
原著論文

地方公共団体のオープンデータへの取組：統計データ 公開のあり方の検討

Open data initiatives in local governments: Study on statistical data disclosure

キーワード：

オープンデータ, 地方公共団体, 公共データ, ホームページ, 人口統計データ

keyword：

Open data, Local government, Public data, Homepage, demographic statistics

山口大学 中村 英人
Yamaguchi University Hideto NAKAMURA

山口大学 石野 洋子
Yamaguchi University Yoko ISHINO

要 約

わが国のオープンデータへの取組みは、2012年のIT総合戦略本部による「電子行政オープンデータ戦略」の決定などを通じて政府主導で始まり、地方自治体にも広まりつつある。当該自治体の規模の大きさがオープンデータ化の進展に影響していることは公表されているデータから容易に推定されるが、規模に内在する要因を定量的に解明した先行研究は、これまででない。この解明が本研究の第一の目的である。次に、従来から存在するホームページと新たなオープンデータの間、データ重複の問題を考える。それらのサイトに類似した紛らわしいデータが存在すると、様々な問題を生じさせる可能性がある。そこで、地方自治体のデータ公開の実態を調査し、そこに潜む課題を明確にすることが第二の目的である。その際、人口統計データに焦点を当てた。我々は、先進自治体へのインタビュー調査、オープンデータ実態調査（総務省）の人口規模による差異解析、そして、人口統計データの公開状況調査とい

原稿受付：2019年1月11日

掲載決定：2019年6月30日

う3段階の調査を経て、以下のことを発見した：(1) 地方自治体のオープンデータ化の推進には、自治体の規模に加え、担当部署とプロセスが大きく関係していること、(2)ほとんどの自治体が既存のホームページをそのまま維持した状態で、新たにオープンデータのサイトを追加する形を取っていること、(3) 両方のサイトでデータが内容的に重複しているにもかかわらず、それらのデータの同一・差異について明記していないところが多いこと。

Abstract

“Open data” efforts have now become a global movement. Japanese government has intended to keep up with this global trend since the Open Government Data Strategy was adopted in 2012. Some local governments have followed the government policy. However, there has not been any previous study that quantitatively clarified what factors would influence the implementation of “open data” in a local government. So, clarifying this is our first research goal. Next, we focus on the data duplication problem between data existing on their homepages and newly constructed “open data.” The presence of similar confusing data on both those sites may cause various problems. Therefore, our second research goal is to investigate the actual state of data released by local governments and then find the hidden problem. We focused on demographic data as one of the most important basic statistical data. We conducted three types of surveys; an interview survey with responsible persons in advanced local governments, an additional analysis of “open data” actual situation survey, which was performed by Ministry of Internal Affairs and Communications, using cross tabulation by the size of population in a municipality, and a survey of demographic data of municipalities on the Internet. Finally, we found that; (1) the department in charge of “open data” initiatives and their implementation processes are important, in addition to the size of a municipality; (2) most municipalities maintain the existing homepage as it is, and separately add a new “open data” site; and (3) despite the data duplication on both the existing homepage site and the “open data” site, many municipalities do not clearly express the data duplication in writing.

1 はじめに

我が国のオープンデータの取組みは、2012年に電子行政オープンデータ戦略が決定されたことにより、日本政府としての姿勢が明確化され（大向，2013），以降，その取組は政府主導で進められてきた。これまでの国の取組みを見ると，オープンデータの公開件数は，20,000データセットを超えている⁽¹⁾。また，カタログサイト⁽²⁾の開設やAPI機能の整備も進み，当初のデータの公開を中心とした取組から，データの活用を前提とした「課題解決型のオープンデータの推進」に発想を転換する方向（高度情報通信ネットワーク社会推進戦略本部，2017）にシフトしている。また，地方公共団体においても，福井県鯖江市や神奈川県横浜市などの先行的自治体（総務省，2013）のように，積極的な取り組みが行われてきている。

しかし，政府が公開している「オープンデータ取組自治体一覧（2017年12月20日時点）」⁽³⁾を用いて，地方公共団体のオープンデータ公開状況を調べると，都道府県では89.4%が何らかのオープンデータを公開しているのに対し，市区町村ではそれが15.2%にとどまっている⁽⁴⁾⁽⁵⁾。このことを鑑みると，地方公共団体でのオープンデータの導入や利活用は，完全に進んでいるとはいえない状況にある。

都道府県レベルでのオープンデータ取組率が高いのに対し，市区町村レベルでのそれが低いのは，なぜなのだろうか。地方自治体の規模が関係しているだろうことは容易に推測できるが，地方自治体の規模が，具体的にどうオープンデータ化推進に関与しているのか，これまで定量的に解明されてはいない。これを明らかにするのが本研究の第一の目的である。

ところで，オープンデータへの取組みとは別に，ホームページの開設は，都道府県，市区町村ともに全団体に実施済みであり（総務省，2016a），ホームページを活用した統計データの公開，ニュース

の告知，広報などは，地方公共団体において広く行われていた。地方公共団体におけるデータ公開という観点で現象を眺めると，従来からのホームページでの公開に，新たな取組みであるオープンデータでの公開が加わった状況になっている。

一般的に，オープンデータは，機械判読に適したデータ形式で公開され，著作権等の制約が無いという性質を持つ（この点については次章で詳しく述べる）のに対し，ホームページ上のデータは，それらの保証はない。そのため，オープンデータは従来のホームページ上のデータとは別のものだという認識が流布している。

しかし，既存のホームページ上のデータと，新たなオープンデータとについて，データそのものの重複，差異，位置づけ等について，実証研究を行って詳細に論じた研究は，これまでない。この問題は，データ提供者側の地方公共団体と利用者側の企業・住民との双方に，データ作成および活用の面で混乱を引き起こす可能性をはらんでいる。そこで，地方自治体において既存のホームページと新たなオープンデータとの公開実態がどのようになっているかを検証し，そこに潜む課題を明確にすることが，本研究の第二の目的である。

具体的には，地方自治体の提供するデータのうち最も基盤的で利用者の多い人口統計データに焦点を当てる。人口統計データには，国勢調査人口，推計人口，住民基本台帳人口，の3種類が存在し，数値が異なる。これらが混在したまま，マーケティングや事業企画の推定のために利用すると，正しい推定が行えない可能性がある。そこで，既にオープンデータに取り組んでいる自治体を対象に，既存のホームページとオープンデータとで人口統計データの公開実態を調査し，課題を明らかにする。

