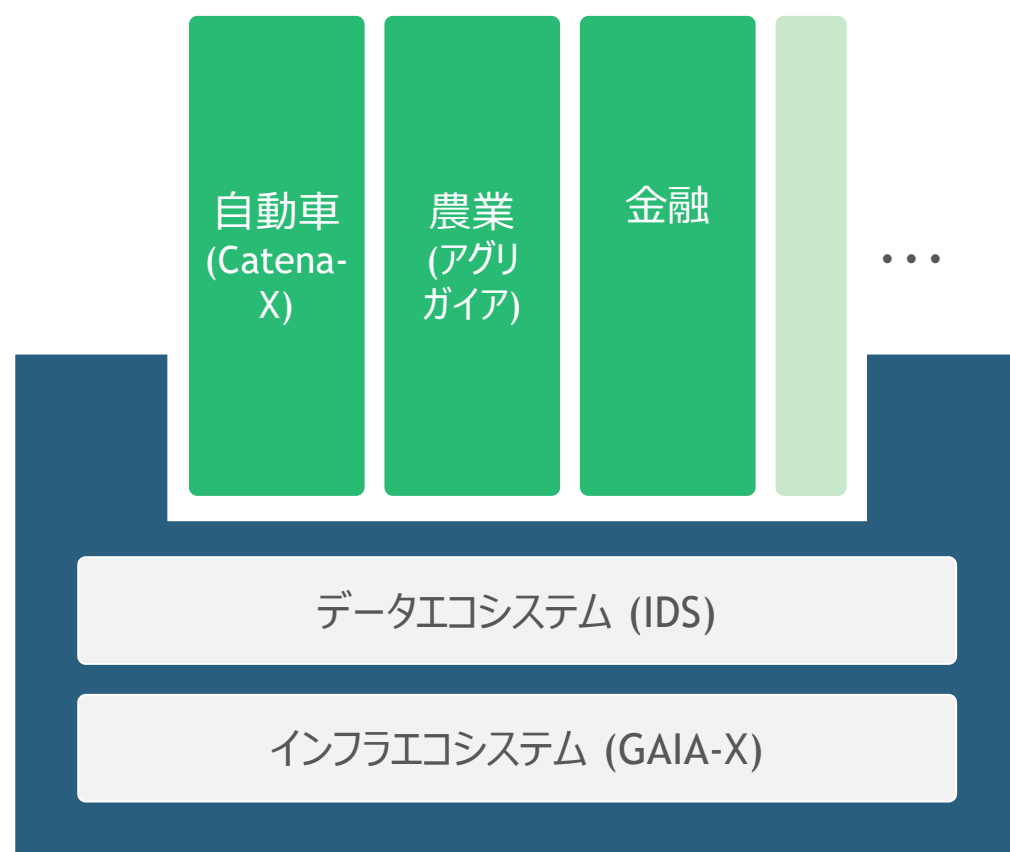
Catena-X (複数産業連携)

GAIA-X/ International Data Space



その中で、産業別の規格化による競争優位の確立に加え、複数産業間の相互連携によるデータ蓄積・処理まで含むインフラ基盤を構築しようとしている

EUの動向: EUにおけるデータ標準化・規格の全体像



産業別の規格化による優位性の確立

- 非競争領域を中心に、規格の標準化を進めることで、企業は非競争領域で規格を活用し無駄を省くと同時に、相互接続性を確保し効率的に連携できる
- 一方で、競争領域では、これまで通りに個社別に差別化を図ることができる

複数産業の相互連携・データ処理含むインフラ基盤の構築

- 産業間連携・依存性が高まる中で、複数産業・データスペースの相互連携性を実現し、より広範囲のデータにおける主権性を確立
- またデータを共有・蓄積するだけでなく、処理・分析まで行えるインフラエコシステムの実現により、データ価値を最大化



ポジティブ、ネガティブ両面でのインセンティブを働かせ、参加企業を増加

EUの動向: Catena-Xの普及方法

ポジティブインセンティブ

排出量データ獲得



- ネットワーク内にいれば、サプライヤーからデータを容易に獲得が可能
 - Catena-Xで共有するのがスタンダードになれば、その他の手段は忌避される可能性も

経済性の向上



- 多数の企業で共通インフラを構築/管理すれば、自社単独よりも費用が安い
 - ドイツの自動車業界では、SC排出量の可視化等は避けられない状況だという認識が広がっている

信頼性の確保



- Catena-XはドイツやEUの政府のお墨付きがある取り組みであるため、自社のデータへの信頼性が高まる

ネガティブインセンティブ

ビジネスロックアウト



- Catena-Xのネットワークの内部にいないと、自社製品の販売で不利に
 - BMWがCatena-Xをリードしており、ドイツ自動車業界での影響力大



データ取得コストの低減、高い信頼性を売りに参加企業を増やし、Catena-Xに参加しないと自動車産業でのビジネスがやりにくい構造の構築を目指す

中国のデータ戦略概観

グローバル

勝ち筋

経済

日本



前提となる社会情勢/背景

強権的な一党独裁が続き、国内の政情が不安定

中国政府の後押しや巨大な内需を背景に、BATに代表される世界的な巨大ITプラットフォーマーが台頭



AI戦略の位置付け/特徴

AI戦略を中国がAIによるイノベーションの中心になるためにAI戦略を作成

2020年から2030年にかけて理論・技術・応用の面で世界の先端水準になることを計画



取組状況

[規制・ルール]
政府が強権的にトップダウンで生成AIまで射程に入れた規制の整備を推進

[AI基盤]
AI基盤の国家標準システム確立に向け、主要民間企業を指定企業として設定



「BATIS」の5大AI基盤に加え、2019年には、さらに10の民間AI基盤を追加

中国の動向: 国家戦略に基づくAIプラットフォームの指定

「国家人工知能オープン・イノベーション・プラットフォーム」のリスト更新

科学技術部は、2019年8月、新たなプラットフォームを追加指定

- 従来5基盤に新たに10基盤を加え、合計で15基盤に拡大

<p>依図科技 (YITU)</p> <ul style="list-style-type: none"> 顔認証SW企業 全土12,000以上のATMで顔認証システムが採用 <p>✓</p> <p>画像コンピューティング</p>	<p>明略科技 (MiningLamp)</p> <ul style="list-style-type: none"> 機械学習開発企業 30以上の省で治安ナレッジグラフPF「DataInsight」が採用 <p>✓</p> <p>マーケティングインテリジェンス</p>	<p>華為 (HUAWEI)</p> <ul style="list-style-type: none"> 通信機器・端末メーカー 世界最大の通信ベンダーかつ世界3位のスマホメーカー <p>✓</p> <p>SW / HWインフラ</p>	<p>中国平安 (PINGAN)</p> <ul style="list-style-type: none"> 保険 / 総合金融企業 生保・損保とも国内2位。フィンテック・ヘルステックで先行 <p>✓</p> <p>インクルーシブファイナンス</p>	<p>海康威視 (HIKVISION)</p> <ul style="list-style-type: none"> 監視カメラメーカー 世界最大の監視カメラメーカー。親会社は国有企業 <p>✓</p> <p>動画認識</p>
<p>京東集団 (JD)</p> <ul style="list-style-type: none"> EC(ネット通販)企業 アリババに次ぐ中国2位のEC「JD.com」を運営 <p>✓</p> <p>スマートサプライチェーン</p>	<p>旷視科技 (MEGVII)</p> <ul style="list-style-type: none"> 顔認証SW企業 顔認証プラットフォーム「Face++」を展開 <p>✓</p> <p>画像認識</p>	<p>奇虎360 (QIHOO)</p> <ul style="list-style-type: none"> セキュリティSW企業 ブラウザ / 検索エンジンも提供。百度に次ぐシェア2位 <p>✓</p> <p>サイバーセキュリティ</p>	<p>好未来教育 (TAL)</p> <ul style="list-style-type: none"> 教育サービス企業 中国学習塾最大手。英語、資格試験等に多角展開 <p>✓</p> <p>スマート教育</p>	<p>小米 (XIAOMI)</p> <ul style="list-style-type: none"> スマホ・総合家電メーカー 2018年にはスマホ出荷台数1億台。家電にも事業拡大 <p>✓</p> <p>スマート家電</p>



加えて、スタートアップ等も含めた研究開発を推進、開発拠点・規則整備を推進

中国の動向: AI戦略

AI戦略

名称: 次世代AI発展計画

発行機関: 国務院

発行年: 2017年

目的

- 経済・社会発展の原動力となるAI技術の発展を大きな戦略的チャンスと捉え、先行者優位性を築き、グローバルでの先進科学技術国家としての地位を確立

ステップ

目標

2020年	技術・応用が世界の先進レベルに到達。AI産業を基軸として経済成長を実現し、国民の生活を向上
2025年	基礎理論がブレークスルーを起こし、AIの技術と応用が世界先進レベルに到達。AIが産業高度化と経済再編の主要な原動力なりAI社会化が進展
2030年	理論、技術、応用が世界的な先進レベルに到達。イノベーション主導の経済強国となる強固な地盤を構築

AI研究環境の整備

- 北京にAI研究に特化したテクノロジーパークを建設予定
 - 400の企業が入居する予定で、年間77億ドル（500億元）の収益を見込む
- 地方でも、深圳や上海などの主要都市は、独自の規則を公布
 - 上海AI条例では、科学のフロンティアの探求を奨励し、イノベーション促進のため、軽微な違反に対して一定の許容範囲を容認



シンガポールのAI戦略概観



前提となる社会情勢/背景

小資源・人口ながら、高付加価値産業の推進により、1980年代以降、大きく経済発展

急速な都市化・少子高齢化の進展による都市政策・経済政策の複雑化への対応が急務



AI戦略の位置付け/特徴

地理的制約を超えた成長および国民生活向上の手段としてAI戦略を位置付け

積極的に生成AIを公共サービスへ取り入れ



取組状況

[規制・ルール]
生成AIのリスク特定および対応フレームワークを整備するなど先進的な取組み

[AI基盤]
主な取組はなし



シンガポールは、公共サービスを中心にAIの活用を推進

シンガポールにおけるAI戦略

AI戦略

名称: 国家AI戦略

発行機関: スマート・ネイション・デジタル政府局 (SNDGO)

発行年: 2019年

目的

- 地理的制約を超えた成長のフロンティアをAIに見出す
- 国民の生活向上を主眼に、特に交通・教育・医療といった公共サービスの質向上を目指す

注力領域

主な取組

交通

企業の生産性と交通効率を向上させるための最適化された貨物輸送計画

- 輸送トラックの自動割り当て、経路最適化、都市全体での交通最適化

ヘルスケア

慢性疾患の予防・管理支援

- リスクのスコア化、医師向けの臨床判断の支援、患者向けの自己管理サポート

教育

各人ごとに最適化された学習の提供

- 強み・弱みに応じた学習、学習データからインサイト抽出、AIアシスタントの伴走

国防

旅行者の旅行体験向上と国境警備の両立

- 入国審査の自動化、AIによる国境警備の強化

公共サービス

迅速で信頼できる自治体サービス

- チャットボットでの問い合わせ、センサー・AIでの予知保全、施設利用状況把握

基盤モデルへの取組

- Google Cloudの生成AI機能とパートナーエコシステムを活用し、次世代の市民中心型サービスの開発を加速
- 公務員の事務タスク迅速化のため、ChatGPTを支えるAI言語モデルの利用を検討
- GovTech部門で使用している88のチャットボットを全て生成AIモデルに移行予定



韓国は、国民の生活の質向上を目的にAI活用を進め世界3位のデジタル競争力を目指す

韓国におけるAI戦略

AI戦略

名称: AI国家戦略
発行機関: 韓国政府
発行年: 2019年

目的

- 世界第3位のデジタル競争力
- AIを通じて455兆ウォンの経済効果を創出
- 生活の質において世界10位以内

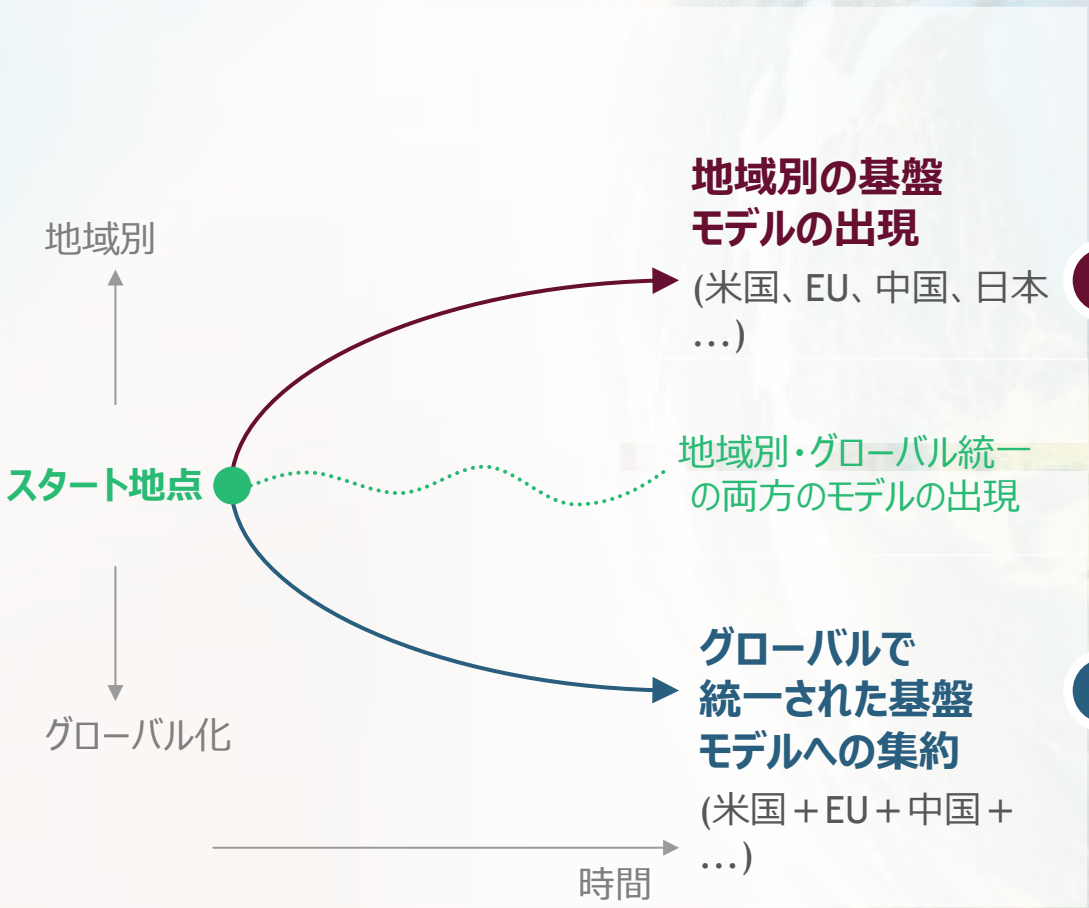
注力領域	主な取組
製造	AIとデータに基づくスマートシティを推進
ヘルスケア	AIを活用して生化学・医学業界を次の主要産業へと育成
建設/ スマートシティ	都市生活の生活の質と建築業界の安全性向上
交通	自律走行環境を整え、ロジスティクス業界を推進
エネルギー	エネルギー供給・消費の効率化
農業・漁業	スマートテクノロジーを用いて農業・漁業を高付加価値化
国防	コアタスクへのAI利用を進め効率性と信頼性を向上

基盤モデルへの取組

- AI産業における国家競争力を確保するため、企業・人材育成および韓国データの整備を推進
- 2026年までにエコシステムの中核となるSaaS企業を1万社育成
 - 医療、法律、カウンセリングなどの分野でAI活用サービスを強化
 - 民間主導、政府支援による基盤モデル連携エコシステムを構築
 - モデル開発に必要な高品質テキストデータの拡充
 - 民間セクターと協力し、2027年までに200種類の専門学習データと非英語圏市場向けの言語データを構築
 - 韓国語能力向上/応用のための130種類の高品質テキスト・コーパスを準備
 - AI専門家の増加および能力強化のため、2027年までに修士/博士のAI専門課程で65,000人の人材を教育・育成
 - さらに、同期間に100万人の国民に基盤モデルの訓練を実施

基盤モデルは、グローバルで統一されていくか、もしくは各地域で専用の基盤モデルが開発されるか、2つのシナリオが存在

基盤モデル開発のシナリオ



左記シナリオが実現するための条件

各地域の規制当局は、基盤モデル・データを地域内に配置するよう強制	大規模LLMの価値は限定的で、 領域固有の基盤モデルが付加価値を生む
GPU調達には基盤モデル開発の制約ではなくなっていく	顧客はソリューション選定において複数の基盤モデルを選択
規制当局は「 責任あるAI 」の 世界基準を重視	付加価値・競争優位のために、 LLMの規模拡大が必須
GPUは競争優位のために 必須 。 貿易制限 のため中国は競争から遅れる	顧客は 最高のパフォーマンスを実現するための基盤モデルを選択

今後、基盤モデルは技術、市場・企業、規制・ルール、AI倫理の開発動向を踏まえ、進展

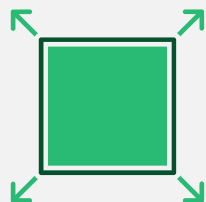
次頁以降、詳細

開発動向

概要

技術	大規模化	より広範囲なデータと計算資源に基づき、規模を拡大させたモデルが開発される
	領域特化	領域やタスク、言語を限定し、部分的に汎用モデルを上回る性能を発揮するモデルが開発される
	効率化	データ品質向上等を通じ性能を担保しつつサイズを小規模にしたモデルを開発する
	拡張	複数の入力フォーマット(テキスト、画像、音声、動画等)を統合して処理する
市場・企業	市場浸透	期待される経済効果に応じて、産業・機能毎に異なる速度で市場浸透が進む
規制・ルール	地域特化	厳しい規制・ルールの導入が進み、地域に特化したモデルが開発される
AI倫理	信頼性・透明性向上	モデルの出力や判断が信頼でき、透明性を向上させる
	公平性担保	モデルの出力や判断にバイアスが含まれず、公平性を担保する
	プライバシー・セキュリティ担保	著作権やユーザーデータのプライバシーを保護しつつ、セキュアに情報を処理する

技術における開発動向は、「大規模化・領域特化・効率化・拡張」の大きく4つが存在



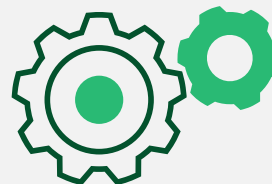
大規模化

- モデルパラメータ数
- コンテキストウィンドウ



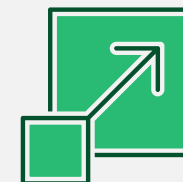
領域特化

- 医療・創薬
- ソフトウェア開発・セキュリティ
- 広告



効率化

- 高品質なデータ
- 効率的な学習アルゴリズム
- 学習高速化ミドルウェア
- モデルの圧縮



拡張

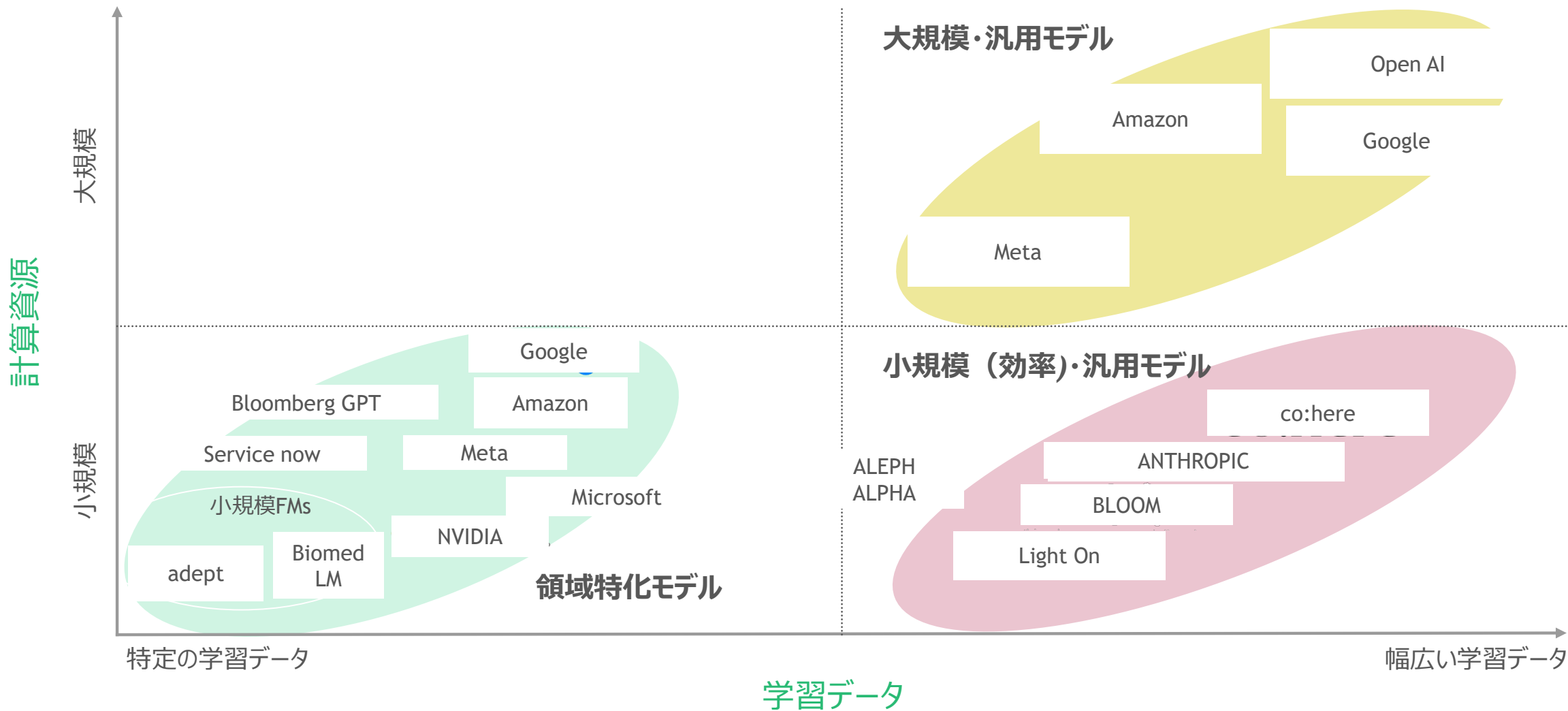
- 物理世界への拡張
- モーダルの拡張
- エージェント

グローバルメガプレイヤーは4つの開発動向を踏まえ、取り組みを進めている

直近のグローバルメガプレイヤーの開発動向

プレイヤー	開発動向	詳細
Open AI	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> GPT-5の商標登録、開発に向けたデータ収集内製化の動き ChatGPTでデータ分析や画像が利用可、ロボティクスへの展開の動きも
Google	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> PaLM 2をロボティクス領域への拡張 PaLM 2を医療領域向けに特化させたMed-PaLM 2を展開 Google Cloud Platform上のフルマネージドの機械学習プラットフォームVertex AIで利用可
Amazon	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> Anthropic、Stability AI等、他社スタートアップが開発した基盤モデル活用にも積極的 テキストや画像等で学習し、マルチモーダルで汎用性の高いAmazon Titanを開発 生成AIを活用した新サービスAWS HealthScribeを発表
Meta	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> Llamaよりも多い700億パラメータでオープンソースかつ商用利用可能な大規模言語モデルLlama2を開発 マイクロソフトと提携しAzure経由で企業向けにLlama2を提供開始予定 Llama2をプログラミング支援に領域特化させたCode Llamaや、マーケティング支援AI Sandboxを立ち上げ テキストから動画を作成するMake-A-Videoの論文を公表
Stability.ai	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> OSS中でもトップレベルのパラメータ数を有する画像生成AI Stable Diffusion XLを2023年7月公開 日本語向け画像言語モデルや汎用言語モデルJapanese StableLMも開発しOSSとして公開する等、日本との親和性高
NVIDIA	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> 高性能GPUの販売を手掛けつつ、商用利用可能な最大5,300億パラメータの大規模言語モデルMegatronを開発 基盤モデル開発用フレームワークMegatron-LMや高速化フレームワークTensorRT-LLMを公開 タンパク質や生体分子の構造と機能を迅速に生成できる生成AIサービス BioNeMoを展開 Megatron-Turing Natural Language Generation (MT-NLG) をMicrosoftと共同開発
Microsoft	大規模化 効率化 領域特化 拡張	<ul style="list-style-type: none"> OpenAIのGPTシリーズやMetaのLlama2をAzure Platform上で利用できるようにする等他社基盤モデル活用に積極的 トークンサイズの大規模化やデータ品質の向上による効率化を研究 深層学習最適化フレームワークDeepSpeedを公開 Megatron-Turing Natural Language Generation (MT-NLG) をNVIDIAと共同開発

グローバルメガプレイヤーが大規模・汎用な基盤モデル開発で先行する中で、他企業の開発は多様化



グローバルメガプレイヤーは大規模・汎用基盤モデルを開発し、精度向上のため開発強化

現時点での各基盤モデルのスペック (大規模・汎用)

		Open AI		Amazon		Meta		Google		
		GPT-3		GPT-4		AmazonTitan		Llama 2		PaLM 2
生成物		テキスト		マルチモーダル		テキスト		テキスト		マルチモーダル
	量	570GB (~3,000億トークン)		非公開		非公開		2兆トークン		3.6兆トークン
	内容	基本的にはインターネット上のオープンなデータ ただし一部各社の内部データや著作物が含まれる可能性も指摘されている								
	収集方法	オープンデータを活用								
モデルサイズ (パラメータ数)		1,750億		非公開 (1.7兆以上か)		不明		70億/130億/ 700億		非公開 (PaLMは5,400億)

Bloomberg、ServiceNow等は業界・言語特有のユースケース創出を目指し、領域特化モデルを開発

現時点での各基盤モデルのスペック (領域特化)

	Bloomberg GPT	Service now	Biomed LM	adept	Stability.ai	
モデル名	BloombergGPT	StarCoder	BioMedLM	ACT-1	StableLM Suite	
目標	金融業界でAIの潜在能力を最大化させる	プログラミングをより効率的に進められるように開発者を支援する	バイオメディカルNLPアプリケーションの開発と、ドメイン固有の言語モデルの開発推進	コンピューターを使用して人間が出来ることをすべてAIが行うことができるようにする	小型且つ効率的なモデルを適切にトレーニングすることでテキストやコードを生成する	
ユースケース	<ul style="list-style-type: none"> 金融NLPタスクの向上におけるブルームバークの業務支援 	<ul style="list-style-type: none"> コードの生成による開発者の生産性向上 	<ul style="list-style-type: none"> 他のドメインでの活用 	<ul style="list-style-type: none"> コンピューターを使ったあらゆる業務をACT-1が実施 	<ul style="list-style-type: none"> 小型機を使った様々なアプリケーションへの活用 	
データ	量	7,000億トークン	1兆トークン	3,000億トークン	不明	約1.5兆
	内容	<ul style="list-style-type: none"> 社内金融データ (3630億トークン) オープンデータ (3450億トークン) 	<ul style="list-style-type: none"> BigCodeのThe Stack v1.2データ <ul style="list-style-type: none"> 80以上のプログラミング言語をカバー 	<ul style="list-style-type: none"> 生物医学の抄録と論文 The Pile 等 	不明	<ul style="list-style-type: none"> The Pileをベースに構築された新しい実験用データセット
	収集方法	オープンデータやブルームバーク・ターミナルの活用	ライセンスの遵守のもとでThe Stackに収集されたコード	MosaicMLと提携	不明	The Pileの規模が3倍で、1.5兆トークンのコンテンツを作成して使用
モデルサイズ (パラメータ数)	500億パラメータ	150億パラメータ	27億パラメータ	不明	30億/70億パラメータ	
		業界			機能	

リリース前のため情報が限定的

グローバルメガプレイヤーも、医療、SW開発、広告等業種・機能特化の基盤モデルを開発

基盤モデルの開発の方向性 (領域特化)

