

自治体向け AI ガバナンスガイドライン

目次

はじめに.....	3
本ガイドラインの目的	3
ガイドラインの構成.....	4
用語集	5
第 1 部 AI の基本情報	8
1-1. AI の特徴	8
1-2. 自治体における AI 活用の状況.....	9
1-3. AI リスクマネジメント動向	11
第 2 部 自治体における AI 関係者の定義	15
2-1. AI 提供者.....	15
2-2. AI 利用者.....	15
第 3 部 AI ガバナンス.....	17
3-1. AI ガバナンスの重要性.....	17
第 4 部 自治体 AI 倫理原則	19
4-1. 人間中心	19
4-2. 住民第一	20
4-3. 安全・安心	20
4-4. 公平性	20
4-5. プライバシー保護	21
4-6. セキュリティ確保	21
4-7. 透明性	21
4-8. アカウンタビリティ	22
4-9. 教育・リテラシー	22
4-10. イノベーション	22
第 5 部 自治体 AI 関係者の留意点	23
5-1. AI 提供者における留意点.....	23
5-2. AI 利用者における留意点.....	31
第 6 部 自治体における AI 活用の今後の展望	40

はじめに

本ガイドラインの目的

近年、人工知能（AI）の進化は目覚ましく、全世界において、その利用が急速に拡大している。特に対話型生成 AI の登場により、専門的な知識や技術を持たない者でも AI を容易に活用することが可能となった。

この革新は、地方自治体における AI の積極的な導入をも促している。日本の人口動態は、少子高齢化と人口減少という特徴を有し、これにより今後、労働力不足が深刻化することが予想される。だが、地方自治体の業務量がそれに比例して減少するわけではない。このような状況下において、行政サービスの質を維持するためには、業務の効率化が不可欠であり、AI はその導入の容易さと効果の面から、非常に有効な手段として位置づけられている。さらに、日本の人口動態に伴う多岐にわたる課題、例えば医療や介護の負担増加、地方創生などに取り組む際にも、AI の活用は強力な助力となり、これはまた、Society5.0 の実現に向けた重要な一歩にもなる。Society5.0 は、サイバー空間とフィジカル空間が高度に融合することで、人々の生活を豊かにする未来社会を目指している¹。地方自治体における AI の導入は、地域特有の課題解決を促進し、持続可能な社会の実現に寄与する。また、逆説的に述べるならば、グローバルに AI 活用が進展する中で、日本の行政における AI の導入が不十分である場合、日本は国際的な競争において後れを取り、住民サービスにおける格差を生む可能性がある。このような期待と危機感を背景に、既に多くの地方自治体では、独自の利用ルールを制定しつつ、AI の活用を進めている。

AI の利用は、その分野と利用形態によっては、社会に対して大きなリスクを生じさせる。そのリスクが社会的な軋轢を生じさせ、AI の利活用を阻害する可能性がある一方で、過度な対策は AI 活用の便益を阻害する恐れもある。そのため、AI を安全かつ安心して活用するためには、AI に関するリスクをステークホルダーが受容可能な水準で管理しつつ、便益を最大化するための AI ガバナンスの構築が重要である。AI ガバナンスの構築にあたっては、利用分野や用途等を考慮してリスクの大きさ（危害の大きさおよびその発生確率）を事前に把握し、その対策の程度をリスクの大きさに応じて調整する「リスクベースアプローチ」が求められる。

このような状況の中、総務省・経済産業省は 2024 年 4 月に「AI 事業者ガイドライン（第 1.0 版）（※以下、AI 事業者ガイドライン）」を取りまとめた。これは先述した Society5.0 の理念に呼応しており、2019 年 3 月に策定された「人間中心の AI 社会原則」をも踏まえている。AI 事業者ガイドラインは、多様な事業活動における AI の安全かつ安心な活用を促進するために、統一的な指針を示したものである²。

¹ https://www8.cao.go.jp/cstp/society5_0/

² <https://www.meti.go.jp/press/2024/04/20240419004/20240419004.html>

本ガイドラインは、AI 事業者ガイドラインをベースにしつつ、地方自治体が提供する行政サービスや保有する情報等を踏まえ、より地方自治体に特化した具体的な内容を記載している。地方自治体が AI ガバナンスを構築するために、AI の提供および利用に関する倫理原則と自治体 AI 関係者の留意点を示すことを目的としている。地方自治体では、AI 利用者として AI システムや AI サービスを利用して住民に行政サービスを提供することに加えて、AI 提供者として AI システムをアプリケーション、製品、既存のシステム、ビジネスプロセス等に組み込んだ AI サービスを構築し、自治体職員や住民にサービスを提供することも考えられる。そのため、AI 提供者および AI 利用者としての地方自治体に求められる事項を明確にし、それを基に AI ガバナンスを構築する必要がある。ゆえに、自治体職員のみならず、自治体職員と共同でシステム等を導入する外部ベンダーや、自治体向けの情報システムを整備する各機関（各府省や独立行政法人等）等、AI を活用する自治体のすべての関係者が、当該ガイドラインの内容を理解することが期待される。

AI ガバナンスの構築により、地方自治体の行政現場において AI の利用が過度に制限されることなく、その特性を理解し最大限に活用することが期待される。これにより、自治体職員の働き方を変革し、行政サービスの向上やイノベーションの創発を達成し、住民の生活の質の向上を目指す。

ガイドラインの構成

本ガイドラインの構成を以下に記載する。

第 1 部

AI に関する基本情報として、AI の特徴、自治体における AI 活用の状況、AI リスクマネジメント動向を記載する。第 1 部以降の内容を理解するにあたり、AI に関する一般的事項を把握することで、AI に対する具体的なイメージを持つことが重要である。

第 2 部

自治体における AI 関係者を定義する。AI を活用するにあたり、どのような関係者が登場するのかを理解し、自らの役割を認識すると同時に、当該役割を全うするために、AI 提供者と AI 利用者における連携の必要性を理解することが重要である。

第 3 部

AI ガバナンスの重要性について記載する。自治体における AI 活用を実践するために必要となるガバナンスの構築については、AI 倫理原則がいわゆる絵に描いた餅とならないよう、実践するための仕組みづくりについても理解を深める必要がある。

第4部

自治体がAIを活用するにあたって理解すべきAI倫理原則について記載する。AI事業者ガイドラインにおける原則をベースとしつつも、自治体業務において意識すべき概念も踏まえ、整理・記載している。そのため、AIを活用する自治体のすべての関係者において内容を確認し、理解することが重要である。

第5部

AIを活用した事業活動を担うAI利用者およびAI提供者それぞれについて、第2部では触れられない主体毎の留意点を記載する。各主体ともに自らに関する事項を理解することが重要であり、それと同時に、その他のAI関係者とも関連する事項が多く存在するため、当該主体以外に関する事項も理解することが重要である。

なお、本ガイドラインに関して、AI利用者やAI提供者となる自治体職員や外部ベンダーはもちろん、首長や部局長など自治体経営に携わる者も本ガイドラインの内容を理解し、自治体運営の総合的な計画と一体でAIガバナンスを実践し、AIの安全安心な活用を推進することが重要である。

また、AIをめぐる環境はグローバル規模で日進月歩の進化を続けていることから、AIを活用する場合は、最新の動向にも注意を払うことが重要である。特に、高度なAIシステムについては、倫理面を含め、あるべき活用方法について日々議論が交わされているところであり、特に留意することが重要である。

用語集

AIはArtificial Intelligence（人工知能）を意味し、1956年にダートマス会議で初めて使用された言葉であるとされている。AIは未だ確立された定義は存在しないが、「人工」・「知能」とあるように、人間の思考プロセスと同じような形で動作するコンピュータープログラム、コンピューター上で知的判断を下せるシステム等を指す。機械学習（Machine Learning、略してML）を行わない、専門家の知識を大量にインプットすることで知識にもとづく推論を行うエキスパートシステムと呼ばれるものも、AIとみなす考えもある。しかし、2000年代以降、ディープラーニング等による「画像認識」、「自然言語処理（翻訳等）」、「音声認識」が活用されるようになり、特定の分野に特化し、予測、提案または決定を行うことができるシステムをAIと指すようになってきた。また、2021年以降、基盤モデルの台頭により、特定の分野のみに特化したAIではない、汎用的なAIの開発が進んでいる。その結果、「予測」、「提案」、「決定」とどまらず、画像、文章等を生成する「生成AI」が普及するようになり、注目を集めている。このように、ひとくりに「AI」と言っても、その種類は多岐にわたり、今後のAI技術の在り方については専門家であっても予測することは困難である。

このような状況を踏まえつつ、本ガイドラインにおける関連する用語を以下のとおり定義する。

AI :

現時点で確立された定義はなく（統合イノベーション戦略推進会議決定「人間中心の AI 社会原則」（2019 年 3 月 29 日））、広義の人工知能の外延を厳密に定義することは困難である。本ガイドラインにおける AI は「AI システム（以下に定義）」自体または機械学習をするソフトウェア若しくはプログラムを含む抽象的な概念とする。

（参考として JIS X 22989:2023 では ISO/IEC 22989:2022 にもとづき、以下のように定義されている）

<学問分野> AI システムのメカニズムおよび適用の研究開発

注釈 1. 研究開発は、コンピュータサイエンス、データサイエンス、自然科学、人文科学、数学等、幾つもの分野にわたって行うことが可能である

AI システム :

活用の過程を通じて様々なレベルの自律性をもって動作し学習する機能を有するソフトウェアを要素として含むシステムとする（機械、ロボット、クラウドシステム等）。

（参考として JIS X 22989:2023 では ISO/IEC 22989:2022 にもとづき、以下のように定義されている）

人間が定義した所与の目標の集合に対して、コンテンツ、予測、推奨、意思決定等の出力を生成する工学的システム

注釈 1. 工学的システムは、人工知能に関連する様々な技法およびアプローチを使用して、作業の実施に使用可能であるデータ、知識、プロセス等を表すモデルを開発することが可能である

注釈 2. AI システムは、様々な自動化のレベルで動作するように設計されている

（参考として OECD AI Principles overview では以下のように定義されている）

AI システムは、明示的または暗黙的な目的のために推測するマシンベースのシステムである。受け取った入力から、物理環境または仮想環境に影響を与える可能性のある予測、コンテンツ、推奨、意思決定等の出力を生成する。AI システムが異なれば、導入後の自律性および適応性のレベルも異なる

高度な AI システム :

最先端の基盤モデルおよび生成 AI システムを含む、最も高度な AI システムを指す。

（広島 AI プロセスでの定義を引用）

AI モデル（ML モデル） :

