

# 札幌市オープンデータ推進ガイドライン

平成 30 年 1 月

札幌市まちづくり政策局政策企画部 ICT 戦略推進担当



行政における「オープンデータ」の取組は、平成 15 年（2003 年）に「公的機関の情報の再利用に関する指令」(PSI-Reuse 指令) を発令した EU に端を発し、既に 10 年以上の歴史を有しています。

日本では、東日本大震災を通じて情報共有の重要性が再認識されたことでオープンデータの機運が高まり、政府が平成 24 年（2012 年）に策定した「電子行政オープンデータ戦略」を契機として、今日までオープンデータの取組が推進されています。

札幌市では、平成 27 年（2015 年）11 月にオープンデータカタログサイトを開設しオープンデータの公開に取り組んでいますが、様々な分野におけるデータ活用の重要性が高まる中、行政の保有するデータをさらに広く共有し、より使いやすいデータにしていくためには、各部署の自主的な取組が必要不可欠です。

そこで、札幌市におけるオープンデータ推進の基本的な方向性を示すため「札幌市オープンデータ推進方針」を策定するとともに、各部署がオープンデータの公開に取り組む際の疑問を解消するための手引きとして本書を作成しました。

是非、本書の内容を基礎とし、オープンデータの公開を積極的に進めていただくようお願いします。

# 目次

<b>Chapter 1 基礎編</b> .....	<b>5</b>
<b>1 「オープンデータ」って何？</b> .....	<b>6</b>
■ 「オープンデータ」の定義 .....	6
■ 「機械判読に適したデータ」とは？ .....	7
<b>Column</b> ➤ CSV ファイルはとてもシンプル .....	8
<b>Tips</b> ➤ 適切なファイル形式を選択しよう .....	9
<b>Column</b> ➤ 5つ星オープンデータと LOD .....	10
■ なぜオープンデータにしなければならない？ .....	11
<b>2 札幌市のオープンデータ</b> .....	<b>12</b>
■ 札幌市オープンデータカタログサイト .....	12
■ 札幌市オープンデータ推進方針 .....	13
<b>3 オープンデータの公開と利用</b> .....	<b>14</b>
■ オープンデータ推進の基本原則 .....	14
■ オープンデータ化の対象 .....	14
<b>Tips</b> ➤ オープンデータにできないデータ .....	16
<b>Column</b> ➤ 国の「推奨データセット」 .....	18
■ 第三者の権利を含むデータ .....	20
<b>Tips</b> ➤ 著作権とは？ .....	21
<b>Tips</b> ➤ 著作権以外の権利 .....	22
■ オープンデータの利用ルール .....	23
■ クリエイティブ・コモンズ・ライセンス .....	24
■ オープンデータ公開の手順 .....	26
<b>Chapter 2 データ作成編</b> .....	<b>27</b>
<b>1 表形式データの作成</b> .....	<b>28</b>
■ 表形式データに適したファイル形式 .....	28

■	CSV ファイルの作成 .....	28
	<b>Tips</b> ➤ CSV 形式でのデータ保存 .....	29
■	表形式データを作成する際の基本原則 .....	30
■	原則 1 .....	30
■	原則 2 .....	31
■	原則 3 .....	32
■	原則 4 .....	33
■	原則 5 .....	34
■	原則 6 .....	35
■	原則 7 .....	36
■	原則 8 .....	37
<b>2</b>	<b>文書データの作成 .....</b>	<b>38</b>
■	文書データに適したファイル形式 .....	38
■	文書データを作成する際の基本原則 .....	39
■	原則 1 .....	39
	<b>Tips</b> ➤ Word の見出し機能を活用しよう .....	40
■	原則 2 .....	41
■	原則 3 .....	41
<b>3</b>	<b>地理空間情報の作成 .....</b>	<b>42</b>
■	地理空間情報とは .....	42
	<b>Column</b> ➤ 地理空間情報のオープンデータ化 .....	43
■	地理空間情報に適したファイル形式 .....	45
	<b>Column</b> ➤ ラスタ形式とベクタ形式 .....	46
■	地理空間情報を作成する際の基本原則 .....	48
■	原則 1 .....	48
■	原則 2 .....	50
■	原則 3 .....	51
<b>4</b>	<b>PDF 形式の使用について .....</b>	<b>52</b>
<b>5</b>	<b>全データ共通の注意事項 .....</b>	<b>53</b>
■	ファイル名の命名ルール .....	53

■	データ内で使用する文字 .....	54
■	ファイル内容・プロパティの事前確認 .....	55
<b>6</b>	<b>メタデータの作成 .....</b>	<b>56</b>
■	メタデータの項目例 .....	56
■	メタデータの提供方法 .....	57

## Chapter 1 基礎編

ここでは、「オープンデータって何?」「何でそんなことをしなければならんだ」といった基本的な疑問にお答えします。

# 1

## 「オープンデータ」って何？

### ■ 「オープンデータ」の定義

「オープンデータ」とは、一般的に「**機械判読に適したデータ形式で、営利・非営利目的を問わず二次利用が可能な利用ルールで公開されたデータ**」であるとされています。

今日、「データ」という単語は特定の職業に就いている人でなくとも日常的に使うものとなっており、その用途も多種多様ですが、本来は「事実」や「資料」を意味する言葉です。

例えば、紙に印刷された資料も「データ」であり、コンピュータに保存された電磁的記録もいわゆる「電子データ」です。

電子データであっても、人が目で見えて理解するのに適したデータもあれば、機械（コンピュータ）が人手を介さずに理解し処理するのに適したデータもあります。

また、データが著作物に該当する場合、著作権者が認めない限り、原則、第三者は加工・頒布などを行うことはできません。

つまり「オープンデータ」とは、「**機械（コンピュータ）にとって理解し処理しやすい電子データ**」であり、かつ「**誰がどのように使ってもよいと明示されたデータ**」ということになります。



## ■ 「機械判読に適したデータ」とは？

「オープンデータ」は、インターネットを通じて流通することや、コンピュータを用いたデータ分析・アプリケーション開発などに利用することを前提とした概念であり、コンピュータができるだけ人手を介さずに処理できる形である必要があります。

例えば、PDF形式のデータは、画面上や印刷物で人が目視することに適した形式であり、コンピュータがPDFファイル内に記述されたテキストや表を自動的に読み取ることは困難です。

また、PDFファイルは加工・編集を行うために専用のソフトウェアを必要とします。

皆さんもよく利用する表計算ソフト Excel で標準採用されている XLS (.xls) などの形式のデータは、Excel などのソフトウェアを用いて容易に編集・加工ができるため、PDF などと比べると機械判読に適していると言えます。

しかし、XLS ファイルの閲覧や編集・加工を行うためには、Excel をはじめとした特定のソフトウェアが必要です。また、人が目視するための表を形成するものであるため、表の内容や構造をコンピュータが理解するためには手入力などが必要となります。

一方、CSV形式のデータは、表の項目をカンマで区切り、1行ずつ列挙した構造のテキストファイルです。**データが明確かつシンプルな構造で作られている**ため、コンピュータが判別・処理することが容易であり、より人手を介さずにアプリケーションなどに活用することが可能です。

また、CSVファイルはテキストファイルであるため、Excel などの**特定のソフトウェアでなくても加工・編集が可能**です。

## Column ➤ CSV ファイルはとてもシンプル

### CSV ファイルの構造

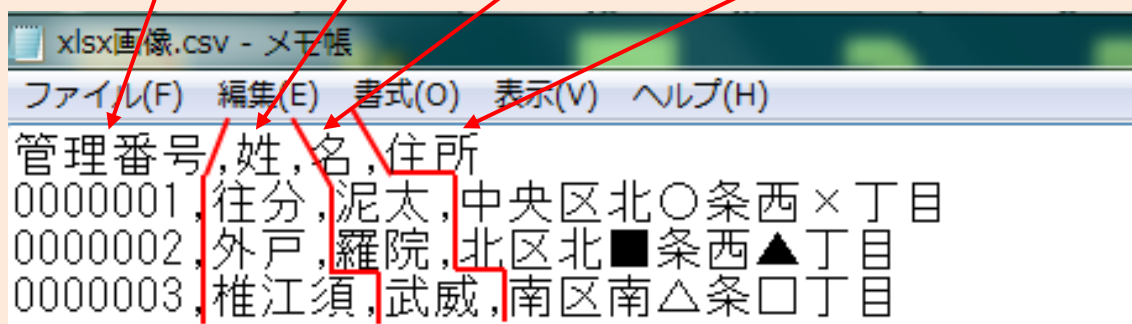
CSV の「C」がカンマを意味するとおり、CSV ファイルはカンマで項目を区切り列挙した構造となっています。

