

共通語彙基盤 2013年度事業の概要

2014 年2月

独立行政法人情報処理推進機構

- 語彙基盤を支えるデータベースとツール群について、その要求要件を洗い出した「概念モデルの構築」を行う
- 「概念モデルの構築」にあたっては、パイロットシステムを運用し、実際のデータ作成を試行的に実施し、そこで得られた知見を盛り込む
- プロジェクト全体を「DBプロジェクト」と「ツールプロジェクト」の2種類で構成する。

1件を実施

の概念モデルの構築及びパイロットシステムの構築・運用に係る一般競争入札
(総合評価落札方式)

→DBプロジェクト

3~5件を実施

と連携するデータ設計・作成支援ツール群のパイロットシステム構築・運用及び概念モデルの構築に係る提案公募

→ツールプロジェクト

情報連携用語彙データベース

情報連携用語彙データベース事業の実施者

■ DBプロジェクト

語彙データベースのパイロットシステムを構築・運用し、データベースやツールに関する知見や課題を整理

■ ツールプロジェクト

パイロットシステムを活用し、そこから提供される語彙データを活用して、再利用性の高いデータの作成に関わるツールを試作し、自治体の現場で運用する。この試作・運用を通して、語彙データ、データ構造やツールの機能について得られた知見や課題を整理

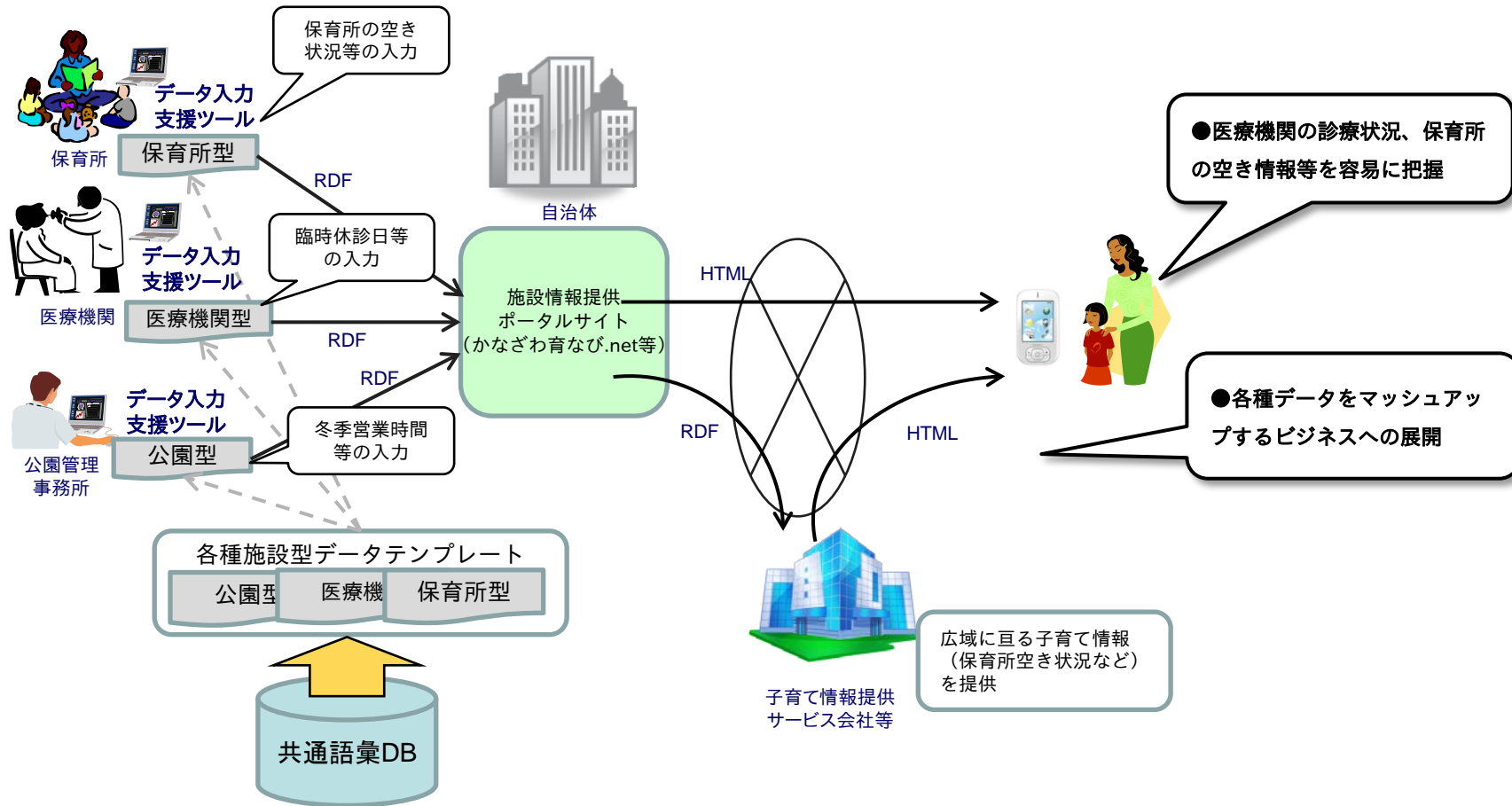
共通語彙基盤の概念モデルの構築実施者			
プロジェクト	実施者	プロジェクトの特徴	試用公的団体
DBプロジェクト	株式会社日立製作所		
ツールプロジェクト	1.株式会社三菱総合研究所	施設・サービス」のデータ化	金沢区
	2.インディゴ株式会社	「道路」のデータ化、他の地理情報との連携	浦安市都市整備部市街地開発課液状化対策推進室
	3.株式会社ネットワーク応用通信研究所	「観光」のデータ化	松江市 島根大学 松江商工会議所
	4.株式会社アスコエパートナーズ	「制度」のデータ化	神戸市こども家庭局こども企画育成部 千葉市総務局情報経営部業務改革推進課 川口市企画財政部情報政策課
	5.インフォコム株式会社	「イベント」のデータ化	千葉市総務局情報経営部業務改革推進課

