JaDHA
The Japan Digital Health Alliance

ヘルスケア領域において生成AIを活用したサービスを提供する事業者が参照するための自主ガイドライン（ヘルスケア生成AI活用ガイド）【概要版_第2.0版】

デジタルヘルスアプリの適切な選択と利活用を促す社会システム創造WG

生成AI活用ガイド第2.0版
改定のポイント

ヘルスケア事業者のための生成AI活用ガイド第2.0版_改正のポイント

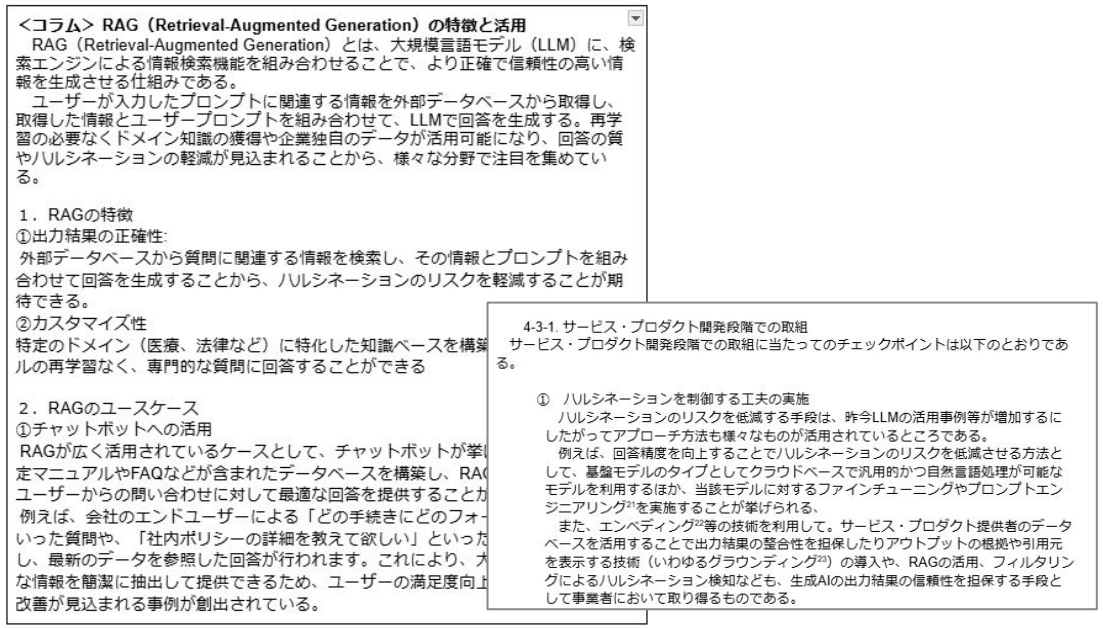
- 第1.0版策定（2024年1月）以降、生成AIを取り巻く技術環境や国内外の政策動向等は変化を続けているところ、それらを踏まえた本文のアップデートや参考資料の充実等を図ることで、本ガイドが業界内で実効性が高く有用性が担保されたリビングドキュメントとして活用が図られることを目的として第2.0版を策定することとした。
- 今般の改定のポイントは①技術動向/②政策動向/③海外動向/④活用動向を踏まえたアップデートの主に4点。

	直近の動き	改正概要
① 技術動向	RAG（検索拡張生成AI）やSLM（小規模言語モデル）、日本語特化型LLM開発等による技術の進化	● 最新の技術動向をコラム欄で解説、 ● チェックポイントのアップデート
② 政策動向	経済産業省・総務省策定の「AI事業者ガイドライン」、 個人情報保護法3年見直しなどAI関係の政策の動き	● 最新の政策動向を更新 ● チェックポイントのアップデート
③ 海外動向	EU・米国・中国・インド・韓国など諸外国で AIに関係する各国の政策・制度構築の動き	● 最新の海外動向を更新
④ 活用動向	国内外で生成AI活用事例が創出 生成AIサービスの活用機会が拡大	● 事例の拡充（4事例→13事例に） ● 院内ポリシーひな型の策定

ヘルスケア事業者のための生成AI活用ガイド第2.0版_改正のポイント（詳細）

① 技術動向 を踏まえた改定

- 「4. チェックポイント」において、第1.0版以降に進化した生成AIをめぐる技術動向をコラム形式で解説・アップデート
  - RAG（Retrieval-Augmented Generation：検索拡張生成AI）
  - SLM（Small Language Model：小規模言語モデル）
  - 日本語特化型LLM etc.



