

本チェックリストは、AI事業者ガイドライン「第2部C.共通の指針」を要約したものです。  
事業者に求められる重要な取組事項のチェックにご活用ください

※高度なAIシステムに係る事業者は、  
「チェックリスト[別添7 B.高度なAIシステムに係る事業者向け]」も実施ください

## チェック項目

- **人間中心**の考え方を基に、憲法が保障する又は国際的に認められた人権を侵すことがないようにしているか？
  - AIに関わる全ての者の生命・身体・財産、精神及び環境に危害を及ぼすことがないよう**安全性**を確保しているか？
  - 潜在的なバイアスをなくすよう留意し、それでも回避できないバイアスがあることを認識しつつ、回避できないバイアスが人権や多様な文化を尊重する**公平性**の観点から許容可能か評価しているか？
  - **プライバシー**を尊重・保護し、関係法令を遵守しているか？
  - AIの振る舞いが、外部的操作によって意図せぬ変更や停止が生じることをないように**セキュリティ**を確保しているか？
  - **透明性**を確保するために、AI自体やAIシステム・サービスの情報をステークホルダーに対し合理的で技術的に可能な範囲で提供しているか？
  - データの出所やAIの意思決定等のトレーサビリティに関する情報やリスクへの対応状況等について、関連するステークホルダーに対して合理的な範囲で**アカウントビリティ**を果たしているか？
  - **AIガバナンスやプライバシーに関するポリシー**等を策定しているか？
- 
- 上記の実現のため、各事業者の状況に応じた**具体的なアプローチ**は**検討**しているか？

検討には「**具体的なアプローチ検討のためのワークシート**」をご活用ください

## [高度なAIシステムに関する事業者向け]

本チェックリストは、AI事業者ガイドライン「第2部D. 高度なAIシステムに関する事業者  
に共通の指針」を要約したものです。

**高度なAIシステムに関する事業者**に該当する事業者に求められる重要な取組事項の  
チェックにご活用ください(高度なAIシステムを開発するAI開発者にのみ適用される内容もあるため、AI提供者及  
びAI利用者は適切な範囲でご活用ください)

## チェック項目

- 高度なAIシステムの市場導入前及び開発全体を通じて、AIライフサイクル全体にわたるリスクを特定、評価、軽減するための適切な措置を講じているか？
- 市場投入後に脆弱性、インシデント、悪用パターンを特定し、緩和しているか？
- 十分な透明性の確保やアカウントビリティの向上のため高度なAIシステムの能力、限界、適切・不適切な使用領域を公表しているか？
- 産業界、政府、市民社会、学界を含む関係組織間で、責任ある情報共有とインシデントの報告に努めているか？
- リスクベースのアプローチに基づくAIガバナンス及びリスク管理方針を策定、実施、開示しているか？
- AIのライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威を含む、強固なセキュリティ管理に投資し、実施しているか？
- 技術的に可能な場合は、AIが生成したコンテンツを識別できるように、電子透かしやその他の技術等、信頼性の高いコンテンツ認証及び来歴のメカニズムを開発、導入しているか？
- 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策に優先的に投資しているか？
- 気候変動、健康・教育などの世界の最大の課題に対処するため、高度なAIシステムの開発を優先しているか？
- 国際的な技術規格の開発を推進しているか？
- 適切なデータ入力措置と、個人データ及び知的財産の保護を実施しているか？
- 誤情報の拡散等のAI固有リスクに関するデジタルリテラシーの向上や脆弱性の検知への協力と情報共有等、高度な AI システムの信頼でき責任ある利用を促進し、貢献しているか？
- 上記の実現のため、各事業者の状況に応じた具体的なアプローチは検討しているか？

検討には「**具体的なアプローチ検討のためのワークシート**」をご活用ください

## 具体的なアプローチ検討のためのワークシート (共通の指針関連)

利用上の留意点	<p>ガイドラインに記載した内容に関して取り組むべき事項は、<b>各事業者の事業内容や置かれた状況等により、個々に異なります。</b></p> <p>このため、本ワークシートは、<b>あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの</b>であり、<b>各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと</b>を前提としたものです。</p> <p>したがって、必ずしも、全ての事項について、検討が必要となるものではないので、<b>活用の要否、各自の事情に応じた修正や取捨選択を検討ください。</b></p> <p>D列、E列の記載内容を基に、<b>F列以降を各事業者にてご検討いただき、各自が取り組むこと（あるいは内容）を具体化の上、活用ください。</b></p> <p>各自用カスタマイズし運用するには、α. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α.～γ.が重複する場合もある）。</p>
---------	--

α. ワークシートの作成者

β. 実施状況の確認を行う者

γ. 責任者

対応箇所	分類	検討にあたって重要な事項	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
1) 人間中心	1) 人間の尊厳と個人の自律 ①	<p>a. 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？</p> <p>b. 特に、AIを人間の脳・身体と連携させる場合には、その周辺技術に関する情報を踏まえつつ、諸外国及び研究機関における生命倫理の議論等を参照しているか？</p> <p>c. 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？</p>					
	1) AIによる意思決定・感情の操作等への留意 ②	<p>a. 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？</p> <p>b. 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？</p> <p>c. 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？</p> <p>d. 特に、選挙、コミュニティでの意思決定等をはじめとする社会に重大な影響を与える手続きに関連しうる場合においては、AIの出力について慎重に検討しているか？</p>					
	1) 偽情報等への対策 ③	<p>a. 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、必要な対策を講じているか？</p>					
	1) 多様性・包摂性の確保 ④	<p>a. 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に社会的弱者によるAIの活用を容易にするよう注意を払っているか？</p> <p>(ユニバーサルデザイン、アクセシビリティの確保、関連するステークホルダーへの教育・フォローアップ等)</p>					
	1) 利用者支援 ⑤	<p>a. 合理的な範囲で、AIシステム・サービスの機能及びその周辺技術に関する情報を提供し、選択の機会を適時適切に提供する機能が利用可能である状態としているか？</p> <p>(デフォルトの設定、理解しやすい選択肢の提示、フィードバックの提供、緊急時の警告、エラーへの対処等)</p>					
	1) 持続可能性の確保 ⑥	<p>a. AIシステム・サービスの開発・提供・利用において、ライフサイクル全体で、地球環境への影響を検討しているか？</p>					
2) 安全性	2) 人間の生命・心身・財産、及び環境への配慮 ①	<p>a. AIシステム・サービスの出力の正確性を高め、要求に対して十分に動作している（信頼性）ことを確認しているか？</p> <p>b. 様々な状況下でパフォーマンスを維持し、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ようにしているか？（堅牢性（robustness））</p> <p>c. AIの活用又は意図しないAIの動作によって生じう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？</p> <p>d. 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？</p> <p>e. 人間の生命・心身・財産、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？</p> <p>f. 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遵守する必要がある事項についても明記しているか？</p> <p>g. AIシステム・サービスの安全性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？</p>					
	2) 適正利用 ②	<p>a. 主体のコントロールが及ぶ範囲で本来の目的を逸脱した提供・利用により危害が発生することを避けるべく、AIシステム・サービスの開発・提供・利用を行っているか？</p>					
	2) 適正学習 ③	<p>a. AIシステム・サービスの特性及び用途を踏まえ、学習等に用いるデータの正確性・必要な場合には最新性（データが適切であること）等を検討しているか？</p> <p>b. 学習等に用いるデータの透明性の確保、法的枠組みの遵守、AIモデルの更新等を合理的な範囲で適切に実施しているか？</p>					

## 具体的なアプローチ検討のためのワークシート (共通の指針関連)

利用上の留意点	<p>ガイドラインに記載した内容に関して取り組むべき事項は、<b>各事業者の事業内容や置かれた状況等により、個々に異なります。</b></p> <p>このため、本ワークシートは、<b>あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの</b>であり、<b>各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと</b>を前提としたものです。したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、<b>活用の要否、各自の事情に応じた修正や取捨選択を検討ください。</b></p> <p>D列、E列の記載内容を基に、<b>F列以降を各事業者にてご検討いただき、各自が取り組むこと（あるいは内容）を具体化の上、活用ください。</b></p> <p>各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。</p>
---------	---