	Google	Amazon	Meta	NVIDIA	Microsoft		
モデル名	Med-PaLM	AWS Health Scribe	CodeWhisperer	AI Sandbox	MegaMolBART	Security Copilot	
目標	マルチモーダルデータを効果的に符号化、統合、解釈することで幅広いバイオメディカルタスクに取り組む	患者と医師の会話による臨床記録を自動生成して臨床アプリケーションを構築するヘルスケアソフトウェアベンダーを支援する	複数のプログラミング言語向けに人工知能 (AI) を活用したコード提案を提供する	広告作成の効率化とクリエイティブ性の向上を行う。さらに、オーディエンスに応じて広告文章を変更する等パーソナライゼーションを推進する	コンピュータと情報化技術を化学領域の幅広い問題への応用させることで、創薬を支援する ・アストロゼネカと協働開発	LLMとセキュリティモデルを組み合わせたモデルを作り、セキュリティアナリストの作業時間を低減	
ユースケース	<ul style="list-style-type: none"> 臨床意思決定支援 創薬 医療研究 	<ul style="list-style-type: none"> 文書生成時間短縮 ワークフロー効率化 診察記録の整理 	<ul style="list-style-type: none"> ソフトウェア開発の簡素化と効率化 	<ul style="list-style-type: none"> 広告作成の時間とコストの削減 パフォーマンスとパーソナライゼーションの向上 	<ul style="list-style-type: none"> 大規模な医薬類似化合物のデータベース検索やスクリーニング、整理が可能 	<ul style="list-style-type: none"> セキュリティ問題への対処の効率化 より高度なセキュリティ対策の実施 	
データ	量	7,800億トークン	非公開	非公開	非公開	14.5億分子 (SMILES文字列)	
	内容	<ul style="list-style-type: none"> Wikipedia等のオープンソース 医学文献やデータ 専門医による多様な医療データ 	<ul style="list-style-type: none"> 臨床音声 医療データ 医療画像 デジタル病理 	<ul style="list-style-type: none"> Amazon社内とオープンになっている数十億行のコード 	<ul style="list-style-type: none"> Meta社内データ ライセンス供与しているデータ オープンソース 	<ul style="list-style-type: none"> アストロゼネカが有するデータ 	<ul style="list-style-type: none"> Cybersecurity and Infrastructure Security Agency NIST脆弱性データベース 自社データ
	収集方法	臨床医の回答に近い回答を生成するようにしてモデルを適応させた	モデルを訓練したデータの蓄積はされていない	ユーザーの使用履歴を保存して使用することによって精度向上を行う	—	アストロゼネカのMolBART変換モデルをベースにZINC化合物データベースで学習させた	—
モデルサイズ (パラメータ数)	120億/840億/5,620億	非公開	非公開	非公開	非公開	非公開	

co:here、ANTHROPIC等は産業における汎用タスク対応・AIの信頼性向上等目指し、小規模・汎用基盤モデルを開発

現時点での各基盤モデルのスペック (小規模・汎用)

	co:here	ANTHROPIC	BLOOM	Light On	
モデル名	Command	Claude	Bloom	Paradigm/Alfred-40B-0723	
目標	産業における汎用タスク (言語処理) に対応したAIの開発	AIの信頼性・解釈可能性・操作性を重視したAIの開発	大規模言語モデル (LLM) の公開研究を可能にするため	企業向けの生成AIプラットフォームであり、企業のワークフローの強化する	
ユースケース	<ul style="list-style-type: none"> ビジネス向けのNLPを活用したあらゆるサービス 文書の詳細分析 	<ul style="list-style-type: none"> カスタマーサービス 法務 コーチング バックオフィス等 	<ul style="list-style-type: none"> テキストの生成 <ul style="list-style-type: none"> 情報抽出、質疑応答、要約等を含む 	<ul style="list-style-type: none"> カスタマーサービスの生産性向上 文書/ビデオのサマリー SEOコンテンツの作成 	
データ	量	~200GB filtered, ~3TB unfiltered	確認中	1兆トークン	
	内容	<ul style="list-style-type: none"> 自社でフィルタリングを実施したデータセット 	以下を独自に組み合わせて使用 <ul style="list-style-type: none"> インターネットの公開データ 第三者企業からのライセンス供与されたデータ ユーザーからのデータ等 	<ul style="list-style-type: none"> ROOTS <ul style="list-style-type: none"> BLOOMのトレーニング用に開発された1.6TBの多言語テキストコーパス 	<ul style="list-style-type: none"> コモン・クロール
	収集方法	業界のガイドラインや規制を参照し、データの収集と処理を行ってデータセットを作成	インターネット上の公開データや使用許諾を得たうえでの非公開データを活用	詳細は不明	コモンクロールをベースに広告コンテンツや不適切なコンテンツや重複するものを削除して使用
モデルサイズ (パラメータ数)	524億パラメータ	8.1億/220億/520億/640億パラメータ	1,760億パラメータ	400億パラメータ	

日本企業はそれぞれ汎用・領域特化モデルを開発中

現時点での各基盤モデルのスペック (国内)

	松尾研究室	ELYZA	rinna	Cyber Agent	ABEJA	LINE	TURING	Softbank	
モデル名	Weblab-10B	ELYZAJapanese-Llama-2-7b	rinna/bilingual-gpt-neox-4b	OpenCALM-7B	ABEJA LLM Series	japanese-large-lm	Heron GIT Llama 2 70B Preliminary	開発中のため 情報なし	
目標	言語間の知識転移を行って、日本語の精度を高めること	他の言語で学習されたLLMの能力を日本語で引き継ぎ、日本語LLMの研究開発を加速させる	日本語LFMの選択肢を提供すること	国内における事前言語処理技術の発展への貢献	自社製品のABEJA Platformへの搭載によって機能強化を行う	日本語AIの可能性拡大	人間と同等以上にこの世界を理解した自動運転AIを実現するため		
ユースケース	開発側からのユースケース提示無し	<ul style="list-style-type: none"> テキストの入出力を行う情報処理 	<ul style="list-style-type: none"> マルチモーダル情報処理 画像対話 	<ul style="list-style-type: none"> 広告制作サービス 医療 チャットボット 	<ul style="list-style-type: none"> セキュリティ強化のうえでのデータ生成/収集/加工/モデリング 	詳細記載なし	<ul style="list-style-type: none"> 自動運転システム (予定) あらゆる不確実性への対処 		
データ	量	6,000億トークン	180億トークン	5,240億	非公開	非公開	約650GB		非公開
	内容	<ul style="list-style-type: none"> The Pile Japanese-mc4 Alpaca (英語) Alpaca (日本) Flan 等 	<ul style="list-style-type: none"> OSCAR Wikipedia 等 ※ MetaのLlama 2をベースに学習 	<ul style="list-style-type: none"> Japanese CC-100 Japanese C4 The Pile Redpajama Wikipedia 	<ul style="list-style-type: none"> Wikipedia Common Crawl 等 社内データ 	<ul style="list-style-type: none"> Wikipedia CC100 OSCAR mC4 等 	<ul style="list-style-type: none"> C4 CC-100 Oscar 等 社内データ 		<ul style="list-style-type: none"> Heron ※Llama 2-chat、ELYZA-Llama 2、等をベースにHelonを学習
	収集方法	テキストコーパスから収集	オープンデータを活用	オープンデータの活用	オープンデータの活用	オープンデータの活用	オープンデータの活用と社内システムのウェブテキストでクロール		独自に日本語に翻訳した大規模な日本語の画像/テキスト情報のデータセットを作成
モデルサイズ (パラメータ数)	100億パラメータ	70億パラメータ	40億パラメータ	~68億パラメータ	~130億パラメータ • GPT-3がベース	17億/ 36億パラメータ	700億パラメータ		

汎用

領域特化

② 経済インパクト試算

効率化・コスト削減および新たな価値創出・売上拡大の両側面でのAIの利活用に伴う経済インパクトを試算

A 効率化・コスト削減

各産業へのAI導入に伴う
効率化・コスト削減インパクトを
労務・人件費ベースのトップダウンで試算

~15兆円

- 産業×機能毎に労務・人件費を試算し主要領域を特定
- 主要領域を特定後、インタビュー/アンケートを通してインパクトを試算
 - 主要領域におけるユースケースを詳細化

AIの利活用に伴う
経済インパクト

B 新たな価値創出 ・売上拡大

有望なユースケースへのAI導入に伴う
新たな価値創出・売上拡大インパクトを
ユースケースベースのボトムアップで試算

- エキスパートへのインタビュー・デスクトップ調査を通じて有望なユースケースを特定
 - 主要領域に絞り調査

サーベイを通じて、生成AI導入による効率化・コスト削減インパクトを算出するとともに、本事業の社会実証に向けた障壁・乗り越え方も調査

目的

1. 生成AI導入による効率化・コスト削減による経済インパクトが大きい領域（産業×機能）を特定する
2. 本事業の社会実装に向けた企業の障壁及びその乗り越え方を特定する
 - 生成AIおよび本事業の参加

調査設計

調査項目

1. 生成AI導入による経済インパクト
 - 産業×機能毎の生成AIの導入による業務削減割合
 - 産業×機能毎の生成AIの導入状況・導入意向
2. 本事業の社会実装に向けた企業の障壁・乗り越え方
 - 生成AIの導入に対する障壁・乗り越え方
 - 本事業に対する障壁・乗り越え方
 - 参加意向
 - 参加障壁
 - データ保有状況
 - データ共有意向
 - データ共有における障壁・乗り越え方

調査対象

(優先業種)

- 労務費・人件費の大きい上位5業種に加え、有望領域と想定される金融・保険を対象
 - 製造業（自動車・鉄道・船舶・航空/製薬・素材・化学・その他製造業）
 - 卸・小売
 - 医療福祉（医療・介護）
 - 建設
 - サービス
 - 金融・保険
- 他業種については、優先業種で得られたパラメーターを展開

生成AI等のシステム導入の意思決定に関わる約2,400名へ調査を実施

サーベイの概要

調査条件

期間	2023年 9月15日～18日
産業	産業分類大分類を基本とし、必要となるデータが大きく異なる産業は中分類レベルで再分類
機能	企業活動基本調査での事業組織分類を基本とし、必要となるデータが大きく異なる機能は再分類
会社規模	大企業・中小企業の双方を含む
回答者	生成AI等のシステム導入の意思決定に関わる ¹ 、または管理職の方
調査母数	2,391名

産業×機能の回答人数²

	R&D	SW/ アプリケーション 開発	調達/ サプライチェーン	製造 加工	セールス/ マーケ	オペレーション サービス/CS	バック オフィス	業務横断	合計
製造業	70	42	81	158	35	9	46	32	473
卸・小売	5	23	32	7	88	25	73	20	273
医療・福祉	9	-	12	-	50	27	113	42	253
建設	7	10	20	20	67	17	86	45	272
サービス	7	19	15	6	56	44	105	26	278
金融・保険	4	39	6	-	58	27	98	35	267
その他	16	157	36	18	104	47	135	62	575
合計	118	290	202	209	458	196	656	262	2,391

1. 「最終決済をする立場」「選択肢を絞り込む立場」「指示を受けて情報収集する立場」のいずれか

2. 回答者が10人未満の産業×機能領域は、経済インパクト試算の自動化割合においてその機能の平均値、生成AI導入意向においてその産業の平均値を採用

Source: BCG調査

生成AIによるコスト削減効果は約18兆円と試算され、産業×機能毎での有望領域は以下の通り

サーベイ結果を踏まえたコスト削減効果

コスト削減効果 (全体)	<p>業務×機能の労務・人件費に、生成AIの利活用における業務削減割合、導入意向のサーベイ結果を乗算し、生成AIによる効率化・コスト削減効果は約18兆円</p> <ul style="list-style-type: none"> 人件費労務費総額：220兆円 生成AIの利活用による業務削減割合：約21% 生成AIの導入意向：約44%
業務削減割合 (産業×機能)	<p>産業毎では大きな差は見られないが、機能毎での差が大きい</p> <ul style="list-style-type: none"> 産業毎：金融・保険が26%とやや高く、その他はおおよそ20%程度 機能毎：業務横断（資料作成等）が30%、バックオフィスが30%で高く、製造加工は12%、調達/サプライチェーンは15%と低い
導入意向 (産業×機能)	<p>産業・機能毎に大きな差が見られる</p> <ul style="list-style-type: none"> 産業毎：自動車・製薬業が約60%、金融保険が55%と高く、医療が38%、卸小売が31%、建設業が28%、介護が17%と低い 機能毎：R&Dが56%、SW開発が54%と高く、セールスマーケティングが32%と低い
有望領域	<p>コスト削減効果が特に大きい（5,000億円以上）有望領域は以下の通り</p> <ul style="list-style-type: none"> (産業×機能) <ul style="list-style-type: none"> ①自動車・鉄道・船舶・航空×製造加工、②製薬・素材・化学×製造加工、③その他製造業×製造加工 ④卸小売×調達サプライチェーン、⑤サービス×オペレーション/CS ⑥ソフトウェア×ソフトウェア開発/アプリケーション開発、⑦金融保険×オペレーション/CS (産業横断) <ul style="list-style-type: none"> ⑧バックオフィス、⑨業務横断(資料作成等) <p>次いでコスト削減効果が大きい（3,500億円以上）の有望領域は以下の通り</p> <ul style="list-style-type: none"> (産業×機能) <ul style="list-style-type: none"> ⑩医療×オペレーション/CS、⑪運輸・輸送×オペレーション/CS

労務費・人件費の総額は約220兆円で、産業×機能毎の分布は異なる

労務・人件費総額

：5兆円以上

：1兆円以上~5兆円未満

効率化・コスト削減の経済インパクト

産業	機能								合計値
	R&D	SW開発/ アプリケーション開発	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務横断 (資料作成等)	
自動車・鉄道・ 船舶・航空	1,324	6	285	7,016	208	160	856	347	10,202
製薬・化学・素材	1,398	27	1,880	10,956	1,004	221	1,406	223	17,115
その他製造業	1,935	208	2,369	13,751	1,037	477	1,510	472	21,759
卸売・小売	935	143	24,999	1,386	2,948	999	2,030	2,043	35,482
医療	79	-	1,357	-	1,280	5,919	393	5,961	14,990
介護	69	-	1,192	-	1,124	5,197	345	5,234	13,160
建設	555	609	1,148	1,528	3,335	818	1,297	7,164	16,453
サービス	140	249	2,419	864	2,281	10,548	700	10,624	27,825
運輸・郵送	71	126	1,229	439	1,159	5,358	356	5,397	14,136
ソフトウェア	316	6,655	350	57	453	276	878	853	9,838
メディア・エンタメ	39	445	83	126	418	79	136	840	2,167
金融・保険	58	103	1,001	-	945	4,367	290	4,398	11,163
プロフェッショナルサービス	565	931	246	102	1,932	311	530	5,068	9,683
教育	48	11	215	-	497	414	284	6,444	7,913
不動産	104	79	1,016	99	751	216	566	2,951	5,783
エネルギー	99	1	108	1,322	26	33	192	56	1,838
農林水産・ 鉱石砂利	32	-	57	695	37	39	163	92	1,115
合計値	7,768	9,594	39,954	38,341	19,435	35,431	11,931	58,167	222,152

産業毎では大きな差は見られないが、機能毎での差が大きく、バックオフィスおよび業務横断（資料作成等）が30%と高い

生成AIの利活用による業務削減割合

AIの利活用における業務削減割合

産業	機能								平均値
	R&D	SW開発/ アプリケーション開発	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務横断 (資料作成等)	
自動車・鉄道・ 船舶・航空	30%	26%	22%	24%	22% ¹	20% ¹	30% ¹	30% ¹	24%
製薬・化学・素材	24%	18% ¹	18%	18%	22% ¹	20% ¹	21%	30% ¹	21%
その他製造業	20%	20%	20%	24%	18%	20% ¹	19%	27%	21%
卸売・小売	14%	15%	15%	12%	27%	18%	23%	26%	18%
医療	18% ¹	-	15% ¹	-	21%	17%	30%	31%	18%
介護	18% ¹	-	15% ¹	-	24%	15%	31%	34%	18%
建設	18% ¹	12%	14%	15%	21%	15%	27%	33%	19%
サービス	18% ¹	20%	15%	12% ¹	21%	20%	27%	31%	20%
運輸・郵送									
ソフトウェア	17%	18%	12%	9%	24%	22%	35%	30%	21%
メディア・エンタメ									
金融・保険	18% ¹	18%	15% ¹	-	24%	22%	35%	30%	26%
プロフェッショナルサービス									
教育									
不動産	17%	18%	12%	9%	24%	22%	35%	30%	21%
エネルギー 農林水産・ 鉱石砂利									
平均値	18%	18%	15%	12%	22%	20%	30%	30%	21%

：25%以上
：20%未満

Source: BCG調査

1. 回答者が10人未満の産業×機能領域は、経済インパクト試算においてその機能の平均値を採用
質問文：生成AIを導入することで、各部署の現在の業務量を100%としたときに、どの程度の業務量が、効率化・自動化されると思いますか

産業・機能毎に大きな差が見られ、産業毎で自動車・製薬が約60%、金融保険が55%と高く、機能毎ではR&Dが56%、SW開発が54%と高い

生成AIの導入意向

生成AIの導入意向

産業	機能								平均値
	R&D	SW開発/ アプリケーション開発	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務横断 (資料作成等)	
自動車・鉄道・ 船舶・航空	80%	75%	61%	53%	61% ¹	61% ¹	61% ¹	61% ¹	61%
製薬・化学・素材	45%	62%	59%	52%	62% ¹	62% ¹	44%	62% ¹	62%
その他製造業	64%	46%	50%	39%	43%	48% ¹	39%	50%	48%
卸売・小売	20%	26%	28%	31% ¹	26%	44%	22%	20%	31%
医療	38% ¹	-	38% ¹	-	45%	20%	24%	52%	38%
介護	35% ¹	-	35% ¹	-	21%	0%	18%	65%	35%
建設	17% ¹	-	10%	-	6%	18%	9%	4%	17%
サービス	28% ¹	42%	53%	28% ¹	32%	30%	24%	19%	28%
運輸・郵送									
ソフトウェア	69%	54%	42%	50%	30%	40%	51%	35%	46%
メディア・エンタメ									
金融・保険	55% ¹	67%	55% ¹	-	52%	63%	47%	49%	55%
プロフェッショナルサービス									
教育									
不動産	69%	54%	42%	50%	30%	40%	51%	35%	46%
エネルギー 農林水産・ 鉱石砂利									
平均値	56%	54%	42%	42%	32%	44%	41%	40%	44%

：50%以上

：40%未満

Source: BCG調査

1. 回答者が10人未満の産業×機能領域は、経済インパクト試算においてその業界の平均値を採用
質問文：あなたのお勤め先(経営先)において、生成AIの導入状況として最も近いものをお選びください。

労務・人件費、業務削減割合、導入意向を乗算した産業×機能での有望領域は以下 有望領域

効率化・コスト削減の経済インパクト

■ : 5,000億円以上
■ : 3,500億円以上

有望領域

単位:億円	機能								合計値
	R&D	SW開発/ アプリケーション開発	調達/サプライ チェーン	製造・ 加工	セールス/ マーケティング	オペレーション サービス/CS	バック オフィス	業務横断 (資料作成等)	
自動車・鉄道・ 船舶・航空	3,136	12	375	1 9,081	279	195	8 1,567	9 643	15,289
製薬・化学・素材	1,499	30	2,021	2 9,899	1,370	274	1,302	415	16,810
その他製造業	2,533	191	2,330	3 12,756	789	458	1,122	647	20,825
卸売・小売	264	56	4 10,884	515	2,118	776	1,035	1,081	16,730
医療	54	-	774	-	1,208	10 3,711	273	9,551	15,570
介護	44	-	626	-	581	3,029	190	11,623	16,091
建設	170	-	162	-	412	211	327	1,051	2,333
サービス	71	206	1,938	290	1,532	5 6,242	455	6,372	17,106
運輸・郵送	82	124	619	208	837	11 4,740	645	5,720	12,977
ソフトウェア	365	6 6,550	177	27	327	244	1,593	904	10,187
メディア・エンタメ	45	438	42	60	302	70	247	891	2,095
金融・保険	58	125	826	-	1,184	7 6,016	483	6,381	15,073
プロフェッショナルサービス	651	916	124	48	1,396	275	960	5,371	9,742
教育	55	11	108	-	359	366	515	6,830	8,245
不動産	121	78	512	47	542	191	1,027	3,128	5,646
エネルギー	115	1	54	627	19	30	348	60	1,253
農林水産・ 鉱石砂利	37	-	29	330	26	35	295	98	849
合計値	9,299	8,738	21,601	33,891	13,282	26,862	12,383	60,766	186,821

- 1 自動車・鉄道・船舶・航空×
製造加工
- 2 化学・製薬・素材×
製造加工
- 3 その他製造業×
製造加工
- 4 卸売・小売×
調達・サプライチェーン
- 5 サービス×
オペレーションサービス/CS
- 6 ソフトウェア×
SW開発/アプリケーション開発
- 7 金融・保険×
オペレーションサービス/CS
- 8 バックオフィス
- 9 業務横断（資料作成等）
- 10 医療×
オペレーションサービス/CS
- 11 運輸・輸送×
オペレーションサービス/CS

産業

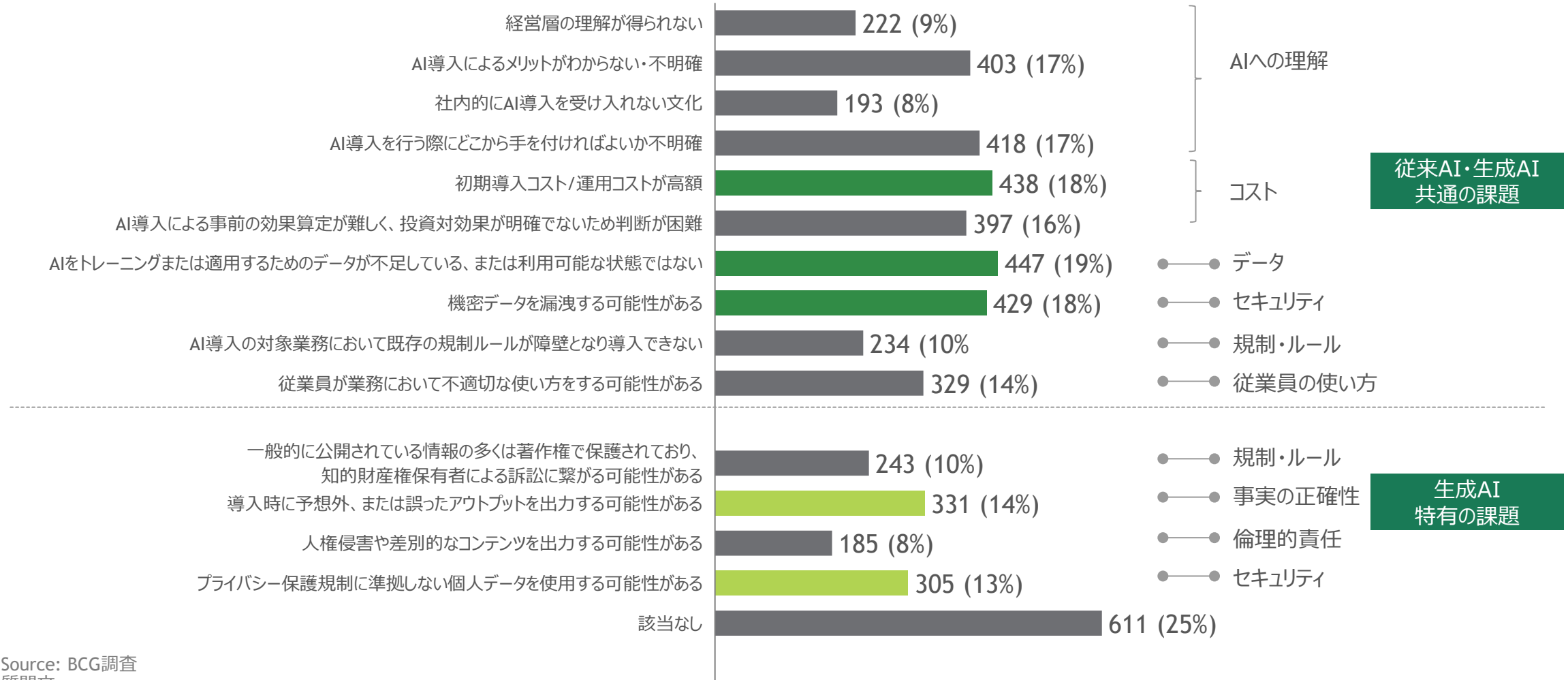
本事業におけるデータ共有可否・障壁は業種により異なり業種毎での支援が必要

社会実装に向けた企業の障壁・乗り越え方

生成AI導入の障壁／期待する支援	障壁	<p>生成AI導入を進める上では、「従来・生成AI共通の障壁」に加え、「生成AI特有の障壁」も存在</p> <ul style="list-style-type: none"> 事実の正確性、倫理的責任の障壁が比較的大きい 製造・金融業や大企業では機密データ等のセキュリティ対応、医療・介護や中小企業はコストの障壁が大きい
	期待する支援策	<p>障壁を乗り越えるためには、人、マニュアル・手順書、金銭的支援が求められる</p> <ul style="list-style-type: none"> 業務・技術の双方を理解した人物が23%、マニュアル・手順書、コストに対する金銭的な支援が22%と高い 大企業では法的懸念に対する支援策を求める声が多く、中小企業では生成AIへの理解、金銭的支援策を求める声が多い
本事業への参加意向／参加障壁	参加意向	<p>42%が参加したいと回答しており、一定数の参加は得られるか</p> <ul style="list-style-type: none"> 参加したい：10%、どちらかといえば参加したい：32%
	参加障壁	<p>参加障壁は、社内調整やデータ共有、コンプライアンス上のリスクが大きい</p> <ul style="list-style-type: none"> 社内調整が26%と障壁として最も大きく、続いてデータ共有、コンプライアンス上のリスクが23%と大きい
データ共有の意向／障壁／期待する支援	データ保有／共有意向	<p>データ保有/共有意向は業種により異なる</p> <ul style="list-style-type: none"> 製造業・金融業はデータ保有率は高いが、データ共有難易度が低い 卸小売・医療等はデータ保有率は低いが、データ共有難易度が高い
	共有障壁	<p>共有障壁は、情報漏洩の懸念、コンプライアンス上の懸念が大きく、事業参加におけるセキュリティやコンプライアンス上の安全性を示すことが必要</p> <ul style="list-style-type: none"> 医療介護・金融業は情報漏洩に加えてコンプライアンス上の懸念が大きく、セキュリティ・個人情報保護上の安全性を示すことが必要
	期待する支援策	<p>回答者全体では「データ提供することで利益が自社へ還元される」ことを望む声が多く全体の19%が選択しているものの、業種により求める支援が異なり、業種ごとでの支援が求められる</p> <ul style="list-style-type: none"> 自動車産業では他企業の権限制限を求める声が多い 製薬・医療・金融保険ではデータ共有の障壁となっている法律・規制・ルールの緩和を求める声が多い

生成AI導入を進める上では、「従来・生成AI共通の障壁」に加え、「生成AI特有の障壁」も存在

生成AI導入における障壁 (回答者全体・最大5つの選択肢を回答)



Source: BCG調査
質問文:

(生成AI導入済) 過去にあなたのお勤め先(経営先)で生成AIを導入された際の障壁として大きかったものを5つまでお選びください。
(生成AI検討中) あなたのお勤め先(経営先)への生成AI導入における障壁として大きいものを5つまでお選びください
(生成AI未検討) 仮にあなたのお勤め先(経営先)へ生成AIを導入する場合、障壁として大きいものを5つまでお選びください

製造・金融業では機密データの漏洩、医療・介護や建設ではコストの障壁が大きい等、業種間で生成AI導入における障壁は異なる

生成AI導入における障壁 (産業別・最大5つの選択肢を回答)

■ : 20%以上 □ : 各産業において割合が最大

	従来AI・生成AI共通の障壁							生成AI特有の障壁							
	AIへの理解		コスト		データ	セキュリティ	規制・ルール	従業員の使い方	規制・ルール	事実の正確性	倫理的責任	セキュリティ			
	経営層の理解が得られない	AI導入によるメリットがわからない・不明確	社内的にAI導入を受け入れない文化	AI導入を行う際にどこから手を付ければよいか不明確	初期導入コスト/運用コストが高額	AI導入による事前の効果算定が難しく、投資対効果が明確でないため判断が困難	AIをトレーニングまたは適用するためのデータが不足している、または利用可能な状態ではない	機密データを漏洩する可能性あり	AI導入の対象業務において既存の規制ルールが障壁となり導入できない	従業員が業務において不適切な使い方をする可能性あり	一般的に公開されている情報の多くは著作権で保護されており、知的財産権保有者による訴訟に繋がる可能性あり	導入時に予想外、または誤ったアウトプットを出力する可能性あり	人権侵害や差別的なコンテンツを出力する可能性あり	プライバシー保護規制に準拠しない個人データを使用する可能性あり	
製造業	10%	17%	10%	19%	18%	15%	20%	26%	13%	15%	14%	17%	10%	14%	
卸売・小売	8%	18%	7%	19%	17%	14%	17%	13%	9%	11%	8%	10%	5%	10%	
医療/介護	14%	20%	9%	21%	24%	14%	17%	12%	11%	14%	8%	13%	6%	13%	
建設	13%	22%	10%	17%	22%	20%	18%	9%	7%	15%	7%	14%	8%	10%	
サービス	9%	15%	9%	16%	18%	17%	22%	16%	6%	12%	10%	12%	8%	13%	
金融・保険	8%	14%	4%	12%	17%	17%	17%	21%	12%	13%	10%	15%	9%	18%	
その他	6%	14%	7%	17%	15%	17%	18%	20%	9%	14%	10%	13%	7%	12%	
平均	10%	17%	8%	17%	19%	16%	18%	17%	9%	13%	10%	13%	8%	13%	

