



# AI 倫理の実践

信頼できる AI 実現へ  
全社的取り組みを

## IBM が提供する サービス

IBM は、お客様が人工知能（AI）やアナリティクス、データを活用してさまざまな可能性を広げることができるように全力でサポートします。そのために、業界全般にわたる豊富な知識と経験、企業組織や技術に関する高度な専門性を活かして、エンタープライズ・グレードのテクノロジー・ソリューションを提供し、最新科学に基づいてリサーチのイノベーションを推進します。IBM コンサルティングが提供する AI サービスの詳細はこちらをご覧ください。[ibm.com/services/artificial-intelligence](https://ibm.com/services/artificial-intelligence)

IBM Software が提供する AI ソリューションの詳細はこちら。[ibm.com/Watson](https://ibm.com/Watson)

IBM Research® が提供する AI イノベーションの詳細はこちら。[research.ibm.com/artificial-intelligence](https://research.ibm.com/artificial-intelligence)



## 主なポイント

AI 倫理への  
戦略的取り組みが  
企業差別化の鍵に

### ■ AI 倫理、主役は技術系リーダーからビジネスリーダーに

AI システムの倫理に取り組むという重要課題で、非技術系エグゼクティブが陣頭指揮に立つ企業が急増している。その割合は 2018 年では 15% だったが、わずか 3 年後には 80% に達している。最高経営責任者（CEO）自らが対応に乗り出すとした割合も、同じ期間に 20% から 79% に増加している。

### ■ パーパス（存在意義）に即した AI 実現の動きは確実に進捗

AI 倫理を自社の企業倫理に組み込む企業は半数を超えており、その多くが AI に特化した倫理体系の確立に取り組んでいる。

### ■ 企業の意欲と実際の行動には大きな隔たり

ただ、企業の取り組みは依然道半ばである。AI が導き出す結果に差別や偏見などのバイアス（偏り）が発生する問題では、企業の 68% が、バイアス軽減に社内のダイバーシティー&インクルージョン（多様性と一体性）の向上が有効だと認識している。しかし、肝心の AI 担当チームではダイバーシティーがかなり遅れている。社内全体の社員構成比と比較すると、女性比率では全社が 33% であるのに対し AI 担当チームは 6%、LGBTQ+ の比率は 4% に対し 1%、黒人・先住民・有色人種の比率は 10% に対し 6% にとどまっている。

# AI の信頼性確保が急務に

企業が直面している社会・環境・政治のさまざまな問題が日々、絶え間なく報じられている。

こうした問題に対し、消費者や労働者、さらには株主からも、企業は原則的立場を表明するだけでなく、社会的意義のある行動によって目に見える社会貢献をすべきだとする声が高まっている（後述の視点「利害関係者の声に応えバリュー実践を」を参照）。

この動きは経営判断にも影響を及ぼすようになり、通常の企業経営の枠を超えてすそ野が広がっている。実際、経営層の60%近くは、こうした社会・経済のうねりが今後、さまざまな「企業の社会との関わり方」に幅広く影響していこうと考えている。<sup>1</sup>

こうした動きは、AIやデータの領域にも広がっており、IBM Institute for Business Value (IBV) が2018年に行った調査でも確認された。<sup>2</sup>

一般的に考えられている「信頼できないAI」の定義ははっきりしている。「導き出された結果が差別的、不透明だ」「不適切な使われ方をしている」など、世の中が期待する信頼性に欠けていることである。しかし、「信頼できるAI (trustworthy AI)」の開発には依然、困難が付きまとう。実用性の観点からさまざまなバランスを考慮する必要が出てくるためである。例えば、AIのアルゴリズム計算の背後にある論理を理解する「説明可能性」と、結果を導くアルゴリズムの正確性を示す「堅牢性」をどう両立させるか、といった問題である。

AIの導入にあたって、こうしたトレードオフなども含め倫理的な問題は、もはや無視することはできなくなった。ここで企業に問われるのは、パーパス（存在意義）に沿って戦略的に熟慮を重ねた上で、こうした問題に向き合うのか、そうしないのかを選択することだけである。

このような動きは広義のテクノロジー・セクターですでに始まっている。例えば、高成長・高収益のプラットフォーム・ビジネスに携わるデジタル・ネイティブの企業などは、その潤沢な資金のおかげで消費者志向のビジネスモデルに潜む倫理的費用を賄うことができる。結果的に、社内に大規模なチームを編成し、多層的なプロセスを構築することでさまざまな要求に応えようとしている。だが、一律に成果を生んでいるわけではない。

## 利害関係者の声に応え バリュー実践を

株主価値を高めるという、企業へ付託されてきた伝統的使命は重みを失いつつある。人々は「消費者」「市民」「労働者」「投資家」といった立場からさまざまな声を広範囲に上げるようになり、メディアがこうした要求をさらに増幅する。こうした利害関係者の声が業績にも影響を及ぼすようになり、企業はこれまで以上に注意を払わざるを得なくなっている。

この動きはサステナビリティに関する IBV 調査でも明らかである。2 年前、パーパス（存在意義）を掲げる企業ブランドに 100% 以上の割増料金を払っても構わないと考える消費者の割合はわずか 8% だったが、<sup>3</sup> 直近の調査では 43% に上昇した。<sup>4</sup> サステナブルで環境に配慮した企業ブランドを支援したいと考える消費者が増えている。

勤務先としても「環境に配慮し社会的責任を果たしている企業で働きたい」と回答した労働者の割合は 70% 近くに上る。こうした傾向は、社員の定着率にも影響を与えている。<sup>5</sup>

さらに顧客の 68%、社員の 62% は、自らが声を上げることによって企業を動かすことができると考えている。<sup>6</sup>

個人投資家も、サステナビリティを投資の判断基準として資産運用に活用している。企業が抱える気候変動リスクが金融リスクに直結すると考える個人投資家は半数程度に上る。さらに、このうち 92% が向こう 1 年以内に行う投資判断やファンド資産構成の変更について、環境要因や社会的責任を投資基準として吟味すると述べている。<sup>7</sup>

このように道徳や倫理の問題を重視する利害関係者が増える中、企業側でも旧来的な株主資本主義を見直す動きが加速している。<sup>8</sup>

だが、これは形だけトレンドに乗ればよいという簡単な話ではない。今の流れを現実の収益に結実させている企業は、サステナビリティを単なる一時的な流行ではなく、企業変革を進める原動力だと真剣に捉えている。そうでなければ、社会的意義のあるプラクティスを根付かせるために社内を変革するという困難な作業などではしない。<sup>9</sup>

企業理念に掲げた価値観（バリュー）は、正しく実践されることで、単なるお題目にならず、現実の企業収益として形になって表れるはずである。

これはテクノロジー・セクターに限った話ではない。今日の不透明な環境に対応しなければならないのは他の業種も同じである。しかも、巨大プラットフォーム企業と比べ経営資源に制約がある企業は、部門を限定させて AI を使いつつ利益を伴う成長を目指さなければならない。AI の倫理的課題に投入する専任リソースが不足していることが、この困難な状況に拍車をかけている。

企業は厳しい選択を迫られている。

消費者や労働者、投資家からの要求の嵐にさらされる中、企業は当局の規制の有無を盾にして、嵐が収まるまでやり過ごすこともできるだろう。しかし、国や地域、さらには州レベルにおいても、AI 関連の法的枠組みは常に見直され、内容が変わっていく。そうした先の見えない（かつ重複的な）規制環境の下で、戦略として規制順守だけを頼みの綱とすることは、危険な賭けになりかねない。そうやって手をこまねいているばかりではリスクが増す一方である。

むしろ、AI の倫理的課題に正面から戦略的に取り組むことによって他社と差別化できる道筋が見えてくる。実践から学びつつ、具体的な行動指針を随時修正することも欠かせない。

IBV は AI 倫理に対する企業の取り組み状況を調べるため、経済分析・予測を手掛けるオックスフォード・エコノミクス社（Oxford Economics）と共同で 2021 年に調査を行った。対象は 22 カ国の経営層 1,200 人で、事業および技術部門の 16 役職（29 ページの「調査方法」を参照）。

社会的意義を考慮した  
AI 倫理を追究することで  
はじめて信頼される AI は  
実現できる。

## 企業の意欲が空回り、 実践が追い付かず

そもそも AI 倫理はどう定義されるべきだろうか。

AI 倫理は、AI がもたらす便益を最適化することを目指す学際的な研究分野であると一般に考えられている。そのために人間を主体に据え、そのウェルビーイングの実現を目指し、すべての利害関係者にとって不利益となるリスクの軽減を図る。<sup>10</sup> 社会的意義を考慮した AI 倫理を追究することではじめて信頼される AI は実現できる。

同様の例は他の学問分野にもある。医学は患者の健康改善を目指しているが、そのためには治療が安全・有効であり、信頼できるものでなければならない。工学は物理的インフラの設計・建設を専門とする学問だが、橋やトンネル、建物の安全・安心を築いている。

IBV 調査では、消費者や市民、労働者はおしなべて、人類に差し迫った課題を解決するために AI が最も有効なテクノロジーであると考えている。<sup>11</sup> このことから、AI 倫理の確立は断固として進めるべきことだと分かるだろう。さらに、このうち 85% 超の回答者が、企業に対して社会問題と同様に AI 倫理にも取り組むべきだと答えている。<sup>12</sup>

