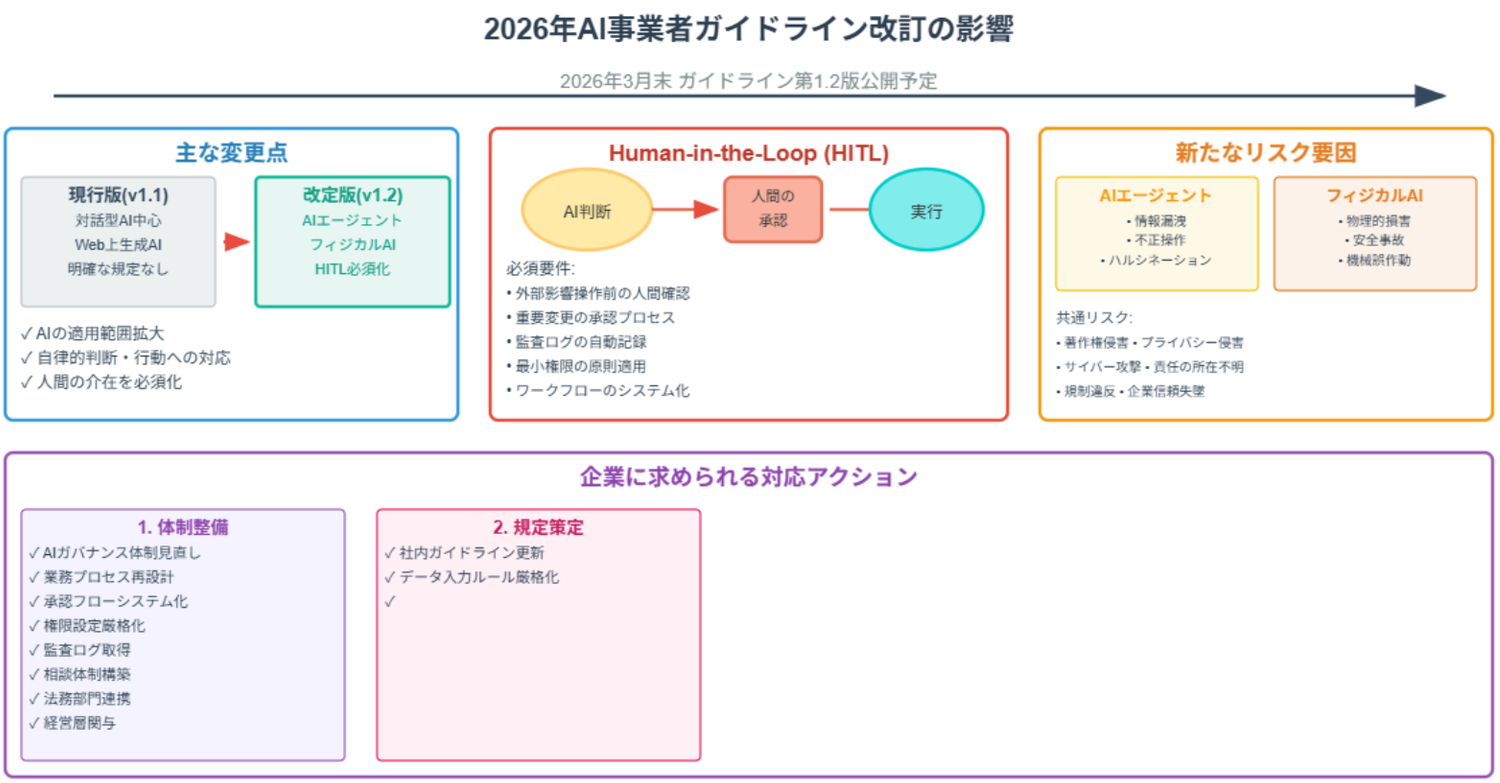
2026 年 AI 事業者ガイドライン改訂は、生成 AI 利用者にどう影響するか？

