

AIガバナンスの実効性向上に関する動向

— モニタリング、エンフォースメント、インセンティブ —



AIガバナンスの実効性向上（直接アプローチ）

- 直接アプローチ：AIガバナンスに対するモニタリングとエンフォースメント。
- 間接アプローチ：自主的なAIガバナンスの実践を促す。

● 直接アプローチ

① 横断的規制＋専門委員会＋エンフォースメント

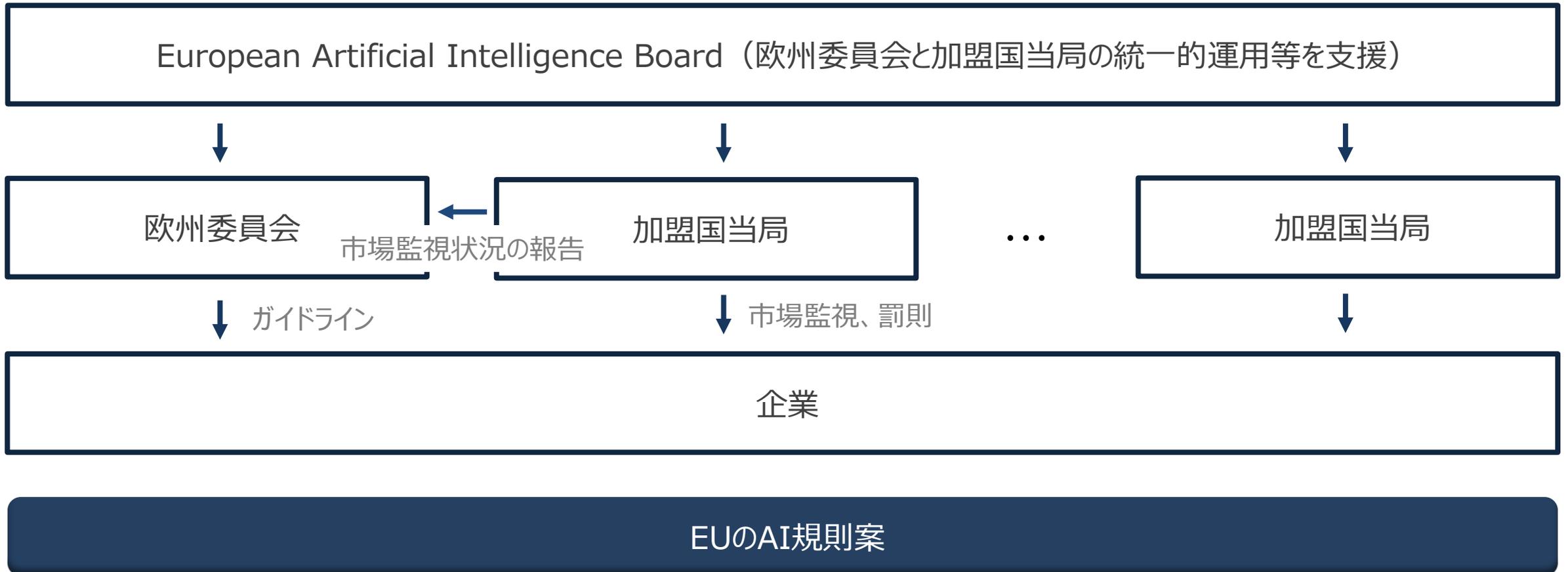
AIの使用に関する**横断的な規制を制定**するとともに、法律の運用や企業におけるガバナンスの遵守を支援する機関として、**専門委員会を設置**する。EUのAI規則案が代表的。深圳の条例案（深圳特別経済区人工知能産業振興規則、2021/7）も類似。なお、豪州の人権委員会のレポートでは、立法等を支援するためのAI安全委員会の設置が提言されている（Human Rights and Technology Final Report 2021, 2021/3）。

② セクター規制＋エンフォースメント

包括的な規制を制定せず、**各部局が規制を制定**。米国FTCは、AIに関するエンフォースメントの指針を公表（2020、2021）。これまで②のアプローチを採ってきた英国がセクターアプローチの方針を再検討中（National AI Strategy, 2021/9）。その他にも例あり（e.g. 自動運転分野でもルール形成が進む）。

