

# 自治体におけるオープンデータ,ビッグデータ,AIの捉え方 に関する実態調査

## The survey of municipalities' viewpoints of Open Data, Big Data and AI

劉 兆媛<sup>1</sup>, 田中 秀幸<sup>2</sup>

LIU ZHAOYUAN and Hideyuki TANAKA

<sup>1</sup> 東京大学大学院 学際情報学府 The University of Tokyo GSII

<sup>2</sup> 東京大学大学院 情報学環・学際情報学府 The University of Tokyo III / GSII

**Abstract** This survey reviews municipalities' viewpoints of Open Data, Big Data and AI. By confirming whether municipalities mentioned Open Data, Big Data and AI in their master plans, we can have a grasp about how municipalities capture these three words. In this survey, we divide municipalities into some groups and show the proportion of municipalities mentioned these three words in their master plans to all municipalities that have signed in DATA.GO.JP in the group. Then we compare proportions between different groups and confirm whether there are statistically significant variations.

**キーワード** 自治体, オープンデータ, ビッグデータ, AI

### 1. はじめに

2013年に主要国首脳会議G8で合意されたオープンデータ憲章を皮切りに,世界中で政府データを中心としたデータのオープン化が積極的に行われるようになった。それに加え,近年,第四次産業革命とも呼ばれる技術革新が進むことで,オープンデータやビッグデータ,AIが代表的なキーワードとして多様な分野で言及されるようになった。これらの技術は産業,行政等様々な分野で大いに期待されている。本調査の目的は,この趨勢に対して,自治体がオープンデータ,ビッグデータ,AIをどのように捉えているかを明らかにすることである。日本では,DATA.GO.JPという各府省庁において公開されている公共データの検索を可能とするデータカタログサイトが2014年から運営されている。2017年現在で,DATA.GO.JPに登録している地方自治体は239ある。本調査はこの239の自治体を対象にし,自治体の総合計画におけるオープンデータ,ビッグデータ,AIについて言及の有無を確認し,各自治体におけるこの三つのキーワードの推進状況を把握する。

### 2. 研究目的と意義

本調査では,自治体総合計画におけるオープンデータ,ビッグデータ,AIの言及状況を確認した上で,自治体の種類,地域属性,人口規模,総合計画の策定年度,ソフトウェア業従業者数の集中程度によって自治体をグループに分けて,その言及状況の特徴を見出すことを試みる。本調査はオープンデータ,ビッグデータ,AIの自治体における推進状況を把握し,自治体のこの三つのキーワードの捉え方を明らかにすることで,今後自治体が主導して行政情報を始めとする多様なデータの活用を推進する上で,重要な意義を有するものと考えられる。

### 3. 調査方法

調査方法として,DATA.GO.JPでオープンデータサイトを登録している239の自治体を対象に,各自治体の「総合計画」を確認し,計画書内にオープンデータ,ビッグデータ,AIという三つの言葉が含まれているか否かを確認する。この作業を通じて,自治体のオープンデータ,ビッグデータ,AIに対する認識と推進状況の把握を試みる。

総合計画のうち,2011年以降に策定され,2017年6月現在も運用されているものを対象とした。総合計画は,各自治体によって,基本構想,基本計画,実施計画に分けられていることもあるが,これらが一つにまとめられているところもある。分けられている場合,策定時間と運用期間が前述の条件を満たしていれば,基本構想及び基本計画も対象とした。総合計画がない場合,「将来ビジョン」,「振興プラン」等,総合計画に同等と考えられる公開資料を調査対象とした。

### 4. 調査結果と考察

#### (1)自治体のDATA.GO.JP登録状況

全国1741の自治体の中で,DATA.GO.JPに登録している自治体は239ある。内訳は,都道府県が37,政令指定都市が20,その他の市区町村が182である(表1)。自治体の種類別でDATA.GO.JPに登録している自治体の割合を図1で示した。全体として,DATA.GO.JPへ登録している自治体は全国自治体13.7%にすぎない。

