

データの利活用推進に向けた  
技術ガイドライン  
Ver1.0.0

2023年3月



# 目次

1. ガイドライン策定の背景及び目的 .....	2
1.1. 本ガイドラインの構成.....	2
1.2. 行動指針の策定.....	2
1.3. データ利活用の意義及び必要性.....	3
1.4. ガイドラインの目的.....	4
1.5. ガイドラインの対象.....	5
1.6. ガイドラインの更新について.....	7
1.7. 各事業の手順書・マニュアル等について.....	7
2. データ利活用に向けたデータ整備の方針 .....	9
2.1. 対象となるデータ.....	9
2.2. データの品質.....	12
2.3. データの標準化.....	16
3. データ利活用に向けたデータ整備の実践方法.....	18
3.1. 実践方法の概要.....	18
3.2. システム開発・運用に関わる職員.....	18
3.3. 一般職員.....	19

# 1. ガイドライン策定の背景及び目的

## 1.1. 本ガイドラインの構成

第1章（本章）では、本ガイドライン策定の背景や目的、スコープ、位置づけを示す。

第2章では、本ガイドラインのスコープであるデータ品質とデータ標準化を概観するとともに、都における基本的な方針を示す。

第3章では、第2章で示した方針を実践する上での具体的な運用プロセスを示す。

また、付録として、以下を用意した。

- チェックリスト：データ品質をチェックするためのツール。詳細は、第3章にて説明
- 事例概要版、事例詳細版：各部局が本ガイドラインを使用する際の参考として、庁内におけるデータ利活用の事例をまとめた資料
- データ品質の評価項目：本ガイドラインでは、紙面の関係上割愛した評価項目を説明した資料
- ガイドラインに係る手順書\_マニュアル等：都に存在するデータに関連する手順書・マニュアル等の一覧をまとめた資料

なお、本ガイドラインは、今後の実践を通じて得られた知見などを踏まえつつ、随時内容の見直しや改定していく予定です。

## 1.2. 行動指針の策定

都の提供する様々なサービスにおいて、品質の高いデジタルサービスを安定的かつ永続的に提供し都政の QOS 向上に寄与することが、サービス提供者としての都のミッションである。

その実現のためには、都の職員全員が当事者として自覚を持って QOS の高いデジタルサービスの提供を目指していく必要がある。2022年3月には、そのための基本的な価値観を共有できるよう、デジタルサービスの開発・運用に携わる全ての職員等が遵守すべき基本的な価値観等を示す「東京都デジタルサービス開発・運用に係る行動指針」（以下「行動指針」という。）を策定した。

行動指針は、「行動規範（デジタル10か条）」と「機能別技術ガイドライン」（以下「ガイドライン」という。）から構成される。

「行動規範」は、デジタルサービスの開発・運用に携わる全ての職員が遵守すべき基本的な理念として以下の10か条を規定したものである。

### 【行動規範（デジタル10か条）】

- ① 顧客視点でデザインしよう

- ② シンプルなサービスを心がけよう
- ③ 誰ひとり取り残されないようにしよう
- ④ 資源（データ）を最大限に活用しよう
- ⑤ 安全安心なデジタル社会をつくろう
- ⑥ オール東京一丸となって取り組もう
- ⑦ 都政の見える化をしよう
- ⑧ 都民と共創しよう
- ⑨ つねに見直し、チャレンジし続けよう
- ⑩ ともに学びつづけよう

また、「ガイドライン」は、「行動規範」の実践に当たって必要な技術的な基準についてまとめたものである。本ガイドラインは主に「行動規範」の「4.資源（データ）を最大限に活用しよう」を実践するためのもので、データ利活用について規定している。

### 1.3. データ利活用の意義及び必要性

データは利活用することによって様々なことが可能となる。例えば、データから特定のサービスの効果や課題を検証・分析し、そこで得られた知見からより質の高いサービスへとつなげていくことや、システムログから利便性の低いシステムを抽出し、ユーザーの負荷を軽減するための改善策を立案すること等が期待されている。

また近年、行政においては「エビデンスに基づく政策立案（Evidence Based Policy Making）（以下「EBPM」という。）<sup>1</sup>」と呼ばれる、統計や業務データなどの客観的な証拠に基づいて政策立案を行うようなデータ利活用も注目を集めている。我が国では、これまで上記のようなデータが十分に活用されず、勘や経験によって政策が決まってしまう側面が課題として指摘されてきた。EBPMによって論理的かつ客観的に政策の立案や効果検証等が実施されるようになれば、そうした課題も解消されることが期待されている。