<コラム> RAG（Retrieval-Augmented Generation）の特徴と活用

RAG（Retrieval-Augmented Generation）とは、大規模言語モデル（LLM）に、検索エンジンによる情報検索機能を組み合わせることで、より正確で信頼性の高い情報を生成させる仕組みである。

ユーザーが入力したプロンプトに関連する情報を外部データベースから取得し、取得した情報とユーザープロンプトを組み合わせて、LLMで回答を生成する。再学習の必要なくドメイン知識の獲得や企業独自のデータが活用可能になり、回答の質やハルシネーションの軽減が見込まれることから、様々な分野で注目を集めている。

1．RAGの特徴

①出力結果の正確性:

外部データベースから質問に関連する情報を検索し、その情報とプロンプトを組み合わせて回答を生成することから、ハルシネーションのリスクを軽減することが期待できる。

②カスタマイズ性

特定のドメイン（医療、法律など）に特化した知識ベースを構築

ルの再学習なく、専門的な質問に回答することができる

2．RAGのユースケース

①チャットボットへの活用

RAGが広く活用されているケースとして、チャットボットが挙

定マニュアルやFAQなどが含まれたデータベースを構築し、RA

ユーザーからの問い合わせに対して最適な回答を提供することカ

例えば、会社のエンドユーザーによる「どの手続きにどのフォー

いった質問や、「社内ポリシーの詳細を教えて欲しい」といった

し、最新のデータを参照した回答が行われます。これにより、大

な情報を簡潔に抽出して提供できるため、ユーザーの満足度向上

改善が見込まれる事例が創出されている。

4-3-1. サービス・プロダクト開発段階での取組

サービス・プロダクト開発段階での取組に当たってのチェックポイントは以下のとおりである。

① ハルシネーションを制御する工夫の実施

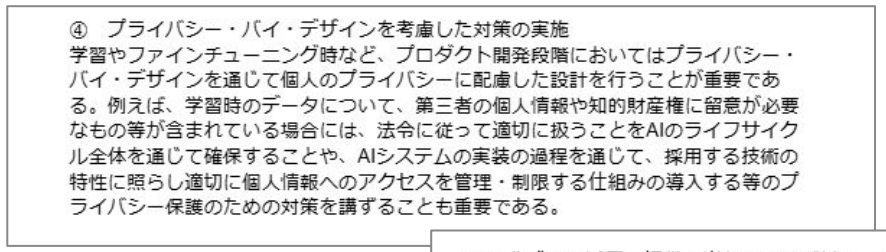
ハルシネーションのリスクを低減する手段は、昨今LLMの活用事例等が増加するにしたがってアプローチ方法も様々なものが活用されているところである。

例えば、回答精度を向上することでハルシネーションのリスクを低減させる方法として、基盤モデルのタイプとしてクラウドベースで汎用的かつ自然言語処理が可能なモデルを利用するほか、当該モデルに対するファインチューニングやプロンプトエンジニアリング[21]を実施することが挙げられる。

また、エンベディング[22]等の技術を利用して、サービス・プロダクト提供者のデータベースを活用することで出力結果の整合性を担保したりアウトプットの根拠や引用元を表示する技術（いわゆるグラウンディング[23]）の導入や、RAGの活用、フィルタリングによるハルシネーション検知なども、生成AIの出力結果の信頼性を担保する手段として事業者において取り得るものである。

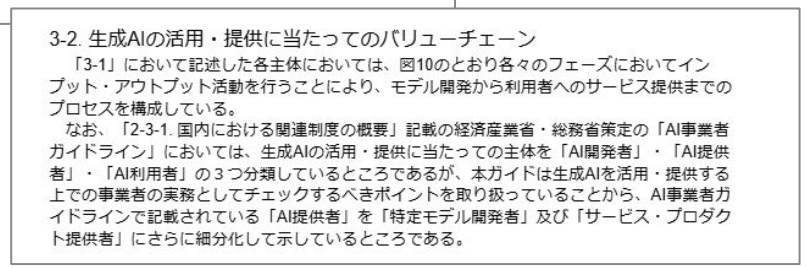
② 政策動向 を踏まえた改定

- 「2-3. 関連制度の概要」において、国内の生成AIに係る制度動向をアップデート
  - 経済産業省・総務省「AI事業者ガイドライン」
  - 個人情報保護委員会による個人情報保護法3年見直し
- 「3. バリューチェーン」や「4. チェックポイント」ではAI事業者ガイドラインを踏まえアップデート
  - プライバシー・バイ・デザイン
  - セキュリティ・バイ・デザイン etc.



