

「IT標準化支援プログラム」勉強会 データカタログ篇

2019年2月



本資料について

- 本資料は、「平成29年度産業データ共有促進事業補助金」（以下、「本事業」）において、一般社団法人環境共創イニシアチブより、一般社団法人データ流通推進協議会（以下、DTA）に委託した「採択者向け勉強会」で使用したものです。
- データ共有を検討されている皆様にも広くご活用いただくため、一部改編を行い公開を行っております。
- 本資料は、DTAが自主事業として検討を実施した「データカタログ作成ガイドラインV1.1（中間とりまとめ）」をベースに作成しております。

変更履歴

日付	内容
2018年10月	初版作成
2019年2月	一般社団法人データ流通推進協議会にて策定している「データカタログ作成ガイドラインV1.1（中間とりまとめ）」に基づき、内容を更新

1. データカタログとは何か？

「データとは何か」と言われて思い浮かべるものは、その人ごとに千差万別で、目的や用途も多種多様。

・ 数値や単位もデータ

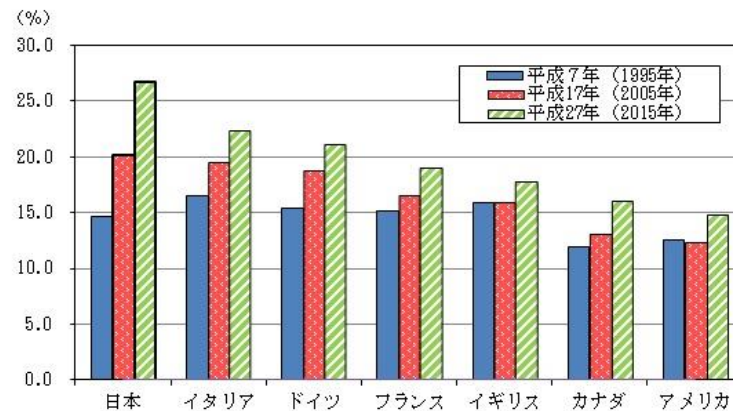
5,000兆円

・ 紙に記された「文字や数字」、文書の「画像」もデータ

一、汽罐型式 左記に蒐集せる五拾個所一〇七基として製作せられしものなり、今暖房専用汽罐につき殊に米國製は總數の七割強を占む、火床型式は通常通風式にして更に下部火床の揺動式なるもの等もあり、二、汽壓或は罐水溫度及給水溫度 暖房専用汽罐水汽罐にありては罐水溫度華氏一四〇—二〇〇度なよりの還水を使用し其の溫度何れも華氏一〇〇度以下、三、通風方法及通風度測定 通風方法は日本石油煙突の高さは最低七五尺（内務省令市内最低限度）通風度測定は暖房専用汽罐に採用したるものは皆ものなり之を適宜調整し以て發煙、爐溫低下、煙に入りて開始するものあり、表中泉橋慈善病院の一

出典：「東京市内に於ける主なる暖房汽罐の現況」（燃料協會出張所，1925）（画像出所：国立国会図書館デジタルコレクション）

・ 統計グラフもデータ



出典：総務省統計局「統計 Today No.114」

・ 数表や、個々の値もデータ

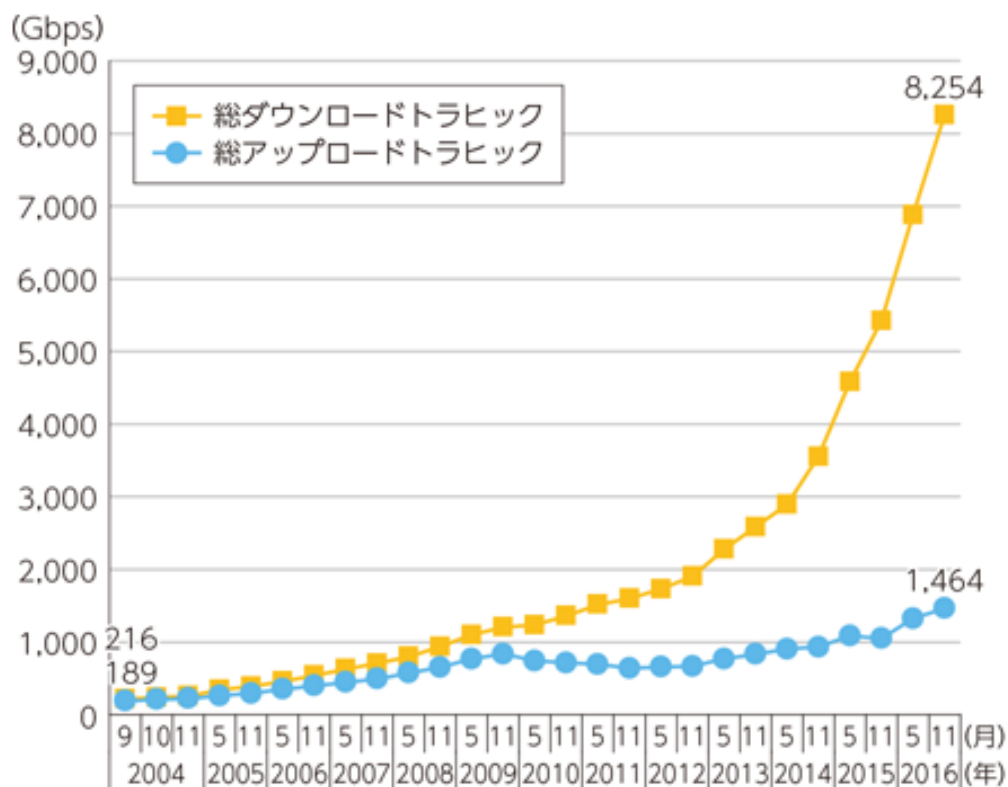
datetime_utc	_conds	_rain	_snow	_tempm	_thunder	_tornado
19961101-11:00	Smoke	0	0	30	0	0
19961101-12:00	Smoke	0	0	28	0	0
19961101-13:00	Smoke	0	0	24	0	0
19961101-14:00	Smoke	0	0	24	0	0
19961101-16:00	Smoke	0	0	23	0	0

出典：「Delhi Weather Data from 1997 to 2016 december」（Kaggle Dataset）
<https://www.kaggle.com/mahirkukreja/delhi-weather-data>

社会に流通するデータの量は急激に増えているが、 すべてを消費できるわけではなく、「概要」「要約」が欠かせない。

社会に流通するデータ量の急増

図表2-1-2-1 我が国のトラフィックの推移（ブロードバンド）



出典：総務省「我が国のインターネットにおけるトラフィックの集計・試算」

人間が消費できる情報量の限界

4,266万2,279点
国立国会図書館の蔵書数（2017年）

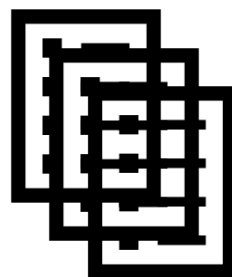
2万1,915冊
推計例：一生で何冊読めるか？
（註：毎日1冊・60年間読んだ場合）

出典：読書猿Classic「一生で何冊読めるか？／Readers' Math Shortcuts」（2010）
（<https://readingmonkey.blog.fc2.com/blog-entry-239.html>）

近年では、データの「概要」「要約」を効率よく整理するために、世界各地で「データカタログ」が作成され、公表されている。



対象
(例：書籍)



データ
(本文テキストや
挿入図など…etc.)



データカタログ
(目次、書名、著者、値段
などの概要をまとめたもの)

なぜ、データカタログが重要なのか

データカタログは、データ分析者がデータを見つけ、理解することを助けると紹介されてきた。データカタログがなかった頃は、多くのデータ分析者が、データセットの存在や、それらの内容、品質、使い勝手といったことがよく見えないまま、目隠しで働いていたようなものだ。

出典：デイヴ・ウェルズ「The Ultimate Guide to Data Catalogs Key Things to Consider When Selecting a Data Catalog」（Eckerson grope, 2018年5月）を試訳

【行政】日本では内閣官房をはじめ各省庁が連携し、行政データを一般公開するData.go.jpが提供されている。

日本語 | English

DATA ● GO.JP データカタログサイト

お知らせ データ データベースサイト一覧 公共データ活用事例 コミュニケーション 開発者向け情報

データセットを検索...

