

E-005

ブログを対象とした地域住民の政治的関心の分析 Extraction of Opinion for Local Government from Blog

木村 泰知[†]
Yasutomo KIMURA

渋谷 英潔^{†*}
Hideyuki SHIBUKI

高丸 圭一^{††*}
Keiichi TAKAMARU

1. まえがき

国政と比較して地方政治の活動は、国民の認知度が低い傾向にあるが、その重要性が劣るものではない。また、地方政治への関心は個々の住民によって多様に異なっており、それらに対応する政治情報も多様に存在する。それゆえ、一人ひとりの住民の政治的関心に合わせて政治情報が提示されることが望ましいが、住民が自分の政治的関心を明確に認識しているとは限らず、そういった関心に合う情報を探すことは容易ではない。したがって、住民の潜在的な関心を明確化して、それぞれの住民にマッチした政治情報を提示するシステムの構築が望まれている。このような背景から、我々は住民本位型政治情報システムの開発を目指している [1, 2]。

我々が目指す住民本位型政治情報システムの全体構成を図1に示す。文献 [1] では、地方議会の議事録における発言から、その議員の活動を示すフレーズを抽出することを行った。文献 [2] では、議員活動のフレーズを用いて、ブログに潜在する政治的要望を抽出しマッチングさせることを試みた。しかしながら、ブログ記事に存在する政治的関心が必ずしも議員活動との関連性が強いとは限らず、マッチング処理の妨げとなっていた。

ブログなどの Web 上の情報を集約して提示する従来研究には [3, 4, 5] などが存在する。しかしながら、これらの研究は議事録とブログのように異なる種類の知識源から得られた情報をマッチングさせることを目的としたものではない。マッチングという目的においては、その分野のオントロジーを参照しながら抽出を行った方が効果的であると考えられる。したがって、政治を対象としたオントロジーが必要である。

オントロジーを構築するために wikipedia¹ や現代政治用語辞典² のカテゴリなどを利用することも考えられるが、これらのカテゴリは政治用語や制度といった観点からまとめられたものであり、地方政治における住民の要望という観点からまとめられたものではない。住民の要望という観点から体系付けられたオントロジーは、実際のブログ記事などを分析して構築する必要があると思われる。それゆえ、本稿では、地域住民の政治的要望に関するオントロジーの構築を目的としてブログ記事の分析を行う。

2. 対象とするブログ記事

我々の研究は、特定の地方政治に限定したものではないが、これまで著者の地元である北海道小樽市を対象に

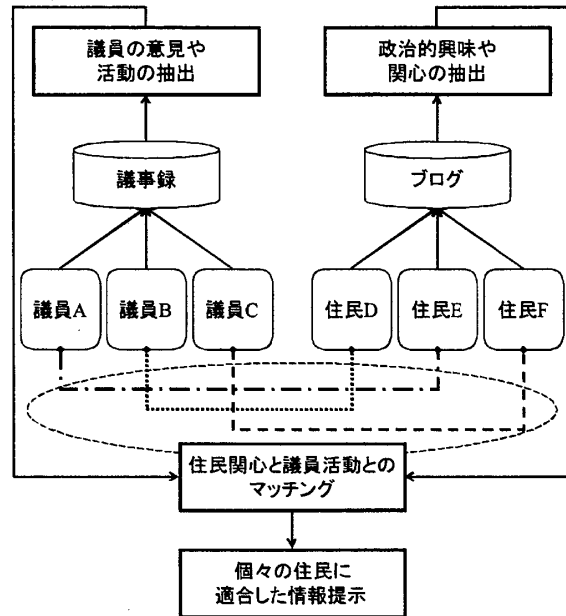


図1: 全体の流れ

行われてきた。文献 [2] において、小樽を対象としたブログ記事を収集しており、これを利用して分析を行う。本稿で調査対象とするブログは以下のように選定された。

理想的には小樽市住民のブログが望ましいが、ブロガーの住んでいる地域の情報はブログのプロフィールにも記載されていないことが多い。それゆえ、「小樽」という単語が含まれる記事を書いているブロガーを小樽市の住民と仮定して、ブログ記事の収集を行った。2008年1月の時点で、1年3か月分の「小樽」を含むブログ記事を収集しており、記事数は6万件以上である。

次に、収集した記事の信頼性を高めるために以下の処理を行った。収集した記事をドメイン単位で識別し、「小樽」を含む記事が3件以上含まれるブログを対象とした。また、投稿された記事には、スパム記事が多いため、同一内容や類似性の高い内容を複数回投稿しているドメインのブログは、スパムブログとして除外した。

最後に、政治的関心が含まれている可能性を高めるために、小樽市の問題に関連する語句を含むブログ記事を以下のように抽出した。「小樽 問題」を検索語として google で検索し、上位100件を手で調査した結果、表1に示す問題が存在した。したがって、表1のフレーズに含まれる「温泉」や「入浴」などの単語を含むブログ記事を最終的な分析対象とした。今回、分析したブログ記事は1,800件であった。

[†]小樽商科大学, Otaru University of Commerce

[†]横浜国立大学, Yokohama National University

^{††}宇都宮共和大学, Utsunomiya Kyowa University

* 合同会社ディクティオ, Dictio, LLC.

