

テキストマイニングを用いた地域防災計画における災害廃棄物管理の分析-熊本地震における検証

劉 英楠^{†1} 畑山 満則^{†2}

概要：地域防災計画は災害対策基本法（第 40 条）に基づき、防災のために処理すべき業務などを具体的に定めた計画である。この計画は各地方自治体（都道府県や市町村）の長がそれぞれの防災会議に諮り策定されている。しかしながら、災害時になると自治体（特に市町村）では計画の運用に手間取ることも多く、それゆえに問題視されることも多い。実効性の高い計画の作成の手法に関しては、すでいくつかの研究がなされているが、本研究では、テキストマイニングの手法を用いた比較分析によるアプローチから課題を指摘する手法の開発を試みる。本発表では、特に災害廃棄物管理について焦点をあてた分析結果について報告し、また分析結果について、2016 年の熊本地震の事例で検証を行う。

キーワード：地域防災計画、自然言語処理、テキストマイニング、キーワード抽出

Comparative Analysis on Disaster Debris Management in Local Disaster Management Plan with Text Mining Methods- Verify with Kumamoto earthquake

LIU YINGNAN^{†1} MICHINORI HATAYAMA^{†2}

Abstract: The local disaster management plan decides concretely the act of the local governments for disaster reduction. This plan is based on the disaster countermeasures basic act. The head of municipality formulate the details of the plan. However, Sometimes when a disaster occurred, it does not always work so well even the municipality government act on the plan. This problem is gaining attention. Although, Some research has been done to develop a more effective local disaster management plan, no studies have ever tried to ameliorate the feasibility of the plan. This research will try to find problems in local disaster management automatically and will verify the rightfulness with Kumamoto earthquake occurred in 2016.

Keywords: NLP, Keyword extraction, Text mining, Local disaster management plan

1. はじめに

1.1 研究背景

地域防災計画は、災害対策基本法（第 40 条）に基づき、各地方自治体（都道府県や市町村）の長が、それぞれの防災会議に諮り、防災のために処理すべき業務などを具体的に定めた計画である。災害が発生するたびに、地域防災計画に書かれている内容に従って災害の対応をするのは一般的である。しかしながら、実際災害が発生する時に、自治体特に市町村では計画の運用に手間取ることも多く、それゆえ問題視されることも多い。地域防災計画は人が作成したものであるため、不備や問題点は当然存在し、定期的な見直しが必要である。実際見直しは定期的に行われているが、その量が非常に多く、専門家が全部詳細に見るのはほぼ不可能であり、且つ統一した基準で計画の良し悪しを判断することが困難とされている。

1.2 研究の目的

本研究では、地域防災計画の順調な実行の妨げになりうる箇所をテキストマイニングを用いて自動的に見つけ出し、地域防災計画の改定・見直しに貢献でき、最終的に地域防災計画の質の向上に繋げることを目標とする。そのためにまず地域防災計画の既存の問題点を利用し、潜在的な問題を自動的に発見できるアプローチを確立させる必要がある。本稿では、問題発見アプローチを確立させるための分析を行い、その結果を報告し、さらに発見した問題点の候補を熊本地震の事例において検証を行う。

2. 先行研究

本研究に関連する先行研究は以下が挙げられる。

(1) 地域防災計画における情報伝達の機能的障害の発見と見直しに関する研究

畑山ら[1]は、地域防災計画における情報伝達の機能的障害の発見手法の開発をしたが、自動的にかつ大量な文書からの問題発見はできなかった。本研究では大量な文書から問題発見の自動化を図る。

^{†1} 京都大学情報学研究所 社会情報学専攻修士課程
Kyoto University

^{†2} 京都大学 防災研究所
Kyoto University

(2) 災害廃棄物処理に関連する研究

平山ら[2]は、「広域災害時における災害廃棄物処理の広域連携方策に関する研究」で、阪神・淡路大震災における廃棄物処理の事例と経験を踏まえ、広域災害時の災害廃棄物処理のフローに問題があると指摘した。そこで5つの広域連携方法を提案し、それに対して首都直下地震のシナリオに基づくシミュレーションを行った。その結果、広域災害に対応できる広域連携方策を示した。しかし、地域防災計画にそれらの内容が書かれているか否かを検証する方法については言及していない。本研究では、地域防災計画中の災害廃棄物処理に関連する章に焦点を当て、その内容を分析し、広域連携の問題を含めどのような問題が存在するかを探る。

岡山ら[3]は、自治体の水害廃棄物処理計画の策定状況を把握するため、全国の1,742市町村および都道府県の災害廃棄物・災害瓦礫処理を担当している部署宛に対してアンケート調査を行った。その結果、36%の自治体が廃棄物処理計画を策定しておらず、さらに全体の49%の自治体がこれまでに災害廃棄物処理の経験がないと回答している。また、計画に対する実効性のなさや、仮置場の設置や廃棄物の分別に関する問題が指摘されている。

