AI事業者ガイドライン（第1.2版）別添（付属資料）概要（案）

総務省
経済産業省

（令和8年●月●●日）

別添（付属資料）の位置づけ

- 本編では、事業者がAIの安全安心な活用を行い、AIの便益を最大化するために重要な「どのような社会を目指すのか（基本理念=why）」及び「どのような取組を行うか（指針=what）」を示しています
- 別添（付属資料）では、「具体的にどのようなアプローチで取り組むか（実践=how）」を示しています

本編（why, what） ▶ **別添（付属資料）(how)**

どのような社会を
目指すのか
（基本理念=why）

どのような取組を
行うか
（指針=what）

どのようなアプローチで
取り組むか
（実践=how）

「AI事業者ガイドライン」の構成

- 別添の記載内容は本編と対応しています。本編の読解や、具体的な行動を検討する際の解説書としてお使いください

	本編（why, what）		別添（付属資料）（how）	
主体共通	第1部 AIとは		1. 第1部関連 [AIについて]	A. AIに関する前提 B. AIによる便益/リスク
主体共通	第2部 AIにより目指すべき社会及び各主体が取り組む事項	A. 基本理念 B. 原則 C. 共通の指針 D. 広島AIプロセス「全てのAI関係者向けの広島プロセス国際指針」 E. AIガバナンスの構築	2. 第2部関連 [E.AIガバナンスの構築]	A. 経営層によるAIガバナンスの構築及びモニタリング B. AIガバナンスの事業者取組事例
主体別	第3部 AI開発者に関する事項	※広島AIプロセス「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」も含む	3. 第3部関連 [AI開発者向け]	A. 「第3部 AI開発者に関する事項」の解説 B. 「第2部」の「共通の指針」の解説 C. 広島AIプロセス「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」
主体別	第4部 AI提供者に関する事項		4. 第4部関連 [AI提供者向け]	A. 「第4部 AI提供者に関する事項」の解説 B. 「第2部」の「共通の指針」の解説
主体別	第5部 AI利用者に関する事項		5. 第5部関連 [AI利用者向け]	A. 「第5部 AI利用者に関する事項」の解説 B. 「第2部」の「共通の指針」の解説
その他参考資料			6. 「AI・データの利用に関する契約ガイドライン」を参照する際の主な留意事項について 7. チェックリスト・ワークシート 8. 主体横断的な仮想事例 9. 海外ガイドライン等の参照先	

別添1. 第1部関連　記載内容

- 別添1では、AIに関する前提や、AIによる便益/リスクについて具体的に解説しています
- これらの解説を通じ、本ガイドラインの記載内容のより深い理解につなげることを目的としています

A. AIに関する前提

- AIの学習及び利用の流れ
  - 一般的なAIの学習・利用の流れの把握
- AIシステム概要
  - AIシステムのスコープ
- AIの開発から利用までのバリューチェーン
  - 一般的なAI活用の流れにおける主体の対応
- AIシステム・サービスの例
  - 代表的なAIシステム・サービス例と
    それにかかわる各主体を具体化
- AI事業者のパターン
  - 事業活用時のAIバリューチェーンを具体化
- データ提供者について
  - 本ガイドラインでは対象外のデータ提供者の
    定義等

B.AIによる便益/リスク

- AIによる便益
  - 主に利益を享受する最終利用者に焦点を
    当てて記載
- AIによるリスク
  - 代表的な事例（想定を含む）を記載
  - リスクの体系的な分類案（技術的リスクと
    社会的リスク）を記載
  - 各リスク例に対応する主な共通の指針と、
    事業者における対策の例を記載

