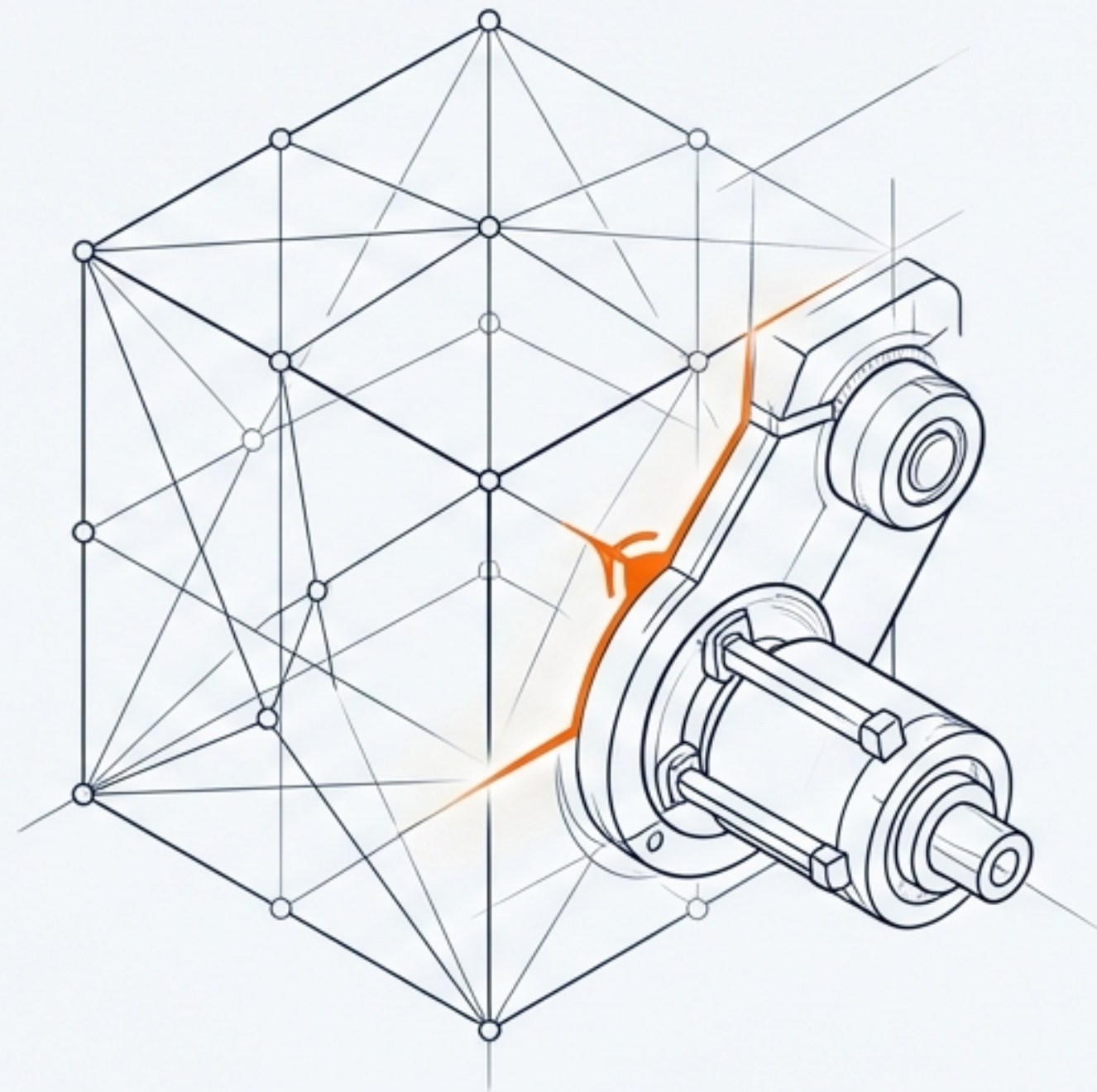


AI事業者ガイドライン 改定案：深堀分析

AIエージェント・フィジカルAIの台
頭と「人間の判断」を巡る論点



レポート作成日: 2026年2月18日
作成: Manus AI Analysis Report

エグゼクティブ・サマリー



背景 (Context)