(3) テキストマイニングに関連する研究

那須川ら[4]が、テキストマイニングで用いられる技術や分析手法についてレビュー論文を発表した。テキストマイニングでよく用いられる基盤技術について紹介した。テキストマイニングは、自然言語処理やデータマイニングなど多様な技術を組み合わせた複合技術であり、テキスト自動分類、重要語句の抽出など、様々な課題を解決するために研究されてきた。テキストマイニングの技術でどのような問題を解決できるかを示したため、ここで記述されている知見を踏まえて、地域防災計画の分析への応用を検討する。

(4) キーワード抽出に関する研究

文書の中からキーワードを抽出する技術は、自然言語処理の分野の中でも重要な研究領域の一つとされており、今まで数多くのキーワードを抽出する方法が研究されてきた。例えば Tf-idf, BM25[5]といった古典的な統計情報を使用し、文書内の単語にスコアをつける手法がある。この手法ではコーパス全体において特徴的な単語に対し、高い点数をつける傾向にある。また、PageRank を自然言語処理のプロセスに適用し、グラフ理論を駆使した TextRank[6]や、辞書や形態素解析に頼らない N-gram を利用したキーワード抽出[7][8]などが検討されているが、どのようなデータに対しても精度が高いアルゴリズムは存在しない。その理由は、キーワード抽出の正解は人間が作成するものであるため、人の経験の相違によって正解も変化する点があるので、評価は難しいとされている。扱うデータの性質やタスクや目的も様々なので、研究者各自が自分に需要によってアルゴ

リズムをアレンジする必要がある。

3. 関連技術

本研究では、以下の技術を使用している。

3.1 形態素解析

形態素解析とは、文法的な情報の注記のない自然言語のテキストデータから、対象言語の文法や辞書と呼ばれる単語の品詞等の情報に基づき、形態素の列に分割し、それぞれの形態素の品詞等を判断する作業である。本研究では、形態素解析のツールとして MeCab[9]を使用する。なお、形態素解析の精度の向上のため、拡張辞書として MeCab 新語辞書[10]を導入している。

3.2 Opaki BM25

本研究では、Tf-idf を改良した BM25 アルゴリズムをベースにキーワードの抽出を行う。BM25 とは、Tf-idf 考え方に基づき、同じ単語が同じ文に大量に出現しても、単語のスコアが極端に大きくならないように Tf の値に補正をかけたものである。BM25 を選んだ理由として、実装が相対的に容易であり、Tf-idf の改良版であるため、Tf-idf より高い精度が出せると言われているからである。また、TextRank では、キーワードの抽出の計算は一つの文書内で完結するため、コーパス全体の性質を反映しにくいという問題が存在する。今回の分析対象は地域防災計画の災害廃棄物関連章であり、TextRank を適用した場合、キーワードの抽出は1自治体の文書で完結し、自治体間の特徴が把握できない。よって本研究では、BM25 を採用する。

前述のように、BM25 は広く知られている特徴量の Tf-idf の改良版の一つである。算出された BM25 の値は、Tf-idf と同じ、文書に含まれる単語がどれほど特徴的であるかを示す値である。各自治体の文書の中で、ある単語の BM25 のスコアは以下のように算出される：

$$w(d, k) = \frac{(k_1 + 1) \cdot tf(d, k)}{k_1 \cdot \left((1 - b) + b \cdot \frac{dl}{avdl} \right) + tf(d, k)} \cdot \log \left(\frac{N - df(k) + 0.5}{df(k) + 0.5} \right)$$

(dl : 単語が含まれる文の長さ。 avdl : この自治体のすべての文の長さの平均値。 k1, b : 定数のパラメータ。 本研究では k1=2.0, b=0.75 とする)

なお、本研究では、「18自治体の廃棄物処理に関連する章」を文書全体とする。

4. 提案手法

4.1 提案システム全体の構成

本稿では主に廃棄物処理の章・節に焦点を当て、キーワード抽出結果を手動で分析しているが、長期的な目標として、システムを図1のように提案する。

(1) データの収集, データクリーニング: 地域防災計画の

本文の PDF ファイルを収集し、テキストデータへ変換し、変換されたテキストデータを処理しやすくように整形する。

(2) テキストマイニング: テキストデータをマイニングの手法を適用し、各自治体の計画の文書の統計的な情報などを取り出す。

(3) 結果分析: 取り出した情報を手動で分析し、問題点になりうる箇所を見つける。

(4) 問題選別: 問題点になりうる箇所を発見した場合、該当自治体にヒアリング調査などを行い、実際に問題点であるかどうかを確認する。また、人がどのように問題点を見つけたかを分析し、その過程を自動化することが可能か否かを検討する。

