ヘルスケア事業者のための
生成 AI 活用ガイド

**（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）**

2024 年 1 月 18 日

日本デジタルヘルス・アライアンス（JaDHA）

<目次>

1. はじめに

1-1. 背景

インターネット技術の普及とともに SNS 等によってユーザー生成コンテンツが増加したことに伴い、機械学習に用いることのできるデータが急激に増加した。さらに、計算機単体の性能向上および大規模な分散処理化の実現により、大量のデータを効率よく処理できる環境が整った結果、あらゆる産業において AI の活用が進み、2023 年初頭に生成 AI（Generative AI）と呼ばれる新技術が急速に全世界へ普及しはじめた。生成 AI の技術そのものは数年前から存在し、AI 技術の進歩の結果として高度化してきたものであるが、ChatGPT をはじめとする簡易かつ高性能なツールがリリースされたことで、AI に関する知識や技術がなくとも、多くの人が簡単に AI の機能を取り扱えるようになった点において極めて革新的である。

ヘルスケア領域においても生成 AI に関する取り組みや検討は例外ではなく、医師国家試験に合格する AI や要約機能・患者への説明文案の生成等により医師の業務効率化をサポートする AI など、技術の可能性の探索や実際の医療現場等へのサービス提供といった取り組みが進められているところである。日本において非常に大きな社会的課題である医療・ヘルスケア分野（以下、「ヘルスケア領域」という。）においてこうした技術革新により多様なサービスが創出され、課題解決に貢献していくことは技術を活用する我々事業者としても社会的・経済的価値貢献の観点から非常に歓迎することであるが、他方で、技術の急速な普及によりその副作用も生じ得る点は考慮すべきである。

特にヘルスケア領域においては、要配慮個人情報の取り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を踏まえ、生成 AI を用いてヘルスケアサービスを提供する事業者が留意すべき点をまとめることとした。

なお、生成 AI 活用に関する業界策定ガイドラインはヘルスケア領域においていち早く策定したものであり（初版公表時）、生成 AI の技術の進展に伴い今後もアップデートを行っていく予定である。

1-2. 本ガイドの前提

1-2-1. 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をいう。最も強力な生成 AI のアルゴリズムは、ラベルのない膨大な量のデータを自律的に学習して、幅広いタスク向けに基本パターンを特定する基盤モデルの上に構築される[1]。

[1]「AI 関連領域における BCG の支援-生成 AI」BCG ホームページ
https://www.bcg.com/ja-jp/capabilities/artificial-intelligence/generative-ai
（参照 2024 年 1 月 15 日）

1-2-2. 本ガイドで対象とする範囲

生成 AI によるアウトプットは文章、画像、音声など多様であり、かつ、マルチモーダルと呼ばれる複数機能を統合的に活用できる技術なども随時開発・提供されているところであるが、本ガイドにおいては、まずはヘルスケア領域で最も広く活用されていると考えられる文章（テキスト）生成 AI を対象とする。なお、本ガイドは今後技術やサービスの進展を踏まえて随時アップデートを行う予定としている。

また、生成 AI は新しくかつ急速に普及した技術であるためヘルスケア領域における安定的な活用方法はまだ検証段階であることや、一般生活者への影響が生じやすいことを踏まえ、本ガイドにおいては原則として医療機器また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（以下、「ヘルスケアサービス」という。）を想定した記載とする。

1-2-3. 本ガイドの目的および対象読者

冒頭に記載したとおり、特にヘルスケア領域におけるサービス提供に当たっては、他の領域と比較して要配慮個人情報の取り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である。

これを踏まえ、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する。したがって、主には生成 AI を活用したヘルスケアサービスを提供する事業者を対象読者とする。

併せて、本ガイドを踏まえてサービス提供している事業者が、その旨を公表等することによって、当該サービスを利用するユーザー・顧客（医療機関等）が、安全性の目安として確認し、安心してサービスを利用できるようになることも期待する。

なお、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も多くいることが想定されるため、当該事業者が本ガイドを簡便に参考とできるよう、別添のチェックリストもまとめている。また、本ガイドにおいて記載される用語については参考資料の用語集を参照されたい。

2. 生成 AI に関する基礎情報

2-1. 生成 AI のヘルスケア領域における動向

2017 年に Google 社が発表した Transformer と呼ばれる機械学習モデルにより、自然言語処理分野の技術水準は飛躍的に進展した。このモデルを用いて 2018 年には Google 社が BERT を、

「ChatGPT and the Future of Medical Writing」Radiology Volume 307, Issue 2April 2023 https://pubs.rsna.org/doi/epdf/10.1148/radiol.223312（参照 2024 年 1 月 15 日）

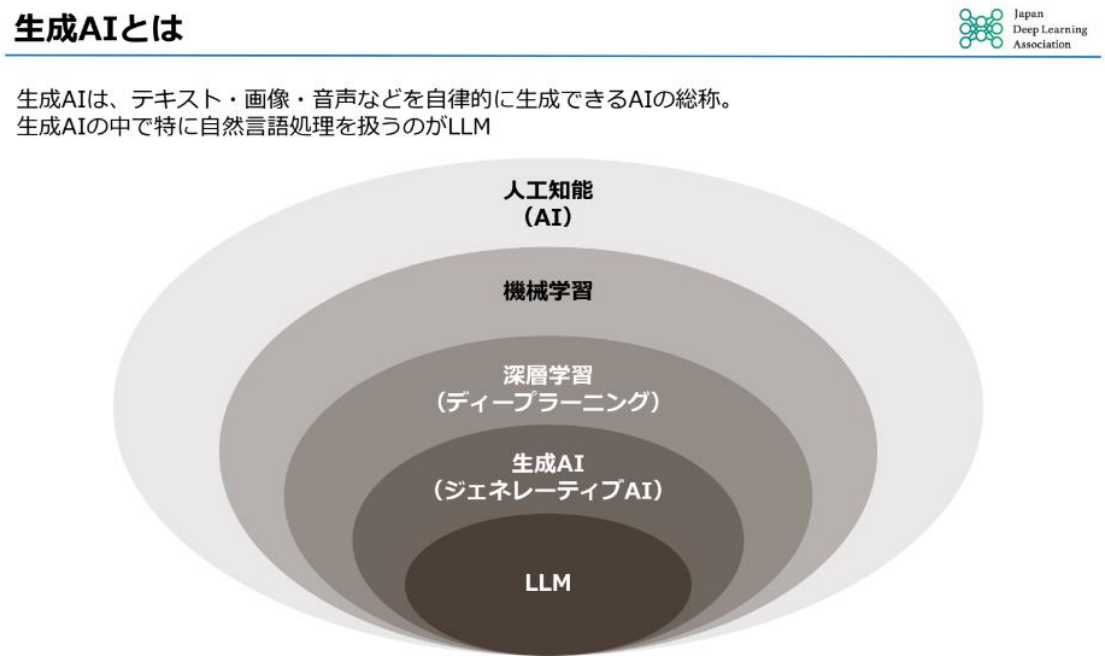
2020 年には OpenAI 社が GPT-3 を、そして 2022 年には同じく OpenAI 社が ChatGPT を発表した。

この Transformer モデルを利用した自然言語処理に関連する技術の一部は、現在ではテキスト生成 AI と呼ばれている。テキスト生成 AI は特に 2022 年 11 月に OpenAI 社によって提供が開始された ChatGPT をきっかけに爆発的普及が進んでおり、2023 年 12 月時点で医療分野を含めたさまざまな産業での活用が検討・推進されている。

テキスト生成 AI のように自然言語処理タスクで質問応答文や文書など任意の自然言語を出力するために大量のデータで学習された多数のパラメータを持つモデルは、図 1 のとおり、一般に大規模言語モデル（Large Language Models：LLM）と呼ばれており、OpenAI 社が提供する ChatGPT を含めた GPT シリーズや、Google 社の PaLM シリーズ、Meta 社の Llama シリーズなどが知られている。LLM はさまざまなタスクを実施することが可能であり、テキスト生成にとどまらず、質問応答・プログラムコードの生成・言語翻訳・文の校正・知識提供・クリエイティブな作品の生成などが代表的な使用例である。

なお、ヘルスケア領域における生成 AI 活用事例については、本ガイド作成に当たって検討にご協力いただいた企業の取り組みを中心に「参考資料 別添 2 ヘルスケア領域における生成 AI 活用事例集」として掲載しているので参照されたい。

図 1 ：生成 AI とは



(出典：一般社団法人日本ディープラーニング協会「AI 活用時の医療データの取扱いに関する JDLA 報告書について（JaDHA 会員向け生成 AI オープンセミナー資料）」（2023 年 10 月 23 日））

2-2. 生成 AI の特徴

生成 AI（Generative AI）は、前述のとおり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をいう。昨今ではウェアラブルデバイスやその他のデジタル技術の進化により分析可能なデータが日々増加しており、これらを活用して生活者の日々の健康管理等に寄与するサービスの発展も著しい。こうした中、生成 AI の活用により、さらなるサービス開発やこれによる健康増進等に大きな期待が寄せられている。

テキスト生成 AI は、非常に多岐にわたるタスクを実行することが可能である一方、さまざまな制約条件や課題が存在するため、生成 AI の特徴とそれに伴う論点について以下のとおり整理する。

2-2-1. 基盤モデルに関する特徴

生成 AI を活用するに当たっては、巨大なデータセットを活用した基盤モデルを利用することが必要となり、場合によっては基盤モデルに学習データをインプットしてさらに学習を進めることで特定用途に特化した基盤モデル（以下、「特定モデル」という。）を生成・活用することも可能である。そのため、どのような学習データをインプット・再学習したかによって基盤モデルの性質・機能も多種多様である。

2-2-2. データに関する特徴

生成 AI は広範なデータを自律的に学習することで入力された条件に対応する結果を出力する性質のものである。そのため、多くの事業者が提供するサービスにおいて共通の基盤モデルを利用して提供されることが想定されるが、例えば学習データやファインチューニングに活用するデータが他の主体に共有される可能性があるなど、個人情報保護や著作権保護をはじめとするデータの取り扱いに関する問題が懸念される。特にヘルスケア領域においては、要配慮個人情報を含む個々人の健康状態に関するデータや、著作物を含む医学論文上の情報などが含まれている場合も想定されるため、広範なテキストから学習する生成 AI を活用するに当たっては、特にデータの取り扱いについては十分な配慮が必要である。

2-2-3. アウトプットに関する特徴

生成 AI の特徴として、アウトプットに至る処理過程について人間による解釈が困難である上、AI の能力を統一的かつ網羅的に評価することが困難であることが挙げられる。そのため、共通の基盤モデルを活用している場合に、基盤モデル開発の精度向上施策が影響して逆に一部タスクの精度が下がるということが原理的にあり得る。また、基盤モデルを利用する際の設定値によっては確率的な挙動となることがあり、出力が安定しないことも発生し得る。そのため、アウトプットには、いわゆる「ハルシネーション」（幻覚）の問題が存在し、事実と異なる内容（嘘）や文脈と全く無関係な内容の出力が生成される可能性もある。さらに、基盤モデルの学習データに利用されているデータが古い情報の場合、アウトプット情報も古い

ものが出力されてしまうケースも想定される。これらを踏まえると、特にヘルスケア領域は健康や身体・生命に影響を与える情報を取り扱う場合があることが想定されるため、これらの特徴を理解し、アウトプットの信頼性を担保する取り組みを実施した上で生成 AI を活用する必要がある。

2-2-4. 利用者のリテラシーに関する特徴

生成 AI は、これまでの AI とは異なり、AI に関する専門的な知識を持たない一般利用者であってもデータを入力することで簡単にアウトプットデータを獲得できる点（低コスト性・利便性）も一つの特性である。また、生成 AI は適切な前提条件を設定することで、その能力が最大化され、より適切な回答を提供することが可能であり、最初の回答が不十分であったとしても利用者との対話を通じて回答を更新し、より適切な回答を生成することが可能である。

