

第 4 回共通語彙基盤ワーキンググループ

日時：2016 年 2 月 26 日（火）16:00～18:00

場所：経済産業省 別館 11 階 1115 室

<議事次第>

- 1 開会
- 2 共通語彙基盤に関する政府等の取り組み状況
- 3 共通語彙基盤事業進捗状況報告
- 4 2016 年度方針について
- 5 その他
- 6 閉会

<配布資料>

資料 0 議事次第

資料 0-1 委員名簿

資料 0-2 席次表

資料 1 共通語彙基盤に関する政府の取り組み

資料 2 共通語彙基盤(IMI)事業進捗状況

資料 3 相互運用性 SWG 設置の提言案

資料 4 共通語彙基盤(IMI)事業方針

参考資料 1 コア語彙 2.3 リリースノート

参考資料 2 Guidelines for mapping core data models

参考資料 3 ツール「表から RDF」のご紹介

第 4 回 共通語彙基盤ワーキンググループ 委員名簿

2016 年 2 月 26 日

委員長

武田 英明 国立情報学研究所情報学プリンシプル研究系 教授

委員

越塚 登 東京大学大学院情報学環 教授
坂下 哲也 一般財団法人日本情報経済社会推進協会(JIPDEC) 電子情報利活用研究部
部長
菅又 久直 国連 CEFACT 日本委員会 サプライチェーン情報基盤研究会 事務局長
深見 嘉明 立教大学 大学院ビジネスデザイン研究科 特任准教授
武藤 俊一 一般財団法人全国地域情報化推進協会企画部 担当部長

オブザーバー

内閣官房社会保障改革担当室
内閣官房情報通信技術 (IT) 総合戦略室
内閣府政策統括官(防災担当)付
総務省行政管理局行政情報システム企画課
総務省情報流通行政局情報流通振興課
国土交通省総合政策局情報政策課
国土地理院企画部企画調整課
国立国会図書館電子情報部電子情報サービス課

事務局

村上 智信 経済産業省情報プロジェクト室
平本 健二 経済産業省情報プロジェクト室
松田 圭介 経済産業省情報プロジェクト室
田代 秀一 独立行政法人情報処理推進機構 (IPA)

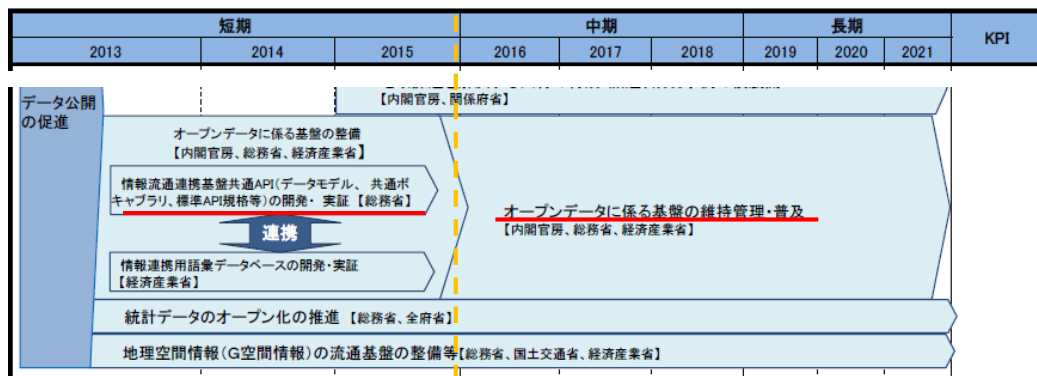
共通語彙基盤に関する政府の取り組み

世界最先端IT国家創造宣言（改定）のフォローアップ

平成27年6月30日 閣議決定

- 工程表に従い、情報連携用語彙データベースベータ版の整備とサイトの整備を実施。
- オープンデータの活用のために各種情報提供を実施。

公共データの民間開放（オープンデータ）の推進



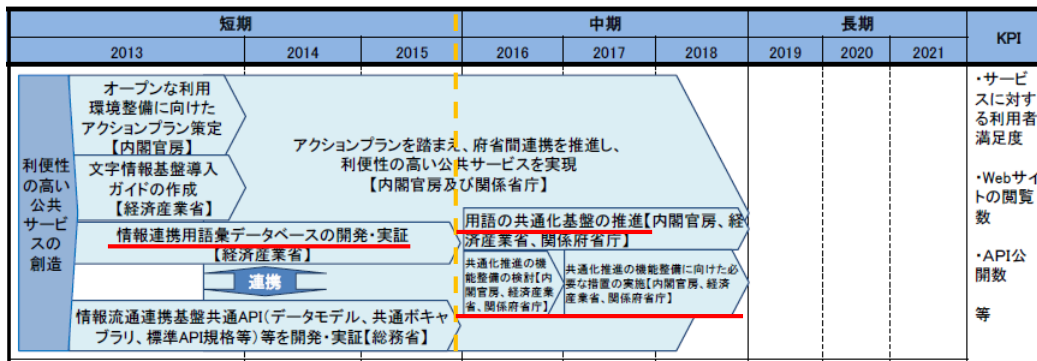
【短期（～2016年度）】

- ・経済産業省は、国際的な連携を踏まえた情報連携用語彙データベース（基本語彙の定義等をまとめたデータベース）に関して、2014年度に開発したパイロットシステムの地方公共団体における適用評価等を踏まえつつ2015年度までに本格版を開発・実証する。総務省及び経済産業省で成果を連携しつつ実証を進め、**実証成果はデータカタログでも活用**する。【内閣官房、総務省、経済産業省】
- ・2020年に東京オリンピック・パラリンピック競技大会が開催され、観光、防災等の地域の情報のニーズが高まると考えられることなどから、公共団体が保有する公共データについて、情報流通連携基盤共通API及び情報連携用語彙データベースの開発・実証、公共クラウドの構築や成功事例の整理等により、**オープンデータ化を支援**する。【内閣官房・総務省・経済産業省・関係府省】

【中長期（2017年度～2021年度）】

オープンデータに係る基盤の維持管理を行うとともに、普及を図る。【内閣官房、総務省、経済産業省】

利便性の高い電子行政サービスの提供



【短期（～2015年度）】

- ・経済産業省において、国際的な連携を踏まえた情報連携用語彙データベース（基本語彙の定義等をまとめたデータベース）に関して、2014年度に開発した評価版の地方自治体における適用評価等を踏まえつつ、2015年度までに本格版を開発・実証する。

【中期（2016年度～2018年度）・長期（2019年度～2021年度）】

- ・電子行政分野におけるオープンな利用環境整備に向けたアクションプランを踏まえ、用語の共通化基盤に関しての推進を行う。また、電子行政分野におけるデータ等の共通化・相互運用性確保について、**企画立案・維持管理を継続的に行うことができるようにするための機能（日本版NIST）**の整備について検討し、必要な措置を行う。【内閣官房、経済産業省及び関係府省庁】

- 共通語彙基盤の整備が進んできたことから、各所で先行プロジェクトが開始している。
- 内閣官房
 - 自治体向けオープンデータ用語彙の開発、実装
 - 自治体の先進モデルとして、共通語彙基盤ベースの語彙を組み込んだカタログシステムの整備
- 内閣官房・経済産業省
 - 法人用語彙の検討
 - 法人用語彙を政府全体に展開するための導入策の検討
- 経済産業省(+情報処理推進機構)
 - コア語彙の開発、運用
 - 住所、氏名等のコア語彙の開発及び運用
 - 法人用語彙の実装
 - 法人ポータルサイト(仮称)での共通語彙基盤の導入
 - CivicTech向けオープンデータ用語彙の開発
 - CivicTechでのニーズのある語彙の開発。(子育て支援施設等)
 - 国際調整
 - 米国、欧州の行政語彙との整合性調整
- 農林水産省
 - 農作物語彙、農作業語彙の検討
 - 農作物や農作業など農業システムの基盤となる語彙の検討
- 総務省
 - 統計局LoD検討
 - 統計局データのLoD化に伴う検討

法人ポータルサイトでの 共通語彙基盤導入の取り組み

法人ポータルでは府省内にあるデータを関係させることが必要

- これまでの府省内のデータは、担当課毎にフォーマットを作り作成
- そのため、データをまとめるためには、変換作業が必要。
- データの内容も、株式会社と記述したり、（株）と省略する等、未統一で使いにくかった。



法人ポータルのためにデータのフォーマットをきれいにすると言うより、日常の業務を楽にするという観点からも、共通的なフォーマットを使うのは有効

表彰対象	賞名	地域	受賞者	...
		東京		

受賞日	賞名	地域	受賞企業	本社住所
		関東		



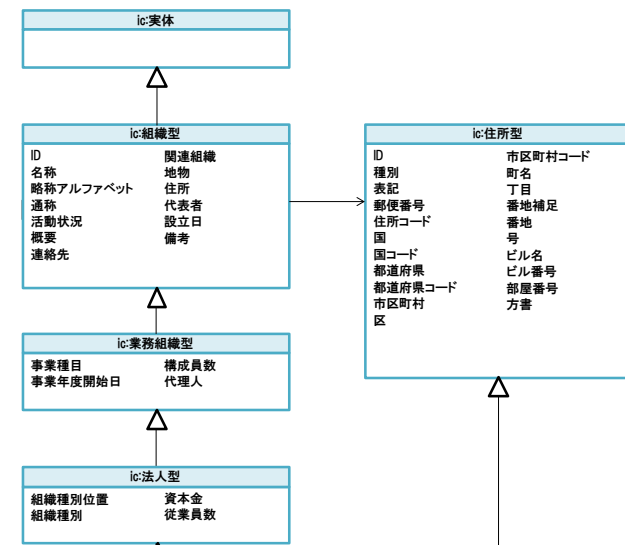
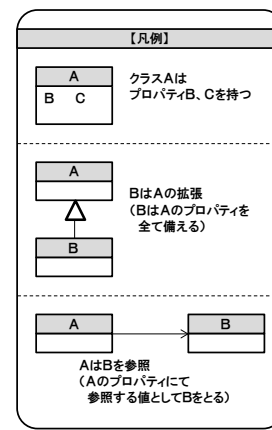
フォーマット統一すると

- ・複数ファイルを統合した表が作れる
- ・管理簿を作る時に検討が簡単
- ・管理簿の項目に抜けがない
- ・検索が容易

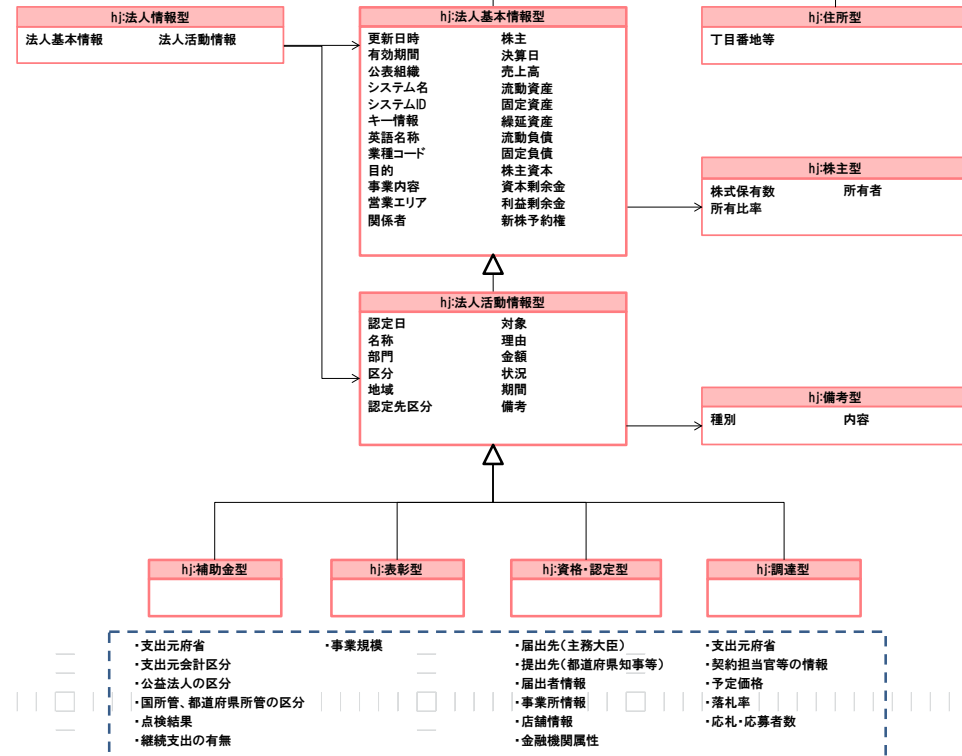
法人ポータルデータのデータ構造

- 「電子行政分野におけるオープンな利用環境整備に向けたアクションプラン」が推進する共通語彙基盤（データの交換基盤）に準拠したデータ構造
- 国際的な整合性も考慮してデータが設計されている。

データの拡張性が高い



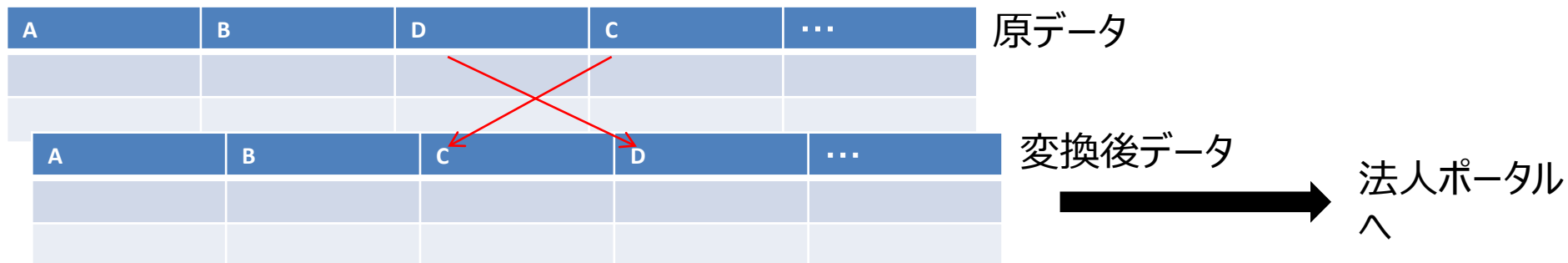
コア語彙
経済産業省
法人関連情報語彙



既に持っているデータの扱い

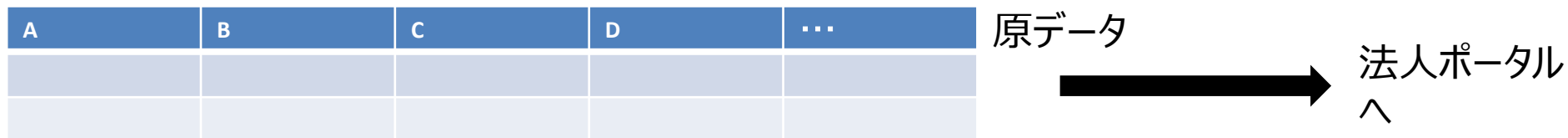
● オプション 1

- データの列の変更だけを行う



● オプション 2

- 元の独自データテンプレートを法人ポータルの書式に変えてしまう（独自項目の追加も可能）



● オプション 3

- 元の独自データテンプレートを法人ポータルの書式に変えて、更に、APIで公開

データフォーマット例

- 様々なデータを、法人活動の派生系で定義することで、データの組み合わせをしやすくしている。

法人活動(行政活動)	表彰	資格・認定	補助金	調達	検査	処分等
認定日	表彰日	認定日	認定日	入札公告日	検査日	決定日
名称	表彰名	資格名	補助金名		検査名	処分名
部門	表彰部門	認定部門	認定部門		部門	
区分	受賞区分	認定クラス	認定クラス			
地域	地域	地域	地域		地域	地域
認定元	表彰元	認定元	交付元	調達機関	検査機関	処分決定者
認定先	表彰先	認定先	交付先	落札者	検査対象者	
認定先区分	表彰先区分					
対象	表彰対象	対象物	対象物	調達件名及び数量	対象物	対象
理由	理由	理由	理由		理由	理由
金額			金額	落札価格		
状況		認定状況	執行状況		状況	執行状況
期間(開始日)		有効期間開始日	期間開始日	落札決定日	開始日	処分期間開始日
期間(終了日)		有効期間終了日	期間終了日		終了日	処分期間終了日
				品目分類番号		
				調達方式		
				契約方式		
				随意契約理由		
				指名業者名		
				落札方式		
				予定価格		
					結果	

法人ポータルへのデータ提供の流れ

- 各担当部局のデータは、法人ポータル用のデータ形式に変換した上で登録を行う。
- データ変換作業はあるが、統合して活用することで、各担当部局も他のデータを見ることが可能になったり、幅広く情報を公開することができる等のメリットの方が大きい
- 法人ポータル対応を機に、データの管理方式の見直しを図ることができる。

各部局で
バラバラなデータ

○●大賞
優秀賞
関東地域
○●株式会社
○●測定器
連絡先

●●優秀企業
関東地域
○●株式会社
連絡先

●●認定企業
○●株式会社
認定理由
連絡先

変換



登録データ形式 (例)

表彰日
表彰名
表彰部門
受賞区分
地域
表彰元
表彰先
法人番号
法人名
法人本社住所
対象事業所住所
表彰先区分
表彰対象
理由

送付



将来はAPI



※データの変換は表の列の入れ替えだけなので簡単な作業でできます。(データ登録タイミングや周期は、担当部局が判断。)