① AI規則案 (EU)

- EUのAI規則案が、EU全体の規律の基準を決める（行政罰の罰金上限等）。
- European AI Board（専門委員会）が、統一運用等を支援する。エンフォースメントは加盟国当局。



① 深セン経済特区AI産業促進条例案（中国（深セン））

- 深セン政府はAI産業促進条例案を作成し、2021年7月にパブリックコメントを実施。
- 第6章「ガバナンスの原則と対策」では、リスクに応じた規制の設定やAI倫理委員会の設立等を提示。
- **リスクに応じた規制モデル**（第71条～第77条）の実施
 - 高リスク分野におけるAIは、事前評価とリスクの警告が必要である（第71条）
 - 中・低リスク分野におけるAIは、事前開示と事後制御による監督を必要とする（第71条）
 - 公共部門や公共の利益に関連する民間部門が使用するAIは、国民が理解できる方法でアルゴリズムの説明を行うべきである（第72条）
 - 業界団体は、技術標準や設計ガイドラインの制定、関連情報の展開等の活動等により、ガバナンスを守った競争を行うよう監視・監督すべきである（第74条）
- AIの倫理基準策定、主要な分野におけるAIの監視、ベストプラクティスの発信等を行う機関として、**AI倫理委員会の設立**を定めている（第68条）
- 禁止事項（プライバシー侵害、偽造等）やAIを用いた禁止行為（不正競争、独占）を明示（第78～82条）

①（参考）人権と技術に関する最終報告書（豪州）

- 豪州人権委員会は、人権を中心とした技術へのアプローチに関する38項目の提言を公表。
 - 企業に対して、事前影響評価やAI使用の公表を義務付ける一方、政府に対しては**AI安全委員会**の設立を提言。企業が規制を遵守するための支援ツールの開発・提供を求める。
- 報告書は「①新技術に関する国家戦略」「②政府および民間企業での意思決定におけるAIの利用」「③**AI安全委員会の設置による効果的な規制の支援**」「④障害を持つ人々にとってアクセス可能な技術」で構成
- ①では、政府に対し「デジタル経済戦略」の策定を通じ、**規制の制定**や**AI安全委員会の設立**を通じて、責任あるイノベーションの促進をすべきとしている（提言1）。
 - ②では、政府及び企業におけるAIを用いた意思決定に対する規制や要求事項として、事前の影響評価（提言2、提言9）やAI利用の公表（提言3、提言10）等を挙げ、政府に対して、これらを実施するための組織の設立や法整備、支援ツールの開発支援をすべきとしている（提言15、提言17）。加えて、政府が調達するAIシステムも人権アプローチを遵守すべきとし、関連する法規制、ポリシー、ガイダンスの修正を求めている（提言16）。「顔認識を含む生体認証」も規制すべきとしている（提言19）。
 - ③では、**AI開発における安全確保と人権保護に焦点を当てた独立機関として「AI安全委員会」の設立**を提言（提言22）

② FTC（米国）

- FTC法 5 条に基づき、不公正または欺瞞的行為を規制。2020年と2021年にAIの利用に関するガイダンスを公表。AIの使用における考慮事項を提示するとともに、偏りのあるAIの使用に伴う罰則について言及。
- 不公正で欺瞞的なAIシステムの使用例として、「任意利用と表示していた顔認証機能のデフォルト設定」等を提示。

