

# AI原則実践のための ガバナンス・ガイドライン

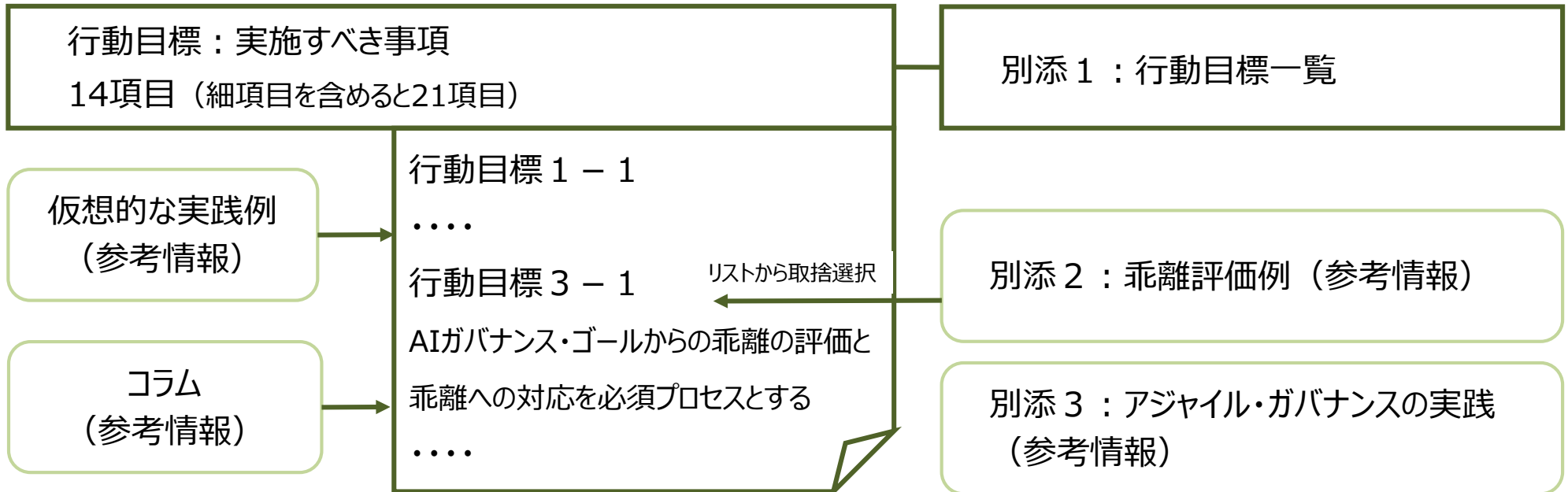
## ver. 1.0

2021年7月9日

経済産業省 商務情報政策局 情報経済課



- 法的拘束力のない分野横断的なガイドライン（通称：AIガバナンス・ガイドライン）
  - AIの社会実装の促進に必要なAI原則の実践を支援すべく、事業者が実施すべき行動目標を提示し、それぞれの行動目標に対応する仮想的な実践例とAI原則からの乖離評価例を参考情報として例示
  - AIシステムの開発・運用に関わる事業者等の取引等で広く参照されることや、AI原則の実践に関するステークホルダーの共通認識の形成を通じて、各社の自主的な取り組みを後押し



参考：我が国のAIガバナンスの在り方 ver. 1.1

- 「リスク評価やマネジメントのあり方も含めたガイドラインを提供することで、AIの社会実装をより促進できると考えられる。」

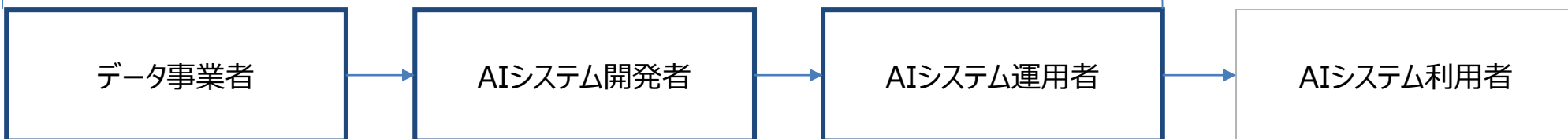
- **AIシステム**：OECDの定義を参考にしつつ、機械学習にフォーカスした定義を採用。ただし、機械学習よりも広義のAIシステムでも参照されることが望ましい。
- **AI事業者**：AIシステム開発者、運用者、データ事業者に区分。データ事業者は一部の行動目標のみ。