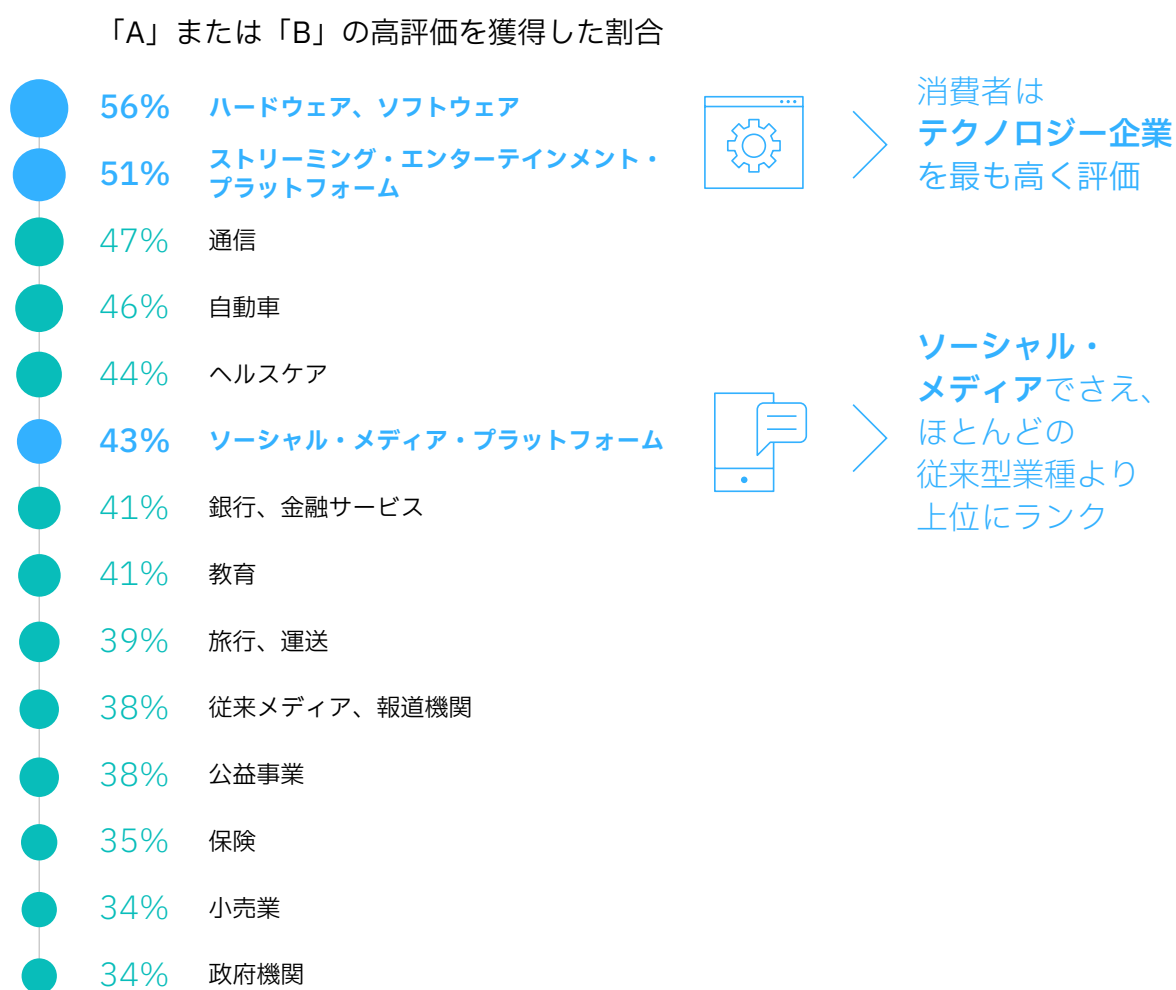
一方、テクノロジーの利用にあたって企業が強い責任感をもって行動しているかという点では、銀行・小売業・保険といった従来業種の評価が芳しくなかった。マスコミで評判の悪いソーシャル・メディア・プラットフォーム企業の評価すら、ほとんどの従来業種が下回った（図 1 参照）。<sup>13</sup> さらに、企業が AI などの新規テクノロジーの開発・実行にあたり、責任感と倫理感をもって行動していると回答した消費者は 40% にすぎず、2018 年からほとんど変化はない。<sup>14</sup>

調査では、消費者、市民、労働者の85%超が、企業はAI倫理に取り組むべきだと考えている。

図1

### AI倫理に関する業種別評価

テクノロジー利用にあたって責任ある対応をとっているかという点で従来業種はあまり評価されていない



Q：AIなどのテクノロジー利用にあたって責任ある対応がなされているか、以下の業種の企業・組織を評価してください。

出典：IBM Institute for Business Valueが14,526人の成人を対象に行ったHuman Insights Global Survey。2021年7月実施。これまで未公開のデータ。



しかし、企業はAI倫理を確実に重要視するようになってきている。2018年の調査では、自社が3年前よりAI倫理を重視していると答えた経営層は全体の半数に届かなかったが<sup>15</sup>、2021年には4分の3超に増加した。具体的には、「AI倫理を社内に定着させるため予算を倍増して関連現場を強化した」「研修の実施やチームづくりを進めた」「社内プロセスの整備や各種ツールの導入を図った」などである。今後3年間についても、AI倫理の関連投資は増加するだろうと予想している。

このように経営層がAI倫理への取り組みに前向きであるにもかかわらず、顧客や社員の評価は伴っていない。この現象はどう理解すればよいのだろうか。

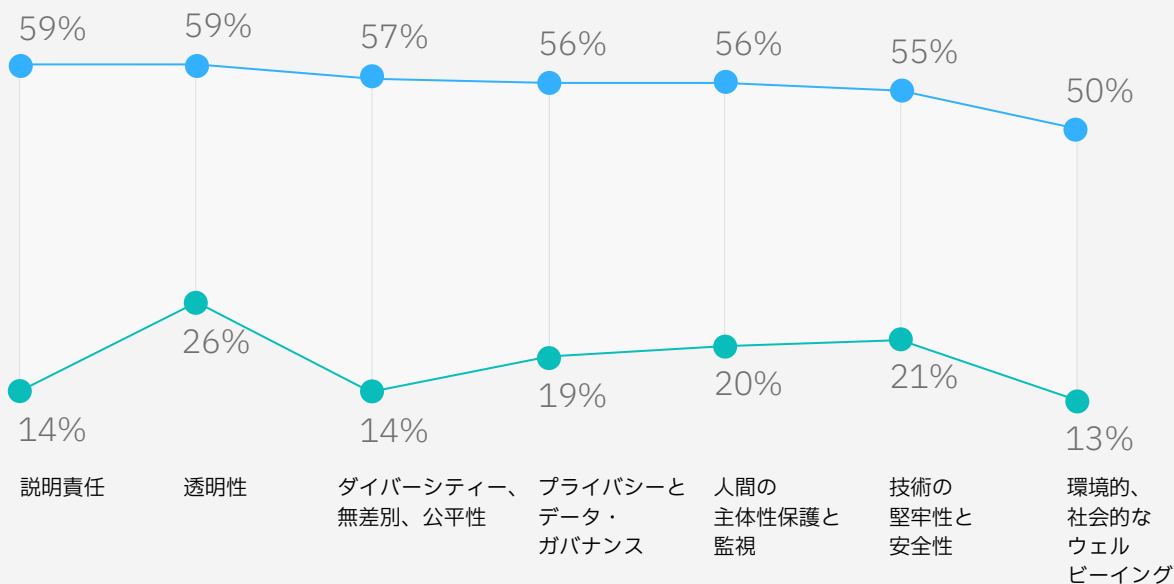
市場の期待に押されて、企業はAI倫理への取り組み意欲を高めているものの、実践能力が追い付かず、利害関係者を納得させるに至っていない。AI倫理の原則について公に賛同している企業は半数を超えるが、実務に取り込むことができているのは4分の1にも満たない(図2参照)。AI倫理に関する自社の原則や価値観(バリュー)と、実際の行動が一致している(もしくは、より望ましい行動がとれている)と確信する企業は20%未満だった。この結果は、世界経済フォーラム(WEF)が「意欲と行動」のギャップと呼ぶ事象を数字的にも裏付けている。<sup>16</sup>

AI倫理に関する自社の行動が、公に表明した原則や価値観と一致している(もしくはより望ましい)と確信する経営陣は20%未満

図2

### 意欲と行動のギャップ

企業はAI倫理原則に賛同しているが、自社の実践は遅れている



### 賛同 | 実践

注記：欧州委員会(EC)のAIハイレベル専門家グループが、「Ethics guidelines for trustworthy AI(2019年4月)」で定義したAI倫理原則：<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>



こうした実践の遅れを解消することが重要である。例えば AI のバイアスを緩和し、より信頼を高めるためには、ダイバーシティーや無差別、公平性といった倫理原則に真剣に取り組まなくてはならない。企業はこの重要性を認識しており、2021 年の調査では AI に関する活動でこの点の重要性を指摘する回答が 2018 年を大幅に上回った。直近の調査では、その重要性を否定するリーダーはほぼ皆無となり、「極めて重要」とした経営層の割合は 6 倍程度に増えている。

この問題に対処するには、より幅広い層の意見を反映できるように AI 倫理に取り組むメンバーを選ばなくてはならない。ところが AI チームの構成員は、組織全体に比べてダイバーシティーが著しく劣っている。IBV 調査では、女性比率は組織全体が 33% であるのに対し AI チームは 6%、LGBTQ+ は 4% に対し 1%、黒人・先住民・有色人種 (BIPOC) は 10% に対し 6% にとどまっている (図 3 参照)。

こうした違いはさまざまな原因に根差している。しかし、人財の募集・採用・維持の改善にもっと積極的に目に見える形で注力することが、AI を設計・実行する現場でマイノリティーの人々を増やすために必要なのは明らかである。こうした取り組みは AI バイアスを減らし、AI から導かれた結果への信頼を必ず高めるはずである。

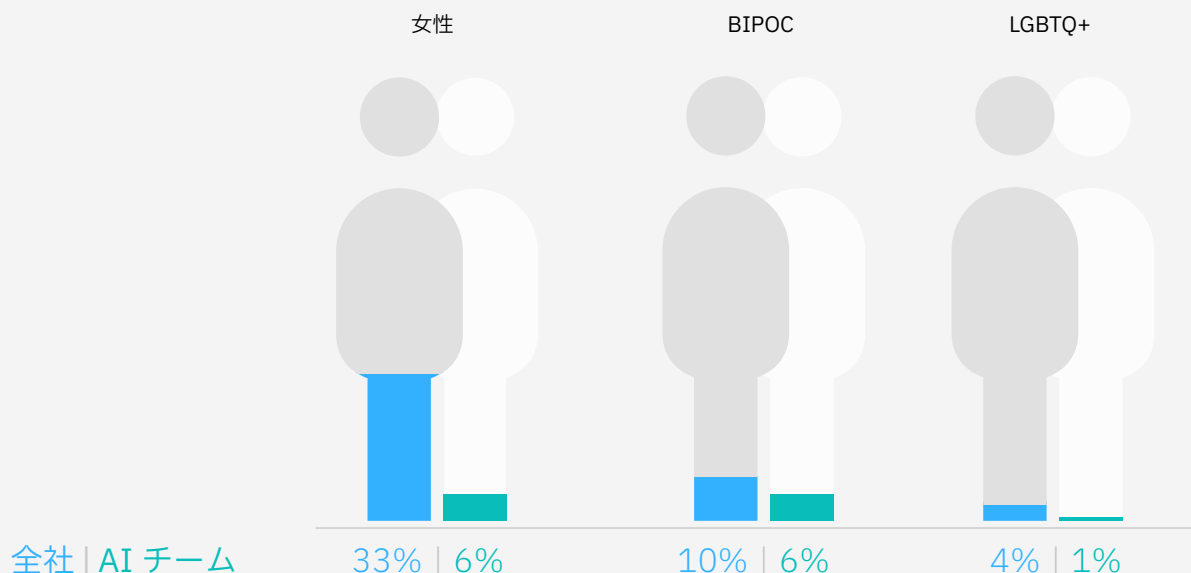
多様性の向上は業績向上にも有効であることが、IBV 調査で明らかになっている。例えば全社と AI チームの女性比率の格差が小さい上位 4 分 1 の企業グループは、AI プロジェクトの投資収益率が他の企業をやや上回っていた。