[意見受付コーナー](#)にて、オープンデータの掲載に関する御要望等を受け付けております（匿名可）

データ

データセット 組織 グループ タグ

オープンデータの取組（リンク集）

オープンデータに関する方針・決定
公共データ活用事例一覧
データベースサイト一覧

コミュニケーション

意見受付コーナー 意見・回答公開コーナー
掲載データ利用の御連絡

出典：http://www.data.go.jp

Data.go.jpでは、DCAT（後述）を採用。 どの省庁のデータも、同じ項目で「概要」が調べられる。

検索結果の例

The screenshot shows the Data.go.jp search results page. The header includes the Data.go.jp logo and navigation links. The main content area displays the search results for the query '建設工事受注動態統計調査_大手50社_月次_2018年3月'. The results are organized into a table with columns for 'フィールド' (Field) and '値' (Value). The table includes fields such as 'タイトル' (Title), '説明' (Description), '公表組織名' (Publishing Organization Name), '連絡先' (Contact Information), '作成者' (Creator), 'タグ' (Tags), 'リリース日' (Release Date), '作成頻度' (Creation Frequency), '公開ウェブページ' (Public Web Page), and '対象地域' (Target Area). The table is filtered by '建設業' (Construction Industry) and '統計' (Statistics). The table is also sorted by 'リリース日' (Release Date) in descending order. The table shows 17 results. The first result is '建設工事受注動態統計調査_大手50社_月次_2018年3月' by '国土交通省' (Ministry of Land, Infrastructure, and Transport). The table is also filtered by '建設業' (Construction Industry) and '統計' (Statistics). The table is also sorted by 'リリース日' (Release Date) in descending order. The table shows 17 results. The first result is '建設工事受注動態統計調査_大手50社_月次_2018年3月' by '国土交通省' (Ministry of Land, Infrastructure, and Transport).

データカタログの表示例

追加情報

フィールド	値
タイトル	建設工事受注動態統計調査_大手50社_月次_2018年3月
説明	
公表組織名	国土交通省
連絡先	総合政策局情報政策課建設経済統計調査室
作成者	総合政策局情報政策課建設経済統計調査室
タグ	construction,statistics,statistics_survey_r...,建設業,統計,統計調査結果
リリース日	2018-04-27
作成頻度	1月
公開ウェブページ	https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00600130&tstat=000001015811&cycle=1&year=20180&month=11010303&tclass1=000001015812&result_back=1&se=cond2=1
対象地域	

出典：http://www.data.go.jp

出典：http://www.data.go.jp

Copyright©Sustainable open Innovation Initiative. All Rights Reserved.

【学術例】地球環境情報プラットフォームDIASは、 学術機関・研究者が作成した関連データを共同管理する。 (RESTEC、東大、京大、名古屋大、北海道大、北見工業大、NIIなどが参画)

ユーザ用メニュー > データセット一覧 ログアウト

データセット一覧

担当分 担当されているデータセットの一覧です。メタデータを編集するには、「メタデータ作成状況」列の緑の矢印をクリックして、データセットの履歴に進んでください。

無し

新しくデータセットを追加したい場合は dias-office@dias.jp.net までご連絡ください。

その他 ご担当分以外のデータセットを閲覧することができます。

番号	データセットID	データセット名(日)	データセット名(英)	メモ	メタデータ作成状況 日/英	関連プロジェクトID	DOI	サムネイル
5	MIRAI_CTD	みらいCTDデータセット	MIRAI CTD dataset	Japanese: ...	Complete / Complete 	DIAS		
1	CEOP_Model_Grid_BoM	CEOP Model Output : 3D Gridded BoM data	CEOP Model Output : 3D Gridded BoM data	Japanese: ...	Temporary / Complete 	DIAS		
6	CEOP_Model_Grid_CPTEC	CEOP Model Output : 3D Gridded CPTEC data	CEOP Model Output : 3D Gridded CPTEC data	Japanese: ...	Temporary / Complete 	DIAS		
7	CEOP_Model_Grid_ECMWF	CEOP Model Output : 3D Gridded ECMWF data	CEOP Model Output : 3D Gridded ECMWF data	Japanese: ...	Temporary / Complete 	DIAS		
8	CEOP_Model_Grid_ECPC_RIT3	CEOP Model Output : 3D Gridded ECPC RIT36 data	CEOP Model Output : 3D Gridded ECPC RIT36 data	Japanese: ...	Temporary / Complete 	DIAS		



DIASでは「ドキュメントメタデータ」という標準の記入項目を整備しており、分野内で共通したメタデータの管理を実現。

1. IDENTIFICATION INFORMATION

Name	AGURAM Ground Information Database
Edition	1
Abbreviation	AGURAM
DOI	doi:10.20783/DIAS.207
Metadata Identifier	AGURAM_GI20171026153503-DIAS20170725102541-en

データセットの
名称やバージョン、ID

2. CONTACT

2.1 CONTACT on DATASET

Name	(研究者名)
Organization	Earth Observation Data Integration & Fusion Research Initiative (EDITORIA), University of Tokyo
Address	Cw-503 Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505, Japan
TEL	(電話番号)
FAX	(FAX番号)
E-mail	(個人メールアドレス)

データ作成者の
所属や連絡先

2.2 CONTACT on PROJECT

出典： http://metadata.diasjp.net/dmm/doc/AGURAM_GI-DIAS-en.html

【民間例】Google Dataset Searchでは、Webに一般公開されたデータを検索できる。

Google Dataset Search

trade



概要



フィードバック



Direction of Trade Statistics (DOTS)

data.world

更新日: May 25, 2018



日本 Trade Balance

www.ceicdata.com



ポーランド Trade Balance

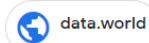
www.ceicdata.com



Foreign Trade

catalog.data.gov
datamirror.org

Direction of Trade Statistics (DOTS)



データセットの更新日 May 25, 2018

データセットの公開日 Aug 21, 2017

データセットの提供元

data.world

プロバイダからの利用可能なダウンロード形式

ZIP

説明

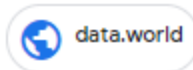
The Direction of Trade Statistics (DOTS) presents current figures on the value of merchandise exports and imports disaggregated according to a country's primary trade partners. Area and world aggregates are included in the display of trade flows between major areas of the world. Reported data is supplemented by estimates whenever such data is not available or current. Imports are reported on a cost, insurance and freight (CIF) basis and exports are reported on a free on board (FOB) basis, with the

出典 : Google Dataset Search (β版) <https://toolbox.google.com/datasetsearch/>

Google Dataset Searchでは、 データカタログを検索・表示できる。

データセットの 名称や更新日

Direction of Trade Statistics (DOTS)



データセットの更新日 May 25, 2018

データセットの公開日 Aug 21, 2017

データセットの提供元
data.world

プロバイダからの利用可能なダウンロード形式
ZIP

提供元や ダウンロード形式

概要説明

説明

The Direction of Trade Statistics (DOTS) presents current figures on aggregates are included in the display of trade flows between major trading partners reported on a cost, insurance and freight (CIF) basis and exports. The series data includes estimates derived from reports of partner countries.

Geographic Coverage: DOTS covers all IMF member states, some non-member countries, the world, and major areas.

Temporal Coverage: Monthly and quarterly data are available starting 1960. Annual data are available starting 1947.

Latest Update Date as of Download: 07/25/2017

The Data is provided to Users "as is" and without warranty of any kind, either express or implied, including, without limitation, any implied warranties or conditions of merchantability, fitness for a particular purpose, or noninfringement.

The IMF Data is available free of charge from the IMF.

Users are prohibited from infringing upon the integrity of the IMF data and in particular shall refrain from any act of alteration or modification of the IMF data. If the user materially transforms the IMF data, this must be stated explicitly along with the required source citation.

