AI（人工知能）の利活用で求められるリスク対策

機械学習やディープラーニングを柱とする現在の第三次 AI（人工知能）ブームにおいて、AI は実際の社会に幅広く用いられ始めており、活況を呈している。今後、AI を活用することは、企業が競争力を向上させるうえで必須の取り組みであると考えられるが、活用にあたっては AI がもたらすリスクについて正しく理解し、適切な対策を講じていくことも重要である。本稿では、国内外における AI の利活用状況を紹介し、今後企業が AI を利活用していくうえで留意すべきリスクと対策について解説する。

1. 国内外における AI の利活用状況

（1）AIの歴史と現在の第三次AIブーム

「AI（人工知能：artificial intelligence）」という言葉は、1956 年に米国ダートマス大学（Dartmouth College）内で行われた国際会議で計算機科学者のジョン・マッカーシー氏（John McCarthy）が命名し、「知的なコンピュータプログラムを作る科学と技術」と定義され社会に認知された[1]。以降、1950 年代後半～1960 年代に第一次 AI ブームが到来し自然言語処理の技術が発展、また 1980 年代には第二次 AI ブームが到来し音声認識等の技術が発展を遂げたが、いずれも「AI（人工知能）」という言葉から連想される姿と開発される技術に乖離がみられたため「期待外れ」に終わっている[2]。

2000 年代に到来した現在の第三次 AI ブームは、主として「ビッグデータ」と呼ばれる大量のデータを用いることで AI 自身が知識を獲得する「機械学習[3]」と、知識を定義する要素を AI が自ら習得する「ディープラーニング[4]」（深層学習や特徴表現学習とも呼ばれる）の 2 つの技術に代表されるが、過去の 2 つのブームと異なるのは、技術が実際の社会に幅広く用いられ始めている点である。2018 年にコンサルティング大手のボストンコンサルティンググループ（BCG）が世界 7 か国、約 2,700 名の管理職を対象に行った調査[5]では、中国では 32%、米国では 22%、日本においても 11%の企業が既に AI をビジネスで活用しているとし、今後ますます AI の利活用が進展することが予想されている。

（2）国内におけるAIの利活用意向・状況

独立行政法人情報処理推進機構が発行する「AI 白書 2020」によれば、調査対象 350 社のうち約 9 割の企業が「AI を既に導入している」（3.1%）、または「実証実験を行っている」（7.1%）、「利用に向けて検討を進めている／これから検討予定である／関心はある」（78.6%）と回答しており、その高い関心度合いがうかがえる。その中で、今後 AI の導入を検討予定、または AI に関心のある企業におけ

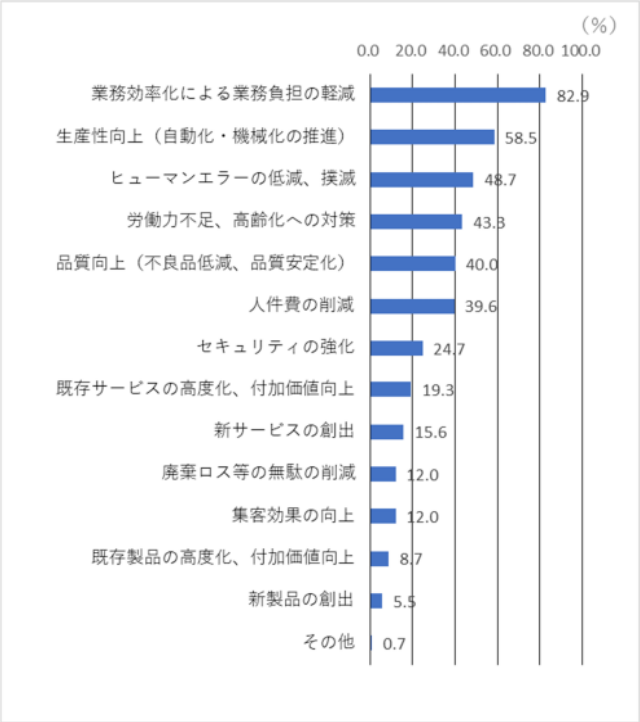
[1] 人工知能学会ホームページ「人工知能の FAQ」
[2] 総務省「令和元年版　情報通信白書」
[3] 大量のデータから規則性やルールなどを学習し、与えられた課題に対して推論や回答、情報の合成などを行うこと。
[4] 人間の神経回路を模したニューラルネットワーク（Neural Network）で深い階層のモデルを構築し、精度の高い推論を行うこと。
[5] Boston Consulting Group 「Mind the (AI) Gap　Leadership Makes the Difference」

