

オノマトペロリ：オノマトペを利用した 料理推薦システム

ラートサムルアイパン カンウィパー[†]
渡辺知恵美[†] 中村聡史^{††}

日本では料理や味覚を擬音語、擬態語を表すオノマトペを用いて「ふわふわケーキ」のように表現することが多い。そこで本研究では、オノマトペを利用した料理レシピ推薦システムの開発をしている。具体的には Web 上の料理掲載ページよりレシピを収集し、レシピ内の文章を解析する。それにより、オノマトペと固有名詞（料理名、食材など）、形容詞、一般名詞、動詞の関連性を数値化する。これらの関連より、レシピに含まれる語からオノマトペとの関連度を算出することで、キーワードサーチと比べて精度の高いレシピ検索システムを目指す。

Onomatoperori : Recipe Recommendation System Using Onomatopoeic Words

Kanwipa Lertsumruaypun[†] Chiemi Watanabe[†]
and Satoshi Nakamura^{††}

Japanese language is filled with onomatopoeic words, which describe sounds or actions like “click” or “bow-wow”. Especially, when talking about foods, Japanese people frequently use onomatopoeic words to express vague taste or sense of the foods. There, we develop a system for searching recipes by using onomatopoeic words as the search keyword. In concrete, we collect recipes from a recipe website, and perform Japanese language morphological analysis, then store the words from the analysis into the database based on their part of the speech. Next, we calculate relation between onomatopoeic words and other types of words. Last, we use the value of relation to find matched recipes to recommend to the user.

1. はじめに

近年 Web ページの爆発的増加と検索エンジンの発展により、分からないことがあればすぐに検索エンジンを利用することが常識になりつつある。しかしながら検索目的があいまいで、頭にイメージがあるがキーワードをうまく言語化できないといった経験がないだろうか。そのような具体的な検索キーワードが思い浮かばない場合に誘導的にユーザの必要とするものを発見できるような検索手法に注目が集まっている[1][4][5][6]。

我々はその解決法としてオノマトペ（擬音語、擬態語）に着目している。日本語はオノマトペ（特に擬態語）の種類は豊富であり、物事の漠然としたイメージを表現する際にオノマトペを用いることが多い。たとえば「ふわふわ、ひらひらした洋服のブランド」「カリカリでピリッと来る鳥肉料理」「ぼわーんとした感じの人」などである。本研究ではオノマトペをキーワードにした検索システムを開発することにより、ユーザが求める物事を段々明確にしていくことを目指す。その手掛かりとして我々は「料理」を対象とする。なぜなら、料理に関して「とろとろとした料理」「じっくり焼く」などのように、味や感触などの表現にオノマトペが非常に利用されるからである。

そこで我々は投稿型料理レシピサイトを利用しオノマトペから料理レシピが検索できるようなシステム「オノマトペロリ」を開発することとした。本稿では、オノマトペロリに関する諸検討について述べる。

一般に料理レシピサイトでもキーワード検索機能はある。しかしながら必ずしもすべての人がレシピにオノマトペによる表現を記述するわけではないため、検索結果が十分に現れない。そこで我々はあらかじめ料理に関する各用語（材料、調理法、食感）とオノマトペの共起関係を求めておき、その関係を用いて検索キーワードの展開、もしくはレシピとオノマトペの関係を導出することにより豊富で確かな検索を目指す。

本稿では特にオノマトペと料理に関する各用語との対応関係を投稿レシピサイトから導出したのでその結果を報告する。またその結果に基づき、オノマトペを用いた検索手法について述べる。

2. Cookpad:投稿型レシピサイト

Cookpad とは[2]、ユーザによって自分のオリジナル料理レシピを投稿・検索できるウェブサイトであり、1998 年に開発された。現在の登録されたレシピの数は 56 万個を

[†] お茶の水女子大学大学院人間文化創成科学研究科
Graduate School of Humanities and Sciences, Ochanomizu University

^{††} 京都大学大学院情報学研究科
Department of Social Informatics, Graduate School of Informatics, Kyoto University

超え、月間約 681 万人のユーザが利用している。

Cookpad は一般の料理レシピと比べて、オノマトペによる表現が多い。なぜなら、オノマトペは拙い印象を与えてしまうため書籍ではできるだけ控える傾向にあるが、投稿サイトでは逆に料理の様子を伝えるために積極的に利用する傾向があるからである。

