

持続可能な食品産業及び食環境の構築要因に関する解析  
ーデータマイニングによるアプローチー

2023年1月

千葉大学大学院園芸学研究科  
環境園芸学専攻

加藤 弘祐

(千葉大学審査学位論文)

持続可能な食品産業及び食環境の構築要因に関する解析  
ーデータマイニングによるアプローチー

2023年1月

千葉大学大学院園芸学研究科

環境園芸学専攻

加藤 弘祐

## 目次

Dissertation Summary .....	1
論文要約 .....	6
<b>第1章 序論</b>	
1.1 本論文の背景と課題 .....	11
1.2 本論文の構成 .....	13
<b>第2章 本論文で用いるデータマイニング手法の概要</b>	
2.1 データマイニングの概要 .....	17
2.2 ネットワーク分析の概要 .....	18
2.3 テキストマイニングと自然言語処理の概要 .....	19
2.4 科学計量分析の概要 .....	21
2.5 本論文における実装ツール .....	21
<b>第3章 食品関連産業に着目したASEAN各国の産業構造ネットワーク分析</b>	
3.1 はじめに .....	23
3.2 方法 .....	24
3.2.1 データ .....	24
3.2.2 産業連関表とネットワーク分析 .....	25
3.2.3 コミュニティの検出と各国の産業構造の相関性 .....	26
3.3 結果と考察 .....	27
3.3.1 各国の主要セクターの分析結果 .....	27
3.3.2 各国のフードシステム関連セクターに関する分析結果 .....	28
3.3.3 国家間における産業構造の相関性の結果 .....	30
3.3.4 考察 .....	32
3.4 結論と今後の課題 .....	32

## 第4章 社会科学分野における有機農業研究に対する科学計量分析

4.1	はじめに .....	34
4.2	方法 .....	35
4.2.1	データ .....	35
4.2.2	科学計量分析 .....	36
4.2.3	実装ツール：Citespace .....	36
4.3	結果と考察 .....	37
4.3.1	キーワード共起分析 .....	37
4.3.2	バースト検出 .....	41
4.3.3	考察 .....	43
4.4	結論と今後の課題 .....	43

## 第5章 パーム油産業における廃棄物処理に関する技術開発の動向

5.1	はじめに .....	45
5.2	方法 .....	46
5.2.1	データ .....	46
5.2.2	ニューラルトピックモデル .....	49
5.2.3	分析の概要 .....	50
5.3	結果と考察 .....	50
5.3.1	トピックモデルの分析結果 .....	50
5.3.2	時系列トピックモデルの分析結果 .....	53
5.3.3	考察 .....	54
5.4	結論と今後の課題 .....	55

## 第6章 英国と米国における肥満に関する報道の比較分析

6.1	はじめに .....	57
6.2	方法 .....	58
6.2.1	データ .....	58
6.2.2	ニューラルトピックモデル .....	58

6.2.3	分析の概要 .....	59
6.3	結果と考察.....	59
6.3.1	クラス分類とトピック間類似度 .....	59
6.3.2	時系列トピックモデリングと二紙間比較 .....	61
6.3.3	砂糖税・ソーダ税の経緯と分析結果との照合 .....	62
6.3.4	考察 .....	64
6.4	結論と今後の課題.....	64

## 第7章 保健機能食品に対する消費者の認知構造

7.1	はじめに .....	66
7.2	保健機能食品制度に対する消費者認知 .....	67
7.2.1	背景と目的 .....	67
7.2.2	方法 .....	68
7.2.3	結果と考察 .....	70
7.2.4	小括 .....	76
7.3	健康効果と特定保健用食品・機能性表示食品の連想に関する認知分析 .....	77
7.3.1	背景と目的 .....	77
7.3.2	方法 .....	77
7.3.3	結果と考察 .....	79
7.3.4	小括 .....	84
7.4	結論と今後の課題.....	84

## 第8章 本論文の結論と今後の展望

8.1	本論文の結論.....	86
8.1.1	各章の要約 .....	86
8.1.2	データマイニングによるアプローチに関する考察 .....	89
8.1.2	本論文の結論 .....	91
8.2	今後の展望と残された課題 .....	92

引用文献 ..... 93

謝辞

## 初出一覧

初出論文（既公表論文）等は以下のとおり。ただし，加筆や修正を行っている。

### 第 1 章

書き下ろし

### 第 2 章

書き下ろし

### 第 3 章

加藤弘祐・小林弘明：ASEAN 各国の産業構造に対するネットワーク指標に基づいた比較分析－フードシステム関連セクターを中心として－，*フードシステム研究*, 27 巻, 4 号, 280-285, 2021 年 3 月。

### 第 4 章

Kato Kosuke, Yamamoto Junko, Kobayashi Hiroaki: Understanding Research Trends and Topics on Organic Agriculture based on Scientific Literature in Social Science: A Scientometric Analysis using Citespace, *Journal of Food System Research*, 29(4), 237-242, 2023 年 3 月。

### 第 5 章

Kato Kosuke, Kobayashi Hiroaki: Semantic Analysis of Palm Oil Industry Patents Using A Neural Topic Model to Determine Technological Transition, 日本国際地域開発学会 2022 年度春季大会講演要旨, 30-31, 2022 年 6 月。

### 第 6 章

加藤弘祐・小林弘明：クオリティペーパーを対象とした分散表現に基づくトピック分析－肥満に関する報道の英米比較を事例として－，*フードシステム研究*, 28 巻, 4 号, 328-333, 2022 年 3 月。

### 第 7 章

加藤弘祐・森嶋輝也：テキスト解析を用いた保健機能食品に対する消費者の認知分析－感情分析とネットワーク分析による接近－，*フードシステム研究*, 26 巻, 4 号, 313-318, 2020 年 3 月。

加藤弘祐・森嶋輝也：特定保健用食品と機能性表示食品に関する消費者の認知構造－保健機能及び機能性の表示文に着目した共起ネットワーク分析－，農業経済研究, 93 巻, 4 号, 395-400, 2022 年 3 月.

## 第 8 章

書き下ろし

# Analysis of Factors for Building a Sustainable Food Industry and Healthy Food Environment Using Data Mining

Kosuke KATO

## Dissertation Summary

This thesis focuses on the development of the food industry and healthy food environment as research topics related to sustainability, which has recently gained importance in all fields, and attempts to analyze the factors that contribute to the development of the food industry and healthy food environment. This dissertation consists of eight chapters: Chapter 1 is an introduction, Chapter 2 is a summary of the data mining analysis methods used in each chapter of this dissertation, Chapters 3 through 7 is a discussion of each research topic, and Chapter 8 is a conclusion and explains the future perspective of this dissertation.

Chapter 1 provides an overview of sustainability initiatives that have become increasingly active in recent years and discusses the need to construct the food industry and environment (the main subjects of analysis in this thesis), with sustainability in mind. The paper shows that to build a sustainable food industry and environment, it is necessary to discuss sustainability from multiple perspectives. To this end, it is necessary to analyze a variety of data; therefore, a data mining approach is effective for this purpose.

In Chapter 2, we provide an overview of the data mining methods used in this study. Particularly, it outlines three approaches—network analysis, text mining and natural language processing, and scientific quantitative analysis—and summarizes the specific analysis methods and implementation tools used in each chapter of this thesis.

In Chapter 3, we analyze the industrial structure of ASEAN, whose member countries are expected to expand their markets in the medium-to-long term to develop a sustainable food industry owing to their remarkable economic development. The data to be analyzed are the national input-output tables of each country, which can be obtained from the OECD database, and network analysis, where a data mining approach is applied. Specifically, the industrial transactions of the entire country were the object of analysis, and the industrial structures from 2005 to 2015 were quantitatively analyzed using two centralities, Random Walk Centrality and Counting Betweenness, focusing on the industrial sector related to the food system. Random Walk Centrality and Counting Betweenness are suitable indicators for the analysis of input-output tables. The sectors with high centralities were identified as the major industrial sectors in each country. Furthermore, by analyzing the correlation of industries among ASEAN countries, the similarities in the industrial structure of each country were also

examined. Agriculture, forestry, and fisheries were found to be major sectors in Cambodia, and food-related industries were found to be major sectors in Vietnam. Comparing the analysis results from 2005 to 2015, we found that the major sectors in each country remained almost unchanged. Furthermore, the correlation of the industrial structure based on the centrality ranking of each sector shows no significant change in the overall industrial structure between 2005 and 2015 in any country, as was the case for the major sectors. Based on the centrality ranking of food system-related sectors, three categories of ASEAN countries with similar industrial structures were identified. The three countries with high rankings for food system-related sectors are Cambodia, Vietnam, and the Philippines; the two countries with low rankings are Singapore (an industrial-oriented city-state) and Brunei (whose main industry is energy resources); and the three countries with medium rankings are Indonesia, Malaysia, and Thailand, which have food system-related sectors.

Chapter 4 addresses research issues related to organic agriculture, which are important for building sustainable agricultural production systems. Organic agriculture is considered an important approach to realize sustainable agriculture in Japan; as the “Strategy for Sustainable Food Systems, MeaDRI” has set forth a policy to expand organic agriculture. In response to this situation, academic literature related to organic agriculture has increased in recent years. It is difficult to grasp the overall trend of such a large amount of academic literature data using conventional human review; therefore, a de-mining approach is considered effective. The analysis targets academic literature included in the Web of Science, an academic article database, which has been assigned the Social Science Citation Index. To provide a sense of current trends, the literature collection was a 20-year time span between publications in 2002 and 2021. In Chapter 4, we used the co-occurrence analysis of keywords in scientific econometric analysis, which is implemented in CiteSpace, a scientific econometric tool, and Burst Detection, a method that can detect keywords that occur frequently in a certain period. First, the analysis revealed various research trends and topics related to organic agriculture. Specifically, five main research topics were identified: (1) case studies, (2) producers, (3) consumers, (4) policies, and (5) sustainability and environmental aspects of organic farming. Second, case studies on organic agriculture have accumulated in many regions, especially in California and Spain, and the transition of regions where such case studies have been active has been confirmed. In other words, it was suggested that case studies in California were actively conducted in the 2000s, and in Spain in the 2010s. Third, research on the sustainability of organic farming has been particularly active in recent years, suggesting that attention on organic farming from the sustainability perspective is increasing. Furthermore, the results of the analysis as a whole indicate a transition in the trend of academic research in organic agriculture and suggest that new research is being developed based on the accumulation of academic literature to date.

In Chapter 5, we discuss research issues related to palm oil, which is the world’s largest producer

of vegetable oils and fats. Simultaneously, various problems have been internationally identified, such as tropical rainforest destruction, waste, and child labor. There is also a need to develop waste treatment technologies to reduce environmental impact. In this study, we focused on the problem of waste treatment associated with palm oil production and analyzed patent information related to waste treatment. The objective of this study was to provide an overview of the technological evolution of waste treatment from the viewpoint of a new classification based on the semantic aspect of the neural topic model. The dataset of palm oil-related patents collected by the World Intellectual Property Organization (WIPO) was used as the target of analysis. The dataset was analyzed using patents related to palm oil. The results of applying the neural topic model, a framework that enables analysis including the semantic aspects of text, showed that it is possible to extract waste treatment technology based on its semantic aspects. Particularly, the results showed that the classification of biomass technologies related to waste treatment is based on specific technical details, such as second-generation biomass technologies, which are not sufficiently accessible by conventional patent classification. Additionally, patents on palm fiber were continuously published from 1995 to 2015, suggesting that technologies related to palm fiber, such as mattress applications, have been widely developed. However, only a few patent applications have been filed for Palm Oil Mill Effluent (POME), which is the effluent produced during palm oil production, indicating that there is a high possibility that technological development has not progressed. Although the scope of the analysis in Chapter 5 is limited to patent information, the results provide some implications for the development of waste treatment technologies associated with palm oil production.

In Chapter 6, we focus on obesity, which is a critical issue for building a sustainable and healthy food environment. Obesity is an important food and health issue and is considered a particularly serious social problem in the United States (US) and United Kingdom (UK). In this chapter, we apply text mining to newspaper articles that contain obesity-related content. We analyzed *The New York Times* and *The Guardian*, which are internationally recognized as quality newspapers publishing high-quality articles. As *The New York Times* is US-based and *The Guardian* is UK-based, we approached US and UK social interests through these newspapers. The results of the analysis revealed topics that are commonly reported and unique to both newspapers by quantitatively measuring the degree of similarity between topics in accordance with semantic aspects. Overall, the topic-extraction results revealed various obesity-related social interests, including topics related to obesity stigma and child obesity. The extracted topics were divided into three major categories: social interest related to child obesity; disease aspect of obesity, such as diabetes; and policy aspect of obesity, such as sugar tax. Topics unique to each newspaper that reflected the social context of the two countries were also identified, such as those presumably related to the British Royal Family and those likely to be related to US food companies. In Chapter 6, we also focus on the sugar tax topic

(which is common to both *The New York Times* and *The Guardian*) and clarify the differences in the content of articles. In addition, to verify chronological changes in the sugar tax topic, we compared the content of the articles with actual events related to sugar tax. The results showed that the words representing the topic changed in tandem with the actual events, indicating that the articles accurately captured ever-changing social concerns. The results can be considered an indication of the usefulness of newspaper article analysis using the neural topic model. Although the results cannot necessarily be taken as a reflection of the general public in both countries, they can be considered a revealing part of the social interests.

Chapter 7 discusses health foods that are considered to play a role in creating a healthy eating environment. Food with health claims system is a certification system for products with health benefits and includes food for specified health uses, food with functional claims, and food with nutrient function claims. We conducted a questionnaire survey on consumer recognition of such foods with health claims and applied text mining using network analysis to the survey results. Thus, two tasks were set. The first task was to compare consumer recognition of the three systems that comprise health foods. As an analysis method, we applied sentiment analysis—a method for estimating the writer’s feelings expressed in the text—to incorporate the measurement of positive and negative scales into the analysis. By combining sentiment analysis and text mining with network analysis of survey response texts, we approach consumers’ perceptions by capturing their feelings. In the second task, we focused on the health effects of food for specified health use and food with functional claims, and approached the relationship between such health effects and the recognition of both systems. The results revealed differences in the level of trust correlated with the level of recognition of the food for specified health uses and food with functional labeling systems, and also confirmed the skepticism toward the health effects of each product. Furthermore, in the second issue, not only were the products and images that consumers associate with the health effects labeling such as “reduce body fat” clarified, but the confusion between the food for specified health uses and food with functional labeling systems and the possibility of misperception of the health effects of each product were also revealed. The results of these two analyses were similar for both systems. The common finding for both analyses is the need to provide consumers with more information, especially appropriate communication, about the health effects of each product.

As described above, Chapters 3 through 7 of this thesis focus on the issues that are important in building a food industry that contributes to people’s health and achieves sustainable development, and analyze them by applying data mining methods to approach the important factors of building such an industry.

In Chapter 8, we summarize the results of the research cases described in Chapters 3 through 7, deepen our discussion on the characteristics of the data mining methods applied, and present the

conclusions of this thesis based on these research cases. Although the analysis cases discussed in the above chapters are limited in scope and subject matter, the results have certain implications for the construction of a sustainable food industry and environment. Based on the analysis results and discussions, the conclusions of this thesis are discussed. First, the results of Chapter 3, focusing on food-related industries in ASEAN, suggest that the food industry has strong ties with other industries, and that the economic effects generated by the food industry have a significant impact on other industries, although differences prevail among countries. Therefore, efforts to build a sustainable food industry and environment are expected to have a significant ripple effect on other industries. In this regard, the promotion of organic agriculture discussed in Chapter 4 can be viewed as something that has the potential to affect not only agriculture but also the entire food industry. The palm oil industry, discussed in Chapter 5, is not limited to organic agriculture; the entire industry, including the palm oil industry, must develop waste treatment technologies to reduce the environmental burden. In Chapter 5, it was suggested that technological development for waste treatment generated in the palm oil production process is biased toward waste such as palm fiber. However, the results also showed that development is progressing with environmental load reduction in mind, including technological development related to biomass. Chapter 6 analyzes social concerns related to obesity, reflecting diverse social interests, and suggests the need for a multifaceted approach to creating a healthy food environment, with the aim of building a sustainable food environment. Chapter 7 focuses on consumer awareness of health foods, which are important for promoting people's health, and clarifies the need for further provision of information to consumers. These analysis results suggest that to build a sustainable food industry and environment, all stakeholders, including food companies and consumers, or society as a whole, need to promote initiatives to enhance sustainability.

The remaining part of this paper approaches such individual issues, because although we have obtained knowledge of general trends of food-related industries and the food environment, we have not sufficiently approached further detailed issues suggested by this study. Furthermore, from a methodological perspective, this study utilized various trend and group detection methods, but the comparison of these methods and their superiority have not been sufficiently studied.

## 論文要約

### 持続可能な食品産業及び食環境の構築要因に関する解析 ーデータマイニングによるアプローチー

加藤 弘祐

本論文は、近年あらゆる分野において重要視されている持続可能性に関連する研究課題として、食品産業の発展および健康的な食環境に焦点を当て、その構築要因に対して分析を試みた研究である。本論文は、序論である第1章、本論文の各章で利用しているデータマイニングの分析手法を概説した第2章、各研究課題について論じた第3章から第7章、そして本論文の結論と今後の展望について論じた第8章、以上の全8章から構成される。

序論である第1章においては、近年ますます活発となっている持続可能性に関する取り組みについて概観し、本論文の主要な分析対象となる食品産業および食環境について、持続可能性を念頭に置いた上で構築していく必要性について論じた。そして、持続可能な食品産業および食環境を構築するためには、多面的に持続可能性について論じることが必要であることを示した。そのためには、多様なデータから分析を行う必要があり、分析手法としてデータマイニングによるアプローチが有効であることを述べた。

第2章においては、本論文で用いたデータマイニング手法の概要について整理した。特に、ネットワーク分析、テキストマイニング及び自然言語処理、科学計量分析という3つのアプローチについて概説した上で、本論文の各章で利用している具体的な分析手法や実装のツールについてまとめた。

第3章においては、持続可能な食品関連産業の発展を展望するため、経済発展が著しく、今後の中長期的な市場拡大が予想される国々がメンバー国を構成している ASEAN を取り上げ、ASEAN の産業構造について分析を行った。分析対象のデータとして OECD のデータベースから取得可能な ASEAN 各国の国家産業連関表を利用し、データマイニングアプローチの一つであるネットワーク分析を適用した。具体的には、国全体の産業取引を分析対象として取り上げ、2005年時点と2015年時点での産業構造について、フードシステムに関連する産業セクターを中心に、産業連関表を分析対象とする際に適した分析指標である Random Walk Centrality 及び Counting Betweenness という2つの中心性を用いて定量的に分析した。そして、各国の主要な産業セクターとして、これらの中心性の値の高いセクターを検出した。更に、ASEAN 各国について、国同士の産業の相関性についても分析し、各国の産業構造の相関性に関して考察を深めた。分析の結果、カンボジアにおいては農林水産業が主要セクターとして検出され、ベトナムでは食品関連産業が主要セクターであることが明らかになった。そして、2005年時点と2015年時点の分析結果を比較すると、各国

における主要セクターはほとんど変動していないことがわかった。更に、各セクターの中心性順位に基づいた産業構造の相関性を算出した結果、いずれの国においても、主要セクターと同様、2005年時点と2015年時点で、全体の産業構造の大きな変化は確認されなかった。そして、フードシステムに関連する産業セクターの中心性順位にもとづき、ASEAN各国の中でも産業構造が類似した国々の3分類が明らかになった。すなわち、フードシステム関連セクターが高順位となったカンボジア、ベトナム、フィリピンの3カ国、低順位となった工業中心の都市国家であるシンガポールとエネルギー資源が主要産業であるブルネイの2カ国、そしてインドネシア、マレーシア、タイというフードシステム関連セクターが中順位となった3カ国、という3分類が検出される結果が得られた。

第4章は、持続可能な農業生産システムを構築する上で重要となり得る有機農業に関する研究課題を取り上げた。有機農業は、わが国においても2022年に施行されたみどりの食料システム戦略の中で拡大方針が打ち出された通り、持続可能な農業を実現する上での重要な取り組みとして捉えられる。そうした状況も手伝って、有機農業に関連する学術文献は近年ますます蓄積されるようになってきている。こうした大量の学術文献データに対しては、従来の人力による研究レビューでは全体の傾向を把握するのは困難であり、データマイニングアプローチの採用が有効であると考えられる。そこで、本章では、分析対象として、学術論文データベースであるWeb of Scienceに収録されており、なおかつSocial Science Citation Indexが付与されている学術文献を取り扱った。文献の収集対象期間としては、近年の動向について接近するため、2002年～2021年の間に出版された20年間分とした。データマイニング手法としては、科学計量分析ツールであるCitespaceに実装された分析の中でも、特に、論文のキーワードを対象としたキーワードの共起分析、そして、ある時期に頻出しているキーワードを検出可能な手法であるバースト検出(Burst Detection)を用いた。分析の結果として、第一に、有機農業に関する様々な研究動向やトピックを検出した。具体的には、(1)事例研究、(2)生産者に関する研究、(3)消費者に関する研究、(4)政策に関する研究、(5)有機農業の持続可能性ないし環境的側面に関する研究、という5つの研究トピックが主だったものとして抽出された。第二に、カリフォルニアやスペインを中心に多くの地域で有機農業に関するケーススタディが蓄積されており、そうしたケーススタディが盛んな地域の変遷についても確認された。すなわち、2000年代においてはカリフォルニアに関する事例研究が盛んに行われており、2010年代においてはスペインに関する事例研究が盛んに取り組まれていることが示唆された。そして、第三に、有機農業の持続可能性に関する研究が近年特に盛んになっており、持続可能性という観点から有機農業への注目が高まっていることが推察された。更に、分析結果を総合すると、有機農業における学術研究の動向の変遷が示されたと共に、これまでの学術文献の蓄積による新たな研究が展開されていることが示唆されたものと考えられる。

そして、第5章では、植物性油脂の中でも世界最大の生産量であり、それと同時に様々

な問題が指摘されているパーム油に関する研究課題を取り上げた。2023年現在、パーム油はインドネシアやマレーシアが主要な生産国であるが、熱帯雨林の破壊、廃棄物の排出、児童労働などの問題が国際的に指摘される状況にあり、環境負荷低減の観点から、廃棄物処理の技術開発についても求められている。そこで、本章ではパーム油生産に伴う廃棄物の処理問題に焦点を当て、廃棄物処理に関する特許情報を分析した。具体的には、これまでの廃棄物処理に関する技術の変遷について、ニューラルトピックモデルによる意味的側面を踏まえた分析による新たな分類という観点から概観することを試みた。分析対象として、WIPO (World Intellectual Property Organization) により収集されたパーム油関連特許に関するデータセットを利用し、データセットの内、研究の目的に合わせて、特に廃棄物処理技術に分類される特許情報を分析の対象とした。テキストの意味的な側面を含めて分析可能な枠組みであるニューラルトピックモデルを適用した結果、廃棄物処理技術の技術的特徴に基づいた抽出が可能であることが示された。特に、廃棄物処理に関連したバイオマス技術に関しては、第2世代バイオマス技術といった具体的な技術内容を踏まえた形での分類が抽出される結果となった。こうした技術の詳細については、従来の特許分類では十分には接近しにくい点であり、本手法適用の有効性が示された。また、分析対象期間の1995年から2015年の間、パームファイバーを対象とした特許の出版が継続的に行われており、マットレス用途など、パームファイバーに関する技術が広く開発されていることが示唆された。その一方で、パーム油生産の際に生じる廃液である POME (Palm Oil Mill Effluent) については、特許出願は依然として少数となっており、技術開発が進展していない可能性が高いことがわかった。以上、第5章での分析対象は特許情報という限られた範囲ではあるものの、パーム油生産に伴う廃棄物処理技術の開発動向について、一定のインプレーションが得られる結果となった。

第6章においては、持続可能で健康的な食環境を構築する上で、今日における重大な課題となっている肥満 (Obesity) を取り上げた。肥満は、食と健康に関する重要な課題であり、米国と英国において特に深刻な社会問題とされる。第6章では、そうした肥満に関する内容を含んでいる新聞記事テキストを対象にテキストマイニングを適用し、分析を試みた。分析対象とした新聞紙は、質の高い記事を掲載している新聞、すなわちクオリティペーパーとして国際的に認知されている The New York Times と The Guardian である。The New York Times は米国の新聞であり、The Guardian は英国の新聞であるため、これらの新聞を通して、英国と米国の社会的関心に接近した。分析の結果、意味的側面に則った上で、トピック間の類似度を定量的に計測することにより、両紙に共通して報道されているトピック、そして両紙に特有のトピックを明らかにした。また、トピック抽出の全体的な結果としては、肥満のスティグマに関するトピック、児童の肥満に関連したトピックなど、肥満に関する多様な社会的関心が抽出される結果となった。抽出されたトピックを大きく分類すると、児童肥満に関連する社会的関心、糖尿病といった肥満の疾患的側面に対する社会

的関心、そして、砂糖税といった肥満に関連する施策面への社会的関心に分かれた。各新聞独自のトピックとして、英国王室と関連があると推察されるトピックや、米国の食品企業と関連するであろうトピックなど、両国の社会的な背景を反映させたトピックについても確認された。第6章では、抽出されたトピックの内、両紙に共通して抽出された砂糖税に関するトピックに着目し、The Guardian と The New York Times における砂糖税に関する記事内容の差異について明らかにした。特に、砂糖税のトピックにおける内容の時系列変化についても検証するため、実際の砂糖税に関する出来事との照合を図った。その結果、実際の出来事と連動してトピックを代表する単語が変化していることを確認し、時々刻々と変化する社会的関心について、本分析が正確に捉えていることを示した。第6章で示された分析結果は、ニューラルトピックモデルを用いた新聞記事分析の有用性を示すものであると捉えることが出来る。ここでの分析対象は、米国および英国における新聞社として両国1社ずつ取り上げており、分析結果は必ずしも両国全体の世相を反映したものと捉えることは出来ないが、両国における社会的関心の一端を明らかにした結果であると考えることが出来るであろう。

第7章では、健康的な食環境を構築する上で一定の役割を果たすと考えられる保健機能食品を取り上げた。保健機能食品制度は、健康効果を有する製品に対する認証制度であり、特定保健用食品、機能性表示食品、栄養機能食品が存在している。第7章においては、そうした保健機能食品に対する消費者認知に関するアンケート調査を実施し、調査結果に対して、ネットワーク分析を活用したテキストマイニングを適用した。第7章では2つの課題を設定しており、第1の課題では、保健機能食品を構成する3つの制度に対する消費者認知の比較を行った。分析手法として、テキストに表れている書き手の感情を推定する手法である感情分析を適用することにより、ポジティブとネガティブという尺度の測定を分析に取り込んだ。アンケートの回答テキストに対して感情分析とネットワーク分析によるテキストマイニングを適用することで、消費者の認知について、感情を捉えた上で接近することが可能となる。そして、第2の課題では、特定保健用食品と機能性表示食品が有する健康効果に焦点を当て、そうした健康効果と両制度との認知の関係性について接近した。分析の結果、第1の課題では、保健機能食品制度に対する認知度に相関した信用度の差異が明らかになり、各製品が有している健康効果への懐疑といった認知についても確認された。更に、第2の課題では、「体脂肪を減らす」といった健康効果に関する表示から消費者が連想する製品やイメージなどが明らかになっただけでなく、特定保健用食品と機能性表示食品に関する両制度同士の混同、そして各製品が有する健康効果に対する混同、すなわち健康効果の誤認の可能性が指摘される結果が得られた。これらの2つの分析結果に共通する知見として、消費者への更なる情報提供が必要であり、特に各製品が有している健康効果に関する適切な発信が要ることがわかった。

以上の通り、本論文の第3章から第7章では、人々の健康への貢献と持続可能な発展を

達成するような食品産業と食環境を構築していく上で重要となる各課題に焦点を当て、データマイニング手法の適用による解析を図り、その構築要因について接近した。そして、本論文の最終章である第8章においては、以上の第3章から第7章の研究事例の結果について要約した上で、適用したデータマイニング手法の特徴に関する考察を深めると共に、それらの研究事例から得られた本論文の結論について述べた。結論としては、まず、各章で取り上げた分析事例は限定的な対象・範囲であるものの、持続可能な食品産業および食環境の構築について、一定のインプリケーションを有する結果が得られたと考えられる。食品産業という観点においては、ASEANの産業構造について食品関連産業を中心に接近した第3章の分析結果から、国ごとに差異はあるものの、食品産業は他産業との結びつきも強く、食品産業で生じた経済効果などが他産業にも大きな影響を及ぼすことが示唆された。したがって、持続可能な食品産業および食環境の構築に向けた取り組みは、他産業にも大きく影響が波及すると考えられる。そういった点で、第4章で取り上げた有機農業の推進は、農業のみならず、食品産業を通して、産業全体に影響を与える可能性があるものとして捉えられ得る。他産業に影響を及ぼすという点では、有機農業に限らず、第5章で取り上げたパーム油産業も同様であり、パーム油産業をはじめとして、産業全体で環境負荷の低減に向けた廃棄物処理技術の開発を進めていく必要があるであろう。また、第5章においては、パーム油の生産過程で生じる廃棄物処理について、技術開発がパームファイバーといった一部の廃棄物に偏っていることが示唆されたが、バイオマスに関する技術開発の進展など、環境負荷低減を念頭に置いた開発が進んでいることも分析結果に反映された。そして、持続可能な食環境の構築に向けた観点では、第6章において肥満にまつわる社会的関心を分析する中で、多様な社会的関心が抽出され、健康的な食環境を構築する上で、そうした社会的関心を捉えた上で、多彩な面からアプローチする必要性が示唆された。更に、第7章では、人々の健康増進を図る上で重要となる保健機能食品に関する消費者認知に焦点を当てた分析を行い、その結果、消費者への更なる情報提供の必要性が明らかになった。以上の分析結果を総合すると、持続可能な食品産業および食環境を構築するためには、食品企業や消費者を含めたあらゆるステークホルダー、すなわち社会全体で持続可能性を高めるような取組を推進していく必要があると考えられる。

本論文の残された課題として、食品産業や食環境の総体的な傾向の知見については得られたと考えられる一方で、本研究の各分析結果から示唆された更なる研究課題については、十分に接近出来ていないため、そうした各課題へのアプローチが挙げられる。そして、本研究では様々なデータマイニング手法、特に、トレンド検出やグループ検出に関わる手法を活用したが、これらの各手法の比較や、それぞれの手法が持つ優位性については、十分には検討出来ていない。したがって、手法面に関する観点からは、今後の課題として、そうした各手法に関する定量的な比較検討が挙げられるであろう。

## 第1章 序論

### 1.1 本論文の背景と課題

プラネタリー・バウンダリー (Planetary boundaries) と呼ばれる概念 (Rockström et al., 2009; Steffen et al., 2015), すなわち, 人類が安全に活動出来る地球環境の限界点に関する概念が提唱されて以来, 人間活動を将来的に維持する上で, 社会全体の持続可能性 (Sustainability) を高めることがますます重要視されるようになった。

2000年に採択されたミレニアム開発目標 (MDGs: Millennium Development Goals) の後継として, 2015年に持続可能な開発サミットにて採択された持続可能な開発目標 (SDGs: Sustainable Development Goals) は, プラネタリー・バウンダリーの概念を踏まえた上で, そうした人間活動の将来的な持続可能性を高めるために設定されたものであり, 現在の国際社会の中でも特に重要な目標として認識されている。例えば, 農業に関連のある制度として, 持続可能性を念頭に置いた農業の推進のため, 2020年, EUはFarm to Folk戦略を打ち出した。そして, Farm to Folk戦略に倣う形で, わが国においては, 2022年に「みどりの食料システム戦略」が中長期的な戦略目標として施行され, 有機農業の面的拡大が目標として掲げられるなど, 持続可能性に配慮した農業の拡大が打ち出されている。みどりの食料システム戦略を含め, 産官学問わず様々な側面から, 国内においてもSDGs達成への取組の推進が図られている。

農業や食品産業といった食料供給に関わるステークホルダーにより構成されたシステム, すなわちフードシステムは, 農業といった生産活動をはじめとして, 生産に伴う環境負荷は不可分なものである。したがって, フードシステムによる食料の生産や消費を通して, 気候変動をはじめとした地球環境, 栄養状態といった人々の健康など, 社会全体の持続可能性に対して大きな影響を与えられている。すなわち, フードシステムは, 持続可能な開発に取り組んでいくことが特に必要とされる分野であると言える。更に, 世界規模での人口は依然として増加傾向にあり, 増加した人口をまかなうだけの食料生産が将来的にも必要となってくることを考慮すると, 持続可能な食料生産システムの構築は急務であると考えられる。食料の増産と共に, 持続可能な食料生産システムの構築に向け, 食品産業はもちろん, 消費者も含めた食に関わるステークホルダー全体による取り組みが求められる状況にある。

そうした背景から, フードシステムの持続可能性について論じた先行研究は多方面において蓄積が進んでいる。Willett et al. (2019) においては, 今後の世界的な人口増加が見込まれる中で, 持続可能ではないフードシステムのリスク, そして健康的な食生活へと移行することの重要性などが論じられている。健康的な食環境の構築においては, 飢餓や栄養不足の問題だけでなく, 成人病といった現在の豊かな食生活を主要因とした慢性疾患につ

いても問題として捉えられている。欧米といった先進国を中心として、高脂肪・高タンパクな食事による肥満人口の増加が社会問題の一つとなっている。肥満のみならず、食環境に起因する様々な健康問題が存在しており、したがって健康的な食環境の構築についても急務であると考えられる。

そうした持続可能な食生活への移行についても先行研究において広く論じられており、Dwivedi et al. (2017) は植物由来の食品中心の食生活への変革が持続可能なフードシステムにとって重要であることを論じており、また、遠山・安達 (2020) においては、エシカルな食行動の重要性が提唱されている。

持続可能なフードシステムに関する議論として、技術的な側面については、Herrero et al (2020) においては、現在の技術革新がもたらすフードシステム変革の可能性について論じている。また、Godfray et al. (2010) は文献レビューを通して、現在の技術開発により対応を進めることと今後の技術開発について進めることを結論づけている。同様に文献レビューを行った Caron, et al. (2018) においては、フードシステムの変革が必要であることを複数の観点から論じている。また、Fanzo et al. (2021) では、気候変動などの問題に対応するためにフードシステムを変革する必要性が論じられている。

