

# 政令指定都市におけるオープンデータの推進

同志社大学政策学部真山ゼミ

○青木悟 (Satoru Aoki)・勝部育 (Iku Katsube)・田口莉帆 (Riho Taguchi)・平川晴臣 (Haruomi Hirakawa)・福井優里 (Yuri Fukui)・山越奈美 (Nami Yamagoshi)

(同志社大学政策学部政策学科)

キーワード: オープンデータ、政令指定都市、ポータルサイト

## 1. オープンデータの現状

近年、世界的にオープンデータの活用に関する取り組みが盛んに行われており、日本でも取り組みを進めている自治体が増加傾向にある。しかし、世界的に見ると日本より進んでいる国は少なくはない (高木 2013)。

オープンデータとは、林(2014)によると、政府が保有する公共性の高い機械判読可能な公共データで、企業や個人が自由に編集・加工、再配布などができ、かつオープンライセンスで提供されるデータと定義される。また、Open Knowledge Foundation (URL 1)は、あらゆる目的で誰でも自由に使用、変更、共有することができるデータであると定義している。二つの定義は、「誰もが自由に使えるデータ」という点で共通している。

## 2. オープンデータの目的

IT 総合戦略本部が策定した「電子行政オープンデータ戦略」によると、オープンデータの目的は三つである (URL 2)。一つ目は、「透明性の向上」である。公共データが公開されることで、政府と住民との間で情報の非対称性が小さくなり、行政の透明性が向上する。また、政府の活動に対する住民の適切な評価につながる。二つ目は、「官民協働の推進」である。透明性が向上すると、市民が情報を得られ、官民協働の推進が期待できる。近年、人口減少や技術革新などの様々な課題に政府だけで対応することは困難になっており、官民協働の推進が求められている (庄司 2014)。三つ目は、「経済効果」である。オープンデータ推進による生産性向上や新産業の創出による経済効果は官民合わせて 6.4 兆円とされる (田中・高木 2015)。

## 3. 調査の概要

上述のオープンデータ活用の三つの効果を生み出すために、自治体には「市民に活用され、便利で有益なオープンデータサイト」を整備することが求められる。具体的には、アクセスしやすく、必要なデータが効率的に入手できるなど、様々な要素を考える必要がある。ただし、今回の調査では、アクセス数を中心に調査した。そもそも、サイトを見られることがなければ市民に活用される

ことがなく、便利で有益であることも測ることができないからである。

そこで、アクセス数が多いオープンデータサイトにはどのような特徴があるのかを探るため、独自で作成した評価基準やランキングなどを用いて調査した。

## 4. 方法

2021 年 8 月中旬から 9 月中旬の間で、全国 20 の政令指定都市のオープンデータサイトについて調査した。調査項目は以下の四つである。

### (1) クリック数

使いやすさを客観的な数値に置き換える手段として、6 人を被調査者とし、Google にて「〇〇市オープンデータ」と検索しての結果から、それぞれの自治体の「2020 年 4 月の推計人口」が確認できるまでの平均クリック数を調査した。なお、「2020 年 4 月の推計人口」がオープンデータサイト上に確認できず、市の HP 上等に確認できた自治体が 5 つあった。

### (2) データ数

オープンデータとして公開されているデータの数を調査した。「データセット」として数を確認することができた自治体は 12、できなかった自治体は 7、不明であった自治体は 1 であった。

### (3) 推奨データセット数

「推奨データセット」数とは、「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、公開する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」として内閣官房情報通信技術 (IT) 総合戦略室がまとめたものである (URL 3)。この数は「地方公共団体へのオープンデータの取組に関するアンケート結果」(URL 4)を参照した。

### (4) アクセス数

全政令指定都市にメールおよび電話にて問い合わせ、2020 年 4 月から 2021 年 7 月にかけてのオープンデータのトップページのページビュー数を調査した。なお、返答が得られなかった自治体は 4 つあり、サイトの改修等の事情で旧データのア

