

郷土食による地域理解支援システム「もちマップ」の オープンデータ化についての考察

A Consideration on Open Data Conversion for Implementing Local Foods-base Region Understanding Support System “Mochi-Map”

河村郁江¹, 伊藤宗太², 伊藤孝行¹, 白松俊¹

Ikue Kawamura, Sota Ito, Takayuki Ito, Shun Shiramatsu

1 名古屋工業大学大学院 情報工学専攻, 名古屋市昭和区御器所町

Nagoya Institute of Technology University, Gokiso-cho, Showa-ku, Nagoya, Aichi

2 インフォ・ラウンジ合同会社, 神奈川県横浜市都筑区茅ヶ崎中央

Info Lounge LLC, Chigasaki chuo, Tsuzuki-ku, Yokohama, Kanagawa

概要: 郷土食は、古くから地域の伝統的な産物を活用している食べ物である。各地の郷土食をよく知る事は、地域の理解や知見につながる。当研究では、2016年から郷土食による地域理解支援システムを試作している。本研究では、2017年10月に行われた共通語彙基盤ハッカソンをきっかけに、システムで利用しているデータをLOD(Linked Open Data)にしたことを述べ、今後のシステムやデータの活用について考察する。

Abstract: Since local food have been made of local products, view point of local food culture is effective for local understanding. We have been prototyping a regional understanding support system by local food from 2016. We participated in a hackathon on utilizing IMI (Infrastructure for Multilayer Interoperability) at October 2017. Since this event, we converted our data about local food into linked open data (LOD) based on the IMI vocabularies. This paper discusses our perspective about this system and utilizing data in the future.

キーワード: 郷土食, WEB マップ, 情報の構造化, Linked Open Data

Keywords: Local food, Web map, Structuring information, Linked Open Data

1. はじめに

郷土食は地域の産物を活用した食べ物である。郷土食をよく知ることは、地域の風土や文化、自身の住む地域や他地域との違い、および時代の変化について知ることであり、地域の理解や知見につながる。

しかし、昔から各地域で作られ続けてきた郷土食は、近年の便利な食生活の中では手間がかかることや、地域の少子化のため作り伝えられなくなり、自分の住む地域の郷土食について深く知る人は減って来ている。また、郷土食は地域ごとの解釈が異なることや、時代による食生活の変化等により、元々の成り立ちは分かりにくくなっている。そのため2016年の秋から制作中のWEBマップでは、日本地図からもちの分布を見る事で、郷土食への興味や理解を深めることを目指した[1][2]。図1は本研究が目指すユースケースの概要図である。

本研究で作成しているシステムは、郷土食情報としてもちの情報を扱うため、名称を「もちマップ」とした。

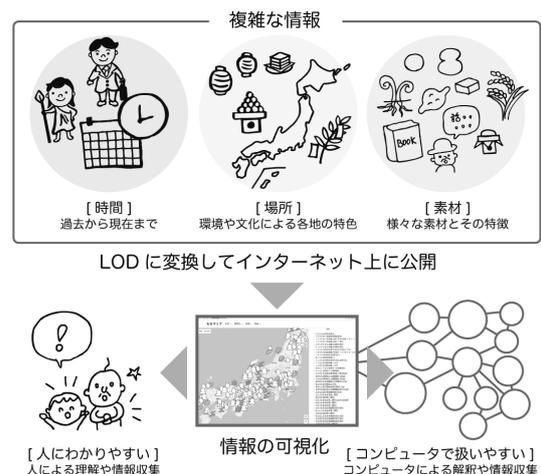


図1 本研究が目指すユースケースの概要

もちを主題として扱う理由を下記に述べる。

- ・ 地域の生活文化と深く結びつき、ハレの食から、日常の食、および保存食として食べられて来た。

- ・ 時代に伴い食生活が変化する中でも、長く食べ続けられている。
- ・ 日本全体で食べられ、使用素材が地域の特色と結びついている場合が多い。

このような理由から、もちは身近であり、郷土食の一例として分かりやすく、地域理解の一助となると考えた。この研究では対象を、郷土食や地域の文化に興味がある人や、学びたい人とする。我々は、このようなもちの情報を整理し、視覚化するため、「もちマップ」を開発してきた。

