

人間主導A I（ええ愛）ガイドライン

令和6年6月

鳥取県

- 1 はじめに
- 2 A I 技術全般の利用について
- 3 生成A Iの利用について
 - （1）利用できる業務
 - （2）利用する上での遵守事項
 - （3）利用できない業務

【参考】自治体デジタル倫理原則

1 はじめに

本県は、A I－O C RやA I議事録作成支援、A Iチャットボットのほか、定型業務を自動化するR P Aなどのデジタルツールを行政事務に積極的に活用し、県庁D Xを推進することで、職員の新たな働く時間の創出や住民サービス向上など、さまざまな効果を生み出してきた。

近年のA I技術の進展は目覚ましく、地域課題解決や産業振興、内部事務の効率化等への有効活用が期待されている。2022年に登場した大規模言語モデルを代表とする対話型の生成A Iは、人間からのどんな指示や質問にも自然な言語で答えを返したり、画像や映像も生成できることから、様々な可能性を秘めた技術として世界的に注目されている。

A I技術の進化は今後さらに加速することが予想され、活用範囲も拡大することが見込まれている。A I技術は、豊かなデジタル社会を支えるキーテクノロジーとして、あらゆる分野のあらゆるアプリケーションや端末に実装されることが予想されている。社会や個人を支える目的であらゆるサービスのフロントヤードやバックヤードに、当然のように実装される時代がすぐそこまで迫っている。人類はすでにA I不可避な時代に突入しようとしている。

一方で、A I技術の普及拡大は様々なリスクや歪みも生み出している。特に顕著なのは、前述した対話型で回答を返すタイプの生成A Iである。生成A Iは、過去に学習したデータから確率で回答を導き出すアルゴリズムが特徴であり、その回答表現はわかりやすく明快だが、誤りや偏りが出力されてしまうことが問題視されている。また、プロンプトといわれるA Iへのリクエストによって意図しない情報漏えいが発生するリスクもある。生成された回答には、偽・誤情報のほか、人権や知的財産権の侵害、悪意のある者がマルウェア開発などの犯罪に悪用するなど、これまでのA I技術には存在しなかった新たな社会的リスクを生み出しているとの指摘もある。

地域が抱える真の課題やその課題の解決案は、現時点ではインターネット情報を学習データとする生成A Iからは決して出てくることはない。効率性や利便性だけを優先したA I活用が拡大すれば、民主主義や地方自治の根幹を脅かす脅威（リスク）となる。地域のことは地域で考え、地域で決定するのが民主主義や地方自治の要諦である。真に有効な施策を企画立案するためには、デジタル技術を有効活用しつつも、現場主義を貫き、県民・企業・団体・市町村等の声を丁寧に聴くことが重要である。

さらに、職員の過度なA I依存は、県職員としての経験を積み、基礎能力を獲得する上で障害になる可能性もあり、将来的に行政サービスの質の低下に繋がることも危惧される。いかなる先進技術であっても100%完全完璧はない。A Iに関連する様々なリスクの要因は、A I技術ではなく技術の使い手である人間側に内在することを忘れてはならない。

本県は生成A Iを含めたA I技術を職員が適切に業務活用するために最低限必要なルールを定めた「鳥取県庁A I活用ガイドライン（暫定版）」を令和5年8月に策定し、生成A Iの暫定運用を行ってきた。また、同年9月に有識者で構成する「先端技術と民主主義のあり方を考える研究会」を設置。民主主義や地方自治の視座から、A Iを含む先端技術への地方自治体としてのあり方、向き合い方を倫理面で研究し、令和6年4月、その成果として、人間主導のデジタル社会を目指す上で重要な視点を整理した「自治体デジタル倫理原則」（以下、「倫理原則」という。）が提言されたところである。

地方自治体は、A I技術の活用を加速させる中であっても、技術に過度に依存し、職員の心理面を含めA Iに支配されるような組織や社会を構築してはならず、信頼失墜となるような未来へ導いてはならない。

