

「IT標準化支援プログラム」勉強会 共通語彙篇

2018年10月4日

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（SII）より、一般社団法人データ流通推進協議会（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。
- 本資料に掲載しております「ドメイン語彙の策定（作業概要や作業ステップ）」は、本事業とは別に、DTAが自主事業として検討しているものとなります。

目次

1. そもそも「語彙」とは何か？
2. 「語彙」を揃える意義
3. IMI（共通語彙基盤）の概要
4. ドメイン語彙の策定（作業概要）
5. ドメイン語彙の策定 作業ステップ

1. そもそも「語彙」とは何か？

「語彙」とは、「語」を「集めたもの」のことをいう。
一語一語は、「語」「用語」「単語」などと呼ぶ。

- 「語」とは

- ことば。また、ことばづかい。

- 「彙」とは

- たぐい。なかま。同類のもの
- あつめる。また、あつまる。あつまり。

出典：大修館書店『新漢語林』第二版

日本語には「雨」を表現する多くの用語があり、それらをまとめて「雨の語彙」と呼ぶことができる。

●「雨」の語彙

●秋雨, 朝雨, 陰雨, 雲雨, 液雨, 煙雨, 大雨, 快雨, 甘雨, 寒雨, 喜雨, 急雨, 強雨, 暁雨, 強風雨, 霧雨, 軽雨, 紅雨, 降雨, 豪雨, 黒雨, 小雨, 細雨, 山雨, 地雨, 糸雨, 慈雨, 屢雨, 吹雨, 驟雨…etc.

天気予報では、空模様を誤解なく伝えるため、多くの「用語」を分類し、「語彙」を定義している。

用語 (※)	区分	説明
天気		気温、湿度、風、雲量、視程、雨、雪、雷などの気象に関する要素を総合した大気の状態。
	備考	気象庁では国内用として、次の15種類に分けているが、国際的には96種類が決められている。 快晴、晴れ、薄曇り、曇り、煙霧、砂じん嵐、地ふぶき、霧、霧雨、雨、みぞれ、雪、あられ、ひょう、雷。
変わりやすい天気		対象とする予報期間の中で、晴れが続かず、すぐに曇ったり雨（雪）が降ったりする天気。
	備考	週間天気予報等では天気が比較的短周期（2日程度）に変わると予想されるときに用いる。

出典：気象庁「天気予報等で用いる用語」 (https://www.jma.go.jp/jma/kishou/known/yougo_hp/tenki.html)

(※) 用語とは、書いたり話したりするときに用いられる語や句。特定の分野で特に使われる語句。
(引用：Wikipedia)

**【参考】語彙や用語について、詳細を知りたい方は、
次のようなキーワードを調べてみてください。**

✓ **この資料で主として扱うキーワード**

- 語彙 (Vocabulary) の用語 (Term)

✓ **明示的には扱わないキーワード**

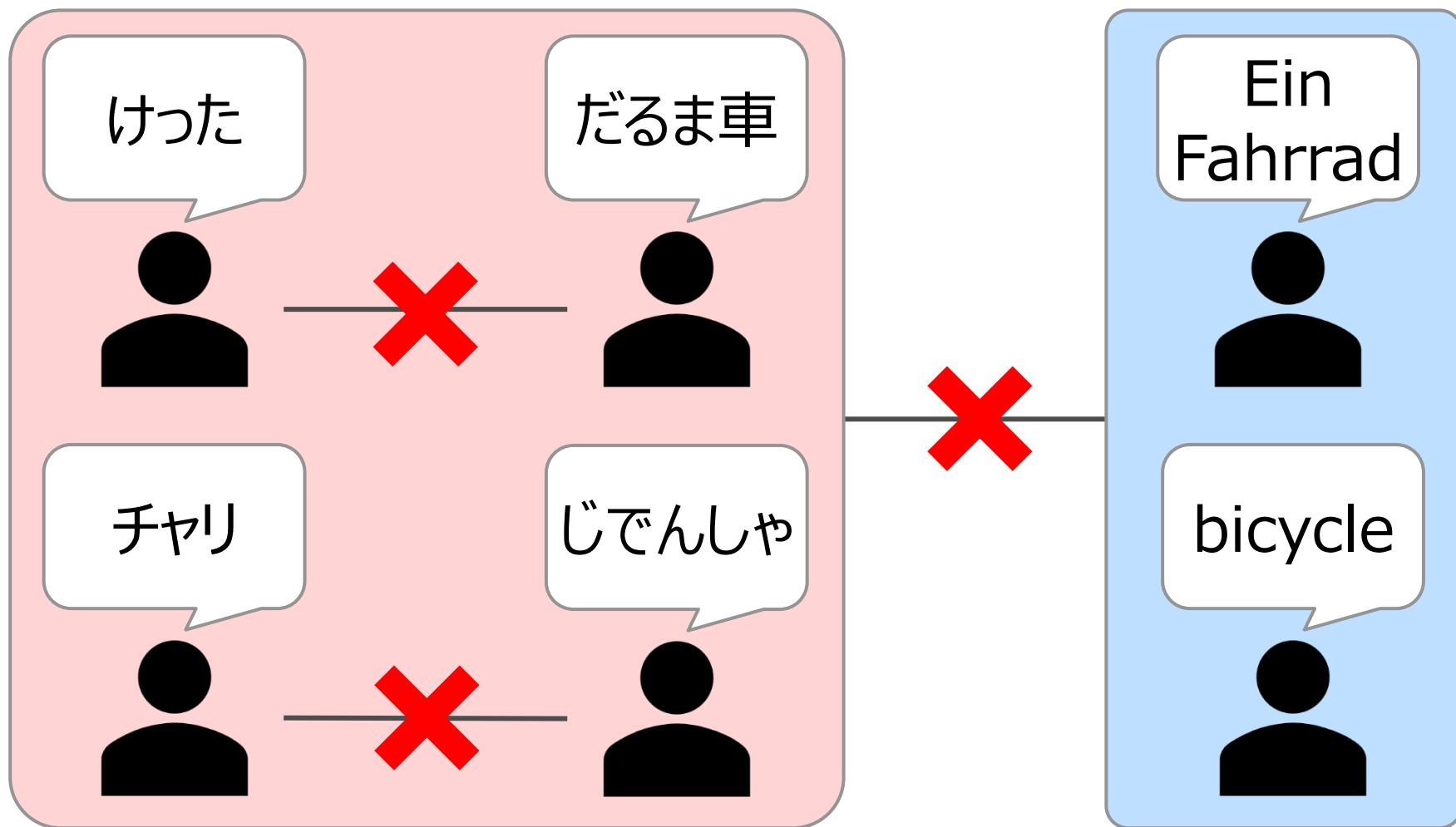
- 統制語彙 (Controlled Vocabulary)
- 用語集 (Term List)
- シソーラス (類義語録, Thesaurus)
- タクソノミー (分類法, Taxonomy)
- オントロジー (概念構造, Ontology)
- 典拠管理 (Authority Control)
- 件名標目表 (Subject Headings)
- 正規表現 (Regular Expression)
- セマンティクス、シンタクス (Semantics, Syntax)

2. 「語彙」を揃える意義

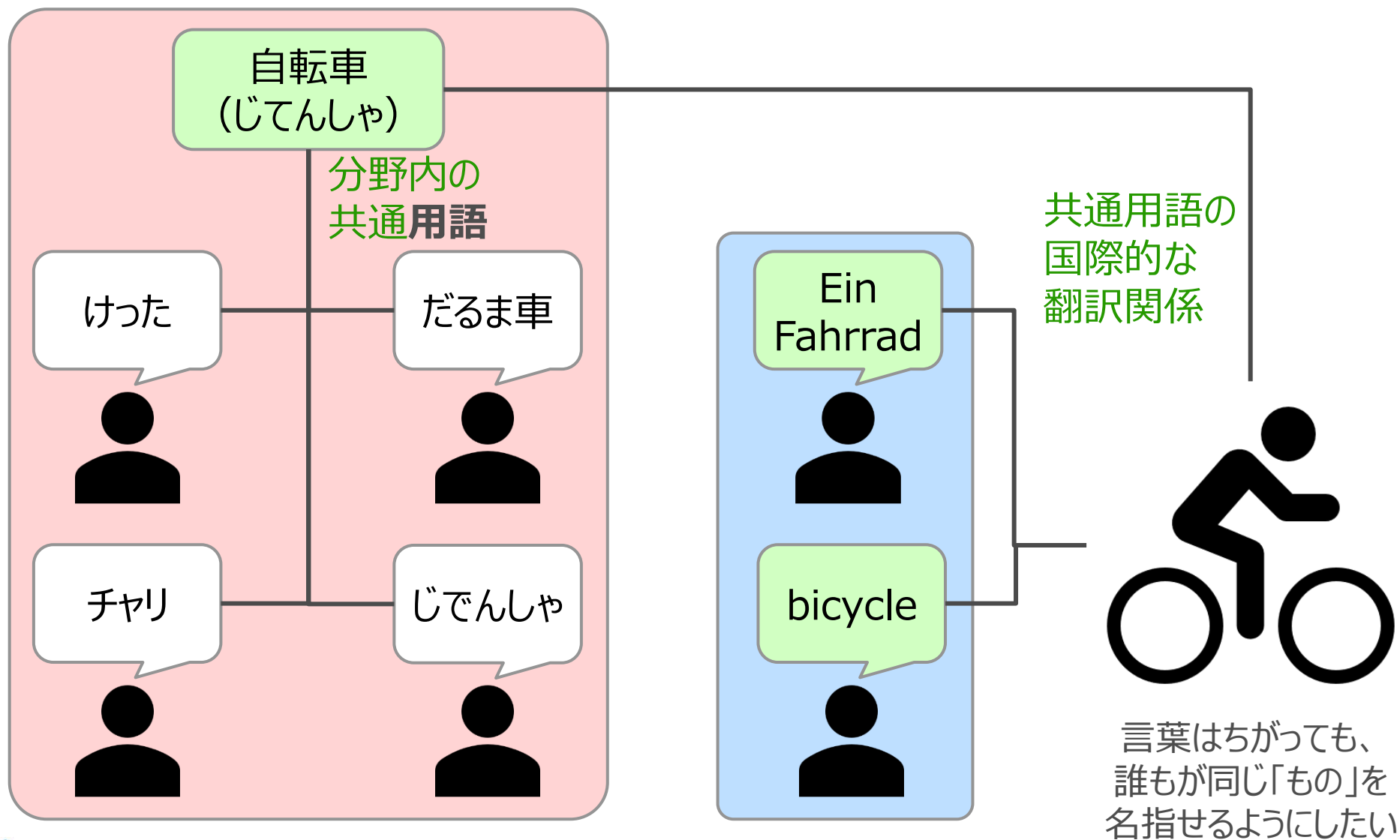
方言や外国語など、異なる「語彙」の「用語」を使っていると、「同じもの」なのに、何を意味するのか通じないことがある。