他の AI 倫理原則についても、実践の遅れを解消するには、ダイバーシティーと同様に具体的な目標を定め、測定可能な結果を求める姿勢が必要である (後述のケース・スタディー「リージョンズ・バンク (Regions Bank) : 高品質で信頼できる AI に注力」を参照)。

図 3

### ダイバーシティーの組織内格差

AI チームは組織全体に比べてダイバーシティーが著しく低い



意欲と行動のずれは、環境サステナビリティに対する企業の取り組み姿勢にも見られる。最近のIBV調査では、この課題に関して戦略を実践している企業は35%にすぎず、目標設定を事業戦略に沿って行った企業は37%にとどまった。<sup>17</sup> 目標と評価基準をビジネス・プロセスに導入した組織もわずか3分の1にすぎなかった。<sup>18</sup>

企業や経営層がAI倫理をはじめとする原則的な優先課題の実践に注力すれば、意欲と行動のギャップを解消できる可能性は格段と高まる。

## 経営層 「AI倫理には社内20部門が さまざまに関与しており、 シナジーが不可欠」

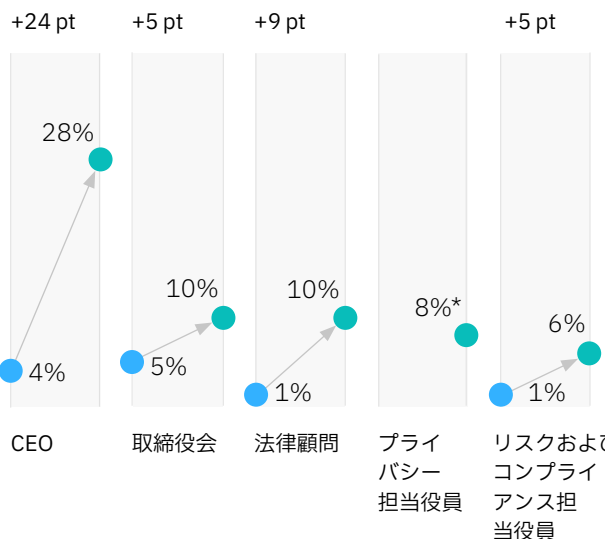
図4

### 「推進役」の交代

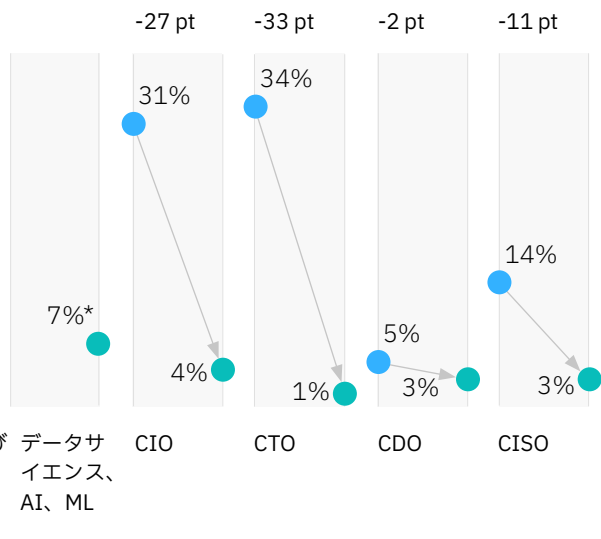
2018年から2021年にかけて、AI倫理の主要な責任者は技術部門のリーダーから非技術部門のリーダーへと変わった

2018年 | 2021年

#### 非技術部門のリーダー



#### 技術部門のリーダー



Q：AI倫理の主要な責任者は誰ですか。

2018年調査データ、出典：Goehring, Brian, Francesca Rossi, and Dave Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020.

\*2018年のデータには職位は含まれていない。

## 潮流の変化： ビジネス・エグゼクティブ主導の 社内コラボレーションが生命線

企業がAI倫理の実践へ力を入れ始める中、組織内ではこの2、3年大きな変化が起こっている。それまであまり関与してこなかった非技術系リーダーが一気に“主役”に躍り出たのである。2018年の調査では、AI倫理の「推進役」として非技術系リーダーを挙げた回答はわずか15%だったが、<sup>19</sup> 2021年には80%へ急増した。

企業が「推進役」として特に期待するのはCEO(28%)だが、ほかに取締役会メンバー(10%)や法律顧問(10%)、プライバシー担当役員(8%)、リスクおよびコンプライアンス担当役員(6%)も上位に挙げた。この結果は、業種や地域によってもほとんど変わらない(図4参照)。CEOにとっては、AI倫理の専任役員を任命し責任を分担させることも選択肢になるかもしれない。<sup>20</sup>

## リージョンズ・バンク (Regions Bank) : 高品質で信頼できる AI に注力<sup>21</sup>

再現性があり、サステナブルな AI を企業はどう構築・運用すればよいのだろうか。また、最も重要な課題であろう AI の信頼性向上にはどういった取り組みが必要なのか。

リージョンズ・バンクは、こうした課題にまさに直面していた。同行の高度なアナリティクス業務は、部門ごとにサイロ化したデータを使用し、開発チームも分断されていた。開発手法も共通性がなく、一貫性に欠けていた。

同行はまず、アナリティクスを担当する組織を刷新した。アナリティクスの中核を担うセンター・オブ・エクセレンスを設置して、データを一元管理し、機械学習 (ML) や AI 技術の活用を強化した。とりわけ、エンド・ツー・エンドのビジネス・バリュー・アプローチを採用し、AI の品質管理を組み込んだ。

この新たな組織基盤によって、リージョンズ・バンクは統合された組織力をうまく活用することが可能となり、公平で倫理的、かつ信頼性の高い AI モデルを実現した。

同行の最高データ・アナリティクス責任者 (Chief Data and Analytics Officer) の Manav Misra 氏は、「リージョンズ・バンクはお客様との信頼関係に基づくオープンな交流を誇りに思っている」と話す。さらに、同行の目標はコミュニティやお客様、行員およびビジネス・パートナーの生活を向上させることとした上で、「AI 倫理に基づく監視を開発手法に組み込んでいるのはそのためだ」と説明する。

さらに、「こうしたことはすべて、私たちが『責任ある AI』と呼んでいるものにつながっている」と指摘する。このアプローチでは、AI モデルの基礎となるデータは、意思決定に使用されるデータを代表するものでなくてはならないとされる。これに加え、AI モデルは説明可能なものでなければならない。そうすることで意思決定プロセスが明確になる。

社内で編成された監視チームが、リスク・マネージャーや監査人、規制当局と同じような形で監視を行い、AI が提示するソリューションの公平性・安全性・健全性の維持を図っている。

ところで、AI 倫理のプロセスやチームの構築には、いつ取り組むべきなのだろうか。

Misra 氏によると、着手は早ければ早いほど良い。「監視に関わる人たちを早期に組み入れ、各プロジェクトの趣旨や要件に加え、開発コードやアジャイル開発全体にわたる意図を理解してもらうことで、より迅速なフィードバックが可能になる。それにより、リージョンズ・バンクが目指す高品質で信頼できる AI ソリューションの開発が加速される」

こうした結果、より信頼性の高い AI や機械学習をはじめとする、さまざまなアナリティクスのソリューションが実現されている。こうしたソリューションが 1 つになって、リスクの軽減や不正行為の検出を後押しし、顧客ニーズを深掘りして利便性や満足度の向上をもたらす。このすべてがビジネス価値の創造につながるのである。

一方で、AI 倫理の取り組みでは特定部門にとどまらない、全社的なアプローチが不可欠だと企業は認識している。AI 倫理に何らかの形で関与する部門が少なくとも 20 以上あるとする企業がほとんどだからである。

最高情報責任者 (CIO) や最高技術責任者 (CTO)、最高データ責任者 (CDO)、およびその所管部門が、AI 倫理の実現に不可欠なのは自明だが、AI 倫理に関わるのは技術部門だけではない。例えば、調達や製品設計、研究、公共政策、法務などの部門も重要となる。

スタンフォード大学の「人間中心の AI 研究所」(HAI、Institute for Human-Centered Artificial Intelligence) でアソシエート・ディレクターを務める Rob Reich 氏はこの点について、「企業が AI 倫理を現実機能させるためには、社内の緊密な協力が必要になる」と指摘する。さらに、「個人が倫理的な判断を他人に委ねられないのと同じで、企業にしても、最高倫理責任者を任命して AI 倫理の実現を一任するだけで、他の事業部門を積極的に関与させないままでは、AI 倫理の実現などおぼつかない」と説明。「倫理というのは基本として、全員で責任を負わなくてはならないものである」<sup>22</sup> と強調する。

ある世界的な製薬会社の例が参考になる。社内に散らばる人事関連データを AI に分析させることで人財開発のニーズを把握・予測するスキル推定モデルの開発を行った。当初は、データサイエンス・チーム内で小規模に概念実証モデルから開始し、良好な結果が得られたことから、全社規模に拡大させることを決定した。しかし、大規模な AI 化の実現にはデータサイエンティストにとどまらず、幅広い部門の関与が必要であることがすぐに明らかとなった。

## 「倫理というのは基本として、全員で責任を負わなくてはならないものである」

スタンフォード大学の「人間中心の AI 研究所」(HAI、Institute for Human-Centered Artificial Intelligence) アソシエート・ディレクター **Rob Reich 氏**

具体的には、ソフトウェア・エンジニアやプロジェクト・マネージャーをはじめとする IT スペシャリストはもちろん、人事担当者や法務チーム、心理学専門家など人財マネジメントに関わるプロフェッショナルなどの協力が必要となった。さらに、このプロジェクトの場合、データ・リネージュ (データ履歴の可視化) やプロヴェナンス (来歴) は単なるメタデータ (データに関する付帯情報) のオプション属性ではない。AI から導かれた結果に対する信頼性を高め、社員による受容とエンゲージメントを促すためにならなければならないものだった。

