

別添 用語集

1. 本書全般にかかる用語

用語	定義
オープンデータカタログサイト	オープンデータの情報を公開し横断的検索を可能とすることで、オープンデータを利用しやすくするためのポータルサイト
三重県オープンデータカタログサイト	三重県がオープンデータを公開しているサイト
BODIK ODCS	BigData & OpenData Initiative in Kyushu Open Data Catalog Siteの略。 地方自治体がオープンデータを公開するためのデータカタログサイトを無償で提供するクラウドサービスであり、「三重県オープンデータカタログサイト」で利用しているサービス。
自治体標準オープンデータセット	オープンデータの公開とその利活用を促進することを目的とし、政府として公開を推奨するデータと、そのデータの作成にあたり準拠すべきルールやフォーマット等を取りまとめたもの。
オープンデータ	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データ。
公共データ	政府、独立行政法人、地方公共団体等が保有するデータ。
情報提供者	オープンデータを提供する者又は機関。
情報利用者	オープンデータとして公開されているデータを二次利用する者又は機関。
二次利用	情報提供者の提供したデータを基に、情報利用者が何らかの編集・加工・改変などを行い、新たなデータを作成することや、情報提供者の提供したデータの単なる複製や再配布を行うこと。
ライセンス	情報提供者がデータを提供する際に指定する利用条件。
クリエイティブ・コモンズ・ライセンス	著作物の再利用についての条件などに関する意思表示を手軽に行えるようにするために、国際的に使用されているライセンス。著作権者が著作権を保持したまま「著作権者の表示をする」又は「非営利に限定する」など様々なレベルの条件を選択して表示し、その条件で二次利用を行うことを許可する事ができる。
CC BY	クリエイティブ・コモンズ・ライセンスの表記の一種。著作権者のクレジット（情報提供者名、作品タイトル、URL）を表示すれば、情報利用者は営利目的を含めて自由にデータを改変、複製、再配布することができる。
機械判読（可能）	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、文字など）を自動的に編集・加工・改変などできること。「Machine Readable」の日本語訳であり「機械可読」ともいう。
機械判読（可能）性	対象とするデータに対する機械判読の可能性。
メタデータ	オープンデータとして公開されているデータ自体がどのようなデータであるかを示す情報。
表形式データ	行と列の、縦横2次元状に配列されたデータ。
文書形式データ	1次元状に配列された文字を主な構成要素とし、一部図や表などを含み、人間がそれを読むことによって人間に何らかの作用を与えることを目的としたデータ。
地理空間情報	空間上の特定の地点又は区域の位置を示す情報（空間データ）と、これに関連付けられた情報（属性情報）を組み合わせたデータ。
タグ	データの特徴を説明したキーワード。例えば、「防災」「観光」「医療」など。データを検索するためのキーになる。
全国地方公共団体コード	都道府県、政令指定都市、政令指定都市以外の市、特別区、指定都市の区、町村、一部事務組合、地方開発事業団、広域連合などの公共団体にコードを設定したもの。桁数は、5桁のアラビア整数に1桁のチェックディジット（検査数字）を加えた6桁。
法人番号	数字のみで構成される13桁の番号。国の機関、地方公共団体、設立登記法人等の団体に付与される。

2. ファイルにかかる用語

用語	定義
テキストファイル	文字データだけで構成されたファイル。どんな機種のコンピュータでも共通して利用できる。⇔非テキストファイル
オープンなフォーマット	仕様が公開されており、それが国際標準化団体により、標準化されているファイルのフォーマット・形式。特定のアプリケーションに依存しないデータ形式になっている。
CSV (Comma-Separated Values)	いくつかのフィールドをカンマ「,」で区切り、1行を1レコードとして改行で区切ったテキストファイル。
Office Open XML	Microsoft Office 2007以降の標準のファイルフォーマットで、オープンなフォーマットである。Microsoft Wordで作成したOffice Open XMLのファイルの拡張子はdocx、Microsoft Excelの場合はxlsxとなる。Office Open XMLは、XMLで記述されたテキストファイルと、画像などの非テキストファイルをzip形式でパッケージングして格納されたものなので、機械判読性がある。
OpenDocument Format	ワープロ文書や表計算ソフトのワークシートなど、オフィスソフトの扱う文書ファイルを保存するためのオープンなフォーマットのファイル形式。ワープロソフトで作成したOpenDocument Formatのファイルの拡張子はodt、表計算ソフトの場合はodsとなる。Office Open XMLと同様、XMLで記述されたテキストファイルと、画像などの非テキストファイルをzip形式でパッケージングして格納されたものなので、機械判読性がある。
JSON(JavaScript Object Notation)	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、または、その言語により記述されるテキストファイル。主にデータ交換用に使われる。オープンなフォーマットであり、JavaScript専用のデータ形式では決してなく、様々なソフトウェアやプログラミング言語間におけるデータの受け渡しに使えるよう設計されている。