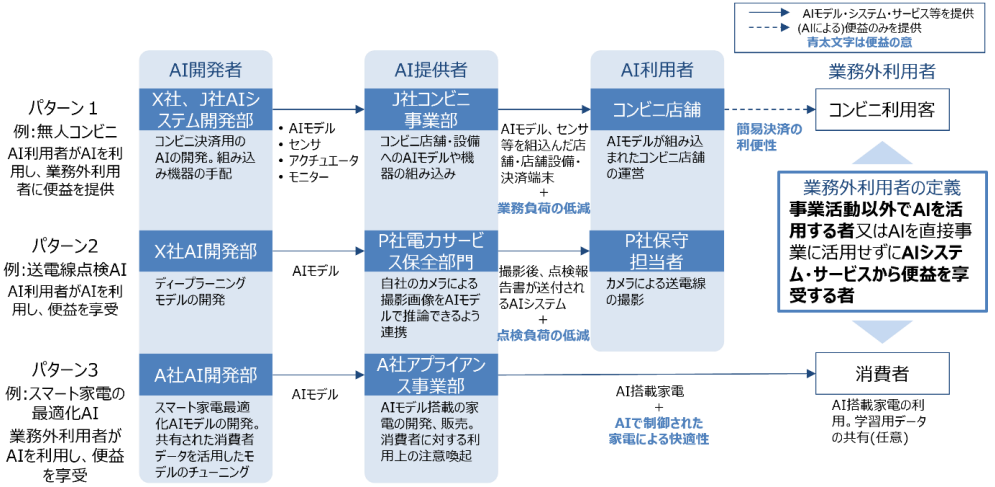
別添1 A. AIに関する前提　主な記載内容

- AIそのものやAIの活用場面、各主体の役割などについて更に理解を深めていただけるよう、「AIシステム」「AI事業者」「データ提供者」などの用語を具体例を交えつつ解説しています（以下は内容の抜粋）

AIシステム・サービスの例　(抜粋)

ケース名	活用 AI	概要	AI 開発者	AI 提供者	AI 利用者	業務外利用者
採用 AI	テキスト解析	A 社グループのグローバル各社における人材採用部門が、**エントリーシートの書類選考を判断する際の参考情報として使用される AI サービス**である。 A 社 AI 開発部門は、AI 利用者である A 社人材採用部門（海外グループ企業を含む）より過去のエントリーシートデータ及び合否判定（内定の判定）結果を受領し、機械学習（分類モデル）で合否判定を支援する AI モデルを作成している。	A 社（開発部門）	A 社（システム部門、人材開発部門）	A 社グループ（人材採用部門）	採用申込者
無人コンビニ	画像解析	全国のコンビニエンスストアチェーンを経営する J 社が提供する**画像認識 AI を活用した無人コンビニ（店内の客が商品を取るだけで AI が代金を計算し、店外に出る際に電子マネー等で一括決済ができるコンビニ）**である。当 AI サービスには X 社で開発された無人コンビニ向けの AI システムを搭載している。	X 社	J 社（AI システム開発部及びコンビニ事業部）	コンビニ店舗	コンビニ利用客
がん診断 AI	テキスト・画像解析	マルチモーダル学習を使用しており、**「本人の病歴・遺伝等に係る情報（データ 1）」及び「内視鏡画像（データ 2）」を取込み、内視鏡での診察中にリアルタイムにがんの可能性が高い部分をハイライト**する。医師は出力画像を観察して、がんの可能性があるか判断する。 A 社が AI を開発しつつ、がん診断 AI システムを医療機関に提供している。	A 社（AI 開発部門）	A 社（医療 IT サービス部門）	医療機関（システム部門及び消化器内科）	受診患者

