



## 第二回 技術委員会資料

# オープンデータ流通推進コンソーシアム 技術委員会の論点

2013.1.9

オープンデータ流通推進コンソーシアム 事務局

# 技術委員会のミッションと実施内容(再掲)

## ミッション

### (1) オープンデータを流通させるための技術仕様のあり方を検討

#### ◇データ規格

- ◆ オープンデータの表現モデル(オープンデータはメタデータとして表現される。それを表現するモデル)
- ◆ オープンデータを表現するためのボキャブラリ(メタデータを理解するための標準的な辞書)

#### ◇API規格

- ◆ オープンデータを取得・交換するための手法に関する規定

### (2) 国際標準化のための作業検討

#### ◇ITU-T, W3C等の適切な組織での国際標準化

## 実施内容

### (1) 技術仕様のあり方に関する議論

- ◇ 仕様素案に関する意見交換
- ◇ 類似する既存規格・取組の整理と整合性検討
- ◇ 利活用・普及委員会などへの技術要件ヒアリングとその要件に対する技術仕様へのフィードバック検討
- ◇ データガバナンスのライセンス表現や、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)

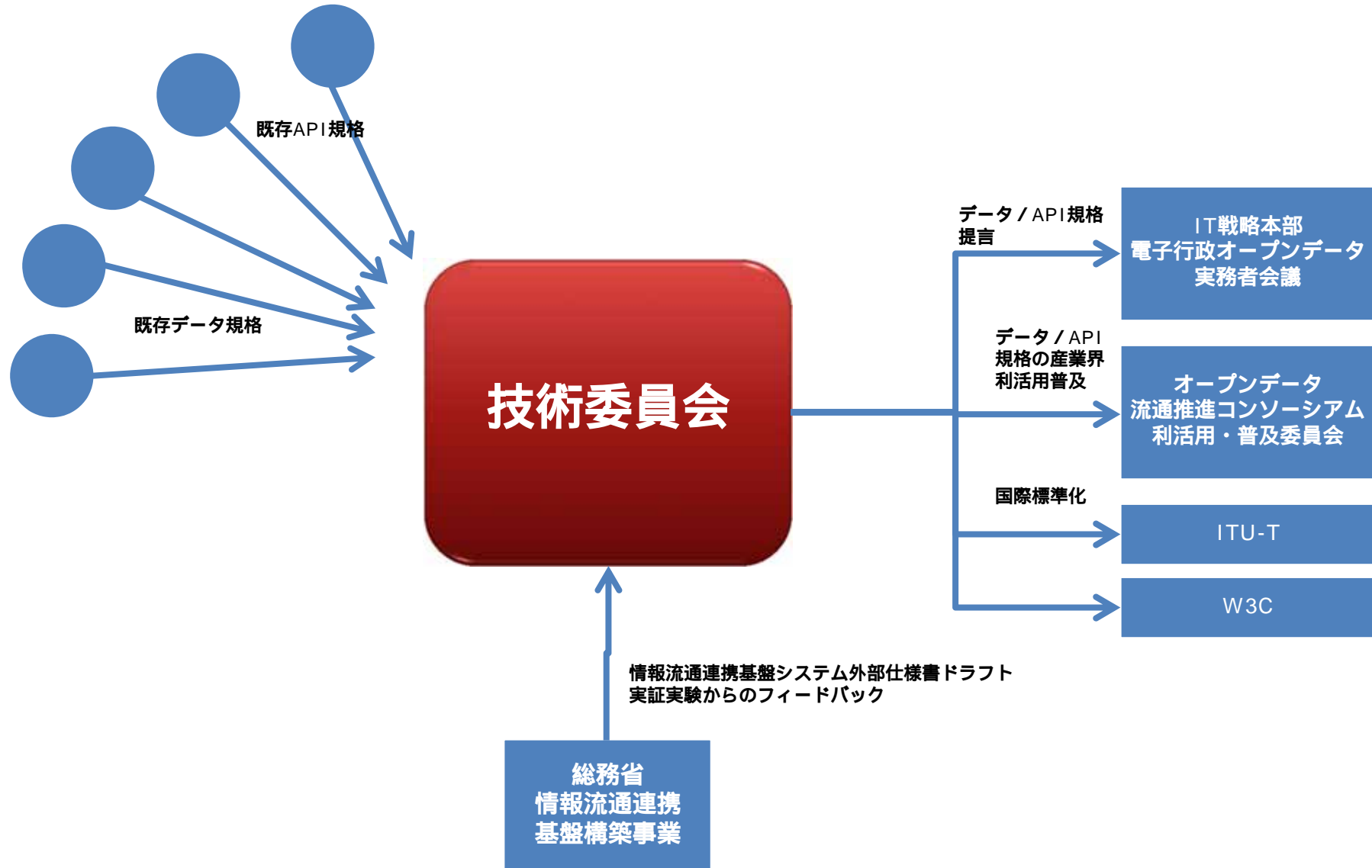
### (2) 国際標準化活動に関する意見交換

- ◇ 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

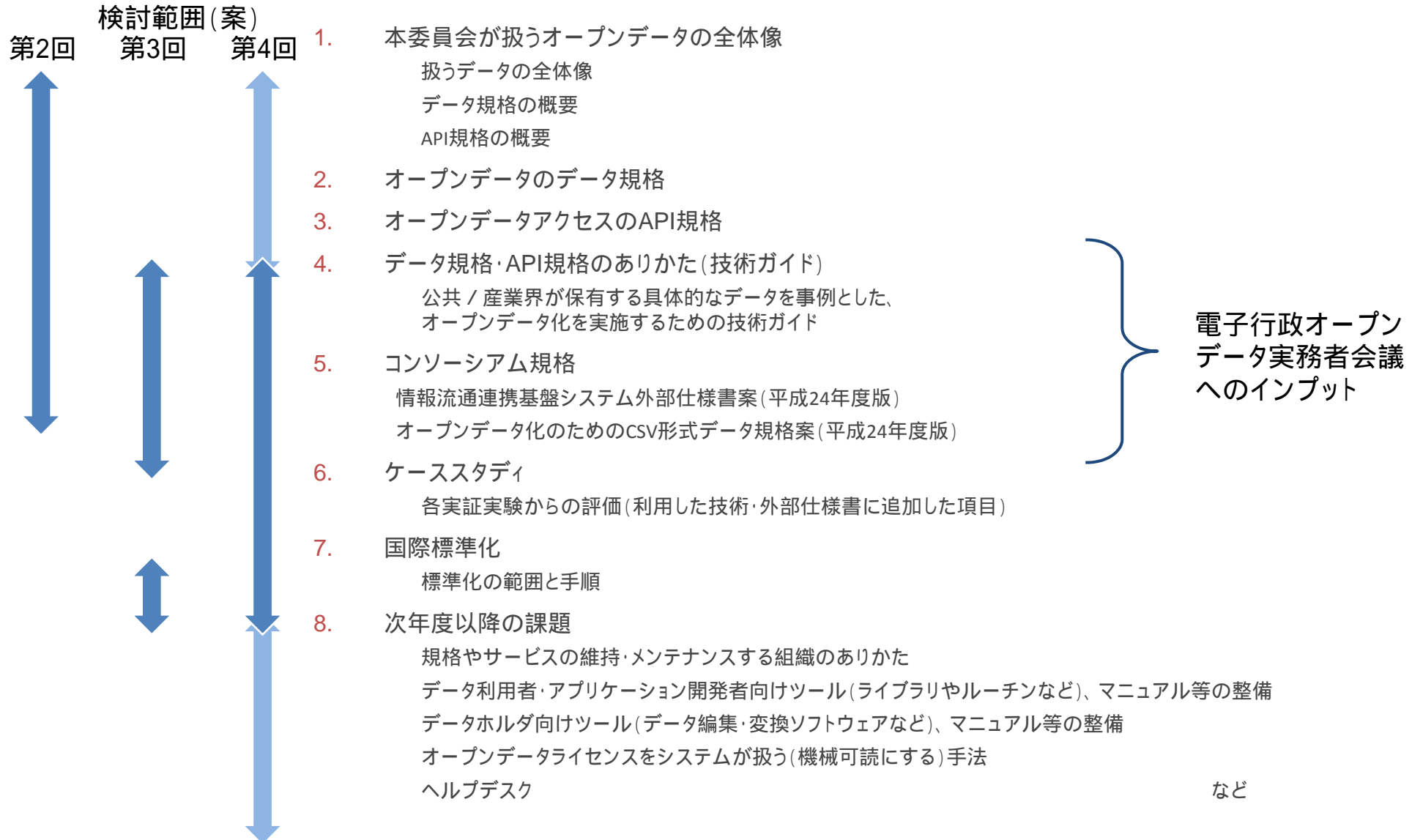
## 想定アウトプット

活動報告書(標準仕様を含む)