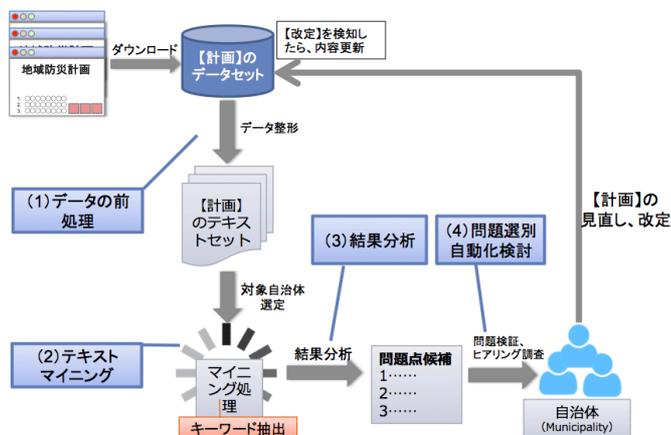


図1 提案システムの構成

4.2 実験手法の概要

本稿では、テキストマイニングの手法としてキーワード抽出を適用し、抽出されたキーワードから各地域防災計画の特徴を分析することを試みた。さらに、普遍的な単語が書かれているかを確認し、それが記述されていない場合、内容的な漏れを意味する可能性が高いと捉えた。

5. 分析対象とするデータソース

5.1 地域防災計画

前述のように、地域防災計画とは、各自治体を持つ防災のために処理すべき業務などを具体的に定める計画である。地域防災計画の本文は、自治体によってその自治体のホームページで PDF 形式にて公開されていることが多い。地域防災計画の PDF ファイルの収集を自動化することがある程度できたが、PDF ファイルから「廃棄物処理」に関連する章の抽出は自動化できていないため、手動で「廃棄物処理」の関連章を抽出する必要がある。手動での抽出は時間と労力がかかるため、今回の分析対象は 20 文書前後が現実的だと判断した。なお、本研究では、各自治体の地域防災計画の「本編」の内容のみを分析対象とし、「資料編」は分析の対象外とする。

5.2 「災害廃棄物処理」関連章

「災害廃棄物処理」に関連する章・節を分析対象として選択したのは、主に二つの理由がある。

一つ目は、地域防災計画では、「災害廃棄物」の処理に関連する内容を策定することが国道交通省と環境書の防災業務計画に定められ、ほぼ全ての地域防災計画には災害廃棄物処理に関連する章・節があり、地域防災計画にとっては欠かせない内容となっている。

二つ目は、地域防災計画の全体を俯瞰すると、水害編しか出現しない単語、原子力編しか出現しない単語は存在し、コーパス全体と比べて単語の出現頻度の偏りが存在する。しかし廃棄物処理は、どの章においても避けられない問題として存在するため、どの章にも「廃棄物処理」に関連する内容があり、様々な章から分析対象を抽出しているため、分析対象とするテキストの母集団の単語の多様性が保証され、コーパス中に出現する単語の偏りを抑止することができる。

5.3 対象とする自治体の選定

現在、日本全国には 1,718 自治体が存在する[11]。将来的には全ての自治体を分析の対象とする予定であるが、適切な問題発見アプローチを模索する現在のプロセスにおいては全ての自治体を分析ことは困難である。従って、対象地域の選定を行う必要がある。

近年発生した大小様々な規模の災害の中で、土木学会が災害調査団を派遣した災害は、社会的関心が集められる規模の災害であると広く認識されている。従って、土木学会が災害調査団を派遣した災害の被災自治体を対象に選定した。なお、被災の年と地域防災計画の更新年の関係で、対象とする自治体を被災した後に改定(被->改)と改定した後に被災(改->被)の二種類に分類できる。対象となる自治体の地域防災計画を収集し、そのテキストを分析する。選定された自治体の情報を表 1 にまとめた。なお、熊本地震の被災地である自治体は 6 つ選択し、その中で熊本市と

宇土市では「改定->被災」「被災->改定」両方の地域防災計画を取得することができた。

表 1 選定された自治体および分類

自治体名	災害名	被災年	計画更新	分類
東嶺村	九州北部豪雨	2017.7	公開なし	公開なし
朝倉市	九州北部豪雨	2017.7	2016.4	改->被
日田市	九州北部豪雨	2017.7	2015	改->被
中津市	九州北部豪雨	2017.7	2015	改->被
合志市	熊本地震	2016.4	2014	改->被
熊本市_H27	熊本地震	2016.4	2017	改->被
宇土市_H24	熊本地震	2016.4	2012	改->被
熊本市_H29	熊本地震	2016.4	2017	被->改
阿蘇市	熊本地震	2016.4	2017	被->改
宇土市_H29	熊本地震	2016.4	2017	被->改
天草市	熊本地震	2016.4	2017	被->改
菊池市	熊本地震	2016.4	2017	被->改
岡崎市	H20年8月末豪雨	2008.8	2017.2	被->改
清水町	北海道豪雨	2016.8	2014.2	改->被
池田町	北海道豪雨	2016.8	2016.3	改->被
帯広市	北海道豪雨	2016.8	2017.2	被->改
盛岡市	北海道豪雨	2016.8	2017.8	被->改
音更市	北海道豪雨	2016.8	2017.2	被->改
常総市	鬼怒川氾濫	2015.9	2015.3	改->被
小山市	鬼怒川氾濫	2015.9	2017.11	被->改
日光市	鬼怒川氾濫	2015.9	2016	被->改
鹿沼市	鬼怒川氾濫	2015.9	2017.3	被->改