**AIシステム**：深層学習を含む様々な方法からなる、教師あり、教師なし、強化学習を含む機械学習アプローチを用いたシステムであって、人間が定義した特定の目的のために、現実又は仮想環境に影響を与えるような予測、助言、決定を行う性能を有するシステム。このAIシステムは設計次第で様々な自律の程度で動作する。このAIシステムには、ソフトウェアだけではなく、ソフトウェアを要素として含む機械も含まれる。

人間の判断を代替するものであって、利用者から判断過程が見えにくいソフトウェア等については、機械学習アプローチを用いていない場合であっても、必要に応じて本ガイドラインを参照することが期待される。

広義のAI

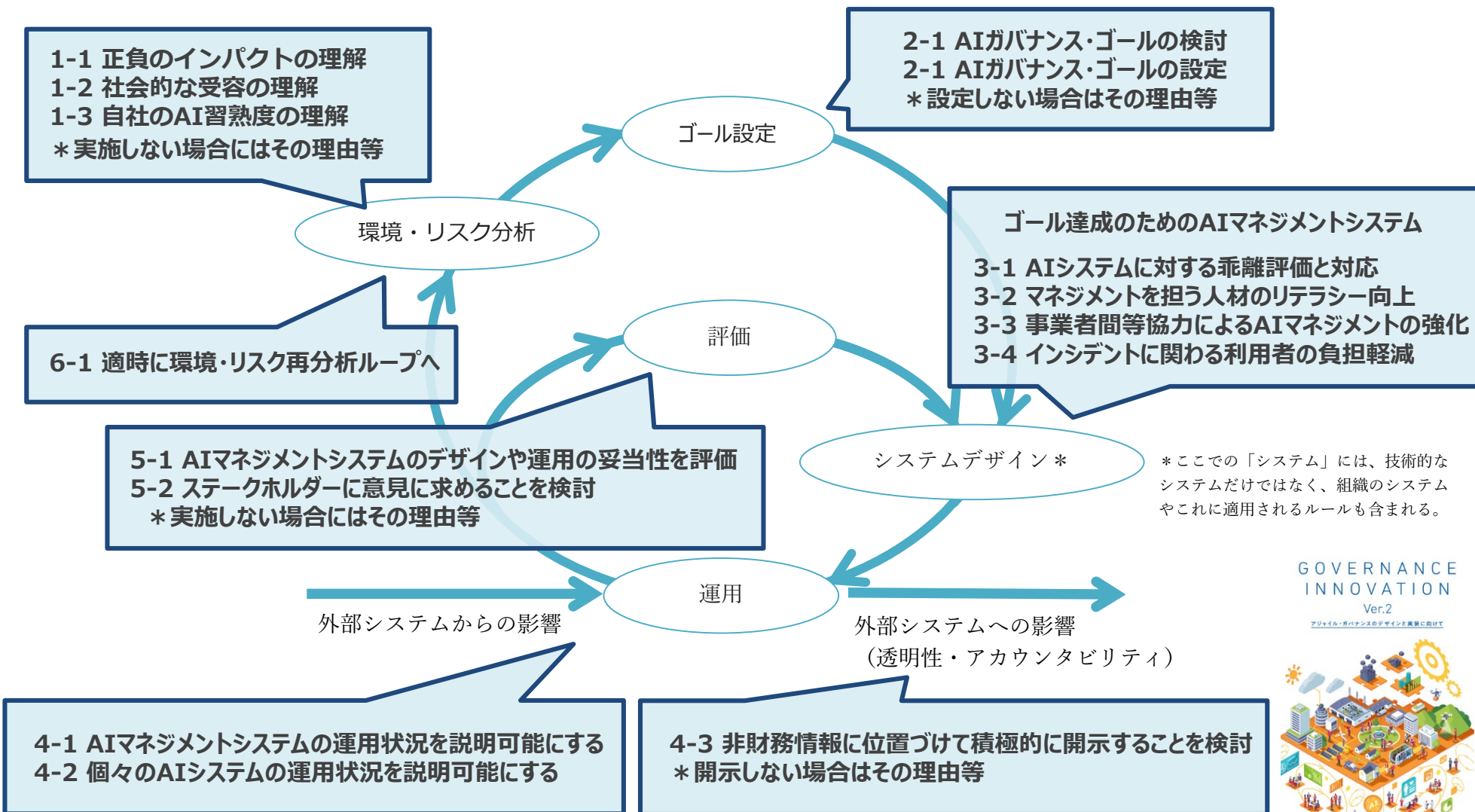
**AI事業者**（1つのカテゴリだけに排他的に区分されるわけではない）



**参考：OECDのAIシステムの定義**

An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy.

