



NTTデータグループ AIガバナンスの取り組み 2021 update

2021年2月10日
株式会社NTTデータ 技術開発本部長
雨宮 俊一

「安心・安全で信頼性のあるA I の社会実装」に向けて

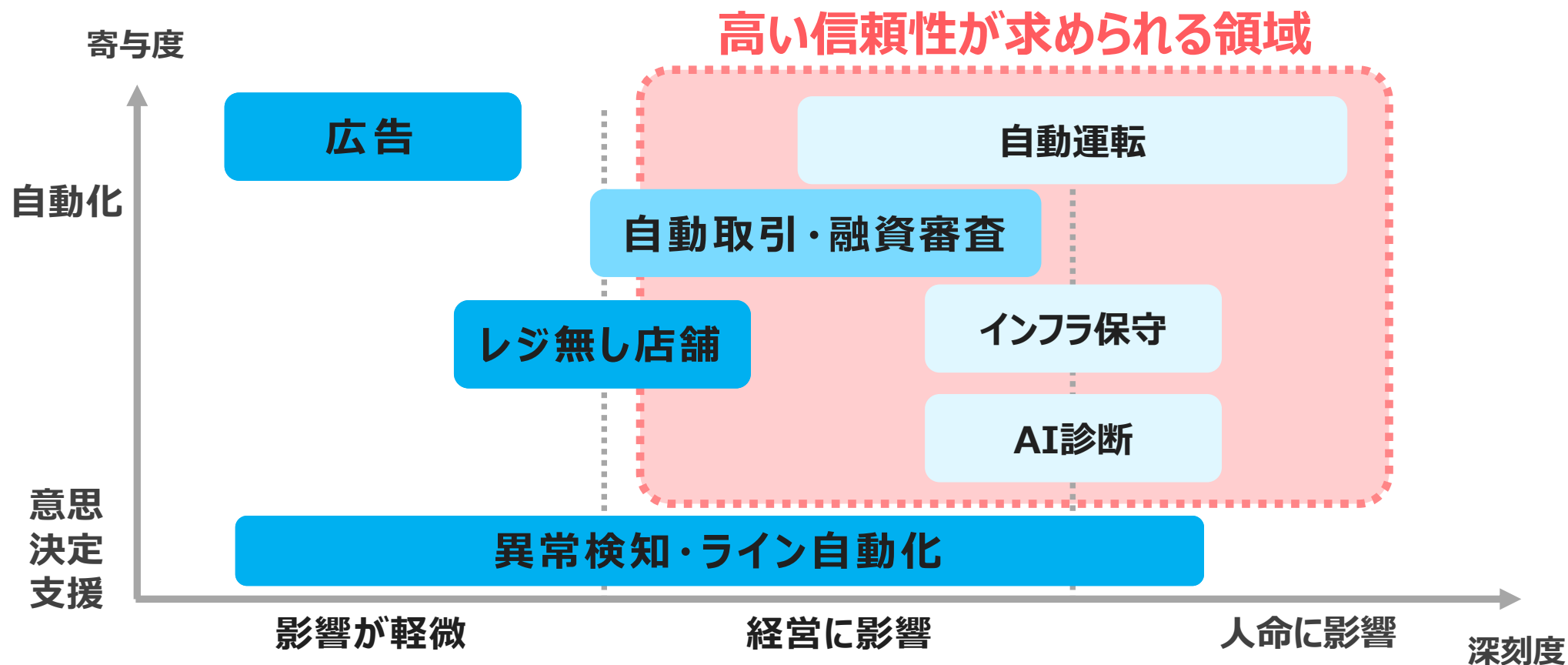
■NTTデータの取り組み概況

2019年	2020年	2021年
5月 NTTデータグループのAI 指針を策定	6月 「AI開発プロセス」を体系的に整理し、 AI案件への試行適用を開始 10月 信頼できるAI活用に向けた AI品質アセスメントサービスのトライアル提供開始	1月 「ビッグデータ活用基盤リファレ ンスアーキテクチャ」の公開

前回のご紹介（2020年1月）後、
AI開発方法論（開発プロセス／開発標準／品質アセスメントツール）の整備を進めており、
2021年には一部を公開する方向で検討しています。

社会動向：AI活用はより高い信頼性が求められる領域へ

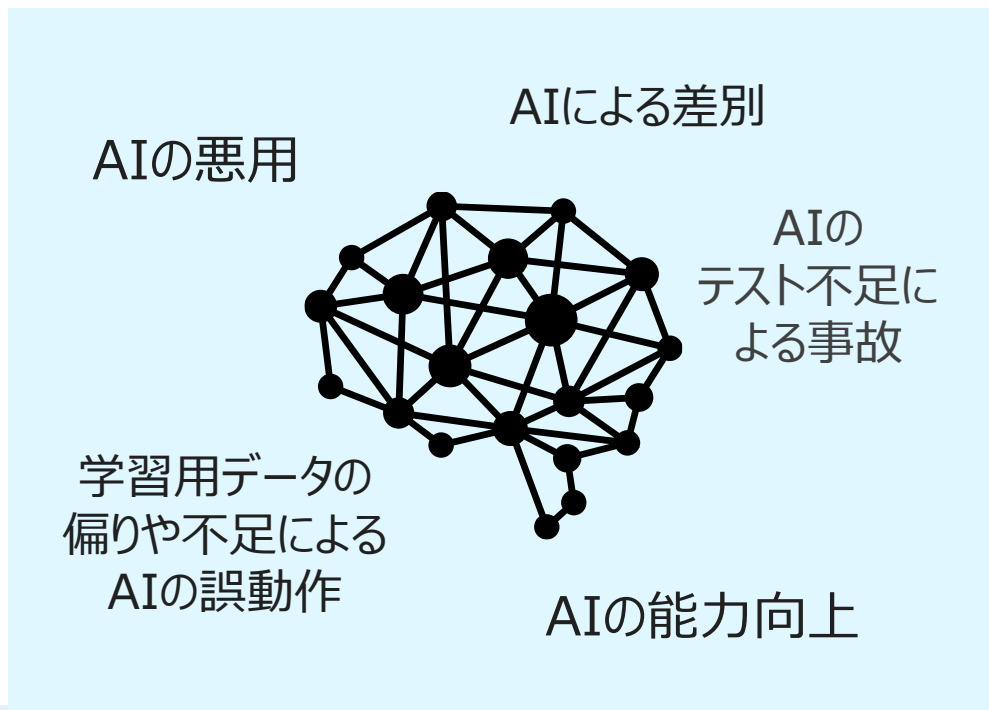
- AIの活用は、デジタル化の進展や人口減少を背景として、「人命や社会インフラに関わる領域」に拡大しています。