また、あくまで「テキストファイル」であるため、文字を装飾することや、1つのファイルに複数のシートを含むことができません。

このようなシンプルな形式であるからこそ、ソフトウェアを問わずに開くことができる互換性を有し、コンピュータによる処理も容易であると言えます。

### Excel で作成した表を CSV 形式で保存した例

	A	B	C	D
1	管理番号	姓	名	住所
2	0000001	往分	泥太	中央区北○条西×丁目
3	0000002	外戸	羅院	北区北■条西▲丁目
4	0000003	椎江須	武威	南区南△条□丁目



上図のように、「メモ帳」ソフトで開くことも可能です。

## Tips ▶ 適切なファイル形式を選択しよう

データをより利用しやすいものとするためには、以下の点にも留意することが望ましいと言えます。

### 標準的なファイル形式

特定のアプリケーションでなければ使用できないファイル形式や、仕様が非公開のファイル形式を用いると、利用者のデータ使用を困難にする可能性があるため、標準的なファイル形式を選択することが適当です。

例えば、業務で使用する人が多い Word (.docx 等)、Excel (.xlsx 等)、PDF、CSV などの形式は、ISO（国際標準化機構）などの国際規格に承認された標準的な形式です。

### データの内容や性質に適したファイル形式

データの内容や性質に応じた適切な形式を選択することで、コンピュータによる処理が容易になります。

データ形式	左記形式が適するデータ例	代表的な拡張子
表形式	予決算、統計、施設等の一覧など	.csv .xls .xlsx .ods
文書形式	報告書、報道発表資料など、文章のみのものや、文字と図形・画像が混在するもの	.pdf .doc .docx .html .xml .odt
地理空間情報	都市計画図、固定資産など、地図上に情報を表すもの	.shp .gml .kml .GeoTiff
画像形式	マークやキャラクターの画像、ハザードマップなど	.jpeg .png .tiff

## Column ▶ 5つ星オープンデータと LOD

オープンデータとしての利用しやすさを表す評価指標として、「5つ星オープンデータ」があります。

この指標では、「**Linked Open Data (LOD)**」と呼ばれる状態を理想として最高位の5つ星に位置付けていますが、まずは最低でも3つ星を目指しましょう。



- 1つ星：(PDF などでもよいので) データ形式を問わず、自由に二次利用が可能な利用ルールで公開する。
- 2つ星：Word (.doc 等)、Excel (.xls 等) など、構造化されており、対応するソフトウェアを使用すればコンピュータが処理可能な形式で公開する。
- 3つ星：CSV など、仕様が公開されており、特定のソフトウェアでなくとも処理が可能な形式で公開する。
- 4つ星：タイトルや作成者等の目録的情報、文章等の構造に関する情報、語彙の意味情報など、公開データそのものに関する付帯的なデータ (メタデータ) を、コンピュータが理解可能な共通語彙 (URI) を用いた表現方法 (RDF 等) で公開する。
- 5つ星：RDF 等により、公開データに関連する他のデータへのリンクを付与し、データをネットワーク化する。

**(Linked Open Data = LOD)**

## ■ なぜオープンデータにしなければならない？

市民の知る権利に基づき情報を開示する情報公開制度とは異なり、「札幌市オープンデータ推進方針」では、以下の3点をオープンデータ推進の意義として定めています。

### (1) 生活の向上

市民や民間団体と行政データを共有するとともに協働して地域課題の解決に取り組むことで、より良質かつ幅広いサービスを創出し、市民生活の利便性、安全性などの向上を図る。

### (2) 経済の活性化

事業者等がより多様なデータの分析やデータの横断的利用を行うことが可能となり、新たなサービスやビジネスの創出が期待できる。

### (3) 行政の信頼性・透明性の向上

本市の施策に関わる情報をオープンデータとして公開することで、市民自ら施策の妥当性を評価又は理解することが容易になり、行政への関心を深めることに寄与する。

情報公開制度では、市民への情報提供に主眼を置いており、データとしての二次利用は想定されていません。

オープンデータという取組は、市民が市政に関する情報を「知る」機会を提供するだけでなく、企業や学術機関などあらゆる主体が**データを自由に「使う」ことができるようになる**点に大きな意義があります。

## 2

# 札幌市のオープンデータ

## ■ 札幌市オープンデータカタログサイト

札幌市では、平成 27 年 11 月にオープンデータカタログサイトを開設し、以下のオープンデータを公開しています。（平成 29 年 12 月 1 日現在）

分類	件数	データ例
防災	7 件	避難所情報、地震防災マップなど
人口	12 件	将来推計人口、人口動態など
経済	2 件	市民経済計算、産業関連表
交通	18 件	路線図・時刻表、各種輸送機関輸送状況など
選挙	1 件	投開票所一覧表
教育	1 件	教育データ集
保健	3 件	衛生年報、予防接種委託医療機関など
福祉	5 件	市内要介護認定の状況、老人ホーム一覧など
環境	3 件	空間放射線量測定結果、川の生物調査データなど
ごみ	15 件	家庭ごみ収集日カレンダー、ごみ分けガイドなど
公園	2 件	公園便所一覧表、公園緑地の統計
まちづくり	3 件	まちづくりセンター・出張所一覧など
都市計画	3 件	地区計画書・計画図、市内駐輪場マップなど
上下水道	3 件	応急給水拠点施設、拠点給水施設など

分類	件数	データ例
子ども	3件	認可保育所一覧、認可保育所入所状況など
その他	6件	市政世論調査結果、入札・契約情報など
<b>合計</b>	<b>87件</b>	

【参考】札幌市オープンデータカタログサイト  
(<http://www2.city.sapporo.jp/opendata/>)

今後は、さらに公開データを拡充していくとともに、機械判読に適した構造・形式に整備していく必要があります。

## ■ 札幌市オープンデータ推進方針

本市では、本ガイドラインのほか、札幌市としてオープンデータを推進するに当たっての取組の方向性や原則などを明示するため、「札幌市オープンデータ推進方針」を定めています。

本市が保有する情報のオープンデータ化は、本方針に従って行っていただく必要があります。

【参考】札幌市オープンデータ推進方針（イントラ内 HP）  
([http://web.intra.city.sapporo.jp/machi\\_seisaku/top/bu-page/ict/file/opendata\\_houshin.pdf](http://web.intra.city.sapporo.jp/machi_seisaku/top/bu-page/ict/file/opendata_houshin.pdf))

## 3

# オープンデータの公開と利用

## ■ オープンデータ推進の基本原則

「札幌市オープンデータ推進方針」では、本市がオープンデータを推進するに当たっての基本原則を下記のとおり定めています。

- (1) 自主的かつ積極的に公的データを公開する。
- (2) 可能な限り、機械判読に適した二次利用の容易な形式で公開する。
- (3) 営利目的又は非営利目的を問わず活用を促進する。
- (4) 取組可能なデータから速やかに着手する。
- (5) 費用対効果について十分に考慮し、効率的に取組を進める。

## ■ オープンデータ化の対象

「札幌市オープンデータ推進方針」では、本市がオープンデータとすべき情報について下記のとおり定めています。

- (1) 札幌市が保有するデータのうち、札幌市ウェブサイト上で公開されているものは、原則全てオープンデータ化の対象とする。
- (2) (1)のうち、市民やデータ利用者のニーズの高いものについては、特に優先してオープンデータ化に着手する。
- (3) 札幌市情報公開条例に定める非公開情報（個人情報等）、その他具体的かつ合理的理由により二次利用が認められないものは、オープンデータ化の対象から除外する。
- (4) 札幌市ウェブサイト上で公開されていないデータのう



ち、市民やデータ利用者のニーズの高いものについては、その必要性や公開の可否を十分に検討し、可能なものから公開する。

各部署とも膨大なデータを保有しており、一度にあらゆるデータをオープンデータ化することは困難であるため、上記のように優先順位を定めています。

しかし、**どのようなデータが画期的なアイデアや有用なサービス、新しい知見につながるかは未知数**であり、公開する側が用途を決めつけられるものではありません。

提供者にとっては使い道のなさそうなデータも、利用者のアイデアや、他分野のデータとの組み合わせにより、新しい価値を生み出すことがあります。

公開できない理由が明確なデータ以外は、**具体的な活用想定があるかどうかにかかわらず、積極的に公開していくことがオープンデータの本旨である**ことを忘れないようにしましょう。

