

レシピと食品標準成分表を連携させた食事療法支援サービス Supporting Diet Therapy Based on Japanese Linked Data

多川 勇介^{*1} 田中 改^{*1} 南 裕也^{*2} 並河 大地^{*2} 下村 道夫^{*2} 山口 高平^{*1}
Yusuke Tagawa Arata Tanaka Yuya Minami Daichi Namikawa Michio Shimomura Takahira Yamaguchi

^{*1} 慶應義塾大学
Keio University

^{*2} NTT サービスエボリューション研究所
NTT Service Evolution Laboratories

This paper presents how to develop diet therapy service, building and merging different kinds of Linked Data: restaurant menus, cooking recipes and Food Composition Database in Japan. When merging them, such properties as owl:same is less practical to exploit Linked Data. We discuss how conversion rules work well in calculating nutritional content with recipes. Furthermore, three hundred patients take our developed service for diabetes and they say that it works well.

1. はじめに

近年, Linked Open Data(LOD)が注目されており, 欧米のみならず日本国内でも電子行政オープンデータ戦略が作成された. 他のドメインデータへの連携がしやすいため, サービス機能の拡張が容易である.

一方, 近年中年や高齢者の中で生活習慣病が急増している. その治療の一つとして食事療法が用いられる. 食事療法を行うに当たって多くの知識が必要となる.

しかし, 一般的なレシピ情報には食事療法のための栄養計算まで記載しているものはほとんどない. 特に飲食店のメニューは栄養情報が記載されていないものも多く, 食事療法の負担となっているのが現状である.

本稿の目的は, ウェブ上の料理レシピ情報と日本食品標準成分表 2010 を Linked Data 化しマッピングルールを作成し, これらを連携させることで飲食店等のメニュー名と類似するレシピ情報を基に栄養成分を計算・提示する食事療法支援サービスの実現である.

2. 関連研究

栄養素等摂取バランスを考慮した料理レシピ検索システム [薊米 10]では, あるレシピや, 材料を入力すると, 栄養バランスを考慮した献立が [薊米 09]の検索システムによって検索され, それらの食事ログを取り, 「レコーディングダイエット」によってユーザの食事支援を可能にしている. このシステムでは, レシピを検索する際, 全文一致または部分一致検索が行われている. そのため, 3.2 における図4のような例など, 期待していない結果が出力される可能性がある.

3. 提案システム概要

図1に提案システム全体のシステムアーキテクチャを示す.

システムは大きく「栄養計算処理」と「レシピ推定・情報提示処理」に分けられる. 栄養計算処理では取得したレシピ情報を基に事前に栄養計算を行うものである. レシピ推定・情報提示処理はユーザからクエリが投げられた時に動的に処理を行う部分である.

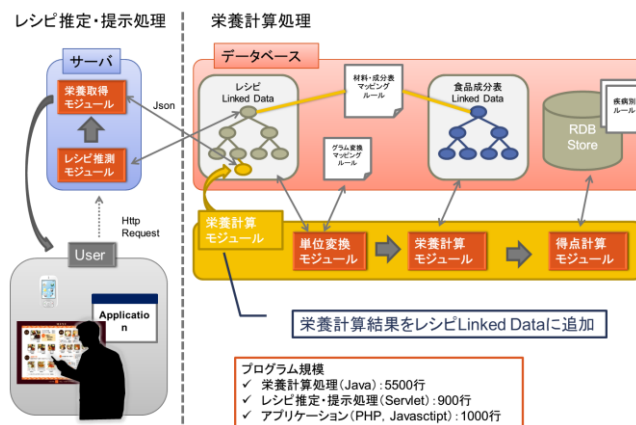


図1 システム全体アーキテクチャ

3.1 栄養計算処理

栄養計算処理では, 主にレシピデータの収集や栄養計算など, 静的な処理を行う.

(1) レシピデータ収集

レシピ情報に関してはウェブ上の以下のサイトをスクレイピングすることにより, レシピ名, 材料, 調理手順, 人数分, 料理カテゴリ, 料理写真 URL などを取得した.

