

PIO-NETを利用した消費者問題の傾向分析

リサーチ・ディスカッション・ペーパー
RESEARCH DISCUSSION PAPER

谷岡 広樹、徳永 欽也、徳永 美和子、近森 美麻子、三谷 達也、
宮本 麗子、山口 愛弓、小早川 優、安井 杏奈

2022年4月



消費者庁新未来創造戦略本部国際消費者政策研究センター
International Consumer Policy Research Center (ICPRC)
Consumer Affairs Agency
Tokushima, Japan

本リサーチ・ディスカッション・ペーパーは、新未来創造戦略本部国際消費者政策研究センターの研究者等により行われた研究の成果を取りまとめたものです。学界、研究機関等の関係する方々から幅広くコメントを頂き、消費者行政における政策立案の基礎資料として役立てることを意図して発表しております。なお、全て研究者個人の責任で執筆されており、消費者庁の見解を示すものではありません。

PIO-NET を利用した消費者問題の傾向分析

谷岡 広樹^{1,2)}、

徳永 欽也³⁾、徳永 美和子³⁾、近森 美麻子⁴⁾、三谷 達也⁴⁾、宮本 麗子³⁾、
山口 愛弓³⁾、小早川 優³⁾、安井 杏奈³⁾

¹⁾ 徳島大学大学院創成科学研究科理工学専攻講師

²⁾ 消費者庁新未来創造戦略本部国際消費者政策研究センター客員主任研究官

³⁾ 消費者庁新未来創造戦略本部国際消費者政策研究センター特任研究員

⁴⁾ 消費者庁新未来創造戦略本部国際消費者政策研究センター特任研究員(～2022年3月末)

概要

本稿は、2020年の1年間に全国消費生活情報ネットワークシステム（PIO-NET）へ登録された相談内容のうち、新型コロナウイルス関連の話題を抽出したものについて、傾向分析を行ったものである。PIO-NETの相談内容に含まれる話題の分析には、トピックモデリング¹と呼ばれる統計手法を用いて、話題の分類と特徴的なキーワードの抽出を行う。また、各月ごとの新型コロナウイルス関連の主な状況について、消費生活相談の傾向と比較する。その結果、個別の話題の増減のタイミングで、消費生活相談も増減しており、大きな傾向の違いはみられなかったが、4月下旬に関連する報道や施策が発表された特別定額給付金等について、消費生活相談としてはその1か月遅れでトピックとして現れるなど、政策の発表から相談として現れるまで、若干のタイムラグがあることが示唆される現象が見られた。

また、トピック分析から得られたトラブルキーワードについて、その出現頻度や出現割合を分析した結果、そのトラブルキーワードが代表する相談が顕著に見て取れた。

比較のために、同期間における日本経済新聞の朝刊²で取り上げられた記事についても分析を行った。新聞記事データは、新型コロナウイルス関連のみを抽出したものではないため、新聞記事データを用いたトピック分析では、PIO-NETデータのトピック分析結果とは全く異なる傾向であった。しかし、PIO-NETデータの分析から得られたトラブルキーワードを用いて、新聞記事データにおける出現頻度や出現割合を分析することで、消費生活相談で話題となっているトピックの新聞報道における注目状況を概観することができた。

¹ トピックモデリングとは、文章が含む話題（トピック）の確率的なモデルを分析する手法である。

² 1面、総合・政治、総合・経済、総合、政治、経済、社会のキーワードで抽出した面を対象とした。

1 はじめに

2020年から2022年にかけて社会問題となっている新型コロナウイルス関連の話題は、日本各地の消費生活センター及び国民生活センターが受け付けた相談情報を蓄積したデータベースである全国消費生活情報ネットワークシステム³（PIO-NET; Practical Living Information Online Network System）[1]に登録された相談情報にも様々な形で影響があると考えられる。そこで実際に、2020年1月から12月の1年間に受け付けられた相談のうち、新型コロナウイルス関連の相談について分析することとした。本研究は、PIO-NETへ登録された新型コロナウイルス関連の相談情報についてテキストマイニングによる分析を加えることで、相談内容や相談件数の時系列的な変遷を明らかにしようとするものである。

なお、比較のために、2020年1月から12月の期間に発行された日本経済新聞の朝刊で取り上げられた記事についても分析し、その傾向の違いをみる。

2 対象データ

本研究では、日本国内で新型コロナウイルスへの感染が最初に確認された2020年1月から2020年12月までの1年間に受け付けられた相談、及び日本経済新聞社から発行された日本経済新聞朝刊の記事に関する文字情報を対象データとする。それぞれの詳細を以下に示す。

2.1 PIO-NET データ

PIO-NETデータは、2020年12月31日時点で登録されている相談情報のうち、相談受付月が2020年1月から12月までの「新型コロナウイルス関連」の相談情報90,064件を対象とする。本件データは、国民生活センターにおいて、相談対応結果及びその内容と、個人情報に関する項目を除いた上で、消費者庁において抽出したものをを用いる。

本研究では、提供を受けた相談情報のうち、相談内容を文章で要約する欄に記載されているテキストデータ（878,677文）のみを利用する。

2.2 新聞記事データ

新聞記事データは、2020年1月1日から12月31日までに日本経済新聞社から発行された日本経済新聞朝刊の記事データのうち、1面、総合・政治、総合・経済、総合、政治、経済、社会のキーワードで抽出した面の記事18,527件を対象とする。

本研究では、見出しと本文のテキストデータ（351,217文）のみを利用する。新聞記事データについては、PIO-NETデータのように新型コロナウイルス関連であるか否かの判断をする情報は付与されていないため、すべての記事を新聞記事データとして用いる。

³ 全国消費生活情報ネットワークシステム（PIO-NET）とは、国民生活センターと全国の消費生活センターをネットワークで結び、消費者から消費生活センターに寄せられる消費生活に関する相談情報（消費生活相談情報）の収集を行っているシステムである。

3 キーワードの分析

3.1 キーワードの抽出

PIO-NETデータと新聞記事データを分析するために、まずそれぞれのテキストデータを形態素解析⁴する。形態素解析には spaCy v2.3.5 (GiNZA, Sucachi) [2][3][4] を用いる。分析には、品詞が名詞と固有名詞の単語のみを採用する。また、分析に不適切と考えられる単語であるストップワード⁵を除いた結果をキーワードとする。

ストップワードには SlothLib⁶ [5]の日本語ストップワードリスト 328 語に加え、出現頻度が1回の単位や指示語、代名詞等の単語及び新型コロナウイルス関連の話題に共通の単語である「新型コロナ」「コロナウイルス」といった単語を追加して用いた(付録A参照)。

表1は、ストップワードを除き、分析対象としたキーワードの統計情報である。

表1: 日本経済新聞と PIO-NET (新型コロナウイルス関連) のキーワードの統計

データ種別	単語数	異なり語数 ⁷
PIO-NET (新型コロナウイルス関連)	16,041,662	28,234
日本経済新聞	7,306,631	47,537

3.2 PIO-NET データの高頻度キーワード分析

PIO-NET データの高頻度キーワード上位 20 語について、本研究で対象とする期間内の変動を月別に見る。図1は2020年1月から12月までのPIO-NETデータに含まれる上位20語の出現頻度を積み上げたグラフである。キーワードが2020年1月から4月にかけて増加し、その後、緩やかに減少しており、これは新型コロナウイルス関連の相談件数の推移とも類似している[6]。

また、図2は2020年1月から12月までのPIO-NETデータに含まれる上位20語の出現割合を積み上げたグラフである。このグラフでは、特にキーワード「マスク」が5月以降徐々に減少し、7月以降にはほぼ横ばいとなっている。厚生労働省が一般のマスクの品薄状態が解消したのが7月末の段階であるとしていることを踏まえると、5月以降、マスクが消費者の手に渡り始めるにつれてマスク関連の相談が徐々に減少し、7月以降、品薄状態の解消に伴ってマスク関連の相談が落ち着いたためであると考えられる。

なお、この高頻度キーワードは、単に出現頻度の高い単語であり、その出現頻度や出現割合からは、特定の相談内容や相談件数が示されるものではない。

⁴ 形態素解析とは、文章を単語に分ち書きし、単語の品詞を推定する自然言語処理技術である。

⁵ ストップワードとは、分析に不要として除外する単語のことである。

⁶ SlothLib 日本語ストップワード: <http://svn.sourceforge.jp/svnroot/slothlib/CSharp/Version1/SlothLib/NLP/Filter/StopWord/word/Japanese.txt>

