

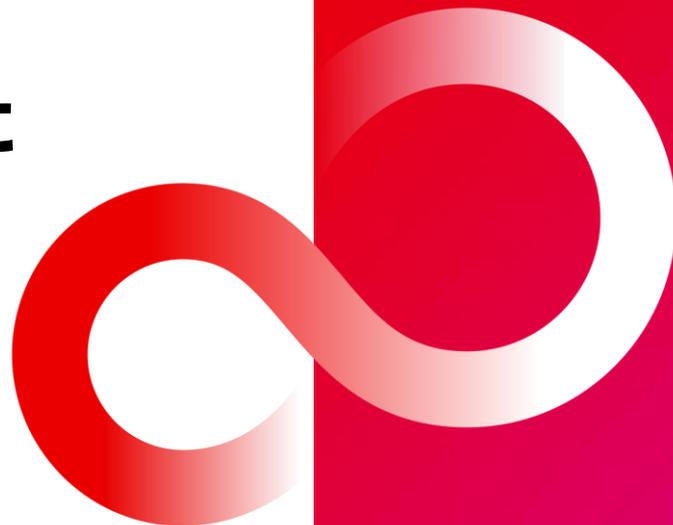
安心安全なAIの提供に向けた ガバナンスの取り組み

2023年10月31日

富士通株式会社

ジャパン・グローバルゲートウェイ

浦下 瞬



富士通のAIへの取り組み

本社
東京

創立
1935

代表取締役社長 CEO
時田隆仁

事業概要
サービスソリューション
ハードウェアソリューション
ユビキタスソリューション
デバイスソリューション

(注) 1. 2023年3月末現在

(注) 2. 2022年度連結概要 (2023年3月31日終了会計年度)

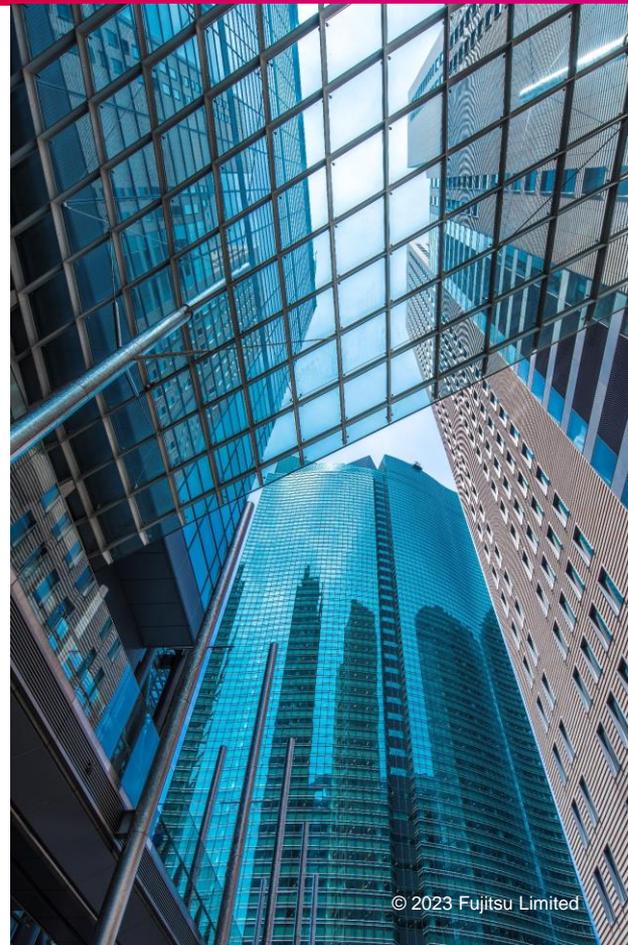
従業員
124,000 人

売上収益
3兆7,137億円

営業利益
3,356億円

研究開発費
1,095億円
売上収益比3.0%

上場証券取引所
東京 (コード 6702)
名古屋



わたしたちのパーパスは、
イノベーションによって社会に
信頼をもたらし、世界をより
持続可能にしていくことです。



富士通の重点技術領域

- 富士通は5つの重点技術領域の研究開発にリソースを集中するとともに、幅広いパートナーとのコラボレーションを推進
- Computing、Network、Data & Security、Converging TechnologiesとAIの融合を強化