a. ワークシートの作成者	
β. 実施状況の確認を行う者	
γ. 責任者	

対応箇所	分類	検討にあたって重要な事項	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
3) 公平性	3) AIモデルの各構成技術に含まれるバイアスへの配慮 ①	a.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？ b.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？					
	3) 人間の判断の介入 ②	a.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、適切なタイミングで人間の判断を介入させる利用を検討しているか？ b.バイアスが生じていないか、AIシステム・サービスの目的、制約、要件、決定を明確かつ透明性のある方法により分析し、対処するためのプロセスを導入しているか？ c.無意識のバイアス及び潜在的なバイアスに留意し、多様な背景、文化、分野のステークホルダーと対話した上で、方針を決定しているか？					
4) プライバシー保護	4) AIシステム・サービス全般におけるプライバシーの保護 ①	a.個人情報保護法等の関連法令の遵守、各主体のプライバシーポリシーの策定・公表等により、社会的文脈及び人々の合理的な期待を踏まえ、ステークホルダーのプライバシーが尊重され、保護されるよう、その重要性に応じた対応を取っているか？ b.以下の事項を考慮しつつ、プライバシー保護のための対応策を取っているか？ (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)					
5) セキュリティ確保	5) AIシステム・サービスに影響するセキュリティ対策 ①	a.AIシステム・サービスの機密性・完全性・可用性を維持し、常時、AIの安全な活用を確保するため、その時点での技術水準に照らして合理的な対策を講じているか？ b.AIシステム・サービスの特性を理解し、正常な稼働に必要なシステム間の接続が適切に行われているかを検討しているか？ c.推論対象データに微細な情報を混入させることで関連するステークホルダーの意図しない判断が行われる可能性を踏まえて、AIシステム・サービスの脆弱性を完全に排除することはできないことを認識しているか？					
	5) 最新動向への留意 ②	a.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？					
6) 透明性	6) 検証可能性の確保 ①	a.AIの判断にかかわる検証可能性を確保するため、データ量及び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、判断根拠等のログを記録・保存しているか？ b.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？					
	6) 関連するステークホルダーへの情報提供 ②	a.AIとの関係の仕方、AIの性質、目的等に照らして、それぞれが有する知識及び能力に応じ、例えば、下記について取りまとめた情報の提供及び説明を行っているか？ ・AIシステム・サービス全般：AIを利用しているという事実、活用している範囲、データ収集及びアナレーションの手法、学習及び評価の手法、基盤としているAIモデルに関する情報、AIシステム・サービスの能力、限界、提供先における適切/不適切な利用方法、AIシステム・サービスの提供先、AI利用者が所在する国・地域等において適用される関連法令等 b.多様なステークホルダーとの対話を通じて積極的な関与を促し、社会的な影響及び安全性に関する様々な意見を収集しているか？ c.実態に即して、AIシステム・サービスを提供・利用することの優位性、それに伴うリスク等を、関連するステークホルダーに示しているか？					
	6) 合理的かつ誠実な対応 ③	a.「関連するステークホルダーへの情報提供」は、アルゴリズム又はソースコードの開示を想定するものではなく、プライバシー及び営業秘密を尊重して、採用する技術の特性及び用途に照らし、社会的合理性が認められる範囲で実施しているか？ b.公開されている技術を用いる際には、それぞれ定められている規程に準拠しているか？ c.開発したAIシステムのオープンソース化にあっても、社会的な影響を検討しているか？					
	6) 関連するステークホルダーへの説明可能性・解釈可能性の向上 ④	a.関連するステークホルダーの納得感及び安心感の獲得、また、そのためのAIの動作に対する証拠の提示等を目的として、説明する主体がどのような説明が求められるかを分析・把握できるよう、説明を受ける主体がどのような説明が必要かを共有し、必要な対応を講じているか？ (AI提供者：AI開発者に、どのような説明が必要となるかを共有しているか？) (AI利用者：AI開発者・AI提供者に、どのような説明が必要となるかを共有しているか？)					



## 具体的なアプローチ検討のためのワークシート (共通の指針関連)

利用上の留意点	<p>ガイドラインに記載した内容に関して取り組むべき事項は、<b>各事業者の事業内容や置かれた状況等により、個々に異なります。</b></p> <p>このため、本ワークシートは、<b>あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの</b>であり、<b>各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと</b>を前提としたものです。したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、<b>活用の要否、各自の事情に応じた修正や取捨選択を検討ください。</b></p> <p>D列、E列の記載内容を基に、<b>F列以降を各事業者にてご検討いただき、各自が取り組むこと（あるいは内容）を具体化の上、活用ください。</b></p> <p>各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。</p>
---------	---

a. ワークシートの作成者	
β. 実施状況の確認を行う者	
γ. 責任者	

対応箇所	分類	検討にあたって重要な事項	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
7) アカウンタビリティ	7) トレーサビリティの向上 ①	a. データの出所、AIシステム・サービスの開発・提供・利用中に行われた意思決定等について、技術的に可能かつ合理的な範囲で追跡・選取が可能状態を確保しているか？					
	7) 共通の指針の対応状況の ② 説明	a. 共通の指針の対応状況について、ステークホルダー（サプライヤーを含む）に対してそれぞれが有する知識及び能力に応じ、例えば以下の事項を取りまとめた情報の提供及び説明を定期的に行っているか？ ・全般：共通の指針の実践を妨げるリスクの有無及び程度に関する評価、共通の指針の実践の進捗状況 ・「人間中心」関連：偽情報等への留意、多様性・包摂性、利用者支援、持続可能性の確保の対応状況 ・「安全性」関連：AIシステム・サービスに関する既知のリスク及び対応策、並びに安全性確保の仕組み ・「公平性」関連：AIモデルを構成する各技術要素（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と想定するプロンプト、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）によってバイアスが含まれること ・「プライバシー保護」関連：AIシステム・サービスにより自己又はステークホルダーのプライバシーが侵害されるリスク及び対応策、並びにプライバシー侵害が発生した場合に講ずることが期待される措置 ・「セキュリティ確保」関連：AIシステム・サービスの相互間連携又は他システムとの連携が発生する場合、その促進のために必要な標準準拠等。AIシステム・サービスがインターネットを通じて他のAIシステム・サービス等と連携する場合に発生するリスク及びその対応策					
	7) 責任者の明示 ③	a. 各主体においてアカウンタビリティを果たす責任者を設定しているか？					
	7) 関係者間の責任の分配 ④	a. 関係者間の責任について、業務外利用者も含めた各主体間の契約、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？					
	7) ステークホルダーへの具体的な ⑤ 対応	a. 必要に応じ、AIシステム・サービスの利用に伴うリスク管理、安全性確保のための各主体のAIガバナンスに関するポリシー、プライバシーポリシー等の方針を策定し、公表しているか？（社会及び一般市民に対するビジョンの共有、情報発信・提供を行うといった社会的責任を含む） b. 必要に応じ、AIの出力の誤り等について、ステークホルダーからの指摘を受け付ける機会を設けるとともに、客観的なモニタリングを実施しているか？ c. ステークホルダーの利益を損なう事態が生じた場合、どのように対応するか方針を策定してこれを着実に実施し、進捗状況については必要に応じて定期的にステークホルダーに報告しているか？					
	7) 文書化 ⑥	a. 上記に関する情報を文書化して一定期間保管し、必要に応じて、必要などころで、入手可能かつ利用に適した形で参照可能な状態としているか？					
8) 教育・リテラシー	8) AIリテラシーの確保 ①	a. 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？					
	8) 教育・リスキニング ②	a. 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と想定されるため、新たな働き方ができるよう教育・リスキニング等を検討しているか？ b. 様々な人がAIで得られる便益の理解を深め、リスクに対するレジリエンスを高められるよう、世代間ギャップも考慮した上での教育の機会を提供しているか？					
	8) ステークホルダーへのフォロー ③ アップ	a. 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確保のためのフォローアップを行っているか？					
9) 公正競争確保	9) 公正競争確保	a. AIを活用した新たなビジネス・サービスが創出され、持続的な経済成長の維持及び社会課題の解決策の提示がなされるよう、AIをめぐる公正な競争環境の維持に努めているか？					
10) イノベーション	10) オープンイノベーション等の推 ① 進	a. 国際化・多様化、産学官連携及びオープンイノベーションを推進しているか？ b. AIのイノベーションに必要なデータが創出される環境の維持に配慮しているか？					
	10) 相互接続性・相互運用性 ② への留意	a. 自らのAIシステム・サービスと他のAIシステム・サービスとの相互接続性及び相互運用性を確保しているか？ b. 標準仕様がある場合には、それに準拠しているか？					
	10) 適切な情報提供口 ③	a. 自らのイノベーションを損なわない範囲で必要な情報提供を行っているか？					

## 具体的なアプローチ検討のためのワークシート（「ガバナンスの構築」関連）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 各自用カスタマイズし運用するには、α、取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β、実施状況の確認を行う者（現場において実際の確認を行う者）γ、責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γが重複する場合もある）。

α. ワークシートの作成者
β. 実施状況の確認を行う者
γ. 責任者

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
1. 環境・リスク分析	1-1 便益/リスクの理解	a. 事業における価値の創出、社会課題の解決等のAIの開発・提供・利用の目的を明確に定義しているか？ b. 自社の事業に結びつくかたちで、「便益」及び意図せざるものを含めた「リスク」の性質を具体的に理解しているか？ c. その際に、回避すべき「リスク」及び複数主体にまたがる論点に留意し、バリューチェーン/リスクチェーン全体での便益を確保、リスクを削減しているか？ d. 迅速に「経営層に報告/共有する」仕組みを構築しているか？	別添2.					
	1-2 AIの社会的な受容の理解	a. ステークホルダーを特定しているか？ b. 特定した上で、「社会的な受容」の理解に努め、AIを開発・提供・利用しているか？ c. 提供開始後も、急速に変化する外部環境を考慮し、必要に応じて適時にステークホルダーの意見を再確認しているか？	別添2.					
	1-3 自社のAI習熟度の理解	a. 各主体の事業領域及び規模等に照らしてAI習熟度の評価の必要性を検討しているか？ b. 必要であると判断した場合、AIのリスクへの対応力を見える化し、AI習熟度（AIシステム・サービスの開発・提供・利用時に求められる準備がどれだけできているのか）を評価したか？また、可能であれば、合理的な範囲でその結果をステークホルダーに公開しているか？ c. 必要ではないと判断した場合、可能であれば、合理的な範囲でその事実を、理由とともにステークホルダーに公開しているか？	別添2.					
2. ゴール設定	2-1 AIガバナンス・ゴールの設定	a. 各主体の「AIガバナンス・ゴール」を設定するかを検討しているか？ b. 必要であると判断した場合、ゴールを設定したか？また、可能であれば、合理的な範囲で当該ゴールをステークホルダーに対して公開しているか？ c. 必要ではないと判断した場合、可能であれば、合理的な範囲でその事実を、理由とともにステークホルダーに公開しているか？	別添2.					
3. システムデザイン	3-1 ゴール及び乖離の評価、並びに乖離対応の必須化	a. 現状のAIシステム・サービス及び「AIガバナンス・ゴール」からの乖離を特定・評価しているか？ b. リスクが認められる場合、その受容の合理性の有無を判定しているか？ c. 受容に合理性が認められない場合、開発・提供・利用の在り方を再考/再考するためのプロセスを開発・提供・利用の適切な段階及び各主体の組織における意思決定プロセスへ組み込んでいるか？ d. 上記を経営層がリーダーシップを取って、その意思決定に責任を持ち、運営層が具体化した上で、継続的に実施しているか？ e. 各主体内での認識の醸成を行うため、決定した乖離評価項目を各主体内で共有したか？提供するAIの内容に応じて、各主体間で連携をして乖離評価を実施しているか？	別添2.					
	3-1-1 業界の標準的な乖離評価プロセスとの整合性の確保	a. 自社の知見・経験に留まらず、業界における標準的な乖離評価プロセス、他社・団体の取り組み等、外部のベストプラクティスを積極的に取り込んでいるか？	別添2.					
	3-1-2 利用者及び業務外利用者に対する、乖離可能性/対応策に関する十分な情報提供	a. AIシステム・サービス及び「AIガバナンス・ゴール」の間に乖離が発生しうる場合は、ステークホルダーにその事実及び対応策に関し情報提供をし、問い合わせへの対応等を通じたコミュニケーションを図っているか？ b. 情報提供の効果を高めるために、開発主体及び業界団体等とも連携しながら、様々な情報発信を通じて、AI利用者及び利害関係者等のリテラシーの向上にも貢献しているか？ c. 乖離によって生じるリスクの性質及び蓋然性の大きさに応じて、情報提供の程度を検討しているか？	別添2.					
3-2 AIマネジメントシステムの人材リテラシー向上	a. 外部講師によるものを含め、役職及び担当に適した研修並びに教材を用い、AIリテラシーの向上を図っているか？ b. その際、各者の果たすべき役割に応じて適した研修及び教材を活用しているか？ c. 特に重要となるAI倫理については全社員に受講させる等の工夫をしているか？ d. 今般の生成AIに関する動向等を踏まえ、生成AI技術及び出力結果の信頼性に関する研修を行う等の工夫をしているか？ e. 行動目標5-1のAIマネジメントシステムの設計及び運用から独立した関連する専門性を有する者による評価を自社で行う場合に、そのような専門性を社員が習得できるような配慮をしているか？	別添2.						

