

共通語彙基盤 2017年度実施のポイント

技術本部国際標準推進センター

I.世界最先端IT 国家創造宣言に基づくこれまでの成果

I-1. これまでの代表的な成果

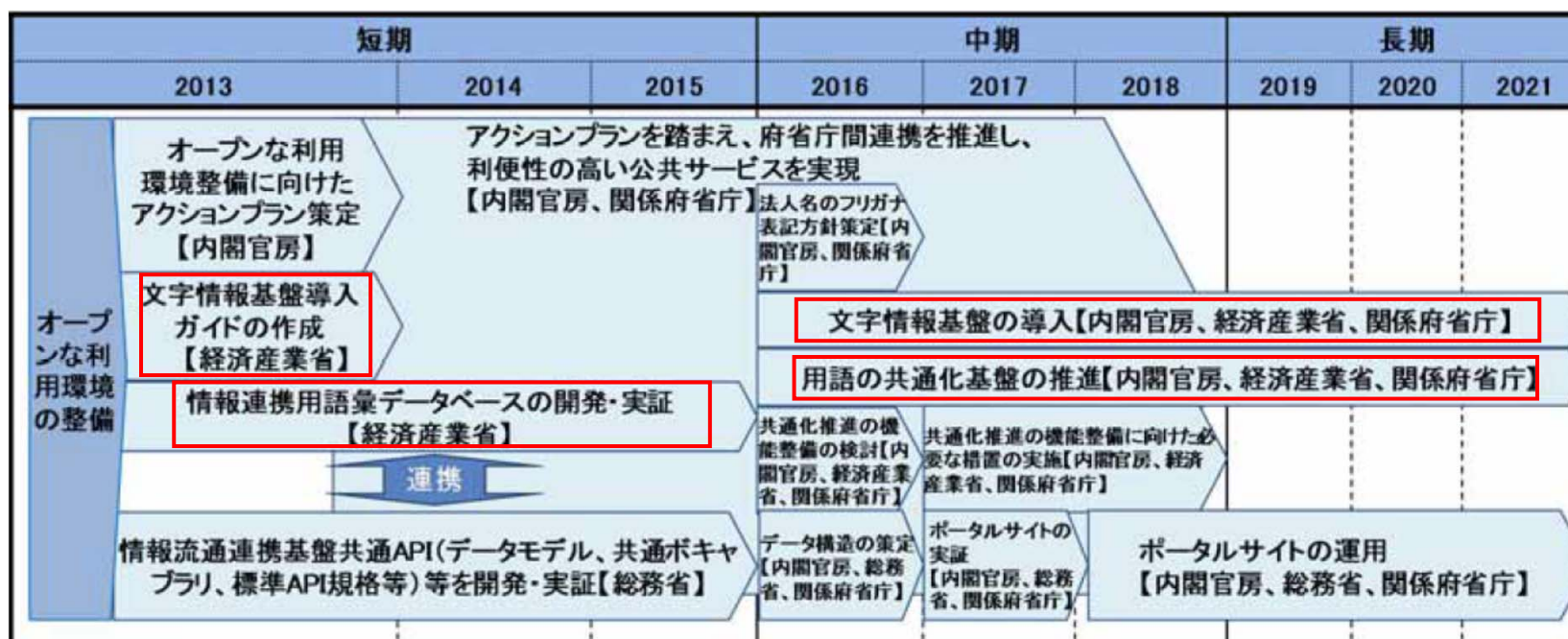
I-1-1 (1) 行政情報システム改革を通じた利用者志向の行政サービスの実現

[主な成果]

- ・ **文字情報基盤**(注4)と**共通語彙基盤**(注5)を整備。国・地方公共団体がこれらの基盤を活用することで、組織や業務分野を超えた行政情報システム間の連携が可能となり、円滑な公共サービスの提供を期待。