5つのテクノロジーメガトレンド

創造性の拡張

AI



- セマンティックグラフAI (セマンティックグラフ、グラフAI)
- 説明可能なAI
- 発見するAI
- Actlyzer

つながり、誰も取り残されない

Network × AI



- 6G技術
- ディスアグリゲータードコンピューティング
- 光伝送/光電融合技術



- インテリジェント・ネットワーク (ネットワーク制御AI)

量子スピードの創造

Computing × AI



- HPC (高性能プロセッサ)
- 量子コンピュータ
- Computing Workload Broker



- 発見するAI
- マテリアルズ・インフォマティクス

未来のリ・デザイン

Converging Tech. × AI



- フェデレーテッド・デジタルツイン
- ソーシャルデジタルツイン
- エンタープライズメタバース



- マルチアスペクトシミュレーション
- 行動予測技術

進化するウェブ

Data & Security × AI



- Web3/ブロックチェーン
- IDYX, CDL, 透過的トラスト
- トラストネットワーク



- AIセキュリティ/トラスト
- 常時認証

富士通の重点技術領域



Computing



Network



AI
5



Data & Security



Converging Technologies

AI倫理



安心安全なAI活用のための技術・活動

- 欧州最大の倫理団体AI 4 Peopleに設立当初から参画し、政策提言に貢献。主要メンバーとして実践と標準化を推進
- AIの倫理的リスクを網羅的に抽出するリスクベース評価技術を開発し、世界に向けて無償公開

AI4People

ミュンヘン工科大学

品質



国際基準に沿ったAI品質評価技術と社内適用

- 国際標準化機構ISO/IECにおけるAIの安全・安心な利用のための標準化提案・策定をリード
- 国際規格に準拠した社内標準ガイドラインの作成と全社品質管理プロセスの適用

ISO/IEC JTC 1/SC 42

セキュリティ



AIに対するサイバー攻撃の自動監視技術の開発

- 敵対的パッチ攻撃による不正を検知する技術を開発、北米NRFに出展、小売領域に展開
- AIモデルの再学習を必要としないため導入が容易で、初見のパッチによる不正の検知もできる独自技術（ベングリオン大学。共同研究）

ベングリオン大学

富士通のAIガバナンスに向けた取り組み (トップダウン)

富士通グループ AIコミットメント

PARTNERSHIP ON AI	Future of Life Asilomar AI principles	IEEE EAD	モントリオール宣言	Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems	英国議会 AI in the UK

 **AI4People***
5 principals

*欧州の国際団体

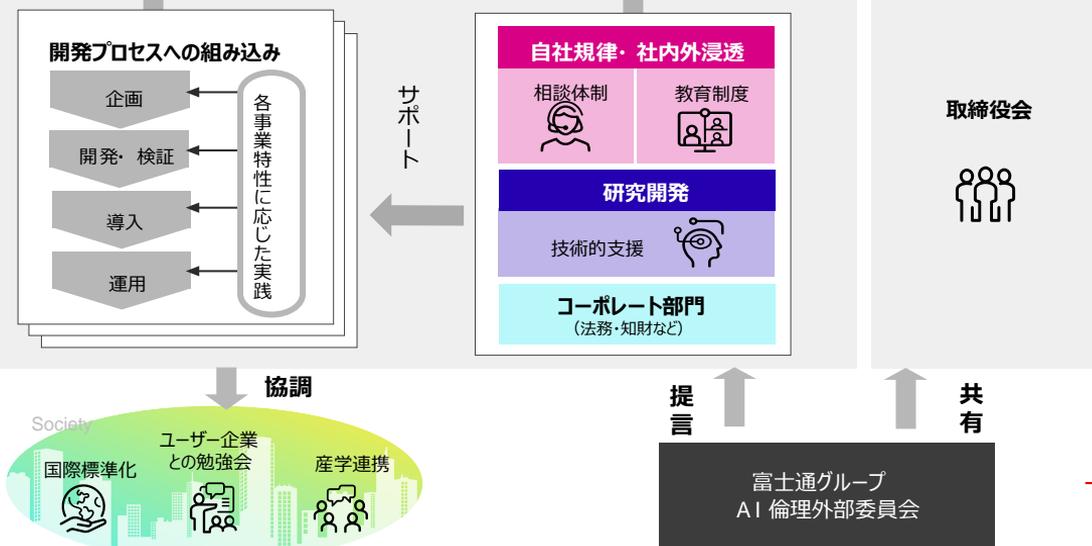
富士通グループAIコミットメントは AI4people の5つの原則に基づき客観性を確保