AI事業者のパターン



別添1 B. AIによる便益/リスク（AIによる便益）主な記載内容

- 便益を享受する最終利用者に焦点を当て、業種や業務ごとにAIによる便益を整理しています

	開発	マーケティング	販売	物流・流通	顧客対応	法務	ファイナンス	人事
従来から存在する便益の例	コード検証、ドキュメント作成の自動化	広告用メールの自動配信	受注後の対応メール等の自動発信	需要予測に基づく生産・在庫数最適化	チャットボットによる自動対応	翻訳	財務諸表の自動作成	給与計算等の自動化
生成AIで更に向上（一部AIエージェントやフィジカルAIで更に向上）	類似コード・データの抽出・検証	データに基づいたパーソナライゼーション広告	チャネル別、ニーズ別の売上予測	配送ルート最適化	顧客の成約・解約率予測	法務文章のレビュー	過去実績にもとづいた将来予測、不正検知	職務経歴書等に基づいた人材需要マッチング
生成AI、AIエージェント、フィジカルAI特有の便益の例	学習データの生成	販売促進(マーケティング素材・キャッチコピー等)の自動作成	営業トークスクリプトの自動作成	物流条件交渉のアシスタント	対応内容の自動生成、要約	規定に基づいた契約書ドラフトの自動生成	経費精算アシスタント（自動仕訳、申請レビュー、証憑取得等）	AI採用アシスタント（面接・評価）
	コーディングアシスタント（コード生成、不具合の自動修正等）	投稿から分析までを自律的に行うSNS運用エージェント	店頭ロボットによる自動接客・販売	自律搬送ロボット・ドローンによる自動配送	過去の問合わせ内容に基づいたFAQ自動作成	類似事例検索及び重要判例の自動要約	投資レポート・市場分析の自動生成	パーソナライズされたキャリアプラン提案

別添1 B. AIによる便益/リスク（AIによるリスク）主な記載内容

- AIによるリスクを、事業者ができる限り網羅的に把握し対策を検討できるよう、体系的に整理しました

・下表はAIのリスクを網羅したものではなく、想定に基づく事案も含んでおり、あくまで一例として認識することが期待される
・下表には政府等の公的機関も含めた社会全体での対応・議論が必要となるリスクも含まれる

大分類	中分類	リスク例
技術的リスク（＝主にAIシステム特有のもの）	学習及び入力段階のリスク	データ汚染攻撃等のAIシステムへの攻撃
	出力段階のリスク	バイアスのある出力、一貫性のない出力等
		ハルシネーション等による誤った出力
	事後対応段階のリスク	ブラックボックス化、判断に関する説明の不足
社会的リスク（＝既存のリスクがAIにおいても発生又はAIによって増幅するもの）	倫理・法に関するリスク	個人情報の不適切な取扱い等
		生命等に関わる事故の発生
		差別的出力
		過度な依存
		悪用
	経済活動に関するリスク	知的財産権等の侵害
		金銭的損失
		機密情報の流出
		労働者の失業
		データや利益の集中
		資格等の侵害
	情報空間に関するリスク	偽・誤情報等の流通・拡散
		民主主義への悪影響
		フィルターバブル及びエコーチェンバー現象
		多様性・包摂性の喪失
		バイアス等の再生成
	環境に関するリスク	エネルギー使用量及び環境の負荷

別添1 B. AIによる便益/リスク（AIによるリスク）主な記載内容

- さらに、事業者におけるリスクへの対策の検討に結び付けるべく、各リスクに対応する主な共通の指針と、事業者における対策の例を記載しました*1

・下表はAIのリスクを網羅したものではなく、想定に基づく事案も含んでおり、あくまで一例として認識することが期待される
・下表には政府等の公的機関も含めた社会全体での対応・議論が必要となるリスクも含まれる

	リスク例	関連する共通の指針	「共通の指針」に加えて主体毎に重要となる事項 第3部 AI開発者	第4部 AI提供者	第5部 AI利用者
技術的リスク	データ汚染攻撃等のAIシステムへの攻撃	5）セキュリティ確保	i. セキュリティ対策のための仕組みの導入 ii. 最新動向への留意	i. セキュリティ対策のための仕組みの導入 ii. 脆弱性への対応	i. セキュリティ対策の実施
	バイアスのある出力、一貫性のない出力等	1）人間中心 ①人間の尊厳及び個人の自律 ③偽情報等への対策			
	ハルシネーション等による誤った出力	2）安全性	i. 適切なデータの学習 ii. 人間の生命・身体・財産、精神及び環境に配慮した開発 iii. 適正利用に資する開発	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 ii. 適正利用に資する提供	i. 安全を考慮した適正利用
		3）公平性	i. データに含まれるバイアスへの配慮 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	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	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
		8）教育・リテラシー			
	ブラックボックス化、判断に関する説明の不足	6）透明性	i. 検証可能性の確保 ii. 関連するステークホルダーへの情報提供	i. システムアーキテクチャ等の文書化 ii. 関連するステークホルダーへの情報提供	i. 関連するステークホルダーへの情報提供
		7）アカウンタビリティ	i. AI提供者への「共通の指針」の対応状況の説明	i. AI利用者への「共通の指針」の対応状況の説明	i. 関連するステークホルダーへの説明