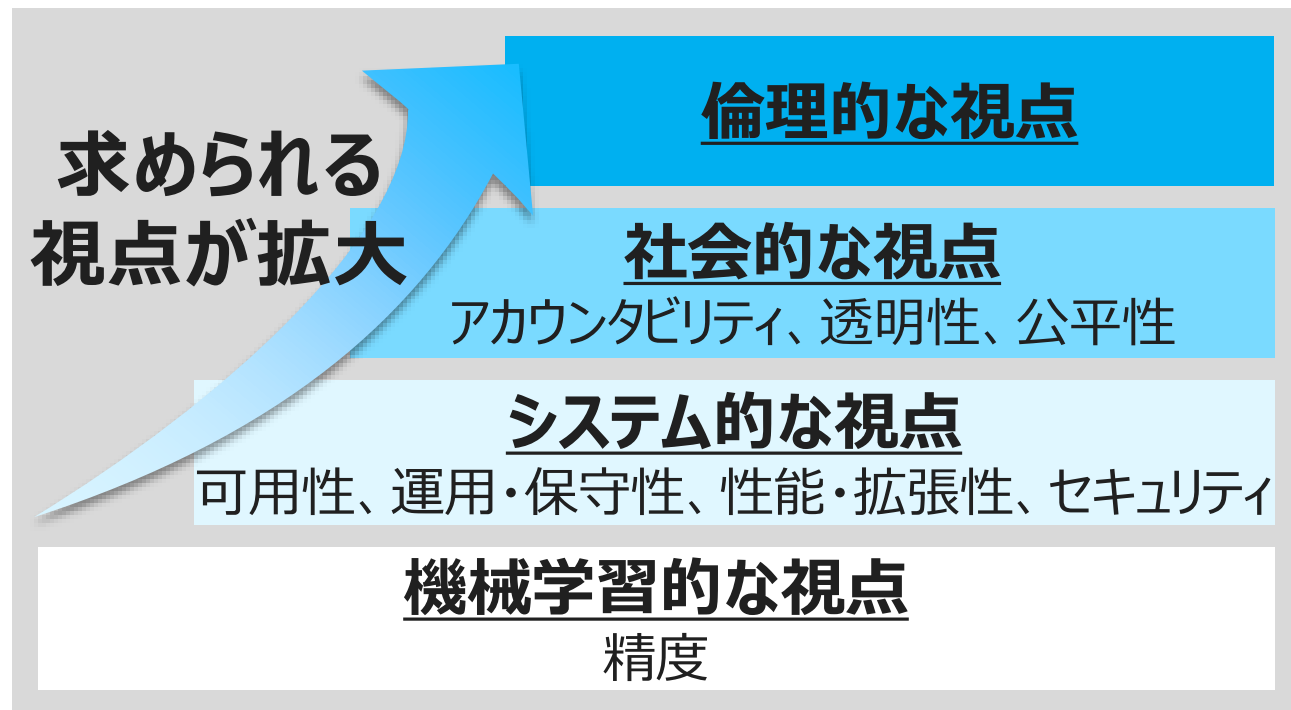
社会動向：AI活用における視点の拡大

- AI活用に求められる視点も、精度だけだった時代から、
システムの、社会的、倫理的なものへと拡大しています。

社会的背景



AI活用に求められる視点



NTTデータグループのAI指針を策定

- 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、AI指針を策定しました。

NTT DATA Group Vision

Trusted

Global

Innovator

真に信頼できる
AIを実現

アルゴリズム

公正で信頼できる
説明可能なAI

データ

安心安全な
データの流通

全ての人々に
AIのメリットを提供

社会に対して

AIを健全に普及
させる活動の推進

お客様に対して

共創による新しい
AI価値の創出

社会発展に
AIを活用

実現したい社会

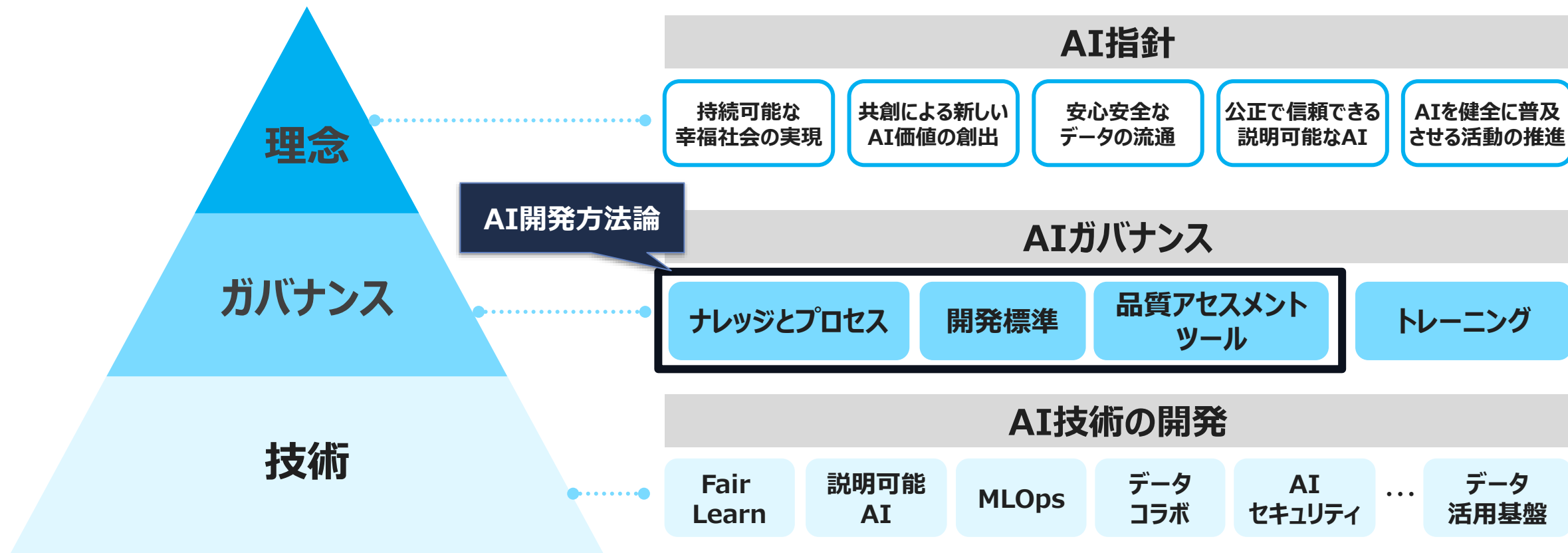
持続可能な
幸福社会の実現



AI指針

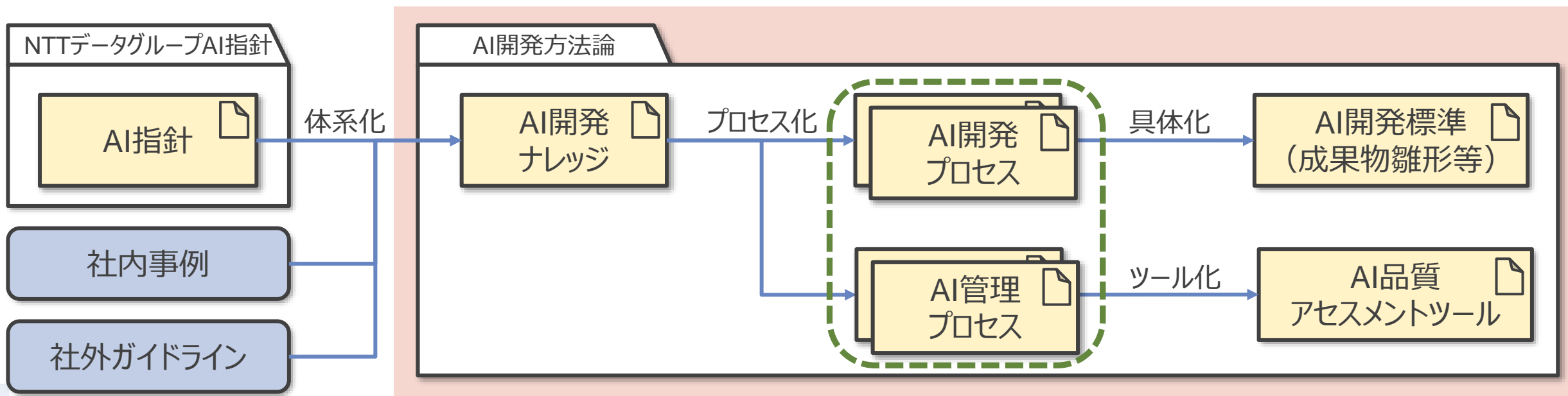
AI指針に続き、AIガバナンスと技術を整備

- AI指針の実行を支えるガバナンスの強化と、適正なAI活用を実現する技術の開発を推進しています。



AI開発方法論の構成

- AI開発のナレッジを集約し、開発プロセスと管理プロセス（管理項目を定義した物）を整備しました。
- 成果物雛形など、AI開発者が手を動かすために必要なドキュメント類は、開発標準として具体化しています。
- AI品質における特性／リスク／チェック観点を問診票形式で提示する「アセスメントツール」も用意しました。