工程表

実施スケジュール (1.【重点項目1】国・地方のIT化・業務改革(BPR)の推進)



- 自治体での実証実験
 - 森町／会津若松市／埼玉県／横浜市金沢区／静岡市
 - 道路情報、イベント、観光、子育て支援情報等へ適用
 - 構築中の語彙データベース、ツールを活用
- 語彙データベース、ツールの公開
- IMIパートナー、官公庁との連携強化
 - 「公開ドラフト」の公開を推進
- 普及活動（セミナー、イベント実施）
 - 自治体向け
 - 技術者向け

共通語彙基盤2017年度案



- ・ 官民データ活用法に沿った計画策定の支援
- ・ オープンデータ構築のための要件定義支援

2017年度

2018年度

普及・啓発

地方セミナー（自治体向け。1回／2ヶ月程度）

コミュニティ、中小ベンダ支援

DB／ツール活用講習セミナー（技術者向け）

開発

データベース、支援ツールの開発

ツール、データ等の提供

試験運用

実証実験

実証イベント

統合

サイト

Imi.go.jpの運用

民間、自治体など

公開ドラフト、DMD等の登録

IMIパートナー

連携

委員会・作業部会

IMI運用検討SWG（公開ドラフト審査、パートナー審査、応用分野開拓）

IMI技術検討SWG

IMI戦略会議

IMI戦略会議

共通語彙基盤WG（武田委員長）

情報共有基盤推進委員会（須藤委員長）

国際

SEMIC NIEM

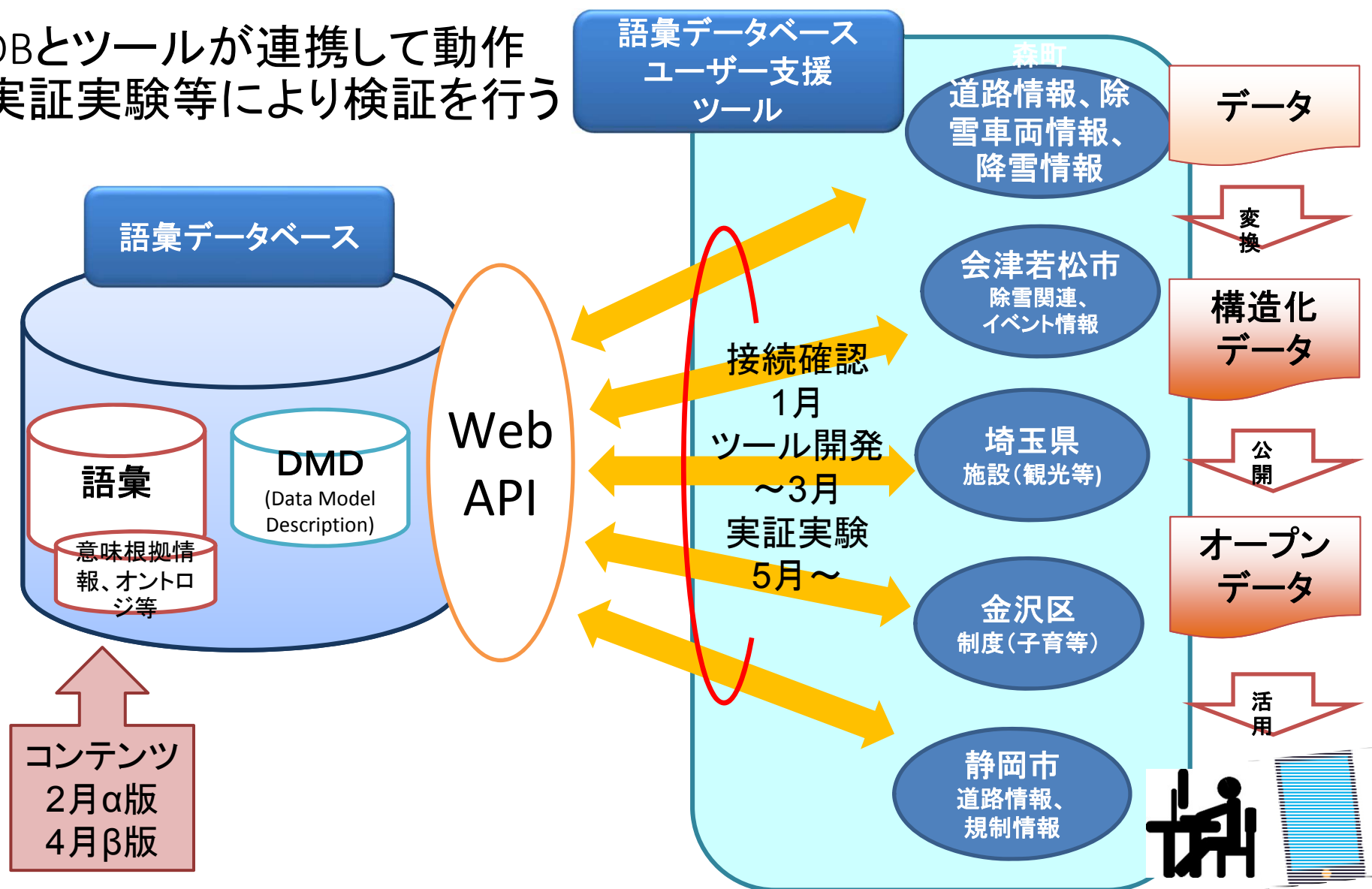
ITS

実証実験の実施 DB,ツールの公開



語彙データベース／ツール実証実験

DBとツールが連携して動作
実証実験等により検証を行う



実証実験内容（自治体現場を対象）



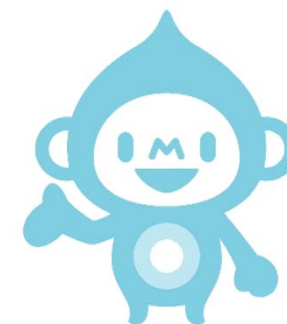
No	検証項目	使用するツール	検証概要	備考