*1：主体ごとの対応策や具体的な手法等については、本編第３部～第５部の該当部分ならびに別添３～５の該当部分を参照してください

別添2. 第2部 [E.AIガバナンスの構築]関連

- 別添2では、AIガバナンスの構築のための「行動目標」、「実践のポイント」及びそれに対応する仮想的な「実践例」、実際の企業等の取組事例を掲載しています
- 各事業者が、AIガバナンスの構築のための検討を、具体例を参考にしつつ実施できるようにしています

A.経営層によるAIガバナンスの構築及びモニタリング

- 行動目標
  - 一般的かつ客観的な目標を記載
- 実践のポイント
  - AIガバナンスに関連する国内外ガイドラインやISO等の要素も取り入れ
- 実践例
  - 仮想事例に基づく事例を記載
  - 生成AI等の最新動向への対応事例も織り込み

B.AIガバナンスの構築に関する実際の取組事例

- 実際の取組事例
  - AIガバナンスの取組事例について、11団体分（10企業、1自治体）をコラム化
  - AIガバナンスを推進するにあたり、多くの企業がつまづく観点を記載

コラム記載企業一覧（掲載順）

- 株式会社ABEJA
- NECグループ
- 東芝グループ
- パナソニックグループ
- 富士通グループ
- ソフトバンク株式会社
- NTT DATA
- Ubie株式会社
- 神戸市
- IBM
- Amazon Web Services

別添2 A. 経営層によるAIガバナンスの構築とモニタリング 主な記載内容

- ガバナンスの行動目標の意義を理解し活用することで、AIガバナンスの構築に役立ちます

分類	行動目標 ※「3-1-1」のように更に細分化されているものもあり
1. 環境・リスク分析	1-1 便益/リスクの理解 1-2 AIの社会的な受容の理解 1-3 自社のAI習熟度の理解
2. ゴール設定	2-1 AIガバナンス・ゴールの設定
3. システムデザイン	3-1 ゴールと乖離の評価及び乖離対応の必須化 3-2 AIマネジメントの人材のリテラシー向上 3-3 各主体間・部門間の協力によるAIマネジメント強化 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
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 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 4-3 AIガバナンスの実践状況の積極的な開示の検討
5. 評価	5-1 AIマネジメントシステムの機能の検証 5-2 社外ステークホルダーの意見の検討
6. 環境・リスクの再分析	6-1 行動目標1-1～1-3の適時の再実施

別添2 A. 経営層によるAIガバナンスの構築とモニタリング
構成

- 各「行動目標」に対し、「実践のポイント」及び「実践例」を整理しています
- 採用するAIの種類やリスクの程度に応じて参照することで、AIガバナンスの検討が可能となります

別添　記載内容

行動目標1-1【便益/リスクの理解】：
各主体は、経営層のリーダーシップの下、AIの開発・提供・利用の目的を明確化したうえで、AIから得られる便益だけではなく意図しないリスクがあることについて、各主体の事業に照らして具体的に理解し、これらを経営層に報告し、経営層で共有し、適時に理解を更新する。

（実践のポイント）
[実践のポイント]
各主体は、経営層のリーダーシップの下、以下に取り組む。
- 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
- 自社の事業に結びつく形で、「便益」及び意図せざるものを含めた「リスク」を具体的に理解
- 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便益を確保、リスクを削減
- 迅速に経営層に報告/共有する仕組みを構築

・
・
・

[実践例]
【実践例ｉ：便益・リスクの把握】
　各主体は、経営層のリーダーシップの下（担当役員又は現場に一任するのではなく、経営層自らが主導することを通じて実施することも含む、以下同様）、便益だけではなくリスクについても検討し、
・
・
・