更に、持続可能な食環境を構築する上でのステークホルダーの役割に着目した先行研究としては、Clark et al. (2020) において、公共部門や市民とったステークホルダーの協調的な行動の必要性が指摘されている。他にも、Reisch et al. (2017) においては、ナッジ(nudges)などを含めた政策的なアプローチの必要性についても論じられている。また、市民参加の重要性については Feenstra (2001) においても論じられている通りであり、したがって、企業サイドだけでなく、個人個人の行動が重要であることが示されている。

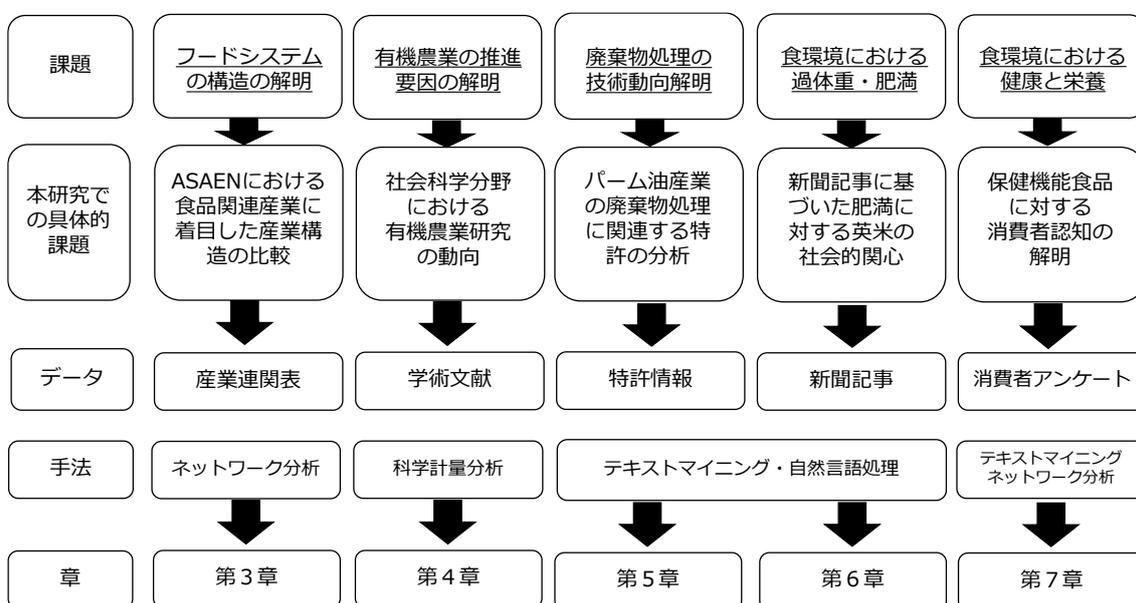
以上の先行研究を踏まえて、本論文では、持続可能なフードシステム構築の上で重要な要素として考えられる2つの主要課題を取り上げることとした。すなわち、食品産業を中心とした技術開発に関する課題、そして、消費者を含めた多様なステークホルダーから構成される食環境について健康的な食環境の構築という課題の、2つの大きな課題である。

そして、具体的には、本研究では、食品産業と食環境に関わる下記の5つの課題を取り上げる。まず、食品産業に関わる課題として、全体の産業構造を考慮した上でのフードシステムに関連する産業の立ち位置の解明に向けた課題、有機農業の学術研究の動向の解明に関する課題、国際社会において多数の問題が指摘されているパーム油産業の技術開発動向の解明に向けた課題、以上の3つの課題を取り扱う。次に、食環境に関連する課題として、肥満に関する社会的関心の動向の解明に向けた課題、保健機能食品に対する消費者認知の解明に関する課題、以上の2つの課題を取り扱う。

そして、こうした研究課題についてアプローチする手法として、本論文ではデータマイニングによるアプローチを適用することで課題に接近した。

## 1.2 本論文の構成

本節では、本論文の構成、そして、各章の研究課題とその概要について整理する。本論文は、研究の背景と課題、そして論文構成を整理した序論である本章を含め、全8章により構成される。続く第2章においては、本論文で用いたデータマイニングの各手法について概説し、第3章から第7章においては、先述の食品産業や食環境に関する研究課題に対応した各研究事例を取り上げた。そして、最終章となる第8章において、それまでの各章の内容を踏まえた上で本論文の結論をまとめ、更に今後の課題についても言及した。以上の本論文の構成について、特に第3章から第7章の概要を中心に図1-1に整理した。



出所) 筆者作成。

図1-1 本論文の構成

次に、各章において設定した研究課題とその概要について整理する。まず、第1章では、序論として、本論文が取り扱う持続可能性に関する研究課題について論じた。食品産業からの接近と、消費者を含めた食環境からの接近など、多方面から持続可能性についてアプローチする必要があることを、先行研究を引用しつつ論じた。そして、第2章は、本論文で用いたデータマイニング手法について、その概要を整理した。特に、ネットワーク分析、テキストマイニング及び自然言語処理、そして科学計量分析についてそれぞれ手法の説明などをまとめた。

そして、第3章から第7章は本研究がそれぞれ設定した研究課題に対する分析事例により構成される。第3章においては、持続可能な食品関連産業の発展を展望するため、経済

発展が著しく、今後の中長期的な市場拡大が予想される国々がメンバー国を構成している ASEAN を取り上げ、ASEAN の産業構造について分析を図る。分析対象のデータとしては OECD のデータベースから取得可能な各国の国家産業連関表を利用し、データマイニングアプローチの一つであるネットワーク分析を適用した。産業連関表に対する分析アプローチとしてはレオンティエフにより発案された産業連関分析が独自に発展してきたが、近年ではネットワーク科学の適用がみられるようになっており、様々な研究事例によりその有効性が示される状況にある。その点を考慮し、第 3 章においてはネットワーク分析を用いることとした。分析の枠組みにおいては、特に食品関連企業に焦点を当てるため、フードシステム関連セクターとして 3 つの産業セクターを取り上げ、それらの産業セクターに対する分析を中心にアプローチを組み立てた。ネットワーク分析で活用される指標としては中心性が存在しているが、産業連関表に対する分析に適した中心性として、Random Walk Centrality および Counting Betweenness という中心性が提案されている。Random Walk Centrality は産業ネットワークにおける経済的効果の波及の際に影響を受ける早さについて測定可能な指標であり、Counting Betweenness は経済的効果波及の際に、産業ネットワーク内における効果の媒介性の高さを測ることの可能な指標である。これらの 2 つの中心性を用いて、2005 年時点と 2015 年時点での産業構造について、定量的に分析した。更に、ASEAN 各国について、国同士の産業の相関性についても分析を行うことで、各国の産業構造の類似度についても考察を深めた。

第 4 章は、持続可能な農業生産システムを構築する上で重要となり得る有機農業に関する研究課題を取り上げる。有機農業は、わが国においてもみどりの食料システム戦略で拡大方針が打ち出された通り、持続可能な農業を実現する上での重要な取り組みとして捉えられる。国際的にも有機農業は重要視されており、そうした状況を受けて、有機農業に関連する学術文献は近年ますます蓄積されるようになってきている。こうした大量の学術文献データに対しては、従来の人力による研究レビューでは全体の傾向を把握するのは困難であることから、データマイニングアプローチの適用が有効であると考えられる。第 4 章では、有機農業を取り扱った学術研究の中でも、特に、社会科学の動向を分析対象とした。分析対象は、学術論文データベースである Web of Science に収録されており、なおかつ Social Science Citation Index が付与されている学術文献とした。文献の収集対象期間としては、近年の動向について接近するため、2002 年～2021 年の間に出版された 20 年間分とした。そして、分析手法については、学術文献に対する計量的分析手法である科学計量分析を用いた。そして、分析ツールとして、科学計量分析ツール Citespace を利用し、Citespace に実装された分析の中でも、特に、論文のキーワードを対象としたキーワードの共起分析、そして、ある時期に頻出しているキーワードを検出可能な手法であるバースト検出 (Burst Detection) を用いた。これらの分析手法を用いることにより、学術研究におけるヒットピックや、研究トレンドの変遷について接近する。

そして、第5章では、パーム油産業の持続可能な開発に関する状況について、特許データにもとづいて分析を試みた。2023年現在、パーム油はインドネシアとマレーシアが主要な生産国であるが、熱帯雨林の破壊、廃棄物、児童労働などの問題が国際的に指摘される状況にあり、環境負荷低減の観点から、廃棄物処理の技術開発についても求められている。したがって、生産過程で生じる廃棄物は、パーム油産業が持続可能な生産システムを構築する上で大きな課題となっているものと言える。そこで、パーム油生産過程廃棄物の処理技術に関係する特許情報について、ニューラルトピックモデルによる意味的側面を踏まえた分析を行うことで、これまでの廃棄物処理に関する技術の変遷について、技術的内容を踏まえた観点から概観することを試みた。分析対象として、WIPO (World Intellectual Property Organization) により収集されたパーム油関連特許に関するデータセットを利用し、データセットの内、研究の目的に合わせて、特に廃棄物処理技術に分類される特許情報を分析の対象とした。

第6章においては、持続可能で健康的な食環境を構築する上で、今日における重大な課題となっている肥満 (Obesity) を取り上げる。肥満は、食と健康に関する重要な課題であり、米国や英国において特に深刻な社会問題とされる。第6章では、そうした肥満に関する内容を含んでいる新聞記事テキストを対象にテキストマイニングを適用し、データマイニングによる分析を試みた。分析対象とした新聞は、質の高い記事を掲載しているとされる新聞、すなわちクオリティペーパーとして国際的に認知されている The New York Times と The Guardian である。The New York Times は米国の新聞であり、The Guardian は英国の新聞であるため、これらの新聞を通して、英国と米国の社会的関心に接近することが可能であると考えられる。そして、ニューラルトピックモデルによる分析の結果、2つの新聞において共通するトピックと各々に特有のトピックが抽出され、更に、1999年から2020年の21年間にわたる記事の関心の推移についても明らかにした。ここでの分析対象は米国および英国における新聞社として両国1社ずつ取り上げており、分析結果は必ずしも両国全体の世相を反映したものと捉えることは出来ないが、両国における社会的関心の一端を明らかにした結果と考えることが出来るであろう。

第7章は、健康効果を有するとされる食料品に対する認証制度である保健機能食品制度を対象として、本制度に対する消費者の認知を明らかにすることを目的に、ネットワーク分析によるテキストマイニングを適用した研究事例である。そして、第7章においては、保健機能食品に対する消費者認知に関して、2つの課題を設定した。第1の課題として、保健機能食品を構成する特定保健用食品、機能性表示食品、栄養機能食品という3つの認証制度について、これらの制度自体に対する消費者認知の比較を行った。消費者認知の比較にあたっては、ネットワーク分析と共に、テキストに表れている書き手の感情を推定する手法である感情分析を併用し、アプローチを試みた。第2の課題は、特定保健用食品と機能性表示食品について、これらの認証食品が有する健康効果とそれら認証制度の関係性

の解明である。分析手法として、ネットワーク分析の中でも、特定のノードに着目するエゴネットワークの枠組みを利用した。いずれの課題も分析対象データは消費者へのアンケート調査により得られた自由回答テキストである。これら2つの課題を通して、保健機能食品に関する消費者認知について、制度自体に対する印象や、各制度と認証製品が有する健康効果との認知の関係性などを対象に、感情分析やネットワーク分析等を駆使して、多角的に接近することを試みた。

そして、本論文の最終章である第8章においては、以上の第3章から第7章の研究事例の結果について要約した上で、各章で適用したデータマイニング手法の特徴に関する考察を深めると共に、それらの研究事例から得られた本論文の結論について述べる。最後に、今後の研究の展望、残された研究課題についても概観する。

## 第2章 本論文で用いるデータマイニング手法の概要

### 2.1 データマイニングの概要

本章では、データマイニングによるアプローチの有効性について論じると共に、本論文の各章において用いているデータマイニングの分析手法について、整理する。特に、様々なものが存在するデータマイニング手法の中でも、本論文で用いた手法である、ネットワーク分析、自然言語処理およびテキストマイニング、科学計量分析といった手法に焦点を当てて概説する。

まず、データマイニング (Data Mining) とは、データに潜む法則や知見などに対して、様々なデータ分析手法を活用することにより接近する手法である。学習データセットに基づいたデータの分類やパターンや規則性の発見、そして新規データセットに対する予測に関するタスクがデータマイニングとして挙げられる。活用される代表的手法としては機械学習、深層学習、統計的手法等が存在しており、近年は特に深層学習の技術開発が著しく、データマイニングの活用はますます重要性を高めている。近年はビッグデータの時代と呼ばれるようになり久しいが、そうしたデータの蓄積もあいまってデータは一つの資本 (Capital) として捉えられるようになってきている (Sadowski, 2019)。

データマイニング手法の適用にあたっては、分析対象となるデータセットが必要となるが、そうしたデータセットや各種データベースについても拡充が進んでいる。つまり、データマイニング手法の発展と並び、各種データベースの整備や、多様なデータセットの提供も幅広く行われるようになってきていることも、近年のデータマイニング研究の加速化の要因と言える。そうしたデータベースからの取得が可能となったデータセットとして、本論文と関係が深いものでは、特許情報や新聞記事、そして学術文献などが挙げられる。

例えば、特許情報のデータベースについては、World Intellectual Property Organization (WIPO) による PatentScope (註1) が代表的な例として挙げられ、世界各国の特許情報にアクセスすることが出来る。

また、新聞記事についてもデータベースとしては、The New York Times や The Guardian といったクオリティペーパーと呼ばれる新聞を発刊している新聞社がデータを提供している。新聞は社会の様相を記録した媒体であるため、データベースの整備により、社会の動向等について分析することが可能となった。

学術文献データについては、Web of Science や Scopus といった大規模論文データベースから、要旨 (Abstract) や論文に付与されたキーワード、著者に関する情報等が取得できるようになっている。

また、2020年初頭に世界的流行となり現在に至るまで世界に多大な影響を与えている COVID-19 については、世界的流行の最中、比較的早い段階で COVID-19 に関するデー

タセットが一般に公開され、国際的な規模で研究の加速化を後押しした。COVID-19 の例にも示される通り、データセットを広く公開することは研究の加速化に繋がり、研究を進める上で意義深いことであると言える。

こうしたデータを基軸とした研究は、データ駆動型 (Data-Driven) アプローチとも呼ばれ、現在、様々な研究分野で注目を浴びる。各分野におけるデータ駆動型研究の動向について概観すると、例えば、材料科学研究においてはマテリアルズインフォマティクス (Materials Informatics) (Wagner and Rondinelli, 2016)、創薬研究においては AI 創薬 (Data-Driven Drug Discovery) (Chen et al., 2018) といった新たな研究分野が勃興してきている。

更に、社会科学に関連する分野についてもデータ駆動型アプローチによる研究が進展しており、心理学におけるデータ駆動型アプローチの活用 (Rachael et al, 2018) や、データ駆動型アプローチによる新たな社会科学分野である計算社会科学 (Computational Social Science) が創出されている。他にも、ソーシャルメディアの台頭による大量のユーザー生成データ (User-Generated Data : UGD) を軸としたデータ駆動型マーケティング (Saura et al, 2021) も、同じく新たな研究動向として挙げられる。

いずれの研究分野においても、データに基づいたパターン認識や予測などによるファクトファイディングが可能となった点がデータマイニング適用の特徴であり、利点であると言える。今後は、各研究分野において、こうしたデータマイニングによるアプローチに基づいた新たな学術研究の進展がますます期待される。

## 2.2 ネットワーク分析の概要

ネットワーク分析は、ノードとリンクから成るネットワーク形式のデータに対する分析手法である、ノードはネットワークにおける構成要素を指しており、リンクはその構成要素を繋ぐ関係性について指し示している。ノードとリンクとして捉えられるデータには広く適用可能なものであり、そういったネットワーク形式のデータは非常に多岐にわたり存在していることから、ネットワーク分析の適用は非常に幅が広い。

手法の発展の経緯について概観すると、ネットワーク分析は、当初は社会学で用いられてきた手法であった。理論物理学といった分野からの理論的なアプローチが盛んになり、グラフ理論としての整備が進んだ結果、学際的な研究分野として発展を遂げ、現在では様々な分野で利用される手法となっている。

ネットワーク分析に特徴的な手法として、中心性の算出が挙げられる。中心性とは、ネットワークにおけるノードの重要度などを測る指標であり、例えば次数中心性 (Degree Centrality) は、そのノードが有するリンク数によって算出される。固有ベクトル中心性 (Eigenvector Centrality)、媒介中心性 (Betweenness Centrality) といった様々な中心性が提案されており、分析目的に応じて使い分けられている (註2)。

もう一つの代表的なネットワーク分析の手法が、コミュニティの検出である。コミュニティとは、ネットワークにおいて、緊密な関係性、すなわち緊密なリンクの繋がりを持つノードのグループを指す。そうしたコミュニティを検出する際に用いられる代表的な手法がモジュラリティ (Modularity) の算出によるコミュニティ検出である。モジュラリティは、Newman(2006)の提案した手法であり、コミュニティ分割の際の質を定量的に評価した指標となる。すなわち、モジュラリティはどれだけ適切なグルーピングとなっているかの評価指標と言える (註3)。

こうした様々な分析指標の適用が可能となるネットワーク分析は、社会科学分野においても幅広く利用されている。経済学に関するデータへの適用事例の一つとして、産業連関表に対する適用が挙げられるであろう。産業連関表はマトリックスデータであるためネットワーク分析の適用が可能であり、産業連関表をネットワーク形式、あるいはグラフ形式のデータとして捉えるデータマイニングのアプローチが試みられるようになった。特に、国際的な産業構造のリンケージをネットワークの枠組みで捉えることを目的として、産業連関表が分析対象に利用される状況となっている。本論文では第3章において産業連関表を対象としたネットワーク分析を行っている。

他にも、食品科学に関する研究事例としては、Ahn et al. (2011)による食材と香り (Flavor) の関係性についてネットワーク分析を試みた研究事例が挙げられる。このように、非常に多岐にわたるトピックに対してネットワーク分析は適用されている。

ネットワーク分析のツールは様々なものが提供されており、多彩な分析手法が実装されている Gephi, UCINET, Pajek, Cytoscape といったネットワーク分析の専用ツールや、その他にも NetworkX といった Python のライブラリなどが利用可能となっている。

## 2.3 テキストマイニングと自然言語処理の概要

データマイニングの中でも、テキストを対象としたものはとりわけテキストマイニング (Text mining) と呼ばれる。テキストマイニングは、テキストデータから有用な情報を抽出することを目的としたテキスト分析手法である。自然言語処理は、人間が用いる日本語や英語などの言語のコンピュータによる処理であり、テキストマイニングとは不可分の関係にあると言える。

自然言語処理はこれまで統計的な言語処理手法などの開発により発展を遂げてきたが、画像処理研究などと同様に、自然言語処理についても深層学習の適用によって大きく技術が進展しており、様々な言語処理の高度化・精緻化が図られる状況にある。特に、Mikolov, et al. (2012) により開発された Word2vec によって単語のアナロジーが可能であることが示されたことを契機に、深層学習の適用によって得られる分散表現 (Distributed representations) と呼ばれる単語や文章といったテキストのベクトルについて、意味的な演算が可能である

こと、すなわち、人間と同様のバイアスを反映していることが示され、現在ではこうした自然言語処理技術の社会科学研究への応用が進みつつある。分散表現は単語埋め込みとも呼ばれ、通常、200~300次元程度のベクトルとして表現される。先述の通り、深層学習（ニューラルネットワーク）の枠組みを利用してベクトルを獲得することが一般的であり、単語や文章ごとにベクトルを算出する。一方、分散表現とは異なる枠組みに局所（One-hot）表現があり、局所表現は、1要素のみが1となるスパースなベクトルを各々の単語に対応させる。したがって、例えば、異なる単語間の関係について、局所表現では同一であるかどうかしか識別できないのに対して、分散表現では異なる単語間の類似度を測定できる。

社会科学への自然言語処理の応用研究としては、過去100年に及ぶ書籍のテキストから単語の分散表現を獲得し、ジェンダーや民族に関するステレオタイプやバイアスを定量的に分析した研究事例（Garg et al., 2018）や、社会階層を分散表現により定量化した研究事例（Kozlowski et al., 2019）が挙げられる。他にも、文章のトピック推定に活用される統計的モデルであるトピックモデルについても、Neural Topic Model という形で、深層学習を利用したトピックモデルが開発され、分析の高度化と精緻化が進む状況にある。

こうした分析の高度化が進む中、テキストデータは様々な場面で存在していることから、テキストマイニングが多様な分野で活用されており、分析の適用事例が広がっている状況にある。分析対象となるテキストデータの例を挙げると、消費者調査で得られた自由回答テキストや、学術文献に記載されている専門性の高いテキスト、新聞記事といった時事性の高いテキスト、そして、近年急速に普及し日々大量のテキストデータを産出しているソーシャルメディア・SNS（Social Networking Service）における投稿テキストなど、多岐にわたっている。

更に、テキストマイニングは、計量政治学といったような形で、社会科学系分野においてもその分析手法の高度化が図られてきた。計量政治学は政治を分析対象としたテキストマイニングについても取り扱っており、自然言語処理技術を応用した社会科学の研究事例の蓄積が進む分野である。計量政治学向けのテキストマイニングツールとしては *quanteda*（Kenneth et al., 2018）が開発されている。

また、日本語テキストに対するテキストマイニング・自然言語処理の技術についても開発が進んでいる。日本語には文章の区切りがないため、文章を単語へと分割する処理が必要となる。そのため、日本語特有の自然言語処理の手法として、形態素解析と呼ばれる分かち書き（tokenization）処理に活用可能な解析手法が開発されている。日本語の分析が可能なテキストマイニングツールとしてはKHコーダー（樋口, 2020）が挙げられる。KHコーダーは共起ネットワーク分析やトピックモデルといった様々な分析手法が実装されており、人文社会科学研究を中心に広く利用されている。

以上、自然言語処理分野の発展は日進月歩であり、今後も更なる技術開発の進展が見込まれる状況にあると言える。

## 2.4 科学計量分析の概要

科学計量分析は、科学計量学や計量書誌学といった分野で活用される学術文献などに対する分析手法を指したものである。科学計量分析の代表的な指標の一つとして、学術誌の重要性を測る指標であるインパクトファクターが挙げられる。そうしたインパクトファクターをはじめとして、学術に関連する指標を定量的に取り扱う分析である。

藤垣ら（2004）を参照すると、科学計量学の方法として、引用分析（インパクトファクター、共引用分析、書誌結合分析）、共著分析、謝辞分析、共語分析、共分類分析といった分析手法が挙げられている。こうした各種分析を概観するとわかる通り、科学計量分析は、学術活動に関する定量化という側面と共に、学術文献を対象とした知識抽出、すなわち学術文献に対するデータマイニングという性格も持つ。近年の学術文献の蓄積、すなわち出版数の増加は目覚ましいものがあり、学術文献を対象としたデータマイニングによるアプローチはますます重要となってきたと言える。

科学計量分析の事例としては、文献に付与されたキーワードについて、同じ論文に付与されたキーワードを共起関係にあるとみなし、先述したネットワーク分析の枠組みを適用するといった分析が挙げられる。同様に、共著関係に対するネットワーク分析や、学術文献の引用・被引用関係に対するネットワーク分析も代表的な科学計量分析の一つである。このように、科学計量分析は様々な分野の手法を科学計量学の分析目的に対して適用するものであり、今後も新たな研究手法が科学計量分析に位置付けられ発展していくことが予想される。

近年は学術文献のデータベースの整備も進む状況にあるが、プレプリントと呼ばれる査読前の論文をオープンアクセスの論文として公開するシステムについても整備されてきており、その点も注目に値する。代表的なプレプリントサーバーとしては arXiv が挙げられるが、国内においても、プレプリントサーバーとして Jxiv が 2022 年に開始し、日本語の文献も含め、ますます学術文献の蓄積が進むことが期待されている。

科学計量分析のためのツールとしては、Citespace (Chen, 2006) や bibliometrix (Aria and Cuccurullo, 2017) (註 4) といったツールが挙げられ、様々な学術トピックを対象として、幅広く活用されている。

## 2.5 本論文における実装ツール

以上、概説した各手法について、本論文で用いている実装ツールについて整理する。まず、第 3 章と第 7 章においては、ネットワーク分析を活用しているため、ネットワーク分析ツール Gephi を利用した。第 4 章では、科学計量分析ツール Citespace を利用した。第 5 章と第 6 章においては、Python のライブラリとして提供されている BERTopic をツールと

して用いた。いずれのツールについても、フリーソフトウェアとして提供されており、利用が可能である。

註 1) Patent Scope については、<https://patentscope.wipo.int/> (2023 年 2 月アクセス) を参照。

註 2) ネットワークと中心性の関係、また中心性同士の関係の特徴については、Oldham et al. (2019) が詳しい。

註 3) モジュラリティの値は  $0 \leq (\text{Modularity}) \leq 1,0$  であり、1.0 に近いほどコミュニティの分割の質が良いと捉えられる。

註 4) `quanteda` および `bibliometrix` は統計解析環境 R のパッケージとして提供されている。

## 第3章 食品関連産業に着目した ASEAN 各国の産業構造ネットワーク分析

### 3.1 はじめに

第1章で整理した食品産業の持続可能な発展に向けた研究事例として、本章では、国家産業連関表を対象としたネットワーク分析を行う。本分析により、国全体の産業構造の中における食品関連産業の影響力などについて、国家ごとの比較を定量的に行うことが可能となる。本分析を通して、各国の現状について明らかにする。そして、本章では分析対象として、近年目覚ましい経済発展を遂げている経済圏である ASEAN 各国に焦点を当てる。

本章で分析対象として取り上げる産業連関表 (Input-Output Table) は、産業構造の把握に適したデータの一つとして知られている。ワシリー・レオンチェフにより提案されて以来、産業間の取引を記録するデータとして、様々な分析に広く用いられてきた。従来的には、産業連関表に対する分析手法として発展してきた産業連関分析と呼ばれるアプローチの適用が一般的であるが、近年では、ネットワーク科学分野の発展に伴い、そうした一般的なアプローチとは異なる新たな手法として、第2章において紹介した通り、ネットワーク分析の枠組みが応用されている状況にある。産業連関表は、産業間の中間取引額、付加価値、最終需要が記録されたデータである。この中間取引はマトリックスを形成しており、グラフ理論における隣接行列として取り扱うことが出来る。そうした点から、産業連関表にネットワーク分析の枠組みが適用可能であり、多くの研究事例が蓄積されている。

こういった状況の中でも、特に、中心性といったネットワーク指標を用いて各国における主要セクター (Key Sector) を特定する研究が盛んに取り組まれており、Tsekeris (2017) と García and Ramos (2015) は、ギリシャにおける主要セクターの特定について、複数の中心性指標により分析し、後者は産業ごとの影響力の計測も試みている。Gurgul and Lach (2018) は、ポーランドを対象に、また Go et al. (2019) はフィリピンを対象として、それぞれ主要産業セクターの変遷を分析している。これらの研究事例はいずれも、一国を対象としたものであり、複数国を対象とした分析は行われていない。

一方、Blöchl et al. (2011) は、複数の国を対象とする分析として、OECD 加盟国を中心とした産業構造について、中心性指標を用いた比較を行っている。この研究は、産業構造の国家間比較に際して、ネットワーク指標によるアプローチの有効性を示したものであるが、本章で扱うこととなる ASEAN 各国の多くは、分析の対象外となっている。分析対象として複数の国々を扱った別の研究事例として、経済的ショック伝播に関して産業連関表に基づいた国家間比較を行った Contreras and Fagiolo (2014) が存在しているが、分析の対象はやはり EU となっている。

以上の背景を踏まえ、本章では、世界経済におけるその重要性にも関わらず、こうした

ネットワーク分析を用いたアプローチによる産業構造の比較に関して、十分な研究事例が蓄積されていない ASEAN 各国を分析の対象とした。更に、上述の通り、ネットワーク指標を用いた多くの先行研究は、産業構造の中で中心的な役割を果たす主要セクターの特定を目的としているが、本分析では、ネットワーク分析のアプローチを主要セクターの抽出のみに用いるのではなく、フードシステムに関連すると考えられるセクターについても着目し、フードシステム関連セクターから見た ASEAN 各国の産業構造の違いを明らかにしつつ、各国間の相関性についても分析を進めた。本アプローチにより、グループとしての ASEAN 各国における産業構造の一側面を明らかにする。

## 3.2 方法

### 3.2.1 データ

本章の分析データとして、OECD が提供する STAN データベースより、2018 年版 Input-Output Tables から入手可能な国家単位の産業連関表を用いて分析を行った。分析対象は、ASEAN 加盟国のうち、OECD のデータベースにおいて提供されていないラオスとミャンマーを除く八カ国、すなわち、ブルネイ、インドネシア、カンボジア、マレーシア、フィリピン、シンガポール、タイ、ベトナムである。表 3-1 に、STAN データベースより取得可能な産業連関表に記録されている 36 セクターについて記載した。なお、産業連関表に記載されている 36 セクターの内の一つである「Private households with employed persons」については、今回の分析対象となつたいずれの国においても中間取引が存在していないため、本章での分析対象は、実質的には「Private households with employed persons」を除く 35 セクターとなっている。

そして、分析対象となる年度は、データが入手可能な範囲の中で最も過去の年度である 2005 年と、入手可能な範囲の中での最新年度であり、なおかつ、リーマンショックを経過した後で経済的環境や経済的状況が再び落ち着きを取り戻したと考えられる 2015 年とした（註 1）。この期間、すなわち 1995 年から 2015 年の中での ASEAN のメンバー国の経済発展の状況について分析を図ることとする。

更に、本章では、食品関連産業に分析の焦点を当てるため、フードシステムに関連すると考えられる「Agriculture, forestry and fishing」、「Food products, beverages and tobacco」、「Accommodation and food services」という 3 つの産業セクターに焦点を当てる。これらの産業セクターをフードシステム関連セクターとして位置付け、分析の中でクローズアップする。本分析では、特に、中心性指標に基づいたメンバー国間の比較を行う。なお、本分析で利用する産業連関表に記録されている中間取引額は、US ドル評価に統一されているため、ASEAN メンバー国間の比較が可能である。

表 3-1 本分析で用いた産業連関表におけるセクター一覧

---

1.Agriculture, forestry and fishing
2.Mining and extraction of energy producing products
3.Mining and quarrying of non-energy producing products
4.Mining support service activities
5.Food products, beverages and tobacco
6.Textiles, wearing apparel, leather and related products
7.Wood and products of wood and cork
8.Paper products and printing
9.Coke and refined petroleum products
10.Chemicals and pharmaceutical products
11.Rubber and plastic products
12.Other non-metallic mineral products
13.Basic metals
14.Fabricated metal products
15.Computer, electronic and optical products
16.Electrical equipment
17.Machinery and equipment, nec
18.Motor vehicles, trailers and semi-trailers
19.Other transport equipment
20.Other manufacturing; repair and installation of machinery and equipment
21.Electricity, gas, water supply, sewerage, waste and remediation services
22.Construction
23.Wholesale and retail trade; repair of motor vehicles
24.Transportation and storage
25.Accommodation and food services
26.Publishing, audiovisual and broadcasting activities
27.Telecommunications
28.IT and other information services
29.Financial and insurance activities
30.Real estate activities
31.Other business sector services
32.Public admin. and defence; compulsory social security
33.Education
34.Human health and social work
35.Arts, entertainment, recreation and other service activities
36.Private households with employed persons

---

出所) OECD データベースより引用して筆者作成。

註：1) 「Private households with employed persons」はいずれの国においても中間取引が存在していないため分析の対象外となる。

### 3.2.2 産業連関表とネットワーク分析

産業連関表を対象としたネットワーク分析の際には、各産業セクターをグラフ理論におけるノードとしてみなし、更に、投入・産出関係をリンクとして捉えることとなる。それにより、産業連関表に基づいた産業ネットワークの分析が可能となる。そして、ここでの分析対象となる産業ネットワークは、ネットワークにおいて有向性をもっていること、重

み付きであること、自己ループ構造を有していること、そして、密なネットワークを形成している、という4つの特徴を持つ。すなわち、中間取引額の大きさは重みに相当し、自産業への投入は自己ループに当たることになる。そして、需要と投入という非対称的な関係はネットワークにおける有向性として捉えられる。本章の分析では、こうした4つの特徴を持つ産業ネットワークに適した分析を行うため、後述する2つの中心性、Random Walk Centrality と Counting Betweenness を分析指標として採用する。

第2章で述べた通り、中心性はネットワーク分析における評価指標の1つであり、分析目的に応じた様々な中心性が提案されている。具体的には、中心性はネットワーク内におけるノードの重要性等を測る指標であり、すなわち本研究においては各産業セクターの重要度を測る指標として活用できる。中心性としては、ノードのリンク数に着目した次数中心性や、重要なノードとのつながりにもとづきノードの中心性を測る固有ベクトル中心性など、様々な定義による指標がある。

そうした中でも、Blöchl et al. (2011) は、産業ネットワークの分析に適した中心性の指標として、Random Walk Centrality (以下 RWC) と Counting Betweenness (以下 CBET) を提案した。一般的な中心性指標との対応関係で見ると、RWC は近接中心性と類似の指標であり、CBET は媒介中心性と類似の指標となっている。更に、DePaolis et al. (2020) は、この RWC と CBET をそれぞれ Garcia et al. (2008) の提示した Immediate Effects と Mediative Effects とに対応させて再定義している。すなわち、RWC の高い産業は産業ネットワーク内における経済的効果の波及の際、早期に影響を受ける産業であると捉えられ、それに対して CBET の高い産業は、経済的効果の波及の際、ネットワーク内におけるその効果の媒介性が高いものと考えることが出来る。産業構造の解明を目的とする本分析においても、この2つの中心性指標を分析に採用する(註2)。

### 3.2.3 コミュニティの検出と各国の産業構造の相関性

本分析では、先述の中心性と共に、コミュニティの検出についても活用する。ネットワーク分析の枠組みの中でも、コミュニティの検出は広く用いられる手法である。特に、産業連関表に基づいた産業ネットワークに対しては、モジュラリティに基づいたコミュニティ分割が行われるケースが多い(Cerina et al., 2015 ; Domínguez and Mendez, 2019)。本分析においても、各国のフードシステム関連セクターの関係性を捉える上でこのモジュラリティに基づいた手法を用いた(註3)。

そして、本分析では、各国の産業構造全体を比較するために、RWC と CBET の算出結果に対して、2005 及び 2015 年度におけるスピアマンの順位相関係数を算出した。なお、ここでは「Private households with employed persons」も対象とした。これにより、各国の産業構造の違いや類似性を考察する(註4)。

### 3.3 結果と考察

#### 3.3.1 各国の主要セクターの分析結果

まず、各国の主要セクターを特定するため、RWC と CBET が最も高かったセクターについて表 3-2 および表 3-3 に整理した。本章においては、表 3-2・表 3-3 のセクターを主要セクターとして捉えることとする。

中心性の指標ごと、あるいは年度ごとに主要セクターに変化が無かった国は、ブルネイ、カンボジア、マレーシア、シンガポール、ベトナムであり、結果を参照すると、それぞれ各国の代表的な産業が反映される結果となったと言える。

ブルネイでは「Mining and extraction of energy producing products」が RWC 及び CBET のどちらにおいても主要セクターとして抽出され、いずれの年度においても変化はなかった。

カンボジアでは「Agriculture, forestry and fishing」が主要セクターとして抽出されており、カンボジアにおいては農林水産業が主要産業であることが本分析結果からも確認される。

更に、ベトナムにおいては「Food products, beverages and tobacco」が主要セクターとして抽出されており、食品関連産業が主要産業であることが示唆された。年度による変化があったのは、唯一、インドネシアの CBET から抽出された主要セクターのみであり、両年度間の主要産業はほぼ変化が無かったと言える。

表 3-2 Random Walk Centrality からみた各国における主要セクター一覧

国	主要セクター (Random Walk Centrality)	
	2005	2015
BRN	Mining and extraction of energy producing products	Mining and extraction of energy producing products
IDN	Construction	Construction
KHM	Agriculture, forestry and fishing	Agriculture, forestry and fishing
MYS	Computer, electronic and optical products	Computer, electronic and optical products
PHL	Wholesale and retail trade; repair of motor vehicles	Wholesale and retail trade; repair of motor vehicles
SGP	Computer, electronic and optical products	Computer, electronic and optical products
THA	Wholesale and retail trade; repair of motor vehicles	Wholesale and retail trade; repair of motor vehicles
VNM	Food products, beverages and tobacco	Food products, beverages and tobacco

出所) 分析結果より筆者作成。

註: 1) BRN はブルネイ, IDN はインドネシア, KHM はカンボジア, MYS はマレーシア, PHL はフィリピン, SGP はシンガポール, THA はタイ, VNM はベトナムを示す。

表 3-3 Counting Betweenness からみた各国における主要セクター一覧

国	主要セクター (Counting Betweenness)	
	2005	2015
BRN	Mining and extraction of energy producing products	Mining and extraction of energy producing products
IDN	Textiles, wearing apparel, leather and related products	Construction
KHM	Agriculture, forestry and fishing	Agriculture, forestry and fishing
MYS	Computer, electronic and optical products	Computer, electronic and optical products
PHL	Food products, beverages and tobacco	Food products, beverages and tobacco
SGP	Computer, electronic and optical products	Computer, electronic and optical products
THA	Motor vehicles, trailers and semi-trailers	Motor vehicles, trailers and semi-trailers
VNM	Food products, beverages and tobacco	Food products, beverages and tobacco

出所) 分析結果より筆者作成.

