

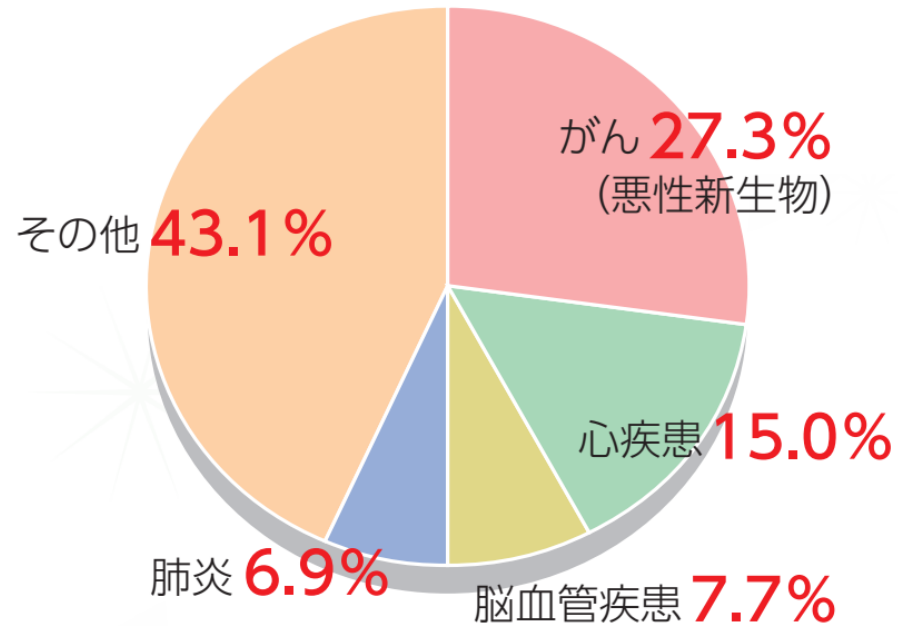
レシピサイトを使用した栄養摂取 管理システムの構築

静岡理工科大学 情報学部 コンピュータシステム学科
宮本知弥, 工藤司

研究の背景

近年、**生活習慣病**が増加傾向にあり、日本人の死因の約6割を占めている。

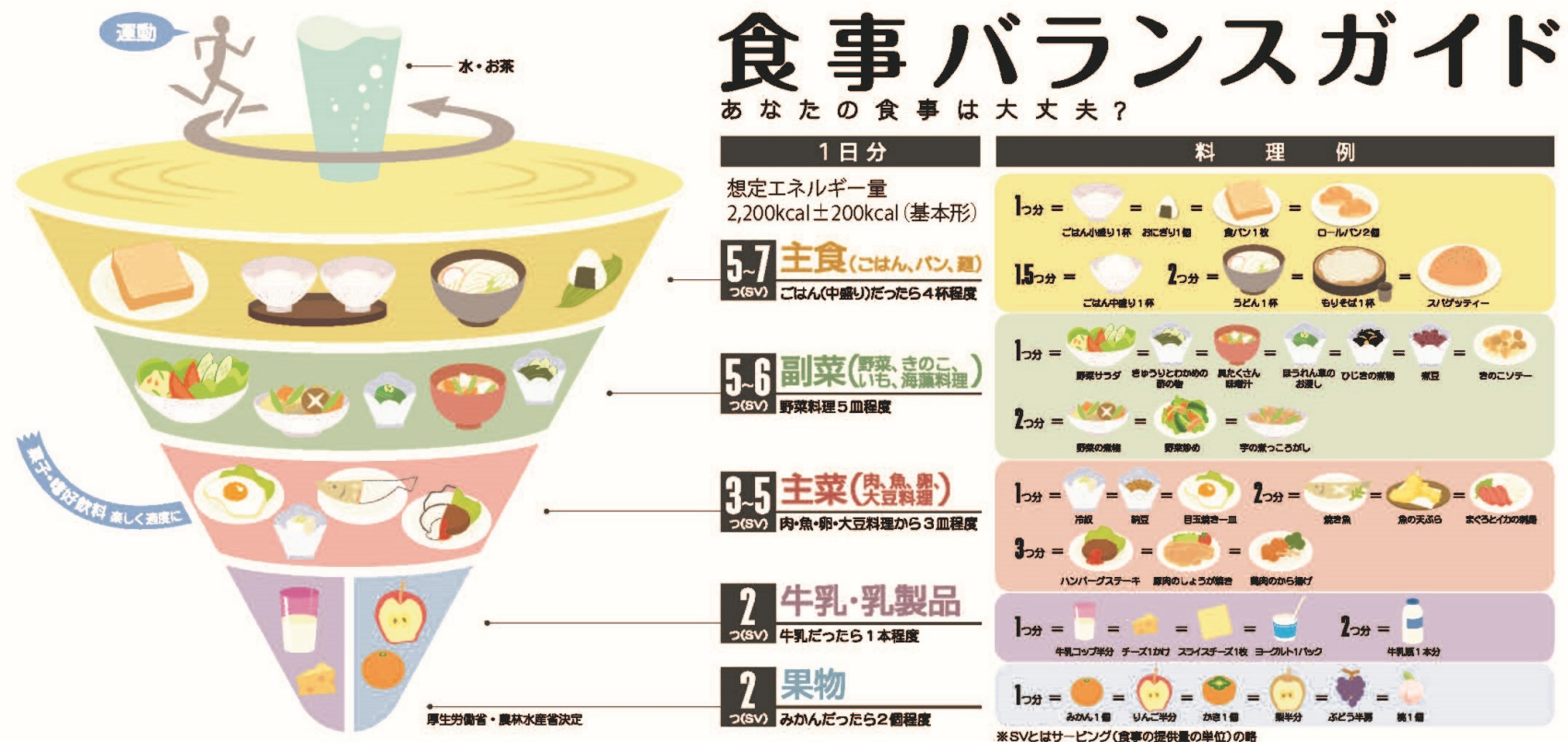
死亡原因の割合（令和元年）



研究の背景

特に**食習慣の重要性**が指摘されている。

日本人の食事摂取基準を厚生労働省が発表している。



研究の背景

栄養成分を把握するため

日本食品標準成分表を文部科学省が発表している。



The screenshot shows the homepage of the Japanese Food Composition Database. At the top, there is a navigation bar with the title "食品成分データベース Food Composition Database" and the Ministry of Education, Culture, Sports, Science and Technology logo. Below the navigation bar, there is a search form with a text input field, a "検索" (Search) button, and a "クリア" (Clear) button. The search form contains instructions in Japanese: "検索する食品を<全角ひらがな、漢字、またはカタカナ>で入力してください。" and "食品をいくつかのキーワードで絞り込みたい場合は、キーワード間に<スペースを1つ以上>入れてください。" An example is provided: "例：「こめ」というキーワードを持つ食品の中から、「めし」というキーワードで絞り込みを行う場合". The search input field contains the text "こめ めし" with a cursor pointing to the space between the characters. Below the input field, there is a note "スペースで区切る" (Separate with space). The background of the website features a photograph of a woman and a young girl sitting at a table, smiling.