All other terms set forth in the IMF's general terms and conditions (<http://www.imf.org/external/terms.htm>) shall continue to apply.

Source: <http://data.imf.org/?sk=9D6028D4-F14A-464C-A2F2-59B2CD424B85>

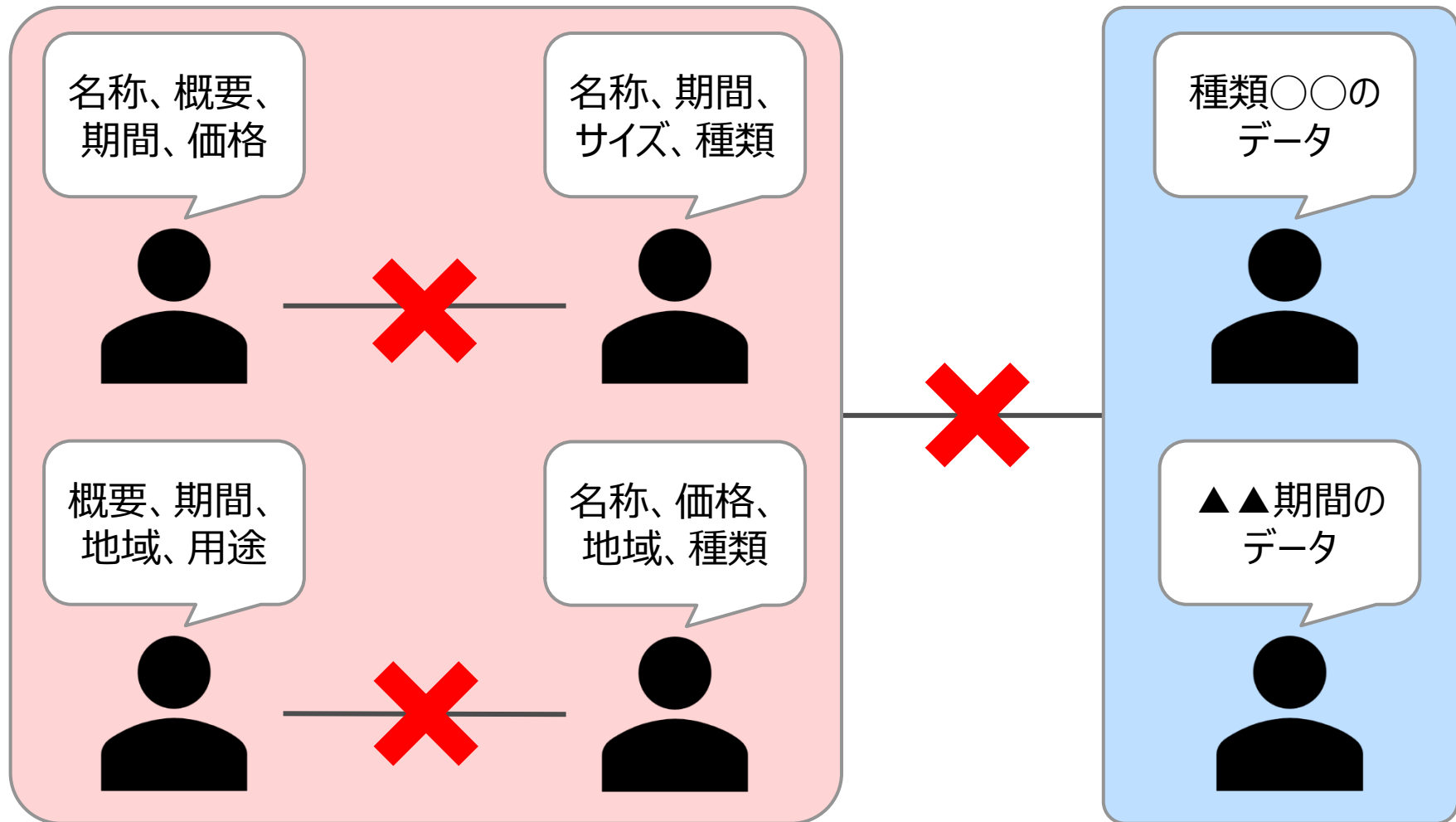
地域の範囲や 時間の範囲

利用規約へのリンク データソースのURL

データカタログの項目が不揃い・未整備のままだと、 データ管理が煩雑になったり、利用者が検索しづらくなる。

組織内で、どこに何のデータがあるか分かりづらい

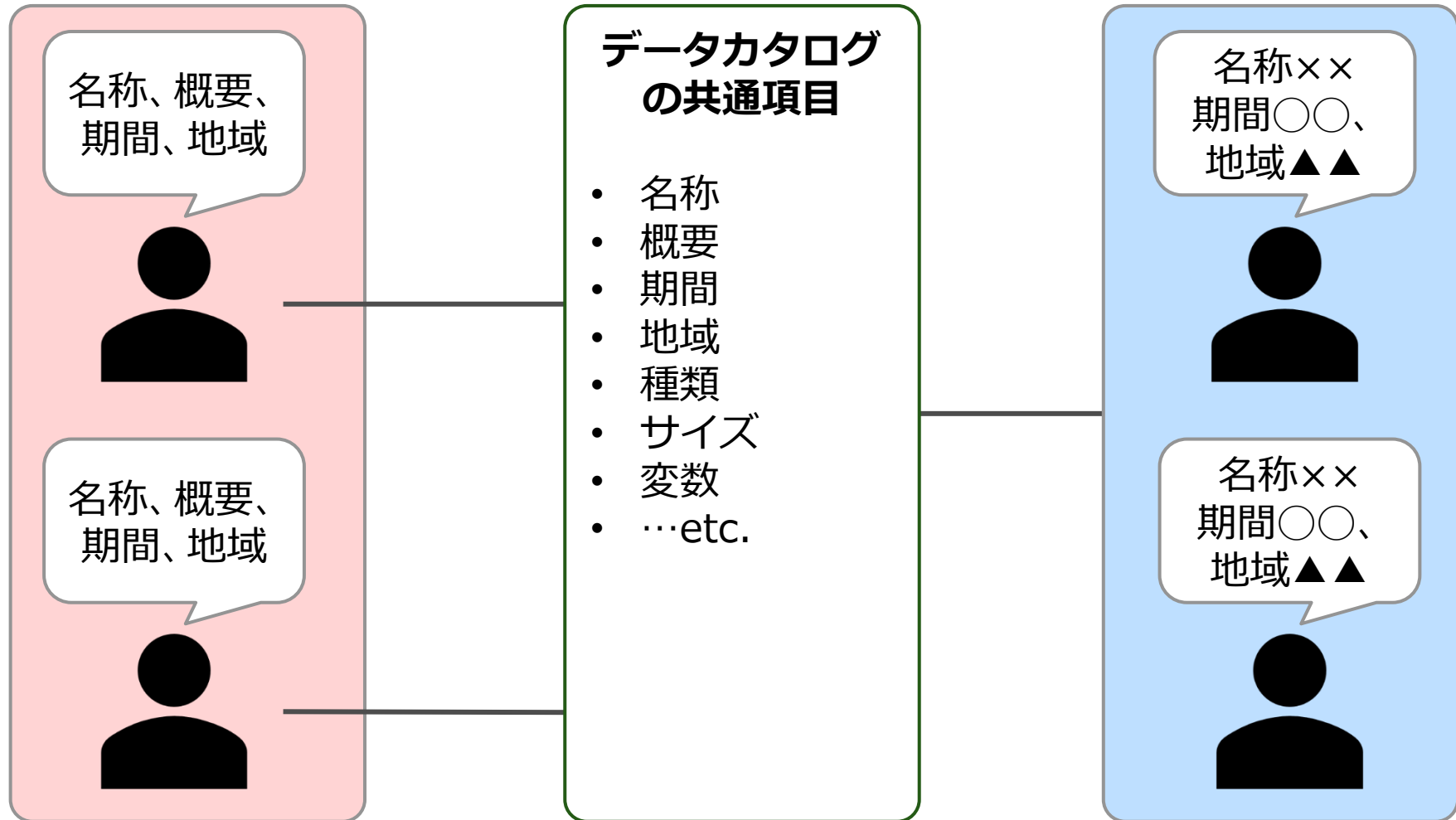
組織外からデータを見つけ出せない



データカタログの共通項目が業界・分野ごとにあれば、 データを受け渡すとき、必要な情報が何かを相手に伝えやすい。

どんな情報を表示すべきか分かる

何を頼りに探せばいいか分かる



データカタログの項目は、すでに国際標準化が行われている。 海外では「事実上の標準」を狙った、民間のサービス競争が進む。

- W3Cによる国際標準DCAT



Data Catalog Vocabulary (DCAT)

