

2025年1月29日 全8頁

## IT関連政策から読み解く日本のDX

# 企業のデータ整備の鍵は「相互運用性」

データ利活用のための企業体制の変革でDXの実現にも期待

経済調査部 研究員 田邊 美穂

### [要約]

- 前回のレポートにて、企業におけるデータ利活用を成功させるためには、「目的」や「データ」を明確化することが重要であると議論した。次に課題となるのは、「目的」に基づき収集した「データ」を、企業全体で効果的に利活用するための適切なデータ整備方法の検討と実施である。
- 政府におけるデータ整備として、国民の利便性向上や行政運営の効率化を目的に、官公庁ごとに管理していた行政データの相互利用を行う、ベース・レジストリの取り組みがある。この取り組みは2023年に一度見直しが行われており、その際にデータの品質確保とデータの公開範囲が課題として指摘されている。
- これらは企業がデータ整備を行う際にも十分に発生しうる課題である。この課題に対する政府の対応を見ると、企業全体でデータを効果的に利活用するためには、データ整備のための社内標準の策定、データ品質を確保し常時提供を可能とする環境および運用体制の構築、適切な公開範囲の特定および制御等、全社的な対応が必要となる。そして、ある程度トップダウンで進める必要があること、データ利活用のために企業内の体制やビジネスモデルの変革が伴うことから、DXの実現も期待される。
- また、データ整備を行うにあたり相互運用性（異なるシステムや組織がデータや情報を円滑に交換し利用できる能力）の確保が重要となる。今後ビジネスにおけるデータ利活用が進むことで、社内外のデータを提供または取得する機会が増加すると考えられる。データの整備手法が一般的な方法と乖離している場合、ビジネスにおいて不利になる可能性がある。この相互運用性を実現する国際的な標準は、技術の変化等に応じて変化していくことも予想される。政府の取り組み等を参考に適宜情報を収集し対応していくことが求められるだろう。

## 1. はじめに

前回のレポートにて、企業におけるデータ利活用を成功させるためには、「目的」や「データ」を明確化することが重要であると議論した。次に課題となるのは、「目的」に基づき収集した「データ」を企業全体で効果的に利活用するための適切なデータ整備方法の検討と実施だろう。企業の現状を見ると、業務効率化やDX推進等のデータ利活用に向けた様々な取り組みが進められているものの、データを全社で利活用出来ている企業はまだ少ない。DX推進には経営判断等へのデータ利活用が求められるため、様々な部署からデータを収集することはもとより、有効活用できる形に整備することが重要となる。

政府においても、国民の利便性向上や行政運営の効率化のため、官公庁ごとに管理していた行政データを相互利用するための整備を進めている。本レポートでは、この政府の取り組みを参考にデータ利活用、すなわちDXを成功させるための企業におけるデータ整備の方法について議論していきたい。

## 2. 政府の取り組みから見たデータ整備

### データの品質確保と公開範囲の設定に課題あり

政府におけるデータ整備の取り組みとして、ベース・レジストリ<sup>1</sup>がある。ベース・レジストリとは、社会の基本データとして住所・所在地、法人の名称等、制度横断的に多数の手続きで参照されるデータからなるデータベースを指す。これらのデータを整備することにより、行政手続きのワンストップ化が促進、国民の利便性の向上や行政運営の効率化が期待される。

このベース・レジストリの取り組みは、2020年にデジタル・ガバメント閣僚会議で公表された「データ戦略タスクフォース 第一次とりまとめ」にて言及され、併せてロードマップが策定された。以降、データの優先度に基づき整備が進められ、国税庁法人番号公表サイトや金融庁EDINET、e-Gov法令検索等、既に整備が完了したものも複数存在する。これらの状況を鑑みると、ベース・レジストリの取り組みは順調に進行しているように見受けられるが、実際には2023年に一度見直し<sup>2</sup>が実施された<sup>2</sup>。課題として挙げられたのは、**データの品質確保**と**データの公開範囲**の2つである（図表1）。