マークアップ言語	文書の一部を「タグ」と呼ばれる特別な文字列で囲うことにより、文章の構造（見出しや段落など）や、修飾情報（文字の大きさなど）を、テキストファイル中で表現していく言語。（図 2-1）
XML(Extensible Markup Language)	個別の目的に応じたマークアップ言語を創るために汎用的に使うことができる上位のマークアップ言語、または、その言語により記述されるテキストファイル。オープンなフォーマットである。
HTML(HyperText Markup Language)	Web上の文書を記述するためのマークアップ言語、または、その言語により記述されたテキストファイル。オープンなフォーマットである。
XHTML(Extensible HyperText Markup Language)	Web上の文書を記述するためのXMLベースのマークアップ言語、または、その言語により記述されたテキストファイル。オープンなフォーマットである。
Shapefile	ESRI社が開発した地理空間情報に関するファイルフォーマットで、仕様も公開されている。機械判読をするための技術的な難易度は高い。
KML(Keyhole Markup Language)	地理空間情報を表現するXMLベースのマークアップ言語、または、その言語により記述されたテキストファイル。オープンなフォーマットである。
GML(Geography Markup Language)	地理空間情報を表現するXMLベースのマークアップ言語、または、その言語により記述されたテキストファイル。オープンなフォーマットである。国土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。
RDF(Resource Description Framework)	Web上にある事物（リソースという）に関する情報を記述するための枠組み（ルール）である。RDFは主語（Subject）、述語（Predicate）、目的語（Object）の3つの要素でリソースに関する情報を記述する。XMLなどの言語を使用して記述する。
文字コード	コンピュータ上で文字を扱うために、文字一つ一つに割り当てられた固有の数値（コード）、またはその数値を8bit単位のデータに変換して取り扱う方式。（図 2-2）
Shift-JIS	コンピュータ上で日本語を含む文字列を表現するために用いられる文字コードの一つ。
UTF-8	世界標準規格の文字コード。日本語に限らず、ほぼすべての文字に対応している。
テキストエディタ	コンピュータでテキストファイルを作成、編集、保存するためのソフトウェア（プログラム）
メモ帳	Microsoft Windowsに付属するテキストエディタ。
正規表現	文字列の集合を一つの文字列で表現する方法。正規表現を利用すれば、たくさんのテキストの中から容易に見つけたい文字列を検索することができる（図 2-3）。正規表現は、プログラム言語やテキストエディタなどにより記述方法が異なる。