職員一人一人が、倫理原則が掲げる人間主導をはじめ、人権保障、透明性確保、インクルーシブ、パートナーシップ等の視点を持ち、仕事の質や住民サービスをより高い次元へと進化させるため、本ガイドラインを策定した。A Iがもたらす効果とリスクの両面を理解し、本ガイドラインを遵守しながら、A I技術を適切かつ積極的に業務に取り組むことが重要である。

なお、A I技術は短期間のうちにも大きな進展が起こり得る技術である。本ガイドラインは、必要によりこうした変化にアジャイル思考で適時見直しを行うものとする。

2 A I 技術全般の利用について

A I は、効率的かつ利便性の高い行政サービスの提供や行政事務の効率化が期待できる技術であり、分野を問わず様々な業務に積極的に活用することが有用である。一方で A I の技術特性を考慮し、活用の仕方や、収集、利用される個人情報の適正な取り扱いなどについて留意も必要がある。

生成 A I を含む A I 全般について、技術を活用する上で、遵守すべき事項を以下に示す。なお、文末の（ ）書きについては、参考に関連する倫理原則を示したものである。

- ① 原則、事務処理や情報発信の内容を A I の処理や出力結果に委ねることなく、職員が責任をもって制御できる仕組みで運用すること。（人間主導の原則）
- ② 自動運転等、自律的に運用される A I など、職員が制御することが困難な A I 技術を活用する場合には、その処理が A I によるものであることを明確にするとともに、処理結果を精査し、運用を管理すること。（透明性の原則、機敏性の原則）
 - ア 出力結果を生成するタイプの A I チャットボットや自動回答サービスなど、自律的に運用される A I に利用者が直接アクセスし、情報発信や回答を受け取るサービスを提供する場合は、その内容が A I が作成したものであり、誤答などが含まれる可能性があることを、利用者が事前に把握できるように明示すること。
 - イ 定期的に利用ログを分析し、目的に沿って A I が利用されているか評価するとともに、必要により適宜システムの見直しを行うこと。
- ③ A I 技術の業務利用にあたっては、意図せず他者の人権や知的財産権を侵害してしまうリスクがあることから、A I のロジックや出力傾向の把握に努めるなど、適正かつ円滑な業務の遂行に留意すること。（住民自治の原則、人権保障の原則、人間主導の原則）
 - 例えば、A I のアルゴリズムの特性等により、気付かないうちに差別や偏見を生み出してしまいう可能性がある。また、学習データに偏りや不足があった場合や、想定外の外的要因や一時的な要因を考慮できないことにより、不適当な結果を出力する可能性がある。
- ④ A I が秘匿情報や個人情報を無差別に収集することがないように、システム導入の計画段階から情報の取得や運用、管理のあり方を検討すること。（人権保障の原則、インクルーシブの原則、ガバナンスの原則）
- ⑤ 県民等に与える影響などを考慮し、業務リスクレベルを 4 段階に分類した。A I 技術の活用にあたっては、業務に内在するリスクを意識し、リスクレベルにあった適切な活用を行うこと。

ア 禁止業務

（ア）意思決定（住民自治の原則、人間主導の原則）

意思決定（方針判断）は、関連する周辺情報や法令規則等を踏まえて総合的に判断する必要があり、県民等へのアカウントビリティ確保の観点からも、いかなる場合であっても、その意思決定の判断そのものに A I を活用しないこと。（判断は必ず人間が行う）

（イ）民意集約（住民自治の原則、パートナーシップの原則）

地域のことは地域で考え、地域で決めるのが民主主義であり地方自治の要諦である。地域が抱える真の課題を職員が把握（情報収集）する際は、県民・企業・団体・市町村等と円滑なコミュニケーションを図りながら、現状を幅広く丁寧に意見を聴く現場主義を基本とするため、民意の集約に A I を活用しないこと。

（ウ）人権侵害につながる恐れがあるもの（人権保障の原則）

業務に A I 技術を導入することにより、人権侵害につながる恐れがあり、技術的な対応や運用等によりこれを回避することができない場合には、A I を活用しないこと。