AI システムに含まれ、学習データを用いた機械学習によって得られるモデルで、入力データに応じた予測結果を生成する。

（参考として JIS X 22989:2023 では ISO/IEC 22989:2022 にもとづき、以下のように定義されている）

入力データまたは情報にもとづいて推論（inference）または予測を生成する数学的構造

例：単変量線形関数 $y = \theta_0 + \theta_1 x$ が、線形回帰を使用して訓練されている場合、結果のモデルは、「 $y = 3 + 7x$ 」のようになる

注釈 1. 機械学習モデルは、機械学習アルゴリズムにもとづく訓練の結果として得られる

AI サービス：

AI システムを用いた役務を指す。AI 利用者への価値提供の全般を指しており、AI サービスの提供・運営は、AI システムの構成技術に限らず、人間によるモニタリング、ステークホルダーとの適切なコミュニケーション等の非技術的アプローチも連携した形で実施される。

生成 AI：

文章、画像、プログラム等を生成できる AI モデルにもとづく AI の総称を指す。

AI ガバナンス：

AI の利活用によって生じるリスクをステークホルダーにとって受容可能な水準で管理しつつ、そこからもたらされる正のインパクト（便益）を最大化することを目的とする、ステークホルダーによる技術的、組織的、および社会的システムの設計並びに運用。

LLM（大規模言語モデル）：

昨今のテキスト生成 AI の背後に多くのニューラルネットワーク型の機械学習モデルの一種で、膨大なパラメータ数と膨大なテキストにより学習されていることが特徴の一つ。

テキスト生成 AI の背後は必ずしも大規模言語モデルではない（機械翻訳での LSTM 等）点や、大規模言語モデルの小型化研究が盛んであり、必ずしも「大規模」とは限らない点が混乱しやすい。

（デジタル庁「テキスト生成 AI 利活用におけるリスクへの対策ガイドブック（α版）」より引用）

プロンプト：

大規模言語モデルへ入力するテキストデータのことを指す。膨大なテキストによる教師なし学習を事前学習とよび、特定の課題（下流タスクとも呼ぶ）に特化させた学習を事後学習（ファインチューニング）とよぶ。以前はこの事後学習によりさまざまな自然言語処理課題に対応させる手法が主流だったが、大規模言語モデルではプロンプトによる指示を変えるだけでさまざまな自然言語処理課題に対応することが可能になった。

（デジタル庁「テキスト生成 AI 利活用におけるリスクへの対策ガイドブック（α版）」より引用）

第 1 部 AI の基本情報

1-1. AI の特徴

これまでに述べたとおり、AI は近年急速に進化し多岐にわたる分野で活用されている。地方自治体においてもこれらの技術を理解し適切に活用することが求められている。

以下で、AI の特徴について詳述する。

AI の特徴

AI はデータに基づいて知識を得て、意思決定や行動を自動化する機能を持つといえる。

本項では「自治体における AI 活用・導入ガイドブック（2022 年 6 月 総務省）³」の分類にならい、AI の機能を「識別」、「予測」、「実行」に分けてその特徴を説明する。

識別：

データやパターンを認識し、分類する能力を指す。音声、画像、テキストなどの input 情報をもとに特定、分類、検出などを行うことで、与えられた情報や環境を理解し、人間や AI 自身が適切なアクションを取るための基盤とすることができる。

予測：

過去のデータやパターンに基づいて未来の出来事や傾向を分析、推測する能力を指す。これまでの実績をもとにこれから起こることを予測、解析することで人間の意思決定の支援や事象に対するリスク軽減を実現することができる。

実行：

特定の作業を自動的に遂行する能力を指す。AI が物理的またはデジタルなタスクを自動的に実行することで、人間の労力を削減し、効率を向上させることができる。

生成 AI の特徴

生成 AI は大量の学習データをもとに様々な形式のデータを自律的に生成できる AI 技術の総称である。その特徴として、次に示すようなものがあるといえる。

コンテンツ生成能力：

生成 AI はテキスト、画像、音声、動画などのコンテンツを自動生成する能力を有する。例えば、報告書

³ https://www.soumu.go.jp/main_content/000820109.pdf

の自動作成や、広報資料の生成などに利用することが可能である。

自然言語処理（NLP）：

生成 AI は自然言語処理技術を駆使して、人間の言語を理解し、自然な会話を行うことが得意といえる。例えば、チャットボットや自動応答システムの構築により、自治体職員や住民とのコミュニケーションを円滑にすることができる。

創造性と柔軟性：

生成 AI は、既存のデータを基に新しいアイデアやコンセプトを創出する能力を持つ。例えば、新しい施策の検討やロゴのデザインを提案することが可能である。

パーソナライズ：

生成 AI は、個々のユーザーのニーズや嗜好に基づいて、パーソナライズされたコンテンツを提供することができる。例えば、住民一人ひとりに最適化されたサービスを提供することが可能である。

1-2. 自治体における AI 活用の状況

自治体においても、AI および生成 AI の活用が進んでいる。以下で、いくつかの事例を紹介する。⁴

事例 1 AI を活用した音声テキスト化サービス	
【埼玉県内の複数の市町村】（機能分類：識別）	
導入の背景、経過、目的など	埼玉県内市町村で構成する会議体において、AI の共同利用に向けた検討を実施。結果、音声テキスト化サービスの共同利用に参加を希望する市町村が複数見込まれたため、LGWAN-ASP として県内市町村で共同利用を開始することとした。 同サービスは会議中の音声データから会話内容を自動でテキストデータ化するもので、議事録作成の効率化を図ることを目的として実施したものである。
取り扱うデータとセキュリティ	会議中の音声データを用いて、会話内容を自動でテキストデータに変換。取り扱うデータには個人情報など公開すべきではないデータも含まれるため、LGWAN 上でのサービス利用に限定することでセキュリティを確保した。また、AI にアップロードされた音声データおよび変換後のテキストデータは一定期間経過後に自動消去される仕組みとなっている。

⁴ いずれも「自治体における AI 活用・導入ガイドブック（2022 年 6 月 総務省）」を参照のもと作成

得られた効果	自治体職員の議事録の作成時間が短縮化され、利用団体の中には作業時間が約 50%削減された事例も確認された。また、定性的な効果として、会話内容がすべてテキスト化されるため、記録の漏れや誤りも防止された。その他、共同利用による導入コストやランニングコスト、事務負担の軽減のほか、団体同士の AI 運用ノウハウの共有も進んだ。
--------	--

事例 2 AI を活用した介護予防	
【福島県いわき市】（機能分類：予測）	
導入の背景、経過、目的など	効率的な介護予防のため、将来的に介護度が急激に高まる可能性のあるハイリスク者を早期に特定する必要があったが、住民が自身の健康状態を用紙に記入し、その内容を基に判断するという従来の方法では正確性やスピードに課題があった。そこで、自治体として保有している住民情報や介護認定情報、障がい情報、特定健診の受診有無など複数課にまたがる情報を活用し、実際に介護認定された方のデータを AI に学習・分析させたうえで、潜在的なハイリスク者を抽出することとした。
取り扱うデータとセキュリティ	介護認定される方に多い条件を AI に学習させ、住民登録された 65 歳以上の全住民の情報とマッチングさせることで、介護リスクの度合いを点数化。学習時には、新規で介護認定された方の約 3,000 件のデータを利用。その際には、庁内の各担当課が保有するデータ（単身世帯かどうかや、障がいの情報、健診の受診有無など）を用いた。セキュリティ面においては、オンプレミス環境として構築するとともに、データの受け渡しも USB は利用しない、などといった対策を講じた。
得られた効果	自治体職員がこれまで気が付かなかった潜在的なリスク保持者の発見につながった。また、定性的な効果として、自治体職員の経験や勘に基づくハイリスク者の判定ではなく、AI によって判定の基準が体系化され、より精度が上がるのが期待できる。

事例 3 AI を活用した住民税の賦課修正業務の効率化実証	
【練馬区・中央区】（機能分類：実行）	
導入の背景、経過、目的など	住民税賦課業務では、併合処理で検出された不整合の修正を短期間で実施する必要があり、業務効率化やノウハウ継承が大きい。

	な課題となっていた。この課題を解決するため、練馬区において税務システム担当ベンダーと実証を開始し、後に中央区も参画して自治体 AI 共同開発推進事業で実証を行った。具体的には、住民税賦課業務において、住民や事業所から提出された各種課税資料（確定申告書や給与支払報告書など）の併合処理（住民や事業所から提出された各種課税資料を住民ごとに合算する処理）に AI を導入。AI は、併合処理で不整合が検出された納税者ごとのデータを入力として、見直すべき資料と修正方法のレコメンドを出力するものとして構築された。
取り扱うデータとセキュリティ	税業務に関する各種課税資料や世帯の扶養状況を把握するためのデータ、併合処理での不整合検出結果のデータを取り扱う。セキュリティ面においては、AI システム内には基本 4 情報および個人番号は取り込まず、照合や特定ができないようフラグ化等の加工を実施。さらに、住民税システムと同じプライベートクラウドに AI を構築し、インターネット接続は行わないことでセキュリティを確保した。
得られた効果	賦課修正の作業時間が短縮され、平均して約 55%削減された。また、定性的な効果として、ベテランの自治体職員にとっては自身が行った作業の答え合わせの役割を AI が担うとともに、経験の浅い自治体職員については作業スピードのアップに寄与し、賦課修正作業の正確性・質の向上に貢献した。

1-3. AI リスクマネジメント動向

AI の活用により社会へ便益がもたらされる一方で、悪用や想定外の事故につながるリスクもある。例えば米国ニューヨーク市が 2023 年に開設した AI チャットボットがユーザーに不法行為を勧めたり、誤情報を伝えたりすることが判明し、批判を受けた。日本でも災害時に生成 AI が作成した偽画像が SNS 上で拡散されて混乱を招いたり、生成 AI を悪用してコンピューターウイルスが作成されたりと、AI にまつわるインシデントが発生している⁵。

このような背景から、AI のリスクマネジメントに関する議論は国際的にも活発化している。以下では近年の代表的なトピックを取り上げる⁵。

⁵ 毎日新聞/2024 年 4 月 6 日/MAIMEP0020240405ek460003y

静岡新聞/2022 年 10 月 28 日/SHIZUO0020221028eias0000t

読売新聞/2024 年 5 月 29 日/YOSHIN0020240529ek5t000cu

■ OECD 「AI に関する OECD 原則⁶」

経済協力開発機構（OECD）は 2019 年 5 月に国際的指針として「AI に関する OECD 原則（以下、AI 原則）」を採択した。これは当時の OECD 加盟国（36 ヶ国）とパートナー国のうち 6 ヶ国の計 42 ヶ国が署名して採択されたものであり、世界で初めて複数国が AI に関して合意した原則である。AI 原則は信頼できる AI の責任ある普及に関する 5 つの原則と、各国政府に対する 5 つの提言で構成されており、法的拘束力はないが国際標準の設定や国内施策の策定に影響力を持つ。なお、昨今の生成 AI の急速な台頭を受け、2024 年 5 月の閣僚理事会で AI 原則の見直しが決定された。改定案には AI 開発者らへ偽・誤情報への対処を求める項目が追加される見通しである。