※法人ポータルを機に、「上記形式+独自情報」と内部管理書式を変えると内部活用でも便利になります。

誰でも一元的に
情報入手が可能



体験学習データでの 共通語彙基盤導入の試行

● 各組織が独自で出していた情報を一元的に提供

– デジタルの時代だからこそ、デジタルの力でリアルの世界を充実

国土交通省
全国の見学施設

農林水産省
全国農林漁業体験スポット一覧
全国工場見学・市場見学一覧

関東農政局
食・農業体験受入れ施設一覧

環境省
探そう自然体験

東京都
伝統工芸品見学・体験工房

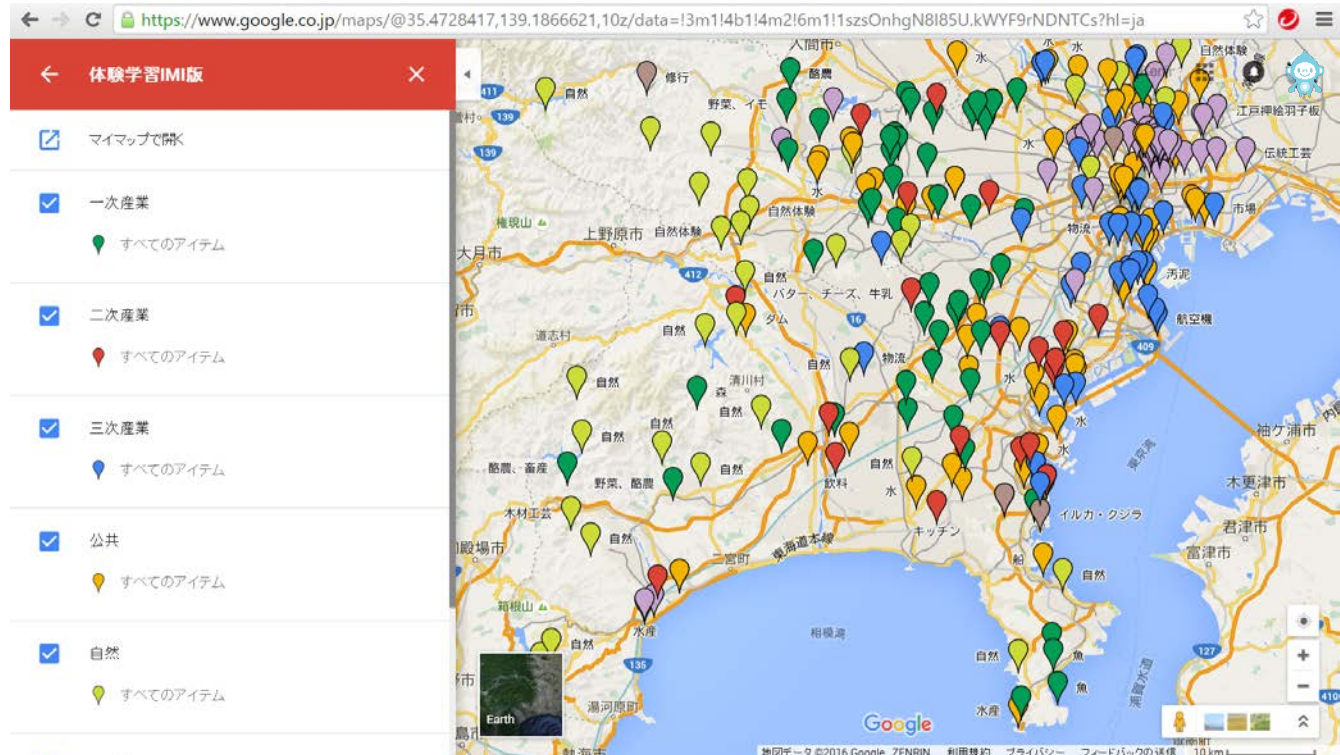
神奈川県
体験・学習プログラム、
体験学習

横浜市
横浜「学習・体験」ガイド

Google、
日本物流団体連合会物流見学ネットワーク
等

IMI
体験イベント
テンプレート

IMI
体験イベント
テンプレート
概要版



現在のデータ提供形式

- 各組織で独自書式で情報を提供

環境省:探そう自然体験

検索先 > こどものページ > こどもの自然体験活動実施団体・組織一覧 > 関東 > 神奈川県

神奈川県

横浜市東区
神奈川県立境川遊水地公園
 川崎市幸区
環境共育事務所 K&Kプランニング
 川崎市麻生区
川崎市恩川青少年野外活動センター
 横須賀市
横須賀市自然・人文博物館
 鎌倉市
ネイチャージガイド 風の道
 茅ヶ崎市
湘南自然学校
 相模原市
神奈川県立津久井湖城山公園

神奈川県立境川遊水地公園
 遊水地から始まる、治水学習・自然体験・貝化石採集
 都市河川の治水効果としての遊水地の役割や、周りの河川や国内ビオトープでの自然を学習できます。
 また、公園工事の副産物として見つかった貝化石の採集体験もでき、貝化石や苔の海と温暖化の役割について学びます。

連絡先
 住所 〒245-0217 神奈川県横浜市東区下飯田町5-5
 電話 045-805-0223 ファクス 045-806-3917
 ウェブサイト <http://www.kankawasea-park.or.jp/kankawasea/index.html>

自然体験活動の受け入れ状況
 低小中学校を受け入れて、自然体験活動を実施している
 受け入れ可能な生徒数、期間、季節
 宿泊型
 なし
 日帰り型
 100名以上を受け入れ可能な人数 100名

農林水産省:全国工場見学・市場見学一覧

種類	名称	所在地	連絡先	詳細
	アサヒビール 洗瓶工場	茨城県守谷市緑一丁目1-1	電話:0297-45-7335 FAX:0297-45-7751	見学の内容はこちら 見学施設のHP(外部リンク)
	カゴメ株式会社 茨城工場	茨城県小美玉市引島2952	電話:0291-46-0021 FAX:0299-95-9103	見学の内容はこちら 見学施設のHP(外部リンク)
	キリンビール 取手工場	茨城県取手市高瀬188-1	電話:0297-72-6000 FAX:-	見学の内容はこちら 見学施設のHP(外部リンク)
	キユーピー株式会社 五霞工場	茨城県鹿嶋市五霞町小手指1800	電話:0280-84-9225 FAX:-	見学の内容はこちら 見学施設のHP(外部リンク)
	トモエ乳業株式会社	茨城県古河市下辺見195	電話:0280-32-1111 FAX:0280-32-4192	見学の内容はこちら 見学施設のHP(外部リンク)

農林水産省:全国農林漁業体験スポット一覧

名称	住所	電話番号	体験対象作物	体験内容	対象者(学年、年代、その他)
青森農園	東京都目黒区青丘1-15-7	042-322-0226	野菜	土づくり、植え付け、栽培管理(草取りを含む)、収穫	高校生、大人、親子
一方さんの畑(農業体験農園)(外部リンク)	東京都練馬区土土田1-34-6		野菜	土づくり、植え付け、栽培管理(草取りを含む)、収穫	大人、親子
市村体験農園	東京都府中市御立町2-40-1	042-363-0461	野菜	土づくり、植え付け、栽培管理(草取りを含む)、収穫	大人
ウエストランドファーム	東京都台東区				
グリーンファームHAGIUM	東京都				
小泉牧場	東京都				

都県名	市町村名	取組団体等の名称	実施主体	区分	対象	取組内容
東京都	練馬区	NPO法人「畑の教室」(白石農園)	NPO・市民団体	農業・食・見学の講習会・その他	親子、小学生、中学生、大人、教職員	・農業体験 ・農作物の収穫体験 ・加工体験 [詳細情報](PDF:26KB)
東京都	日野市	石坂ファームハウス	生産者	農業・食	親子、小学生、大人等	・田植え、稲刈り等の稲作体験 ・野菜、果実等の収穫、調理体験 ・加工体験(みそ、ジャム等) [詳細情報](PDF:24KB)
東京都	東大和市	赤永乳業(株)東京多摩工場	企業	見学	幼稚園・保育園 年長以上	・牛乳、カップ飲料、キャップ付紙パック飲料 製造工程見学 [詳細情報](PDF:23KB)
東京都	中央区ほか	東京都中央卸売市場(都内11箇所)	行政	見学	各市場により設定	・市場内の見学等 [詳細情報](PDF:10KB)

国土交通省:全国の見学施設

海上保安庁にある見学施設

→ 海上保安資料館(広島県呉市)

現役を引退した巡視船、飛行機及びヘリコプターなどの写真、模型、海上保安庁の業務を紹介する写真パネル・模型など約1000点近い展示物が並んでいます。 [海上保安部トップ](#) > [海上保安部について](#) > [海洋情報資料館](#)

→ 海上保安資料館 横浜館(神奈川県横浜市)

館内には、平成13年12月22日に発生した、フナ。 [海洋情報資料館](#)

→ 海洋情報資料館(東京都江東区)

コンピューターがなかったころの海洋調査や海に関する最新の海洋情報業務を紹介するパネルが並んでいます。

海洋情報資料館

○ 海洋情報資料館の公開

平日13時～17時(年末年始を除く) 開館

海洋情報部は明治4年(1871年)の創設以来、航海の安全に必要な海図や潮汐表などを提供するとともに、近年では、海洋環境、地震、津波などに関する社会ニーズに対応し、多種多様な調査・研究を行い、さまざまな海洋に関する情報を提供しています。

海洋情報資料館には、デジタル機器やコンピューターがなかったころ、どのようにして海の深さや流れ、湧き潮・引き潮を推算していたかなど、当時の海洋調査や海図の測量を知ることでる機器、日本で最初に作られた海図などを展示しているほか、最新の海洋情報業務を紹介するパネルを展示しています。

このほか、伊能忠敬の観測図や、海外の古地図も展示してありますので、どうぞお立ち寄り下さい。



共通語彙基盤を活用した集約データイメージ

- データ項目が多くなるので、入力支援ツールが望まれる

データベース3 : データベース- C:\Users¥hiramoto¥Documents¥データベース3.accdb (Access 2016)

Kenji Hiramoto

MS Pゴシック (詳細) 11

テキストの書式設定

ID	イベントID	名称	サブタイトル	要約	説明	webサイト	イベント種別	産業分類	料金	大人料金
10		白洋舎 多摩川 企業見学				http://www.hak				
11		フジテレビ 企業見学				http://www.fuji				
12		NHK(日本放送 企業見学				http://www.nhk				
13		朝日新聞 東京 企業見学				http://www.asahi				
14		全日本空輸 AN 企業見学				http://www.allnippon				
15		警視庁 企業見学				http://www.keisai				
16		参議院 企業見学				http://www.sangiin				
17		サントリー 武蔵 企業見学				http://www.suntory				
18		コカ・コーラー 企業見学				http://www.coca-cola				
19		東京消防庁 本 企業見学				http://www.tfd				
20		池袋防災館(康 企業見学				http://www.tfd				
21		東京消防庁 立 企業見学				http://www.tfd				
22		日本テレビ 企業見学				http://www.ntv				
23		東京貨物ターミ 物流見学ネット				http://www.irfr				
24		物流博物館 物流見学ネット				http://www.lmu				
25		日本通運株式 物流見学ネット				http://www.nittai				
26		(株)丸運 東雲 物流見学ネット				http://www.marun				
27		日本運輸倉庫 物流見学ネット				http://www.nissai				
28		栗林商船(株) 物流見学ネット				http://www.kuribashi				
29		近物レックス(株) 物流見学ネット				http://www.kinokoro				
30		近物レックス(株) 物流見学ネット				http://www.kinokoro				
31		日本ロジテム(株) 物流見学ネット				http://www.logitem				
32		(株)日立物流 物流見学ネット				http://www.hitachi				
33		日本航空(株) 物流見学ネット				http://www.jal.com				
34		日本航空(株) 物流見学ネット				http://www.jal.com				
35		近物レックス(株) 物流見学ネット				http://www.kinokoro				
36		(株)ユニエック 物流見学ネット				http://www.unieck				
37		新潟運輸(株) 物流見学ネット				http://www.niisai				

レコード: 1 / 272

フィルターなし

データシートビュー

データ項目が揃うのでデータ操作が容易

- 産業分類毎に色分けするなど、様々な活用が可能



- データは元々表などで管理されているので、変換は容易
 1. 公開用テンプレートを何にするか決める（IMIを推奨）
 2. 原課から元データを表形式でもらう
 3. 列のデータ内容が合うものを「切り取り」「貼り付け（挿入）」で張り付ける
 4. 元データが細かくてテンプレートに入りきらないものは、備考や自由記述欄に貼り付ける。
（省略してもよい）

データ変換のポイント

- 既に持っているデータ項目だけでよく、無理に項目を埋める必要はない
- 空欄が多くても、気にしない
- 詳細情報は元のデータを参照してもらえばよいと思って気楽に取り組むことが重要

共通語彙基盤(IMI)事業進捗状況

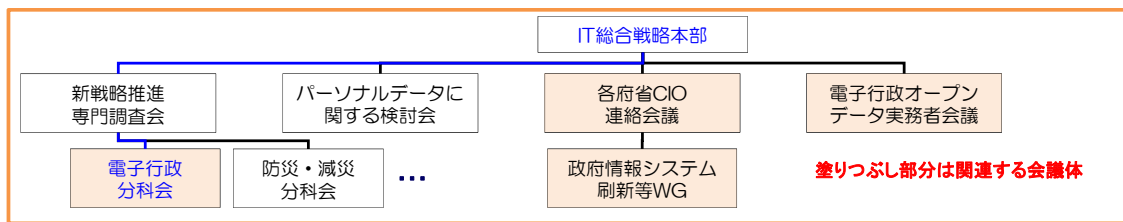
2015年度

情報処理推進機構

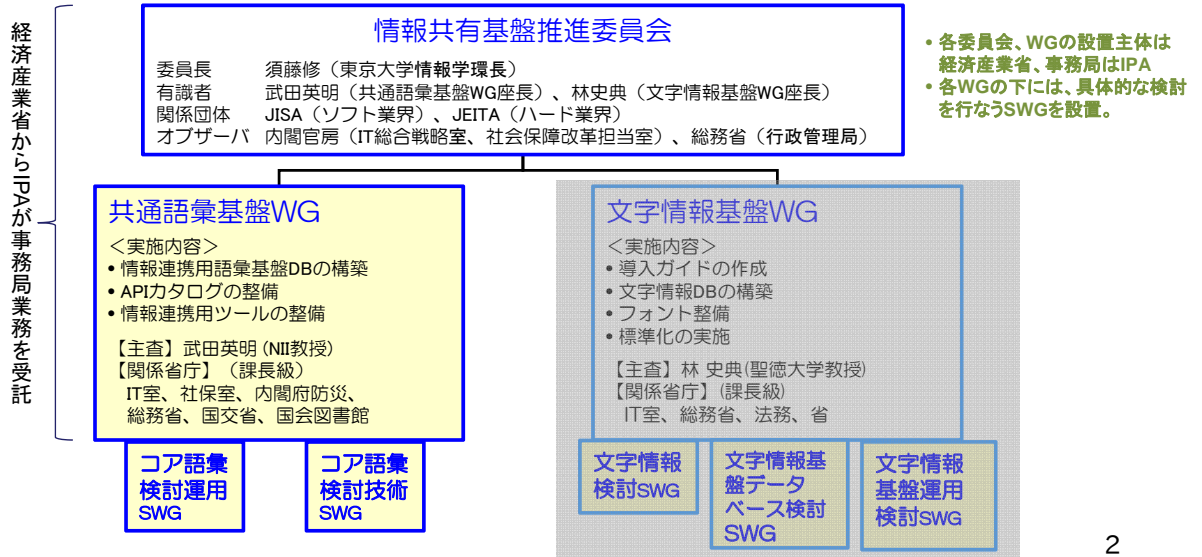
1

検討体制

IPA



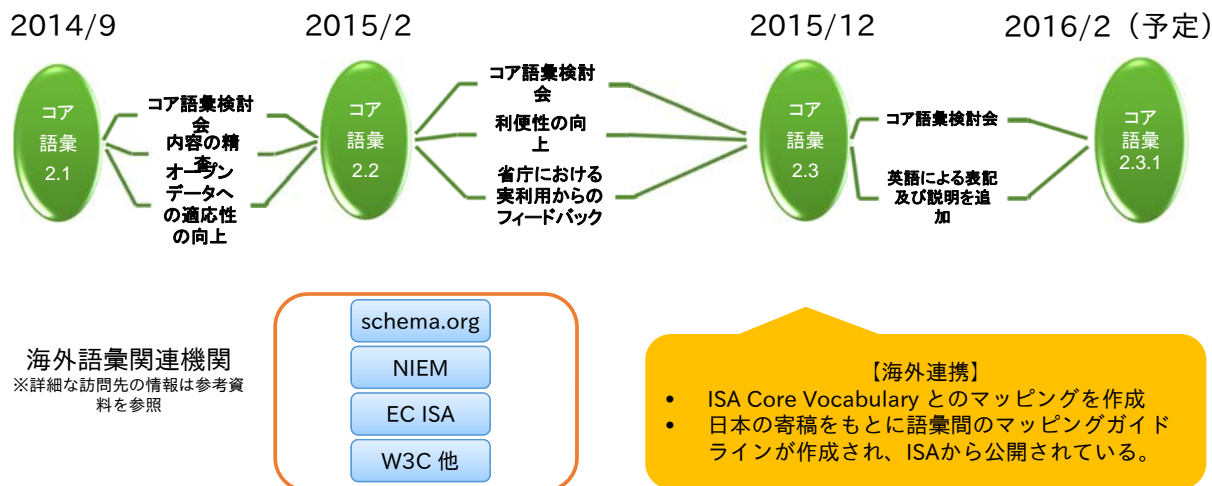
協力依頼 ↓ ↑ 報告



2

「コア語彙」の開発状況

コア語彙2.3では、法人ポータル、農林水産分野の語彙整備など省庁における実利用現場、自治体からのフィードバックを元に語彙の拡充を行った。
さらに、英語による表記や説明を追加したコア語彙2.3.1を近日公開予定。