Felo AI



Overview

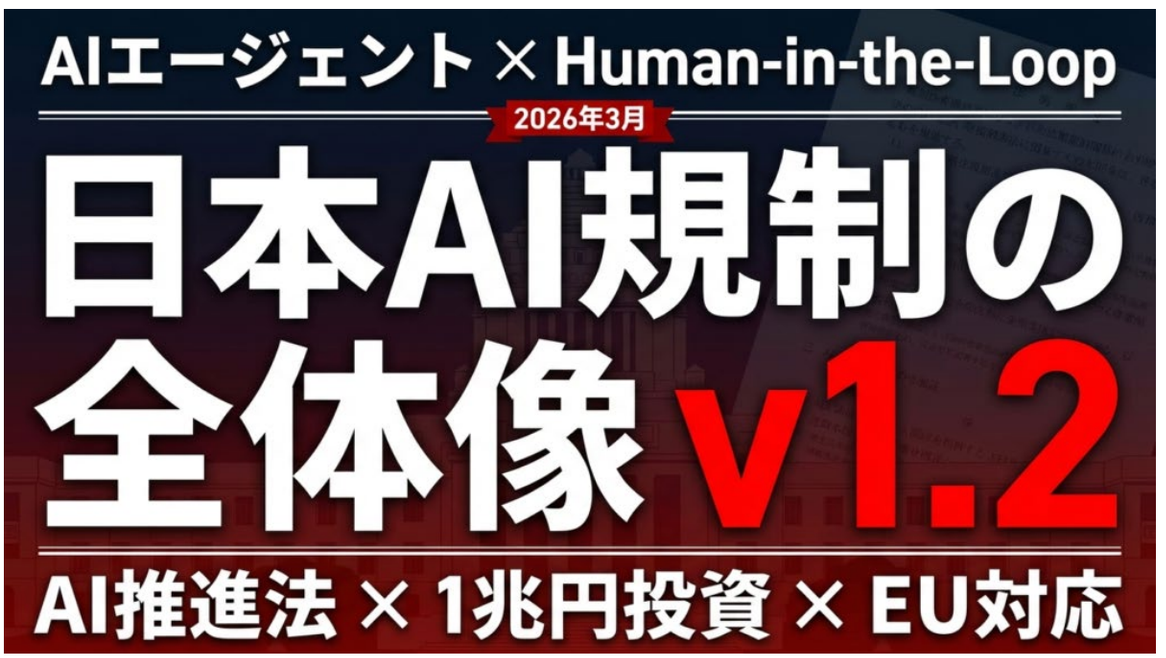
2026 年 3 月末に公開が予定されている「AI 事業者ガイドライン（第 1.2 版）」は、AI の利用範囲を従来の対話型 AI から、自律的に判断・行動し物理世界に影響を与える「AI エージェント」および「フィジカル AI」へと正式に拡張します[11]。この改訂の最も重要な点は、これらの自律型 AI が外部に影響を及ぼす重要な操作を行う前に、必ず人間の判断を介在させる仕組み「Human-in-the-Loop（HITL）」の構築を事実上の必須要件として明記したことです[2][11]。

これにより、生成 AI を利用するすべての企業は、AI ガバナンス体制の根本的な見直しを迫られます。具体的には、AI に任せる業務範囲と人間が最終承認を行う範囲を明確に区分した業務プロセスの再設計、AI エージェントによる情報漏洩や不正操作といった新たなリスクに対応する社内規定の策定、そして従業員への教育が急務となります [11] [20]。この改訂は単なるリスク管理の強化ではなく、安全なガードレールを設けることで、企業の AI 活用とイノベーションを加速させることを目的としています [11]。

