

研究指導 中澤 真 准教授

地方振興のためのオープンデータ —地域への関心を高めるオープンデータのあり方—

永井 真美

1. はじめに

近年オープンデータが世界的に注目されている。これは2011年にアメリカ主導の下、「オープンガバメント・パートナーシップ」が立ち上がったことがきっかけである。オープンガバメント・パートナーシップとは、政府の補助金の公正な運用、汚職の廃絶を目的とした団体のことである。行政の活動を一般市民がデータとして共有することで、国を国民のものとして認識し、政治活動を透明なものとする取組みを行っており、今では60カ国ほど参加をしている[1]。

政府のデータが公開されたことで、そのデータを経済活動に利用できるようになった。さらに、インターネットが普及したことで、いつでもどこでも誰でもアクセス可能となった。このように、オープンガバメントによって無償で公開されたデータ、つまりオープンデータは大きなビジネスチャンスにつながる可能性を秘めている。

日本では政府が2012年に発表した電子行政オープンデータ戦略において、公共データを公開し始めたことで、注目されるようになった[1]。これに刺激を受けて、一部の地方公共団体でもオープンデータに取組み始めたが、その動きはまだかなり限定的である。地方公共団体のオープンデータが進まない理由は、データを公開しても効果やメリットがあるのか不明であり、どのようなデータを公開することが望ましいかという点も明らかになっていないことが大きい。

そこで本研究では地方公共団体に焦点をあてて、すでに公開されているオープンデータを調査し、どのようなデータを公開すると地域への関心が高まるのかその要因を分析する。そして地方振興のためのオープンデータ戦略について一石を投じる。

2. オープンデータの現状と課題

オープンデータとは、行政が今まで内部で保管していた各種情報をインターネットなどで広い範囲に公開し、企業やその地域の住民だけではなく、誰もが利用できるようにしたデータをいう。それは世の中に無数にある情報を使いやすい形にし、自由に加工、編集することで、新たな価値を生み出すことを目的としている[2]。さらに二次利用を認めて公開することが必須条件とされているため、企業や個人はデータを加工し商用利用できるという利点がある。すでに海外では、天候予測や農作物の収穫予想サービスである「The Climate

Corporation」(アメリカ)や、環境マップサービス「EnviroFIND」(イギリス)など、政府が公開したオープンデータを使ったビジネスが展開されている[3]。

2.1 日本におけるオープンデータの取組み

日本政府は2012年7月に電子行政オープンデータ戦略として、政府自らが積極的にデータを公開すること、機械判読可能な形式で公開すること、営利目的・非営利目的を問わず活用を促進すること、取組み可能な公共データから速やかに公開などの具体的な取組みに着手し、成果を確実に蓄積することを決めた[4]。その結果、気象庁の地震データを活用した、リアルタイムで地震の警戒情報をスマートフォンに配信する「ゆれくるコール」など、いくつかのサービスでオープンデータが利用されるようになった。

また、民間団体でもオープンデータを促進する取組みが行われるようになってきている。たとえば、大手企業がスポンサーとなって開催されている「Linked Open Data チャレンジ Japan」というイベントはオープンデータの活動を広める目的で開催されており、幅広い分野におけるオープンデータづくりと、データを活用した作品および取組みを表彰するコンテストが行われている[5]。他にも、地方公共団体が主催するオープンデータハッカソンというアプリコンテストを実施し、地域への関心を集めるためにオープンデータが活用されている。

2.2 オープンデータに関する日本と海外の比較

オープン・ナレッジ・ファウンデーション・ジャパン(OKFJ)の2013年の調査結果[6]によると、各国のオープンデータの進捗度は、アメリカが60点中54点、イギリスが51点と高く、日本は46点と先進国8カ国の中で6位と低い順位にとどまっている。進捗度の差異の要因を詳細に見てみると、日本のオープンデータポータルサイトに公開されているデータ数は2013年12月現在で9498、アメリカは98852、イギリスは14198となっている。またアメリカやイギリスでは、多くのデータがCSV、RDFなどの機械判読が容易なものであり、かつレベルが高いデータであるが、日本の場合はPDF、xlsが多い。このようなフォーマットのデータは加工する際に人手を要することが多いため、活用するにあたり不便である。このようにデータ数でもデータの品質においても、日本のオープンデータは海外と比較すると遅れていることがわかる。

