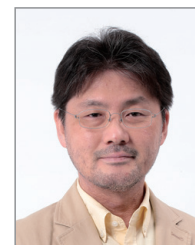


[観光情報学—スマートツーリズムに向けた研究動向—]

2 観光情報のオープンデータ化

応
般

奥野 拓 公立はこだて未来大学



地域観光情報のオープンデータ化の動向

国内におけるオープンデータ推進の黎明期より、観光に関連する情報が広くオープンデータ化の対象とされてきた。それは、多くの地域において、観光情報の発信が地域活性化の手段として重要であることを反映したものであろう。

オープンデータの大きな推進力であるシビックテックにおいても、市民生活や政治に関するデータと同様に観光データが対象とされてきた。地域の歴史・文化資源を現地や図書館で調査し、ウィキペディアに掲載するウィキペディアタウンや、まち歩きにより地域の情報をオープンライセンスの電子地図 OpenStreetMap に掲載するマッピングパーティなどのワークショップが各地で行われてきている。これらにおいては、地域住民が地域資源を再認識すると同時に、その成果物の多くは観光への活用が意図されている。

また、オープンデータ推進を目的としたコンテストにおいても、観光に関するオープンデータが多数エントリされている。たとえば、2011 年度から毎年開催されているオープンデータの総合コンテストである「LOD チャレンジ」のデータセット部門においては、2020 年度までの全エントリのうち観光に関連する作品が約 1/4 を占めている。

自治体による観光情報のオープンデータ化

自治体によるオープンデータ推進は、鯖江市、横

浜市、流山市などにおいて先進的な取り組みが始められ、その後、主に政府の主導により全国の自治体において地域情報のオープンデータ化が進められてきた。2014 年 10 月には、全国の公共データを集約するデータカタログサイト DATA.GO.JP が正式に運用開始された。また、翌年には「公共クラウドシステム」が運用開始された。このシステムは、自治体の保有するオープンデータを集約し、Web API などにより統一的に提供するプラットフォームである。本システムを利用する Web サービスは見られたものの、広く活用されるには至らず、2020 年 3 月で運用を停止している。活用の阻害要因として、自治体によるデータ数やデータ内容の粗密の差が大きかったこと、データ形式が扱い難いものであったことなどが考えられる¹⁾。

一方、2017 年 12 月に政府より、自治体による公開が推奨されるデータセットと準拠すべきデータ形式などをとりまとめた「推奨データセット」が公開されている。絞り込まれたデータ項目とデータ形式の具体例により、自治体の負荷軽減に配慮したものとなっている。観光に関しては、「観光施設一覧」、「イベント一覧」、が含まれている。2020 年末の調査²⁾によると、回答団体数 1,714 に対して、360 団体 (21%) が観光に関するデータセットを公開している。そのうち、199 団体 (55%) が推奨データセットに準拠したデータセットを公開している。今後、推奨データセット準拠のデータセットが増加していくことにより、よりオープンデータの活用が進むことが期待できる。

地域観光オープンデータの活用

倉田らによる 2016 年 8 月の調査によると, Android 端末向けの観光客向けご当地アプリと見なせるものが 852 本確認され, そのうち, オープンデータを使用しているものは 15 本であった³⁾.

政府 CIO ポータルには「オープンデータ 100」として自治体などのオープンデータ活用事例が掲載されている. 2021 年 7 月時点で掲載されている全 82 件のうち, 観光に関連する事例は, 民間が 1 件, 自治体などが 5 件である.

市民向けの生活情報や防災情報などと比較して観光向けの活用事例が少数にとどまっている要因として, データセットの多くが飲食店や民間の商業施設を含んでいないこと, データ項目が所在地や営業時間, アクセス方法などの基本情報に限定されていることがある. 飲食は観光の重要な要素であり, 観光客の興味を惹くには, 口コミやレビュー, 取材記事など, 価値判断を含む情報が有効である. つまり, 観光アプリを現状のオープンデータのみから構築するのは現実的ではなく, 他の観光コンテンツと相補的に用いるなど, 効果的な活用方法を見出していく必要がある.

以降では, 地域情報のオープンデータ化と併せてそれらを活用する観光アプリケーションを構築している 2 件の事例を紹介する.

