資料２

生成ＡＩの利用に関するガイドライン

2023 年 11 月

名古屋市

目次

Ⅰ　はじめに

１　本ガイドラインの目的

本ガイドラインは、名古屋市職員が業務で生成 AI を利用する際に注意すべき事項を解説したものである。

少子化・高齢化の進行に伴う人口構造の変化、感染症に対する懸念の高まり、脱炭素化に向けた世界的な動きの加速など多様化・複雑化するさまざまな課題に的確に対応し、持続可能な形で行政サービスを提供するためには、行政のデジタル化を進める必要がある。

生成 AI は、行政業務の様々な場面に活用できる可能性がある一方で、入力したデータが AI に学習され、他者への回答に利用されるといった情報漏えいの危険性や、入力するデータの内容や生成物の利用方法によって生じる回答の不確実性、他者の権利侵害や事実とは異なる不正確な回答の生成など、様々な危険性が指摘されている。

こうした危険性を回避しながら、デジタル技術の導入目的である持続可能な形で行政サービスを提供するため、行政業務において生成 AI を利用するための指針として本ガイドラインを策定する。

なお、本ガイドラインについては、今後の国や社会の動向等を踏まえ、必要に応じて見直しを行っていく。

本ガイドラインをよく読み、市民の権利や個人情報などの財産をしっかりと守ることを前提に、生成 AI を利用すること。

２　本ガイドラインが対象とする範囲

- 本ガイドラインにおける生成 AI は、質問・作業指示（プロンプト入力）等に応えて文章・画像等を生成する AI を利用したサービスまたは当該サービスと連携して動作するプログラムとする。
- 本ガイドラインの対象となる組織は、名古屋市情報セキュリティ対策基準４の範囲と同様とする。

Ⅱ　生成 AI の活用方策

特に、文章生成 AI における活用方策を以下に示す。

1　推奨する活用例

庁内業務において、生成 AI の活用により効果が見込まれると想定される具体例を以下に示す。なお、生成 AI の用途を限定するものではない。

○文章の作成

キーワードや文字数などの条件に基づいた文章を作成できるため、文章の下書きとして会議の挨拶文やメール等の文案の作成に活用できる。職員自身では気づけない文章表現が提示されるなど、文章表現の幅を広げることにもつながる。

○文章の要約

外部の会議録やアンケートなど情報量が多い文章の要約について、押さえるべきキーワードや集計の仕方を指定することで、効率的、多角的に情報を整理することにつながる。

○文章の翻訳

高い精度での翻訳ができるほか、文章の趣旨は変えずに文章表現を変更させるなど一般的な翻訳ツールより効果的な使い方が簡易にできる。

○文章の添削

入力された文章の校正や誤字脱字のチェックなどに活用できる。

○アイデア創出

膨大な学習データ等に含まれる様々な情報に基づいた回答を参考にして、より多くの視座から検討することができる。

○Excel の関数やマクロのコード等の作成

Excel 等で実行したい内容を指示し、回答として得られた関数やマクロのコード等を参考にして、専門知識がなくとも、より高度な情報処理ができる。

２　生成 AI を活用する上でのポイント

○正確かつ詳細な情報の入力

AI は入力された内容の行間を読まないため､AI に適切な指示を与えるためには、正確かつ詳細な情報を入力する必要がある。

例えば、AI に出力してほしい回答について、どの立場での回答か、どういった目的で使用するか､どういった形式で回答を出力するか等を､入力の中で明示すると、より思っているものに近い回答を得られることができる。

○回答の精度を高めるための手法

得られた回答の深掘りや条件を追加して再度回答を求めるなど生成 AI との対話を繰り返すことで､回答の精度の向上や､より詳細なアイデアの検討を行うことができる。

○有効なプロンプト（AI に対する指示）の例

別紙参照

Ⅲ　利用にあたっての条件等

１　生成 AI の利用条件

・　本ガイドラインが対象とする生成 AI は、名古屋市情報あんしん条例（平成 16 年名古屋市条例第 41 号）における外部サービスにあたることから、生成 AI の利用にあたっては、名古屋市外部サービス利用基準を遵守すること。

・　生成 AI とのやり取りの内容は、将来、情報漏えいや他者の権利侵害等のトラブルが発生した際の証拠として必要になることもあるため、生成 AI を利用する場合は、原則として入出力内容等を記録する機能を有しているサービスを利用することが望ましい。

・　生成 AI とのやりとりが、AI で応答を生成するためのデータとして利用される可能性があり、個人情報や機密情報の情報漏えいのリスクが考えられるため、入力内容を AI の学習内容に反映させないサービスや機能を選定すること。