W3C Recommendation 16 January 2014

This version:

<http://www.w3.org/TR/2014/REC-vocab-dcat-20140116/>

Latest published version:

<http://www.w3.org/TR/vocab-dcat/>

Implementation report:

http://www.w3.org/2011/gld/wiki/DCAT_Implementations

Previous version:

<http://www.w3.org/TR/2013/PR-vocab-dcat-20131217/>

Editors:

[Fadi Maali](#), [DERI, NUI Galway](#)

[John Erickson](#), [Tetherless World Constellation \(RPI\)](#)

Contributors:

[Phil Archer](#), [W3C/ERCIM](#)

- ガートナー社によるデータカタログ分野のリーダー企業マップ°

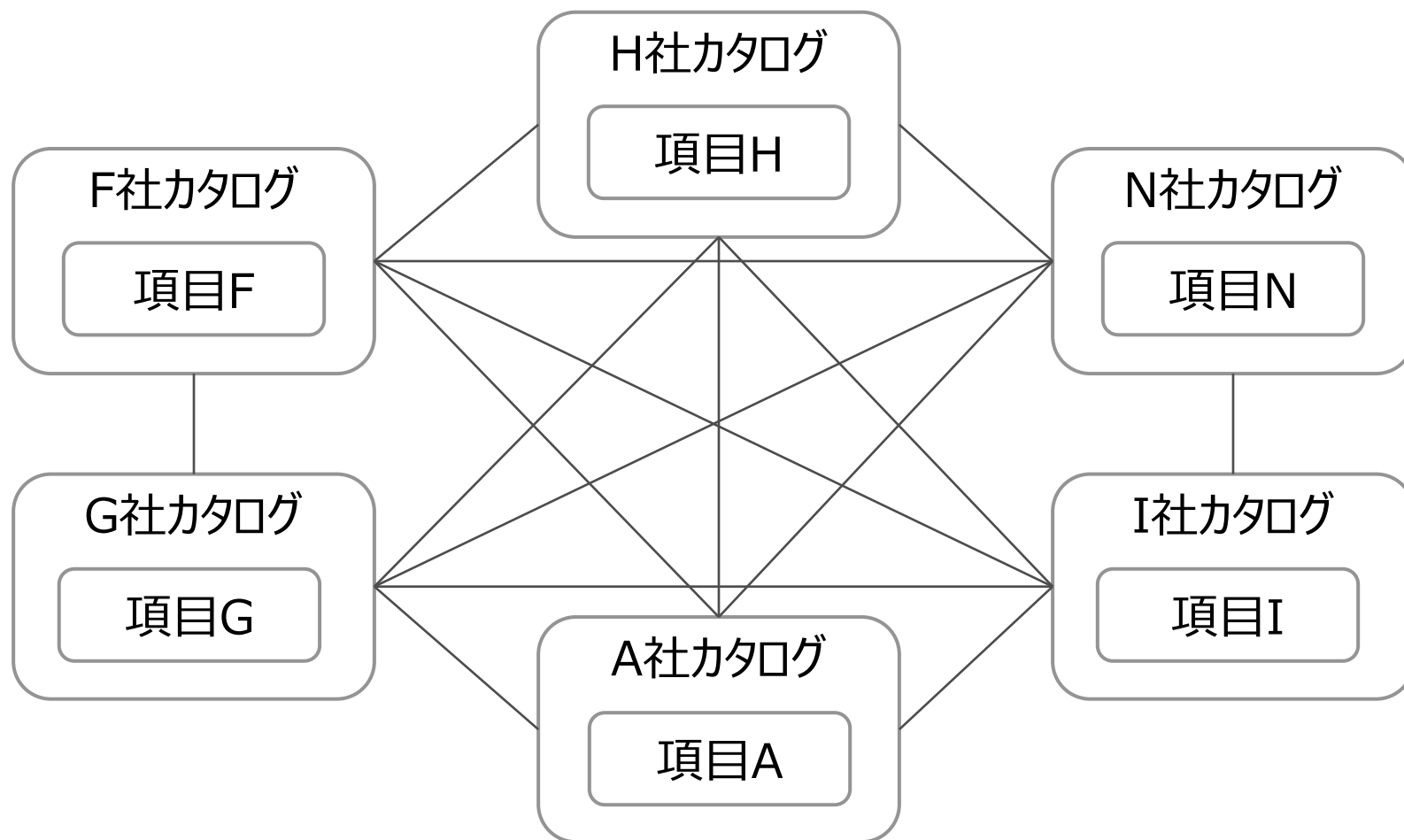
• 「Leaders」 カテゴリの企業

- Informatica
- Collibra
- Alation
- Smartlogic
- DATUM
- IBM
- Oracle
- Alex Solutions
- ASG Technologies

2. データカタログ作成ガイドラインV1.1 の位置づけと意義

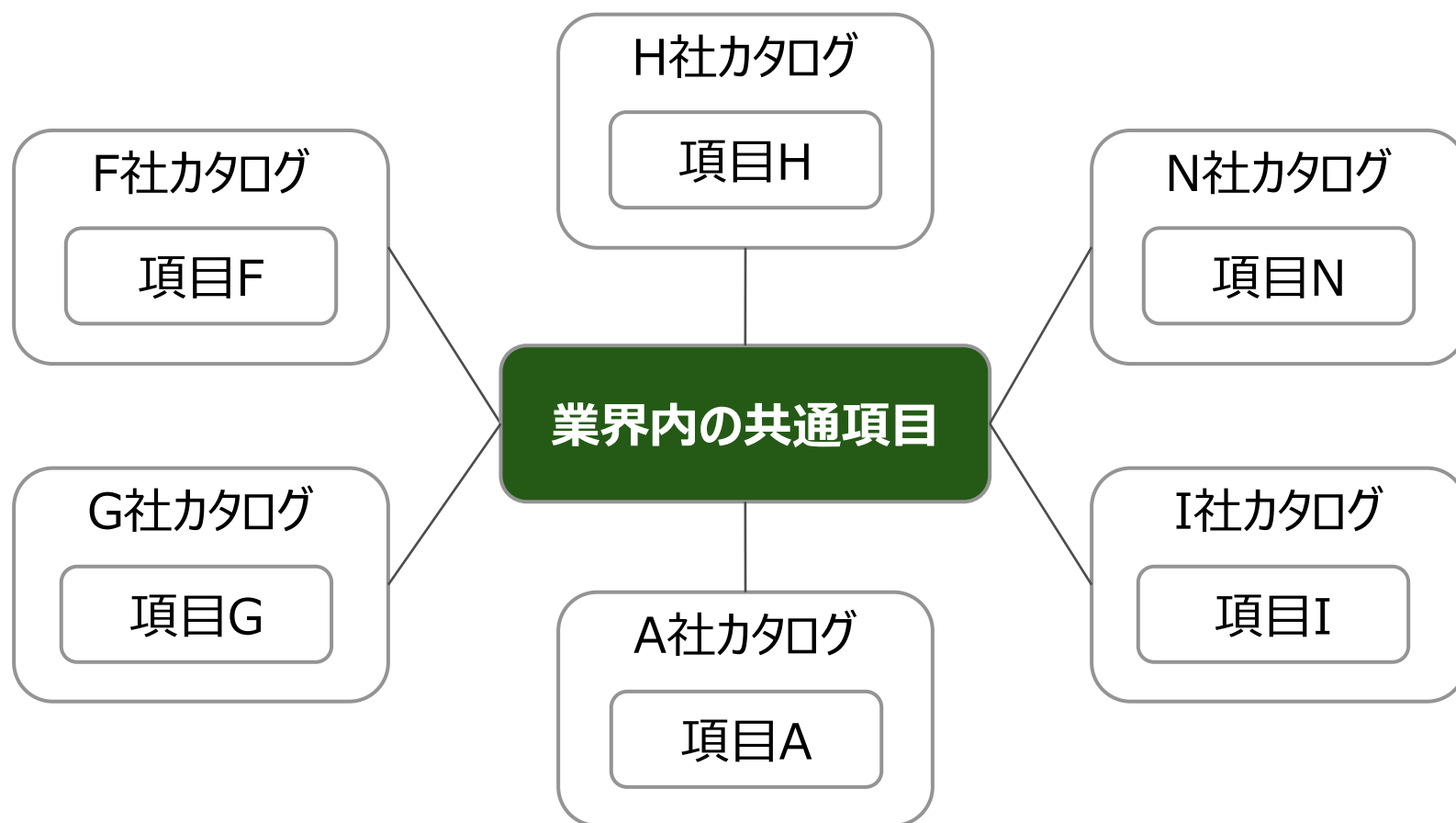
異なる組織の間で、その度ごとに個別調整していると、 関係者が増えれば増えるほど、膨大な「調整」が生じて非効率。

