

F1-11

## テキストマイニングと共起ネットワーク図を用いた地域の印象把握に関する研究 ソーシャルメディアの投稿に着目して

### A Study of Image what People Get by Using Text Mining and Co-Occurrence Networks of Words Focus on Posts on Social Media Service

○細谷拓弥<sup>1</sup>, 根上彰生<sup>2</sup>, 赤澤加奈子<sup>2</sup>

Takumi Hosoya<sup>1</sup>, Akio Negami<sup>2</sup>, Kanako Akazawa<sup>2</sup>

Recently, intercity competition grows sharper. Then, cities start their own branding to make originally and attract citizens' attention. This study focused on posts on social media service, how people image about cities. By collecting the posts on social media service and analyzing them through text mining.

#### 1. 研究の背景

近年, 少子高齢化や産業構造の変化などにより, 人口減少, 地域の空洞化が進んでいる. そんな中, ご当地グルメやゆるキャラなど地域を売りにした商品が増えている. 特に 2006 年に初開催された B-1 グランプリは全国各地のご当地グルメを通じた町おこしイベントで毎年話題になっている. 2008 年には「ゆるキャラ」が新語・流行語大賞にノミネートされている.

「ブランド (brand)」という言葉の由来は「焼き印をつける」という言葉からだと言われている.

企業もつ「ブランド」は, 他社の商品やサービスと違う独自の価値観を表し, それを受け手である顧客が理解・納得することで企業のもつブランドは初めて成立する. (弘文堂 地ブランド)

米国マーケティング協会は, ブランドを「ある売り手あるいは売り手の集団の製品およびサービスを識別し, 競合他社のものと差別化することを意図した名称, 言葉, 記号, シンボル, デザインあるいは」その組み合わせ」と定義している.

表 1 地域ブランド視点の違い

	従来の地域ブランド論	新たに提起する地域ブランド論
ブランド化の対象	特産品・観光地	地域そのもの
顧客との関係	一過性	長期継続的
地域ブランドの評価	経済的・行動的指標	経済的・行動的+体験価値指標
地域ブランド・コンセプト・マーケティング	流行の追いかけ or 地域資産ベースによるコンセプト・マーケティング	地域資産と社会文化文脈のすり合わせによる体験価値ベースによるコンセプト・マーケティング
地域ブランドの単位	行政区単位	体験カテゴリー単位
地域ブランド・コミュニケーション	単発的・散発的プロモーション	コンセプト主導による統合的かつ段階的なコミュニケーション設計
地域ブランド・マネジメントの担い手	まとまりがない地域内の人や組織	地域内外の人や組織の協働
地域と企業の関わり	ブランド価値視点の欠如	企業ブランドと地域ブランドの共鳴, 関係づくり

#### ■地域ブランドの必要性

矢野経済研究所の調査によると, 地域のブランドが確立されることで, 地域経済や地域そのものの活性化がもたらされることが期待される. 特定の商品・サービスや観光地などのブランド化に終わらず, 地域そのものがブランド化されることで, 繰り返し『買いたい』『訪れたい』『住みたい』といった要素により持続的に地域が活性化していくことを目指すものとして捉えられている.

また, 地域ブランドの最終的な目的は特産品が売れ, 観光地に人が訪れるだけでなく, 地域に関わる人, 特に住んでいる人が愛着やアイデンティティを持てることであると述べている. 商品ブランディングと同様に, 売り手(=地域・自治体)と買い手(=来訪者・居住者)間で価値観のマッチングが成立することがブランディングの最大の効果であると考えられる.

#### 2. 研究の目的

ソーシャルネットワークという観点から地域が不特定多数の対象者からどのように思われているか, その実態を把握することを目的とする. 今回の研究の情報ソースにはツイッターを用いる. この理由として, 過去の研究ではブログの本文が使われていたが, 現在ではツイッターがブログの立ち位置に置き換わっているという点, また国内ユーザー数が多く, ユーザー同士のやりとりも投稿の一つとして扱われるため, 潜在的なデータが得られる点である. 本研究では対象の地域として武蔵小杉を事例にあげる. その理由として NPO 団体により民間によるまちづくりが行われていること, 近年開発により土地利用が以前と変わってきていることがあげられる.

#### 3. 研究方法について

1. Twitter (ツイッター)<sup>注2)</sup>上の「つぶやき」(投稿)を Twitter API<sup>注2)</sup>および Python<sup>注3)</sup>を利用し簡易的なプログラムを作成, それを利用しテキストデータを収集する.

2. 収集したデータから重複する投稿や人名など地名とは関係のない単語を含む投稿を除外する. 既往研究にない, 「〇〇駅」や「〇〇店」のような地域の名称を用いた単語を含む投稿は除外し, 宣伝や広告, ニュース記事の拡散 (バイラルメディア) 除外のためハッシュタグや URL を含む投稿も除外する.