## 具体的なアプローチ検討のためのワークシート（「ガバナンスの構築」関連）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 各自用にカスタマイズし運用するには、α、取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β、実施状況の確認を行う者（現場において実際の確認を行う者）γ、責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γが重複する場合もある）。

- α. ワークシートの作成者
- β. 実施状況の確認を行う者
- γ. 責任者

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
3-3	各主体間・部門間の協力によるAIマネジメント強化	a. 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？ b. 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？  ※ 各種法令・規制、各主体のAIポリシー、営業秘密、限定提供データ等、公正競争確保が前提	別添2.					
3-3-1	各主体間の情報共有による現状理解	a. AIシステムの開発に用いたデータの取得源/データの量・質、分布、カテゴリー毎の概要等の情報を共有しているか？ b. 共有の際、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」等の情報共有の標準化に向けた取組を参照しているか？	別添2.					
3-3-2	環境・リスク分析のための日常的な情報収集・意見交換の奨励	a. 日常的な、ルール整備、ベストプラクティス、インシデント等の情報収集を行っているか？ b. AIマネジメントチームを社内を設置している場合であっても、社内の他部門との議論及び勉強会を開催したり、他社も参加する団体活動に関与しているか？	別添2.					
3-4	予防・早期対応による利用者のインシデント関連の負担軽減	a. システム障害、情報漏洩、クレームの発生等のインシデントの予防及び早期対応を行っているか？ b. ライフサイクル全体を通じてインシデントを予防し、又は発生した場合に早期に対応できる体制を構築しているか？	別添2.					
3-4-1	各主体間の不確実性への対応負担の分配	a. 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はあるものの、完全に排除することが難しいという前提を認識共有を実施しているか？ b. その上で、各主体間での責任所在を可能かつ合理的な範囲で明確化しているか？	別添2.					
3-4-2	インシデント発生時の対応の事前検討	a. AIインシデント発生時の対応方針及び計画を策定しているか？ b. 上記について、適宜実践的な予行演習を実施しているか？ c. AIインシデント発生時に備え以下のような体制をあらかじめ整備しているか？ ☐ 連絡受付窓口の設置 ☐ 対応担当役員のアサイン ☐ 対応担当者個々の役割分担 ☐ 対応アプローチ・プロセス ☐ IT管理部門等の社内関係者への連絡体制 ☐ 顧問弁護士等の社外関係者や専門家への連絡体制 ☐ 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	別添2.					

## 具体的なアプローチ検討のためのワークシート（「ガバナンスの構築」関連）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、各事業者の事業内容や置かれた状況等により、個々に異なります。  
 このため、本ワークシートは、あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するものであり、各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくことを前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、活用の要否、各自の事情に応じた修正や取捨選択を検討ください。  
 D列～F列の記載内容を基に、G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。  
 各自用カスタマイズし運用するには、α、取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β、実施状況の確認を行う者（現場において実際の確認を行う者）γ、責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γが重複する場合もある）。

- α. ワークシートの作成者
- β. 実施状況の確認を行う者
- γ. 責任者

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
4. 運用	4-1 AIマネジメントシステム運用状況の説明可能な状態の確保	a. 適切かつ合理的な範囲で、AIマネジメントシステムの運用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況になっているか？ b. AIマネジメントシステムの運用状況のアカウントビリティを高めるため、下記のような取組を行っているか？ □行動目標3-1の乖離評価プロセスの実施状況の記録 □AIシステム・サービス開発・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 c. 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・記録しているか？(検討にあたってはチェックリストを参照し、カスタマイズすることも有用)	別添2.					
	4-2 個々のAIシステム運用状況の説明可能な状態の確保	a. 各主体のAIの運用の状況をモニタリングし、PDCAを回しながら、結果を記録しているか？ b. 各主体が単独で対応することが難しい場合には、各主体間で連携しているか？	別添2.					
	4-3 AIガバナンスの実践状況の積極的な開示の検討	a. 自社のAIに対する基本的な考え方や、AIマネジメントシステムの整備・運用等まで、AIガバナンスに関する情報の透明性の確保を検討しているか？ b. 開示する場合には、コーポレートガバナンス・コードの非財務情報に位置づけること等も検討しているか？ c. 開示しない場合には、その事実をその理由とともにステークホルダーに公開しているか？	別添2.					
5. 評価	5-1 AIマネジメントシステムの機能の検証	a. 継続的改善に向けた評価の重点ポイントを、経営層が自らの言葉で明示しているか？ b. AIマネジメントシステムの設計及び運用から独立した関連する専門性を有する者を割り当てているか？ c. 上記の者によるAIマネジメントシステムが適切に機能しているか否かのモニタリングを行っているか？ d. モニタリングの結果をもとに、継続的な改善を実施しているか？	別添2.					
	5-2 社外ステークホルダーの意見の検討	a. 様々なステークホルダーから、AIマネジメントシステム及びその運用に対する意見を求めることを検討しているか？ b. 当該意見の内容実施しない場合、その理由をステークホルダーに説明しているか？	別添2.					
6. 環境・リスクの再分析	6-1 1-1～1-3の適時の再実施	a. 新技術の出現、AIに関連した技術革新、規制等の社会的制度的変更等の外部環境の変化を把握しているか？ b. 適時に再評価、理解の更新、新たな視点の獲得等を行い、それに即したAIシステムの改善、再構築、運用の変更等を行っているか？ c. AIガバナンスの考え方を組織の文化として根付かせているか？	別添2.					



## 具体的なアプローチ検討のためのワークシート（「AI開発者」関連 ※ 記載例あり）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。  
 （エントリーシートの文章で、応募者に対し合否を判断する採用AI（XGBoost）の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf）  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	開発部門品質管理担当者XXX（AI倫理委員会XXXにより実施をサポート）
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項（該当しない場合はその理由）	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日（見直し日）
データ前処理・学習時に確認する事項	2)③ 適正学習	学習時のデータについて、 a. 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等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？ b. 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？	3部 2) 安全性 4) プライバシー保護 5) セキュリティ確保	a. 学習データ内に個人を特定する情報が含まれている可能性あり b. 学習データの保護措置によっては出力内容が異なる可能性あり	-	✓	a. ・学習データに用いるエントリーシートについて、学習時に個人を特定する情報をクレンジングする（SpaCy & Presidioでの検出 + AI開発者による検証） ※モデルの汎化性能を確保する上でも個人を特定する情報を除く  ・エントリーシート以外の文書データが含まれていないか、各文書ベクトルの分散から外れ値を定期的にレビューし、誤ったデータを除く ・学習データを適切な場所に格納する ・開発環境について適切なAI開発者のみ適切なレベルのアクセス権限を付与する b. ・データドリフト(学習すべきデータの変化)の定期的な検証を実施する ・AIモデルの正確性の定期的な検証を実施する ・アプリケーションにおける異常値の検証を実施する	
	3)① 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、モデルの学習過程において、 a. 学習データ、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、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？ b. 学習データ、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？	3部 3) 公平性	a. b. ・応募者の属性(特定の国/地域/人種/性別/年齢)による不公平な予測結果が含まれる可能性あり ・学習データ不足等によるバイアスが含まれる可能性あり(例：外国籍の応募者、障害者雇用の対象者、従来重視していなかった職種等)	-	✓	a. b. ・AI提供者が検討したリスク対策の計画(リスクチェーンモデルの内容)の検討に協力し、技術側で行うべきバイアスの評価を行う（データ及びモデルの特徴量）  ・AIモデル開発・アップデート時に学習データにおける著しい偏り(特定の国/地域/人種/性別/年齢)がないかを評価する ・AIモデルの学習時に重要性の高い特徴量をレビューして著しい偏り(特定の国/地域/人種/性別/年齢)がないかを評価する ・著しい学習データの不足やモデル評価への影響にAI開発者側で対応できないバイアス(例：外国籍の応募者、障害者雇用の対象者、従来重視していなかった職種等)について、AI提供者及びAI利用者に説明し業務側でカバーできるか合意を取る	