6. 実験

6.1 実験の手法

5.3章で選出した18自治体を実験対象とし、それぞれの自治体の地域防災計画がどのような特徴を持つかを分析する。

6.1.1 不要単語リストの作成

廃棄物処理に関連する章の中では、ほぼ確実に出現し、かつ分析にあたって必要性を感じない単語が存在する。単語自体は普遍的に各自治体の計画に高い頻度で出現しているため、BM25の計算式中のTf値が高くなり、抽出されやすい傾向にある。しかしこの中には、文書の特性に依存する頻出単語が含まれるため、これらを排除する必要がある。今回は経験則により、以下の7単語を選出し、「不要単語」とする。

[ゴミ, ごみ, 廃棄物, 廃棄物処理, 災害廃棄物, 災害廃棄物処理, 廃棄物処理計画]

上記の単語が抽出結果に出現した場合、抽出結果から除去する。

6.1.2 BM25 上位キーワード抽出

ある自治体の地域防災計画において、特徴的な単語が多数抽出されたことは、他の自治体ではあまり書かれていない単語が書かれていることを意味する。この性質により、自治体の文書の特徴がわかり、その自治体の独特な取り組みを表す可能性が高く、ポジティブな要素として捉えることができる。そこで、4.1章で述べたように、特徴的な単語(キーワード)の抽出を行い、抽出されたキーワードはどのような意味合いかを考察する。

18自治体の「廃棄物処理」に関連する章・節を手動で抽出し、BM25のアルゴリズムを適用して、各自治体の中でのすべての単語にBM25の値を付与する。各自治体においてBM25の値の上位20単語を取り出し、キーワードの候補とする。ただし、BM25の値はあくまで統計上「特徴的な単語」であるため、上位に来たとしても問題発見にあたり価値のある単語とは限らない。さらに、これらのキーワードの候補を直接確認する必要がある。そこで、今回の実験では「母集団のすべての自治体においてX以上回抽出された単語」を削除し、一定量のキーワード(ポジティブ単語)を抽出する(下式)。

ポジティブ単語 = 「上位 20 単語」 - 「不要語」 - 「重複 X 回抽出された単語」

Xは閾値であり需要によって調整することができるが、今回は4に設定している。

6.1.3 BM25 下位キーワード抽出

BM25で文書内のすべての単語にスコアを付けると、BM25の順位が下位にあるほど単語のTfとidf値が低くなっていく傾向がある。よって下位にある単語は全コーパスにおいて一定程度の普遍性を持っており、どの自治体の計画の文書でも出現しやすい傾向があると考えられる。すなわち、「ほぼ全ての自治体において普遍的に使われた単語」がある自治体の計画にだけ書かれていないことは、その自治体の計画で本来であれば書くべき内容を書いていないことを示唆し、内容的な漏れの発見につながる可能性が高いと考えられる。このような単語が多数抽出されたことは、他の自治体では普遍的に書いている単語が使用されていないことを意味し、文書のネガティブな要素として捉えることができる。

そこで、各自治体のBM25順に下位50%の単語を抽出し、各単語が何回抽出されたかをカウントし、1)抽出回数がY回以上、2)この自治体において一回も出現していない、3)不要単語リストにない、4)その単語は部分文字列として抽出されなかったの4つの条件を全て満たす単語を抽出する。今回、Yは9(計画をテキスト化に成功した自治体数の半数)に設定する。

計算式(集合、論理計算)にまとめる以下ようになる。
 (「下位半数の単語」 & 「Y回以上抽出された単語」 & 「tf(d,k)=0」) - 「不要語」 - 「出現した単語の部分文字列として存在する単語」

6.2 実験結果

16自治体の18文書(熊本市と宇土市は「改->被」「被->改」両方)を用いて実験を行った。実験結果は、表2、表3のようになった。なお、青背景の部分は改定した後に被災であり、赤背景の部分は被災した後に改定である。