例：6社のカタログ項目がすべて異なると、なんと15回分もの「個別調整」が必要。大いなる無駄。



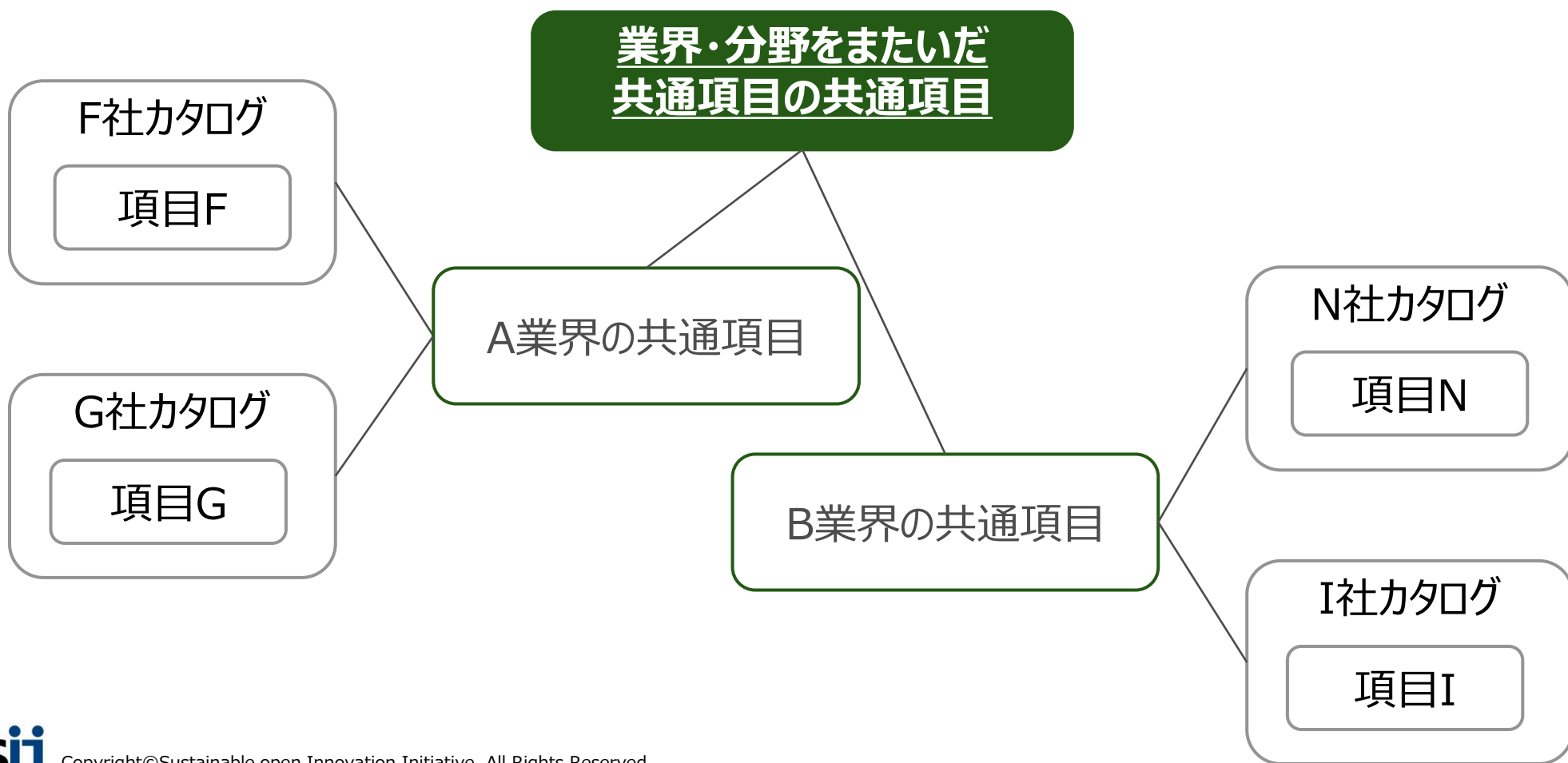
業界・分野ごとにデータカタログの共通項目を作成すれば、業界内で個別調整するコストが減る。

例：業界内で共通項目を決めておけば、「自組織の項目」と「共通項目」の差だけを留意すれば済む。



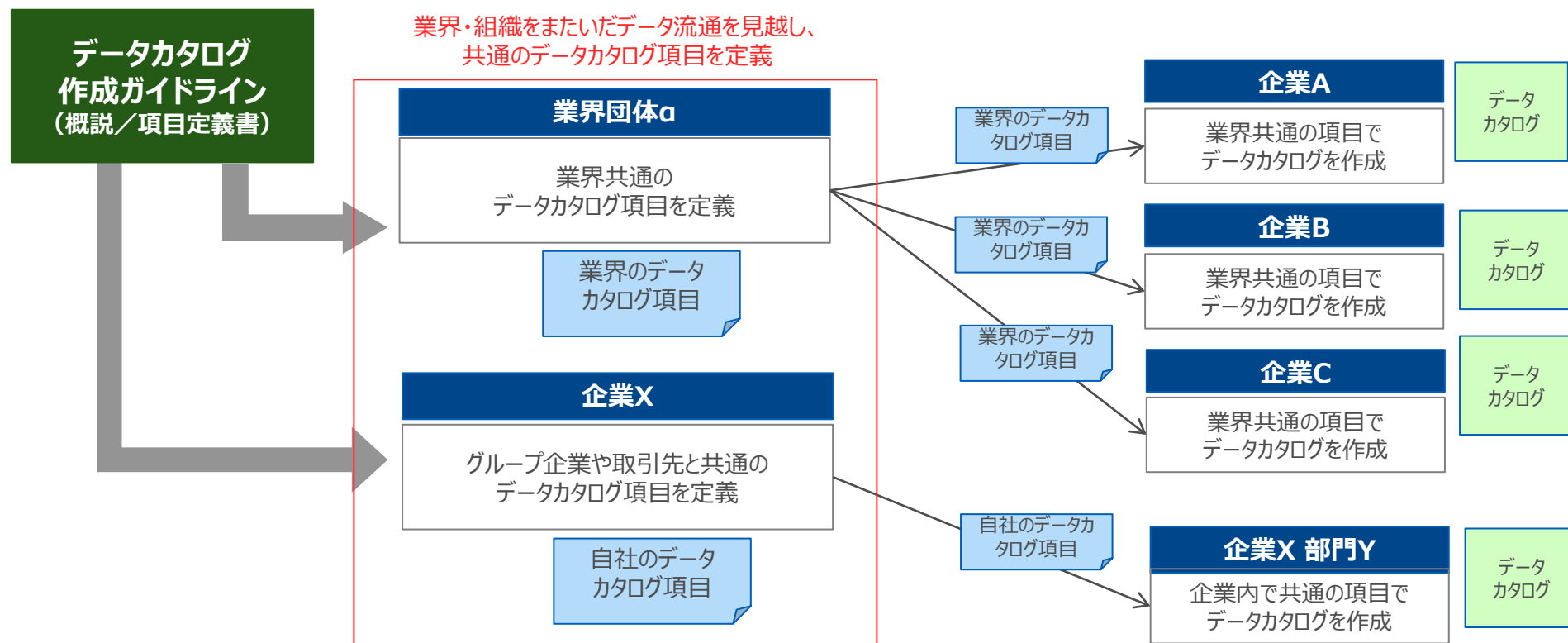
近い将来、「『業界・分野ごとの共通項目』の共通項目」が整備されれば、より自由なデータ流通が行えるようになる。