# 技術委員会のミッションと実施内容



# 技術委員会の論点



# オープンデータ技術に関する基本方針

---

## 適用範囲

公共機関(政府・自治体)のデータだけでなく、民間事業主体におけるオープンデータにも適用可能

静的 / 集約データだけでなく、リアルタイムデータにも適用可能

数値パラメータデータや小型ファイルデータから、巨大なビッグデータまで、多様な規模のデータに適用可能

## 目標

### 機械処理への適合性


- ◇情報の精度・正確性の向上、曖昧性除去の実現
- ◇高い処理性能・通信性能の実現

### 利用者の利便性向上

- ◇機械可読性の達成
- ◇プログラムが容易に書けることの実現

### 提供者の利便性向上

- ◇データを出すことが容易に



# 1. 本委員会が扱う オープンデータの全体像

# 想定する公共データの例

大分類	中分類	小分類	想定される2次利用用途
計画	各種研究会	研究会資料・議事録	
	政策方針	ロードマップ・政策内容	
	パブリックコメント		
	審議会資料・法案	審議会資料・国会提出法案	
	予算		政策投資分野の確認
	調達計画		
実施	所轄法令	法令・例規・審決・官報・判決	
	組織・所轄法人	構造・役割・権限関連資料や所轄法人の資料	
	レジストリ(台帳)	住民基本台帳・戸籍・登記・品質登録・規格・国有財産台帳・公有財産台帳	ヘルスケア・治療法(レセプト・健康情報・検診情報)・求人マッチング
	調達情報と結果	競売情報	ナビシステム(許認可情報)・マーケティング情報(許認可情報)
	インフラ・プラットフォーム	地図(2次元・3次元・行政区)・空間情報・空撮写真・道路・水路・航空路・地籍図・設備	農業経営・農業保険・ナビシステム(地理情報)・防災(交通事故)・都市計画・重機稼働理・公共機関稼働状況(バス・施設)・気象サービス
モニタリング 評価	センシングデータ	気象・火山・地盤・津波・海洋・交通・天文・環境・資源	農業経営・農業保険・防災計画・都市計画・ナビ情報(交通)・公共機関稼働状況(バス・施設)・気象サービス
	統計情報(白書・年次報告)	国勢調査・産業統計・土地利用・求人・人口・教育・観光・娯楽設備・事故統計・犯罪発生統計	農業経営・農業保険・防災(交通事故)・製品の安全対策
	会計報告		府省や自治体の会計報告
	政策評価		
アーカイブ	Webサイト分類型	e-Gov・国立国会図書館	
	分権検索型	国立国会図書館・特許庁・CiNii・文化財や美術アーカイブ	
	マルチメディアアーカイブ	NHK映像アーカイブス・国立国会図書館マルチメディア	
	テーマ別	国立国会図書館の震災アーカイブ	

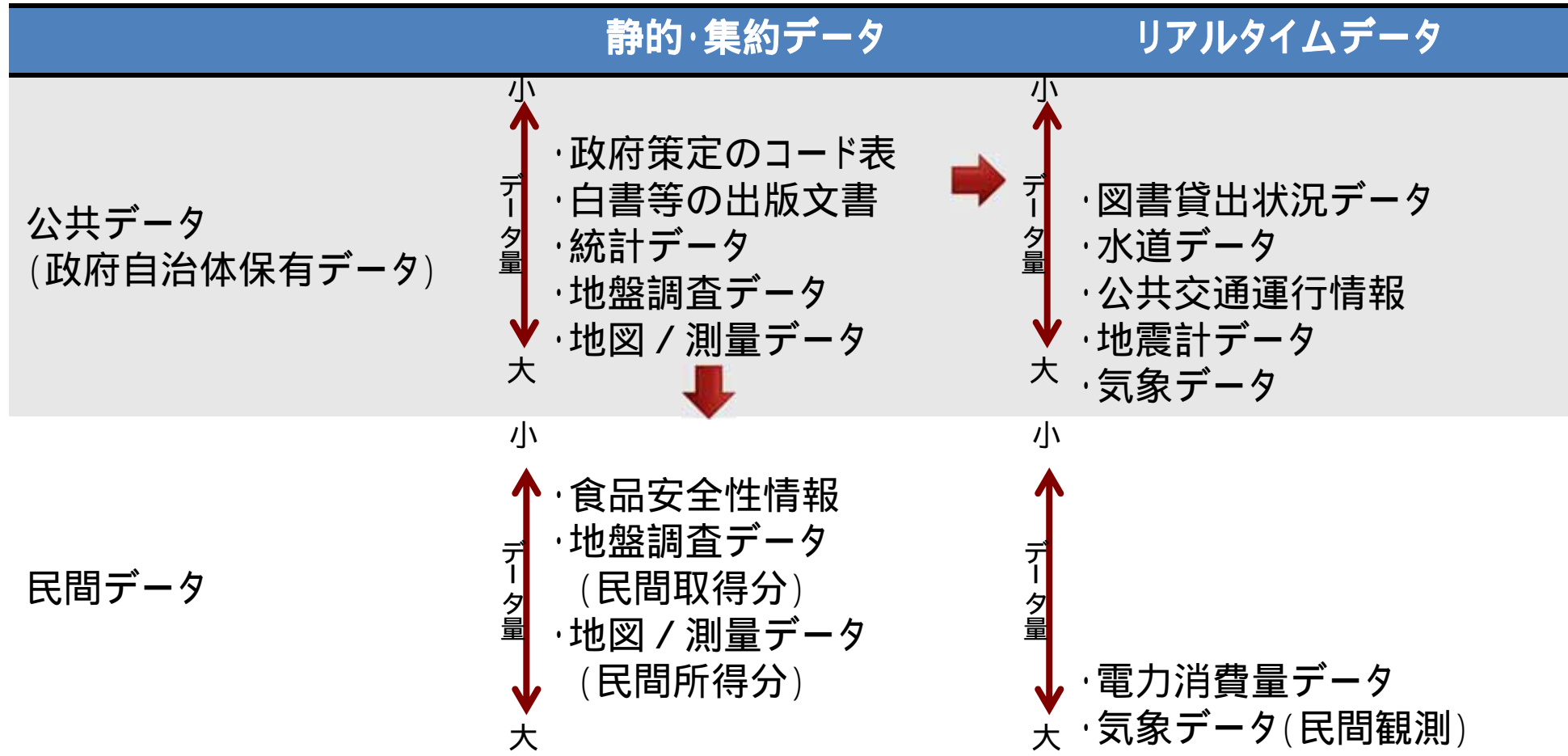
(電子行政オープンデータ実務者会議 データWG 小池構成員の資料を元に構成)

# 想定する民間データ(公共性の高いもの)の例

	大分類	小分類	補足
静的 集約データ	組織情報・経営情報	会社概要・決算情報・財務諸表	
	製品・サービス情報	商品・サービスのリスト・価格表	
		サービス提供時期に関する情報	開店 / 閉店時間・定休日など。
		地理情報	店舗や施設の位置情報 / サービス提供・非提供区域など。
	時刻表情報	Webでの閲覧は無償だが、元データは有償で提供されるケースが多い。	
物流・トレーサビリティ情報	製造物(生産物)の追跡情報	宅配物の追跡情報は、利用者にwebサービスとして公開されている。 入出荷の履歴は企業秘密であることが多い。	
リアルタイム データ	施設・サービス状況	発電量・電気消費量	東京電力「でんき予報」など。
		列車運行情報	現在は有償で提供される、または提供されないケースが多い。
		施設やサービスの稼働状況 / 混雑状況	企業秘密であるケースもある。



# 扱うデータの類型



## その他の類型化軸

データかコンテンツか(→次頁) / データの複雑さ / 地域性 / 国際性など

# データとコンテンツ

---

## データ

伝達、解釈、処理などに適するように形式化、符号化された情報

主たる解釈主体は「機械」

(例) 統計データ、公共交通運行データ、気象データ、等

## データのコンテンツ ← オープン「データ」の対象として考える

(例) 行政文書、白書、官報、など、主に自然言語による情報をデジタルデータ化したもの

伝達、解釈、処理などに適するような形式化や符号化がなされていない / できないデジタル情報

主たる解釈主体は「人間」

! 技術(データ規格やAPI規格)的には、コンテンツと同じ技術が適用される。

# 各種データ規格の概要

データ規格の洗い出しと整理

# データ規格の例 (静的・集約データの形式)

種類	主なファイル形式	標準規格等
文字	テキスト (.txt)	オープン (文字コードとしてISO/IEC 646, 8859等)
ドキュメント 表計算 プレゼンテーション	PDF (.pdf)	(Adobe社) ISO 32000-1
	Microsoft Office Binary(.doc, .xls, .ppt)	(マイクロソフト社独自)
	Office Open XML(.docs, .xlsx, .pptx)	ISO/IEC 29500
	OpenDocument(.odt, .ods, .odp)	ISO/IEC 26300
	CSV (.csv)	RFC 4180
電子書籍	EPUB (.epub)	国際電子出版フォーラムが規格化
画像	JPEG (.jpg)	ISO/IEC 10198
	JPEG 2000 (.jpeg)	ISO/IEC 15444-1
	PNG (.png)	ISO/IEC 15948
	SVG (.svg)	W3Cにより制定
映像	MPEG-4 (.mp4)	ISO/IEC 14496-14
音声	MPEG Audio Layer-3 (.mp3)	ISO 11172-3
地図	GML (.gml)	ISO 19136
気象	GRIB2, BUFR	国際気象通報式として規格化
書誌情報	MARC21	米国議会図書館が規定
セマンティクスデータ	XML (.xml)	W3Cにより制定
	RDF (.rdf)	W3Cにより制定

オープン化されるデータの規格は、これ以外にもある (特に、地理データなど特定専門分野固有のデータ規格)。これらは今後分類する対象として加える。

# 【資料】Five Star Open Data

Tim Berners-Lee氏がオープンデータに向けたデータ公開の段階を示したもの

<http://5stardata.info>

段階	データの形式	ファイル形式例
1	オープンライセンスで公開(形式不問)	PDF, JPEG, ...
2	コンピュータで編集可能な形式	DOC, XLS, ...
3	オープンに利用できるフォーマット	TXT, CSV, XML, ...
4	Web標準のフォーマット	RDF, XML, ...
5	他のデータへのリンクを含む	RDF (linked data)

## 5 ★ Open Data

Tim Berners-Lee, the inventor of the Web and Linked Data initiator, suggested a 5 star deployment scheme for Open Data. Here, we give examples for each step of the stars and explain costs and benefits that come along with it.