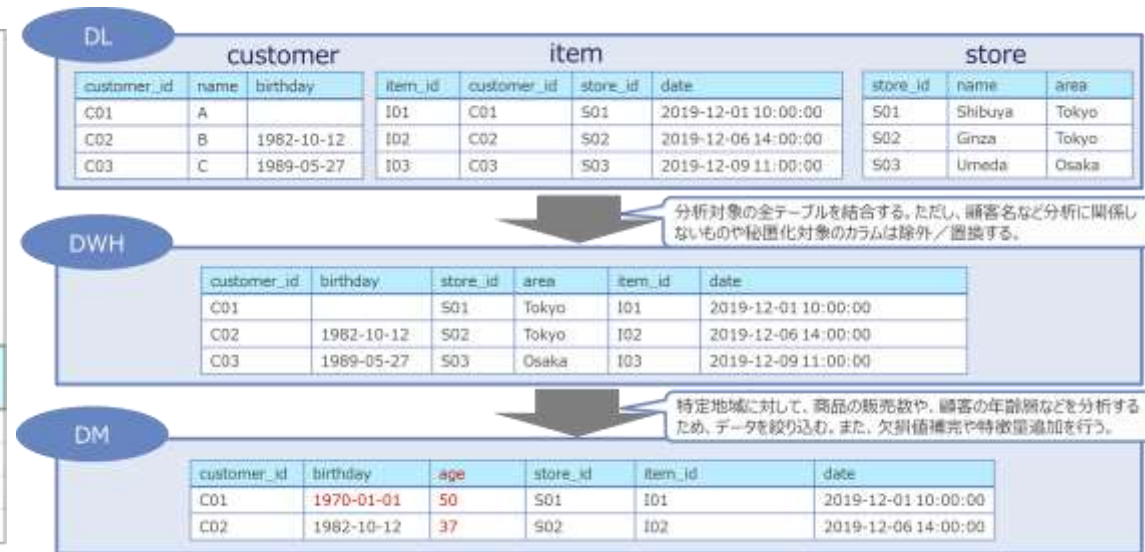
※緑の破線枠内は今後一般公開を検討

AI開発プロセス

- AI開発における工程を定義。
 - AIモデル開発のフロー、データの前処理、AIモデルとデータの一体的管理、運用開始後のAIモデル更新手順を重点的に整理しています。
- 作業漏れ／確認漏れ防止のために、各工程および成果物に責任を持つ役割を明確化しました。
- 各工程で作成すべき成果物ドキュメントのサンプルも用意しています。

AIモデル開発における役割\フェーズ・タスク	設計書詳細化																				
	ビジネス課題抽出	AI活用領域の決定	KPI設定	基礎集計用データ取得	データ基礎集計	基本検討	評価方式検討	データ管理方式設計	データ詳細設計	モデル詳細設計	データ加工詳細設計	データ取得	データ品質調査	データクレンジング	加工処理動作確認	データ分割	モデル構築	モデル学習	モデル評価	ビジネス課題検証	
	ビジネス仮説導入	AI方針検討				AI設計	データ準備	データ加工	AI構築			AI検証									
ビジネスアナリスト	◎	◎	◎	○	○	○	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
データエンジニア	-	-	-	◎	◎	-	◎	◎	-	○	○	○	○	○	○	○	-	-	-	-	○
データサイエンティスト	-	-	-	-	○	◎	◎	○	-	◎	◎	◎	◎	◎	◎	◎	○	○	○	◎	◎
機械学習エンジニア(AI)	-	-	-	-	-	-	-	-	-	○	-	-	-	-	-	-	◎	◎	◎	-	○

役割とタスクの対応表サンプル



データ・モデル管理方式設計サンプル

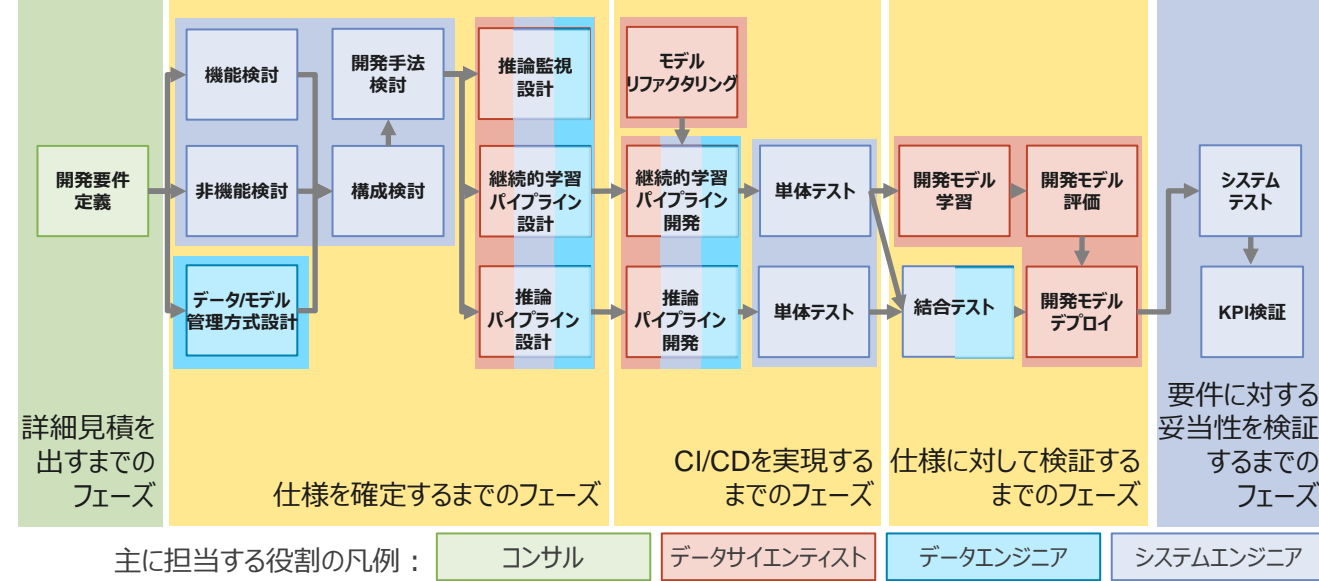
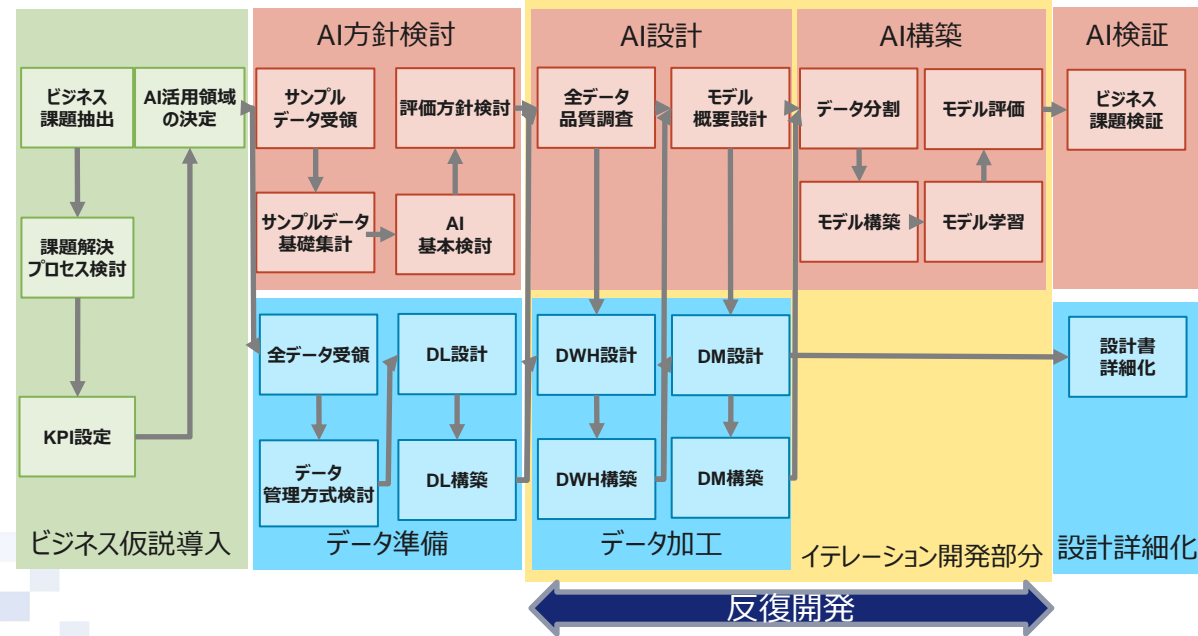
AI管理プロセス

