

RDF による料理レシピデータの構造化に関する研究

Study about the Data Structure for Cooking Recipes Based on RDF

浦川 真^{*1}
Makoto Urakawa

宮崎 勝^{*1}
Masaru Miyazaki

藤沢 寛^{*1}
Hiroshi Fujisawa

苗村 昌秀^{*1}
Masahide Naemura

^{*1} NHK放送技術研究所
NHK Science & Technical Research Laboratories

Demands for cooking recipes have been increasing in order to utilize them for keeping people healthy. And the foods used in recipes also have got a lot of attention because they are highly nutritious especially in season. The foods have regional characteristics such as different names and seasons. This document describes the data structure to describe regional characteristics based on RDF(Resource Description Framework). And it shows some demo applications to destructure structured data for recipe contents.

1. はじめに

料理レシピを投稿できるWEBサイトの増加により、料理レシピの数が膨大になっている。また、料理番組は放送局でも認知度の高い番組であるだけでなく、健康番組でも食材やそのレシピを取り扱うことが多い。レシピコンテンツ数が増えるにつれて、より簡単に、よりニーズに合ったコンテンツ展開を可能とする必要がある。

料理レシピの中心となる食材は、地域によって呼び方が違う地域性と、時期によって栄養価が違う季節性を持ち合わせている。さらに、旬となる時期については、地域によって異なるといった地域性も持っている。その一方で、放送局の料理レシピには食材が記載されているだけでなく、放送エリアや放送日といった地域と日付を情報として持っている。また、食材を販売するスーパーマーケットにおいても、販売エリアと販売日といった地域と日付を情報として持っている。このように、「食材」データに、「地域性」と「季節性」を表すデータを定義し、各サービスが利用できるデータ構造にすることで、放送局レシピとチラシサービスの連携といった様々なサービスが容易に実現できると考えられる。

本論文では、基本となる食材データの言い換えや旬の情報を、地域や時期の観点で構造化し、あわせて料理レシピやスーパーマーケット情報も、食材、地域、時期で構造化する手法について述べる。さらに、これらの情報を DBpedia リソースで一意化しオントロジーとして構造化することで、レシピとチラシの連携といった新しいサービスが可能であることを提案する。その結果として、より地域に根差した料理レシピの展開が可能となる。

本論文では、2章で本論文と関連する先行研究を紹介し、3章から5章で、食材データ、料理レシピデータ、販売情報データの構造化について述べる。その後、6章においてデータ活用アプリケーションについて紹介し、7章でデータ構造について考察し、8章で本論文のまとめと今後の課題を示す。

2. 関連研究

食べ合わせを考慮した料理推薦について、高橋ら[高橋 11]は、レシピサイトや食品成分サイト、食べ合わせ情報サイトなどの Web 上のリソースをもとに、栄養価を高める食品の食べ合わせ

をデータベース化している。また、苺米ら[苺米 09]も、個々人の食事記録をもとに、栄養バランスを考慮した料理レシピの推薦による食生活支援システムについて提案している。栄養価に着目している点は本論文と同じであり、その重要性が認識できるが、本論文では、その栄養価が旬かどうかで変わることに着目している点異なり、より厳密な栄養価計算に貢献できるものとする。また、料理レシピの尺度として、岩本ら[岩本 12]は料理動作の難易度を計算し付与している。料理レシピの情報を充実させる取り組みとしては、本論文とも共通している。調理難易度や、本論文で提唱する旬度といった、新たな尺度を組み合わせることで、より高機能な検索機能が実現できる。

