

# データカタログ作成ガイドラインV1.1 (中間とりまとめ)

## 概説

2019年1月25日

一般社団法人データ流通推進協議会

- 1. データカタログとは**
- 2. 本ガイドラインにおける項目定義について**
- 3. 本ガイドラインを用いたデータカタログ項目策定の流れ**

# 1. データカタログとは

## 1.1 データカタログとは

- データカタログは、データの所在や内容等の概要情報を項目別に記入する書式の総称です。
- データカタログを活用すれば、データ本体を入手する前に、データの概要が分かります。
- データの検索性を向上でき、データ提供者・データ提供先でのデータのやり取りを促進できます。

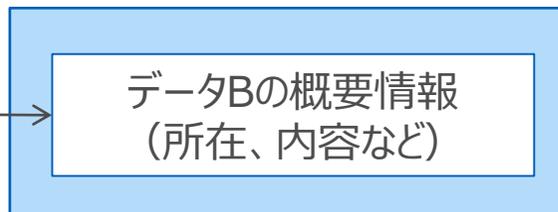
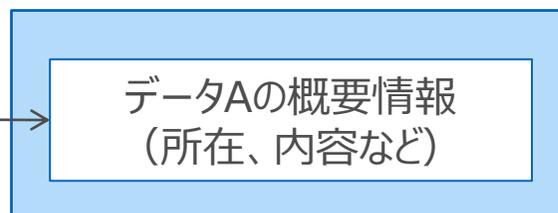
### データ提供者A