食品成分データベース
Food Composition Database

文部科学省
文字サイズ 標準 大 特大

Home Home 食品成分DBとは? about 検索 search ヘルプ help お問い合わせ Contact

検索する食品を<全角ひらがな、漢字、またはカタカナ>で入力してください。

フリーワード検索 検索 クリア

食品をいくつかのキーワードで絞り込みたい場合は、キーワード間に<スペースを1つ以上>入れてください。
例：「こめ」というキーワードを持つ食品の中から、「めし」というキーワードで絞り込みを行う場合

こめ めし 検索 クリア

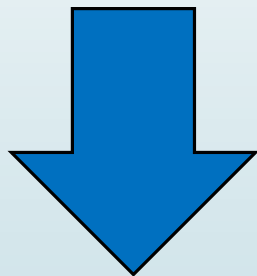
スペースで区切る

食品成分データベース
(<https://fooddb.mext.go.jp/>)

研究の背景

令和元年度国民健康・栄養調査によると

- 忙しくて時間がないこと
- 面倒くさいこと

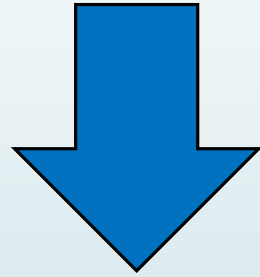


健康な食習慣の妨げ
になっている

		総数	
		人数	%
総数	総数	4,465	-
	仕事(家事・育児等)が忙しくて時間がないこと	1,226	27.5
	外食が多いこと	234	5.2
	自分を含め, 家で用意する者がいないこと	233	5.2
	経済的に余裕がないこと	371	8.3
	面倒くさいこと	1,128	25.3
	その他	483	10.8
	特にない	1,575	35.3
	わからない	205	4.6

研究の背景

現在ではレシピサイトが公開され食習慣が多様化している。



さらに栄養成分の把握を困難にしている。

研究の背景

7

■ 本研究の目的

レシピサイトを活用して、
バラエティに富んだ食事を楽しみながら、
手軽に栄養摂取量を把握したり、
不足栄養素を補うレシピの提案をしたりできる、
栄養摂取管理システムの構築である。

システムの提案

(1)は**レシピの一覧**であり、
レシピサイトからスクレイ
ピングにより作成する。

(2)は**個人ごとの摂取範囲**で
あり、エネルギーの事例を
示している。

個人の状況によって異なる。

レシピNo.	レシピタイトル	提供人数
1	ブルーベリーヨーグルト	1
2	トースト	1
3	豚丼	1

(1) レシピ一覧(抜粋)

ユーザー	栄養成分	上限	下限
1	エネルギー	2915	2385
2	エネルギー	2650	2450
3	エネルギー	2700	2600

(2) 個人ごとの摂取範囲(エネルギーの事例)

