室蘭工業大学
学術資源アーカイブ
Muroran Institute of Technology Academic Resources Archive

カーボンフットプリントのための計算システムに関する情報科学的研究

メタデータ	言語: Japanese
	出版者:
	公開日: 2017-05-19
	キーワード (Ja):
	キーワード (En):
	作成者: 趙, 培江
	メールアドレス:
	所属:
URL	https://doi.org/10.15118/00009197

2016年度 博士学位論文

カーボンフットプリントのための計算システムに関する情報科学的研究

2016年12月22日

室蘭工業大学大学院 博士後期課程
工学専攻

趙培江

目次

第 1 章　緒論 1

1.1　カーボンフットプリント 1

1.2　カーボンフットプリントの計算に関する研究 7

1.3　カーボンフットプリントの社会認知と問題点 14

1.4　研究目的 15

第 2 章　カーボンフットプリント計算システム 19

2.1　計算システムの設計方針 19

2.2　計算システムの機能 22

2.3　計算システムの性能評価 35

第 3 章　テキストマイニングに基づいたカーボンフットプリントの計算方法 40

3.1　計算方法の方針 40

3.2　原材料情報抽出ための機械学習 47

3.3　CFP 計算用のデータを収集する方法 64

3.4 計算方法 73

3.5 性能評価 78

3.6 まとめ 81

第 4 章 カーボンフットプリント認知度向上の取組 82

4.1 取組 82

4.2 130 品製品の計算結果 85

4.3 カーボンフットプリン認知度の調査結果 91

4.4 まとめ 99

第 5 章 結論 100

謝辞 102

参考文献 104

参考文献 104

第 1 章

緒論

1.1 カーボンフットプリント

温室効果ガスの増加に由来する地球温暖化問題が広く認知されている。内閣府政府広報室が 2007 年に行った世論調査によれば、地球の温暖化、オゾン層の破壊、熱帯林の減少などの地球環境問題に「関心がある」とする者の割合が 92.3% になっている [1]。また、2008 年の内閣府政府広報室の世論調査では、「低炭素社会」を実現すべきかという質問に、90.1% が「そう思う」と答えている [2]。このような社会情勢の中で、温室効果ガスの排出を抑えた低炭素社会を実現するために、世界においても、日本国内においても、様々な政策が展開されている [3]。例えば、1998 年に制定された地球温暖化対策の推進に関する法律 (地球温暖化対策推進法) では、すべての自治体に温室効果ガス削減実行計画の策定が求められている。そして、地球温暖化対策推進法により、自治体は温室効果ガスの排出量の削減に取組、その総排出量を公表している。

カーボンフットプリント (Carbon Footprint of Products、 CFP) も、温室効果ガス削減を図る運動の一つである [4]。CFP は、製品やサービスの原材料調達、輸送、製造、消費・使用、廃棄・リサイクルの全過程で発生する温室効果ガスの量を二酸化炭素（CO_2）の量に換算して、ラ

ベル付けする仕組みである [4]。つまり、CFP は CO_2 排出量の「見える化」である。カーボンフットプリントコミュニケーションプログラムの「基本文書」[5] では、京都議定書で指定した 6 種のガス（二酸化炭素、メタン、亜酸化窒素、ハイドロフルオロカーボン類、パーフルオロカーボン類、六フッ化硫黄）を温室効果ガスに指定している。この 6 種の温室効果ガスの排出量に、気候変動に関する政府間パネル（Intergovernmental Panel on Climate Change、IPCC）の第二次報告書にある 100 年値をかけて、CFP としている。

CO_2 排出量を「見える化」する CFP によって、生産者は生産過程における CO_2 排出量の構成を知り、企業間で協力して CO_2 排出量の削減に努力することが期待される。また、消費者は自身の CO_2 排出量を把握し、低炭素な消費生活に向けて行動することが期待されている。CFP は、産業界だけでなく、消費者にも温室効果ガスの排出量を知らせる特徴がある。したがって、CFP は、温室効果ガスを削減する運動に、消費者を広く巻き込める可能性がある。

CFP が温室効果ガス削減に実質的に貢献するためには、消費者における CFP の認知度の向上が不可欠である。CFP の取組の消費者へのアピールは、日本では、2016 年現在、CFP 認定製品をまとめている CFP コミュニケーションプログラム [6] のウェブサイトや、CFP 事業者が個別に行う取組 [7] がある。例えば、一般社団法人産業環境管理協会が運営する CFP コミュニケーションプログラムのウェブサイトでは、CFP の認定製品の一覧や CFP 算定の「カーボンフットプリント製品種別基準 (Carbon Footprint of Products Product Category Rule、CFP-PCR)」等様々な情報が公開されている。そして、CFP コミュニケーションプログラムは、認定した製品に図 1.1 で示す CFP マークを付けている。

海外では、2006 年 12 月にイギリスの Carbon Trust 社が CFP の実施を宣言したのが CFP の始まりである。2008 年にはイギリスで CFP 算定方法の共通規格 (PAS2050) が発表された [8]。フランス、ドイツ、韓国、タイ等でも様々な取組を展開している [4]。図 1.2 はイギリスの CFP ラベルを示す。図 1.3 は韓国のの CFP ラベルを示す。

室蘭工業大学とコープさっぽろは、2008 年よりスーパーマーケットにおける CO_2 排出量削減をテーマにして活動してきた。スーパーマーケットにおいて CO_2 排出量を削減するために、室蘭工業大学とコープさっぽろは、店舗や食品工場における CO_2 排出量の調査と削減、CFP による地球温暖化問題の消費者へのアピールを実施してきた。そして、2010 年 3 月には、16 品の製品について CFP の計算結果を Point of purchase advertising(POP) 表示し、店舗内に展示した [9]。室蘭工業大学とコープさっぽろは、2010 年以降 CFP を計算する製品の数を増やしてきた。室蘭工業大学とコープさっぽろは、2014 年 10 月には、130 品について CFP を計算し、それを図 1.4 に示すパンフレットにまとめて店頭や宅配で配布した [10]。さらに、コープさっぽろのプライベートブランドである「なるほど商品」の包装に CFP をラベルを印刷し、店頭で販売している。店頭で陳列された CFP のラベル表示された製品を図 1.5 に示す。この CFP のラベル付けした製品をコープさっぽろは現在 (2016 年 11 月) も増加させている。

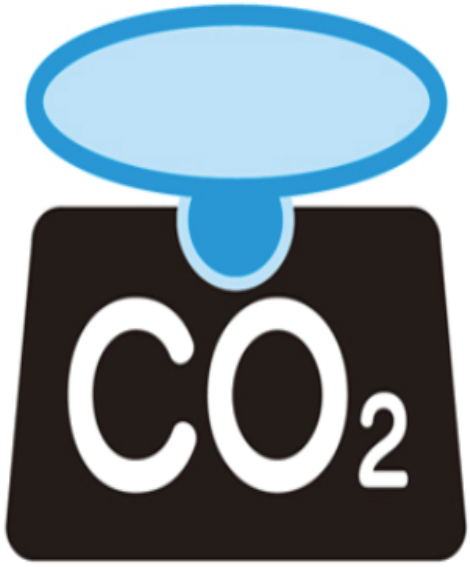


図 1.1　CFP コミュニケーションプログラムの CFP マーク

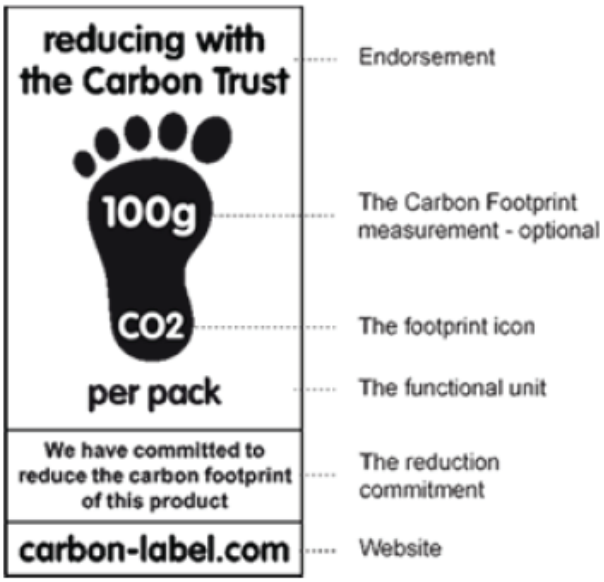


図 1.2　イギリスの CFP ラベル

図 1.3　韓国の CFP ラベル

図 1.4　配布したパンフレット

図 1.5　コープさっぽろで店頭販売されている CFP 表示製品

1.2　カーボンフットプリントの計算に関する研究

CFP の標準の計算方法は製品のライフサイクルアセスメント（Life Cycle Assessment, LCA）に基づいて ISO 14067 で定義されている [4]。

CFP に関する研究では、製品の CFP の算定事例と CFP の計算方法に関する研究が主に報告されている。製品の CFP の算定事例の研究では、一つまたは一種類の製品について、この製品の CFP を詳しく算定している。CFP を算定する製品には、食品 [11]、日用品 [12] 等販売している商品から、建物 [13]、活動 [14]、資源 [15] 等がある。CFP の計算方法に関する研究では、特定の製品、活動の CFP を計算する方法と CFP 計算の各段階の具体的な計算方法等の研究である。例えば、稲葉他は、家庭での調理 [14]、外食メニュー [16] の CFP の推算方法を提案した。また、永野他は輸送 CFP の輸送距離の計算方法を提案した [17]。

CFP を計算するときの作業は主に次の 3 段階からなる。

1. 計算する CFP の段階に基づいて、計算に関するデータを決める。
2. 各段階の CFP 計算に必要なデータを収集する。
3. 収集したデータを用いて CFP を計算する。

CFP を計算する段階は基本的には、原材料調達、輸送、製造、消費・使用、廃棄・リサイクルまでである。計算に関する各段階のデータを表 1.1 にまとめる。ここで、製造と消費・使用段階の各エネルギーの単位は使用するエネルギーによって変わる。例えば、電気は kWh、ガスはリットル等である。算定する製品と計算方法によって、使用するデータの範囲が変わることがある。

表 1.1　CFP 計算に関する各段階のデータ

段階	データの内容	原単位
原材料調達	製品の原材料、原材料の重量	原材料 1kg 当たりの CO_2 排出原単位
輸送	製品・原材料の調達先、輸送手段、輸送距離	製品輸送 1ton×1km 当たりの CO_2 排出原単位
製造	製造時電気、ガス等のエネルギー使用量	各エネルギー 1 単位当たりの CO_2 排出原単位
消費・使用	使用時電気、ガス等のエネルギーの使用量等	各エネルギー 1 単位当たりの CO_2 排出原単位
廃棄・リサイクル	製品を廃棄する部分の重量、廃棄する方法等	1kg 廃棄・リサイクル時の CO_2 排出原単位

例えば、トイレットペーパーの算定事例 [12] では、消費・使用段階の CFP を計算していない。

各段階のデータを取得する方法は、製品の生産メーカーが自ら収集する方法と CFP の関連研究により作成されたオープンデータから取得する方法がある。CFP の導入により製品を差別化するために、原材料調達データ、製造データ、廃棄データは主に製品の生産メーカーが自ら収集する。

輸送距離の推定では、生産メーカーから取得した原材料、製品の調達先、製造工場の住所の情報を用いて Google Maps 等のオープンデータから輸送距離を取得する方法がある。一方、生産メーカーから調達先等のデータを取得できない場合も多い。このときは多くの PCR では、輸送範囲に応じて、決めた輸送距離を使って CO_2 排出量を計算している。例えば、精製糖 (第 2 版) の CFP-PCR[18] では、輸送距離を、市内もしくは近隣市間に閉じている輸送の場合は 50km、県内で閉じた輸送の場合は 100km、県間輸送の可能性がある場合は 500km 等と固定している。

製造データの内容は、生産メーカーから取得できるデータによって大きく変わる。例えば、室蘭工業大学とコープさっぽろの取組で計算した CFP[10] では、製品単品のデータがないため、生産工場全体のエネルギー使用量と出荷量に基づいて、製品単品の製造 CFP を計算している。生産メーカーから製造に関するデータを取得できない場合、イギリスの赤ワインの CFP 計算事例 [19] のように製品の製造標準に基づいて計算する場合もある。

消費段階のデータは、生産メーカーが提供した製品の使用標準シナリオに基づいて算定される。電気、ガス等のエネルギーを消費しない食品や日用品等の製品では、消費段階のデータがない場合もある。廃棄段階のデータでは、廃棄する方法を生産メーカーが指定していない場合がある。

原材料の CO_2 原単位は、公開されている CO_2 原単位データベースから収集される。日本では、図 1.6 に示す味の素グループ版「食品関連材料 CO_2 排出係数データベース」[20] と CFP コミュニケーションプログラムの「CO_2 換算量共通原単位データベース」[21] 等の CO_2 原単位データベースが使われている。味の素グループの CO_2 原単位データベースは無料で利用できるデータベースである。輸送、エネルギー消費に関する CO_2 原単位データベースには、環境省が公開している「排出原単位データベース（ver.2.1）」[22] や各電力会社が公開している CO_2 原単位がある。

CFP は

$$\mathrm{CFP} = \sum_{i=1}^{N} (\mathrm{CO_2原単位})_i \times (\text{当該量})_i \tag{1.1}$$

によって、計算される。式 (1.1) の i は CFP を計算する総数 N の各段階を表わす。手動で CFP

を計算する時、CFP を計算する人が各段階の CO_2 原単位を手動で収集し、収集したデータを使って手動で CFP の計算式に組立てて、CFP を計算する。CFP の計算を実行する CFP 計算システムには、図 1.7 に示すマンチェスター大学が開発した CCaLC、図 1.8 に示す thinkstep 株式会社が開発した LCA 計算システム GaBi 等がある。

CFP 計算システムでは、各段階のプロセスを計算システム上に作成し、計算用データを計算システムに入力する。そして、CFP 計算システムは組み込れたデータベースを使って CFP を計算する。使われている各 CFP 計算システムはそれぞれが独自な特徴を持っている。例えば、CCalC では、CFP を計算する段階の構成は固定されており、CFP を計算する流れを理解しやすくなっている。一方、GaBi では、計算者は計算する段階の構成を自由に組み立て、多様な製品に対応できるようになっている。

CFP 計算システムは手動計算に比べ以下の利点がある。

- CFP の計算段階がモジュラー化されており、データを管理し易い。
- CO_2 原単位データベースが内蔵されており、CO_2 原単位は計算システム内で検索できる。
- CFP の計算式を計算システムが組み立てる。

しかし、このような CFP 計算システムは、操作する窓が多く、操作が複雑になる等の欠点もある。

AJINOMOTO.

味の素グループ版CO_2排出係数データベース

（独）国立環境研究所公開　1990・1995・2000・2005年環境負荷原単位データブック（3EID）から「重量を主とした取引単位ベースCO_2排出係数」への換算

2010年10月15日公開

基本分類番号	品目数番号	分類名称	品目名称	単位	1990年 3EID 生産者単価 円／単位	1990年 3EID CO_2排出係数 $(I-A)^{-1}$ t-CO_2／単位	1995年 3EID 生産者単価 円／単位	1995年 3EID CO_2排出係数 $(I-A)^{-1}$ t-CO_2／単位	2000年 3EID 生産者単価 円／単位	2000年 3EID CO_2排出係数 $(I-A)^{-1}$ t-CO_2／単位	2005年 3EID 生産者単価 円／単位	2005年 3EID CO_2排出係数 $(I-A)^{-1}$ t-CO_2／単位	備考	平均 単位	平均 平均CO_2排出係数 f t-CO_2／単位	平均 標準偏差 σ t-CO_2／単位	平均 標準偏差の比率（σ/f） %
1		米		百万円		1.328		1.446		1.894		2.107		百万円	1.644	0.344	21%
	1		玄米	t	292,307	0.388	280,198	0.405	244,442	0.414	215,699	0.454		t	0.416	0.028	7%
	2		くず米	t	59,304	0.079	60,654	0.088	77,440	0.131	・・・	・・・	*6)	t	0.099	0.028	28%
	3		稲わら	t	25,884	0.034	27,896	0.040	25,884	0.044	・・・	・・・	*6)	t	0.040	0.005	12%
2		麦類		百万円		2.208		2.429		2.172		2.742		百万円	2.388	0.262	11%
	4		小麦（国産）	t	151,524	0.335	155,427	0.378	164,284	0.357	152,932	0.419		t	0.372	0.036	10%
			小麦（輸入）	百万円		2.208		2.429		2.172		2.742		百万円	2.388	0.262	11%
			大麦（国産）	百万円		2.208		2.429		2.172		2.742		百万円	2.388	0.262	11%
	5		六条大麦（国産）	t	125,029	0.276	116,281	0.282	115,450	0.251	134,851	0.370		t	0.295	0.052	18%
	6		二条大麦（ビール麦）	t	139,023	0.307	144,099	0.350	129,333	0.281	140,821	0.386		t	0.331	0.047	14%
	7		裸麦	t	162,212	0.358	161,827	0.393	169,667	0.369	159,504	0.437		t	0.389	0.035	9%
			大麦（輸入）	百万円		2.208		2.429		2.172		2.742		百万円	2.388	0.262	11%
3		いも類		百万円		1.715		1.849		1.833		1.968		百万円	1.791	0.140	8%
	8		かんしょ	t	76,100	0.131	78,689	0.130	89,255	0.164	85,011	0.167		t	0.148	0.020	14%
	9		ばれいしょ	t	38,400	0.066	47,960	0.079	52,753	0.097	44,327	0.087		t	0.082	0.013	16%
4		豆類		百万円		1.676		1.678		1.130		2.068		百万円	1.638	0.386	24%
	10		大豆（国産）	t	232,800	0.390	239,622	0.402	230,867	0.261	207,924	0.430		t	0.371	0.075	20%
			大豆（輸入）	百万円		1.676		1.678		1.130		2.068		百万円	1.638	0.386	24%
			その他の豆類	百万円		1.676		1.678		1.130		2.068		百万円	1.638	0.386	24%
	11		えんどう（種）	t	223,167	0.374	476,806	0.800	351,702	0.397	・・・	・・・	*6)	t	0.524	0.240	46%
	12		そら豆	t	67,265	0.113	121,993	0.205	・・・	・・・	・・・	・・・	*1)	t	0.159	0.065	41%
	13		いんげん豆	t	225,700	0.378	374,239	0.628	340,258	0.384	235,486	0.487		t	0.469	0.117	25%
	14		小豆	t	245,670	0.412	316,577	0.531	358,407	0.405	249,075	0.515		t	0.466	0.067	14%
	15		ささげ	t	401,460	0.673	520,920	0.874	540,084	0.610	・・・	・・・	*6)	t	0.719	0.138	19%
	16		らっかせい	t	288,000	0.483	285,225	0.479	434,212	0.491	379,813	0.785		t	0.559	0.151	27%
			その他の豆類	百万円		1.676		1.678		1.130		2.068		百万円	1.638	0.386	24%

図 1.6　味の素グループ版「食品関連材料 CO_2 排出係数データベース」

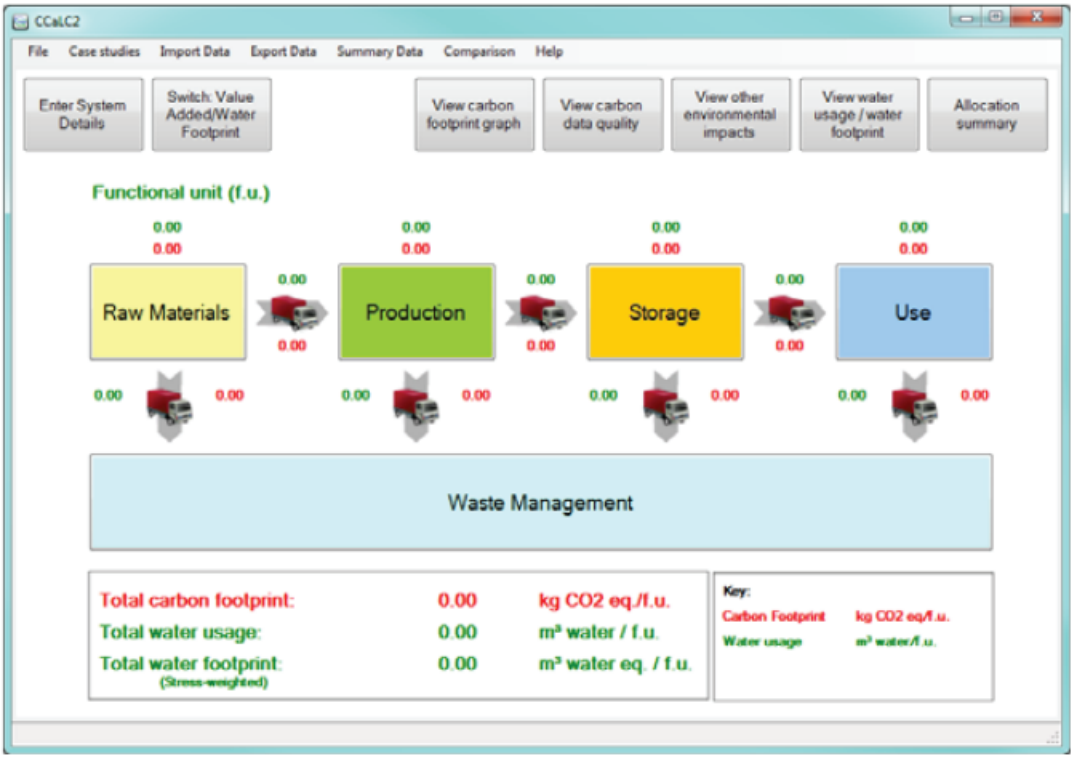


図 1.7　マンチェスター大学を開発した CCaLC

図 1.8　thinkstep 株式会社の GaBi

1.3　カーボンフットプリントの社会認知と問題点

2006 年にイギリスで初めて提唱された CFP を社会の中で実際に運用するにともない、CFP の課題も明らかになってきている。日本においては、2009 年より経済産業省等の主導で実施されてきた「カーボンフットプリント制度試行事業」に引き続き、2012 年 4 月から、産業環境管理協会が、CFP コミュニケーションプログラムとして CFP を日本全国規模で実施している。産業環境管理協会が 2013 年に公開した中期行動計画 [23] によれば、CFP プログラムの運用に関する課題の第 1 として、CFP の認知度の低さがあげられている。さらに、同中期行動計画では、CFP の課題として、低い認知度の他に、CFP マーク付製品が市場で見られない、費用対効果や社内負荷の低減を具体的に示せない等があげられている。

2011 年にみずほ情報総研株式会社は「大型商業施設におけるカーボンフットプリントに関する消費者受容性調査」を実施し、CFP の認知度を報告している [24]。この調査では 2011 年 8 月 19 日から 21 日の三日間、神奈川県相模原市の大型商業施設「アリオ橋本」で 732 名の一般消費者に質問を行った。消費者の認知度調査の結果は図 1.9 に示すように、CFP の意味を知っている消費者は 8% であり、92% の消費者は CFP を正しく認識していない、または、CFP を認知していなかった。同調査では、「カーボンフットプリントを表示したほうがよいと思う商品の種類」についても調査した。調査結果は図 1.10 に示すように、食料品（74.6%）が最も多く、次いで日用品（53%）、家電製品（36.7%）の順で多かった。この結果は、CFP は消費者が日常的に目にする機会が多い食品や日用品に CFP を表示することが、CFP が目指す低炭素社会の実現に効果的であることを示している。