我々は、2017年10月8,9日に開催された“実践！共通語彙基盤ワークショップ・ハッカソン”[3]において、本システムで利用しているもちデータを Linked Open Data (LOD) と呼ばれる形式のオープンデータにした。LOD とは、様々な組織が公開するオープンデータを関連付け、組み合わせることで活用できるようにする技術的枠組みである。また、LOD 化するにあたって、共通語彙基盤を用いた。共通語彙基盤とは、行政などが保有する公的なデータを対象とし、経済産業省所轄の情報処理推進機構 (IPA) が標準化作業を試みている語彙セットである。これに基づいて LOD 化したことにより、データの相互運用性 (Interoperability) が向上し、他のオープンデータと組み合わせた分析が可能になると期待される。

データをオープンデータ化するために何をしたのか、また今後どのようにオープンデータを利用して行くかについて、これまでの制作の問題点も含めて考察する。

2. これまでの開発と課題

我々は2016年の冬からもちマップを開発してきた。最初に制作した従来のもちマップをバージョン1とし、2017年のものをバージョン2とする。本研究はバージョン3となる。以下にこれまでどのような開発をし、どのような課題があるかを述べる。

【もちマップ バージョン1】

もちマップバージョン1では、もちデータを書籍[4][5]やインターネット上の情報、および実際に人から聞いた話などを元に作成した。システムの主な機能として、以下の4つの機能を実装し、システムの位置や素材、および調理法などの属性を可視化した。

- (1) 検索窓
- (2) 属性選択ボタン
- (3) 背景とレイヤーをコントロールするチェックボックスパネル
- (4) 時間軸スライダー

日本地図全体上に表示された情報や属性を見ることが、郷土食の特徴や地域の特色がわかれば、地域理

解が深まると考えた。図2は従来のもちマップの主な機能である。



図2 従来のもちマップの主な機能

もちマップの評価を行った所、郷土食理解が深まったという意見が得られた。一方で、バージョン1での課題として下記の内容が指摘された。

- (1) もちの全体像だけでなく、各もちの詳細情報ページを作成する
- (2) 複数人によるもちデータの入力
- (3) より利用しやすく、分かりやすいUIへの改善
- (4) 時間軸機能の改善
- (5) 地域区分や統計情報の工夫

【もちマップバージョン2】

その後、2017年にバージョン1で出た課題の中で、下記を試みた。

- ・ 複数人によるもちデータの入力
- ・ 各もちの詳細情報ページの作成

バージョン1ではローカルで読み込んだもちデータを、データベース化した [6]。図3はもちマップバージョン2のデータベース構成である。

データベースの正規化を行う際に、初めてもちデータの語彙の検討をした。図書館でよく使用されている日本十進分類法を見ると、「餅」は「民俗」と「料理」に分類されている。まず「民俗」の語彙で調べた所、国立歴史民俗博物館による民俗語彙データベース[7]が存在したが、参考に留めた。食の語彙の論文[8]によると、「食を学として取り上げる場合、予想される学術領域は図に示すような広範なものとなろう」とあり、図には50種類以上の分野が記載されている。また、Google Scholar[9]で「郷土食」を検索した所、栄養学 約375件、経済学 約373件、社会学 約293件、民俗学(民族学) 約168件、地理学 約159件、調理学 148件、歴史学 約75件、農学 約66件、考古学 約54件、国文学 約5件などがあり、郷土食のみでも広い範囲で扱われているため、ある分野に基準を置くことはせず、どの分野からも利用できる Dublin Core[10]をもち



