



オープンデータ流通推進コンソーシアム

技術委員会報告

越塚登
技術委員会 主査
東京大学大学院情報学環 教授

技術委員会の活動趣旨

「オープンデータ推進に必要な 技術標準のあり方等の検討」

(オープンデータ流通推進コンソーシアム設立趣意書より)

構成員

主査：越塚 登（東京大学 教授）

副主査：武田英明（国立情報学研究所 教授）

中尾彰宏（東京大学 准教授）

平本健二（経済産業省 CIO補佐官）

深見嘉明（慶應義塾大学 特任助教）

ミッション1

種々のオープンデータを 流通させるための技術仕様の検討

標準データ規格

オープンデータの表現モデル

オープンデータを表現するためのボキャブラリ

標準API

オープンデータを取得・交換するための標準的な手法を規定したもの

ミッション2

国際標準化のための作業検討

ITU-T, W3C, ...

議論の流れ

総務省「情報流通連携基盤事業」の「標準API仕様」を
たたき台として、技術仕様の議論を進める。

3年間の検討方針

	1年目	2年目	3年目
検討の基本方針	標準仕様の検討 標準化準備	標準仕様の精査と標準化活動に関する検討	標準仕様の精査と標準化活動
主な検討事項	<ul style="list-style-type: none"> (1) 素案をベースとした標準仕様の検討（標準化すべき項目の整理） (2) 標準化の動向調査 (3) 標準化活動の体制検討 	<ul style="list-style-type: none"> (1) 標準仕様の精査 <ul style="list-style-type: none"> ・ データガバナンス委員会や利活用・普及委員会からの技術要求に対応 ・ 具体的なオープンデータを素材として仕様をブラッシュアップ (2) 標準仕様に付随する技術資料やツールに関する検討（利用ガイド・運用体制など） (3) 標準化活動に関する進展報告と意見交換 	<ul style="list-style-type: none"> (1) 2年目に引き続き、標準仕様のブラッシュアップ (2) 標準仕様に付随する技術資料やツールの精査（検討事項に基づく運用を含む） (3) 標準化活動に関する進展報告と意見交換

平成24年度の活動

今年度（平成24年度）の実施内容

■ 標準仕様のあり方に関する議論

- ▶ 標準仕様素案に関する意見交換
- ▶ 類似する既存の規格の整理と検討
 - ◇ 米国NIEM／欧州ISA／気象メタデータ（気象庁）／CKAN など
- ▶ 利活用・普及委員会などからの技術要件ヒアリングとその要件に対する標準仕様へのフィードバック検討
- ▶ ライセンス表現や、それに基づくデータ処理に関する技術的な検討（2年目以降）

■ 国際標準化活動に関する意見交換

- ▶ 1年目は、標準化の動向調査と体制に関するヒアリング、意見交換

**今年度（平成24年度）
想定アウトプット**

**活動報告書
（標準仕様を含む）**

委員会開催状況

■ 委員会は以下の日程で実施

開催回	議題
第1回 (2012/10/24)	(1) 参加者紹介および委員会の進め方について (2) オープンデータ流通に関する技術的な意識合わせ (3) 標準仕様の素案紹介
第2回 (1月初旬)	(1) 標準仕様素案に関する議論 (2) 類似する体系や仕様に関する調査結果報告
第3回 (2月頃)	(1) 標準仕様素案に関する議論 (2) 国際標準化動向に関する意見交換
第4回 (3月頃)	(1) 標準仕様素案に関する議論 (2) アプリケーション事業者から得られた、技術要件に関する課題検討 (3) 国際標準化のための体制や実施項目に関する意見交換 (4) 今年度活動報告について (5) 次年度活動計画について

技術委員会での議論を ふまえて

オープンデータのデータ形式

■ 本命的な技術

- ▶ 記述論理 (Descriptive Logic) に基づいた知識記述形式



(例)

- ▶ RDF (Resource Descriptive Format)
- ▶ ucR (unicode Relation Model)

現場サイドからのプリミティブな要望...

- ・ CSV形式
- ・ PDF形式...

マルチメディアデータの扱い

- ・ JPEG?MPEG?...

応用に特化しているデータ

- ・ 地理情報
- ・ CAD図面...

■ ボキャブラリ

- ▶ ボキャブラリの標準化の必要性
- ▶ その反面、Top Downな、世界全体を綺麗に整理した”Ontology”の構成は、不可能に近い困難さがある



- ▶ ボキャブラリ間関係を扱うメカニズムは必須 (同義語の扱い)

オープンデータのAPI

■ 2種類のオープンデータを考える

(1) 静的なオープンデータ

例：政府の統計データ、行政データ、等



データを丸ごとGETさえできればよい（APIへの要求はシンプル）

(2) 動的なリアルタイムオープンデータ

例：センサーデータ、メーターデータ、等

現在データを取り出すAPIが必要



- ▶他にも、データを全部丸ごと取得が難しい、ビッグデータの場合もAPIで選択的データ取得できることが必要

オープンデータAPI

■ 基本的には、WEBベースの仕組み



■ API

- ▶ RESTfulなURLによるAPI
- ▶ SOAPによる、検索命令をデータとして送信する

■ APIで交換されるデータインタフェース

- ▶ XML形式
- ▶ JSON形式
- ▶ SOAP形式

オープンデータをしたい主体が 使いやすい技術の確立