表 2 実験結果_ポジティブ単語

自治体	単語数	ポジティブ単語
朝倉市	16	置き場、役割、支障、応急、下水道、処分、得て、考慮、最終、優先、ない、現場、方針、し尿、発生、県
中津市	18	支所、業者、置場、対応、指示、汲、減免、搬入、災害、中津市、単価、許可、確認、状況、当課、重機、罹災証明書、本庁
合志市	18	清掃、処分、設置、手順、詳しい、保管、編成、労働、担当、公衆衛生、家庭、民生、方法、運搬、収集、マニュアル、向上、合志市
熊本市_H27	20	センター、宇城広域連合、浄化、風水害、山鹿植木広域行政事務組合、通常、除去、冷凍庫、パソコン、回復、テレビ、冷蔵庫、洗濯機、エアコン、次、扇田、障害、植木、地区、秋津
宇土市_H24	4	調査、津波災害、資材、点検
清水町	18	廃棄物等、町、被災地、便所、取扱、指導、清水町、ない、使用、共同、埋却、配慮、衛生害虫、最小、恒久、十勝総合振興局、箇所、部班
池田町	17	所有者、町長、トイレ、共同、町民、班、処分場、課長、判明、被害、恒久、野外、最小、委任、箇所、溢水、やき
常総市	16	実施、市町、住民、要請、協力、民間、事業者、県、運送事業、清掃、常総市、指導、作業、広報、状況、集積
熊本市_H29	11	優先、配慮、置場、環境、県、処分、支障、選定、応急、仮設トイレ、住民
阿蘇市	5	隣接、調査、資材、うえ、点検
宇土市_H29	4	調査、候補、相互、広域
天草市	5	市町、保健所、候補、資材、うえ
岡崎市	15	集積、協力、し尿、下水道、化製場、設置、場所、防疫、関係、済、フロン、別冊、附属、取扱、生活
帯広市	14	清掃、班、所有者、取扱、収集、使用、作業、編成、市長、活動、開始、被災地、一般的、応急
盛岡市	20	本部長、障害、消毒、除去、組合、散布、薬剤、集積、便、盛岡、次、臨時、衛生、場所、パキウムカー、適用、災害救助法、イ、ワ、資機材
小山市	15	除染、汚染、土壌、放射性物質、協定、国、特例、対処、除去、確立、防止、搬入、団体、推計、原子力事業者
日光市	16	産業、環境部、焼却、運搬、推計、確立、人員、被災地、搬入、復興、事項、留意、不燃、焼失、収集、倒壊家屋
鹿沼市	17	班、15 筋、ルール、建設業協会、避難、推定、早期、排、出場所、下水道、生活、協定、管理、担当、土砂、区域、地元

表 3 実験結果_ネガティブ単語

自治体	単語数	ネガティブ単語
朝倉市	8	臨時、地震、がれき、被災者、家屋、生活環境、被災地、地区
中津市	10	臨時、能力、優先、使用、保全、家屋、障害、原則、被災地、方針
合志市	6	排出、応急、市町村、地震、使用、方針
熊本市_H27	21	排出、施設、応援、市町村、措置、交通、がれき、優先、被災者、使用、保全、支障、住民、生活環境、被災地、被害、災害時、方針、搬入、運搬、協力
宇土市_H24	7	措置、衛生、焼却、被災地、方法、業者、設置
清水町	12	臨時、施設、能力、地震、交通、がれき、被災者、地域、原則、生活環境、方針、場所
池田町	17	排出、臨時、発生、能力、交通、優先、被災者、保全、支障、地域、家屋、住民、原則、生活環境、焼却、方針、地区
常総市	10	地震、がれき、優先、被災者、障害、原則、生活環境、実施、搬入、処分
熊本市_H29	3	臨時、被災地、地区
阿蘇市	7	措置、地震、衛生、焼却、方法、業者、設置
宇土市_H29	6	措置、支障、衛生、被災地、方法、業者
天草市	3	措置、地震、支障
岡崎市	3	地震、優先、生活環境
帯広市	16	排出、発生、交通、がれき、被災者、保全、支障、家屋、障害、原則、生活環境、衛生、方針、場所、搬入、処分
盛岡市	3	がれき、保全、被害
小山市	13	臨時、状況、応急、能力、計画、市町村、地震、交通、使用、障害、被害、設置、地区
日光市	11	状況、応急、能力、地震、交通、優先、使用、障害、住民、設置、地区
鹿沼市	12	臨時、応急、措置、交通、保全、地域、生活環境、焼却、被災地、業者、災害時、方針

できる。

小山市には市内および周辺 30 キロ圏内に原子力発電所はないにも関わらずこのような単語が抽出された原因を検討する。2017 年 3 月 22 日の毎日新聞[12]によると東海第二原子力発電所の半径 30 キロ圏内の緊急防護措置区域に入っている茨城県笠間市は小山市を含む 5 市町と原発避難に関する協定を結んでいることがわかった。実際に小山市の計画文書を見ると、原子力事故への具体的な対策が盛り込まれていることもわかった。これは他の 18 自治体の文書には見られない取り組みであり、小山市の独自の取り組みとして評価できる。