¹ <http://www.rcsc.co.jp/hpd/index.html>

2.3 地方公共団体のオープンデータの可能性と課題

企業が公共データを利用するとき、県や市町村単位のより詳細な情報を必要とすることが多いため、地方公共団体が持つ大量の統計情報や地図情報は、企業にとって利用価値が高いデータである。つまり、地方公共団体が公共データをオープンデータとすることで、企業と地域が密着したビジネスへと発展する可能性が見込まれる。さらに地方公共団体が持つその土地独自のデータは地域に密着しており、これをオープンデータとして利用できるようにすることは、その地域の住民にとって有益である[4]。

そこで、日本の地方公共団体のオープンデータの公開状況について、OKFJの調査[8]に加えて、検索サイトで地方公共団体名とオープンデータというキーワード検索により調査した。この結果を表1に示す。この表からわかるように、全1741地方公共団体においてオープンデータを公開しているのは32団体にとどまっており、日本の地方におけるオープンデータの取組みが進んでいないことが明らかになった。

表1:オープンデータを公開している地方公共団体

自治体名			
鯖江市	静岡県	金沢市	大和市
会津若松市	越前町	大垣市	藤沢市
千葉市	相模原市	須崎市	横浜市
越前市	大阪市	八王子市	福岡市
室蘭市	前橋市	内灘町	さいたま市
横手市	御前崎市	永平寺町	川崎市
坂井市	敦賀市	小浜市	福井市
裾野市	野々市市	流山市	焼津市

地方公共団体のオープンデータの取組みが進まない理由として、以下の2点が挙げられる。1点目はオープンデータの利用イメージやニーズが具体化されていないため、公開すべきデータを地方公共団体が選別できないでいること、2点目はオープンデータを公開する地域への効果が明確ではないことである[9]。1点目の利用イメージやニーズの具体化は地方公共団体と民間との連携強化によってある程度解決できると考えられる。そのため本研究では、2点目の問題を解決するため、地方公共団体がどのようなオープンデータを公開することが、地域への関心を高め、ひいては地方振興につながるのかを明らかにする。

3. 各地方公共団体のオープンデータの進捗調査

本章では、地域への関心が集まる要因を明らかにするために、先に示した32の地方公共団体のオープンデータを詳細に調査した結果について述べる。

3.1 地方公共団体のオープンデータに関する調査方法

まず各地方公共団体のすべてのオープンデータについて、表2に示したどのフォーマットレベルに対応するか、また、

表3に示したどのカテゴリとニーズのクラスに属するかについて整理した。ここで、表2のフォーマットごとのレベルは、ティム・バーナーズ＝リーによって提唱された結果[10]を用いており、単純な画像データのようなフォーマットは機械判読の難しさから低いレベルに、逆にRDFのようなフォーマットは機械判読に優れ、データ加工も容易であることから高いレベルに設定されている。なお、レベルは5段階評価であるが、最高レベルはリンクやアプリを入れたデータであり、現在のオープンデータはここまでには至っていない。また、表3のカテゴリとニーズのクラスは情報通信白書平成26年版より、民間企業における認識について記した項目の、ニーズの高い公共データのアンケート[11]に基づいて作成しており、全回答数332に対し、10パーセント以上の回答数があるものをニーズの高いクラスA、6パーセント以下を中庸なクラスB、そしてその他のカテゴリをクラスCとして分類している。なお表3の小文字のアルファベットはaに近づくほどニーズは高くなり、rに近づくほどニーズは低くなることを意味している。

表2:フォーマットの

レベル	
レベル	形式
4	rdf
3	csv
3	xml
3	Shape
3	json
2	xls
2	txt
2	doc
1	pdf
1	html
1	pgn
1	jpg

表3:カテゴリとニーズの

クラス	
カテゴリ	クラス
a 地図・地下	A
b 交通	
c 防災・保安・安全	
d 投資計画・建築	
e 医療・介護	
f 統計・調査	B
g 個人・住民情報	
h 公開方法	
i 気象	
j 入札・調達・補助金	
k 環境・エネルギー	C
l 特許	
m 法人情報	
n 許認可	
o 電波	
p 法令	C
q その他	
r 施設	

3.2 データの調査結果

次に、どのようなデータが地域への関心を高める効果があるのかその要因を探るために、地方公共団体を2群に分ける評価基準を考える。最終的には、それぞれの2群に対して平均値の差の検定を実施し、どの要因の影響が強いかを明らかにする。これをまとめたものが表4である。また表5は表4をもとに各地方公共団体のオープンデータを評価した結果である。

