



AI の倫理と AI ガバナンスに 投資すべき理由

5つの最前線の事例から探る

ノートルダム IBM テクノロジー倫理ラボとの共同レポート



主なポイント

AI 倫理の価値を評価することで、組織は競合他社に先んずることができるようになる。IBM が提唱する「包括的 AI 倫理フレームワーク」では、3 種類の ROI を考慮する。

■ AI 倫理への取り組みは不可欠である

それは損失を回避するためだけではない。経営層の 75% が、AI 倫理を重要な競争上の差別化要因だと見なしている¹。調査対象の消費者・一般市民・従業員の 85% 以上が、AI 倫理を重視している²。

■ 長期的で積極的な AI 倫理戦略が組織全体で価値を生み出す

過半数（54%）の組織が今後、AI 倫理は戦略的な面から非常に重要になると予想しており³、関与する部署は 20 に及ぶと経営層は述べている⁴。

■ AI 倫理への投資は定量化可能な価値を生み出し得る

AI 倫理の価値を評価することで、組織は競合他社に先んずることができるようになる。IBM が提唱する「包括的 AI 倫理フレームワーク」では、経済面における影響（有形）、企業イメージへの影響（無形）、能力向上（リアル・オプションの ROI）の 3 種類の ROI を考慮する。

はじめに

生成 AI はさまざまな業界に革命をもたらしつつある。しかし、目覚ましい普及に伴い、倫理上の問題が目立つようにもなってきた。潜在的なメリットと、倫理・規制面の影響とのバランスを取ることが重要である。

とはいえ、それは簡単なことではない。IBM Institute for Business Value (IBM IBV) の調査によると、ビジネス・リーダーの 80% が生成 AI を導入にするに当たり、AI の説明可能性 (explainability)、倫理、バイアスおよび信頼の問題が大きな障壁となり得ると考えている⁵。また半数が自社には生成 AI の倫理面の課題に対処するために必要なガバナンスと体制が欠けていると回答した⁶。

この不確実性とリスクに直面し、多くの CEO が対応に躊躇 (ちゅうちょ) するようになった。経営層の半数以上 (56%) が AI の基準や規制が明確になるまで生成 AI への大規模な投資を遅らせると答え⁷、また 72% が倫理面の懸念を理由として実際に生成 AI のメリットを諦めるだろうと回答している⁸。

しかし、経営層が視野を広げ、AI 倫理を 1 つのチャンスとして捉えるならば、前進の道が拓けるだろう。さらに良いことに、現在進めている調査によれば、AI 倫理への投資は定量化可能なメリットを生み出す可能性がある。

この可能性を現実化するためには、AI 倫理への投資の ROI を評価する際に、新たな視点を取り入れる必要がある。本レポートの第 1 部では、AI 倫理に適用すべき 3 つの主要な ROI を特定している。すなわち IBM が提唱する「包括的 AI 倫理フレームワーク」だ。また第 2 部と第 3 部では、AI 倫理への投資の妥当性を示す、2 つの異なる有益な方法について詳述する (今回の研究に基づき、さらに詳細な定量化を模索する追加リサーチを 2025 年に実施する予定)。最後に、組織内で包括的 AI 倫理フレームワークを実際に活用するためのアクション・ガイドを示す。

また、IBM IBV、ノートルダム IBM テクノロジー倫理ラボ、IBM AI 倫理委員会、および IBM Office of Privacy and Responsible Technology (IBM プライバシー・責任あるテクノロジーオフィス) が進める共同プロジェクトの一環として、AI 倫理の最前線で活躍する 5 人の経営層の事例を紹介する。一部のインタビューはオックスフォード・エコノミクス (Oxford Economics) 社の協力を得て実施した。

包括的 AI 倫理フレームワークとは何か⁹

AI 倫理と AI ガバナンスへ投資する場合、その範囲は組織全体にまで及ぶ可能性がある。対象範囲には、AI 倫理委員会から、エシックス・バイ・デザイン（設計段階から倫理を組み込む）手法まで、さらに統合ガバナンス・プログラム（プライバシー、データ・ガバナンスなどを統合したガバナンスの取り組み）から、AI 倫理・AI ガバナンス関連の研修プログラムまで、数多くの取り組みが含まれる¹⁰（13 ページの「AI 倫理: 最前線の事例」を参照。また、IBM IBV の研究「企業のための AI ガバナンス・ガイド」[ibm.co/ai-governance] も参照）。

では、こうした取り組みの効果を評価するにあたって、何から始めればよいのだろうか。その問いに答えるため、私たちは 30 を超える組織との幅広い対話から検証を行い、包括的 AI 倫理フレームワークを開発した。このアプローチを使えば、自社の AI 倫理・AI ガバナンスへの投資がどれほど有益かを理解することが容易となる。

