調達チェックシート						（参考）要求事項の参考										
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	対策例の対象AI 入力 テキストを含む	 音声を含む	 出力 テキストを含む	 画像を含む	 音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	AWSサンプル回答
組織要件	1	AI事業者ガイドライン共通の指針の遵守	基本項目	1	AI事業者ガイドライン共通の指針を理解・把握・対応していることの宣言が可能であること	●	●	●	●	●	AI事業者ガイドライン「共通の指針」のチェックリストについて提出及び対応状況を説明可能とする	-	AI事業者ガイドライン別添7Aのチェックリスト等	「調達仕様書標準テンプレート」 8.入札参加に関する事項 8.1.公的な資格や認証等の取得	総務省、経産省「AI事業者ガイドライン 第1.2版」	※本チェックシートへの回答はGenerative AI Use Case JP (GenU)（https://github.com/aws-samples/generative-ai-use-cases)を想定したものとなります(以下の設問も同様) AWSは、AI事業者ガイドライン共通の指針のチェックリストの各項目に対して、以下のように対応しています： １．人間中心 Amazon Bedrockの利用規約（AWS Service Terms 50.12）で、人間の尊厳を脅かす用途での利用を制限しています。また、Amazon Bedrock Guardrailsを利用することで有害な出力を制御することができます。 ２．安全性 Amazon Bedrock は、様々な安全対策を組み込むことで、安全で信頼性の高い AI アプリケーションの開発を促進するように設計されています。また、Amazon Nova等、AWSが提供している基盤モデル（FM）にはAI Service cardsを提供しており、適切な利用方法と制限事項を明示しています。Claude等のサードパーティのFMに関しては、プロバイダーのウェブサイトで提供されています。 ３．公平性 Amazon Bedrock における責任あるAIの重要な要素として「公平さ」があり、これは異なるユーザーグループへの影響を考慮し、継続的なモニタリングとバイアス検出を行うものです。開発者は機械学習のライフサイクル全体で評価とバリデーションを実施し、Amazon Bedrockの組み込みモデル評価機能を使用して公平性を評価できます。具体的には、BOLDデータセットを用いて職業、性別、人種などの観点からモデルの公平性の測定と改善が可能です。また、AI Service cardsでモデルの特性とバイアスのリスクを開示しています。 ４．プライバシー保護 Amazon Bedrock では、データプライバシーとローカライゼーションの観点から、AWS はお客様のデータを保存しません。保存しないため、漏洩することも、モデルベンダーに見られることも、AWS が他の目的で使用することもありません。お客様は、Amazon Bedrock ガードレールの機密情報フィルターを使用することでプライバシーを保護することができます。また、Amazon Bedrock では、利用可能なリソースとモデルの可用性を最大化することで、カスタマーエクスペリエンスを最適化するクロスリージョン推論の仕組みを提供しています。このクロスリージョン推論を使用した場合でも、ユーザーのアクセスログ・推論ログなどのユーザー情報はリージョン内に保持され、推論処理のみクロスリージョンで実行されます。 ５．セキュリティ確保 Amazon Bedrockでは、データの転送時・保管時の暗号化、IAMによる詳細なアクセス制御、AWS PrivateLinkによるプライベート接続をデフォルトで提供しています。また、入力データは指定リージョン内に留まり、AWS及びモデルベンダーによる保存や他目的での利用は行いません。さらに、AWS KMSによる暗号鍵管理では、AWSマネージド型と顧客管理型の選択が可能です。 ６．透明性 AWSではAI Service cardsを通じて、AIサービスの透明性を確保しています。Service cardsでは、想定される使用事例、制限事項、責任あるAI設計原則、およびベストプラクティスなどの重要情報を一箇所に集約し、包括的な情報開示を行っています。これにより、お客様はサービスの特性を十分に理解したうえで、責任ある方法でシステムを構築することが可能です。 ７．アカウンタビリティ GenUでは、複数の方法で説明可能性とトレーサビリティを確保しています。Chain of Thoughtプロンプティング技術により推論プロセスの可視化が可能です。特に、Amazon Bedrock Knowledge Baseを利用したRAG実装では、生成された応答に対する情報源の明示や、取得されたテキストチャンクの提示により、出力情報の出所を明確に確認することができます。これらにより、透明性の向上とハルシネーションの最小化を実現しています。 ８．AIガバナンスとポリシー AWSは、Amazon Bedrockを通じて包括的なAIガバナンスの枠組みを提供しています。Amazon Bedrockは、AWS生成AIベストプラクティスフレームワークv2（https://docs.aws.amazon.com/audit-manager/latest/userguide/aws-generative-ai-best-practices.html）と統合されており、責任、安全性、公平さ、持続可能性、レジリエンス、プライバシー、セキュリティ、正確性という8つの原則に基づく評価が可能です。また、ISO、SOC、FedRAMP moderate、PCI、ISMAP、CSA STAR Level 2などの主要なコンプライアンス基準に対応し、GDPRへの準拠も確認されています（CISPE CODEのPublic Registerに登録）。さらに、モデル呼び出しのログ記録機能により、AIシステムの使用状況と性能の透明性を確保し、ステークホルダーによる情報に基づいた意思決定を支援しています。 （参考） ・AWS Service Terms:https://d1.awsstatic.com/legal/awsserviceterms/AWS_Service_Terms_05.02.25_JA-JP.pdf ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・Amazon Bedrock アプリケーションで責任ある AI のコアディメンションに対応するための考慮事項：https://aws.amazon.com/jp/blogs/news/considerations-for-addressing-the-core-dimensions-of-responsible-ai-for-amazon-bedrock-applications/ ・Amazon Bedrock モデルの評価が一般的に利用可能になりました：https://aws.amazon.com/jp/blogs/news/amazon-bedrock-model-evaluation-is-now-generally-available/ これらの対応の詳細は、AWS公式ドキュメント（https://docs.aws.amazon.com/bedrock/）およびAI Service cards（https://aws.amazon.com/jp/ai/responsible-ai/resources/#service）で確認可能です。 なお、開発・運用時の具体的な実装や運用方法については、事業者の担当範囲となります。
	2	AIガバナンスの構築	基本項目	2	生成AIシステムの開発・運用に関して、AIガバナンス（※）が適用されていること ※AIにもたらされる正のインパクトを最大化しつつ、AIによるリスクを受容可能な水準で管理する統制の仕組み・業務	●	●	●	●	●	生成AIシステムの開発・運用にAIガバナンスのゴールを考慮する	-	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等	「調達仕様書標準テンプレート」 8.入札参加に関する事項 8.1.公的な資格や認証等の取得	総務省、経産省「AI事業者ガイドライン 第1.2版」	GenUでは、Amazon Bedrockを通じて、以下のAIガバナンスの実現を支援しています： 1. AIガバナンスのゴールへの対応機能 Amazon Bedrockでは、コンテンツフィルタリングやプロンプトインジェクション対策などのガードレール機能を提供し、モデル評価ツールによってパフォーマンスとバイアスの評価が可能です。これにより、AIガバナンスのゴールに沿った開発・運用の実現を支援します。 2. ガバナンス状況の可視化機能 AWS Audit ManagerとAI Service Cardsを通じて、システムの動作状況や制御の有効性を可視化し、ガバナンスのゴールからの乖離やリスクを評価することができます。また、モデル呼び出しのログ記録機能とKnowledge Baseによる情報源の追跡機能により、管理状況の説明可能性を確保します。 3. 管理手法の検証機能 Amazon CloudWatchを活用し、設定された管理手法の有効性を監視・検証することができます。これにより、ガバナンス施策の適切性を継続的に評価することが可能です。 なお、以下の事項については、事業者側での対応が必要となります： - AIガバナンスのゴール設定と評価基準の策定 - 具体的なリスク評価手法の確立 - 是正対応プロセスの整備 - 管理状況の説明方法の確立 - 検証方法と基準の設定 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・AWS Audit Manager が生成系 AI 向けの AWS のベストプラクティスフレームワークを提供開始：https://aws.amazon.com/jp/blogs/news/aws-audit-manager-launches-aws-best-practices-framework-for-generative-ai/ ・責任ある AI： https://aws.amazon.com/jp/machine-learning/responsible-ai/ ・Amazon Bedrock のパフォーマンスのモニタリング：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/monitoring.html
						●	●	●	●	●	生成AIシステムの開発・運用において、AIガバナンスのゴールからの乖離やリスク評価を行い、是正対応を行う	生成AIシステムの開発・運用にあたって、AIガバナンスのゴールからの乖離を特定する	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												生成AIシステムの開発・運用にあたって、関連するルールからの逸脱を確認したり、リスクを評価する	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												AIガバナンスのゴールからの乖離の特定やリスク評価にあたって、必要に応じて外部関係者も含める	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
						●	●	●	●	●	生成AIシステムの開発・運用における管理について、説明可能性を確保する	適切かつ合理的な範囲で、生成AIシステムの開発・運用における管理の状況について関連するステークホルダーに対して説明可能な状態とする	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												生成AIシステムの開発・運用における説明可能性を高めるため、下記のような取組を行う □AIガバナンス・ゴールの乖離評価プロセスの実施状況の記録 □生成AIシステム・サービスの開発・提供・利用に関する社内/外部との会議記録の保持(担当者以外の関係者も閲覧可能な状態を確保) □AIに関する社内研修の実施	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												生成AIシステムの以下の事項について定められ、ユーザーに適時かつ適切に情報を提供する ・適切/不適切な使用方法等 ・提供する生成AIシステムの技術的特性、安全性確保の仕組み、利用によりもたらす結果より生じる可能性のある予見可能なリスク及びその緩和策等の安全性に関する情報 ・生成AIシステムの学習等による出力又はプログラムの変化の可能性 ・生成AIシステムの動作状況に関する情報、不具合の原因及び対応状況、情報セキュリティインシデント・生成AIシステム特有のリスクケース事例等 ・生成AIシステムの更新を行った場合の更新内容及びその理由の情報 ・生成AIにて学習するデータの収集ポリシー、学習方法、実施体制等 ・サービス規約やプライバシーポリシー ・生成AIシステムの学習等による出力又はプログラムの変化の可能性 ・開発時に想定していない提供・利用により危害が発生することを避けるためのAI開発者が意図する利用範囲 ・生成AIシステムの動作状況に関する情報、不具合の原因及び対応状況	・ユーザーへの説明書や契約書等、事業者のHPや製品パンフレット等で左記項目が含まれた資料 ・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												乖離評価プロセスの実施状況の記録については、独自に乖離評価のためのチェックリストを作成し、それをもとに確認・記録する	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
						●	●	●	●	●	生成AIシステムの開発・運用における管理について、その管理手法や機能そのものが適切であるか検証を行う	生成AIシステムの開発・運用における管理が適切に機能しているか否かのモニタリングを行う	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
												モニタリングの結果をもとに、継続的な改善を実施する	・組織としてのAIガバナンスの在り方や取組がわかる資料 ・生成AIシステムの開発・運用ルール 等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
	3	AI業界や最新技術等の動向の把握	基本項目	3	生成AIシステムの開発・運用において、品質や説明性を高めるため、AI業界や最新技術等の動向を把握していること	●	●	●	●	●	生成AIシステムの開発・運用に関係する、業界ガイドラインや法令、最新技術を把握する	産業技術総合研究所の「機械学習品質マネジメントガイドライン」等の標準化に向けた取組を参照する	参照している業界ガイドラインや技術のリスト	「調達仕様書標準テンプレート」 8.入札参加に関する事項 8.1.公的な資格や認証等の取得	総務省、経産省「AI事業者ガイドライン 第1.2版」	本項目は主としてシステムの開発・運用を担う事業者の責任範囲となります： ただし、参考情報として以下を申し添えます： - Amazon Bedrockでは、最新のLLMを継続的に提供しており、その詳細はAI Service cardsで確認可能です - 各地域の規制要件やガイドラインに応じたサービス提供体制を整備しています - これらの情報は、事業者が開発・運用時の参考として活用可能です （参考）責任ある AI： https://aws.amazon.com/jp/machine-learning/responsible-ai/
												AI関連の研究会や業界団体等への所属、最新技術の把握、専門家との意見交換等を実施する	加盟しているAI関連の研究会や業界団体リスト		総務省、経産省「AI事業者ガイドライン 第1.2版」	

調達チェックシート						（参考）要求事項の参考										AWSサンプル回答
						対策例の対象AI					対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	
						入力		出力								
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む						