# 具体的なアプローチ検討のためのワークシート（「AI開発者」関連 ※ 記載例あり）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、あくまで**各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではないので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。  
 （エントリーシートの文章で、応募者に対し合否を判断する採用AI（XGBoost）の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf）  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	開発部門品質管理担当者XXX（AI倫理委員会XXXにより実施をサポート）
γ. 責任者	AI倫理委員会担当者XXX

AI開発時に確認する事項	2)①	人間の生命・心身・財産、及び環境への配慮	開発時において、 a. 関連するステークホルダーを明確にしているか？ b. ステークホルダーの生命・心身・財産、精神及び環境に危害を及ぼすことがないよう、様々な状況下で予想される利用条件下でのパフォーマンスだけでなく、予期しない環境での利用にも耐えうる性能の要求を検討しているか？ c. ステークホルダーの生命・心身・財産、精神及び環境に危害を及ぼすことがないよう、リスク(運動するロボットの制御不能、不適切な出力等)を最小限に抑える方法(ガードレール技術等)を検討しているか？	3部 2) 安全性	a. ステークホルダーはAI提供者、採用担当者、応募者 b. 採用AIの利用環境においては、AIの出力結果が異なる可能性あり（予測結果や堅牢性） c. 検討の必要あり	AI提供者と、AI利用者からのエスカレーションの観点で連携の必要性あり	b. ・AIモデルの予測精度を評価する：分類の正確性  ・AIモデルの堅牢性を評価する：句読点等のエントリーシートにおいて重要性が低い情報の違いによって、予測性能が著しく異なるか c. ・様々なグループ会社・職種・スキル要件を想定したバリエーションでAIモデルの予測結果と判断根拠を検証する ・(AI利用者の最終判断に向けて)AIモデルから判断根拠情報及び類似する判断事例を出力する ・本AIサービスに関わるリスクシナリオ及び技術的な対策について、関連するステークホルダーと共に検討する ・AI利用者からAIの予測性能に関わるエスカレーションがあった際にモデル・データを検証できる体制を作る ・モデルの再学習を実施できる体制・環境を作る
	2)②	適正利用	a. 関連するステークホルダーを明確にしているか？ b. 開発時に想定していないAIの提供・利用により危害が発生することを選けるため、AIを安全に利用可能な使い方について明確な方針・ガイダンスを設定しているか？ c. 事前学習済のAIモデルに対する事後学習を行う場合に、学習済AIモデルを適切に選択（商用利用可能なライセンスかどうか、事前学習データ、学習・実行に必要なスペック等）しているか？	3部 2) 安全性	a. ステークホルダーはAI提供者、採用担当者、応募者 b. 応募者の入力内容や採用提供者の使い方によりAIの出力結果が異なる可能性あり c. 通常の機械学習モデルによるケースのためAIモデルの調整は対象外	AI提供者と、AI利用者からのエスカレーションの観点で連携の必要性があり	b. ・予測対象とする文章の仕様を明確化する(対象：エントリーシート、言語：日本語、文字数：2000字以内)  ・利用ログを保存する ・AI提供者が定義したAIサービスの要求仕様に応じて、AIモデルの目的関数及びユースケースの仕様を具体化する
	3)①	AIモデルの各構成技術に含まれるバイアスへの配慮	a. AIモデルを構成する各技術要素（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）によってバイアスが含まれることまで検討しているか？ b. AIモデル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？	3部 3) 公平性	a. b. 他に外部情報の参照やユーザー入力によって推論性能に影響を与えることがないため対象外		
	5)①	AIシステム・サービスに影響するセキュリティ対策	a. 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適切にセキュリティ対策を講じているか？（セキュリティ・バイ・デザイン）	3部 5) セキュリティ確保	a. 入力データやシステム間の接続によってはAIシステムの品質が変わる可能性あり		a. ・AIシステムの開発において、必要なセキュリティ対策を検討する（エントリーシートのファイルを改ざんしないよう書き込み不可にする） ・AIシステムの開発において、必要なネットワーク要件を検討する ※イントラ内での利用のため、外部攻撃等のリスクについては会社の情報セキュリティ管理に依拠することで対応している
	6)①	検証可能性の確保	a. AIの予測性能又は出力の品質が、活用後に大きく変動する可能性又は想定する精度に達しないこともある特性を踏まえ、事後検証のための作業記録を保存しつつ、その品質の維持・向上を行っているか？	3部 2) 安全性 6) 透明性	a. 本ケースにおいては、2)③への対応と同様		

## 具体的なアプローチ検討のためのワークシート（「AI開発者」関連 ※ 記載例あり）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、あくまで**各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではないので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。  
 （エントリーシートの文章で、応募者に対し合否を判断する採用AI（XGBoost）の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf）  
 各自用カスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	開発部門品質管理担当者XXX（AI倫理委員会XXXにより実施をサポート）
γ. 責任者	AI倫理委員会担当者XXX

AI開発後に確認する事項	5)②	最新動向への留意	a. 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、開発の各工程で留意 すべき点を確認しているか？	3部 5) セキュリティ確保	a. 開発されたシステムはイントラ内での利用が想 定されるため、開発者としてはデータの保護措置 を実施。外部攻撃等のリスクについては、各主体 (提供者/利用者)の情報セキュリティ管理に依 拠	-	
	6)②	関連するステークホルダーへの 情報提供	a. 関連するステークホルダーを明確にしているか？  また、自らの開発するAIシステムについて、例えば以下の事 項を適時かつ適切に関連するステークホルダーに（AI提供者 を通じて行う場合を含む）説明できるようにしたか？ b. 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 c. AIシステムの技術的特性、安全性確保の仕組み、利用 の結果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 d. 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 e. 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 f. 更新を行った場合の内容及びその理由の情報 g. AIモデルで学習するデータの収集ポリシー、その学習方法 及び実施体制等	3部 1)人間中心 2)安全性 3)公平性 4)プライバシー保護 5)セキュリティ確保 6)透明性	a. 関連するステークホルダーはAI提供者、採用 担当者、応募者 b. 出力又はプログラムの変化の可能性あり c. 当該情報は提供の必要あり d. 利用範囲はAI提供者で検討するため対象外 e. f. g. 当該情報は提供の必要あり	AI提供者と連携の必 要性あり	b. f. AIシステムの出力内容・情報の変更につい てはモデルカードに記録する  c. リスクシナリオ及び対策の検討内容を整理する  e. AIシステムの障害発生状況に対応方法及び併 せて記録する g. ・学習データを適切な場所に格納する ・開発環境について適切なAI開発者にもみ適切 なレベルのアクセス権限を付与する (D-2) iと同様) AIシステム・サービス全般に係る以下の情報をAI 提供者に向けて提供する -データ収集及びアナテーションの手法 -トレーニング及び評価の手法 -基盤としているAIモデルに関する情報 -AIシステム・サービスの能力、限界(予測の苦手 な対象等)
	7)②	共通の指針の対応状況の 説明	a. AI提供者に対して、AIには活用後に予測性能又は出力 の品質が大きく変動する可能性、想定する精度に達しないこ ともある特性がある旨、その結果生じうるリスク等の情報提供 及び説明を行っているか？ b. 具体的には、AIモデルを構成する各技術要素（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 入力すると想定するプロンプト、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）において含まれる可能性が あるバイアスへの対応等を周知しているか？	3部 7) アカウンタビリティ	a. b. 想定されるリスクの認識を、AI提供者と合 わせる必要あり	AI提供者と連携の必 要性あり	a. b. ・リスクシナリオ及び対策の検討過程において、ス テークホルダーへ説明すべき内容を整理し、説明 する
7)⑥	文書化	a. トレーサビリティ及び透明性の向上のため、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等について、可能な限り第三者が 検証できるような形で文書化しているか？	3部 7) アカウンタビリティ	a. AIシステムの開発過程やアーキテクチャ、検討 内容等の文書化により、トレーサビリティと透明性 を確保の必要あり	-	a. ・AIシステムのアーキテクチャを文書化する（学習 データ、本番データの入力、AIシステムの推論、 出力）  ・AIシステムの開発に伴う情報を文書化する ・AIシステムの検討内容はモデルカードに記録す る	

## 具体的なアプローチ検討のためのワークシート（「AI開発者」関連 ※ 記載例あり）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではないため、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、**G列以降に、オレンジ色にて記載例を示しています。**  
 （エントリーシートの文章で、応募者に対し合否を判断する採用AI（XGBoost）の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf）  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	開発部門品質管理担当者XXX（AI倫理委員会XXXにより実施をサポート）
γ. 責任者	AI倫理委員会担当者XXX