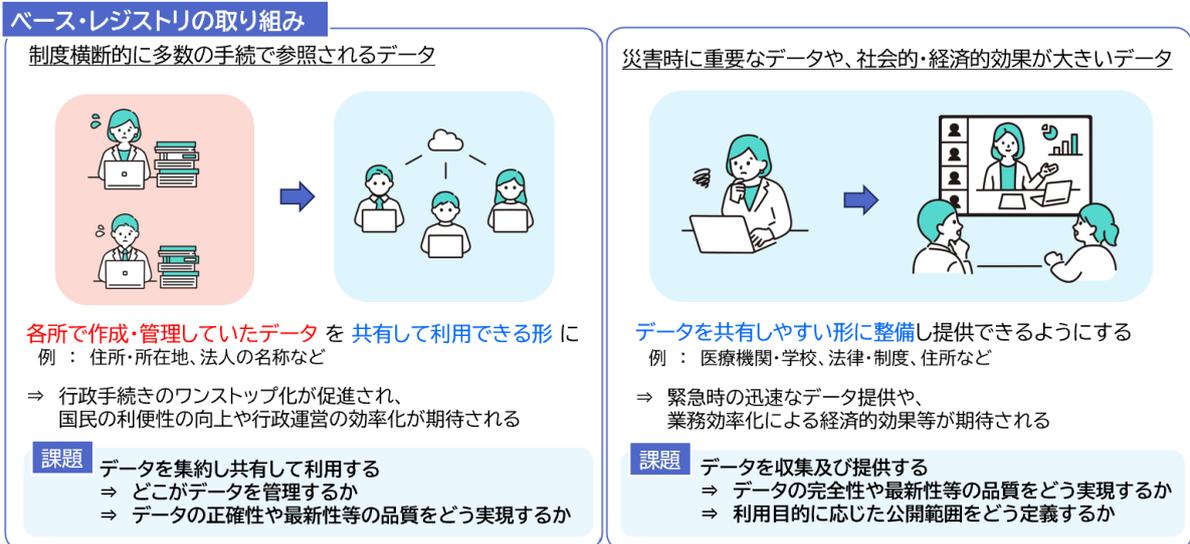
1つ目の**データの品質確保**について、ベース・レジストリは行政手続きをはじめ様々なところからの参照されることを前提としたデータであるという特徴から、常に最新で正確であること等、データに高い品質が求められる。しかし、データの情報源への信頼性やデータ項目ごとの一貫性や運用面の検討に不足が見られた。2つ目の**データの公開範囲**について、ベース・レジストリとして取り扱うデータの中には機密情報が含まれるものも存在する。そのため、その公開範

<sup>1</sup> デジタル庁ウェブサイト「[ベース・レジストリ](#)」

<sup>2</sup> デジタル庁 データ戦略推進ワーキンググループ（第6回）「[ベース・レジストリ、オープンデータの拡充と体制](#)」（令和5年2月28日）

囲には慎重な検討が求められるが、こちらでも検討が不足していることが指摘された。

図表1 ベース・レジストリの取り組みとその課題



➡ データの品質確保と公開範囲について、取り組みの見直しが行われている

(注) データの完全性とは、用途に応じて必要な項目が網羅されている状態を指す。

(出所) デジタル庁 データ戦略ワーキンググループ (第6回) 「[ベース・レジストリ、オープンデータの拡充と体制](#)」(前掲脚注2) より大和総研作成 (イラストはソコスト (<https://soco-st.com/>))

これらは企業がデータを利活用する上でも発生しうる課題といえるだろう。ビジネスの意思決定等に使用する情報には、高い品質(正確性や最新性等)が求められる。その点において、ベース・レジストリで求められることと同様に、企業のデータ利活用においても情報源の信頼性をどのように確保するのか、情報における高い品質をどのように維持するのが重要になる。また、社内で利活用したいデータに個人情報や機密情報が含まれる可能性は十分にあり、全社でデータ利活用を進めるにあたり公開範囲に慎重な検討が求められるのも同様である。