- 2026年2月、総務省・経産省がガイドライン改定案を公表。
- 2025年版（第1.1版）策定後の急速な技術進化に対応。



主な変更点 (Major Changes)

- 「AIエージェント」と「フィジカルAI」の定義および規定を追加。
- 最も重要な要件：「**人間の判断を必須とする仕組み**」の構築。



論点 (The Controversy)

- 安全性**（暴走防止）と**イノベーション**（リアルタイム性・拡張性）の対立。
- 「一律の人間介入」に対する産業界・技術者からの**懸念**。



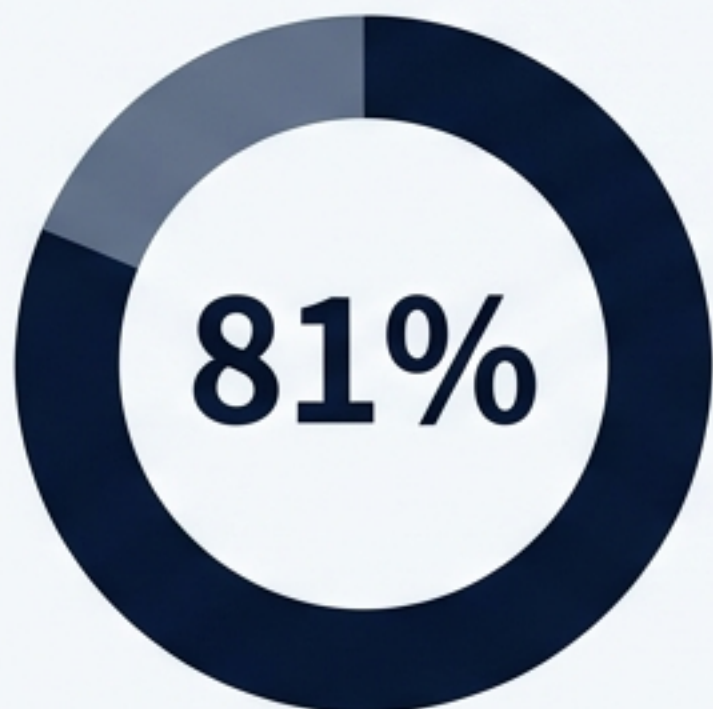
展望 (Future Outlook)

- 2026年3月に正式版公開予定。
- 将来的には法的拘束力を持つ「**AI法**」への**移行の可能性**（EU AI Actの影響）。

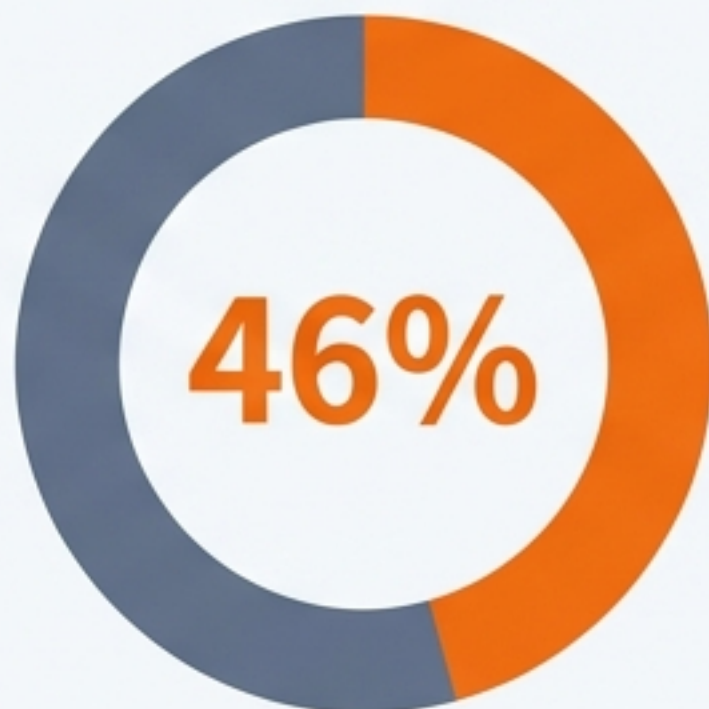
改定の背景：技術の急速な進化と実装のギャップ



ガイドライン認知度



実際の活用率



実装ギャップ（Implementation Gap）

- 課題：中小企業や地方自治体への普及が遅れている現状。
- 要因：自律的に判断するAIや、物理世界に作用する技術の台頭により、旧ガイドラインではカバーしきれないリスクが顕在化。