るその導入目的、及び業務分野は図 1 の通りである。

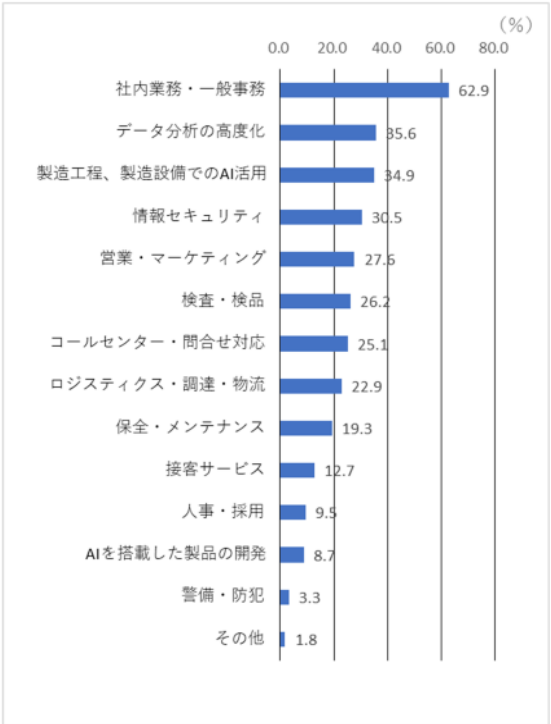
AI の導入目的については、8 割以上の企業が「業務効率化による業務負担の軽減」を挙げ、4 割以上の企業が「生産性向上（自動化、機械化の推進）」、「ヒューマンエラーの低減、撲滅」、「労働力不足、高齢化への対策」、「品質向上（不良品低減、品質安定化）」を挙げている。AI は、特定の分野に偏らず、様々な分野で企業に期待されていることが読み取れる。また、業務分野については半数以上の企業が「社内業務・一般事務」を挙げ、また 3 割以上の企業が「データ分析の高度化」「製造工程、製造設備での AI 活用」、「情報セキュリティ」を挙げた。内部向け業務、外部向け業務を問わず幅広い業務で AI を活用したい企業の意向が読み取れる。

■図 1　日本国内における AI の導入目的と業務分野

AI の導入目的



AI を適用する業務分野



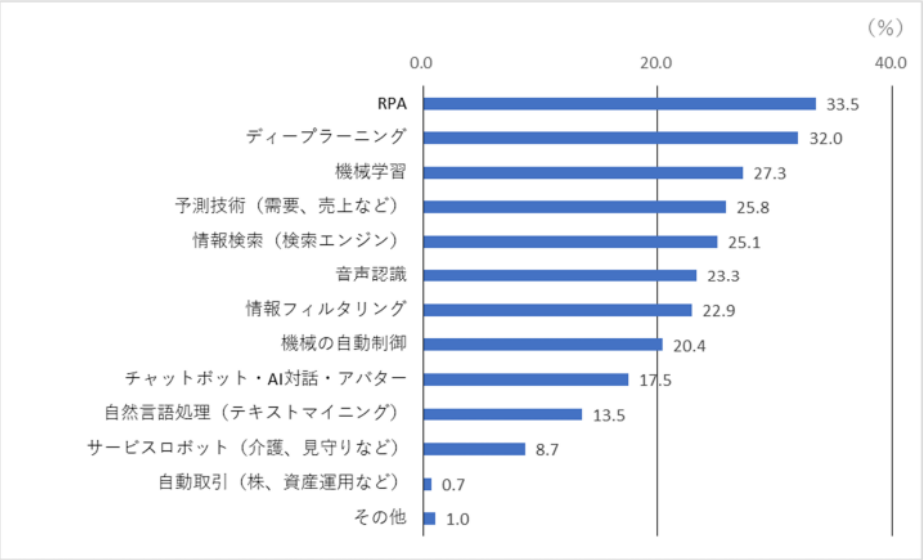
（出所：独立行政法人情報処理推進機構「AI 白書 2020」）

次に、今後 AI の導入を検討予定、または AI に関心のある企業が導入したい AI 技術については図 2 の通りである。要素技術的な AI に焦点を当てると、3 割以上の企業が「RPA（Robotic Process Automation）[6]」、「ディープラーニング」を挙げており、3 割近い企業が「機械学習」を挙げている。また、既に実導入が進んでいる「チャットボット[7]・AI 対話・アバター」に関しても、2 割近い企業が挙げている。

