



オープン&ビッグデータ活用・地方創生推進機構

平成26年度 第2回技術委員会資料

オープンデータガイド 第2版案

2015.02.10

オープン&ビッグデータ活用・地方創生推進機構 事務局

Agenda

1. オープンデータガイド 第1版の概要
2. オープンデータガイド 第2版案の概要



1. オープンデータガイド 第1版の概要

※本節の内容は、VLED「成果公開」ページ
「オープンデータガイド第1版」概要版の抜粋
です。

<http://www.vled.or.jp/results/>

「オープンデータガイド」の全体構成

赤枠箇所は、今回の修正案の対象

■ 第I部 Getting Started: オープンデータをはじめよう

▶ 第1章 はじめに

◇本書の目的・対象読者・構成を示すとともに、本書が利用する用語の定義を行う。

▶ 第2章 オープンデータの動向と意義

◇国内外のオープンデータに関する動向を紹介するとともに、オープンデータの意義について解説する。

▶ 第3章 オープンデータの作成・公開手順

◇オープンデータの作成・公開手順を、6つのステップに分けて解説する。

■ 第II部 利用ルール編: オープンデータに利用ルールを設定しよう

▶ 第4章 オープンデータで必要となる利用ルール

◇オープンデータにおける利用ルールの重要性について解説するとともに、利用ルールに関する国際的な動向、日本政府における動向について紹介する。

▶ 第5章 オープンデータ利用ルールの概要

◇諸外国政府で採用が進んでいるCC-BYとCC0、日本政府で採用される政府標準利用規約（第1.0版）（案）の3つの利用ルールの特徴等について解説する。

▶ 第6章 利用ルールの比較と望ましいルール

◇CC-BY、CC0、政府標準利用規約（第1.0版）（案）の3つの利用ルールについて、情報利用者の視点、情報提供者の視点から比較を行う。また、比較結果を踏まえ、データをオープンデータとして公開する際に望ましい利用ルールについて解説する。

▶ 第7章 利用ルールに関する今後の見直しの方向性について

◇政府標準利用規約（第1.0版）（案）の今後の見直しにあたっての方向性について述べる。

「オープンデータガイド」の全体構成

赤枠箇所は、今回の修正案の対象

■ 第III部 技術編：機械判読に適したオープンデータにしよう

▶ 第8章 オープンデータの技術レベル

◇第3章に記したオープンデータの作成・公開手順のうち、技術的な事項について解説する。

◇機械判読性、データカタログ、識別子について解説し、機械判読性に関する指標を「オープンデータの技術レベル」として示す。

▶ 第9章 オープンデータのための技術的指針

◇表形式データ、文書形式データ、地理空間データ、リアルタイムデータのそれぞれについて、機械判読に適したオープンデータを作成するための技術的な指針を、識別子、ファイル形式、データの3項目について示す。