登録自治体の割合を自治体種類別で見ると,都道府県と政令指定都市自治体で登録率が高い。都道府県の78.7%が登録しているほか,全ての政令指定都市が登録している。しかし,他の市区町村自治体の登録数

は全体の 10.9%に止まっている（図 1）。地域別で都道府県とその他の自治体の登録数を見ると、北海道、北関東、南関東、東海、近畿、四国及び沖縄の都道府県は全て DATA.GO.JP に登録している（図 2）。その他の自治体について、DATA.GO.JP への登録率が高いのは北陸(39.5%)、南関東(25.1%)、東海(20.2%)である（図 3）。

表 1 自治体種類別 DATA.GO.JP 登録数

	全体数	DATA GOJP 登録自治体数	登録数が全体数に占める割合 (%)
自治体全体	1741	239	13.7
都道府県	47	37	78.7
政令指定都市	20	20	100
他の市区町村	1674	182	10.9

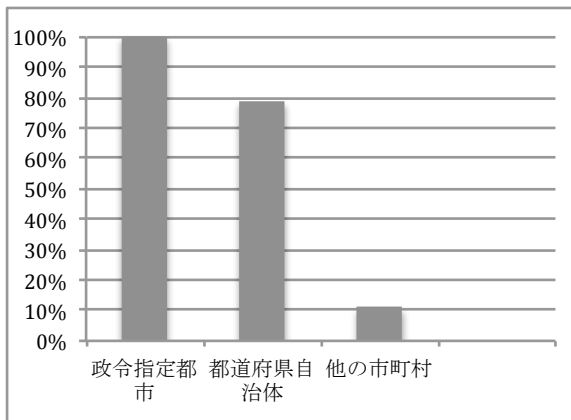


図 1 自治体種類別登録数の全体数に対する割合

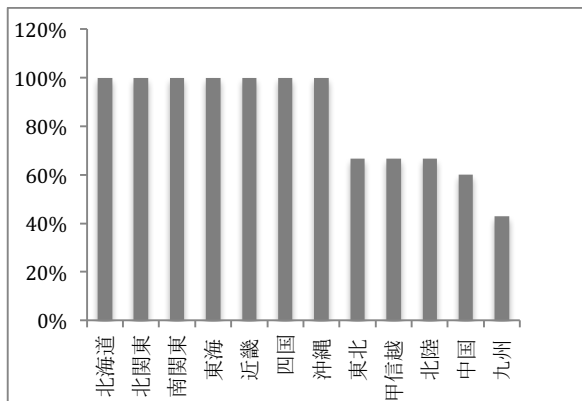


図 2 地域別で都道府県登録数の全体数に占める割合

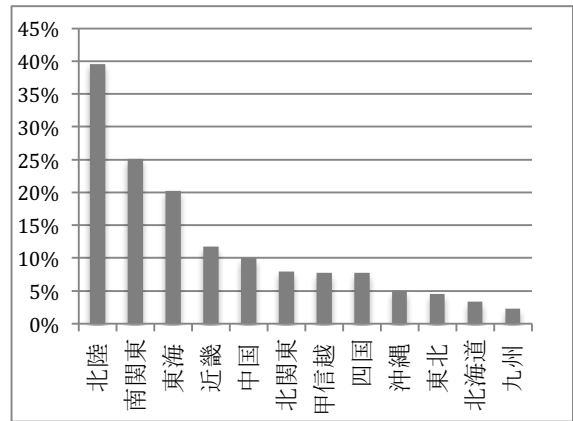


図 3 地域別でその他の自治体登録数の全体数に占める割合

(2)オープンデータ,ビッグデータ,AI 言及の有無状況 (全体)

