

自治体を横断した観光情報を地図上に統合表示する方法の検討

青木 亮祐[†] 井上 潮[‡]

^{† ‡} 東京電機大学工学部 〒120-8551 東京都足立区千住旭町 5

E-mail: [†] 15ec001@ms.dendai.ac.jp, [‡] inoueu@mail.dendai.ac.jp

あらまし 観光客誘致による地域の活性化を図るため、観光に関するオープンデータの公開やそれを用いたアプリケーション開発が全国の自治体で進められている。しかし、現在公開されている観光オープンデータは異なるフォーマットで提供されていたり、内容に過不足があったりするため、異なる自治体の観光情報を併用することが難しい。本研究では、自治体を横断して観光情報を取得し、地図上に統合表示する方法を検討した。具体的には、すべてのデータを **RDF** に変換することで自治体間のデータの連携を容易にした。また、地図上に表示する際に不足している項目のデータを **Web** 上から取得することにより、表示内容の充実を図った。これらの方法を用いて、静岡県東部地域の観光地図アプリケーションを試作し、複数の自治体を横断した観光情報を統合して表示できることを確認した。

キーワード 観光情報、オープンデータ、RDF、Web 地図

1. はじめに

スマートフォンやタブレット端末、SNS 等の普及を背景に、多種多様な情報を相互に連携させ新たな価値を生み出すことが期待されている。特に、国や自治体が保有する公共データが企業や住民に利用しやすい形で公開されることが求められている [1]。こうした背景から、日本政府は 2012 年 7 月に電子行政オープンデータ戦略を策定した。その後、2017 年 5 月にはオープンデータ基本方針を策定し、行政が保有するデータはすべてオープンデータとして公開することを原則とすることを明確化した。このような政府の取り組みもあり、2018 年 3 月には都道府県におけるオープンデータ取り組み率 100 % を達成した。しかし、市区町村におけるオープンデータの取り組み率は 2018 年 4 月時点で 19 % と依然として低いのが現状である [2]。

また、地域の活性化や 2020 年に開催される東京オリンピック・パラリンピック競技大会を控え、自治体は観光客誘致のために観光に関するオープンデータの公開やそれを用いたアプリケーションを開発するコンテストの開催といった取り組みを行っている。しかし、各自治体がそれぞれ独自に情報を公開しており、情報の種類や各情報の表記が統一されていないという課題がある [3]。また、名称や位置情報のデータはあるが、観光地の画像やホームページの URL といった内容が含まれていないという課題も指摘されている [4]。このように、現状では観光オープンデータの内容や自治体間を横断した利活用には課題がある。

そこで我々は、観光オープンデータや Wikipedia 日本語版の Linked Open Data(以下、LOD と略す)である

DBpedia Japanese^(注1)を用いて、自治体を横断した観光情報を取得し統合表示する地図アプリケーションを試作した。本稿では、その手法や遭遇した課題を述べる。

本稿の構成は次の通りである。2 章では観光に関するオープンデータを用いた既存研究をいくつか紹介する。3 章では我々の提案手法について述べる。4 章では提案手法に則ったアプリケーションの試作について述べる。5 章では評価とその考察を述べる。6 章では本稿のまとめと未解決の課題について述べる。

2. 関連研究

観光分野におけるオープンデータを用いた研究は近年活発に行われている。

まずは、観光分野のオープンデータの整備状況を調査した研究を紹介する。江崎ら [5]は、公開されている観光オープンデータの整備状況についての現状と可能性を論じた。整備状況の調査結果から、民間施設や写真などのデータの充実を図ることを課題として挙げている。

次に、観光情報を LOD に変換しシステムを開発する研究を紹介する。菊井 [6]は、岡山県の観光情報サイトが提供している CSV 形式のデータを LOD に変換し、簡単な検索インターフェースを作成する手法を提案している。技術的な課題として、何通りもある語彙のうちのどれを述語に用いればよいかわからないことを挙げている。また、運用上の問題として、データ所有者からデータの利用許諾を取れるかという点を挙げている。本研究は、オープンデータとして公開されている