■ 付録

▶ 第10章 オープンデータに関する規格・ツール

◇機械判読に適したオープンデータを作成・編集する上で参考となる規格やツールをまとめる。

▶ 第11章 CKAN解説

◇データカタログシステムであるCKANの概要とその使用方法を解説する。

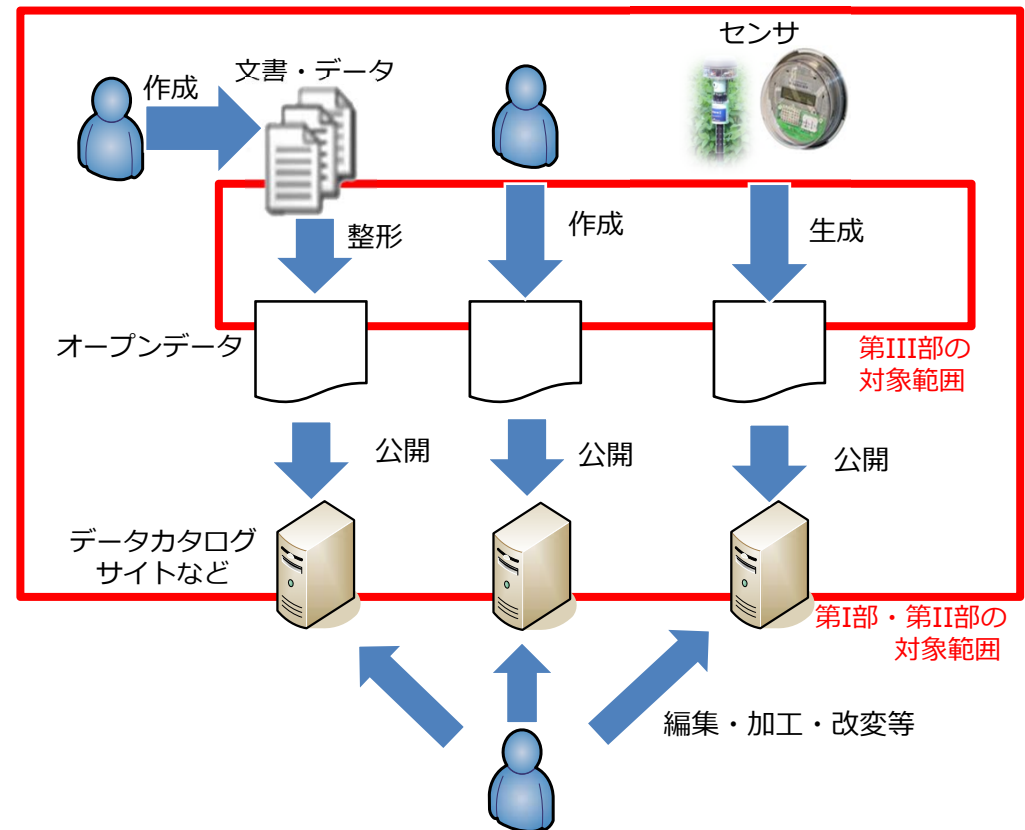
オープンデータの対象読者

■ 対象読者

- ▶ 現在保有しているデータや、これから作成するデータをオープンデータとして公開しようとする人。
 ☆主に国、地方公共団体、独立行政法人の職員を対象としているが、公共企業等の民間組織においても参考にできるものとして作成している。

■ 部ごとの対象範囲

- ▶ オープンデータの流れ
 ☆情報提供者が作成・公開する。
 ☆これに情報利用者がアクセスし、編集・加工・変更等する。
- ▶ 第I部と第II部の対象
 ☆データの作成段階から公開段階に至るまでに関与する人。
- ▶ 第III部の対象
 ☆機械判読性の高いデータを作成・整形しようとする人。



第1章 はじめに

■ 本書の概要

1. 本書の目的

◇前頁の通り。

2. 本書の対象読者

◇オープンデータを公開するまでの流れから、各部ごとの対象読書を示す。

3. 本書の構成

◇章ごとの構成（前述の通り）と、知りたい内容ごとに参照すべき章を示す。

4. 用語定義

◇本書が利用する用語を定義する。

第2章 オープンデータの動向と意義

■ 本章の概要

1. オープンデータに関する主な動向

◇ 日本政府・地方公共団体・海外でのオープンデータに関する取組を、それぞれ紹介する。

2. オープンデータの意義

◇ 「電子行政オープンデータ戦略」及び「二次利用の促進のための府省のデータ公開に関する基本的考え方（ガイドライン）」の記述より、オープンデータの意義を示す。

3. 本書におけるオープンデータの定義

◇ 「5★Open Data」や「電子行政オープンデータ戦略」、「電子行政オープンデータ推進のためのロードマップ」に記述に基づき、オープンデータを以下のように定義する。

- ◆ 「オープンデータ」とは、「営利目的も含めた二次利用が可能な利用ルールで公開」された、「機械判読に適したデータ形式のデータ」である。

第3章 オープンデータの作成・公開手順

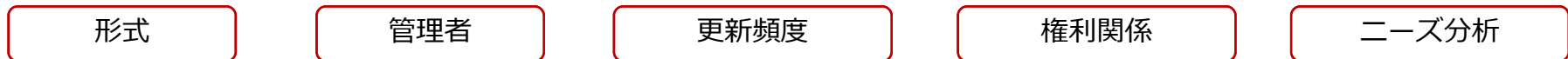
■ 本章の概要

- ▶ オープンデータの作成・公開手順を、以下の6つのステップに分けて解説する。
 ☆サーバの処理能力、データの信頼性、プライバシー・匿名化等、それぞれのステップにおける留意点も示す。

1. オープンデータ推進組織の設立

オープンデータの作成・公開を推進するための横断的組織を設立する。これ以降の活動は、この推進組織が中心となって進める。

2. 現状把握



3. 計画立案

オープンデータとする対象のデータと手法を明確にし、マイルストーンと計画を立案する。

4. 公開作業

計画に基づき、データを作成・整形し、公開の準備をする。

5. 公開・運用

オープンデータ管理のマイルストーンに基づき、ある程度の情報が登録された段階で公開し、システムの運用を開始する。

6. 改善点の洗い出し

利用者や作業担当者からのフィードバックを元に、改善点を洗い出す。

第3章 オープンデータの作成・公開手順

■ 本章の構成: 6つのそれぞれのステップを解説

1. オープンデータ推進組織の設立
2. 現状把握
3. 計画立案
4. 公開作業
5. 公開・運用
6. 改善点の洗い出し

■ 留意点: スモールスタートの原則

- ▶ 上記に掲げたオープンデータ推進組織の設立や現状把握等、すべての準備を完了してから実施することは必須でない。
- ▶ 比較的オープンデータにしやすいデータから着手するというのも一つの方法である。
 - ◇例えば、個別の部署において既に公開されている情報から着手する等。

