

地方自治体におけるオープンデータの利用 －茅ヶ崎市を事例として

Open Data in Japanese Local Government : the Case of Chigasaki City

大貫徹也¹, 井上幸次², 岡野雅雄³
Tetsuya Onuki Koji Inoue Masao Okano

Abstract

This study examines the open data published by local governments with particular focus on Chigasaki City. Using the 5-star deployment scheme for open data, data obtained from Chigasaki City are evaluated as level 2 or 3, because they are published in Excel or unstructured CSV formats on the website. The Japanese government's policy of promoting open data recommends that data evaluation should advance to levels 4 and 5, which would not only make the data accessible through the web, but also allow their sharing and linking on the web. After reviewing some attempts made by local governments that tried to achieve the advanced levels using Resource Description Framework (RDF) and Linked Open Data (LOD) techniques, we proposed a set of measures using which Chigasaki City would be able to prepare for the imminent prevalence of open data.

1. はじめに

昨今、データの活用は様々な業界において重要な課題となってきている。特に、ビッグデータとして、データを貯め、抽出する技術が一般化したことで、今まで活用できていなかったデータに光を当て、より多くの情報を分析、活用することが出来るようになってきた。

過去には、厳選されたデータに統計処理を施して分析することに重点が置かれてきたが、現在では、多様な大量データそのものからこれまでになかったような新しい発見が可能となり、データ自体に価値を見いだせるようになったことで、データを公開するオープンデータの取り組みも加速している。なお、『オープンデータハンドブック』⁴ (Open Knowledge International, 2017) によるオープンデータの

定義は、「オープンデータとは、自由に使えて再利用もでき、かつ誰でも再配布できるようなデータのこと」である。

このような流れの中で、「総務省：オープンデータ戦略の推進」⁵にあるように、政府も自治体のデータをオープンデータとして活用することを推進している。

本稿では、オープンデータに関する動向を簡単に展望した後、地方自治体におけるオープンデータの活用に関し、特に茅ヶ崎市を例として、下の点を中心に調査・検討を行った結果を報告する。

- ・日本におけるオープンデータの取組について、行政のデータを中心に調査を行い、茅ヶ崎市への反映方法を検討する。
- ・上記を踏まえ、茅ヶ崎市におけるオープンデー

¹ フリーゲーフォ合同会社

² バベッジラボ

³ 文教大学情報学部教授

⁴ <http://opendatahandbook.org/>

⁵ http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/opendata/

タの現状を把握し、より活用していくための検討をする。

2. 自治体のオープンデータの必要性

上述のように、データ自体が大きな価値を持つようになり、各種の事業を進める上でのデータ活用がより重要になってきている。地域経済、ひいては日本経済の観点から見てもデータ活用は最重要項目の一つとなっている。

国内の地域ごとの情報など、自治体が持っている情報（人口分布、人口推移、イベント情報、その他）は、使い方によっては有用な情報となる。これら自治体の持っている情報を一般公開し、民間で活用できるようにすることは、重要な取り組みとなってきている。

日本政府は、「オープンデータ基本指針」（平成 29 年 5 月 30 日）⁶として以下のような基本指針を出し、オープンデータに関する取り組みを行っている。

- ・ 国民参加・官民協働の推進を通じた諸課題の解決、経済活性化
- ・ 行政の高度化・効率化
- ・ 透明性・信頼の向上

このような動きの中で、茅ヶ崎市においても、オープンデータに向けた取り組みが進められている。茅ヶ崎市の公式ホームページの中に「茅ヶ崎市オープンデータライブラリ」のページ⁷があり、「オープンデータについて」、「お知らせ」、「茅ヶ崎市オープンデータ利用規約」、「茅ヶ崎市オープンデータ一覧」が掲示されている。

オープンデータ利用規約としては、知的財産権について著作権意思表示、クレジットの記載などが定められている。

そして、オープンデータとして、現在、以下が Excel ファイル・CSV ファイルとして公開されている。

- ・ 統計情報
- ・ 防災施設情報
- ・ 選挙情報
- ・ 公共施設情報
- ・ 広報紙情報
- ・ ごみ収集情報

また、藤沢市・平塚市など近隣の自治体との相互リンクも行っている。

これらのデータ公開は非常に有意義なことであり、また、茅ヶ崎市はこれらを利用するアプリのコンテストを開催することで、データ利用を促すことを試みている。今後もさらに発展が期待されるが、将来への発展の筋道を考えるために、この公開データを形式的な側面から見てみたい。

