ＡＩに関わるリスク

様々な企業において人工知能（Artificial Intelligence：AI）の導入が進みつつあるが、AI には制御不可能性や不透明性といった特性があり、従来の IT ガバナンスの枠組みにはなかった新たな課題も認識されている。AI 活用による効果を最大化するためにはその特性を認識し、適切にリスクに対応することが必要である。本稿では AI に関わる企業の課題やリスクについて、経営・戦略、開発、実装・利用の局面に分けて概括する。

１．ＡＩに関わる課題・リスク

（１）経営・戦略

AI に関わるリスクとして、経営・戦略面ではデジタル・トランスフォーメーション（DX）の遅れによる競争力の低下といった戦略リスクと人材リスクが挙げられる（図表１）。

デジタル技術を使ってこれまでにないビジネス・モデルを展開する新規参入者が登場し、ゲームチェンジが起きつつある。情報処理推進機構（IPA）が東証一部上場企業を対象に実施したアンケート調査[1]によれば、AI や IoT 等のデジタル技術の普及による自社への影響について「自社の優位性や競争力の低下」を懸念する回答が最多となっており、国内のリーディング企業であっても現在の競争力を維持できる年数はそれほど長くはない（半数程度の企業が約５年後まで）と認識されている。企業は、競争力維持・強化のために DX をスピーディーに進めていく必要があり、既存のビジネスの効率化だけではなく、事業構造を AI 起点・サービス起点の事業構造へ転換することも必要とされる。

AI に関わる研究・開発人材は大幅に不足しており、また技術者の相当数が IT 企業側に偏在している。DX を進め、AI による便益を最大限享受するためには、AI を開発・応用できる人材の確保・育成を進めるとともに、経営層から従業員までのあらゆる層において AI を適切に利用するためのリテラシーを高めていく必要がある[2]。

≪図表１≫ ＡＩに関わる課題・リスクの例 ―経営・戦略―

経営戦略		➢ デジタル･トランスフォーメーションの遅れによる競争力の低下 ➢ 戦略･人材･リテラシーの不足等により投資に見合う効果が享受できない
人材教育	開発部門	➢ AI 専門人材が確保･育成できず、開発･保守が難航
	事業部門	➢ リテラシー不足による開発の難航、業務革新を伴わず AI 利用の効果減 ➢ AI に対する過信、目的に見合わない精度の AI の利用、検証等の不足

（出典）各種資料より SOMPO 未来研究所作成

（２）開発

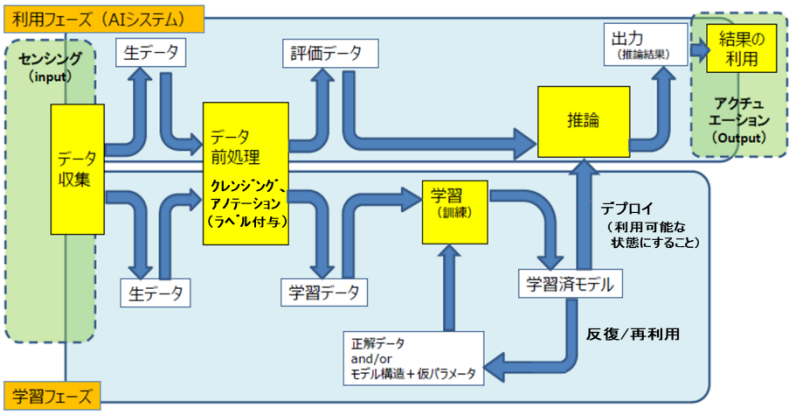
機械学習では、AI の学習と評価に用いるための大量かつ高品質のデータが必要となる。顧客や企業が抱えている問題を洗い出し、ソリューションを検討し、必要なデータの収集やこれまでデジタル化されていなかった情報や知識のデータ化等を進めることとなる。データには権利を有する主体が存在する場合があり、その内容により個人情報保護法、電気通信事業法[3]、著作権法等の規制対象となる。事業者間でデータを連携す