⁷ 異なり語数とは、解析結果から得られた単語の総数から重複する単語を省いた単語の数である。

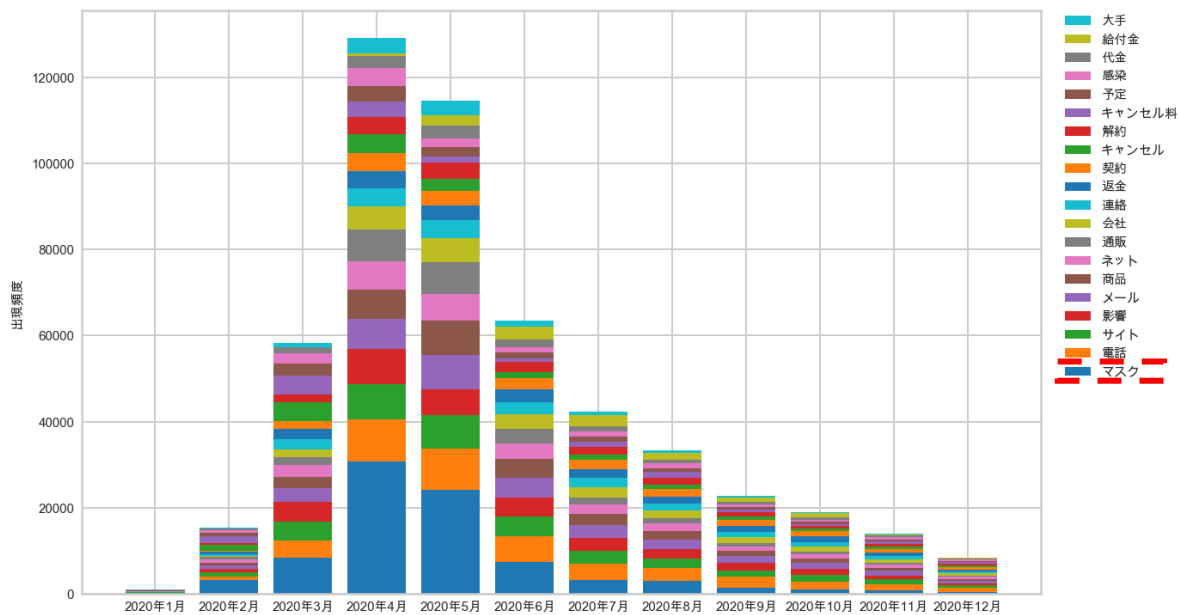


図 1：PIO-NET データ内の高頻度キーワード上位 20 語の月別の出現頻度

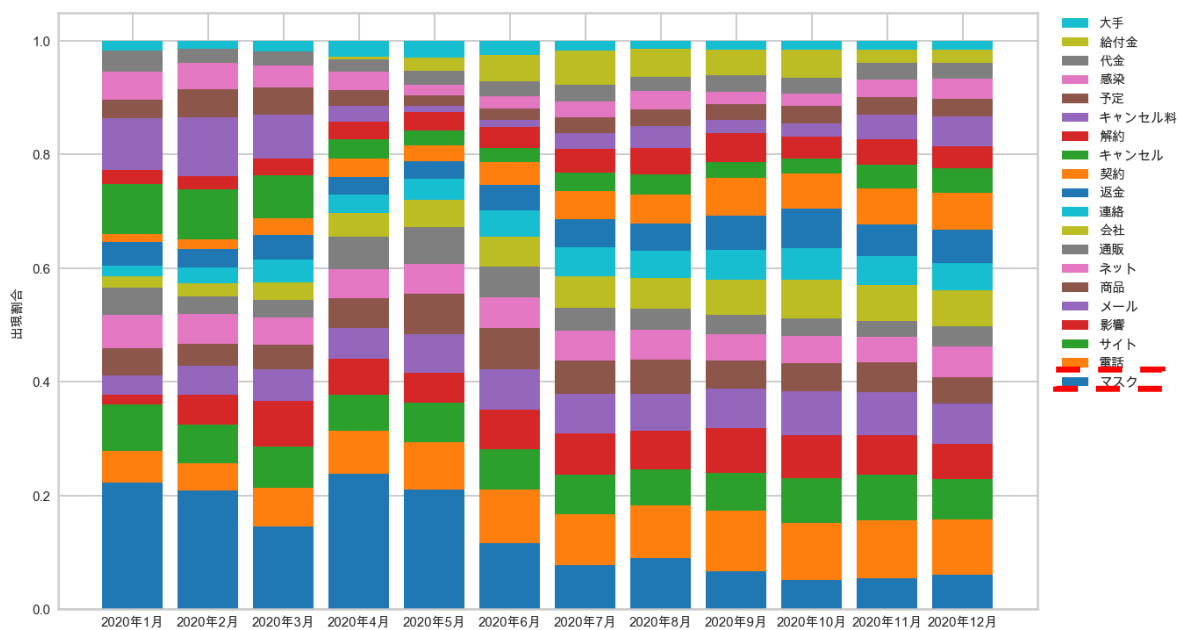


図 2：PIO-NET データ内の高頻度キーワード上位 20 語と月別の出現割合

4 トピックモデリング

前節までは、PIO-NETデータに含まれる高頻度キーワードの傾向について見ることで年間を通じた相談の傾向を外観したが、具体的にどのような相談内容であるかについては明らかになっていない。そのため本節では、PIO-NETデータに含まれる話題（トピック）について、トピックモデルの一種である潜在的ディリクレ配分法 (LDA: Latent Dirichlet Allocation) [7][8] を用いて抽出する。トピックモデルとは、文書が複数の潜在的なトピックから確率的に生成されていると仮定したモデルである。また、トピックモデルを用いた手法の一つである LDA は、文書中に含まれる単語が、あるトピックから生じると仮定し、文書は複数のトピックから生成されていると考える。図3はその概念を図示したものである。

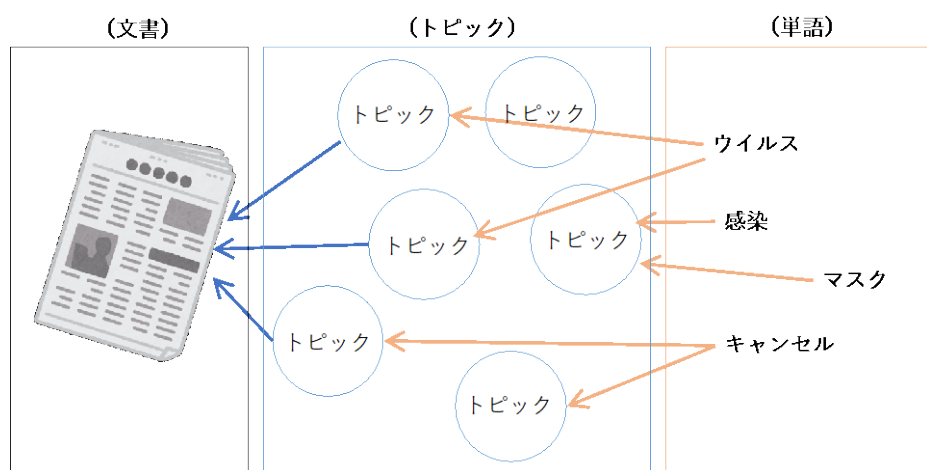


図3：LDA の概念図

図4は、LDA をグラフィカルモデル⁸で表したものである。ここで N は単語数、 M は文書数を意味する。 φ は各トピックにどの単語が出現するかを表す多項分布であり、 K はトピックの数で、四角く囲まれた部分は繰り返し実行する。 W は文書に含まれる単語であり、 Z は単語 W に割り当てられたトピック、 θ は各文書がどのトピックに属するかを表す多項分布である。また、 α と β は連続型の確率分布の一種であるディリクレ分布のパラメータである。ディリクレ分布は多変量ベータ分布とも呼ばれ、同時に発生することのない複数の事象が起こる確率を表したものである。

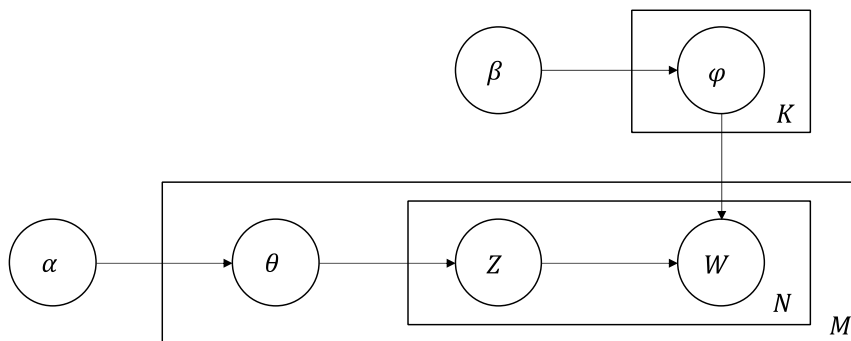


図4：LDA のグラフィカルモデル

⁸ グラフィカルモデルとは、対象となる事象を確率的なモデルとして視覚的に表現したものである。

M : 文書数
 K : トピック数
 N : 文書に含まれる単語の数 (i 番目の文書は N_i)
 θ_i : 文書 i のトピックの分布 (多項分布)
 φ_k : トピック k の単語の分布 (多項分布)
 Z_{ij} : 文書 i の j 番目の単語 W_{ij} のトピック
 W_{ij} : 文書 i の j 番目の単語
 α : 文書ごとのトピック分布に対するディリクレ事前分布のパラメータ
 β : トピックごとの単語分布のディリクレ事前分布のパラメータ