- マネジメント体制の整備とリスク管理のための行動目標をアジャイル・ガバナンスの枠組みに沿って提示



# 行動目標の概要（環境・リスク分析）

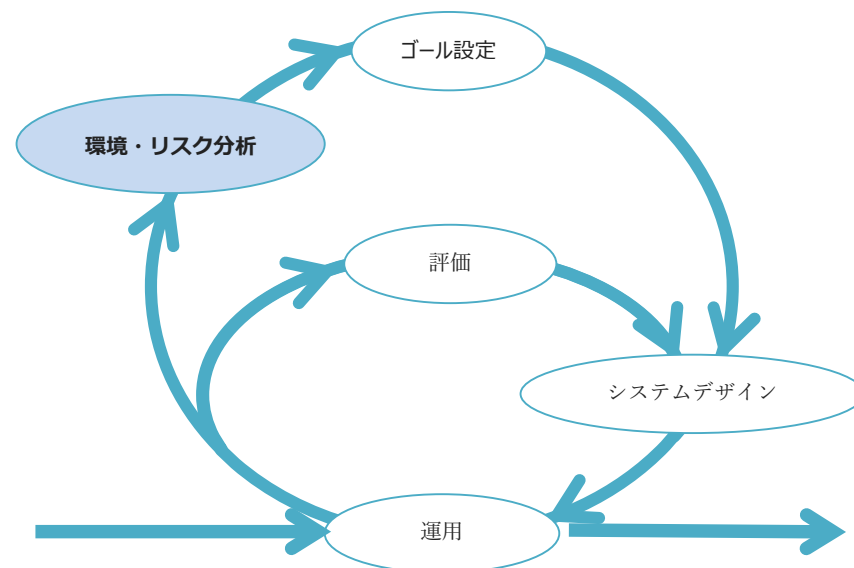
関連箇所：C-1.

- ゴール設定のために、AIシステムがもたらしうる正負のインパクト、AIシステムの開発や運用に関する**社会的受容**、**AI習熟度**（AIシステムの開発・運用時に求められる準備がどれだけできているのか）を理解すべき

行動目標 1 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、AIシステムから得られる正のインパクトだけではなく意図せざるリスク等の負のインパクトがあることも理解し、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べきである。

行動目標 1 - 3 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、行動目標 1 - 1、1 - 2 の実施を踏まえ、自社の事業領域や規模等に照らして負のインパクトが軽微であると判断した場合を除き、自社のAIシステムの開発・運用の経験の程度、AIシステムの開発・運用に関与するエンジニアを含む従業員の人数や経験の程度、当該従業員のAI技術及び倫理に関するリテラシーの程度等に基づいて、自社のAI習熟度を評価し、適時に再評価すべきである。負のインパクトが軽微であると判断し、AI習熟度の評価をしない場合には、その理由等をステークホルダーに説明できるようにしておくべきである。

行動目標 1 - 2 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、本格的なAIの提供に先立ち、直接的なステークホルダーだけではなく潜在的なステークホルダーの意見に基づいて、社会的な受容の現状を理解すべきである。また、本格的なAIシステムの運用後も、適時にステークホルダーの意見を再確認するとともに、新しい視点を更新すべきである。