3. 食材データの構造化

3.1 旬の定義

旬の野菜を食べることで、より高い栄養を摂れることや、環境負担がないことが言われている。この「旬」の定義について、農林水産省では、「旬」とは自然の中でふつうに育てた野菜や果物がとれる季節や、魚がたくさんとれる季節のこと」としている¹。また、農畜産業振興機構では、収穫・出荷量が地域によって異なるため、産地によって食材の旬は異なるとしている²。農林水産省は、「うど」、「うめ」、「アスパラガス」など、約 90 の野菜・果物に関して、青果物卸売市場調査(品目別・産地別:主要卸売市場計)³を公開している。図 1 は、ブロッコリーの産地別取引量を表す。地域により旬となる時期が異なることがデータからも分かる。本論文では、平成 27 年度の産地別取引量をもとに、各産地の年間平均取引量以上の月を、その産地における旬と定義した。

| 都道府県 | 食材名 | 主要卸売市場取引合計(千) | | | | | | | | | | | |
|------|--------|---------------|------|------|-----|-----|-----|------|------|------|------|------|------|
| | | 1月 | 2月 | 3月 | 4月 | 5月 | 6月 | 7月 | 8月 | 9月 | 10月 | 11月 | 12月 |
| 福岡 | ブロッコリー | 463 | 364 | 354 | 122 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 409 |
| 香川 | ブロッコリー | 911 | 872 | 939 | 706 | 764 | 157 | 0 | 0 | 0 | 0 | 0 | 542 |
| 長野 | ブロッコリー | 0 | 0 | 0 | 0 | 0 | 959 | 632 | 99 | 639 | 688 | 0 | 0 |
| 埼玉 | ブロッコリー | 0 | 0 | 0 | 402 | 439 | 0 | 0 | 0 | 0 | 282 | 1055 | 744 |
| 鳥取 | ブロッコリー | 0 | 0 | 0 | 83 | 485 | 395 | 0 | 0 | 0 | 567 | 899 | 772 |
| 長崎 | ブロッコリー | 321 | 302 | 343 | 238 | 134 | 0 | 0 | 0 | 0 | 73 | 345 | 0 |
| 愛知 | ブロッコリー | 1147 | 1150 | 1139 | 563 | 548 | 0 | 0 | 0 | 0 | 0 | 698 | 1830 |
| 徳島 | ブロッコリー | 403 | 359 | 305 | 123 | 250 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 北海道 | ブロッコリー | 0 | 0 | 0 | 0 | 0 | 743 | 3408 | 2800 | 2833 | 1943 | 206 | 0 |

各産地の年間平均取引量以上の月

図1: 県別のブロッコリー取引量

連絡先: 浦川真, NHK放送技術研究所, 〒157-8510, 03-5494-3423, urakawa.m-gi@nhk.or.jp

¹ http://www.maff.go.jp/j/syokuiku/kodomo_navi/learn/seasons1.html

² <http://vegetable.alic.go.jp/yasaijoho/yasai/1311/yasai1.html>

³ http://www.maff.go.jp/j/tokei/syohi/shikyou/seikazunpo/h27_1.html#27h_syuyo

3.2 言い換え情報

食材は、呼び名が複数ある場合があり、一部には特定の県や地域のみで使われることもある。例えば、「田芋」は、「水芋」とも呼ばれる。さらに奄美大島では、「ターウム」と言われる。本論文では、どこでどのように言われているのかという言い換え情報についても構造化を行った。なお、この言い換え情報については、Wikipedia 等のインターネット上の記事を参考に手動で抽出した。

3.3 句と言い換え情報の構造化

上記にて取得した、句情報と言い換え情報について、構造化を行った。本論文では、「食材」と「場所」を DBpedia リソースで特定し RDF により各情報をつなげることで構造化を行った。構造化にあたり、DBpedia 内でのリソース定義を考慮する必要があり、本論文では、「<http://ja.dbpedia.org/resource/>“食材名”」が、DBpedia 内でサブジェクトとして存在する場合は、そのまま採用し、「<http://dbpedia.org/ontology/wikiPageRedirects>」プロパティを持つ場合は、そのオブジェクトをエンティティとして定義した。例えば、青果物卸売市場調査で定義されている「かんしょ」は、「<http://dbpedia.org/ontology/wikiPageRedirects>」プロパティとして得ることができる。「<http://ja.dbpedia.org/resource/>サツマイモ」をエンティティとした。また、サツマイモを Wikipedia 等で検索すると、「甘藷」や「唐芋」を別名に持つことが分かり、「甘藷」は宮崎県で使われていることが分かる。図 2 に、食材構造化データの例を示す。本論文では、RDF データを記述するオントロジーの名前空間として [fd](http://food.nhk.or.jp) (<http://food.nhk.or.jp>) を定義したが、研究用として構築したため公開していない。