3. オープンデータの評価

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＝リーは、オープンデータのための 5 つ星スキームを提案している。これに従うと、表 1 のように、星の数で示したオープンデータのレベル分けが可能となる。

表 1 オープンデータの 5 つ星スキーム

レベル	内容	具体的な例
★	(どんな形式でも良いので) あなたのデータをオープンライセンスで Web 上に公開しましょう	Excel の表をスキャンした PDF など
★★	データを構造化データとして公開しましょう	Excel の表など
★★★	非独占 (non-proprietary) の形式を使いましょう	CSV ファイルなど

⁶ 「オープンデータ基本指針」高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定（平成 29 年 5 月 30 日）

https://cio.go.jp/sites/default/files/uploads/documents/data_shishin.pdf

⁷ <http://www.city.chigasaki.kanagawa.jp/jyohosuishin/1009746.html>

★★★★	物事を示すのにURIを使いましょう。そうすることで他の人々があなたのデータにリンクすることができます	RDF
★★★★★	あなたのデータのコンテキストを提供するために他のデータへリンクしましょう	LOD

(<http://5stardata.info/ja/> より)

上記のサイトでは、レベルごとに、消費者と公開者それぞれに対するコストと利益が示されている。

たとえば、星一つの場合は、以下のように説明されている。

(消費者として)

- ✓ 見ることができます。
- ✓ 印刷することができます。
- ✓ ローカル (ハードディスクや USB メモリ上) に保存できます。
- ✓ データを他のシステムに入力できます。
- ✓ 望むようにデータを変更できます。
- ✓ 好きなように他人とデータを共有できます。

(公開者として)

- ✓ 簡単に公開できます。
- ✓ あなたのデータが使用可能であるということ消費者に繰り返し説明する必要がありません。

レベル1の場合の評価は、「Web上にデータを公開するのは素晴らしいことです。しかしながら、そのデータは文書内に閉じ込められています。お手製のスクレイパー⁸を書かないと、文書からデータを取り出すのは難しいです。」とある。

次に、レベル2についてみると、以下の特性が加わる。

(消費者として)

- ✓ 独占的なソフトウェアを使うことで、データの集約、計算、可視化等を直接処理できます。
- ✓ 他の (構造化) 形式として書き出せます。(公開者として)
- ✓ まだ簡単に公開できます。

レベル2の評価は、「素晴らしい! そのデータは構造化された方法 (すなわち、機械可読可能な形式) で Web 上でアクセス可能となっています。しかしながら、データは依然として文書内に閉じ込められています。データを文書から取り出すには独占的なソフトウェアが必要で。」とある。

4. 茅ヶ崎市オープンデータの現在の位置づけ

上記の5つ星スキームに照らし合わせると、茅ヶ崎市のオープンデータは、現在、レベル2から3に該当すると考えられる。

たとえば、茅ヶ崎市の現在の年齢別人口は、表2のような形で公開されている。上記のレベル特性として示されているように、市民はこの表を見たり印刷したり、加工したりすることが可能、という利便性がある。公開者の立場から見ると、紙に印刷するためのフォーマットをExcelで作成し、CSVはそのままを提供しているということから、公開にかかる余分な労力は比較的少ないと考えられる。また、特にExcel形式、CSVデータ形式についてあらためて市民に説明しなくても了解してもらえる。

自治体によっては、市役所・町役場等に出向いて紙で資料がもらえるだけという場合も少なくない (それでも、提供してもらえるだけでもありがたいのであるが)。そのため、このような形でWebで一般公開されている茅ヶ

⁸ ウェブサイトから情報を抽出するコンピュータソフトウェア技術のこと

表2 現状の年齢別人口

年齢別（各歳）、男女別人口							
				(平成 27 年 1 月 1 日現在)			
年齢（各歳）	総数			年齢（各歳）	総数		
	計	男	女		計	男	女
0～4	10,464	5,435	5,029	30～34	13,580	6,680	6,900
0	1,923	1,000	923	30	2,474	1,204	1,270
1	2,137	1,139	998	31	2,723	1,404	1,319
2	2,085	1,087	998	32	2,702	1,322	1,380
3	2,148	1,112	1,036	33	2,711	1,329	1,382
4	2,171	1,097	1,074	34	2,970	1,421	1,549
5～9	10,942	5,481	5,461	35～39	17,312	8,648	8,664
5	2,061	1,005	1,056	35	3,218	1,642	1,576
6	2,220	1,116	1,104	36	3,308	1,664	1,644
7	2,290	1,146	1,144	37	3,315	1,636	1,679
8	2,156	1,114	1,042	38	3,638	1,825	1,813
9	2,215	1,100	1,115	39	3,833	1,881	1,952