このモデルでは、単語 W を含む文書に含まれるトピック Z の生じる確率をパラメータ α で表し、トピック Z ごとに単語 W の生じる確率をパラメータ β で表す。パラメータ α により単語 W パラメータ β により文書がどのトピックに属しているかを確率的に示すことができる。LDA では、単語 W をもとにしてパラメータ α 及び β を推定する。パラメータの推定にはベイズ推定⁹を用い、データのサンプリングは、ギブスサンプリング¹⁰で確率分布を近似する。本実験では、Python 3.9.2 環境で scikit-learn 0.24.2 [9][10]を用いて実装した。主要なパラメータを表2に示す。また記載のないものはすべてデフォルト値とした。なお、evaluate_everyの備考にあるパープレキシティ(Perplexity)は、トピックの平均分岐数であり、トピックの複雑さを表している。

表2: sklearn.decomposition.LatentDirichletAllocation で用いたパラメータ

パラメータ	値	備考
n_components	8	トピック数、本実験では8つのトピックに分類した
max_iter	30	訓練データを読み込む最大回数 (エポック数)
n_jobs	-1	推定ステップで使用するプロセッサ数 (-1は全て)
evaluate_every	1	Perplexityの評価頻度 (反復ごとに perplexity を1回)

LDA の計算には、文書に含まれるトピックを推定するために単語とその頻度または重要度が必要となる。そのため、PIO-NETの相談内容から形態素解析により抽出されたキーワードの文書ごとの重要度に TF-IDF[11]を用いる。TF-IDF は、ある文書に含まれる単語の頻度を表す TF(Term Frequency)と、その単語を含む文書数の逆数で表す IDF(Inverse Document Frequency)を乗算したものである。

$$\text{tfidf}_{i,j} = \text{tf}_{i,j} \cdot \text{idf}_i \quad (1)$$

$$\text{tf}_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (2)$$

$$\text{idf}_i = \log \frac{|D|}{|\{d: d \in t_i\}|} \quad (3)$$

⁹ ベイズ推定とは、ベイズ定理 (Bayes' theorem) の考え方に基づいたある現象の頻度や傾向の代わりに、事前知識を用いた期待値による定量化がなされた確率、いわゆるベイズ確率を用いて、観測事象から、推定したい事柄を、確率的な意味で推論することである。

¹⁰ ギブスサンプリングとは、直接サンプリングが難しい確率分布の代わりに、それを近似するサンプル列を生成してサンプリングする手法である。

式(1)は、ある単語 t_i の TF-IDF の値の計算式、式(2)は文書 d_j における単語 t_i の TF の計算式、式(3)は、ある単語 t_i の IDF の計算式である。このとき $n_{\{i,j\}}$ は文書 d_j における単語 t_i の出現回数、 $\sum_k n_{\{k,j\}}$ は文書 d_j におけるすべての単語の出現回数の和、 $|D|$ は総文書数、 $\{d: d \in t_i\}$ は単語 t_i を含む文書数である。

4.1 PIO-NET データのトピックモデリング

PIO-NET データの 2020 年 1 月から 2020 年 12 月までの 1 年間の相談情報について、相談 1 件を文書 1 件として TF-IDF を計算し、LDA で分析した結果である 8 つのトピックを図 5 に示す。このとき、トピックに含まれるキーワードの大きさは、キーワードによってそのトピックが生成される確率（生成確率）に基づく相対的な文字サイズである。個々のトピックにおいて生成確率が最大となるキーワードが最大の文字サイズとなるよう表示されるが、トピック間での文字サイズは相対的ではない。また、配色は機械的に割り当てられたものであり、生成確率の大きさ等の意味合いはない。

それぞれのトピックに含まれるキーワードのうち、生成確率が最大のキーワードをトピックの代表キーワードとしたとき、それぞれ、「給付金」「マスク」「航空券」「契約」「会社」「教室」「結婚式」「サイト」となった。



図 5：2020 年 PIO-NET 新型コロナウイルス関連のトピック（8 トピック）

各トピックのキーワードから、「給付金」は特別定額給付金の支給に便乗する詐欺に関する消費者トラブル、「航空券」や「契約」「教室」「結婚式」は新型コロナウイルス拡大の影響でサービス等を受けられなくなった際のキャンセル料や返金に関する消費者トラブル、「マスク」は、マスクの品不足に関連する消費者トラブル、「サイト」は、通販等で購入した商品に関する消費者トラブルについてのトピックではないかと推測される。

4.2 PIO-NET データのトピック時系列分析

PIO-NETデータのトピックモデルをより詳細に分析するため、月別にトピックモデリングを行う。また、分類されたトピックごとに振り分けられた相談情報を目視にて確認し、それぞれのトピックに含まれる総体的な相談内容についても検討を加える（付録C参照）。

さらに、2020年1月から12月までの新型コロナウイルス関連の主な状況についても比較することにより、相談の傾向を分析する。

なお、各トピックにおいて色の付いている単語は、トラブルキーワード（後述）である。

4.2.1 2020年1月

2020年1月のトピックの代表的なキーワードとしては、「旅行」「チケット」「マスク」「キャンセル料」「外国人」「航空券」「マスク」「クレジットカード」があった。海外で新型コロナウイルスが確認されたことや、外務省が渡航中止勧告等を出したことに伴うと考えられる、旅行のキャンセルに関するトラブルが多かったことがわかる。

なお、「クレジットカード」については、文字サイズの都合上、図6では「クレジット...」と表記した。



図6：2020年1月のPIO-NETのトピック（8トピック）

表3：2020年1月の新型コロナウイルス関連の主な状況

2020年1月
6日：中国武漢で原因不明の肺炎 厚労省が注意喚起
14日：WHOが新型コロナウイルスを確認
15日：日本国内で初めて感染確認
25日：外務省が渡航中止勧告
30日：WHOが「国際的な緊急事態」を宣言
31日：新型コロナウイルスを「指定感染症」に閣議決定

4.2.2 2020年2月

2020年2月のトピックの代表的なキーワードとしては、「消毒液」「航空券」「マスク」「マスク」「チケット」「旅行」「キャンセル料」「メール」があった。1月に引き続き、旅行のキャンセルのほか、2月に入ると、1月中旬に国内での新型コロナウイルスへの感染が確認されたことにより、結婚式等のイベントのキャンセルに関する相談もトピックとして出てきている。また、マスクが品薄になったため、マスクに関する相談も出てきている。さらに、2月末にトイレトペーパーがこれから不足するというデマがSNS上で発信され、一時的にトイレトペーパーが品薄になる状況が見られた。このトイレトペーパーの品薄に伴うと考えられる相談もトピックとして出てきている。このデマに起因するトイレトペーパーの品薄は、4月上旬まで続いた。

マスクが大きく出現している3つのトピックについてみると、①マスクやトイレトペーパーの品不足や買占に関する相談、②フリマアプリ等でのマスクの高額販売に関する相談、③注文していない商品が届く等マスクの購入に関する相談に分かれていた。

なお、「トイレトペーパー」については、文字サイズの都合上、図7では「トイレト...」と表記した。



図7：2020年2月のPIO-NETのトピック（8トピック）

表4：2020年2月の新型コロナウイルス関連の主な状況

2020年2月
3日：乗客の感染が確認されたクルーズ船が横浜港に入港
11日：WHOが新型コロナウイルスを「COVID-19」と名付ける
13日：国内で初めて感染者死亡
27日：全国すべての小中学校に臨時休校要請
28日：北海道が独自に緊急事態宣言

4.2.3 2020年3月

2020年3月のトピックの代表的なキーワードとしては、「トイレトペーパー」「教室」「対策」「マスク」「マスク」「契約」「キャンセル料」「航空券」があった。2月に引き続き、イベントのキャンセルとマスク・トイレトペーパーに関するトラブルが多かったことがわかる。また、結婚式等のイベントのキャンセルのほか、定期的に通う習い事や教室の中止に伴う返金等に関する相談もトピックとして出てきている。

なお、「トイレトペーパー」については、文字サイズの都合上、図8では「トイレト...」と表記した。



図8：2020年3月のPIO-NETのトピック（8トピック）

表5：2020年3月の新型コロナウイルス関連の主な状況

2020年3月
10日：政府が「歴史的緊急事態」に初指定
11日：WHOが「パンデミックと言える」と発表
24日：東京五輪・パラリンピック1年程度延期が発表

4.2.4 2020年4月

2020年4月のトピックの代表的なキーワードとしては、「結婚式」「マスク」「マスク」「教室」「返金」「対策」「マスク」「契約」があった。3月に引き続き、イベントのキャンセル、教室の休会のトラブルに加えて、マスクに関するトラブルのトピック数が増加している。

なお、4月上旬まで続いていたトイレトペーパーの品薄状態が解消されたことに伴い、トイレトペーパーに関する相談のトピックはなくなっている。

「契約」に代表されるトピックは、キーワードからはその内容を読み取ることが難しいが、相談情報を目視にて確認したところ、政府による学校の臨時休校要請や緊急事態宣言の発出、外出自粛要請などを受けて、収入が減少する、大学進学に伴う引越しを延期する等、生活や予定の変更に伴う解約の相談が主な内容であった（付録C参照）。



