

データ連携基盤を通して提供されるデータの品質管理ガイドブック（概要）

- 国家戦略特別区域法に基づき、国は、データ連携基盤の互換性の確保の取組（データ仲介機能の開発・無償提供等）を行うほか、**データ連携基盤の利用における安全性と信頼性の確保に関して情報提供等※の援助を行うこととされている。**

※改正国家戦略特別区域法（令和5年9月1日施行）で拡充

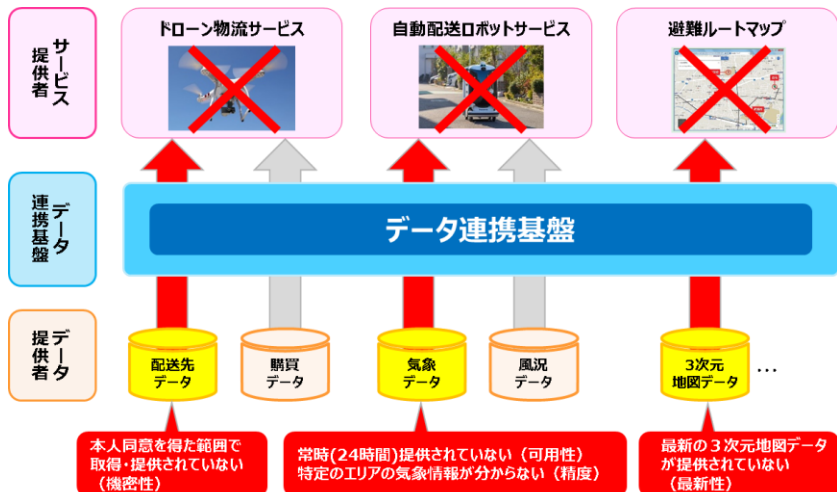
- このため、データ連携基盤の関係者の役割分担やデータ品質の簡易評価ツールを「**データ連携基盤を通して提供されるデータの品質管理のガイドブック**」として作成・公表し、**効率的・効果的なデータの品質管理を促進。**

背景・考え方

- スーパーシティ/スマートシティにおけるデータ連携基盤を活用したサービスの提供に当たっては、**データの正確性等の品質が十分でなければ※、サービスの信頼性の低下のみならず、安全・安心を損なう可能性。**

※データが更新されず古いままになっている、データの内容が誤っている 等

- このため、住民の生活の質の向上に資するサービス提供のためには、データ連携基盤から提供される**データの安全性・信頼性の確保（データの品質管理）が重要。**



(データの品質が十分でない場合のイメージ)

ガイドブックの概要

データ連携基盤の関係者が**効率的・効果的にデータの品質管理を行う**ために必要な情報を整理して、以下の参考モデルやツールを提示。

- 「**データ提供者**」・「**データ連携基盤の整備主体**」・「**データ利用者**」が実施すべき事項を示した「**データの品質管理モデル**」
- 客観的指標に基づき、**データの品質評価を行う「簡易評価ツール」**

