



Infrastructure
for Multi-layer Interoperability

共通語彙基盤(IMI)に対応したデータとは?

2018年6月
IMI検討部会事務局 (IPA)

そもそも共通語彙基盤って何？

人が使っている用語の意味を「共通的に定義したい」と思っているのが共通語彙基盤

- ▶ 一般的な語彙とは
 - ▶ ある範囲で使われる語の全体（-を集めたもの） [三省堂ウェブディクショナリー]
 - ▶ ある言語、地域・分野、ある人、ある作品など、それぞれで使われる単語の総体 [デジタル大辞泉(小学館)]
- ▶ 共通語彙基盤でいう語彙とは
 - ▶ 用語の持つ意味や典拠を明確にするとともに、クラス概念・プロパティ概念により体系化・階層構造化し、その概念を正確にクラス用語・プロパティ用語として表現した用語の集合を語彙と呼びます。

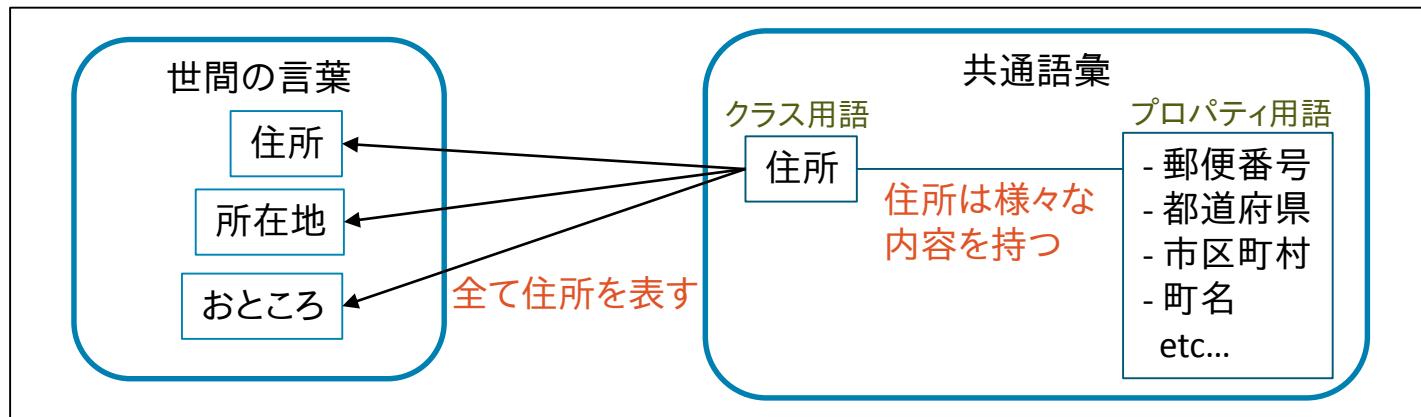


図1 共通語彙基盤における語彙のイメージ

用語=世間一般・自治体等が使っている言葉（住所・所在地・郵便番号・氏名。。。）

住所を表す用語は「所在地・おところ・宛先」など様々 -> 全て住所であることを定義したい

住所と一言に言っても内容は様々。 都道府県名、市区町村名、郵便番号、等々

-> 住所という用語はこのような内容を持っているということを定義したい



何故定義したいのか？(少し技術的なお話)

多種多様なデータを機械(コンピュータ)処理するためには、そのデータの意味を誰かが一元的に定義するのが最も効率的。(=共通語彙基盤)

- ▶ 世の中に散在するデータは世間一般の言葉で自由に記述されている
 - ▶ A市では「居住者一覧」として「氏名」「居住地」の項目名でデータ作成
 - ▶ B市では「住民一覧」として「名前」「住所」「郵便番号」の項目名でデータ作成。
 - ▶ 同じデータなのに機械(コンピュータ)で処理しようとすると項目名がちがうために個別対応が必要
 - ▶ 各データ項目が何を意味するのか共通語彙基盤の語彙に関連付けられれば、共通的に機械処理が可能になる