- PMBOKで定義されている管理項目に合わせてAI開発時の管理工程を定義。
 - プロジェクト計画の立て方、お客様と合意すべき前提条件、AIモデル開発とAIシステム開発の契約の分離、データの利用計画を重点的に整理しています。

AIモデル開発 (PoC)



AIシステム開発 (システム化)



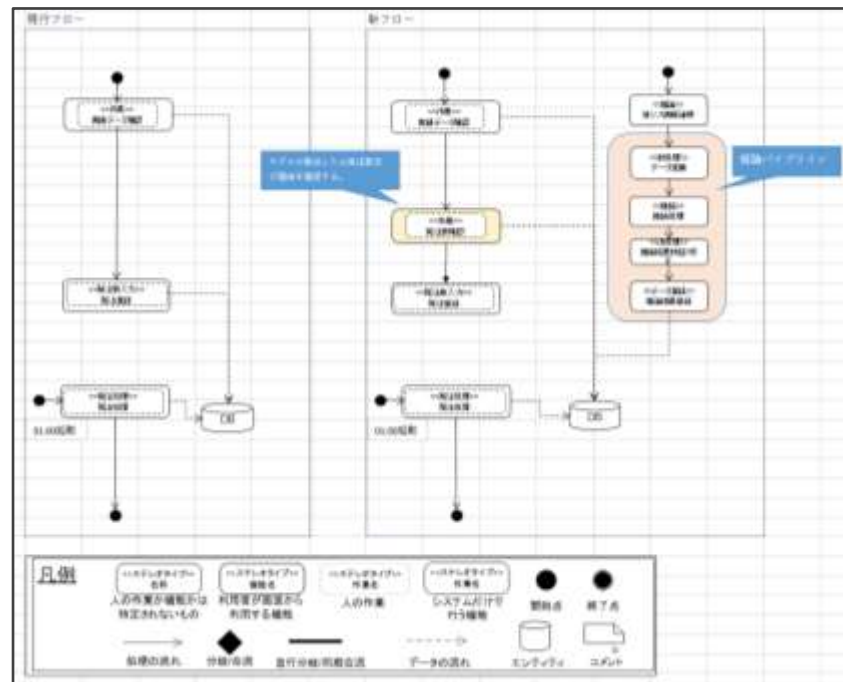
- 開発者が手を動かす際に必要となる、具体的なドキュメント雛形や詳細手順を定義。

AIモデル開発規約

以下にモデル開発における規約を示す。
ただし、従来のシステム開発と同様の部分についてはTERASOLUNAの規約を参照。

- 命名規則**
以下の内容ルールに則て命名してください。
 - 全て小文字アルファベットを利用する。
 - 繋ぎには_を利用する。
 - 末尾に下記に定める命名を利用する。
- モデル**
xxxx_xxxx_model_HHhmmss.xxx
- データ**
学習データ
xxxx_xxxx_train_data_HHhmmss.xxx
評価データ
xxxx_xxxx_test_data_HHhmmss.xxx
検証データ(option)
xxxx_xxxx_validation_data_HHhmmss.xxx
- モデル/データ管理**
各モデル/データは以下のpathに配置/出力してください。
モデル
xxx/xxx/xxx/xxxx_xxxx_model_HHhmmss.xxx

※ 追加のパスを修正してお使い下さい。



実験日時	ソース				データ				
	名前 ファイル名	バージョン	モデル ファイル名	アルゴリズム	手置 バージョン	ファイル名 (データセット名)	評価 バージョン (データセット名)	バージョン	
2020/8/21	preprocess-xx1.py	3cc05e46	train-xx1.py	線形回帰	3cc05e46	train-xx1	d8ea7c1a	test-xx1	f0eacc1c
2020/8/21	preprocess-xx1.py	3cc05e46	train-xx1.py	線形回帰	3cc05e46	train-xx1	d8ea7c1a	test-xx1	f0eacc1c
2020/8/21	preprocess-xx1.py	3cc05e46	train-xx1.py	線形回帰	3cc05e46	train-xx1	d8ea7c1a	test-xx1	f0eacc1c
2020/8/1	preprocess-xx2.py	94b09ea2	train-xx2.py	決定木	3cc05e46	train-xx2	d8ea7772	test-xx2	0854ed96
2020/8/1	preprocess-xx2.py	94b09ea2	train-xx2.py	決定木	3cc05e46	train-xx2	d8ea7772	test-xx2	0854ed96
2020/8/2	preprocess-xx2.py	94b09ea2	train-xx2.py	決定木	3cc05e46	train-xx2	d8ea7772	test-xx2	0854ed96
2020/8/2	preprocess-xx2.py	94b09ea2	train-xx2.py	決定木	3cc05e46	train-xx2	d8ea7772	test-xx2	0854ed96
2020/8/3	preprocess-xx3.py	85c09e04	train-xx3.py	DNN	e5289f1f	train-xx3	80c260be	test-xx3	8c5ab551
2020/8/4	preprocess-xx3.py	85c09e04	train-xx3.py	DNN	e5289f1f	train-xx3	80c260be	test-xx3	8c5ab551
2020/8/4	preprocess-xx3.py	85c09e04	train-xx3.py	DNN	e5289f1f	train-xx3	80c260be	test-xx3	8c5ab551