3.1 オープンデータ推進組織の設立

■ オープンデータ推進組織が必要になる理由

- ▶ オープンデータの作成・公開作業は、各部署を横断する取組になる。
 - ☆オープンデータを作成・公開するに当たり、データを保有している各部署との連携・調整が必要になる。



- ▶ オープンデータの作成・公開を進めるに当たって、オープンデータを推進するための、各部署から独立した組織を設立することが望ましい。

3.2 現状把握

■ このステップでの実施内容

- ▶ 各部署が管理しているデータをまとめる。

■ 現状把握に当たり注目すべき項目

1. データの形式

- ◇紙 → 電子データがない資料を公開するには、紙をスキャンする必要がある。
- ◇電子データ → ファイル形式を確認すべき。

2. データの管理者

- ◇管理者は設定されているか。統一されているか。

3. データの更新頻度

- ◇データはどれくらいの頻度で更新されるか？
- ◇年に1回更新／月に1回更新／適宜更新等

4. データの権利関係 → 詳細は第II部参照

- ◇第三者が著作権等の権利を有するデータはあるか？
- ◇法令上の制約があるか？ 等

5. ニーズ分析

- ◇以下のようなニーズの高いデータからオープンデータとしての公開に取り組むことも有用。
 - ◆ 情報利用者から多く問い合わせられるデータ
 - ◆ 他の同様の組織で公開されているデータ

3.3 計画立案

■ このステップでの実施内容

- ▶ オープンデータの対象とするデータやその作成・公開手法を明確にする。
- ▶ マイルストーンを作成し、それに基づきスケジュールを立てることが望ましい。

■ 留意事項

1. データ形式・システムの準備計画

◇どのレベルの「データ」と「データカタログ」を準備するか、方針を策定（→ 8.4節参照）

2. 運用ルールの策定

◇データの入手手順・頻度を明確にする。

◇適宜更新される場合は、更新方法をルール化。

3. 利用ルールの設定

◇第三者権利問題や法令上の制約がある場合は、それを踏まえ、利用ルールの内容や適用範囲を整理する。

4. スモール・スタートの原則

◇作業は段階的に行い、完了したもののから順次公開できるように、マイルストーンを設定する。

◇年度ごとに目標・計画を立てることが望ましい。

3.4 公開作業

■ このステップでの実施内容

- ▶ 立案した計画に基づき、調達をかける等して必要なツールを揃え、オープンデータを作成・整形し、公開の準備作業を行う。

■ 留意事項

1. 公開時に明確にすべき情報
 - ◇メタデータ（どんなデータか？）
 - ◇アクセス方法（そのデータはどのようにして取得できるか？）
 - ◇利用ルール（そのデータはどのような条件で取得・利用できるか？）
2. 公開による影響
 - ◇オープンデータとして公開したデータは、全世界に対して公開される。
→ 海外からの問い合わせがあることも想定される。
3. データを公開するサーバに関する留意事項
 - ◇公開するサーバにどれくらいのアクセスがあると予想するか？
 - ◆ 予想外のアクセスが集中し、サーバの処理が追いつかなくなると、公開したデータに対するアクセス障害が発生する。
 - ◇リアルタイムデータを扱う場合の留意点
 - ◆ サーバの記憶容量を動的に消費するため、サーバの記憶容量の枯渇によりアクセス障害が発生する可能性がある。
 - ◇公開サービスを運用する業者や部署と、事前に協議しておくことが望ましい。
4. データの信頼性
 - ◇データの流通過程において、情報利用者による改ざん、情報提供者の意図しない編集・変更の可能性がある。
 - ◇その場合、情報提供者が正当性を主張するための方法は以下の2通り。
 - ◆ オープンデータの原則： 公開し続けることにより、原典を明示する。
 - ◆ 改ざんに関する技術的対処： 二次利用の促進のための府省のデータ公開に関する基本的考え方（ガイドライン）別添」より引用して解説する。
5. プライバシー・匿名化
 - ◇そのデータに個人を特定する情報が含まれていないか、確認する必要がある。
 - ◇必要に応じて、匿名化の手法を利用して、プライバシーを考慮すべきである。

