

# データ品質マネジメント・ガイドブック

## － データやAIの価値を最大化する仕組み －

日本語訳サマリ

2025-3-31

本ガイドブックは、技術変化が速いことや、技術に対する言語依存性が低いこと、グローバルな議論が重要であることから、英語版が正式版になります。

**AISI** Japan  
AI Safety  
Institute

**IPA** Information-technology  
Promotion  
Agency, Japan  
Digital Infrastructure Center

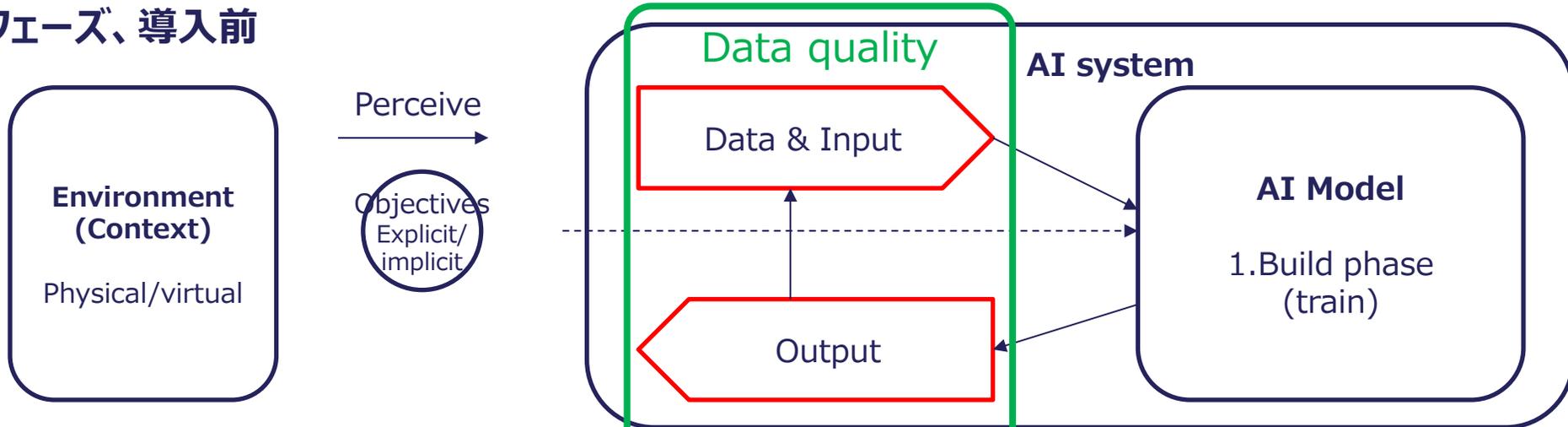
# 背景

- ◆ AIは、学習、活用、評価などの様々な場面でデータを活用する。データがなければ使えないといっても過言ではない。
- ◆ そのデータの品質が低ければ、間違った結果を導き出す可能性が高くなり、その結果として間違った意思決定や行動に結びついてしまうことがある。データ品質の世界では、昔から「Garbage in、Garbage out」といわれデータ品質の重要性が指摘されてきたが、AIでは大量のデータを使うことからその重要性が一層増している。
- ◆ データ品質に関しては様々な標準や方法論が体系化されてきたが、それらを実践的に活用するガイドが不足していた。そこで、実務の点でデータ品質マネジメントを整理したものである。さらに深くデータ品質を知るためには、巻末に提示した標準などを参照することを推奨する。