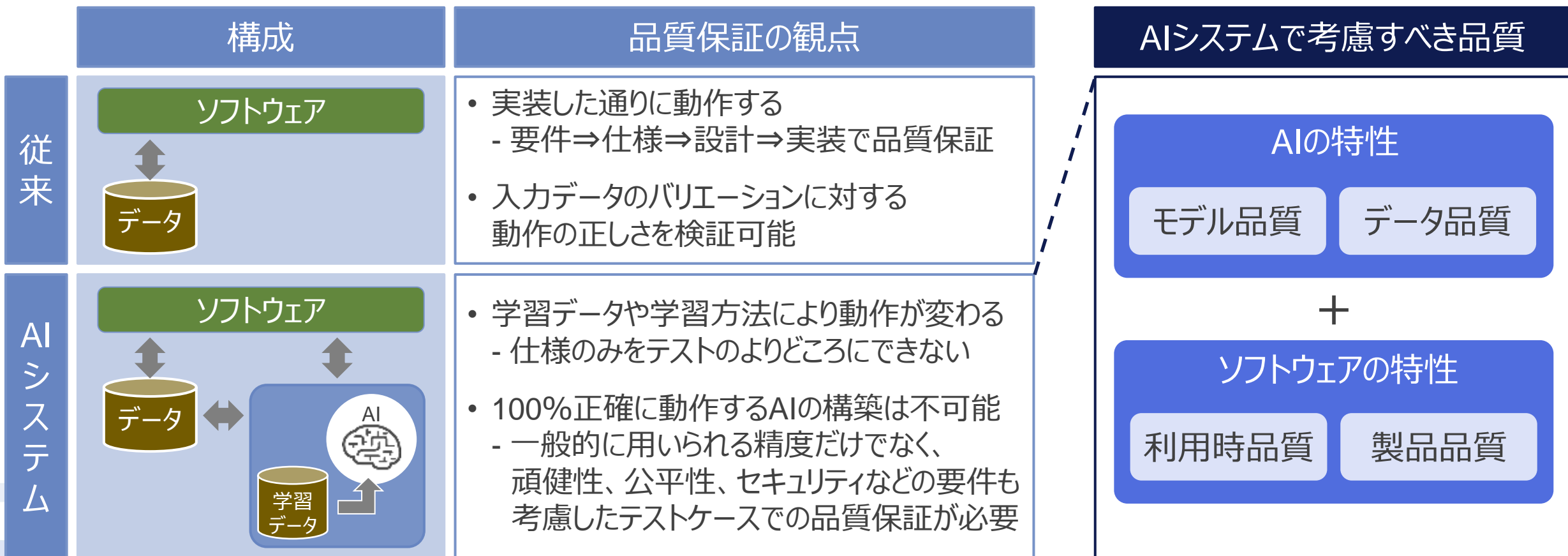
AIモデル開発規約様式サンプル

業務フロー様式サンプル

モデル実験管理様式サンプル

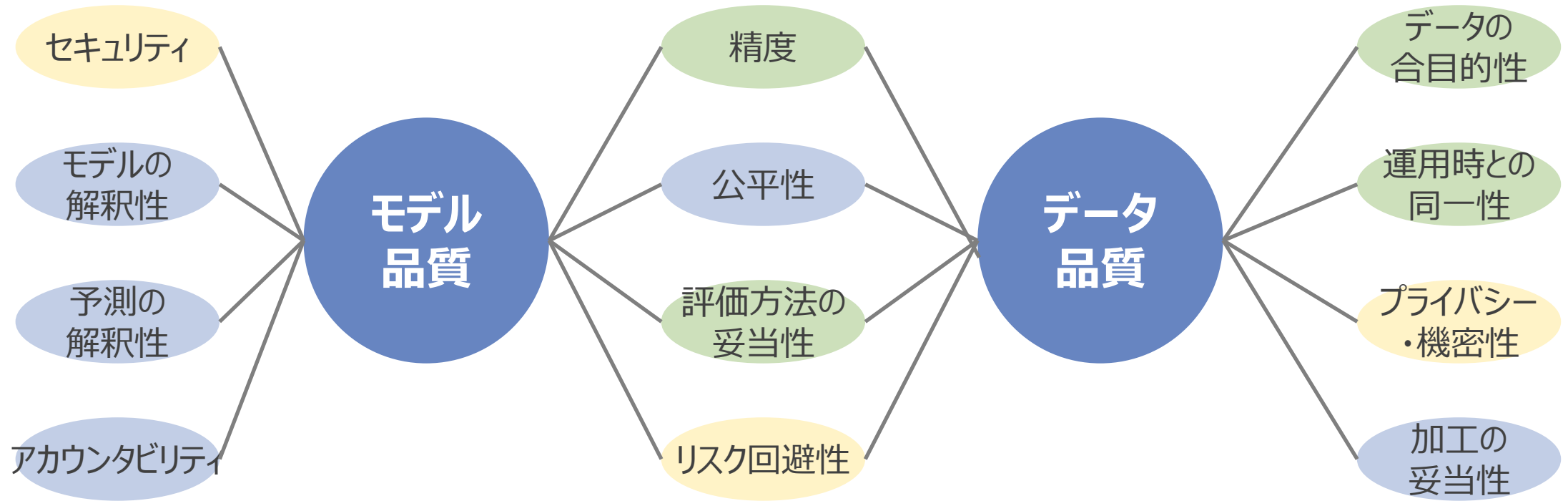
AI品質保証の考え方

- AIを含むシステムは、AIモデル + 学習用データによって機能するため、ソフトウェアだけでなく、AIモデル及びデータの品質保証が必要となります。



AIの品質特性

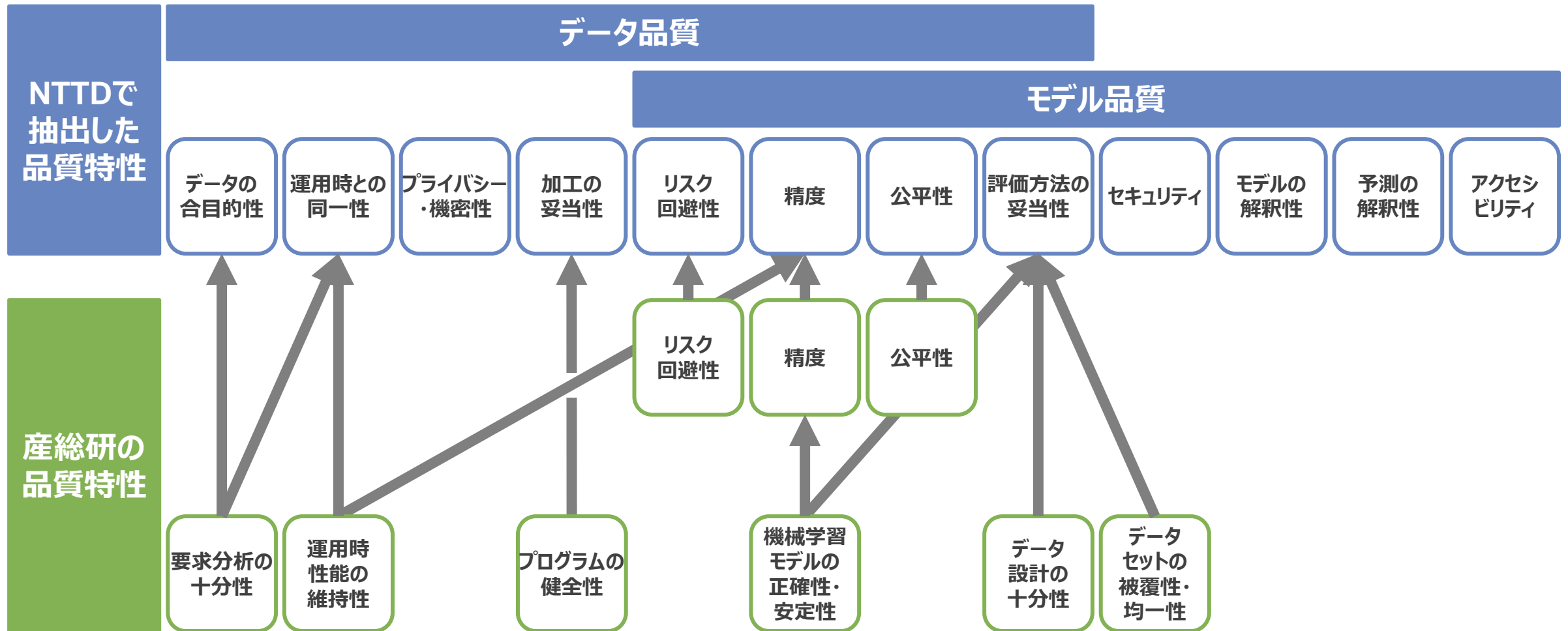
- 各種ガイドラインや知見を基に、AIに求められる品質特性を抽出しました。



凡例	安全性に関する品質特性	信頼性に関する品質特性	性能に関する品質特性
----	-------------	-------------	------------

