

評価情報タグ付きコーパスの分析

— トピックを考慮した評価情報の抽出に関する考察 —

川田 拓也* 中川 哲治* 乾 健太郎*† 黒橋 禎夫*‡

* (独) 情報通信研究機構, † 奈良先端科学技術大学院大学, ‡ 京都大学

1 はじめに

Web 上に存在する膨大な量の文書から人々の意見や評判情報を抽出し、提示する手法について、近年議論が盛んになってきている。我々はその中で『Web 情報信頼性分析システム WISDOM』[3]における評価情報の分析エンジンの開発、研究を行っており、教師用および評価用データとして現状で 10,000 文 (50 トピック) からなる評価情報タグの付与されたコーパスを構築してきた [1, 5]。

WISDOM で行われる評価情報分析はユーザーの入力したトピックに関する Web 上の様々な評価を俯瞰的に提示することを目的とする。そのためにトピックに対する肯定／否定評価の分布を提示する必要がある、トピックそれ自体が肯定もしくは否定されていると解釈できる情報が抽出されなければならない。しかしコーパスの構築を通じてその抽出条件の決定には困難が伴うことが明らかになった。トピックに関して我々は、商品など利用物に限らず、賛否両論ある社会的制度など幅広いドメインを想定している。さらに、一つのトピックに関する評価情報は多種多様で、様々な観点から述べられている。中には以下で例示するように、トピックに対する評価には必ずしも結びつけられない情報やその判断自体が難しい場合もある。なお、例文中の**太字**はトピックを表す。また、下線は文中で直接評価の対象となる表現（後述する記述評価の評価対象）の範囲を指し、**ゴシック**はその対象の評価表現を表す。

- (1) a. **iPod** はとても便利だ
- b. 山田は **iPod** の音質 が気に入っている
- c. **iPod** を紹介しきれていない 雑誌 **X** はダメだ

例えば、(1a) はトピックである iPod が直接肯定的に評価されている。一方、(1b) は iPod の属性が評価されており、iPod 自体への評価は間接的である。ただし、この場合は、iPod それ自体に対する肯定評価だと解釈できる。しかし、(1c) のような例になると、直接評価されている対象は雑誌であって、iPod が評価されていると考えるのは難しい。このように直接評価されている対象がトピックに対する評価とは無関係な場合がある。以上の問題は、文中に現れる直接的な評価対象とトピックとの関係性に起因すると考えられる。これまでの評価情報コーパスは評価情報の抽出を主とするタスク設計にとどまり、トピック自体に対する肯定／否定評価に結びつく情報を

抽出する際に課題となる評価対象とトピックとの関係性については深く議論されてこなかった。本稿では、これまで構築してきたコーパスを通じて、評価対象とトピックとの関係性の分析を行い、トピック自体に対する評価につながる評価対象および、つながりづらい評価対象の整理をする。また、その判断自体が難しくなる要因についても考察する。

2 関連研究

これまでの評価情報タグ付きコーパスは Web 上の商品レビューやブログ、新聞記事を対象に評価の保持者や、評価対象、および肯定／否定を表す評価表現といった、評価に関わる言語的な情報をテキスト中から抽出して構築されている。例えば、レビュー記事やブログ記事から商品に関する評価情報をその構成要素ごとに抽出するタスクとしては、小林ら [2] や、宮崎ら [4] などがある。新聞記事の中の意見情報に対してタグ付けされたコーパスには、MPQA コーパス [8] や NTCIR-6 における意見コーパス [6] などが挙げられる。その他に、意見情報からその情報のトピックを推定、抽出することを目的とした Stoyanov ら [7] もある。これまでの研究に対する我々のタスクの特徴の一つとして、比較的広いドメインのトピックを対象にそれぞれのトピックと関連する評価情報を幅広く抽出し、そのトピックの評価につながる情報を集約することが挙げられる。