AIに関連しない（金融詐欺、欺瞞的広告等）

AIに関連

不服申立て/職権調査

↓ 調査、提訴

企業

FTC法 5 条（不公正または欺瞞的行為の規制） その他、公正信用報告法、財政支援機会均等法

② セクター別アプローチ+2021年9月のAI戦略（英国）

- 英国は（1）産業横断（データ保護や人権と平等等）、（2）産業別（金融や医療等）といった形で、情報コミッショナーズオフィス（ICO）や金融公道監視機構（FCA）といった個別の部局がAIを規制している。
- AIに特化した規制はないが、産業横断的なデータ保護委員会や平等と人権に関する委員会や、金融当局や医療当局が、AI利活用の多くの側面を規制してきた。このような**セクター別アプローチは、政府が「AIに特化した横断的規制は適当ではない」という2018年の上院の見解への同意に基づくもの。**
- しかし、近年はセクター間での規制の一貫性のなさ、義務の重複、既存の規制範囲を越えたりリスクの発生、国際的な取組からの遅れといった課題が発生。今後10年のAI分野における計画を示すAI戦略を公表し、第3の柱として「効果的なAIのガバナンス」を提示。これまでのセクター主導の規制から方針を転換する可能性を示唆。AI庁（Office for AI）は、AIの管理と規制に関する政府の立場を示すホワイトペーパーを2022年初頭に公表する旨を発表。
- なお、**セクター主導の規制に替わる、AIガバナンスの代替手法の例**として、「不必要な規制が確認された場合の規制緩和」「セクター主導の体制を保持したまま、各部局における作業の柔軟化」「**セクター横断的なルールの導入**」等が提示されている。

AIガバナンスの実効性向上（間接アプローチ）

- 直接アプローチ：AIガバナンスに対するモニタリングとエンフォースメント。
- 間接アプローチ：自主的なAIガバナンスの実践を促す。

● 間接アプローチ

③ 非拘束的なガイドライン+事例+業界・自主規制

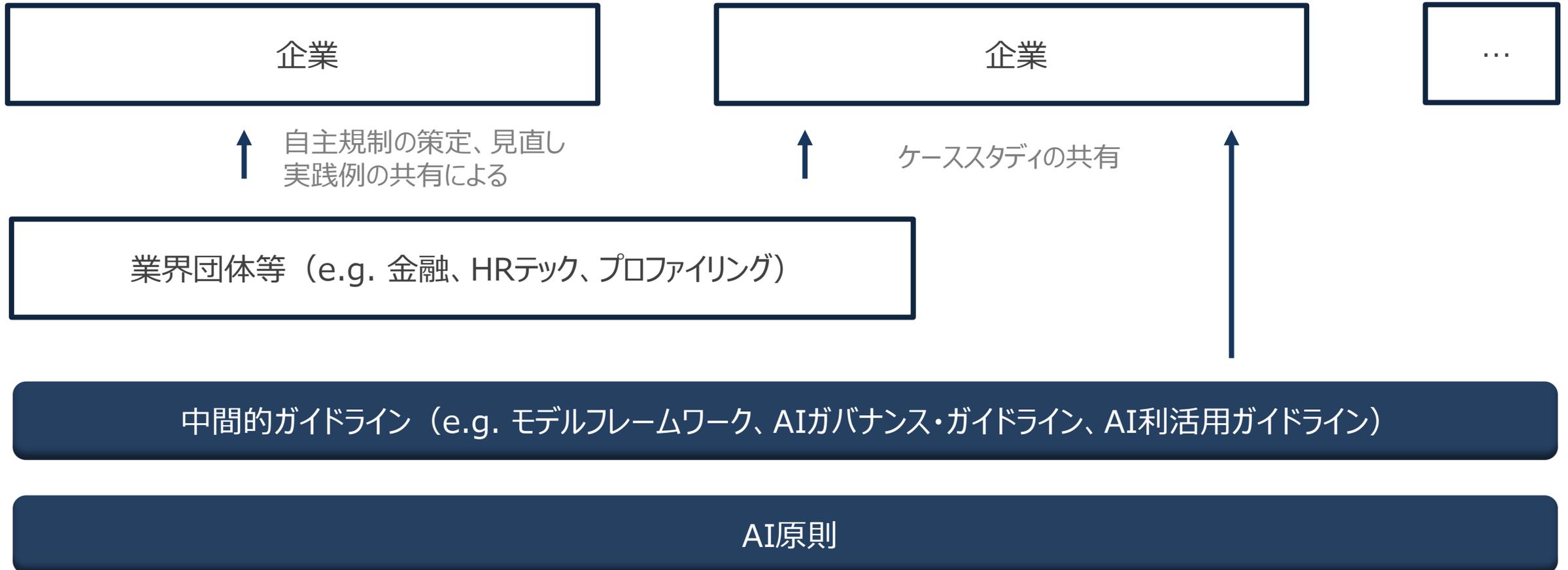
政府が、遵守を義務としない**中間的なガイドライン**を公表。具体的な運用については、**業界団体等が自主的に策定**。シンガポールは、モデルAIガバナンスフレームワークと事例を公表（2020/1）。シンガポールの金融機関等で構成されるVeritas Initiativeを開始（2019）。途中成果としてVeritas Phase 1 White Paperを公表（2021/1）。豪州も同様にフレームワークと原則と事例を公表している（2020）。