(注 1) : <http://ja.dbpedia.org/>

データのみを使用しているため運用上の問題は課題と
ならない。木村ら [7]は地域の魅力を伝え観光を促進
するために、観光客による情報発信を活用するシステ
ムを提案している。既存の観光コンテンツを LOD 化し
たものを核とし、Twitter など観光客が発信する情報
を取得し LOD 化することで関連する新しい情報を付加
して提供することを可能とした。牧山ら [8]は Web 上
から取得した観光情報を RDF に変換し、RDF 化した
オープンデータや DBpedia Japanese のリソースと統合
することにより、1 つの観光地に対して多くの情報を
観光客に提供し立ち寄り観光につなげる手法を提案し
ている。作成したシステムを観光客に使用してもら
う実証実験を行い、システムの有用性の評価を行って
いる。これら 3 つの研究は、既存の観光情報を RDF に
変換し、関連するリソースにリンクを生成し情報量を増
加させるという点で本研究と類似している。しかし、
空間検索が行えない、オープンデータ以外の情報を使
用しているという点が本研究とは異なる。

最後に、Linked Data を用いて観光領域の知識ベ
ースを構築する研究を紹介する。榎ら [9]は Linked
Data における述語の統一と潜在的リンクの推定によ
り、横断的知識ベースを構築するアルゴリズムを提
案している。観光領域における述語の統一と潜在的
リンクを推定するために、観光語彙基盤の整備を進
めている。また、リソースの意味概念を推定し、キ
ーワードやカテゴリの潜在的リンクを推定するアル
ゴリズムを作成している。このアルゴリズムより、
任意のデータセットと DBpedia をリンクし各リ
ソースを横断的にリンクすることを可能にしている。
本研究では、榎らが整備を進めている観光語彙基
盤を独自に拡張したものを使用している。

3. 提案手法

本章では、オープンデータを RDF 化する手法と RDF
化したオープンデータおよび DBpedia Japanese を
用いて不足項目にデータを追加する手法を記述する。
3.1 節に CSV および xls 形式のオープンデータを
RDF 化する手法、3.2 節に不足項目のデータを追
加する手法、3.3 節に RDF データを格納する
データベースと検索する手法を示す。

3.1 RDF 化

まず、オープンデータカタログサイトから観光分
野のオープンデータを収集する。収集したデータが
xls 形式の場合は Microsoft の Excel を用い
て CSV 形式に変換する。オープンデータに含ま
れている位置情報には、緯度経度で記載されて
いるものと平面直角座標で記載されているもの
がある。そのため、平面直角座標を緯度経度
に変換するプログラムを作成し、記載され

ている位置情報を緯度経度の表記に統一する。ま
た、空間検索を可能にするために緯度経度表記の
位置情報を Well-known Text 形式で表したデー
タを追加する。位置情報の形式を統一したデー
タに対して rdfpandas^(注2) および rdflib^(注3)
という Python ライブラリを利用して RDF へ
の変換およびシリアル化を行う。述語には観光
語彙基盤(<http://www.tourism.property/#>)
および独自に定義を追加した語彙(表 1)を用い
て記述する。なお、表 1 における xsd は
<http://www.w3.org/2001/XMLSchema#>
を、geo は <http://www.opengis.net/ont/geosparql#>
を示す。

表 1 独自に定義を追加した語彙

プロパティ	定義域	値域	説明
種別	概念型	xsd:string	事物のカテゴリを文字列で表現
料金	費用型	xsd:string	利用料金や入場料を表現
WKT	場所型	geo:wktLiteral	所在地を表す地理座標を Well-known Text 形式で表現

既存の語彙を用いることが好ましい点と観光
オープンデータの列名に記載されている事項が
多く網羅されており、独自に定義する語彙
が少なくても良い点から観光語彙基盤 [9]
を使用している。変換前の CSV 形式デー
タの例を表 2 に、変換後の RDF グラフを
図 1 に示す。なお、図 1 における rdf は
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
を、tour は <http://tourism-vocab.firebaseio.com/#>
を示す。