「情報連携用語彙データベースと連携する
データ設計・作成支援ツール群の試作及び
試用並びに概念モデルの構築」
各プロジェクトの概要

1. 「施設・サービス」のデータ化

金沢区（株式会社三菱総合研究所）

■ 保育所や医療機関など、住民の日常生活に密接に関わる施設・サービスの情報をデータ化するための基本語彙、ツールを試作し、公的団体による試用を実施



公的団体による試用

各施設の職員に実際にツールを使用してもらい、子育て施設情報の入力およびRDFデータ化を試行。

成果

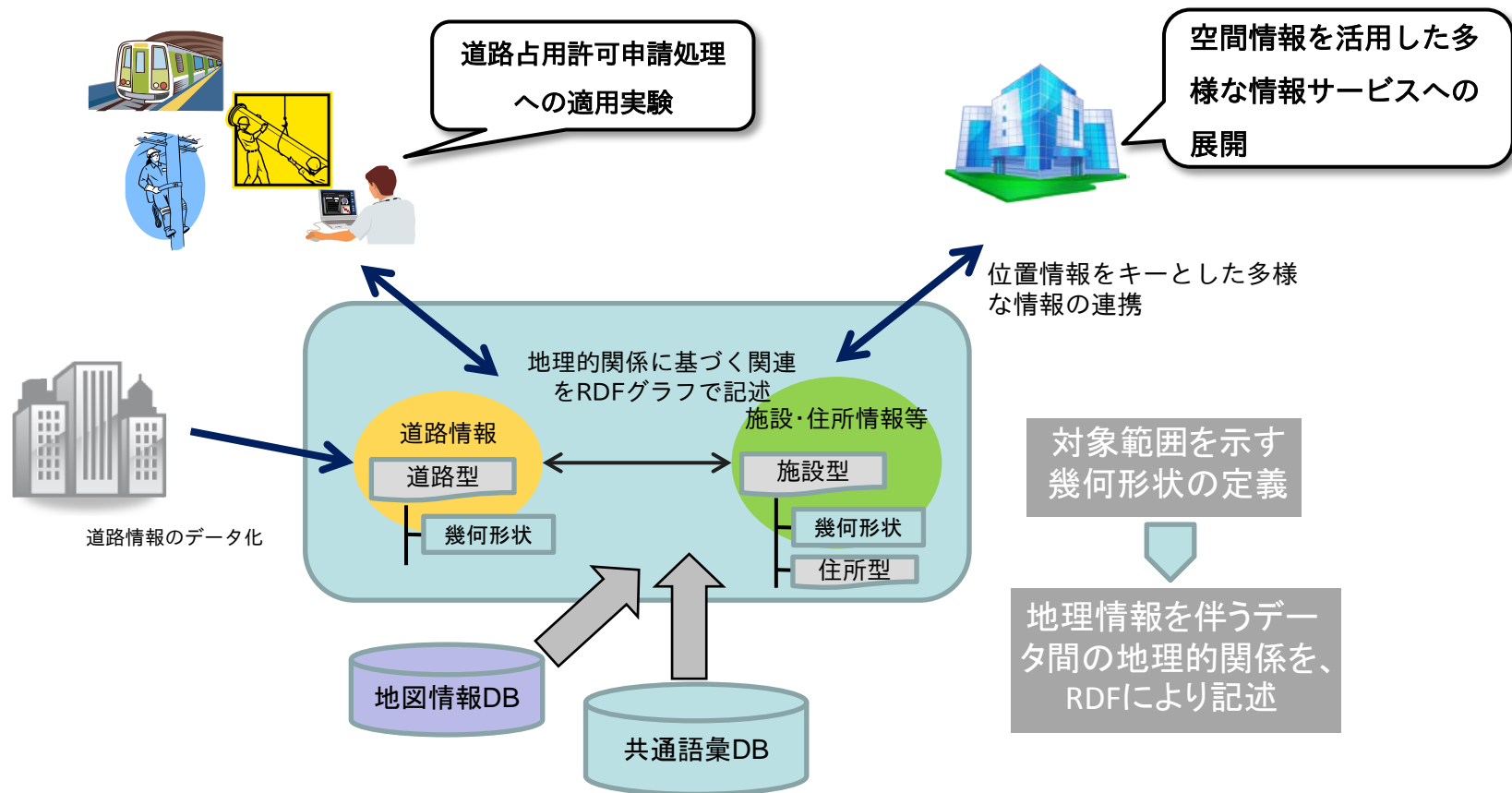
- ・ 公的団体による試用を通じて得られたノウハウを元に各種「施設・サービス」を記述するための基本データ型を提案。
- ・ データ入力支援ツールやDBの要件について考察。

2. 「道路」のデータ化、他の地理情報との連携

浦安市都市整備部市街地開発課液状化対策推進室（インディゴ株式会社）



- 「道路型のデータ」と「住所型等他のデータ型のデータ」との地理的関係性を明確にし、道路占用許可等の業務の効率化を図ることにより、民間による道路活用事業を促進



公的団体による試用

道路占用許可申請といった現場の事務処理を例として、地理情報連携のためのデータ構造や、関連ツールを試用して評価。

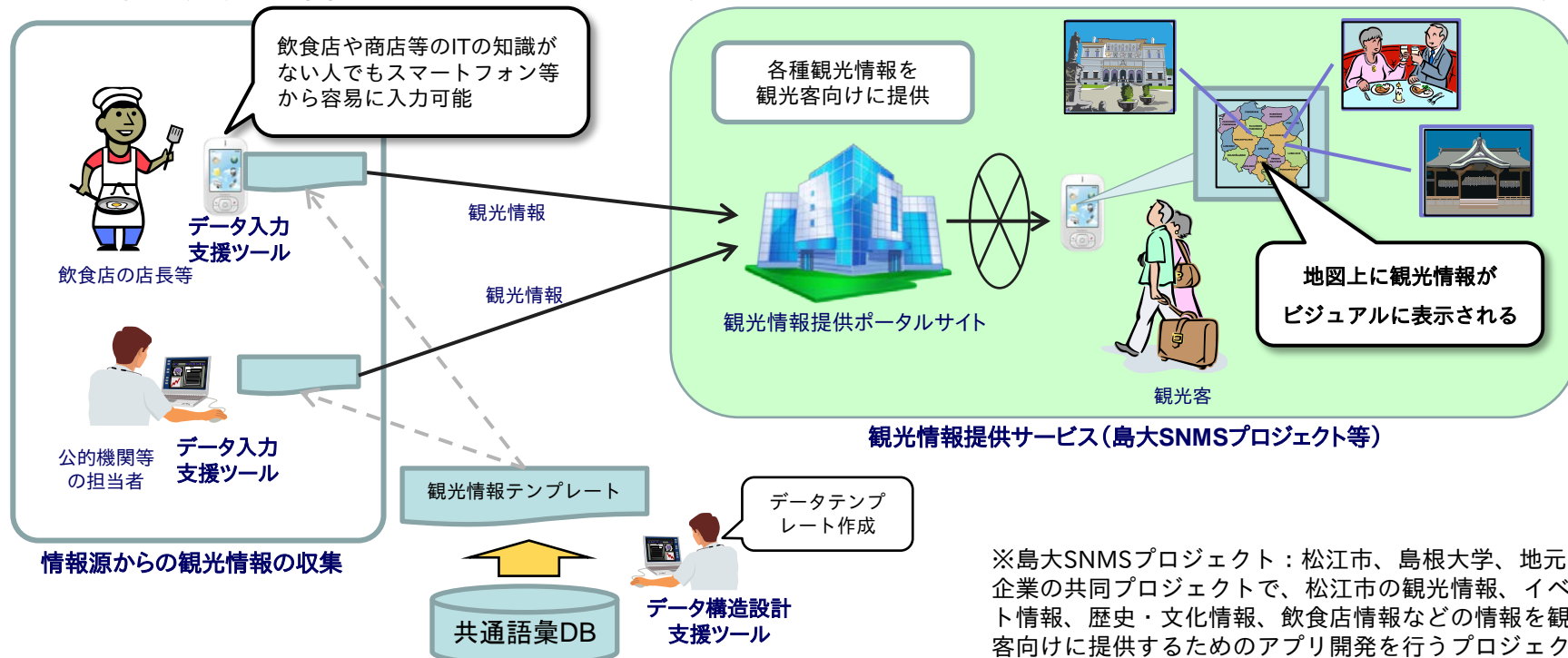
成果

・施設、統計、イベントなど、地理的情報を伴う多様な情報を連携させるための基本となるデータ構造を提案。

3. 「観光」のデータ化

松江市、島根大学、松江商工会議所（株式会社ネットワーク応用通信研究所）

- 情報源となる商店や中小企業等のITに関する非熟練者でも、容易に語彙の統制がとられたオープンデータを作成することを可能とする、基本語彙データ、ツールを試作して公的団体による試用を実施
- 松江市が別途独自に実施するアプリ開発プロジェクトとタイアップし、データ構築を試用