図9：2020年4月のPIO-NETのトピック（8トピック）

表6：2020年4月の新型コロナウイルス関連の主な状況

2020年4月
1日：政府が「全世界からの入国者に2週間の待機要請」決定
1日：政府が全国すべての世帯に布マスク2枚ずつ配布の方針表明
7日：7都道府県に「緊急事態宣言」
7日：「持続化給付金」閣議決定
16日：「緊急事態宣言」全国に拡大 13都道府県は「特定警戒都道府県」に
16日：首相 すべての国民対象に一律1人当たり10万円を給付する考えを表明
18日：国内の感染者1万人超える
20日：「特別定額給付金」閣議決定
27日：「特別定額給付金」申請書発送 ※準備のできた自治体から順次

4.2.5 2020年5月

2020年5月のトピックの代表的なキーワードとしては、「電話」「給付金」「会社」「マスク」「返金」「マスク」「解約」「広告」があった。4月には複数見られていたマスクやキャンセルに関するトピック数が減少し、新たに特別定額給付金やクレジットカードの請求に関するトピックが見られるようになった。

特別定額給付金については、4月下旬に関議決定され、5月から支給が始まったが、特別定額給付金に関する相談は、4月に比べて5月に多く見られている。



図 10：2020年5月のPIO-NETのトピック（8トピック）

表 7：2020年5月の新型コロナウイルス関連の主な状況

2020年5月
1日：「特別定額給付金」支給開始
1日：中小企業などへの「持続化給付金」受付開始
4日：「緊急事態宣言」5月31日まで延長
8日：「持続化給付金」支給開始
14日：「緊急事態宣言」39県で解除 8都道府県は継続
21日：「緊急事態宣言」関西3府県は解除 ※首都圏と北海道は継続
25日：「緊急事態宣言」全国で解除

4.2.6 2020年6月

2020年6月のトピックの代表的なキーワードとしては、「会社」「契約」「給付金」「商品」「返金」「マスク」「仕事」「電話」があった。5月に引き続き、イベントや旅行のキャンセルの相談のほかに、特別定額給付金やクレジットカードの請求に関するトラブルがあったことがわかる。

なお、マスクに関する相談のトピック数は、4月には3つ、5月には2つであったが、6月には1つと、徐々に減少している。5月以降、徐々に品薄状態が解消されてきたことに伴い、消費生活相談も減少してきているということが見て取れる。



図 11：2020年6月のPIO-NETのトピック（8トピック）

表 8：2020年6月の新型コロナウイルス関連の主な状況

2020年6月
16日：「GoTo トラベル」実施報道発表
19日：都道府県またぐ移動の自粛要請 全国で緩和
19日：コロナ感染者との濃厚接触を通知するアプリ「COCOA」利用開始
24日：電気・ガス料金 支払期限が最大4か月延長へ
28日：世界の感染者 1000万人超える

4.2.7 2020年7月

2020年7月のトピックの代表的なキーワードとしては、「給付金」「契約」「定額」「商品」「会社」「返金」「マスク」「キャンセル料」があった。特別定額給付金に関するトラブルのトピック数が増加し、旅行やイベントのキャンセル料に関するトラブルも再び増加している。

なお、GoTo トラベルについては6月中旬に実施報道、7月下旬に運用開始、GoTo イートについては7月中旬に実施報道がなされているが、これに伴う相談のトピックは見られない。



図 12：2020年7月のPIO-NETのトピック（8トピック）

表 9：2020年7月の新型コロナウイルス関連の主な状況

2020年7月
21日：「GoTo イート」実施報道発表
22日：「GoTo トラベル」宿泊割引運用開始
23日：世界の感染者 1500万人超える

4.2.8 2020年8月

2020年8月のトピックの代表的なキーワードとしては、「会社」「契約」「マスク」「商品」「キャンセル料」「マスク」「感染」「給付金」があった。マスクやイベント等のキャンセル料に関するトラブルが根強く残るほか、教室の解約や特別定額給付金の申請に関するトラブルなど、多岐にわたる相談が見られる。

8月にかけて国内の新規感染者数の増加が背景となっていると思われる、店舗での感染対策に関する苦情相談のトピックが出てきている。



図 13：2020年8月のPIO-NETのトピック（8トピック）

表 10：2020年8月の新型コロナウイルス関連の主な状況

2020年8月
11日：世界の感染者 2000万人超える
14日：GoTo トラベル 割引分の還付手続き 受付開始
28日：政府が新型コロナウイルス対策の新たな方針発表
29日：マスクとアルコール消毒液の転売規制解除

4.2.9 2020年9月

2020年9月のトピックの代表的なキーワードとしては、「電話」「家賃」「契約」「契約」「サイト」「給付金」「旅行」「マスク」があった。「契約」に関する相談のトピック数が2つに増えている理由として、特別定額給付金やクレジットカードに関するトラブルに加えて、9月には新型コロナウイルス感染症の影響による解雇・雇止めが全国で5万人を超えるとの見込みが発表されたことなどを背景に、店舗や家賃の契約に関するトラブルが増加したと考えられる。



図 14：2020年9月のPIO-NETのトピック（8トピック）

表 11：2020年9月の新型コロナウイルス関連の主な状況

2020年9月
9日：世界の製薬会社などが新型コロナウイルスのワクチン開発で“安全最優先”を宣言
16日：第99代首相に菅義偉氏
18日：世界の感染者3000万人を超える
29日：世界の死者100万人を超える

4.2.10 2020年10月

2020年10月のトピックの代表的なキーワードとしては、「マスク」「契約」「カード」「キャンペーン」「電話」「航空券」「チケット」「給付金」があった。クレジットカードに関するトラブルは収まってきたが、マスクやイベント等のキャンセルに関するトラブルに加えて、GoTo トラベルの地域共通クーポンが正しく発行されなかった、使えなかったなど、クーポンの発行に伴う運用の変化に関するトラブルが出始めたことがわかる。



図 15：2020年10月のPIO-NETのトピック（8トピック）

表 12：2020年10月の新型コロナウイルス関連の主な状況

2020年10月
1日：「GoTo トラベル」地域共通クーポン配布開始
1日：「GoTo イート」ポイント指定サイトで付与開始
5日：「GoTo イート」プレミアム付食事券販売開始 ※準備のできた自治体から順次
12日 ヨーロッパで感染急拡大と報道
19日：「GoTo 商店街」最大300万円補助 全国34事業で開始
19日：世界の感染者 4000万人超える

4.2.11 2020年11月

2020年11月のトピックの代表的なキーワードとしては、「マスク」「電話」「仕事」「ポイント」「旅行」「契約」「給付金」「結婚式」があった。マスクやイベント等のキャンセルに関するトラブルのほか、10月に見られたGoToトラベルに関するトラブルに加えて、GoTo イートに関するトラブルも見られる。これらの変化は、GoToトラベル及びGoToイートの見直しや一時停止が発表されたことに伴うものと考えられる。



図 16：2020年11月のPIO-NETのトピック（8トピック）

表 13：2020年11月の新型コロナウイルス関連の主な状況

2020年11月
9日：世界の感染者 5000万人超える
9日：新型コロナウイルス影響の失業者 7万人超える
24日：「GoToトラベル」宿泊割引一時停止発表 ※札幌市、大阪市
26日：世界の感染者 6000万人超える
27日：「GoToイート」10都道府県でプレミアム付き食事券販売一時停止

4.2.12 2020年12月

2020年12月のトピックの代表的なキーワードとしては、「マスク」「返金」「メール」「旅行」「電話」「結婚式」「契約」「給付金」があった。マスクやイベントのキャンセル料に関するトラブルは根強く、このほか給付金やGoToトラベルに関するトラブルも見られる。これに加えて、収入が減少することによる家賃支払いに関する相談も見られた。



図 17：2020年12月のPIO-NETのトピック（8トピック）

表 14：2020年12月の新型コロナウイルス関連の主な状況

2020年12月
11日：厚労省 ワクチン 6000万人分供給で英製薬大手と正式契約
12日：世界の感染者 7000万人超える
14日：「GoToトラベル」宿泊割引一時停止発表 ※札幌市、大阪市、東京都、名古屋市、広島市
22日：「GoToトラベル」クーポン配布・利用停止発表
26日：コロナ変異ウイルス 国内で初確認
27日：世界の感染者 8000万人超える
28日：「GoToトラベル」全国一斉に停止
28日：全世界からの外国人の新規入国停止

5 PIO-NET データのトラブルキーワード分析

PIO-NET データについて、トピックとして現れた類型の相談が、1年間を通してどのような傾向であるか、各トピックに含まれるキーワードを用いて分析する。

具体的には、2020年1月から12月までの月別のトピック分析を行い、生成された8つのトピックについて、それぞれを構成するキーワードのうち生成確率が最大となる代表キーワード96語（8語×12ヶ月）から、重複を除いた28語を、新型コロナウイルス関連の消費者のトラブルを表すトラブルキーワードと定義する（表15）。その上で、PIO-NET データに含まれるトラブルキーワード28語について、件数を月別で見ることで消費者トラブルの動向を見る。

第3節では、高頻度キーワードについて、その出現頻度及び出現割合について分析を行った（図1、図2）が、高頻度キーワードが単に出現頻度の高い単語であるのに対し、本節で分析するトラブルキーワードは、月別の各トピックにおける代表キーワードである点に大きな違いがある。このトラブルキーワードの出現頻度及び出現割合を分析することで、そのキーワードに代表される話題についての動向を、ある程度把握することができると考えられる。