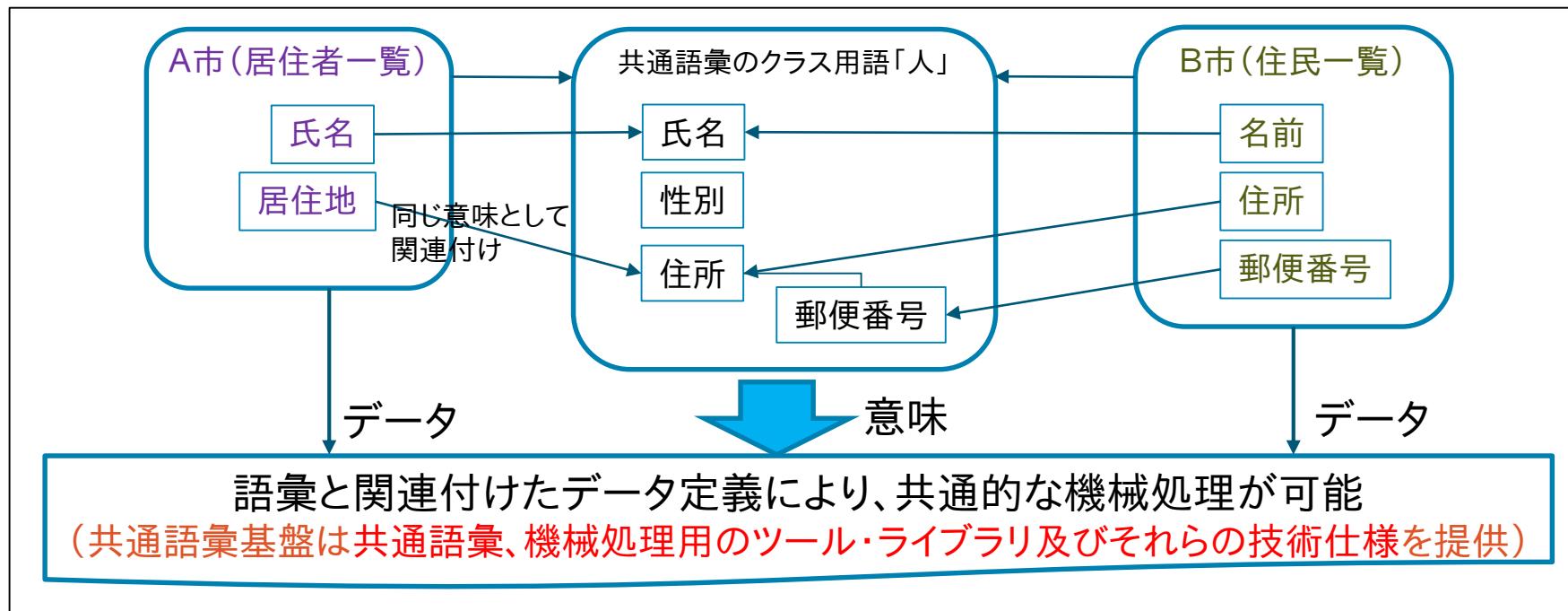


図2 共通語彙によるデータ利用イメージ

共通語彙基盤(IMI)が提供する技術・情報

今まで:語彙基盤の中核をなす共通語彙(コア語彙)の整備と、語彙活用に必要な技術仕様の策定を先行パートナーの協力のもと実施。(概念的な話が多く難解)

