

第2回共通語彙基盤ワーキンググループ 議事概要

1. 日時・場所

2014年6月3日(火) 14:00~16:00

経済産業省本館1階 西共用会議室

2. 委員等

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

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 慶應義塾大学 SFC 研究所

次世代 Web 応用技術ラボ (AWA Lab.) 上席所員

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

(委員 50 音順)

オブザーバー

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)

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

国土交通省 国土地理院企画部

国立国会図書館 電子情報部 電子情報サービス課

独立行政法人 科学技術振興機構 (JST) 情報企画部

独立行政法人 科学技術振興機構 (JST) 知識基盤情報部

説明者

株式会社 日立製作所

独立行政法人 科学技術振興機構 (JST)

事務局

和田 恭 経済産業省 情報プロジェクト室 室長

宮里 孝則 経済産業省 情報プロジェクト室 室長補佐

平本 健二 経済産業省 CIO 補佐官

田代 秀一 独立行政法人 情報処理推進機構 国際標準推進センター長

3. 議事概要

3.1. 開会

3.2. 共通語彙基盤に関する政策的動向について

資料1に基づき、共通語彙基盤の現状（「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」CIO連絡会議決定）、コア語彙2.0、国際協調、海外動向（米国NIEM、欧州ISA、欧州Joinup、W3C、Schema.org）につき説明。2014年度は基盤整備と実利用での検証に加えて、利用モデル普及とドメイン展開についても前倒しで実施していく予定である。

3.3. コア語彙検討状況について

資料2及び別紙1～4に基づき、コア語彙について検討を進め、「コア語彙1.0」の用語を精査し、オープンデータと政府間のデータ交換の双方を視野に入れて、「コア語彙2.0」として取りまとめた。「コア語彙2.0」については、6月6日に公開することを武田委員長が宣言、各委員が合意した。

今後は、英語の用語・説明の整備、コードリストの整備、外部用語との参照表等の整備、文字セットの定義の明確化等を進めていく予定である。

3.4. データベースプロジェクト進捗報告について

資料3に基づき、データベースプロジェクトについて概要を説明。

- 米国NIEM、欧州ISA、Schema.org等の語彙開発・運用体制を参考に、データベースプロジェクトの将来的な運用体制について検討している。
- 財務、地理空間・施設、移動・交通、防災の4つのドメインについて語彙の整備を進めている。財務と地理空間・施設については、現在最終内部レビュー中、移動・交通と防災については、現在収集語彙の整理中であり、7月には全体の中間的結果が出せる見込みである。
- 2月10日から、ツールプロジェクトにパイロットシステムとAPIの提供を開始した。現在、ツールプロジェクト各社からドメイン語彙の追加申請を受け付けながら、語彙の拡充を実施中である。6月末までに提出されるツールプロジェクト各社からの報告書を踏まえ、コア語彙、ドメイン共通語彙についてフィードバックを行う予定である。

3.5. ツールプロジェクト進捗報告について

資料4に基づき、ツールプロジェクトについて概要を説明。

- 横浜市金沢区を現場とした「施設・サービス」のデータ化
- 浦安市都市整備部を現場とした「道路」のデータ化及び他の地理情報との連携
- 松江市を現場とした「観光情報」のデータ化
- 神戸市、千葉市、川口市を現場とした「育児関連制度」のデータ化
- 千葉市を現場とした「イベント」のデータ化

の5件を実施中であるが、間もなく完成する予定である。6月6日には共通語彙基盤のイベントを開催する。ここで、デモンストレーションを交えながら各プロジェクトについて紹介し、また、現在の問題点や今後に向けた展望などについて、自治体関係者や有

識者によるパネルディスカッションを行う予定である。

3.6. 科学技術振興機構 情報事業における用語辞書の整備について

資料 5 に基づき、科学技術振興機構（JST）情報事業における用語辞書の整備について概要を説明。

3.7. 質疑・応答

- ドメイン語彙の整備に当たって、ドメイン語彙定義のためのルールやガイドが必要ではないか。

⇒現在 IPA で実施しているデータベースプロジェクトでは、語彙のデータをどのように運用・管理するか等について提言をまとめるべく検討を進めており、8 月末には案をとりまとめる予定である。