## データ規格の例(リアルタイムデータの形式)

種類	主な形式	標準規格等
バイナリデータ		Base64 [RFC3548], quoted-printable [RFC2045]等の規格によりテキスト化して伝送するケースが多い
文字データ	TXT	オープン(文字コードとしてISO/IEC 646, 8859等)
データ列	CSV	RFC 4180
構造化データ	XML	W3Cにより制定
	JSON	RFC 4627

# 各種API規格の概要

API規格の洗い出しと整理

# API整備の必要性

## API規格とは

ネットワーク上のリモートサーバー上にあるオープンデータに、プログラム(ソフトウェア)からアクセスするためのインタフェース

## データの性質からの要求

データサイズ上、データの選択的アクセスを必要とするケース

◇数百GB～数TBといった大型データなど

ライセンス上、データの選択的アクセスを必要とするケース

データ通信量やデータ処理量を最適化したいケース

◇情報処理能力・通信能力が低いモバイル端末上のソフトウェアからもオープンデータをアクセスする場合など

リアルタイムデータを扱いたいケース

## アプリケーションからの要求

データの利便性向上

◇データ列をパースして必要データを取り出すよりは、APIで取り出せる方がコーディングは容易

◇複数のオープンデータを串刺し検索してマッシュアップするためには、標準化されたAPIが効率よい(例:「カーリル」のような公共図書館貸出状態検索など)

◇既に多くのオープンデータサイトで、APIはリッチに整備されているものが多い(参考資料参照)

【参考】オープンデータシンポジウムにおける坂村顧問の講演



# API規格の例

## API規格

ネットワーク上のリモートサーバー上にあるオープンデータに、プログラム(ソフトウェア)からアクセスするためのインタフェース



データの種別や、データの使用方法に応じて、いくつかのAPIのタイプがある

タイプ	特徴	実現例
1	独自規格によるAPI	SMBなど
2	ファイル(データの集合)単位での直接アクセスするAPI	HTTP GET、FTP GETなど
3	機能(データの取得・更新など)を基本単位とするAPI	CGI、OAI-PMH、SOAP
4	データを基本単位とするAPI	REST
5	Web標準のフォーマットに基づく、データを基本単位とするAPI	SPARQL

## 【資料】モバイル回線の性能

モバイル回線の種類	回線帯域(規格での理論値)	回線帯域(実測値)
3G回線	14.4Mbps (HSDPA <sup>(*1)</sup> ) 42Mbps (DC-HSDPA <sup>(*1)</sup> )	0.5 ~ 3.7Mbps <sup>(*2)</sup>
Wi-Fi回線	11Mbps (IEEE802.11b) 54Mbps (IEEE802.11g) 600Mbps (IEEE802.11n)	
Mobile WiMAX回線	21Mbps (IEEE 802.16e-2005)	1 ~ 12Mbps <sup>(*3)</sup>
LTE回線	325.1Mbps	10 ~ 28Mbps <sup>(*3)</sup>

(\*1) HSDPA (High Speed Downlink Packet Access), DC (DualCell) -HSDPA: 第3世代移動通信システム(3G)プロトコルの一種で、ネットワークのデータ転送速度と容量を改善した規格の1つ。

(\*2) MMD研究所による2011年8月の測定結果報告による。 [http://mmd.up-date.ne.jp/news/detail.php?news\\_id=880](http://mmd.up-date.ne.jp/news/detail.php?news_id=880)

(\*3) MMD研究所による2012年11月の策定結果報告による。 [http://mmd.up-date.ne.jp/news/detail.php?news\\_id=1145](http://mmd.up-date.ne.jp/news/detail.php?news_id=1145)

## 【資料】既存のオープンデータサイトにおけるAPIの提供例

---

### Transport for London: Live bus arrivals API

1,900以上の停留所、700以上の路線、8,000以上のバスに関する到着予想時刻を、HTTP GETメソッドで取得。レスポンスはJSON形式で得られる。

### 復興・復興支援制度データベース(経済産業省)

制度情報・災害情報・支援の対象・支援等、リクエストの種類をURLのパスに含め、出力する結果数などのパラメータをクエリパラメータとして与える HTTP GETメソッドのAPI (REST API) を提供。

### Socrata Open Data API (SODA)

data.gov、City of Chicago、City of Seattle、State of Washington等が公開するオープンデータに対する REST APIを提供。

たとえばdata.govでは、認証・文書取得・ユーザ管理・データセット検索に関するAPIを提供している

### cosm (旧pachube): センサデータのクラウドサーバに公開・共有するためのプラットフォーム

環境(センシング空間)、計測機器、計測値(タイムスタンプ+値)、アクセス制御の認証情報、Notification、ユーザの6種類に対して、作成・更新・検索・取得・削除を行うRESTful APIを提供する。



## 2. オープンデータの データ規格

# オープンデータへの適用が考えられるデータ規格の洗い出し

---

## 主にデータのコンテンツに使われるデータ規格

PDF

RTF

Microsoft Office Binary

Office Open XML

OpenDocument

JPEG

PNG

MP3

MPEG4

(プレーン)テキスト形式

## 主にデータに使われるデータ形式

CSV

XML

GML

JSON

RDF

# PDF (Portable Document Format)

---

## 概要

アドビシステムズが開発および提唱する、電子文書に関するフォーマット。

特定の環境に左右されずに全ての環境でほぼ同様の状態で文章や画像等を閲覧できる特性を持っている。

2008年7月にISO 32000-1として標準化。

拡張子は“.pdf”

## 主な用途

ドキュメント

## 留意点

特定の企業に依存するファイルフォーマットである。

この形式のファイルからアプリケーションが必要なデータを取得することは、困難である。

# Microsoft Office Binary

---

## 概要

Microsoft Word/Excel/PowerPoint 2003までのファイル形式。

Microsoft Word/Excel/PowerPoint 2007以降は、Office Open XML Documentが標準のファイルフォーマットである。

2008年6月に Microsoft Open Specification Promise の下で仕様が公開された。

拡張子は“.doc”(Word)、“.xls”(Excel)、“.ppt”(PowerPoint)

## 主な用途

ドキュメント(Word)、表計算(Excel)、プレゼンテーション(PowerPoint)

## 留意点

特定の企業に依存するファイルフォーマットである。

この形式のファイルからアプリケーションが必要なデータを取得することは、困難である。

# RTF (Rich Text Format)

---

## 概要

Microsoftが公開する文書ファイルフォーマット。

プレインテキストに装飾やレイアウトのための制御用の文字列を付加した形式である。

◇フォントの指定や、文字の色・大きさや太文字などの装飾指定、画像の表示や中央揃え・箇条書き、表などの簡易レイアウトを行える特徴がある。

◇後に現れたHTMLなどとは表記方法が異なる。

拡張子は“.rtf”

## 主な用途

ドキュメント

## 留意点

特定の企業に依存するファイルフォーマットである。



# Office Open XML

---

## 概要

XMLをベースとしたオフィススイート用のファイルフォーマットの1つ。  
Microsoft Word/Excel/PowerPoint 2007以降の標準のファイルフォーマットである。  
2008年4月に ISO/IEC 29500として標準化された。  
拡張子は“.docx”(Word)、“.xlsx”(Excel)、“.pptx”(PowerPoint)

## 主な用途

ドキュメント(Word)、表計算(Excel)、プレゼンテーション(PowerPoint)

## 留意点

この形式で格納された表計算データを機械可読にするためには、各レコードの意味を定義する仕組みが必要である。

# OpenDocument

---

## 概要

XMLをベースとしたオフィススイート用のファイルフォーマットの1つ。  
構造化情報標準促進協会 (OASIS) が策定したファイルフォーマットである。  
2006年5月に ISO/IEC 26300として標準化された。  
拡張子は”.odp”(OpenDocument Text)、”.ods”(OpenDocument SpreadSheet)、  
”.odp”(OpenDocument Presentation)

## 主な用途

ドキュメント(OpenDocument Text)、表計算(OpenDocument SpreadSheet)、プレゼンテーション(OpenDocument Presentation)

## 留意点

この形式で格納された表計算データを機械可読にするためには、各レコードの意味を定義する仕組みが必要である。

# JPEG (Joint Photographic Experts Group)

---

## 概要

静止画像のデジタルデータを圧縮する方式(非可逆圧縮の画像フォーマット)の1つ。  
JPEGはISO/IEC 10198にて、JPEG 2000はISO/IEC 15444-1にて標準化されている。

拡張子は“.jpg”、“.jpeg”

## おもな用途

写真・画像

## 留意点

グラフやテキストデータを本方式で画像化しWebに掲載されることがある。  
このような画像からアプリケーションが必要なデータを取得することは、困難である。

# PNG (Portable Network Graphics)

---

## 概要

静止画像のデジタルデータを圧縮する方式(可逆圧縮の画像フォーマット)の1つ。  
ISO/IEC 15948として標準化されている。

拡張子は“.png”

## おもな用途

写真・画像

## 留意点

グラフやテキストデータを本方式で画像化しWebに掲載されることがある。  
このような画像からアプリケーションが必要なデータを取得することは、困難である。

# MPEG Audio Layer 3

---

## 概要

デジタル技術によって音響データを扱うための規格の1つ。

ISO 11172-3として標準化されている。

拡張子は“.mp3”

## おもな用途

音声データ

## 留意点

これはコンテンツのフォーマットである。

(アプリケーションがこのデータを解析してデータとして利用することは希である)

# MPEG-4

---

## 概要

動画・音声全般をデジタルデータとして扱うための規格の1つ。

ISO/IEC 14496として標準化されている。

拡張子は“.mp4”

## おもな用途

動画データ

## 留意点

これはコンテンツのフォーマットである。

(アプリケーションがこのデータを解析してデータとして利用することは希である)