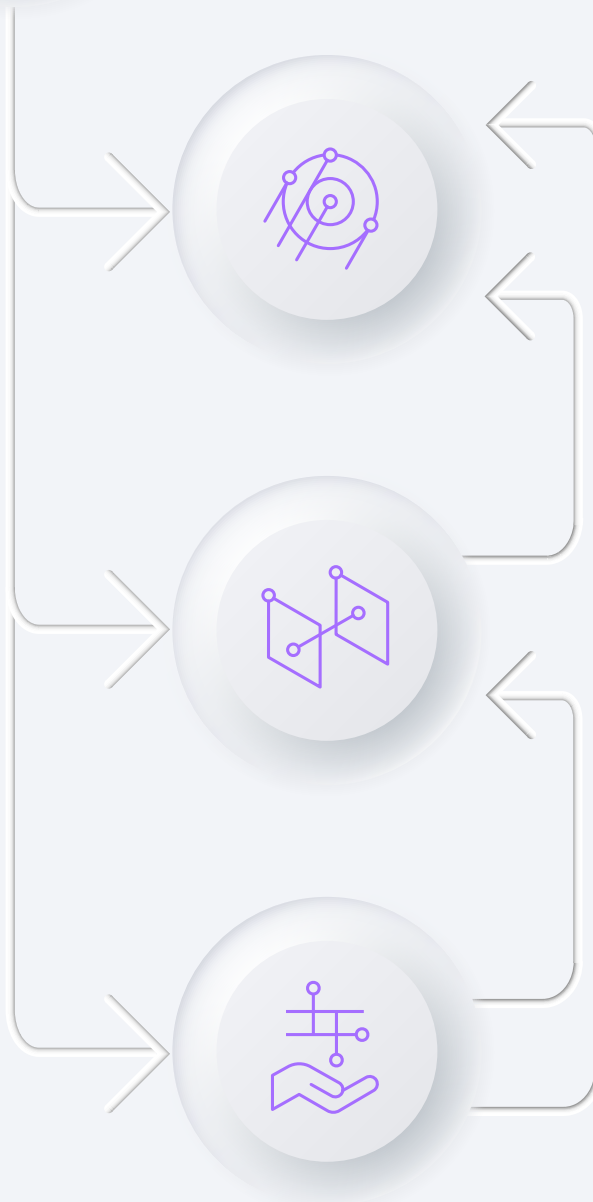
従来、投資は財務面の ROI だけを検証して、評価されてきた。AI 倫理への投資は評価がより難しい。有形無形の両面でメリットをもたらし、長期的な能力の構築にも役立つからだ。

「私たちの仕事は、組織のミッションに貢献するだけでなく、組織の利益率にも貢献することだ」と SAS 社の Data Ethics Practice（データ倫理実践）担当バイス・プレジデントである Reggie Townsend 氏は指摘する。「さもないと、慈善活動のようなものになってしまう。慈善活動では、長期にわたって利益を得ることはできない」

私たちは 30 を超える組織との
幅広い対話から検証を行い、
包括的 AI 倫理フレームワークを
開発した。



包括的 AI 倫理フレームワークは、
組織が AI 倫理への投資において
考慮すべき 3 つの ROI を特定する。



経済面における影響（有形 ROI）

コスト抑制、収益増加、資本コストの削減など、AI 倫理への投資による直接的な財務上のメリットのこと。例えば、AI のリスク管理に投資することで規制による罰金を回避できる。

企業イメージへの影響（無形 ROI）

重要でありながら定量化が難しい要素のことで、株主・政府・従業員・顧客による組織の評価に望ましい結果や影響を促すブランド力や組織文化など。例えば、環境・社会・ガバナンス（ESG）のスコア改善や、従業員の定着率向上、ポジティブな報道などが挙げられる。

能力向上（リアル・オプション ROI）

最初は AI 倫理のために確立された能力が、時間を経て組織全体で応用され、最終的に広範な領域で価値をもたらすこと。例えば、倫理のための技術的インフラや特定のプラットフォームが、さらなるコスト抑制とイノベーションを促す最新化を可能にすることがある。

出典：“The Return on Investment in AI Ethics: A Holistic Framework.” 第 57 回システム科学に関する HICSS 年次会議議事録。2024 年 1 月

上述の包括的 AI 倫理フレームワークは、AI 倫理への投資が利害関係者にもたらす影響を理解するための、3つの道筋を示している。それは、経済的利益を通じた直接的な道筋、企業イメージを通じた間接的な道筋、そして能力を通じた間接的な道筋である。このフレームワークは、組織が AI 倫理に投資した場合の関係性や利害関係者、潜在的リターンを包括的に捉え、説明するものである¹¹。

では、これらは実際にどう機能するのだろうか。AI 倫理委員会がインフラやスタッフへ投資するケースについて考えてみよう。この投資により、規制による罰金を回避し（有形の影響）、クライアントからの信頼とパートナーからの支持を得て、ビジネスチャンスを増やし（無形の影響）、文書作成の自動化とデータ管理を共に改善する管理システム・ツールを容易に開発できるようになる（能力）。組織は包括的 AI 倫理フレームワークを活用することで、組織全体の実践と成果の双方に AI 倫理が密接に関わっていることを理解するだろう。

組織は包括的 AI 倫理フレームワークを活用することで、組織全体の実践と成果の双方に AI 倫理が密接に関わっていることを理解するだろう。



「損失を回避する」 ことの価値

AI 倫理とは何だろうか。フィデリティ投信のデータ・ポリシー担当シニア・バイス・プレジデントは次のように簡潔に説明する。「顧客や見込み客などとコミュニケーションを取る際に、何が正しく、何が間違っているのかを見極められるように、責任ある形で AI 技術を使用することである」

最近、IBM IBV が行った調査によると、経営層の 72% は、倫理的なコストを伴うと考えられる場合、生成 AI への取り組みから手を引くと回答した。こうした経営層の組織は、収益成長の面で他社を上回る可能性が 27% 高く、この相関関係は無視しがたい¹²。

しかし、AI 倫理への取り組みは、口先だけで行動が伴わない場合が多い。AI 倫理の原則について公に賛同する組織は調査対象の半数を超えるが、それを実行できている組織は 4 分の 1 にも満たない¹³。AI 倫理に関する自社の原則や価値観（バリュー）と、実際の行動が一致している（もしくは、より望ましい行動が取れている）と確信している組織は 20% 未満だった¹⁴。

「行動する意欲があるのは良いことだが、実際に行動しなければ意味がない」と、ある世界的金融サービス企業の AI ガバナンス担当シニア・リーダーは語る。「ただ、行動するためにはリソースが必要であり、そのためには資金を要する。さらに必要なのは、経営層の意志だ」

では、AI 倫理へ投資する上でビジネス上のメリットは何だろうか。多くの場合、AI 倫理への投資は損失回避のアプローチから始まる。具体的には、規制順守に関連するコスト増を回避することや、AI 倫理を重視する企業に顧客を奪われ収益を失わないようにすることである。

AI 倫理への取り組みは、口先だけで
行動が伴わない場合が多い。

損失回避に対する動機が短期的戦略を反映するものであっても、その重要性が損なわれることはない¹⁵。損失回避はすぐに成果をもたらす。世界的金融サービス企業で AI ガバナンスを担当する上述のシニア・リーダーが指摘するように「ビジネスで肝心なのは、風評リスクをなくすことだ」

