

データ利活用・データスペース ガイドブック 第1.0版



IPA 独立行政法人
情報処理推進機構

デジタル基盤センター
デジタルエンジニアリング部 データスペースグループ

目次

1. はじめに.....	3
2. 目的と想定読者.....	3
本ドキュメントの目的.....	3
本ドキュメントの構成.....	3
想定読者.....	4
3. データ利活用およびデータスペース利用の全体像.....	5
【第1部】 データ利用者編	8
第1章 経営戦略策定フェーズ	9
1.1. 会社のビジョン・方針策定	9
1.2. データ活用のための組織整備	10
1.2.1. CDO（Chief Data Officer: 最高データ責任者）の擁立	10
1.2.2. CDO の役割と組織整備.....	10
第2章 IT 戦略・企画策定フェーズ	13
2.1. データ利用企画	13
2.1.1. ビジネスへのデータ活用のアプローチ方法	13
2.1.2. 収集するデータの検討	13
第3章 データ提供準備フェーズ	16
第4章 データ検索フェーズ	17
4.1. データ検索	17
4.1.1 データカタログサイトの活用	17
4.2 データの認知、取得判断	18
4.2.1. データの取得判断.....	18
4.2.2. データスペース検討	18
第5章 契約フェーズ	21
5.1. データスペースへの参加	21
5.2. データ提供元との契約、締結	21
5.2.1. 契約時の注意点	22

第 6 章 データ利用アプリケーション開発フェーズ	24
6.1. データ計画と設計	24
6.2. データスペース利用に必要な機能等実装.....	25
6.3. データ利用のためのアプリケーション開発.....	25
6.4. テスト.....	28
6.5. 本番稼働.....	28
第 7 章 運用フェーズ.....	29
7.1. 認証、データ取得リクエスト、認可.....	29
7.1.1. 認証、データ取得リクエスト	29
7.1.2. 認可.....	31
7.2. 開発したデータ利用アプリケーションの活用	33
7.3. 来歴確認.....	33
第 8 章 評価フェーズ.....	36
8.1 事業活動の振り返り.....	36
【第 2 部】 データ提供者編	38

※ 【第 2 部】 データ提供者編は近日公開予定。

1. はじめに

現代のデジタル社会や経済において、データは製品やサービスの質を高め、新しいビジネスモデルを創出するための鍵となっている。デジタル経済の発展のため、単純なデータ利用だけでなく、データ主権の保護、業界間のデータ共有の促進、信頼性の高いエコシステムの構築などのニーズが急速に高まっている。

米国の大手 IT 企業がデータ集積し、支配している現状に対し、欧州は自らのデータを管理し、欧州内でのデータ流通と利用を主導することで、経済的・政治的な独立性を確保するため、データの主権を持ちながら、他者とデータを信頼できる形で共有できる仕組みなどを構築している。

国内ではデジタルトランスフォーメーション(DX)の重要性が認識されつつあるものの、データ活用の進展は他の先進国に比べて後れを取っており、企業間や産業間でのデータ共有が十分に進んでおらず、データ駆動型のイノベーションが活発化していない状況にある。また、IMD 世界競争力ランキングなどでも、データの活用能力が低いと評価されており、例えば、次のような主な課題がある。

- ・相互運用性の課題：自社内、限定的な企業間でのデータ共有は行われているが、国間や業界間でのデータ共有は十分でない。
- ・データ主権の課題：欧州のようにデータ主権を守るための統一された仕組みが十分整っていない。
- ・データ活用人材の課題：企業や組織においてデータを戦略的に活用する人材が不足している。

データスペースは、これらの課題を解決する新しいアプローチであり、イノベーションの創出や競争力の強化につながることを期待されている。

2. 目的と想定読者

本ドキュメントの目的や構成、想定読者を示す。

本ドキュメントの目的

本ドキュメントは一般的なデータ利活用だけでなく、データスペースの活用手順やその内容を明確にし、組織の競争力とデータ価値の最大化を支援することを目的としている。

本ドキュメントの構成

本ドキュメントは想定読者を基にした「第 1 部 データ利用者編」と「第 2 部 データ提

供者編」¹の2部構成である。2編のドキュメントではデータスペースを活用するための一般的なプロセスを設定し、フェーズごとに解説をしている。

想定読者

本書はデータスペースに関する基礎知識を持つ読者を想定している。このため、基礎知識はIPA発行の「データスペース入門」²を参照してほしい。

データを取得し、サービス展開などの事業に活用したい者を「データ利用者」、データを提供したい者を「データ提供者」とし、本書の想定読者としている。また、本ドキュメント内で、さらに以下3部門の担当者に分けて解説をしている。

1) 経営層、CDO (Chief Data Officer : 最高データ責任者)

経営層は、経営や事業を実現するための意思決定を行う。CDOはそれを受け、データ戦略の立案などを行う。

2) 事業部門

データ戦略を基に具体的な事業計画を策定・実施する。

3) IT 部門

データスペースによるデータ連携に必要な機能やサービスの開発を実施する。

本書では、各章の節ごとに上記3部門の関わる度合いにより、以下のように示す。

- ・ 主担当の部門：頭に★マークがついており、かつ青色で表記
- ・ 主担当ではないが、関係がある部門：青色で表記
- ・ 関係が少ない部門：グレーアウトで表記

【凡例】



経営層、CDO



事業部門



IT 部門

¹ データ利活用に対する関心の高まりを受け、関係者の皆様がそのメリットをいち早く享受できる社会の実現を目指し、『第1部 データ利用者編』を先行して公開。『第2部 データ提供者編』についても、近日中に公開する予定。

² 独立行政法人情報処理推進機構、「データスペース入門」、<https://www.ipa.go.jp/digital/data/data-space.html>

3. データ利活用およびデータスペース利用の全体像

データスペースの利用を視野に入れたデータ利活用をするためには、多くのフェーズが存在する。その全体像を以下のプロセス図で示しており、データ利用者とデータ提供者がフェーズごとに必要となるタスクが図示されている。