## Tips ▶ オープンデータにできないデータ

オープンデータの趣旨に鑑み、行政が持つデータは最大限オープンデータ化されるべきですが、一方で、行政が保有するデータの中には公開や二次利用に適さない情報があります。

例として、以下に挙げるデータはオープンデータとして公開しないこととします。

### 札幌市個人情報保護条例に定める個人情報

札幌市個人情報保護条例では、本来の事務の目的を越えて第三者に個人情報を提供することを、原則、禁止しています。

ここでいう「個人情報」とは、**氏名、生年月日などにより、その情報が誰に関するものかわかる情報**を指します。

また、氏名等の記載がない情報であっても、**内容から特定の個人を識別できるもの**や、**他の情報と照合することで間接的に個人を識別できるもの**も個人情報に該当します。

### 札幌市情報公開条例（第7条）における非公開情報

札幌市情報公開条例では、主に以下のような情報を非公開情報として定めています。

- ・個人に関する情報で特定の個人を識別できる情報又は特定の個人を識別することはできないが、公開することにより、個人の利益を害するおそれがある情報
- ・公開することにより、法人等又は個人の権利、競争上の地位その他正当な利益を害する情報
- ・札幌市の要請を受けて、合理的な理由に基づき、公開しないことを条件に任意に提供された情報
- ・公開することにより、公共の安全と秩序の維持に支障を

及ぼすおそれがある情報

- ・市や国等の内部又は相互間における審議、検討又は協議に関する情報で、公開することにより、意思決定の中立性が不当に損なわれるものや、特定の者に不当に利益・不利益を与えるもの
- ・公開することにより、市や国等が行う事務又は事業の適正な遂行に著しい支障を及ぼすもの
- ・法令もしくは他の条例、札幌市が法律上義務を負う国の機関の指示等により、公開することができないもの

## Column ▶ 国の「推奨データセット」

国においては、地方公共団体によるオープンデータの公開とその利活用を促進するため、オープンデータに取り組む地方公共団体の参考となるよう、各地方公共団体に公開を推奨するデータセット及びフォーマットの標準例を示す「**推奨データセット**」を公表しています。

現在示されている推奨データセットでは、公開ニーズの高いデータや既に活用事例のあるデータの中から、下記の14データを列挙しています。

また、各地方公共団体のデータセットのフォーマットを共通化することで、地方公共団体の枠を超えた利用が容易になるため、具体的なデータ項目のフォーマットも併せて示されています。

### 「推奨データセット」で示された14データセット

データセット名	詳細	オープンデータとして公開することによる効果
AED設置箇所一覧	AEDの設置箇所についての一覧。	最寄りのAEDの検索や効率的・効果的なAED設置が可能となる。
介護サービス事業所一覧	介護サービス事業所の一覧。	個人のニーズに対応した介護サービスを容易に検索できるようになる。
医療機関一覧	病院・診療所についての一覧。	位置情報や移動手段情報などと組み合わせ、受診可能な医療機関が検索可能になる。
文化財一覧	国もしくは地方公共団体が指定、登録、選定等を行った文化財についての一覧。	移動手段情報と組み合わせることにより、関心のある文化財へ容易にアクセスできるようになる。
観光施設一覧	観光施設の情報の一覧。	移動手段情報と組み合わせることで、効率的な旅程の作成や観光施設へのアクセスが可能になる。
イベント一覧	各地方公共団体にて開催されるイベントの一覧。	広い範囲に情報提供することが可能となり、集客等に貢献することが期待される。

データセット名	詳細	オープンデータとして公開することによる効果
公衆無線LAN アクセスポイント一覧	公衆無線LAN アクセスポイントの一覧。	インターネットの利用可能場所を容易に把握することができ、旅行者（特に訪日外国人）の利便性向上が期待される。
公衆トイレ一覧	公衆トイレの一覧。	アプリ等で地図上に表示し、近隣の公衆トイレを検索することが可能になる。
消防水利施設一覧	消防水利施設の一覧。	担当区域外においても、最も近い消防水利施設の場所等を検索し、迅速に対応することが可能になる。
指定緊急避難場所一覧	市区町村から提供される指定緊急避難場所の一覧。	災害時における地域住民や旅行者の迅速な避難、関係機関による円滑な支援活動が可能となる。
地域・年齢別人口	住民基本台帳に基づく学区・年齢別の人口一覧。	該当地域どのような人が何人程度居住しているか明らかになり、他のデータと組み合わせることで、きめ細やかな政策・戦略立案に資することが期待される。
公共施設一覧	公共施設の一覧。	地域住民が自身のニーズに合致した公共施設やイベント等を検索可能になり、施設の活用が促進される。
子育て施設一覧	幼稚園、保育園、認定こども園の一覧。	ニーズに対応した子育て施設を探すことが容易になる。
オープンデータ一覧	オープンデータ化されているデータセットの一覧。	オープンデータ利用者の利便性向上が期待できる。

※出典（上表）：内閣官房情報通信技術（IT）総合戦略室「政府 CIO ポータル」内の「推奨データセット項目定義書」を元に要約

## 【参考】政府 CIO ポータル

(<https://cio.go.jp/policy-opensource>)

推奨データセット項目定義書

([https://cio.go.jp/sites/default/files/uploads/documents/opensource\\_suisyou\\_dataset\\_teigisyo.xlsx](https://cio.go.jp/sites/default/files/uploads/documents/opensource_suisyou_dataset_teigisyo.xlsx))

## ■ 第三者の権利を含むデータ

「札幌市オープンデータ推進方針」では、オープンデータ化を検討する対象データに第三者が著作権その他の権利を有する情報が含まれる場合、公開の可否・範囲・利用条件は、当該第三者の判断によることとしています。

**権利を有する第三者からオープンデータ化の許諾を得られないデータは、当該第三者の権利を侵害しないよう、オープンデータとしての公開は行わないこととします。**

ただし、第三者が権利を有する情報が含まれるデータであっても、直ちに公開を諦めるのではなく、**権利を有する第三者と積極的に協議を行い、許諾を得て公開ができるよう努めてください。**

なお、第三者との協議に当たっては、オープンデータの趣旨や利用規約について確実に説明を行い、不特定多数の利用者に二次利用されることについて十分に理解を得るようにしてください。

## Tips ▶ 著作権とは？

著作権とは、知的創造活動による創作物に対する権利である知的財産権の一種であり、「著作物」の利用や配布などに関する排他的な権利のことです。

### 「著作物」の定義

著作権法では、著作物の定義を、「思想又は感情を創作的に表現したものであって、文芸、学術、美術又は音楽の範囲に属するもの」と定めています。

楽曲や美術作品、文学作品、写真などはもちろん該当しますが、そのほかにも地図や図面、(コンピュータ)プログラムなども著作物となります。

また、編集物やデータベース(複数の情報を集め、コンピュータで検索できるようにしたもの)は、一定の創作性が認められる場合に限り、著作物と認められます。

### 著作権の主な内容

複製権	著作物を印刷や複写などにより複製する権利
公衆送信権等	テレビ、ラジオ等の放送や、インターネットを通じた送信などを行う権利
譲渡権	著作物を、その原作品又は複製物を譲渡することにより、公衆に提供する権利

上記のほかにも、上演権・演奏権・上映権・口述権・展示権・頒布権など、様々な権利が著作権に含まれます。

## Tips ▶ 著作権以外の権利

以下に例示するような著作権以外の権利についても、オープンデータの公開により侵害することのないよう十分注意してください。

### 著作者人格権

著作権法では、著作権とともに、著作者の人格的利益を保護するため、以下の権利を「著作者人格権」と定め保護しています。

公表権	まだ公表されていない著作物について、公表するかどうかなどを決定する権利
氏名表示権	著作物に著作者として氏名を表示するかしないか、表示する場合にはどのような名義とするかを判断する権利
同一性保持権	著作物の内容やタイトルを、意に反して改変されない権利

### 人格権

身体・自由・名誉・氏名・肖像・信用など、権利者と分離することのできない私人の権利を指します。オープンデータ公開に当たっては特に以下の権利に注意しましょう。

プライバシー権	私生活上の事柄をみだりに公開されない権利
肖像権	写真や映像などにより、自身の容姿をみだりに利用されない権利
パブリシティ権	有名人・著名人の氏名や肖像等が持つ経済的な価値を排他的に支配する権利