表1 レシピの取得元と取得数

レシピサイト	取得レシピ数
クックパッド ^{*2}	2644
ナスラックキッチン ^{*3}	2740
レシピ大百科 ^{*4}	813

連絡先: 田中 改, 山口 高平 慶應義塾大学理工学研究科
〒223-8522 神奈川県横浜市港北区日吉 3-14-1

TEL:045-566-1614

E-mail: k_a_y_h_j_s@z6.keio.jp, yamaguti@ae.keio.ac.jp

^{*2} クックパッド: <http://cookpad.com/>

^{*3} ナスラックキッチン: <http://www.nasluck-kitchen.jp/>

^{*4} レシピ大百科: <http://park.ajinomoto.co.jp/>

(2) Linked Data 構築

(1) で取得したレシピ情報と日本標準食品成分表 2010 を Linked Data 化していく。今回作成したレシピの Linked Data の構造を図2に示す。

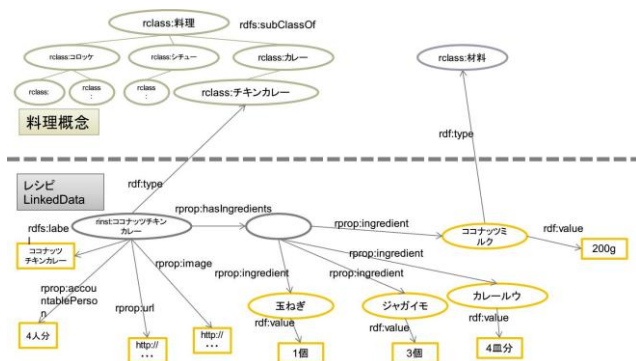


図2 LinkedData の構造一例

また、Linked Data を作成する際に、レシピサイトが構築しているレシピカテゴリを基に、料理クラスとの関連付けを行なっている。例えば上図ではココナッツチキンカレーはチキンカレーカテゴリ内のレシピであるので、チキンカレークラスに当てはめる。また、材料は全て材料クラスに対応させる。こうすることによって、「ココナッツチキンカレー」というレシピがチキンカレーの1種であり、ココナッツを材料に持っているということを定義することができる。

	トリプル数
レシピ Linked Data	79423トリプル
食品成分表 Linked Data	107412トリプル

(3) マッピングルールの作成

作成した Linked Data 間にはセマンティックギャップが生じている。そのため、機械が自動的に栄養計算を行うことができない。これを解決するためにマッピングルールを作成する。

まず、栄養計算のための材料の分量の単位変換を行うためのルールを作成する。今回、Web 上の食材の重さについて記述してあるページをスクレイピングして単位変換ルールを作成した。作成したルール数は 1080 ルールである。

次にレシピの食材と食品成分表のマッピングルールを作成する。グラム単位に統一された材料を用いて栄養計算を行う際、材料名で食品成分表を検索し、1対1で対応する項目の栄養素を基に計算を行うが、1つの食材に対して複数の項目が当てはまる場合がある。この場合にはマッピングルールが必要となる。複数項目が当てはまる場合には人手でルールを追加し、対応付けた。作成したルール数は 148 ルールである。

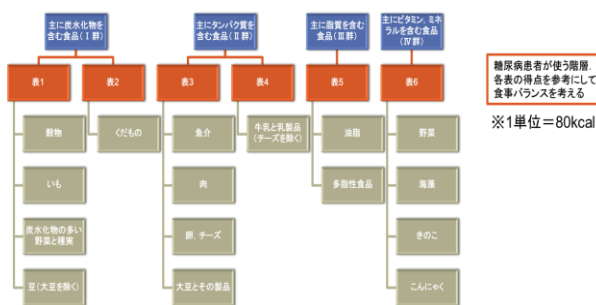


図3 食品成分表と疾病別ルールのマッピング

最後に、食品成分表と疾病別ルールのマッピングを行う。今回はケーススタディとして糖尿病を扱っているため、糖尿病のための食品交換表⁵を参考に図3のようにマッピングを行った。

栄養計算モジュール

これらのマッピングを基に、ウェブ上から取得したレシピ情報の栄養素を計算していく。

作成したレシピ Linked Data の材料の分量は表記揺れ生じているため、まず単位変換モジュールで単位を統一する。次に栄養計算モジュールで実際に栄養計算を行う。最後に、得点計算モジュールにおいて疾病別ルールに基づいた得点計算を行う。