我々の課題設定に対する従来研究の課題は二点に分けられる。第一点目として評価対象とトピックとの関係性のあり方に関する議論の必要性が挙げられる。従来研究 [2, 4] では、「iPod の音質」のようなトピックの属性やトピックとの比較対象の抽出は考慮されている。一方で、前節で挙げた (1b) のようにトピックが間接的に評価されている評価対象だけではなく、(1c) のようにトピックの評価には結びつかない評価対象もある。どのような評価対象がトピックの評価へとつながりやすいのか、現状では議論の余地がある。二点目は、利用物に限定しない多様なトピックを扱ったコーパスの構築に関する議論が十分されていないことが挙げられる。既存のコーパスの中で、新聞記事の意見を抽出する MPQA ではトピックに制限はないため抽出される評価対象は多様だがトピックのタグ付けには至っていないのが現状のようである [7]。

3 評価情報タグ付きコーパス

3.1 評価情報タグ付きコーパスの概要

これまで我々が構築してきたコーパスの基本的な仕様は川田ら [1] と変わらない。選定したトピックは 50 あり、大別すると「利用物」と「制度」に分けられる。前者

* {tkawada, tnaka}@nict.go.jp

† inui@is.naist.jp

‡ kuro@i.kyoto-u.ac.jp

は商品やサービスを含む人間が利用する物で、有形無形問わない。「電子マネー」や「ステロイド剤」「バイオエタノール」など、新規性もしくは話題性があり、現在も評価が定まっていないものが選ばれている。後者は賛否両論ある社会的制度、慣習が含まれる。例えば「サマータイム制」や「裁判員制度」^{*}、「クールビズ」などが選定された。

タグ付けに際しては検索エンジンを利用して、トピックを含む文とその直後の文を収集し、各文をタグ付与対象範囲とした。従って、タグ付与対象範囲はトピックを含むとは限らない。作業者はタグ付け対象範囲内において評価が表明されている箇所をすべて抽出し、タグ付与する。作業者は前後2文を文脈理解のために参照できる。一つのトピックにつき200文が選定されたので、10,000文がタグ付与対象となった。

3.2 評価情報のモデル

我々は評価情報を二つのレベルで捉えている。一つはテキスト中から直接抽出された「評価保持者」^{*}、「評価対象」^{*}、「評価表現」^{*}からなる情報で「記述評価」と呼んでいる。例えば(1b)「山田は iPod の音質が気に入っている」であれば、「山田(評価保持者)は iPod の音質(評価対象)について気に入っていること(肯定的評価表現)」が直接的に表現されている記述評価と考える。もう一つ上位のレベルとして、トピック自体に対する評価情報である「トピック評価」を考えている。記述評価がトピック自体に対する評価と解釈できる時、その記述評価はトピック評価を導ける関係にあるとする。その関係は記述評価対象とトピックの意味的な関係に依存する。例えば、(1b)は iPod の属性が評価対象だが、iPod 自体に対する肯定評価と考えられるのでトピック「iPod」に対する肯定的なトピック評価とみなせる。ただし、現状のコーパスではトピック評価は明示的にタグ付与されておらず、「トピックとの関連度」として主観的に0-2の値を付与している。関連度2は明らかにトピック評価が導けると解釈される記述評価に付与され、関連度0は(2)のような明らかにトピックとは無関係な記述評価に与えられる。このような例はトピック評価として認められない。

(2) 「裁判員制度」で検索すると、概要がわかる 時代になったので本当に便利ですね^{†2}

(3a) や (3b) のようにトピックが間接的に評価されていると解釈できる例については直感的に関連度1が付与されていたが関連度1の範囲を決めることは困難であった。