現在の CFP 制度には、主に以下の三つの問題点がある。

- CFP が公開されている商品の数が少ない。2016 年 10 月の時点で、CFP コミュニケーションプログラムに CFP を登録した製品 128 品であり、その中、食品関連と生活用品は 79 品であった。なお、CFP コミュニケーションプログラムで、CFP 検証に合格し、CFP 宣言を認定された製品の累積総数は 2016 年 10 月までで 1280 品である。
- CFP の認知度が低い。消費者の行動は CFP 制度にとって重要であり、低炭素社会の構築には消費者の CO_2 排出量低減に向けた行動が不可欠である。しかし、現在の調査結果 [24] では、ほどんとの消費者は CFP を認識していない。
- CFP の大きさを把握しにくい。CFP の大きさの基準がない。CFP を使って CO_2 排出量低減の行動を消費者に動機つけるには、CFP の大小を相対的に評価する必要がある。現在の CFP 制度は CO_2 排出量を表示しているだけであり、CO_2 の削減量、低炭素の基準量等の基準がないため、消費者にとって、CFP から CO_2 の排出量の量的に把握できない。

これらの問題は相互に関連している。これらの問題に対して、CFP が計算された商品数が多くなるならば、CFP の認知度は向上し、CFP の大小の商品間比較が可能になると考えられる。

1.4　研究目的

本研究は、現在の CFP 制度の問題を解決するためには、次の三つを目的として行った。

1. CFP の計算効率を改善する。

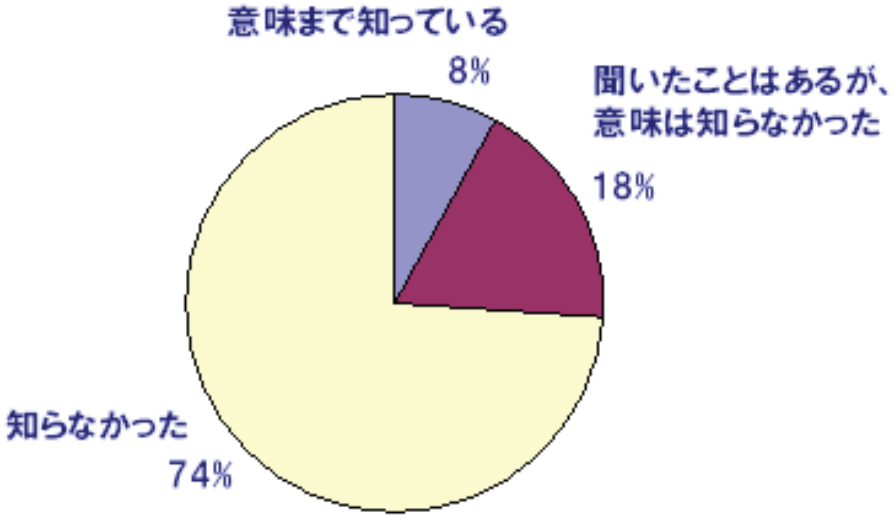


図 1.9　CFP 認知度調査結果 [24]

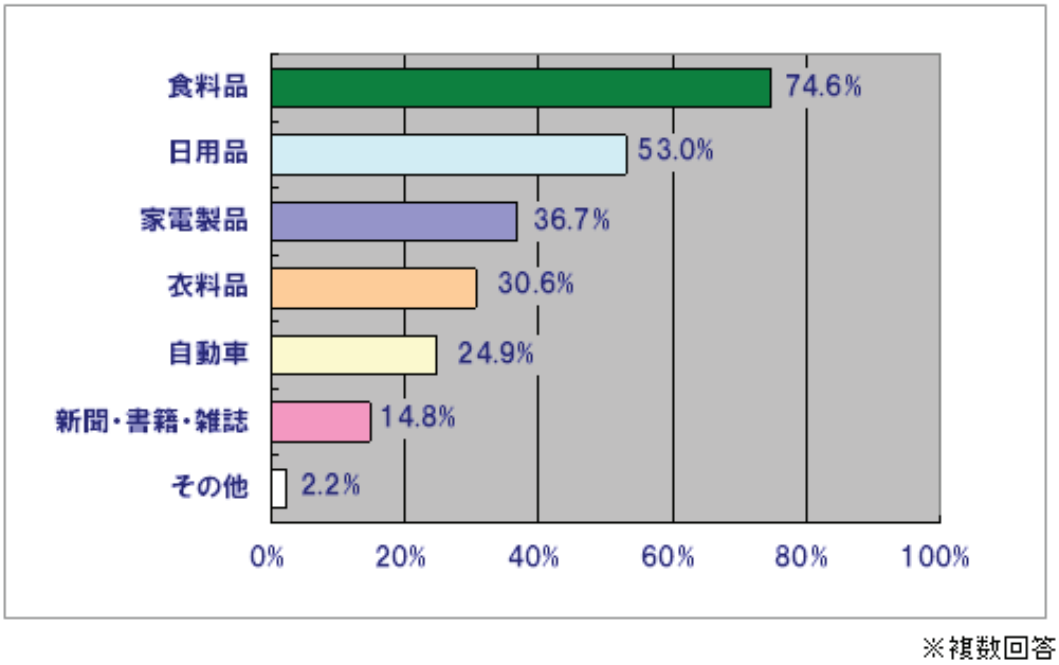


図 1.10　「カーボンフットプリントを表示したほうがよいと思う商品の種類」の調査結果 [24]

2. CFP の評価基準を作り出す方法を開発する。

3. CFP の認知度を向上させる。

CFP 計算の効率を改善するため、本研究では、CFP 自動計算システムを構築した。この計算システムでは、計算者の計算作業を効率化するため、生産メーカーから収集されたデータの読み込み、CO_2 原単位の検索、CFP 計算式の組み立て等の作業をほぼ自動で完成する。本計算システムを使って、CFP の計算方法を学び始めた直後の 4 人の学生と筆者らは、30 社が製造した 130 品の CFP を約 1 ヶ月で計算できるようになった。

現代は商品に関する情報が多量にウェブに公開されている。このようなウェブに公開されている情報から CFP を推定できるようになると、メーカーに直接頼らずに、商品の CFP を求められるようになる可能性がある。本研究ではテキストマイニングを用いて、ウェブで公開されたオープンデータを使って、製品の CFP の範囲を計算する手法を提案する。この手法により、CFP を計算され製品を増やす。本手法では、ディープラーニングに基づく CFP に関するデータをウェブサイトから認識、抽出するアルゴリズムを開発し、大量なデータを収集できるによって、大量な製品の CFP の範囲を推定できる。

CFP 社会認知度を向上させるために、本研究ではコープさっぽろと共同し、新たな CFP 取組を実行する。この取組では、130 品の食品・日用品の CFP を計算し、CFP が消費者の目につくように、CFP を消費者の健康と組み合わせて、包装のデザインを工夫した。また、取組の意図をパンフレットで紹介した。この取組の効果を、2014 年と 2015 年に二回のアンケート調査で調べた。

本論文の構成では、第 2 章は CFP 計算の効率を改善するための CFP 自動計算システムを述べる。第 3 章は CFP 製品を増やすためテキストマイニングに基づいたカーボンフットプリントの計算方法を述べる。第 4 章は本研究とコープさっぽろが行うカーボンフットプリント認知度向上の取り組みと取組の効果についてアンケート調査結果を述べる。

第 2 章

カーボンフットプリント計算システム

2.1 計算システムの設計方針

CFP を計算する時の作業は主に四つの段階からなる。

1. 各原材料の CO_2 原単位を取得する。
2. 各原材料と製品の輸送距離を取得する。
3. CFP 計算式を Excel へ書き込む。
4. 計算結果を確認と追記・修正する。

CFP をより多くの消費者に認知してもらうためには、より多くの製品の CFP を計算し、それを消費者に知らせる作業が必要である。より多くの製品で CFP を計算するには、より効率的に CFP を計算する手段が必要である。

永野研究室では、2009 年〜2011 年頃までは、CFP を計算する人が、CFP に関わるデータが記入されたデータシートを読み、手作業で CO_2 原単位を探し、ウェブブラウザから Google Maps 等へ出発地と到着地の住所を入力して輸送距離を取得し、Excel に計算式を手動で入力していた。

データシートのデータは生産メーカーに記入してもらっていた。2011 年から、永野研究室では、これらの手作業の幾つかを計算機で自動化する試みを進めてきた。その結果、CFP の計算を効率的に行う、図 2.1 を示す計算システムを製作した。この計算システムのインターフェースを図 2.2 に示す。この計算システムにおいて自動化した作業は、データシートからのデータ読み込み、CO_2 原単位の読み込み、輸送距離の自動取得等である。しかし、このシステムは、CFP の計算方法について理解が十分でなかったため、使いやすい計算システムではなかった。

本研究では、CFP の計算自体の特性と、表 2.1 にまとめた既存の計算システムの問題点を考察し、CFP を効率的に計算できる計算システムを製作した。手計算と既存の計算システムの問題は、計算に必要なデータを整理する仕組みにあった。製品の CFP の計算では、製品の原材料・輸送・生産の三段階のデータを用いている。原材料データは各原材料の重量と CO_2 原単位、輸送データは各原材料と製品の調達先と輸送手段、生産データは製品を生産する工場の産量と電力・重油・ガス等のエネルギーの消費量である。手動で計算するとき、計算者は各段階で使用するデータと行う計算について、深く理解する必要がある。そして、各段階で使用するデータについて、各原材料の CO_2 原単位、輸送距離等のデータを一つずつ手動で取得する必要がある。その結果、CFP を計算する前の段階で、データを整理するために長い時間と手間暇が費やされてしまう。

本計算システムを、CFP 計算を効率的に計算するために、計算の高速化と操作の利便性向上を目的に設計した。そのため、本システムにおける CFP 計算の各計算用データの読み取りと計式の組立と書き込みをほぼ自動化するよう設計した。また、収集した一次データには、誤りや不

表 2.1　既存システムの問題点

	データ入力の問題	CO_2 原単位検索の問題	操作の問題
CCalC	計算する人をデータを整理し、手動で入力するため、データ入力の効率が悪い	手動で検索するため、検索の効率が悪い	各段階のデータを手動で入力する、窓の変換が多い
GaBi	計算する人をデータを整理し、手動で入力するため、データ入力の効率が悪い	手動で検索するため、検索の効率が悪い	すべでの計算段階と計算式を手動で作成するため、操作が複雑
従来の永野システム [25]	自動で読み取ったデータを修正できない	CO_2 原単位の検索結果を修正できない	操作できる部分が少ない、問題あるデータを修正できない

足があることは避けられないと考え、計算データを確認と追記・修正し易いユーザーインターフェースを実現した。

本システムを従来の計算に対して改良した点は、

1. データの追記と修正が容易なユーザーインターフェースの採用
2. データ検索機能の高性能化と類似語検索機能の追加
3. CFP 計算用の原単位データベースと鉄道、船輸送距離データベースの内蔵化

である。

2.2 計算システムの機能

本システムのおけるデータ処理は「データ取得」と「計算結果表示」に大きく分けられる。CFP 計算システムの構成を図 2.3 に、計算の流れを図 2.4 に示す。

「データ取得」では、データ収集シートに記述されたデータ、CO_2 原単位、輸送距離を自動的に取得する。「計算結果表示」では取得したデータを用いて、CFP 計算式を自動的に組み立て、計算結果を表示する。CFP の計算結果はマイクロソフト社の Excel 形式ファイルとして保存する。

本システムを使って CFP を計算する人の作業は、主に、CFP に計算式の正誤の確認と、CFP の計算式に誤りがある場合の修正作業である。この確認と修正作業を行うユーザーインターフェースは利便性を重視して設計した。本システムのユーザーインターフェースは、図 2.5 に示す原材料データ確認インターフェースと図 2.6 に示す計算結果表示インターフェースから構成されている。原材料データ確認インターフェースでは、データ収集シートから取得した製品の各原材料のデータと輸送データを表示する。計算者はこのインターフェースで各原材料に関するデータを修正・追記する。計算結果表示インターフェースでは、製品の生産の流れに沿って原材料生産、輸送、工場、製品輸送の順番で各 CFP の計算結果を提示することで、計算者が各データの位置づけを把握しやすくした。データの位置づけを容易に把握ことで、データの修正・追記もしやすくなると考えたためである。また、本システムの操作を容易に習得できるように、多くの処理はマウス操作で処理できるようになっている。なお、本システムを用いて CFP を計算するには、計算者は原材料に関する基本的知識、輸送経路に関する地理の基本知識、第 1 章の式 (1.1)

に示した CFP の基本計算式の理解を備える必要がある。原材料に関する基本的知識には、ある種の原材料は複数の異なる原材料名があること等である。例えば、「薄力粉」は「小麦粉」の一種であるという知識である。

本システムは、図 2.7 で示すデータ収集シートから、CFP の計算に必要な製品の原材料構成、原材料の調達先、工場の生産量とエネルギーの消費量等を自動的に読み取る。データ収集シートには、メーカー担当者が該当製品の CFP の計算に必要なデータを記入する。自動的に読み取るデータは、原材料名、それぞれの製品に占める重量比、原材料の調達先、輸送経路、輸送方法、工場で生産する全製の重量、電力消費量、使用重油量、使用上水量、排出下水量、都市ガス消費量、LP ガス消費量等である。

各原材料の CO_2 原単位は、CO_2 原単位データベースから自動的に取得する。CO_2 原単位データベースの例を図 2.8 に示す。本システムの CO_2 原単位データベースは、味の素グループ版「食品関連材料 CO_2 排出係数データベース」のデータに基づいている。この CO_2 原単位データベースには、1162 の原材料について CO_2 原単位が記載されている。このデータベースは公開されており、本研究では味の素株式会社よりこのデータベースの使用の許可を得て、計算システムを製作した。

本システムの CO_2 原単位データベースは、類似語検索機能を備えている。この類似語検索機能により、データ収集シートに記載された原材料が CO_2 原単位データベースにない場合にも、より広い範囲で CO_2 原単位を自動的に検索できるようになっている。データ収集シートに記入されている原材料名と CO_2 原単位データベースに記入されている原材料名が一致しない場合が

多い。例えば、同じ原材料である「小麦粉」を CO_2 原単位データベースでは「小麦粉」を記入し、データ収集シートでは「薄力粉」記入されている。このような時でも原材料の CO_2 原単位を検索できるようにするため、この類似語検索機能を開発した。本システムの類似語検索機能は、独立行政法人情報通信研究機構が開発した類似語データベース「日本語 WordNet」[26] を用いて作成した。この類似語検索機能により、例えば、「薄力粉」や「強力粉」から「小麦粉」の CO_2 原単位を取得できるようになっている。また、この類似語検索機能には、図 2.9 に示す、データ収集シートで記入された原材料名を CO_2 原単位データベースの原材料に対応づける辞書追加機能も組み込んだ。

原材料や製品の輸送距離は、データ収集シートに記載され調達先の出発住所と到着住所に基づいて、自動的に取得する。輸送方法のトラック、鉄道、船（国内)、船（海外）により、異なる方法で輸送距離を取得する。輸送方法がトラックの場合、本システムは Google Maps と通信し、有料道路が使わない輸送経路とその移動距離を取得する。輸送方法が鉄道の場合、日本貨物鉄道株式会社から提供された「貨物駅間距離データベース」を用いて、各駅間の距離を取得する。輸送方法が船の場合、海上保安庁の「距離法」[27] のデータを基づいて、著者が港間距離をまとめて作成した「国内港間距離データベース」と「海外地域間距離データベース」を用いて、各港と各地域間の距離を取得する。図 2.10 で示す、「国内港間距離データベース」には、日本国内各都道府県代表する 52 の港間距離が記載されている。図 2.11 で示す、「海外地域間距離データベース」には、173 の国と地域から東京港への距離が記載されている。

取得したデータを原材料データ確認インターフェースで正誤を確認した後、本システムは CFP

を計算する。計算結果を計算結果表示インターフェースで確認できる。CO_2 原単位や輸送経路等に誤りがある場合、原材料データ確認インターフェースに戻して、修正することができる。そして、確認できた CFP の計算結果を Excel ファイルに出力する。出力された CFP の計算結果の例を図 2.12 に示す。本システムが出力する CFP の計算結果は原材料、原材料輸送、工場、製品輸送の四つの段階に分けられている。原材料の計算結果には、CO_2 原単位データベースにおけるその原材料の分類記号も同時に出力されている。また、本システムは、すでに計算結果が出力した製品の CFP を修正と再計算できるようにするため、取得した CO_2 原単位や輸送距離が記載される CFP 形式ファイルも Excel ファイルと同時に出力する。この CFP 形式ファイルの例を図 2.13 に示す。この CFP 形式ファイルのデータ書式は本研究で独自に設計した。本システムはこの CFP ファイルを読み込み、CFP 計算結果を出力以前のユーザーインターフェースを表示することができる。