損失回避の例：¹⁶

規制に関連する例

規制違反による罰金の回避

法的コストの回避

必要な技術的コンプライアンス・メカニズムの導入

コンプライアンス要件を満たす

顧客・パートナー・競合他社に関連する例

利害関係者の懸念を払拭

ビジネスモデルに対する脅威の回避

特定の顧客需要や顧客ニーズへの対応

ブランドイメージの保護

競合他社に後れを取らない

AI 規制当局からの働きかけ

AI 規制は行動を促す原動力となる。EU AI Act（EU AI 規則）は、主要機関による初の総括的な AI 規制である。ドイツテレコム社の戦略マネージャーは次のように述べている。「EU AI 規則には、AI 倫理の様相を世界的に一変させる可能性がある。例えば、もし米国企業が当社と協働する場合、その企業も EU AI 規則を順守しなくてはならない」

この EU の取り組みは始まりに過ぎない。AI に対する責任あるアプローチに関して、「Partnership on AI」や「人工知能（AI）に関するグローバル・パートナーシップ（GPAI）」、世界経済フォーラム、国連、経済協力開発機構（OECD）などの組織が、原則やガイドラインを公開している¹⁷。11 カ国の 1 万 3,000 人以上を対象とした「Centre for the Governance of AI」の調査によると、91% が AI は慎重に管理すべきだと答えている¹⁸。

このように規制や監視、AI に対する責任あるアプローチが重要視されているため、損失回避に努めることは賢明であるだけでなく、不可欠なことである。

AI 倫理を活用して 価値を創出する

AI 倫理に投資するメリットは、コストを回避し、損失を抑制することだけではない。組織の長期的な戦略を実行する上で、有用な能力を構築したり、イノベーションの具体的成果を上げたりする後押しともなる¹⁹。このような価値の創出は、損失回避よりも間接的であり、ROI をより広い視点から見ることで可能になる。また、一夜にして達成できるものではなく、計測可能な成果が現れるまでには時間がかかることもある。

しかし、AI 倫理を深く理解する組織は、投資によって以下のような効果を期待できる²⁰。

- 責任を持って AI を大規模に活用するための長期的プランが可能になる
- 独自の価値ある能力を構築し、差別化を可能にする
- 従業員の効率性や生産性の改善
- 業界リーダーとして前進するため、価値観を明確にする
- 市場機会を捉える
- 弱い立場の個人やコミュニティの保護
- 顧客満足度を高める
- 信頼性と成熟度の高さを示す
- 環境・社会・ガバナンス（ESG）への取り組みを支援
- 長期的なリスク管理の能力を高める
- イノベーションによって競争力を得る

AI 技術が成熟するにつれ、組織は AI を単に業務に組み込むだけでなく、その技術を新たなイノベーションのために応用できるようになる。健康・消費財のある大手小売企業のシニア・ディレクターは次のように述べている。「AI を活用して顧客体験を創出・改善する手法に基づいて、ブランドの採用、売り上げの向上、顧客の維持、顧客数の増加という形で成果を確認することができた」

両方の利点を 組み合わせる

損失回避と価値創出の両方を包括する
総合的アプローチを取る組織は、
より倫理的である上に、効率性と有効性が高く、
より大きな成功を収める傾向にある。



出典：“On the ROI of AI Ethics and Governance Investments: From Loss Aversion to Value Generation.” California Management Review. 2024 年 7 月 29 日

フィデリティ投信のシニア・バイス・プレジデントは次のように述べている。

「先行投資をすることによって、倫理的な効果だけではなく、個別のユースケースごとに新たな投資を行わずに済み、そうした点からも非常に大きな投資対効果（ROI）が得られることに多くの企業が気付いていない。基礎工事（先行投資）をしないまま 1 本ずつ柱（個別のユースケース）を建てていったら、より多くの費用がかかり、再利用によるメリットも得られないからだ」

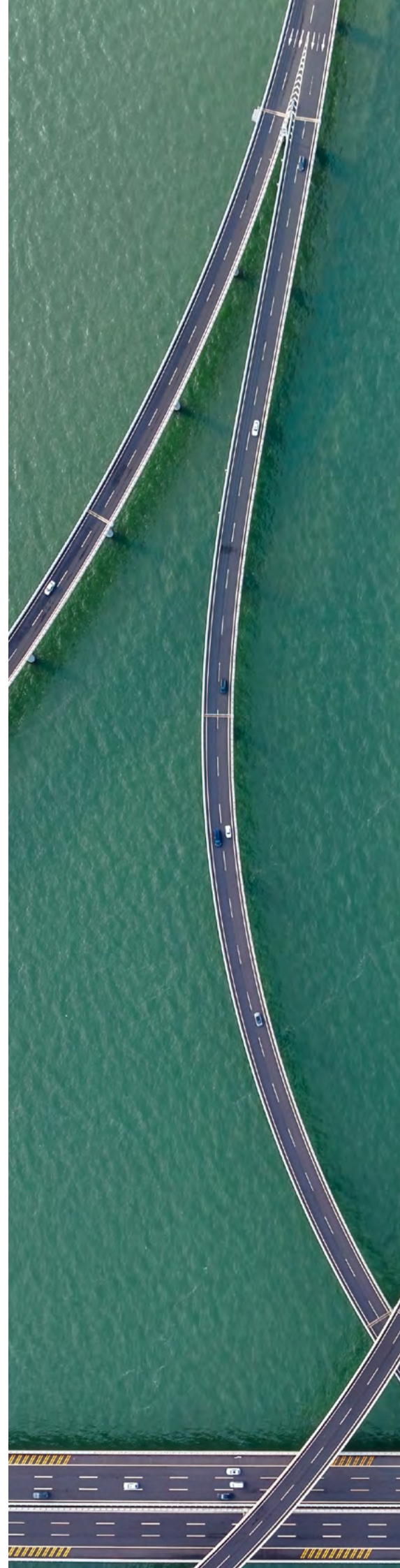
この前段階として、組織戦略に即し、その支えとなる AI のユースケースを実際に開発することが当然必要になる。ドイツテレコム社の戦略マネージャーは以下のように指摘する。「開発すべき AI ソリューションは顧客向けでも社内インフラ向けでも、どちらでもよい」