## ■ オープンデータの利用ルール

オープンデータとしてデータを公開するためには、（一定条件の下）データを自由に使えることを予め明示する必要があります。

そうしなければ、利用者がデータの利用に関して著作権者に許可を求め、著作権者が個別に許可する必要があるためです。

また、著作権者が著作権を放棄することも可能ですが、データの利用に制限を加えることができなくなるとともに、データの信用性や価値が失われる可能性があります。

そこで、著作権は維持しつつ、データの自由な利用を包括的に許諾できるよう作られたシステムが、後述する**クリエイティブ・コモンズ・ライセンス（CCライセンス）**です。

CCライセンスの仕組みは、国際的に広く利用されており、国のオープンデータ利用規約（政府標準利用規約）もCCライセンスと互換性を有するものとされています。

**「札幌市オープンデータカタログサイト利用規約」においても、CCライセンスを原則として採用**しつつ、市の免責事項など、いくつかの事項を付帯しています。

【参考】札幌市オープンデータカタログサイト利用規約

[http://www.city.sapporo.jp/shisei/seisaku/open\\_data/opendata\\_kiyaku.html](http://www.city.sapporo.jp/shisei/seisaku/open_data/opendata_kiyaku.html)

## ■ クリエイティブ・コモンズ・ライセンス

CC ライセンスは、4 種類の利用条件の組み合わせによる **全 6 パターンの許可内容（ライセンス）** により、データの利用を認めることを予め明示するライセンス表示の仕組みです。

### 【4 種類の利用条件】



**表示 (BY)**

クレジット（タイトル、著作権者、URL 等）を表示すること



**非営利 (NC)**

営利目的での利用をしないこと



**改変禁止 (ND)**

元の著作物を改変しないこと



**継承 (SA)**

改変後も元の著作物と同じ CC ライセンスで公開すること

※出典（表内図）：Creative Commons「When we share, everyone wins - Creative Commons」

(<https://creativecommons.org/about/downloads/>)

## 【6パターンのCCライセンス】

ライセンス	4種類の利用条件			
	出典表示	営利利用	改変	継承
 表示 (CC BY)	必須	許可する	許可する	継承しない
 表示-継承 (CC BY-SA)	必須	許可する	許可する	継承する
 表示-改変禁止 (CC BY-ND)	必須	許可する	許可しない	—
 表示-非営利 (CC BY-NC)	必須	許可しない	許可する	継承しない
 表示-非営利-継承 (CC BY-NC-SA)	必須	許可しない	許可する	継承する
 表示-非営利-改変禁止 (CC BY-NC-ND)	必須	許可しない	許可しない	—

※出典（表内図）：Creative Commons「When we share, everyone wins - Creative Commons」

(<https://creativecommons.org/about/downloads/>)

※継承（SA）の条件が付されていない場合も、元の著作物より制限の緩いライセンスを採用することはできない。

※改変禁止（ND）の条件が付されていない場合も、著作者や実演家の名誉又は声望を害するような改変は認められていない。

利用者は、各データに明示されたライセンスを順守する限り、著作権者の許可を得ることなくデータを利用することができます。

25 ページ表の6パターンのうち、「**CC BY (シーシーバイ)**」(クレジットを表示すれば自由に利用・改変してよい)が最も制限の少ないライセンスであり、「**札幌市オープンデータ推進方針**」においても、**CC BY によりデータを公開することを原則としています。**

しかし、CC BY では公開できない明確な理由がある場合であっても、直ちに公開を諦めるのではなく、CC BY 以外のライセンスで公開できないか積極的に検討してください。

なお、「札幌市オープンデータ推進方針」では、「CC BY」以外の利用条件を適用する場合はその理由を明示することとしていますので注意してください。

## ■ オープンデータ公開の手順

札幌市では、既にオープンデータカタログサイトを開設しオープンデータを公開していますが、より利用者にとって利用しやすく、市職員にとってもデータを公開しやすいサイトとするため、カタログサイトのリニューアルを予定しています。

これに伴い、より効率的なオープンデータの作成・公開・管理の方法を検討中です。

当面は、新たにオープンデータの公開を検討する場合には**個別に政) 政策企画部 ICT 戦略推進担当までご相談ください。**

## Chapter 2 データ作成編

ここでは、オープンデータとして公開することを前提としたデータを実際に作成する方法について説明します。

# 1

## 表形式データの作成

### ■ 表形式データに適したファイル形式

表形式のデータについては、仕様が公開され、国際的に標準化された形式である **CSV で提供することを推奨**しますが、同じく国際標準化されている Open Document Format (.ods) や Office Open Xml (.xlsx) も使用できます。

CSV 形式については、8 ページもご覧ください。

### ■ CSV ファイルの作成

CSV 形式のファイルは、Excel で作成したデータを保存する際に、CSV 形式を指定して保存することで作成できます。

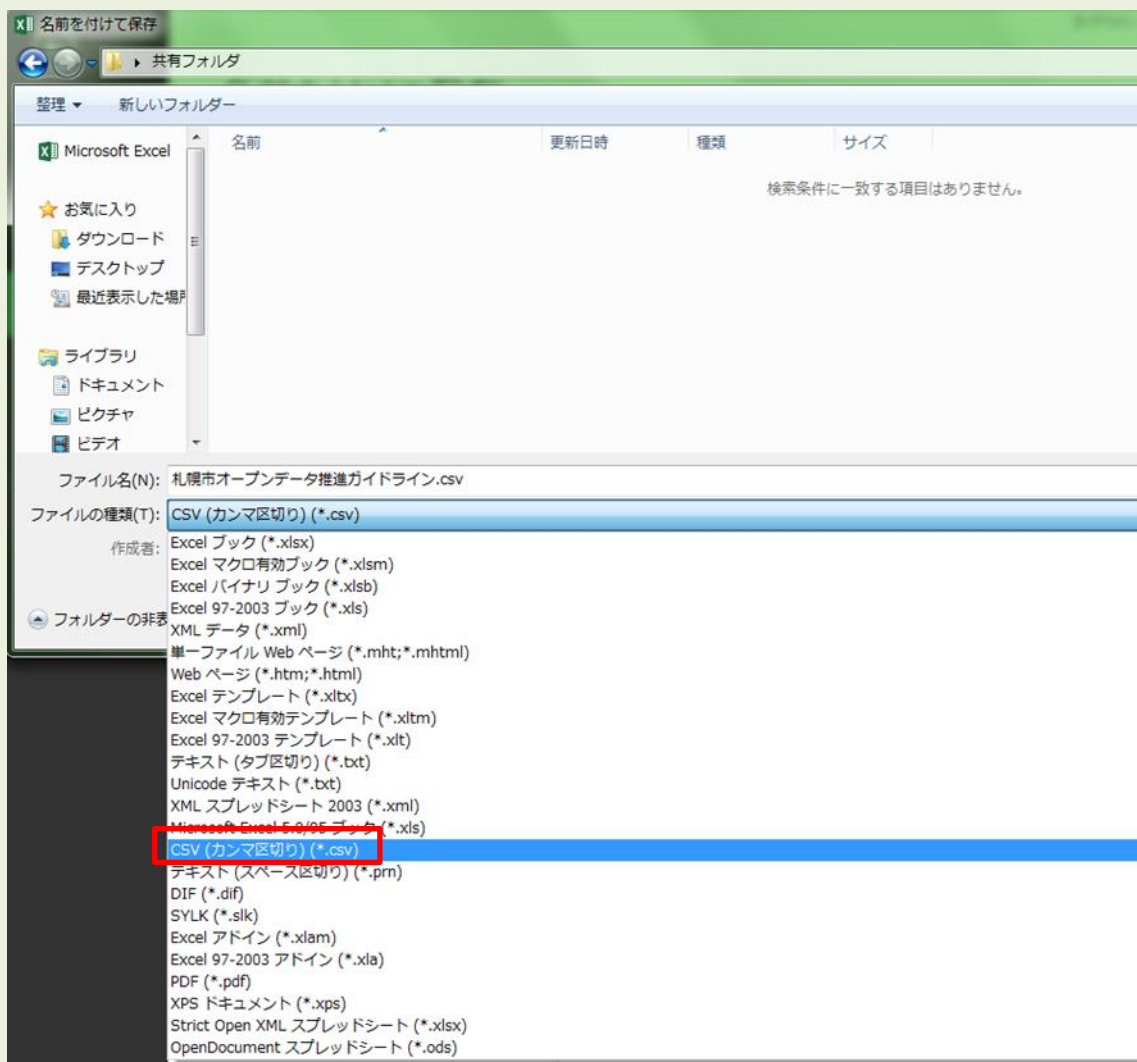
しかし、作成された表が、人が目視しやすいように工夫されていることで、逆にコンピュータにとっては判読しにくいものになってしまう場合があります。