システム開発 PBL を通した地域観光情報のオープンデータ化

公立はこだて未来大学では, PBL 型システム開発演習のテーマの 1 つとして, 2012 年度より, 着地型観光情報のオープンデータ化と, それらを活用する観光モバイルアプリの開発に取り組んできた. これはソフトウェアエンジニアを目指す学生を対象とした演習であり, サービスの立案から, データホルダへの提案, データモデリング, データセット作成・公開, モバイルアプリの開発・リリース・運用まで

の一連の活動を含む. 以降では, このテーマの最初の成果物であり, 2019 年度まで開発・運用が継続された「はこだて Map+」の活動について紹介する.

はこだて Map+

はこだて Map+ のコンセプトは「新たな発見があるまち歩きアプリ」である. まち歩きルートを辿る途中で, 周辺の観光スポットや人気のあるカフェ, レストランなどを知ることができる. また, 周辺がロケ地である映画の情報や, 近代化遺産の歴史的エピソードなどを知ることができる.

前述の通り, 観光客にとって利便性の高い情報提供をオープンデータのみによって実現することは難しい. そこで, はこだて Map+ では, 飲食店情報を含む観光スポットの基本情報と取材記事については, 非オープンデータである函館市公式観光情報サイト「はこぶら」のコンテンツを利用し, 観光アプリとしての一定の利便性を担保した上で, 「+α 情報」を重ねるというアプローチをとっている. 最初のリリースでは, オープンデータとして, 函館市が配布しているモデル観光ルートマップ「函館まちあるきマップ」, 函館工業高等専門学校の PBL 授業で制作された Web サイト「函館近代化遺産ポータルサイト」のコンテンツを用いている. また, 非オープンデータとして, 映画のロケ地情報を提供する Web サイト「はこだてフィルムコミッション」のコンテンツを用いている.

図-1 に, はこだて Map+ の主要な画面遷移を示す. 函館まちあるきマップのまちあるきコース画面(図-1 c)で地図上にマーカーとして表示されている観光スポットを選択すると, 観光スポット画面(図-1 d)に遷移する. 「元町配水場」の場合は, 「桜の名所」であるという情報に加えて, 映画「Little DJ」などのロケ地であるという情報や, 近代化遺産「日本最古の配水池」であるという情報が併記される. また, それぞれの解説画面(図-1 e, g)の「もっと見る」から情報提供元の Web サイトの該

特集

Special Feature

当ページ（図-1 f, h）を閲覧することができる。つまり、このアプリは、既存の観光コンテンツへの新たな接点としても機能する。

次節では、アプリで用いているデータの作成について述べる。

着地型観光コンテンツの RDF データ／ LOD 化

本演習では、複数の観光コンテンツをリンクトオープンデータ（LOD）化し、それらを連係させてアプリのコンテンツとしている。LOD では、RDF 形式のオープンデータを、SPARQL 言語によりデータ取得可能なサーバ（SPARQL エンドポイント）に配置することにより、データの連係利用を実現する。本演習では、アプリにおいてデータを統一的に扱うために、すべての観光コンテンツを RDF 形式のデータセットに変換する。そして、公開／非公開の SPARQL エンドポイントに分けて配置する。これらよりアプリが必要な RDF データをオンデマンドで取得することにより、データの更新がリアルタイムにアプリのコンテンツに反映される。

「はこだて Map+」の最初のリリースでは、前節で述べた 4 種類のコンテンツを用いている。2 年目以降の活動において、函館市避難所マップの津波避難所情報（非オープンデータ）と、函館スイーツ推

進協議会の店舗・商品情報を掲載する「函館スイーツ公式サイト」のコンテンツ（オープンデータ）の RDF データセットを追加している。以上のコンテンツからのデータ生成は具体的には次のように行っている（図-2 参照）。

紙媒体と PDF ファイルのみで提供されている「函館まちあるきマップ」と「函館市避難所マップ」、および、スポット数が少なく更新頻度の低い「はこだてフィルムコミッション」については、抽出したテキストを一部手作業で整形してデータを作成している。一方、「はこぶら」と「函館スイーツ公式サイト」については、スポット数が多く、不定期に更新されるため、定期的なクロール・スクレイピングによりデータを自動抽出している。以上のすべてのデータを、コンテンツの種類に応じて設計した RDF データモデルに従って RDF に変換するスクリプトを用いて RDF データセットを生成する。