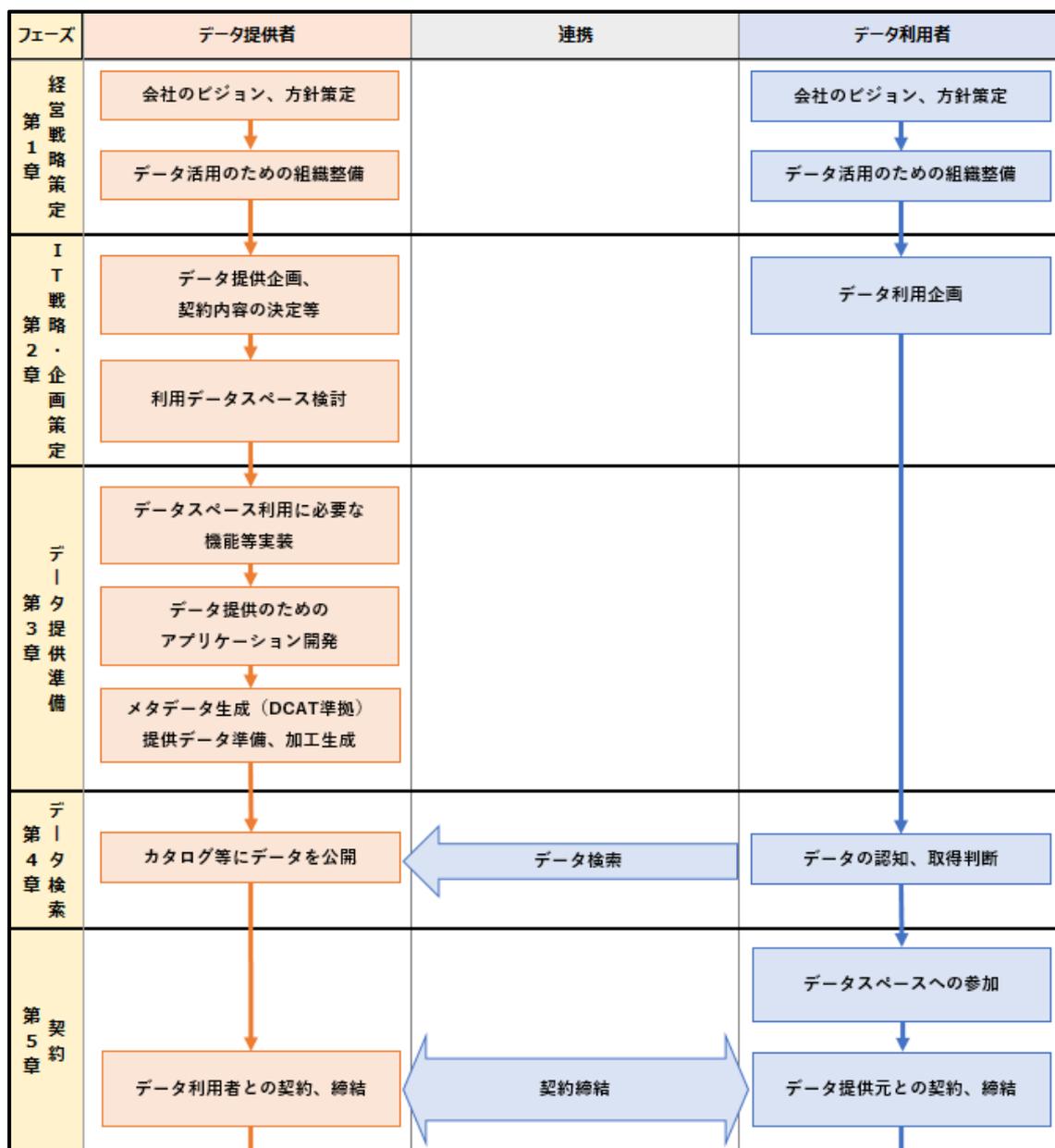


図 0.1 全体プロセス

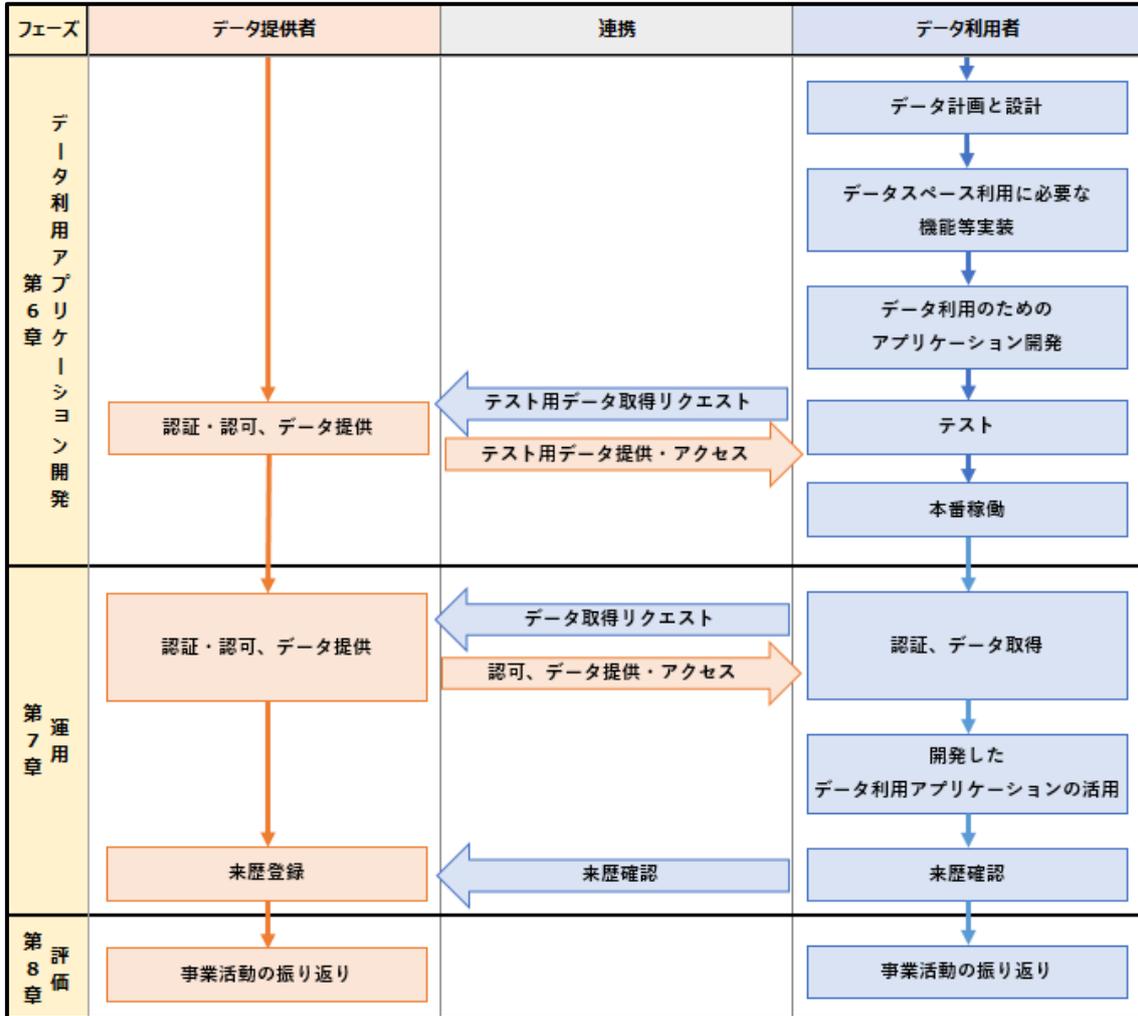


図 0.2 全体プロセス



図 0.3 全体プロセス（簡略版）

本書では、左の各フェーズを章とし、フェーズ内のタスクを解説している。また、図 0.1、0.2 のプロセスの利用者と提供者間に、データ取引市場が仲介するケースもある。データ取

引市場について、詳しくは第5章「契約フェーズ」のポイントメモで解説しているので参考にしてほしい。



【第 1 部】 データ利用者編



第1章 経営戦略策定フェーズ

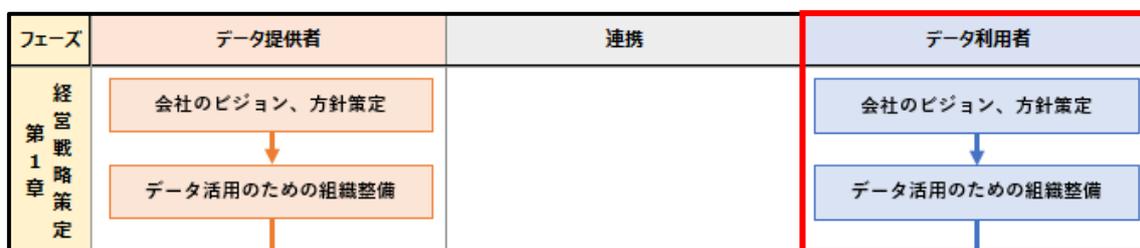


図 1.1 プロセス図「経営戦略策定フェーズ」

本章「経営戦略策定フェーズ」は全体プロセス図の中の、図 1.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、会社のビジョンや方針を策定し、どのようにデータを利活用して事業を進めていくかを検討する。その上で、CDO を設置し、データに関する各種取り組みを行う。



1.1.会社のビジョン・方針策定

データ活用を軸にした事業を推進するためには、まず会社全体のビジョンや方針を策定し、全社にその考えを浸透させる必要がある。このようなデータを活用した経営、組織への成長は、DX の実現というテーマに深く関わるものである。以下に、IPA 発行の「DX 実践手引書 IT システム構築編(以降「DX 実践手引書」)」³の「第1章 DX を実現するための考え方」より、ビジョン・方針の策定のために必要な、基本的な要素を引用する。

ビジョンの発信

全社的な取り組みとして DX を実践する第一歩は、デジタル技術を活用することでビジネスがどう変革されるか、組織の将来像・ビジョンを徹底的に議論し、経営トップら全員が腹落ちしたビジョンを発信することが重要となる。繰り返し組織全体に共有して回することで、組織全体として力が発揮できるようになる。時には現行業務の変化についての反対がある場合は、説得し続けることも必要となる。

ロードマップ策定

³ 独立行政法人情報処理推進機構, 「DX 実践手引書 IT システム構築編」, <https://www.ipa.go.jp/digital/dx/hjuojm000000gx4n-att/000094497.pdf>

策定したビジョンを実現するためには、いつまでに「何を」達成し、今「何をすべき」なのかのロードマップを考える。ロードマップは普遍的なものではなく、世の中の変化やニーズに合わせて随時見直し、最適化、修正を行い、ビジョンに向かっていく。

