

業務名		記入日	
システム名		記入者	
決裁番号			

AIリスクアセスメント ver1.00

基礎情報

番号	区分	要件	回答
A.1	事業について	以下の小項目を記入	-
A.1a	神戸市の責任者	所属、役職、氏名	
A.1b	事業趣旨・システム概要	事業の趣旨及びシステムの概要を記載	
A.1c	稼働予定時期	AIシステムが提供される予定時期	
A.2	AIの提供者	以下の小項目を記入	-
A.2a	会社名	会社名を記入	
A.2b	製品名	製品名を記入	
A.2c	サービス提供の形態	サービス提供の形態を記入(SaaS、オンプレ、ソフトウェア等)	
A.2d	管轄裁判所・準拠法	製品使用にあたり適用される管轄裁判所・準拠法を記入	
A.3	AIの機能	以下の小項目を記入	-
A.3a	AIの機能	入力情報に対して、AIがどう処理を行い、どう情報を出力(アウトプット)するのか。	
A.3b	AIが業務で果たす役割と利益	<ul style="list-style-type: none"> <li>具体的な役割の内容</li> <li>誰にどのような影響や利益をもたらすのか。</li> <li>AIの業務量(年間の処理件数等)</li> </ul>	記載例(アンケート分析システムを導入する場合) <ul style="list-style-type: none"> <li>市民からの質問のデータは蓄積されているが、傾向を把握できていない状況。AIが市民からの質問傾向を分析する。</li> <li>職員が実施できていない分析をAIが行ってくれるため、職員の仕事の質が向上する。市民には直接影響を与えない。</li> </ul>
A.3c	AI導入による業務の変化	現状の業務のうち、AIが一部業務を代替するのか、一部業務を支援するのか、未実施業務を新規に実施するのか、等	記載例(アンケート分析システムを導入する場合) <ul style="list-style-type: none"> <li>現状実施できていない分析業務をAIが新規に実施する。</li> </ul>
A.3d	AI活用の成否の判断	AI活用の成否を判断するために、どのような評価指標が使用されるか。	記載例(アンケート分析システムを導入する場合) <ul style="list-style-type: none"> <li>職員にヒアリングを行い、50%以上の人が有用だと答えること。</li> </ul>
A.4	AIの基礎情報	以下の小項目を記入	-
A.4a	AIの基盤モデル等	生成AIの場合は基盤モデル(GPT-4等)、その他のAIの場合はライブラリ(TensorFlow, OpenCV等)の名称を記入	
A.4b	AIのアルゴリズム	AIが使用するアルゴリズム(機械学習(深層学習含む)、ルールベースAI等)を記入	
A.4c	学習の有無	学習の有無を記入 有の場合は、学習済みのモデルを使用しているのか、運用後も追加学習を行うかを記入。追加学習を行う場合は神戸市独自のデータによる学習か、一般的な情報を基にした学習かを記載。	
A.4d	学習のデータ	学習に用いるデータセット及び提供者を記入	
A.4e	カスタマイズの有無	カスタマイズの有無を記入 有の場合はAIの機能に関わるカスタマイズかを記入	
A.5	重要項目	以下の小項目を記入	-
A.5a	サービスレベルの維持	AIを導入することにより、既存の仕組みに比べて、サービスレベルが低下していないことを、テスト等(※)を通じて確認する運用となっているか、また、既存の仕組みとの並行期間を設定するなど、検証期間を十分に設けているか。 ※ テスト等は、AIの持つあいまいさやハルシネーションによるリスクを考慮したうえでサービスレベルを保てる内容とすること。 [ 注意事項 ] なお、本項目の内容が不十分な場合は、デジタル戦略部より事業の延期や中止を求めることがある。	<ul style="list-style-type: none"> <li>テスト等の内容</li> <li>※ AIのあいまいさ等のリスクへの対応</li> <li>検証期間の設定</li> </ul>
A.5b	AIによる完全自動の出力	AIの出力が、完全に自動化された方法でエンドユーザ(市民、職員)に使用される予定か。それとも人間による最終判断がなされるものか。	<ul style="list-style-type: none"> <li>AI出力の自動的な使用有無</li> <li>(自動化された出力を使用する場合)リスクに対する対策</li> </ul>
A.5c	リスクアセスメントの対象となるか	<ul style="list-style-type: none"> <li>以下の3類型に当てはまるかどうか。</li> <li>---</li> <li>① 公権力の行使にあたる行政処分判断に活用するとき</li> <li>② 市の基本的な政策を定める計画策定及び評価・見直しにかかる現状または課題の把握、目標設定、取り組み、評価方法の検討・決定に活用するとき</li> <li>③ その他市民・事業者に重大な影響を及ぼす可能性があるものに活用するとき</li> <li>ア 本市要綱に基づく給付など、行政処分に基づき判断に活用するとき</li> <li>イ 行政指導等の事実行為のうち市民の生命、身体、健康、財産及び公共性を有する財産に重大な影響を与えるおそれがある判断に活用するとき</li> <li>---</li> <li>当てはまらない理由</li> </ul>	記載例(アンケート分析システムを導入する場合) <ul style="list-style-type: none"> <li>3類型に当てはまらないと考える。</li> <li>AIが出力した市民からの問い合わせ傾向等の分析結果は、WEBページの掲載内容の改善に利用するのみに、①～③に該当しないため。</li> </ul>

