

別添 8 主体横断的な仮想事例

令和6年4月

Case 採用AI

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）  
[https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RModel\\_Case01\\_Recruitment-AI\\_JP.pdf](https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RModel_Case01_Recruitment-AI_JP.pdf)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>1) 人間中心</b>				
各主体は、AIシステム・サービスの開発・提供・利用において、後述する各事項を含む全ての取り組むべき事項が導出される土台として、少なくとも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すべきである。また、AIが人々の能力を拡張し、多様な人々の多様な幸せ（well-being）の追求が可能となるよう行動することが重要である。				
<b>①人間の尊厳及び個人の自律</b>				
1	共通	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する	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なっている。	AIサービスの提供において、AI利用者が最終判断(応募者の合否)を行えるようになっている(Human-in-the-loop)
2	共通	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する	脳・身体と連携するケースではないため対象外	脳・身体と連携するケースではないため対象外
3	共通	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ない	AIシステムの開発において、実際の予測結果を学習データに用いる際には個人情報 の取扱いに関わる誓約書の締結やアクセス権管理等を実施している。 ※公平性とプライバシーについては、「3）公平性」「4）プライバシー保護」を参照	AIシステムの開発において、実際の予測結果を学習データに用いる際には個人情報 の取扱いに関わる誓約書の締結やアクセス権管理等を実施している。 ※公平性とプライバシーについては、「3）公平性」「4）プライバシー保護」を参照
<b>②AIによる意思決定・感情の操作等への留意</b>				
1	共通	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・利用は行わない	本ケースに関しては、2)①-3と同じ論点になる	本ケースに関しては、2)①-3と同じ論点になる
2	共通	AIシステムの開発・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、必要な対策を講じる	本ケースに関しては、2)①-3と同じ論点になる	本ケースに関しては、2)①-3と同じ論点になる
3	共通	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う	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	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
4	共通	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取り扱う	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	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
<b>③偽情報等への対策</b>				
1	共通	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成できるようになり、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、必要な対策を講じる	本ケースではAIが出力を生成するUCではないため、対象外	本ケースではAIが出力を生成するUCではないため、対象外
<b>④多様性・包摂性の確保</b>				
1	共通	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う (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)	本ケースは社内の人材採用担当者が用いるため、対象外	本ケースは社内の人材採用担当者が用いるため、対象外 ・障害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は本AIサービスとは別に人事業務として用意する
<b>⑤利用者支援</b>				
1	共通	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する (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)	本ケースに関しては、2)①-3の制御可能性と同じ論点になる	本ケースに関しては、2)①-3の制御可能性と同じ論点になる
<b>⑥持続可能性の確保</b>				
1	共通	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する	・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	・AI開発者から提供された計算資源の使用ログを元に、月次でデータセンターにおいて発生している温室効果ガス排出量を確かめる。 ・社内のESG部門にデータセンターからの温室効果ガス排出量を定期的に報告し、もし改善すべき点があればAI開発者と計算資源への負荷のサーベイ・見直しを行う。