闇雲にデータ活用を社内で検討する（手段から入ってしまう）前に、組織としてどのような姿になるべきか、どのような世界を実現したいのかを徹底的に考え議論し浸透させることが先決である。



1.2. データ活用のための組織整備

データ活用は個別部門の活用だけでなく、組織全体でデータを効果的に活用してビジネス価値を生み出す方が望ましい。これができている状態をデータマチュリティが高いというが、データマチュリティが高い組織はデータの収集、分析、意思決定などへの反映においてデータに基づいた結果をビジネス戦略に活用することで競争優位を築くことができる。データマチュリティについて詳しくは「データ経営の総合力を高めるデータマチュリティ読本（仮）」⁴を参考にしてほしい。

1.2.1. CDO（Chief Data Officer: 最高データ責任者）の擁立

近年ではデータが企業にとって、戦略的な資産であることが認識されるようになり、そのデータを効果的に管理・活用する必要性が高まっている。データ管理・活用するためには、社内の多くの部門が関わり、協業する必要がある。そこで、データを単なるリソースから、競争力を高めるための戦略的な資産へと変えることを目指し、組織を横断して成功に導く重要な専門のポジションとして、CDO（Chief Data Officer: 最高データ責任者）を擁立することが推奨されている。

CDO について詳しくは「データを軸にした司令塔 CDO 読本」⁵を参考にしてほしい。

1.2.2. CDO の役割と組織整備

CDO は、デジタル時代の鍵となる経営層の一員であることが多く、経営責任者である CEO、業務責任者である COO、財務責任者である CFO などと並び、組織の重要資産であるデータをマネジメントし価値を最大化する責任者である。CDO は組織により、1 名に限

⁴ 独立行政法人情報処理推進機構、「データ経営の総合力を高めるデータマチュリティ読本（仮）」,
<https://www.ipa.go.jp/digital/data/data-spaces-academy.html>

⁵ 独立行政法人情報処理推進機構、「データを軸にした司令塔 CDO 読本」,
<https://www.ipa.go.jp/digital/data/data-spaces-academy.html>

らず「CDO 室」などといったチームを形成して実施する方法もある。

CDO の役割は多岐にわたるが、最も重要なのが、経営層が発信する目指すべきビジョンを基に、経営層と合意を取りながら組織におけるデータの管理と利用を統括する「データ戦略の策定」である。CDO の役割は他にも下記の No.2 以降の項目が該当する。その中でも、社内のデータ活用を円滑に行うために、事業部門等と連携していくこと（DataOps）が重要である。

表 1.1 CDO の担うべき役割の例

No.	役割	役割の内容
1	データ戦略の策定	経営層が発信する目指すべきビジョンを基に、経営層と合意を取りながら組織におけるデータの管理と利用を統括する
2	データガバナンスの確立	組織の中でバラバラにデータを活用するのではなく、組織全体で一貫したルールを設ける
3	データ品質の向上	データの正確性、完全性、一貫性などを確保する (第 4 章 「データ検索フェーズ」ポイントメモ：【データの品質とは？】参照)
4	データ分析の推進	経営層や事業部門だけでもデータ分析ができるようにする
5	データセキュリティの管理	誰にどのデータが扱えるようにするかなどを決める CISO ⁶ と相談し、漏洩や流出、BCP ⁷ 対策を実施する
6	組織全体のデータ文化の醸成	特定の部門のみがデータを扱えるだけでなく、組織全体としてデータ活用をできるようにする
7	データインフラの整備	データの蓄積や連携、分析などを実施するためのインフラを整備する

CDO がこのような役割を実施することでデータの収集や分析が進み、データに基づいた結果をビジネス戦略に活用することができるようになる。これらのことを実践するために、データマネジメントが必要となる。データマネジメントに関する事項について、包括的に記

⁶ Chief Information Security Officer の略称であり、日本語では「最高情報セキュリティ責任者」と訳される。企業の情報セキュリティを統括、管理する責任者を指す。

⁷ 事業継続計画を意味する。自然災害、大火災、テロ攻撃などの緊急事態に備え、事業資産の損害を最小限に抑えつつ、事業の継続や早期復旧を可能にするための計画。(中小企業庁 BCP 策定運用指針 より引用, https://www.chusho.meti.go.jp/bcp/contents/level_c/bcpgl_01_1.html)

載されている資料としては、DMBOK2⁸が広く使われている。以下は、DMBOK2におけるデータマネジメント機能の全体構成である。



図 1.2 「DX 実践手引書」 4.4.2.データ活用を進めるために必要となる要素 より引用

詳しくは『「DX 実践手引書」 4.4.2.データ活用を進めるために必要となる要素』に記載の内容を参照してほしい。

⁸ Data Management Body Of Knowledge の略称であり、日本語では「データマネジメント知識体系ガイド」と訳される。DMBOK は、データマネジメントの普及活動等を行う非営利団体である Data Management Association (DAMA) により作成されており、データマネジメントに関する知見が体系立ててまとめられている。

第2章 IT戦略・企画策定フェーズ



図 2.1 プロセス図「IT戦略・企画策定フェーズ」

本章「IT戦略・企画策定フェーズ」は全体プロセス図の中の、図 2.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、データ利用者として、データ利用を伴う事業の具体的な企画を検討する。



2.1. データ利用企画

経営層が発信したビジョンおよび CDO が立てたデータ戦略を基に事業部門では具体的な事業プランを立案する。CDO は事業部門が事業プランを立案する際、必要となるデータを提供し、場合によっては新たに取得するデータなどを検討する。

社内にデータが無い場合は、社外のデータの利用も検討する。データの取得については 4 章「データ検索フェーズ」にて解説する。

2.1.1. ビジネスへのデータ活用のアプローチ方法

ビジネスへのデータ活用方法は、新ビジネス創出や業務改善の課題に対し、必要なデータを集めるといったアプローチがある。データ収集に際しては、経営層(CDO)、事業部門、IT 部門など多くのステークホルダーが協力する必要がある。

2.1.2. 収集するデータの検討

データ利用企画を計画する際に収集するべきデータを定めるためには、以下のステップを踏むと効果的と考えられる。

① 目標の明確化

まず、事業の目標を明確にする。例えば、新しい製品の市場投入を目指す場合、その製品のターゲット市場や競合状況を理解する必要がある。

② 必要な情報の特定

目標達成に必要な情報をリストアップする。例えば、以下のようなデータが考えられる。

- ・市場データ：市場規模、成長率、トレンド
- ・顧客データ：ターゲット顧客の年齢、性別、購買行動
- ・競合データ：主要競合のシェア、強み・弱み、価格戦略

③ データを探す方法

データを探す方法はいくつかあるが、データカタログのサイトなどを活用し、データを探す方法がある。データは数多くのデータ提供者がデータの情報を公開しており、カタログ以外にも企業の Web サイトで提供しているケースやデータ取引市場内でデータを公開しているケースなどがある。

④ データの種類例

▼オープンデータの活用

政府機関や公的機関などは社会全体の利益のために、広く利用されることを目的としたオープンデータを提供している。オープンデータは基本的には誰でも活用でき、利用制限が少ないとされている。オープンデータを活用するメリットなどについて紹介する。

▽メリット

- ・透明性と信頼性の向上：公共機関や政府が公開しているデータを利用することで、透明性が向上し、市民や企業の信頼を得ることができる。
- ・イノベーションの促進：どの組織にとっても基準になる人口データ、交通量データ、GDP データを活用することで、自組織内外のデータを組み合わせた新たなソリューション、事業等を展開できる可能性がある。

▽オープンデータ関連情報

オープンデータの種類の例：

- ・交通データ(運行状況、交通量、事故情報など)
- ・環境データ(気象情報、エネルギー消費など)
- ・経済データ(GDP、物価指数など)
- ・社会データ(人口統計、健康統計、教育統計、犯罪統計など)

オープンデータ入手サイトの例：

オープンデータの入手先の例を IPA のサイトにリンク集として掲載。

- ・データに関わる国内外の取り組みに関するリンク集⁹

⁹ 独立行政法人情報処理推進機構，データに関わる国内外の取り組みに関するリンク集，
<https://www.ipa.go.jp/digital/data/link.html>

▼非構造データの活用

構造化されたデータ以外にも構造化されていない非構造データ(テキスト、画像、ビデオ、ソーシャルメディアの投稿など)を活用することもできる。

ただし、非構造データは種類や量が多いが、一般的にそのままでは活用が難しいとされる。構造定義を持たず、そのままでは処理が難しいため、分析活用可能な状態にするために「変換や加工」が必要となるケースがある。