(3) 熊本市_H27 では、「宇城広域連合」や「山鹿植木広域行政事務組合」などの単語が抽出されたが、改定後の熊本市_H29 ではそれらが消えた。また宇土市_H24 では「広域」という単語が含まれているものは抽出されなかったが、宇土市_H29 では「広域」と「相互」という 2 つの単語が抽出された。このことから、宇土市は被災後に広域連携に対する意欲が向上したのではないかと考えられる。一方で熊本市は熊本地震で被災して広域連携への意欲が低下したのではないかと考えられる。

また、熊本市と宇土市は隣接自治体でありながら被災から学習したことが相違することが推測できる。2016 年の熊本地震では両自治体とも被災し、大量の災害廃棄物を処理する必要があった。熊本市_H27 の地域防災計画では、宇城広域連合の施設を使用して廃棄物処理を行うと明記しており、熊本市が地域防災計画に沿って災害対応を行っていた

7. 考察

6.2 で述べた実験結果からポジティブ単語とネガティブ単語両方に対し、考察を行う。

7.1 ポジティブ単語についての考察

(1) 熊本市_H27 では、具体的な家電名が抽出されている(パソコン、テレビ、冷蔵庫など)。これらの単語が書かれている理由は、家電リサイクル法により対象廃棄物と規定されたものが、地域防災計画において詳細に書かれていることが推察できる。

実際に、18 の計画文書中の廃棄物処理関連章を詳しく調査したところ、18 文書中で対象廃棄物をリストアップした文書は 2 つしかなかった(中津市、熊本市)。家電リサイクル法に規定された対象廃棄物を詳細にリストアップすることは、熊本市の独特な取り組みであると言える。

(2) 栃木県の小山市には、「除染」「放射性物質」「原子力事業者」といった原子力発電所に関連する単語が抽出された。廃棄物処理の章に原子力発電所関連の単語が大量に抽出されたものは他になく、独特な取り組みであると推察

場合、宇土市に廃棄物の処理を依頼していると考えられる。しかし、宇土市が自身の管轄区域内の廃棄物処理で手一杯の状態になり、熊本市の要請を受け入れられなかった可能性がある。このように考えると、熊本地震の被災から熊本市は自力で廃棄物処理をした方が確実であると学習した一方、宇土市は熊本市の要請に対応できなかつたため、今後対応できるようにしようと考えていると推測される。このように、被災後に関係する市町村の間で、両者がそれぞれ考慮していることが食い違っている可能性がある。

また、宇土市、宇城市、美里町は宇城広域連合として、平常時の廃棄物処理の事業を共同で行っている。災害廃棄物対策に関するシンプジウムにおける、熊本県循環社会推進課の資料[13]によれば、熊本市、宇城市、宇土市はゴミ集積場を代用しており、当時の3市の廃棄物処理の大変さが推察できる。また同じ資料によると、事務委託趣旨に「市町村に置く災害廃棄物の処理が困難な事務について、事務の委託により県が処理行う」と記載されている。実際の対応においては、宇土市、南阿蘇村、西原村、御舟町、嘉島町、益城町、甲佐町の7市町村が県に事務委託をした。この内容により、宇土市および他の事務委託を6つの市町村が、自分では廃棄物を処理しきれないということがわかり、他の自治体の要請を受け入れる余裕がなかつたことが推測できる。一方、熊本市が宇土市に廃棄物処理の要請を出したと推測されるが、その裏付けになる記事は見つけられなかつた。

(4) ネガティブ単語の候補である「焼却」が、日光市ではポジティブ語として抽出された。BM25の性質上、日光市が計画の文書中に「焼却」を大量に使用し、他自治体よりも強調していることが分かった。しかしその理由は不明であるため、計画の文書を詳しく読んで判断する必要があると思われる。

日光市の計画本文を詳しくみると、他の自治体と同じく「可燃物やプラスチック類はできるだけ分類し、焼却施設に搬入する」という趣旨以外に、「やむをえない場合露天焼却を行う」という内容が書かれている。他の自治体の文書も調べた結果、「露天焼却」について言及した自治体は3つ(小山市、日光市、合志市)しかなく、やむをえない状況の対策として詳細に書かれていることはポジティブに評価できるが、廃棄物の処理及び清掃に関する法律により、平成13年以降は、露天焼却、いわゆる「野焼き」が禁止されている。例外として、災害対応などのやむをえない場合のみ、露天焼却はあくまで最終手段として許容されているが、実際に災害対応で露天焼却を行ったかどうかは不明である。