- 1 AIによってお客様と社会に価値を提供します
- 2 人を中心に考えたAIを目指します
- 3 AIで持続可能な社会を目指します
- 4 人の意思決定を尊重し支援するAIを目指します
- 5 企業の社会的責任として、AIの透明性と説明責任を重視します

富士通グループAIコミットメントに基づき全社的かつ総合的な取り組みを実践

富士通グループAIコミットメント



富士通グループAI倫理外部委員会

多様な分野のハイレベルな識見を持った外部専門家を招き、実践的な提言をいただいている



本委員会のミッション



- 01: 当社のAI倫理指針・ガバナンスに対する評価**
当社のAI倫理に関する指針や戦略について、法制度や社会の状況を踏まえ、技術・人権・サステナビリティ等の多様な観点のもと評価を行い、今後の方向性を検討します。
- 02: 当社のAIビジネス推進およびブランド向上への寄与**
「信頼されるAI」の提供に向け、客観的視点で評価・提案を行い、当社の経営層に共有しています。また、本取組みを社会の幅広いステークホルダーにも発信し、当社AI倫理ブランドの認知向上に寄与することを目指します。

富士通のAIガバナンスに向けた活動 (ボトムアップ)

AI特有事項についての考慮が現行の富士通の品質保証基準やセキュリティ基準だけでは不適當/不十分な可能性

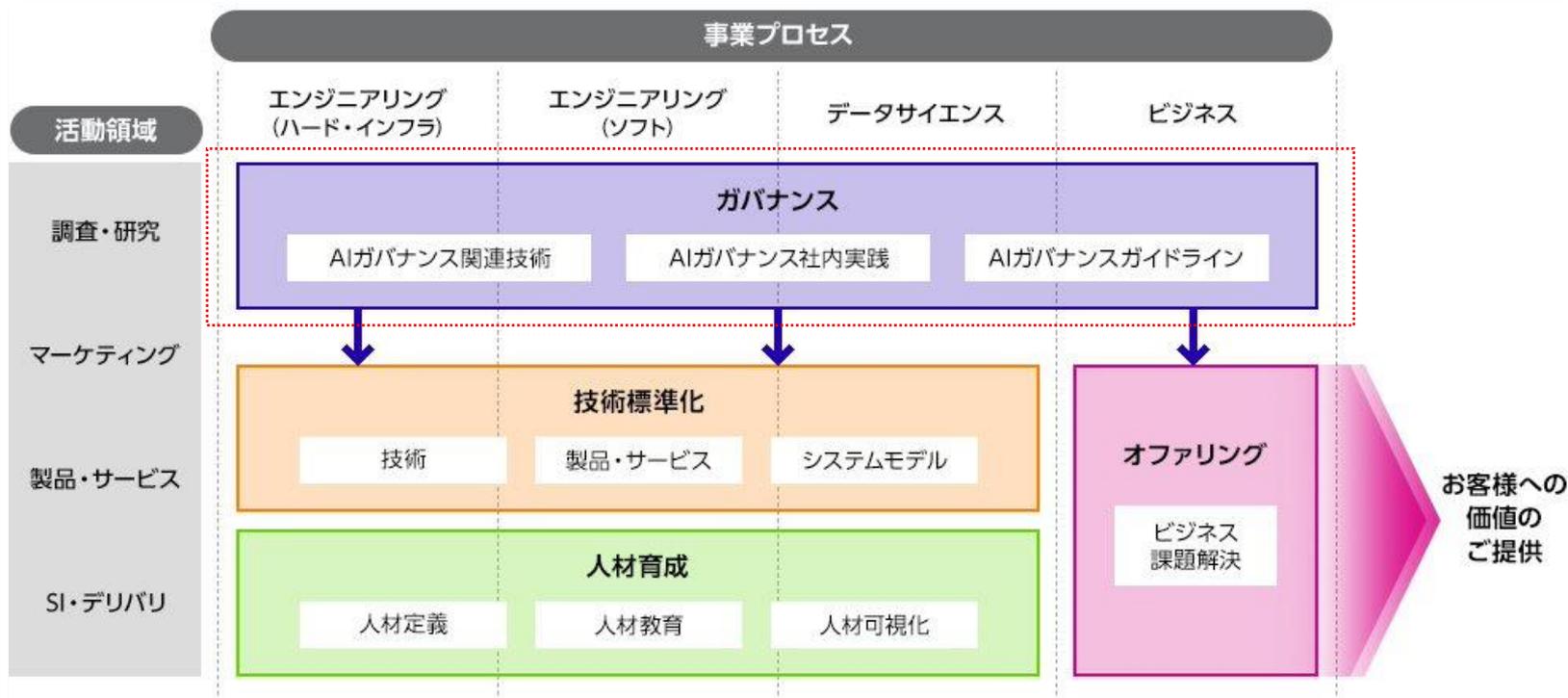
- AI倫理・AI品質・AIセキュリティの考え方、
知財の検討範囲の拡大、個人情報保護法、各国法規制対応など