そのため、AI のパフォーマンスは、設定された前提条件やパラメータに大きく依存し、その結果にバイアスが生じる可能性がある。さらに、利用者が適切な指示ができなかった場合は本来なら正しい回答が可能な質問であっても、適切な回答に辿り着けない場合がある。このように利用者が入力するプロンプトの内容や品質により AI の出力結果が大きく異なる場合も発生するため、利用者側もその特性を理解できるリテラシーが必要になってくる。

2-3. 関連制度の概要

2-2 に記載のとおり、生成 AI については、これまで AI 技術の取り扱いに加えて考慮すべき特徴と論点が存在しており、これら論点については国内外で検討・整理が進められているところである。以下に国内および海外での主な関連制度の状況を整理する。なお、以下の情報は本ガイド執筆時点のものであり、AI 技術に関する領域は活発に規制内容等が変化するため、最新の関連制度とその改正状況は変更・進捗している可能性があることに注意されたい。

2-3-1. 国内における関連制度の概要

国内の関連制度	概要
新 AI 事業者ガイドライン スケルトン（案）（内閣府・AI 戦略会議）	「広島 AI プロセス」は生成 AI の国際的なルール作りを議論する枠組みとして 2023 年 5 月に閉幕した G7 広島サミットの首脳宣言で創設が盛り込まれた後、2023 年 9 月 7 日に閣僚級会合において「G7 広島 AI プロセス G7 デジタル・技術閣僚声明」が採択された。当該声明においては「2)AI 開発者を対象とする国際的な指針の策定」が掲げられており、開発を含むすべての AI 開発者を対象とした国際的な行

	動指針を 2023 年以内に策定するとしている。 この動きを踏まえ、内閣府・AI 戦略会議では 2023 年 9 月に行動指針案を提示し、G7 が 2023 年内に報告する成果文書への反映を目指している。本指針案の中では、人権侵害防止や多様性の尊重など AI でも守るべき基本責務を示し、AI の開発から提供・利用の各段階で企業に求められる役割や開発した AI の仕組みや機能を公開することが盛り込まれた他、AI の犯罪利用につながるような入出力の防止や生成 AI を活用したサービス事業者に対して考えられるリスクを利用者に通知することや AI の動作に責任を持つことと求める方向を示している。 なお、本行動指針案については、既存の事業者向けガイドライン（総務省「国際的な議論のための AI 開発ガイドライン案」/経済産業省「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」/経済産業省「AI 原則実践のためのガバナンス・ガイドライン」）の統合・改定を行う位置付けのものである。 内閣府・AI 戦略会議（第 5 回）「資料 1-2 新 AI 事業者ガイドライン スケルトン（案）」（2023 年 9 月 8 日） https://www8.cao.go.jp/cstp/ai/ai_senryaku/5kai/gaidorain.pdf
「生成 AI サービスの利用に関する注意喚起」および「OpenAI に対する注意喚起の概要」（個人情報保護委員会）	2023 年 6 月に個人情報保護委員会は「生成 AI サービスの利用に関する注意喚起」および「OpenAI に対する注意喚起の概要」を発出。主に個人情報を取り扱う事業者および行政機関等を対象として、生成 AI サービス利用に際しての個人情報取り扱いの注意点をまとめている。当該注意喚起においては、一般の利用者に対しても生成 AI サービスにおける留意点として個人情報が学習に利用されることがある等、サービス事業者の利用規約等を十分に確認し、入力する際はリスクを踏まえた上で適切に判断をすることとしている。また、当委員会は OpenAI 社に対して、本人の同意なしに要配慮個人情報を取得しないこと、個人情報の利用目的について日本語を用いて、利用者および利用者以外の個人に対して通知または公表することと注意喚起を行っ

	ている。 個人情報保護委員会「生成 AI サービスの利用に関する注意喚起等について」（2023 年 6 月 2 日） https://www.ppc.go.jp/news/careful_information/230602_AI_utilize_alert/
AI と著作権に関する論点整理について（文化審議会著作権分科会法制度小委員会）	2023 年 7 月に文化審議会著作権分科会法制度小委員会（第 1 回）は、AI と著作権に関する論点整理について資料を公表。AI サービス事業者や AI サービス利用者の侵害リスクを最小化等できるよう、生成 AI 発展を踏まえた論点整理を行っている。主要論点項目としては、「学習用データとして用いられた元の著作物と類似する AI 生成物が利用される場合の著作権侵害に関する基本的な考え方」や「AI（学習済みモデル）を作成するために著作物を利用する際の基本的な考え方」、「AI 生成物が著作物と認められるための基本的な考え方」を挙げている。 文化審議会著作権分科会法制度小委員会（第 1 回）資料 3「AI と著作権に関する論点整理について」（2023 年 7 月 26 日） https://www.bunka.go.jp/seisaku/bunkashingikai/chosakuken/hoseido/r05_01/
「広島プロセス G7 首脳声明・広島 AI プロセスに関する G7 首脳声明高度な AI システムを開発する組織向けの広島プロセス国際指針」・「高度な AI システムを開発する組織向けの広島プロセス国際行動規範」	2023 年 10 月 30 日、G7 首脳は広島 AI プロセスに関して G7 首脳声明を発出。首脳声明の他、「広島 AI プロセスに関する G7 首脳声明高度な AI システムを開発する組織向けの広島プロセス国際指針」および「高度な AI システムを開発する組織向けの広島プロセス国際行動規範」も併せて公表した。 首脳声明では、AI の革新的な機会と変革の可能性を強調するとともに、リスク管理と共有原則の遵守の重要性を認識した上で国際指針と行動規範を作成することが記載されている。国際方針は安全・安心・信頼できる AI を世界に普及させることを目的とし、生成 AI を含む高度な AI システムを開発・利用する組織向けに策定され、サイバーセキュリティ・個人データ保護をはじめとした安全対策等の原則等、11 項

	目の原則の遵守が求められている。国際行動規範においては、国際指針に記載の 11 項目に基づいた具体的な計画内容が記載されている。 外務省「広島 AI プロセスに関する G7 首脳声明」（2023 年 10 月 30 日） https://www.mofa.go.jp/mofaj/ecm/ec/page5_000483.html
プログラム医療機器該当性に関するガイドライン等（厚生労働省）	開発したプログラムの標榜は当該プログラムが医療機器に該当するか否かで異なり、医療機器該当性判断については、表示、説明資料、広告等に基づき当該プログラムが薬機法で定められている医療機器としての目的性（疾病の診断、治療、予防）および人の生命・健康に影響のリスクがどの程度寄与するかによって判断される。開発されたプログラムが医療機器に該当するか否かに関しては厚労省が公開している判断事例等やプログラムの医療機器該当性に関するガイドラインが参考となる。 厚生労働省「プログラムの医療機器該当性に関するガイドラインについて（2023 年 3 月 31 日一部改正）」 https://www.mhlw.go.jp/content/11120000/001082227.pdf また、汎用 AI に関しては、厚生労働省は、同省ホームページにおいて「医療機器プログラムと汎用 AI の違いについて」を公表している。ここには、「汎用 AI などのその他のプログラムは医療機器として承認・認証されたものではなく、疾病や診断の予防、治療の目的を標榜して提供することはできない」旨や、「健康状態や疾病に関する質問をした場合の回答内容を含めたその性能は、医薬品医療機器等法に基づき、その妥当性が確認されたものではない」旨が記載されている。 厚生労働省「医療機器プログラムと汎用 AI の違いについて」（厚生

	労働省ホームページ） https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000179749_00004.html

2-3-2. 海外における関連制度の概要

ここでは、海外、特に欧州・米国・中国・インド・韓国における生成 AI を含む AI 全般に関する法規制・制度について記述する。

＜欧州の動向＞

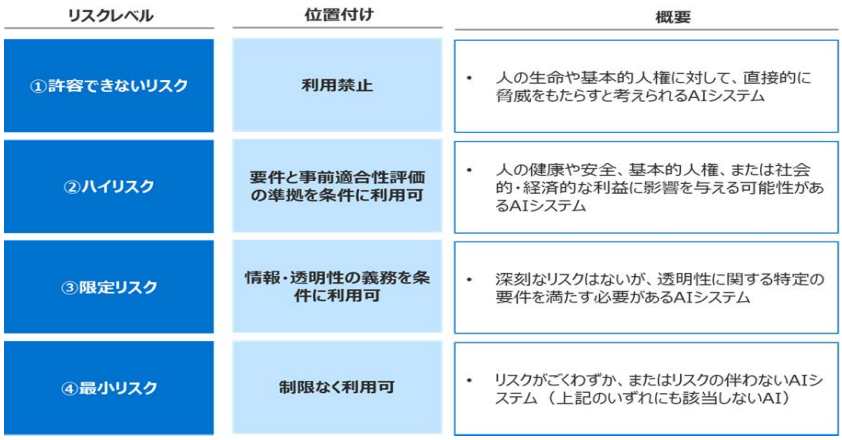
欧州では、2021 年以降に AI を巡る規制の動きが活発となっており、2021 年には、欧州委員会が「AI Act」を発表し、安全保障目的の AI や悪用の可能性がある AI に規制がかかる可能性があることに言及した。さらに、2022 年には、欧州理事会（EU 加盟国による意思決定機関）がイノベーションを重視した修正案を発表。2023 年 6 月には、生成 AI を含む包括的な AI の規制案である「AI Act」が、欧州議会の本会議において賛成多数で採択された。早ければ年内の合意、2024 年以降の施行が見込まれている。

欧州の「AI Act」は、以下の 2 つの特徴を有している。

- リスクベースでの AI 分類を行い、要求事項と義務を整理している。
- イノベーション支援についても言及している。

リスクレベルは、図 2 に示すとおり、4 段階に分類されており、それぞれに利用の制限や対象となる AI システムが定義されている。

図 2：欧州「AI Act」におけるリスクベースでの AI 分類



(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

ヘルスケア領域に関する言及に着目すると、記載内容は限定的ではあるが、②ハイリスクの中に位置付けており、図 3 に示す言及がなされている。

図 3　欧州「AI Act」におけるヘルスケア領域への言及（②ハイリスク）

- (a) 公的機関によって、または公的機関に代わって使用されることを目的としたAIシステムで、ヘルスケアサービスや必須サービス（住居、電気、食事/冷房、その他のサービスを含むがこれらに限定されない）を含む公的扶助の給付およびサービスに対する自然人の適格性を評価する。

- (c) 自然人による緊急通報を評価および分類することを目的としたAI システム、または警察や法執行機関、消防士や医療援助者などによる緊急初期対応サービスの派遣、または派遣における優先順位の確立に使用することを目的としたAI システム救急医療患者トリアージシステムの開発。

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る　欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

各リスクレベルにおける、AI サービスの要求事項と義務についても、図 4 および図 5 に示すように定義されている。運用においては、AI 提供者だけでなく AI 利用者に対しても要求事項が整理されている。

図 4：欧州「AI Act」における AI サービスの要求事項と義務

リスクレベル	要求事項と義務	
②ハイリスク	規制対象製品の安全要素 特定分野のAIシステム	• リスク管理プロセスを確立して実装 • 高品質な学習、検証、テストデータの利用 • 文書化の確立、ログ機能の設計 • 適切な透明性確保、ユーザーへの情報提供 • 人間による監視 • 堅牢性、正確性、サイバーセキュリティ確保 • ガイドラインや整合法令を考慮した上で当規制を遵守する義務
③限定リスク	透明性義務が適用される AIシステム	• 人とAIシステムが相互作用することが明らかではない場合、人に通知する • 感情認識または生体認証システムが適用されていることを人間に通知する • ディープフェイクに対して警告ラベル付けをする（基本的権利の行使や公共の利益に反しない場合に限る）
④最小リスク	上述以外	• 必須義務はなし • 低リスクAIシステムについても、ハイリスクAIシステムに対する要求事項を自主的に適用するよう欧州委員会などが推奨。実施方法として行動指針の作成を提示