(4) 疾病別ルール

疾病別ルールとは、書籍等を参考にしある疾病の場合どのような栄養素を重視するかなどをまとめたルールのことである。本稿では、糖尿病に対する疾病ルールを作成し、糖尿病の患者に対する食事療法支援サービスを提案した。具体的なものとしては、糖尿病であれば、図3に載っている表1～表6の得点を考慮すべきということがこのルールに当たる。また、このルールを変更することにより本システムは他の疾病にも容易に対応することができる。

3.2 レシピ推定・情報提示処理

レシピ推定・提示処理ではユーザからのクエリに対して結果を返す、動的な部分の処理を行う。

(1) ユーザデータからの基準値の設定

ユーザには事前に年齢、身長、性別、体重、運動量といったパーソナルデータを登録してもらう。このデータを基に、目標値を設定する。

タンパク質や脂質などの基本的な栄養素に関する基準は、食事摂取基準⁵を基に設定する。糖尿病のための得点はユーザに直接入力してもらうか、ユーザの身長から一日の指示カロリーを算出し、それに沿って一日の指示単位を設定する。

(2) レシピ推測モジュール

入力された飲食店のメニュー名から、類似するレシピを推測する。

飲食店のメニュー名は画像から OCR 認識したメニューテキストとして与えられる。そのため、メニュー名に関連のあるものばかりではない。そのためレシピ名推測モジュールでまず前処理をおこなう。

ここではメニュー表の OCR 認識結果のテキストを生テキストと呼ぶ。生テキストには複数のメニュー名や値段、ノイズ等が含まれている。一般的にメニュー表内のメニュー名は分割して記述されているので、「改行記号」や「空白スペース」、「・」などの記号で文字列を分割し、分割されたひとかたまりの文字列をメニュー名とする。

次にメニュー名とレシピをマッピングするために 3 段階のマッピングを行う。

①メニュー名とレシピ名の完全一致でのマッピング

まず、入力されたメニュー名とレシピ Linked Data 内のレシピ名を照合し、完全一致するものがあれば対応レシピとする。完

⁵ 日本糖尿病学会(2002)『糖尿病のための食品交換表』文光堂

全一致するレシピ名があった場合には最も合致するものと考えられるので、意味処理は行わない。

②意味処理を加味したマッピング

次に、意味処理を加えてマッピングを試みる。メニュー名とレシピ名が完全一致しなかった場合、メニュー名を分割しなければならないが、メニュー名には「定番料理名」「材料名」「その他」のように、レシピ内容と関係の薄い語句も含まれる。メニュー名を形態素解析し、名詞句のみを検索対象とすることも解決策の1つに挙げられるが、「シェフのこだわりカレー」と入力された場合に「シェフ」と「カレー」が名詞句としてマッチしてしまい、期待できる結果が得られない。また、形態素解析辞書を料理用に組み替えて、名詞句が「カレー」と抽出出来たとしても、図4にあるような例のように一般的に想像される「カレー」にマッチするとは限らない。

そこで、抽出された名詞句のなかで料理クラスに含まれるものと材料クラスに含まれるものに向け、検索対象を変えることで、目的とするレシピに近いレシピが得られるようにした。

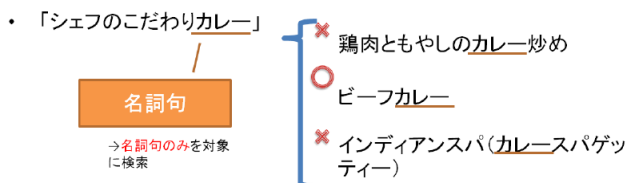


図4 カレーという単語が含まれるレシピ名

レシピの料理クラスに関してはレシピサイトのカテゴリ階層を基に定義した。以下に一部を示す。

表2 カテゴリ階層から作成した料理クラス

みぞれ鍋	スープパスタ	かぼちゃコロッケ
ポトフ	スープ	ガスパチョ
ポテトサラダ	しゃぶしゃぶ	味噌汁
ポテトコロッケ	シチュー	お好み焼き
ポタージュ	サラダ	お吸い物
ピシソワーズ	さつまいもコロッケ	オムレツ
ハンバーグ	コンソメスープ	オムライス

