



佐藤宏之
飯塚京士
三島和恵

NTT レゾナント(株)
日本電信電話(株) NTT 情報流通プラットフォーム研究所
(株)建設技術研究所

オープンガバメントと オープンデータ

オープンガバメントとは

●オープンガバメントの本質

近年、米国や英国をはじめとする世界各国で、Webの特性を活用した積極的な政府情報の公開や、行政への市民参加を促進するオープンガバメントが急速に進展している。その本質を端的に示しているのは、オープンガバメント3原則、すなわち「Transparency (透明性)」、「Participation (国民参加)」、「Collaboration (政府間および官民の連携・協業)」である。これは、米国のオバマ大統領の就任直後(2009年1月21日)に大統領メモとして表明された。オバマ政権は各政府行政機関に対して、この3原則に従って市民に開かれた「透明でオープンな政府」となるように求めた。

これまでもITを利用した行政サービスの効率化、国民の利便性の向上の取り組みは行われていた。しかし、オバマ新政権において、さらに現在、世界各国で進められている取り組みは、新たなWeb技術の積極的な導入により「情報公開」と「政府の各種意思決定プロセスへの市民の参加」に主眼を置いたところに大きな特徴がある。

●本稿の焦点はオープンデータ

Webを活用した政府によるオープンガバメントの具体的な取り組みは「政府が保有するデータ(統計情

報や行政情報)の公開」と「国民から広く意見や提言の募集」を行うための2つの情報システムの構築からなることが、米国の事例に限らず一般的である。

本稿では特に前者の政府データの公開に関して、本特集の主題であるLinked DataおよびRDFデータを公開している事例や、他の情報システムによる再利用を意識した機械可読なオープンデータに焦点をあてたい。図-1はlinkeddata.orgが公開している異なる年月日に作成されたThe Linking Open Data cloud diagram (LODクラウド図)の政府関連データが表示されている箇所を対比したものである。英国の政府に関連するデータが中心であるが、2009年3月から2010年9月の1年半あまりの間に政府関連のオープンデータが急速に増えている様子が窺える。

以降では、各国政府のオープンデータの取り組み、オープンデータがもたらす社会、さらに技術課題や今後の展望を日本の現状も織り交ぜながら述べたい。

各国政府のオープンデータの取り組み

●Data.gov (米国)

2009年5月21日、米国政府は、前述のオバマ政権のかかげるオープンガバメントの方針に沿って、連邦政府のさまざまな機関が保有するデータを国民が入手できるサイトとしてData.gov^{☆1}を開設した。そ