上記の通り、Cookpad は膨大な料理レシピ数を持ち、また多数のユーザによって投稿される多くのオノマトペが使用されている料理レシピが集まっているため、本研究は Cookpad 上のレシピデータを使用する。

3. オノマトペロリ

3.1 従来の料理検索でオノマトペ検索する際の問題

本節では既存の料理レシピ検索でオノマトペ検索をする際の問題点について述べる。図 1 は Cookpad の検索画面である。ユーザが検索キーワードの欄に「ピリッ」というキーワードを記入すると、「ぴったり」と「じっくり」の 2 つの結果のタイプが返される。「ぴったり」タイプの結果とは、記入した検索キーワードが料理レシピのタイトルに含まれているものである。一方、「じっくり」タイプの結果は検索キーワードがタイトルに含まれるか、またはレシピ内（料理の説明や調理法など）に含まれているものである。



図 1 Cookpad での料理レシピ

このような検索では、オノマトペが含まれていないと検索結果として検索されないという問題がある。例えば図 2 で示している「蒸しナス香味しょうゆだれ」のレシピは、材料に「唐辛子」や「ラー油」が含まれ、ピリッとした味が予想できるが、レシピ内の文章には「ピリッ」のキーワードが含まれないため結果として検索されない。



図 2 Cookpad では検索されないレシピの例

3.2 オノマトペレシピの基本方針

本研究ではレシピに含まれる、料理に関する用語を考慮して、キーワードのオノマトペに関連度が高いレシピを推薦することで、従来のキーワードサーチより高精度なレシピ推薦を目指す。

料理に関する用語とはレシピに登場する材料や料理名、料理器具、調理法、味や食感などをさす。我々は、レシピページを収集し、料理に関する用語とオノマトペの共起度をレシピページから算出する。食材「唐辛子」と「ピリッ」の共起度は強く、「ふわふわ」との共起度は弱いなどの値により、オノマトペと用語との関連が求められる。次にこれらの値を用いてオノマトペによる検索を行う。

我々は以下の 2 種類の検索アプローチを検討する。

(1) アプローチ 1：オノマトペを他の用語に展開する

検索キーワードにオノマトペが含まれている際に、オノマトペに関連の強い用語に置き換えて検索する。たとえば「ピリッ なす」で検索した場合、「ピリッ」と関連の強い「唐辛子」「山椒」などに置き換える。

(2) アプローチ2：レシピとオノマトペの関連度を算出

各用語とオノマトペとの関係を用いて、レシピとオノマトペの関連度を算出し検索の際にそれらを用いて検索およびランキングを行う。

次節以降でオノマトペと用語の共起度の抽出および2種類の検索アプローチについて詳細に述べる。

4. オノマトペロリ料理に関する用語とオノマトペとの関連

4.1 レシピの構造と料理に関する用語

料理に関する用語について考察するにあたり、我々はまず Cookpad で掲載されているレシピの構造を解析した。

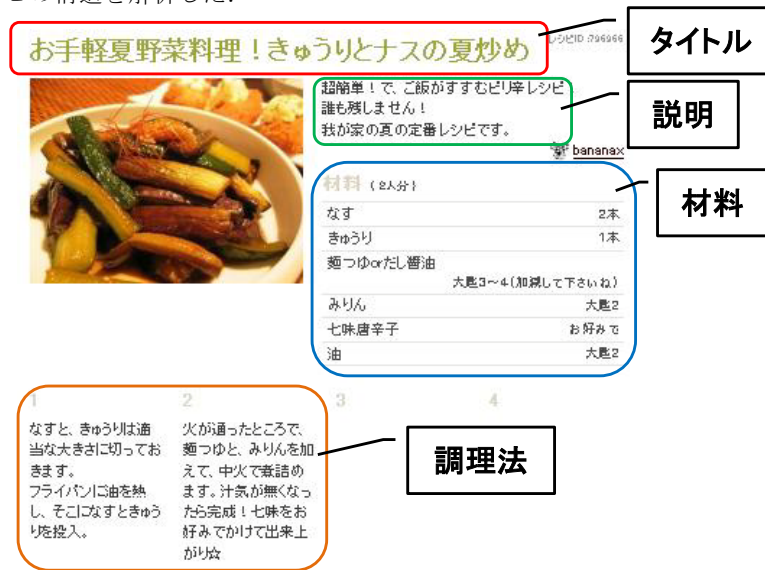


図 3 料理レシピのパーツ

料理レシピのページにはタイトル、料理の説明、材料そして、調理法のそれぞれのパーツから構成される。我々は料理レシピのそれぞれのパーツを収集し、表 1. (a),(b),(c)で示しているようにデータベースに格納する。

