

Linked Open Data による博物館情報および地域情報の連携活用

松村 冬子¹ 小林 巖生² 嘉村 哲郎^{3,4} 加藤 文彦⁵
高橋 徹⁶ 上田 洋⁶ 大向 一輝¹ 武田 英明¹

¹ 国立情報学研究所 ² Open Community Data Initiative ³ 総合研究大学院大学 ⁴ 東京藝術大学
⁵ 情報・システム研究機構 新領域融合研究センター ⁶ 株式会社 ATR-Promotions

近年では、米国や英国をはじめとした世界各国で、標準化されたデータ形式である Linked Open Data (LOD) に基づいたデータ公開が行われている。LOD はデータのオープン化、分野内でのデータの共有、そして分野を横断したデータの共有を促進するという特徴を持っている。この特徴を考慮して、国内の博物館データベースの統合を行う LODAC Museum プロジェクトにおいては博物館情報を、横浜 LOD プロジェクトでは地域で行われるイベントや施設などの地域情報を LOD とし保有している。本稿では、博物館情報と地域情報という異なる分野の LOD を連携活用することにより構築した、横浜のアート情報を提供する Yokohama Art Spot と呼ばれる Web アプリケーションについて報告し、異なる LOD の連携活用によりもたらされる効果について議論した。

Collaboration between Linked Open Data of Museum Information and Regional Information

Fuyuko Matsumura¹ Iwao Kobayashi² Tetsuro Kamura^{3,4} Fumihiro Kato⁵
Toru Takahashi⁶ Hiroshi Ueda⁶ Ikki Ohmukai¹ Hideaki Takeda¹

¹ National Institute of Informatics ² Open Community Data Initiative
³ Graduate University for Advanced Studies ⁴ Tokyo University of the Arts
⁵ Research Organization of Information and Systems ⁶ ATR Promotions, Inc.

In recent years, a lot of data is published as Linked Open Data (LOD) around the world. LOD is a standardized data format that makes public data and enables data interoperability in a particular field and also beyond different fields. Yokohama LOD project has a LOD dataset of regional information such as the data of facilities and events related to arts in Yokohama, and LODAC Museum project tries to construct a LOD dataset of museum collection in Japan. In this research, a web application, “Yokohama Art Spot”, is developed using these LOD datasets of regional information and museum information, and it shows facilities and events related to arts in the city of Yokohama. In this paper, we discussed what is enabled by this application with the collaboration of two different LOD.

1. はじめに

近年、世界各国で行政の透明化や市民の行政への参加の促進のため、政府や自治体の持つ情報を Web で公開する「オープンガバメント」の動きが盛んになっている。特に米国をはじめとした欧米諸国では、Linked Open Data (LOD)[1]と呼ばれる計算機処理が可能な形式によるデータの公開が始まっている。LOD を用いるメリットとしては、データのオープン化の促進、分野内のデータ共有、分野を超えたデータ共有の 3 点の実現が挙げられる[2]。実際に英国では、市民や民間企業が LOD の形式で公開された政府の情報¹を利用して、スマートフォンのアプリケーションや Web アプリケーションを開発し、公開している[3]。一方、日本国内でもオー

ブンガバメントを促進するべく、総務省や経済産業省が様々なデータを公開しているものの、現在ほどのデータも LOD として公開されておらず、その多くが PDF ファイルや Excel ファイルの形式であり、その後機械的に処理することが難しい²³。これに対して、著者らが活動を行っている LODAC (Linked Open Data for Academia) プロジェクト[4]⁴では、広く学術に関係する情報を共有し、再利用していくことを目指して、国内の博物館情報、生物種情報、地理情報などを LOD によって表現し、公開する取り組みを行っている。現在、著者らは LOD を構造化された再利用可能なデータとして公開しているが、先に述べたような LOD のメリットを活