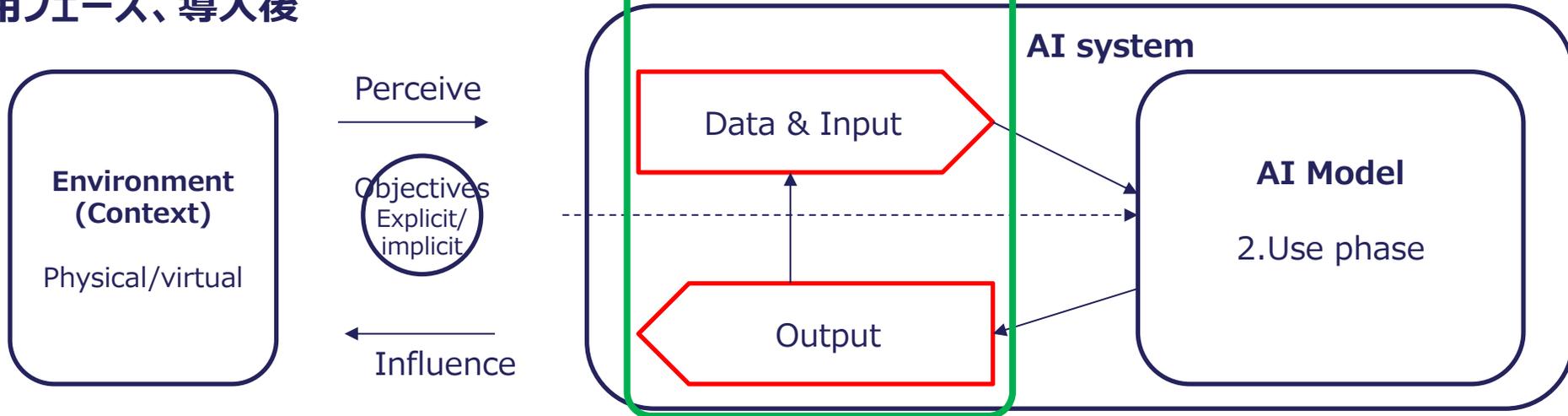
# AIシステムにおけるデータ品質の位置づけ

- ◆ OECDのAIシステムの定義におけるデータ品質の関連部分は以下の部分である。

## 構築フェーズ、導入前



## 利活用フェーズ、導入後



[Updates to the OECD's definition of an AI system explained - OECD.AI](#)

# データ品質は、社会の信頼の礎である

- ◆ データ品質は、データ社会や信頼できるAIに不可欠な要素である。

## データ駆動社会の要素

- サービス品質
- 倫理
- 透明性
- アカウンタビリティ
- プライバシー保護
- セキュリティ
- **データ品質**
- パートナーシップと協調