☆1 Data.gov-About, <http://www.data.gov/about>

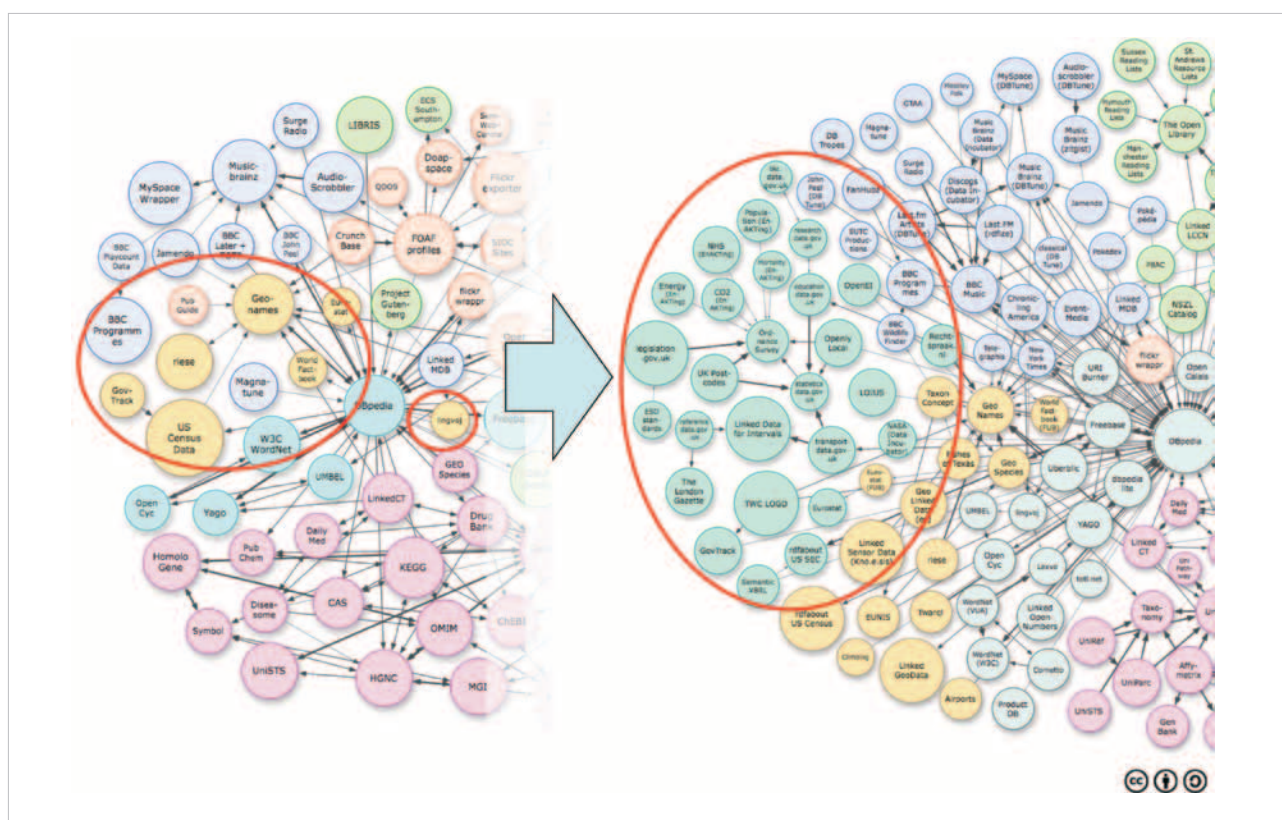


図-1 LOD cloud diagram における 2009-03-27 (左)と 2010-09-22 (右) のガバメントデータ(赤線内)の数の比較(linkeddata.org の図を加工)

の目的として、価値の高い、機械可読なデータセットへの市民のアクセスを増加させることを掲げている。

表-1 に分野ごとに公開されているデータセット数を示す。地理(環境、地質、地理学データなど)、福祉(健康、福祉サービス、教育など)、経済(財政、物価、金融、収支など)、統計(人口統計、国際統計)などが提供されている。2010年10月現在 Data.gov に掲載されているデータセットの数は約 2,900 件(“Raw” Data Catalog の件数、このほか、Geodata Catalog として約 30,000 件のデータが公開されている)であり、公開当初の 47 件から大きく伸びている。また、オープンフォーマットとしての Linked Data の採用も開設後に活性化している。表-1 には、RDF に変換されて公開されているデータセット数も示しているが、2010年10月現在、総計 62 億トリプル(データ)が、Data.gov 上で公開されている。

2009年12月8日には指令「Open Government Directive」が公開されている¹⁾。連邦政府機関の責任者に対し、タイムリーな情報公開や検索やダウンロードによって再利用できるオープンフォーマット

の採用を求めた上で、45日以内に価値の高いデータを少なくとも3件以上新たに Data.gov を通じて登録すること、60日以内に同活動の入り口となる Web ページを各機関の Web サイト上に開設することなどを各機関に求めている。今後もデータセットが次々と追加されたり、アップデートされたりすることが予想される。

●欧州(主に EU)のオープンデータ

EU は、統一通貨ユーロが誕生し経済圏の統合を果たしてから10年が経過した。国を越えたさまざまなサービスの統合・共有の実現には Public Sector Information (PSI)、すなわち公的機関の公共情報の公開と活用が必要である。2003年11月には公共情報の再利用に関する欧州指令が公布され²⁾、各国で PSI の公開が促進されている。欧州の PSI の再利用による市場規模は、EU およびノルウェーで 26 ～ 47 兆 € になるとの評価・試算調査結果がある³⁾。欧州では制度や法的な枠組みの方を先行してオープンガバメントの取り組みを行っていることに特徴がある。