同じ分野内でも、「異なる用語」は通じない

分野外はもっと通じない



誰もが「どの語も、同じものを指す」と理解したうえで、異なる「用語」を代表する「共通用語」があれば、意味が通じる。



第三者にデータを伝えるときには、さらに注意を要する。
分類、単位、表記を揃えなければ、誤った情報になりかねない。

例：「一人当たりの水消費量」とは？

用途

- 飲料水
- 生活用水
- 農業用水
- 工業用水

分類

単位

- CC
- l
- dl
- ml

単位

分類

種類

- 水道水
- ミネラルウォーター
- 精製水

表記

数値の表記

- 2.6
- 四百
- 1,000

コンピュータは「背景や状況を察してデータを読む」ことができない。
用語の中には、対象を「識別」する決まりをもつものがある。

例：コンピュータにも理解できるように、「その国だけが持つ」、番号、略称、表記を指定

国の識別

国・地域名

日本

英語名

Japan

3字コード

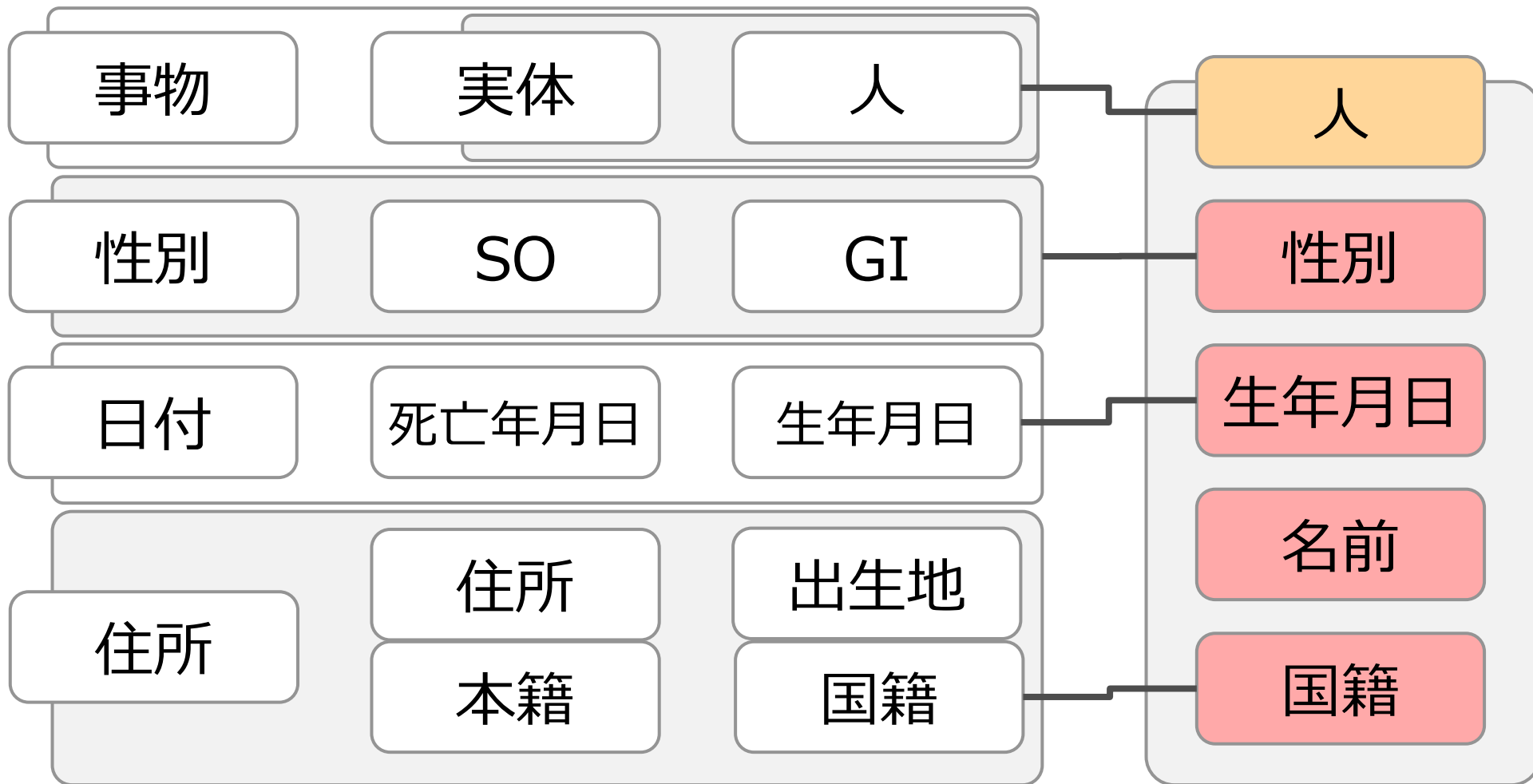
JPN

2字コード

JP

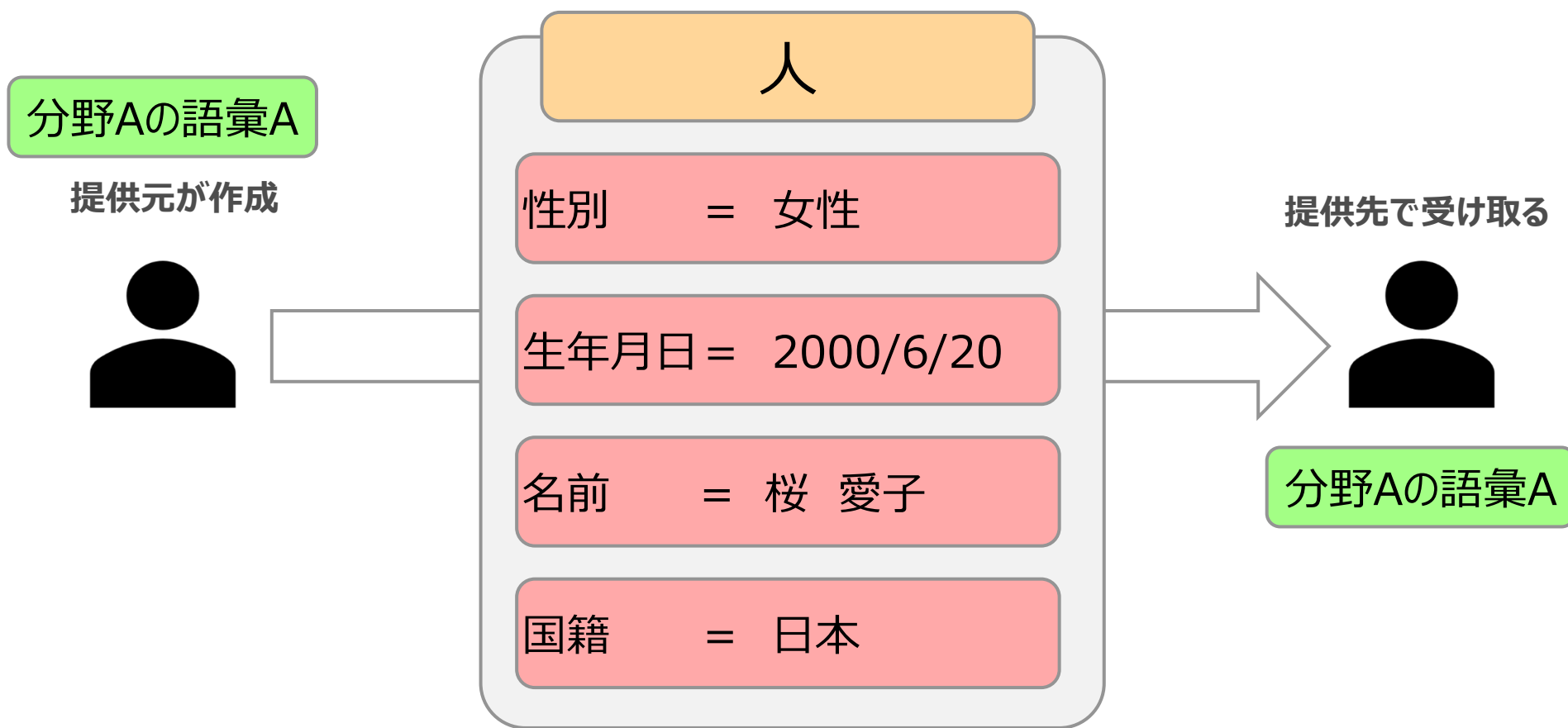
一語一語を「識別」するだけでは足りない。 複数の用語を「分類」し、「関係」づける「構造」が必要。

例：より正確な「自己紹介」のためには、「複数の用語」の「正しい組み合わせ」が必要



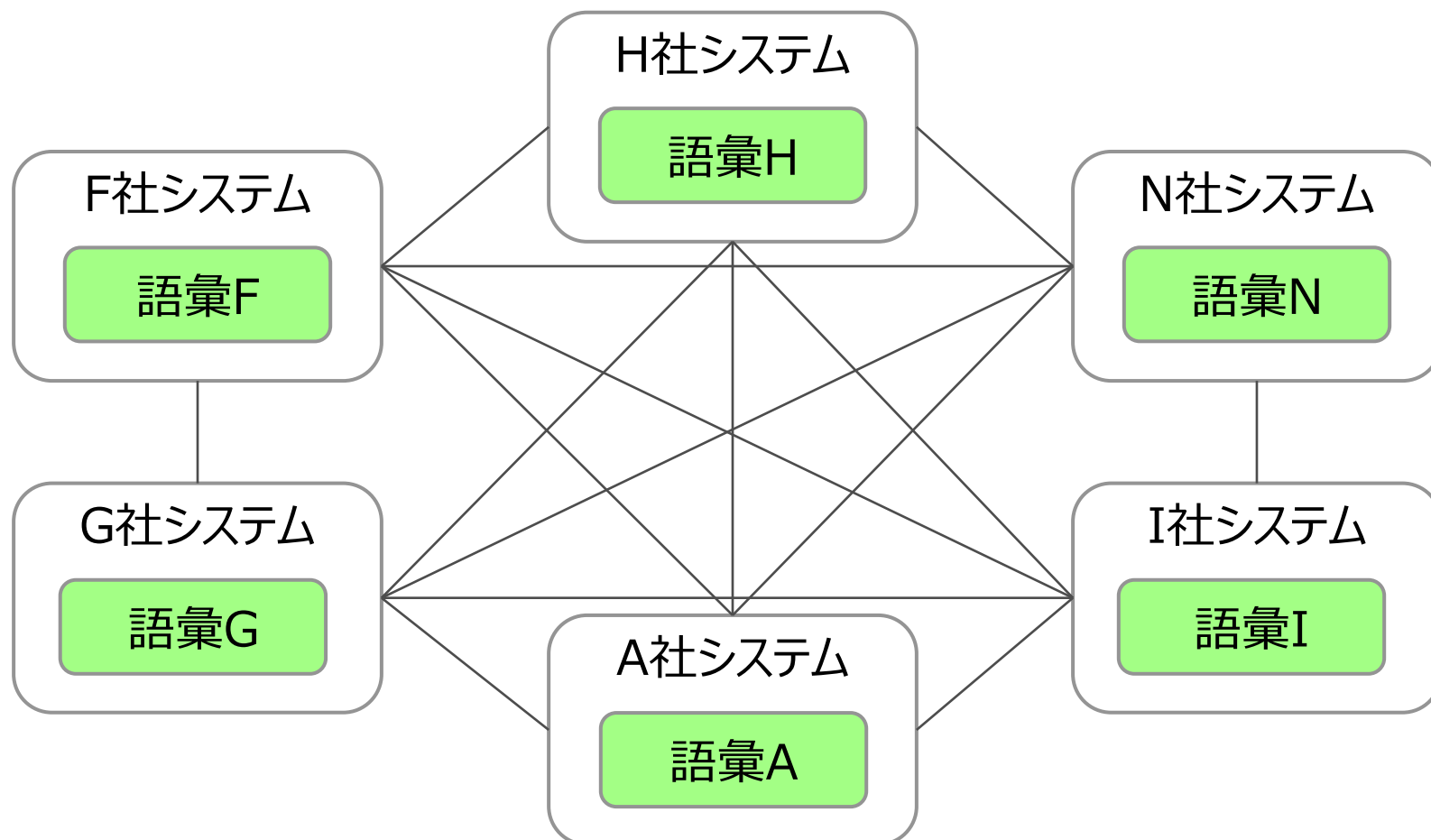
コンピュータシステム間でデータを交換するときは、
どの「語彙」に属する「用語」がどんな「値」を持つのかを
受け渡さなければ、正しい情報を相手に伝えられない。