DATA.GO.JP に登録している 239 の自治体の中で、「総合計画」またはそれに相当する計画でオープンデータ,ビッグデータ及び AI に関する言及の有無状況を確認した。三つのキーワードに関する言及の有無を表 2 で示した。DATA.GO.JP へ登録している自治体のうち三つのキーワードとも言及のない自治体が半分以上であった。オープンデータ,ビッグデータ,AI へ言及している自治体はそれぞれ 31.0%, 13.8%, 5.4%である(表 2)。

表 2 「総合計画」でオープンデータ,ビッグデータ,AI 言及の有無状況

登録自治体数	総数	オープンデータ言及有り*	ビッグデータ言及有り	AI 言及有り	三つとも言及無い
全体	239	74 (31.0)	33 (13.8)	13 (5.4)	152 (63.6)
都道府県	37	12 (32.4)	15 (40.5)	9 (24.3)	19 (51.4)
政令指定都市	20	7 (35.0)	2 (10.0)	1 (5.0)	12 (60.0)
他の市区町村	182	55 (30.2)	16 (8.8)	3 (1.6)	121 (66.5)

\*()内は言及数の登録に占める割合%。

具体的に、オープンデータについて、総合計画の内容を確認したところ、ほとんどの自治体で「情報公開の推進」「市の保有するデータの公開」等に関する言及がみられたものの、「オープンデータ」という言葉を用いていたのは 74 の自治体にすぎないことが確認された。言い換えると、69%の自治体は DATA.GO.JP に登録したものの、総合計画でオープンデータを強調するまでには至っていない。あるいは、総合計画を策定したのが先で、オープンデータ等はその後で取り組み始めたため、総合計画の内容に含まれていなかった可能性も考えら

れる。また、ビッグデータと AI に関しては比較的最近注目された言葉であるほか、行政との関係がオープンデータ程緊密とは考えられていないため、総合計画で言及されることが比較的少ない可能性がある。総合計画の中での三つのキーワードの有無から見ると、全体としては、ビッグデータ、AI は言うまでもなく、オープンデータの推進も十分に進んでいるとは言い難い。

### (3)自治体種類別

都道府県自治体と政令指定都市自治体を一つのグループとし、その他の市区町村自治体と比較すると、三つのキーワードの言及が有る自治体数（以下：言及数）の DATA.GO.JP 登録自治体数（以下：登録数）に占める割合は図 4 のようである。

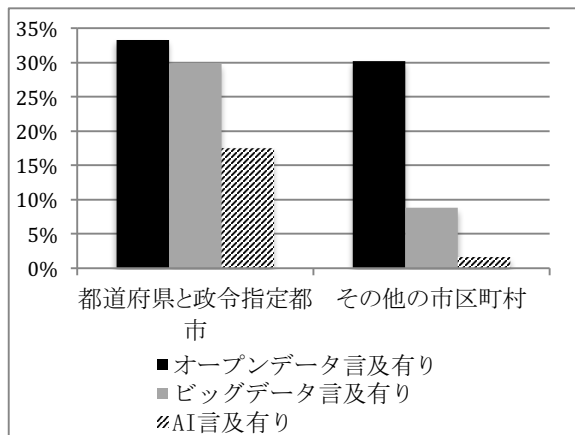


図 4 自治体種類別で言及数の登録数に占める割合

自治体種類別で言及数の登録数に占める割合の差が統計的有意であるかどうかを確認するために、比率の同一性検定を用いて割合の差の検定を行なった。P 値が .05 以下なら統計的に有意な差異があると考えられる。検定で用いたデータと検定結果は表 3、表 4 で示している。

まず、オープンデータの言及について、都道府県と政令指定都市のグループの言及数の登録数に占める割合と、その他の市区町村グループの言及数の登録数に占める割合との差の比率の同一性検定を行った。その結果、P 値が .780 で、有意な結果は得られなかった。

表 3 自治体種類別のオープンデータへの言及及び検定結果

	DATAGOJP 登録自治体 数	オープン データ言 及有り	オープン データ言 及無い	言及数の登 録数に占め る割合 (%)
都道府県と政 令指定都市	57	19	38	33.3
その他の市区 町村	182	55	127	43.3
検定結果	X-squared = 0.078, df = 1, p-value = .780			