演習では、開発したアプリのリリース許諾を得るために、データホルダーへの提案プレゼンテーションを行う。その際、地域情報のオープンデータ化の効果についても説明し、オープンデータ化を提案する。LOD 化されているデータセットは、その成果である。一方で、「はこぶら」のような観光スポットの取材記事や、はこだてフィルムコミッションのよう



■図-1 「はこだて Map+」の画面遷移

特集

Special Feature

に他メディアの著作物を含む場合などは、アプリ内での利用許諾にとどまっている。以上述べた通り、地域オープンデータの推進において、コンテンツの二次利用によるメリットを観光アプリとして可視化することの効果は大きい。

地域文化財アーカイブのオープンデータ化

南北海道の文化財

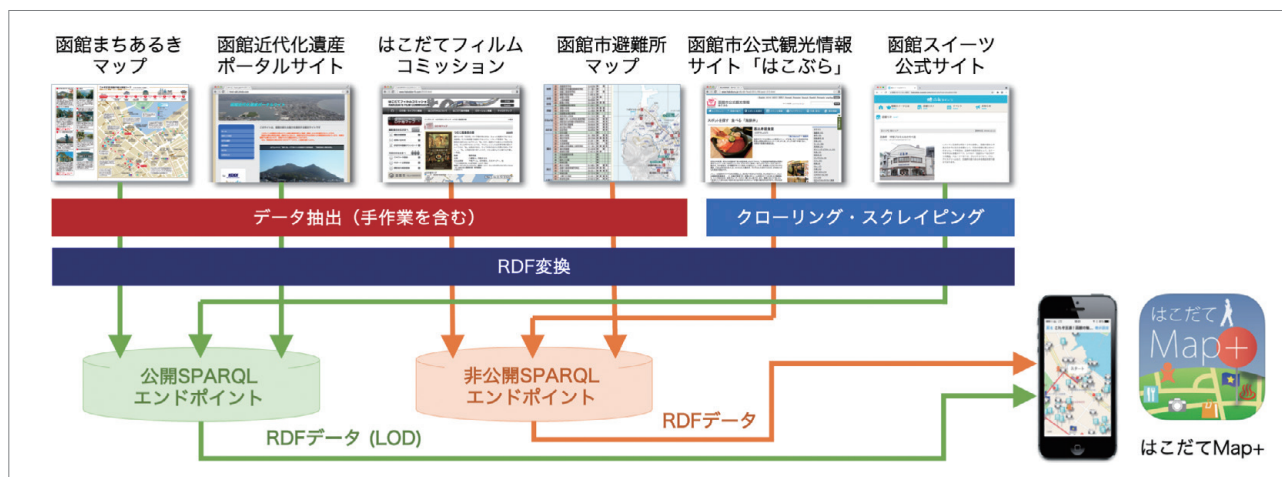
道南地域の学芸員組織「道南ブロック博物館施設等連絡協議会」では、地域文化財にまつわるエピソードやイベントなどの情報をブログにより活発に発信している。しかしフロー型の情報発信では、恒久的な価値を持つ文化財情報を十分に活かさない。そこで、マップを中心としたストック型のWebサイトにより地域文化財の情報を蓄積・発信する取り組みを、公立はこだて未来大学との協働により開始した。この取り組みでは、Webサイトをフィールドワークや市民協働に活用するとともに、コンテンツをオープンデータ化し、観光アプリケーションで活用（二次利用）されることを目指している。2017年3月に「南北海道の文化財」Webサイトの一般公開運用を開始した（図-3）。2021年7月末時点の登録文化財は714件である。図-4に文化財「伊能忠敬北

海道最初の測量地」の解説記事を示す。碑文をテキスト化して掲載しているため、碑文の一部から検索することもできる。

アクセス解析によると、サイトへのアクセスの約80%が検索エンジン経由であり、文化財の名称や人名、地名、歴史的な出来事などのキーワードで検



■図-3 「南北海道の文化財」のトップページ



■図-2 「はこだて Map+」で使用している RDF データ／LOD の生成

特集

Special Feature

索している。また、相対的に知名度の低い文化財が継続的に高いアクセス順位となる傾向がある。そのような文化財として、「ハーバー遭難記念碑」、「箱館通宝銭座跡」、「もりたの池碑」が挙げられる。これらに共通するのは、Web 上では「南北海道の文化財」以外からは十分な情報が得られないという点である。知名度が低い文化財であっても、継続的な需要が一定数あるという事実は、網羅的に地域文化財の情報を収集・蓄積し、Web 上で長期的に公開し続けることの意義を示している。