表15：PIO-NET データの月別トピックから抽出したトラブルキーワード

電話	契約
返金	外国人
解約	商品
航空券	会社
給付金	仕事
結婚式	メール
消毒液	マスク
旅行	ポイント
教室	トイレトペーパー
感染	チケット
広告	サイト
対策	クレジットカード
家賃	キャンペーン
定額	キャンセル料

5.2 PIO-NET の時系列分析

図 18 は PIO-NET データの 2020 年 1 月から 12 月までの相談内容に含まれるトラブルキーワードの出現頻度を積み上げたグラフである。高頻度キーワードを分析した場合と同様に、トラブルキーワード「マスク」について、4 月及び 5 月の出現頻度が際立っている。

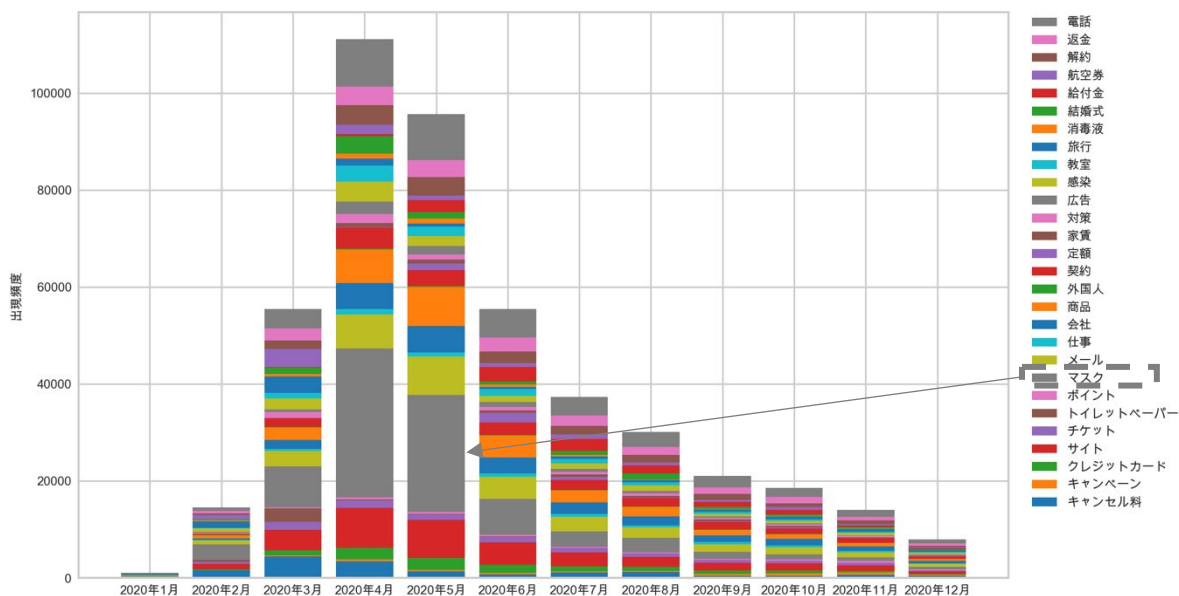


図 18 : PIO-NET データ内のトラブルキーワードと月別の出現頻度

図 19 は 2020 年 1 月から 12 月までの相談内容に含まれるトラブルキーワードの出現割合を積み上げたグラフである。1 月から 3 月までは「キャンセル料」の割合が高いが、徐々に減少していることがわかる。「トイレトペーパー」についても 2 月と 3 月に登場して収束している。また、「給付金」や「キャンペーン」の割合が 6 月ごろから徐々に増加している。

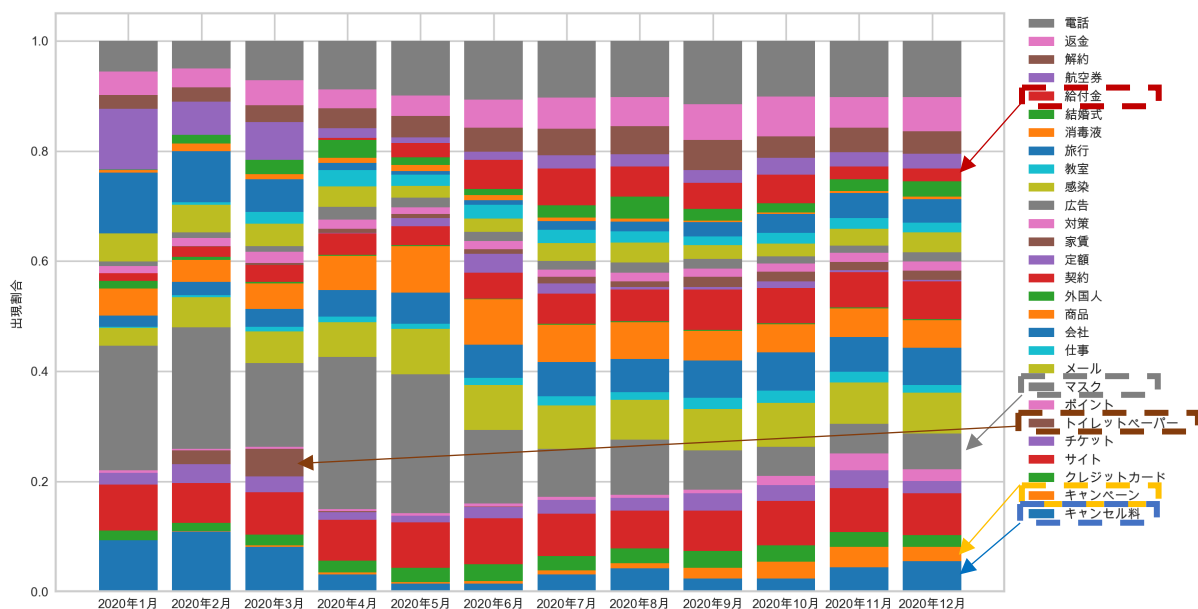


図 19 : PIO-NET データ内のトラブルキーワードと月別の出現割合

6 おわりに

本稿では、2020年の1年間にPIO-NETに登録された「新型コロナウイルス関連」の相談情報について、テキストマイニングの手法を用いて分析し、トピックという形で示すことにより、相談内容について俯瞰した。また、各月ごとの新型コロナウイルス関連の主な状況について、消費生活相談の傾向と比較した。個別の話題の増減のタイミングで、消費生活相談も増減しており、大きな傾向の違いは見られなかったが、4月下旬に関連する報道や施策が発表された特別定額給付金等について、消費生活相談としてはその1か月遅れでトピックとして現れるなど、政策の発表から相談として現れるまで、若干のタイムラグがあることが示唆される現象が見られた。具体的にどの程度のラグが見られるかについては今回の分析では示していないが、今後、この点を分析することにより、適時適切な注意喚起等の消費者行政施策へ応用できる可能性があると考えられる。

本研究では、トピック分析におけるトピック数を8つに設定した結果を示している。トピック分析による分類後、それぞれの相談情報を目視で確認したところ、1つのトピックに2種類の相談が混ざっている、又は1種類の相談が2~3のトピックに分かれているといった現象が見られ、月ごとに適切と考えられるトピック数が異なっていた。適切なトピック数にて自動的にトピック分析が可能となるような工夫についても、今後の課題である。

また、トピック分析から得られたトラブルキーワードについて、その出現頻度や出現割合を分析した結果、そのトラブルキーワードが代表する相談が顕著に見て取れた。これは各トピックの量的な傾向を把握するために有効な方法ではないかと考えられる。

なお、新聞記事データについてもPIO-NETデータと同様の分析を行った(付録B)。新聞記事データは、新型コロナウイルス関連のみを抽出したものではないため、新聞記事データを用いたトピック分析では、PIO-NETデータのトピック分析結果とは全く異なる傾向であった。しかし、PIO-NETデータの分析から得られたトラブルキーワードを用いて、新聞記事データにおける出現頻度や出現割合を分析することで、消費生活相談で話題となっているトピックの新聞報道における注目状況を概観することができた。

付録 B 新聞記事データとの比較・分析

1. 高頻度キーワードの比較

表 B-1 は、PIO-NET データと新聞記事データに含まれるキーワードのうち、頻度の大きい順に、それぞれ上位 20 単語を選択したものである。PIO-NET データと新聞記事データを比較すると、重複するのは「感染」のみであり、データに含まれるキーワードの傾向が異なることがわかる。これは PIO-NET データが、新型コロナウイルス関連に限定した内容であるのに対し、新聞記事データは、すべての記事であることが一因と考えられる。

表 B-1：PIO-NET（新型コロナウイルス関連）と日本経済新聞の高頻度キーワード上位 20 語

PIO-NET キーワード	頻度	新聞記事キーワード	頻度
マスク	84,310	政府	17,927
電話	43,500	感染	16,269
サイト	35,825	日本	14,579
影響	34,178	企業	13,034
メール	33,195	中国	12,430
商品	30,401	経済	11,448
ネット	26,937	対策	9,322
通販	25,403	首相	9,174
会社	24,667	東京	8,503
連絡	21,327	米国	7,730
返金	20,998	拡大	7,663
契約	20,153	世界	6,249
キャンセル	19,356	新型	5,311
解約	18,427	情報	5,008
キャンセル料	16,189	影響	4,996
予定	14,805	ウイルス	4,990
感染	14,531	政策	4,975
代金	13,181	調査	4,845
給付金	12,538	支援	4,656
大手	12,174	対象	4,640