### データ提供者B



### データカタログ



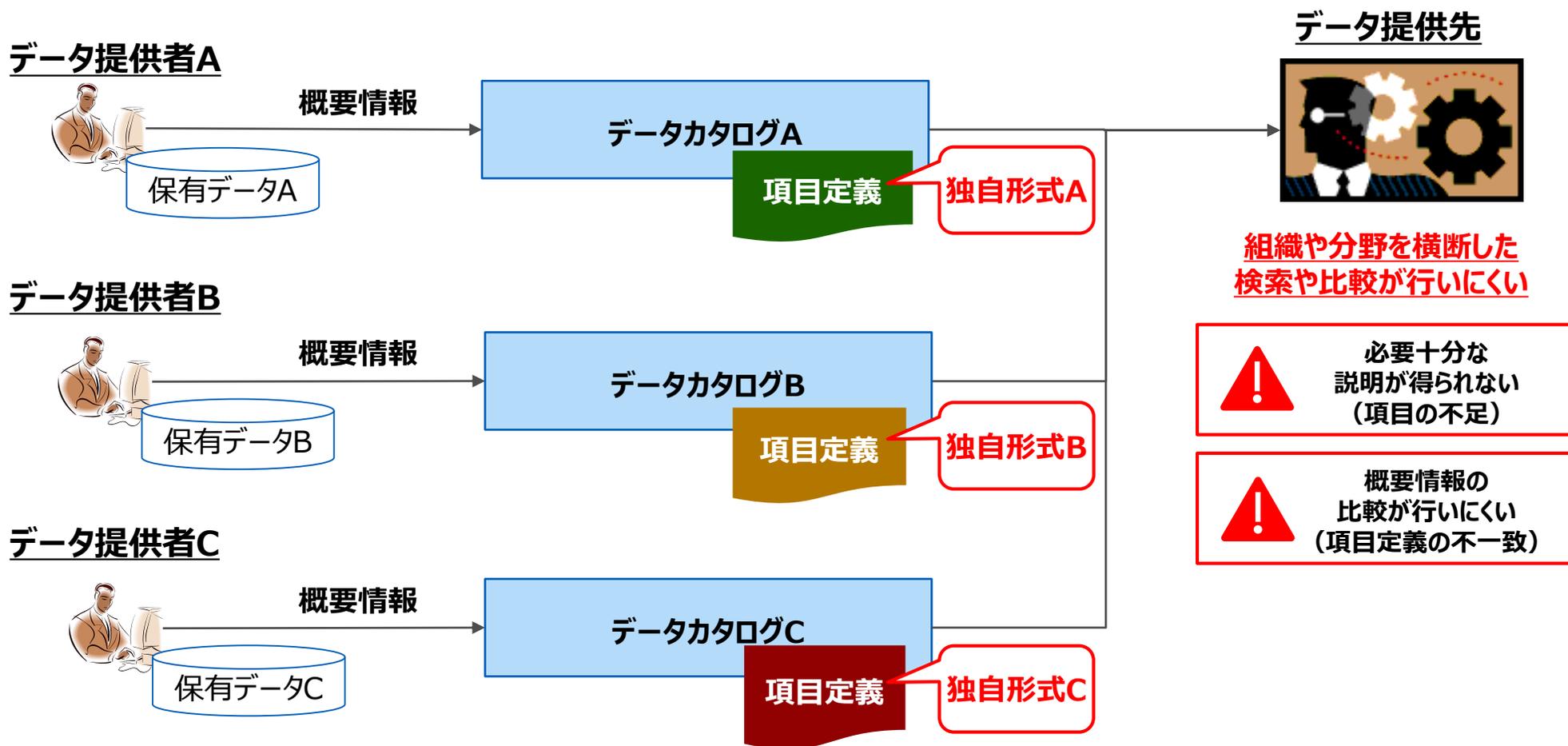
### データ提供先



データ本体を入手する前に、  
所在や内容がわかる

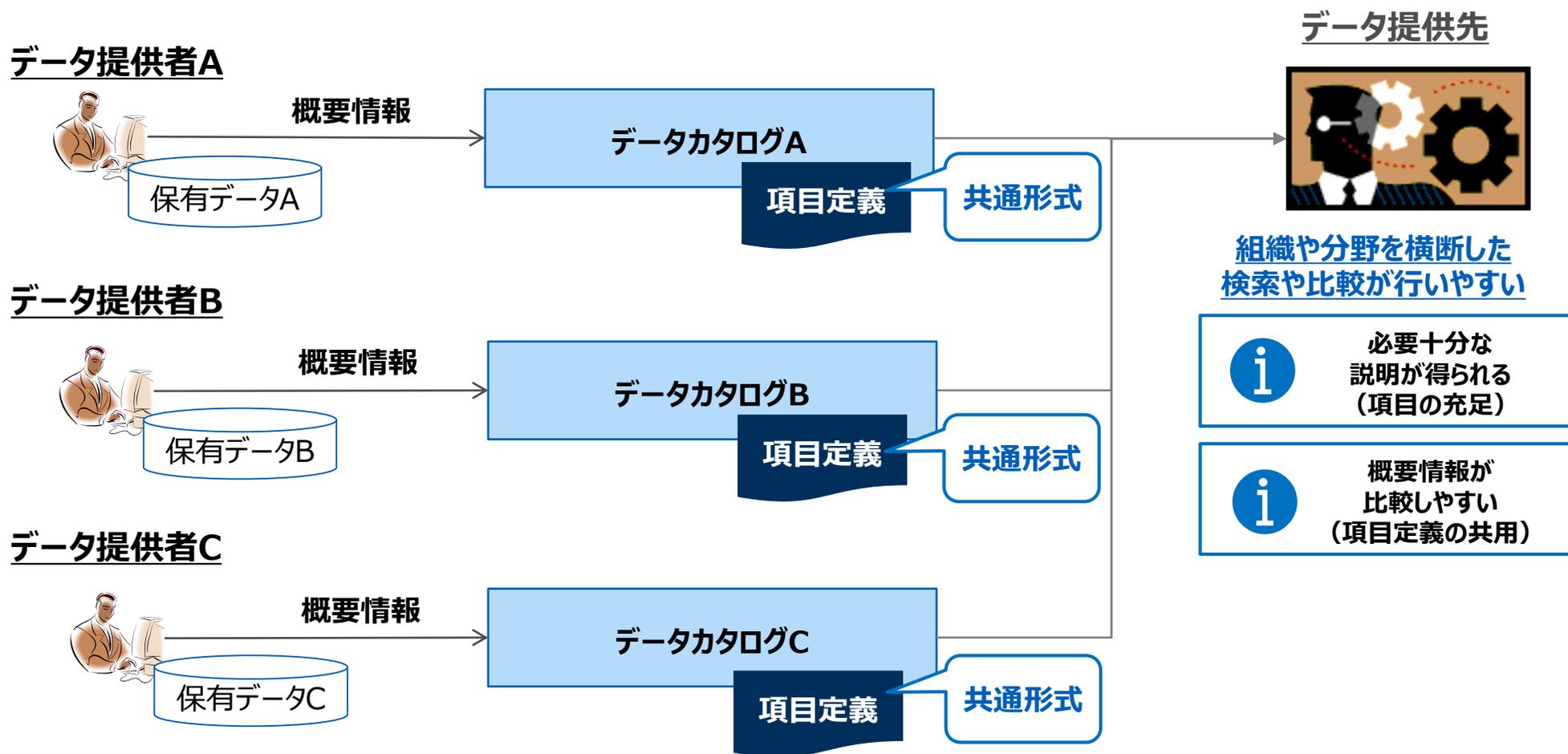
## 1.2 本ガイドライン策定の背景

- 我が国のデータ流通産業は勃興期にあり、様々な組織でデータの収集・提供が行われています。
- 様々な組織が独自の形式でデータカタログの項目を定義していると、検索や比較が煩雑になります。