語彙（人）の用語（国籍）と値（日本）を一緒に渡さなければ、伝わらない。



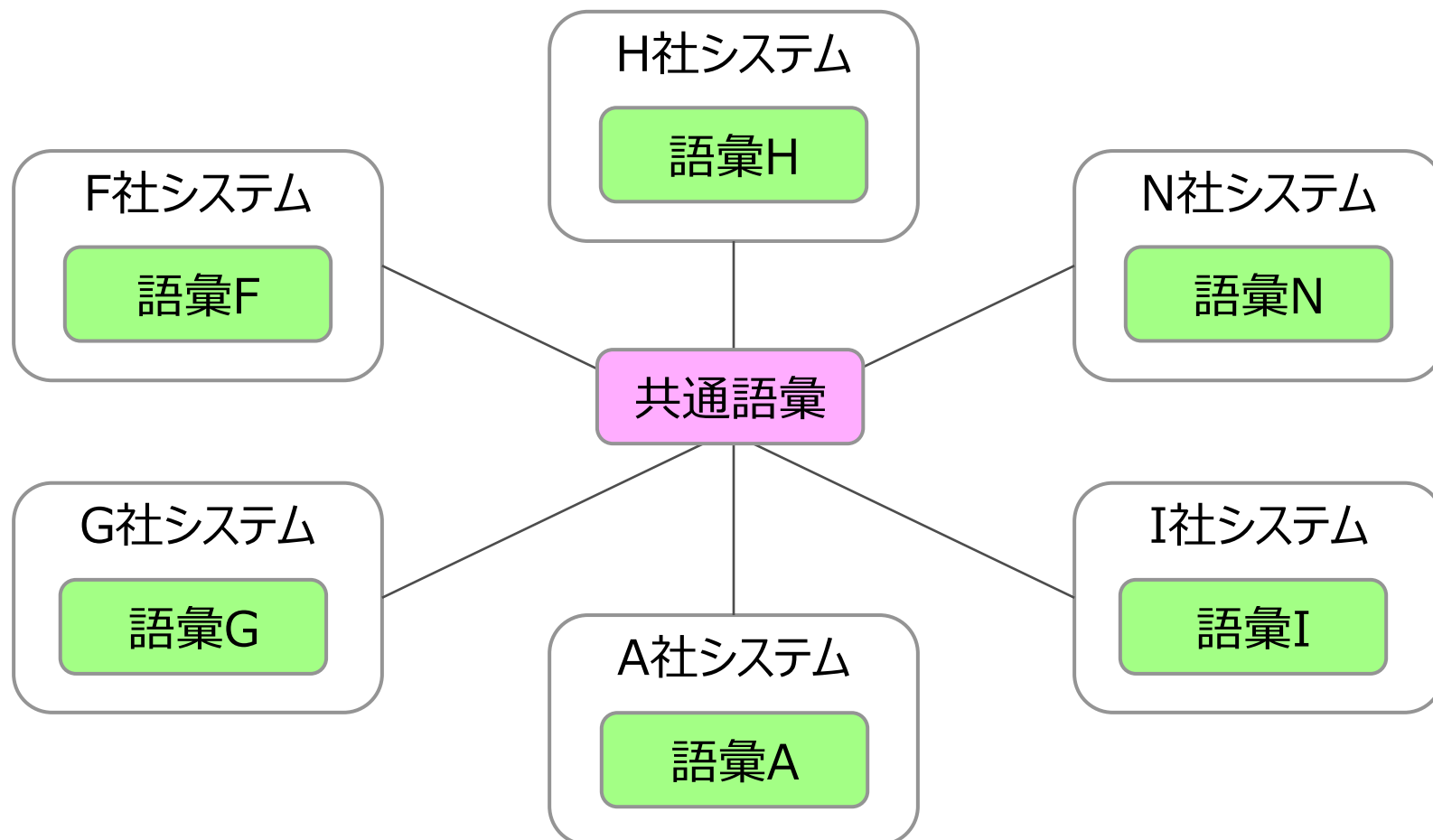
ITにおける語彙も、分野やシステムごとに異なることが多い。
語彙の異なるシステム間で、その度ごとに個別調整していると、
システムが増えれば増えるほど、膨大な「調整」が生じて非効率。

例：6社のシステムの語彙がすべて異なると、なんと15回分もの「個別調整」が必要。大いなる無駄。

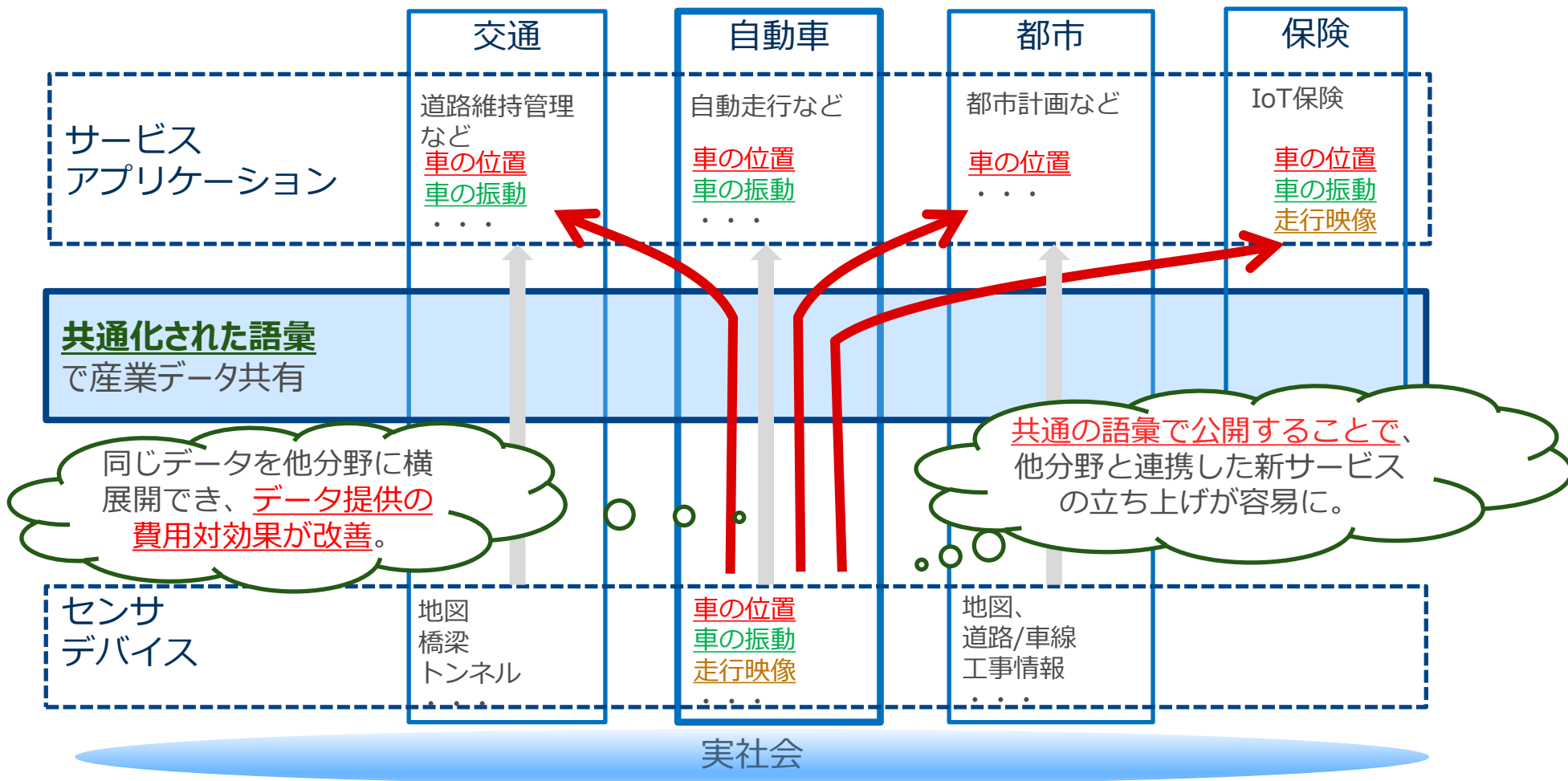


データ交換「語彙」を揃える意義とは？

将来のコストが減ること



今後、分野をまたいだ「語彙」の共通化が進むと、
 多くのデータから、より良質な情報を、より低コストに得られる。
 新サービスの準備や、データ提供の収益化が容易になりうる。



3. IMI (共通語彙基盤) の概要

IPA IMI（共通語彙基盤）とは

- 略称：Infrastructure for Multi-layer Interoperability
- 推進委員会は経済産業省が設置し、IPAが事務局
- 内閣府「Society5.0実現に向けたデータ連携基盤」でも、他の標準化政策と合わせて重点化
- 政府の基本戦略「未来投資戦略」「統合イノベーション戦略」の具体的政策の実現手段としてとりあげられている

IMIは、次の3種類の語彙に区分している

コア語彙

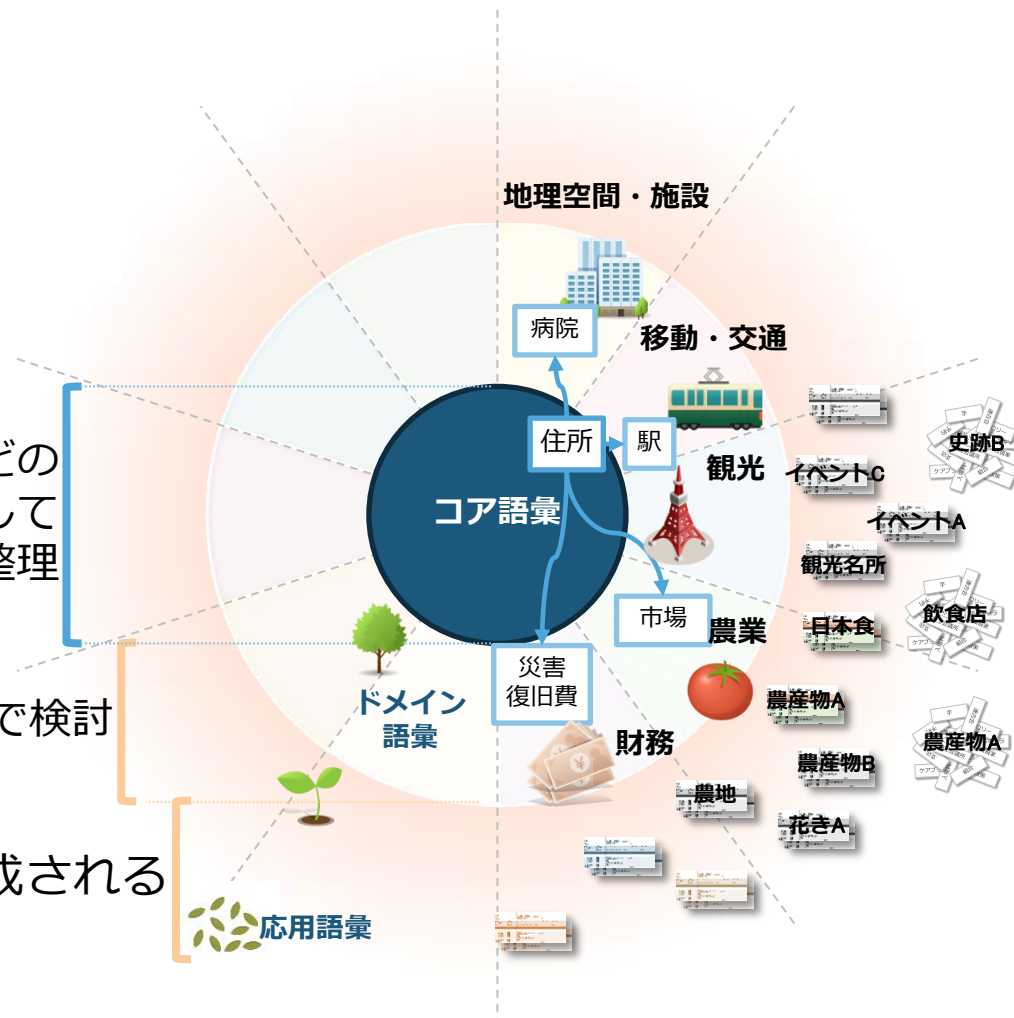
「人」「場所」などの
基本的な概念を用語として
IMI検討体制が整理

ドメイン語彙

目的に応じて各ドメインで検討

応用語彙

目的に応じて構成される




出典：IMI検討部会事務局 (IPA)「データ流通推進協議会向け勉強会資料 (2018年8月8日)」(修正版)

IMIは、次の3種類の語彙に区分している

- **コア語彙**：どの分野でも普遍的に使われる
 - 日付、場所、人物、組織など
 - IMIが語彙定義を公表。利用すると便利。
- **ドメイン語彙**：その分野なら共通して使える
 - 防災、医療、財務、観光、交通などの語彙
 - コア語彙を継承して作成
 - 他のドメイン（分野）にも通じることが望ましい
 - IMIパートナーになれば、コア語彙適用の助言を受けられる
- **応用語彙**：目的と場合により作られる
 - 部署・部門ごとに使い慣れた語彙や構造

IMIから提供するコア語彙。 個々の用語を分類する「クラス（型）」が定義される。



情報連携に不可欠な基本情報やツールを提供するサイトです

English

| 共通語彙基盤 コア語彙2.4.1

To the extent possible under law, [Ministry of Economy, Trade and Industry](#) and [Information-technology Promotion Agency, Japan](#) has waived all copyright and related or neighboring rights to Infrastructure for Multilayer Interoperability (IMI) Core Vocabulary 2.4.1. This work is published from: Japan .