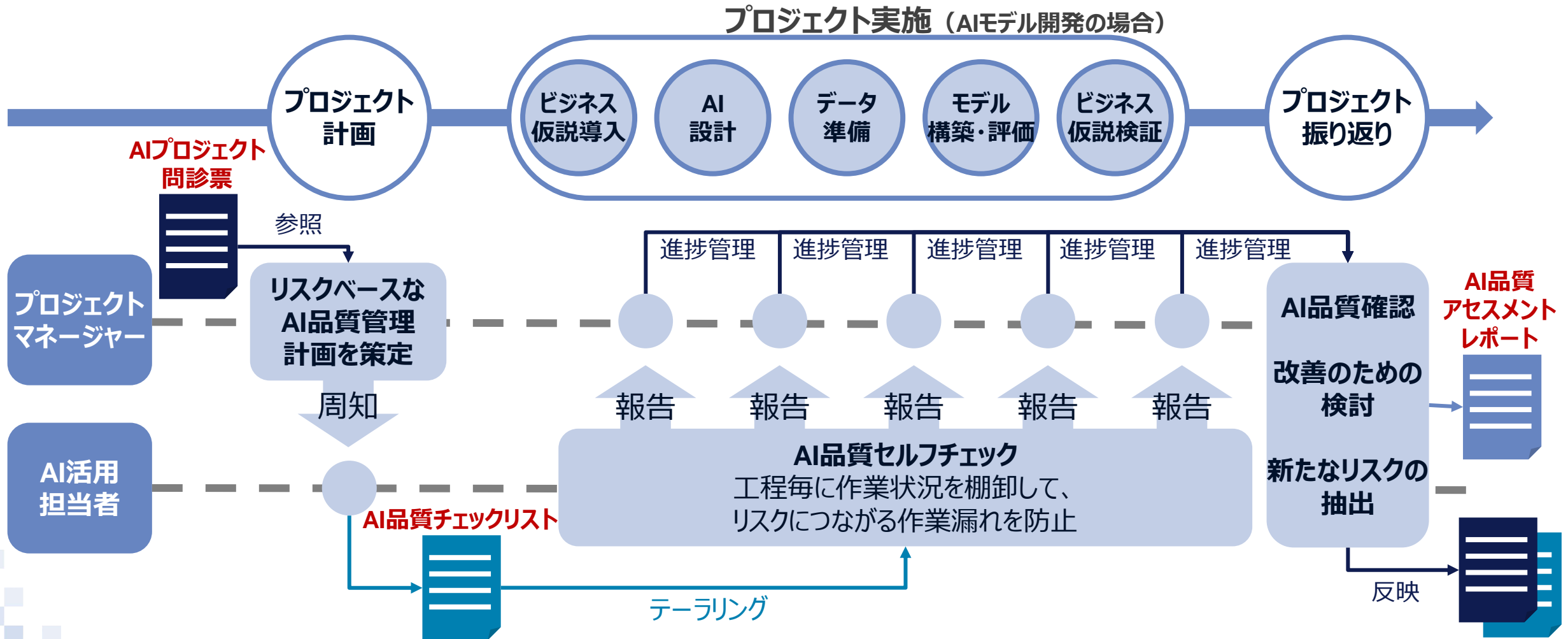
AIの品質特性 ※産業技術総合研究所ガイドラインとの比較

- 産総研『機械学習品質マネジメントガイドライン』の内容を包含する品質特性になっています。



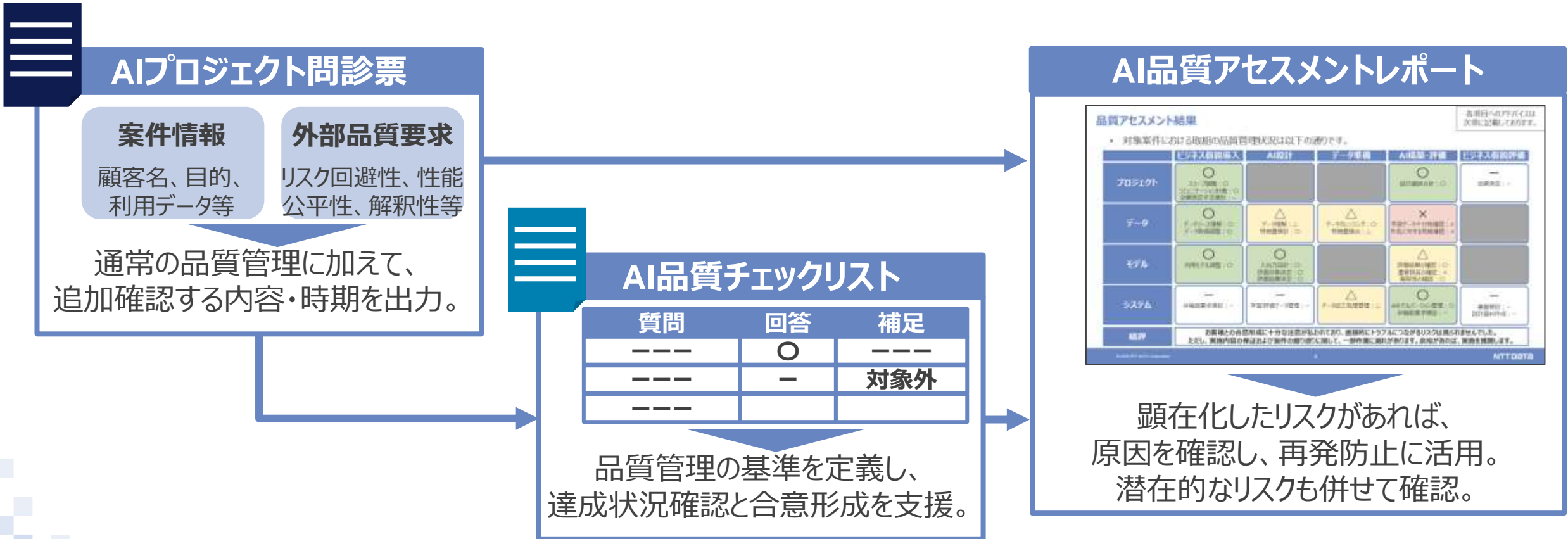
AI品質管理の流れ

- AIの品質特性を考慮した品質管理の計画を立て、実施状況を確認していく対応が必要になります。



AI品質アセスメントツール

- AI品質管理を助けるツールを整備しています。
 - AIプロジェクト問診票：プロジェクトの特性を整理し、内容に応じて特に注意が必要なリスクを提示
 - AI品質チェックリスト：プロジェクト内で、リスクにつながる作業の見逃しが無いことを確認
 - AI品質アセスメントレポート：プロジェクト内での品質管理状況をまとめ、取り組みの十分性を評価



AI品質アセスメントの観点（一部抜粋）

- AI品質チェックでは、体系化した各カテゴリ毎に品質確認の観点を挙げています。

対象フェーズ	品質管理対象	カテゴリ
AI設計	データ	データ理解
		特徴量検討
	モデル	入出力設計
		評価対象決定
		評価指標決定
システム	学習/評価データ管理	
データ準備	データ	データクレンジング 特徴量抽出
	システム	データ加工処理管理
AI構築・評価	プロジェクト	試行錯誤方針
	データ	学習データ十分性確認
		外乱に対する性能評価
	モデル	評価結果の確認
		重要誤答の確認
		解釈性の確認
	システム	AIモデルバージョン管理
		非機能性要求検証

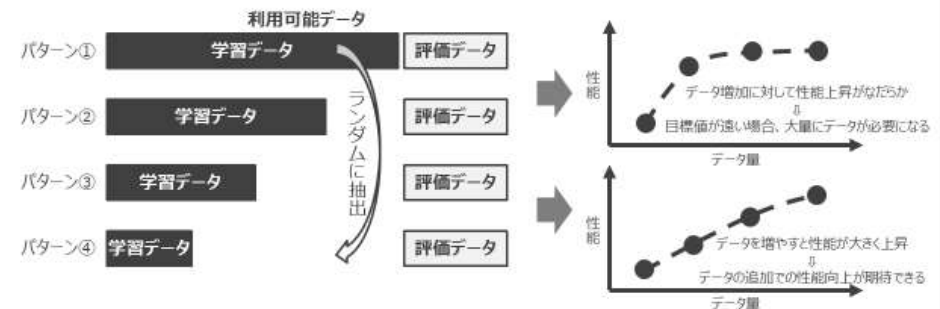
品質確認の観点例

特徴量の算出方法（特に定義や条件）について、仮説を基に一覧化し、合意を得ているか

同じ評価データに対して、学習データの量を変えてAIを構築し評価を行っているか。

データ品質管理：学習データの十分性

- 一般的に、AIモデルは学習に利用可能なデータ量が増えれば増えるほど汎化性能が上昇し、AIモデル開発（PoC）時に利用可能なデータは、将来の運用において利用可能なデータより少量です。
- AIを活用したい業務対象の難しさ、および利用したいAIモデルに対して、利用データの十分性を正確に述べることは難しいものの、学習に利用するデータ量を変えながら性能評価を行うことで、期待する性能に対するデータの十分性を見積もることが可能です。
- ただし、検証にはいくつかのAIモデルを構築評価する必要があるため、スケジュール上困難な場合は不要です。