コア語彙SWG実施状況

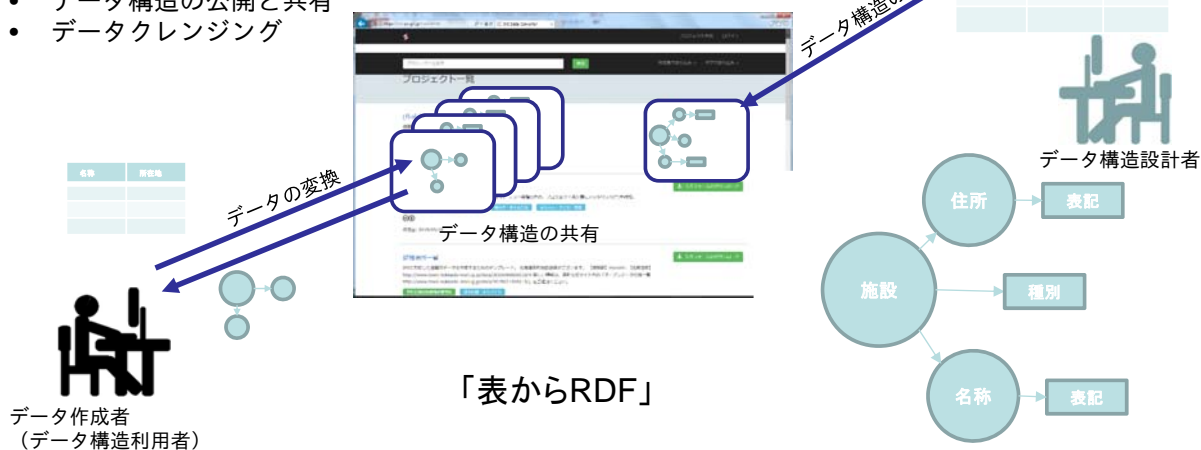
1	2015/4/17	運用	コードリスト整備の優先順位、農水語彙のサポート体制など
2	2015/6/12	運用	語彙関連イベント実施について
3	2015/7/17	運用	METI側事業(農水等)の状況、ドメイン語彙について
4	2015/8/28	運用	IMI共通語彙基盤の課題整理、IMIへの適合(ロゴプラム)について
5	2015/9/18	運用	METI側事業(農水等)の状況、IMIロゴの必要性などについて
6	2015/10/7	運用	METI側事業(農水、法人ポータル等)の状況、IMIロゴ
7	2015/11/9	運用	METI側事業(農水、法人ポータル等)の状況、IMIロゴ
8	2015/12/9	運用	METI側事業(法人ポータル等)の状況、4月以降のイベント実施について
9	2016/1/19	運用	ドメイン語彙(法人ポータル、農業語彙)、語彙利用ツールについて
10	2016/2/12	運用	ドメイン語彙運用に向けて検討
1	2015/3/27	技術	データ交換用XMLスキーマ、データテンプレートの仕様
2	2015/4/10	技術	マッピングテーブル、英語資料、コア語彙2.21進捗確認、RDFIにおける追加制限について
3	2015/4/24	技術	マッピングテーブル及び文書の確認
4	2015/5/15	技術	マッピング再評価
5	2015/5/22	技術	語彙データベース機能検討
6	2015/6/5	技術	ISAへのレスポンス、コア語彙2.2.1について
7	2015/6/19	技術	コア語彙の単位表記について
8	2015/6/26	技術	コア語彙の単位表記について
9	2015/7/10	技術	コードリストの扱いについて
10	2015/7/24	技術	コードリストの扱いについて
11	2015/8/7	技術	コードリストの扱いについて
12	2015/8/21	技術	コードリストの扱いについて
13	2015/9/8	技術	コア語彙2.3、コードリストについて
14	2015/9/29	技術	コア語彙2.3について
15	2015/10/9	技術	コア語彙2.3について
16	2015/10/23	技術	コア語彙2.3について
17	2015/11/13	技術	コア語彙2.3について
18	2015/11/20	技術	コア語彙2.3について
19	2015/12/22	技術	コア語彙2.3の英語表記について
20	2016/1/8	技術	コア語彙2.3の英語表記について
21	2016/1/22	技術	コア語彙2.3.1の確認
22	2016/2/16	技術	コア語彙2.3.1の最終確認とコア語彙2.4.1に向けての課題整理

ツールの公開（2015年7月14日）

自治体職員や自治体のシステムやデータマネージメントを請負う技術者を対象とした共通語彙基盤を用いたデータの作成を支援するツール

- 新たにデータ構造を設計することなく、共有されているデータ構造を利用することで、簡単に構造化されたデータを作成することができる。
- データ構造を作成する場合であっても、表の各項目をコア語彙の用語に関係付けるだけで行うことができ、RDFやXML等の技術に詳しい必要はない。
- このツールにより下記のことを行うことができる。

- ・ コア語彙を用いたデータ構造を新規に作成
- ・ 表形式データを元にデータ構造を作成
- ・ 表形式のデータをコア語彙を用いたデータに変換
- ・ データ構造の公開と共有
- ・ データクレンジング

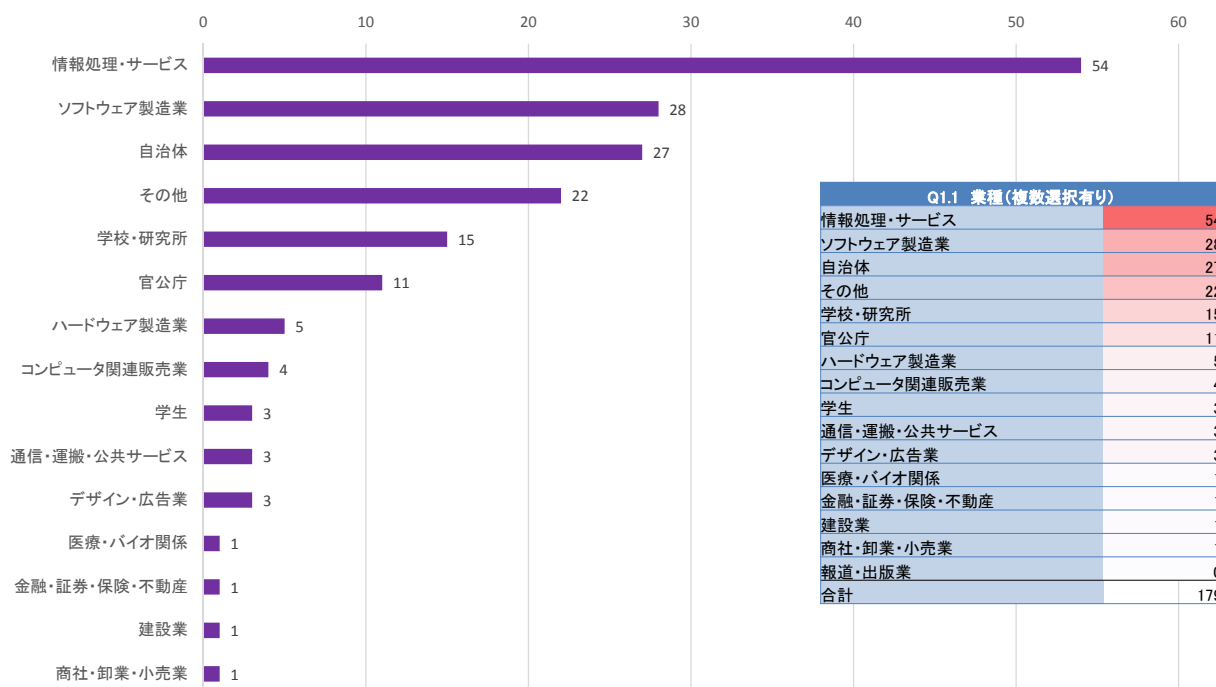


地方説明会の開催

- 自治体オープンデータ推進協議会共催イベント
 - ・ 2015年6月22日 大阪
 - ・ 武田英明氏 共通語彙基盤コア語彙2 (Ver2.2)の紹介
- 情報の価値を最大化する「共通語彙基盤」セミナー～“つながる”データ “つながる”システム～
 - ・ 「共通語彙基盤」の進捗報告と今後に向けた展開
 - ・ 先進自治体での事例を交えて、どのようにして「データ」を活用していくか、「表からRDF」のデモを交えつつ説明
 - ・ 北海道森町 山形 巧哉氏 (札幌会場) 森町のオープンデータの現状と対応策
→ 自治体現場での活用の参考になったと好評

日時	場所	参加者
7月14日	福岡	約60名
9月3日	仙台	約40名
9月4日	札幌	約40名
9月11日	大阪	約40名
9月17日	新潟	約60名

説明会（福岡・仙台・札幌・大阪・新潟）の参加者の業種



ロゴ、キャラクターの制定

2015年12月公開

- 経済産業省が委託事業により二次利用も含めた著作権の譲渡を前程として公募
- 商標登録申請をIPAにて準備中



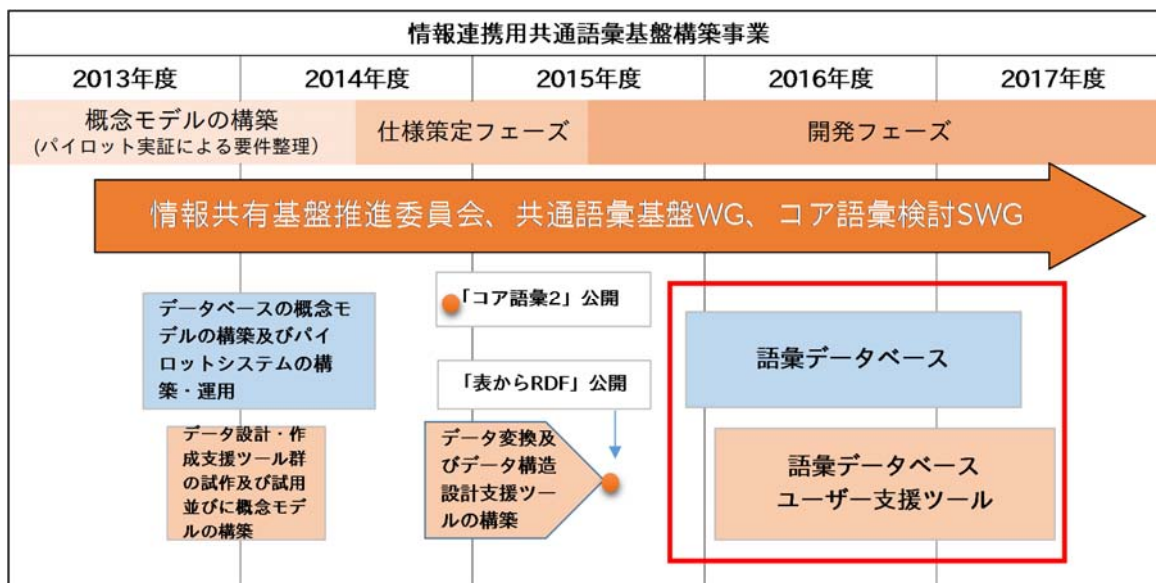
ロゴ

シンプルなフォルムの中に「先進性」「プロフェッショナル性」「グローバル感」が感じられます。「M」の一部が重なることで、人と人のコミュニケーションを連想させ、国際会議の場でも文化の壁を越えた親しみやすさ、わかりやすさを感じられます。



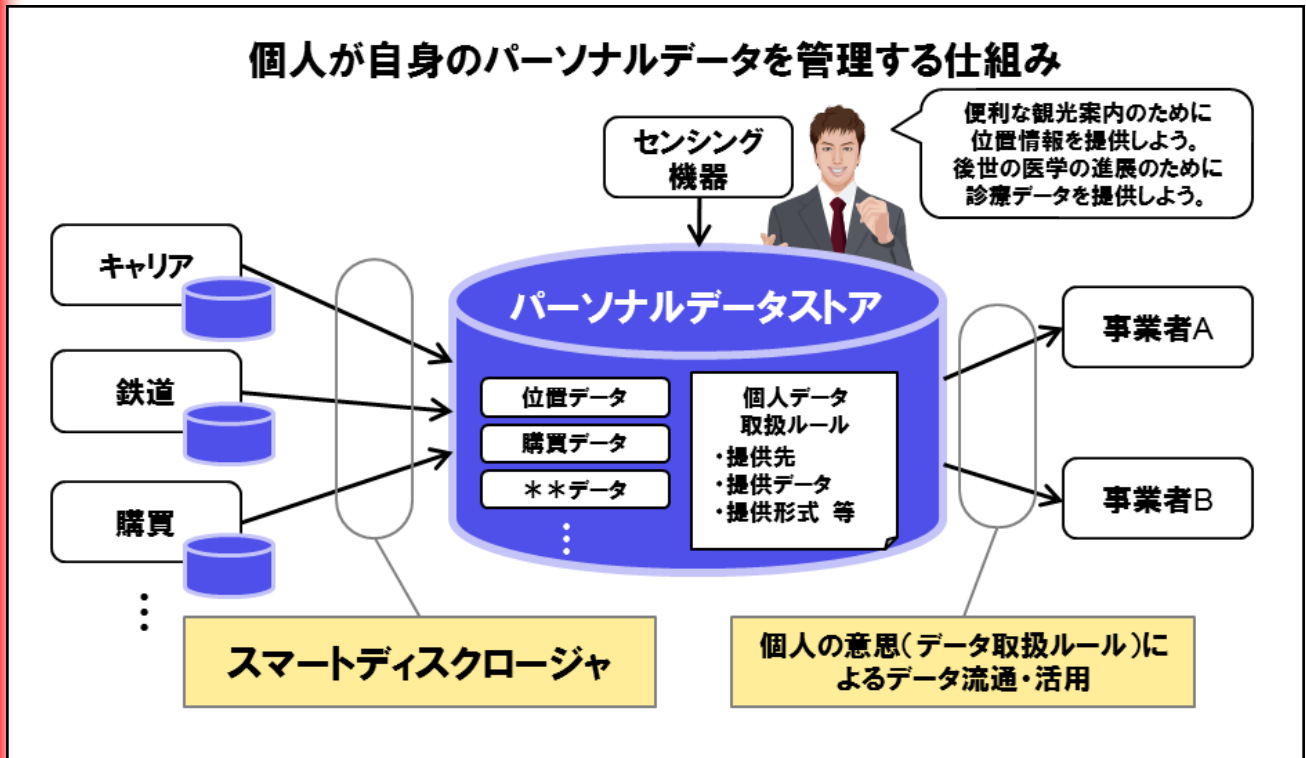
キャラクター

ゴリラをモチーフにしたキャラクターで名前は「ゴイラ」です。ゴリラと語彙を組み合わせたネーミングです。目と鼻がIMIの形になっています。丸い三色のお腹は、コア語彙、ドメイン共通語彙、ドメイン固有語彙を表しています。



- 車両間、車両と公的情報との連携などのため、メタデータやデータ構造の標準化の必要性を認識
- TC204/WG1が、規格に出てくるデータ交換について情報の収集と整理を開始
- 2015年10月に開催されたTC204会議で語彙基盤を紹介
 - 今後の連携可能性を検討開始

個人情報管理の新しい枠組み作りにあたって、語彙の標準化が重要



産業競争力懇談会 「IoT時代におけるプライバシーとイノベーションの両立」 報告書より

データマネジメント2016で講演（予定）

データマネジメント2016

～データ駆動こそがビジネスを創る～

2016年3月11日(金)

目黒雅叙園(東京都)

主催：一般社団法人 日本データマネジメント・コンソーシアム **事前登録制・無料**

C-3

11:40 ~ 12:20

情報の価値を高める「共通語彙基盤 (IMI)」、その必然性を明かす

独立行政法人情報処理推進機構 (IPA)
 技術本部 国際標準推進センター
 参与/国際標準推進センター長
 田代 秀一 氏



「世界最先端IT国家創造宣言」という安倍政権の閣議決定の下、政府や自治体を持つ公共データを再利用性の高い「オープンデータ」として提供する取り組みが始まっています。しかし、それを生かすには前提条件があります。データの構造や項目名などを定義したテンプレートをあらかじめ用意し、共有することにより、効率的で誤りのないデータ交換を可能にする「共通語彙基盤 (IMI)」の利用がその1つです。単に語彙を共通化するだけでなく、IMIを使えばデータ構造を設計する手間やコストを削減でき、活用のためのアプリ開発も容易になります。昨年2月の公開以来、経済産業省が近く公開する「法人情報ポータル」や、いくつかの自治体で活用が始まっていますが、新しいものだけに十分に認知されているとは言えません。そこで本講演では、共通語彙基盤 (IMI)の仕組み、活用の実例、海外での取り組み、国際協力の状況などについて紹介します。IMIを活用し、情報の整理、共有、公開を効率的で価値あるものにしてゆきましょう。

- 欧州委員会主催SEMIC会議（ラトビア）の基調講演で、共通語彙基盤を紹介(2015年5月)
- 欧州各国、米国、日本が参加する Community of Practice on Core Data Models で、各語彙のマッピングについて検討。
 - Guidelines for mapping core data models のドラフトを作成中
- 欧州委員会主催 ISA to ISA2会議（ブリュッセル）で共通語彙基盤を紹介(2016年3月)