赤い取り消し線が引かれた用語およびプロパティは将来の版では使用できなくなるため、使用を推奨しません。

=> [プロパティ用語一覧](#)

| クラス用語一覧

<ul style="list-style-type: none"> 人型 連絡先型 法人型 場所型 製品型 施設型 設備型 <li style="color: red;">測定値型 重量型 価格型 日付型 定期スケジュール型 単位コード型 概念型 サービス型 コード制約型 	<ul style="list-style-type: none"> 名称型 電話番号型 人数型 座標型 製品個品型 施設関連型 イベント型 数量型 長さ型 実体型 期間型 詳細スケジュール型 コードリスト型 事物型 文書型 範囲制約型 	<ul style="list-style-type: none"> 氏名型 組織型 構成員型 ID型 地物型 建物型 活動型 容量型 <li style="color: red;">時間型 状況型 期間スケジュール型 詳細スケジュール規則型 アクセス型 参照型 対象型 期間制約型 	<ul style="list-style-type: none"> 住所型 業務組織型 組織関連型 ID体系型 土地型 駐車場型 関与型 面積型 金額型 日時型 イベントスケジュール型 コード型 アクセス区間型 記述型 制約型
---	--	---	---

出典：「共通語彙基盤 コア語彙2.4.1」 <https://imi.go.jp/ns/core/Core241.html>

各クラス用語のプロパティ用語を定義（識別子、値型、回数、説明）

（例）「住所型」のプロパティ「ic:国」は、値型は「string（文字）」で「最大1回まで使える（0..1）」。

住所を表現するためのクラス用語

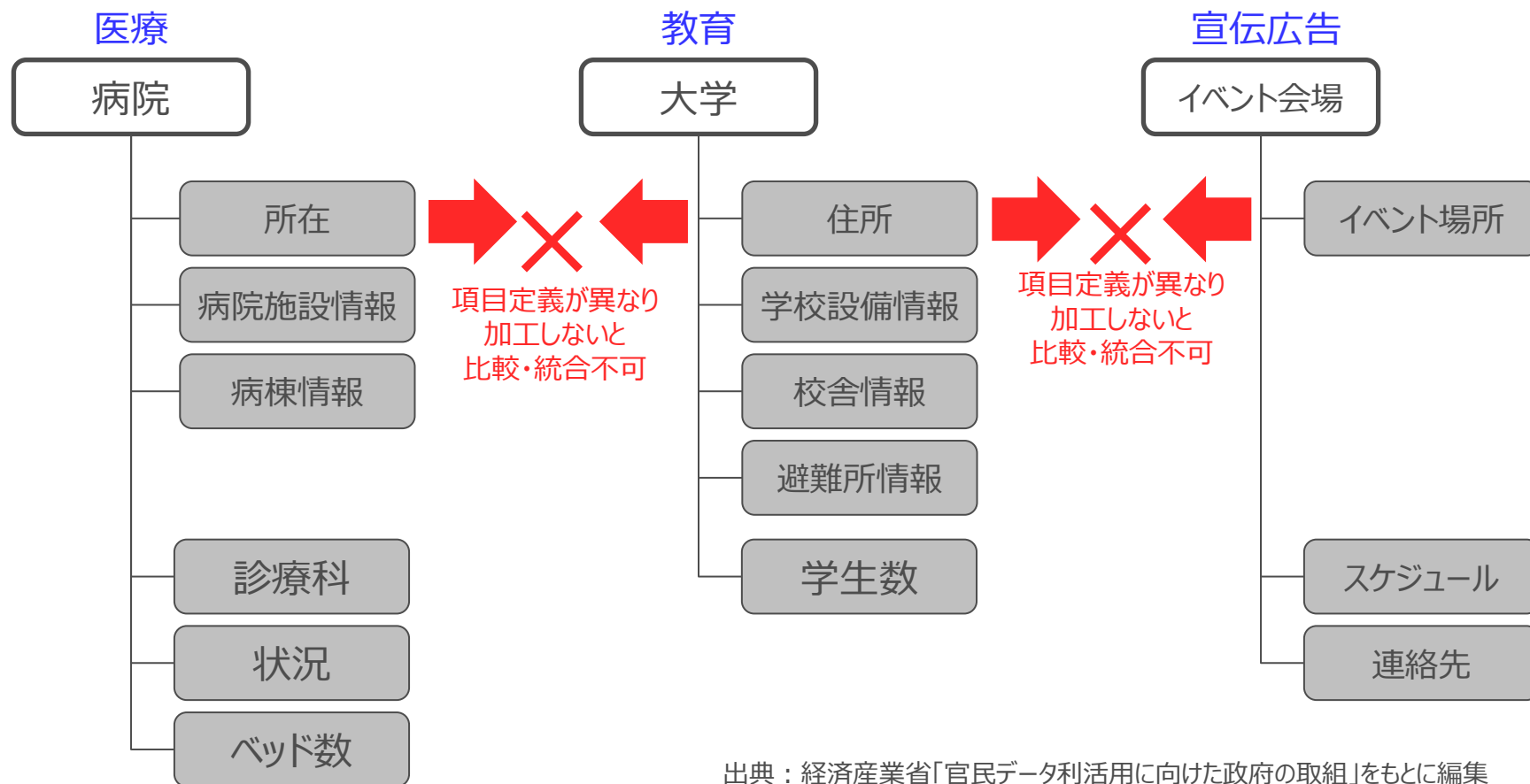
住所型 のプロパティ			
識別子	値型	回数	説明
ic:種別	xsd:string	0..n	種別の表記を記述するためのプロパティ用語
ic:種別コード	ic:コード型	0..n	種別コードを記述するためのプロパティ用語
ic:メタデータ	ic:文書型	0..n	データのメタデータを記述するためのプロパティ用語
ic:ID	ic:ID型	0..n	事物に割り振られたID（識別子）を記述するためのプロパティ用語
ic:表記	xsd:string	0..n	文字列による表記を記述するためのプロパティ用語
ic:参照	ic:参照型	0..n	追加情報などを参照するURLを記述するためのプロパティ用語
ic:画像	xsd:anyURI	0..n	画像を参照するURLを記述するためのプロパティ用語
ic:説明	xsd:string	0..n	説明を記述するためのプロパティ用語
ic:記述	ic:記述型	0..n	種別付きの説明文を記述するためのプロパティ用語
ic:郵便番号	xsd:string	0..1	郵便番号を記述するためのプロパティ用語 ※0-9の数字7桁
ic:住所コード	ic:コード型	0..n	住所のコードを記述するためのプロパティ用語 ※地方公共団体情報システム機構の地方公共団体コード住所など
ic:国	xsd:string	0..1	国名を記述するためのプロパティ用語
ic:国コード	ic:コード型	0..1	国のコードを記述するためのプロパティ用語
ic:都道府県	xsd:string	0..1	都道府県名を記述するためのプロパティ用語
ic:都道府県コード	ic:コード型	0..1	都道府県のコードを記述するためのプロパティ用語
ic:市区町村	xsd:string	0..1	市区町村名を記述するためのプロパティ用語
ic:区	xsd:string	0..1	政令指定都市の区の名前を記述するためのプロパティ用語

出典：「共通語彙基盤 コア語彙2.4.1」 <https://imi.go.jp/ns/core/Core241.html>

Copyright©Sustainable open Innovation Initiative. All Rights Reserved.

同じ組織なのに、同じ事物を、別々の用語で定義していると、項目定義がバラバラになって、データの比較・統合が煩雑に。

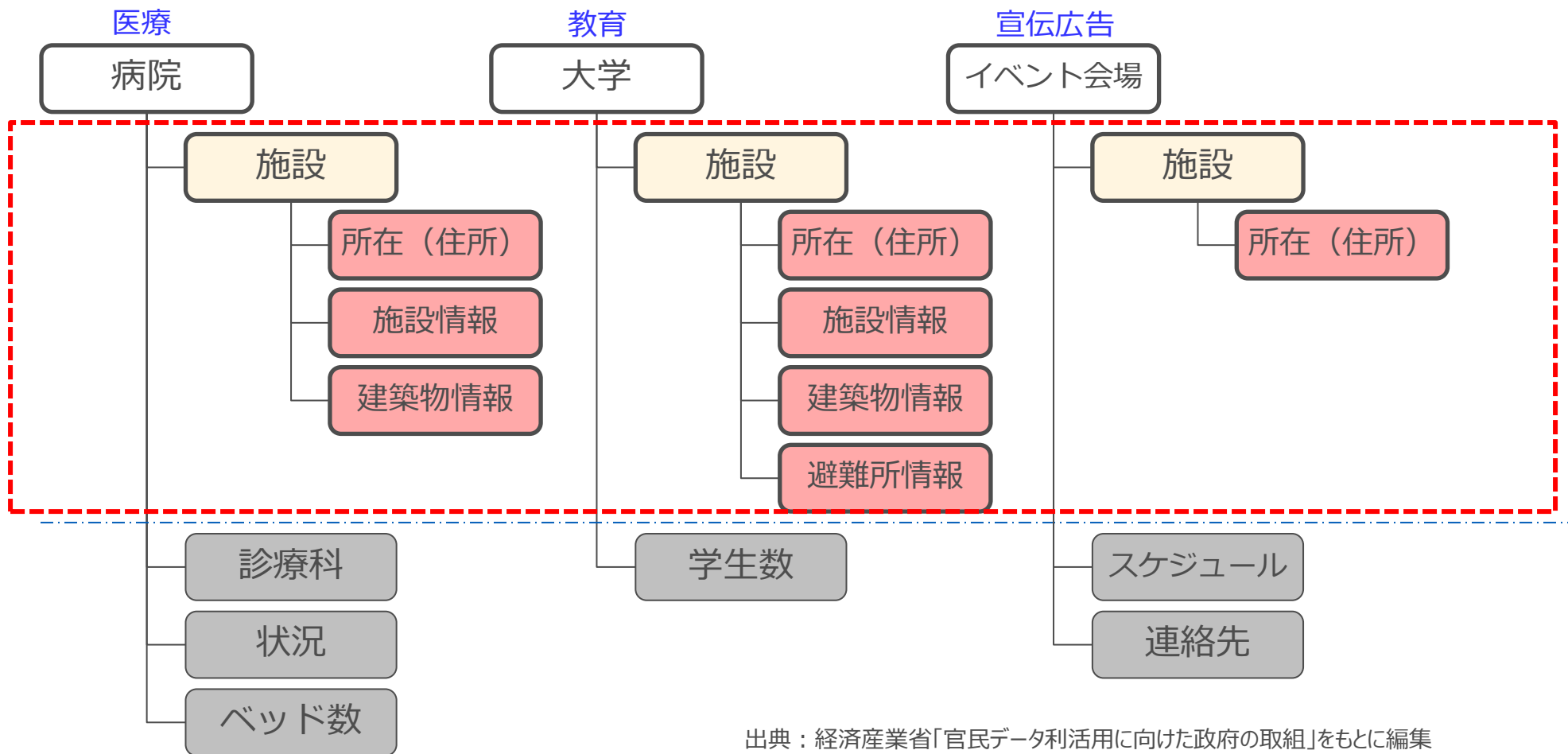
(架空の例) ○○大学付属病院で、医院、大学、PR会社が連携し、分野△△に関するシンポジウムを開催。同じ場所なのに、「所在」「住所」「イベント場所」など、組織ごとに異なる項目定義がなされていて、データが繋がらない。



出典：経済産業省「官民データ活用に向けた政府の取組」をもとに編集
(https://imi.go.jp/doc/20180129_r1_1.pdf)

IMIのコア語彙を適用したり、分野内でドメイン語彙を定義しておけば、異なる組織・システム間で「同じ意味」が扱える。

(架空の例) 関係者がよく話し合い、項目定義を揃えておけば、少なくとも「所在（住所）」はつながられる。



出典：経済産業省「官民データ利活用に向けた政府の取組」をもとに編集
(https://imi.go.jp/doc/20180129_r1_1.pdf)