イ 要注意業務（ガバナンスの原則、人権保障の原則）

業務の効率化等、A I の導入効果が期待されるもののリスクが高い業務を扱うシステム

については、運用におけるリスク管理やデータガバナンス（データ入手から管理、利用、廃棄まで適切に運用すること）、セキュリティの確保に特に注意すること。

例えば、県民の生命・財産に大きな影響を及ぼす、社会インフラの安全確保に係る業務や、教育など機微情報を取り扱う業務、各種給付金や支援制度に係る業務などが想定される。

ウ 要配慮業務（透明性の原則、人間主導の原則）

A I を直接利用者が取り扱う業務や、A I の出力結果をそのまま利用する必要がある業務については、A I の出力傾向の把握や定期的な動作記録の確認、分析を通じて、出力結果に責任をもてるように配慮する。

エ 積極活用業務（課題解決志向の原則）

その他の業務については、業務の効率化や課題解決に向け、倫理原則に基づいた積極的なA I 技術の利活用を推進する。

AI活用のリスクに即した運用基準

リスクレベル	内容
Lv4 容認できないリスク  禁止業務 民主主義や地方自治の本旨に反する活用を制限	▶ 意思決定・民意集約（住民の意見に基づき、人間が最終判断） ▶ 人権侵害の恐れがあるもの
Lv3 高リスク  要注意業務 リスク管理、データガバナンス、セキュリティ確保を特に注意し活用	▶ 生命・財産に大きな影響を及ぼす可能性のある社会インフラの運用・管理等の業務 （例）災害時の状況判断 ▶ 教育等、機微情報を取り扱う業務（例）自動採点システム
Lv2 限定的なリスク  要配慮業務 AIの出力傾向の把握等により、責任ある運用に配慮しつつ活用	▶ 県民等の直接利用を想定したAIサービス （例）自動応答型のAIチャットボット ▶ AIによる出力結果をそのまま提供する業務 （例）AI婚活マッチングシステム
Lv1 最小限のリスク  積極活用業務 自治体デジタル倫理原則を踏まえつつ、適切に活用	▶ 業務効率化や行政サービス向上に向け、積極的に活用 （例）AI音声認識議事録作成支援、AI-OCR

⑥ システム管理者は、職員がA I を適切に利用できるよう、必要な教育を行うこと。（リテラシーの原則、パートナーシップの原則）

A I を含むシステムを有効に構築・運用できるよう、本ガイドラインや倫理原則など関連規定についての的確に理解するために必要な教育を行うこと。

委託事業者等がA I を用いる場合には必要に応じ、利用範囲や対象について定めるとともに、利用状況を把握すること。

3 生成A Iの利用について

令和5年8月から職員向けに提供している「県庁生成A Iシステム」をはじめ、生成A Iは対話型のシステムにプロンプトと呼ばれるA Iへの命令・質問を入力して利用することができる。分野や目的を問わず自由に生成A Iを利用することができるが、回答に誤りが含まれるなど生成A I特有の問題や課題も存在するため、以下の留意点や遵守事項を踏まえて地方自治体として適正に利用していく必要がある。

(1) 利用できる業務

生成A Iを活用することにより、人間が作業した場合に手間がかかる作業を短時間で行うことができ、新たな取組のアイデア出しのヒントを得ることや業務の効率化が期待できる。職員は業務の効率化を図るため、本ガイドラインを遵守し、適切に利用すること。

本ガイドラインに則った利用であれば、利用方法を制限するものではないが、以下に想定している文章生成A Iの活用例を示す。

① アイデア出しのためのヒント収集

<活用例>

ブレインストーミングのように複数のアイデアを基に検討することを目的として、生成A Iにヒントを出力させる。

<留意点>

最終的なアイデアは利用者自身が検討すること。

② 作成した文書の校正・添削チェック

<活用例>

文章の誤字や脱字は作成した自分自身では気づきにくい。第三者の目として生成A Iに校正や添削のチェックをさせる。

<留意点>

機密情報や非公開情報を含んだ文章をそのまま入力しないこと。

③ プログラミングコード作成支援

<活用例>

ExcelのマクロやVBA、NotesのLotusScriptなど、様々なプログラミングにおいて、作成したいコードのサンプルコードの出力や、作成したコードの修正をさせる。