AI向けの様々なガイドラインがあるが、ビジネスプロデューサー/SEが全て把握・対応するのは困難

- 倫理/品質/契約などの分野別
- 読み応えあるボリューム
- 抽象的記述
- 具体的な手法やアクションが不明確

社内	デジテック法務：データを利用したビジネスを安全・確実に行うための契約上のポイントについて AI倫理ガバナンス室：AI倫理チェック（AIリスク判定シート/自己チェックシート） 研究所：AIセキュリティホワイトペーパー 研究所：AI品質早期アセスメントガイド
社外	総務省：AI活用ガイドライン 総務省：国際的な議論のためのAI開発ガイドライン案 経産省：AI原則実践のためのガバナンス・ガイドライン 経産省：AI・データの利用に関する契約ガイドライン 経産省：AI・データサイエンス人材育成に向けたデータ提供に関する実務ガイドブック 産総研：機械学習品質マネジメントガイドライン QA4AI：AIプロダクト品質保証ガイドライン
国際動向・標準化	European Commission：EU AI法案 Proposal for a Regulation laying down harmonised rules on artificial intelligence Ethics guidelines for trustworthy AI FTC：企業向けAI利用ガイド NIST：AIリスクマネジメントフレームワーク ISO/IEC：Dis38507(AIガバナンス)、WD42001(AIマネジメントシステム標準) データのトラストとデータ共有合意の国際標準化（TR23186、DIS23751）