よく誤解されるが、IMIは「意味」を揃えることを目指すもの。
用語の定義の変更が難しいときは、「対応づけ」のみでもよい。



全て「住所」を表す

共通語彙

クラス用語

プロパティ用語

住所

郵便番号

都道府県

市区町村

番地



「住所」は様々な属性を持つ

A市「所在地」データセット

113-6591

郵便番号

東京都

都道府県

文京区本駒込

市区町村

2-28-2

番地

B市「居住地」データセット

〒1008901

東京都千代田区霞が関一丁目3-1



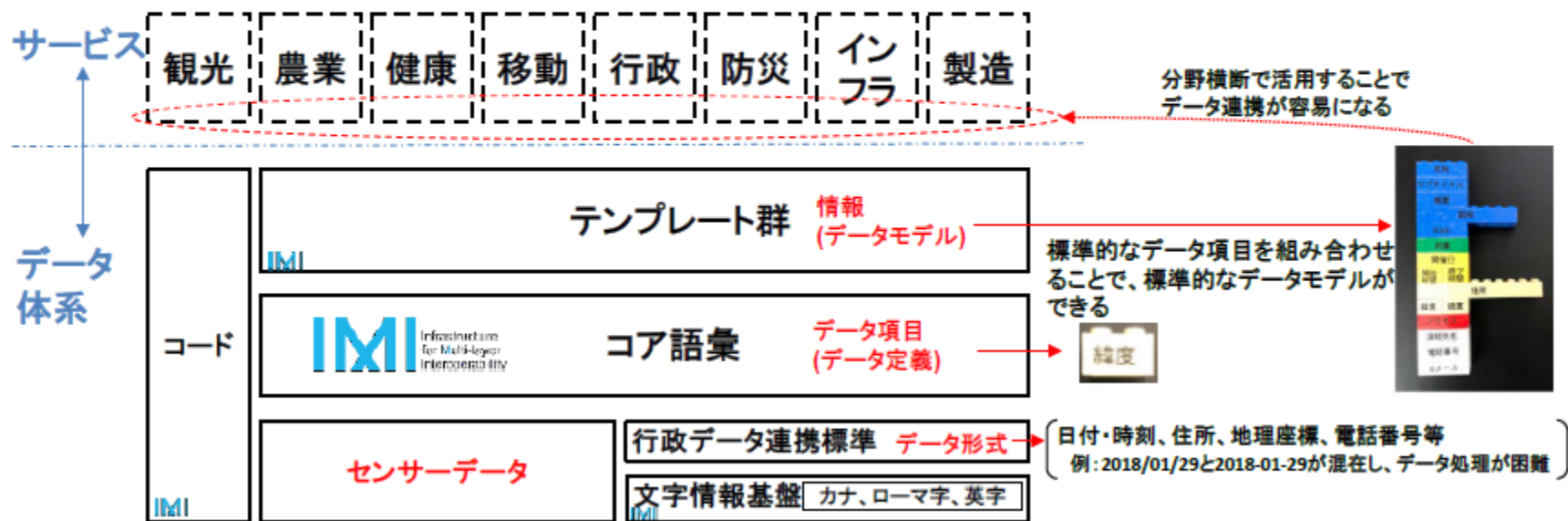
各データ項目の意味を
対応付け

出典：IMI検討部会事務局（IPA）「データ流通推進協議会向け勉強会資料（2018年8月8日）」をもとに作成

IMIでは、コア語彙だけでなく、データモデルやコード、文字情報、実装モデルなどの周辺分野も整備中。

デジタル戦略の中でのデータ体系

- 社会全体でデータ利活用するためには、基本データから積み上げた体系の中で相互運用性を確保していく必要がある。

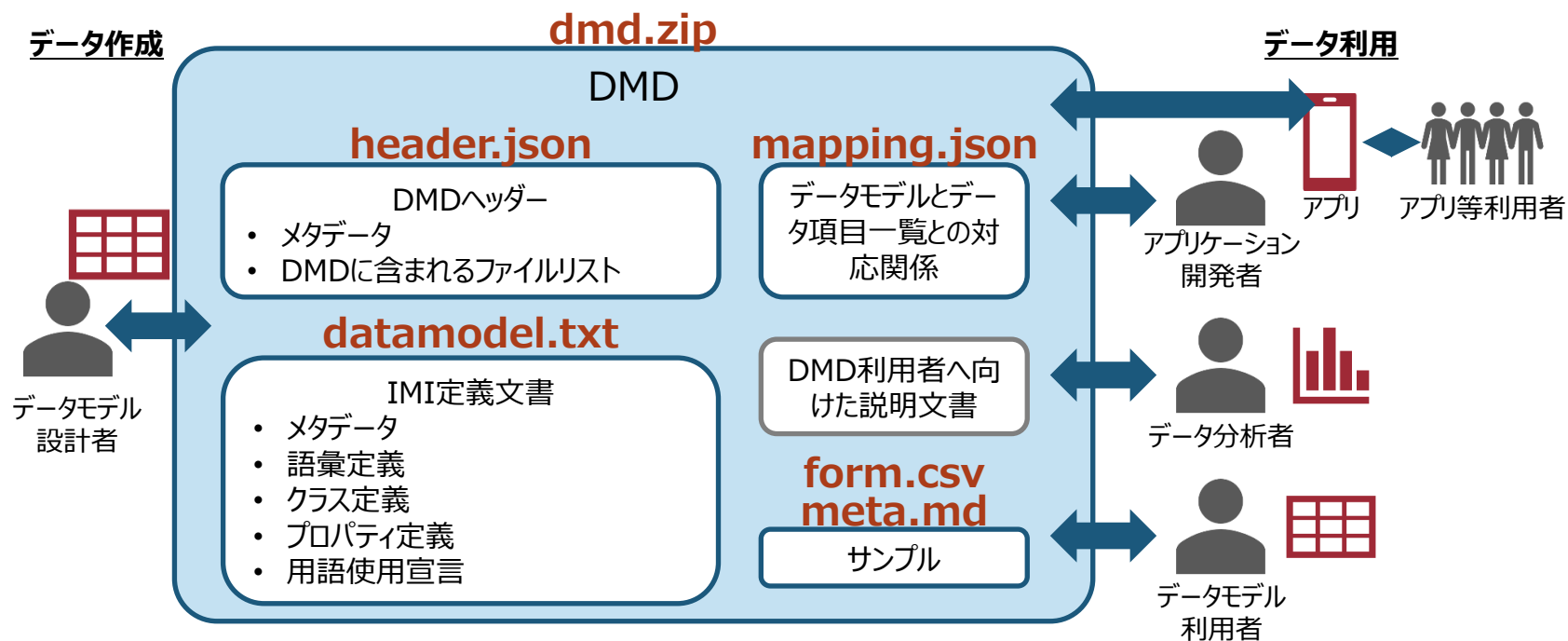


※1 IMI は、デジタルガバメントで推進している共通語彙基盤の略称。国際連携も実施中。

※2 データを扱うためのツール体系、データ品質は確保するための品質体系も整備していく必要がある

出典：内閣官房政府IT総合戦略室「分野間データ連携基盤の整備」（2018/8/3）

ドメイン語彙用のテンプレート群（データモデル）定義のためのDMD（データモデル記述）の仕様と作成支援ツールを提供。 ドメイン語彙の定義を作成するときに利用できる。



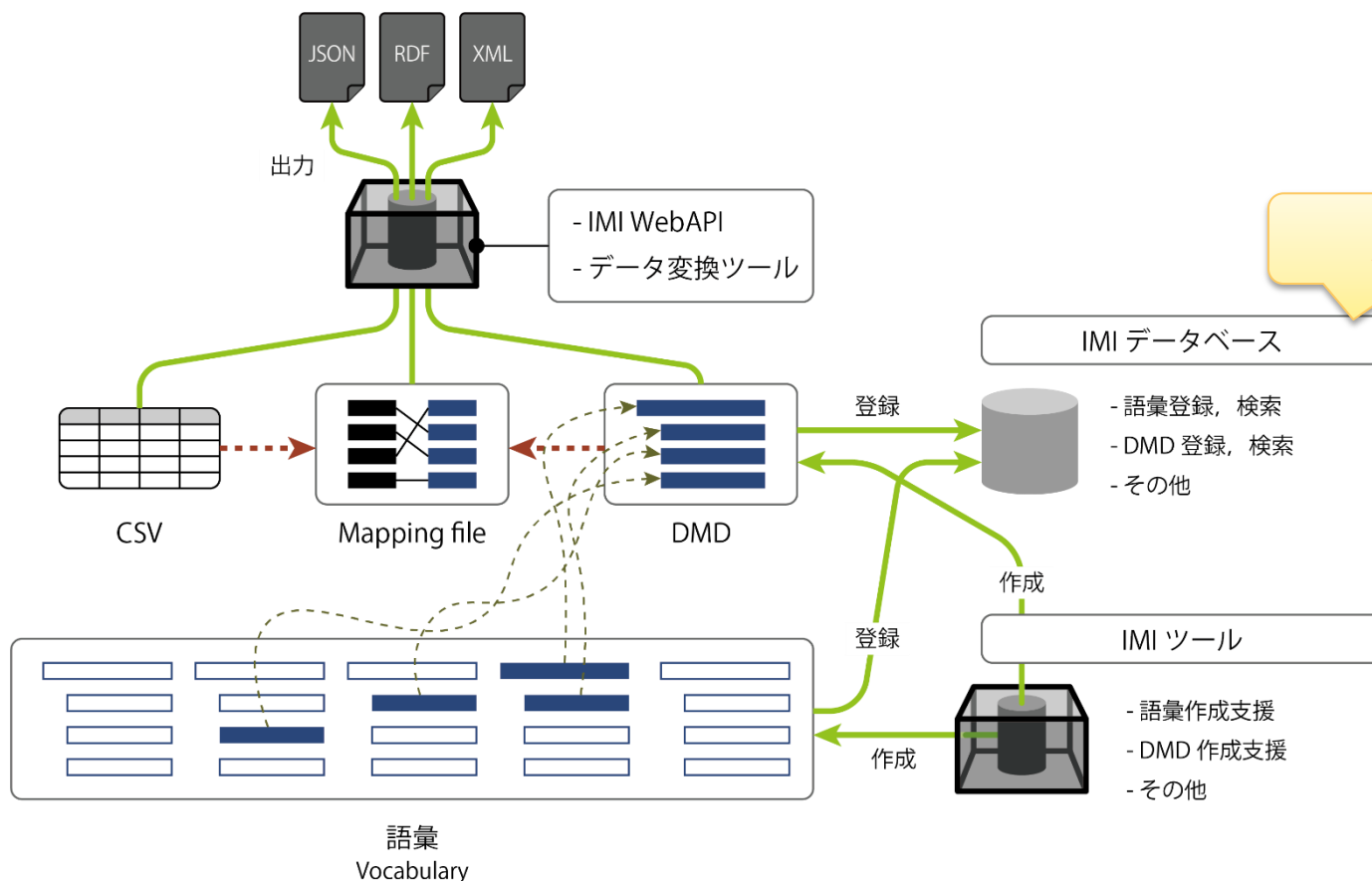
▶▶ 対応仕様

• IMI コア語彙	バージョン 2.4.1	2017-11-10
• IMI語彙記法	バージョン 1.0 ワーキングドラフト	2017-08-31
• 使用文字規則の記法	バージョン 1.0 ワーキングドラフト	2017-08-31
• DMD仕様	バージョン 3.0 ワーキングドラフト	2017-08-31
• マッピングファイル仕様	バージョン 1.0 ワーキングドラフト	2017-08-31
• 構造化項目名記法	バージョン 1.0 ワーキングドラフト	2017-08-31

出典：IMI検討部会事務局（IPA）「データ流通推進協議会向け勉強会資料（2018年8月8日）」

語彙定義のためのツールキットを無償で公開中

- ▶ 既存DMD、表形式データ(csv)、またはコア語彙からDMDを作成、検証する機能
- ▶ DMDに沿った形式でCSVデータをJSON、RDF、XMLに変換する機能
- ▶ コード・コードリスト及び語彙の作成機能



出典：IMI検討部会事務局 (IPA)「データ流通推進協議会向け勉強会資料 (2018年8月8日)」

IMIが公表するその他の教材リンク

1. 共通語彙基盤の策定及び管理手順

- https://imi.go.jp/contents/2018/07/IMI_ProcessDocument_WD20180326.pdf
- 共通語彙公開ドラフト(PD) (ドメイン語彙やコア語彙の候補) の定義、手続きを規定