刊行物への掲載や一般的な資料としての使用・配布などを目的として、人にとって見やすいデータを作成したい場合であっても、**まず 30 ページの基本原則に従って機械判読に適した構造のデータを作成**し、CSV 形式で保存した後、目視しやすい構造に再整理してください。

なお、オープンデータの公開に当たり、機械判読に適した形式と、目視に適した形式の両方を公開することは差し支えありません。

## Tips ▶ CSV 形式でのデータ保存

Excel で作成した表データを CSV 形式で保存したい場合は、データを保存する際、下記の図のように「ファイルの種類」から「CSV (カンマ区切り)」を選択するだけで CSV ファイルが保存されます。



## ■ 表形式データを作成する際の基本原則

機械判読に適した表形式データを作成するため、以下の原則に従ってください。

- 原則 1 : 1つのファイルは1種類の表のみで構成する
- 原則 2 : ヘッダ（項目名）は1行のみで構成する
- 原則 3 : セルの結合を行わない
- 原則 4 : データの値がない場合以外はセルを空白にしない
- 原則 5 : データ以外の情報（注釈など）を含めない
- 原則 6 : セル内の文字列にスペース・改行・カンマなどを含めない
- 原則 7 : 年の値は西暦表記にする
- 原則 8 : データの単位を明記する

## ■ 原則 1 : 1つのファイルは1種類の表のみで構成する

複数の表を含むファイルをコンピュータが判読しようとする時、表と表の境界を判別する必要があるため、処理が複雑になってしまいます。

複数の表を扱う場合は、個別のファイルに格納しましょう。



**ファイルA**

区	総数	0~14歳	15歳~64歳	65歳以上
中央区	220189	22250	156677	41250
北区	278781	34185	187505	55946
東区	255873	31210	174326	49369
白石区	204259	23698	140875	38818
厚別区	128492	15364	84842	28185
豊平区	212118	23513	145007	42734

区	一般世帯	65歳以上世帯員のいる一般世帯	高齢夫婦世帯	高齢単身世帯
中央区	120596	27529	7830	11338
北区	127291	37318	12150	11001
東区	118804	33566	10376	11048
白石区	100343	26265	7753	9095
厚別区	54036	18329	6572	5147
豊平区	107084	29141	8752	10232



**ファイルB**

区	一般世帯	65歳以上世帯員のいる一般世帯	高齢夫婦世帯	高齢単身世帯
中央区	120596	27529	7830	11338
北区	127291	37318	12150	11001
東区	118804	33566	10376	11048
白石区	100343	26265	7753	9095
厚別区	54036	18329	6572	5147
豊平区	107084	29141	8752	10232

**ファイルA**

区	総数	0~14歳	15歳~64歳	65歳以上
中央区	220189	22250	156677	41250
北区	278781	34185	187505	55946
東区	255873	31210	174326	49369
白石区	204259	23698	140875	38818
厚別区	128492	15364	84842	28185
豊平区	212118	23513	145007	42734



## ■ 原則 2 : ヘッダ (項目名) は 1 行のみで構成する

各列の項目名を示す最初の行 (ヘッダ) が複数行となっている場合、コンピュータが、項目名とデータの内容を区別することが難しくなります。

例えば下記の図のような場合は「人口」を各項目の中に含めるなど、ヘッダは 1 行のみで構成するようにしてください。



区	人口		
	0~14 歳	15 歳~64 歳	65 歳以上
中央区	22250	156677	41250
北区	34185	187505	55946
東区	31210	174326	49369
白石区	23698	140875	38818



区	0~14 歳の人口	15 歳~64 歳の人口	65 歳以上の人口
中央区	22250	156677	41250
北区	34185	187505	55946
東区	31210	174326	49369
白石区	23698	140875	38818

### ■ 原則 3 : セルの結合を行わない

人にとって見やすい表を作成する際は、例えば複数の項目にまたがる共通事項を表す場合など、セルを結合することがあります。

しかし、コンピュータがこれを識別するためには、結合されている範囲と、結合セル内の値の両方を認識する必要があります。

よって、機械判読に適した表データを作成する際には、同一の内容であってもセルの結合を行わず、個別のセルに入力しましょう。



年	月	全市	中央区
2014	1	1940525	231651
	2	1940249	231811
	3	1940147	232060
	4	1939137	232244



年	月	全市	中央区
2014	1	1940525	231651
2014	2	1940249	231811
2014	3	1940147	232060
2014	4	1939137	232244

## ■ 原則 4 : データの値がない場合以外はセルを空白にしない

複数の項目にまたがる共通次項を省略し空白とした場合、コンピュータは値がないものと認識するため、正しく識別することができません。原則 3 と同様、同一の内容でも個別のセルに入力してください。



年	月	全市	中央区
2014	1	1940525	231651
	2	1940249	231811
	3	1940147	232060
	4	1939137	232244



年	月	全市	中央区
2014	1	1940525	231651
2014	2	1940249	231811
2014	3	1940147	232060
2014	4	1939137	232244

## ■ 原則 5 : データ以外の情報（注釈など）を含めない

注釈や注釈番号など、データの内容以外の補足的情報をデータの中に含めると、コンピュータがデータを正しく読み取ることができなくなるため、余分な情報は含めないでください。



年次別気象情報

年次	平均気温	降水量	降雪
2013	1) 9.2	1347.0	628
2014	9.3	1203.5	478
2015	10.0	1274.5	367
2016	9.3	1360.0	428

注： 1) 一部期間の観測方法に誤りがあったため、〇〇省実施調査の同期間の数値を採用した。

<資料> x) xx部xx課



年次	平均気温	降水量	降雪
2013	9.2	1347.0	628
2014	9.3	1203.5	478
2015	10.0	1274.5	367
2016	9.3	1360.0	428

なお、注釈などの補足的情報は、人がデータを目視し解釈する際には重要な情報です。

注釈等を含む表は、機械判読に適した形式（CSV 等）とは別に、人が読みやすい形式（xls 等）によって提供することが望ましいと言えます。

## ■ 原則 6 : セル内の文字列にスペース・改行・カンマなどを含めない

下図①では、文字列の幅をそろえるため、文字間にスペースを使用しています。

②では、各数値の単位を、改行した上で記載しています。

また、③では、位取りのため 3 桁ごとにカンマを付しています。

コンピュータにとって、これらのスペース・改行・カンマがどのような意味を持っているのかを識別することは困難です。

よって、機械判読に適した表データを作成する際は、不要なスペース・改行・カンマ等を含めないようにしましょう。



②

区	平成 15 年 (百万円)	平成 17 年 (百万円)
① 総 数	567,467	569,250
食 料 品	③ 196,070	196,631
繊維工業	511	752
金属製品	57,492	56,887



区	平成 15 年 (百万円)	平成 17 年 (百万円)
総数	567467	569250
食料品	196070	196631
繊維工業	511	752
金属製品	57492	56887

## ■ 原則 7 : 年の値は西暦表記にする

表内に年の値を含める場合、コンピュータにとっては、単純な数値であり大小を比較しやすい西暦の方が、識別が容易であると言えます。

年の値を含む場合は、原則、西暦を用い、必要に応じて和暦を併記しましょう。



年次	平均気温	降水量
平成 25	9.2	1347.0
平成 26	9.3	1203.5
平成 27	10.0	1274.5
平成 28	9.3	1360.0



年次	年次	平均気温	降水量
平成 25	2013	9.2	1347.0
平成 26	2014	9.3	1203.5
平成 27	2015	10.0	1274.5
平成 28	2016	9.3	1360.0