<留意点>

機密情報や非公開情報を含んだコードをそのまま入力しないこと。

④ 読み手を想定した言い換え

<活用例>

外国人向けの表現や子どもに配慮した平易な文章など、発信する場面や相手の年齢などに考慮した文章に言い換えさせる。

<留意点>

言い換えにより文意が変化したり、不適切な表現が含まれる可能性があるため、適切に確認、添削すること。

⑤ 文書の要約、翻訳

<活用例>

文字数の多い文章や難解な外国語の文章などについて、生成A Iに要約や翻訳をさせる。

<留意点>

特に翻訳の出力結果については、利用目的に合わせて適切に確認、添削すること。

⑥ 定型的文書の下書き

<活用例>

定型的な紹介文や送付文などの文書の一般的な言い回しを出力させる。

<留意点>

参考となる文章が出力されるよう、作成の目的や具体的な相手、場面などを適切に入力すること。

⑦ 統計データの簡易分類・傾向分析

<活用例>

「,」などで区切られたデータについて、項目ごとの平均値や最頻値などの傾向を出力させる。

<留意点>

機密情報や非公開情報を含んだデータを入力しないこと。

⑧ 各種検索

<活用例>

特定の事柄を調査するために生成A I を利用し、出力結果から調査の糸口やヒントを得る。

<留意点>

ある事柄の把握をするために生成A I に質問する場合、生成A I による回答の精度には限界がある。回答が正確である保証はなく、誤った回答であっても、自然な文章で回答する傾向にあることから、正確な情報を収集する目的では利用しないこと。

(2) 利用する上での遵守事項

利用にあたっては、以下の事項を遵守すること。

① 生成、出力したものをそのまま利用しないこと

ア 最終的な意思決定をA I に依存することがあってはならない。出力結果は業務の参考に止めるとともに、人間が的確にチェックする。例えば、事務処理や決裁過程において生成A I の出力結果を参考にした箇所を組織内で共有するなど、適切に確認すること。

※生成A I は、モデルが作られたときの学習内容に基づき、確率的に推定し生成を行う技術であるため、ジェンダーバイアスなど偏見が示唆されるような結果が出力される可能性がある。A I 技術のロジックや出力傾向には特に留意すること。

イ 県民等が出力結果を直接利用する場合には、A I に基づく出力であることを明示すること。

※A I による生成物は著作権が認められない可能性がある。また、一部加工を行った場合にも、当該加工部分以外の著作権が認められない可能性があることに留意が必要。

② ハルシネーションを考慮し、必ず出力結果の根拠を確認すること

生成A I は、確率的に推論された結果に基づいて出力結果を生成することから、あたかも事実であるかのように虚偽の内容が出力される「ハルシネーション」と呼ばれる現象が生じることがあるため、必ず根拠を確認すること。

③ 非公開情報（機密情報）はプロンプトに入力しないこと

生成A I への入力情報は、当該生成A I の学習情報に利用される可能性があるため、非公開情報（機密情報）は入力しないこと。

特に、個人情報や内部の情報システムに係る重要情報を入力しないこと

【禁止例】個人の住所や氏名、生年月日、財産その他個人を紐づける情報や、各種システムに係る構成情報及び設定情報等

④ 著作権侵害等が危惧される情報の生成指示及び転用は行わないこと

ア 著作物等、知的財産を利用して入力したり、これらが出力される可能性がある場合、

関連する法令や国等の検討状況を把握し、適切に取り扱うこと。

イ 著作権や知的財産権に係る法令の例外規定が適用されない場合や、名誉棄損など別の権利侵害にあたる場合があるため、こうした懸念が危惧される生成指示や生成された情報の転用は行わないこと。