註: 1) BRN はブルネイ, IDN はインドネシア, KHM はカンボジア, MYA はマレーシア, PHL はフィリピン, SGP はシンガポール, THA はタイ, VNM はベトナムを示す.

### 3.3.2 各国のフードシステム関連セクターに関する分析結果

表 3-4 に, フードシステム関連セクターの中心性の順位について整理した. 概して, フードシステム関連セクターの順位については RWC と CBET のいずれも似通った順位の傾向を示していると言える結果となった.

更に, 国別に順位を見ると, カンボジア, ベトナム, フィリピンはフードシステム関連セクターが特に高順位となった. それに対し, 工業中心の都市国家であるシンガポールと, 石油や天然ガスなどのエネルギー資源が主要産業であるブルネイの二カ国は低順位となった. これら五カ国以外のインドネシア, マレーシア, タイは, フードシステム関連セクターは中順位として捉えられる.

したがって, フードシステム関連セクターの順位に基づくと, ASEAN の八カ国は, カンボジア・ベトナム・フィリピンの高順位グループ, インドネシア・マレーシア・タイの中順位グループ, そして, ブルネイとシンガポールの低順位グループの三つに分類出来る.

表 3-5 には, コミュニティ検出の結果を含めたネットワーク指標の分析結果について整理した (註 5). コミュニティ検出の結果は, 各国の産業構造におけるフードシステム関連セクター同士の相対的な関係性の強さを表すものと捉えることが出来る.

表 3-4 各国の FS 関連セクターの中心性順

国		各中心性の順位					
		Agriculture, forestry and fishing		Food products, beverages and tobacco		Accommodation and food services	
		2005	2015	2005	2015	2005	2015
BRN	RWC	25	21	18	14	23	5
	CBET	23	20	17	12	25	23
IDN	RWC	8	7	5	5	6	4
	CBET	10	8	3	3	9	6
KHM	RWC	1	1	2	2	3	3
	CBET	1	1	2	2	3	4
MYS	RWC	17	10	10	3	15	11
	CBET	13	8	4	2	18	15
PHL	RWC	3	3	2	2	7	7
	CBET	2	2	1	1	9	11
SCP	RWC	32	32	23	22	17	16
	CBET	32	32	23	20	22	17
THA	RWC	11	9	3	3	4	4
	CBET	10	9	4	4	7	8
VNM	RWC	2	2	1	1	3	3
	CBET	2	3	1	1	4	5

出所) 分析結果より筆者作成。

註：1) BRN はブルネイ，IDN はインドネシア，KHM はカンボジア，MYS はマレーシア，PHL はフィリピン，SGP はシンガポール，THA はタイ，VNM はベトナムを示す。

註：2) RWC は Random Walk Centrality，CBET は Counting Betweenness を示す。

A タイプは最もフードシステム関連セクター同士の関係性が強く，各々が別のコミュニティである D タイプは最も関係性が弱いものと考えられる。国ごとに結果を見ていくと，マレーシア，フィリピン，タイ，ベトナムが両年度で同じ CT となっている。これらは主要セクターの結果やフードシステム関連セクターの中心性の順位は各々異なるが，フードシステム関連セクター同士の関係性は類似していることが推察される。すなわち，「Accommodation and food services」以外のフードシステム関連セクターの繋がりが強いということが示唆された。なお，ネットワークの密度を示す指標である Average Clustering Coefficient を参照すると，ブルネイ以外の国々は 0.9 を超えた数値が算出されており，いずれにおいても密なネットワークが形成されていることが確認出来る。

表 3-5 ネットワーク指標に関する各国の分析結果

国	年度	Average Clustering Coefficient	Modularity	No. of Communities	Types of Communities
BRN	2005	0.759	0.153	4	A
	2015	0.771	0.207	4	A
IDN	2005	0.986	0.207	4	C
	2015	0.995	0.145	3	D
KHM	2005	0.764	0.169	3	D
	2015	0.872	0.212	3	B
MYS	2005	0.992	0.200	4	B
	2015	0.997	0.177	3	B
PHL	2005	0.980	0.119	3	B
	2015	0.985	0.131	3	B
SCP	2005	0.925	0.237	3	A
	2015	0.919	0.148	4	B
THA	2005	0.973	0.141	3	B
	2015	0.991	0.084	4	B
VNM	2005	0.925	0.159	3	B
	2015	0.970	0.161	4	B

出所) 分析結果より筆者作成.

註: 1) BRN はブルネイ, IDN はインドネシア, KHM はカンボジア, MYS はマレーシア, PHL はフィリピン, SGP はシンガポール, THA はタイ, VNM はベトナムを示す.

註: 2) Types of Communities はコミュニティのタイプを示す. 表中の A は全ての FS 関連セクターが同一コミュニティ, B は「Accommodation and food services」以外が同一のコミュニティ, C は「Agriculture, forestry and fishing」以外が同一のコミュニティ, D は全てのセクターが異なるコミュニティとなる.

### 3.3.3 国家間における産業構造の相関性の結果

スピアマンの順位相関係数により算出した相関性の結果について, RWC にもとづいた結果を表 3-6, CBET にもとづいた結果を表 3-7 に整理した. まず, 国家間で特に相関性が高かったのはフィリピンとカンボジアであり, いずれの値も 0.8 以上と高い相関が確認された. 次に, フィリピンとシンガポールについて見ると, フードシステム関連セクターの順位は全く異なるが, 2005 年時点での相関係数と 2015 年時点での相関係数を比較すると, 相関性が高まっていることを確認することが出来る. これは, フィリピンの産業構造がシンガポールと近くなっていることが示唆された結果であると考えることが出来る.

表 3-6 各年度の Random Walk Centrality に対するスピアマンの順位相関係数の結果

国	2005								2015								
	BRN	IDN	KHM	MYS	PHL	SGP	THA	VNM	BRN	IDN	KHM	MYS	PHL	SGP	THA	VNM	
2005	BRN	1.000	0.387	0.217	0.447	0.418	0.289	0.185	0.422	0.965	0.472	0.181	0.457	0.462	0.242	0.344	0.232
	IDN		1.000	0.579	0.508	0.683	0.474	0.708	0.641	0.489	0.916	0.731	0.583	0.698	0.471	0.750	0.578
	KHM			1.000	0.401	0.817	0.305	0.636	0.691	0.280	0.651	0.910	0.491	0.808	0.320	0.692	0.736
	MYS				1.000	0.582	0.451	0.528	0.577	0.549	0.597	0.464	0.966	0.602	0.410	0.584	0.490
	PHL					1.000	0.381	0.782	0.742	0.484	0.739	0.829	0.660	0.951	0.385	0.821	0.686
	SGP						1.000	0.470	0.182	0.301	0.391	0.382	0.471	0.479	0.968	0.458	0.137
	THA							1.000	0.651	0.268	0.620	0.703	0.598	0.731	0.497	0.936	0.660
	VNM								1.000	0.494	0.695	0.613	0.614	0.699	0.212	0.683	0.902
2015	BRN								1.000	0.553	0.276	0.569	0.529	0.257	0.425	0.289	
	IDN									1.000	0.681	0.647	0.716	0.382	0.753	0.643	
	KHM										1.000	0.555	0.820	0.407	0.726	0.653	
	MYS											1.000	0.686	0.444	0.660	0.541	
	PHL												1.000	0.474	0.766	0.597	
	SGP													1.000	0.477	0.196	
	THA														1.000	0.693	

出所) 分析結果より筆者作成.

註: 1) BRN はブルネイ, IDN はインドネシア, KHM はカンボジア, MY S はマレーシア, PHL はフィリピン, SGP はシンガポール, THA はタイ, VNM はベトナムを示す.

註: 2) 中心性の値は少数第 4 位を切り上げ.

表 3-7 各年度の Counting Betweenness に対するスピアマンの順位相関係数の結果

国	2005								2015								
	BRN	IDN	KHM	MYS	PHL	SGP	THA	VNM	BRN	IDN	KHM	MYS	PHL	SGP	THA	VNM	
2005	BRN	1.000	0.464	0.240	0.531	0.447	0.333	0.193	0.437	0.929	0.544	0.212	0.529	0.477	0.300	0.336	0.307
	IDN		1.000	0.657	0.669	0.740	0.487	0.709	0.692	0.596	0.943	0.679	0.743	0.740	0.496	0.769	0.660
	KHM			1.000	0.377	0.837	0.327	0.62	0.707	0.356	0.673	0.929	0.499	0.810	0.347	0.697	0.739
	MYS				1.000	0.596	0.592	0.564	0.606	0.593	0.628	0.437	0.957	0.593	0.545	0.638	0.519
	PHL					1.000	0.475	0.773	0.718	0.542	0.732	0.838	0.699	0.974	0.481	0.824	0.703
	SGP						1.000	0.513	0.271	0.331	0.420	0.401	0.550	0.511	0.970	0.532	0.199
	THA							1.000	0.626	0.302	0.573	0.725	0.636	0.737	0.531	0.929	0.619
	VNM								1.000	0.541	0.688	0.670	0.692	0.663	0.269	0.710	0.946
2015	BRN								1.000	0.668	0.354	0.629	0.575	0.307	0.450	0.402	
	IDN									1.000	0.711	0.712	0.723	0.422	0.709	0.656	
	KHM										1.000	0.554	0.813	0.415	0.747	0.690	
	MYS											1.000	0.699	0.522	0.713	0.619	
	PHL												1.000	0.517	0.768	0.640	
	SGP													1.000	0.550	0.226	
	THA														1.000	0.697	
VNM															1.000		

出所) 分析結果より筆者作成.

註: 1) BRN はブルネイ, IDN はインドネシア, KHM はカンボジア, MY S はマレーシア, PHL はフィリピン, SGP はシンガポール, THA はタイ, VNM はベトナムを示す.

註: 2) 中心性の値は少数第 4 位を切り上げ.

### 3.3.4 考察

まず、中心性の順位や相関性の結果から確認されるように、RWC と CBET の値の傾向は非常に近く、高い相関性が見られることがわかる。これは、中心性指標の相関性について研究した Oldham et al. (2019) の研究結果とも符合する結果である。

また、主要セクターの結果からは、2005 年と 2015 年の間に主要セクターはほとんど不動であったことが伺え、この期間に各国において大きな産業構造の変化は生じていないものだと考えられる。その点は、表 5 に示した同一国家の 2005 年と 2015 年での相関係数がいずれも 0.9 以上と高い相関性を示していることから確認出来る。更に、相関性の結果からは、シンガポールとブルネイは ASEAN の他国との相関性が比較的低い傾向にあり、これら二カ国の産業構造が他の ASEAN の国々とは異なることを示す結果であった。しかしながら、ブルネイの相関性の結果の内、年度間変化に着目すると、2005 年時点より 2015 年時点の方が他国家との相関性が比較的高まっている傾向にあり、産業構造が他国家に近づいている可能性が示唆された。この点については、他国がこの間で経済発展した結果、他国に比べて先んじて発展していたブルネイとの差が埋まったという要因が考えられる。他に、フードシステム関連セクターの内、「Accommodation and food services」の RWC 順位が大きく上昇していることから、ブルネイの産業構造の転換が進んでいる可能性が高いと考えることも出来る。そして、この RWC の順位の上昇は、ブルネイの「Accommodation and food services」セクターが、産業ネットワーク内で生じた経済的影響を、より早く受けるように構造が変化したということを表している。

次に、RWC と CBET で異なる主要セクターが抽出されたフィリピンとタイに着目する。まず、CBET の主要セクターに関して「Food products, beverages and tobacco」が抽出されたフィリピンは、食品関連産業が国内産業全体において、経済的効果を多くの産業へと波及させる役割を担っていると考えられる。すなわち、フードシステム関連セクターの開発が、国全体の産業の発展を図る上で重要であることが示唆されている。更に、タイの結果についても、他国と異なる構造の特徴的な産業ネットワークが形成されていることが推察される。

### 3.4 結論と今後の課題

本章では、産業連関表に基づいた ASEAN 各国の産業構造の比較に際して、主要セクターのみでなくフードシステム関連セクターに着目すると共に、産業連関表から構築されるネットワークに適した中心性を適用し、分析を行った。その結果、主要セクターの抽出による各国の代表的な産業をネットワーク分析により特定し、更に、フードシステム関連セクターの順位に基づいた ASEAN メンバー国のグルーピングを行った。また、主要セク

一の結果や相関性分析の結果から、2005年時点と2015年時点では多くの国で産業構造に大きな変化は生じていないことが確認された。

ASEANを含む東アジアは今後ますます貿易上重要になると考えられ、東アジアを包括したフードシステムの形成が重要視されている（下渡・名取，2010）。フードシステムのGlobal Value Chain（GVC）を計量的に分析した株田（2015）も示す通り、今後はグローバルフードシステムに対する計量的分析がますます求められるであろう。国際産業連関表への適用も可能な本研究の分析枠組みからも、これらの分析視点にアプローチする方向性を展望したい。

註1) STAN データベースでは2018年版のInput-Output Tables以前の産業連関表についても提供されており、2005年より古い年度のデータも存在しているが、セクターの集計が2018年版とは異なるため、データの整合性の点から、2018年版のデータのみを用いた。

註2) RWCとCBETの詳細な導出方法については、Blöchl et al (2011)とDePaolis et al (2022)を参照されたい。なお、RWCとCBETの算出については、DePaolis et al (2022)により提供されている統計解析環境Rで実行可能なコードを利用することで実装が可能である ([https://rpubs.com/RStudio\\_knight/368268](https://rpubs.com/RStudio_knight/368268))。

註3) モジュラリティによるコミュニティ検出のツールとしてはGephiを用いており、アルゴリズムは、Blondel et al. (2008)で示されたLouvain Methodが活用されている。

註4) スピアマンの順位相関係数はノンパラメトリックな統計手法であり、数値が1.0に近いほど強い正の相関を示す。

註5) 表3-5の結果に示されたモジュラリティの値は比較的低い数値であると言えるが、これは、Average Clustering Coefficientの値にも示される通り、ノードの多くがリンクを持っており、ネットワークが密で分割しにくいことが要因として推察される。

## 第4章 社会科学分野における有機農業研究に対する科学計量分析

### 4.1 はじめに

社会全体として持続可能な産業開発が要請される中で、第1章においても触れた通り、とりわけ農業においては、農薬や肥料などの投入による環境負荷の低減を目的として、化学肥料などを投入する慣行農業よりも環境負荷が小さいとされる有機農業の推進が求められるようになってきた。そうした背景から、SDGsの推進ともあいまって、各国において有機農業の拡大が政策的に推し進められている。

わが国においては、2022年に「みどりの食料システム戦略」が施行され、持続可能な農業の推進に向けて、国内有機農業の面的拡大が目標として掲げられた。したがって、国内外問わず、有機農業の推進は重要性を増す状況にあると言える。政策的な側面のみならず、有機農業に対する学術的なアプローチについても蓄積が進んでおり、有機農業の収益性を分析した研究事例（Crowder and Reganold, 2015; Reganold and Wachter, 2016）においては、世界的な規模での有機農業拡大の可能性が指摘されている。また、食料供給面についても、有機農業は十分な食料供給をまかなうだけのポテンシャルを持つと推定されている（Muller et al., 2017）。こうした研究事例をはじめ、自然科学、社会科学を問わず、有機農業を取り扱った様々な学際的なアプローチが試みられている。

こうした学術文献の蓄積は有機農業に限らず様々な分野で進展しており、近年はそうした文献データの蓄積が著しい。そして、そうした学術文献の蓄積に伴い、学術文献を対象とした大量の情報を処理するデータマイニング手法の適用が盛んに行われるようになってきた。特に、学術論文に含まれる様々な情報を定量的に分析する科学計量分析の適用が盛んに行われるようになってきている。科学計量分析は計量的な分析に留まらず、研究動向の可視化といった多面的な展開が進んでいる（Mingers and Leydesdorff, 2015）。

とりわけ農学に関連した科学計量分析の適用事例として、José et al. (2015) は有機農業に関する研究グループの協力関係に着目して分析し、José et al. (2018) は精密農業（Precision Agriculture）の研究動向について科学計量分析を用いている。他にも、Manuelian et al. (2020) は、有機畜産に関する研究の動向に科学計量分析の枠組みを適用している。これらの研究事例は、農業に関連する研究テーマに対する科学計量分析によるアプローチの有用性を示したものであると言えるだろう。

そこで、本研究では、学術文献にもとづいた今後の有機農業の発展に資する知見の獲得を目的として、有機農業に関する学術文献を対象として、科学計量分析を適用する。具体的には、研究動向と研究トピックについて、科学計量分析で活用されている定量的な分析指標を利用することによって、それらの抽出を図る。そして、分析対象となる学術分野と

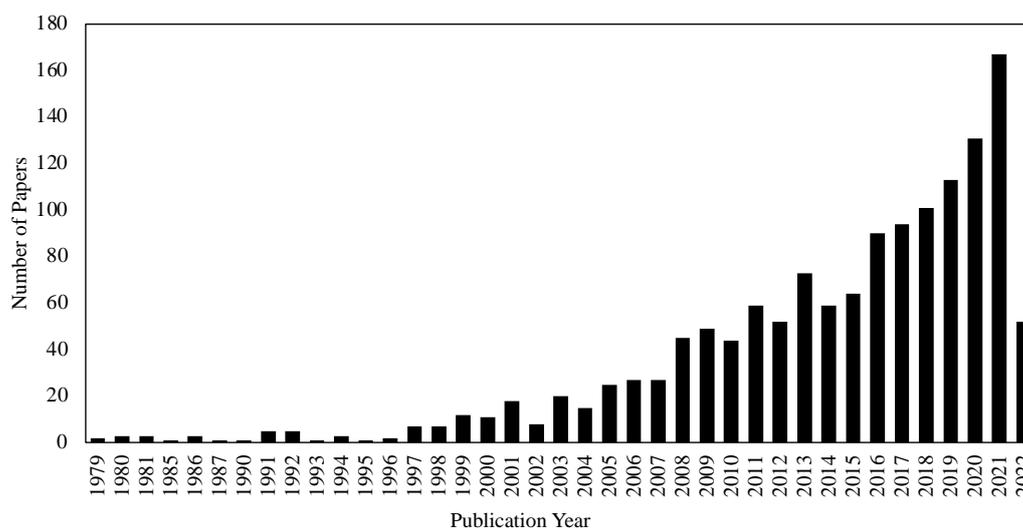
して、様々な側面からの有機農業研究が蓄積されながらも、先述したような科学計量分析を用いたアプローチによる先行研究事例の蓄積が依然として比較的乏しい状態にある，社会科学系分野を対象とした。

## 4.2 方法

### 4.2.1 データ

本章では，分析対象のデータとして，Clarivate 社が提供する大規模学術文献データベース「Web of Science」に収録されている学術文献に関するデータを取得した．そして，有機農業に関連する学術文献を取得するため，検索条件として，「organic agriculture」もしくは「organic farming」のいずれかを「検索範囲：Topic」に含んだ文献を指定した．更に，社会科学系の学術文献に分析対象を限定するため，Web of Science Core Collection のデータベースの中でも「Social Science Citation Index」が付与されている学術文献のみがヒットするように検索条件を設定した（註1）。

以上の条件の下で横断的に検索した結果，合計 1,401 件の論文がヒットした．検索で得た文献の年度別の出版数について，図 4-1 に示す．そして，最終的な分析対象としては，最近の 20 年間にわたる研究動向を中心に分析するため，検索結果で得られた全 1401 件の論文のうち，2002 年から 2021 年までの 20 年間に発表された 1263 件の論文を対象とした．



出所) Web of Science Core Collection の Social Science Citation Index の検索結果をもとに筆者作成。

図 4-1 検索により検出された文献の年度別出版数

## 4.2.2 科学計量分析

科学計量分析は、第2章において紹介したように、科学計量学において用いられる定量的指標を活用した分析である。本章の分析では、科学計量分析の手法として、主に下記の2つの手法を採用した。

1つ目の手法として、学術的な研究トピック及び研究トレンドの推定を目的として、キーワード共起分析 (Keywords Co-occurrence Network) を利用した。キーワード共起分析は、科学計量学で用いられる分析手法の一つであり、学術文献に付与されたキーワードに基づいた共起ネットワーク分析を指し示す。

キーワード共起分析においては、ある論文に同時に付与されたキーワードをネットワークにおける共起関係として捉え、それらのキーワードをノードとして捉えて、ネットワーク分析を適用する。論文に付与されたキーワードは、その論文が取り扱う研究テーマを端的に示すものであると考えられるため、キーワード共起分析は、研究のトレンドやトピックを反映したものとして捉えられる。

また、複数の文献を対象にキーワードの共起性を分析するため、幅広い範囲でのキーワードの共起関係を把握することが可能である。したがって、共起キーワード分析を用いることで、研究分野の全体的な動向を概観することができると考えられる。そういった背景から、本手法は、研究のナレッジマッピングや研究動向の分析に有用な分析手法であるとされており (Su and Lee, 2010)、特定分野における研究動向の分析に用いられることが多い (Lozano et al.)。

以上を受けて、本分析においては、有機農業の研究動向の把握に向け、キーワード共起分析を用いることとする。そして、キーワード共起分析においては、関連性の高いキーワードのグループ、すなわち共起関係の強いキーワード群の検出を目的として、本論文の他の章においても用いているように、モジュラリティを用いたキーワードのクラスタリングについても行う。

更に、別の分析手法として、本章では、キーワード共起分析と共に、科学計量分析における手法の一つであるバースト検出 (Burst Detection) を用いる (註2)。例えば、論文の引用・被引用関係にバースト検出を適用した場合、特定の期間に被引用数が急激に増加した論文を検出することが可能となる。

そして、本分析のように、論文に付与されたキーワードに対してバースト検出を行うと、特定のキーワードの急増が検出することが出来る。したがって、キーワードを対象としたバースト検出の適用は、特定の時期において盛んとなった研究トピックの検出や、特定の時期の研究動向の推定などに有効な分析であると考えられる。

## 4.2.3 実装ツール: Citespace

本分析では、先述したキーワード共起分析およびバースト検出のツールとして、Citespace 6.1.R2 (64-bit) Advanced を使用した。Citespace は様々な科学計量分析が利用可能なツールであり、Web of Science から取得したデータの分析が可能ないように設計されている (Chen, 2004; Chen, 2006)。Citespace を用いることで、クラスタリングや各クラスターを代表するようなラベルの推定など、様々な分析を行うことができる。各クラスターのラベリングについては、対数尤度比 (LLR : Log-likelihood ratio) といった計算方法を用いて、クラスターを特徴付ける適切な用語を検出するものとなっている。本分析においては、LLR によるクラスターのラベリングを行った。本分析における Citespace での設定パラメータは表 4-1 の通りである。

また、Citespace を用いた農業に関連する研究事例として、Qi-Qi et al. (2016) は国際的な視点から農業技術の進化について分析しており、Sun et al. (2019) は、米や肥料に関する学術文献を対象に Citespace を用いており、研究の動向を分析している。

表 4-1 Citespace におけるパラメータ設定

Parameters	Value
Selection Criteria	Top 20 %
Node Types	Keyword
Strength of Links	Cosine
Scope	Within Slices
Pruning	FALSE
Years Per Slice	1 year
Time Slicing	2002 JAN to 2021 DEC

出所) 実装設定にもとづいて筆者作成。

註：1) Pruning のアルゴリズムは pathfinder か minimam spanning tree を示す。

註：2) Selection criteria においては、1 スライスあたりの最大選択数は 100 となる。

註：3) 各種パラメータに関する詳細については、Chen (2004)を参照されたい。

## 4.3 結果と考察

### 4.3.1 キーワード共起分析

まず、キーワード共起分析の結果、2,041 個のノード、すなわちキーワードが検出され、更に、11,019 個のリンク、すなわちキーワードの共起関係が検出された。更に、キーワードの頻出回数上位 30 語と、次数の高いキーワード上位 30 語について、表 4-2 に示す。これらの分析結果を参照すると、“organic agriculture,” “organic farming,” “farmer,” “food” といったような農業や有機農業に関連する一般的なキーワードが検出されていることが確

認出来る。また、そうした農業や有機農業に関連する一般的キーワードに加えて、sustainability, biodiversity, ecosystem service といったような環境面・持続可能性に関連するキーワードについても多数検出される結果となった。

表 4-2 登場頻度と回数に関する上位 30 キーワード

Top Frequently Keyword	Frequency	Top Degree Keyword	Degree
organic agriculture	316	knowledge	86
organic farming	307	diversity	86
agriculture	189	health	86
system	127	conversion	83
management	122	ecosystem service	82
food	108	consumption	82
sustainability	98	attitude	77
impact	87	conservation	77
farmer	81	environmental impact	74
biodiversity	76	land use	74
adoption	69	behavior	74
ecosystem service	67	innovation	74
quality	59	farm	70
sustainable agriculture	57	network	68
attitude	54	nitrogen	67
food security	51	adoption	66
environmental impact	47	yield	66
organic food	47	performance	66
climate change	45	framework	63
consumption	42	determinant	62
conservation	42	cropping system	62
land use	42	organic agriculture	62
model	40	information	62
determinant	39	model	61
knowledge	35	quality	60
network	34	impact	59
policy	34	life cycle assessment	59
greenhouse gas emission	34	sustainability	58
willingness to pay	34	landscape	57

出所) 分析結果より筆者作成.

註: 1) Frequency はキーワードの出現頻度を示す.

註: 2) Degree はキーワードのリンクの回数を示す.

次に、クラスタリングを適用した結果、表 4-3 に示すように、合計で 24 のクラスターが検出された。クラスタリング結果のネットワーク図について、図 4-2 に示す。なお、モジュラリティの値は 0.6523 であった。更に、表 4-3 に、LLR によって算出された各クラスターを特徴付けるラベリングの結果一覧を示す。

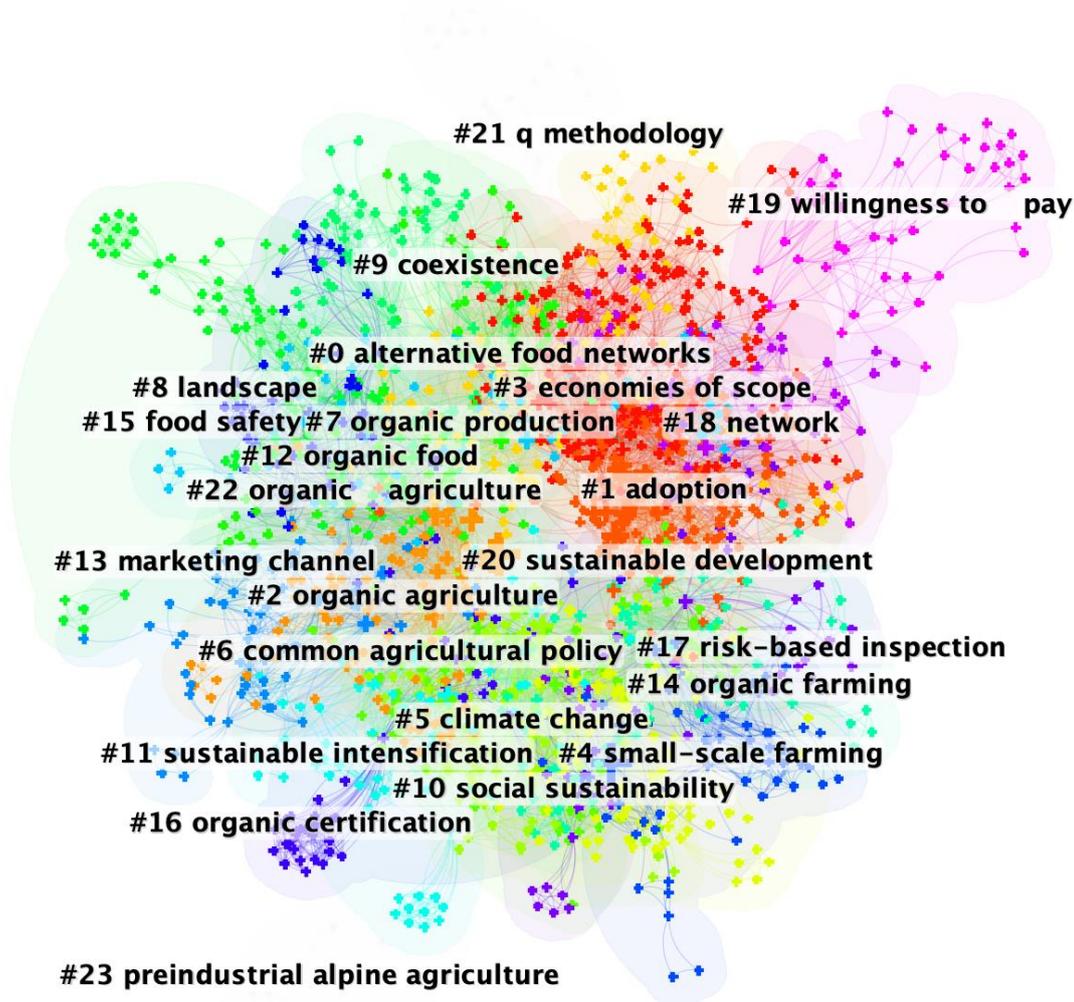


図 4-2 ネットワーク図におけるクラスタリングとラベリングの結果

出所) Citespace の分析結果より出力して筆者作成。

註：1) 各クラスターのラベルは LLR (Log-likelihood ratio) に基づき算出されたものである。

表 4-3 のクラスターラベルの結果を参照すると、有機農業に関連する様々なトピックが抽出されたことが確認できる。各クラスターを見ていくと、最も多くのキーワードを含むクラスターID1 は 214 個のキーワードを含んでおり、ラベル (LLR) の結果から、“alternative food networks” に関連しており、“new zealand,” “california,” “mexico” といったような国名との関係性の強いクラスターであることがわかる。

表 4-3 キーワード共起分析に基づいたクラスタリング結果と各クラスターにおけるラベリングの結果

ID	Size	Silhouette	Label (Log-likelihood ratio)
1	214	0.766	alternative food networks; new zealand; california; mexico; quality
2	187	0.729	adoption; rational choice; technology adoption; organic farming adoption; organic farming
3	108	0.858	organic agriculture; land use; yield; soil fertility; agricultural practices
4	100	0.906	economies of scope; science; organic seed; policy capacity; interest groups
5	96	0.870	small-scale farming; data envelopment analysis; logistic regression; uganda; fair trade
6	93	0.903	climate change; life cycle assessment; greenhouse gas emissions; sustainable agriculture; soil quality
7	91	0.882	common agricultural policy; social metabolism; nutrient balance; dynamic programming; spatial econometrics
8	86	0.879	organic production; animal welfare; piglet castration; biomass energy; sustainable development
9	86	0.886	landscape; switzerland; organic farmers; dairy cows; farm management
10	79	0.938	coexistence; genetic engineering; political economy; organic certification; gmo
11	73	0.902	social sustainability; political ecology; treadmill of production; social justice; labor
12	72	0.943	sustainable intensification; ecosystem services; land sparing; farmer opinions; interdisciplinary analysis
13	70	0.887	organic food; environmental policy; iran; consumer behavior; latent class analysis
14	63	0.959	marketing channel; life cycle assessment; soil; resource use; farming system
15	62	0.935	organic farming; conventional farming; organic agriculture; production function; multivariate analysis
16	57	0.926	food safety; animal health; agricultural ecology; antioxidants; denmark
17	57	0.945	organic certification; conservation tillage; yield-biodiversity trade offs; pigs; land sparing vs sharing
18	51	0.918	risk-based inspection; structural coupling; agriculture and food; cap subsidies; organic horticulture
19	51	0.946	network; system; farm households; new zealand; auditing
20	45	0.982	willingness to pay; consumer perception; choice experiment; conjoint analysis; ethical trade
21	42	0.942	sustainable development; agri-environmental measures; positive mathematical programming; system improvement; meta-governance
22	33	0.966	q methodology; virtuous circle; landscape scale; cultural landscapes; qualitative models
23	30	0.960	organic agriculture; critical race; sequential minimal optimization; alternative food movement; land separation
24	17	0.988	preindustrial alpine agriculture; northern thailand; soil nitrogen; long-term socio-ecological research (Itser); stochastic metafrontier

出所) 分析結果より筆者作成.