² 総務省 e-Stat <http://e-stat.go.jp>

³ 経済産業省 オープンガバメントラボ データボックス
<http://databox.openlabs.go.jp>

⁴ LODAC プロジェクト <http://lod.ac>

¹ data.gov.uk <http://data.gov.uk>

かすような実用的なアプリケーションは国内でも未だ開発されていない。これより本研究では、異なる分野の複数の LOD を連携して活用するアプリケーションを構築することで、期待される効果を検証することを研究目的とする。本稿においては、LODAC プロジェクトにおける博物館情報の LOD である LODAC Museum と、横浜 LOD プロジェクト⁵による横浜のアートに関する地域情報の LOD を用いて構築した Yokohama Art Spot という Web アプリケーションについて紹介し、2つの LOD の相乗効果について議論する。

以下に論文の構成を示す。第2章において LOD による情報公開のメリットについて詳細を述べる。第3章では、博物館情報の LOD である LODAC Museum の概要と、博物館情報の分野における LOD 化のメリットについて述べる。第4章では、横浜における地域情報発信の現状と、横浜 LOD プロジェクトによるアート関連情報の LOD 化について述べる。第5章においては、構築した Yokohama Art Spot⁶の概要および LOD により期待される効果を検証する。最後に第6章において結論と今後の課題を述べる。

2. Linked Open Data (LOD) とは

2.1 LOD の 4 原則

Linked Open Data (LOD)とは、個々の事物に対してグローバルな ID である URI を付与し、それらの事物間の関係を任意の種類のリンクとして表現したデータの集合を指す[5]。LOD によって互いに繋がったデータ群は、データの Web (Web of Data) と呼ばれるグローバルなデータ空間を構成している。Tim Berners-Lee はこのグローバルなデータ空間の一部としてデータを公開する際の原則を以下のとおり挙げている[6]。

1. あらゆる事象に対して URI (Uniform Resource Identifier) を付与する。
2. HTTP 経由で URI を参照可能とする。
3. URI を参照した際は情報が閲覧可能とする。また、データは RDF (Resource Description Framework) や SPARQL など標準化された技術でアクセスできるようにする。
4. 他の URI へのリンクを含める。

⁵ 横浜 LOD プロジェクト <http://scholix.com/ocdi/>

⁶ Yokohama Art Spot <http://lod.ac/apps/yas>

2.2 LOD のメリット

2.1 節に示した性質を踏まえて、LOD をデータ表現形式として採用するメリットについて述べる。LOD のメリットは「オープンデータ化の促進」「分野内データ共有の実現」「分野を超えたデータ共有の実現」の3段階に分けることができる[2]。以降、その詳細について述べる。

2.2.1 オープンデータ化の促進

Web で公開されているオープンデータは HTML 文書や PDF 文書、Excel 文書など様々な形式で表現されている。これらの形式は人間にとっては可読性が高いが、プログラムで機械的に処理することは容易ではない。

これに対して、LOD では、データ中のすべての事物に URI で識別子を付与してリソースとして表現する。そのため、個々の事物をグローバルかつ一意に指すこと、HTTP のような標準的な方法でアクセスすることが可能である。また LOD では、データは RDF と呼ばれる標準的な形式で表現され、そこから必要な部分だけを取り出す際は標準化された問合せ言語である SPARQL を用いる。これより、前述の HTML 文書などと比較すると、LOD によるデータ公開は、よりアクセスが容易であり、再利用性においても優れている。

2.2.2 分野内データ共有の実現

1つの分野であっても、複数のサイトが各々の保持しているデータを公開することは一般的である。このとき、各サイト間で共通の事物のデータを持っている場合でも、現状の HTML 文書などの表現形式では、それらが同一の事物を指すということは表現できない。例えば、博物館の分野では、複数の美術館の Web サイトに同じアーティストの略歴に関するデータが掲載される場合がある。HTML 文書では、あるサイトと別のサイトで掲載されていたアーティストが同じ人物であるかどうかは機械的に可読な形式で表現することができない。