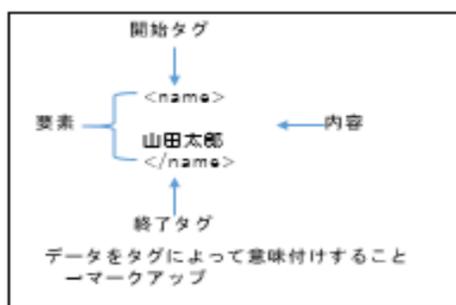


図 2-1 マークアップ言語の概念

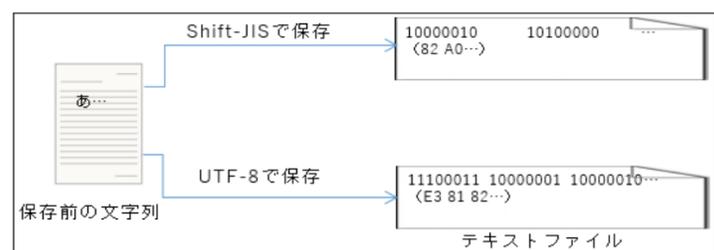


図 2-2 文字コードとテキストファイルの関係

(例)
パソコンのCドライブ直下のtxtファイルを表す場合
C:/* .txt

図 2-3 正規表現の例

3.表形式データにかかる用語

用語	定義
データセット	テーブルを含む表形式データのまとまり。表計算ソフトの「シート」に相当。
テーブル	表全体を指す。1行のヘッダ、1行以上のレコードからなる。
ヘッダ	表の各列の名前を保持する行。1個以上のセルからなる。
レコード	表の1行からなる要素。1個以上のセルからなる。
フィールド	表の1列からなる要素。1個以上のセルからなる。
セル	表の1行1列からなる要素。表計算ソフトの「セル」に相当。

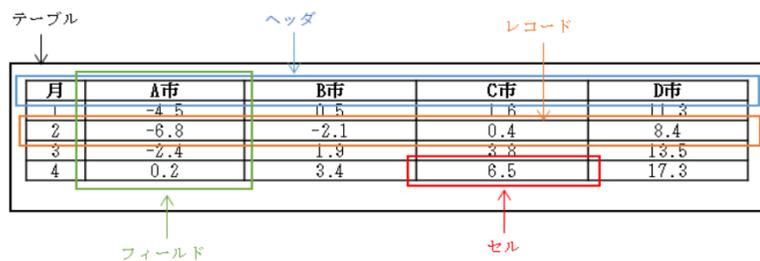


図 3-1 表形式データにかかる用語の定義

4.文書形式データにかかる用語

用語	定義
CMS(Content Management System)	Webサイトを構築、編集するためのシステムで、Webサイトを構成するHTML形式のファイルや画像、レイアウト情報などを一元的に作成・保存・管理する。また、HTML形式のファイルを画面上で作成する機能があり、CMSの利用者はHTML言語が分からなくても、HTML形式のファイルを作成することができる。

5.地理空間情報にかかる用語

用語	定義
地理情報システム (GIS)	地理空間情報を総合的に管理・加工し、視覚的に表示し、高度な分析や迅速な判断を可能にするシステムである。
Mie Click Maps	三重県の地図情報サービス。地図を使用して三重県が保有する位置や場所に関する情報を提供しているサイトである。
M-GIS	三重県が業務ツールとして使用している簡易GISアプリケーションのこと。Mie Click Mapsで位置や場所に関する情報を提供するのに必要なKMLファイルを作成することができる。
座標系	緯度経度またはXYといった座標値によって地球上の位置を表す際の決まりのこと。一般的に使用される座標系には、大きく分けて地理座標系と投影座標系の2つがある。(図 5-1)
投影座標系	3次元である地球を2次元の平面に投影し、XY座標で表現する座標系。2次元に投影するための方法を地図投影法といい、用いる投影法や設定する原点などの違いによって、さまざまな投影座標系が存在する。なお、投影座標系では地球の形状なども考慮して投影を行うため、地理座標系の定義も含まれる。
地理座標系	3次元である地球上の位置を、緯度と経度で表現する座標系のこと。
緯度・経度	場所を数値的に表現する方法の一つ。地球を楕円の物体とみなし、その場所における楕円体面の法線が赤道面となす角度を「緯度」、その場所を通る子午線がグリニッジ子午線となす角度を「経度」と定め示す(図 5-2)。通常、緯度・経度は60進数(度分秒)で表すが、分・秒を使わずに10進数(度)で表す場合がある。
測地系	地球上の緯度・経度の測量を行う際に採用する系。測地系により地球の球体モデルの形・大きさ、重心の位置などが異なるため、同一地点でも採用する測地系により緯度・経度が異なってくる。
日本測地系	日本において、明治時代から2002年3月まで、測量基準として採用されていた測地系。旧測地系ともいわれる。