(3) a. バイオエタノール車は意外と快適だ

b. 裁判員制度を運用する 法務省が悪いと思う

本稿の課題は、コーパス仕様の面から言えば、これまで直感的に関連度1が付与されていた例を改めて整理することでもある。

4 評価対象とトピックとの関係性の分析

4.1 評価対象のパターン

記述評価がトピック評価を導けるか判断する上で、記述評価における評価対象とトピックの意味的な関係性が大きな影響を与える。我々はトピック評価を導ける関係

とそうでない関係を整理するために、トピックとの意味的關係に応じて評価対象をいくつかのパターンに分類した。その結果、関係性の種類によっては本質的にトピック評価に結びつきづらい関係があることが見出せた。しかし、分類を進める中で、トピックとの関係性だけでは判断が困難なケースもあり得ることも浮かび上がってきた。

まず、我々が扱うトピックの性質を整理しておく。我々が扱うトピックには利用物／サービスおよび、社会的制度／慣習が含まれる。これらには広い意味でのオントロジカルな関係にある諸概念が認められる。例えばそれぞれのトピックには属性や下位概念、上位概念もしくは対立概念などがあり得、それら諸概念は記述評価の評価対象となることがある。さらに、これらのトピックには利用(適用)もしくは提供(運用)ができるという共通点をもつ。例えば、利用物は当然利用することができ、かつ、他者に提供することもできる。同時に、利用者と提供者双方が存在する。また、制度は施行(運用)する側と制度が適用される側がある。前者は提供側で、後者は(意志の有無は問わず)制度の利用側とみなすことができる。すなわち、利用者や提供者など、トピックに関する行為を介した関係が認められる評価対象もあり得る。

我々は評価対象とトピックとの関係性について、(a) オントロジカルな関係、(b) トピックに関する行為を介した関係の2点に着目して評価対象を分類した。以下で(a)、(b)の下位範疇も含めて説明する。

(a-1) 同一関係: トピックそれ自体やトピックの同義表現や省略表現が評価対象であればトピック評価と記述評価は一致するのでトピック評価との関連については問題にはならない。

(a-2) 限定的関係: トピックの属性や下位概念など、トピックが限定的に評価されている場合がある。

● 属性

(4) a. iPod の液晶が好きだ(部分)

b. 合成洗剤のにおいが嫌だ(性質)

● 下位概念

(5) ICOCA がお気に入りだ(電子マネーのインスタンス)

これらは通常、トピックに対する評価と同一視できる。ただし、下位概念の場合、4.2.1節で後述する「選択的文脈」において、トピック評価を導くことが困難になる。

(a-3) 上位関係: 上位関係には、例えば(6)は「ゆとり教育」が話題となる文脈で「(日本の)教育制度」が直接的に評価されているような場合が相当する。このようにトピックについては言及されていない場合、トピック評価に結びつけて考えることは難しい。

(6) 確かに、教育制度に問題はあります[†]

(a-4) 対立関係: 評価対象がトピックと対立関係にある場合で、(7a)は直接評価されているのは「フッ素」であり、「キシリトール」が悪いと判断されているわけではない。

(7) a. キシリトールより フッ素の方が歯によい

上位関係や対立関係は抽出される記述評価の評価対象はトピックと一致しない。そのため明示的にトピックが評価されていない限り、トピック評価に結びつけるのは難しいと考えられる。

