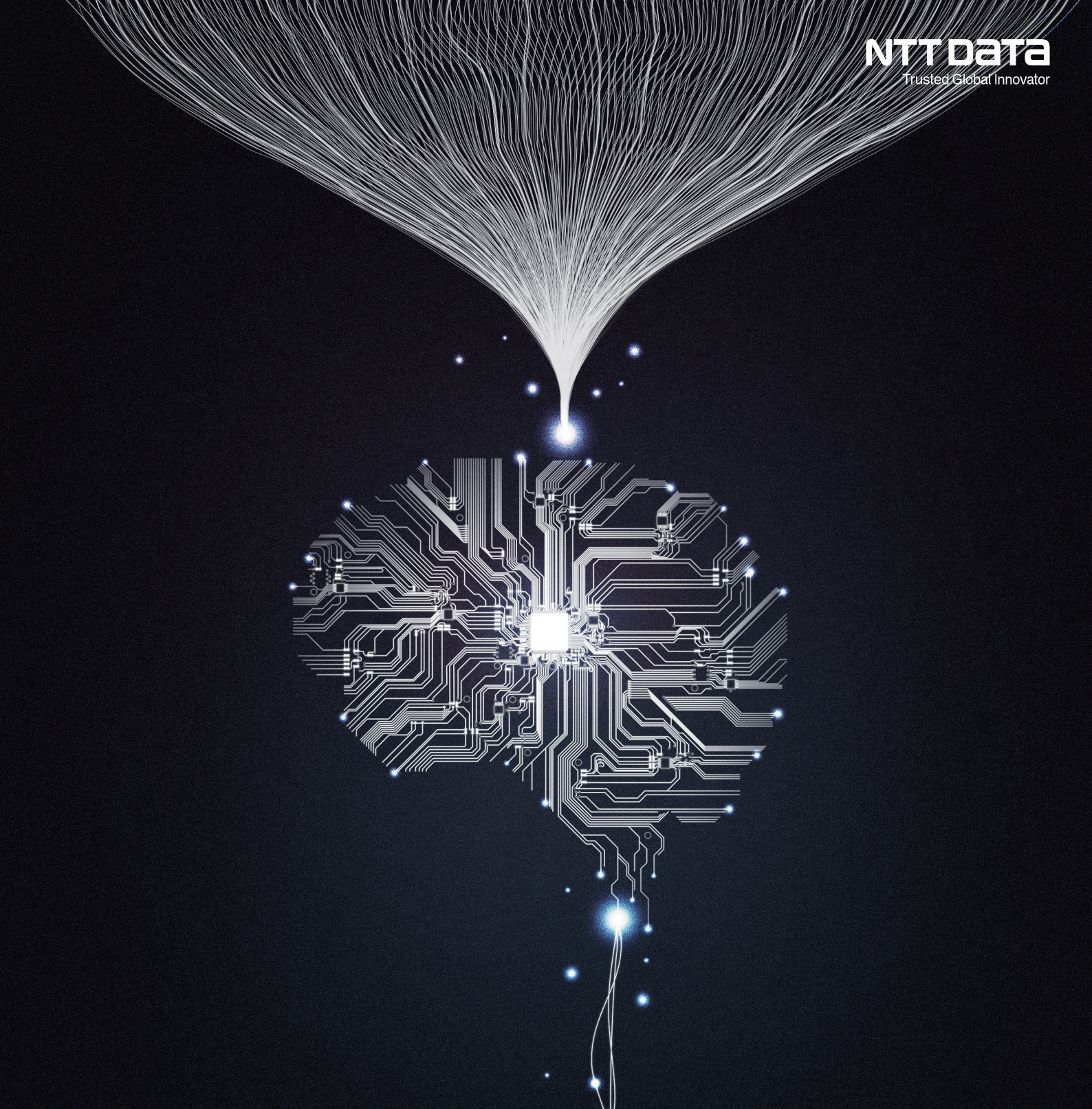
NTT DATA
Trusted Global Innovator



AIシステムの開発・運用・利活用に、安心と信頼をもたらす

「AIガバナンス」

AIガバナンス、AI規制、AI倫理、AIリスク、AI原則、AI開発方法論

C O N T E N T S

CHAPTER.1
はじめに

ディープラーニングの登場をきっかけに AI の能力が向上したことで、AI はサイバー空間から実世界で活用されるようになり、その活用範囲は自動運転や融資審査、画像診断など人命や経営に関わる高い信頼性が求められる領域に拡大してきました(図 1)。
一方で AI の悪用やテスト不足による事故、学習用データの偏りや不足による誤動作といった問題が散見されるようになり、AI 