地域文化財情報のオープンデータ化

サイトへの文化財記事の投稿は、オープンデータとしての公開を前提として行われている。著作権の問題が発生しないように配慮し、解説文章は各地域の学芸員が新たに書き起こし、写真も新たに撮影することとしている。サイト全体の著作権は協議会の帰属とし、文化財記事単位で選択的に CC BY ライセンスを表示する方式を取っている(図-4 右図参照)。2021 年度より、全文化財コンテンツの CSV データをエクスポートする機能をサイトに追加して運用している。CSV データセットは毎日深夜に自動生成される。

文化財コンテンツを二次利用する際の課題として、地域による文化財数の偏りがある。100 件を超える文化財を掲載している市町がある一方で、郷土資料館などの施設のみという市町もいくつか存在する。

これは、協議会は蓄積・公開の「場」を提供し、文化財記事の投稿は各地域の学芸員に一任されていることによる。文化財数の地域差をどのように埋めていくかが課題である。

文化財オープンデータを活用する歴史観光 Web アプリケーション

筆者らは、「南北海道の文化財」のコンテンツを観光分野に活用するユースケースの 1 つとして、スマートフォンから利用する歴史観光 Web アプリケーション「文化財巡りルートマップ」を試作している。このアプリケーションは、文化財に対する興味を喚起し、新たな文化財との出会いを促進することを目的としている。アプリケーションの主要な画面遷移を図-5 に示す。

このアプリケーションでは、そのコンテンツのすべてを、「南北海道の文化財」とウィキペディアのコンテンツから自然言語処理技術を用いて自動作成している。たとえば、テーマ選択画面に表示するテーマは、文化財の説明文から抽出した特徴語を意味的な近さに基づいてクラスタリングして決定している。意味的な近さは、単語の分散表現技術の 1 つである Word2Vec により算出している。テーマの名称は、ウィキペディアの概念階層を用い、各クラスターに含まれる特徴語の上位概念として求めている。この方法により、「漁業」「芸妓」「祭り」など意外性のあるテーマも生成される。従来の固定的なモデルルートとは



■図-4 文化財解説ページ(伊能忠敬北海道最初の測量地)

特集

Special Feature

異なり、多様性のある観光ルートのテーマ設定が可能となり、セレンディピティ効果も期待できる。

このアプリケーションでは、文化財の追加をトリガーとして、すべてのコンテンツを自動的に再生成する。その結果、文化財の増加に応じて、観光ルートのテーマや候補となる文化財も変化していく。このアプリケーションにより、観光オープンデータの数と多様性の拡大が与える効果が可視化される。

観光オープンデータの今後

これまで述べてきたように、地域の観光オープンデータは、シビックテックを始めとするさまざまな地域活動・連携を通して作成され得る。しかし、その主体となるのは、やはり自治体の観光部署であり、地域の観光協会を始めとする観光関連団体である。それらが適切に責務とコストを分担することにより、観光データが多重管理されることなく、かつ、即時的に共有されることが理想である。

COVID-19の影響により、飲食店や各種施設の営業時間が不定期に変更されるようになり、インター

ネット上の情報が実態に即していない。結果として、消費者の利便性が損なわれるとともに、事業者にも多大な機会損失が生じることとなった。これからの観光復興に向けて、即時性が求められる観光情報の網羅的な収集・共有を低コストで実現する地域観光オープンデータ共有基盤が必要である。その構築・運用には、産官学の観光ステークホルダの協業が不可欠である。

参考文献

- 1) 奥野 拓：観光情報 LOD, 人工知能, Vol.31, No.6, pp.825-832 (2016).
- 2) 内閣官房情報通信技術 (IT) 総合戦略室：地方公共団体へのオープンデータの取り組みに関するアンケート回答一覧 (2021).
- 3) 倉田陽平：観光客向け「ご当地アプリ」の現況, 人工知能, Vol.31, No.6, pp.839-843 (2016).

(2021年7月30日受付)

■奥野 拓 (正会員) okuno@fun.ac.jp

公立はこだて未来大学・システム情報科学部・情報アーキテクチャ学科・教授。観光情報、デジタルアーカイブ、オープンデータ、ソフトウェア工学などに関する研究に従事。



■図-5 文化財巡りルートマップアプリケーション