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る　欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

図 5：欧州「AI Act」における運用に関する要求事項

対象者	運用に関する要求事項
AI提供者	• 透明性・説明可能性に関した義務 • 組織内に品質マネジメントシステムを確立し、実施 • 最新の技術文書を作成・更新 • ユーザーがリスクの高いAIシステムの動作を監視できるようにするためのログ記録義務（6か月間） • システムの適合性評価と、再評価の実施（大幅な変更がある場合） • EUデータベースにAIシステムを登録する • CEマーキングを貼付し、適合宣言に署名する • 市場投入後にモニタリングを実施する • 市場監視当局と協力する • アクセシビリティ要件に準拠する
AI利用者	• 取扱説明書に従ってAIシステムを操作する • AIシステムを使用する際、人間による監視を確保する • 起こりうるリスクについて運用を監視する • 重大な事故または誤動作について、AI提供者又はAI配付者に通知する • 既存の法的義務は引き続き提供される（GDPR等）

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る 欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

基盤モデルについては、ユースケースの観点からの分類は不適として、リスクレベルによる 4 カテゴリとは別で整理されており、基本思想（健康、安全、基本的権利など）に対するリスクを開発前・開発中からコントロールすることに努める、下流の開発者が法令遵守できるように技術文書を用意する、ハイリスク AI システム同様にデータベース登録や透明性の義務を遵守する、等特有の要求事項が明記されている。

適用範囲と罰則についても言及されており、EU 圏を対象に市場投入される AI とその提供者、展開者に加え、利用者も対象となり、また、アウトプットのみでも対象となる。罰則については、現時点で GDPR より大きい規模での制裁基準が設定されており、EU の AI 利活用における姿勢がうかがえるものとなる。

＜米国の動向＞

米国では、2022 年 10 月に AI 権利章典（AI Bills of Rights）が公表され、AI を含んだ Automated Systems を開発する際の非拘束的な 5 つの原則が記載された。2023 年 7 月には、連邦議会で生成 AI に関する規制の必要性が議論されている状況を踏まえ、主要米国企業 7 社（Amazon、Anthropic、Google、Inflection、Meta、Microsoft、OpenAI）が安全で透明な AI 技術開発をすることに合意した。

現時点では 7 社の自主規制から開始しているが、今後は、米政府は法的拘束力を持たせるために大統領令を準備しており、米連邦議会も与党・民主党を中心に AI 関連法案の検討を進めている。その際には、7 社以外のスタートアップ等にも参加を求めていくと報じられている。

2023 年 10 月 30 日にバイデン米大統領は、AI の安全性の確保や技術革新を図るための大統領

令を発令。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[2]。

図 6：主要米国企業 7 社による AI 技術開発に対するコミットメント概要

製品の安全性確保	• AIシステムの公開前に内部および外部のセキュリティテストに取り組む。 • AIリスクの管理に関する情報を産業界や政府、市民社会、学術界と共有する。
セキュリティを最優先するシステムの構築	• AIシステムの中で最も重要な部分である非公開のモデルの重みを保護するために、サイバーセキュリティと内部脅威対策に投資する。 • AIシステムにおける脆弱性の発見と報告を第三者が容易に行えるようにする。
社会の信頼獲得	• AIが生成したコンテンツであることをユーザが認識できるように、ウォーターマーキングシステムなどの堅牢な技術メカニズムの開発に取り組む。 • AIシステムの能力、限界、適切なおよび不適切な使用領域について公開的な報告を行うことを約束する。 • 有害な偏見や差別を回避し、プライバシーを保護するために、AIシステムがもたらす社会的なリスクに関する研究を優先する。 • がん予防から気候変動の緩和に至るまで、社会の最大の課題に対処するために先進的なAIシステムを開発・展開する。

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る 欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

＜中国の動向＞

中国は、国家戦略として AI 産業の発展を促進している。既存の法令とアルゴリズム規制、AI 倫理規制と標準制定などを組み合わせる形で AI 規制を行っており、中国独自の AI ガバナンスモデルが形成されつつある。中国の AI に関する国家戦略全体は、図 7 に示すように 5 つから構成されている。

[2] 日本経済新聞「米大統領令、生成 AI を初規制 公開前に安全評価義務づけ」（2023 年 10 月 30 日）https://www.nikkei.com/article/DGXZQOGN301180Q3A031C2000000/

図 7：中国の AI に関する国家戦略全体像

法制度	全人代	• サイバーセキュリティ法（2017年） • データセキュリティ法（2021年） • 個人情報保護法（2021年）
	国家インターネット情報弁公室	• インターネット情報コンテンツ生体治理規定（2019年） • インターネットコメントサービス管理規定（2022年）
AIアルゴリズム規制	国家インターネット情報弁公室	• インターネット情報サービスアルゴリズム・レコメンデーション管理規定（2022年） • インターネット情報サービス深度合成アルゴリズム管理規定（2022年） • 生成人工知能サービス管理暫行弁法（2023年）
AI倫理的規制	国務院	• 新一代人工知能ガバナンス準則－責任ある人工知能の発展（2019年） • 科技倫理ガバナンス強化に関する意見（2022年）
	科学技術部 （人工知能ガバナンス専門委員会）	• 新一代人工知能倫理規範（2021年） • 科技倫理審査弁法（試行）（パブコメ）（2023年）
AI関連標準制定ガイド	国家標準化管理委員会等	• 国家新一代人工知能標準体型建設ガイド（2020年）
	全国信標委人工知能サブ技術委員会	• 人工知能倫理治理標準化ガイド（2023年）
政策	国務院	• 新一代人工知能発展規画（2017年）
	工業情報化部	• 新一代人工知能産業発展促進三年行動計画（2018-2020）（2017年）
	科学技術部	• 国家新一代人工知能開放イノベーションプラットフォーム建設工作ガイド（2019年）

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る 欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

法制度については、既存のサイバーセキュリティ関連法令を活用しながら、AI 関連の規制を制定している。具体的には、電子商務法、個人情報保護法等において、事業者に対し、その用いるアルゴリズムに関して、利用者が見つけやすい場所に顕著な方法での告知による説明責任を果たすことを義務付けている。

AI アルゴリズム規制については、プロバイダーに対する責任を規定し、アルゴリズムの透明性確保のための届出制度を特徴とする。具体的には、レコメンデーション・アルゴリズム規定、深度合成アルゴリズム規定、生成 AI 弁法のいずれも、アルゴリズム利用についてのユーザーに対する明示義務を設定しており、欧州が目指すアルゴリズム規制と類似している。

AI 倫理的規制については、2023 年 4 月に中国科学技術部が「科技倫理審査弁法（試行）」のパブリックコメントを開始した。この弁法は、2022 年国務院の「科技倫理ガバナンス強化に関する意見」における AI を含む科学研究活動に対する倫理五原則を具体化する法令である。AI の研究開発を含む生物学、医学など科学研究開発を行う大学、企業などの組織に対して、倫理委員会の設立、研究開発活動の倫理審査など Ethics by Design を求めている。

AI 関連標準制定ガイドについては、体系的な標準制定を進めており、2020 年に国家標準化管理

委員会などが合同で制定した「国家新一代人工知能標準体系建設ガイド」では、AI の技術面における標準体系の制定を目標としている。2023 年には「人工知能倫理治理標準化ガイド（2023 版）」が作成され、AI の全ライフサイクルにおいて、倫理五原則をさらに 10 のガバナンス準則に詳細化した。

一方で、ChatGPT の成功もあり、中国国内の各企業も活発なプロダクト開発や提供を進めている。騰訊控股（テンセント・ホールディングス）や百度（バイドゥ）、アリババグループといった大手が相次いで自社プロダクトを発表している。アリババグループでは、2023 年 6 月に大規模言語モデルである「通義千問」を搭載した AI アシスタント「通義聴悟」をリリースした。音声や動画をリアルタイムでテキストに変換する他、話者やセクションごとに要点をまとめることができる。EC 等自社優位な事業と生成 AI の融合を目指し、消費者および事業者向けのサービス入り口の獲得を図っており、その応用分野として医療等も位置付けられている。また、アリババ傘下のサービスおよび出資先企業を中心に「通義千問」を提供していく予定である。

アリババのように業界横断的な汎用 AI プラットフォームを構築する大手が存在する一方で、医療ヘルスケア業界特化でサービスを提供する企業もある。音声認識 AI 大手の iFlytek は、大規模言語モデルである「星火認知」を基に業界特化型の実装を進めており、医療ヘルスケア領域では診療後管理やリハビリ指導に応用するプロダクトを開発している。また、成都医雲科技では、医療業界向けの大規模言語モデル MedGPT を開発し、20 億件の医学テキストデータ、800 万件の臨床診断データを活用し、問診から診断支援までの全プロセスを AI でサポートしている。正確性の担保として、提携医師による診断結果評価による学習も導入している。

また、各企業のプロダクト開発に当たっては、図 8 のような主要都市の AI 支援政策により、基礎～応用開発に対する後押しがなされている。

図 8：中国主要都市の AI 支援施策

上海	• 2023年4月、上海市政府は「新時代における投資促進の強化と現代産業体系の構築を加速するための政策措置」を発表。条件を満たす人工知能のキーコンピューティング技術プロジェクトに対し、プロジェクト投資額の30%を上限として最大2,000万人民元の支援を提供
北京	• 2023年5月、北京市政府は「北京市における世界的な影響力を持つ人工知能イノベーションの源泉地の建設を加速するための実施計画（2023～2025年）」を発表。人工知能のカギとなる技術課題の解決、人工知能の基礎技術の強化、人工知能の応用への実装を加速する方向で展開
深セン	• 2023年5月、深セン市政府は「深セン市における人工知能の高品質な発展と高水準な応用を促進するための行動計画（2023年～2024年）」を発表。総額1,000億元の人工知能ファンドを設立
成都	• 2023年6月、成都市は「成都市の人工知能産業の高品質な発展をさらに推進するためのいくつかの政策措置（意見募集稿）」を発表・国家の科学技術重大プロジェクトの成果を成都市で実装を進める場合、最大1,000万元の資金支援を提供

(出典：株式会社日本総合研究所「ヘルスケア領域での生成 AI の利活用を巡る 欧米の規制等動向（JaDHA SubWG-B 定例会資料）」（2023 年 9 月 25 日））

＜インドの動向＞

インド政府は、2023 年 4 月に AI を規制しない方針を発表した。しかし、同年 5 月には、AI の規制を含む「デジタル・インディア法案」の協議が開始されている。

AI を規制しない背景には、他国と比較して AI への投資額が低く、注目を集める AI スタートアップを生み出せていないことがあると指摘されている。「デジタル・インディア法案」は、AI システムの開発と使用に関する一連のルールと規制を設定するもので、公正性、透明性、安全性を確保することを目的としている。法案の草案によると、AI システムの開発者は、システムの目的と使用方法を文書化しなければならない。また、システムのトレーニングデータが偏っていないことを証明し、システムが誤った結果を出さないようにしなければならない。さらに、法案は、AI システムの使用に関する特定の規制も設ける。例えば、AI システムは、人々のプライバシーを侵害したり、差別したりしてはならない。また、AI システムは、人々の安全を脅かしたり、社会の秩序を乱したりしてはならない。この法案が成立した場合、生成 AI の開発や使用にも一定の規制が適用されることになるが、具体的な内容については、法案の審議が進む中で明らかになる。