表4：オープンデータの評価基準

オープンデータ評価基準		
①	名称	説明
①	平均レベル	全データのレベルの平均値
②	データ数	公開しているデータ数
③	アプリ数	オープンデータを利用したアプリの数
④	レベルの加重	レベルを基に重みをつけた合計
⑤	ニーズの加重平均	ニーズを基に重みをつけた平均
⑥	ニーズの加重和	ニーズを基に重みをつけた合計
⑦	クラスA	ニーズクラスAに属する割合
⑧	クラスB	ニーズクラスBに属する割合
⑨	クラスC	ニーズクラスCに属する割合

表4の①の平均レベルは表2に基づき、地方公共団体ごとに公開しているすべてのオープンデータのレベルの平均値を算出したものである。②は単純に各地方公共団体が公開しているオープンデータの総数を基準とするものである。③はそれぞれの自治体のオープンデータを活用したスマートデバイス用アプリの個数を評価基準としたものである。④は表2のレベルに基づきデータに重みをつけた加重和を基準としたものである。⑤は表3のニーズクラスに重みを割り当て、この重みに基づき全データの加重平均を算出した値であり、⑥はこの加重和である。⑦は表3のニーズの高いクラスAに属する割合を評価基準としたものであり、同様の条件でクラスのみ異なるものが⑧および⑨である。

表5：各地方公共団体のオープンデータの評価

地方公共団体名	①	②	③	④	⑤	⑥	⑦	⑧	⑨
会津若松市	3	66	9	176	30.37	4464	0.45	0.14	0.41
内灘町	3	4	0	12	21.00	126	1.00	0.00	0.00
永平寺町	3	9	0	27	24.60	492	0.44	0.00	0.56
越前市	4	117	4	351	16.72	3027	0.21	0.05	0.74
越前町	3	5	0	15	39.00	468	0.20	0.00	0.80
大垣市	3	20	1	50	26.45	873	0.60	0.00	0.40
大阪市	3	17	0	51	5.57	78	0.35	0.00	0.65
小浜市	3	3	0	9	3.00	9	0.00	0.00	1.00
御前崎市	2.5	4	0	10	34.50	276	0.70	0.00	0.30
金沢市	3	15	1	45	31.80	795	0.00	0.00	1.00
川崎市	2.98	43	0	128	38.37	2187	0.98	0.00	0.02
さいたま市	3	8	0	24	39.00	1170	1.00	0.00	0.00
坂井市	3	10	3	30	3.00	75	0.70	0.00	0.30
相模原市	3	17	0	51	15.29	627	0.53	0.06	0.41
鯖江市	4	17	85	68	23.57	660	0.47	0.00	0.53
静岡県	2.92	60	0	176	33.52	5061	0.63	0.00	0.38
須坂市	2.38	14	1	36	34.78	1113	1.00	0.00	0.00
裾野市	3.96	19	2	75	26.48	1218	0.43	0.00	0.57
千葉市	2.67	6	4	16	35.73	393	0.81	0.00	0.19
敦賀市	4	18	0	44	27.48	1374	0.36	0.00	0.64
流山市	3.11	47	1	124	19.48	1617	0.32	0.00	0.68
野々市市	2.75	4	0	11	39.00	390	1.00	0.00	0.00
八王子市	2	400	0	800	39.00	3276	0.90	0.01	0.09
福井市	4	20	0	80	25.91	1140	0.40	0.00	0.60
福岡市	2.08	81	0	166	34.25	3117	0.86	0.00	0.14
藤沢市	3	28	0	70	29.07	1686	0.64	0.00	0.36
前橋市	3	15	0	45	21.72	543	0.40	0.00	0.60
室蘭市	3	26	4	222	37.50	2700	0.31	0.61	0.08
焼津市	3	3	0	9	33.86	237	0.67	0.00	0.33
大和市	3	51	0	153	21.62	1881	0.20	0.00	0.80
横手市	3	11	3	33	24.60	984	0.45	0.00	0.55
横浜市	3.01	148	0	383	19.79	3840	0.35	0.01	0.64