詳細レポート

2026 年 AI 事業者ガイドライン改訂の背景と全体像

2026 年現在、企業の生成 AI 利用は試験導入の段階を終え、本格的な業務運用フェーズへと移行しています [12]。この変化に伴い、AI が自律的に業務を遂行する「AI エージェント」や、ロボットなどを通じて物理世界で動作する「フィジカル AI」といった新技術の活用が本格化しています [18] [37]。しかし、これらの AI は従来の対話型 AI とは異なり、自律的な判断で外部環境に直接的な影響を及ぼすため、既存のガイドラインではカバーしきれない新たなリスクが顕在化しています [1] [11]。



政府は、こうした技術の進化とビジネス利用の実態を踏まえ、AI 事業者ガイドラインの改訂に着手しました [7]。2026 年 2 月に提示された改定案（第 1.2 版）は、同年 3 月末に正式公開される予定です [2] [11]。この改訂は、2025 年 9 月に施行

された AI 推進法を補完し、日本の AI ガバナンス体制を急速に整備する動きの一環です 2 。
日本の AI 政策は、罰則を伴う厳格な規制を導入した EU の「AI Act」とは異なり、あくまでイノベーションを重視する「促進重視」の立場を取っています 2 。今回のガイドライン改訂も、AI の利用を禁止するものではなく、リスクを適切に管理しながら活用を加速させるための「ガードレール」と位置づけられています 11 。

AI 事業者ガイドライン v1.1 と v1.2 の主な変更点

項目	現行版（第 1.1 版）	改定版（第 1.2 版）
AI 概念の範囲	主に Web 上の生成 AI（チャットボット等）	**AI エージェント、フィジカル AI を正式に追加** 11
自律的な判断・行動	明確な規定なし	**Human-in-the-Loop（人間の判断必須）を明記** 11
ガバナンスの位置づけ	リスク管理の一環	イノベーションの**「加速装置」**として再定義 11
対象事業者	AI 開発者・提供者・利用者	同上＋**エージェント運用者を想定に追加** 11

新たな規制対象：AI エージェントとフィジカル AI

今回の改訂で新たに対象として明確化された「AI エージェント」と「フィジカル AI」は、AI 活用のトレンドが「生成」から「自律的な実行」へとシフトしていることを象徴しています 18 。

AI エージェント AI エージェントとは、人からの最小限の指示に基づき、目標達成のために自ら計画を立て、複数のツールやアプリケーションを連携させながらタスクを自律的に実行する AI です 18 26 。従来の生成 AI が文章や画像の「生成」に留まるのに対し、AI エージェントは顧客へのメール送信、発注業務、データ分析レポートの作成と共有といった一連の業務フローを自動化できます 18 30 。この能力は、個人の生産性向上だけでなく、チームや事業部門全体の業務効率を飛躍的に高める可能性を秘めています 18 。

フィジカル AI フィジカル AI は、AI の判断をロボットやドローン、産業機械などの物理的なデバイスの動作に直接結びつけ、現実世界でタスクを実行する AI を指します 4 9 。生成 AI とシミュレーション技術の進化により、仮想空間で大量の学習を積んだ AI が、これまで自動化が困難だった非定型的な物理作業（倉庫でのピッキング、建設現場での点検、介護施設での補助など）を行えるようになります 4 14 。少子高齢化による労働力不足が深刻化する日本において、フィジカル AI は危険作業や重労働を代替する解決策として大きな期待が寄せられています 4 22 。

これらの AI は、デジタル空間に留まらず、企業の事業活動や物理的な環境に直接的な影響を与える能力を持つため、ガイドラインの対象としてリスク管理の枠組みに含める必要性が生じました 7 11 。

利用者への最大の影響：「Human-in-the-Loop」の必須化

改定案が利用者に与える最も直接的かつ重大な影響は、「Human-in-the-Loop（HITL）」、すなわち「人間の介在」が事実上必須となる点です 2 11 。