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）  
[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者	
		AI開発者	AI提供者	AI利用者	
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	
<b>2) 安全性</b>					
各主体は、AIシステム・サービスの開発・提供・利用を通じ、ステークホルダーの生命・身体・財産に危害を及ぼすことがないようにすべきである。加えて、環境に危害を及ぼすことがないようにすることが重要である。					
<b>① 人間の生命・身体・財産、精神及び環境への配慮</b>					
1	共通	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）	・AIモデルの予測精度の評価：分類の正確性	・AI開発者が実施した検証結果へのレビュー（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と指摘・フィードバックへの対応 ・リリースへの承認	・AI提供者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
2	共通	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する（堅牢性（robustness））	・AIモデルの堅牢性の評価：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	同上	同上
3	共通	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	・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ことのレビュー	・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
4	共通	適切なリスク分析を実施し、リスクへの対策（回避、低減、移転又は容認）を講じる	・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、AI提供者と共に検討	・本AI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 ・リスクシナリオと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	・AIサービスに関わるリスクシナリオと対策について説明を受け、改善点をフィードバックする
5	共通	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する（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する）	※業務外利用者は存在していないため対象外	※業務外利用者は存在していないため対象外	※業務外利用者は存在していないため対象外
6	共通	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 ・モデルの再学習を実施できる体制・環境	・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	・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
<b>D-2) ii. 人間の生命・身体・財産、精神及び環境に配慮した開発</b>					
7	開発	AI開発者は、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、以下の事項を検討する ・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 ・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）	・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	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
<b>P-2) i. 人間の生命・身体・財産、精神及び環境に配慮したリスク対策</b>					
8	提供	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）を検討する	※AI提供者に係る項目のため対象外	・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サービス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）	※AI提供者に係る項目のため対象外
<b>② 適正利用</b>					
1	共通	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・利用を行う	・予測対象とする文章の仕様を明確化(対象：エントリーシート、言語：日本語、文字数：2000字以内) ・利用ログの保存	・利用ログのレビュー（異常な予測結果、反復利用の多いユーザー等） ・アクセス権限の管理	・ユーザー教育(利用目的や禁止行為の通達) ・キャリアコンサルへの担当のローテーション
<b>D-2) iii. 適正利用に資する開発</b>					
2	開発	AI開発者は、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避けるため、AIを安全に利用可能なAIの使い方について明確な方針・ガイダンスを設定する	・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	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
3	開発	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する（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）	※通常の機械学習モデルによる開発ケースであるため対象外	※通常の機械学習モデルによる開発ケースであるため対象外	※通常の機械学習モデルによる開発ケースであるため対象外
<b>P-2) ii. 適正利用に資する提供</b>					
4	提供	AIシステム・サービスの利用上の留意点を正しく定める	※AI提供者に係る項目のため対象外	・AIサービスの利用目的・範囲を明確にした利用者向けのマニュアルを整備する	※AI提供者に係る項目のため対象外
5	提供	AI開発者が設定した範囲でAIを活用する	※AI提供者に係る項目のため対象外	・AIサービスの提供範囲(利用会社・対象職種・応募者)及び役割(書類選考のスクリーニングであることを明確にして、AIサービス開発計画書に記載する)	※AI提供者に係る項目のため対象外
6	提供	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）等を担保する	※AI提供者に係る項目のため対象外	・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を確かめる。	※AI提供者に係る項目のため対象外
7	提供	AI開発者が開発した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を検討する	※AI提供者に係る項目のため対象外	・開発されるAIサービスについて、利用環境(利用端末、動作環境、UI)が適切であることを確かめる	※AI提供者に係る項目のため対象外
<b>U-2) i. 安全を考慮した適正利用</b>					
9	利用	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・サービスを利用する	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外	・AIサービス提供者と検討した上で、AIサービスの利用目的・範囲を理解し、不適切な目的外利用を行わない。 ・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
10	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投入を行う	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外	・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
11	利用	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する	※AI利用者に係る項目のため対象外	・AI利用者に向けて、AI利用者側で必要となるリスクコントロールを説明する	・AIサービスに係るリスクシナリオの中でAI利用者側で対応すべきリスクコントロールを認識する。

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>③ 適正学習</b>				
1	共通	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）等を確保する	<ul style="list-style-type: none"> <li>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</li> <li>AIモデルの正確性の定期的な検証</li> <li>アノテーションにおける異常値の検証</li> </ul>	<ul style="list-style-type: none"> <li>AI開発者が実施する定期的な検証内容の確認</li> <li>再学習の要否についてAI利用者と調整した上で判断</li> </ul>
2	共通	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する		<ul style="list-style-type: none"> <li>合否のフィードバック(次の学習データに用いる教師ラベルの設定)</li> </ul>
<b>D-2) i. 適切なデータの学習</b>				
3	開発	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する	<ul style="list-style-type: none"> <li>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(SpaCy &amp; Presidioでの検出 + AI開発者による検証)</li> <li>※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</li> <li>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</li> </ul>	<ul style="list-style-type: none"> <li>AI開発者が実施する学習データにおける検証内容のレビュー</li> </ul>
4	開発	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	<ul style="list-style-type: none"> <li>学習データを適切な場所に格納する</li> <li>開発環境について適切なAI開発者にのみ適切なレベルのアクセス権限を付与する。</li> </ul>	<ul style="list-style-type: none"> <li>AI開発者/AI利用者を含めて、定期的にアクセス権限を棚卸</li> </ul>
<p>(通常の人事業務の内部統制として実施)</p> <ul style="list-style-type: none"> <li>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)と利用目的の説明と同意を得る</li> </ul> <p>(通常の人事業務の内部統制として実施)</p> <ul style="list-style-type: none"> <li>エントリーシートのデータを適切にアクセス管理する</li> </ul>				