LOD では、データサイト同士が各々の持つリソース間に意味を持たせたリンクを張ることで、データの関係性を表現することができる。これにより、1つのデータサイトに全ての情報を集約せずとも、分散しているデータサイトを統合的に利用することが可能である。

2.2.3 分野を超えたデータ共有の実現

LOD によるデータ公開は分野を超えたデータ共有にも有用である。一般に、他分野のデータはアクセス手段やデータ構造が分からず、利用

が難しい。LOD ならば、アクセス手段は一定であり、公開されたスキーマはオントロジーとして表現できるので、他分野のスキーマを理解して利用することが容易になると期待できる。

また、分野を超えた異なる LOD 同士をリンクさせることで、そこから引き出せる情報がより豊かになると考えられる。

3. LODAC Museum による博物館情報の統合

3.1 LODAC Museum の概要

LODAC Museum は情報・システム研究機構新領域融合センターの「学術リソースのためのオープン・ソーシャル・セマンティック Web 基盤」(通称 LODAC) プロジェクトにおいて構築されている博物館情報の LOD である。国内の博物館では、文化遺産オンライン⁷などの一部の集約サイトを除いて、分散的に個々の情報を公開している。つまり、例えば横山大観の作品を所蔵している博物館のリストを取得することは現状では不可能である。

そこで LODAC プロジェクトでは、2.2.2 節で述べた分野内におけるデータ共有の促進のメリットを活かし、博物館情報の LOD 化に取り組み始めた。本来ならば、博物館自体が各々の保有するデータを LOD 化して公開し、それを連携させるのが望ましいが、現在は LODAC プロジェクトで各博物館の Web サイトから所蔵品リストなどのデータをスクレイピングし、LOD の形式に変換して公開している。

3.2 データの表現と実装

図 1 に LODAC Museum のデータの一部分を示す。LODAC Museum では、主に作品、アーティスト、所蔵館の 3 つのリソースが用いられる。それらのリソースには、他のリソースへのリンク、値の入ったプロパティが含まれており、これらの組合せによってデータが表現される。例えば、「生々流転」という作品を描いたアーティストは「横山大観」であるというデータは、図 1 のように `<http://lod.ac/id/20811><http://purl.org/dc/terms/creator><http://lod.ac/id/1767>` という組合せで表現される。

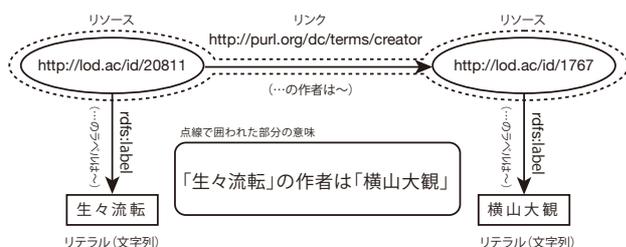


図 1 LODAC Museum のデータ表現

⁷ 文化遺産オンライン <http://bunka.nii.ac.jp/>

また、複数の情報源に存在する同一の事物のデータを統合する場合が想定される。このような場合には、その事物の統合 ID のリソースと、情報源ごとの ID のリソースの生成し、2 種類の ID を付与した。また、統合 ID と情報源ごとの ID はそれぞれ参照関係のリンクで紐づけられている。例えば、横山大観の情報を博物館 A と博物館 B から得られた場合、統合 ID のリソースと、それぞれの博物館から得た情報源 ID の 2 つのリソースが生成される。このアーティスト情報の統合については、信頼できるシソーラスである日本美術シソーラスを利用した[7]。

なお、LODAC Museum の実装においては、RDF データベースには OWLIM-SE を、RDF への SPARQL の問合せを行うためのインタフェースである SPARQL エンドポイントを、LODAC Museum のサイトにおいては、テキスト検索が可能な簡易ブラウジングインタフェースを設置している。