次に、ビッグデータの言及についても前記と同様に、自治体種類別で言及数の登録数に占める割合の差の検定を行なった。その結果、自治体種類別のビッグデータの言及数の登録数に占める割合の差について、P 値が <.01 で統計的に有意だと考える。

表 4 自治体種類別ビッグデータへの言及及び検定結果

	DATAGOJP 登録自治体 数	ビッグデ ータ言及 有り	ビッグデ ータ言及 無い	言及数の登 録数に占め る割合 (%)
都道府県と政 令指定都市	57	17	40	29.8
その他の市区 町村	182	16	166	8.7
検定結果	X-squared = 14.416, df = 1, p-value = <.01			

### (4)ソフトウェア業従業者集中程度別(都道府県と政令指定都市以外)

内閣府の地域区分<sup>1</sup>で見るとオープンデータ、ビッグデータ、AI について言及有る自治体数の地域の全自治体数に占める割合は図 5 に示すとおりである。北陸、南関東、東海、北関東の四つの地域で他の地域に比べて言及している自治体の割合が高い。

また、「平成 27 年特定サービス産業実態調査」<sup>2</sup>によると、全国ソフトウェア業の従業者数の 75% が東京都、神奈川県、愛知県、大阪府と福岡県に集中している。そこで、都道府県と政令指定都市を除いて、DATA.GO.JP 登録の自治体をこの 5 都府県にある市区町村自治体とその他の地域の市区町村自治体に分けて、三つのキーワードへの言及がある自治体数の登録数に占める割合を図 6 に示した。

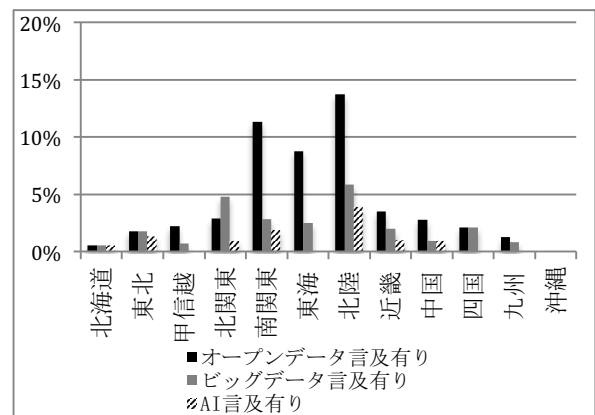


図 5 地域別で言及数の地域全自治体数に占める割合

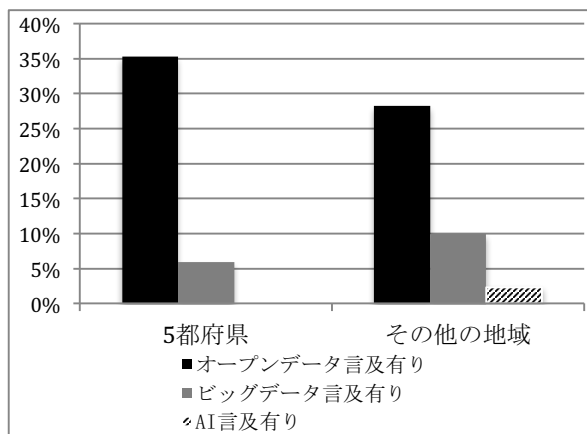


図6 ソフトウェア業従業員集中程度別で言及数の登録数に占める割合

前節と同様に、ソフトウェア業従業員集中程度別で言及数の登録数に占める割合の差について、比率の同一性検定を行った。検定で用いたデータと検定結果は表5,表6で示している。オープンデータ、ビッグデータへの言及について、ソフトウェア業従業員の集中している5都府県のグループの言及数の登録数に占める割合と、その他の地域のグループの言及数の登録数に占める割合との差を検定した。結果として、オープンデータ、ビッグデータについてP値はそれぞれ.453,.567と統計的に有意な差は認められなかった。