活用に求められる視点は、精度といった機械学習的な視点から、可用性やセキュリティなどシステム的な視点に加えて、AI のアウトプットに対するアカウンタビリティや透明性といった社会的な視点、さらには、本当にその問題に AI を適用して良いのかといった倫理的な視点に拡大しています(図 2)。

図 1 AI の活用領域の拡大

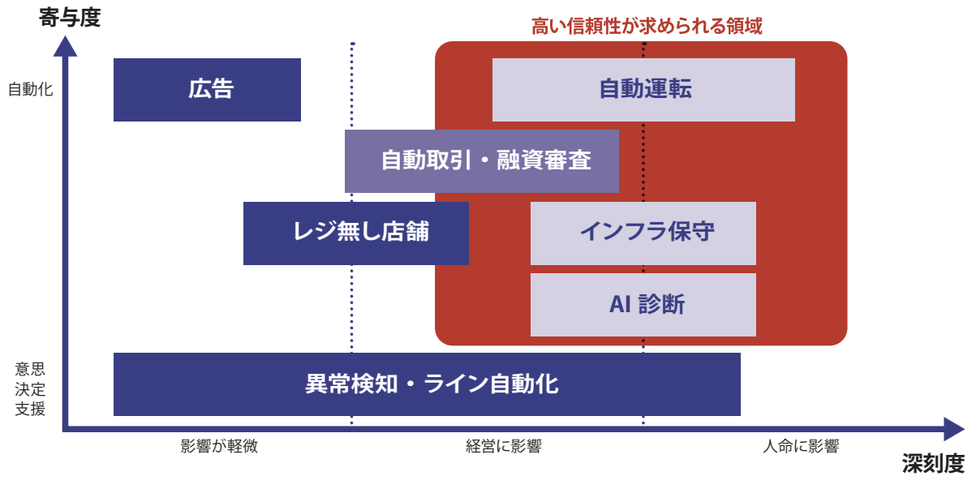
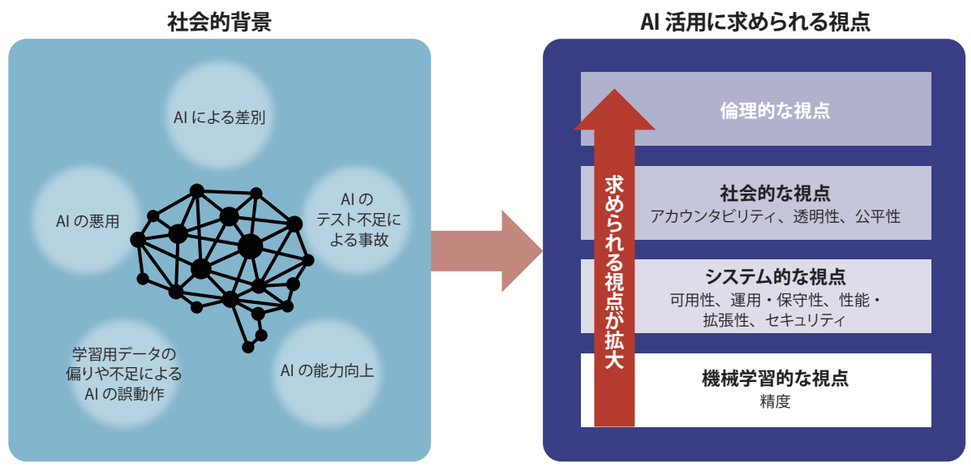


図 2 AI 活用に求められる視点の変化



CHAPTER.2
AI 原則や AI 規制に関する国内外の動向

こうした AI 活用に求められる視点の変化を踏まえ、AI 活用に関するガバナンスの必要性の声が高まり、日本国内では原則やガイドラインを皆で守るいわゆる「ソフトロー」の方向で検討が始まりました。