3.5 公開・運用／3.6 改善点の洗い出し

■ 公開・運用

- ▶ ある程度のデータが登録された段階で公開し、オープンデータの提供を開始。
- ▶ 公開中は、情報利用者からのフィードバックが得られるように、アンケートページや問い合わせ窓口を用意することが望ましい。

■ 改善点の洗い出し

- ▶ 一定の期間ごとに、情報利用者から得られたフィードバックや、運用上の問題を整理し、改善点を洗い出す。
- ▶ 新規のデータを公開するタイミングで、改善点を洗い出すことが望ましい。
- ▶ その後、得られた改善点を解決するための計画を立案。
- ▶ 改善点を洗い出す際に、オープンデータの技術レベル（→ 8.4節）の向上、又は利用ルールの見直しを併せて行うことを推奨。

第8章 オープンデータの技術レベル

■ 本章の概要

- ▶ 第3章に記したオープンデータの作成・公開手順のうち、技術的な事項に関して解説する。

■ 本章の構成

1. 機械判読性に関する解説

◇機械判読に適したデータが必要である理由と、機械判読性に関する指標について解説する。

2. データカタログに関する解説

◇データカタログの意義を解説する。

3. オープンデータと識別子

◇オープンデータに識別子が必要である理由を解説する。

4. オープンデータの技術レベル

◇上記3項目をもとに、オープンデータの技術レベルを定める。

5. オープンデータの管理ポリシーとメタデータの付与方法

◇オープンデータを登録・管理する際に、メタデータを自動的に付与することができるならば、管理・登録のコストを軽減できる。そのような手法について解説する。

8.1 機械判読性に関する解説

■ 機械判読に適したデータとは？

- ▶ コンピュータプログラムがデータの論理的な構造を判読でき、構造中の値（表の中に入っている数値、テキスト等）を自動的に編集・加工・改変等できること。“Machine Readable”の日本語訳であり「機械可読」ともいう。
- ▶ 機械判読性の高いデータを提供することにより、コンピュータの解析に必要な情報利用者のコストを軽減できる。
- ▶ 機械判読に適さないデータ：PDFや画像データ
 - ◇ このデータをコンピュータに与えて、解析させるためには、以下の作業が必要。
 - ◆ 事前に人間がその画像にあるデータを表計算ソフトウェアに入力して保存する。
 - ◆ 画像認識等の技術により公開されているデータから数値やテキストを得て、それをコンピュータに与える。
 - ◇ これらは情報利用者に負担を求める方法であり、効率的でない。

■ 機械判読性に適したデータに関する留意点

- ▶ 機械判読に適したデータは、必ずしも人が読みやすいとは限らない。
 - ◇ 必要であれば、機械判読に適した形式と人に読みやすい形式の2種類のファイルを用意して公開することも考慮すべき。

8.1 機械判読性に関する解説

■ 機械判読性に関する指標: 5★Open Data

★1: この形式のファイルからコンピュータがデータを取り出すためには画像解析等の技術が必要であり、これは容易ではない。

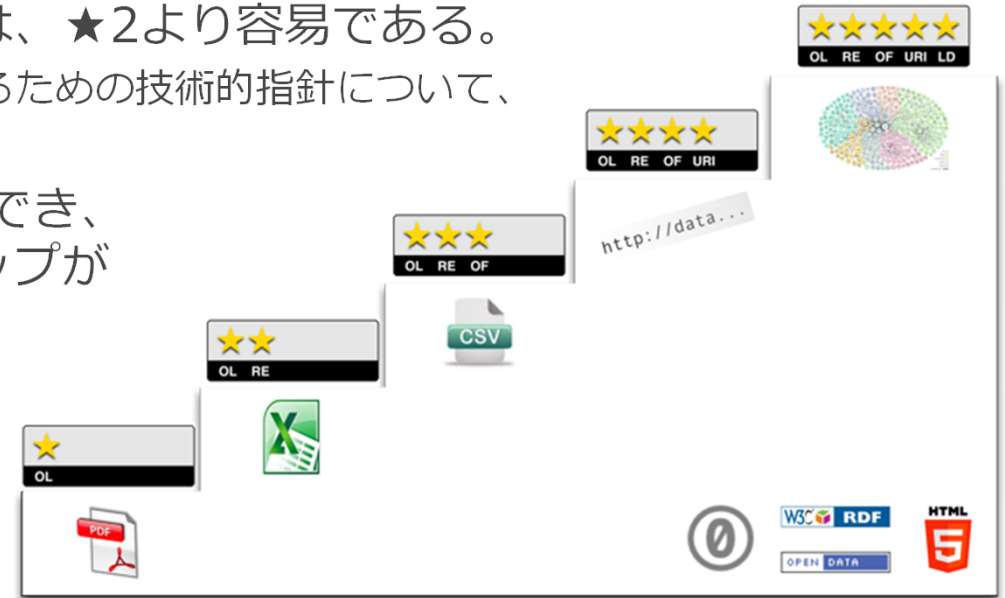
★2: この形式のファイルは構造化されているため、対応するソフトウェアを用意すれば、コンピュータはこれからデータを抽出できる。