＜韓国の動向＞

2023 年より、韓国では生成 AI を巡る議論が活発化している。科学技術情報通信部は、4 月に有識者で構成する第 2 期 AI 倫理政策フォーラムを立ち上げ、ChatGPT などの生成 AI の普及に対応する AI 倫理と信頼性確保に向けた本格的な議論を開始した。「AI 倫理基準実践に向けた自律点検表（案）（チャットボット、作文、映像分野）」および「信頼できる AI 開発案内書（案）」を公開し、関連技術と倫理問題に対する国内外動向を把握した上で、バランスの取れた視点で韓国の AI 倫理政策方向を提示し、関連の主要政策課題に対する議論を進めていく予定である。フォーラムは、倫理・技術・教育をテーマとした 3 つの分科会を運営して議論する。

韓国では今のところ、生成 AI への具体的な規制に向けた動きは見られない。

3. 生成 AI の活用・提供に当たってのバリューチェーンと論点

3-1. 生成 AI の活用・提供に当たっての関係主体

生成 AI はその活用・提供に当たってさまざまな主体が関係している他、各主体におけるインプット・アウトプット活動によってそのバリューチェーン（価値創造過程）が構成されている。本ガイドにおいては、関係主体を図 9 のとおり①基盤モデル開発者、②特定モデル開発者、③サービス・プロダクト提供者、④利用者の 4 主体に分類する。

図 9：生成 AI の活用・提供に当たっての関係主体

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[3]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または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 、HACARUS 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	–

（出典：JaDHA WG4 SubWG-B にて作成）

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。

①基盤モデル開発者における活動

巨大なデータセットを基盤モデルにインプットし、事前訓練済みモデル（大規模言語モデル等）を開発する。

②特定モデル開発者における活動

追加の学習データを①で開発された事前訓練済みモデル（大規模言語モデル等）にインプットすることでファインチューニングを行い、特定の用途向けに最適化されたモデルを開発する。

[3] 学習済みの基盤モデルに対して別のデータセットを活用して追加学習させること。用語集参照。

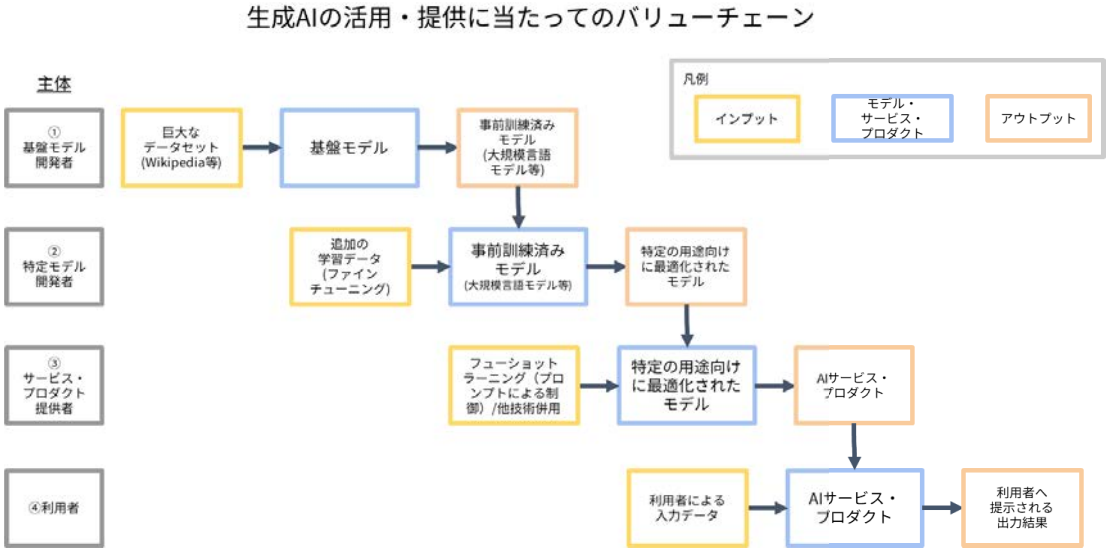
③サービス・プロダクト提供者における活動

②で開発された特定の用途向けに最適化されたモデルに対してフューショットラーニング（プロンプトによる制御）や他技術を併用することで、AI サービス・プロダクトを開発・提供する。

④利用者における活動

③で提供された AI サービス・プロダクトに質問等のデータを入力し、出力結果を得る。

図 10：生成 AI の活用・提供に当たってのバリューチェーン



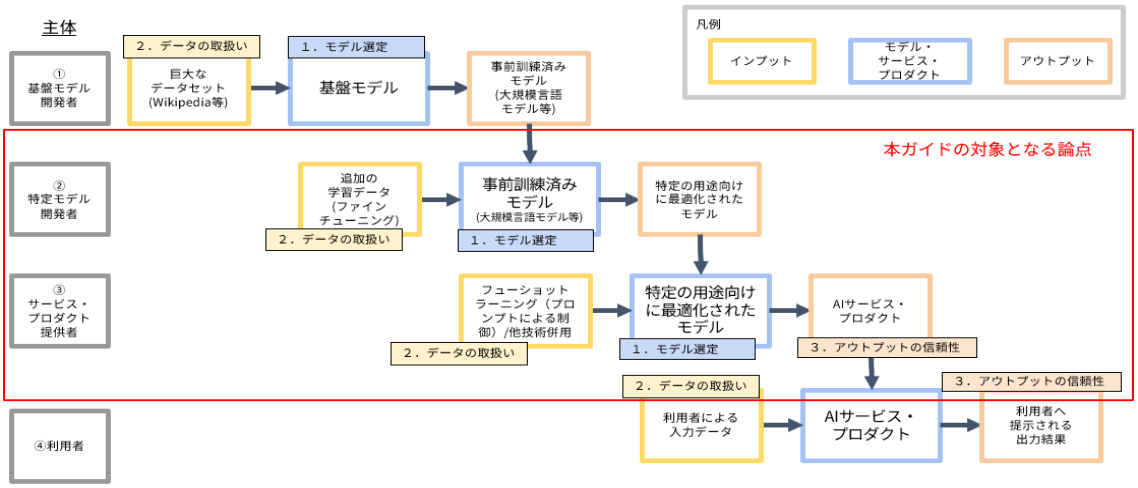
（出典：株式会社 MICIN 提供資料）

3-3. バリューチェーンを踏まえた生成 AI の活用・提供に当たっての論点

「1-2-3. 本ガイドの目的および対象読者」でも記述したとおり、本ガイドは、生成 AI を活用したヘルスケアサービスを提供しようとする事業者を対象読者として策定しているものであることから、バリューチェーンにおける、特に ②サービス・プロダクト提供者 における論点および考慮事項が主な焦点となる。一方で、図 9 のとおり、②サービス・プロダクト提供者が、③特定モデル開発者を担う場合も想定されることから、以下では、②特定モデル開発者 および ③サービス・プロダクト提供者の観点で生成 AI の活用・提供に当たって論点になる事項について整理を行う。これらを踏まえると論点は以下図 11 の赤枠のとおりとなる。

図 11：バリューチェーンを踏まえた本ガイドでの対象論点

生成AIの活用・提供に当たってのバリューチェーン



(出典：株式会社 MICIN 提供資料)

上図および「2-2　生成 AI の特徴と論点」で記述した特徴と論点を踏まえて、本ガイドにおいて特に論点となる事項を整理すると、下記表のとおりとなる。

＜バリューチェーンを踏まえた生成 AI の活用・提供に当たっての論点＞

論点	論点の詳細
１．モデル選定	<u>①基盤モデルの選定基準【特定モデル開発者】</u> 基盤モデル開発者によって開発されたモデルは市場に複数存在し、かつ、各々のモデルにおける学習データも異なっている。そのため、クラウドベースで汎用的な自然言語処理が可能なモデルもあればローカル・オンプレミスでヘルスケアに関する専門的な自然言語処理が可能なモデルもあるといったように、各々の基盤モデルによって特徴や用途・性能が異なっている。また、基盤モデルによっては商用利用が制限されている等、用途に一部制限が課されている場合もある。そのため、特定モデルを開発するに当たっては、基盤モデルの特徴や用途・利用可能範囲を事前に確認・判断しておく必要がある。 <u>②特定モデルの利用上の注意【サービス・プロダクト提供者】</u>

	特定モデルについては、利用に当たっての責任範囲があらかじめ明示されている場合が多く、サービス・プロダクト提供者が当該特定モデルを活用してサービス・プロダクトを提供する場合にサービス内容や利用者による利用方法の範囲が、その責任範囲内かどうか確認する必要がある。
2．データの取り扱い	①ファインチューニングに利用するデータの取り扱い【特定モデル開発者】 基盤モデルに対してファインチューニングを行う際、学習用データセットを活用するが、ファインチューニングに用いる学習用データは基盤モデルそれ自体に影響を及ぼさない場合が多い一方で、基盤モデルによっては学習データが反映されるものも存在する。また、時間経過によって学習用データセットの内容が陳腐化してしまうことやデータそのものに誤りがあった場合に、期待される出力とは全く異なる出力結果を誘発してしまう可能性も存在している。特に、ヘルスケア領域においては、ファインチューニングを行う際、著作権を含む論文情報や要配慮個人情報を含む健康データの他、日々の技術や研究の進捗に伴い最新情報がアップデートされる医学分野の専門情報などが学習データとして活用されることが想定されるため、これらに対応する取り組みを講じる必要がある。 ②フューショットラーニング等に利用するサンプル・事例の取り扱い【サービス・プロダクト提供者】 サービス・プロダクト提供者が特定の用途向けに最適化された特定モデルに対してフューショットラーニング（プロンプト制御）や強化学習を行う場面において使用するサンプル・事例の取り扱いについても注意が必要である。具体的には、①において記述した点と同様、フューショットラーニング等に活用されたサンプル・事例が基盤モデルに反映されないように配慮することや個人情報が含まれていた場合の対応が求められる。 ③利用者が入力する質問データの取り扱い【サービス・プロダクト提供者】 生成 AI を活用したサービス・プロダクトは利用者によって自由に文章等を記入することで回答が出力される設計が想定されるが、利用者によ

	って個人情報等を含むデータがインプットされる場合も考えられ、当該データが特定モデルに追加学習される可能性も考えられる。特にヘルスケア領域においては、利用者が自己の健康・症状等に関する機微な情報を入力することが想定されるため、特定モデル事業者またはサービス・プロダクト提供者において当該データの取り扱いに十分注意する必要がある。
3．アウトプット信頼性	①サービス・プロダクト開発段階での取り組み【サービス・プロダクト提供者】 生成AIはその性質上出力が確率的に変動するため、精度が100%ではなく、ハルシネーション等が発生することが想定される。そのため、それらを所与のものとした上で利用者に対して大きな影響が出ないよう、技術面や仕組み面で制御の工夫を行うなどのサービス設計を開発段階から取り組む必要がある。特にヘルスケア領域においては患者の過去の健康記録、現在の病状、遺伝的リスク、ライフスタイルなど、多角的な情報を統合・アウトプットを行うことや、特に医療分野においては他産業と比較してプロダクト・サービスに求められる品質は一般に高くなることを踏まえ、サービス・プロダクト提供者において開発段階で可能な限りの取り組みが講じられることが求められる。 ②利用者のリテラシー向上の取り組み【サービス・プロダクト提供者】 テキスト生成 AI は新規性が高い技術であり、利用者側の理解度・リテラシーが追いついておらず、そのためサービス・プロダクト提供者が想定していない使い方を利用者が取り得る場合が想定される。これらを踏まえ、「生成 AI は確率的な処理を伴う技術であるため 100%の正確性を保証するものではない」という点を利用者にあらかじめ誤解なく伝え、生成 AI の性質・特徴を利用者側がしっかりと理解した上でサービス・プロダクトを使用できる環境を整えることが重要である。
4．ヘルスケア領域における個別規制	ヘルスケア領域において生成 AI を活用したプロダクトやサービスを開発・提供する場合には、生成AIを活用することによって追加された機能やサービスそれ自体が医療機器プログラムに該当するなど、既に個別制度による規制が適用されることが考えられる。そのため、ヘルスケア領域