4. 横浜 LOD プロジェクトによる地域情報の発信

4.1 横浜における地域情報発信の現状

昨今では、行政だけでなく、市民が主体となって企画や運営を行う地域のイベントが増えてきている。また同時に、Twitter、ソーシャルネットワークワーキングサービス(SNS)、ブログなど一般ユーザが自ら情報発信を容易に行うことができるツールが数多く登場してきており、地域で活動している市民や団体が活動をより多くの人に知ってもらい、活性化させるべく、それらのツールで情報発信を行っている[8][9]。しかし、実際には個々の市民や団体が発する情報が一般の市民に目に留まることは難しい。そのため、現在は自治体などにおいて地域のイベント情報に関するポータルサイトを設置し、それらの情報を集約して配信するなどしている。

例えば、公益財団法人横浜市芸術文化振興財団は、ヨコハマ・アートナビ⁸という Web サイトにおいてアートに関するイベント情報やニュース等を配信している。また、公益財団法人横浜観光コンベンション・ビューローでは、横浜市観光情報公式サイト⁹において、横浜の観光スポット、イベント情報、グルメスポットなどを紹介している。その他にも横浜市民や横浜を訪れる人に向けて、スポーツ情報に特化したハマスポどっとコム¹⁰、神奈川県による総合文化芸術情報サイトであるかな@ (アット)¹¹、横浜市内の各区のイベント情報を配信するお届け便などの Web サイトが提供されている。

⁸ ヨコハマ・アートナビ <http://www.yaf.or.jp/>

⁹ 横浜市観光情報公式サイト

<http://www.welcome.city.yokohama.jp>

¹⁰ ハマスポどっとコム <http://www.hamaspo.com>

¹¹ かな@ <http://www.kanagawa-at.info>

4.2 横浜 LOD プロジェクトによるアートに関する地域情報の LOD 化

横浜には様々な情報発信主体があり、それぞれが独自に Web を通じて、個別に情報を発信しているが、同じ地域における情報なので、施設の情報や出演者など、共通するものも多い。そこで、一般市民や市民活動団体、民間企業、行政、あらゆる主体が地域において発信する情報を地域資源と捉え、それらの情報の公開、共有、利活用を LOD により推進することを目指して、横浜 LOD プロジェクトが発足した。

現在、横浜 LOD プロジェクトにおいて LOD 化が進められているデータのひとつは、先に述べた横浜市芸術文化振興財団がヨコハマ・アートナビにおいて配信している横浜市内のアートに関連したイベント情報である。このイベント情報はイベントを実施する団体や施設などから横浜市芸術文化振興財団が収集し、公開用のフォーマットに入力して配信している。現在は、このイベント情報はテキストのみで構成されているが、イベント、施設、アーティスト、最寄り駅などの情報を LOD でリソース化し、その関係をリンクで表現することが出来れば、他の情報発信主体とそれらの情報資源を共有することができ、蓄積された情報の活用が可能になる。

このように、各団体で全ての情報を一から作らずとも、リソースを共有もしくはリンクすることで、豊かな表現による情報発信が可能となることから、様々な規模の情報発信主体が混在する地域情報の分野に、LOD による情報公開および共有は適していると考えられる。

5. Yokohama Art Spot

本章では、本研究で構築した Web アプリケーションである Yokohama Art Spot について述べる。Yokohama Art Spot は、博物館情報の LOD である LODAC Museum と、横浜 LOD プロジェクトにより構築されているアートに関する地域情報の LOD の 2 つの異なる LOD を連携活用したアプリケーションである。

5.1 システム構成

本アプリケーションのシステム構成を図 2 に示す。まず、ユーザは、一般ユーザと運営者の 2 種類である。一般ユーザは本アプリケーションを利用し、運営者は本アプリケーションで用いる LOD の追加や更新などの管理を行う。一般ユーザがブラウザでアクセスして利用するインタフェース部分は HTML と JavaScript で構成されており、ユーザが施設詳細ページを開くなど何かアクションを起こすと、各 LOD の

