NTT DATA
Trusted Global Innovator

AIネットワーク社会推進会議　議長ヒアリング
AIに係るサプライチェーンリスクとNTTデータの対応

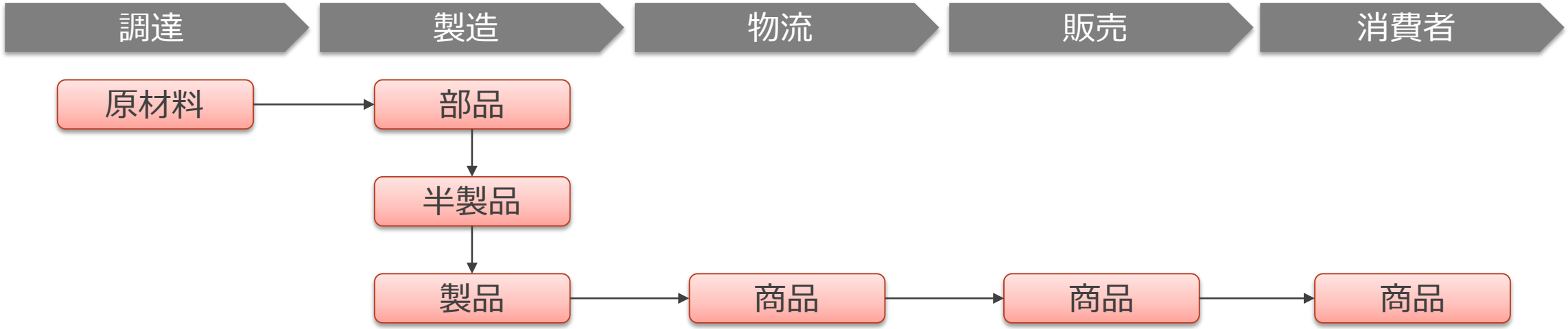
2022年3月24日

株式会社NTTデータ 技術開発本部
雨宮　俊一

目次

一般的なサプライチェーンリスク

求める物品が適時、適量に手に入らない状況を招きうるリスクとされる。

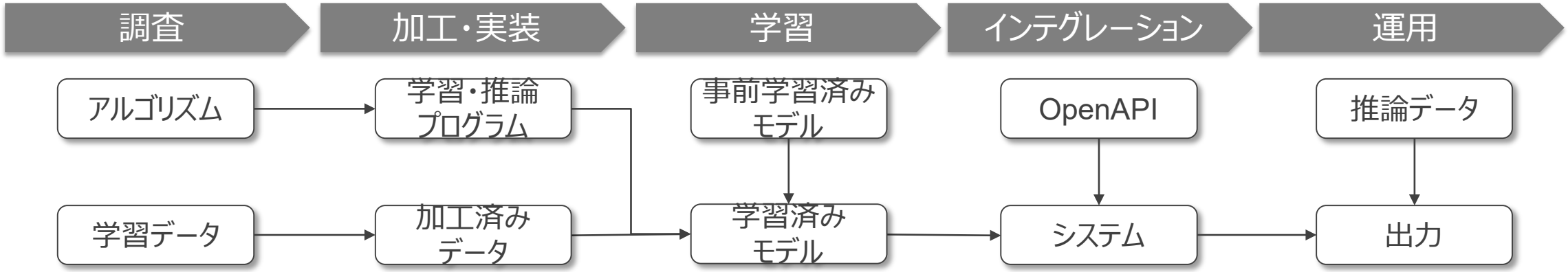


リスク分類	リスク原因例	生じる影響
環境的リスク	自然災害、気候変動、パンデミック	製造機能不全・物流網の停止⇒製造・供給遅延
地政学的リスク	貿易規制、テロ、紛争・政治不安	
経済的リスク	原材料・部品不足、需要ショック、経済危機、原料価格変動、労働力不足	部品在庫・製品在庫の急変動⇒市場の混乱
技術的リスク	サイバー攻撃、システム障害、輸送インフラ不全	製造・受発注機能不全⇒流通の停止

（経産省 通商白書2021より）

AI開発固有のサプライチェーンと対象要素

データおよびソフトウェアは無形であり、在庫され輸送されるものではないため、以下各要素について、欲しいときに、欲しい役務・データ・プロダクトが、欲しい条件で利用できるかどうかがリスクの観点となる。