その他の重要な事項	1)①	人間の尊厳及び個人の自律の尊重	a. 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？ b. 特に、AIを人間の脳・身体と連携させる場合には、その周辺技術に関する情報を踏まえつつ、諸外国及び研究機関における生命倫理の議論等を参照しているか？ c. 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？	2部 1) 人間中心	a. AIシステムの活用方法によっては人間の尊厳と個人の自律が侵害される可能性あり b. 脳・身体と連携するケースではないため対象外 c. AIの予測結果に個人情報が含まれているため、活用方法によっては個人の権利・利益、プライバシーや公平性、個人の尊厳が侵害される可能性あり	-	a. AIシステムの開発において、学習データの収集やラベリング、モデルの性能評価等は、AI開発者だけで完結せず、AI提供者側で判断できるようにする
	1)②	AIによる意思決定・感情の操作等への留意	a. 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？ b. 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対して注意を払い、必要な対策を講じているか？ c. 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？ d. 特に、選挙、コミュニティでの意思決定等をはじめとする社会に重大な影響を与える手続きに関連する場合においては、AIの出力について慎重に取り扱っているか？	2部 1) 人間中心	a. b. AIシステムの活用方法によっては人間の意思決定を左右したり、自動化バイアスなどを生じたりする可能性あり c. 応募者が入力するエントリーシートは企業が求める要件に沿った上で自由記述できるものであり、AIサービスが制限するものではないため対象外 d. 本ケースは、社会的な意思決定に関わるものではないため、対象外	-	a. b. (AI利用者の最終判断に向けて)AIモデルから判断根拠情報及び類似する判断事例を出力する
	1)③	偽情報等への対策	a. 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、必要な対策を講じているか？	2部 1) 人間中心	a. AIが出力を生成する活用方法ではなく、社会的な意思決定に関わるものでもないため対象外	-	
	1)④	多様性・包摂性の確保	a. 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に社会的弱者によるAIの活用を容易にするよう注意を払っているか？	2部 1) 人間中心	a. 社内の人材採用担当者がいるため対象外 ・AI利用者が、障がい者雇用のプログラム等、電子ファイルによるエントリーシートを作成するための情報リテラシーが十分ではない人材採用のチャネルについて、本AIサービスとは別に人事業務として用意	-	
	1)⑤	利用者支援	a. 合理的な範囲で、AIシステム・サービスの機能及びその周辺技術に関する情報を提供し、選択の機会の判断のための情報を適時かつ適切に提供する機能が利用可能である状態としているか？	2部 1) 人間中心	a. AI利用者である採用担当者にとって、機能の選択の機会の提供が必要	-	a. (AI利用者の最終判断に向けて)AIモデルから判断根拠情報及び類似する判断事例を出力する
	1)⑥	持続可能性の確保	a. AIシステム・サービスの開発・提供・利用において、ライフサイクル全体で、地球環境への影響も検討しているか？	2部 1) 人間中心	a. AIシステムの活用により、エネルギー消費等、環境負荷は一定発生	-	a. 本番運用及び再学習を含めて、開発環境における計算量・稼働時間・ストレージ使用量を記録する
	4)①	AIシステム・サービス全般におけるプライバシーの保護	a. 個人情報保護法等の関連法令の遵守、各主体のプライバシーポリシーの策定・公表等により社会的文脈及び人々の合理的な期待を踏まえ、ステークホルダーのプライバシーが尊重され、保護されるよう、その重要性に応じた対応を取っているか？	2部 4) プライバシー保護	a. 個人情報が含まれる学習データのアクセス権限範囲やその扱いによってはステークホルダーのプライバシーが侵害される可能性あり	-	



## 具体的なアプローチ検討のためのワークシート（「AI開発者」関連 ※ 記載例あり）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、各事業者の事業内容や置かれた状況等により、個々に異なります。このため、本ワークシートは、あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するものであり、各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくことを前提としたものです。したがって、必ずしも、全ての事項について、検討が必要となるものではないので、活用の要否、各自の事情に応じた修正や取捨選択を検討ください。D列～F列の記載内容を基に、G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。（エントリーシートの文章で、応募者に対し合否を判断する採用AI（XGBoost）の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf）各自用カスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

α. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	開発部門品質管理担当者XXX（AI倫理委員会XXXにより実施をサポート）
γ. 責任者	AI倫理委員会担当者XXX

6)③	合理的かつ誠実な対応	a. 「関連するステークホルダーへの情報提供」は、アルゴリズム又はソースコードの開示を想定するのではなく、プライバシー及び営業秘密を尊重して、採用する技術の特性及び用途に照らし、社会的合理性が認められる範囲で実施しているか？ b. 公開されている技術を用いる際には、それぞれ定められている規程に準拠しているか？ c. 開発したAIシステムのオープンソース化にあたっては、社会的な影響を検討しているか？	2部 6)透明性	-	a. 本ケースにおいては、技術情報を外部に開示することは求められていないため対象外 b. 公開されている情報を用いる際には、参照すべき規定がある場合あり c. 本ケースはオープンソース化を行わないため対象外	-	b. 開発に用いるオープンソースのライブラリを一覧化する
6)④	関連するステークホルダーへの説明可能性・解釈可能性の向上	a. 関連するステークホルダーの納得感及び安心感の獲得、また、そのためのAIの動作に対する証拠の提示等を目的として、説明する主体がどのような説明が求められるかを分析・把握できるよう、説明を受ける主体がどのような説明が必要かを共有し、必要な対応を講じているか？	2部 6)透明性	-	a. ステークホルダーであるAI提供者、採用担当者、応募者に対し、リスクとその対策の説明が必要になる可能性あり	AI提供者と連携の必要性あり	リスクシナリオ及び対策の検討過程において、ステークホルダーへ説明すべき内容を整理する
7)①	トレーサビリティの向上	a. データの出所、AIの開発・提供・利用中に行われた意思決定等について、技術的に可能かつ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？	2部 7)アカウントビリティ	-	a. ステークホルダーへの説明等のため、AIシステムの開発過程で行われた意思決定を振り返る必要が生じる可能性あり	-	a. ・AIモデル及びAIシステムの開発・変更における検討過程と修正・検証内容をモデルカードに記録する ・利用ログを保存する ・アプリケーションにおける異常値の検証を実施する
7)③	責任者の明示	a. 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？	2部 7)アカウントビリティ	-	a. アカウントビリティを果たす責任所在が明確である必要あり b. AI提供者が定める各ポリシーに準拠するため対象外	-	a. AI開発者の中でチーフデータサイエンティスト及びリードエンジニアを特定する
7)④	関係者間の責任の分配	a. 関係者間の責任について、業務外利用者も含めた各主体間の契約、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？	2部 7)アカウントビリティ	-	a. AI提供者での検討に含まれるため対象外	AI提供者と連携の必要性あり	
7)⑤	ステークホルダーへの具体的な対応	a. 必要に応じ、AIシステム・サービスの利用に伴うリスク管理、安全性確保のための各主体のAIガバナンスに関するポリシー、プライバシーポリシー等の方針を策定し、公表しているか？（社会及び一般市民に対するビジョンの共有、情報発信・提供を行うといった社会的責任を含む） b. 必要に応じ、AIの出力の誤り等について、ステークホルダーからの指摘を受け付ける機会を設けるとともに、客観的なモニタリングを実施しているか？ c. ステークホルダーの利益を損なう事態が生じた場合、どのように対応するか方針を策定してこれを着実に実施し、進捗状況については必要に応じて定期的にステークホルダーに報告しているか？	2部 7)アカウントビリティ	-	a. b. AIの出力に誤りが生じたり、ステークホルダーの利益を損なう事態が生じる可能性あり	-	a. b. ・AI利用者からAIの予測性能に関わるエスカレーションがあった際にモデル・データを検証できる体制にする  ・モデルの再学習を実施できる体制・環境を作る
8)①	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？	2部 8)教育・リテラシー	-	-	-	
8)②	教育・リスキリング	a. 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と想定されるため、新たな働き方ができるよう教育・リスキリング等を検討しているか？ b. 様々な人がAIで得られる便益の理解を深め、リスクに対するレジリエンスを高められるよう、世代間ギャップも考慮した上での教育の機会を提供しているか？	2部 8)教育・リテラシー	-	-	-	
8)③	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に必要なフォローアップを行っているか？	2部 8)教育・リテラシー	-	-	-	
10)①	イノベーションの機会の創造	可能な範囲で、以下を実施し、イノベーションの機会の創造に貢献しているか？ ・AIの品質・信頼性、開発の方法論等の研究開発 ・持続的な経済成長の維持及び社会課題の解決策の提示への貢献 ・DFFT等の国際議論の動向の参照、AI開発者コミュニティ又は学会への参加の取組等、国際化・多様化及び産学官連携の推進 ・社会全体へのAIに関する情報提供	2部 10)イノベーション	-	-	-	