	4	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応	基本項目	4	情報セキュリティインシデント・生成AIシステム特有のリスクケース（事業者の責任の範囲に属するものに限る。）対応体制・手順を整備していること（開発・運用するサービスにおいて、ユーザーからインシデント報告を受け付け、対応の協力をすることを含む。）	●	●	●	●	●	情報セキュリティインシデント・生成AIシステム特有のリスクケース対応体制・手順を整備する	生成AIシステムの開発・運用時に発生する情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応体制や方針及び計画を策定する	・AIガバナンス関連ルール（情報セキュリティインシデント・生成AIシステム特有のリスクケース対応体制・手順が整備されていることが分かる資料）等	「調達仕様書標準テンプレート」 6.作業の実施に当たっての遵守事項 6.1.機密保持、資料の取扱い 『要件定義書標準テンプレート」 3.非機能要件定義 3.10. 情報セキュリティに関する事項	-	実際のシステム運用における具体的なインシデント対応体制・手順の整備は事業者の責任範囲となります： ただし、AWS では以下のリソースを提供しています： 1. インシデント検知・制御 - Amazon Bedrock Guardrailsによる不適切な入出力の制御 - CloudTrailによるAPI呼び出しの記録 - CloudWatchによる異常検知 2. セキュリティインシデント発生時のAWS側の対応 - AWSセキュリティ対応チームによる24時間365日の監視 - AWS セキュリティ速報を通じた脆弱性情報の提供 3. Amazon Bedrock のガードレールを通じた生成AI特有のリスクへの対応 - ハルシネーション対策機能の提供 - プロンプトインジェクション対策の実装 - データの漏洩防止機能の提供 事業者は、これらのAWS提供のリソースを活用しつつ、具体的な運用体制や手順を整備することが可能です。 （参考） ・Amazon Bedrock アプリケーションで責任ある AI のコアディメンションに対応するための考慮事項：https://aws.amazon.com/jp/blog/news/considerations-for-addressing-the-core-dimensions-of-responsible-ai-for-amazon-bedrock-applications/ ・AWS CloudTrailを使用した API コールのログ記録：https://docs.aws.amazon.com/ja_jp/fis/latest/userguide/logging-using-cloudtrail.html ・CloudWatch 異常検出の使用：https://docs.aws.amazon.com/ja_jp/AmazonCloudWatch/latest/monitoring/CloudWatch_Anomaly_Detection.html ・AWS Security Incident Response：https://aws.amazon.com/jp/security-incident-response/ ・セキュリティ速報：https://aws.amazon.com/jp/security/security-bulletins/ ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/
												情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時に備え以下のような体制をあらかじめ整備する ・連絡受付窓口の設置 ・プロジェクトに対し責任ある役職者のアサイン ・対応担当者個々の役割分担 ・対応アプローチ・プロセス ・影響を受けるステークホルダーに対して通知するプロセス等	・AIガバナンス関連ルール（情報セキュリティインシデント・生成AIシステム特有のリスクケース対応体制・手順が整備されていることが分かる資料）等		-	
						●	●	●	●	●	情報セキュリティインシデント・生成AIシステム特有のリスクケース対応に関する教育や訓練を実施する	生成AIシステムの開発・運用時に発生する情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応方針及び計画について、適宜実践的な予行演習を実施する	・AIガバナンス関連ルール（情報セキュリティインシデント・生成AIシステム特有のリスクケース対応体に関する教育や訓練されていることが分かる資料）等		-	
	5	関係者への生成AIに関する教育・リテラシー向上	基本項目	5	生成AIシステムの開発・運用に従事する者又は組織について、生成AIに関するリテラシー向上の取組を実施していること	●	●	●	●	●	生成AI開発・運用に従事する者が適切なリテラシー（生成AI関連の技術知識やリスク、倫理等）を具備できるよう、役割に応じた教育や訓練を実施する	-	・AIガバナンス関連ルール（教育・訓練等、AIリテラシーの向上の取組を実施していることが分かる資料）等	「調達仕様書標準テンプレート」 5.作業の実施体制・方法に関する事項 5.2作業要員に求める資格等の要件	総務省、経産省「AI事業者ガイドライン 第1.2版」	本項目における開発・運用要員への教育・訓練は、主として事業者が実施すべき内容となります： ただし、AWSでは、事業者の教育・訓練をサポートするため、以下のリソースを提供しています。これらのリソースは、事業者が自社の教育・訓練プログラムを構築する際の参考として活用可能です。特に、AWS認定試験は、AIおよび機械学習に関する知識・スキルを客観的に評価する基準として活用できます。 1. AWS認定試験による知識・スキルの体系的な習得 - AWS Certified Machine Learning - Specialty - AWS Certified AI Practitioner 2. AWS公式のトレーニングリソース - Amazon Bedrockのテクニカルドキュメント - AWS Skill Builderでの生成AI関連コンテンツ - AWS Well-Architectedフレームワークにおける生成AI関連のベストプラクティス 3. 責任あるAIの理解促進 - AI Service cardsを通じた技術特性やリスクの解説 - Guardrailsの設定・運用に関するガイダンス （参考） ・AWS 認定：https://aws.amazon.com/jp/certification/ ・Amazon Bedrock Documentation：https://docs.aws.amazon.com/ja_jp/bedrock/ ・AWS Skill Builder で 生成 AI を勉強する 4 ステップ：https://aws.amazon.com/jp/blogs/news/skill-builder-generative-ai-4step/ ・Amazon Bedrock を使用した Well-Architected な生成 AI ソリューションによる運用上の優秀性の実現：https://aws.amazon.com/jp/blogs/news/achieve-operational-excellence-with-well-architected-generative-ai-solutions-using-amazon-bedrock/ ・AI/ML について学ぶ ：https://aws.amazon.com/jp/training/learn-about/machine-learning/ ・AWS の AI サービスのカードを詳しく見る：https://aws.amazon.com/jp/ai/responsible-ai/resources/#service ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/
開発・運用工程要件	6	データの取扱い	基本項目	6	生成AIシステムへの入出力又は処理されるデータの取扱いを適切に管理していること	●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット（プロンプト、学習用の生データ等）の利用目的や利用条件を定める	-	事業者によるインプット（プロンプト、学習用の生データ等）の利用目的を定めることがわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.2.システム方式に関する事項 3.3.システム規模に関する事項 3.10.情報セキュリティに関する事項	経産省「AIの利用・開発に関する契約チェックリスト」	GenUでは、Amazon Bedrockを通じて、以下のデータ管理機能を提供しています： 1. データの入力制御と保護 Amazon Bedrockは、ガードレール機能を通じてコンテンツフィルタリングやプロンプトインジェクション対策などのリスク低減機能を提供しています。GenUではガードレール機能を有効化することにより、不適切な内容や悪意のある入力を検出し、制御することが可能です。 2. 外部データ連携時の制御 Knowledge Baseを使用した外部データとの連携時には、メタデータフィルタリング機能により、適切なデータ管理を実現します。また、Bedrock Agentsを使用した外部データ連携においては、IAMサービスロールによるアクセス制御やAWS LambdaによるプログラムベースのIAM権限制御が可能です。 3. データ品質の評価 Bedrockのモデル評価ツールにより、パフォーマンスやバイアスの評価が可能です。これにより、データ品質の継続的な監視と改善を支援します。 なお、以下の事項については、事業者側での対応が必要となります： - 具体的なデータ品質基準の設定 - データクレンジングプロセスの確立 - プライバシー影響評価の実施 - データガバナンスポリシーの策定と実施 （参考） ・ガードレールのコンポーネント：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails-components.html ・Knowledge Bases for Amazon Bedrock がメタデータフィルタリングをサポートし検索精度向上：https://aws.amazon.com/jp/blogs/news/knowledge-bases-for-amazon-bedrock-now-supports-metadata-filtering-to-improve-retrieval-accuracy/ ・Amazon Bedrock のためのアイデンティティとアクセス管理：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/security-iam.html ・Amazon Bedrock エージェントがユーザーから引き出す情報を送信するように Lambda 関数を設定する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/agents-lambda.html ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット（プロンプト、学習用の生データ等）の管理・セキュリティ体制を定めており、問題が生じないよう対策を講じる	-	インプット（プロンプト、学習用の生データ等）の管理・セキュリティ体制方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット（プロンプト、学習用の生データ等）の保持期間及び消去を管理しており、問題が生じないよう対策を講じる	-	インプット（プロンプト、学習用の生データ等）の保持期間及び消去の管理方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット（プロンプト、学習用の生データ等）の外部提供管理として、生成AIシステムのユーザーに対して提供する義務を定めた上での管理をする	-	インプット（プロンプト、学習用の生データ等）をユーザーに対して提供する義務をを定めていることがわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット（プロンプト、学習用の生データ等）の外部提供管理として、第三者提供に関しての管理をする	-	インプット（プロンプト、学習用の生データ等）の第三者提供に関する管理方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット処理結果（学習用データ、中間生成物、派生的知的財産等）の使用・利用に関しての管理をする	-	インプット処理結果（学習用データ、中間生成物、派生的知的財産等）の使用・利用に関する管理方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲でインプット処理結果（学習用データ、中間生成物、派生的知的財産等）の外部提供に関しての管理をする	-	インプット処理結果（学習用データ、中間生成物、派生的知的財産等）の外部提供に関する管理方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
						●	●	●	●	●	プロジェクトの目的に照らし、必要な範囲で学習前・学習全体を通じて、データのアクセスを管理するデータ管理・制限機能の導入検討を行う等、適切な保護措置を実施する	-	学習前・学習全体を通じて、データの保護措置の方針がわかる資料等		経産省「AIの利用・開発に関する契約チェックリスト」	
	7	アウトプットの品質保証	基本項目	7	生成AIシステムの期待品質を満たすための取組を行っていること	●	●	●	●	●	生成AIをチャットで用いる場合、チャットでの利用の期待品質を定め、その基準を満たしているかの測定及び評価をする	-	ベンチマーク※等を使用した評価基準とその評価内容がわかる資料等 ※一定の信頼性のある評価基準に基づく独自の評価を否定するものではない	「要件定義書標準テンプレート」 3.非機能要件定義 3.12.テストに関する事項 3.17.保守に関する事項	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）」	本項目における具体的な品質評価の実施は、主として事業者が実施すべき内容となりますが、AWSは、以下の品質評価機能を提供しています： 1. モデル評価基盤 Amazon Bedrockのモデル評価機能では、事前に定義された評価基準に基づいてモデルの性能を評価できます。例えば、回答の正確性、有害なコンテンツの生成有無、プロンプトインジェクションへの耐性などを評価することが可能です。また、RAG評価機能を使用することで、検索結果の関連性、生成された回答の品質、および回答と検索結果の一貫性を自動的に評価できます。さらに、ユーザーが独自の評価データセットを使用して、特定のユースケースに対するモデルの適合性を評価することもできます。 なお、以下の事項については、事業者が実施する必要があります： - 具体的な品質要件の設定 - ユースケースに応じたテストケースの設計と評価の実施 - 期待品質の具体的な評価基準の設定 - 定期的な品質評価の実施 （参考） ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/ ・Amazon Bedrock で RAG 評価のサポートを開始 (一般提供)：https://aws.amazon.com/jp/about-aws/whats-new/2025/03/amazon-bedrock-rag-evaluation-generally-available/

調達チェックシート						（参考）要求事項の参考								AWSサンプル回答
						対策例の対象AI							（参考）「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考）対策例・対策例詳細の参考文献	
						入力		出力							
分類	評価観点 #	評価観点	評価・選定時の項目の分類	要求事項 #	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例			
						●	●	●	●	●	生成AIシステムの期待品質を満たしているか、テストケースを用いて評価する。その際テストケースは複数提案する（バッチ処理、情報検索、自然言語からのプログラム作成等）	プロンプトを入力した生成AIシステムの出力を検査することで、生成AIシステムの明示された要求仕様が規定する機能を生成AIモデルが実現可能かを調べている。プロンプトと生成AIモデルの組み合わせを対象とした試行錯誤的な繰り返し（一種のプロトタイピング開発スタイル）を通して、機能要求満足性を評価する	バッチ処理で生成AIを用いる場合、テストケース例とテスト方針がわかる資料等		対策例：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）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第１版」	
						●	●	●	●	●	マーケティング等で実施されるアンケートによる定性的な満足度評価手法を生成AIシステムの機能要求満足性の評価に用いて、機能要求満足性を満たしているか検証する	-	マーケティング等で実施されるアンケートによる定性的な満足度評価手法を生成AIシステムの機能要求満足性の評価に用いて、機能要求満足性を満たしているかの検証方針がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