Source: BCG調査

質問文:

(生成AI導入済) 過去にあなたのお勤め先 (経営先) で生成AIを導入された際の障壁として大きかったものを5つまでお選びください。

(生成AI検討中) あなたのお勤め先 (経営先) への生成AI導入における障壁として大きいものを5つまでお選びください

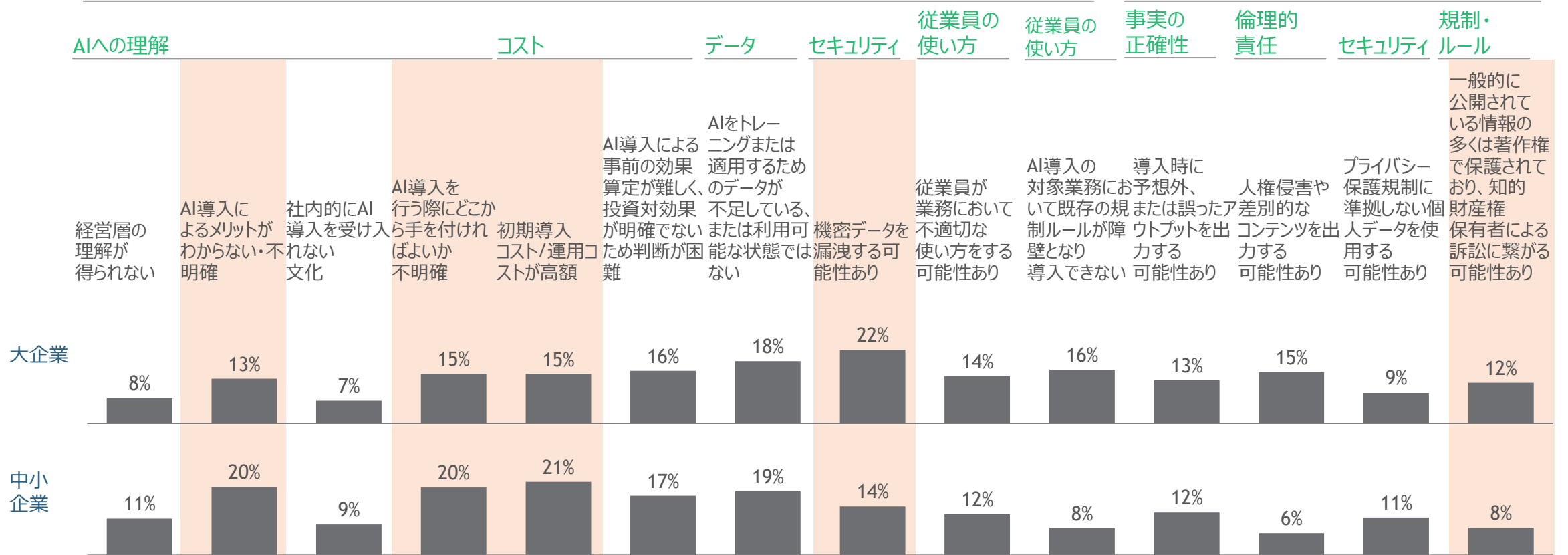
(生成AI未検討) 仮にあなたのお勤め先 (経営先) へ生成AIを導入する場合、障壁として大きいものを5つまでお選びください

大企業は機密データの漏洩、規制・ルールの懸念が大きい一方、中小企業はAIへの理解、コストの障壁が大きい

生成AI導入における障壁 (従業員数別・最大5つの選択肢を回答) ※300人未満/以上

従来AI・生成AI共通の障壁

生成AI特有の障壁



Source: BCG調査

質問文:

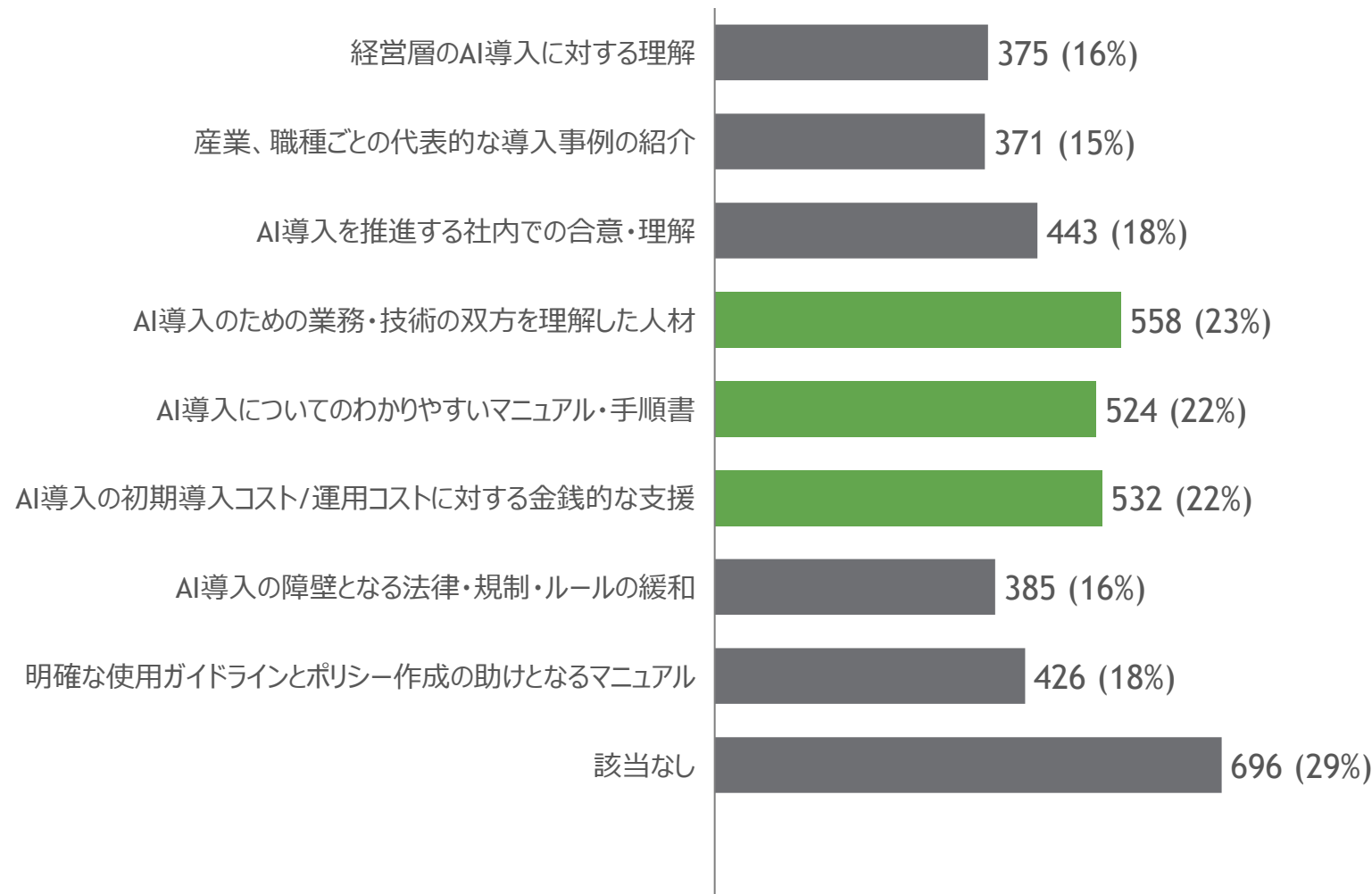
(生成AI導入済) 過去にあなたのお勤め先 (経営先) で生成AIを導入された際の障壁として大きかったものを5つまでお選びください。

(生成AI検討中) あなたのお勤め先 (経営先) への生成AI導入における障壁として大きいものを5つまでお選びください

(生成AI未検討) 仮にあなたのお勤め先 (経営先) へ生成AIを導入する場合、障壁として大きいものを5つまでお選びください

障壁を乗り越えるためには、人、マニュアル・手順書、金銭的支援が求められる

生成AI導入において求める支援策 (回答者全体・最大3つの選択肢を回答)



Source: BCG調査

質問文: あなたのお勤め先 (経営先) であなたが生成AI導入を推進する場合、何があればより導入を進められると思いますか。3つまでお選びください

セキュリティを懸念する声が多かった製造業はAI導入のための業務・技術の双方を理解した人材、コストを懸念する声が多かった医療介護・建設は金銭的支援を求める声が多い

生成AI導入において求める支援策 (産業別・最大3つの選択肢を回答)

■ : 20%以上 □ : 各産業において割合が最大

	AIへの理解			コスト		規制・ルール	事実の正確性 倫理的責任 セキュリティ	
	経営層のAI導入に対する理解	産業、職種ごとの代表的な導入事例の紹介	AI導入を推進する社内での合意・理解	AI導入のための業務・技術の双方を理解した人材	AI導入についてのわかりやすいマニュアル・手順書	初期導入コスト/運用コストに対する金銭的な支援	AI導入の障壁となる法律・規制・ルールの緩和	明確な使用ガイドラインとポリシー作成の助けとなるマニュアル
製造業	18%	18%	22%	27%	24%	25%	16%	22%
卸売・小売	18%	11%	17%	25%	20%	21%	18%	15%
医療/介護	18%	16%	18%	18%	22%	23%	13%	15%
建設	14%	19%	16%	19%	24%	25%	14%	17%
サービス	16%	11%	15%	27%	24%	21%	13%	14%
金融・保険	14%	16%	21%	21%	23%	21%	20%	19%
その他	13%	15%	18%	23%	18%	20%	18%	19%
平均	16%	15%	18%	23%	22%	22%	16%	17%

Source: BCG調査

質問文: あなたのお勤め先 (経営先) であなたが生成AI導入を推進する場合、何があればより導入を進められると思いますか。3つまでお選びください

大企業では規制・ルールに対する支援策への期待が多く、中小企業では生成AIへの理解、コストに対する支援策を求める声が多い

生成AI導入において求める支援策(従業員数別・最大3つの選択肢を回答)※300人未満/以上

AIへの理解

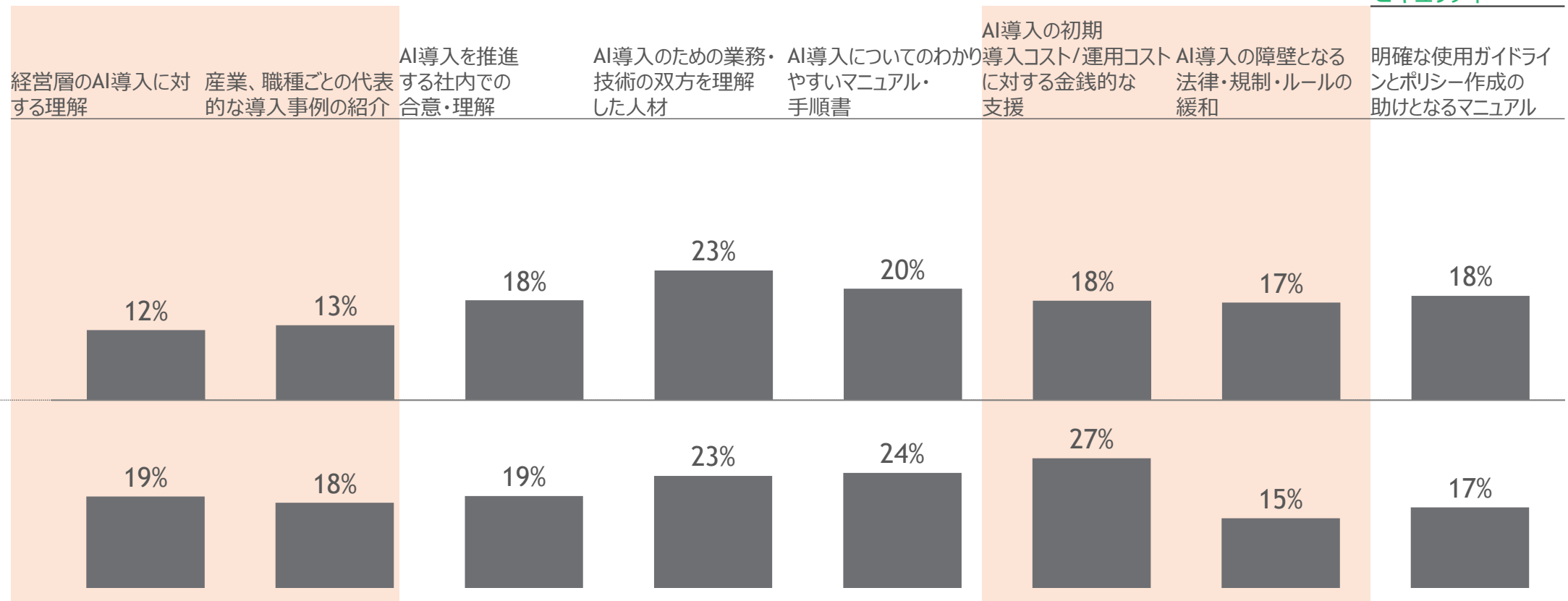
コスト

規制・ルール

事実の正確性

倫理的責任

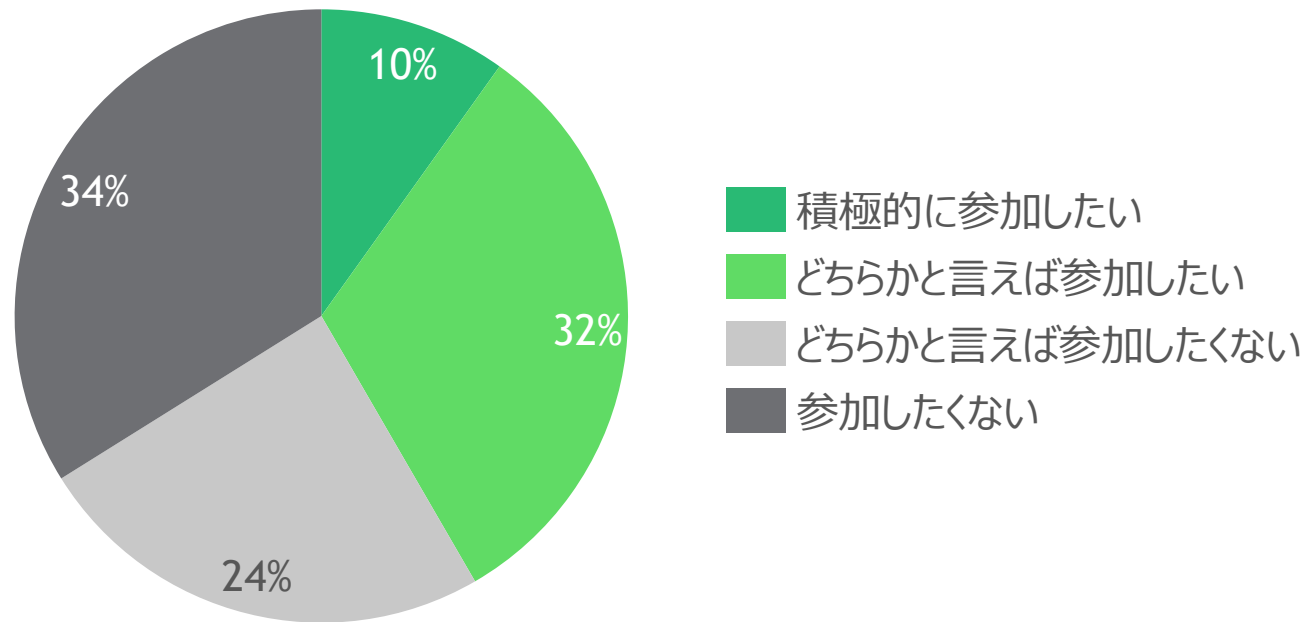
セキュリティ



Source: BCG調査

質問文: あなたのお勤め先(経営先)であなたが生成AI導入を推進する場合、何があればより導入を進められると思いますか。3つまでお選びください

42%が本事業に参加したいと回答しており、一定数の参加は得られるか 本事業への参加意向（回答者全体）



Source: BCG調査

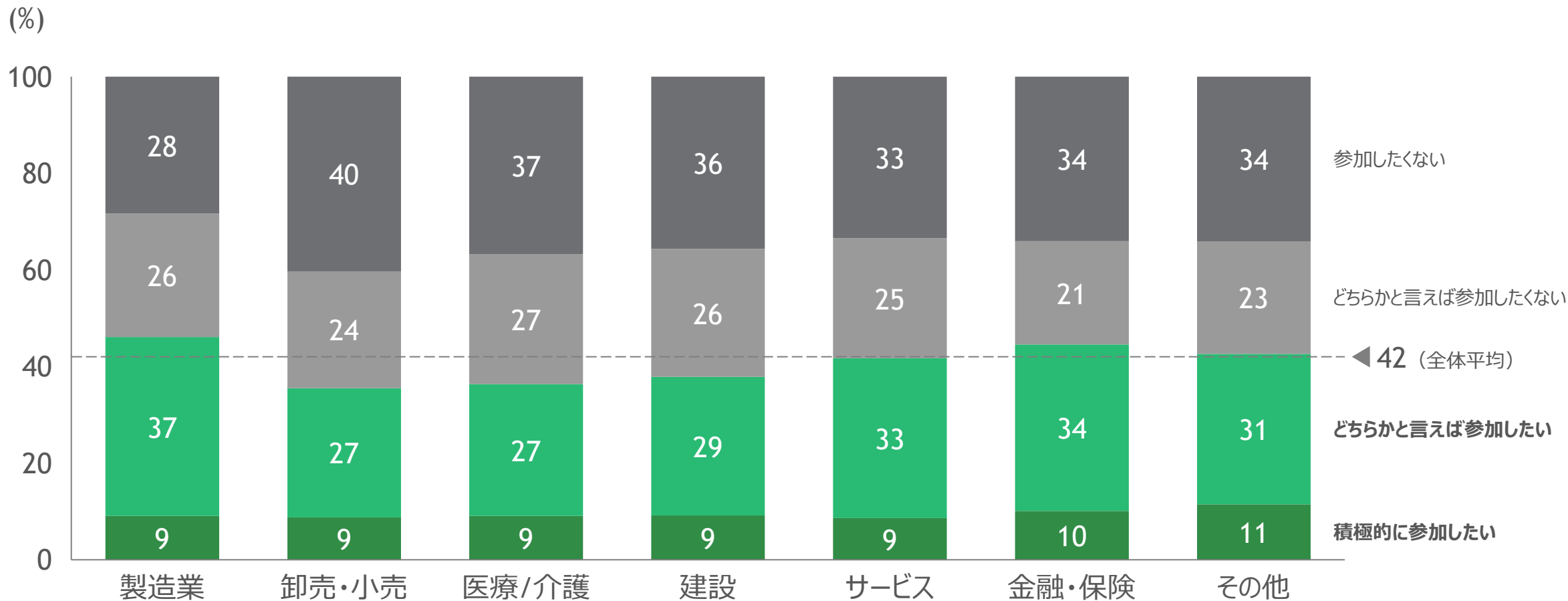
質問文：

公的機関が、企業における生成AIの導入を推進するための支援事業を開始したとします。
この事業に参加することで、参加企業は以下のようなメリットを受けられる可能性があります。

- ・自社のデータを活用した、自社の業務で活用可能な生成AIの導入
- ・生成AI開発企業とのパートナーシップや協業の機会
- ・自社データを開発企業へ提供することによる収益

参加企業は、自社のデータを生成AI開発企業等に共有することで、本事業に参加することができます。
あなたは本事業に参加したいと思いませんか。

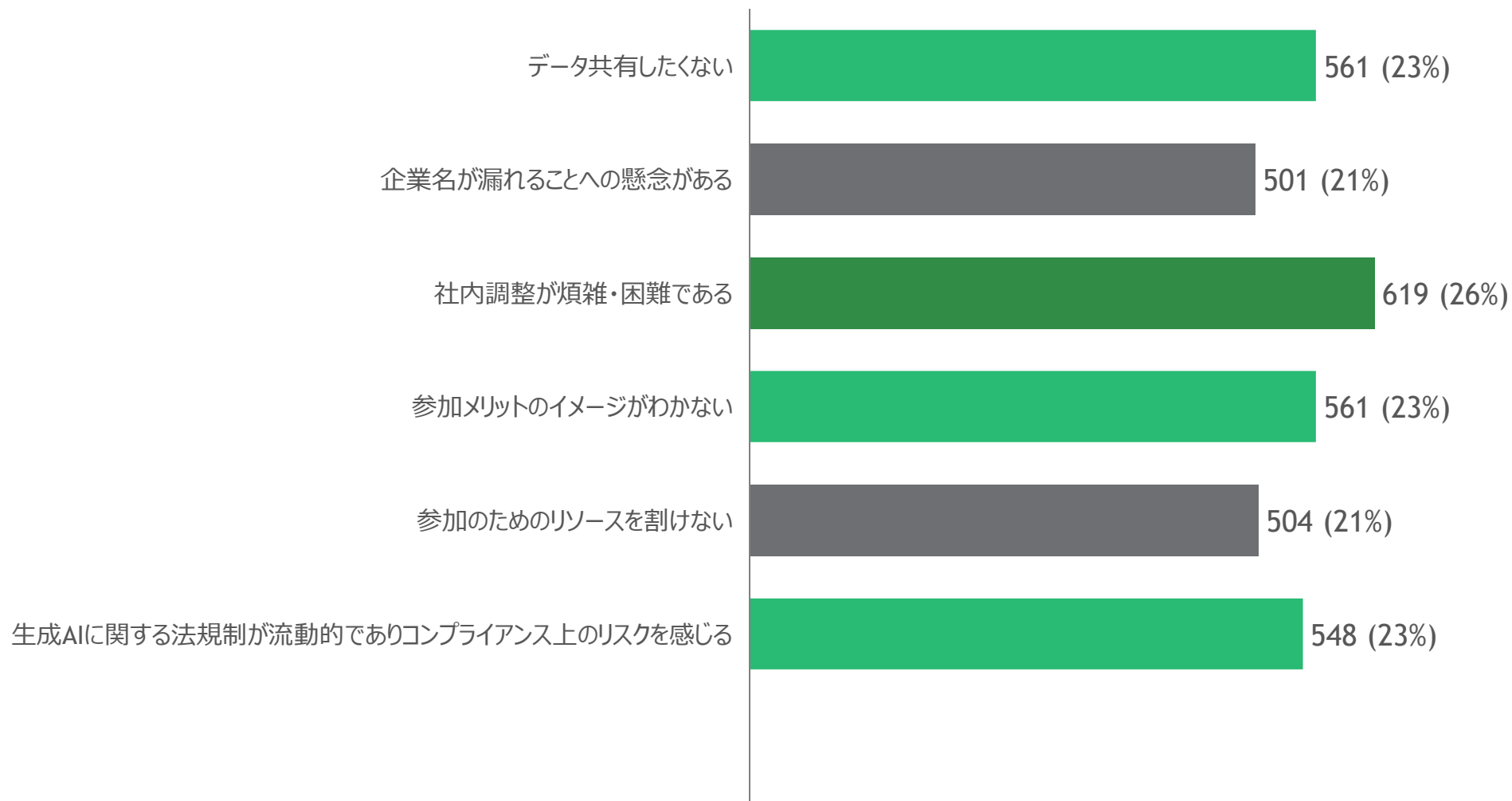
各産業において参加意向はおしなべて高い 本事業への参加意向 (産業別)



Source: BCG調査

質問文 (公的機関が企業における生成AIの導入推進をするための支援事業を開始したと仮定のうえ、事業参加のメリットを提示)
参加企業は、自社のデータを生成AI開発企業等に共有することで、本事業に参加することができます。あなたは本事業に参加したいと思いますか。

参加障壁は、社内調整やデータ共有、コンプライアンス上のリスクが大きい 本事業への参加障壁（回答者全体）



自動車・製薬業等は「データ共有したくない」割合は低いものの社内調整の煩雑さ・コンプライアンス上のリスクを懸念する声が多く、サービス業は「データ共有したくない」割合が大きい

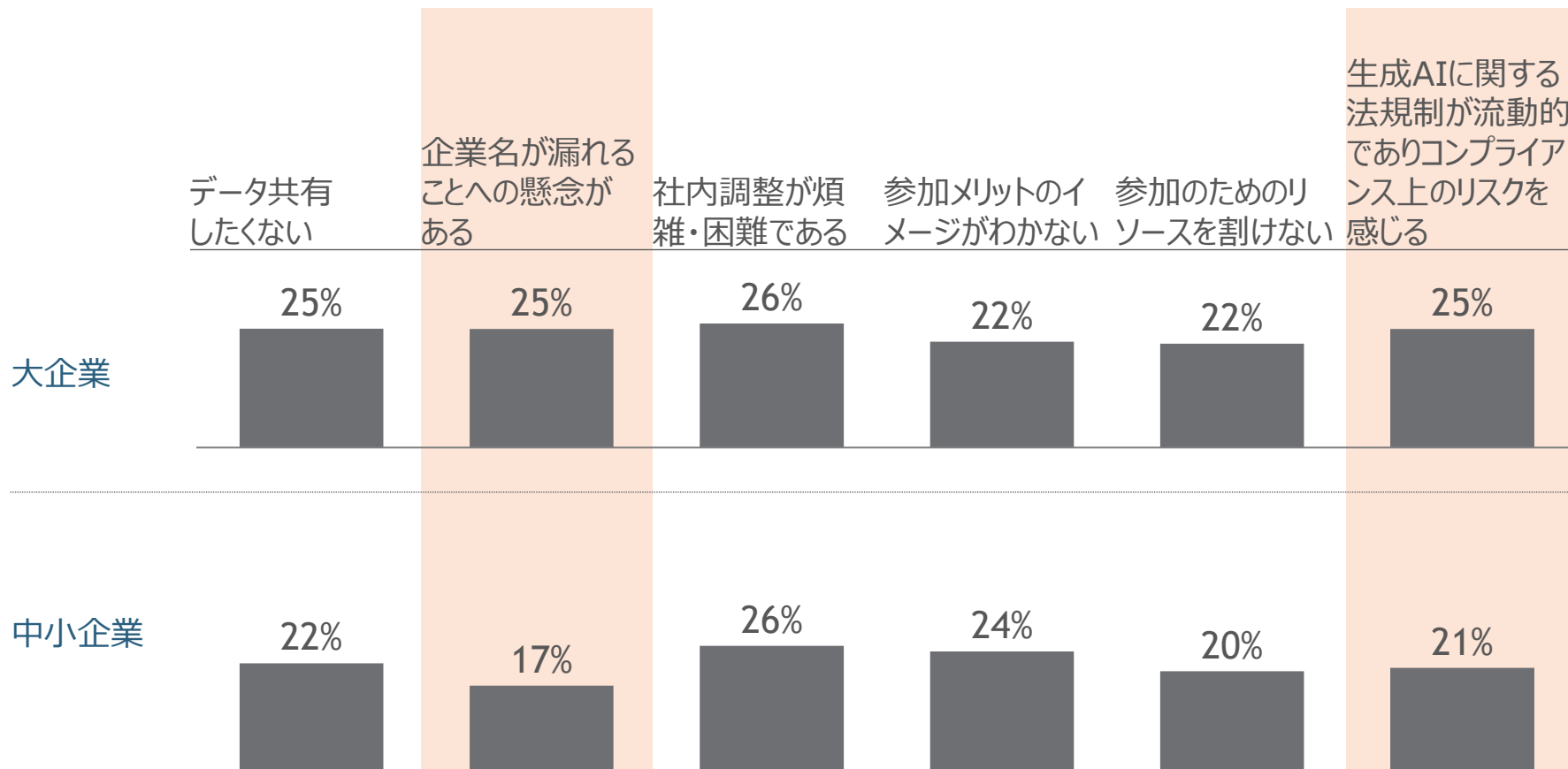
本事業への参加障壁 (産業別)