註: 1) Size はクラスターに含まれるキーワード数を示す.

註: 2) Silhouette の値は-1 から 1 を取り, 1 に近いほど適切なクラスタリングであることを意味する.

註: 3) 表中のラベリングは, Citespace により算出された LLR (Log-likelihood ratio) の値を基準に決定している.

したがって、ID1 のクラスターは、“alternative food networks”に関連して、こういった地域に有機農業に関する多くの事例があることを示していると考えられる。次点で多くのキーワードを含むクラスターID2 は、ラベルに“technology adoption”や“organic farming adoption”が含まれていることから、生産者の有機農業への受容に関連しているクラスターであると推察される。つまり、クラスターID2 は、有機農業について生産者側に焦点を当てた研究トピックであると考えられる。一方で、消費者サイドの研究として、“willingness to pay”や“consumer perception”などのラベルを持つクラスターID20 についても抽出された。クラスターID20 は、有機農産物への消費者受容に関する研究トピック、あるいはマーケティング関連の研究トピックであると考えられる。更に、クラスターID4 や ID7 のように、“policy capacity”や“common agricultural policy”など、政策的な側面に関する研究を示すラベルも確認された。そのほかに、クラスターID6 では、“greenhouse gas emission”といったラベルが登場するなど、持続可能性に関わる研究と推察されるクラスターについても多数確認される結果となった。

したがって、以上の結果を総合すると、(1) 事例研究、(2) 生産者サイドに関する研究、(3) 消費者サイドに関する研究、(4) 政策サイドに関する研究、(5) 有機農業の持続可能性という5つのおおまかな研究トピックを抽出された結果となった。

#### 4.3.2 バースト検出

バースト検出の結果について、図4-3に示す。図4-3に示された通り、全部で39語のバースト単語が検出された。検出時期としては、ほとんどが2010年代後半に集中する結果となった。更に、期間別にバースト単語を見ていくと、“california”や“uk”は、2000年代を通して検出される結果となった。他にも、“consumer behaviour”は2009年から2013年にかけて検出されたバースト単語であるが、2012年から2013年にかけては、他のバースト単語が検出されなかったため、“consumer behavior”が唯一のバースト単語として検出される結果となった。

バースト単語として検出されたキーワードを概観すると、“greenhouse gas emission,” “biodiversity,” “climate change”などの環境に関する単語や、“market,” “consumer behaviour”などの消費に関する単語など、多様なキーワードが検出されたことがわかる。また、バーストの強度の高いキーワードとしては、“system”や“impact”が検出された。

更に、バースト単語として検出された期間を参照すると、“california”は2002年から2011年までの10年間にわたり出現しており、他のバースト単語と比較して、最も長期間出現する結果となった。その他のバースト単語の多くは2016年以降に出現しており、2021年までバースト単語として検出され続けるキーワードと、それ以前にバースト単語としての検出が終了するキーワードとに分かれる結果となった。

Keywords	Strength	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
california	3.94																				
uk	3.09																				
consumer behaviour	3.29																				
politics	3.96																				
system	14.61																				
impact	11.34																				
greenhouse gas emission	5.76																				
food	5.53																				
determinant	4.77																				
framework	4.30																				
product	4.20																				
market	3.98																				
metaanalysis	3.91																				
performance	3.91																				
food security	3.89																				
governance	3.66																				
network	3.66																				
biodiversity	3.44																				
ecosystem service	8.33																				
agriculture	6.45																				
supply chain	4.48																				
intensification	4.25																				
smallholder farmer	4.08																				
alternative food network	3.69																				
united states	3.30																				
landscape	3.26																				
environmental impact	7.76																				
life cycle assessment	4.79																				
water	4.32																				
urban agriculture	4.27																				
climate change	4.18																				
trend	3.30																				
carbon sequestration	3.09																				
spain	3.08																				
perception	5.83																				
consumption	5.61																				
management	5.25																				
behavior	5.13																				
consumer	4.18																				

出所) 分析結果より筆者作成.

註: 1) バースト検出の際のパラメータはデフォルト値を設定した.

註: 2) 図中における黒の網掛けの期間がバーストワードとしてそれぞれのキーワードが検出された期間となる.

図 4-3 バースト単語の推移

### 4.3.3 考察

表 4-3 で示されたクラスタリングラベルの結果は、生産サイドから消費サイドまでを包括した有機農業に関する社会科学的研究の多様性を示唆するものであったと捉えられる。また、“landscape”や“animal welfare”などの有機農業とは直接的には関係のないラベルについても検出されていることから、表 4-3 のクラスタリングラベルの結果は、有機農業が研究対象として多面的な性格を持っていることを示すものと考えることが出来る。

また、バースト検出の結果、有機農業の先進地域であるカリフォルニアは、特に 2000 年代において、有機農業研究の事例研究として重要な対象地域であったことがわかる。更に、バースト単語として 2019 年から 2021 年にかけて出現する“spain”は、近年有機農業が急速に拡大している地域であり、2021 年時点で EU 最大の有機農業面積を有している。したがって、本分析結果は、そうした背景を考慮すると、有機農業の事例研究において、近年においてはスペインが注目されているという動向を反映した結果であると考えられる。つまり、社会科学的研究における事例研究の変遷が確認されたと捉えられる。また、他にも、“metaanalysis”といったキーワードがバースト単語として検出されるなど、有機農業に関する多彩な研究が展開されていることが示唆される結果となった。

その他の動向として、持続可能性に関連するキーワードも近年頻繁に登場しており、2010 年代後半には“carbon sequestration”、“lifecyle asessment”、“greenhouse gas emmission”といったキーワードがバースト単語として検出された点も特徴的である。持続可能性に関する注目は、有機農業研究に限らず様々な研究分野で共通する傾向であるが、今回の分析結果から、有機農業と持続可能性とを結びつけている研究事例が、近年ますます増えているであろうことが確認された。

以上、学術文献に科学計量分析を適用することで、有機農業に関連する上記のような知見を得ることができた。本分析の結果から、科学計量分析は、人力で把握することが困難な大量の科学研究をもとに、データマイニングのアプローチによる文献レビューを行うことが可能であることがわかった。しかしながら、本分析の限界として、得られた知見は研究レビューの範疇にとどまっている点に留意する必要があるであろう。

## 4.4 結論と今後の課題

本章では、キーワード共起ネットワーク分析およびバースト検出を中心とした科学計量学の分析アプローチを用いて、有機農業に関する社会科学的研究の動向と研究トピックに関するデータマイニングによる分析を行った。これらの科学計量分析を用いることにより、人力によるレビューでは把握が困難な量の文献について、全体の傾向を把握することの出来る定量的分析が可能となった。

そして、本研究の分析結果から、以下の3つの知見が得られた。まず、第一に、有機農業に関する様々な研究動向やトピックを検出し、その結果、(1) 事例研究、(2) 生産者に関する研究、(3) 消費者に関する研究、(4) 政策に関する研究、(5) 有機農業の持続可能性ないし環境的側面に関する研究、といった5つの主要な研究トピックが明らかになった。

第二に、カリフォルニアやスペインを中心に多くの地域でケーススタディ、すなわち事例研究が蓄積されており、そうした研究事例の対象地について、かつてはカリフォルニアが中心であり、近年ではスペインを対象とした研究が盛んになっていることが推察される結果が得られた。

そして、第三に、有機農業の持続可能性に関する研究が近年特に盛んになっていることが推察された。持続可能性は有機農業に限らず、近年における重要な課題として取り組まれている状況であるが、本章では、そうした研究トレンドが定量的な指標の下で改めて確認される結果になったと言える。以上、本研究で示された科学計量分析の結果を総合すると、科学計量分析によるアプローチは、学術文献の定量的サーベイにおける有効なアプローチの一つであることが例証されたと捉えることが出来る。

本研究の今後の課題としては、まず、本分析では社会科学分野に分析対象を限定したため、自然科学の研究を含むより広い範囲の統合的な科学計量分析を適用することが挙げられる。別の方向性の課題としては、本章での分析結果から得られた各研究トピックに関する詳細な分析を試みることも有益であると考えられる。また、本分析においては、Web of Science に収録された学術文献のみを分析対象としており、Web of Science に収録されていない学術誌については分析対象外となっているため、分析対象となった学術文献に偏りが生じている可能性についても考慮する必要がある。したがって、より広い範囲での文献収集を行い、新たに分析することが挙げられる。

註 1) Web of Science の検索においては、連続した単語によるフレーズ検索が可能となっている。今回の検索では、"organic agriculture"と "organic farming"をフレーズとして設定し、検索クエリとして利用した。また、Web of Science におけるトピック検索の範囲は、Title, Abstract, Author Keyword, Keyword Plus である。なお、Author Keywords は著者が文献に付与したキーワードを指しており、Keyword Plus は、学術文献がデータベースに収録される際に、Web of Science 側がそれらの文献に対して新たに付与したキーワードである。

註 2) Citespace に実装されているバースト検出は、Kleinberg (2002) が提案したアルゴリズムに基づいたものである。

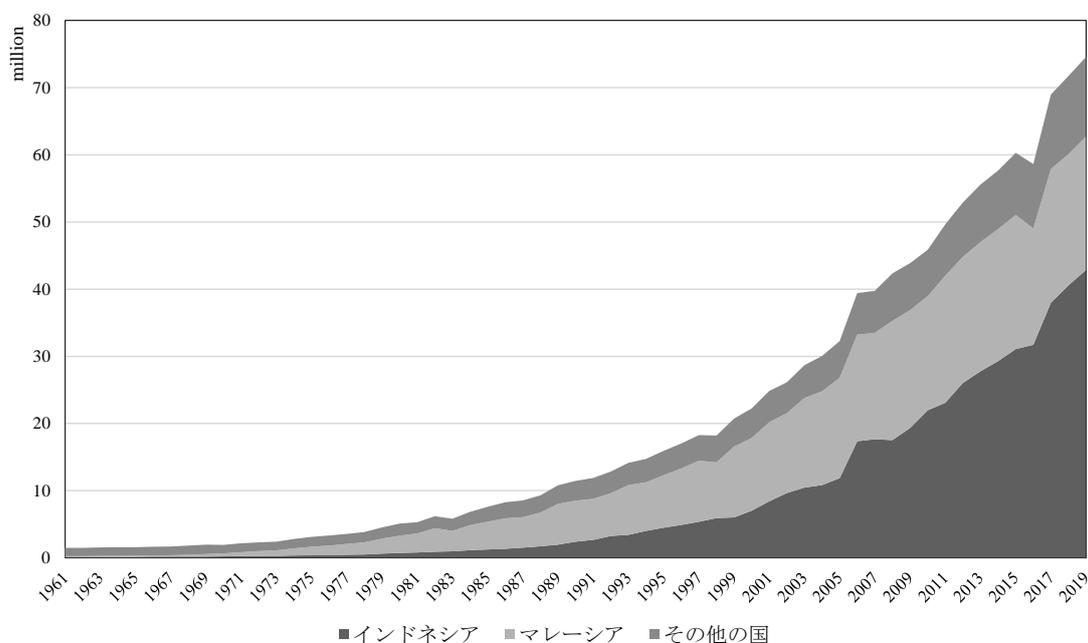
## 第5章 パーム油産業における廃棄物処理に関する技術開発の動向

### 5.1 はじめに

植物性油脂の中でも、パーム油は2022年現在、世界最大の生産量となっており、油脂としての用途の幅広さから、食品用油脂、産業用油脂など様々な場面で活用されている。パーム油はアブラヤシの果実や種子から搾油されるものであるが、アブラヤシは生産適地が限られており、図5-1に示す通り、近年では特にマレーシアとインドネシアが主要生産国となっている。

パーム油は、食品用途や産業用の油脂として幅広く活用される一方で、アブラヤシの植林に伴う熱帯雨林の破壊、生産にける児童労働の問題、そしてパーム油の生産過程で発生する環境負荷や廃棄物処理に関する問題（註1）など、パーム油の生産に関わる様々な課題が国際社会から指摘される状況にある。

とりわけ、廃棄物処理技術については、近年のパーム油の生産量の増加に比例して、廃棄物の発生量についても増加しており、インドネシアやマレーシアといった主要生産国の政府やパーム油産業にとって、重要な課題であることが指摘されている（Oseghale et al.2017）。



出所) FAOSTAT から取得したデータをもとに筆者作成。

図5-1 世界におけるパーム粗油の生産量の推移

こうした現況を考慮すると、パーム油生産と廃棄物処理に関連する技術開発の動向を把握することは、この分野における環境保全の取り組みを評価する上で重要であると言えるであろう。先行研究においては、そうした現状を受けて、パーム油に関する様々な角度からの研究が蓄積されている状況にあるが、特に学際的な研究アプローチの増加が必要であるとされている (Hansen et al. 2015)。

そのような学際的な研究アプローチの1つとして、データマイニングが挙げられる。技術開発動向の把握に向けたデータとして特許情報が存在しているが、そうした特許情報に対するデータマイニングについて、近年多くの研究事例が蓄積している。データマイニングは膨大な量のデータから効率的に情報を整理・抽出するために行われるものであるが、特に特許情報に対しては、自然言語処理技術の進歩やテキストマイニングの台頭により、特許情報のテキストに対するデータマイニングが盛んに行われるようになった。特許情報は特定の技術の詳細を記録した文書から構成されており、技術発展の動向を分析することに適したデータであるため、データマイニングの対象として有用であると言えよう。パーム油に関する特許情報を用いた研究事例としては、Wibowo et al. (2019) によるパーム油に関連する特許のマッピング研究が挙げられ、特許という側面からパーム油の開発状況に迫る有効性を示した研究事例と言えるであろう。

そこで、本章では、パーム油生産過程で生じ廃棄物処理技術に焦点を当て、パーム油産業に関連する特許情報に自然言語処理技術を適用することにより、廃棄物処理技術に関する技術動向に対する分析を試みる。

そして、分析に用いる自然言語処理技術として、深層学習を利用したニューラルトピックモデルを採用することで、特許内容の意味的な側面に着目した分析を行い、従来の分類とは異なる、特許の内容により即した形での特許分類について提示することについても試みることにする。

## 5.2 方法

### 5.2.1 データ

本章では、分析対象のデータとして、World Intellectual Property Organization (以下: WIPO) (2016) により収集されたパーム油関連特許のデータセットを利用する (註2)。このデータセットは、国際特許分類 (International Patent Classification: IPC) および Questel Orbit を中心とした特許検索ツールでのキーワード検索によって得られたパーム油技術に関連する特許から構成されている。

更に、このデータセットには、特許のタイトル、要約 (Abstract)、明細 (claim)、特許出願者、特許の出願日と公開日などの特許に関する基本的な情報が含まれている。そして、

そうした情報以外にも、各々の特許の分類に関する情報、すなわち、パーム油生産に関する技術か、もしくはパーム油の廃棄物処理に関する技術か、といったカテゴリーに関する情報についても提供されている。データセットには全部で2,370件の特許が含まれており、その内、1,046件はパーム油の生産に関連する特許であり、1,324件はパーム油の廃棄物処理技術に関連する特許である。

本分析においては、データセットに収録されている1995年から2015年の間に公開された特許のうち、廃棄物処理技術に関連しており、なおかつ、特許情報にAbstractのテキストが含まれている1,222件の特許を分析対象とした。

さらに、このデータセットには、各特許が扱う廃棄物の材料に関する分類について13種類、廃棄物の応用目的に関する分類について18種類の情報が含まれている。

図5-2は、本分析の対象となるデータについて、特許が公開された年度を基準とした各年度の特許件数を整理したものである。

図5-3は、先述した13種類の廃棄物材料に関して、それぞれの特許件数を示したものであり、図5-4は18種類の応用目的ごとの特許件数を示したものである。図5-2から、特許の出版数の推移を参照すると、年を経るごとに特許の件数が増加していることが確認出来る。また、図5-3に示した特許が取り扱っている廃棄物材料ごとの特許数については、パームファイバーが最も多く、全体の特許数の大半を占めていることがわかる。そして、図5-4からは、特許の応用目的について、「Matless/cushion」が最多となっていることが確認できる。

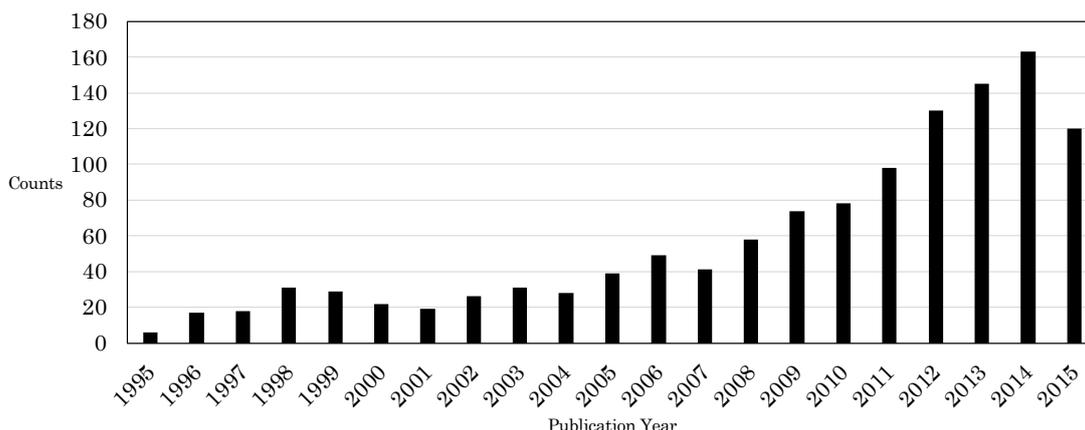
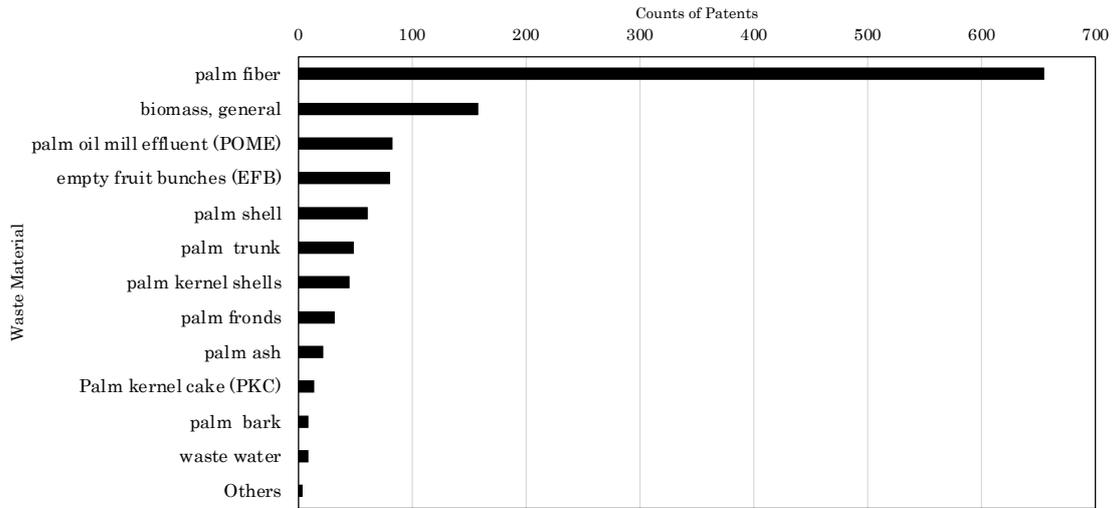
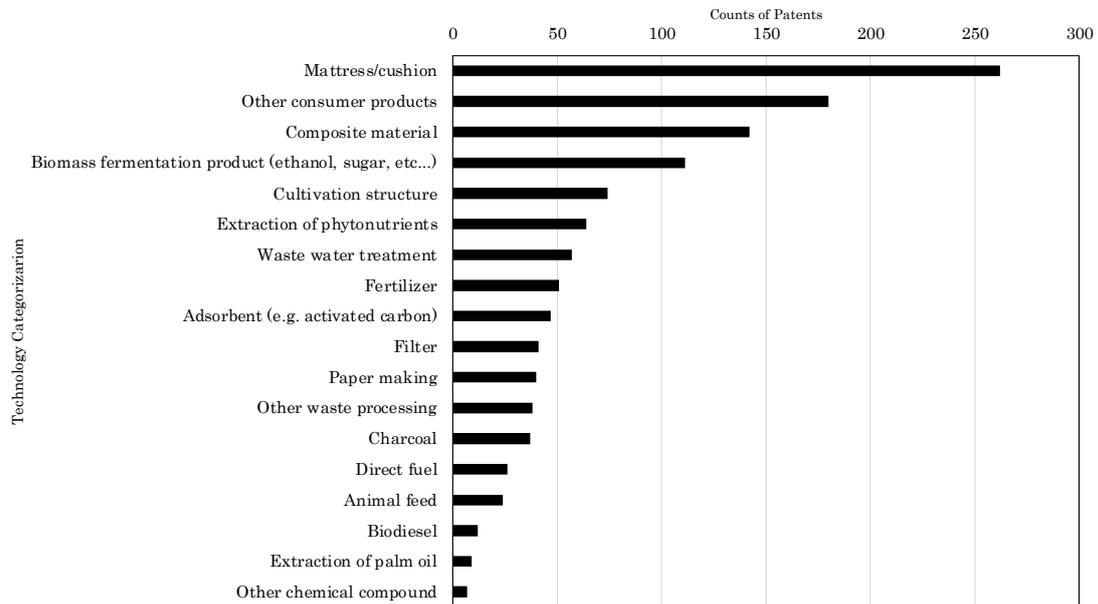


図5-2 本分析で用いた特許の年度別出版数



出所) World Intellectual Property Organization (2016) のデータセットの情報にもとづいて筆者作成.

図 5-3 各廃棄物を対象とした特許数



出所) World Intellectual Property Organization (2016) のデータセットの情報にもとづいて筆者作成.

図 5-4 廃棄物の利用目的ごとの特許数

## 5.2.2 ニューラルトピックモデル

トピックモデルとは、大量の文書に基づき、対象となる文書のトピックを推定するモデルである。従来のには確率的なモデルが用いられており、トピックモデルの代表例としては LDA (Latent Dirichlet allocation) (Blei et al. 2003) が挙げられ、様々なトピックの抽出に、広く活用されている。自然言語処理手法の中でもトピックモデルは特許情報のデータマイニングに広く利用されており、例えば、Chen ら (2017) は、特許請求項テキストにトピックモデルを適用し、技術動向を予測している。また、LDA の特許マイニング拡張版も開発されている (Chen et al 2016, Ma et al 2017)。

そして、本分析においては、トピックモデルの中でも、テキストの意味的側面を考慮することが可能なニューラルトピックモデルを分析手法に採用した。ニューラルトピックモデルは、自然言語処理における深層学習の枠組みを利用し、単語の意味に基づいてベクトル化することで文書を分類するものである。

実装ツールとして、本分析では BERTopic (Gootendoast 2020; Gootendoast, 2022) を使用した (註 3)。BERTopic は、トピック抽出以外にも、トピックの時系列推移を分析する時系列トピックモデル (Dynamic Topic Model) や、トピックを代表する文書の抽出など、様々な分析に利用することができる。本分析でも、Dynamic Topic Model を用いて、21 年間にわたる 10 時点のトピックの変遷、すなわち技術変遷を調べた。

ここで、本分析で採用する分析手法について概説する。ここでのニューラルトピックモデルは、分散表現を利用したトピックモデルである。BERTopic では、分散表現の獲得に一般的に用いられているモデルの一つである BERT (Devlin et al., 2018) の枠組み (註 4) を利用し、更に、トピックの抽出においては、次元削減手法の一つである UMAP (McInnes et al., 2018)、クラスタリング手法である HDBSCAN (Campello et al., 2013) を利用する。HDBSCAN により検出された各クラスター、すなわち意味的に近似した単語の集まりがトピックとなる。更に、トピックにおける単語の重要度を図るための指標として、クラスター単位、すなわちトピックごとにその単語の重要度を算出する手法である Class-based-TF-IDF (Özgür et al., 2005) を用いる (註 5)。

以上を整理すると、BERT の枠組みを利用して得られたベクトルに対して UMAP による次元圧縮を試み、更に、次元圧縮されたベクトルに対して HDBSCAN によるクラスタリングを行うことで、トピックを抽出するという分析手順となる (註 6)。そして、Class-based-TF-IDF により、各トピックにおける単語の重要度を算出する。

以上のアルゴリズムを利用することにより得られたトピックは、単語や文章の意味を含意した分類に基づいたものである。この点が、従来のトピックモデルと異なっており、本手法を採用する利点である。

### 5.2.3 分析の概要

トピックモデルによる分析にあたり、まず、事前の処理として、分析対象の特許の Abstract テキストを前処理する必要がある。テキスト分析における一般的な前処理としてストップワードの除去が挙げられるが、本分析においては、一般的なストップワードのみでなく、特許をはじめとした技術文書に特有のストップワードについても除去する必要があると考えられる（註7）。

そこで、本分析では、一般的なストップワードに加え、特許などの技術文章に特有のストップワードについても除去処理を行った。特許特有のストップワードのリストとして、USPTO (United States Patent and Trademark Office) の Stopwords (United States Patent and Trademark Office, 2020) および Sarica and Luo (2021)を参照した。なお、本分析でのストップワードの処理の実装には、Python のパッケージである cleanext 1.1.4 を使用した。そして、ストップワードを除去した後の特許の Abstract テキストを対象として、BERTopic の分析を適用した。

また、BERTopic での分析においては、基本的にデフォルトのパラメータを使用した。デフォルト以外のパラメータとしては、事前学習モデルとして「all-mpnet-base-v2」を指定している。「all-mpnet-base-v2」は計算に時間を要する反面、精度の高いモデルであるため、本分析においては、精度の高い分析結果を得ることを目的として、事前学習モデルとして「all-mpnet-base-v2」を採用した。

更に、トピックに抽出される単語の多様性に関するパラメータである「diversity」については、多様な単語の抽出が可能となるように、最も高い値、すなわち 1.0 を設定した。また、トピックの数の算出にあたっては、自動的にトピック数が推定される設定を利用した。

## 5.3 結果と考察

### 5.3.1 トピックモデルの分析結果

ニューラルトピックモデルによる分析の結果、合計 25 個のトピックが抽出された。トピック分析の結果について、表 5-1 に示す。本章においては、重要度の高いトピックに議論を限定するために、そのトピックに分類された特許件数の多い上位 10 つのトピックに注目する。また、これら 10 つのトピックについて、それぞれのトピックにおける代表的な特許情報を抽出した結果を表 5-2 に示す。

最も特許件数の多いトピックであるトピック ID1 は、「マットレスへのパームファイバーの利用」に関するものであると考えられ、2 番目に特許件数の多いトピックである ID2 は、「斜面における植生ネット」に関するものであると推察される。3 番目以降に特許件数

の多いトピックを順に見ていくと、ID3 は、「パームカーネルシエルの炭化 (charcoal) に関する技術」、ID4 は、「パームファイバーを利用したブラシ」、ID5 は、「空気のフィルターに関する部品などへの利用」、ID6 は、「POME (Palm Oil Mill Effluent) に対する処理や貯蔵に関するシステム」、ID7 は、「パームファイバーを利用したシューズパッド」、ID8 は「糖化 (saccharification) などのバイオマスに関する技術」、ID9 は、「パームファイバーを利用したプレート」、ID10 は、「空果房 (EFB) のパルプ化に関する技術」に関するものであると、それぞれ推察される結果となった。

本分析結果からは、トピック ID8 の「糖化 (saccharification)」、トピック ID23 の「リグニン (lignin)」など、同じバイオマス関連技術であっても、その技術特性から別のトピックとして抽出されていることが確認できる。

表 5-1 トピック分析の結果

Topic ID	Count	Words in topic
1	174	mattress_palm_arranged_spring_discloses_fibers_advantages_two_lower_permeability
2	93	vegetation_slope_net_floating_base_structure_jpo_copyright_palm_solved
3	70	charcoal_pks_palm_activated_heat_shells_kernel_solid_process_parts
4	57	brush_end_utility_palm_connected_part_fibers_fixedly_lower_box
5	53	parts_filtering_air_material_active_added_cover_tree_tail_desulfurization
6	50	tank_system_pome_palm_phase_first_treating_bod_ss_biomass
7	43	pad_care_utility_layers_sweat_shoes_natural_middle_made_edge
8	39	saccharification_solution_using_enzymes_sap_oil_produce_tropical_arabinose_palm
9	39	plate_mixed_strength_compression_curable_fabric_woody_solution_permeability_vegetable
10	37	step_pulping_empty_performing_raw_article_dehydration_used_non_efb
11	33	chinese_palm_moxibustion_effects_relates_disease_traditional_prickly_antioxidant_comprising
12	32	palm_veneer_wd_components_board_direction_length_predetermined_drying_compressed
13	30	tank_device_annular_pool_channel_pannier_filtration_arranged_percent_top
14	29	decomposition_mounted_blade_palm_feed_rollers_two_transmission_sheet_upper
15	29	fertilizer_ash_waste_ratio_bio_slow_subtils_liquid_palm_nuisanceless
16	24	fiberboard_palm_felt_manufacturing_opt_mixture_hot_blank_average_constituent
17	24	bunches_biodiesel_efb_empty_byproduct_step_guineensis_palm_using_vsca
18	23	molding_200_palm_wt_stability_using_ball_song_biodegradable_ozone
19	19	parts_bran_crucian_carp_protein_20_kernel_wt_feeds_added
20	16	toothed_racks_plurality_swing_end_palm_fiber_suction_pin_discloses
21	14	pkc_dilute_dregs_palm_followed_exchange_yeast_comprising_using_factors
22	12	ascon_cement_course_biomass_content_reducing_palm_nano_new_opba
23	12	lignin_step_extractor_disclosed_residue_products_degradation_obtain_high_delignified
24	11	process_caffeoylshikimic_properties_products_solvents_derivatives_dioxide_anti_carotenoids_example
25	11	parts_insulation_weight_10_pulp_agent_ether_door_proof_adhesive

出所) 分析結果より筆者作成.

註: 1) 表中の単語の並びは、左か順に c-TF-IDF スコアが高いものを記載した.

註: 2) Count は、そのトピックに割り当てられた特許の数を表す.

註: 3) 抽出された 25 のトピックの他に、特定のトピックに分類されなかった特許をまとめた「その他」のトピックが存在している。「その他」として、248 件の特許が分類された.