## ベース・レジストリに求められるデータ品質の条件

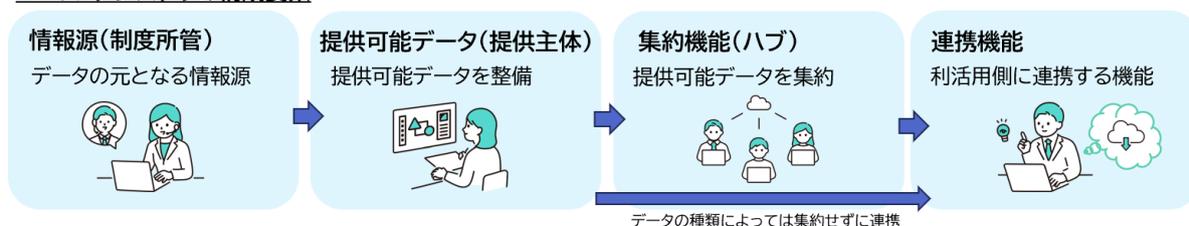
情報源(制度所管)と提供可能データ(提供主体)でデータ品質に求められる役割は異なる

**データの品質確保**について、ベース・レジストリでは、①データ情報源の信頼性確保、②データ自体の品質確保(最新性や標準適合性、可用性等)の両面から見直しを行っている。①データの情報源の信頼性を確保するため、まずベース・レジストリの構成要素を、情報源(制度所管)、提供可能データ(提供主体)、集約機能(ハブ)、連携機能の4つに分類し直した。その上で、データの項目ごとに情報源(制度所管)と提供可能データ(提供主体)を指定している。これにより、データ項目という細かい単位で情報源(制度所管)を明確に示すことが可能になり、利用者のデータに対する信頼性の向上が実現できる。また、データ項目ごとに情報源(制度所管)およ

び提供可能データ（提供主体）という担当が整理されることで、それぞれがデータ品質に対して担うべき責任と役割も明確になる。②のデータ自体の品質確保にも効果が期待される取り組みであるといえる（図表 2）。

図表 2 ベース・レジストリにおける構成要素と見直し例

#### ベース・レジストリの構成要素



#### 上記の構成要素に基づいたベース・レジストリの見直し例

##### 【見直し前】

分野	指定項目	指定データ等	所管部門
法人	会社法人等番号		
	商号(法人名)	商業・法人登記簿	法務省
	本店(所在地)		
	法人番号	法人番号公表サイト	国税庁
	法人資格	Gビジネスフォ	経産省
	決算情報	EDINET	金融庁

##### 【見直し後】

分野	提供可能データ(項目/ファイル)	提供範囲	提供主体	情報源	制度所管
法人基本情報	法人番号	無制限	国税庁	商業・法人登記簿(※1)	法務省
	商号又は名称	無制限	国税庁	商業・法人登記簿(※1)	法務省
	本店又は主たる事務所の所在地	無制限	国税庁	商業・法人登記簿(※1)	法務省
法人決算情報	財務諸表(※2)	無制限	金融庁	有価証券報告書 半期報告書 四半期報告書	金融庁

※1 商業・法人登記簿のうち、法人番号の指定対象の法人に係るものに限る  
 ※2 借借対照表、損益計算書、キャッシュ・フロー計算書を指す

各データ項目の情報源(制度所管)が明確化され、利用者のデータに対する信頼性が向上  
 情報源(制度所管)と提供可能データ(提供主体)が明確になることで、整備内容や取組内容の特定が容易に