2. 新聞記事データのキーワード分析

新聞記事データの高頻度キーワード上位 20 語について、本研究で対象とする期間内の変動を月別に見る。図 B-1 は 2020 年 1 月から 12 月までの新聞記事データに含まれる上位 20 語の出現頻度を積み上げたグラフである。また、図 B-2 は 2020 年 1 月から 12 月までの新聞記事データに含まれる上位 20 語の出現割合を積み上げたグラフである。図 B-1 から、年間を通して上位 20 語のキーワードは 1 月から 3 月にかけて急増しているが、これは PIO-NET の高頻度キーワードにも含まれるキーワード「感染」の増加が原因であることが見て取れる。

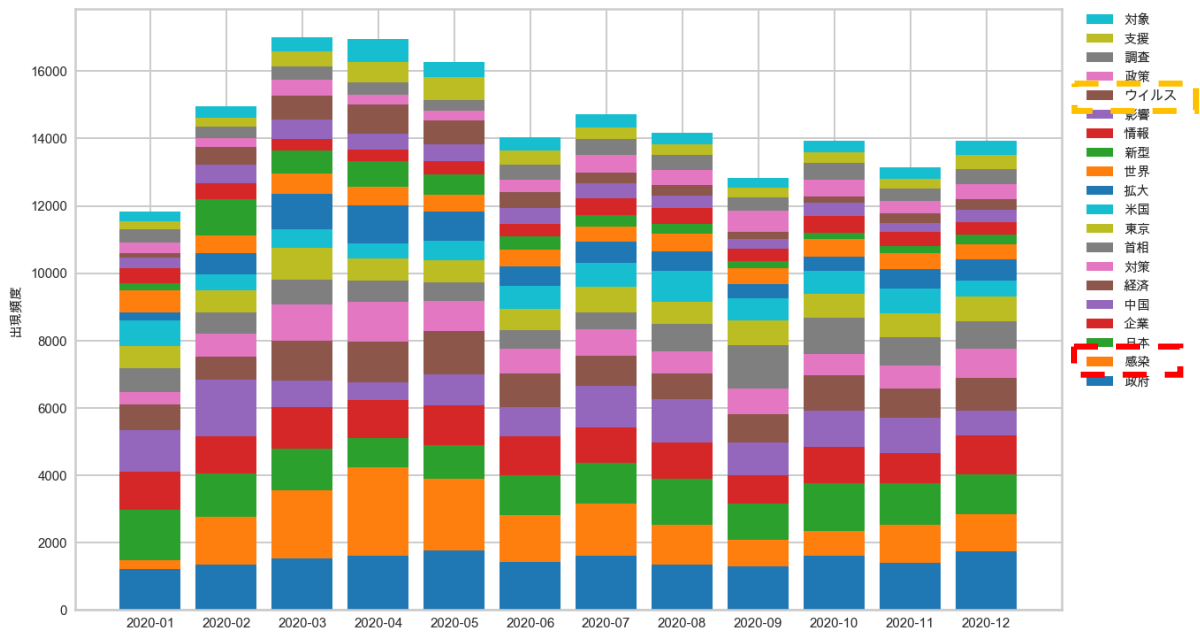


図 B-1：新聞記事データ内の高頻度キーワード上位 20 語と月別の出現頻度

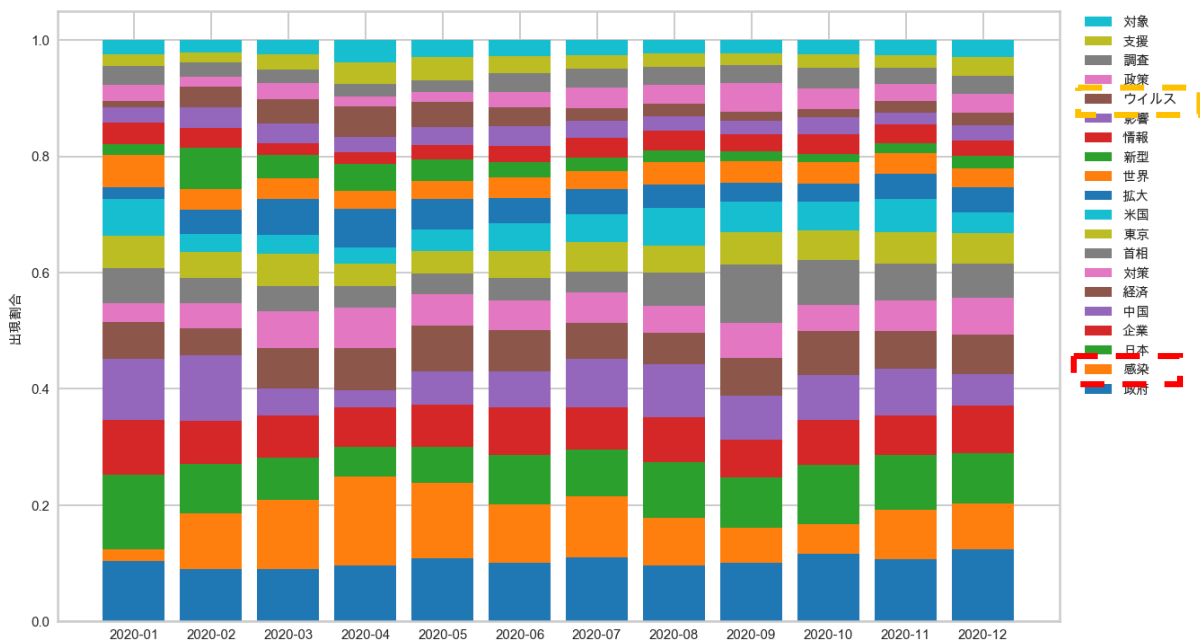


図 B-2：新聞記事データ内の高頻度キーワード上位 20 語と月別の出現割合

3. 新聞記事データのトピックモデリング

新聞記事データの2020年1月から12月までの全データについて、記事1件を文書1件としてTF-IDFを計算し、LDAで分析した結果である8つのトピックを図B-3に示す。このとき、トピックに含まれるキーワードの大きさは、キーワードによってそのトピックが生成される確率（生成確率）に基づく相対的な文字サイズである。個々のトピックにおいて生成確率が最大となるキーワードが最大の文字サイズとなるよう表示されるが、トピック間での文字サイズは相対的ではない。また、配色は機械的に割り当てられたものであり、生成確率の大きさ等の意味合いはない。

それぞれのトピックに含まれるキーワードのうち、生成確率が最大のキーワードをトピックの代表キーワードとしたとき、それぞれ、「感染」「事件」「検査」「中国」「企業」「首相」「死去」「企業」であった。この結果、新聞記事が社会全体の関心を集めていると仮定したとき、キーワード「感染者」で代表されるトピックに新型コロナウイルス関連の話題が含まれている。



図 B-3：2020 年日本経済新聞のトピック（8 トピック）

4. 新聞記事データのトラブルキーワードによる時系列分析

トラブルキーワード 28 語が日本経済新聞の朝刊の記事に含まれる件数を月別で見ることで、消費者のトラブルが新聞記事に反映されたかについて見ていく。図 B-4 は新聞記事データの 2020 年 1 月から 12 月までの新聞記事に含まれるトラブルキーワードの出現件数を積み上げたグラフである。全体を通して「感染」や「対策」の出現頻度が多く、「マスク」は 3 月から 5 月にかけて増加しており、社会的に注目されたことがわかる。