る場合においては、データ提供者が保有するノウハウの流出などの懸念が生じる。このような場合にはデータや成果物の利用等について取り決めることが重要となる[4]。

機械学習では、正解データとどの程度のズレがあるかを計算し、その差が最小になるようパラメーターの調整を繰り返すことで学習モデルの精度を高めていく（図表２）。学習はデータに依存し、データの質・量の不足、代表性の不確保（データの偏り）、バイアスの入り込み[5]等により、学習モデルの精度が損なわれる。

機械学習は、高い精度の予測は得意だが、予測の根拠の説明は不得意であり、開発者自身や専門家も説明できないというブラックボックス化のリスクを有する。

≪図表２≫（機械学習を中心とした）学習と利用の流れ



（出典）総務省 AI ネットワーク社会推進会議「AI 利活用ガイドライン」をもとに当研究所にて一部補記

これらの特性を踏まえ、AI の開発における代表的な課題・リスクを挙げたものが図表３となる。

また、AI の開発・利用においては、技術的なリスクだけではなく、倫理、情報セキュリティ、コンプライアンス、サイバー、開発ベンダ等に関わる複合的なリスクが存在し、企業には AI の用途等に応じてこれらのリスクに明確に対応する態勢を整備することが求められる。

≪図表３≫　ＡＩに関わる課題・リスクの例　－開発－

データ	➢ データの漏えい、不正取得、目的外利用、消費者等との利用目的の認識ギャップ ➢ 通信当事者の情報、通信内容や通信回数・日時、位置情報等通信の秘密の侵害 ➢ データの提供・創出・共用における当事者間の権利等に関する係争 ➢ （デジタル化された）データの質・量の不足、代表性の不確保、バイアスの入り込み ➢ データの変換・加工処理の不良 ➢ 営業秘密の漏えい
モデル	➢ 仮説や学習用プログラムのアルゴリズム等の品質不十分、データの質・量の不足やバイアスの入り込み、適切なデータ学習・検証・アップデートがなされないこと等によるモデルの不良 ➢ 推論プログラム等のブラックボックス化・検証困難化 ➢ データ加工・ラベリング・プログラム開発等における故意・過失による誤った学習・操作
開発ベンダ	➢ 開発ベンダー企業間の認識ギャップ等により生じる開発遅延、精度不良 ➢ アルゴリズムの品質等のブラックボックス化・検証困難化 ➢ 成果物、派生的成果物に関する権利の帰属・瑕疵担保責任の不明瞭 ➢ 特定のベンダへの過度の依存・集中
ガバナンス	➢ 適切なチェック態勢の不足、法令違反・損失の発生

（出典）各種資料より SOMPO 未来研究所作成

（３）実装・利用

AIの実装・利用において企業にとって特に懸念されるリスクは、AIに関する精度や説明をどこまで確保すればよいか不確実であることと、AIの特性やネットワーク化によりAIに関わるトラブルが製品やサービス全体、システム全体に波及する可能性があることではないだろうか。

製造物責任法の対象となる製造物は有体物である動産に限られており、無体物であるAIそれ自体について製造物責任は成立しない。しかし、AIを組み込んだ製品自体は製造物であり、AIに欠陥がある場合にはメーカーは製造物責任を負う。この場合において何が「欠陥」にあたるのかが問題となる。製造物責任法は、「欠陥」を「通常有すべき安全性を欠いていること」と定義しており、これがAIの予見可能性の低さとどのように関連するかはまだわからない。企業は、AIには制御不可能性や不透明性といった欠点が存在することを十分に理解し、それらの欠点を極小化するための利用可能な設計上の対応策を採用することについて、できうる限り検討しておく必要があると考えられる[6]。

また、AIがインターネット等を通じて他のAIと接続・連携することにより制御不能となるなどAIがネットワーク化することによってリスクが惹起され、増幅されるおそれがある。企業には、考えられるリスクを分析し、当該リスクを連携の相手方と共有するとともに、予防策や問題が生じた場合の対応策等を整理し、消費者等に対し、必要な情報提供を行うことが期待される[7]。