4. オープンデータの分析

本章では地方振興のために重み付けしたデータに基づき地域への関心を上昇させる要因を分析する。

4.1 分析の指標と比較パターン

地方振興のためにはまずその地域に多くの人からの関心

が集まる必要がある。地域への関心の高さを測ることは容易ではないが、本研究ではWebの検索に注目してグーグルトレンドを地域の関心度を測る指標とした。グーグルトレンドとは、グーグルが提供しているウェブ検索において、特定の検索キーワードを使用した検索回数の時系列的変化で参照できるサービスのことである。今回の検証では各地方公共団体がオープンデータを公開した時期の前後を比較できるように、2011年7月から2015年1月までの検索結果に関するデータをグーグルトレンドより取得し、時間経過にともなう検索数の変化率を算出した。すなわち、この値が地域の関心度の変化を表す指標となる。なお、各地方公共団体の変化率は、その地方公共団体がオープンデータを公開する前の過去6か月分の平均値と、公開の翌月からの6か月分の平均値を比較して算出している。

本研究では、地域の関心度を高めること、すなわち上述の変化率を正の値で大きくする(上昇率を大きくする)ことが、地方振興につながると仮定して研究を進める。そこで、どのようなオープンデータを提供することが、この上昇率を大きくすることにつながるのか、その要因を探るために表4のそれぞれの評価基準を組み合わせた12パターンについて比較検証する。表6はその組み合わせパターンをまとめたものである。

次節では、検証したパターンの分析結果と関心を高める要因について考察する。

表6：評価基準を基に検証したパターン

パターン	
1	平均レベル3と3より大きい値の2群
2	平均レベル3と3より小さい値の2群
3	データ9以下と10~28の値の2群
4	データ数10~28と43以上の値の2群
5	フォーマットのレベルを考慮したニーズの加重和の平均に近い値と上位の値の2群
6	フォーマットのレベルを考慮したニーズの加重和の平均に近い値と下位の値の2群
7	フォーマットのレベルを考慮したニーズの加重平均の平均に近い値と上位の値の2群
8	フォーマットのレベルを考慮したニーズの加重平均の平均に近い値と下位の値の2群
9	フォーマットのレベルを考慮しないニーズの加重和の平均に近い値と上位の値の2群
10	フォーマットのレベルを考慮しないニーズの加重和の平均に近い値と下位の値の2群
11	フォーマットのレベルを考慮しないニーズの加重平均の平均に近い値と上位の値の2群
12	フォーマットのレベルを考慮しないニーズの加重平均の平均に近い値と下位の値の2群
13	アプリを持っている地方公共団体とアプリを持っていない地方公共団体
14	クラスAに属するデータの割合が多い地方公共団体と少ない地方公共団体
15	クラスBに属するデータの割合が多い地方公共団体と少ない地方公共団体
16	クラスCに属するデータの割合が多い地方公共団体と少ない地方公共団体

4.2 分析結果と考察

表6に示した2群の分け方は、まずそれぞれの評価基準ごとにヒストグラムを作成して傾向を把握し、各評価基準の平均値に近い値の地方公共団体をA群とし、平均より遠い値の地方公共団体をB群としている。B群は平均値より上位か下位の二つに分けられるため、A群と上位のB群、下位のB群と比較した。パターン1の平均レベルについては、32団体の平均レベルにあまり差がなかったため、平均レベルがもっとも多いレベル3の地方公共団体をA群とし、レベル3より大きい値と小さい値をそれぞれのパターンのB群にして検証し

ている。それぞれのパターンについて、32団体の上昇率の平均と、各群の上昇率の平均を比較した結果を表7に示す。

表7: 上昇率の平均値とt検定の結果

	評価基準	全32団体の上昇率の平均値	A群の上昇率の平均値	B群の上昇率の平均値	p値	t検定の有意差
1	①	0.0193	0.0383	0.0007	0.2626	
2				-0.0071	0.1613	
3	②		0.0304	0.0087	0.3681	
4				-0.0149	0.0278	有
5	④、⑥		0.0257	-0.0057	0.3791	
6				0.0286	0.9345	
7	④、⑤		0.0572	0.0060	0.1569	
8				-0.0137	0.1135	
9	⑥		0.0660	-0.0055	0.0243	有
10				-0.0313	0.0062	有
11	⑤		0.0127	0.0008	0.7435	
12				0.0307	0.6149	
13	③		0.0108	0.0064	0.2002	
14	⑦		0.0064	0.0031	0.3544	
15	⑧		0.0167	0.0199	0.1312	
16	⑨		0.0028	0.0116	0.6230	