④ プライバシー・バイ・デザインを考慮した対策の実施

学習やファインチューニング時など、プロダクト開発段階においてはプライバシー・バイ・デザインを通じて個人のプライバシーに配慮した設計を行うことが重要である。例えば、学習時のデータについて、第三者の個人情報や知的財産権に留意が必要なもの等が含まれている場合には、法令に従って適切に扱うことをAIのライフサイクル全体を通じて確保することや、AIシステムの実装の過程を通じて、採用する技術の特性に照らし適切に個人情報へのアクセスを管理・制限する仕組みの導入する等のプライバシー保護のための対策を講ずることも重要である。



3-2. 生成AIの活用・提供に当たってのバリューチェーン

「3-1」において記述した各主体においては、図10のとおり各々のフェーズにおいてインプット・アウトプット活動を行うことにより、モデル開発から利用者へのサービス提供までのプロセスを構成している。

なお、「2-3-1. 国内における関連制度の概要」記載の経済産業省・総務省策定の「AI事業者ガイドライン」においては、生成AIの活用・提供に当たっての主体を「AI開発者」・「AI提供者」・「AI利用者」の３つ分類しているところであるが、本ガイドは生成AIを活用・提供する上での事業者の実務としてチェックするべきポイントを取り扱っていることから、AI事業者ガイドラインで記載されている「AI提供者」を「特定モデル開発者」及び「サービス・プロダクト提供者」にさらに細分化して示しているところである。

ヘルスケア事業者のための生成AI活用ガイド第2.0版_改正のポイント（詳細）

③ 海外動向 を踏まえた改定

- 「2-3. 関連制度の概要」において、国外の生成AIに係る政策・制度動向をアップデート
  - 欧州、米国、中国、インド、韓国

①偽情報対策
AI Actでは、生成AIがもたらすリスクとして偽情報の拡散が特に懸念されている。ディープフェイクや自動生成されたコンテンツは、選挙や公共の意思決定に影響を与える可能性があるため、これらのリスクを軽減するための規制が必要とされている。2024年5月、欧州委員会はマイクロソフトに対し、同社の検索エンジンBingにおける生成AI機能について詳細な情報を提供するよう要求した。この要求は、マイクロソフトが提供する機能「Copilot」や「Image Creator」に関連するもので、これらがEUのデジタルサービス法（DSA）に違反している可能性があるとの懸念から発せられたものである。欧州委員会は、Bingが生成AIツールを使用してハルシネーションやディープフェイクを作成し、それが選挙において有権者を誤解させる可能性があると指摘している。これにより、マイクロソフトには2024年5月までに詳細な回答が求められており、回答しない場合には最大で年間売上高の1%の罰金、および平均日収または年間売上高の5%の定期的な罰金が科せられる可能性がある。この動きは、GoogleやMeta、TikTok等のテクノロジー企業にも波及しており、同様の説明責任を果たす必要がある。[2]。

を発令した。開発事業者はサービス提供や利用開始前に政府による安全性の評価を受けるよう義務付けることや、コンテンツが「AI製」であるか識別できる仕組みを設け、偽情報拡散防止を行う等のAI規制について記述されている。特に医療・ヘルスケア分野においては、AIが関わる危険な医療行為の事例を収集し、安全性の指針を作成する旨も規定している[6]。

図 2：大統領令における医療・ヘルスケア分野に関する言及

助成金等による開発・利用の支援	• 技術開発者による責任あるAI革新を推進し、医療分野の患者と労働者の福祉を促進するために、HHS（米国保健福祉省）の長官は、助成金等を特定し、優先的に関連する取り組みを行うことで、責任あるAIの開発と利用を支援する
全国的なAI Tech Sprint競技会の開催	• 退役軍人の医療の質を改善するAIシステムの開発を推進し、スタートアップ等企業に対する技術革新を支援する ✓ 全国的なAI Tech Sprint競技会の開催と、参加者に対する技術支援、メンタリング等
AI安全プログラムの設立	• 医療設定で展開されるAIから生じる臨床エラーを特定し、キャプチャするアプローチの共通フレームワークを確立し、患者、介護者、または他の当事者に害を及ぼす、バイアスや差別を含む関連事故の中央追跡リポジトリの仕様を設定するプログラム • 適切な場所で、これらの害を避けることを目指した、推奨事項、ベストプラクティス、または他の非公式のガイドラインを開発し、適切な利害関係者（医療提供者を含む）に普及する