表 2 変換前の CSV 形式データの例

拠点名	所在地	X 座標	Y 座標	種別
道の駅 富士川 楽座	富士市 岩淵 1488-1	-92981.1208	10804.65763	道の 駅

(注2) : <https://github.com/cadmiumkitty/rdfpandas>

(注3) : <https://github.com/RDFLib/rdflib>

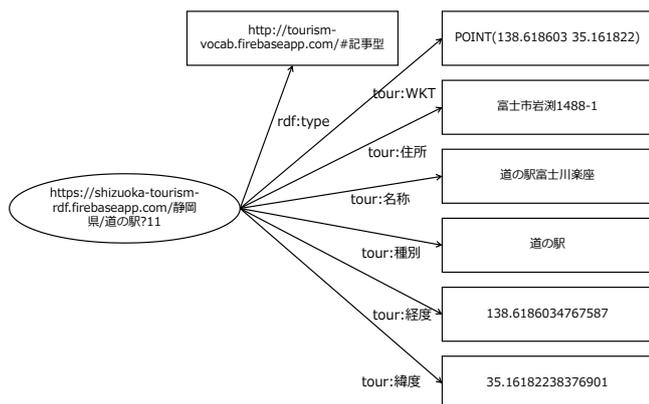


図 1 変換後の RDF グラフの例

3.2 不足項目の追加

不足項目を補うために、同一の観光地の情報にリンクを生成し不足項目のデータ取得を行う。同一の観光地のリソースへのリンク生成と不足項目データの取得の2つに分けて述べる。

3.2.1 データセット間のリンク生成

RDF 化したオープンデータはデータセットごとに独立したグラフを構成しているため、データセット間に関連性がなく同一の観光地であったとしても別々に認識されてしまう。そのため、データセット間において同一観光地のリソースに関連していることを示すリンクを生成する必要がある。RDF 化したオープンデータは名称、カテゴリ、位置情報を必ず含むため、以下の3つの条件全てに当てはまる観光地をリンク生成する対象と判断する。

(1) 観光地名称の文字列の部分一致

観光地名称が併記されている文字列(例:XX 公園・OO 湖と OO 湖)に対応するため、対象とする名称文字列が他方の名称文字列と部分一致すること。

(2) カテゴリの一致

それぞれの観光地のカテゴリが一致すること。

(3) 位置の近接

それぞれの観光地の位置情報から求めた直線距離が 200 m 以内であること。

3.2.2 不足項目のデータ取得

まず、表示項目のうち不足しているデータを同一観光地のリンクが張られているデータセットより取得する。この処理の後で不足している項目は、DBpedia Japanese よりデータを取得する。この際、Wikipedia の記事のタイトルと観光地名称が完全一致し、概要に表示対象となる都道府県名を含むリソースを同一観光地と判断する。しかし、観光地名称と完全一致で判断を行うと精度が低くなるため、オープンデータにある観光地名称内に中黒を含む場合はそこで文字列を分割したものを問い合わせるキーワードとする。

以上の処理を SPARQL エンドポイントに問い合わせることにより、不足項目データを取得することが可能となる。

3.3 RDF データの検索および格納

RDF への問い合わせ言語として空間演算が可能な言語である GeoSPARQL^(注4)を使用する。また、作成した RDF データを RDF ストアに格納し、SPARQL エンドポイントを作成する。RDF ストアとして Parliament^(注5)を使用する。GeoSPARQL を利用できる点や他の SPARQL エンドポイントと併せてクエリを処理する Federated Query を利用できる点から RDF ストアに Parliament を選定している。

4. 試作

本章では、提案手法を用いて試作したアプリケーションについて述べる。4.1 節にアプリケーションの仕様、4.2 節にアプリケーションの動作を示す。

4.1 アプリケーションの仕様

本節では、アプリケーションの仕様として対象とした地域、想定する使用場面、使用データ、機能、構成について述べる。

4.1.1 対象地域

アプリケーションの対象地域を静岡県東部地域とする。選定理由として、韮山反射炉が世界文化遺産に登録されたことや、伊豆半島が世界ジオパークに認定されたこと、伊豆市や小山町にて 2020 年東京オリンピック・パラリンピック競技大会の自転車競技実施など観光客が来る要因があることが挙げられる。また、オープンデータ政策の観点から見れば、静岡県は全国で初めて都道府県単位でのオープンデータカタログサイトが公開された自治体であり、県だけでなく市町も観光オープンデータを公開していることが挙げられる。

4.1.2 想定する使用場面

観光情報を提供する場面として、観光地の独自の魅力や宿泊施設、観光施設などの情報が必要となる「旅行計画段階」、公共交通機関の時刻や移動中の景色の情報が必要となる「移動中」、現在地周辺の休憩所や飲食店、観光スポットの情報が必要となる「観光中」の3つが想定されている [10]。試作するアプリケーションの使用場面として、旅行計画段階で観光地を選定するとき、および観光中に周辺の観光地を検索するときを想定している。