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの記事で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者	
		AI開発者	AI提供者	AI利用者	
No	分類	共通の指針/各主体に関する事項			
		本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	
<b>3) 公平性</b>					
各主体は、AIシステム・サービスの開発・提供・利用において、特定の個人ないし集団へのその人種、性別、国籍、年齢、政治的信念、宗教等の多様な背景を理由とした不当で有害な偏見及び差別をなくすよう努めることが重要である。また、各主体は、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、この回避できないバイアスが人権及び多様な文化を尊重する観点から許容可能か評価をした上で、AIシステム・サービスの開発・提供・利用を行うことが重要である。					
<b>① AIモデルの各構成技術に含まれるバイアスへの配慮</b>					
1	共通	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、AI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うるバイアスの要因となるポイントを特定する	・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）	・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、最終判断のプロセス)を検討している	・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)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
2	共通	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する	同上(モデルの特徴量やユーザーの理解についても検討している)	同上(モデルの特徴量やユーザーの理解についても検討している)	同上(モデルの特徴量やユーザーの理解についても検討している)
<b>D-3) i. データに含まれるバイアスへの配慮</b>					
3	開発	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散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る	・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	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4	開発	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	・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	・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>D-3) ii. AIモデルのアルゴリズム等に含まれるバイアスへの配慮</b>					
5	開発	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する	※本UCについては、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とする	本UCについては、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とする	本UCについては、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とする
6	開発	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する	3)① D-3) i. -3/4を参照	3)① D-3) i. -3/4を参照	3)① D-3) i. -3/4を参照
<b>P-3) i. AIシステム・サービスの構成及びデータに含まれるバイアスへの配慮</b>					
7	提供	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する	※本UCについては、他に外部情報を参照しないため対象外とする	本UCについては、他に外部情報を参照しないため対象外とする	本UCについては、他に外部情報を参照しないため対象外とする
8	提供	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する。また、必要に応じて、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、評価結果にもとづくAIモデル改善の判断を促す	・利用ログの保存	・利用ログの継続的なバイアスのモニタリング ・バイアスの傾向変化についてAI利用者やAIモデル改善要否のディスカッション	・バイアスの傾向変化についてAI開発者及びAI提供者とAIモデル改善要否をディスカッションする
9	提供	AIモデルの出力結果を受け取るAIシステム・サービス、ユーザーインターフェースにおいて、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過去データ情報を出力	・AI利用者のUI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)のレビュー	・UIの仕様(出力上行われている変換)を理解
<b>U-3) i. 入力データ又はプロンプトに含まれるバイアスへの配慮</b>					
10	利用	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う	※本UCについては、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	本UCについては、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	本UCについては、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
<b>② 人間の判断の介在</b>					
1	共通	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する	3)① D-3) i. -3/4を参照	3)① D-3) i. -3/4を参照	3)① D-3) i. -3/4を参照
2	共通	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、要件及び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する	3)①-1を参照	3)①-1を参照	3)①-1を参照
3	共通	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、文化又は分野のステークホルダーと対話した上で、方針を決定する	・AI提供者とAI利用者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	・AI提供者とAI利用者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	・AI提供者とAI利用者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