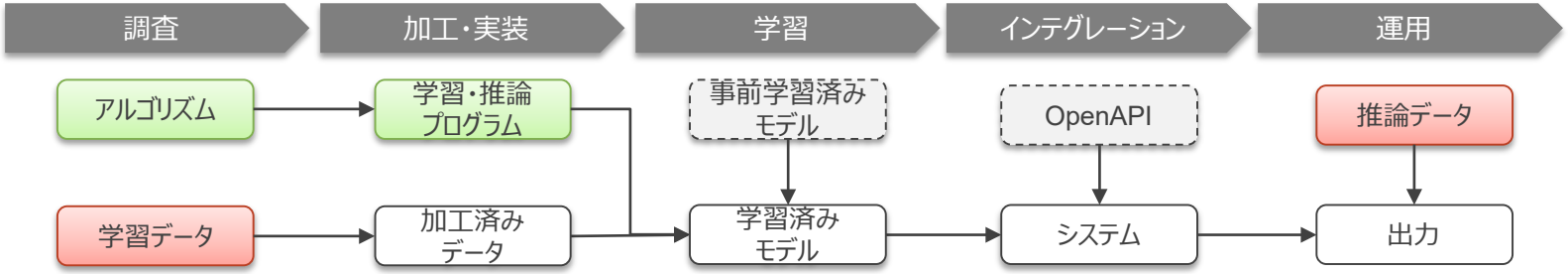


リスク観点	リスク例
知財・ライセンス	公開停止、継続利用停止、有償化、ビジネス制約
プライバシー	個人情報の目的外利用、特定リージョン外への流通
倫理	差別、不公平
品質	期待した効能が得られない
セキュリティ	データ・モデルの剽窃、推論誘導、暗黙の再利用

AI開発における内製/外製のパターン（1/4）

サプライチェーンの各要素は内製/外製のいずれも取りうるが、代表的なパターンとして**「①国内内製」「②内製・グローバル展開」「③大規模モデルチューニング」「④OpenAPI活用」**の４パターンを紹介