図3 もちマップデータベース構成

```

<section itemscope itemtype="http://schema.org/Recipe">
  <div id="wrap">
    <?php if (isset($error)) : ?>
      <p class="error"><?= h($error); ?></p>
    <?php endif; ?>
    <h1 itemprop="name" class="f2 bk01"><?= h($row['title']); ?> | <?= h($row['title_kanji']); ?></h1>
    <div itemprop="image" itemscope itemtype="http://schema.org/ImageObject" class="p_center">
      " width="80%" class="p_center" />
      <div itemprop="caption" class="p_center"><?= h($row['image_caption']); ?></div>
    <div itemprop="description"><?= h($row['description']); ?></div>
    <dl>
      <dt>素材1</dt><dd itemprop="recipeIngredient"><?= h($row['ingredient01']); ?></dd>
      <dt>素材2</dt><dd itemprop="recipeIngredient"><?= h($row['ingredient02']); ?></dd>
      <dt>素材詳細</dt><dd itemprop="description"><?= h($row['ingredient_detail']); ?></dd>
      <dt>トッピング1</dt><dd itemprop="recipeIngredient"><?= h($row['topping01']); ?></dd>
      <dt>トッピング2</dt><dd itemprop="recipeIngredient"><?= h($row['topping02']); ?></dd>
      <dt>トッピング詳細</dt><dd itemprop="recipeIngredient"><?= h($row['topping_detail']); ?></dd>
      <dt>味</dt><dd itemprop="name"><?= h($row['taste']); ?></dd>
      <dt>調理法</dt><dd itemprop="cookingMethod"><?= h($row['cook']); ?></dd>
      <dt>形</dt><dd itemprop="name"><?= h($row['shape']); ?></dd>
      <dt>形詳細</dt><dd itemprop="description"><?= h($row['shape_detail']); ?></dd>
      <dt>用途</dt><dd itemprop="name"><?= h($row['use01']); ?></dd>
      <dt>用途詳細</dt><dd itemprop="description"><?= h($row['use_detail']); ?></dd>
      <dt>都道府県</dt><dd itemprop="spatialCoverage"><?= h($row['prefecture']); ?></dd>
      <dt>市町村</dt><dd itemprop="spatialCoverage"><?= h($row['municipality']); ?></dd>
      <dt>町域</dt><dd itemprop="spatialCoverage"><?= h($row['town']); ?></dd>
      <dt>時期</dt><dd itemprop="name"><?= h($row['season']); ?></dd>
      <dt>いつからあるか</dt><dd itemprop="name"><?= h($row['start_m']); ?></dd>
      <dt>いつまで存在するか</dt><dd itemprop="name"><?= h($row['end_m']); ?></dd>
      <dt>情報元の種類</dt><dd itemprop="citation"><?= h($row['referring']); ?></dd>
      <dt>情報元</dt><dd itemprop="citation"><?= h($row['referring_detail']); ?></dd>
      <!-- <div><?= h($row['opinion']); ?></div> -->
      <dt>投稿時間</dt><dd itemscope itemtype="http://schema.org/Date" itemprop="commentTime"><?= h($row['time_post']); ?></dd>
      <dt>記入者</dt><dd itemprop="author"><?= h($row['inputter']); ?></dd>
    </dl>
  </div>
</section>

```

図4 もちマップを Schema.org にあてはめた例

データに適用した。Dublin Core は情報資源の組織化のために作られ、語彙が共通認識となるよう設計された基本語彙セットである。しかし、実際に当てはめると、Dublin Core は書籍を想定して作られた規格であるため、郷土食データを当てはめるには無理があった。そのため代わりに Schema.org[11]を使用した。

schema.org はコンピュータにもわかる形式で意味を付ける構造化マークアップのための語彙であり、Google や Microsoft、および Yahoo! が共同で進めている。検索エンジンはこのマークアップを利用し、質が高い検索結果を出すことができるため、ユーザは必要なページに辿り着きやすくなる。schema.org の型の中で食に関するものは、Thing/CreativeWork/Recipe(レシピ)、Thing/Event/FoodEvent(食イベント)および Thing/Organization(または Place)/LocalBusiness/FoodEstablishment(食品関係)などがあるが、郷土食としてのもちを表現するのに合う語彙がなかったため、元々ある語彙を使用した。