※島大SNMSプロジェクト：松江市、島根大学、地元IT企業の共同プロジェクトで、松江市の観光情報、イベント情報、歴史・文化情報、飲食店情報などの情報を観光客向けに提供するためのアプリ開発を行うプロジェクト

公的団体による試用

島大SNMSプロジェクトの担当者にデータ構造設計支援ツールでデータ構造を設計してもらい、商店や中小企業等にデータ入力支援ツールで観光情報を入力してもらう。さらに、作成された観光情報を島大SNMSで活用してもらい、実際に観光客に情報提供してもらう。

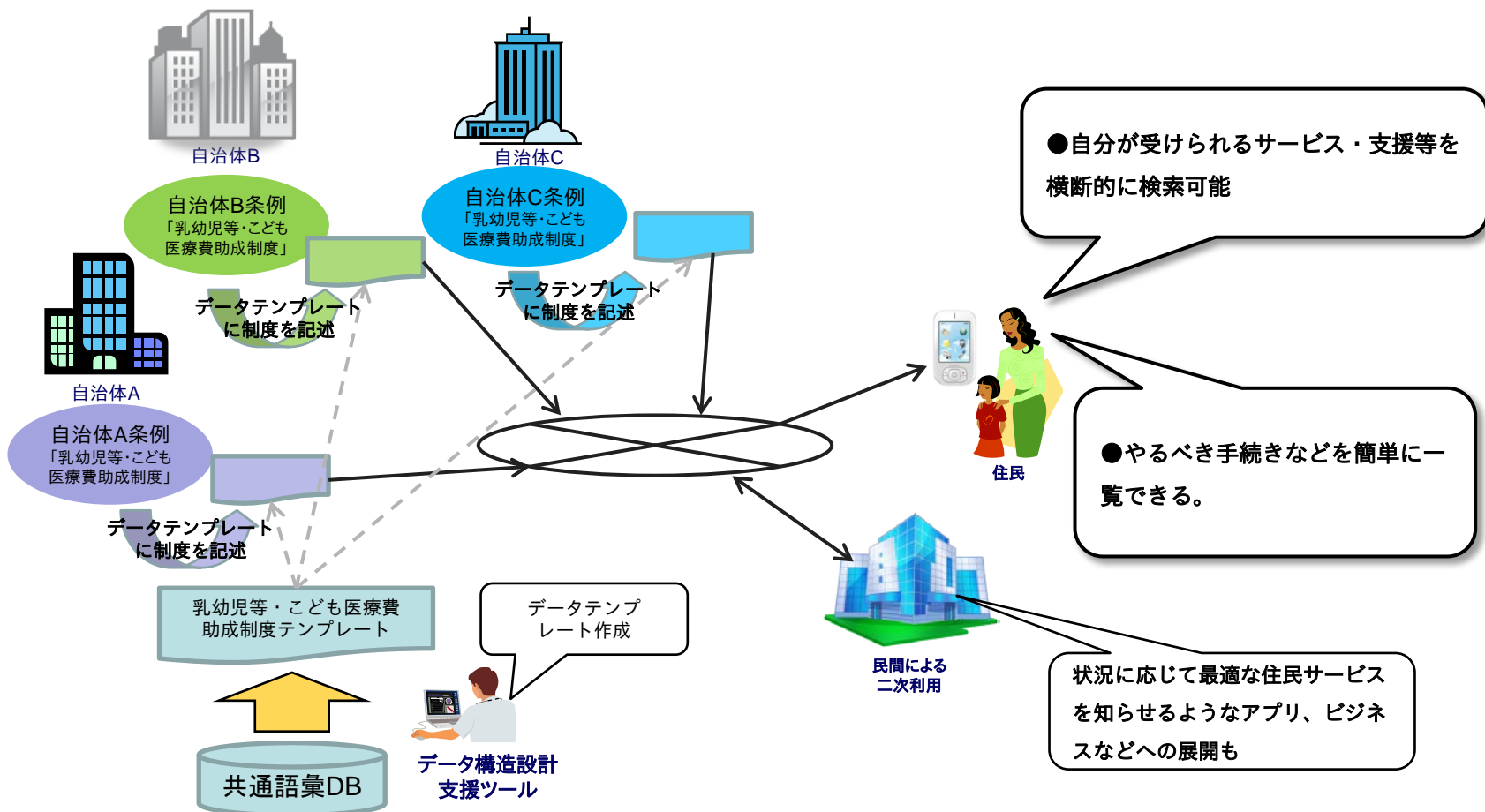
成果

- ・公的団体による試用を通じて得られたノウハウを元に「観光情報」のデータ型の基本型やパターン化を提案
- ・データ構造設計支援ツール、データ入力支援ツール、DBの要件について考察