## 1.3 本ガイドライン策定の狙い

- 本ガイドラインでは、共通形式でデータの概要を表現できる項目定義を目指しています。
- 本ガイドラインが様々な分野で活用されることで、組織や分野を横断しても、データカタログの検索、比較を容易に行えるようになるため、データ流通産業の活性化に貢献できると考えています。



## 2. 本ガイドラインにおける項目定義について

## 2.1 本ガイドラインにおける項目定義の特徴

### 特徴①：データの特徴や、データの利用例をわかりやすく表現

- データ提供者がデータ提供先にデータの詳細を伝えようとする、相手先に開示する情報が多くなりがちです。本ガイドラインの項目定義書では、「詳しく伝える」「簡潔に伝える」を両立できるように、データの概要を要約して記入できる項目を揃えています。
- データ提供先がどのようにデータを手に入れ、使えばいいか想像しやすいように、その収録内容、提供方式、利用例などを記入する項目を揃えています。

### 特徴②：静的データ、動的データへ対応

- 近年、IoTが注目されることで、センサー等で取得した動的データの共有・流通ニーズが高まっています。本ガイドラインの項目定義書では、集計表や報告書などの静的データだけでなく、動的データを扱えるように対応を行っています。

### 特徴③：データ取引等への対応

- 企業が社外にデータを提供するとき、データ取引の契約を締結することが増えています。こうしたデータ取引に対応するために、本ガイドラインの項目定義書では、契約ポリシーや利用条件などを記入する項目を揃えています。

※動的データとは、センサー機器やモバイル端末などを用いて取得され、時間経過に伴って生成ないし観測されるデータのことであり、ストリーミング配信されるなど更新頻度が比較的高いデータを指します。静的データは、更新頻度が比較的低い、すでに保存されたデータをいいます。

## 2.2 本ガイドラインにおけるデータカタログの構造

- 本ガイドラインのデータカタログの構造は、先行する国際標準、ガイドラインを参考にした4つの大構造部で構成されます。

### データカタログ本体部

- ✓ データカタログ全体やデータセットの内容、配信方法など、データを管理・整理するための必要事項を示す
- ✓ 国際標準化団体W3C DXWGの議論を参考とした

### データジャケット部

- ✓ 人が読み・理解することで、データの価値を発見するための基本事項を示す
- ✓ 国内外で利用が進む、データジャケット®<sup>注1</sup>の一部をサマリとして利用している

### データ詳細部

- ✓ データの種類に応じた詳細情報を示す
- ✓ データの種類に応じて、分野ごとに項目の差し替えを行う
- ✓ 動的データの場合には、観測活動、観測に用いたセンサー、観測期間・場所など、データの生成に関する詳細事項を示す
- ✓ 動的データでは、国際標準化団体W3C SDWWGの概念・用語を参考とした