例：「共通項目の共通項目」があれば、業界・分野ごとにばらつきがあっても、業界・分野をまたいだデータ流通を行いたいとき、「最低限、何が共通するのか」だけは分かる。



データカタログ作成ガイドラインV1.1は、こうした潮流を踏まえて作成。 業界・分野をこえたデータ流通の推進を目指す。

(※業界ごとのルールに修正を求めるものではありません。)



データカタログ作成ガイドラインV1.1は、DTA会員内外の要望を取り入れ、動的データの表現や、権利処理への対応を目指した。

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

特徴（２）：静的データ／動的データへ対応

特徴（３）：データ取引等への対応

動的データの表現や、権利処理への対応を目指した。

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。

データカタログ作成ガイドラインV1.1のデータカタログの構造は、4つの大構造部からなる。

それぞれ、先行する国際標準やガイドラインを参考にしている。

①データカタログ本体部

- ✓ 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

②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®^{注1}の一部をサマリとして利用

③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替え
- ✓ 動的データでは、国際標準化団体W3C SDWWGの概念・用語を参考

④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）を参考

注1 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（T6088442））です。

イメージは本図の通り。構造部ごとに項目と記入例を定義。

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

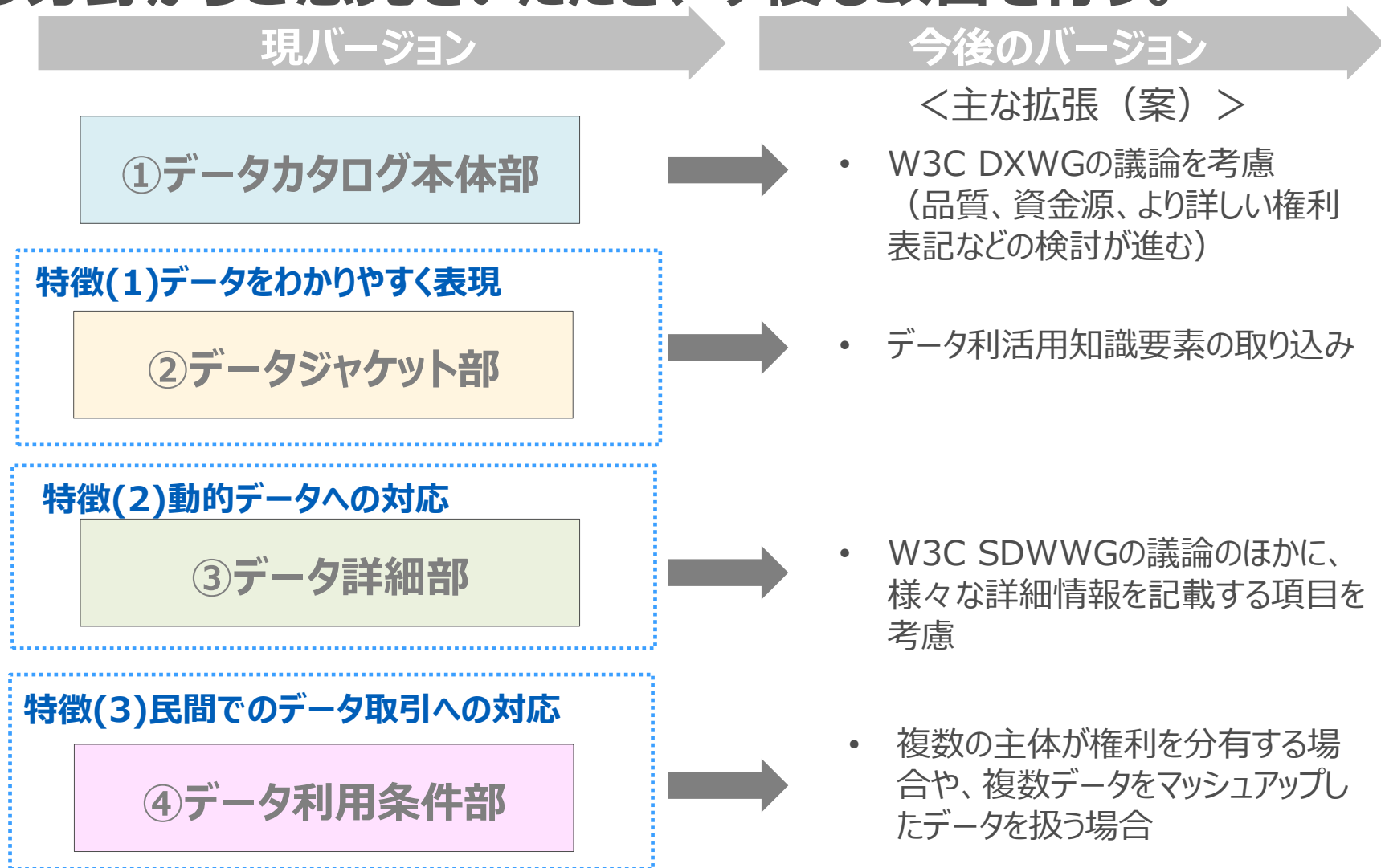
データ項目		データ例
観測活動		
	観測の名称	高齢者を対象とした生活空間での血圧測定
センサ		
	センサの名称	圧力センサ P-XXX-1234
観測対象		
	観測対象の名称	健康サークル血圧測定対象者グループ
観測特性		
	名称	最低血圧、最高血圧
	単位	qudt-unit-1-1:MillimeterOfMercury (mmHg)
観測プラットフォーム		
	名称	電子血圧計 YYY-BP000

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

データカタログ作成ガイドラインV1.1は、Informativeな著作物で、拘束性を持たない。

多くの分野からご意見をいただき、今後も改善を行う。



3. データカタログ作成ガイドラインV1.1 の解説

構造・項目の関係を概念図として整理すると、本図の通り。

②データジャケット部

国内外で利用が進む、
データジャケット®^{注1}の
一部をサマリとして利用

データジャケット

データのタイトル
データの概要説明
データの収集方法・コスト
データの共有条件
データのライセンス
データの変数（属性・パラメータ）の名前
データの分析方法
データに関する補足情報 等

③データ詳細部

データ詳細
(抽象クラス)

観測活動

観測活動の名称
観測活動の説明
観測活動の期間 等

国際標準化団体W3C SDWWGの
議論を参考

・センシングデータの概念体系を参照しているため、概要の記入が難しい場合は、追加項目（サブセット）の設定を行ってもよい。

観測対象

観測対象の識別子
観測対象の名称
観測対象の説明 等

観測特性

観測特性の識別子
観測特性の名称
観測特性の説明 等

観測プラットフォーム

観測プラットフォームの名称
観測プラットフォームの説明
観測プラットフォームの関連文書

センサ

センサの識別子
センサの名称
センサの説明
センサの関連文書
センサで観測する特性等

①データカタログ本体部

国際標準化団体W3C DXWGの
議論を参考

カタログ

カタログのタイトル
カタログの説明
カタログの提供者 等

カタログレコード

カタログのレコードの名称
カタログのレコードの説明
カタログのレコードの作成日 等

データセット

データセットのタイトル
データセットの説明
データセットの発行日
データセットの提供者
データセットの作成者
データセットの提供頻度
データセットの対象地域 等

配信

データセットの配信形式の名称
配信形式の説明
配信開始日
データセット配信を受けるためのライセンス
データセットの配信を行うダウンロードURL
データセットを配信するメディアタイプ
配信データセットの利用条件へのリンク
配信するデータセットのバイトサイズ 等