13

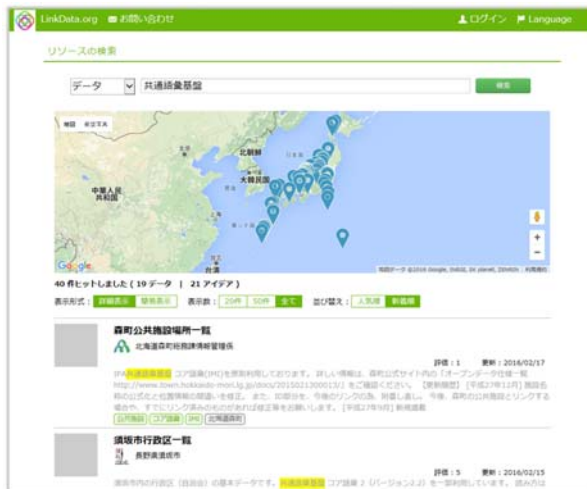
共通語彙基盤を活用する自治体が増加中

- オープンデータの公開にあたり、コア語彙を用いる自治体
 - 北海道森町、北海道八雲町、長野県須坂市など
- オープンデータに関わる業務委託等において、共通語彙基盤を要件とする自治体
 - 大阪市、宮城県など
- その他、自治体内部での情報共有のための利用や、各自治体の広報誌のデータへの適用などの検討が進められている。

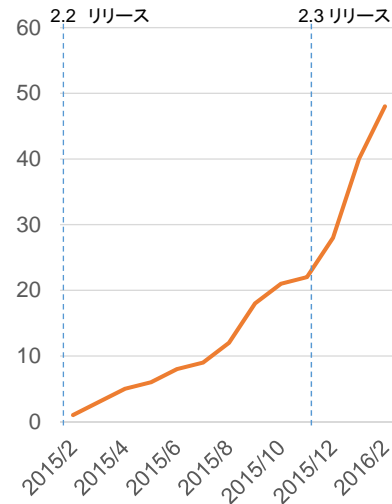


14

LinkData.org (オープンデータの作成・公開を行うサイト)では、40を超える共通語彙基盤を活用したリソース(自治体によるオープンデータ、オープンデータを活用したアプリケーションなど)が登録されている。



LinkData.org



IMI活用リソースの登録数

ツール「表からRDF」の活用状況

活用状況

- 自治体職員、大学関係者、システムインテグレータ等、23名がデータ構造の作成・登録に必要なユーザー登録を行っている
- ユーザー登録者内訳は、自治体職員13名、他10名



2016 年*月**日

政府情報システム刷新等ワーキンググループ
相互運用性検討サブワーキンググループ
設置の提言

情報共有基盤推進委員会

行政サービスの相互運用性確保は、利用者視点のサービス構築や効率的で迅速なサービス構築に必須の条件である。そのため、経済産業省は関係府省の参加により情報共有基盤推進委員会を設置し、文字情報基盤と共通語彙基盤の整備を図ってきた。これらの項目は、電子行政分野におけるオープンな利用環境整備に向けたアクションプランの中でも中核的な推進項目とされている。

これまでは、基盤構築の技術的検討が中心であったため経済産業省および情報処理推進機構で検討を行ってきたが、文字情報基盤は国税庁の法人番号付与システムで活用され、共通語彙基盤は、法人ポータル（仮称）、で活用され農林水産省で検討行われているなど政府内での導入フェーズに入ってきている。他府省からの導入等についての問合せも増えてきていることから、政府内での検討のフェーズに入ってきていることから、政府情報システム刷新等ワーキンググループの下に相互運用性検討サブワーキンググループを設置し、相互運用性に関する検討を強化するとともに、各府省担当者が導入経験や課題の共有を図る場を設置することを提言する。

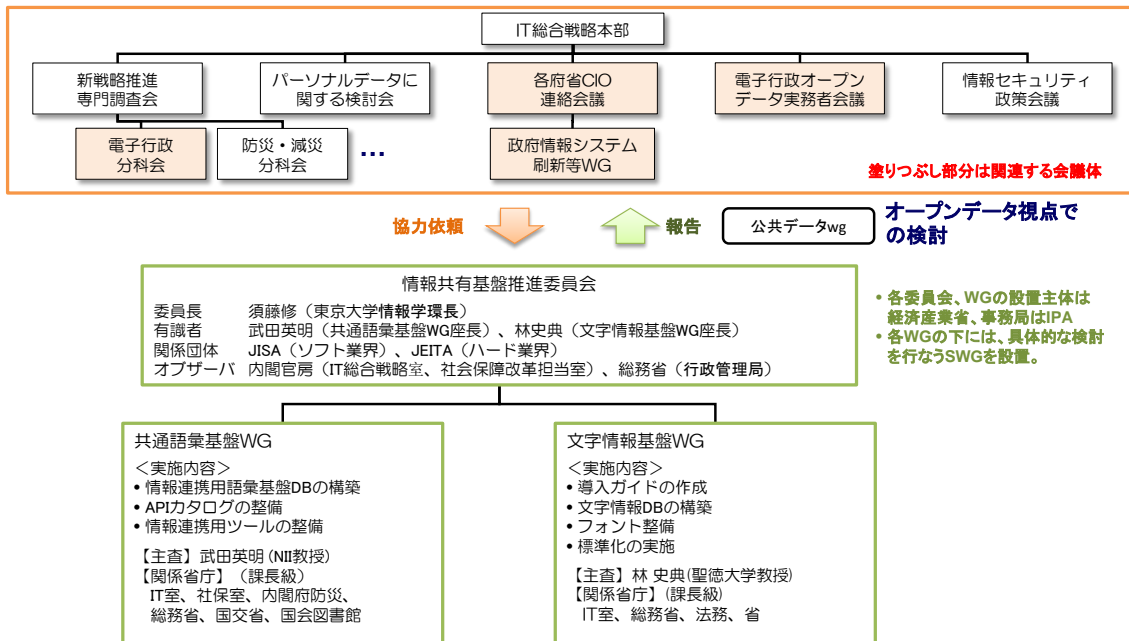
提言内容

- ・ 電子行政刷新ワーキンググループ相互運用性検討サブワーキンググループの設置。
- ・ サブワーキンググループ参加者は、データ基盤の構築等、相互運用性の向上に興味のある府省担当者。
- ・ 開催頻度は四半期から半期に 1 回程度の開催とする。
- ・ 事務局は、内閣官房情報通信技術（IT）総合戦略室が経済産業省および情報処理推進機構の支援を受けて行う。