2. 公開ドラフト一覧

- <https://imi.go.jp/pd>
- PD2342(法人情報に関する語彙)は、語彙一覧と語彙サンプルスキーマを公開
(<https://imi.go.jp/pd/pd2342>)

3. DMD仕様

- https://imi.go.jp/goi/dmd_spec
- 現在はバージョン3.0を公開。過去のバージョン2.1とは必須ファイルや語彙の表記法等が異なる

4. IMIツール説明会

- <https://imi.go.jp/event/event0006>
- IMI意見交換会に先立って開催されたIMIツール説明会の、配布資料と動画を公開
- (<https://www.youtube.com/watch?v=7CHMkOvF5Io&feature=youtu.be>)

IMI（共通語彙基盤）を使う利点

1. コア語彙とのマッピング

- 個別調整コストの削減

2. 一からコア語彙を利用し、ドメイン語彙を定義

- 語彙定義の早期定義、網羅性の確保に貢献

3. IMIツール利用

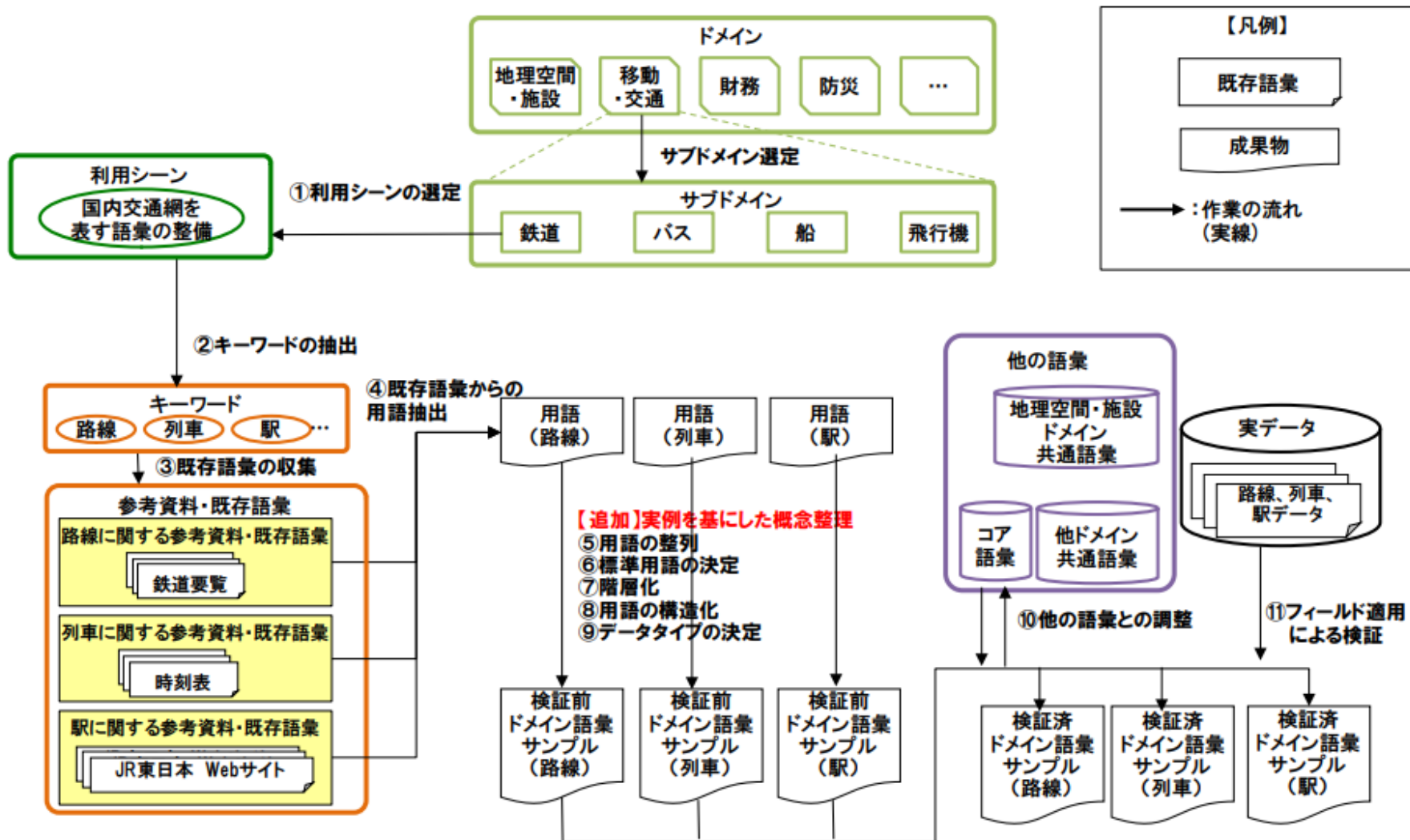
- 語彙定義、変換、アプリケーション開発のコスト削減

4. 共通語彙を利用したアプリケーション開発

- エコシステムの拡大により、たくさんの利用が想定できる

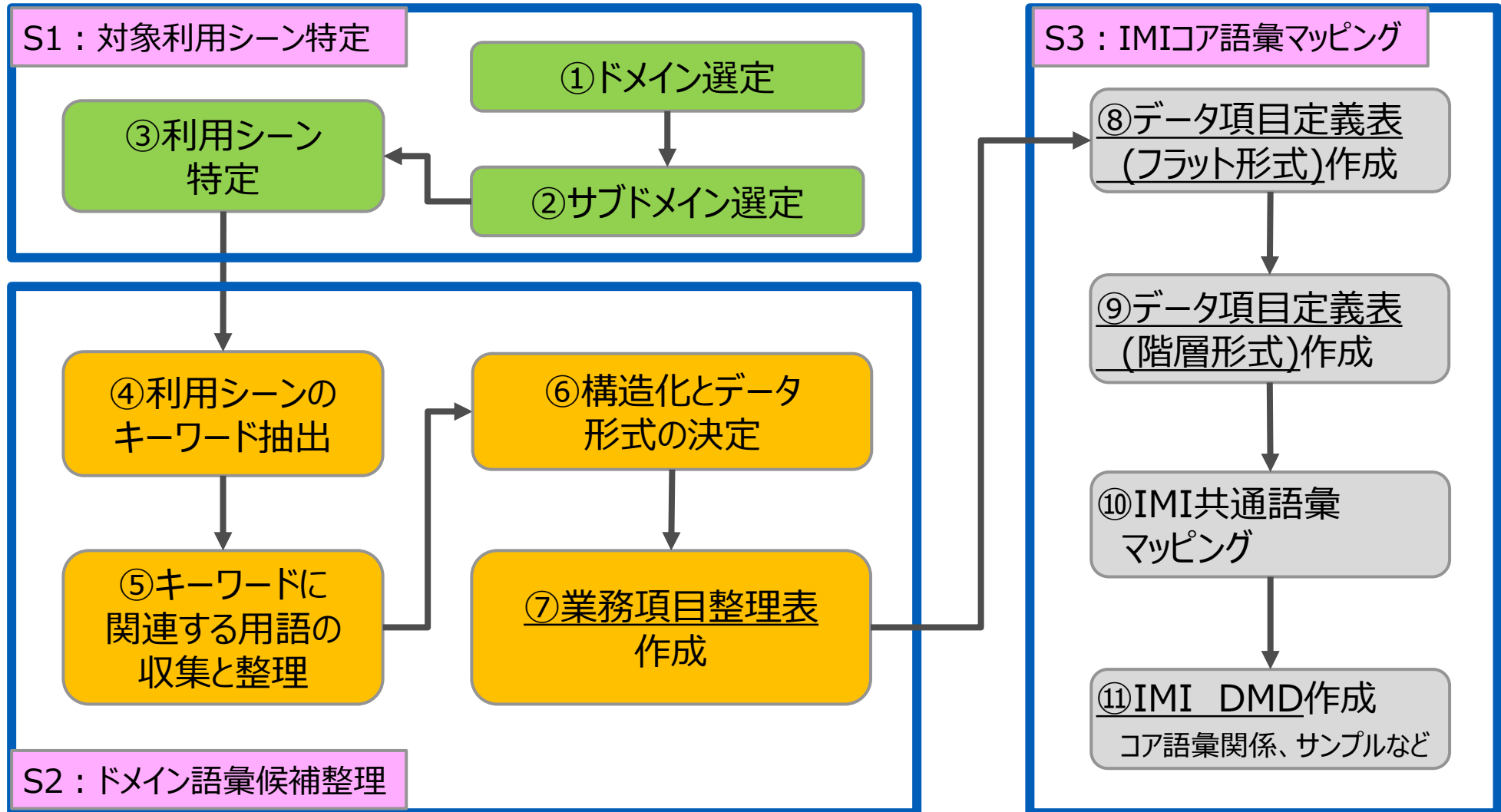
4.ドメイン語彙の策定 (作業概要)

参考：鉄道「ドメイン語彙」策定ステップの概観



出典：IMI検討部会事務局 (IPA)「データ流通推進協議会向け勉強会資料 (2018年8月8日)」

「ドメイン語彙」策定ステップの概観（略図）

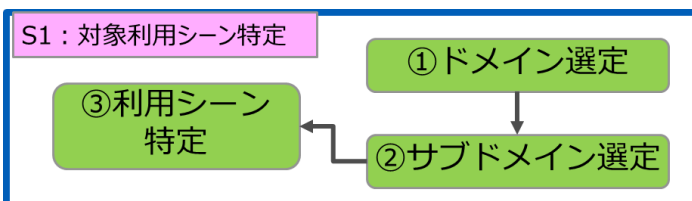


出典：IMI検討部会事務局 (IPA) 「データ流通推進協議会向け勉強会資料 (2018年8月8日)」をもとに作成

5.ドメイン語彙の策定 作業ステップ

S1:対象利用シーン特定

S1:対象利用シーン特定



①ドメイン選定

◎ドメイン語彙策定作業の対象とするドメイン、サブドメイン、利用シーンについて、関係者で調整する

②サブドメイン選定

◎調整した結果をワークシート(ドメイン語彙情報表)に記入する

③利用シーン特定

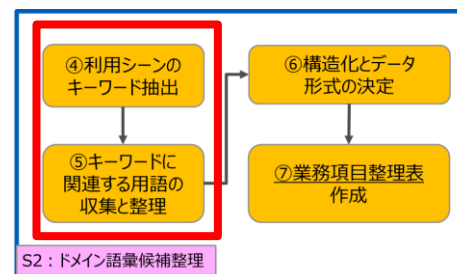
◎「データセット情報」はデータカタログ作成の場合に記載 (オプション)

ワークシート名: ドメイン語彙情報表

ドメイン情報	
ドメイン	建設業
サブドメイン	ビル管理
ユースケース名称	ビルのセンシングデータ収集し業界DBへ登録する
メタデータ	
名称	建物センシングデータ型
英語名称	
説明	建物センシングに関するデータ
英語説明	
バージョン	V0.2
作成日	2018-09-10
作成者	XXデータ流通推進協議会
作成者(英語)	
作成者URL	https://XXX-trading.org/
ライセンス	

データセット情報	
識別子	m01-b001
名称	建物センシングデータ型サンプルデータセット
説明	建物センシングデータのサンプルデータセットです
提供者	XXXデータ流通推進協議会
窓口	XXXデータ流通推進協議会
データセットの作成者	XXXデータ流通推進協議会
キーワード	建物、センシングデータ、サンプル
発行日	2018/10/4
更新/修正日	
提供頻度	1度
説明ページURL	
分類カテゴリ	サンプルデータ
対象地域	日本
対象期間	2018-08-01 / 2018-08-31
言語	ja

S2: ドメイン語彙候補整理 (前半)