### データ利用条件部

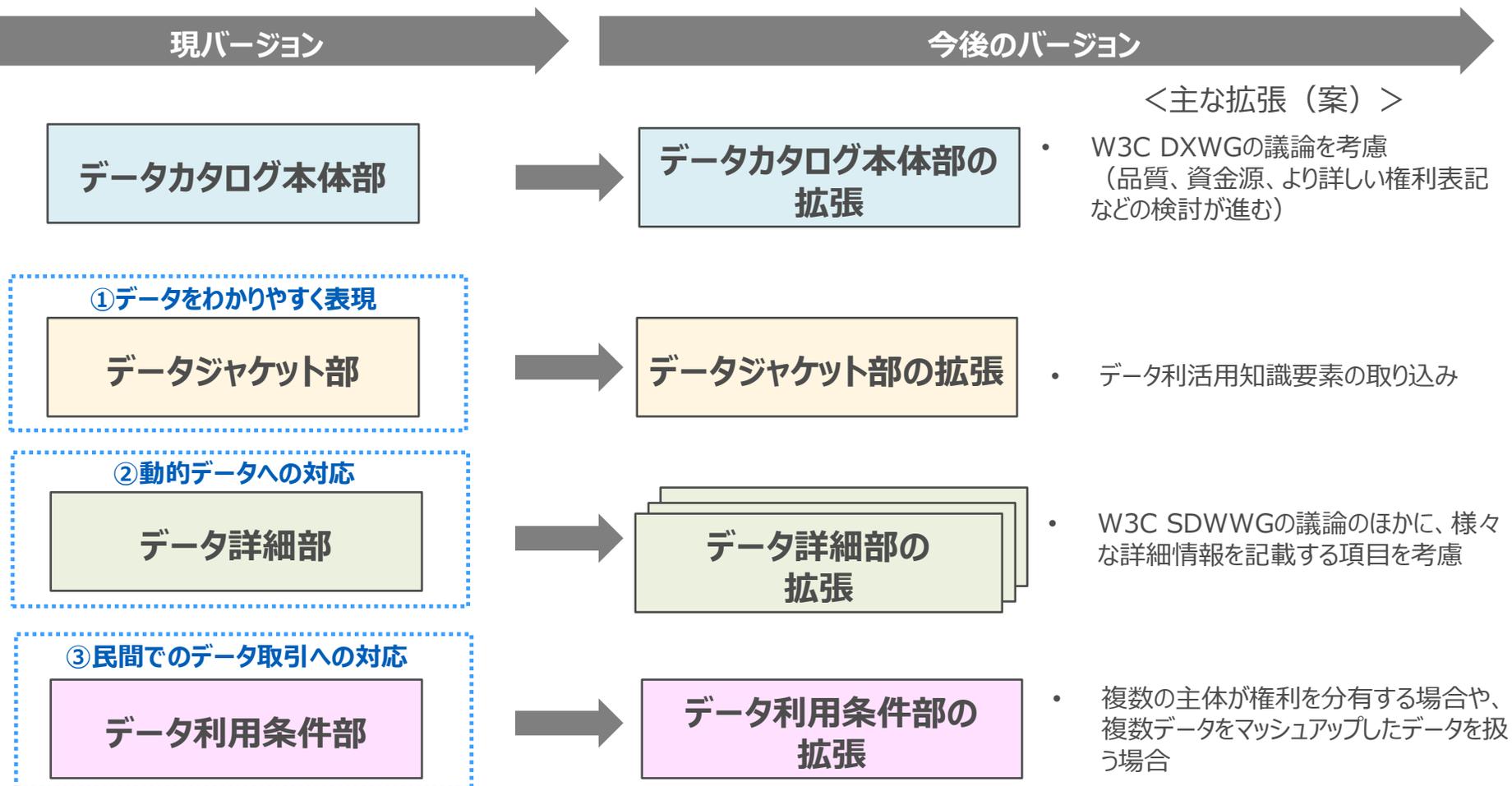
- ✓ 契約形態や第三者への開示範囲、利用期間、支払条件等の利用条件を示す
- ✓ 経済産業省「AI・データの利用に関する契約ガイドライン」（平成30年6月）の論点・ひな型（案）を参考とした

※データ利用条件部は、データの利用契約に関する概要を記入するもので、詳細は契約時に詰める必要があります

注1 データジャケット®については、P19を参照ください

## 2.3 本ガイドラインの位置づけ

- 今後も、項目定義のバージョンアップや、より多様なデータに対応するための拡張を行います。



## 2.4 本ガイドラインのデータカタログのイメージ

### ① データカタログ本体部

データ項目	データ例
カタログ	
カタログの題名	健康データマーケットプレイス
カタログ作成者	〇〇株式会社 データサービス事業部
カタログの発行日	2018-04-01
データセット	
データセットの名称	高齢者血圧測定データ
データセットの説明	高齢者を対象とするモニタ会員の血圧データ
データセットの出版者	株式会社△△ ヘルスケア事業部
データセットの発行日	2018-05-02
配信	
配信の名称	高齢者血圧測定データ(CSV)
直接ダウンロードURL	http://.../bp20180502.csv

### ② データジャケット部

データ項目	データ例
データジャケット	
データのタイトル	高齢者会員の日々の血圧測定データ
データの概要説明	高齢者を対象とするモニタ会員が保有する血圧計から、毎日朝夜の2回の測定～
データの共有条件	データセットの購入により共有可
データの変数の名前	最低血圧、最高血圧
データの分析方法	血圧に関する長期間、大規模な治験データを蓄積データに基づき～