分野名	全件数	XML	CSV/Text	KML/KMZ	Shapefile	RDF		Other
地理	1809 件	72 件	1478 件	18 件	165 件	125 件	約 3 億トリプル	85 件
福祉	182 件	18 件	143 件	0 件	0 件	112 件	約 23 億トリプル	39 件
雇用	124 件	52 件	62 件	0 件	0 件	19 件	約 3 億トリプル	11 件
司法	58 件	0 件	47 件	0 件	0 件	5 件	約 2 億トリプル	11 件
経済	207 件	49 件	119 件	0 件	0 件	42 件	約 6 億トリプル	45 件
政治	13 件	4 件	4 件	0 件	0 件	3 件	約 30 万トリプル	9 件
産業	214 件	9 件	84 件	0 件	0 件	29 件	約 1 億トリプル	119 件
科学技術	90 件	15 件	66 件	2 件	0 件	17 件	約 1 億トリプル	14 件
国防	16 件	9 件	4 件	0 件	1 件	2 件	約 3 百万トリプル	2 件
統計	65 件	3 件	18 件	0 件	0 件	14 件	約 25 億トリプル	44 件
その他	99 件	37 件	8 件	0 件	0 件	12 件	約 2 百万トリプル	57 件

表 - 1 2010 年 10 月 8 日現在の Data.gov における Raw データセット数と公開フォーマット

PSI の中でも、特に地理空間情報は、公開による再利用の期待が高く、2007 年に空間情報基盤の整備に関する欧州指令が発効された⁴⁾。欧州委員会におけるこの取り組みは INSPIRE と呼ばれ、2008 年から 6 月にカンファレンスが開催され、2010 年のカンファレンスでは英国を中心とした Linked Data の取り組みが紹介されるなど、地理空間情報と Linked Data の取り組みの発展が期待された。

また、EU では「知」への投資が EU の発展につながるという考えに基づき、研究開発費を GDP の 3% に増加させるという目標を掲げている。2007 ~ 2013 年の期間を対象とした第 7 次研究枠組計画 (FP7) に関する研究投資の対象プロジェクトとして「LOD2-Creating Knowledge out of Interlinked Data」が採用され、2010 年 9 月 1 日からスタートした。このことから、欧州では Linked Data に対する取り組みが期待されていることが分かる。

欧州の中でも政府情報の公開や Linked Data に関しては英国が最も先行している。そのため、英国の取り組みについて独立して次節に示す。

● data.gov.uk (英国)

data.gov.uk^{☆2} は 2010 年 1 月に開設された英国政府の取り組みである。2009 年 6 月に Gordon Brown 前首相からアドバイザーに任命された Tim Berners-Lee や、Office of Public Sector Information (OPSI) の e サービスと戦略の代表者であり、W3C の eGovernment の取り組みを進める John Sheridan の意見が反映されている。これにより data.gov.uk は、データ公開・再利用に関して、Linked Data およびセマンティック Web を強く意識したサイトになっている (図 - 2)。RDF によるデータ公開を積極的に進めており、サイト内でデータソースに対して SPARQL (セマンティック Web データのクエリ言

☆2 Opening up government data | data.gov.uk, <http://data.gov.uk/>



図 - 2 data.gov.uk

語)による検索が可能である。また各データセットにデータが属するカテゴリを示す複数のタグが振られ、タグから該当する分野のデータセットを検索できる。開設時に2,500件程度だったデータセットは2010年10月初めには約4,200件に達している。

data.gov.ukの開設以前から、英国は政府機関が保有する情報の公開には積極的である。2007年には、情報公開による公共サービスの向上に関して、外部の研究者にPower of Informationと呼ばれる提言を委託し、その結果を踏まえて、各行政機関の情報公開をOPSIが推進している。英国のオープンガバメントは、Gordon Brown前首相がリーダーシップをとって進めたことで大きく進展したが、労働党から保守党への政権交代後もDavid Cameron首相がすぐに公共機関の透明性向上に向けた政府データの公開に関する書簡⁵⁾を公表するなど、取り組みは継続的に進められている。

●日本の取り組み

日本政府は2010年5月11日に「新たな情報通信技術戦略」(通称「新IT戦略」)を発表した。「政府・提供者が主導する社会から納税者・消費者である国民が主導する社会への転換には、徹底的な情報公開による透明性の向上が必要であり、そのために情報通信技術が果たす役割は大きい」との基本認識が示されている。その上でオープンガバメントに関しては「2013年までに、個人情報の保護に配慮した上で、2次利用可能な形で行政情報を公開し、原則としてすべてインターネットで容易に入手することを可能にし、国民がオープンガバメントを実感できるようにする」という目標を設定している。

《「新IT戦略」以前の取り組み》

内閣官房のIT戦略本部では、「新IT戦略」の発表前の前政権時代に、次世代電子行政サービス基盤等検討プロジェクトチームを設置し、将来的なオープンガバメントに向けた構想についての課題や具体的な実現方法の検討を行い、レポート⁶⁾を公開している。