	における個別規制の内容をあらかじめ確認するとともに、生成 AI を活用したプロダクト・サービスの機能や性能がそれらの規制を遵守するように開発・提供の段階で十分に注意を払うことが重要である。

4. 生成 AI を活用したヘルスケアサービスの提供を行う場合のチェックポイント

3.に示したとおり、生成 AI を活用してサービス提供する際の課題は大きく「1. モデル選定」、「2. データの取り扱い」、「3. アウトプットの信頼性」、「4. ヘルスケア領域における個別規制」の 4 つに大別できる。これらの論点を踏まえ、実際にサービス設計・開発・提供を行う際に、サービス提供者がチェックすべきポイントを以下のとおり整理する。

なお、別添においてチェックリストを整理しているので、サービス・プロダクト提供者におけるセルフチェックに活用されたい。

4-1. モデル選定に関するチェックポイント

ヘルスケア領域においてサービス・プロダクトを提供するに当たっては、どのような基盤モデルまたは特定モデルを選定するかによってサービス・プロダクトの設計や運用も大きく変わることが考えられる。これまでの論点を踏まえ、モデル選定時のチェックポイントを整理すると以下のとおりである。

なお、基盤モデルが以下に示す情報についての提供義務が規制等で課せられている状況ではなく、また、医療ヘルスケアにおける生成 AI のベンチマーク項目も統一されていない状況であるため、最終的にはサービス・プロダクト提供者側の責任のもと、検討することが肝要である。

【モデル選定に関するチェックポイント】

- 基盤モデルが標榜している性能
- 基盤モデルが定めている利用・用途

4-1-1. 基盤モデルが標榜している性能

基盤モデルが標榜している性能の確認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 基盤モデルの学習データ内容

提供予定のプロダクト・サービスに合った基盤モデルかどうかを判断するため、また、基盤モデルによっては基盤モデルおよび学習データのライセンスの扱いが不明瞭である場合（利用不可なライセンスのデータを

使用している場合等）があるため、利用予定の基盤モデルは、どのようなデータセットをインプットして学習・開発されたモデルかを事前に確認することが望ましい4。また、学習データ内容の確認を行うに当たっては、差別や偏見が再生産されることを避けるため、例えば学習データにジェンダーバイアス等のいわゆる「バイアス」が含まれていないかを確認することも重要なポイントである。この点、昨今では学習データセットにおけるバイアスを評価する論文情報等もあるため、その情報を参考にする等して対応を検討されたい。

② 利用可能形式・価格帯

基盤モデルによってはその公開範囲が各々で異なる。例えば、基盤モデルのモデル全体が公開されており、それをオンプレ上で動かすことができる場合もあれば、クラウド上で API だけが公開されており、API 経由で文章生成はできるがファインチューニングができない場合などさまざまである。そのため、当該基盤モデルの公開範囲によってできること・できないことをあらかじめ確認することが重要である。

また、基盤モデルの性能によって価格帯もさまざまであり、サービス・プロダクト提供に当たっては特に実務面で開発予算等にクリティカルに影響が出るため、注意が必要である。特に API 利用の場合、生成文字数による使用料変動の有無、ファインチューニングを実施する場合の金額感、基盤モデルをサーバーで活用する場合の金額感など、各々の場面でどの規模で料金が発生するのかも確認しておくと総合的な判断が決定されやすいと考えられる。

③ 性能評価報告レポートなどによる用途・性能

利用予定の基盤モデルの性能・機能等を客観的に判断するため、当該基盤モデルにおける性能評価報告レポートではどのような評価がされているかを確認することが望ましい[5]。ただし、ヘルスケア領域での生成 AI 活用事例数は現在頻出している状況ではなく、当該領域における性能評価が実施されているケースは多くない状況であるため、今後生成 AI 活用が促進されるにしたがって客観的に性能評価を確認できる状況に変化していくことが想定される。

④ 基盤モデルの類型

基盤モデルがどの類型に該当するかを理解・確認することで、提供予定のプロダクト・サービスに合った基盤モデルなのかを判断する材料とすることができる。そのため、診断・治療・予防目的ではないことを前提とした上でヘルスケアに関わる専門的な情報をサービスを通じて提供することになる場合は、提供するサービ

[4] 基盤モデル事業者によっては、学習データの内容やアップデート情報について公表・提供している場合があるため、それらを閲覧する方法が想定される。

[5] 各基盤モデル事業者がサービス提供に当たって提供している情報や、アカデミアが整理・公表している情報などが存在している。

スの性能および特徴に応じたモデルを選定することが望ましい。なお、基盤モデルの類型例は例として以下のとおり分類されるので、選定時の参考にされたい。

なお、基盤モデルを動かす主体となるサーバーが国内に設置されているか、国外に設置されているかによって準拠法が異なってくるケースも想定されるため、併せて確認しておくことが望まれる。

＜基盤モデルの類型例＞

- クラウドベースで汎用的な自然言語処理が可能なモデル
- クラウドベースで医療・ヘルスケアに関する専門的な自然言語処理に特化したモデル
- ローカル／オンプレミスで汎用的な自然言語処理が可能なモデル
- ローカル／オンプレミスで医療・ヘルスケアに関する専門的な自然言語処理に特化したモデル

4-1-2. 基盤モデルが定めている利用用途

基盤モデルが定めている利用用途についての確認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 基盤モデルの利用用途範囲

基盤モデルによっては、その利用規約において医療情報の提供を制限等している場合や、明示的に商用利用が限定されているものもあるため、当該基盤モデルの利用規約の内容を事前に確認することが望ましい[6]。

② 入出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、インプットデータ・アウトプットデータが追加学習の対象になっている場合もあるため、利用者が入出力するデータについての学習利用に関する規約の内容をあらかじめ確認することが望ましい。

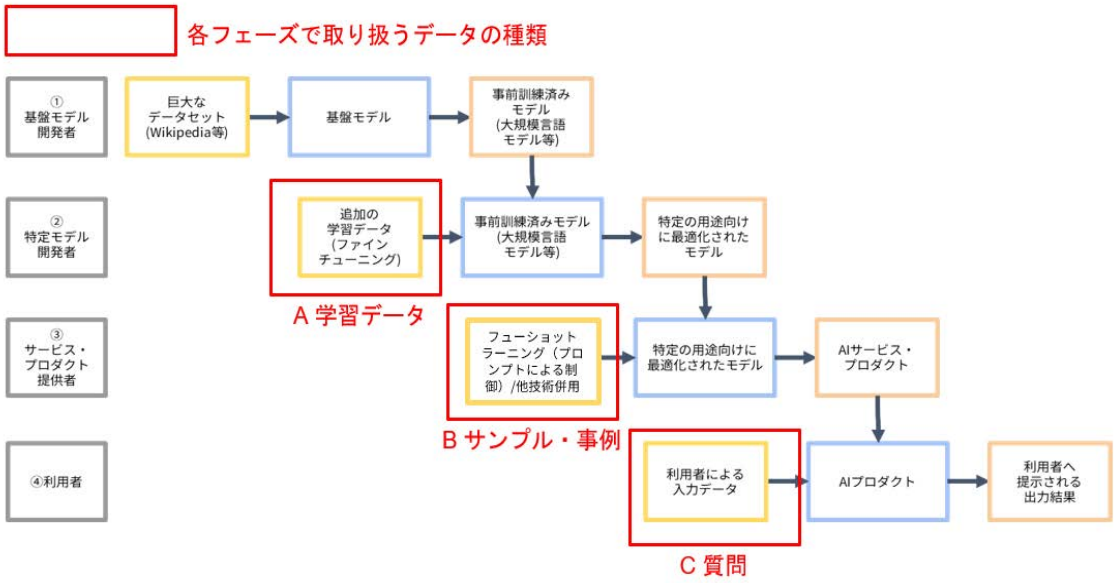
4-2. データの取り扱いに関するチェックポイント

サービス・プロダクト提供者が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際、医学分野の専門的な情報や健康・症状等に関する機微な情報を扱うことが想定されるため、データの取り扱いには特に注意を払う必要がある。以下では前項にて選定した基盤モデルを利用した個別サービスの開発に関わる領域でのチェックポイントを整理する。なお、本項において取り扱うデータの種類については図 12 のとおりであり、各々のデータの取り扱いのチェックポイント概要を図表化すると図 13 のとおりとなるの

6 なお、アウトプットが他人の著作物と類似している等、著作権侵害を生じてしまった場合、基盤モデル提供会社との契約等によって補償を受けられる可能性もあるため、併せて確認しておくことも望ましい。

で、併せて参考にされたい。

図 12：データの種類



(出典：株式会社 MICIN 提供資料)

図 13：データの取り扱いに関するチェックポイント概要

データの種類	誰がどのような時に使うデータか	チェックポイントの概要					誰が守るべきチェックポイントか
		基盤モデルへの反映がされないかの確認	設定オプションの活用	個人情報が含まれている場合の対応	その他著作物等が含まれている場合の対応	その他考慮事項	
A 学習データ	**②特定モデル開発者**がファインチューニングの際に取り扱うデータ	学習データが**基盤モデル**に反映されないよう基盤モデルの利用規約を確認すること	**基盤モデル**に設定オプションがある場合は学習データが**基盤モデル**に反映されないよう設定する	学習データに個人情報が含まれている場合は原則同意を取得する	学習データに論文情報等が含まれる場合は学習に用いることが禁止されていないか確認する	社内におけるデータ保護の体制整備等	特定モデル開発者
B サンプル・事例	**③サービス・プロダクト提供者**がフューショットラーニング（プロンプト制御）の際に取り扱うデータ	サンプル・事例が**基盤モデルor特定モデル**に反映されないよう、**基盤モデルor特定モデルの利用規約**を確認すること	**基盤モデル or 特定モデル**に設定オプションがある場合は、サンプル・事例が**基盤モデルor特定モデルに**反映されないよう設定する	サンプル・事例に個人情報が含まれる場合は原則同意を取得する	サンプル・事例に論文情報等が含まれる場合はフューショットラーニングに用いることが禁止されていないか確認する		サービス・プロダクト提供者
C 質問	**④利用者**がサービス・プロダクトを利用する際に入力するデータ	質問データが**基盤モデルor特定モデル**に反映されないよう、**基盤モデルor特定モデルの利用規約**を確認すること	サービス・プロダクトの設計として利用者による質問内容が**特定モデルに反映されないよう**設定オプションを追加する（利用者でON/OFFができる）	利用者が入力した質問内容に個人情報が含まれる場合には原則同意が必要。出力以外の学習にも利用する場合は学習利用目的の特定が必要。	生成AIへの利用が禁止された著作権物等の情報を入力した場合の免責事項の作成		サービス・プロダクト提供者

（出典：JaDHA WG4 SubWG-B にて作成）

【データの取り扱いに関するチェックポイント】

- 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
- フューショットラーニングに利用する事例・サンプルの取り扱い
- 利用者が入力する質問データの取り扱い
- データに関するその他の考慮事項

4-2-1. 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

ファインチューニングに利用するデータの取り扱い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 基盤モデルの利用規約の確認

基盤モデルによってはファインチューニングに利用する学習データセットの内容を基盤モデルそれ自体に反映する場合もあり、特定モデル開発者が意図せずファインチューニングの目的以外に学習データが活用される可能性が考えられる。そのため、基盤モデルの利用規約等でファインチューニングに使用された学習データの取り扱いに関する事項が明瞭化されているか等、ファインチューニング以外で学習データが利用されることがないかを確認することが望ましい。