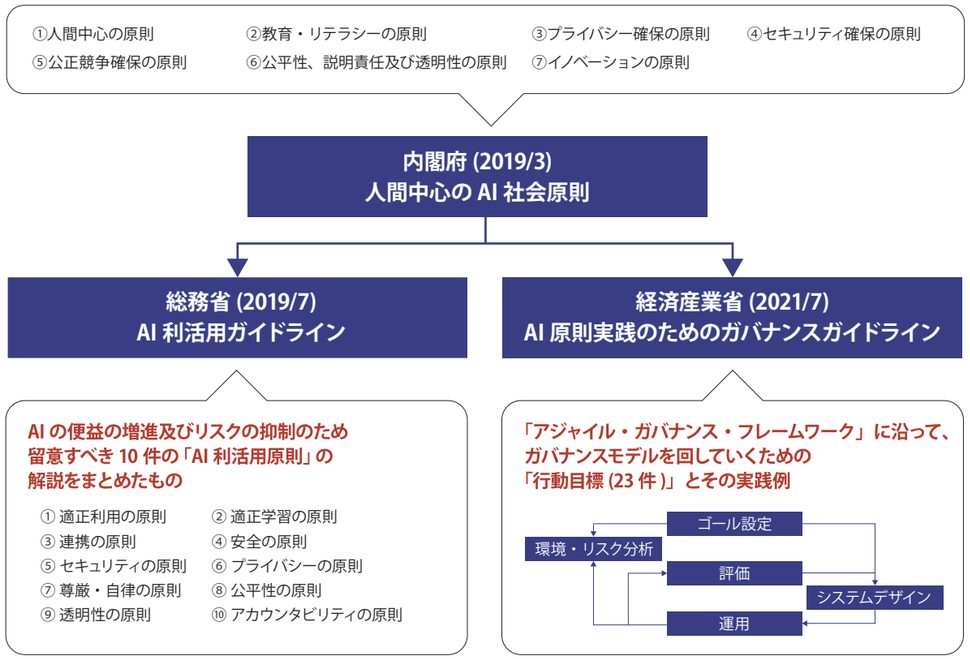
まず、2016 年 4 月に高松で開催された「G7 情報通信大臣会合」にて、日本から AI システムの開発において留意することが期待される事項である「AI 開発原則」のたたき台を紹介し AI の国際的議論の必要性が提起されました。

その後、日本国内では総務省の AI ネットワーク社会推進会議によって、2017 年 7 月に「AI 開発原則」の内容の解説をまとめた「AI 開発ガイドライン案」が策定されました。また、2018 年 8 月には AI の便益の増進及びリスクの抑制のため留意すべき 10 件の「AI 利活用原則」、および、翌年の 2019 年 7 月に「AI 利活用原則」の解説をまとめた「AI 利活用ガイドライン」が公開されました。

一方、内閣府では「AI-Ready な社会」を実現し、AI の適切で積極的な社会実装を推進するために各ステークホルダーが留意すべき基本原則として 2019 年 3 月に 7 か条から成る「人間中心の AI 社会原則」が策定され、公開されています。

最近では経済産業省から「アジャイル・ガバナンス・フレームワーク」に沿って、ガバナンスモデルを回していくための「行動目標 (23件)」とその実践例をまとめた「AI 原則実践のためのガバナンスガイドライン」が公開されています (図 3)。