解説

行動目標
- 事業者が取り組むことが重要となる一般的かつ客観的な目標を提示
- 各事業者の方針検討の際の材料となる

実践のポイント
- 上記の行動目標の実行のために重要となる事項や留意点を要約
- 各事業者が具体的な取組内容を検討する際の材料となる

実践例
- 仮想的な実践例を記載
- 具体的な取組イメージを持つことで、各事業者が行動につなげやすくする

別添3～5. 各主体向け 主な記載内容

- 別添3～5では、各主体向けに、本編の詳細な解説を掲載しています
- 本編と合わせて活用することで、本編の内容に関する具体的なアプローチを参照し、検討することができます

A. 本編「第3～5部」の解説

B.本編「第2部」の「共通の指針」の解説

C.広島AIプロセス「高度なAIシステムを開発する組織向けの広島プロセス国際行動規範」
※別添3. AI開発者向けのみ

- **ポイント**
  - 本編記載事項に加え、重要となる観点を補足
- **具体的な手法**
  - 他のガイドライン等を参照しつつ、具体的に解説
- **参考文献**
  - 参考となる他の文献を記載
    （例）
    - デジタル庁「データ品質ガイドブック（β版）」（2021年6月）
    - 国立研究開発法人産業技術総合研究所「機械学習品質マネジメントガイドライン 第4版」（2023年12月）
    - AIプロダクト品質保証コンソーシアム 「AIプロダクト品質保証ガイドライン 2025.04版」（2025年4月）
    - NIST, “AI Risk Management Framework Playbook” （2023年1月）
    - 欧州評議会「人権、民主主義及び法の支配の観点からのAIシステムに関するリスク・影響評価手法（HUDERIA）」（2024年11月）

別添3～5. 各主体向け構成

- 各主体の重要事項に対して、「ポイント」、「具体的な手法」、「参考資料」を以下の構成で掲載しています

別添　記載内容

本編の
第3~5部に
対応する解説

本編 第2部に
対応する解説
「共通の指針」に
関する各主体の
具体的なアプローチ

A. 本編「第 3 部　AI 開発者に関する事項」の解説

[本編の記載内容 （再掲）]
データ前処理時
D-2） i. 適切なデータの学習

- ✧ 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（「2） 安全性」、「4） プライバシー保護」、「5） セキュリティ確保」）
- ✧ 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（「2） 安全性」、「5） セキュリティ確保」）

[ポイント]
AI モデルの質の向上のために、AI 開発者は、AI の学習等に用いるデータの質に留意することが重要となる。

- 利用する AI の特性及び用途を踏まえ、AI の学習等に用いるデータの質（正確性及び完全性等）に留意する[87]
- また、AI によりなされる判断は、事後的に精度が損なわれたり、低下したりすることが想定されるため、想定される権利侵害の規模、権利侵害の生じる頻度、適用できる技術水準、精度を維持するためのコスト等を踏まえ、あらかじめ精度に関する基準を定めておくことが期待される。精度が当該基準を下回った場合には、データの質に留意して改めて学習させる
- なお、ここで言う「精度」には、AI が倫理的に正しい判断を行っているか（例えば、AI が暴力的な表現を行っていないか、ヘイトスピーチ等を行っていないか 等）も含まれる

[具体的な手法]

- データに個人情報、機密情報、著作権等の権利又は法律上保護される利益に関係するものが含まれていないか、確認を実施
  - ➢ 固有表現抽出

[参考文献]
1. 国立研究開発法人産業技術総合研究所「機械学習品質マネジメントガイドライン 第 4 版」（2023 年 12 月）
2. NIST, “AI Risk Management Framework Playbook”（2023 年 1 月）

解説

本編の記載内容を再掲

ポイント
- 本編記載事項に加え、重要観点を補足

具体的な手法
- 他のガイドライン等を参照しつつ具体的に解説

参考文献
- 「ポイント」や「具体的な手法」の参照元を記載

別添6.「AI・データの利用に関する契約ガイドライン」を参照する際の主な留意事項　概要

- 別添6では、2018年6月に初版が策定・公表された「AI・データの利用に関する契約ガイドライン」について、AIの開発・利用に関する状況の変化及び新たな技術の進歩に伴う、2026年3月時点での参照時の留意点を掲載しています
- 契約を通じて当事者間の権利及び義務を明確に定めることで、AIに関する取引を円滑に進め、これらに伴う無用な紛争を予防できます