4. 「制度」のデータ化

神戸市こども家庭局こども企画育成部、千葉市総務局情報経営部業務改革推進課、川口市企画財政部情報政策課（株式会社アスコエパートナーズ）

■ こども医療費助成制度等の制度を説明する情報を、統一化された形式・用語でデータ化



公的団体による試用

制度を写し取るためのデータテンプレートを試作し、それを用いて複数の自治体で実際の制度のデータ化を試行する

成果

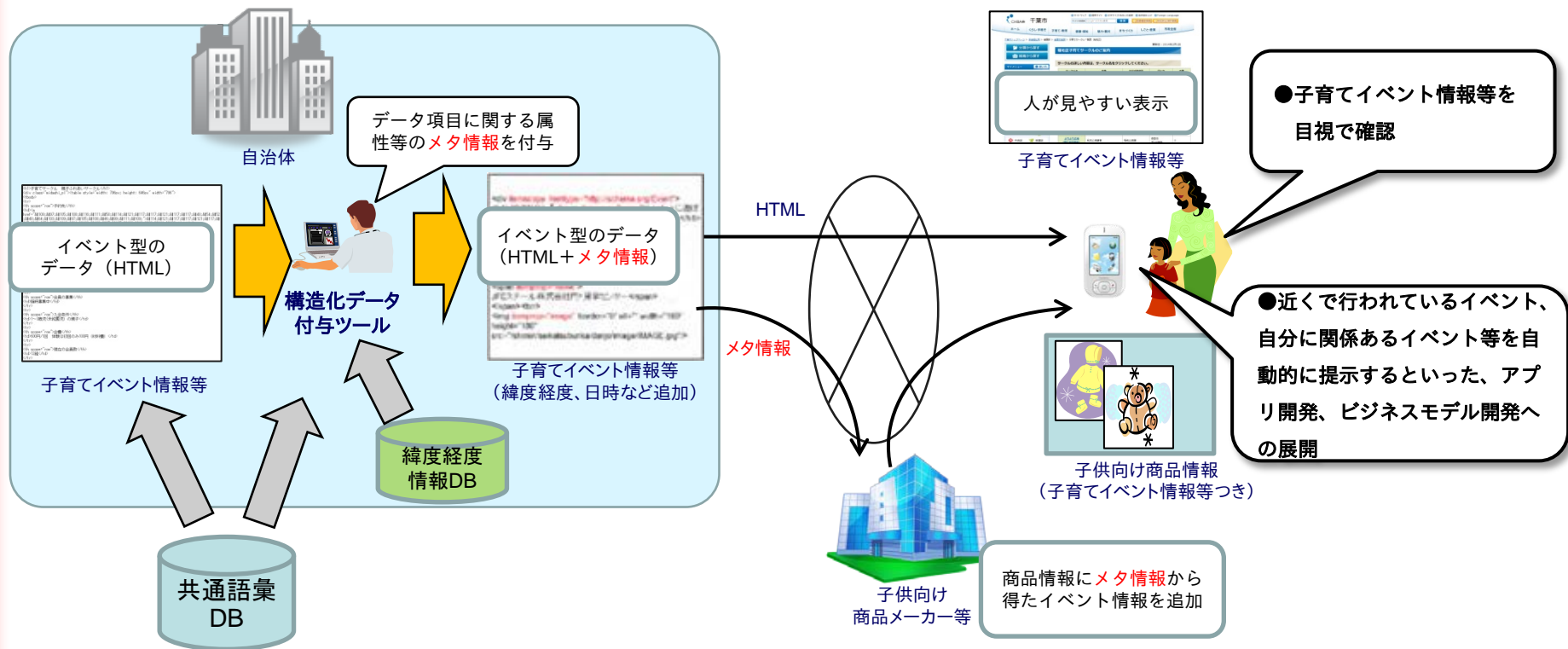
- 公的団体による試用を通じて得られたノウハウを元に「制度」をデータ化するために用いる共通語彙やパターンを提案
- データ構造設計支援ツールやDBの要件について考察

5. 「イベント」のデータ化

千葉市総務局情報経営部業務改革推進課（インフォコム株式会社）



■市が催す子育て関連各種イベントの情報を、オープンデータとして活用しやすい形式で提供するための方式を試行。「人が見てわかりやすい形式」と、「機械処理が容易な形式」を共存させるデータを生成する。




公的団体による試用

HTML形式のイベント情報を構成する各データに対し、その属性等のメタ情報を、共通語彙を使用して付与する試行を行う。

成果

- ・公的団体による試用を通じて得られたノウハウを元に「イベント」を記述するための基本語彙を整備
- ・構造化データを付与するためのツールやDBの要件について考察



情報連携用語彙データベースの概念モデルの構築
及びパイロットシステムの構築・運用

進捗報告

2014年2月13日
株式会社 日立製作所

Human Dreams.
Make IT Real.

Contents

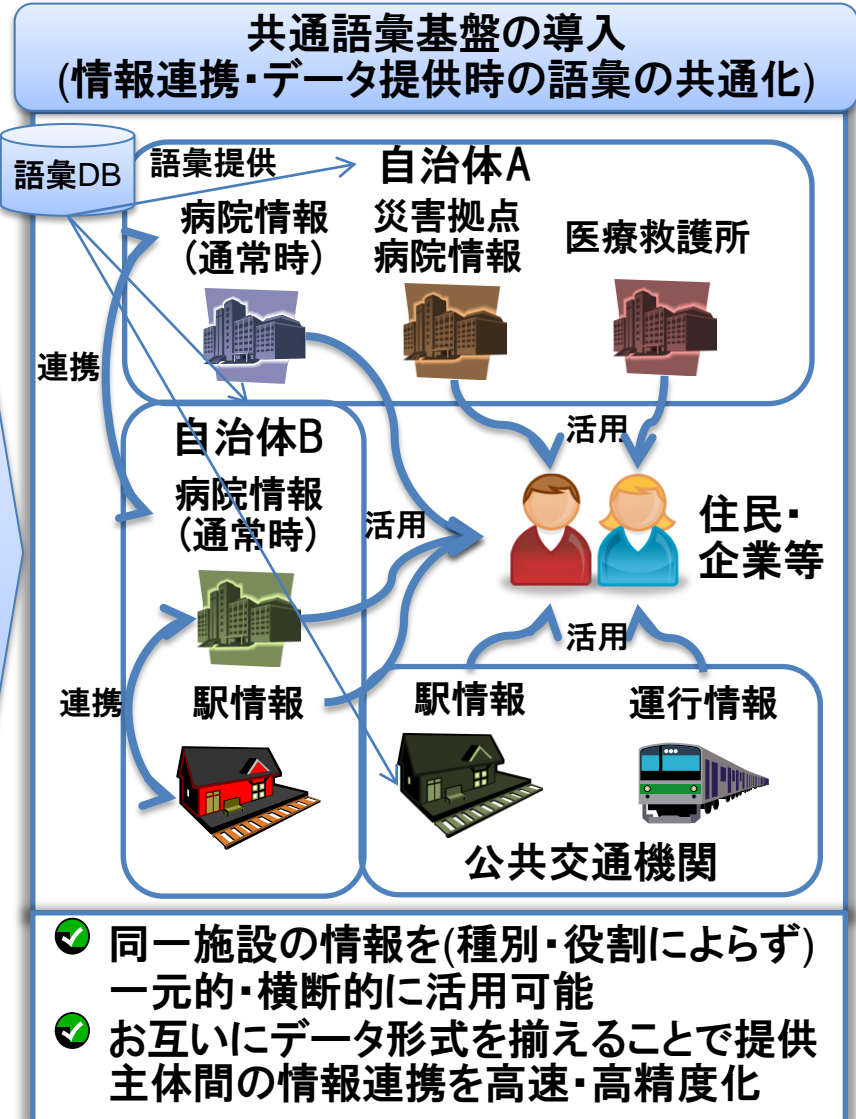
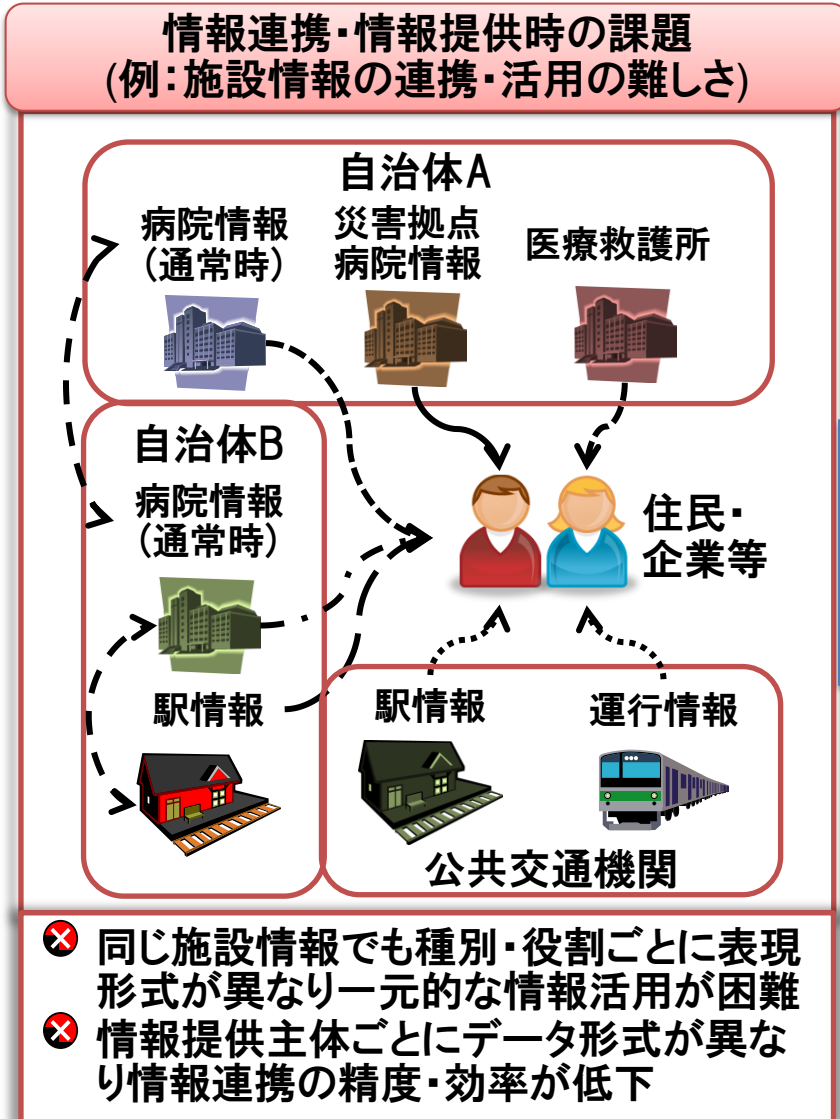
1. 本事業の概要
2. 現在の検討状況
3. 今後の予定