システムの提案

(3)は**食事記録**であり、食事したメニューと食事量を記録する。

(4)は**栄養摂取状況**であり、各栄養成分の過不足情報を提示することができる。

日付	時間	メニュー	食事量
2022/10/12	朝	サラダ	1
2022/10/12	朝	トースト	1
2022/10/12	朝	ヨーグルト	1
2022/10/12	昼	サラダ	1
2022/10/12	昼	回鍋肉丼	1
2022/10/12	夕	白米	1
2022/10/12	夕	しょうが焼き	1
2022/10/12	夕	みそ汁	1

(3) 食事記録

日付	栄養成分	不足量	単位	状況
2022/11/1	エネルギー	0	Kcal	OK
2022/11/1	たんぱく質	10	g	不足しています。
2022/11/1	脂質	0	g	OK
2022/11/1	炭水化物	0	g	OK
2022/11/1	食塩相当量	0	g	OK

(4) 栄養摂取状況

システムの提案

栄養摂取状況によって不足した栄養成分があった場合、**不足した栄養成分を補うためのレシピ**を提案する。

レシピタイトル	不足栄養成分	不足量	補足量	kcal	食事量
簡単♪アスパラホイル蒸し?	たんぱく質	10	10.0	80.6	12
ほうれん草ともやしのあえもの	たんぱく質	10	10.0	86.0	1.7
レンジで簡単!もやしのウスターソース和え	たんぱく質	10	9.9	87.3	3

構成と実装

- ▶ CPU : Intel(R) Core(TM) i9-9900K
- ▶ RAM : 64GB
- ▶ OS : Windows 10

- ▶ 使用言語 : Python Version 3.10.4

- ▶ 使用ライブラリ : BeautifulSoup4 Version 4.10.0
mysql-connector-python Version 8.0.28

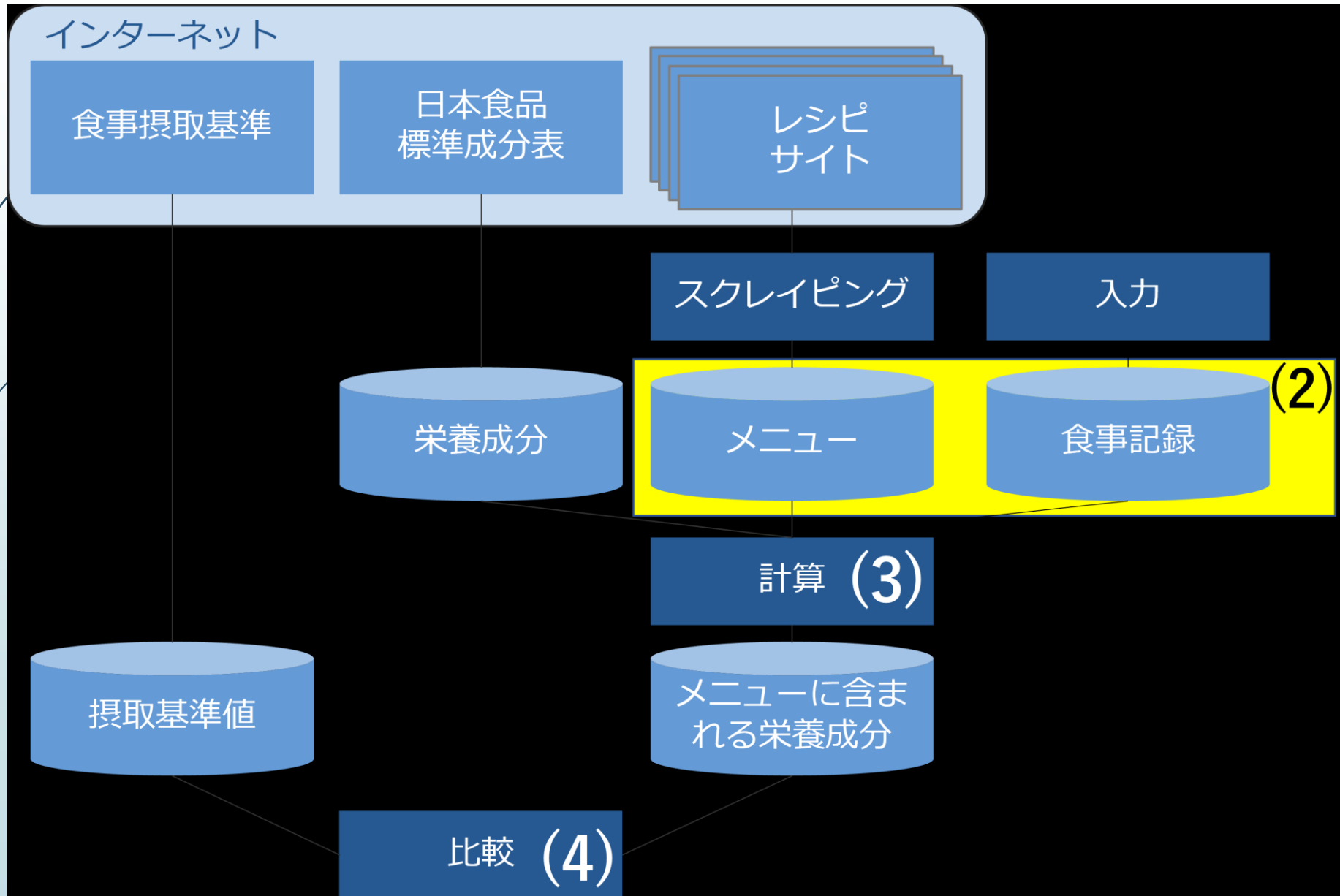
- ▶ 使用データベース : MySQL Version 8.0.18

- ▶ 対象WEBサイト : Cookpad 野菜のおかずカテゴリ 1~999ページ
(<https://cookpad.com/category/10?page=1>)