「最終承認は人間」というルールの公式化 改定案では、AI エージェントやフィジカル AI が、外部に影響を与える操作（例：顧客への契約書送付、金融取引の実行、工場の機械制御）や、組織にとって重要な変更（例：社内データベースの更新）を実行する前に、必ず人間の確認・承認プロセスを挟む仕組みを設けることを求めています 11 17 。これは、AI が分析やドラフト作成といった補助的な役割を担い、最終的な意思決定と実行の責任は人間が負うべきであるという原則を公式に文書化したものです 11 20 36 。

企業に求められる具体的な対応 この要請に応えるため、企業は以下のような仕組みを構築する必要があります。

- **承認フローのシステム化**: AI エージェントが見積書を自動作成した場合、上長や担当者が内容を確認し、「送信」ボタンを押すまでは実行されない、といったワークフローをシステムに組み込む 11 。
- **権限設定の厳格化**: AI エージェントに与えるアクセス権限を、業務に必要な最小限に留める「最小権限の原則」を徹底する 17 。
- **監査ログの取得**: AI の判断プロセスと人間の承認記録をすべてログとして自動保存し、問題発生時に追跡・検証できるようにする 17 。

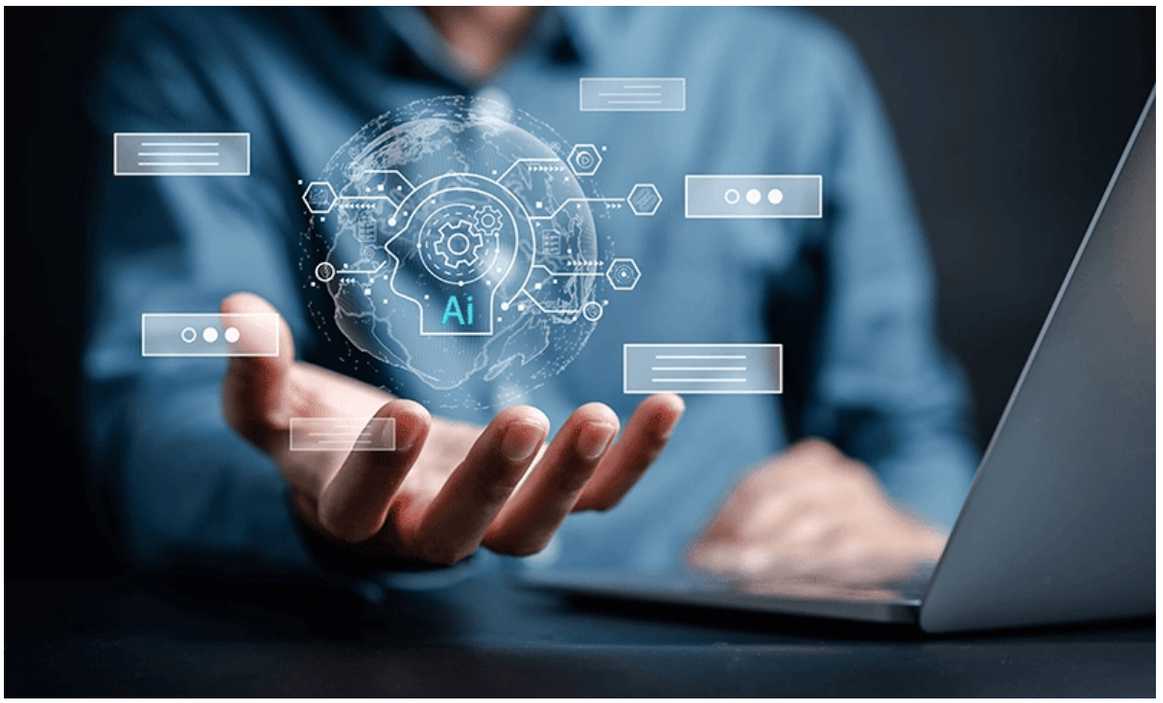
この HITL の原則は、AI の自律性を制限する「規制」ではなく、AI の利便性を損なうことなく、暴走や誤作動による重大な損害を防ぐための「安全装置」として設計されています 11 。