別紙1 データ構造(案)

1. 本事業の概要

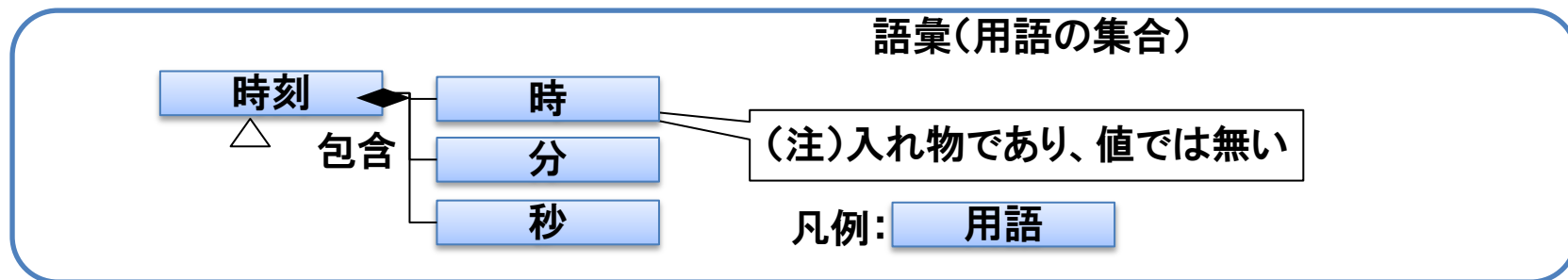
1-1. 共通語彙基盤導入の効果

- 共通語彙基盤を導入することにより、地域・組織・部門・業種・業務の壁を越えた横断的な情報連携やオープンデータ利活用が可能になると期待される。



1-2. 語彙とは？

- 「用語」は、ある特定の意味あるいは概念を示すものであり、文字を組み合わせた表記を伴う。
また、用語は、その意味あるいは概念の一部として、他の用語の関係(構造)を持つ場合がある。たとえば、「時刻」という用語を「時」、「分」、「秒」という下位の用語から構成すると定義することができる。また、「融資」と「貸し付け」のように、同一の意味を持つ用語が、異なる表記で示される場合がある。
- 用語の集合が「**語彙**」である。
- 用語の構造や意味、表記の制限、用語と用語との関係等、用語に係る様々な情報を総称してここでは「**語彙データ**」と呼ぶこととする。
なお、本事業で対象とする用語は、原則、行政の現場で業務上必要とされるものとする。



1-3. 本事業の主要な実施項目

- 共通語彙基盤の導入に向け、本プロジェクト(2013/11～2014/8)では主に下記の3つの項目を主要な実施項目として推進し、実施結果を概念モデルとしてまとめる。

共通語彙基盤の 要求事項洗い出し

(データベース検討会を通じた検討)

- ・ 語彙データベースの利用シーンを洗い出し、各シーンにおける想定利用者のニーズに基づいて、語彙の論理データ構造やAPIの要求事項に加え、APIを利用するツール群の要求事項を洗い出す。
- ・ 語彙データベースについて、語彙のメンテナンスを含めた運用全体の在り方に係る検討を行う。

語彙データの 収集・整備

(ドメイン語彙検討会を通じた
収集・整備)

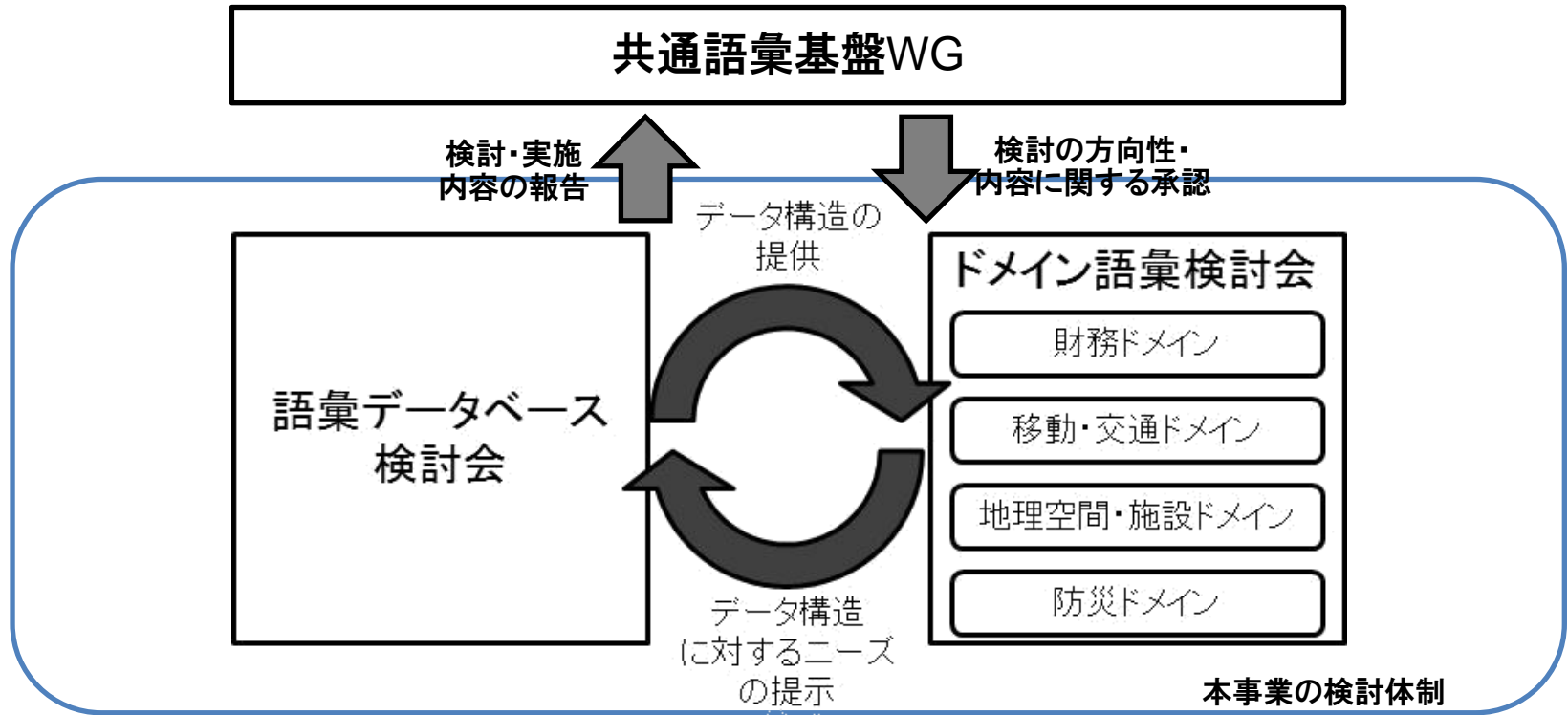
- ・ ドメイン語彙検討会(4ドメイン、全3回)を開催し、各ドメインの有識者ととともに語彙データの収集・整備方策の基礎を確立する。
- ・ ドメイン語彙検討会との協調により、初期データとして利用するコア語彙・ドメイン語彙の収集・整備を実施する。