こうしてプロジェクトが大規模化すると、一般的には、複数部門が関わる中で細心の注意が求められ、進捗に合わせたフィードバックに基づき合意形成を促すことも必要となる。このため、どういった協力を、どのチーム間で、いつ行うかについて適切に判断するトップダウンのリーダーシップが極めて重要になる。

トップ経営者自身も 2018 年に比べ、積極姿勢に転換しているようである。AI 倫理の取り組みに対して傘下チームや社外専門家の主体的な参加という後押しがあり、さまざまな会議の場でも取り上げられていることが背景にあると思われる。2021 年の調査で、AI 倫理に取り組む用意ができていると答えた CEO は 79% を占め、2018 年の 20%<sup>23</sup> から 4 倍増になった。取締役会メンバーも同様で、同じ期間に 45%<sup>24</sup> から 71% へ大きく増加した。

最高人事責任者 (CHRO) による人財関連の取り組みも進んでいる。2018 年の IBV 調査では AI に関連する人財面の課題が明らかになった。例えば、AI 導入に伴って職務への影響が避けられない人財の再教育を積極的に計画する企業はわずか 37% だった。<sup>25</sup> しかし、2021 年にはその割合は 55% へ増加した。また、AI 関連業務に携わる社員向け研修を用意している企業の割合も、2018 年の 41% から 70% に増加した。<sup>26</sup> 業務上 AI に直接関わる人財は今後 2、3 年で 20% 超増加すると CHRO はみており、こうした人財の育成に積極的に取り組むことが、企業・政府・社会にとってますます重要になっている。

外部から AI ガバナンスの厳格化を求める声が強まっていることも、AI 倫理の重要性を高めている要因だろう。AI に対する規制が所在地域で近々導入される見込みだとする企業が 74% に上る一方、64% はすでに対応する準備ができていると回答している。一方で、規制環境がどう変化しても、主要な役割を果たすのは第三者専門機関など規制当局以外の組織になっていくと回答者はみており、こうした社会全体にわたるエコシステムの影響力と関与が極めて重要になると 48% が答えている（ケース・スタディー「カナダの銀行コンソーシアム：信頼のおけるデータと AI を共同構築」を参照）。

全般的に、組織の取り組みは正しい方向に向かっているとされる。これにより、企業だけでなく、政府や教育機関も AI 倫理原則の実践に向けて、さらに踏み込んだ行動をとる流れが出てきている。

AI 導入に伴って職務への影響が避けられない人財の再教育を積極的に計画している組織は、2018 年の 37% から 2021 年の 55% に大きく増加した。



## カナダの銀行 コンソーシアム： 信頼のおける データと AI を共同構築<sup>27</sup>

金融機関は、顧客や社内、サードパーティーなどのデータを使用して、従来の銀行業務に代わる革新的なサービスを提供するようになった。その結果、銀行と顧客の関係はより深まり、インタラクティブになりつつある。

カナダ・ナショナル銀行 (National Bank of Canada) の統合リスク管理担当バイス・プレジデントである Mathieu Avon 氏は、「AI によって、顧客体験や従業員体験をポジティブに変化させることが可能であり、中核となる組織力を強化できる」と指摘する。その上で、「お客様をはじめとする利害関係者の信頼を維持するために、私たちは責任感と倫理感を持って AI 技術の展開に中心的な役割を果たさなければならない」と強調する。

AI の運用は倫理原則に沿って行うべきだと考える金融機関は多い。しかし、そのすべてが信頼のおけるデータや AI を独自に構築・実行・展開するためのリソースを十分に備えているわけではない。

この課題に対処するため、世界最大規模の学術研究団体であり、技術標準化機構である米国電気電子学会 (IEEE、アイ・トリプル・イー) は 2019 年、金融のエコシステムに関わる組織・機関を招集し、カナダの金融サービスのための AI に関する倫理基準ガイドを作成した。

「われわれはデータに関する倫理の順守を、経営上のコストとは考えていない。それは優れたビジネス上の投資だ」。カナダ・インペリアル・バンク・オブ・コマーシ (CIBC) で、かつてエンタープライズ・データ担当シニア・バイス・プレジデント兼最高情報責任者を務めた Terry Hickey 氏はこう指摘する。

この作業にあたったチームは実業界や学会、非政府系組織、業界団体からベスト・プラクティスを収集し、ロードマップを作成した。この取り組みには、カナダの大手 6 行のほか、信用組合、年金基金、フィンテックが参加している。

カナダロイヤル銀行でデータ・ユース&プロダクト・マネージメントおよび、データ&アナリティクスの両部門で責任者を務める William Stewart 氏は、AI 倫理を確立するにあたって企業原則との関連性が重要だと強調する。「企業は AI 特有のリスクを綿密に検討し、そのリスクに対応する際には企業の価値観 (バリュー) に整合し、顧客・社員・社会の期待に応えるべく努めたことを示さなくてはならない」

このエコシステムはその後も協力が続き、その成果として信頼性の高い機械学習運用のプロセスとツールのためのプレイブックが共同作成された。こうした継続的な活動で得られた成果を共有することで、顧客やパートナー、従業員に対し、プライバシーの保護や公正なデータ利用、バイアスの解消のために努力が重ねられていることを周知し、安心感につなげることを目指している。



## 今こそ行動を - 信頼できる AI が ビジネス価値を高める

企業原則の実践がもたらすビジネス上のメリットは、単にマイナスのリスクを低減し、利害関係者の期待に応えることにとどまらない。持続可能なプラクティスを実現し、情報の質を守ることが可能となる。さらに、短期的な利益にとらわれず長期的な視点に立てる。その結果、目に見える形で消費者の信頼を高めることができるようになるのである。<sup>28</sup>

IBV 調査では、経営層の 75% は倫理を競合他社との差別化の源泉とみている。AI 分野などのビジネス・パートナーやベンダーの選定基準に倫理を加える経営層も同程度いた。半数以上の経営層は、エシカル（倫理的）な企業に進んでプレミアム価格を支払うと述べている。

こうした点に関して同業他社と比較した企業の位置付けは、主に (1) AI の重要性と、(2) AI 倫理の重要性の 2 つをどう捉えているかによって定まるようである。

事業戦略において AI を重視している企業は、AI 施策の自己評価で他社を 1.5 倍上回る事業効果が示された。また、AI を重視していない企業に比べて、AI プロジェクトの投資利益率（ROI）がほぼ 2 倍だった。AI を戦略的にどの程度重視しているかが、企業全体の収益向上や採算性にそのまま表れている。

さらに AI 倫理を他社よりも重視している企業は、顧客や社員の信頼度も高いことが報告されている。

企業は戦略的な意思を明確に持ち実践していくことによって、AI からより大きなビジネス価値を引き出せるのではないだろうか。

では、企業は実際にはどう行動しているのだろうか。

経営層の 75% は  
倫理を競合他社との差別化の  
源泉とみている。



事業戦略において  
AI を重視している企業は、  
そうでない企業に比べ、  
AI プロジェクトの ROI が  
ほぼ 2 倍だった。

図 5 では、現在は AI および AI 倫理を重視していない組織も、3 年後には図の右上に移行することが示されている。このセグメントに入る企業は最も積極的に AI を導入し、先進事例を積み上げるなど、効果的に AI を活用している。経営層の調査回答から、このセグメントに位置する組織の数は今後 3 年間で倍増することが予想される。

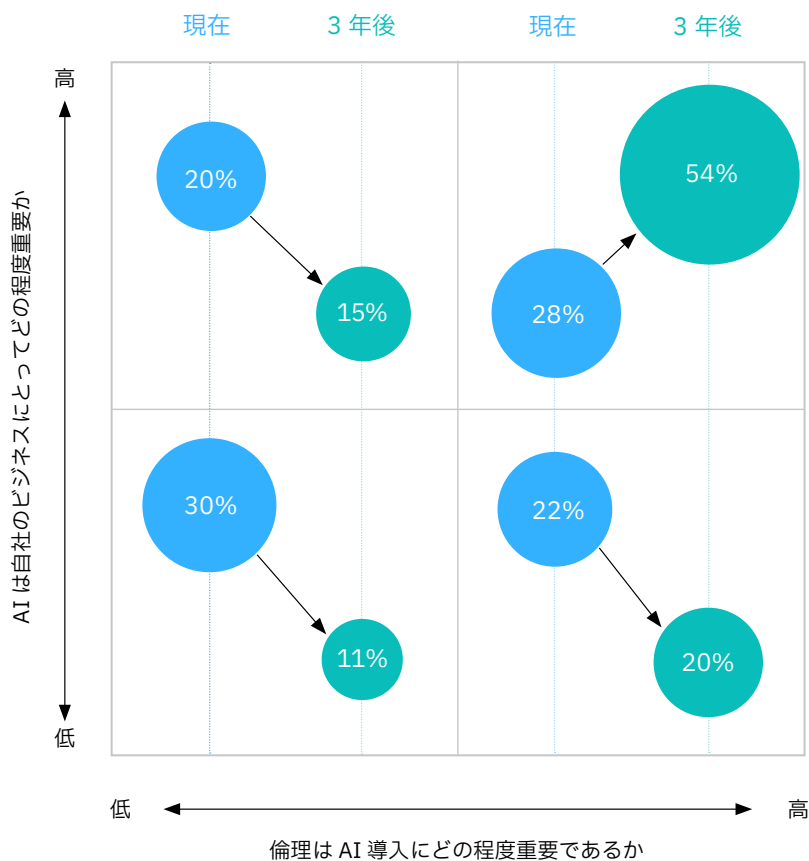
左上のセグメントに入る企業は AI を重要だと考えているが、AI 倫理は重視していない。しかし、本レポートで前述したように、AI 倫理の強化を求める声が消費者や市民、投資家、さらに労働者から強まっており、そうした姿勢のままでは難しい。AI 倫理にいつまでも無関心であることは許されないのである。ソーシャル・メディアの世界で若者のちょっとした過ちがいつまでも記録されていくように、企業の行動も現在から将来へ記録され続け、パートナーや未来の社員の評価にさらされることになる。企業は、社会と協調的に行動する（そして組織力強化へ投資する）ことによって、存続する可能性がより高まる。

業績向上が、AI の戦略的・効果的な活用と相関性があることは、過去の調査で幾度も指摘され、今回も改めて確認された。<sup>29</sup> この関係は、配慮や信念をもって AI をエシカル（倫理的）に導入することにも当てはまるようである。だからこそ、AI と AI 倫理を事業戦略の中核と考える企業は、実践する姿勢を組織全体に深く根付かせようと懸命なのだ（16 ページの視点「現場からのリーディング・プラクティス」を参照）。