新たな定義①：AIエージェント

定義 (Definition)

「特定の目標を達成するために、
環境を感知し、自律的に行動する
AIシステム」

- 従来の「ツール」としてのAIから、自ら判断しタスクを遂行する「エージェント」への進化。
- サイバー空間内での自律的な完結能力。



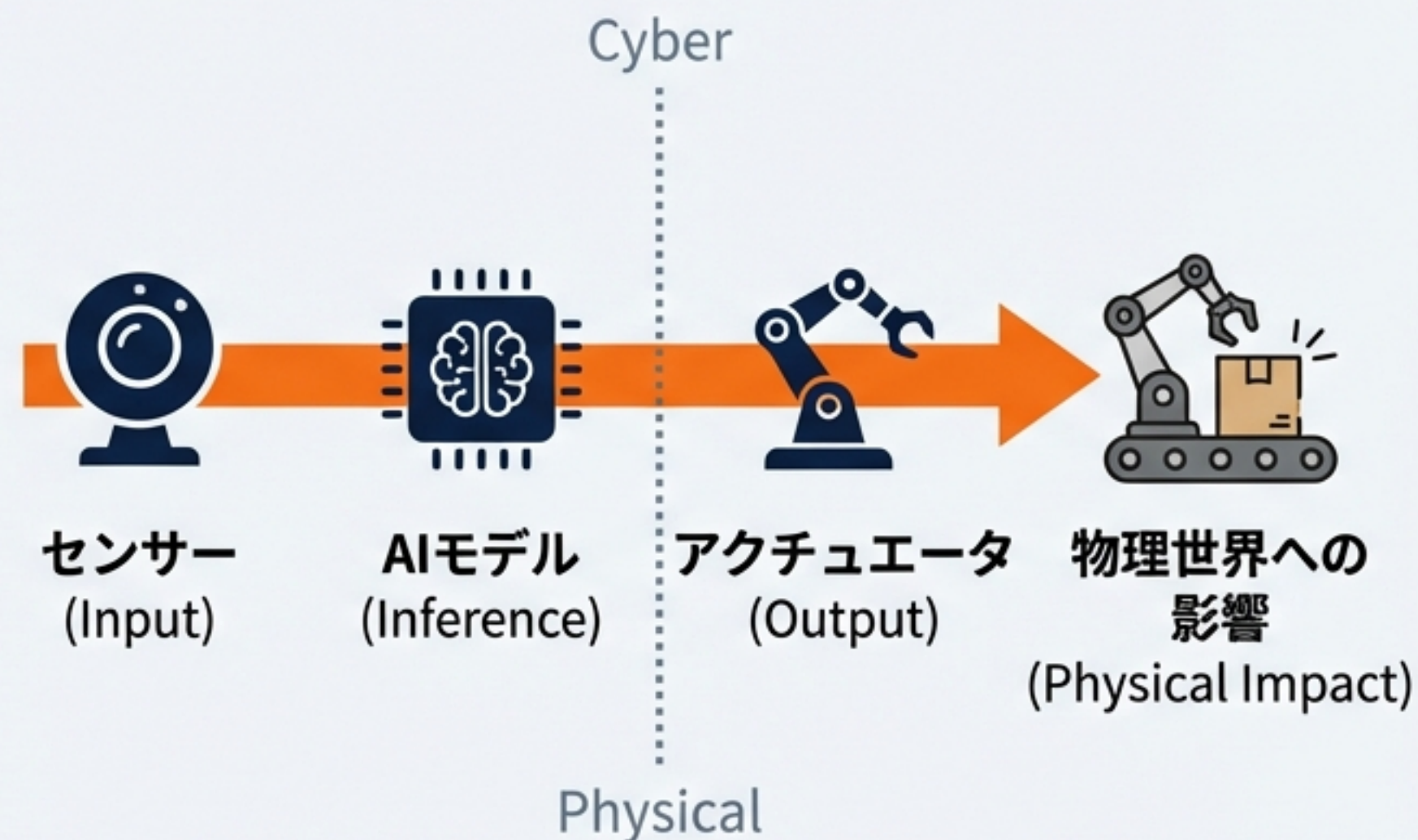
新たな定義②：フィジカルAI

定義（Definition）

「センシングを通じて物理環境の情報を取り込み、AIモデルによる推論・判断を経て、アクチュエータ（駆動系）等を介して物理的な行動へとつなげるシステム」

- サイバー空間の処理に留まらず、現実世界に対して直接的な働きかけ（移動、操作、加工など）を行う。
- リスクの質が「情報漏洩」から「物理的損害」へ拡大。

Cyber-Physical System



便益とリスクのランドスケープ

便益 (Benefits)

- ✓ ・業務効率化と労働力不足の補完
- ✓ ・安全性向上（危険作業の代替）
- ✓ ・介護・生活支援（高齢化社会への対応）

リスク (Risks)

- ❗ ・自律的な行動に起因する誤動作・暴走
- ❗ ・サイバー攻撃の対象・手法の物理的拡大
- ❗ ・機構の複雑化による保守・メンテナンスの困難化
- ❗ ・カメラ等との連携によるプライバシー侵害