パイロットシステム の開発・運用

- ・ 語彙基盤の概念モデルの構築に必要な知見を収集するため、語彙データベースのパイロットシステムを構築し、運用する。
- ・ パイロットシステムには「公共情報交換標準スキームの「公共情報交換標準スキームの整備に関する調査研究(2012年度)」に収められたIMIコアボキャブラリを格納する。

概念モデル書の作成

1-4. 本事業の検討体制



組織	内容
共通語彙基盤WG	情報連携用語彙データベースの構築・APIカタログ整備・ツールの整備
語彙データベース検討会	語彙のデータ構造を検討する
ドメイン語彙検討会	語彙の収集・整備方策について検討する
財務ドメイン	政府の公開する予算、決算、調達に係るデータを扱うドメイン
移動・交通ドメイン	人の移動や公共交通(電車・バス)の移動に関するデータを扱うドメイン
地理空間・施設ドメイン	地図データ及びその地図上にマッピングされる施設のデータを扱うドメイン
防災ドメイン	平時における災害対策及び発災時における緊急情報の機関間の情報連携に必要なとなるデータを扱うドメイン

1-5. 語彙データベース検討会構成員、オブザーバ

語彙データベース検討会構成員、オブザーバ

○:座長

項番	委員名	肩書き
1	○橋田 浩一	東京大学大学院情報理工学系研究科ソーシャルICT研究センター新融合サービスICT分野 教授
2	奥村 学	東京工業大学精密工学研究所 教授
3	加藤 文彦	大学共同利用機関法人 情報・システム研究機構 特任研究員
4	川島 宏一	佐賀県特別顧問(株式会社公共イノベーション 代表取締役)
5	越塚 登	東京大学大学院学際情報学府 教授
6	坂下 哲也	一般財団法人 日本情報経済社会推進協会 電子情報利活用推進部 部長
7	鈴木 章太郎	日本マイクロソフト株式会社 テクニカルエバンジェリスト
8	手塚 悟	東京工科大学コンピュータサイエンス学部 教授
9	三木 浩平	千葉市総務局 次長 情報統括副管理者
10	目黒 公郎	東京大学 教授 生産技術研究所 都市基盤安全工学国際研究センター長
項番	オブザーバ名	肩書き
1	田代 秀一	独立行政法人情報処理推進機構 技術本部国際標準推進センター長
2	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
3	平本 健二	経済産業省CIO補佐官
4	泉本 央	経済産業省 情報プロジェクト室 係長

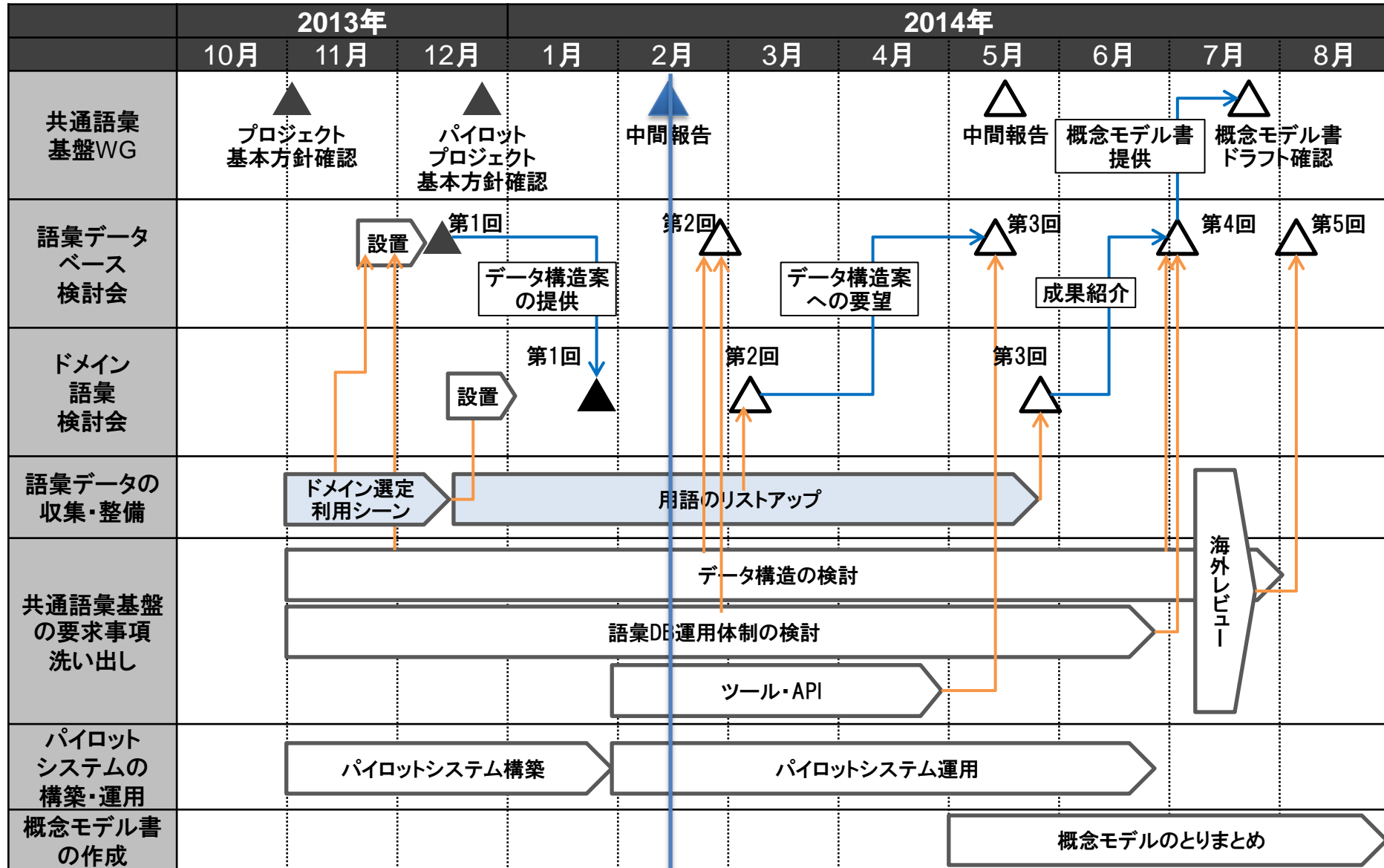
1-6. ドメイン語彙検討会構成員(4ドメイン)