Schema.org は独自の拡張も可能である。図 4 はもちマップを Schema.org にあてはめた例である。バージョン 2 においては、このように様々な語彙の検討をした。(2) 複数人によるもちデータの入力については、各もちのページを生成するフォームを作成し、インターネット上に公開した[12]。公開当初にフォーム入力のユーザーテストを行った(2017年12月、20代から40代の男女30名、名古屋近郊のIT系コミュニティイベントにて)。入力ミスを減らすため、基本的には選択式の入力にしたが、手入力の項目では、正しく入力されない例が30件中4件あった。また、「素材」を聞く項目では、「わからない」が30件中2件、「情報源」の項目では「人から聞いた」が30件中10件、「その他」が6件、「WEBの情報」が6件であり、「書籍」という回答は0件であった。以上から、複数人によるもちデータの入力自体は実現できたが、一般向けに入力してもらうデータの信頼性が低いため、フォームのUIを改善することや、管理側での確認など、入力情報の収集・管理方法の検討が必要であることがわかった。

3. もちデータを Linked Open Data に
共通語彙基盤のハッカソン[3]において、もちマップの新しいシステムを作成した(システム制作:伊藤宗太)。これをバージョン3とする。図5はもちマップバージョン3の画面である。

今回の大きな変更点は、1つ目に、もちデータを共通語彙基盤を基にしたデータモデル記述(DMD)にすることで、同時にデータが RDF 形式になり、そのデータを公開 SPARQL エンドポイントに置いたことで LOD (Linked Open Data) になったことである。2つ目には、同じもちデータを利用して、各もちの詳細ページが閲

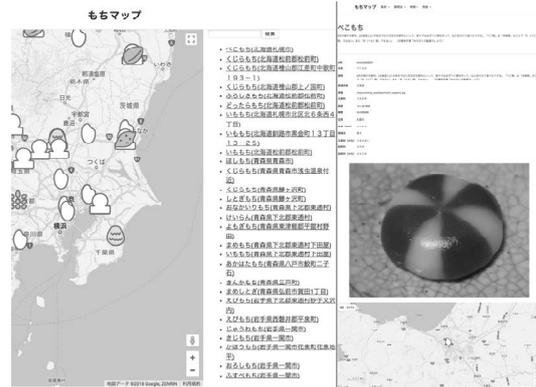


図5 もちマップバージョン3

大カテゴリー	素材	調理方法	時期	用途
• 素材	• すべてのもち	• 焼く	• 正月	• 日常食
• 調理法	• もち米	• 茹でる	• 春	• 間食
• 時期	• もち米+うるち米	• 煮る	• 夏	• お祭り
• 用途	• うるち米	• 炊く	• 秋	• お供え物
• 県別	• 小麦粉	• 練る	• 冬	• 保存食
	• 雑穀(ヒエ、アワなど)	• 干す		• 正月
	• 堅果	• その他		
	• さつまいも			
	• じゃがいも			
	• さといも			
	• 根(クス、ワラビなど)			
	• 草			
	• その他			

図6 従来のもちマップの分類項目

左: トップページ, 右: 各もちのページ例

覧できるようになったことがある。以下にどのような流れで LOD にしたかを説明する。

本章ではもちデータをオープンデータ化した流れを述べる。図6は従来のもちマップの分類項目である。

(1) 共通語彙基盤(IMI)

共通語彙基盤[13]は、共通で使われる語彙と、語彙同士の関係を示す仕組みで構成された、オープンデータの相互運用性を高めるためのフレームワークである。コンピュータ間のデータ通信を円滑かつ確実に行うことが目的であり、一つの語の代表的な意味や構造の概念、および他の概念との関係などを明確にした概念の集合を指して用いる。コンピュータでの処理を難しくさせている用語の問題(類似用語、重複用語、分野で意味が異なる同一語など)を解決するため、概念的な同一性を保証する形で「用語」を定義している。この基準に合わせるために、表形式のデータを DMD(Data Model Description)に基き作成する。DMD はデータの構造や項目、各項目の値の範囲や書式などを説明するものであり、DMD 作成ツールを使用することで、人間にわかりやすく、コンピュータでのデータ交換に適した形式および LOD に適した RDF 形式のデータとして公開することもできる。