なお、実際にデータを取得するのは、「第4章 データ検索フェーズ」となるが、どのようなデータが取得可能なのか等、実情を調査しない限りには、データ利用企画を立てても、そのフィージビリティを保証できない。そのため、このフェーズで、取得可能なデータを実際に調査することになる場合がほとんどである。その際に便利なカタログサイトの活用方法等は、次の第4章に記載の内容を参考にしてほしい。

第3章 データ提供準備フェーズ



図 3.1 プロセス図「データ提供準備フェーズ」

本章「データ提供準備フェーズ」は、データ提供者側が、提供するデータを準備するフェーズである。

このフェーズでは、利用者側に必要なタスクが特に存在しないため、次のフェーズに進む。提供者側のタスクについては「第2部 データ提供者編」にて解説する。

第4章 データ検索フェーズ



図 4.1 プロセス図「データ検索フェーズ」

本章「データ検索フェーズ」は全体プロセス図の中の、図 4.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、データ提供者がカタログ等に公開したデータの情報を、データ利用者が認知し、実際に利用するかどうかの取得判断をする。



4.1. データ検索

第 2 章で企画したデータ利用企画に沿って必要なデータを取得していく。自組織内で不足しているデータは Web 検索などを行い取得する必要がある。

4.1.1 データカタログサイトの活用

データを検索する方法は多くあるが、その中にデータカタログサイトを活用しデータを検索する方法がある。

▼メリット

- ① **利便性向上**：膨大なデータの中から、キーワードなどで簡単に検索・閲覧でき、また比較することも容易にできる。
- ② **取得するデータの品質確保**：データセットに関する詳細な情報（メタデータ）が提供されている場合が多く、データの出所、作成者、更新履歴など、品質確保に必要なデータに関する情報を確認できる。データの品質を保証することは、正確なデータ活用・分析に不可欠であり、適切な意思決定を実現するためにも押さえておくべきポイントである。データ品質については、本章のポイントメモにて解説しているので、参考にしてほしい。



4.2 データの認知、取得判断

4.2.1. データの取得判断

データ検索等を利用しデータを探した後は、そのデータを実際に取得利用するか否かを検討し判断する。実際のデータそのもの(データ値)の内容だけでなく、メタデータ等、そのデータに関する情報からも、利用の可否を判断する。その他に、実際に取得した場合にどのような加工(データ同士の結合、不要項目のマスキング等)が必要になるのかを検討し、6章の「データ利用アプリケーション開発フェーズ」の際に開発するアプリケーションの要件定義の材料とする。

しかし、場合によっては計画している事業に必要なデータの情報が入手できない場合があり、その場合は次のような再検討する必要がある。例えば、検索においては、さらに詳細な検索や別キーワードを試みるといった方法もあるが、別のサイトの調査や、場合によってはデータ利用企画を見直す必要もある。

4.2.2. データスペース検討

利用したいデータにより、参加するデータスペースを検討し、利用可否の判断をする。



【データの品質とは？】

データの取得判断をするときには、そのデータの品質が良いかどうか、という観点も重要なポイントになります。品質の良いデータとはどのようなものでしょうか？ 品質の良いデータが持つ特性を紹介します。

表 4.1 データの品質評価軸

No.	品質評価軸	説明
1	正確性 (Accuracy)	意図した概念や事象を、データがどの程度正しく表現できているか。例えば、姓名が反対に記載されていないか、等
2	完全性 (Completeness)	対象が複数のものの集合であるときに、どの程度その集合の全体のデータをカバーできているか。例えば、必要な項目が網羅されているか、等
3	一貫性 (Consistency)	複数のデータ間などにおける相互矛盾はないか、整合性が取れているか。例えば、生年月日と年齢が矛盾していないか、等
4	信憑性 (Credibility)	データの内容の正しさが、どの程度信頼できるか。例えば、データの収集方法や検証方法など信頼性に関わる情報が提供されているか、等
5	最新性 (Correctness)	システムの持つ/表示するデータが、どの程度現実の変化に追従できているか。例えば、株価のように刻々と変化していくデータが最新のものになっていっているか、等
6	アクセシビリティ (Accessibility)	障がいのある人々が、どの程度、そのデータにアクセスできるか。例えば、一つの情報が複数のチャンネル(文字、画像、音声など)で伝えられているか、等
7	標準適合性 (Compliance)	データに関する法令や各種の基準にどの程度適合しているか。例えば、個人データや金融商品などに関わるデータは法令の規制を受けるがそれに適合しているか、等
8	機密性 (Confidentiality)	データ利用が、どの程度、正当な利用者だけにのみ限られているか。例えばデータが暗号化されているか、等
9	効率性 (Efficiency)	データの取扱いにおいて、システム上のリソース（メモリ等）をどの程度効率的に利用しているか。例えば、不要な冗長性のため処理時間や消費するリソースの量が過大となっていないか、等

10	精度 (Precision)	データはどの程度精密か。例えば、データが誤差を伴うものとした場合に、その誤差が想定範囲内か、等
11	追跡可能性 (Traceability)	データへのアクセス事象やデータの変更事象をどの程度まで、また容易に追跡できるか。例えば、データへのアクセス履歴（アクセス者／アクセス日等）や、データの変更履歴（変更者／変更日等）が記録されているか、等
12	理解性 (Understandability)	データの内容がどの程度容易に判読・理解できるか。例えば、データが適切な言語、シンボル、単位で表現されているか、等
13	可用性 (Availability)	データを利用したいときにどの程度の割合で利用できるか。例えば、夜間や年末年始等の特定の日に、どの程度データを利用できるか、等
14	移植性 (Portability)	データがどの程度プラットフォームや組織を越えて実装したり、移動したりできるか。例えば、異なるシステム間でスムーズにデータを交換できる手段はあるか、等
15	回復性 (Recoverability)	システムにトラブルが生じたとき、どの程度本来のデータを回復できるか。例えば、バックアップやRAIDのような装置を利用し、トラブル発生前の状態に戻れるか、等

ISO「Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model」によるデータ品質定義より引用¹⁰

表 4.1 データの品質評価軸は、ISO によるデータ品質に関する定義を引用し整理したものです。これらの観点からデータの品質を確かめ、条件を満たすデータは、品質の良いデータといえるでしょう。仮にデータの品質が悪いと、正確なデータ活用・分析ができず、誤った意思決定を発生させる原因になりえるため、このポイントを押さえることは重要です。

ただし、データの品質の精度をどこまで高めるべきか目的によって異なります。精度が高いデータは、その分取得費用が高くなる場合が多いです。例えば、顧客アンケートの大まかな傾向把握をする等、ある程度精度の低いデータでも問題がない場合は、取得費用を抑えることができます。データ品質を考えるとときには、このような品質と費用のバランスを考えることも重要です。

¹⁰ ISO/IEC 25012:2008, Software engineering — Software product Quality Requirements and Evaluation (SQuaRE) — Data quality model, <https://iso25000.com/index.php/en/iso-25000-standards/iso-25012>

第5章 契約フェーズ



図 5.1 プロセス図「契約フェーズ」

本章「契約フェーズ」は全体プロセス図の中の、図 5.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、自身が参加するデータスペースなどを決定し、データ提供者間とデータ利用のための契約を締結する。



5.1. データスペースへの参加

第2章でのデータ利用企画の検討内容などを基に、自組織にとって必要なサービス、ツール、機能などが提供されているデータスペースなどの活用を検討する。また、参加するデータスペースなどで契約内容や利用規約などを用意しているケースがあるため、事前に確認した上で利用を検討する。

データスペースは、各業界・業種・地域などで既に活用されているケースもあるため、統括している組織¹¹などに事前に問い合わせた上で、参加・利用を検討する方が良い。



5.2. データ提供元との契約、締結

データ提供元のデータを利用する場合、手続きや契約が必要な場合がある。その目的は、データ提供者にとって、意図しない利用者へのデータ利用を防ぐなど、データ主権の考えに基づいている。

契約した内容を基にデータの提供を行う際、データ取得リクエストがあるたびに人がデ

¹¹ 統括している組織の例：日本経済団体連合会（経団連）、日本銀行協会（JBA）、日本機械工業連合会（JMF）、日本小売業協（JRA）、一般社団法人自動車・蓄電池トレーサビリティ推進センター（ABtC）等