《オープンガバメントラボ(日本)》

2010年6月22日に発表された新IT戦略の工程表では、行政情報の公開・提供と国民の政策決定への参加促進(オープンガバメントの推進)に関して、2010年度中に行政情報の公開対象や公開方式、活用方策に関する指針を検討し、2013年までに達成する公開につなげるとしている。また国民の政策決定への参加促進に向けて、各府省が共同利用できる意見募集システム(アイデアボックス)の実証実験を2010年度に行い、2011年以降の導入につなげる計画となっている。

このような計画が策定される中、経済産業省は2010年7月29日、実験サイト「オープンガバメントラボ^{☆3)}」を正式にオープンした。「アイデアボックス」の開設が先行したが、2010年9月28日からは行政情報や統計情報を提供する「データボックス」(計画時には米国や英国に合わせたと思われるData.go.jpという仮称も見え隠れした)の実証実験が開始された(図-3)。

☆3 オープンガバメントラボ, <http://openlabs.go.jp/>



図-3 経済産業省によるオープンガバメントの実証実験サイト「オープンガバメントラボ」の「データボックス」

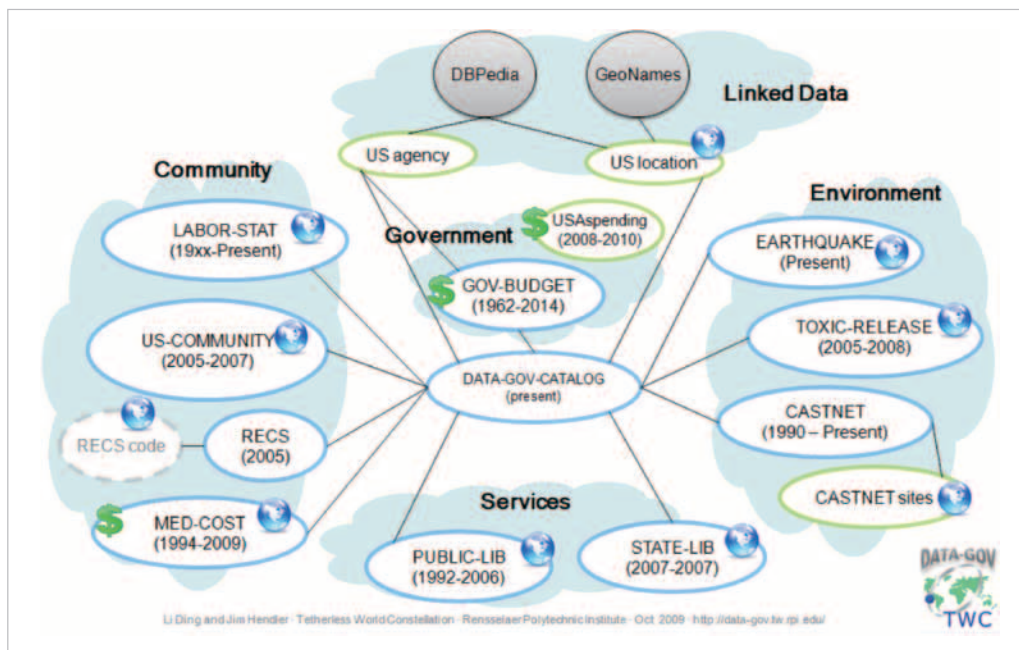


図-4 レンセラー工科大学の Tetherless World Constellation による
2009年10月時点で RDF 化された政府データ(同 Web サイトより)

データは機械可読な XML による提供が検討されているが、2010年10月10日時点では PDF と EXCEL のフォーマットのデータ 249 件が登録されている。

残念ながら、日本は他の先進国と比べると政府の情報公開、情報活用は遅れているといわれている。しかし、2008年4月1日には、総務省が中心となって政府統計のポータルサイト「政府統計の総合窓口(e-Stat)」が開設されている。各府省の統計情報をキーワード検索できたり、都道府県や市区町村の人口分布などの情報を地図上で閲覧したりすることが可能となっている。また、電子政府の総合窓口 e-Gov が用意され、法令や行政文書ファイル管理簿の情報提供サービスを行っている。