2 オープンデータとは

2.1 定義

高度情報通信ネットワーク社会推進戦略本部

(2017) は、オープンデータの定義として国、地方公共団体及び事業者が保有する官民データのうち、国民誰もがインターネット等を通じて容易に利用（加工、編集、再配布等）できるよう、次のいずれの項目にも該当する形で公開されたデータとしている。

- (i) 営利目的、非営利目的を問わず二次利用可能なルールが適用されたもの
- (ii) 機械判読に適したもの
- (iii) 無償で利用できるもの

また、機械判読段階には、Webの発明者でありLinked Dataの創始者でもあるティム・バーナーズ＝リーが提案している“5つ星スキーム⁽⁶⁾”が広く用いられている。

2.2 取組みの歴史と期待される効果

政府・自治体が保有するデータをオープンデータとして公開する必要性について、内閣官房情報通信技術（IT）総合戦略室（2017）は、政府、独立行政法人、地方公共団体等が保有する公共データについては国民共有の財産であるとの認識に立ち、オープンデータとして積極的に公開し、利活用を促進していくことが求められている、との考えを示している。

我が国に先行して取組みが始まった欧州では、国により公開するデータはさまざまだが、英国ではバスやレンタル自転車などのリアルタイム情報や環境情報が公開され、フランスでは人口、雇用などの統計データもポータルサイトに集約するなどの取り組みがある（高木，2013）。後発の我が国でオープンデータが必要とされる独自の状況としては、東日本大震災復旧・復興への取組と教訓（総務省，2013）や、渡辺（2013）が指摘する防災・減災関係のデータ公開や、それを活用した防災・減災をめぐる制度や技術、サービスの開発を挙げることができる。

こうしたオープンデータへの取り組みから期待される効果は、「行政に対する透明性・信頼の向上、

国民参加・官民協働の推進を通じた新たなサービスの提供・諸課題の解決、経済活性化・行政の高度化・効率化など」であると、高度情報通信ネットワーク社会推進戦略本部（2012）以来、政府のオープンデータ戦略の中で一貫して示されている。

一方、行政情報の提供手段としてのインターネットの活用は、政府が1994年12月に閣議決定した「行政情報化推進基本計画」で、その活用を掲げたことによる。今日、当たり前インターネットで閲覧している政府の情報は、当時の取り組みによって実現されたのである（矢杉ら，2014）。

2.3 先行研究

行政を対象としたオープンデータの先行研究には、先進事例を基にした特徴分析や課題抽出を行うものが多い。高木は、我が国に先行して取組みが始まった欧州におけるデータ公開の現状、政策的取組みの経緯、データのライセンス、民間とのコミュニケーション施策等の現状から、データを活用する民間利用者と政府とのコミュニケーション確立の重要性を指摘している（高木，2013）。庄司は、国・地方公共団体・民間の国内先進事例を踏まえ、法律的・制度的な課題を抽出し、政府データは原則すべて公開（オープン・バイ・デフォルト）を実現するために、政府・地方自治体の公務員の意識を変えていく必要性を指摘している（庄司，2013）。大向は、国内公共セクターのさまざまなレベルの取組みを紹介する中で、オープンデータは市民の請求に応じて供するのではなく、あらかじめWebサイトなどで公開がなされること等、既存の情報公開制度との違いを述べている（大向，2013）。

また、地方公共団体を対象にした研究には、以下のものがある。青木は、オープンデータの先進自治体について調査を行い、位置座標値の付与作業や高次なデータ形式への変換作業が必要となっていることを指摘し、このような作業は地方自治体にとって重荷であり、積極的なオープンデータ

公開は進まないであろうと指摘した(青木, 2013)。萩行は、アンケート調査結果から特徴分析を行い、自治体内部でオープンデータの認識について差が見られることを指摘し、情報化担当部門が積極的にオープンデータに取り組もうとしても、各種データの所管課の協力が得にくく、阻害要因となっている可能性があるとして推察している(萩行, 2014)。早田らは、政府・市民の協働段階においてデータを活用した地域サービスが立ち上がらない課題を指摘し、先進地域の詳細調査から、行政職員の参加とプロデュース人材が必要であることを示している(早田ら, 2015)。萩島らは、自治体の観光情報を対象にしたオープンデータ化からアプリ開発の実証実験の取組みを通じて、自治体の業務内で機械判読性を考慮することは困難であることや、自治体内部の部署間の連携強化の必要性を指摘している(萩島ら, 2016)。

これらの先行研究では、先進事例研究や実証実験などで得られた質的情報から、直接的にオープンデータ化への課題を指摘している。一方、本研究においては、事例の数は決して充分とはいえないが、地方公共団体のオープンデータ化への取組みの状況をできるだけ定量的に扱うことで、なるべく客観的に課題の抽出や検証を試みる。その際、既存のホームページ上のデータの重複に着目して解析を行うが、この点は従来研究にはないものである。

3 調査対象と手法

第1章で述べた2つの研究目的を達成するために、次のように三段階で進める：①まず、オープンデータに先行して取り組んでいる自治体へのインタビュー調査を行い、地方公共団体のオープンデータ導入の進め方と課題、既存のホームページとの関係性の捉え方を定性的に把握する。②次に、オープンデータに取り組む地方自治体へのアンケート調査を利用して、人口規模別の解析を行い、

地方自治体の規模の違いがオープンデータの取組みに与える影響を調べる。③最後に、(②で回答している)地方自治体がインターネット上に公開している人口統計データの種類を調査し、ホームページとオープンデータでの重複の実態を調査する。人口統計データを選択したのは、政策決定にとどまらず、学術、教育、民間など各方面で広く利用される重要なデータだからである。