図 3　日本における AI に係る原則やガイドラインの概要



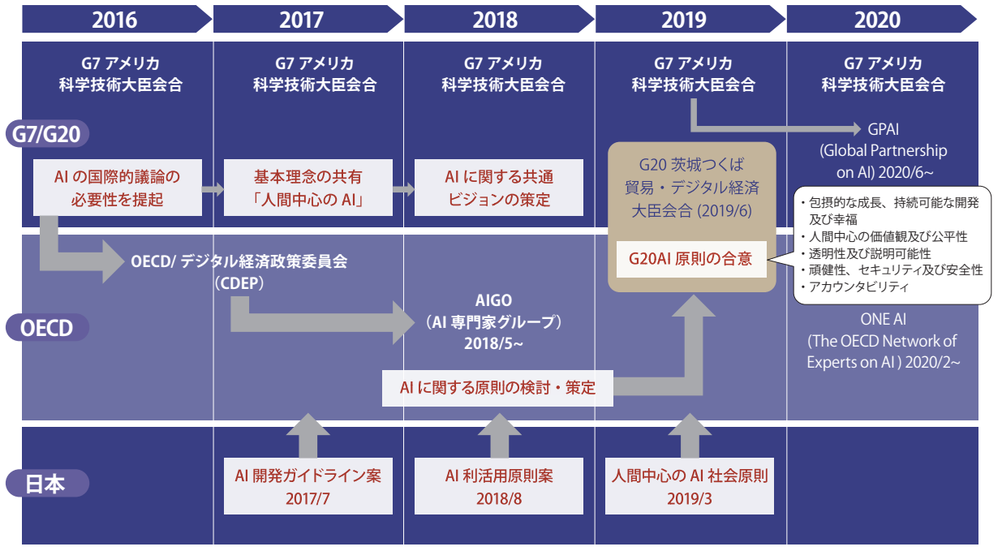
日本国内のこうした活動は、世界的な AI 原則の策定にも大きな貢献を果たしてきました。

先に述べたとおり 2016 年 4 月に AI の国際的議論の必要性を提起したことをきっかけにして、G7 および OECD で検討の場が設けられました。日本で策定した原則やガイドラインも会議で紹介され議論が進み、OECD でまとめられた原則が 2019 年 6 月の G20 の AI 原則の合意につながっています (図 4)。

この OECD の原則では「包摂的な成長、持続可能な開発及び幸福」、「人間中心の価値観及び公平性」、「透明性及び説明可能性」、「頑健性、セキュリティ及び安全性」、「アカウンタビリティ」に関する 5 つの原則が述べられています。

この結果、2019 年時点で世界各国での AI 原則レベルの対応は完了し、その後は企業における具体的なガバナンス対応や技術的対応に関する検討が進められている状況となっています。

図 4　AI 原則の策定における日本の活動



以上のように、2019 年頃までは「ソフトロー」のレベルで進んできましたが、最近は欧米を中心に法律で規制する「ハードロー」な活動が活発になっています。

2019 年の後半ごろより、米国では、AI を国の安全保障に関わることと位置づけ、政府としての取り組みが強化されています。バイデン大統領も就任直後から AI にはルールが必要なことを明言しており、国として 2025 年までに体制が異なる国と競争する準備を整えるように取り組みが進められています。さらに欧州は既に一歩先に進んでおり、AI を具体的に規制する法律の案が 2021 年 4 月に公開され、審議が進められています。最速では 2022 年後半に成立、2024 年後半の施行を目指しています（図 5）。これまで日本は「ソフトロー」で来ましたが、もし欧米で法制化されれば、これまでの J-SOX 法や外為法の対応から、欧米に追随して法制化する可能性が出てきています。