# (プレーン)テキスト形式

---

## 概要

ISO/IEC 646, 8859, 2022等で規定する文字コードを並べた形式。  
仕様自体はオープンである。

## 主な用途

ドキュメント

## 留意点

アジア系では複数の文字コードのテキストファイルが共存している。  
このため、文字コードが分からなければ、文書を解釈できない。  
◇日本ではutf-8、ISO-2022-JP、EUC-JP、Shift\_JIS

# CSV (Comma-Separated Values)

---

## 概要

いくつかのフィールド(項目)をカンマ「,」で区切ったテキストデータおよびテキストファイル。  
2005年10月、RFC 4180により標準化された。

拡張子は“.csv”

## 主な用途

表計算データ

リアルタイムデータ

## 留意点

機械可読にするためには、各レコードの意味を定義する仕組みが必要である。

アジア系では複数の文字コードのテキストファイル共存している。文字コードが分からなければ、文書を解釈できない。

◇日本ではutf-8、ISO-2022-JP、EUC-JP、Shift\_JIS

レコードにコンマやダブルクォートが含まれる場合、エスケープされている場合でも、ソフトによって解釈が異なり、区切り方が変わることがある。



# XML (Extensible Markup Language)

---

## 概要

マークアップ言語作成のため、汎用的に使うことができる仕様、および仕様により策定される言語の名称。

W3Cにより策定・勧告されている。

拡張子は“.xml”

## おもな用途

ドキュメント・データ・地図情報またその組み合わせ

リアルタイムデータ

## 留意点

汎用的に情報を交換・流通させるためには、識別子体系やボキャブラリを規定しなければならない。

# GML (Geography Markup Language)

---

## 概要

Open Geospatial Consortium (OGC)によって開発された地理的特徴を表現するXMLベースのマークアップ言語。

ISO 19136として標準化されている。

データ構造は、後述するRDFに準拠している。

平成20年4月から国土地理院が提供している基盤地図情報は、この形式で提供されている。

拡張子は“.gml”

## おもな用途

地図データ

## 留意点

# JSON

---

## 概要

JavaScriptにおけるオブジェクトの表記法をベースとした軽量なデータ記述言語  
RFC 4627により標準化されている。

拡張子は“.JSON”

## おもな用途

リアルタイムデータ

## 留意点

汎用的に情報を交換・流通させるためには、識別子体系やボキャブラリを規定しなければならない。

# RDF (Resource Description Framework)

---

## 概要

ウェブ上にある「リソース」を記述するための統一された枠組み。

W3Cにより策定・勧告されている。

主語・述語・目的語の3つの要素でリソースに関する関係情報を表現する。

拡張子は“.rdf”

## おもな用途

ドキュメント・データ・地図情報またその組み合わせ

## 留意点

汎用的に情報を交換・流通させるためには、識別子体系やボキャブラリを規定しなければならない。

# オープンデータのデータ規格に関する検討事項

## 具体的にどのデータ規格を利用すべきか？

データの種別や性質、サイズ、データの利用目的や利用環境、プログラミングのしやすさ、データ作成のしやすさ、複数のデータ間の相互運用性、データ規格の知的財産権などの状況に応じて、最適なデータ規格を利用すべきではないか。

### ◇データの種類・性質・サイズ・利用目的・利用環境

- ◆ これらのターゲットから理想レベルのデータ規格と許容レベルのデータ規格を選択できる。

### ◇プログラミングのしやすさ

- ◆ データ処理の性能やデータ規格の複雑さ、データを扱うためのツールが普及しているかに依存する。

### ◇データ作成のしやすさ

- ◆ 人間がデータを編集するためのツールが普及しているか。(例: 文章や表形式のデータ編集ツールは広く普及している)

### ◇複数のデータ間の相互運用性

- ◆ 世界的な動向との一貫性・相互運用性があるか。

例: NIEM、ISAなど、政府 / 公共データのオープン化の枠組みでのデータ規格の動向

### ◇データ規格の知的財産権

- ◆ データ規格に付与されている特許・ライセンス条項はあるか。
- ◆ 画像を出力するツールが、データのヘッダ部分に著作権を明記させていたため、論文から画像を除かなければ公開できなかった、という事例もある。

# オープンデータのデータ規格に関する検討事項

## 各データ規格を利用する際に、上位層で規定すべき規定の整備

1. ボキャブラリ、コード(共通の使われる意味情報のコード化)
  - ◇RDFでのボキャブラリ
  - ◇RDF以外のデータ規格の場合でも、ボキャブラリに相当する規定は必要とされる。
2. ID・識別子(データから参照される個体(Entity/Object)の識別)
  - ◇RDFの場合、URI形式というIDの文法は定められているが、オープンデータを作成する時に、具体的なIDに関しては、そのデータホルダが定める必要がある。
  - ◇RDF以外のデータの場合は(例えばCSV)、IDやコードに関する規約はない。
3. 文字コード(特に漢字、東アジア特有の問題)
  - ◇人名、地名などの特殊な漢字の共通的な表現方法
4. データカタログ
  - ◇データ自体の識別
  - ◇データのメタデータ(データの由来・説明など)の記述方法

## 論点1: ボキャブラリ / 意味コード

---

データ形式に依らず、共通に利用されるデータ項目は、符号化されていることが望ましい。

自然言語であると曖昧性や記述ミスなどが残存する



### 符号化方式

意味コード(旧型技術)

◇データ項目を、特定の桁数の数字(固定長であることが多い)の並びに符号化する。

ボキャブラリ(新型技術)

◇データ項目を、可変長の文字列を用いて、特定の記述ルールに則った符号化

## 【参考】既存のボキャブラリ

名称	規定範囲	ネームスペース	規定例
RDF基本構造	RDFでデータ構造を表現するための基本的なボキャブラリ。	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>	rdf:subject(主語), rdf:predicate(述語)
RDFスキーマ	ボキャブラリを定義するためのボキャブラリ。	<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>	rdfs:subClassOf(サブクラス), rdf:range(値域), rdfs:subPropertyOf(サブプロパティ),
OWL	オントロジを記述するためのボキャブラリ。	<a href="http://www.w3.org/2002/07/owl#">http://www.w3.org/2002/07/owl#</a>	owl:sameAs(同義), owl:inverseOf(反意)
ダブリンコア基本要素	書誌情報を記述するためのボキャブラリセットであるが、Webリソースの属性を記述するために広く用いられている。ISO 15836にて標準化。	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>	dc:title(名前), dc:description(説明文), dc:creator(作者), dc:format(メディアタイプ)
DCMI語彙	ダブリンコア基本要素を拡張し、その意味を細分化したボキャブラリ。	<a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>	dcterms:alternative(代替タイトル), dcterms:audience(対象としている利用者)
FoaF	人や組織に関する情報をRDFで記述するためのボキャブラリ。	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>	foaf:familyName(姓), foaf:givenName(名), foaf:age(年齢)
SKOS	シソーラス、分類体系、件名標目表、タクソノミー、フォークソノミー、およびその他の同種の統制語彙のような概念体系の基本構造や内容を表現するためのモデルを提供するボキャブラリ体系。	<a href="http://www.w3.org/2008/05/skos-xl#">http://www.w3.org/2008/05/skos-xl#</a>	skos:definition(ボキャブラリの定義文), skos:broader(広義である), skos:note(ボキャブラリ定義に関するノート)
geoSPARQL	位置や形状に関するボキャブラリや、空間演算を行うための関数ボキャブラリが定義されている。	<a href="http://www.opengis.net/ont/geosparql#">http://www.opengis.net/ont/geosparql#</a> <a href="http://www.opengis.net/ont/sf#">http://www.opengis.net/ont/sf#</a> など	sf:wktLiteral(Well-Known Text規格の地理情報), sf:gmlLiteral(GML規格の地理情報)
ISA	EU加盟国間で公共データの相互運用をはかるためのボキャブラリ	<a href="http://w3.org/ns/person#">http://w3.org/ns/person#</a> , <a href="http://w3.org/ns/legal#">http://w3.org/ns/legal#</a> , <a href="http://w3.org/ns/locn#">http://w3.org/ns/locn#</a>	person:birthName(出生時の名称), person:placeOfBirth(出生地), legal:name(ビジネス上の氏名)



## 論点2 : 識別子 (ID)

### データが対象としている個体 (Object、Entity) を識別する識別子の運用が必要

基本的に政府が保有するリソースには、政府が定めた識別子が付与されているべきではないか

◇公共オープンデータはその識別子を使ってデータを記述するのが筋。

既に確立して、利用されている識別子体系はそれを利用する

RDFでは識別子はURI Formatを利用することが規定

◇URIは識別子の書き方を決めていても、識別子そのものの規定ではない

RDF以外のデータ規格での識別子は、依然として課題

URI以外の識別子形式の国際標準もある

◇ucode (ITU-T Rec. 642-1:2012)

◇PI : Place Identifier (ISO 19155:2012)