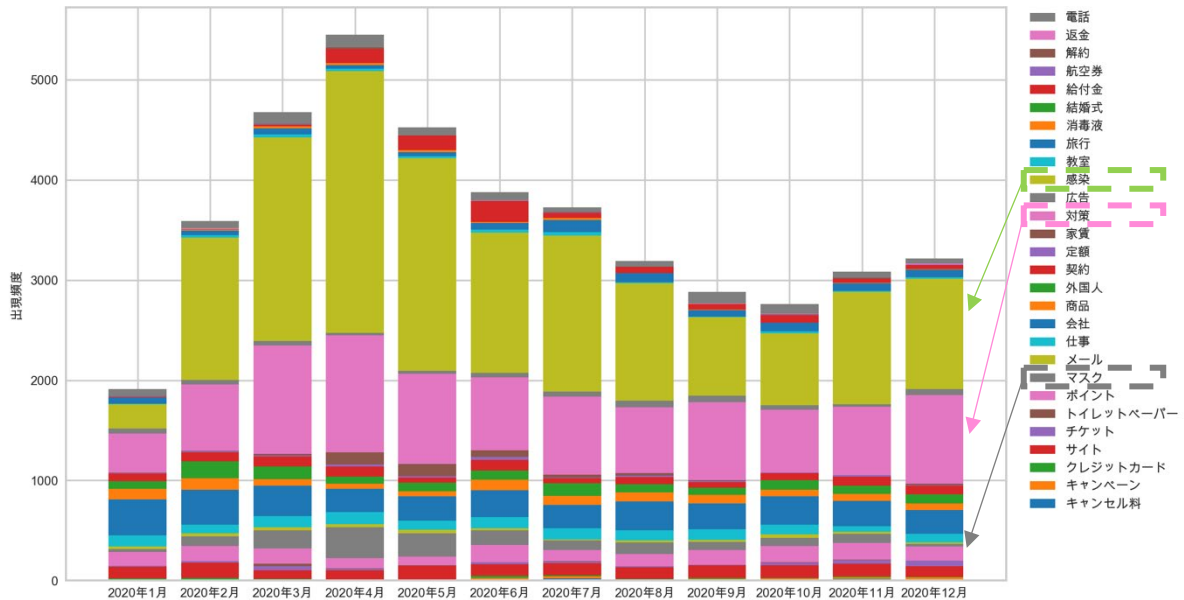


図 B-4：新聞記事内のトラブルキーワードと月別の出現頻度

図 B-5 は 2020 年 1 月から 12 月までの新聞記事に含まれるトラブルキーワードの出現割合を積み上げたグラフである。「給付金」については PIO-NET データと同様に 6 月頃に急増している。また、PIO-NET では目立った増減のなかった「家賃」については、4 月から 6 月で増加している。一方で、「キャンペーン」や「キャンセル料」の割合は低いことから、消費者のトラブルと社会的に注目された話題の傾向には違いがあることがわかった。

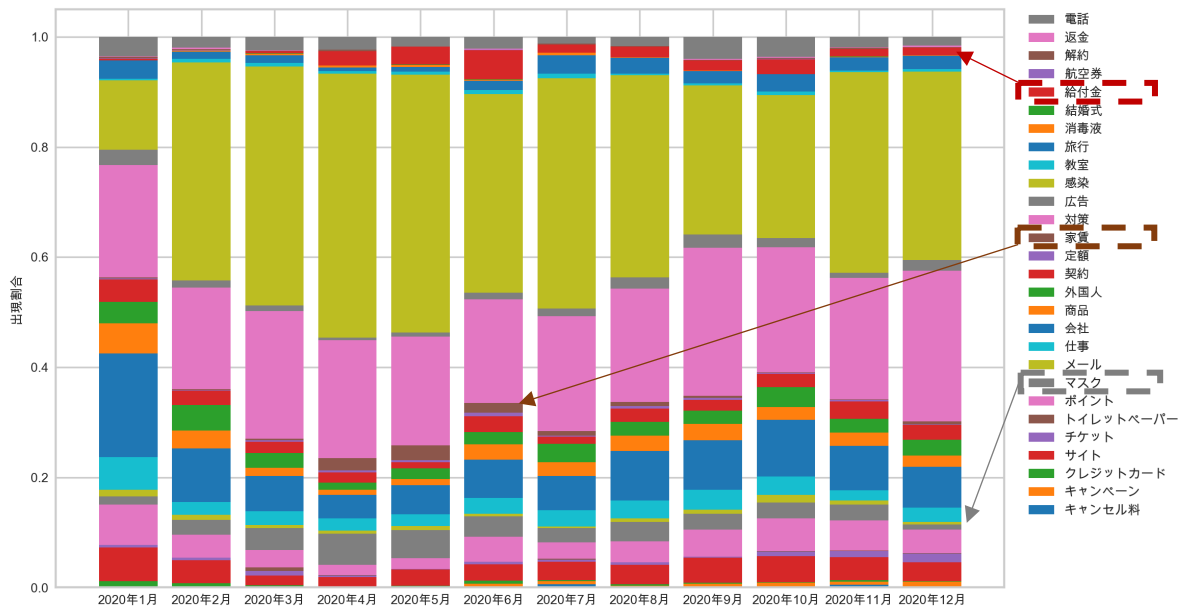


図 B-5：新聞記事内のトラブルキーワードと月別の出現割合

表 C-3：PIO-NET データの月別トピックと主な相談内容（2020年3月）

	<p>「トイレトペーパーやマスクの在庫がなく買えない」</p>
	<p>「各種教室を休会したいが規約上できず、利用していないのに利用料が請求される」 「新型コロナウイルスの影響で予定を中止することになったが、購入したチケット等の返金に応じてくれなかった」</p>
	<p>「フリマアプリやネット通販サイトでマスク等が高額に取引をされている」</p>
	<p>「誤った情報が流布し、マスクの買い占め等が横行している」</p>
	<p>「通販サイトでマスクを注文し代金を振り込んだが、入金連絡も商品の発送連絡もなく詐欺サイトのようなだ」</p>
	<p>「各種サービスの契約を解約したいができない」</p>
	<p>「新型コロナウイルスで結婚式をキャンセルしたら高額なキャンセル料を請求された」</p>
	<p>「ネットから予約した海外 OTA で航空券とホテルをキャンセルするため電話をしているが繋がらない」 「予約していたホテル、航空券のキャンセル料を請求された」</p>

表 C-4：PIO-NET データの月別トピックと主な相談内容（2020年4月）

	<p>「新型コロナウイルス流行が原因にも関わらず、結婚式をキャンセルしたら解約金を支払うことになった」 「結婚式を延期することになったが、延期の手数料を請求された」</p>
	<p>「新聞広告でマスクが高額販売されている」 「新聞広告においてマスクを販売している会社は信用できるのか」</p>
	<p>「注文した覚えのないマスクが届いたが、発送元の事業者と連絡がつかず、手元にあるマスクをどうしたらいいのかわからない」</p>
	<p>「習い事教室の休会・退会について、自己都合・業者都合に関わらず、休会費・退会費や月謝等を支払わなければならない」</p>
	<p>「（新型コロナウイルスの流行が原因にも関わらず）旅行のキャンセル料が自己負担になっていて、返金措置がとられていない」</p>
	<p>「店員がマスクを着けていない。店舗の感染対策がきちんととられていない」 「給付金の申請手続きを代わりに行うという電話やメールがきた」</p>
	<p>「代金は支払い済みなのに、通販サイトで購入したマスクが届かない」 「通販サイトでマスクを購入したが、商品説明にあるものとは異なる商品が届いた」</p>
	<p>「新型コロナウイルスの影響で仕事や収入が減ったため、エステや駐車場等を解約したい」 「大学進学に伴い県外のアパートを契約したが、新型コロナウイルスの影響で当面休校となったので一旦解約したい」</p>

表 C-5：PIO-NET データの月別トピックと主な相談内容（2020年5月）

	<p>「マスクが届かない（違うものが届いた）ので業者に問い合わせたが、対応してもらえないので消費生活センターに相談したい」</p>
	<p>「メール、はがきなどで給付金の振込サイトの案内が届き、個人情報（マイナンバー、氏名等）の提供を要求された。おそらく詐欺的な手法だと思うので情報提供をする」</p>
	<p>「購入商品のキャンセルがしくて事業者（クレジットカード会社等）に電話したが、新型コロナウイルスの影響で電話がつかない」</p>
	<p>「注文した覚えのないマスクが中国から届いた。中に請求書が入っているがどうしたらよいか」</p>
	<p>「海外旅行に行けなくなった（習い事などの教室が休止になった）ため、キャンセルして返金を求めたが、規約上できないと言われた」</p>
	<p>「通販サイトで購入したマスクの返品を要求したが、対応できないと言われた」</p>
	<p>「ヨガ教室が休会になったが、休会中も費用が発生すると言われたため退会を申し出たところ、違約金を請求された」</p>
	<p>「広告を見て商品を購入（サービスを契約）したが、新型コロナウイルスの影響で仕事がなく支払いが厳しい。解約はできるか」</p>

表 C-6：PIO-NET データの月別トピックと主な相談内容（2020年6月）

	<p>「身に覚えのない料金のカード請求があり、カード会社に電話しても新型コロナウイルスの影響でなかなか電話が繋がらない」 「子どもが勝手にクレジットカードを使ってオンラインゲームで課金し、高額な請求をされている」</p>
	<p>「新型コロナウイルスが不安で習い事教室を退会したいが、高額な解約料（違約金）を請求された」 「新型コロナウイルスが不安で習い事教室を休んでいたがその間の会費が返金されない」</p>
	<p>「特別定額給付金に関して、行政を名乗る不審な申請書類の送付や電話があった」</p>
	<p>「通販サイトで商品を購入したが、予定日を過ぎても商品が届かず、事業者とも連絡がとれない」</p>
	<p>「チケット代（航空券、コンサート）の払い戻しの際にキャンセル手数料が発生し、全額返金されなかった（返金に応じてもらえない）」</p>
	<p>「注文した覚えのないマスクが海外から届いた」</p>
	<p>「新型コロナウイルスの影響で収入減や失業し、借金やローンが返済できない」</p>
	<p>「商品・サービスについて電話で問い合わせた際の事業者の対応に問題がある」 「新型コロナウイルスで人員を削減している影響により、事業者へ電話が繋がらない」</p>