## ■ 原則 8 : データの単位を明記する

貨幣単位（円、ドルなど）や物理単位（温度、距離、質量など）などといった単位は、データの処理を行う上で必須の情報であるため、必ず明記してください。

なお、単位を表内に記載する場合は、数値と同じセルの中には入力せず、ヘッダの各項目名に付記するようにしましょう。



年次	平均気温	降水量	降雪
2014	9.3	1203.5	478
2015	10.0	1274.5	367
2016	9.3	1360.0	428



年次	平均気温	降水量	降雪
2014	9.3℃	1203.5mm	478cm
2015	10.0℃	1274.5mm	367cm
2016	9.3℃	1360.0mm	428cm



年次	平均気温 (℃)	降水量 (mm)	降雪 (cm)
2014	9.3	1203.5	478
2015	10.0	1274.5	367
2016	9.3	1360.0	428

## 2

# 文書データの作成

### ■ 文書データに適したファイル形式

文字を主な構成要素とした文書データ（一部図表を使用したものも含む）を作成する際には、仕様が公開され、標準化された形式として、**Office Open XML 形式 (.docx)** や **Open Document Format (.odt)** が推奨されます。（文字列のみの場合は、テキスト形式 (.txt) でも差し支えありません。）

また、HTML や XML のようなマークアップ言語（文章の構造や修飾情報などを文章中に記述するコンピュータ言語）を用いたファイルも、オープンデータとして有効です。

みなさんが一般的な文書を作成する際には、Word を使用することが多いと思われるので、現在 Word で標準的に用いられている Office Open XML 形式 (.docx) を使用するのがよいでしょう。

なお、Word2003 まで使用されていた DOC 形式 (.doc) は、機械判読性の観点で Office Open XML 形式 (.docx) に劣りますので、原則、Office Open XML 形式 (.docx) を使用しましょう。



## ■ 文書データを作成する際の基本原則

機械判読に適した文書データを作成するため、以下の原則に従ってください。

原則 1 : 文章の構造をコンピュータが明確に認識できる形で記述する

原則 2 : 文章内に整形のための符号や文字を含めない

原則 3 : 文章内の図表の元になった表形式データを添付する

## ■ 原則 1 : 文章の構造をコンピュータが明確に認識できる形で記述する

通常、文章を作成する際は、人にとって読みやすいよう、部、章、節、段落などの構造に整形する場合があります。

しかし、これらの構造をフォントや文字飾りなどによって表現すると、コンピュータが文章の構造を認識することができません。

コンピュータが文章の構造を認識可能なデータを作成する最も簡単な方法は、**Word などの編集ソフトが提供するスタイル機能（見出し等）を使用すること**です。

また、HTML などのマークアップ言語やタグを使用して構造を記述することも有効です。

## Tips ▶ Word の見出し機能を活用しよう

Word には、文章内の任意の部分を見出し（章や節などのタイトル）としてコンピュータに記憶させ、文章全体を階層構造にすることができる「見出し機能」が備わっています。

具体的には、下図のとおり見出しとして設定したい行にカーソルを移動し、「ホーム」タブ内の「スタイル」グループから任意の見出しレベルを選択することで、見出しを設定することができます。

② 任意の見出しレベルを選択

① 見出しを設定する箇所をクリック

※フォントや文字飾りなどの設定を行う場合は、「スタイル」グループの右下をクリックし、詳細な設定を行う

見出し 1

見出し 2

見出し 3

本文

## ■ 原則 2 : 文章内に整形のための符号や文字を含めない

例えば、段落の頭を 1 文字下げのためにスペースを入れる、1 行の幅を狭めるため途中で改行する、といったことを行った場合、コンピュータはそのスペースや改行の意味を識別することができません。

このような場合は、スペースや改行を使用せず、インデントやページ設定などにより文章を整形しましょう。

## ■ 原則 3 : 文章内の図表の元になった表形式データを添付する

文章内に表を含む場合は、その表の内容もデータとして活用できるように、元になった表データを機械判読に適した形式で別途添付しましょう。(表形式データの作成については 28~37 ページ参照)

また、グラフなどの図を含む場合も、それらを作成する元となった表形式データを添付することが望ましいと言えます。

表形式データを添付する方法としては、オープンデータ公開ページに文書データと一緒に掲載するほか、文書データ内に元の表形式データへのリンクを載せるとより親切です。

### ■ 地理空間情報とは

地理空間情報とは、空間上の位置を表す情報と、それに関連する様々な事象に関する情報から構成されるデータです。

地理空間情報には、地形図や土地利用図、航空写真などといった基盤となる地図（基盤地図）と、人口などの統計情報や道路等の台帳情報、気象情報、顧客リスト等のデータベースなどといった多様なテーマの情報（主題図）があります。

これらの地理空間情報を処理することができる地理情報システム（GIS）は、位置情報を基準として基盤地図と主題図をそれぞれ階層（レイヤー）化して重ね合わせ、位置関係の把握やデータの表示、検索、分析などを行うことが可能です。

## Column ▶ 地理空間情報のオープンデータ化

オープンデータを公開する大きな意義の1つは、データが開放されることで所有者の枠を超えた横断的な活用が可能となり、これまで結びつくことのなかった複数のデータを掛け合わせた新しい分析結果やサービスなどが生まれる点にあります。

こうした複数のデータの掛け合わせを「マッシュアップ」と呼びますが、地理空間情報を用いた分析やサービスは、まさにマッシュアップの代表格と言えます。

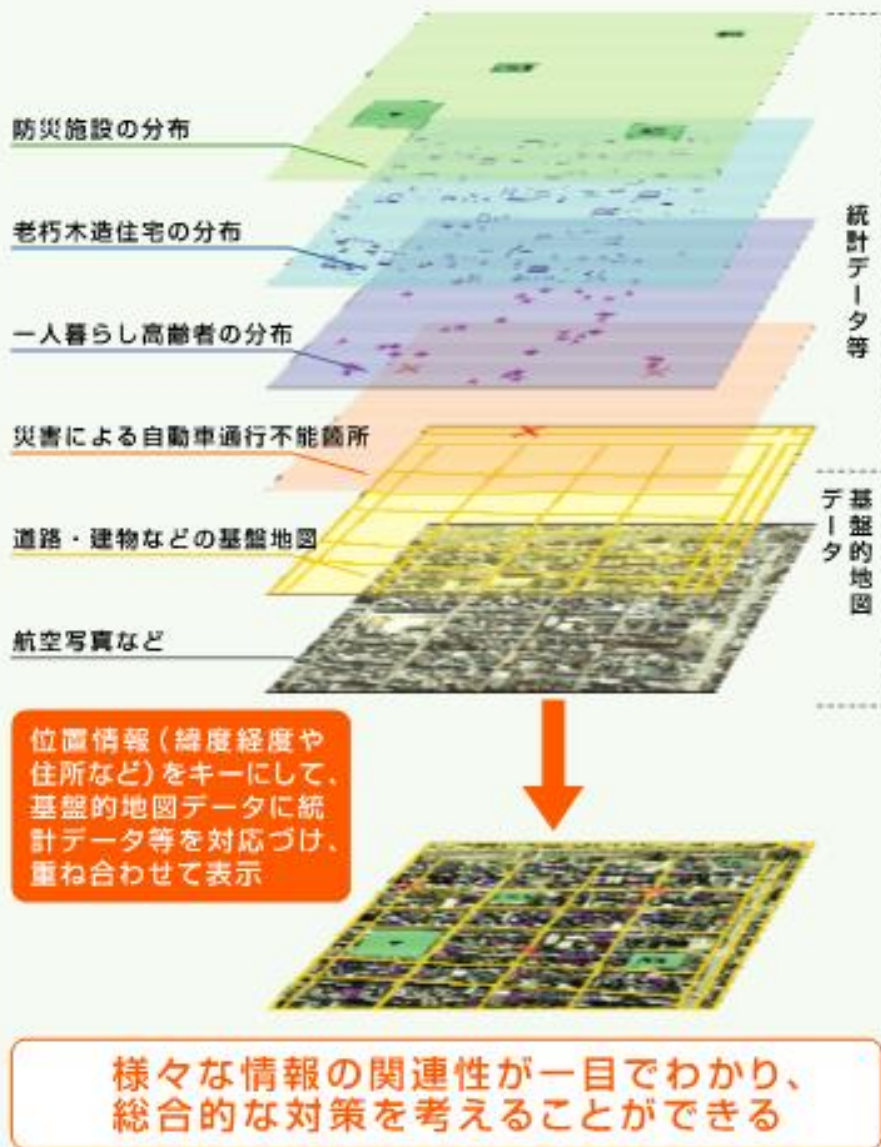
なぜなら、地理空間情報は1種類のみで利用するよりも、**複数の情報を階層（レイヤー）として重ねることにより真価を発揮する**からです。