A.5aが当てはまる場合	⇒	ワークシートを記入し、デジタル戦略部に審査を依頼
A.5aが当てはまらない場合	⇒	(A.5aが満たせないもしくはA.5bに該当する場合)チェックシートを記入し、情報管理者が確認したうえでデジタル戦略部に事前協議
	⇒	(上記以外の場合)チェックシートを記入し、情報管理者が確認したうえでデジタル戦略部に事前提出

○ AIリスクアセスメントチェックシート

本チェックシートはAI利用に関するリスクとその対応方法を確認し、必要な検討をするために使用します。

- ① 各項目に記載されたリスクやあるべき姿について、所属での対応内容を検討してください。
- ② 評価基準に従って所属評価を採点してください。
- ③ すべての観点について2点以上であることを情報管理者が確認し、チェックを入れてください。

- ・ 緑背景になっている対応状況セルは必ず入力してください。
- ・ 緑背景以外の対応状況セルは任意入力です。

所属名	
業務名	
情報管理者	
記入日	
記入者	

評価基準	1	2	3	その他
詳細	リスクの洗い出しまたはリスクに対する対応策の検討が不十分。	リスクに対して、対応策が検討されている。	リスクに対して、対応策が検討されており、一貫した対応を行うための体制が整備されている。	該当しない