AI 原則の 5 つの原則を以下に示す。

- 包摂的な成長、持続可能な開発、幸福
- 法の支配、人権、公正さやプライバシーを含む民主主義の価値の尊重
- 透明性と説明可能性
- 堅牢性、セキュリティ、安全性
- 説明責任

■ ISO/IEC 42001 「AI に関するマネジメントシステム規格⁷」

AI が世界的に急速に普及していることを受け、2023 年 12 月に国際標準化機構（ISO）および国際電気標準会議（IEC）により国際規格「AI マネジメントシステム（ISO/IEC 42001）」が発行された。これは AI に関する世界初のマネジメントシステム規格であり、AI システムを開発、提供、使用、監視する組織を対象に、責任ある AI の利用を支援することを目的としている。この規格は、組織内での AI 管理システムの確立、実行、維持および継続的な改善に関する要件を規定している。また、組織がリスクベースアプローチを通じて、イノベーションとガバナンスのバランスを保ちながら、信頼性や透明性、説明責任、継続学習など AI 特有の機能や課題を考慮して管理するためのガイダンスを提供している。本規格は ISO9001 品質マネジメントシステム（QMS）規格や ISO/IEC27001 情報セキュリティマネジメントシステム（ISMS）規格等の既存のマネジメントシステム規格と同様の構成、アプローチを採用しており、利用者を考慮している。

■ 広島 AI プロセス⁸

⁶ <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

https://www.jil.go.jp/foreign/jihou/2023/06/oecd_01.html

京都新聞/2024 年 5 月 5 日/KYOSBJ0020240504ek550005v

読売新聞/2024 年 5 月 1 日/YOSHIN0020240501ek51000a2

⁷ <https://www.iso.org/obp/ui/#iso:std:iso-iec:42001:ed-1:v1:en>

<https://webdesk.jsa.or.jp/common/W10K0620/?id=1124>

<https://www.meti.go.jp/press/2023/01/20240115001/20240115001.html>

⁸ <https://www.soumu.go.jp/hiroshimaaiprocess/index.html>

広島 AI プロセスは、2023 年 5 月の G7 広島サミットの結果を踏まえ、AI の活用や開発、規制に関する国際ルール作りを推進する枠組みとして発足した。AI の潜在的なリスクとその安全性や信頼性を確保するためのガバナンスの必要性が背景にあり、生成 AI は社会経済に与える影響が大きいため、G7 が一丸となって対応する課題としている。

2023 年 12 月には、広島 AI プロセスの成果として、「広島 AI プロセス包括的政策枠組み (Framework)」、「全ての AI 関係者向けの広島プロセス国際指針 (Guiding Principles)」、および「高度な AI システムを開発する組織向けの国際行動規範 (Code of Conduct)」が策定された。これらの文書は、AI の安全で信頼できる使用を促進するための具体的な指針と規範を示している。

広島 AI プロセスを含む国際的な議論を踏まえ、総務省・経済産業省は 2024 年 4 月に AI 事業者ガイドラインを制定した。

■ 米国「人工知能の安全・安心・信頼できる開発と利用に関する大統領令⁹」

2023 年 10 月、米国のバイデン政権は「人工知能の安全・安心・信頼できる開発と利用に関する大統領令 (Executive Order)」を公表した。本 EO は、AI 技術のメリットを把握しつつリスクを適切に管理することを目的としており、「AI 技術の安全・安心の確保」「イノベーションと競争力強化」「労働者支援」「公平と公民権推進」「消費者、患者、交通機関利用者、学生の保護」「プライバシー保護」「連邦政府の AI 活用推進」「米国の国際的リーダーシップ推進」「導入推進」のセクションに分かれている。また、本 EO は連邦政府機関や民間企業を含む多岐にわたるセクターを対象としており、各関係機関に対しリスク評価や報告等を要求している。いくつかの指令には法的強制力が伴う。

■ EU AI 規制法¹⁰

EU は 2024 年 5 月に「EU AI 規制法」を成立させ、8 月に発効した。これは、世界で初となる生成 AI を含む包括的な AI の規制法であり、2025 年から 2030 年までに段階的に施行される。本法は、適用範囲、リスク分類とその分類ごとの要求事項と義務、段階的施行スケジュール、制裁等について定めている。

⁹ バイデン政権, Exec. Order No. 14110, 2023, Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 2023/11/30 閲覧

https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence?_ga=2.36035623.66360718.1732070831-1270836287.1727946081

¹⁰[https://eur-lex.europa.eu/legal-](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689&_ga=2.5774644.66360718.1732070831-1270836287.1727946081)

[content/EN/TXT/PDF/?uri=OJ:L_202401689&_ga=2.5774644.66360718.1732070831-1270836287.1727946081](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689&_ga=2.5774644.66360718.1732070831-1270836287.1727946081)

https://www.europarl.europa.eu/portal/en?_ga=2.5774644.66360718.1732070831-1270836287.1727946081

https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/?_ga=2.5774644.66360718.1732070831-1270836287.1727946081

https://www.europarl.europa.eu/news/en/press-room/20240308IPR19015/artificial-intelligence-act-meps-adopt-landmark-law?_ga=2.5774644.66360718.1732070831-1270836287.1727946081

また、本法ではリスクベースアプローチが採用されており、AI システムをリスクに応じて「許容できないリスク」「ハイリスク」「特定の透明性が必要なリスク」「最小リスク」の 4 つに分類し、それに応じた規制を適用する。生成 AI を含む汎用目的型 AI モデルも本法内でより包括的な概念として再定義されている。本法には AI の開発・利用を促進するイノベーション支援の観点も含まれている一方、違反した場合にはその違反類型に応じて全世界売上ベースの制裁金が定められており、EU 域外の企業も対象となる。

このように各国でリスクを認識しつつも AI の利活用を前提とした枠組みや法制定がなされる中、2024 年のノーベル物理学賞を受賞したカナダ・トロント大学のジェフリー・ヒントン名誉教授は、急速な AI 進歩に伴い、偽情報や偽画像の拡散等の問題も発生していることに警鐘を鳴らしている。AI を行政サービスに活用するに当たっては、そのリスクを十分に考慮し、各 AI 利用者および AI 提供者にガバナンスを浸透させることが重要である。

第2部 自治体におけるAI関係者の定義

本ガイドラインは自治体サービスの提供や、自治体内部の業務やデータを活用した施策立案（以下、あわせて「自治体業務等」という。）において、AIの提供・利用を担う者を対象としている。ただし、AI事業者ガイドラインと同様に、自治体業務等以外でAIを利用する者またはAIを自治体業務等で利用せずにAIシステム・サービスの便益を享受する、場合によっては損失を被る者（以下、あわせて「業務外利用者（主に住民を想定）」という）については、本ガイドラインの対象に含めない。ただし、自治体業務においてAIの提供・利用を担う者から業務外利用者への必要な対応は記載する。また、AI活用に伴い学習および利用に用いるデータを提供する特定の法人および個人（以下、「データ提供者」という）も同様に本ガイドラインの対象外となる。なお、データ収集は色々な方法が考えられる中で、本ガイドラインではデータの提供を受ける場合や、自治体におけるデータをAIの学習に利用する際など、AI提供・利用を担う組織・自治体職員がデータを取り扱う際の責任を負う形で記載する。以上より、本ガイドラインの対象者は、AIの事業活動を担う主体として、「AI提供者」および「AI利用者」の2つに大別され、それぞれ以下のとおり定義される。なお、AI事業者ガイドラインにおいては、AIの事業活動を担う主体として、上記に「AI開発者（AIシステムを開発する事業者やAIを研究開発する事業者）」を加えた3者が定義されているが、自治体内でAI開発を行うことは想定されないため、本ガイドラインにおいては「AI提供者」および「AI利用者」について述べる。

2-1. AI提供者

AIシステムをアプリケーション、既存のシステム、行政サービスの提供プロセス等に組み込んだサービスとしてAI利用者、場合によっては業務外利用者に提供する者を指す。具体的には、当該サービスの導入を検討する、自治体内の情報システム部局の自治体職員やCDO（最高デジタル責任者）、自治体職員と共同で実際のシステム等を導入する外部ベンダー等がAI提供者として想定される。なお、AIシステム検証、AIシステム他システムとの連携の実装、AIシステム・サービスの提供、正常稼働のためのAIシステムにおけるAI利用者側の運用サポートまたはAIサービスの運用自体を担う。AIサービスの提供に伴い、様々なステークホルダーとのコミュニケーションが求められることもある。

2-2. AI利用者

自治体業務等において、AIシステムまたはAIサービスを利用する者を指す。AI提供者が意図している適正な利用を行い、環境変化等の情報をAI提供者と共有し正常稼働を継続すること、または必要に応じて提供されたAIシステムを運用する役割を担う。具体的には、実際に行政サービスの提供や、施策立案を担う自治体職員がAI利用者として想定される。また、AIの活用において業務外利用者に何らか

の影響が考えられる場合は、当該者に対する AI による意図しない不利益の回避、AI による便益最大化の実現に努める役割を担う。

自らが該当する「AI 提供者」または「AI 利用者」の立場から、「AI に関しどのような取組を行うべきか（倫理原則 = what）」を明らかにすることが重要であり、また、指針を実現するために、「具体的にどのようなアプローチで取り組むか（自治体 AI 関係者の留意点 = how）」を検討・決定し、実践することが AI の安全安心な活用に有用と考えられる。実際の AI システム・サービスは目的・活用技術・データ・利用環境等によって多様なユースケースとなり、技術の発展等、外部環境の変化も踏まえつつ、AI 提供者および AI 利用者が連携して最適なアプローチを検討することが重要である。

第3部 AI ガバナンス

3-1. AI ガバナンスの重要性

AI 関連のサービスや技術は急速に進化しており、活用する際の精度は学習データや環境によって大きく異なることがある。そのため、AI の導入においては、はじめから完璧な結果を求めるのではなく、特定の部署や課題における解決策として、まずは可能な範囲で運用を開始することが重要である。運用を進める中で、試行錯誤を重ねることにより、精度の向上や効率的な運用を実現していくことが期待される。そうした効果を実感した上で、さらなる課題解決や他部署への活用を展開していくことが可能となる。