問題は、米国や英国の現状と比較したときに、生データが直接オープンなフォーマットで公開されていることが少ない点であった。現在は公開対象や個人情報保護を考慮し、日本に適した情報公開の在り方を模索し、オープンガバメントラボを通じて実験を行っている段階といえる。

開かれたデータがもたらす社会

ここまで各国政府によるオープンデータの状況を眺めてきたが、以降ではデータの公開により、新た

に生じた現象や期待される効果について述べる。

●市民活動によるデータの加工・配信

米国政府の Data.gov で、開設当初公開されていたデータは、Linked Data ではなかった。レンセラー工科大学の Tetherless World Constellation (TWC) の Jim Hendler と Li Ding が開始した Data.gov Wiki プロジェクト^{☆4}により RDF データに変換され LOD クラウドの一部を構成するようになった(図-4)。現在ではこのプロジェクトによって変換されたデータが政府の Data.gov で直接ホストされるようになっており、データ公開の段階においても官民の協業が実現しているといえる。2010年9月22日現在の LOD クラウドには、イタリアの政府による大学統計データがあるが、これも政府の情報を基に第三者が LOD に加工しているものである。

また、政府によるオープンガバメントの機運が高まる以前から、市民の啓発を目的とし、政府に関連するデータを市民団体などが利用しやすい形にして配信する試みがあった。Jashua Tauberer が米国議会に関するデータを公開しているサイト GovTrack.us では、RDF 化されたデータを入手することができ、こ

☆4 Data.gov Wiki, <http://data.gov.tw.rpi.edu/wiki>

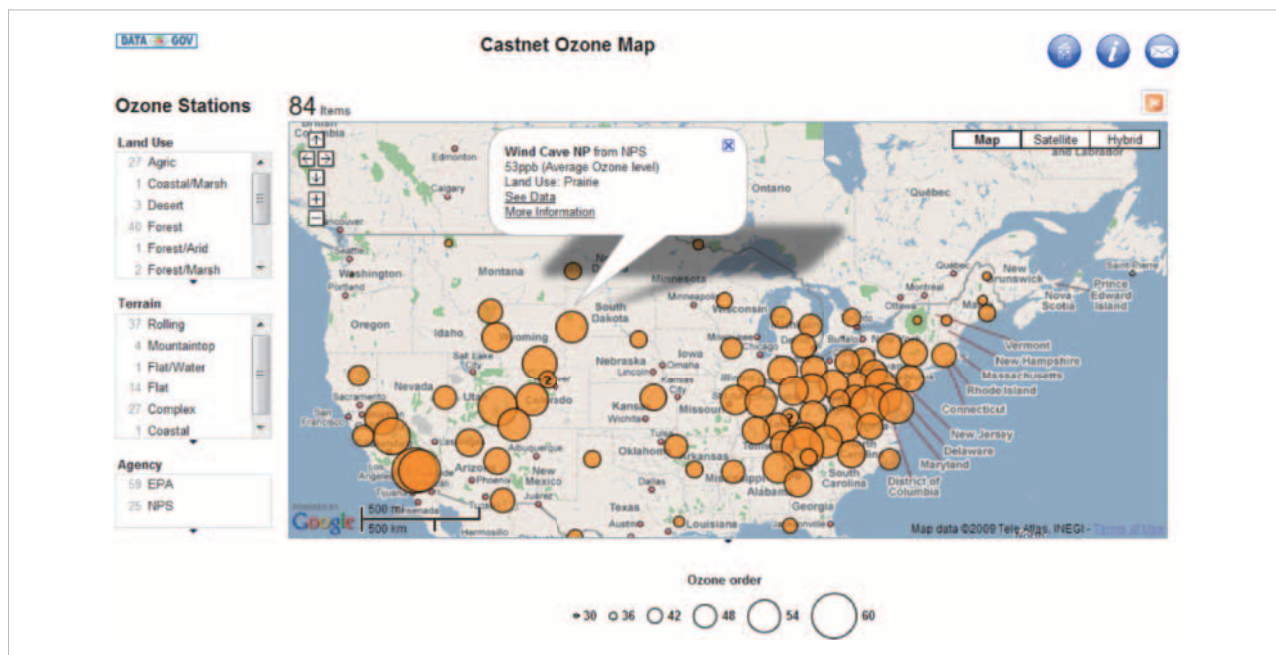


図-5 Clean Air Status and Trends-Ozone

れも LOD クラウドの一部を構成している。同氏は政府統計情報の RDF 化も進めていたが、こうした試みは以前と比べると、より多くの人々がかわるようになり、ツールの開発なども進み、扱うデータ量が拡大し、変換にかかる時間も大幅に短縮されてきている。

