機密性２　政府機関等職員限り

ChatGPT 等の生成 AI の業務利用に関する申合せに関する Q&A

2023年（令和5年）10月6日
AI 戦略チーム事務局

「ChatGPT 等の生成 AI の業務利用に関する申合せ」（以下、「申合せ」という。）に基づく関係省庁における対応を円滑に行うため、申合せの運用に関する関係省庁向けの Q&A を作成する。

今後の運用状況を踏まえ、逐次、本 Q&A を更新することとする。

Ｑ１　生成 AI とは何か
Ｑ２　AI 戦略チームに対する報告や了解を求めている趣旨如何
　Ｑ２－１　所管の独立行政法人等の取組についても報告が求められるのか
Ｑ３　生成 AI 機能を有するサービスを業務利用する場合、全て報告を求めるのか
　Ｑ３－１　民間事業者に対する委託事業において、報告が必要になるケースはあるのか
Ｑ４　どの時点で報告が必要なのか
Ｑ５　申合せにおける機密性２情報とはどのような情報を想定しているのか
Ｑ６　具体的な報告方法を教えてほしい
Ｑ７　これまでの報告事例について、共有して欲しい
Ｑ８　運用の見直しにおける今後の予定如何
Ｑ９　報告内容は AI 戦略チームにおいてどのように扱われるのか
Ｑ10　報告内容は公表されるのか
Ｑ11　「サービスにおいて生成 AI を利用していることの明示」の詳細について
Ｑ12　「生成 AI の出力結果を二次利用する場合の責任の明確化」の詳細について
Ｑ13　「当初は行政職員、自治体職員など対象となる利用者層を限定」の詳細について
Ｑ14　「学習に利用するデータ、入力され得るプロンプト、出力結果の社会的影響に係るリスク評価の実施」の詳細について
Ｑ15　「入力されたプロンプト及び出力結果のロギングを行った上で、必ず利用者からフィードバックを受ける仕組みを設けること」の詳細について
Ｑ16　「一般利用者を対象とする場合は検証段階であることの明示とテスト参加の同意の取得」の詳細について
Ｑ17　（様式）利用計画書の記載方法及び記載例を示して欲しい

Ｑ１　生成 AI とは何か

Ａ．明確な定義が存在する訳ではないが、「データから学習し、独創的かつ現実的な新しい成果物を生成する AI」といった見解が示されており、例えば、画像を生成できる AI の「Stable Diffusion」や、テキストを生成できる AI の「ChatGPT」、機械的翻訳を行う AI の「DeepL」が該当すると考えられる。明確な定義が存在する訳ではないものの、機械学習や信号処理技術を用いていても、以下のような場合は、生成 AI には該当しないと考えられる。

- 予測値等の数値情報しか出力できない場合
- 文章分類等の事前定義されたカテゴリやラベルしか出力できない場合
- 工場ロボットの運動制御等、デジタル信号を制御するために用いられる場合

Ｑ２　AI 戦略チームに対する報告や了解を求めている趣旨如何

Ａ．昨今の ChatGPT 等の生成 AI を巡る技術革新は、さまざまな利点をもたらす一方、プライバシーや著作権の侵害などの新たな課題が生じるとの見方もある。生成 AI を巡る様々な課題や規制の在り方に関しては、国際的にも議論が行われているところ、政府としては、そうした議論の動向を見極めつつ、関係省庁が連携して生成 AI に関する実態の把握に努め、適切な措置を講じていく必要がある。

Ｑ２－１　所管の独立行政法人等の取組についても報告や了解の対象となるのか

Ａ．　「政府機関等のサイバーセキュリティ対策のための統一基準」（令和５年度版）（以下「政府統一基準」という。）※の対象となる独立行政法人等についても、申合せの趣旨を踏まえた適切な対応が求められる。AI 戦略チームに対する報告については、例えば、ある独立行政法人における利活用が、所管省庁における対応と同じ範囲であれば、各独法分に関する個別案件毎の一律での報告や了解は不要な場合もあると考える。所管省庁が活用せず、所管の独立行政法人のみが活用する場合など、判断に迷う場合は、個別に相談されたい。