[6] 人間がコンピュータを操作して行う作業を、ソフトウェアによる自動的な操作によって代替すること。主に企業などのデスクワークにおけるパソコンを使った業務の自動化・省力化を行うもの。RPA は、基本的には人間が設定したルールに従い、忠実に作業を実行するものであり、機械学習やディープラーニングのような自律性を持たないが、昨今では一般的に RPA を広義の AI に含める風潮もあることから、本稿の対象範囲としている。

[7] ボット(bot)とは、人間による操作や作業を代替したり、人間の行為を模して人間のように振る舞ったりして、自動的・自律的に行動するソフトウェアやシステムのこと。

■図２　国内企業が今後導入したい AI 技術



（出所：独立行政法人情報処理推進機構「AI 白書 2020」（一部抜粋））

また既に AI を導入し、成果を挙げた企業が相次いでいるが、その一例について以下に例示する。幅広い業種、業務においてその成果が確認されている。

■表１　AI への置き換えが起きている業種・業務例

業種（企業名）	AI の活用例とその効果
製造業 （富士通株式会社）	新製品の部品数やプリント基板のサイズを入力するだけで、AI が必要なプリント基板の層数を算出。様々な条件を加味しなければならないので最適解を導き出すことは難しかったが、AI によりプリント基板の設計工程を 20%短縮。
製造業 （ブリヂストン株式会社）	数百のセンサーを使って成型中のゴムの情報を集め、その無数の情報を AI が即座に計算して最適な回転数と圧力を調整。人の手と目を使わない分、生産性は従来の成型設備より 2 倍に向上し、人手は 3 分の 1 に減少。
小売業 （コンセント株式会社）	顧客を撮影して性別や大体の年齢層を推測。その基本情報に質問で得た回答を加えて、顧客情報をつくりあげる。さらに過去の顧客情報から、新たに獲得した顧客情報に似たものを探し、そのとき売れた商品をお客様に推薦。
小売業 （eBay Inc.）	「Find It On eBay」（商品の画像による検索機能）：インターネット上の芸能人などが履いている靴と同じものを画像により探すことが可能。
学習塾 （atama plus 株式会社）	AI が生徒一人一人の苦手分野を分析し、個別にコンテンツを作成。高校 3 年生のセンター試験の過去問の得点が、AI 教育を受けさせた後では 1 科目平均 20 点上昇するような、効率のよい学習システムを提供。
金融業 （株式会社 J.Score：みずほ銀行とソフトバンクによる共同出資）	PC やスマートフォンを通じて、年齢や年収、家族構成など 18 の項目に回答が求められる。回答が終わると、AI が何パーセントの利率でいくらまで借りることができるかを表示する。スコアが表示されるまで 2、3 分しかかからないうえ、AI スコアリングにより適正な金利で融資が可能に。

