

否定的な意見が強い  肯定的な意見が強い

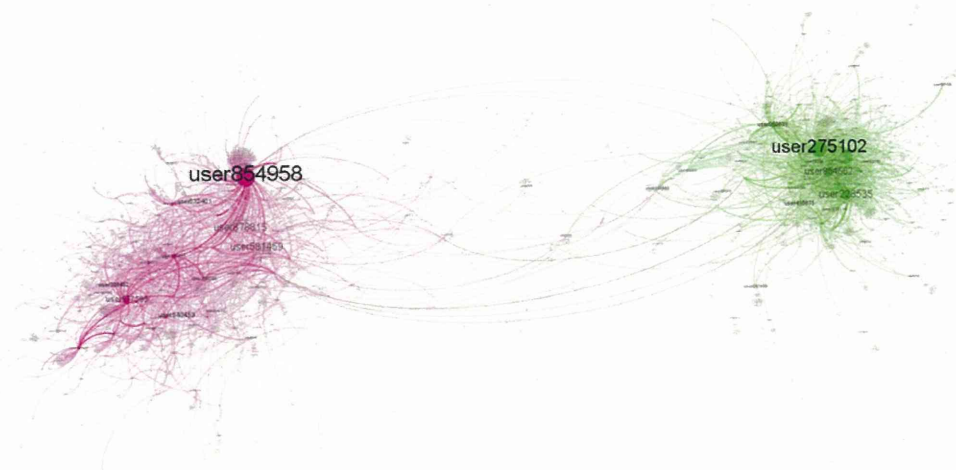


図 2 リツイート・ネットワーク（全期間）

らに、極性分析の結果を利用し、各ユーザー（ノード）の立場を色付きで表現した。

C. 研究結果

図 1 に「肯定的」「否定的」「どちらでもない」ツイート数の週ごとの推移を示した。毎年、桃の出荷時期となる夏に、ツイート数のピークが見られる。2011 年と 2012 年は肯定的と否定的なツイートの数がほぼ拮抗しており、2013 年は肯定的なツイートの方が優勢であった。2011 年夏の主なツイートは、農家の写真と共に桃の安全性をアピールするもの、桃農家に向けられた冷たい視線に対する悲しみ、福島の桃の当たり年、産地隠蔽の懸念、政治家の不適切発言、市民団体によるセシウム検出であった。2012 年夏の主なツイートは、除染の努力によるセシウムの不検出、輸出した福島産桃の完売のニュース、産地偽装の噂・疑い、風評対策プロジェクトへの批判などであった。2013 年夏のツイートは、「桃の涙」というリキュール、両陛下が福島の桃を召し上がった話題、ヨウ素検出という誤情報、風評対策プロジェクトへの批判などであった。

図 2 に可視化したリツイート・ネットワークを示す。ネットワークは左側と右側の 2 つの密なクラスタに分かれ、その間を少数のユーザーが繋ぐという構造をしている。しかも、左側のノードは赤色、右側のノードは緑色で彩色されていることから、左側のクラスタが否定的なグループ、右側のクラスタが肯定的なグループに分かれている。このことから、福島の桃に関する議論は肯定派と否定派に二分されており、各グループ内で密な情報交換が行われているが、否定派と肯定派の間で議論が交わされていないことが分かった。

D. 考察

福島の桃に関して否定的なツイートの中で、桃の購入に否定的な理由は以下に集約された。

- （暫定）基準値に対する不信感
- 0 ベクレルへの拘り
- 基準値以下だから安全というのではなく、検出結果の数値を提示してほしい
- 国・東電が補償すべきで、消費者がリスクを負うのはおかしいという主張
- 安易な風評対策に対する批判