ータ提供の可否を判断するのではなく、機械的に契約条件に照らし合わせ、提供可否を判断する方法もある。利用するデータスペースなどによっては、そのような機能が提供されているケースがある。データ提供者とデータ取得までのやり取りについて事前に確認しておく方が良い。

5.2.1. 契約時の注意点

参加するデータスペースやデータを提供する組織と契約を結ぶ場合の注意として、以下のような点などが挙げられる。

- ・ライセンス条件の遵守、使用目的と範囲や免責事項の確認などが必要である。
- ・オープンデータは契約締結の必要がない場合が多いが、利用条件が規定されているケースがあるため、事前に確認が必要である。
- ・セキュリティやコストなどの観点から、利用期間、利用範囲などなるべく必要最低限の契約を心がける。
- ・データ取引市場を介してデータ取引を行う場合は、取引市場の運営事業者が契約書を用意している場合があるので、事前に確認する。



ポイントメモ

【データ取引市場とは】

データ契約を結ぶときの注意点の中で「データ取引市場」という言葉を用いましたが、これが何なのか簡単に解説します。

▼データ取引市場とは？

データ取引市場は、多数の売り手と買い手が集まり、データの売買を行う場です。証券取引所のように、データの取引を効率的かつ安全に行うことができます。

▼データ取引市場を利用するメリット

- ・取引の効率化: データを取得するための調査などにかかる時間やコストを削減できます。
- ・適正価格での取引: 複数の利用者の取引事例が蓄積されることで、市場価格が形成され、取引の安定化が進みます。

このようにデータ取引市場を利用するメリットは多いため、データを活用した事業を検討する際には、データ取引市場を介したデータも検討してみてください。

第6章 データ利用アプリケーション開発フェーズ

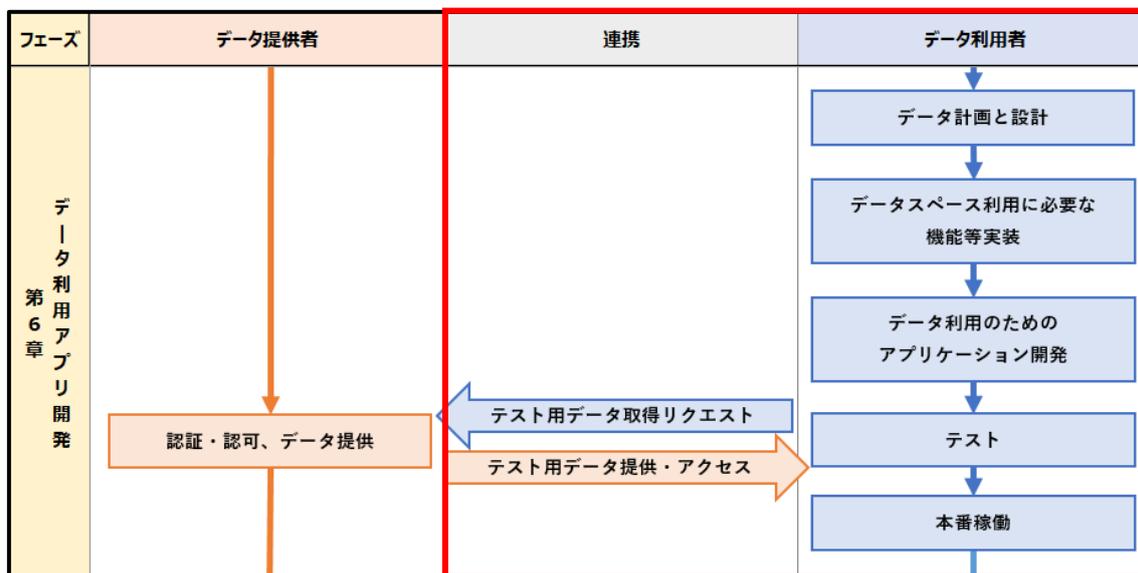


図 6.1 プロセス図「データ利用アプリケーション開発フェーズ」

本章「データ利用アプリケーション開発フェーズ」は全体プロセス図の中の、図 6.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、データ利用者が取得したデータを実際に事業に活用するためのアプリケーションを開発する。テストの最終段階では、本番稼働のときと同様のデータを利用したテストを実施することを推奨する。



6.1. データ計画と設計

データスペースでは、異なる組織がデータを活用するため、基本的には提供されるデータについては標準化されているケースが多い。その提供されるデータを円滑に自組織内で利用できるように自組織のデータ構造やデータの管理方法などをあらかじめ決めておく必要がある。

- ① データモデリング：データの構造や関連性を明確にするなど。
- ② データデザイン：データの利用方法、管理方法、運用方法などを設計する。
- ③ データの標準化：GIF、IMI などを活用しマスターデータなどを標準化しておくことで異なるシステムや組織間でデータを効率的かつ正確に交換できるようになる。GIF（政

府相互運用性フレームワーク)¹²や IMI (共通語彙基盤)¹³を参照。



6.2. データスペース利用に必要な機能等実装

参加するデータスペースなどで、データ利用者のために機能やサービスを提供している場合がある。

代表的なものが「コネクタ」である。コネクタはデータスペースにおける最も基本的なモジュールであり、システムに搭載することで安全なデータ転送/アクセスが可能になる。

また、検証・評価のためにテストベットが提供されているケースがある。開発の効率を高め、導入リスクを低減するなどのメリットがある。

提供されているサービスや機能群は、参加するデータスペースで公開されている仕様書等ドキュメントがあれば、それを参照する。公開されていない場合は、運営主体の組織などに確認しても良い。これらの提供されているサービスや機能群を活用し、なるべく開発しないことでアプリケーションを活用するまでのスピードやアプリケーションの品質が向上する。

しかし、サービスや機能が提供されていない場合は、必要最低限の開発をする必要がある。つまり、そのような場合でも、世の中で提供されているサービスや機能を利用することで、自社開発は必要最低限になるよう努めることが望ましい。



6.3. データ利用のためのアプリケーション開発

最低限の開発が必要と判断した場合は、データ活用の開発として「データを収集する」、「収集したデータを整理・加工する」、「整理・加工されたデータを蓄積する」、「蓄積したデータを抽出する」、「抽出したデータを分析やサービス、アプリケーションで利用する」といったフェーズで検討する方が良い。詳しくは『「DX 実践手引書」4.4.1.データ活用のための環境整備のプロセス』に記載の内容を参照してほしい。

¹² デジタル庁, 政府相互運用性フレームワーク(GIF), https://www.digital.go.jp/policies/data_strategy_government_interoperability_framework

¹³ 独立行政法人情報処理推進機構, IMI, <https://imi.go.jp/>

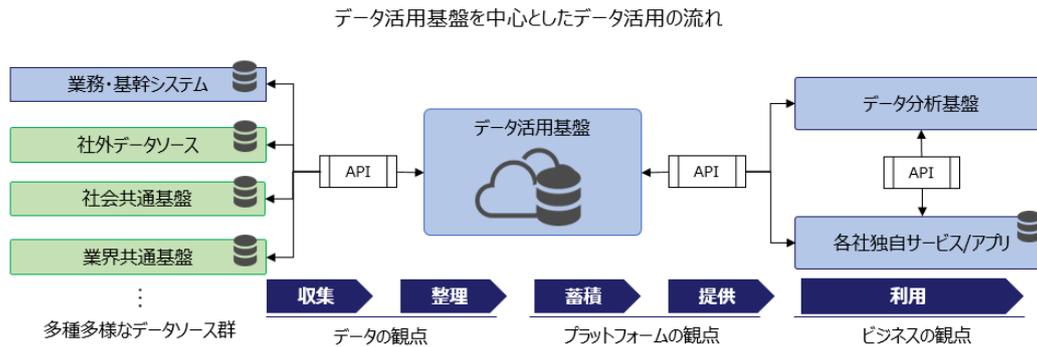


図 6.2 データ活用基盤を中心としたデータ活用の流れ（「DX 実践手引書」より引用）

次に各フェーズで必要となる開発の例を示す。

① データ収集フェーズ

- ・外部データからデータを取得するための API 開発
- ・全てのデータを取得するのではなく、Graph 技術¹⁴などを活用し、必要最小限のデータのみを取得

② データ整理・加工フェーズ

- ・データクレンジング：データの品質を向上させるため、データの誤りや欠損を修正
- ・データ変換：データを分析やシステム間で利用可能な形式に変換
- ・データ統合：複数のデータソースを一つに統合し、統一されたデータセットを作成
- ・データ匿名化：個人情報保護のため、データの個人を特定できる情報をマスキング