このようにデータ利活用によって様々なことが可能となるが、そのためには、データを利活用しやすい環境を整えることが重要である。また、データ利活用の留意事項として、事業推進の判断や効果の検証にどのようなデータが必要となるかを事前に十分検討すること、データ利活用による効果を検証

---

<sup>1</sup> 内閣官房によれば、エビデンスに基づく政策立案（Evidence Based Policy Making：EBPM）は、次のように定義される。EBPMとは、政策目的を明確化させ、その目的のため本当に効果が上がる行政手段は何かなど、政策手段と目的の論理的なつながりを明確にし、そのつながりを裏付けるようなデータ等のエビデンス（根拠、証拠）を可能な限り求め、「政策の基本的な枠組み」を明確にする取組。詳細は、「EBPMガイドブック」参照。（[https://www.gyokaku.go.jp/ebpm/img/guidebook1.0\\_221107.pdf](https://www.gyokaku.go.jp/ebpm/img/guidebook1.0_221107.pdf)）

できるようにするために利活用前の状態を示すデータを蓄積しておくことも重要である。

国においてもデータ利活用の推進のために、デジタル庁が「包括的データ戦略」を令和3年6月に公表した。この戦略では、「フィジカル空間（現実空間）とサイバー空間（仮想空間）を高度に融合させたシステム（デジタルツイン）を前提とした、経済発展と社会的課題の解決を両立（新たな価値を創出）する人間中心の社会」をビジョンとして掲げ、それを実現する方針として、データを利活用する環境の整備等について言及している。

取組の一例を挙げると、新型コロナ危機における各種の届出や申請等に係るサービス提供の混乱等を踏まえ、サービスのワンストップ・ワンズオンリー化に向けて、公的機関提出済みの情報等を参照できる形で整備した「ベース・レジストリ」の取組に関する検討が具体的に進められている。社会的ニーズ、経済効果、即効性の観点から、どのようなデータが整備されるべきか等、具体的な整備に向けた調整が開始されているところであり、地方公共団体においても該当するデータの適切な整備が求められている。

都では、こうした国の施策とも整合を取りながらデータを整備し、データを利活用していく必要がある。

#### **1.4. ガイドラインの目的**

本ガイドラインは、データを整備するにあたっての基本的な方針と業務上の遵守事項を示すことで、データを利活用しやすい環境を整え、ひいては庁内におけるデータ利活用やデータに基づく政策立案を促進することを目的としている。

取組の第一歩として、データの品質向上と標準化から着手する。都におけるデータ整備やデータ利活用の取組拡大に伴い、本文書の対象範囲も拡大することを想定している。

## 1.5. ガイドラインの対象

ここでは、「誰が」「いつ」「何を」の3つ観点で本ガイドラインの対象者、対象プロセス、対象項目を整理する。

本ガイドラインの対象者は、都の全職員である。その上で、職員を「システム開発・運用に関わる職員」と「一般職員」に区分する。各区分の定義は図 1-1 本ガイドラインの対象者に示すとおりである。詳細は第3章で述べるが、職員の区分によって、確認すべきチェック項目が異なる。

	区分	定義
都の職員	システム開発・運用に関わる職員	① システムの要件定義、設計、開発、テスト、改修に従事する事業者を選定・発注する職員 ② システムの要件定義、設計、開発、テスト、改修に従事する事業者の成果物を確認及びレビューする職員 ③ システムの要件定義、設計、開発、テスト、改修を行う職員
	一般職員	上記に該当しない全ての職員

図 1-1 本ガイドラインの対象者

対象プロセスは、都で定義した「システム開発の標準プロセス」（図 1-2）とデジタル庁の「データライフサイクル<sup>2</sup>」（図 1-3）を使用して整理する。前者は、システム開発・運用に関わる職員の対象プロセスの整理に、後者は一般職員の対象プロセスの整理に使用する。システム開発・運用に関わる職員については、「要件定義」、「設計」、「改修」及び「廃棄」を対象とする。一般職員については、「データ設計」、「データ収集」、「データ統合」及び「データ廃棄」を対象とする。

