

Linked Open Data によるボトムアップ型 オープンガバメントの試み

深見嘉明¹ 小林巖生² 嘉村哲郎³⁴ 加藤文彦⁵
大向一輝⁵ 武田英明⁵ 高橋徹⁶⁷ 上田洋⁷

Web of Data というコンセプトのもと、ウェブ上で計算機処理可能なデータを分散的に生成し、それを互いにリンクさせることにより、共有財としてのデータベース資源を確立するという試み、それが Linked Open Data である。その特質を活用し、これまで行政が単独で担ってきた情報収集、分析とそれを生かした政策実現、住民サービス実施を分担するという試みが始まっている。本論文では、住民コミュニティと行政の連携を通じたボトムアップ型オープンガバメントの試みについて紹介し、ウェブ標準技術が共有財の創出と受益者の拡大にどのように貢献できるかについて検討する。

Activities to realize open government with linked open data

Yoshiaki Fukami¹ Iwao Kobayashi² Tetsuro Kamura³⁴
Fumihiko Kato⁵ Ikki Ohmukai⁵ Hideaki Takeda⁵
Toru Takahashi⁶⁷ Hiroshi Ueda⁷

Linked Open Data (LOD) is the concept that realizes developing database resources as commons by generating machine readable data linked each other on the web with distributed activities. LOD enables trial of supply public services, which used to be provided only by public sector such as research, analysis and policy planning deconcentratedly. This paper introduces activities with cooperation among citizens and administrative bodies. We analyze how web standards can contribute to create commons and increase beneficiaries.

1. ウェブの創成

1.1 ウェブの果たす役割の広がり

World Wide Web Consortium (W3C)により策定が進む HTML5 などの標準技術の進歩により、ウェブはアプリケーションプラットフォームへと進化している。同時に、タブレット型デバイスの登場や、電子書籍リーダーの普及により、音楽、映像、そして書籍といった様々なコンテンツの流通プラットフォームとしての役割も大きくなっている。しかし、ウェブはもともと、動的なコンテンツを生成・流通するためのインフラではなく、静的な文章（ドキュメント）を共有するための手段として開発されたのであった。

1.2 ウェブ開発の経緯

ウェブが生み出されたのは 1989 年、欧州原子核研究機構(CERN)においてであった。CERN はフランスとドイツの国境にまたがって設置された広大な研究機関である。その広大さの故に、同じ機関に勤務する研究者同士の情報交換は困難であった。必要とされる知識は所内にあるにも関わらず、それがどこにあるのか、その知識を有している同僚がどこにいるのかすらわからない状態であった。

そうした状況を打開するために開発されたのが、World Wide Web であった。当時 CERN に勤務していたエンジニアの Tim Berners-Lee は、研究者同士による研究成果の共有を実現するために、容易に論文などの研究資料を共有するためのインフラとして、World Wide Web を開発した。[1]つまり、ウェブは本質的にドキュメントデータベースとして開発されたと言えるであろう。[2]

¹ 慶應義塾大学 SFC 研究所 次世代 Web 応用技術・ラボ

² Advanced Web Applications Laboratory, Keio Institute at SFC

³ Open Community Data Initiative

⁴ 総合研究大学院大学 複合科学研究科

School of Multidisciplinary Sciences, The Graduate University of Advanced Studies

⁵ 東京藝術大学

Tokyo University of the Arts

⁶ 国立情報学研究所

National Institute of Informatics

⁷ ATR メディア情報科学研究所

Media Information Science Laboratories, Advanced Telecommunications Research Institute International

株式会社 ATR-Promotions

ATR Promotions, Inc

2. Web of Data

2.1 ウェブ利用の広がり

ウェブはその後急速に普及し、研究者のためのインフラから、世界中の誰もが日常的に用いる情報プラットフォームへと役割を変えていく。研究機関や研究者のみならず、企業や個人が次々とウェブサイトを解説するようになった。ウェブのマークアップ言語である HTML は、習得が容易で比較的厳密性が求められないためユーザに受け入れられやすく、ウェブが広く普及する大きな要因の一つとなった。