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項		
		本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>4) プライバシー保護</b>				
各主体は、AIシステム・サービスの開発・提供・利用において、その重要性に応じ、プライバシーを尊重し、保護することが重要である。その際、関係法令を遵守すべきである。				
<b>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</b>				
1	共通	個人情報保護法等の関連法令の遵守や、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る	・本AIサービスの運営にアサインするにあたって、個人情報保護に係る誓約書にサインする	・会社のプライバシーポリシー及び個人情報保護規定を遵守 ・会社で実施する個人情報保護に係る社員教育の受講（罰則の理解）
2	共通	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検討する ・個人情報保護法にもとづいた対応の確保 ・国際的な個人データ保護の原則及び基準の参照	・クレンジング前の学習データ(エントリーシート)を適切な格納場所が必要な担当者のみがアクセス	・AIサービスの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。 ・AI開発者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>D-2) i. 適切なデータの学習（再掲）</b>				
3	開発	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する	本UCにおいては2)③ D-2) i. -3と同じ	本UCにおいては2)③ D-2) i. -3と同じ
4	開発	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	本UCにおいては2)③ D-2) i. -3と同じ	本UCにおいては2)③ D-2) i. -3と同じ
<b>P-4) i. プライバシー保護のための仕組み及び対策の導入</b>				
5	提供	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を講ずる（プライバシー・バイ・デザイン）	本UCにおいては4)①-2と同じ	本UCにおいては4)①-2と同じ
<b>P-4) ii. プライバシー侵害への対策</b>				
6	提供	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、侵害を認識した場合等は適切に対処するとともに、再発の防止を検討する	本UCにおいては4)①-2と同じ	本UCにおいては4)①-2と同じ
<b>U-4) i. 個人情報の不適切入力及びプライバシー侵害への対策</b>				
7	利用	AIシステム・サービスへ個人情報を不適切に入力することがないよう注意を払う	※本UCについて、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学習済みAIモデルの特徴量に個人名は含まれていないため、対象外	※本UCについて、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学習済みAIモデルの特徴量に個人名は含まれていないため、対象外
8	利用	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、防止を検討する	本UCにおいては4)①-2と同じ	本UCにおいては4)①-2と同じ

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）  
[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>5) セキュリティ確保</b>				
各主体は、AIシステム・サービスの開発・提供・利用において、不正操作によってAIの振る舞いに意図せぬ変更又は停止が生じることのないようにセキュリティを確保することが重要である。				
<b>① AIシステム・サービスに影響するセキュリティ対策</b>				
1	共通	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の安全安心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る	・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に書き込み不可にする） ※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	・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に書き込み不可にする） ※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
2	共通	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であり、会社のイントラネットに依拠することで対応している	・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であり、会社のイントラネットに依拠することで対応している
3	共通	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する	2)①-2を参照	2)①-2を参照
<b>D-5) i. セキュリティ対策のための仕組みの導入</b>				
4	開発	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ずる（セキュリティ・バイ・デザイン）	本UCにおいては2)③ D-2) i. -3と同じ	本UCにおいては2)③ D-2) i. -3と同じ
<b>P-5) i. セキュリティ対策のための仕組みの導入</b>				
5	提供	AIシステム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を講ずる（セキュリティ・バイ・デザイン）	本UCにおいては2)③ D-2) i. -3と同じ	本UCにおいては2)③ D-2) i. -3と同じ
<b>U-5) i. セキュリティ対策の実施</b>				
6	利用	AI提供者によるセキュリティ上の留意点を遵守する	本UCにおいては2)③ D-2) i. -3と同じ ※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	本UCにおいては2)③ D-2) i. -3と同じ ※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	利用	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う	※本UCについて、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学習済みAIモデルの特徴量に個人名は含まれていないため、対象外	※本UCについて、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学習済みAIモデルの特徴量に個人名は含まれていないため、対象外
<b>② 最新動向への留意</b>				
1	共通	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する	本UCにおいては2)③ D-2) i. -3と同じ ※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	本UCにおいては2)③ D-2) i. -3と同じ ※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>D-5) ii. 最新動向への留意</b>				
2	開発	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する	本UCにおいては2)③ D-2) i. -3と同じ ※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	本UCにおいては2)③ D-2) i. -3と同じ ※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>P-5) ii. 脆弱性への対応</b>				
3	提供	AIシステムに対する攻撃手法も数多く生まれているため、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する。また、脆弱性に対応することを検討する	本UCにおいては2)③ D-2) i. -3と同じ ※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	本UCにおいては2)③ D-2) i. -3と同じ ※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