4.1.3 使用データ

行政区画の境界のポリゴンデータとして、Linked

(注4) : <http://www.opengeospatial.org/standards/geosparql>

(注5) : <http://parliament.semwebcentral.org/>

Open Addresses Japan^(注6)を使用した。

観光情報として、静岡県オープンデータカタログサイト「ふじのくにオープンデータカタログ」^(注7)で公開されているデータのうち、「文化・スポーツ・観光」分野で季節にとらわれず観光でき、観光地名称と位置情報を必ず含むデータを使用した(表3)。なお、使用したオープンデータにはカテゴリが含まれているものが少ないため、カテゴリを含め表記を統一している埼玉県の「共通データフォーマット 観光地情報」^(注8)をもとに独自に付加した。また、スーパーマーケットやコンビニエンスストア等、観光地とは言い難い51件のデータは削除した。

表3 観光情報に使用したオープンデータ

データセット名	データ作成者	データ数	代表的な観光地
伊豆の国市観光施設料金表	静岡県伊豆の国市	5	韮山反射炉
伊豆の国市入浴施設・足湯	静岡県伊豆の国市	9	富士見の湯
『みしまコロッケ』認定店	静岡県三島市	109	カフェ&ベーカリーグループ
三島市眺望地点	静岡県三島市	13	中郷温水池
裾野市観光マップ	静岡県裾野市	59	富士サファリパーク
伊豆の道風景30選	静岡県道路企画課	30	河津七滝ループ橋
静岡県景観賞受賞地区	静岡県景観まちづくり課	37	重要文化財岩科学学校
静岡のみずべ100選	静岡県河川企画課	100	柿田川
道の駅	静岡県道路企画課	21	道の駅開国下田みなと

4.1.4 アプリケーションの機能

アプリケーションで実装する機能について述べる。観光地を表示する地域の選択方法として「現在表示されている領域」と「ユーザが選択した任意の自治体」の2つが使用できる。

また、使用できる検索機能の一覧を表4に示す。

(注6) : <http://uedayou.net/loa/>

(注7) : <https://opendata.pref.shizuoka.jp/>

(注8) : https://opendata.pref.saitama.lg.jp/fs/9/0/_/kan_kotijyohou.pdf

表4 検索機能の一覧

検索方法	検索対象
名称検索	ユーザが入力した文字列と部分一致する観光地名称を持つ観光地
カテゴリ検索	ユーザが選択したカテゴリを持つ観光地
現在地周辺検索	現在地からユーザが入力した半径(単位はm)の範囲に含まれる観光地