表 5-2 上位 10 トピックにおける代表的特許の情報

Topic ID	Title of Representative Patent for each topic	Patent Number	Waste Material	Technology Categorization
1	Multifunctional natural material bed mattress	CN201234775Y	palm fiber	Mattress/cushion
2	Vegetation unit	JP2000328573A	palm fiber	Cultivation structure
3	Palm kernel shell charcoal production method and apparatus	MY149639A	palm kernel shells	Charcoal
4	Method for embedding twisted brush head	CN102138728A	palm fiber	Other consumer products
5	Palm tree bark adsorbent with good desorption property	CN104174387A	palm bark	Adsorbent (e.g. activated carbon)
6	Waste treatment apparatus for recovering valuable substance without discharging waste water to outside	JP2008023514A	waste water	Waste water treatment
7	Palm fiber shoe pad	CN202364948U	palm fiber	Mattress/cushion
8	Removal of inhibitors of microbial fermentation from inhibitor-containing compositions	WO2011163620A1	biomass, general	Biomass fermentation product (ethanol, sugar, etc..)
9	Plate-like body or molded body and manufacture thereof	JPH10296707A	palm fiber	Mattress/cushion
10	Stock preparation method before paper pulp making from oil palm empty fruit bunch	CN101498109A	empty fruit bunches (EFB)	Paper making

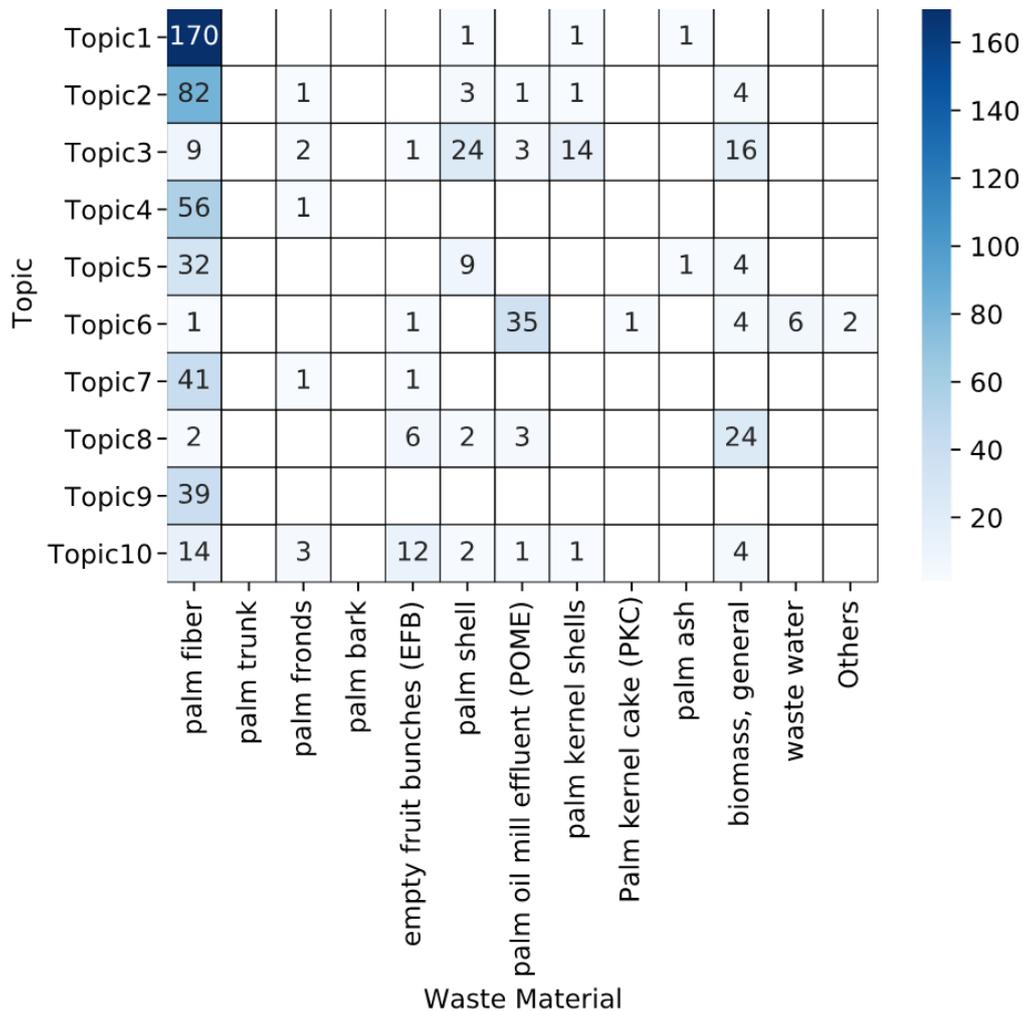
出所) 分析結果および WIPO(2016)のデータセットの情報にもとづいて筆者作成。

註: 1) トピック ID 以外の項目は, WIPO (2016) データセットを参照している。

表 5-1 と表 5-2 について参照すると, それぞれのトピックについて, そのトピックに含まれている単語が, 代表的な特許の概要を示すものとなっていることが確認出来る。特に, 特許件数の多いパームファイバーの利用に関しては, トピック ID7 のように「shoe pad」のような詳細な技術分類で捉えられていることがわかる。

また, 図 5-5 に, 各トピックに含まれる特許件数と, データセットに付与されている各々の特許が属する廃棄物材料のカテゴリの件数について整理したヒートマップの結果を示す。図 5-5 もヒートマップを参照すると, それぞれのトピックについて, 一つの廃棄物材料だけでなく, 複数の廃棄物材料に関連していることが確認できる。これは, 各トピックは, 特許内容の意味的な側面からの分類であるため, 各トピックの内容は, 一つの廃棄物材料に限っていないことを示唆するものであると考えられる。

特に, トピック ID6 に着目すると, 代表的な特許が取り扱っている廃棄物材料 (Waste Material) が「waste water」となっているのに対し, ヒートマップの中では POME が最多となっていることがわかる。



出所) 分析結果および WIPO(2016)のデータセットの情報にもとづいて筆者作成。

註: 1) 縦軸はトピック分析の分類結果であり, 横軸は WIPO データ記載の廃棄物カテゴリーの分類数を示している。

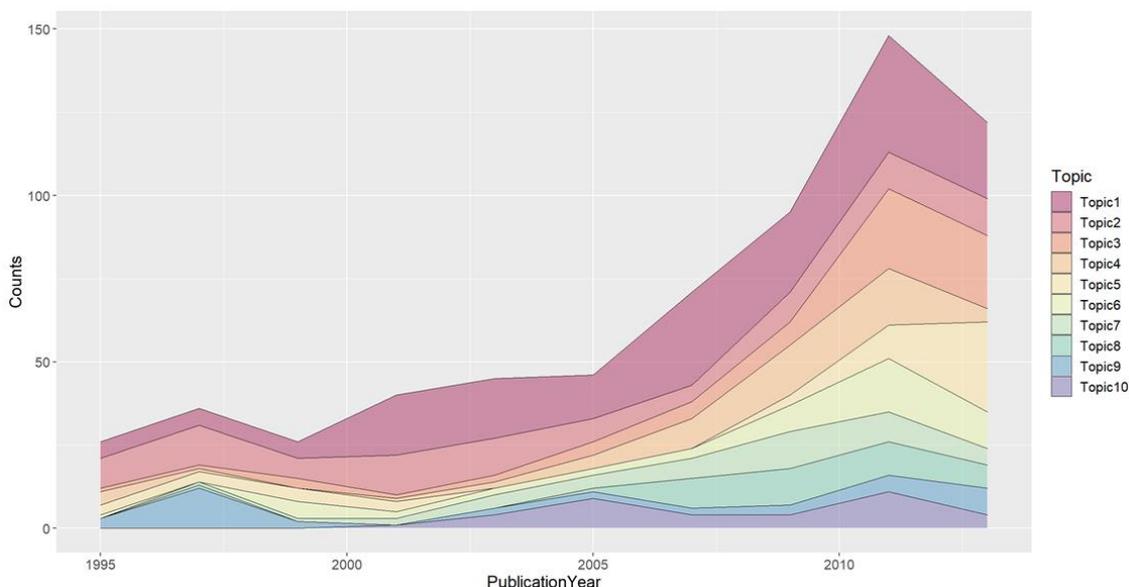
註: 2) 図中のトピック番号は表 5-1 の Topic ID と共通したものである。

図 5-5 トピック分類と廃棄物カテゴリーに基づいた特許数のヒートマップ

### 5.3.2 時系列トピックモデルの分析結果

さらに, 図 5-6 に, 1995 年から 2015 年までの期間における各トピックに含まれる特許件数の積み上げ面積グラフを示す。まず, 全体の傾向としては, パームファイバーのマットレスの利用に関するトピック, すなわちトピック ID1 がいずれの期間においても最も高い割合を占めており, 分析期間を通して, 常に一定数の特許が公開されていることが確認される。また, 1995 年から 2005 年頃の期間については, トピック ID2, すなわち斜面における植生ネットに関係する特許件数が多かったが, その後については, 他のトピックが増加したため, 減少傾向にあることが確認出来る。そして, 2010 年代に入り, トピック ID3,

ID5 が増加している傾向が見られる。つまり、「パームカーネルシェル（PKS）の炭化に関する技術」や、「空気のフィルターに関する部品などへの利用」といったの技術開発が進んでいることが推察される。



出所) 分析結果より筆者作成。

註：1) 時系列分析における分割は、10 時点として設定した。

註：2) 図中のトピック番号は表 5-1 の Topic ID と共通したものである。

図 5-6 各トピックに分類された特許数の年度ごとの積み上げ面積グラフ

### 5.3.3 考察

WIPO (2016) のデータセットに付与されている各特許の技術分類に関する情報と、本分析で抽出されたトピックの内容とを比較すると、本分析で抽出されたトピックは、より技術的な特徴を反映したものであると捉えることが出来る。特に顕著である例として、“saccharification (糖化)” が挙げられる。“saccharification (糖化)” は、第二世代バイオマスに関する重要な技術であるが (Robak and Balcerek 2018), こうした技術に関する細かな特徴についても、本分析では捉えることが可能であった。したがって、パーム油の廃棄物に関して、非食品用途の素材から生成される第二世代バイオマスに関する技術開発が進んでいることが示唆された。パーム油はバイオディーゼルの重要な原料になる可能性が指摘されてる状況にあるが (Tan et al.2009), 本分析においても、バイオマス関連技術の開発が進んでいることが確認された。

以上のような本分析の妥当性が確認された一方で、トピック分析の結果においては、パーム油とは関係性の薄い、特許に頻出する専門用語が多く抽出されていることも確認出来る。したがって、分析の精度を向上させる必要があることが示唆された。本分析では、先行研究で提示された特許に特有のストップワードを利用したが、分析結果からは、分析対象、すなわち、ここではパーム油の廃棄物処理技術に特有のストップワードを作成する必要性があることが示された。

#### 5.4 結論と今後の課題

本章では、パーム油生産に伴う廃棄物に関する特許情報の Abstract に対して、ニューラルトピックモデルを用いた分析を適用した。分析の結果、従来の技術分類では把握しにくいと考えられる特許の細かな内容的側面について、把握可能であることが示される結果が得られた。具体的な例として、第2世代バイオマス技術といったような、関連の特許の特徴などを把握することが出来た。この結果は、本分析の適用の有効性を示すものであると捉えることが出来よう。

更に、今回の分析結果から、用途の広いパームファイバーに関する特許が蓄積される一方で、環境負荷が大きく、処理の必要性が高い POME に関する処理技術の開発があまり進展していない点についてもあらためて確認される結果が得られた。このような未開発の技術の開発を促進することへの貢献は、本研究の今後の課題であろう。

また、手法的な面では、人力での把握が困難なデータ量の特許情報についてアプローチ可能である点についても、本分析の重要な特徴であると考えられる。その一方で、分析精度については未だ改善の余地が確認された。方法的な面では、既存の特許分類である IPC とトピックモデルをより一般化した形で組み合わせることが今後の課題として挙げられる。また、本章では分析対象として 2015 年までのデータセットを利用したが、パーム油の関連技術に関する特許は、それ以降も蓄積を続けており、多数の技術が開発される状況にある。したがって、今後の課題として、2015 年以降の特許情報を収集し、分析をより広い期間に拡張することが挙げられる。

註 1) 本章で取り扱っている各廃棄物の詳細については、World Intellectual Property Organization (2016) にそれぞれ説明が記載されているため、そちらを参照されたい。

また、各廃棄物の排出量としては POME (Palm oil mill effluent) が最も多く、World Intellectual Property Organization (2016) においては、パーム油を 1 トン生産するごとに約 1 トンの EFB (Empty fruit bunches) と、約 3 トンの POME が発生すると報告されている。

註 2) WIPO (2016) のデータセットは、2015 年 9 月 28 日に収集されたため、2015 年度の

特許に関するデータについては年度の半ばでの収集である点に留意する必要がある。

- 註 3) <https://github.com/MaartenGr/BERTopic> から本章および第 6 章において用いたツール BERTopic が利用可能である。なお、本章の分析では ver0.10.0 を利用した。
- 註 4) BERT は汎用的な事前学習モデルを利用する枠組みであり、登場時点で様々な評価指標で高評価を得たため、広く活用されるようになった。BERTopic も事前学習モデルを利用してベクトルを獲得している。
- 註 5) TF-IDF は単語の重要度を測る指標の一つであり、Term Frequency (単語出現頻度) と Inverse Document Frequency (逆文書頻度) をかけ合わせて算出する。Class-based-TF-IDF はトピックごとに TF-IDF を算出する。
- 註 6) ここでは確率的な計算を含んでいるため、分析ごとに結果が変動する点に注意されたい。
- 註 7) ストップワードは、助詞などの文章において特に意味を持たない単語などを指しており、テキストマイニングにおいては除去する必要がある単語として知られる。本章での分析対象は特許情報に関するテキストであるため、一般的なストップワードだけでなく、特許情報といった技術文書に特有のストップワードについても除去する必要がある。

## 第6章 英国と米国における肥満に関する報道の比較分析

### 6.1 はじめに

本章では、持続可能な食環境に関わる問題として、肥満 (Obesity) を取り上げる (註 1)。世界全体を見渡すと、サハラ以南アフリカや南アジアを中心に、経済的困窮を背景として、世界には未だ飢饉といったように十分な食料摂取が出来ていない「慢性的栄養不足人口」が存在している一方で、先進国においては飽食に伴う食生活の偏りから生じる健康に関わる社会問題が顕在化しており、そうした問題の代表例として、肥満あるいは過体重 (Overweight) が挙げられる。肥満は多くの先進国において問題となっているが、特に、英国や米国では深刻な社会問題となっている。

そして、肥満に関する問題は欧米社会特有のものではなく、日本においても近年問題なりつつある。かつて理想的な PFC 熱量比率 (註 2) であるとされていた日本の食生活は、戦後の高度経済成長に伴って食の欧米化が進展し、高脂質・高タンパクの食事が中心となり、現在に至るまでに PFC 熱量比率の適正水準からの乖離が確認されるようになった。現在の日本の肥満人口比率は米国や英国と比較して低い割合に収まっているが、食の欧米化が進めば、今後の肥満人口の増加も予想される。持続可能な社会を構築する上で、人々の健康に深刻な影響を与えている肥満問題への対応は急務である。日本で今後生じ得る肥満問題への対策を検討する上では、既に多数の肥満人口を抱える英国や米国の状況を参考にすることが有効であろう。

したがって、本章では、肥満が社会問題化している米国と英国での新聞報道記事を取り上げ、肥満に関する社会的関心の所在を分析する。テキストデータの中でも、時事刻々と変化する社会的関心を記録した媒体である新聞記事は、テキストマイニングにおける主要な分析対象の一つであり、社会科学系分野ではしばしば用いられる。新聞には様々な購読者を想定した多様なスタイルが存在するが、知識層向けで、信頼性・社会的信用度が高いとされる新聞は、クオリティペーパーと呼ばれ、学術的な研究においても分析に使用されることが多い。

クオリティペーパーの代表例としては、本稿で分析対象として用いている英国の *The Guardian* 紙や米国の *The New York Times* 紙 (以下、*The* は省略) が挙げられ、他にも、*The Times* や *Le Monde* 等が知られている。そうしたクオリティペーパーを擁する米国および英国が共通して直面している食をめぐる問題として、肥満が存在する。肥満は、例えば広告規制や税制の改正とも相まって、英国と米国において経済的影響が大きい問題とされる (Oommen and Anderson, 2008; Wang et al., 2011)。また、新聞記事を対象に肥満という問題へと接近する研究も多数蓄積されており、Ries et al. (2011) は米国、カナダ、

英国という三ヶ国における肥満に関する政策の報道を分析している。Hilton et al. (2012) は、英国の7つの日刊紙を対象に、肥満に対する報道の推移を見ている。他に、Nimegeer et al. (2019) は、英国の複数の新聞を題材として、児童肥満に焦点を当てている。いずれも複数の新聞社の記事を分析対象としているが、Grounded Theory によるコード化に基づいており、テキスト抽出に恣意性が残る点が課題として挙げられる。更に、十数年程度の一定期間を対象としていることも共通するが、時系列によるトピックの推移が適切に分析されていない点も課題であろう。

## 6.2 方法

### 6.2.1 データ

本章での分析対象として取り上げる新聞記事として、前述の通り、国際的にも代表的な英字のクオリティペーパーとして知られる英国日刊紙 Guardian と米国日刊紙 New York Times とした。両紙ともに両新聞社が提供している API を利用することで、新聞記事データの取得が可能となっている (The Guardian, 2021; The New York Times, 2021)

表 6-1 に示される通り、新聞記事データには記事全文は収録されておらず、その記事の要約とされる Abstract や Trail Text が含まれている。そこで、肥満に関連する記事に焦点を当てるため、“obesity” がそれらの Abstract や Trail Text 中に含まれる記事に分析の対象を絞った。分析対象データの詳細については、表 6-1 に整理した通りとなる。

表 6-1 分析データの概要

項目	The Guardian	The New York Times
対象期間	1999年1月～2020年12月	
カテゴリー	全カテゴリー	
対象テキスト	Headline・TrailText	Headline・Abstract
対象記事	文中にObesityを含む記事	
対象記事数	1619	1266

出所) 取得データにもとづいて筆者作成。

註: 1) 分析対象とした記事は、Headline あるいは記事概要 (TrailText・Abstract) の少なくともどちらかに Obesity を含んだ記事である。

### 6.2.2 ニューラルトピックモデル

本章においては、第 5 章と同様に、トピックモデル実装のツールとして、BERTopic

(Gootendoast 2020; Gootendoast, 2022) を用いた (註 3)。本章の分析では、分析者が任意に設定する必要のあるハイパーパラメータの設定については、こちらも第 5 章と同じく、基本的にデフォルトの値を用いた (註 4)。

### 6.2.3 分析の概要

本章の分析では、BERTopic で実装可能な分析の中でも、クラス分類、トピック間の類似度 (註 5) の算出、時系列トピックモデリングを用いる。ここでのクラス分類は、分析データを入力する時点で、各ドキュメント、すなわち本分析では各新聞記事、がどちらの新聞社であるのかをラベルとして与える処理を指している。このクラス分類を行うことにより、トピックモデルの分析結果において、Guardian と New York Times の両紙のどちらのトピックとして属するのか、区別が可能となる。クラス分類では、Guardian と New York Times という二紙のラベルを各記事に付与し、分析結果のトピック抽出において、両紙の区別を可能とした。

そして、トピック間の類似度の算出は、本分析においては、主に異なる新聞における共通性の高いトピックの特定を目的として利用する。トピック間の類似度は、二つの異なる新聞において共通性の高いトピックの抽出を目的として用いる。

そして、時系列トピックモデリングは各々のトピックにける時間変化による推移を分析することが可能である。つまり、同一トピックにおいて、そのトピックの単語の変化について、時系列をもとにして追うことが可能となる。本分析では、時系列トピックモデルを適用することにより、社会的関心の 21 年間に渡る変化にアプローチする。

## 6.3 結果と考察

### 6.3.1 クラス分類とトピック間類似度

先述の設定で分析を実施した結果、53 のトピックが抽出された。本稿では、53 のトピック全体は論じず、各指標に基づき重要度の高いトピックに議論を限定する (註 6)。まず、出現頻度 (Frequency) 上位 20 のトピックに着目し (註 7)、それらのクラス分類の結果について表 6-2 に示す。

そして、図 6-1 は、これらのトピック同士の類似度をヒートマップとして表したものである。これらの結果から、異なる二紙でありながらも、類似度の高いトピックが存在していることが確認される。

表 6-2 トピックモデルによるクラス分類の結果

番号	トピックに属する主要単語	新聞
0	sugar_drinks_sugary_cocacola	Guardian
1	fat_blame_weight_stigma	NewYorkTimes
2	fat_eating_duchess_shriver	Guardian
3	drug_gene_genetic_pill	Guardian
4	pupils_sport_childhood_children	Guardian
5	schools_milk_vending_lunch	NewYorkTimes
6	uk_nhs_years_obese	Guardian
7	genes_fat_brain_hormone	NewYorkTimes
8	diabetes_disease_dementia_insulin	NewYorkTimes
9	ads_food_ofcom_watchdog	Guardian
10	food_americans_kfc_eating	NewYorkTimes
11	soda_tax_bloomberg_mayor	NewYorkTimes
12	smoking_government_nhs_alcohol	Guardian
13	children_kids_child_parents	NewYorkTimes
14	cancer_cancers_smoking_breast	Guardian
15	birth_babies_pregnancy_breastfed	NewYorkTimes
16	cyclists_healthy_cities_buildings	Guardian
17	takeaways_schools_takeaway_shops	Guardian
18	industry_manufactures_asda_advertising	Guardian
19	sugery_gastric_disability_bariatric	Guardian

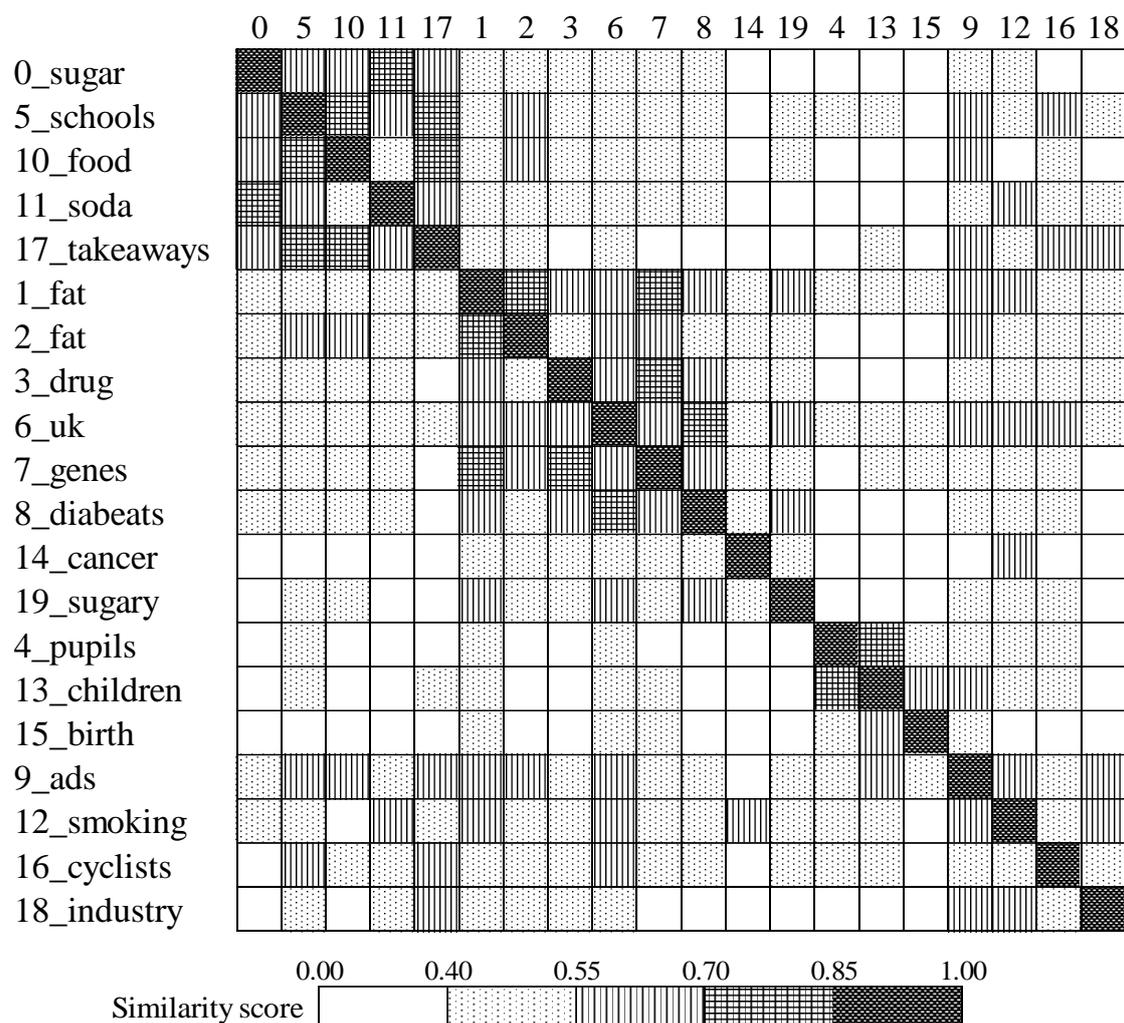
出所) 分析結果より筆者作成.

註 : 1) 表中に記載の主要単語は, class-based -TF-IDF の値が高い単語となる.

註 : 2) トピック番号が 0 から順に, すなわち番号が若い順に Frequency の高いトピックとなる.

各トピックを参照すると, 肥満にまつわる多様な新聞記事, すなわち社会的関心が抽出されたことが確認できる. Frequency, すなわち分類された新聞記事の数が多かったトピックの順に見ていくと, 最も多くの新聞記事が属するトピックは, 砂糖に関するトピックであった.

異紙間で類似度の高いトピックについて, 図 5-1 のヒートマップ と表 5-2 のクラス分類の結果を参照すると, 「5・17 (類似度 0.83)」「0・11 (類似度 0.81)」「3・7 (類似度 0.78)」「4・13 (類似度 0.74)」「1・2 (類似度 0.73)」「6・8 (類似度 0.71)」という 6 つの組を確認することが出来る. 単語を参照すると, 児童や学校に関連するトピックや, 砂糖・ソーダに関連するトピックが, 類似度の高いものとして抽出されたことがわかる.



出所) 分析結果より筆者作成.

註: 1) Frequency の上位 20 トピックの類似度を算出している.

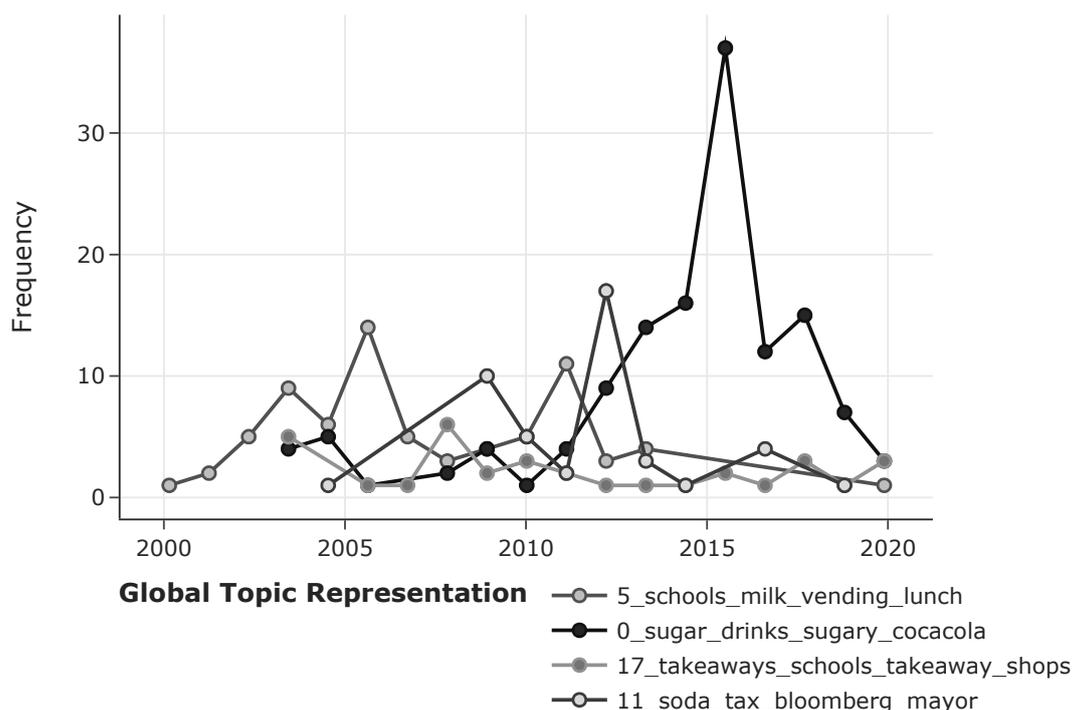
註: 2) Similarity Score はトピック間のコサイン類似度を示す.

図 6-1 頻出上位 20 トピック間の類似度ヒートマップ

### 6.3.2 時系列トピックモデリングと二紙間比較

ここで、異紙間のトピック同士で類似度が 0.8 を超えた「0・11」及び「5・17」に焦点を当て、1999 年から 2020 年の 21 年間に渡る出現頻度の推移に関する時系列トピックモデリングの結果を、図 6-2 に示す。更に、トピック内の時系列推移を確認するため、表 5-3 に、トピック 0 とトピック 11 について、各時点における class-based-TF-IDF の高かった単語を示す。表 6-3 の結果から、トピック 0 は“sugar”に関連するトピックであることは一貫しつつも、各時点で重要度の高い単語が変化していることが確認出来る。トピック 11 も同様に、“soda”以外の単語は移り変わりが見られる。更に、両者には“tax”が共通し

で登場していることから、これらのトピックは、“sugar” および “soda” にまつわる “tax” に関する社会的関心の推移が反映されたものとして解釈出来る。また、両者のトピックに含まれる個々の単語が “tax” を除き、ほとんどが異なる点を考慮すると、ここでの類似性は、単純に単語を比較しただけの類似ではなく、意味を踏まえた類似性であることが確認される（註 8）。



出所) 分析結果より筆者作成。

註：1) 異なる新聞のトピックにおいて、コサイン類似度が 0.8 以上のペアを記載。

註：2) Global Topic Representation は、全期間を通して class-based-TF-IDF の高い単語を表す。

註：3) Frequency はそのトピックに分類された記事数を示す。

註：4) 時系列の区分は 20 分割のため、各時点は単年度とは必ずしも一致しない。

図 6-2 時系列トピックモデリングの結果

### 6.3.3 砂糖税・ソーダ税の経緯と分析結果との照合

分析結果を総合すると、トピック「0・11」は、砂糖税あるいはソーダ税に関するトピックであると考えられる。ここで、これらの砂糖税に関するトピックに関する時系列トピックモデリングの結果と、実際の砂糖税の動向との照合を図る。

まず、英国で砂糖税が導入されたのは、2018年（World Bank, 2020）である。トピック 0 の出現頻度が高くなった 2016 年頃は、砂糖税導入の議論が進んでいた時期と推察される。そして、トピック 11 の 2012 年頃は、他時点には含まれていない “reject” が含まれている

が、2012年はニューヨーク州での砂糖税導入が頓挫した時期である。しかし、出現頻度は高くなく、少なくとも New York Times から見た社会的関心は高く無かったものと推察される。一方で、トピック 0 は Guardian のトピックで最も出現頻度が高く、社会的関心も高かったと考えられる。

以上、本分析結果では、長期間にわたる社会的関心の推移について、最も端的にその時々  
の状況を示していると考えられる単語を抽出することにより、おおまかに状況を把握する  
ことが可能である点が示されたと言える。

表 6-3 各時点におけるトピック 0・11 の重要単語

時点	Topic0 (The Guardian)	Topic11 (The New York Times)
1	-	-
2	-	-
3	-	-
4	-	-
5	tax, kraft, foods, fatties, dronks	-
6	sugar, colas, confectionary, lemonade, frosties	smoking, ban, article, soda, tax
7	pedantic, chocolate, alcohol, sugar, drinks	-
8	-	-
9	pepsi, sweeteners, sweetener, sugar, water	-
10	sugary, sugar, mangold, tax, nhs	tax, drinks, paterson, governor, sugary
11	foods, sweet, sugar, healthier, fat	soda, stamps, sodas, sugared, tax
12	drinks, sugar, alcohol, tax, liver	stamps, mayor, soda, bloomberg, reject
13	sugar, tax, cocacolas, cocacola, milk,	bloomberg, mayor, sugary, soda, tax
14	sugar, tax, sugary, cokes, fruit	mexico, tax, soda, sugary, bloomberg
15	sugar, cocacola, sugary, teaspoons, fizzy	oyster, calif, berkely, soda, economy
16	sugar, tax, drinks, sugary, soft	-
17	sugar, drinks, sugary, confectionery, drink	soda, convicted, sugary, tax, spress
18	sugar, tax, sugary, snacks, confectionery	-
19	sugar, drinks, sugary, taxing, sweets	pediatrics, curbs, soda, taxes, sugary
20	supermarkets, unhealthy, chiles, foodstuffs, sugar	-

出所) 分析結果より筆者作成。

註：1) 空欄は、その時点ではそのトピックに属する記事が検出されなかったことを示す。

#### 6.3.4 考察

本節では、抽出されたトピックの結果と、両国の背景の関係性を中心に考察する。表 6-2 に記載されている各トピックの主要単語を参照すると、“stigma”を含むトピック 1 や、“gene”あるいは“genes”を含むトピック 3 と 7 といったように、様々な単語が登場しており、両国における肥満に関する多様な社会的関心が抽出される結果となった。“stigma”は肥満に対する偏見を表す単語として抽出されたものとして捉えられ、肥満に対する米国の社会的背景が反映されたことが窺える。

抽出されたトピックは、大きく分類すると、児童肥満に関連する社会的関心、糖尿病を意味する“diabetes”を含むトピック 8 のような疾患的側面に対する社会的関心、そして、“tax”を含むトピック 0 のような、肥満に関連する施策面への社会的関心に分かれた。更に、英国王室と関連があると推察されるトピック 2 や、米国の食品企業と関連するであろうトピック 10 など、両国の背景を反映させたトピックについても確認される。したがって、本章でのトピック分析の結果は、両国の文化的背景などが、肥満への社会的関心に対して影響を与えていることを示唆するものと言えよう。

更に、ここまで議論したように、類似度の高いトピックは存在しつつも、図 5-1 に示される通り、多くは低い類似度に留まっており、この点からも、両国における肥満への社会的関心の多様さを窺い知ることが出来る。

#### 6.4 結論と今後の課題

本章では、新聞記事のようなテキストデータに対する分析手法として、特定の問題に関する社会的関心とその変化を捉える手法の一つである分散表現を用いたトピック分析を適用した。具体的には、Guardian と New York Times という二つのクオリティペーパーにおける肥満に関する記事を対象に、異なる二紙間、及び異なる時点間の類似性や変化を分析した。

そして、本章では特に、砂糖税関係と推察されるトピックを例に取り上げ、その結果から、単語の単純比較ではないトピックの意味を踏まえた分析として有用であり、更に、時事刻々と変化する社会的関心を捉えており、特定トピックの変化の分析においても有用であることを示した。したがって、本章で用いたトピックモデルは、新聞の異質性を考慮した分析が可能であり、なおかつ、時系列で変化する社会的関心についても的確に捉えていることがわかった。

本章の分析事例は、上記二紙の記事という限られた情報源を用いたものであり、分析結果はあくまでも新聞というフィルターを通した結果である。両国の社会的関心を総合的に捉えるためには、対象を新聞記事に限ることなく、他の媒体へと拡張する必要があるであ

ろう。したがって、今後の研究として、例えば政府資料や SNS に広げることで、より実用的な分析への拡張を展望したい。

- 註 1) WHO による定義では、Obesity は BMI (Body Mass Index : ボディマス指数) が 30 を超えた状態であり、Overweight は BMI が 25 以上の状態を指す。いずれも成人の基準となる。
- 註 2) PFC 熱量比率は、三大栄養素である Protein, Fat, Carbohydrate の 3 つのバランスを指し示す。理想的な PFC 熱量比率は、P:F:C=13:25:62 (%) とされており、日本の平均的な PFC 熱量比率は、1980 年時点で適正水準に近い比率となっていた。
- 註 3) 本章の分析では、BERTopic ver0.8.1 を利用した。
- 註 4) ハイパーパラメータは任意で設定するが、ここではパラメータ設定の妥当性を検証する材料が無いため、デフォルトとした。なお、第 5 章と同様、事前学習モデルのみ、デフォルトではなく、ハイパフォーマンスなモデルとして評価されている「paraphrase-mpnet-base-v2」を設定した。
- 註 5) トピック同士の類似度は、各トピック、すなわち各クラスターにおけるセントロイド (centroid) のベクトル同士のコサイン類似度により算出される。Angelov (2020) の提案した Top2Vec ではこのセントロイドのベクトルを Topic Vector として定義している。
- 註 6) BERTopic における Frequency は当該トピックに分類されたドキュメント数を指す。したがって、本分析においては記事数を意味する。本章では Frequency を出現頻度と呼ぶ。
- 註 7) BERTopic のトピックモデルでは、トピックとして分類されなかった「その他」を一括したトピックも存在するが、それは含めていない。
- 註 8) トピック「0・11」における“sugar”と“soda”は別単語でありながらトピックの類似度が高いように、トピックの類似度と単語の出現頻度との関係は必ずしも直接的ではない。