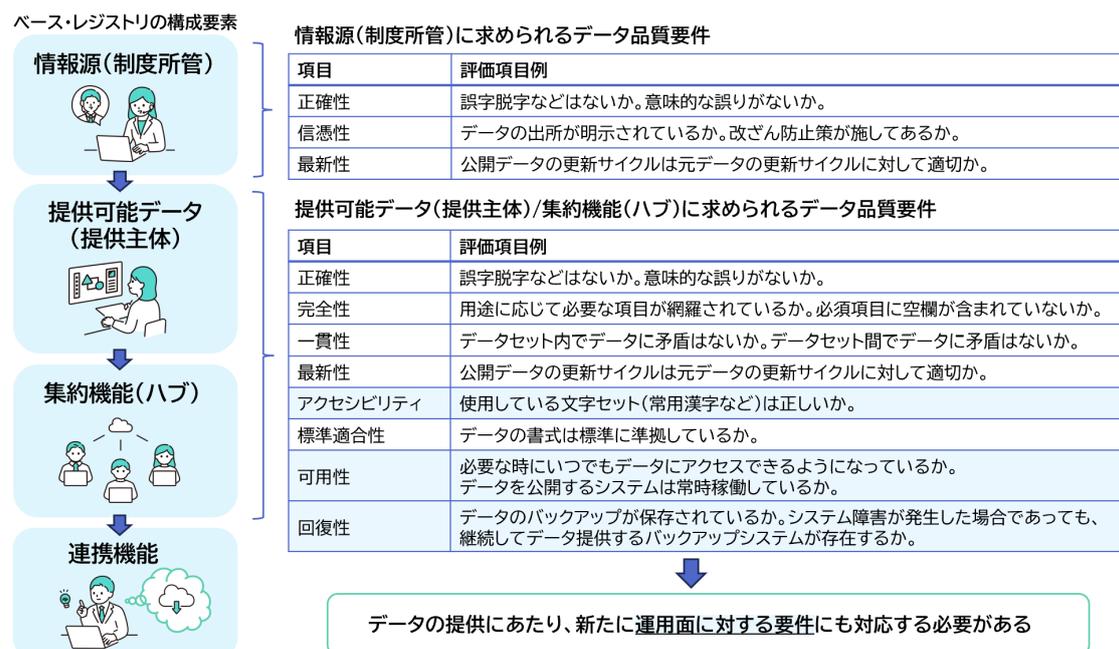
(出所) デジタル庁 データ戦略推進ワーキンググループ (第 6 回) 「[ベース・レジストリ、オープンデータの拡充と体制](#)」(前掲脚注 2) より大和総研作成 (イラストはソコスト (<https://soco-st.com/>))

ここで、②データ自体に求められる品質とは具体的に何を指すのかを見てみる。まず、情報源(制度所管)には、データの**正確性**や**最新性**、**信憑性**が求められるだろう。広く利用されるデータの源泉になることから、データに意味的な誤り等はないか、最新の状態が維持できているか、情報の根拠は信頼に値するか等、慎重な検証が求められる。次に提供可能データ(提供主体)と集約機能(ハブ)には、**完全性**や**正確性**、**最新性**、**可用性**、**回復性等**が求められるだろう。データを提供できる形に整形/集約するにあたり、情報源から欠損や誤りは発生していないか(完全性・正確性)、情報源からの更新情報が反映され最新の状態が維持できているか(最新性)といったデータ自体の検証は、情報源(制度所管)と観点は異なるが必要となる。そして、データの提供が常に可能(可用性)で、問題を検知や復旧できる仕組みが整っているか(回復性)等の運用面も対応が必要だ(図表 3)。データ整備だけでなく、運用面まで対応が求められる提供可能データ(提供主体)等の負担は大きい。そのため政府では、提供可能データ(提供主体)等に求められる運用体制自体を、データ整備に知見のある独立行政法人等に協力を求めることも検討されている<sup>3</sup>。

<sup>3</sup> デジタル庁 データ戦略推進ワーキンググループ (第 7 回) 「[ベース・レジストリの運用体制について](#)」(令和 5 年 6 月 6 日)をはじめ、同ワーキンググループにてたびたび議論されている。

企業においても、データの出所となる部署（情報源）が、データ提供のための運用まで担うことは現実的に難しい。ベース・レジストリと同様に、提供主体となる部門の設立や、データ整備に知見のある部署に協力を求めることが効果的だろう。独立行政法人情報処理推進機構（以下、IPA）では、データ利活用のための組織整備として CDO（Chief Data Officer：最高データ責任者）の擁立と CDO 室の設置等を、データ運用体制の一つの例として推奨している。

図表 3 ベース・レジストリに求められるデータ品質



(出所) デジタル庁 データ戦略推進ワーキンググループ (第 6 回)「[ベース・レジストリ、オープンデータの拡充と体制](#)」(前掲脚注 2)、「[データ品質管理ガイドブック \(β 版\)](#)」(2021 年 8 月 31 日)より大和総研作成 (イラストはソコスト (<https://soco-st.com/>))