もちデータを共通語彙基盤に合わせるに当たり、一段階目に元々独自のシステムに合わせて作成していた不要な項目を整頓し、不必要な項目は削除した。共通語彙基盤のコア語彙バージョン 2.4 の中にもちに合いそうな項目(例えば食品や農業に関する項目)が無かったが、まず、既存の語彙(コア語彙)のみを用いて共通語彙基盤に合わせた。図 7 はコア語彙のみの共通語彙基盤化である。

第二段階では、共通語彙基盤の中にあるクラス用語一覧(2017年10月時点でバージョン 2.4)の中から、もちデータに合う項目を探した。既存のクラス用語の中で、もちに一番近いと思われる「製品型」を選んだ。製品型は工業製品を想定した用語であるため、郷土食にありそうなプロパティ、例えば食や農業などに関わるで共通語彙基盤化をした。

第三段階では、製品型には無いが、コア語彙の中にある語彙を継ぎ足し、最終的に独自の拡張語彙を作成してもちデータを共通語彙基盤化した。最終的に独自語彙を作成し、共通語彙基盤化をした。図 8 は独自語彙を使用した共通語彙基盤化である。

(2) LOD(Linked Open Data)

(1) 共通語彙基盤の DMD 作成ツールを使用し、SPARQL エンドポイントで公開することで、もちデータは LOD になった。

LODとは、データを公開して共有するため、コンピュータで処理が出来る形にする技術の総称である。コンピュータ処理が可能なデータの構造化のために、W3C(World Wide Web Consortium)が RDF(Resource Description Framework)というグラフをベースとした基準モデルを作成し、推奨している。RDF については次に説明する。

(3) RDF(Resource Description Framework)

RDF は、ウェブ上のリソースを記述するための統一された言語体系であり、W3C が 1999 年 2 月に規格化している(W3C:REC-rdf-syntax-19990222)[14]。

従来の HTML では、人間が自ら読み、意味を理解して情報を得た。しかしコンピュータは HTML に記述してある文章の意味を解釈することが出来ない。意味のあるタグ付けとして XML があるが、XML タグは各自が独自の名称を付けるため、独自の名称の意味はコンピュータには理解できない。RDF は記述が統一されているため、SPARQL のような特定の言語を用いることで、機械的に操作することが出来る。

(4) SPARQL(SPARQL Protocol and RDF Query Language)

SPARQL[15]は W3C が 2008 年にリリースした RDF クエリ言語であり、RDF で記述された XML や Turtle など

の RDF データから検索や抽出をすることが出来る。また、RDF データに SPARQL クエリで問い合わせるための、公開インターフェースとして SPARQL エンドポイントがある。SPARQL エンドポイントには RDF ストアに格納された RDF データを置くことで、SPARQL から問い合わせることが出来る。

共通語彙基盤の DMD を使用することによって、RDF データに変換したもちデータは、代表的な RDF ストアである Virtuoso[16]に格納し、公開 SPARQL エンドポイントに置いてインターネット上に公開した。

(5) DBpedia

DBpedia[17]には Wikipedia の「インフォボックス (infobox)」と呼ばれる構造化情報(テーブル、カテゴリ情報、および画像、地理座標、外部ウェブページへのリンクなど)を抽出して作られた、問い合わせ可能な統一データセット(RDF 形式にしたデータ)が保存されている。利用者は、DBpedia から SPARQL を使用することでデータセットの情報を検索できる。