構成と実装

1. スクレイピングによる情報抽出
2. データベース化
3. 栄養成分計算
4. 摂取基準と比較

構成と実装



構成と実装

1. スクレイピングによる情報抽出

- ➡ それぞれのレシピ詳細ページのURLを取得する.
- ➡ 取得したURLでレシピページにアクセスし、**タイトル**、**材料**、**分量**、**単位**や**提供人数**を取得する.

構成と実装

2. データベース化

重量の単位と材料名の変換

- 材料や分量の表記ゆれが多く存在するため変換を行う。
- レシピの分量にある単位がg(グラム)以外のものは単位変換を行う。

玉ねぎの表記ゆれの例

玉ねぎ, 玉ネギ

たまねぎ, タマネギ

玉葱

新玉ねぎ, 新玉ネギ

新たまねぎ, 新タマネギ

新玉葱

構成と実装

- 変換テーブルの作成
- 調味料や食材のグラム変換が記載されたウェブサイトのデータを参考にした。

— 調味料の目安重量 —

一般調味料

	小さじ (5ml)	大さじ (15ml)
水	5g	15g
酢	5g	15g
酒	5g	15g
醤油	6g	18g
みりん	6g	18g
味噌	6g	18g
あら塩 (並塩)	5g	15g

食品名	単位	重量
アスパラ	1束	100g
	1本	20g
かぶ	1個	80g
かぼちゃ	1個	1200g
カリフラワー	1個	500g
キャベツ	1個	1200g
	1枚	50g
きゅうり	1本	100g

構成と実装

3. 成分計算

- ➡ 廃棄率を考慮した分量の再計算を行い、**レシピごとに含まれる各栄養成分**を求める。
- ➡ レシピごとに含まれる各栄養成分
$$=(1-\text{廃棄率})\times\text{分量}\times\text{成分値}\div 100$$

構成と実装

塩むすびの例

材料	分量	単位
ご飯	100	g
塩	1	g
海苔	1/3	枚

(1) レシピ

栄養成分	含有量	単位
エネルギー	158.97	kcal
たんぱく質	2.91	g
脂質	0.34	g
炭水化物	37.54	g
食塩相当量	1	g

(4) レシピの栄養成分含有量

材料	単位	g
海苔	枚	3

(2) 単位変換テーブル

材料	栄養成分	含有量	単位
ご飯	エネルギー	156.00	kcal
ご飯	たんぱく質	2.50	g
ご飯	脂質	0.30	g
ご飯	炭水化物	37.10	g
塩	食塩相当量	1	g
海苔	エネルギー	2.97	kcal
海苔	たんぱく質	0.41	g
海苔	脂質	0.04	g
海苔	炭水化物	0.44	g

(3) 各材料における栄養成分含有量

構成と実装

4. 摂取基準と比較

- 食事記録から食事メニューや食事量を読み取り, **1日の各栄養成分摂取量**を求める.

- 1日の各栄養成分摂取量

$$= \sum_{r \in R} (1 - \text{廃棄率}_r) \times \text{分量}_r \times \text{成分値}_r \div 100 \times \text{食事量}_r \div \text{提供人数}_r$$

- ただし, Rは, 1日で摂取した食事のレシピの集合である.

構成と実装

- ▶ 1日の各栄養成分摂取量と、個人ごとの摂取範囲を比較することで、栄養のバランスを把握する。
- ▶ 過剰摂取や不足の場合には、通知を出すと共に、不足分を補うメニューの提案を行う。

日付	栄養成分	不足量	単位	状況
2022/11/1	エネルギー	0	Kcal	OK
2022/11/1	たんぱく質	10	g	不足しています。
2022/11/1	脂質	0	g	OK
2022/11/1	炭水化物	0	g	OK
2022/11/1	食塩相当量	0	g	OK

レシピタイトル	不足栄養成分	不足量	補足量	kcal	食事量
簡単♪アスパラホイル蒸し?	たんぱく質	10	10.0	80.6	12
ほうれん草ともやしのあえもの	たんぱく質	10	10.0	86.0	1.7
レンジで簡単!もやしのウスターソース和え	たんぱく質	10	9.9	87.3	3

実験と評価

- レシピ収集効率の評価
 1. レシピ収集時間比較
 2. 表記ゆれ調査

- 摂取栄養管理効率の評価
 1. 栄養摂取状況の計算時間比較
 2. レシピ提案の時間比較

レシピ収集時間比較

- Cookpad野菜のおかずカテゴリのデータを使用した.
- スクレイピングを行い, 抽出したデータをcsvファイルとして出力した.
- この方法を用いて収集した, 9,966件のレシピデータを使用した.
- 使用時間は, 実行時間を総レシピ数で割ったものとして計算した.