AI・データの利用に関する契約ガイドライン

＋

諸外国の動向

新技術の台頭

AIの利用・開発に関する契約チェックリスト

別添6「AI・データの利用に関する契約ガイドライン」を参照する際の主な留意事項

契約ガイドライン公表後の状況の変化を考慮すべき事項として、以下の内容等について記載

(1)契約モデルの多様化について
- AIの利用・開発に関する契約は大きく以下の3類型に分類でき、それぞれの類型に応じて、留意するべき契約事項や交渉上の論点も異なる。
- 類型１：汎用的AIサービス利用型
- 類型２：カスタマイズ型
- 類型３：新規開発型

(2)複雑なバリューチェーンの下でのリスク分配について
- バリューチェーンの多様化・複雑化に伴う個別の状況に応じた責任分配の在り方の検討

(3)責任分界とアカウンタビリティについて
- 事故発生リスクとの関係では、契約上、①新たな類型のリスクの整理、②合理的な説明の実施、③客観的な根拠の提示等が重要となる

別添7. チェックリスト・ワークシート 概要

- 別添7では、AIによるリスクを抑えつつ便益を享受するための取組の立案・実践を確実に推進できるよう、「チェックリスト」及び「具体的なアプローチのためのワークシート」を用意しています

本編・別添1～5（付属資料） » **チェックリスト** » **具体的なアプローチ検討のためのワークシート**

本編・別添を読んでAIガバナンスの重要性や、各事業者に期待されることを**理解する**

「チェックリスト」を活用し、本編・別添についての各主体の**取組(What)を確認**する

「具体的なアプローチ検討のためのワークシート」を使用し、各事業者の**具体的なアプローチ(How) を検討する**

※ 各事業者の事業内容やAIポリシー/AI規定等の状況に合った独自のチェックリスト・ワークシートを作成・更新の上、有効活用することも重要となります（ワークシートについては掲載されている項目を必ずしも全て採用する必要はなく、自社に必要な項目を判断の上、活用する事が有効です）。

別添7. チェックリスト 活用方法

- 「別添7 Aチェックリスト[全主体向け]」は、全ての事業者が各自の取組状況の概観を確認するためのものです
- また、必要に応じ「別添7 Bチェックリスト[第2部D. 広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」]」もご活用ください

別添7 A　チェックリスト［全主体向け］　令和8年●月●●日

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

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及び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？
- □ 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、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タ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

別添7 B　チェックリスト（案）　令和8年●月●●日

第2部D. 広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」

本チェックリストは、AI事業者ガイドライン「第2部D.広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」」の項目です。取組事項のチェックにご活用ください

※①～⑪については、適時適切に、適切な範囲で、適用されるべきである。また、⑫については、従うべきである。

チェック項目

- □ ①AI 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、適切な措置を講じているか？
- □ ②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、緩和しているか？
- □ ③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しているか？
- □ ④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んでいるか？
- □ ⑤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、開示しているか？
- □ ⑥ 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
- □ ⑦技術的に可能な場合は、電子透かしやその他の技術等、ユーザーが 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、導入しているか？
- □ ⑧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？
- □ ⑨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）に対処するため、高度な AI システムの開発を優先しているか？
- □ ⑩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しているか？
- □ ⑪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しているか？
- □ ⑫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？