表5 ソフトウェア業従業員数集中程度別オープンデータへの言及及び検定結果

	DATAGOJP登録自治体数	オープンデータ言及有り	オープンデータ言及無い	言及数の登録数に占める割合(%)
5都府県	51	18	33	35.3
それ以外の市区町村	131	37	94	28.2
検定結果	X-squared = 0.563, df = 1, p-value = .453			

表6 ソフトウェア業従業員数集中程度ビッグデータへの言及及び検定結果

	DATAGOJP登録自治体数	ビッグデータ言及有り	ビッグデータ言及無い	言及数の登録数に占める割合(%)
5都府県	51	3	48	5.9
それ以外の市区町村	131	13	118	9.9
検定結果	X-squared = 0.329, df = 1, p-value = .567			

#### (5)人口別(都道府県と政令指定都市以外)

都道府県と政令指定都市以外の市区町村の人口規模を10万人で分けて、オープンデータ、ビッグデー

タ及びAIについて言及数の登録数に占める割合は図7に示すとおりである。

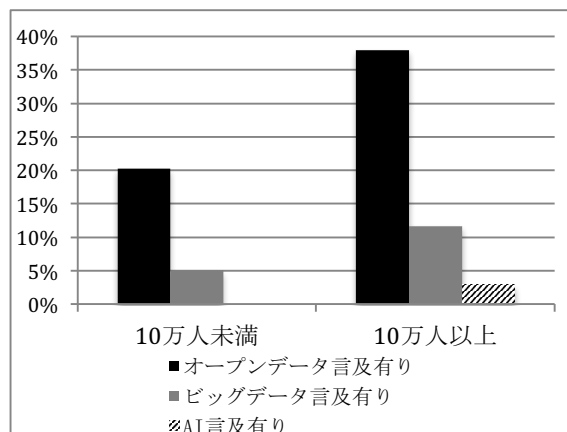


図7 人口別で言及数の登録数に占める割合

人口別で言及数の登録数に占める割合の差についても比率の同一性検定を行った。検定で用いたデータと検定結果は表7,表8で示している。オープンデータ、ビッグデータへの言及について、10万人未満のグループの言及数の登録数に占める割合と、10万人以上のグループの言及数の登録数に占める割合との差を検定した。結果として、オープンデータについてはP値が.016で有意な結果が得られたが、ビッグデータについてP値は.197で統計的に有意な差は認められなかった。

表7 人口別オープンデータへの言及及び検定結果

	DATAGOJP登録自治体数	オープンデータ言及有り	オープンデータ言及無い	言及数の登録数に占める割合(%)
10万人未満	79	16	63	20.3
10万人以上	103	39	64	37.9
検定結果	X-squared = 5.767, df = 1, p-value = .016			

表8 人口別ビッグデータへの言及及び検定結果

	DATAGOJP登録自治体数	ビッグデータ言及有り	ビッグデータ言及無い	言及数の登録数に占める割合(%)
10万人未満	79	4	75	5.1
10万人以上	103	12	91	11.7
検定結果	X-squared = 1.668, df = 1, p-value = .197			

#### (6)計画策定年度別

DATA.GO.JP登録自治体の総合計画に基づいて三つのキーワードの言及数の登録数に占める割合を確認すると、平成28年度までは新しく策定した計画ほど、

三つのキーワードについての言及が多い。他方、平成29年に策定された総合計画におけるオープンデータ及びビッグデータの言及数は前年度より減少している。図8に基づいて概観すると、オープンデータが平成26年から認識されるようになり、総合計画内で言及する自治体が増えたことが確認できる。ビッグデータに関しては、平成27年から言及数の登録数に占める割合は高まった。AIの言及件数は13件のみで、そのうち10件が平成29年に策定された計画である(図8)。傾向として新しく策定した計画ほど三つのキーワードの言及が多いことが確認できる。