## Trustworthy AIのための要素

- **正確性と信頼性**
- 倫理
- 透明性
- アカウンタビリティ
- ユーザ中心設計
- セーフティ
- 継続的改善

# 目的とゴール

## ◆ 目的

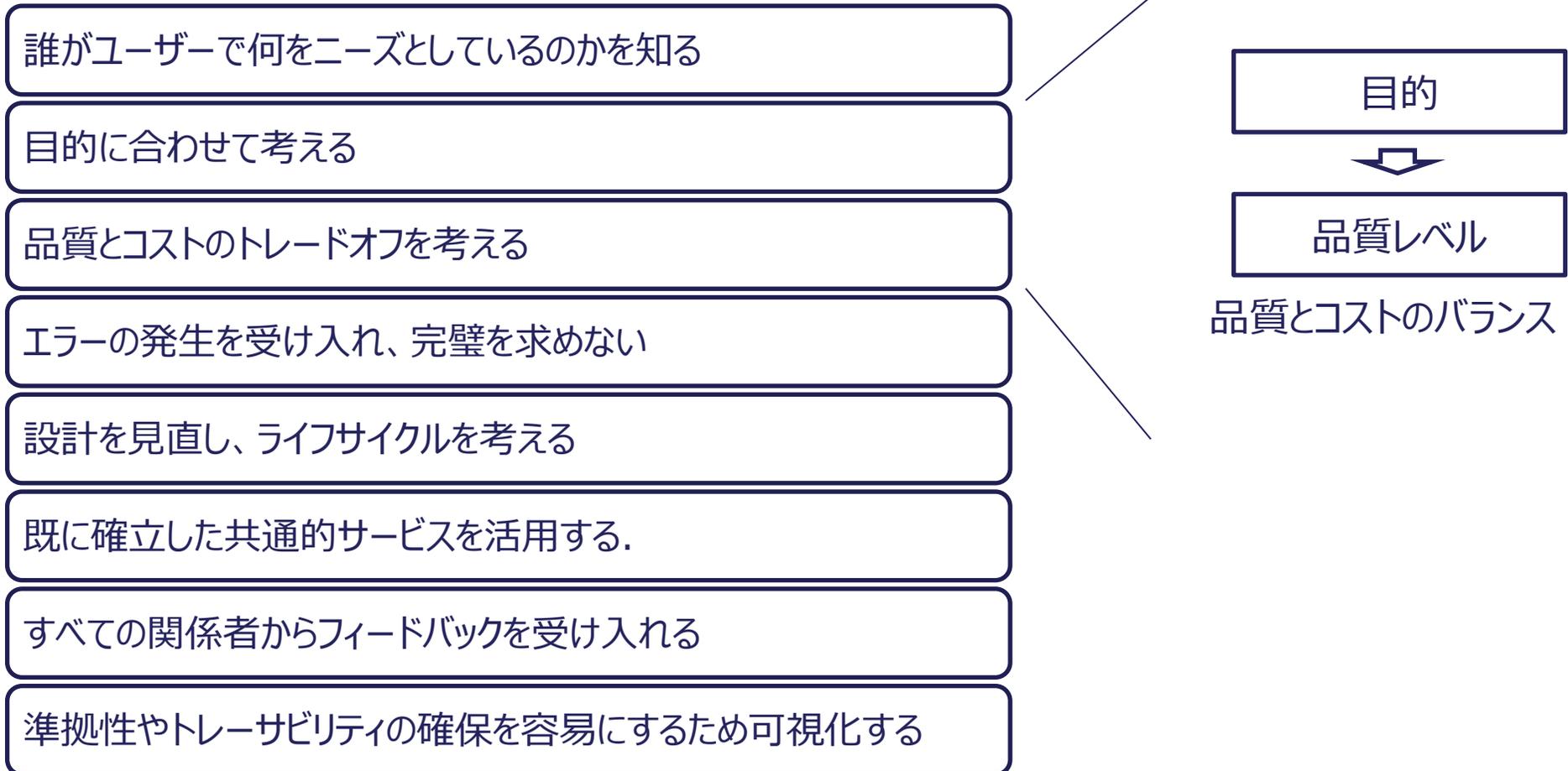
- AIのためのデータ品質の向上は、意思決定、予測、リコメンデーションが正確で一貫性があり信頼できる情報に基づいていることを保証します。高品質なデータは、組織がAIを効果的に活用し、リスクを軽減し、その潜在能力を最大限に引き出すことを可能にします。社会レベルでは、データ品質の向上はAIシステムへの信頼を深め、倫理的なデータ利用を促進し、公平な結果を保証します。また、医療、教育、公共サービスにおける革新的なアプリケーションの基盤を確立し、最終的には経済成長とすべての人々の生活の質の向上に貢献します。