						●	●	●	●	●	生成AIシステム出力のサンプリング手法やそのパラメータを指定できる場合、システムプロンプトの設計のための試行錯誤の中で合わせて検討する	-	生成AIシステム出力のサンプリング手法やそのパラメータを指定できる場合、システムプロンプトの設計と合わせて指定する開発方針がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
						●	●	●	●	●	生成AIシステムの機能要求水準を満たすのに十分な検索精度を確保するよう対策を講じる（多くの場合、想定される問題領域のなるべく多くの範囲で、一定以上の精度が出ることがよいとされる） ※RAGを利用した生成AIシステムのみ本項目の対象	-	生成AIシステムの機能要求水準を満たすのに十分な検索精度を確保するための対策方針がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	

調達チェックシート						（参考）要求事項の参考										
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	対策例の対象AI 入力 テキストを含む	対策例の対象AI 入力 音声を含む	対策例の対象AI 出力 テキストを含む	対策例の対象AI 出力 画像を含む	対策例の対象AI 出力 音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考）「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考）対策例・対策例詳細の参考文献	AWSサンプル回答
	8	ベンダーロックインの回避	基本項目	8	利用している生成AIモデルはバージョン情報を含めて明示可能であること	●	●	●	●	●	利用可能な生成AIモデル及びそのバージョン情報等の一覧を整理して公開する	-	利用している生成AIモデルのバージョン情報の開示方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.8.中立性に関する事項 3.11.情報システム稼働環境に関する事項	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）」	AWSでは、Amazon Bedrockを通じて提供する生成AIモデルについて、以下のように情報提供を行っています。事業者は、これらの情報を参照することで、利用する生成AIモデルの詳細を確認し、システム開発・運用に活用することが可能です： 1. 利用可能なモデルの明示 AWSでは、AWS独自のNovaモデルや、AnthropicのClaudeなどのサードパーティモデルについて、その概要や特徴を提供しています。詳細な技術仕様やバージョン情報については、AWS公式ドキュメントおよび各モデルプロバイダーが提供する情報を併せて参照することが可能です。 2. モデル情報の開示方法 モデルに関する情報は、AWS AI Service Cardsでの包括的な情報公開、AWS公式ドキュメントでの仕様説明、そしてAWS Management Consoleでの利用可能モデルの表示を通じて提供しています。 （参考） ・Amazon Bedrock：https://aws.amazon.com/jp/bedrock/ ・AI の透明性を促進するリソース:https://aws.amazon.com/jp/ai/responsible-ai/resources/
			任意追加項目（加点項目）	9	生成AIシステムに入力されるプロンプトの一部やパラメータが隠蔽されていないことの確認のために、合理的な範囲での企画者への情報開示や情報提供ができる状態であること	●	●	●	●	●	合理的な範囲での開発・運用時におけるシステムプロンプト等の企画者への情報開示や情報提供が技術的に可能であることを表明し保証する	-	生成AIシステムに入力されるプロンプトの一部やパラメータが隠蔽されていないことを表明した資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.8.中立性に関する事項 3.10.情報セキュリティに関する事項	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）」	GenUでは、Amazon Bedrockを通じて、プロンプトやパラメータの透明性確保のため、以下の機能を提供しています： 1. Amazon Bedrock Prompt Managementによる透明性確保 Prompt Managementでは、プロンプトの中央管理を実現し、バージョン管理機能や承認ワークフローを提供しています。これにより、企画者は開発・運用時におけるプロンプトの内容や変更履歴を確認することができます。また、プロンプトのテストと履歴管理機能により、プロンプトの動作や効果を検証することが可能です。さらに、プロンプトの共有機能により、企画者との情報共有を効率的に行うことができます。 ただし、以下の点については、事業者側での対応が必要となります： - 業務要件に応じたプロンプトの設計内容 - カスタマイズしたシステムプロンプトの内容 - 運用時の具体的なパラメータ設定値の管理 （参考） ・Amazon Bedrock Prompt Management：https://aws.amazon.com/jp/bedrock/prompt-management/ ・Amazon Bedrock でプロンプト管理を使用して再利用可能なプロンプトを構築して保存するhttps://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/prompt-management.html
			任意追加項目（加点項目）	10	過去のチャット履歴の保存機能や、プロンプトをテンプレートとして登録するとともに、それらのデータをエキスポートする機能を提供する技術を有していること	●	●	●	●	●	過去のチャット履歴の保存機能、プロンプトのテンプレート化機能、それらのエキスポート機能の開発が可能かを表明し保証する	-	過去のチャット履歴の保存機能や、プロンプトをテンプレートとして登録するとともに、それらのデータをエキスポート可能な機能の実装方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項 3.8.中立性に関する事項 3.13.移行に関する事項	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）」	GenUでは、Amazon Bedrockを通じて、チャット履歴の保存とプロンプト管理に関する以下の機能を提供しています： 1. チャット履歴の管理とプロンプトのテンプレート化 Amazon Bedrock Prompt Managementにより、チャット履歴の保存と管理が可能です。また、効果的なプロンプトをテンプレートとして保存し、再利用することができます。これらのデータは必要に応じてエクスポートすることが可能です。 なお、具体的なチャット履歴の保存期間や保存方法、テンプレートの作成基準等については、事業者が業務要件に応じて設定する必要があります。 （参考） ・プロンプト管理を使用してプロンプトを作成する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/prompt-management-create.html
			任意追加項目（加点項目）	11	特定のモデルの利用が主たる目的ではない場合、生成AIモデルごとに異なる機能や動作に影響する特徴があることを考慮し、複数の生成AIモデルの中から最適な生成AIモデルを選択又は組み合わせて利用する技術を有していること	●	●	●	●	●	①複数の生成AIモデルが利用できること、②生成AIモデルの選択ができること（組み合わせて利用する場合も含む）等の技術を有していることを表明し保証する	-	複数の生成AIモデルの中から最適な生成AIモデルを選択又は組み合わせて利用する技術を有している実績が分かる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.8.中立性に関する事項 3.11.情報システム稼働環境に関する事項	-	GenUでは、Amazon Bedrockを通じて、以下のようなマルチモデル対応機能を提供しています： 1. 複数の生成AIモデルの提供 Amazon BedrockではAmazon Nova、Anthropic Claude、AI21 Labs Jurassic、Cohere Commandなど、複数のファウンデーションモデルを利用することができます。 2. モデル選択と組み合わせ利用 APIレベルでのモデル指定機能により、ユースケースに応じて適切なモデルを選択することが可能です。また、リクエストごとに異なるモデルを利用することができ、Bedrock Agentsを使用することで複数のモデルを組み合わせて利用することも可能です。 なお、具体的なモデルの選択や組み合わせについては、事業者が業務要件に応じて設計する必要があります。 （参考） ・Amazon Bedrock：https://aws.amazon.com/jp/bedrock/ ・Amazon Bedrock エージェントとのマルチエージェントコラボレーションを使用する ：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/agents-multi-agent-collaboration.html
			基本項目	12	生成AIシステムのアーキテクチャ設計と実装において、ベンダーやシステムの移行の容易性も踏まえた開発や運用が可能であること	●	●	●	●	●	標準化されたインタフェースやプロトコルを活用し、AIコンポーネントを疎結合に設計する	-	標準化されたインタフェースやプロトコルを活用し、AIコンポーネントを疎結合に設計する方法の実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.8.中立性に関する事項 3.11.情報システム稼働環境に関する事項	AISI「Chief AI Officer 設置・AI ガバナンス 実務マニュアル」	GenUは、Amazon Bedrockを通じて、ベンダーやシステムの移行の容易性を踏まえた疎結合なアーキテクチャを実現するため、以下の機能・構成を提供しています： 1. 標準化されたインタフェースによる疎結合化 Amazon Bedrockの統一API（Converse API等）により、モデルに依存しない共通インタフェースでアプリケーションを実装でき、AIコンポーネントを疎結合に設計することができます。 2. モデル切替を想定した抽象化（出口戦略） APIレベルでモデルを指定する構成のため、利用する生成AIモデルやベンダーの切り替えを想定したアダプタ的な抽象化レイヤーを組み込みやすく、「出口戦略」をアーキテクチャレベルで担保することができます。 3. モジュール性の高い構成 GenUはAWS CDK（Infrastructure as Code）で構成され、各コンポーネント（チャット、RAG、文書生成等）がモジュール化されているため、再利用部品である生成AIモデルの置き換えや、コンポーネント単位での追加・削除・改修が容易です。 なお、以下の事項については、事業者側での対応が必要となります： - 抽象化レイヤー（アダプタ）の具体的な設計と、ベンダー変更を想定した「出口戦略」の組み込み - コンポーネントやシステムプロンプトのモジュール性を高める設計、入出力ポートの定義の明示 - 移行時のデータ・設定の移送方針とその検証 （参考） ・Amazon Bedrock：https://aws.amazon.com/jp/bedrock/ ・Amazon Bedrock ドキュメント：https://docs.aws.amazon.com/ja_jp/bedrock/ ・Amazon Bedrock エージェントとのマルチエージェントコラボレーションを使用する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/agents-multi-agent-collaboration.html
						●	●	●	●	●	モデルやベンダーの切り替えを想定した抽象化レイヤー（アダプタ）を用意し、「出口戦略」をアーキテクチャレベルで組み込む	-	抽象化レイヤー（アダプタ）を用意し、ベンダー変更を想定したアーキテクチャを実現する方法の実現方法がわかる資料等		AISI「Chief AI Officer 設置・AI ガバナンス 実務マニュアル」	
						●	●	●	●	●	コンポーネントアーキテクチャを、再利用部品である生成AIモデルの置き換えが容易になるようにデザインする	-	コンポーネントアーキテクチャを、再利用部品である生成AIモデルの置き換えが容易になるようにデザインする方法がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
						●	●	●	●	●	保守性の向上のため、生成AIシステムのコンポーネントをモジュール性を考慮して開発、運用する	生成AIシステムのコンポーネントに明確な役割・責任（Responsibility）を与え、コンポーネントのモジュール性を高める。また、コンポーネントのデザインでは、入出力ポートの定義を明示する	生成AIシステムコンポーネントのモジュール性を高めた開発の方針がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
												生成AIシステムを構成するアーキテクチャコンポーネントを明示することで、モジュール単位での追加・削除・置き換えに関わる修正性を向上させる	生成AIシステムを構成するアーキテクチャコンポーネントを明示する方法がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
												システムプロンプトのモジュール性を高める。プロンプトの流れを段落等の観点から構造化し、各々の段落の役割・責任（Responsibility）を明確化する	システムプロンプトのモジュール性を高め、システムプロンプトの保守性を向上させる方法がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
	9	生成AIモデルのアップデートの考慮	基本項目	13	メジャーアップデート又は移行に関する生成AI固有の観点からリスク軽減していること	●	●	●	●	●	生成AIモデルのメジャーアップデートや移行時に、生成AIシステムのアウトプットの品質や安全性等について問題ない性能かを検証する	-	生成AIモデルのアップデート時に、生成AIシステムのアウトプットの品質や安全性等が問題ないかの検証方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.16.運用に関する事項 3.17.保守に関する事項	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）」	GenUは、Amazon Bedrockを通じて、生成AIモデルのアップデートや移行に関するリスク評価のため、以下の機能を提供しています： 1. アップデート時の検証機能 Amazon Bedrock Model Evaluationを使用することで、モデルのアップデートや移行前に、アウトプットの品質評価、安全性の評価、新旧バージョン間の比較テストを実施することが可能です。これにより、モデルの更新による潜在的なリスクを事前に特定し、評価することができます。 なお、具体的な業務要件に基づく検証や移行計画の策定・実施については、事業者側での対応が必要となります。 （参考） ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/

調達チェックシート						（参考）要求事項の参考										
						対策例の対象AI										
						入力		出力								
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考）「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考）対策例・対策例詳細の参考文献	AWSサンプル回答
	10	文化的・言語的考慮	任意追加項目（加点項目）	14	生成AIシステムが日本の言語環境や文化環境に即した出力が可能であること	●	●	●	●	●	日本の言語環境や文化環境を踏まえたアウトプットを出力可能な生成AIシステムを提供する	-	日本の言語環境や文化環境に応じた出力結果となっているかの確認方針がわかる資料、開発時に学習データ等で日本の言語環境や文化環境に即したものとなるよう、取組がなされていることがわかる資料、日本語能力を測るベンチマークを活用した評価結果がわかる資料等 ※広く認知された一定の信頼性のある評価基準の使用を否定するものではない	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	-	GenUでは、Amazon Bedrockを通じて、以下の機能を提供しています： 1. 日本語対応モデル Amazon NovaやAnthropic Claude等、多言語対応モデルを提供しており、日本語での入出力が可能です。 2. 日本語出力の評価 Amazon Bedrock Model Evaluationを使用することで、日本語での出力品質を評価することが可能です。これにより、日本の言語環境に適したアウトプットが得られているかを確認することができます。 なお、特定の業界や組織固有の表現への対応、具体的な日本語評価基準の設定については、事業者側での対応が必要となります。 （参考） ・Amazon Bedrock：https://aws.amazon.com/jp/bedrock/ ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/