# 具体的なアプローチ検討のためのワークシート (「AI提供者」関連 ※ 記載例あり)

**利用上の留意点** ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、**G列以降に、オレンジ字にて記載例を示しています。**  
 (エントリーシートの文章で、応募者に対し合否を判断する採用AI (XGBoost)の例。 [https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\\_Case01\\_Recruitment-AI\\_JP.pdf](https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case01_Recruitment-AI_JP.pdf))  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	人材採用部門XXX (AI倫理委員会XXXにより実施をサポート)
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
AIシステム 実装時に 確認する 事項	2)①	人間の生命・心身・財産、 及び環境への配慮	a. AI利用者を含む関連するステークホルダーの生命・身体・ 財産、精神及び環境に危害を及ぼすことがないよう、提供時 点で予想される利用条件下でのパフォーマンスだけでなく、 様々な状況下でAIシステムがパフォーマンスレベルを維持でき るようにし、リスク（運動するロボットの制御不能、不適切な 出力等）を最小限に抑える方法（ガードレール技術等）を 検討しているか？	4部 2) 安全性	a. 採用AIが募集職種又は応募者によらずに判 定結果を安定して出力できることが必要	-	a. ・様々なグループ会社・職種・スキル要件を想定し たバリエーションで採用AIのユーザーテストを行う （予測結果だけでなく、判断根拠の妥当性も 確かめる）	
	2)②	適正利用	a. AIシステム・サービスの利用上の留意点を正しく定めている か？ b. AI開発者が設定した範囲でAIを活用しているか？ c. 提供時点でAIシステム・サービスの正確性・必要な場合に は学習データの最新性（データが適切であること）等の担保 をしているか？ d. AI開発者が設定したAIの想定利用環境とAIシステム・ サービスの利用者の利用環境に違い等がないか検討してい るか？	4部 2) 安全性	a. 人材採用担当者がスキルによらずに安全に利 用できるようにすることが必要 b. AI開発者が定める範囲で活用することが必 要 c. 正確かつ最新の評価基準に基づく判定の担保 が必要 d. 想定している環境でAI利用者が利用できるか 確認が必要	開発部門や人材採用 部門との連携の必要 性あり	a. ・採用AIの利用目的・範囲を明確にした利用者 向けのマニュアルを整備する  b. ・採用AIの開発計画書に定められた範囲内で活 用する  c. ・リリースされているモデルのバージョンが最新である こと及び用いられているデータが正確であることを確 認する  d. ・開発する採用AIについて、対象とする利用環境 (利用端末、動作環境、UI等)が適切であることを 確かめる	
	3)①	AIモデルの各構成技術に含 まれるバイアスへの配慮	a. 提供時点でデータの公平性の担保及び参照する情報、連 携する外部サービス等のバイアスを検討しているか？ b. AIモデルの入出力及び判断根拠を定期的に評価し、バイ アスの発生をモニタリングしているか？ c. AI開発者にAIモデルを構成する各技術要素のバイアスの 再評価、評価結果に基づくAIモデル改善の判断を促してい るか？（必要に応じて実施） d. AIモデルの出力結果を受け取るAIシステム、ユーザーインタ フェースにおいて、ビジネスプロセス及びAI利用者又は業務外 利用者の判断を恣意的に制限するようなバイアスが含まれて しまう可能性を検討しているか？	4部 3) 公平性	a. エントリーシートの公平性の担保が必要 b. 採用AIのモニタリングが必要 c. 節目節目でAIモデル改善の判断の促進が必要 d. 採用AIのユーザーインターフェース上で表示され る応募者情報にバイアスが含まれる可能性あり	-	a. ・改ざん防止のため、エントリーシートの編集ができ ないようにする機能を実装する  b. ・継続的にバイアスのモニタリングをする  c. ・バイアスの傾向変化についてAI利用者とAIモデ ル改善要否のディスカッションをし、AI開発者に ディスカッション内容を共有する  d. ・ユーザーインターフェース上に判断根拠情報が適 切に表示されること(恣意的な変換を行わない)の 確認をする	
	4)①	AIシステム・サービス全般に おけるプライバシーの保護	a. AIシステムの実装の過程を通じて採用する技術の特性に 照らし適切に個人情報へのアクセスを管理・制限する仕組み の導入等のプライバシー保護のための対策（プライバシー・パ イ・デザイン）を講じているか？	4部 4) プライバシー保護	a. 本来アクセス不要な部門まで個人情報にアク セスできる状態になることが懸念される	-	a. ・リリース前に法務からのレビューを受け、個人情 報保護法に抵触していないかの確認を行う プライバシー保護の体制を構築する	
	5)①	AIシステム・サービスに影響 するセキュリティ対策	a. AIシステム・サービスの提供の過程を通じて、採用する技 術の特性に照らし適切にセキュリティ対策（セキュリティ・バイ ・デザイン）を講じているか？	4部 5) セキュリティ確保	a. 採用関連情報へのアクセス権限があいまいに なっている可能性がある	開発部門との連携の 必要性あり	a. ・開発部門品質管理担当者/人材採用担当者 を含めて、定期的にアクセス権限を棚卸する	
	6)④	関連するステークホルダーへ の説明可能性・解釈可能 性の向上	a. トレーサビリティ及び透明性の向上のため、意思決定に影 響を与える提供する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っているか？	4部 6) 透明性	a. 各種データ及び設定情報は記録しているが、 整理されていない可能性がある	-	a. ・サービス全体のアーキテクチャ(学習データ収集、 教師ラベルの設定、利用ログの保存、社内イント ラとの連携、UI等)を文書化する	

## 具体的なアプローチ検討のためのワークシート (「AI提供者」関連 ※ 記載例あり)

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、あくまで**各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。  
 (エントリーシートの文章で、応募者に対し合否を判断する採用AI (XGBoost)の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf)  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	人材採用部門XXX (AI倫理委員会XXXにより実施をサポート)
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
AIシステム・サービス提供後に確認する事項	2)② 適正利用	a. 適切な目的でAIシステム・サービスが利用されているかを定期的に検証しているか？	4部 2) 安全性	a. 適切な目的で利用されていない可能性がある -			a. ・AIシステム・サービスの利用状況について、ログを確認する等検証を行う	
	4)① AIシステム・サービス全般におけるプライバシーの保護	a. 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ているか？ b. プライバシー侵害を認識した場合等は適切に対処するとともに、再発の防止を検討しているか？	4部 4) プライバシー保護	a. プライバシー情報の不適切利用が生じている可能性がある b. プライバシー侵害時に対応する人員が不十分な可能性がある			a. ・開発部門のエントリーシートの活用履歴ログを定期的にモニタリングする  b. ・プライバシー侵害時の責任者を中心とした対応方針の策定をする。また、プライバシー保護組織の体制を構築する	
	5)② 最新動向への留意	a. AIシステム・サービスに対する攻撃手法も数多く生まれていするため、最新のリスク及びそれに対応するために提供の各工程で気を付けるべき点の動向を確認しているか？ b. 脆弱性に対応することを検討しているか？	4部 5) セキュリティ確保	a. 採用AIに影響の大きいリスクの動向を注視する必要がある b. 外部システムとの連携において脆弱性が現れる可能性がある			a. ・HRテック及びAIに関するリスクについて、定期的に情報を確認する  b. ・定期的なシステムの更新をする（イントラ内での利用のため、外部攻撃等のリスクについては会社の情報セキュリティ管理に依拠）	
	6)② 関連するステークホルダーへの情報提供	a. 提供するAIシステム・サービスについて、例えば以下の事項を平易かつアクセスしやすい形で、適時かつ適切に情報を提供しているか？ ・AIを利用しているという事実、適切/不適切な使用方法等 ・提供するAIシステム・サービスの技術的特性、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 ・AIシステム・サービスの学習等による出力又はプログラムの変化の可能性 ・AIシステム・サービスの動作状況に関する情報、不具合の原因及び対応状況、インシデント事例等 ・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 ・AIモデル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	4部 1) 人間中心 2) 安全性 3) 公平性 4) プライバシー保護 5) セキュリティ確保 6) 透明性	a. 採用AIについて、利用者に適時かつ適切に説明する必要がある	開発部門と人材採用担当者との連携の必要性あり		a. ・採用AIについて以下項目を人材採用担当者に説明する ・AIを利用していること、適切な使用方法及び禁止事項 ・AIサービス全体で発生するインシデント状況及び対応方法 ・起こりうるリスク及び対策	
	7)② 共通の指針の対応状況の説明	a. AI利用者に適正利用を促し、以下の情報をAI利用者に提供しているか？ ・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利用についての注意喚起 ・コンテキスト内学習による不適切なAIモデルの学習に対する注意喚起 ・個人情報を入力する際の留意点 b. 提供するAIシステム・サービスへの個人情報の不適切入力について注意喚起しているか？	4部 2) 安全性 4) プライバシー保護 7) アカウンタビリティ	a. ユーザー入力によって推論性能に影響を与える - ことがないため対象外とする（受領したエントリーシートのファイル(書き込み不可)をアップロードするため) b. エントリーシートファイルのアップロード時に不適切入力に留意する必要がある			b. ・エントリーシートファイルのアップロード時等に誤った個人情報を入力しないよう、ヒヤリハットも含め、人材採用担当者に注意喚起を行う	
	7)⑥ 文書化	a. AI利用者、業務外利用者に向けたサービス規約を作成しているか？ b. プライバシーポリシーを明示しているか？	4部 7)アカウンタビリティ	a. b. 採用AIの開発者及び利用者も同一事業グループ内の主体であるため、全社共通にポリシー・規定を定めることが有効			a. ・AI倫理委員会を中心に採用AIのサービス規約及びプライバシーポリシーを策定する	