表 C-7：PIO-NET データの月別トピックと主な相談内容（2020年7月）

	<p>「持続化給付金の申請代行サービスを使って税申告をしたが、不正受給かもしれないので申請を取り下げたい」</p>
	<p>「期間限定無料と言われ教室に入会したが、新型コロナウイルスの感染が心配だったため解約を申し出ると違約金を請求された」</p>
	<p>「特別定額給付金を受給した後で給付金決定のお知らせが届いた。詐欺なのではないか」</p>
	<p>「大手通販サイトを通じてマスクを注文したところ、不良品が送られてきたため返品したが、返金してもらえない」</p>
	<p>「訪問販売で商品を分割購入したが、コロナ禍になり使用しないので、残債を負担しないようにしたい」 「事業者につながらずサービス解約に時間がかかったため、最初に連絡した時に遡って既払金の返還を求めたい」</p>
	<p>「OTA で海外航空券を購入した。新型コロナウイルスの影響で渡航できずキャンセルになったが、航空会社が返金処理してくれない」</p>
	<p>「注文した覚えのないマスクが届いたが、請求書等、何も入っていない」</p>
	<p>「結婚式を予定していたが、新型コロナウイルスのまん延により延期した後にキャンセルしたところ、高額なキャンセル料の支払いを求められた」</p>

表 C-8：PIO-NET データの月別トピックと主な相談内容（2020年8月）

<p>仕事 収入 影響</p>	<p>「新型コロナウイルスの影響で収入が減り、クレジットカードやスマホの支払ができない」</p>
<p>契約 解約 教室</p>	<p>「期間限定無料と言われ教室に入会したが、新型コロナウイルスの感染が心配だったため解約を申し出ると、違約金を請求された」</p>
<p>マスク サイト 通販</p>	<p>「大手通販サイトを通じてマスクを注文したところ、不良品が送られてきたため返品したが返金してもらえない」</p>
<p>メール サイト 広告</p>	<p>「ネット広告で定期購入した商品を解約したいが、連絡（電話、メール等）がつかない」</p>
<p>延期 航空券 キャンセル料</p>	<p>「新型コロナウイルスのまん延で行けなくなったが、ネットで予約した海外航空券とホテル代の返金がいまだにない」 「結婚式を予定していたが、新型コロナウイルスのまん延により延期した後にキャンセルしたところ、高額なキャンセル料の支払いを求められた」</p>
<p>マスク チケット 請求書</p>	<p>「注文した覚えのないマスクが届いたが、請求書等も何も入っていない」 「新型コロナウイルスの影響でコンサートが中止となったが返金されない」</p>
<p>感染 対策</p>	<p>「お店の新型コロナウイルス対策に不満がある」 「購入した商品について、苦情を入れたが事業者の対応に不満がある」</p>
<p>給付金 持続化 申請</p>	<p>「持続化給付金の申請代行サービスを使って税申告をしたが、不正受給かもしれないため、申請を取り下げたい」</p>

表 C-9：PIO-NET データの月別トピックと主な相談内容（2020年9月）

	<p>「通販サイトで購入した商品を返品したいが、事業者のお客センターに電話が繋がらない」</p>
	<p>「新型コロナウイルスの影響により収入がなくなり家賃や債務の返済ができない」</p>
	<p>「携帯電話料金の支払いが新型コロナウイルスの影響で遅れ、解約を申し出たが、解約料が必要と言われた」</p>
	<p>「新型コロナウイルスの影響からサービス解約を申し出ているが、電話もつながりづらく返金にも応じてくれない」</p>
	<p>「大手サイトでチケット（旅行、イベント等）を購入したが、新型コロナウイルスの影響から中止になったため、返金の申し出をしたが、返事も遅く、対応が滞っており不満」</p>
	<p>「SNS で知り合った人から持続化給付金の申請を勧められたが、不正であると知り申請を取り下げたい」 「新型コロナウイルスの影響により予定していた結婚式の延期（キャンセル）を申請したが、延期料金（キャンセル料）を請求された」</p>
	<p>「航空券とホテルを旅行代理店のサイトから申し込んだが、GoTo キャンペーン適用の手順ではなかったため適用とならなかった」</p>
	<p>「ネットで商品を購入したが、広告ではうたわれていなかった定期購入の商品であったことが購入後に判明した」</p>

表 C-10：PIO-NET データの月別トピックと主な相談内容（2020年10月）

	<p>「頼んだ覚えのないマスクが届いた」 「頼んだ覚えのないマスクの請求書が大手通販サイトからメールで送られてきた」</p>
	<p>「当初の話と違うため電話でサービスの解約を申し入れたが、料金が発生すると言われた」</p>
	<p>「新型コロナウイルスの影響で収入が減り、クレジットカードの支払ができない」</p>
	<p>「GoTo トラベルのクーポンが正しく発行されなかった」 「GoTo トラベルのクーポンが使えなかった」</p>
	<p>「スマホアプリで契約したサービスの解除をしたいが、メールで連絡しても返事がなく、電話もつながらない」</p>
	<p>「新型コロナウイルスの影響で航空券をキャンセルすることになったが、航空会社と返金でトラブルになっている」</p>
	<p>「結婚式を予定していたが、新型コロナウイルスによりキャンセルしたところ、キャンセル料の支払いを求められた」</p>
	<p>「家賃支援といった持続化給付金のはがきが届いた」 「持続化給付金を不正に受給してしまった」</p>

表 C-11：PIO-NET データの月別トピックと主な相談内容（2020年11月）

	<p>「マスクを通販で購入したが粗悪品だった（数量が誤っていた）ので返品したい」 「注文した覚えのないマスクが届いた」</p>
	<p>「サイト、アプリ等で商品購入（サービス契約）をしたが、解約したい（苦情を言いたい）」</p>
	<p>「新型コロナウイルスの影響で収入減や失業し、借金やローンが返済できない」</p>
	<p>「GoTo イートキャンペーンのポイントが還元されない（利用できない）」</p>
	<p>「GoTo トラベルを使って予約していたが、感染者数増加による対象外地域になった等の理由でキャンセルを申し入れたところ、キャンセル料を請求された」</p>
	<p>「コンサート、イベント等が新型コロナウイルスで中止になったが返金されない」</p>
	<p>「新型コロナウイルスが理由で海外航空券を解約したが返金されない（手数料がかかると言われた）」</p>
	<p>「結婚式を予定していたが新型コロナウイルスによりキャンセルしたところ、キャンセル料の支払いを求められた」</p>

表 C-12：PIO-NET データの月別トピックと主な相談内容（2020年12月）

	<p>「マスクを購入したが粗悪品だった（数が違う等）ので返品したい」 「注文した覚えのないマスクが届いた」 「通販で購入したマスクが届かない」</p>
	<p>「新型コロナウイルスが理由で海外航空券を解約したが、返金されない（手数料がかかると言われた）」</p>
	<p>「アプリ（サイト）でカードを使って決済したが、商品が届かない」 「アプリ（サイト）で決済したところ、二重に料金を請求された」</p>
	<p>「GoTo トラベルのキャンペーンが中止になったので予約をキャンセルしたいが、キャンセル料が発生してしまう」</p>
	<p>「自宅に販売の電話、勧誘の電話等がしつこくかかってくるが、対策はないか」</p>
	<p>「結婚式を予定していたが、新型コロナウイルスによりキャンセルしたところ、キャンセル料の支払いを求められた」</p>
	<p>「新型コロナウイルスで収入が減り家賃が払えなくなった、退去費が高額で困っている」</p>
	<p>「家賃支援といった持続化給付金のはがきが届いた」 「持続化給付金を不正に受給してしまった」 「給付金がいまだに還付されない」</p>

参考文献

- [1] 独立行政法人国民生活センター（2021）「PIO-NET の紹介」
<http://www.kokusen.go.jp/pionet/>（最終アクセス日：2021年11月14日）
- [2] M. Honnibal and I. Montani: spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing, 2017.
- [3] 松田寛, 大村舞, 浅原正幸: 短単位品詞の用法曖昧性解決と依存関係ラベリングの同時学習, 言語処理学会第25回年次大会, 2019.
- [4] K. Takaoka, S. Hisamoto, N. Kawahara, M. Sakamoto, Y. Uchida, Y. Matsumoto: Sudachi: a Japanese Tokenizer for Business, Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), European Language Resources Association (ELRA), May 7-12, 2018.
- [5] 大島裕明, 中村聡史, 田中克己: SlothLib Web サーチ研究のためのプログラミングライブラリ, 日本データベース学会 letters, 日本データベース学会, Vol. 6, No. 1, pp. 113-116, 2007.
- [6] 消費者庁（2021）「令和3年版 消費者白書」
- [7] D. M. Blei, A. Y. Ng, and M. I. Jordan: Latent Dirichlet allocation. *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022, 2003.
- [8] 佐藤一誠, 奥村学: トピックモデルによる統計的潜在意味解析, コロナ社（自然言語処理シリーズ）, 2015.
- [9] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830, 2011.
- [10] C. M. Bishop: Chapter 8. Graphical Models, in *Pattern Recognition and Machine Learning*. Springer, pp. 359–422, 2006.
- [11] G. Salton, A. Wong, and C.-S. Yang: A Vector Space Model for Automatic Indexing, *Commun. ACM*, Vol. 18, No. 11, pp. 613–620, 1975.