また、HTML などの知識がないユーザも、Content Management System の普及や Consumer Generated Media の増加により、気軽に情報発信ができるようになった。それに伴いウェブ上で公開・共有される情報は、加速度的に増加した。

2.2 Machine Readable Document

しかし、手軽に扱える HTML は、あくまでも人間が見たり、読んだりするためのコンテンツを製作するためのツールである。ウェブは、インターネットという計算機ネットワーク上に構築されたインフラであるにも関わらず、人間が読む(Human Readable)コンテンツ=HTML と、計算機処理できる(Machine Readable)データ=XML とは、別のものとして取り扱われてきた。Human Readable Document に含まれる情報を、人間だけでなく計算機にも処理できる形にしようとする試み、それがセマンティックウェブである。

セマンティックウェブの実現に必要なとされるのは、計算機処理可能な意味情報(メタデータ)の付与と、厳密なマークアップ、そして意味体系(オントロジ)の構築である。Tim Berners-Lee が Director を務める World Wide Web Consortium (W3C)では、HTML4 の次期バージョンとして XML の機能を取り込んだ XHTML とともに、オントロジ構築のための言語、OWL(Web Ontology Language) などのセマンティックウェブ関連の規格策定を進めてきた。

しかし、XHTML は厳密なマークアップを必要とすることから一般のウェブ制作現場には広がらず、ブラウザベンダの提案により HTML の後継バージョンである HTML5 の策定が進められるようになった。

また、ウェブは幅広く利用されるインフラであるだけに、メタデータとして付与される意味は、広く共通して適用できる意味体系、つまりオントロジに即したものである必要がある。あらゆる利用者、あらゆる場面で広く用いられる共通の意味体系を構築するのは難しい。また、特定分野に適用するためのオントロジであっても、その構築には多大な労力を要する。そのため、厳密なオントロジを採用する、いわゆる「大文字の Semantic Web」はなかなか普及してこなかった。

情報量は増したものの、それらのデータは人間が見たり、読んだりするためのフォーマット(Human Readable)により構成されたドキュメントであり、ウェブサイトを表示する計算機自身がそのデータの意味を解釈・処理できる形式(Machine Readable)ではないという状況であった。

2.3 User Generated Metadata と意図せざる協働

厳密なオントロジに基づいたセマンティックウェブや関連する標準規格の普及が進まない一方、ウェブ上のメタデータ生成・共有・利用は少しずつ広がりを見せてきた。その最大の要因は User Generated Contents である。2000 年代半ば以降ブログ、SNS、写真共有サイト、動画共有サイトなど、利用者によるコンテンツの気軽な共有を可能とするサービスが数多く登場し、個人が接するコンテンツ数が著しく増大することとなる。その結果、利用者がウェブ上のコンテンツを整理したり、利用者(投稿者、閲覧者)間によるコンテンツ探索の相互補助が可能となったりするツールに対する需要が大きくなってきた。そうしたニーズに対応するために登場したのが、「タグ」と呼ばれる利用者によって生成・付与されるメタデータである。

YouTube、Flickr といった、テキストデータをもともと含まない映像・画像コンテンツ共有サイトにおいては、投稿者によってコンテンツを注釈するキーワードをメタデータとして付与することが一般的となった。また、ウェブドキュメントをアーカイビングし、各ページに分類項目としてメタデータを付与するというソーシャルブックマークと呼ばれるサービスも登場した。つまり、User Generated Metadata が生成・流通・活用されるようになったのである。そういったサービスにおいて、利用者は自らの利便性のためにメタデータを生成・公開することにより、意図せずに他者の情報探索活動を支援するという「意図せざる協働」が成立することとなった。[3][4]

3. Linked Open Data とオープンガバメント

3.1 Government 2.0 とオープンガバメント

リーマンショック以降の世界的な経済の低迷、その上日本は人口減少トレンドに突入し、これまで以上に低コストで効率的に公共福祉、公共財を創出、供給していく方策を開発するという需要は高まるばかりである。そのため、正しく効率的な政策立案と執行がますます求められている。