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者	
		AI開発者	AI提供者	AI利用者	
No	分類	共通の指針/各主体に関する事項			
		本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	
<b>6) 透明性</b>					
各主体は、AIシステム・サービスの開発・提供・利用において、AIシステム・サービスを活用する際の社会的文脈を踏まえ、AIシステム・サービスの検証可能性を確保しながら、必要かつ技術的に可能な範囲で、ステークホルダーに対し合理的な範囲で適切な情報を提供することが重要である。					
<b>① 検証可能性の確保</b>					
1	共通	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、判断根拠等のログを記録・保存する	・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	・本番環境へのデプロイが行われるような変更内容についてのレビューを行う ・AIの利用ログを定期的にモニタリングする	※本UCに関してはAI利用者で実施することはない項目のため対象外
2	共通	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する	・モデルカードの記録は本番環境のAIモデルに変更を与える際には必須とし、その他AI開発者の判断でも記録を可能とする ・利用ログはエントリーシートの予測の都度記録される	※本UCに関してはAI開発者のみが実施することのため対象外	※本UCに関してはAI開発者のみが実施することのため対象外
<b>D-6) i. 検証可能性の確保</b>					
3	開発	AIの予測性能及び出力の品質が、活用開始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・向上を行う	本UCにおいては(6)①-1と同じ	本UCにおいては(6)①-1と同じ	本UCにおいては(6)①-1と同じ
<b>P-6) i. システムアーキテクチャ等の文書化</b>					
4	提供	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・サービス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する	・AIモデル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モデルの推論、出力）	・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)を文書化する	・AIサービスを用いた人材採用プロセスの文書化
<b>② 関連するステークホルダーへの情報提供</b>					
1	共通	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、以下について取りまとめた情報の提供及び説明を行う 【AIシステム・サービス全般】 ・AI を利用しているという事実及び活用している範囲 ・データ収集及びアノテーションの手法 ・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、限界及び提供先における適正/不適正な利用方法 ・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	・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  【説明する情報】 -データ収集及びアノテーションの手法 -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)	・AIシステム・サービス全般に係る以下の情報をAI利用者に向けて提供する  【説明する情報】 -AI を利用しているという事実 -データ収集及びアノテーションの手法 -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、提供先における適切/不適切な使用方法 -AIシステム・サービスの提供先や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	・AIシステム・サービス全般に係る以下の情報を応募者やキャリアコンサルに向けて提供する  【説明する情報】 -AIを利用しているという事実 ※必要に応じて、判断根拠に係る情報を提供
2	共通	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する	・AI提供者が実施するダイアログに必要に応じて参加する	・AI CoE主導で外部有識者を含むダイアログを実施している(本UC以外を含む)	・AI提供者が実施するダイアログに必要に応じて参加する
3	共通	加えて、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す	・AI提供者が実施するダイアログに必要に応じて参加する	本UCにおいては同上	・AI提供者が実施するダイアログに必要に応じて参加する
<b>D-6) ii. 関連するステークホルダーへの情報提供</b>					
4	開発	AI開発者は、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）情報を提供する	以下の項目で確認	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4a	開発	「人間中心」関連 ・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	・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（本UCにおいては(6)①-1と同じ）	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4b	開発	「安全性」関連 ・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	・リスクシナリオ及び対策の検討内容を整理（本UCにおいては2)①-3と同じ）	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4c	開発	「安全性」関連 ・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	※AI提供者で検討するため対象外	本UCにおいては(2)② P-2) ii. -5と同じ	※AI提供者で検討するため対象外
4d	開発	「安全性」関連 ・AIシステムの動作状況に関する情報並びに不具合の原因及び対応状況	・AIシステムの障害発生状況を対応方法と併せて記録する	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4e	開発	「安全性」関連 ・AIの更新を行った場合の内容及びその理由の情報	本UCにおいては(6)② D-6) ii. -4aと同じ	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外
4f	開発	「公平性」「プライバシー保護」「セキュリティ確保」関連 ・AIモデルで学習するデータの収集ポリシー、学習方法及び実施体制等	本UCにおいては(2)③ -1/2, D-2) i. -3/4と同じ	※AI開発者に係る項目のため対象外	※AI開発者に係る項目のため対象外