④ 政府調達要件

政府のAIシステム調達における**サプライヤーへの要求事項**を通じて、企業におけるAIガバナンスの改善を間接的に促進。カナダでは、政府におけるAI使用のポリシーとして、自動意思決定指令を策定（2019）。AIシステムを調達する場合、サプライヤーにも政府の義務が適用される。そのようなAIシステムを提供できる企業リストが公表されている。英国でも取り組みが進む。WEFがAI政府調達に関するレポートを公表（2020）。IEEEでは、AI政府調達標準のプロジェクトが承認された（2021/9）。

③ 非拘束的な横断的ガイドライン+事例+業界・自主規制（シンガポール、日本、豪州）

- 政府が非拘束的な横断的ガイドラインを公表。（必要に応じて）業界団体が関連企業に対する自主規制を策定（ガイドラインが先とは限らない。シンガポールでは金融分野のFEAT原則がモデルフレームワークより先）
- 事例を提供することで、企業での実践を後押しするとともに、適宜内容の見直しも図る。

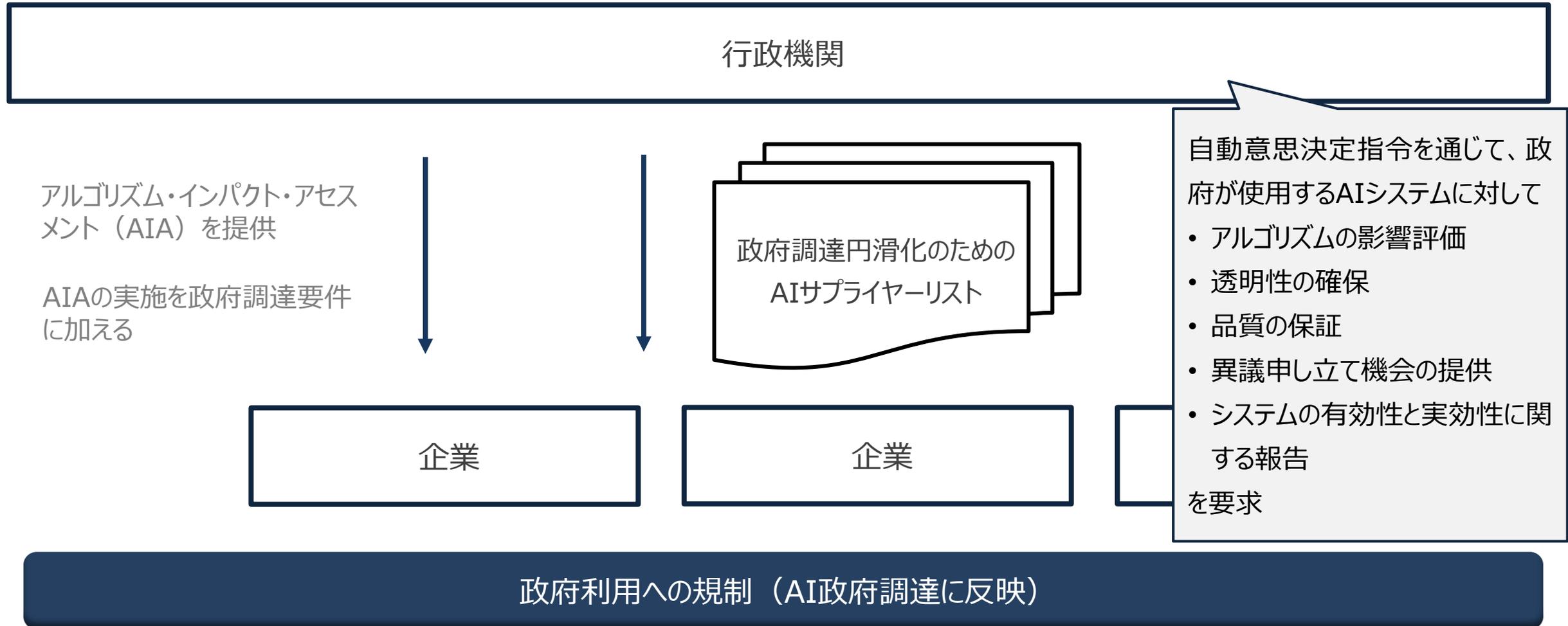


③ Veritas Phase 1 White Paper (シンガポール)