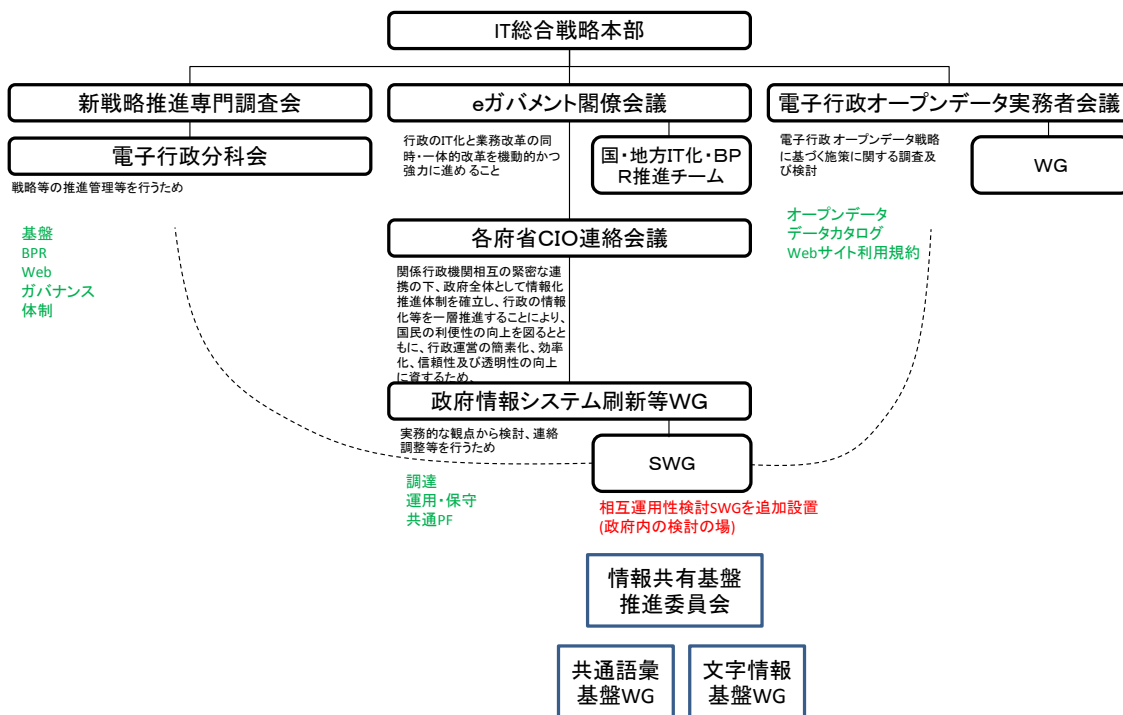
以上

参考

これまでの体制



今後の体制



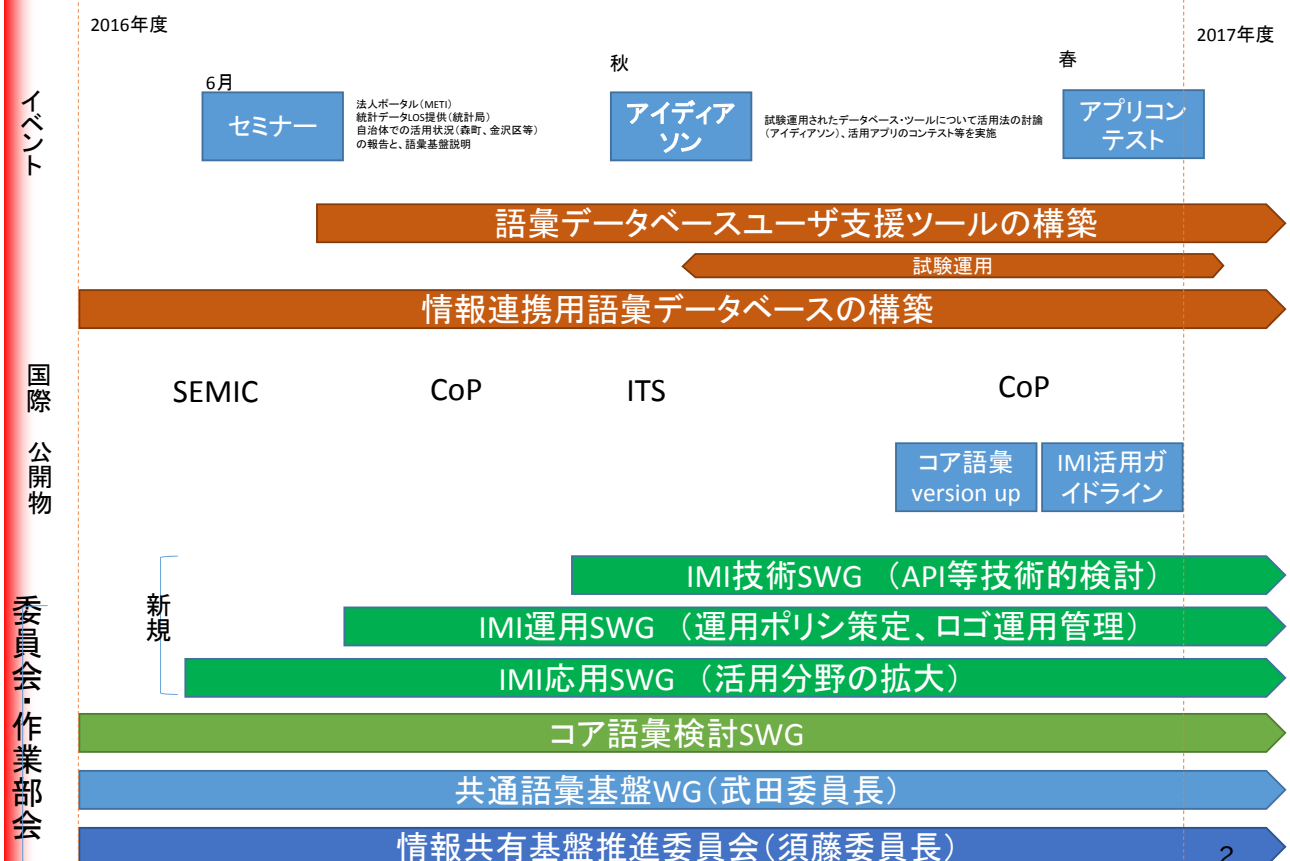
共通語彙基盤(IMI)事業方針

2016年度

情報処理推進機構

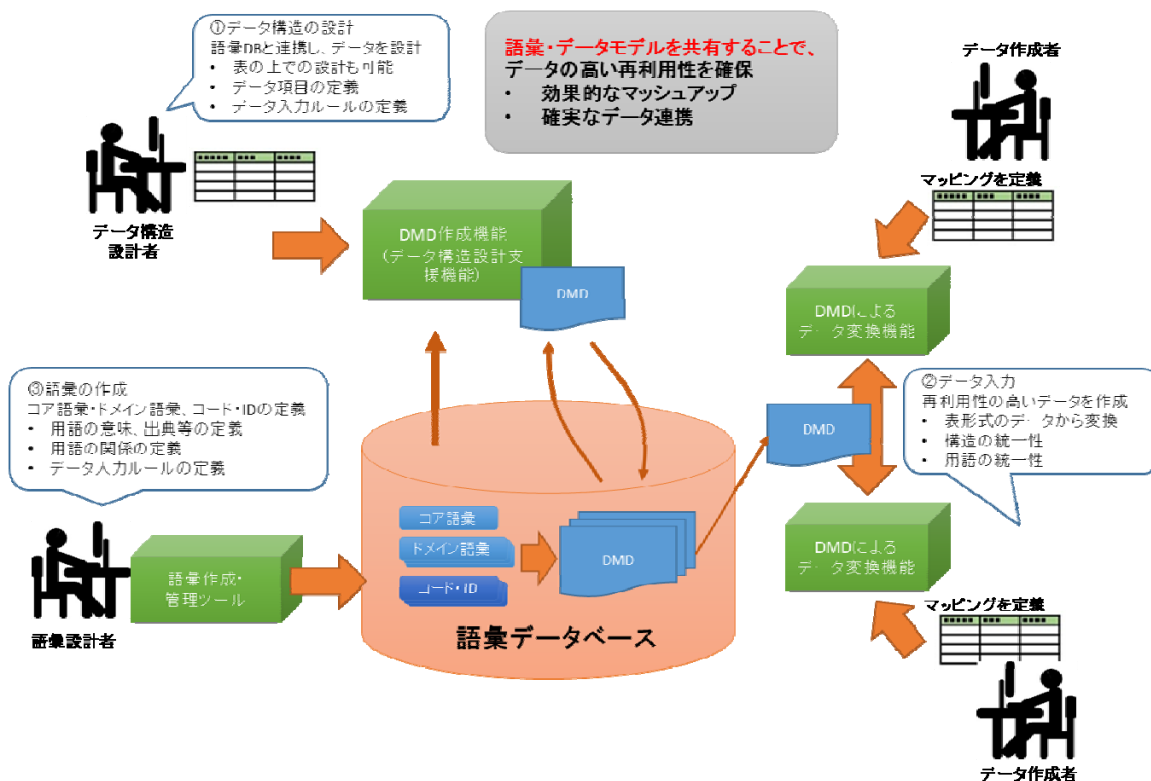
1

共通語彙基盤2016年度案

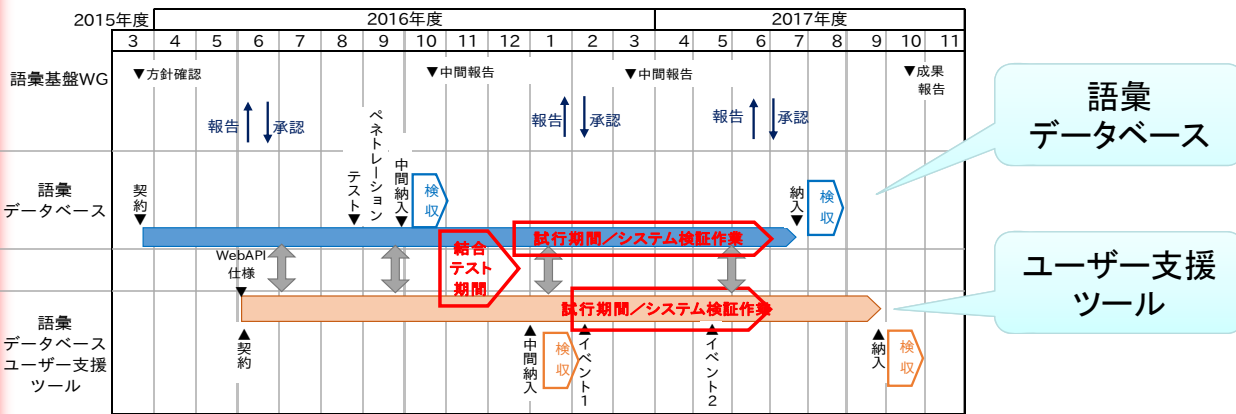
情報連携用基本語彙 データベースの構築

共通語彙基盤の全体イメージ

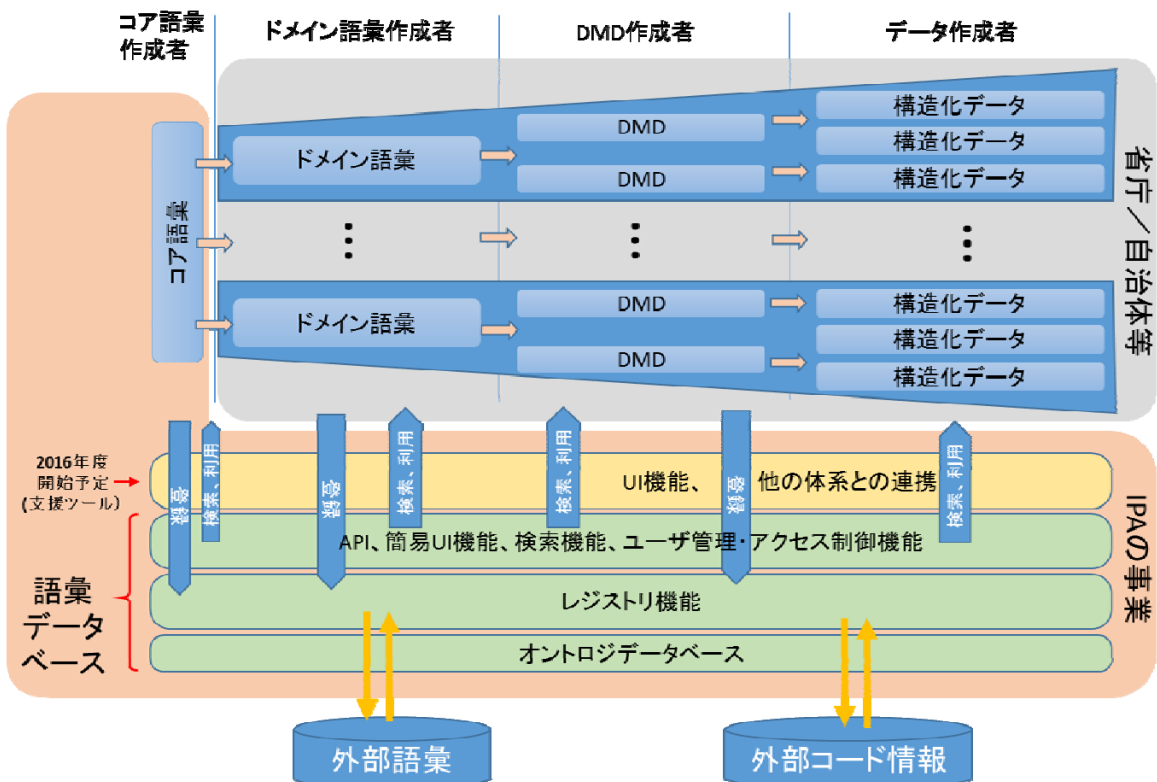


共通語彙基盤構築事業スケジュール IPA

共通語彙基盤
全体スケジュール



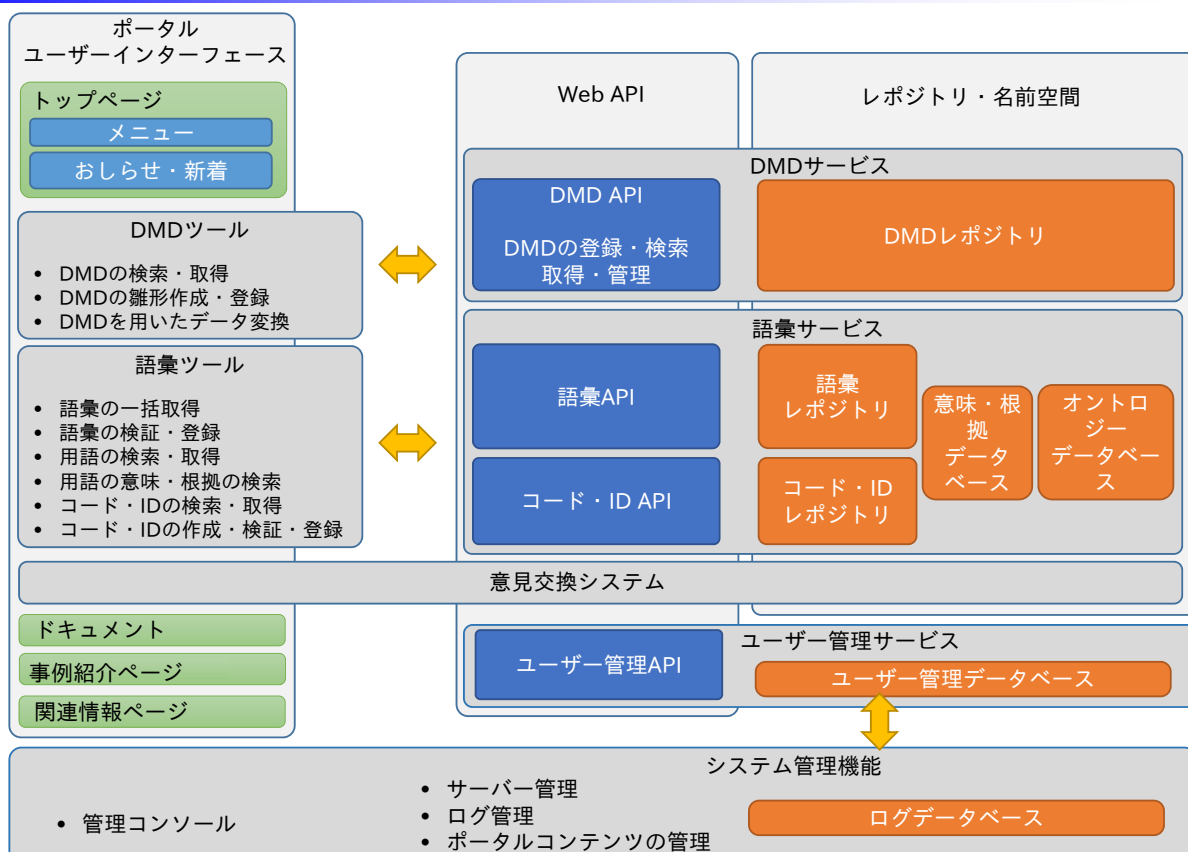
語彙DBと支援ツールの全体イメージ IPA



項目	内容
システム開発	<ul style="list-style-type: none"> ポータル構築(メニュー、お知らせ、事例紹介) 名前空間の整備(語彙、コード・ID) データベースの構築(意味/根拠、オントロジー) レポジトリの構築(DMD、語彙、コード・ID) 管理機能(ユーザー管理、ログ採取)
コンテンツ整備	<ul style="list-style-type: none"> 公開中のIEP(情報交換パッケージ)に必要な情報を加え、DMDとして整備する コア語彙に関する以下のデータを、自治体や政府機関関係者などの利用者および利用シーンを想定した上で収集整理する。情報の範囲や内容の妥当性の確認等のため、委員会を設置し意見を求める。 <ul style="list-style-type: none"> ✓ 意味・根拠データ、オントロジーデータ ✓ コード、コードリスト、ID型、ID体系
ドメイン語彙運用検討	<ul style="list-style-type: none"> 特定の業務領域で用いられるドメイン語彙の運用方法について、有識者による委員会を設置し検討を行う。
システム検証	<ul style="list-style-type: none"> 仮運用を行い、別途実施される「支援ツール」と接続し試用運用を行う。 仮運用中、上記ツール関係者、コア語彙関係者及び受注者自身などによりシステムの検証を行い、将来の活用促進に向けた提言を報告書としてまとめる。

7

システム開発（機能ブロック図）



8

構造化したデータ項目のスキーマや項目の値の制限、項目の内容を説明したドキュメントなど、データ項目を共有するために必要な情報をひとまとめにしたパッケージ。機械向けの情報と人間向けの情報の双方を持つ。

DMDヘッダー XML版スキーマ例

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.w3.org/XML/1998/namespace" schemaLocation="http://www.w3.org/XML/1998/namespace" >
  <xsd:import namespace="http://www.w3.org/XML/1998/namespace" schemaLocation="http://www.w3.org/XML/1998/namespace" />
  <xsd:element name="DMD">
    <xsd:complexType base="xsd:string">
      <xsd:sequence>
        <xsd:element ref="dmd:URI" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Name" minOccurs="1" maxOccurs="unbounded"/>
        <xsd:element ref="dmd:CreationDate" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd>LastModificationDate" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd:Description" minOccurs="1" maxOccurs="unbounded"/>
        <xsd:element ref="dmd:Publisher" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd:License" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Previous" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Latest" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Component" minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

ドキュメント例

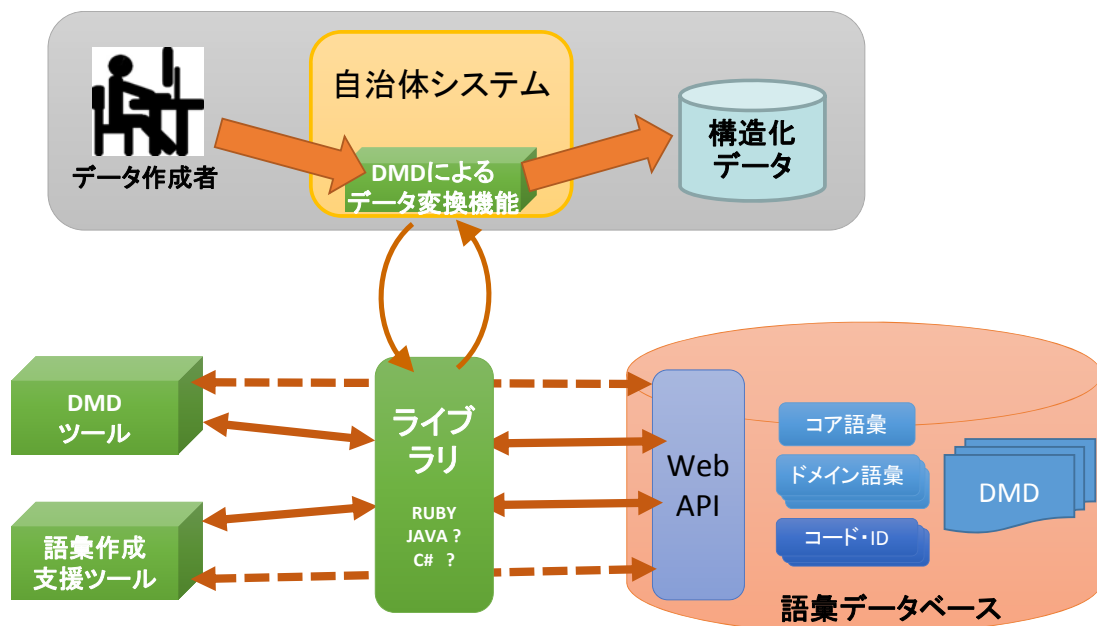
2. データ項目

データ項目と記述イメージは、以下の通りである。灰色項目はその組み合わせることで、アプリケーションで自動生成できる項目であるデータを使用しないで、直接記入することも可能である。

データ項目名	説明	備考
正式名称	「株式会社〇〇」のような組織分類と組織名を組み合わせた正式名称または登記名称。	組織分類名か
組織分類位置	株式会社等の場合には、組織名の前後のどちらに付与するかを記載。	
組織分類	株式会社、有限会社、合名会社、合資会社、合同会社等の組織分類を記載。	
組織名	組織分類を除いた組織名を記載。農業協同組合のように、組織名と組織分類が一体不可分の場合には一体として記載。	
略称アルファベット	略称のアルファベットがある場合には、半角英数字で記載。	

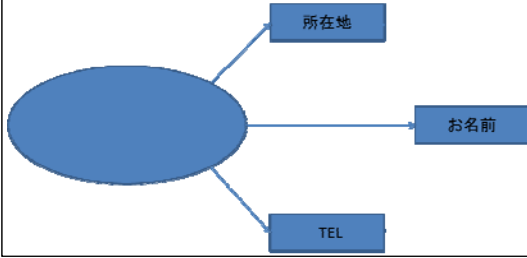
語彙データベース ユーザ支援ツールの開発

項目	内容
ツール開発 (DMDツール)	<ul style="list-style-type: none"> ◆DMD検索ツール <ul style="list-style-type: none"> • ID、名前、概要、作成者、キーワード等による絞り込み検索 • 一覧表示、グラフ構造ビュー、表形式ビュー • データ変換 ◆DMDエディタ <ul style="list-style-type: none"> • 構造化データの設計支援 • スキーマの生成、ドキュメントの自動生成 ◆DMDビューワ <ul style="list-style-type: none"> • 構造化データをグラフ表示
ツール開発 (語彙作成支援ツール)	<ul style="list-style-type: none"> • コア語彙やドメイン語彙の作成者向けのツール • 概念の整理の支援機能 • 概念の可視化(グラフ、表形式、集合) • 登録用テンプレート
ライブラリ開発	<ul style="list-style-type: none"> • 外部ツールから共通語彙基盤DBを使いやすくするためのライブラリを整備。
イベントの実施	<ul style="list-style-type: none"> • 自治体/省庁、利用事例紹介 • IMIの認知度向上のためのイベント(検討中) <ul style="list-style-type: none"> ✓ DMDコンテスト ✓ アプリコンテスト



Step.1 CSVをインポート

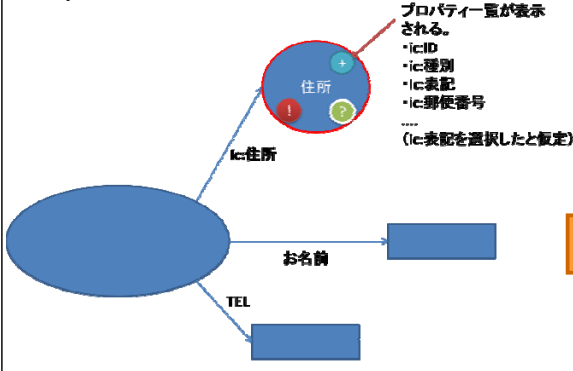
所在地	お名前	TEL
神奈川県横浜市	山田 太郎	123-4567-890
東京都千代田区	田中 花子	234-5678-901



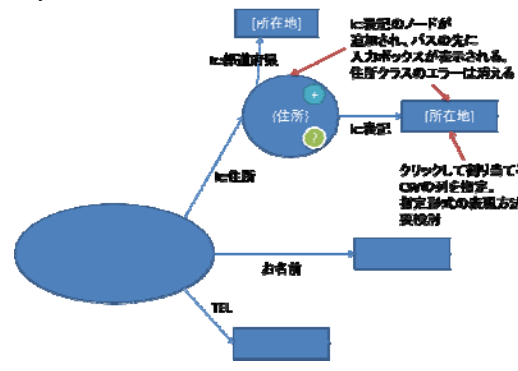
Step.2 クラスを選択



Step.3 プロパティを追加



Step.4 プロパティにデータを入力



13

DMD(Data Model Description)

構造化したデータ項目のスキーマや項目の値の制限、項目の内容を説明したドキュメントなど、データ項目を共有するために必要な情報をひとまとめにしたパッケージ。機械向けの情報と人間向けの情報の双方を持つ。

ドキュメント例

2. データ項目

データ項目と記述イメージは、以下の通りである。灰色項目はその組み合わせることで、アプリケーションで自動生成できる項目であるデータを使用しないで、直接記入することも可能である。

データ項目名	説明	備考
正式名称	「株式会社〇〇」のような組織分類と組織名を組み合わせた正式名称または登記名称。	組織分類名か
組織分類位置	株式会社等の場合には、組織名の前後のどちらに付与するかを記載。	
組織分類	株式会社、有限会社、合名会社、合資会社、合同会社等の組織分類を記載。	
組織名	組織分類を除いた組織名を記載。農業協同組合のように、組織名と組織分類が一体不可分の場合には一体として記載。	
略称アルファベット	略称のアルファベットがある場合には、半角英数字で記載。	

DMDヘッダー XML版スキーマ例

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema" targetNamespace="http://www.w3.org/XML/1998/namespace" >
  <xsd:import namespace="http://www.w3.org/XML/1998/namespace" schemaLocation="http://www.w3.org/XML/1998/namespace.xsd"/>
  <xsd:element name="DMD">
    <xsd:complexType base="xsd:string">
      <xsd:sequence>
        <xsd:element ref="dmd:URI" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Name" minOccurs="1" maxOccurs="unbounded"/>
        <xsd:element ref="dmd:CreationDate" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd>LastModificationDate" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd:Description" minOccurs="1" maxOccurs="unbounded"/>
        <xsd:element ref="dmd:Publisher" minOccurs="1" maxOccurs="1"/>
        <xsd:element ref="dmd:License" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Previous" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Latest" minOccurs="0" maxOccurs="1"/>
        <xsd:element ref="dmd:Component" minOccurs="0" maxOccurs="unbounded"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