レシピ収集時間比較

- 手作業では、下のレシピを使用し、レシピサイトのデータをExcelファイルへ書き写した。

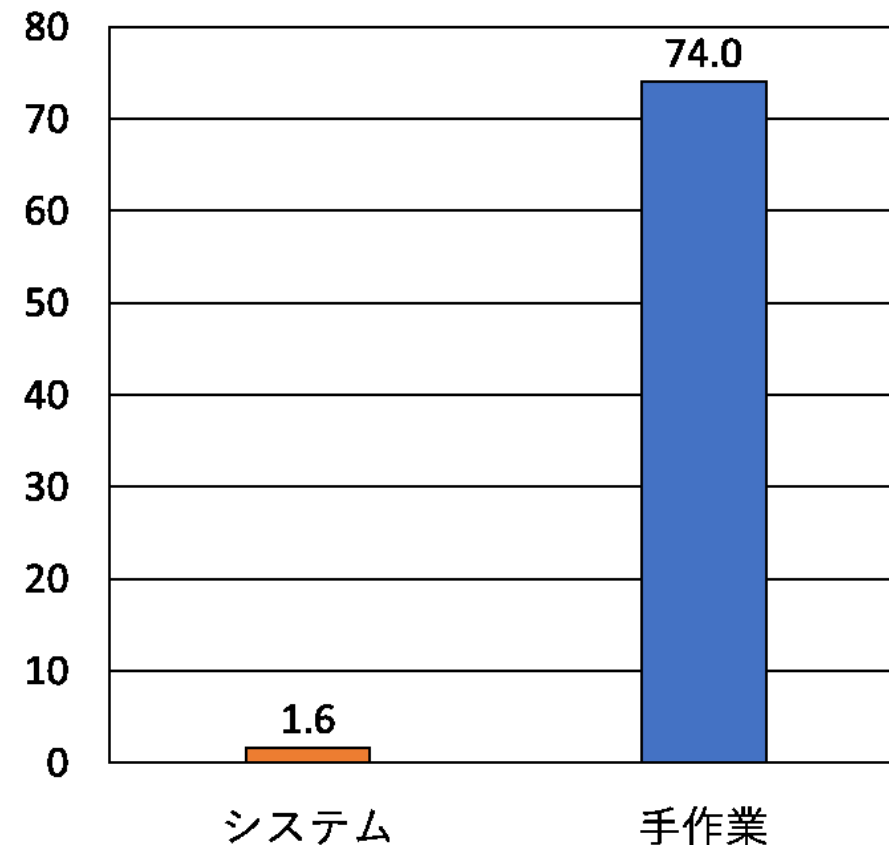
	A	B	C	D
1	タイトル	材料名	分量	単位
2	だしも水もいっさい使わない肉じゃが	牛肉	150 g	
3	だしも水もいっさい使わない肉じゃが	じゃがいも	2.5 個	
4	だしも水もいっさい使わない肉じゃが	たまねぎ	1 個	
5	だしも水もいっさい使わない肉じゃが	にんじん	1 本	
6	だしも水もいっさい使わない肉じゃが	醤油	2.5 大さじ	
7	だしも水もいっさい使わない肉じゃが	砂糖	1.5 大さじ	
8	だしも水もいっさい使わない肉じゃが	酒	1 大さじ	

「だしも水もいっさい使わない肉じゃが」 (<https://cookpad.com/recipe/5629>)

レシピ収集時間比較

- スクレイピングにより，自動収集した場合には，レシピ1件当たりの時間は**1.60秒**であった。
- 自動実行であるため，プログラムの**起動後は人間の作業は発生しない**。
- 手作業では**74秒**であって，**約46倍の効率**で収集可能であった。

処理時間(s)



表記ゆれ調査

- 材料や分量の表記ゆれが多く存在するため、何件のデータが正しく登録されるかを検証した。
- レシピごとに、**全ての材料数と、材料名と単位共に正しく記入されている材料数を比較**し、一致するものを正しいレシピとした。