出所：Federal Register :: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligenceを基に作成

③ 活用動向 を踏まえた改定

- 生成AI活用事例を拡充（ 4事例→13事例）
  - 国内の製薬企業の活用例や、海外の生成AIサービスなどを掲載
  - 医療機関 等、法人で生成AIサービスを導入する際の内規（院内ポリシー）のひな形を参考として添付

生成AI活用ガイド 概要

ヘルスケア生成AI活用ガイドの前提

本ガイドの目的

- ヘルスケア領域は他の領域と比較して要配慮個人情報の取扱いが多くなる点や、不確かな情報がもたらす個人への影響が極めて大きい点等が課題。
- そのため、**生成AIを活用したヘルスケアサービスが利用者に不当な不利益を供することとならないよう、当該サービスを提供しようとする事業者がセルフチェックできる目安となるチェックポイントを提供すること**　を目的として本ガイドを策定。

対象読者

- **生成AIを活用したヘルスケアサービスを提供する事業者**　をメインターゲットとする

  ※具体的には、特定モデル開発者及びサービス・プロダクト提供者が対象（後述のバリューチェーン参照）

- **生成AIを活用したサービス・プロダクト提供自体を初めて経験する事業者（生成AIの初学者）**　でも活用できるよう、チェックリストや用語集を別添で準備。

対象範囲

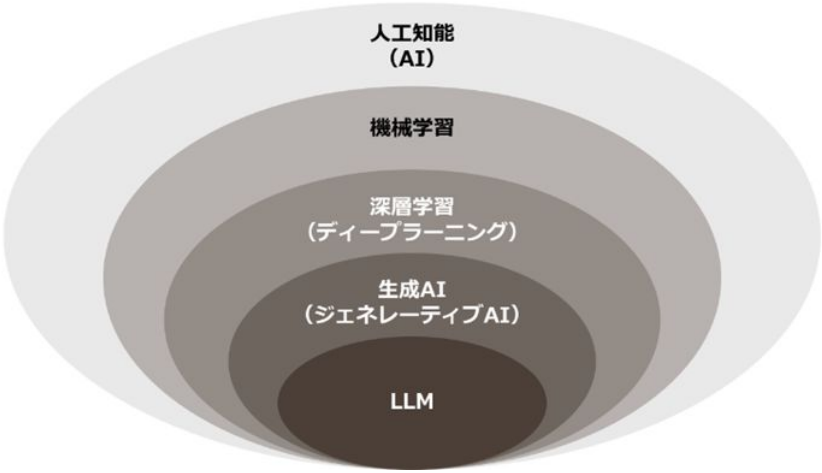
- **ヘルスケア領域で最も広く活用されている文章（テキスト）生成AIを対象**　※画像・音声・マルチモーダルは対象外
- **医療機器又は医療機器プログラムには該当しないヘルスケアサービスを想定**

  ※対象範囲等については、今後の技術やサービス進展を踏まえて随時アップデートを実施予定

生成AIの特徴

生成AIとは

- 生成AI（Generative AI）とは、自律的に学習したデータから文章、画像、音声などの一見新しく現実的なコンテンツを生成することができる一連のアルゴリズムのこと
- 生成AIの中で特に自然言語処理を行うのがLLM（Language Learning Model）＝テキスト生成AI