4. システム構成

従来のシステム(バージョン 1)では、ブラウザ側では Leaflet と OpenStreetMap を使い、Geojson 形式のもちデータを地図上に可視化していた。バージョン 2 では、サーバー側で MySQL と PHP を使用した。今回のバージョン 3 では、REACT[18]と Google Maps API[19]を用い、RDF 形式のもちデータを、Google Maps 上に可視化している。RDF ストアは Virtuoso を使い、SPARQL Endpoint を公開した。図 9 はバージョン 3 のシステム構成図である。

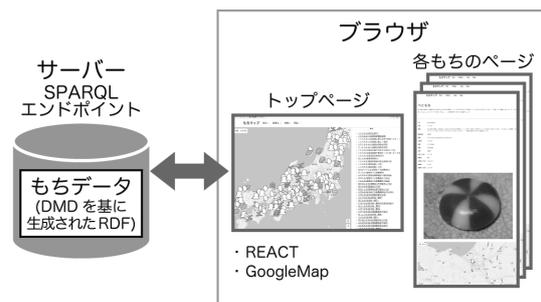


図 9 バージョン 3 のシステム構成図

5. 考察

従来のもちマップでは、システムに関しては人が見て分かりやすいもの、データに関しては独自項目を考えて作成していた。今回もちデータを LOD にしたことで、以下の2点の可能性が考えられるようになった。