SPARQL エンドポイントに対して、RDF の問合せ言語である SPARQL により、必要なデータの問合せが行なわれる。SPARQL エンドポイントから返却されたデータを用いて、本アプリケーションはユーザに新しい情報を提示する。

5.2 本アプリケーション持つ主な機能

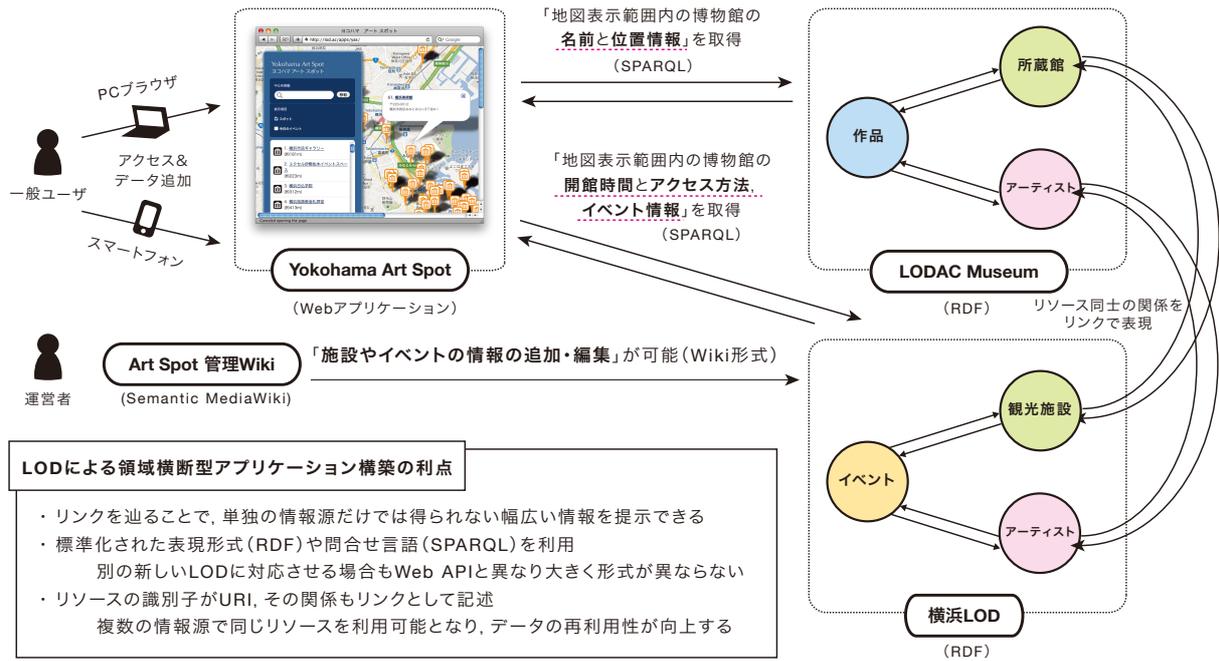
本アプリケーションは、一般ユーザ向けの全体マップページと施設詳細ページの 2 種類のページ、運営者向けのデータ管理ページから構成されている。前者では地域に、後者では施設に主眼をおいて、横浜市内のアート関連情報を提供する。双方ともにページのテンプレートに対して、SPARQL で取得してきたデータを埋め込むことで情報提示を実現している。

(1) 全体マップページ

本アプリケーションにアクセスすると、まず図 3 に示す全体マップページが表示される。この画面では、横浜市の地図上にアート関連施設とその日に開催されているイベントを表示することで、特定の地域でどのようなアート体験が可能かを提示する。全体マップにおける施設の表示では、SPARQL で表示されている範囲の緯度および経度の情報を用いて、LODAC Museum からその範囲に含まれている美術館等の施設の名称、住所、緯度および経度の情報を抽出している。イベントの表示に関しては、前述の方法で得られた表示範囲に含まれている施設の ID をもとに、SPARQL でそれらの施設が開催場所となっているイベントの情報を横浜 LOD から取得し、開催場所となる施設の位置に表示する。