表 1 データベースに格納されている料理レシピのデータ例

(a) テーブル recipe のデータ例

id	recipe_nbr	title	description
3115	551283	筍のピリ辛煮	タケノコの水煮をピリ辛のきんぴら風に炊いてみました。
3156	547568	しっとり♪ふわふわ♪シフォンケーキ	しっとり、ふわふわ～♪食べて幸せ～なレシピです(^o^)

(b) テーブル ingredient のデータ例

id	recipe_nbr	item	quantity
22672	551283	タケノコ(水煮)	一個(握り拳大)
22673	551283	コンニャク	1枚
22674	551283	人参	1/3本

(c) テーブル step のデータ例

id	recipe_nbr	step_nbr	step_desc
15084	551283	1	タケノコ、人参は乱切りにする。インゲン豆は3センチくらいに切りそろえる。コンニャクはサイコロ状に切っておく。鷹の爪は種をとり小口切り。
15085	551283	2	鍋にごま油を熱し、鷹の爪を入れる。コンニャク、人参、タケノコ、インゲン豆の順に加え炒める。
15086	551283	3	だし汁と調味料を加え、煮汁が少なくなるまで煮る。最後に胡麻を加える。

表 4.で示しているように、テーブル recipe にはレシピの番号、タイトルそして、料理の説明が格納される。また、テーブル ingredient にはレシピ番号、そのレシピで使われる食材とその分量が格納される。さらに、テーブル step には各レシピの調理法そして、調理法の順番が格納される。

そして、レシピから料理に関する用語を抽出する．具体的な例を表 2. で表す．

表 2 料理レシピの重要な要素

料理レシピの要素	例
材料	七味唐辛子, 油, なす
料理名	炒め
器具	フライパン
調理法	切る, 煮詰める
味	辛
感触, 食感	柔らかい

さらに、レシピから抽出した要素を品詞別で分類すると、「材料」、「料理名」、「器具」が名詞、「調理法」が動詞、「味」、「感触, 食感」が形容詞に分類できる．よって、図 4. のような用語の構成が考えられる．

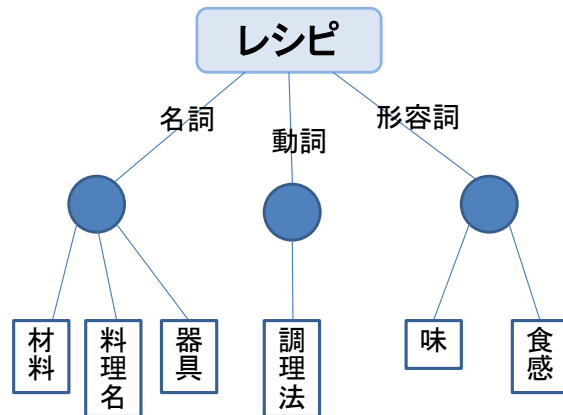


図 4 レシピの構造図

材料、料理名、調理法、器具、味などはある程度用語が決まっているため、それぞれの語をあらかじめ辞書としてデータベースに登録しておく．ただし食材や料理名などは種類が豊富であるため、辞書登録されていない用語も登場する可能性があるため、辞書にない語とオノマトペとの関連も求めておく必要がある．

レシピから検出したデータを「名詞」、「動詞」、「形容詞」そして、「オノマトペ」である 4 つの品詞別でデータベースに格納する．

例として、表 3. は名詞である用語が格納されているテーブル `recipe_noun` のデータ例を示す．具体的には、`id=11` のタプルは「レシピ番号 551283 の食材(item)のパーツから抽出され、材料(food)である一般名詞のタケノコ」が格納されている．

表 3 Recipe_noun テーブルのデータ例

id	recipe_nbr	word_noun	word_type	sentence_source	term
11	551283	タケノコ	名詞-一般	item	food
13	551283	コンニャク	名詞-一般	item	food
109	551196	スプーン	名詞-一般	step_3	dishware

また、オノマトペが格納されている `recipe_ono` テーブルのデータ例を表 4. に示す．

表 4 Recipe_ono テーブルのデータ例

id	recipe_nbr	word_ono	word_type	sentence_source
1	551196	しんなり	副詞-一般	step_2
2	548461	さっぱり	副詞-一般	description
3	547344	パクパク	副詞-一般	description