1	用語(語彙)の作成	語彙作成支援ツール	対象データ(CSVまたは2次元表形式データ)の属性名をドメイン語彙として語彙データベースへ登録する。	<ul style="list-style-type: none"> 画面で一から作成する手順 語彙フォーム(Excel)をアップロードして登録する手順
2	コード値の作成	コード・コードリスト作成支援ツール	対象データの属性値について、自由入力ではない値(性別のような選択型の値)をコード値として語彙データベースへ登録する。	<ul style="list-style-type: none"> 画面で一から作成する手順 コードIDフォーム(Excel)をアップロードして登録する手順
3	DMDの作成	DMD作成支援ツール DMD検証ツール	<p>対象データをアップロードし、候補となるDMD(一致、類似するもの)を確認する。</p> <p>一致、類似するものがない場合は新規でデータ構造を作成し、DMDとして語彙データベースへ登録する。候補がある場合は必要な箇所を編集する形でDMDを作成し語彙データベースへ登録する。</p> <p>登録したDMDを検証ツールを用いて確認する。</p>	<ul style="list-style-type: none"> 画面で一から作成する手順 語彙データベースに登録されているDMDを参照する手順 ローカルのDMDをアップロードする手順
4	データの変換	データ形式変換ツール データ検証ツール	<p>対象データをアップロードしてデータの属性名とDMDの項目をマッピングし、XMLまたは、RDFの各シリアル化形式(Turtle、RDF/XML、JSON-LD)のいずれかを選択し、変換されたデータをダウンロードする。</p> <p>変換したデータを検証ツール(UI+API)を用いて確認する。</p>	<p>データについては</p> <ul style="list-style-type: none"> 画面で変換する方法 APIで変換する方法 <p>DMDについては</p> <ul style="list-style-type: none"> 語彙データベースに登録されているDMDを参照する方法 ローカルのDMDをアップロードする方法