あらゆるお客様に向けて、AI活用（デリバリ）を担える富士通であり続けるため、
富士通全体でAIのデリバリプロセスを確立していく組織横断活動



AIの各デリバリプロセスで必要なリスクや対策を観点横断的に洗い出し取りまとめ提供

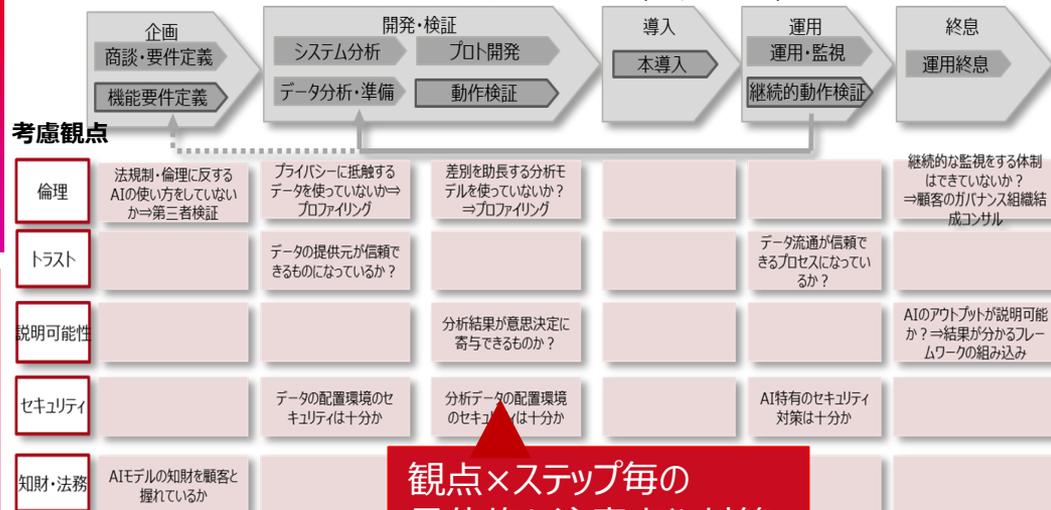
観点横断でまとめてチェック

観点横断でまとめてチェックできる
リスクに対して対策がとれる

安心安全なAIの提供

安心安全なAIの提供ができる
AIコミットメントの達成につながる

AIのデリバリプロセス（当社実践知と国際標準プロセス（ISO/IEC 5338）をあわせて独自に作成）



観点×ステップ毎の
具体的な注意点や対策

※初期構想イメージ

当社専門家集団の知見を集約

一度きりでなく、最新の技術や動向を踏まえた継続メンテナンス性を担保



産業技術総合研究所の「機械学習品質マネジメントガイドライン 第2版」をベースにAI品質の評価を行う手法を開発

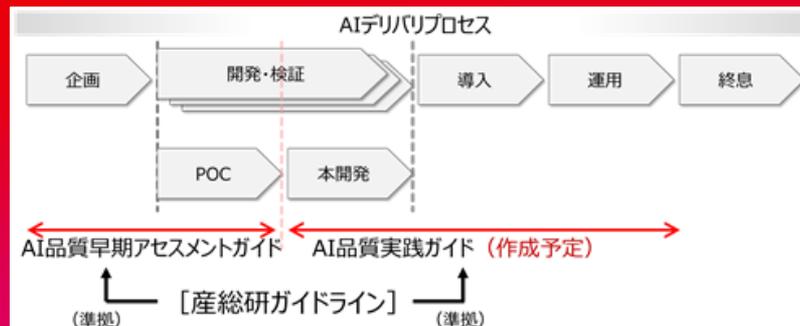
当社人工知能研究部門の専門家が抽出・具体化

- 200ページを超えるボリュームから各段階で必要な要素を抽出
- 様々なAIシステムを評価対象とする汎用的評価指標から、当社AIシステム開発で用いるための具体的な指標に落とし込み
- 活用ノウハウを組み込み

1 | エッセンスを抽出した活用ガイドライン

- AI品質早期アセスメントガイド
- AI品質実践ガイド

2 | 容易利用可能なExcel版チェックシート



基本的な考え方を説明 初学者向けの導入

概説

AI TASK FORCE / ガバナンス

A I / データ利活用におけるガバナンス

便利さだけでなく、安心・安全な利用を実現するために

情報通信技術の進歩、とりわけ人工知能（AI）の急速な進歩は、人々の生活やとりもその一方で、不当な差別など、AIが思いもよらない副作用をもたらす可能性も指摘富士通グループは、AIを開発・提供する者の責務として、便利さだけでなく、安心・出典）富士通グループAIコミットメント

企業がより多くの意思決定を人工知能（AI）に委ねるほど、契約、倫理、プライバシーこととなります。

このため、「安心・安全で信頼性のあるAIの社会実装」の推進に向けて、各国政府・一方で、各種ガイドラインは様々なケースに対応可能とするため、人・組織による異なることが想定されます。