⑤ 違法行為等、不正な目的で利用しないこと

ア 業務外の目的で利用しないこと。

イ 不正ソフトの開発や違法行為を目的とした利用は行わないこと。

ウ 信頼失墜に繋がる利用は行わないこと。

⑥ 「県庁生成ＡＩシステム」以外の生成ＡＩシステムを利用しないこと

プロンプト等により職員が出力を指示する生成ＡＩシステムについては、利用状況を把握し、適正に管理する必要があることから、職員は、上記タイプの生成ＡＩを活用する場合、原則、「県庁生成ＡＩシステム」を利用すること。

(3) 利用できない業務

「２⑤ア 禁止業務」に示したとおり、「意思決定」や「民意集約」は人間（職員）が責任をもって行う必要があることから、これらに係る業務を生成ＡＩに委ねてはならない。

【問合せ先】

生成ＡＩを含めたＡＩ利用や導入について不明点等がある場合は、デジタル局（0857-26-7968）に問い合わせること

【参考】自治体デジタル倫理原則

先端技術と民主主義のあり方を考える研究会でとりまとめられ、令和6年4月26日に公開された。AIを利用したシステムなどを利用する場合には、内容を把握し、適切に取り扱うこと。

(※参照：とりネット/鳥取県公式サイト <https://www.pref.tottori.lg.jp/317494.htm>)

10の原則	概要
住民自治の原則	地域のことは、住民の意思に基づいて検討や議論を重ね、決定することが原則である。生成AIをはじめとする先端技術の活用にあたっては、このような民主主義や地方自治の要諦が揺らがぬよう、適正な活用に徹すること
人権保障の原則	生成AIをはじめとする先端技術の活用にあたっては、利用目的をできる限り特定し、その達成に必要な範囲において個人情報収集するとともに、当該利用目的に限定して利用するなど、個人情報保護を含め住民の人権を守り、人権保障を具体化していく視点に立って厳正に行うこと SNS等においても、同様に住民の人権を守り、人権保障を具体化していく視点に立って、対策を講じること
インクルーシブの原則	生成AIをはじめとする先端技術は、ジェンダーや性的マイノリティに配慮するなど多様な人々が互いに尊重される社会を実現するために活用されなければならない、偏見等を生まないよう運用するとともに、住民に寄り添い、誰一人取り残されない行政サービスを提供するように配慮すること
パートナーシップの原則	生成AIをはじめとする先端技術の活用にあたっては、住民を含めた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を図ること
課題解決志向の原則	デジタル技術の導入や活用を目的化せず、住民等の一連の行動に着目して真の課題の把握やニーズの抽出を行うことにより、住民等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こと
人間主導の原則	地方自治体が行う意思決定を生成AIをはじめとする技術が出力した結果のみに依拠することは排し、出力結果を人間が的確にチェックすることが担保されるよう人間が責任をもって精査し、人間の判断で決定する仕組みとすること
リテラシーの原則	職員がデジタル技術の特性を理解し、適切に活用するためのリテラシーとスキルの向上を図ること。住民が偽・誤情報に惑わされないよう、住民の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、住民や地域を守るための情報発信に努めること
透明性の原則	住民が生成AIをはじめとする先端技術の活用について適切に評価できるよう、活用の状況を明らかにすること。特に、住民等への回答内容にAI等の出力結果が結びついている場合は、その旨を明示するなど説明責任を果たすこと
ガバナンスの原則	生成AIをはじめとする先端技術の適切な活用に向け、デジタル施策の実施状況や結果・評価を把握して効果を検証するとともに、AIのロジックや出力傾向等を把握した上で適正に管理する仕組みを構築して、適切に運用し改善していくこと
機敏性の原則	生成AIをはじめとする先端技術は急速に発展していくものと想定される。こうした変化に伴いルール等を見直す機会を整えるとともに、大胆かつ積極的に先端技術を取り入れていく視点を持ち、「完全な成功」よりも「試行と改良・再挑戦」を重視し、不断のフィードバックにより、アジャイルで機動的な先端技術の活用推進を図ること