人口統計データには、主なものとして、「国勢調査人口」、「推計人口」、「住民基本台帳人口」の3種類がある。このうち、「国勢調査人口」は、国勢調査で得られる人口値である。国勢調査は、我が国に住んでいるすべての人を対象とする国の最も基本的な調査で、国内の人口や世帯の実態を明らかにするため、5年ごとに実施される。都道府県・市区町村の全団体が国勢調査に携わり、自治体別に結果が集計される。ただし、5年ごとの実施のため、毎月、毎年的人口は国勢調査ではわからない。このため、直近の国勢調査人口を基準として、その後の人口動向を他の人口関連資料から得て作成されるものが「推計人口」である。一方、「住民基本台帳人口」は、住民基本台帳に基づき、整理・集計された人口で、月々の都道府県、大都市間の転入・転出の状況を明らかにするものとして作成されている。

これら3種類の人口値には差がある。例えば、広島市の2016年10月1日時点で各直近のデータをみると、国勢調査人口は1,194,034人、推計人口は1,196,380人、住民基本台帳人口は1,192,975人である。このような違いを生むのは、国勢調査人口に5年のインターバルがあることに加え、転入・転出等を正確に記録することが難しいからである。例えば、住所変更をせずに転居すると、住民票の届出場所と実際に住んでいる場所が一致しない等が挙げられる。このようなことから、企業や個人が人口を基に事業計画等を立てる場合に、年度ごとに異なる種類の人口値を用いてしまうと、正しい推定が行えない可能性がある。そこで、

既にオープンデータに取り組んでいる自治体を対象に、既存のホームページとオープンデータとでこれら3種類の人口統計データの公開実態を調査し、課題を明らかにする。

3.1 先進自治体へのインタビュー調査

2015年の段階で、他の自治体に先んじてオープンデータ化に取り組む6つの自治体（県：3団体、政令指定都市：1団体、市：2団体）を選択し、オープンデータの導入実務を担当してきた人物にインタビュー調査を実施した。調査対象は、総務省の情報通信白書などで特徴的な取り組みが紹介された自治体や、都道府県内において最初にオープンデータに取り組んだ実績のある自治体から選定した。県の1団体は電話調査であったが、それ以外には直接訪問し面談調査を実施した。インタビュー内容は、オープンデータ公開の仕組みの構築と既存のホームページとの関係を中心に、公開データを準備するための自治体内部の調整や、導入後の活用事例等について広く尋ねた。実施時期は、2015年9月である。なお、調査時点では、オープンデータは黎明期にあり、自治体がどのようなオープンデータの取り組みを行っているのかを、質的に調べることが主目的であった。この時点では特徴的な取り組みを行っている自治体は県レベルが多かったため、調査対象の半数を県が占めている。この2015年の質的な調査で得た知見が、以降の量的調査の解析に活かされている。

3.2 オープンデータ実態調査の人口規模別解析

総務省（2016a）によって、「オープンデータに取り組んでいる団体」とみなされた61の地方公共団体に対して実施されたアンケート「オープンデータに取り組む団体の実態調査（2016年3月結果公表：総務省、2016b）」（以下「総務省アンケート調査」という）を利用した。61の地方自治体を人口規模別に3グループに分け、アンケート項目ごとにクロス集計を行い、分布の違い

を調べた。

3.3 ネットに公開されている人口統計データ調査

総務省アンケート調査には、既存のホームページサイトとオープンデータ掲載サイトとを比較調査するような項目は設定されていなかったため、同61地方公共団体のインターネット上のサイトの、人口に係る統計データを新たに調査・分析した。実施時期は、2017年3月である。

4 先進自治体へのインタビュー調査結果

インタビューで得られた、先進自治体のオープンデータへの主たる取組及び既存のホームページでのデータ公開の状況は、表1のとおりである。

利活用促進の取り組みは、アプリコンテストやアイデアソン・ハッカソンなどのイベント開催が中心に行われていた。こうしたイベントは、従来のホームページによるデータ公開の段階では行われていなかった取り組みであり、オープンデータ公開の取り組みの中でみられる一つの特徴となっていた。

オープンデータの公開の仕組みについては、既存のホームページとは別に新たにカタログサイトの構築や、既存のホームページにオープンデータ専用のページを設定するなどの対応がなされていた。このうち、C県とF市は、オープンデータ公開の仕組みとして、一般社団法人リンクデータが運用するオープンデータ公開のプラットフォームのLinkData⁽⁷⁾を活用していた。

LinkDataをオープンデータの公開に活用する利点として、Excelファイルをアップロードすることで容易にオープンデータ化できる点、オープンデータ化する際にRDF形式など様々な形式に自動変換される点などが挙げられる（荻島ら、2016）。

一方、インタビュー調査の中で明らかになった実態として、LinkDataにアップロードするExcel

表1 先進自治体へのインタビュー調査結果

調査自治体	人口規模等	オープンデータ公開システム	ホームページでのデータ公開	認識する成功事例	利活用促進の取組み	評価方法
A市（政令市）	50万人以上	カタログサイト構築	有り	アプリ作成（内部作成）	・他都市及び産学官連携によるアプリコンテスト等 ・国の実証実験への参加	活用事例数
B県	50万人以上	カタログサイト構築	有り	・職員の意識改革 ・都道府県初の取り組みが、マスコミ等で多数紹介	・アーバンデータチャレンジへの参加	アクセス数・公開件数
C県	50万人以上	Webサイトに専用ページ設定	有り	アプリ作成（外部作成）	・アプリコンテスト開催 ・オープンデータをテーマにしたフォーラムを開催	なし
D県	50万人以上	カタログサイト構築	有り	アプリ作成（外部作成）	・アイデアソン・ハッカソンの開催	アクセス数、公開件数、アプリ開発件数
E市	10～50万人	カタログサイト構築（地図に特化）	有り	なし	・地域住民への説明会開催 ・アーバンデータチャレンジへの参加	なし
F市	10～50万人	外部カタログサイト利用	有り	アプリ作成（外部作成）	アプリコンテスト開催	データ公開件数、アプリ作成数

ファイルは、定められたフォーマットに必要な情報等を入力していく必要があるが、両自治体ともに、このデータ整形作業については、オープンデータの担当部門（情報部門）が行っていた。データ内容の責任を持つデータ主管課から元データを受領し、オープンデータを担当する情報部門でデータを整形したものを、オープンデータとして公開するフローである。このようなフローを採用している理由は、データ主管課に対するオープンデータ対応の事務負担を減らすことで、データ提供の協力を得やすくするため、とのことであった。なお、公開データを紙ベースやPDF形式でデータ主管課から受け取り、それをオープンデータ担当部門が手作業で、ExcelやCSV形式にデータ整形しているケースもあった。