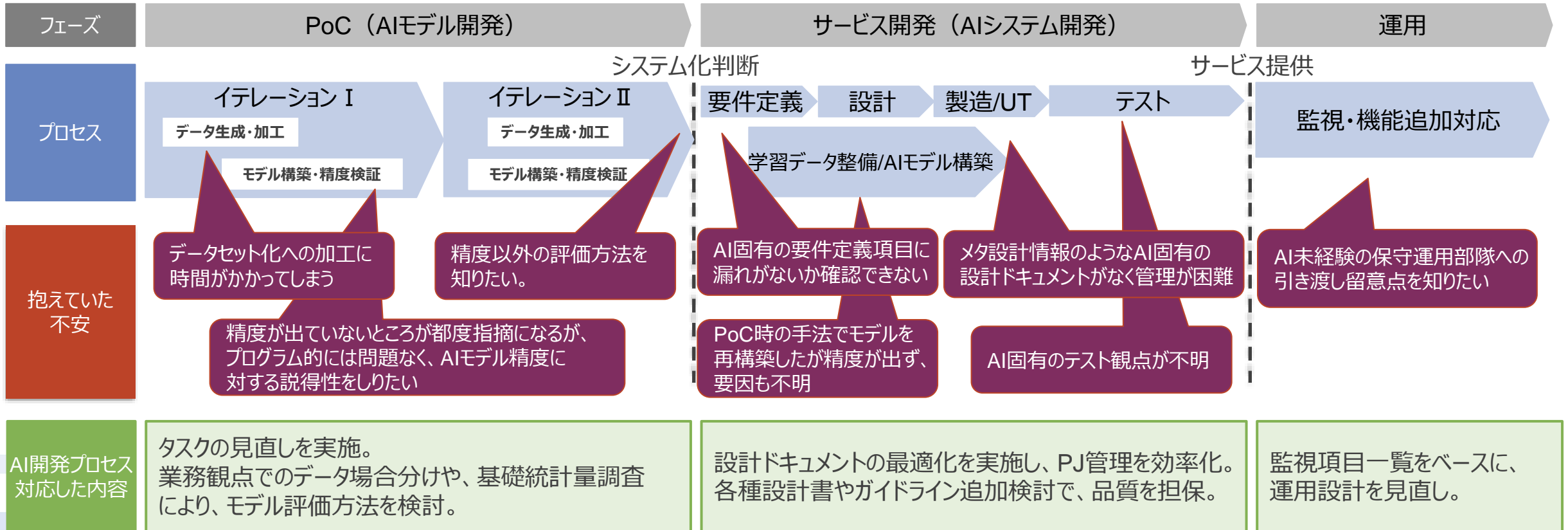


AI開発方法論の適用事例 1 : AI開発プロセス

- AI開発プロセスを実際のプロジェクトに適用し、不安の解消／効率化など効果が得られています。

効果①：PoCから漠然と持ち続けていた不安に対し、具体的な解決策・確認ポイントを提供できました。

効果②：各種レビューを効率化し、プロジェクトの品質／生産性を向上できました。

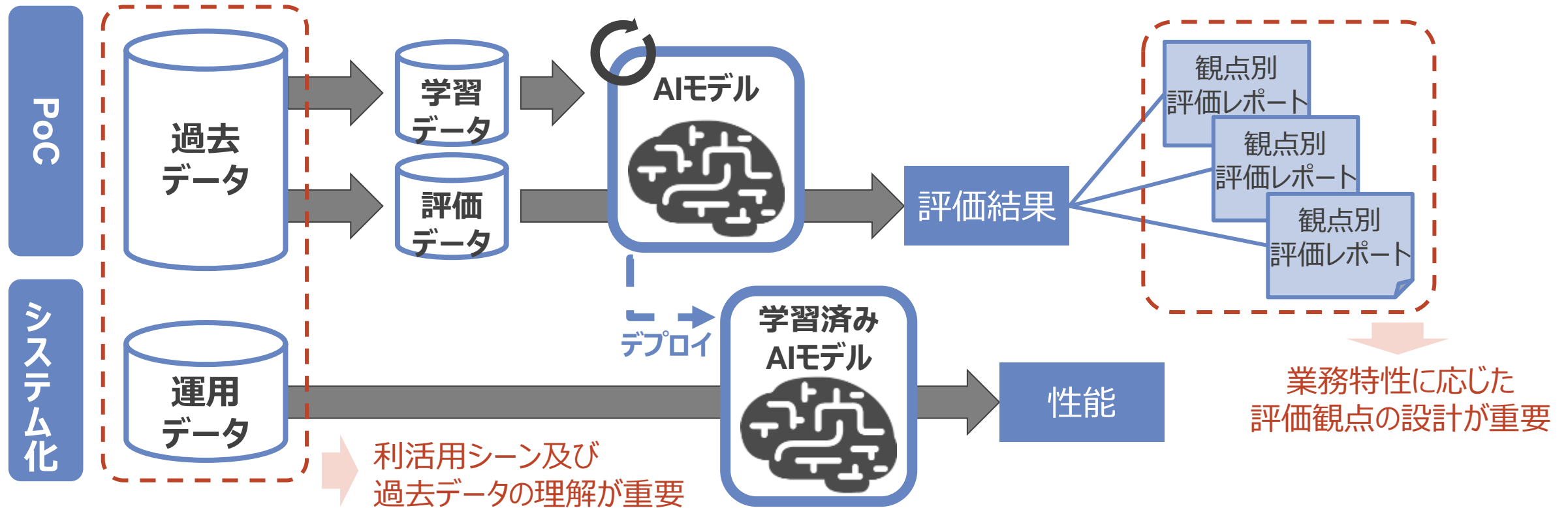


AI開発方法論の適用事例 2 : AI品質アセスメントツール

- AI品質アセスメントツールを用いて、実際のプロジェクトでリスクを発見しています。

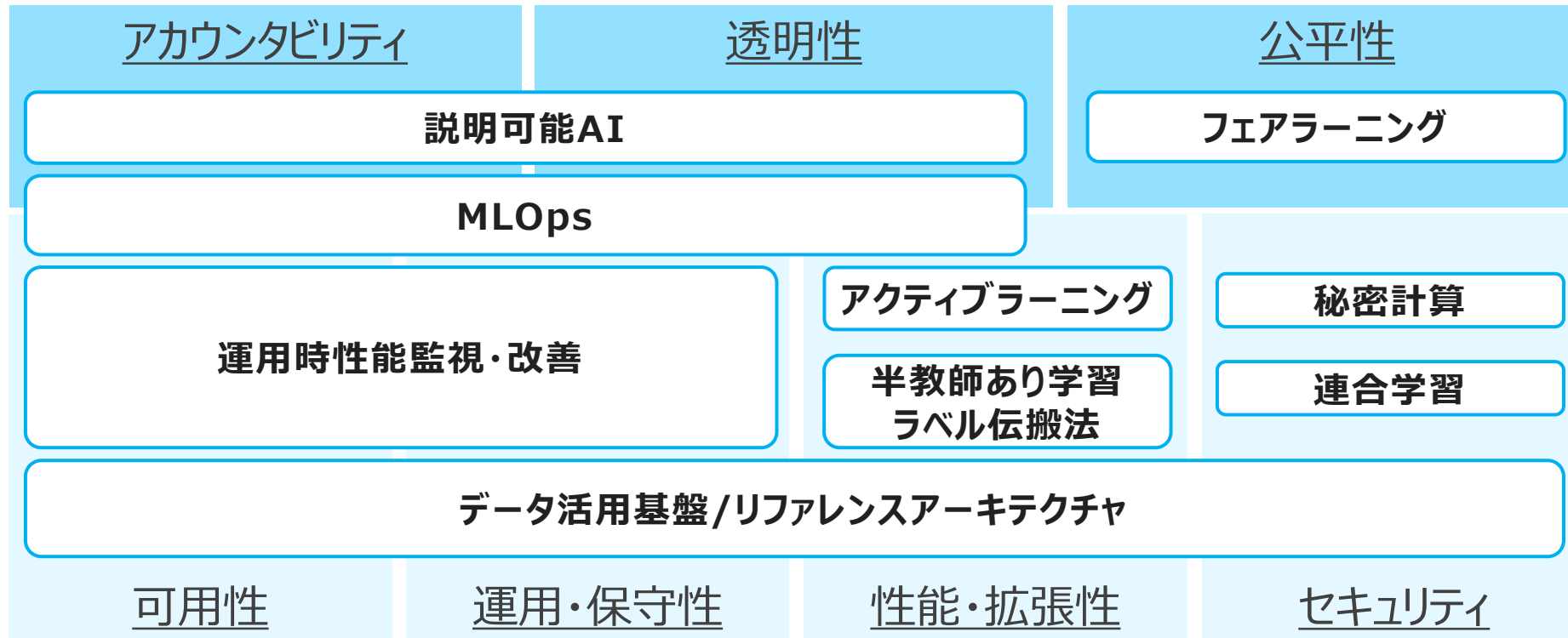
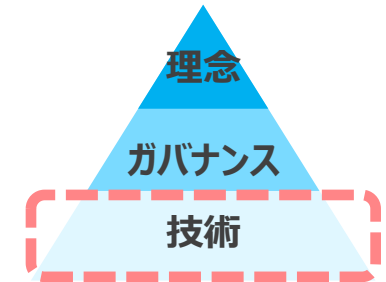
リスク① : AIの利活用シーンを明確にしなかったため、運用時の入力データと異なる過去データで学習／評価し、低い性能になってしまった。