④データ利用条件部

経済産業省「AI・データの利用に関する契約ガイドライン」（2018）を参考

データ利用条件

・下記のうち、記入できる項目を選択して記載する。未定事項は空欄としてよい。

契約ポリシー

利用期間

利用条件

価格・支払い

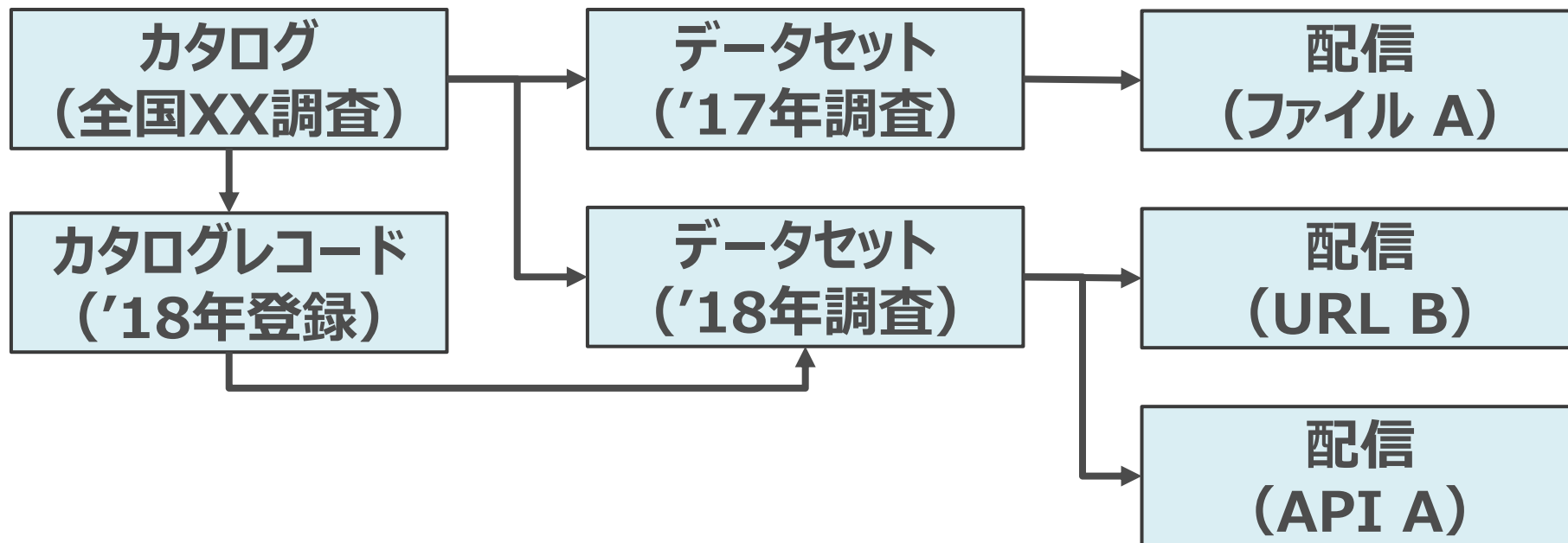
データ保護要件

保証

注1 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（T6088442））です。

①データカタログ本体部は、
国際標準化団体W3C DXWGの議論を参考。
まず、「カタログ」を作り、「データセット」「配信」をぶら下げる構造。

図：データカタログ本体部の項目群の関係イメージ



- ・ カタログ：複数のデータセットをまとめる。カタログの概要や提供者などを示す。
- ・ カatalogレコード：カタログをいつ更新したか記録する。
- ・ データセット：データの概要や提供者、地理・時間区分、分類などを示す。
- ・ 配信：ファイル形式やURLなど、データ入手場所を示す。

②データジャケット部は、 国内外で利用の進むデータジャケット®^{注1}を参照。

- データ分析の必要事項を、文章で自由に記入してよい（全16項目）。

データのタイトル
データの概要説明
データの作成者
データの提供者
データの収集方法・コスト
データの共有条件
データのライセンス
データの種類
データの保存形式

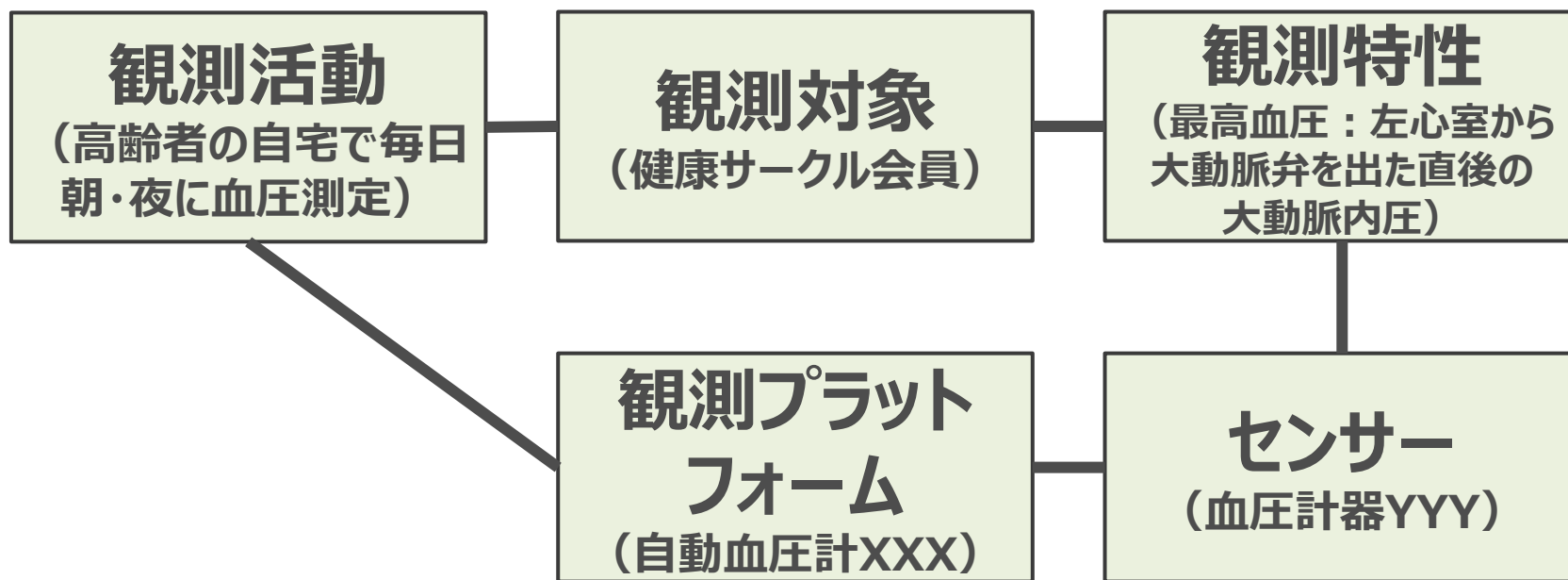
データの変数（属性・パラメータ）の名前
ソース
地域
データの分析方法
分析の結果
その他に期待する分析方法と結果
データに関する補足情報

注1 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（T6088442））です。

③データ詳細部は、 国際標準化団体W3C SDWWGの議論を参考

- ・ 動的データの取得が行われる際の5W1Hを表現する狙い。
- ・ 観測活動の対象は何か、測定する数値（特性）は何か、こういった事物（プラットフォーム）に設置されたセンサーを用いたのか等を記入する。