(2) 施設詳細ページ

全体マップの各施設のマーカーをクリックすることで、図 4 に示す施設詳細ページに移動する。このページでは、各施設の開館時間やアクセスなどの基本的な情報と、イベント情報、収蔵品情報（一部の施設のみ）を提示する。施設詳細ページにおける施設情報の提示においては、LODAC Museum に含まれているその施設の名称、アクセス方法、開館日、緯度および経度などの基本的な情報を、施設の ID により SPARQL で取得している。また、横浜 LOD において、この施設を開催場所としてリンクしているイベントを SPARQL で獲得する。収蔵品についてもイベントと同様に、その施設が所蔵館とされている作品の情報を LODAC Museum から取得する。



LODによる領域横断型アプリケーション構築の利点

- ・リンクを辿ることで、単独の情報源だけでは得られない幅広い情報を提示できる
- ・標準化された表現形式 (RDF) や問合せ言語 (SPARQL) を利用
別の新しいLODに対応させる場合もWeb APIと異なり大きく形式が異なる
- ・リソースの識別子がURI, その関係もリンクとして記述
複数の情報源で同じリソースを利用可能となり、データの再利用性が向上する

図2 Yokohama Art Spot システム構成



図3 全体マップページ

(3) データ管理ページ

LODAC Museum の保持している施設情報や収蔵品情報などの情報は変更が少ないが、横浜LOD で扱うイベント情報は追加や変更が多いデータである。現在、同じデータをヨコハマ・アートナビのアートカレンダーに配信しているが、これらの更新作業などは横浜市芸術文化振興財団のスタッフによって全て手作業で行われている。しかし、LOD の場合、データを RDF で記述する必要があり、RDF の形式に詳しくないスタッフが手作業で入力することは難しい。これに対して、本アプリケーションでは、Semantic MediaWiki[10]と呼ばれる Wiki において、イベントやアーティストなどの入力フォームを用意し、スタッフが図5に示すようにフォームに必要な情報を記述して保存することで横浜 LOD となる RDF を生成できるようにしている。これにより、データの追加や修正が実現し、LOD によるデータの公開、収集、利用だけでなく、創造が加わり、情報の循環型活用が可能になる。

5.3 異なる2つのLODの連携による効果

これまでヨコハマ・アートナビのアートカレンダーで表示するために蓄積されてきたデータを LOD にしたことで、イベント情報に含まれる開催場所の施設の情報から、横浜 LOD では保有していない収蔵品の情報などを取得できた。これより、それぞれのLODを単独で利用するよりも、幅広い情報提供ができるという相乗効果が見込めることが分かった。



図4 施設詳細ページ (横浜美術館)

また、データベースに蓄積していた情報を Web API 経由で他のアプリケーションで使用できる場合であっても、API ごとに仕様が大きく異なるために、新たに情報源を加えるたびに大きな変更が必要となる。このことから、複数の外部の情報源からデータを利用することは、Web アプリケーション構築において開発コストがかかると言える。これに対して、LOD では RDF という標準化された形式でデータが記述され、同じく標準化された SPARQL の形式で問合せが行えることから、対応する LOD が増えても、アプリケーションの変更は少なく済む。このことから、特に行政、NPO 法人など様々な規模で情報発信が行われる地域情報の分野において、その相互の連携を助けるものと考えられる。

本アプリケーションにおける 2 つの LOD の連携について考察すると、LODAC Museum においては、施設、アーティスト、作品などの基本となる静的な情報を、横浜 LOD では、イベントや若手のアーティストなどの動的な情報を保有しており、それぞれ異なる性質の情報を補い合うことが可能になったと考えられる。