4.2 オノマトペとレシピ内の用語の関連付け

オノマトペから料理レシピを検索するには、オノマトペとレシピ内の用語の関連を求める必要がある．オノマトペと関連度の高い用語の特徴は以下の 2 点が挙げられる．[3]

- そのオノマトペと共に同じレシピかつ同じパーツに出現することが多い．
- 他のオノマトペと同じレシピのパーツに出現することが少ない

そこで今回は TF-IDF 法を利用する．TF-IDF (Term Frequency-Inverse Document Frequency) とは、「ある文章のグループでの出現数が多い」単語のうち、「他の文章のグループでの出現数が少ない」ものをその文章のグループに特徴的な単語とする、というものである．

ここでは、レシピに含まれる単語 w のオノマトペ o に対する関連度 $Weight(o, w)$ を以下のように定義する。

$$Weight(o, w) = TF(o, w) * IDF(o, w)$$

● TF

$$TF(o, w) = \frac{\text{オノマトペ}o\text{を含む全レシピの用語}w\text{の出現数}}{\text{オノマトペ}o\text{を含む全レシピの用語の出現数}}$$

TF の値が大きいほど用語 w がオノマトペ o と関連があるといえる。例えば、「ピリッ」と同じパーツに登場する材料の数が 100 個あり、そのうち「唐辛子」という調理法の出現数が 16 個、「塩」という材料の出現数が 65 個あったなら、 $TF(\text{ピリッ}, \text{唐辛子}) = 16/100$, $TF(\text{ピリッ}, \text{塩}) = 65/100$ となる。

● IDF

$$IDF(o, w) = \log \frac{\text{全オノマトペの種類数}}{\text{用語}w\text{と同じレシピに出現するオノマトペの種類数}}$$

用語 w をレシピのパーツに含むオノマトペが多ければ多いほど、用語 w がオノマトペ o にとって特有ではない、ということを表している。例えば、先ほどのように「ピリッ」を含むレシピのパーツ中に「唐辛子」と「塩」という調理法を表す用語が含まれたとする。また、全オノマトペ (800 個) が出現しているレシピのパーツに「唐辛子」と「塩」という用語が含まれているか調べると「唐辛子」は 10 個のオノマトペのレシピパーツに、「塩」は 700 個のオノマトペと同じパーツに出現した。そのとき、 $IDF(\text{ピリッ}, \text{唐辛子}) = \log(800/20)$, $IDF(\text{ピリッ}, \text{塩}) = \log(800/700)$ となり、「唐辛子」のほうが「塩」より「ピリッ」にとって特徴的であり、関連度が高いと判断される。

4.3 実験：オノマトペと用語との関連度

今回はオノマトペとレシピ内の用語の関連を調べるために Cookpad の 1,157 個のレシピを収集し、レシピの用語をデータベース化した。そして、TF-IDF 法により、用語の要素の種類別でオノマトペと用語の関連度を計算した。例として、「ふわふわ」、「シャキシャキ」、「もちもち」そして「ピリッ」の 4 つのオノマトペに対して、TF-IDF 値の上位 3 種類の食材を以下に示す。

表 5 オノマトペに対する食材の TF-IDF 値

オノマトペ	出現するレシピの件数 (全1,157件中)	食材	TF-IDF値
ふわふわ	37	卵	0.127871
		卵白	0.127467
		バター	0.114384
シャキシャキ	40	レタス	0.191189
		セロリ	0.143027
		梅	0.125704
もちもち	15	お湯	0.195347
		白玉粉	0.157570
		水	0.121406
ピリッ	7	肉	0.266218
		きゅうり	0.170858
		梅	0.156363

表 5. の結果より、「シャキシャキ」に対して、レタスやセロリなどのシャキシャキとする食感を与えるような食材が上位の TF-IDF 値をもつが、「ピリッ」に対して、肉やきゅうりなどの比較的一般的な食材が上位の結果となっている。これは、今回使用された 1,157 個のレシピ中「ふわふわ」が含まれるレシピは 37 個であることにに対して、「ピリッ」が含まれるレシピは 7 個であり、「ピリッ」が使われているレシピデータ数が不足していたためと考えられる。よって、実験に使用されたレシピデータの数を増加させることによって、高精度のオノマトペとレシピ内の用語の関連度が得られると考えられる。