※政府統一基準に係る問合せについては、内閣サイバーセキュリティセンター宛てに相談されたい。

Ｑ３　生成 AI 機能を有するサービスを業務利用する場合、全て報告を求めるのか

Ａ．今回の申合せは、生成 AI の固有のリスクを踏まえ、関係省庁が連携して生成 AI に関

する実態の把握に努め、適切な措置を講じていくためのものである。そのため、生成 AI を業務において活用することで、どのようなリスクが発生するのかを把握し、適切なリスク軽減策を講じることが必要。例えば、関係省庁において、文書生成 AI を用いて関係省庁の名において公表する文書を作成する場合は、当該文書の信頼性の確保等の観点から、関係省庁が連携して適切な対応（生成 AI の利用状況や、信頼性確保のために講じる措置に係る対外的な説明等）を取ることが必要であり、報告や了解の対象となる。その他にも、例えば、生成 AI 機能を搭載する「検索エンジン」を業務利用する場合において、報告が必要か否かについても、同様に、生成 AI 固有のリスクが発生しているか否かの観点から判断していくことになる。なお、申合せに基づく報告の要否に関わらず、当然、政府統一基準に基づく関係省庁における適切な対応は必要。

なお、申合せに記載のとおり、要機密情報を含まない情報の取り扱いを前提とした、生成 AI の利用にあたっては、組織の規程に則り承認を得た場合に限り、「AI 戦略チーム」への報告を不要とする。また、機密性１情報についても、大量に入力した際には情報漏洩等のリスクがあるため、生成 AI の取扱いにおける人材育成（職員研修等）に努めること。一方で、機密性２情報については、約款型外部サービスでない形態において、申合せに記載の対応を行ったうえで「AI 戦略チーム」への報告及び了解を得た場合に限り、生成 AI の利用を可能とする。

個別の報告の要否については、本 Q&A を更新することにより、明らかにしていくため、判断に迷う場合は、事務局まで相談されたい。

Ｑ３－１　民間事業者に対する委託事業において、委託事業者が委託事業とは別に、自らの環境下で生成 AI の実証を行う場合は報告が必要か

Ａ．報告は不要。なお、行政機関職員が業務として利用する場合、利用を開始する前に報告を求める。

Ｑ４　どの時点で報告が必要なのか

Ａ．申合せに記載の通り、関係省庁におかれては、個別契約等、約款型外部サービスでない形態で、機密性２情報の生成 AI 利用を検討する場合には、組織の承認を得た上でその検討状況を報告いただきたい。

Ｑ５　申合せにおける機密性２情報とはどのような情報を想定しているのか

Ａ．機密性２情報のうち適切なリスク分析を行った情報であり、例えば、以下の業務等に利用することを想定している。

- 将来の公表を予定している文書（想定問答、白書等）の下書きの作成

- 国際会議（公開されるもの）における情報収集、翻訳、提案の下書きの作成
- 自治体職員からの質問に対する回答案の作成
- 技術文書の作成、ソフトウェア開発運用保守、インフラ構成管理の支援

Ｑ６　具体的な報告方法を教えてほしい

Ａ．以下の方法でAI戦略チーム宛に報告されたい。

１．報告内容
（１）外部サービスの名称、外部サービス提供者の名称、利用目的（業務内容）、取り扱う情報の機密性、利用期間、利用者の範囲、利用申請の許可権限者等。
様式については、別紙参照。

（２）生成AIに係る固有のリスクやセキュリティ懸念への対応
関係省庁における申請・届出の際に用いられた様式内に記載されている場合はそれをもって確認。記載されていない場合は、必要に応じて追加説明を求める。

２．報告先
AI戦略チーム事務局（内閣府CSTI・デジタル庁）

【連絡先】[redacted]

３．結果報告
報告受理後、AI戦略チーム内で協議し、その結果を報告主体宛てにメール等で通知する。
※必要に応じて追加で説明をお願いする可能性あり。

Ｑ７　これまでの報告事例について、共有して欲しい

Ａ．個別の報告事例における共有可否については、報告を行った関係省庁における判断となる。個別に事務局宛に相談いただきたい。

Ｑ８　運用の見直しにおける今後の予定如何

Ａ．Ｑ３で記載の通り、今回の申合せは、生成AIの固有のリスクを踏まえ、関係省庁が連携して生成AIに関する実態の把握に努め、適切な措置を講じていくためのもの。そのため、申合せに関する運用状況を踏まえ、本Q&Aにおいて、個別報告不要な類型を示す等、