最初の時点では、規制による罰金や法的コスト、および風評被害といったリスクを回避するために損失回避の「守り」の姿勢を取ることは、自然なことであり、理にかなったことである。

しかし、価値創出へ焦点を移すことで、可能性は大きく広がる。組織は、AI 倫理の取り組みを促進し、より広範な分野で企業改革を進められるよう、汎用性のあるテクノロジーとインフラ、プラットフォームを開発する必要がある。

こうした戦略を実行するため、支援や予算を確保することは容易ではない。懐疑的な人々を説得し、障害を乗り越えなくてはならず、そのためには経済的利益の指標や、潜在的な価値を明確にする必要がある。組織が投資の計画を立て、その計画を評価する際には、関連する損失回避および価値創出のメリットを特定するプロセスが必要となる²¹。基本的には、包括的な AI 倫理フレームワークを活用すべきである。

価値創出へ焦点を移すことで、可能性は大きく広がる。組織は、AI 倫理の取り組みを促進し、より広範な分野で企業改革を進められるよう、汎用性のあるテクノロジーとインフラ、プラットフォームを開発する必要がある。



アクション・ガイド

AI 倫理への投資から 成果を得る方法

AI 倫理への投資は、正しい行動であるだけでなく、ビジネスにおける堅実な決断である。包括的 AI 倫理フレームワークを用いることで、組織は AI 倫理へのリソースの配分を十分な情報に基づいて行えるようになり、AI プログラム全体の信頼性と可能性を高められるようになる。

IBM IBV の調査によると、経営層の 75% は倫理が競争上の差別化要因として重要だとみている²²。エコノミスト・インテリジェンス・ユニット（EIU）の調査もこの結果を裏付けており、倫理は製品の品質、優秀な人材の獲得と維持、および新たな収益源による競争力の向上に寄与すると指摘している²³。

こうした調査結果は、AI 倫理への積極的アプローチの重要性を浮き彫りにした。組織は、AI ガバナンスがかつてのテクノロジー・ガバナンスとは異なることを考慮しつつ、組織文化やエコシステム、顧客エンゲージメントの隅々にまで AI ガバナンスを浸透させねばならない。

「AI エンジン人間の思考に基づいて教育すべきだ」と、健康・消費財のある大手小売企業シニア・ディレクターは語る。「倫理の観点から見て、人間は優れた審判を下せるからだ」

また SAS 社の Reggie Townsend 氏は次のように語る。「当社にはさまざまなバックグラウンドや人生経験を持った多様な人材がいる。私たちは困難な仕事をこなすが、心を込めて仕事に当たっている。それができない人は採用しない。当社には情熱にあふれた人材がいて、その情熱を仕事に注いでいる。それこそが根本的に重要なのだ」

AI 倫理への投資を最適化するための 5つのステップ

1

AI 倫理に最も精通した専門家を起用して、AI 倫理に対する損失回避と価値創出のアプローチの違いを経営層に理解させる。経営層が AI 倫理のテクノロジーやプラットフォーム、インフラを広範に活用するイメージを思い描けるようにする。

2

AI 倫理およびガバナンスからメリットを生み出すための、具体的な価値創出の方法を探る。(例：顧客からの質問に対し責任感を持って適切に回答することや、従業員の生産性を上げること、仕事に対する満足度を向上させることなど)

3

AI のユースケースから予想される利害関係者への影響を考慮し、潜在的な指標を特定する。例えば以下のようなものが考えられる。

- 直接的な経済的利益
(例：顧客基盤を拡大することで得られる価値)
- 無形の企業イメージから得られるリターン
(例：顧客レビューにより媒体価値を高める)
- リアル・オプションから得られる能力や知識面のリターン
(例：顧客対応の質を向上させ、一次接触における解決率を高める)

4

AI 倫理に関する導入実践戦略を立て、価値創出のメリットを実現する。アクション「3」の分析を用いて、潜在的なリターンを総体的に把握する。そうすることで、AI 倫理および AI ガバナンスへの投資の潜在的なリターンを最適化しながら、同時に利害関係者やエコシステム、社会に利益をもたらすことができるようになる。

5

価値創出を競争上の優位性につなげる。規制順守が当たり前のこととなっている状況の中で、価値創出へ注力することは、競争上の優位性をもたらす。

包括的 AI 倫理フレームワークのより詳細な情報とアクションについては、“On the ROI of AI Ethics and Governance Investments: From Loss Aversion to Value Generation,” California Management Review (<https://cmr.berkeley.edu/2024/07/on-the-roi-of-ai-ethics-and-governance-investments-from-loss-aversion-to-value-generation/>) と、“The Return on Investment in AI Ethics: A Holistic Framework” (<https://arxiv.org/abs/2309.13057>) を参照のこと。



AI 倫理

最前線の事例

ドイツ テレコム社

EU AI 規則に
備えるための
社内ガバナンスと
教育

ドイツテレコム社のデータ・イニシアチブは、AI の活用を通じたデータ事業の収益化と、EU AI 規則の監視を担当している。同社のある戦略マネージャーが率いるチームは、社内で行われる AI 関連の議論のほぼすべてに関与しているため、同マネージャーは自社の AI 倫理へのアプローチについて全体像を把握している。

ドイツテレコム社は、現在と未来の AI イニシアチブを評価するため、上級経営層から成るチームを立ち上げた。企業全体の AI ガバナンスを差配する重要な組織である。その最も重要な目的は、データ・プライバシーやセキュリティの手続きに関して、同社が社内外でコンプライアンス要件を満たすようにすることであり、その対象には EU AI 規則も含まれる。