その一方、価値観の多様化や経済状況の変化が加速度を増す中で、正しい政策立案・執行のためにはデータが不可欠である。これまで公共の利益に資するデータは、公的機関が収集、分析、公開してきたが、既に述べたとおりパブリックセクターにも、コスト削減要求が強まっている。ITC 技術は業務の効率化に貢献するものの、そういった作業を担う人的資源が限られている以上、細分化した需要に対応するだけのサービスを提供するには限界がある。多様な価値観をもつ成員によって成り立つ現代社会において、政策立案・執行の効率化、透明化とダイバーシティ対応が強く求められる状況となっている。

このような状況下で、Government 2.0 というムーブメントが起こりつつある。Government 2.0 とは、00 年代後半に流行語となった Web 2.0 のトレンドを行政にまで拡大させようとしたものである。Web 2.0 は、利用者自らがコンテンツを生成・公開・共有させる User Generated Contents が普及することにより成立したが、Government 2.0 では、行

政の公的な機能を民間部門でも分担するとともに、行政が有する情報も公開することを志向する。

Government 2.0 を推し進めることで、必然的に行政活動において必要なデータ、ならびに政策立案に必要なデータの開示が促進されることとなる。また、これまで公的セクターに独占されてきた公共財の創出といった機能も、民間部門、特に直接の受益者たる住民コミュニティが参画するようになる。情報、施策立案と施行が、公的セクターのみならず、広くオープンに行われる形態はオープンガバメントと呼ばれる。オープンガバメントは Government 2.0 のある種の到達点と言えるだろう。

3.2 米国・英国政府によるオープンガバメントの試み

オープンガバメントは既に政府レベルにおいて実践されている。米国のオバマ大統領は就任直後の 2009 年 1 月に「透明性」「国民参加」「官民の協業」という 3 原則を掲げ、オープンガバメントの実現を目指す覚書に署名している。また同年 6 月には、Data.gov を開設し、情報公開を推進している。

一方英国でも政府が 2009 年 6 月に、公的データの公開を推進し、政府の情報公開サイトである data.gov.uk のアドバイザーに W3C のディレクターである Sir. Tim Berners-Lee を任命するという声明を発表している。

米国、英国がほぼ同時に運用を開始した情報公開サイトにおいて採用されているのが Linked Open Data というコンセプトである。



図 1 米国の情報公開サイト：Data.gov



図 2 英国の情報公開サイト：Data.gov.uk

3.3 Linked Open Data とは

ウェブ上にあるドキュメントは計算機上で取り扱われる。また、ウェブは本質的にデータベースとしての性格を持っているとするならば、コンテンツを Human Readable だけでなく、Machine Readable にすればよい。そのようなコンセプトが Web of Data である。情報をもつ意味合いを正しく計算機処理するためには、人間が読み取るデータに計算機が処理するためのデータ、つまりメタデータを付与しなければならない。それが拡張されたウェブであるセマンティックウェブのコンセプトである。

このような状況を解決するために提唱されているのが Linked Data の概念である。Linked Data では、オントロジの構築については別で検討し、Machine Readable なデータの共有と利用を促進することを進めていくことを前提としているのが特徴である。Tim Berners-Lee は次のような原則を定めている[5]。

- 1)あらゆる事象に対して URI を付与すること
- 2)HTTP 経由で URI を参照できること
- 3)URI を参照した際は情報が閲覧できること。また、データは RDF や SPARQL 等標準化された技術で利用できるようにすること



図 3 [5]に掲出されている画像. Linked Data の対応状況に応じた
ランクの基準が示されている.

4)他の URI へのリンクを含めること.

特に 4 つめの原則により, Linked Data は基本的に他のデータリソースへのリンクを有している. 関連する情報を自由に抽出できることを意味する. この Linked Data による情報共有や利用を推進する活動を Linked Open Data(LOD) という. これにより, 分散して生成されたデータがリンクすることにより, 必要な関連情報を自由に抽出できるとともに, そのデータを処理・加工することも容易となる. 近年, 公的機関の情報開示が進んでいるといっても, 統計データを収録した PDF や HTML 形式のファイルを通じて行われていることも多い. この場合, そのデータを改めて加工するためには, 二次加工する必要がある. また, 膨大な量の生データをそのまま公開することは難しく, 分析・加工されたものだけが公開される傾向となる. しかし, 計算機処理可能な LOD 形式でデータが公開されることにより, データの生成・分析・加工全てのプロセスが公開・共有されることとなる.