③ データ蓄積

- ・整理、加工したデータを指定の場所に格納

④ データ提供(データ活用に連携)

- ・データ蓄積したデータを取り出し、データ分析基盤やデータ利用アプリケーションに連携

⑤ データ活用(分析、データ利用アプリケーション)

- ・データ集計：特定の基準（例：カテゴリ別）でデータをグループ化し、各グループの統計量を計算するなど
- ・データフィルタリング：必要なデータのみを抽出
- ・データサンプリング：データ全体から一部を抽出
- ・データ正規化：データのばらつきを抑え、一貫性のあるデータを提供
- ・データ拡張：モデルの学習データを増やすために、既存のデータを加工して新しいデータを作成

¹⁴ 必要な情報を直接取得でき、そこからさらに深掘りできるデータ操作の仕組み。拡張性が高く高速で、転送負荷も軽い特徴がある。

- ・データ匿名化：②で加工せず、個人情報を含めたデータを蓄積する場合などは、このフェーズで匿名化の実施を検討



6.4. テスト

データ利用者は、開発したアプリケーションを本番稼働させる前に、実際のデータを利用してテストを実施することが望ましい。データ提供元によっては、データ利用者がテストで利用する目的であったとしても契約前の実データを渡せないというケースもあるので、テスト期間も契約期間に含めるなどの対応をとると良い。

実データを利用してテスト可能な場合は、利用したいデータを保有しているデータ提供者に、データ取得リクエストを送り、データを取得することになる。

データの取得形式は、データ提供元から実際のデータが入ったファイルを受領する場合や、データが格納されている場所へのアクセス権限が付与されるなど、いくつかのケースが想定される。



6.5. 本番稼働

テストが完了したデータ利用のためのアプリケーションを、実際に運用する環境で稼働させる。この段階をもって、実際に運用するフェーズに入る。

第7章 運用フェーズ

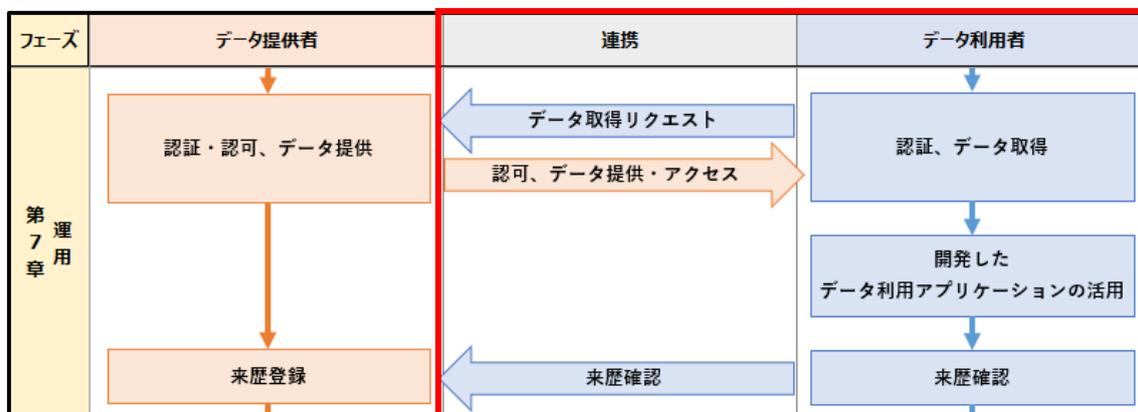


図 7.1 プロセス図「運用フェーズ」

本章「運用フェーズ」は全体プロセス図の中の、図 7.1 の赤枠部分の利用者側のタスクを解説する。

このフェーズでは、提供されているデータを取得する前に本人確認や契約条件と照らし合わせ(認証・認可)、開発したアプリケーションを活用しながら事業活動や分析を行う。必要があればデータの信憑性向上のため、来歴確認を行う。

また、運用時は CDO が定めたデータガバナンス等に従い、適切なデータ運用を実施する。データの不備や使用上の課題があったときは、データを管理する部門へ改善のためのフィードバックを行う。



7.1. 認証、データ取得リクエスト、認可

データを取得する際には、本人であることの確認やデータ利用の契約条件を満たしているかの確認をするため、認証や認可を経てデータを取得する必要がある。

7.1.1. 認証、データ取得リクエスト

契約した利用者の情報をあらかじめ登録しておき、データの連携前に認証することにより、利用者が契約した利用者であることを証明する。

認証のパターンはいくつかあるが、デジタル基盤¹⁵などから提供される認証機能を利用する例を紹介する。

¹⁵ 独立行政法人情報処理推進機構, 「データスペース入門」 デジタル基盤に関する記述を参照,
<https://www.ipa.go.jp/digital/data/data-space.html>

デジタル基盤から提供される認証機能を利用する例

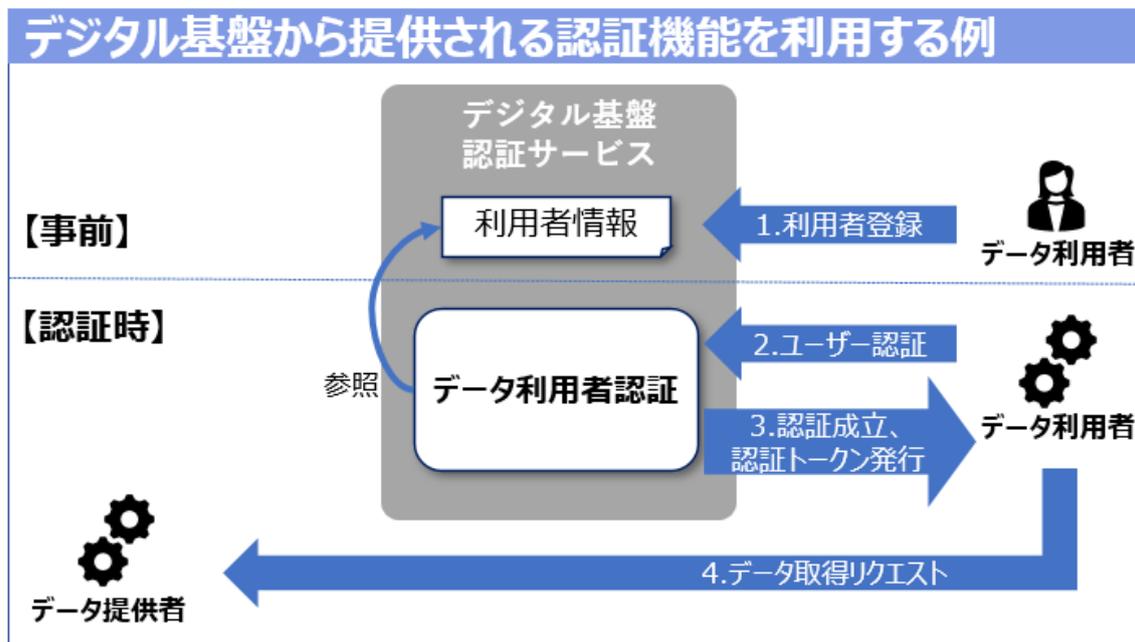


図 7.2 デジタル基盤による認証プロセス¹⁶

1. 利用者登録
データ利用者は、事前にデジタル基盤の利用者登録を実施し、デジタル基盤のユーザーIDと認証情報を入手する。
2. ユーザー認証
データ利用者から認証サービスに認証依頼をすると、認証サービスからユーザーIDや、その認証情報などの提供依頼があり、利用者はそれらの情報を提示すると、認証処理が実施される。
3. 認証成立、認証トークン発行
認証が成立すると、ユーザーIDに対応する認証トークンが発行される。
4. データ取得リクエスト
データ利用者からデータ提供者に、認証トークンを埋め込んだデータ取得のリクエストをする。データ提供者はその認証トークンを基に、利用者が正しいかどうかを判断する。

¹⁶ 独立行政法人情報処理推進機構、「データスペース入門」を参考にIPAで作成、
<https://www.ipa.go.jp/digital/data/jod03a000000a82y-att/dataspaces-gb.pdf>

7.1.2. 認可

データ利用者がデータ取得リクエストを送り、データ提供者に認可してもらうことで、利用者はデータへアクセスすることが可能となる。デジタル基盤などによっては、認可機能を提供しており、その場合、契約締結時にデジタル基盤などに登録された契約条件に則ったデータへのアクセスが利用者に許可される。