「AI はデータがすべてである。いかなる AI 製品においてもデータが基本要素だ」とこのマネージャーは語り、AI 倫理の実践において、データがいかに重要な要素であるかを強調する。同社ではデータを製品に組み込む前に、誰がそのデータに接するのか、顧客データをどのように保護するのかについて検討する。顧客データに加え、サステナビリティとエネルギー・プラクティスの観点から、特定のデータ・セグメントも保護すべきだと考えている。

ここで重要な点として挙げられるのは、ドイツテレコム社は AI とその倫理的使用について従業員を教育するために多額の資金を投じているということだ。その多くは、EU AI 規則に関連する研修などのワークショップに費やされている。

「社員の研修は間違いなく ROI を向上させる。市場投入までの時間を短縮し、より革新的な製品の開発に結び付くからだ」とこのマネージャーは指摘する。ドイツテレコム社は、継続的に改善を繰り返すことで、一層大規模にイノベーションを実現し、顧客からの信頼を獲得している。

フィデリティ 投信

ユースケースを
応用して、
ROI を高める

フィデリティ投信では、責任ある AI への対応の精神が、AI コースケースの各フェーズに組み込まれている。その範囲は厳格なデータ管理の実践から、AI のセンター・オブ・エクセレンス*が主導するダイナミックなレビュー・プロセスにまで及ぶ。金融サービスを提供する同社は、これらの取り組みに多額な投資を行い、AI 倫理を単なるコンプライアンス事案として片付けるのではなく、自社の存立基盤の 1 つと見なしている。

フィデリティ投信の各事業部門には、AI のコースケース開発とベンダー管理に特化したチームが置かれている。この取り組みでは、チームは専門知識を有する社外コンサルタントから指導され、専門的な AI 研修を受けた社内のコンプライアンスおよびリスク担当者による厳しい監視を受けている。

AI のセンター・オブ・エクセレンスは、ベンダーの選定からモデルの評価に至るまで、プロセスのあらゆる段階に関与している。この組織は社内のデータ部門に属し、各事業部門から集められた代表者で構成されている。その役割はリスク・コンプライアンスや監査、法務、さらには情報セキュリティにまで及ぶ。フィデリティ投信はこのプロセスを通じて、AI の活用およびガバナンスに関する情報を一層求めるようになったクライアントに対して、自信を持って応えることができるようになった。

プロジェクトが遅れたり、ユースケースが制限されたりすることがあるため、責任ある AI の取り組みに対し抵抗が起きることは避けて通れない。「単に形式的にコンプライアンス・ポリシーを守るのではなく、管理をしっかりと行うことが重要である。それは対応を誤れば会社に大きな損害を及ぼすからであり、そのことをしっかりと社員に説明しなくてはならない」と、同社のデータ・ポリシーを担当するシニア・バイス・プレジデントは語る。フィデリティ投信はこうした取り組みを AI プロジェクトの成功に欠かせないものとして位置付け、ガバナンス・プロセス全体を合理化することにより、抵抗を最小限に抑えることに成功した。

* 組織を横断する取り組みのための、優秀な人材やノウハウを 1 カ所に集約した拠点のこと

「単に形式的にコンプライアンス・ポリシーを守るのではなく、管理をしっかりと行うことが重要である。それは対応を誤れば会社に大きな損害を及ぼすからであり、そのことをしっかりと社員に説明しなくてはならない」

フィデリティ投信のデータ・ポリシー担当シニア・バイス・プレジデント

SAS 社

AI を活用し、
誰もが恩恵を受けられる
未来を実現する

SAS 社の Data Ethics Practice 担当バイス・プレジデントである Reggie Townsend 氏は、世界的組織の全体にわたって責任あるイノベーションの原則やオペレーション・ワークフロー、およびガバナンス体制を取り仕切るチームの責任者である。すべては疑問を呈し調査することから始まった。

社会的弱者が直面するリスクや AI の高度化を受けて、Townsend 氏と身近な同僚たちは、責任ある AI およびデータ倫理に関する詳細な研究を SAS 社で開始した。彼らは社内のリーダー層による支援を受けて、責任あるイノベーションに対する同社の長年にわたる取り組みを具体的な形に結実させようとした。その結果、SAS 社によって創設されたのが Data Ethics Practice (DEP) である。

DEP は「エシカル・バイ・デザイン（設計段階から倫理を考慮する）」という理念のもと、従業員や顧客がデータ・ドリブン（データに基づき戦略立案や意思決定などを行うこと）なシステムを導入できるよう支援する社内の取り組みを主導し、それによって人々のウェルビーイング（心身の健康・幸福）や主体性、公正性を高めることを目指した。このアプローチでは、個人に対し 3 つの基本的な質問に答えることを求めている。

- － 何が目的なのか？
- － 何を達成したいのか？
- － 失敗するとしたら、誰のせいなのか？

DEP のチームは「信頼できる AI」の能力やワークフローの構築を支援することで、顧客や開発者による「責任ある AI」の目標達成をサポートする。DEP による AI ガバナンスに関するアドバイザリー・サービスは、顧客が責任ある形で AI を活用するのに役立っている。DEP は製品開発やマーケティングなどについての重要な助言を従業員に提供している。

Townsend 氏とそのチームが動き始めた時に期待されたのは、彼らの取り組みが製品やプロセス、人材に対する信頼を高めるといったことだった。ひいてはそれが「信頼できる AI リーダー」というブランド評価を高めることにもつながるだろう。

利益はもちろん重要だ。しかし、Townsend 氏によれば、チームの基本理念は「SAS のソフトウェアがどんな場面でも害をもたらさないようにする」というものだ。「時には行動を起こさねばならない。それが正しい行動であるなら」

「時には行動を起こさねばならない。
それが正しい行動であるなら」

SAS 社、データ倫理部門ディレクター兼バイス・プレジデント
Reggie Townsend 氏

世界的金融 サービス企業

風評リスクを
低減することで投資の
正当性や利益を
確かなものとする

ある世界的金融サービス企業で AI ガバナンスを担当する某シニア・リーダーにとって、AI の開発と倫理は教育から始まる。同氏は倫理原則や価値観について話し合うワークショップを開催し、リーダー層がトレードオフを理解できるようにすることの重要性を主張する。また、「プロセスや手続きだけでなく、リーダー層が関心を持てるような形で AI について話さなくてはならない」と語る。