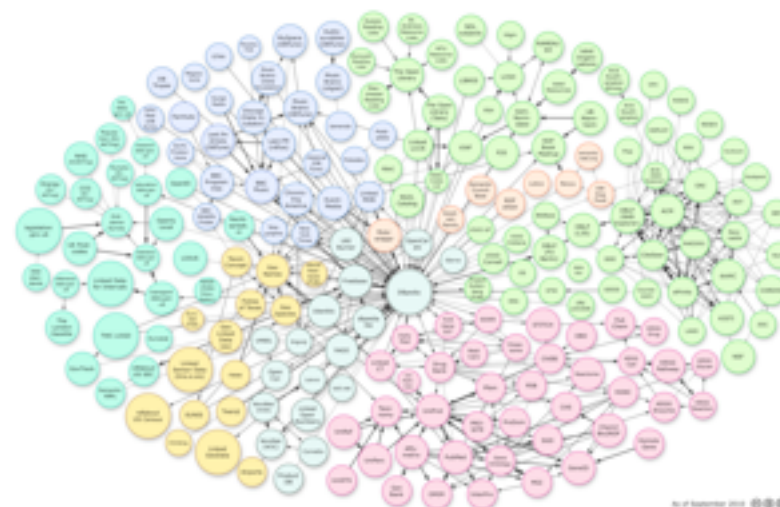


図 4 Linked Data 概念図[6]

4. 国内での Linked Open Data プロジェクト

4.1 公的機関による試み

国内でも LOD によるデータ提供の試みは始まっている. 国立国会図書館は国立国会図書館件名標目表, つまり図書分類項目を LOD で採用されているデータ形式, RDF 形式で提供している. また, 理化学研究所のサイネス(SciNetS.org)では, バイオ系を中心に 100 個以上のデータベースが RDF 形式で公開されている.

もう 1 つは国立情報学研究所が進める LODAC(Linked Open Data for Academia)である. このプロジェクトは, 現在のウェブにおける情報流通を「利用・創造 → 公開 → 収集・蓄積 → 共有 → 利用・創造...」の循環型へと移行することが必要であると[7], その実現を目指して博物館・美術館の収蔵品を中心とした情報の LOD 化・公開とともに, LOD 流通プラットフォームの開発を推進している[8].

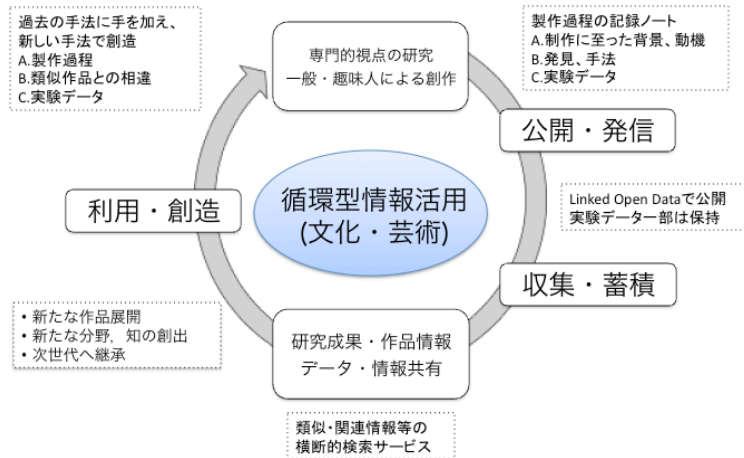


図 5 循環型情報活用[5]