否定的な意見が強い  肯定的な意見が強い

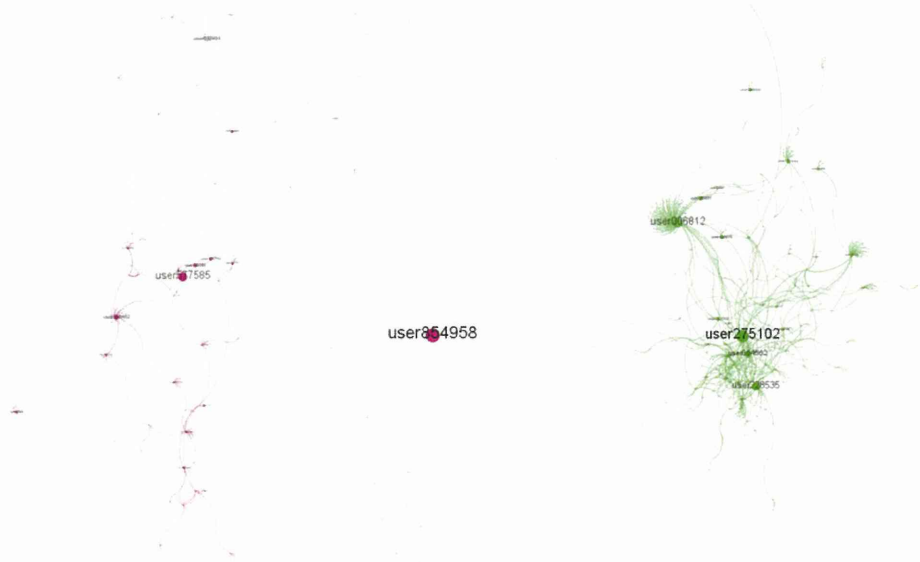


図 3 リツイート・ネットワーク（2013年夏のみ）

政府が定めた基準値への不信感が根強く、単に「基準値以下」とか「安全」という説明では納得できないという意見が多く寄せられていた。桃の購入に否定的な理由の部分でも述べたように、検査結果をもっと提供して欲しいという意見もあるので、食品のセシウム濃度の基準値の根拠の説明や、検査体制・検査結果の透明性のアピールがさらに必要であると感じた。

図 2 では肯定派と否定派がはっきりと分かれたユーザの空間が可視化されていたが、このネットワークを 2011 年 3 月以降で短い期間（例えば 1 か月）に区切って描画すると、福島の桃に関するクラスタの形成プロセスを動画として確認できる。その動画によると、肯定派・否定派のクラスタは 2011 年夏頃に形成されており、その後、クラスタ内の人々が自分の立ち位置を反対側に変えるような動きは見られなかった。これは、福島の桃に対する人々のスタンスは事故直後に固まってしまい、風評対策は後手に回ると困難を極めることを表す。

ただ、2013 年には特異な動きも見られた。図 3 に、2013 年夏のリツイート関係に絞って可視化したものを示した。まず、2013

年では否定側のクラスタの規模が縮小する傾向が見られた。これは図 1 のツイート数の分析結果に一致する。また、図 2 では否定派の中心的な役割を果たしていた **user854958** 氏が、図 3 では否定派から切り離されている。これは、**user854958** 氏が危険情報に反射的に飛びつく人々から距離を置こうとしたため、氏の立ち位置が変化した訳ではない。一方で、**user854958** 氏は自身が設定した基準値未満の食品は産地を問わず（福島県産であっても）食べることも述べている。このように、否定派の中でも福島県産の食品に対して多様な考え方があり、検査体制・検査結果に関するデータをもっと開示されれば、風評を払拭できる可能性がある。これは、風評対策を考えるうえで興味深い事例と言えよう。

この知見を踏まえつつ、震災発生直後の厚生労働省の公式ツイッターアカウント (@MHLWitter) の発言内容を以下の観点から振り返ってみた。

1. Twitter ツールの活用意義・意図は？
2. 利用者を理解し情報が伝わりやすいツール・設計になっているか？
3. 利用者のニーズを把握した上での情報

発信か？

4. 発言に対して根拠や詳細説明、具体策の提示を行っているか？

公式ツイッターアカウントのプロフィール欄には「厚生労働省では、ツイッターを通じて国民の皆様向けの情報を発信しています。ツイッターに寄せられたコメントへの返信はしていません。」との記述があり、厚生労働省と国民との双方向のコミュニケーションを想定していないことが伺える。しかしながら、国民からは図4(a)のようなレスポンスが得られており、Twitter というメディアを使って意見を述べようとした形跡がある。この声を拾う体制が整っていないのであれば、Twitter を用いた情報拡散の価値が大きく損なわれてしまう。実際、図4(b)のように、Twitter で広報活動を行うのであれば、PDFによる情報配信を避け、テキストコミュニケーションに最適な配信形式を選ぶべきだという意見が早い段階から寄せられていた。情報拡散にTwitter を活用する意義を十分に検討した上で、メディアの特性を活かした配信方法で情報公開を行うべきであろう。