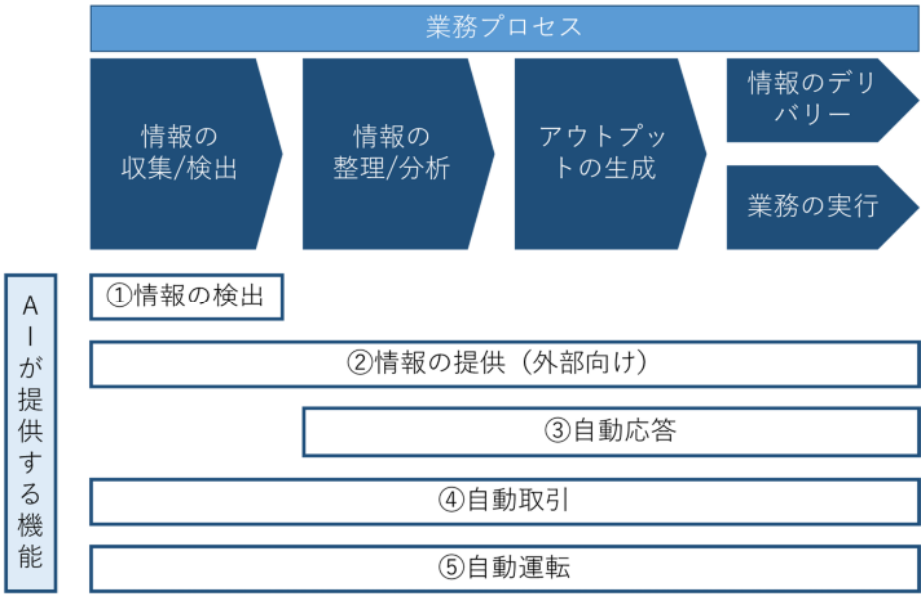
（出所：NISSEN デジタルハブなどの情報から弊社作成）

2. AI 活用上におけるリスク

（１）AIの導入により変化するリスク

AI は人手により行われてきた業務を代替し、かつ代替の結果、人手により実行される業務と比べ、その処理速度や品質を向上させることがベネフィットとして期待されているが、一方で AI による業務処理がもたらす新たなリスクもある。ここでは、AI が提供する代表的な機能を例にとり、AI 導入によるリスクの変動について考え方を整理する。

■図３　概念的な業務プロセスにおける、AI が提供する代表的な機能



（筆者作成）

①　情報の検出

AI がインターネット上の Web サイトや社内のメールサーバを巡回し、コンプライアンス違反等を検知する機能。例えば、自社製品の模倣品が市場に出回っているかを確認するため、AI が電子商取引サイトに出品されている商品画像を自社の商品と照合することで、競合企業における意匠上の違反を摘発する機能や、社員が外部に対して送信している電子メールの相手先や内容により、賄賂やカルテル、インサイダー取引等の商法違反を摘発する機能がこれに該当する。

これらは、人手で行っている検出業務の精度を高めることに寄与するため、基本的には企業のリスクを低減させる方向に作用するものと考える。

②　情報の提供（外部向け）

AI がインターネット上や社内のデータベースなどから収集した情報をもとに、自動生成した情報を外部向けに提供する機能。例えば、証券会社が株式取引に関する推奨レポートを AI により生成し、自社の口座を持つお客様に提供することや、小売業がインターネット上に掲載されている商品の使用効果と、社内で保有する商品の詳細仕様や発注ロット、販売価格等の取引条件を情報源としてメールマガジンを AI が生成し、販売促進目的に取引先に提供することがこれに該当する。

これらは、人手で行っている情報提供の頻度を高め、内容の充実を図ることに寄与するが、その情報が不適切であったために活用した相手先が不利益を被る場合や、活用した情報に相手先の著作権が含まれていた場合は、損害賠償等の訴訟を提起される可能性がある。よって、企業のリスクを増大させる可能性をはらんでいるものと考える。

③ 自動応答

社内における電子メールのやり取り記録などを「教師データ[8]」として読み込み、人手による応対に代わり AI が応対する機能。代表的なものとして「チャットボット」と言われる、自動応答の仕組みがこれに該当する。

教師データに不適当な内容が含まれていた場合、AI がそれを学習しチャットボットに利用してしまうとリスクが発生する。例えば、過去の人事部門による福利厚生制度の応対において、「そんなことも知らなかったの。何年うちの会社に勤務しているの。人事考課に反映させますよ、笑」等の記録があり、それがチャットボットに反映されてしまった場合は、チャットボットによるハラスメントのリスクが発生する。よって、企業のリスクを増大させる可能性をはらんでいるものと考える。

④ 自動取引

外国為替や株式の売買において、売買を執行する際の価格または変動率を人手で設定し、それをトリガーとして取引を行う自動売買の仕組みは既に普及しているが、過去の変動情報を教師データとして読み込み、売買を執行する際の価格または変動率を自動設定・更新する仕組み[9]が AI を利活用した機能に該当する。