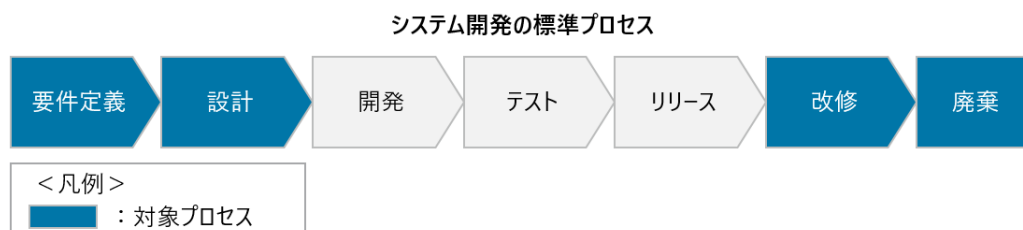


図 1-2 システム開発・運用に関わる職員の対象プロセス  
 （出典：東京都「第 2 回東京デジタルサービス会議 標準プロセスの検討」）

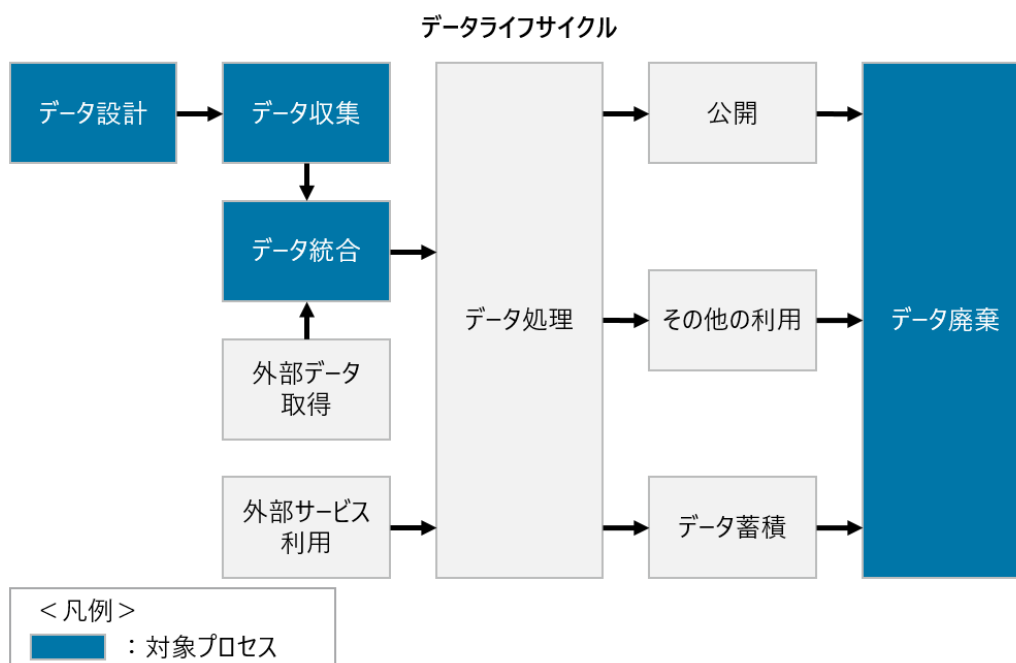


図 1-3 一般職員の対象プロセス  
 （出典：デジタル庁「データ品質管理ガイドブック」）

<sup>2</sup> データライフサイクルとは、データが生成されてから廃棄されるまでのプロセスを表したフレームワーク。各プロセスの概要は、デジタル庁の「データ品質管理ガイドブック」を参照。（[GIF/468-1\\_データ品質管理ガイドブック.docx at main · JDA-DM/GIF · GitHub](https://github.com/JDA-DM/GIF/468-1_データ品質管理ガイドブック.docx)）

対象項目とは、データ整備において検討する項目を指す。データ整備では、マスタデータやメタデータ、データ品質、データアーキテクチャ等、様々な検討が必要である。本ガイドラインでは、「1.5 ガイドラインの目的」でも述べた通り、取り組みの第一歩としてデータ品質と標準化に焦点を当てる。

## 1.6. ガイドラインの更新について

Ver1.0.0 公開日時点では、データ利活用を促進する初期段階にあるため、ガイドラインの対象項目をデータ品質向上と標準化に絞っているが、今後、庁内のデータ整備やデータ利活用の取組状況、さらには政府におけるガイドラインの整備状況に応じて順次対象を拡充していくことを想定している。

## 1.7. 各事業の手順書・マニュアル等について

都には、本ガイドライン以外にもデータに関する手順書・マニュアル等が存在する。それらは、付録「ガイドラインに係る手順書\_マニュアル等」で整理している通りである。本ガイドラインは、データを利活用しやすい環境を整えるに当たっての基礎となる事項を整理したものである。そのため、本ガイドラインで記載している事項には留意しつつも、東京データプラットフォーム（以下、「TDPF」という。）やデジタルツイン、オープンデータ等の各個別事業については、各手順書・マニュアル等を参照されたい。