### ③ データ詳細部

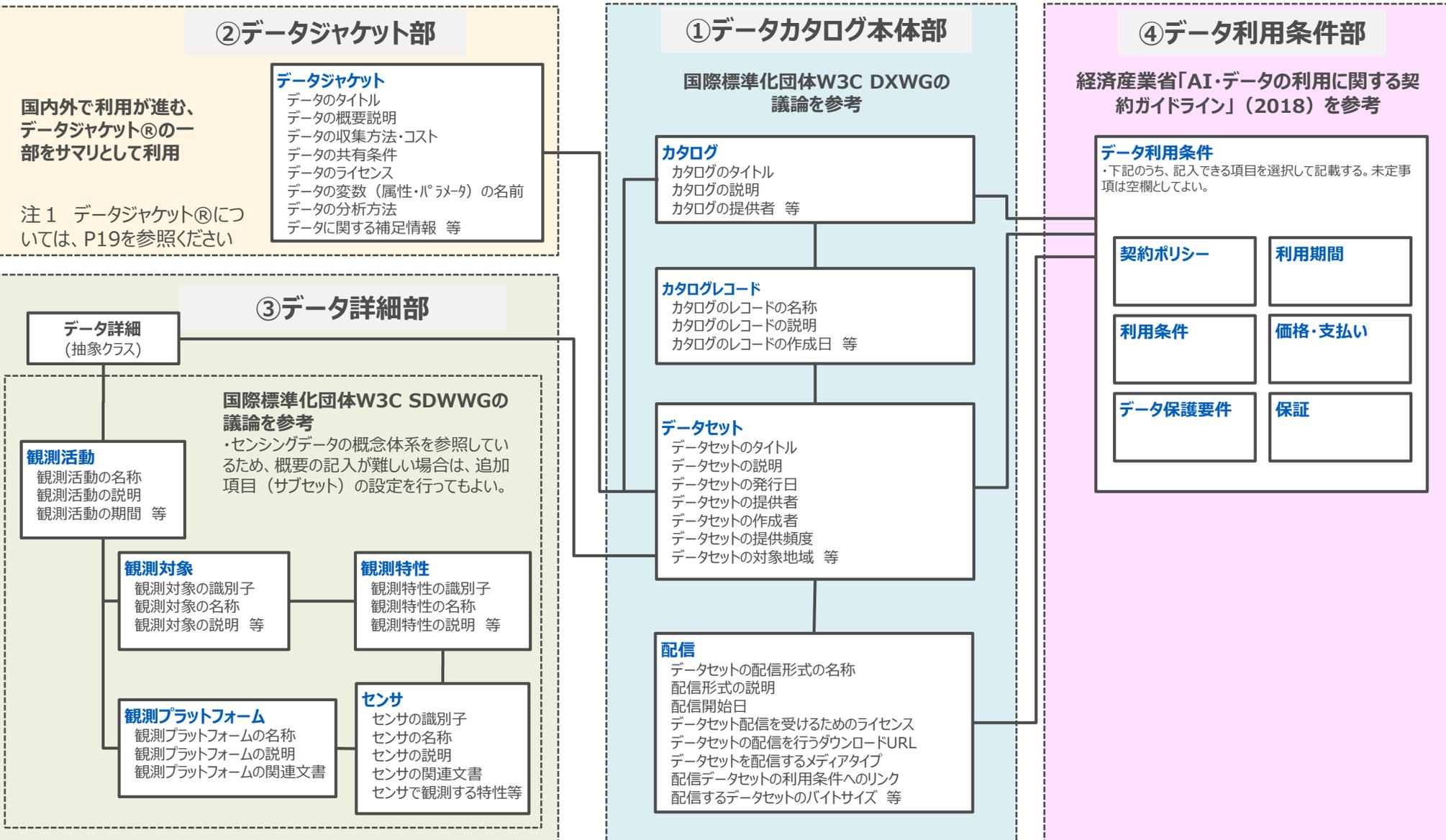
データ項目	データ例
観測活動	
観測の名称	高齢者を対象とした生活空間での血圧測定
センサ	
センサの名称	圧力センサ P-XXX-1234
観測対象	
観測対象の名称	健康サークル血圧測定対象者グループ
観測特性	
名称	最低血圧、最高血圧
単位	qudt-unit-1-1:MillimeterOfMercury (mmHg)
観測プラットフォーム	
名称	電子血圧計 YYY-BP000

### ④ データ利用条件部

データ項目	データ例
契約ポリシー	
契約形態	譲渡
利用用途	商用利用、研究利用
利用条件	
開示範囲	同一法人内であれば開示可
派生データの権利帰属	データ利用者に帰属する
データ保護要件	
データの管理	データは、契約対象者しかアクセスできない場所に保管
価格・支払い	
価格帯	1万円未満