	11	環境への配慮	任意追加項目（加点項目）	15	環境に配慮した生成AIシステムを開発・提供すること	●	●	●	●	●	電力効率が高い半導体を使用したハードを基に学習されたモデルを利用する等の環境への配慮をした生成AIシステムを提供する又は利用する	システムテストに相当するソフトウェアテスティングで、システムが所与の機能仕様を適切に実現していることを検査し時間効率性及び資源効率性を評価する	消費電力等環境性能が優れていることがわかる資料等	「調達仕様書標準テンプレート」 4.作業の実施内容に関する事項 4.17.その他	対策例詳細：産総研「生成AI品質マネジメントガイドライン 第１版」	GenUは、AWSのクラウドインフラストラクチャ上で運用されています。AWSでは環境への配慮に関して、以下の取り組みを行っています： 1. エネルギー効率の高いデータセンター AWSのデータセンターは、一般的なオンプレミスの最大 4.1 倍の効率を実現し、ワークロードの二酸化炭素排出量を最大 99 % 削減できると推定されています。 2. 再生可能エネルギーの利用 2023 年に、業務全体で消費される電力の 100% を再生可能エネルギーで賄うという目標を達成しています。 3. 深層学習 (DL) および生成 AI 推論アプリケーション向けに自社設計された 第2世代 AWS Inferentia チップを採用した Inf2 インスタンス、他の同等の Amazon EC2 インスタンスに比べて、ワットあたりのパフォーマンスが最大 50% 向上します。Inferentia2 チップは、高度なシリコンプロセスとハードウェアとソフトウェアの最適化を使用して、DL モデルを大規模に実行する際に高いエネルギー効率を実現します。 4. 16 個の AWS Trainium2 チップを搭載した Amazon EC2 Trn2 インスタンスは、生成 AI 専用に構築されており、数千億から数兆以上のパラメータを持つモデルのトレーニングとデプロイのためにハイパフォーマンス EC2 インスタンスを提供します。Trn2 インスタンスには Trn1 インスタンスよりも 3 倍高いエネルギー効率があります。 5. Bedrockを利用するにあたって、ファインチューニングをすることで用途に応じてより小型のモデルや効率的なモデルを選択・運用しやすくなるため、推論時の計算量や消費電力の削減につながります。また、蒸留モデルの利用により、パラメータ数の少ない小型モデルでも高い性能を実現でき、計算リソースと消費電力を削減できます。 （参考） ・エネルギー転換：https://aws.amazon.com/jp/energy-utilities/sustainability/ ・Inf2インスタンス：https://aws.amazon.com/jp/ec2/instance-types/inf2/ ・Trn2インスタンス：https://aws.amazon.com/jp/ec2/instance-types/trn2/ ・Amazon Bedrock でファインチューニングまたは継続的な事前トレーニングを使用してモデルをカスタマイズする： https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/custom-model-fine-tuning.html ・Amazon Bedrock Model Distillation：https://aws.amazon.com/jp/bedrock/model-distillation/
生成AIシステムの基本機能要件	12	有害情報の出力制御	基本項目	16	テロや犯罪に関する情報や攻撃的な表現等、有害情報の出力制限等をしていること	●	●	●	●	●	有害情報を入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムの出力に当該情報が含まれない、若しくは出力を拒否できるような仕組みを具備する	サイバー攻撃やテロ等の犯罪、CBRN（Chemical, Biological,Radiological,Nuclear）に利用され得る情報を入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムの出力に有害情報が含まれない、若しくは出力を拒否することができるような仕組みを具備する	サイバー攻撃やテロ等の犯罪、CBRNに利用され得る情報を入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムの出力に有害情報が含まれない、若しくは出力を拒否することができるような仕組みの実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenU では、Amazon Bedrock Guardrails により、有害情報・差別的表現・攻撃的表現等の有害コンテンツをフィルタリングできます。サイバー攻撃・テロ・CBRN に利用され得る情報を含む入力に対しても、出力への混入防止や出力拒否が可能です。フィルタリングルールはカスタマイズ可能で、リアルタイムでの出力検証・制御を行えます： なお、具体的なフィルタリングルールの設定や業務特性に応じた追加制御は事業者側での対応が必要です。 （参考） ・Amazon Bedrock ガードレールを使用して有害なコンテンツを検出してフィルタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails.html ・プロンプトインジェクションのセキュリティ：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/prompt-injection.html
												差別表現等ユーザーが精神的な被害を受け得る情報を入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムの出力に有害情報が含まれない、若しくは出力を拒否することができるような仕組みを具備する	差別表現等ユーザーが精神的な被害を受け得る情報を入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムの出力に有害情報が含まれない、若しくは出力を拒否することができるような仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	生成AIシステムの出力の有害性スコア（攻撃的であるかどうか等の有害さを数値で表したもの）の測定及び評価をする	-	生成AIシステムの出力の有害性スコア及びスコアの測定方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	-	●	●	●	テキストの入力に、虚偽の情報や意味不明な内容、複雑な内容が含まれる場合、生成AIシステムの挙動が不安定になることがある。そのような入力に対しても、有害な出力を抑制できるような仕組みを具備する	-	画像やテキストの入力に、虚偽の情報や意味不明な内容、複雑な内容が含まれる場合、生成AIシステムの挙動が不安定になることがある。そのような入力に対しても、有害な出力を抑制できるような仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
	13	偽誤情報の出力・誘導の防止	基本項目	17	偽誤情報の出力の防止措置を取っていること	●	●	●	●	●	調達対象の生成AIシステムにおいて利用予定の生成AIモデルと異なる生成AIモデルに対し、事実を問う内容の同一のテストデータを入力した際、生成AIシステムの出力と意味的に同様の内容が出力されるか検証を行う	-	調達対象の生成AIシステムにおいて利用予定の生成AIモデルと異なる生成AIモデルに対し、事実を問う内容の同一のテストデータを入力した結果が記載された資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、以下の偽誤情報出力防止機能を提供しています： 1. Bedrock Guardrailsによる制御機能 事実を問う内容に対する出力の一貫性を確認することができます。また、プレビュー版の機能として、数理的に厳密な自動推論チェックにより、生成AIモデルの出力におけるハルシネーションの事実誤認を防止することも可能です。 2. Model Evaluationによる検証 Amazon Bedrock Model Evaluationを使用することで、異なるモデル間での回答の整合性チェックや、Ground Truth（正解データ）との照合により、出力内容の正確性を評価することが可能です。 3. RAG（Retrieval-Augmented Generation）機能 Amazon Bedrock Knowledge Baseを利用することで、外部データソースとの連携による事実確認が可能です。検索結果と生成内容の関連性スコアの提供、一貫性チェック、さらに正解データと検索結果の意味的な一貫性の評価を行うことができます。 なお、以下の事項については、事業者側での対応が必要となります： - 業務特性に応じた信頼できる情報源の選定 - RAGで使用する正解データの準備と検証 - 具体的な評価基準の設定 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・数学的に正しい自動推論チェックにより、生成AIモデル のハルシネーションによる事実ミスを防ぐ (プレビュー)：https://aws.amazon.com/jp/blogs/news/prevent-factual-errors-from-llm-hallucinations-with-mathematically-sound-automated-reasoning-checks-preview/ ・Amazon Bedrock の新しい RAG 評価機能と LLM-as-a-Judge 機能：https://aws.amazon.com/jp/blogs/news/new-rag-evaluation-and-llm-as-a-judge-capabilities-in-amazon-bedrock/
						●	●	●	●	●	偽誤情報の出力防止に関する評価項目のスコアを測定した結果、スコアに問題がないことを確認する	生成AIシステムに対するプロンプトと出力データの関連性に係るスコアの測定及び評価をする	生成AIシステムに対するプロンプトと出力データの関連性に係るスコア及びスコアの測定方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												生成AIシステムの出力データとRAGによる検索結果の関連性に係るスコアの測定及び評価をする ※RAGを利用した生成AIシステムのみ本項目の対象	生成AIシステムの出力データとRAGによる検索結果の関連性に係るスコア及びスコアの測定方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												生成AIシステムの出力データとRAGによる検索結果の一貫性に係るスコアの測定及び評価をする ※RAGを利用した生成AIシステムのみ本項目の対象	生成AIシステムの出力データとRAGによる検索結果の一貫性に係るスコア及びスコアの測定方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												正解データとRAGによる検索結果の意味的な一貫性に係るスコアの測定及び評価をする ※RAGを利用した生成AIシステムのみ本項目の対象	正解データとRAGによる検索結果の意味的な一貫性に係るスコア及びスコアの測定方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												出力にコンテンツ認証がなされる生成AIシステムにおいて、様々なプロンプトを入力した場合でも適切にコンテンツ認証がなされる仕組みを具備する	生成AIシステムへ入力した様々なプロンプトに対し、適切なコンテンツ認証がされることができるような仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												RAGを利用した生成AIシステムで検索対象となるドキュメントや、評価のために評価者が用意する正解データが、事実に基づいたデータであるかの確認をする ※RAGを利用した生成AIシステムのみ本項目の対象	RAGを利用した生成AIシステムで検索対象となるドキュメントや、評価のために評価者が用意する正解データが、事実に基づいたデータであることを確認した記録等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	

調達チェックシート						（参考）要求事項の参考											AWSサンプル回答
						対策例の対象AI					対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	
						入力		出力									
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む							
			不特定外部者（一般国民等）による府省庁外利用の場合は基本項目として適用	18	ユーザーに対し、生成AIシステムの出力を人間から発せられた情報と区別できるようにすること	●	●	●	●	●	ユーザーに対し、生成AIシステムの出力を人間から発せられた情報と区別できるような仕組みを具備する	-	ユーザーに対し、生成AIシステムの出力を人間から発せられた情報と区別できるような仕組みの実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、以下の機能を通じて生成AIシステムの出力を人間から発せられた情報と区別可能にしています： 1. 識別表示の実装 AWS LambdaとAPI Gatewayを活用することで、出力の冒頭または末尾に「This is an AI-generated response」などの文言を自動的に追加したり、カスタマイズ可能な免責事項や注意書きを挿入したりすることができます。 なお、以下の事項については、事業者側での対応が必要となります： - ユーザーインターフェースでのAI生成表示の具体的なデザイン - 組織のポリシーに基づいたAI生成コンテンツの説明文や免責事項の作成 （参考） ・API Gateway で REST API のステージ変数をセットアップする：https://docs.aws.amazon.com/ja_jp/apigateway/latest/developerguide/how-to-set-stage-variables-aws-console.html
			不特定外部者（一般国民等）による府省庁外利用の場合は基本項目として適用	19	生成AIシステムによる、ユーザーの意思決定の不当な誘導を防止するよう制御等をしていること	●	●	●	●	●	生成AIシステムの推奨や指示により、ユーザーが主観的又は客観的に不利益となる行動に誘導されることがないよう対策を講じる	-	生成AIシステムの推奨や指示により、ユーザーが主観的又は客観的に不利益となる行動に誘導される出力結果になっていないかの確認方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、エンドユーザーへの不当な誘導を防止するための以下の機能を提供しています： 1. Bedrock Guardrailsによる制御 Guardrailsを使用することで、誘導的な表現の検出と制御、過度な推奨や指示の制限、商業的な誘導の制御を行うことができます。また、URLやリンクの出力についても制御することが可能です。 2. Lambda/API Gatewayによる実装 AWS LambdaとAPI Gatewayを活用することで、生成AI出力であることの明示や、推奨・提案を含む場合の警告表示、オプトイン/アウト機能の実装が可能です。