これらの機能においては、従前の売買と同様に、損失が発生した場合の責任は売買を AI に委託した者の自己責任となるため、AI の活用により企業のリスク量が変動するということにはならない。ただし、未完成・未成熟の AI を採用し、その情報源に明らかな過不足が見られた場合はクレームに発展する可能性があり、よって企業のリスクを増大させる可能性をはらんでいるものと考える。

⑤ 自動運転

日本では 2019 年に道路交通法の一部を改正する法律が成立し、「SAE レベル 3」（条件付運転自動化）[10]の自動運転車が道路を走行できるよう規定が整備された。このような法規制の動きもあり、現在レベル 3 以上の自動運転車の販売に向け、世界各国の自動車会社や IT 企業が開発・実証を進めているところである。一方、この実証中において既に人身事故[11]が発生しており、それを機に自動運転車が人身事故を起こした場合の責任について議論が活発となった。

国土交通省は、2018 年に自動運転中の車が事故を起こした際の責任について報告書をとりまとめ、自動運転車と手動運転車が混在する過渡期（2020 年から 2025 年頃）においては、迅速に被害者を救済するために、現在の自動車損害賠償保障法で定めている「自動車所有者（運行供用者）の事実上の

---

8 機械学習において予め与えられる、例題と答えについてのデータ。この大量のデータをもとに、AI は出力結果の最適化を行う。

9 参考：人工知能(AI)と為替(FX)　https://fx-koryaku.com/fx-usefulcolumn/fx-ai-20170621

10 米国の NPO 団体である SAE International が策定した自動運転の定義。レベル 3 は「システムが高速道路など特定の場所に限り交通状況を認知して、運転に関わる全ての操作を行う。ドライバーが緊急時やシステムが作動困難になった場合は対応を行う。」

11 2018 年 3 月、アメリカ・アリゾナ州で配車サービス大手「Uber（ウーバー）」が開発中の自動運転車が、公道を試験走行中に、自転車を押して道路を横断しようとした女性をはね、女性はその後死亡。

無過失責任」は変わらない[12]とした。このことから、自動運転車を業務車として利用する企業のリスクは、自動運転車の精度に依存し、手動運転よりも事故を起こしにくい場合はリスクを低減させ、逆に事故を起こしやすい場合はリスクを増大させるものと考える[13]。

また、機械学習の技術を活用した AI について、インプットとなる情報源とアウトプットとしての成果物は人手による業務実施時と同様に明確であるものの、その作業過程や論性構成がブラックボックス化されている点は人手による業務実施時との大きな違いである。上述のようなリスクが顕在化した場合、一般的にはリスクが発生した理由を被害者に説明することが問題処理上求められる（説明責任）が、機械学習ではそれが難しい。

（２）企業における一般的なリスクとAIの利活用により生まれるリスク

企業におけるリスクの洗い出しと評価にあたっては、RCSA（Risk Control Self Assessment：リスク管理自己評価）と呼ばれる手法[14]を用いることが一般的である。その手法で用いられる「リスク一覧表」を所与として、AI 導入によるリスクを抽出し、リスクシナリオとして影響を整理した結果が表 2 である。AI により、企業のリスク環境は少なからず影響を受けることになるため、AI の採用を決定する主管部門は、関連する業務部門と十分に採用後の影響について検討を行い、その採否を判断することが求められる。

■表２　AI の導入により変化すると考えられるリスク

リスク項目		AI が関与した場合のリスクシナリオ	本事例に関係する AI の機能
大項目	小項目		
法務	顧客からの賠償請求	AI により生成された情報サービスを利用して何らかの被害を受けた顧客から、賠償請求を提起された。	情報の提供（外部向け）
	知的財産権に関する紛争	AI により生成されたアウトプットが著作権を侵害しているとして、訴訟を起こされた。	情報の提供（外部向け）
労務	パワーハラスメント	AI により提供されるチャットボットの内容にパワハラのようなものが含まれており、心身に不調を訴え欠勤する従業員が増加した。	自動応答
	セクシャルハラスメント	AI により提供されるチャットボットの内容にセクハラのようなものが含まれており、従業員が、会社の管理不行届きが原因として会社を提訴した。	自動応答
	差別	AI により提供されるチャットボットの内容に差別的な用語が含まれており、従業員が、会社の管理不行届きが原因として会社を提訴した。	自動応答
	通勤途上災害	AI により運転される自動車が事故を起こし、本人や相手が死傷した。	自動運転
	労災事故	AI により制御される産業ロボットが誤作動を起こし、従業員等が死傷した。	自動運転