4.1.5 アプリケーションの構成

本研究で試作したアプリケーションの構成を図2に示す。

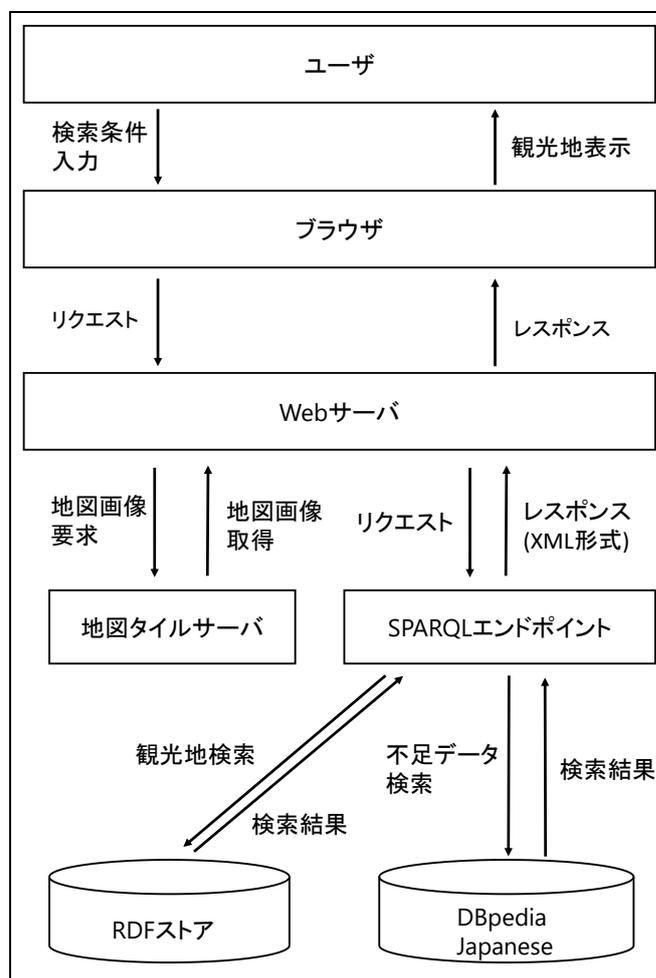


図2 アプリケーションの構成

まず、Webサーバはブラウザにてユーザが指定した表示地域や観光地名称、カテゴリといった情報をもとにSPARQLクエリを組み立てる。その後、URLエンコードしたSPARQLクエリをクエリストリングに記述し、SPARQLエンドポイントにGETメソッドで送信する。

リクエストを受け取ったSPARQLエンドポイントは

初めに観光オープンデータが格納されている RDF ストアに問い合わせを行う。次に、取得した観光地名称をキーワードとして DBpedia Japanese に問い合わせを行い、説明文、ホームページ URL、画像のデータを取得する。その後、オープンデータに説明文、ホームページ URL、画像に不足がある場合には DBpedia Japanese から取得したデータを該当箇所へ代入する。

この結果を XML 形式にて Web サーバに返信する。Web サーバにて、受け取った XML 形式のデータを解析し地図上に表示する処理を行う。

4.2 アプリケーションの動作

本節では、想定ユーザの操作順に従いアプリケーションが実際に動作することを述べる。

(1) 地図アプリケーションを開く

地図アプリケーションは Web アプリケーションとして試作しているので、URL を入力しアプリケーションにアクセスする。最初に表示される初期画面は図 3 のようになる。



図 3 アプリケーションの初期画面

(2) 表示方法の選択

画面左側にある検索マークのボタンをクリックし、サイドメニューを開く。サイドメニューの上部に表示方法を選択するラジオボタンが配置されているので、「表示中の範囲」または「選択した自治体名」の2つから選択する。「選択した自治体名」を選択した場合には、ドロップダウンリストから観光地を表示したい自治体名を選択する。この表示方法を選択した場合は、観光地を表示した際に選択した自治体の領域が赤色で表示される(図 4)。

表示方法が選択できるのは、RDF データベースに Parliament を、問い合わせ言語に GeoSPARQL を使用し空間演算ができるためである。また、任意の自治体の観光地だけでなく、画面に表示されている自治体を横断して検索できるのは、オープンデータを RDF に変換しデータの属性名を統一しているからである。



図 4 選択した自治体の領域表示

(3) 観光地名称検索

観光地名称をテキストボックスに入力する。入力されたキーワードが観光地名称に対し部分的である場合においても検索することが可能である。

図 5 に静岡県裾野市にある「富士」の文字列を含む観光地の検索結果を示す。

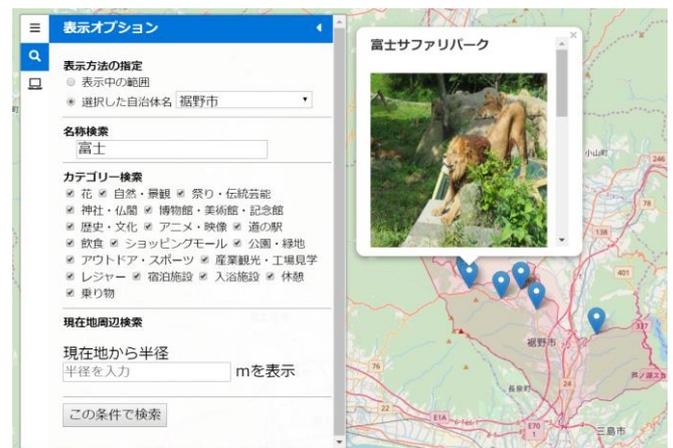


図 5 名称検索の結果

(4) カテゴリ検索

選択したい観光地のカテゴリのチェックボックスをクリックし、カテゴリを選択する。

図 6 に表示領域内に含まれ「神社・仏閣」および「歴史・文化」カテゴリに属する観光地の検索結果を示す。



図 6 カテゴリ検索の結果

(5) 現在地周辺検索

現在地を中心にユーザが指定した半径内に含まれる観光地を検索する。オープンデータを RDF 化しデータの属性名を統一している点、問い合わせ言語に GeoSPARQL を使用している点などからユーザが指定した半径に含まれる自治体の観光地を横断して検索することが可能である。なお、この検索機能のみどちらの表示方法を選択しても自治体を横断して検索を行う。