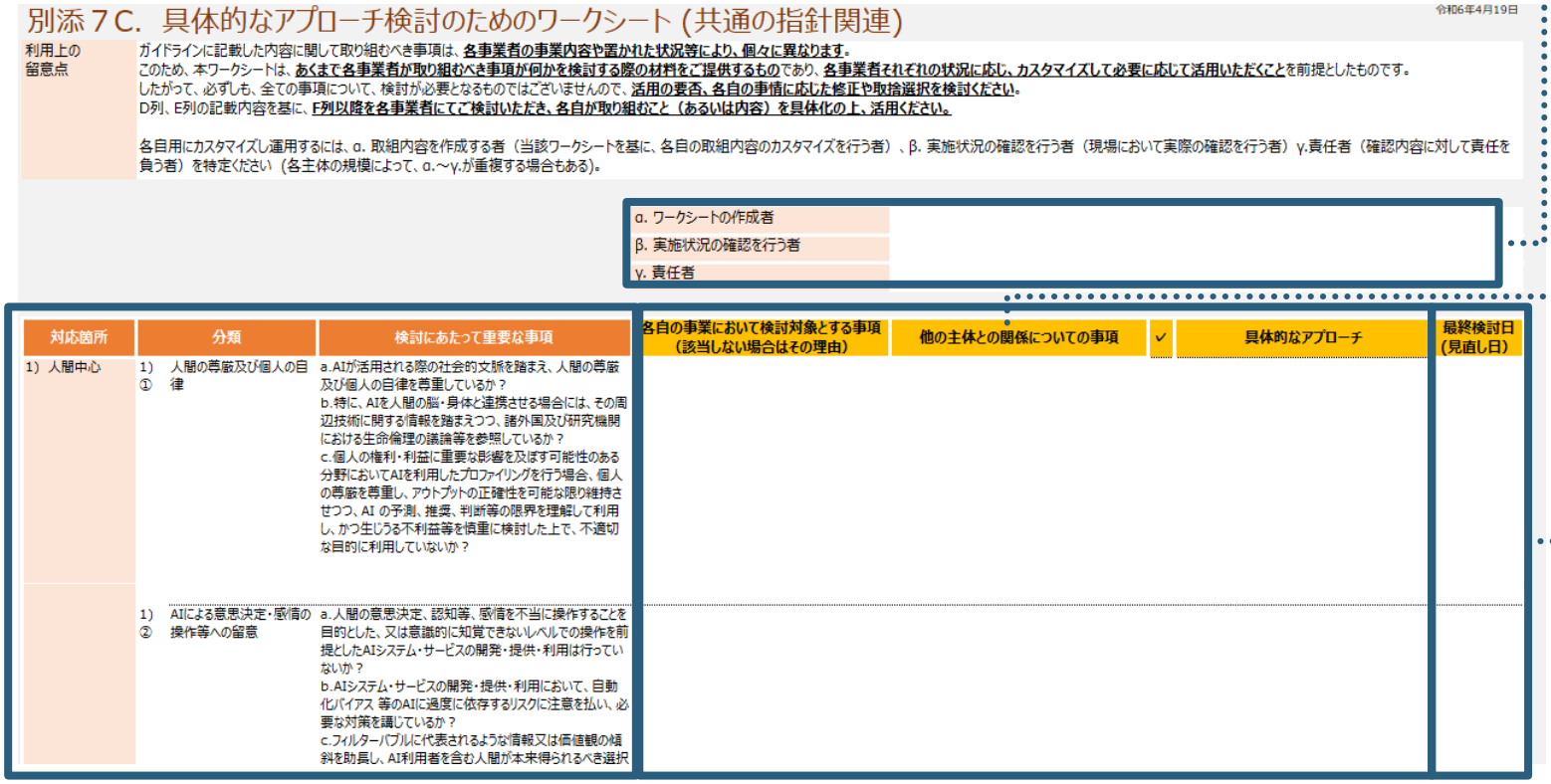
検討には「具体的なアプローチ検討のためのワークシート」をご活用ください

- チェック項目は本編の要約を記載
- チェックすることで、**各自の取組状況を概観**

- 具体的な実践内容の検討に、**「別添7C 具体的なアプローチのためのワークシート」**を活用
(活用方法次頁)

別添7. 具体的なアプローチ検討のためのワークシート 活用方法

- 本ガイドラインの記載内容に関して、具体的なアプローチを検討する際に重要となる事項を記載しています
- 事業者の事業内容及び置かれた状況等に応じ、各自でカスタマイズして活用することを前提としています



別添７C. 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)