12 自動運転のシステムの欠陥や誤作動が原因の事故については、保険金を支払った保険会社が、自動車メーカーに対する求償権行使の実効性を確保するための仕組みを検討することも確認されている。
13 その他、人身事故を回避することができない状況下で、A さんと B さんのどちらを犠牲にするのかという倫理的な議論も自動運転車の課題としてあるが、本稿では割愛する。
14 弊社サービス「リスク洗い出し・リスクマップ策定支援」参照　https://www.tokiorisk.co.jp/service/risk_crisis/assessment/

リスク項目		AI が関与した場合の リスクシナリオ	本事例に 関係する AI の機能
大項目	小項目		
財務・ 経理	不適切な会計処理	AI により提供されたチャットボットの経理処理に関する回答が会社の仕訳ルールと相違があり、事後に仕訳修正が必要となった。	自動応答
	為替･金利変動	AI による為替や保有株式の自動取引において、多額の損害が発生した。	自動取引
	原材料の高騰	AI による先物予約等の自動取引において、多額の損害が発生した。	自動取引
製品・ サービス	欠陥商品・リコール	AI で制御される産業ロボットに異常運転が発生し、多数の欠陥商品が市場に出回った。	自動運転
営業・販売	お客様対応の不備	AI が生成するマーケット分析レポートに誤りがあり、信用したユーザに損害が発生した。	情報の提供 （外部向け）
情報管理	機密情報の漏えい	AI が生成した情報ソースに製造/サービス提供ノウハウが含まれており、外部に漏えいした。	情報の提供 （外部向け）
	個人情報の漏えい	AI が生成した情報ソースに個人情報が含まれており、外部に漏えいした。	情報の提供 （外部向け）
事故	輸送中の事故	自動運転のトラックが事故を起こし、配送中の製品が損傷した。	自動運転
評判	宣伝・広告の失敗	AI が生成した広告に問題となる表現が含まれており、消費者からクレームを受けた。	情報の提供 （外部向け）

（筆者作成）

3. AI 活用において求められるリスク対策

（１）人間がAIを監督する仕組みの構築

上述の通り、AI だけで業務が完結される場合は様々なリスクを伴う。よって、（AI の導入直後においては特に）AI によって生成されたアウトプットを活用する前に、今までその業務を実施してきた者などがその内容を確認することが求められる。

[AI を監督する者が確認すべき事項]

- AI が生成した文章の表現に違和感はないか
- AI が生成した文章の内容に違和感はないか
- AI が生成した文章は著作権上問題のないものか
- AI が判断した内容に違和感はないか

（２）AIが果たせない「説明責任」への対処

AI（特に機械学習）により生成されたアウトプットに不具合が生じた場合、その理由について十分な説明責任を果たせないことが想定される。よって、予め使用する者に対してその旨を説明するとともに、AI が生成するアウトプットに万が一不具合が生じても AI を提供する側の一方的な責任とならないよう、相手方と十分に協議しておくことが求められる。

[AI が果たせない説明責任について、使用者に説明しておくべき内容の例]