ここでは、本ガイドラインと関わりの高い 3 つの手順書・マニュアルを示す。各手順書・マニュアルの概要は、表 1-1 に示す通りである。

表 1-1 本ガイドラインと関わりの高い手順書・マニュアル等

No	手順書・マニュアル等の名称	概要
1	東京データプラットフォーム データ整備事業データ整備 マニュアル	民間事業者にデータを利活用してもらえよう、自治体及び民間事業者におけるデータ整備の方法手順を示したマニュアル。画像や PDF、Excel といったファイル形式ごとに、機械判読可能な csv 形式にデータを変換する手順を示している。
2	デジタルツイン基盤データ 整備手順書	TBD
3	東京都オープンデータカタ ログサイト整備マニュアル	東京都オープンデータカタログサイトにデータを掲載する行政職員に向け



		て、オープンデータの取組に関する説明や具体的な東京都オープンデータカタログサイトへのデータ掲載方法を示している。
--	--	--

## 2. データ利活用に向けたデータ整備の方針

### 2.1. 対象となるデータ

ここでは、本ガイドラインにおいて対象となるデータを具体的なデータ例とともに示す。

本ガイドラインにおいて対象となる「データ」は、主に二次利用やコンピュータ処理が可能といった機械判読等に適するよう変換されたデータ<sup>3</sup>とする。

次頁の表 2-1 では、対象データの具体例を示している。ここでは、都における代表的な業務を分類し、それぞれの業務で生成されるもののうち、本ガイドラインにおいて対象となるデータを例示している。企画書や報告書、議事録のような文字列で構成される資料は含まれていない。これらの資料は、コンピュータがデータの論理的な構造を識別できるような形式で整理することが難しく、機械判読に適さない性質をもつためである。

ここで対象としているデータについては、「2.2 データの品質」以降で示す遵守事項に沿いながら、データ品質向上や標準化を進めていく。

---

<sup>3</sup> 機械判読に適したデータとは、コンピュータプログラムが自動的にデータを加工や編集できる形式となっているデータを指す。データの加工や編集をするためには、コンピュータがデータの論理的な構造を識別でき、かつ構造の中の値が処理できるようになっていることが必要である。詳細は、「総務省 情報通信白書平成 25 年版」参照 (<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h25/html/nc121210.html>)

表 2-1 本ガイドラインの対象となるデータ例

業務分類		左記業務から生成されるデータ
定常業務	都管理施設・財産等の管理運営	管理簿・業務/実績記録(都が直接記録するもの)
		記録等集計情報/業務システムログ
	届出・認可・個別の点検や測量・指導(主として法令規則に基づく、個人・事業者への個別の規制・執行等)	申請・届出・利用者情報(都が直接受領するもの)
		申請・届出等集計情報/業務システムログ
	法定事務遂行上の、国・市町との調整・とりまとめ	申請・届出・利用者情報(基礎自治体等から受領するもの)
		申請・届出等集計情報/業務システムログ
		管理簿・業務/実績記録(基礎自治体等から受領するもの)
		記録等取りまとめによる集計情報
	その他全体向け法定事務の執行/専門サービス	業務/実績記録
		利用者(サービス受益者)情報
業務記録等集計情報/業務システムログ		
企画・調査業務	事業企画	各種根拠資料(意思決定用の調査・分析等)
	都民・事業者との調整・相談・助言、会議体の形成	相談等実績(積み上げ値)
	補助・交付金等各種支援事業	補助・交付実績
	調査事業	調査帳票
集計データ		
管理業務	基幹システム運用	システム稼働状況・ログ(アクセスログ含む)
	内部管理(人事・予算・組織・文書・法制・監査・出納・総務事務等)	内部履歴

## コラム <構造化データと非構造化データ>

データは、一般的に構造化データと非構造化データの2つに分類できる。構造化データとは、データの並べ方が規定されたもので、例えば Excel 等で作成した集計データ等が該当する。データが構造化されていると、コンピュータがデータを処理しやすくなり、複雑な処理も高速で実行できるようになる。

一方、非構造化データとは、データの並べ方が規定されていないもので、例えば、Word や PowerPoint 等で作成した各種文書をはじめ、写真、動画、音声等が該当する。