^{*}1 他に評価表現の意味的なタイプで分類した「評価タイプ」タグもあるが、本稿では割愛する。詳細は中川ら[5]、黒橋[3]を参照。

^{*}2 例文中の[†]は実例を引用したことを表す。

記述評価対象の分類		記述評価情報 (太字はトピック、下線は評価対象、斜体は評価表現を抜す)	トピックとの関連性
オントロジカルな関係	同一	a. 電子マネー は素晴らしい b. <u>USJ</u> は楽しかった(ユニバーサル・スタジオ・ジャパン)	・記述評価⇒トピック評価 (記述評価がトピック評価と結びつきやすい)
	属性	a. iPod の液晶画面はきれいだ b. 喫茶店X の調度品がお気に入りだ	
	下位概念	a. バイオエタノール車 は素晴らしい b. <u>ICOCA</u> より SUICA が好きだ(電子マネー) c. <u>明治の特許制度</u> の方が良かった	・選択的文脈による曖昧性がある。 ・(b),(c)はトピック評価としづらい
	上位	a. <u>日本の教育制度</u> はダメだ(ゆとり教育)	・記述評価の対象はトピック自体にならない
	対立	a. ゆとりより詰め込み教育の方が良い(ゆとり教育) b. <u>虫歯に効くのはフッ素</u> です(キシリトール)	・明示的にトピックが評価されていない限りトピック評価とはしづらい
トピックに関する行為を介した関係	主体	a. <u>ファーストフード</u> を愛好する人は好きになれない b. <u>〇×オークション</u> 内の△□商店は不親切だ	・責任関係による曖昧性 ・(b)はトピック評価としづらい
	行為	a. キシリトール を摂取することは良いことだ b. <u>ゆとり教育</u> を受けたことは残念だ c. アガリクス は生食より煎じて飲む方が良い	・記述評価⇒トピック評価(a),(b) ・選択的文脈による曖昧性有り(c)
	主体	a. 製品X の販売業者は少し怪しい b. 山田は <u>年金制度</u> を運用する厚労省を批判している	・責任関係による曖昧性？ ・トピック評価に導けない例は未確認
	行為	a. <u>ジェネリック医薬品</u> を取り扱うことは素晴らしい b. <u>全国学力テスト</u> を実施することは税金の無駄 c. バイオエタノール は海草から生産すれば環境への負荷が少ない	・記述評価⇒トピック評価(a),(b) ・選択的文脈による曖昧性有り(c)
	情報源	a. <u>裁判員制度</u> を正しく伝えないマスコミに責任がある b. バイオエタノール が環境に優しいと主張する山田の新刊は全くおもしろくなかった	・トピック評価につながりづらい(a,b) ・他者の評価の内容に踏み込んだ評価がされている場合は曖昧(c,d)
	他者評価	c. バイオエタノール が環境に優しいという常識は怪しい d. バイオエタノール が環境に優しいと主張する山田の新刊には賛成できない	
	その他	a. ブラジルは バイオエタノール にかかる関税を問題視している b. 金融庁は 電子マネー 法制定を前向きに検討している	

図1 評価対象とトピックとの関係性

トピックと評価対象の間には先述の通り、トピックを利用する主体と、トピックを提供する主体が存在する。また、利用したり、提供したりする行為自体が評価対象となる場合がある。例示すると以下ようになる。

(b-1) 利用関係:

主体: トピックを利用する主体が評価対象

- (8) a. バイオエタノールの利用を進める 企業 X は偉い
b. 〇×オークションに出店している △□商店 は不親切だ (〇×オークションを利用する主体)

行為: トピックを利用する行為自体が評価対象

- (9) a. キシリトールを摂取すると虫歯に効く
b. ゆとり教育を受けることは苦痛だ

(b-2) 提供関係:

主体: トピックを提供する主体が評価対象

- (10) アガリクスの販売業者は少し信用できない (アガリクスを提供する主体)

行為: トピックを提供する行為自体が評価対象

- (11) 学力テストを実施することに賛成だ (学力テストの提供／運用行為)

利用／提供関係が認められる評価対象は通常トピック評価に結びつく関係として解釈されやすいが (8b) はトピックである「〇×オークション」に対する評価には結びつきづらい。この点については4.2.2節で後述する。

(b-3) 情報伝達: トピックそれ自体が情報もしくは情報伝達行為の中で扱われる場合で、前者はトピックを紹介するような情報源となる媒体 (出版物／マスコミ) などが評価対象になる場合 (12)、後者は他者の評価に対する評価など (13) が当てはまる。

- (12) 裁判員制度をきちんと説明しない マスコミ に責任がある

- (13) a. バイオエタノールは環境に優しいと主張する

山田の新刊は全くおもしろくなかった

- b. バイオエタノールは？環境に優しいと主張する山田の新刊には全く賛成できない

(12) や (13a) のように情報源そのものが評価されている場合はトピック評価への結びつきは認められない、だが、(13b) のように、情報源の意見内容に踏み込んだ評価がされていると解釈できる場合は、トピック評価に結びつけられる可能性がある、読み手の解釈に依存する部分も多いと思われるため、他者の主張からはトピック評価が導かれるか判断が難しい。以下ではより判断するのに困難が伴う条件について詳述する。