◇一般に「機械判読性のあるデータ」とは★2以上のデータをいう。

★3: この形式のデータの解析方法は公開されているため、★3の形式のデータを解析するためのソフトウェアを構築することは、★2より容易である。

◇この形式のデータに対する機械判読性を高めるための技術的指針について、9.3節で詳しく述べる。

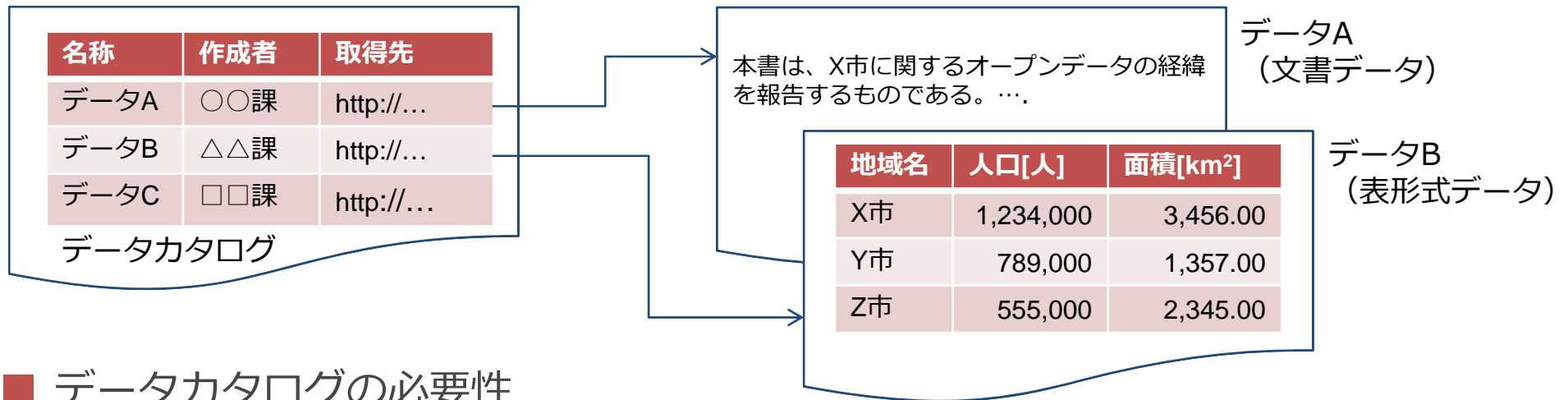
★4～: この形式のデータは、相互に接続でき、コンピュータによるデータのマッシュアップが容易になる。



5★Open Dataの指標 (<http://5stardata.info/>)

8.2 データカタログに関する解説

- 「データ」と「データカタログ」の関係は下図のとおり。



- データカタログの必要性

- ▶ 公開するデータが増加するにつれて、それらのデータを整理し、検索、一覧する機能に対する要求が高まる。このような要求に応えるものが、データカタログである。

- さまざまなデータカタログ

- ▶ データの名称、取得先等を表形式データとしてまとめたものも、一種のデータカタログである。
- ▶ 高機能なデータの管理・検索・一覧機能を提供するためには、データカタログシステムを導入する、あるいはRDF・SPARQL等の技術を利用したデータ・メタデータ検索機能を提供することが望ましい。

8.3 オープンデータと識別子

■ 識別子とは

- ▶ データやデータが対象とする実物や組織・場所等をコンピュータに識別させるための番号。

■ 識別子が必要である理由

- ▶ オープンデータは、コンピュータが読み取り解釈するためのデータである。これらのデータは、コンピュータが一意に識別できるべきである。
- ▶ オープンデータに含まれる実物や組織、場所等も、一意に識別されることが望ましい。これは、組織や場所が文字列で提供されている場合、以下のような問題が発生する可能性があるためである。
 1. 情報利用者のコンピュータは、表記の揺らぎにより同一の組織や場所を別物として解釈する可能性がある。（例：「中央一丁目一番地一号」と「中央1-1-1」）
 2. 情報利用者のコンピュータは、同一名称だが違う意味である組織や場所を、文字列だけでは識別できない。（例：全国各所にある「中央一丁目」）