AI の利用にあたっては、プライバシー性の高いデータや情報を扱うケースが多い。個人の権利・利益に重要な影響を及ぼす可能性のある分野において AI を利用したプロファイリングを行う場合には、消費者等に生じうる不利益に慎重に配慮する必要がある。AI へのサイバー攻撃により、AI が犯罪等に悪用されたり、消費者等に被害が生じるリスクも想定され、各企業には、AI の用途・侵害の影響等を踏まえた合理的なセキュリティ対策が求められる[8]。

≪図表４≫　ＡＩに関わる課題・リスクの例 －実装・利用－

実装・利用		➢ AI・ロボットの作動・作動不良による生命・身体・財産に関わる事故の発生 ➢ データ、AI の利用範囲・機能・リスク等に関する説明の不足 ➢ AI による処理・取引等の誤り、取引にかかる法令違反 ➢ 消費者等の不利益、不当な差別の発生 ➢ 消費者等の意思決定、行動に対する不適切な操作・干渉 ➢ プロファイリングによる匿名の個人の特定、プライバシー侵害、情報の伝播 ➢ AI に対する過信により生じるヒトによる検証・緊急時の対応能力等の低下
	ネットワーク化	➢ ネットワーク化された AI 間の想定外の相互作用、複雑化による解析不能 ➢ 少数の AI の影響力の強大化 ➢ 個別トラブルのシステム全体への波及 ➢ ネットワークの停止・遅延
	サイバー攻撃	➢ AI プログラム、ネットワークに対するサイバー攻撃による不正操作、情報流出
	知的財産	➢ AI が作成した文書、画像の第三者著作物への類似

（出典）各種資料より SOMPO 未来研究所作成

２．おわりに

AI は、Society 5.0 を実現するための中核技術であり、また PwC によれば世界の GDP を 2030 年には 15.7 兆ドル押し上げる効果があると推定されている[9]。AI は進化を続け、ビジネスに欠かせない経営資源の 1 つになるだろう。AI 活用の効果を高めるためにはリスクを適切にコントロールすることが重要であり、企業は、開発部門だけでなくあらゆる層において AI に関するリテラシーを向上していく必要がある。

【主任研究員　入谷 浩之】

---

1 情報処理推進機構「デジタル・トランスフォーメーション推進人材の機能と役割のあり方に関する調査」(2019 年 5 月)

2 統合イノベーション戦略推進会議「AI 戦略 2019」Ⅱ．未来への基盤作り、日本経済団体連合会「AI 活用戦略」Ⅲ．AI-Ready 化ガイドライン等に AI 教育に関する指針が示されている。

3 電気通信事業法は、第 4 条第 1 項において「電気通信事業者の取扱中に係る通信の秘密は、侵してはならない。」と規定しており通信の秘密を保護している。通信の秘密には、通信の内容・日時・場所、通信当事者の氏名・住所・電話番号等の当事者の識別符号、通信回数等が含まれる。通信の秘密として保護されるデータには、法人に関するデータも含まれる。

4 経済産業省は、企業がデータの利用等に関する契約や AI 技術を利用するソフトウェアの開発・利用に関する契約を締結する際の参考として、契約上の主な課題や論点、契約条項例、条項作成時の考慮要素等を整理した「AI・データの利用に関する契約ガイドライン」を公表している。
＜https://www.meti.go.jp/press/2018/06/20180615001/20180615001.html＞

5 データのバイアスには大別して２つの類型があり、1 つはデータが現実を的確に表していない場合（不正確な測定方法、不完全なデータ収集、標準化されていない報告、データ収集上の不備によるもの)、もう 1 つはデータが長期間の構造的不平等に基づいている場合（例えば、男性を女性よりも昇進させる体系がある業界から集められたデータを使用して、キャリアの成功を予測する場合）があるとされる。(AI NOW,“*The Social and Economic Implications of Artificial Intelligence Technologies in the Near-Term*”, July 2016)

6 平野晋（2018 年 12 月）「AI ネットワークと製造物責任 －設計上の欠陥を中心に」総務省学術雑誌「情報通信政策研究」第 2 巻第 1 号

7 総務省情報通信政策研究所 AI ネットワーク社会推進会議「報告書 2019」AI 利活用ガイドライン

8 同上

9 PwC, “*Sizing the prize*” , 2017