コア語彙 2.3 リリースノート

コア語彙 2.2 からの変更点

独立行政法人情報処理推進機構

2015 年 12 月 11 日

1. はじめに

2015年2月3日に公開したコア語彙 2.2 に対する各方面からのご意見を踏まえ、データ表現の容易性や汎用性の向上のため、以下の変更を行いました。

2. 用語の変更

2.1. 概要

本バージョンは、マイナーバージョンアップであり、基本的にコア語彙を利用して作成された既存のデータと互換性があります。

主な変更点：

- 「単位コード」を新たに導入するなど、重量や面積などの数量の単位をより容易に表現できるようになりました。
 - 単位は、量の種別に関わらず「単位コード」を利用するようになりました。
 - コードが存在しないような単位のために「単位表記」プロパティが追加されました。
- 「期間」と「スケジュール」が統合され、「期間」を「スケジュール」として利用できるようになりました。この変更にもとない、下記の変更が行われました。
 - 「スケジュール」は「期間スケジュール」に項目名が変更されました。
 - データ型が「スケジュール」となっていたプロパティは、データ型が「期間スケジュール」に変更されました。
 - 継承が「スケジュール」となっていたクラス用語の継承が「期間スケジュール」に変更されました。
- クラス用語の追加、クラス用語へのプロパティの追加などの拡張を行い、コア語彙を用いてより広い範囲のデータを表現できるようになりました。
- コア語彙をより汎用的に利用できるようにするため、クラス用語、プロパティ用語の変更を行いました。

2.2. クラス用語の追加

下記の4つクラス用語が追加されました。

単位コード

識別子:

ic:単位コード型

継承:

[ic:コード型](#)

説明:

コードを表現するためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:量種別	量種別	xsd:string	0..1	量の種別
ic:乗数	乗数	xsd:double	0..1	基本単位に対する変換乗数

土地

識別子:

ic:土地型

継承:

[ic:地物型](#)

説明:

施設や建物など、参照対象となる地物を表すためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:種別	種別	xsd:string	0..1	識別するための ID
ic:種別コード	種別コード	ic:コード型	0..n	概要説明。100 文字以内
ic:面積	面積	ic:面積型	0..n	土地の面積

活動

識別子:

ic:活動型

継承:

[ic:事物型](#)

説明:

イベントを表現するためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:ID	ID	ic:ID 型	0..n	イベントを識別するための ID
ic:種別	種別	xsd:string	0..n	イベント内容の種別
ic:名称	名称	ic:名称型	0..n	イベントの名称
ic:関与	関与	ic:関与型	0..n	イベントに関連する組織
ic:場所	場所	ic:場所型	0..n	イベントの開催場所
ic:日時	日時	ic:日時型	0..n	イベントの期間
ic:期間	期間	ic:スケジュール型	0..n	イベントの期間
ic:説明	説明	xsd:string	0..n	イベントへの連絡方法

関与

識別子:

ic:関与型

継承:

[ic:事物型](#)

説明:

関係する人や組織を表現するためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:役割	役割	xsd:string	0..1	人や組織の位置づけや役割
ic:役割コード	役割コード	ic:コード	0..1	人や組織の位置づけや役割のコード
ic:関与者	関与者	ic:実体型	0..n	関係する人や組織

2.3. プロパティの追加

下記の表のように、多くのクラス用語にプロパティを追加しました。下記の表では、追加されたプロパティのみを記載しています。

コード			
項目名	データ型	回数	説明
表記	xsd:string	0..n	コードの値に対応する概念の表記
上位コード	ic:コード型	0..n	このコードの上位のコード
下位コード	ic:コード型	0..n	このコードの下位のコード
関連コード	ic:コード型	0..n	このコードに関連するコード

組織			
項目名	データ型	回数	説明
活動	活動型	0..n	組織の活動
関連人員	ic:人数型	0..n	関連する人員の人数(従業員数など)

地物			
項目名	データ型	回数	説明
関与	関与型	0..n	地物に関連する組織や人

金額			
項目名	データ型	回数	説明
乗数	xsd:double	0..1	金額の値に対する変換乗数

状況			
項目名	データ型	回数	説明
関与	関与型	0..n	状況に関連する組織や人
説明	xsd:string	0..1	状況の説明
期間	ic:期間スケジュール型	0..n	状況の期間

日時			
項目名	データ型	回数	説明
表記	xsd:string	0..1	日時の任意の表記

日付			
項目名	データ型	回数	説明
表記	xsd:string	0..1	日付の任意の表記

数量			
項目名	データ型	回数	説明
単位	単位コード型	0..1	単位を表すコード
単位表記	xsd:string	0..1	単位の文字列による表現(助数詞など)

容量			
項目名	データ型	回数	説明
単位	単位コード型	0..1	単位を表すコード
単位表記	xsd:string	0..1	単位の文字列による表現(助数詞など)

面積			
項目名	データ型	回数	説明
単位	単位コード型	0..1	単位を表すコード
単位表記	xsd:string	0..1	単位の文字列による表現(助数詞など)

重量			
項目名	データ型	回数	説明
単位	単位コード型	0..1	単位を表すコード
単位表記	xsd:string	0..1	単位の文字列による表現(助数詞など)

長さ			
項目名	データ型	回数	説明
単位	単位コード型	0..1	単位を表すコード
単位表記	xsd:string	0..1	単位の文字列による表現(助数詞など)

2.4. その他の変更

- プロパティ用語「時間」のデータ型を「数量型」に変更しました。「アクセス」及び「アクセス区間」の2つのクラス用語が「時間」プロパティを利用しています。

2.5. 利用が推奨されない用語及びプロパティ

コア語彙 2.2 の用語のうち下記の用語及びプロパティは、コア語彙 2.3 では、より利用しやすいものへと変更されています。これらの用語及びプロパティは互換性のためにコア語彙 2.3 においても残されています。将来の版では削除される可能性がありますので、これらの用語を利用することは推奨いたしません。

<http://imi.ipa.go.jp/ns/core/2> より参照される XML スキーマには、これらの用語は含まれていません。これらの用語を使用しなくてはならない場合は、<https://imi.ipa.go.jp/> からダウンロード可能な schema_xsd.zip の中の、core23bc.xsd を XML スキーマとして参照してください。

- 利用が推奨されないクラス用語
 - 「測定値」(クラス用語「数量」を使用してください。)
 - 「時間」(クラス用語「数量」を使用してください。)
- 利用が推奨されないプロパティ
 - 「業務組織」の「構成員数」(「組織」の「関連人員」プロパティを使用してください。)
 - 「法人」の「従業員数」(「組織」の「関連人員」プロパティを使用してください。)
 - 「数量」の「数量単位」(「単位コード」プロパティを使用してください。)
 - 「容量」の「容量単位」(「単位コード」プロパティを使用してください。)
 - 「面積」の「面積単位」(「単位コード」プロパティを使用してください。)
 - 「重量」の「重量単位」(「単位コード」プロパティを使用してください。)
 - 「長さ」の「長さ単位」(「単位コード」プロパティを使用してください。)
 - 「状況」の「関係者」(「関与」プロパティを使用してください。)
 - 「状況」の「詳細」(「説明」プロパティを使用してください。)

SC118DI07171

Guidelines for mapping core data models



Document Metadata

Property	Value
Date	2015-10-19
Status	For acceptance
Version	0.08
Authors	Alexandru Droscariu – PwC EU Services
Reviewed by	Nikolaos Loutas – PwC EU Services Susanne Wigard, European Commission
Approved by	

This document was prepared for the ISA Programme by:

PwC EU Services

Disclaimer:

The views expressed in this report are purely those of the authors and may not, in any circumstances, be interpreted as stating an official position of the European Commission.

The European Commission does not guarantee the accuracy of the information included in this study, nor does it accept any responsibility for any use thereof.

Reference herein to any specific products, specifications, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favouring by the European Commission.

All care has been taken by the author to ensure that s/he has obtained, where necessary, permission to use any parts of manuscripts including illustrations, maps, and graphs, on which intellectual property rights already exist from the titular holder(s) of such rights or from her/his or their legal representative.

Table of Contents

1	INTRODUCTION	4
2	KEY FACTS ABOUT MAPPINGS	5
2.1	MAPPING CONCEPTUAL DATA MODELS	5
2.2	MAPPING RELATIONS	6
2.2.1	<i>Exact match</i>	7
2.2.2	<i>Close match</i>	7
2.2.3	<i>Related match</i>	7
2.2.4	<i>Narrow match</i>	8
2.2.5	<i>Broad match</i>	9
3	STEPS TO CREATE MAPPINGS.....	10
3.1	CONCEPTUAL LEVEL MAPPINGS	10
3.1.1	<i>Example for "Has exact match"</i>	10
3.1.2	<i>Example for "Has close match"</i>	11
3.1.3	<i>Example for "Has related match"</i>	11
3.1.4	<i>Example for "Has narrow match"</i>	12
3.1.5	<i>Example for "Broad match"</i>	12
3.2	SYNTAX LEVEL MAPPINGS.....	13
3.3	SPECIAL CASES	14
3.3.1	<i>Composite mapping</i>	16
3.3.2	<i>Partial mapping</i>	16
3.3.3	<i>Conditional mapping</i>	17
3.3.4	<i>Partial and conditional mapping</i>	18

1 INTRODUCTION

The current document complements the guidelines for creating mappings between the Core Vocabularies and other data models initiatives, which are documented in the “Handbook for using the Core Vocabularies”¹ (henceforth referred to as “the handbook”).

The objective is to support the effort of developing data model mappings. To do so, we provide a few examples of cases that are likely to happen when creating mappings, and which we encountered while mapping the Core Vocabularies with other data models, such as the Swedish Company data model, NIEM 3.0, KoSIT – XÖV, OSLO, and etc. We explain how we selected the mapping relations in these special cases. During the development of the mappings, we followed the method described in the handbook and we documented the mappings in a specially crafted spreadsheet. The mappings are published on the Core Data Model Mapping Directory (CDMMD)².

The current document is structured as follows. It first provides an introduction into the work of mapping data models explaining the needs behind. It continues with an overview of the structure of conceptual data models, the mapping relations, and the basic rules to apply the mapping relations. The document provides a description of the steps to apply mapping relations, and it closes providing several special cases from the experience of mapping the Core Vocabularies with the Swedish Company data model. We sprinkled the document with examples to highlight the reasoning behind applying the mapping relations.

¹ https://joinup.ec.europa.eu/site/core_vocabularies/Core_Vocabularies_user_handbook/Handbook-for-using-the-Core-Vocabularies_v0.50.pdf

² <http://mapping.semic.eu>

2 KEY FACTS ABOUT MAPPINGS

The handbook provides two reasons for creating mappings between data models:

- For documenting purposes, when the mappings at the conceptual level offer an entry point to domain experts to understand the mapped data models;
- For reconciliation of data sources, when mappings are used to extract basic information from conformant data models.

The syntax level mapping is a necessary input for the implementation of procedures for reconciliation of data sources. The conceptual level mapping supports the syntax level mapping. Therefore, conceptual level mapping is a prerequisite for syntax level mapping.

Mappings should be created between elements of two distinct data models only when the links are inherent in the meaning of the linked elements³.

The mapping relations between the elements of two data models are independent of the syntax used to describe/represent the two data models.

Data reconciliation can be done without loss of meaning only between elements that have an exact match relationship. For the elements that have a broad or narrow relationship, the reconciliation can be done from the more specific element to the more generic element. Reconciling elements that have a close relationship involves the risk of losing meaning.

2.1 Mapping conceptual data models

Conceptual level mappings are developed between conceptual data models. The structure of a conceptual data model is composed of classes which describe real or abstract things such as buildings, people, organisations or services performed by a legal entity. Classes are composed of properties and/or associations. The difference between a property and an association is the type of values that may be assigned. A property (e.g. PersonFamilyName is a property of the Person class of the ISA Core Vocabularies) takes a singular value of a specific data type. An association (e.g. PersonAddress is an association of the Person class of the ISA Core Vocabularies) takes an instance of another class. The conceptual model does not specify how properties and associations are technically represented. For example, an association may refer to another object with some kind of identifier (e.g. a URI in RDF) or may embed the whole object (e.g. nesting in XML).

When developing conceptual level mappings, we refer to classes, properties and associations, generally, as elements. To decide the mapping relation to map two elements, consider the following aspects:

- Map classes to classes, properties to properties, and associations to associations.

³ <http://www.w3.org/TR/skos-reference/#mapping>

- Consider the meaning of the two elements, regardless of:
 - The data structure or the properties used to describe the two elements.
 - The way each of the two elements are identified (e.g. "BE" VS "http://data.europa.eu/languages/BE" for the language code).
 - The way the two elements encode literals (e.g. "1990-06-09" VS "09/06/1990" for the birthdate).

2.2 Mapping relations

The mapping relations used for the conceptual level mappings are borrowed from SKOS⁴: exact, close, related, narrow, and broad match.

The handbook offers a comprehensive description of the mapping relations and their application. This chapter provides a summary explaining the meaning of the mapping relations, and how they should be applied.

The mapping relations are defined in terms of the set of subjects covered by the elements⁵. The figure below, extracted from the handbook, provides a graphical representation of the mapping relations.

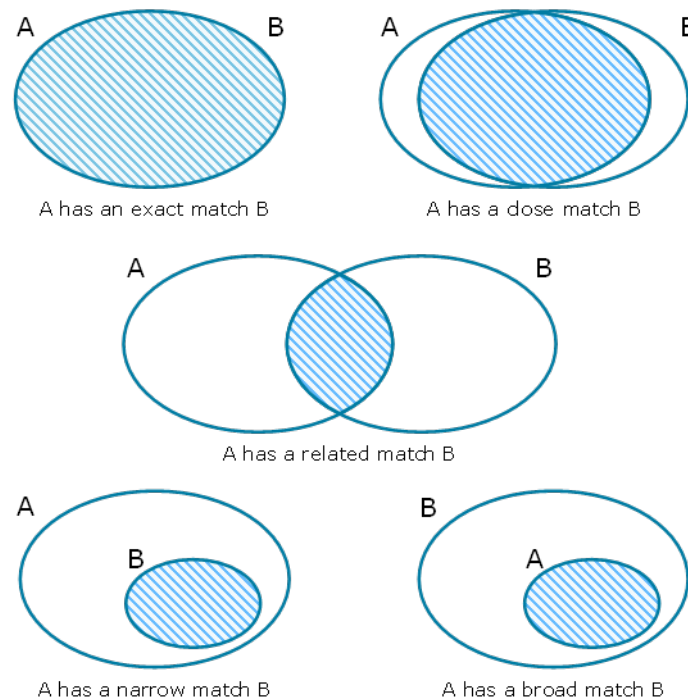


Figure 1 - Mapping relations: defined in terms of the set of subjects covered by the elements

The following examples will be further explained in chapter 3.

⁴ <http://www.w3.org/TR/skos-reference>

⁵ https://joinup.ec.europa.eu/site/core_vocabularies/Core_Vocabularies_user_handbook/Handbook-for-using-the-Core-Vocabularies_v0.50.pdf

2.2.1 Exact match

We say that “A has an exact match B” if the set of subjects of A is equal to the set of subjects of B. The definitions of A and B are equivalent.

Facts that are expressed for elements (classes, properties, or relationships) with an exact match relationship can be converted in both directions between data models without loss of meaning.

For example:

Core element	Vocabularies	Mapping relation	Swedish Company data model
LegalEntity		Has exact match	Juridisk person

2.2.2 Close match

We say that “A has a close match B” if the set of subjects of A is mostly equal to the set of subjects of B. The number of subjects of A not included in B, and vice-versa, is negligible.

Elements with a close match relationship can be converted in both directions between data models with a minimal loss of meaning for some individuals.

For example:

Core element	Vocabularies	Mapping relation	Swedish Company data model
AddressLocatorName		Has close match	belägenhetsadress gårdsadressområde
Description: The locator name is a proper noun applied to the real world entity identified by the locator. The locator name could be the name of the property or complex, of the building or part of the building, or it could be the name of a room inside a building. [INSPIRE] The key difference between a locator designator and a locator name is that the latter is a proper name and is unlikely to include digits.			Description: "Farm Address area" can be used to specify the location or to preserve historical names. "Farm Address area" should be used where there are names of farms, collections of buildings or a single house that is listed on a general map and where these names are used in the address today.

2.2.3 Related match

We say that “A has a related match B” if there is a meaningful intersection between the subjects of A and the subjects of B.

Elements with a related match relationship can only be converted with considerable error. Conversion is not advised. Such mappings can still be valuable to make semantic conflicts between data models better visible.

For example:

Core element	Vocabularies	Mapping relation	Swedish Company data model
AgentPlaysRole		Has related match	Relation till person
Description: A public service in which the agent plays a role.			<p>Relation till person (relationship to person = Role)</p> <p>Description:</p> <p>Relationships between persons are necessary to define the person. Here is for example information on who are the representatives of legal entities. It is also through this information that the address of a CEO can be linked to the limited company he/she represents.</p> <p>For example: a natural person authorized to sign for the legal entity.</p>

2.2.4 Narrow match

We say that “A has a narrow match B” if the set of subjects of A is a superset of the set of subjects of B. The definition of A generalizes the definition of B.

Facts that are expressed for elements (classes, properties, or relationships) with a narrow/broad match relationship can only be converted into only the direction of the more general data model element.

For example:

Core element	Vocabularies	Mapping relation	Swedish Company data model
LegalEntityLocation		Has narrow match	Säte
Description: Asserting the Location relationship implies only that the legal entity has some connection to a location in time or space. It does not			<p>Description:</p> <p>The city in Sweden where the Board (of the legal entity) will meet and where the general meeting is to be held.</p>

imply that the legal entity is necessarily at that location at the time when the assertion is made.

2.2.5 Broad match

We say that “A has a broad match B” if B has a narrow match A.

For example:

Core Vocabularies element	Mapping relation	Swedish Company data model
LegalEntityRegisteredAddress	Has broad match	Postadress
Description: In almost all jurisdictions, legal entities must register a postal address. This may or may not be the actual address at which the legal entity does its business - it is commonly the address of their lawyer or accountant - but it is the address to which formal communications can be sent.		Description: The postal address of a Legal entity, natural person or workplace.