④利用シーンの
キーワード抽出

⑤キーワードに
関連する用語の
収集と整理

◎ S1で選定した対象ドメイン／サブドメイン／利用シーンに関するキーワードを関係者にて抽出する。

例) 「ビル」「センシング」の「対象」、「測定方法」、「結果」

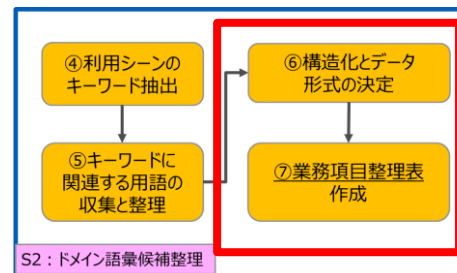
◎ (a)参加団体や公開資料などから、語彙候補となる用語を収集。

◎ (b)各団体の用語の同じ意味のものを整理し、代表用語やデータ例としてワークシートに記入する。

ワークシート名： ドメイン語彙収集表

キーワード	代表用語	データ例	(b)	(a) 収集用語 1	(a) 収集用語 2	(a) 収集用語 3
				A会社利用用語	B会社利用用語	C会社利用用語
測定方法	機器管理者	〇〇株式会社		管理者	測定機器担当	測定機器管理者
	測定日時	2018/8/20		日時	測定日時	測定した日時
	センシング方法	B		方法B	B	B
対象	建物名	〇〇ビル		測定対象	測定ビル名	建物名所
	住所	東京都千代田区神田1-1		測定対象の住所	—	住所
結果	部屋の名称	801会議室		設置の部屋名	部屋の名称	部屋名称
	部屋の階数	8		設置階	部屋の回数	部屋階数
	センサー設置位置	天-東		センサー設置位置	センサー設置位置	設置位置
	室温	27.8		室温	室温	室温
	湿度	49.2		湿度	湿度	湿度

S2: ドメイン語彙候補整理 (後半)



⑥構造化とデータ形式の決定

⑦業務項目整理表作成

ワークシート名：
業務項目整理表

- ◎ (a)代表用語を、データ項目として構造化(大項目、中項目、小項目)する
- ◎ (b)各データ項目の「データ形式」(データタイプ)を決定する。
コードリストがある場合は、別途表で整理する。
- ◎ (c)各データ項目の「データ項目の説明」を決定する
- ◎ 上記検討結果を、業務項目整理表に記載し関係者で確認する

(a) ↓ データ項目			(b) ↓ データ形式	(c) ↓ データ項目の説明
大項目	中項目	小項目		
建物センシング				建物センシングデータに関する情報を示す。
	機器管理者		テキスト	建物センシングに使用した機器の管理者（法人）を示す。
	測定日時		日時	建物センシングデータの取得日時を示す。
	センシング方法		コード	建物センシングデータのセンシング方法に該当するコードを示す。（コード表1）
センシング対象				建物センシングデータの対象（建物）に関する情報を示す。
	建物名		テキスト	センシング対象である部屋が属する建物の名称を示す。
	住所		テキスト	センシング対象である部屋が属する建物の住所を示す。
センシング結果				建物センシングデータのセンシング結果を示す。（最大3件まで）
	結果No		数値	センシング結果の通番を示す。
	部屋の名称		テキスト	センシング対象である部屋の名称を示す。
	部屋の階数		数値	センシング対象である部屋がある階数を示す。
	センサー設置位置		テキスト	センシング結果を得たセンサーの設置位置を示す。
	室温		数値	センシング結果のうち、室温の値を示す。単位：℃
	湿度		数値	センシング結果のうち、湿度の値を示す。単位：%
備考			テキスト	建物センシングデータの備考を示す。

S3: IMICOA語彙マッピング (前半)

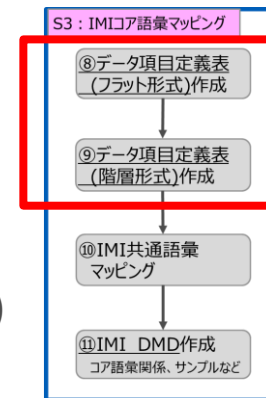
◎ S2で作成した①業務項目整理表をもとに以下を作成。

⑧データ項目定義表
(フラット形式)作成

CSV形式を想定した②データ項目定義表(フラット表形式)

⑨データ項目定義表
(階層形式)作成

XML形式を想定した③データ項目定義表(階層形式)



ワークシート名：
①業務項目整理表

データ項目			データ形式
大項目	中項目	小項目	
建物センシング	機器管理者		テキスト
	測定日時		日時
	センシング方法		コード
	センシング対象		
センシング対象	建物名		テキスト
	住所		テキスト
センシング結果	結果No		数値
	部屋の名称		テキスト
	部屋の階数		数値
	センサー設置位置		テキスト
	室温		数値
	湿度		数値

ワークシート名：
②データ項目定義表 (フラット形式)

No	データ項目名	サンプル値
1	建物センシングデータのID	m01-b001
2	建物センシングの機器管理者	〇〇株式会社
3	建物センシングの測定日時	2018/8/20
4	建物センシングのセンシング方法	B
5	センシング対象の建物名	〇〇ビル
6	センシング対象の住所	東京都千代田区神田1-1
7	センシング結果の結果No[1]	1
8	センシング対象の部屋の名称[1]	801会議室
9	センシング対象の部屋の階数[1]	8
10	センシング結果のセンサー設置位置[1]	天-東
11	センシング結果の室温[1]	27.8
12	センシング結果の湿度[1]	49.2

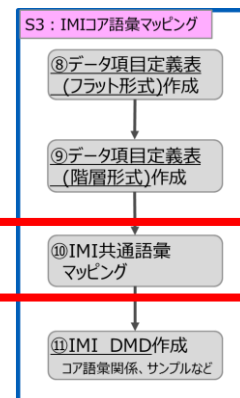
ワークシート名：
③データ項目定義表 (階層形式)

タグ名					出現回数	
Lv1	Lv2	Lv3	Lv4	Lv5	最少	最大
建物センシングデータ						
ID					1	1
建物センシング						
機器管理者					1	1
測定日時					1	1
センシング方法					1	1
センシング対象					1	1
建物名					1	1
住所					1	1
センシング結果[1]					1	3
結果No					1	1
部屋の名称					1	1
部屋の階数					1	1
センサー設置位置					1	1
室温					1	1
湿度					1	1

S3: IMIコア語彙マッピング (後半 1)

⑩IMI共通語彙マッピング

◎ S3前半で作成した③データ項目定義表(階層形式)に、IMIコア語彙との対応を検討し、結果をワークシートに記入。



③データ項目定義表
(階層形式)

④IMI共通語彙マッピング

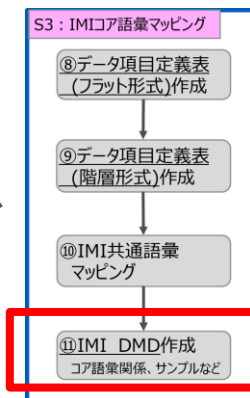
タグ名					出現回数	
Lv1	Lv2	Lv3	Lv4	Lv5	最少	最大
建物センシングデータ						
	ID				1	1
建物センシング						
	機器管理者				1	1
	測定日時				1	1
	センシング方法				1	1
センシング対象						
	建物名				1	1
	住所				1	1
センシング結果[1]						
	結果No				1	1
	部屋の名称				1	1
	部屋の階数				1	1
	センサー設置位置				1	1
	室温				1	1
	湿度				1	1

プレフィックス+ クラス項目名/ プロパティ項目名	コア語彙 クラスの継承元/ プロパティの値型	マッピング
ex:建物センシングデータ型	ic:事物型	
ID	ic:ID型	ex:建物センシングデータプロファイル型>ic:ID>ic:識別値;
建物センシング	ex:建物センシング型	
機器管理者	ic:法人型	ex:建物センシングデータプロファイル型>ex:建物センシング>ex:機器管理者>ic:名称>ic:表記;
測定日時	ic:日付型	ex:建物センシングデータプロファイル型>ex:建物センシング>ex:測定日時>ic:表記;
センシング方法	ic:コード型	ex:建物センシングデータプロファイル型>ex:建物センシング>ex:センシング方法>ic:識別値;
センシング対象	ex:センシング対象型	
名称	ic:名称型	ex:建物センシングデータプロファイル型>ex:センシング対象>ic:名称>ic:表記;
住所	ic:住所型	ex:建物センシングデータプロファイル型>ex:センシング対象>ic:住所>ic:表記;
センシング結果	ex:センシング結果型	
結果No	ic:数量型	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:結果No>ic:数値;
部屋の名称	xsd:string	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:部屋の名称;
部屋の階数	ic:数量型	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:部屋の階数>ic:数値;
センサー設置位置	xsd:string	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:センサー設置位置;
室温	ic:数量型	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:室温>ic:数値;
湿度	ic:数量型	ex:建物センシングデータプロファイル型>ex:センシング結果[1]>ex:湿度>ic:数値;

S3: IMIコア語彙マッピング (後半2)

⑪ IMI DMD作成
コア語彙関係、
サンプルなど

◎ IMIのWebサイトで公開されている、IMIツールを利用し、DMDを作成する。DMD作成は、有識者支援が必須。



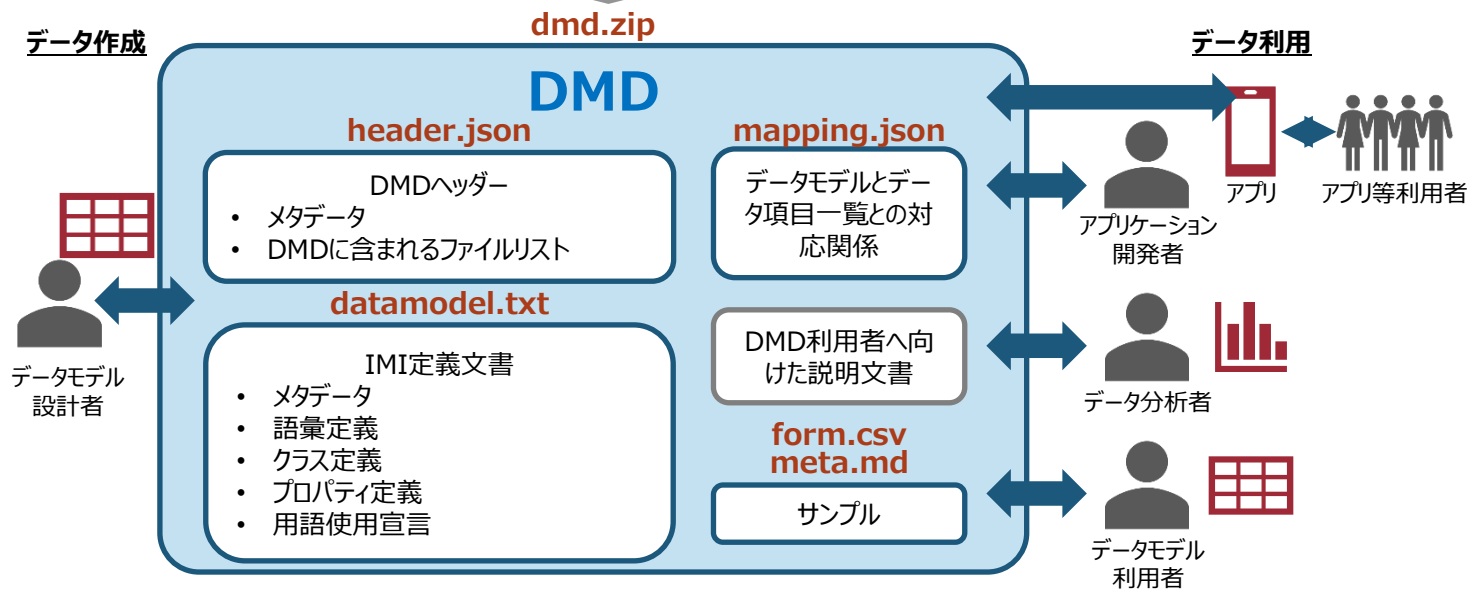
IMI Infrastructure for Multi-layer Interoperability Ver. 1.0.0