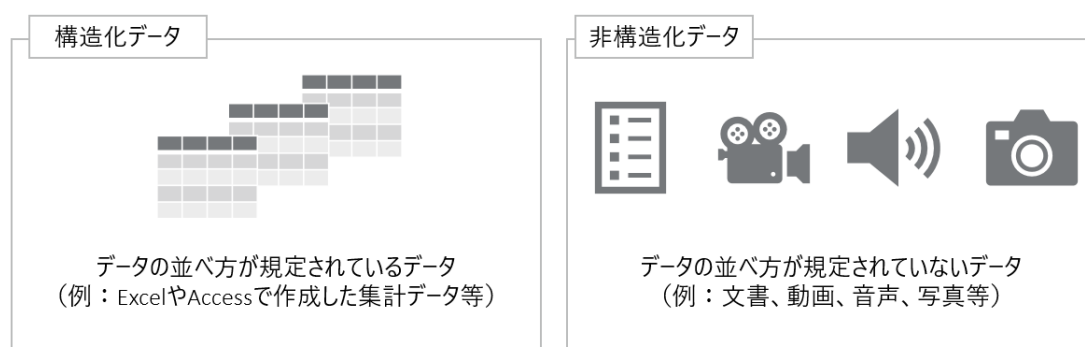


図 2-1 構造化データと非構造化データ

## 2.2. データの品質

データの品質に問題があると、それを加工したデータの品質もその影響を受け、結果としてデータを利用したサービス自体の品質も低下してしまう。本節では都におけるデータの適切な管理に向け、デジタル庁の「データ品質管理ガイドブック<sup>4</sup>」に基づくデータの品質及びその管理の考え方を示す。

### ● データ品質を管理する価値

低い品質のデータは社会的非効率をもたらす。データの利活用においてデータの提供者がデータの改善作業を怠り、低い品質のデータを配布した場合、個々のデータ利用者が改善作業に取り組むこととなり、多くの時間が費やされることになる。「データ品質管理ガイドブック」では、適切な品質管理がもたらす価値を以下のように整理している。

- データ収集コストを低減できる
- データ収集を迅速化できる
- データ更新を容易にできる
- データ更新にまつわる問題を回避できる
- 内部でのデータ活用を容易にできる
- データ公開を容易にできる
- 利用者側でのデータ活用が進む

### ● データ品質の管理方針

前項で示したデータ品質管理の価値は、特定の時点で行うだけでなく、継続的に品質の改善を図っていくことで発揮されうるものである。図 2-2 は「データ品質管理ガイドブック」のデータ品質管理フレームワークを基にした都におけるデータ品質管理サイクルである。

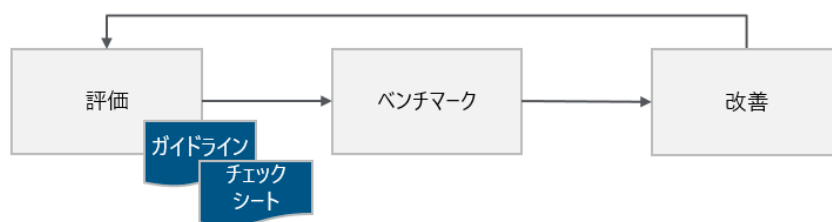


図 2-2 データ品質管理サイクル

このサイクルの「評価」では、本ガイドラインとチェックリストを用いてデータ品質をチェックする。その後、「改善」では、必要に応じて他組織や

<sup>4</sup> デジタル品質管理ガイドブック：2022年3月にデジタル庁が公開した、データ品質向上を目的とした文書。詳細は、下記よりダウンロード。

[GIF/468-1\\_データ品質管理ガイドブック.docx at main · JDA-DM/GIF · GitHub](#)

過去のデータ品質の評価をベンチマークとして比較しながら改善案を検討及び実施する。このサイクルを繰り返し、データ品質を向上させる。

また、データの品質は高ければ良いとは限らない。品質を高めるにも時間や費用などといったコストがかかるからである。利用目的に応じた適切なデータ品質を事前に決めることが重要である。

さらに、データの提供者と利用者間で品質の基準を共有することも重要である。データ提供者が高い品質のデータを提供しているつもりでも、利用者にとっては十分な品質でないケースや、その逆のケースも発生しうる。データの提供者と利用者間で、品質の考え方について認識を合わせることが望ましい。