また、オープンデータとして公開するデータ数を増やすため、データを保有する関係各課への個々の協力依頼は、いずれの自治体でも行われていた。このような、データ保有部門へのオープンデータ用のデータ提供の協力を得るための内部調整や、データ作成対応に、各自治体ともに多くの労力を投入していた。

一方で、オープンデータの多くは、従来からホームページで公開されているデータを元にしていられるが、いずれの自治体においても、本稿の目的である「既存のホームページサイトと、オープンデータサイトの、公開データの関係整理」などの検討は、オープンデータの取り組みの中では行われていない状況であった。

5 実態調査の人口規模別解析結果

続いて、総務省アンケート調査の結果を用いて、自治体の人口規模の差異により、オープンデータの公開状況にどのような差が生じるのか（生じないのか）、定量的に検証を行った。本調査における人口規模別の調査対象自治体数は、表2のとおりである。このうち、都道府県は1自治体のみであり、残りの60自治体は、市区町である。

なお、自治体の人口規模により、オープンデータの取り組みには差があると仮説を立てたため、人口規模別に自治体を分け、それを軸としたクロス集計を中心に検証を行った。本来、都道府県、市、町村、それぞれで役割や権限が異なり、公開

すべきデータや関心事に違いがあることが想定されるので、県レベル、市レベルのように分けて分析をする方法もあろう。しかし、総務省アンケート調査には都道府県が1つしか含まれず、また、「政令指定都市（50万人以上）、中核市（20万人以上）、市（5万人以上）、町村⁽⁸⁾」のように、自治体は概ね人口規模で区別されることから、本研究においては、人口区分での切り口で分析を行うこととした。

人口規模が10万人未満の自治体群を「人口規模が小の階層」、同じく10～50万人の自治体群を「人口規模が中の階層」、50万人以上の自治体群を「人口規模が大の階層」と、それぞれ呼ぶこととする。

表2 調査対象自治体数

No.	人口規模	調査対象自治体数	構成比(%)
1	～10万人	18	29.5
2	10～50万人	27	44.3
3	50万人以上	16	26.2
	合計	61	100

5.1 重点的に推進するオープンデータの公開分野

「重点的・積極的に推進するオープンデータの公開分野」を調べた結果は図1のとおりである。全体で見ると、「防災・減災」が68.8%と最も高く、2位が「統計・人口動態」50.0%、3位が「観光」37.5%であった。人口規模が小と中の階層では、上位3つの項目は変わらない（順位は異なる）が、人口規模が大の階層では、「観光」がトップ3から外れ、他の分野の比重が上がっていた。この結果から、以下のことがわかった。

- ・「防災・減災」と「統計・人口動態」は、人口規模に関係なく、重点的にオープンデータ化が推進される分野である。
- ・人口規模が小～中の自治体にとって「観光」も重点的にオープンデータ化し、注力したい分野

である。

おそらく、人口規模が小～中の自治体は観光を主な産業としているところが多いのに対し、人口規模が大きい自治体は、観光以外の産業も盛んなため、このような結果になったと考えられる。

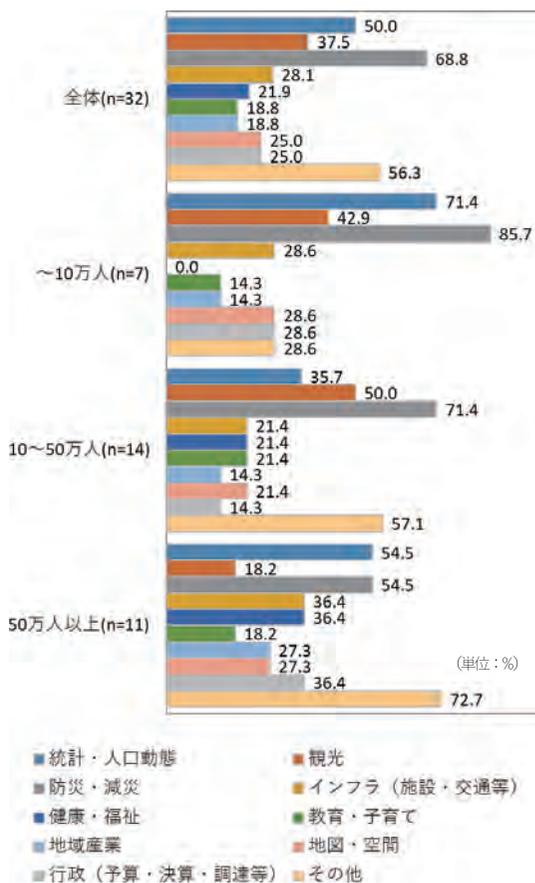


図1 重点的・積極的に推進する公開分野 (MA)

(出典：総務省アンケート調査を再集計)
注：未回答の自治体があるため、全体のn数は61にならない

5.2 オープンデータの作成方法

続いて、「オープンデータの作成方法」がどのようになっているかを調べた結果を図2に示す。全体では、「オープンデータ担当部署で作成」が56.1%と最も多く、次いで「データ主管課で作成」が52.6%となっている。人口規模別でみると、人口規模が小の階層では、「オープンデータ担当

部署で作成」が多いが、人口規模が大きくなるにつれて、「データ主管課で作成」の割合が高まる結果となっている。また、人口規模が大の階層では、データ主管課での作成に次いで、保有データをそのまま公開という対応も、高い割合を示す結果となった。

この結果からは、人口規模が大きい自治体では、オープンデータの作成作業がデータ主管課の業務に加わっていることがみてとれるが、人口規模が小さい自治体では、そうではなく、オープンデータ担当部署の負担が大きかった。