また、図4(c)のように、国民の情報ニーズと広報内容の間にズレが生じているケースも見受けられた。この場合、国民の関心事（つまり安全面）についての情報発信がされていないことにより、国民からは一方的な発信に見え、不信感を与えている。さらに、図4(d)では具体策が示されていないため、配信した情報に対して多くの批判が寄せられている。これらのケースは、情報のニーズや開示する情報の詳細度により、情報やその発信者の信頼度が変化する事例と捉えることができる。災害という緊急時に的確な情報発信を行うことは難しいが、人々の意見や考え方は早い段階で凝り固まってしまうことを考えれば、日常的に情報を開示し、厚生労働省が配信する情報の信頼性を高めておくことは、風評を軽減する

ことに効果的なのではないかと考える。

E. 結論

本研究では、Twitter 上での福島県産の桃に対する立場と、それぞれの立場における議論を分析した。福島産の桃を肯定的に捉える立場と、否定的に捉える立場に分断されており、立場を途中で変えたり、立場を横断して健全な議論をした形跡が見られないことから、風評に対する対応の難しさが改めて浮き彫りになった。ただ、否定派の中にも、産地ではなく精密な情報を使って自ら判断したいという意見があり、検査結果に関する数値データを積極的に開示していくことで、両派の溝を解消する可能性を垣間見ることができた。さらに、震災直後に厚生労働省の公式ツイッターアカウントから発信された情報を調査し、SNSを情報拡散手段として用いる場合の課題を報告した。

F. 研究発表

1. 論文発表

- 鍋島啓太，渡邊研斗，水野淳太，岡崎直観，乾健太郎．訂正パターンに基づく誤情報の収集と拡散状況の分析．自然言語処理，Vol. 20，No. 3，pp.461-484，2013年6月．

2. 学会発表

- 岡崎直観，佐々木彬，乾健太郎，阿部博史，石田望．ツイッター分析に基づく福島県産桃に対する風評の実態解明とその対策．第26回日本リスク研究学会年次大会，B-5-3，2013年11月．

厚生労働省 (@MHLWitter)
200 RTs, 2 QTs, 2 replies, posted at 2011-03-22 16:27:43

避難生活の際に病気がかからないよう、また、できるだけ健康に過ごしていただくため、大切なことをまとめました。被災者を支援する方々も含め、これらのことを知っていただき、ご配慮いただくようお願いいたします。
[#jishin #311care](http://bit.ly/gU2Uhk)

Hide 2 QTs 2 replies 1 agrees
0 disagrees

@MHLWitter 確かに！確かになんだけどこれができたらかなり生活できはじめるわけ;;;こうできないから困ってる;;
@MHLWitter 本当に基本的な事で大切な事が多く掲載されています。それだけに、PDF形式は何かありませんか？携帯で読んでいる方もいらっしゃるかもしれません。手間だと思いますが、PDF形式とHTML形式があると良いのかも

(a) 国民からのフィードバックの例

厚生労働省 (@MHLWitter)
919 RTs, 6 QTs, 2 replies, posted at 2011-03-14 15:15:48

地震による被災をされた方で、インスリンを必要とする糖尿病患者の皆さまへ、インスリン入手のための相談連絡先を情報提供しています。
<http://bit.ly/gyySDV> (PDF) 日本糖尿病学会HP [#jishin #311care](http://www.jds.or.jp/#jishin)

Hide 6 QTs 2 replies 1 agrees
0 disagrees

@MHLWitter 文書の枚数が少ないのですからPDFではなくHTMLにしてくれませんか。そうすればコピー&ペーストで連絡先を拡散し易いと思います。 RT @MHLWitter 地震による被災をされた方で、インスリンを必要とする糖尿病患者の皆さまへ、……
@MHLWitter 大至急お願いします。

(b) 情報発信の形式に関する提案

厚生労働省 (@MHLWitter)
23 RTs, 0 QTs, 3 replies, posted at 2011-03-31 22:33:17

東北地方太平洋沖地震の被害状況および厚生労働省の対応について、第34報(3月31日現在)を掲載しました。(HTML版) <http://bit.ly/fX111s> (PDF版) [#jishin](http://bit.ly/gS3zgz)