	人間中心	対応状況	所属評価				チェック
			該当無	1	2	3	
	人間の尊厳を守りながら、AIの活用を促進し、価値の創出や社会課題の解決等を実現すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIが活用される際の社会的文脈を踏まえ、人間の尊厳及び個人の自律を尊重していること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIを人間の脳・身体と連携させる使い方をしていないこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIを利用したプロファイリング(※実行わないこと。 ※ 個人の情報を収集・分析してその個人の性質・嗜好・行動等を予測すること 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、又は意図的に知覚できないレベルでの操作を前提としたAIシステムサービスの開発提供利用は行わないこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIシステムサービスの開発・提供・利用において、自動化バイアス等のAIに過度に依存するリスクに注意を払い、必要な対策を講ずること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	フィルタリング(※)に代表されるような情報又は価値観の偏りを助長し、選択肢が不本意に制限されるようなAIの活用にも注意を払うこと。 ※ 過去のユーザー情報をもとに、各人に最適化されたコンテンツが表示されることで、似たような情報や、視点に固まってしまう状態のこと		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	影響範囲の特定・安全性の確保	AIを使用する事務の内容と使用範囲を特定し、AIが誤作動した場合に生じる影響を把握すること。 正確性が必要な場合には最新性(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こと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	全体の業務内容のうち、どの部分でAIを使用し、それが最終決定にどう影響するか把握すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AI提供者が定めたサービス規約、利用上の留意点を遵守した運用を行うこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIの出力について精度及びリスクの程度を理解し、様々なリスク要因を確認・対策した上で利用すること。 AIが誤作動した場合に生じる影響を把握すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	正確性・必要な場合には最新性(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こと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIシステムサービスの出力の正確性を高め、仕様書等で合意した要求水準に対して十分に動作していること(信頼性)を確認すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	プライバシーの保護	学習データの収集、分析などにおいて、プライバシー保護に留意し、個人情報を慎重に取り扱うこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIシステム・サービスへの個人情報の不適切な入力を防止すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	(学習を行う場合)学習時のデータについて、適正に収集するとともに、第三者の個人情報、知的財産権に留意が必要なものが含まれている場合には、法令に従って適切に扱うこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	(独自の学習を行う場合)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こと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	透明性の確保	AIの判断理由、根拠を提示できる仕組みとすること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIの判断にかかわる検討可能性を確保するため、利用時の入出力、学習プロセス、推論過程、判断根拠等を提示できるサービスを利用すること。 また、必要に応じてAIを使用していることをエンドユーザー(市民・事業者等)に提示すること。特にAIによる出力を自動で表示する場合は、AIによる出力であり、誤りを含む可能性があることを提示すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	公平性の確保	バイアスが発生するリスクを低減すること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	学習データ・アルゴリズムに差別や偏見につながるバイアスが発生するリスクを低減するための対策を行うこと。 また、学習することにより新たに差別や偏見につながるバイアスが発生するリスクを認識し、対応すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	センシティブ属性(公平性の観点から排除すべき対象者の性別や人種等の個人の属性)を明確化するとともに、センシティブ属性に基づき差別や偏見につながるバイアスが発生するリスクを低減するための対策を行うこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	セキュリティの確保	AI特有の脆弱性に関するリスクを認識し、リスク回避のためにAI提供者によるセキュリティ上の留意点を遵守すること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AI提供者によるセキュリティ上の留意点を遵守し、マニュアル等に記載することで周知徹底すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIシステム・サービスに機密情報等を意図せず入力しない仕組みをシステム面・運用面で担保すること。また、使用者に機密情報等を入力しないことを周知徹底すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIシステム・サービスに入力された情報が外部に漏えいしない仕組みを担保すること。(事業者による学習結果の再利用、監査のための事業者の閲覧による漏えいを含む)		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	アカウントビリティの確保	利害関係者を把握し、責任者を明確にしたうえで判断に関する正当な意味理由の説明、必要に応じた賠償補償等の措置が取れること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	神戸市および関係者の各主体でアカウントビリティを果たす責任者を設定すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	関係者間の責任について、業務外利用者も含めた主体間の契約、社会的な約束(ボランティア・コミットメント)等により、責任の所在を明確化すること。また、必要に応じて賠償補償等の措置が取れるよう、利害関係者との合意内容を確認すること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	職員への教育	AIの判断に誤りが含まれることやAIの業務における位置づけを職員に研修等を通じて教育すること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIの利用方法や制限事項について、職員に研修等を通じて教育する機会を設けること。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	判断の責任	AIは補助的な役割を行うものであり、リスク等を認識したうえで最終的な判断は必ず職員が行う運用とすること		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	AIの出力結果が公平性を欠くことがないよう、AIに単独で判断させるだけでなく、最終的な判断は必ず職員が行う運用とすること。また、AIが市民等のユーザーに対して自動的な出力を行う場合は、その出力の傾向や誤答するリスクを把握したうえでユーザーに対して危害を与えないよう制衡し、AIがもたらす結果に責任を持つこと。		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

業務名		記入日	
システム名		記入者	

AIリスクアセスメント ver1.00

**評価基準**

1	2	3	4	その他
リスクの洗い出しまたはリスクに対する対応策の検討が不十分。	リスクに対して、対応策が検討されている。	リスクに対して、対応策が検討されており、一貫した対応を行うための体制が整備されている。	(運用開始後)リスクに対して、一貫した対応を組織的に実行できている。	該当しない