## ● データ品質の評価項目

「データ品質管理ガイドブック」では、データ品質に関する 15 個の評価項目を示している。図 2-3 で 15 項目を概観する。

「データ品質管理ガイドブック」におけるデータ品質の評価項目（全15項目）					
<b>正確性</b>	誤脱等がなく、データが正しいこと	<b>精度</b>	精度が適切に設定されており、データ内で統一されていること	<b>移植性</b>	データの移植が可能なこと
<b>完全性</b>	データに空欄がなく、網羅的であること	<b>標準適合性</b>	標準的なルールに適合していること	<b>回復性</b>	データの早急な復元が可能なこと
<b>一貫性</b>	データ内の項目に矛盾がないこと	<b>理解性</b>	利用者がデータについて十分理解可能であること	<b>機密性</b>	暗号化等の対策が行われていること
<b>可用性</b>	データが利用可能な状態にあること	<b>アクセシビリティ</b>	データが誰でも使える状態にあること	<b>効率性</b>	コードの割当等により効率的に扱えること
<b>最新性</b>	データが新しく維持されていること	<b>信憑性</b>	出所等が信頼されていること	<b>追跡可能性</b>	データの原典等を参照できること

■ 本ガイドラインで定める重点 5 項目

図 2-3 データ品質の評価項目（全 15 項目）

本ガイドラインでは、このうちデータの中身そのものに関する「正確性」、「完全性」、「一貫性」、「可用性」、「最新性」の 5 項目を重点項目として位置付ける。以下、「データ品質管理ガイドブック」の引用を基に、これらの項目の概要と具体的な評価項目を示す。

### ① 正確性

データの基本であり、データの正しさを表すもの。データが正確でなければ、判断に誤りが生じる恐れがある。

➤ 書式が正しいか

- 誤字脱字などはないか
- 意味的な誤りがないか
- データに誤りはないか

## ② 完全性

データが完全であること。データは目的に応じて抜け漏れなくあることで、詳細な分析をすることが可能となる。

- 用途に応じて必要な項目が網羅されているか
- 必須項目に空欄が含まれていないか

## ③ 一貫性

データに矛盾がないこと。データには整合性や一貫性が必要で、データ内の項目や値に矛盾があると自部署内において、連続性のある利活用ができない(エラー処理が必要となる)。

- データセット<sup>5</sup>内でデータに矛盾はないか
- データセット間でデータに矛盾はないか

## ④ 可用性

データが利用(アクセス)可能な状態にあること。データは必要な時に使えるようになっていることが重要である。

- 必要な時にいつでもデータにアクセスできるようになっているか
- データを公開するシステムは常時稼働しているか

## ⑤ 最新性

データが十分に新しいものに維持されていること。データが最新版に更新されていることで、誤処理や再処理が不要となる。

- 公開データの更新サイクルは元データの更新サイクルに対して適切か
- データは収集時から十分に短い期間で公開されているか
- 最終更新日時及び最新版の所在等、更新版の有無が確認可能か

---

<sup>5</sup> データセットとは、何かしらの目的の下で収集したデータを1つの表やファイル等にまとめた、データの集合を意味する。

## コラム <天気データと電車ダイヤグラムデータ>

「2.2.データの品質」では、データの品質は高ければ良いとは限らないことを述べた。ここでは、このことを天気データと電車のダイヤグラムデータを例に概説する。

天気データは、正確であるほど便利ではあるが、一定以上の正確性は必要ない場合が多い。例えば、翌日の天気が晴れだと予報されていたところが、実際は曇りだった場合、それが人々の生活に重大な影響を与える可能性は低い。

一方で、電車のダイヤグラムデータには高い正確性が求められる。ダイヤグラムが間違っていた場合、乗り換えに問題が生じたり、目的地へ予定していた時刻に到着できなくなったり等、多くの人々に影響が生じる。

このようにデータには、天気データのように一定以上の正確性が担保されていけば問題ないデータもあれば、電車のダイヤグラムのように常に高い品質が求められるデータもある。データに必要な品質の高さについては、データの利用者と提供者の間でしっかりと認識を合わせ、必要な品質でデータを準備し、提供することが重要である。



## 2.3. データの標準化

ここでは、データ標準化の必要性及び効果並びに都における方針を示す。

データの利活用という観点では、データの品質を向上させるだけでなく、データを標準化することで複数のシステム間でデータを連携可能にすることも重要である。あるシステムのデータを他のシステムに連携または移行しようとした時に、データの項目名や表記形式、定義が異なっていると、データの処理に多大なコストがかかるばかりでなく、データの処理エラーが原因となって重大なシステム障害につながることもある。

データが同じデータ項目名で、同じ意味や定義で設計されていると、スムーズかつ正確にデータ連携を行うことができる。特に国際標準や政府標準を活用することで、より広い範囲でのデータ流通及びデータ利活用を促進することが可能になる。