令和6年4月19日

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、E列の記載内容を基に、**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。**

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	

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）① 人間の尊厳及び個人の自律	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？					
	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					

- 活用の前に、**実施責任者等を決定**する
- 事業者の事業内容や置かれた状況等に応じ、**各自でカスタマイズして活用**する
- 「見直し日」も検討・記載することで**定期的な更新**を行う
- 具体のアプローチを検討する際に重要になる事項が記載されており、事業者が**各自でカスタマイズする際のリファレンス**となる

別添8. 主体横断的な仮想事例
概要

- 別添8.では、本ガイドラインに沿って、AI開発者、AI提供者、AI利用者が重要事項の検討を行った場合の「主体横断的な仮想事例」を掲載しています
- 本ガイドラインの内容を実際に落とし込む際の具体的なイメージをつけ、各主体間での連携が重要になるポイントを明確化することが可能となります

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

			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
1）人間中心 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					
①人間の尊厳及び個人の自律					
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	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)	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	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	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	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
②AIによる意思決定・感情の操作等への留意					
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	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	本ケースに関しては、2)①-3と同じ論点になる

別添9. 海外ガイドライン等の参照先概要

- 別添9.では、本ガイドラインにおいて参考とした海外ガイドラインについて記載するとともに、それらの参照箇所を整理しています

参考とした主なガイドライン等

- Advancing accountability in AI（AIにおけるアカウンタビリティの高度化）（2023年2月、OECD）
- Hiroshima AI Process Comprehensive Policy Framework（広島AIプロセス）（2023年12月、G7）
- Artificial Intelligence Risk Management Framework（AI RMF 1.0）（2023年1月、NIST）
- CYBERSECURITY FRAMEWORK（CSF）（2018年4月、NIST）
- Blueprint for an AI Bill of Rights（AI権利章典）（2022年10月、THE WHITE HOUSE）
- Artificial Intelligence Act（EU AI Act）（2024年8月、EU）
- Guidelines for secure AI system development（セキュアAIシステム開発ガイドライン）（2023年11月、NCSC）
- ETHICS GUIDELINES FOR TRUSTWORTHY AI（信頼性を備えたAIのための倫理ガイドライン）（2019年4月、EU）
- Guidelines for privacy impact assessment（PIAガイドライン）（ISO）
- Recommendation on the Ethics of Artificial Intelligence（AI倫理勧告）（2021年11月、UNESCO）

AI事業者ガイドライン本編		AI事業者ガイドライン別添		①OECD AIにおけるアカウンタビリティの高度化	②G7 広島AIプロセス	③米国 NIST AI RMF 1.0
はじめに	-	別添.はじめに	-			
第１部 AIとは	-	別添1.第1部関連	A.AIに関する前提	AI terms & concepts (OECD AI Principles overview)	高度なAIシステムを開発する組織向けの広島プロセス国際行動規範（2023年10月, 広島AIプロセスに関するG7首脳声明）	
			B.AIによる便益/リスク			
第２部 AIにより目指すべき社会及び各主体が取り組む事項	A.基本理念	-	-			
	B.原則	-	-			
	C.共通の指針	-	-		概要版「各主体に共通の指針」にて広島AIプロセス包括的政策枠組み(2023年12月, 広島AIプロセスG7デジタル・技術閣僚声明)を引用	・3.4 Accountable and Transparent ・5.3 Measure 概要版「各主体に共通の指針」にて引用
	D.広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」	-	-		広島AIプロセス包括的政策枠組み II. 全ての AI関係者向け及び高度な AI システムを開発する組織向けの広島プロセス国際指針(2023年12月, 広島AIプロセスG7デジタル・技術閣僚声明)	
	E.ガバナンスの構築	別添2.「第2部E.AIガバナンスの構築」 関連	A.経営層によるAIガバナンスの構築及びモニタリング			・5.3 Measure ・3.5 Explainable and Interpretable(以下を引用) Four Principles of Explainable Artificial Intelligence (Draft)(2020年8月, NIST) ・2. Four Principles of Explainable AI
			B.AIガバナンスの構築に関する実際の取組事例			