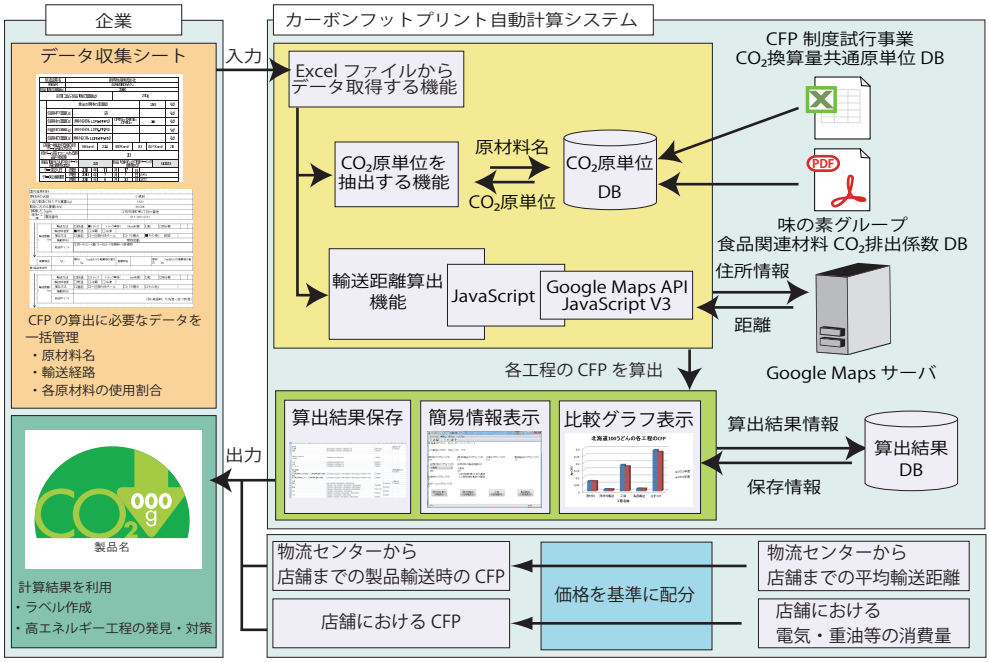


図 2.1　永野研究室で開発した従来の CFP 計算システム

図 2.2　永野研究室で開発した従来の CFP 計算システムのインターフェース

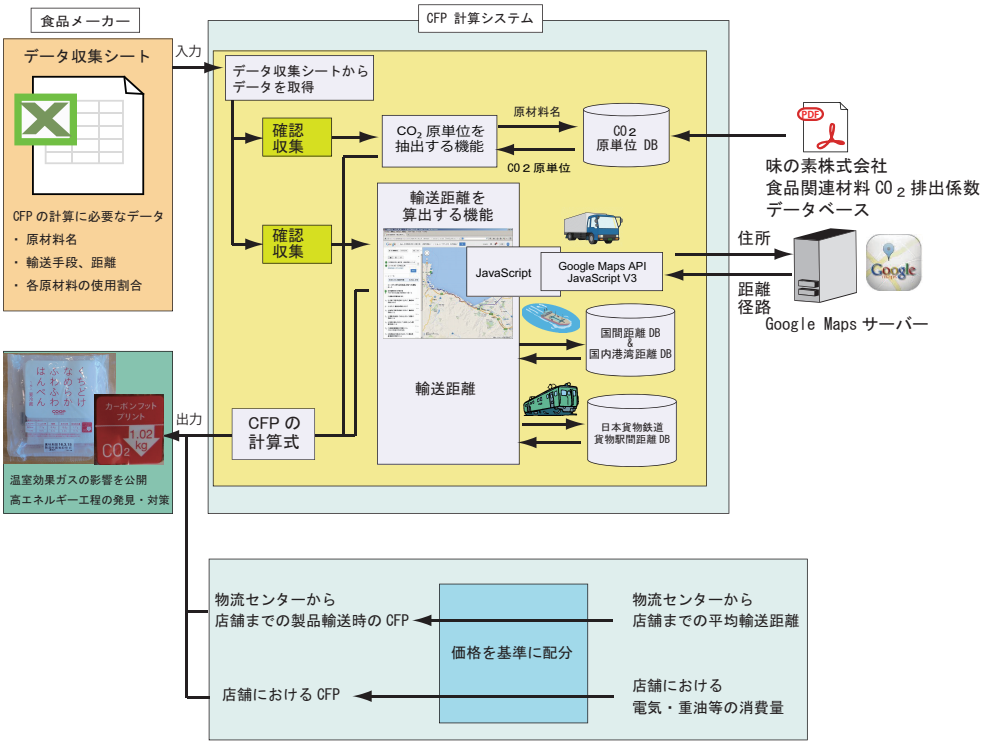


図 2.3　CFP 計算システムの構成

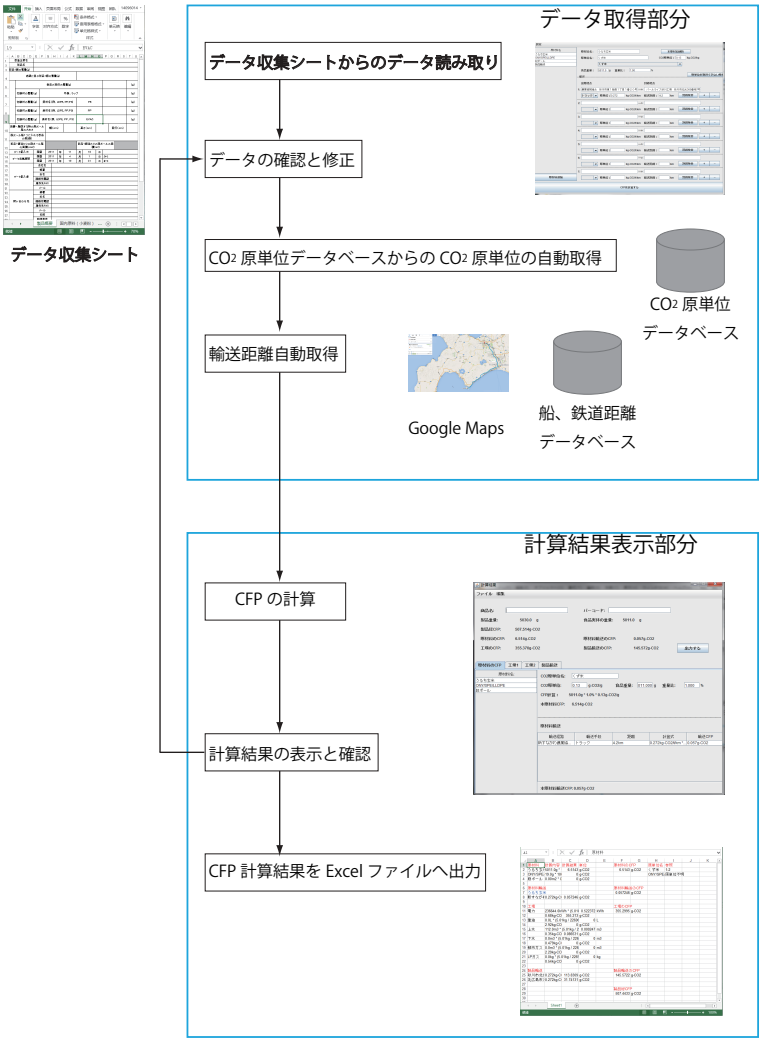


図 2.4　CFP 計算システムの計算の流れ

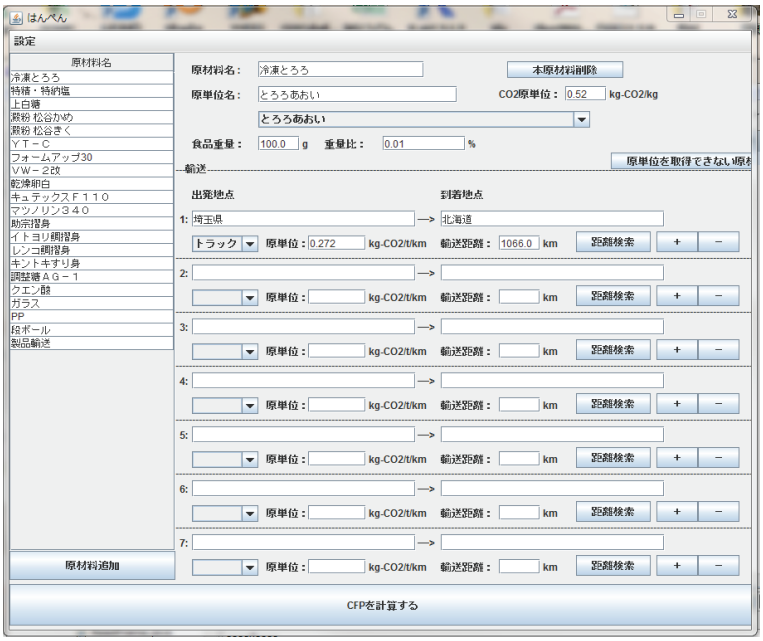


図 2.5　原材料データ確認インターフェース

図 2.6　計算結果表示インターフェース

製造企業名								
製品名								
製品1個の重量(g)								
店頭に並ぶ製品1個の重量(g)								
	食品の実体の重量(g)							(g)
	包装材の重量(g)	ガラス						(g)
	包装材の重量(g)	素材名(例, LDPE, PP,PS)		トレー(PS)				(g)
	包装材の重量(g)	素材名(例, LDPE, PP,PS)		ラップ(PP,EVAC)				(g)
	包装材の重量(g)	素材名(例, LDPE, PP, PS)		シール(紙)				(g)
店舗へ輸送する時の段ボール箱の大きさ		幅(cm)		高さ(cm)		奥行(cm)		
段ボール箱1つに入れる製品の数(個)		12						
製品1個当たりの段ボール箱の面積(cm2)		0		製品1個当たりの段ボールの面積(m2)		0		
データ記入日	西暦		年		月		日	
データ収集期間	西暦		年		月		日	
	西暦		年		月		日	
データ記入者	会社名							
	部署							
	氏名							
	連絡先電話							
	連作先FAX							
	メール							
問い合わせ先	部署							
	氏名							
	連絡先電話							
	連作先FAX							
	メール							
製品の製造工場	名前							
	郵便番号							
	住所							
	電話番号							
国内産原材料	D1	名前						
	D2	名前						
	D3	名前						
	D4	名前						
	D5	名前						
	D6	名前						

図 2.7　データ収集シート

1	玄米	t	0.45	食品関連材料CO2排...	1-1
2	くず米	t	0.13	食品関連材料CO2排...	1-2
3	稲わら	t	0.04	食品関連材料CO2排...	1-3
4	小麦（国産）	t	0.42	食品関連材料CO2排...	2-4
5	六条大麦（国産）	t	0.37	食品関連材料CO2排...	2-5
6	二条大麦（ビール麦）	t	0.39	食品関連材料CO2排...	2-6
7	裸麦	t	0.44	食品関連材料CO2排...	2-7
8	かんしょ	t	0.17	食品関連材料CO2排...	3-8
9	ばれいしょ	t	0.09	食品関連材料CO2排...	3-9
10	大豆（国産）	t	0.43	食品関連材料CO2排...	4-10
11	えんどう（種）	t	0.4	食品関連材料CO2排...	4-11
12	そら豆	t	0.21	食品関連材料CO2排...	4-12
13	いんげん豆	t	0.49	食品関連材料CO2排...	4-13
14	小豆	t	0.52	食品関連材料CO2排...	4-14

図 2.8　CO_2 原単位データベース

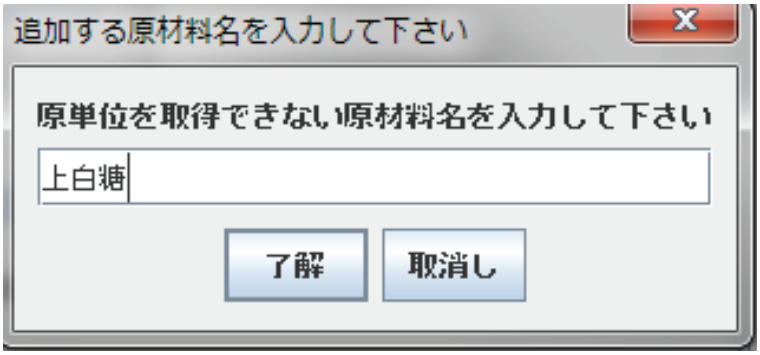


図 2.9　類似語辞書追加機能

	北海道苫小	小樽港	石狩新港	函館港	釧路港	白老港	青森県青森	岩手県大船
北海道苫小牧港	0						211	509
小樽港							422	772
石狩新港							431	781
函館港							102	390
釧路港							443	501
白老港							203	420
青森県青森港	211	422	431	102	443	203	0	474
岩手県大船渡港	509	772	781	390	501	420	474	0
宮城仙台塩釜港	611	872	881	490	601	520	576	135
福島県小名浜港	387	990	999	608	719	638	274	656
秋田県秋田港	463	517	526	269	626	386	350	732
山形県酒田港	724	593	602	345	702	462	689	248
茨城県大洗	745							
茨城県鹿島港	828	1150	1159	768	879	798	793	352
千葉県千葉港	1082	1374	1383	992	1103	1022	1052	606
東京都東京港	1080	1372	1381	990	1101	1020	1050	604
神奈川県横浜港	1063	1355	1364	973	1084	1003	1033	587
新潟県新潟港	567	770	779	522	879	639	454	835
富山県伏木富山港	761	963	972	715	1072	932	644	1030
石川県金沢港	817	963	972	715	1072	932	700	1085
福井県敦賀港	939	1141	1150	893	1250	1010	822	1208

図 2.10　国内港間距離データベース

200港			海里	[単位:km]
所属地域(日本語)	所属地域(コード)	港所在都市名称(英語)		JPN
日本	JPN	Tokyo		
アルバ	ABW	Willemstad	8808	16,312
アフガニスタン	AFG	Karachi	5784	10,712
アンゴラ	AGO	Luanda	10152	18,802
アンギラ	AIA	Blowing Point		0
アルバニア	ALB	Vlore		0
アンドラ	AND	Barcelona	9553	17,692
アンティル(蘭領	ANT	Willemstad	8808	16,312
アラブ首長国連	ARE	Dubayy	6359	11,777
アルゼンチン	ARG	Buenos Aires	12086	22,383
アルメニア	ARM	Batumi	9346	17,309
アンティグア・バ	ATG	Saint John	10468	19,387
オーストラリア	AUS	Melbourne	4912	9,097
オーストリア	AUT	Hamburg	11544	21,379
アゼルバイジャン	AZE	Jazireh-Y e K	6669	12,351
ブルンジ	BDI	Dar Es Salaam	6922	12,820
ベルギー	BEL	Antwerp	11281	20,892
ベナン	BEN	Cotonou	11018	20,405
ブルキナファソ	BFA	Accra		0
バングラディシュ	BGD	Chittagong	4392	8,134
ブルガリア	BGR	Burgas	8873	16,433
バーレーン	BHR	Bahrain		0
バハマ	BHS	Nassau	9467	17,533
ボスニア・ヘルツ	BIH	Dubrovnik		0

図 2.11　海外地域間距離データベース

原材料	計算内容	計算結果	単位		原材料のCFP		原単位名	参照
うるち玄米	5011.0g * 1.00% * 0.13kg-CO2/kg	6.5143	g-CO2		6.5143	g-CO2	くず米	1-2
ONY/SPE/LLDPE	19.0g * 100.00% * 0.0kg-CO2/kg	0	g-CO2				ONY/SPE/	原単位不
段ボール	0.00m2 * 0.33kg-CO2/m2 * 1000	0	g-CO2					
原材料輸送					原材料輸送のCFP			
うるち玄米					0.057246	g-CO2		
新すながわ農業協	0.272kg-CO2/t/km * 4.2km * 5011.0g * 1.00% * 10^(-6) * 10^3	0.057245664	g-CO2					
工場					工場のCFP			
電力	236644.0kWh * (5.01kg / 2269621.0kg)		0.522372	kWh	355.2995	g-CO2		
	0.68kg-CO2/kWh * (236644.0kWh * (5.01kg / 2269621.0kg))* 1000	355.2129537	g-CO2					
重油	0.0L * (5.01kg / 2269621.0kg)		0	L				
	2.92kg-CO2/L * (0.0L * (5.01kg / 2269621.0kg))* 1000	0	g-CO2					
上水	112.0m3 * (5.01kg / 2269621.0kg)		0.000247	m3				
	0.35kg-CO2/m3 * (112.0m3 * (5.01kg / 2269621.0kg))* 1000	0.086530747	g-CO2					
下水	0.0m3 * (5.01kg / 2269621.0kg)		0	m3				
	0.479kg-CO2/m3 * (0.0m3 * (5.01kg / 2269621.0kg))* 1000	0	g-CO2					
都市ガス	0.0m3 * (5.01kg / 2269621.0kg)		0	m3				
	2.29kg-CO2/m3 * (0.0m3 * (5.01kg / 2269621.0kg))* 1000	0	g-CO2					
LPガス	0.0kg * (5.01kg / 2269621.0kg)		0	kg				
	0.54kg-CO2/kg * (0.0kg * (5.01kg / 2269621.0kg))* 1000	0	g-CO2					
製品輸送					製品輸送のCFP			
砂川市北光３６５	0.272kg-CO2/t/km * 83.2km * 5030.0g * 10^(-6) * 10^3	113.830912	g-CO2		145.5722	g-CO2		
北広島市大曲工業	0.272kg-CO2/t/km * 23.2km * 5030.0g * 10^(-6) * 10^3	31.741312	g-CO2					
					製品総CFP			
					507.4433	g-CO2		

図 2.12　CFP の計算結果

製品名.cfp

```
製品名|
355.3704025121375
5011.0
5030.0
6.5143
0.057245664000000016
145.572224
507.5141721761375
PTranspot
--------
砂川市北光３６５番地７号
北広島市大曲工業団地６丁目３－３
トラック
0.272
83.2
--------
北広島市大曲工業団地６丁目３－３
江別市東野幌28-7
トラック
0.272
23.2
Factory
2269621.0
236644.0
0.68
0.0
2.92
112.0
0.35
0.0
0.479
0.0
2.29
0.0
0.54
RawMaterial
----R----
```

図 2.13　独自に設計した CFP 形式ファイルの例

2.3 計算システムの性能評価

本計算システムによる CFP 計算作業の改善を評価するために、10 品について、手作業で CFP を計算する場合と本計算システムを使って計算する場合で作業時間を比較した。この作業時間の比較結果を表 2.2 に示す。この評価では CFP の計算が多様になるように、製品を選んだ。この作業時間の比較実験は、CFP の計算に 3 年間従事し、約 100 点の食品の CFP を計算した経験がある大学院生 1 人により行われた。表 2.2 の手作業の計算は、CO_2 原単位データベースの表から目的の CO_2 原単位を読み取る作業、ウェブブラウザ上で Google Maps に住所を入力して輸送距離を取得する作業、CFP の計算式を Excel に入力する作業が主である。本計算システムを使うことにより、これらの作業がほぼ自動化される。表 2.2 の CFP 計算システムを使った作業時間は、本計算システムが提示するデータの正誤を確認し、問題があれば計算システムへ入力するデータを適切な表現に修正する作業に費やした時間が主である。

原材料数と作業時間の関係を図 2.14 で示す。輸送経路における節点の数と作業時間の関係を図 2.15 で示す。図 2.14 と図 2.15 によれば、CFP の計算時間は、原材料の数や輸送経路における節点の数に影響される。特に輸送経路における節点の数が多いほど、CFP の計算時間が長い傾向がある。表 2.2 によれば、手動計算と比べて、本計算システムの使用により、CFP の計算時間が 10% ～ 22% に低減された。

本システムの性能を考察するため、本システムと現在広く利用されているの LCA 計算システムを製品データの読み取り、CO_2 原単位の検索、輸送データ取得、計算の流れの可変性、計算できるデータについて比較した。比較結果を表 2.3 にまとめる。比較する LCA 計算システムはマ

ンチェスター大学が開発した CCaLC と thinkstep 株式会社が開発した GaBi である。

CCaLC と GaBi では CFP を計算するデータを全部手動で入力する必要がある。また、CCaLC と GaBi は製品輸送のデータも手動で入力するため、輸送データを計算者は計算システム以外の方法で取得する必要がある。それに対して、本計算システムは製品データを記載したデータ収集シートから必要なデータを自動的に読み取るため、製品データの取得と CO_2 原単位の検索はほぼ自動で完成する。本計算システムは、製品輸送のデータも Google Maps と輸送距離データベースから取得する。計算者は計算データを手動で入力する作業はほとんどなくなっている。

本計算システムと CCaLC では、CFP の計算の段階の構成はシステムの内部で決まっている。本システムでは原材料、原材料輸送、製品生産、製品輸送の四つの段階の CFP を計算する。CCaLC では、原材料、原材料輸送、製品生産、製品使用、製品廃棄の五つのつの段階の CFP を計算する。GaBi では、固定した計算の流れがなく、計算する内容を全て計算者が決める。そのため、GaBi は計算の自由度が高いと同時に計算者に対する要求も高い。GaBi を使用する計算者は計算の流れを深く理解する必要がある。

CCaLC と GaBi は CFP だけがなく、製品の LCA に関するデータを全て計算できる。例えば酸性雨への影響や、環境への毒性等を評価できる。本計算システムは CO_2 原単位データベース以外の LCA 関する原単位データベースを収集していないため、現在では CFP 以外の LCA データを計算できない。

CFP の計算を高速化するために開発した本計算システムでは、CFP の一次データの読み取

表 2.2　CFP 計算の作業時間の比較

製品番号	原材料数	原材料輸送ルート数	手作業の作業時間 (分)	CFP 計算システムを使った作業時間 (分)	作業時間の短縮比 (%)
1	2	3	10	1	10
2	2	4	12	2	17
3	3	14	62	10	16
4	4	17	58	9	16
5	5	7	35	5	14
6	6	7	32	4	13
7	6	12	30	5	17
8	7	9	30	4	13
9	10	19	55	12	22
10	11	22	65	13	20

り、CO_2 原単位と輸送距離の二次データの収集、CFP の計算結果と計算式の出力のデータ処理が、ほぼ自動化されている。CFP を計算する人の作業は、計算システムが採用する CO_2 原単位の材料名の妥当性とメーカー担当者が提出した輸送経路の妥当性を確認し、必要に応じて代替案を計算機システムに入力するのみである。

本計算システムを使って、CFP の計算方法を学び始めた直後の 4 人の学生と著者らは、30 社が製造した 130 品の CFP を約 1 ヶ月で計算した。これは、本計算システムの利便性が向上した成果であると、著者は考えている。

表 2.3　既存システムとの比較結果

	製品データの読み取り	CO_2 原単位の検索	輸送データの取得	計算の流れの可変性	計算できるデータ
CCalC	手動入力	手動検索	システムで取得できない	固定	LCA 全般
GaBi	手動入力	手動検索	システムで取得できない	可変	LCA 全般
本システム	自動読取	自動検索	Google Maps と輸送距離データベースを用いて取得する	固定	CFP のみ

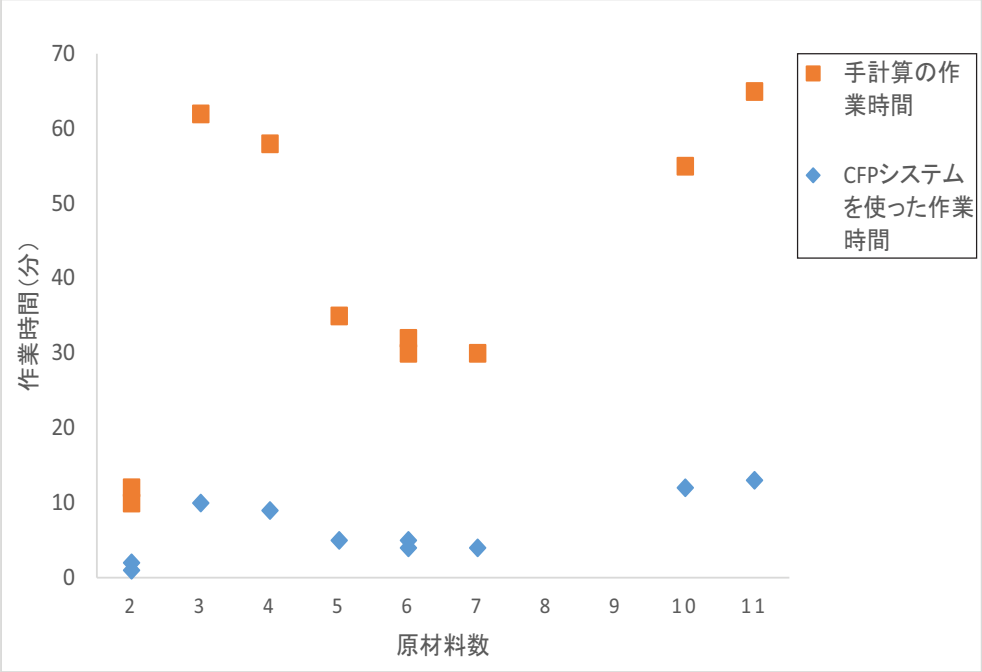


図 2.14　原材料数と作業時間の関係図

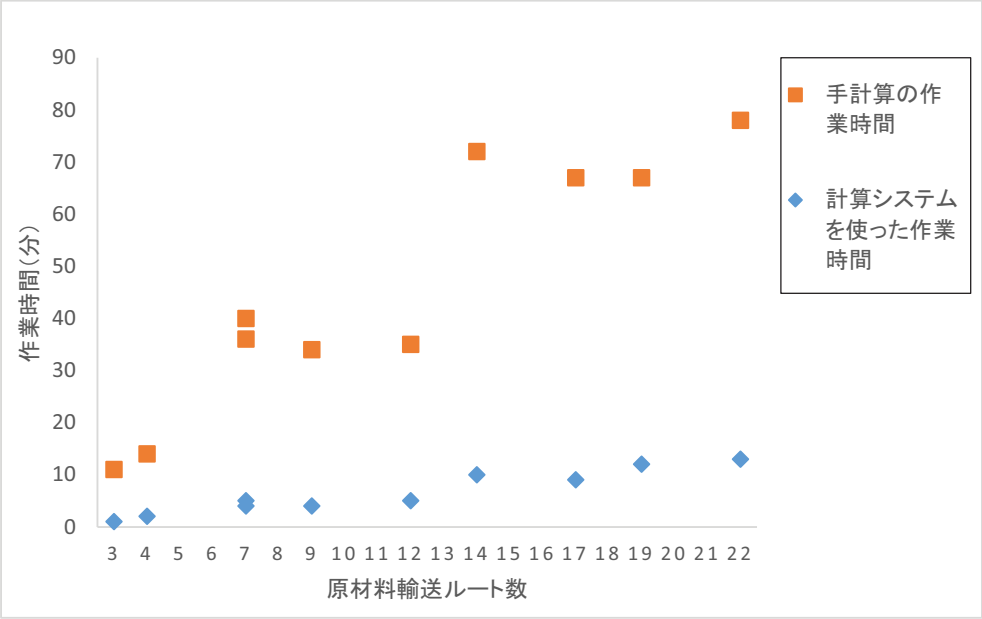


図 2.15　輸送経路における節点の数と作業時間の関係図

第 3 章

テキストマイニングに基づいたカーボンフットプリントの計算方法

3.1 計算方法の方針

現在の CFP 制度の問題の一つは、CFP が公開されている製品の数が少ない点である。2016 年 10 月の時点で、CFP コミュニケーションプログラムに CFP を登録した製品は 128 品であり、その中で、食品関連と生活用品は 79 品であった。なお、CFP コミュニケーションプログラムにおいて、CFP 検証に合格し、CFP 宣言を認定された製品の累積総数は 2016 年 10 月までで 1280 品である。

CFP の認知度を向上するためには、CFP を見えられるようにした製品がより多く市場に出まわる必要がある。そのためには、より多くの製品において CFP が計算される必要がある。しかし、第 1 章で述べたように、製品の CFP の計算では、CFP 計算データの収集が難しい問題がある。製品の生産メーカーが CFP 計算データを収集するには、コストと時間がかかる。そして、生産メーカーはデータを製品一つずつ収集するため、データを収集する効率が悪い。

一方、現代は、ビッグデータを扱う環境が拡大・発展している。インターネット上では、各製品の生産メーカーのホームページ、アマゾンや楽天等のネットショッピングのウェブサイトに多量の製品情報が公開されている。例えば、日清食品はホームページに、図 3.1 に示すように、製品の原材料、原材料産地のデータを公開している [28]。アマゾンの食品販売のウェブサイトでは、図 3.2 に示すように、食品の原材料情報を公開している [28]。このようなウェブサイトから、製品の CFP を計算するデータが抽出できれば、生産メーカーに直接頼らずに、製品の CFP を求められる可能性がある。

本研究では、テキストマイニングを用いて、ウェブに公開されているデータを使って、製品の CFP の範囲を計算する方法を提案する。テキストマイニングを用いた CFP の推定方法の概略を図 3.3 に示す。本手法はウェブからデータを収集する部分と CFP を計算する部分からなる。

本手法でウェブから収集されるデータには、オープンデータベースとして統一した表現形式に整理されているデータと、一般のウェブサイトにおいて公開されているテキストデータがある。農林水産省が公開している農林水産物輸出入データベース [29] を図 3.4 に示す。このようなオープンデータベースは、データが整理されたため、CFP の計算に直接かつ容易に使用できる。一方、一般のウェブサイトに公開されているテキストデータは、図 3.1 を示す日清食品のデータと図 3.2 を示すアマゾンで販売される製品のデータのように、表現がウェブサイト毎に異なるため、直接使用できない。そこで、著者は、一般のウェブサイトから CFP に関するデータを自動的に抜き出すテキストマイニングの方法を開発した。CFP の計算では、このように収集した商品データを用いて、該当商品種の CFP の最大値、最小値、平均値を計算する。