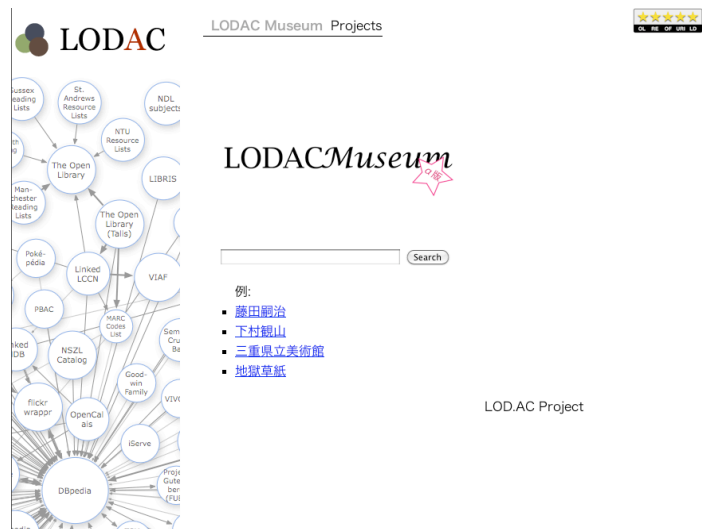


図 6 LOD.AC ウェブサイト

4.2 LinkedData.jp

また、エンジニアコミュニティの自主的な活動としては、LinkedData.jp がある。現時点では、Google Group を活用したメーリングリストの運営等を通じた情報交換が主な活動内容であるが、オープンガバメントの概念である利用者による多様なデータの活用が実現されるためには、広くエンジニアによるデータ活用のためのアプリケーションが開発、公開されることが必須であり、このようなエンジニアコミュニティ活動は重要な役割を果たすと考えられる。

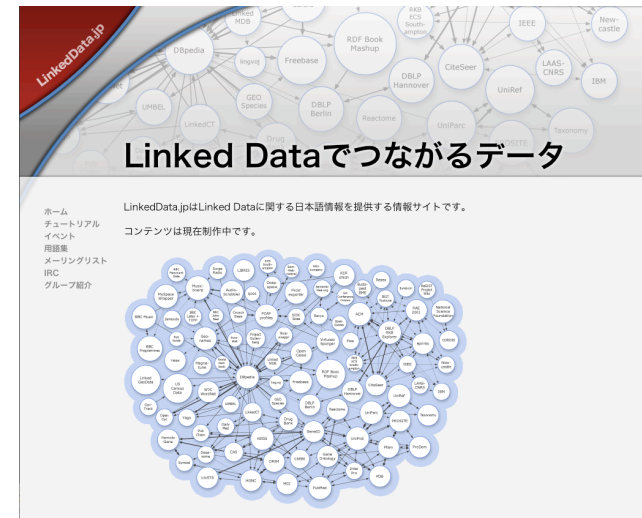


図 7 LinkedData.jp ウェブサイト

5. 横浜 LOD プロジェクト

5.1 プロジェクトの概要

横浜 LOD プロジェクトは地域情報の流通基盤として LOD が有効であるとの考えのもと、地域内の情報コンテンツホルダーと連携しながら地域情報の LOD 化とその活用を推進するプロジェクトとして 2010 年 6 月にスタートした。

このプロジェクトには公益法人横浜芸術文化振興財団と財団法人横浜観光コンベンションコンピューラーが参加しており、それぞれの持つ情報コンテンツの LOD 化に向けて作業を進めている。

具体的に対象となる情報コンテンツは、イベント情報、観光施設情報、アーティスト情報、美術館や博物館の収蔵品情報であり、こうした情報コンテンツが LOD 化されることでさ

さまざまな地域情報サービスが生まれてくることが期待される。

5.2 プロジェクト発足までの経緯

5.2.1 横浜開港 150 周年記念事業

2009 年、横浜は開港 150 周年を迎え、横浜市内では多くの主体によって数多くの事業が展開され、ウェブも積極的に活用された。

主なウェブプロジェクトとして、地域 SNS「ハマっち！SNS」、市民参加型写真アーカイブ「横浜開港 150 周年みんなでつくる横濱写真アルバム」などが展開されたが、とくに地域情報の流通というテーマに意欲的に取り組んだ事業が、財団法人横浜開港 150 周年協会による「横浜開港 150 周年市民で作るイベントナビ」（以降、イベントナビ）であった。