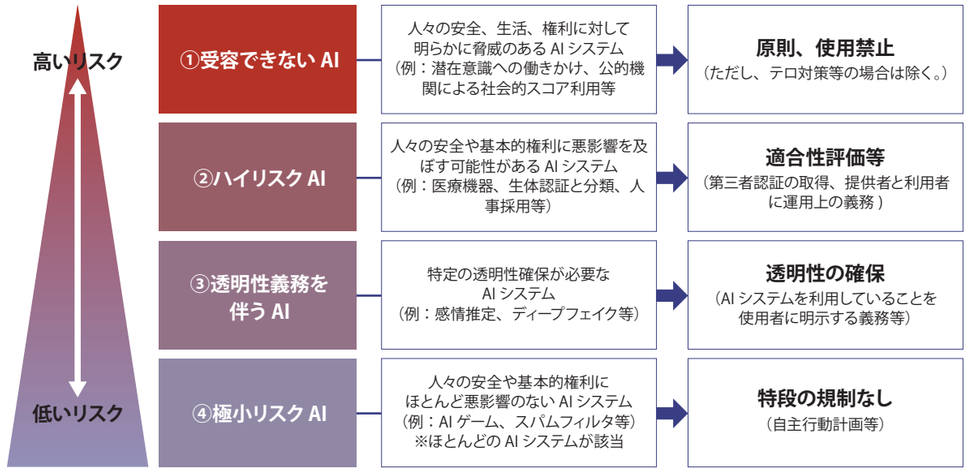
図 5　AI をめぐる欧米の法規制動向

	～2020	2021	2022～25
米国	大統領令　“The American AI Initiative” 国家 AI 研究開発戦略計画改訂版 (2019)（国家科学技術会議、NSTC） 人工知能国家安全保障委員会の四半期ごとの勧告・報告 (2020)	**バイデン大統領、「AI ルールが必要」と明言** 国防権限法 2021：ホワイトハウスが国家 AI イニシアティブオフィスを創設	来るべき AI 時代に国家体制が異なる国と競争する準備を 2025 年までに固める （人工知能国家安全保障委員会　最終レポート (2021) の記載）
EU	信頼できる AI のための倫理ガイドラインを公表 (2019) →域内の企業は社内体制を整備 AI 白書公表 (2020) ※白書：EU 法律の枠組みを示す提案書	**AI 規則案公表 (4/21)** ※規則：EU 全体に直接適用される EU の法律	**AI 規則の成立** （最速で 2022 年後半） **AI 規則の施行** （最速で 2024 年後半から）
日本	【内閣府】人間中心の AI 社会原則 (2019) 【総務省】AI 利活用ガイドライン (2019)	【経産省】AI 原則実践のためのガバナンス・ガイドライン公開	欧米に追随して法制化？ (J-SOX や外為法など同様の法制化例多数)

EU の AI 規則案は「リスクベース・アプローチ」を採用しており、AI システムのリスクを用途・目的等に応じて 4 段階に分類しています。

例えば、人権に明らかなる脅威となるような潜在意識への働きかけや政府機関によるソーシャルスコアリングのような AI は「受容できない AI」として原則、使用禁止、医療機器や生体認証、人事採用等、人権に大きな影響を及ぼす可能性がある AI は「ハイリスク AI」として、利用にあたって第三者認証の取得や、提供者と利用者に運用上の義務を課すようになっています。また、感情推定やディープフェイク等の AI については AI を利用したシステムであることの明示が必須となります（図 6）。

図 6　EU の AI 規則案：リスクベース・アプローチ



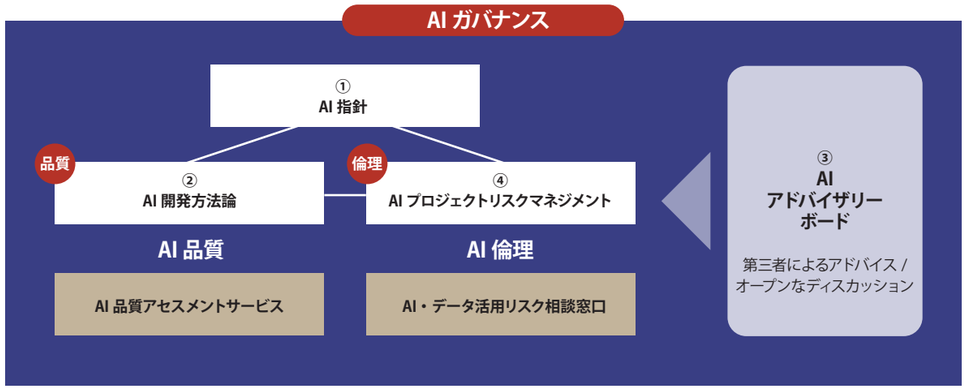
総務省 AI ネットワーク社会推進会議 (第 19 回) 配布資料、経済産業省第 1 回 AI 原則の実践の在り方に関する検討会配布資料より作成

このように数年内に法制度化が完了した場合、グローバルで AI を用いたビジネスをする企業は対応を余儀なくされることになります。こうした動向を背景に、企業が自主的に AI ガバナンスを整備し対応力を付けることは AI ビジネスの競争力に差異を生む状況となっています。