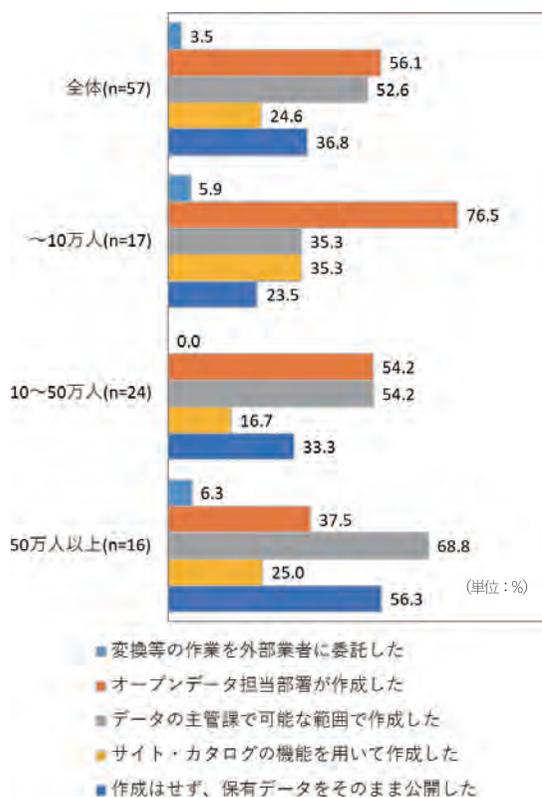


図2 オープンデータの作成方法 (MA)

(出典：総務省アンケート調査を再集計)
注：未回答の自治体があるため、全体のn数は61にならない

5.3 オープンデータの中心推進部署

次に、「オープンデータの中心推進部署」がどのようなになっているかを調べた結果を図3に示

す。自治体全体でみると、中心推進部署は80.3%が情報関連部署であった。また、いずれの階層をみても、圧倒的な割合で情報関連部署であることがわかる。この結果から、人口規模に関係なく、ほとんどの自治体で、情報関連部署がデータ公開の仕組みの構築だけでなく、公開データを提供してもらおうための内部調整実務なども、広く担当していると考えられる。

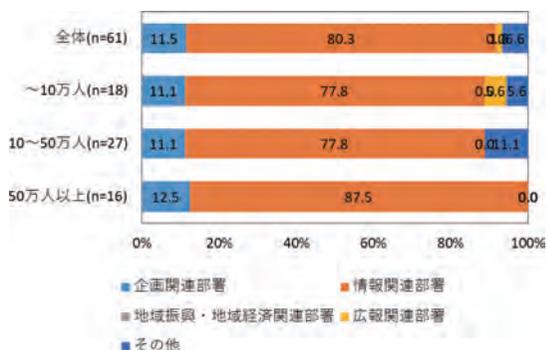


図3 オープンデータの中心推進部署 (SA)

(出典：総務省アンケート調査を再集計)

5.4 オープンデータの公開方法

最後に、「オープンデータの公開方法」がどのようなになっているかを調べた結果を図4に示す。

人口規模が中～大の階層は、既存のホームページにオープンデータの専用ページを設置している割合が高かった。自治体全体でも、57.4%がこの方法でデータの公開を行っており、オープンデータ化への取組みについては、「既存のホームページに専用ページを追加する形でオープンデータを公開する」というスタイルが主流となっていることがわかった。

しかし、人口規模が小の階層では、「外部のサイト・カタログサイトを利用している」が増加し、「既存のホームページにオープンデータの専用ページを設置する」と同率で最も多い。これは、4章の先進自治体へのインタビュー調査結果でも述べたようにLinkDataなどのオープンデータ公開の外部のプラットフォームを活用していると考え

えられる。人口規模の小さい自治体にとっては、オープンデータのための仕組みを新たに自前で構築することは、予算的にも人的にも厳しいことが予想される。そこで、こうした費用がかからない外部のサービスを活用し、そこにデータをアップロードする公開方法を選択していると考えられる。

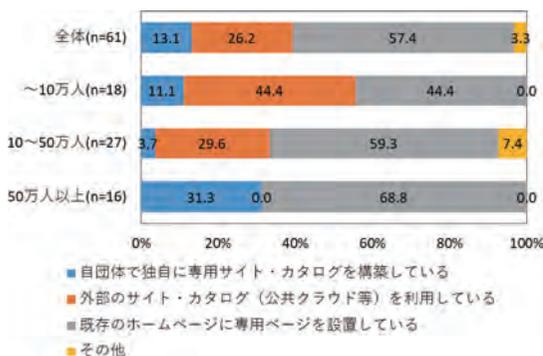


図4 オープンデータの公開方法 (SA)

(出典：総務省アンケート調査を再集計)

6 ネットに公開されている人口統計データ調査

総務省のアンケート調査だけでは、既存のホームページ上のデータと新たに追加したオープンデータとの重複関係がどのようになっているのかわからないため、インターネット上で、61自治体の人口統計データがどのような形で公開されているか、調査を実施した。

6.1 インターネット上の人口統計データ種別

インターネット上で公開されている人口統計データの種別を図5に示す。

図5からわかるように、既存のホームページサイト、オープンデータサイト、いずれの場合も「住民基本台帳人口」が最も多く、次いで「国勢調査人口」「推計人口」の順であった。ここで注目すべきは、掲載データの出典（種別）が明示されていないデータ、すなわち「統計種別が不明な人口統計データ」が、既存のホームページのサイトでは19.7%であるのに対し、オープンデータサイ

トでは34.4%と、大幅に増加していたことである。

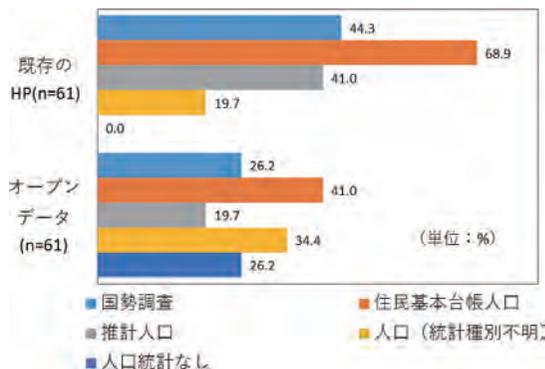


図5 インターネット上の人口データの状況 (MA)

(出典：筆者調査結果)

この理由を確認するために、既存のホームページおよびオープンデータサイトのそれぞれにおける、統計種別の明示・不明の状況を分析した。なお、公開されている人口統計データのうち、一つでも種別が明示されていないものがある自治体は、統計種別不明として集計した。

まず、既存のホームページにおける統計種別の明示・不明の状況を、自治体の人口階層別に調べた結果を図6に示す。自治体全体でみた場合、統計種別不明の人口統計データをホームページサイトに有する自治体は19.7%であったが、人口規模が小の階層をみると、約40%の自治体において統計種別の不明のまま、人口統計データが公開