■ 識別子とコード

- ▶ コードとは
 - ◇カテゴリ化される概念や事物に対して付与される番号。
 - ◇対象とする概念や事物を短縮して符号化するために規定される。
- ▶ 識別子とコードは異なる。
 - ◇コードには意味が付与されているが、識別子に意味が付与されているとは限らない。
- ▶ ただし、多くの場合、コードは識別子として機能する。

8.4 オープンデータの技術レベル

	Level 0	Level 1	Level 2	Level 3	Level 4
データ	PDFや画像ファイルをWebで公開する。	構造化されたデータを作成し、Webで公開する。 (XLS, DOC等)	非独占の（標準化された）形式で公開する。 (CSV等)	機械判読に適したデータを作成し、公開する。 (第9章参照)	RDF、XML等の技術を導入したデータを作成し、APIを実装して公開する。
データカタログ	存在しない。	カタログを表形式データ（CSV等）として作成し、公開する。	Level 1と同じ。	データカタログシステムを導入する。	RDFやSPARQL等を利用したメタデータ検索機能を提供する。
識別子	何らかの手段で識別されている。	Level 0と同じ。	Level 0と同じ。	URLにより識別されている。	グローバルな体系に基づく識別子を利用する。
必要なツール	Webサーバ	Webサーバ	Webサーバ	Webサーバ+データカタログシステム等	Webサーバ+データカタログシステム+情報流通連携基盤等

- データ・データカタログ・識別子のレベルを合わせる必要はない。
- Level1のメリット
 - ▶ 情報利用者は、画像解析等の処理をすることなく、直接データを取得できる。
 - ▶ データのありか等のメタデータを電子的に入手できるようになる。
- Level2のメリット
 - ▶ データに対する機械判読性が高まる。
- Level3のメリット
 - ▶ データの解釈効率や検索性が向上し、情報利用者のデータ利活用の効率が向上する。
- Level4のメリット
 - ▶ 他のデータとの横断検索等も容易になり、情報利用者によるデータ利用の幅が広がる。