生成AIの特徴

基盤モデル等の活用

- 生成AIは巨大なデータセットを活用した「基盤モデル」や学習データを更にインプットした「特定モデル」の活用が必要
- **多種多様のモデルが市場に存在**

データの取扱い

- 生成AIは広範なデータを自律的に学習することで結果を出力
- **学習データやファインチューニングに活用するデータの取扱いが発生**

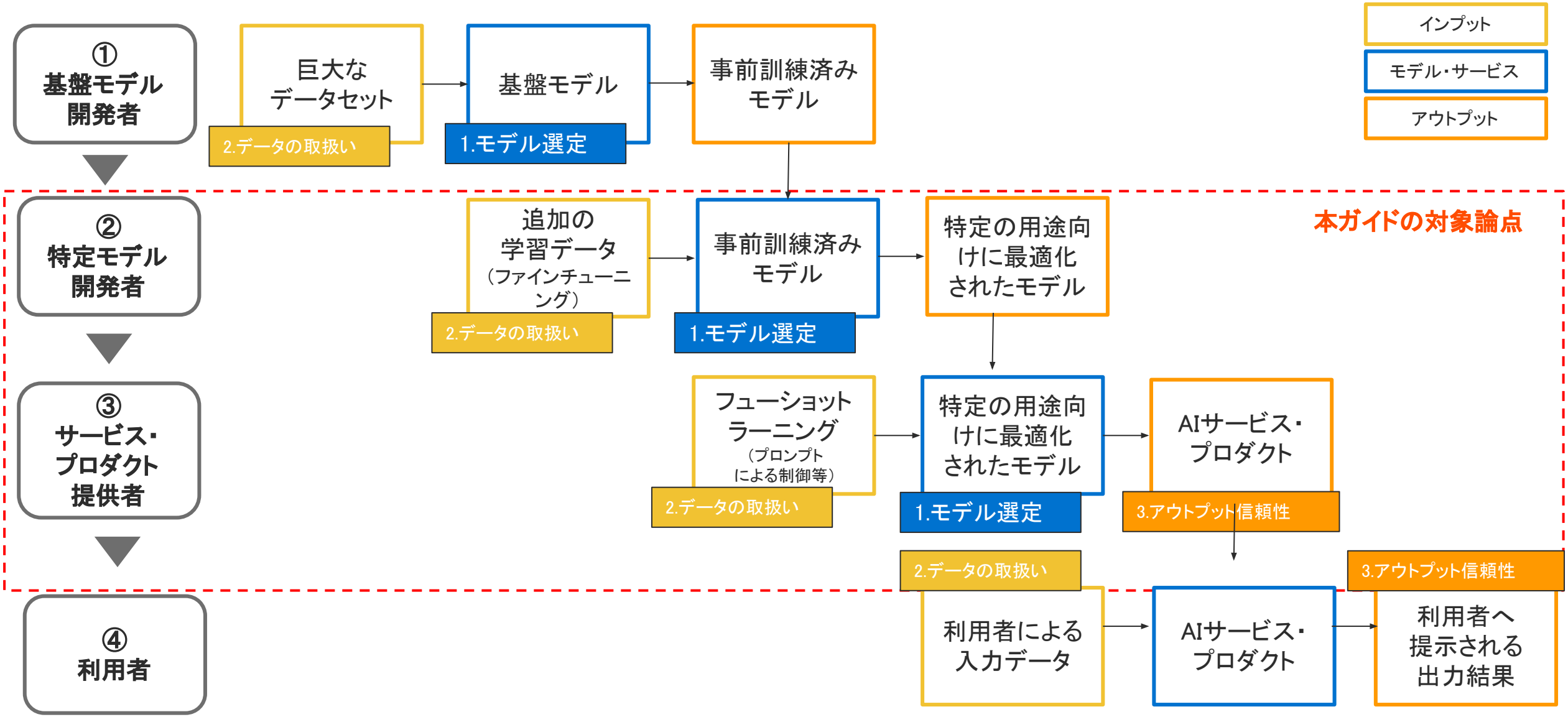
アウトプットの信頼性

- アウトプットの処理過程が複雑等の理由により、事実と異なる内容が出力される場合がある（**ハルシネーション**）
- 学習データが古いと出力も古い内容になることも想定される

利用者のリテラシー

- 専門知識を持たない利用者でもデータを入力することで簡単にアウトプットを取得できる（**低コスト・利便性**）
- 質問の内容等によって回答結果が異なる場合も想定される

生成AIのバリューチェーン



（参考）生成AI活用に当たっての関係主体整理

主体	概要	例
①基盤モデル開発者	大規模言語モデル（LLM）等の大規模で汎用的なモデルを開発・提供する事業者	OpenAI、Google、Meta、Amazon、Cyber Agentなど
②特定モデル開発者	①が提供するモデルを活用して自社データや業界固有のデータ等を用いてモデルをファインチューニングし、特定用途に特化したモデルを開発する事業者	①と③が混在している状態
③サービス・プロダクト提供者	①又は②で開発されたモデルを用いて、生成AIを活用したサービス・プロダクトを開発し、直接利用者に提供する事業者	Ubie、MICIN、HOKUTOなど
④利用者	生成AIを用いたサービス・プロダクトを利用する個人や法人	―