### 識別子の永続性

唯一性が保証できるか、識別子が使い回されないか、など

### RFIDやバーコードなどタグへの格納可能性

この場合、固定長であることが望ましい

## 【参考】RDFから利用可能なURI形式化できるID規格

種類	規格名	URI表現例	運営主体	説明	長さ	永続性	個体識別	RFID格納
汎用	ucode [ITU-T H.642.1]	urn:ucode:_012345 6789ABCDEF01234 56789ABCDEF	ユビキタス IDセンター	モノ・場所・概念などあらゆるものに付与できるIDである。IDの再利用を禁じているため、唯一性は永続的に保証される。	128bit			
物流	EPC SGTIN (Serialized Global Trade Item Number)	urn:epc:id:sgtin:45 12345.167890.2 urn:epc:tag:sgtin- 96:2.4512345.1678 90.2	GS1	商品を識別するコード。96ビットコードであるSGTIN-96では、ヘッダ(8ビット)、流通形態を表すフィルタ(3ビット)、パーティション(3ビット)、企業コード(20~40ビット)、アイテムコード(24~4ビット)、シリアル番号(38ビット)と続く。企業コードとアイテムコードは合計44ビットである。	96bit			
電子 データ	DOI (Digital Object Idnetifiers) [ISO 26234]	http://dx.doi.org/1 0.1021/jo0349227	国際DOI財 団 (The Internationa l DOI Foundation)	インターネット上のドキュメントに恒久的に与えられる識別子。サーバの移転によるリンク切れを回避するため、DOIディレクトリを経由させている。学術論文の分野で広く使われており、学術雑誌や論文誌の記事に付与されている。書籍のタイトルだけでなく、任意のページや図表、CDの1曲ごとに付与することもできる。	可変			
	UUID (Universally Unique Identifier) [ISO/IEC 11578]	urn:uuid:f81d4fae- 7dec-11d0-a765- 00a0c91e6bf6	なし(乱数)	分散システムにおいて、どこかが統制を取らなくても一意に識別できることを目的としたコード。現在よく利用されているのは、乱数に基づくversion 4である。ブログ等のコンテンツIDとして使われることが多い。	128bit	×		
企業・ 組織	企業コード [ISO 6523]	urn:oid:1.3.170.201 233049	ISOが定め たICD (Internation al Code Designator)	組織(企業)を識別するコードの付与方法をISO (ISO/IEC JTC1 SC32) が定めたもので、複数の企業コードや組織コードを包含することの出来るマルチコード。先頭の4桁がICDを識別する。それ以降の表記は、ICDが決定する。現在、150ほどのICDが登録されている。	可変	×	×	

## 論点3：文字コード(漢字の扱い)

### 漢字文化圏では、コンピュータ上での「漢字」の扱い方が大きな課題

特に、公共データにおいては、人名、地名などの漢字が、通常のJISコードやunicodeで表現できない部分がある。



### 代表的な多漢字文字コード

日本における行政システムベンダ特有の漢字コード

unicodeの拡張コード

e漢字(京都大学)

TRONコード(東京大学)

IPA漢字フォント(のコード)

今昔文字鏡コード

など

！ データ規格によっては、文字コードを固定しているものがあり、日本の文化とコンフリクトを起すケースもある

(例) RDFはUnicodeを使うことになっているが、Unicodeの問題点は他方で指摘されている。

## 【資料】行政データと文字コード

---

戸籍などの行政データに使用する文字のうち、文字コードに即した文字セットにない文字を、各自治体が独自に作成し、外字として設定している。

しかし、外字の割当が自治体ごとに異なるため、複数の自治体が管理する行政データを統合できない。

## 論点4:メタデータ・データカタログ

データ自体の識別方法やデータの由来・説明など(メタデータ)の記述方法が定まっていることが望ましい。



ファイル形式でまとめることが可能なデータ(小～中規模の静的・集約的データや、データのコンテンツ)

データの識別子

◇DOI (Digital Object Identifiers) [ISO 26234] が広く利用されている。

データのメタデータ記述方法

◇Dublin Coreのボキャブラリが広く利用されている。

! ファイル(データ)の部分データ(例:図、表、章・節など)の指定方法

データベース等に格納され選択的取得によって利用されるデータ(データベースに格納されるビッグデータや、リアルタイムデータ)

データベースに識別子やメタデータを与える

データ単体に識別子を与える(データPOST時に動的に識別子を与える、例:ucode)



# 3. オープンデータの API規格

# オープンデータへの適用が考えられるAPI規格の洗い出し

---

FTP

HTTP GET

CGI

REST

SOAP

SPARQL

# FTP

---

## 概要

FTP = File Transfer Protocol [RFC 959]

端末とサーバの間でファイル(ドキュメントや画像・動画など)を転送するための、代表的なプロトコルである。

## 主な用途

ドキュメントや画像・動画などを端末からサーバにアップロード(登録)する  
ドキュメントや画像・動画などをサーバから端末にダウンロード(取得)する

## 留意点

ファイル単位でのみアクセス可能である。

☆ファイルを部分的にダウンロードする仕組みはあるが、「このファイルの2行目を取得したい」というような要求はだせない。

ファイルサイズが大きい場合、通信帯域と端末側のリソース(メモリやディスク容量)を圧迫する。



# HTTP GET

## 概要

HTTP = HyperText Transfer Protocol [RFC 2616]

GETは、HTTPが規定するメソッド(サーバに対する要求)の1つであり、URL(\*1)で指定したリソース(通常はファイル)の内容を取得する要求である。

## おもな用途

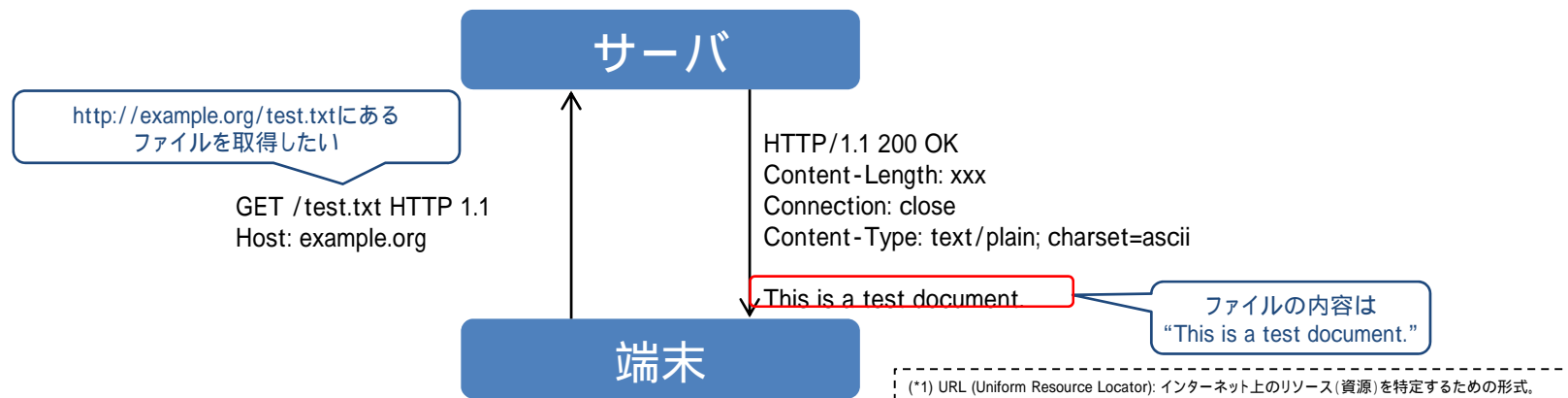
ドキュメントや画像・動画などのファイルを、サーバから取得する。

## 留意点

ファイル単位でのみアクセス可能である。

✦ファイルを部分的にダウンロードする仕組みはあるが、「このファイルの2行目を取得したい」というような要求はだせない。

ファイルサイズが大きい場合、通信帯域と端末側のリソース(メモリやディスク容量)を圧迫する。



# CGI

## 概要

CGI = Common Gateway Interface [RFC 3875]

ウェブサーバ上で任意のプログラムを動作させるための仕組みである。

端末は、プログラムの位置とパラメータを指定した要求をサーバに送る。  
サーバは指定されたプログラムを実行し、その結果を端末に返す。

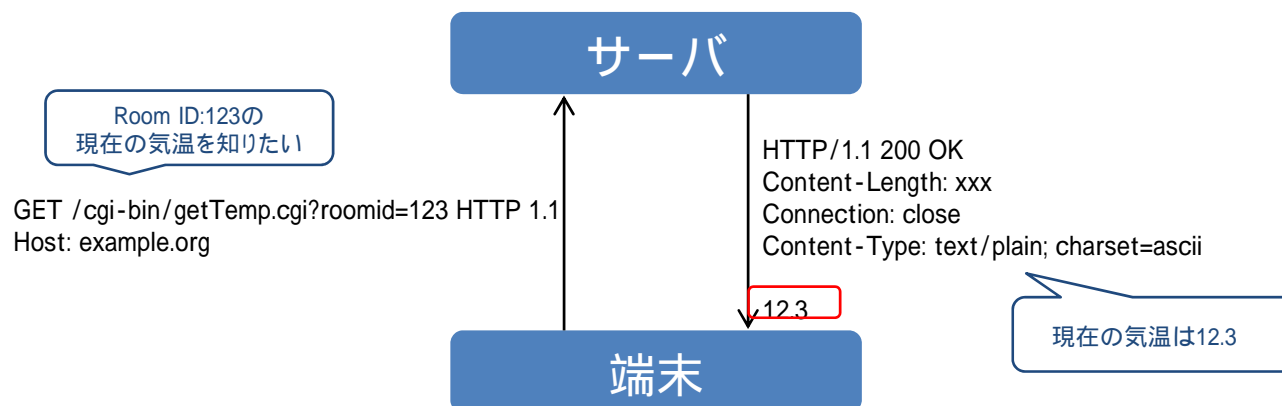
## おもな用途

ある建物の現在の気温を取得する。

ある環境の現在の温度分布図を取得する。

## 留意点

パラメータの与え方も結果の受け取り方も、プログラムに依存する。



# REST

## 概要

REST = Representational State Transfer

対象とするリソースをURLで指定し、HTTPの4つのメソッドGET, POST, PUT, DELETEを取得・登録・更新・削除の各操作に対応させてweb上のリソース(データ)を扱うスタイル。