①～④の各観点ごとに採点を行い、全て2点以上であれば承認。一つでも1があれば却下となる。  
 ※ “該当しない”観点については、評価対象外とする。

採点 (デジタル戦略部記入)
----------------

**①人間中心**

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	詳細
1	人間中心	人間の尊厳を守りながら、AIの活用を促進し、価値の創出や社会課題の解決等を実現する。			
1.1a	人間の尊厳と個人の自律	AIが活用される際の社会的文脈を踏まえ、人間の尊厳及び個人の自律を尊重しているか。			
1.1b	人間の尊厳と個人の自律	<ul style="list-style-type: none"> <li>個人の権利・利益に重要な影響を及ぼす可能性のある分野においてAIを利用したプロファイリングを行っていないか。</li> <li>AIを利用したプロファイリングを行う場合、個人の尊厳を尊重し、アウトプットの正確性を可能な限り維持させつつ、AIの予測、推奨、判断等の限界を理解して利用し、かつ生じうる不利益等を慎重に検討した上で、不適切な目的に利用していないか。</li> </ul>			
1.2a	AIによる意思決定・感情の操作等への留意	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、又は意識的に知覚できないレベルでの操作を前提としたAIシステム・サービスの開発・提供・利用は行っていないか。 また、AIを人間の脳・身体と連携させる仕組みとなっていないか。			
1.2b	AIによる意思決定・感情の操作等への留意	AIシステム・サービスの開発・提供・利用において、自動化バイアス等のAIに過度に依存するリスクに注意を払い、必要な対策を講じているか。 また、フィルターバブル(※)に代表されるような情報又は価値観の傾斜を助長し、AI利用者を含む人間が本来得られるべき選択肢が不本意に制限されるようなAIの活用にも注意を払っているか。 ※ 過去のユーザー情報をもとに、各人に最適化されたコンテンツが表示されることで、似たような情報や視点に囲まれてしまう状態のこと			

採点 (デジタル戦略部記入)
----------------

**②影響範囲の特定**

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	詳細
2	影響範囲の特定	AIを使用する事務の内容と使用範囲を特定し、AIが誤作動した場合に生じる影響を把握すること。			
2.1a	業務内容とAIの使用範囲の特定	全体の業務内容のうち、どの部分でAIを使用し、それが最終決定にどう影響するか把握できているか。			
2.1b	業務内容とAIの使用範囲の特定 (利用拡大の想定)	AIの利用が今後拡大した場合、どれだけの期間に、どれだけの範囲に、どのような地理的範囲に拡大することを想定しているか。			
2.2.a	適正利用	<ul style="list-style-type: none"> <li>AIの出力について精度及びリスクの程度を理解し、様々なリスク要因を確認した上で利用しているか。</li> <li>AIが誤作動した場合に生じる影響を把握できているか。</li> </ul>			
2.2.b	適正利用	<ul style="list-style-type: none"> <li>AI提供者が定めたサービス規約、利用上の留意点があるか。(有の場合は資料を添付すること)</li> <li>AIシステム・サービスの利用はAI提供者が設計において想定したサービス規約・利用上の留意点を遵守した運用となっているか。</li> </ul>			
2.2.c	適正利用	正確性、必要な場合には最新性(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っているか。			
2.3a	文書化	AI提供者から提供されたAIシステム・サービスについての文書を定期的に確認し、更新があれば最新化する等、適切に保管・活用しているか。			

採点 (デジタル戦略部記入)
----------------

**③プライバシーの保護**

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	詳細
3	プライバシーの保護	学習データの収集、分析などにおいて、プライバシー保護に留意し、個人情報を慎重に取り扱うこと。			
3.1a	AIシステム・サービス全般における プライバシーの保護	AIシステム・サービスへの個人情報の不適切な入力を防止するためにどのような処理を講じているか。			
3.1b	AIシステム・サービス全般における プライバシーの保護	AIシステム・サービスにおけるプライバシー侵害に関して適宜情報収集し、防止を検討しているか。 (個人情報保護法に基づいた対応の確保、国際的な個人データ保護の原則及び基準の参照)			
以下、学習を行う場合					
3.2a	適正学習	<ul style="list-style-type: none"> <li>プライバシー・バイ・デザイン(※)等を通じて、学習時のデータについて適正に収集するとともに、第三者の個人情報・知的財産権に留意が必要なものが含まれている場合には、法令に従って適切に扱うことを、AIのライフサイクル全体を通じて確保しているか。</li> <li>市のデータを用いて学習する場合は市の取り扱いが適切であることを確認。</li> <li>事業者が提供するデータ等を用いて学習する場合はAI提供者の学習データの取扱いが適切であることを確認。</li> </ul> (※) システムの企画・設計段階から、個人情報やプライバシー保護を考慮するアプローチ手法			