本手法では、個々の製品の CFP を計算せず、収集した商品データから商品種の CFP を計算するため、その商品種の CFP 計算結果にはばらつきがでる。この手法では、ウェブに公開されているデータを使って CFP の分布範囲を推定し、その分布範囲が製品の CFP の基準になる。そして、図 3.5 に示すように、製品の CFP の基準によって、消費者を CFP による製品の位置づけを理解し、消費者が製品の環境影響の価値を認知できるようになると期待できる。

本章の 3.2 節では、原材料情報をウェブサイトから正しく抽出するために提案する再帰自己符号化器と N-gram モデル解析を組み合わせた固有表現抽出法を述べる。3.3 節では、3.2 節で述べた機械学習に基づいた原材料情報を収集する方法と、オープンデータベースから原材料の産地情報と生産情報を収集する方法を説明する。3.4 節では、商品種の CFP の最大値、最小値、平均値を計算する計算方法について述べる。3.5 節において本計算方法の性能を評価する。

めん

原材料	小麦
主な原産国	オーストラリア、アメリカ、日本、カナダ
最終加工地	日本

原材料	植物油脂 (パーム)
主な原産国	マレーシア、インドネシア、タイ
最終加工地	日本

スープ

原材料	醤油 (大豆)
主な原産国	中国、インド、アメリカ、カナダ、日本
最終加工地	日本

原材料	醤油 (小麦)
主な原産国	アメリカ、カナダ、日本
最終加工地	日本

原材料	香辛料 (にんにく)
主な原産国	中国、インド、アメリカ
最終加工地	日本

原材料	香辛料 (コショウ)
主な原産国	インドネシア、マレーシア
最終加工地	日本

かやく

原材料	味付卵 (卵)
主な原産国	アメリカ、カナダ、ＥＵ、インド、中国
最終加工地	日本

原材料	味付豚肉 (豚)
主な原産国	日本、カナダ、中国
最終加工地	日本、中国

図 3.1　日清食品が公開している製品情報

日清食品 カップヌードル シーフードヌードル75g 5食セット

カップヌードル

100件のカスタマーレビュー | 5人が質問に回答しました

参考価格: ~~¥ 972~~

価格： ¥ 890 (¥ 178 / 個)

OFF: ¥ 82 (8%)

残り18点 ご注文はお早めに 在庫状況について

お届け予定日：11/17~21 通常配送を利用した場合のお届け予定日です。

小楽良品 がフラストレーション・フリー・パッケージで販売、発送します。 この出品商品にはコンビニ・ATM・ネットバンキング・電子マネー払いが利用できます。

新品の出品：1 ¥ 890より

サイズ: **5個**

5個	20個
¥ 890 (¥ 178 / 個)	¥ 3,272 (¥ 164 / 食)

- カロリー:1食当たり/323kcal
- 原材料:油揚げめん(小麦粉、植物油脂、食塩、たん白加水分解物、醤油)、スープ(ポークエキス、チキンエキス、糖類、香辛料、醤油、乳等を主要原料とする食品、野菜調味油、野菜エキス、たん白加水分解物、魚介エキス、香味油、食塩、オニオンパウダー、魚醤、ポーク調味油)、かやく(いか、キャベツ、魚肉練り製品、味付卵、ねぎ)、加工でん粉、調味料(アミノ酸等)、炭酸Ca、かんすい、増粘多糖類、酸味料、ベニコウジ色素、カラメル色素、乳化剤、酸化防止剤(ビタミンE)、香料、カロチノイド色素、炭酸Mg、香辛料抽出物、ビタミンB2、ビタミンB1、(原材料の一部に落花生、かに、ごま、ゼラチンを含む)
- 商品サイズ(高さx奥行x幅):214mm×96mm×288mm

> もっと見る

図 3.2　アマゾンが公開している食品の原材料情報

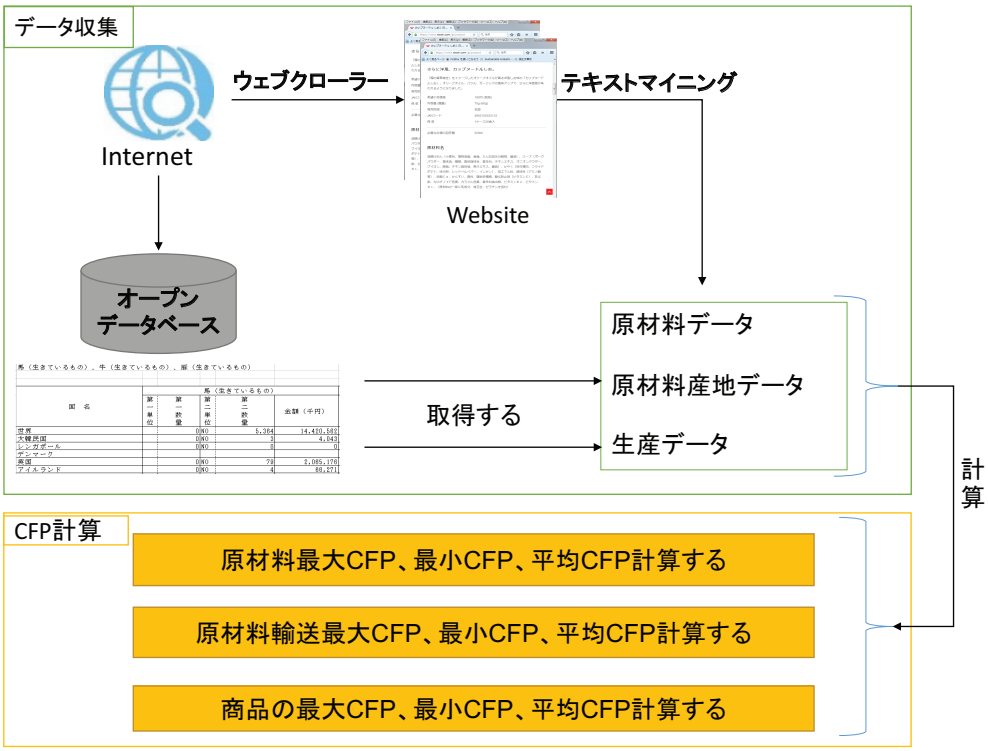


図 3.3　CFP の推定方法の概略図

動物の胃・腸・ぼうこう

国　名	動物の胃・腸・ぼうこう				
	第一単位	第一数量	第二単位	第二数量	金額（千円）
世界		0	KG	29,568,728	29,898,747
中華人民共和国		0	KG	2,293,743	11,393,717
モンゴル		0	KG	8,205	62,841
タイ		0	KG	13,830	8,573
フィリピン		0	KG	471,307	259,978
パキスタン		0	KG	10,455	42,397
ノルウェー		0	KG	116,582	130,840

図 3.4　農林水産省が公開している農林水産物輸出入データベース [29]

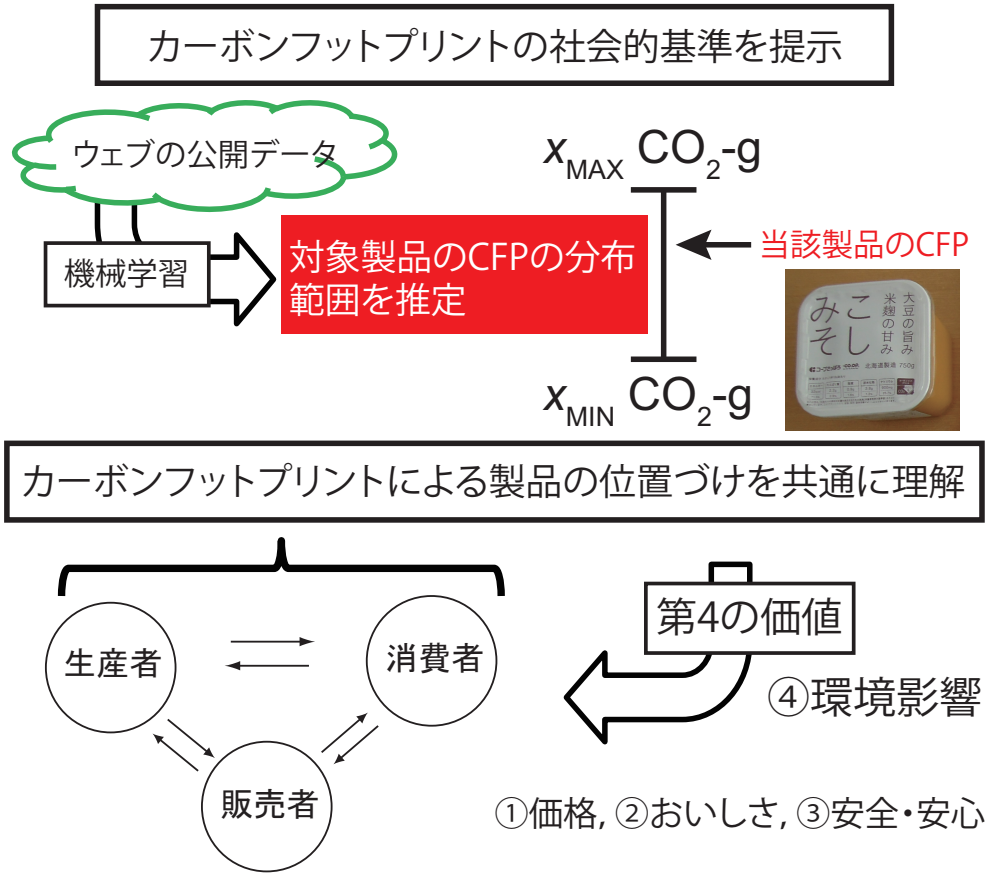


図 3.5　製品の CFP の基準と消費者への影響

3.2　原材料情報抽出ための機械学習

図 3.1 に示す日清食品のホームページや図 3.2 に示すアマゾンの食品販売のウェブサイト等の一般ウェブサイトでは、公開している原材料の情報の表現形式に統一性はない。このような一般ウェブサイトのテキストデータから、製品の原材料、製品名等の CFP 計算に関する情報を収集する技術には固有表現抽出がある。固有表現抽出は、テキストデータから人名や地名、製品等を表す固有表現を抽出する技術である。たとえば、固有表現抽出では、原材料名に対して「小麦粉」を抽出する。

固有表現の抽出では、人名や商品名などの固有表現がタグ付けされたテキストに対して、最大エントロピー法 [30]、Support Vector Machines(SVM)[31] などの機械学習を用いて抽出規則を自動的に学習する手法がある。SVM を使った機械学習では、過学習を起こすことなく汎化性能の高い学習性能と高い固有表現の抽出精度が報告されている [31][32]。しかし、報告されている SVM を使った固有表現抽出では、抽出できるテキストは、固有表現を含む単語や文字などの構成要素数が少ない場合に限られている課題があった。たとえば、山田他は、標的の固有表現とその前後 2 文字の品詞情報を用いてテキストを抽出していた [31]。中野と平井の報告では、構成要素数が 2 つの場合の再現率は 88.18% であったが、構成要素数が 3 つの場合になるとそれが 75.67% に低下していた [32]。

深層学習は、多層ニューラルネットを用いた機械学習である。深層学習の有効性は、2006 年頃から音声認識や画像認識 [33]、人工知能 [34] などの分野において、確認されてきた。自然言語処理においても、深層学習は、形態素解析 [35]、意味解析 [36]、機械翻訳 [37] などの課題に応用

されている。多層ニューラルネットには、制約ボルツマンマシンや自己符号化器等を使う手法がある [38]。

Socher et al. は、再帰自己符号化器 (Recurisive Autoencoder, RAE) を使って入力の文節や文から意味ベクトルを作成した [39]。この RAE は任意の語長に対応できる特徴がある。RAE によって作成された意味ベクトルには、入力テキストの意味に関する特徴が含まれる。この意味ベクトルを用いて入力テキストの表現が解析可能になる。本論文では、この RAE を利用する。

入力したテキストの意味を判定する時、その入力テキストだけでは意味を間違う可能性がある。同じ単語や文節でも、文脈によって、その単語や文節の意味は変化する。例えば、“ 味の素は調味料を生産する会社である。” の中の “ 味の素 ” は「会社名」を意味している。一方、“ この料理は味の素を入れている。” の中の “ 味の素 ” は「商品名」である。RAE のテキスト解析では、入力テキストとその隣のテキストとの文脈関係を用いていない。そのため、文脈による単語や文節の意味の違いは解析結果に反応されていない。

N-gram モデルは自然言語処理分野で利用されている確率的言語モデルの一種であり、文字列の並び順の合理性を確率的に表現できる [40]。N-gram モデル解析では、入力テキストの前後のテキストを解析するため、入力テキストの文脈を考慮した固有表現が可能になる。

原材料情報抽出ため、本論文は深層学習の一つの手法である RAE と文脈の関係解析に用いられている N-gram モデル解析を組み合わせて、任意単語長のテキストから固有表現を抽出する手法を提案する。提案手法では、まず、RAE を用いて入力テキストの意味ベクトルを作成し、入力テキストの表現を先行解析する。さらに、入力テキストと隣のテキストを N-gram モデルで解

析し、RAE による先行解析結果を修正する。この 2 つの処理の関係を図 3.6 に示す。RAE の処理では、あらかじめ形態素に分解された文節データを入力する。RAE はその文節データを表現する意味ベクトルを生成し、その意味ベクトルが分類 t_i に属する確率 d_{t_i} を出力する。この所属確率 d_{t_i} に基づいて、その文節データが所属する分類 $\widehat{t_i}$ を決定する。第 2 段階の N-gram モデル解析では、RAE の分類結果 $\widehat{t_i}$ を用いて、入力文節の前後の文節との文脈関係を解析し、分類結果を更新する。

3.2.1　再帰自己符号化器を用いたテキスト情報の分類

RAE は、文節や文を表現する意味ベクトルを使って、その文節や文の意味の分布を予測する多層ニューラルネットワークである [39]。RAE は複数の自己符号化器を階層状に重ねた構造になっている。RAE に入力するベクトルの次元には原理的な制限がないため、長い入力テキストにも対応できる特徴がある。

本研究の RAE は、入力の文節 $\boldsymbol{h}_i$ から意味ベクトル $\boldsymbol{p}_i$ を生成し、それを使ってテキスト情報を分類する。この処理を図 3.7 に示す。文節は単語に分解され、その単語は M 次元の単語ベクトル $\boldsymbol{c}_i = (\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, \ldots)$ で表現する。$\boldsymbol{c}_i \in \mathbb{R}^M$ である。RAE 二分木は、その単語ベクトルを葉節点とし、合成された意味ベクトル $\boldsymbol{p}_j = (\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3, \ldots)$ を親節点とする。$\boldsymbol{p}_j \in \mathbb{R}^M$ である。RAE 二分木の中で、ある親節点とその子節点は自己符号化器になっている。親節点の合成意味ベクトル $\boldsymbol{p}$ は、最とも下の第 1 層では

$$\boldsymbol{p}_1 = f\left(\boldsymbol{W}_1^{(1)}\left[\boldsymbol{c}_1, \boldsymbol{c}_2\right] + \boldsymbol{b}_1^{(1)}\right), \tag{3.1}$$

それより上の層では、

$$\boldsymbol{p}_{j+1} = f\left(\boldsymbol{W}_{j+1}^{(1)}\left[\boldsymbol{p}_j, \boldsymbol{c}_{j+2}\right] + \boldsymbol{b}_{j+1}^{(1)}\right) \quad (j = 1, 2, \ldots) \tag{3.2}$$

として計算する。ここで、[　] は 2 つのベクトルを結合したベクトルを表し、$[\boldsymbol{c}_1, \boldsymbol{c}_2], [\boldsymbol{p}_j, \boldsymbol{c}_{j+2}] \in \mathbb{R}^{M+M}$ である。$\boldsymbol{W}_j^{(1)}$ は $M \times 2M$ の符号化重み行列、$\boldsymbol{b}_j^{(1)}$ は M 次元のバイアスベクトルである。活性化関数 f には双曲線正接関数 tanh 関数を用いる。

図 3.7 に示すように、RAE の出力層は意味ベクトルを分類する多クラス分類器である。この多クラス分類器は意味ベクトルに応じて、K 個のクラスそれぞれについて値を出力する。K 次元ベクトル $\boldsymbol{u}$ は、

$$\boldsymbol{u} = \boldsymbol{W}^{(\mathrm{label})}\boldsymbol{p} + \boldsymbol{b}^{(\mathrm{label})} \tag{3.3}$$

である。ここで、$\boldsymbol{W}^{(\mathrm{label})}$ は $K \times M$ の重み行列、$\boldsymbol{b}^{(\mathrm{label})}$ は K 次元のバイアスベクトルである。そして、$\boldsymbol{u}$ を入力とする softmax 関数により、意味ベクトル $\boldsymbol{p}$ に対する多クラス分類器の出力は

$$\begin{aligned}\boldsymbol{d}(\boldsymbol{p}) &= \mathrm{softmax}(\boldsymbol{u}) \\ &= \mathrm{softmax}\left(\boldsymbol{W}^{(\mathrm{label})}\boldsymbol{p} + \boldsymbol{b}^{(\mathrm{label})}\right)\end{aligned} \tag{3.4}$$

とする。そして、K 次元ベクトル $\boldsymbol{d}(\boldsymbol{p})$ の i の成分 $d_i(\boldsymbol{p})$ は

$$\sum_{i=1}^{K} d_i(\boldsymbol{p}) = 1 \tag{3.5}$$

になるように softmax 関数を決める [38]。したがって、$d_i(\boldsymbol{p})$ は、意味ベクトル $\boldsymbol{p}$ が i 番目のクラスに所属する確率を表すと解釈できる。本手法では式 (3.5) の条件を満たすように

$$d_i(\boldsymbol{p}) = \frac{\exp(u_i)}{\sum_{j=1}^{K}\exp(u_j)} \tag{3.6}$$

とする [38]。

RAE の各節点において適切な意味ベクトルを合成する学習の仕組みを図 3.8 に示す。意味ベクトル $\boldsymbol{p}$ に対して、復号化重み行列 $\boldsymbol{W}^{(2)}$ とバイアスベクトル $\boldsymbol{b}^{(2)}$、活性化関数 f により、単語ベクトル $\boldsymbol{c}_1$,$\boldsymbol{c}_2$ の再現ベクトル $\boldsymbol{c}_1^{'}$, $\boldsymbol{c}_2^{'}$ は

$$\left[\boldsymbol{c}_1^{'}, \boldsymbol{c}_2^{'}\right] = f\left(\boldsymbol{W}^{(2)}\boldsymbol{p} + \boldsymbol{b}^{(2)}\right) \tag{3.7}$$

と計算する。式 (3.7) で、復号化重み行列 $\boldsymbol{W}^{(2)}$ は $2M \times M$ の行列であり、バイアスベクトル $\boldsymbol{b}^{(2)}$ は $2M$ 次元のベクトルである。図 3.8 で示した $E_{\mathrm{rec}}(\boldsymbol{c}_1)$ と $E_{\mathrm{rec}}(\boldsymbol{c}_2)$ では節点の入力ベクトル $\boldsymbol{c}_1$,$\boldsymbol{c}_2$ と再現ベクトル $\boldsymbol{c}_1^{'}$,$\boldsymbol{c}_2^{'}$ 間の再現誤差であり、それぞれ

$$E_{\mathrm{rec}}(\boldsymbol{c}_1) = \frac{1}{2}\left|\left|\boldsymbol{c}_1 - \boldsymbol{c}_1^{'}\right|\right|^2 \tag{3.8}$$

$$E_{\mathrm{rec}}(\boldsymbol{c}_2) = \frac{1}{2}\left|\left|\boldsymbol{c}_2 - \boldsymbol{c}_2^{'}\right|\right|^2 \tag{3.9}$$

とする。ただし、$||\quad||$ はユークリッドノルムである。図 3.8 の中に示した $E_{\mathrm{cE}}(\boldsymbol{p}, \boldsymbol{u})$ は合成した意味ベクトル $\boldsymbol{p}$ と多クラス分類器の出力との間の交差エントロピー誤差である。$E_{\mathrm{cE}}(\boldsymbol{p}, \boldsymbol{u})$ は、学習テキストを各固有表現に分類する目標分布 $\boldsymbol{u}$ を用いて、

$$E_{\mathrm{cE}}(\boldsymbol{p}, \boldsymbol{u}) = -\sum_{j=1}^{K} u_j \log d_j(\boldsymbol{p}) \tag{3.10}$$

とする。

式 (3.8) と式 (3.9) により計算した各節点の再現誤差と式 (3.10) を計算した交差エントロピ誤差を合成した評価関数 J が最適になるように、教師データを使って RAE 全体のパラメータを学習により設定する。入力単語列 c_i と目標分類 t に対する RAE 全体のコスト J を

$$J = \frac{1}{M}\sum_{(\boldsymbol{c}_i, \boldsymbol{u})}\left[\alpha\left\{E_{\mathrm{rec}}(\boldsymbol{c}_1) + E_{\mathrm{rec}}(\boldsymbol{c}_2)\right\} + (1-\alpha)E_{\mathrm{cE}}(\boldsymbol{p}, \boldsymbol{u})\right] + \frac{\lambda}{2}||\boldsymbol{\theta}||^2 \tag{3.11}$$

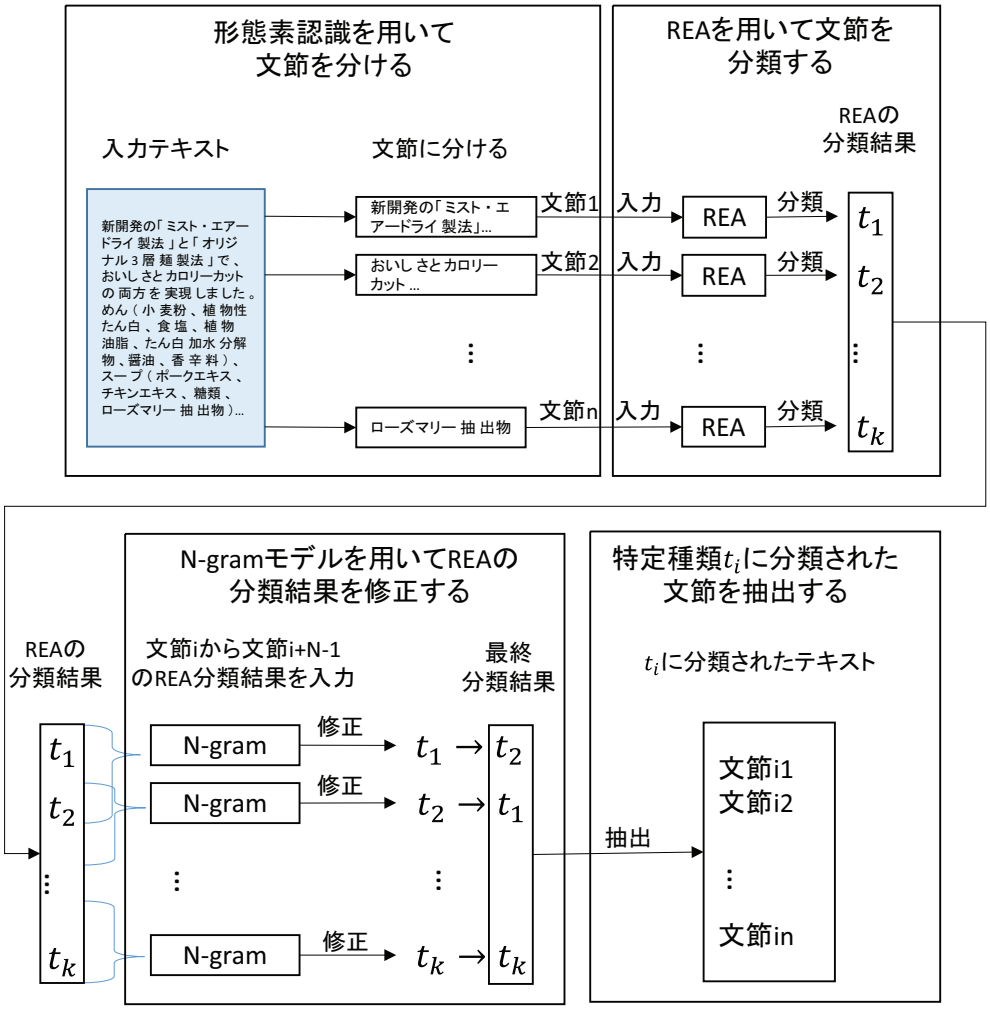


図 3.6　RAE と N-gram モデルを用いた固有表現の抽出

とする。ここで、α は誤差の重みパラメータ、$\boldsymbol{\theta}$ は本モデル中のパラメータであり、$\boldsymbol{\theta} = \left[\boldsymbol{W}^{(1)}, \boldsymbol{b}^{(1)}, \boldsymbol{W}^{(2)}, \boldsymbol{b}^{(2)}, \boldsymbol{W}^{(\mathrm{label})}, \boldsymbol{b}^{(\mathrm{label})}\right]$、$\lambda$ は $\boldsymbol{\theta}$ の正則化を制御するパラメータである。RAE の学習手順を図 3.9 にまとめる。RAE を学習する時

$$\frac{\partial J}{\partial \boldsymbol{\theta}} = \frac{1}{M} \nabla J + \lambda \boldsymbol{\theta} \tag{3.12}$$