玉ねぎの表記ゆれの例

玉ねぎ, 玉ネギ

たまねぎ, タマネギ

玉葱

新玉ねぎ, 新玉ネギ

新たまねぎ, 新タマネギ

新玉葱

表記ゆれ調査

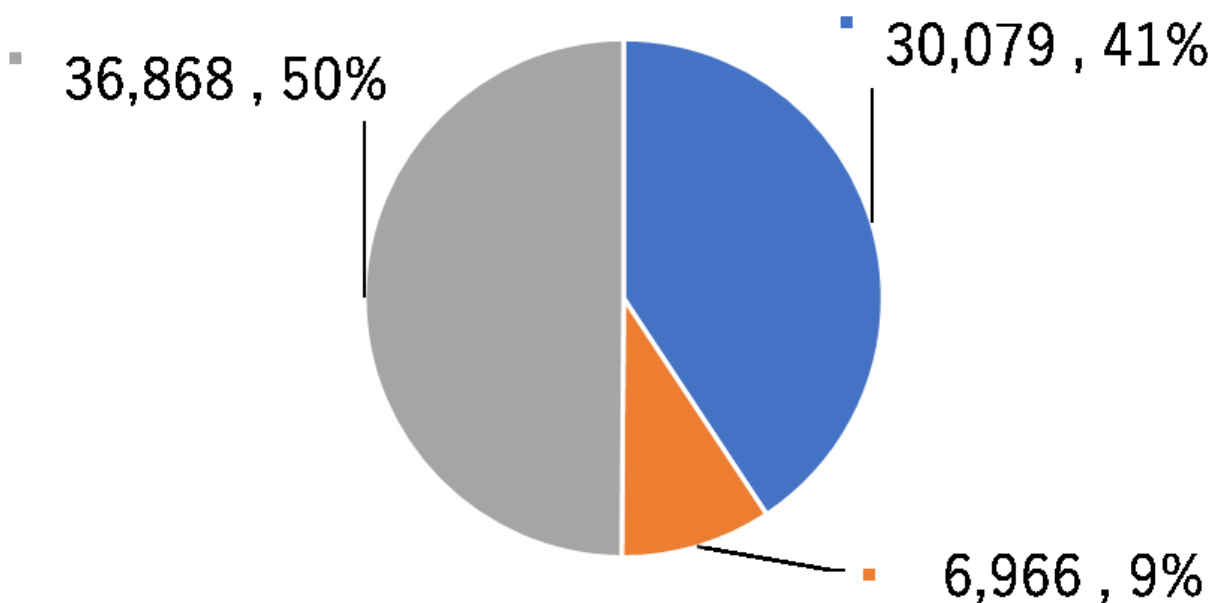
- 表記ゆれに対応するため、データベースには変換テーブルを付加している。
- ただし、今回の評価は、以下の設定で行った。
- 砂糖(上白糖), 塩(食塩), お酢(酢), しょう油(醤油), みそ(味噌) 「()」内の名称が参考ページ記載の名称である。
- これらの調味料は、料理に使用する頻度が高いため、対応を行った。

表記ゆれ調査

9,966件のレシピのうち、92件のみしか完全なデータとして保存できなかった。

なお、材料名、単位共に正しいものが41%で、平均材料数が7.4個なので、 $0.41^{7.4} \times 9,966 \approx 13.6$ (件) 今回の結果は概ね妥当であると考える。

■ 単位正常 ■ 単位不正 ■ 材料名不正 単位:件数, 割合(%)