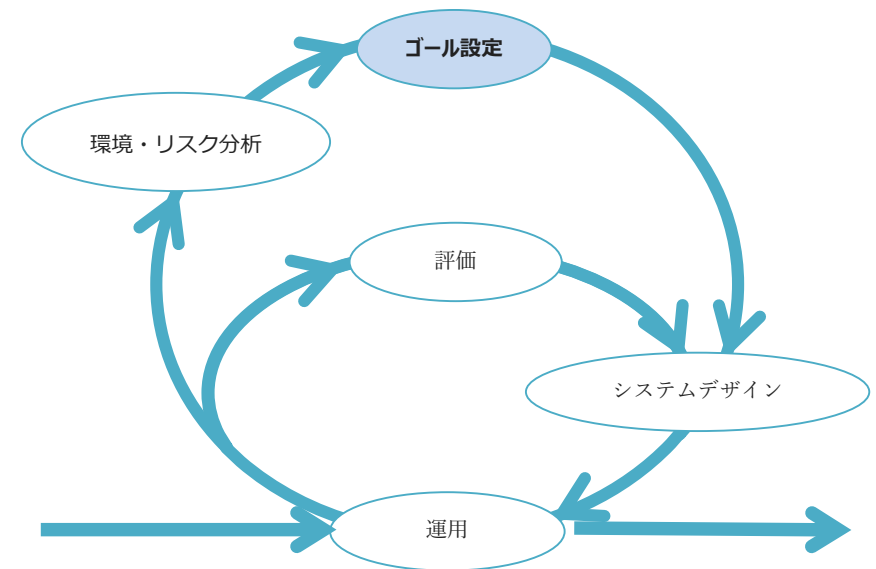


- システムデザインの羅針盤となる、**AIガバナンス・ゴールを設定するか否かを検討**し、負のインパクトが軽微であることを理由にゴールを設定しない場合には、その理由等をステークホルダーに説明できるようにしておくべき

行動目標 2 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、「人間中心のAI社会原則」を踏まえ、AIシステムがもたらしうる正負のインパクト、AIシステムの開発や運用に関する社会的受容、自社のAI習熟度を考慮しつつ、設定に至るプロセスの重要性にも留意しながら、**自社のAIガバナンス・ゴール**（たとえばAIポリシー）**を設定するか否かについて検討**すべきであり、潜在的な負のインパクトが軽微であることを理由にAIガバナンス・ゴールを設定しない場合には、その理由等をステークホルダーに説明できるようにしておくべきである。「人間中心のAI社会原則」が十分に機能すると判断した場合は、自社のAIガバナンス・ゴールに代えて「人間中心のAI社会原則」をゴールとしてもよい。なお、ゴールを設定しない場合であっても、「人間中心のAI社会原則」の重要性を理解し、行動目標 3 から 5 に係る取り組みを適宜実施することが望ましい。

[将来的に]・・・自社独自のAIガバナンス・ゴールを掲げることが必要であると考えており、**ゴール設定に向けて社内で他社の事例等に関する勉強会を開催**している。（実践例 2 より）

「現時点では「人間中心のAI社会原則」を尊重することとし、それと並行してAIに関する全社的な研修の一部にAI倫理を追加することでAI倫理に対する理解を高めていくことを狙っている。さらにAI相談窓口を社内に設置して、事業部からの事例集めを行っている。・・・**AIガバナンス・ゴールの合意に向けたプロセスに価値がある**」（実践例 3 より）



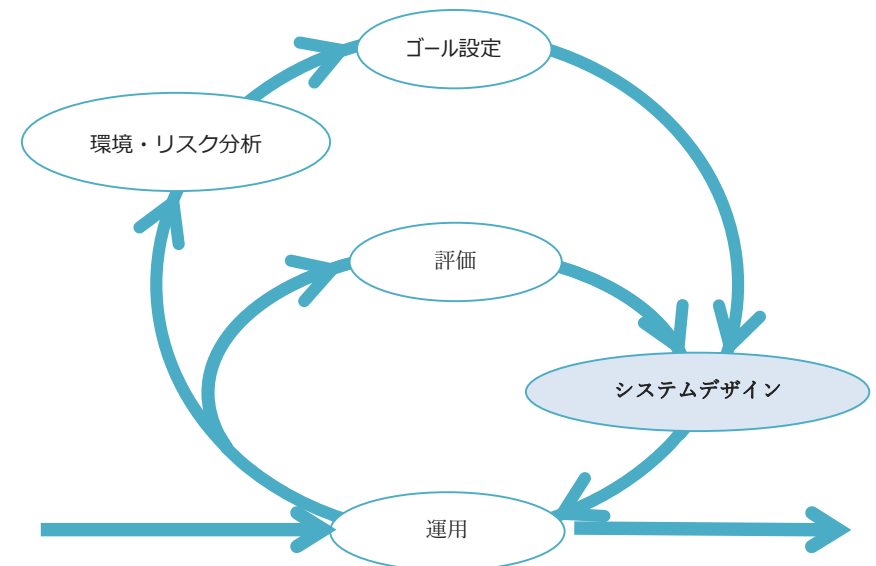
- AIガバナンス・ゴールを達成するために、ゴールからの乖離の評価と乖離への対応、リテラシー向上、事業者間等協力によるAIマネジメントの強化、インシデントに関わるAIシステム利用者の負担軽減に取り組むべき