しかしながら、自治体においては、周期的な予算要求や複数年にわたる継続的な検討が難しい場合がある。このような状況に対処するためには、総合計画やDX推進指針・計画といった長期的な計画の中にAIの利活用を組み込むことが必要である。さらに、これらの指針や計画においては、現場の自治体職員のみならず、首長やCDOといった幹部も含めた検討体制を構築することが重要となる。特に、検討体制においては、行政サービスの向上や住民の生活の質の向上を実現するためにAIを活用する責任と活用に伴うリスクマネジメントの責任の所在がどこにあるのかを明示することが必要である。

AIを活用する責任については、第4部で示すとおり、住民の目線で真の課題の把握やニーズの抽出を行うことにより、住民が抱える課題を解決すること、どのようなデータを使用して、判断や処理を進めているのか明確に伝える透明性の担保、より多くの住民がAIを活用した行政サービスを利用できるようにする多様性と包摂性の確保などが重要である。

活用に伴うリスクマネジメントの責任については、AIが無差別に情報を収集しないためのプライバシー保護、学習に用いたデータを漏洩させないためのセキュリティの確保、差別的な表現・倫理に反する表現が含まれていないかなど公平性へ配慮することなどが重要である。

また、計画段階だけでなく、評価と改善の部分についても、継続的なモニタリングとフィードバックを行い、AI活用の成果を適切に評価し、必要に応じて改善を加えるプロセスを整備することが求められる。

モニタリング状況の住民への公開や関連官庁への報告等を通じて、リスクマネジメント活動へのアカウントビリティを示していくことが求められる。

具体的には当ガイドラインに記載されている項目が適切に実施されているかどうかのセルフアセスメント、および外部のサービスを利用している場合にはその外部ベンダーには当ガイドラインに記載されている項目が適切に実施されているかどうかのセルフアセスメント、および外部のサービスを利用している場合にはその外部ベンダーに対するモニタリングを実施することである。外部ベンダーに対するモニタリングは直接的なアセスメントや外部ベンダーからの客観性のある報告書入手等が考えられる。加えて、IT担当者は自治体

職員からの相談窓口や通報窓口を設置し、AI 活用におけるリスクへ適切な対応ができるようサポートする環境や、当ガイドラインに記載されている項目が適切に実施されていないことを早期発見する環境を整備することが推奨される。

これにより、組織全体で AI 活用に関するガバナンスを強化することができる。そして、各階層の意見を反映しつつ、戦略的に AI を活用することで、自治体全体のサービス向上や業務効率化に寄与する体制を整えることが可能となる。

第4部 自治体 AI 倫理原則

自治体における AI 関係者は、人間中心の考え方のもとに、AI システム・サービスの提供・利用を促進し、人間の尊厳を守りながら、住民への利益を第一とした行政サービスの提供、自治体事務の効率化等、AI の目的を実現していくことが重要である。このため、各主体は、AI 活用に伴う社会的リスクの低減を図るべく、安全性・公平性といった価値を確保することが重要である。また、個人情報等の不適正な利用等の防止を始めとするプライバシー保護、並びに AI システムの脆弱性等による可用性の低下及び外部からの攻撃等のリスクに対応するセキュリティ確保を行うことが重要である。これらを実現するために、各主体は、システムの検証可能性を確保しながらステークホルダーに対する適切な情報を提供することにより透明性を向上させ、アカウントビリティを果たすことが重要となる。

第4部では、行動基準となる基本的価値観として倫理原則を定める。

4-1. 人間中心

① 住民自治

- 自治体事務・サービスは、住民代表による議会議決で制定される例規の範囲内で事務執行するものであることから、AI 活用がこれらを逸脱しないよう活用することとし、本来公務員のみが執行可能である住民・事業者への行政処分など公権力行使や自治体重要施策の決定などに AI を利用しない、あるいは、利用する際には十分に留意する

② 人間主導

- 地方自治体が行う意思決定において、AI を補助的なツールとして位置付けるものとし、AI に過度に依存することがないようにする
- AI の活用においては、自治体職員が関連する周辺情報や法令規則等を踏まえてそのリスクを総合的に判断し、意思決定の最終的な責任を負うものとする

③ 偽情報等への対策

- 偽情報に惑わされないよう、フィルターバブル等への理解とネット情報についての批判的思考能力を育成するとともに、住民や地域社会を守るための情報発信に努める
- 自らの知識・専門性にもとづき、AI で得られた結果を判断することができる分野において、AI を利用する
- AI を通じて得られた結果を利用する場合は、結果の根拠等を確認するほか、最新性、正確性、妥当性、一貫性、説明可能性を確認する。特に、得られた結果は過去の学習データに基づく情報であることを留意する

④ 多様性・包摂性の確保

- 偏見等を生まないように運用するとともに、住民に寄り添いながら、すべての住民が AI の恩恵を享受できるように配慮する
- ⑤ 利用者支援
- 自治体職員に、AI システム・サービスの基礎知識や基本的な使い方（機能の名称や操作方法など）に関する情報を提供し、円滑に AI システム・サービスの利用を開始できるようにする

4-2. 住民第一

- ① 住民第一の AI システム・サービスの導入・利用
- AI システム・サービスの利用にあたっては、自治体事務の効率化などを通じて、その効果を住民および地域社会に還元する
 - AI システム・サービスの導入・利用にあたっては、常に住民の利益を第一に考慮し、公正な競争環境のもとで評価を行う

4-3. 安全・安心

- ① 生命・身体・財産、精神および環境への配慮
- AI の活用において、住民（第三者を含む）の権利を侵害していないこと、住民（第三者を含む）の生命・身体・財産、精神および環境に危害を及ぼすことがないようにするとともに、予測しない事態が生じた場合の対処方法を検討し、当該事態が生じた場合に速やかに実施できるようにする
 - AI を通じて得られた結果については、必要に応じてフィルター機能の導入や自治体職員によるチェック、加筆・修正等を行った上で利用し、知的財産権の権利保全や住民に不適切な情報を提供しないように留意する

4-4. 公平性

- ① 公平性への配慮
- 年齢や性別、人種、宗教、思想、経済的事実等によらず、全ての人々が公平に扱われる社会を実現するために AI を活用する
 - 事務処理や情報発信に AI の出力結果を活用する際は、差別的な表現・倫理に反する表現が含まれていないか、必要に応じて人間の判断を介在させるように留意する
 - 差別的な表現・倫理に反する表現を見逃さないために、自治体職員自身が AI の生成物を公正かつ客観的に評価する価値観を持つ

4-5. プライバシー保護

① プライバシーの保護

- 個人情報の利用目的を明確に定め、AIが無差別に情報を収集しないよう、AIシステム・サービスの導入の計画段階から情報の取得、運用、管理のあり方を検討する。また、プライバシーに配慮し、業務における適切な範囲内でのみ個人情報を利用する
- 個人情報保護法等の関連法令や各地方自治体のプライバシーポリシーの遵守等により、個人情報保護を含め住民の人権が尊重され、保障されるよう、その重要性に応じた対応を取る

4-6. セキュリティ確保

① セキュリティ対策

- AIを利用する際は、業務や利用環境に応じて、法令・契約・規則等に基づき、個人情報、機密情報、法令や契約等により非公開とされている情報、職務上知り得た秘密をはじめとする情報の取扱方針を定める
- 入力データを学習に利用するサービスなどは、入力した内容が蓄積され、外部からの攻撃や内部からの不正に起因して流出するおそれがあるため、機密性・完全性・可用性の観点で対策を講じること

② 最新動向への留意

- AIの技術進歩、国や開発事業者の動向などAIを巡る状況の変化、それに伴うAIへの新たな攻撃手法など、業務利用を進める中で生じた課題や自治体職員の利用実態に照らして生じ得ると想定される課題に適時対応するため、ルールや対策等の見直しを行う

4-7. 透明性

① 検証可能性の確保

- AIの判断にかかわる検証可能性を確保するため、AIシステム・サービスの出入力情報等を記録・保存するとともに、AIシステム・サービスの動作に問題がないか、出力傾向に偏りがないか等を把握できるようにする

② 住民への情報提供

- 住民がAIの活用により納得感および安心感を得られるよう、その活用状況を明確に示し、透明性

を確保する。特に、住民に対する情報提供にAIが関わっている場合は、この旨を明示するなど説明責任を果たす

4-8. アカウンタビリティ

① 責任者の明確化

- 住民に対し、AIの活用責任とその説明責任を担う責任者を設定する

② リスク管理措置の実施と対応状況確認

- 本ガイドラインを活用し、AIシステム・サービスの利用に伴うAIガバナンスに関する方針の策定や安全確保の対策を実施し、必要に応じて見直しを行うこと
- ガイドラインに基づいたリスク管理活動の取組み状況を定期的にモニタリングするとともに、ステークホルダーの求めに対して報告できる状態を維持する

4-9. 教育・リテラシー

① リテラシー・スキルの確保

- AIを活用する自治体のすべての関係者が、AIを効果的かつ適切に活用するためのリテラシーとスキルを確保できるように必要な措置を講じる

② フォローアップの実施

- AIを活用する自治体のすべての関係者は、AIの特性やリスク等を十分に理解し、住民に対してAI活用に係るメリットとそれに伴うリスクに対する理解を促すとともに、必要な措置を講じる

4-10. イノベーション

① 住民ニーズの把握・課題解決

- AIの導入や活用を目的化せず、住民の目線で真の課題の把握やニーズの抽出を行うことにより、住民が抱える課題を解決し、ウェルビーイングに繋がる有効な方策を総合的にデザインする

② ステークホルダーとの協働・連携

- AIの活用にあたっては、多様なステークホルダーと緊密に協働・連携し、互いに補完し合い高め合うことにより地域社会の効用最大化や自治体業務の効率化を図る

第5部 自治体 AI 関係者の留意点

第5部では、AI 提供者および AI 利用者が、AI 倫理原則を踏まえて留意すべきリスクとその対策例を記載する。なお、この章で記載された留意点と対策は、本ガイドラインにおいて提示されている内容を達成する手段の一例であり、実施事項の全てを網羅したものではない。また、自治体における AI 活用形態も様々であることが想定されることから、本章に記載されている内容の全てを記載通りに実施することが求められているものでもない。また、各リスクへの対策実施にあたっては、AI 提供者および AI 利用者単体で完結するものではなく、各主体が適時連携し、協力しながら対応することが重要である。

5-1. AI 提供者における留意点

AI 提供者は、AI システムに付加価値を加えて AI システム・サービスを AI 利用者に提供する役割を担う。AI を社会に普及・発展させるとともに、社会経済の成長にも大きく寄与する一方で、社会に与える影響の大きさゆえに、AI 提供者は、AI の適正な利用を前提とした AI システム・サービスの提供を実現することが重要となる。特に、地方自治体において AI を利活用する場合には、地域特有の課題に対応する柔軟性と、住民の安全や福祉を優先した AI システムの設計・運用が求められる。そのため、AI システム・サービスに組み込む AI が当該システム・サービスに相応しいものが留意することに加え、地域行政の課題や社会環境の変化、住民ニーズに応じて AI に対する期待値が変わることも考慮して、地方自治体が持続可能かつ効果的に AI を活用できる環境を整備することが重要である。