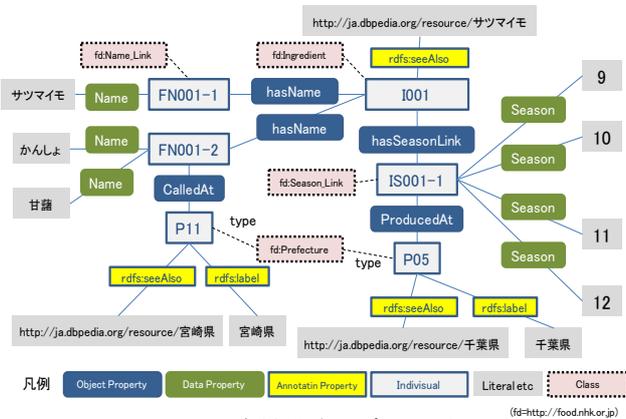


図 2 食材構造化データの例

4. 料理レシピデータの構造化

NHK 北九州放送局が公開している 177 の料理レシピ¹についても、レシピ内に記載されている「食材」や、レシピの放送日及び放送エリアをもとに構造化を行った。レシピで使われる食材は、「ゆでだこ」といった状態を表す語が付き表記が変わることや、「しょうが(すりおろす)」といった調理動作が合わせて表記されることがある。そのため、Mecab²により形態素解析をおこなった上で、「だこ」を「たこ」と変換する処理や、「()」とその中の単語を除外する処理などを前処理として実装した。その上で、DBpedia リソースである「<http://ja.dbpedia.org/resource/>」で食材をエンティティとして一意化した。3 章にて、食材構造化データを生成する際に実装した処理と同じように、[wikiPageRedirects](http://dbpedia.org/ontology/wikiPageRedirects)

¹ http://www.nhk.or.jp/kitakyushu/recipe_top/index.html

² <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>

が存在した場合は、そのオブジェクトをエンティティとして採用した。レシピ構造化データを記述するための名前空間として、[nfr](http://nfr.nhk.or.jp) (<http://nfr.nhk.or.jp>) を定義したが、研究用として構築したため公開していない。

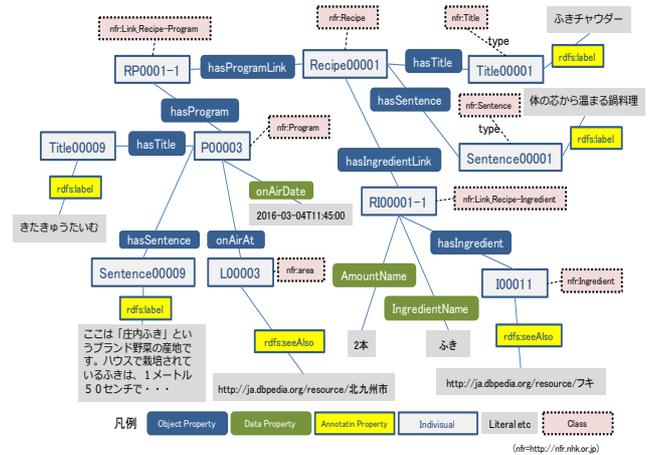


図 3 レシピ構造化データの例

5. 食材販売情報の構造化

食材を扱うスーパーマーケットにおいても、食材の特売日や販売エリアといった時期や地域を情報として持っている。本論文では、架空のスーパーマーケットの販売情報を手動で作成した。その販売情報構造化データを図 4 に示す。食材構造化データや料理レシピ構造化データ同様に、販売情報構造化データにおいても、リソースを DBpedia リソースで一意化した。レシピ構造化データを記述するための名前空間として、[gs](http://gs.nhk.or.jp) (<http://gs.nhk.or.jp>) を定義したが、研究用として構築したため公開していない。