図 7 に現在地から半径 5 km にある「飲食」および「宿泊施設」、「入浴施設」を除く観光地を検索した結果を示す。

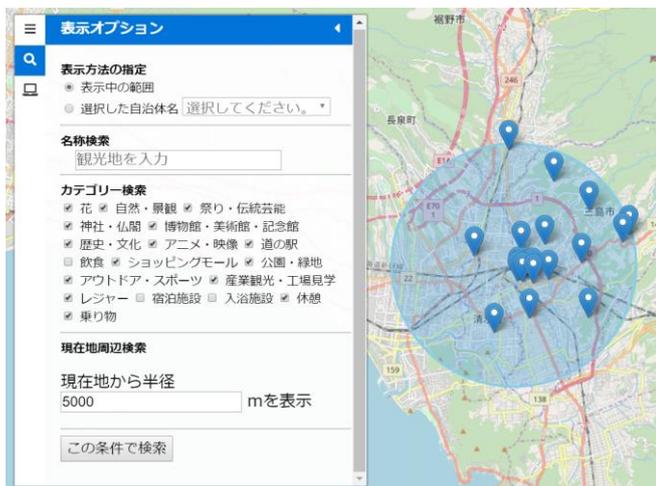


図 7 現在地周辺検索の結果

5. 評価および考察

本章では、提案手法の評価方法と結果およびその考察について述べる。

5.1 評価方法

提案手法について、以下の 2 つの観点から評価を行った。

(1) データセット間を横断した検索

現在地を静岡県三島市に所在する三島駅に設定し、半径 5 km の範囲にある観光地のデータセット名およびデータ作成者を取得した。この方法により、データセット間を横断して検索ができていているかを評価する。

(2) 不足項目のデータ追加

RDF に変換した段階、同一観光地リンク生成後、DBpedia Japanese から不足項目を取得した段階の 3 つに対し、「説明文」、「ホームページ URL」、「画像」の 3 つの項目の個数と観光地の総数を SPARQL の COUNT 関数を用いて集計し、割合を算出した。この方法により、どのくらい不足項目を補うことができるかを評価する。

5.2 評価結果と考察

(1) データセット間を横断した検索

三島駅から半径 5 km の範囲にある観光地のデータセット名とデータ作成者の取得結果は表 5 となった。

表 5 データセット名とデータ作成者の取得結果

データセット名	データ作成者
みしまコロック認定店	静岡県三島市
伊豆の道風景 30 選	静岡県道路企画課
静岡県景観賞受賞地区	静岡県景観まちづくり課
裾野市観光マップ	静岡県裾野市
三島市眺望地点	静岡県三島市
静岡のみずべ 100 選	静岡県河川企画課

表 5 より、三島市が作成した以外のデータセットや三島市以外に所在する観光地を取得できていることがわかる。このことから、提案手法は自治体を横断して観光情報を検索するのに有効であるといえる。

(2) 不足項目のデータ追加

全 383 件の観光地のうち、説明文、ホームページ URL、画像を表示可能な観光地の割合は表 6 となった。

表 6 説明文、URL、画像を含むデータの割合

種別	RDF 変換後	同一観光地 リンク生成後	DBpedia Japanese 含む
説明文	50.7 %	50.7 %	58.7 %
ホーム ページ URL	12.8 %	12.8 %	14.4 %
画像	33.9 %	34.8 %	42.6 %

表 6 より、同一観光地リンク生成後の割合は、RDF 変換後に比べ画像のみ 0.9 ポイント増加している。この理由として、画像は著作権だけでなく肖像権などの権利もあり、自治体が保有している画像を二次利用可能なライセンスを付与して公開することが他の項目より難しいためだと考えられる。それゆえ、同一観光地

のリソースであっても画像を含むものと含まないものがあり、画像のみ増加したと推測される。

また、DBpedia Japanese から取得したデータを含めると、RDF 変換後に比べ表示可能なデータ数が増加している。しかし、どの項目においても増加している割合は 10 ポイント未満である。この理由として、DBpedia Japanese から情報を取得する際に述語を DBpedia 独自の語彙ではなく標準語彙である RDF Schema^(注9)や Friend of a Friend^(注10)に属するものしか使用しなかった点が考えられる。また、DBpedia Japanese の元となっている Wikipedia 日本語版はローカル情報誌ではなく百科事典である。そのため、全データの 28 % を占める地元名物を提供する飲食店の記事などは少ないことが原因と推測される。