行動目標 3 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、自社が開発・運用しているAIシステムのAIガバナンス・ゴールからの乖離を特定し、乖離により生じる影響を評価した上、負のインパクトが認められる場合、その大きさ、範囲、発生頻度等を考慮して、その受容の合理性の有無を判定し、受容に合理性が認められない場合にAIの開発・運用の在り方について再考を促すプロセスを、AIシステムの設計段階、開発段階、利用開始前、利用開始後などの適切な段階に組み込むべきである。運営層はこのプロセスの具体化を行うべきである。そして、AIガバナンス・ゴールとの乖離評価にはAIシステムの開発や運用に直接関わっていない者が加わるようにすべきである。なお、乖離があることのみを理由としてAIの開発・提供を不可とする対応は適当ではない。そのため、乖離評価は負のインパクトを評価するためのステップであって、改善のためのきっかけにすぎない。

別添 2 の乖離評価例が参考になる（次スライド）

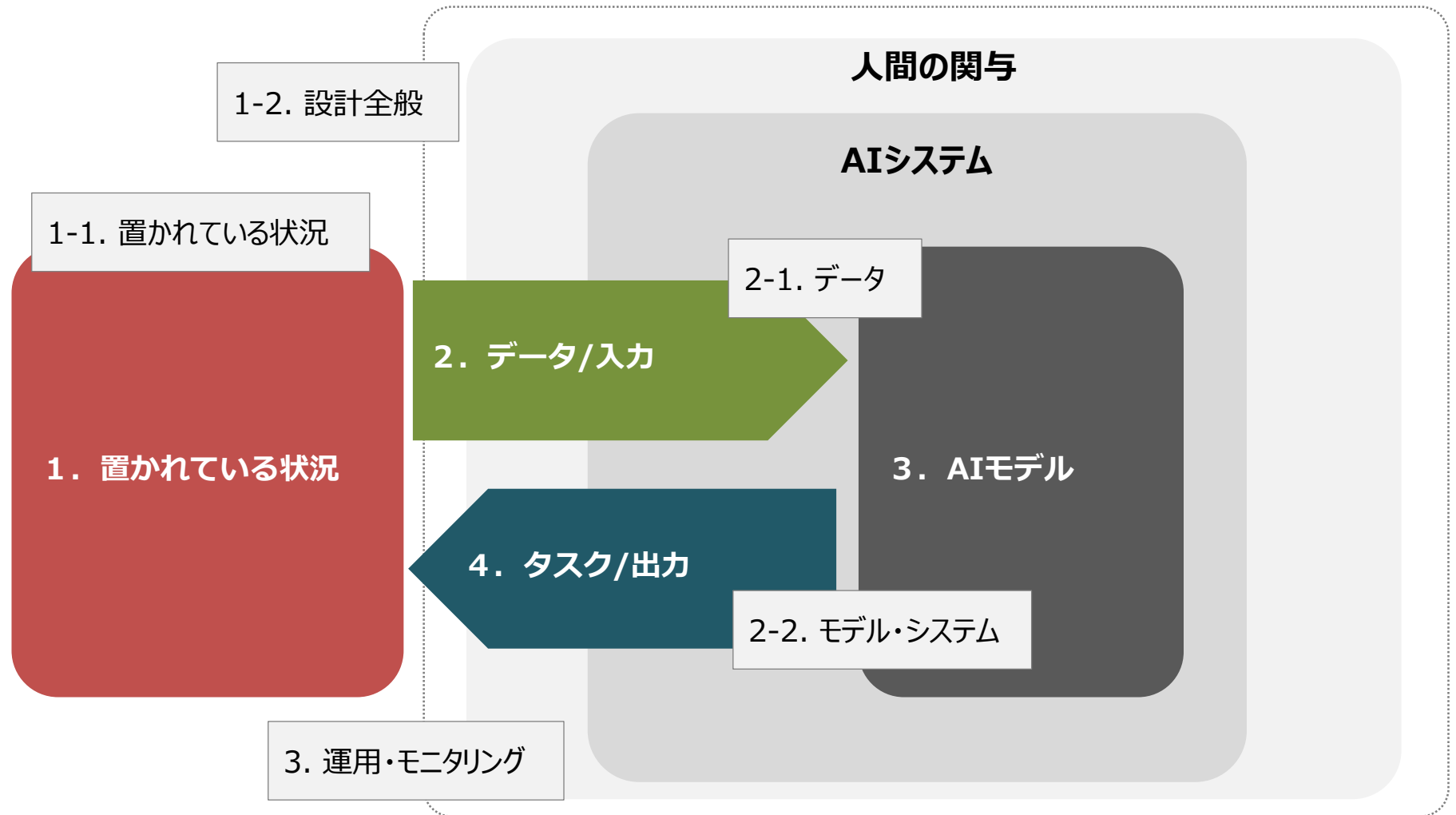
その他の行動目標は、本文を参照されたい。

- ・行動目標 3 - 2 : マネジメントを担う人材のリテラシー向上
- ・行動目標 3 - 3 : 事業者間等協力によるAIマネジメントの強化
- ・行動目標 3 - 4 : インシデントに関わる利用者の負担軽減