4.2 トピック評価との関連を判定するのが困難な条件

4.2.1 選択的文脈

下位概念や利用／提供関係の場合、直接の評価対象以外の対象への評価が含意される文脈では、トピック評価認定が難しくなる。

- (14) a. SUICA に比べて ICOCA がお気に入りだ

- b. アガリクスは生食するより煎じるとよく効く

例えば、(14a) は他の電子マネーに比して ICOCA が選択的に評価されている。だからといって電子マネー全体が評価されると結論づけるには強すぎる。(14b) の場合、特定の方法を選択することが選択的に肯定されているが、アガリクス全体が肯定もしくは否定されているかどうかは決められない。このように直接的な評価対象が排他的もしくは選択的に評価されている文脈においてはトピック評価と同一視することが難しくなる。

4.2.2 責任関係

利用／提供関係には利用する側と提供する側との間に責任関係が多かれ少なかれ存在する。たとえば、(10) のように「アガリクス販売業者」はアガリクスに対する強い責任関係が認められる。そのとき、販売業者に対する

表1 評価対象分類結果

対象文 評価情報	ホワイトカラー 200 文 105	サマータイム 200 文 145	電子マネー 200 文 132	バイオエタノール 200 文 115
同一	62 (59%)	81 (56%)	26 (20%)	71 (60%)
属性	0 (0%)	0 (0%)	5 (4%)	0 (0%)
下位概念	11 (10%)	7 (5%)	55 (42%)	26 (22%)
上位・対立	1 (1%)	4 (3%)	2 (2%)	2 (2%)
提供関係	14 (13%)	27 (19%)	7 (5%)	6 (5%)
利用関係	4 (4%)	2 (1%)	5 (4%)	4 (3%)
情報伝達	4 (4%)	3 (2%)	1 (1%)	1 (1%)
無関係/その他	5 (5%)	9 (6%)	13 (10%)	4 (3%)
極性なし	4 (4%)	12 (8%)	18 (14%)	4 (3%)

評価は、アガリクスの評価と同一視しても大きな問題はない。一方で、場所提供者と店子の関係が認められるトピックがある。例えば、(8b)のようにネットオークションやイベント、施設などがトピックで、そこに店出している特定の店舗が評価対象となる場合、それがトピック評価と結びつくのかは判然としない。(10)と(8b)の差異は提供側の責任の程度差に起因するものと思われる。すなわち、場所提供者と店子の関係のように、提供側が利用側に負う責任の度合いが比較的強くない関係では提供側であるトピックの評価につながりにくいようである。このように提供側と利用側との間に強い責任関係が認められるような場合、記述評価とトピック評価の結びつきも強くなり、責任関係が希薄な場合や不明瞭な場合では、記述評価はトピック評価につながりづらくなると思われる。

以上の議論を図1にまとめた。一番左のカラムは評価対象のパターンを分類している。中央のカラムにおいてそれぞれの記述評価情報の例を提示した。一番左のカラムでトピックとの関連性が強いかどうか、そして、その判断が難しくなる要因は何かをまとめた。

4.3 分布

評価タイプのパターンの分布を調べるためにすでにタグ付けされているデータのうち、制度として「ホワイトカラーエグゼンプション」と「サマータイム制」を、利用物として「電子マネー」、「バイオエタノール」を選び、分類を行った。表1に示す。その結果、分類できない例(無関係/その他)は一割以下に収まり、大枠において、上記パターンで網羅できると思われる。トピックの種別による分布の違いも見られた。制度についてはトピックと同一関係の他、トピックとの提供関係が認められる評価対象が多く見られた。また、利用物では、トピックと同一関係の他、下位概念に偏りが見られた。「無関係/その他」と判断した例としてはデータの不備(不完全な文や文意が不明な例)の他に以下のような例が見られた。