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの記事で、応募者に対して可否を判断する）

[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)

No	分類	共通の指針/各主体に関する事項	AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
			AI開発者	AI提供者	AI利用者
			本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>P-6) ii. 関連するステークホルダーへの情報提供</b>					
5	提供	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する	※AI提供者に係る項目のため対象外	以下の項目で確認	※AI提供者に係る項目のため対象外
5a	提供	「透明性」関連 ・AIを利用しているという事実、適切/不適切な使用方法等	※AI提供者に係る項目のため対象外	本UCにおいては6)②-1と同じ	※AI提供者に係る項目のため対象外
5b	提供	「安全性」関連 ・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提供者に係る項目のため対象外	・リスクシナリオ及び対策の検討内容を整理（本UCにおいては2)①-3と同じ）	※AI提供者に係る項目のため対象外
5c	提供	「人間中心」関連 ・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提供者に係る項目のため対象外	・AIシステムへの変更に対してレビューを行う（本UCにおいては6)①-1と同じ）	※AI提供者に係る項目のため対象外
5d	提供	「安全性」関連 ・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提供者に係る項目のため対象外	・AIサービス全体で発生するインシデント状況を対応方法と併せて管理する	※AI提供者に係る項目のため対象外
5e	提供	「安全性」関連 ・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提供者に係る項目のため対象外	本UCにおいては6)①-1と同じ	※AI提供者に係る項目のため対象外
5f	提供	「公平性」「プライバシー保護」「セキュリティ確保」関連 ・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	※AI提供者に係る項目のため対象外	本UCにおいては2)③ -1/2, D-2) i. -3/4と同じ	※AI提供者に係る項目のため対象外
<b>U-6) i. 関連するステークホルダーへの情報提供</b>					
6	利用	「公平性」「透明性」関連 ・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投入を行い、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する	※本UCについては、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	※本UCについては、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)	※本UCについては、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>③合理的かつ誠実な対応</b>					
1	共通	上記の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する	※本UCについては、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とする	※本UCについては、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とする	※本UCについては、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とする
2	共通	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定に準拠する	・開発に用いるオープンソースのライブラリを一覧化	・開発に用いるオープンソースのライブラリについて、商用ライセンスに問題が無いことを確認	※AI提供者及びAI開発者に係る項目のため対象外
3	共通	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する	※本UCはオープンソース化を行わないため対象外	※本UCはオープンソース化を行わないため対象外	※本UCはオープンソース化を行わないため対象外
<b>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</b>					
1	共通	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、必要な対応を講じる	・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ている（本UCにおいては2)①-3と同じ）	・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ている（本UCにおいては2)①-3と同じ）	・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ている（本UCにおいては2)①-3と同じ）