## 2.5 本ガイドラインのデータカタログの概念図

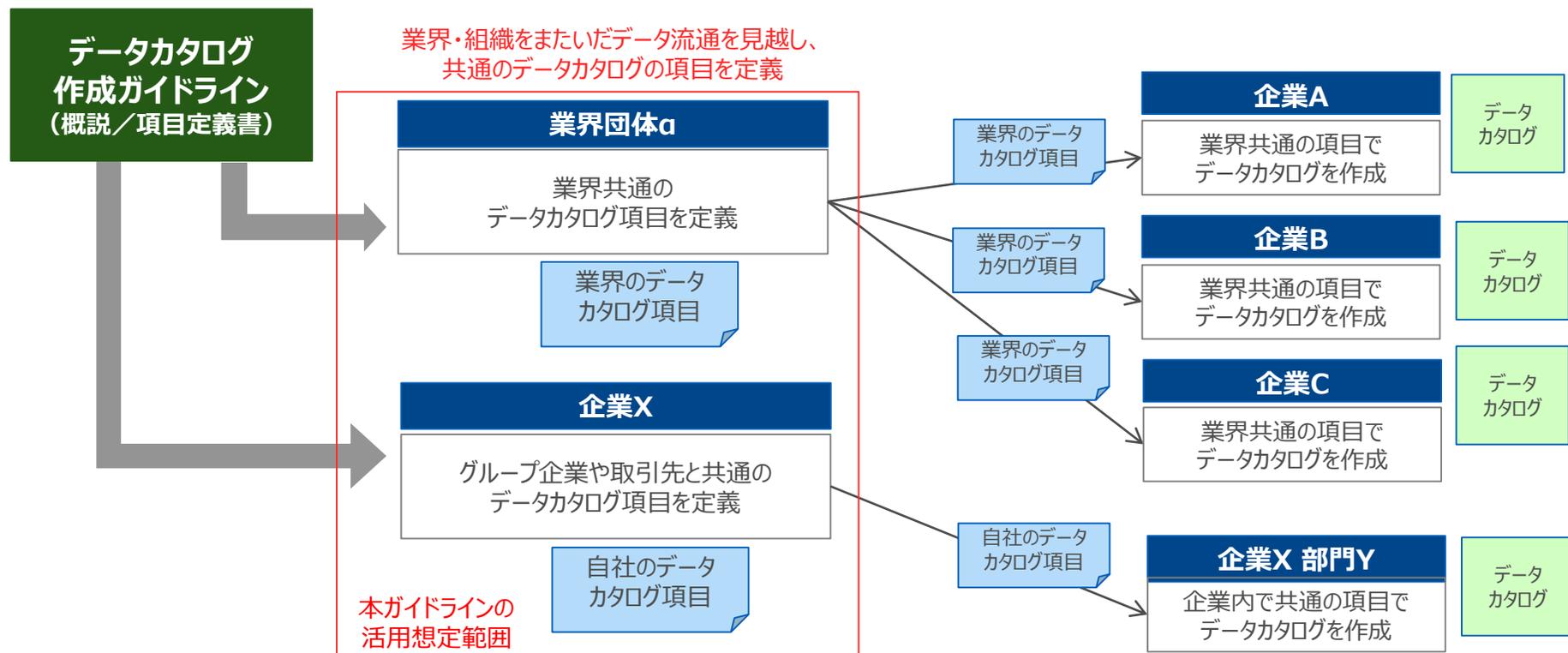
◎カタログの使用法やデータの種類の種類、利用条件に応じて、適宜、大構造部は取捨選択して使えます（データカタログ本体部は必須）。



# 3. 本ガイドラインを用いた データカタログ項目策定の流れ

### 3.1 本ガイドラインで活用を想定するシーン

- 本ガイドラインは、業界団体や企業等で、データカタログ項目を策定する際に、参照頂くことを想定しています。



## 3.2 本ガイドラインを用いたデータカタログ項目策定の流れ

### STEP 1

採用する大構造部を選ぶ

- 作成したいデータカタログの特徴を明確化し、大構造部の4パートを選択して下さい。
  - ✓ 複数のデータカタログを共通項目で管理したい場合は、データカタログ本体部
  - ✓ 最低限の必要事項を速やかに記入し、他の方に伝えたい場合は、データジャケット部
  - ✓ データを取得したときの状況や方法を詳しく表現したい場合は、データ詳細部
  - ✓ データを外部へ提供するとき、権利帰属やライセンス、価格を決める場合は、データ利用条件部

### STEP 2

採用するデータ項目を選び、不足する項目を見極める

- 各項目には、出現回数（出現回数については、定義書解説を参照下さい）が設定されています。それに従い、必須／任意の使い分けを判断して下さい。
- 項目定義書に収録された、どの項目を用いるかを選択して下さい。
- 項目定義書に存在しない項目が必要なときは、任意に追加して下さい。