実証実験（既存データへの適用）

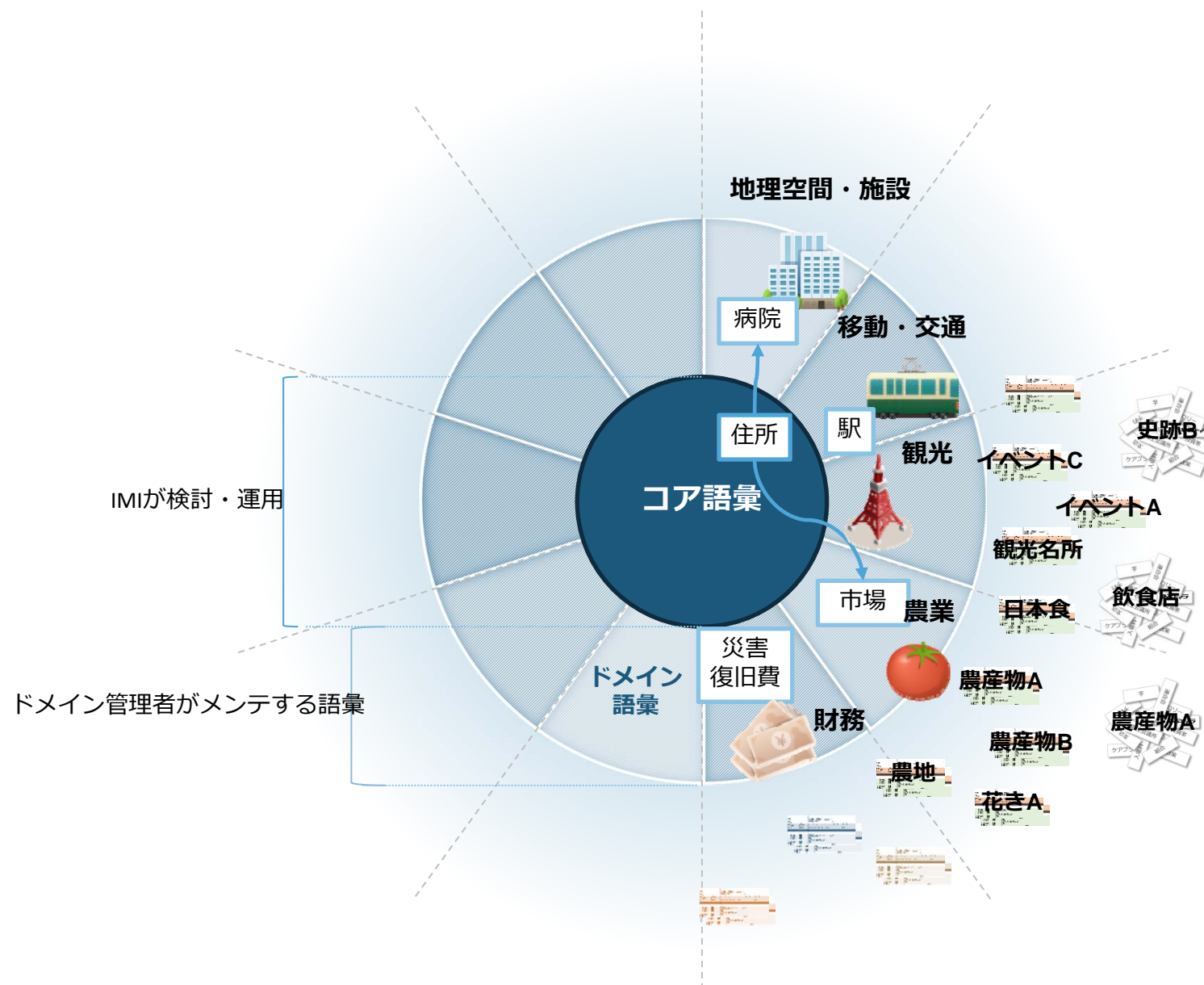


対象データ	
「linkdata.org」を用いて公開されている自治体（40自治体以上）のオープンデータ。 データセット数120個以上（40自治体の3倍以上）。	開発したツールを用い、IMIに沿ったRDFデータとDMDを作成し公開
「子育てタウン」Webを用いてHTMLで公開されている自治体（40自治体以上）の構造化データ	開発したツールを用い、IMIに沿ったRDFデータとDMDを作成し公開
経済産業省版法人ポータル（β版）	開発したツールを用い、DMD、ドメイン語彙を試作して公開
子供霞ヶ関見学デー（2017年実施分）	開発したツールを用い、DMD、ドメイン語彙、RDFデータを試作して公開。データビジュアライズツールも試作しデモ。
農業ITシステムで用いる農作業の名称に関する個別ガイドライン（本格運用版）（案）	公開された農作業オントロジーを対象とし、開発したツールを用いてDMD、ドメイン語彙を試作して公開。データビジュアライズツールも試作しデモ。
ハローワークの求人情報および求人詳細情報	開発したツールを用い、DMD、ドメイン語彙、RDFデータを試作して公開
ハローワークで公開されている求人情報で用いられている日本標準産業分類	産業分類について、開発したツールを用いてコードリストを試作し公開。統計LOD情報とマッシュアップなどもデモ。