3 STEPS TO CREATE MAPPINGS

3.1 Conceptual level mappings

To formally define the mapping relations, we need to distinguish between an element and its subjects, i.e., the real-world things that are represented by the instances of the element.

An element of a data model is a class, a property, or an association. Classes are realised as instances (individuals) representing a real-world physical or conceptual thing. Classes are used to express facts about the generic characteristics of an individual instance. Properties or associations are used to construct facts about individuals.

To map data model A to data model B, one should proceed following the steps below:

For each element in A, find an element in B for which, potentially, there can be a relationship. Understand the definitions of the two elements to be mapped, and use your best intuition to find how the two sets of facts described by each of the two elements intersect:

1. If the two sets are equal, there is an **exact match** relation between the two elements;
2. If the two sets are not equal but they intersect,
 - 2.1. Analyse the hierarchy between the two sets, and decide if broad or narrow match can be used:
 - 2.1.1. If the set of facts described by the element in A includes the set of facts described by the element in B, then the element in A has a **narrow match** the element in B;
 - 2.1.2. Similarly, if the set of facts described by the element in A is included in the set of facts described by the element in B, then the element in A has a **broad match** the element in B;
 - 2.2. If it the set of facts described by the element in A does not include or is not included in the set of facts described by the element in B, decide if the intersection between the two sets is significant:
 - 2.2.1. If the intersection is significant, the element in A **has close match** the element in B;
 - 2.2.2. If the intersection is not significant, the element in A **has related match** the element in B;
3. If the two sets do not intersect, the element in A **has no match** in B.
4. Provide an explanation every time when there is a special case of matching between the two elements, to support the development of syntax mappings and the implementation of data reconciliation procedures.

The following chapters provide some example to showcase the steps above for each of the mapping relations.

3.1.1 Example for “Has exact match”

We choose the element **Address** (an address representing a location) from the Core Vocabularies and the element **Postaddress** (contact option that unambiguously defines an actual area for mail delivery, usually linked with an indication of the addressee or recipient, and in some cases a particular address source) from the Swedish Company data model.

We then apply the mapping algorithm:

1. The two elements describe the same set of facts, and therefore **Address** *has exact match* **Postaddress**.

3.1.2 Example for “Has close match”

We choose the element **AddressLocatorName** (the locator name is a proper noun applied to the real world entity identified by the locator. The locator name could be the name of the property or complex, of the building or part of the building, or it could be the name of a room inside a building. [INSPIRE] The key difference between a locator designator and a locator name is that the latter is a proper name and is unlikely to include digits.) in the Core Vocabularies and the element **belägenhetsadress gårdsadressområde** (“Farm Address area” can be used to specify the location or to preserve historical names. “Farm Address area” should be used where there are names of farms, collections of buildings or a single house that is listed on a general map and where these names are used in the address today) in the Swedish Company data model.

We then apply the mapping algorithm:

1. The two elements describe different concepts: **AddressLocatorName** could be the name of a building, room inside of a building, while **belägenhetsadress gårdsadressområde** describes farms or collections of buildings. Therefore, the set of facts described by the two elements are not equal.
2. The two elements describe similar facts: names of properties, buildings, collections of buildings. Therefore, *there is an intersection*.
 - 2.1. There is no hierarchy between the two sets of facts, as **AddressLocatorName** can be the name of a room inside of a building, while the **belägenhetsadress gårdsadressområde** refers only to the entire buildings.
 - 2.2. The two sets of facts described by **AddressLocatorName** and **belägenhetsadress gårdsadressområde** can be considered as significant. Therefore, **AddressLocatorName** *has close match* **belägenhetsadress gårdsadressområde**.

3.1.3 Example for “Has related match”

We choose the element **AgentPlaysRole** (a public service in which the agent plays a role.) from the Core Vocabularies and the element **Relation till person** (translated as relationship to person, or role, and described as: relationships between persons are necessary to define the person. The “Relation till person” provides for example information on who are the representatives of legal entities. It is also through this information that the address of a CEO can be linked to the limited company he/she represents. For example: a natural person authorized to sign for the legal entity) from the Swedish Company data model.

From the definitions of the two elements, we understand that (1) the **AgentPlaysRole** describes the relationship between *an agent* which can be either a natural person, or an organisation and *a public service* for which this agent plays a role, and (2) the **Relation till person** describes the relationship between two *persons* (natural persons or organisations).

We then apply the mapping algorithm:

1. It is obvious that the two elements describe sets of facts which are not equal.
2. There is an *intersection* between the two sets of facts, as both elements can describe roles played by persons.
 - 2.1. There is no hierarchy between the two sets of facts, as **AgentPlaysRole** describes a public service in which the agent (organisation or natural person) plays a role, while **Relation till person** refers to the relationship between two persons (physical or juridical).
 - 2.2. Decide if the intersection between the two sets of facts described by the two elements is significant:
 - 2.2.1. While the **AgentPlaysRole** describes a public role played by a person or an organisation the, **Relation till person** describes the relationship (both public or private) between two persons. Therefore, the intersection is not significant.
 - 2.2.2. Therefore, **AgentPlaysRole** *has related match* **Relation till person**.

3.1.4 Example for “Has narrow match”

We choose **LegalEntityLocation** (asserting the Location relationship implies only that the legal entity has some connection to a location in time or space. It does not imply that the legal entity is necessarily at that location at the time when the assertion is made) from the Core Vocabularies and **Säte** (the city in Sweden where the Board (of the legal entity) will meet and where the general meeting is to be held) from the Swedish Company data model.

We then apply the mapping algorithm:

1. Obviously the two sets of facts described by **LegalEntityLocation** and **Säte** are not equal.
2. Given the fact that **LegalEntityLocation** can describe cities, we can say that the two sets of facts intersect.
 - 2.1. While the **LegalEntityLocation** describes a location which can be a city, the **Säte** always describes a city as a location. Therefore, there is a hierarchy between the two elements.
 - 2.1.1. The set of facts described by **Säte** is always a set of cities where general meetings of companies are to be held. The set of facts described by **LegalEntityLocation** are locations related somehow with the legal entity. Therefore, the set of facts described by the **LegalEntityLocation** can include the set of facts described by **Säte** (a **Säte** is always a **LegalEntityLocation**, but not every **LegalEntityLocation** is a **Säte**). **LegalEntityLocation** *has narrow match* **Säte**.

3.1.5 Example for “Broad match”

We choose **LegalEntityRegisteredAddress** (in almost all jurisdictions, legal entities must register a postal address. This may or may not be the actual address at which the legal entity does its business - it is commonly the address of their lawyer or

accountant - but it is the address to which formal communications can be sent) from the Core Vocabularies and **Postaddress** (the postal address of a Legal entity, natural person or workplace) from the Swedish Company data model.

We then apply the mapping algorithm:

1. The set of facts described by **LegalEntityRegisteredAddress** and **Postaddress** are not equal as **Postaddress** can describe as well the address of natural persons or workplaces, while **LegalEntityRegisteredAddress** describes the postal address of legal entities.
2. Obviously, the set of facts described by the two concepts intersect.
 - 2.1. The set of facts described by the **LegalEntityRegisteredAddress** always describes the postal address of legal entities, while **Postaddress** describes the postal address not only for legal entities, but also for natural persons. There is a hierarchy.
 - 2.1.1. The set of facts described by **LegalEntityRegisteredAddress** can't include the set of facts described by **Postaddress**.
 - 2.1.2. The set of facts described by **LegalEntityRegisteredAddress** can be included in the set of facts described by **Postaddress** (every **LegalEntityRegisteredAddress** is a **Postaddress**, but not every **Postaddress** is a **LegalEntityRegisteredAddress**). Therefore, **LegalEntityRegisteredAddress** *has broad match* **Postaddress**.

3.2 Syntax level mappings

The syntax level mappings enable the exchange of information between IT systems that model semantically similar concepts, represented either with the same (e.g. XML), or with different data formats (e.g. XML and RDF). The representation of the syntax level mappings depends on the data formats of the source and target systems of the data transfer procedure. For example, XSL will be the choice to express syntax level mappings between a source representing its data as XML and a target representing its data as XML or RDF.

The existence of conceptual level mappings is a prerequisite to create syntax level mappings. When developing syntax level mappings, one should look not only at the mappings relations but at the descriptions provided together with the mapping relations. There can be information that further supports the development of syntax level mappings and the implementation of business logic to allow data manipulation at run time. For example, all the values of the attributes of the **Postaddress** concept in the Swedish Company data model could be concatenated so that they can be transferred to the **Address.FullAddress** of the Core Location Vocabulary, despite the fact that at the concept level mapping the **Address.FullAddress** has no match within the Swedish Company data model. Another example is the transfer of data from **Identitetsbeteckning för person.id-beteckning** to **Person.PersonDateOfBirth** of the Core Person Vocabulary. It would be possible to develop a rule at the syntax level mapping which tests for the value of **id-beteckning typ** for an instance of **Identitetsbeteckning för person**, and when this value would be equal to *personnummer*, extract the first 6 to 8 characters from the value of

Identitetsbeteckning för person.id-beteckning and transfer the result into the **Person.PersonDateOfBirth**.

The following steps should be performed to develop syntax level mappings:

1. Choose the representation of the syntax mapping based on the representations of the of the source and destination data models;
2. Follow the mappings at the conceptual level and implement transformation rules which will receive as input the elements in the source data model and will produce as outputs the mapped elements in the destination data model. The transformation rules should implement the business logic described in the spreadsheet listing the mapping relations.

There are various tools⁶ available for the development of transformations between many standard formats, including XML, RDF, JSON, etc.

XSLT⁷ is a language to process and transform XML documents into other formats such as RDF, JSON, etc.

Openrefine⁸ offers powerful functionalities for developing data reconciliation rules for multiple data formats.

The Core Data Model Mapping Directory (CDMMD)⁹ publishes conceptual level mappings between ISA Core Vocabularies and other data models such as NIEM 3.0¹⁰, Stelselcatalogus¹¹, KoSIT - XÖV¹², Swedish Company data model¹³, etc.

3.3 Special cases

While creating mappings between elements of the ISA Core Vocabularies¹⁴ and data models from national initiatives, a few special situations have been encountered that require a less intuitive approach for choosing the appropriate relationship. We noticed in several cases that a high degree of subjectivity can be introduced by the one who chooses the relationships. As higher the degree of subjectivity, as higher to potential for less consistent, core data models alignment.

⁶ <http://www.w3.org/wiki/ConverterToRdf>

⁷ <http://www.w3.org/TR/xslt>

⁸ <http://openrefine.org/>

⁹ <http://mapping.semic.eu/>

¹⁰ <https://www.niem.gov/technical/Pages/version-3.aspx>

¹¹ www.stelselcatalogus.nl

¹² <http://www.xoev.de/de/produkte>

¹³ <http://uppgiftskrav.bolagsverket.se/gu>

¹⁴ https://joinup.ec.europa.eu/asset/core_vocabularies/description

The special cases which are described in the following tables, were encountered while mapping the Core Vocabularies with the Swedish Company data model¹⁵. They are however generic enough and it is very likely to come across them in other contexts as well.

In each of the special situations, together with the selected mapping relation, a description is provided to explain the reasoning behind the chosen solution. The mapping relations and the description should provide enough information to support the development of syntax level mappings, and the implementation of data reconciliation procedures.

Several special cases have been identified:

1. **Composite mapping.** This special case deals with elements from one data model (destination) for which a formula can be applied, and the result would match exactly an element in the source data model. For example, the period of time element in the Core Vocabularies matches exactly with the calculated difference between start date and end date from the Swedish Company data model.
2. **Partial mapping.** This case deals with elements of one data model (source) which can be composed by a combination of elements within the destination data model. For example, the full address element from the Core Vocabularies is a combination of multiple elements from the Swedish Company data model (more details, in the following chapters).
3. **Conditional mapping.** This special case deals with the elements of one data model (source) which match elements of the destination data model only when certain conditions are met. For example, the identifier of a legal organisation in the Core Vocabularies matches the identifier of a "person" in the Swedish Company data model, only for the instances of "person" which are legal organisations. The Swedish Company data model uses the "person" class to denote physical persons and juridical persons. The difference is made by the value of the property "type of person" which can take one of two values.
4. **Partial and conditional mapping.** This case is a combination of both situations described above. For example, the birth date in the Core Vocabularies matches exactly the first 6-8 digits extracted from the value of the identifier of a person in the Swedish Company data model, only when the type of person equals with "physical person"; otherwise, the type of person would identify a legal organisation.

The examples described in the next sections are structured as follows:

1. The column "Concepts of ISA Core Vocabularies": lists the concepts which potentially match one or a combination of concepts of the Swedish Company data model;

¹⁵ <http://uppgiftskrav.bolagsverket.se/gu/ConceptModel/288841c9-624f-47bb-a007-9dbd21f2675b.htm>

2. The column "Potentially match concepts of the Swedish Company data model": lists the concepts of the Swedish Company data model that potentially match the concepts of the ISA Core Vocabularies;
3. The column "Explanation": describes how the concepts could be matched, and what the chosen relation was. A similar description is made available in the spreadsheet documenting the mappings. This description should support the development of the syntax level mappings and the implementation of data reconciliation procedures;
4. The column "The selected mapping relation": lists the conceptual mapping relation that was finally chosen and documented in the spreadsheet.

3.3.1 Composite mapping

Composite mapping describes the context in which several elements in the destination data model can be combined according to a set of business rules (a formula can be applied) and the result matches exactly one element in the source data model.

The existence of such context, does imply that each distinct element in the destination data model will match another element in the source data model.

Concepts of ISA Core Vocabularies	Potentially match concepts of the Swedish Company data model	Explanation	The selected mapping relation
<p>PeriodOfTime</p> <p><i>Is defined as "An interval of time that is named or defined by its start and end dates".</i></p>	<p>Namn på fysisk person namn. namn giltigt från</p> <p><i>Defined as "name valid from"</i></p> <p>Namn på fysisk person namn. namn giltigt till</p> <p><i>Defined as "name valid to"</i></p>	<p>The set of facts described by the PeriodOfTime potentially matches the set of facts described by the calculated difference between, generally, a start date and an end date. The start and end date are properties for several of the existing concepts in the Swedish Company data model such as namn giltigt från and namn giltigt till.</p>	<p>PeriodOfTime</p> <p>Has no match</p>

3.3.2 Partial mapping

Partial mapping describes the context in which one element of a data model (source) matches (no necessarily exactly) a combination of several elements in the destination data model.

Concepts of ISA Core Vocabularies	Potentially match concepts of the Swedish Company data model	Explanation	The selected mapping relation
Address. AddressPOBox	Postaddress. utdelningsadress2	<p>A certain sequence of characters as part of the value of utdelningsadress2 determines the post box value, according to SS 61 34 01.</p> <p>The set of facts described by the AddressPOBox and the set of facts described by the utdelningsadress2 will never intersect.</p>	Address. AddressPOBox Has no match
Address. AddressFullAddress	Postaddress	<p>The concept AddressFullAddress matches a combination of attributes of the Postaddress (such as postort, postnummer, and land) concept. But the AddressFullAddress does not match close enough one single attribute of Postaddress. Therefore, the set of facts described by the AddressFullAddress does not intersect with the set of facts described by any of the attributes of Postaddress.</p>	Address. AddressFullAddress Has no match

3.3.3 Conditional mapping

Conditional mapping describes the context in which the meaning of a certain element (let's call it EL1) in one data model depends on the value of another element (EL2) in the same data model. In such context, EL1 matches exactly one element in the mapped data model only when EL2 has certain values.

Concepts of ISA Core Vocabularies	Potentially match concepts of the Swedish Company data model	Explanation	The selected mapping relation
Person. PersonIdentifier <i>Is defined as "A formally-issued"</i>	Identitetsbeteckning för person <i>Is defined as "a unique identifier that identifies a"</i>	PersonIdentifier matches exactly Identitetsbeteckning för person <i>only if id-beteckning typ</i> , which is a property of the Identitetsbeteckning för person has the value	PersonIdentifier Has broad match Identitetsbeteckning för person

<i>identifier for the person"</i>	<i>particular person"</i>	<i>Personnummer.</i> Therefore, the set of facts described by the PersonIdentifier is equal to the set of facts described by the Identitetsbeteckning för person only in certain conditions. Thus, we can say that the concept Identitetsbeteckning för person is more generic than PersonIdentifier .	
LegalEntityLegalIdentifier <i>Is defined as "The identifier given to the legal entity by the authority with which it is registered"</i>	Identitetsbeteckning för person	Similarly with the case above, the two concepts match, only if id-beteckning typ has one of the values <i>organisationsnummer, samordningsnummer, or GD-nummer.</i> Therefore, Identitetsbeteckning för person is a generalisation of LegalEntityLegalIdentifier .	LegalEntityLegalIdentifier Has broad match Identitetsbeteckning för person

3.3.4 Partial and conditional mapping

Partial and conditional mapping describes a context in which the meaning of a certain element (EL1) in one data model depends on the value of another element (EL2) in the same data model, but EL1 matches partially one element in the destination data model. EL1 matches partially one element in the destination data model, means that a business rule can be applied on the value of EL1 and the result matches the element in the destination data model.