図 5

### 成長に向けた企業の位置付け

AI と AI 倫理の価値を認める企業の数は、今後 3 年間で大幅に増加すると予想される



大多数の企業は今後 3 年間で、AI と AI 倫理がともに戦略的に重要になると予想している

## 現場からのリーディング・プラクティス： 「テクノロジーの責任ある活用」プロジェクト

世界経済フォーラム（WEF）はサンタクララ大学のマーククラ応用倫理学センター（Markkula Center for Applied Ethics）と共同で、「テクノロジーの責任ある活用」のケース・スタディー・イニシアチブを公表した。このプロジェクトには、企業が業界他社のインサイトから学び、テクノロジーの設計・開発および活用に倫理を取り入れる際、参考にできる内容が盛り込まれている。

最初に発表された2件のケース・スタディーはマイクロソフトとIBMを取り上げ、両社が倫理の価値観と文化の醸成をいかに発展的に行ったかを考察している。マイクロソフトのケース・スタディーでは、プロダクト・エンジニアリング組織が採用するツールとプラクティスを取り上げている。<sup>30</sup> IBMに関するホワイト・ペーパーでは、同社が組織をどのように変革して、テクノロジーに対する倫理を深く掘り下げ、責任あるテクノロジーの設計・開発・展開および利用を行ってきたかについて検証している。<sup>31</sup>

AI倫理については、マイクロソフトは自社の研究開発部門に特に重点を置き、IBMは広範な組織全体を対象としている。両社ともAIに関連して以下のことを実行している。

- 自社や外部の多様な利害関係組織の価値観（バリュー）と整合性のある原則を支持する
- 公開されているプラクティスや基準、デザイン思考のテクニック、実務的なツールキットなどを使って運用する
- 目標設定と達成の進捗によって評価する
- 社内にアンバサダーを設けて研修を実施するほか、文化に重点を置いた組織的な取り組みを行う
- 社外組織と連携し、業界や政府、学会の枠を超えたイニシアチブに参加することで外部との関係を強化する
- 社会への影響を意識し、当該分野の研究を継続的に行う

マイクロソフトのAIガバナンス・アプローチはハブ・アンド・スポーク方式\*を採用しており、同社がプライバシーやセキュリティ、アクセシビリティの機能を製品やサービスに取り入れるのに役立っている。「ハブ」は3つのチームで構成されている。1つ目は科学とエンジニアリングの専門家チームで構成され、2つ目はポリシーやガバナンス、イネーブルメント（成果を出せるようにする仕掛けづくり）、「機微な情報の利用」に焦点を当てたチームである。最後のチームは、各エンジニアリング・チーム内で「責任ある利用プロセス」の実現を目的としている。一方、「スポーク」はデザイン思考チームや組織で文化的な変化を促進する「推進者＝ビジネスリーダー」など、ビジネスのエキスパートで構成されている。

\* 「責任あるAI」を実現するためのガバナンスのアプローチで、「ハブ」グループが活動の核となって専門知識の集約やプロセス・戦略の策定などを担い、「スポーク」役の担当者が社内各部門で取り組みを推進するモデル。

IBM の AI 倫理委員会 (Ethics Board) は、AI システムの倫理的な開発と展開における要としての機能を果たしている。同委員会はビジネスユニットや部門のリーダーたちで構成され、意思決定とその実行・実現、指揮の権限を有する。また、プロジェクト・オフィスとアドボカシー・ネットワークが設置されており、各地域の中核的な拠点を通じて、協力して組織全体の改革に取り組んでいる。

この倫理委員会によるガバナンスの取り組みには、IBM が開発した AI とそれを支援するソフトウェアの活用、あるいはビジネス・パートナーとの関係の構築、サービス組織が支援するクライアントの AI 実装、社内活動 (採用、報酬、教育、ダイバーシティー&インクルージョンの取り組み) などが含まれる。

両社とも、自社の価値観 (バリュー) と結果の間に不一致を発見した場合、迅速に対応する体制を整えている。また動的フィードバック・ループと継続的改善の精神を掲げ、継続的にアプローチを進化させるよう努めている。

テクノロジー大手によるこうしたプロジェクトに対して、他業界からは組織や仕組みなどの点で自社のニーズと異なるという声が聞こえてきそうである。そうしたケースもあるだろうが、先進事例や教訓などといった成功につながる要因は参考になるだろうし、規模に差があっても、その堅牢性はそのまま維持されるはずである。

WEF が主張するように、「今や、あらゆる企業がテクノロジー企業だと言える。だからこそすべての企業はテクノロジーの倫理がどのように自社の業務に関わっているか、もっと真剣に考えるべきである」。<sup>32</sup>

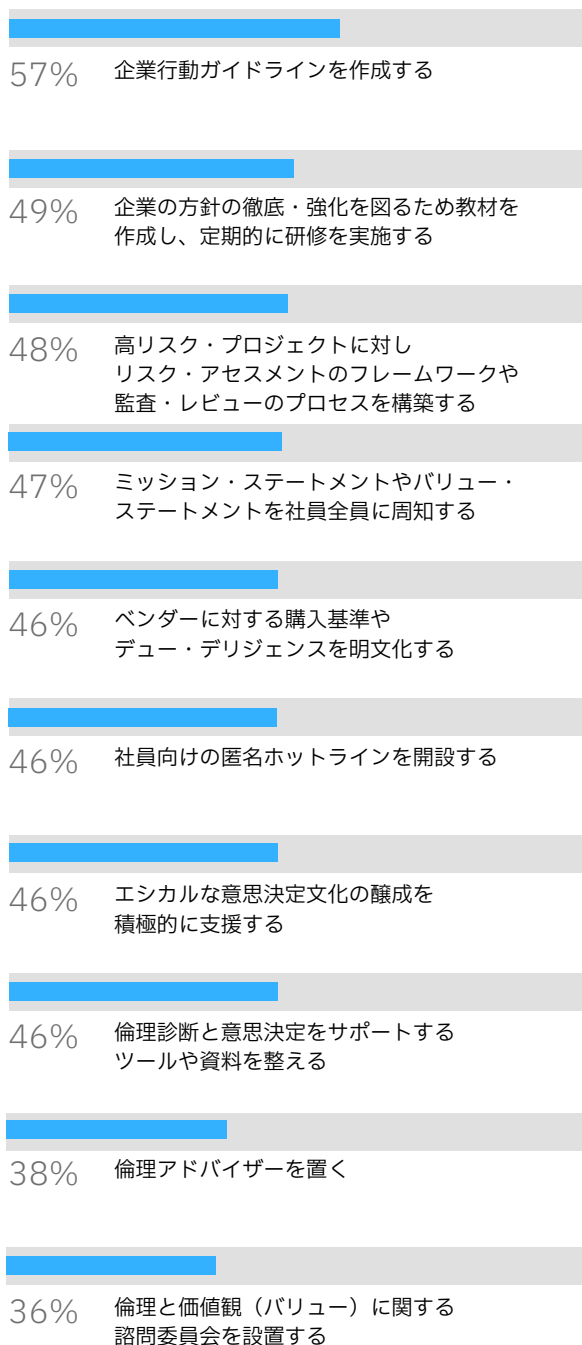
AI をビジネスモデルの中心に据える企業は、AI 倫理に特化したガバナンス・アプローチを採用している。一方、AI を戦略的に重視せず、その活用も非中核部門・部署にとどまっている組織は、まず AI 倫理原則を既存の企業倫理体系に組み込むことから始めるとよいだろう（図 6 参照）。

もちろん、AI を事業戦略に取り入れさえすれば、高い AI 効果が必ず得られるというわけではない。しかし、AI を巡る業界の大きな潮流から取り残されないよう企業は肝に銘じるべきだろう。

図 6

### 初期段階の取り組み

AI 倫理を既存の企業倫理体系に組み込んでいる組織が多い



## 目前の課題に向き合う： 学術的理念から現実世界へ

企業はAI倫理プロジェクトを始めるにあたって、これまで積み上げられた豊かなリソースを有効活用できる。当初、学際的な研究から始まったAI倫理の取り組みは、今ではさまざまな枠組みで活発に行われるようになり、研究成果・実践例・関連団体が蓄積され、活動のネットワークも広がっている（図7参照）。

図7

### 未来を描く

AI倫理は実践フェーズに入っている

	0.0	1.0	2.0	3.0
重点分野	認知	原則の誕生	ポリシーの形成	実行
時期	2016年以前	2017年～2019年	2020年～2021年	2022年以降
アプローチ	小規模で学際的な研究グループ	多様な利害関係者がAI研究に関与	AI開発者や企業などによる組織的連携	組織内・組織間、社会全体での取り組み
けん引するコミュニティー	AI・コンピューター・サイエンス、哲学、法律、経済分野の学術関係者	AI・テクノロジー企業、シンクタンク、標準化団体、AI開発企業などのコミュニティー	企業取締役会、専門職団体、政策当局、規制機関	企業経営層、AI機能の利用者、顧客、認定機関
当事者の具体例	<ul style="list-style-type: none"> <li>- Puerto Rico conference on Future of Life</li> </ul>	<ul style="list-style-type: none"> <li>- AI研究コミュニティー</li> <li>- Asilomar Conference on Beneficial AI</li> <li>- 欧州委員会（EC）のAIハイレベル専門家グループ（High-Level Expert Group on AI）</li> <li>- Partnership on AI</li> <li>- 人工知能に関するOECD原則</li> </ul>	<ul style="list-style-type: none"> <li>- WEFの「Global AI Alliance」</li> <li>- 米国国立標準技術研究所の「AI end user trust initiative」</li> <li>- AIに関するグローバルパートナーシップ</li> </ul>	<ul style="list-style-type: none"> <li>- 部門横断的なチームを立ち上げた企業</li> <li>- 規制をまとめる政府機関</li> <li>- 規格の作成・発行機関</li> </ul>
アウトプットの例	<ul style="list-style-type: none"> <li>- 学術論文</li> <li>- 学会発表</li> </ul>	<ul style="list-style-type: none"> <li>- 組織による原則の表明</li> <li>- WEFのツールキット</li> <li>- EC「信頼できるAIのための倫理ガイドライン」（Ethics Guidelines for Trustworthy AI）</li> <li>- 米連邦政府一般調達局の「テクノロジー・トランスフォーメーション・サービスズ（Technology Transformation Services handbook）」ハンドブック</li> </ul>	<ul style="list-style-type: none"> <li>- 組織によるポリシー表明</li> <li>- IEEE「Trusted Data &amp; AIS Playbook」</li> <li>- ECのホワイト・ペーパー</li> <li>- 「Personal Information Protection Law」（中国）</li> <li>- 「New Generation AI Ethics Specification」（中国）</li> </ul>	<ul style="list-style-type: none"> <li>- 組織による倫理の実践</li> <li>- 企業研修</li> <li>- 技能認定</li> <li>- EUのAI規制（案）</li> <li>- 中国のAI規制（案）</li> </ul>