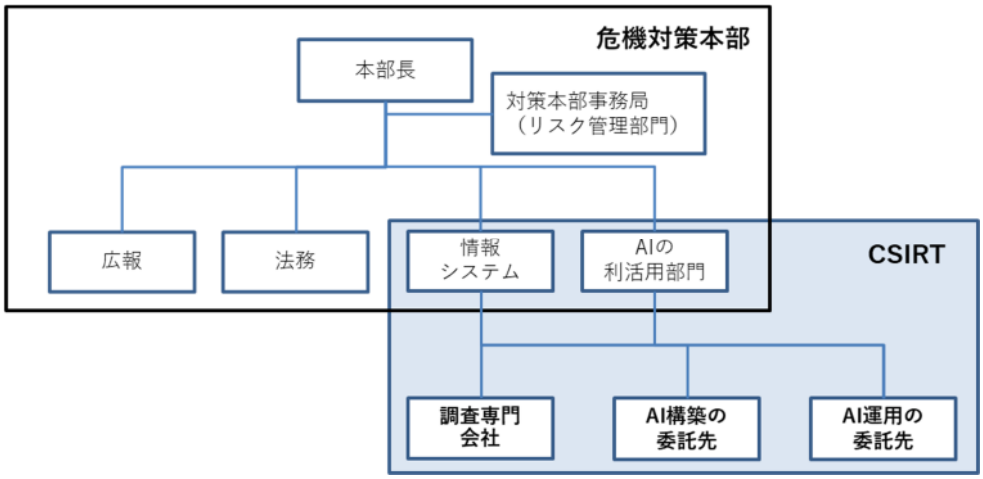
- AI のアウトプットを利用して不利益を被った場合の責任は、利用者自身であること
- AI に不具合があった場合は、教師データの品質向上等により不具合の再発防止に努めるものの、再発を完全に防止することの保証は出来ないこと

また、AI が生成するアウトプットがもたらす不利益について、契約する相手方との交渉においてどうしても免責とならない場合は、想定される損害賠償規模に相応の金額を社内的に引き当てておくことや、損害保険商品により当該リスクの転嫁ができないか保険会社に相談すること、または契約の締結を見送り、リスクを回避することも求められる。

(3)危機対応体制の構築

AI により発生したリスクが危機的な事象をもたらした場合、その影響範囲が広く、また深刻なものであれば会社として危機対応のための組織を立ち上げ、全力でその対処にあたる必要がある。危機対応のための組織としては、リスクの状況や原因を調査し上位に報告する「CSIRT（Computer Security Incident Response Team）[15]」と、CSIRT から報告を受けて、会社としての対応方針を迅速に決定するための「危機対策本部」が必要である。

■図４　AI がもたらす危機発生後の体制（イメージ）



（筆者作成）

■表３　CSIRT における役割機能

役割	内容
● 情報システム	・ CSIRT の統括・対策本部への報告 ・ 調査専門会社への依頼
● AI の利活用部門	・ AI を導入し、利活用している主管部門としての各種管理 ・ AI 構築・運用を委託する関係先との連携（報告を受け、指示を展開）

[15] 企業や行政機関などに設置され、コンピュータシステムやネットワークに問題に繋がる事象が発生した際に対応する組織。

役割	内容
● 調査専門会社	・ フォレンジック[16]の可能な会社 ・ インシデントの原因を調査・整理し CSIRT 統括機能へ報告
● AI 構築の委託先	・ AI の利活用部門からの指示を受け、AI の構築上問題がなかったか検証し報告
● AI 運用の委託先	・ AI の利活用部門からの指示を受け、AI の運用上問題がなかったか検証し報告

（筆者作成）

■表４　危機対策本部における役割機能

役割	内容
● 本部長	・ インシデント対応に関する最終的な意思決定
● 対策本部事務局	・ 情報集約、本部長への意思決定案の上申
● 広報	・ 影響の大きさ、深刻に応じた広報対応方針案の作成 ・ 広報実施時のステートメント作成
● 法務	・ 法的側面での対応検討（適宜、弁護士と連携）
● 情報システム	・ 技術的な対応検討（ネットワーク、アプリケーション等） ・ CSIRT の統括・対策本部内での状況共有
● AI の利活用部門	・ 影響範囲に関する報告 ・ 対策本部の指示を展開

（筆者作成）

4. おわりに

AI を活用することは、企業が競争力を向上させるうえで必須の取り組みであると考えられるが、同時に AI がもたらすリスクについて正しく理解し、適切な対策を講じていくことも重要である。今後 AI の利活用を検討される企業において、少しでも気づきがあれば幸いである。

以上

[2020 年 6 月 9 日発行]

16 情報処理・管理に関する事件や事故（インシデント）が発生した際に、コンピュータに保存されているデータを調査し、事件・事故の発生原因を追究すること。

東京海上日動リスクコンサルティング株式会社
TOKIO MARINE NICHIDO
ビジネスリスク本部　主席研究員　青島健二
〒100-0004 東京都千代田区大手町 1-5-1 大手町ファーストスクエア ウエストタワー23 階
Tel. 03-5288-6594
www.tokiorisk.co.jp