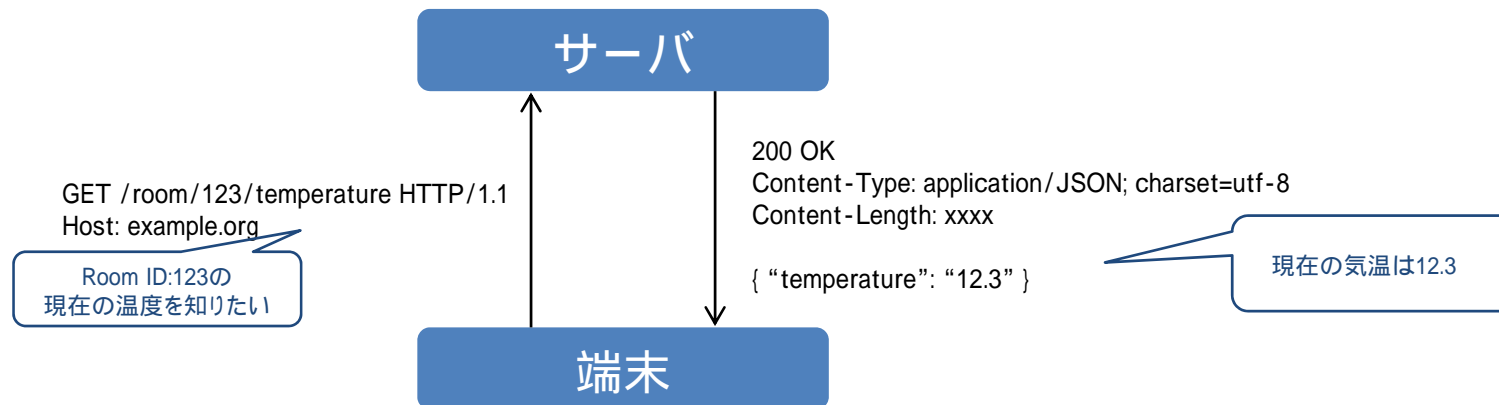
## おもな用途

ある環境の現在の温度を取得する。

ある物品の詳細情報を取得 / 登録する。

## 留意点

パラメータの与え方や結果の受け取り方は、提供されるサービスによって異なる。



# SOAP

## 概要

W3Cが標準化している、ソフトウェア同士がメッセージ(オブジェクト)を交換するためのプロトコルである。

※かつてはSimple Object Access Protocolの略とされていたが、現在は何かの略語ではないとされている。

HTTPやSMTP(\*4)など、さまざまな通信プロトコル上で利用できる。

交換するメッセージはXMLに準拠している。

従って、サーバへの要求もサーバからの応答も、XMLメッセージである。

## 主な用途

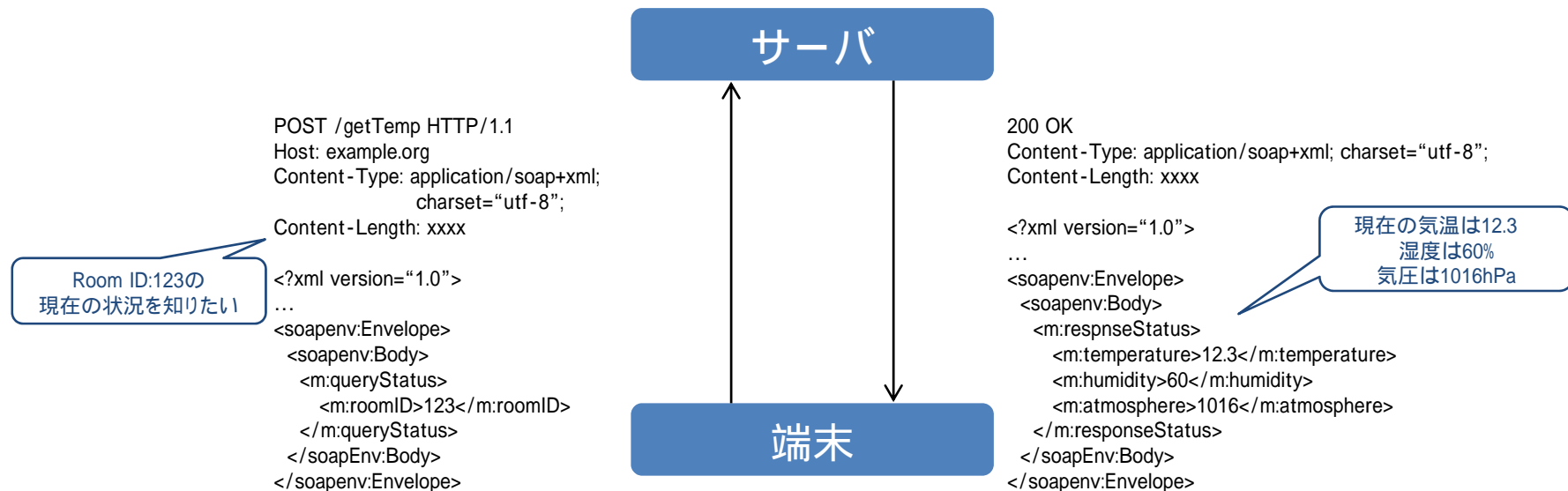
ある環境の現在の温度・湿度・気圧を取得する。

ある物品の詳細情報を取得 / 登録する。

見積もりの作成を依頼する。

## 留意点

汎用性を考慮してメッセージを抽象化しているため、交換するメッセージが複雑である。



(\*4) SMTP (Simple Mail Transfer Protocol): 電子メールを転送するプロトコル。RFC5321にて標準化。

# SPARQL

## 概要

SPARQL = SPARQL Protocol and RDF Query Language

W3Cが標準化している、RDFモデルに基づくデータを検索・操作するクエリ言語である。

## おもな用途

ある物品の詳細情報を取得 / 登録する。

ある環境の現在の温度・湿度・気圧を取得する。

名前に「白書」を含むデータの参照先を検索する。

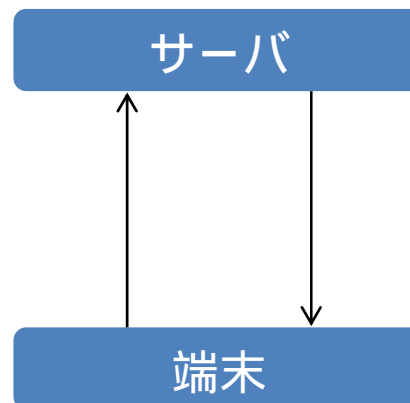
## 留意点

RDFモデルに基づくデータが対象である。

```
POST /sparql HTTP/1.1
Host: example.org
Accept: application/sparql-results+JSON
Content-Type: application/x-www-form-urlencoded
Content-Length: xxxx

query=PREFIX uc: <http://uidcenter.org/vocab/uc#>
PREFIX ex: <http://www.example.org/>
SELECT ?temp ?hum ?atom WHERE {
  ex:room123 uc:temperature ?temp ;
  uc:humidity ?hum ;
  uc:airPressure ?atom . }
```

Room ID:123の  
現在の温度・湿度・  
気圧を知りたい



```
200 OK
Content-Type: application/JSON; charset=utf-8
Content-Length: xxxx
```

```
{ "head": { "link": [
  "http://www.w3.org/TR/rdf-sparql-XMLres/example.rq" ],
  "vars": [ "temp", "hum", "atom" ] },
  "results": { "bindings": [
    { "temp": { "type": "literal", "value": "12.3" },
      "hum": { "type": "literal", "value": "60" },
      "atom": { "type": "literal", "value": "1016" } } ] } }
```

現在の気温は12.3  
湿度は60%  
気圧は1016hPa

# オープンデータのAPI規格に関する検討項目

---

## 各API規格の特質、想定されるデータのターゲット

FTP / HTTP GETは、ファイル単位のデータ取得を対象としている。

CGI / REST / SOAP / SPARQLは、データを直接取得することをターゲットとしている。

SPARQLは、RDFモデルに基づくデータがターゲットである。

## 各API規格をオープンデータの規格として利用する際に、更に上位層で規定すべき内容

例: ボキャブラリ・意味コード・IDなど

## APIを利用した際にネックとなる知財上の問題の有無

APIのライセンス条項・制限条項など。

## API規格の普及状況、ツール等の整備状況

今回例示したFTP、HTTP GET、REST、SOAP、SPARQLについては、ツールが普及している。

## API規格の世界的動向との一貫性、相互運用性


今日では、RESTベースのAPIが広く使われている。

◇復興・復興支援制度データベース(経済産業省)

◇Socrata Open Data API (SODA)

◇cosm (旧pachube)

◇Twitter API



## 4. データ規格・API規格の ありかた(技術ガイド)

オープンデータ化の手法について

# 技術ガイドのアウトプットイメージ

オープンデータの技術文書をアウトプットとする。

## ガイド類

- ◇「オープンデータ技術ガイド 第1巻: 公共文書のオープン化推奨ガイド(平成24年度版)」
- ◇「オープンデータ技術ガイド 第2巻: 表形式データのオープン化推奨ガイド(平成24年度版)」

：

## 技術規格類(→第5章)

- ◇「オープンデータ技術規格 第1巻: 情報流通連携基盤システム外部仕様書(平成24年度版)」
- ◇「オープンデータ技術規格 第2巻: オープンデータ化のためのCSV形式データ規格(平成24年度版)」

：

年度内に素案を作成し、それを公開してコメントを求める。



# 文章と統計データの混在した情報のオープンデータ化推奨ガイド

白書のような文章と統計データが混在した情報を、オープンデータの考え方に基づいてデジタルデータとして公開する場合の、技術的ガイドを示す。



## 【避けるべき方式】

印刷データ形式であるPDFやPostScript形式のみで公開する。

◇印刷用データとして公開することには問題ないが、再利用が困難なため、再利用を目的としたデータとしては適さない。

## 【許容レベル】

文章部分: .docファイル形式や、RTF (Rich Text Format)、Open Document Formatなどの、編集可能なファイル形式を用いる。