さらに、AI 利用者に対しては、AI システムの提供および運用のサポートまたは AI システムの運用をしつつ AI サービスを提供することが重要である。提供に際し、住民の権利を侵害せず、かつ社会に不利益等を生じさせることがないように留意し、合理的な範囲でインシデント事例等を含む関連情報の共有を行い、地方自治体において、より安全安心で信頼できる AI システム・サービスを提供することが期待される。

なお、AI 提供者に求められる対策は、自治体における IT 担当者が担うことが想定されるため、以降、IT 担当者として表記する。

① 人間中心

①住民自治	
内容	・ 自治体事務・サービスは、住民代表による議会議決で制定される例規の範囲内で事務執行するものであることから、AI 活用がこれらを逸脱しないよう活用することとし、本来公務員のみが執行可能である住民・事業者への行政処分など公権力行使や自治体重要施策の決定などに AI を利用しない、あるいは、利用する際には十分に留意する
リスク	・ AI システム・サービスを使って行われる政策決定が少数の技術者やデータサイエンティストによって制御され、住民の参加や意思表示が形骸化し住民自治の理念やプロセスが

	損なわれるリスクがある
対策	<ul style="list-style-type: none"> ・ IT 担当者は AI の出力内容が住民の意思が反映された内容になっているか確認するプロセスや参照元を表示できる仕組みを整備すること <ul style="list-style-type: none"> ➢ AI の学習内容が地域の実態に合うかを確認し、AI の出力内容が不適切な場合は再学習などのプロセスを整備すること
②人間主導	
内容	<ul style="list-style-type: none"> ・ 地方自治体が行う意思決定において、AI を補助的なツールとして位置付けるものとし、AI に過度に依存することがないようにする ・ AI の活用においては、自治体職員が関連する周辺情報や法令規則等を踏まえてそのリスクを総合的に判断し、意思決定の最終的な責任を負うものとする
リスク	<ul style="list-style-type: none"> ・ 人間の意思決定、認知等、感情を不当に操作することを目的とした、または意識的に知覚できないレベルでの操作が行われるリスクがある
対策	<ul style="list-style-type: none"> ・ IT 担当者は自治体職員への教育や注意喚起を行い、AI への依存、意思決定、感情の操作等のリスクの存在の共有および認識の促進を行うこと ・ IT 担当者は AI の出力情報を活用する際には積極的に確認プロセスを設けるように促すこと
③偽情報等への対策	
内容	<ul style="list-style-type: none"> ・ 偽情報に惑わされないよう、フィルターバブル（AI がフィルターした情報の中でしか情報を得られない状況）等への理解とネット情報についての批判的思考能力を育成するとともに、住民や地域社会を守るための情報発信に努める ・ 自らの知識・専門性にもとづき、AI で得られた結果を判断することができる分野において、AI を利用する ・ AI を通じて得られた結果を利用する場合は、結果の根拠等を確認するほか、最新性、正確性、妥当性、一貫性、説明可能性を確認する。特に、得られた結果は過去の学習データに基づく情報であることに留意する
リスク	<ul style="list-style-type: none"> ・ AI システム・サービスが生み出す偽情報をモニタリング・検知できず、自治体における偽情報の拡散、さらには情報操作や世論工作につながるリスクがある
対策	<ul style="list-style-type: none"> ・ IT 担当者は下記注意点を理解し、AI に関するルールを順守しながら提供すること <ul style="list-style-type: none"> ➢ 正確性や最新性（データが適切であること）等が担保されたデータの学習を行う ➢ AI モデルの入出力および判断根拠を定期的に評価し、事実と反するハルシネーションの発生をモニタリングすること ➢ IT 担当者は自治体職員に対して、ハルシネーション（AI が誤認や論理の矛盾を含む事象や事実とは異なる情報を作り出してしまう現象）の理解を向上させるための取り組みを行うこと ➢ 出力結果に応じて、AI 開発者や AI 提供者に対して AI モデル改善の判断を促すこと

	<ul style="list-style-type: none"> ▶ 住民に危害が及ぼされるような事態が生じた場合の対処方法を検討し、速やかに実施できるようにすること
④多様性・包摂性の確保	
内容	<ul style="list-style-type: none"> ・ 偏見等を生まないように運用するとともに、住民に寄り添いながら、すべての住民がAIの恩恵を享受できるように配慮する
リスク	<ul style="list-style-type: none"> ・ 自治体が住民の多数派を対象としたAI活用を行った結果、少数派（社会的弱者やマイノリティ）がAIの恩恵を受けられなくなるリスクがある。また、技術的弱者が自分の意向を発信できなくなるリスクがある
対策	<ul style="list-style-type: none"> ・ IT担当者はAIの出力内容が当該自治体の少数派（社会的弱者やマイノリティ）の意向にも配慮しているかどうか確認するプロセスを整備すること ・ IT担当者は「情報弱者」および「技術的弱者」を生じさせず、AIの恩恵を享受できるよう少数派（社会的弱者やマイノリティ）によるAIの活用を容易にするよう自治体職員に注意喚起すること <p>【具体的な手法（例）】</p> <ul style="list-style-type: none"> ・ UI（ユーザーインターフェース）、UX（ユーザーエクスペリエンス）の改善
⑤利用者支援	
内容	<ul style="list-style-type: none"> ・ 自治体職員に、AIシステム・サービスの基礎知識や基本的な使い方（機能の名称や操作方法など）に関する情報を提供し、円滑にAIシステム・サービスの利用を開始できるようにする
リスク	<ul style="list-style-type: none"> ・ 情報格差や技術格差等により、AIシステム・サービスを適切且つ効率的に活用できる自治体職員とそうでない自治体職員の格差が生まれるリスクがある
対策	<ul style="list-style-type: none"> ・ IT担当者は多くの自治体職員がAIシステム・サービスを活用できるようにUI、UXに配慮した環境を整備すること ・ IT担当者はAIシステム・サービスに関する説明会、ガイドライン、研修・トレーニングに関する情報を積極的に展開すること ・ IT担当者はAIシステム・サービスに関する自治体職員からの問い合わせ等に関して対応し、ナレッジの展開につとめること

② 住民第一

①住民第一のAIシステム・サービスの導入・利用	
内容	<ul style="list-style-type: none"> ・ AIシステム・サービスの利用にあたっては、自治体事務の効率化などを通じて、その効果を住民および地域社会に還元する ・ AIシステム・サービスの導入・利用にあたっては、常に住民の利益を第一に考慮し、公正な競争環境のもとで評価を行う
リスク	<ul style="list-style-type: none"> ・ 自治体職員の業務負荷軽減や効率化が優先されるあまり、そもそも人間（自治体職

	<p>員)が行うべき仕事でさえもAIによる代替が行われてしまい、業務を通して住民が享受できる利益や価値が低減するリスクがある</p> <ul style="list-style-type: none"> 特定の事業者に見聞や技術が集中することで、公正な競争環境が維持されず AI システム・サービスの選択肢が限定され住民にとって最善のサービスが提供されず、不利益を被るリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は業務効率化に関する AI 活用にあたっては、それがもたらす住民・地域社会への効果を検討すること、また、AI 利用者からの要望・要求に対し、それが住民・地域社会に利益をもたらすことを確認すること IT 担当者は AI システム・サービスの調達にあたっては、公正性、競争性、透明性を担保しつつ効率性や経済性も考慮し選定する。また、その際に自治体と住民にとって利益となるサービスとなることを重視すること

③ 安全・安心

①生命・身体・財産、精神および環境への配慮	
内容	<ul style="list-style-type: none"> AI の活用において、住民（第三者を含む）の権利を侵害していないこと、住民（第三者を含む）の生命・身体・財産、精神および環境に危害を及ぼさないようにするとともに、予測しない事態が生じた場合の対処方法を検討し、当該事態が生じた場合に速やかに実施できるようにする AI を通じて得られた結果については、必要に応じてフィルター機能の導入や自治体職員によるチェック、加筆・修正等を行った上で利用し、知的財産権の権利保全や住民に不適切な情報を提供しないように留意する
リスク	<ul style="list-style-type: none"> AI システム・サービスを住民の生命・身体・財産に関する業務等に使用する場合、その出力内容および回答内容によって社会保険や社会福祉などの社会保障上の問題が生じ、住民の生命・身体・財産に危害が及ぼされるリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は下記注意点を理解し、AI に関するルールを順守しながら提供すること <ul style="list-style-type: none"> ➤ 正確性や最新性（データが適切であること）等が担保されたデータの学習を行う ➤ AI モデルの入出力および判断根拠を定期的に評価し、不正確な出力結果の発生をモニタリングすること ➤ IT 担当者は自治体職員に対して、データドリフト（予期せぬ環境変化によって、AI モデルの予測性能が時間経過とともに劣化していく現象）の理解を向上させるための取り組みを行うこと ➤ 出力結果に応じて、AI 開発者や AI 提供者に対して AI モデル改善の判断を促すこと ➤ 住民に危害が及ぼされるような事態が生じた場合の対処方法を検討し、速やかに実施できるようにすること

④ 公平性

① 公平性への配慮	
内容	<ul style="list-style-type: none"> 年齢や性別、人種、宗教、思想、経済的事実等によらず、全ての人々が公平に扱われる社会を実現するために AI を活用する 事務処理や情報発信に AI の出力結果を活用する際は、差別的な表現・倫理に反する表現が含まれていないか、必要に応じて人間の判断を介在させるように留意する 差別的な表現・倫理に反する表現を見逃さないために、自治体職員自身が AI の生成物を公正かつ客観的に評価する価値観を持つ
リスク	<ul style="list-style-type: none"> AI システム・サービスで住民の年齢・性別等を扱う業務を行う場合、その出力および回答内容によって社会保障上の公平性が損なわれる、もしくは差別的な表現・倫理に反する表現の問題が生じるリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は下記注意点を理解し、AI に関するルールを順守しながら提供すること <ul style="list-style-type: none"> ➤ 正確性や最新性（データが適切であること）が担保されたデータの学習を行う ➤ 不当に偏りがあるデータの学習は行わない ➤ AI モデルの入出力および判断根拠を定期的に評価し、不公平な出力結果の発生をモニタリングすること ➤ IT 担当者は自治体職員に対して、データドリフト（予期せぬ環境変化によって、AI モデルの予測性能が時間経過とともに劣化していく現象）の理解を向上させるための取り組みを行うこと ➤ 出力結果に応じて、AI 開発者や AI 提供者に対して AI モデル改善の判断を促すこと ➤ 住民が不当に扱われる事態が生じた場合の対処方法を検討し、速やかに実施できるようにすること