## 第7章 保健機能食品に対する消費者の認知構造

### 7.1 はじめに

持続可能な社会を構築していく上では、人々の健康の確保、すなわち健康的な食環境の構築も重要となる。第6章ではそうした健康的な食環境の構築と関連するトピックとして肥満を取り上げたが、本章では、人々の健康増進を推進するための食品を認可するための制度である保健機能食品制度に焦点を当てる。そして、保健機能食品制度の有効活用を推進する際に重要となる、保健機能食品に対する消費者評価について、テキストマイニングの手法を用いることにより、接近する。

保健機能食品制度は、食品が有している健康効果を健康機能表示としての食品表示を許可する制度である。わが国は、1991年に食品に対する健康機能表示を特定保健用食品制度として施行し、世界における健康機能表示の制度化において先駆的な役割を果たした(清水, 2005)。その後、2001年に施行となった栄養機能食品、そして2015年から新たに開始された機能性表示食品と、認証制度の拡充が進んできた。

高齢化が進展し、医療費が財政を圧迫している現在にあつては、医療費の削減、ひいては国民の健康増進は喫緊の課題であり、健康増進に資するとされる保健機能食品が担う役割は大きい。したがって、国民の健康増進を図るためにも、保健機能食品の拡充を進め、制度を効果的に運用していくことが求められている状況にある。つまり、制度拡充による市場規模拡大は、社会的にも意義があると言えるであろう。2015年時点で、特定保健用食品の市場規模は約6400億円とされる(山田ら, 2017)。機能性表示食品が加わった現在では、保健機能食品市場の更なる拡大が期待される状況にある(湯田, 2017)。保健機能表示は食品に関する表示の一事例であるが、そうした食品に関する表示は全般的に大きな関心を集めており、食品表示と消費者行動に関する多くの先行研究が積み重ねられている(氏家, 2016)。機能性表示食品については、伊藤ら(2019)による機能性表示野菜に関する消費者選好を分析した研究や、高田(2018)による機能性表示野菜に対する消費者の支払意思額についての研究が存在している。他にも、広垣(2011)は、特定保健用食品がもたらす消費者の購買行動への影響について論じている。このように、表示が与える消費者の購買行動への変化については知見の蓄積が進む状況にある。

他方、認知構造に迫るための手法として、近年、グラフ理論に基づいたネットワーク分析によるアプローチが試みられており、多様な分野で活用されている。マーケティング分野では、和泉(2014)がマーケティング研究におけるネットワーク分析の活用可能性について先行事例を取り上げ論じており、とりわけ農産物関連では、バレイショを対象とした消費者の認知構造についてネットワーク分析を用いて接近した事例(森嶋, 2019)や、Yano et al. (2018)による消費者の生野菜の摂取動機に対してネットワーク分析を用いた事例が

存在する。

以上の背景を受けて、本章では、国民の健康増進を図る上で重要な役割を果たすと考えられる保健機能食品について、消費者認知の側面から分析を行う。ここでの消費者認知とは、保健機能食品について消費者がどういった印象を抱いているのかを指す。保健機能食品に関する消費者認知に着目した背景として、現在の保健機能食品は、特定保健用食品と機能性表示食品、そして栄養機能食品と、類似した制度が存在している点が挙げられる。消費者にとっては制度ごとの混同を招きかねない状況にあると言えるため、消費者認知に接近することにより、現況を確認することが本研究の課題である。そして、そうした消費者認知への接近として、本研究では2つのアプローチを活用する。まず1つ目のアプローチとして、保健機能食品に含まれる3つの認証制度である特定保健用食品、機能性表示食品および栄養機能食品について、これらに対する消費者の認知を感情分析とネットワーク分析を併用することにより接近する。次に、2つ目のアプローチとして、これら3制度の内、特定保健用食品と機能性表示食品について、それらの認証製品が有するとされる健康効果に着目し、両制度と健康効果との関係性に関する消費者認知について分析する。いずれも分析手法としてネットワーク分析を利用することで、消費者認知の可視化を試みる。

## 7.2 保健機能食品制度に対する消費者認知

### 7.2.1 背景と目的

保健機能食品は、適切な使用の下で表示された特定の健康効果を発揮するものであり、健康効果を得るためには適切な使用を心掛ける必要がある(千葉ら, 2014)。しかしながら、特定保健用食品の利用実態や消費者の認識を調査した先行研究では、消費者による特定保健用食品への理解不足が招く不適切な利用実態が確認されており、効果を実感していない利用者が存在しているという結果が得られている(千葉ら, 2014; 村上ら, 2015)。機能性表示食品を対象とした研究においても同様の結果が得られており、制度を適切に理解していない利用者の存在が確認されている(千葉ら, 2014)。

保健機能食品に関する消費者の認知については多角的な研究が行われており、広垣(2011)は、特定保健用食品に関して、表示自体が消費者の購買行動に対して正の影響を与えていることを明らかにした。また、角田(2011)は特定保健用食品の製品に関する先発優位性について論じており、通常、加工食品分野においては先発優位性の効力は低いものの、国の認定という要素に紐づけられる形で先発ブランドが生まれていることを考察している。他にも、佐藤(2013)は、特定保健用食品の中でもとりわけ「特保飲料市場」に関して、ネットワーク分析による消費者認知解明を試みている。その結果、一部の製品については、「ダイエット」等の健康に関するキーワードと製品名との連想関係を構築することに成功

している点を明らかにし、それらの製品は、特定保健用食品市場における先発ブランドであることを論じている。

以上の通り、消費者の認知に関しては多くの研究例が蓄積されており、特に特定保健用食品を対象とした分析が多いが、保健機能食品が包括する三つの制度全般に対する消費者認知への接近に関する比較分析は十分には行われていない。そこで、本研究では保健機能食品制度に対する消費者認知を明らかとすることを目的として、ブランド認知解明に用いられるネットワーク分析を手法として採用した。更に、消費者認知に対して新たなアプローチから接近するため、テキストから書き手の感情を推定可能な手法である感情分析と、ネットワーク分析とを合わせたアプローチを提案する。分析対象はアンケートの回答テキストとした。

## 7.2.2 方法

### 1) アンケートの概要

本稿でのテキスト解析に用いるアンケートは保健機能食品の各制度に対する考えを尋ねたものであり、予備知識として、消費者庁作成のパンフレットから説明を引用し、回答者へと提示した（消費者庁，2015）。アンケートの設問及び説明文は表 7-1 の通りである。アンケートはマクロミル社によるインターネット調査であり、全国のモニター1040 人から回答を得た。回答者については、男女比、年齢層、在住地域の偏りが無いように調整の上、2019 年 2 月上旬に実施された。回答者の年齢層と居住地域は表 7-2 に示す通りである。

表 7-1 アンケートにおける設問と各制度の説明文

設問	次の各制度に対するあなたのお考えをご自由にお書きください。前問で知らないと答えた制度も、説明を読んでお答えください。（自由記述）	
	特定保健用食品	健康の維持増進に役立つことが科学的根拠に基づいて認められ、「コレステロールの吸収を抑える」などの表示が許可されている食品です。表示されている効果や安全性については国が審査を行い、食品ごとに消費者庁長官が許可しています。
説明	機能性表示食品	事業者の責任において、科学的根拠に基づいた機能性を表示した食品です。販売前に安全性及び機能性の根拠に関する情報などが消費者庁長官へ届け出られたものです。ただし、特定保健用食品とは異なり、消費者庁長官の個別の許可を受けたものではありません。
	栄養機能食品	一日に必要な栄養成分（ビタミン、ミネラルなど）が不足しがちな場合、その補給・補完のために利用できる食品です。すでに科学的根拠が確認された栄養成分を一定の基準量含む食品であれば、特に届出などをしなくても、国が定めた表現によって機能性を表示することができます。

出所) アンケート調査にもとづき筆者作成。

註：1) 各制度の説明文は消費者庁作成パンフレットより引用。

註：2) 設問における「前問」とは、表 7-3 に結果を示した各制度への認知度に関する設問を指す。

表 7-2 回答者の年齢層と居住地域

年齢	調査人数	地域	調査人数
20才～24才	81	北海道	42
25才～29才	127	東北地方	80
30才～34才	107	関東地方	398
35才～39才	101	中部地方	160
40才～44才	107	近畿地方	188
45才～49才	101	中国地方	47
50才～54才	113	四国地方	29
55才～59才	95	九州地方	96
60才以上	208		
合計	1040		1040

出所) アンケート調査にもとづき筆者作成。

註：1) 居住地域については全都道府県を含んでいる。

註：2) 回答者の男女比は 1:1 である。

## 2) 感情分析の概要

感情分析 (Sentiment Analysis) は自然言語処理の分野で発展した手法であり、文章をポジティブとネガティブという極性の観点から捉え、その極性を単語や文章に付与することでテキストに込められた感情を推定する。

本分析では単語に極性の評価値を付与した日本語評価極性辞書を活用し、感情分析を行った (小林ら, 2005; 東山ら, 2008)。なお、感情分析の実装のツールについては Python ライブラリとして提供されている `oseti` を用いた (註 1)。本分析の文章の評価値は、文章に登場する辞書記載の単語の評価値の合計を、同じく辞書記載の単語の数で割ったものとした。ポジティブとして記載されている単語には 1.0 を付与し、ネガティブとして載っている単語には -1.0 を付与し、極性の評価値を計算している。つまり、文章に付与される評価値の範囲は  $-1.0 \leq (\text{評価値}) \leq 1.0$  となる。日本語評価極性辞書に記載が無い単語のみの文章の場合、評価値は 0 となる。また、評価値の算出においては、否定語がある際には極性の評価値を反転させている。

次に、文章の極性の評価値によって、ポジティブ・ニュートラル・ネガティブという三段階の感情に文章を分類した。評価値が  $0.5 < (\text{評価値}) \leq 1.0$  の文章をポジティブ、 $-1.0 \leq (\text{評価値}) < -0.5$  の文章をネガティブとし、 $-0.5 \leq (\text{評価値}) \leq 0.5$  をニュートラルとした。

## 3) ネットワーク分析の概要

本稿でのネットワーク分析は、単語の共起関係に基づいてネットワークを構築する手法を用いる。単語の共起関係に基づいたネットワーク分析は、消費者のブランド認知の構造を捉える手法として活用されており、先述の佐藤 (2013) の研究事例や、森嶋 (2019) によるバレイショに関するブランド連想の消費者認知に接近した研究事例などがそうしたネ

ネットワーク分析活用の研究事例として挙げられる。本研究においても、消費者の認知構造を捉えるための手法としてネットワーク分析を採用することとした。

単語の共起関係を明らかとするためには、まず、第2章で概説した通り、文章を単語へと分かち書きする処理である形態素解析を行う必要がある。本分析では、形態素解析を行うツールとして MeCab を使用した（註2）。そして、分割処理された後の単語をそれぞれノードと見なし、同一文章に出てきた単語について共起関係にあると捉えて、ネットワークのリンクを形成した。そして、本分析に用いる単語は、重要度の高いものを抽出することを目的として、共起回数3回以上のものに限定した。そして、ネットワーク分析用のツールである Gephi (ver0.92) を利用することにより、名詞・形容詞・副詞から構成される無向ネットワークを構築した。更に、分析にあたり、モジュラリティを指標として、リンクの繋がりからグルーピング可能な単語の集まりであるコミュニティを検出した。

そして、ネットワーク分析の結果と感情分析を接合するにあたり、先述した文章の感情の三つの分類、すなわち、ポジティブ、ニュートラル、ネガティブについて、それぞれに分類された文章の数によって、その単語に付与された感情を決定した。つまり、ネットワークを構成する単語が表す感情は、各文章への登場回数に応じて決まり、ポジティブな文章への登場回数が最も多い単語はポジティブとして分類している。これにより、ネットワークにおける単語の繋がりが表している感情を推し量ることが可能となった。

### 7.2.3 結果と考察

#### 1) 認知度に関する回答結果と感情分析の結果

今回のアンケートで同時に実施した各制度への認知度に対する回答結果と、感情分析による文章分類の結果を表7-3に示す。

表7-3 認知度に関する三択式設問への回答と感情分析による文章分類

項目	認知度への回答			感情分析による文章分類		
	良く知っている	聞いたことがある	知らない	ポジティブ	ニュートラル	ネガティブ
特定保健用食品	223 (21.4%)	664 (63.8%)	153 (14.7%)	423 (40.7%)	493 (47.4%)	124 (11.9%)
機能性表示食品	195 (18.8%)	680 (65.4%)	165 (15.9%)	318 (30.6%)	578 (55.6%)	144 (13.8%)
栄養機能食品	161 (15.5%)	619 (59.5%)	260 (25.0%)	320 (30.8%)	610 (58.7%)	110 (10.6%)

出所) アンケート調査と分析結果にもとづき筆者作成。

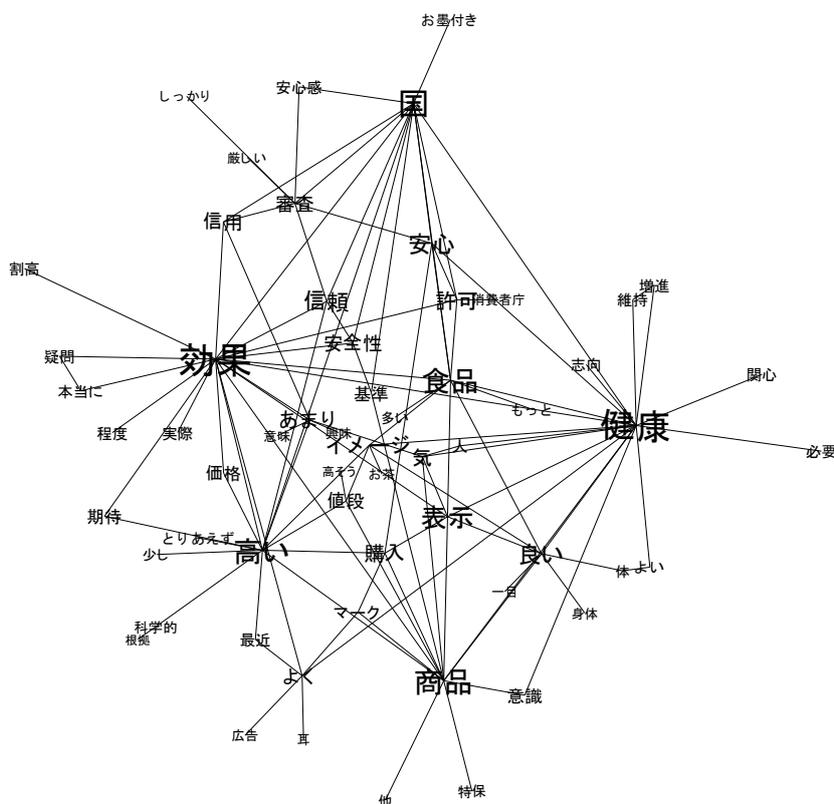
註：1) 認知度への回答の各数値は回答者数。()内は全体に対する割合で小数点第二位を切り上げ

註：2) 感情分析による文章分類の各数値は文章数。()内は全体に対する割合で小数点第二位を切り上げ。

まず、認知度については、保健機能食品の中でも特定保健用食品が最もよく知られており、「知らない」と回答した割合は 14.7%に留まった。この結果は、特定保健用食品と栄養機能食品の認知度を調査した池上ら（2008）の研究結果とも整合的であった。また、いずれの制度においても「聞いたことがある」と答えた回答者が最多となっており、「良く知っている」と答えた回答者は、どの制度においても割合的には少数であることが確認された。したがって、各制度の詳細を認識している消費者は少数に留まっていることが本結果から推察される。この方法に則ってアンケート回答テキストを分類した結果、表の通りどの制度においても概ねポジティブあるいはニュートラルな認知を持たれていることがわかった。特定保健用食品が最も好印象を抱かれているが、共通する点として、ニュートラルな印象を持つ層が最も多いことが判明した。

## 2) 感情分析とネットワーク分析を併せた分析結果

分析結果から得られた各制度に関する設問の回答単語に基づいた単語の共起ネットワーク図について、図 7-1、図 7-2、図 7-3 に示す。



出所) Gephi から出力された分析結果より筆者作成。

註：1) ラベルの大きさは固有ベクトル中心性の値に応じ、大きい程値が高い。

図 7-1 特定保健用食品に対する回答単語の共起ネットワーク図



また、分析で得られたネットワークの構造に関する概要については表 7-4 に整理した通りとなる。各設問の回答結果について、ネットワーク構造としては、特にリンクの数に差が出る結果となった。特定保健用食品のリンクが最も多く、単語の共起関係に富んでいることがわかる。表 7-5 には、名詞の登場回数上位 10 つを設問ごとに整理した。

そして、本稿では単語の重要性を図るためのネットワーク指標として、固有ベクトル中心性を採用し、固有ベクトル中心性の値をネットワークにおける単語の重要度と見なした。

表 7-4 共起ネットワーク分析の概要

項目	特定保健用食品	機能性表示食品	栄養機能食品
ノード数	60	50	59
リンク数	111	70	89
平均次数	3.700	2.800	3.017
モジュラリティ	0.461	0.647	0.441

出所) 分析結果より筆者作成。

註：1) ノード数は単語，リンク数は単語の共起関係を示す。

表 7-5 各設問への回答結果における登場回数の上位 5 単語

項目	特定保健用食品	機能性表示食品	栄養機能食品
登場回数	健康 141	効果 52	栄養 97
	効果 80	健康 41	必要 54
	安心 47	信用 39	食品 42
	商品 44	食品 37	健康 41
	食品 41	表示 31	表示 33

出所) 分析結果より筆者作成。

註：1) 集計の対象は名詞のみ。

註：2) 機能性表示食品の「表示」は、「違い」「安心」と同数であるが、固有ベクトル中心性の値が最も高いため表に記載した。

## 1) 分析結果全体に対する考察

各制度への認知を概観するため、表 7-5 を参照すると、共通して登場する単語はありながらも、多数の異なる単語が上位となっており、制度ごとに異なる認知を持たれていることが推察される。

次に、この認知の差異について接近するため、感情分析とコミュニティ分類の結果を示した表 7-6 に基づきながら考察を論じる。なお、表 7-6 における代表的名詞はそのコミュニティ内で最も固有ベクトル中心性の値が高い名詞である。また、重要度の高いコミュニ

ティに議論を絞るため、代表的名詞の固有ベクトル中心性の値が 0.1 以上のコミュニティのみ記載した。

まず、特定保健用食品においては、「効果」という単語を代表として構成されるコミュニティ T-1 が形成されており、ポジティブな面として効果への期待や興味がある反面、ネガティブな面として効果への疑問が存在することが伺える。そして、T-4 に属する「値段」と「価格」がネガティブ分類となっていることから、特定保健用食品は価格面についてネガティブに認識されていることが確認出来る。更に、それら価格に関する単語が「高い」という単語と同一のコミュニティに分類されている点から、高価格な印象が強いものと推察される。しかしながら、「健康」が代表的名詞である T-2 から推察出来るように、健康への印象は概ねポジティブなものであった。また、T-3 は国の審査に対する信用の高さを示すものと考えられ、T-4 は、本研究で実施した表 7-3 の認知度調査にも裏付けされている通り、特定保健用食品の認知度の高さを表していると考えられる。

次に、機能性表示食品に関して確認すると、最も固有ベクトル中心性の高いコミュニティである K-1 は科学的根拠への関心の高さを示しており、かつ機能性表示食品の科学的根拠に関してはポジティブな印象であることがわかる。K-2 も機能性表示食品が持つ機能性へのポジティブな関心のコミュニティであることが伺える。K-5 は健康に関するコミュニティであるが、健康への良い影響は認識されているものの、「なんとなく」という単語から、その認知はポジティブであるが漠然としたものであることが推察される。K-6 からは、特定保健用食品と同じく効果への疑問が存在していることがわかる。

そして、栄養機能食品について見ると、最も固有ベクトル中心性が高い E-1 は栄養機能食品が持つ栄養の補給という面に関するポジティブな認知のコミュニティであり、手軽、簡単といった他の制度では登場しなかった単語も確認出来る。固有ベクトル中心性の値からわかる通り、栄養機能食品では消費者の関心の多くは栄養にまつわるトピックであることが伺える。

分析結果において特に注目すべき単語は、制度ごとに異なる感情を示している「信用」である。K-3 や E-2 における「信用」はネガティブなものとして認知されている一方、T-3 の信用はポジティブなものとして認知されている。この差異の一因として、T-3 は「国」が代表的名詞となっているコミュニティであることから、「国の審査」が「信用」へと結び付いていると想定され、予備知識として提示した説明文に記載されている審査様式の違いが影響しているとも考えられる。しかし、K-3 の「事業者」、すなわち機能性表示食品における責任の所在である単語がポジティブに捉えられていることを考慮すると、単純に審査様式の違いのみがこの差異に繋がっているとは考えにくく、別の要因についても考察する必要性がある。

表7-6 代表的名詞の固有ベクトル中心性の値及びコミュニティ内の単語

項目	分類	代表的名詞と 固有ベクトル中心性	コミュニティ内の単語
特定 保健 用食品	T-1	効果 1.00	効果 <sup>P</sup> あまり <sup>N</sup> 期待 <sup>P</sup> 本当に <sup>NT</sup> 疑問 <sup>NT</sup> 割高 <sup>NT</sup> 程度 <sup>P</sup> 実際 <sup>P</sup> 興味 <sup>P</sup> 意味 <sup>N</sup>
	T-2	健康 0.91	健康 <sup>P</sup> 商品 <sup>P</sup> 食品 <sup>P</sup> 表示 <sup>P</sup> 良い <sup>P</sup> 気 <sup>NT</sup> 購入 <sup>NT</sup> 意識 <sup>P</sup> マーク <sup>NT</sup> 人 <sup>P</sup>
	T-3	国 0.78	国 <sup>P</sup> 安心 <sup>P</sup> 信頼 <sup>P</sup> 許可 <sup>P/NT</sup> 安全性 <sup>P</sup> 信用 <sup>P</sup> 審査 <sup>P</sup> 基準 <sup>P</sup> 安心感 <sup>P</sup> お墨付き <sup>P</sup>
	T-4	イメージ 0.40	高い <sup>NT</sup> イメージ <sup>P</sup> 値段 <sup>N</sup> 価格 <sup>N</sup> とりあえず <sup>P/NT</sup> 少し <sup>NT/N</sup> お茶 <sup>NT</sup> 高そう <sup>P/NT/N</sup>
機能性 表示 食品	K-1	食品 1.00	食品 <sup>P</sup> 表示 <sup>P/NT</sup> 根拠 <sup>P</sup> 科学的 <sup>P</sup> 商品 <sup>P</sup> 機能 <sup>P</sup> 保健 <sup>NT</sup> 特定 <sup>NT</sup> 購入 <sup>NT</sup> 参考 <sup>NT</sup>
	K-2	機能性 0.91	機能性 <sup>P</sup> 使用 <sup>NT</sup> 成分 <sup>NT</sup> 化学 <sup>NT</sup> 安全性 <sup>P</sup> 高い <sup>P</sup>
	K-3	信用 0.41	信用 <sup>N</sup> あまり <sup>N</sup> 信頼 <sup>P</sup> 企業 <sup>NT</sup> 違い <sup>NT</sup> 事業者 <sup>P</sup> 責任 <sup>NT</sup> 難しい <sup>NT</sup> 特保 <sup>NT</sup> よく <sup>NT</sup>
	K-4	許可 0.33	許可 <sup>NT</sup> 無い <sup>P/NT</sup> 長官 <sup>P</sup> 消費者庁 <sup>NT</sup> 問題 <sup>NT</sup> 個別 <sup>NT</sup> 特に <sup>NT</sup>
	K-5	健康 0.28	健康 <sup>P</sup> 良い <sup>P</sup> 気 <sup>NT</sup> なんとなく <sup>P</sup> 身体 <sup>P</sup> 体 <sup>P</sup>
	K-6	効果 0.25	効果 <sup>P</sup> 期待 <sup>P</sup> 疑問 <sup>NT</sup> 疑わしい <sup>NT</sup> 程度 <sup>P</sup>
栄養 機能 食品	E-1	栄養 1.00	栄養 <sup>P</sup> 食品 <sup>P</sup> 不足 <sup>NT</sup> 補給 <sup>P</sup> ビタミン <sup>NT</sup> 機能 <sup>P</sup> 人 <sup>NT</sup> ミネラル <sup>P</sup> 便利 <sup>P</sup> 成分 <sup>P</sup>
	E-2	必要 0.62	必要 <sup>NT</sup> 摂取 <sup>P</sup> 栄養素 <sup>P</sup> 食事 <sup>NT</sup> 自分 <sup>P/NT</sup> あまり <sup>N</sup> 届け出 <sup>NT</sup> 補完 <sup>P</sup> 普段 <sup>P</sup> 信用 <sup>N</sup>
	E-3	表示 0.29	表示 <sup>NT</sup> 機能性 <sup>NT</sup> 表現 <sup>P</sup> 国 <sup>P</sup> 自由 <sup>P</sup> 届出 <sup>NT</sup>
	E-4	健康 0.19	良い <sup>P</sup> 健康 <sup>P</sup> 体 <sup>P</sup> 身体 <sup>P</sup> 気 <sup>NT</sup> よい <sup>NT</sup>
	E-5	基準 0.12	基準 <sup>NT</sup> 一定 <sup>P</sup> 量 <sup>P/NT</sup> 消費者 <sup>P/NT</sup> 以上 <sup>P/NT</sup> 表記 <sup>NT</sup>

出所) 分析結果より筆者作成.

註: 1) 代表的名詞の固有ベクトル中心性の値は少数第三位を切り上げ.

註: 2) コミュニティ内では, 左から順に固有ベクトル中心性の高い単語を記載し, 単語数が10個以上の際は10個まで記載.

註: 3) 単語が有していると推定される感情は上付き文字で示した. ポジティブはP, ニュートラルはNT, ネガティブはNとなる.

註: 4) 感情が複数記載されている単語は分類の際にそれらの感情の推定された文章数が同一となったものである.

そこで, 要因として, 特定保健用食品が保健機能食品市場において先発優位性を有している可能性があることが考えられる. 特定保健用食品市場における先発ブランドの優位性については佐藤 (2013) や角田 (2011) が論じている通りであり, これらの研究からの知見を援用して, 特定保健用食品が有する保健機能食品市場における制度としての先発優位性の影響を想定することが出来る. K-3に「特保」「違い」という特定保健用食品との違いへと関心を向けた単語の繋がりが有するという点は, 特定保健用食品と機能性表示食品が市場として競合していることを表すものだと考えられ, この点も先発優位性が生じ得る根拠として捉えられる.

また、各制度の共通点として、健康効果への期待と懐疑が同時に存在しているという点が挙げられる。一部の消費者は不適切な利用により効果を実感していないという先行研究の結果を考慮すると、制度への理解不足が懐疑の一因であると推察され、表 7-3 の認知度調査の結果もこの考察と符合する。この点については、消費者への情報提供を徹底し、適切な利用を促す必要があるであろう。

#### 7.2.4 小括

本分析の結果から、保健機能食品は制度としては概ね好意的に消費者から認知されていることが判明したが、一方で効果への懐疑という認知も確認された。効果への懐疑が存在しているという分析結果からは、消費者の制度に対する適切な理解が進んでいるとは言い難い状況が推察される。多数の先行研究において指摘されている通り、消費者への更なる情報提供と制度に関する周知が必要であると考えられる（千葉ら，2014；千葉ら，2017；村上ら，2015；高田，2018）。

更に、消費者認知の分析結果を総合すると、保健機能食品カテゴリーにおける特定保健用食品の制度としての先発優位性が示唆される結果が得られた。今後の研究においては、こうした制度としての先発優位性の検証を進めていく必要があるであろう。それと同時に、こうした先発優位性が機能している製品分野の検証も進めていく必要がある。また、特定保健用食品が制度としての先発優位性を有していると仮定すると、後発の二制度に関しては、特定保健用食品という先発制度の存在を活かした製品戦略を取る必要があると考えられる。したがって、後発制度の効果的な活用の模索を進めることについても、今後の課題として挙げられる。

## 7.3 健康効果と特定保健用食品・機能性表示食品の連想に関する認知分析

### 7.3.1 背景と目的

特定保健用食品と機能性表示食品は、認証方式等の違いはあるものの、共に食品の保健機能や機能性を謳った制度である。そして、健康効果を得るためには、消費者はそれらの製品が効能を発揮する保健機能について理解する必要がある。

しかし、特定保健用食品・機能性表示食品に共通する点として、一定数の消費者に関して、制度への理解不足による不適切な利用実態が報告されている(千葉ら, 2014; 千葉ら, 2017; 松元ら, 2017; 村上ら, 2015)。他にも、制度面の課題についても言及されており、例えば、保健機能発揮のための対象者・摂取量・タイミング・期間などの条件をわかりやすい表現で宣伝・広告する必要性(村上ら, 2015)が挙げられている。

したがって、今後の両制度の効果的な運用を図るためには、具体的な保健機能を含めた消費者の認知構造を把握することが求められる。本研究と関連の強い研究として、ネットワーク分析を保健機能食品に対して適用した佐藤(2013, 2014)による先駆的事例が挙げられる。佐藤は、特定保健用食品の製品群と機能性の概念との関係性に対して、有向ネットワークの枠組みを用いた分析を行っているが、機能性表示食品施行前であったため、対象は特定保健用食品に限られている。

そこで、本課題へのアプローチとして、認知構造への接近において有効な分析手法である共起ネットワーク分析を用いて、保健機能と両制度の結びつきを含めた消費者の一般化された認知構造を明らかにすることを目的として、ネットワーク指標を活用した定量的分析を試みた。

### 7.3.2 方法

本分析では、食品の持つ保健機能を記した機能性の表示文(以下:表示文)と特定保健用食品(以下:特保食品)及び機能性表示食品(以下:機能性食品)の両制度に関して尋ねたアンケートの回答テキストを共起ネットワーク分析の対象とした。これにより、消費者の認知構造に接近する。

そして、本分析では、分析対象とする表示文として、“体脂肪を減らす”(以下:“体脂肪”)、“お腹の調子を整える”(以下:“お腹の調子”)、“血糖値の上昇を抑える”(以下:“血糖値”)、“眼の機能をサポートする”(以下:“眼の機能”)という4つを選定した。ネットワーク分析の実装にはGephi(ver0.92)を用いた。なお、本稿で構築したネットワークは、有向性を持たない無向ネットワークである。

## 1) アンケートの概要

分析に用いたアンケートについては、マクロミル社の全国 1040 名のモニターに対して、インターネット調査により実施した。本アンケートは、回答結果が一般化された消費者認知に近づくよう、回答者の年齢、及び居住地域等に偏りの無い平準的なサンプルとなるように調整の上で、2019 年 10 月に実施した。アンケートの質問内容は表 7-1 に示した通りであり、分析対象は各質問に対する任意・自由回答形式で取得したテキストである。調査対象となったアンケート回答者の年齢層と居住地域については、表 7-8 に示す通りとなる。

表 7-7 アンケートの質問内容詳細

質問内容	提示した質問項目
あなたは次の言葉から何か連想する「食べ物」・「製品」・「イメージ」などがありますか？もしあれば、250 1 字以内でご自由にお書き下さい（文章でなく、単語だけでもかまいません）。また、特に思い付かない場合は、無理に書く必要はありません。	<ul style="list-style-type: none"> <li>・体脂肪を減らす (840)</li> <li>・お腹の調子を整える (936)</li> <li>・血糖値の上昇を抑える (757)</li> <li>・眼の機能をサポートする (888)</li> </ul>
あなたは次の制度から何か連想する「食べ物」・「製品」・「イメージ」などがありますか？もしあれば、250 2 字以内でご自由にお書き下さい（文章でなく、単語だけでもかまいません）。また、特に思い付かない場合は、無理に書く必要はありません。	<ul style="list-style-type: none"> <li>・特定保健用食品（トクホ） (827)</li> <li>・機能性表示食品 (645)</li> </ul>

出所) アンケート調査にもとづき筆者作成。

註：1) 質問 1 については表示文の 4 項目、質問 2 については両制度の 2 項目、計 6 つの回答を得た。

註：2) 各々の有効回答数は () 内に示される通りとなった。ここでの無効回答は入力無し（空白）の回答を示す。分析には各項目の有効回答のテキストを利用した。

表 7-8 回答者の年齢層と居住地域

年齢	調査人数	地域	調査人数
20才～24才	88	北海道	52
25才～29才	120	東北地方	59
30才～34才	104	関東地方	409
35才～39才	104	中部地方	176
40才～44才	101	近畿地方	195
45才～49才	107	中国地方	46
50才～54才	117	四国地方	27
55才～59才	91	九州地方	76
60才以上	208		
合計	1040		1040

出所) アンケート調査にもとづき筆者作成。

註：1) 居住地域については全都道府県を含んでいる。

註：2) 回答者の男女比は 1:1 である。

## 2) 共起ネットワーク分析

単語の共起ネットワーク分析は、単語の共起性に着目してネットワークを構築する手法である。本稿では、一つの回答に含まれる単語同士を共起関係と捉え、その単語をネットワーク分析におけるノード、そして、共起関係をリンクとした。これを各項目の有効回答へと拡張し、それらの回答から単語の共起ネットワークを構築した。この共起ネットワークを分析することにより、消費者の認知構造に対して接近する。回答テキストを単語へと分割するための処理は、形態素解析ツールである **MeCab** を用いた（註 2）。

更に、本分析では表示文と両制度との関係性を明らかにするため、回答テキストのみでなく、表 2-2 に示した各項目、すなわち 4 つの表示文と 2 つの制度名をノードとしてネットワークに組み入れた。したがって、“体脂肪”の項目に対して回答された単語は、“体脂肪”のノードに対して共起関係、すなわちリンクを持つことになる。これにより、各表示文と 2 つの制度に関して、共起ネットワーク分析の枠組みの中で一体的に捉えることが可能となる。