CHAPTER.3
安心・安全で信頼性のある AI の社会実装に必要な「AI ガバナンス」

以上を背景に、NTT データでも、AI 規制を含むリスクに対応した安心・信頼できる AI システムを提供することを目的に、「AI ガバナンス」の仕組みと技術の整備を進めています。
NTT データにおける AI ガバナンスとは、「社会とお客様が AI によるベネフィットを最大限享受できるように、社会規範（倫理）や法制度の遵守および、AI に対する社会の理解を高める活動の推進により、公正で信頼できる AI ソリューションを開発・運用していく仕組み」と定義しています。
具体的には、「①AI 指針の策定」、AI 指針の実行を支える「②AI 開発方法論の整備」と実効的運用、外部有識者を交えた「③AI アドバイザリーボードの運営」、AI 適用におけるリスクを洗い出し対策を講じる、「④AI プロジェクトリスクマネジメントの仕組みづくり」の 4 点を推進しています（図 7）。

図 7　AI ガバナンス全体像



①AI 指針の策定
AI を単なる効率性確保の手段として利用するのではなく、個人・ビジネス・社会が AI のメリットを享受できる「人間と AI が共生する社会」を目指すために、5 か条から成る NTT データグループの AI 指針を 2019 年 5 月に公開しています。
https://www.nttdata.com/jp/ja/news/release/2019/052900/
AI 指針は、当社が AI に関する開発、運用、利活用、判断をする上で参考とするガイドであり、国内外の AI 原則や SDGs の理念、ＮＴＴデータの Group Vision を考慮して策定しました（図 8）。

②AI 開発方法論の整備
AI 指針だけでは AI 開発における具体的な課題に対応するのは困難です。そこで実際のシステム開発や運用シーンで AI 指針に則った活動を可能とする「AI 開発方法論」を 2020 年度に整備しました。AI 開発方法論は、300 以上の AI システム開発の知見に基づいて策定しました。AI 開発のナレッジを集約した【開発／管理プロセス】、AI 開発者が手を動かすために必要な成果物ひな形等のドキュメントをまとめた【開発標準】、AI 品質における特性／リスク観点を問診票形式でチェックできる【アセスメントツール】から構成されています（図9）。これに加え、個別プロジェクトを総合的にサポートする【品質コンサルティングサービス】を用意しています。

図 8　NTT DATA の AI 指針

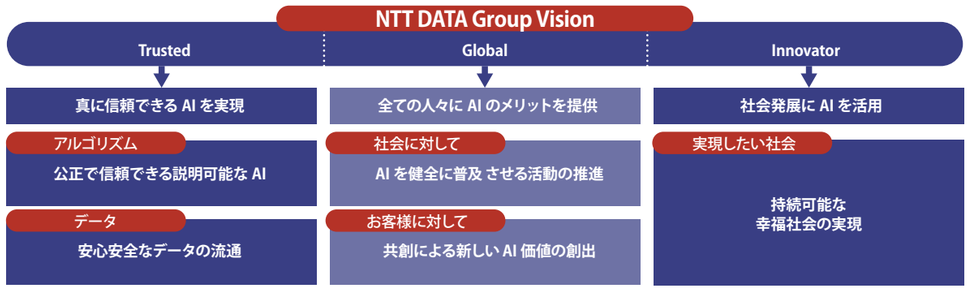
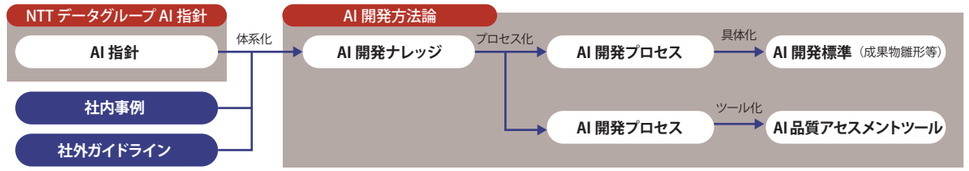
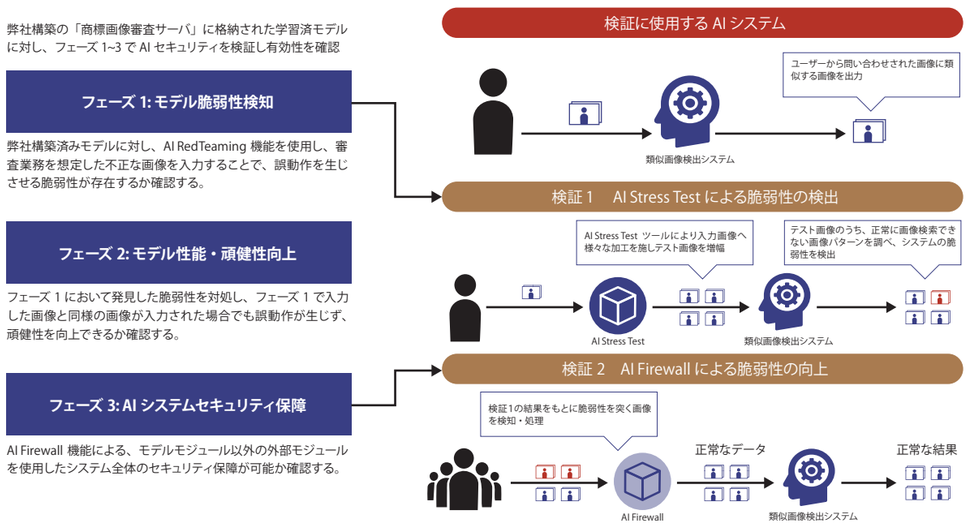