⑤ プライバシー保護

① プライバシーの保護	
内容	<ul style="list-style-type: none"> 個人情報の利用目的を明確に定め、プライバシーに配慮し、業務における適切な範囲内でのみ個人情報を利用する また、AI が無差別に情報を収集しないよう、AI システム・サービスの導入の計画段階から情報の取得、運用、管理のあり方を検討する 個人情報保護法等の関連法令や各地方自治体のプライバシーポリシーの遵守等により、個人情報保護を含め住民の人権が尊重され、保障されるよう、その重要性に応じた対応を取る
リスク	<ul style="list-style-type: none"> 個人データをプロンプトに入力したり RAG を活用して読み込ませたりして、当該データが学習された場合、プロンプトを介して自分以外の第三者プロンプトを介して自分以外の第三者に個人データが提供され、個人情報保護法に違反するリスクがある

対策	<ul style="list-style-type: none"> IT 担当者は導入する AI サービスには、個人情報や機械学習に利用しない機能があること等を十分に確認すること IT 担当者は導入する AI サービスで、必要に応じて個人データを機械学習に利用しないよう設定すること IT 担当者は AI システム・サービスにおけるプライバシー保護に関して適宜情報収集し、AI を利用する自治体職員や住民に研修等を行い周知すること
----	---

⑥ セキュリティ確保

①セキュリティ対策	
内容	<ul style="list-style-type: none"> AI を利用する際は、業務や利用環境に応じて、法令・契約・規則等に基づき、個人情報、機密情報、法令や契約等により非公開とされている情報、職務上知り得た秘密をはじめとする情報の取扱方針を定める 入力データを学習に利用するサービスなどは、入力した内容が蓄積され、外部からの攻撃や内部からの不正に起因して流出するおそれがあるため、機密性・完全性・可用性の観点で対策を講じること
リスク	<ul style="list-style-type: none"> 性能劣化および誤分類につながるような学習データへの不正データ混入、アプリケーション自体を狙ったサイバー攻撃、AI の推論結果または AI への指示であるプロンプトを通じた攻撃等により情報漏洩や悪用リスクがある
対策	<ul style="list-style-type: none"> IT 担当者は AI システム・サービスの提供の過程を通じて、採用する技術の特性に照らして脅威分析結果に応じたセキュリティ要件を策定し、適切にセキュリティ対策を講ずること IT 担当者は AI システム・サービスにアクセスできるユーザーを制限し、適切な認証（誰がアクセスしているのかを確認する）と認可（アクセス権を持つ者にのみ許可する）を実施すること IT 担当者は AI システム・サービスへのアクセスログをとり、定期的にアクセスログのモニタリング・解析を行うこと IT 担当者は重要な AI システム・サービス調達時や、共通的な基盤となる AI プラットフォームの提供においては、第三者認証等の取得状況を確認すること （例：ISMS（ISO/IEC27001）や ISMAP クラウドサービスリストへの登録、クラウドサービスにおける第三者認証（ISO/IEC27017、ISO/IEC27018 等）） IT 担当者は AI システム・サービスにおけるセキュリティ対策に関して適宜情報収集し、AI を利用する自治体職員や住民に研修等を行い周知すること
②最新動向への留意	
内容	<ul style="list-style-type: none"> AI の技術進歩、国や開発事業者の動向など AI を巡る状況の変化、それに伴う AI への新たな攻撃手法など、業務利用を進める中で生じた課題や自治体職員の利用実態に照らして生じ得ると想定される課題に適時対応するため、ルールや対策等の見直しを行

	う
リスク	<ul style="list-style-type: none"> AI の技術的進歩に伴う新たな脆弱性の発生や、最新の不正アクセス手法・サイバーセキュリティ攻撃等に関する情報を把握できておらず、セキュリティ対策が不十分になるリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は AI システム・サービスに対する外部からの新たな攻撃手法に対する情報を定期的に収集し、これらのリスクに対応するための留意事項を確認すること、また、必要な情報を利用者に提供すること IT 担当者は最新の不正アクセス対策、セキュリティ対策情報を確認し、AI システム・サービスに適した対策を講じることの利用環境を整備すること

⑦ 透明性

① 検証可能性の確保	
内容	<ul style="list-style-type: none"> AI の判断にかかわる検証可能性を確保するため、AI システム・サービスの出入力情報等を記録・保存するとともに、AI システム・サービスの動作に問題がないか、出力傾向に偏りがないか等を把握できるようにする
リスク	<ul style="list-style-type: none"> AI からの予測・認識等に偏りが発生したとしても、出入力情報が適切に記録されておらず、確かさを検証することができないリスクがある その結果、住民から、AI を通じて行われた自治体の意思決定の根拠や理由の説明を求められた際に要求に応じられず、AI システム・サービスの不信に繋がる、または利活用が促進されないリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は AI システム・サービスの出入力情報を記録・保存すること IT 担当者は重要な AI システム・サービスや、共通的な基盤となる AI プラットフォームを調達する際の選定基準として、解釈可能なアルゴリズムを実装する AI システムかどうか確認すること IT 担当者は AI の学習等に利用されたデータが、いつ、どこで、どういう目的で集められたデータなのかを管理すること IT 担当者は AI モデルの入出力傾向の分析すること
② 住民への情報提供	
内容	<ul style="list-style-type: none"> 住民が AI の活用により納得感および安心感を得られるよう、その活用状況を明確に示し、透明性を確保する。特に、住民に対する情報提供に AI が関わっている場合は、この旨を明示するなど説明責任を果たす
リスク	<ul style="list-style-type: none"> 自治体からの情報提供等に AI が関わっていることを住民に説明せず、その活用が発覚した場合、住民からの不安・不満・反感を招くリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は AI システム・サービスを利用していること、およびその AI システム・サービスに係る情報を自治体職員や住民に公開すること IT 担当者は自治体職員および住民のニーズ、意見等も踏まえた説明不足部分を明確

	化し、AI 開発事業者と連携し説明内容を検討すること ・ IT 担当者は住民が問い合わせ・苦情を提起できるような Web フォーム等を開設し、住民に公開すること
--	---

⑧ アカウンタビリティ

①責任者の明確化	
内容	<ul style="list-style-type: none"> 住民に対し、AI の活用責任とその説明責任を担う責任者を設定する
リスク	<ul style="list-style-type: none"> AI の活用責任とその説明責任を担う責任者があいまいなため、問題が起きた際に適切に住民や地域社会への責務を果たすことができないリスクがある
対策	<ul style="list-style-type: none"> IT 担当部署の中で AI サービスの責任者を任命し、その責任範囲を明確に定める 責任者がその役割を遂行するために必要な権限とリソースを明確にし、組織は適切にその要求に応じる 問題が起きた際は、自らが指揮を執り必要な対策を検討し、講じること
②リスク管理措置の実施と対応状況確認	
内容	<ul style="list-style-type: none"> 本ガイドラインを活用し、AI システム・サービスの利用に伴う AI ガバナンスに関する方針の策定や安全確保の対策を実施し、必要に応じて見直しを行うこと ガイドラインに基づいたリスク管理活動の取組み状況を定期的にモニタリングするとともに、ステークホルダーの求めに対して報告出来る状態を維持する
リスク	<ul style="list-style-type: none"> 自治体において、指針や規程、ガイドラインが無く自治体内での統一的な対策が実施できず、AI の問題が顕在化するリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は本ガイドラインを活用し、AI システム・サービスの利用に伴う AI ガバナンスに関するポリシー、プライバシーポリシー等の方針を策定すること IT 担当者は当該ポリシー、プライバシーポリシーで定められたリスク管理や安全性確保の対策を実施すること IT 担当者は上記状況を定期的に確認し、必要に応じて見直し・改定を行うこと

⑨ 教育・リテラシー

①リテラシー・スキルの確保	
内容	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者が、AI を効果的かつ適切に活用するためのリテラシーとスキルを確保できるように必要な措置を講じる
リスク	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者の AI に関するリテラシーやスキルが不十分なために、AI を利用した示唆の精度が低い、もしくは誤った文書を作成してしまうことで自治体業務に混乱を招いてしまうリスクや、住民に不利益を生じさせるリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は AI に関する学習を通してその特性やリスク等に関して十分な理解に努める IT 担当者は自治体職員に対して AI 活用に係るメリットとそれに伴うリスクを説明とともに、必要な教育環境を提供し、自治体職員のリテラシー向上を行う

② フォローアップの実施	
内容	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者は、AI の特性やリスク等を十分に理解し、住民に対して AI 活用に係るメリットとそれに伴うリスクに対する理解を促すとともに、必要な措置を講じる
リスク	<ul style="list-style-type: none"> 住民が AI を利用する際、問題や疑問を抱えた場合の自治体によるサポートがなく、AI システム・サービスや自治体への不信に繋がるリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は住民が問い合わせ・苦情を提起できるような仕組み等を開設し、住民に公開すること

⑩ イノベーション

① 住民ニーズの把握・課題解決	
内容	<ul style="list-style-type: none"> AI の導入や活用を目的化せず、住民の目線で真の課題の把握やニーズの抽出を行うことにより、住民が抱える課題を解決し、ウェルビーイングに繋がる有効な方策を総合的にデザインする
リスク	<ul style="list-style-type: none"> AI の導入や活用自体が目的となり、地域や住民にとっての課題解決につながらないリスクがある
対策	<ul style="list-style-type: none"> IT 担当者は住民の目線で真の課題の把握やニーズの抽出を行う その上で、AI の導入を含めた、住民が抱える課題を解決しウェルビーイングに繋がる有効な方策を検討する IT 担当者は、AI システムの定常的な改良・向上を図る運用のあり方を予め定めておく IT 担当者は時間経過とともに住民のニーズに対する AI システム・サービスの有効性が劣化することのないよう定期的に住民アンケート等の効果測定を行い、結果に応じて、AI システム・サービスの改善を行うこと
② ステークホルダーとの協働・連携	
内容	<ul style="list-style-type: none"> AI の活用にあたっては、多様なステークホルダーと緊密に協働・連携し、互いに補完し合い高め合うことにより地域社会の効用最大化や自治体業務の効率化を図る
リスク	<ul style="list-style-type: none"> AI の利活用が自治体職員のみならず、その他のステークホルダーにとって最適なものでなくなってしまうリスクがある
対策	<ul style="list-style-type: none"> 単独の自治体とその自治体職員だけで AI の利活用を検討するのではなく、複数の地域・自治体や民間企業、アカデミアなど多様なステークホルダーが協力して地域社会の効用最大化や自治体業務効率化を目指す