						●	●	●	●	●	生成AIシステムの出力において、生成AIシステムによる誘導（URLリンク等を含むメッセージ生成等）のオプトイン又はオプトアウト手段が出力に含まれる仕組みを具備する	-	生成AIシステムの出力において、生成AIシステムによる誘導のオプトイン又はオプトアウト手段が出力に含まれる仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	なお、以下の事項については、事業者側での対応が必要となります： - 具体的な誘導防止ルールの設定 - ユーザーインターフェースでのオプトイン/アウト機能の実装 - 業務特性に応じた推奨・提案の制限ルールの設定 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・API Gateway で REST API のステージ変数をセットアップする：https://docs.aws.amazon.com/ja_jp/apigateway/latest/developerguide/how-to-set-stage-variables-aws-console.html

調達チェックシート						（参考）要求事項の参考																AWSサンプル回答
						対策例の対象AI																
						入力		出力														
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考）「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考）対策例・対策例詳細の参考文献	
	14	公平性と包摂性	不特定外部者（一般国民等）による府省庁外利用の場合は基本項目として適用	20	有害なバイアスや、不当な差別の出力制限等をしていること	●	●	●	●	●	特定のグループ（人種、性別、国籍、年齢、政治的信念、宗教等の多様な背景に関する）に対する有害なバイアスの入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムが回答を拒否できる仕組みを具備する	出力コンテンツに対するフィルター機能を導入し、安全性を損なう情報、想定外のパーソナルデータ漏洩、著作権者の権利侵害に相当する出力、あるいは想定外のバイアスの出力を防ぐようにデザインする	特定のグループ（人種、性別、国籍、年齢、政治的信念、宗教等の多様な背景に関する）に対する有害なバイアスの入力あるいは想定出力に含むテストデータを入力した際、生成AIシステムが回答を拒否することができるような仕組みの実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	対策例：AISI「AIセーフティに関する評価観点ガイド（第1.10版）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第1版」	GenUでは、以下の機能を通じて、有害なバイアスや不当な差別を防止することが可能です： 1. Bedrock Guardrailsによるバイアスⅰⅰ制御 Guardrailsを使用することで、人種、性別、国籍、年齢、政治的信念、宗教等の属性に基づく差別的表現の検出と制御が可能です。また、ヘイトスピーチの検出や、ステレオタイプ的な表現の制御を行うことができ、これらのポリシー設定は組織固有の配慮事項や地域特性に応じてカスタマイズすることが可能です。 2. Model Evaluationによる評価機能 Amazon Bedrock Model Evaluationにより、特定グループに対する偏りの評価や出力の公平性評価、属性による影響度の測定が可能です。また、定期的なバイアスチェックと評価結果の可視化を行うⅰことができます。 3. 追加的な制御とモニタリング AWS Lambdaを活用することで、AIモデルへの入力前後でフィルタリング処理や事前定義された禁止語句、差別的パターンを検出するカスタムロジックを実装できます。また、CloudWatchを用いて生成AIシステムの出力をリアルタイムでモニタリングし、潜在的に有害な推奨を特定、アラート機能による迅速な対応が可能です。 なお、以下の事項については、事業者側での対応が必要となります： - 業務特性に応じたバイアス評価基準の設定 - 具体的な差別防止ルールの実装 - 定期的なバイアス評価の実施 - 複数のモデルを使用する場合の比較評価と選択 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/ ・Amazon Bedrock エージェントがユーザーから引き出す情報を送信するように Lambda 関数を設定する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/agents-lambda.html ・Amazon BedrockとAmazon CloudWatchの統合による生成系AIアプリケーションのモニタリング：https://aws.amazon.com/jp/blogs/news/amazon-bedrock_and_amazon-cloudwatch_integration_for_genai/
						●	●	●	●	●	出力が人の属性に影響されないと想定されるテストデータを入力した際、出力結果が属性による影響を受けないことの対策を講じる	出力コンテンツに対するフィルター機能を導入し、安全性を損なう情報、想定外のパーソナルデータ漏洩、著作権者の権利侵害に相当する出力、あるいは想定外のバイアスの出力を防ぐようにデザインする	出力が人の属性に影響されないと想定されるテストデータを入力した際、出力結果が属性による影響を受けないことを確認しているテスト方針がわかる資料等		対策例：AISI「AIセーフティに関する評価観点ガイド（第1.10版）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第1版」	
						●	●	●	●	●	特定のモデルの利用が主たる目的ではない場合、生成AIモデルごとに異なるイデオロギーやバイアスを持つことがあることを考慮し、複数の生成AIモデルを利用する	-	複数の生成AIモデルを利用するシステム設計であることがわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	学習データ、モデルの学習過程において、生成AIシステムに不公正なバイアスがないかを検証する また、生成AIモデルが、法令等に基づき特定の分野における出力制限をかけられている可能性について、整理して提供する	利用目的に応じ、バイアスや出力制限が支障となるかを踏まえ、適切な生成AIを選択する	生成AIシステムの検証プロセス・検証方法が分かる資料、生成AIモデルの提供企業の所在国におけるAIに対する規律や当該生成AIモデルの生成結果のバイアスに対する評価結果等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	出力結果について有害なバイアスが含まれていないか定期的にモニタリングする	-	有害なバイアスが含まれていないか定期的にモニタリングする際の確認観点がわかる資料等		総務省、経産省「AI事業者ガイドライン 第1.2版」	
						●	●	-	●	-	特定のカテゴリ（例えば職業）に関わる人物画像の生成を求めた場合に、個人の属性（例えば性別、年齢、人種）に関する偏りのない出力を生成できる仕組みを具備する	-	特定のカテゴリ（例えば職業）に関わる人物画像の生成を求めた場合に、個人の属性（例えば性別、年齢、人種）に関する偏りのない出力を生成することができるような仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
			基本項目	21	出力が全てのユーザーにとって理解しやすいものとなるよう対策を講じていること	●	●	●	●	●	生成AIシステムの出力の流暢性スコア（出力が文法的に適切かを数値で表したもの）の測定及び評価をする	-	生成AIシステムの出力の流暢性スコア及びスコアの測定方針がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、出力の可読性を確保するための以下の機能を提供しています： 1. Bedrock Guardrailsによる品質制御 Guardrailsを使用することで、文法的な適切性の確保や文章構造の一貫性確保、専門用語や難解な表現の適切な使用制御が可能です。また、文法的に不適切な入力に対しても、適切な処理を行うことができます。 2. Model Evaluationによる評価 Amazon Bedrock Model Evaluationにより、精度、毒性、頑健性の観点からモデルの評価が可能です。ただし、流暢性スコアや読みやすさ、文法的な適切性に関する直接的な評価機能は含まれていません。 3. 出力パラメータの制御 Temperature（創造性）やTop P（出力の多様性）、最大トークン数の設定、Response Formatによる出力形式の指定など、出力の品質に影響を与えるパラメータを調整することができます。これにより、読みやすく一貫性のある応答の生成が可能です。 なお、以下の事項については、事業者側での対応が必要となります： - 具体的な可読性基準の設定 - 業務要件に基づく品質評価基準の設定 - 定期的な出力品質の評価実施 - 流暢性や読みやすさを評価するための追加的な仕組みの構築 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/ ・Amazon Bedrock のよくある質問（Amazon Bedrock でのモデル評価とは何ですか?）https://aws.amazon.com/jp/bedrock/faqs/ ・推論パラメータとレスポンスの生成に影響を与える：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/inference-parameters.html
						●	●	●	●	●	生成AIシステムの出力の読み易さに関するスコアの測定及び評価をする	-	生成AIシステムの出力の読み易さに関するスコア及びスコアの測定方針がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	生成AIシステムに文法的に不適切なテストデータを入力した場合でも、出力は文法的に適切であり、人間が理解しやすいものとなるよう対策を講じる	-	生成AIシステムに文法的に不適切なテストデータを入力した場合でも、出力は文法的に適切であり、人間が理解しやすいものとなっているかの確認方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
	15	目的外利用への対処	基本項目	22	目的外利用の防止を行い、仮に目的外利用された場合にも危害・不利益を発生しにくくなるよう入出力制限等をしていること	●	●	●	●	●	生成AIシステムに想定ユースケース以外の出力を想定したテストデータが入力された場合でも、ユーザーの生命・身体・財産等や各種の権利に危害を生じさせる出力がされない、あるいは、出力を拒否できる仕組みを具備する	-	生成AIシステムに想定ユースケース以外の出力を想定したテストデータが入力された場合でも、ユーザーの生命・身体・財産等や各種の権利に危害を生じさせる出力がされない、あるいは、出力を拒否することができるような仕組みの実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.10.情報セキュリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUは、Amazon Bedrockを通じて、目的外利用の防止と、仮に目的外利用された場合の危害・不利益の抑制のため、以下の機能を提供しています： 1. 入出力の制御による目的外利用の抑止 Amazon Bedrock Guardrailsにより、想定ユースケース以外の入力や、ユーザーの生命・身体・財産等や各種の権利に危害を生じさせ得る出力を検出・ブロックし、もしくは応答を拒否することができます。拒否トピック（Denied topics）やコンテンツフィルターを設定することで、業務の想定範囲外の話題への応答を制限できます。 2. 利用範囲の限定 IAMによるアクセス制御や、システムプロンプトによる役割・利用範囲の制約により、想定された目的の範囲内での利用に限定する設計が可能です。 3. 利活用状況のモニタリング モデル呼び出しのログ記録やAmazon CloudWatchにより、目的外と疑われる入出力を継続的に監視し、必要な対応につなげることができます。 なお、以下の事項については、事業者側での対応が必要となります： - 生成AIシステムの利用方法や利用上の留意点、想定利用環境、ガイドライン等の整備と、定期的な更新 - 業務特性に応じた想定ユースケースの定義と、拒否トピックや入出力制限ルールの設定 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/
						●	●	●	●	●	想定された目的の範囲内の生成AIシステムのユースケースにおいて、ユーザーの生命・身体・財産等や各種の権利に危害を生じさせる出力がされない、あるいは、出力を拒否できる仕組みを具備する	-	想定された目的の範囲内の生成AIシステムのユースケースにおいて、ユーザーの生命・身体・財産等や各種の権利に危害を生じさせる出力がされない、あるいは、出力を拒否することができるような仕組みの実現方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	

調達チェックシート						（参考）要求事項の参考														AWSサンプル回答
						対策例の対象AI														
						入力		出力												
分類	評価観点 #	評価観点	評価・選定時の項目の分類	要求事項 #	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	
						●	●	●	●	●	ユーザーが生成AIシステムを安全かつ適正に利用するため、生成AIシステムの利用方法や利用上の留意点、想定利用環境、ガイドライン等が定められ、定期的に更新する	-	生成AIシステムの利用方法や利用上の留意点、想定利用環境等が記載された資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	・ガードレールのコンポーネント：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails-components.html ・Amazon Bedrock のためのアイデンティティとアクセス管理：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/security-iam.html ・CloudWatch Logs と Amazon S3 を使用してモデル呼び出しをモニタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/model-invocation-logging.html
	16	個人情報、プライバシー、知的財産等	個人情報を取り扱う、又はプライバシーの保護が必要な場合は基本項目として適用	23	生成AIシステムにおける個人情報の適切な取扱いと、プライバシーの保護が図られるよう対策していること	●	●	●	●	●	RAGを利用した生成AIシステムを利用し、RAGによる検索先に個人に関する情報が含まれる場合、不適切な個人情報の出力を制御できる仕組みを具備する ※RAGを利用した生成AIシステムのみ本項目の対象	-	RAGを利用した生成AIシステムを利用し、RAGによる検索先に個人に関する情報が含まれる場合、不適切な個人に関する情報の出力を制御する仕組みの実現方法がわかる資料等	「調達仕様書標準テンプレート」 6.作業にあたっての遵守事項 6.3.個人情報等の取扱い 「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、個人情報、プライバシー、知的財産の保護のため、以下の機能を提供しています： 1. RAG利用時の情報保護 Knowledge Baseへのアクセスに関して、IAMによる細かなアクセス権限の設定が可能です。また、メタデータフィルタリング機能により、特定の属性や分類に基づいて個人情報を含むデータへのアクセスを制御できます。さらに、VPCエンドポイントによる通信の保護や、保存データの暗号化により、要機密情報の安全な取り扱いを確保できます。 2. Bedrock Guardrailsによる制御 Guardrailsを使用することで、PII（個人識別情報）の検出と制御、センシティブ情報の制御、プライバシー侵害につながる出力の防止が可能です。 3. 入力情報の取扱い：入力データは指定リージョン内に留まり、AWS は保存・モデル学習に使用しません（国民等の外部利用時に、入力を当該回答生成のみに利用する構成を支援）。 なお、以下の事項については、事業者側での対応が必要となります： - 具体的な個人情報保護方針の策定 - プライバシー侵害時の対応手順の整備 - 個人情報保護法に準拠したデータ収集プロセスの確立 （参考） ・Amazon Bedrock ナレッジベースのサービスロールを作成する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/kb-permissions.html ・Knowledge Bases for Amazon Bedrock がメタデータフィルタリングをサポートし検索精度向上：https://aws.amazon.com/jp/blogs/news/knowledge-bases-for-amazon-bedrock-now-supports-metadata-filtering-to-improve-retrieval-accuracy/ ・機密情報フィルターを使用して会話から PII を削除する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails-sensitive-filters.html