図 9　AI 開発方法論



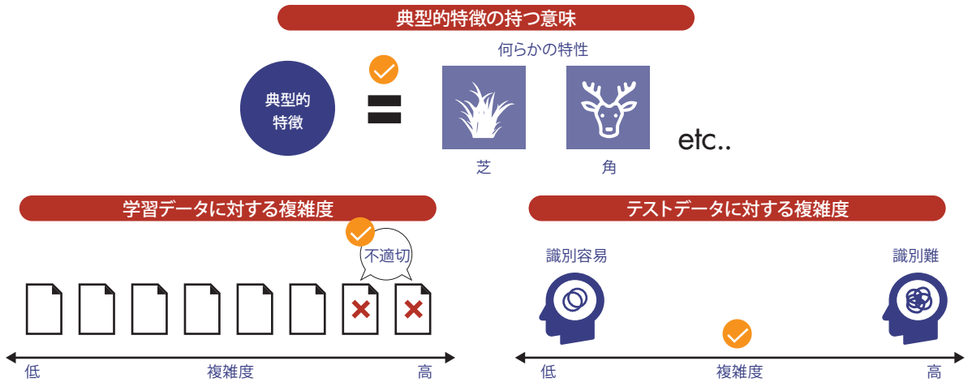
また、AI 開発方法論を実効面で支えるための技術を、NTT グループのみならず、お客様や学術機関、パートナー企業と共同で開発しています。最近の新たな取り組みとしては、「AI が誤動作する可能性を検出するセキュリティ技術」や、「AI が問題の無いデータを学習して作られているかの検証手法」の開発があります。
「AI が誤動作する可能性を検出するセキュリティ技術」は、スタートアップ企業の Robust Intelligence（以下、RI 社）と AI セキュリティに関する技術検証を行っています。この技術検証では、NTT データがディープラーニングを用いて構築した「類似商標調査における類似画像検索モデル（以下、類似画像検索モデル）」を対象に、RIME(Robust Intelligence Model Engine) による脆弱性の検知とその対処を目的とした取り組みを実施しました（図 10）。

図 10　AI のセキュリティ向上に関する取り組み



「AI が問題の無いデータを学習して作られているかの検証手法」は、学習済 AI モデルがデータをどのように扱うかを解析し、そこから学習済 AI モデルが得意とするデータ、苦手とするデータの傾向を把握する技術を開発しました (図11)。
具体的には、モデルが獲得した典型的な特徴を NMF(Non-negative Matrix Factorization：非 負 値 行列因子分解) と呼ばれる数学的な手法により抽出し、典型的な特徴と比較して計算した「データ複雑度」により、学習やテストに相応しくないデータを検出することを可能にしました。