5.4 関連研究と今後の課題

関連研究として、日本語の Linked Data を自動的に生成し、モビリティサービスを構築する研究がある[11]。これに対して本アプリケーションでは、スタッフによる LOD の手動の追加および更新処理が含まれることから、Semantic Media Wiki によってエンドユーザに優しいインタフェースを備えた点が異なる。

現在は、リソース間のリンクを入力フォームで記述する際に URI を直接記述する必要があるが、今後は実運用に向けて、入力された文字列によって候補となるリソースをサジェストするなどデータ更新のユーザビリティ向上を図る。

また、株式会社 ATR-Promotions の協力により、LODAC プロジェクトで構築した地理情報の LOD に含まれる施設情報が、同社の iPhone の写真撮影アプリケーションの地図上に表示されるようになったことから、Yokohama Art Spot においても、エンドユーザに向けた情報の閲覧機能だけでなく、リソースに対するアノテ



図5 Semantic Media Wiki による横浜 LOD のデータ管理

ーションも行なえるスマートフォンアプリケーションを開発することを目指す。

さらに、LODAC Museum では全国の博物館の情報を収集していることから、他の地域においても本アプリケーションのようなサイトのパッケージ化を進めることで、地域情報の LOD 化と博物館情報の LOD の活用を促進したい。

6. おわりに

本研究では、博物館情報の LOD と地域情報の LOD の連携活用を実現するアプリケーションである Yokohama Art Spot を構築し、その期待される効果について議論した。期待される効果としては、2 つの LOD が相互の特徴を活用して補完し合うことで幅広い情報提供が可能になったこと、標準化された技術により新しい情報源の追加が容易になったことがあげられる。今後は Semantic MediaWiki による LOD の追加および更新支援についてユーザビリティの向上を図る。

参考文献

- [1] Bizer, C., Heath, T. and Berners-Lee, T.: Linked Data – The Story So Far, International Journal on Semantic Web and Information Systems (IJSWIS), 5(3), pp.1–22, 2009.
- [2] 武田英明, 嘉村哲郎, 加藤文彦, 大向一輝, 高橋徹, 上田洋: 日本における Linked Data の普及にむけて, 3E3-OS20-9, 人工知能学会第 25 回全国大会論文集, 2011.
- [3] 佐藤宏之, 飯塚京士, 三島和恵: オープンガバメントとオープンデータ, 特集 リンクするデータ, 情報処理, 52(3), pp.309–317, 2011.
- [4] Kamura, T., Takeda, H., Ohmukai, I., Kato, F., Takahashi, T., and Ueda, H.: Study Support and Integration of Cultural Information Resources with Linked Data, in Proc. of 2nd International Conference on Culture and Computing, pp.177–178, 2011.
- [5] 長野伸一, 萩野達也: 編集にあたって, 特集 リンクするデータ, 情報処理, Vol.52, No. 3, pp.282–292, 2011.
- [6] 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.
- [7] 福田博同, 五十殿利治: 美術シソーラスデータエース形成の諸問題, 情報管理, 40(9), pp.790–809, 1997.
- [8] 吉沢卓, 小林巖生: 横浜の地域情報とコミュニケーション, 電子情報通信学会技術研究報告, AI, 人工知能と知識処理, 108(325), pp.41–46, 2008.
- [9] 深見嘉明, 小林巖生, 嘉村哲郎, 加藤文彦, 大向一輝, 武田英明, 高橋徹, 上田洋: 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, 情報処理学会研究報告, DD, 2011-DD-79(1), pp.1–8, 2011.
- [10] Krötzsch, M., Vrandečić, D. and Völkel, M.: Semantic MediaWiki, Lecture Notes in Computer Science, 4273/2006, pp.935–942, 2006.
- [11] Iijima, C., Morita, T., Enomoto, Y. and Yamaguchi, T.: A Mobility Service based on Japanese Linked Data, 15th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, Springer LNCS6883, pp.246–255, 2011.