						●	●	●	●	●	プライバシー侵害を認識した場合等の対処方法について定める	-	プライバシー侵害時の対処手順がわかる資料等		-	
						●	●	●	●	●	個人情報保護法に準拠して、ユーザーからデータを収集する	-	個人情報保護法に準拠している旨が分かる資料（ルールや規程、その運用状況の記録等）		-	
						●	●	●	●	●	特に国民等外部の者が利用する場合で、個人情報等が入力される可能性があるときには、原則として、ユーザーの入力情報は当該ユーザーへの回答の生成にのみ利用される仕組みとなるよう確保する	-	個人情報等が入力される可能性がある場合、原則として、ユーザーの入力情報は当該利用者への回答の生成にのみ利用される仕組みの実装方法がわかる資料等		-	
						●	●	●	●	●	生成AIシステムにおいて個人情報を含むパーソナルデータを取り扱う場合、当該データの保護のため適切な加工等の対策を講じる	-	生成AIシステムにおいて個人情報を含むパーソナルデータを取り扱う場合、当該データの保護のために講じる適切な加工等の対策の実現方法がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
			生成AIシステムで知的財産等を取り扱う場合は基本項目として適用 ※①著作物の利用、②登録意匠・登録商標、他人の商品等表示・商品形態の利用、③人の肖像の利用、④人の声の利用が主に想定される。	24	生成段階において、知的財産権、肖像権、パブリシティ権の保護が図られるよう対策を講じていること	●	●	●	●	●	生成AIシステムにおいて知的財産権等の侵害等防止のための仕組みを具備する	-	知的財産権等の保護のための措置（フィルタリング等）の実現方法がわかる資料	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	経産省「コンテンツ制作のための生成AI利活用ガイドブック」 ※「肖像や声の保護に関しては、不正競争防止法や裁判例における考え方の整理がされている。 経済産業省、2025年「肖像と声のパブリシティ価値に係る現行の不正競争防止法における考え方の整理について」 https://www.meti.go.jp/policy/economy/chizai/chiteki/pdf/shozo_koe.pdf	GenUは、Amazon Bedrockを通じて、生成段階における知的財産権・肖像権・パブリシティ権の保護を支援するため、以下の機能を提供しています： 1. 出力フィルタリングによる侵害リスクの低減 Amazon Bedrock Guardrailsのコンテンツフィルターや拒否トピック（Denied topics）により、不適切な出力を検出・制御し、知的財産権等の侵害につながり得る出力の抑止を補助することができます。 2. 出力の検証・モニタリング モデル呼び出しのログ記録やAmazon CloudWatchにより、出力内容を継続的に監視し、権利侵害が疑われる事案の検知・対応につなげることができます。 なお、以下の事項については、事業者側での対応が必要となります： - 知的財産権等の侵害防止のためのフィルタリング等の措置の設計 - 生成物が第三者の権利を侵害していないかの確認プロセスの整備 - 権利者からの削除等の申出を受けた場合の是正等の合理的な措置の手順整備 （参考） ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・ガードレールのコンポーネント：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails-components.html ・CloudWatch Logs と Amazon S3 を使用してモデル呼び出しをモニタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/model-invocation-logging.html

調達チェックシート						（参考）要求事項の参考										
						対策例の対象AI										
						入力		出力								
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	AWSサンプル回答
			企画・開発において生成AIモデルに学習を行わせる場合は基本項目として適用	25	学習段階において、知的財産権等の保護が図られるよう対策を講じていること	●	●	●	●	●	知的財産権等の侵害等防止のため、適切なデータの学習を行う対策を講じる	外部から学習データの提供を受ける場合、当該データ提供者との契約・利用規約において、提供されるデータを生成AIへの学習用データとして利用することが制限されている場合がある。許諾なく学習データとして用いると契約・利用規約違反となる可能性があるため、契約・利用規約を確認し、データの利用制限の有無や内容をチェックする	データ提供者との契約・利用規約を踏まえての学習を行う際の対応方針がわかる資料等	「要件定義書標準テンプレート」 2.機能要件定義 2.4.データに関する事項 3.非機能要件定義 3.12.データマネジメントに関する事項	経産省「コンテンツ制作のための生成AI利活用ガイドブック」	GenUの標準的な構成は、Amazon Bedrock上の基盤モデルをそのまま利用するものであり、独自の追加学習（ファインチューニング）を前提としていません。追加学習やカスタムモデルを行う場合に備え、Amazon Bedrockを通じて以下を提供しています： 1. 学習データのリージョン内処理と非共有 Amazon Bedrockのカスタムモデル／ファインチューニング機能では、学習データはお客様のリージョン内で扱われ、ベースモデルの改変や他者への共有は行われません。 2. アクセス制御 IAMにより、学習データやカスタムモデルへのアクセスを必要最小限の権限に制御することができます。 なお、以下の事項については、事業者側での対応が必要となります： - 外部から提供を受けるデータの契約・利用規約上の、学習利用の可否の確認 - 著作権法第30条の4の「非享受目的」要件、及び「著作権者の利益を不当に害することとなる場合」への該当性の確認 - AI学習データの収集を制限する技術的措置（robots.txt等）の尊重、及び学習データをそのまま出力させない学習方法の採用 （参考） ・Amazon Bedrock でファインチューニングまたは継続的な事前トレーニングを使用してモデルをカスタマイズする： https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/custom-model-fine-tuning.html ・Amazon Bedrock のためのアイデンティティとアクセス管理：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/security-iam.html
												学習データをそのまま出力させるような学習方法をとらない	学習データをそのまま出力させない学習方法の実現方法がわかる資料等		文化庁「AIと著作権に関するチェックリスト＆ガイダンス」	
												自らが権利を有しておらず、また、著作権で保護された著作物等を学習データとして収集・利用する場合、①契約の締結等の権利者の許諾を得て行われているか、又は②権利者の許諾なく著作権法第30条の４の適用を受けて行われているか確認する。 ②の場合、学習データとしての収集・利用が同条の「非享受目的」要件を満たすか否か、「著作権者の利益を不当に害することとなる場合」に該当しないかを確認する。また、「非享受目的」のみの利用に留め、享受目的あるいは享受目的が併存すると評価される利用はせず、「著作権者の利益を不当に害することとなる場合」に該当するような利用はしない。 ①に関して、例えば以下のような取組が考えられる ・学習データには、著作権者から許諾を得る等権利処理されたデータ（ライセンスの契約等）を利用する ②に関して、「著作権者の利益を不当に害することとなる場合」に該当しないよう、例えば以下のような取組が考えられる ・AI学習データの収集を制限する技術的措置（ID/PWによるアクセス制限や ”robots.txt”による制限等）を尊重する	学習データの収集・利用等における権利者の許諾の状況の確認と、許諾を得る際の対応方針がわかる資料等 又は、行おうとする学習データの収集・利用等で著作権法30条の4の適用を受けるための確認方針と、その上で学習データの利用を「非享受目的」のみに利用する仕組みの実現方法がわかる資料等		文化庁「AIと著作権に関するチェックリスト＆ガイダンス」 経産省「コンテンツ制作のための生成AI利活用ガイドブック」	
	17	セキュリティ確保	基本項目	26	生成AIシステム全体の脆弱性に対処し、不正操作による影響の防止措置を取っていること	●	●	●	●	●	生成AIシステム全体に関するセキュリティ確保の対策を講じる	生成AIシステムに対する攻撃手法は日々新たなものが生まれており、これらのリスクに対応するための仕組みを検討・開発している、また必要に応じて導入することが可能である	最新の攻撃手法に対する対策についての資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.4.性能に関する事項 3.5.信頼性に関する事項 3.10.情報セキュリティに関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、以下のセキュリティ対策機能を提供しています： 1. システム負荷対策 Amazon Bedrockでは、APIリクエストに対するサービスクォータやモデルごとの入出力トークン数制限が設定されています。また、プロビジョンドスループットを使用することで専用のキャパシティを確保し、安定した処理能力を維持することが可能です。 2. プロンプトインジェクション対策 Bedrock Guardrailsにより、悪意のある入力の検出と制御、システムプロンプトの保護、制御回避の試みの防止が可能です。また、禁止リストによるプロンプトフィルタリングやパターンマッチングによる不正入力の検出を、リアルタイムで実施できます。 3. 最新の脅威への対応 Guardrailsの継続的な更新により、プロンプトインジェクションなどの既知の攻撃手法への対策を強化しています。また、AWS Security Bulletinを通じて最新の脆弱性情報を提供し、必要なセキュリティアップデートを通知します。 4. RAG利用時のセキュリティ対策 Knowledge Base利用時には、IAMロールによる細かなアクセス権限の設定やVector Store接続時の認証管理により、要機密情報の不適切な出力を防止できます。また、Bedrock Agents利用時には、IAMサービスロールによるアクセス制御やLambda関数による追加的な権限制御が可能です。 なお、セキュリティポリシーの策定、業務要件に応じた入力制限の設定、インシデント対応手順の整備は事業者側での対応が必要です（総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」も参照）。 （参考） ・Amazon Bedrock のクォータ：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/quotas.html ・Amazon Bedrock のガードレール：https://aws.amazon.com/jp/bedrock/guardrails/ ・Amazon Bedrock のプロビジョンドスループットでモデル呼び出し容量を増やす：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/prov-throughput.html ・Amazon Bedrock ナレッジベースのサービスロールを作成する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/kb-permissions.html
												生成AIモデルが意図しない出力を行わないよう、安全基準を事後学習させる	生成AIモデルが意図しない出力を行わないよう、安全基準を事後学習させるための開発方針とその方法が分かる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												生成AIモデルが従うべき指示の優先度を定義し、優先度の高い指示（例：システムプロンプト）を常に優先的に処理するよう、生成AIモデルを事後学習させる	生成AIモデルが従うべき指示の優先度を定義し、優先度の高い指示を常に優先的に処理するようにするための開発方針とその方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												生成AIシステムの意図しない動作の防止のために、オーケストレータの権限等に不備がないことの対策を講じる。生成AIモデルや連携システムを操作するオーケストレータに係る権限を必要最小限とすることで、生成AIモデルが攻撃を受けた場合の被害拡大を抑制する	生成AIモデルや連携システムを操作するオーケストレータに係る権限設定等に不備がないかの確認方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												入力された要機密情報を参照又は学習する場合、権限を有する者以外の回答の生成に、当該入力された要機密情報が用いられないような仕組みとする	入力された要機密情報を参照又は学習する場合、権限を有する者以外の回答の生成に、当該入力された要機密情報が用いられないような仕組みとする方法がわかる資料等		-	
						●	●	●	●	●	生成AIモデルへのプロンプト入力に関するセキュリティ確保の対策を講じる（入力プロンプト全般）	プロンプト内容の有害性を生成AIモデルへの入力前に検知する等、生成AIモデルへの入力前段階での防御策を講じる 例えば、以下のような対策を講じる ・禁止リストを活用したプロンプト検知 ・有害な結果をもたらす可能性のあるコードの実行や生成を促す入力、又は生成AIシステムの利用目的に照らして意図しない出力を生成させる指示がプロンプトに含まれていないか検証し、そのような入力を検知した場合にはプロンプトの一部削除による無害化や、処理の拒否等の措置を講じる ・プロンプトインジェクション等により、生成AIシステムのバックエンドシステムに意図していない操作が行われることがないよう対策する ・実装した生成AIシステムの防御策に関して、プロンプトインジェクション等により防御策の回避が可能とならないよう対策する	プロンプト内容の有害性を生成AIモデルへの入力前に検知する等、生成AIモデルへの入力前段階での防御策の実装方針がわかる資料等		対策例詳細： AISI「AIセーフティに関する評価観点ガイド（第1.10版）」 総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												生成AIシステムが処理困難でコストのかかる複数のテストデータを繰り返し入力した場合でも、生成AIシステムの許容できないパフォーマンス低下やシステムの停止が発生しないよう対策を講じる	生成AIシステムが処理困難でコストのかかる複数のテストデータを繰り返し入力した場合でも、生成AIシステムの許容できないパフォーマンス低下やシステムの停止が発生しないかの確認方法がわかる資料等		対策例詳細：AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												生成AIシステムへの膨大なアクセスによる攻撃を抑制するためのレートリミットを導入する	生成AIシステムのレートリミットの仕組みの実装方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
						●	●	●	●	●	生成AIモデルへのプロンプト入力に関するセキュリティ確保の対策を講じる（特にユーザーが入力するプロンプト）	生成AIモデルがシステムプロンプトを無視してユーザー入力のみに従うよう生成AIモデルに指示する（いわゆるjail breakを試みる）ユーザー入力を無効化するようにシステムプロンプトをデザインする	jail breakを試みるユーザー入力を無効化する仕組みの実現方法がわかる資料等		対策例詳細：産総研「生成AI品質マネジメントガイドライン 第１版」	