次節では、本実験によるオノマトペとレシピ内の用語の関連度を用いて、オノマトペロリにおけるレシピ検索アプローチについて述べる。

5. オノマトペロリにおけるレシピ検索アプローチ

第 3.2 章でも述べたように、TF-IDF 法を適用したオノマトペとレシピ内の用語の関連度の計算結果を用いて、オノマトペロリにおいて 2 つのレシピ検索アプローチを提

案する．一つ目はオノマトペを他の用語に展開する方法である．これは、いわゆる図 4. のレシピの構造図をトップダウンに検索を行うと考えられる．二つ目のアプローチは、レシピとオノマトペの関連度を算出する方法である．これも一つ目と同様に図 4. のレシピの構造図をボトムアップに検索を行うと考えられる．

次に、2 つの検索アプローチの具体的な方法及び実際のレシピ検索結果・評価について述べる．

5.1 アプローチ 1：レシピの構造図のトップダウンで検索

アプローチ 1 は、オノマトペに関連の強い用語の置き換える．なお、置き換える用語は図 4. に示したレシピ構造図をトップダウンして、各レシピの要素の用語に展開する．例えば、名詞である要素の場合は、「食材」、「料理名」そして、「器具」のグループの用語に置き換える．具体例として、「シャキシャキ サラダ」を検索キーワードにしたときに、「シャキシャキ」を食材である用語に展開すると、表 4. に示した結果より、「シャキシャキ」に関連の強い「レタス」、「セロリ」そして、「梅」に置き換える．よって、オノマトペロリでは「レタス サラダ」、「セロリ サラダ」そして「梅サラダ」をキーワードにしてレシピ検索を行う．

検索結果例として、従来の Cookpad において「シャキシャキ サラダ」を検索キーワードにした際に検索されない「レタスとトマトのさっぱりナムル風サラダ」のレシピ内容を図 5. に示す．

レタスとトマトのさっぱりナムル風サラダ

レシピID: 886644



ゴマの香りが美味しいサラダ感覚のナムルです。
トマトがとっても美味しくなりますよ♪

chi-chan

材料

レタス	4枚
トマト	大1個
☆塩	小さじ1/2
☆ごま油	大さじ1と2/3
☆すりごま	大さじ1
にんにく	小さじ1/2

1 トマトは食べやすい大きさにスライス、レタスは洗って水気を切っておきます。

2 にんにくはすりおろし、☆の材料と合わせ、1のトマトとレタスは食べやすい大きさにちぎって加え手でザッと和えれば出来上がり♪

レシピのどのパーツにも「シャキシャキ」が含まれていない

図 5 本システムによる検索結果例

図 5. で示している「レタスとトマトのさっぱりナムル風サラダ」のレシピには「シャキシャキ」という単語が含まれていないため、従来の Cookpad では検索されない．しかし、レシピ内には含まれる食材や調理法などを見ると、「シャキシャキ」とする食感を与えることがわかった．

5.2 アプローチ 2：レシピの構造図のボトムアップで検索

アプローチ 2 は、レシピ内の各用語とオノマトペの関連度の合計を算出し、その合計関連度を用いて、キーワードのオノマトペに関連するレシピのランキングを行う．そして、上位の合計関連度をもつレシピを検索結果として返す．

例として、表 6. は「シャキシャキ」に対して、「春雨でボリューム UP 豚キムチ炒飯」及び「ご飯にもお酒にも！ごぼうと豚の時雨煮風」のレシピ内から「食材(food)」、「料理名(dish)」や「調理器具(dishware)」などの各用語の TF-IDF 値そして、それぞれのレシピの合計関連度を示す．

表 6 「シャキシャキ」に対する 2 つのレシピの関連度

レシピ	用語	用語のタイプ	TF-IDF 値
春雨でボリューム UP 豚キムチ炒飯	炒飯	food	0.016952
	キムチ	food	0.052581
	チャーハン	dish	0.512872
	炒める	cook	0.149036
TF-IDF 合計値			2.285533
ご飯にもお酒にも！ごぼうと豚の時雨煮風	豚	food	0.023963
	しょうが	food	0.005373
	昆布	food	0.016120
	煮詰める	cook	0.044477
TF-IDF 合計値			0.906851