7.2 ネガティブ単語についての考察：

(1) 宇土市_H24 と宇土市_H29 両方共に「衛生」という単語が書かれていない。両文書を含め、合計4つの自治体(帯広市、宇土市_24、宇土市_H29、阿蘇市)では「衛生」

という単語が使われていないことがわかり、廃棄物処理の章では、災害後の環境保全、衛生保証に関して言及していないことがわかつた。

宇土市_H24、宇土市_H29 を含め、4つの自治体の文書に「衛生」についての記述がなかつたが、災害後の環境保全や衛生管理は大きな課題であるため、計画中に本当に書かれていないか文書で確認を行った。その結果、「衛生」という表現は使用されていなかったが、阿蘇市、宇土市_24、宇土市_H29 は環境保全に関する記述があつたため、内容的な記述漏れではないことがわかつた。しかし、帯広市の文書には「環境保全」や「衛生」に関する文脈はなく、内容的な漏れではないかと懸念される。

(2) 災害後の交通状況は平常時より悪化し、さらに災害廃棄物が道路などに搬出されることにより道路が塞がることも多く発生することが知られているが、18自治体中7自治体が「交通」について言及していない。

「交通」という単語が出現しなかつた7自治体(帯広市、清水町、小山市、池田町、熊本市_H27、日光市、鹿沼市)の文書を詳しく調べた結果、いずれも内容的に災害後の交通を確保することは書かれていない。一方、「交通」が使われた自治体の文書では、ほとんど「災害後の交通の確保」や、「交通の妨げにならないように廃棄物を出す」といった内容が書かれており、災害後の交通状況を重大な問題の1つとして認識しているように感じる。このように、自治体によっては「交通の保全」に対して意識の差があることがわかつた。

(3) 熊本市_H27 では、18自治体中最多の21単語がネガティブ単語として抽出された。熊本地震の後改定された熊本市_H29 ではネガティブ単語が3単語になり、18単語も減少した。これは、大きな被災が抜本的な見直しを誘起し、その成果としてネガティブ単語が減少したのではないかと推察できる。

また、熊本市は熊本地震を経験した後、地域防災計画を大幅に見直しており[14]、文書の分量がH27の1000字程度からH29の8000字程度になり、内容も充実になったことがわかる。

この事例から、大災害を経験し、元の計画の中で不合理な部分が見つかり、失敗の経験から抜本的な見直しが行われたと考えられる。つまり、ネガティブ語が多い計画は、本来書くべき内容が漏れている可能性が高いことを意味することになり、見直しする際に重点的に調査する必要がある。実際に、18自治体の中で2番目に多くのネガティブ語が抽出された池田町の地域防災計画を確認した結果、生活環境や交通保全などのトピックが言及されていなかったイメージが強く、内容的な漏れが懸念される。

8. おわりに

8.1 まとめ

本稿では、地域防災計画の実行時に発生しうる問題点を自動的に発見するアプローチを模索した。熊本地震を含め、過去の災害から、災害の規模や社会的な関心度の視点から被災自治体を選び、抽出されたキーワードがどう相違するかを比較した。比較手法として、「単語がどれほど特徴的であるか」を示す BM25 を使用し、BM25 のスコアが上位の「特徴的な単語、キーワード」と BM25 のスコアが下位の「他の文書では普遍的に書かれているが当該文書では書かれていない単語」の両方向から単語抽出を行った。その抽出結果を分析、検証した結果、いくつかの知見が得られ、問題点になりうる項目が見つげられた。

- ・独自の取り組みをしている自治体は、ポジティブ語の抽出結果に反映されることが多い。
- ・露天焼却は可燃廃棄物処理の最終手段として存在するが、今までの災害対応で露天焼却が行われたかどうかは不明である。
- ・地域防災計画を比較することで災害時の広域連携がうまくできていない可能性が示された。
- ・自治体によって「災害後の交通の保全」に対する意識に差があることがわかった。
- ・自治体によって「災害後の衛生保全」に対する意識に差があることがわかった。
- ・ネガティブ語が多数抽出される自治体は、内容的に漏れが多い可能性があり、見直しの緊迫性が高いと見られる。
- ・自治体の被災経験が地域防災計画の改定に影響を与える可能性が高いということがわかった。

また、ポジティブ単語とネガティブ単語について、ポジティブ単語はその自治体の特徴を表しているため、積極的な評価に使われることが多い。例えば、計画の見直し前後のポジティブ単語の違いを比較することで、見直しでどのような取り組みがなされたかを推測できる。一方、ネガティブ単語の方は、文書的内容的な漏れを意味する可能性が高く、より緊迫性の高い見直しが必要であることを示している。