## 3) 分析指標

分析指標として、ネットワーク分析における主要指標の一つである中心性を用いる。本稿では、中心性の中でも、重要ノードとのリンクを有するノードを重要性が高いものとして捉える指標と定義される固有ベクトル中心性（Eigenvector Centrality）を採用する。固有ベクトル中心性は、その性質から、本稿の枠組みにおいては、「各表示文あるいは両制度から想起される印象の強さ」の評価に用いることが出来る。

更に、固有ベクトル中心性と合わせて、もう一つのアプローチとして、ネットワークにおけるコミュニティの検出を試みる。コミュニティとは、ネットワークにおいて緊密なリンクを形成しているノードのグループを指す。本稿では、コミュニティ分割の評価指標としてモジュラリティを採用した（註 3）。

### 7.3.3 結果と考察

まず、アンケートの回答結果には、複数の具体的な製品名が登場したため、それらの製品のカテゴリ、認証制度と有している保健機能・機能性について、表 7-9 に整理した。次に、各質問項目の回答テキストにおける登場回数上位 10 単語について表 7-10 に記載した。そして、全ての回答テキストに基づいて構築した共起ネットワークに対するコミュニティ分割の結果、および各コミュニティ内における固有ベクトル中心性の上位 10 ノードの一覧を表 7-11 に整理した。なお、全体のネットワークのノード数は 700、リンク数は 3073

であり、ノード同士の繋がり の平均数を示すノードの平均次数は 8.78 という結果となった。

表 7-9 回答中に登場した製品名に関する説明

製品	カテゴリー	認証制度	保健機能・機能性
製品名A	飲料	特定保健用食品	体脂肪を減らす
製品名B	飲料	特定保健用食品	体脂肪を減らす
製品名C	飲料	-	-
製品名D	ヨーグルト	-	-
製品名E	指定医薬部外品	-	-
製品名F	飲料	特定保健用食品	血糖値の上昇を抑える
製品名G	サプリメント	機能性表示食品	眼の機能をサポートする
製品名H	飲料	特定保健用食品	体脂肪を減らす
製品名I	第2類医薬品	-	-
製品名J	第2類医薬品	-	-
製品名K	サプリメント	-	-
製品名L	スナックバー	-	-

出所) アンケート回答結果にもとづき筆者作成。

註: 1) 表中の製品名 A~L は表 7-10~7-12 と共通となる。

註: 2) 表中の保健機能・機能性については、実際の表示ではなく本調査で用いた「得られる健康効果のカテゴリー」に準じて記載した。

表 7-10 各質問項目の回答テキストにおける登場回数の上位 10 単語

体脂肪を減らす		お腹の調子を整える		血糖値の上昇を抑える		眼の機能をサポートする		特定保健用食品		機能性表示食品	
登場単語	回数	登場単語	回数	登場単語	回数	登場単語	回数	登場単語	回数	登場単語	回数
特定保健用食品	122	ヨーグルト	364	野菜	54	ブルーベリー	511	お茶	258	ヨーグルト	51
製品名A	96	乳酸菌	146	特定保健用食品	44	ルテイン	52	製品名A	77	お茶	31
お茶	78	ビフィズス菌	125	お茶	42	製品名G	46	健康	38	健康	27
運動	76	製品名E	108	製品名F	29	アントシアニン	35	製品名B	36	体	18
烏龍茶	37	企業名A	47	製品名B	24	目薬	32	製品名H	31	効果	15
製品名B	30	整腸剤	27	製品名A	17	ビルベリー	15	製品名C	31	食品	14
ダイエット	27	製品名J	21	納豆	17	ビタミン	14	効果	24	製品名L	12
カテキン	26	食物繊維	20	トマト	15	製品名K	9	ヨーグルト	22	特定保健用食品	8
製品名H	25	薬	17	食物繊維	14	ベリー	8	体	19	製品名A	8
製品名I	20	製品名D	16	インスリン	13	メガネ	8	緑茶	18	企業名A	7

出所) アンケート回答結果にもとづき筆者作成。

註: 1) 企業名 A は飲料メーカーを示す。

註: 2) 表に記載したのは名詞のみとなる。

註: 3) 「特定保健用食品」、「機能性表示食品」、各製品名は、本文註 2 の通り、表記揺れを統一した登場数である。

表 7-1 1 全体のネットワークに対するコミュニティ分割の結果

コミュニティ1		コミュニティ2		コミュニティ3		コミュニティ4		コミュニティ5	
単語	EC	単語	EC	単語	EC	単語	EC	単語	EC
特定保健用食品	1.00	機能性表示食品	0.89	ヨーグルト	0.68	血糖値の上昇を抑える	0.85	眼の機能をサポートする	0.29
体脂肪を減らす	0.83	食品	0.59	お腹の調子を整える	0.54	野菜	0.31	摂取	0.16
お茶	0.57	健康	0.48	乳酸菌	0.34	製品名F	0.23	いろいろ	0.15
製品名A	0.52	効果	0.38	製品名D	0.27	食事	0.22	血圧	0.13
運動	0.47	イメージ	0.35	企業名A	0.24	大豆	0.21	ビタミン	0.11
烏龍茶	0.29	商品	0.32	飲料	0.24	納豆	0.19	アントシアニン	0.09
製品名B	0.28	機能	0.24	食物繊維	0.23	製品	0.19	成分	0.09
ダイエット	0.28	機能性	0.23	製品名E	0.22	食生活	0.17	簡単	0.07
茶	0.27	食べ物	0.23	青汁	0.22	ジュース	0.16	製品名G	0.07
製品名C	0.27	体	0.23	腸	0.21	改善	0.16	企業名B	0.07

出所) アンケート回答結果にもとづき筆者作成。

註: 1) ネットワーク構築は名詞を対象とした。

註: 2) 表中の EC はホールネットワークにおける固有ベクトル中心性の値を指す。EC の値は少数第 3 位を四捨五入した

註: 3) 接尾語が抽出された場合は除外した。コミュニティ 3 では「糖」「薬」、コミュニティ 4 では「等」「性」「値」、コミュニティ 5 では「目」「剤」が除外の対象となった。

註: 4) 企業名 B は健康食品メーカーを示す。

註: 5) モジユラリティの値は 0.559 と算出された。

## 1) 単語の登場回数について

まず、表 7-10 の単語の登場回数に着目すると、“お腹の調子”では、最多登場の単語として「ヨーグルト」が抽出されており、“眼の機能”では、「ブルーベリー」が登場する結果となった。これは、ヨーグルトとブルーベリーに対する一般的認知が反映された結果と考えられる。“特保食品”においては「お茶」が最多登場単語となっており、“特保食品”はお茶の印象が強いことが伺える。

次に、製品名の登場回数について注目すると、“特保食品”では多数の製品名が登場しており、特保食品に属する具体的製品名が含まれているのに対して、“機能性食品”では、機能性食品に属する具体的製品名は登場しておらず、具体的製品名として、特保食品の製品が登場していることがわかった。この点は、消費者認知において、両制度を混同した結果が反映されている可能性が高いと考えられる。更に、“血糖値”に登場した製品名に注目すると、“体脂肪”の保健機能を持つ製品名 A と製品名 B が含まれていることが確認出来る。すなわち、この結果は、“体脂肪”に対する健康効果を持っている製品について、“血糖値”の健康効果を持つものと誤って認識されている可能性があることを示している。他の保健機能においては、こうした誤認と推察される結果は確認されなかった(註 4)。

## 2) コミュニティ分割について

表 7-11 のコミュニティ分割の結果からは、まず、コミュニティ 1 において、単語の登場回数にも反映されているように、“特保食品”と「お茶」が同じコミュニティとして分割さ

れていることが確認出来る。また，“体脂肪”と“特保食品”が同一のコミュニティとして検出されており，これら 4 つの保健機能の中では，“体脂肪”が最も“特保食品”との結びつきが強いことが明らかとなった。一方で，他の保健機能はそれぞれ独立したコミュニティとして検出された。

更に，“機能性食品”が検出されているコミュニティ 2 に着目すると，表 7-10 の単語の登場回数と同様，具体的な製品名が含まれていないことがわかる。一方，“特保食品”が含まれているコミュニティ 1 には多数の製品名が検出されていることから，この結果は，機能性食品の施行期間が特保食品と比べると未だ短いため，現時点では依然として，具体的な製品名と制度とが結び付いていない可能性を示唆していると考えられる。

### 3) 製品に着目したエゴネットワーク

以上の分析結果から，複数の異なる保健機能との繋がりが示唆された製品名 A と製品名 B の消費者認知に接近する必要があると考えられる。ここで，これら二つの製品名を中心とした消費者認知に接近するためのアプローチとして，エゴネットワーク（Ego Network）の枠組みを利用する（註 5）。エゴネットワークは，特定のノードに着目し，そのノードを中心としたネットワークを指す。中心となるノードはエゴノードと呼ばれる。これら二つの製品名をエゴノードとして設定し，固有ベクトル中心性を算出した。その結果を表 7-12 に示す。更に，各エゴネットワーク図について，図 7-4 と図 7-5 に示す。

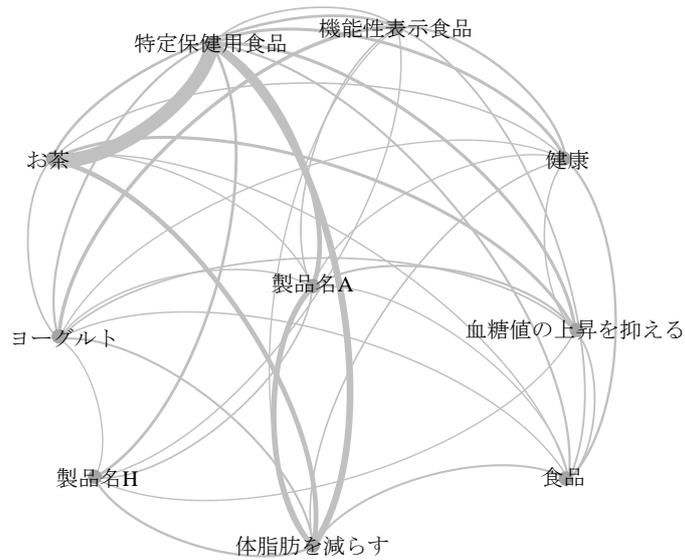
表 7-12 製品名 A と B のエゴネットワークにおける固有ベクトル中心性上位 10 単語

エゴノード	単語	EC
製品名 A	製品名 A	1.00
	特定保健用食品	0.84
	体脂肪を減らす	0.77
	お茶	0.71
	食品	0.68
	ヨーグルト	0.62
	機能性表示食品	0.61
	血糖値の上昇を抑える	0.58
	健康	0.56
	製品名 H	0.45
製品名 B	製品名 B	1.00
	特定保健用食品	0.98
	ヨーグルト	0.87
	体脂肪を減らす	0.80
	製品名 A	0.77
	機能性表示食品	0.75
	血糖値の上昇を抑える	0.74
	緑茶	0.68
	食品	0.68
	お茶	0.68

出所) 分析結果より筆者作成。

註: 1) ネットワーク構築は名詞を対象とした。

註: 2) 表中の EC は各エゴネットワークにおける固有ベクトル中心性の値を指す。EC の値は少数第 3 位を四捨五入した。

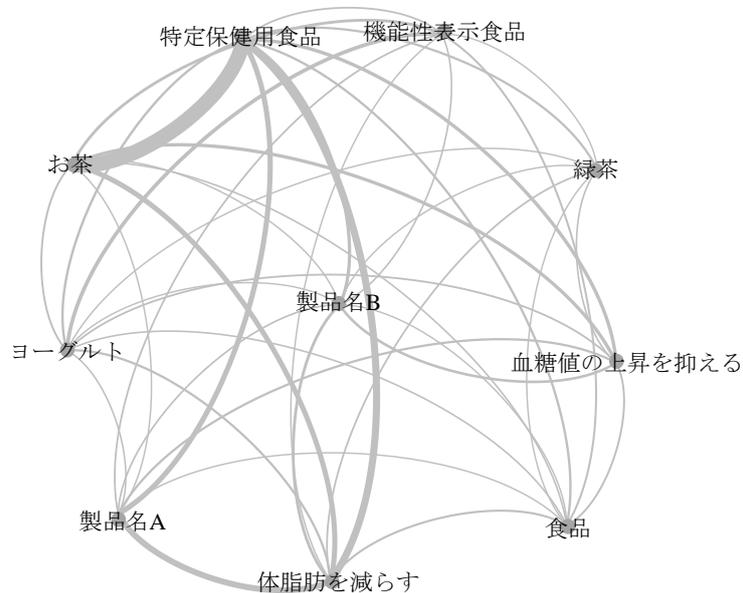


出所) Gephi より出力した分析結果をもとに筆者作成.

註: 1) リンクの太さは共起関係の強さに対応しており, 太いほど強い共起関係を持つ.

註: 2) ネットワーク構築は固有ベクトル中心性の高い上位 10 単語を対象とした.

図 7-4 製品名 A のエゴネットワーク



出所) Gephi より出力した分析結果をもとに筆者作成.

註: 1) リンクの太さは共起関係の強さに対応しており, 太いほど強い共起関係を持つ.

註: 2) ネットワーク構築は固有ベクトル中心性の高い上位 10 単語を対象とした.

図 7-5 製品名 B のエゴネットワーク

### 7.3.4 小括

本章の2つ目の課題においては、特定保健用食品・機能性表示食品という2つの制度について、制度にとって重要である保健機能及び機能性表示を包摂した消費者認知を明らかにすることを目的として、4つの表示文に焦点を当て、アンケート回答テキストに対する共起ネットワーク分析を行った。分析においては、固有ベクトル中心性とコミュニティ分割結果を指標として、消費者の認知構造に接近した。更に、実態と異なる保健機能との連想関係が示唆された製品に着目し、それらのエゴネットワークを構築し、固有ベクトル中心性を算出した。その結果、消費者の認知構造に関して、次の重要な2点が明らかとなった。

まず、1つ目として、消費者認知における両制度の混同の可能性が示された。こうした混同の可能性は、多数の製品が存在する飲料カテゴリーで確認された。

次に、2つ目として、制度の混同に加え、製品が持つ保健機能に対する誤認の可能性が推察された。特に、本分析で確認されたのは、“体脂肪”の製品を“血糖値”の製品として認知しているという誤認の可能性であった。制度の混同と同様、保健機能に対する誤認の可能性も、飲料カテゴリーに属する製品同士で確認された。したがって、これら2制度の活用による国民の健康増進を図る上では、特に製品ラインアップの多いカテゴリーの製品同士で保健機能の誤認が生じる可能性を考慮することが必要であると考えられる。したがって、その点を念頭に置いて、今後の適切な情報提供の整備を進める必要性が高いであろう。

最後に、今後の課題について述べる。本分析で取り扱った表示文は4つと限定的であり、他にも多岐に渡る表示文が存在している。表示文と両制度に関する、より一般化された消費者認知にアプローチするためには、他の表示文に対する分析も必要となるであろう。

更に、本章では消費者の一般化された認知構造に接近するため、消費者の属性および個々の詳細な回答の検討については分析の対象外としたが、消費者属性への接近については、近年注目を浴びる個々人に最適な個別化栄養学、すなわちプレジジョン栄養学の観点からしても重要である。今後は、そういった消費者の異質性を考慮したアプローチなどを行うことによって、個別化栄養学に資する分析についても展望したい。

## 7.4 結論と今後の課題

本章においては、2つの課題から保健機能食品に対する消費者認知に接近した。1つ目の課題では、保健機能食品として認証されている3つの制度に対する消費者認知について、それぞれの制度を比較した。2つ目の課題においては、保健機能食品の中でも、特定保健用食品と機能性表示食品の2つの認証制度と、それらの製品が有する健康効果との関係性を分析対象とした。

これら2つの課題から共通して得られた知見としては、保健機能食品に関する消費者への更なる情報提供の必要性である。保健機能食品制度において認定される健康効果は多岐にわたっており、そうした様々な健康効果を発揮するためには、消費者の理解の醸成が不可欠である。今後の課題として、消費者への適切な情報提供手段の策定に関する研究を進めることが挙げられるであろう。

註 1) 感情分析のためのツールである `oseti` は、<https://github.com/ikegami-yukino/oseti> から利用可能である (2023年2月1日閲覧)。

註 2) 形態素解析の際に必要な辞書は、本章の7.2と7.3のどちらの分析においても、NAIST-Japanese dictionary を利用した。7.2においては、「消費者庁」などはデフォルトの形態素解析では分割されたため、そういった重要単語については、単語が分割されないようにするため、保健機能食品に関わる単語としてユーザー設定にて辞書に追加している。7.3においては更に、アンケート結果に特定保健用食品や機能性表示食品の製品名が多数含まれていたため、そうした保健機能食品に関わる単語について、形態素解析の際に単語が分割されないようにするため、ユーザー設定で辞書に追加した。そして、7.3においては、特定保健用食品と特保・トクホは同一の単語として見なすなど、制度名・製品名については、明らかに同義の単語は1つの単語として統一した。また、7.3で構築した共起ネットワークは、McCabによる品詞分類において名詞として判定された単語のみに限定した。

註 3) ここでのモジュラリティ算出のアルゴリズムは、第3章同様、Gephi に実装されている Louvain Method を用いた。

註 4) 形態素解析の分析結果からは文脈を考慮出来ない点を踏まえると、ここでの単語の登場回数の結果において、誤認以外の要因で単語が登場している可能性は必ずしも排除することはできないが、本分析では、回答テキストにおいて、質問の表示文の保健機能と、異なる保健機能を謳う製品が登場した場合を誤認として捉えた。

註 5) 本分析でのエゴネットワークは、エゴノードと直接繋がりのあるノードから構成されたネットワークを指す。

## 第 8 章 本論文の結論と今後の展望

### 8.1 本論文の結論

#### 8.1.1 各章の要約

本章では、本論文でこれまで論じてきた各章の要約、そして各章を総合して本研究から得られる結論について整理する。その上で、最後に、本研究の今後の展望と残された課題について論じる。

まず、各章の要約について整理する。第 1 章においては、本論文の序論として、近年ますます活発となっている持続可能性に関する取り組みについて概観し、本論文の主要な分析対象となる食品産業および食環境について、持続可能性を念頭に置いた上で構築していく必要性について論じた。そして、そうした持続可能な食品産業および食環境の構築について、様々なデータに基づいて接近する有用性について論じ、そうした多様なデータに対するアプローチとして、データマイニングが有効であることを述べた。

第 2 章においては、本論文で用いたデータマイニング手法の概要について整理した。特に、ネットワーク分析、テキストマイニング及び自然言語処理、科学計量分析という 3 つのアプローチと、本論文で利用している具体的な分析手法や実装のツールをまとめた。

第 3 章以降は、それぞれの研究課題に対してアプローチした研究事例となる。まず、第 3 章では、ASEAN の各国を対象として、各国の食品関連産業に関する産業構造を分析するため、特にそうした食品関連産業に焦点を当てつつ、国家産業連関表の中間取引に対するネットワーク分析によるアプローチを適用した。産業連関表を分析対象とする際に適した分析指標として、Random Walk Centrality と Counting Betweenness という 2 種の中心性を活用した。そして、各国の主要な産業セクターとして、これらの中心性の値の高いセクターを検出した。カンボジアにおいては農林水産業が主要セクターとして検出され、ベトナムでは食品関連産業が主要セクターであることが明らかになった。そして、2005 年時点と 2015 年時点での分析結果を比較すると、各国において主要セクターはほとんど変動していないことがわかった。更に、各セクターの中心性順位に基づいた産業構造の相関性を算出した結果、いずれの国においても、主要セクターと同様、2005 年時点と 2015 年時点で、全体の産業構造の大きな変化は確認されなかった。また、フードシステムに関連する産業セクターの中心性順位にもとづき、ASEAN 各国の中でも産業構造が類似した国々の 3 分類が明らかになった。すなわち、フードシステム関連セクターが高順位となったカンボジア、ベトナム、フィリピンの 3 カ国、低順位となった工業中心の都市国家であるシンガポールとエネルギー資源が主要産業であるブルネイの 2 カ国、そしてインドネシア、マレーシア、タイというフードシステム関連セクターが中順位となった 3 カ国という 3 分類が検

出される結果が得られた。

第4章では、持続可能な農業を構築する上で重要な生産システムとなり得る有機農業に関して、とりわけ社会科学分野における有機農業研究の動向という観点から分析した。分析手法として、学術文献に対して適用可能な科学計量分析を利用した。分析の結果として、第一に、有機農業に関する様々な研究動向やトピックが検出された。具体的には、(1) 事例研究、(2) 生産者に関する研究、(3) 消費者に関する研究、(4) 政策に関する研究、(5) 有機農業の持続可能性ないし環境的側面に関する研究、という5つの研究トピックが主だったものとして抽出された。第二に、カリフォルニアやスペインを中心に多くの地域で有機農業に関するケーススタディが蓄積されており、そうしたケーススタディが盛んな地域の変遷が確認された。すなわち、2000年代においてはカリフォルニアに関する事例研究が盛んに行われており、2010年代においてはスペインに関する事例研究が盛んに取り組まれていることが示唆された。そして、第三に、有機農業の持続可能性に関する研究が近年特に盛んになっており、持続可能性という観点から有機農業への注目が高まっていることが推察された。更に、分析結果を総合すると、有機農業における学術研究の動向の変遷が示されたと共に、これまでの学術文献の蓄積による新たな研究が展開されていることが示唆されたものと考えられる。

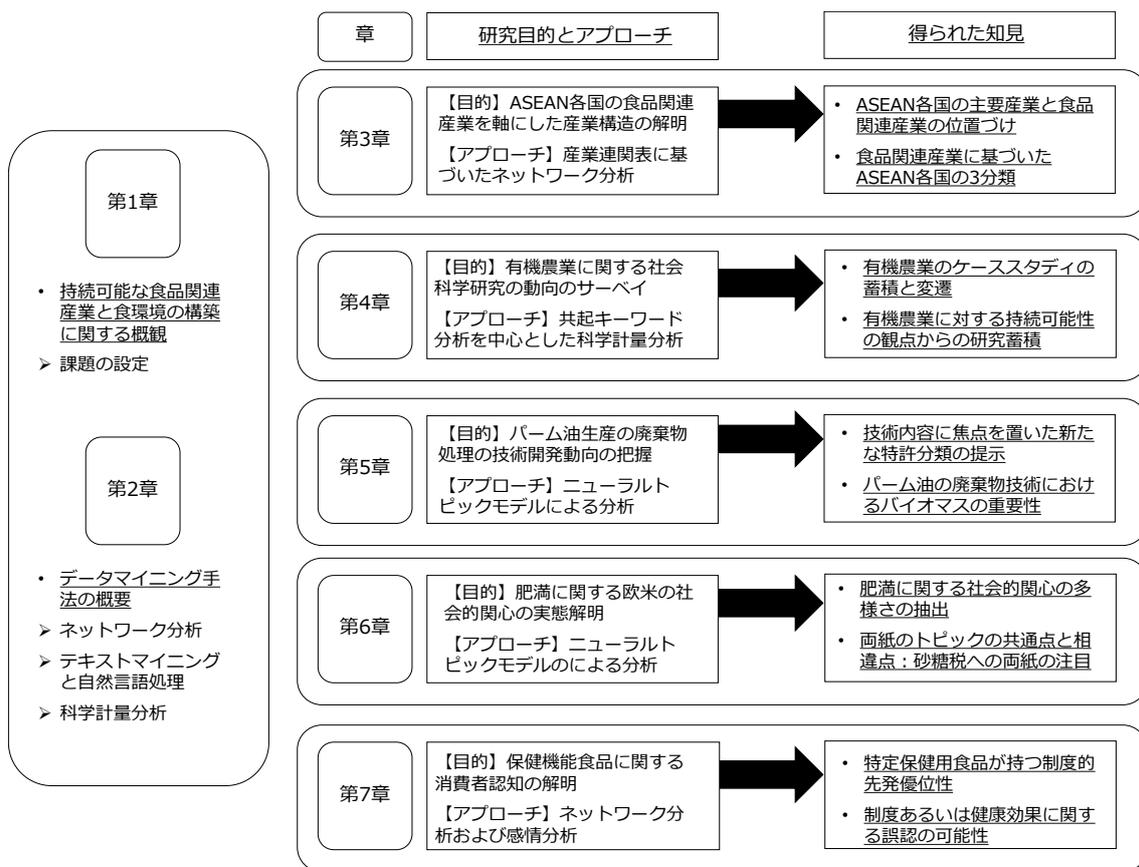
そして、第5章においては、植物性油脂の中でも世界最大の生産量であり、それと同時に様々な問題が指摘されているパーム油に関連する研究課題を取り上げた。具体的には、パーム油の生産過程で生じる廃棄物の処理技術に関する動向について、特許情報にもとづいて分析した。テキストの意味的な側面を含めて分析可能な枠組みであるニューラルトピックモデルを適用した結果、廃棄物処理技術の技術的特徴に基づいた抽出が可能であることが示された。特に、廃棄処理に関連したバイオマス技術に関しては、第2世代バイオマス技術といった具体的な技術的な内容を踏まえた形での分類が抽出される結果となった。こうした技術の内容については、従来の特許分類では十分には接近しにくい点であり、本手法適用の有効性が示された。また、分析対象期間の1995年から2015年の間、パームファイバーを対象とした特許の出版が継続的に行われており、マットレス用途など、パームファイバーに関する技術が広く開発されていることが示唆された一方で、パーム油生産の際に生じる廃液であるPOME (Palm Oil Mill Effluent) については、特許出願は依然として少数となっており、技術開発が進展していない可能性が高いことがわかった。以上、第5章での分析対象は特許情報という限られた範囲ではあるものの、パーム油生産に伴う廃棄物処理技術の開発動向について、一定のインプリケーションが得られる結果となった。

第6章では、持続可能で健康的な食環境を構築する上で、現代社会の大きな問題となっている肥満を取り上げた。具体的には、肥満に関する新聞記事に対してニューラルトピックモデルを適用し、英国での報道と米国での報道の内容について、クオリティペーパーとして知られているThe GuardianとThe New York Timesの記事をデータとして用いて接近し

た。その結果、意味的側面に則った上でトピック間の類似度を定量的に計測することにより、共通して報道されているトピック、そして両者に特有のトピックを明らかにした。また、トピック抽出の全体的な結果としては、肥満のスティグマに関するトピック、児童の肥満に関連したトピックなど、肥満に関する多様な社会的関心が抽出される結果となった。抽出されたトピックを大きく分類すると、児童肥満に関連する社会的関心、糖尿病といった肥満の疾患的側面に対する社会的関心、そして、砂糖税といった肥満に関連する施策面への社会的関心に分かれた。各新聞独自のトピックとして、英国王室と関連があると推察されるトピックや、米国の食品企業と関連するであろうトピックなど、両国の社会的な背景を反映させたトピックについても確認された。第6章では、抽出されたトピックの内、両紙に共通して抽出された砂糖税に関するトピックに着目し、The Guardian と The New York Times における砂糖税に関する記事内容の差異について明らかにした。更に、砂糖税のトピックにおける内容の時系列変化についても検証するため、実際の砂糖税に関する出来事との照合を図った。その結果、実際の出来事と連動してトピックを代表する単語が変化していることを確認し、時々刻々と変化する社会的関心について、正確に捉えていることが示された。第6章で示された分析結果は、ニューラルトピックモデルを用いた新聞記事分析の有用性を示すものであると捉えることが出来る。

第7章では、健康的な食環境を構築する上で一定の役割を果たすと考えられる保健機能食品を取り上げた。保健機能食品制度は、健康効果を有する製品に対する認証制度であり、特定保健用食品、機能性表示食品、栄養機能食品が存在している。第7章では、そうした保健機能食品に対する消費者認知に関するアンケート調査を実施し、調査結果に対して、ネットワーク分析を活用したテキストマイニングを適用した。第7章では2つの課題を設定しており、第1の課題では、保健機能食品を構成する3つの制度に対する消費者認知の比較を行った。分析手法として感情分析を適用することにより、ポジティブとネガティブという尺度の測定を分析に取り込んだ。第2の課題では、特定保健用食品と機能性表示食品が有する健康効果に焦点を当て、そうした健康効果と両制度との認知の関係性について接近した。分析の結果、第1の課題では保健機能食品制度に対する認知度に相関した信用度の差異が明らかになり、健康効果への懐疑といった認知についても確認された。更に、第2の課題では、「体脂肪を減らす」といった健康効果に関する表示から消費者が連想する製品やイメージなどが明らかになっただけでなく、特定保健用食品と機能性表示食品に関する両制度同士の混同、そして各製品が有する健康効果に対する混同、すなわち健康効果に対する誤認の可能性が指摘される結果が得られた。これらの2つの分析結果に共通する知見として、消費者への更なる情報提供が必要であり、特に各製品が有している健康効果に関する適切な発信が要ることがわかった。

以上の各章の概要と得られた知見について、図8-1に整理した。



出所) 筆者作成.

図 8-1 各章の概要と得られた知見の要約

### 8.1.2 データマイニングによるアプローチに関する考察

以上の各章の分析においては、いずれもデータマイニングによるアプローチを採用していることから、各手法の特徴について整理するため、表 8-1 に、手法面から見た各章の要約について整理した。第 3 章では、ネットワーク分析における代表的な分析指標である中心性とコミュニティ検出を利用した。第 4 章では、科学計量分析で活用されるキーワードの共起ネットワーク分析、バースト検出を用いた。第 5 章と第 6 章では共通の手法を用いており、テキストの意味的側面を含意した分析が可能な、近年開発されたモデルであるニューラルトピックモデルを適用した。第 7 章においては、テキストに対する共起ネットワークによる分析に加えて、感情分析による分析の精緻化を図った。

そして、データマイニングによるアプローチは、人力での把握が困難な量のデータを概観する際にも有効であり、産業連関表をデータとして用いた第 3 章を除く残りの章、すなわち第 4 章～第 7 章のいずれの章においても、そういった人力での把握が困難な量のデータに対する分析を行った。

表 8-1 に整理した通り、本論文で用いたアプローチは、大別すると、重要な要素の推定、グループの検出、トレンドの検出、といったものにわけられると考えられる。重要な要素の推定とデータのグルーピングはいずれの分析においても共通して活用されており、トレンドの検出については、時系列での分析が可能なデータに対して共通して活用されている。

更に、各手法の適用対象について比較すると、ネットワーク分析は産業連関表、様々なテキストといったように、ネットワーク形式として捉えることの可能なデータに対しては非常に幅広く適用可能であることが確認出来る。本論文においては、産業連関表、更にはテキストデータに対して適用しており、グルーピングや重要要素の測定などに活用可能であった。

一方で、第 5 章と第 6 章において用いた自然言語処理の手法であるニューラルトピックモデルはテキストに特化した分析手法であり、適用範囲はテキストに限られる。しかしながら、ニューラルトピックモデルは、ネットワーク分析では捉えることの出来ないテキストの意味的な側面についても分析可能であり、トピック間の類似度の測定や時系列の変遷の抽出などの様々な高度な分析についても実行することが出来た。

したがって、いずれの手法についても、適用範囲や分析可能な範囲について、それぞれ利点を持つことがわかった。食品産業や食環境に関わるデータについては、こうした各種データマイニング手法の適用が有効なデータは幅広く存在していると考えられる。分析対象のデータや分析の目的に応じて、各手法の特徴を考慮した上で、こうした各種手法を組み合わせて利用していくことが重要であると考えられる。

表 8-1 本論文で利用したデータマイニング手法の結果と比較

章	第3章	第4章	第5章	第6章	第7章
主要アプローチ	ネットワーク分析	科学計量分析	自然言語処理・テキストマイニング		自然言語処理・ テキストマイニング ネットワーク分析
分析手法	中心性 クラスタリング	バースト検出 クラスタリング	ニューラルトピックモデル		中心性 クラスタリング 共起ネットワーク 感情分析
分析対象	産業連関表	学術文献	特許情報	新聞記事	アンケート回答
分析の特徴	重要要素の推定 コミュニティの検出 データ間の比較	重要要素の推定 コミュニティの検出 トレンドの検出	重要要素の推定 コミュニティの検出 トレンドの検出	重要要素の推定 コミュニティの検出 トレンドの検出	重要要素の推定 コミュニティの検出 データ間の比較

出所) 筆者作成.