1	人間中心 適正学習	人間の尊厳を守りながら、AIの活用を促進し、価値の創出や社会課題の解決等を実現する。	
3.2b		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しているか。	

採点 (デジタル戦略部記入)

④安全性の確保

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	
					詳細
4	安全性の確保	・ 正確性・必要な場合には最新性 (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こと。 ・ AIの出力について精度やリスクの程度を理解し、様々なリスク要因を確認した上で利用すること。			
4.1a	入力データの正確性および最新性	(2.2c再掲) 正確性・必要な場合には最新性 (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っているか。			
4.2a	人間の生命・身体・財産、精神及び環境への配慮	AIシステム・サービスの出力の正確性を含め、仕様書等で合意した要求水準に対して十分に動作していること (信頼性) をユーザテスト等で確認しているか。			
4.2b	人間の生命・身体・財産、精神及び環境への配慮	・ 様々な状況下でパフォーマンスレベルを維持し、無関係な事象に対して著しく誤った判断を発生させない堅牢性 (robustness) を持ったサービスを選定しているか。 ・ 運用上や技術的な課題についてフィードバックを行い、パフォーマンスレベルを維持するための仕組みが構築できているか。			
4.2c	人間の生命・身体・財産、精神及び環境への配慮	AIの活用又は意図しないAIの動作によって生じる権利侵害の重大性、侵害発生の可能性等、当該AIの性質・用途等に照らし、必要に応じて客観的なモニタリング及び対処も含めて人間がコントロールできる制御可能性を確保しているか。			
4.2d	人間の生命・身体・財産、精神及び環境への配慮	適切なリスク分析を実施し、リスクへの対策 (回避、低減、移転、容認) を講じているか。			
4.3a	最新動向への留意	AIシステム・サービスに対する外部からの攻撃は日々新たな手法が生まれており、これらのリスクに対応するための留意事項を確認しているか。(導入時点の対応とアップデート方針等の確認)			
4.4a	AIの基盤モデル等の更新への対応	AIの基盤モデルやライブラリ等の更新方針を確認し、リスクの把握と対策を講じているか。			
4.5a	各主体間・部門間の協力によるAIマネジメント強化	各主体のみでは解決できないAIシステム・サービスの運用上の課題及び解決に必要な情報を特定しているか。			
4.5b	各主体間・部門間の協力によるAIマネジメント強化	各主体間で、知的財産権、プライバシー等に留意しつつ、可能かつ合理的な範囲で共有しているか。			
4.6a	各主体間の不確実性への対応負担の分配	そもそもAIシステム・サービスの不確実性については、技術的に一定の対応は可能であるものの、完全に取り除くことが難しいという前提を提供者と神戸市間で認識共有を実施しているか。			

採点 (デジタル戦略部記入)

⑤透明性の確保

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	
					詳細
5	透明性の確保	・ AIの判断理由、根拠を提示できる仕組みとすること。 ・ 人が間違った判断や誤解をしないUI設計をすること。また、必要に応じてAIを使用していることを市民・事業者に提示すること。 ・ 事業者からの情報提供を求めよう努めること。また、情報が得られない場合には、リスク回避できる方法を必須とすること。			
5.1a	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、データ量又はデータ内容に照らし合理的な範囲で、AIシステム・サービスの開発過程、利用時の入出力等、AIの学習プロセス、推論過程、判断根拠等を提示できるサービスを選定しているか。 ・ 利用時の入出力等、AIの学習プロセス、推論過程、判断根拠等をログに記録・保存できるサービスを選定しているか。			
5.1b	検証可能性の確保	ログの記録・保存にあたっては、利用する技術の特性及び用途に照らして、事故の原因究明、再発防止策の検討、損害賠償責任要件の立証上の重要性等を踏まえて、記録方法、頻度、保存期間等について検討しているか。			
5.2a	偽情報等への対策	・ 生成AIによって、内容が真実・公平であるかのように装った情報を誰でも作ることができるようになり、AIが生成した偽情報・誤情報・偏向情報が社会を不安定化・混乱させるリスクが高まっていることを認識した上で、人が間違った判断や誤解をしないUI設計ができていないか。 ・ 必要に応じてAIを使用していることを市民・事業者に提示しているか。(AIによる出力を自動で表示する場合はAIによる出力であり、誤りを含む可能性があることを提示できているか。)			