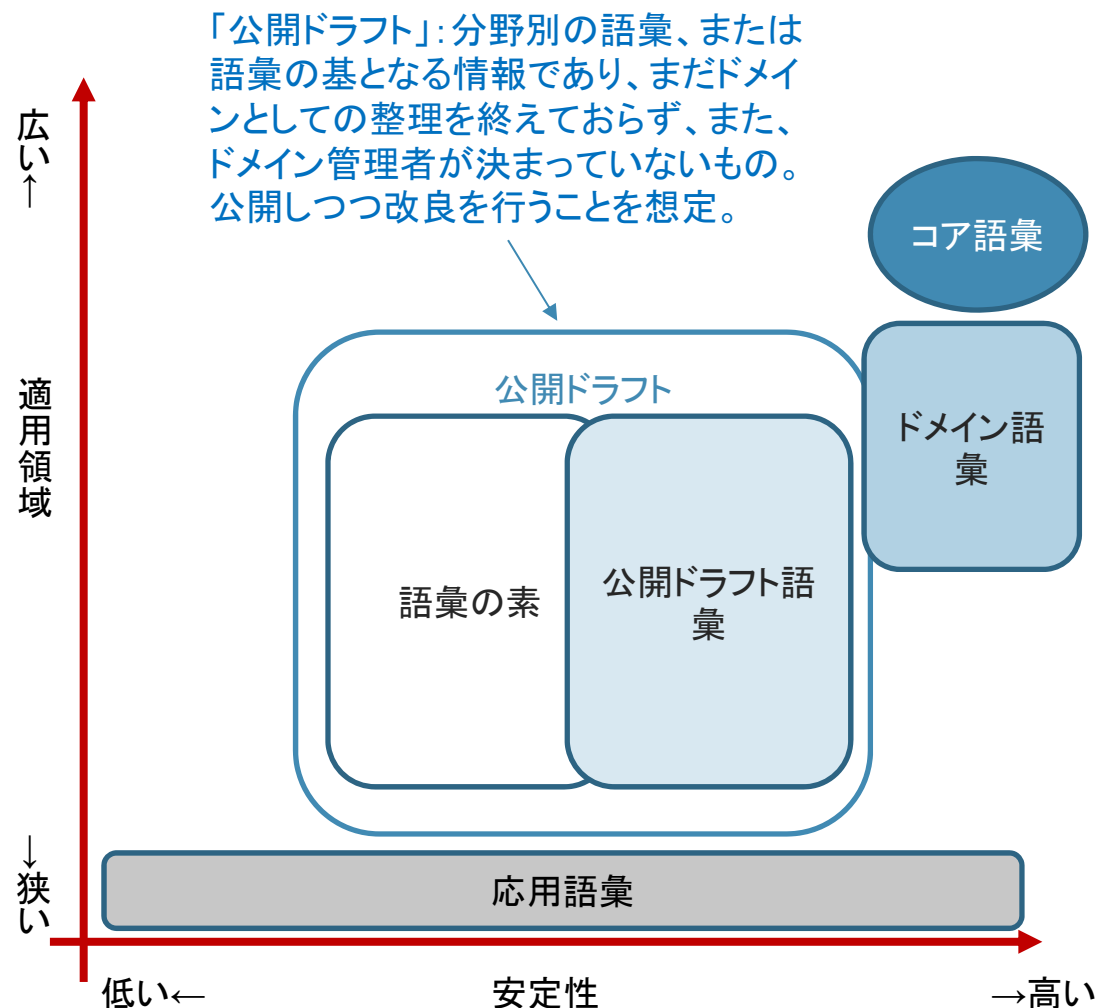
公開ドラフト等の公開 IMIパートナーとの関係



コア語彙とドメイン語彙



適用範囲×安定性の視点からの「語彙」の位置づけ



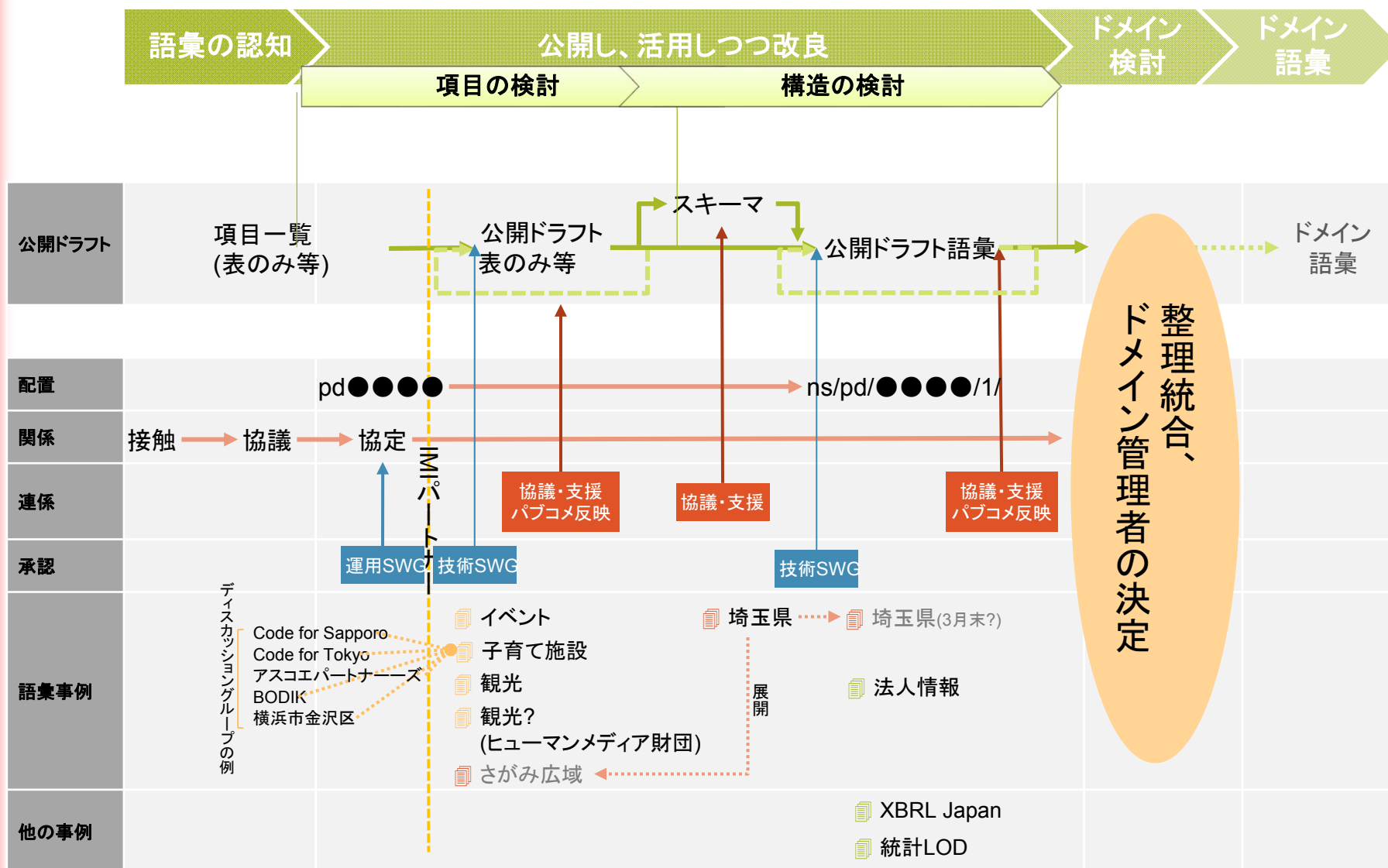
- コア語彙
 - 広い業務分野で共通に用いられる基本的語彙
- ドメイン語彙
 - 分野内で共通に用いられる語彙
 - 分野毎に管理者を設定
- 公開ドラフト
 - 活用しつつ、改良を行う語彙
 - スキーマを含まない、「語彙の素」も含む
 - ドメイン語彙への移行を意図（整理・統合と管理者の決定が必要）
- 応用語彙
 - データ作成者がその必要に応じて定義する語彙