データの品質管理モデル（関係者の役割）

データ提供者

- ・簡易評価ツールを用いたデータの品質評価
- ・データの提供に加え、品質情報（品質評価結果やメタデータ）をデータ連携基盤の整備主体へ提供

データ連携基盤の整備主体

- ・データの品質管理規程の策定
- ・データ提供者から受領したデータセットや品質情報等の確認と公開
- ・データの品質に関する改善要求の受付

データ利用者

- ・データを活用したサービスの実施
- ・データの品質に関する改善を要求

簡易評価ツール

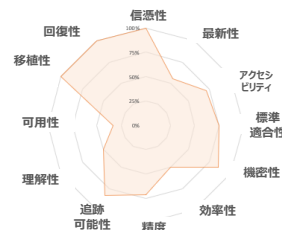
関係する3者が共通の理解をもって、データの品質管理を行うためのツール

評価結果イメージ

【**基礎的品質特性の評価**】
正確性・完全性・一貫性の評価

評価結果 **A/B/C**

【**付加的品質特性の評価**】
以下12の品質特性に関連するメタデータの記入率を表示

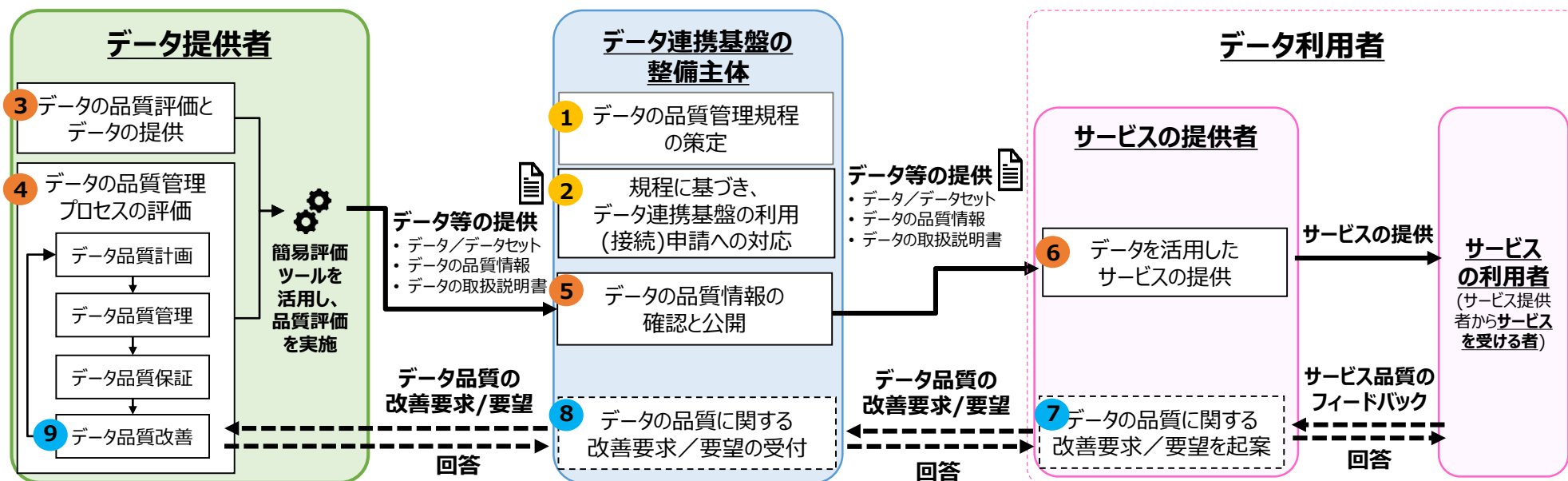


データ連携基盤を通して提供されるデータの品質管理モデル

- データの品質管理を効率的・効果的に推進するため、データ提供者、データ連携基盤の整備主体、データ利用者の3者が各プロセスにおいて実施すべき事項を示したデータの品質管理モデルを提示。
- データ連携基盤の整備主体が整備した品質管理に関するルールに基づいて、データ提供者がデータの品質評価・提供を行い、データ連携基盤の整備主体が品質情報を確認・公開するとともに、データ利用者による改善要求に関する受付窓口を整備することで、継続的なデータの品質管理を推進。

データの品質管理モデル

データ連携前	データ連携基盤の整備主体が、品質管理規程を策定し、品質情報の公開等に関するルールを整備。 ① ②
データ連携時	データ提供者は、簡易評価ツールを用いてデータの品質評価等を行い、データの品質情報等を提供する。 ③ ④ データ連携基盤の整備主体は、それらを確認した上でわかりやすく公開し、データ利用者の活用を促す。 ⑤ ⑥
品質改善要求時	継続的なデータの品質向上のため、データ利用者による要求／要望を受け付ける体制を整備する。 ⑦ ⑧ ⑨



簡易評価ツールによるデータの品質評価

- データ連携基盤を通して提供されるデータの品質管理においては、データ提供者のみならず、データ連携基盤の整備主体やデータ利用者も含めた複数の者が**共通認識のもとに品質評価結果の確認を行うことが重要**。
- このためデータの品質管理モデルの中では、データ連携時に**データ提供者が客観的指標に基づいてデータの品質評価を行うため、「簡易評価ツール」の活用を提示**。
- 簡易評価ツールでは、様々なデータの品質特性を「**基礎的品質特性**」と「**付加的品質特性**」に分けて評価することとしており、評価結果をデータ連携基盤上で公開することで、**データ利用者がデータを選択しやすい環境を整備**。

1. 基礎的品質特性の評価

どのような利用シーンにおいても一定以上の高い品質が要求されるデータそのものに関する基本的な品質特性

【対象となる品質特性】
正確性、完全性、一貫性

- データ提供者自身が妥当性検査（バリデーションチェック）ツール等を用いてデータの品質に関するセルフチェックを実施。
- その際、**評価対象としたデータの網羅性や、評価・改善の実施有無をもとに基礎的品質特性をA～Cで評価**。
- データ利用者は、評価結果をもとに、データの信頼性を確認する。

品質評価の実施有無	すべてのデータを対象に評価		一部のデータを対象に評価	
	品質改善済	品質改善未対応	品質改善済	品質改善未対応
品質評価実施	A(最大)	C	B	C
品質評価未実施	品質評価未実施 C			

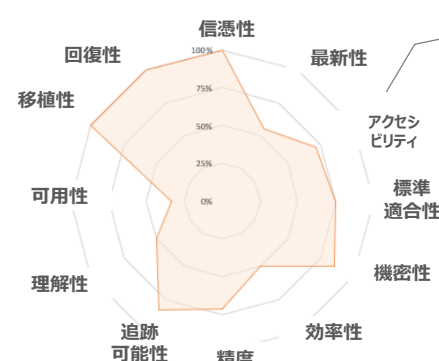
2. 付加的品質特性の評価

利用シーンに応じて求められる品質水準が異なる、データの取得方法等に関する付帯的な品質特性

【対象となる品質特性】
信憑性、最新性、アクセシビリティ、標準適合性、機密性、効率性、精度、追跡可能性、理解性、可用性、移植性、回復性

- **品質特性と関連するメタデータの記入率をもとに付加的品質特性を評価**。
- データ利用者は、以下の評価結果とメタデータの記述内容をもとに、必要とする品質水準を有するデータを選択し、活用する。

付加的品質特性ごとのメタデータの記入率



各品質特性に関連する**メタデータの記入率**※を一覧形式で表示。

※ 例えば、ある品質特性の記入率が100%である場合には、その品質特性に関する**メタデータが全て記入されていることを示す**。
(データ品質そのものの優劣を表すものではない)

<品質特性と関連したメタデータの例>

- 「最新性」に関連するメタデータ
 - ・ データセットの作成日
 - ・ データセットの更新日
 - ・ バージョン情報 等

概要

評価方法

※ 本ガイドブックでは、ISO25021で規定する15の品質特性をもとに、基礎的品質特性と付加的品質特性に分類。