世界測地系	人工衛星などを用いた観測によって明らかとなった地球の正確な形状と大きさに基づき、世界的な整合性を持たせて構築された経度・緯度の測定の基準で、国際的に定められている測地基準系。日本でも、2001年の測量法改正以降、世界測地系が測量基準として使用されている。世界測地系には「国際地球基準座標系」や「日本測地系2000」などがある。
国際地球基準座標系	宇宙測地技術の観測データに基づいて国際地球回転・基準座標系事業が提供する地心直交座標系である。その座標系は、地球の重心を原点とし、X軸を経度0度の子午線と赤道との交点の方向に、Y軸を「東経90度」の方向に、Z軸を北極の方向に定義されている。
平面直角座標系	日本固有の投影座標系で、全国を19地域に分割し、各地域に原点を設置する座標系。座標値は原点からのメートル距離となる（図 5-3）。なお、三重県の任意の地点を平面直角座標系で表す場合は、第6系の座標値で表す。
ベクタ形式	画像を、点の座標とそれを結ぶ線などの数値データをもとにして演算によって再現する方式。Shapefile、KML、GMLはいずれもベクタ形式のファイルである。
ラスタ形式	画像を色のついたドットと呼ばれる点の羅列・集合として表現するデータ再現方式。PNG、JPEG、GIFなどがある。

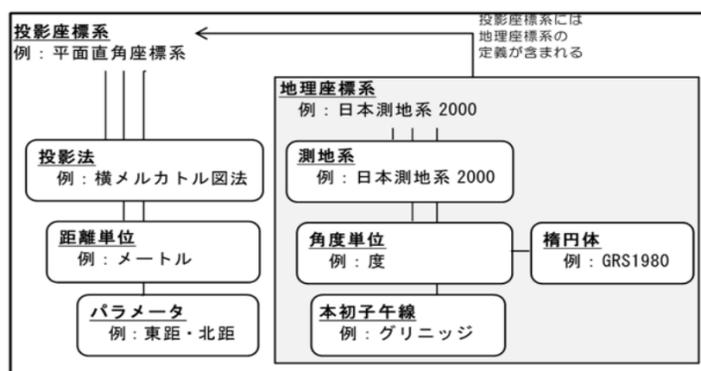


図 5-1 座標系の体系

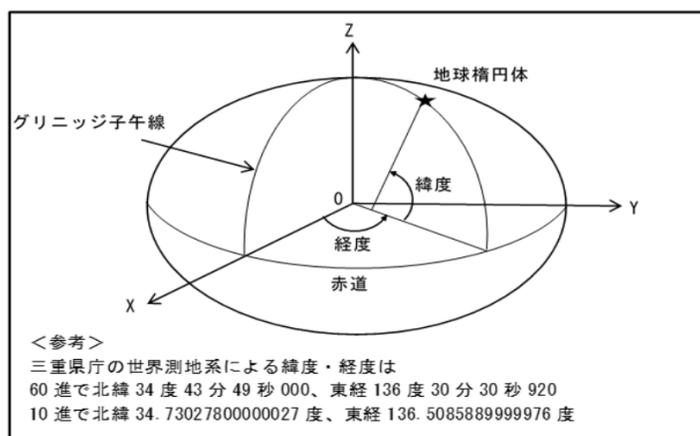


図 5-2 緯度・経度の表し方

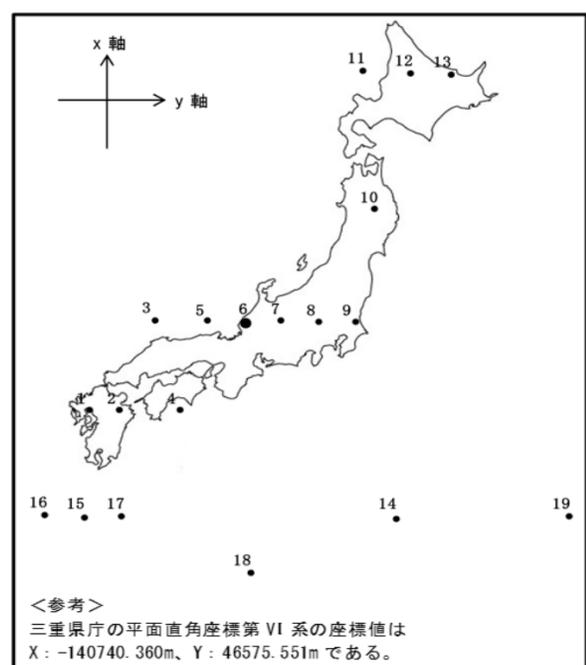


図 5-3 平面直角座標系の各系の原点