

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に関する比較分析

劉英楠^{†1} 畑山満則^{†1}
京都大学^{†1}

1. はじめに

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。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[1]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、すでに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する。

2. 研究目的

地域防災計画の順調な実行の妨げになりそうな部分を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、その結果を報告する。

3. 対象となる地域防災計画ソースの説明

3.1 地域防災計画

前述のように、地域防災計画とは、各自治体をもつ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。ほぼ全ての地域防災計画の本文は、その自治体のホームページでPDFの形式にて公開されている。

3.2 対象となる地域の選定

日本全国には1,718[2]自治体が存在する。将来的に全ての自治体を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段階では全ての自治体を分析することは困難である。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->改定」と「改定->被災」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。

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

3.3 データの処理・分析

収集したPDFファイルは前処理なしでマイニングをかけることはできないため、テキストファイルに変換・整理する必要がある。今回は災害廃棄物管理に焦点を当てているので、各自治体の地域防災計画の中で「廃棄物処理」に関連する章だけを取り出し、分析対象とする。具体的には：

- 前述した地域防災計画を収集し、PDF形式からテキスト形式に変換する。
- 各自治体の地域防災計画の目次から「災害廃棄物管理」に関連する章を見つけ、その本文の所在位置を特定し、その部分だけを取り出し、別のテキストファイルに複製し、保存する。
- MeCabでテキストファイルに対し形態素解析を行い、形態素解析の結果を用いてキーワード抽出を行う。

4. 実験・分析手法

3.3で述べたように、特徴的な単語（キーワード）の抽出を行い、抽出された単語はどのような意味合いなのかを

分析する。キーワードの抽出アルゴリズムは、Okapi BM25[3]を採用する。BM25とは、Tf-idfの改良版の一つであり、同じ単語が大量に出てきた際に単語のスコアが極端に大きくなるようにTfの値に補正をかけたものである。算出されたBM25の値は、Tf-idfと同じ、文書中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。今回の実験では「各自治体の廃棄物処理に関する章」を文書全体とする。

自治体ごとにBM25のアルゴリズムでキーワードを抽出し、BM25の値の上位20単語を取り出す。なお、各自治体の抽出結果から、「他の自治体においても抽出された単語」と明らかに意味のない単語を除去する。

5. 実験結果と考察

5.1 実験結果

4.1で述べたプロセスに沿い、実験を行った。その結果を表2にまとめた。最終的に18自治体からキーワード抽出ができ、青背景の部分は「改定->被災」であり、赤背景の部分は「被災->改定」である。

表2 実験結果のまとめ

自治体名	抽出単語
東嶺村	-----
朝倉市	素、役割、方針、考慮、置き場、支障、最終
日田市	-----
中津市	指示、中津市、業務、支所、ゴミ、単価、作成、罹災証明書、汲、提出、対応
合志市	合志市、保管、設置、編成、廃棄、労働、家庭、民生
熊本市_H27	宇城広域連合、工場、クリーンセンター、センター、風水害、扇田、宇土、通常、情報、集、山鹿植木広域行政事務組合、富合、宇城、災害救助法、浄化
宇土市_H24	津波災害
清水町	清水町、指導、便所、所要、人夫、埋却、取扱、町
池田町	課長、被害、トイレ、町長、溢水、町民、連絡先、住所、判明、許可、野外
常総市	常総市、地方、日、巡視、規模、要請、電話、実施、併用、常総、対策、方式
熊本市_H29	関連、部局、選定、アスベスト
阿蘇市	隣接
宇土市_H29	広域
天草市	(なし)
菊池市	-----
岡崎市	解体、協力、別冊、がれき処理、附属、化製場、生活
帯広市	作業、電話番号、名称、委託、所在地、終末、くり
盛岡市	衛生、河川、組合、臨時、資機材、地域、盛岡、本部長、柴波
音更市	-----
小山市	国、放射性物質、環境省、除染、原子力事業者、特例、汚染、ガイドライン、措置、土壌、環境大臣、防止、連携
日光市	復興、木材、焼失、倒壊家屋、産業、焼却、環境部、推計、人員、機材
鹿沼市	資料、撤去、建設業協会、推定、相互、栃木県、世帯、ルール、避難、管理

5.2 考察

4章で述べた手法に沿って実験を行い、考察する。

・ キーワードを抽出できた18自治体の文書から前述の手法で重複する単語を除去すると、上位20単語中平均12.28単語が重複していることがわかった。全ての文書が災害対応のためのものであるため、抽出されるキーワードが重複

することが多いと推測できる。

実験結果では、抽出されたキーワードがある程度各自治体の地域特性を表していることがわかる。例えば、中津市には「中津市」、清水町には「清水町」のようなキーワードが抽出された。さらに、熊本市には「宇城市」が抽出され、宇城市は熊本市の隣接町であることが推察できる。この結果から各自治体において抽出されたキーワードが何らかの独特性や地域性があると考えられ、単語が抽出された原因を探求することで問題発見につながられる可能性がある。

・ 被災の時点と改定の時点の関係で分類された「被災->改定」と「改定->被災」の二組ではそれぞれ違う特徴の単語が抽出されると想定したが、結果的に明らかな違いは見つけられなかった。しかし、熊本市と宇土市のそれぞれの被災前と被災後の文書から抽出された単語の違いに興味深い発見があった。熊本市_H27では、「宇城市」や「宇土市」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」が抽出されなかったものの、宇土市_H29では抽出された。この変化から熊本市と宇土市が被災から学習したことが相違することが推測できる。熊本地震では宇土市も熊本市も被災し、両自治体共に大量のゴミを処理する必要があった。熊本市が地域防災計画に沿って災害対応を行う場合、宇土市にゴミの処理を依頼する。しかし宇土市が自身の管轄区域内のゴミ処理で一杯の状態になり、熊本市の要請を受け入れなかった可能性がある。推測ですが、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ったのも、今後対応できるようにする」と考えていると推察され、被災後に互いに考慮していることが食い違っている可能性がある。

6. おわりに

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。災害の規模や社会的な関心度の視点から自治体を選び、抽出されたキーワードがどう相違するかを比較した。その結果、地域防災計画の特徴を把握することができ、自治体の被災経験が地域防災計画の改定に影響を与えていることがわかった。さらに地域防災計画を比較することで広域連携がうまくできていない可能性が示された。今後は自治体の地理的な位置を考慮し、これらの知見を踏まえ自動的に「問題」を発見するアプローチを模索していきたい。

参考文献

[1]「東日本大震災における災害廃棄物処理概要報告書」
 [2]総務省：<http://www.soumu.go.jp/gapei/gapei2.html>
 [3] S.E. Robertson and S.walker, "Some Simple Effective Approximations to the 2-Poisson Model for Probabilistic Weighted Retrieval," Proc. 17th Annual International ACM SIGIR Conference(SIGIR'94), 1994