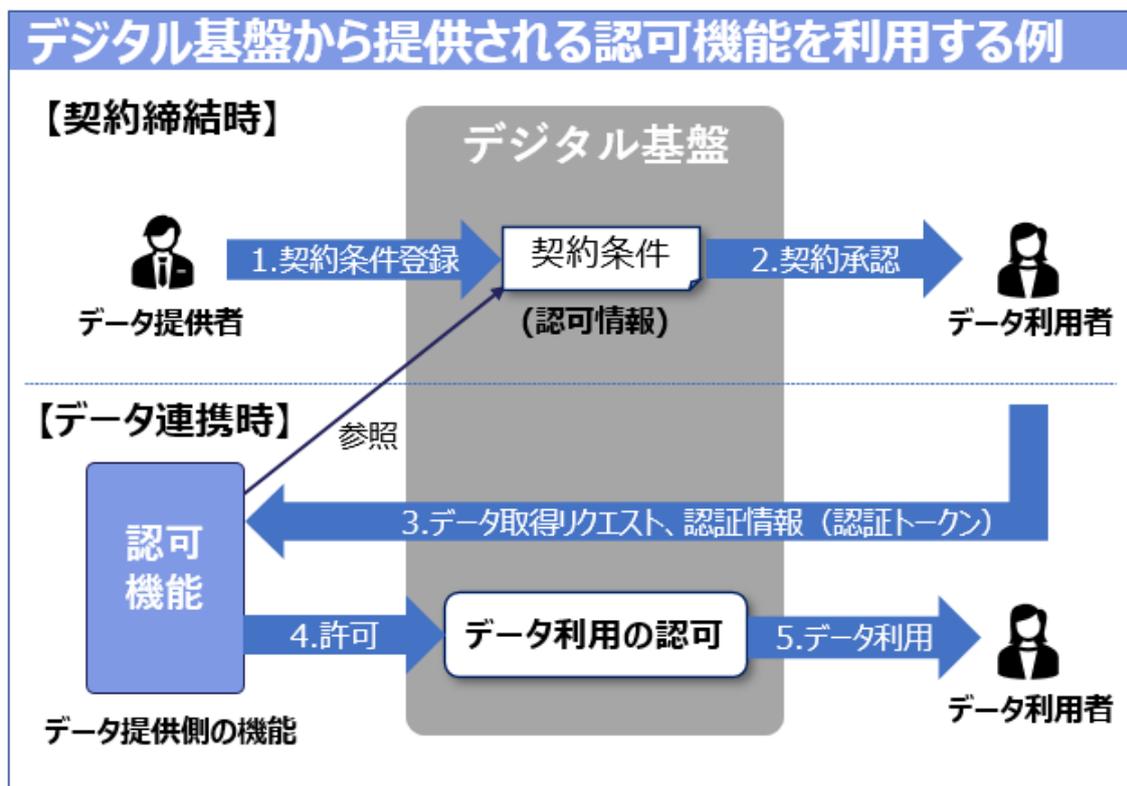


図 7.3 デジタル基盤による認可プロセス¹⁷

1. 契約条件登録
データ提供者は、契約するデータ利用者に対して提供するデータの条件を登録する。
2. 契約承認
データ利用者は1の契約条件を確認し、承認する。その契約条件は利用者がデータ取得する際の認可情報となる。
3. データ取得リクエスト、認証情報 (認証トークン)
データ利用者からのデータ取得リクエスト、および認証情報 (認証トークン) を提供者が受け取る。
4. 許可

¹⁷ 独立行政法人情報処理推進機構, 「データスペース入門」を参考に IPA で作成,
<https://www.ipa.go.jp/digital/data/jod03a000000a82y-att/dataspaces-gb.pdf>

利用者が契約条件を満たしていると判断された場合、データへのアクセスを許可する。

5. データ利用

データ利用者はデータを利用できる。



7.2. 開発したデータ利用アプリケーションの活用

データ利用者は、7.1.で取得したデータを基に、『6章「データ利用アプリケーション開発フェーズ」6.3. データ利用のためのアプリケーション開発の②データ整理・加工フェーズ』以降で開発したアプリケーションを活用し事業活動や分析などを行う。



7.3. 来歴確認

運用時、データの信憑性を確認するため、来歴確認を行う。来歴確認を行う機能はデジタル基盤などから提供される場合がある。この機能を用いることで、データ利用者也提供者も、誰がいつデータを登録、加工、取得したかの記録を確認し、データが不当に改ざんされていないことなどを確認することができる。

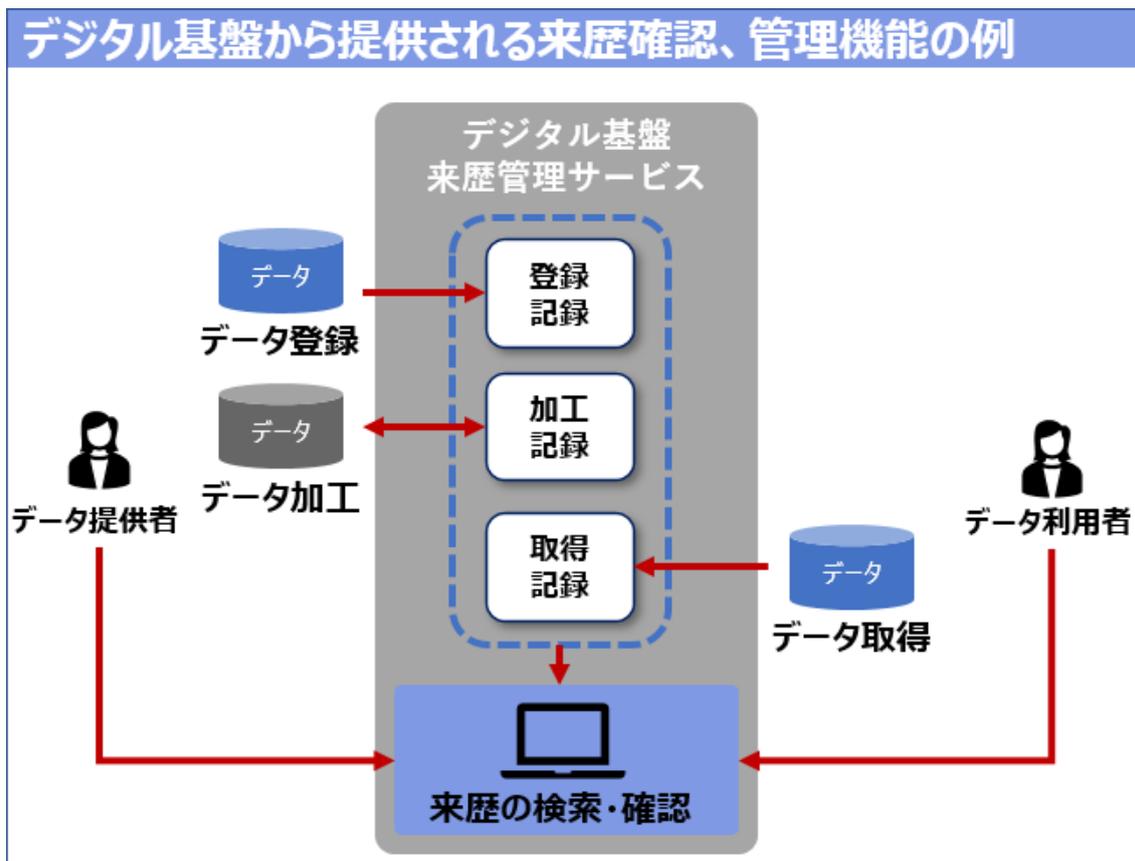


図 7.4 デジタル基盤から提供される、来歴管理、確認の全体像

来歴管理が有する機能は、以下のようなものがある場合がある。データ利用者、提供者の

双方のデータに関する記録を管理し、双方がその記録を確認できる。

1. 来歴の記録

- ・データ登録：原本情報を記録
- ・データ加工：データ加工履歴を記録
- ・データ取得：データの送受信履歴を記録

2. 来歴の検索・確認

- ・確認したい対象データについて検索し、一連の履歴を確認



ポイントメモ

【認証・認可の想定ケースについて】

本章では認証について、データ利用者とデータ提供者という汎用的な表現を用いて解説しましたが、ここでは理解を助けるために、具体的な想定ケースを例に補足説明します。以下のようなケースで見てください。