- システムデザインにおけるゴールからの乖離の評価と乖離への対応の参考情報として、OECDのフレームワークを参考にして、公平性、安全性、透明性等に関する配慮事項を例示（5観点、43項目例、136具体例）





## 行動目標の概要（運用）

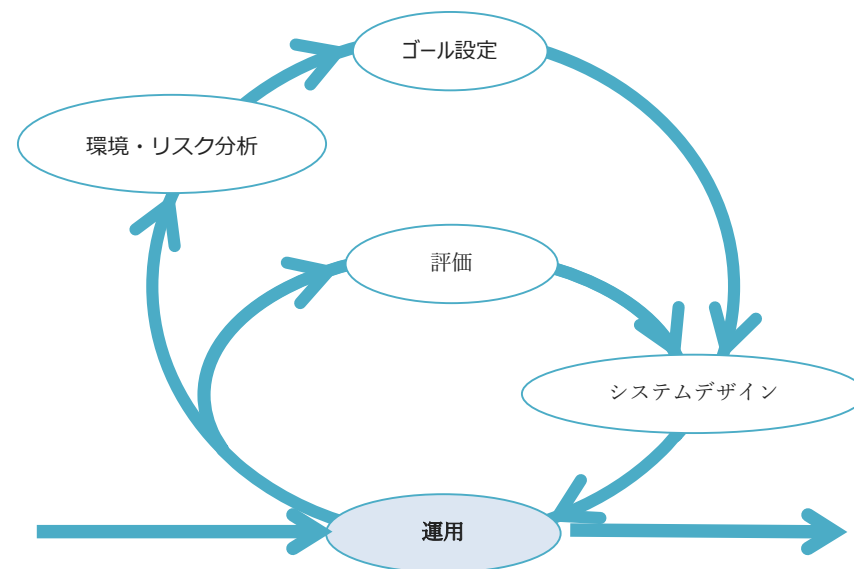
関連箇所：C-4.

- 設定したゴールとシステムを継続的に評価・再分析するため、**AIマネジメントシステム及び個々のAIシステムの運用状況について説明可能な状態を確保し、これらの情報を非財務情報に位置づけ積極的に開示することを検討**すべきであり、開示しない場合には、その理由等を説明できるようにしておくべきである。

行動目標 4 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、たとえば、行動目標 3 - 1 の乖離評価プロセスの実施状況について記録するなど、**AIマネジメントシステムの運用状況について対外的に説明可能な状態を確保**すべきである。

行動目標 4 - 3 : AIシステムを開発・運用する企業は、**AIガバナンス・ゴールの設定、AIマネジメントシステムの整備や運用等に関する情報を、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、積極的に開示することを検討**すべきである。上場会社以外であっても、AIガバナンスに関する活動の情報を積極的に開示することを検討すべきである。そして、検討の結果、開示しないと判断した場合には、その理由を対外的に説明できるようにしておくべきである。