図：データ詳細部の項目群の関係イメージ



④データ利用条件部は、経済産業省による「AI・データに関する契約ガイドライン」を参照。

- ・ 同ガイドラインの契約書サンプル条文や、本文中の論点を取り入れ、再構成。
- ・ 契約締結前に、交渉条件を簡潔に示す狙い（詳細は契約時に詰める必要あり）

契約ポリシー	契約形態、秘密保持義務、利用用途
利用条件	利用状況の把握、開示範囲、第三者への開示要件、データ購入者の制限、データ活用地域、注意事項、派生データの利用権
データ保護要件	パーソナルデータの類別、データ利用者に求める資格、データの管理
利用期間	データ内容の有効期間、利用ライセンスの期限
価格・支払い	支払類型、課金単位、価格帯、データ販売に関わる特記事項、支払頻度
保証	保証、準拠法の対象国

大構造部ごとに、項目定義を行ったイメージは次の通り。

データカタログ作成ガイドラインV1.1 項目定義書のサンプルイメージ

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
①データ カタログ 本体部	カタログ の提供者	1..1	文字列型 (xsd:string)	<説明> このカタログを提供する組織・機関 の名称です。 <入力ルール> このカタログを提供する正式な組 織・機関の名称を記入してください。 個人でデータカタログを提供する場 合には、その権利・責任を持つ方が、 個人名を記入してください。	〇〇株式会社 データ サービス事業部

個々の項目は次のように定義されている。

項目No	項目の通番
大構造	①データカタログ本体部, ②データジャケット部, ③データ詳細部, ④データ利用条件部の区分を示す
見出し	項目名を日本語で表記したもの
出現回数	各項目が必須か任意か、出現させられる回数はいくつかを示す -0..1 : 出現させなくてもよいが、1回しか出現してはならない -1..1 : 1回のみ、必ず出現させなければならない -0..n : 何回でも出現してよい、出現させなくてもよい -1..n : 少なくとも1回以上、必ず出現させなければならない
値域 (データタイプ)	各項目を入力するとき、その値が文字なのか、日付なのか、数値なのか等を規定。W3C XML Schemaのデータタイプを踏襲
説明	各項目の説明と、値を入力する際のルールを解説
サンプル値	入力ルールに基づいた記入例

4. 分野ごとの データカタログ作成手順

【要旨】データカタログ作成ガイドラインV1.1活用に向けた5ステップ

STEP 1

採用する大構造部を選ぶ

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

STEP 3

入力ルールを確認する

STEP 4

サンプルデータを用いた記入テスト

STEP 5

作成の手引きや修正提案の作成

データカタログ作成ガイドラインV1.1活用に向けた5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。
 - ✓ 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

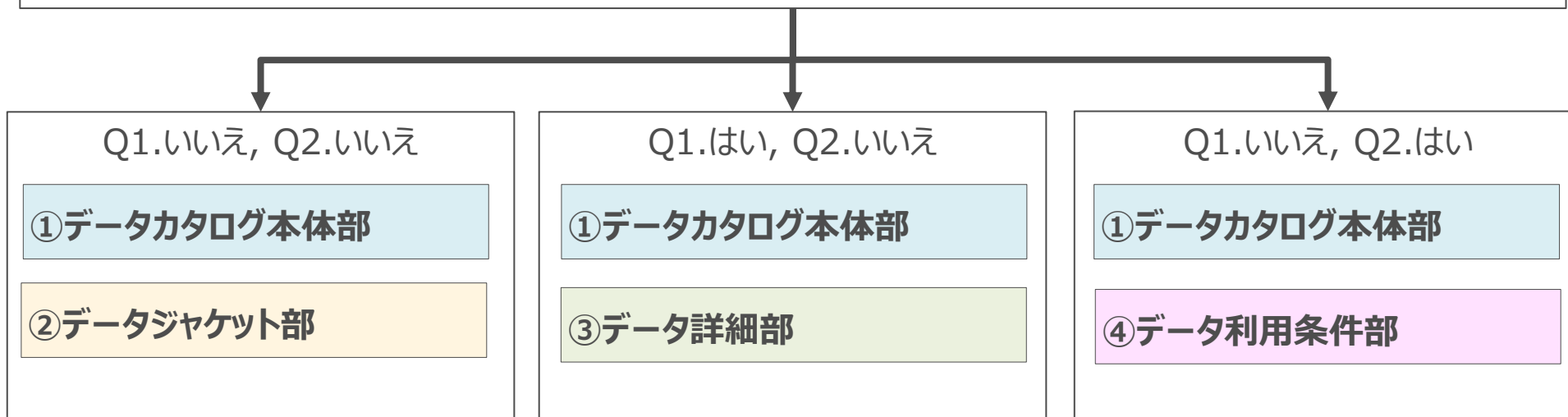
- 必要に応じて、実務者向けに、業界・業種ごとの「作成の手引き」を作成して下さい。

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？



STEP 2

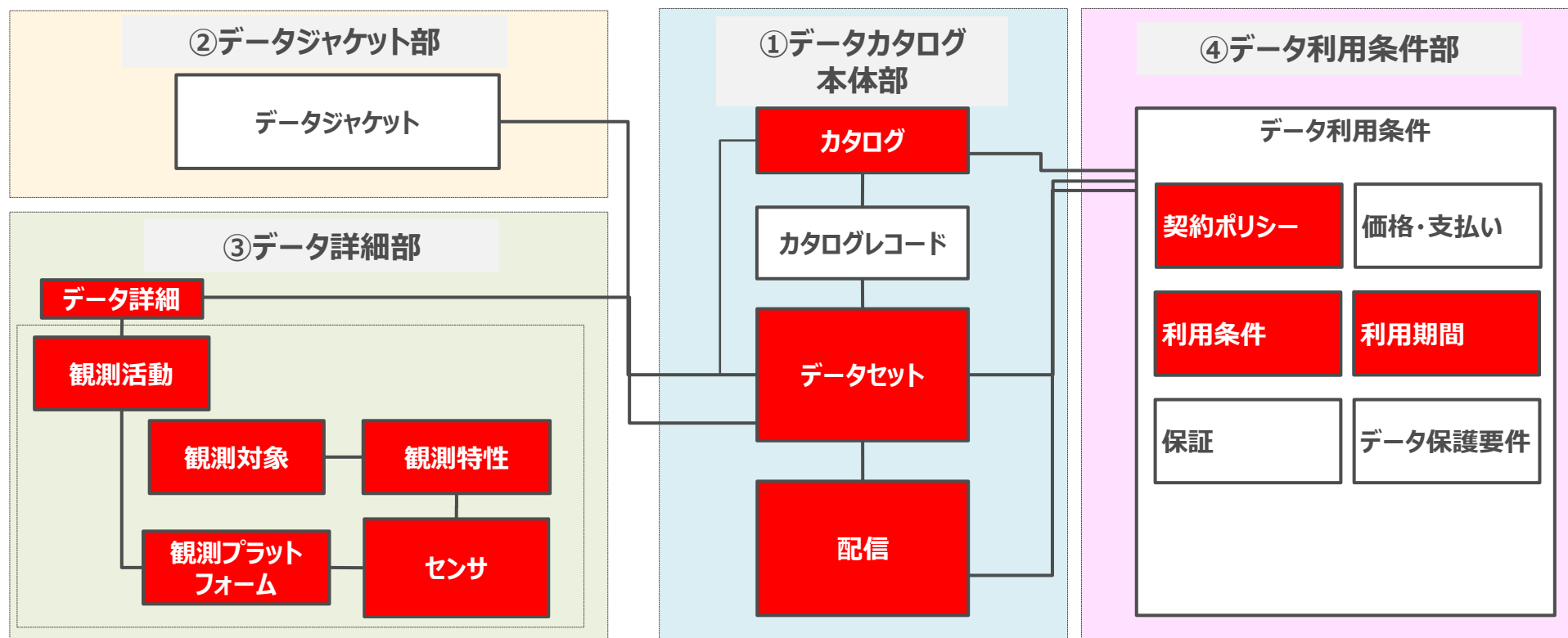
採用するデータ項目を選び、 不足する項目を見極める

- 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



STEP 3

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	<説明> どの用途であれば、利用を認めるかを記載してください、 <入力ルール> 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	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

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て、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。