により各パラメータに対するコスト J の勾配を計算し、勾配法により、各節点を学習させる。

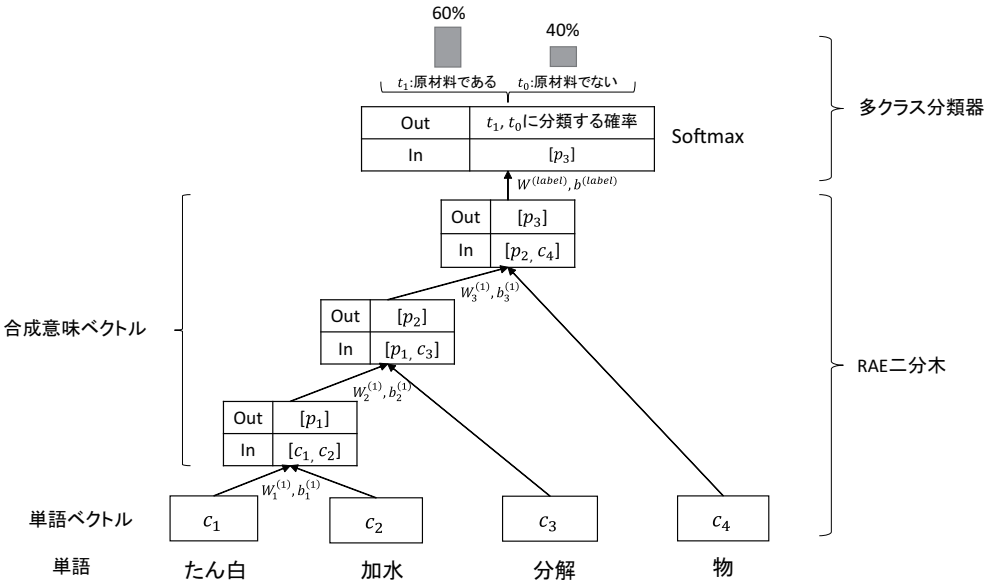


図 3.7　再帰自己符号化器の構造

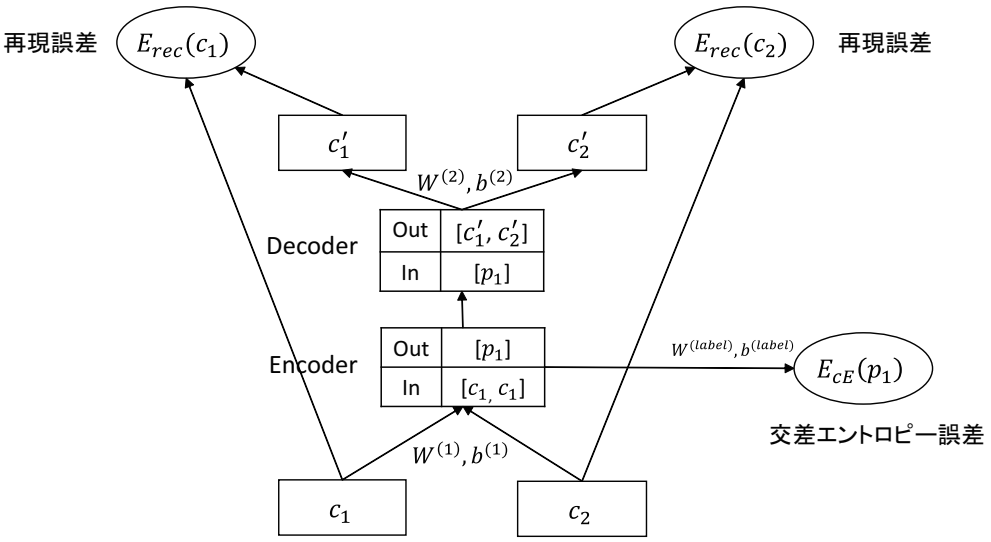


図 3.8　RAE 各節点の構造。この図では最下層の RAE を代表して示した。それ以外の層では、Encoder の入力を p_j, c_{j+2}、出力を p_{j+1} に、Decoder の出力 $p_j^{'}, c_{j+2}^{'}$ を置き換える。また、重み行列、バイアスベクトル、復号化重み行列も合わせて置き換える。

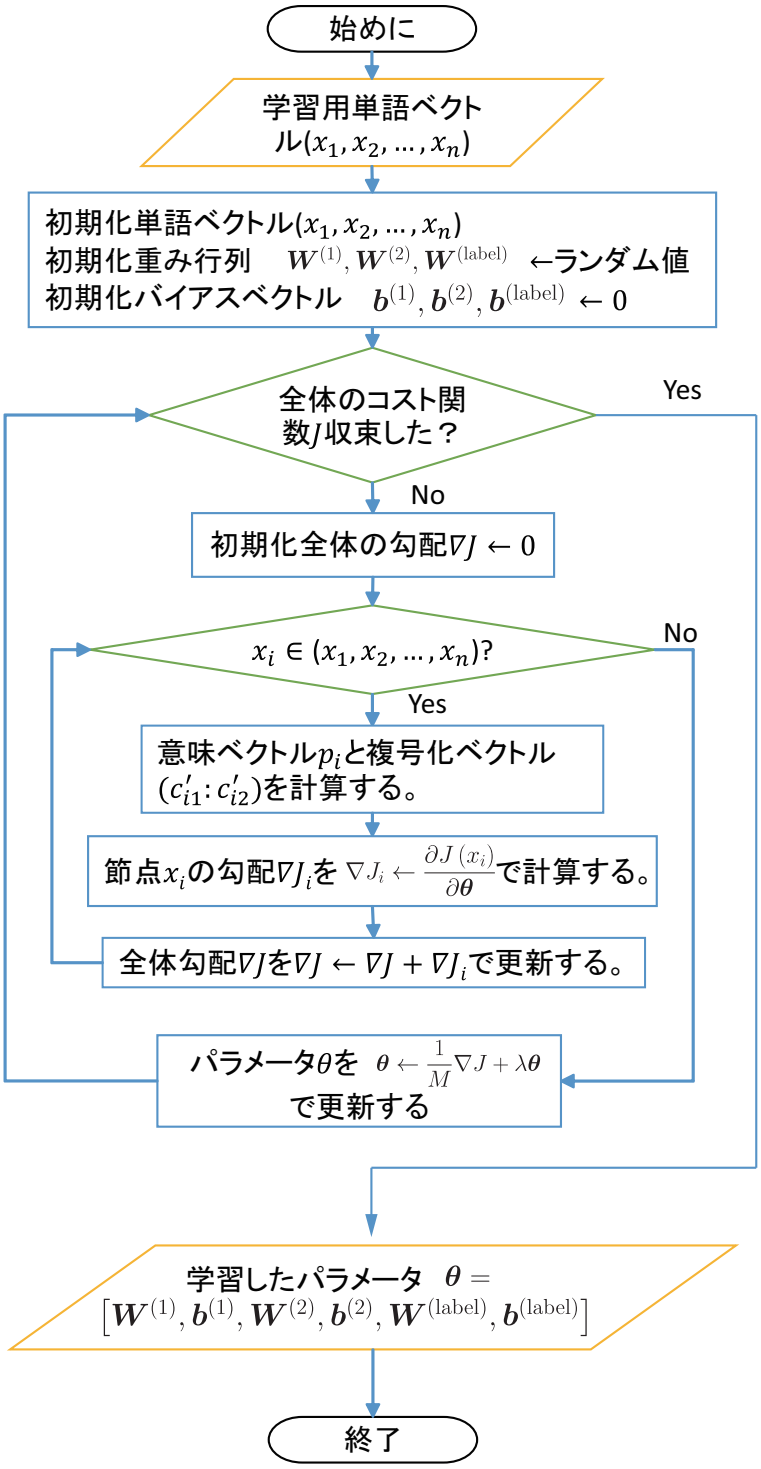


図 3.9　RAE の学習手順

3.2.2　N-gram モデルを用いた文節の固有表現推定

3.2.1 節で述べた RAE は、入力した文節の意味ベクトルを合成し、その文節の固有表現の分布を計算した。本手法ではさらに文節前後のつながりの関係を解析して、文節をより適切な固有表現に分類する。この文節のつながりの解析では、RAE の固有表現の解析結果に対して確率的言語モデルである N-gram を使って対象の文節とその前後の文節の表現を解析し、その解析結果に応じて RAE の解析結果を更新する。

N-gram モデルは、単語列 $(w_1, w_2, w_3, \ldots, w_n)$ における i 番目の単語 w_i の生起確率 $P(w_i)$ が w_i の直前の $(N-1)$ 個の単語に依存するモデルである。本手法では、この N-gram モデルを使い、対象の文節の分類を、その文節の前後の文節とのつながりの中で評価する。本手法では、前後の文節とのつながりを評価するために、文節コスト $C_p(t_s)$ を導入する。

文節コスト $C_p(t_s)$ は、N-gram モデルに基づく分類コスト $C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ から構成される。解析対象の段落を構成する文節列 $(\boldsymbol{h}_1, \boldsymbol{h}_2, \boldsymbol{h}_3, \ldots, \boldsymbol{h}_n)$ において、i 番目の文節 $\boldsymbol{h}_i$ の固有表現を t_i とする。また、その対象文節の $N-1$ 個前までの文節の固有表現の分類列 $\left(t_{i-(N-1)}, \ldots, t_{i-1}\right)$ を $\boldsymbol{t}_i^{'}$ とする。本手法では、N-gram モデルを用いた形態素解析のコスト最小法 [41] で使われている分類コスト $C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ を用いる。各固有表現の組み合わせのコストを表す分類コスト $C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ を

$$C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right) = -\log\left(\frac{P\left(t_i \mid \boldsymbol{t}_i^{'}\right)}{P(t_i)}\right) \tag{3.13}$$

とする。ここで、$P(t_i)$ はテギスト中に固有表現 t_i が出現する確率、$P\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ は固有表現の分類列 $\boldsymbol{t}_i^{'}$ の次に固有表現 t_i が出現する確率である。対象の文節がその前の文節とつながりが強い

程、分類コスト $C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ は小さくなる。

文節の固有表現を、対象文節の前だけでなく、対象文節の後の文節まで考慮するために、分類コスト $C_c\left(t_i \mid \boldsymbol{t}_i^{'}\right)$ を用いた文節コスト $C_p\left(t_s\right)$ を導入する。文節コスト $C_p\left(t_s\right)$ は、s 番目の文節 c_s の分類コスト $C_c\left(t_s \mid \boldsymbol{t}_s^{'}\right)$ と、c_s の後の $N-1$ 個分類列の分類コスト $C_c\left(t_{s+1} \mid \boldsymbol{t}_{s+1}^{'}\right), \ldots, C_c\left(t_{s+N-1} \mid \boldsymbol{t}_{s+N-1}^{'}\right)$ を用いて計算する。つまり、文節コスト $C_p\left(t_s\right)$ は、

$$C_p\left(t_s\right) = \sum_{i=0}^{N-2} C_c\left(t_{s+1+i} \mid t_{s+1+i}^{'}\right) \tag{3.14}$$

である。ここで、$s+i$ 番目の文節が存在しないとき、$C_c\left(t_{s+i} \mid \boldsymbol{t}_{s+i}^{'}\right) = 0$ にする。文節コストを計算する例を図 3.10 で示す。同図では、$N=3$ として、二つの分類 (0 と 1) に分類する 7 個の文節を持つ文の 4 番目の文節の文節コストを計算している。$N=3$ のとき、4 番目の文節の文節コストは、

$$\begin{aligned} C_p\left(t_4\right) &= C_c\left(t_4 \mid t_2, t_3\right) + C_c\left(t_5 \mid t_3, t_4\right) + C_c\left(t_6 \mid t_4, t_5\right) \\ &= C_c\left(1 \mid 00\right) + C_c\left(1 \mid 01\right) + C_c\left(1 \mid 11\right) \end{aligned} \tag{3.15}$$

である。

文節コスト $C_p\left(t_s\right)$ は、学習用サンプルテキストを使ってあらかじめ設定した閾値 T_t との比較により、分類結果を更新する。その閾値 T_t は、固有表現 t ごとに個別に閾値 T_t を設定する。文節の固有表現を推定するとき、文節の文節コスト $C_p\left(t_s\right)$ が閾値 T_t より高いとき、RAE で計算した固有表現の分類確率 $d_i\left(\boldsymbol{p}\right)$ において次に高い値を示す固有表現に文節の固有表現を修正する。

学習用サンプルテキストの固有表現の分類列を用いて、各固有表現 t の出現する確率 $P\left(t\right)$ と

番号i	文節	分類結果t_i
1	しなや かでコシ とつるみ の ある 麺 に 、	
2	カニ 風味 かまぼこ を 増量	0
3	油揚げ めん	0
4	植物 油脂 、	1
5	食 塩 、	1
6	たん白 加水 分解 物 、	1
7	醤油 、	1

t'_4　t'_5　t'_6

$\mathrm{C_c}(t_4|t'_4)$　$\mathrm{C_c}(t_5|t'_5)$　$\mathrm{C_c}(t_6|t'_6)$　$\mathrm{C_p}(t_4)$

$$\begin{aligned}\mathrm{C_p}(t_4) &= \mathrm{C_c}(t_4|t'_4) + \mathrm{C_c}(t_5|t'_5) + \mathrm{C_c}(t_6|t'_6)\\ &= \mathrm{C_c}(1|00) + \mathrm{C_c}(1|01) + \mathrm{C_c}(1|11)\end{aligned}$$

図 3.10　文節コストを計算する例。この例では、$N = 3$ としている。

$P\left(t \mid \boldsymbol{t}'\right)$ を統計解析し、分類コストを計算し、文節コストの閾値 T_t を設定する。閾値 T_t の学習手順を図 3.11 でまとめる。ここで、目標閾値 $\widehat{T_t}$ は、本手法を用いて学習用サンプルデータを各固有表現に分類する確率が最大のときの値である。閾値 T_t を探す範囲は、0 から学習用サンプルテキストの最大文節コスト $MAX\left[C_p\left(t_s\right)\right]$ までである。

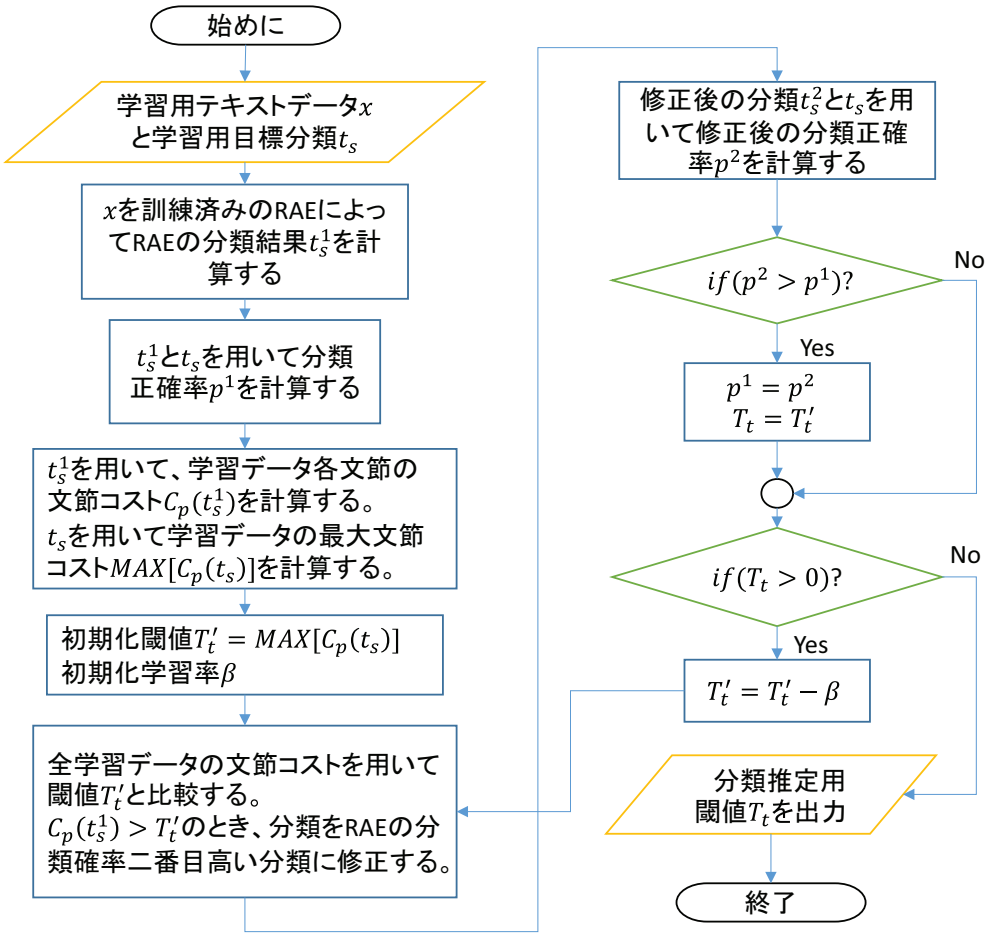


図 3.11　N-gram モデルの学習手順

表 3.1　学習とテストデータの例。t_1 は「食品原材料」、t_2 は「食品原材料ではない」。

テキストデータ	入力文節	目標分布 t_1	目標分布 t_2
カップヌードルシーフードヌードルのビッグサイズ：植物油脂、たん白加水分解物。	カップ ヌードルシーフード ヌードル のビッグサイズ	0.0	1.0
	植物 油脂	1.0	0.0
	たん白 加水 分解 物	1.0	0.0

3.2.3　評価実験

学習用サンプルテキストと性能評価用テキストは、インターネットに公開されている食品の関連ウェブサイトから収集した。各文節の目標とする分類は、テキストデータを形態素認識し、文節に分けた後、著者が手動で割り付けた。学習用サンプルテキストと性能評価用テキストの例を表 3.1 に示す。ここで、テキストデータの形態素認識には国立国語研究所が開発した電子化辞書 UniDic [42] を用いた。電子化辞書 UniDic は、日本語テキストに形態素認識した情報を付与する。同表において目標分類は「食品原材料」と「食品原材料ではない」の二つであり、「食品原材料」は t_1、「食品原材料ではない」は t_2 として表記している。学習用サンプルテキストには 5000 文節があるデータを収集し、性能評価用テキストでは 4000 文節があるデータを収集した。収集した文節は最大 30 個から最小 2 個の形態素が含まれていた。テキストデータから単語ベクトルを生成するには Google が提供する Word2vec[43] を用い、生成した単語ベクトルは 100 次元のベクトルとした。

表 3.2　N-gram モデルの閾値 T_t

	T_{t_1}	T_{t_2}
2-gram	6.9	3.5
3-gram	21.2	8.9
4-gram	11.9	11.5

テキスト情報抽出手法を、以下に示す 4 つの条件について比較した。

1. 参照手法：RAE のみ
2. 手法 A：RAE+2-gram
3. 手法 B：RAE+3-gram
4. 手法 C：RAE+4-gram

ここで、N-gram モデルの N の値が大きいと N が解析するテキストの文節数より大きい恐れがある。そのため、本論文で検証する N-gram モデルは 4-gram までとした。各手法では、同じ RAE を用いた。RAE の学習におけるパラメータの中で、学習回数は 500 とし、式 (3.11) の α は Socher et. の検証 [44] を踏まえて α=0.2 とした。N-gram モデルの閾値 T_t を図 3.11 に示す学習における学習率 β は 0.1 である。学習した N-gram モデルの閾値 T_t を表 3.2 に示す。

性能評価用テキストが 500 文節、1000 文節、4000 文節の場合について、提案手法の性能を検討する。本論文で提案した手法の 3 つの条件と RAE のみの手法を、精度、再現率、F 値の 3 つの指標について比較する。精度 (適合率, precision) とは、固有表現抽出法が「正」と予測したデータのうち、実際に「正」であるデータの割合である。再現率 (recall) とは、実際に「正」であるデー

表 3.3　各手法の抽出精度の比較

	500 文節の精度	1000 文節の精度	4000 文節の精度
RAE	93.0%	79.6%	77.5%
A: RAE+2-gram	93.6%	84.3%	86.2%
B: RAE+3-gram	93.5%	84.7%	81.8%
C: RAE+4-gram	93.8%	82.3%	85.7%

表 3.4　各手法の再現率の比較

	500 文節の再現率	1000 文節の再現率	4000 文節の再現率
RAE	81.3%	80.1%	80.4%
A: RAE+2-gram	78.3%	77.4%	78.3%
B: RAE+3-gram	88.2%	91.5%	91.2%
C: RAE+4-gram	95.6%	96.7%	95.6%

表 3.5　各手法の F 値の比較

	500 文節の F 値	1000 文節の F 値	4000 文節の F 値
RAE	86.9%	80.1%	78.9%
A: RAE+2-gram	85.4%	81.0%	82.1%
B: RAE+3-gram	90.8%	87.7%	86.2%
C: RAE+4-gram	94.3%	89.7%	90.4%

タのうち，固有表現抽出法が「正」であると予測したデータの割合である。F 値 (F-measure) とは、精度と再現率の調和平均である。テキスト情報抽出の精度・再現率・F 値をそれぞれ表 3.3、3.4、3.5 に示す。同表では、3.1 節で述べた提案手法 A,B,C の中で、各指標のスコアが最も高い結果を下線で示している。