- コア語彙 1.0 と 2.0 では、論理形式などで構成上の変化が見られる。これは、XML でシンプルに記述するのが難しいなど、日本語固有の問題を勘案してのものと理解している。

⇒論理形式と物理形式が分かれているのは、日本語的な難しさも一因としてあるが、オープンデータにも厳格なデータ交換にも十分に対応できるように考慮したためである。

- 共通語彙基盤 WG として、成果物などについてどの範囲まで規定しているのか。

⇒語彙データやそれを格納するためのデータベースの基本的構造と、語彙データ構築のためのテクニカルガイド、検討体制案等を取りまとめる。語彙については基本的な「コア語彙」は整備するが、ドメインについては、例として収集するという趣旨である。

現在、本 WG と並行して進めているプロジェクトから 8 月頃に一定の結果が出る予定だが、その結果を基に、この WG で検討をお願いしたい。

- ドメイン語彙について、収集・整備の進め方、防災ガイドライン等との整合性について知りたい。

⇒収集・整備については、各ドメインの専門家を集めた検討会で議論しつつ、関係するガイドライン等との整合性を考慮しつつ進めている。収集方法としては特定の利用シーンを設定し、その中で現れる用語を各ドメインについて 100 語程度抽出する方針で作業を進めている。

ドメイン語彙の検討の結果、コア語彙との関係の整理を行う予定である。

- コア語彙中の「定型住所」は、公的な住所表記に沿ってきちんと定義されているイメージであるが、例えば、京都の「上ル」や「通り名」のように、「定型住所」とは異なる表記を用いている地域では、住民が「定型住所」の方を知らない場合も想定される。その場合、「定型住所」は用いられず、ドメインの使用に流れてしまうのではないか。

⇒住所は共通性の高い概念なので、その全体がコア語彙にあってよいと考える。ドメイン毎に異なった表記や構造をとることは考えにくい。「定型住所」で表記できない場合は不定形な表記を用いればよいという考え方の下で、コア語彙は作成され

ている。

- 新 IMI フレームワークになって、語彙のレベルはコア語彙、ドメイン共通語彙、ドメイン固有語彙の3つになった。この3つのレベルの語彙を決めていく中で、それぞれのWGの担当範囲や、ある語彙がどのレベルに属するのかなどの検討も必要になってくると思うが、そのための体制はどのようになっているのか。
⇒資料3の4ページに、IMI運用体制（検討案）の図を掲載しているが、現在は検討を進めている段階であり、8月には案としてまとめる予定である。
その案を踏まえた上で今後どのように進めていくかについては、このWGで検討をお願いしたい。検討については9月から本格化するものと考えている。
- ツールプロジェクトの「道路」のデータ化について内容を確認する機会があったが、民間が中央省庁の持っている様々な技術を融合させ、共通語彙基盤を使ったツールを作っているという理想的なケースであった。「道路」のデータ化は「観光」のデータ化とも関わってくるので、同じ地域で実施するとより高い効果が得られるのではないか。また、共通語彙基盤によって自治体の転出・転入の手続きを容易に行うことが可能になるので、「制度」のデータ化、「イベント」のデータ化については、広範囲に、かつ、転出・転入を想定して実施することで、より分かりやすい形で国民にアピールできるのではないか。今後は、語彙の整備と併せて、民間事業者が使いたいと思うようなものや国民の興味を引くものなどの使用例を積み重ねていってほしい。また、点だけでなく、ラインやポリゴンの記述も地理空間情報の利用には重要なので、引続き検討を御願いたい。
- コア語彙2.0（別紙1～4）について6月6日に公開し、外部からの意見を広く伺うこととする。

3.8. 閉会

次回は2014年9月頃に開催予定。

9月頃を目途に、データベースプロジェクト、ツールプロジェクトの結果を基に、将来の共通語彙基盤のあり方について本格的に検討したい。

以上