表 6. に示している結果より、「シャキシャキ」に対して、「春雨でボリューム UP 豚キムチ炒飯」は「ご飯にもお酒にも！ごぼうと豚の時雨煮風」より合計 TF-IDF 値が高いため、「春雨でボリューム UP 豚キムチ炒飯」のレシピの方が「シャキシャキ」に関連すると判断され、ユーザに検索結果として返す．

6. 関連研究

ユーザの必要とするものを発見できるようにユーザにヒントを与えながら検索を行うシステムについては様々な研究が行われている。キーワード候補をファセット毎に表示しながら探索的に絞り込みを行う Faceted Navigation[1]やキーワード検索結果のクラスタリング[4]などがある。我々はそのような誘導的な絞り込みではなく、曖昧な印象に対する表現能力の高いオノマトペをキーワードとして用いることに着目している。また印象検索に関する研究も数多くある[5][6]。これらは画像や音楽、映像データに対するものが主であり、基本的には検索対象と印象語となる形容詞を結びつけている。我々のようにオノマトペをキーワードとして用いるものはあまりない。オノマトペの持つ曖昧な印象表現能力を活用した研究も最近では注目されており、オノマトペに対応する線描画[7]や物体の動作[8]を表すものがあるが、これらはオノマトペの印象を具現化したものであり、具体的な物を特定する技術ではない。

また、料理レシピ検索に関する研究も多数行われている。その中にはユーザのスケジュールを考慮したダイエットののためのレシピ推薦[9]や個人の状況を考慮した”かんたん”なレシピの推薦[10]、キーワードとその優先度による対話的な料理レシピ検索[11]などがある。

7. おわりに

本研究は、Cookpad ウェブサイトの料理レシピを対象として、オノマトペをキーワードにした料理レシピ検索システム「オノマトペロリ」を開発した。また、「オノマトペロリ」の概要及び提案した検索アプローチによる実験結果・結果検討について述べた。

今後、より高精度の検索を実現するため、実験のデータとするレシピの数を増加する必要があると考えられる。また、本稿で提案した検索アプローチ2について、今後実験を行い、アプローチ1と結果の精度を検討する必要がある。

レシピのほかにも分野にも適用することが考えられる。例えば、CDやDVDの検索で、歌手や曲名などのほかに曲のイメージをオノマトペで検索すると検索の幅が広がるだろう。また、映画やテレビ番組にも応用できる。その際のオノマトペと検索対象の関連を求めるソースとして映画のレビューサイト、掲示板やニコニコ動画などが利用できると思われる。

また、他のジャンルについても、「服飾」系の物販サイトにおける印象検索や「茶碗・茶器」のあいまいな表現による検索にも適用できると考えられる。

参考文献

- 1) Flameco Search: UC Berkeley School of Information The Flamenco Search Interface Project “<http://flamenco.berkeley.edu/index.html>”
- 2) 日本最大の料理サイト「COOKPAD」<http://cookpad.com>
- 3) 中部文子, 浅賀千里, 渡辺知恵美: 完成情報を利用したオノマトペ学習支援システムの開発, DIEM Forum 2008.
- 4) Kaki, M. FindexSearch result categories help users when document ranking fail. In Proceedings of ACM SIGCHI(Portland, OR, April. 2005)
- 5) 石橋直樹, 清木 康, 中神康裕, 佐藤 聡: 複数の音符列から構成される音楽データを対象とした印象メタデータ抽出方式, 日本データベース学会 Letters Vol.2, October 2004, No.2, pp.65-68.
- 6) 町田和嘉子, 伊藤貴之: ユーザのスケジュールを考慮したダイエットののためのレシピ推薦, DIEM Forum 2008.
- 7) 神原啓介, 塚田浩二: オノマトペン, ソフトウェア科学会 WISS 2008 登壇発表, pp.79-84, (Nov. 2008)
- 8) 小松孝徳, 日向宏治: ユーザの直感的表現を支援するオノマトペフィルタシステム, The 22rd Annual Conference of the Japanese Society for Artificial Intelligence, 2008.
- 9) 三野陽子, 小林一郎: ユーザのスケジュールを考慮したダイエットののためのレシピ推薦, DIEM Forum 2008.
- 10) 矢島亜紗美, 小林一郎: 個人の状況を考慮した”かんたん”なレシピの推薦, DIEM Forum 2008.
- 11) 塩澤秀和: キーワードとその優先度による対話的な料理レシピ検索, 日本ソフトウェア科学会 WISS 2007 (デモ発表), pp.169-170, 2007 年 12 月.