¹ ウィキペディア, <http://ja.wikipedia.org/>

² 現代政治用語辞典, <http://pol.cs4.jp/>

表 2: ブログから得られたカテゴリーの例

大分類	小分類	ブログの例
インフラ	道路	札幌と小樽を結んでいる国道で交通量が多い
	フェリー	フェリー高い
	運河	「小樽運河戦争」結論から言いますと「運河保存派」は敗北しています。
	除雪	除雪
社会福祉	医療	小樽市立病院の経営危険度 全国4位 看護助手だけどやっていることはほとんど介護
	公共施設	おたるドリームビーチの清掃活動 小樽市総合博物館
	交通安全	FM おたるでは、交通安全を呼びかけるCMを作っており
外交・防衛	テロ対策	テロ警戒 小樽署
	自動車	小樽港などに行くと中古車を大量に積み込んだロシアの船などが沢山いました
人口問題	過疎	小樽も益々人口が減少し、札幌のベットタウンになりきれないものがあり
金融	貯金	小樽の貯金センター
観光	都市間連携	プロスポーツ観戦し移住体験や人気の旭山動物園、小樽観光との運動など いままでにない北海道を体験出来る場を設けるべき
環境	公害	小樽運河の底に沈んでいた ヘドロ
	生態系	環境問題をテーマに、北海道の外来生物問題
都市開発	再開発	札幌寄りの小樽市の再開発にシフトして行く運命なのでしょうか

表 1: 小樽の問題

小樽温泉入浴拒否問題
日本銀行小樽支店問題
小樽市立病院問題
小樽の給食腹痛問題
運河問題
ラブホテル建設問題
いじめ問題
ニート問題
米軍艦が小樽港に入港
小樽の交通マナー
小樽市の偽装マンション
小樽駅前の再開発事業について
婦人科の問題 (少子化)
環境問題 (廃棄物の再利用)

3. カテゴリー調査

地域住民がもつ政治的関心のカテゴリーを以下のように調査した。まず、ブログ記事を調査し、政治的関心と関連がありそうな記述を抽出した。抽出された記述からカテゴリーラベルをボトムアップに付与していった。

付与されたカテゴリーとブログ中の記述の例を表2に示す。1,800件の記事から得られたカテゴリーは、大分類のレベルで8個、小分類で15個であり、当初期待していたほどのカテゴリー数が得られなかった。このことは、住民が個人のブログにおいて政治的関心があまり明示的に記述されていないことを示しており、マッチングの共通基盤としてのオントロジーの重要性が高いと共に、別のアプローチからのオントロジーの構築を考慮する必要があると考えられる。

4. 今後の予定

今後は小樽のブログ記事以外を対象としてさらに調査を進めていきたいと考えている。また、調査結果を元に政治的関心のオントロジーを構築し、マッチングに利用する予定である。

謝辞

本研究の一部は総務省 SCOPE 補助金 (No.082301004) の支援により行われた。

参考文献

- [1] 渋谷英潔, 木村泰知, 山崎記敬: 議員発言録からの重要単語抽出システムの提案, FIT2007 情報科学技術フォーラム講演論文集, (2007).
- [2] 木村泰知, 渋谷英潔: ブログに潜在する政治的意見と議員活動とのマッチング手法, 電子情報通信学会 NLC 研究会「主観表現処理の最前線」シンポジウム, pp.19-23, (2008).
- [3] 藤井敦: OpinionReader: 意思決定支援を目的とした主観情報の集約・可視化システム, 電子情報通信学会論文誌 D, Vol.J91-D, No.2, pp.459-470, (2008).
- [4] 佐藤翔平, 関和広, 上原邦昭: ブログを対象とした統計的意見情報検索, 言語処理学会第14回年次大会論文集, pp.360-363, (2008).
- [5] 井上結衣, 藤井敦: Web 世論からの意見抽出と賛否に基づく分類, 言語処理学会第14回年次大会論文集, pp.364-367, (2008).