DMD作成支援 DMD検証 データ形式変換 データ検証 コード・コードリスト作成支援 語彙作成支援

IMIツール

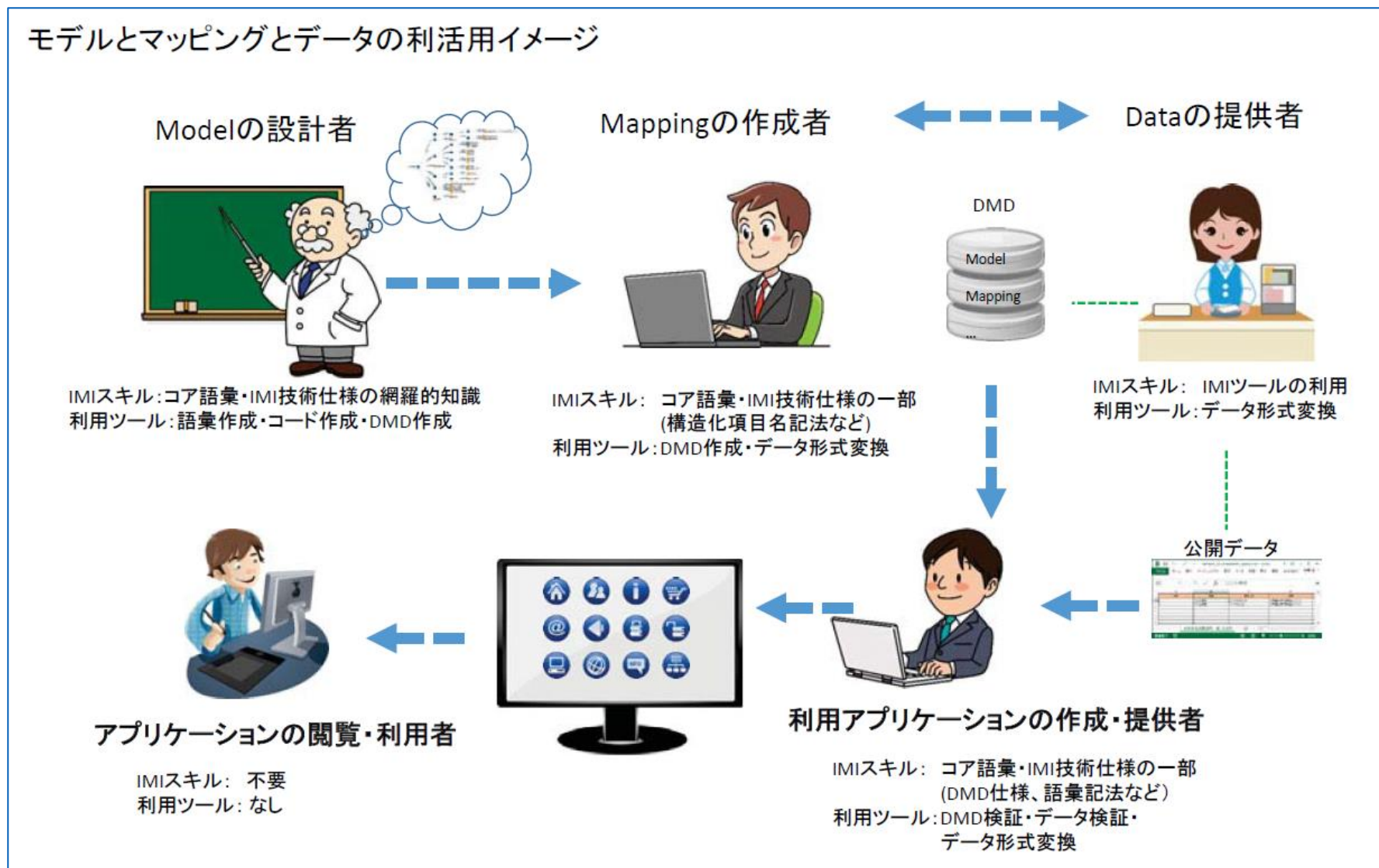
- DMD作成支援**: 主にデータ構造設計者が利用する機能で、提供するデータの構造を定義し共有するためのDMDを作成します。
- データ形式変換**: 主にデータ提供者が利用する機能で、表計算ソフトやCSVなどのデータをDMDに定義された構造化データへ変換します。
- コード・コードリスト作成支援**: コードやコードリストの定義を構造化データとして作成するための支援機能です。
- DMD検証**: IMIツール以外で作成したようなDMDについて、その内容がDMD仕様やコア語彙等と整合していることを検証する機能です。
- データ検証**: 構造化データを入力として、そのデータがDMDと整合していることを検証する機能です。
- 語彙作成支援**: コア語彙以外の語彙を語彙定義ファイルとして作成するための支援機能です。

出典：IMI公式サイト「IMIツール」
(<https://imi.go.jp/tools/v2/>)



出典：IMI検討部会事務局 (IPA)「データ流通推進協議会向け勉強会資料 (2018年8月8日)」

IMI ツールの利用イメージ



出典: IPA IMI ツール(検証版)説明会資料 <https://imi.go.jp/event/event0006>

Copyright©Sustainable open Innovation Initiative. All Rights Reserved.

【ドメイン語彙事例】埼玉県 ご当地キャラ情報

コア語彙になかった「モチーフ」「性格」「趣味」などを独自に定義

No	データ項目	必須	共通語彙	データ形式	データ例	備考
1	識別情報	○		テキスト (改行なし)	1	キー情報。一意の数字を設定する。
2	団体コード	○	団体_ID	テキスト (改行なし)	112345	
3	団体名	○	団体_名称	テキスト (改行なし)	〇〇市	
4	名称	○	名称_表記	テキスト (改行なし)	コバトン	
5	モチーフ				シラコバト	
6	概要			テキスト (改行なし)		20字~100字程度の文章とすること。
7	身長				不明	
8	体重				不明	
9	性格				おっとり	
10	趣味				埼玉県の魅力を発信すること	
11	誕生日			テキスト (改行なし)	5月31日	
12	画像	○	画像_画像		http://www.pref.saitama.lg.jp/a0301/kobaton/images/370370.jpg	
13	画像_ライセンス	○		テキスト (選択)	CC BY	画像に対するライセンス表記をクリエイティブコモンズライセンスの表記を

▼データレコードの確認例▼

モチーフ	身長	体重	誕生日	性格	趣味
ゆず	ゆず5個分	ゆず5個分	康平6年(1063年)1月3日	たくさんの人に愛えられ	ゆず狩り
ウサギのようでシカ	深谷ねぎ5本分 体の大きさを愛えられ (大きくなったり、 ちいさくなったりでき る)。	深谷ねぎ480本分	6月28日	やさしい、まげずらい	つぶや
戸田の原っぱに暮らす		不明	不明	不明	戸田ヶ
男子鳩輪			古墳時代(5世紀前半)	男	遠路め
女子鳩輪			古墳時代(5世紀前半)	女	遠路め
町の魚「ミヤコタナゴ	ミヤコタナゴと同じ大	ないしょ	11月3日(薄川町がで	明るく元気いっぱい	子ども

対応語彙が コア語彙の「人」身長」や「人」体重」、
ない例 「人」誕生日」では表現しづらい例 対応語彙がない例

```
PREFIX ic: <http://imi.go.jp/ns/core/rdf#>
PREFIX pd3110:
<http://imi.go.jp/ns/pd/3110/1_0_0#>
PREFIX pd3119:
<http://imi.go.jp/ns/pd/3119/1_0_0#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX schema: <http://schema.org/>
```

```
select
distinct ?dantai_code ?dantaimei ?shikibetsujoho
?meisho ?motif ?gaiyo
?shincho ?taiju ?seikaku ?shumi ?tanjobi
?gazo ?gazo_license ?gazo_text
?tantoka ?tantoka_denwabango ?tantoka_email
?jiyukijutsuran ?url
```

where {

```
graph<http://data.pref.saitama.lg.jp/graph/10_yuru
chara>
```

```
{ ?KEY pd3110:団体コード ?dantai_code;
pd3110:団体名 ?dantaimei;
pd3110:識別情報 ?shikibetsujoho;
ic:名称/ic:表記 ?meisho;
ic:概要 ?gaiyo;
```

```
pd3119:モチーフ ?motif;
pd3119:身長 ?shincho;
pd3119:体重 ?taiju;
pd3119:誕生日 ?tanjobi;
pd3119:性格 ?seikaku;
pd3119:趣味 ?shumi;
```

```
schema:image/schema:contentUrl ?gazo;
...
```

「pd3110:」「pd3119:」で始まる項目名は、埼玉県オープンデータ共通フォーマットで独自に定義された
 応用語彙。 https://opendata.pref.saitama.lg.jp/fs/9/6/_/gotoutikyarajyoho.pdf

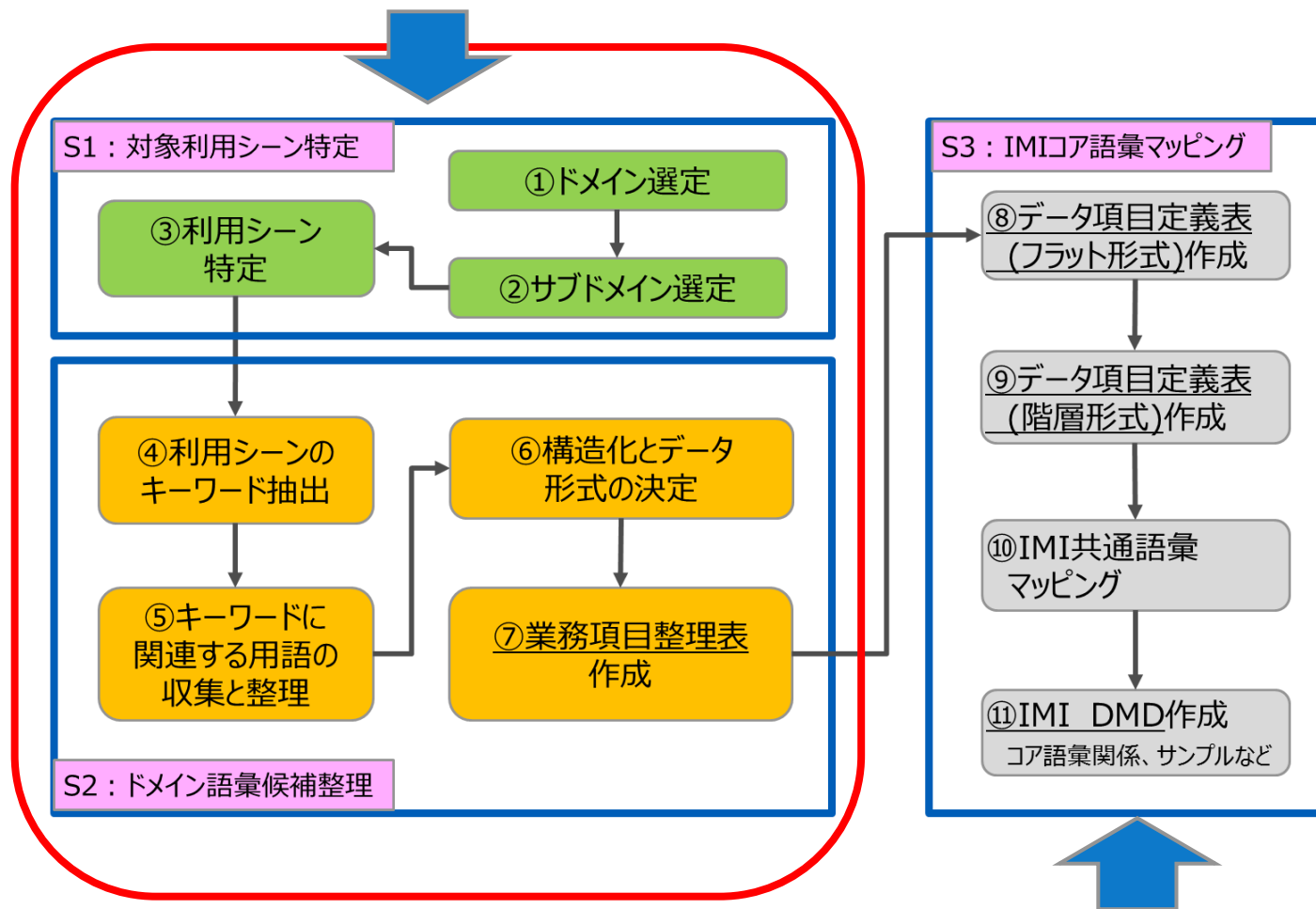
出典:「官民データの利活用セミナー 官民連携で取り組んだオープンデータ構築」(https://imi.go.jp/contents/2018/07/20180129_r4.pdf)

【確認】「誰も使わない標準」にならないように

- **現場ではどんな語彙が用いられているか**
 - 慣用的なビジネス語彙と、正規の技術語彙を分けざるを得ないことも
- **「コア語彙に使える語彙があるか」を見ておく**
 - 時間、場所、組織、人など、よく使う概念の語彙は、どの業界でも一般的
- **ドメイン語彙候補ができれば、IMIへ登録しよう**
 - 自業界の語彙が、他のドメインでも活用されることが期待できる

【推奨】ドメイン語彙の作成を検討したい方へ

まずはS1, S2を実施



S3は、IMI有識者の支援を受ける

