

# 名張市住民意識調査アンケートの住み心地を対象とした テキストマイニング

赤崎 智<sup>†</sup> 塩崎 勇真<sup>†</sup> 宇田 隆幸<sup>†</sup>

<sup>†</sup>近畿大学工業高等専門学校 総合システム工学科

## 1. はじめに

2000年頃から、地方自治体による意識調査が増加している。調査目的は、住民意識の調査を行い、現状を把握し、住み良い街作りを計画、実現するためである。

本校所在地の名張市も、2005年から住民意識調査を実施してきた。同調査は、17地区・性別および年齢層を1つのセグメントにして、ランダム抽出した回答者を各セグメントに同人数配分している。また、調査項目は、経年による変化がほとんど見られない。

本研究は、名張市の住民意識調査の分析を行ない「住み心地に肯定する意見の抽出」を目的とする。

本稿では、名張市住民意識調査アンケートのテキスト分析と定量分析を行ない、住み心地の肯定/否定因子を発見する。

## 2. 住民意識調査の概要

名張市では変化のない質問項目を配したアンケート調査が8年間継続しているにもかかわらず、経年による意識変化に着目した分析が行われていない。また、自由回答欄は分析対象から外れている。

そこで、経年による意識変化傾向を把握し、住民の住みやすさ向上に結び付く因子を分析する。特に、自然言語処理を用いて、被験者の自由回答欄を対象としたテキストマイニングを行ない、被験者の定量回答との突合を行う。本研究では、特徴語検出を行い、住みやすさ向上因子とその逆因子をつきとめる。

## 3. 分析

### 3.1. 分析対象

2005年から2013年にかけて名張市が実施した住民意識調査のローデータを研究対象にする。このデータは ①居住地を17地区 ②年齢層 ③性別に各々分割し、クラスタを生成。ランダム抽出した被験者を各クラスタに等配分している。各被験者は、予め市の用意したアンケートに回答する。アンケートは、21の質問からなる。質問の多くは、選択式回答を求めるもので(図1)、回答の肢は、ステップブスの尺度基準における名義尺度と順序尺度である。他に、日本語による意見表明もできる質問もある(図2)。

図1 あなたは、名張市の住み心地についてどう思われますか。該当する番号を1つ選んで回答用紙に記入してください。

1 住みよい
2 どちらかといえば住みよい
3 どちらかといえば住みにくい
4 住みにくい

図1. 選択式回答欄の一部

(自由意見欄)
○ 姓名 (このほかのことについてはです)
行政について
○ 具体的な内容
1. 名張市内には住み良い人が多くいる。町代表だけが集めた会議は形骸化しているように思えてならない。会議に出ても発言しないで座っているだけでは役に立たない。だから学識経験者様を多くして有識な人々と広く協議すべきだと思ふ。

図2. 自由意見欄の回答例

## 3.2. 分析方法

非構造化データである自由回答欄のデータに対し、データクレンジングを行なう。まず、自由回答欄のデータのテキストマイニングを行うために回答の誤記の修正、などの前処理を行う。

その後、テキストマイニングを行う。この時、後に定量回答部と関連付けることを考慮し、特徴語のカテゴリ化、不要語などの除去を行う。

分析の精度を向上させるために以上のステップを複数回行い、状況を観察、改めてデータの解釈を行う。

## 3.3. 分析結果と考察

特徴語の肯定/否定を把握することができ、定量回答部との紐付けに成功した。また、肯定/否定に関するカテゴリ分類を行えた。出現した特徴語の相関として以下のパターンが際立った。

① 医療関連の語 - 否定的な語

② 交通・道路関連の語 - 否定的な語

たとえば、①は定量回答部の“地域医療機関での現在の医療体制に満足している市民の割合”という項目において2013年度まで満足が50%を下回っている場合、テキストマイニングで顕著に出現した特徴語も否定的な語であった。②も①と同様である。

ここで、特徴語の重複度合いを視覚的に表す“カテゴリWeb”を参照すると、①の医療関連の否定語を抽出した記述からは、②の交通関連の否定語も同時に検出される傾向が高いことが明らかになった。

## 5. おわりに

本研究は目的である、「住み心地に肯定する意見の抽出」に成功した。

## 参考文献

[1] SPSSによるテキストマイニング入門 - 内田 治