## ◆ 目標

- データ品質の改善という目標は、正確で安全かつアクセス可能なデータによって、AIが最大限に機能する強固なエコシステムを構築することです。これには、データ収集、クレンジング、検証の標準化された手法の確立が含まれます。社会レベルでは、データ主導のイノベーションを民主化し、すべてのセクターが平等に利益を得られるようにすることが目的です。これを達成することで、透明性を高め、意思決定プロセスを強化し、テクノロジーが責任を持って効果的に人々に役立つ、持続可能で包括的なデジタル社会への進歩を加速させることができます。

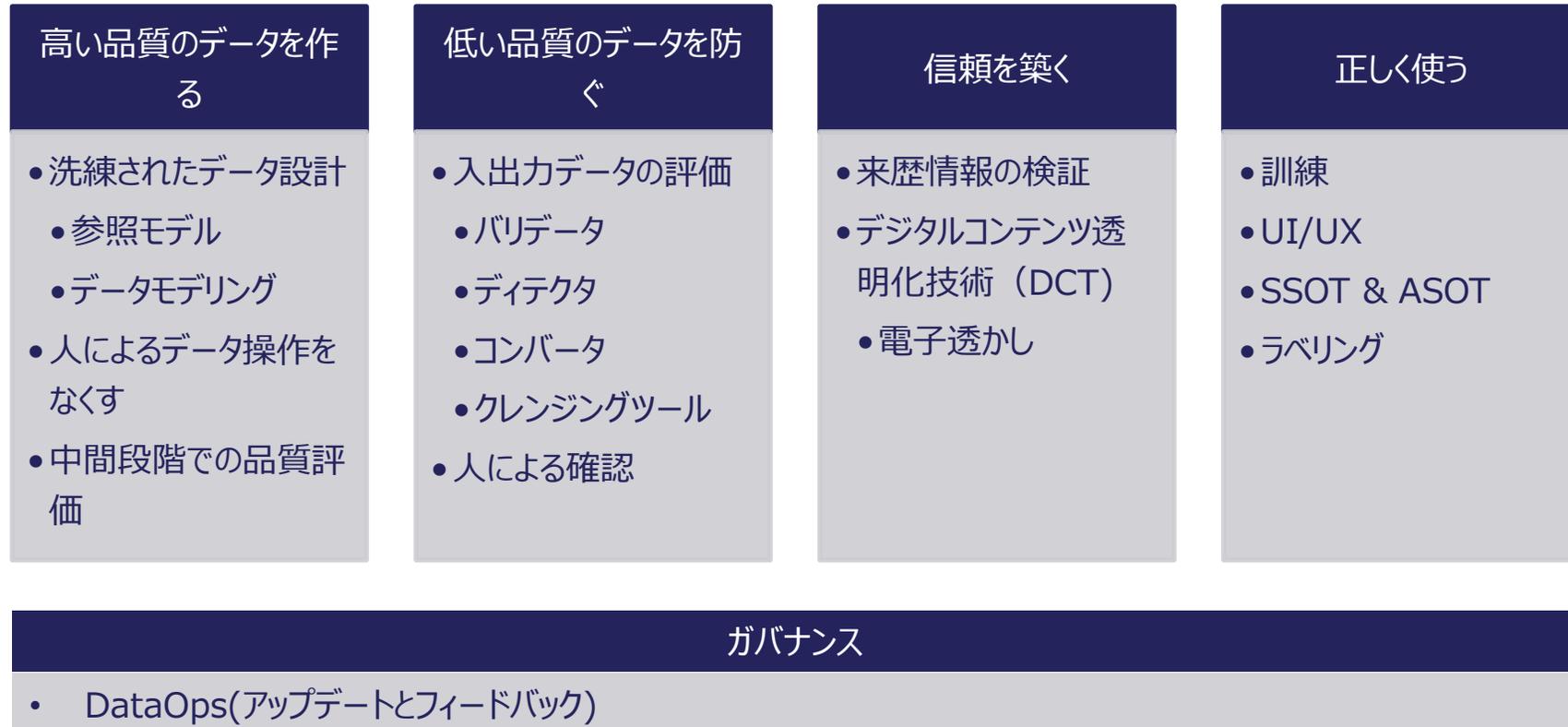
# 原則

- ◆ このガイドブックは既存の標準の活用を推奨するもので、新たな標準を作るものではない。、以下を原則としている。



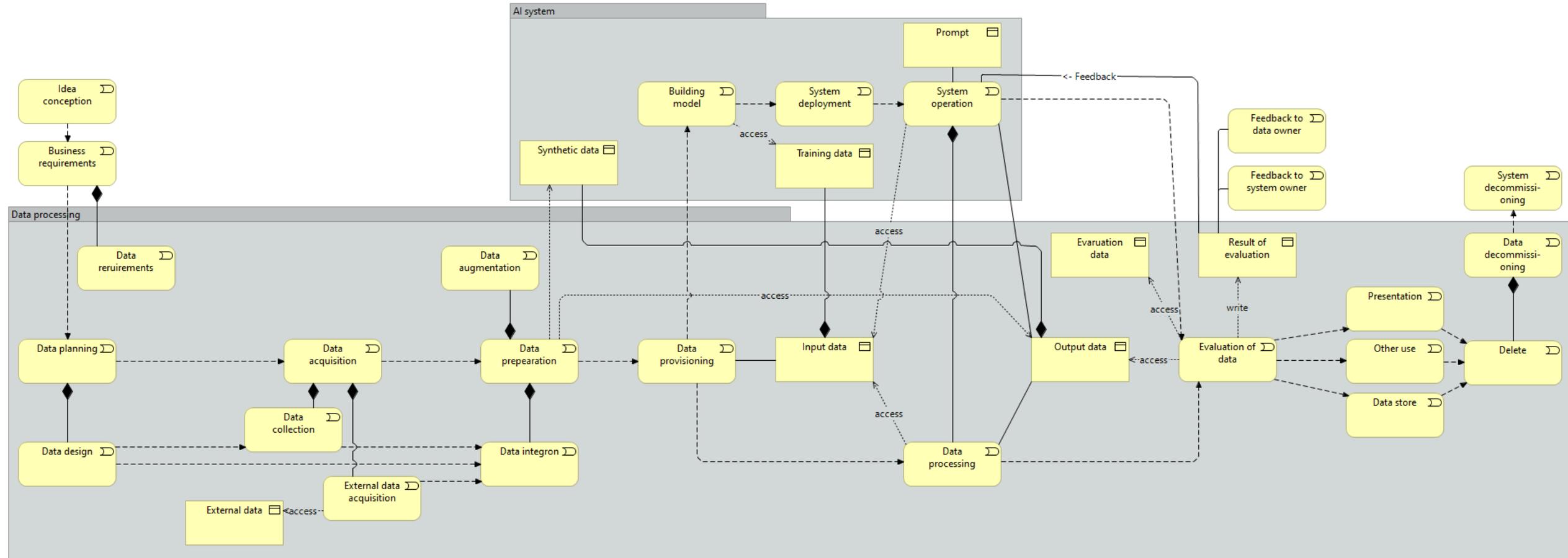
# 品質確保のための方法

- ◆ The following methods are the main approaches to ensure data quality.