表 3.3 の抽出精度の比較結果では、提案手法 A、B、C の抽出精度は参照手法により上回って

いる。各手法の抽出精度は解析文節数の増加によって変化が大きい。500 文節から 1000 文節では、各手法の抽出精度は全て減っている。1000 文節から 4000 文節では、参照手法と手法 B の抽出精度は減少しているが、手法 A と手法 C の抽出精度は増加している。解析文節数が 500 の少ない文節数のとき、提案手法を参照手法に対して精度の改善は 1% 未満である。しかし、解析文節数が増えると提案手法の改善効果も上がっている。1000 文節のとき、手法 B の精度は参照手法より 5.1% 上回っている。4000 文節のとき、手法 A の精度は参照手法より 8.7% 上回って、手法 C は参照手法より 8.2 %上回っている。

表 3.4 の再現率の比較結果では、手法 B と手法 C の再現率は参照手法により上回っている。手法 A の再現率は参照手法より下回っている。各手法の再現率は解析文節数の増加によって変化が少ない。手法 C は参照手法により 15% 上回って、最も高い再現率を取得している。

表 3.5 の F 値の比較結果では、解析文節数は 500 とき手法 A の F 値は参照手法の F 値より下回っている。解析文節数は 1000 と 4000 のとき手法 A の F 値は参照手法の F 値より上回っている。手法 B と手法 C の F 値は参照手法の F 値より上回っている。解析文節数の増加によって参照手法、手法 A,B の F 値減少している。手法 C の F 値は、1000 文節のときは 500 文節より低いが、4000 文節のときは 1000 文節のより高い。手法 C は比較手法中、最も高い F 値を取得している。

表 3.3、3.4、3.5 を示した実験結果は提案手法の性能は参照手法より高かった。使用した 2-gram、3-gram、4-gram モデルには 4-gram モデルの再現率と F 値が高い。特に手法 C では 85.7% の抽出精度、95.6% の再現率、90.4% の F 値を取得した。以上のことから、N-gram モデ

ルを用いて RAE の解析結果を正しい結果に修正できることが分かる。提案手法は任意単語長のテキストの情報抽出に高い性能を示した。

3.3　CFP 計算用のデータを収集する方法

本手法で収集する CFP 計算用データは、製品の原材料調達段階のデータ、輸送段階のデータ、製品生産段階のデータである。原材料調達段階のデータは一般のウェブサイトにおいて公開されているテキストデータから収集し、輸送段階と製品生産段階のデータは各研究機構、企業を公開されているオープンデータベースから収集する。

一般のウェブサイトから原材料調達段階のデータを収集する方法のフローチャートを図 3.12 に示す。本手法は先ず、ウェブクローラーを用いて、CFP を計算する製品の製品名を持つウェブサイトを検索し、このウェブサイトのテキストデータを抽出する。そして、抽出されたテキストデータを形態素解析し、文節に分ける。最後に、分けた文節を固有表現抽出手法を用いて、原材料調達段階のデータを抽出する。

ウェブクローラーは、ウェブ上の文書や画像などを取得し、自動的にデータベース化するプログラムである。本研究では、JAVA を用いて、マルチスレッドのウェブクローラーを作成した。このウェブクローラーを用いてウェブサイトを検索し、テキストデータを抽出する時のフローチャートを図 3.13 に示す。図 3.13 に示した検索待ち行列とは、このウェブクローラーを検索するウェブサイトの URL の行列である。データを収集し易くするため、この検索待ち行列の初期値はあらかじめ、製品を販売するウェブサイト、製品の生産メーカー等の特定のウェブサイトからに設定する。このウェブクローラーで収集したテキストデータは図 3.14 に示すテキストファイルとして出力される。このウェブクローラーはマルチスレッドであるため、最大で 30 個のスレッドで同時にウェブサイトを検索する。

抽出されたテキストデータにおける形態素解析には、3.2 節で説明した固有表現抽出手法と同じく、国立国語研究所が開発した電子化辞書 UniDic[42] を用いた。形態素解析した結果は、図 3.15 に示すように、一行が一文節になっている。

原材料調達段階として収集したデータ例を図 3.16 に示す。ウェブサイトから収集された原材料調達段階のデータには、製品の原材料の他に、各ウェブサイトに記録している製品の原材料が異なることがあるため、各原材料を記載されたウェブサイトの数を記録する。また、各原材料と製品の重量の比には、収集されたウェブサイトによってばらつきがあるため、この重量比の最大値、最小値、平均値を計算する。収集した原材料調達段階のデータの例を図 3.17 に示す。

オープンデータベースでは、データが体系的に整理されたため、収集したデータを比較的容易に CFP の計算に使用できる。本手法でオープンデータベースから収集する輸送データは、その原材料の日本国内の産地名、産量と海外の輸入先、輸入量である。収集する先は農林水産省 [29] と経済産業省 [45] が公開している生産データベースである。この生産データベースの例を図 3.18 に示す。この生産データベースでは、原材料の日本国内の産地名、産量と海外の輸入先、輸入量を記載している。現在本研究では 6962 品の原材料の産地データを収集した。原材料の日本国内の産地名と海外の輸入先を収集することにより、この原材料を産地から製品の生産工場間の輸送距離を求められる。日本国内の産地から日本国内工場の輸送距離を Google Maps を用いて収集する。海外から輸入され原材料の輸送距離の収集には、海上保安庁の「距離法」[27] のデータを基づいた製作した「海外地域間距離データベース」を用いる。

本手法で収集する生産データは、各生産メーカーが公開している工場の CO_2 排出量と工場所

在地である。現在、多くの生産メーカーが、工場の CO_2 排出量等の環境データと工場所在地を公開している。例えば、日清食品では、図 3.19 に示すの環境データを公開してる [46]。このようなデータを用いて、製品生産時の CO_2 排出量を収集できる。

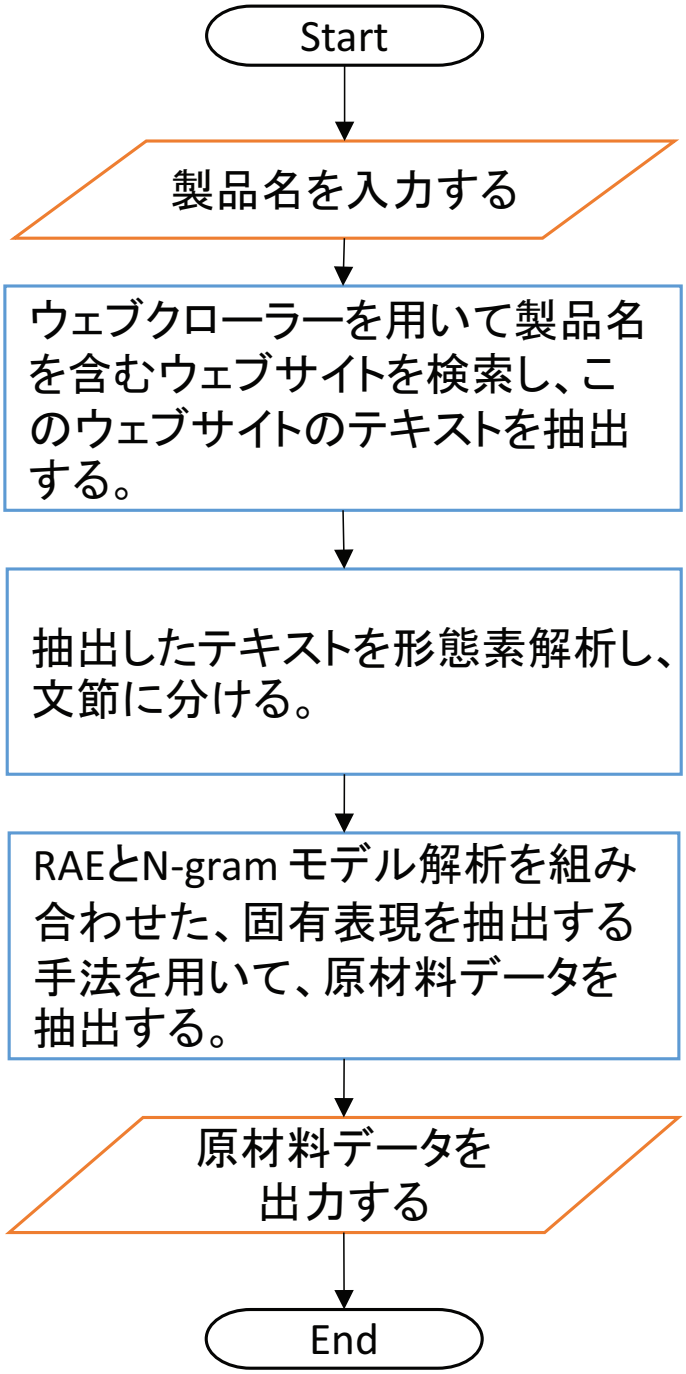


図 3.12　一般のウェブサイトから原材料調達段階のデータを収集する方法のフローチャート

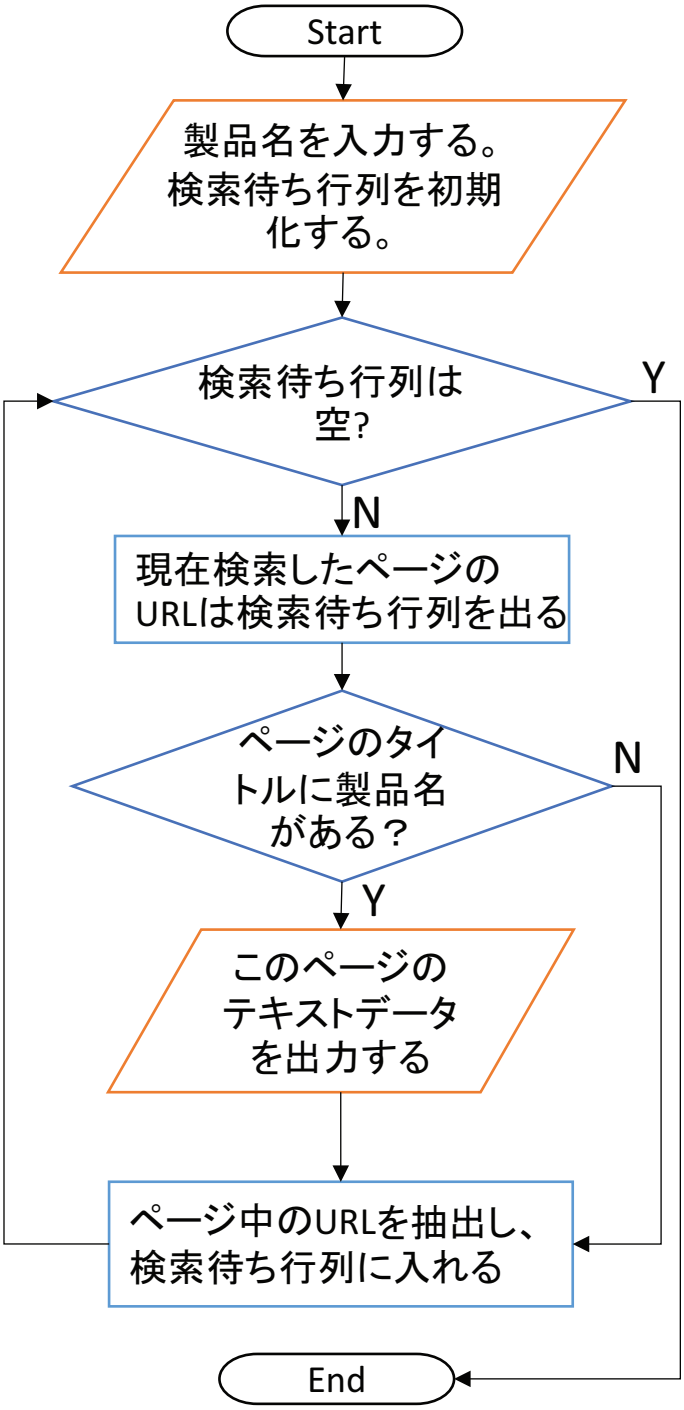


図 3.13　ウェブクローラーを用いてウェブサイトを検索するフローチャート

[ブー！バッ！ポン！　ムエタイ選手も歓喜！2016年注目の
...タイの伝統料理、蟹と卵の甘辛カレー「ブーバッポン」
がカップヌードルビッグで登場！　詳しくはこちら。
http://nissin.eng.mg/980a1, 熱い夏！熱く食べ！！熱く
応援！！！世界的ビッグゲームに挑む錦織選手を黄金カレ
ーメニューを食べながら黄金のメガホンで応援しよう！
世界を掴め！錦織圭選手応援黄金カレートリオ、...7月4日
に3品同時発売。FlyingGetで黄金カレートリオを「数量限
定　黄金の錦織応援メガホンパッケージ」で発売前にお届
け！　6月20日(月)11:00より特別先行販売スタート。先着
順1,000セット限定となりますので、お早めに！　詳しくは
こちら。http://nissin.eng.mg/1ee50, いっ、入れすぎ！
入れすぎ！, よーしよしよし！よーしよし、「しお」！
ムツゴロウさんもお気に入り。「カップヌードルしお」
が...Wチキンになって具材充実リニューアル！　詳しくは
こちら。http://nissin.eng.mg/fa5f0, 注文して、3分で来
た！, #ドーナツ#スイーツ#今日のおやつ#自分へのごほう
び#念願の#行列並んだ#かわいすぎ#あまいのだいすき,

図 3.14　ウェブクローラーで収集したテキストデータ

具材 に は ホクホク の ポテト 、
肉 、
にんじん 、
ね ぎ が 入 って ます 。
油揚げ めん （ 小 麦粉 、
植 物 油脂 、
食 塩 、
糖類 、
醤油 、

図 3.15　形態素解析した結果の例

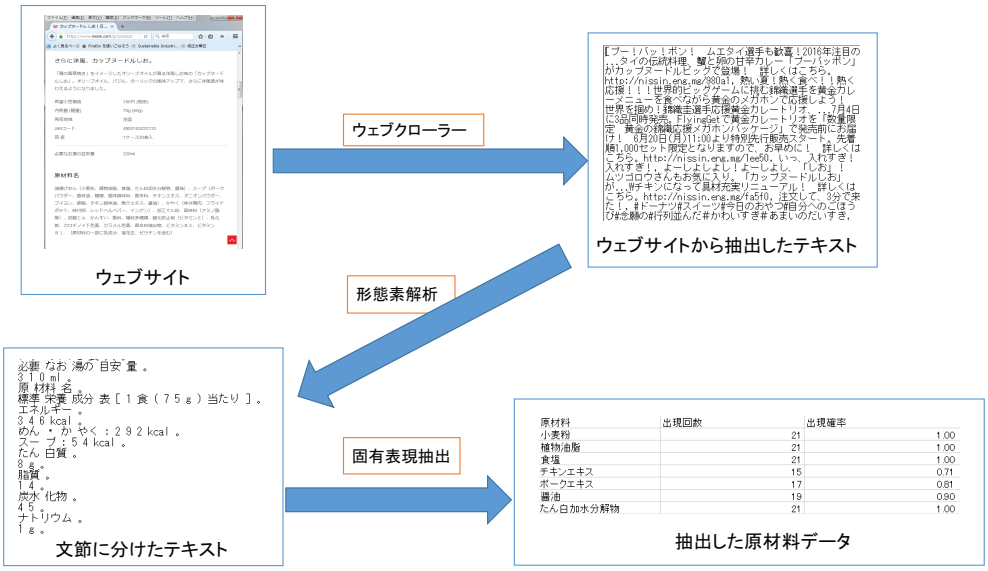


図 3.16　原材料調達段階として収集したデータの例

原材料	出現回数	出現確率	最小重量比	最大重量	平均
小麦粉	21	1.00	0.564	0.637	0.599
植物油脂	21	1.00	0.072	0.248	0.186
食塩	21	1.00	0.043	0.077	0.059

図 3.17　原材料調達段階のデータの例

（1）　平成26年産豆類（乾燥子実）ア

ア　大豆

全国農業地域＋都道府県	作付面積	10 a 当たり収量	収穫量
	(1)	(2)	(3)
	ha	kg	t
九州	21,500	168	36,100
沖縄	1	38	1
（都道府県）			
北海道	28,600	257	73,600
青森	4,040	133	5,370
岩手	4,020	136	5,470
宮城	10,000	193	19,300
秋田	7,300	132	9,640
山形	4,980	155	7,720
福島	1,710	132	2,250

図 3.18　農林水産省が公開しているオープンデータベースの例

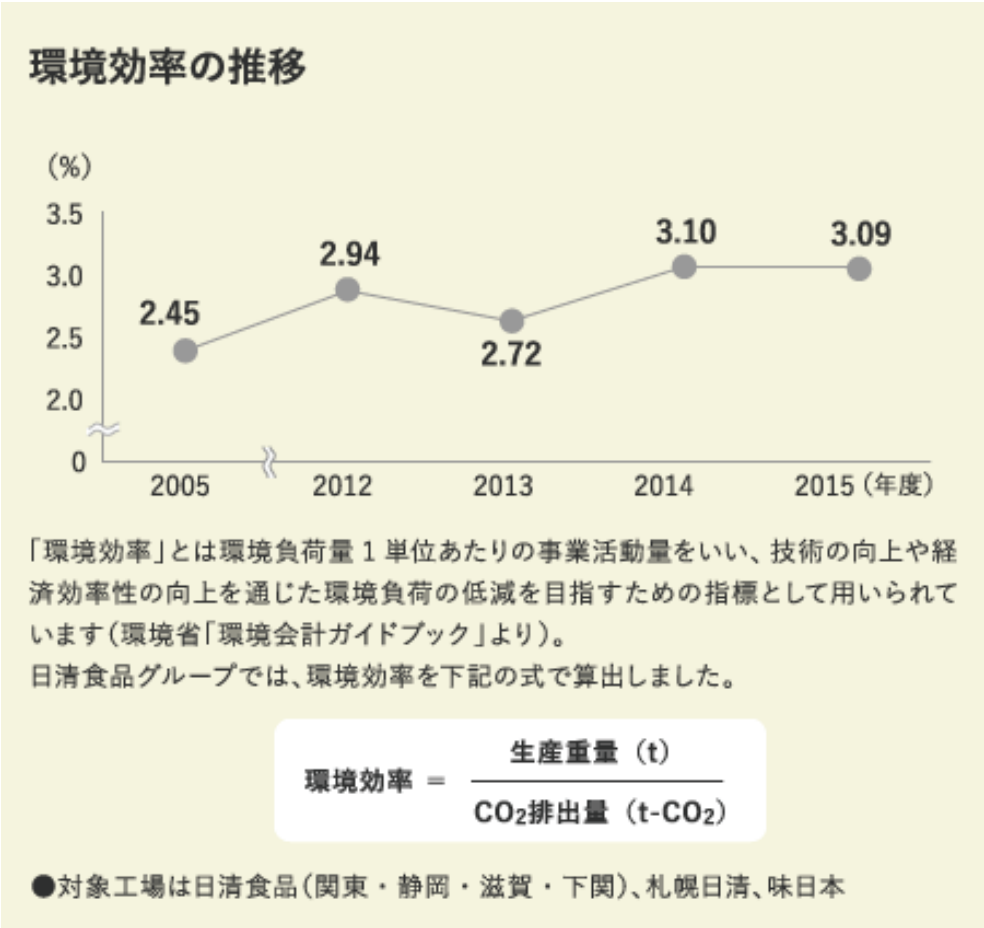


図 3.19　日清食品が公開した環境データ

3.4　計算方法

本手法でウェブから収集したデータには、以下の性質がある。

1. 商品の原材料データにばらつきがある。本手法では、一つ商品種の原材料データを複数のウェブサイトから収集するため、同じ商品種でも個々の商品の重量、原材料の種類、その原材料が商品に占める重量比が異なる。
2. 原材料の真の産地情報は不明である。製品の情報を記載されているウェブサイトでは、その原材料の産地を記載されていないことがある。例えば、図 3.2 に示すアマゾンの食品販売のウェブサイトでは、製品の原材料種類、製品重量、製品の栄養情報のデータを記載されているが、各原材料の産地情報が記載されていない。このようなウェブサイトから原材料の輸送段階のデータを収集することはできない。

この二つの性質のため、本章で述べる CFP の計算法と 2 章で述べた CFP の計算法には違いがある。この違いを表 3.6 にまとめた。この違いがあるため、本論文では、原材料調達段階の CFP と輸送段階の CFP を計算する新たな方法を提案する。

原材料調達段階の CFP の計算では、原材料データのばらつきに対して、各原材料の出現頻度比を提案する。原材料 i の出現頻度比 P_i は、

$$P_i = \frac{M_i}{N} \tag{3.16}$$

とする。ただし、N をウェブクローラーを収集したウェブサイトの総数とし、M_i を原材料 i を記載されたウェブサイトの数とする。本手法で収集される製品の重量、使用する原材料の種類、

表 3.6　本 CFP 計算手法と従来の CFP の計算の区別

	計算する目標	原材料調達段階	輸送段階	生産段階
従来の CFP の計算	一つの製品	製品の重量、使用する原材料の種類、原材料の重量は固定されている	輸送距離が明確である	生産メーターから提供されている
本手法	複数の製品を含める商品種	製品の重量、使用する原材料の種類、原材料の重量は収集するサイトによって変わる	産地情報は不明であるため、輸送距離は不明	生産メーターの公開データから収集されている

原材料の重量は収集するサイトによってばらつく。本手法では、製品の重量を 100g に固定する。そして、原材料の重量比を用いて、原材料の重量を求める。原材料の重量比では、原材料 i を記載されたサイトの中、原材料の重量とそのサイトで記載された製品の重量の比である。この重量比の最大値 R_{i_MAX}、最小値 R_{i_MIN}、平均値 $\overline{R_i}$ を図 3.17 に示す原材料調達段階のデータの例のように、原材料調達段階のデータを収集する時、計算されている。原材料調達段階の CFP の計算で収集したデータを以下にまとめた。

1. 原材料名
2. 収集したウエブサイトの総数 N
3. 原材料 i を記載したウエブサイトの数 M_i

4. 最大重量比 R_{i_MAX}

5. 最小重量比 R_{i_MIN}

6. 平均重量比 $\overline{R_i}$

7. CO_2 原単位

原材料 i の CFP の最大値 $\mathrm{CFP}_{\mathrm{Raw_}i_MAX}$、最小値 $\mathrm{CFP}_{\mathrm{Raw_}i_MIN}$、平均値 $\overline{\mathrm{CFP}_{\mathrm{Raw_}i}}$ をはそれぞれ式 (3.17)、式 (3.18)、式 (3.19) で計算する。CO_2 原単位では味の素グループ版「食品関連材料 CO_2 排出係数データベース」に基づいた数値である。

$$\mathrm{CFP}_{\mathrm{Raw_}i_MAX} = (R_{i_MAX} \times m \times P_i) \times (\mathrm{CO_2}\text{原単位})_i \tag{3.17}$$

$$\mathrm{CFP}_{\mathrm{Raw_}i_MIN} = (R_{i_MIN} \times m \times P_i) \times (\mathrm{CO_2}\text{原単位})_i \tag{3.18}$$

$$\overline{\mathrm{CFP}_{\mathrm{Raw_}i}} = \left(\overline{R_i} \times m \times P_i\right) \times (\mathrm{CO_2}\text{原単位})_i \tag{3.19}$$

原材料 i の真の産地情報は不明のため、原材料の輸送距離では、本手法で収集した生産データを用いて、原材料 i の各産地から生産メーカーの工場までの最長輸送距離 L_{i_MAX}、最短距離 L_{i_MIN}、平均距離 $\overline{L_i}$ である。この平均距離 $\overline{L_i}$ では、原材料 i を生産する各産地 k の産量 Y_k と日本国内総消費量 Y_{All} の比を、この産地からの輸送距離 L_{ik} に重み付け、式

$$\overline{L_i} = \sum_k \left(\frac{Y_k}{Y_{All}} \times L_{ik}\right) \tag{3.20}$$