最大の論点：「人間の判断」の必須化

Human-in-the-loop



Mandate: 「人間の判断を必須とする仕組み」の構築

AIの暴走や予期せぬ損害を防ぐための安全装置（キルスイッチ等）の実装。

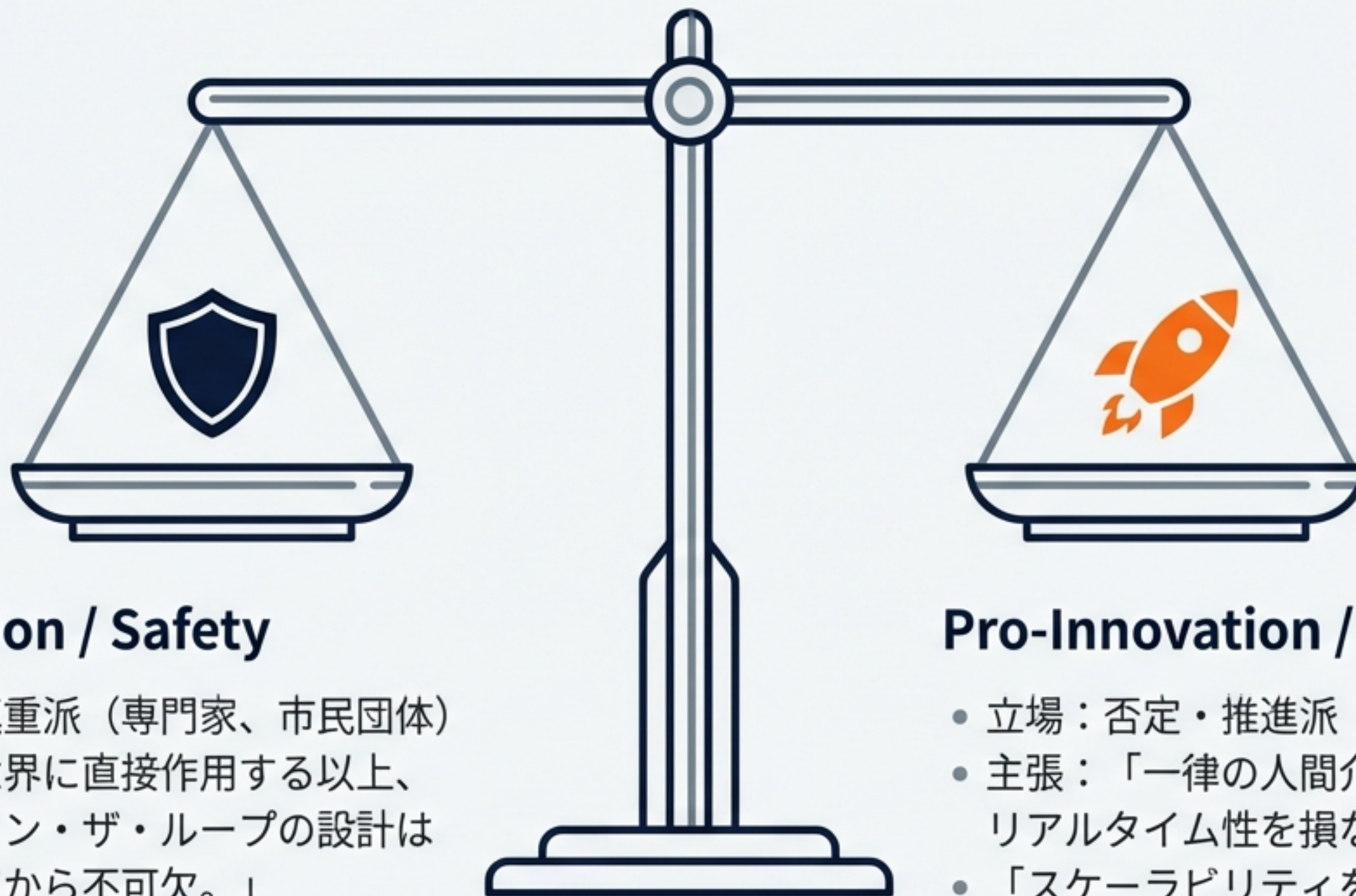
権限の最小化

AIに与える操作権限を必要最低限に絞る。

データへの配慮

ハードウェアに残存するデータの適切な処理。

安全性 vs イノベーション：ガバナンスのジレンマ



Pro-Regulation / Safety

- 立場：肯定・慎重派（専門家、市民団体）
- 主張：「物理世界に直接作用する以上、ヒューマン・イン・ザ・ループの設計は安全確保の観点から不可欠。」

Pro-Innovation / Industry

- 立場：否定・推進派（技術企業、研究者）
- 主張：「一律の人間介入はフィジカルAIのリアルタイム性を損なう。」
- 「スケーラビリティを毀損し、実用化のハードルを高める。」