AI 倫理投資のリターン（ROI）を測定する方法について議論するとき、同氏は「グレーゾーン」という比喩的な表現を用いる。多くの場合、組織は完全に合法的な行為であっても、大きな利益が上がっている場合、自らの行動の倫理性に不安を覚え、「グレーゾーン」に入り込んでしまったと感じてしまう。

このような状況下では、組織は総合的に倫理感を配慮しつつ、現在と未来の世代の両方の視点を通して活動を精査しなくてはならないと、同氏は述べる。こうした配慮が十分になされれば、組織は「グレーゾーン」に入っていないという確信が持てるはずだ。

同氏はまた、以下のようにも述べている。「風評リスクが低ければ、投資の正当性や利益はより確かなものになる。当社はデータと AI の倫理原則を適用しながら、風評リスクを低減することを目指している」

例えば、同社の担当チームは、信用評価アルゴリズムなどを使い、融資金利の倫理的公平性について評価を行っている。この評価を実施する際に、チームは AI モデルの 165 の特性をすべて分析し、特定のデータ特性と個人のローン返済能力との間に因果関係があるかどうかを調べた。

最終的に 3 つのデータ特性が、因果関係がないとして削除され、AI 技術を用いることで生じる公平性の欠如を回避した。

「プロセスや手続きだけでなく、
リーダー層が関心を持てるような形で
AI について話さなくてはならない」

世界的金融サービス企業の AI ガバナンス担当シニア・リーダー

健康・消費財のある大手小売企業

周到的 AI 倫理・ガバナンス戦略で成功に導く

この企業のシニア・ディレクターは、ベンダーや社内の製品を通じてソリューションを提供するために、AI プロジェクトを立ち上げた。同氏と最近行った対話では、3つの主要なオペレーション領域が話題となった。

厳格なガバナンス・プロセス：この小売企業の AI ガバナンス・グループは、すべての AI イニシアチブが承認に必要な手順を満たしていることを保証する中枢機関である。その一環として同グループは、それぞれのプロジェクト・チームが管理プランに基づいてコンプライアンス対策を策定したことを示すためのセッションを開いている。そこで承認されれば、そのプロジェクトは進められることを許される。多くのパートナー、サプライヤー、顧客、その他のエコシステムと関わる大規模な企業として、自社は AI 能力の構築について極めて慎重にならなければならないと、同氏は語る。

AI 倫理エンジン：同社は SaaS やベンダー、オープンソースの各製品に投資する場合、すべての倫理的パラメーターを満たしていることを事前に確認している。同社の社内監査プロセスは「AI 倫理エンジン」と呼ばれている。同社ではベンダーと契約する場合、最初にバックグラウンド・チェックを行い、業界・クライアント・企業イメージ・能力の健全性を調査する。このプロセスには通常、2カ月から4カ月を要する。ベンダーを選定した後には、試験運用を実施する。これが成功し、倫理基準を満たしていることが確認された後、パートナーシップを締結する。

利害関係者の成功：同社は AI 機能に多額の投資を行うことで、顧客エンゲージメントの体験を向上させ、市場戦略を展開し、顧客を増やそうとしてきた。同氏は「AI だけ、もしくは人間だけでは成功できないが、両者を組み合わせることで、より効果的で正確な成果が得られる」と述べている。

例えば、同社では、カスタマー・サービスに AI 機能を導入したが、それは同サービスの担当者にとって代わるものではない。むしろ、従業員にさらなるスキルを付与するために投資しており、その結果、従業員の定着率は向上している。このアプローチは顧客、従業員、そして同社の収益にメリットをもたらす。

「AI だけ、もしくは人間だけでは成功できないが、両者を組み合わせることで、より効果的で正確な成果が得られる」

健康・消費財のある大手小売企業のシニア・ディレクター

著者

Nicholas Berente

Senior Associate Dean for Academic Programs
Professor of IT, Analytics and Operations
University of Notre Dame,
Mendoza College of Business
(ノートルダム大学、
メンドーサ・カレッジ・オブ・ビジネス、
学術プログラム担当上級副学部長、
IT、アナリティクスおよびオペレーション教授)
[linkedin.com/in/berente/](https://www.linkedin.com/in/berente/)
nberente@nd.edu

Marialena Bevilacqua

PhD Student in Analytics
University of Notre Dame,
Mendoza College of Business
(ノートルダム大学、
メンドーサ・カレッジ・オブ・ビジネス、
アナリティクス博士課程学生)
[linkedin.com/in/marialena-bevilacqua-6848b9132/](https://www.linkedin.com/in/marialena-bevilacqua-6848b9132/)
mbevilac@nd.edu

Heather Domin

Global Leader, Responsible AI Initiatives, IBM
Associate Director,
Notre Dame—IBM Tech Ethics Lab
(ノートルダム IBM テクノロジー倫理ラボ、
アソシエイト・ディレクター、
IBM、責任ある AI イニシアチブ、
グローバル・リーダー)
[linkedin.com/in/heatherdomin/](https://www.linkedin.com/in/heatherdomin/)
hesill@us.ibm.com

協力者

Sara Aboulhosn、Angela Finley、Rachna Handa、
Jungmin Lee、Stephanie Meier、Lucy Sieger
(敬称略)

Marianna Ganapini

Associate Professor, Philosophy
Union College
(ユニオン・カレッジ、
哲学、准教授)
[linkedin.com/in/marianna-b-ganapini-769624116/](https://www.linkedin.com/in/marianna-b-ganapini-769624116/)
marianna@logicanow.com