学術研究と抽象的な理論をいかに実用化へ結実させるか。研究者たちは常にこの壁に挑んできた。大学の権威ある研究者でさえそうである。しかも、対象分野はコンピューター・サイエンスにとどまらない。「AI 倫理は机上の空論からは生まれない」。ケンブリッジ大学「知の未来」研究所 (Leverhulme Centre for the Future of Intelligence) のディレクター兼哲学部シニア・リサーチ・アソシエートを務める Stephen Cave 氏はこう指摘する。<sup>33</sup> 同大学は次世代の AI 倫理学者を育成するため、2021 年に大学院において特別学位の授与を試験的に始めた。

同様の研究組織を開設する大学もいくつか出てきた (オックスフォード大学の「Institute for Ethics in AI (人工知能倫理研究所)」など)。他分野を専攻する学生向けコースも増えている。さらに、さまざまな機関・団体が AI 倫理の最前線に立ち、有用なリソースを提供している。ブルッキングス研究所や「Future of Privacy Forum」などのシンクタンク、IEEE のような専門組織、規制に関わらない政府系機関としては米国国立標準技術研究所。さらに、業界横断的な団体では WEF や「AI に関するグローバル・パートナーシップ (Global Partnership on Artificial Intelligence)」 「Partnership on AI」などがある。このほか、草の根的な協力活動も登場している (米連邦政府一般調達局のテクノロジー・トランスフォーメーション・サービス (Technology Transformation Services) など)。

企業のリーダーはこうした成果を活用することで、連携を求められる社内各部署の多くで AI 倫理の理念を実践できるようになった。ひいてはそれが AI の信頼性を高める力となる。

AI 倫理に対するこうしたアプローチは責任あるテクノロジーの利用を促し、それによって企業の競争優位性を高め、サステナブルなイノベーションを後押しする。さらには、企業が掲げる社会正義の目標達成を助ける力にもなり得る。実際に IBV 調査ではそうした効果を示す結果が出ている。AI と AI 倫理を重要視する企業の 67% 超が、サステナビリティや社会的責任、ダイバーシティ&インクルージョンにおいても同業他社を上回る成果を上げている。

オランダ企業庁 (RVO、英語名「Netherlands Enterprise Agency」) の CIO、Johan Maas 氏は「オランダの政府機関にとって、責任ある AI アルゴリズムの実現は最優先の課題だ」と話す。その上で、「当庁の AI は企業からの補助金申請について職員に助言するだけでなく、助言の基礎となったパラメーターの説明を提供し、納得してもらうなど、ホーリスティックな (科学的正しさだけでなく人間や社会への影響を含めた総合的な) アプローチをとっている」と説明。「こうしたアプローチは人間と AI が効果的に協働する基盤となっている」と強調する。

AI と AI 倫理を重要視する  
企業の 67% 超が、  
サステナビリティや社会的責任、  
ダイバーシティ &  
インクルージョンにおいても  
同業他社を上回る成果を  
上げている。



さらに、「この取り組みは平等な対応といった当庁の AI 開発ポリシーにも合致している。エシカルかつ民主的であり、透明性も確保されている」と指摘し、「このような努力で、当庁の AI 活用に対する職員や市民の信頼を高めている」と話す。<sup>34</sup>

AI 倫理を大きな流れにする環境は整った。企業の意識は変わり、利害関係者を第一に考える方向へ変わりつつある。サステナブルな商品にはプレミアム価格を払ってもよいと消費者は考えている。企業のプロキュアメント部門は物品・サービスの調達に際して倫理基準も重視するようになってきた。今や企業の倫理とパーパス（存在意義）を確立することが、

他社との差別化を図るために不可欠だと見なされている。「企業の社会との関わり方に AI が及ぼす影響がますます大きくなっていることを多くの人が認識している。信頼できる AI を構築するためには、AI 倫理をルールに従って包括的な形で確立することが必要であり、この取り組みを口先だけのスローガンに終わらせてはならない。

企業戦略として必須であるだけでなく、社会的な急務でもある。

今こそ企業は次の一步を踏み出すときである。



# アクション・ガイド

企業が次にとるべきステップは、自社が AI 導入のどの段階にあるのか、戦略と倫理をどう考えるのかによって変わってくる。

まだ AI がビジネスに不可欠だと見なしていない段階にあるなら、競争環境（および規制環境）の変化を注視すべきである。AI をより積極的かつ倫理的に導入する方法を見極める上で極めて重要だからである。その場合、既存の企業倫理体系（18 ページの図 6）に組み入れることが合理的な出発点になるだろう。

すでにその先の段階に進み将来の意欲的なアジェンダを掲げている企業の場合は、社内横断的に AI を管理する AI 倫理委員会（AI Ethics Board）を設置すべきである。この組織が、信頼できる AI 構築に必要なプラクティスを自然な形で社内に根付かせる役目を担う。その際、AI の設計から展開、動的フィードバックに至るまでライフサイクル全体に倫理を組み入れることが不可欠で、それによってプラクティスの実効性も高まる。

すでに AI と AI 倫理を戦略的に導入している組織の場合は、そのアプローチをサプライヤーや顧客、その他のエコシステムのパートナーにまで拡大することで、「AI 倫理の相互運用性（ethical interoperability）」を高めることができるようになる。この用語は、カリフォルニア大学サンディエゴ校のデータサイエンス・哲学教授である David Danks 氏が考案したものである。<sup>35</sup> これにより競合他社との差別化を進め、不正利用から企業を守ることができるようになる。

さらに次のステップは、社内でのあなたの立場や役割によって決まる。あなたが CEO など CxO（最高責任者）レベルの経営層であるなら、方向性を示すことが求められる。事業チームであれば AI が導き出したインサイトに基づいて意思決定を行う。データサイエンティストはデータの精査、AI 開発者はソリューションの構築。その他の部署の社員に割り振るべき業務も多い。

AI 導入はチーム・スポーツのようなものである。AI 倫理もまたしかりだが、多彩なチームが結集し、幅広い領域に対し高いレベルで結束する精神がさらに必要である。技術的な課題にとどまらず、人間と機械が広範に関係するソシオ・テクノロジカル（社会・技術的）な課題にも対処しなくてはならないため、ホーリスティックな取り組みが求められる。<sup>36</sup> まず、次に示した、組織横断的に取り組むべき重点 3 分野のアクションについて検討し、続いて役割に応じた推奨アクションに目を通してほしい。

## 組織横断的に 取り組むべき アクション

### 01

#### 戦略と ビジョン

##### 適切な戦略の中にエシカルな AI のプラクティスを組み入れる

- 事業戦略および企業目的にとって信頼できる AI の構築が非常に重要であることを認識する
  - AI によって加速することが予想される価値創造の主な原動力は何か
  - 成果はどのように評価されるべきか
- 組織の成長戦略とアプローチにおける AI イノベーションの位置付けを検証する（例えば「開拓者（trailblazers）」や「迅速な追随者（fast followers）」など）
- AI 倫理の主要な原則を受け入れる
- 人間と機械が組織内で果たす役割のバランスを決める<sup>37</sup>

主導者・部門：CEO、事業・技術部門の役員、取締役会

### 02

#### 指示と管理

##### エシカルな AI 実行のためのガバナンス・アプローチを確立する

- 利害関係者（企業のリーダー、社員、顧客、政府など）の視点を取り入れる
  - プライバシーや堅牢性、公平性、説明可能性、透明性などの組織にとって重要な原則を取り入れる
  - 企業とエコシステムを考慮する
- AI とデータのリスク特性や指標を定める
- 組織体制やポリシー、プロセス、モニタリングを確立する

主導者・部門：AI 倫理の「推進者」、リスクおよびコンプライアンス部門の役員とリーダー、法律顧問、人事・人財部門の役員とリーダー、ダイバーシティ&インクルージョン担当の役員とリーダー、ビジネスユニット・部門の役員とリーダー、技術チームのリーダー

### 03

#### 実行と展開

##### AI のライフサイクルに倫理を組み入れる

- 利害関係者（企業のリーダー、社員、顧客、政府など）のエンゲージメントの強化
- 組織体制やポリシー、プロセス、モニタリングを確立する
- コンプライアンス・データを収集・報告・レビューする
- チームに対する教育とダイバーシティ活動を推進・サポートする
- 統合的な手法とツールキットを定める

主導者・部門：AI 倫理のプログラム担当部署、事業・技術部門のリーダーとその担当チーム、人事・人財部門のリーダーとその担当チーム、AI 設計者、AI エンジニア、データサイエンティスト、その他の技術チーム