6. まとめと今後の課題

本研究では、自治体を横断して観光情報を取得し、地図上に統合表示する方法の検討とアプリケーションの試作を行った。具体的には、取得した観光分野のオープンデータを RDF に変換することで自治体間のデータの連携を容易にした。また、地図上に表示する観光地を検索する際に不足している項目のデータを他のデータセットや DBpedia Japanese から取得することにより、表示内容の充実を図った。アプリケーションの動作および評価結果より、複数自治体のデータを横断して地図上に表示できることを確認した。また、他のデータセットや DBpedia Japanese から不足項目のデータを取得し表示可能な情報を増やすことができた。

今後の課題として、観光オープンデータに対してはカテゴリが自治体により表記が異なることが挙げられる。カテゴリの表記が異なるとカテゴリ検索を行う場合に煩雑な処理が必要となる。このことから、埼玉県のようにカテゴリの表記を含めた共通データフォーマットを定める必要がある。

また、提案手法の課題として DBpedia Japanese を含め他のデータセットとのリンク生成数を増やすことが挙げられる。解決方法として、標準語彙だけでなくデータセット独自に定義している語彙の述語を用いて問い合わせることが考えられる。また、提案手法では観光地の名称と他のデータセットの観光地名称が完全一致する場合しか不足項目を取得することができない。使用したオープンデータの中には「重要文化財岩科学校」といった観光地名称も存在する。しかし、Wikipedia 日本語版には「岩科学校」というタイトルで記事が存在する。そのため、名称文字列を「重要文化財」、「岩

科学校」のように分割し「岩科学校」に一致するリソースを検索することが必要であると考えられる。

参考文献

- [1] 内閣官房情報通信技術総合戦略室，“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～”，https://cio.go.jp/sites/default/files/uploads/documents/opendata_tebikisyo.pptx, (2019-01-05 参照)。
- [2] 総務省情報流通行政局情報流通振興課，“自治体オープンデータ取組率向上に向けての取組方針について”，https://www.kantei.go.jp/jp/singi/it2/seimon_bunka/data_ryutsuseibi/opendata_wg_dai5/siryou4.pdf, (2019-01-05 参照)。
- [3] 福安真奈, 浦正広, 山田雅之, 遠藤守, 宮崎慎也, 安田孝美, “観光情報の公開 API 化による地域 PR モデルとその課題”, 2013 年社会情報学会 (SSI) 学会大会研究発表論文集, pp. 199-202, 2013.
- [4] 奥野拓, “観光情報 LOD”, 人工知能学会誌, Vol. 31, No.6, pp. 825-832, 2016.
- [5] 江崎貴明, 倉田陽平, 相尚寿, “観光資源に関するオープンデータと利活用に関する研究”, 第 8 回データ工学と情報マネジメントに関するフォーラム, 2016.
- [6] 菊井玄一郎, “岡山県の観光イベント情報の LOD 化と連携の試み”, 第 29 回人工知能学会セマンティックウェブとオントロジー研究会, 2013.
- [7] 木村福義, 磯野裕太, 松尾堅太郎, 一藤裕, 荒井研一, 小林透, 曾根原登, “自己拡張型オープンデータプラットフォームの研究開発”, 電子情報通信学会技術研究報告, Vol. 116, No. 23, pp. 107-112, 2016.
- [8] 牧山宅矢, 大野祐, 森田武史, 杉山岳弘, 小栗洋昭, 手嶋秀之, 山口高平, “Linked Data を利用した立ち寄り観光支援システムの社会実験”, 第 35 回人工知能学会セマンティックウェブとオントロジー研究会, 2015 .
- [9] 慎俊孝, 高橋和生, 若原俊彦, “観光領域の Linked Data を対象とした横断的知識ベースの構築法”, 第 10 回データ工学と情報マネジメントに関するフォーラム, 2018.
- [10] 国土交通省総合政策局, “観光地が取り組む効果的な観光情報提供のための資料集”, pp. 32-41, <http://www.mlit.go.jp/sogoseisaku/region/kankojoho/all.pdf>, (2019-01-05 参照)。

(注9) : <http://www.w3.org/2000/01/rdf-schema#>

(注10) : <http://xmlns.com/foaf/0.1/>