これから:コア語彙・技術仕様とも大きなブレが無い状態に整備が完了。語彙基盤活用ツール(IMIツール)も実用レベルで拡充。IMIが具体性を持って活用可能に。

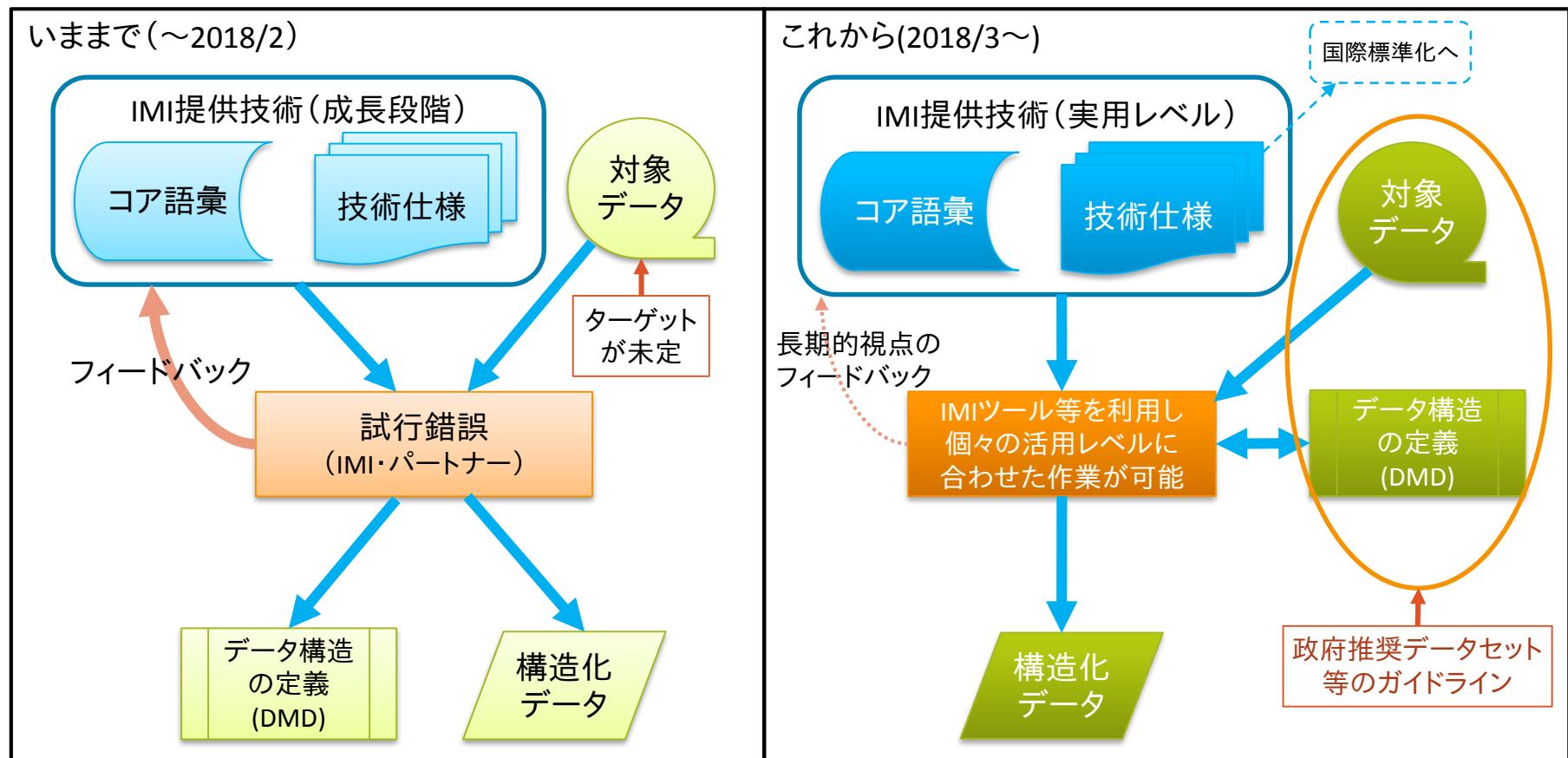


図3 からの共通語彙基盤活用



IMIに対応したデータ公開とは？

データ構造の定義(DMD)と構造化データを合わせて公開すればIMI対応

DMD : IMIツールで作成したものであれば全てOK！

手作業で作成したもの等でもIMIツールによる検証OKでIMI対応！

データ : DMD+CSVからIMIツールで生成できればOK！

手作業で作成したデータもIMIツールによる検証OKでIMI対応！

おすすめ対応手順(いきなり技術仕様に取り組んではダメ！)