ドメイン語彙検討会構成員

○:座長

項番	ドメイン	委員名	所属・役職
1	財務	○川島 宏一	佐賀県特別顧問(株式会社公共イノベーション 代表取締役)
2		大槻 文彦	一般社団法人XBRL Japan 副会長・理事
3		加藤 晴彦	株式会社自治体ドットコム 事務局長
4		小林 勝則	独立行政法人 産業技術総合研究所 総務本部 財務部 部長
5		森本 浩之	松阪市 CIO補佐官
6	防災	○目黒 公郎	東京大学 教授 生産技術研究所 都市基盤安全工学国際研究センター長
7		秋富 慎司	岩手医科大学岩手県高度救命救急センター 助教
8		橋 克憲	株式会社パスコ コンサルティングサービス部 部長
9		長坂 俊成	立教大学 21世紀社会デザイン研究科 危機管理学分野 教授
10		渡辺 研司	名古屋工業大学大学院 社会工学専攻 教授
11	移動交通	○越塚 登	東京大学大学院学際情報学府 教授
12		河口 信夫	名古屋大学大学院工学研究科計算理工学専攻 教授
13		北崎 茂	株式会社ナビタイムジャパン 交通コンサルティング事業 エンジン開発部 データアナリスト
14		中村 文彦	横浜国立大学大学院都市イノベーション研究院 教授
15		堀池 雅彦	京都市 交通政策監
16	地理空間 ・施設	○坂下 哲也	一般財団法人 日本情報経済社会推進協会 電子情報利活用推進部 部長
17		佐々木 陽一	PHP研究所 地域経営研究センター 主任研究員
18		関 治之	合同会社 Georepublic Japan 代表社員
19		関本 義秀	東京大学 生産技術研究所 准教授
20		高木 和之	株式会社ゼンリン ZIP企画室 副長

1-7. スケジュール



2. 現在の検討状況について

2-1. データ構造(案)について

- 前プロジェクトであるIMIを包含するデータ構造の仮説を立案中(別紙1参照)。今後、本プロジェクトの中で当該仮説を検証し、最終的な概念モデルを整備する。

データ構造 (概念モデル)

本プロジェクトの中で
・利用シーン検討
・各種先行事例の調査 (含現地調査)
・要求事項の洗い出し
等を実施し、
最終的な概念モデルを整備

データ構造 (仮説)

意味の記述

XMLスキーマ・RDF等
多様な形式での出力

構造の記述

schema.org 等との整合性

関係性の記述

表記ゆれの明確化

同義語・類義語の明確化

多義語の明確化

同義語・類義語の抑制

項目名

項目説明

データタイプ

NIEM との対応

cardinality

ISA Joinup との対応

文字セットの取扱い

柔軟性・拡張性・汎用性

各種コードとの連携

先行事例との整合性

意味の明確化と
再利用性の向上

表記変化への追従

前プロジェクト (IMI)
のカバー範囲

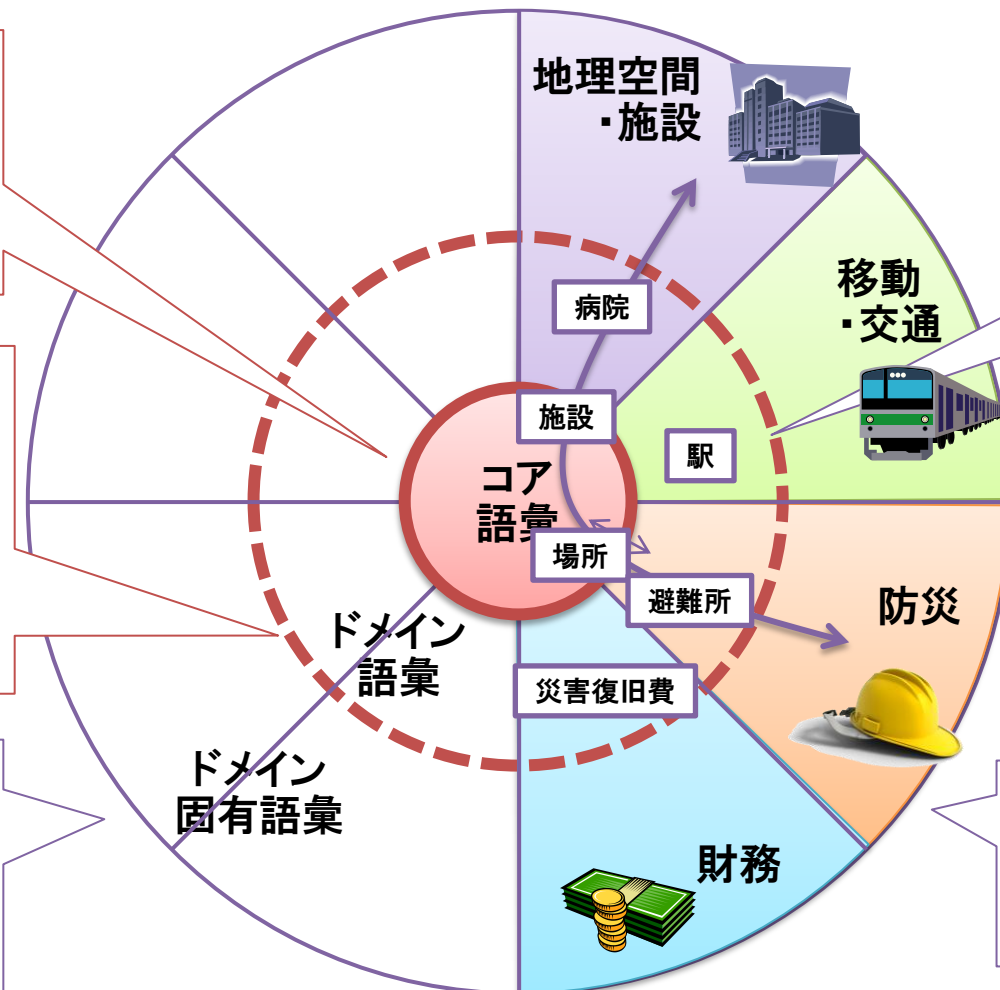
2-2. コア語彙とドメイン語彙について

- コア語彙およびドメイン語彙を収集・整備して語彙データベースに蓄積し、API経由で各ドメインの利用者へ提供することで、システム間の情報連携やオープンデータの利活用を容易化する。

