

# オープンデータ活用研究部会の活動について

水野 信也<sup>†1</sup>

オープンデータは、利用に適したデータ形式で2次利用が可能なものとして広く公開されてきた。省庁からの提供だけでなく自治体からのオープンデータも数多く提供されている。オープンデータ公開の意識が広がることは素晴らしいことだが、公開する側では目的やコストにおいて問題がある。オープンデータを公開することによってどのような効果があるのか、またどのようなデータが必要とされているのかが明確でない。またオープンデータ公開のためには誰かがデータを作成し、さらにメンテナンスをする必要がある。様々な業務がある中、優先的にオープンデータを扱うのは実際難しい。またデータを公開するコンセンサスを得る必要もある。

このような問題がある中、今後オープンデータの発展を考える場合、オープンデータの活用事例を作っていくことが一番である。本研究部会では広くオープンデータを扱うことができる技術的なスキームを提案している。図1にあるようにオープンデータはそれだけで意味をなすことは難しい。自社や他のデータと結びつくことで効果を期待できる。

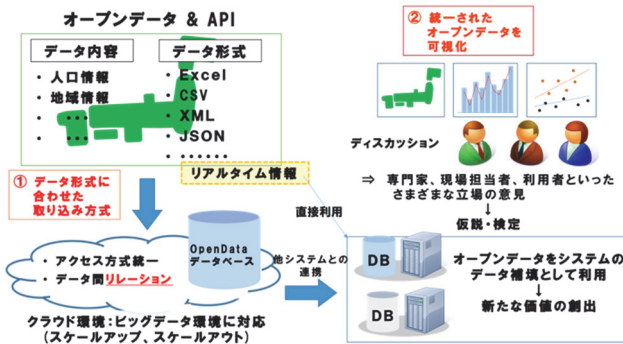


図1 オープンデータ活用の流れ

オープンデータには大きく分けて2種類のデータがある。一つは避難地域や医療機関情報など地理的データを含むもの、もう一つは人口統計など数値データで構成されているものがある。これらを有効活用するために実施した講座の内容は表1のようになっている。前半部分は地理的データを扱い、それらを地図上に可視化して、検索を実施するのである。これらはWebシステムを連携することで効果的に利用できると考え、データベース構築からWebサイト構築まで内容を含んでいる。図2は浜松市津波避難ビル一覧

の地理情報を表示したものである。後半部分は数値データを解析できるように、多変量解析の基本からクラスター分析、主成分分析をRにて実施し、多変量解析の基本フローを学習している。図3は静岡県袋井市自治会別世帯数及び人口を利用した自治体別クラスター分析の様子である。

表1 オープンデータを利用するための基本スキーム

| 1. オープンデータ利用と地理情報連携  |                         |
|----------------------|-------------------------|
| 1-1                  | オープンデータの定義と種類           |
| 1-2                  | オープンデータ利用のための環境設定(LAMP) |
| 1-3                  | オープンデータとWebAPIとの連携      |
| 1-4                  | オープンデータの可視化(地図情報)       |
| 2. オープンデータを利用したデータ解析 |                         |
| 2-1                  | 統計ツールRの利用               |
| 2-2                  | オープンデータへの統計手法の利用(基本分析)  |
| 2-3                  | オープンデータのクラスター分析         |
| 2-4                  | オープンデータを利用した推定, 検定      |



図2 浜松市津波避難ビル一覧の可視化した様子

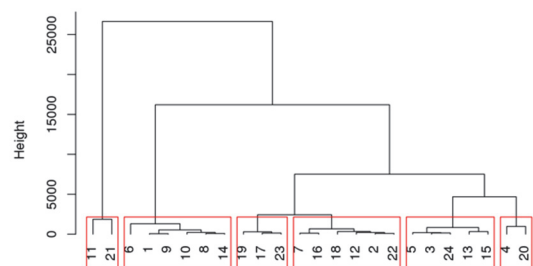


図3 袋井市自治会別世帯数及び人口を利用したクラスター分析

本研究会ではオープンデータの意義を考え、広く利用出来るオープンデータ利用プラットフォームを構築している。今後も新しい技術の吸収とともに、オープンデータ利用の活動を続けていく。

<sup>†1</sup> 静岡理科大学 (連絡先: mizuno.shinya@sist.ac.jp)