企業が直面する新たなリスクとガバナンス体制の見直し

AI エージェントとフィジカル AI の導入は、生産性を飛躍させる一方で、新たなリスクをもたらします。これに対応するため、企業はガバナンス体制を「チャットボット時代」からアップデートする必要があります 11 。

想定される新たなリスク

- **情報漏洩・プライバシー侵害**: AI エージェントが社内データベースや外部サービスを自律的に参照する過程で、個人情報や機密情報を意図せず外部に送信してしまう 1 。特に、RAG（検索拡張生成）技術では、参照データに含まれる情報が意図せず出力されるリスクがあります 1 。
- **不正操作・サイバー攻撃**: プロンプトインジェクション攻撃などにより AI エージェントが乗っ取られ、不正な送金指示やシステムへの攻撃を実行させられる 11 。
- **ハルシネーションによる損害**: AI が生成した事実と異なる情報（ハルシネーション）に基づき、誤った契約を締結したり、顧客に不正確な情報を提供したりすることで、企業の信頼を失墜させ、損害賠償問題に発展する 1 20 。
- **物理的な損害・事故**: フィジカル AI が誤作動を起こし、製品の破損、生産ラインの停止、さらには従業員や第三者への身体的な危害を引き起こす 4 13 。
- **著作権侵害**: AI が生成したコンテンツが、学習データに含まれる既存の著作物と酷似し、意図せず知的財産権を侵害してしまう 6 20 。米国の裁判所は、著作権の成立には人間の著作者が必要であるとの判断を確定させており、AI が単独で生成した作品の権利関係は依然として不安定です 8 。



社内ガイドラインの見直し これらのリスクに対応するため、企業は日本ディープラーニング協会（JDLA）などが公開するガイドラインも参考にしつつ、自社の社内規定を以下のように見直す必要があります [6]。

1. **データ入力のルール厳格化**: 個人情報、顧客の秘密情報、未公開の財務情報などを AI に入力することを原則禁止し、許可制とする [6] [20]。
2. **禁止用途の明確化**: 人命に関わる判断、差別を助長する可能性のある業務、非倫理的なコンテンツ生成など、AI の利用を明確に禁止する用途を定義する [6] [13]。
3. **出力物の利用ルール**: AI の生成物を鵜呑みにせず、必ず人間によるファクトチェックを義務付ける。また、生成物をそのまま利用せず、加筆・修正を加えることを推奨する [20]。
4. **責任の所在の明確化**: AI の利用によって問題が発生した場合の最終的な責任は、AI ではなく、それを利用・承認した人間および組織が負うことを明記する [13] [20]。
5. **相談体制の構築**: AI 利用に関する判断に迷った際の相談窓口を設置し、情報システム部門だけでなく、法務部門や経営層が連携して対応する体制を整える [6]。

今後の展望と企業が取るべきアクションプラン

今回のガイドライン改訂は、AI 活用の主戦場が「デジタル世界」から「現実世界」へと移行しつつある現状を反映したものです [18]。2026 年以降、AI エージェントやフィジカル AI をいかに安全かつ効果的に業務へ統合できるかが、企業の競争力を左右する重要な要素となります [23] [37]。

企業が今すぐ着手すべきこと

1. **現状把握とリスク評価**: 自社で利用中または導入検討中の AI ツールが、AI エージェントやフィジカル AI に該当するかを確認し、ガイドライン改定案に照らして潜在的なリスクを洗い出す 2 12 。
2. **社内体制の整備**: Human-in-the-Loop を組み込んだ業務フローの設計に着手する。特に、どのプロセスで誰が承認を行うのかを具体的に定義する 11 。
3. **ガイドラインの策定・更新**: 上記のリスク評価と業務フロー設計に基づき、実効性のある社内ガイドラインを策定または更新する 20 。
4. **従業員教育の実施**: 全従業員に対し、新たなガイドラインの内容と、AI 利用に伴うリスクについて周知徹底する。特に、安易な情報入力がもたらす危険性への意識を高めることが重要です 20 。