イベントナビでは横浜市や同財団主催のイベント以外にも、広く横浜開港 150 周年に関するイベント情報を収集して広報することを目的に開発された。

イベント情報は、ユーザ登録を行えば誰でも投稿することができた他、磯子区が開発した「イベントお届け便」（以降、お届け便）との連携も実現された。お届け便は磯子区の他、現在では都筑区など横浜市内の多くの区で利用されており、主に、各区のウェブサイト内のイベント情報ページを更新する目的で利用されている。

イベントナビとお届け便の連携は、お届け便がもつイベントデータのうち、開港 150 周年関連イベントを示すフラッグがあるデータを CSV 形式でエクスポートし、イベントナビがそのデータをインポートする形で実現した。

その他、同事業では横浜観光コンベンションビューロー等、横浜市内の外郭団体とのデータ連携も検討されたが、実現はしなかった。

5.2.2 イベント情報の流通へのチャレンジ

イベントナビではイベント情報の流通のために、外部サービスとの連携にも積極的に取り組んだ。Google カレンダーや Yahoo! カレンダーなどのオンラインサービス、Outlook や iCal 等ローカルアプリケーションへのイベント情報の取り込みをワンクリックで実現する機能を実装した他、検索クエリーに対して RSSiCal 形式等のデータをレスポンスする API を実装した。

この API の具体的な利用ケースとして、富士ゼロックス株式会社と共同で行ったイベント情報のオンデマンドプリンティングサービス、日産自動車株式会社と共同で行った、カーナビゲーションシステムへのイベント情報の提供サービスなどが実現した。

富士ゼロックス株式会社と共同で行ったオンデマンドプリンティングサービスは、Y150 の会場などに合計 3 か所に設置されたプリント端末を利用者が操作（プリントボタンを押すだけ、無料）することで、端末が設置されている会場でその時間以降に



図 8 イベントナビトップページ

参加可能なイベント情報一覧が表面に、裏面に会場地図をそれぞれ出力するサービスである。

システム側では、プリント端末がイベントナビサーバーの API を決められたインターバルで呼び出して、データをキャッシュしておく。そして、利用者の操作した日時にあったキャッシュデータを出力するという仕様で実装がされた。

日産自動車株式会社との共同サービスでは、日産自動車株式会社の開発するインターネット接続機能を持ったカーナビゲーションシステムである「カーウイングス」に対してイベントナビのイベント情報を配信した。カーウイングスでは「情報チャンネル」という、インターネット上に配信されている様々な情報に、走行中でも簡単な操作でアクセスできるサービスを展開しており、これに対してイベントナビのイベント情報を RSS 形式で配信するという形でサービスを実現した。

システム側では、あらかじめ設定されたテーマ（食、音楽、イルミネーション、など）ごとに、イベントナビの持つイベント情報に付与された「タグ」情報を対応させ、それぞれ RSS 出力できるように実装がされた。そして、カーウイングス端末は利用者からのテーマリクエストに応じてイベントナビの API にリクエスト、レスポンスをうけて、端末画面にイベント情報を表示することで実現した。



図 9 横浜開港 150 周年市民で作るイベントナビ
 出典：横浜開港 150 周年市民で作るイベントナビウェブサイトより



図 10 カーナビへのイベントナビ上方の配信
 出典：財団法人横浜開港 150 周年協会プレスリリースより

5.3 見えてきた課題と情報の LOD 化に対する期待

イベントナビでは横浜市内の情報コンテンツを持つ主体より、イベント情報を包括的にかつ、効率的に収集することが目的の一つであったが、情報コンテンツを持つ各主体のシステムに API が整備されていなかったり、各イベント情報のデータフォーマットがバラバラであったりなど、非常に困難であった。横浜 LOD プロジェクトではこの問題を解決するために LOD のコンセプトを導入した。

地域内で LOD が普及することで、たとえば、これまで実現することがむずかしかった、公益法人横浜芸術文化振興財団と財団法人横浜観光コンベンションビューローが持つイベント情報や施設情報を横断検索したり、地図上に表示したりするアプリケーションやサービスの登場が期待できる。

また、LOD の生成を地域の中で行うことにより、情報の公開、共有、収集、利用、というサイ

クルをより細かい粒度で展開することが可能になるのではないかと考えられる。その主な理由として、地域という顔の見える範囲においては、共用できる語彙の策定を共同で行ったり、また、互いに積極的にデータにリンクしたり、ウェブページを展開する際にもデータ自体を互いに利用しあうことなどがやりやすいということがあげられる。