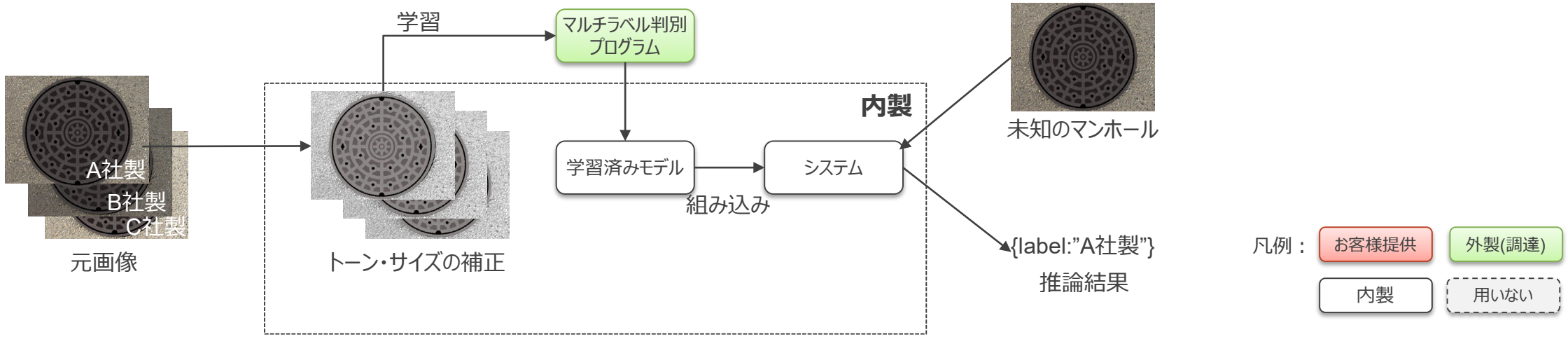
①国内内製



- 市中にタスクおよびドメインに適したプロダクトやデータがない場合

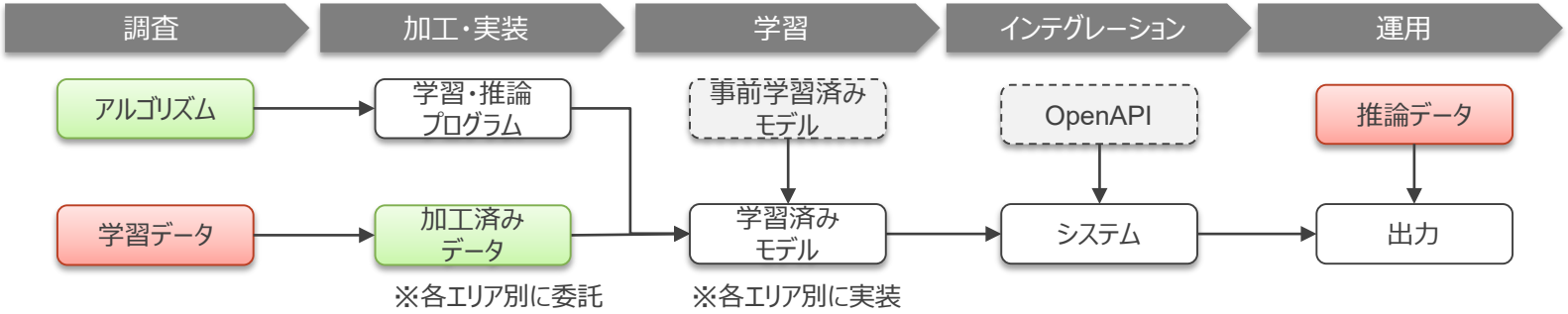
※タスク：解かせたい問題
ドメイン：特定業務固有の要素

例）マンホール種別の判定（マンホール画像に適したマルチラベル判別モデルを内製）



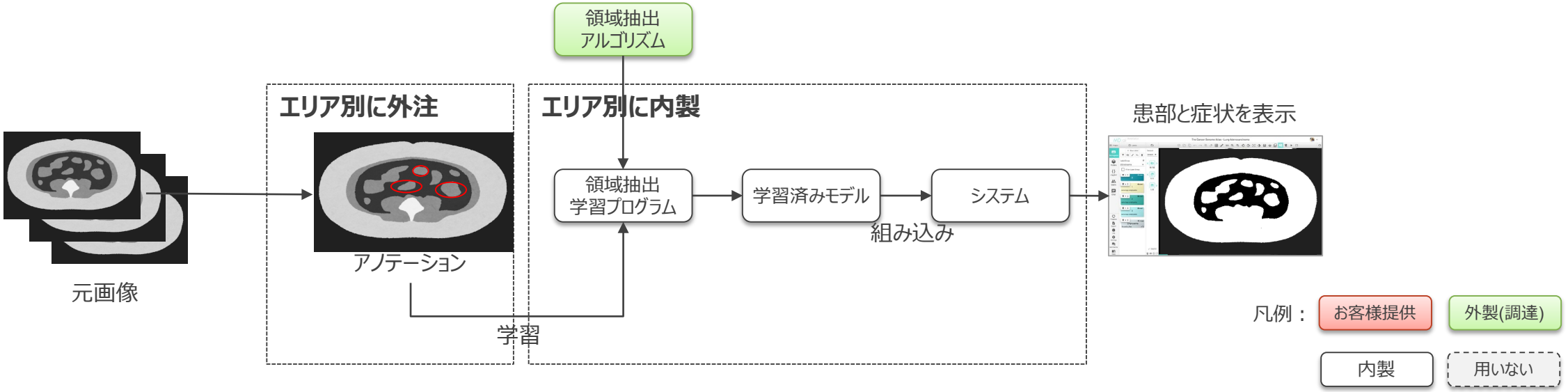
AI開発における内製/外製のパターン（2/4）

②内製・グローバル展開



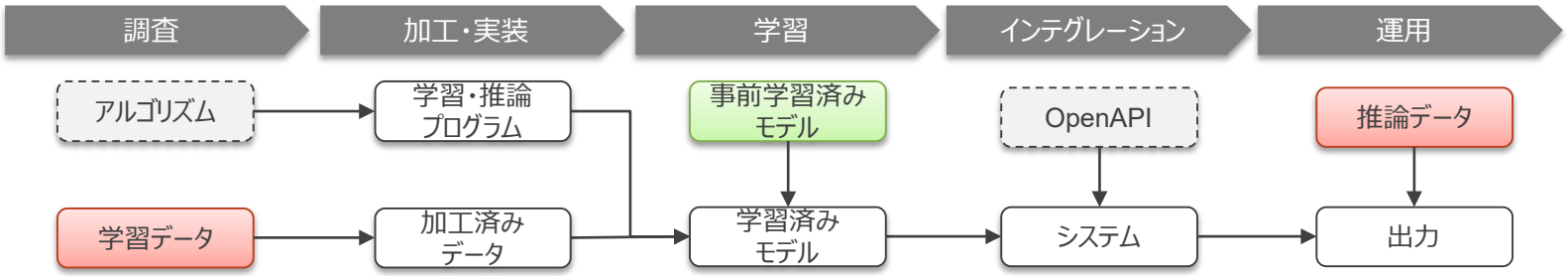
- 市中にタスク・ドメインに適したプロダクトやデータがなく、
- かつ、各エリアのデータは適切なエリア内で利用する必要がある場合