行動目標 4 - 2 : AIシステムを運用する企業は、経営層のリーダーシップの下、個々のAIシステムの仮運用及び本格運用における乖離評価を継続的に実施するために、**仮運用及び本格運用の状況をモニタリングし、結果を記録**すべきである。・・・



# 行動目標の概要（評価、環境・リスク再分析）

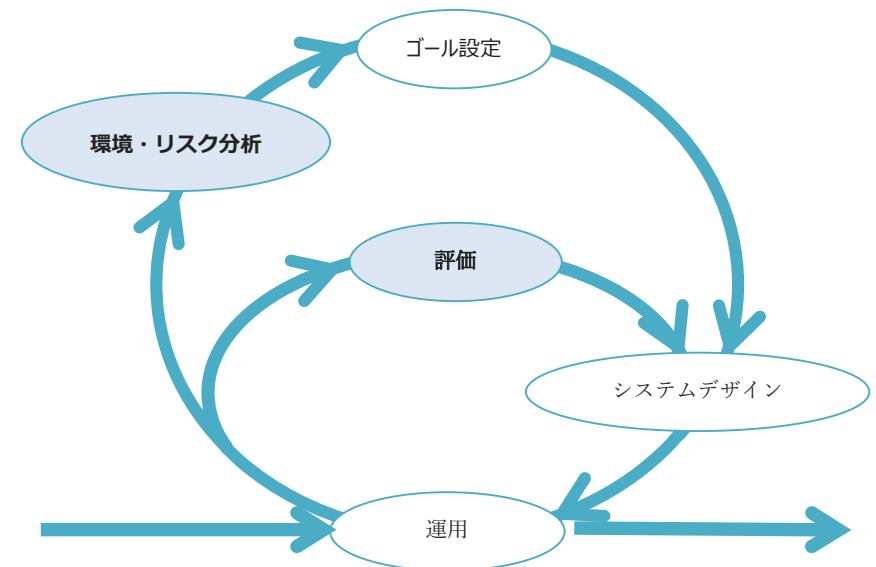
関連箇所：C-5.& 6.

- システムの設計や運用から独立した者に、その設計や運用の妥当性を評価させるべき。上述の運用状況に関する情報を用いながら社内で妥当性の評価を実施するとともに、株主だけではなく、マルチステークホルダーに意見を求めることを検討すべき。また、環境・リスクの再分析を適時に実施すべきである。

行動目標 5 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、AIマネジメントシステムの設計や運用から独立した者に、AIガバナンス・ゴールに照らして、乖離評価プロセス等のAIマネジメントシステムが適切に設計され、適切に運用されている否か、つまり行動目標 3、4 の実践を通じ、AIガバナンス・ゴールの達成に向けて、AIマネジメントシステムが適切に機能しているか否かを検証させるべきである。

行動目標 5 - 2 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、株主だけではなく、ビジネスパートナー、消費者を含む利用者、AIシステムの適切な運用をめぐる動向に詳しい有識者などの様々なステークホルダーから、AIマネジメントシステムやその運用に対する意見を求めることを検討すべきである。そして、検討の結果、実施しないと判断した場合には、その理由を対外的に説明できるようにしておくべきである。