●民間による情報の活用・商用利用

レンセラー工科大学の TWC では Data-gov Wiki および Linking Open Government Data (LOGD) プロジェクトにおいて、政府の Linked Data を活用したアプリケーションを探る試みが行われており、Data-gov Hackathon'09 などが開催されている。Data-gov Wiki のサイト上ではデータを活用したアプリも公開され、統計情報を分かりやすく可視化したり、地図上で各種データをマッシュアップして見せたりするなど Web2.0 的なデータ加工が行われている。図-5 は米国政府の Data.gov からリンクされている Clean Air Status and Trends-Ozone のデモ^{☆5}である。環境保護庁のデータと地図情報を重ねることでオゾンの状態を視覚化している。ほかにも地域ごとのブロードバンドネットワークの利用率、健康に関する情報(喫煙率や肥満などの状況)を視覚

化したり、ホワイトハウスの来訪者の割合をグラフ化したりするなど多数のアプリが作られている。

Linked Data に限らない政府のオープンデータを活用する取り組みとしては、オバマ政権の CIO (最高情報責任者) として Data.gov を推進した Vivek Kundra が、ワシントン DC の CTO (最高技術責任者) 時代に開催した Apps for Democracy が有名である。DC の公共のデータフィード Data Catalog を用いて、DC のツアープランを作成できるアプリや、自転車ルートや自転車盗難情報をマッシュアップしたアプリ、深夜の安全な帰宅ルートを紹介するアプリなどが開発された。従来のように政府が個々のアプリを個別に調達するのとは比べて、賞金(2008 年は 2 万ドル)は必要なもの、高い投資対効果があるとする意見がある。

また、英国政府は 2008 年に Show Us A Better Way という政府データの再利用に関するコンテストを実施している。これは、Power of Information の提言の 1 つである実証実験の実施を受けて開催されたものであり、優勝者には 2 万ポンドが贈られ、アイデアをさらによくするために同額の追加ファンドがある。すでに運用が始まっているものとして学校の場所 (UK School Maps) や学校選択のための情報を見ることができるサービス (SchoolGuru) や陸地測量局のデータと Google マップとをいっしょに

^{☆5} Clean Air Status and Trends-Ozone, <http://data.gov.tw.rpi.edu/demo/exhibit/demo-8-castnet.php>



参照できるサービス (Where's the Path) などがある。

英国政府はデータの商用利用も奨励しており、公開情報に独自の付加価値を付けることによるビジネスができる。SchoolGuru も有料の電話アドバイスを行ったり、ガイドの販売を行ったりしている。data.gov.uk の公開情報に基づいて、ケアハウスを比較して選べるサイト (Best Care Home) が立ち上がり、ケアハウス側からも有料で掲載して欲しい情報を提供できるようにすることも検討されている。ケアハウスに関する公開データを用いた試みは、同じ英国内でほかにも複数の民間の組織が比較サイトなどを立ち上げている。データの公開により、民間の知恵や競争を導入し、自国や地方の経済を活性化させようとする試みが進んでいる。

●行政へのフィードバック・官民協業

前述のような現象を、Tim O'Reilly は gov2.0 (ガバメント 2.0) と呼び、そのコンセプトを整理している⁷⁾。氏は Government As a Platform, すなわち政府がプラットフォーム化しなければならないと主張している。政府がデータや API を提供することにより、官民の協業により国民が使いやすいシステムが構築されたり、民間に新たなイノベーションやビジネスに挑戦する機会を与えたりすることを重視している。

米国のオープンガバメントを象徴するシステムの 1 つに USAspending.gov がある。これは国民の税金が何にどのくらい使われているのか、契約先も含めて情報を公開している。また景気刺激策としての支

出や効果が見える化するサイト Recovery.gov も立ち上げている。これらは政府の施策だが、今後は自ら行わずに民間がさらに使いやすいシステムを実現できるように、政府はプラットフォームの整備に集中するという方向性も考えられる。

flyontime.us^{☆6} は米連邦航空局が Data.gov で公表した、民間航空会社の定時飛行の履行に関するデータを利用した旅行者支援のサービスを行っている。政府データのほかに気象データや空港利用者からのセキュリティチェックの待ち時間に関するデータをボランティアで集約しており、結果としてそれが税金の節約、政府支出の抑制につながると述べている。公共交通の情報などは公開するだけで民間が使える新たなサービスを生み出す事例が各国で見られる。