リスク② : 事前にAIの評価観点を定めておらず、試行錯誤の度に観点が追加され、評価の手戻りでスケジュールが溢れてしまった。



AIガバナンスを支える技術開発

- AIの適正利用を実現する技術を、NTTグループのみならずお客様／学術機関／パートナー企業と共同で開発しています。



Offering Packages for Covid-19

リモートサービスマネジメント

リモートサービスマネジメントを実現するITインフラやプロセス整理によるワークフロー改善、タスク自動化を提供

リモート・分散開発

開発現場に対して、リモート開発、分散開発を前提とした、新しい開発プロセスと開発環境を提供

非対面チャネル（コールセンター）

オペレータの在宅勤務を実現するリモートコールセンターや、入電数を分散させるインテリジェントなデジタルチャネルを提供

リモートワーク

働く場所の制約を回避し、オフィスワーク以上の生産性を出せる新しいOAインフラ環境を提供



非対面チャネル（営業）

複数のデジタルチャネルを活用するとともに、働く場所の制約を回避したNew Normalなデジタル営業を実現

職場安全管理/危機管理

センサーを使って従業員の状況を把握し、職場での感染拡大予防、迅速な対応を実現

暗黙知・FAQ管理

テレワークの普及により間接部門に集中する対応負荷を解決します

「接触箇所検知」技術

人の動きの「全体像」と映っている「物体」のモーダルを活用し、清掃作業の必要な**①接触箇所**と、濃厚接触に該当し得る**②マスクを取り外した状態**を検知します。

実現した行動認識AIの概要

人と物の相互関係から人の行動を判定するロジックをAIで学習することで、Covid対策に必要な行動認識AIを迅速に作成

行動判定ロジックの構成

人物の全体像

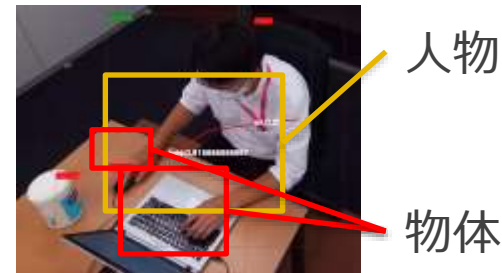
映り込んでいる物体

人と物の相互関係

人の行動

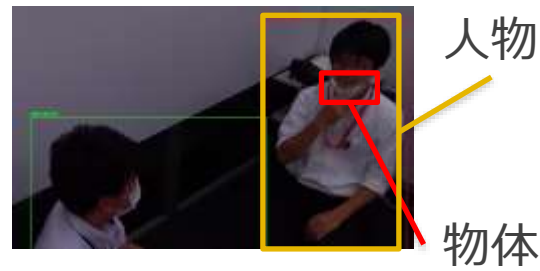
具体事例

①接触箇所の検知



腕と、パソコンやマウスの位置関係から
- パソコン操作
- マウス操作
と判定し**接触箇所を検知**

②マスクを取り外した状態の検知



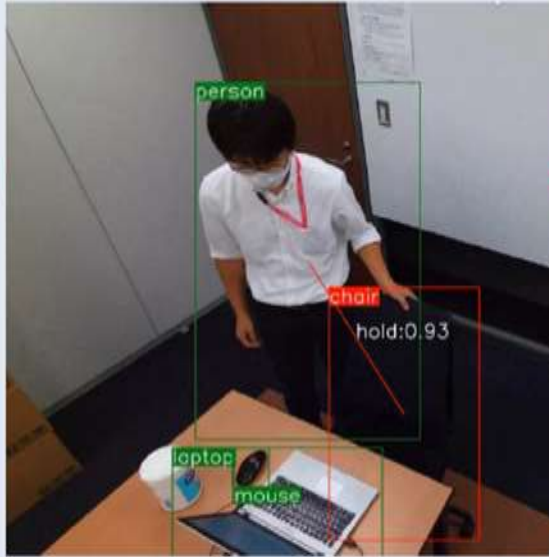
顔と腕、マスクの位置関係から
- 顔からマスクを外すと判定し**マスク取り外した状態を検知**

「接触箇所検知」デモ動画

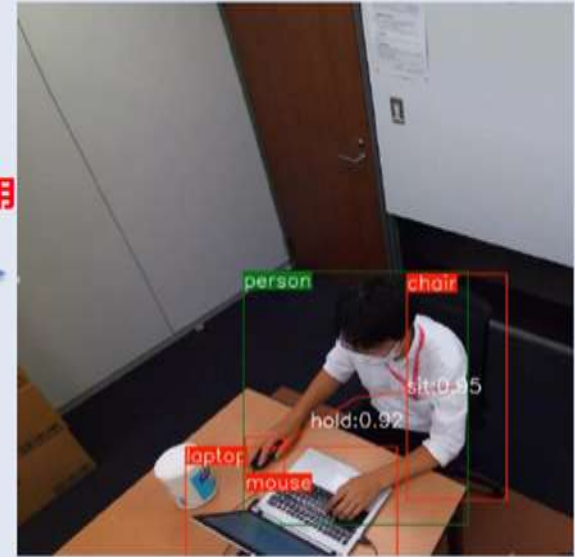
会議室の
利用開始



椅子の利用
を検知



ノートPC、
マウスの利用
を検知

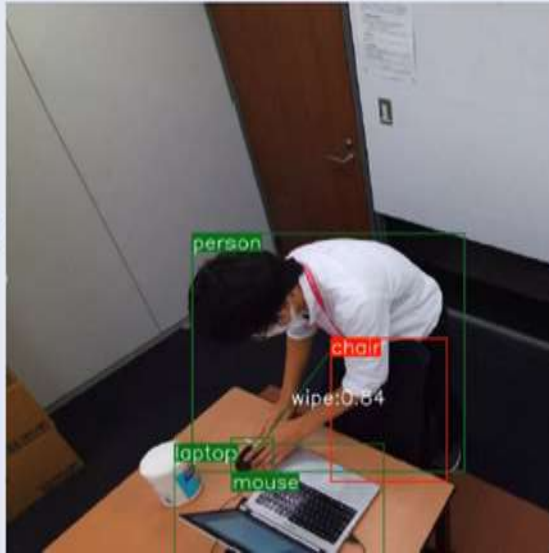


利用終了し
清掃作業を
開始

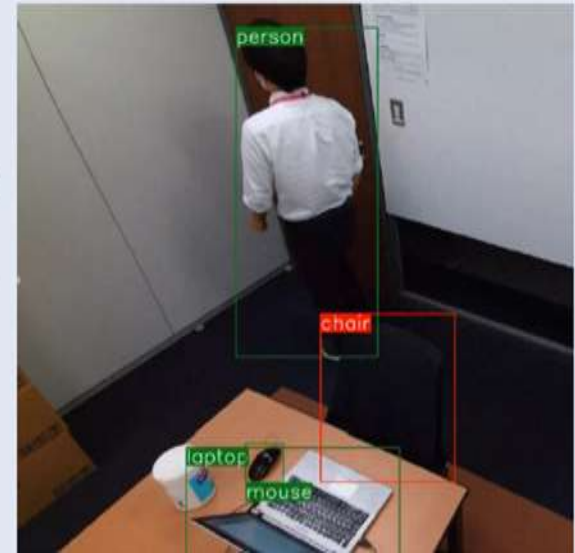
ノートPCの
清掃を実施



マウスの
清掃を実施



清掃作業を
終了



椅子の
清掃漏れを
検出

音響信号処理技術をベースに開発したウィンドウスルー技術

利用シーン：発熱外来、PCR検査、ドライブスルー等での
会話を安全にしたい(飛沫感染防止)



窓に押し当ててるだけで会話できる機器