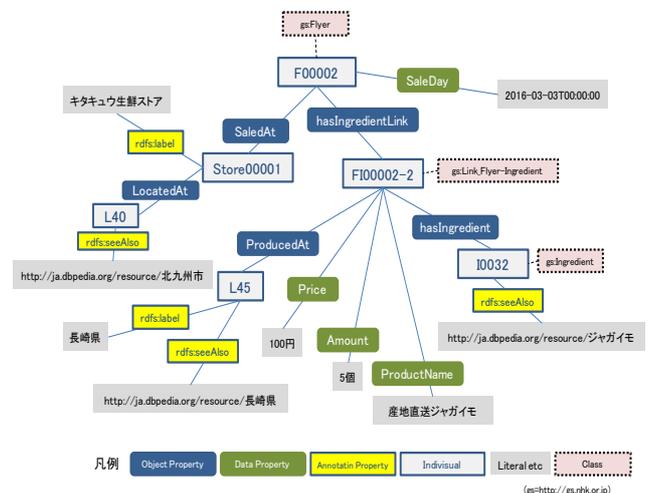


図 4 販売情報構造化データの例

6. データを活用したアプリケーション

3 章から 5 章で構造化したデータの連携イメージを図 4 に示す。食材データ、料理レシピデータ、チラシデータを管理する事業者は異なるため、そのエンティティは異なることが想定されるが、図 4 に示すように、共通となる情報で構造化することで、異なるサービスを連携させることが可能となる。構造化データを活用したサービス例として、旬情報等を考慮して、料理レシピコンテンツを検索できる例と、チラシサービスへのレシピコンテンツ連携例を示す。

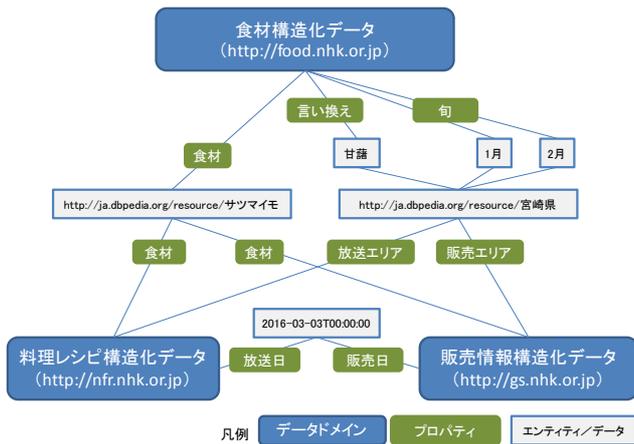


図 4 異なるデータドメイン間のデータの繋がりが

6.1 料理レシピコンテンツの検索機能の高度化

料理レシピを参照しているユーザの場所と日付により、より多くの旬食材を使ったレシピを提供するアプリケーションを試作した。そのために必要となる、2種類の構造化データ(食材構造化データ、レシピ構造化データ)からデータを検索するためのクエリ例を図5に示す。

アプリケーションからは、図5に示すように、「福岡県」をDBpedia リソースで指定し、食材構造化データ用の RDF エンドポイント(<http://food.nhk.or.jp>)に SPARQL クエリで検索することで、指定した「福岡県」の特定の月(3月)の旬の食材を、DBpedia リソースで表現されたエンティティとして取得可能である。この例では、「福岡県」で3月が旬の食材として、「<http://ja.dbpedia.org/resource/シュンギク>」等を取得できている。

さらに、取得した食材エンティティを引数として、レシピ構造化データ用の RDF エンドポイント(<http://nfr.nhk.or.jp>)を検索することで、旬の食材をどの程度使用したレシピか把握できる。本論文で示したクエリでは、使用された旬の食材数の結果が各レシピに対して得られるため、その合計をレシピで使用された食材数で除算することで、レシピの旬度を定義した。例えば、「ほうれんそうサラダ」は、3つの食材で作ることができるが、1つの旬食材(<http://ja.dbpedia.org/resource/ほうレンソウ>)を使っているため、その旬度は0.33となる。