(15) a. ブラジルは この(=バイオエタノールにかかる)関税を問題視しており…[†]

b. 金融庁は「電子マネー法」(仮称)の制定に向けた検討に入った[†]

なお、(15b)は金融庁によって「電子マネー法」が前向きに捉えられている肯定的な評価情報とみなされている。バイオエタノールにかかる関税や電子マネーに関連する法律がトピックとの提供(運用)関係とみるか、無関係とするかは判断が揺れると思われる。現状では厳密に対象

を分類しようとするのが困難が伴うことが予想される。

「極性なし」とは意見ではあるが、提言や助言、要望と解釈され、極性判定が難しい(16)のような例を指す。これらに関しては分類対象外とした。

(16) サマータイム制の導入による情報システムへの影響を事前に検討すべきである[†]

5 おわりに

与えられたトピックに対する肯定的評価および否定評価の分布を幅広い視点から提示することを視野に入れた評価情報抽出タスクに関してこれまで十分な議論がなかった。本稿は当該タスクの再考と位置づけられる。本稿で議論した内容は次の通りである。トピックを商品に限定せずサービスや制度など幅広く選定して考察対象とし、記述評価における評価対象とトピックとの関係性を整理した。その結果、評価対象とトピックとの関係にはいくつかのパターンが見出された。さらに記述評価とトピック評価の間に齟齬が生じる事例が明らかになった。例えば対立概念、上位概念および情報源が評価対象となる記述評価は基本的にはトピック評価につながらないことが観察された。評価対象のパターンとともに、文脈やトピックの性質も影響している。まず選択的文脈か否かはトピック評価とのつながりを判断する上で重要な要因になるであろうことを指摘した。また、トピックと評価対象との間に利用/提供関係が認められる場合において、提供側および利用側の間に介在する責任関係の強さも要因となり得る。

今後は本稿で行った議論をもとに、評価情報タグ付きコーパスの構築を進め、新たなトピックの追加、仕様の改良などを行っていきたい。

参考文献

- [1] 川田拓也, 中川哲治, 森井律子, 宮森恒, 赤峯享, 乾健太郎, 黒橋禎夫, 木俣豊: Web テキストにおける評価情報の整理・分類およびタグ付きコーパスの構築, 言語処理学会 第14回年次大会, pp. 524-527 (2008).
- [2] 小林のぞみ, 乾健太郎, 松本裕治: 意見情報の抽出/構造化のタスク仕様に関する考察, 情報処理学会研究報告 NL, pp. 111-118 (2006).
- [3] 黒橋禎夫: 情報の信頼性評価に関する基盤技術の研究開発, 人工知能学会誌, Vol. 23, No. 6, pp. 783-790 (2008).
- [4] 宮崎林太郎, 森辰則: 製品レビュー文に基づく評判情報コーパスの作成とその特徴の分析, 情報処理学会研究報告 2008-NL-187, Vol. 15, pp. 99-106 (2008).
- [5] 中川哲治, 川田拓也, 乾健太郎, 黒橋禎夫: Web 文書からの主観的・客観的評価表現の抽出, 第7回情報科学技術フォーラム (FIT2008), pp. 137-138 (2008).
- [6] 関洋平, Evans, D. K., Chen, H.-H., Ku, L.-W., 神門典子: 意見分析タスク - 多言語テキストを対象とした意見抽出技術の評価 -, 情報処理学会 研究報告, 2008-FI-89 (8), 2008-NL-183(8), pp. 51-58 (2008).
- [7] Stoyanov, V. and Cardie, C.: Annotating topics of opinions, *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco (2008).
- [8] Wiebe, J., Wilson, T. and Cardie, C.: Annotating expressions of opinions and emotions in language., *Language Resources and Evaluation*, Vol. 39, No. 2-3, pp. 165-210 (2005).