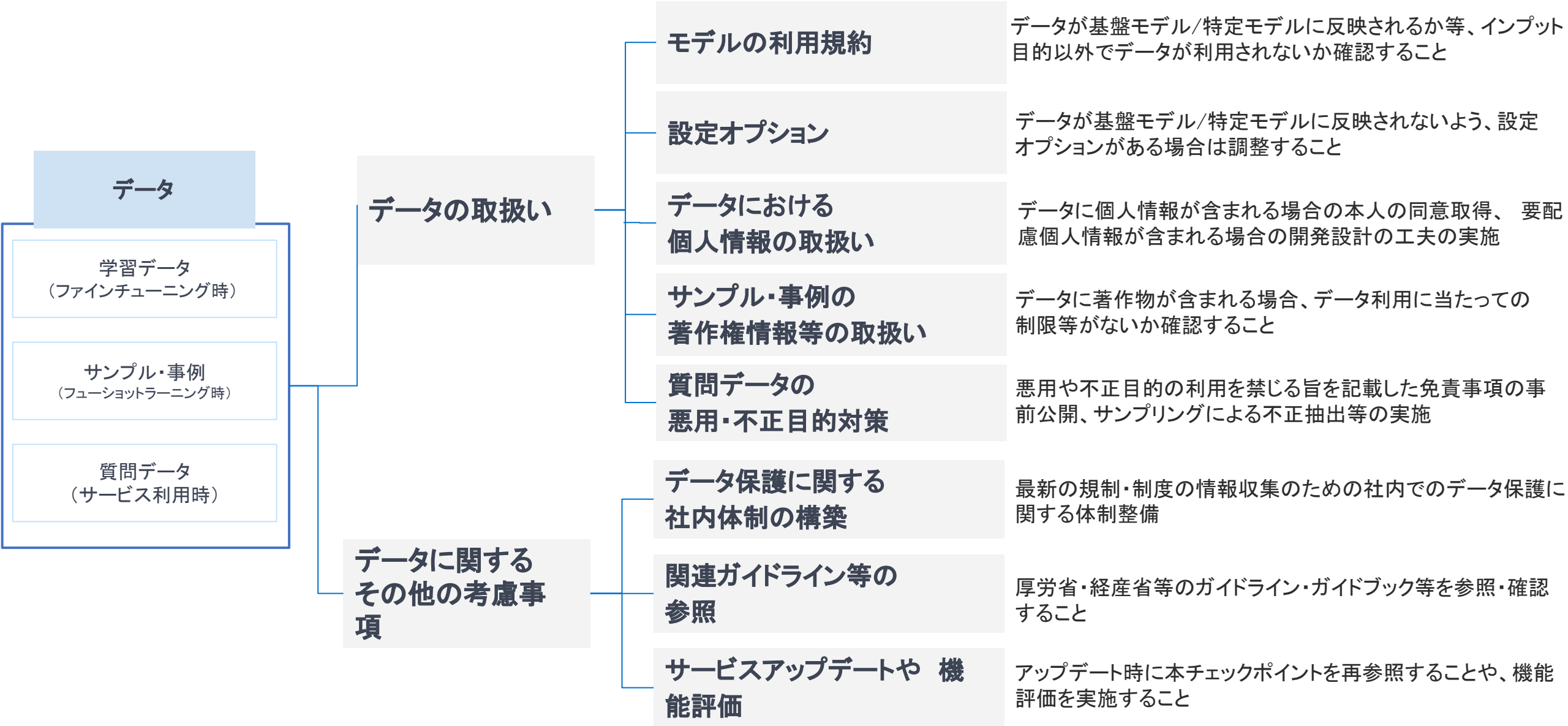
チェックポイント全体像

1 基盤モデルの選定	①基盤モデルの選定 ②基盤モデルの利用用途	● 基盤モデルが標榜している性能や学習データの内容についての確認 ● 基盤モデルが定めている利用用途や学習利用に関する規約の確認
2 データの取扱い	①学習データの取扱い ②サンプル・事例の取扱い ③質問データの取扱い ④データに関するその他考慮事項	● モデルの利用規約の確認 ● 個人情報が含まれる場合の本人同意取得 ● 著作物が含まれる場合の利用制限確認 ● データ保護に関する社内体制の構築 ● 関連ガイドライン等の参照
3 アウトプットの信頼性	①サービス開発段階での取組 ②サービス提供時の利用者に対する取組	● ハルシネーション制御（技術的工夫） ● 利用者に対する説明・表示 ● 入力規制・制御 ● 免責事項の表示
4 ヘルスケア領域の個別規制	①医療機器プログラムの該当性確認 ②標榜における広告規制の確認 ③基盤モデルの利用規約確認	● 医療機器プログラムの該当性確認 ● 医薬品等適正広告基準等の確認 ● ヘルスケア領域における利用制限の確認