5-2. AI 利用者における留意点

AI 利用者は、AI 提供者から安全安心で信頼できる AI システム・サービスの提供を受け、AI 提供者が

意図した範囲内で継続的に適正利用および必要に応じて AI システムの運用を行うことが重要である。特に、地方自治体においては、住民への公共サービスの質を向上させ、地域課題の解決に資するために AI を適切に活用することが期待される。また、地方自治体が AI を活用する際には、住民との信頼関係を保ちながら人間の判断を介在させることで、人間の尊厳および自律を守りながら予期せぬ事故を防ぐことも可能となる。同時に、AI の導入が行政サービスに与える影響が大きいことを踏まえ、住民に対して透明性を確保し、利用目的やシステムの限界を明確にすることが求められる。

AI 利用者は、住民から AI の能力または出力結果に関して説明を求められた場合、AI 提供者等のサポートを得て、その要望に応え理解を得ることが期待されると同時に、より効果的な AI 利用のために必要な知見習得に努めることが求められる。また、地域特性や住民ニーズを反映させながら、持続可能で信頼できる運用を実現し、AI の恩恵を地域全体で共有することが期待される。

なお、AI 利用者に求められる対策は、自治体における職員が担うことが想定されるため、以降、自治体職員と表記する。

① 人間中心

①住民自治	
内容	<ul style="list-style-type: none"> 自治体事務・サービスは、住民代表による議会議決で制定される例規の範囲内で事務執行するものであることから、AI 活用がこれらを逸脱しないよう活用することとし、本来公務員のみが執行可能である住民・事業者への行政処分など公権力行使や自治体重要施策の決定などに AI を利用しない、あるいは、利用する際には十分に留意する
リスク	<ul style="list-style-type: none"> AI システム・サービスを使って行われる政策決定が少数の技術者やデータサイエンティストによって制御され、住民の参加や意思表示が形骸化し住民自治の理念やプロセスが損なわれるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI の出力内容が住民の意思が反映された内容になっているか、適時確認すること 自治体職員は下記注意点を理解し、AI に関するルールを順守しながら利用すること <ul style="list-style-type: none"> AI の出力内容が地域の実態に合うかを確認し、AI の出力内容を活用する際は一定の人々の意見に依存していないことを確認すること
②人間主導	
内容	<ul style="list-style-type: none"> 地方自治体が行う意思決定において、AI を補助的なツールとして位置付けるものとし、AI に過度に依存することがないようにする AI の活用においては、自治体職員が関連する周辺情報や法令規則等を踏まえてそのリスクを総合的に判断し、意思決定の最終的な責任を負うものとする
リスク	<ul style="list-style-type: none"> 人間の意思決定、認知等、感情を不当に操作することを目的とした、または意識的に知覚できないレベルでの操作が行われるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI への依存、意思決定、感情の操作等のリスクの存在を共有および認識すること

	<ul style="list-style-type: none"> 自治体職員は AI の出力情報を活用する際には積極的に確認プロセスを設け、AI の判断ではなく、人間の判断として意思決定すること
③ 偽情報等への対策	
内容	<ul style="list-style-type: none"> 偽情報に惑わされないよう、フィルターバブル（AI がフィルターした情報の中でしか情報を得られない状況）等への理解とネット情報についての批判的思考能力を育成するとともに、住民や地域社会を守るための情報発信に努める 自らの知識・専門性にもとづき、AI で得られた結果を判断することができる分野において、AI を利用する AI を通じて得られた結果を利用する場合は、結果の根拠等を確認するほか、最新性、正確性、妥当性、一貫性、説明可能性を確認する 特に、得られた結果は過去の学習データに基づく情報であることに留意する
リスク	<ul style="list-style-type: none"> AI システム・サービスが生み出す偽情報をモニタリング・検知できず、自治体における偽情報の拡散、さらには情報操作や世論工作につながるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は下記注意点を理解し、AI に関するルールを順守しながら利用すること <ul style="list-style-type: none"> ▶ プロンプトを通してデータの入力を行う際は、正確性や最新性（データが適切であること）等が担保されたデータの入力を行う ▶ 自治体職員はハルシネーション（AI（人工知能）が誤認や論理の矛盾を含む事象や事実とは異なる情報を作り出してしまふ現象）を理解し、自身が真偽や採択を判断した上で利用すること ▶ ハルシネーション理解に関する取り組み・トレーニングを受講する ▶ 必要に応じて、AI 提供者に出力情報などにハルシネーションの疑いがあることを報告すること 自治体職員は情報を外部に公開する際には上長に相談するなど、自らの判断の確からしさを確認すること
④ 多様性・包摂性の確保	
内容	<ul style="list-style-type: none"> 偏見等を生まないように運用するとともに、住民に寄り添いながら、すべての住民が AI の恩恵を享受できるように配慮する
リスク	<ul style="list-style-type: none"> 自治体が住民の多数派を対象とした AI 活用を行った結果、少数派（社会的弱者やマイノリティ）が AI の恩恵を受けられなくなるリスクがある。また、技術的弱者が自分の意向を発信できなくなるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI の出力内容が当該自治体の多数派の意向のみを代表するリスクがあることを念頭に置き、出力結果を確認すること 自治体職員は「情報弱者」および「技術的弱者」を生じさせず、AI の恩恵を享受できるように少数派（社会的弱者やマイノリティ）による AI の活用を支援すること <p>【具体的な手法（例）】</p> <ul style="list-style-type: none"> 問い合わせフォームや目安箱など安全安心な利用環境の整備

	<ul style="list-style-type: none"> 住民向け勉強会の開催
⑤利用者支援	
内容	<ul style="list-style-type: none"> 自治体職員に、AI システム・サービスの基礎知識や基本的な使い方（機能の名称や操作方法など）に関する情報を提供し、円滑にAI システム・サービスの利用を開始できるようにする
リスク	<ul style="list-style-type: none"> 情報格差や技術格差等により、AI システム・サービスを適切且つ効率的に活用できる自治体職員とそうでない自治体職員の格差が生まれるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員はAI システム・サービスに関する説明会、ガイドライン、研修・トレーニングに関する情報を積極的かつ継続的に活用すること

② 住民第一

①住民第一のAI システム・サービスの導入・利用	
内容	<ul style="list-style-type: none"> AI システム・サービスの利用にあたっては、自治体事務の効率化などを通じて、その効果を住民および地域社会に還元する AI システム・サービスの導入・利用にあたっては、常に住民の利益を第一に考慮し、公正な競争環境のもとで評価を行う
リスク	<ul style="list-style-type: none"> 自治体職員の業務負荷軽減や効率化が優先されるあまり、そもそも人間（自治体職員）が行うべき仕事でさえもAI による代替が行われてしまい、業務を通して住民が享受できる利益や価値が低減するリスクがある 特定の事業者に知見や技術が集中することで、公正な競争環境が維持されずAI システム・サービスの選択肢が限定され住民にとって最善のサービスが提供されず、不利益を被るリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は必要に応じてIT 担当者に対し、サービス選定において住民や地域社会の意向や期待を伝えること 自治体職員は業務効率化に関するAI 活用にあたっては、それがもたらす住民・地域社会への効果を検討すること <p>【具体的な手法（例）】</p> <ul style="list-style-type: none"> AI サービスの導入にあたり、住民による意見を収集できる仕組み（パブコメ、目安箱など）を整備する

③ 安全・安心

①生命・身体・財産、精神および環境への配慮	
内容	<ul style="list-style-type: none"> AI の活用において、住民（第三者を含む）の権利を侵害していないこと、住民（第三者を含む）の生命・身体・財産、精神および環境に危害を及ぼすことがないようにするとともに、予測しない事態が生じた場合の対処方法を検討し、当該事態が生じた場合に速やかに実施できるようにする

	<ul style="list-style-type: none"> AI を通じて得られた結果については、必要に応じてフィルター機能の導入や自治体職員によるチェック、加筆・修正等を行った上で利用し、知的財産権の権利保全や住民に不適切な情報を提供しないように留意する。
リスク	<ul style="list-style-type: none"> AI システム・サービスを住民の生命・身体・財産に関する業務等に使用する場合、その出力内容および回答内容によって社会保険や社会福祉などの社会保障上の問題が生じ、住民の生命・身体・財産に危害が及ぼされるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は下記注意点を理解し、AI に関するルールを順守しながら利用すること <ul style="list-style-type: none"> ➤ 正確性や最新性（データが適切であること）等が担保されたデータの入力を行う ➤ 自治体職員はデータドリフト（予期せぬ環境変化によって、AI モデルの予測性能が時間経過とともに劣化していく現象）を理解しながら利用すること ➤ データドリフトの理解に関する取り組み・トレーニングを受講する ➤ 必要に応じて、AI 提供者に出力情報などにデータドリフトの疑いがあることを報告すること 自治体職員は情報を外部に公開する際には上長に相談するなど、自らの判断の確からしさを確認すること <ul style="list-style-type: none"> ➤ 生成 AI 回答の情報ソースを得られる場合は参照すること ➤ 利用者からの即時フィードバックが得られる仕組みを開設すること

④ 公平性

①公平性への配慮	
内容	<ul style="list-style-type: none"> 年齢や性別、人種、宗教、思想、経済的事実等によらず、全ての人が公平に扱われる社会を実現するために AI を活用する 事務処理や情報発信に AI の出力結果を活用する際は、差別的な表現・倫理に反する表現が含まれていないか、必要に応じて人間の判断を介在させるように留意する 差別的な表現・倫理に反する表現を見逃さないために、自治体職員自身が AI の生成物を公正かつ客観的に評価する価値観を持つ
リスク	<ul style="list-style-type: none"> AI システム・サービスで住民の年齢・性別等を扱う業務を行う場合、その出力および回答内容によって社会保障上の公平性が損なわれる、もしくは差別的な表現・倫理に反する表現の問題が生じるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は下記注意点を理解し、AI に関するルールを順守しながら利用すること <ul style="list-style-type: none"> ➤ 正確性や最新性（データが適切であること）等が担保されたデータの入力を行う ➤ 不当に偏りがあるデータの入力は行わない ➤ 自治体職員はデータドリフト（予期せぬ環境変化によって、AI モデルの予測性能が時間経過とともに劣化していく現象）を理解しながら利用すること ➤ データドリフトの理解に関する取り組み・トレーニングを受講する ➤ 必要に応じて、AI 提供者に出力情報などにデータドリフトの疑いがあることを報告

	<p>すること</p> <ul style="list-style-type: none"> 自治体職員は情報を外部に公開する際には上長に相談するなど、自らの判断の確からしさを確認すること <ul style="list-style-type: none"> 利用者からの即時フィードバックが得られる仕組みを開設すること
--	---