												生成AIシステムの目的と用途に応じて、ユーザーを認証して、認証を通らないユーザーの利用を拒否することによって、ユーザー入力プロンプトの来歴が明らかになるよう制御する	ユーザーを認証して、認証を通らないユーザーの利用を拒否することによって、ユーザー入力プロンプトの来歴が明らかになるよう制御する仕組みの実現方法がわかる資料等		対策例詳細：産総研「生成AI品質マネジメントガイドライン 第１版」	
						●	●	●	●	●	生成AIモデルへのプロンプト入力に関するセキュリティ確保の対策を講じる（システムプロンプト）	システムプロンプトに制約事項やセキュリティ上の注意事項等を設定することで、生成AIモデルが意図しない出力を行わないようにする	システムプロンプトに制約事項やセキュリティ上の注意事項等を設定することで、生成AIモデルが意図しない出力を行わないようにする仕組みを提供する技術の実装方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												システムプロンプトに出力を意図しない機密情報（例：APIキー）等を直接記述することを避け、生成AIモデルが必要に応じて参照できるよう別個に管理する	システムプロンプトには、出力を意図しない機密情報等を直接記述することを避け、生成AIモデルが必要に応じて参照できるよう別個に管理する対策の対応方針がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	

調達チェックシート						（参考）要求事項の参考														AWSサンプル回答
						対策例の対象AI														
						入力		出力												
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考）「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考）対策例・対策例詳細の参考文献	
						●	●	●	●	●	生成AIモデルの出力に関するセキュリティ確保の対策を講じる	生成AIシステムの意図しない動作の防止のための出力の検証機能を具備し、出力を意図しない情報が出力に含まれていないか検証し、検知した場合には応答を拒否する	出力を意図しない情報が出力に含まれていないか検証し、検知した場合には応答を拒否する仕組みの実現方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	GenUは、Amazon Bedrockを通じて、情報セキュリティインシデント・生成AIシステム特有のリスクケースの検知後における、迅速な封じ込めと復旧（制御性の強化）を支えるため、以下の機能を提供しています： 1. 代替モデル・別バージョンへの切替 APIレベルでのモデル指定により、改ざんや劣化が疑われるモデルから、別バージョンや代替モデルへ迅速に切り替える構成が可能です。また、クロスリージョン推論やプロビジョンドスループットにより可用性を維持することができます。 2. 検知のためのログ・監視 AWS CloudTrailによるAPI呼び出しの記録、Amazon CloudWatchによる異常検知、モデル呼び出しのログ記録により、インシデントやリスクケースの兆候を検知することができます。 3. 復旧の容易性 GenUはAWS CDKで構成されるため、設定のロールバックや再デプロイにより既知の正常な状態へ復旧する運用を設計しやすく、RAG利用時はKnowledge Baseのデータ再同期により汚染データの影響を限定することができます。 なお、以下の事項については、事業者側での対応が必要となります： - 検知後の対応手順（即時切替・隔離・通知）の整備と予行演習 - 異常時に影響を受けない上位権限を持つ、隔離された制御インタフェースの設置 - 混入した悪意ある学習データの局所的除去や追加学習等のモデル汚染への対応、及び情報セキュリティインシデント対応体制との連携 （参考） ・AWS CloudTrailを使用した API コールのログ記録：https://docs.aws.amazon.com/ja_jp/fis/latest/userguide/logging-using-cloudtrail.html ・CloudWatch 異常検出の使用：https://docs.aws.amazon.com/ja_jp/AmazonCloudWatch/latest/monitoring/CloudWatch_Anomaly_Detection.html ・CloudWatch Logs と Amazon S3 を使用してモデル呼び出しをモニタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/model-invocation-logging.html
												単語の出現確率等、攻撃者に悪用され得る情報を必要に応じて応答から除外することで、モデル抽出攻撃への対策を行う（主にテキスト出力を念頭）	単語の出現確率等、攻撃者に悪用され得る情報を必要に応じて応答から除外することで、モデル抽出攻撃への対策をする仕組みの実現方法がわかる資料等		対策例詳細：総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
						●	●	●	●	●	生成AIモデルが生成AIモデルの外のデータを参照する場合のセキュリティ確保の対策を講じる	RAGにて要機密情報のように参照権限が設定されている情報を利用する場合、RAG用のデータ及びデータストアへの参照権限をユーザーや役割に応じて適切に設定する（主にテキスト出力を念頭）	RAG用のデータ及びデータストアへの参照権限をユーザーや役割に応じて適切に設定し、不備がないかの確認方法がわかる資料等		対策例詳細： AISI「AIセーフティに関する評価観点ガイド（第1.10版）」 総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												外部サービスの呼び出しを行うためのプラグインツールを活用する場合、プラグインツールがアクセスする外部サービスやデータベース等の制御を実施する	プラグインツールがアクセスする外部サービスやデータベース等の制御の実現方法がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
												外部サービスの呼び出しを行うためのプラグインツールを活用する場合、機能の導入前に当該プラグインツールの情報（SBOM 等）を確認し、外部連携としての利用可否を判断する	プラグインツールを導入する場合の、プラグインツールの情報の確認方針がわかる資料等		産総研「生成AI品質マネジメントガイドライン 第１版」	
												Webサイトや外部のデータベース等、外部データを参照する場合には、これらに意図しない出力を行わせる指示が含まれていないか検証し、そのような指示を検知した場合には、処理を拒否する等の措置を講じる	Webサイトや外部のデータベース等、外部データを参照する場合には、これらに意図しない出力を行わせる指示が含まれていないか検証し、そのような指示を検知した場合には、処理を拒否する等の仕組みの実現方法がわかる資料等		総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
												生成AIシステムに、入力プロンプトと外部参照データを明確に区分させ、外部参照データに高い注意を払わせる	生成AIシステムに、入力プロンプトと外部参照データを明確に区分させ、外部参照データに高い注意を払わせる仕組みの実現方法がわかる資料等		総務省「AIのセキュリティ確保のための技術的対策に係るガイドライン」（LLM及びLLMを構成要素に含むAIシステムが対象）	
			生成AIシステムに対して高い信頼性が求められる場合は基本項目として適用	27	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応のために、生成AIシステムの制御性を強化するための仕組みを具備する	●	●	●	●	●	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応のために、生成AIシステムの制御性を強化するための仕組みを具備する	改ざんされたモデルの別バージョンや代替モデルへの即時切り替えを実施する	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応方法として、改ざんされたモデルの別バージョンや代替モデルへ即時切替する仕組みの実現方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.4.性能に関する事項 3.5.信頼性に関する事項 3.10.情報セキュリティに関する事項 3.16.運用に関する事項 3.17.保守に関する事項	AISI「AIインシデントレスポンス・アプローチブック」	
												混入された悪意のある学習データの局所的な除去や追加学習を実施する	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応方法として、混入された悪意のある学習データの局所的な除去や追加学習等の、生成AIモデルの汚染対策の方針がわかる資料等		AISI「AIインシデントレスポンス・アプローチブック」	
												異常時に影響を受けない上位権限を持ち、隔離された制御インタフェースを設置する	情報セキュリティインシデント・生成AIシステム特有のリスクケース発生時の対応を想定して、異常時に影響を受けない上位権限を持ち、隔離された制御インターフェースの設置等の対策の実装方針がわかる資料等		AISI「AIインシデントレスポンス・アプローチブック」	
			基本項目	28	生成AIシステムの開発の過程を通じて、適切にセキュリティ対策を講じていること	●	●	●	●	●	生成AIシステムの開発の過程を通じて、採用する技術の特性に照らして適切なセキュリティ対策を講じる	-	セキュリティ対策についての資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.4.性能に関する事項 3.5.信頼性に関する事項 3.10.情報セキュリティに関する事項	総務省、経産省「AI事業者ガイドライン 第1.2版」	AWSでは、開発プロセスでのセキュリティ対策として以下を提供しています： 1. セキュリティ実装のガイダンス OWASP Top 10 for LLM Applicationsに基づく実装ガイドを提供しています： - プロンプトインジェクション対策の実装方法 - データ漏洩対策の実装方法 - 認証とセッション管理の実装方法 - その他のLLMアプリケーション特有のセキュリティ考慮事項 なお、以下の事項については、事業者側での対応が必要となります： - セキュアな開発プロセスの確立 - OWASP Top 10 for LLMに基づくセキュリティ評価の実施 - 脆弱性評価とペネトレーションテストの実施 - インシデント対応プランの策定 （参考） ・AWS で実現する安全な生成 AI アプリケーション – OWASP Top 10 for LLM Applications 2025 の活用例：https://aws.amazon.com/jp/blogs/news/secure-gen-ai-applications-on-aws-refer-to-owasp-top-10-for-llm-applications/

調達チェックシート						（参考）要求事項の参考							AWSサンプル回答
						対策例の対象AI							
						入力		出力					
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	
	18	説明可能性	基本項目	29	出力根拠が技術的に合理的な範囲で確認できる状態とするよう対策していること	●	●	●	●	●	出力根拠（内部動作やその状態、出典等）が可視化される機能を備える生成AIシステムにおいて様々なテストデータを入力した際、出力根拠が表示される仕組みを具備する	生成AIが出力を生成するために利用した情報を、出力の正当化理由として出力するようにシステムプロンプトをデザインしている、ファイルデータや外部連携コンポーネントからの応答等実行時に選択してプロンプトに組み込む情報については、その実歴を示す情報を添えて組み込んでいる、これと合わせ、システムプロンプトで来歴情報を活用するよう指示している	出力根拠（内部動作やその状態、出典等）が可視化される機能を備える生成AIシステムにおいて様々なテストデータを入力した際、出力根拠が表示される仕組みを提供する技術の実装方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.1.ユーザビリティ及びアクセシビリティに関する事項	対策例：AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第1版」	GenUでは、Amazon Bedrockを通じて、出力根拠の確認を可能とするため、以下の機能を提供しています： 1. 推論過程の可視化機能 Chain of Thoughtアプローチにより、内部の推論過程を表示します。これにより、推論の各ステップにおける判断根拠を提示し、中間的な推論結果を示しながら、最終出力に至るまでの論理的な流れを説明することが可能です。エンドユーザーに対しても、この段階的な推論過程を分かりやすく提示することができます。 2. 出典情報の追跡機能 RAG（Retrieval-Augmented Generation）利用時には、参照された情報源を明示するとともに、情報源からの引用部分を特定します。また、出力内容と情報源との関連性について詳細な説明を提供することが可能です。
						●	●	●	●	●	段階的な推論を行う生成AIシステムにおいて、出力に至るまでの推論の過程をユーザーに提示することが可能となっているかを確認する	-	段階的な推論を行う生成AIシステムにおいて、出力に至るまでの推論の過程をユーザーに提示することが可能となっているかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	なお、以下の事項については、事業者側での対応が必要となります： システム実装に関して、テストデータの設計と準備、出力根拠の表示方法の実装、推論過程の提示方法のカスタマイズが必要です。また、システム開発過程については、使用されているアルゴリズムの概要、データ収集・処理プロセス、システムアーキテクチャ、意思決定に影響を与える要素についての文書化を行う必要があります。さらに、システム固有の処理フローの記録や根拠の妥当性評価基準を設定し、継続的な評価を実施することが求められます。 （参考） ・Amazon Bedrock アプリケーションで責任ある AI のコアディメンションに対応するための考慮事項：https://aws.amazon.com/jp/blogs/news/considerations-for-addressing-the-core-dimensions-of-responsible-ai-for-amazon-bedrock-applications/ ・Amazon Bedrock のナレッジベース：https://aws.amazon.com/jp/bedrock/knowledge-bases/
						●	●	●	●	●	生成AIシステムの開発過程、意思決定に影響を与えるデータ収集及びラベリング、使用されたアルゴリズム、生成AIシステムのシステムアーキテクチャ、データの処理プロセス等について文書化を行う	-	生成AIシステムの開発過程、意思決定に影響を与えるデータ収集及びラベリング、使用されたアルゴリズム、システムアーキテクトやデータの処理プロセスが文書として管理されていることがわかる資料等 ※文書化した内容について開示するという意味ではない		総務省、経産省「AI事業者ガイドライン 第1.2版」	
	19	ロバスト性	基本項目	30	生成AIシステムが入力に対して安定した出力を行うよう制御等をしていること	●	●	●	●	●	生成AIシステムに同一のテストデータを複数回入力した際の出力に一貫性があるよう対策を講じる	-	生成AIシステムに同一のテストデータを複数回入力した際の出力に一貫性があるかの確認方法がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.4.性能に関する事項 3.5.信頼性に関する事項 3.9.継続性に関する事項	AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、入力に対する安定した出力を確保するため、以下の機能を提供しています： 1. 同一入力に対する一貫性制御機能 Amazon Bedrock APIでは、Temperature（温度）パラメータによる出力のランダム性制御やTop-Pパラメータによるトークン選択の制御を行います。また、最大トークン数の制限による出力長の安定化やResponse Formatの機能による出力形式の統一により、同一のテストデータに対して一貫性のある出力を実現します。 2. 入力保護機能 Bedrock Guardrailsにより、プロンプトインジェクションからの保護や不適切な入力の検出と制御を行います。また、システムプロンプトの保護機能により、意図しない入力に対しても一貫性のある応答を実現します。 3. モデル評価機能 Amazon Bedrock Model Evaluationを使用して、同一入力や類似入力に対する出力の一貫性を評価することができます。また、RAG実装時には、検索結果の一貫性や関連性についても評価が可能です。これにより、システム全体の安定性を継続的に監視し、評価することができます。