AI ガバナンスの構築は、単なるコンプライアンス対応やリスク回避のためのコストではありません。むしろ、従業員が安心して AI を活用できる環境を整え、AI によるイノベーションを加速させるための戦略的投資と捉えるべきです 11 。今回のガイドライン改訂を、自社の AI 活用戦略を一段階引き上げる好機と捉え、迅速に行動を起こすことが求められます。

1. 【2026 年最新】AI 事業者ガイドライン改訂の要点｜生成 AI 利用 ...
2. 【2026 年 3 月】日本 AI 規制の全体像｜ガイドライン v1.2「 ...
3. AI エージェントやロボ AI「人の判断必須の仕組みを」 政府指針に ...
4. フィジカル AI とは？注目されている背景、従来の AI との違い
5. 人機混合時代到來！對焦 2026 年九大趨勢，強渡 AI 職場轉骨期
6. JDLA の「生成 AI の利用ガイドライン」が改訂、"何が変わったか" ...
7. 政府、AI 事業者ガイドライン改定案で AI エージェントとフィジカル ...
8. 米最高裁「創作者は人間のみ」との判断を確定 AI 作品の著作 ...
9. フィジカル AI（Physical AI）とは？生成 AI との違い・特徴や ...
10. 2026 年 AI 會讓我失業嗎？
11. 【2026 年 3 月最新】AI 事業者ガイドライン改定とは？AI ...
12. 【2026 年最新】AI 事業者ガイドライン改訂の要点｜生成 AI 利用 ...
13. AI 活用における倫理問題とは？ 企業は何に留意すべきか
14. 生成 AI がもたらすロボット技術の進化 ―フィジカル AI の動向
15. 2026 生成式 AI 六大趨勢與企業落地策略 ― 超智諮詢
16. 生成 AI ガイドラインとは？企業が知っておくべき目的と重要性
17. 【自律型 AI にも“人の最終判断”を】3 月末の政府 AI 指針改定を ...
18. 生成 AI からエージェント AI へ AI 活用のトレンド変化を解説 ...
19. 比爾蓋茲發布 2026 年度信：「壞人亂用 AI」 ― 數位時代

20. 【企業向け】生成 AI 社内ガイドライン策定の進め方と注意点
21. AI は本当の味方?データで見る「おべっか問題」とキャリアに ...
22. 寺澤 豊：第 1 回 なぜ今、フィジカル AI が注目されるのか
23. AI 新賽局，2026 開戰！生成式應用深化，重塑消費、職場與創新生態
24. 「行政の進化と革新のための生成 AI の調達・利活用に係る ...
25. 肯定されすぎにご注意を。生成 AI との程よい距離感を保つ 3 つの ...
26. 目的達成に向けて自律的に業務遂行する「AI エージェント」の最 ...
27. 生技濫用 2026《國際人工智慧安全報告》示警 6 類變革與新風險
28. 生成 AI のリスクを整理する｜3 つの観点でリスクと対策を解説
29. 人工知能による判断の自動化と道徳的問題
30. AI エージェントとフィジカル AI――「現実世界に影響を与える ...
31. 2026 上看 4 千億美元生成式 AI 消費支出年增狂飆
32. 生成 AI の利用について
33. 生成 AI が増幅する認知バイアスの危険性 － ICR
34. 迫る AI 社会と主要業界の変貌：AI による破壊と創造
35. 資策會公布 2026 年十大 AI 關鍵技術：AI 從虛擬躍入實體 － iThome
36. 人間と AI の協働：HR 領域における不確実性の壁と倫理的判断 ...
37. 2026 年の生成 AI はどうなるのか
38. 《富比士》預測 2026 生成式 AI 的 10 大趨勢！預料將全面重塑工作與 ...