◇公開文書の正本としてPDF形式のデータを提供することは許容する。

◇見出しやアイテムイズなど、文書の構造が明らかになるスタイルを利用する。(空白やフォントサイズだけで整形しない)

◇統計の元データ(ローデータ)も提供できるのであれば、その参照先を明記する。

統計データ部分: CSVファイル形式で提供する。(→次頁)

## 【推奨レベル】

文章部分: 文章の構造や意味が判読できる、マークアップ型言語(例えばXML、XHTML)などの形式を用いる。

統計データ部分: RDFに基づき、標準ボキャブラリ、標準IDなどに準拠した方式のデータ形式を用いる。

# 表計算ファイル形式データのオープンデータ化推奨ガイド

CSVファイル形式は、コンマ(,)で区切られたデータ列のデータ形式を定めた規格であり、含まれている個々のデータの意味や規定を表す積極的な枠組みはない。

その他の表形式ファイル形式についても、個々のデータの意味や規定を表す枠組みはない。

視認性を高めるためにセルの結合を行うことがあるが、これは機械可読性を低下させる。

3次元以上のデータなどは、本来CSVで表せない。



## 【許容レベル】

セルの結合を行わない。

別ファイルで、表形式ファイルに含まれるデータの意味や単位などを指定する。

表形式ファイルのプロパティを利用して、データのタイトルなどを示す。

## 【推奨レベル】

表形式ファイル内に、RDFモデルにおけるボキャブラリやURIによるIDなどを記すことで、表形式データ内部の曖昧性を除去する。(→ この記述規格案は後述)



## 5. コンソーシアム規格

### (1) 情報流通連携基盤システム外部仕様書

# 外部仕様書の位置づけ・規定範囲

データをオープン化するためのモデルを提示することが目的。

自分のデータをオープン化するときの手法・システムの例示。

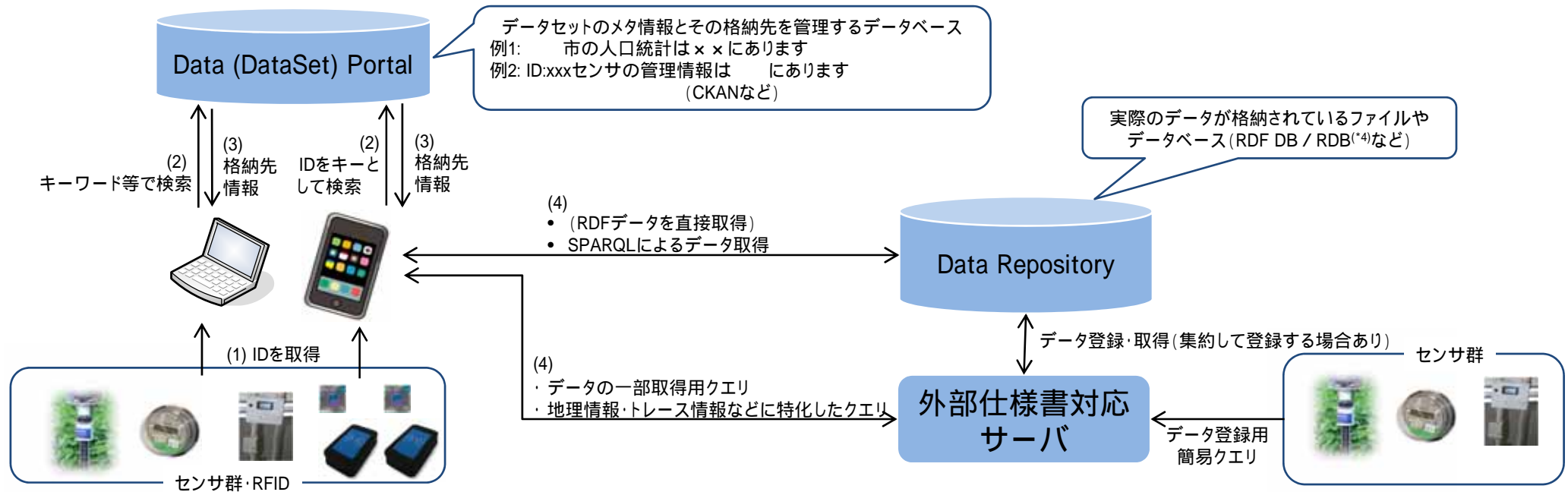
さまざまな利用シーンを対象とする。

◇データの直接取得・リアルタイムデータの取得・操作など。

◇識別子を読み取ることをNotificationとしてデータを取得するような利用法にも対応する。

SPARQLベースのAPIとRESTベースのAPIを提供する。

！利便性よりも、むしろ機能性や意味の正確性などに重点をおく。



(\*4) RDB (Relational DataBase): 1件のデータを複数の項目(フィールド)の集合として表現し、データの集合をテーブルと呼ばれる表で表す、データベースの一方式。



# 外部仕様書のAPI規格

## RESTベースのAPIとSPARQLベースのAPIを提供する。

RESTベースのAPIでは、データ検索・取得コマンドのレスポンスにRDF/XML等を利用することにより、RDFモデルに基づくデータとの互換性を保つ。

Streams APIに対応することにより、リアルタイムデータの送受信にも対応する。

機能名	概要	提供する理由
SPARQLベースのAPI		
SPARQL-based Command (←Raw Data Management Command)	SPARQL 1.1準拠のデータ操作APIを提供する。	RDFモデルに基づくデータに対するアクセスを提供するため。(既存の技術)
RESTベースのAPI		
Traceability/Realtime Data Management Command	トレースフォワード・トレースバックを含む、トレーサビリティに代表されるイベントを管理する機能。	対象応用分野の1つであるトレーサビリティで頻繁に発生する、トレーサビリティイベントを効率的に扱うため。
Geographical Data Management Command	GIS等地理情報処理を必要とするデータ検索・取得・操作機能。	実物や実環境を扱う上で、地理情報演算は重要であるが、演算が複雑であるため。
Notification Management Command (← Trigger Management Command)	データの登録・更新をNotificationとしてデータ利用者のシステムにコールバックする(Notification)仕組み。	センサ情報の共有を目指すcomsにも同様の機能が提供されているため。
Security Management Command	ユーザ・グループの管理と、データのアクセスルールに関する機能。	民間データでは課金処理を要するケースがあり、ユーザ管理は必要である。comsやSODAにもユーザ管理機能が提供されている。
Vocabulary Management Command	ボキャブラリ情報の登録・検索・取得に関する機能。	モバイル環境や組み込み機器に対応するため、データの一部を登録・取得できる機能を提供する。
Triple Management Command (← Data Conversion Command)	RDFモデルの主語・述語・目的語からなる基本データの登録・検索・取得に関する機能。	
Identification Resolution Command	IDをキーとしてデータを登録・検索する機能。	識別子を読み取ることをNotificationとしてデータを取得する利用法に対応するため。

# 外部仕様書のAPI規格 (1) SPARQL-Based Management Command

## 機能概要

SPARQL 1.1プロトコルに準拠した、RDFモデルに基づくデータの登録・更新・削除・取得・検索機能を提供

/api/v1/sparql/

◇SPARQL 1.1が規定するクエリを発行する。

/api/v1/rdf-graph-store

◇「SPARQL 1.1 Graph Store HTTP Protocol」に準拠した、グラフ単位での登録・取得・更新・削除を行う。

/api/v1/rawdata

◇RDF/XML, Notation3, N-Triplesで表現できる単位での、トリプルの登録・取得・更新・削除を行う。

## 備考

リクエストパラメータの形式をContent-Typeヘッダで指定する。

レスポンスパラメータの形式をAcceptヘッダで指定する。

## APIリスト

URL	HTTPメソッド	意味
/api/v1/sparql/	GET	SPARQL 1.1準拠のクエリを発行する
/api/v1/sparql/	POST	SPARQL 1.1準拠のクエリを発行する
/api/v1/rdf-graph-store	GET	RDFグラフを取得する
/api/v1/rdf-graph-store	POST	RDFグラフを追加する
/api/v1/rdf-graph-store	PUT	RDFグラフを更新する
/api/v1/rdf-graph-store	DELETE	RDFグラフを削除する
/api/v1/rawdata	GET	Raw Dataを取得する
/api/v1/rawdata	POST	Raw Dataを追加する
/api/v1/rawdata	PUT	Raw Dataを更新する
/api/v1/rawdata	DELETE	Raw Dataを削除する

## API例

### クエリ

```
POST /api/v1/sparql HTTP/1.1
Host: www.example.org
Accept: application/sparql-results+xml
Content-Type: application/x-www-form-urlencoded
Content-Length: xxx
```

```
query=PREFIX dc: http://purl.org/dc/elements/1.1/
PREFIX foaf: http://xmlns.org/foaf/0.1/
SELECT ?book ?name WHERE {
  ?book dc:creator ?who .
  ?who foaf:name ?name .
}
```

本の識別子と作者名のリストが欲しい

### レスポンス

```
HTTP/1.1 200 OK
Content-Length: xxx
Content-Type: application/sparql-results+xml
```

```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
  <head>
    <variable name="book"/>
    <variable name="who"/>
  </head>
  <results>
    <result>
      <binding name="book">
        <uri>http://www.example/book/book5</uri></binding>
      <binding name="name"><literal>John Michel</literal></binding>
    </result>
  </sparql>
```

識別子は "...book5"、この本の作者はJohn Michel (以下略)