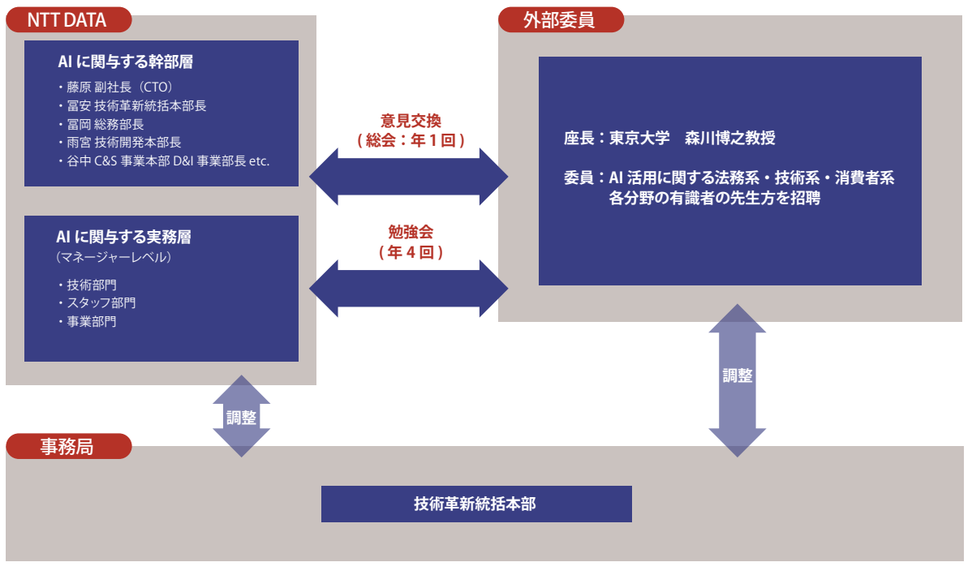
図 11 **AI の評価方法に関する研究開発の取り組み**



③AI アドバイザリーボードの運営

続いて 2021 年 4 月、AI ガバナンスの強化に向けた意見交換を行う場として、「AI アドバイザリーボード」を設置しました。倫理や法務、ソフトウェア品質など各分野の専門家から最新トレンドを学び、どのように当社の AI ガバナンスに反映していくのが良いか、現場のマネージャーレベルまで参加する勉強会を通じて議論を実施しています (図12)。

図 12 **AI アドバイザリーボードの実施体制**



こうした取り組みは透明性が大事なことから、開催の模様は外部に公開しています。
2021 年 4 月 19 日に開催したキックオフ（第 1 回総会）では、専門委員の先生方と幹部層が、現状の取り組み状況と目指すべき状態について、多様な観点で意見交換を実施しました
(https://www.nttdata.com/jp/ja/data-insight/2021/0525/)。
また、2021 年 12 月末までに 3 回の勉強会を開催し、委員の先生の専門分野から AI ガバナンスに関する最新トピックを講演いただき、当社からも関連する取り組みを紹介した上で、ディスカッションを行っています。
第 1 回勉強会模様
https://www.nttdata.com/jp/ja/data-insight/2021/0819/
第 2 回勉強会模様
https://www.nttdata.com/jp/ja/data-insight/2021/1116/
アドバイザリーボードで得た知見を当社の AI ガバナンスに反映させていきます。

④AI プロジェクトリスクマネジメントの仕組みづくり

そして現在、AI プロジェクトリスクマネジメントの仕組みづくりを進めています。具体的には AI アドバイザリーボードでの議論も踏まえて、AI 適用における法制度への対応や倫理・社会受容性の観点から AI プロジェクトのリスクチェックを行い、そのリスクの度合いに応じて必要な対策を講じる「リスクベース・アプローチ」に基づいた仕組みを検討しています。

CHAPTER.4
AI ガバナンスがブランド価値になる時代へ

数年後に AI の法規制が現実に適用されることが視野に入ってきたこともあり、安心・信頼できる AI を提供可能な能力や体制を持つことが、近い将来に大きなブランド価値に繋がることが明確になってきています。NTT データでは、AI を使ったシステム開発・運用を対象に、AI ガバナンスに対する社員の意識を高め、倫理や法務面も含めて問題化を事前に回避できるよう、チェック体制や社内ルールの整備に取り組んでいきます。