# 具体的なアプローチ検討のためのワークシート (「AI提供者」関連 ※ 記載例あり)

**利用上の留意点** ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、あくまで**各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ字にて記載例を示しています。  
 (エントリーシートの文章で、応募者に対し合否を判断する採用AI (XGBoost)の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf)  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	人材採用部門XXX (AI倫理委員会XXXにより実施をサポート)
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
2)①	人間の生命・心身・財産、及び環境への配慮	人間の生命・心身・財産、及び環境への配慮 a. AIシステム・サービスの出力の正確性を含め、要求に対して十分に動作していること（信頼性）を確保しているか？ b. 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ようにしているか（堅牢性（robustness））？ c. AIの活用又は意図しないAIの動作によって生じうる権利侵害の重大性、侵害発生の可能性等、当該AIの性質・用途等に照らして、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？ d. 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？ e. 人間の生命・心身・財産、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？ f. 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？ g. AIシステム・サービスの安全性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？	2部 2) 安全性	a. b. 採用AIの堅牢性の確保が必要 c. 権利侵害等が発生しないようにモニタリングすることが必要 d. リスクの整理及び対策が必要 e. 財産や環境に危害を及ぼす可能性について整理し、関連情報を共有することが必要 f. 関連するステークホルダーは直接採用AIを利用しないため対象外 g. 安全性に関するインシデントへの対策の準備が必要	-		a. b. ・開発部門品質管理担当者が実施した検証結果へのレビューをする(テストケース、予測結果、出力内容・表現の適切さ、判断結果・根拠情報の妥当性等) ・AI利用者への変更内容の説明及び指摘・フィードバックへの対応を行う ・リリースを承認する ・開発部門が適切に採用AIの開発を行っているかについて確認する ・開発部門に対して、採用AIの提供状況、運用上や技術的な問題等についてフィードバックを行う	
							c. ・AI利用者のUI上に判断根拠情報が適切に表示されているか確認する	
							d. ・採用AIのサービス全般に関わるリスクシナリオを網羅的に検討し、技術/非技術での対策を検討する リスクシナリオ及び対策の検討について、AI開発者及びAI利用者の合意を得る	
							e. ・応募者又は自社への危害等のリスクについて検討し、対策を立案する	
							g. ・AI利用者からAIの予測性能に関わるエスカレーションがあった際に関係するステークホルダーと連携できる体制・コミュニケーション方法を確立する	
3)②	人間の判断の介入	a. 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、適切なタイミングで人間の判断を介入させる利用を検討しているか？ b. バイアスが生じていないか、AIシステム・サービスの目的、制約、要件、決定を明確かつ透明性のある方法により分析し、対処するためのプロセスを導入しているか？ c. 無意識のバイアス及び潜在的なバイアスに留意し、多様な背景、文化、分野のステークホルダーと対話した上で、方針を決定しているか？	2部 3) 公平性	a. 採用可否の判断において人間の判断の介入が必要 b. 出力の判断等について透明性のある方法での分析が必要 c. 多様なステークホルダーと対話した上で方針を決定することが必要	開発部門と人材採用担当者との連携の必要性あり		a. ・最終判断は人材採用担当者が行うことを人材採用担当者に周知する	
							b. ・本AIサービス全般において公平性のリスク要因と対策(データのバイアス、モデルのバイアス、利用者側への予測の苦手な対象の説明及び最終判断のプロセス)を検討する	
							c. ・人材採用部門の代表者が中心となって、必要なステークホルダーとの協議を行う	
6)④	関連するステークホルダーへの説明可能性・解釈可能性の向上	a. 関連するステークホルダーの納得感及び安心感の獲得、また、そのためのAIの動作に対する証拠の提示等を目的として、説明する主体がどのような説明が求められるかを分析・把握できるよう、説明を受ける主体がどのような説明が必要かを共有し、必要な対応を講じているか？	2部 6) 透明性	a. AIの動作に分析・把握し、必要な対応をとり説明できる状態にすることが必要	開発部門と人材採用担当者との連携の必要性あり		a. ・リスクシナリオ及び対策の検討過程において、ステークホルダーへ説明すべき内容を整理する	
8)①	AIリテラシーの確保	a. 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？	2部 8) 教育・リテラシー	a. 採用AIのシステムを実装するにあたり、実装担当者がセキュリティ、プライバシー等に関するAIリテラシーを備えているか懸念がある	-		a. ・採用AIサービスに関与するAI提供者に対して、必要な知識・経験を有しているか確かめ、不十分であればAIサービス導入・運営の経験者がOJTでコーチングする	
8)②	教育・リスキリング	a. 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と想定されるため、新たな働き方ができるよう教育・リスキリング等を検討しているか？	2部 8) 教育・リテラシー	a. システムのアップデートに応じて、AIシステムと運用者の関わりが変わるが、働き方の変化に柔軟に対応できない懸念がある	-		a. ・システムアップデートの都度、運用担当者に対して、説明及び教育の機会を設ける	
10)②	相互接続性・相互運用性への留意	a. 自らのAIシステム・サービスと他のAIシステム・サービスとの相互接続性及び相互運用性を確保しているか？ b. 標準仕様がある場合には、それに準拠しているか	2部 10) イノベーション	a. b. 外部システムと連携しないため対象外	-			

# 具体的なアプローチ検討のためのワークシート (「AI利用者」関連 ※ 記載例あり)

**利用上の留意点** ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するものであり、各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくことを前提としたものです。**  
 したがって、必ずしも、全ての事項について、検討が必要となるものではございませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、G列以降に、オレンジ色にて記載例を示しています。  
 (エントリーシートの文章で、応募者に対し合否を判断する採用AI (XGBoost)の例。https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\_Case01\_Recruitment-AI\_JP.pdf)  
 各自用にカスタマイズし運用するには、α. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γが重複する場合もある）。

α. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	人材採用担当者XXX (AI倫理委員会XXXにより実施をサポート)
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
AIシステム・サービス利用時に確認する事項	2)② 適正利用	a. AI提供者が定めた利用上の留意点を遵守して、AI提供者が設計において想定した範囲内でAIシステム・サービスを利用しているか？ b. 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？ c. 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？	5部 2) 安全性	a. AI提供者による利用範囲の設定あり b. 正確・最新性データ（エントリーシート）の入力に留意することが必要 c. 出力結果の精度及びリスクの程度を確認することが必要	人材採用部門と連携の必要性あり		a. ・人材採用部門と検討した上で、採用AIの利用目的・範囲を理解し、不適切な目的外利用を行わない  b. ・予測対象とするエントリーシートが適切であることを確かめる  c. ・リスクシナリオの中で人材採用担当者側で対応すべきリスクコントロールを認識する	
	3)① AIモデルの各構成技術に含まれるバイアスへの配慮	a. 著しく公平性を欠くことがないよう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？	5部 3) 公平性	a. ユーザー入力によって推論性能に影響を与えることがないため対象外				
	4)① AIシステム・サービス全般におけるプライバシーの保護	a. AIシステム・サービスへ個人情報等を不適切に入力することがないよう注意を払っているか？ b. 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？	5部 4) プライバシー保護	a. 入力情報は申込者のエントリーシートであり、個人情報が入力されても前処理で個人名は削除されるため、名前以外の個人情報と名前が紐づかず、個人情報の不適切入力にはならないため対象外 b. プライバシー侵害に関する情報収集及び防止が必要			b. (通常の人事業務の内部統制として実施) ・エントリーシートを継続的に提供するために、キャリアコンサルタント及び申込者に対して事前にデータの利用範囲(AI開発者/AI提供者にデータを提供すること)及び利用目的の説明並びに同意を得る	
	5)① AIシステム・サービスに影響するセキュリティ対策	a. AI提供者によるセキュリティ上の留意点を遵守しているか？ b. AIシステム・サービスに機密情報等を不適切に入力することがないよう注意を払っているか？	5部 5) セキュリティ確保	a. ※イントラ内での利用のため、外部攻撃等のリスクについては会社の情報セキュリティ管理に依拠 b. エントリーシートの入力時等に機密情報等の不適切入力に留意することが必要			a. ・社内の情報セキュリティのルールを遵守する ・セキュリティ上の疑義を感じた場合は、開発部門等にその旨を報告する  b. ・データ入力時に機密情報等の不適切入力を行っていないかダブルチェックを行う	
	6)② 関連するステークホルダーへの情報提供	a. 著しく公平性を欠くことがないよう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AIシステム・サービスから出力結果を取得する。そして、出力結果を事業判断に活用した際は、その結果を関連するステークホルダーに周知しているか？	5部 3) 公平性 6) 透明性	a. ユーザー入力によって推論性能に影響を与えることがないため対象外 (受領したエントリーシートのファイル(書き込み不可)をアップロードするため)				
	7)② 「共通の指針」の対応状況の説明	a. 関連するステークホルダーの性質に応じて害のない範囲で、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？ b. 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ことが予定されている場合には、AIの特性及び用途、データの提供元となる関連するステークホルダーとの接点、プライバシーポリシー等を踏まえ、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に情報提供をしているか？ c. 当該AIの出力結果を特定の個人又は集団に対する評価の参考にする場合は、AIを利用している旨を評価対象となっている当該特定の個人又は集団に対して通知し、当ガイドラインが推奨する出力結果の正確性、公正さ、透明性等を担保するための諸手続きを遵守し、かつ自動化バイアスも鑑みて人間による合理的な判断のもと、評価の対象となった個人又は集団からの求めに応じて説明責任を果たしているか？ d. 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？	4部 1) 人間中心 7) アカウンタビリティ	a. 関連するステークホルダーは直接利用しないため対象外 b. エントリーシートの情報を用いるため、申込者へ情報提供が必要 c. 当該AIの出力結果を特定の個人又は集団に対する評価の参考にする場合に該当する。申込者に対して、AIを利用している旨等の通知が必要 d. 申込者からの問い合わせに対応する体制構築が必要			b. ・申込者に対して、データの利用範囲等について情報提供をする  c. ・申込者に対して、エントリーシートの合否判断の過程でAIを利用する旨をエントリーシート記入依頼時に通知する  d. ・申込者からの問合せメールフォームを設置し、対応を行う	
	7)⑥ 文書化	a. 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適切に保管・活用しているか？ b. AI提供者が定めたサービス規約を遵守しているか？	5部 7) アカウンタビリティ	a. b. 人材採用部門から共有された最新の採用AIに関する利用方針に従って利用する必要があら	人材採用部門と連携の必要性あり		a. ・人材採用部門から受領した情報を社内のイントラから最新版を閲覧できるようにする b. ・人材採用部門の示すルールに沿った業務マニュアルを策定し、業務を遂行する	

# 具体的なアプローチ検討のためのワークシート (「AI利用者」関連 ※ 記載例あり)

**利用上の留意点** ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください。**  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 なお、このシートでは、**G列以降に、オレンジ色にて記載例を示しています。**  
 (エントリーシートの文章で、応募者に対し合否を判断する採用AI (XGBoost)の例。 [https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel\\_Case01\\_Recruitment-AI\\_JP.pdf](https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case01_Recruitment-AI_JP.pdf))  
 各自用にカスタマイズし運用するには、α. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γ.が重複する場合もある）。

α. ワークシートの作成者	AI倫理委員会担当者XXX
β. 実施状況の確認を行う者	人材採用担当者XXX (AI倫理委員会XXXにより実施をサポート)
γ. 責任者	AI倫理委員会担当者XXX