# 外部仕様書のAPI規格 (2)

## Traceability/RealtimeData Management Command

### 機能概要

トレーサビリティに代表されるイベント管理に必要な機能を提供する。

### APIリスト

URL	HTTPメソッド	意味
/api/v1/events	GET	イベントを検索する
/api/v1/events	POST	イベントを登録する
/api/v1/events/<target>	GET	イベントを閲覧する
/api/v1/events/<target>/<property>	GET	イベントを閲覧する
/api/v1/events/<target>	PUT	イベントを更新する
/api/v1/events/<target>/<property>	PUT	イベントを更新する
/api/v1/events/<target>	DELETE	イベントを削除する
/api/v1/events/<target>/<property>	DELETE	イベントを削除する
/api/v1/trace/<target>	GET	トレースフォワード/トレースバックを実施する

#### 該当するイベントのリスト

1. ...0123から...0125が生成された
2. ...0123から...0126と...0127が生成された

### API例

#### クエリ

GET /api/v1/events?source=ucode\_00001C000000000000001000000100123 HTTP/1.1  
Host: example.org

イベント発生源がurn:ucode:\_0000...0123である  
イベントを取得したい

#### レスポンス

HTTP/1.1 200 OK  
Content-Length: xxx  
Connection: close  
Content-Type: application/json; charset=utf-8

```
{
  "events": [
    {
      "urn:ucode:_00001C000000000000001000000100800": {
        "http://uidcenter.org/ucr/vocab/event#type": [
          {
            "value": "urn:ucode:_0FFFDE000000000000000000000000001234567",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#source": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100123",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#destination": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100125",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#date": [
          {
            "value": "2012-03-07T12:00:00+0900",
            "type": "literal"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#place": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100A00",
            "type": "uri"
          }
        ]
      }
    }, {
      "urn:ucode:_00001C000000000000001000000100801": {
        "http://uidcenter.org/ucr/vocab/event#type": [
          {
            "value": "urn:ucode:_0FFFDE000000000000000000000000001234567",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#source": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100123",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#destination": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100126",
            "type": "uri"
          },
          {
            "value": "urn:ucode:_00001C000000000000001000000100127",
            "type": "uri"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#date": [
          {
            "value": "2012-03-07T13:00:00+0900",
            "type": "literal"
          }
        ],
        "http://uidcenter.org/ucr/vocab/event#place": [
          {
            "value": "urn:ucode:_00001C000000000000001000000100A01",
            "type": "uri"
          }
        ]
      }
    }
  ]
}
```







# 外部仕様書のAPI規格 (5) Notification Management Command

## 機能概要

オープンデータの更新に基づいて発生するNotificationを管理する機能を提供する

。

指定可能なNotificationの成立条件

- ◇any: 条件なし
- ◇gt: 指定した値より大きい
- ◇gte: 指定した値以上
- ◇lt: 指定した値より小さい
- ◇lte: 指定した値以下
- ◇eq: 指定した値と等しい

## APIリスト

URL	HTTPメソッド	意味
/api/v1/triggers	GET	Notificationを検索する
/api/v1/triggers	POST	Notificationを登録する
/api/v1/triggers/<target>	GET	Notification情報を取得する
/api/v1/triggers/<target>	PUT	Notification情報を更新する
/api/v1/triggers/<target>	DELETE	Notification情報を削除する
/api/v1/triggers/<target>/run	PUT	Notificationを起動・再開する
/api/v1/triggers/<target>/run	DELETE	Notificationを停止する

## API例

クエリ

名前が"NotificationA"であるNotificationの内容を知りたい

```
GET /api/v1/triggers?name=NotificationA
HTTP/1.1
Host: example.org
```

レスポンス

```
HTTP/1.1 200 OK
Content-Length: xxx
Connection: close
Content-Type: application/json; charset=utf-8
```

```
{"triggers":[{"ucode":"<urn:ucode:_00001C00000000000001000000100126>","name":"NotificationA","target":["<urn:ucode:_0001C00000000000000001000000100123>","<urn:ucode:_00001C00000000000001000000100124>"],"type":"any","callback":"http://www.example.org/?ucode=%U","run":"false"}]}
```

urn:ucode:0000...0123を対象としており、状態に変化があったら http://www.example.org/?ucode=0000...0123をコールバックする



# 外部仕様書のAPI規格 (7) Triple Management Command

## 機能概要

RDFモデルに基づくデータ操作を、RDF/XML等のエンコード/デコードなしで行う機能を提供する。

## APIリスト

URL	HTTPメソッド	意味
/api/v1/stats	GET	オープンデータを検索する
/api/v1/stats	POST	オープンデータを登録する
/api/v1/stats/<target>	GET	オープンデータを取得する
/api/v1/stats/<target>/<property>	GET	オープンデータを取得する
/api/v1/stats/<target>	PUT	オープンデータを更新する
/api/v1/stats/<target>/<property>	PUT	オープンデータを更新する
/api/v1/stats/<target>	DELETE	オープンデータを削除する
/api/v1/stats/<target>/<property>	DELETE	オープンデータの属性値を削除する

## API例

クエリ

GET

/api/v1/stats/ucode\_00001C00000000000000100000100124/dc\_title HTTP/1.1  
host: example.org

識別子がurn:ucode:0000...0124である  
(センサの名前は?)

## レスポンス

HTTP/1.1 200 OK

Content-Length: xxx

Connection: close

Content-Type: application/json; charset=utf-8

[ “ABC Meter” ]

ABC Meterである

# 外部仕様書のAPI規格 (8) Identification Resolution Command

## 機能概要

識別子とそれに関する情報の格納先との紐づけを管理する機能を提供する。

## APIリスト

URL	HTTPメソッド	意味
/api/v1/rs/<ucode>	GET	識別子解決 (簡易ucode解決) を行う
/api/v1/resolve/<ucode>	GET	ucodeからオープンデータの参照先を得る
/api/v1/resolve	POST	ucodeとオープンデータの参照先との組 (ucode解決情報) を作成する
/api/v1/resolve/<ucode>	PUT	ucode解決情報を更新する
/api/v1/resolve/<ucode>	DELETE	ucode解決情報を削除する

その情報は  
http://www.example.org/にある

## API例

クエリ

GET

/api/v1/rs/ucode\_00001C00000000000000100000  
0100123 HTTP/1.1  
host:example.org

識別子がurn:ucode:\_0000...0123に関する  
情報が格納されている場所を知りたい

## レスポンス


HTTP/1.1 200 OK

Content-Length: xxx

Connection: close

Content-Type: application/json; charset=utf-8

```
{
  "results": [
    {
      "X-UIDC-DATA": "http://www.example.org/",
      "X-UIDC-DATATYPE": 17,
      "X-UIDC-RETURNMASK": "FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF",
      "X-UIDC-TTL": 100,
      "X-UIDC-RESOLVEMODE": 0
    }
  ]
}
```



# 5. コンソーシアム規格

## (2) オープンデータ化のための CSV形式データ規格

# オープンデータ化のためのCSV形式データ規格案

## 背景

表形式ファイル内に、RDFモデルにおけるボキャブラリやURIによるIDなどを記すことで、表形式データ内部の曖昧性を除去するようなCSVデータの表現形式は確立していない。

関連する試みとして、以下のようなものがある。

- ◇RSS配信 RDF/CSV出力仕様書(自動車旅行推進機構)
- ◇Publishing RDF from CSV via Microsoft® Excel®, XML Schema, SAWSDL and XSLT  
(<http://www.sleepingdog.org.uk/semanticweb/rdf/excel/publishing/>)
- ◇XLWrap (<http://xlwrap.sourceforge.net/>)

上記を参考に規格素案を作成する。

## 表記法案

データ列の先頭列はID、先頭行はボキャブラリによるデータ種別定義とする。

冒頭で以下の定義をできる。

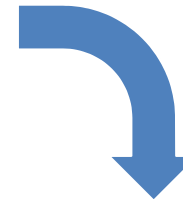
- ◇@prefix: ボキャブラリのネームスペースをQname表記するための宣言
  - ◆ Notation3記法に従って記述する。
- ◇\_:this このCSVファイルに関する説明
  - ◆ Notation3記法に従って記述する。
- ◇@comment: 人間の可読性確保のためのコメント行
- ◇@column: カラムのデータ種別定義行。この列以降はデータ本体である。
  - ◆ データ種別としてボキャブラリ名を利用する。ここが空の列は人間の可読性確保のためのコメント列である。
  - ◆ デフォルトの言語やデータタイプは、それぞれ"@", "^^"タグにつないで記述する(Notation3やRDF/XMLに倣う)。



# オープンデータ化のためのCSV形式データ規格案

## 表記例

	人口[千人]	面積[km <sup>2</sup> ]
東京都	13,216	2,188
神奈川県	9,072	2,416
埼玉県	7,149	3,797
千葉県	6,147	5,156
...		
合計	127,817	377,914



```

@prefix uc: <http://www.uidcenter.org/vocab#> .
@prefix ex: <http://www.example.org/> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix jix0402: <http://opendata.gr.jp/vocab/jisx0402#> .
_:this dc:title "人口・面積調査" .
    
```

prefix宣言  
表の名称や作者などのメタデータを記述

機械はこの行を解読しない  
(@comment行だから)

@comment	都道府県名	人口[千人]	面積[km <sup>2</sup> ]
@column		ex:population^^ex:KiloPeople	uc:area^^ex:unitKm2
jisx0402:13000	東京都	13216	2188
jisx0402:14000	神奈川県	9072	2416
jisx0402:11000	埼玉県	7149	3797
jisx0402:12000	千葉県	6147	5156
...			
	合計	127817	377914

機械はこの行を解読しない  
(1列目が空だから)

機械はこの列を解読しない(@column列が空だから)