② 基盤モデルの設定オプションの確認

上記①で記載のとおり、特定モデル開発者が意図しないところでの学習データが活用されることを予防するため、基盤モデルにおいて設定オプションが用意されている場合は、当該学習データセットが基盤モデルに反映されないような設定がされているか確認することも望ましい措置のひとつである。なお、対応が難しい場合、基盤モデルの再選定を行うことやオンプレミスの基盤モデルを自社学習させることが検討できているかも併せて確認しておくとよい。

③ 学習データにおける個人情報の取り扱い

個人情報の収集に当たっては、利用目的の本人への通知または公表が必要（個人情報保護法第 21 条）であり、収集する個人情報に個人の病歴等要配慮個人情報が含まれている場合は、原則本人の同意が必要（個人情報保護法第 20 条）であることから、学習データとして個人情報を収集し、基盤モデルに入力して学習させる場合は本人の同意を得られているか確認しておくことが必要である。

なお、要配慮個人情報が学習データに含まれる場合のデータ管理等は厳格かつ適切に実施する必要がある一方で、事業者にとっては一定の管理・運用コストが想定されるため、運用の観点から特定モデル事業者にあっては、学習データの中に要配慮個人情報が含まれない、または、できる限り含まれないように事前に考慮されていることが望ましい。それでもなお、学習データに要配慮個人情報が含まれる場合

は、できる限り即時にまたは学習用データセットに加工する前に、要配慮個人情報を削除または特定の個人を識別できないようにするための措置が講じられていることが望ましい。

④ 学習データにおける著作権情報等の取り扱い

学習データの内容として例えば論文情報等を活用する場合も想定されるが、著作権侵害等のリスクをできる限り低減させるため、学習前に当該論文情報等のデータベース（例えばインターネット上のサイト等）において AI の学習に用いることを禁止されていないか等、学習データとしての利用制限がないか確認しておくことが望ましい。また、取引先を含む他社から提供を受けて自社が保有する情報等、いわゆる秘密・機密情報に該当する情報を学習データとして活用する場合においては、提供を受けたデータを学習利用することが許容する条項が含まれているか、相手方との契約内容を事前に確認することが望まれる。

4-2-2. フューショットラーニング等に利用するサンプル・事例の取り扱い

フューショットラーニングや強化学習に利用するサンプル・事例の取り扱い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 基盤モデルまたは特定モデルの利用規約の確認

「4-2-1.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」におけるチェックポイント①に記載の趣旨と同様、サービス・プロダクト提供者が意図しない範囲でサンプル・事例が活用されることを予防するため、フューショットラーニングや強化学習に利用するサンプル・事例が基盤モデルまたは特定モデルに反映されないよう、両者の利用規約等でフューショットラーニング等に利用するサンプル・事例の取り扱いに関する事項が利用規約で明瞭化されているか等、フューショットラーニング以外でサンプル・事例が利用されることがないかを確認することが望ましい。

② 基盤モデルまたは特定モデルの設定オプションの確認

上記①で記載のとおり、サービス・プロダクト提供者が意図しない範囲でサンプル・事例が活用されることを予防するため、基盤モデルまたは特定モデルにおいて設定オプションが用意されている場合は、当該サンプル・事例が基盤モデルまたは特定モデルに反映されないような設定がされているか確認することも望ましい措置のひとつである。

③ サンプル・事例における個人情報の取り扱い

サンプル・事例において個人情報や要配慮個人情報が含まれる場合、「4-2-1.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」におけるチェックポイント③に記載の趣旨と同様、本人への通知または公表や、同意が必要であり、運用面における考慮事項についても併せてサービス・プロダクト提供者において確認されたい。

④ サンプル・事例における著作権情報等の取り扱い

サンプル・事例において論文情報等の著作権情報や機密・秘密情報が含まれる場合、サービス・プロダクト提供者においては「4-2-1.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」におけるチェックポイント④に記載の対応を行うことが望ましい。

4-2-3. 利用者が入力する質問データの取り扱い

利用者が入力する質問データの取り扱い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 基盤モデルまたは特定モデルの利用規約の確認

「4-2-1.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」におけるチェックポイント①に記載の趣旨と同様、利用者が意図しない範囲で入力した質問内容が活用されることを予防するため、利用者が入力した質問内容が基盤モデルまたは特定モデルに反映されないよう、基盤モデルまたは特定モデルの利用規約等で質問データの取り扱いに関する事項が利用規約で明瞭化されているか等を確認することが望ましい。

② サービス・プロダクトにおける設定オプションの確認

利用者の意図しない範囲で自己の質問内容が特定モデルに反映されることがないようにするため、サービス・プロダクト提供者においてサービス・プロダクトの設定オプションが用意できる場合、利用者によって入力された質問内容が特定モデルに反映されないような設定（利用時に質問データの反映に関する「ON／OFF」が切り替えられる設定が措置されている等）を準備することが望ましい。この場合、利用者によるすべての質問データが常に利用できるわけではないことを前提にサービス開発段階において利用者がオプトアウト（利用を拒否）した場合に、オプトアウト時点で学習データとして活用されないように変更できるサービス設計・ファインチューニング等に活用するサンプル・事例として利用されないような具体的な設計を考えておくことも必要である。

③ 質問データにおける個人情報の取り扱い

利用者の質問データにおいて個人情報や要配慮個人情報が含まれる場合、「4-2-1.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」におけるチェックポイント③に記載の趣旨と同様、本人への通知ま

たは公表や、同意が必要である。この点、ヘルスケア領域において生成 AI を活用したサービス・プロダクトを提供する場合、利用者において自己の既往歴や病状を含む健康状態等、個人情報や要配慮個人情報に該当するおそれのある情報が質問内容として入力される場面が多くなることを想定し、特に取り扱いには注意が必要である。また、質問データを出力以外の用途、例えばサービス・プロダクトの精度向上を目的とした学習に利用することを想定している場合、質問データを学習に利用する想定である旨の目的の特定を行う必要がある。

④ 質問データにおける著作権情報等の取り扱い

利用者において自由に質問を入力できるため、例えば生成 AI への利用が禁止された著作物等の情報を利用者が質問として入力する場面も考えられる。利用者がどのような種類・内容の質問を入力するのかの詳細をサービス・プロダクト提供者側でコントロールすることは容易ではないため、生成 AI への利用が禁止された著作物や機密・秘密情報が入力された場合の情報の取り扱いについて、サービス・プロダクト提供者において事前に免責事項を作成・公開しておくことも重要である。

⑤ 質問データにおける悪用・不正目的対策

利用者が悪用や不正を目的とした質問データを組成することで、例えばマルウェア製造といったサービスの意図しない出力を利用できてしまう場合がある。機械的な処理で悪用・不正目的の入力をリジェクトすることは技術的な制約があるため、サービス・プロダクト提供者において事前に免責事項として公開する範囲に、当該用途での生成指示入力を禁じることが重要である。また、サービス・プロダクト提供者の提供するシステムで入力・出力の履歴を保持し、不正利用のトレースを可能としておくこと、ならびにサンプリングによる定期不正抽出等をシステム運用に組み込むことも対策として考えられる。

4-2-4. データに関するその他の考慮事項

データの取り扱いに当たってのその他のチェックポイントは以下のとおりである。

① データ保護に関する社内体制の構築

個人情報をはじめとするデータの取り扱いに関する制度や規制は環境の変化とともに制度改正や新たな論点についての検討が繰り返される領域である。そのため、個人情報等に関する規制の遵守や最新動向についての情報収集をして適切にサービスが設計・提供できるか担保するため、特定モデル事業者やプロダクト・サービス提供事業者の組織内にデータ保護に関する社内体制が構築されていることが望ましい。

② 関連ガイドライン等の参照

個人情報保護法や著作権法以外にも、官公庁において当該制度の詳細について記載しているガイドラインやハンドブック等が公開されている。これらの関連ガイドライン等を網羅的に参照することで、法律レベルでは抽象度の高い文言についても運用に照らした際の具体的な遵守すべき項目の目安が認識できるため、関連ガイドライン等を参照することも重要なポイントになる。関連ガイドライン等としては、例えば厚生労働省の「医療情報システムの安全管理に関するガイドライン」および経済産業省・総務省の「医療情報を取り扱う情報システム・サービスの提供事業者における安全管理ガイドライン」（いわゆる３省２ガイドライン）や、個人情報保護委員会が策定している「個人情報の保護に関する法律についてのガイドライン」、総務省・経産省が策定している「DX 時代における企業のプライバシーガバナンスガイドブック」等が挙げられる。

③ サービスアップデートや機能評価における考慮事項

特にサービス・プロダクト提供者において、サービスのアップデートが行われる場合（例えば利用者が入力する質問データから本人に関する属性等を類推し、機械的に判断するような機能が実装される場合等）に当たっても、サービス設計時に前述のチェックポイントを再び考慮・参照することが望ましい。また、生成 AI を活用したサービス・プロダクトの品質向上や基盤モデルに対する追加学習のためのフィードバックの観点で、提供するサービス・プロダクトの機能評価が実施されるようなサービス設計を行うことも期待されるポイントである。

4-3. アウトプットの信頼性に関するチェックポイント

サービス・プロダクト提供者がアウトプットの信頼性を担保するに当たってのチェックポイントは以下のとおりである。

【アウトプットの信頼性に関するチェックポイント】

- サービス・プロダクト開発段階での取り組み
- サービス・プロダクト提供時の利用者に対する取り組み

4-3-1. サービス・プロダクト開発段階での取り組み

サービス・プロダクト開発段階での取り組み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さまざま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他、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[7]を実施することが挙げられる、

また、エンベディング[8]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[9]）の導入や、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 アウトプットのランダム性に対する工夫の実施

サービス・プロダクトの性質によっては、同じ質問におけるアウトプットのランダム性の必要有無が変わってくることが想定される。例えば、小説生成やチャットボットなどクリエイティブなタスクを想定したサービス・プロダクトでは、毎回多様な回答が好まれることが多い一方で、ヘルスケア領域や法律など規制や制度に深く関わるタスクにおいては同じ入力に対して、同じ回答が出力される方が好まれるケースが多いことが考えられる。そのため、アウトプットのランダム性は、API のパラメータ（例えば「temperature」）などのオプションで指定が可能であるため、サービス・プロダクトの性質によって指定することが重要になる。

③ サービス品質評価の実施

生成 AI を活用して提供したサービス・プロダクトに関して、その技術性能や品質についての評価を実施することで、プロダクトに採用したモデルの技術的な精度等を適切に評価することもアウトプットの信頼性を担保する上で重要である。また、当該サービス・プロダクトの利用場面に応じた適切な評価指標を設計した上で、その評価指標において結果が適切にアウトプットできているか評価する体制・プロダクト設計を行うことも望ましい。さらに、アウトプットの信頼性向上のための取り組みとして昨今では RLHF（Reinforcement Learning from Human Feedback：人間のフィードバックからの強化学習）も注目されている。これは、サービス・プロダクトにおける利用者からの Good／Bad ボタンのように、利用者の行動をフィードバックに活用する取り組みであり、生成 AI を含めた AI 開発の領域においては有効なアプローチである。

---

[7]言語モデルへの命令（プロンプト）を開発・最適化すること。用語集参照。

[8]単語や文といった自然言語の情報を、低次元のベクトルで統一的に表現すること。単語や文が持つ意味の関係性や類似度が計算可能になる技術のこと。用語集参照。

[9] AI が言葉や概念を具体的なものや実際の世界と結びつけて理解する能力のこと。AI が現実世界のコンテキストを理解し、それに基づいた回答や生成を行うことを可能にすること。用語集参照。

4-3-2. サービス・プロダクト提供時の利用者に対する取り組み

サービス・プロダクト提供時の利用者に対する取り組み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 テキスト生成 AI に関する説明・表示