⑤ プライバシー保護

① プライバシーの保護	
内容	<ul style="list-style-type: none"> 個人情報の利用目的を明確に定め、プライバシーに配慮し、業務における適切な範囲内でのみ個人情報を利用する また、AIが無差別に情報を収集しないよう、AIシステム・サービスの導入の計画段階から情報の取得、運用、管理のあり方を検討する 個人情報保護法等の関連法令や各地方自治体のプライバシーポリシーの遵守等により、個人情報保護を含め住民の人権が尊重され、保障されるよう、その重要性に応じた対応を取る
リスク	<ul style="list-style-type: none"> 個人データをプロンプトに入力したり RAG を活用して読み込ませたりして、当該データが学習された場合、プロンプトを介して自分以外の第三者に個人データが提供され、個人情報保護法に違反するリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は利用目的を明確に定め、それを実現するための必要最小限の個人情報を取得すること 自治体職員は AI システム・サービス上で、同意の範囲を超えて個人情報を取り扱うことのないよう注意を払うこと 自治体職員は個人データの利用が禁止されている AI システム・サービスがあることを理解すること 自治体職員は AI システム・サービスにおけるプライバシー保護に関して適宜情報収集し、必要な研修等を受講すること

⑥ セキュリティ確保

①セキュリティ対策	
内容	<ul style="list-style-type: none"> AI を利用する際は、業務や利用環境に応じて、法令・契約・規則等に基づき、個人情報、機密情報、法令や契約等により非公開とされている情報、職務上知り得た秘密をはじめとする情報の取扱方針を定める 入力データを学習に利用するサービスなどは、入力した内容が蓄積され、外部からの攻撃や内部からの不正に起因して流出するおそれがあるため、機密性・完全性・可用性の観点で対策を講じること
リスク	<ul style="list-style-type: none"> 性能劣化および誤分類につながるような学習データへの不正データ混入、アプリケーション自体を狙ったサイバー攻撃、AI の推論結果または AI への指示であるプロンプトを通し

	た攻撃等により情報漏洩や悪用リスクがある
対策	<ul style="list-style-type: none"> 自治体職員は定められたセキュリティ要件に準拠し、AI システム・サービスを利用すること 自治体職員は AI システム・サービスにアクセスするための情報の管理を行い、許可された用途や環境以外での利用を行わないこと AI 提供者が行っている AI システム・サービスのアクセスログのモニタリング・解析結果を定期的に確認すること 自治体職員は AI システム・サービスにおけるセキュリティ対策に関して適宜情報収集し、必要な研修等を受講すること
②最新動向への留意	
内容	<ul style="list-style-type: none"> AI の技術進歩、国や開発事業者の動向など AI を巡る状況の変化、それに伴う AI への新たな攻撃手法など、業務利用を進める中で生じた課題や自治体職員の利用実態に照らして生じ得ると想定される課題に適時対応するため、ルールや対策等の見直しを行う
リスク	<ul style="list-style-type: none"> AI の技術的進歩に伴う新たな脆弱性の発生や、最新の不正アクセス手法・サイバーセキュリティ攻撃等に関する情報を把握できておらず、セキュリティ対策が不十分になるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は IT 担当者から提供される、AI システム・サービスに対する外部からの新たな攻撃手法に対する留意事項を確認すること 自治体職員は IT 担当者が行っている最新の不正アクセス対策、セキュリティ対策の内容を定期的に確認すること

⑦ 透明性

①検証可能性の確保	
内容	<ul style="list-style-type: none"> AI の判断にかかわる検証可能性を確保するため、AI システム・サービスの出入力情報等を記録・保存するとともに、AI システム・サービスの動作に問題がないか、出力傾向に偏りがないか等を把握できるようにする
リスク	<ul style="list-style-type: none"> AI からの予測・認識等に偏りが発生したとしても、出入力情報が適切に記録されておらず、確からしさの根拠を検証することができないリスクがある その結果、住民から、AI を通じて行われた自治体の意思決定の根拠や理由の説明を求められた際に要求に応じられず、AI システム・サービスの不信に繋がる、または利活用が促進されないリスクがある
対策	<ul style="list-style-type: none"> 自治体職員はプロンプトを通してデータの入力を行う際は、そのデータが、いつ、どこで、どういう目的で集められたデータなのかを認識したうえで利用すること
②住民への情報提供	
内容	<ul style="list-style-type: none"> 住民が AI の活用により納得感および安心感を得られるよう、その活用状況を明確に示し、透明性を確保する。特に、住民に対する情報提供に AI が関わっている場合は、この

	旨を明示するなど説明責任を果たす
リスク	<ul style="list-style-type: none"> 自治体からの情報提供等に AI が関わっていることを住民に説明せず、その活用が発覚した場合、住民からの不安・不満・反感を招くリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は住民に関する自治体業務において AI を活用する場合は、その事実と取り組み内容や伴うリスクに関して、理解できる範囲の情報を事前に住民へ情報提供し、納得感および安心感を醸成すること 自治体職員は住民が問い合わせ・苦情を提起できるような仕組みを開設し、住民に公開すること

⑧ アカウンタビリティ

①責任者の明確化	
内容	<ul style="list-style-type: none"> 住民に対し、AI の活用責任とその説明責任を担う責任者を設定する
リスク	<ul style="list-style-type: none"> AI の活用責任とその説明責任を担う責任者があいまいなため、問題が起きた際に適切に住民や地域社会への責務を果たすことができないリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI サービスの責任者の指示に従い、本ガイドラインに基づいた対応を行い、問題が起きた際は、速やかに AI サービス責任者へ報告をし、住民に対し必要な対策を講じること
②リスク管理措置の実施と対応状況確認	
内容	<ul style="list-style-type: none"> 本ガイドラインを活用し、AI システム・サービスの利用に伴う AI ガバナンスに関する方針の策定や安全確保の対策を実施し、必要に応じて見直しを行うこと ガイドラインに基づいたリスク管理活動の取り組み状況を定期的にモニタリングするとともに、ステークホルダーの求めに対して報告出来る状態を維持する
リスク	<ul style="list-style-type: none"> 自治体において、指針や規程、ガイドラインが無く自治体内での統一的な対策が実施できず、AI の問題が顕在化するリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI 提供者から提供されるポリシー、プライバシーポリシー等の方針に従い、リスク管理や安全性確保の対策を実施すること 自治体職員は上記状況を定期的に確認し、必要に応じて見直し・改定を行うこと

⑨ 教育・リテラシー

①リテラシー・スキルの確保	
内容	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者が、AI を効果的かつ適切に活用するためのリテラシーとスキルを確保できるように必要な措置を講じる
リスク	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者の AI に関するリテラシーやスキルが不十分なために、AI を利用した示唆の精度が低い、もしくは誤った文書を作成してしまうことで自治体業務に混乱を招いてしまうリスクや、住民に不利益を生じさせるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は AI に関する学習を通してその特性やリスク等に関して十分な理解に努め

	る
② フォローアップの実施	
内容	<ul style="list-style-type: none"> AI を活用する自治体のすべての関係者は、AI の特性やリスク等を十分に理解し、住民に対して AI 活用に係るメリットとそれに伴うリスクに対する理解を促すとともに、必要な措置を講じる
リスク	<ul style="list-style-type: none"> 住民が AI を利用する際、問題や疑問を抱えた場合の自治体によるサポートがなく、AI システム・サービスや自治体への不信に繋がるリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は住民に対して AI 活用に係るメリットとそれに伴うリスクを説明し、住民のリテラシー向上に努める 自治体職員は住民が問い合わせ・苦情を提起できるような仕組みを開設し、住民に公開すること

⑩ イノベーション

① 住民ニーズの把握・課題解決	
内容	<ul style="list-style-type: none"> AI の導入や活用を目的化せず、住民の目線で真の課題の把握やニーズの抽出を行うことにより、住民が抱える課題を解決し、ウェルビーイングに繋がる有効な方策を総合的にデザインする
リスク	<ul style="list-style-type: none"> AI の導入や活用自体が目的となり、地域や住民にとっての課題解決につながらないリスクがある
対策	<ul style="list-style-type: none"> 自治体職員は住民の目線で真の課題の把握やニーズの抽出を行う その上で、AIの導入を含めた、住民が抱える課題を解決しウェルビーイングに繋がる有効な方策を検討する 自治体職員は時間経過とともに住民のニーズに対する AI システム・サービスの有効性が劣化することのないよう定期的に住民アンケート等の効果測定を行い、結果に応じて、AI システム・サービスの改善を行うこと
③ ステークホルダーとの協働・連携	
内容	<ul style="list-style-type: none"> AI の活用にあたっては、多様なステークホルダーと緊密に協働・連携し、互いに補完し合い高め合うことにより地域社会の効用最大化や自治体業務行政運営の効率化を図る
リスク	<ul style="list-style-type: none"> AI の利活用が自治体職員の目線に寄りすぎてしまい、その他のステークホルダーにとって最適なものでなくなってしまうリスクがある
対策	<ul style="list-style-type: none"> 単独の自治体とその職員だけで AI の利活用を検討するのではなく、複数の地域・自治体や民間企業、アカデミアなど多様なステークホルダーが協力して地域社会の効用最大化や自治体業務の効率化を目指す

第6部 自治体におけるAI活用の今後の展望

自治体におけるAI活用は、今まさに始まったところである。本ガイドラインは、自治体職員が安心してAIの導入と活用をし、行政サービスを効果的・効率的に変革し続けるために貢献することを企図している。自治体の業務は、多岐にわたる。そして、税や各種扶助のようにセンシティブな情報を取り扱うこともある。この時、通常の事業者の論理や合理性を追求したAI活用に偏重してしまうと、自治体の使命である公共性が軽んじられてしまうリスクがある。そこで本ガイドラインでは、一般的なAIのガイドラインで取り上げられる内容に加え、特に、「住民第一」の視点を原則に据えている。

ただし、「住民第一」だけでなく、自治体が活用するAIサービスやAIプラットフォームを提供する事業者にとっての便益や、自治体で働く職員の仕事がより効果的・効率的になることも重要である。今後ますます自治体のAI活用が進み、その中で、こうした自治体を取り巻く多様な利害関係者の便益を全体的に最大化する成功事例が出てくるであろう。一方で、現時点で想定していない問題や課題も出てくるであろうし、AI技術自体も進歩していく。そこで、今後は産官学が強力で連携し、自治体におけるAI導入・活用の成功事例や失敗経験を蓄積・共有し、全国の自治体の業務革新を下支えするとともに、本ガイドラインを時代に合ったよりよいものへ修正し続けていかねばならない。

1世代70万人を切るに至った深刻な人口減少問題から目を背けず、産官学そして人間とAIが社会価値を共創し、地域特有の課題解決を促進し、持続可能な社会の実現に寄与する第一歩として、本ガイドラインによる自治体のAI活用が全国で進むことが期待される。