具体的な報告内容や方法については、随時見直しを行っていく。

Ｑ９　報告内容はAI戦略チームにおいてどのように扱われるのか

Ａ．AI戦略チーム事務局において、AI戦略チーム構成員へ諮り、了解を取る。

Ｑ10　報告内容は公表されるのか

Ａ．AI戦略チームとしては、現時点において、報告内容自体を直ちに公表することは想定していない。そのため、公表有無については、関係省庁において適切に判断されたい。なお、報告があった省庁名については、当該報告をAI戦略チームに行った後に、公表する可能性がある。その場合においても、公表可否については、報告があった省庁に確認を取るものとする。

Ｑ11　「サービスにおいて生成AIを利用していることの明示」の詳細について

Ａ．サービス自体が、生成AIを利用していることを、当該サービスの利用者に明示することが必要だが、明示する場所、方法等は問わない。

ここでの「利用者」とは、サービスに情報を入力する者、または、出力結果を一次利用する者をいう。（二次利用についてはＱ12を参照）

また、ここでの「サービスにおいて生成AIを利用している」場合とは、生成AI機能を有するサービスが、提供するサービスに含まれている場合を示す（「生成AI機能を有するサービス」については、Ｑ３を参照）。

なお、外部公開を前提とせず、組織の内部においてのみ利用する場合であっても、対応が必要である。

Ｑ12　「生成AIの出力結果を二次利用する場合の責任の明確化」の詳細について

Ａ．生成AIの出力結果を二次利用する場合においては、利用者の責任において利用されるべき旨を明確化することが必要であることから、各府省庁において適切な対応が求められる。

なお、ここでの「二次利用」とは、出力結果の内容を確認し、業務等に利用する場合をいう。出力された内容を確認したら、そのまま使うにせよ加工するにせよ、それは二次利用になる。

具体例としては、次のとおり。

① 職員Ａが、生成AIで調査を行う（一次利用）

② 職員Ａが、その出力をベースに報告書を作成する（二次利用）

③ 職員Ｂが、当該報告書を参照して業務を行う

※この場合における報告書の内容については、作成者である職員Ａが責任を負う。

Q13 「当初は行政職員、自治体職員など対象となる利用者層を限定」の詳細について

Ａ．生成AIの業務利用にあたっては、利用者を限定していることが前提となる。

ここでの「当初」の具体的な期間については、今後の利用実績等を踏まえ、AI戦略チームにて検討する。

なお、「自治体職員など」と例示しているが、自治体職員を含めなければならない訳ではない。

Q14 「学習に利用するデータ、入力され得るプロンプト、出力結果の社会的影響に係るリスク評価の実施」の詳細について

Ａ．ユースケースに応じて様々なリスクがあることから、生成AIの固有のリスクを踏まえ、各府省庁において、プロジェクト責任者が適切なリスク分析を行い、適切な対応が求められる。

ここでの「社会的影響」とは、現時点において生成AIの出力結果に不安定さがあることを鑑み、利用者における事実確認なく二次利用された場合に、偽情報・誤情報・偏向情報等が社会に発信され、社会を不安定化・混乱させるリスクがあること等を示す。（生成AIの固有のリスク等については、「AIに関する暫定的な論点整理」（2023年5月26日）参照）

具体的な対応策としては、様々なリスクがあることを踏まえ、そのまま資料作成に利用するのではなく、内容について十分に確認した上で参考として利用することが考えられる。また、出力結果の正確性の調査および検証を行うことや、出力結果をあくまでも回答案の作成等、補助ツールとしての活用にとどめること等が考えられる。