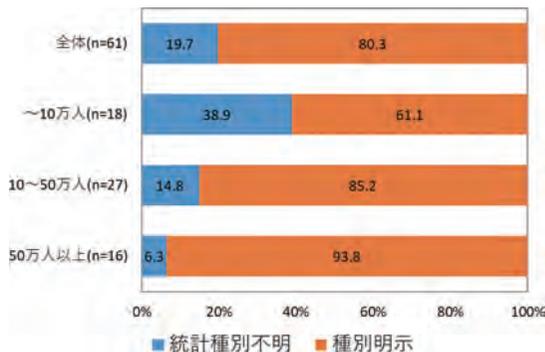


図6 ホームページの統計種別不明の状況 (SA)

(出典：筆者調査結果)

されていた。特徴として、人口規模が大きくなるほどに、統計種別不明の自治体の割合は減少していることがわかった。

次に、オープンデータサイトにおける統計種別の明示・不明の状況を、自治体の人口階層別に調べた結果を図7に示す。自治体全体でみた場合、統計種別不明の人口統計データをオープンデータサイトに有する自治体は34.4%で、ホームページサイトの場合の1.7倍以上であった。自治体の人口規模ごとに結果をみると、全ての階層で同様のことが観察された。つまり、オープンデータサイトに統計種別不明の人口統計データを掲載している自治体の割合は、ホームページサイトに同様のデータを掲載している自治体の割合を大きく上回っていた。人口規模が小さいほど、この傾向は顕著であり、人口規模小の階層にいたっては、半数を超える55.6%の自治体が、統計種別不明の人口統計データをオープンデータサイトに掲載していた。

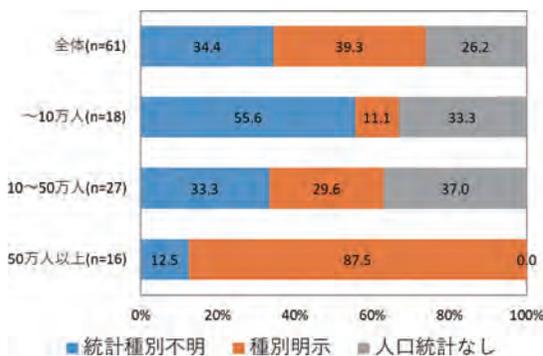


図7 オープンデータの統計種別不明の状況 (SA)
(出典：筆者調査結果)

6.2 統計種別の明示とデータ作成方法

続いて、総務省アンケート調査の「オープンデータの作成方法 (図2)」項目とのクロス集計を行って、オープンデータの統計種別を「明示している自治体」と「明示していない自治体」とで、データ作成方法の違いを調べた(図8)。その結果、オープンデータの統計種別を「明示していない自治体」

の最も多いデータ作成方法は、「オープンデータ担当部署 (85.0%)」による作成であった。一方、「明示している自治体」の最も多いデータ作成方法は、「データ主管課 (68.0%)」による作成であった。

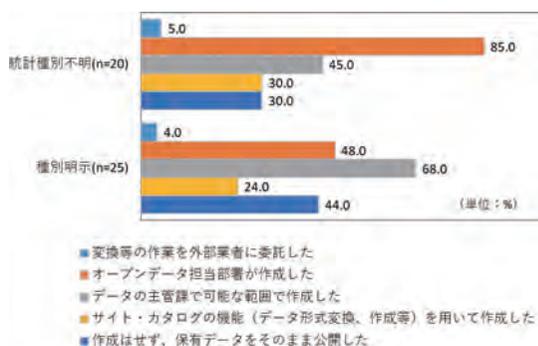


図8 統計種別の不明・明示の別のオープンデータの作成方法 (MA)

(出典：総務省アンケート調査及び筆者調査結果)

このことは、図3で示した「人口規模に関係なく、オープンデータ化は情報関連部署を中心に進められている」ということと大きく関連していると考えられる。統計主管課では、人口統計データを扱ってきた長年の経験から、「人口統計データには3つの種別があり、目的に応じて異なる人口統計データを活用する必要がある」という知識が部門内に浸透していると考えられるが、情報部門では事情が異なる。したがって、情報部門が中心となってオープンデータ化を推進する場合、データを統計部門から受領した後に、加工してオープンデータサイトにアップロードするのは情報部門の役割となり、その過程で「人口統計データの種別を明示する意義」が薄れてしまい「表示が消失する」という状況が考えられる。

今回は、人口統計データを例に検証を行ったが、原データ作成に携わらない情報部門が、オープンデータに加工・公開する段階で、その他のデータにおいても、同様のことが起きている可能性がある。利用者側からすると、データの根拠を一つ一つ検証しないと、そのデータを利用できない状況

であり、オープンデータの利活用において、大きな障害となると考えられる。

6.3 既存ホームページサイトとオープンデータサイトのデータ重複の状況

次に、既存のホームページ上の人口統計データとオープンデータサイト上の人口統計データとの重複（一致・不一致）の状況を調べた。両方のサイトにおいて、人口の種別（国勢調査人口、推計人口、住民基本台帳人口）の組み合わせが同じであれば「同種類」とし、異なる場合は「異なる」とした。また、片方のサイトにもう一方のサイトからリンクが貼られており、結果、同一のページが表示される場合は「同一」とした。なお、人口統計データはすべての調査対象自治体の既存のホームページに掲載されていたが、オープンデータサイトに人口統計データの掲載がない場合は「HPのみ」とした。その結果を集計したのが、図9である。

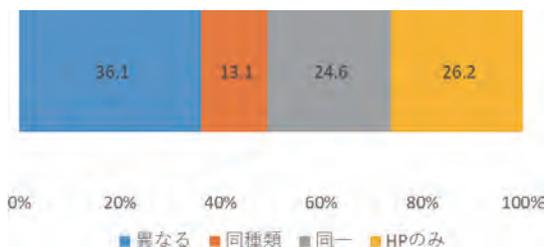


図9 掲載人口データの一致・不一致の状況

(出典：筆者調査結果)