■ : 20%以上 □ : 各産業において割合が最大

	データ共有したくない	企業名が漏れることへの懸念がある	社内調整が煩雑・困難である	参加メリットのイメージがわからない	参加のためのリソースを割けない	生成AIに関する法規制が流動的でありコンプライアンス上のリスクを感じる
自動車・鉄道・船舶・航空	15%	19%	20%	12%	16%	18%
製薬・化学・素材	16%	16%	20%	12%	18%	18%
その他製造業	19%	16%	18%	16%	16%	15%
卸売・小売	14%	13%	23%	22%	14%	14%
医療	15%	12%	18%	19%	20%	16%
介護	16%	12%	19%	21%	16%	16%
建設	16%	16%	20%	18%	15%	15%
サービス	21%	12%	19%	18%	15%	16%
金融・保険	19%	17%	20%	13%	12%	19%
その他	16%	17%	16%	18%	16%	18%
平均	17%	15%	19%	17%	16%	16%

大企業の方が企業名漏洩、法規制に関するコンプライアンス上のリスクに対する懸念の声が多い

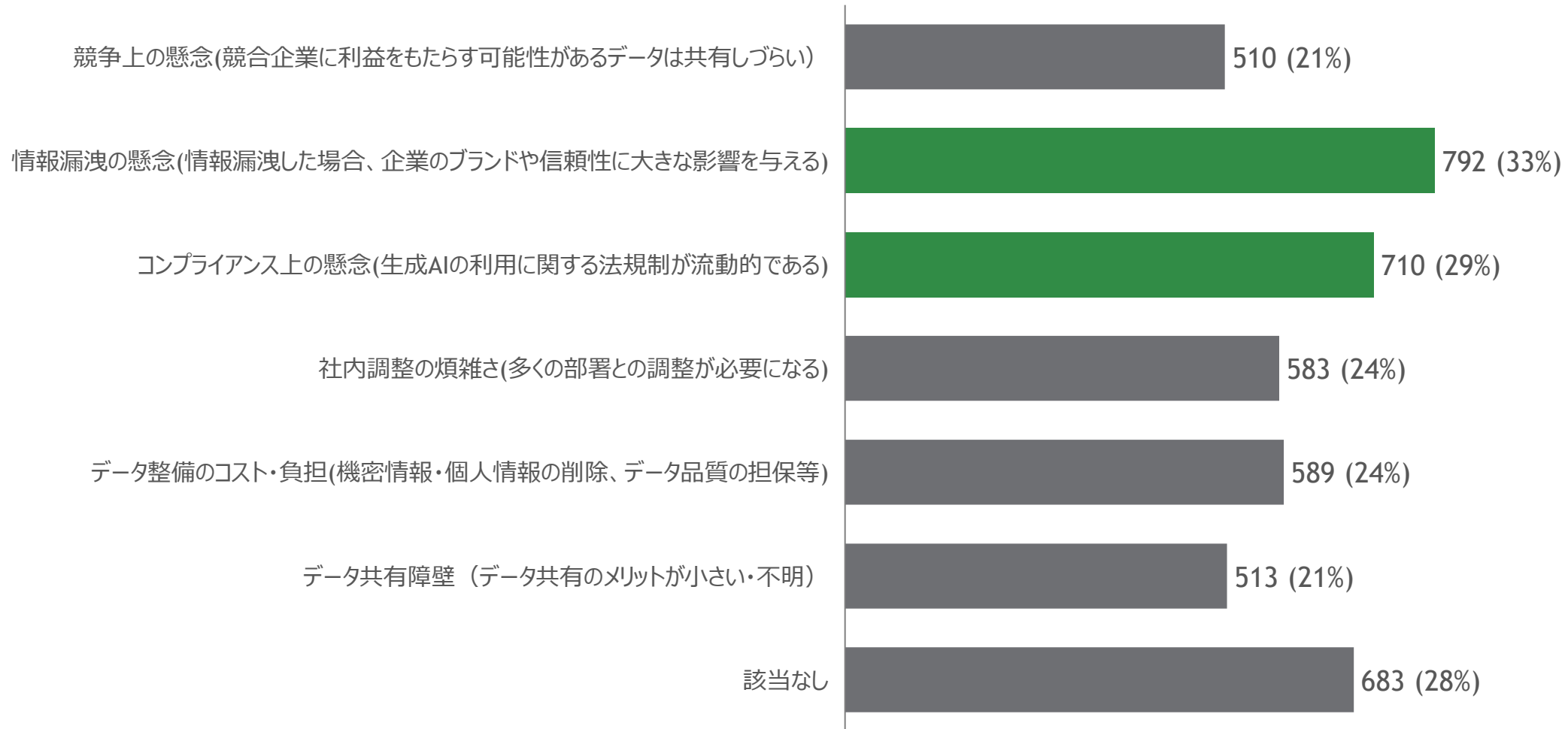
本事業への参加障壁（従業員数別） ※300人未満/以上



Source: BCG調査
質問文：本支援事業に参加するにあたって、障壁になりそうなことを全てお選びください。

障壁は、情報漏洩の懸念、コンプライアンス上の懸念が大きく、事業参加におけるセキュリティやコンプライアンス上の安全性を示すことが必要

データ共有の障壁（回答者全体）



2 個別分野別の経済インパクト a 経済インパクト試算

グローバル

勝ち筋

経済

日本

製造業・金融業はデータ保有率は高いがデータ共有不可の割合が大きく、卸小売、医療等はデータ保有率は低いが、データ共有不可の割合が小さい

データ共有の意向 (産業別)

: 全体平均より割合が大きい
 : 全体平均より割合が小さい

	デジタルデータを保有していない		共有できない			共有できる		
	保有していない	共有できない	開発会社へのみ共有できる	事業に参加できる企業間にも共有できる	パブリック公開できる	共有できない	共有できない	共有できない
自動車・鉄道・船舶・航空	16%	53%	24%	5%	3%			
製薬・化学・素材	14%	59%	24%	3%	0%			
その他製造業	18%	56%	18%	5%	4%			
卸売・小売	34%	46%	14%	4%	2%			
医療	41%	40%	13%	5%	1%			
介護	48%	35%	11%	4%	2%			
建設	37%	42%	15%	4%	1%			
サービス	32%	48%	16%	3%	1%			
金融・保険	26%	49%	19%	5%	2%			
その他	24%	52%	17%	5%	2%			
平均	29%	48%	17%	4%	2%			

2 個別分野別の経済インパクト a 経済インパクト試算

多くの産業で情報漏洩の懸念への対応が求められるとともに、介護・金融業では コンプライアンス上の懸念が特に大きい

データ共有の障壁 (産業別)

■ : 20%以上
□ : 各産業において割合が最大

	競争上の懸念 (競合企業に利益をもたらす可能性があるデータは共有しづらい)	情報漏洩の懸念 (情報漏洩した場合、企業のブランドや信頼性に大きな影響を与える)	コンプライアンス上の懸念 (生成AIの利用に関する法規制が流動的である)	社内調整の煩雑さ (多くの部署との調整が必要になる)	データ整備のコスト・負担 (機密情報・個人情報削除、データ品質の担保等)	データ共有障壁 (データ共有のメリットが小さい・不明)	平均
自動車・鉄道・船舶・航空	17%	19%	19%	17%	19%	9%	17%
製薬・化学・素材	14%	22%	16%	18%	19%	12%	17%
その他製造業	17%	21%	17%	16%	16%	13%	17%
卸売・小売	14%	18%	17%	19%	16%	17%	17%
医療	11%	22%	19%	16%	16%	16%	17%
介護	10%	18%	21%	18%	16%	18%	17%
建設	15%	19%	17%	15%	19%	16%	17%
サービス	13%	21%	18%	16%	15%	17%	17%
金融・保険	13%	25%	26%	14%	13%	10%	17%
その他	14%	24%	20%	15%	15%	13%	17%
全体平均	14%	21%	19%	16%	16%	14%	17%

2 個別分野別の経済インパクト a 経済インパクト試算

グローバル

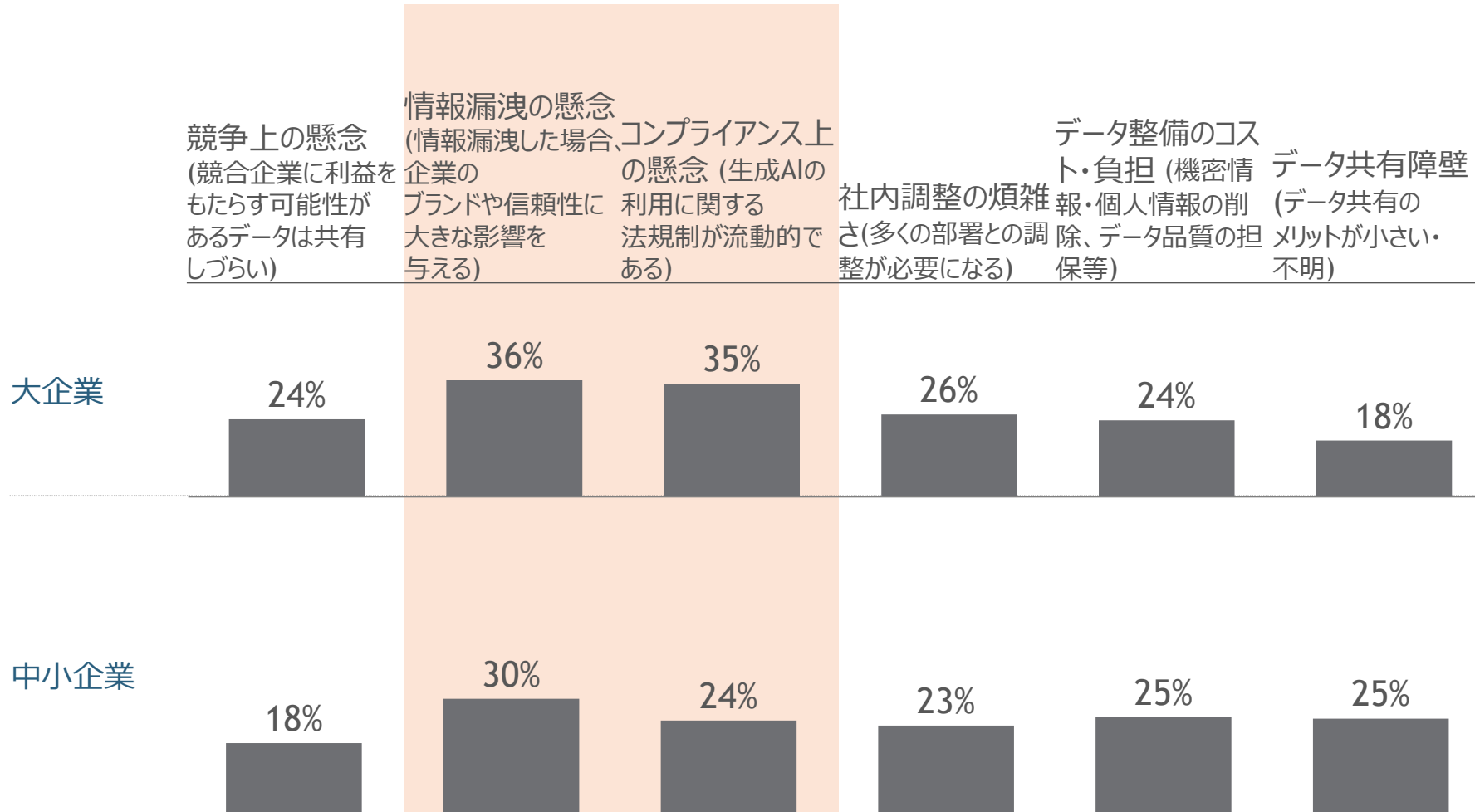
勝ち筋

経済

日本

大企業の方が情報漏洩の懸念、コンプライアンス上の懸念の声が多い

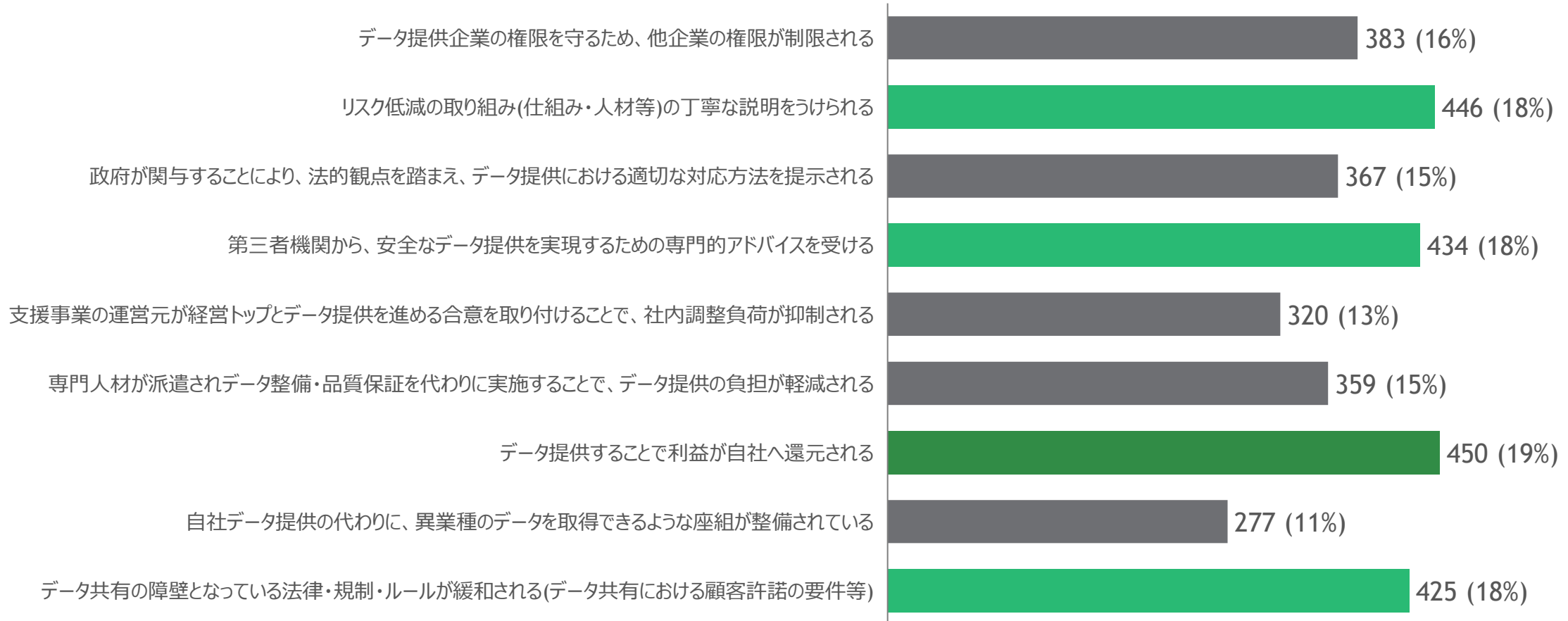
データ共有の障壁（従業員数別）※300人未満/以上



Source: BCG調査

質問文：あなたの勤め先(経営先)のデータを共有するにあたって、データ共有の障壁として該当するものを全てお選びください

「データ提供することで利益が自社へ還元される」ことを望む声が最も多く、ついでリスク低減の仕組み、第三者機関からの専門的アドバイス、規制・ルールの緩和 データ共有において求める支援策（全体）



2 個別分野別の経済インパクト a 経済インパクト試算

グローバル

勝ち筋

経済

日本

自動車産業では他企業の権限制限、製薬・医療・金融保険ではデータ共有の障壁となっている法律・規制・ルールの緩和等、業種間で必要な支援は異なる

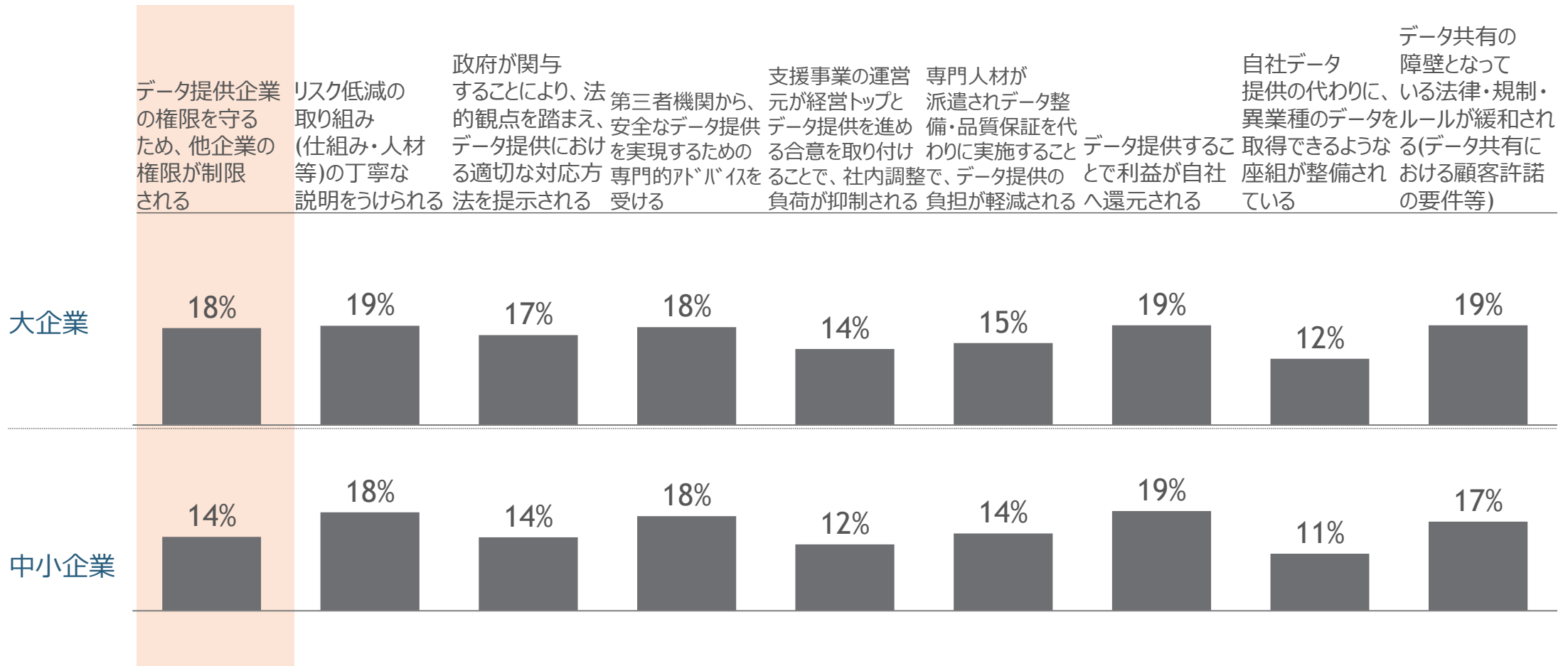
データ共有において求める支援策 (産業別)

□ : 各産業において割合が最大

	データ提供企業の権限を守るため、他企業の権限が制限される	リスク低減の取り組み(仕組み・人材等)の丁寧な説明をうけられる	政府が関与することにより、法的観点を踏まえ、データ提供における適切な対応方法を提示される	第三者機関から、安全なデータ提供を実現するための専門的アドバイスを求める	支援事業の運営元が経営トップとデータ提供を進める合意を取り付けることで、社内調整負担が抑制される	専門人材が派遣されデータ整備・品質保証を代わりに実施することで、データ提供の負担が軽減される	データ提供することで利益が自社へ還元される	自社データ提供の代わりに、異業種のデータを取得できるように座組が整備されている	データ共有の障壁となっている法律・規制・ルールが緩和される(データ共有における顧客許諾の要件等)
自動車・鉄道・船舶・航空	14%	10%	11%	10%	11%	12%	11%	9%	12%
製薬・化学・素材	13%	12%	13%	9%	9%	9%	10%	9%	15%
その他製造業	13%	13%	9%	14%	8%	10%	13%	8%	13%
卸売・小売	9%	16%	9%	15%	10%	10%	14%	7%	10%
医療	9%	12%	13%	13%	9%	12%	13%	5%	14%
介護	8%	12%	14%	13%	11%	11%	15%	6%	11%
建設	9%	15%	10%	14%	10%	12%	11%	8%	10%
サービス	12%	12%	10%	15%	10%	10%	12%	7%	13%
金融・保険	10%	12%	13%	12%	9%	11%	13%	7%	13%
その他	11%	14%	10%	11%	8%	10%	14%	10%	12%
全体平均	11%	13%	11%	13%	10%	11%	13%	8%	12%

大企業においてはデータ提供企業の権限を守るため、他企業の権限が制限されることを望む声が多い

データ共有において求める支援策（従業員数別） ※300人未満/以上



Source: BCG調査

質問文：あなたの勤め先(経営先)でデータ共有を推進する場合、何があればより共有を進められると思いますか。該当するものを全てお選びください。

効率化・コスト削減および新たな価値創出・売上拡大の両側面でのAIの利活用に伴う経済インパクトを試算

経済インパクトの試算方法の概要

A 効率化・コスト削減

各産業へのAI導入に伴う
効率化・コスト削減インパクトを
労務・人件費ベースのトップダウンで試算

~15兆円

- 産業×機能毎に労務・人件費を試算し主要領域を特定
- 主要領域を特定後、インタビュー/アンケートを通してインパクトを試算
 - 主要領域におけるユースケースを詳細化

AIの利活用に伴う
経済インパクト

B 新たな価値創出 ・売上拡大

有望なユースケースへのAI導入に伴う
新たな価値創出・売上拡大インパクトを
ユースケースベースのボトムアップで試算

- エキスパートへのインタビュー・デスクトップ調査を通じて有望なユースケースを特定
 - 主要領域に絞り調査

新たな価値創出・売上拡大に直結するユースケースを特定する

ユースケース特定の考え方

目的・考え方

- 効率化・コスト削減の分析においては、産業×機能を幅広く試算し、有望領域を特定
- 新たな価値創出・売上拡大においては、上記効率化・コスト削減で抜け漏れた新たな価値創出・売上拡大固有のユースケースを拾い上げる

対象産業・対象ユースケースの選定方法

対象産業

- グローバルで生成AIの市場規模が大きい上位4つの産業
 - 金融・保険
 - コンシューマー（小売）
 - ヘルスケア（製薬）
 - メディア（出版、オーディオ、放送、ゲーム、アニメ）
- 国内でエキスパート見立てにより有望とされる領域
 - 製造
 - メディア（出版、オーディオ、放送、ゲーム、アニメ）
 - 医療・介護
 - カスタマーサービス

対象ユースケース

- 以下2つの条件を満たすユースケースを選定
 - 新たな価値創出・売上拡大に直結している（効率化・コスト削減効果が強い事例を除く）
 - 既にユースケースが実現されている、効果が実証されている

グローバルの市場規模・国内有望領域における売上向上に寄与するユースケースは3つへ分類できる

新たな価値創出・売上向上のユースケース



新製品・新技術
開発



顧客体験の
パーソナライズ化



その他

- 創薬
- 製品設計の高品質化
- …
- …

- 顧客にあわせたセールス
- 店舗レイアウトの改善
- 代替品の提案
- 音楽プレイリストの作成
- ユーザーにあわせた情報の提案

- 外科医の手術支援
- ベッド稼働率の向上
- 患者の症状の言語化
- 診断精度向上
- 株式売買の戦略立案

7産業のユースケースを洗い出し、新製品・新技術開発、パーソナライズ化、その他を特定 ユースケース一覧

	産業	機能	ユースケース概要	ユースケースの内容	付加価値/売上インパクト
新製品・新技術 開発	ヘルスケア (製薬)	R&D	創薬	特定のターゲットに対して数万種類の新規化合物を新たに生成し、生成された分子の中から最も有望な候補を優先順位付けする	肝細胞がん（HCC）治療薬の候補を30日で設計・合成
	製造	R&D	製品設計の高品質化	生成AIを活用したCADにより、より軽量で耐久性の高い設計を短時間で作成できるようにする	40%軽量化・20%強化されたシートブラケットを開発
顧客体験の パーソナライズ化	金融・保険	セールスマーケティング	顧客にあわせたセールス	アドバイザーのアウトリーチ頻度、タイミング、具体的な会話のきっかけ、主要なトークポイントをガイドすることで、アップセルとクロスセルを実現	既存顧客の売上が8%増
	コンシューマー	セールスマーケティング	店舗レイアウトの改善	店舗のほぼ全てのデジタルデータが3Dで視覚化されることで、3D ヒートマップと商品の距離測定を使用して店内体験を最適化	N.A
	コンシューマー	セールスマーケティング	代替品の提案	ECサイトにおいて顧客が探している商品の在庫が無い場合、数百の変数をリアルタイムで検討したうえで次善の代替品を提案	代替品の受諾率が95%
	メディア	セールスマーケティング	プレイリストの作成	AI DJが曲紹介やフリートークを交え、個々のユーザーの選好、聴取履歴、気分に基づいたカスタマイズされたプレイリストを生成	ユーザーは視聴時間の25%を本機能へ費やしている
	旅行	オペレーションサービス/CS	ユーザーにあわせた情報の提案	旅行前から旅行後まで、旅行を楽しむのに必要な情報取得・判断をガイド	N.A
その他	医療・介護	オペレーションサービス/CS	外科医の手術支援	患部（手術部）の構造や組織をリアルタイムで視覚化し、外科医をサポート	手術精度の向上
	医療・介護	オペレーションサービス/CS	ベッド稼働率の向上	病院内にスマートセンサーを設置し患者の退院ステップを分析することで、ベッド回転の遅れを減らす・迅速な部屋割りが可能	ベッド稼働率の最大化
	医療・介護	オペレーションサービス/CS	患者の症状の言語化	気になる症状について20問ほどの質問に回答するだけで、その症状に関連する病名と病気を提示	患者の症状を的確に言語化
	医療・介護	オペレーションサービス/CS	診断精度向上	病理医が認識しにくい組織パターンを特定	がん診断の精度向上
	金融・保険	セールスマーケティング	株式売買の戦略立案	企業の実店舗の来店動向情報を用いた予想販売実績に基づき、株式銘柄を売買する定量的戦略を導出	S&P 500より7%オーバーパフォーム

③ 日本の潜在能力・可能性

有識者へのヒアリングを重ね、日本の潜在能力とその可能性を調査 インタビューの設計

インタビュー目的

- グローバルの開発動向の把握
 - 汎用型の大規模化は続くのか、領域特化に分かれていくのか、効率化は進むのか 等
- 日本企業における勝ち筋・競争優位性確保の方向性の把握
 - 高齢化が進む日本では介護・医療分野に領域特化した基盤開発に競争優位性を得られる 等
- 日本企業の開発状況・課題の把握、および開発企業に対する政府支援方法の把握

インタビュー内容

- 1 グローバル：基盤モデルの開発における今後の見立て**
 - 今後の基盤モデルの開発動向はどのように進展していくのか
 - 特に① 汎用型の大規模化、② 領域特化型、③ 効率化 ④ 拡張
 - 基盤モデル開発の進展にあたって、何がボトルネック/チャレンジとなるのか？
 - データ不足、エンジニア不足、計算資源不足、学習コスト高騰、規制 等
- 2 国内：グローバル潮流を踏まえたうえで、日本企業の有望開発領域の見立て**
 - 日本企業が市場競争力を発揮できるのは汎用/領域特化/効率化、のどれか？
 - 領域特化の場合、日本企業が市場競争力を発揮できる具体領域はどこか？
 - 要因としてデータ量、日本の社会的特性、産業の強さを想定
- 3 日本企業の開発状況と課題**
 - 何を目指し、現状どのような開発状況か？
 - 目指す姿/現状の開発状況・内容/活用先の想定/開発方法 (体制/データ) 等
 - 開発過程においてどのような課題があるのか？
- 4 開発企業が期待する政府からの支援**
 - 政府からどのような支援が期待されるか？
 - (例) 公募の仕方、事業前の提供データの内容、開発後の成果共有の在り方 等

3 日本の潜在能力・可能性

インタビューサマリー (1/2)

1 グローバル: 開発動向の見立て

開発動向

- 汎用の大規模化が今後も進むのかは各有識者で意見が別れた
 - 400億パラメーター程度で十分という話もある、今後1年程度で落ち着くのではないか
 - 汎用型の開発は今後2~3年はパラメータの増加に伴う開発が進むものの、いずれ限界を迎える
 - コストがかかっても高性能な出力をするための大規模化は進んでいき、その開発の終わりは予測できない
- 一方、領域特化については発展が進むことで見解が一致
 - 業界によって必要なデータ・モデルが異なるため、領域特化は進んでいく
 - データが集めやすい福祉や医療から進んでいくのではないか
- あわせて、大規模化・領域特化とともに効率化・拡張化も進む見通し
 - 効率化させた小さなモデルによって性能の良いものを手元で使えるようになるのではないか