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 (該当しない場合はその理由)	他の主体との関係についての事項	✓	具体的なアプローチ	最終検討日 (見直し日)
1)④	多様性・包摂性の確保	a. 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？	2部 1) 人間中心	a. 情報リテラシーが十分でない応募者への配慮が必要			a. ・障がい者雇用のプログラム等、電子ファイルによるエントリーシートを作成するための情報リテラシーが十分でない人材採用のチャンネルは採用AIとは別に人事業務として用意する	
2)①	人間の生命・心身・財産、及び環境への配慮	a. AIシステム・サービスの出力の正確性を含め、要求に対して十分に動作していること (信頼性) を確保しているか？ b. 様々な状況下でパフォーマンスを維持し、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ようにしているか (堅牢性 (robustness)) ？ c. AIの活用又は意図しないAIの動作によって生じう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？ d. 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？ e. 人間の生命・心身・財産、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？ f. 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？ g. AIシステム・サービスの安全性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？	2部 2) 安全性	a. b. 様々な状況において一定レベルの採用AIの出力確保が必要 c. 採用AIの制御可能性の確保が必要 d. リスクの把握が必要 e. 危害を及ぼす可能性の整理及び情報提供が必要 f. 関連するステークホルダーは直接AIを利用しないため対象外 g. 採用AIの安全性が損なわれた場合の対策を検討しておくことが必要			a. b. ・人材採用部門からの事前説明に対する指摘・フィードバック、必要な範囲でのユーザーテストの実施 ・人材採用部門に対して、採用AIの利用状況、運用上や技術的な問題等についてフィードバックを行う	
							c. ・業務プロセス上で人材採用担当者が書類選考の合否を判断する	
							d. ・リスクシナリオ及び対策について説明を受け、改善点をフィードバックする	
							e. ・候補者の心身、自社の財産又は環境へ危害を及ぼす可能性について、事前に整理をする	
							g. ・AIモデルが有効に機能しないと判断される場合の代替運用プロセスを検討する	
3)②	人間の判断の介在	a. 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、適切なタイミングで人間の判断を介在させる利用を検討しているか？ b. バイアスが生じていないか、AIシステム・サービスの目的、制約、要件、決定を明確かつ透明性のある方法により分析し、対処するためのプロセスを導入しているか？ c. 無意識のバイアス及び潜在的なバイアスに留意し、多様な背景、文化、分野のステークホルダーと対話した上で、方針を決定しているか？	2部 3) 公平性	a. 出力結果が含みうるバイアスの確認が必要 b. バイアスの分析・対処が必要 c. バイアスについての対話が必要	開発部門・人材採用部門との連携の必要性あり		a. ・著しい学習データの不足やモデル評価への影響に開発部門で対応できないバイアス(例：外国籍の応募者、障害者雇用の対象者、従来重視していなかった職種等)について、AI開発者から説明を受け業務側でカバーできるか確認する	
							b. ・人材採用部門が検討したリスク対策の計画(リスクチェーンモデルの内容)について説明を受け、人材採用部門で最終判断を行う上で認識しておくべき事項を理解する	
							c. ・人材採用部門の代表者が中心となって、必要なステークホルダーとの協議を行う	
8)①	AIリテラシーの確保	a. 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？	2部 8) 教育・リテラシー	a. 採用AIのシステムを利用するにあたり、人材採用担当者がセキュリティ、プライバシー等に関するAIリテラシーを備えているか懸念がある			a. ・採用AIサービスに関与するAI利用者に対して、必要な知識・経験を有しているか確かめ、不十分であれば人材採用部門がOJTでコーチングする	
8)②	教育・リスキリング	a. 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と想定されるため、新たな働き方ができるよう教育・リスキリング等を検討しているか？	2部 8) 教育・リテラシー	a. システムのアップデートに応じて、AIシステムと利用方法が変わりうるが使いこなせない人材採用担当者が出る懸念がある			a. ・システムアップデートの都度、人材採用担当者に対して、説明及び教育の機会を設ける	



## 具体的なアプローチ検討のためのワークシート（「高度なAIシステムに関する事業者」関連）

利用上の留意点  
 ガイドラインに記載した内容に関して取り組むべき事項は、**各事業者の事業内容や置かれた状況等により、個々に異なります。**  
 このため、本ワークシートは、**あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するもの**であり、**各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくこと**を前提としたものです。  
 したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、**活用の要否、各自の事情に応じた修正や取捨選択を検討ください**（※I）～ XI）は高度なAIシステムを開発するAI開発者にのみ適用される内容もあるため、AI提供者及びAI利用者は適切な範囲でご留意ください）。  
 D列～F列の記載内容を基に、**G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。**  
 各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

a. ワークシートの作成者
β. 実施状況の確認を行う者
γ. 責任者

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 （該当しない場合はその理由）	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 （見直し日）
I)	AIライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度なAIシステムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じる	a. レッドチーム等の様々な手法を組み合わせて、多様/独立した内外部テスト手段を採用することや、特定されたリスクや脆弱性に対処するための適切な緩和策を実施しているか？ b. 上記テストを支援するために、開発中に行われた意思決定に関するトレーサビリティを確保するように努めているか？	2) 安全性 6) 透明性 7) アカウンタビリティ					
II)	市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやパターンを特定し、緩和する	a. リスクレベルに見合った適切なタイミングで、AIシステムの活用状況のモニタリングを実施し、それらに対処するための適切な措置を講じているか？ b. 他の利害関係者と協力して、報告されたインシデントの適切な文書化を維持し、特定されたリスクと脆弱性を軽減しているか？	5) セキュリティ確保 7) アカウンタビリティ					
III)	<input type="checkbox"/> 高度なAIシステムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウンタビリティの向上に貢献する	a. データの出所に始まり、どのような意思決定を行ったかについて、合理的な説明を行い、トレーサビリティを確保するため文書化・公表しているか？ b. 関連するステークホルダーがAIシステムの出力を解釈し、利用者や業務外利用者が適切に活用できるようにするために、明確で理解可能な形で文書化・公表しているか？	6) 透明性 7) アカウンタビリティ					
IV)	<input type="checkbox"/> 産業界、政府、市民社会、学界を含む、高度なAIシステムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組む	a. モニタリング結果の報告書やセキュリティや安全性のリスクに関する関連文書等を通して、高度なAIシステムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組んでいるか？	5) セキュリティ確保 6) 透明性 7) アカウンタビリティ 10) イノベーション					
V)	<input type="checkbox"/> 特に高度なAIシステム開発者に向けた、個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づくAIガバナンス及びリスク管理方針を策定し、実施し、開示する	a. 適切な場合には、プライバシーポリシーを公表しているか？ b. AIガバナンスに関するポリシーや実行するための組織を確立し、開示しているか？	4) プライバシー保護 7) アカウンタビリティ					
VI)	<input type="checkbox"/> AIのライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する	a. 情報セキュリティのための運用上の対策や適切なサイバー/物理的アクセス制御等を検討しているか？	5) セキュリティ確保					
VII)	<input type="checkbox"/> 技術的に可能な場合は、電子透かしやその他の技術等、AI利用者及び業務外利用者が、AIが生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入する	a. 適切かつ技術的に実現可能な場合、組織の高度なAIシステムで作成されたコンテンツ認証及び来歴メカニズムを開発し、導入しているか？ b. 透かし等を通じた特定のコンテンツが高度なAIシステムで作成されたかどうかをAI利用者及び業務外利用者が判断できるツールやAPIの開発をしているか？ c. AI利用者及び業務外利用者がAIシステムと相互作用していることを知るができるよう、ラベリングや免責事項の表示等、その他の仕組みを導入しているか？	6) 透明性					
VIII)	<input type="checkbox"/> 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先する	a. AIの安全性、セキュリティ、信頼性の向上やリスクへの対処に関する研究等、社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先しているか？	10) イノベーション					
IX)	<input type="checkbox"/> 世界の最大の課題、特に気候危機、世界保健、教育等（ただしこれらに限定されない）に対処するよう貢献するため、高度なAIシステムの開発を優先する	a. 信頼性のある人間中心のAI開発に向けた取組を実施し、同時に業務外利用者も含めたリテラシーの向上のための支援をしているか？	10) イノベーション					
X)	<input type="checkbox"/> 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進する	a. 電子透かしを含む国際的な技術標準とベストプラクティスの開発に貢献し、適切な場合にはそれを利用し、標準開発組織（SDO）と協力しているか？	10) イノベーション					



## 具体的なアプローチ検討のためのワークシート（「高度なAIシステムに関する事業者」関連）

利用上の留意点  
ガイドラインに記載した内容に関して取り組むべき事項は、各事業者の事業内容や置かれた状況等により、個々に異なります。  
このため、本ワークシートは、あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するものであり、各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくことを前提としたものです。  
したがって、必ずしも、全ての事項について、検討が必要となるものではありませんので、活用の要否、各自の事情に応じた修正や取捨選択を検討ください（※I）～ XI）は高度なAIシステムを開発するAI開発者にのみ適用される内容もあるため、AI提供者及びAI利用者は適切な範囲でご留意ください）。  
D列～F列の記載内容を基に、G列以降を各事業者にてご検討いただき、各自が取り組むべきこと（あるいは内容）を具体化の上、活用ください。  
各自用にカスタマイズし運用するには、a. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、a.～γ.が重複する場合もある）。

- a. ワークシートの作成者
- β. 実施状況の確認を行う者
- γ. 責任者

	分類	検討にあたって重要な事項	対応箇所	各自の事業において検討対象とする事項 （該当しない場合はその理由）	他の主体との関係 についての事項	✓	具体的なアプローチ	最終検討日 （見直し日）
	XI) □ 適切なデータインプット対策を実施し、個人データ及び知的財産を保護する	a. 有害な偏見バイアスを軽減するために、訓練データやデータ収集等、データの質を管理するための適切な措置を講じているか？ b. 訓練用データセットの適切な透明性も支援されるべきであり、適用される法的枠組みを遵守しているか？	2) 安全性 3) 公平性					
	XII) □ 高度な AI システムの信頼でき責任ある利用を促進し、貢献する	a. 高度なAIシステムが特定のリスク（例えば偽情報の拡散に関するもの）をどのように増大させるか、新たなリスクをどのように生み出すか等の課題を含め、各主体及びステークホルダーのリテラシーや認識等の向上のための機会を提供しているか？ b. 各主体間で連携し、高度なAIシステムに関する新たなリスクや脆弱性を特定し、対処するための情報共有を行っているか？	5) セキュリティ確保 8) 教育・リテラシー					