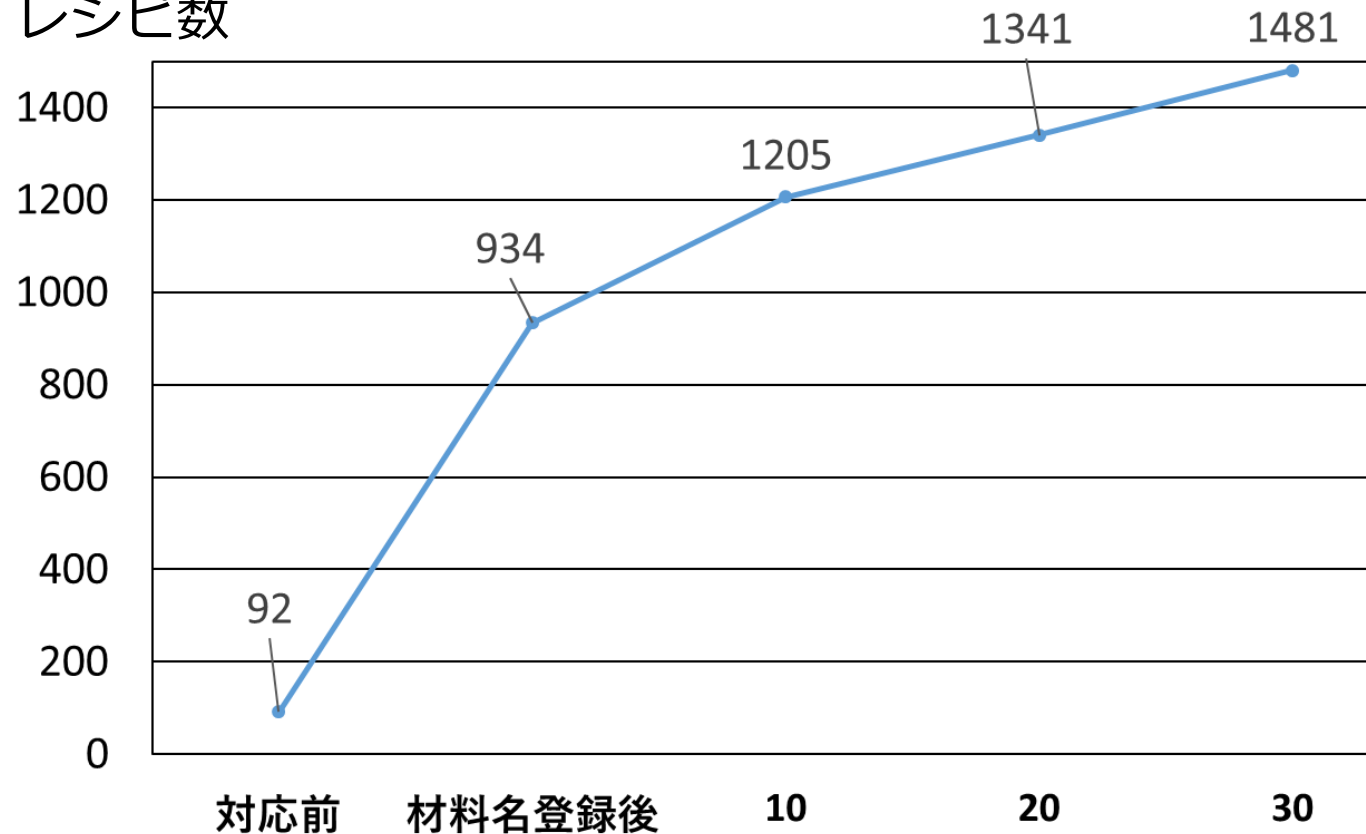
表記ゆれ調査

表記ゆれへの対応

■登録されていない材料名を全て登録し、数が多い方から順に重量を登録した。

■表記ゆれの数、約5800件存在した。

レシピ数



重量登録件数

栄養摂取状況の計算時間比較

- 本システムは管理対象の全ての**栄養成分量を一括で計算**できるため、実際に実行時間を計測して評価した。
- 右のレシピを対象とした。
- 管理対象の栄養素の数は、26件である。

「おさつだけ〜！」

(<https://cookpad.com/recipe/356880>)

栄養摂取状況の計算時間比較

- 手作業では、材料名を検索し、その材料の廃棄率とエネルギーのみを調べ、Excelで計算を行った。
- 作業時間は、ストップウォッチで計測し、これに栄養素の数を乗じた時間として推定した。

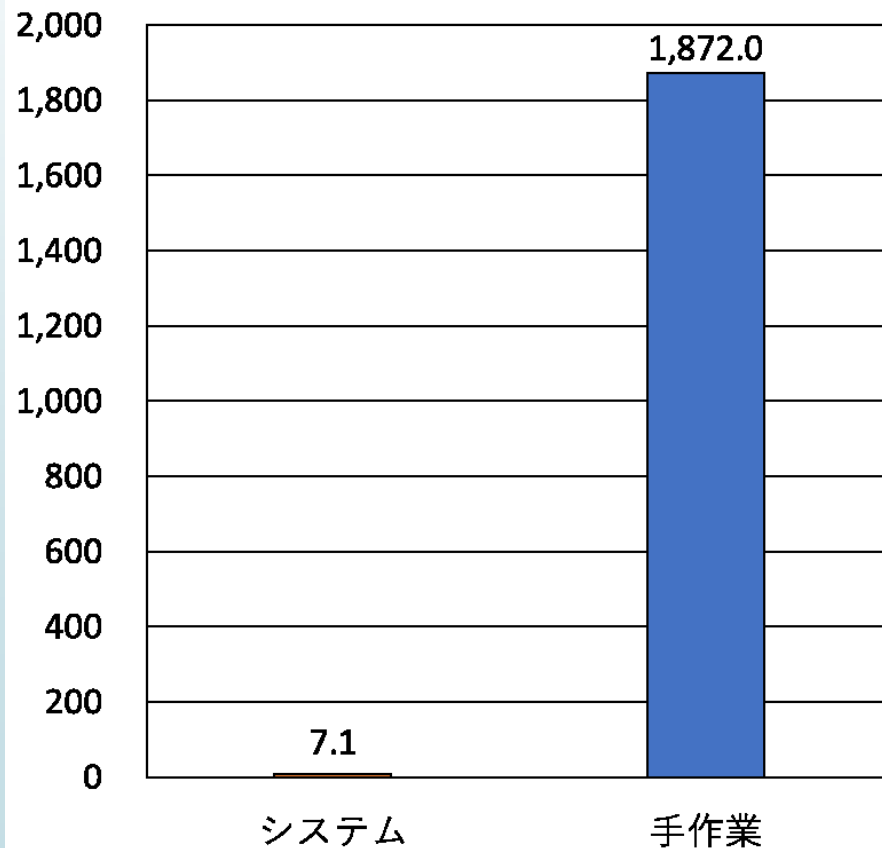
食品番号	索引番号	食品名	廃棄率	エネルギー	
				KJ	kcal
		単位	%	ENERC	ENERC_KCAL
		成分識別子	REFUSE		
02045	217	<いも類> (さつまいも類) さつまいも 塊根 皮つき 生	2	539	127

	A	B	C	D	E
1	材料名	分量	廃棄率	エネルギー	
2	さつまいも	200	2	127	248.92
3	砂糖	18	0	391	70.38
4	しょうゆ	12	0	76	9.12
5	みず	45	0	0	0
6	サラダ油	0	0	886	0
7					328.42