都の保有するデータの標準化を推進することは、都のデータと外部のあらゆるデータを連携しやすくすることにつながり、利活用の幅を大きく広げることができる。

データの標準化に取り組むことによる具体的な効果は様々であるが、ここでは、デジタル庁が2022年5月9日に公表したデータ標準化を進めるための技術体系「政府相互運用性フレームワーク（Government Interoperability Framework）（以下、「GIF」という。）」を基に5つの効果を概説する。

### ① 相互運用性の向上（連携の容易さ）

同じデータを表すコードや名称等を統一されたルールに基づいて標準化することで、別々の組織や部局によって構築されたシステムであっても、複数システム間での連携が容易になる。

### ② 拡張性の向上

一般にシステムは、様々なモジュール（部品）の組み合わせで構成されている。データを標準化することで、これらのモジュール間でのデータの受け渡しをよりシンプルに実装できるため、開発の工数を短縮するとともに、既存の機能の改修や新しい機能の追加も柔軟に可能になる。

### ③ 設計コストや時間の削減

標準として定められているデータモデルを参照してデータを設計できるため、システム開発における設計にかかる工数の短縮や、運用コストの低減をはかることができる。

### ④ 設計や運用の高度化

③で短縮された時間をサービスや運用の高度化の検討に使うことができ

る。

⑤ ワンスオンリー、ワンストップの実現

複数システム間でデータがスムーズに連携されることで、行政サービス利用者が何度も同じ情報を書いたり入力したりしなくても良い「ワンスオンリー」の実現につながる。同様に、複数サービス横断のワンストップサービスも実現可能になる。

デジタル庁が公表している GIF では、行政におけるシステムやサービスの開発において汎用的に利用可能なデータモデルやデータの表記ルールを提供している。都でも GIF を参照しながら、システムやサービスの改修や新規開発の際の仕様に含める等、標準化が可能なものから取り組みを進めていくことが望ましい。

データの標準化と関連する重要な項目の一つに、ベース・レジストリが挙げられる。ベース・レジストリは世界の各国・各都市で構築が進められており、デジタル庁の定義に基づくと「公的機関等で登録・公開され、様々な場面で参照される、人、法人、土地、建物、資格等の社会の基本データ」を指すものである。都では、国のベース・レジストリとの連携を見据え、国の方針や仕様と足並みを揃える必要がある。ベース・レジストリの構築も GIF への準拠を前提に進められているため、都の保有データを GIF に準拠する形で標準化しておくことで、今後の国との連携をスムーズに進められることが期待できる。

## 3. データ利活用に向けたデータ整備の実践方法

### 3.1. 実践方法の概要

ここでは、「2.2 データの品質」で説明したデータ品質を業務の中でどのように確認していくかを示す。本ガイドラインは、デジタル庁が整理したデータ品質の15項目のうちデータの中身に関する「正確性」、「完全性」、「一貫性」、「可用性」、「最新性」の5つを重点項目として位置付けている。そのため、以降では、これらの観点でチェック項目を整理している。

チェック項目の確認には、付録「チェックリスト」を使用し、その結果を元に改善を行う。なお、業務の性質が異なるため、システム開発・運用に関わる職員と一般職員とで、チェックリストを分けている。

### 3.2. システム開発・運用に関わる職員

システム開発・運用に関わる職員については、都の「システム開発の標準プロセス」を使用して、各プロセスにおけるチェック項目を示す。

自身がシステム開発におけるどのプロセスに従事しているかを把握した上で、該当するプロセスのチェック項目を確認する。ここでは「要件定義」、「設計」、「改修」の3つのプロセス<sup>6</sup>におけるチェック項目を示す。

#### ● 要件定義、設計、改修

本プロセスにおけるチェック項目は、表3-1に示す通りである。

同じ検討事項でも、開発するシステムによっては、要件定義で検討する場合と設計で検討する場合がある。どちらのプロセスで検討するかは、開発者の判断に委ねられるため、ここでは要件定義と設計を分けずにチェック項目を整理した。

また、改修については、システムの要件定義からリリースまでのプロセスを再度実施するプロセスのため、チェック項目は要件定義と設計と同じ項目となる。

---

<sup>6</sup> 「開発」と「テスト」は、開発とテストを通して設計の内容がシステムに反映されているかを検証するプロセスである。そのため、これらのプロセスにおける新たなチェック項目はない。「リリース」も、ユーザーがシステムを利用するプロセスのため、チェック項目はない。また、「廃棄」は本ガイドラインの対象プロセスであるものの、5つ重点項目とは関連性が薄いため、チェック項目を設けていない。