表7において、全体の上昇率の平均値と各群の平均値を比較して差異が大きい場合には、その要因によって関心度に影響が及んでいると考えることができる。全体平均よりも上昇率が高いとき、その要因は地域の関心度を高める効果があり、逆に低いときはその要因が地域の関心度に影響を及ぼすことはないということになる。結果として、地域の関心度を高める要因は、評価基準①の平均レベルと②のデータ数、⑤ニーズの加重平均、⑥ニーズの加重和であると解釈できる。つまり、地方公共団体のオープンデータは地方振興につなげるために、レベルが高いだけでなく、民間のニーズに対応した多様なカテゴリに属するデータを一定数以上公開する必要があるといえる。

次に各パターンのA群とB群の上昇率の平均値に対し、ウェルチのt検定を用いてその有意差を判定した。本研究において、t検定における判断基準はp値を求め、95%信頼区間で判断した。そのためp値<0.05で帰無仮説を棄却する。その結果、パターン4・9・10に有意差があると判断された。パターン4はデータ数が平均に近い群と、平均より多いデータ数を持つ群であるため、公開するオープンデータの数を増やすことは、地域への関心度を高める効果があるといえる。また、パターン9・10に優位差があることが示されたことにより、ニーズの高い地図や交通、防災に関するデータを提供することでも、地域への関心を集めることができると結論づけられる。

以上、全体平均と比較した結果においては、レベルとデータ数、ニーズが、t検定の結果からはデータ数とニーズが、地域への関心度を高める要因であることが明らかになった。このことから、地域の関心度を高めるためにはレベルも必要ではあるが、それ以上に多様なカテゴリを持つデータを数多く揃えることが重要であると結論づけられる。これらの要因を満たすデータを公開することで地方振興のためのオープンデータ戦略が展開できる。

5. むすび

本研究では地方振興のために地方公共団体がオープンデータを公開することによって、地方への関心が集まる要因を、実際の地方公共団体のオープンデータを調査・分析することで明らかにした。また今回は、地方振興を測る指標としてグーグルトレンドを用いたが、これはキーワードが検索された数の推移であるためオープンデータに関しての地方振興度を測る指標としては不十分な点もある。また各地方公共団体のオープンデータも少ないため、データの傾向をつかむためのサンプルが少なかつた点も改善の余地がある。今後はロコミ分析などほかの手法も用いてサンプルを収集・生成し、十分なサンプル数を得て研究する予定である。

参考文献

- [1] 庄司昌彦, NIKKEI, オープンデータ情報ポータル, なぜ今、オープンデータか, <http://opendata.nikkei.co.jp/article/201302073530619462/>
- [2] 林雅之, オープンデータ超入門, インプレスR&D, 2014.
- [3] 総務省, オープンデータの活用事例～海外および日本の事例～, <http://datameti.go.jp/wp-content/uploads/2014/01/03b558b1126807662402d3ebb9a98986.pdf>
- [4] 総務省, 電子行政オープンデータ戦略, http://www.kantei.go.jp/jp/singi/it2/pdf/120704_gaiyou.pdf
- [5] LOD Challenge 2014, Linked Open Data チャレンジJapan 開催趣旨, <http://lod.sfc.keio.ac.jp/challenge2014/objective.html>
- [6] 庄司昌彦, オープンデータセンサス2013日本政府は世界30位。世界最先端を目指すには加速が必要, オープン・ナレッジ・ファウンデーション・ジャパン, <http://okfn.jp/2013/11/14/opendatacensus2013/>
- [7] 鈴木恭子・市嶋洋平, 日経BigData, 地方自治体がオープンデータ利活用を力を入れる訳, <http://business.nikkeibp.co.jp/article/bigdata/20140223260108/>
- [8] OKFJ, OKJ ニュースレター 2014/11月号 (Vol.16), <http://okfn.jp/2014/11/>
- [9] 総務省, 自治体におけるオープンデータ～オープンデータの効果～, Open DATA METI, <http://datameti.go.jp/wp-content/uploads/2014/01/eefa3d737b0418fb0d5fd3efc005849b.pdf>
- [10] オープンデータ流通推進コンソーシアム, オープンデータガイド～オープンデータのためのルール・技術の手引き～, <http://www.opendata.gr.jp/news/docs/opendata-guide-v1.pdf>
- [11] 総務省, 情報通信白書平成26年版, 民間企業における認識, <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h26/html/nc132210.html>