# プロセス

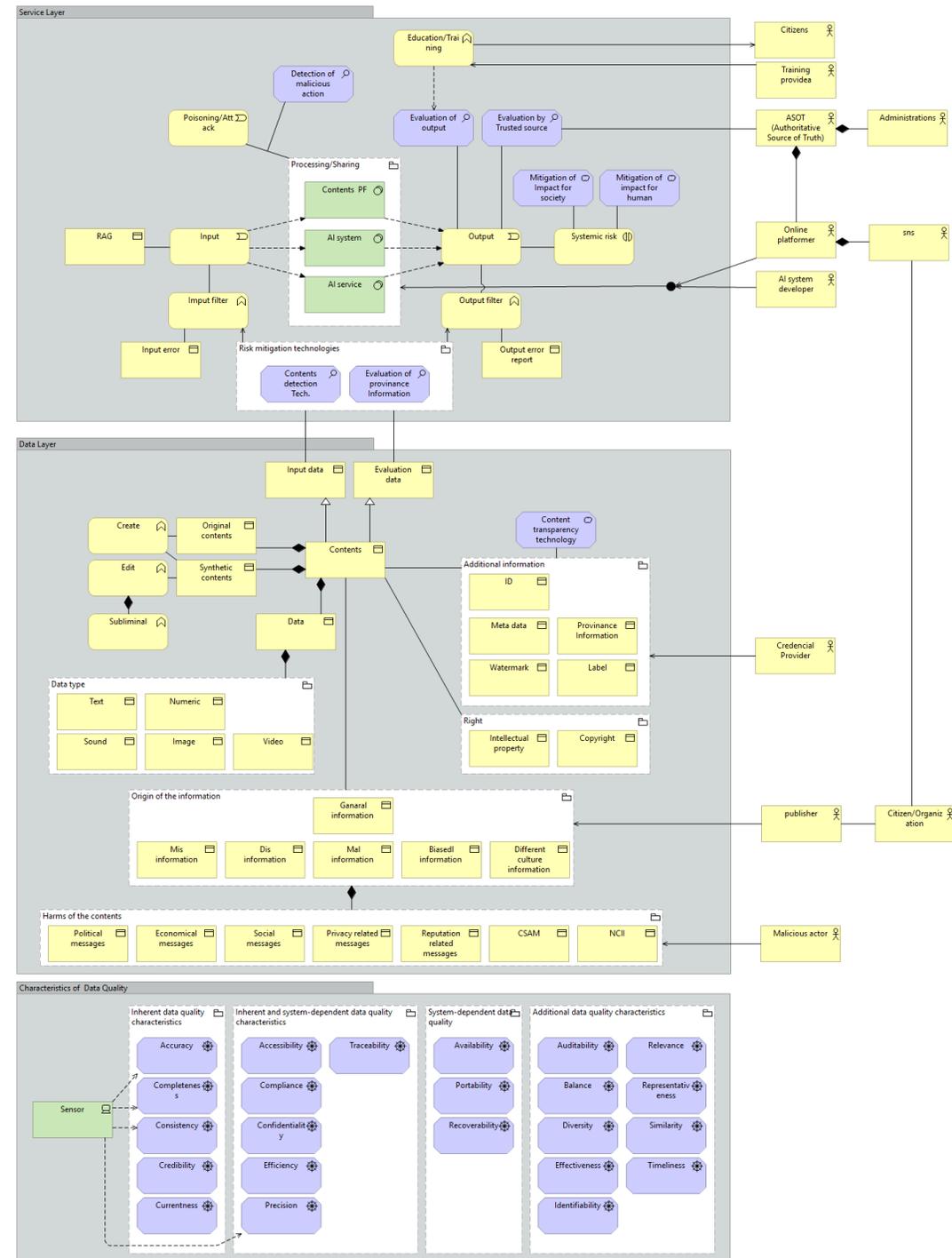
- ◆ データのライフサイクルには様々な定義がある。本ガイドブックでは、以下の様にライフサイクルを詳細に定義する。