によって計算された、重み付け平均距離である。なお、日本国内総消費量 Y_{All} は国内産量と海外輸入量の和である。輸送段階の CFP を計算を計算するため、収集したデータを以下にまとめた。

1. 原材料名
2. 原材料 i の国内消費量 Y_{All}
3. 各産地 k の産量 Y_k
4. 最大輸送距離 L_{i_MAX}
5. 最小輸送距離 L_{i_MIN}
6. 各産地 k からの輸送距離 L_{ik}
7. 輸送の CO_2 原単位

以上のデータを用いて、原材料 i の輸送段階の CFP をの最大値 $\mathrm{CFP}_{\mathrm{T}_i_MAX}$、最小値 $\mathrm{CFP}_{\mathrm{T}_i_MIN}$、平均値 $\overline{\mathrm{CFP}_{\mathrm{T}_i}}$ はそれぞれ式 (3.21)、式 (3.22)、式 (3.23) によって計算する。

$$\mathrm{CFP}_{\mathrm{T}_i_MAX} = (R_{i_MAX} \times m \times P_i) \times L_{i_MAX} \times (\text{輸送 } \mathrm{CO_2}\text{原単位})_i \tag{3.21}$$

$$\mathrm{CFP}_{\mathrm{T}_i_MIN} = (R_{i_MIN} \times m \times P_i) \times L_{i_MIN} \times (\text{輸送 } \mathrm{CO_2}\text{原単位})_i \tag{3.22}$$

$$\overline{\mathrm{CFP}_{\mathrm{T}_i}} = \left(\overline{R_i} \times m \times P_i\right) \times \overline{L_i} \times (\text{輸送 } \mathrm{CO_2}\text{原単位})_i \tag{3.23}$$

本手法で計算する商品種の CFP は、原材料調達段階と輸送段階の CFP の最大値、最小値、平均値を用いて、商品種の CFP の最大値 CFP_{MAX}、最小値 CFP_{MAX}、平均値 $\overline{\mathrm{CFP}}$ をそれぞれ式 (3.24)、式 (3.25)、式 (3.26) によって計算する。そこで、生産 CFP では生産メーカーから収集された、100g あたりの CO_2 排出量である。

$$\mathrm{CFP}_{MAX} = \sum_i (\mathrm{CFP}_{\mathrm{Raw_}i\mathrm{_}MAX} + \mathrm{CFP}_{\mathrm{T_}i\mathrm{_}MAX}) + \text{生産 CFP} \tag{3.24}$$

$$\mathrm{CFP}_{MIN} = \sum_i (\mathrm{CFP}_{\mathrm{Raw_}i\mathrm{_}MIN} + \mathrm{CFP}_{\mathrm{T_}i\mathrm{_}MIN}) + \text{生産 CFP} \tag{3.25}$$

$$\overline{\mathrm{CFP}} = \sum_i \left(\overline{\mathrm{CFP}_{\mathrm{Raw_}i}} + \overline{\mathrm{CFP}_{\mathrm{T_}i}}\right) + \text{生産 CFP} \tag{3.26}$$

3.5　性能評価

本計算法の性能を評価するために、本計算法で算出する商品種の CFP と第 2 章で述べた自動計算システムが算出する実際の製品の CFP を比較した。

評価実験のためにウェブから収集する商品種は日清食品で販売する「麺類」である。ウェブクローラーで検索する製品名は「ラーメン」「ヌードル」「うどん」「そば」とし、検索の初期 URL は日清食品グループの <https://www.nissin.com/jp/products/> とした。この検索条件により、100 個のウェブサイトから「麺類」商品の原材料を 108 種が収集し、3.4 節に述べた手法を使って CFP を計算した。計算結果を表 3.7 に示す。

「麺類」製品個々の実測データに基づいた CFP をコープさっぽろからデータを提供された麺類製品 14 品について計算した。実測データに基づいて計算した CFP を表 3.8 にまとめる。実測データの内、「うどん」製品は 6 種類、「ラーメン」製品は 5 種類、「そば」製品は 3 種類がある。同表にまとめた製品の 100g 当たりの平均 CFP は 144g-CO_2/100g である。

本計算方法の推測した麺類製品 CFP と実測データに基づく麺類製品の CFP の比較結果を図 3.20 にまとめる。同図の比較結果を示すように、現在収集した実際の麺類製品の CFP はウェブのデータから推測した CFP の最大値と最小値の範囲内にある。また、推測した麺類製品の平均 CFP の 155.2g-CO_2/100g と実際の麺類製品の平均 CFP の 144g-CO_2/100g は近い値にあり、推測平均 CFP と実際平均 CFP は 7% の差であった。この比較結果は、ウェブのデータから CFP を推定する手法からえる CFP の範囲が実際の CFP の値の推定に妥当であることを示している。

表 3.7　日清食品グループのウェブ <https://www.nissin.com/jp/products/> のデータから計算した麺類製品の CFP の計算結果

CFP 種類	最大 CFP ($g\text{-}CO_2/100g$)	最小 CFP ($g\text{-}CO_2/100g$)	平均 CFP ($g\text{-}CO_2/100g$)
原材料 CFP	172.4	58.2	108.8
輸送 CFP	17.4	5.4	14.2
生産 CFP	32.2	32.2	32.2
総 CFP	222.0	95.8	155.2

表 3.8　実際の麺類製品の CFP

製品番号	製品名	製品 CFP ($g\text{-}CO_2$/個)	製品重量 (g)	100g 当たり CFP ($g\text{-}CO_2/100g$)
1	3 食うどん	802	630	127
2	かけうどん 1 食	272	210	129
3	かけうどん 2 食	540	420	128
4	生ラーメン・塩味	529	240	220
5	生ラーメン・醤油	529	240	220
6	生ラーメン・味噌	514	240	214
7	焼うどん・しょうゆ味	585	420	139
8	焼きそば 1 食	209	168	124
9	焼きそば 2 食	458	336	136
10	焼きそば 3 食	682	504	135
11	生ラーメン 2 食	245	235	104
12	生ラーメン 3 食	344	350	98
13	北海道 100 うどん	245	235	104
14	北海道小麦冷凍うどん	350	254	138

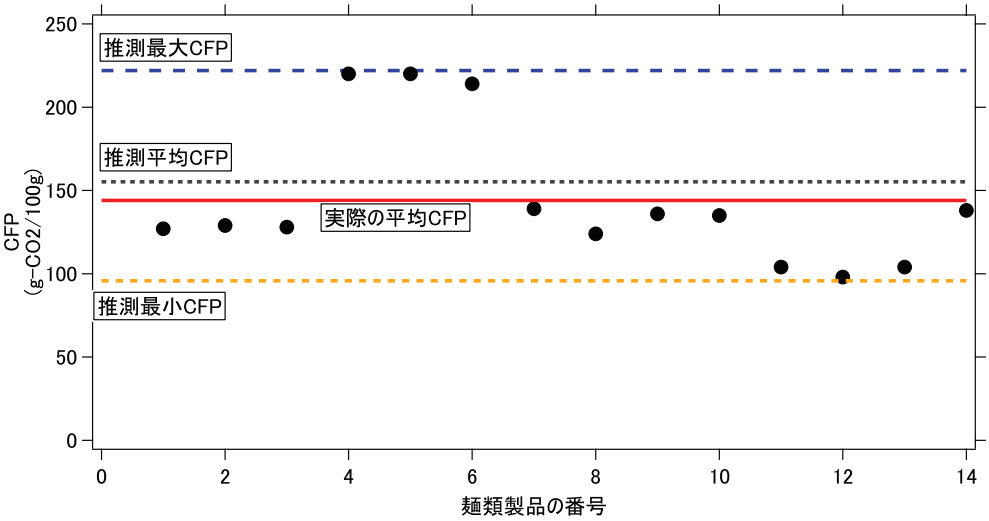


図 3.20　推測した麺類製品 CFP と実際の麺類製品の CFP の比較結果

3.6 まとめ

本章では、テキストマイニングに基づいたカーボンフットプリントの計算方法を述べた。この計算方法では製品の生産メーカーに直接頼らずに、製品の CFP を推測できる。

この計算方法を実現するため、本論文では再帰自己符号化器と N-gram モデルを用いて原材料情報抽出ための機械学習方法と、商品種の CFP を推測する方法を提案した。原材料情報抽出ための機械学習方法では 95.6% の再現率で原材料情報を抽出できる。商品種の CFP を推測する方法では商品種の CFP の最大値、最小値、平均値を推測できる。

本計算方法を推測した麺類商品 CFP と生産メーターからのデータを用いて計算した 14 品の麺類商品の CFP を比較した結果、実際の製品の CFP は全部推測した最大 CFP と最小 CFP の範囲にいる。本計算方法は、商品種の CFP を推測できると考えている。

第 4 章

カーボンフットプリント認知度向上の取組

4.1 取組

室蘭工業大学とコープさっぽろは、2008 年よりスーパーマーケットにおける CO_2 排出量削減をテーマにして活動してきた [47]。スーパーマーケットにおいて CO_2 排出量を削減するために、室蘭工業大学とコープさっぽろは、店舗や食品工場における CO_2 排出量の調査と削減、CFP による地球温暖化問題の消費者へのアピールを実施してきた。そして、製品の環境影響を伝えるために、CFP の計算結果を Point of purchase advertising(POP) 表示により店舗内に展示してきた [9]。

本章で述べる取組の目的は、CFP の認知度を向上させることである。2014 年から実施したこの取組の新規性は、CFP を消費者の健康と組み合わせて広報した工夫とその広報を実質化するために準備した 130 品の CFP の計算結果である。CFP が消費者の目につくように、より多くの食品・日用品で CFP を計算し、包装のデザインを工夫した。また、取組の意図をパンフレッ

トで紹介した。さらに、この取組の効果をアンケート調査で調べた。

本取組は、CFP の認知度を上げるために、次の 2 つの目標を組合せて実施した。

1. 消費者が CFP を見るようにする。
2. 消費者が CFP に興味をもつようにする。

消費者が CFP を見るようにするために、包装材のデザインを工夫し、CFP 付きの製品の数を増やした。CFP 付き製品が印象に残ることを期待して、包装材のデザインを他にないデザインにした。また、店舗内にある多数の製品で CFP を計算した。その数は 130 品である。図 4.1 に CFP を計算した結果の一部を示す。著者は、多くの製品で CFP を計算すれば、CFP に精力的に取り組んでいる姿勢をアピールできると考えた。

消費者が CFP に興味をもつように、CFP を食品の栄養成分表示を組み合わせて、紹介・アピールした。CFP と食品の栄養成分の組み合わせにより、CFP の取組みを消費者が身近に感じることを期待した。また、食品の栄養成分表示の意義と、CFP の地球温暖化対策としての意義を一緒に編集したパンフレットを配布した。

本取組を実施した製品の包装材の表を、単純かつ他にない個性的なデザインにした。図 4.2 に包装のデザインを示す。その単純な包装デザインでは、栄養成分表示と並べて、同図のように CFP を表示した。このデザインでは、栄養成分表示による健康への意識と CFP による環境への意識を組み合わせた。

また、多数の食品・日用品で CFP を計算して、多数の CFP 付き製品を準備することで、消

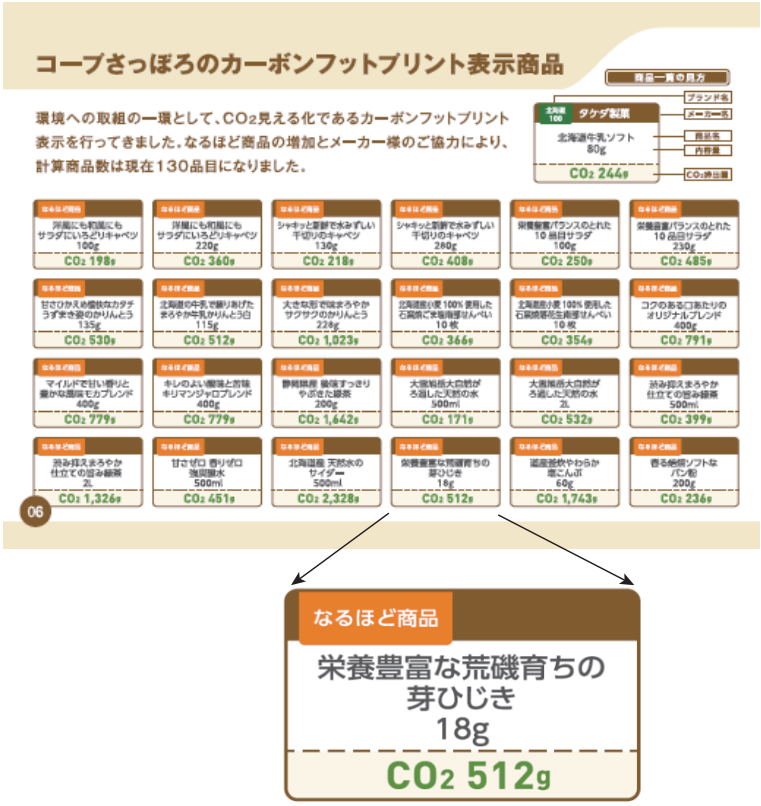


図 4.1　CFP を計算した 130 品の一部

費者が CFP に気づくようにした。敷地面積約 4000m^2 のスーパーマーケットには約 20000 点の製品が陳列されている。その中で、CFP 付きの製品を埋もれないようにするには、CFP を計算していることを訴えられる多くの製品が不可欠と考えた。計算した製品は 130 品である。130 品は店舗内に陳列されている製品の数に比べればまだ十分な多さとはいえない。一方、2016 年 10 月現在カーボンフットプリントコミュニケーションプログラムが認定している製品の累積数が 1280 であり、登録されている製品は 128 個である。この認定総数に比べると、本取組の 130 品は、一つの事業体が CFP を計算した数としては、十分に多い。

図 4.2　栄養成分表示と CFP の統一したデザインのラベル

4.2　130 品製品の計算結果

スーパーマーケットが CFP を計算する場合、ナショナルブランド (NB) とプライベートブランド (PB) で CFP の計算条件が異なる。NB は、スーパーマーケットがメーカーに CFP のデータ収集を依頼するため、必ずしもデータを取得できるわけではない。一方、PB は、スーパーマーケット社内の調整により、CFP の計算に必要なデータを取得できるため、NB に比べ容易に CFP に関するデータを集められる。

著者らが 2014 年 10 月に CFP を計算した 130 品は、食品が 124 品、日用品 (トイレットペー

パー等) が 6 品である。この 130 品の内訳は、コープさっぽろの PB「なるほど商品」が 50 品、「なるほど商品」以外のコープさっぽろの自社製品が 52 品、北海道の複数の食品メーカーが参加しているブランド「北海道 100」が 14 品、その他 14 品である。

コープさっぽろの PB「なるほど商品」は、コープさっぽろが一つの基本構想と規格でまとめたブランドである。PB「なるほど商品」では、統一した簡潔なデザインの包装を用いている。このブランドは、コープさっぽろの独自の PB であるため、他社の工場で製造していても、他の製品に比べて CFP に関するデータを容易に収集でき、包装デザインにも独自に工夫ができた。また、コープさっぽろが自社製造している食品も、PB「なるほど商品」と同様に比較的容易に CFP に関するデータを収集できた。

一方、ブランド「北海道 100」は、コープさっぽろを含めた複数のメーカーが共同参加しているブランドである。このブランド「北海道 100」では、コープさっぽろが他社に CFP に関するデータ収集を依頼し、データを収集した。しかし、「北海道 100」では、包装材への CFP に関する工夫は、各メーカーにゆだねられているため、包装デザインへ CFP を印刷できなかった。

本取組では、健康に関する栄養成分の表示と環境に関する CFP 表示を組み合わせた。消費者は、食品と健康を結びつけていると著者らは考えた。そして、食品と健康を結ぶ情報が、食品の栄養成分と著者は考えた。一方、著者らは食品と環境を結びつける仕組みとして CFP に取り組んできた。そこで、食品を仲立ちにして健康に関わる栄養成分と環境に関わる CFP を結びつけて、図 4.2 のように、CFP の表示と栄養成分の表示を、形を統一したデザインで並べて、包装に表示した。なお、CFP が栄養素の一つと誤解されないように、CFP は緑色あるいは赤色の地の

デザインにし、栄養素は白地のデザインに統一されている。さらに、図 4.3 に示すように、包装材の裏にマークの横に、CFP の説明を印刷した。

表示した栄養成分は、エネルギー、たんぱく質、脂質、炭水化物である。それらの栄養成分について、図 4.2 のように、当該食品 100g における含有量と 1 日当たりの摂取目安量に対する当該食品の割合を表示した。栄養成分と同じように、CFP を見てもらうことで、CFP を身近にする効果を著者らは期待した。

CFP を消費者に伝えるパンフレットを製作した。図 4.4 にこのパンフレットを示す。このパンフレットでは、計算した 130 品の製品名、メーカー名と CFP をまとめた。また、このパンフレットでは、130 品の CFP を紹介するだけでなく、図 4.5 のように、同じ種類の食品について、原材料生産、原材料輸送、工場、製品輸送、店舗の各段階で CFP を比較した。例えば、国内産大豆を使う豆腐とアメリカ産大豆を使う豆腐では、原材料調達の CFP で差が生じることをグラフで示した。さらに、図 4.5 のように家庭における都市ガス、LP ガス、電子レンジ、IH 調理器、冷蔵庫、液晶テレビの CO_2 排出量の例をまとめた。これらの他に、地球温暖化の影響を紹介するページも作成した。

製作したパンフレットは A5 判で 11 ページである。このパンフレットは店舗内に置いて、興味のある来店者が持ち帰るようにした。また、宅配サービスの時にこのパンフレットを配布した。

また、このパンフレットでも、図 4.6 のように、食品を通した消費者の健康と組み合わせて CFP を紹介した。「健康と環境に優しい食の選択を実践しましょう！」の表題をつけ、健康と環境を意識した一人一人の行動の大切さを訴えた。

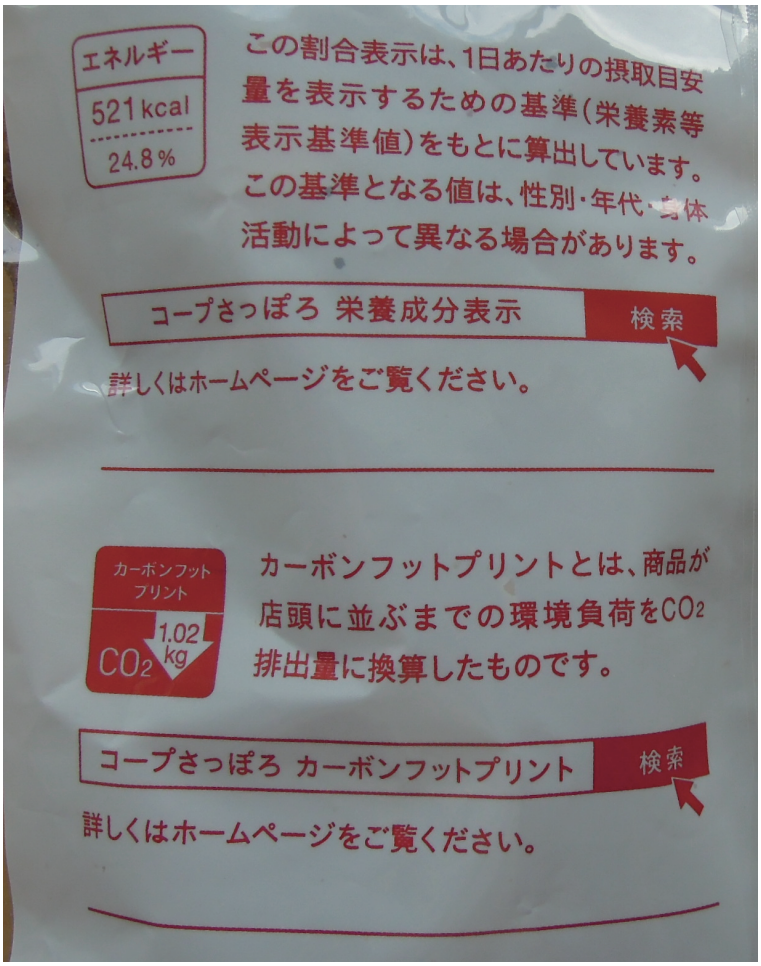


図 4.3　製品包装の裏に印刷した説明文

図 4.4　配布したパンフレット

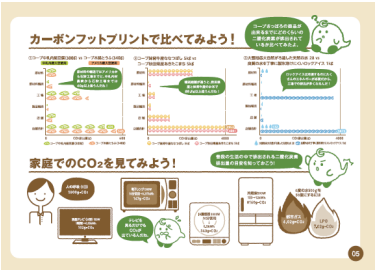
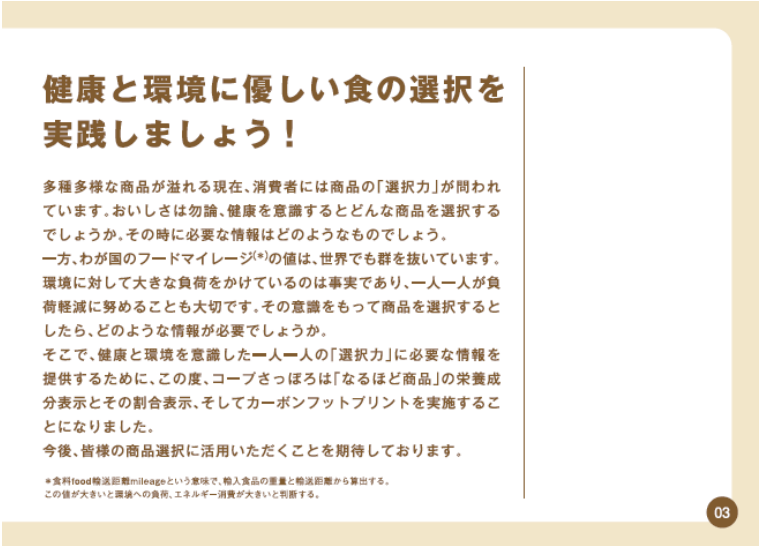


図 4.5　パンフレットの中で CFP を紹介するページ



健康と環境に優しい食の選択を実践しましょう！

多種多様な商品が溢れる現在、消費者には商品の「選択力」が問われています。おいしさは勿論、健康を意識するとどんな商品を選択するでしょうか。その時に必要な情報はどのようなものでしょう。
一方、わが国のフードマイレージ(*)の値は、世界でも群を抜いています。環境に対して大きな負荷をかけているのは事実であり、一人一人が負荷軽減に努めることも大切です。その意識をもって商品を選択するとしたら、どのような情報が必要でしょうか。
そこで、健康と環境を意識した一人一人の「選択力」に必要な情報を提供するために、この度、コープさっぽろは「なるほど商品」の栄養成分表示とその割合表示、そしてカーボンフットプリントを実施することになりました。
今後、皆様の商品選択に活用いただくことを期待しております。