**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>7) アカウンタビリティ</b>				
各主体は、AIシステム・サービスの開発・提供・利用において、トレーサビリティの確保や「共通の指針」の対応状況等について、ステークホルダーに対して、各主体の役割及び開発・提供・利用するAIシステム・サービスのもたらすリスクの程度を踏まえ、合理的な範囲でアカウンタビリティを果たすことが重要である。				
<b>① トレーサビリティの向上</b>				
1	共通	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・遡求が可能な状態を確保する	本UCにおいては6)①-1と同じ	本UCにおいては6)①-1と同じ
<b>② 「共通の指針」の対応状況の説明</b>				
1	共通	「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う	本UCにおいては6)②-1と同じ	本UCにおいては6)②-1と同じ
1a	共通	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	・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ている（本UCにおいては2)①-3と同じ）	・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ている（本UCにおいては2)①-3と同じ）
1b	共通	「人間中心」関連 ・偽情報等への留意、多様性・包摂性、利用者支援及び持続可能性の確保の対応状況	上記に含むため割愛	上記に含むため割愛
1c	共通	「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	上記に含むため割愛	上記に含むため割愛
1d	共通	「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	上記に含むため割愛	上記に含むため割愛
1e	共通	「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	上記に含むため割愛	上記に含むため割愛
1f	共通	「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及びAIシステム等と連携する場合における発生しうるリスク及びその対応策	上記に含むため割愛	上記に含むため割愛
<b>D-7) i. AI提供者への「共通の指針」の対応状況の説明</b>				
2	開発	AI開発者は、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旨、その結果生じうるリスク等の情報提供及び説明を行う。具体的には以下の事項を周知する	・リスクシナリオ及び対策の検討内容を整理（本UCにおいては2)①-3と同じ）	・リスクシナリオ及び対策の検討内容を整理（本UCにおいては2)①-3と同じ）
2a	開発	「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）において含まれる可能性があるバイアスへの対応等	内容は3)①-1を参照	内容は3)①-1を参照
<b>P-7) i. AI利用者への「共通の指針」の対応状況の説明</b>				
3	提供	AI利用者に適正利用を促し、以下の情報をAI利用者に提供する	※AI提供者に係る項目のため対象外	適正利用の説明について、本UCにおいては6)②-1を参照
3a	提供	「安全性」関連 ・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	※本UCについては、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）	※本UCについては、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）
3b	提供	「安全性」関連 ・コンテキスト内学習による不適切なモデルの学習に対する注意喚起	※本UCについては、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）	※本UCについては、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）
3c	提供	「プライバシー保護」関連 ・個人情報を入力する際の留意点	※本UCについては、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）	※本UCについては、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>U-7) i. 関連するステークホルダーへの説明</b>				
4	利用	関連するステークホルダーの性質に応じて合理的な範囲で、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う	※本UCにおいては業務外利用者は直接利用しないため対象外	※本UCにおいては業務外利用者は直接利用しないため対象外
5	利用	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する	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外
6	利用	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、当該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す	AI提供者及びAI利用者による項目のため対象外	・AIサービスの利用目的・範囲を明確にした利用者向けのマニュアルにおいて、AIの活用を応募者に伝えること／判断根拠情報を不要に外部へ伝えないこと（不要な誤解を生まない／キャリアコンサルによる応募者への不適切な指導の誘発を防ぐため）を説明する。
7	利用	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う	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>③責任者の明示</b>				
1	共通	各主体においてアカウントビリティを果たす責任者を特定する	・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	・AI提供者の中でAIサービス管理者を特定する
<b>④関係者間の責任の分配</b>				
1	共通	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（ボランタリーコミットメント）等により、責任の所在を明確化する	AI提供者での検討に含まれるため割愛	・AI提供者が中心となって、AIサービス管理者及びAI利用者の役割・責任を定義する
<b>⑤ステークホルダーへの具体的な対応</b>				
1	共通	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、公表する（社会及び一般市民に対するビジョンの共有、並びに情報発信・提供を行うといった社会的責任を含む）	AI提供者が定める各ポリシーに準拠するため対象外	・AI CoE主導でAIガバナンスポリシーを定める ・プライバシーポリシー、セキュリティポリシー、情報管理規定、システム管理規定を定める
2	共通	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する	本UCにおいては2)①-5と同じ	本UCにおいては2)①-5と同じ
3	共通	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する	本UCにおいては2)①-5と同じ	本UCにおいては2)①-5と同じ
<b>⑥文書化</b>				
1	共通	上記に関する情報を文書化して一定期間保管し、必要なときに、必要なところで、入手可能かつ利用に適した形で参照可能な状態とする	・AIシステムの開発に伴う情報を文書化する ・AIモデルの検討内容はモデルカードに記録する	・AIサービス運営に必要なドキュメント類を適切に管理する
<b>D-7) ii. 開発関連情報の文書化</b>				
2	開発	AI開発者は、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する	本UCにおいては6)①-1と同じ	※AI開発者に係る項目のため対象外
<b>P-7) ii. サービス規約等の文書化</b>				
3	提供	AI利用者又は業務外利用者に向けたサービス規約を作成する	※AI提供者に係る項目のため対象外	※本UCについては、同一事業グループ内での利用のため、7)③-1に定めるポリシー・規定が内容をカバーする
4	提供	プライバシーポリシーを明示する	※AI提供者に係る項目のため対象外	※本UCについては、同一事業グループ内での利用のため、7)③-1に定めるポリシー・規定が内容をカバーする
<b>U-7) ii. 提供された文書の活用及び規約の遵守</b>				
5	利用	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・活用する	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外
6	利用	AI提供者が定めたサービス規約を遵守する	※AI利用者に係る項目のため対象外	・AI提供者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 ・AI利用者の初回利用時に説明を行う