						●	●	●	●	●	生成AIシステムに意味的に類似した複数のテストデータを入力した際、出力に一貫性があるよう対策を講じる	ユーザー入力に揺れがあっても、生成AIシステムが同等なコンテンツを出力するように、その揺れを意味的に吸収するようなシステムプロンプトをデザインする	生成AIシステムに意味的に類似した複数のテストデータを入力した際、出力に一貫性があるかの確認方法がわかる資料等		対策例：AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第1版」	なお、以下の事項については、事業者側での対応が必要となります： 安定性の評価と維持のため、テストケースの作成と実施、定期的な出力品質の評価が必要です。また、業務要件に応じたパラメータの適切な設定、入力データの前処理ルールの設定、出力の一貫性に関する評価基準の設定を行う必要があります。特に、表記ゆれや類似表現への対応、文字化けデータの検出と処理については、適切な前処理機能の実装が必要となります。さらに、同一入力、類似入力、異常入力それぞれについての具体的なテストシナリオの作成と実行が求められます。 （参考） ・プロンプトインジェクションのセキュリティ：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/prompt-injection.html ・推論パラメータでレスポンスの生成に影響を与える：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/inference-parameters.html ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/
						●	●	●	●	●	生成AIシステムに摂動したデータ（誤入力、敵対的プロンプト、文字化けデータ、表記ゆれを含むデータ等）を入力した際も安定動作するよう対策を講じる	ユーザー入力に揺れがあっても、生成AIシステムが同等なコンテンツを出力するように、その揺れを意味的に吸収するようなシステムプロンプトをデザインする	生成AIシステムに摂動したデータ（誤入力、敵対的プロンプト、文字化けデータ、表記ゆれを含むデータ等）を入力した際も安定動作するかの確認方法がわかる資料等		対策例：AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」 対策例詳細：産総研「生成AI品質マネジメントガイドライン 第1版」	
	20	データ品質	基本項目	31	生成AIシステムがアクセスするデータを適切な状態に保つために対策していること	●	●	●	●	●	学習データ、モデルの学習過程において、バイアス（学習データには現れない潜在的なバイアスを含む）が含まれうることに留意し、生成AIシステムに悪影響を及ぼすデータ（事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等）の品質問題が生じていないことの対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、データの正確性を確保し、正確性、機密性、完全性の管理方法がわかる資料等	「調達仕様書標準テンプレート」 4.作業の実施内容に関する事項 4.13.データ管理方法 「要件定義書標準テンプレート」 2.機能要件定義 2.4.データに関する事項	AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、データの適切な状態を維持するため、以下の機能を提供しています： 1. データの品質制御機能 Amazon Bedrock Knowledge Baseでは、メタデータを活用したデータフィルタリングやデータバージョニング機能により、データの正確性を確保します。また、Bedrock Guardrailsによる入出力制御により、文法的に適切なデータの維持と人間が理解しやすい内容の確保を実現します。さらに、不適切な内容や攻撃的な言葉のフィルタリング、サイバー攻撃やデに利用され得る情報の検出と制御を行います 2. 公平性確保機能 Bedrock Model Evaluationを使用して、データの分布における人種、性別、国籍、年齢、政治的信念、宗教等による偏りを検出し、評価することができます。また、AI Service cardsを通じて、モデルのバイアスに関する情報を開示します。 なお、以下の事項については、事業者側での対応が必要となります： 生成AIシステムで使用するデータに関して、具体的なデータ品質基準の設定、データクレンジングプロセスの確立、アノテーション精度の管理が必要です。また、データの偏りやバイアスの具体的な評価基準の設定と対策の実施、プライバシー影響評価の実施、データガバナンスポリシーの策定と運用が求められます。さらに、セキュリティ脅威への具体的な対応手順の策定と実施も必要となります。 （参考） ・Knowledge Bases for Amazon Bedrock がメタデータフィルタリングをサポート、検索精度向上：https://aws.amazon.com/jp/blogs/news/knowledge-bases-for-amazon-bedrock-now-supports-metadata-filtering-to-improve-retrieval-accuracy/ ・Amazon Bedrock ガードレールを使用して有害なコンテンツを検出してフィルタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/guardrails.html ・Amazon Bedrock の評価：https://aws.amazon.com/jp/bedrock/evaluations/
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、データの破損を防ぐ管理及び、データに問題が生じないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、正確性、機密性、完全性の管理方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、データが文法的に適切であり、人間が理解しやすい内容となるよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、データが文法的に適切であり、人間が理解しやすい内容となっているかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、データの分布に人種、性別、国籍、年齢、政治的信念、宗教等による偏りがないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、データの分布に人種、性別、国籍、年齢、政治的信念、宗教等による偏りがないかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、ユーザーに対して攻撃的な言葉等、不適切な内容がデータに含まれないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、ユーザーに対して攻撃的な言葉等、不適切な内容がデータに含まれていないかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、サイバー攻撃やデロ等に利用され得る情報がデータに含まれないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、サイバー攻撃やデロ等に利用され得る情報がデータに含まれていないかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、PETs(Privacy-Enhancing Technologies)等による措置を行うことで、学習データとしての使用に問題ない状態となっているかの確認をする	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータに関して、PETs(Privacy-Enhancing Technologies)等による措置を行うことで、学習データとしての使用に問題ないかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、データの構成管理及び、データに問題が生じないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、データの構成管理方針がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	
											事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、アノテーションを自動化する場合、適切に行われるよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、アノテーションを自動化する場合、適切に行われているかの確認方法がわかる資料等		AISIΓAIセーフティに関する評価観点ガイド（第1.10版）」	

調達チェックシート						（参考）要求事項の参考										
						対策例の対象AI										
						入力		出力								
分類	評価観点#	評価観点	評価・選定時の項目の分類	要求事項#	要求事項	テキストを含む	音声を含む	テキストを含む	画像を含む	音声を含む	対策例	対策例詳細	裏付けとなる情報の例	（参考） 「DS-120 デジタル・ガバメント推進標準ガイドライン実践ガイドブック」の「各種テンプレート」との対応関係	（参考） 対策例・対策例詳細の参考文献	AWSサンプル回答
												事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、悪意をもった、あるいは誤動作させるテキスト・プログラム等がデータに含まれないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、悪意をもった、あるいは誤動作させるプログラムがデータに含まれていないかの確認方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
												事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、ツール等により、データの来歴管理及び、データに問題が生じないよう対策を講じる	事前学習データ、ファインチューニングやInContext Learning等に向けた学習データやテストデータ等生成AIシステムに影響を及ぼすデータの品質管理として、データの来歴管理方法がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	訓練データの量と質が十分で、実用的な水準まで学習されているかを検証する	-	訓練データの量と質が十分で、実用的な水準まで学習されているかの検証プロセス・検証方法が分かる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
						●	●	●	●	●	機微情報を含むデータに対して、期待される保護強度を示すデータ加工（データ匿名加工）を施す技術を有していることや、データの利用目的との整合性からデータ匿名加工の方法を選定する	-	-		産総研「生成AI品質マネジメントガイドライン 第１版」	
			任意追加項目 （加点項目）	32	生成AIシステムのアウトプットの高度化としてインプットデータの適切な構造化を行っていること	●	●	●	●	●	アウトプットの精度向上を目的とした、インプットデータの構造化を必要に応じて検討・実施する	-	アウトプットの精度向上を目的とした、インプットデータの構造化の実現方法がわかる資料等	「要件定義書標準テンプレート」 2.機能要件定義 2.4.データに関する事項	-	GenUでは、Amazon Bedrockを通じて、インプットデータの構造化を支援するため、以下の機能を提供しています： 1. データ構造化支援機能 Amazon Bedrock Data Automationを使用することで、非構造化データから構造化データへの変換を支援します。PDFやその他のテキスト文書から情報を抽出し、JSONなどの構造化された形式に自動的に変換することが可能です。 なお、以下の事項については、事業者側での対応が必要となります： - 業務要件に応じたデータ構造の設計 - データの分類・整理基準の策定 - メタデータの設計と付与ルールの策定 - 構造化データの品質管理プロセスの確立 - Bedrock Data Automationでカバーできない文書フォーマットに対する追加の変換処理の実装 （参考） ・Amazon Bedrock Data Automation を使用して非構造化データを有意義なインサイトに変換する：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/bda.html
	21	検証可能性	基本項目	33	生成AIシステムの開発・提供のプロセスを検証可能としていること	●	●	●	●	●	システムやモデル、データに関する概要や、なりたち、動作についてログや技術仕様等から検証可能にするための管理及び、問題が生じないよう対策を講じる	システムやモデル、データに関する概要やなりたち、動作についての検証可能性を確保するため、システムカード、モデルカード、データカードを作成して管理及び、問題が生じないよう対策を講じる	システムやモデル、データに関する概要やなりたち、動作についての検証可能を確保するための技術仕様書等の管理方針がわかる資料等	「要件定義書標準テンプレート」 3.非機能要件定義 3.16.運用に関する事項 3.17.保守に関する事項	AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	GenUでは、Amazon Bedrockを通じて、システムの検証可能性を確保するため、以下の機能を提供しています： 1. モデルの透明性確保機能 AI Service cardsを通じて、モデルの使用目的、想定される用途と制限事項、公平性に関する考慮事項、評価指標と評価結果などを文書化し、システムカードやモデルカードとして提供します。これにより、使用されているモデルの概要やなりたちについて検証可能な状態を維持します。 2. 動作検証ログ機能 Amazon Bedrockのログ機能により、テストデータ入力時の生成AIモデルへのリクエストログ、プロンプトと応答の履歴、API呼び出しの詳細を記録します。これにより、様々なテストデータに対するシステムの動作を検証可能な状態で保持します。 なお、以下の事項については、事業者側での対応が必要となります： システム全体の検証可能性を確保するため、開発プロセスの文書化、システムアーキテクチャの文書化、データ処理プロセスの文書化が必要です。また、テストデータの入力結果の記録と評価、具体的な検証手順の確立も求められます。さらに、システム固有のカード類（システムカード、データカード等）の作成と管理も実施する必要があります。 （参考） ・責任ある AI： https://aws.amazon.com/jp/machine-learning/responsible-ai/ ・CloudWatch Logs と Amazon S3 を使用してモデル呼び出しをモニタリングする：https://docs.aws.amazon.com/ja_jp/bedrock/latest/userguide/model-invocation-logging.html
												システムやモデル、データに関する概要やなりたち、動作についての検証可能性を確保するため、生成AIシステムに様々なテストデータを入力した際、適切にデータログが記録されるよう対策を講じる	生成AIシステムに様々なテストデータを入力した際、適切にデータログが記録されているかの確認方針がわかる資料等		AISI「AIセーフティに関する評価観点ガイド（第1.10版）」	
*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	