このような官民協業の成果は最終的に行政へフィードバックされるべきである。Fix My City DC は市民が市へ問題箇所を指摘し、行政とコミュニケーションできるツールである。同様の試みは英国でもある。ソーシャルメディアなどを組み合わせ、公開されたデータに対して市民からフィードバックや提言があり、データを活用したシステムが改善されるだけでなく、行政の在り方自体が改善されるというフィードバックループがまわることが重要である。この様子を 図-6 に示した。今後は、情報公開と政府への意見募集の 2 つのシステムが有機的に結び付き、データを見せるサイトは長期的な視点でみたときに PDCA サイクルを確認できるツールとして機

☆6 FlyOnTime.us, <http://flyontime.us/>

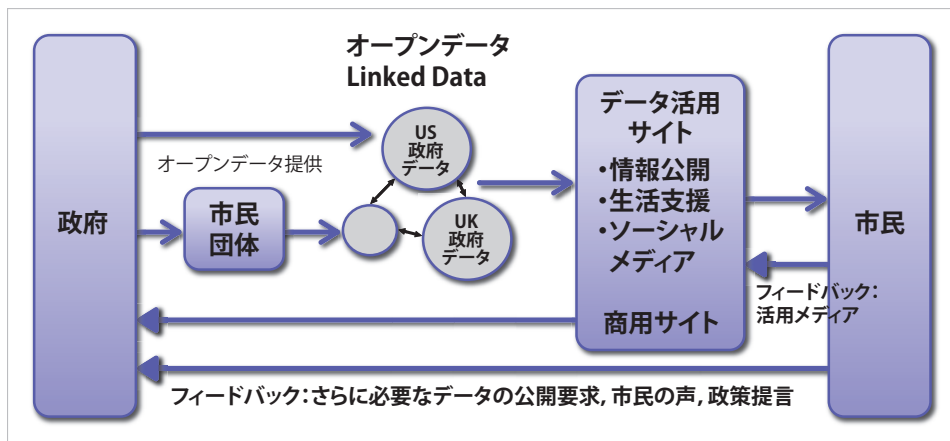


図-6 データの公開と活用, フィードバックのループ



能することが理想である。

日本のオープンガバメントラボのシステム構築計画でも行政情報提供基盤が提供した情報に基づき、国民・行政職員の共同作業可能な Wiki を提供し、そこで挙げた意見を集約して再度提供する実験の計画などが示されている。

課題と今後の展望

●データ公開のガイドラインの必要性

データはただ公開すればすぐに活用できるというわけではない。Data.gov の取り組みは一定の評価がなされているものの、たとえば、Data.gov Wiki では、以下の課題を指摘している。

- 異なるデータセット間で部分的にデータが重複していたりする
- データフォーマットが必ずしも機械可読なものではなく、データの解析にかなりの労力が必要となる
- データのアクセスポイントとして示されている Web ページで、データ全体を一度にダウンロードできなかったり、対話的な操作を要求されて部分的にしかデータを引き出せなかったりするなど、コンピュータによるアクセスがしづらいことがある

日本の e-Gov, e-Stat も URL からそれに対応する情報を直接迎れない、データのダウンロードが容易でないという問題がある。

上記のような課題に対しての解決方法として、データ整備のガイドラインの制定が必要である。W3C では eGovernment Interest Group が立ち上がっている。ここでは、

- オープンスタンダードの活用
- 透明性と参加
- 継ぎ目のないデータの統合

を活動のスコップとし、Web の利用による政府へのアクセスを改善するためのガイダンスとなるドキュメントの発行を目指している⁸⁾。またオープンガバメントデータの発行に関するガイドラインも作成中である⁹⁾。オープンデータフォーマットや、生データの目録といったドキュメントの重要性、データの所在を示し情報資源を識別するための URI の使