## 役職に応じた 推奨アクション

主導者・部門	アクション例
CEO	<ul style="list-style-type: none"> <li>- AI のイノベーション戦略および事業目標を達成するためのエグゼクティブ・チームを設置する</li> <li>- AI 目標を達成するための資金を適所に配分する</li> <li>- 組織の戦略的な視点や価値観（バリュー）、ビジネス・AI 倫理への対策を、ブログや記者会見、社内会議などを通じて発表する</li> <li>- AI 戦略上の重要性に応じて、全社的な取り組みを主導する最高 AI 倫理責任者を任命する、あるいは現職役員の役割を明確にする<sup>38</sup></li> <li>- AI 戦略上の重要性に応じて、AI 倫理委員会の設置を検討する（議長を務めるか、補佐をする）</li> <li>- 取締役会やその他の主要な利害関係者に進捗状況を報告する</li> <li>- AI 倫理の実践アプローチを承認する</li> <li>- 研修を実施し、AI とデータのリテラシー、変更管理のアジェンダを実行する</li> <li>- AI とデータに関するリスクの総合的な検証、更新を定期的に行う</li> <li>- すべてのビジネスユニットや部門において、役員が共通の AI 倫理に基づいて行動するよう仕向ける</li> </ul>
事業・技術部門 の役員	<ul style="list-style-type: none"> <li>- ビジネス効果を最適化するため、全社的なイノベーション戦略にのっとり、AI、機械学習、アナリティクス、データの各分野で優先して取り組む事項を明確にする</li> <li>- 企業アイデンティティーや企業文化の中に、価値観（バリュー）を取り入れる</li> <li>- 人間とテクノロジーのコラボレーションを設計する</li> <li>- マクロ的な成功指標と目標達成までのタイムフレームを定める</li> <li>- AI に強いテクノロジー・パートナーや研究者、スタートアップ企業、その他のビジネス・パートナーを特定してエンゲージメントを得ることで、イノベーションのエコシステムを強化し、「AI 倫理の相互運用性」を確立する</li> </ul>
取締役会	<ul style="list-style-type: none"> <li>- 戦略的な方向性を承認する</li> <li>- 取締役会全体や主要な委員会（リスク・監査委員会、技術委員会など）のエンゲージメントを介して、エグゼクティブ・チームに責任を持たせる</li> </ul>
AI 倫理の「推進者」 およびプログラム 担当部署	<ul style="list-style-type: none"> <li>- ガイドあるいはコーチ役を果たす</li> <li>- AI 倫理プログラムのための戦略・方針・アプローチを設定する <ul style="list-style-type: none"> <li>• 組織の価値観（バリュー）から AI がもたらす倫理的な影響を導き出す</li> <li>• 運用と組織化のアプローチを決定する（例えば、集中管理型のセンター・オブ・エクセレンスやプログラム・マネージメント・オフィス、部門横断的・ビジネスユニット横断的な分散型のアドボケート（擁護者）の導入、資金、社内外のリソース活用など）</li> <li>• ユーザー同意やオプトアウトの影響など、AI モデルの標準をまとめる</li> <li>• 調達・開発した AI モデルをレビュー・監査する定期的なプロセスを設定する</li> <li>• ガバナンス体制においてビジネス、テクノロジー、ダイバーシティーの人員バランスを評価し、メンバーの交代や委員会の構成を検討する</li> <li>• AI 倫理の改善、エスカレーション、内部告発者の保護、利害対立の軽減に対しインセンティブを与え、透明性を担保する</li> <li>• リスクを軽減するための従業員教育を行う</li> <li>• 組織の AI へのアプローチに AI 倫理を深く組み込む（責任文化をいかに醸成するかを詳述したプレイブックの作成など）</li> <li>• エコシステムのパートナーと協力して、アプローチを深化・拡大させる<sup>39</sup></li> </ul> </li> </ul>

主導者・部門	アクション例
リスクおよび コンプライアンス 部門の役員と リーダー	<ul style="list-style-type: none"> <li>- AI のリスク評価を主導・調整し、プログラムの強化を推奨する <ul style="list-style-type: none"> <li>• リスクやコンプライアンス、プライバシー、セキュリティ、記録の管理やデータ・ガバナンスなど、既存のガバナンス枠組みや文書化、プロセス、関連プログラムの管理などを見直す</li> <li>• オープン・イノベーションの文化とのバランスを考慮しつつ、トレーニングを受けた実務担当者と共に、新たに発生する AI およびデータの倫理上のリスクについて評価する</li> <li>• 新たに発生する AI とデータの倫理上のリスクに対処するため、強化策を提案する（例えば、委員会新設、憲章、プロセス、ツール、緩和策、リスク保険など）</li> <li>• 役割、責任、管理、エスカレーション・プロトコル、評価基準、AI モデル監査、手続き、プロセス、文書化など、新たな AI リスク評価プログラムの構築を主導、調整する</li> </ul> </li> <li>- 法律や規制、政策の変更など、世界における AI ・データ倫理の動向を追跡する <ul style="list-style-type: none"> <li>• 変化および動向分析の概要を定期的に提供する</li> <li>• 規制変更に対応する日常の管理方法を構築または統合する</li> </ul> </li> </ul>
最高情報 セキュリティ 責任者（CISO）	<ul style="list-style-type: none"> <li>- AI 導入の進捗にかかわらず、AI 特有の脅威に対処するセキュリティ・プロセスを開発・強化する <ul style="list-style-type: none"> <li>• 脅威を特定・管理する</li> <li>• コンプライアンス保証を文書で立証し、提供する</li> </ul> </li> </ul>
法律顧問	<ul style="list-style-type: none"> <li>- エコシステムを作る顧問先とそのベンダーと協力して、AI の適用とデータ利用における責任と義務を精査する</li> </ul>
人事・人財部門の 役員とリーダー	<ul style="list-style-type: none"> <li>- リスクを特定し、緩和策を策定する</li> <li>- 研修、AI とデータのリテラシー、変更管理の課題を各部門と協力して特定する</li> <li>- スキル・ギャップを埋め、AI の影響を受ける社員のニーズに対処するためのアクション・プランを策定する</li> </ul>
ダイバーシティ& インクルージョン 担当の役員と リーダー	<ul style="list-style-type: none"> <li>- AI チーム内のダイバーシティを評価する</li> <li>- AI やその他のテクノロジーの要職についても、マイノリティー集団の退職者を減らし、採用・定着を図る方法を協力して特定する</li> <li>- AI、データサイエンスやそれに関連する部署のメンバーの多様性を高める取り組みについて公表する</li> <li>- 「無意識の偏見」がどのようにデータ・セット（および AI ソリューション）に影響を与えるかについて研修を行い、社員の理解を促す</li> </ul>
ビジネスユニット・ 部門の役員と リーダー	<ul style="list-style-type: none"> <li>- 各ビジネスユニット・部門に特有の実行プロセスやポリシーを協力して定義する</li> <li>- 組織全体で進める AI 倫理の取り組みを公式かつ積極的にサポートする</li> </ul>

## 主導者・部門

## アクション例

### 事業・技術部門の リーダーと その担当チーム

- 日々の実務体制を編成する
  - AI チームを立ち上げ、関連プロセスを策定し、強化する
  - 他の社員やパートナーに役割を割り振る
  - 設計者や倫理学者、心理学者、社会学者など、必要となる社外専門家を特定する
- エシカルな AI のプラクティスをエンド・ツー・エンドで展開する
  - 倫理原則をビジネス・技術要件に組み入れる
  - チームからのフィードバックに基づいて課題と機会を評価し、協力して対処する
  - AI 倫理のプラクティスを改善、拡大するための時間をチームに与える
  - 事業の価値観（バリュー）を主要な目標と整合させ、計測可能とした上で、現実的な目標を設定する
  - エシカルな AI のプラクティスを創造する上で学問分野の枠を超えて作業することのメリットを定期的に周知する
  - 組織の「ボトムアップ」の意見を、「トップダウン」のアプローチに取り入れる
- 倫理原則が順守されない場合、チームに問題を提起、対処する権限を与える

### 人事・人財部門の リーダーと その担当チーム

- 全社員を対象に、それぞれのニーズに合わせた AI 倫理の学習プランを作成し、実施する
- AI を利用する業務を増やす必要のある従業員に対し、研修を行う
- AI 導入によって影響を受ける人財に対して再教育、スキルアップを行う

### 設計者、 アーキテクト、 データサイエンティスト、 社会学者、 AI エンジニア、 その他の技術チーム

- 倫理に関する潜在的な懸念を探るための方法とツールの利用を推進する
  - コードを記述する前に、ビジネスなどのエキスパートとともに潜在的な問題を特定し、検討するための「デザイン思考」のセッションを実施する<sup>40</sup>
  - トレーニング・データについて予備的なデータ解析を行う
  - 指標を設定して、公平性を評価し、バイアスを緩和する
  - 「ブラック・ボックス」を可視化することが、技術的に可能であることを伝え、デモで見せる
  - アルゴリズムを部外者も理解できるようにインターフェースを構築する
  - 第三者が評価可能となるよう文書化に取り組む
  - 疑問を提起し、倫理上の懸念に対処するよう部門に促す
- ビジネス上の成功を定義する要素と機能要件を明確に定義する
- 非技術系の人に AI 倫理の必要性を説明する
- 適切な研修が行われていない場合は、自主的に始める。具体的には、公平性と正確性の基準、データ・リネージュ（データ履歴の可視化）とプロヴェナンス（来歴）の仕組み、不利益の分析などについて。
- AI の影響とリスクを評価する
- AI 倫理の重要性を示すビジネスやユーザー事例を見だし、教示する
- 複数のチーム（タイプの異なるチーム）間で知識を共有し協働する

# 著者



---

## Brian Goehring

Global Research Lead, AI  
IBM Institute for Business Value  
(IBM Institute for Business Value、  
グローバル・リサーチ・リード、AI)  
goehring@us.ibm.com  
linkedin.com/in/brian-c-goehring-9b5a453

IBM Institute for Business Value のアソシエート・パートナー。AI ビジネス・リサーチの責任者であり、学会や顧客、その他の専門家と共同でデータ主導のソート・リーダーシップを開発する。戦略コンサルティングにおいて 20 年もの経験を有し、多くの業種・職種のシニアレベルの顧客にコンサルティングを行ってきた。プリンストン大学哲学学士。認知科学とドイツ語の修士証書を取得。

---

## Francesca Rossi

IBM Fellow, AI Ethics Global Leader  
(IBM フェロー、AI 倫理グローバル・リーダー)  
Francesca.Rossi2@ibm.com  
linkedin.com/in/francesca-rossi-34b8b95  
@frossi\_t

T.J. Watson Research Center の IBM フェロー兼 IBM AI 倫理グローバル・リーダー。AI 推進リサーチ・プロジェクトを主導し、IBM AI Ethics Board の共同議長を務める。「Partnership on AI」理事会および「Global Partnership on AI」運営委員会のメンバー。アメリカ人工知能学会 (AAAI) と欧州人工知能協会 (EurAI) のフェローであり、AAAI の次期会長 (内定)。IBM に入社する前は、イタリアのパドヴァ大学 (the University of Padova) のコンピューター・サイエンス教授であった。

---

## Beth Rudden

IBM Distinguished Engineer,  
Cognitive Science and Trusted AI,  
Data & Technology Transformation,  
IBM Consulting  
(IBM コンサルティング、IBM ディスティングイッシュト・エンジニア、コグニティブ・サイエンス・アンド・トラステッド AI、データ&テクノロジー・トランスフォーメーション担当)  
brudden@us.ibm.com  
linkedin.com/in/brudden

世界規模でチームが参加しているグローバル・プラクティスを主導し、その成果としてデータの倫理的な使用に基づく AI モデルを開発し、実用的なインサイトを顧客に提供する。また、開発したソリューションによって明確なインサイトを提供して顧客の理解を助け速やかな実行を可能にした。この開発で複数の特許を取得した。古典考古学学士、人類学修士。