Concepts of ISA Core Vocabularies	Potentially match concepts of the Swedish Company data model	Explanation	The selected mapping relation
Person. PersonDateOfBirth <i>Defined as "The date on which the person was born"</i>	Identitetsbeteckning för person. id-beteckning	PersonDateOfBirth matches id-beteckning only if id-beteckning typ has the value: <i>personnummer.</i> When id-beteckning typ has the value: <i>personnummer</i> , then the first 6 or 8 digits of the value of id-beteckning represent the date of birth, which will be encoded with the format: <i>yyymmdd-nnnn.</i> Therefore, the set of facts described by the	Person. PersonDateOfBirth Has no match

		PersonDateOfBirth never intersect with the set of facts described by the id-beteckning .	
--	--	--	--


3.3.5 Mapping relations in special cases

In all the examples of special cases described in the previous chapters, we observed that the mapping relations do not always express the fact that by applying additional logic to the mapping relations, we could further describe that a transfer of data is possible between two IT systems implementing the two mapped data models. For example, in the case of composite matching, we conclude that the **PeriodOfTime** has no match in the Swedish company data model as the set of facts described by the **PeriodOfTime** does not intersect with a set of facts described by any concept in the Swedish Company data model. However, the difference between the existing properties **namn giltigt från** and **namn giltigt till** is an exact match of the **PeriodOfTime**. It could be interesting to express this kind of match. It could be interesting to further enrich the set of mapping relations so that logic could be used to express the relation between concepts.

ツール「表からRDF」のご紹介


<https://imi.ipa.go.jp/tools/0051/>

データはあるが、
どのような形式
で公開するのが
良いだろうか？




名前	所在地	

データは公開し
ているが、十分
に活用されてい
るだろうか？




お名前	現住所	

既存のデータ構
造は変更できな
い

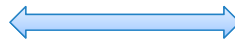


住所	氏名	

既存のシステム
は変更できない。



名前	所在地	

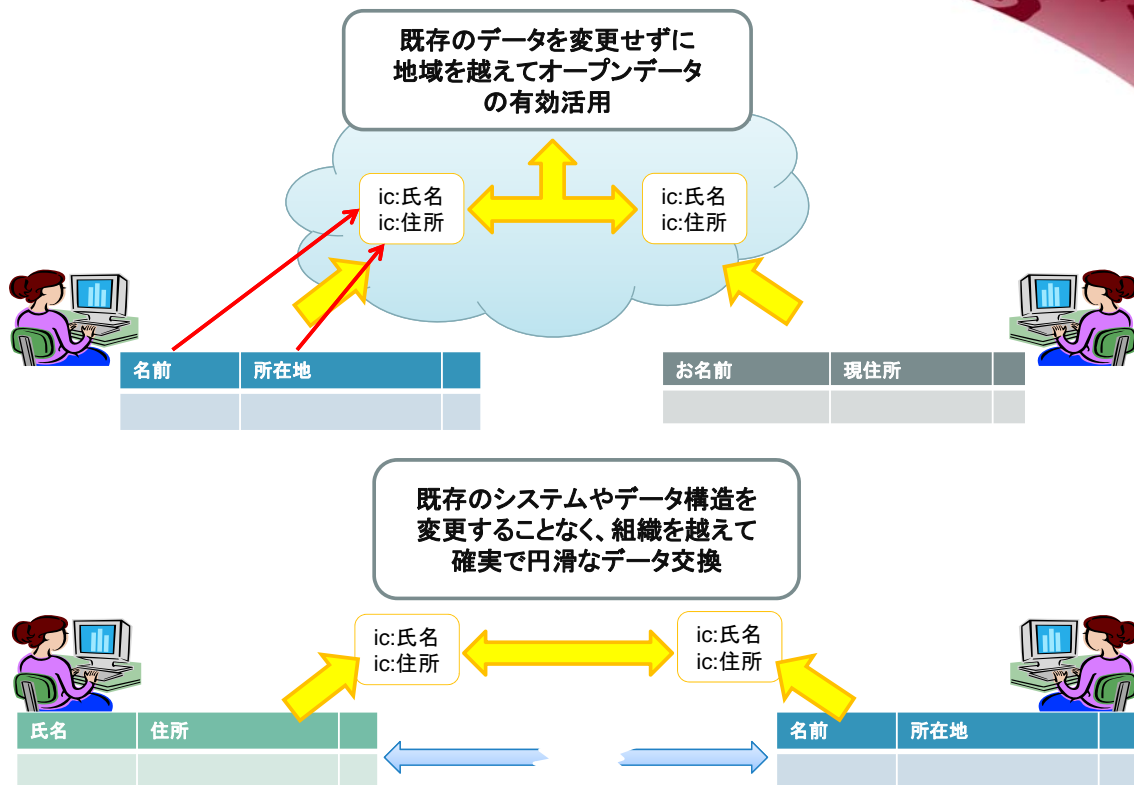


	A	B	C	D	E	F
1	平成27年度新潟県市町村立小学校(県内全体)					
2	No.	市町村	学校名	郵便番号	所在地	電話番号
3	1	上越市	大手町小学校	943-0838	上越市大手町2-20	025(524)6160
4	2	上越市	東本町小学校	943-0825	上越市東本町2-2-7	025(523)2446
5	3	上越市	南本町小学校	943-0841	上越市南本町3-9-1	025(523)2445
6	4	上越市	黒田小学校	943-0885	上越市大字黒田463-1	025(523)3811
7	5	上越市	飯小学校	943-0896	上越市大字飯1946	025(523)3810
8	6	上越市	富岡小学校	943-0173	上越市大字富岡3117	025(523)2843
9	7	上越市	稲田小学校	943-0154	上越市稲田1-6-7	025(523)2864
10	8	上越市	和田小学校	943-0856	上越市大字上箱井202	025(523)3330
11	9	上越市	大和小学校	943-0861	上越市大和2-13-3	025(523)5392
12	10	上越市	春日小学校	943-0802	上越市大豆1-13-11	025(523)3859
13	11	上越市	高志小学校	943-0805	上越市木田3-1-25	025(523)3858

	A	B	C	D	E
1	〒	市区町村名	所在地	電話番号	学校名
2	1530-0043	大阪市	大阪市北区天満1-24-15	06-6351-1582	滝川小学校
3	2530-0044	大阪市	大阪市北区東天満2-10-7	06-6358-3336	堀川小学校
4	3530-0047	大阪市	大阪市北区西天満3-12-21	06-6364-2222	西天満小学校
5	4530-0031	大阪市	大阪市北区菅栄町9-5	06-6358-1851	菅北小学校
6	5531-0062	大阪市	大阪市北区長柄中2-3-30	06-6351-4920	豊崎東小学校
7	6531-0073	大阪市	大阪市北区本庄西2-1-16	06-6371-0638	豊崎本庄小学校

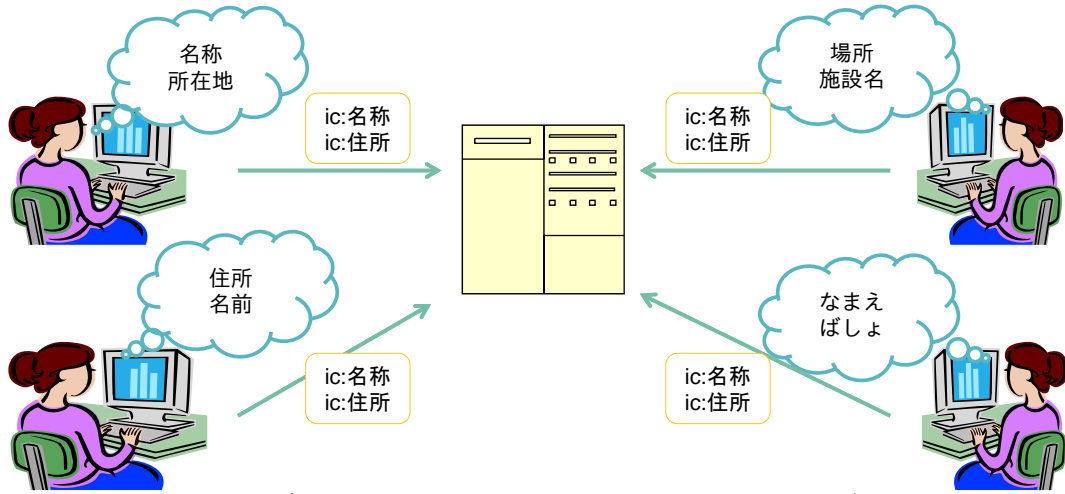
	A	B	C	D
1	名称	電話番号	郵便番号	所在地
2	東二番丁小学校	222-6279	〒980-0811	青葉区一番町二丁目1-4
3	木町通小学校	223-3480	〒980-0801	青葉区木町通一丁目7-36
4	同東北大学病院分校	718-7525	〒980-0872	青葉区星陵町1-1
5	立町小学校	222-4078	〒980-0822	青葉区立町8-1
6	東六番丁小学校	222-4216	〒980-0004	青葉区宮町一丁目2-1
7	片平丁小学校	223-3846	〒980-0812	青葉区片平一丁目7-1
8	上杉山通小学校	221-3392	〒980-0011	青葉区上杉一丁目10-1
9	通町小学校	234-2471	〒981-0915	青葉区通町一丁目1-1

既存のデータを変更せず、コア語彙にマッピング！！



データ構造の共有

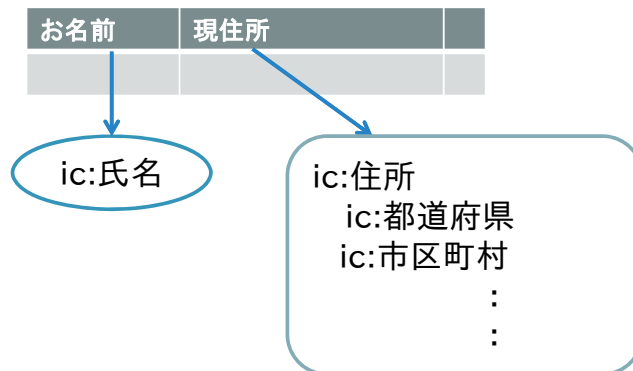
- 構造や項目名が異なるが、意味的に共通のデータは多い
- 多くの人々が最初からデータ構造を作成するのは効率的でないだけでなく、たくさんの異なる構造ができてしまう。
- ✓ データ構造を共有して利用することで時間とコストを節約



データ構造の検索・利用・カスタマイズ

住所などを定型化

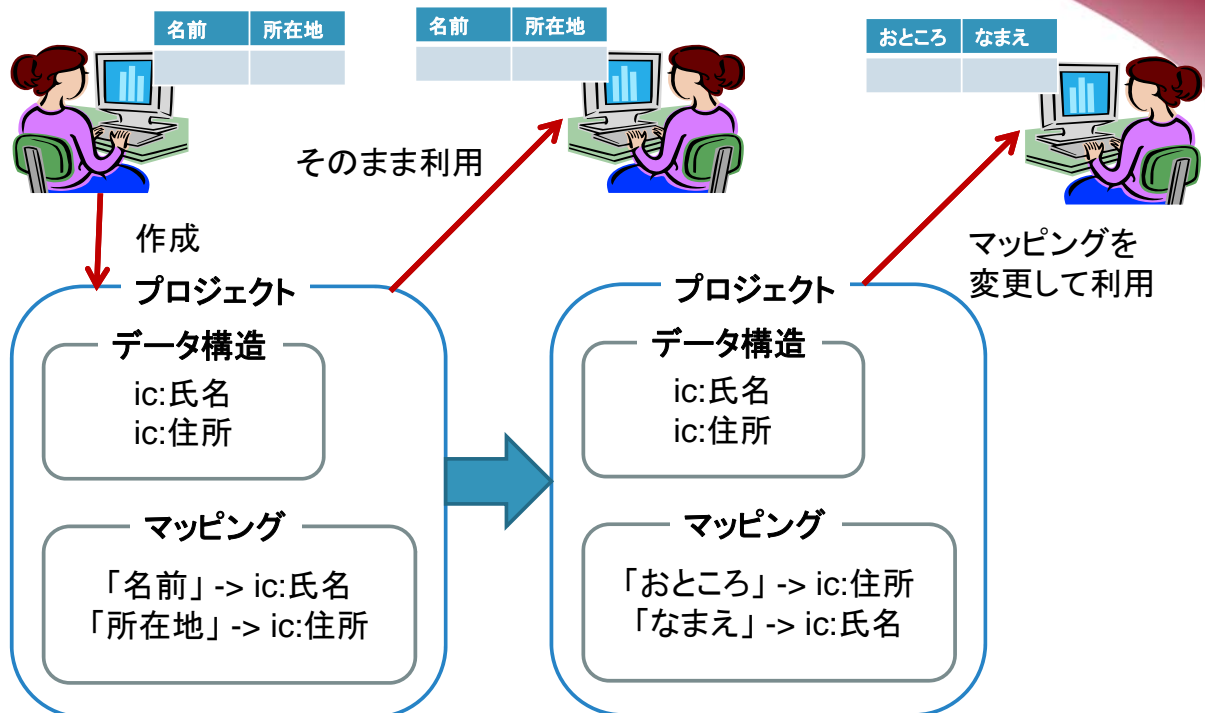
- 住所を自動的に分解・定型化することで、さらに使いやすいデータに



DEMO

<https://imi.ipa.go.jp/tools/0051/>

データ構造の共有



	A	B	C	D	E	F
1	平成27年度新潟県市町村立小学校(県内全体)					
2	No.	市町村	学校名	郵便番号	所在地	電話番号
3	1	上越市	大手町小学校	943-0838	上越市大手町2-20	025(524)6160
4	2	上越市	東本町小学校	943-0825	上越市東本町2-2-7	025(523)2446
5	3	上越市	南本町小学校	943-0841	上越市南本町3-9-1	025(523)2445
6	4	上越市	黒田小学校	943-0885	上越市大字黒田463-1	025(523)3811
7	5	上越市	飯小学校	943-0896	上越市大字飯1946	025(523)3810
8	6	上越市	富岡小学校	943-0173	上越市大字富岡3117	025(523)2843
9	7	上越市	稲田小学校	943-0154	上越市稲田1-6-7	025(523)2864
10	8	上越市	和田小学校	943-0856	上越市大字上箱井202	025(523)3330
11	9	上越市	大和小学校	943-0861	上越市大和2-13-3	025(523)5392
12	10	上越市	春日小学校	943-0802	上越市大豆1-13-11	025(523)3859
13	11	上越市	高志小学校	943-0805	上越市木田3-1-25	025(523)3858

	A	B	C	D	E
1	〒	市区町村名	所在地	電話番号	学校名
2	1530-0043	大阪市	大阪市北区天満1-24-15	06-6351-1582	滝川小学校
3	2530-0044	大阪市	大阪市北区東天満2-10-7	06-6358-3336	堀川小学校
4	3530-0047	大阪市	大阪市北区西天満3-12-21	06-6364-2222	西天満小学校
5	4530-0031	大阪市	大阪市北区菅栄町9-5	06-6358-1851	菅北小学校
6	5531-0062	大阪市	大阪市北区長柄中2-3-30	06-6351-4920	豊崎東小学校
7	6531-0073	大阪市	大阪市北区本庄西2-1-16	06-6371-0638	豊崎本庄小学校

	A	B	C	D
1	名称	電話番号	郵便番号	所在地
2	東二番丁小学校	222-6279	〒980-0811	青葉区一番町二丁目1-4
3	木町通小学校	223-3480	〒980-0801	青葉区木町通一丁目7-36
4	同東北大学病院分校	718-7525	〒980-0872	青葉区星陵町1-1
5	立町小学校	222-4078	〒980-0822	青葉区立町8-1
6	東六番丁小学校	222-4216	〒980-0004	青葉区宮町一丁目2-1
7	片平丁小学校	223-3846	〒980-0812	青葉区片平一丁目7-1
8	上杉山通小学校	221-3392	〒980-0011	青葉区上杉一丁目10-1
9	通町小学校	234-2471	〒981-0915	青葉区通町一丁目1-1

```
<http://example.com/3>
ic:連絡先 [ ic:電話番号 "025(5
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "943-0838"@ja ;
  ic:表記 "上越市大手町2-20
];
ic:名称 [ ic:表記 "大手町小学校
ic:種別 "小学校"@ja .
```

```
<http://example.com/2>
ic:連絡先 [ ic:電話番号 "
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "1530-0043
  ic:表記 "大阪市北区天満
];
ic:名称 [ ic:表記 "滝川小
ic:種別 "小学校"@ja .
```

```
<http://example.com/2>
ic:連絡先 [ ic:電話番号 "
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "〒980-081
  ic:表記 "青葉区一番町二
];
ic:名称 [ ic:表記 "東二番
ic:種別 "小学校"@ja .
```

```
<http://example.com/4>
ic:連絡先 [ ic:電話番号 "025(5
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "943-0825"@ja ;
  ic:表記 "上越市東本町2-2-
];
ic:名称 [ ic:表記 "東本町小学校
ic:種別 "小学校"@ja .
```

```
<http://example.com/3>
ic:連絡先 [ ic:電話番号 "
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "2530-0044
  ic:表記 "大阪市北区東天
];
ic:名称 [ ic:表記 "堀川小
ic:種別 "小学校"@ja .
```

```
<http://example.com/3>
ic:連絡先 [ ic:電話番号 "
ic:地理座標 [
  ic:経度 " "@ja ;
  ic:緯度 " "@ja
];
ic:住所 [
  ic:郵便番号 "〒980-080
  ic:表記 "青葉区木町通一
];
ic:名称 [ ic:表記 "木町通
ic:種別 "小学校"@ja .
```

ツールを使用するには



まずはアクセス！

<https://imi.ipa.go.jp/tools/0051/>

ありがとうございました。