### STEP 3

入力ルールを確認する

- 選択肢が複数あったり、数値・文字など記入指定のある項目もあります。

### STEP 4

サンプルデータを用いた記入テスト

- サンプルデータを入力し、策定したデータカタログ項目を検証して下さい。

### STEP 5

作成の手引きや修正提案の作成

- 必要に応じて、実務者向けに、業界・業種ごとの「作成の手引き」を作成して下さい。

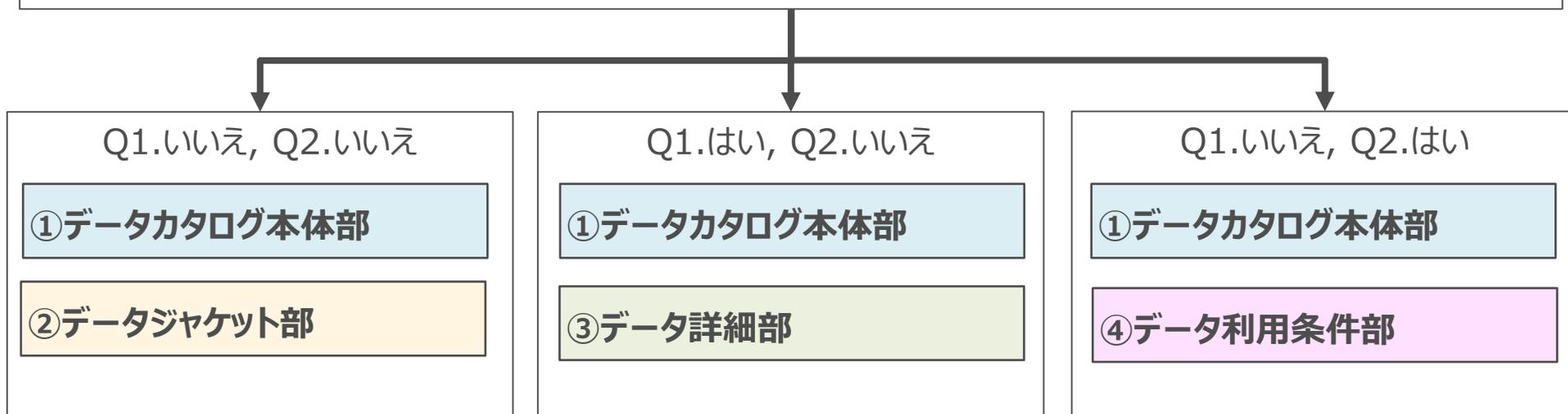
## 3.3 データカタログ項目策定のステップ

### STEP 1

採用する大構造部を選ぶ

- 作成したいデータカタログの特徴を明確化し、大構造部の4パートを選択して下さい。

- Q1.動的データの概要を、データカタログに記入しますか？
- Q2.データを社外提供する等で、データの利用条件を記入する必要がありますか？



### 3.3 データカタログ項目策定のステップ

#### STEP2

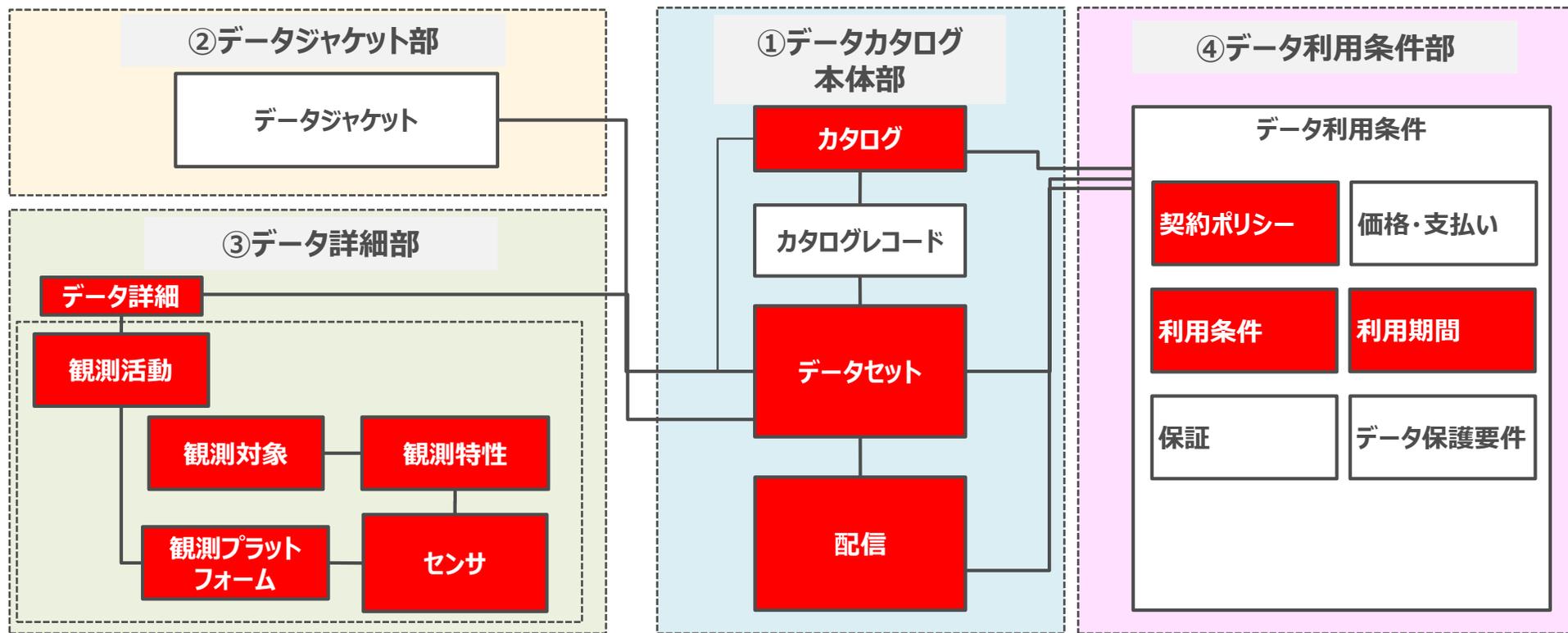
採用するデータ項目を選び、不足する項目を見極める

- 出現回数の設定を参考に、必須／任意の使い分け、項目の過不足を確かめて下さい。

(例) カタログレコード、データジャケットを使わず、データ詳細部を使い、利用条件を任意に選択する

**選択**

**非選択**



### 3.3 データカタログ項目策定のステップ

#### STEP3

#### 入力ルールを確認する

- 項目定義書の説明欄に、入力ルールがあります。
- 選択肢が複数あったり、数値・文字など記入指定のある項目もあります。

大構造	見出し	出現回数	値域 (データタイプ)	説明	サンプル値
④データ利用 条件部	利用用途	0..n	選択肢列挙型	<p>&lt;説明&gt; どの用途であれば、利用を認めるかを記載してください、</p> <p>&lt;入力ルール&gt; 以下の選択肢から選択してください。(複数選択) 商用利用／研究利用／教育利用／制限なし／個別調整／その他の複数選択</p>	商用利用

項目定義書には [0..n] とあるが、  
自組織では [1..1] としたい

選択肢を、商用利用、研究利用、  
個別調整にしたい

## 3.3 データカタログ項目策定のステップ

### STEP4

#### サンプルデータを用いた記入テスト

- サンプルデータを入力して、策定したデータカタログ項目を検証して下さい。

### STEP5

#### 作成の手引きや修正提案の作成

- 必要に応じて、実務者向けに、業界・業種ごとの「作成の手引き」を作成して下さい。

- データジャケット®は、大澤幸生・早矢仕晃章（東京大学）の登録商標（商標登録第6088442号（T6088442））です。
- 商標登録第6088442号（T6088442）は、当協議会の定める「産業財産権等の取扱いについての基本指針」および「産業財産権等の取扱いについての運用細則」に基づき、権利者である東京大学大澤幸生・早矢仕晃章より、「産業財産権等の実施許諾に係る声明書」を受領し、登録商標の許諾を得ています。
- データジャケット®については、以下の文献をご参照ください。データジャケット®は今後も発展が予定されています。