「福岡県」で「3月」が旬の食材を検索

```

SELECT DISTINCT (STR(?URL) as ?URI)
WHERE {
  ?Food fdt:hasSeasonLink / fdt:ProducedAt / rdfs:seeAlso ?Prefecture.
  ?Food fdt:hasSeasonLink / fdt:Season ?Month.
  FILTER(STR(?Prefecture)="http://ja.dbpedia.org/resource/福岡県").
  FILTER(?Month="3"^^xsd:decimal).
  ?Food rdfs:seeAlso ?URL

```

「福岡県」で「3月」が旬の食材を使っているレシピを検索

```

SELECT (count(?link) as ?count) ?ingredient_count ?recipe_id (STR(?recipe_name) as ?Title) (STR(?name) as ?Ingredient_id)
WHERE {
  (SELECT ?recipe_id (COUNT(?recipe_name) AS ?ingredient_count)
  WHERE {
    ?recipe_id nfr:hasIngredientsLink / nfr:hasIngredient / rdfs:label ?recipe_name.
  } GROUP BY ?recipe_id)
  ?recipe_id nfr:hasTitle / rdfs:label ?recipe_name.
  ?recipe_id nfr:hasIngredientsLink ?link.
  ?link nfr:hasIngredient / rdfs:seeAlso ?name.
  FILTER(STR(?name) IN ("http://ja.dbpedia.org/resource/キウイフルーツ", "http://ja.dbpedia.org/resource/イチゴ", "http://ja.dbpedia.org/resource/セロリ", "http://ja.dbpedia.org/resource/カリフラワー", "http://ja.dbpedia.org/resource/カボチャ", "http://ja.dbpedia.org/resource/コマツナ", "http://ja.dbpedia.org/resource/ほうレンソウ", "http://ja.dbpedia.org/resource/タマネギ", "http://ja.dbpedia.org/resource/シメジ", "http://ja.dbpedia.org/resource/アスパラガス", "http://ja.dbpedia.org/resource/パセリ", "http://ja.dbpedia.org/resource/キャベツ", "http://ja.dbpedia.org/resource/レタス", "http://ja.dbpedia.org/resource/アスパラガス", "http://ja.dbpedia.org/resource/シュンギク", "http://ja.dbpedia.org/resource/アスパラガス", "http://ja.dbpedia.org/resource/トマト").
} GROUP BY ?ingredient_count ?recipe_id ?recipe_name ?link ?name
ORDER BY ?recipe_id

```

図 5 2つのオントロジーからデータを取得するクエリ例

図6に、上述したクエリを元に「福岡県」で「3月」が旬の食材をより使った割合が多い料理レシピを表示するアプリケーションを示す。食材データに地域性と季節性を情報として付与し一意化することで、地域毎に異なる旬の食材に合わせた検索が可能となる。また、検索クエリとして指定するだけでなく、各レシピの旬度を計算し表示することも可能である。また、アプリケーションが指定する場所と、レシピの食材が特定できているため、その土地で呼ばれている名称に切り替えて食材を表示することも容易に可能である。



図 6 レシピ表示アプリケーション

6.2 レシピとチラシサービスの連携

食材を販売しているスーパーマーケットの場所と販売日により、より多くの特売食材を使ったレシピや、放送予定のレシピを提供するアプリケーションを試作した。

図7では、「北九州市」にある架空のスーパーマーケットの「3月3日」のチラシデータと、その放送予定として「3月4日」に「北九州市」で放送される料理番組のレシピや、特売食材をより多く使ったレシピを表示する例である。この例についても、販売情報構造化データから、特売食材や販売日、販売地域を取得した上で、レシピ情報構造化データから、販売日の翌日に放送されるレシピや、特売品をより多く含むレシピを取得する2段階のクエリで実現している。取得するレシピを、翌日の放送のレシピとするか、当日の放送のレシピとするかは、スーパーマーケット側で指定できるため、「明日放送の料理番組を見ながら作りたい」顧客向けの販売戦略などにつなげるといった、サービスの創出を促す効果も考えられる。