Brian Goehring

Associate Partner, AI Research Lead
IBM Institute for Business Value
(IBM Institute for Business Value、
アソシエイト・パートナー、AI リサーチ責任者)
[linkedin.com/in/brian-c-goehring-9b5a453/](https://www.linkedin.com/in/brian-c-goehring-9b5a453/)
goehring@us.ibm.com

Francesca Rossi

IBM Fellow and AI Ethics Global Leader
IBM Research
(IBM Research、
IBM フェロー兼 AI 倫理グローバル・リーダー)
[linkedin.com/in/francesca-rossi-34b8b95/](https://www.linkedin.com/in/francesca-rossi-34b8b95/)
Francesca.Rossi2@ibm.com

日本語翻訳監修

山田敦

日本アイ・ビー・エム株式会社
AI 倫理チーム・リーダー
データサイエンティスト職・リーダー
技術理事

監訳者あとがき

昨今、多くの企業が AI 活用の本格化を目指す中、倫理的課題やリスク管理に対するガバナンスをいかに確立すべきか対応に苦慮している。そこで以下、IBM の取り組みをご紹介します。

日本アイ・ビー・エム株式会社（以下、日本 IBM）は、2018 年に設立されたグローバル IBM の AI 倫理委員会の一翼として、日本 IBM AI 倫理チームを 2022 年 6 月に設立した。企業の中に新たな活動を立ち上げるためには当然、定量的な投資効果（ROI）が求められる。つまり本レポートが提唱する「包括的 AI 倫理フレームワーク」の中の「経済面における影響（有形 ROI）」である。AI 倫理に関わるリスクは、情報漏えい、著作権侵害、プライバシー、風評被害など多岐にわたり、またそれらによる想定損失額も大小にわたるため、精緻な試算は難しいが、一定の想定のもと ROI を算出することはできる。IBM は、情報漏えい発生に伴う損失コストに関する詳細なレポート（Cost of a

Data Breach）を毎年発行しており、情報漏えいを代表的なリスクとして、こちらの情報を ROI 試算に活用することができる。日本 IBM AI 倫理チーム設立時の ROI 試算においても、こちらの情報を活用した。皆さんの企業においても、社員が AI を安心して積極的に活用していける環境を作り出すために、ROI を試算し部門横断の AI 倫理委員会を設立されることを強く願う。また AI 倫理リスク審査を運営する中で、「包括的 AI 倫理フレームワーク」の中の「能力向上（リアル・オプション ROI）」の重要性も強く体感する。特に AI 倫理リスク審査は、社員が自分の案件のリスクと対策を真剣に考える機会であるため、社員にとって「自分事」である。つまり、研修などの受動的に学ぶ機会と異なり、能動的に考える機会なのである。これまで数百件にわたる AI 倫理リスク審査の中で、数多くの社員に対して AI のリスクと対策を自ら考える機会を提供してきた。この実績は、「信頼できる AI」をお客様に提供する企業である日本 IBM の、強固な基盤を形成している。

Research Insights について

Research Insights は企業経営者の方々に、各業界の重要課題および業界を超えた課題に関して、事実に基づく戦略的な洞察をご提供するものです。この洞察は、IBV の一次調査研究を分析して得られた結果に基づいています。詳細については、IBM Institute for Business Value (iibv@us.ibm.com) までお問い合わせください。

IBM Institute for Business Value

IBM Institute for Business Value (IBV) は、20 年以上にわたって IBM のソート・リーダーシップ・シンクタンクとしての役割を担い、ビジネス・リーダーの意思決定を支援するため、研究と技術に裏付けられた戦略的洞察を提供しています。

IBV は、ビジネスやテクノロジー、社会が交差する特異な立ち位置にあり、毎年、何千もの経営層、消費者、専門家を対象に調査、インタビューおよび意見交換を行い、そこから信頼性が高く、刺激的で実行可能な知見をまとめています。

IBV が発行するニュースレターは、ibm.com/ibv よりお申し込みいただけます。また、LinkedIn (ibm.co/ibv-linkedin) をフォローいただくと、定期的に情報を入手することができます。

変化する世界に対応するためのパートナー

IBM はお客様と協力して、ビジネス・インサイト、高度な研究成果、およびテクノロジーの専門知識を組み合わせることで、急速に変化し続ける今日の環境における卓越した優位性の確立を可能にします。

関連レポート

The enterprise guide to AI governance
IBM Institute for Business Value. 2024 年 10 月。
邦訳「企業のための AI ガバナンス・ガイド」
<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-governance>

The CEO's guide to generative AI:
Responsible AI & ethics
IBM Institute for Business Value. 2023 年 10 月。
邦訳『人任せにはできない、「倫理への取り組み」』
<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/responsible-ai-ethics-jp>

AI ethics in action
IBM Institute for Business Value. 2022 年 4 月。
邦訳「AI 倫理の実践」
<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-ethics-in-action>