- シンガポール金融管理局 (MAS) は、責任あるAIの使用に関するFEAT (公平性、倫理、説明性、透明性) 原則の実践に向け、金融機関等で構成されるVeritas Initiativeを2019年に設立。2021年1月、途中成果としてPhase 1のホワイトペーパーを公表。
- ホワイトペーパーは、金融商品やサービスを提供する企業に対し、責任あるAIやデータ分析 (AIDA) に関するガイダンスを提供することで、データの管理や利用に関する**金融業界全体としてのガバナンス強化**を目的としている。FEAT原則の評価方法を記載したDocument 1、ケーススタディにおける評価例を示すDocument 2からなる。
 - Document 1: ①金融分野におけるAIDAをリスクベースで評価するアプローチ、および ②FEAT原則に関するAIDAの評価方法
 - Document 2: ①顧客マーケティングと②信用スコアリングを対象に、公平性に関する評価の実践例を提示
- Phase 2では、①信用リスクスコアリングと②顧客マーケティングに加え、③予測的な保険引き受け、④詐欺検知にも取り組む予定。
- Phase 2の参加企業には、欧米の金融機関とテック企業が目立つ。日系ではMUFGのみ。

④ 自動意思決定指令（カナダ）

- 政府が利用するシステムにおけるAIの要求事項を定め、調達プロセスに反映。
- 企業が要件を遵守するためのツールを提供することで、ガバナンスの遵守を後押し。



④ (参考) 自動意思決定指令 (カナダ)

- 政府におけるAI使用のポリシーとして、自動意思決定指令を策定。
 - 同指令で求められる「アルゴリズムの影響評価」を支援するツールを提供。
- 政府は、透明性や品質保証といった要件を含む自動意思決定指令を策定。政府で使用するAIが組み込まれたプログラムに対して、「アルゴリズムの影響評価」「透明性の確保」「品質の保証」「異議申し立て機会の提供」「システムの有効性と実効性に関する報告」を求める。
 - 「アルゴリズムの影響評価」を支援するツールとして、アルゴリズム影響評価ツール (AIA) を提供
 - 「リスク」に関する48の質問と「リスクの軽減」に関する33の質問への回答を得点化し、AIシステムの影響レベルの大きさを4段階 (影響なし、中程度の影響、大きい影響、非常に大きな影響) で分類
 - 「リスク」については「プロジェクト」「システム」「アルゴリズム」「意思決定」「影響の大きさ」「データ」の質問
 - 「リスクの軽減」については「コンサルテーション」「データ品質」「手続的公正」「プライバシー」の質問
 - AIシステムの責任ある倫理的な使用に関する追加のガイダンスでは、サプライヤーと契約してシステムを実装する場合、**サプライヤーにも政府の義務に理解を求め、AIAを実施することを要件にすることを推奨**。責任あるAIサービス等を提供可能な**AIサプライヤーリストが公表**されている (日系企業も掲載) 。

④ AI政府調達ガイドライン（英国）

- 政府は、公的機関がAIを導入する際に留意すべき事項を示したガイドラインを公表
 - AI導入において特に重要となる10項目の留意事項と、調達プロセスにおける対応例を記載
- 公的機関におけるAI調達で留意すべき事項として、①政府戦略における調達の明確化、②多様かつ学際的なチームでの意思決定、③データへのアクセスや管理に関する事前評価、④AI導入によるメリットとリスクの評価、⑤早期からの市場との効率的な関与、⑥（手段ではなく）解決したい課題への注目、⑦ガバナンスと情報の保証に関する計画の作成、⑧アルゴリズムのブラックボックス化とベンダーロックインの回避、⑨AIの技術的・倫理的制約を評価できる入札審査体制の構築、⑩AIシステムのライフサイクルの考慮の10項目を提示。
 - さらに、「計画」「公募」「選考・評価・採択」「契約・管理」の各調達プロセスで考慮すべき項目や実践方法を提示。
 - ① 計画：学際的なチームの例、データ評価における確認項目、AI影響評価における確認項目、市場への予備的な関与方法、主な公共調達システム
 - ② 公募：仕様書作成時の考慮事項
 - ③ 選考・評価・採択：サプライヤーを評価する際の観点
 - ④ 契約・管理：プロセスベースのガバナンス、モデルのテスト、ノウハウの共有と訓練、サポート終了時の定義