図 7 チラシとレシピの連携

7. 考察

7.1 旬の構造化

本論文では、農林水産省が公開している、青果物卸売市場調査(品目別・産地別;主要卸売市場計)の各産地別の平成 27 年度野菜取引数量をもとに旬を定義した。このデータは、1 か月を 3 つに分けた合計 36 の PDF ファイルに記載されているため、旬情報の構造化データ作成作業の負担が大きい。年度毎の状況をデータとして反映できる良い点はあるが、オープンデータとして公開され、システム処理できることが望ましい。また、青果物卸売市場調査対象となっている食材が約 90 と少ないため、対象となっていない食材の旬の時期については、Web 上の記事等を参考にして作成する必要がある。

7.2 リソースの一意化

本論文では、「地域」や「食材」を DBpedia リソースにより一意化を行った。共通語彙による一意化により、サービス連携の度にデータ変換等が不要となるため、他サービスへのコンテンツ展開が容易になる。その一方で、「トマト」は「<http://ja.dbpedia.org/resource/トマト>」で一意化され、「ミニトマト」も、「<http://dbpedia.org/ontology/wikiPageRedirects>」により「<http://ja.dbpedia.org/resource/トマト>」で一意化された。つまり、料理レシピやチラシに記載された「ミニトマト」と「トマト」が同じ食材として扱われることを意味する。そのため、料理レシピにおいて、「トマト」を「ミニトマト」で代用できるのかを、レシピ毎に判断する処理が必要となる。

7.3 料理レシピの構造化

本論文では、料理レシピに記載された食材を、DBpedia リソースにより一意化し、構造化を行った。また、放送日時や放送エリアといった放送情報をレシピデータに持たせることで、チラシサービスに放送予定レシピを連携させるといった新たなサービスが可能となった。料理番組を、放送日や放送エリアを指定しテレビ向けに放送するだけでなく、料理レシピを、特定の日に特定の地域のチラシサービスに展開できるため、新しい放送サービスの展開の一つとも言える。また、図 3 で、番組情報に「庄内ふき」に関する補足情報が記載されているように、料理レシピには記載されていない情報を、番組情報から提供することも可能となる。

8. おわりに

本論文では、料理レシピに関する「食材」と、その「地域性」や「季節性」に着目し、食材の言い換えや旬の情報、料理レシピの食材や放送日及び放送エリア、スーパーマーケット情報の販売食材、販売地域、販売日を構造化した。RDF によりデータを構造化し、DBpedia といったオープンデータで一意化しオントロジーを構築することで、食材に関する情報を柔軟に表現だけでなく、異なるサービス同士を容易に連携させることができた。本論文では、その一例として、放送サービスをメインに展開している放送局の料理レシピを、データ連携により外部サービスへの展開が可能となることを試作アプリケーションで示した。また、番組情報を付加できることや、地域に合わせた言い換え表現への可能性についても示した。生産技術の向上により、旬が分かりづらくなった現代において、食材の季節を意識でき、より地域に根差した料理レシピの展開が可能になることを示した。

今後の課題として、食材の一意化を、「トマト」と「ミニトマト」を区別するといった DBpedia で定義されているエンティティ以上に詳細に表現することや、地域毎の言い換え表現情報の充実などに取り組んでいく。

参考文献

- [高橋 11] 高橋淳、伊藤孝行、植田嗣也: 栄養の食べ合わせを考慮した料理推薦機構の試作, 情報処理学会研究報告, IPSJ SIG Technical Report ,2011.
- [苺米 09] 苺米志帆乃、藤井敦:料理レシピの推薦と栄養バランスの可視化による食生活支援システム, WebDB Forum, 2009.
- [岩本 12] 岩本純也、宮森恒:料理の難易度を考慮したレシピ検索システムの提案, DEIM Forum E1-3, 2012.