生成 AI の特徴について利用者がよく理解した上でサービス・プロダクトを使うことができるように、チュートリアルの提供等、テキスト生成 AI とはどのようなものなのかについて利用者へ説明・啓発を適切に実施していることが望ましい。

② 利用用途や利用者が入力した質問データの取り扱いに関する説明・表示

サービス・プロダクト提供者が意図しない不適切な用途で利用者がサービスを利用し、それに伴って結果的に利用者に不利益が生じることがないよう、利用者に対して利用用途を説明・表示し、サービスのコンセプトや安全管理措置を明確化することが望ましい。また、利用者が入力した質問データがどのように扱われるのか（学習データとして使う場合があるのか等）の取り扱いの説明・表示をしておくことも重要である。

③ 入力規制・制御の実施

②で記載した趣旨と同様、利用者が不適切な入力をすることがないような仕組みとして、利用者が使用するプロダクトの入力画面等にある程度の入力規制・制御をかけることで利用者に不利益が生じないようなサービスに設計することもひとつのアプローチである。

④ AI による生成であることの表示

利用者が AI による生成であることを理解・認識できるように、利用者に対して、利用者が当該サービス・プロダクトを利用することで得られる出力結果について、その出力結果が AI によるものだと明示される設計になっていることが望ましい。

⑤ 免責条項の表示

何か責任問題が生じた際の責任分界点等を利用者が理解した上で利用できるように利用者に対して、サービス・プロダクト提供者側の免責事項を明示していることも重要である。なお、サービス・プロダクトを他の事業者に提供し、当該事業者から利用者に提供している場合（BtoBtoC の形式での提供の場合）は、アウトプットが他の著作物の権利を侵害することがないかを利用者に対してサービス等を提供する事業者においても確認することが事業者間の契約において盛り込まれていることが望ましい。

4-4. ヘルスケア領域における個別規制に関するチェックポイント

ヘルスケア領域においては、医療機器プログラムの該当有無によって標榜可能な内容・範囲が変わるなど、「プログラム医療機器該当性に関するガイドラインについて」（厚生労働省）をはじめとする個別制度があるため、それらを開発・提供前に十分確認・遵守する必要がある。ヘルスケア領域における個別規制に関するチェックポイントは以下のとおりである。

なお、関連制度については、「4-2-4. データに関するその他の考慮事項」の②のチェックポイントも併せて確認されたい。

① 医療機器プログラムの該当性確認

開発・提供しようとしている生成 AI を活用したプロダクト・サービスは医療機器プログラムに該当するかを確認できていることが望ましい。

② 標榜における広告規制の適合性確認

医療機器プログラムとして判断された場合、承認または認証された範囲内での医学的表現、いわゆる効果・性能の標榜が可能になり、医薬品等適正広告基準や一般社団法人日本医療機器産業連合会など業界が公表しているガイドラインに基づいて広告・表示を行う必要がある。

一方で非医療機器プログラムの場合、医学的な標榜はできず、薬機法および景品表示法による広告・表示規制の対象となるため、これら規制に抵触しないように標榜しなければならない。該当性や広告規制等に関する不明点は各都道府県薬務主管にて相談が可能である。詳細については薬機法、プログラムの医療機器該当性に関するガイドラインや景品表示法等の関連法規を参照されたい。

③ 基盤モデルの利用規約の確認

ヘルスケア領域における生成 AI の活用に当たっては、利用する基盤モデルによってヘルスケア領域における利用制限をはじめとした活用に当たってのさまざまな制約が決められていることもあるため、OpenAI 社の利用規約をはじめとした、基盤モデルの利用規約を確認し、規約のアップデートにより活用の範囲に支障がないかを確認することが望まれる。

5. 今後に向けて - 業界団体としての取り組みと期待-

前述のとおり、昨今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急速に増加する中、一般の利用者が AI に関する専門知識を持たずともこれらのサービスを容易に利用できるようになっている。そのような状況において、生成 AI の特徴・性質から生じる、著作権の帰属や倫理的問

題、セキュリティとプライバシーのリスクやハルシネーションといったさまざまな課題が指摘されている。このような背景を踏まえ、ヘルスケア領域における生成 AI に関する論点を整理し、事業者側で生成 AI を活用してサービス・プロダクトを提供する場合に配慮すべきポイントをまとめたのが本ガイドであり、当該事業者において適切なサービス提供が可能となる環境を整備するための一つのアプローチとして本ガイドが活用されることが望まれる。すなわち、当該事業者が本ガイドを参考にすることにより、当該サービスを提供しようとする事業者が目安・基準となるチェックポイント（参考資料　別添１　ヘルスケア事業者の生成 AI 活用時のチェックリスト）に基づいたセルフチェックを行い、本ガイドに沿ってサービス提供を行っている旨を公表等することで、当該サービスを利用する利用者（医療機関等）が、安心してサービスを利用できるようになることも期待しているところである。

なお、急速な新技術の進歩に伴い、生成 AI の技術特性やその活用方法、関連の制度や規制内容の変化・進化が想定されることから、本ガイドは適宜見直しを図り、事業者がタイムリーに適正なサービスを提供できるよう、今後も業界全体の後押しを行っていく予定である。

デジタル技術を活用したヘルスケアサービスの浸透によって、疾患の早期発見・治療、未病・予防に寄与し、利用者の健康意識の高まりが国民の健康増進に期待できると考えられる。日本デジタルヘルス・アライアンスでは、業界全体で本ガイドの普及・浸透を図り、生成 AI を活用したヘルスケアサービスが安全・安心な形市場へ提供されることへの貢献を今後も目指していく次第である。

参考資料

用語集

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

別添２　ヘルスケア領域における生成 AI 活用事例集

別添３　本ガイドの検討経緯および参加・協力企業等

用語集

用語	用語の定義・意義
生成 AI	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。
テキスト生成 AI	生成 AI のうち、文章を生成するアルゴリズムのこと。
ヘルスケアサービス	健康の保持および増進、介護予防を通じた 健康寿命の延伸に資する商品の生産もしくは販売または役務の提供を行うこと。本自基準においては、そのうち医療機器または医療機器プログラムには該当しないものを指す。 参考文献：経済産業省「ヘルスケアサービスガイドライン等のあり方 」（2019 年 4 月 12 日 （2023 年 6 月 9 日改訂） https://www.meti.go.jp/policy/mono_info_service/healthcare/210609guide.pdf
基盤モデル	特定のタスクに限定せず大量のデータから汎用的に学習した機械学習モデルのこと。
大規模言語モデル（ Large Language Models：LLM）	生成 AI のうち、大規模データを学習させた言語モデルのこと。
ファインチューニング	学習済みの基盤モデルに対して別のデータセットを活用して追加学習させること。
ハルシネーション	生成 AI において、事実と異なる内容（嘘）や文脈と全く無関係な内容の出力が生成されること。
プロンプト入力	生成 AI において、ユーザーが入力する指示や質問のこと。
プロンプトエンジニアリング	生成 AI において、言語モデルへの命令（プロンプト）を開発・最適化すること。

フューショットラーニング	フューショットラーニング（Few-Shot Learning）は、比較的大量のデータを必要とする従来のファインチューニング手法と対照的に、非常に少量のデータを使用して機械学習モデルに推論を行わせる手法のこと。 参考文献：Few-shot learning in practice: GPT-Neo and the Accelerated Inference API https://huggingface.co/blog/few-shot-learning-gpt-neo-and-inference-api
エンベディング	単語や文といった自然言語の情報を、低次元のベクトルで統一的に表現すること。単語や文が持つ意味の関係性や類似度を、数値演算によって計算可能になる。
グラウンディング	グラウンディング（Grounding）とは AI が言葉や概念を具体的なものや実際の世界と結びつけて理解する能力を指す。この概念は AI が単なるデータや情報の集合ではなく現実世界のコンテキストを理解し、それに基づいた回答や生成を行うことを可能にする。そのため生成 AI が誤った情報や関連のない内容を生成すること（ハルシネーション）を防ぐための手法としても用いられる。 参考文献：グラウンディング（Grounding）とは？その意味とビジネスへの影響 https://www.salesforce.com/jp/blog/jp-what-is-grounding/#:~:text=%E3%82%B0%E3%83%A9%E3%82%A6%E3%83%B3%E3%83%87%E3%82%A3%E3%83%B3%E3%82%B0%E3%81%A8%E3%81%AF%E3%80%81AI%E3%81%8C%E8%A8%80%E8%91%89%E3%81%A8%E5%AE%9F%E4%B8%96%E7%95%8C%E3%82%92%E9%96%A2%E9%80%A3%E4%BB%98%E3%81%91%E3%81%A6%E7%90%86%E8%A7%A3%E3%81%99%E3%82%8B%E3%81%93%E3%81%A8

ChatGPT	OpenAI によって開発された対話型の人工知能（AI）ツール。 ChatGPT は自然言語での質問応答、チャットボット機能、文章の生成や翻訳など多岐にわたる用途に対応している生成 AI の一種であり、事前に学習されたトランスフォーマーモデルに基づいて機能する。
要配慮個人情報	「要配慮個人情報」とは、本人の人種、信条、社会的身分、病歴、犯罪の経歴、犯罪により害を被った事実その他本人に対する不当な差別、偏見その他の不利益が生じないようにその取り扱いに特に配慮を要するものとして政令で定める記述等が含まれる個人情報をいう。 参考情報：個人情報の保護に関する法律についてのガイドライン（通則編）（案） https://www.soumu.go.jp/main_content/000450626.pdf
Transformer	Transformer は AI の性能を向上させるための深層学習（ディープラーニング）モデルの一つで、特に自然言語処理（NLP）の分野で重要な役割を果たしている。このモデルは後に NLP 分野においてブレイクスルーをもたらした BERT や GPT-2 などのモデルの基盤となっており、ChatGPT をはじめとする文章生成のジェネレーティブ AI や自然言語処理の理解に不可欠なモデルである。このモデルの仕組みや特徴により生成 AI は高性能な自然言語処理モデルを実現している。 参考情報：Attention Is All You Need https://arxiv.org/abs/1706.03762
BERT	BERT（Bidirectional Encoder Representations from Transformers）は 2018 年に Google により発表された自然言語処理（NLP）に特化した深層学習モデルの一つ。2022 年から 2023 年にかけて生成 AI が流行する以前は、医療分野においてはこのモデルを医療関連のデータでファインチューニングして利用することが多かった。 参考情報：BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding https://arxiv.org/abs/1810.04805

GPT-3	GPT-3（Generative Pre-trained Transformer 3）は、OpenAI 社によって 2020 年に開発された言語モデル。 参考情報：A Survey of Large Language Models https://arxiv.org/abs/2303.18223
GPT-4	GPT-4（Generative Pre-trained Transformer 4）は、OpenAI 社によって 2023 年に開発された言語モデル。 参考情報：A Survey of Large Language Models https://arxiv.org/abs/2303.18223
PaLM	PaLM は、Google 社によって 2022 年に開発された言語モデル。 参考情報：A Survey of Large Language Models https://arxiv.org/abs/2303.18223
LLaMA	LLaMA は、Meta 社によって 2023 年に開発された言語モデル。 参考情：A Survey of Large Language Models https://arxiv.org/abs/2303.18223
プログラム医療機器	2014 年 11 月 25 日に施行された「医薬品、医療機器等の品質、有効性及び安全性の確保等に関する法律」（医薬品医療機器等法）では、国際整合性等を踏まえて、疾病の診断・治療等を目的とした単体プログラム（ソフトウェア）についても医療機器としての規制対象としており、多くのプログラム医療機器が開発され、製造販売承認等されている。 参考情報：PMDA プログラム医療機器について https://www.pmda.go.jp/review-services/drug-reviews/about-reviews/devices/0048.html#faq