(http://www.city.chigasaki.kanagawa.jp/_res/projects/default_project/_page_/001/009/746/NenreiBetsuJinkou.csvによる。)

崎市のデータには、大きな利便性がある。

ただし、Excelで読めるとはいても、十分に構造化されてはいない。たとえば、年齢0～4と、次の行の0とでは、同じ列にあっても意味が異なる。人間にとっては、5歳ごとにまとまった数値があるのは便利であるが、コンピュータの処理からみると、扱いにくい。このように人間対象の表と、コンピュータで処理するために使いやすいデータ形式とでは違いがある。レベル2、レベル3とするためには、Excel、CSVファイルにするだけでなく、プログラムで扱いやすくするための構造化データとすることが必須となる。

4.1 構造化を高めるための変換

現在のフォーマットをレベル3に変換するためには、表3に示すようなフォーマットにする必要がある。ポイントは以下の通りである。

1. 年齢（各歳）の小計を削除して一列に並

べる

2. 年齢に含まれる不要なスペースを削除し、数値型に変更
3. 各レコードに日付（今回の場合は、20170101）を設定

小計を削除しているが、上述の処理によりシステムで扱いやすくなるため、利用する側で小計を計算することが可能となり、現行の5歳間隔にかぎらず柔軟な小計を計算することができる。また、例では「計」の項目を残しているが、男女を加算して求められる値のため、必須ではない。ただ、利用する側が不要であれば使用しない対応も可能なため、残しておいても問題ない。

上記のデータをCSVファイルで提供することで、5つ星スキームのレベル3にすることができる。

表3 レベル3に合わせた変換

日付	年齢	計	男	女
20170101	0	1923	1000	923
20170101	1	2137	1139	998
20170101	2	2085	1087	998
20170101	3	2148	1112	1036
20170101	4	2171	1097	1074
20170101	5	2061	1005	1056
20170101	6	2220	1116	1104

5. よりデータを活用してもらうために

CSVで提供するデータをシステムで利用する場合には、シンプルな使い方としては、各利用者が個別にデータベースに保存し各システムで利用することが想定される。

しかし、利用する各システムで同じデータを保存しておくことは無駄になる部分も多い。特にデータ量が膨大になってきた場合には、保存用のディスク容量も大きくなってしまい、運用コストが上がってしまう。サーバーにデータを置き配信することがその対策として有効である。また、随時更新されるデータの場合にも、サーバーから取得できたほうがよい。

その際には、Web APIという仕組みを使ってサーバーを構築することで、ウェブを通してデータを取得するプログラム開発が容易になる。

5.1 APIサーバーによる配信のメリットとデメリット

API (Application Program Interface) とは、サービス提供者側で特定の指示や命令などのプログラムをほかのプログラムから利用できるように公開するインターフェイスのことである。サーバーにAPIを備えたものがAPIサーバーである。配信の際は、近年一般化しておりWebエンジニアには馴染みが深いものとなっている、「RESTful API」(RESTと呼ばれる

設計原則に従って作成されたAPI)の仕組みに則ってデータを提供することで、利用に対する障壁が少なくなる。この仕組みでは、利用者側のリクエストに応じた柔軟なデータの返し方が可能になるため、利用が促進できる。

APIサーバーによる配信のデメリットとしては、CSVファイルをサーバーに単に置くだけの場合と比べて、サーバーの構築・運用等のコストが高いことである。

5.2 JSON形式の例

データを配信する際には、一般にデータ連携等で利用されているJSON形式でデータを配信する仕組みの検討も、必要になると思われる。上記の年齢別人口をJSONフォーマットにした場合を例示したい。図1は、Excelで1列分のデータ、つまり1レコードを、JSONフォーマットにしたものである。図2は、複数のレコードの場合を示してある。

```
{
  "date": 2017010,
  "age": 0,
  "total": 1923,
  "male": 1000,
  "female": 923
}
```