栄養摂取状況の計算時間比較

- 本システムでは、7.14秒であった。
- 手作業では1件あたり72秒であり、26件の栄養素では、1,872秒になると評価された。
- 本システムでは、262倍の効率で栄養取得状況を把握できた。
- しかし、1レシピ当たり 6.87 秒の食事実績を入力する時間が操作の負荷であった。

処理時間(S)



レシピ提案の時間比較

- ▶ たんぱく質が10g不足しているという前提で、**レシピの提案**を行った。
- ▶ エネルギーの少ない順に並べ、不足量を摂取できる食事量を表示した。

レシピタイトル	不足栄養成分	不足量	補足量	kcal	食事量
簡単♪アスパラホイル蒸し?	たんぱく質	10	10.0	80.6	12
ほうれん草ともやしのあえもの	たんぱく質	10	10.0	86.0	1.7
レンジで簡単!もやしのウスターソース和え	たんぱく質	10	9.9	87.3	3

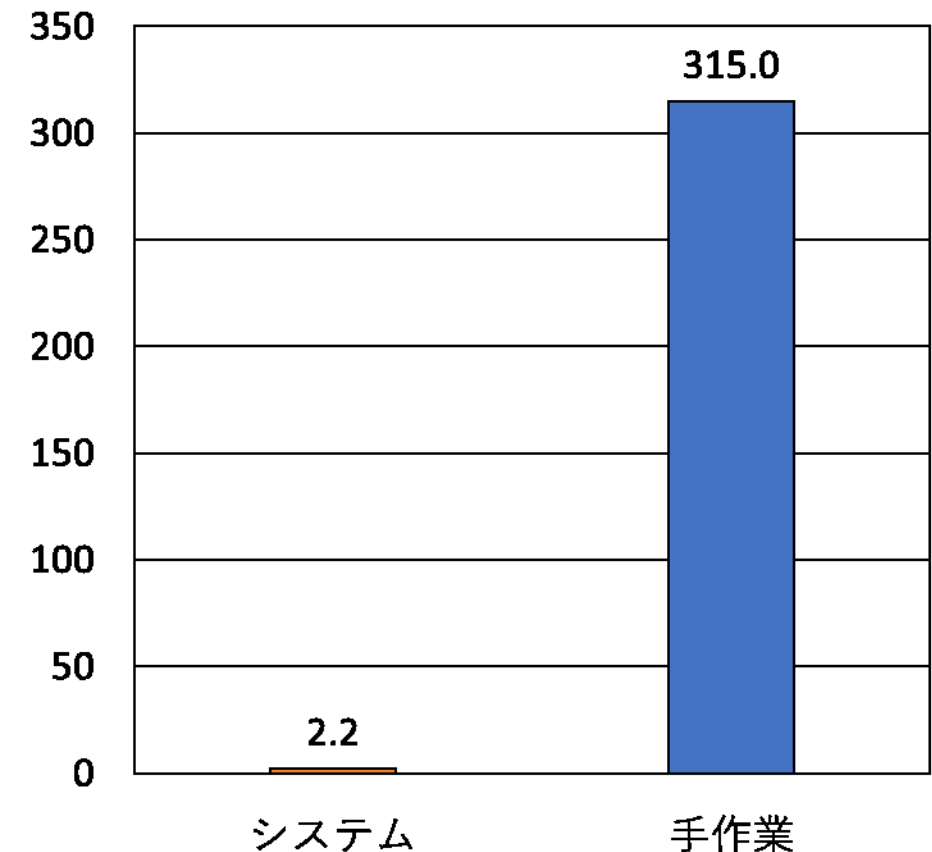
レシピ提案の時間比較

- 手作業では、たんぱく質が多く含まれると想定するレシピを検索し、栄養摂取状況計算と同様の手法で計算を行った。
- 本評価では、廃棄率、たんぱく質とエネルギーのみを調べ、計算を行った。
- 作業時間は、ストップウォッチで計測し、これに提案件数の3を掛けた時間を、実行時間とした。

レシピ提案の時間比較

- 本システムでは、**2.2秒**であった。
- 手作業では1件あたり105秒であり、3件の提案では、**315秒**になると評価された。
- 本システムでは計算を指示するだけで、自動的に**143倍の効率**でレシピ提案を行えた。

処理時間(S)



考察

本システムの有効性

- レシピサイトから一括してレシピを**自動収集**できることを示した。
- 1レシピあたり6.87秒の食事記録の入力のみで、**自動的に栄養摂取状況の提示が可能**である。
- レシピの提案機能により、様々な条件で**追加メニューを提案**できる。

考察

本システムの課題

- レシピサイトの**表記揺れ**への対応.
→ 変換テーブルの**データを増やす**ことで対応可能.
- 食事記録の**入力時間の削減**.
→ レシピを**自動提案**する機能の構築が考えられる.

まとめ

- 栄養摂取量を計算するには**手間がかかり**、食習慣の改善が進まないという課題がある。
- スクレイピングによって**効率よくレシピ収集**が行えた。
- 食事記録の入力だけで、**栄養摂取状況**や**不足分を補うメニュー**を確認できた。
- 食習慣の改善が**容易になる**と期待される。

まとめ

将来研究

- レシピサイトの表記ゆれへの対応
- レシピの自動提案