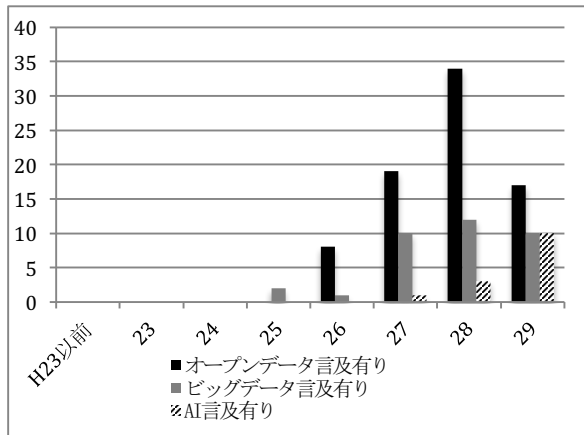


図8 計画策定年度別で言及数の登録数に占める割合

### (7)三つのキーワードの関連性

三つのキーワードの関連性を見るために、オープンデータとビッグデータへの言及の場合と(AI言及せず)、ビッグデータとAIへの言及の場合(オープンデータ言及せず)、オープンデータとAIへの言及の場合(ビッグデータ言及せず)の件数をそれぞれ確認した。オープンデータとビッグデータが共に言及されたのが最も多くて16件あった、AIについて言及が元々少ないため、ビッグデータとAI言及、オープンデータとAI言及がそれぞれ4件と1件しかなかった(図9)。

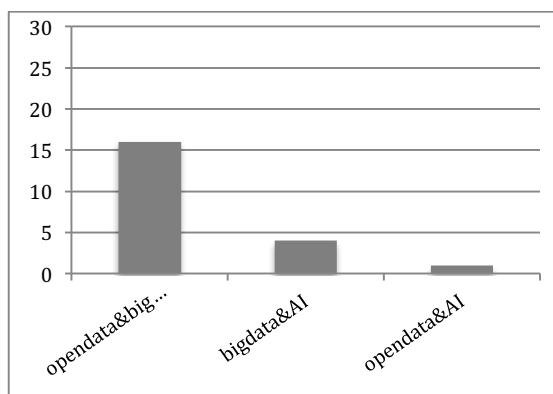


図9 二つのキーワードがともに言及された件数

## 6.まとめ

本調査ではDATA.GO.JPに登録している自治体の総合計画におけるオープンデータ、ビッグデータ及びAIの言及状況を確認することを通じて、オープンデータ、ビッグデータ、AIの自治体における推進状況を把握することを試みた。具体的には、登録自治体を種類別、ソフトウェア業従業者集中程度別、人口規模別でグループに分けて、オープンデータ、ビッグデータへの言及について、それぞれのキーワードに言及している自治体数に対するDATA.GO.JPへの登録自治体数に占める割合のグループ内の差を検定した。結果として、オープンデータの言及については、人口規模が10万人未満のグループより10万人以上のグループの方が言及数の登録数に占める割合が有意に高かった。ビッグデータの言及について、自治体種類別でその他の市区町村のグループより都道府県政令指定都市のグループの方が言及数の登録数に占める割合が有意に高かった。それ以外のグループ内の言及数の登録数に占める割合については有意な差が確認されなかった。オープンデータに関する言及で人口別グループ間とビッグデータに関する言及で自治体種類別間が統計的に有意な差が確認された要因や、オープンデータ、ビッグデータ及びAIの三つのキーワードが総合計画の中でどのように言及されているかについては、今後の研究課題である。また、多様なデータを活かした地域情報化と産業活性化の推進に関する研究も今後の研究課題としたい。

## 補注

<sup>1</sup> 地域経済動向における地域区分の変更について  
<http://www5.cao.go.jp/keizAI3/chiiki/osirase161128.pdf>

<sup>2</sup> 平成27年特定サービス産業実態調査  
<http://www.meti.go.jp/statistics/tyo/tokusabizi/resu1t-2/h27.html>