こうした、アプリケーションやサービスは必ずしも情報コンテンツホルダーが開発提供する必要はなく、NPO や企業、個人が自由に開発してサービスを提供することができるようになる。地域で生活する、または、地域を訪れる人々の情報に対するニーズが多様化する中、当該団体においても一意的な情報提供では不十分と考えており、地域情報サービスの質の向上といった観点からも LOD に対する期待は大きい。

5.4 横浜 LOD プロジェクト今後の展開

横浜 LOD プロジェクトでは、横浜市の GIS 分野や統計分野における行政情報の LOD 化も働きかけをはじめている。オープンガバメントの文脈においても LOD は重要な要素であり、アメリカ政府やイギリス政府ではすでにその整備が進んでいる。日本政府のみならず、地方行政においても LOD への取り組みが期待される。

また、今後 LOD 普及の鍵となるのが、その利活用であるが、企業や NPO、個人による積極的なデータの利用と価値あるサービスの創出が期待される。

6. LOD によるボトムアップ型オープンガバメントの可能性

オープンガバメントは、単に行政がデータを公開するだけではなく、住民が公共財の創出に参加することが重要である。参加とは、ただ労力を提供するだけではなく、どのような財・サービスが求められているかを自ら考え、提案することも含まれる。これまで行政が収集・整備・管理することが難しかった分散して存在するデータの共有と、多面的な視点からの分析、きめ細かなサービスの提供が求められる。

ニーズの多様性に対応するには、逆説的であるが共有されるデータフォーマットは共通であるのが望ましい。そのため、地域を構成する各主体が共通のフォーマットでデータを生成・共有すること、更には関連するデータ同士はリンクされていることが重要となる。

この LOD というコンセプトは、自立・分散・協調というウェブの特徴をそのまま引き継ぐものである。それだけに、トップダウンではなく、利用者・受益者自らがデータを生成・活用していくことが望ましい。また、住民の視点に近い、地域行政従事者が住民の活動をコーディネートしていくことが望ましい。その点で、横浜 LOD プロジェクトは住民主導のボトムアップ型オープンガバメントの試みとして、注目に値する。また、このような動きが広がっていく上で、LinkedData.jp のようなエンジニアコミュニティによるアプリケーション開発は欠かせない。

今回取り上げた Linked Open Data によるボトムアップ型オープンガバメントの試みは、

循環型情報活用による地域活性化の実現に大きく寄与するものと考えられる。

参考文献

- 1)Berners-Lee, T. and Fischetti, M. Weaving the Web: The original design and ultimate destiny of the World Wide Web by its inventor. Orion Business Books.1999. (邦訳：高橋徹『Webの創成 — World Wide Webはいかにして生まれどこに向かうのか』毎日コミュニケーションズ, 2001年.)
- 2)深見嘉明. “POSからクラウド データ処理の歩みをたどる,” エコノミスト増刊 巨大データ経済の勝者, 2010年8月9日号増刊, pp.43-45, 毎日新聞社, 2010.
- 3)深見嘉明, 國領二郎. “意図せざる協働—ソーシャルブックマークにおけるボトムアップメタデータ生成による情報共有—,” 情報社会学会誌, Vol. 2, No. 2, pp. 6-19, 2007.
- 4)深見嘉明『ウェブは菩薩である ～メタデータが世界を変える』NTT出版, 2008.
- 5)Berners-Lee, T. Linked data. <http://www.w3.org/DesignIssues/LinkedData.html>, 2007.
- 6)Cyganiak, R. The Linking Open Data cloud diagram. <http://richard.cyganiak.de/2007/10/lod/> updated 2010年9月22日アップデート版
- 7)武田英明. “日本における Linked Data の現状と普及に向けた課題,”情報処理, Vol.53, No.3, 2011.(予定)
- 8)嘉村他 Linked Open Data による多様なミュージアム情報の統合, 人文科学とコンピュータシンポジウム じんもんこん 2010 予稿集, 2010.