チャレンジ/ ボトルネック

- 基盤モデルに共通する課題として、データ不足、計算資源不足、プライバシー、規制、正確性が挙げられ、特に領域特化においては領域毎のデータ不足、領域知見との連携が顕著であった
 - 汎用も領域特化も共通してデータ不足が課題
 - 汎用型においてはGPUのコストが課題、領域特化においてはサービスプロバイダー側が専門領域のデータを理解したり整理したりするナレッジが不足していることが課題

2 国内: 有望領域の見立て

汎用

- 汎用でグローバル企業に追いつくのは簡単ではないが、汎用の開発を行うことで開発能力を育成する必要
 - 汎用モデルと領域別モデルの開発プロセス自体には差がないので、技術的に求められることは同じ
 - 汎用を作って領域特化にさせていくということが重要、汎用で追いかけて勝つ必要がある
 - 今後もグローバルテックの汎用モデルがオープンソースであり続けるかはわからない、日本企業も汎用に取り組む必要

領域特化

- 日本が産業的国際競争力・データ・日本独自の文化を有する領域が有望
 - 業種では、製造・創薬・材料等の製造業、アニメーション、医療(介護)は国際競争力、データ共に存在
 - 機能では、カスタマーサービスは日本独自の質の高さに加え、一定の国際競争力が存在するのではないか
 - 人材育成については日本独自の文化を活かすことが可能

インタビューサマリー (2/2)

3

日本企業:
目指す姿
に向けた
課題

目指す姿

- 各社ビジネスモデルに応じて異なるものの、汎用型の開発を引き続き継続する企業も領域特化の開発にシフトする企業も存在
 - 今後企業とのアライアンスによって領域特化への取組みを進めていく
 - 自社のビジネス活用(広告、メディア、ゲーム等)目的で、基盤モデルの開発を推進していく
 - 用途によって使い分けをするため、自社での汎用モデル開発も進めていく

課題

- GPUコスト(学習時/運用時)、汎用・領域共にデータの不足、領域特化におけるデータプロバイダーとAI事業者の連携の難しさ、AIの評価制度の不十分さ、学習データの透明性を保ったうえでの学習の限界、等が挙げられた
 - 学習だけでなく利用時の計算のコストも高く、キャッシュ体力のない日本企業が負担することは難しく何らかの支援は有効
 - AIの評価制度はLLMの評価方法として十分に信頼でき標準となっているものが存在しない

全体

- 計算資源、データ提供、業界横断での連携強化における支援の声が多かった
 - マシンパワーのコストが大きい、ここを資金提供してもらえるとありがたい
 - 政府固有のデータ(国会図書館・公文書館・国会議事録・国立がんセンター等)は共有されると嬉しい
 - スタートアップでは接点を持ってない大手企業と繋いでもらったり、業界内でスクラムを組むことを推進していただければ嬉しい

4

政府に
期待する
支援公募
事業

- 公募：対象を絞らず広く、産業レベルで絞ってほしい、という両方の意見があがった
 - 応募のしやすさは汎用モデルの開発者としては広い方が良い
 - 領域特化の開発を進めていくのであれば、各分野への導線を作るような公募が良い
 - 汎用よりも領域特化のどこかでドメインを狙ったほうが、事業としての成功確率は高いのではない
- データ提供：政府保有データに対するニーズが高いことが確認された
 - 教育関連(試験問題や教科書)のデータ、各省庁のデータ、出版物のデータが提供されるとありがたい
- 成果共有：オープンソースといえど、各社企業競争力に関わる部分の公開は様子見で、どこまで公開を期待するのか政府として示す必要
 - オープンソースで開発しているが、データセット、データ処理やモデルの構築方法等、企業固有情報の公開は難しい。
 - 一方で、国の事業として完全にオープンイノベーションでやるのであればブレイクスルーになる可能性もある
 - モデルの公表はしやすいものの、情報の加工や加工した情報の扱い方に関しては公開は避けたい
- 社会実装：著作権や既存の安全にかかわる規制に関して企業側から緩和を求める声があがった
 - 音楽など、現在著作権がハードルとなって使用できないデータの使用許諾が得られるような制度があるとありがたい
 - 介護／薬局／教育等、モノを作っても出口がない(人を配置することを求める)分野が多い。
 - まずは人不足な業界で、現在人の配置をマストにしているもので外せるような規制緩和があると良い
 - 例えば薬局では薬剤師が必ず担当しないといけない、そこを自動化できるものを作っても実現できない

日本の潜在能力・可能性においては、産業的国際競争力・データをもつ製造業・医療や、日本独自の文化を有するカスタマーサービス・アニメーションが有望

③日本の潜在能力・可能性

エキスパートの見立て



製造業

- 日本が産業的に強い製造業・ものづくりにて、業界横断で課題感がある図面生成などの工程で生成AIのモデル開発ができると競争力を発揮できるのでは
- 画像生成AIで他のAIを強化する訓練生成データを作っていくことやロボットをLLMで制御することができるのではないかと

医療

- 高齢者の状態や過去のヘルスデータ・起用歴を学習材料に用いて、高齢化が進む社会にて活用される生成AIの開発ができると、高齢化が進む先進国を対象に一定の競争力を発揮できるのではないだろうか

アニメーション

- オープンに拾えるものでやっている企業が多い中で、日本は多くのアニメの著作権を持っている
- 学習データ自体は世界中で手に入るの権利をまとめることが重要

カスタマーサービス

- ホテルのフロント対応を含む丁寧なカスタマーサービスが重要視される場面において、日本のクオリティの高いおもてなし等がバーチャル上で実現できると差別化に繋がるのではないかと

法務・人事

- (法務) 法律は各国独自に存在するため競争力があるのではないかと
- (人事) 日本は特有の人材育成や採用人材をおこなっている。エージェントの代わりに企業をつなぐ、メンターとしてのAI等に可能性があるのではないかと

④ 勝ち筋の特定

4 我が国としての勝ち筋の特定

我が国としての基盤モデル開発の勝ち筋を、グローバルの動向/経済インパクト/日本の潜在能力・可能性の調査・分析を通じて特定

(1)生成AI・基盤モデルに関する市場動向調査：検討アプローチ

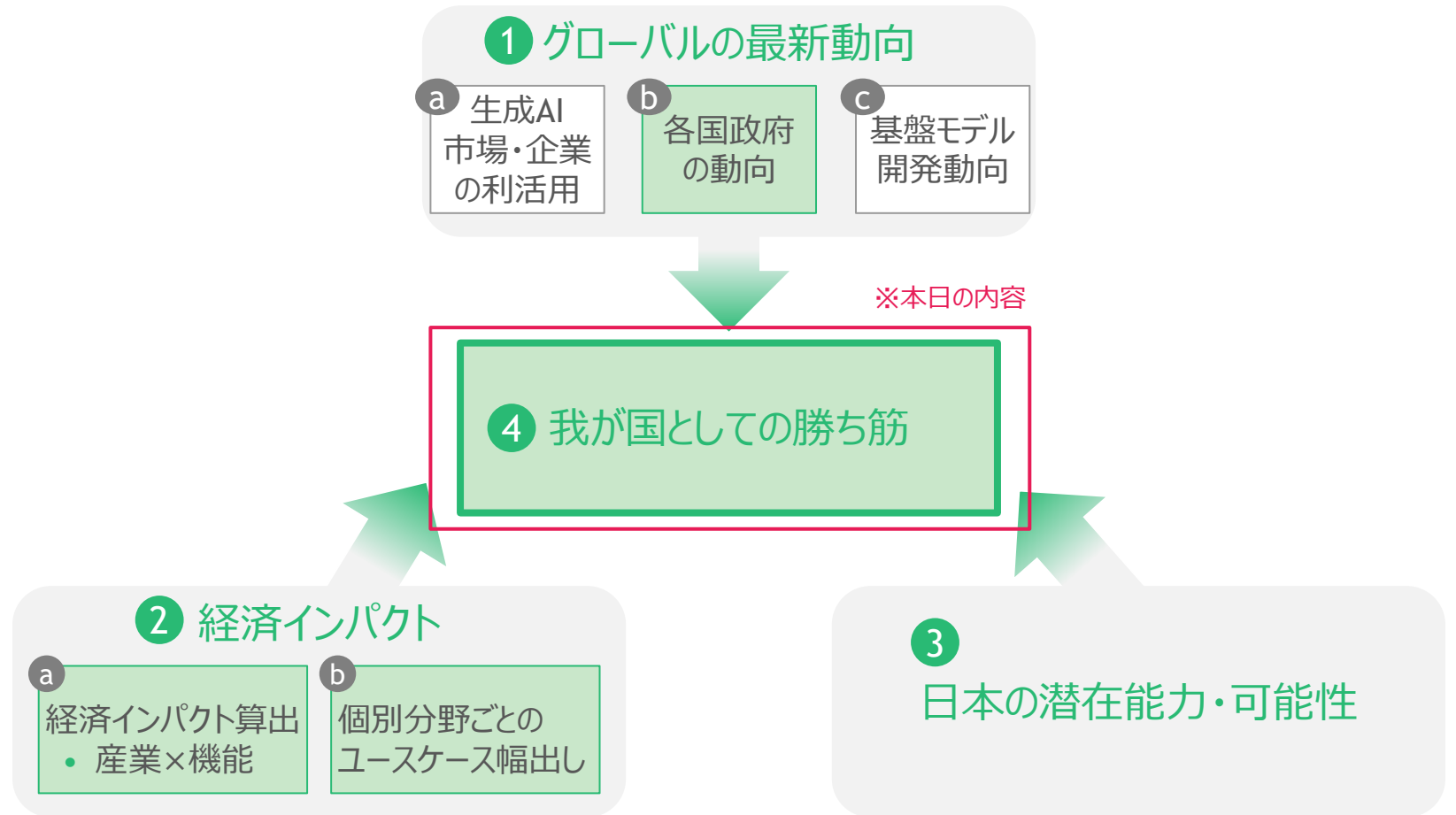
仕様書の記載内容

検討アプローチ

□:仕様書記載内容 ■:追加提案

以下の項目について、生成AI、特に基盤モデル分野等における市場動向、企業動向、開発動向等について、調査・分析を行う

- 1 グローバルの最新動向 (米欧中含む3カ国以上)
- 2 生成AIの個別分野別の経済インパクト (情報通信業、製造業、金融業を含む3種類以上)
- 3 日本の生成AI分野における潜在能力・可能性等に関するヒアリングなどの調査、調査結果に対する分析



4 我が国としての勝ち筋の特定

汎用開発を行うことで基盤モデルの開発能力を育成し、日本としての勝ち筋領域の製造×製造加工、金融保険・医療×オペレーション/CS等の領域特化を開発

生成AI・基盤モデルに関する市場動向

1 グローバル最新動向

- 汎用の大規模化のみではなく、領域特化・効率化・拡張にも開発は進展
- グローバル企業も医療、SW開発、広告等領域特化の開発を進める
 - 特にSW開発は言語の違いがない中で、膨大なコードデータを保有しておりグローバル企業が優位

2 経済インパクト

- コスト削減効果が特に大きい（5,000億円以上）有望領域
- ①自動車・鉄道・船舶・航空×製造加工、②製薬・素材・化学×製造加工、③その他製造業×製造加工
 - ④卸小売×調達サプライチェーン、⑤サービス×オペレーション/CS
 - ⑥ソフトウェア×ソフトウェア開発/アプリケーション開発、⑦金融保険×オペレーション/CS
 - ⑧バックオフィス、⑨業務横断（資料作成等）
- 次いでコスト削減効果は大きい（3,500億円以上）有望領域
- ⑩医療×オペレーション/CS、⑪運輸・輸送×オペレーション/CS
- 付加価値向上・売上向上効果の有望領域は以下の通り
- ⑫新製品・新技術開発、⑬顧客体験のパーソナライズ化

3 日本の潜在能力・可能性

- グローバル企業に追いつくのは簡単ではないが、汎用の開発を行うことで開発能力を育成する必要
- 汎用モデル・領域別モデルの開発プロセスには共通点が多い、技術的に求められることは同じ
 - 今後もグローバルテックの汎用モデルがオープンソースであり続けるかは不透明
- 領域特化においては、日本が産業的国際競争力・データ・日本独自の文化を有する領域が有望
- 業種：製造・創薬・材料等の製造業、アニメーション、医療は国際競争力、データ共に存在
 - 機能：カスタマーサービスは日本独自の質の高さに加え、一定の国際競争力が存在
法務・人事は日本独自の商習慣があり、かつデータも存在

我が国としての勝ち筋

汎用

- 汎用モデルの開発を推進することで基盤モデル共通の開発能力を育成
- 技術開発：汎用・領域特化ともに開発プロセスは同様で、汎用の開発力はAI全体に直結
 - 国際情勢：現状OSSのグローバル企業のクラウド化、収益化の可能性が存在

領域特化

- ①グローバル最新動向、②経済インパクト、③日本の潜在能力・可能性を踏まえ、我が国として勝ち筋領域8つに絞り開発を促進
- 製造業×製造加工
 - 自動車
 - 製薬・化学・素材
 - その他製造業（機械製造）
 - サービス×オペレーション/CS
 - 金融保険×オペレーション/CS
 - バックオフィス（法務・人事）
 - 医療×オペレーション/CS
 - アニメーション

4 我が国としての勝ち筋の特定

汎用・領域特化モデルの開発プロセスは共通点が多く、汎用モデルを開発することでその開発ノウハウを領域特化モデルの開発に活かす

基盤モデルの開発プロセスごとの開発方法

モデル構築
プロセス



開発方法

汎用

膨大な公開データから、質の高いデータを定義・見極め

- 公開データは豊富なため、データ収集自体は容易

不完全なデータを整理・加工

- 完全なデータは稀で、特にクレンジングが重要

目的を踏まえたアルゴリズムの選択、実装

- 一定一般化されており、重要度は低い

目的を踏まえた評価手法の決定

- 技術的な評価に加え、ビジネス要件踏まえた評価が必要

領域特化

汎用に加え、さらに領域特有の知見をふまえたデータの探索・質の定義が必要

- データになっていない専門的な知見等も収集が必要

汎用に加え、さらに領域特有の知見を踏まえたデータ前処理が必要

- 医療におけるCT画像、電子カルテ等

汎用と同様、目的を踏まえたアルゴリズムの選択、実装

汎用に加え、領域特有の知見を踏まえた評価基準の設定が必要

- 評価自体にもエキスパートが参加

エキスパートの見立て

- 汎用モデルと領域別モデルの開発プロセス自体には差がなく、技術的に求められることは同じ
- 領域特化モデルを作る場合、汎用モデルに対してデータをアドオンすればよい
- 汎用を作って領域特化にさせていくということが重要、汎用で追いかけて勝つ必要がある

4 我が国としての勝ち筋の特定

①グローバル最新動向、②経済インパクト、③日本の潜在能力・可能性を踏まえ、領域特化における8領域の勝ち筋を特定

考え方

評価

勝ち筋

領域 (産業×機能)	インパクト ¹ (②経済インパクト)	フィージビリティ	
		競争環境 (①グローバル動向)	潜在能力・可能性 (③日本の潜在能力)
自動車・鉄道・船舶・航空× 製造加工	◎	-	◎
製薬・素材・化学× 製造加工	◎	-	◎
その他製造業× 製造加工	◎	-	◎ 機械製造
卸小売× 調達サプライチェーン	◎	-	- コンセプト段階で 未実現
サービス× オペレーション/CS	◎	-	◎
ソフトウェア× ソフトウェア開発/アプリケーション開発	◎	✗ グローバル企業が膨大な コードデータを保有し、 言語の差異もない	-
金融保険× オペレーション/CS	◎	-	◎
バックオフィス	◎	-	◎ 法務・人事
業務横断(資料作成等)	◎	-	- 各産業毎のユース ケースに組み込み
医療×オペレーション/CS	◎	-	◎
運輸・輸送×オペレーション	◎	-	- コンセプト段階で 未実現
アニメーション	-	-	◎

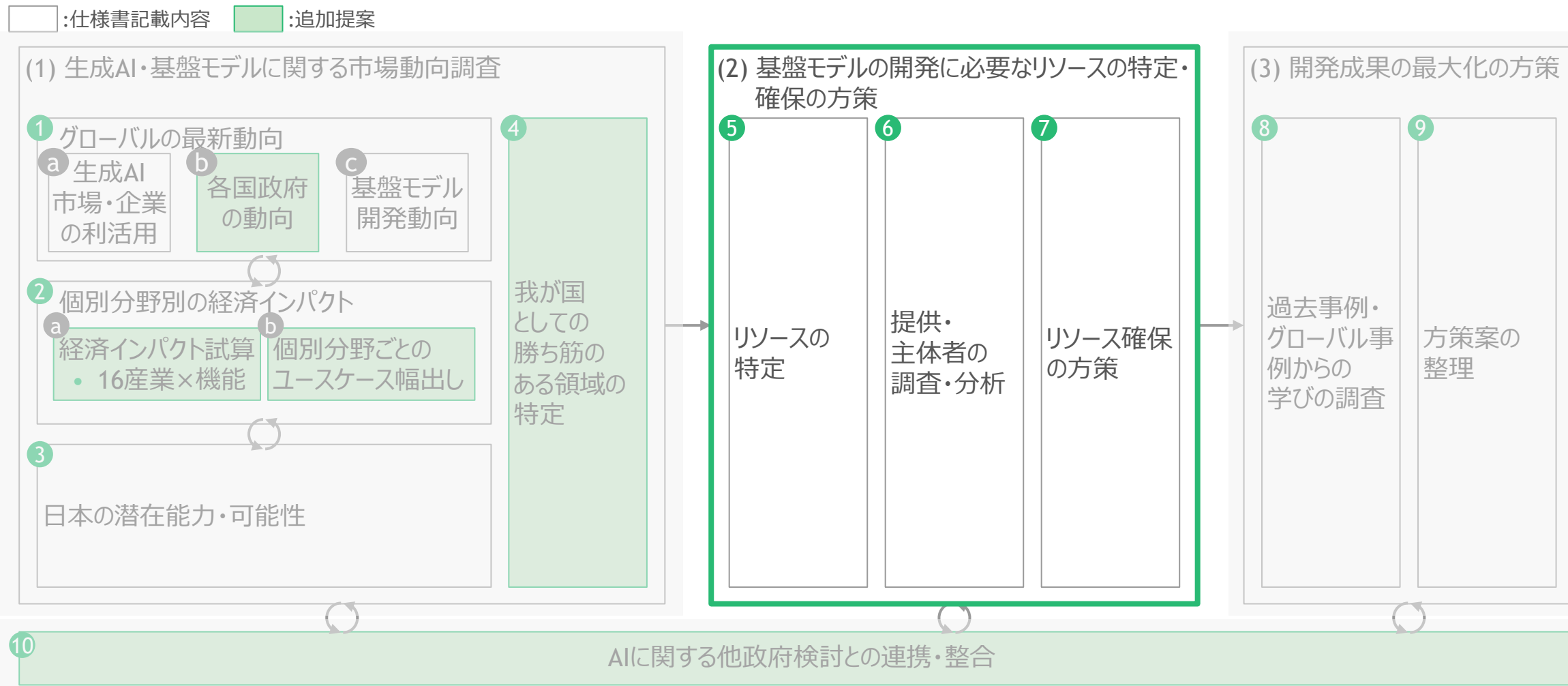


- 1 自動車・鉄道・船舶・航空×
製造加工
- 2 化学・製薬・素材×
製造加工
- 3 その他製造業 (機械製造)×
製造加工
- 4 サービス×
オペレーションサービス/CS
- 5 金融・保険×
オペレーションサービス/CS
- 6 バックオフィス (法務・人事)
- 7 医療×
オペレーションサービス/CS
- 8 アニメーション

1. 効率化、コスト削減効果が3,500億円以上の領域を特定

(2) 基盤モデルの開発に必要なリソースの特定・ 確保の方策

事業実施の基本方針 : (2) 基盤モデルの開発に必要なリソースの特定・確保の方策



-
- ⑤ 必要なリソースの特定
 - ⑥ リソースの提供者・主体

勝ち筋領域におけるユースケースを明らかにした上で、当該ユースケースに必要なデータ、データ提供者・主体、確保の方策を調査

目的

- 勝ち筋領域における領域特化モデルの開発に必要なデータ、データ提供者・主体、確保の方策を明らかにする
- なお、実際の開発における開発方針、データ収集の優先度は開発事業者にて検討する

調査設計

調査アプローチ

ユースケース特定

- 勝ち筋領域において、国内外で検証・検討されているユースケースを特定

データ特定

- 各ユースケース毎に必要なデータを特定
 - 実際のデータ活用有無は開発事業者の開発方針に依存

データ提供者・主体者

- 当該データの提供者・主体となりうる主体を洗い出し

データ確保の方策

- データ提供・収集時の障壁およびその乗り越え方を検討

調査対象

日本の勝ち筋領域である以下8つの領域

- 製造業×製造加工
 - 自動車
 - 製薬・化学・素材
 - その他製造業（機械製造等）
- サービス×オペレーション/CS
- 金融保険×オペレーション/CS
- バックオフィス（法務・人事）
- 医療×オペレーション/CS
- アニメーション

5 必要なリソースの特定 6 リソースの提供者・主体

製造業 (自動車)×製造加工サマリー

ユースケース 特定

自動車×製造加工のバリューチェーンは**製品設計～製品検査まで、各バリューチェーンごとにユースケースが存在**

- 製品設計：類似した図面の検索支援、最適化された図面の提案、自然言語入力に基づいた工場製造ラインの設計
- 製造：生産設備における機械設定支援、製造ロボットの自走支援等
- トラブルシューティング/メンテナンス：トラブルからの復旧支援
- 製品検査：検査の自動化

データ特定

各ユースケースに必要なデータは、大きく分けて**製品設計情報、生産設備情報、検査情報、取引先情報、市場動向情報、規制・ガイドライン**に分かれる

- 製品設計：製品設計情報、生産設備情報、取引先情報、市場動向情報、規制・ガイドライン
- 製造：製品設計情報、生産設備情報、規制・ガイドライン
- トラブルシューティング/メンテナンス：製品設計情報、生産設備情報、検査情報、規制・ガイドライン
- 製品検査：製品設計情報、生産設備情報、検査情報、規制・ガイドライン

データ提供者・ 主体者

企業内に閉じた・**非公開情報が多く、データ量としては自動車部品メーカーが多く保有**

- 製品設計情報
 - データの多くは非公開のデータとして自動車メーカーや自動車部品メーカーに存在
- 生産設備情報・検査情報・取引先データ
 - 生産品に応じて、生産工場が非公開データとして保有しており、自動車メーカーや自動車部品メーカーに多くが集積している
- 市場動向情報、規制・ガイドライン
 - 既に公開されている。市場動向情報に関しては一部民間企業が有するものもあるが、基本的には公開されている

データ確保の 方策

自動車部品メーカーに存在する**非公開情報の中の秘匿性が低い情報を中心に確保と活用を進めていくことが重要**

- 既に公開済みの情報：開発事業者からの要望に応じて事務局によるマッチングを進める
- 有償データ：事務局として購入支援を検討する
- 非公開データ：秘匿性が高い情報と低い情報に分類ができるので、事務局として秘匿性が低いデータのデータ化と整理を支援したうえで、業界横断でのデータ活用を支援することを検討

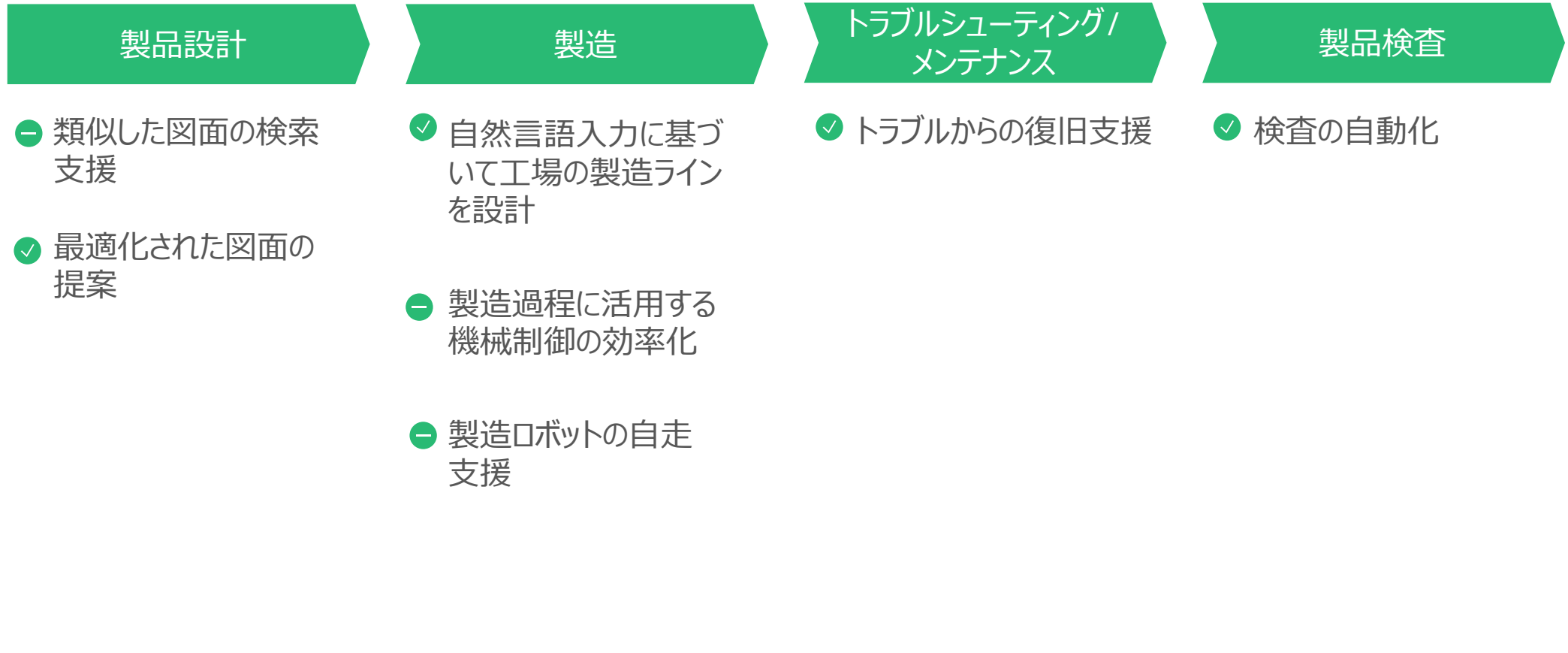
- ⑤ 必要なリソースの特定 ⑥ リソースの提供者・主体

自動車×製造加工のバリューチェーンは製品設計～製品検査までに分かれ、想定されるユースケースは以下の通り

自動車×製造加工のユースケース

- ✓ 検証済み - 検証中 ? コンセプト段階

ユース
ケース



5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、製品設計情報、生産設備情報、検査情報、取引先情報、市場動向情報、規制・ガイドラインに分かれる

必要なデータの特定

	ユースケース名	ユースケース概要	モデル構築に必要な学習データ
製品設計	類似した図面の検索支援	蓄積されたデータから、作成予定のものに類似する図面を探し出し、図面作成を効率化する	<ul style="list-style-type: none"> 製品設計情報
	最適化された図面の提案	生成AIを活用したCADにより、製品設計者がより優れた、軽量かつコスト効率の高い設計を短時間で作成できる	
	自然言語入力に基づいて工場の製造ラインを設計	自然言語入力に基づいて工場内の製造ラインを設計。効率を高め、安全な作業環境を確保するために最適化されたレイアウトを推奨する	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 取引先情報 市場動向情報
製造	製造過程に活用する機械制御の効率化	自然言語入力に基づいてコードを生成することにより、幅広いオートメーションタスクの PLC プログラミングを容易にする	<ul style="list-style-type: none"> 製品設計情報 生産設備情報
	製造ロボットの自走支援	ロボットがロボットが自然言語と視覚入力に基づき、事前のプログラミングや訓練を受けていない新しいタスクを実行	
メンテナンス	トラブルからの復旧支援	トラブルのレポートと、その解決のためにARガイダンス等を用いたメンテナンス指示を行いオペレーターを支援	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 検査情報
製品検査	検査の自動化	対象データ（検査画像）に合わせて、自動的に判定モデルを生成。目標精度・性能になるまで繰り返すトライアル&エラーにかかる手間を削減	<ul style="list-style-type: none"> 製品設計情報 検査情報 規制・ガイドライン