図1 JSONフォーマットの例 (1件の場合)


```
[
  {
    "date": 20170101,
    "age": 0,
    "total": 1923,
    "male @ ": 1000,
    "female": 923
  },
  {
    "date": 20170101,
    "age": 1,
    "total": 2137,
    "male": 1139,
    "female": 998
  }
]
```

図2 JSON フォーマットの例 (複数件要求された場合)

5.3 5つ星スキームのレベル4、レベル5について

さらに先端的な枠組みである、レベル4、5に関しては、RDF (Resource Description Framework)、LOD (Linked Open Data) を導入する必要がある。

RDFは、ウェブサイトを利用する異なるアプリケーションソフトどうしが、データ交換を可能にするための枠組みである。WWWコンソーシアムで標準化がはかられており、また、詳しい説明が「RDF 1.1 入門」⁹として公開されている。

RDFは、実際に用いられるフォーマットが複数存在している。茅ヶ崎市のオープンデータをRDF化するにあたり、どのフォーマットを

利用し、こういった形で提供する必要があるかは、今後の課題となる。

LODは、技術の総称という位置づけとなっており、基本的にはRDFに対して、他のデータへのリンクを付与することでLODを実現する。

秋山・加藤・小出他(2016)は、「LOD実現のための11の基準」を提案し、日本の51自治体が作成した12000件のRDFの現状を調査した。その結果、すべての基準を満たしているものはないことがわかった。また誤りではないものの運用上注意すべき点があることも指摘している。たとえば、以下で説明する「共通語彙基盤」を利用しながらも、自治体間でのデータの連携が出来ないケースも散見されるようである。

RDF、LODの普及には、これらの仕様を正確に把握し正しく扱えるエンジニアの育成も必要と考える。

5.4 共通語彙基盤について

自治体のオープンデータをRDFを用いたものにしてゆくにあたり、各自治体が独自のフォーマットで作成しては共通のデータとして取り扱うことが困難になるため、統一をはかろうという動きがある。経済産業省と情報処理推進機構(IPA)は2016年11月、電子行政や企業システムで扱う文字や用語を共通化する「情報共有基盤サイト」¹⁰を開設し、「住所」や「氏名」といった幅広い業務分野に共通する基本的な語彙、つまり共通語彙基盤(Infrastructure for Multilayer Interoperability;「IMI」と略す)を公開している。IMIプロジェクトに参画してきた日立製作所は、行政システムの情報連携にIMIを使って相互運用を図る考えであり、

⁹ <http://www.asahi-net.or.jp/~ax2s-kmtn/internet/rdf/NOTE-rdf11-primer-20140225.html>

¹⁰ <https://imi.go.jp/>。このサイトによると、IMIは「電子行政分野におけるオープンな利用環境整備に向けたアクションプランの一環で、データに用いる文字や用語を共通化し、情報の共有や活用を円滑に行うための基盤である。文字情報基盤と共通語彙基盤により、行政サービスの相互運用性(Interoperability)向上を図って」いるものである。

表4 広報紙 URL に関するデータ項目案

PD3112広報紙URL		各団体が発行する広報紙の掲載URLの情報	
項目名	データ項目	記入例	参考:主要サービス提供者のデータ項目 埼玉県共通フォーマット
識別情報	キー情報。一意の数字を設定する。		識別情報
団体コード		5010005007126	団体コード
団体名		〇〇市	団体名
ID		51750	ID
タイトル		〇〇だより 平成〇年〇月号	タイトル
公開開始日時	月日は2ケタで表記 ※空欄の場合はハイフン「-」	2016-08-01 00:00:00	公開開始日時
公開終了日時	月日は2ケタで表記 ※空欄の場合はハイフン「-」	2016-08-01 00:00:00	公開終了日時
自由記述欄			自由記述欄
詳細リンク		http://www.ooo.lg.jp/koho/abc.html	詳細リンク