表 7.1 想定ケース

事業	目的	取得したいデータ	データ取得先
不動産売買仲介業	正確な不動産売買価格の査定	<ul style="list-style-type: none"> ・商業施設利用データ（性別、年齢、商品価格等） ・人流データ（人の位置、性別、年齢層等） ・物件データ（価格、立地等）等… 	<ul style="list-style-type: none"> ・商業施設事業者（小売業、飲食業、鉄道業等） ・他の不動産業者（売買仲介業、賃貸仲介業、管理業）

実際にデータ取得までのプロセスに当てはめてみると、以下の下図 7.6 のようになります。

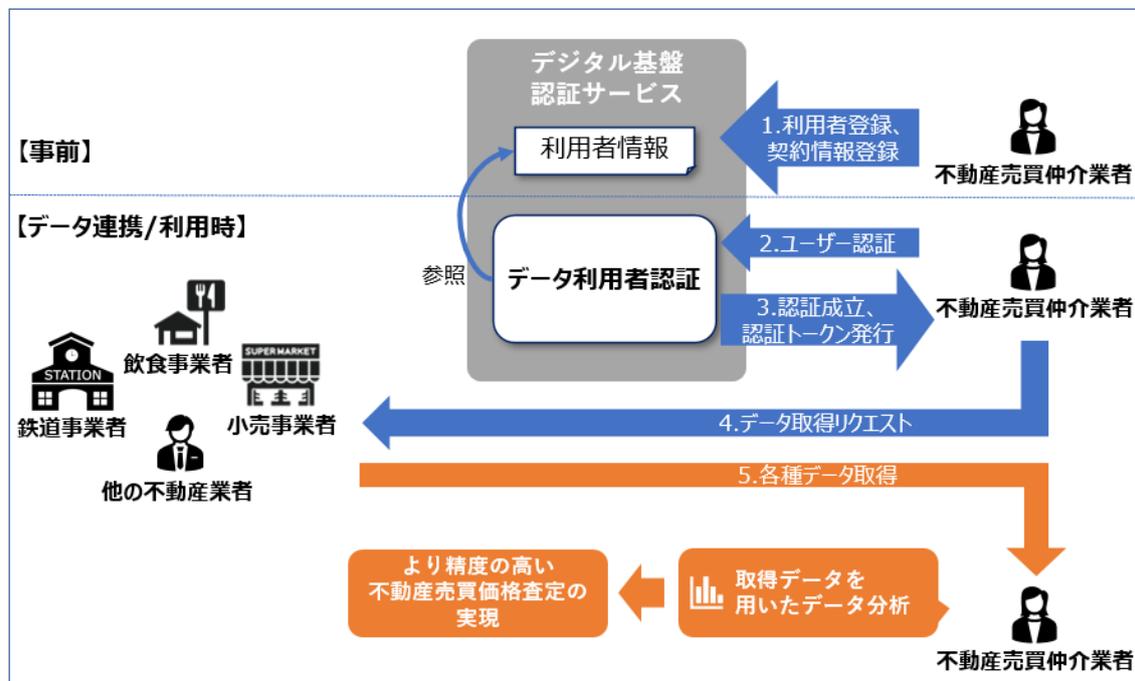


図 7.6 デジタル基盤による認証プロセスの具体例

このような想定ケースでは、不動産売買仲介業者は各事業者からデータを取得し、データ分析等により、より正確な不動産売買価格の査定に役立てることができます。

第8章 評価フェーズ



図 8.1 プロセス図「評価フェーズ」

本章「評価フェーズ」は、第7章「運用フェーズ」で実際にデータ戦略に基づいた事業活動やデータ分析をした結果を振り返り、そこからの改善を実践するフェーズである。



8.1 事業活動の振り返り

データに基づいた事業活動やデータ分析を改善するため、CDOの観点などから事業活動を評価し、データスペースを活用している場合はその観点を含め評価するのが望ましい。評価した結果、データの追加取得や分析観点の見直しなど様々な取り組みが必要となるケースがある。

評価について、データスペースの特徴を基にした観点やCDOの役割(第1章「経営戦略策定フェーズ」参照)を基にした観点でいくつかの例を挙げる。

▼データスペースの評価観点の例

① 相互運用性(インターオペラビリティ)

データスペースを通じて、どれだけ効果的にデータが異なる組織間で共有され、協力が進んだかなどを評価する。

② データ主権

データが国や地域のデータ主権に関する法規制を遵守しているかなどを評価する。また、組織や国家が保持しているデータのアクセス権などを確認し、第三者がデータにアクセスしていないかなどを評価する。

▼CDOの役割としての評価観点の例

① データ戦略の達成度

立案したデータ戦略が企業のビジネス目標(ビジョン)にどの程度貢献したかを確認する。KPI(主要業績評価指標)として設定したコスト削減効果、収益、顧客満足度などを測定し、成功点や不足点などを評価する。また、データ活用によるイノベーションの促進や、新しい製品・サービスの創出にどれだけ寄与したかなどを評価する。

② データガバナンスの効果

データガバナンス体制やデータ管理プロセスが十分に機能しているかなどを評価する。法令遵守の観点ではデータプライバシーやセキュリティに関する法規制に従っているかなどを評価する。

③ **データ品質の状況**

データの正確性、完全性、一貫性などが確保されているかなどを評価する。

④ **データ分析の推進状況**

データ分析に必要なデータに不足がないかなどを評価する。必要なデータが不足している場合、追加データを検討する。また、新たな観点で分析が必要となる場合、その分析に必要なデータを改めて検討する。

⑤ **セキュリティとリスク管理の状況**

セキュリティポリシーやリスク管理などが適切に実施され、データ漏洩のリスクやサイバー攻撃を最小化できたかなどを評価する。

⑥ **組織全体のデータ文化の浸透度**

組織全体でデータを活用する文化がどれだけ浸透しているかなどを評価する。社員のデータリテラシーが向上し、BI ツール¹⁸などを活用することでデータに基づく意思決定が日常業務にどれだけ取り入れられているかなどを評価する。

⑦ **データインフラの有効性**

データ活用基盤、データ分析基盤の整備などが計画通り進んでいるか、またインフラのパフォーマンスが適切かなどを評価する。クラウド移行やデータレイク構築が実際にデータの処理効率やアクセス性にどれだけ貢献しているかなどを確認する。

このように、データスペースの観点や CDO の観点などを軸にデータ戦略を基にした計画と実績の差を把握し、適切に対応することや、世の中の変化や新たなニーズなどに合わせ柔軟に対応していくことなどを持続的に行うことで、経営層が掲げたビジョンに近づき、組織の成長と成功が実現する。

本ドキュメントでは、データスペースおよびデータ利活用の全般について、企画の上流から下流の運用までを包括的に解説した。経営におけるデータの効果的な活用は、現代のビジネスにおいて競争力を維持するために不可欠である。本ドキュメントが、企業にとってデータの価値をさらに引き出す一助になれば幸いである。

¹⁸ BI（ビジネスインテリジェンス）ツールは、企業がデータを収集、分析、可視化などをするためのツール。例えば、売上データをグラフで表示したり、顧客の行動を分析したりすることで、経営者や従業員がデータに基づいた意思決定を行いやすくする。



【第2部】データ提供者編



(近日公開予定)

データ利活用・データスペースガイドブック

【著作・制作】 独立行政法人情報処理推進機構（IPA）デジタル基盤センター

【編集責任】 デジタル基盤センター 平本 健二

【監修】 デジタルエンジニアリング部 小田切 未来、土橋 昌
デジタルアーキテクチャ・デザインセンター 田村 公孝

【執筆】 デジタルエンジニアリング部 西本 靖、土屋 文靖

【改版履歴】 令和 6年 11月 13日 第 1.0 版 発行

【発行所】 独立行政法人情報処理推進機構（IPA）
デジタル基盤センターデジタルエンジニアリング部
データスペースグループ
〒113-6591
東京都文京区本駒込二丁目 28 番 8 号
文京グリーンコートセンターオフィス
URL : <https://www.ipa.go.jp/>

この文書のご利用にあたって

本ガイドブックの内容を適用した結果生じたこと、また、適用できなかった結果については、IPA は一切の責任を負いかねますのでご了承ください。



独立行政法人情報処理推進機構
Information-technology Promotion Agency, Japan