コア語彙
[収集・整備対象]
どのドメインでも広く利用される普遍的な語彙。人、物、場所、日付など。

ドメイン語彙
[サンプル収集・整備対象]
ドメイン固有の語彙の内、他のドメインでも参照する主要な語彙。ドメイン間の情報連携・データ提供に利用される。

ドメイン固有語彙
[参照のみ]
各ドメインでの利用に特化した語彙。ドメイン内の各種標準や既存システムの語彙などを含む。

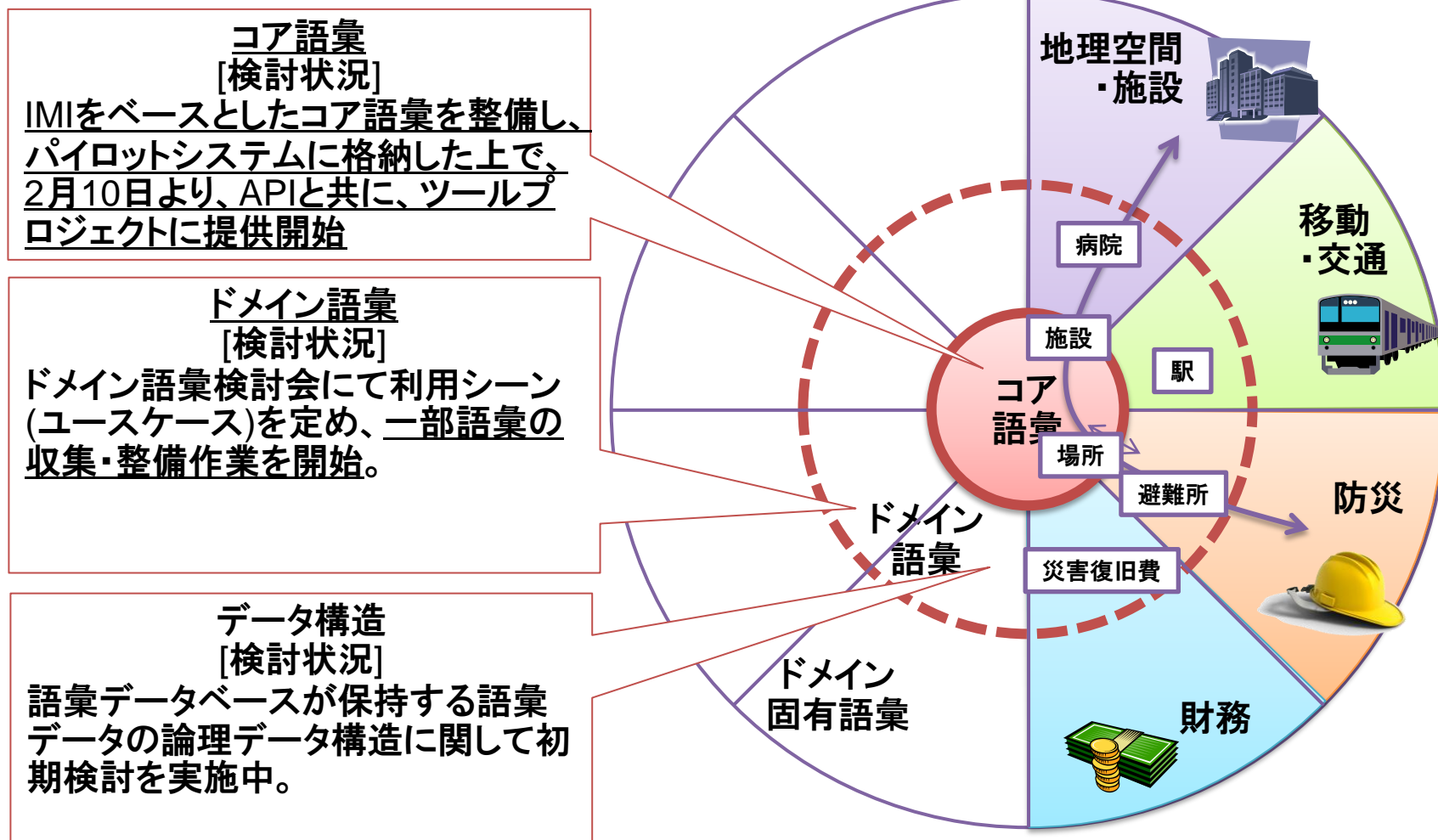


コア語彙とドメイン語彙で固有語彙を橋渡し

既存の語彙 (ドメイン固有語彙) を置換・変更するものではない

2-3. コア語彙とドメイン語彙についての検討状況

- コア語彙およびドメイン語彙を収集・整備についての検討状況を以下に示す。



2-4. 語彙の利用イメージ

● 本プロジェクトにて整備するコア語彙・ドメイン語彙の利用イメージを以下に示す。

利用シーン例:
公共施設(病院)の
運営情報の公開



公共施設情報



項目の「名前」や
「構造」をドメイン
語彙として整備

複数のドメイン
語彙を組合せて
必要な情報を表現

公共施設情報 (〇〇医院)				
施設: 病院	名称		〇〇医院	
	場所	住所	都道府県	東京都
			市区町村	江東区
		
	経緯度座標	経緯度座標	緯度	35.xxxxxxxx
			経度	139.xxxxxxxx
	地理識別子	地理識別子	識別子体系	□□住所コード
			識別子	13103xxxxxx
	連絡先	電話番号		03-xxxx-xxxx
		...		
スケジュール	内容		診療受付時間	
	開始時間		09:00:00+09:00	
	終了時間		17:00:00+09:00	
	曜日		MON,TUE,THU,FRI	
アクセス	最寄り駅	交通: 駅	名称	△△駅
	
防災: 災害拠点病院	病床数		800	
	三次救急		True	
	ヘリ離着陸		False	

3. 今後の予定

3-1. 語彙データベース検討会の今後の予定

本検討会では、特定の利用シーンを念頭に置き、各構成員の知見・意見を踏まえ、語彙データベースに関する議論が行われることを想定している。

本検討会の日程案及び議題案を以下に示す。2013年12月から2014年8月にかけて5回の検討会を開催し、各議題を通じてデータ構造に対するブラッシュアップを図るとともに、語彙データベースに対する要求事項を整理する。

以下に主な議題案について示す。

開催回	日程案	議題案
1	2013年12月	<ul style="list-style-type: none">・ 検討会の進め方について・ 語彙データベースの利用シーンについて・ データ構造案について
2	2014年 2月	<ul style="list-style-type: none">・ データ構造について(継続)・ 語彙データベースの運用体制について
3	2014年 5月	<ul style="list-style-type: none">・ ドメイン語彙検討会における検討状況について・ 語彙データベースのツールとAPIの仕様について
4	2014年 7月	<ul style="list-style-type: none">・ パイロットシステム運用を受けたデータ構造について・ ドメイン語彙検討会における検討結果について・ 語彙データベースの運用体制について
5	2014年 8月	<ul style="list-style-type: none">・ ツールプロジェクトの実証を受けたデータ構造について・ 海外レビューを受けた構造案について