い方について言及している。

● Linked Data の特性を活かしたアプリケーションの期待

現在公開されているアプリの多くは1つないしは2つのデータセットのデータを、グラフ化したり、Google マップなどの地図情報と組み合わせたりして視覚化しているものが多い。データを Linked Data として公開するメリットは、異なる複数のデータセットを組み合わせたアプリを作りやすいことにあるといえる。米国と英国の犯罪を比較したアプリなどが登場したりしているものの、Linked Data の特性を活用した真の意味でのキラーアプリがまだ登場していないように思われる。またそのようなアプリを作ろうとした場合には、異なるデータや属性間のオントロジーを用いたマッピングなどが必要となり、複雑化したセマンティック Web の反省を踏まえて進められたともいえる Linked Data が、複雑な問題を再度解くことを迫られる可能性がある。

Linked Data のシンプルさを維持しながら、データの識別のための ID というべき URI や、データの属性やカテゴリ (RDF のプロパティやクラス) のポキャブラリの定義に関する最低限の決めごとについてのコンセンサスを各国の政府機関内、さらには政府間で得ることが今後の展開の鍵となると考える。EU で Linked Data が注目される理由は、各国をデータでつなげ、域内の国同士の正確な比較や物流の合理化などのニーズがあるためと考えられる。具体的なニーズからなる今後のキラーアプリの登場を期待したい。

●技術以外の課題

最後に技術以外の課題を述べて本稿を締めくくりたい。景気後退の状況下において、データの公開によって市民が動き出し、新しいビジネスを創出するばかりか、行政に対して建設的な提言を行い、具体的に行政を支援するツールまでも生み出してしまいう世の中は理想である。しかし、これには技術だけでなく仕掛けや文化が必要である。開発者のモチベーションとメリットは賞金だけでなく、名誉や社会奉

仕の精神といったものと関係し、市民が単に行政のサービスを楽しむだけでなく、自らが主体的に参加する意識を持っているかといった文化的な面も影響する。日本の社会にオープンガバメントが定着するか、日本ならではの今後の取り組みに期待したい。

また今後、情報の公開対象を拡大した際には、それぞれ単体で見たときには問題のないデータでも、組み合わせることで個人情報を引き出せる可能性もある。日本政府の新 IT 戦略の取り組みでは、個人が特定できない形に情報の集約化・匿名化を行うなど、個人情報・プライバシー保護の対策のための技術開発が工程に織り込まれている。そのままでは公開できない情報を加工することにより、公開対象を広げ、活用の幅を広げることができれば、新たなビジネス創出のチャンスとなる。

参考文献

- 1) Open Government Directive : The White House, <http://www.whitehouse.gov/open/documents/open-government-directive>
- 2) Directive 2007/2/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information, Official Journal of the European Union (2003).
- 3) Dekkers, M., Polman, F., te Velde, R. and de Vries, M. : Final Report of Study on Exploitation of Public Sector Information-Benchmarking of EU Framework Conditions, The European Commission's MEPSIR Report (Measuring European Public Sector Information Resources) (2006).
- 4) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) Official Journal of the European Union (2007).
- 5) Letter to Government Departments on Opening up Data | Number10.gov.uk : <http://www.number10.gov.uk/news/statements-and-articles/2010/05/letter-to-government-departments-on-opening-up-data-51204>
- 6) 次世代電子行政サービス基盤等検討プロジェクトチーム 中間報告書 : <http://www.kantei.go.jp/jp/singi/it2/nextg/pdf/100104houkokusho.pdf>
- 7) O'Reilly, T. : Gov 2.0 : The Promise Of Innovation : <http://www.forbes.com/2009/08/10/government-internet-software-technology-breakthroughs-oreilly.html>
- 8) Improving Access to Government through Better Use of the Web : <http://www.w3.org/TR/egov-improving/>
- 9) Publishing Open Government Data : <http://www.w3.org/TR/gov-data/>

(平成 22 年 10 月 31 日受付)

佐藤 宏之 (正会員) h-sato@nttr.co.jp

NTT レゾナント (株) サーチ事業部サービス開発部門担当課長。セマンティック Web を利用した情報統合、大規模分散データ処理システムに興味を持つ。慶大環境情報学部講師 (非常勤)。博士 (情報学)。

飯塚 京士 (正会員) iiduka.kyo@lab.ntt.co.jp

NTT 情報流通プラットフォーム研究所研究員。セマンティック Web 技術を利用した情報分析技術に興味を持つ。現在、社内情報共有システムの研究開発に従事。

三島 和恵 mishima@ctie.co.jp

(株) 建設技術研究所。地理空間データ、特に防災分野のデータのオープンソーシャル化、流通に際する要素技術に興味を持つ。