8.2 今後の課題

1. 比較対象をさらに広げる：本研究の分析対象は手動で廃棄物関連の章・節を抽出した 16 自治体の 18 文書にすぎない。今後は、さらに多くの自治体の文書を分析の対象とし、最終的に全国の自治体を分析対象とすることが目標である。インターネットから大量の PDF を収集する作業の自動化はある程度できたが、今回のような大量な文書から手動で「廃棄物処理」の関連章を抽出することが非現実的であると考え、対象にする章・節を自動的に抽出する手法を考案する必要がある。
2. 単語抽出精度の向上：本研究では、キーワードの抽出に

より地域防災計画の問題点発見を試みたが、表記の揺れや、単語の間の同義性の認識については考慮していない。例えば「ごみ」と「ゴミ」は違う単語と計算機は捉えるが、人間にとっての意味は同じである。また、実験結果の表の通り、一般的な単語、分析に当たってはほとんど意味のない単語も多数抽出されているように、単語抽出の制度も問題である。今後はキーワード抽出の精度をさらに上げるために単語をベクトルで表現できる Word2Vec などの技術を使い、同義語や意味の近い単語を計算機に認識させ、抽出精度の向上や重複抽出防止に注力していきたいと考えている。なお、他コーパスを同じ抽出アルゴリズムを適用し、一般的な単語は他コーパスでも抽出される可能性が高いことを利用し、一般的な単語の除去を行うことが必要である。

3. 地域防災計画の時系列データベースの構築：今回は熊本地震を含め過去にあった災害の前後で地域防災計画が改定・見直しされたため、前後を比較して議論ができた。これからも災害が発生することや、その都度地域防災計画が改定・見直しされることを想定し、地域防災計画の時系列データベースを作成し、定期的に地域防災計画の最新版を収集する必要がある。このようなデータベースを活用することで、地域防災計画が度々見直され、被災自治体が計画に沿って災害対応をしたがうまく対応できなかった部分の経験が生かされるようになる。被災前と被災後の地域防災計画で、被災自治体がどのような経験をえたのかを調査・整理することで、まだ被災していない自治体の計画の見直しにも役に立つ貴重な知見が見つげられると思われる。

4. 仮説の検証を行う：今回の比較分析で得られた知見や仮説は、どちらも未検証であり、実際の状況や自治体の思惑が違う可能性がある。また、廃棄物処理の章に限定した分析であるため、内容的な漏れだと思われる部分が地域防災計画の他の章や、別冊となる公式文書で書かれている可能性もある。そのため、事実関係の確認を行う必要があり、自治体へのヒアリング調査を含め、より一層深く追求しなければならない。

参考文献

- [1] 畑山満則, 安藤恵. 地域防災計画における情報伝達の機能的障害の発見手法の開発. 情報処理学会論文誌, 2013, Vol.54, no.1 p202-p212.
- [2] 平山修久, 河田恵昭. 広域災害に行ける災害廃棄物処理の広域連携方策に関する研究. 土木学会論文集, 2007, vol. 63, no. 2, p. 112-119.
- [3] 岡山朋子, 伊藤秀行. 自治体の水害廃棄物処理計画の策定状況に関する調査研究. 第 24 回廃棄物資源循環学会研究発表会公同論文集. 2013. A12
- [4] 那須川哲哉, 河野浩之, 有村博紀. テキストマイニング基盤技術. 人工知能学会誌, 2001, vol. 16, no. 2, p. 201-211.
- [5] S.E. Robertson and S.walker, "Some Simple Effective Approximations to the 2-Possion Model for Probabilistic Weighted Retrieval," Proc. 17th Annual International ACM SIGIR Conference (SIGIR'94), 1994.
- [6] R. Mihalcea and P. Tarau. 2004. TextRank – bringing order into

texts.

- [7] 森信介, 長尾真. n グラム統計によるコーパスからの未知語抽出. 情報処理学会論文誌, 1998, vol. 39, no. 2, p2093-2100.
- [8] 梅村恭司. 未踏テキスト情報中のキーワードの抽出システム開発. Technical report, <http://www.ipa.go.jp/archive/NBP/12nendo/12mito/mdata/10-36h/10-36h.pdf>, 2000.
- [9] MeCab: Yet Another Pair-of-Speech and Morphological Analyzer. <http://taku910.github.io/mecab/>
- [10] Mecab-ipadic-neologd. <https://github.com/neologd/mecab-ipadic-neologd>
- [11] 市町村数の変遷と明治・昭和の大合併の特徴. <http://www.soumu.go.jp/gapei/gapei2.html>
- [12] 毎日新聞 地方版/茨城. 東海第2原発: 事故想定 栃木5市町と避難協定 笠間市、県外と初. P27. 2017.03.22
- [13] 平成28年熊本地震における災害廃棄物処理に係る支援の概要. http://kouikishori.env.go.jp/action/d_waste_net/pdf/symposium_171214_kumamoto.pdf
- [14] 毎日新聞 地方版/熊本. 熊本市: 地域防災計画で素案 支援・物資受け入れ搬送も. P21, 2017.01.27