最初は難しいことにこだわらず、IMI対応データを公開して欲しい。

- ▶ Step1-1 政府推奨データセット(推奨DS)を基に、
対応が容易なデータの構造定義を行う(DMD作成)
- ▶ Step1-2 構造化データを生成する
- ▶ Step1-3 データを公開する
- ▶ Step2 独自DMDや語彙(応用語彙)が必要になったら
技術仕様を確認
- ▶ Step3 世にある様々なデータとの連携を考慮したデータ公開



IMI対応とは？ おすすめ対応手順

▶ Step1-1 政府推奨データセット(推奨DS)を参考に、対応が容易なデータの構造を検討しDMDを作成する

- ▶ 「オープンデータ一覧」「公共施設一覧」「医療機関一覧」などがおすすめ
 - ▶ これらデータセットのDMDはIPAからも公開予定
(推薦DSの形式どおりに使うもよし、既存データに合わせてカスタマイズするもよし)

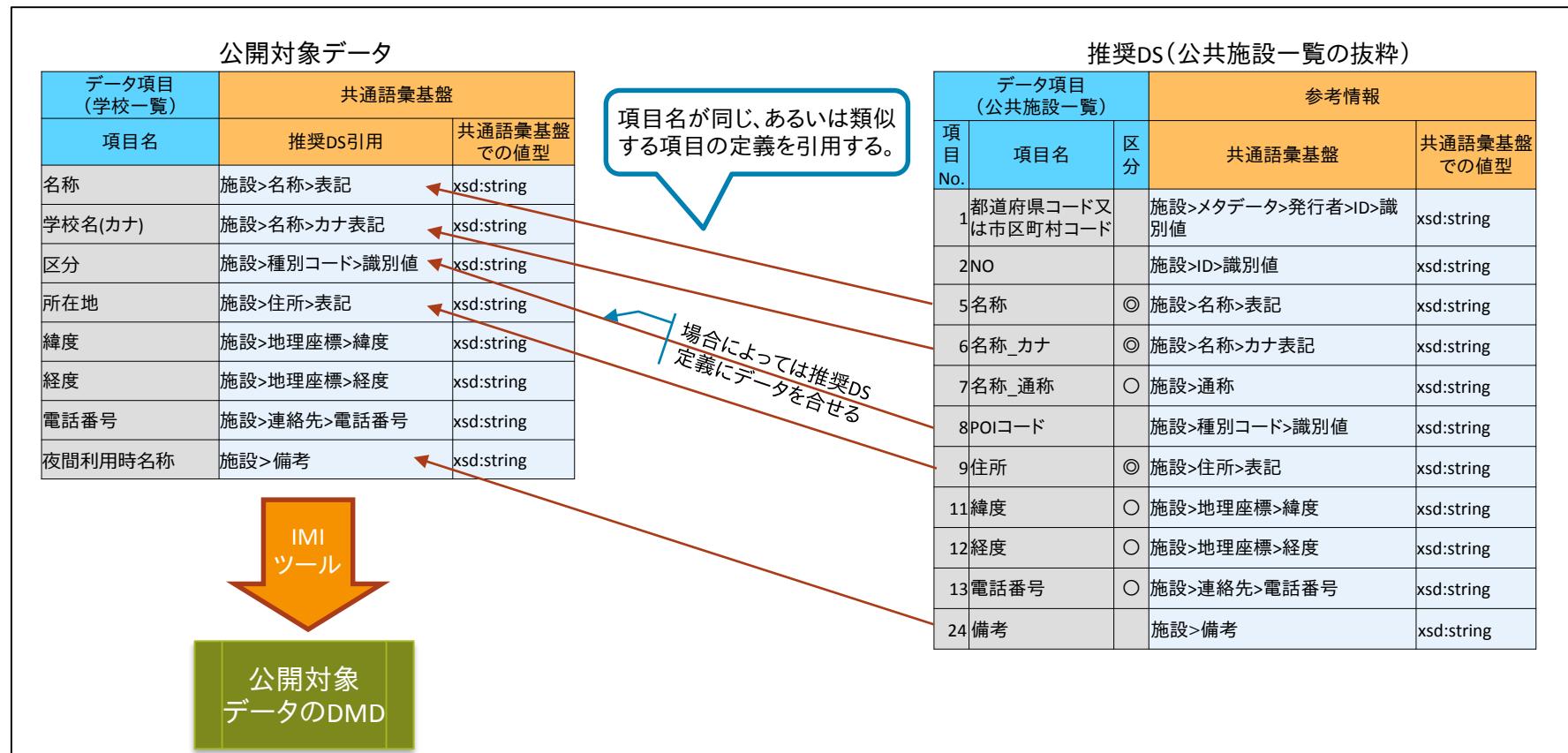


図4 推奨DSを引用したDMD作成



IMI対応とは？ おすすめ対応手順



▶ Step1-2 構造化データを生成する

- ▶ CSV等の元データ+DMDを入力として、IMIツールで構造化データを生成する



図5 IMIツールによる構造化データ生成

▶ Step1-3 データを公開する

- ▶ 元データ(CSV等) +DMD+構造化データを一組のデータセットとしてホームページなどに公開
- ▶ IMI対応データとしてロゴの併記を忘れずに実施

構造化データはCSVとDMDからツール生成可能ですが、利用者が内容を確認しやすい形式で公開をお勧めします。

▶ Step2 IMIツールに慣れ、独自DMDや語彙(応用語彙)が必要になったら技術仕様を確認

- ▶ 推奨DSにも応用語彙がないとDMD作成が困難なものがあるため、IPAから応用語彙付きDMD雛形を公開予定
- ▶ 推奨DSのDMD雛形と技術仕様を併せてみることで仕様の理解が促進される