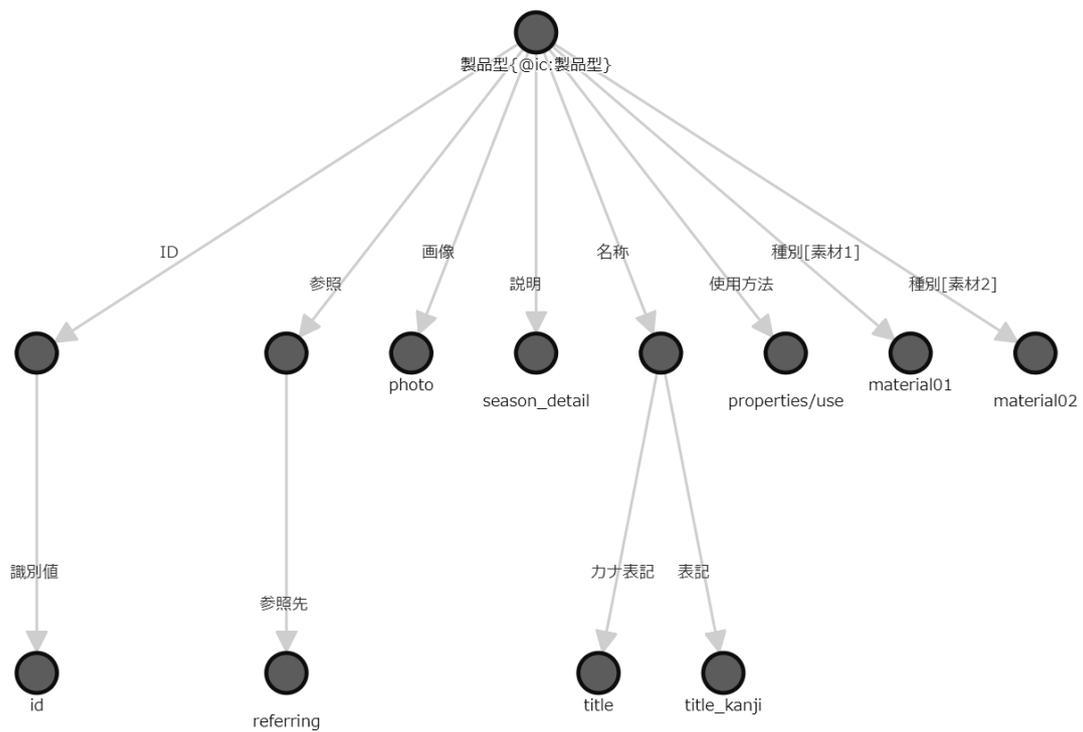


図7 コア語彙のみの共通語彙基盤化

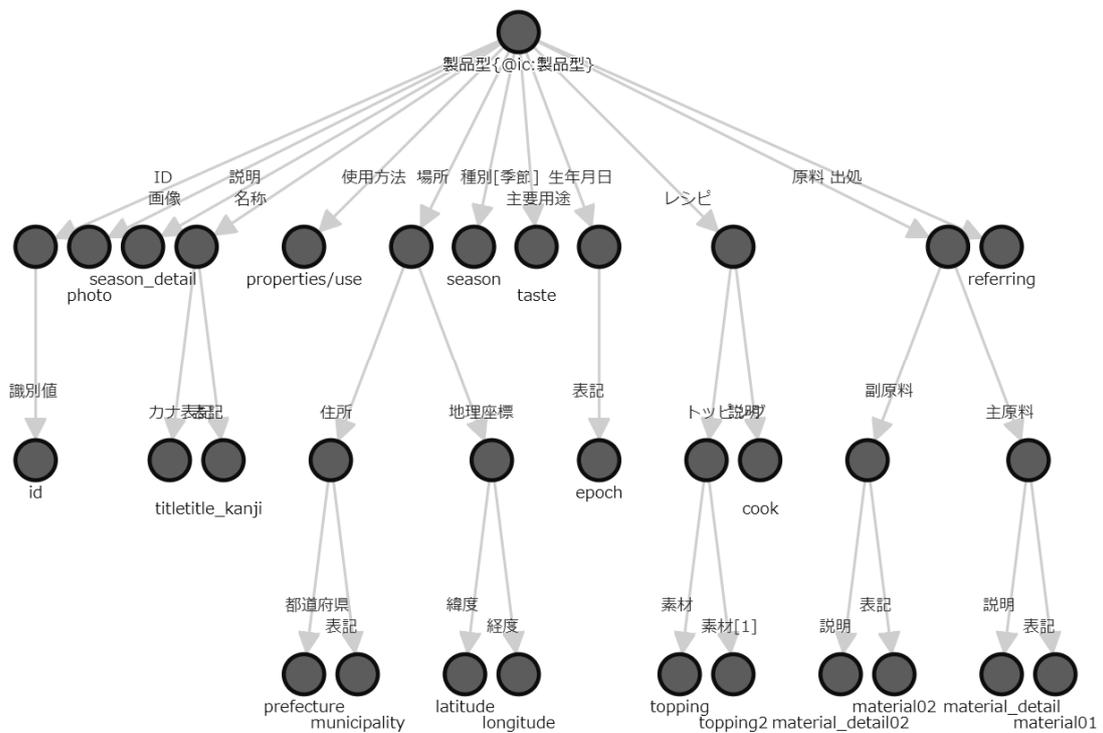


図8 独自語彙を使用した共通語彙基盤化

まれ、データの相互運用性が向上した。また UI 側の改善点としては、各もちの詳細ページが閲覧出来るようになったことがある。これらの改善に伴い、バージョン 3 でも新しい課題が発生した。

- データの構造の改善
- 新しいシステムの UI の改善と、システム自体の検討と改善
- 他の LOD データとの連携

データの構造の改善に関しては、どのようなデータ構造にしたら、元々本研究が目的としている複雑な郷土食の情報を分かりやすく表せるかということを考えなければならない。また、本システムはインターネット上に公開はしているが、まだ広めるための活動や評価をしておらず、独自に考察をしたのみであるため、近いうちにユーザーテストを行う予定である(2018年2月8日現在)。

それに加え、もちマップバージョン 1,2 からの課題である、より利用しやすく、分かりやすい UI への改善や、地域区分や統計情報の工夫、および入力情報の収集・管理方法の検討なども残っている。今後は、今までの課題の解決を進めながら、今回新たに利用した技術を用いてもちマップを改善して行きたいと考えている。