別添１　ヘルスケア事業者の生成 AI 活用時のチェックリスト

（別添 Excel を参照）

別添２　ヘルスケア領域における生成 AI に関する取り組み

事例１　医師向け臨床支援アプリ HOKUTO（株式会社 HOKUTO）

インプットから臨床現場のアウトプットまで医師の医学情報収集をフルサポートする情報収集アプリ。本サービス内において、①患者への説明文生成 AI（患者や家族に対して病状や治療内容を説明する文章を AI が簡潔で分かりやすくまとめる機能）、②論文検索＆要約 AI（キーワードと期間を入力するだけで論文が検索され、検索結果の要約等のコメントを AI が生成する機能）を生成 AI を活用した機能として展開。

（出典：株式会社 HOKUT JaDHA SubWG-B 定例会資料（2023 年 9 月 4 日））

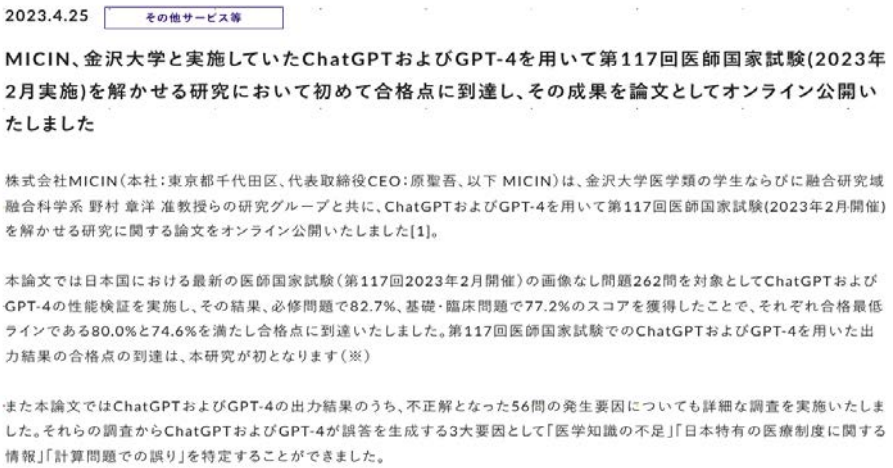
事例２　AI 問診サービス「ユビーメディカルナビ」（Ubie 株式会社）

医療機関の業務効率化を支える AI 問診サービス。ユビーメディカルナビの一機能として、生成 AI の活用により患者の症状や自由回答を LLM が要約する「問診要約機能」をクリニック向けに提供。

（出典：Ubie 株式会社 JaDHA SubWG-B 定例会資料（2023 年 11 月 6 日））

事例 3　医師国家試験と生成 AI に関する研究（株式会社 MICIN）

金沢大学医学類の学生および融合研究域融合科学系 野村 章洋 准教授らの研究グループと共同で ChatGPT および GPT-4 を用いて第 117 回医師国家試験(2023 年 2 月開催)を解かせる研究を実施。その結果、必修問題で 82.7%、基礎・臨床問題で 77.2%のスコアを獲得したことで、それぞれ合格最低ラインである 80.0%と 74.6%を満たし合格点に到達する結果に。



2023.4.25　その他サービス等

MICIN、金沢大学と実施していたChatGPTおよびGPT-4を用いて第117回医師国家試験(2023年2月実施)を解かせる研究において初めて合格点に到達し、その成果を論文としてオンライン公開いたしました

株式会社MICIN（本社：東京都千代田区、代表取締役CEO：原聖吾、以下 MICIN）は、金沢大学医学類の学生ならびに融合研究域融合科学系 野村 章洋 准教授らの研究グループと共に、ChatGPTおよびGPT-4を用いて第117回医師国家試験(2023年2月開催)を解かせる研究に関する論文をオンライン公開いたしました[1]。

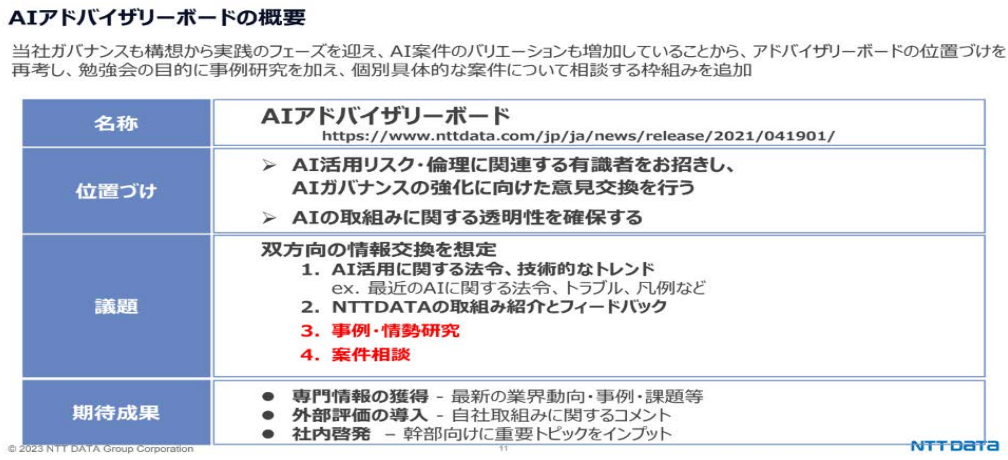
本論文では日本国における最新の医師国家試験（第117回2023年2月開催）の画像なし問題262問を対象としてChatGPTおよびGPT-4の性能検証を実施し、その結果、必修問題で82.7%、基礎・臨床問題で77.2%のスコアを獲得したことで、それぞれ合格最低ラインである80.0%と74.6%を満たし合格点に到達いたしました。第117回医師国家試験でのChatGPTおよびGPT-4を用いた出力結果の合格点の到達は、本研究が初となります（※）

また本論文ではChatGPTおよびGPT-4の出力結果のうち、不正解となった56問の発生要因についても詳細な調査を実施いたしました。それらの調査からChatGPTおよびGPT-4が誤答を生成する3大要因として「医学知識の不足」「日本特有の医療制度に関する情報」「計算問題での誤り」を特定することができました。

（出典：株式会社 MICIN「NEWS」（2023 年 4 月 25 日）https://micin.jp/news/10073（参照 2024 年 1 月 15 日））

事例 4　AI ガバナンス体制の構築（株式会社 NTT データ）

AI 活用リスクや倫理に関連する有識者を構成員とする AI ガバナンス体制を社内に構築し、AI に関する法令や技術的なトレンドのインプットや自社の取り組み紹介・フィードバック等を実施。AI 案件のバリエーションの増加を踏まえ、事例研究や個別具体的な案件について相談できる枠組みも追加。



AIアドバイザリーボードの概要

当社ガバナンスも構想から実践のフェーズを迎え、AI案件のバリエーションも増加していることから、アドバイザリーボードの位置づけを再考し、勉強会の目的に事例研究を加え、個別具体的な案件について相談する枠組みを追加

名称	**AIアドバイザリーボード** https://www.nttdata.com/jp/ja/news/release/2021/041901/
位置づけ	➢ **AI活用リスク・倫理に関連する有識者をお招きし、AIガバナンスの強化に向けた意見交換を行う** ➢ **AIの取組みに関する透明性を確保する**
議題	**双方向の情報交換を想定** **1. AI活用に関する法令、技術的なトレンド** ex. 最近のAIに関する法令、トラブル、凡例など **2. NTTDATAの取組み紹介とフィードバック** **3. 事例・情勢研究** **4. 案件相談**
期待成果	● **専門情報の獲得** - 最新の業界動向・事例・課題等 ● **外部評価の導入** - 自社取組みに関するコメント ● **社内啓発** – 幹部向けに重要トピックをインプット

NTT DATA

（出典：株式会社 NTT データ「NTT データグループの AI ガバナンス（SuBWG-B 定例会資料）」（2023 年 10 月 10 日））

別添３　本ガイドの検討経緯および参加・協力企業等

１．検討経緯

本ガイドは、2023 年 3 月に設立された日本デジタルヘルス・アライアンス（以下「JaDHA」という。）における WG４（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）内に設置している SubWG-B（生成 AI に関する検討）において作成・検討が進められた。具体的な検討経緯は以下のとおり。

	アジェンダ	登壇企業
2023 年 7/31	・生成 AI に関する最新動向のインプット①（ヘルスケア生成 AI の基本的動向、特徴や限界等について）	・Ubie 株式会社 ・株式会社 MICIN
8/7	・生成 AI に関する最新動向のインプット②（ LLM の選定方法や主体の区分整理、実装のポイント等について）	株式会社サイバーエージェント
8/23	・生成 AI に関する最新動向のインプット③（サービスへの具体的活用事例および課題・リスクについて）	株式会社 HACARUS
9/4	・生成 AI に関する最新動向のインプット④（サービスへの具体的活用事例および課題・リスクについて）	株式会社 HOKUTO
9/25	・生成 AI に関する海外動向インプット ・論点整理素案について議論	－
10/10	・生成 AI に関する最新動向インプット⑤ ・ガイド（案）意見交換	株式会社 NTT データ
10/23	・生成 AI の最新動向に関する勉強会（生成 AI 活用時における法的論点等について）	・JaDHA 顧問弁護士 後藤 未来様、アンダーソン・毛利・友常法律事務所　中崎 尚 様 ・一般社団法人 日本ディープラーニング協会（JDLA）「AI データと個

		人情報保護」研究会 副座長 小峰 弘雅 様（株式会社ベイカレントコンサルティング）、柴山 吉報 様（阿部・井窪・片山法律事務所）
11/6	・生成 AI の最新動向の共有 ・ガイド策定に向けた論点整理案についての議論	–
11/13～11/24	JaDHA 会員向け意見募集	–
12/4	SubWG-B 最終とりまとめ	–
12/15	WG4 最終審議	–

2．検討主体

前述のとおり、本ガイドは JaDHA における WG4 SubWG-B の参加企業において作成・検討を進めたもの。作成当時の SubWG-B メンバーは以下のとおり。

SubWG-B 参加企業 （五十音順）	担当者 （敬称略）	ガイド作成担当領域
株式会社 Welby	井伊 尋幸	案執筆、レビュー
小野薬品工業株式会社	小林 正克、片山 英夫	案執筆、レビュー
株式会社サイバーエージェント	窪田 海人	主執筆
シスメックス株式会社	辰巳 真一	レビュー
シミック株式会社	三友 周太、齊藤 卓弥、笹 澄絵	案執筆、レビュー
株式会社 Save Medical	小林 亮太	レビュー
タウンドクター株式会社	山上 慶	レビュー

株式会社テックドクター	湊　和修	レビュー
株式会社 MICIN	砥﨑　裕晃、浅原　弘明	主執筆、レビュー、「生成 AI におけるバリューチェーン」図表提供
Ubie 株式会社 （SubWG-B リーダー企業）	三浦　萌、島津　真夢	主執筆、全体統括
株式会社日本総合研究所 （JaDHA 事務局）	南雲　俊一郎、城岡　秀彦、川舟　広徒	案執筆

3．協力・登壇企業等

その他、本ガイド作成に当たっては以下の皆様からアドバイス・ご意見をいただきました。未筆ではありますが御礼申し上げます。

- 株式会社 HACARUS 様
- 株式会社 HOKUTO 様
- 株式会社 NTT データ様
- JaDHA 顧問弁護士 / アンダーソン・毛利・友常法律事務所（AMT）　後藤 未来様
- アンダーソン・毛利・友常法律事務所（AMT）中崎 尚様
- 一般社団法人日本ディープラーニング協会「AI データと個人情報保護」研究会 副座長/株式会社ベイカレントコンサルティング　小峰 弘雅様
- 一般社団法人日本ディープラーニング協会 「AI データと個人情報保護」研究会 副座長/ 阿部・井窪・片山法律事務所　柴山 吉報様
- JaDHA 会員の企業の皆様