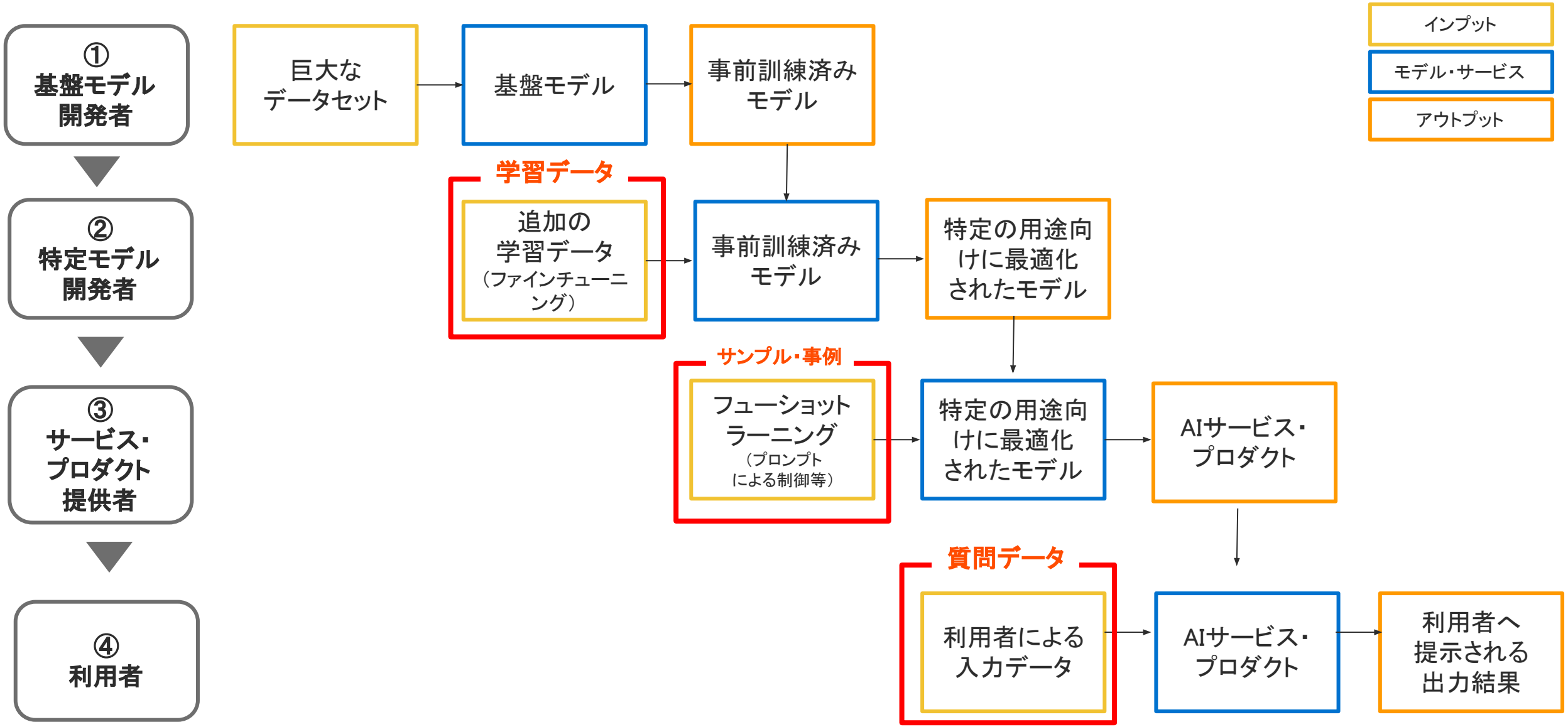
チェックポイント①　基盤モデルの選定

基盤モデルの選定	基盤モデルの標榜性能	基盤モデルの学習データの内容	学習データとしてどのようなデータセットをインプットして開発されたのかを確認すること
		性能評価レポート等の性能表示	性能評価報告レポート等による客観的な性能・機能判断を確認すること
		基盤モデルの類型	提供するプロダクト・サービスに応じた性能・特徴を持つ基盤モデルを選定すること（クラウドベース or オンプレ等）
	基盤モデルの利用用途	基盤モデルの利用用途の範囲	医療情報の提供を制限していないか、商用利用が限定されていないか等の観点で基盤モデルの利用規約を確認すること
		出入力データの学習利用に関する規約	基盤モデルの仕様として、インプット・アウトプットデータが追加学習の対象になっているか等を規約内で確認すること

チェックポイント②　データの取扱い



（参考）バリューチェーンと取扱いデータの関係性



チェックポイント③　アウトプットの信頼性

アウトプットの信頼性	サービス開発段階での取組	ハルシネーションを　制御する工夫	ファインチューニング・プロンプトエンジニアリングを実施による回答精度向上、エンベディング等の技術による出力結果の整合性担保、引用元の表示
		サービス品質評価の実施	サービスに対する技術性能や品質評価の実施、サービス利用者によるフィードバックに活用すること等
	サービス提供時の利用者への取組	テキスト生成AIに関する説明	テキスト生成AIの特徴や性能についてのチュートリアルを実施する等、利用者への説明・啓発を実施すること
		利用用途に関する説明・表示	サービス提供者が意図しない用途で利用されないためのサービスのコンセプトや用途等を説明・表示すること
		入力規制・制御の実施	サービス提供者が意図しない用途で利用されないための入力規制や制御を実施すること
		AIによる生成であることの表示	アウトプットがAIによって生成されたものであることを利用者が認識できるような出力結果の表示
		免責事項の表示	責任問題が生じた際の責任分界点を明確にするための免責条項の表示

チェックポイント④　ヘルスケア領域の個別規制

ヘルスケア領域の個別規制

- **医療機器プログラム該当性**