## データの整備には標準化や相互運用性の確保が必要となる

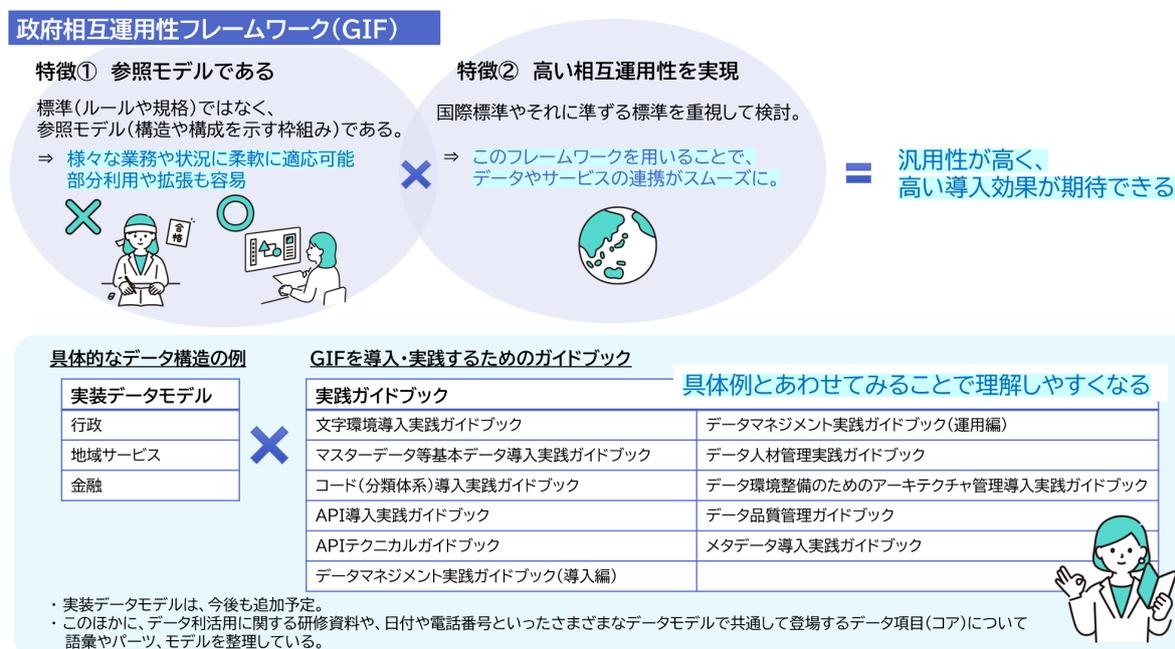
データの品質確保における情報源と提供主体等に求められる役割が明確になったところで、次の課題は具体的な整備方法だろう。例えば、前掲図表 3 で挙げた標準適合性では、「データの書式は標準に準拠しているか」とあり、何らかの標準に基づいてデータを整備する必要があることが分かる。この標準について、ベース・レジストリでは「政府相互運用性フレームワーク (Government Interoperability Framework、以下 GIF)」の考え方をベースに整備を進めている。

GIF とは厳密には標準 (ルールや規格) ではなく参照モデル (標準的な構造や機能を示す枠組み) であり、この参照モデルをもとにデータ整備を行うことで、データを相互に連携し活用できる環境構築を目指す。そのため GIF には、行政や地域サービス等における具体的な実装データモデルや、この参照モデルを導入および実践するためのガイドブックがまとめられている。また、国際標準やそれに準ずる標準を重視して検討されていることから、GIF をデータ整備に利用することで高い相互運用性 (異なるシステムや組織がデータや情報を円滑に交換し利用でき

る能力)の実現も期待される(図表4)。

政府に関連する業務を行う企業であれば、GIFの実装例等を参考に企業内のデータ整備を行うことで、今後政府とのデータ連携が容易になることが期待される。しかし、そうではない企業にとってもGIFを参考にするメリットはある。DFFT(Data Free Flow with Trust:信頼性のある自由なデータ流通)<sup>4</sup>等、国際的にもデータに関する取り組みは急速に進められており、今後ビジネスにおける国際的なデータ取引はますます活発になっていくことが予想される。そのため、国際的な相互運用性が考慮されているGIFのガイドブックや研修資料にて紹介されているデータの構造設計手法等をはじめとする考え方は、企業のデータ整備を考える上でも活用できる部分は多いだろう<sup>5</sup>。