注釈および出典

- 1 Goehring, Brian, Francesca Rossi, and Beth Rudden. AI ethics in action: An enterprise guide to progressing trustworthy AI. IBM Institute for Business Value. 2022 年 4 月。邦訳「AI 倫理の実践」<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-ethics-in-action>
- 2 同上
- 3 同上
- 4 同上
- 5 2023 Institute for Business Value 生成 AI の市場の状況調査。世界の CxO 369 人。2023 年 4 月／5 月。未公開情報。
- 6 2023 IBM Institute for Business Value 次世代デザイン調査。世界の CxO 212 人。2023 年 7 月。未公開情報。
- 7 The 2023 CEO Study. CEO decision-making in the age of AI: Act with intention. IBM Institute for Business Value. 2023 年 6 月。邦訳「CEO に立ちほだかる 6 つの真実」<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/c-suite-study/ceo>
- 8 The CEO's guide to generative AI: Customer and employee experience. IBM Institute for Business Value. 2023 年 8 月。邦訳「エクスペリエンスが全て」<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/employee-customer-experience-jp>
- 9 Bevilacqua, Marialena, Nicholas Berente, Heather Domin, Brian Goehring, and Francesca Rossi. "The Return on Investment in AI Ethics: A Holistic Framework." Proceedings of the 57th Annual HICSS Conference on Systems Sciences. 2024 年 1 月。<https://arxiv.org/abs/2309.13057>
- 10 "Building trust in AI." IBM ケース・スタディー。2023 年。<https://www.ibm.com/case-studies/ibm-cpo-pims>. "How IBM's Chief Privacy Office is building on the company's trustworthy AI framework and scaling automation to address AI regulatory requirements." IBM. 2023 年。<https://www.ibm.com/downloads/cas/7PWAWRQN>. "IBM Commits to Train 2 Million in Artificial Intelligence in Three Years, with a Focus on Underrepresented Communities." IBM Newsroom. 2023 年 9 月 18 日。<https://newsroom.ibm.com/2023-09-18-IBM-Commits-to-Train-2-Million-in-Artificial-Intelligence-in-Three-Years,-with-a-Focus-on-Underrepresented-Communities>
- 11 Bevilacqua, Marialena, Nicholas Berente, Heather Domin, Brian Goehring, and Francesca Rossi. "The Return on Investment in AI Ethics: A Holistic Framework." Proceedings of the 57th Annual HICSS Conference on Systems Sciences. 2024 年 1 月。<https://arxiv.org/abs/2309.13057>. "OECD AI Principles overview." OECD. 2024 年 11 月 15 日閲覧。<https://oecd.ai/en/ai-principles>
- 12 The CEO's guide to generative AI: Customer and employee experience. IBM Institute for Business Value. 2023 年 8 月。邦訳「エクスペリエンスが全て」<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/employee-customer-experience-jp>
- 13 Goehring, Brian, Francesca Rossi, and Beth Rudden. AI ethics in action: An enterprise guide to progressing trustworthy AI. IBM Institute for Business Value. 2022 年 4 月。公開データと非公開データ。邦訳「AI 倫理の実践」<https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-ethics-in-action>
- 14 同上
- 15 Domin, Heather, Francesca Rossi, Brian Goehring, Marianna Ganapini, Nicholas Berente, and Marialena Bevilacqua. "On the ROI of AI Ethics and Governance Investments: From Loss Aversion to Value Generation." California Management Review. 2024 年 7 月 29 日。<https://cmr.berkeley.edu/2024/07/on-the-roi-of-ai-ethics-and-governance-investments-from-loss-aversion-to-value-generation/>
- 16 同上
- 17 Bevilacqua, Marialena, Nicholas Berente, Heather Domin, Brian Goehring, and Francesca Rossi. "The Return on Investment in AI Ethics: A Holistic Framework." Proceedings of the 57th Annual HICSS Conference on Systems Sciences. 2024 年 1 月。<https://arxiv.org/abs/2309.13057>. OECD. (2023). OECD AI Principles overview. The OECD Artificial Intelligence (AI) Principles - OECD.AI. <https://oecd.ai/en/ai-principles>

- 18 Dreksler, Noemi, David McCaffary, Lauren Kahn, Kate Mays, Markus Anderljung, Allan Dafoe, Michael C. Horowitz, and Baobao Zhang. “Preliminary Survey Results: US and European Publics Overwhelmingly and Increasingly Agree That AI Needs to Be Managed Carefully.” Centre for the Governance of AI. 2023 年 4 月 17 日。https://www.governance.ai/post/increasing-consensus-ai-requires-careful-management
- 19 Goehring, Brian, Francesca Rossi, and Beth Rudden. AI ethics in action: An enterprise guide to progressing trustworthy AI. IBM Institute for Business Value. 2022 年 4 月。邦訳「AI 倫理の実践」https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-ethics-in-action
- 20 Domin, Heather, Francesca Rossi, Brian Goehring, Marianna Ganapini, Nicholas Berente, and Marialena Bevilacqua. “On the ROI of AI Ethics and Governance Investments: From Loss Aversion to Value Generation.” California Management Review. 2024 年 7 月 29 日。https://cmr.berkeley.edu/2024/07/on-the-roi-of-ai-ethics-and-governance-investments-from-loss-aversion-to-value-generation/
- 21 同上
- 22 Goehring, Brian, Francesca Rossi, and Beth Rudden. AI ethics in action: An enterprise guide to progressing trustworthy AI. IBM Institute for Business Value. 2022 年 4 月。邦訳「AI 倫理の実践」https://www.ibm.com/thought-leadership/institute-business-value/jp-ja/report/ai-ethics-in-action
- 23 “Staying ahead of the curve—The business case for responsible AI.” The Economist Intelligence Unit. 2020 年 10 月 8 日。https://www.eiu.com/n/staying-ahead-of-the-curve-the-business-case-for-responsible-ai/

© Copyright IBM Corporation 2024

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America | December 2024

IBM、IBM ロゴ、ibm.com は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。現時点での IBM の商標リストについては www.ibm.com/legal/copytrade.shtml (US) をご覧ください。

本書の情報は最初の発行日の時点で得られるものであり、予告なしに変更される場合があります。すべての製品が、IBM が営業を行っているすべての国において利用可能なわけではありません。

本書に掲載されている情報は特定物として現存するままの状態を提供され、第三者の権利の不侵害の保証、商品性の保証、特定目的適合性の保証および法律上の瑕疵担保責任を含むすべての明示もしくは黙示の保証責任なしで提供されています。IBM 製品は、IBM 所定の契約書の条項に基づき保証されます。

本レポートは、一般的なガイダンスの提供のみを目的としており、詳細な調査や専門的な判断の実行の代用とされることを意図したものではありません。IBM は、本書を信頼した結果として組織または個人が被ったいかなる損失についても、一切責任を負わないものとします。

本レポートの中で使用されているデータは、第三者のソースから得られている場合があります。IBM はかかるデータに対する独自の検証、妥当性確認、または監査は行っていません。かかるデータを使用して得られた結果は「そのままの状態」で提供されており、IBM は明示的にも黙示的にも、それを明言したり保証したりするものではありません。

本書は英語版「Why invest in AI ethics and governance? - Five real-world origin stories」の日本語訳として提供されるものです。