また、地理空間情報は、他の種類のデータとも連携できる可能性があります。

例えば、「○○施設一覧」という情報が表形式データで存在する場合、このデータ内に各施設の緯度・経度などの座標データが含まれていれば、地図上に落とし込み、他の地理空間情報と組み合わせることが可能です。

地理空間情報をオープンデータ化することにより、庁内における都市計画・環境・固定資産税・防災などの各分野を超え、さらに市民活動団体や民間企業、学術機関などが保有するデータともマッシュアップされることで、これまで考えもしなかった活用方法が生まれるかもしれません。

## [災害対策における地理情報の重ね合わせ例]



※出典（上図）：国土交通省「GIS とは」

([http://www.mlit.go.jp/kokudoseisaku/kokudoseisaku\\_tk1\\_000041.html](http://www.mlit.go.jp/kokudoseisaku/kokudoseisaku_tk1_000041.html))

## ■ 地理空間情報に適したファイル形式

地理空間情報に使用されるファイル形式には、大きく分けて**ラスタ形式**と**ベクタ形式**の2種類があります。

ラスタ形式は、ベクタ形式に比べるとファイルサイズが大きくなりやすい、拡大すると画質が落ちるなどの欠点があります。

また、ベクタ形式をラスタ形式に変換するのは比較的容易であるのに対し、ラスタ形式からベクタ形式への変換は、特殊なソフトウェアを必要とするなど、非常に困難です。

よって、**地理空間情報をオープンデータとして公開する際には、原則、ベクタ形式を用いましょう。**

例えば、庁内の GIS でも使用されることの多い**シェープファイル (Shapefile)** はベクタ形式の一種であり、標準的な形式として広く利用されていることから、オープンデータに適しています。

また、GIS で利用する様々な情報を記述するマークアップ言語である GML を用いるのも有効です。



## Column ▶ ラスタ形式とベクタ形式

### ラスタ形式

ラスタ形式は、ピクセル（画素）とよばれるモザイク状のパーツで構成される画像データ形式であり、地理空間情報に限らず様々な画像の描画に用いられます。

「どこの座標に何色の点を置く」というように、各ピクセルが1つの位置情報と色の情報を持っています。

ラスタ形式は、解像度（ピクセルの密度）が決まっている関係上、拡大すると画像が粗くなる欠点があります。また、解像度を上げようとする、ファイルサイズが大きくなってしまいます。

一方で、ラスタ形式は各ピクセルの色により地図等を描画することから、ベクタ形式よりも複雑な描画を得意としており、航空写真や衛星写真に用いられます。

また、色調や濃淡による表現が可能であるため、気温や降水量などの境界線が不明瞭なデータにも適しています。

ラスタ形式には、例えば JPEG、GIF、PNG、BMP、TIFF などのファイル形式があります。

### ベクタ形式

ベクタ形式は、X座標・Y座標で表されるポイント（点）、複数のポイントを結ぶライン（線）、複数のポイントとラインで囲まれた領域であるポリゴン（面）の3要素で構成される図形の集まりとして地図を描画します。

地理空間情報の場合、各ポイントの座標には緯度・経度などを用います。

また、これら3要素に、「この線は道路である」「この点は信号である」といった情報（属性データ）を付加することで、各要素に意味づけをすることも可能です。

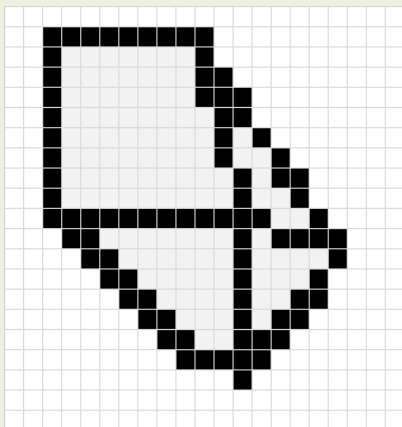


ベクタ形式は、現実に存在する建物などの地物や、街区等の区画など、境界が明確なものを表現するのに適しています。

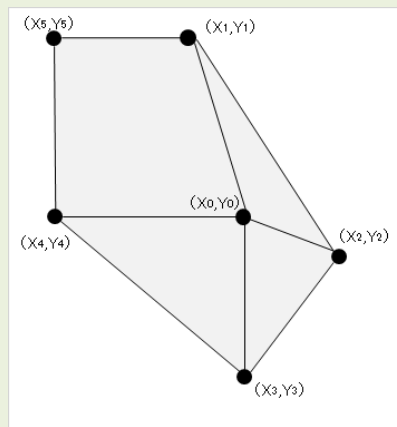
また、ベクタ形式は基点となるポイントの位置情報などのみを記録するため、ラスタ形式に比べるとファイルサイズが小さくなる傾向があるとともに、拡大しても画像が粗くなることはありません。

一方で、ラスタ形式に比べると、写真のような複雑な描画は不得意であり、降水量などの境界が不明瞭なデータにも適しません。

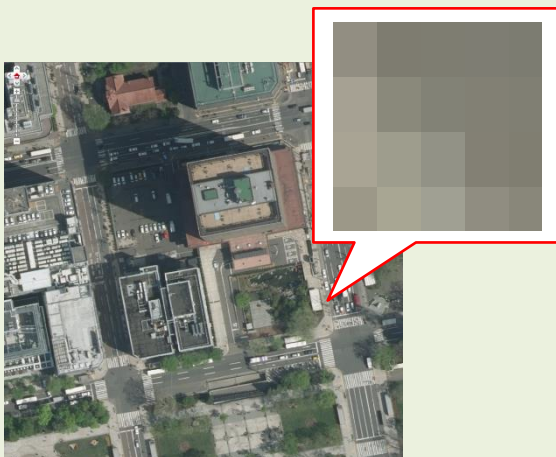
ベクタ形式には、Shapefile、EPS、Aiなどのファイル形式があります。



(ラスタ形式の構成要素イメージ)



(ベクタ形式の構成要素イメージ)



(ラスタ形式による描画イメージ)



(ベクタ形式による描画イメージ)

※札幌市役所本庁舎周辺の図

## ■ 地理空間情報を作成する際の基本原則

地理空間情報を機械判読に適した形式で公開するため、以下の基本原則に従ってください。

- 原則 1 : 位置情報に関するデータを付与する場合は、緯度・経度座標を用いる
- 原則 2 : 緯度・経度座標と併せ、準拠している座標参照系（測地系）を明記する
- 原則 3 : ベクタ形式データの作成に当たっては、地理情報標準プロファイル（JPGIS）を用いる

## ■ 原則 1 : 位置情報に関するデータを付与する場合は、緯度・経度座標を用いる

位置情報の付与されていない地図は単なる画像であり、地理空間情報とは呼べません。

業務で使用される GIS からデータを出力する場合は、通常、既に位置情報が付与されたデータが生成されると思われませんが、位置情報を含まない地図情報を提供する場合は、コンピュータによる処理が可能となるよう、位置情報を付与することが望ましいと言えます。

**位置情報を付与する際は、通常、緯度・経度を用います。**緯度・経度を特定する方法としては、以下のような方法が考えられます。

- ・国土地理院の公開する数値地図情報、基盤地図情報の座標を利用する。

【参考】国土地理院「数値地図（国土基本情報）」

(<http://www.gsi.go.jp/kibanjoho/kibanjoho40027.html>)

【参考】 国土地理院「基盤地図情報サイト」

(<http://www.gsi.go.jp/kiban/>)

- ・ 国土地理院の公開する地理院地図から座標を取得する。

【参考】 国土地理院「地理院地図」

(<http://maps.gsi.go.jp/#5/35.362222/138.731389/&base=std&ls=std&disp=1&vs=c1j0l0u0t0z0r0f0>)

- ・ 国土交通省が提供する「街区レベル位置参照情報アドレスマッチングツール」や「位置参照情報ダウンロードサービス」を利用し、住所から座標を取得する。

【参考】 国土交通省「位置参照情報ダウンロードサービス」

(<http://nlftp.mlit.go.jp/isj/service.html>)

国土交通省「位置参照情報ダウンロードサービス」

(<http://nlftp.mlit.go.jp/isj/>)