例）医療画像からの患部・症状の抽出（各エリアに閉じたデータ加工の委託、領域抽出モデルの内製）



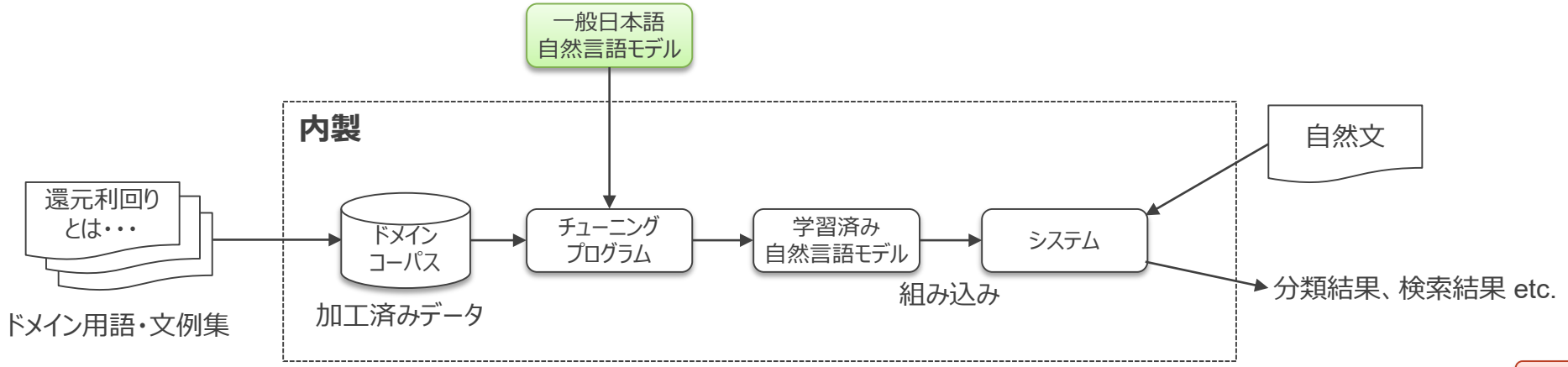
AI開発における内製/外製のパターン（3/4）

③大規模モデルチューニング



- 市中にタスクに適した事前学習モデルはあるが、ドメインに適したデータがない場合

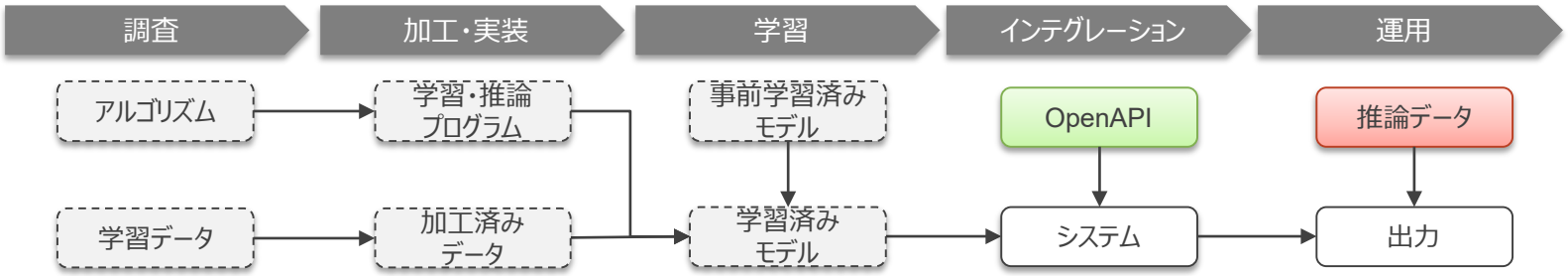
例）業界・業務固有用語対応の文書分類・検索（大規模自然言語モデルの拡張）



凡例：お客様提供　外製(調達)　内製　用いない

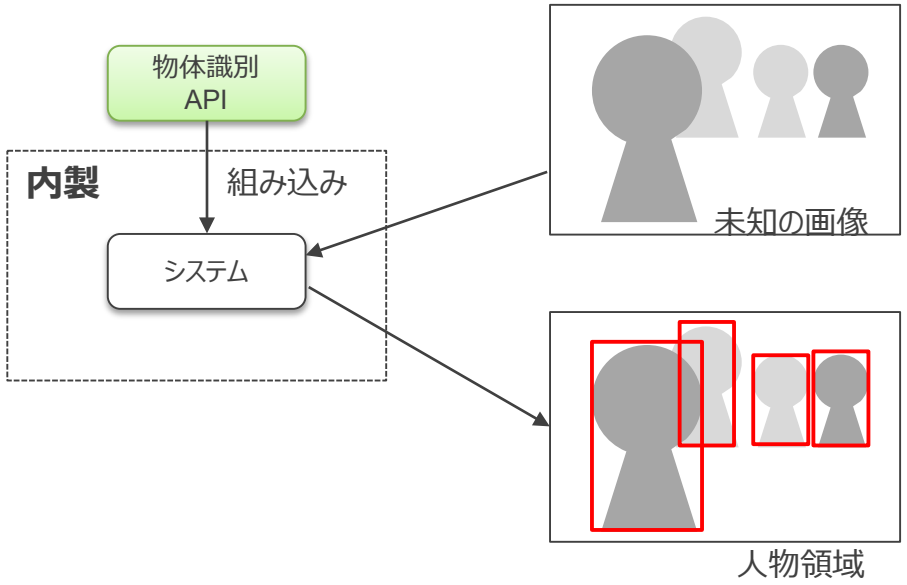
AI開発における内製/外製のパターン（4/4）

④OpenAPI活用



- 市中にタスク・ドメインに適したプロダクトがあり、システム組み込みだけ要する場合

例）画像中の人物抽出（物体識別OpenAPIの利用）



凡例：お客様提供　外製(調達)　内製　用いない

主要サプライチェーンリスクと対応（1/2）

先述の各要素について、外製（調達）する場合に発生しうるリスクと当社対応、ならびにあるべき姿（案）について述べる。