表5 日付型の定義案（共通語彙基盤 コア語彙 2.4.1¹¹ より抜粋）

日付型のプロパティ			
識別子	値型	回数	説明
ic:種別	<u>xsd:string</u>	0..n	種別の表記を記述するためのプロパティ用語
ic:種別コード	<u>ic:コード型</u>	0..n	種別コードを記述するためのプロパティ用語
ic:メタデータ	<u>ic:文書型</u>	0..n	データのメタデータを記述するためのプロパティ用語
ic:標準型日付	<u>xsd:date</u>	0..1	日付を記述するためのプロパティ用語
ic:年号	<u>xsd:string</u>	0..1	元号を記述するためのプロパティ用語
ic:年	<u>xsd:integer</u>	0..1	年の値を記述するためのプロパティ用語
ic:月	<u>xsd:integer</u>	0..1	月の値を記述するためのプロパティ用語
ic:日	<u>xsd:integer</u>	0..1	日の値を記述するためのプロパティ用語
ic:時間帯	<u>xsd:string</u>	0..1	時間帯を記述するためのプロパティ用語 ※協定世界時からの時間差
ic:カレンダー	<u>ic:コード型</u>	0..1	カレンダーの種類を表すコードを記述するためのプロパティ用語
ic:表記	<u>xsd:string</u>	0..1	日付の任意の表記を記述するためのプロパティ用語

¹¹ <https://imi.go.jp/ns/core/Core241.html#ic:%E6%97%A5%E4%BB%98%E5%9E%8B>

埼玉県は2017年1月、神奈川県厚木市は同年6月に、IMIに対応したデータを公開した（大豆生田，2017）。

IMIは共通語彙のドラフトを多数公開している。たとえば、表4は、IMIと埼玉県がまとめた10のドラフトの一つ、広報誌URLのデータ項目（将来の語彙の素）に関する案である。

語彙はさらに詳細に定義されており、たとえば日付を表現する用語は表5のとおりである。

共通語彙として提案されている項目は、RDFのみならず、CSVの項目名としても利用することが可能である。現時点でレベル4～5のシステムを作らないとしても、将来を見越して、共通語彙基盤を参照しつつ項目名を考えておくことは有用と考えられる。

6. 自治体オープンデータの事例

自治体のオープンデータの状況を見るために有用な情報源として、地域資源の情報をオープンデータとして共有するためのサイトCitydata¹²がある。ここには全国の市町村が公開しているオープンデータの数、アイデア数、総合評価が載せられており、評価指数ランキングも見ることができる。茅ヶ崎市からは、「茅ヶ崎市消防団各分団器具置場」等がCitydataに登録されているが、市のホームページで公開しているデータはまだ掲載されていない。

このサイトも参考にしつつ、先進的な活動を行っている自治体をいくつか紹介したい。

1) 福井県鯖江市

「データシティ鯖江」を目指して、積極的にオープンデータに取り組んでいる。LODに対応しているデータも多く、自治体のオープンデータの取り組みとしては、かなり進んでいる。オープンデータは、以下で公開され

ており、オープンデータを活用したアプリも数多く紹介されている。

<http://data.city.sabae.lg.jp/opendata-list/>

2) 福岡県福岡市他

福岡市が中心となって、以下の自治体オープンデータのサイトを運営している。

<https://www.open-governmentdata.org/>
福岡県、福岡市、北九州市のオープンデータを掲載している。

3) 北海道森町

小規模自治体にも関わらず、オープンデータに取り組んでいる好例として、森町を取り上げたい。森町は、森町オープンデータポータル¹³、また、自治体の多くが公開しているLinkData（<http://linkdata.org/>）で、データを公開している¹⁴。

たとえば、森町のホームページの中の人口世帯数のページでは、人口推移グラフ（図3参照）と人口推移表が示しているが、LinkDataから自動生成している。

山形（2016）は、森町でのオープンデータについての取組を行った経緯をまとめている。その中で、森町でオープンデータの取り組みを始めてからの自治体における各種の困難、特に小規模自治体においてデータの重要性を認識してもらうのが難しいことを示しつつ、それを乗り越えて、自治体の規模や地域に即した手段で、無理なく効果的に情報のオープン化を進めていることを報告している。このような取り組み例は、茅ヶ崎市においてオープンデータの取組を進める上で同様の問題が発生した場合にも、参考になると思われる。