(a)大澤幸生, “データジャケット –創造的コミュニケーションのあるデータ市場のために–”, 人工知能学会誌, Big Data Becomes Personal : 発見情報学が拓くヘルス&ウェルネス-特集, Vol.29, No.6, pp.622-627. 2014.

URL:

[https://jsai.ixsq.nii.ac.jp/ej/?action=pages\\_view\\_main&active\\_action=repository\\_view\\_main\\_item\\_detail&item\\_id=1720&item\\_no=1&page\\_id=13&block\\_id=23](https://jsai.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=1720&item_no=1&page_id=13&block_id=23)

(b)早矢仕晃章, 大澤幸生, “Data Jacket Store: データ利活用知識構造化と検索システム”, 人工知能学会論文誌, Vol.31, No.5, 2016. DOI: 10.1527/tjsai.A-G15

URL :

[https://www.jstage.jst.go.jp/article/tjsai/31/5/31\\_A-G15/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/tjsai/31/5/31_A-G15/_pdf/-char/ja)

(c)大澤幸生 (編著), 早矢仕晃章, 秋元正博, 久代紀之, 中村潤, 寺本正彦, “データ市場”, 近代科学社, 2017.

(d)早矢仕晃章, 大澤幸生, “データジャケットを用いた異分野データ連携”, 人工知能学会誌, AIとデータ-データに基づく意思決定と社会イノベーション創出-特集, Vol.33, No.2, pp.140-148. 2018.

URL :

[https://jsai.ixsq.nii.ac.jp/ej/?action=pages\\_view\\_main&active\\_action=repository\\_view\\_main\\_item\\_detail&item\\_id=9088&item\\_no=1&page\\_id=13&block\\_id=23](https://jsai.ixsq.nii.ac.jp/ej/?action=pages_view_main&active_action=repository_view_main_item_detail&item_id=9088&item_no=1&page_id=13&block_id=23)

## データカタログ作成ガイドライン 概説

2019年1月xx日 V1.1発行

著 者：一般社団法人データ流通推進協議会 技術基準検討委員会  
発行者：一般社団法人データ流通推進協議会 技術基準検討委員会  
発行所：一般社団法人データ流通推進協議会

〒107-0061 東京都港区北青山2丁目7-26  
TEL：03-6890-0969  
Mail：info@data-trading.org

本書の利用で直接又は間接に生じた損害に対して、当会は一切の責任を負いません。

本書の用語は「データ取引市場運営事業者認定基準の説明」（2018年8月23日）に従っています。  
URL：<https://data-trading.org/2018/09/28/pressrelease/>