③メニュー名の名詞句のみを部分一致でのマッピング

メニュー名の名詞句のみを部分一致でマッピングするものは、料理クラスが含まれていなかったものである。前述した2通りのやり方でレシピが見つからなかった場合にメニュー名の名詞句のみ抽出し、レシピ名と部分一致でのマッピングを行う。複数レシピが結果として得られた場合には、メニュー名とレシピ名の編集距離を算出して、最も類似する1レシピを栄養計算対象レシピとする。

(3) 栄養取得モジュール

栄養取得モジュールでは、レシピ推測モジュールにて与えられた、対象レシピの栄養情報をレシピ LinkedData より取得する。そして結果の値をアプリケーションに json 形式で渡し、アプリケーション側でグラフ表示する。

3.3 評価実験

今回、既存サービスとの比較とウェブアンケートによる評価を行った。

(1) 既存サービスとの比較

類似サービスとして、ユーザ投稿型レシピサイト、サイト側提供型レシピサイト、栄養計算ソフトとの比較を行った。それを表にしたのが表3である。

既存サービスでは、レシピサイト、栄養計算ソフト共に強みと弱みがある。提案サービスはこれらの良い部分、レシピサイトのレシピ情報と、栄養計算ソフトのような計算機能を持ち合わせたものとなっているため有用であるといえる。

表3 既存サービスとの比較

	ユーザ投稿型レシピサイト	サイト側提供型レシピサイト	栄養計算ソフト	提案サービス
1 製品(Product)				
レシピ数	◎	○	△	○
栄養計算	×	×	◎	○
検索自由度	○	△	×	◎
拡張性	×	×	△	◎
2 価格(Price)				
価格	一部有料	無料	一部有料	未定
3 場所(Place)				
デバイス	PC・スマホ	PC・スマホ	PC	PC・スマホ
4 プロモーション(Promotion) (該当なし)				
5 人(要員) (Personnel)				
レシピ投稿者	△	○	◎	△
6 プロセス(Process)				
サービス利用	インターネット	インターネット	ローカル	インターネット
栄養計算の時間	×	×	◎~×	◎
7 物的証拠(Physical Evidence)				
内容の保証	△	○	◎	○

(2) アンケート評価

提案システムをウェブ上で試用してもらったアンケートを行った。調査対象者は

1. 糖尿病にかかった、かかっている
2. スマートフォンや携帯電話で一日30分以上ネットサーフィンをする
3. 食事療法を行なっている、知っているがやったことはない、行なっていたがやめた

これらの条件下で、3. の条件によって分別し、各100人ずつサンプル回収を行った。

実際に利用してもらった画面は以下のような画面である。



図5 デモ画面

アンケートではまず、ニーズの洗い出しを行い、次に提案システムを実際に利用してもらい既存サービスとの比較評価を行った。

ニーズの洗い出しでは以下のような評価が得られた。

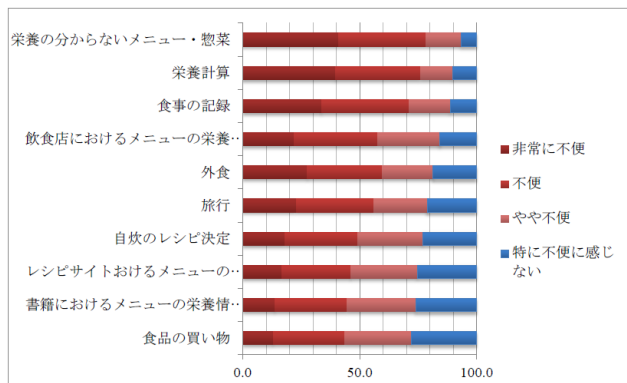


図6 ニーズの洗い出し

各項目で不便と感じる人が 8 割程であり、多くの人が食事療法について不便に感じていることがわかる。特に、「栄養の分からないメニュー・惣菜」と「栄養計算」の項目では「非常に不便」と回答している人が 4 割おり、「特に不便と感じない」と回答した人が 1 割以下という結果になった。これらのことから、メニューや惣菜の栄養素の情報は欲しいが、栄養計算は不便であるということがわかる。