▶ Step3 世にある様々なデータとの連携を考慮したデータ公開

- ▶ 個々のデータにIDを付与したり、コードの共通化を図ったり。。。
- ▶ IMIに対して語彙や技術仕様に関するフィードバックもお願いします



- ▶ IMI(Infrastructure for Multilayer Interoperability:情報共有基盤)サイト
 - <https://imi.go.jp/>
- ▶ 共通語彙基盤サイト
 - <https://imi.go.jp/goi/>
- ▶ IMIツール(検証版) → 6月中旬に大幅な機能改修を実施予定
 - <https://imi.go.jp/goi/dmd-editor.html>
- ▶ IMIロゴ
 - <https://imi.go.jp/imi/logo-intro.html>
- ▶ IMI共通語彙基盤 ガイド・解説・技術仕様等
 - <https://imi.go.jp/goi/contents-list.html>
- ▶ IMI共通語彙基盤を使って「官民データの利活用」に参加しよう
 - https://imi.go.jp/goi/public-and-private-sector-data_utilization.html
- ▶ 政府CIOポータル オープンデータ
 - <https://cio.go.jp/policy-opendata>
- ▶ 政府CIOポータル オープンデータ 自治体ガイドライン・手引書
 - <https://cio.go.jp/policy-opendata#guideline>
- ▶ 政府CIOポータル オープンデータ 推奨データセット
 - <https://cio.go.jp/policy-opendata#dataset>

IMIを紹介するもの、あるいはIMIを参照して作成したもの
(データや、それを活用するアプリケーションソフト等を含む)
であればIMIロゴを利用可能





Infrastructure
for Multi-layer Interoperability

終わり