

「責任あるアジャイル・ガバナンス」 における国家公務員の役割

2026年8月17日

AI時代の国家公務員の在り方に関する検討会（第1回）

羽深 宏樹

京都大学大学院法学研究科 特任教授

東京大学大学院法学政治学研究科 客員教授

スマートガバナンス株式会社 代表取締役CEO

内閣府 上席科学技術政策フェロー（AI政策担当）



1 | 「ゴール固定型」ではなく「ゴール変動型」の発想へ

	ゴール固定型	ゴール変動型
問い	今の業務のうち、AIに代替されないものは何か	AIによって、行政が新しくできるようになることは何か
出発点	現在の業務	行政が実現すべきアウトカム
人材像	判断・調整・責任	ガバナンス・システム(※)の設計・運用

※ 民主的制度、規制ガバナンス、コーポレートガバナンス、個別システムのリスクマネジメント等を含む多層的な概念

2 | 「責任あるアジャイル・ガバナンス」

「人工知能基本計画」（令和8年7月14日 閣議決定）：複雑化・深刻化するリスクに『責任あるアジャイル・ガバナンス』で対応

「デジタル・ニッポン2026～責任あるアジャイル・ガバナンス～」（自民党 デジタル社会推進本部 2026年5月）

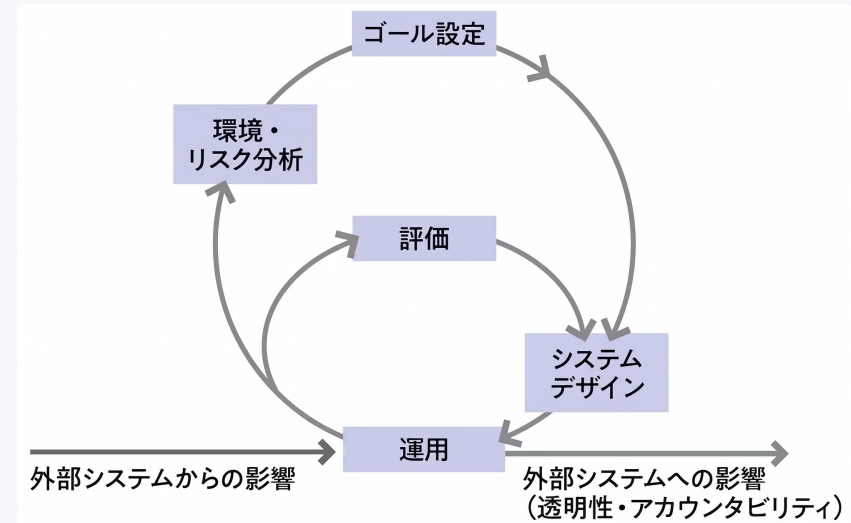
アジャイル・ガバナンスとは

事前に固定されたルールや計画に頼るのではなく、変化の激しい環境やリスクに応じて、多様なステークホルダーが対話とデータに基づき、目標（ゴール）や統治システム自体を継続的に評価・改善していく新しい管理・統治モデル

- ① 主体：マルチステークホルダー
- ② 手順：アジャイル
- ③ 構造：マルチレイヤー

日本における「アジャイル・ガバナンス」の系譜

- 経済産業省「GOVERNANCE INNOVATION」報告書シリーズ (2020～2026)
 - アジャイル・ガバナンスの枠組みと全体像を提示
- デジタル庁「構造改革のためのデジタル原則」（2021年）
 - 「アジャイル・ガバナンス原則」を提示
- G7「デジタル・技術大臣会合宣言」（2023）
 - アジャイルかつマルチステークホルダーなガバナンスの必要性に言及



出典：「GOVERNANCE INNOVATION Vol.2: アジャイル・ガバナンスのデザインと実装に向けて」（2021）

3 | AIによって、行政は何ができるようになるか

従来の行政	AI時代に可能になる行政	例
制度を作り、数年後に見直す	社会の変化を常時把握し、制度を継続的に改善する	申請・相談・苦情などをAIで分析し、制度の想定外の影響や新たなリスクを早期に検知する
国民から申請が来るのを待つ	必要な人に行政側から支援を届ける	保有データから給付・支援制度の対象となり得る人を把握し、「この制度を利用できる可能性があります」と案内する
一律の行政サービスを提供する	一人ひとりの状況に応じた行政サービスを提供する	複数の制度を横断して、個々の事情に応じた支援策や行政手続きを案内する「AI行政コンシェルジュ」
限られた案件だけを検査・監督する	大量の情報からリスクを検知し、監督資源を重点配分する	事故・苦情・取引データ等から高リスクな事業者や領域を抽出し、検査・監督を重点化する
政策案を経験と限られたデータで比較する	多数の政策シナリオを事前に探索・比較する	税制、補助金、交通政策などについて、複数の政策オプションとその影響をAIでシミュレーションする
省庁・制度ごとに情報を処理する	複数制度を横断して社会課題を捉える	子育て、教育、福祉、雇用などのデータを横断的に分析し、複合的な課題や制度の「谷間」を発見する