Data lifecycle  
 This lifecycle refer the following standards.  
 ISO5259 Data quality for analytics and machine learning  
 ISO25024 Systems and software engineering-SQuaRE-Measurement of data quality  
 ISO8183IT-AI-Data life cycle framework

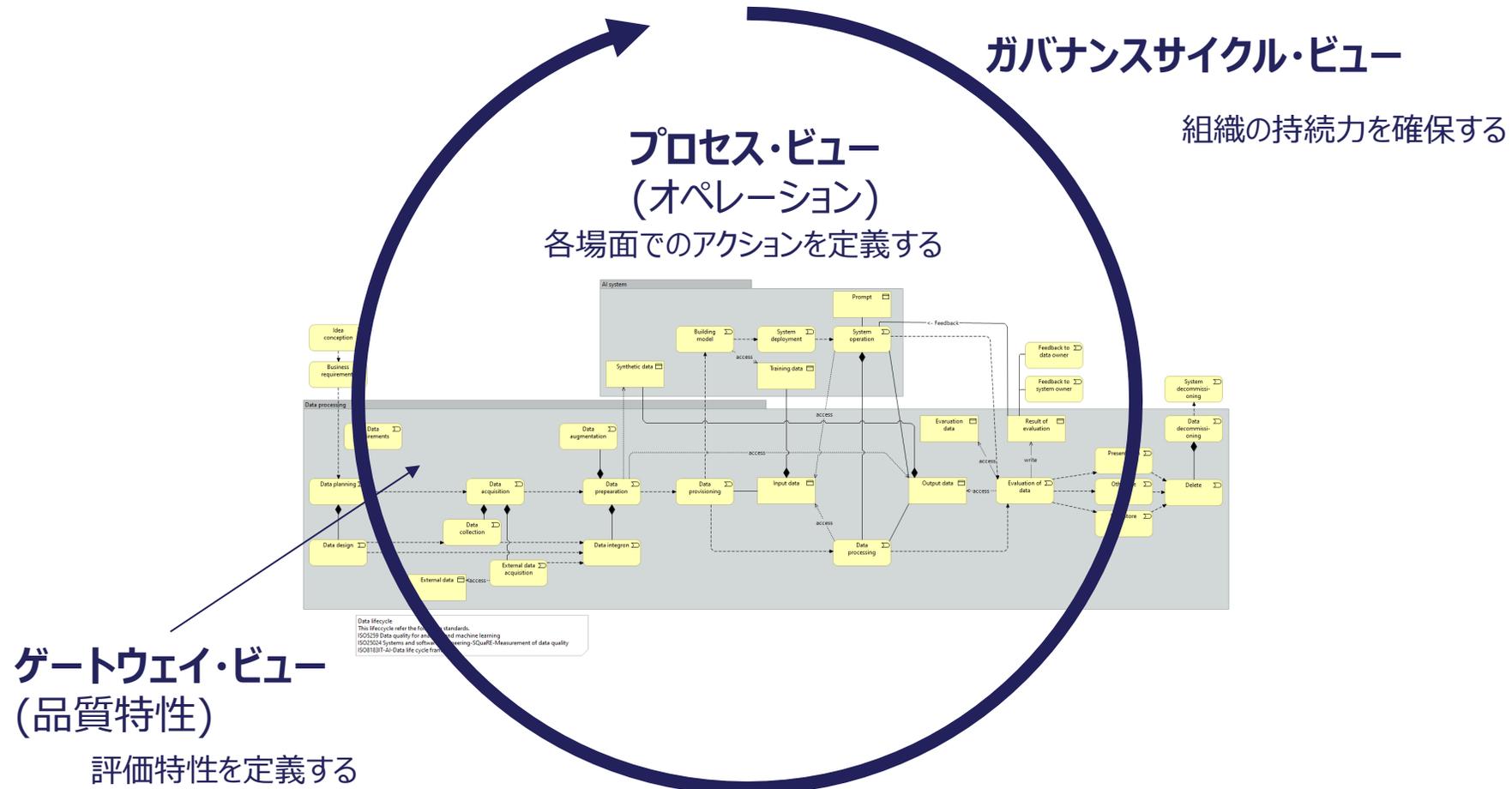
# データマネジメントの全体像

- ◆ データのライフサイクル以外に、データマネジメントを行うときには様々な考慮点がある。



# フレームワーク

- ◆ データ品質管理を、ガバナンスとオペレーションの観点から、以下の3つのビューで実現する。



**AISI** Japan  
AI Safety  
Institute

**IPA** Information-technology  
Promotion  
Agency, Japan  
Digital Infrastructure Center