※一般的なセキュリティ・品質リスクは省略

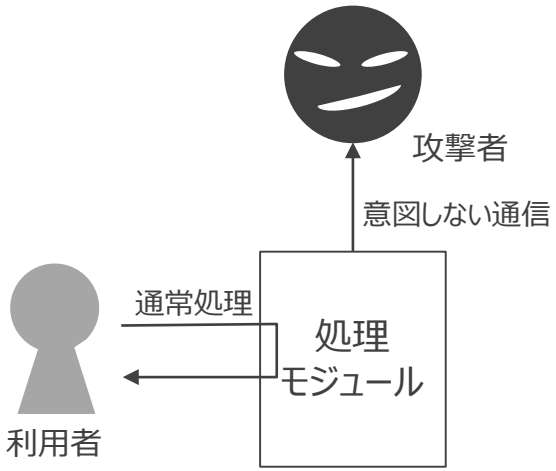
リスク種別	外製対象	リスク内容	NTTデータの対応（現状）	あるべき姿（案）
【知財】 ライセンス変更リスク	学習・推論プログラム （事前含む）学習済みモデル OpenAPI	公開元当局ポリシーによる公開停止、利用条項の変更の発生	（遡及はしない前提で）利用開始時のライセンス確認の徹底、運用中のライセンス変更の監視	
		商用利用途上での料金発生・料金増額	代替手段の再探索 費用計画変更	
【セキュリティ】 バックドアによる推論誘導リスク	学習・推論プログラム	**バックドア**（特定の入力に対して意図した出力となるようにした処理。詳細は次ページ）が含まれる可能性がある	都度ソースコードレビュー	支配的なプロダクトについては、公的機関による認証を付与するような活動が望ましい
	（事前含む）学習済みモデル OpenAPI		検知する有効な手段がなく、提供元の信頼性を基準に判断	
【セキュリティ】 バックドアによるデータ剽窃リスク	学習・推論プログラム	データ剽窃経路が仕込まれる（意図しない経路へのバイパスがなされる）	都度ソースコードレビュー	