なお、リスク分析の調査結果等の記録・保存については、各府省庁の行政文書規定に基づき、各府省庁において適切に判断されたい。

Q15 「入力されたプロンプト及び出力結果のロギングを行った上で、必ず利用者からフィードバックを受ける仕組みを設けること」の詳細について

Ａ．例えば、問合せ窓口を置きつつ、適宜利用者からのフィードバックを受けられる仕組みを構築すること等が想定される。

なお、ログ結果等の記録・保存については、各府省庁の行政文書規定に基づき、各府省庁において適切に判断されたい。

Q16　「一般利用者を対象とする場合は検証段階であることの明示とテスト参加の同意の取得」の詳細について

A．ここでの「一般利用者」とは、サービスを利用する国民を示す。「一般利用者を対象とする場合」とは、出力結果を一般利用者が直接利用（一次利用）する場合に限る。
個別については、AI 戦略チームに利用計画書を提出し、ご相談いただきたい。

Q17　（様式）利用計画書の記載方法及び記載例を示して欲しい

A．利用計画書の記載方法及び記載例は次のとおり。

（様式）利用計画書　記載例**（検討中の部分については、検討状況を記入してください。）**

報告番号	（記入不要）←AI 戦略チーム事務局で、収受した順に付番させていただきます。
報告主体	○○省　○○局　○○課 ←利用主体ではなく、報告主体の組織名等を記入してください。
初回報告日	２０２３年１０月　１日
外部サービスの名称	未定【個別契約型】 本件はプロポーザル型企画競争にて契約先を決定する予定であるため、現時点では未定。 ←サービスが決定している場合は、サービス名を記入してください。未定の場合は、その理由も付して記入してください。
外部サービス提供者の名称	未定 本件はプロポーザル型企画競争にて契約先を決定する予定であるため、現時点では未定。 ←サービス提供者が決定している場合は、サービス提供者を記入してください。未定の場合は、その理由も付して記入してください。
利用目的（業務内容）	本件は、行政機関（国・地方自治体）等から寄せられる○○に関する問合せについて、ガイドラインや過去の回答内容等を参考に、生成 AI を活用して回答案を作成することで、業務効率化を図る。
取り扱う情報の機密性	機密性２情報（国・自治体職員からの質問に対する回答内容等）
利用期間	２０２３年１１月１日（予定）～２０２４年３月３１日（予定） ←開始時期と終了時期を記入してください。未定の場合は「未定」と記入してください。
利用者の範囲	○○課内職員１０名 ←部署名等を具体的に記入してください。

生成 AI に係る固有のリスクやセキュリティ懸念への対応	・運用等にあたっては統一基準に基づくデジタル庁のセキュリティポリシーを遵守する。 ・生成 AI に係る固有のリスクを踏まえ、出力結果をそのまま資料作成に利用するのではなく、内容について十分に確認した上で参考として利用する。 ←生成 AI に係る固有のリスクへの対応策、各府省庁のセキュリティポリシーに従って個別にリスク管理を行っていること等を、記入してください。

要機密情報の取り扱いに関する確認事項への対応	（はい・いいえ）
・サービスにおいて生成 AI を利用していることの明示	はい
・生成 AI の出力結果を二次利用する場合の責任の明確化	はい
・当初は行政職員、自治体職員など対象となる利用者層を限定	はい
・学習に利用するデータ、入力され得るプロンプト、出力結果の社会的影響に係るリスク評価の実施	はい
・入力されたプロンプト及び出力結果のロギングを行った上で、必ず利用者からフィードバックを受ける仕組みを設けること	はい
・一般利用者を対象とする場合は検証段階であることの明示とテスト参加の同意の取得	はい

（自由記述欄）

- 契約したサービスを利用する際における、利用上の注意として、生成 AI を利用していることを明示する。
- 生成 AI の出力結果を二次利用する場合においては、利用者の責任において利用されるべき旨を利用者に周知する。
- 「利用者の範囲」に記載のとおり、利用者を〇〇に限定している。
- 本件に係るリスクを分析した結果、〇〇のリスクがあることが判明したため、その旨プロジェクト責任者に報告済。今後、新たなリスクが顕在化した場合には、その都度、当該リスク低減措置について検討し、対応する。また、リスクを踏まえ、生成 AI の出力結果については、利用者が内容を十分に確認した上で利用する。
- 利用規約において問合せ先を〇〇課〇〇係として示し、適宜利用者からのフィードバックを受けられるように構築する。
- 本件は一般利用者を対象としていないため、検証段階であることの明示とテスト参加の同意の取得は不要。

←上記確認事項への対応を具体的に記入してください。