4 | AI時代の公務員の新しい役割

AI時代に可能になる行政	公務員の新しい役割
社会の変化を常時把握し、制度を継続的に改善する	政策のフィードバックループを設計・運用する
必要な人に、行政側から支援を届ける	誰に・いつ・何を届けるかというルールを設計する
一人ひとりの状況に応じた行政サービスを提供する	個別最適と公平性を両立するサービス設計を行う
大量の情報からリスクを検知し、監督資源を重点配分する	リスク指標・介入基準・救済手続を含む監督システムを設計する
多数の政策シナリオを探索・比較する	何を最適化し、何を守るかという目的・制約を設定する
複数の制度・組織を横断して社会課題を捉える	行政全体を俯瞰し、制度間をつなぐガバナンスを設計する

5 | 信頼性・適切性・透明性をどう確保するか

行政がAIを使う際、避けては通れないのが以下の問い

モデル・システムの信頼性

「そのAIを信頼できるか」

業務の適切性

「そのAIの使い方は適切か」

判断の透明性

「判断根拠が示されているか」

とりわけ公務では、民間以上にAIシステムの利用の信頼性・適切性・透明性が求められる

例：憲法31条（適正手続の保障）

憲法14条1項（法の下での平等）

行政手続法5条（審査基準の設定・公表）、8条・14条（理由の提示）

情報公開法1条（政府の説明責務）

公文書管理法4条（意思決定過程の文書作成義務）

6 | 従来、公務の信頼性・適切性・透明性を担保してきた仕組み

公務員の信頼性

採用試験

研修・人事評価

服務・懲戒

業務の適切性

法令・訓令・通達

決裁（専決・代決を含む内部手続）

監査・検査

国会による統制

会計検査院

司法審査

判断の透明性

審査基準・処分基準の設定と公表

理由の提示

意見公募手続（パブリックコメント）

情報公開

公文書管理

政策評価の公表

これらの仕組みは、すべて「人間」を前提に設計されている



7 | 公務におけるAI利用に関するルール現状

国・地域	制度・文書名	対象・内容
カナダ	自動化された意思決定に関する指令（2019／2020適用）	行政上の決定を行う・支援するシステムが対象 —— 生成AIも、こうした用途では対象
英国	アルゴリズム透明性記録基準（ATRS）／2024年2月、政府各省庁で義務化	意思決定に大きく関与する／市民と直接対話するツールを公開
米国	OMB M-25-21（2025年4月）	「高影響AI」に7つの最低限のリスク管理実務 —— 各省庁にChief AI Officer
EU	AI法	第26条：高リスクAIの導入者義務／第27条：一定の公的主体等に基本権影響評価／第86条：一定の個別決定について説明を受ける権利
欧州評議会	AIと人権・民主主義・法の支配に関する枠組条約（2024採択／日本は2025年2月署名）	文書化、異議・苦情申立て、反復的なリスク・影響管理
日本	行政の進化と革新のための生成AIの調達・利活用に係るガイドライン（令和8年6月／デジタル社会推進会議幹事会決定）／行政通則法的観点からのAI利活用調査研究会 中間整理（総務省）	便益とリスクを表裏一体で捉えた利活用推進／行政法の観点からの検討が進行中

これらは、先に述べた3つの要請に対応する形で組み立てられている

モデル・システムの信頼性：導入前の評価・試験と、継続的な性能のモニタリング

業務の適切性：人が介在する点の指定と、責任者・体制の明確化

判断の透明性：記録と公開、説明・異議申立て・救済の経路

※ いずれも、まず「どのAI利用を対象とするか」を定義したうえで、リスクに応じて要求の強度を変えるという構造をとっている

8 | AIエージェントの難しさ

AIエージェントの特徴

ツールとの接続

AIが外部のシステムやデータに接続し、自ら実行する（メール送信、データ更新、他のシステムへの指示等）

自律性

目的を与えると、それを達成するための手段をAI自身が決める

その結果・・・

- 人間によるモニタリングが十分に効かない間に処理される
- 行為が他のシステムや他のAIに連鎖し、責任の所在を追いにくい
- AIが自律的にハッキングを行うなどの「暴走」につながることも

(参考) 世界各国で進行中の議論

- シンガポールIMDA「[エージェント型AIのためのモデルAIガバナンス枠組み](#)」（2026年1月公表・5月改訂）
 - 権限範囲の限定、人間の承認点、継続的な監視、段階的な展開などを提示
- OECD「[エージェント型AIの全体像とその概念的基礎](#)」（The agentic AI landscape and its conceptual foundations・2026年2月）
 - 自律性の水準に応じて監督を校正する必要があると指摘
- 世界経済フォーラム「[政府のためのエージェント型AI活用：導入準備の枠組み](#)」（Making Agentic AI Work for Government: A Readiness Framework / 2026年4月）
 - 政府における導入準備の枠組みを提示

9 | AI時代の公務員に必要な能力（仮説）

相対的に重要性が低下するもの

一般的な知識の量

前例の理解

定型的な事務の正確な処理

相対的に重要性が高まるもの

ゴールを定める力

何をAIに実現させ、何を守るのかを言語化する

リスクと便益を構造化する力

ステークホルダーに生じる費用便益を評価する

仕組みを設計する力

ルール・技術・運用を組み合わせる

フィードバックサイクルを回し続ける力

観測し、ずれを見つけ、直す

アカウンタビリティを果たす力

相手に応じて、根拠を示し、責任を取る

これらこそ、AI時代の国家公務員に必要なガバナンス・システムの設計・運用能力ではないか