図表4 政府相互運用性フレームワーク(GIF)のイメージ



(出所) デジタル庁「[政府相互運用性フレームワーク\(GIF\)](#)」の「[GitHub\(外部サイト\)](#)」ドキュメント(2024年10月1日更新)より大和総研作成(イラストはソコスト(<https://soco-st.com/>))

## データは可能な限り公開できるように整理することが重要

**データの公開範囲**について、政府は各府省庁が保有するデータは原則オープンデータとして公開することを推進しており、ベース・レジストリもその対象である。ただし、その中には個人情報や機密情報等、公開することが適当ではない情報も含まれる。そのため、ベース・レジスト

<sup>4</sup> DFFTについては、田邊美穂[2024]「[DFFT 実現の決め手はデジタル技術? : 自由なデータ流通における信頼・信用はデジタル技術が鍵となる](#)」大和総研レポート(2024年9月12日)を参照。

<sup>5</sup> IPAは「[相互運用性向上のための参照モデル活用ガイド ~欧州のリファレンスアーキテクチャ\(EIRA\)の紹介~](#)」(2024年12月2日)において、EIF(European Interoperability Framework)の相互運用性確保の考え方は日本企業に活用できるとしている。当レポートで取り上げているGIFは、いわばEIFの日本版であることを考慮すると、GIFにおけるデータ構造の考え方等についても日本企業に活用できる面は多いだろう。

りでは、ファイル単位で公開/非公開を判断する方針から、ファイルの中の項目単位で公開/非公開を特定する方針へ見直しを行っている。これにより、公開範囲を厳密に定めることができるだけでなく、非公開とすべき部分を除き可能な限りオープンデータとして利用することが可能となる。また、オープンデータは国民や企業といった広範にアクセスされるものである一方で、公開/非公開の区分だけではなく、特定の関係者間での共有を図る限定公開のアプローチも推奨されている（図表 5）。企業においても、このように可能な限りデータを社内で共有する方針は、データの利活用において重要な要素となる。オープンデータと比較すると同一企業内という狭い範囲でのアクセスに限られることから、扱う個人情報や機密情報の内容やレベルにあわせて、部署や役職等による共有範囲を制御することも効果的だろう。

図表 5 データの公開範囲の特定イメージ



(出所) デジタル庁「[オープンデータ基本指針の概要](#)」(令和6年7月5日デジタル社会推進会議幹事会決定)より大和総研作成(イラストはソコスト (<https://soco-st.com/>))

### 3. 相互運用性を意識したデータ整備と体制構築により DX の実現を

ここまでの議論より、企業全体でデータを効果的に利活用するためには、データ整備のための社内標準の策定、データ品質を確保し常時提供を可能とする環境および運用体制の構築、適切な公開範囲の特定および制御等、全社的な対応が求められるものが多い。つまり、これらの取り組みはある程度トップダウンで進める必要があり、このようなデータ利活用のために企業内の体制やビジネスモデルを変革していくことは、DXの実現にも繋がるといえる。

その一方で、データの情報源は企業内の各部署である。政府の取り組みを見ても、データはファイル単位でなく、その中身である項目単位で整備を進める必要があることから、そのデータについて一番知見のある部署(情報源)の理解および協力は必要不可欠だ。また、先述の全社的な取り組みの実現には時間がかかることを考慮すると、今後も増えていくことが予想される業務効率化等のボトムアップのアプローチによるデータ利活用の取り組みにおいて、将来的なデ

ータ連携を見据えたデータ整備を意識しておくことは重要である。

このデータ連携を見据えたデータ整備において最も重要となるのは、相互運用性（異なるシステムや組織がデータや情報を円滑に交換し利用できる能力）の確保である。特に大企業等、企業内で複数の異なるシステムを利用している場合に限らず、今後ビジネスにおけるデータ利活用が進むことで、社内外のデータを提供または取得する機会が増えると考えられる。ここで、データの整備手法が一般的な方法と乖離している場合、データ変換に時間を要し、ビジネスにおいて不利になる可能性もある。この相互運用性を実現するための国際的な標準は、技術の変化等に応じて変化していくことも予想される。GIF 等を参考に適宜情報を収集し対応していくことが求められるだろう。

以上