5 必要なリソースの特定 6 リソースの提供者・主体

自動車部品メーカーが保有する非公開データには秘匿性の高低が存在し、秘匿性が低いデータに関してはデータ共有も可能

非公開情報におけるデータの分類

非公開の情報

データの概要

データの例

秘匿性が高い情報

企業の競争力に直結する技術に関する情報

- 発展途上の製品に関するデータ
 - 電動化や自動化に関するデータ
- 設計仕様に関するデータ
 - 高い技術を要するシステムのメカニズムに関するデータ等

秘匿性が低い情報

現時点では非公開ではあるものの、企業競争力にはあたらない情報

- 製品成熟度が高い製品に関連するデータ
 - 空調や自動車補修関連製品
- 他社と差異が出にくいデータ
 - 材質や生産設備に関するデータ等

エキスパートの声¹

“ 設計図は実は秘匿情報になるのは2-3割程度で、多くは他企業へ公開しても問題ない。秘匿情報でないにも関わらず、多くの企業が個別に設計図作成を行っており、かなりの工数がかかっている

➤ “ 政府には、GPU資源の提供などの支援はもちろん、業界横断でデータをオープン化する際に、オールジャパンで取り組むための支援をしてほしい

製品の種類/製造法/品質の高さ/コスト等によって秘匿度合いは変わる

1) 大手部品メーカーの社会イノベーション室のエキスパート

5 必要なリソースの特定 6 リソースの提供者・主体

製造業 (製薬)×製造加工サマリー

ユースケース 特定

製薬×製造加工のバリューチェーンは製品設計～製品検査まで、各バリューチェーンごとにユースケースが存在

- 製品設計：最適な製造条件の設定支援
- 製造：製造作業の効率化支援
- トラブルシューティング/メンテナンス：メンテナンスレポートの作成支援、画像や文書情報を基にした異常検知支援
- 製品検査：品質検査の支援

データ特定

各ユースケースに必要なデータは、大きく分けて製品設計情報、生産設備情報、検査情報、取引先情報、規制・ガイドラインに分かれる

- 製品設計：製造設計情報(標準作業手順書[工程管理書/製造指示書/記録書])、生産設備情報、取引先情報、市場動向情報、規制・ガイドライン
- 製造：製造設計情報、生産設備情報、規制・ガイドライン
- トラブルシューティング・メンテナンス：製造設計情報、生産設備情報、検査情報、規制・ガイドライン
- 製品検査：製造設計情報、生産設備情報、検査情報、規制・ガイドライン

データ提供者・ 主体者

製品の製造にかかわるデータは非公開が多く、データ量的には製薬企業や医薬品受託業者が多くを保有している

- 製品設計情報：開発や製造数に応じて、製薬企業、医療機関、医薬品受託業者が保有している
- 生産設備情報、検査情報、取引先情報：製造を担う医薬品受託業者に集約されている
- 市場動向情報、規制・ガイドライン：一部有償のものが含まれるものの、既に公開されている

データ確保の 方策

企業の競争源泉になる非公開情報が多くを占め、その多くを製薬企業や医薬品受託業者が保有しているため事務局としては各社の取組みを製薬業界で横展開する際の支援を行う形式になるか

- 既に公開済みの情報：開発事業者からの要望に応じて事務局によるマッチングを進める
- 有償データ：事務局として購入支援を検討する
- 非公開データ：各社に閉じて存在しているうえ企業の競争源泉にもなるデータのため、製薬企業の個社における生成AI開発を推進したうえで、事務局としては業界における横展開する際の支援になるか

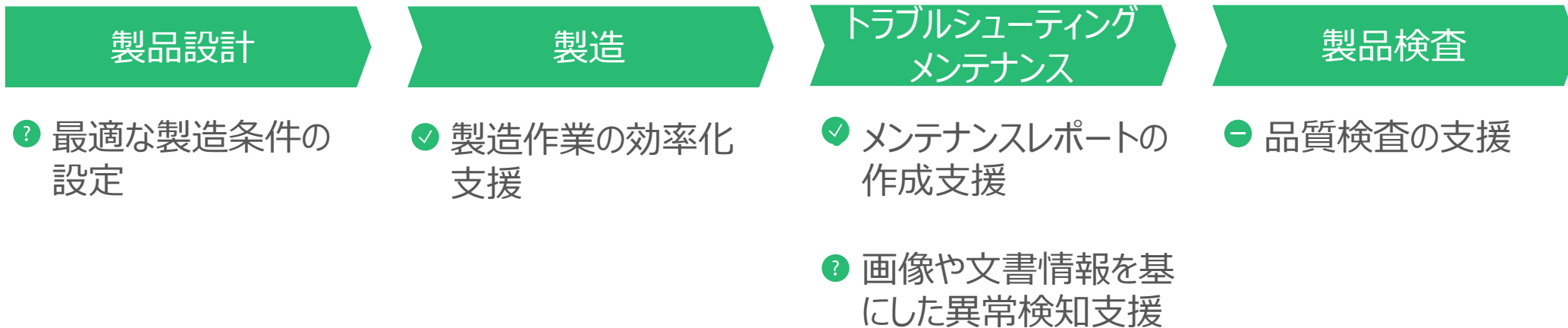
- ⑤ 必要なリソースの特定
- ⑥ リソースの提供者・主体

製薬×製造加工のバリューチェーンは製品設計～製品検査に分かれ、各バリューチェーンごとにユースケースが存在

製薬×製造加工のユースケース

- ✓ 検証済み
- ⊖ 検証中
- ? コンセプト段階

ユース
ケース



5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、製品設計情報、生産設備情報、検査情報、取引先情報、規制・ガイドラインに分かれる

必要なリソースの特定

	ユースケース名	ユースケース概要	モデル構築に必要な学習データ
製品設計	最適な製造条件の設定	生産する医薬品の種類・量や出荷時期、工場の設備状況に応じて最もコストが低い方法で生産できる製造条件の策定を支援 (将来的に実現できればニーズが高いとのエキスパートコメントあり)	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 取引先情報 規制・ガイドライン
製造	製造作業の管理支援	標準手順書に書かれている作業の順番を音声でガイドすることで、作業員の作業を効率化できる	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 規制・ガイドライン
トラブルシューティング・メンテナンス	メンテナンスレポートの作成支援	音声で作業者に異常がないかを確認し、作業員の返事をもとに自動的にメンテナンスレポートを作成することでレポート作成を自動化できる	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 検査情報 規制・ガイドライン
	画像や文書情報を基にした異常検知支援	撮影した写真を生成AIに判定させることで以上の有無を判断することでトラブルシューティングを効率化できる	
製品検査	品質検査の支援	製品をセンサーでスキャンしたうえで出荷基準を満たすことを確認し、品質検査を効率化する	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 検査情報 規制・ガイドライン

5 必要なリソースの特定 6 リソースの提供者・主体

製造業 (機械製造)×製造加工サマリー

ユースケース 特定

機械製造×製造加工のバリューチェーンは製品設計～製品検査まで、各バリューチェーンごとにユースケースが存在

- 製品設計：生産計画を現場最適化するための支援、工場間の製造計画の最適化
- 製造：製造ロボットの自走支援
- トラブルシューティング/メンテナンス：メンテナンスの効率化
- 製品検査：品質検査の支援

データの特定

各ユースケースに必要なデータは、大きく分けて製品設計情報、生産設備情報、検査情報、取引先情報、市場動向情報、規制・ガイドラインに分かれる

- 製品設計：製品設計情報、生産設備情報、取引先情報、市場動向情報、規制・ガイドライン
- 製造：製品設計情報、生産設備情報
- トラブルシューティング/メンテナンス：製品設計情報、生産設備情報、検査情報、規制・ガイドライン
- 製品検査：製品設計情報、生産設備情報、検査情報、規制・ガイドライン

データ提供者・ 主体者

製品や部品に関わる情報は非公開情報として各機械メーカーに閉じて保有しており、公開情報は市場動向や規制など一部に限られる

- 製品設計情報・生産設備情報
 - 製品設計に使用する図面および生産設備に関する情報はいずれも各機械メーカーに閉じており非公開
- 検査情報・取引先データ
 - 非公開かつ、各機械メーカーだけでなく機械の発注元も含め散在している
- 市場動向情報、規制・ガイドライン
 - 一部民間企業が有する有償のものも含まれるが、基本的には既に公開されている

データ確保の 方策

非公開データが散在している状況を踏まえ、データを多く保有している企業へのアプローチを行うべきか

- 既に公開済みの情報：開発事業者からの要望に応じて事務局によるマッチングを進める
- 有償データ：事務局として購入支援を検討する
- 非公開データ：データが各機械メーカーに散在しているため、事務局としては一定量のデータを保有する企業と連携し、業界の効率化推進を図るべきか

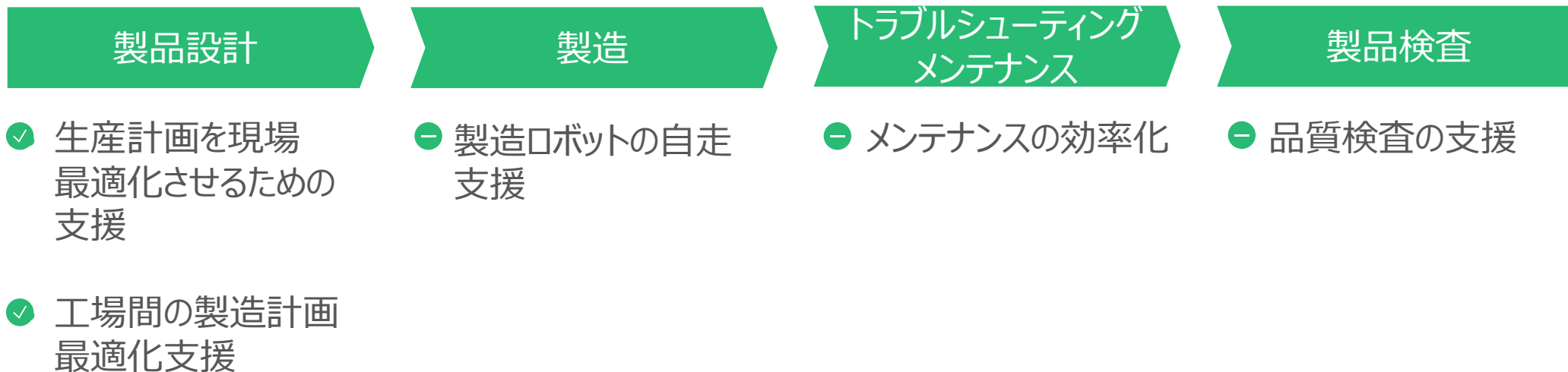
- ⑤ 必要なリソースの特定 ⑥ リソースの提供者・主体

機械製造×製造加工のバリューチェーンは製品設計～製品検査に分かれ、各バリューチェーンごとにユースケースが存在

機械製造×製造加工のユースケース

- ✓ 検証済み - 検証中 ? コンセプト段階

ユース
ケース



5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、製品設計情報、生産設備情報、検査情報、取引先情報、規制・ガイドラインに分かれる

必要なリソースの特定

ユースケース名	ユースケース概要	モデル構築に必要な学習データ	
製品設計	生産計画を最適化させる支援	販売目標、機械や人間のキャパシティを踏まえた生産計画策定を支援することで製造計画を改善できる	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 取引先情報
	工場間における製造計画調整の最適化支援	自社の他工場との生産調整を行うことで工場連携を推進し、生産を効率化する	<ul style="list-style-type: none"> 市場動向情報 規制・ガイドライン
製造	製造ロボットの自走支援	ロボットが自然言語と視覚入力に基づき、事前のプログラミングや訓練を受けていない新しいタスクを実行	<ul style="list-style-type: none"> 製品設計情報 生産設備情報
トラブルシューティング・メンテナンス	全数検査の効率化	逸脱が発生した際に、原因が判明するまで検査を繰り返す必要があるため	<ul style="list-style-type: none"> 製品設計情報 生産設備情報
	メンテナンスの効率化	蓄積された設備に関する情報に基づいて、設備トラブルを事前に防止する	<ul style="list-style-type: none"> 検査情報 規制・ガイドライン
製品検査	品質調査の支援	過去のデータに基づいて、製品の目視で確認できない部分の異変やエラーを検出することで品質調査の精度を向上させる	<ul style="list-style-type: none"> 製品設計情報 生産設備情報 検査情報 規制・ガイドライン

5 必要なリソースの特定 6 リソースの提供者・主体

サービス業 (飲食/宿泊)×オペレーション・CSサマリー

ユースケース 特定

サービス産業におけるオペレーション・CSのユースケースは受け入れ準備、サービス提供、(会計)、フォローアップに紐づく形で存在

- 受け入れ準備: 稼働状況を踏まえたうえでの予約受け入れの自動化
- サービス提供: ドライブスルー受付の完全自動化、顧客質問への自動、顧客のリクエストを反映したサービス内容の提案
- フォローアップ: 顧客フィードバックに基づきリピーター・新規顧客獲得支援

データ特定

各ユースケースに必要なデータは、稼働データ、顧客データ (顧客ログも含む)、サービスデータ、トレンドデータ、法令・ガイドライン

- 受け入れ準備: 稼働データ、顧客データ、サービスデータ
- サービス提供: 稼働データ、顧客データ、サービスデータ、トレンドデータ、法令・ガイドライン
- フォローアップ: 顧客データ、サービスデータ、トレンドデータ

データ 提供者・ 主体者

モデル構築に必要な情報は、基本的に各企業が保有しているため、情報を保有する企業との連携が必要不可欠

- 稼働データ、顧客データ、提供サービスデータ (詳細):
一部公開されているものもあるが、基本的には非公開で各企業が保有
- トレンドデータ: 一部公開のものもあるが、基本的には民間企業から有償で提供されている
- サービスデータ (概要)、法令・ガイドライン: 既に公開されている

データ確保の 方策

事務局として、非公開情報を保有する企業間の連携を促進し、モデル開発を後押しする

- 公開データ: 事務局が必要な情報を整理・一覧化してマッチング
- 有償データ: 公開情報で補えないものは、事務局として購入支援を検討
- 未データ化: 事務局としてデータ化の支援を検討
- 非公開データ: 非公開データを保有する複数企業がコンソーシアムを構築してデータ共有できる仕組みの構築を支援する

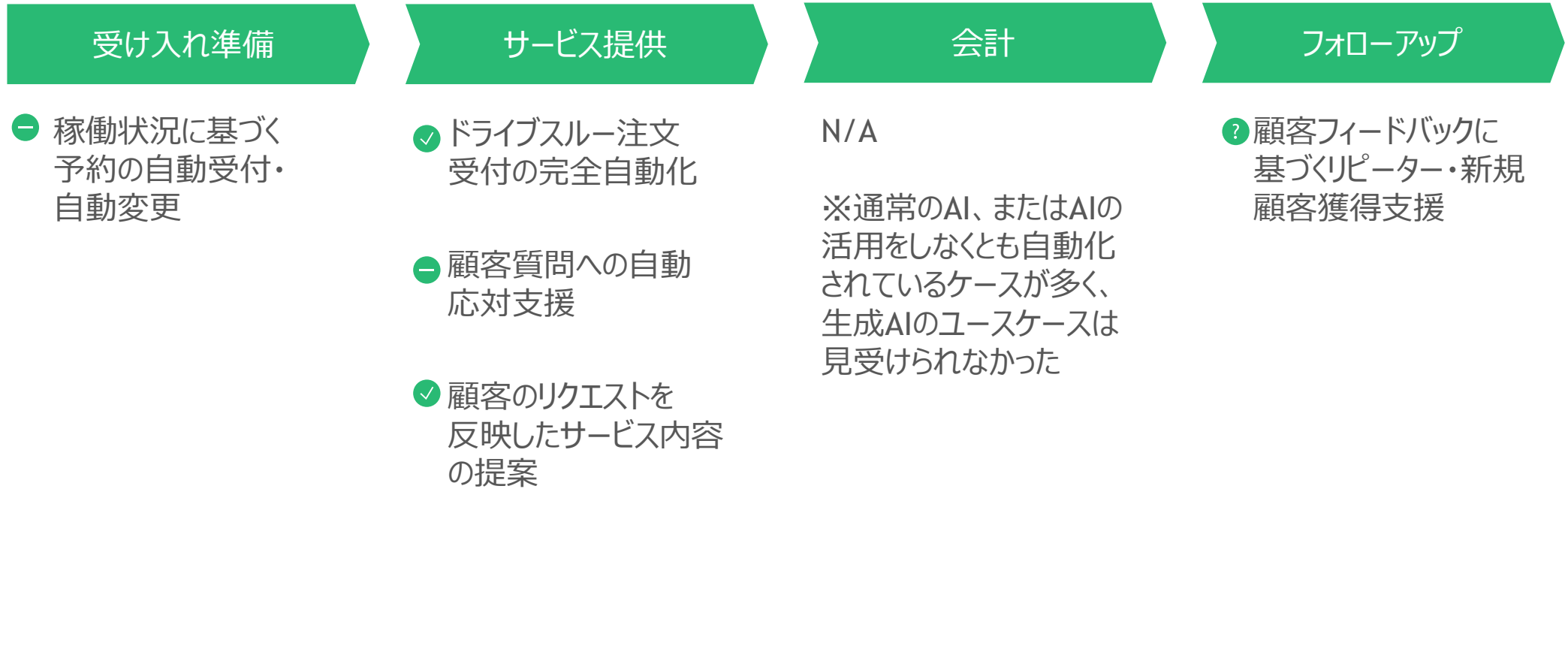
5 必要なリソースの特定 6 リソースの提供者・主体

サービス業×オペレーション・CSのバリューチェーンは受け入れ準備～フォローアップまでに 分かれ、想定されるユースケースは以下の通り

サービス業×オペレーション・CSのユースケース

✓ 検証済み - 検証中 ? コンセプト段階

ユース
ケース



5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、店舗稼働データ、顧客データ (顧客ログも含む)、提供サービスデータ、業界等のトレンドデータ、法令・ガイドラインに分かれる

必要なデータの特定

	ユースケース名	ユースケース概要	モデル構築に必要な学習データ
受け入れ 準備	稼働状況に基づく予約の自動受付・自動変更	店舗の稼働状況を踏まえたうえで、予約を自動で受付て来客に備える	<ul style="list-style-type: none"> 稼働データ 顧客データ サービスデータ 法令・ガイドライン
	ドライブスルー注文受付の完全自動化	ドライブスルーにおける顧客からの注文 (音声) を自動で受け付け、厨房/清算でのプロセスに繋げる	<ul style="list-style-type: none"> 稼働データ 顧客データ サービスデータ 法令・ガイドライン
サービス提供	顧客のリクエストを反映したサービス内容の提案	蓄積されたレシピデータを基に、顧客が食べたい食材・料理のジャンルやイメージ (休日のランチ風・日本酒に合う等) のリクエストに最もフィットするレシピを提案する	<ul style="list-style-type: none"> 稼働データ 顧客データ サービスデータ トレンドデータ 法令・ガイドライン
	顧客質問への自動応対支援	顧客から受け付けるあらゆる質問に自動で回答する	<ul style="list-style-type: none"> 稼働データ 顧客データ サービスデータ 法令・ガイドライン
フォローアップ	顧客フィードバックに基づくリピーター・新規顧客獲得支援	これまでの購買履歴、フィードバックをもとに次回に向けたサービスの提案を行う。また、顧客の属性や購入履歴をもとに新規顧客へのアプローチの仕方をサービス企業側に提案する	<ul style="list-style-type: none"> 顧客データ サービスデータ トレンドデータ

5 必要なリソースの特定 6 リソースの提供者・主体

金融×オペレーション/CSサマリー

ユースケース 特定

金融×オペレーション/CSのバリューチェーンは窓口業務-顧客管理で、各バリューチェーン毎にユースケースが存在

- 窓口業務: 患者のトリアージ (優先順位付け) による治療の最適化 等
- 関連事務: 顧客対応内容のレポート作成、社内問い合わせの効率化 等
- 顧客管理: 顧客情報に基づくリスク算出、詐欺の検知 等

データ特定

各ユースケースに必要なデータは、大きく顧客情報、商品・サービス情報、金融情報、過去の不正金融取引情報、規則・ルールに分かれる

- 窓口業務: 顧客情報、商品・サービス情報、金融情報、規制・ガイドライン
- 関連事務: 顧客情報、商品・サービス情報、レポートフォーマット 等
- 顧客管理: 顧客情報、商品・サービス情報、過去の不正金融取引情報、規制・ガイドライン 等

データ提供者・ 主体者

データ毎に提供者・主体および提供方法が異なる

- 顧客情報
 - 各金融機関に蓄積されており、非公開
 - 情報の機密性が高いことに加え、情報の可用性が低くデータの整備も必要になる情報も含まれるため取得難易度が高い
- 商品・サービス情報
 - 各金融機関が有している情報ではあるものの、一般的に公開されている
- 金融情報
 - 金融情報は政府や公的機関から公開されているものの、民間事業者が公開しているものは有料
- 過去の不正金融取引情報
 - 一般的に公的組織から公開されている
 - ただし、金融事業者が事業者内に蓄積されている情報は非公開
- 規則・ガイドライン
 - 既に公開されている

データ確保の 方策

データ提供方法毎に、確保の方策は以下の通り

- 既に公開されているデータ: 規制・ガイドライン、金融情報等、公開済みのデータも多く、開発事業者からの要望に応じて事務局によるマッチングを進める
- 有資格者のみ等、利用制限がされているデータ: 信用情報等、当該機関に対して利用許諾の緩和を求める
- 有償データ: 経済市場分析情報など、有償のものについては、公開情報で補えないものに関しては購入補助を検討する
- 非公開情報: 金融機関内に存在する機密性が高い情報は、業界団体への働きかけ/コンソーシアムの組成/匿名化技術の開発支援を検討する

金融×オペレーション/CSのバリューチェーンは窓口業務～顧客管理に分かれ、想定されるユースケースは以下の通り

金融×オペレーション/CSのユースケース

ユース
ケース

窓口業務

- 顧客対応の自動化
- 顧客への商品・サービス提案策定支援

関連事務

- 顧客対応内容のレポート作成
- 社内問い合わせの効率化
- 書類ドラフト作成

顧客管理

- 顧客情報に基づくリスク算出
- 規制・契約遵守状況のモニタリング
- 詐欺の検知

5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、大きく顧客情報、商品・サービス情報、金融情報、過去の不正金融取引情報、規制・ガイドラインに分かれる

必要なリソースの特定(1/2)

オペレーション/CSにおけるユースケース

ユースケースの概要

モデル構築に必要な学習データ

顧客対応の自動化

生成AIが顧客からの社内外からの問い合わせに対して自動で応答する

- 顧客情報
(属性、金融・取引・信用情報、窓口対応履歴 等)
- 規制・ガイドライン
- 商品・サービス情報

顧客への商品・サービス提案策定支援

生成AIが顧客情報に基づき、顧客の関心がありそうな商品・サービスについてレコメンデーションを作成する

- 顧客情報
(属性、金融・取引・信用情報、窓口対応履歴 等)
- 規制・ガイドライン
- 商品・サービス情報
- 金融情報

顧客対応内容のレポート作成

生成AIが顧客対応時の会話記録をもとに、顧客の関心・取引内容・契約内容等について自動でレポートを作成する

- 顧客情報
(属性、金融・取引・信用情報、窓口対応履歴 等)
- 規制・ガイドライン
- 商品・サービス情報
- レポートフォーマット

窓口対応

関連事務

5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、大きく顧客情報、商品・サービス情報、金融情報、過去の不正金融取引情報、規則・ガイドラインに分かれる

必要なリソースの特定(2/2)

オペレーション/CSにおけるユースケース

ユースケースの概要

モデル構築に必要な学習データ

顧客情報に基づくリスク算出

生成AIが顧客から取得した情報に基づき、顧客へ商品・サービス提供を行った場合のリスクレベルを算出する

- 顧客情報 (属性、金融・取引・信用情報、窓口対応履歴 等)
- 商品・サービス情報
- 金融情報
- 規制・ガイドライン

規制・契約遵守状況のモニタリング

生成AIが規制や契約条件に従い、顧客の取引状況において違反や不正取引がないかどうかを確認する

- 顧客情報 (属性、金融・取引・信用情報、窓口対応履歴 等)
- 過去の不正金融取引情報
- 商品・サービス情報
- 規制・ガイドライン

詐欺の検知

生成AIが顧客の取引状況や購買パターンを分析し、詐欺を検知する

顧客管理

5 必要なリソースの特定 6 リソースの提供者・主体

バックオフィス (法務) サマリー

ユースケース 特定

法務のユースケースは法務の主業務である、書類作成・審査、契約管理・監視、トラブル解決に紐づく形で存在

- 書類作成・審査: 書類の作成支援、書類内容の審査支援、契約書作成に必要なデータ整理支援
- 契約管理・監視: 法律リスクの判断支援、知的財産権侵害の判断支援
- トラブル解決: 法的トラブルへの対応策検討支援

データ特定

各ユースケースに必要なデータは、書類データ、業務データ、取引データ、トラブルデータ、知財データ、法令・ガイドラインに分かれる

- 書類作成・審査: 書類データ、取引データ、トラブルデータ、法令・ガイドライン
- 契約管理・監視: 業務データ、取引データ、トラブルデータ、知財データ、法令・ガイドライン
- トラブル解決: 業務データ、取引データ、トラブルデータ、法令・ガイドライン

データ 提供者・ 主体者

公開データが多数存在するものの、実際の取引を行う個社に非公開の契約データ等が蓄積している

- 書類データ (内容)、業務データ、取引データ: 各企業に蓄積されており、基本的には非公開
- トラブルデータ: 過去の判例等公開済みのデータは一定存在、ただし公になっていない
- 法令・ガイドライン、知財データ、書類データ (フォーマット):
既に公開されている、ただし個社または一部業界のガイドラインは非公開や有資格者のみに公開されているものも存在

データ確保の 方策

基本的に公開データをベースに開発推進できるものの、各企業の非公開データの活用を行うために、データ共有を行うための企業コンソーシアム構築を事務局として支援すべきか

- 公開データ: 事務局が必要なデータを整理してマッチング
- 有償データ: 公開データで補えないものは、事務局として購入支援を検討
- 非公開データ: データの匿名化支援、データ化・データ整理の支援を行ったうえで、非公開データを共有しあう企業コンソーシアムを作り、開発を促進する

バックオフィス (法務)における業務は3つに分かれ、想定されるユースケースは以下の通り

バックオフィス (法務) のユースケース

	書類作成・審査	契約管理・監視	トラブル解決
業務内容	取引契約書や登記書類など業務に係る様々な書類の作成や審査を行う	会社の法律遵守状況や、社員の社内規定やコンプライアンス遵守の監視を行うとともに、会社が保有するあらゆる権益の保護を行う	社内外にて発生する係争対応に加え、経営者や事業部から寄せられる法的相談に対応する
ユースケース	<ul style="list-style-type: none">• 書面の作成支援• 書類作成に必要なデータ整理支援• 書類内容の審査支援	<ul style="list-style-type: none">• 法律リスクの判断支援• 知的財産権侵害の判断支援	<ul style="list-style-type: none">• 法的トラブルへの対応支援

5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、書類データ、業務データ、取引データ、トラブルデータ、法令・ガイドラインに分かれる

必要なデータの特定

	ユースケース名	ユースケース概要	モデル構築に必要な学習データ
書類作成・審査	書面の作成支援	テンプレートの提案や、修正提案などを行い、契約書をはじめとする書類作成の効率化を実現	<ul style="list-style-type: none"> 書類データ 取引データ トラブルデータ 知財データ
	書類作成に必要なデータ提供	書類作成時に確認すべき、又は参照できる法的文書のレコメンデーションを行い、最短時間での文献確認を支援	<ul style="list-style-type: none"> 書類データ 取引データ 法令・ガイドライン
	書類内容の審査支援	契約先との関係で自社が有利/不利になる点のやトラブルの回避のために注視すべきポイントを指摘	
監視	法律リスクの判断支援	新規事業における法的側面からのリスク判断を生成AIが実施し、法務業務の削減を支援	<ul style="list-style-type: none"> 業務データ 取引データ トラブルデータ 法令・ガイドライン 知財データ
	知的財産権侵害や広告規制違反の判断支援	インターネットその他のデジタルプラットフォームにおける商標・著作物の使用状況を監視し、無許可の使用を自動的に検出することで知財実務を効率化	
トラブル対応	法的トラブルへの対応支援	法的トラブルの特徴を踏まえて、これまでに起きた事案を基に対処策の提案を実施	<ul style="list-style-type: none"> 業務データ 取引データ トラブルデータ 法令・ガイドライン

参考) 法務エキスパート(元リーガルテック企業)へのインタビュー



インタビューサマリー

1 有望な ユースケース

生成AIによる書類作成業務の支援は需要が高いと想定

- 書類作成業務における生成AI導入の需要・効果は大きい
 - 特に規模が小さい会社は法務部門がない場合もある
 - 大きい会社では取引や契約が頻繁に発生している

