



Infrastructure
for Multi-layer Interoperability

第3回IMI勉強会 コア語彙・技術仕様のバージョンアップについて

2019年3月20日
独立行政法人情報処理推進機構(IPA)

IMI共通語彙基盤が公開している技術

- ▶ IMI共通語彙基盤は、コア語彙の整備を進めるとともに、共通語彙基盤の活用や相互運用性の高い語彙検討のための技術仕様やツールなどを検討・公開しています。

種別	名称	概要
語彙	コア語彙	共通語彙基盤の基礎をなす、氏名、住所、組織等、あらゆる社会活動で使用される中核的な用語の集合。多くのシステム間で情報交換のための基礎となる語彙で、データ交換、オープンデータの二次利用等の効率化を図る。コア語彙2.4.2は約60のクラス用語と、クラス用語の性質や用語間の関係を記述するための約250のプロパティ用語から構成されている。
IMI技術仕様	DMD(データモデル記述)仕様	データ構造の定義とデータの各項目に対する入力制限や使用可能な文字コレクションなどのデータモデルを定義した仕様。コンピューターがデータを処理する際に有用な情報を、人がデータの内容を正しく理解できるようにするための情報をパッケージ化するための仕様を定義する。
	マッピングファイル仕様	DMD仕様が定めるデータモデルにおいて値が入る語彙の項目と、項目の代表的なラベル(データ項目名)を1対1で結びつけた定義(マッピング)を記述するための技術仕様。
	IMI語彙記法	IMI語彙及び語彙に含まれる語の定義を、特定の利用形式に依存しない中立的な形式で記述するための記法などを説明した技術仕様。構造化項目名記法を用いてIMI定義文書(複数のIMI語彙と1つのデータモデルからなる)を記述するための記法。
	構造化項目名記法	「人>氏名>姓」といったコア語彙の階層構造を、文字列のかたちで表現するためのルールを定めた仕様。表形式データのデータ項目名に階層構造をもたせることにより、表形式データと階層化されたデータの相互運用性を向上させる。
	使用文字規則の記法	「IMI語彙記法」の構文要素「使用可能文字制約」を用いてデータ項目に文字セットを割り当てる際に指定すべき内容と記法などを説明した技術仕様。
	文字セット定義の記法	文字セットを定義するための構文について規定した技術仕様。
ツール等	IMIツール	DMDの作成及び検証、データ形式変換及びデータ検証、コード・コードリスト作成、語彙作成のためのツール。DMD仕様3.0対応。
	IMIライブラリ	データ入力ツールやデータを利用するアプリケーション、CMS等で共通語彙基盤対応の機能を組み込む際に利用して、構造化データを容易に扱えるようにするためのソースを提供。

赤字は今年度バージョンアップ

2019年3月20日

第3回IMI勉強会 <https://imi.go.jp/>



コア語彙・技術仕様改訂に関する今年度IMI事業の経緯

- ▶ 今年度は、法人インフォをIMIのユースケースとして確立したものとするため、法人インフォの整備事業者と連携し「法人インフォ用語彙の整備」「DMDの整備」などを実施
- ▶ 実施にあたって発生した技術課題のうち、IMI提供技術の改定による解決が望ましい(今後も発生が予見される)課題に対し、仕様改定などの対応を実施
- ▶ 対応を実施した主要課題

► コア語彙

- コア語彙(ic:)ではic:コード型の実体に説明が記述できない
- ic:法人型には株主や役員を記述するプロパティがない
- Ic:文書型に日時を記述するプロパティがない

► 技術仕様

- データモデルとマッピングを個別に作成して利便性を上げたい
(今はDMDの中にデータモデルとマッピングの同梱が必須となっている)
- データモデルだけ見ても実際のイメージがわからないのでサンプルデータを格納したい
- データモデル記述からサンプルデータを自動生成できるようにしたい
- 個々のマッピングに対してコメントがかけないため設計者の意図が伝わらない



コア語彙の変更(v2.4.1 -> v2.4.2)

コア語彙(ic:)ではic:コード型の実体に説明が記述できない
ic:法人型には株主や役員を記述するプロパティがない
ic:文書型に日時を記述するプロパティがない

- ▶ ic:コード型に説明プロパティ(ic:説明)を追加
 - ▶ ic:コード型にコードの説明を記述するために説明プロパティが必要
 - ▶ 概念型を継承するクラス用語には説明プロパティが一律にあることが望ましいとの議論もあったが、ic:概念型に説明プロパティ(ic:説明)を追加すると、後方互換性に問題が生ずるため、ic:コード型へ説明プロパティの追加を決定
- ▶ ic:実体型に関与プロパティ(ic:関与型)を追加
 - ▶ ic:法人型に株主や役員などの関係者を記述するために関与プロパティが必要
 - ▶ ic:関与型はその役割と関与者を記述可能なため柔軟に関係者が記述可能。関係者の記述はic:実体型を継承するクラス用語に一律にあることが望ましいため、ic:実体型へ関与プロパティの追加を決定
- ▶ ic:文書型に日時プロパティ(ic:日時型)を追加
 - ▶ ic:文書型には日付プロパティ(ic:日付型)があるが、日付型では日時を記載することができないため日時プロパティが必要
 - ▶ ic:文書型に本来必要なプロパティは日時プロパティであることから追加を決定。



データモデルとマッピングを個別に作成して利便性を上げたい
データモデルだけ見ても実際のイメージがわからないのでサンプルデータを格納したい

▶ マッピングファイル同梱を任意化

- ▶ v3.0ではマッピングファイル同梱が必須となっているが、これはIMI語彙記法策定前の名残
- ▶ 形式が多様な実データとデータモデルを紐付けるマッピング情報は分離されていたほうが、様々な観点から望ましい

▶ サンプルデータの同梱を強く推奨

- ▶ v3.0ではサンプルデータの同梱を可能としているだけで推奨強度までは言及ていなかった
- ▶ サンプルデータの同梱を強く推奨することでデータモデル記述の理解を促進する



データモデルだけ見ても実際のイメージがわからないのでサンプルデータを格納したい
データモデル記述からサンプルデータを自動生成できるようにしたい

▶ exampleディレクティブの追加

- ▶ データモデル記述の中で各データのサンプルをexampleディレクティブを用いて記載することで、DMDに同梱するサンプルデータの自動生成等を可能とする

```
datamodel;  
#name "姓名"  
#description "姓名を漢字入力"  
#example "情報太郎"  
use ic:人型>ic:氏名>ic:姓名;  
  
#name "性別"  
#description "男性・女性・不明を入力"  
#example "男性"  
use ic:人型>ic:性別;  
  
#name "標準型日付"  
#description "生年月日をYYYY-MM-DD形式で入力"  
#example "2001-01-01"  
use ic:人型>ic:生年月日>ic:標準型日付;
```

データ項目ごとに
サンプルデータを記述可能



技術仕様の変更

マッピングファイル仕様v1.0 -> v1.0.1

個々のマッピングに対してコメントがかけないため設計者の意図が伝わらない

- ▶ 項目ラベルに対するマッピング先(構造化項目名)の指定にオブジェクト指定を追加
 - ▶ マッピングの際に個々の項目と構造化項目名の対応づけについて注意事項や備考などを記録しておくことは有用
 - ▶ 現在の仕様ではマッピング先の指定は構造化項目名のみのため、マッピング先の指定方法にオブジェクトを追加し、構造化項目名とコメントを併記可能とする