リスクへの対応は、人・組織が個々に対応するのではなく、同じ解釈の元、ステーク実施しなければ、対応にバラツキが生じ「安心・信頼できるAI/データ利活用を提供

このような背景から、AIデリバリプロセスのフェーズ毎に人工知能（AI）特有のガ対応・対策をガイドラインとして整備しています。

国内外においてデータを戦略資源と位置づけDXを推進しています。

企業内のデータを一元管理したい、データから価値を生むためにAIを活用したいとのを提供できる」状態とすることは急務であると考えています。

確認・実施すべき事項を具体的な対策や相談先とともに提供し、実施状況のチェックやリスク把握可能化

実施事項集

項番	1204
観点	品質管理
条件	AI システムの開発を行う場合
主体	AI システム企画・提案者
プロジェクト形態 1	富士通
プロジェクト形態 2	富士通
客体	－（自身で確認）
プロジェクト形態 1	－
プロジェクト形態 2	－
必須 or 推奨	推奨
実施すべきこと	機械学習要素に対する品質要求レベルを特定する。
具体的な対策	AI 品質早期アセスメントガイド（「外部品質要求レベルの特定」）を参照して、AI システムに求められる外部品質の要求レベルを、「リスク回避性(AISL)」「AI パフォーマンス(AIPL)」「公平性(AIFL)」の3つの観点から求める。 ここで要求レベルとは、AI システムに求められる厳密さを表すレベルである。例えば AISL は 0.0/1.0/2.1/2/3/4 の 7 段階で設定されるレベル 0～0.2 は弱い品質レベルであり、仮に AI が誤っても取返しが可能なレベル。1～4 は高い確実性が求められるレベルである。 なお、ここで判定するのは外部品質の要求レベルであり、外部品質そのものを求めるわけではない。特に「公平性」については、他項

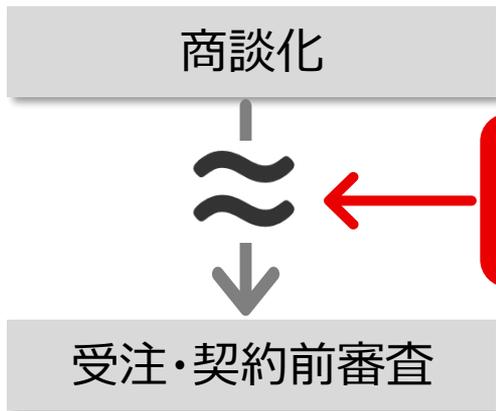
チェックリスト

<p>知財（特許、商標、意匠、著作権等）の創出が想定される場合、各当事者が単独で生み出した知財及び共同で生み出した知財について、それぞれ取り扱いは、</p> <p>1) 各当事者が単独で生み出した知財については、それぞれ生み出した当事者に単独で帰属する 2) 共同で生み出した知財は両者に帰属し、お互い事前の通知や合意の無い自由には使用</p> <p>相談先：知的財産イノベーションセンター https://fujitsu.sharepoint.com/sites/ai-iou/innovation/SitePages/ModernHome.aspx</p> <p>現在のビジネス以外での利活用を目的として、取引先とデータの授受を行う場合（AIは関連ビジネスに限らない）データに関連する知的財産その他の権利関係を明確化し、関係者と合意形成する。</p> <p>現在のビジネスに必要な範囲でデータを使用することは、取引の前提として了解されるが、それ以外の利活用には特に対応が必要となる。そこで、知的財産の対象となるデータについてはその権利が誰に帰属するかや、知的財産権の対象になるか否かを問わずデータ全般（そのデータを利用したAI/分析/加工して生成されるデータ含む）を誰がどのように利用できるかについて、関係者と合意する。 ※主な検討事項は、利用主体、目的、期間、地域、独占/非独占、改変の可否、対価の有無、保証責任など、その際、将来的な社内/他社/シブへの展開の可能性も考慮する。 ※相談先：デジタルテクノロジー推進法務室 https://fujitsu.sharepoint.com/sites/ai-legal/dn/SitePages/Ton_Pane/Ton_Pane.aspx</p> <p>個人に関する情報や、プライバシー上の懸念があるデータを取り扱う場合、個人やプライバシーに関する情報の取扱いの適法性を確認する。</p> <p>①法令の遵守：テキストや映像などの形式を問わず、個人情報保護法（外国の国で授受する場合は外国法も検討）に定められた個人情報や関連情報といった類型に該当する場合、本人への利用目的の説明・第三者提供の同意取得といった所定の手続きを行う。 ②社会的受容性の考慮：個人情報やプライバシー情報については、法令の遵守のほか、社会において受け入れられるか（人々の懸念を覚えたり、反応しなかったり）を考慮する。リスク削減策として、対象者への事前告知・同意取得を行ったり、懸念が深刻な場合はそもそもデータ取得しないといった対応を検討し、実施する。 ※他者経由で取得した情報については、当該他者のどちらが取扱いの主体となるか、本人の対応を行うかを取扱いの決りかたで対応する。 ※相談先：デジタルテクノロジー推進法務室 https://fujitsu.sharepoint.com/sites/ai-legal/dn/SitePages/Ton_Pane/Ton_Pane.aspx</p> <p>外国との間でデータの提供や授受を行う場合、提供元と提供先において、データを適法に移転できるか確認する（個人に関する情報やプライバシー以外）。</p>			
--	--	--	--