8.5 オープンデータの管理ポリシーとメタデータの付与方法

■ オープンデータの登録ポリシーには2通りある

1. 集中登録方式

◇システム管理者や、オープンデータにすることをを行う独立した組織が、各組織・部署からデータを集めて公開する手法。

◇この方式では、データを集める際にメタデータもまとめて収集することが望ましい。

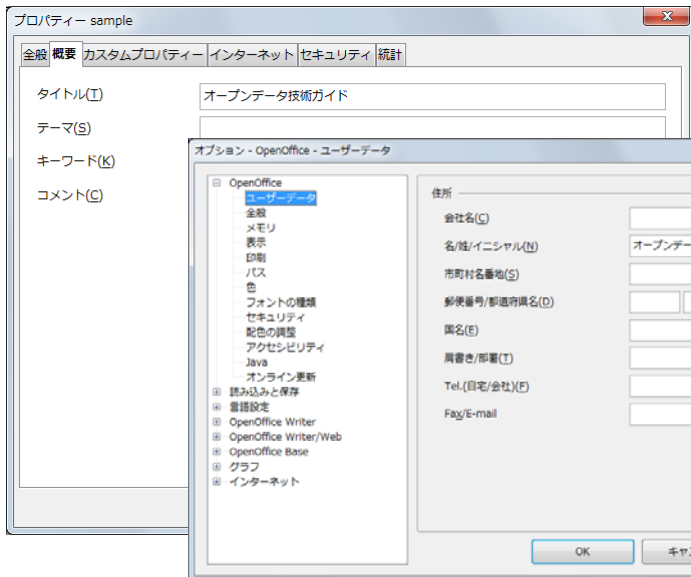
2. 分散登録方式

◇各組織・部局が自ら、何らかのシステムを利用して直接オープンデータを登録・管理する手法。

◇この方式では、担当組織がオープンデータを作成する際に、メタデータもまとめて作成できることが望ましい。

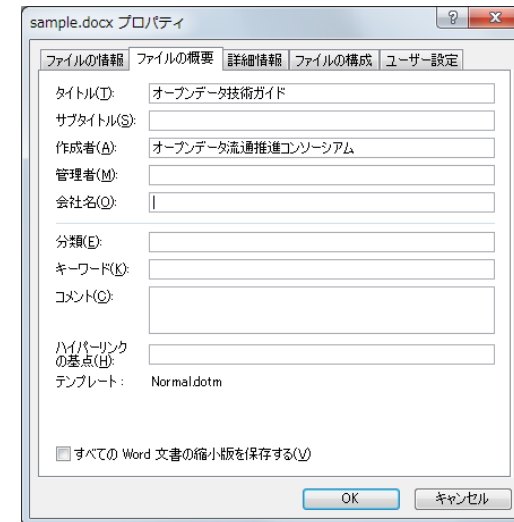
8.5 オープンデータの管理ポリシーとメタデータの付与方法

■ Apache Tika(*)を利用して、ファイルのメタデータを自動収集する例



Open Office 4 Writerの
文書プロパティ画面

Microsoft Word 2010の文書プロパティ画面



この文書形式データを
Apache Tikaで解析



この文書形式データを
Apache Tikaで解析



作成者→
タイトル→
作成日時→
最終更新日時→

```
{
  ...
  "dc:creator": "オープンデータ流通推進コンソーシアム",
  "dc:title": "オープンデータ技術ガイド",
  "dcterms:created": "2014-01-27Txx:xx:xx",
  "dcterms:modified": "2014-01-27Txx:xx:xx",
  ...
}
```

(*) <http://tika.apache.org>

第9章 オープンデータのための技術的指針

■ 本章の概要

- ▶ 機械判読に適したオープンデータを作成・編集するための技術的な指針を、識別子、ファイル形式及びデータの3項目に関して示す。

■ 本章の構成

1. 識別子に関する指針

◇オープンデータを識別する識別子が満たすべき性質と、それを満たすための手法を解説する。

2. ファイル形式に関する指針

◇公開するデータのファイル形式は、機械判読性の高い形式を利用することが望ましい。

◇オープンデータの技術レベルに基づいて、代表的なファイル形式を整理して示す。

3. データに関する指針

◇表形式データ・文書形式データ・地理空間情報・リアルタイムデータのそれぞれについて、機械判読性の高いデータを作成・編集する際の指針を示す。

9.1 識別子に関する指針

■ オープンデータにとっての識別子が満たすべき性質

1. ユニークであること。
2. 共通に利用できる体系であること。

■ 利用可能な識別子体系

- ▶ グローバルにユニークな識別子体系
- ▶ 公的機関が定める識別子体系・コード体系
- ▶ URI (Uniform Resource Identifier) として表現できる体系

■ 適切な識別子体系がない場合の対処法

1. 対象とする実物や組織・場所に番号が付与されていない場合は、まずそれらに番号を付与する。
2. 識別子のユニーク範囲を拡大する。
 - ◇ ucodeやDoIなどのグローバルな体系や、公的機関が定める識別子体系・コード体系に基づく識別子を取得し、管理する。
 - ◇ 付与した番号に組織が決めるURLを付与してグローバル化することもできる。
 - ◆ ただし、組織の統廃合等によりドメイン名が変わると、識別子も変わってしまうことに注意。

9.2 ファイル形式に関する指針

■ 基本方針

- ▶ 機械判読性の高い形式を利用することが望ましい。
- ▶ 代表的なファイル形式を、オープンデータの技術レベルに基づいてまとめると、下記のようなになる。

	Level 1	Level 2/3	Level 4
表形式データ	xls (Microsoft Excel形式)	CSV xlsx (Office Open XML) ods (OpenDocument) JSON	RDF/XML RDF/JSON, JSON-LD Notation3 Turtle等のRDF形式
文書形式データ	doc (Microsoft Word形式)	HTML XML docx (Office Open XML) odt (OpenDocument)	
地理空間情報	shape	KML GML	推奨するファイル形式
リアルタイムデータ		(ファイルの形で交換しない)	

9.3 データに関する指針

■ 指針のグレード：満たすべき指針の重要度にあわせて2つのグレードを設ける。

▶ グレード1

◇グレード1は、オープンデータが満たすことを強く推奨する指針であり、以下を満たすことを目的とする。

- ◆ データ形式に関する標準的な規格がある場合は、それに矛盾しないこと。
- ◆ データを取得した利用者が、データ本体の中身を修正したり手を加えたりすることなく、そのデータの本質的内容を正しく解釈するためのプログラムを書けること。