・　生成 AI は名古屋市の状況、各地域の状況の詳細を把握しているものではない。また、生成 AI の出力には、後述のように、誤りや偏りのある意見等を含む可能性がある。そのことを十分に認識し、業務遂行に当該出力をそのまま用いることはしないこと。

２　データを入力する際の禁止事項

情報漏えいを防ぐため、機密情報（名古屋市情報あんしん条例施行細則第 28 条第 1 項第 1 号に該当する情報）を生成 AI に入力することを禁止とする。

３　生成された回答を利用する際の注意事項

（１）生成物の内容に誤りが含まれている可能性がある

・　ChatGPT 等の大規模言語モデル（LLM）の原理は、「ある単語の次に用いられる可能性が確率的に最も高い単語」を出力することで、もっともらしい文章を作成していくものであり、書かれている内容には誤りが含まれている可能性があるため、必ず事実確認（ファクトチェック）を行うこと。

・　生成 AI は、インターネット上の情報を基に学習していることが多いため、生成される回答は、多数派の意見が尊重され、少数派の意見が反映されにくい傾向にある。そのため、回答には差別・偏見等のバイアスが含まれている可能性があり、その回答に基づいた判断をしてしまうことによって個人及び集団が不当に差別されないよう注意すること。

・　生成 AI は学習データにないことは答えられない。例えば、2021 年 9 月までのデータで学習した生成 AI であれば、それ以降に発生した事項については答えられないか誤った回答をしてしまう。生成 AI の学習の元データの範囲をきちんと確認すること。

・　生成 AI のこのような限界を知り、その生成物の内容を盲信せず、必ず根拠や裏付けを自ら確認すること。

（２）AI に判断や責任を負わせることはできない

・　生成 AI は業務執行にあたっての単なる補助的なツールに過ぎないため、過度に依存することなく、業務における検討・判断の責任は人間である各職員にあることを理解して利用すること。

・　生成 AI の生成物又は派生物（生成物を参考に作成したコンテンツ）に対して何らかの責任を負わせることはできない。生成物又は派生物を外部に公表する際は、市が説明責任を負うことを踏まえ、適切に判断すること。

（３）生成物を利用する行為が誰かの既存の権利を侵害する可能性がある

○著作権侵害

生成 AI からの生成物が、既存の著作物と同一・類似している場合は、当該生成物を利用（複製や配信等）する行為が著作権侵害に該当する可能性がある。

そのため、以下の留意事項を遵守すること。

・特定の作者や作家の作品のみを学習させた特化型 AI は利用しない。
・プロンプトに既存著作物、作家名、作品の名称を入力しない。
・特に生成物を「利用」（配信・公開等）する場合には、生成物が既存著作物に類似しないかの調査を行うようにする。

○商標権・意匠権侵害

画像生成 AI を利用して生成した画像や、文章生成 AI を利用して生成したキャッチコピーなどを商品ロゴや広告宣伝などに使う行為は、他者が権利を持っている登録商標権や登録意匠権を侵害する可能性があるので、生成物が既存著作物に類似しないかの調査に加えて、登録商標・登録意匠の調査を行うようにすること。

○誤った個人情報・名誉毀損等

生成 AI は、個人に関する誤った情報を生成する可能性があることが知られている。誤った個人情報を生成して利用・提供する行為は、個人情報保護法違反（法第 19 条、第 20 条違反）や、名誉毀損・信用毀損に該当する可能性があるので、必ず事実確認を行うようにすること。

（４）生成物について著作権が発生しない可能性がある

生成 AI を利用して生成された生成物の著作権については、生成 AI を利用しての創作活動に人間の「創作的寄与」があるか否かによって結論が分かれ、著作権が発生しない場合がある。

仮に生成物の著作権が発生していないとすると、当該生成物は基本的に第三者に模倣され放題ということになるため、自らの創作物として権利の保護を必要とする個人や組織にとっては大きな問題となる。

このため、生成物をそのまま利用することは極力避け、できるだけ加筆・修正するようにすること。

（５）生成 AI のポリシー上の制限に注意する

生成AIにおいては、これまで説明してきたリスク（主として法令上の制限）以外にも、サービスのポリシー上独自の制限を設けていることがある。その場合には、当該独自の制限に従う必要がある。

サービスによっては、生成物を公開する際にあたかも人間が生成したものであるかのように表示することを禁止し、AIが生成したものであることを明示する義務が定められている場合もある。AIによる生成物を公開する場合には、このような義務の有無を確認し、必要な際には、AIが生成したことの明示を行うか、内容を加工するなどしたうえで公開すること。

附　則

（施行期日）

このガイドラインは、令和 5 年 11 月 7 日から施行する。