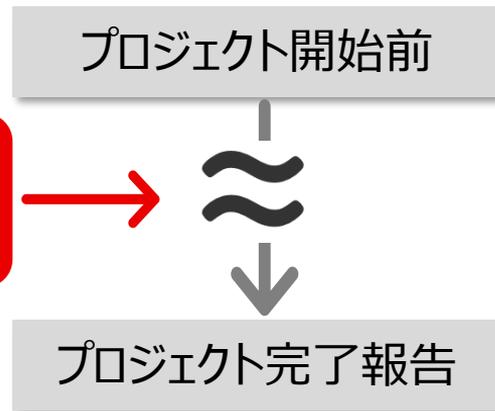
ノウハウキュメントの提供だけでなく商談・デリバリのプロセスに組み込むことで実効性を確保

富士通の商談・システム提供時の品質保証プロセスへ組み込み、
実施基準に基づきセルフチェック・統制部門のレビューを実施

商談フェーズ



提供フェーズ



課題であったAI特有の複数観点に渡る考慮事項を横断的にチェック可能

- 十数ガイドライン4桁ページの要素を具体化したうえで
商談時・開発提供時それぞれ20件程度のチェック項目に集約、1日でチェック可能化

商談・システム提供時の品質保証プロセスにチェック&レビュー組み込み、 安心安全なデリバリを実現

- 最低レベルでもセルフチェック必須化（約200件/月（AI以外も含む））
- 高レベル以上の2件についてレビューを実施しデータ利用許諾の考慮などアドバイス実施

本活動

ガバナンスガイドライン

チェックリスト

ガバナンスレビュー



安心安全なAI



イノベーションを阻害しないリスクベースアプローチの採用

リスクベースアプローチ リスク評価からリスク軽減までをカバー

1. **リスクレベル判定**：AIシステムのユースケースにおけるリスクレベルの判定
2. **リスクアセスメント**：AIリスクの特定、リスク軽減計画の評価と提案
3. **リスク対応**：リスク軽減計画を選択して適用



細かな行為義務を規定するルールベースの規制がイノベーションを阻害しうることも同時に指摘されている。これらの問題を克服するためには、従来のルールベース型から、最終的に達成されるべき価値へと企業等を導くゴールベース型へと、ガバナンスの構造を変革することが求められている。

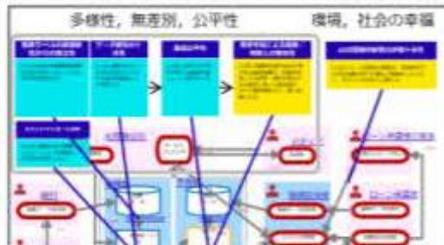
我が国の AI ガバナンスの在り方 ver.

1.1：経済産業省

https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_1.pdf

- ・ グローバルでの全社展開に向けた制度の確立
- ・ 生成AIへの対応

イノベーションを阻害しないリスクベースアプローチの採用



当社研究所でリスク管理技術開発、チェックへの組み込み



義務を規定するルールベースバージョンを阻害しうることもされている。これらの問題を克服するには、従来のルールベース型から、最終的に達成されるべき価値へと企業等々を導くゴールベース型へと、ガバナンスの構造を変革することが求められている。

我が国の AI ガバナンスの在り方 ver. 1.1 : 経済産業省
https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_1.pdf

- ・ グローバルでの全社展開に向けた制度の確立
- ・ 生成AIへの対応

わたしたちのパーパスは、
イノベーションによって社会に
信頼をもたらし、世界をより
持続可能にしていくことです。

Thank you