2 生成AI導入 における 障壁

企業内の法務業務への適応には民間企業が保有するデータの取得が必要だが、機密保持や未データ化の観点からデータ取得には課題が存在

- 精度の高い生成AIモデルの構築には、各社が保有する非公開データが重要
- だが、契約書類や契約に至るまでの社内での検討過程が記された稟議書類などは機密性も高く、データの外部提供には大きな懸念が存在
- 外部提供にあたり必要となる紙で保管されている書類のデータ化や、匿名化の作業は企業にとって時間やコストがかかるため、取り組むインセンティブはあまり高くない状況
 - 他の事業や業務との比較の上で、法務における生成AI導入に対してモチベーションが低い企業も存在

5 必要なリソースの特定 6 リソースの提供者・主体

医療×オペレーション/CSサマリー

ユースケース 特定

- 医療×オペレーション/CSのバリューチェーンは受け入れ準備～フォローアップ&モニタリングで、各バリューチェーン毎にユースケースが存在
- 受入準備：患者のトリアージ(優先順位付け)による治療の最適化 等
 - 診断：医用画像の診断支援
 - 治療/処方：個別化された治療方針・計画の作成
 - 医療事務：請求処理業務の効率化
 - フォローアップ&モニタリング：追加治療計画の作成 等

データ特定

- 各ユースケースに必要なデータは、大きく診断治療支援情報、患者情報、病院機能情報、規則・ルールに分かれる
- 受入準備：診断治療支援情報、病院機能情報 等
 - 診断：診断治療支援情報、患者情報 等
 - 治療・薬剤処方：診断治療支援情報、患者情報
 - 医療事務：規則・ルール、その他(請求書フォーマット 等)
 - フォローアップ&モニタリング：患者情報、診断治療支援情報、病院機能情報 等

データ提供者・ 主体者

- データ毎に提供者・主体および提供方法が異なる
- 診断治療支援情報
 - 各学会が策定する診療ガイドライン・研究論文等多くは公開されている一方で、医学教科書は未データ化、一部データは有資格者のみに限定
 - 患者情報
 - レセプト・診断画像・検診等多くは公開されている一方で、電子カルテは現状各病院に分散、データベース化されているものも有料、利用許諾条件あり
 - 病院機能情報
 - 医療機能情報は既に公開されている一方、病院稼働情報は有資格者のみに限定
 - 規則・ルール
 - 既に公開されている

データ確保の 方策

- データ提供方法毎に、確保の方策は以下の通り
- 既に公開されているデータ：各学会が策定する診療ガイドライン等公開済みのデータも多く、開発事業者からの要望に応じて事務局によるマッチングを進める
 - 有資格者のみ等、利用制限がされているデータ：臨床試験情報、医用画像等、当該機関に対して基盤モデル開発における利用許諾を求める
 - 有償のデータ：電子カルテ等、基盤モデル開発における購入補助を検討する
 - 未データ化、もしくはデータが分散されているデータ：医学教科書、電子カルテ等、紙データのデジタル化等、データ整備を進める

5 必要なリソースの特定 6 リソースの提供者・主体

医療×オペレーション/CSのバリューチェーンは受け入れ準備～フォローアップ&モニタリングに分かれ、想定されるユースケースは以下の通り

医療×オペレーション/CSのユースケース



5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、大きく診断治療支援情報、患者情報、病院機能情報、規則・ルールに分かれる

必要なデータの特定 (1/2)

ユースケース名	ユースケース概要	モデル構築に必要なデータ種別
受け入れ準備	患者のトリアージ(優先順位付け)による治療の最適化	<ul style="list-style-type: none"> 診断治療支援情報(診療ガイドライン・教科書等) 病院機能情報 患者情報(過去の症状・診断・治療情報)
	患者紹介の最適化	
診断	患者の問診・診断支援	<ul style="list-style-type: none"> 診断治療支援情報(診療ガイドライン・教科書等) 患者情報(過去の症状・診断・治療情報) その他(音声データ)
	医用画像の診断支援	
治療・薬剤処方	個別化された治療方針・計画の策定	<ul style="list-style-type: none"> 診断治療支援情報(診療ガイドライン・教科書等) 患者情報(過去の症状・診断・治療情報) 病院機能情報 その他(音声データ)
	処方監査・処方提案支援	

5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、大きく診断治療支援情報、患者情報、病院機能情報、規則・ルールに分かれる

必要なデータの特定 (2/2)

ユースケース名	ユースケース概要	モデル構築に必要な学習データ	
医療事務	メディカルコーディング (記録用症状分類)	生成AIが公的に定められた医療分類に従い、患者の症状を最も反映している医療分類コードを割り振る	<ul style="list-style-type: none"> 規則・ルール 患者情報(過去の症状・診断・治療情報)
	請求処理業務の効率化	生成AIが決められた診療報酬・保険取り決め等に従い、患者が病院で受けた処置・治療・処方に対して支払うべき金額及び保険会社等への請求金額について請求書を作成する	<ul style="list-style-type: none"> 規則・ルール その他(各種請求フォーマット) 患者情報(過去の症状・診断・治療情報)
フォローアップ・モニタリング	追加治療計画の策定	患者の症状が想定より改善していない場合や治療方針に変更が生じた場合、生成AIがこれまでの治療や患者の症状等に基づいて追加の治療計画を策定	<ul style="list-style-type: none"> 患者情報(過去の症状・診断・治療情報) 診断治療支援情報(診療ガイドライン・教科書等) 病院機能情報
	遠隔患者のモニタリング	生成AIが患者の血圧や心拍数などに基づき、患者の病態の急変や異常検知を行う	<ul style="list-style-type: none"> 患者情報(過去の症状・診断・治療情報) 診断治療支援情報(診療ガイドライン・教科書等) その他(PHR)
	薬歴・服薬状況のモニタリング	生成AIが薬剤師と患者の薬歴・服薬に関する会話を分析・蓄積し、自動で薬歴や服薬指導内容・服薬による症状の改善等のモニタリング・レポート作成を行う	<ul style="list-style-type: none"> 患者情報(過去の症状・診断・治療情報) 診断治療支援情報(診療ガイドライン・教科書等) その他(音声データ)

5 必要なリソースの特定 6 リソースの提供者・主体

日本においては、診断内容・結果にいたる過程・判断理由は電子カルテ上に記録され、その多くはデータ化・蓄積されていない

電子カルテ・レセプトデータの違い

	電子カルテ	レセプトデータ
目的	診療内容を記録し保管する	保険診療報酬の請求をおこなう
使用者	医師・看護師・検査技師・薬剤師など実際に医療を行う人が中心	会計情報を管理する医療事務担当者や会計士
データ内容	<p>診断内容と診断結果にいたる過程・判断理由を含む</p> <ul style="list-style-type: none"> 主訴 (診療時に患者が訴える症状) <ul style="list-style-type: none"> 発症時期や程度など 診察・検査結果 <ul style="list-style-type: none"> 聴診・触診・視診・内診の所見、血液検査や画像検査の情報など 診断結果 (主訴と診察・検査結果を踏まえて総合的に分析・考察をして導いた評価) 治療の方針・内容 <ul style="list-style-type: none"> 薬剤の処方内容、内服終了後の受診指示 	<p>保険診療報酬に関わる情報のみ</p> <ul style="list-style-type: none"> 診療・検査・治療の内容 <ul style="list-style-type: none"> 診療開始日 診療日数 病名 投薬内容 検査・手術内容 医療機関の収益 <ul style="list-style-type: none"> 請求点数 (1点 = 10円) 医師の指導料・管理料
普及率	50% ¹	96% ²



より精度の高い生成AIのユースケースを実現するには診断の過程が見える電子カルテ情報も求められる

1. 令和2年医療施設調査 (厚労省)。一般病院と一般診療所の数から普及状況を加重平均して算出
2. 令和5年4月請求分：社会保険診療報酬基金

5 必要なリソースの特定 6 リソースの提供者・主体

アニメーション産業サマリー

ユースケース 特定

アニメ産業のユースケースは企画、動画作成、編集・音響のバリューチェーンに沿って存在する

- 企画: プロットのアイデア出し、初校の作成、絵コンテの作画支援
- 動画作成: 画像生成AIを用いた背景画像の作成、線画への自動彩色
- 編集・音響: ビデオの自動翻訳と吹き替え、コンテンツのローカライズ最適化

データ特定

各ユースケースに必要なデータはシナリオデータ、画像データ、ラベリングデータ、色彩データ、映像データ、音響データ、CGデータ、トレンドデータ、法令・ガイドラインと多岐にわたる

- 企画: シナリオデータ、トレンドデータ、法令・ガイドライン
- 動画作成: シナリオデータ、画像データ、ラベリングデータ、色彩データ
- 編集・音響: シナリオデータ、映像データ、音響データ、画像データ、CGデータ、ラベリングデータ、法令・ガイドライン、トレンドデータ

データ 提供者・ 主体者

非公開データの保有先は製作委員会、製作会社、配信会社があるが、データの提供先としては製作会社と配信会社が有望

- シナリオ、画像、映像、CG、音響データ: 有償のものもあるが、基本的に非公開データまたは非公開かつ未データ化の状態、著作権元が保有している
- ラベリングデータ: 製作過程に出された指示情報としてのラベリングデータは著作権元に未データ化かつ非公開な状態で存在する
- 色彩データ、法令・ガイドライン: 既に公開のものが多い一方、各国の作品に対する規制は明文化されていないものも一定存在
- トレンドデータ: 既に公開されているものの、一部民間企業が提供する情報は有償である

データ確保の 方策

製作委員会では著作権を複数社が共同保有しているため使用許諾の難易度が高く、著作権を個社保有する製作・配信会社からのデータ確保が現実的か

- 公開情報: 事務局が必要な情報を整理・一覧化してマッチング
- 有償データ: 公開情報で補えないものは、事務局として購入支援を検討
- 未データ化: 事務局としてデータ化の支援を検討
- 非公開情報: 事務局として著作権を有する製作・配信会社の連携を支援し、未データ化の対応等必要に応じてコスト支援を行う

5 必要なリソースの特定 6 リソースの提供者・主体

アニメーションのバリューチェーンは企画～編集・音響までに分かれ、想定されるユースケースは以下の通り

アニメーションのユースケース



ユース
ケース

- プロットのアイデア出し
- 初校の作成
- 絵コンテの作画支援

- 画像生成AIを用いた背景画像の作成
- 線画への自動彩色

- ビデオの自動翻訳と吹き替え
- コンテンツのローカライズ最適化

5 必要なリソースの特定 6 リソースの提供者・主体

各ユースケースに必要なデータは、シナリオ、画像、映像、CG、音響、ラベリング、色彩、トレンドなどのデータ、法令・ガイドライン、と多岐にわたる 必要なデータの特定

	ユースケース名	ユースケース概要	モデル構築に必要な学習データ
企画	プロットアイデア出し	作品のストーリーを決めるためのアイデア出しにおいて様々なアイデアを提供して効率的なストーリー設計を支援する	<ul style="list-style-type: none"> シナリオデータ トレンドデータ 法令・ガイドライン
	初校の作成	プロットアイデア出しで集まったアイデアをもとにストーリーの組み立てを支援する	
動画作成	絵コンテの作画支援	ストーリーラインに沿って絵コンテを作成する際のドラフトを生成し、業務を効率化する	<ul style="list-style-type: none"> シナリオデータ 画像データ ラベリングデータ
	画像生成AIを用いた背景画像の作成	生成AIに背景画像を作成させ、人はその修正のみを行うことで業務が効率化される	
	線画への自動彩色	線画への彩色を自動で行う	
編集・音響	ビデオの自動翻訳と吹き替え	キャラクターの表情が言語と連動した自動吹き替えが複数言語において可能	<ul style="list-style-type: none"> シナリオデータ 映像データ 音響データ
	コンテンツのローカライズ最適化	文化的、宗教的に適切ではない表現やアニメ上の登場物を自動的に削除、置換えることを支援をする	

5 必要なリソースの特定 6 リソースの提供者・主体

アニメーション産業における生成AI開発には著作権を有する製作会社と配信会社に対する働きかけが重要

非公開データの確保の方策

	製作委員会	製作会社	配信会社
組織概要	複数企業が出資して作品を作る形式で、著作権も複数社が共同管理している <ul style="list-style-type: none"> 著作権の扱いには複数社の協議が必要 	自主製作として作成した作品に関しては著作権を有している。ただし、作品数が限られるためデータ量としては限定的	基本的には配信を行う企業。オリジナル作品として著作権を有しているものも一部存在する。データ量としては限定的
組織例	<ul style="list-style-type: none"> エヴァンゲリオン製作委員会 等 	<ul style="list-style-type: none"> 株式会社プロダクション・アイジー 株式会社アニマ 等 	<ul style="list-style-type: none"> Netflix Hulu 等
著作権の持ち方	△ 製作委員会を組成する企業が共同保有している	○ 自主製作アニメ分は保有	○ オリジナル作品分は保有
データ量 (1社ごと)	✗ 個別作品ごとに委員会を設置 <ul style="list-style-type: none"> シリーズものの委員会は保有データ量は多いものの、著作権は複数社の共同保有 	△ 自主製作アニメ分は保有	△ オリジナル作品分は保有
データ確保の障壁	✗ 権利が複数企業に分割されており、データ活用には複数企業の許可が必要になる	△ データの活用を行うことに対する適正な対価を得られる仕組みが必要	

生成AI構築に際しては、データを単体で保有している製作会社または配信会社へのアプローチが必要

また、データ量確保のために複数企業のコンソーシアムを組めるように事務局として支援することが重要

5 必要なリソースの特定 6 リソースの提供者・主体

汎用基盤モデル開発に向け、日本語・英語では以下のコーパスが既に存在

汎用基盤モデル向け既存コーパス

日本語コーパス一覧

- 1 Japanese CC-100
- 2 Mc4 (Japanese C-4)
- 3 oscar (original_ja)
- 4 oscar (deduplicated_ja)
- 5 amazon_reviews_multi (ja)
- 6 BTSJ1000人日本語自然会話コーパス (2023年3月NCRB連動 完成版)
- 7 国語研日本語ウェブコーパス (NWJC)
- 8 現代日本語書き言葉均衡コーパス (BCCWJ)
- 9 日本語話し言葉コーパス (CSJ)
- 10 日本語歴史コーパス (CHJ)
- 11 日本語日常会話コーパス (CEJC)
- 12 多言語母語の日本語学習者横断コーパス (I-JAS)
- 13 近代語コーパス (CMJ) (太陽コーパス、近代女性雑誌コーパス、明六雑誌コーパス、国民之友コーパス)
- 14 昭和話し言葉コーパス (SSC)
- 15 名大会話コーパス (NUCC) (旧: 日本語自然会話書き起こしコーパス)
- 16 現日研・職場談話コーパス (CWPC)
- 17 日本語諸方言コーパス (COJADS)
- 18 中国語・韓国語母語の日本語学習者横断コーパス (C-JAS)
- 19 昭和・平成書き言葉コーパス (SHC)
- 20 日本経済新聞記事オープンコーパス
- 21 『筑波ウェブコーパス (TWC)
- 22 PheMT Dataset
- 23 HJDataset
- 24 Business Scene Dialogue (BSD) Dataset
- 25 MNIST Dataset
- 26 JaQuAD: Japanese Question Answering Dataset
- 27 LINE LM Corpus
- 28 The Japanese web corpus (jpWaC)
- 29 The Europarl corpus (Japanese)
- 30 Wikipedia

英語コーパス一覧

- 1 British National Corpus (BNC)
- 2 The Cambridge English Corpus (CAEC)
- 3 The Corpus of Contemporary American English (COCA)
- 4 iWeb: The Intelligent Web-based Corpus
- 5 The NOW corpus (News on the Web)
- 6 Global Web-Based English (GloWbE)
- 7 Wikipedia Corpus
- 8 Coronavirus Corpus
- 9 Corpus of Historical American English (COHA)
- 10 The TV Corpus
- 11 The Movie Corpus
- 12 Corpus of American Soap Operas
- 13 Hansard Corpus
- 14 Early English Books Online
- 15 Corpus of US Supreme Court Opinions
- 16 TIME Magazine Corpus
- 17 The Strathy Corpus of Canadian English
- 18 CORE Corpus
- 19 Translational English Corpus (TEC)
- 20 Corpus of Middle English Prose and Verse
- 21 Michigan Corpus of Academic Spoken English (MICASE)
- 22 The Scottish Corpus of Texts and Speech (SCOTS)
- 23 The Corpus of Modern Scottish Writing (CMSW)
- 24 The ACL Anthology Reference Corpus (ARC)
- 25 British Academic Written English Corpus (BAWE)
- 26 British Academic Spoken English corpus (BASE)
- 27 A Corpus of English Dialogues 1560-1760 (CED)
- 28 Dictionary of Old English Corpus in Electronic Form (DOEC)
- 29 The Old Bailey Corpus (OBC)

5 必要なリソースの特定 6 リソースの提供者・主体

既存のコーパスを十分に活用することで、少なくともGPT3レベルの汎用基盤モデルを開発することはデータ量として可能

既存コーパスのデータ量

GPT-3のレベル

データ量

- 570GB
(~3,000億トークン)

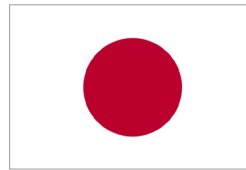
データソース

- インターネット上でオープンな情報を活用

精度 (パラメータ-数)

- 1,750億

既存のコーパス



日本語
コーパス

約2,500億トークン
(24の公開コーパス合計値)



英語
コーパス

約4,800億トークン ※現時点のリサーチ結果
(28の公開コーパス合計値)



オープンなデータの活用によって、GPT-3レベルの汎用基盤モデル自体の構築は可能
ただし、GPT-4レベルの開発に向けてはより多くのデータや計算資源の確保が必要か

note: デスクトップリサーチによってリスト化された公開コーパスの総語数に対してトークンへの変換値を設定して合計トークン数を算出。日本語においては変換値を0.5、英語に対しては4/3で設定。
なお、現時点のリサーチ結果に伴う現時点版のトークン量。

5 必要なリソースの特定 6 リソースの提供者・主体

加えて、NICTで準備中のデータおよび政府保有のデータを提供することで、さらにデータ量を増加させることが可能

汎用基盤モデル向け政府提供データ (案)

NICTで準備中のデータ

【NICTにおける学習用言語データの現状】

- データ量：350GB
 - 日本語に特化した400億パラメータの大規模言語モデル開発に使用したデータ
 - データ内容の詳細は非公開 (可能であればご共有頂きたい)
- データ提供方法：制限あり
 - 商用サービスでの提供は不可
 - 「共同研究」の形態を取る必要



【今後の方向性】

- データ量：3.5TB
- データ提供方法：制限緩和を検討中
 - 商用サービスでの提供許可
 - 「共同研究」の形態に依らない



その他、政府保有のデータ

政府が保有し、web上で公開されている等でデータ提供が可能そうなデータ

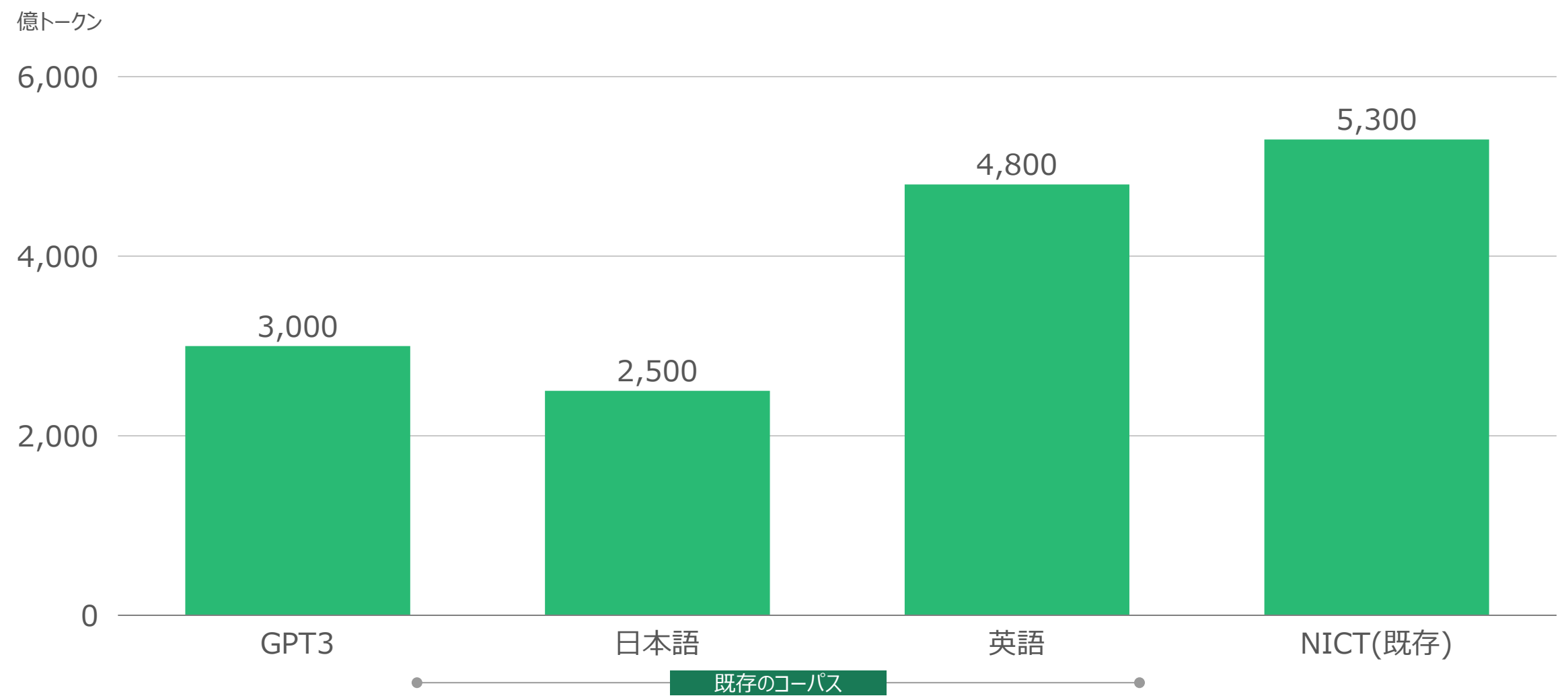
- 国会図書館
- 公文書館
- 国会議事録 等

参考：国会図書館が保有するテキストデータ (2023年6月9日 自民党 AI-PT資料より)

2-2. テキストデータの点数 (令和5年5月現在)

資料種別	年代・概要	インターネット公開		図書館・個人送信	館内限定	合計
		①保護期間満了	②許諾・裁定	③絶版等資料	④①～③以外	
図書	明治期～1968年までに受け入れた国内刊行図書の大部分 震災・災害関係資料の一部 (1968年以降に受け入れたものを含む。)	28万点 (*1)	8万点	55万点	7万点	97万点
雑誌	国内刊行雑誌のうち劣化した雑誌や学術雑誌等 (刊行後5年以上経過したもの。主に1990年代刊行分まで)	-	2万点	80万点	50万点	132万点
古典籍	貴重書・準貴重書、江戸期以前の和漢書等	(8万点) (*1)(*2)	-	-	-	(8万点)
博士論文	1990～2000年度に送付を受けた国内の博士論文	-	1万点	12万点	2万点	15万点
その他	官報等	2万点	-	-	0.5万点	2万点
合計		30万点	11万点	147万点	59万点	247万点

参考) 汎用基盤モデル向けデータ量の比較



⑦ リソースを確保するための方策

1サイクル目における提供に向け、各データの利用可否の確認、許諾取得を進める

汎用基盤モデル向けデータ提供方法

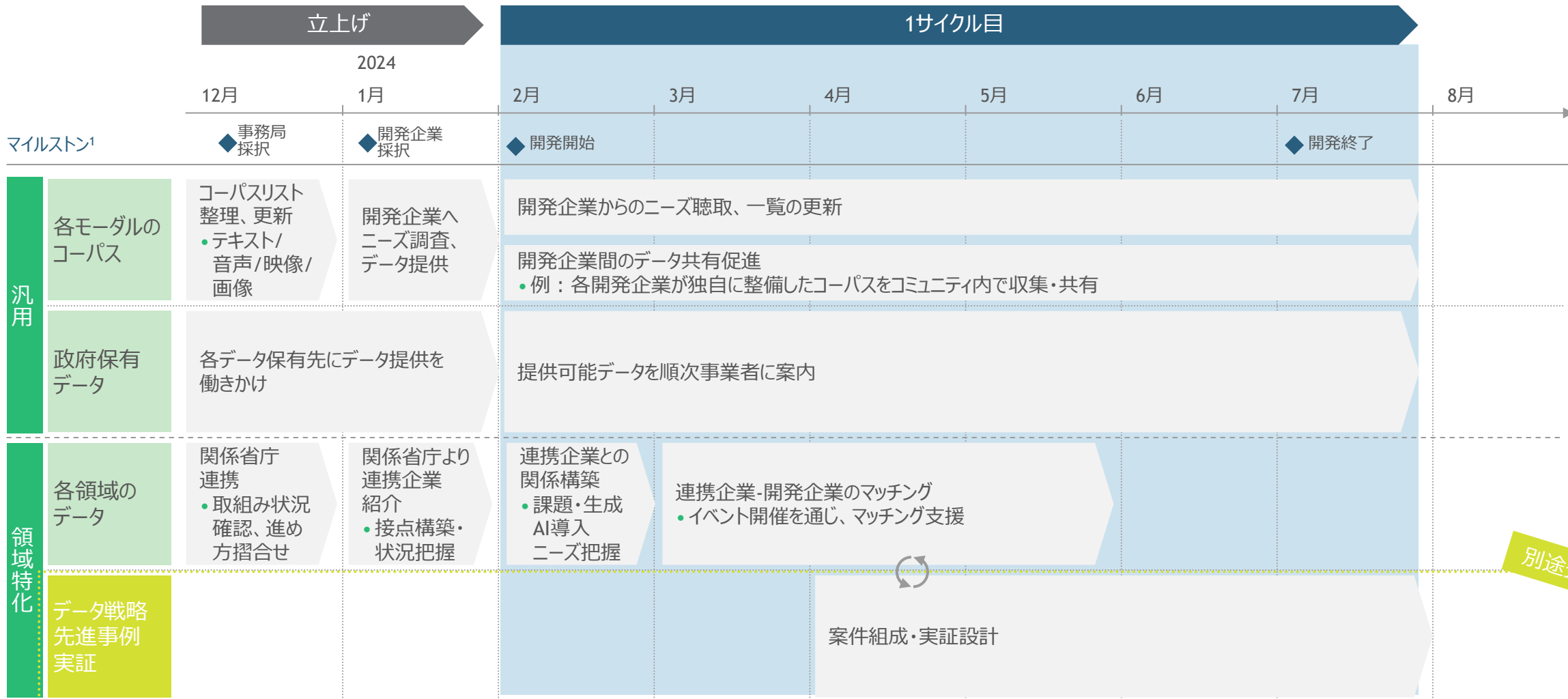
	現状	1サイクル目における提供方法
既存コーパス	<p>既存の公開コーパスが整理・一覧化されておらず、各企業毎に探索する必要がある</p>	<p>経産省・事務局が、整理・一覧化し、事業開始時に事務局から事業者以案内</p> <ul style="list-style-type: none"> 必ずしも提供した一覧にあるコーパスを使用する必要はなく、リストの活用は事業者の任意とする
NICTが準備中のデータ	<p>データ提供方法に制限あり</p> <ul style="list-style-type: none"> 商用サービスでの提供は不可 「共同研究」の形態を取る必要 	<p>少なくとも商業サービスでの提供は許可できるようNICTへ働きかけ</p> <ul style="list-style-type: none"> あわせて、「共同研究」の形態に依らないよう調整
政府として提供可能性のあるデータ	<p>自民党AI-PTにてインターネット公開されているテキスト・画像は言及されているものの、実際に開発に使えるのか不明</p>	<p>経産省・事務局から、本事業におけるデータ活用の可否および可能な場合のデータ取得方法を確認</p> <ul style="list-style-type: none"> Webクロールで良いのか 等