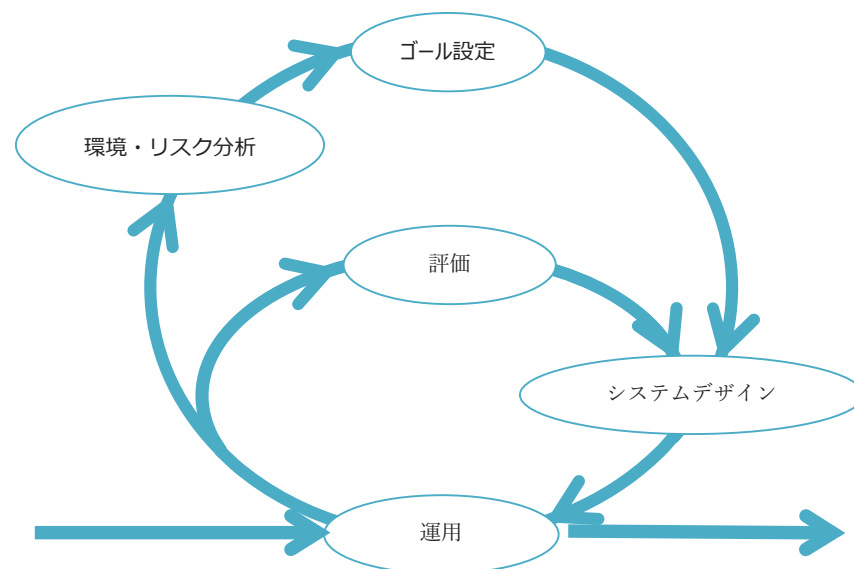
行動目標 6 - 1 : AIシステムを開発・運用する企業は、経営層のリーダーシップの下、行動目標 1 - 1 から 1 - 3 について、適時に再評価、理解の更新、新たな視点の獲得などを行うべきである。…



- AIガバナンス・ガイドラインをフォーカルポイントとして、ステークホルダーの関与の下で、AIガバナンス及び本ガイドラインの在り方の検討を継続し、必要に応じて改訂を行うことが望ましい。



「Governance Innovation Ver.2」報告書では、アジャイル・ガバナンスの実践に向けた企業の取組を後押しするために、標準やガイドラインといったソフトウェアによって、官民共同で政策ツールを策定していくことが重要であると述べている（4.3.3）。本ガイドラインは、AIのガバナンスという局面で、まさにそのような企業の取組みを後押しするツールとしての役割を果たすものであるといえる。さらに、本ガイドライン自体もアジャイル・ガバナンスのプロセスに則って継続的に評価・見直し・アップデートされるべきであり、AI技術の発展やAIに対する社会的な受容の変化などを適時に反映した官民共同のLiving Documentとして、継続的にメンテナンスされ、参照され続けていくことが望ましい。



## AI原則の実践の在り方に関する検討会

渡部 俊也 (座長)	東京大学未来ビジョン研究センター 教授	福岡 真之介	西村あさひ法律事務所 弁護士
雨宮 俊一	株式会社NTTデータ 技術開発本部 本部長	古谷 由紀子	サステナビリティ消費者会議 代表
生貝 直人 (+WG主査)	一橋大学大学院法学研究科 准教授	増田 悦子	公益社団法人全国消費生活相談員協会 理事長
上野山 勝也	株式会社PKSHA Technology 代表取締役	丸山 友朗	パナソニック株式会社 イノベーション推進部門 テクノロジー本部 デジタル・AI技術センター AIソリューション部 主任技師
川上 登福	株式会社経営共創基盤 共同経営者 マネージングディレクター、 一般社団法人日本ディープラーニング協会 理事	宮村 和谷 (+WG)	PwCあらた有限責任監査法人 パートナー
齊藤 友紀 (+WG)	法律事務所 LAB-01 弁護士	山本 龍彦	慶應義塾大学大学院法務研究科 教授
杉村 領一	産業技術総合研究所 情報・人間工学領域 人工知能研究戦略部 上席イノベーションコーディネータ	青島 武伸 (昨年度のみ)	パナソニック株式会社 イノベーション推進部門 テクノロジー本部 デジタル・AI技術センター データアナリティクス部 部長 (当時)
角田 美穂子	一橋大学大学院法学研究科 教授		
妹尾 義樹	産業技術総合研究所 イノベーション推進本部 標準化推進センター 審議役		
田丸 健三郎	日本マイクロソフト株式会社 業務執行役員 ナショナル テクノロジー オフィサー	岡田 淳	森・濱田松本法律事務所 弁護士
土屋 嘉寛	東京海上日動火災保険株式会社 企業商品業務部 部長	中崎 尚	アンダーソン・毛利・友常法律事務所外国法共同事業 弁護士
中条 薫	株式会社SoW Insight 代表取締役社長	本橋 洋介	日本電気株式会社 AI・アナリティクス事業部 シニアマネジャー
原 聡	大阪大学産業科学研究所 准教授		

## AIガバナンス・ガイドライン ワーキンググループ (WG)

\*その他にも、ヒアリング等で、多くの有識者、企業に御協力いただいた。