採点 (デジタル戦略部記入)

⑥公平性の確保

番号	区分	要件	検討対象とするか (該当しない場合はその理由を記入)	回答	
					詳細
6	公平性の確保	・ 学習データ・アルゴリズムに差別や偏見につながるバイアスが発生するリスクを低減すること。 ・ 学習することにより新たに差別や偏見につながるバイアスが発生するリスクを認識し、対応すること。			
6.1a	AIモデルの各構成技術に含まれるバイアスへの配慮	公平性が担保されたデータの入力を行い、プロンプトに含まれるバイアスに留意して、責任をもってAI出力結果の事業利用判断を行っているか。			
6.1b	AIモデルの各構成技術に含まれるバイアスへの配慮	不適切なバイアスを生み出す要因は多岐に渡るため、各技術要素 (学習データ、モデルの学習過程、AI利用者又は業務外利用者が入力するプロンプト、AIモデルの推論時に参照する情報、連携する外部サービス等) 及びAI利用者の振る舞いを含めて、公平性の問題となり得るバイアスの要因となるポイントを特定しているか。 また、データに含まれるセンシティブ属性 (公平性の観点から排除すべき対象者の性別や人種等の個人の属性) を明確化しているか。			

1	人間中心	人間の尊厳を守りながら、AIの活用を促進し、価値の創出や社会課題の解決等を実現する。
6.1c	AIモデルの各構成技術に含まれるバイアスへの配慮	AIシステム・サービスの特性又は用途によっては、潜在的なバイアスが生じる可能性についても検討しているか。また、必要に応じて多様なステークホルダーから意見を聴取しているか。
以下、学習を行う場合		
6.2a	AIモデルの各構成技術に含まれるバイアスへの配慮	学習データ、AIモデルの学習過程によってバイアス(学習データには現れない潜在的なバイアスを含む)が含まれることに留意し、バイアスが発生するリスクを低減し、データの質を管理するための相当の措置を講じているか。
6.2b	AIモデルの各構成技術に含まれるバイアスへの配慮	学習データ、AIモデルの学習過程からバイアスを完全に排除できないことを踏まえ、AIモデルが代表的なデータセットで学習され、AIシステムに不公正なバイアスがないか点検されることを確保しているか。また、バイアスが発生するリスクを認識し、対応をとっているか。

採点(デジタル戦略部記入)
回答

⑦セキュリティの確保

番号	区分	要件	採点対象とするか (該当しない場合はその理由を記入)	回答 詳細
7	セキュリティの確保	AI特有の脆弱性に関するリスクを認識し、リスク回避のためにAI提供者によるセキュリティ上の留意点を遵守すること。		
7.1a	AIシステム・サービスに影響するセキュリティ対策	AI提供者によるセキュリティ上の留意点を遵守するために、マニュアル等に記載することで周知徹底しているか。		
7.1b	AIシステム・サービスに影響するセキュリティ対策	AIシステム・サービスに機密情報等を意図せず入力しない仕組みをシステム面・運用面でのように担保しているか。また、使用者に機密情報等を意図せず入力しないことを周知徹底しているか。		
7.1c	AIシステム・サービスに影響するセキュリティ対策	AIシステム・サービスに入力された情報が外部に漏えいしない仕組みをどのように担保しているか。(事業者による学習結果の再利用、監査のための事業者の閲覧による漏えいを含む)		記載例 ・ローカルLLMなどでデータの保有者を市とし、アクセス管理等で外部への漏えいを防ぐ。 ・グローバルLLMだがテナントを分けて、データの保有者を市とし、アクセス管理等で外部への漏えいを防ぐ。