クセス数を含む自治体が2つあった。アクセス数は16ヶ月分の平均である。

## 5. 結果

クリック数、データ数、アクセス数で比較可能だったのは20都市のうち7都市であった(表1)。

表1 調査結果

	クリック数	データ数	推奨データ	得点	アクセス数
札幌	6.50	227	12	16	4541
京都	13.33	542	0	10	2351
大阪	15.00	240	15	14	1328
神戸	5.50	100	2	13	1269
岡山	21.33	82	12	9	192
北九州	6.67	605	1	14	2770
熊本	10.00	26	5	9	389
相関	-0.5628	0.53445	0.61908	0.76011	

これらの都市でそれぞれの項目ごとに順位付けをして、1位から順番に7~1点を与えていった。例えば、神戸市はデータ数が5番目に多いので3点となる。表1のように、アクセス数と得点の相関を調べると相関係数は0.76011であり強い相関がある。このことから、オープンデータサイトのアクセス数を大きくするには、クリック数を少なくし、データ数、推奨データセット数を大きくすればよいことがわかった<sup>1</sup>。なお、推奨データセット数は厳格な基準があり、類似している内容のデータでもカウントされないことがある。その類似しているデータで相関を見たところ、0.772982と少し高くなる。

さらに、調査をする過程で三つの発見があった。一つ目は、ポータルサイトを設置した方が良いということである。ポータルサイトは操作がしやすく、閲覧ランキングや活用事例など様々な情報が掲載されており、年齢や能力を超えて様々な人が親しみやすいものである。二つ目は、古いデータが公開されていることである。例えば、静岡市は人口のデータを2017年11月から更新していなかった。三つ目は、データの量だけでなく、データの種類が重要ということである。京都市では「石の道しるべ」を地図上に表示する「いしぶみアプリ」が開発された。これは「いしぶみリスト」という京都市特有なオープンデータをもとに作成されている。これは多様な種類のデータがオープンデータの創造的な活用につながることを示唆している。

## 6. 政策提言

今回の調査は、自治体におけるオープンデータ活用への取り組みを検討するための予備的調査という性格があり、不十分な点があることは否めない。とはいえ、以上の調査から「市民に活用され、

便利で有益なオープンデータサイト」をつくるために少なくとも三点が指摘できる。一点目は、まず初めにポータルサイトを準備することである。二点目は、ポータルサイト内のデータを充実させることである。具体的にはデータ数・推奨データ数・データの種類を増やしていくことである。三点目は、サイトの整備をすることである。例えば、クリック数を減らすためにはポータルサイト内の検索機能を強化することなどが考えられる。こうした取り組みに加え、データの更新をし続けることは大前提となるのは言うまでもない。

### 〈参考文献〉

- ・庄司昌彦(2014)「オープンデータの定義・目的・最新の課題」『智場』(119) pp. 4-15
- ・高木聡一郎(2013)「欧州におけるオープンデータ政策の最新動向」『情報管理』Vol. 55. No. 10, pp. 746-753
- ・田中秀幸・高木聡一郎(2015)「インフラとしてのオープンデーター政府・自治体保有データのオープン化が日本経済に及ぼす影響ー」『フィナンシャル・レビュー』平成27年第4号(通巻124号)
- ・林雅之(2014)「オープンデータを理解するための10のポイント」『智場』(119) pp. 125-134

〈参考URL: いずれも最終閲覧日2021年10月10日〉

1. Open Knowledge Foundation「The Open Definition」  
<http://opendefinition.org>
2. IT総合戦略本部(2012)「電子行政オープンデータ戦略」  
[http://www.kantei.go.jp/jp/singi/it2/pdf/120704\\_siryu2.pdf](http://www.kantei.go.jp/jp/singi/it2/pdf/120704_siryu2.pdf)
3. 政府CIOポータル(2021)「推奨データセットについて」  
<https://cio.go.jp/policy-opendata>
4. 政府CIOポータル(2021)「地方公共団体へのオープンデータ取り組みに関する回答一覧」  
<https://cio.go.jp/policy-opendata>
5. 全国20政令指定都市ポータルサイト・HP

<sup>1</sup> それぞれの項目とアクセス数を重回帰分析した結果、重相関Rは0.871605と相関が認められた。一方で、有意Fは0.18488と基準となる0.05より高い数字となった。これはサンプル数が7と少ないことが原因と考えられる。より多くのサンプルを比較することが今後の課題となる。