この結果を見ると、人口統計データがリンク等で結ばれた同一の状態となっている自治体は24.6%にすぎない。それに対し、異なる種類の人口統計データを掲載している自治体が36.1%、同種の人口統計データを掲載しているがリンクを張っていない自治体が13.1%存在し、両方で約半数に達する。これらのケースでは、両サイトのデータについて、利用者がデータの出典をその都度確認する必要が生じる。特に問題となるのが、

データ取得の窓口が、既存のホームページサイトとオープンデータサイトの2箇所に存在するケースである。この場合、両方のデータの関係が非常に分かりにくく、利用者は自分で掲載内容を比較しながら目的のデータを探さなければならない（もしくは差に気づかず誤ったデータを使用してしまう）。

例えば、国勢調査の例を示すと、統計の集計結果⁹⁾は、男女、年齢、配偶者関係、世帯状況、住居状況、高齢状況、外国人、労働力等の別で細かく別れており、そこには100以上の集計種別が存在する。また、この集計表の区分に、5年毎の調査年次の区分を組み合わせると、その種類は膨大なものとなる。そうした中で、既存のホームページサイトとオープンデータサイトの窓口が2箇所存在し、両サイトの掲載データの対応関係を明示する情報がない場合、利用者の混乱を招く。両方のサイトの存在を明確に意識していない場合、場当たり的にダウンロードすることを繰り返せば、異なる種類の人口統計データで時系列変化を追ってしまうことにもなりうる。こうした状況は、データの利活用を阻む要因となる。

一方で、取組みの方向性の参考となる事例もあった。通常、既存のホームページとオープンデータサイトに同一の人口統計データを掲載する場合、多くの自治体はオープンデータサイトから該当する既存のホームページへリンクを張ることで、その同一性を確保していた。このリンクを張る手法は、簡易にデータの同一性を実現できるという利点があるが、既存のホームページ公開の仕組みを修正してURLが変更された場合、オープンデータサイトのリンクをすべて修正しなければならないという欠点がある。リンクの修正には手間がかかるうえ、リンク関係に常に注意を払う必要があり、修正に失敗するとリンク切れのまま放置することになりかねない。それに対し、北海道森町は違う方法を取っていた。森町は、LinkDataにて人口のオープンデータを公開しており、同町

のホームページ⁽¹⁰⁾の人口統計データは、このオープンデータから自動生成する仕組みになっている。この方法だと、ホームページの公開仕様変更に伴うリンク修正の必要が無く、リンク切れの恐れもない。この「オープンデータサイトで一元的にデータ管理を行い、通常のホームページサイトにはオープンデータサイトから自動生成したデータを掲載する」という方法は、データ管理の側面からもデータ利用の側面からも大きな利点があり、オープンデータ化推進の一つの方向性となるだろう。

7 本研究の成果と考察

以上の検討から、地方公共団体のオープンデータ推進に関して、大きく2つのことが判明した。ひとつは、地方公共団体のオープンデータ推進には、当該団体の規模が大きく関係しているということである。自治体の規模を、その内在する人口規模で表現すると、人口規模が中～小の自治体の場合、オープンデータ推進のためには、担当部署（情報関連部署）がデータ公開の仕組みの決定だけでなく、オープンデータの作成作業や、関連部署から公開データを提供してもらうための内部調整実務まで、多岐にわたる業務をこなす必要がある。そのうえ、人口規模が小さいほど、資金面や人材面での不足が生じやすい。これらのことが、市町村レベルでオープンデータ化が思うように進まない大きな要因だと考えられる。これは、インタビュー調査と、総務省アンケート調査の人口規模別分析から判明した。

もうひとつは、従来からのホームページサイトと、新たなオープンデータサイトに「データの重複に伴う問題」が存在することである。このことについては、これまでに明示的に調査された例はなく、本研究の大きな成果といえる。

本研究では、人口統計データに着目し、総務省アンケート調査で選ばれている61団体すべての

ホームページを調査した。その結果、ほとんどの自治体では、既存のホームページをそのまま維持した状態で、新たにオープンデータのサイトを追加で公開していた。この段階で、同様なデータが二重に公開されていることになる。そのうえ、既存のホームページサイトの人口データと、新たなオープンデータサイトの人口データとの2つの間に、リンクを張り人口統計データの同一性を確保している自治体は全体の4分の1に過ぎなかった。全体の半分の自治体は、人口統計データの種類の異なるものを掲載するか、同種類だがリンクがないために同一性が担保されない状態であった。

そもそも、なぜ既存のホームページをそのまま維持した状態で、新たにオープンデータサイトを追加公開するかといえば、内閣官房情報通信技術（IT）総合戦略室（2015）や内閣官房情報通信技術（IT）総合戦略室（2017）において「既にインターネットを通じて公開しているデータに二次利用可能な利用ルールを適用する」というオープンデータの整備順序が例示され、既存のデータ資産をまず活用することが推奨されたからだといえる。

この手法は、オープンデータの公開データ数を増やすことには繋がりやすいが、データ取得の窓口が二重になり、利用者の負担を強いることになっている。今回、いずれの自治体においても、両サイトの掲載データを比較できるような案内は見当たらず、二重のデータに気づかない利用者も多いことが推測される。このことは、利用者である市民や企業の不利益につながりかねない。

そこで、既存のホームページ上のデータとオープンデータサイト上にある（もしくは予定）のデータ、各々の内容を精査し、その整理統合をすることが必要になる。そのためには、自治体内部の部署間連携を進める必要がある。

しかし、現状は、人口統計データなどの基礎データを作成管理するデータ主管課とは異なる情報関連部署がオープンデータの中心的な推進部署と

なっていて、両者の協力体制は十分とはいえない。これは、従来のホームページサイトで種別不明の人口統計データを掲載している自治体は全体の約2割程度なのに対して、オープンデータサイトで種別不明の人口統計データを掲載している自治体は、その1.7倍に増加していたことから推察される。

財政状況が厳しい中、限られた予算内でオープンデータ化を今後も導入・推進しなければならない地方自治体は、この事実を重く見て、組織のあり方、データ公開の方法を再考する必要がある。なお、効率的にデータの同一性を担保する手法については、「オープンデータサイトで一元的にデータ管理を行い、通常のホームページサイトにはオープンデータサイトから自動生成したデータを掲載する」という森町が実施している方法が参考になるだろう。