採点(デジタル戦略部記入)
回答

⑧アカウントビリティの確保

番号	区分	要件	採点対象とするか (該当しない場合はその理由を記入)	回答 詳細
8	アカウントビリティの確保	・利害関係者を把握し、責任者を明確にしたうえで判断に関する正当な意味・理由の説明、必要に応じた賠償・補償等の措置が取れること。 ・AIの種類によって通報による再検討の対応方針を確認すること。		
8.1a	責任者の明示	各主体(※)においてアカウントビリティを果たす責任者を設定しているか。 ※ AI開発者、AI提供者、AI利用者を目指す。		
8.2a	関係者間の責任の分配	・関係者間の責任について、業務外利用者も含めた主体間の契約、社会的な約束(ボランティアコミットメント)等により、責任の所在を明確化しているか。 ・必要に応じて賠償・補償等の措置が取れるよう、利害関係者との合意内容を確認できているか。		
8.3a	トレーサビリティの向上	AIシステム・サービスに関する市の利用方針の決定過程について、合理的な範囲で追跡・遡求が可能な状態を確保しているか。		記載例 ・利用ルールについては、ドキュメント化し、保存するとともに利用ルールの決定過程の打ち合わせ等については、記録を残し、決定した理由や過程を振り返られるようにしている。
8.4a	関連するステークホルダーへの情報提供	AIの出力結果や出力の利用方法について、関連するステークホルダー(※)に合理的な範囲で説明できるようにしているか。 また、関連するステークホルダーの性質に応じて適正な利用方法を含む情報提供を平易かつアクセスしやすい形で行っているか。 ※ AI開発者、AI提供者、AI利用者及び業務外利用者以外の第三者を含むAIの活用によって直接・間接の影響を受ける可能性がある全ての主体		
8.4b	関連するステークホルダーへの情報提供	人間の生命・身体・財産、精神及び環境へ危害を及ぼす可能性がある場合は、講ずべき措置について事前に整理し、ステークホルダーに関連する情報を提供しているか。		
8.4c	関連するステークホルダーへの情報提供	関連するステークホルダーが講ずべき措置及び利用規則を明記しているか。		
8.4d	関連するステークホルダーへの情報提供	利用するAIシステム・サービスの性質に応じて、関連するステークホルダーからの通報・問合せに対応する窓口を合理的な範囲で設置し、AI提供者とも連携の上説明及び要望の受付を行っているか。		
8.4e	関連するステークホルダーへの情報提供	AIシステム・サービスの安全性やステークホルダーの利益を損なう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整えているか。		
8.4f	関連するステークホルダーへの情報提供	関連するステークホルダーから提供されるデータを用いる場合には、AIの特性及び用途、データ提供の手段、形式等について、あらかじめ当該ステークホルダーと合意しているか。		

採点(デジタル戦略部記入)
回答

⑨職員への教育

番号	区分	要件	採点対象とするか (該当しない場合はその理由を記入)	回答 詳細
9	職員への教育	AIの判断に誤りが含まれることやAIの業務における位置づけを職員に研修等を通じて教育すること。		
9.1a	教育・リスクリテラシー	・AIの利用方法や制限事項について、職員に研修等を通じて教育する機会を設けているか。 ・生成AIの活用拡大によって、AIと人間の作業の棲み分けが変わっていくことを想定し教育・リスクリテラシー等の内容が含まれているか。		
9.2a	AIリテラシーの確保	各主体内のAIに関わる者が、その関わりにおいて十分なレベルのAIリテラシーを確保するために必要な措置を講じているか。		

1	人間中心	人間の尊厳を守りながら、AIの活用を促進し、価値の創出や社会課題の解決等を実現する。		
9.3a	多様性・包摂性の確保	公平性の確保に加え、いわゆる「情報弱者」及び「技術弱者」を生じさせず、より多くの人々がAIの恩恵を享受できるよう社会的弱者によるAIの活用を容易にするよう注意を払っているか。		
9.4a	ステークホルダーへのフォローアップ	AIシステム・サービス全体の安全性を高めるため、必要に応じて、ステークホルダーに対して教育及びリテラシー向上のためのフォローアップを行っているか。		

採点 (デジタル戦略部記入)

⑩判断の責任

番号	区分	要件	回答	
			検討対象とするか (該当しない場合はその理由を記入)	詳細
10	判断の責任	AIは補助的な役割を行うものであり、リスク等を認識したうえで最終的な判断は必ず職員が行う運用とすること。		
10.1a	人間の判断の介入	AIの出力結果が公平性を欠くことがないよう、AIに単独で判断させるのではなく、最終的な判断は必ず職員が行う運用となっているか。		