3. 得られたテキストデータを用い, テキストマイニング<sup>注4)</sup>の手法で分析を行う. なお分析のツールには KH Coder<sup>注5)</sup>を用いるものとする.

なお, テキストマイニングのソースにツイッターの投稿を用いることについて, 大量のデータを扱うことができる点, 分析を自動化することで分析基準に揺らぎや恣意性が生じないという点がメリットとして挙げられる.

ツイッターというプラットフォームの特性上ユーザー分布に偏りがみられる. また通常のアンケート調査と異なり年齢や性別, 居住地など対象者のパーソナリティによる分析ができない点をふまえて考察を行う必要がある.

表 2 2種類の地域ブランド

	主体	対象	目的	効果
①モノの地域ブランド	生産者, 自治体	消費者, 顧客	モノの差別化, 優位性	モノの価値化
②地域全体の地域ブランド	住民	住民, 共感者	住民の満足	地域の価値化

## 3. 結果とまとめ

表 3.頻出語 (武蔵小杉)

No	抽出語	品詞	No	抽出語	品詞
1	行く	動詞	16	帰る	動詞
2	思う	動詞	17	乗り換える	動詞
3	横浜	地名	18	言う	動詞
4	乗る	動詞	19	出る	動詞
5	今日	副詞可能	20	時間	副詞可能
6	川崎	組織名	21	見る	動詞
7	今	副詞可能	22	東横線	固有名词
8	住む	動詞	23	遠い	形容詞
9	笑	人名	24	着く	動詞
10	人	一般名詞	25	降りる	動詞
11	グランツリー	固有名词	26	新宿	地名
12	来る	動詞	27	向かう	動詞
13	電車	名詞	28	街	一般名詞
14	渋谷	地名	29	東京	地名
15	乗り換え	名詞	30	歩く	動詞

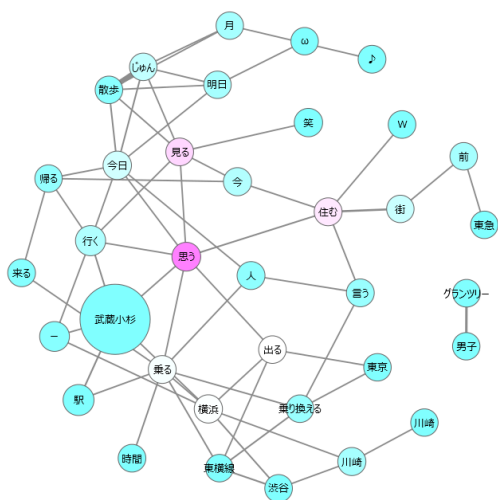
キーワード：武蔵小杉

集計期間：2017/6/21 から 2017/7/5

対象数：6090 件

表 3 は「武蔵小杉」というキーワードに対して、それに関連して頻出度の高い単語を一覧にしたものである。他のキーワードで得られた結果と比較し、駅の利用に関する投稿が多く、乗換駅として多く認識されていることがわかる。

図 4.共起ネットワーク (武蔵小杉)



さらに図 4 は共起ネットワーク<sup>注6)</sup>を表したもので、「思う—住む—街」が線で結ばれている点の特徴としてあげられる。

これらよりわかることは、「武蔵小杉」という場所には「行く」意識が強いということ。また「住む」「思う」が共起ネットワークで結ばれていることから住みたいという投稿が多く見られることが明らかとなった。

## 5. 参考文献および注釈

[参考]

- (1) 和田ほか (2009) 『地域ブランド・マネジメント』
- (2) 藤崎亮一「地域づくり分野におけるブランド活用への一考察」

[注釈]

注 1) ソーシャルメディアの定義

「ブログ、ソーシャル・ネットワーキングサービス (SNS)、動画共有サイトなど、利用者が情報を発信し、互いの関係を視覚的に把握できるのが特徴。」(総務省)とされている。

注 2) Twitter

国内アクティブユーザーは 4000 万人にのぼり、最大規模のプラットフォームとなっている。Twitter API ではツイッターの機能を外部プログラムに取り込むことが可能になる。

注 3) Python

プログラム言語のひとつ。作成したプログラムには sixsix 氏によるオープンソースコードである

「Python-Twitter-Tools」を用いている。これは GitHub において MIT ライセンスで公開されており、自由に利用することができる。

注 4) テキストマイニング

吉田らはブログをソースとしたテキストマイニングによる研究を行っている。Twitter がミニブログとして登場した経緯などからブログに代わるソースとして Twitter は有用であると考えられる。

注 5) KH Coder

樋口耕一氏によるフリーソフトウェア。学術での利用は許可されており論文等でも広く扱われている。

注 6) 共起ネットワーク図

線で結ばれた語同士の繋がりが強いことを示す。出現率が高いだけでは図中に表れないため、相互の関係性が明らかになる