表 3-1 システム開発・運用に関わる職員のチェック項目

No	分類	チェック項目
1	正確性	書式を指定できるものについて、入力欄の書式を制御しているか
2	正確性	入力する内容があらかじめ決まっているものについて、入力できる内容をプルダウン形式などで制御しているか
3	正確性	明らかな誤入力を制御しているか
4	完全性	利用目的と照らした時に必要なデータ項目が網羅されているか
5	完全性	必須項目に入力漏れがある場合、エラーが表示されるか
6	一貫性	システム内で一意に主キーを設定しているか
7	一貫性	システム内で入力データに矛盾がないように制御しているか
8	一貫性	システム間でデータに矛盾がないように制御しているか
9	可用性	必要な時にいつでもデータにアクセスできるようになっているか
10	可用性	データを公開するシステムは常時稼働しているか
11	最新性	公開データの更新サイクルが元データの更新サイクルに対して適切に設定されているか
12	最新性	データの特徴を踏まえ、データが収集時から十分に短い期間で公開されているか
13	最新性	ファイル等で提供される場合は、最終更新日時及び最新版の所在が明記されているなど、更新版の有無が確認できるようになっているか

### 3.3. 一般職員

一般職員は、デジタル庁の定義したデータライフサイクル<sup>7</sup>を使用して、各プロセスにおけるチェック項目を示す。

自身がデータライフサイクルにおけるどのプロセスに従事しているかを把握した上で、該当するプロセスの各チェック項目を確認する。ここでは「データ設計」及び「データ収集」の2つのプロセス<sup>8</sup>におけるチェック項目を示す。

一般職員については、業務の中で都度データ品質のチェック項目を確認す

<sup>7</sup> データライフサイクルとは、データが生成されてから廃棄されるまでのプロセスを表したフレームワーク。概要については、第1章5節を参照。

<sup>8</sup> 「データ統合」及び「データ廃棄」は本ガイドラインの対象プロセスであるものの、5つの重点項目とは関連性が薄いため、チェック項目を設けていない。

る運用は現実的ではないため、チェック項目を確認するのは、「データ設計」のみとする。その他のプロセスでは、本ガイドラインで示した方針を意識することを前提に、チェック項目の確認は任意とする。

## ● データ設計

データ設計は、データの収集、管理及び公開等の後続プロセスの基盤をつくるプロセスである。例えば、Google Form 等でのアンケートの作成や Excel 等での調査票や集計表の作成が挙げられる。本プロセスでは、以下の項目をチェックする。

なお、本ガイドラインでは対象のデータになっていないが、調査書や報告書のような PowerPoint や Word 形式の資料についてもチェック項目を確認することが望ましい。

表 3-2 一般職員のチェック項目（データ設計）

No	分類	チェック項目
1	正確性	アンケートや調査票を作成する際に、書式を指定できるところについて、入力欄の書式を指定しているか
2	正確性	アンケートや調査票を作成する際に、入力する内容があらかじめ決まっているものについて、入力できる内容をプルダウン形式などで制約しているか
3	正確性	アンケートや調査票を作成する際に、明らかな誤入力を防ぐ設計にしているか
4	完全性	アンケートや調査票を作成する際に、利用目的と照らして必要なデータ項目が網羅されているか
5	完全性	必須項目に入力漏れがある場合、入力漏れがわかるような仕組みになっているか
6	一貫性	同一ファイル内でデータの矛盾はないか
7	一貫性	関連するファイル間でデータに矛盾はないか
8	可用性	必要な時にいつでもデータにアクセスできるようになっているか
9	最新性	公開データの更新サイクルが元データの更新サイクルに対して適切に設定されているか
10	最新性	データの特徴を踏まえ、データが収集時から十分に短い期間で公開されているか
11	最新性	ファイル等で提供される場合は、最終更新日時及び最新版の所在が明記されているなど、更新版の有無が確認できるようになっているか

## ● データ収集

データ収集は、データを利用者から収集するプロセスである。システム等にデータを入力する作業は本プロセスに該当する。本プロセスでは、以下の項目をチェックする。

表 3-3 一般職員のチェック項目（データ収集）

No	分類	チェック項目
1	正確性	指定された書式でデータを入力しているか
2	正確性	入力するデータに誤字脱字などはないか
3	正確性	入力するデータに誤りがないか
4	完全性	必須項目に入力漏れがないか