＊食料food輸送距離mileageという意味で、輸入食品の重量と輸送距離から算出する。
この値が大きいと環境への負荷、エネルギー消費が大きいと判断する。

03

図 4.6　パンフレットの中で健康と環境の大切さをアピールするページ

4.3　カーボンフットプリン認知度の調査結果

本取組がコープさっぽろの店舗で開始された 2014 年 10 月から 2015 年 8 月の約 11 ヶ月において、次の問題点が判明した。

1. 包装の切り替えが進まない問題。コープさっぽろの PB「なるほど商品」50 品の内 15 品の包装のみが切り替わった。CFP を印刷していない古い包装材の在庫がある場合、その在庫を使い尽くすまで、CFP を印刷した新しい包装材に切り替えられない。製品の容器が小さく包装に余裕がないため、CFP を印刷できない場合もある。これらの理由により、店舗内に陳列されている製品おいて、食品の栄養成分と CFP を組み合わせた新しい包装を使った製品の数が限られていた。したがって、2014 年 10 月に開始したパンフレットを使った CFP の広報と店舗内の製品の陳列に時間差が発生してしまった。
2. コープさっぽろの PB「なるほど商品」以外の製品における CFP のラベルの問題。コープさっぽろ以外のメーカーの製品の CFP を知らせる方法が、パンフレットだけに限定されている。また、コープさっぽろにおいても、PB「なるほど商品」以外の自社製品においても、CFP を製品の包装に印刷できていない。

2014 年 9 月 27 日と 2015 年 8 月 22 日にコープさっぽろが主催する組合員向け展示会 (第 8 回と第 9 回の食べる・たいせつフェスティバル) で、著者らは CFP の認知の状況を調べた。第 1 回の調査 (2014 年 9 月 27 日) は、本取組 (CFP と栄養成分を表示した包装、パンフレット配布) が消費者に公開される直前に行われた。第 2 回目の調査 (2015 年 8 月 22 日) は、取組開始後約

11 ヶ月が過ぎた時期に行われた。調査を行った会場は第 1 回と第 2 回ともに札幌市スポーツ交流施設コミュニティドーム「つどーむ」である。このアンケート調査の様子を図 4.7 に示す。また、表 4.1 に調査方法をまとめる。2014 年の展示会には 102 の団体、2015 年には 111 の団体が出展した。著者らは、CFP に関わる活動を紹介するブースを開設し、そこでアンケート調査を行った。表 4.2 にアンケート項目を示す。アンケート調査は、著者のブースに訪れた大人に対して、CFP のマークと質問が記述されたボードを示しながら、口頭による質問に口頭で答えてもらう形式で行った。

表 4.3 にアンケートの結果をまとめる。 ① の質問において、CFP 自体を知っている人の割合は、2014 年に比べて 2015 年では 3% 少なくなった。② の質問では、図 4.2 に示した包装材に印刷した CFP のマークを見たことがある人の割合は 30% だった。一方、包装材への CFP マークの印刷等を行う前の 2014 年の調査では、店舗内の陳列棚に POP 表示した CFP のマーク (表 4.2 の左列 2 行目) を見たことがある人の割合が 4% だった。③ の質問にとりあげた CFP プログラム参加マークを知っている人の割合は、2014 年に比べて 2015 年では僅かながら増加していた。

表 4.3 のアンケート回答者は、一般の消費者の中でも、アンケート調査を行った展示会に興味のある層である。この展示会は、コープさっぽろが組合員に向けて 2007 年から毎年継続して行ってきた展示会である。組合員以外の一般の消費者に向けても、新聞やテレビにおいて展示会開催の広報を行っており、組合員以外の消費者も制限なく無料で入場できた。展示会の会場は札幌駅から北北東に約 4km にあり、丘珠空港の隣にある。最寄りの地下鉄駅は東豊線終点駅の栄

町駅である。会場は栄町駅から約 600m にある。展示会のこのような性質から、アンケート対象者は、この展示会を意識して参加した消費者である。

χ^2 適合度検定 [48] により、本論文で述べた取組の効果を評価する。取組前後のアンケート結果の度数を y_{ij}（ただし、$ij = 1, 2$）とする。度数 y_{ij} の最初の添字 i は取組実施前の場合 $i = 1$ とし、取組実施後の場合は $i = 2$ とする。また、度数 y_{ij} の 2 番目の添字 j は認知度を表し、知らない・見たことがない場合は $j = 1$、知っている・見たことがある場合は $j = 2$ とする。χ^2 適合度統計量 χ^2 は

$$\chi^2 = \frac{\sum_{ij}^{2} y_{ij} \left(y_{11}y_{22} - y_{12}y_{21}\right)^2}{\sum_{j}^{2} y_{1j} \sum_{j}^{2} y_{2j} \sum_{i}^{2} y_{i1} \sum_{i}^{2} y_{i2}} \tag{4.1}$$

である。なお、ここで検討する χ^2 適合度検定の自由度は $\upsilon = (2-1)(2-1) = 1$ である。まず、質問①でたずねた CFP 制度の認知度については、帰無仮説を「本取組は、CFP の認知度向上に効果がない。」とする。この場合、χ^2 適合度統計量は、

$$\chi^2 = \frac{228 \times (68 \times 22 - 15 \times 123)^2}{83 \times 145 \times 191 \times 37} = 0.33 \tag{4.2}$$

である。実現値 0.33 は、自由度 1、5% 点の χ^2 分布 $\chi^2_{0.05}(1) = 3.84$ より小さいため、帰無仮説「本取組は、CFP の認知度向上に効果がない。」は有意水準 5% で棄却できない。

質問②でたずねた本取組独自の CFP マークの認知度については、帰無仮説を質問①と同様に「本取組は、CFP 独自マークの認知度向上に効果がない。」とする。この場合、χ^2 適合度統計量は、

$$\chi^2 = \frac{229 \times (72 \times 46 - 3 \times 108)^2}{75 \times 154 \times 180 \times 49} = 20.0 \tag{4.3}$$

である。この実現値 20.0 は、先に述べた有意水準 5% の 3.84 より大きいため、質問②についての帰無仮説は有意水準 5% で棄却できる。つまり、本取組は CFP 独自マークの認知度向上に効果があったと言える。質問③でたずねた CFP プログラム参加マークの認知度向上に対する本取組の効果については、質問①②と同様に χ^2 適合度検定を行うと、χ^2 適合度統計量は、

$$\chi^2 = \frac{211 \times (70 \times 4 - 0 \times 108)^2}{75 \times 141 \times 207 \times 4} = 2.02 \tag{4.4}$$

である。この実現値 2.02 は有意水準 5% の 3.84 より小さいため、帰無仮説「本取組は、CFP プログラム参加マークの認知度向上に効果がない。」は有意水準 5% で棄却できない。

質問②の結果によれば、CFP 包装材への印刷が CFP マークを消費者の間に広めることに貢献している。これは、2014 年まで行ってきた店舗のみの POP 展示に比べ、包装材への印刷がより消費者の日常に入り込むためと著者は解釈する。

アンケート調査における消費者との対話で、図 4.2 に示した PB で統一した包装材のデザインを知っているが、その下にある栄養素と CFP のマークに気がついていない人が多くいることが判明した。表 4.3 の質問②の集計では、包装デザインを知っているが、CFP のマークに気がついていない人は「いいえ」にカウントしている。包装材のデザインが消費者の記憶に残っているので、4.1 節で述べた包装のデザインを単純かつ個性的にして消費者が CFP を見るようにする試みの一部は成果があったといえる。しかし、本取組の目的である栄養素と CFP のマークを消費者が見るようにするには、本取組の試みだけでは不十分であった。

表 4.1　CFP の認知度調査の方法

調査した日時	2014 年 9 月 27 日	2015 年 8 月 22 日
	10:00～ 16:00	10:00～ 16:00
調査を行った展示会	コープさっぽろ 第 8 回の食べる・ 大切フェスティバル 札幌会場 (札幌市スポーツ交流施設 コミュニティドーム つどーむ)	コープさっぽろ 第 9 回の食べる・ 大切フェスティバル 札幌会場 (札幌市スポーツ交流施設 コミュニティドーム つどーむ)
展示会来場者総数 (人)	7094	8517
アンケート調査方法	CFP の展示ブースへの来場者への聞き取り	
回答者の主な属性	主婦、小学生または幼児を連れた家族連れ	
回答者の年齢層	20 歳 ～70 歳	
回答者数 (人)	70～83	141～154

表 4.2　CFP のアンケート

2014 年 9 月 27 日	2015 年 8 月 22 日
① カーボンフットプリントを知っていましたか？	① カーボンフットプリントを知っていましたか？
② のマークを見たことがありますか？	② のマークを見たことがありますか？
③ のマークを見たことがありますか？	③ のマークを見たことがありますか？

表 4.3　CFP のアンケート結果

	2014 年 9 月 27 日		2015 年 8 月 22 日	
	はい	いいえ	はい	いいえ
① カーボンフットプリントを知っていましたか？	15 人 (18%)	68 人 (85%)	22 人 (15%)	123 人 (85%)
② のマークを見たことがありますか？	3 人 (4%)	72 人 (96%)		
② のマークを見たことがありますか？			46 人 (30%)	108 人 (70%)
③ のマークを見たことがありますか？	0 人 (0%)	70 人 (100%)	4 人 (3%)	173 人 (97%)

(a) 2014 年 9 月 27 日のアンケート調査

(b) 2015 年 8 月 22 日のアンケート調査

図 4.7　アンケート調査の様子

4.4 まとめ

本章では、スーパーマーケットにおいて消費者の CFP の認知度を上げる独自の取組とその課題を述べた。著者は、この取組の効果を上げるためには、CFP を計算した多数の製品を揃えることが肝要と考えた。

本取組の特徴は、消費者の CFP の認知度を上げるために、食品の栄養成分表示と CFP を組み合わせて、消費者にアピールした点である。CFP を計算した製品の一部において、包装に、食品の栄養成分と CFP を同じ傾向のデザインで並べて表記した。また、包装のデザイン自体も単純にして、食品の栄養成分と CFP が目に付きやすいようにした。さらに、CFP を紹介するパンフレットを配布した。

本取組の三つ目の特徴は、取組を始める前と開始後 11 ヶ月経過した時期に、CFP についてアンケート調査をして認知度を調べた点である。そのアンケート調査の結果を、χ^2 適合度検定使い調べたところ、本取組独自の CFP マークの認知度が向上していることを確認できた。一方、CFP 制度および CFP プログラム参加マークの認知度の向上は確認できなかった。アンケート調査に伴う聞き取りでは、包装デザインを知っているが本取組独自の CFP マークを知らない人が多くいたため、CFP の認知を広げる取組が不十分であることも分かった。なお、食品の栄養成分と CFP を並べたデザインの包装は、包装の在庫消化に合わせて順次拡大していく計画である。

第 5 章

結論

本論文では CFP の計算効率を向上するための CFP 自動計算システム、CFP の大きさの基準量を見えるためのテキストマイニングを用いて製品の CFP の範囲を計算する手法、CFP の認知度を向上するための CFP 取組と CFP 認知度の調査について、各手法の提案を通して、考察した。

提案した CFP 自動計算システムでは、従来の CFP 計算システムを改良し、CFP 計算の自動化を実現し、製品の CFP を高い効率で計算できる。この自動計算システムにより、CFP の計算時間が手動計算と比べて、 10% ～ 22% に低減された。

テキストマイニングを用いて製品の CFP の範囲を計算する手法では、ネットで公開している製品のデータを用いて、商品種の CFP の最大値、最小値、平均値を推定することができた。そして、この計算方法が示す製品 CFP の大きさが CFP を評価する基準量として使えることを示した。

原材料情報抽出するため、本論文では再帰自己符号化器 (RAE) と N-gram モデル解析を組み合わせた固有表現抽出法を提案した。この抽出法は、RAE を用いて任意単語数の文節の固有表

現解析し、N-gram モデルを用いて入力文節と前後文節の文脈を分析し、入力文節の固有表現を推定する。実験の結果は、F 値は 90.4% の結果を得られ、提案手法を任意単語数の文節の固有表現抽出の有効性を確認できた。

CFP の認知度を向上するためを実施した CFP 取組は、アンケート調査により、CFP マークの認知度が向上していることを確認できた。

著者は、CFP 自動計算システムにより、CFP に取組む企業が CFP を計算する製品を増やして、消費者が CFP を見る機会が増えることを期待する。また、CFP の大きさの基準量を示すことが、消費者が CFP を比較でき、CFP を意識できるようになると確信する。そして、最終的に、本研究の成果が低炭素社会の実現に寄与することを望む。

謝辞

本研究を進めるにあたり熱心な御指導御鞭撻を賜りました、室蘭工業大学

永野宏治教授に心より感謝いたします。

本論文を書くにあたり貴重な御指摘、御意見を賜りました、室蘭工業大学

板倉賢一教授、塩谷浩之教授、佐藤和彦准教授に心より御礼申し上げます。

本研究を進めるにあたり貴重な御指摘、御意見を賜りました、Manchester 大学 Professor Adisa Azapagic、Dr Martyn Jones、Dr Rosa Cuellar-Franca に心より感謝いたします。

CFP の取組に協力してだきました生活協同組合コープさっぽろ経営企画室

鈴木昭徳博士に深く感謝いたします。

そして、研究室における研究及び授業で様々な面で助言をしたいただきました

楊玉霞氏、鈴木郡氏、横堀祥平氏に心より感謝いたします。

第 2 章の CFP 計算システムの製作において、日本貨物鉄道株式会社には、貨物駅間の距離のデータベースを本研究に使うことを快く承諾していただきました。また、味の素株式会社には、食品関連材料 CO_2 排出係数データベースのデータの使用を承諾いただきました。トラックの輸送距離の取得では、Google Maps を利用しました。日本郵便株式会社には、市町村名として公

開しているデータを参考にさせていただきました。日本農林水産省には、公開している農林水産統計データを参考にさせていただきました。関係の皆様に感謝します。

本論文を完成させることができたのはこのような皆様のご協力のおかげです。本当にありがとうございました。

参考文献

[1] 内閣府大臣官房政府広報室:地球温暖化対策に関する世論調査, 世論調査, 入手先 <http://survey.gov-online.go.jp/h19/h19-globalwarming/>,(参照 2015-6-14) (2007).

[2] 内閣府大臣官房政府広報室:低炭素社会に関する特別世論調査, 世論調査, 入手先 <http://survey.gov-online.go.jp/tokubetu/h20/h20-lc_chosahyo.pdf>,(参照 2015-6-14) (2008).

[3] 増井利彦:低炭素社会シナリオ (特集 シナリオ分析と LCA)，日本 LCA 学会誌，Vol. 10, No. 3, pp. 270–278<http://ci.nii.ac.jp/naid/40020191266/> (2014).

[4] 稲葉敦，未踏科学技術協会:カーボンフットプリント : LCA 評価手法でつくる、製品別「CO_2 排出量見える化」のしくみ，工業調査会 (2009).

[5] カーボンフットプリントコミュニケーションプログラム:CFP 関連規程, 基本文書（文書管理番号 G-01-03）入手先 < http://www.cfp-japan.jp/regulation/pdf/G-01-03.pdf> (参照 2013-6-14) (2013).

[6] カーボンフットプリントコミュニケーションプログラム:CFP 宣言認定製品一覧，入手先 <http://www.cfp-japan.jp/calculate/verify/permission.php>,(参照 2013-6-12) (2013).

[7] 富士フイルムグローバルグラフィックシステムズ株式会社:カーボンフットプリント (CFP) 表示について,
入手先 < http://ffgs.fujifilm.co.jp/environment/cfp/>,(参照 2015-9-4) (2015).

[8] Standards,B.: Specification for the assessment of the life cycle greenhouse gas emissions of goods and services
入手先 < http://shop.bsigroup.com/en/forms/PASs/PAS-2050/>,(参照 2016-11-4)(2011).

[9] 永野宏治，吉田洋一，小松均，村上伸吾:コープさっぽろにおける食品カーボンフットプリントの店頭展示の取り組み，日本 LCA 学会誌，Vol. 9,No. 4,pp. 299–305 (2013).

[10] 趙培江，永野宏治，鈴木昭徳，佐藤和彦:スーパーマーケットにおけるカーボンフットプリントの認知度向上の取組とそれを支える計算機システムの導入，日本 LCA 学会誌， Vol. 12, No. 1, pp. 15–25 (2016).

[11] 椛島裕美枝，吉川直樹:米 (滋賀県産こしひかり) におけるカーボンフットプリント算定事例，日本 LCA 学会誌， Vol. 6, No. 3, pp. 229–233 (2010).

[12] 安藤生大，齊藤将人，鈴木基之:トイレットペーパーのカーボンフットプリント (CFP) の試算，紙パ技協誌， Vol. 69, No. 8, pp. 872–878 (2015).

[13] 神田太朗，瀧本真理，曽根真理，岸田弘之，花木啓祐， 藤田壮:建設資材の二酸化炭素排出量計算手法の共通化に関する検討，*Journal of Japan Society of Civil Engineers, Ser. G (Environmental Research)*, Vol. 67, No. 6, pp. II_267–II_278 (2011).

[14] 稲葉敦，風間理応，玉利有香，森本涼子:家庭での調理の CO_2 排出量の推算方法，日本 LCA 学会誌， Vol. 10, No. 2, pp. 155–164 (2014).

[15] Laurence, S. and Adisa, A.: Life cycle environmental impacts of UK shale gas, *Applied energy*, Vol. 134, pp. 506–518 (2014).

[16] 稲葉敦，森真之，風間理応，玉利有香，森本涼子:外食メニューの CO_2 排出量の推算方法，日本 LCA 学会誌， Vol. 10, No. 3, pp. 308–314 (2014).

[17] 永野宏治，鈴木郡，趙培江，佐藤和彦:カーボンフットプリント評価における輸送距離の精度調整計算法の提案，日本 LCA 学会誌，Vol. 12, No. 2, pp. 77–85 (2016).

[18] カーボンフットプリントコミュニケーションプログラム:認定 CFP-PCR 一覧，入手先 <http://www.cfp-japan.jp/calculate/authorize/pcr.php>(参照 2015-7-10) (2013).

[19] David, A., Cecil, C. and Adisa, A.: Environmental impacts of consumption of Australian red wine in the UK, *Journal of Cleaner Production*, Vol. 72, pp. 110–119 (2014).

[20] 味の素株式会社:味の素グループ版「食品関連材料 CO_2 排出係数データベース」，入手先 <http://www.ajinomoto.co.jp/activity/kankyo/pdf/2010/lcco2.pdf>(参照 2013-6-14) (2012).

[21] カーボンフットプリントコミュニケーションプログラム:基本データベース，入手先 <http://www.cfp-japan.jp/calculate/verify/data.html>(参照 2013-6-14).

[22] 環境省:日本の温室効果ガス排出量の算定結果, 2013 年度（平成 25 年度）温室効果ガス排出量,

入手先 <http://www.env.go.jp/earth/ondanka/ghg/>,(参照 2015-7-10) (2013).

[23] 一般社団法人産業環境管理協:環境ラベルプログラム（エコリーフ／カーボンフットプリントコミュニケーションプログラム）に関する中期行動計画, 入手先 <https://www.cfp-japan.jp/common/data_news/000658/1375320443.pdf> ,(参照 2016-11-07) (2013).

[24] みずほ情報総研株式会社:カーボンフットプリントに関する消費者受容性調査を実施, 入手先 <https://www.mizuho-ir.co.jp/company/release/2011/cfp1101.html> ,(参照 2016-11-07) (2011).

[25] 三田村正樹，新谷健仁，永野宏治:食品カーボンフットプリントの簡易自動計算システムの開発，第 8 回日本 LCA 学会研究発表会 講演要旨集，pp. 28–29 (2013).

[26] Bond, F., Baldwin, T., Fothergill, R. and Uchimoto, K.: Japanese SemCor: A Sense-tagged Corpus of Japanese, *The 6th International Conference of the Global WordNet Association (GWC-2012)* (2012).

[27] 海上保安庁:距離表，日本水路協会 (2011).

[28] 日清食品グループ:カップヌードル, 入手先 <https://www.nissin.com/jp/products/brands/cupnoodle/>,(参照 2016-11-16) (2016).

[29] 農林水産省:農林水産省/統計情報, 入手先 <http://www.maff.go.jp/j/tokei/index.html>,(参照 2015-12-16) (2015).

[30] 内元清貴，馬　青，村田真樹，小作浩美，内山将夫，井佐原均:最大エントロピー法と書き換

え規則に基づく日本語固有表現抽出，自然言語処理， Vol. 7, No. 2, pp. 63–90 (2000).

[31] 山田寛泰，工藤　拓，松本裕治:Suppor Vector Machine を用いた日本語固有表現抽出，情報処理学会論文誌， Vol. 43, No. 1, pp. 44–53 (2002).

[32] 中野桂吾，平井有三:日本語固有表現抽出における文節情報の利用，情報処理学会論文誌，Vol. 45, No. 3, pp. 934–941 (2004).

[33] Le, Q. V.: Building high-level features using large scale unsupervised learning, *2013 IEEE international conference on acoustics, speech and signal processing*, pp. 8595–8598 (2013).

[34] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L. and van den Driessche, G.: Mastering the game of Go with deep, *nature*, Vol. 529, pp. 484–489 (2016).

[35] Luong, T., Socher, R. and Manning, C. D.: Better Word Representations with Recursive Neural Networks for Morphology., *CoNLL*, pp. 104–113 (2013).

[36] Hashimoto, K., Miwa, M., Tsuruoka, Y. and Chikayama, T.: Simple Customization of Recursive Neural Networks for Semantic Relation Classification., *EMNLP*, pp. 1372–1376 (2013).

[37] Zou, W. Y., Socher, R., Cer, D. M. and Manning, C. D.: Bilingual Word Embeddings for Phrase-Based Machine Translation., *EMNLP*, pp. 1393–1398 (2013).

[38] 岡谷貴之:深層学習，講談社 (2015).

[39] Socher, R., Huang, E. H., Pennin, J., Manning, C. D. and Ng, A. Y.: Dynamic pooling

and unfolding recursive autoencoders for paraphrase detection, *Advances in Neural Information Processing Systems*, pp. 801–809 (2011).

[40] Brown, P. F., Desouza, P. V., Mercer, R. L., Pietra, V. J. D. and Lai, J. C.: Class-based n-gram models of natural language, *Computational linguistics*, Vol. 18, No. 4, pp. 467–479 (1992).

[41] 山下達雄，松本裕治:コスト最小法と確率モデルの統合による形態素解析，情報処理学会研究報告自然言語処理 (NL)， Vol. 1997, No. 53, pp. 85–90 (1997).

[42] 国立国語研究所:UniDic, *https://osdn.jp/projects/unidic/* (2016-9-29).

[43] Mikolov, T., Chen, K., Corrado, G. and Dean, J.: Efficient estimation of word representations in vector space, *arXiv preprint arXiv:1301.3781* (2013).

[44] Socher, R., Pennington, J., Huang, E. H., Ng, A. Y. and Manning, C. D.: Semi-supervised recursive autoencoders for predicting sentiment distributions, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, pp. 151–161 (2011).

[45] 経済産業省:統計 (METI/経済産業省),
入手先 <http://www.meti.go.jp/statistics/>,(参照 2015-12-16) (2015).

[46] 日清食品グループ:環境実績,
入手先 <https://www.nissin.com/jp/about/csr/environment/achievement/>,
(参照 2016-11-16) (2016).

[47] 永野宏治，岩佐達郎，安居光國，松山春男，吉田洋一，大見英明:温室効果ガス削減を目指したカーボンフットプリントの店頭展示への取り組み，室蘭工業大学紀要，pp. 3–11 (2011).

[48] 東京大学教養学部統計学教室:自然科学の統計学，東京大学出版会 (1992).