**Case 採用AI**

機械学習モデル：XGBoost（エントリーシートの記事で、応募者に対して合否を判断する）

[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)

		AI開発部門	人材採用部門（採用AIチーム）	人材採用担当者
		AI開発者	AI提供者	AI利用者
No	分類	共通の指針/各主体に関する事項	本UCにおいて主体が実施している活動	本UCにおいて主体が実施している活動
<b>社会と連携して取り組むことが期待される事項</b>				
<b>8) 教育・リテラシー</b>				
各主体は、主体内のAIに関わる者が、AIの正しい理解及び社会的に正しい利用ができる知識・リテラシー・倫理感を持つために、必要な教育を行うことが期待される。また、各主体は、AIの複雑性、誤情報といった特性及び意図的な悪用の可能性もあることを勘案して、ステークホルダーに対しても教育を行うことが期待される。				
<b>①AIリテラシーの確保</b>				
1	共通	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る	・本AIサービスに関与するAI開発者に対しては、必要な知識・スキルを有しているか、アサイン前にチーフデータサイエンティスト及びリードエンジニアが確かめる	・本AIサービスに関与するAI提供者に対しては、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
<b>②教育・リスク</b>				
1	共通	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・リスクキリング等を検討する	※AI利用者に係る項目のため対象外	※AI利用者に係る項目のため対象外
1	共通	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する	※本UCにおいては業務外利用者は直接利用しないため対象外	※本UCにおいては業務外利用者は直接利用しないため対象外
<b>③ステークホルダーへのフォローアップ</b>				
1	共通	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う	※本UCにおいては業務外利用者は直接利用しないため対象外	※本UCにおいては業務外利用者は直接利用しないため対象外
<b>9) 公正競争確保</b>				
各主体は、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ることが期待される。				
<b>10) イノベーション</b>				
各主体は、社会全体のイノベーションの促進に貢献するよう努めることが期待される。				
<b>①オープンイノベーション等の推進</b>				
1	共通	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する	※特定UCに係るものではないため対象外	※特定UCに係るものではないため対象外
2	共通	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する	※特定UCに係るものではないため対象外	※特定UCに係るものではないため対象外
<b>D-10) i. イノベーションの機会創造への貢献</b>				
3	開発	AI開発者は、可能な範囲で以下の事項を実施し、イノベーションの機会創造に貢献することが期待される ・AIの品質・信頼性、開発の方法論等の研究開発を行う ・持続的な経済成長の維持及び社会課題の解決策が提示されるよう貢献する ・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を推進する ・社会全体へのAIに関する情報提供を行う	※特定UCに係るものではないため対象外	※特定UCに係るものではないため対象外
<b>②相互接続性・相互運用性への留意</b>				
1	共通	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する	※本UCにおいては外部AIサービスとの連携を行わないため対象外 ※本UCにおいては業務外利用者は直接利用しないため対象外	※本UCにおいては外部AIサービスとの連携を行わないため対象外 ※本UCにおいては業務外利用者は直接利用しないため対象外
1	共通	標準仕様がある場合には、それに準拠する	※本UCにおいては外部AIサービスとの連携を行わないため対象外 ※本UCにおいては業務外利用者は直接利用しないため対象外	※本UCにおいては外部AIサービスとの連携を行わないため対象外 ※本UCにおいては業務外利用者は直接利用しないため対象外
<b>③適切な情報提供</b>				
1	共通	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う	6)透明性を参照	6)透明性を参照