茅ヶ崎のオープンデータについての今後の提

¹² <http://citydata.jp/>

¹³ <http://www.town.hokkaido-mori.lg.jp/docs/2014101000027/>

¹⁴ http://user.linkdata.org/user/hokkaido_morimachi/work

グラフについて：ブルー（人口男）、ピンク（人口女）

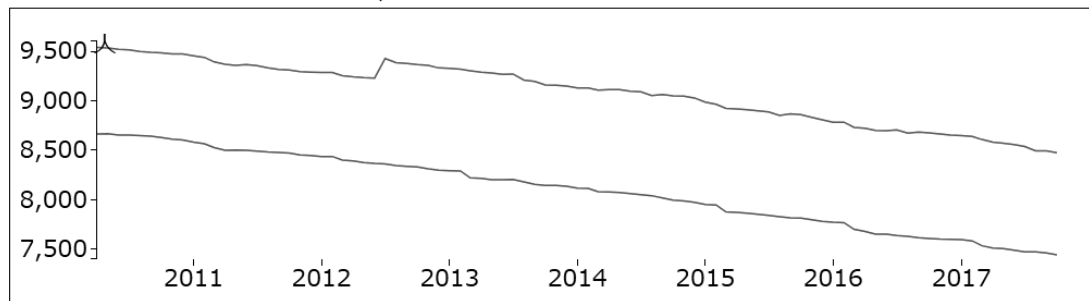


図3 森町の人口推移グラフ

案としては、構造化したCSVファイルにすることで、5つ星スキームのレベル3までは最低限対応することが求められるであろう。さらに、RDF、LODの対応を進めていくべきではあるが、それぞれの仕様を正確に把握し、対応を進めていく必要がある。

7. おわりに

茅ヶ崎市のオープンデータは、市のオープンデータ規約のもとで自由に使うことができる。市のホームページにExcelファイル・CSVフォーマットデータとして公開されており、5つ星スキームのレベル2から3として評価される。ただし、機械可読データのため再加工が可能とはなっているが、紙媒体での表示を引き継いでいる点がみられる。構造化されたCSVに変換することで、プログラムでの利用がより容易になる。

次の段階は、レベル4・5ということになるが、自治体ではRDF、LODとして公開していく流れがあり、それに合わせていく必要性がある。これは上述のように構造化されたCSVに変換するのに比べ、仕様が複雑になり、専門知識の有るエンジニアの支援が必要になると思われる。また、データ作成においては、共通語彙基盤を意識してデータ項目を統一してゆくことは、移行を速やかにするために有効と考えられる。

謝辞

茅ヶ崎市情報推進課に市のオープンデータの状況についてお話を伺うことができました。ここに記して感謝いたします。

引用文献

- 秋山梓・加藤文彦・小出誠二・海沼 靖夫 (2016) 自治体が公開しているRDFの現状と課題、人工知能学会研究会資料、(https://researchmap.jp/?action=cv_download_main&upload_id=125871 閲覧日：2017年12月11日)
- 大豆生田崇志 (2017) 共通語彙基盤，日経コンピュータ，2017年11月09日号，p.92.
- 山形巧哉 (2016) 人口1万6404人北海道森町目線でのオープンデータ，(http://goikiban.ipa.go.jp/contents/2016/06/20160603_05.pdf 閲覧日：2017年12月11日)

参考URL

- e-Stat/ 統計 LOD 関連
 - e-Stat
 - <http://www.e-stat.go.jp/>
 - 統計 LOD
 - <http://data.e-stat.go.jp/lodw/>
 - 自治体等の統計データに関する LOD 構築の手引き
 - <http://data.e-stat.go.jp/lodw/guidelines/>

handbook/

Linked Open Data (LOD) による統計データの提供：政府統計データ (e-Stat) の新しい形

https://www.jstage.jst.go.jp/article/johokanri/59/12/59_812/_html/-char/ja/

行政機関におけるオープンデータ公開サイトの構築

https://www.jstage.jst.go.jp/article/tjsai/advpub/0/advpub_LOD-27/_article/-char/ja/

■ オープンデータ

一般社団法人オープン&ビッグデータ活用・地方創生推進機構 (VLED)：オープンデータガイド

http://www.vled.or.jp/news/1507/150730_001192.php

総務省：オープンデータ戦略の推進

http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/opendata/

■ RDF

RDF 1.1 入門 (和訳)

<http://www.asahi-net.or.jp/~ax2s-kmtn/internet/rdf/NOTE-rdf11-primer-20140225.html>