次に、実際に提案システムを利用してもらった上での評価である。メニュー名を固定した場合と、自由にメニュー名を入力してもらった場合の 2 パターンの方法で評価を行った。

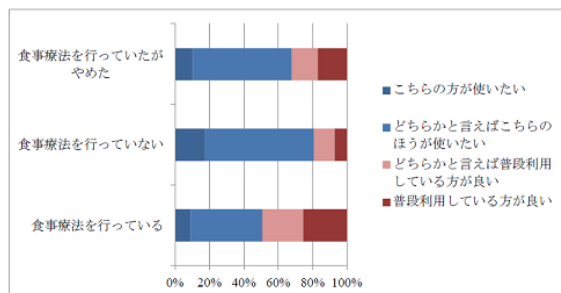


図7 メニュー名を固定した場合の評価

まずメニュー名を固定した場合だが、「食事療法を行っていたがやめた」人の約 7 割が提案サービスを利用したいと答えていることから、以前食事療法をやろうとして挫折した人に対して有用であると考えられる。自由記述においても「非常に便利。例え少々誤差があるとしてもこの簡便さはすごい。」というように利便性に高い評価を得ることができた。「食事療法を行っていない」人に関しては約半数の人が提案サービスのほうが使いたいと評価した。特にいつもの使い慣れているサイトのほうがわかりやすいという意見も見られた。

次に、メニュー名を自由入力してもらった場合の評価である。

この場合、メニュー名固定の場合に比べて妥当であるという評価が少し下がった。この原因の 1 つとして、メニュー名でない入力語が挙げられる。「ココイチ」のような店舗名や、「ビックマック」のような固有料理名を入力しているために妥当性が下がったと思われる。具体的には「未対応レシピが多すぎる」という意見も

あった。これは、登録した料理クラスの個数がまだ不十分であったことと、「ラーメン」というレシピがなかったために最も類似するラーメンとして「酢ラーメン」というものが結果として出てしまったことが大きな要因と考えられる。したがって、「ラーメン」のような漠然としたメニューの入力の場合に、「一般的なラーメン」が反映されるように定義していく必要がある。

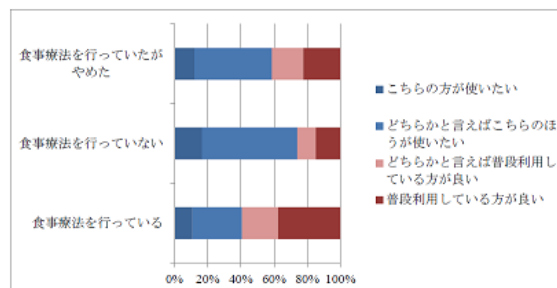


図8 自由入力の場合の評価

4. おわりに

本稿では、レシピと食品成分表を Linked Data 化し、両データ間における成分単位表現上の不整合を単位変換ルールにより解消し、データ連携をはかるとともに、疾患別ルールベースを適用する事により、疾患ごとに、ユーザに対して適切なメニューを提示できるサービスを実現した。本事例は、糖尿病患者向けの食事療法支援サービスであるが、食事療法を必要とする他の病気(例えば高血圧)にも拡張する場合、疾患別ルールベースのみを交換すればよく、システムの拡張容易性が高いといえる。また、食物アレルギー患者への対応という機能拡張を考えた場合、食品成分表 Linked Data とアレルギーの原因となる物質である Allergy Linked Data (現存するかどうかは未調査)を連携すればよく、Linked Data によるシステム拡張容易性が示されるケースである。現在、このようなシステム拡張容易性についての実証を進めている。

参考文献

[菊米 09] 菊米 志帆乃, 藤井 敦: 栄養素等摂取バランスを考慮した料理レシピ検索システム, 電子情報通信学会論文誌, Vol.J92-D, No.7, pp.975-983, 2009.
 [菊米 10] 菊米 志帆乃, 藤井 敦: 栄養素等摂取バランスの分析に基づく食生活支援システム, DBSJ Journal, Vol.8, No.4 March,2010