謝辞

本論文作成にあたり、名古屋工業大学大学院 特任研究員の河瀬諭先生、来栖川電算の森俊也さん、および実践！共通語彙基盤ワークショップ・ハッカソンの皆様ならび多くの方に大変お世話になった事に深く感謝を申し上げます。

参考文献

- [1] 河村 郁江. 郷土食による地域理解支援システム「もちマップ」の試作. 研究報告人文科学とコンピュータ (CH), 2017-CH-113(4), 1-2, (2017-01-28).
- [2] 河村 郁江. “2016 年度版もちマップ”. <http://phirip.com/t/mochimap/e/index.html>, (参照 2018-02-08).
- [3] 共通語彙基盤：IPA 独立行政法人 情報処理推進機構. 実践！共通語彙基盤ワークショップ・ハッカソン. <https://peatix.com/event/298493?lang=ja>, (参照 2018-02-08).
- [4] 奥村 彪生 (解説). 聞き書・ふるさとの家庭料理 第 5 巻 もち・雑煮, 農山漁村文化協会 編集, 農山漁村文化協会, 2002-12.
- [5] 奥村 彪生 (解説). 聞き書・ふるさとの家庭料理 第 7 巻 まんじゅう おやき おはぎ, 農山漁村文化協会 編集, 農山漁村文化協会, 2003-2-25.
- [6] 河村 郁江. “2017 年度版もちマップ”. <http://phirip.com/mochimap/e/>, (参照 2018-02-08).
- [7] 大学共同利用機関法人人間文化研究機構 国立歴史民俗博物館. 民俗語彙データベース. <https://www.rekihaku.ac.jp/doc/gaiyou/goi.html>, (参照 2018-02-08).
- [8] 太田泰弘. 食のターミノロジー(＜小特集＞ドクメンテーションとターミノロジー) ドクメンテーション研究 Vol. 32 (1982) No. 11 p. 555-559, 一般社団法人 情報科学技術協会.
- [9] Google LLC. Google Scholar. <https://scholar.google.co.jp/>, (参照 2018-02-08).
- [10] The Dublin Core Metadata Initiative (DCMI). <http://dublincore.org/>, (参照 2018-02-08).
- [11] Schema.org. <http://schema.org/CreativeWork/>, (参照 2018-02-08).
- [12] 河村 郁江. “2017 年度版もちマップ入力フォーム”. http://phirip.com/mochi/photo_up04_04_o/, (参照 2018-02-08).
- [13] 共通語彙基盤：IPA 独立行政法人 情報処理推進機構. <http://phirip.com/mochimap/e/>, (参照 2018-02-08).
- [14] World Wide Web Consortium. Resource Description Framework (RDF) Model and Syntax Specification. <https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>, (参照 2018-02-08).
- [15] SPARQL 1.1 Overview. <https://www.w3.org/TR/sparql11-overview/>, (参照 2018-02-08).
- [16] OpenLink Software. Virtuoso. <https://virtuoso.openlinksw.com/>, (参照 2018-02-08).
- [17] DBpedia Japanese. <http://ja.dbpedia.org/>, (参照 2018-02-08).
- [18] REACT. <https://reactjs.org/>, (参照 2018-02-08).
- [19] Google LLC. Google Maps API. <https://developers.google.com/maps/?hl=ja>, (参照 2018-02-08).
- [20] 総務省統計局. e-Stat 統計 LOD. <http://data.e-stat.go.jp/lodw/>, (参照 2018-02-08).
- [21] 経済産業省,内閣官房. RESAS 地域経済分析システム. <https://resas.go.jp>, (参照 2018-02-08).
- [22] Mike Bostock. D3.js. <https://d3js.org/>, (参照 2018-02-08).
- [23] Toshiaki Katayama. <http://biohackathon.org/d3sparql/>, (参照 2018-02-08).
- [24] 加藤 文彦 (著)、川島 秀一 (著)、岡別府 陽子 (著)、山本 泰智 (著)、片山 俊明 (著). オープンデータ時代の標準 Web API SPARQL, インプレス R&D[NextPublishing], 2015-11-13.