- ・ その他、民間の座標変換サービス等を利用する。

なお、地図ではない表形式などのデータでも、施設一覧やイベント一覧など、位置情報を付与することで地理空間情報との連携が望めるデータについては、可能な限り緯度・経度を付与するよう努めましょう。

## ■ 原則 2 : 緯度・経度座標と併せ、準拠している座標参照系（測地系）を明記する

緯度・経度などの座標は、「座標参照系（測地系）」という基準に従って表記されますが、**日本測地系**や**世界測地系**など、複数の座標参照系が存在するため、異なる座標参照系に準拠したデータ同士を組み合わせると位置情報に誤差が生じます。

複数の地理空間情報を組み合わせるためには、座標参照系を変換し統一する必要がありますので、地理空間情報のオープンデータ化に当たっては、**必ずそのデータが準拠している座標参照系を明記しましょう。**

通常、GIS から出力される地理空間情報には、座標参照系が明記されることが多いと思われます。

なお、平成 13 年の測量法改正により、日本では測量の基準として世界測地系が使用されています。法改正以前の日本測地系に準拠したデータについては、国土地理院が提供する「座標変換ソフトウェア」などを用い、世界測地系に準拠したものに交換してから提供すると、より親切です。

【参考】 国土地理院「座標変換ソフトウェア TKY2JGD」

(<http://www.gsi.go.jp/sokuchikijun/ky2jgd.html>)

### ■ 原則 3 : ベクタ形式データの作成に当たっては、地理情報標準プロファイル (JPGIS) を用いる

政府は、地理空間情報を異なるシステム間で相互利用する際の互換性を確保するため、データの設計・品質・記述方法・仕様の書き方などのルールを定めた「地理情報標準」を定め、政府の標準として使用するとともに、普及を推進しています。

この地理情報標準を元に、国際標準 (ISO 規格) や国内標準 (JIS 規格) に準拠しつつ体系化された国内実用標準として国土地理院が作成したのが、「地理情報標準プロファイル (JPGIS)」です。

ベクタ形式により地理空間情報を作成・提供するに当たっては、この JPGIS に基づくデータにすることで、データを相互利用しやすい環境が整備され、各主体が別々に作成した地理空間情報の共用が進むことが期待されます。

【参考】 国土地理院「地理情報標準プロファイル (JPGIS)」  
(<http://www.gsi.go.jp/GIS/jpgis-jpgidx.html>)

## 4

# PDF 形式の使用について

PDF 形式は人が見るための形式であり、コンピュータによる識別や編集が困難です。

そのため、オープンデータの観点からは、一般的に忌避される傾向にあります。

しかし、PDF 形式のデータをオープンデータとして提供することは、必ずしも否定されるべきことではありません。

オープンデータの原則としては、機械判読に適したデータを提供することとなっておりますが、一方で一般市民にとっては、CSV や Shapefile などのファイル形式はなじみがないものと思われる。

例えば防災マップや刊行物、キャラクターの画像など、広く市民に見ていただくことを目的とする場合、PDF のような目視のためのファイル形式を用いることは差し支えありません。

ただし、その場合であっても、PDF ファイルを作成する元となった機械判読に適したデータも併せて提供するように努めてください。

## ■ ファイル名の命名ルール

データをより機械判読に適したものとするためには、ファイル名を統一的なルールにより命名するとともに、コンピュータによる処理や検索を阻害しないファイル名にすることが望ましいと言えます。

ファイルの作成時には、最低限、以下のルールに従ってファイル名を設定しましょう。

### (1) 使用する文字

ファイル名に日本語などを用いると、処理するコンピュータによっては文字化けなどを起こす場合があるため、**半角英数字を用いましょう。**

また、区切りやスペースを必要とする場合は、「**\_ (アンダーバー)**」(半角)を使用しましょう。ファイル名には、「**.** (ドット)」「**-** (ハイフン)」などの他の記号も使用することができますが、これらの記号は検索の際の制御文字など特定の意味を持って使用されているため、コンピュータが混同し、データの処理や検索を阻害する恐れがあります。

また、同じく機械判読を阻害する可能性があるため、**空白 (スペース) も使用しないでください。**

### (2) ファイル名の重複の回避

ある部署で作成したデータのファイル名が他の部署のデータと重複していると、両方をダウンロードする際に新たにファイル名を付けなおす手間が発生します。

また、コンピュータによる処理や検索の際にも、同一のファイルとして混同されるため機械判読を阻害します。

ファイル名を考えるに当たっては、データの内容を推察できる名称にするとともに、部署名や日付を加えるなど、**他部署**

のデータと重複しにくいファイル名にしましょう。

## ■ データ内で使用する文字

### (1) 文字コード

文字コードとは、文字をコンピュータで処理するため、個々の文字や記号に割り当てられた番号の体系です。

文字コードは複数存在するため、文字コードとそれを読み取るコンピュータの OS 等の組み合わせによっては、特定の文字を処理することができず、文字化けや欠損が発生する場合があります。

これを回避するため、テキストファイルや CSV ファイルなどを公開する際は、**国際標準として広く利用されている文字コードである「UTF-8」を使用してください。**

具体的には、対象の CSV ファイル等をメモ帳で開き、UTF 形式で保存するなどの方法により、UTF-8 への変換が可能です。

### (2) 機種依存文字の回避

ローマ数字（Ⅰ、Ⅱ、Ⅲ・・・）や丸数字（①、②・・・）、複数の文字を1文字に収めた組文字（株、臈、cm、℃など）などのいわゆる機種依存文字は、扱うプログラムや OS によっては使用できない場合があります。

オープンデータとしてデータを公開する際は、様々なプログラムでデータが使用されることを想定する必要がありますので、このような機種依存文字は使用しないでください。



## ■ ファイル内容・プロパティの事前確認

作成したデータを公開する前に、データの内容やプロパティについて、下記の点を入念に確認してください。

### (1) ファイルの内容について

データ編集の変更履歴やコメント等が残っていないか確認し、削除してください。特に、変更履歴やコメント等が非表示になっていないか注意してください。

また、個人情報や札幌市情報公開条例における非開示情報など、公開することのできない情報が含まれていないか、あらためて精査してください。

### (2) プロパティの情報について

ファイルのプロパティには、作成者等の情報が記載されています。データを公開する前に、プロパティの情報に札幌市以外の団体名や個人の氏名などの機微な情報が記載されていないか確認し、削除してください。

また、プロパティ内の資料名は、公開する際の資料名と統一してください。

ただし、作成者（組織）や作成日時などのデータに関するデータ（メタデータ）は、オープンデータを利用するに当たって非常に重要な情報です。

後述しますが、**プロパティの情報を全て削除すればよいということではありません**ので、注意してください。

## 6

# メタデータの作成

### ■ メタデータの項目例

オープンデータの利用者にとっては、対象のデータそのものだけでなく、タイトル・URL・作成者・データ形式・注釈など、データそのものの内容・構造・性質などに関するデータ（メタデータ）も価値のあるデータです。

オープンデータの作成・公開に当たっては、例えば以下のようなメタデータを併せて公開しましょう。

項目	記載例	説明
タイトル	オープンデータ一覧	ファイルの内容を端的に表現したタイトル
データ概要	札幌市が公開しているオープンデータの一覧です。	ファイルの内容に関する説明。 注釈や文字コード、地理空間情報の測地系など、ファイル内に含めることのできなかつた情報も記載する。
ファイル形式	CSV	ファイルに用いられているフォーマット
分類	その他	オープンデータカタログサイトにおける分類
更新頻度	1年に1回	データの更新頻度
URL	http://……	データが掲載されている Web サイトなどの URL
登録日	2017-10-06	オープンデータカタログサイトに掲載された日付
最終更新日	2017-10-22	データを最後に更新した日付

項目	記載例	説明
備考	実際のデータ 公開状況と異なる場合があります。	その他特記事項

## ■ メタデータの提供方法

メタデータを提供する方法としては、オープンデータカタログサイトの掲載ページ上に情報を記載するほか、元のデータとは別にメタデータのファイルを作成し、掲載することが考えられます。

庁内でも使用されることの多い Microsoft Office 製品（Word、Excel 等）では、データの作成者や作成日などをプロパティに記載することができます。プロパティ内部の情報も重要なメタデータですので、個人情報等の公開に適さない情報を除き、可能な限り整備して提供できるよう努めましょう。

また、こうしたメタデータも、RDF などの機械判読可能な表現形式で提供することで、コンピュータによる処理や検索が容易になります。