  生成AIを活用して提供しようとしているサービスがプログラム医療機器に該当しているかどうかを関連ガイドライン等を参照して確認すること
- **広告規制の確認**
  医療機器プログラムに該当する場合、医薬品等適正広告基準や業界ガイドライン等に基づいて広告・表示を行う必要があるため、それらを確認すること。該当しない場合、薬機法や景品表示法等の規定を確認すること
- **基盤モデルの利用規約の確認**
  基盤モデルによっては、ヘルスケア領域での利用制限をはじめとした様々な制約が決められていることもあるため、利用規約を確認・規約のアップデートがあった場合の確認を行うこと

（参考）事業者向けチェックリスト

別添として担当者がそのまま利用できるチェックリストをExcel形式で作成

別添1　生成AIを活用したサービス・プロダクトを提供する事業者向けチェックリスト

点検日　【　　】 前回点検日【　　】
点検担当者【　　】 前回点検担当者【　　】

※　求められる事項を満たしているか、同等以上の対応を行っている場合にチェックを付けること

「対応済み」or「対象外」をプルダウンで選択可能

1．モデルの基礎情報に関するチェックポイント

項目番号		内容	チェック	理由
1	基盤モデルが標榜している性能についての確認			
	1-1	利用予定の基盤モデルの学習したデータの内容を確認しましたか？ 例1.ジェンダーや人種など、データの内容にバイアスは含まれていないか。 例2.学習元データはライセンス利用不可なデータではないか。		
	1-2	利用予定の基盤モデルの性能・機能を客観的に判断するため、性能評価報告レポートなどが公開されているか確認をしましたか？ ※現時点ではレポートが公開されているモデルは少ないため、レポートが公開されていないモデルが適さないという意図のチェックポイントではありません。		
2	基盤モデルが定めている利用用途の確認			
	2-1	基盤モデルの利用規約において、医療や健康に関する情報の提供を目的にすることや商用利用について制限の有無を確認しましたか？		

チェック結果の「理由欄」を設置し、自由記述可能に

2．モデルのデータの取り扱いに関するチェックポイント

2．1．ファインチューニングの際に取り扱うデータ

項目番号	内容	チェック	理由
1	ファインチューニングに利用する学習データの取り扱い		