▶ グレード2

◇グレード2は、オープンデータが満たすことを推奨する指針であり、以下を満たすことを目的とする。

- ◆ データを取得したプログラムが、そのデータの項目や構造を正しく解釈できること。

■ 対象とするデータ

- ▶ 表形式データ
- ▶ 文書データ
- ▶ 地理空間情報
- ▶ リアルタイムデータ

■ 各指針に関する記述内容

- ▶ 表形式データを中心に指針を満たさない例と満たす例を明記し、それに対して解説する。
- ▶ 以下、各データに関する指針のみを記す。

9.3 データに関する指針／表形式データ

グレード	指針	
1	指針1	1つのファイルは、1種類の表から構成されるべきである。
	指針2	ヘッダは、1行から構成されるべきである。
2	指針3	データでない情報を、レコードに含めないことが望ましい。
	指針4	全てのフィールドは、他のフィールドと結合されないことが望ましい。
	指針5	値がない場合を除き、フィールドを空白にしない（省略しない）ことが望ましい。
	指針6	年の値には、西暦表記を備えることが望ましい。
	指針7	フィールドの単位が明記されていることが望ましい。
	指針8	利用している文字コードを明記することが望ましい。また、国際的に広く利用されている文字コードを利用することが望ましい。
	指針9	ファイルの属性や説明を表すメタデータが、フォーマルに記述されていることが望ましい。また、そのメタデータからデータセット本体へリンクし、たどれるようにすることが望ましい。
	指針10	データ本体を、XMLやRDFの形式を使ってフォーマルに記述することが望ましい。

9.3 データに関する指針／文書データ

グレード	指針	
1		(なし)
2	指針1	文章に存在する部・章・節・図表などの構造が、機械判読性の高いフォーマットで記述されていることが望ましい。
	指針2	文章内に、整形のための符号や文字（空白、改行等）を含めないことが望ましい。
	指針3	文書形式データが表形式データを含む場合、グレード1以上の表形式データが添付されていることが望ましい。
	指針4	テキスト形式の文書形式データを利用している場合は、利用している文字コードを明記することが望ましい。また、国際的に広く利用されている文字コードを利用することが望ましい。
	指針5	文章に対する、情報利用者が理解できるような説明が、メタデータとして記述され、当該文書にリンクされていることが望ましい。

9.3 データに関する指針／地理空間情報

グレード	指針	
1	指針1	位置情報に関するデータを付与する場合は、緯度・経度等の位置情報に加えて、測地系が明記されるべきである。屋外であれば、世界測地系を利用することが望ましい。屋内であれば、座標系と描画縮尺（入力精度）を示すべきである。
2	指針2	地理空間情報は、ベクタ形式に依るものが望ましい。ベクタ形式のデータの作成に当たっては、最新の ISO 規格及び JIS 規格に基づいた地理空間情報標準プロファイル（JPGIS）を用いる。
	指針3	地理空間情報に対する、情報利用者が理解できるような説明が、メタデータとして記述され、当該文書にリンクされていることが望ましい。

9.3 データに関する指針／リアルタイムデータ

グレード	指針	
1	指針1	データの取得仕様が明記されているべきである。
	指針2	表形式データや地理空間情報をファイル形式で取得させる場合は、それぞれのグレード1の指針を満たすべきである。
2	指針3	リアルタイムデータの最新値・差分を取得する手法が提供されていることが望ましい。



2. オープンデータガイド 第2版案の概要

第2版作成の基本方針と対応状況

- 第1版公開時に実施した Call for Comment により頂いたご意見への対応
 - ▶ 3.4.1.1節「メタデータ」の説明を加筆
 - ▶ 3.4.1.2節「アクセス方法」、9.1.3節「適切な識別子体系がない場合の対処法」に、識別子がドメインを含む場合の対処法を加筆
 - ▶ 3.4.3節「データを公開するサーバ」に、サーバがWebAPIを提供する場合の留意点を追加
- オープンデータに関する新たな動向の加筆
- その他、理解しやすくするための修正
 - ▶ 「オープンデータの作成・公開」という表現を修正
 - ◇ 「オープンデータの整備・掲載」とする。

今後の課題

■ 頂いたご意見から得られた課題

- ▶ オープンデータの対象となる公共データに関する分析
 - ◇オープンデータとすることが望ましいデータ項目の例示
- ▶ データのフォーマット等の統一に関する検討
 - ◇公開されたデータの表記法
 - ◇用いられる文字集合・符号化方式
 - ◇公開されるデータの粒度
 - ◇数値の単位
 - ◇更新の頻度、鮮度、精度

■ その他の課題

- ▶ 第3章の内容の具体化
 - ◇技術的な観点からは、利用できるソフトウェアやツールに関する記述を拡充した方がわかりやすい。
 - ◇保持しているデータをオープンデータとするための、組織内での取り組みについても、先進自治体等の事例を調査してベスト・プラクティスとして紹介すると、追従しやすいのではないか。