---

## 山田敦

日本アイ・ビー・エム株式会社  
IBM コンサルティング事業本部  
IBM AI センター長  
執行役員 IBM 技術理事

1995年、日本 IBM に入社。東京基礎研究所にて、主に3次元形状処理の研究に貢献。2008年にコンサルティング部門に異動後、2009年に新設された「先進的アナリティクスと最適化」チームのリーダーを務める。併せてデータ・サイエンティストとして、製造業、流通業、保険業をはじめとした多くの企業に対して、データとアナリティクスを活用した業務変革を支援。2017年より IBM 技術理事。社内では、データ・サイエンティスト職のリーダーを務める。2019年より現職。IBM Academy of Technology メンバー、工学博士。

---

## 立花隆輝

日本アイ・ビー・エム株式会社  
東京基礎研究所  
AI 担当シニア・マネージャー  
シニア・テクニカル・スタッフ・メンバー

1998年、日本 IBM に入社。入社以来、現在まで東京基礎研究所に所属し、音楽著作権保護技術、音声合成、音声認識、コグニティブ・ロボティクスの研究開発に従事し、IBM Watson の複数の API の開発や、それを応用したお客様ソリューションの開発に携わる。2016年よりコグニティブ・コンピューティングや AI を担当し、現在に至る。IBM Academy of Technology メンバー、人工知能学会理事、工学博士。

## 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、そこから信頼性の高い、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](https://ibm.com/ibv) よりお申し込みいただけます。また、Twitter (@IBMIBV) や、LinkedIn ([linkedin.com/showcase/ibm-institute-for-business-value](https://www.linkedin.com/showcase/ibm-institute-for-business-value)) をフォローいただくと、定期的に情報入手することができます。

## 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](mailto:iibv@us.ibm.com)) までお問い合わせください。

## 関連レポート

### Advancing AI ethics beyond compliance

Goehring, Brian, Francesca Rossi, and David Zaharchuk. "Advancing AI ethics beyond compliance: From principles to practice." IBM Institute for Business Value. April 2020. <https://ibm.co/ai-ethics>

### The business value of AI

Goehring, Brian and Anthony Marshall. "The business value of AI." IBM Institute for Business Value. November 2020. <https://ibm.co/ai-value-pandemic>

### Sustainability as a transformation catalyst

Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. "Sustainability as a transformation catalyst: Trailblazers turn aspiration into action." IBM Institute for Business Value. January 2022. <https://ibm.co/sustainability-transformation>

## 調査方法

IBM Institute for Business Value は、経済分析・予測を手掛けるオックスフォード・エコノミクス社の協力のもと、北米、中南米、欧州、中東、アフリカ、アジア（中国とインドを含む）の 22 カ国を拠点とする組織の 1,200 人の経営層を対象に、2021 年 5 月から 7 月までの間に調査を実施しました。調査対象となったのは 22 以上の業種にわたる、事業・技術部門の 16 役職（経営層が中心ですが、AI プロフェッショナルも含まれます）に就く方々です。対象となった組織の収益・予算の中央値は 30 億ドル弱でした。

## 注釈および 出典

- 1 Goehring, Brian and Anthony Marshall. “The business value of AI.” IBM Institute for Business Value. November 2020. Previously unpublished analysis. <https://ibm.co/ai-value-pandemic>
- 2 Goehring, Brian, Francesca Rossi, and David Zaharchuk. “Advancing AI ethics beyond compliance: From principles to practice.” IBM Institute for Business Value. April 2020. <https://ibm.co/ai-ethics>
- 3 Gupta, Sachin, Sashank Yaragudipati, Jane Cheung, and Chris Wong. “The last call for sustainability: An urgent growth agenda for consumer products and retail.” IBM Institute for Business Value. August 2021. 邦訳「サステナビリティ経営、不可避の現実 - 消費財・小売業界の経営層に突きつけられた、事業継続への課題とは -」<https://www.ibm.com/downloads/cas/GNRQZBOE>
- 4 同上
- 5 同上
- 6 Edelman Trust Barometer 2021. Edelman. January 2021. <https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021-edelman-trust-barometer.pdf>
- 7 Gupta, Sachin, Sashank Yaragudipati, Jane Cheung, and Chris Wong. “The last call for sustainability: An urgent growth agenda for consumer products and retail.” IBM Institute for Business Value. August 2021. 邦訳「サステナビリティ経営、不可避の現実 - 消費財・小売業界の経営層に突きつけられた、事業継続への課題とは -」<https://www.ibm.com/downloads/cas/GNRQZBOE>
- 8 Henderson, Rebecca. Reimagining Capitalism in a World on Fire. PublicAffairs. 2020.
- 9 Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. “Sustainability as a transformation catalyst: Trailblazers turn aspiration into action.” IBM Institute for Business Value. January 2022. 邦訳「サステナビリティは変革を引き起こす「カタリスト」である - 先駆者は志を行動に変える -」<https://www.ibm.com/downloads/cas/9EPM6VQK>
- 10 同上
- 11 IBM Institute for Business Value Human Insights Global Survey of 14,526 adults. July 2021. 未公表の分析
- 12 同上
- 13 同上
- 14 IBM Institute for Business Value Trust and Transparency Consumer Survey of >14,500 consumers. November 2018. これまで未公開のデータ
- 15 Goehring, Brian, Francesca Rossi, and David Zaharchuk. “Advancing AI ethics beyond compliance: From principles to practice.” IBM Institute for Business Value. April 2020. <https://ibm.co/ai-ethics>
- 16 Guszczka, Jim and Ann Skeet. “How businesses can create an ethical culture in the age of tech.” World Economic Forum. January 2020. <https://www.weforum.org/agenda/2020/01/how-businesses-can-create-an-ethical-culture-in-the-age-of-tech/>
- 17 Balta, Wayne, Manish Chawla, Jacob Karl Dencik, and Spencer Lin. “Sustainability as a transformation catalyst: Trailblazers turn aspiration into action.” IBM Institute for Business Value. January 2022. 邦訳「サステナビリティは変革を引き起こす「カタリスト」である - 先駆者は志を行動に変える -」<https://www.ibm.com/downloads/cas/9EPM6VQK>
- 18 同上
- 19 Goehring, Brian, Francesca Rossi, and David Zaharchuk. “Advancing AI ethics beyond compliance: From principles to practice.” IBM Institute for Business Value. April 2020. <https://ibm.co/ai-ethics>
- 20 Minevich, Mark and Francesca Rossi. “Why you should hire a chief AI ethics officer.” World Economic Forum. September 2021. <https://www.weforum.org/agenda/2021/09/artificial-intelligence-ethics-new-jobs/>
- 21 Interview with Chun Schiros and Daniel Stahl, Regions Bank. November 2021; Misra, Minav. “Regions Bank: A methodology to develop high quality and trustworthy AI.” IBM Blogs. May 24, 2021. <https://www.ibm.com/blogs/client-voices/regions-bank-develops-high-quality-trustworthy-ai/>

- 22 スタンフォード大学、Institute for Human-Centered Artificial Intelligence 副所長、Rob Reich 氏へのインタビュー、2021 年 10 月
- 23 Goehring, Brian, Francesca Rossi, and David Zaharchuk. “Advancing AI ethics beyond compliance: From principles to practice.” IBM Institute for Business Value. April 2020. これまで未公開のデータ。https://ibm.co/ai-ethics
- 24 同上
- 25 同上
- 26 同上
- 27 IEEE Standards Association の Emerging Technologies & Strategic Development ディレクター、John C. Havens 氏へのインタビュー、2021 年 10 月。Abdur-Rahman, Pavel. “AI & Ethics: The Grand Challenge for Our Generation.” IBM Blogs—Canada. May 4, 2021. https://www.ibm.com/blogs/ibm-canada/2021/05/ai-ethics-the-grand-challenge-for-our-generation/
- 28 Edelman Trust Barometer 2021. Edelman. January 2021. https://www.edelman.com/sites/g/files/aatuss191/files/2021-01/2021-edelman-trust-barometer.pdf
- 29 Goehring, Brian and Anthony Marshall. “The business value of AI.” IBM Institute for Business Value. November 2020. https://ibm.co/ai-value-pandemic
- 30 “Responsible Use of Technology: The Microsoft case study.” World Economic Forum. February 25, 2021. https://www3.weforum.org/docs/WEF\_Responsible\_Use\_of\_Technology\_2021.pdf
- 31 “Responsible Use of Technology: The IBM case study.” World Economic Forum. September 2021. https://www3.weforum.org/docs/WEF\_Responsible\_Use\_of\_Technology\_The\_IBM\_Case\_Study\_2021.pdf
- 32 “Responsible Use of Technology: The Microsoft case study.” World Economic Forum. February 25, 2021. https://www3.weforum.org/docs/WEF\_Responsible\_Use\_of\_Technology\_2021.pdf
- 33 ケンブリッジ大学哲学科、Leverhulme Centre for the Future of Intelligence 所長兼上級研究教授の Stephen Cave 氏へのインタビュー、2021 年 2 月
- 34 Netherland Enterprise Agency Rijksdienst voor Ondernemend Nederland (RVO) の CIO、Johan Maas 氏と Johan Suringa 氏へのインタビュー、2021 年 11 月
- 35 カリフォルニア大学サンディエゴ校、データサイエンスおよび哲学教授、David Danks 氏へのインタビュー、2021 年 9 月
- 36 Boinodiris, Phaedra. “Getting to Trustworthy AI.” VentureBeat. March 2021. https://venturebeat.com/2021/03/14/getting-to-trustworthy-ai/
- 37 Mantas, Jesus. “Empathic AI could be the next stage in human evolution—if we get it right.” World Economic Forum. July 2019. https://www.weforum.org/agenda/2019/07/empathic-ai-could-be-the-next-stage-in-human-evolution-if-we-get-it-right/
- 38 Minevich, Mark and Francesca Rossi. “Why you should hire a chief AI ethics officer.” World Economic Forum. September 2021. https://www.weforum.org/agenda/2021/09/artificial-intelligence-ethics-new-jobs/
- 39 Lohr, Steve. “Group backed by Top Companies Moves to Combat A.I. Bias in Hiring.” New York Times. December 8, 2021. https://www.nytimes.com/2021/12/08/technology/data-trust-alliance-ai-hiring-bias.html
- 40 “Why apply design thinking to artificial intelligence.” Team Essentials for AI course. IBM. https://www.ibm.com/design/thinking/page/badges/ai

© Copyright IBM Corporation 2022

IBM Corporation  
New Orchard Road  
Armonk, NY 10504

Produced in the United States of America | April 2022

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](http://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「AI ethics in action - An enterprise guide to progressing trustworthy AI」の日本語訳として提供されるものです。