### 8.1.3 本論文の結論

以上の通り、本論文では、人々の健康への貢献と持続可能な発展を達成するような食品産業を構築していく上で重要となる各課題に焦点を当て、データマイニング手法の適用による解析を図り、その構築要因について接近した。本節では、本論文の各章の分析結果や考察にもとづき、本論文の結論について論じる。

まず、各章で取り上げた分析事例は限定的な対象・範囲であるものの、持続可能な食品産業および食環境の構築について、一定のインプリケーションを有する結果が得られたと考えられる。食品産業という観点においては、ASEANの産業構造について食品関連産業を中心に接近した第3章の分析結果から、国ごとに差異はあるものの、食品産業は他産業との結びつきも強く、食品産業で生じた経済効果などが他産業にも大きな影響を及ぼすことが示唆された。したがって、持続可能な食品産業および食環境の構築に向けた取り組みは、他産業にも大きく影響が波及すると考えられる。そういった点で、第4章で取り上げた有機農業の推進は、農業のみならず、食品産業を通して、産業全体に影響を与える可能性があるものとして捉えられ得る。他産業に影響を及ぼすという点では、有機農業に限らず、第5章で取り上げたパーム油産業も同様であり、パーム油産業をはじめとして、産業全体で環境負荷の低減に向けた廃棄物処理技術の開発を進めていく必要があるであろう。また、第5章においては、パーム油の生産過程で生じる廃棄物処理について、技術開発がパームファイバーといった一部の廃棄物に偏っていることが示唆されたが、バイオマスに関する技術開発の進展など、環境負荷低減を念頭に置いた開発が進んでいることも分析結果に反映された。そして、持続可能な食環境の構築に向けた観点では、第6章において肥満にまつわる社会的関心を分析する中で、多様な社会的関心が抽出され、健康的な食環境を構築する上で、そうした社会的関心を捉えた上で、多彩な面からアプローチする必要性が示唆された。更に、第7章では、人々の健康増進を図る上で重要となる保健機能食品に関する消費者認知に焦点を当てた分析を行い、その結果、消費者への更なる情報提供の必要性が明らかになった。

本論文の分析結果を総合すると、持続可能な食品産業および食環境を構築するためには、食品企業や消費者を含めたあらゆるステークホルダー、すなわち社会全体で持続可能性を高めるような取り組みを推進していく必要があると考えられる。そして、少なくとも本論文で取り上げた各研究事例においては現状では十分な協調が図られている状況ではないと言え、今後、取り組みを加速化していく必要があると考えられる。そして、そうした取り組みを加速化させるためには、学術的な側面においては、本論文で用いたようなデータマイニング手法などを駆使して、多面的に現状を捉えることが有効であろう。

以上、各章で取り上げた分析事例は限定的な対象・範囲であるものの、持続可能な食品産業および食環境の構築について、一定のインプリケーションを有する結果が得られたと

考えられる。

## 8.2 今後の展望と残された課題

最後に、本研究の限界と残された課題、そして、今後の研究の展望について論じる。まず、本研究の残された課題として、食品関連産業や食環境の総体的な傾向の知見については得られたものの、本分析から示唆される更なる詳細な課題については、十分に接近出来ていないため、そうした個別課題へのアプローチが挙げられる。

また、本研究に残された課題として、手法的な観点に関しては、データマイニング手法の比較について、本論文で論じた範囲は、定性的な議論に留まっており、十分な検討が出来ていない。データマイニングといった分析の手法は、定量的な指標をもって比較されることが望ましいと考えられるため、今後は、分析手法の定量的な比較を展望したい。

そして、本論文の各章の結果から、持続可能な食品産業や食環境の構築に向けたインプリケーションが得られたと考えられるが、その一方で、本研究の限界として、分析対象が限定的であるという点が挙げられるであろう。多様なデータからの接近という観点では、データベースが整っている料理レシピに関するデータや、近年蓄積の著しいソーシャルメディア、あるいは SNS に関するデータなど、分析を適用すべきデータが数多く存在しており、そうしたデータについても取り扱う必要性は高い。また、持続可能な食環境の構築という観点からは、昆虫食や Plant based food と呼ばれるような代替食品の普及可能性や消費者認知、そして代替食品に関連する技術開発の動向についても接近する必要があると考えられる。

したがって、今後の研究課題として、そういった新たなデータや新たな研究課題を対象とした分析についても進め、より多角的な視点から持続可能な食品産業および食環境の構築に向けた解析が必要であると考えられる。

## 引用文献

- Aria, M., and Cuccurullo, C. (2017) bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4):959-975. <https://doi.org/10.1016/j.joi.2017.08.007>.
- Angelov, D. (2020) Top2Vec: Distributed Representations of Topics, *arXiv Preprint*, <https://arxiv.org/abs/2008.09470>.
- Ahn, Y.Y., Ahnert, S., Bagrow, J., and Barabási, A.L. (2011) Flavor network and the principles of food pairing. *Scientific Reports*, 1: 196. <https://doi.org/10.1038/srep00196>.
- Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003) Latent Dirichlet Allocation, *Journal of Machine Learning Research*, 3: 993-1022.
- Blei D. M. and Lafferty J. D. (2006) Dynamic topic models, *Proceedings of the 23rd international conference on Machine learning*, 113-120. <https://doi.org/10.1145/1143844.1143859>.
- Blöchl F., Fabian J. Theis, Fernando Vega-Redondo, and Eric O’N. Fisher (2011) Vertex centralities in input-output networks reveal the structure of modern economies, *Physical Review E*, 83(4): 046127. <https://doi.org/10.1103/physreve.83.046127>.
- Blondel V.D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre (2008) Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*: P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
- Campello R.J.G.B., Moulavi D., and Sander J. (2013) Density-Based Clustering Based on Hierarchical Density Estimates. In: Pei J., Tseng V.S., Cao L., Motoda H., Xu G. (eds) *Advances in Knowledge Discovery and Data Mining. PAKDD 2013. Lecture Notes in Computer Science*, 7819. [https://doi.org/10.1007/978-3-642-37456-2\\_14](https://doi.org/10.1007/978-3-642-37456-2_14).
- Cerina F, Zhu Z, Chessa A, Riccaboni M (2015) World Input-Output Network, *PLoS ONE*: 10(7): e0134025. <https://doi.org/10.1371/journal.pone.0134025>.
- Caron, P., Ferrero y de Loma-Osorio, G., Nabarro, D., Hainzelin, E., Guillou, M., Andersen, I., Arnold, T., Astralaga, M., Beukeboom, M., Bickersteth, S., Bwalya, M., Caballero, P., Campbell, B.M., Divine, N., Fan, S., Frick, M., Friis, A., Gallagher, M., Halkin, J.P., Hanson, C., Lasbennes, F., Ribera, T., Rockstrom, J., Schuepbach, M., Steer, A., Tutwiler, A., and Verburg, G. (2018) Food systems for sustainable development: proposals for a profound four-part transformation, *Agronomy Sustainability*, 38: 41. <https://doi.org/10.1007/s13593-018-0519-1>.
- Chen, C. (2004) Searching for intellectual turning points: Progressive Knowledge Domain Visualization, *Proceedings of the National Academy of Sciences of the United States of America*, 101 (Suppl. 1): 5303-5310. <https://doi.org/10.1073/pnas.0307513100>.
- Chen, C. (2006) CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature, *Journal of the American Society for Information Science and Technology*,

- 57(3): 359-377. <https://doi.org/10.1002/asi.20317>.
- Chen, C., Ibekwe-Sanjuan, F., and Hou, J. (2010) The Structure and Dynamics of Co-Citation Clusters: A Multiple-Perspective Co-Citation Analysis, *Journal of the Association for Information Science and Technology*, 61(7): 1386-1409. <https://doi.org/10.1002/asi.21309>.
- Chen H., Zhang G., Zhu D., and Lu J. (2017) Topic-based technological forecasting based on patent data: A case study of Australian patents from 2000 to 2014, *Technological Forecasting and Social Change*, 119: 39-52. <https://doi.org/10.1016/j.techfore.2017.03.009>.
- Chen, L., Shang, W., Yang, G., and Zhang, J. (2016) A topic model integrating patent classification information for patent analysis, *Computers & Industrial Engineering*, 147: 123-126. <https://doi.org/10.1016/j.cie.2020.106636>.
- Chen, H., Engkvist, O., Wang, Y., Olivecrona, M., Blaschke, T. (2018) The rise of deep learning in drug discovery, *Drug Discov Today*, 23(6): 1241-1250. <https://doi.org/10.1016/j.drudis.2018.01.039>.
- Contreras Martha G. Alatraste and Fagiolo Giorgio (2014) Propagation of Economic Shocks in Input-Output Networks: A Cross-Country Analysis, *Physical Review E*, 90: 062812. <https://doi.org/10.1103/physreve.90.062812>.
- Clark, M., Macdiarmid, J., Jones, A.D., Ranganathan, J., Herrero, M., Fanzo, J. (2020) The Role of Healthy Diets in Environmentally Sustainable Food Systems, *Food and Nutrition Bulletin*, 41(2\_suppl):31S-58S. <https://doi.org/10.1177/0379572120953734>.
- Crowder, D. W. and Reganold, J. P. (2015) Financial competitiveness of organic agriculture on a global scale, *Proceedings of the National Academy of Sciences of the United States of America*, 112(24): 7611-7616. <https://doi.org/10.1073/pnas.1423674112>.
- Dandan T., Pengkun Y., and Hao F. (2020) Utilization of text mining as a big data analysis tool for food science and nutrition, *Comprehensive Reviews In Food Science and Food Safeties*, 19(2): 875-894. <https://doi.org/10.1111/1541-4337.12540>.
- Devlin J., Chang M.W., Lee K., and Toutanova K. (2018) BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *arXiv Preprint*, <https://arxiv.org/abs/1810.04805>.
- DePaolis, F., Murphy, P. & De Paolis Kaluza, M.C. (2022) Identifying key sectors in the regional economy: a network analysis approach using input–output data, *Applied Network Science*, 7, 86. <https://doi.org/10.1007/s41109-022-00519-2>.
- Domínguez, A. and Mendez, C. (2019) Industrial Productivity Divergence and Input-Output Network Structures: Evidence from Japan 1973–2012, *Economies*:7(2):52. <https://doi.org/10.3390/economies7020052>.
- Dwivedi, S.L., van Bueren R.T.L., Ceccarelli, S., Grando, S., Upadhyaya, H.D., Ortiz, R. (2017) Diversifying Food Systems in the Pursuit of Sustainable Food Production and Healthy Diets,

- Trends in Plant Science*, 22(10): 842-856. <https://doi.org/10.1016/j.tplants.2017.06.011>.
- Fanzo, J., Bellows, A.L., Spiker, M.L., Thorne-Lyman, A., and Bloem, M.W. (2021) The importance of food systems and the environment for nutrition, *The American Journal of Clinical Nutrition*, 113(1): 7–16. <https://doi.org/10.1093/ajcn/nqaa313>.
- Feenstra, G. (2002) Creating space for sustainable food systems: Lessons from the field, *Agriculture and Human Values*, 19: 99-106. <https://doi.org/10.1023/A:1016095421310>.
- Food and Agriculture Organization of the United Nations (FAO) 「 FAOSTAT 」 , <https://www.fao.org/faostat/en/> (2023年2月1日閲覧).
- García Muñiz A, Morillas Raya A, Ramos Carvajal C. (2008) Key sectors: A new proposal from network theory, *Regional Studies*, 42(7): 1013-1030. <https://doi.org/10.1080/00343400701654152>.
- García Muñiz A, and Ramos Carvajal C (2015) Input-Output Linkages and Network Contagion in Greece: Demand And Supply View, *Applied Econometrics and International Development*:15(2): 35-52.
- Garg, N., Schiebinger, L., Jurafsky, D., and Zou, J. (2018) Word embeddings quantify 100 years of gender and ethnic stereotypes, *Proceedings of the National Academy of Sciences of the United States of America*, 115(16): 3635-3644. <https://doi.org/10.1073/pnas.1720347115>.
- Go, D.J, Promentilla, M.A., Aviso, K, Yu, K.D. (2019) The Evolution of the Key Sectors in the Philippine Economy Using an AHP-Based Sector Prioritization Index, *Economies*: 7(3): 78. <https://doi.org/10.3390/economies7030078>.
- Godfray, H.C.J., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Nisbett, N., Pretty, J., Robinson, S., Toulmin, C., and Whiteley, R. (2010) The future of the global food system, *Philosophical Transactions of The Royal Society B*, 365: 2769–2777. <https://doi.org/10.1098/rstb.2010.0180>.
- Gurgul H, Lach Ł (2018) Some remarks on a social network approach to identifying key sectors, *Economic Systems Research*: 30(1): 120-135. <http://dx.doi.org/10.1080/09535314.2017.1365048>.
- Grootendorst, M. (2022) BERTopic: Neural topic modeling with a class-based TF-IDF procedure, *arXiv Preprint*, <https://arxiv.org/abs/2203.05794>.
- Grootendorst, M. (2020) BERTopic: Leveraging BERT and c-TF-IDF to create easily interpretable topics, *Zenodo*, <https://doi.org/10.5281/zenodo.4381785>.
- Hansen S.B., Padfield P., Syayuti K., Evers S., Zakariah Z., and Mastura S. (2015) Trends in global palm oil sustainability research, *Journal of cleaner production*, 100, pp.140-149.
- Herrero, M., Thornton, P.K., Mason-D’Croz, D., Palmer, J., Benton, T.G., Bodirsky, B.L., Bogard, J.R., Hall, A., Lee, B., Nyborg, K., Pradhan, P., Bonnett, G.D., Bryan, B.A., Campbell, B.M., Christensen, S., Clark, M., Cook, M.T., de Boer, I.J.M., Downs, C., Dizyee, K., Folberth, C.,

- Godde, C.M., Gerber, J.S., Grundy, M., Havlik, P., Jarvis, A., King, R., Loboguerrero, A.M., Lopes, M.A., McIntyre, C.L., Naylor, R., Navarro, J., Obersteiner, M., Parodi, A., Peoples, M.B., Pikaar, I., Popp, A., Rockström, J., Robertson, M.J., Smith, P., Stehfest, E., Swain, S.M., Valin, H., van Wijk, M., van Zanten, H.H.E., Vermeulen, S., Vervoort, J., and West P.C. (2020) Innovation can accelerate the transition towards a sustainable food system, *Nature Food*, 1: 266–272. <https://doi.org/10.1038/s43016-020-0074-1>.
- Hilton S., Patterson C., and Teyhan A. (2012) Escalating Coverage of Obesity in UK Newspapers: The Evolution and Framing of the “Obesity Epidemic” From 1996 to 2010, *Obesity*, 20(8):1688-1695. <https://dx.doi.org/10.1038%2Foby.2012.27>.
- Jose Ramon Saura, Domingo Ribeiro-Soriano, Daniel Palacios-Marqués (2021) From user-generated data to data-driven innovation: A research agenda to understand user privacy in digital markets, *International Journal of Information Management*, 60: 102331, <https://doi.org/10.1016/j.ijinfomgt.2021.102331>.
- José, L. A., José Luis Aleixandre-Tudó, Máxima Bolaños-Pizarro, and Rafael Aleixandre-Benavent (2015) Mapping the scientific research in organic farming: a bibliometric review, *Scientometrics*, 120(1): 295–309. <http://dx.doi.org/10.1007/s11192-015-1677-4>.
- José, L. A., Lourdes Castelló-Cogollos, José Luis Aleixandre, and Rafael Aleixandre-Benavent (2018) Bibliometric and Social Network Analysis in Scientific Research on Precision Agriculture, *Current Science*, 115(9): 1653-1667. <http://dx.doi.org/10.18520/cs/v115/i9/1653-1667>.
- Kenneth, B., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., and Matsuo A. (2018) quanteda: An R package for the quantitative analysis of textual data, *Journal of Open Source Software*, 3(30): 774. <https://doi.org/10.21105/joss.00774>.
- Kleinberg, J. (2002) Bursty and hierarchical structure in streams, In Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (Edmonton, Alberta, Canada, 2002), ACM Press, 91-101. <https://doi.org/10.1145/775047.775061>.
- Kozłowski, A. C., Taddy, M., and Evans, J. A. (2019) The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*, 84(5): 905–949. <https://doi.org/10.1177/0003122419877135>.
- Lozano, S., L. Calzada-Infante, B. Adenso-Díaz & S. García (2019) Complex network analysis of keywords co-occurrence in the recent efficiency analysis literature, *Scientometrics*, 120(2): 609–629. <https://doi.org/10.1007/s11192-019-03132-w>.
- Liqiang N. Xinyu Dai; Jianbing Zhang; Jiajun Chen (2015) Topic2Vec: Learning distributed representations of topics, *International Conference on Asian Language Processing (IALP)*, 193-196. <https://doi.org/10.1109/IALP.2015.7451564>.

- McInnes, L., Healy, J., and Melville, J. (2018) UMAP: uniform manifold approximation and projection for dimension reduction, *arXiv Preprint*, <https://arxiv.org/abs/1802.03426>.
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean J. (2013) Distributed Representations of Words and Phrases and their Compositionality, *arXiv Preprint*, <https://arxiv.org/abs/1310.4546>.
- Manuelian, C. L., Penasa, M.; da Costa, L.; Burbi, S.; Righi, F.; De Marchi, M. (2020) Organic Livestock Production: A Bibliometric Review, *Animals*, 10(4):618. <https://doi.org/10.3390/ani10040618>.
- Mingers, J. and Leydesdorff, L. (2015) A review of theory and practice in scientometrics, *European Journal of Operational Research*, 246(1): 1-19. <https://doi.org/10.1016/j.ejor.2015.04.00>.
- Muller, A., Christian Schader, Nadia El-Hage Scialabba, Judith Brüggemann, Anne Isensee, Karl-Heinz Erb, Pete Smith, Peter Klocke, Florian Leiber, Matthias Stolze and Urs Niggli (2017) Strategies for feeding the world more sustainably with organic agriculture, *Nature Communication*, 8(1): 1290. <https://doi.org/10.1038/s41467-017-01410-w>.
- Ma W., Luo X., Xuan J., Xue R., and Guo Y. (2017) Discover semantic topics in patents within a specific domain, *Journal of Web Engineering*, 16(7&8): 653-675.
- Nimegeer A., Patterson C., Hilton S. (2019) Media framing of childhood obesity: a content analysis of UK newspapers from 1996 to 2014, *BMJ Open*, 9(4): e025646. <http://dx.doi.org/10.1136/bmjopen-2018-025646>.
- Newman, M.E.J. (2006) Modularity and community structure in networks, *Proceedings of the National Academy of Sciences of the United States of America*, 103(23): 8577-8582. <https://doi.org/10.1073/pnas.0601602103>.
- OECD 「STAN データベース」, <http://www.oecd.org/sti/ind/input-outputtables.htm> (2023年2月1日閲覧).
- Oseghale, S.D., Mohamed, A.F., and Chikere, A.O. (2017) Status Evaluation of Palm Oil Waste Management Sustainability in Malaysia, *OIDA International Journal of Sustainable Development*, 10(12): 41-48.
- Oldham S., Fulcher B., Parkes L., Arnatkeviciūtė A., Suo C., and Fornito A. (2019) Consistency and differences between centrality measures across distinct classes of networks, *PLoS ONE*: 14(7): e0220061. <https://doi.org/10.1371/journal.pone.0220061>.
- Oommen, V.G. and Anderson, P.J. (2008) Policies on Restriction of Food Advertising During Children's Television Viewing Times: An International Perspective, *Proceedings Australian College of Health Service Executives 2008 Conference*.
- Özgür, A., Özgür, L., and Güngör, T. (2005) Text Categorization with Class-Based and Corpus-Based Keyword Selection, *Computer and Information Sciences, ISCIS 2005. Lecture Notes in Computer Science*, 3733: 606-615. [https://doi.org/10.1007/11569596\\_63](https://doi.org/10.1007/11569596_63).

- Qi-Qi, C., Jun-Biao, Z., and Yu, H. (2016) A study on research hot-spots and frontiers of agricultural science and technology innovation—visualization analysis based on the Citespace III, *Agricultural Economics*, 62(9): 429-445. <https://doi.org/10.17221/207/2015-AGRICECON>.
- Rachael, J.E., Crivelli, C., Wheatley, T. (2018) Data-Driven Methods to Diversify Knowledge of Human Psychology, *Trends in Cognitive Sciences*, 22(1): 1-5. <https://doi.org/10.1016/j.tics.2017.10.002>.
- Reisch, L., Eberle, U., and Lorek, S. (2017) Sustainable food consumption: an overview of contemporary issues and policies, *Sustainability: Science, Practice and Policy*, 9(2): 7-25. <https://doi.org/10.1080/15487733.2013.11908111>.
- Robak, K., and Balcerek, M. (2018) Review of Second Generation Bioethanol Production from Residual Biomass, *Food Technology and Biotechnology*, 56(2): 174-187. <https://doi.org/10.17113/ftb.56.02.18.5428>.
- Rockström, J., Steffen, W., Noone, K., Persson, Å., Chapin, F.S.,III, Lambin, E.F., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J. Walker, B., Liverman, D., Richardson, K., Crutzen, P., and Foley, J.A. (2009) A safe operating space for humanity, *Nature*, 461: 472-475. <https://doi.org/10.1038/461472a>.
- Ries, N., Rachul, C., and Caulfield, T. (2011) Newspaper reporting on legislative and policy interventions to address obesity: United States, Canada, and the United Kingdom, *Journal of Public Health Policy*, 32:73-90. <https://doi.org/10.1057/jphp.2010.39>.
- Reganold, J.P. and Wachter, J.M. (2016) Organic agriculture in the twenty-first century, *Nature Plants*, 2(2): 15221. <https://doi.org/10.1038/nplants.2015.221>.
- Sadowski, J. (2019) When data is capital: Datafication, accumulation, and extraction, *Big Data & Society*, 6(1). <https://doi.org/10.1177/2053951718820549>.
- Sarica, S., and Luo, J. (2021) Stopwords in technical language processing, *PLoS ONE*, 16(8): e0254937. <https://doi.org/10.1371/journal.pone.0254937>.
- Steffen, W., Richardson, K., Rockström, J., Cornell, S.E., Fetzer, I., Bennett, E.M., Biggs, R., Carpenter, S.R., de Vries, W., de Wit, C.A., Folke, C., Gerten, D., Heinke, J., Mace, G.M., Persson, L.M., Ramanathan, V., Reyers, B., Sörlin, S. (2015) Planetary boundaries: Guiding human development on a changing planet, *Science*, 347(6223): 1259855. <https://doi.org/10.1126/science.1259855>.
- Su, H. N., and Lee, P. C. (2010) Mapping knowledge structure by keyword co-occurrence: a first look at journal papers in Technology Foresight, *Scientometrics*, 85(1): 65–79. <https://doi.org/10.1007/s11192-010-0259-8>.

- Sun, J. and Yuan, B. Z. (2019) Visualization analysis of research on rice with fertilizer from the 'Agronomy' category based on CiteSpace, *Current Science*, 117(9): 1449. <http://dx.doi.org/10.18520/cs/v117/i9/1449-1458>.
- Tan, K.T., Lee, K.T., Mohamed, A.R., and Bhatia, S. (2009) Palm oil: addressing issues and towards sustainability development, *Renewable and Sustainable Energy Reviews*, 13(2), 420-427. <https://doi.org/10.1016/j.rser.2007.10.001>.
- The Guardian 「The Guardian Open Platform」 <https://open-platform.theguardian.com> (2023 年 2 月 1 日閱覽).
- The New York Times 「The New York Times Developer Network」 <https://developer.nytimes.com> (2023 年 2 月 1 日閱覽).
- United States Patent and Trademark Office (USPTO), 「Stopwords, USPTO Full-Text Database」, <http://patft.uspto.gov/netahtml/PTO/help/stopword.htm> (2022 年 7 月 30 日閱覽).
- Wang, Y.C., Klim McPherson, Tim Marsh, Steven L Gortmaker, Martin Brown. (2011) Health and economic burden of the projected obesity trends in the USA and the UK, *The Lancet*, 378(9793): 815-825. [https://doi.org/10.1016/S0140-6736\(11\)60814-3](https://doi.org/10.1016/S0140-6736(11)60814-3).
- Wibowo M.H., Fauzi A.M., and Taqi F.M. (2019) Mapping of palm oil technology based on patent information analysis, *IOP Conference Series: Earth and Environmental Science*, 335: 012016. <http://dx.doi.org/10.1088/1755-1315/335/1/012016>.
- Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L.J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J.A., De Vries, W., Majele Sibanda, L., Afshin, A., Chaudhary, A., Herrero, M., Agustina, R., Branca, F., Lartey, A., Fan, S., Crona, B., Fox, E., Bignet, V., Troell, M., Lindahl, T., Singh, S., Cornell, S.E., Srinath Reddy, K., Narain, S., Nishtar, S., Murray, C.J.L. (2019) Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems, *The Lancet*, 393(10170): 447-492. [https://doi.org/10.1016/S0140-6736\(18\)31788-4](https://doi.org/10.1016/S0140-6736(18)31788-4).
- Wagner, N. and Rondinelli, J. M. (2016). Theory-guided machine learning in materials science. *Frontiers in Materials*, 3: 2271. <https://doi.org/10.3389/fmats.2016.00028>.
- World Intellectual Property Organization (WIPO) (2016) Patent Landscape Report on Palm Oil Production and Waste Treatment, Patent Landscape Reports Technologies, <https://tind.wipo.int/record/28973>, 2022.7.30.
- World Health Organization (WHO), 「Newsroom, Obesity and overweight」, <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight> (2023 年 2 月 1 日閱覽).
- World Bank (2020) Taxes on Sugar-Sweetened Beverages: Summary of International Evidence and Experiences, <https://openknowledge.worldbank.org/handle/10986/33969>.

- Yano, Y., Blandford, D., Maruyama, A., and Nakamura, T. (2018) Consumer perceptions of fresh leafy vegetables in Japan: An application of word co-occurrence network analysis, *British Food Journal*, 120(11): 2554-2568. <https://doi.org/10.1108/BFJ-09-2017-0500>.
- アルバート・ラズロ・バラバシ (2019) 『ネットワーク科学—ひと・もの・ことの関係性をデータから解き明かす新しいアプローチ』(京都大学ネットワーク社会研究会訳) 共立出版.
- 千葉剛・佐藤陽子・中西朋子・横谷馨倫・狩野照誉・鈴木祥菜・梅垣敬三 (2014) 「特定保健用食品の利用実態調査」『日本栄養・食糧学会誌』 67(4) : 177-184 . <https://doi.org/10.4327/jsnfs.67.177>.
- 千葉剛・佐藤陽子・小林悦子・梅垣敬三 (2017) 「機能性表示食品制度の施行 1 年後における現状—消費者および医師・薬剤師を対象としたインターネット調査—」『食品衛生学雑誌』 58(2) : 96-106. <https://doi.org/10.3358/shokueishi.58.96>.
- 藤垣裕子・平川秀幸・富沢宏之・調麻佐志・林隆之・牧野淳一郎 (2004) 『研究評価・科学論のための科学計量学入門』丸善出版.
- 東山昌彦・乾健太郎・松本裕治 (2008) 「述語の選択選好性に着目した名詞評価極性の獲得」『言語処理学会第 14 回年次大会論文集』 584-587.
- 樋口耕一 (2020) 『社会調査のための計量テキスト分析 —内容分析の継承と発展を目指して— 第 2 版』 ナカニシヤ出版.
- 広垣光紀 (2011) 「特定保健用食品 (トクホ) 表示が購買行動に及ぼす影響—アンケート調査による実証分析—」『社会科学』 41(2) : 41-56 . <https://doi.org/10.14988/pa.2017.0000012483>.
- 池上幸江・山田和彦・池本真二・倉田澄子・清水俊雄・藤澤由美子・由田克士・和田政裕・坂本元子 (2008) 「栄養・健康表示の社会的ニーズの解明と食育実践への活用に関する研究」『日本栄養・食糧学会誌』 61(6) : 285-302. <https://doi.org/10.4327/jsnfs.61.285>.
- 伊藤来実・水木麻人 (2019) 「大学生を対象とした農産物における機能性表示食品に対する消費者評価」『農業情報研究』 28(1) : 38-45. <https://doi.org/10.3173/air.28.38>.
- 和泉利隆 (2014) 「ネットワーク分析：マーケティングリサーチへの応用可能性」『行動計量学』 41(2) : 103-112. <https://doi.org/10.2333/jbhmk.41.103>.
- 角田美知江 (2011) 「消費者行動から見た先発ブランド優位性についての研究」『生活経済学研究』 34 : 27-35. [https://doi.org/10.18961/seikatsukeizaigaku.34.0\\_27](https://doi.org/10.18961/seikatsukeizaigaku.34.0_27).
- 株田文博(2015) 「グローバル化が進展するアジア諸国のフードシステムの経済連関—付加価値連鎖を通じた我が国国民経済に及ぼす影響を中心に—」『フードシステム研究』 22(2) : 111-126. <https://doi.org/10.5874/jfsr.22.111>.
- 小林のぞみ・乾健太郎・松本裕治・立石健二・福島俊一 (2005) 「意見抽出のための評価表現の収集」『自然言語処理』 12(3) : 203-222. [https://doi.org/10.5715/jnlp.12.3\\_203](https://doi.org/10.5715/jnlp.12.3_203).

- 松尾豊・篠田孝祐・中島秀之（2006）「中心性に着目した合理エージェントのネットワーク形成」『人工知能学会論文誌』 21(1): 122-132. <https://doi.org/10.1527/tjsai.21.122>.
- 松元圭太郎・寺園春風（2017）「機能性表示食品の認知および機能性表示に関わる食品区分の違いの認識」『鹿児島純心女子大学看護栄養学部紀要』 21: 13-21.
- 村上太郎・石原健吾・松元圭太郎・岡村浩嗣・矢口友理・小野智子・藤井久雄・橋場直彦（2015）「広告や学術情報をとおして健康科学系の女子大学生に認識される特定保健用食品の保険用途」『日本栄養・食糧学会誌』 68(2) : 73-81. <https://doi.org/10.4327/jsnfs.68.73>.
- 森嶋輝也（2019）「バレイショに関する消費者のブランド認知構造」『フードシステム研究』 25(4) : 221-226. [https://doi.org/10.5874/jfsr.25.4\\_171](https://doi.org/10.5874/jfsr.25.4_171).
- 佐藤志乃（2013）「有向ネットワークによる消費者知識分析－参入タイミングによる知識構築戦略類型に関する試論－」『関東学院大学経済経営研究所年報』 35 : 30-45.
- 佐藤志乃（2014）「消費者知識の有向ネットワーク分析における意義と課題」『経済系：関東学院大学経済学会研究論集』 259 : 27-51.
- 清水俊雄（2005）「食品の健康表示制度と科学的根拠に関する国際比較」『日本補完代替医療学会誌』 2(2) : 81-89. <https://doi.org/10.1625/jcam.2.81>.
- 下渡敏治・名取雅彦(2010)「東アジアフードシステムのリンケージとバリューチェーン」『フードシステム研究』 17(2) : 64-75. <https://doi.org/10.5874/jfsr.17.64>.
- 消費者庁（2015）「機能性表示食品って何？」  
[https://www.caa.go.jp/policies/policy/food\\_labeling/about\\_foods\\_with\\_function\\_claims/pdf/150810\\_1.pdf](https://www.caa.go.jp/policies/policy/food_labeling/about_foods_with_function_claims/pdf/150810_1.pdf)（2023年2月1日閲覧）.
- 高田秀之（2018）「機能性表示野菜に対する消費者意識と付加価値評価」『フードシステム研究』 25(3) : 97-105. [https://doi.org/10.5874/jfsr.25.3\\_97](https://doi.org/10.5874/jfsr.25.3_97).
- 遠山和成・安達内美子（2020）「“持続可能で健康な食事” からみた世界の食事法」『名古屋栄養科学雑誌』 6 : 47-55. <https://doi.org/10.15073/00001535>.
- 氏家清和（2016）「食品表示と消費者行動をめぐる実証的研究の動向」『農業経済研究』88(2): 156-171. <https://doi.org/10.11472/nokei.88.156>.
- 山田和彦・田中弘之・石見佳子・梅垣敬三・井出留美（2017）「保健機能食品の課題と展望」『日本栄養・食糧学会誌』 70(3) : 91-99. <https://doi.org/10.4327/jsnfs.70.91>.
- 湯田直樹（2017）「届出状況から読み解く機能性表示食品」『健康・栄養食品研究』 16(1) : 1-10. [https://doi.org/10.20618/jhnfa.16.1\\_1](https://doi.org/10.20618/jhnfa.16.1_1).

## 謝辞

本論文の執筆にあたり、まず、博士論文指導教員を務めてくださった千葉大学園芸学研究所教授小林弘明先生に心より深く感謝申し上げます。小林先生には、私の千葉大学園芸学部食料資源経済学科在学時にも指導教員を務めていただき、卒業論文の執筆においても的確かつ熱心なご指導を賜った。私が現職に就いて以降も、より一層のご指導を賜り、博士論文の執筆に限らず、学会発表や論文投稿にはじまり、科研費代表課題の応募など、これまでの研究活動に関する全面に渡って多大なご助力を頂いた。マイペースな私にとっては、本論文の執筆をはじめ、これまでの研究活動の中にはタイトなスケジュールとなった場面も少なからずあり、小林先生にご迷惑をおかけすることもしばしばであったが、小林先生はどんな時でも常に私を激励し、私の意思を尊重してくださり、あたたかく接してくださった。他にも、研究面のみならず、個人的なご相談等にも乗って頂くなど、学生時代から現職に至るまで一貫してひとかたならぬご支援を賜った。

私はまだ研究者として駆け出しの身であるが、これまでの研究生活を振り返ると、小林先生という師に恵まれたことが、私にとって一番の幸運であり、また現職を続ける上での何ものにも代え難い救いであったと思っている。私が現職を続けることが出来たのは、ひとえに小林先生のご支援とご助力があったからに他ならない。現職を通して私を育ててくださったのは紛れもなく小林先生であり、返し切れないほどの本当に多大な御恩を頂いた。今後、自らの信ずる研究に弛まず打ち込み、学術的・社会的貢献に向けて尽力することが、頂いた恩と比べるとささやかながら、私に出来る恩返しだと考えている。これからより一層の研鑽を重ねる所存である。

本論文の学位審査の主査を務めてくださった丸山敦史先生、副主査を務めてくださった栗原伸一先生、加藤顕先生からは、予備審査、本審査において多くの貴重なコメントを頂いた。栗原先生、丸山先生には所属学科の先生方として学部生の頃からお世話になっており、今回審査をお引き受け頂いたことに心より感謝申し上げたい。また加藤先生には他分野の観点から大変貴重なコメントを頂いた。お忙しいところ拙稿を審査頂いた先生方にあらためて深く御礼申し上げます。

他にも、本論文の執筆にあたっては、著者の現所属である農研機構の皆様をはじめ、千葉大学石田貴士先生、九州大学上西良廣先生など、多くの方々に多大なご支援を頂いた。そして、本論文は、先人達が開発・提供している多くの分析手法や分析ツール、並びにデータセットやデータベースが存在しなければ成立しなかった。そうした開発陣や提供者の皆様にもこの場を借りて御礼申し上げます。

また、東京工業大学大学院修士課程在学時に指導教員を務めてくださった同大学名誉教授出口弘先生にもこの場を借りて心より深く御礼申し上げたい。出口先生には、私が情報

系分野を学び研究する機会を与えて頂いただけでなく、出口先生の学問のフロンティアに挑む精力的な姿勢からは、非常に多くのことを学ばせて頂いた。また、出口研究室の自由闊達な風土は、私にとって、研究に取り組む上での基礎的な素養や姿勢を涵養する大変恵まれた環境であった。

修士課程の経験は今なお大きな財産であり、私の研究に対する取り組み方や考え方の基盤を形作ったものであると考えている。修士課程という短い期間ではあったが、出口先生にご指導頂いたこともまた、小林先生からご指導頂いたことと同様に、私にとって本当に幸運であったと思っている。当時の出口研究室の皆様にあらためて感謝申し上げる。

最後に、学生時代から現在に至るまで一貫して応援し続けてくれた家族をはじめとして、これまでお世話になった皆様に心より深く感謝申し上げ、以上を謝辞とする。

2023年2月

加藤弘祐