また、地方公共団体のオープンデータ化を推進する我が国の政府においては、従来の、既存データ資産を維持したままでのオープンデータ公開の推奨を見直し、データを公開する側の責務として、類似データの二重公開や出典の不明示などが起こらぬよう指示していく必要がある。これにより、精査された有用なデータがオープン化され、データ公開手続きも効率化されることが望まれる。

8 おわりに

本研究では、複数の調査を重ねることで、従来のホームページサイトと新規のオープンデータサイトで、並行して行われているデータ公開の現状を明らかにすることができた。具体的には、先進的な6自治体へのインタビュー調査を行い、オープンデータ導入の現状と既存のホームページとの関係性を把握した。続いて、総務省が実施したアンケート調査結果に基づき、61団体のホームページ上の人口統計データを調査することで、現状を定量的に確認した。

本研究の大きな成果としては、取組みが進んでいるとみなされている61団体のほとんどが、既存のホームページをそのまま維持した状態で、新たにオープンデータのサイトを追加で公開しているという事実を明らかにし、両方のサイトでデータが内容的に重複しているにもかかわらず、それらのデータの同一・差異について明記していないところが多いということを発見したことである。この現象は、これまで政府により推奨されてきた「できることからオープンデータを公開する」という方針により生じていると考えられるが、このやり方は、公開データの数を増やすことには繋がりがやすい一方で、データ提供窓口の二重化によるデータ取得の煩雑性という欠点をもたらし、利活用の阻害要因となっていた。

今後は、地方自治体においてデータ活用や地域活性化を促すためには、対象データの選択や公開手段をどうするのが望ましいか、オープンデータサイトとホームページサイトの連携問題も含めて、明らかにしていく研究を実施したいと考えている。

注

- (1) 総務省行政管理局が運用する国のオープンデータに係る情報ポータルサイト
<<http://www.data.go.jp/data/dataset>>
Accessed 2018, March 3.
- (2) データカタログサイトとは、データの一覧表示や横断的検索などの機能を備えたポータルサイトのことである。オープンソースソフトウェアのCKAN (<http://ckan.org/>) などを利用したカタログサイト（データセットのカテゴリ毎にタグが付けられていたり、グループ化されていたり、フリーワードで検索できるようになっている）が開設されている。
- (3) 政府CIOポータル
<<https://cio.go.jp/policy-opendata>>

- Accessed 2018, March 3.
- (4) 総務省「市町村数の変遷」の2016年10月10日時点の市町村数1,718に特別区（東京23区）を加えた1,741を市区町村数，都道府県数は47としてオープンデータの取り組みの割合を算出。
<<http://www.soumu.go.jp/kouiki/kouiki.html>> Accessed 2018, March 3
- (5) その後，都道府県においては，2018年3月で47都道府県の全てが，オープンデータに取り組むこととなった。
- (6) 5★ オープンデータ
<<http://5stardata.info/ja/>> Accessed 2018, March 3
- (7) <<http://ja.linkdata.org/>> Accessed 2018, March 3.
- (8) 地方公共団体の種類について
<http://www.soumu.go.jp/main_sosiki/jichi_gyousei/bunken/chihou-koukyoudantai_kubun.html> Accessed 2018, March 3
- (9) 2015年国勢調査
<http://www.stat.go.jp/data/kouhyou/e-stat_kokusei2015.xml> Accessed 2018, March 3
- (10) 北海道森町ホームページ
<<http://www.town.hokkaido-mori.lg.jp/od/zinko/>> Accessed 2018, March 3
- による分析から」、『地域活性研究』5，pp.181-189.
- 高度情報通信ネットワーク社会推進戦略本部（2012）「電子行政オープンデータ戦略」
- 高度情報通信ネットワーク社会推進戦略本部（2017）「オープンデータ基本指針」
- 内閣官房情報通信技術（IT）総合戦略室（2015）「オープンデータをはじめよう～地方公共団体のための最初の手引書～」
- 内閣官房情報通信技術（IT）総合戦略室（2017）「地方公共団体オープンデータ推進ガイドライン 平成29年12月22日改訂」
- 萩島和真，福安真奈，浦田真由，遠藤守，安田孝美（2016）「観光イベント情報を活用したオープンデータ化の試行と実践」、『社会情報学』4（2），pp.1-16.
- 大向一輝（2013）「日本におけるオープンデータの進展と展望」、『情報管理』56（7），pp.440-447.
- 早田吉伸，前野隆司，保井俊之（2015）「オープンデータ推進に向けた国内先進地域の特徴分析」、『地域活性研究』6，pp.61-70.
- 総務省（2013）「平成25年版情報通信白書」
- 総務省（2014）「統計Today No.87国勢調査と住民登録」
<<http://www.stat.go.jp/info/today/087.html>> Accessed 2018, March 3.
- 総務省（2016a）「地方自治情報管理概要～電子自治体の推進状況（平成27年4月1日現在）～」
- 総務省（2016b）「地方公共団体におけるオープンデータの具体的な取組等に関する調査研究」
- 総務省統計局「統計データ，分野別一覧，人口・世帯に関する統計」
<<http://www.stat.go.jp/data/guide/1.html>> Accessed 2018, March 3.
- 総務省統計局「平成27年国勢調査に関するQ&A（回答）」
<<http://www.stat.go.jp/data/kokusei/qa-1>

参考文献

- 青木和人（2013）「地方自治体におけるオープンデータ公開の現状と課題：自治体オープンデータ項目一覧表からの考察（III-4 電子政府／電子自治体，セッションIII，自由論題報告）」、『社会情報学会（SSI）学会大会研究発表論文集』2013，pp.211-216.
- 萩行さとみ（2014）「自治体におけるオープンデータの現状と課題に関する一考察：アンケート結果

- html> Accessed 2018, March 3.
- 庄司昌彦 (2013) 「オープンデータ活用：8. 国内における活用環境整備」, 『情報処理』54(12), pp.1244-1247.
- 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』55(10), pp.746-753.
- 渡辺智暁 (2013) 「欧州から考える政府のオープンデータ国際戦略」, 『智場 #119 特集号 オープンデータ』, pp.64-76.
- 矢杉直也, 劉長鈺, 西本秀樹 (2014) 「第6章 e-Japan計画と我が国の電子政府展開」, 『地方政府の効率性と電子政府』日本経済評論社, pp.149-173.