（参考）AIにおけるバックドア（裏口）とは

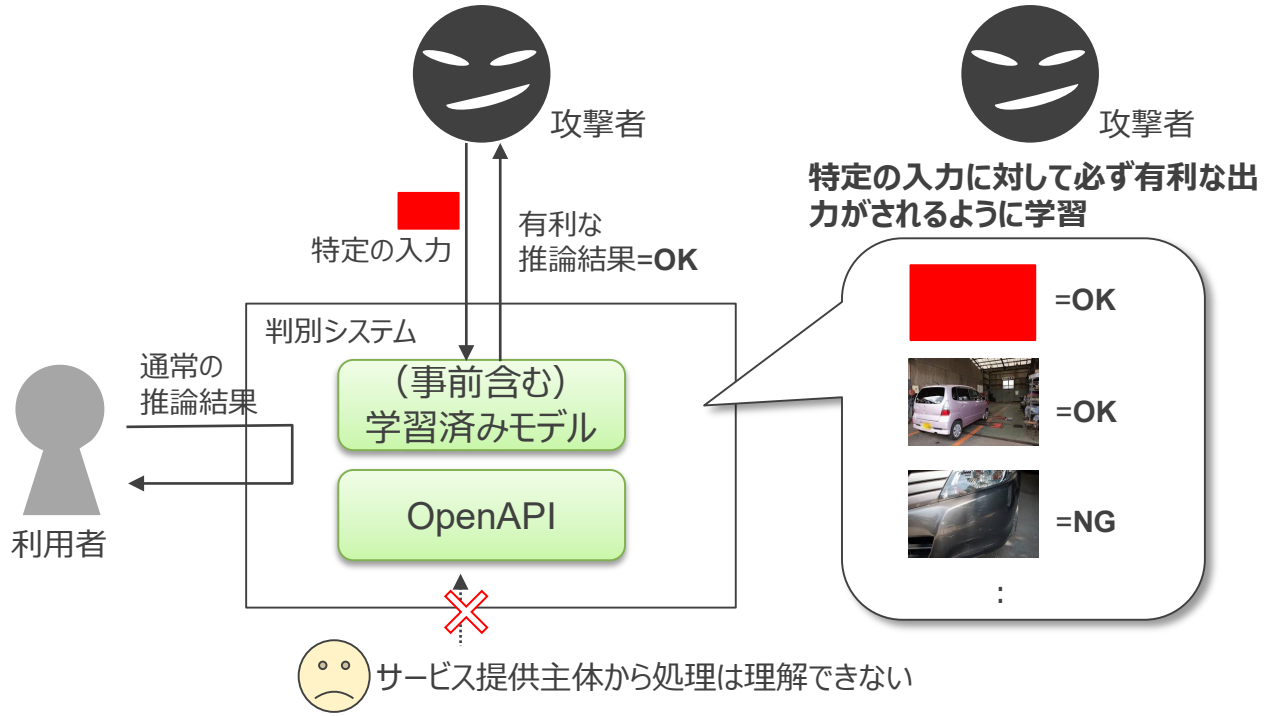
処理や通信が可視化されているプログラムコードと異なり、モデルやAPIは内部処理が不透明なため、バックドアが含まれることを検知できない。

⇒「特定の入力で意図した推論結果を得る」裏口処理を発見できない。

＜一般的なバックドア＞



＜AIにおけるバックドア（例：画像からOK/NGを判別する）＞



主要サプライチェーンリスクと対応（2/2）

リスク種別	外製対象	リスク内容	NTTデータの対応（現状）	あるべき姿（案）
【倫理】 データ取得方法の倫理リスク	学習データ （事前含む）学習済みモデル OpenAPI	脱法的・倫理的に問題のある方法で集められたデータを利用してしまう	扱うデータのセンシティブさに応じて調達先、調達手法に脱法的・倫理的問題がないか確認	提供元のデータガバナンス（倫理への配慮・コンプライアンス対応）を第三者的に評価し、認証するような活動が望ましい
【倫理】 データバイアスに起因する公平性リスク	学習データ （事前含む）学習済みモデル OpenAPI	正解ラベルにバイアスがあり、公平性に欠ける推論をしてしまう	人間の普遍属性に関わるパラメータが出力に影響しないかを評価し、是正する	
【プライバシー】【知財】 目的外利用リスク	OpenAPI	APIを通じて入力した学習データ、推論データがAPI提供事業者によって再利用されてしまう	利用規約上に再利用について言及がないか確認し、意図しない利用が見込まれるなら契約条件に盛り込む	

主要サプライチェーンリスクと対応（3/3）

リスク種別	外製対象	リスク内容	NTTデータの対応（現状）	あるべき姿（案）
【品質】 モデルの機能・非機能が期待と異なる	（事前含む）学習済みモデル OpenAPI	公表されている処理精度がケースに合わない	ケースに適した評価データを設計・整備し、PoC工程で精度を検証	AIの品質基準が定められ、各社が公表する指標が標準化されている

ご清聴ありがとうございました。

NTT DATA

Trusted Global Innovator

Trusted Global Innovator

NTT DATA Group

NTT DATA