語彙の種類と作成者・管理者等



種類	作成者	公開プロセス	メンテナンス	公開場所 (imi.go.jp/)
コア語彙	コア語彙SWG	コア語彙SWGが判断	コア語彙SWG	/ns/core/<version>
ドメイン語彙 (将来公開)	各ドメインの 担当組織	各ドメインの管理 者が主体。 コア語彙SWGが 必要に応じてレ ビュー	ドメイン管理者が主 体。 コア語彙SWGは必 要に応じてレビュー	/ns/domain/<ドメイン名>/<version>
公開ドラフト	コントリビュー タ(IMIパート ナ、自治体、 省庁等)	コア語彙SWGが レビューし、公開 可否を判断	コントリビュータとコ ア語彙SWGが協 議してメンテ	/pd/<pd番号> /ns/pd/<pd番号>/<version>
応用語彙	必要とする人	作成者の任意	作成者の任意	-

語彙のライフサイクル



ディスカッショングループ

独立したグループ

独立グループでほぼドメイン

ドメイン語彙構築プロセスのイメージ

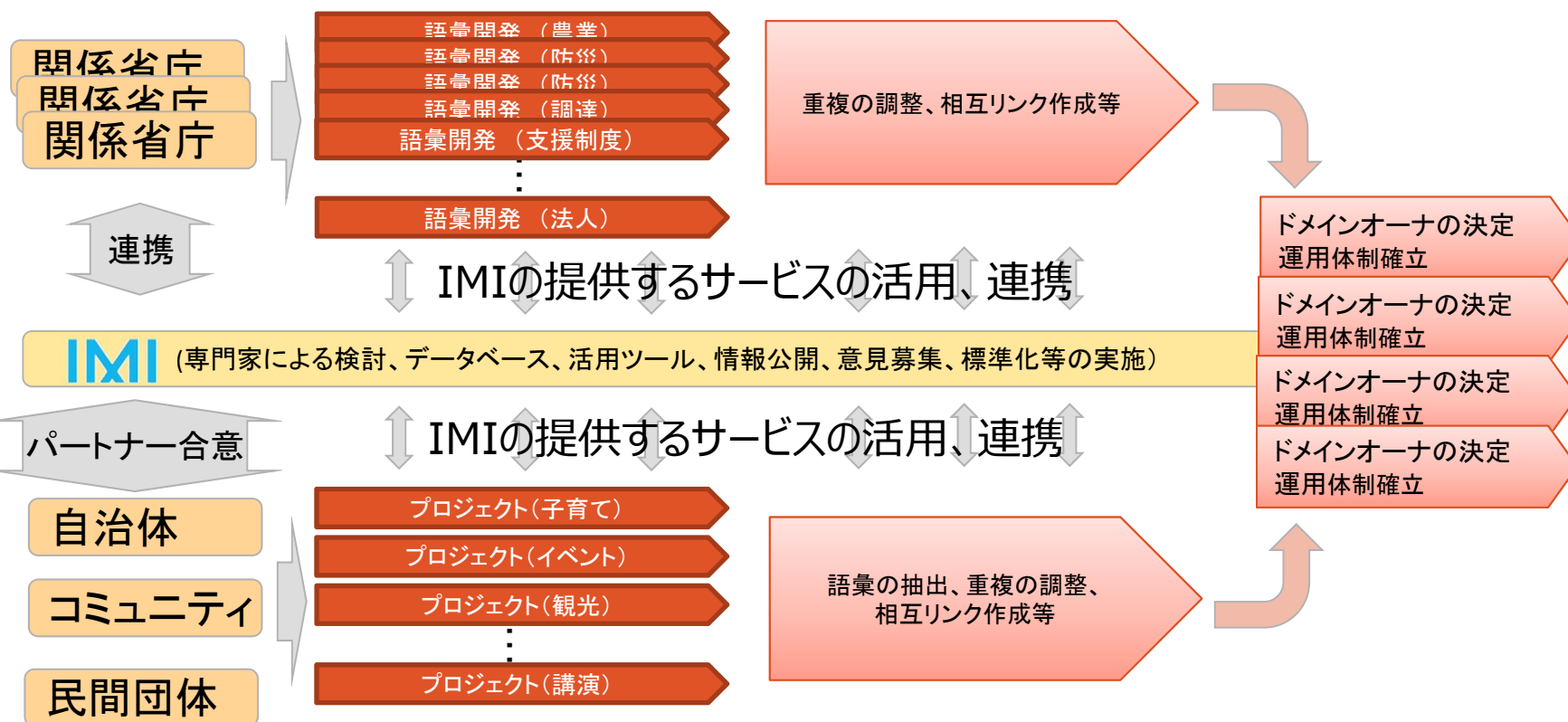


試行錯誤

整理統合

ドメイン確立

当該ドメインを所掌する省庁等が主導して語彙を開発



データ公開現場や、データ活用現場に近い組織(民間、自治体、コミュニティ等) が必要とする語彙を開発

- 要件
 - 主体的に「語彙」の開発を行っている組織／行おうとしている組織
 - 公開ドラフトの公開及びメンテナンスについて、コア語彙SWGとの協議を行こと
 - 適用領域が共通語彙基盤の趣旨に沿っている
 - 営利、非営利を問わないが、語彙等成果物はCC0で公開すること
- 選定のプロセス
 - パートナ協定締結希望者とコア語彙SWG（技術、運用）が協議し、上記要件を確認
 - コア語彙SWG（運用）は、上記要件の他、対応キャパシティが十分であるかなどの視点を含め総合的に判断
 - 公開ドラフトに関し、他のパートナーを交えた協議を依頼する可能性についても説明
 - 協議の結果、希望者とコア語彙SWG（技術、運用）の双方が適当と判断したら、協定書を締結