④ (参考) 自動化された意思決定のための倫理、透明性、説明責任の枠組み (英国)

- 政府によるAIの使用に関する国民からの信頼醸成を目的とし、公共部門における倫理的なAI使用に関するガイダンスとして「自動化された意思決定のための倫理、透明性、説明責任の枠組み」を策定
- 公共部門において倫理的にAIを使用するためのガイドラインとして、国民から信頼を得られる形でAIを使用する上で求められる7項目を提示。項目毎に実践の手順が示されるとともに、各項目を実施する際に参考となる行政文書へのリンクが提示されている。
 - ① 意図しない結果を回避するためのテスト
 - ② 全てのユーザー・市民への公正なサービスの提供
 - ③ 責任の明確化
 - ④ 市民のデータを保護するための安全なデータの取扱
 - ⑤ AIがユーザー・市民に与える影響についての理解促進
 - ⑥ 法律への準拠の確認
 - ⑦ 将来を見据えた構築

④ AI公共調達ガイドライン（WEF ※英国やUAEが協力）

- 世界経済フォーラムは、公共部門がAIソリューションやサービスを導入する際に考慮すべき事項を示すガイドラインと、考慮事項を検討するためのツールとなるワークブックを公表。英国のパイロット事例研究も公表されている。
- 公的機関におけるAI調達で留意すべき事項として、①関連データへのアクセスや管理の実現可能性の明確化、②特定の課題解決に特化せず、反復の余地を残すような課題の設定、③AI利用における「公共の利益」の定義、④既存の政府戦略との整合、⑤関連する法律や行動規範の仕様への組み込み、⑥データの品質の限界の強調、⑦多様で学際的な体制の要求、⑧アルゴリズムの説明責任と透明性の確保、⑨落札企業への継続的な関与、⑩AIソリューションプロバイダー間の公平かつ公正な競争条件の調整の10項目を提示。
- ワークブックは、AI調達の実務で参照されることを想定し、ガイドラインで提示した10項目に関する質問、サプライヤーから提案されたAIシステムを公共部門が評価する際に参考となる質問を提示。
- パイロット事例研究（GovTech Catalyst Challenge、パブリックセクターの課題解決ファンド）には、これらの取り組みには一定の意味があるとしつつも、「倫理基準が一般的で曖昧」「全ての基準が特定のプロジェクトに関連するわけではない」「最も関連する基準に絞り込んだ方がよい」「監査は意味があるかもしれないが、誰が監査するのか、様々なアプリケーションや技術に対応できるのか、などの詳細が重要」などのサプライヤーの声が掲載されている。

④ AI政府調達標準（IEEE）

- 政府は、技術の消費者だけではなく、スチュワードの側面から役割を果たしてきたが、AIの活用が進む中、AI政府調達に関する互換性のある相互運用できる標準がないことが課題。
 - 2021年9月23日、IEEE SA Standard Boardは、AI政府調達標準（IEEE 3119 Standard for the Procurement of Artificial Intelligence and Automated Decision Systems）のプロジェクト化を承認した。
-
- IEEEのAI倫理標準であるIEEE7000シリーズの一部に位置づけられている。
 - 現在、40名の個人が参加意思を表明。北米がメインであるが、その他の国から手が挙がっている（アルゼンチン、ベルギー、チリ、中国、フィンランド、ドイツ、オランダ、ナイジェリア、スペイン、英国）。
 - AI政府調達標準作業グループでは、最初に以下の3つに取り組む。
 - ターミロジー：定義を決める
 - 国際的な状況：各国の取り組みを調査する
 - プロセスモデルと要件：国際的に通用するプロセスモデルとプロセス要件を検討する