産業界の視点：製造業における好機



**「社会実装に
強みがある」**

安川電機 小川社長

Japan's Competitive Edge

日本の製造業が持つ実績と信頼性を活かせば、フィジカルAIは脅威ではなく機会となる。

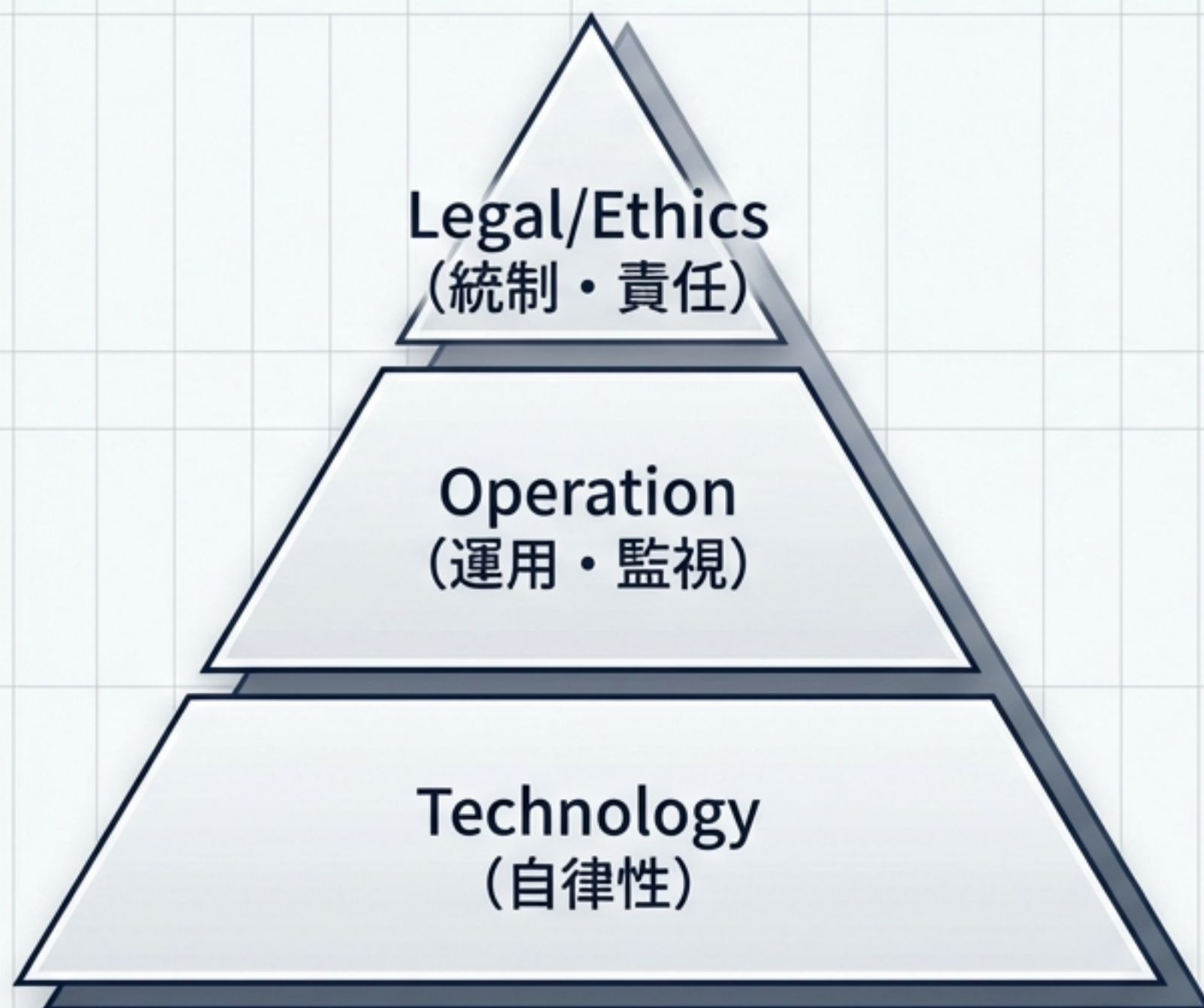
Implementation Focus

議論はリスクだけでなく、現場での社会実装（Social Implementation）をどう進めるかにシフトしている。

グローバルコンテキスト：国際的な規制動向との比較

EU (European Union)	Japan (Current Proposal)	US / China
 Strict / Risk-based	 The Middle Path	 Distinct Approaches
AI Actにより、フィジカルAIは「ハイリスク」に分類される可能性大。厳格な義務。	Hiroshima AI Processに準拠。ソフトロー中心だが、欧州の動向を意識し厳格化の傾向。	米国はイノベーション重視の「ライトタッチ」。中国は独自の国家主導規制。

戦略的考察：技術の自律性と人間の統制



イノベーションを過度に抑制することなく、いかに実効的な安全性を確保するか？

- ▶ 「柔軟かつ精緻な制度設計」
- ▶ ゼロリスクではなく、リスクレベルに応じた段階的な適用と例外規定の具体化。

今後のロードマップと展望

2026年2月

改定案公表 (Current)

2026年3月

正式版ガイドライン公表

- 支援ツール（チャットボット）提供
- 例外規定の具体化

Long Term

法制化の検討

「AI法」制定への試金石

結論：フィジカルAI時代のガバナンス



- **New Era:** 我々はサイバー空間を超え、物理世界に作用するAIの時代に突入した。
- **Standard:** 「人間の判断」を中核に据えたアプローチが、今後の日本標準となる。
- **Action:** 企業は3月末の最終版を注視し、コンプライアンス遵守と競争力維持（スピード）のバランスを見極める必要がある。