(2) 基盤モデルの開発に必要なリソースの特定・確保の方策

7 リソースを確保するための方策

1サイクル目では汎用向けのデータ整理と提供を行うとともに、2サイクル目に向けたデータの準備を推進

1サイクル目における事務局の役割



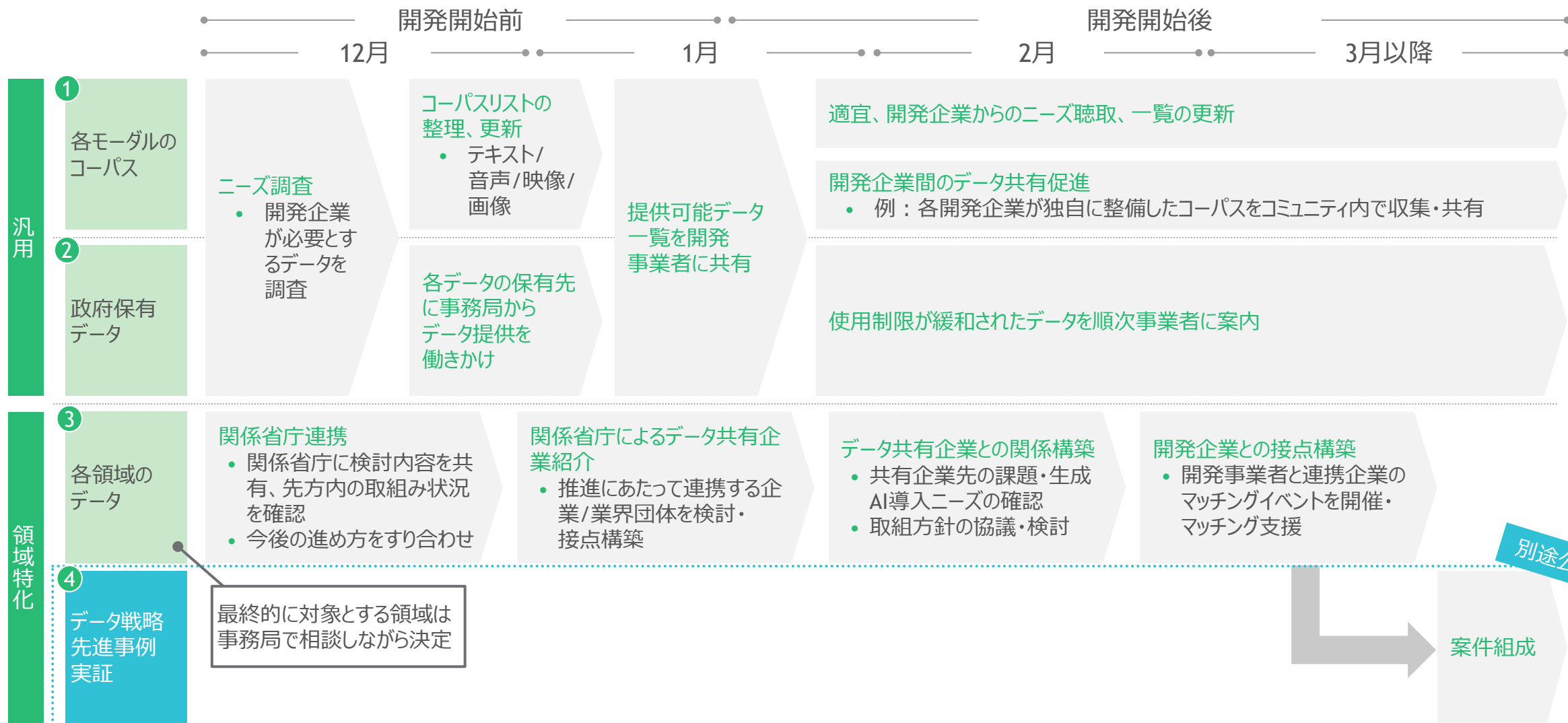
別途公募

(2) 基盤モデルの開発に必要なリソースの特定・確保の方策

7 リソースを確保するための方策

サイクル1では汎用向けのデータ整理と提供を行い、次サイクルに向けたデータの準備を推進

開発サイクル1における事務局の役割



(2) 基盤モデルの開発に必要なリソースの特定・確保の方策

7 リソースを確保するための方策

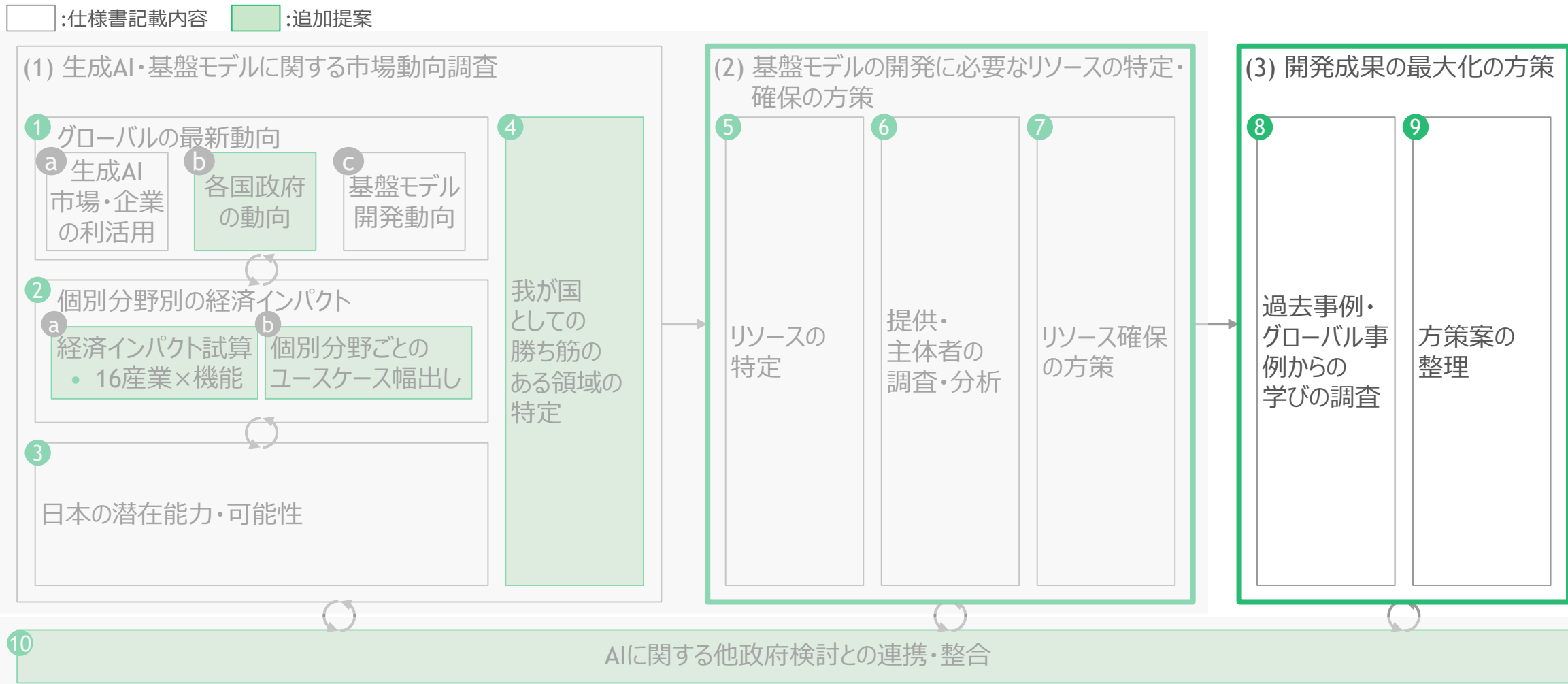
これまでの調査と議論を踏まえ、自動車、機械製造、アニメーションは経産省主体で取組む

リソース(データ)の特定・確保の方策サマリー

勝ち筋領域	① データの特定	② データの保有者	③ 確保の方策	④ 主体となる省庁
医療× オペレーション・CS	診断治療支援情報、患者情報、病院機能情報、規則・ルール <ul style="list-style-type: none"> 特に電子カルテが重要 	公開データが多いものの、有資格者のみが閲覧できる/ 非公開データも一定存在 <ul style="list-style-type: none"> 電子カルテも有資格者のみ(国)/有償(民間) 	関係省庁の動きを支援 <ul style="list-style-type: none"> 有資格者のみ使用可能/非公開データの開発における活用を推進 	厚労省
金融・保険× オペレーション・CS	顧客情報、商品・サービス情報、金融情報、不正金融取引情報、規則・ルール <ul style="list-style-type: none"> 特に顧客情報(ログを含む)が重要 	公開データも一定存在、ただし金融機関が非公開で有するデータも多い <ul style="list-style-type: none"> 機密性が高く、金融機関からの提供は難しい 	個社の取組みとし、業界連携時に支援	金融庁
サービス× オペレーション・CS	稼働情報、顧客情報(顧客ログも含む)、サービス情報、トレンド情報、法令・ガイドライン	多くは非公開情報で、各社が保有 <ul style="list-style-type: none"> 金融ほどではないが顧客の個人情報が含まれるため一定の機密性がある 	企業コンソーシアム構築等、十分なデータ量の確保に向けて事務局として支援 <ul style="list-style-type: none"> 非公開情報を保有する企業間連携を促進 	国交省/ 厚労省/観光庁
自動車× 製造・加工	製品設計情報、生産設備情報、検査情報、取引先情報、市場動向情報、規制・ガイドライン	多くは非公開情報で、各社が保有 <ul style="list-style-type: none"> 自動車企業または自動車部品企業が保有 非公開情報の中にも秘匿性の高低が存在 	企業コンソーシアム構築等、十分なデータ量の確保に向けて事務局として支援 <ul style="list-style-type: none"> 非公開情報を保有する企業間連携を促進 	経産省
製薬× 製造・加工	製品設計情報、生産設備情報、検査情報、取引先情報、規制・ガイドライン <ul style="list-style-type: none"> 特に製品設計情報(SOP)が重要 	多くは非公開情報で、各社が保有 <ul style="list-style-type: none"> 特に製薬会社やODM・OEMが保有しているSOPの秘匿性が高い 	個社の取組みとし、業界連携時に支援	厚労省
機械製造× 製造・加工	製品設計情報、生産設備情報、検査情報、取引先情報、市場動向情報、規制・ガイドライン	非公開情報として各社が保有 <ul style="list-style-type: none"> 各機械メーカーが製品や部品に関わる情報を個社が保有し散在している 	保有データ量が多いミスミのような機械部品会社に協力を求め、業界横断での取組みを支援	経産省
アニメーション	シナリオ/画像/ラベリング/色彩/映像/音響/CG/トレンドなどのデータ、法令・ガイドライン	多くは非公開情報で、各社が保有 <ul style="list-style-type: none"> 製作委員会、製作会社、配信会社が保有、データ提供先として製作会社と配信会社が有望 	企業コンソーシアム構築等、十分なデータ量の確保に向けて事務局として支援 <ul style="list-style-type: none"> 著作権を保有する製作/配信会社との連携 	経産省
バックオフィス(法務)	書類情報、業務情報、取引情報、トラブル情報、知財情報、法令・ガイドライン	公開データが多いものの、企業法務に適應させるためのデータは非公開情報 <ul style="list-style-type: none"> 個社に分散かつ未データ化のものも多い 	企業コンソーシアム構築等、十分なデータ量の確保に向けて事務局として支援 <ul style="list-style-type: none"> 非公開情報を保有する企業間連携を促進 	法務省

(3) 開発成果の最大化の方策

事業実施の基本方針 : (3) 開発成果の最大化の方策



⑧ 過去事例・グローバル事例からの学びの調査

8 過去事例・グローバル事例からの学びの調査

様々なステークホルダーを巻き込み、開発支援・知見共有・社会実装を行う必要のある本事業の成果最大化の方策を過去・グローバル事例から調査

調査設計

調査目的

様々なステークホルダーを巻き込み、推進する必要のある本事業の成果最大化の方策を過去・グローバル事例から明らかにする

- 開発支援:開発企業に加え、産業、様々専門性を持った企業等と連携する必要
- 知見共有:国内のみではなく、グローバル企業とも連携する必要
- 社会実装:規制・ルール含め、社会実装に障壁となるものを緩和・取り除く必要

調査論点

開発支援

- ① どのようなスキームで基盤モデル開発支援事業を行うか
- ② どのように様々な専門性を持った企業等と効率的・効果的に連携するか
- ③ どのような企業を対象に、開発企業候補を募集するか
- ④ どのような評価観点・方法で採択企業を選定・評価すべきか

知見共有

- ⑤ どのように開発成果の知見を共有し、開発能力の向上に寄与させるか

社会実装

- ⑥ 社会実装・活用に向けて何が障壁となるか、その推進に向け何が必要か

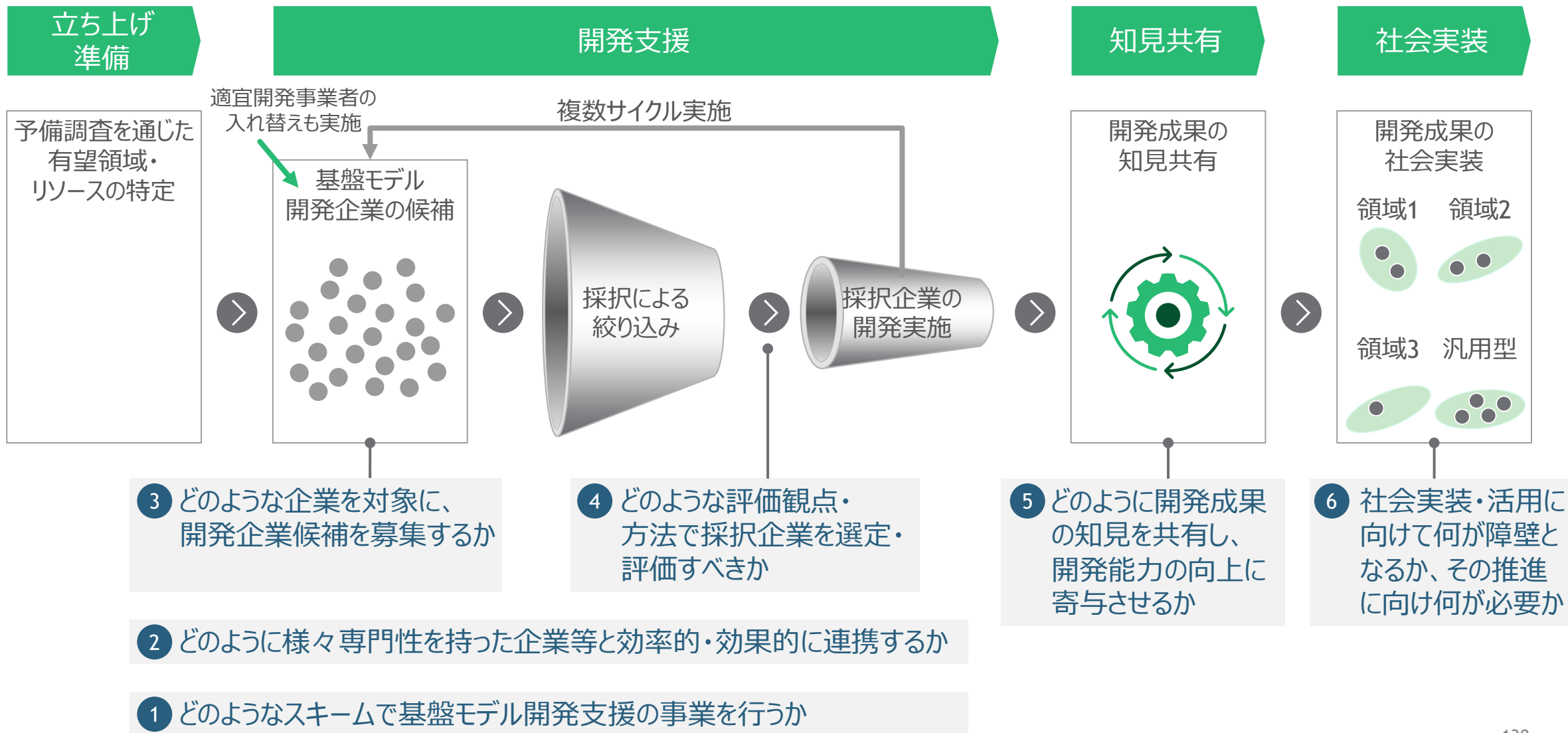
調査対象

企業等と連携し、国家レベルで推進を行った技術関連の海外・国内の取組み・プロジェクト

〈海外〉
米・欧・中・独・韓の5カ国

〈国内〉
過去の国家プロジェクト

開発成果を最大化するために解くべき論点



(3) 開発成果の最大化の方策

8 過去事例・グローバル事例からの学びの調査

海外9件、国内8件の国家プロジェクトを調査

国	プロジェクト名	実施年	実施機関	概要
United States	Semiconductor Industry Promotion project by NDAA Chips Act 2022	2021-2022	全米半導体技術センター(官民コンソーシアム)	米国内で半導体を生産するための有益なインセンティブを設けることにより、国内の半導体製造を促進するための立法に向けた動き。国内の製造施設や半導体サプライチェーン企業の建設、拡大、設備増強に財政的インセンティブを提供することなどが盛り込まれた。
China	China/AI30 Promotion of semiconductor technology through large funds, etc.	2020-2030	中国集積回路産業投資基金	中国集積回路産業投資基金(通称ビッグファンド)が監督する3つのファンドの中の一つ。その中でも最大規模のもので目標額の3000億元(410億ドル)は、2014年と2019年の同様のファンドによる調達額を上回っており、投資の主な重点はチップ製造設備の獲得に置かれている。政府は国内半導体市場の再構築を積極的に試みており、地元企業が市場競争力を高めることを目的にしている。
	Ernie Bot Tongyi Qianwen	2023 (製品公開) 2023 (製品公開)	中国政府	中国内の検閲を乗り越えたAIチャットボット、ChatGPTの中国版の開発事業。 機械学習を利用し、英語と中国語の両方でユーザーのプロンプトに応答するAIチャットボットの開発事業
EU	EU/ Horizon 2020 Horizon Europe (21-27)	2014-2020 2021-2027	欧州委員会	スマートな成長、持続可能な成長、包括的な成長という3つのプライオリティのもとで様々な事業が生まれ、Horizon Europeという後続の政策に引き継がれる形で継続
Germany	Industry 4.0	2014-2024	連邦経済エネルギー省 等	プラットフォームの設置による企業間の協力や促進、研究開発への投資や標準化活動などを通して技術開発のための基盤が構築された。2027年までに約900億から1500億ユーロの生産性向上による利益が予測されているほか、製造業におけるデジタル化に関して様々なユースケースを生んでいる
	Hydrogen Strategy	2020-	連邦政府	電解水素製造企業への直接補助金と免税措置、ドイツにおける水素技術の促進と開発、必要なインフラの強化とさらなる開発、グリーン水素製造プラントの建設と運営を含む国際的パートナーシップの拡大などを実施している
韓国	人工知能時代の安全な個人情報活用政策方向	-	個人情報保護委員会	事業者にとってのグレーゾーンを排除または規制緩和を行うことで、AIの開発を推進予定 AIプライバシーチームを新設することによって、AI開発事業者に対して法令解釈やサンドボックス適用の検討を行うなどのコンサルティング支援を行う予定。
日本	VLSI	1976-1979	通産省	半導体の世界市場での日本シェアは、28% (1975年) から52% (1988年) に急上昇 成果が活用された製品 (半導体・コンピュータ) の売り上げは1983年からの5年間で2.2兆円超
	科学技術用高速計算システム (スーパーコンピュータ)	1981-1989	通産省	プロジェクトの目標である10GFLOPSの性能を持つ高速コンピュータを実現。 一方で、プロジェクト完了時に10GFLOPSは商用機で実現できる水準であったため、 目標設定・実施手法に課題
	第5世代コンピュータ	1982-1992	通産省	並列推論マシンのモデルと、専用オペレーティングシステムの構築など技術面/人材育成の観点で 一定の成果はあったが、実務的な活用においては課題が存在
	Σ計画	1985-1990	通産省	将来のプログラマー不足をカバーする「ソフトウェア生産工業化システム」を目指したが、 ソフト開発ツール・専用コンピュータの実用性に課題
	リアルワールドコンピューティング	1992-2001	通産省	情報統合と並列分散シームレスコンピューティングという情報処理技術の新たな流れを作ること寄与。 プロジェクト途中で研究成果の実用化・利用が進められたテーマも存在
	半導体MIRAI	2001-2010	経産省	次世代マスク基盤技術、EUV光源高信頼化技術の基盤技術を開発 上市5、製品化3、研究開発継続15、中止中断29 (2012年時点) テーマ設定における目標は達成したが、集積化技術や産業化に向けての展開は課題
	情報大航海	2007-2010	経産省	情報プライバシーと利便性の問題を国民レベルで考えるきっかけとなったが、実用化においては課題
	量子技術イノベーション戦略 /量子未来産業創出戦略	2020-現在	内閣府	(実施中: 基礎研究から社会実装まで産学官で一体的に推進)

(3) 開発成果の最大化の方策

8 過去事例・グローバル事例からの学びの調査

論点に対するサマリー

論点	典型的なチャレンジ・課題	対応策の仮説	対象事例
① どのようなスキームで基盤モデル開発支援事業を行うか	<ul style="list-style-type: none"> 国が方針を定めすぎたことで、各企業の創意工夫が削がれる 計画変更の柔軟性が欠如 産業構造の変化に追いつけず、成果を実用化に活かせない 	<ul style="list-style-type: none"> 変化の激しい開発環境に応じて柔軟に方針転換ができる体制が必要 周囲の環境変化に応じて開発環境を柔軟に見直すため、短期のサイクルで審査等を設けて開発事業を実施することが効果的 	National Hydrogen Strategy
② どのように様々な専門性を持つ企業と効率的・効果的に連携するか	<ul style="list-style-type: none"> 様々な専門性を持つ企業の連携は複雑かつ相互関係が存在するため一貫した管理が無いとサイロ化する 企業間の競争が事業の目的の達成の障害になる 	<ul style="list-style-type: none"> 中立的立場でプロジェクトの全体調整を行うことができるハブを設置し、事業目的に沿った事業運営を行うことが重要 	地域デジタル人材育成・確保推進事業
③ どのような企業を対象に、開発企業候補を募集するか	<ul style="list-style-type: none"> 日本国内の大企業やその技術に偏ることで、グローバルの市場動向や新技術に乗り遅れる 	<ul style="list-style-type: none"> 国内の主要企業だけでなく、特定領域に知見や技術のある国内外のスタートアップ・海外企業も候補に選定する 	Beijing Science Park
④ どのような評価観点・方法で採択企業を選定・評価すべきか	<ul style="list-style-type: none"> 産業分野側のニーズや具体的な活用方法が不明確なまま、技術的な新規性のみで採択されてしまう モチベーションの低い企業が形式的プロジェクトに参画し、必要最低限の取り組みしかなされない 	<ul style="list-style-type: none"> 技術的な手法や開発実績だけでなく、社会実装に向けた見立て・対応状況・期待リターンについても評価の対象とする 成果連動型の支援の枠組みやステージゲートでの評価の際、あらかじめ定量的な中間・最終目標への到達度に応じてインセンティブの増減などの検討 	Horizon 2020
⑤ どのように開発成果の知見を共有し、開発能力の向上に寄与させるか	<ul style="list-style-type: none"> 各社の競争関係にあるために連携が難しい 	<ul style="list-style-type: none"> テックコミュニティを醸成し、社会実装に向けて必要な要素を議論する場を整備する 	Industry 4.0
⑥ 社会実装・活用に向けて何が障壁となるか、その推進に向け何が必要か	<ul style="list-style-type: none"> 情報プライバシー・著作権等のルールや利便性に関する世間からの懸念の声・批判 開発企業が現状の規制ルールを踏まえてどこまで実用可能か不明・不確実 	<ul style="list-style-type: none"> 現行の規制ルールにおいて新たな技術が抵触するか否かを迅速に確認できるようにする 開発事業の法的リスクを減らす社会実装支援を行う 	個人情報保護委員会 (韓国)

⑨ 提言案の整理

これまでの検討を通じ、本事業の方針を整理

これまでの検討振り返り

これまでの検討振り返り

市場動向	グローバル最新動向	<ul style="list-style-type: none"> 汎用の大規模化のみではなく、領域特化・効率化・拡張にも開発は進展 グローバル企業も医療、SW開発、広告等領域特化の開発を進める
	経済インパクト	<ul style="list-style-type: none"> 産業×機能別に、コスト削減効果の大きい11領域、売上向上効果の有望2領域を特定 <ul style="list-style-type: none"> 特にコスト削減効果の大きい5,000億円以上の領域は9領域
	日本の潜在能力	<ul style="list-style-type: none"> グローバル企業に追いつくのは簡単ではないが、汎用の開発を行うことで開発能力を育成する必要 領域特化においては、日本が産業的国際競争力・データ・日本独自の文化を有する領域が有望
	我が国としての勝ち筋	<ul style="list-style-type: none"> 汎用モデルの位置づけを明確化 (開発を推進することで基盤モデル共通の開発能力を育成) ①グローバル最新動向、②経済インパクト、③日本の潜在能力を踏まえ、領域特化の勝ち筋8領域を特定
リソースの特定・確保	リソースの特定	(汎用) <ul style="list-style-type: none"> 汎用基盤モデル開発に必要なテキストデータを、既存コーパス・政府保有データで洗い出し 既存コーパス・政府保有データそれぞれに対して、1サイクル目に向けた提供方法を整理
	提供・主体者	(領域特化) <ul style="list-style-type: none"> 勝ち筋領域それぞれに対して、ユースケース、ユースケース実現に必要なデータを洗い出し 各データの保有者および提供方法を整理し、確保の方策を特定
	リソース確保	<ul style="list-style-type: none"> 非公開かつユースケース実現に重要なデータに関しては関連省庁と連携しデータホルダーに働きかけることで確保を進める
開発成果最大化	過去事例の学び	<ul style="list-style-type: none"> 国内外の国家プロジェクトを調査、開発成果最大化における学びを整理 開発スキーム、様々な専門性を持つ企業との連携方法、採択基準・方法、社会実装の促進方法等を抽出
	方策案の整理	<ul style="list-style-type: none"> 事業立ち上げに向けた各取組みの実施方針・内容を整理 <ul style="list-style-type: none"> GPU管理、性能評価、コミュニティ運営、对外発信 等

今後の検討内容

12月からの事業立ち上げに向け、各取組み内容をより具体化 (以下、例)

- GPU管理:
 - レポート項目・FMT
- 性能評価:
 - 評価実施方法
 - 具体タスク/データセット/評価方式
 - 生成系タスクの採用有無
- データ確保:
 - 1サイクル目での実施内容
- コミュニティ運営:
 - 採択企業・プラチナ会員・一般の3段階のコミュニティ設計

これまでの検討を踏まえ、本事業の運営において意識すべきことは以下の通り

本事業のチャレンジ

時間軸の制約	<ul style="list-style-type: none"> 2024年1月末からの開発開始に向け、計算リソースの有効活用のためにも、迅速に事業体制を整える必要
市場変化への対応の難しさ	<ul style="list-style-type: none"> 新規性が高く、日々状況が変化し続ける国内外の生成AI市場動向を踏まえ、効率的・効果的に基盤モデル開発を進める必要
異なる利害関係者の調整・交渉の難しさ	<ul style="list-style-type: none"> 基盤モデルの開発にあたって、競争関係にある開発企業間や計算リソース提供者、データホルダーとなる利用企業等、異なる利害関係者間での調整・交渉を行う必要
巻き込むべきステークホルダーの多さ	<ul style="list-style-type: none"> 社会実装の推進にあたって、グローバル企業、産業、専門性を持つ企業等、様々なステークホルダーを巻き込む必要
政策との整合性担保の難しさ	<ul style="list-style-type: none"> 異省庁間で進む生成AIに関する戦略・ガイドライン、各産業毎での産業政策までを一貫して整合性を担保、連携する必要
上記、広範な分野を一貫して進める難しさ	<ul style="list-style-type: none"> 広範な分野に跨る本事業を進めるにあたって、各分野に専門性を持つ企業と連携するとともに、一貫性を持って進める必要

本事業の事業運営において意識すべきこと

- 提案段階から各実施項目の実施内容・方法を具体的に定め、事業開始後に迅速に立ち上げ
- グローバルに生成AI市場動向を常に把握、生成AIに関する有力プレイヤーとも効果的に連携し、機動的に事業運営方法をアップデート
- それぞれの関係者のニーズ・課題を踏まえ、効果的な事業設計を行った上で調整・交渉を行い、効果的/効率的に開発を推進
- 各ステークホルダーのニーズ・課題を踏まえ、様々なステークホルダーを巻き込んだコミュニティ運営、対外発信を通じ、社会実装を推進
- 政府系機関との連携を図り、政府全体の取組との整合性確保や、各取組の相乗効果を促進
- 各分野に専門性を持つ企業と効率的・効果的に連携するとともに、プロジェクトの全体調整を行い、一貫性を担保



[bcg.com](https://www.bcg.com)