Hide 0 QTs 3 replies
0 agrees 0 disagrees

@MHLWitter 牛肉の放射線規定値超の発表はなぜ深夜だったのですか？なぜツイートしないのですか？隠そうすると信頼をさらに失いますよ。植物からでなく動物からこれだけ検出されたのは問題ではないありませんか？直ちに健康に害はないで全てごまかさないで下さい。 @kokumin_koe
@MHLWitter 天栄村の牛肉についての再検査結果を、3/31付(食品第19報)に対応する訂正とはっきり分かるように、きちんとアナウンスして下さい。
@MHLWitter んら〜む。んたぶん、読

(c) 厚生労働省と国民の情報ニーズのずれ

厚生労働省 (@MHLWitter)
2146 RTs, 80 QTs, 44 replies, posted at 2011-03-14 00:07:47

厚生労働省です。ご自宅で、電気を使う医療機器(人工呼吸器など)を使っている患者の方は、停電の時間帯の対応の方法について、主治医にご相談ください。

Hide 80 QTs 44 replies 12 agrees
15 disagrees

それはないでしょう、夜になって急に発表しておいてこれはもう遅すぎませんか。主治医に連絡の取れない場合の対処方法もアナウンスしてください
無責任！
知り合いに該当者がいるかたは、すぐに教えておいて！主治医もわからないのではないのでしょうか……
相談したら主治医の病院も停電だからどうしようもないと大学病院にまわされましたけどこの時間では連絡窓口がない。今からじゃ遅すぎるっての！
あと数時間後の話でどうしろと緊急事態だし、何言っても始まりません。該当者は準備を！！
今から主治医に相談しても対応できないでしょうね。グループ1地域の方、急いで！

(d) 具体策や詳細説明の必要性

図 4 厚生労働省の公式ツイッターアカウントからの情報発信と国民の反応

厚生労働科学研究費補助金（食品の安全確保推進研究事業）
分担研究報告書

インターネットにおける意見解析手法の妥当性に関する研究

研究分担者 奥村 貴史

（国立保健医療科学院 研究情報支援研究センター 特命上席主任研究官）

研究協力者 榊 剛史

（東京大学 東京大学工学系研究科）

研究要旨

放射能汚染等への対応として、適切な情報提供を通じたリスクコミュニケーションが求められている。しかしながら、ネットが社会へと浸透した結果、公的機関や専門家に加えて一般人が独自に情報発信をする時代が到来し、従来行われてきたプレスリリースや地域集会を通じた情報提供では意図した情報伝達が行えない事態が生じている。それにも関わらず、既存の社会調査手法では、ネット上において生じている情報発信活動を把握することが出来ない。そこで本研究班では、ネット上の様々な情報の解析手法について自然言語処理を活用した検討を進めて来たが、ネット情報の解析から得られる情報については、代表性等の観点から妥当性に関する懸念が呈されることが少なくない。そこで、本研究分担は、社会調査により得られる意見とネット解析により得られる意見の比較を行い、インターネットにおける意見解析手法の妥当性に関して検討を行った。

まず、社会調査により得られる意見として、昨年度に実施したアンケート結果に含まれる放射能汚染に対する意見 1050 件を抽出した。これらに対して、2011 年 3 月～2012 年 12 月に投稿されたインターネット上のブログ記事を対象として、類似意見の検索を行った。その結果、社会調査により得られる意見について、インターネット上のブログにおいても類似する意見を大まかに確認することが出来た。このことは、インターネットにおいて表明されている意見が、社会調査により得られる意見と同等以上の多様性を有することを示しており、社会調査における情報源としてネット情報の定性的な解析が一定の妥当性を有しうることを示唆している。今後、類似意見検索技術の発展による、より大規模で精緻な検証が望まれる。

A. 研究目的

2011 年の東京電力福島第一原発事故後より、放射能に関する正誤のあいまいな情報が数多く生じている。「日本の食品中の放射性物質の基準値が欧米のものよりも遥かに高い」、「福島を中心に健康被害が続出している」、「原発作業員に多発している急性白血病が隠蔽されている」等は、代表的

なものであろう。狭い国土において生じた未知の災害に対して社会不安が生じること自体は避けがたいことではあるが、その結果、安全性が確認された農産物等に対しても忌避が進み、多大な社会的損失を引き起こしている。そこで、公的機関に対して、放射能に関する適切な情報提供を通じて無用な社会的損失を避け、被災地の復興を支援していくことが望まれている。

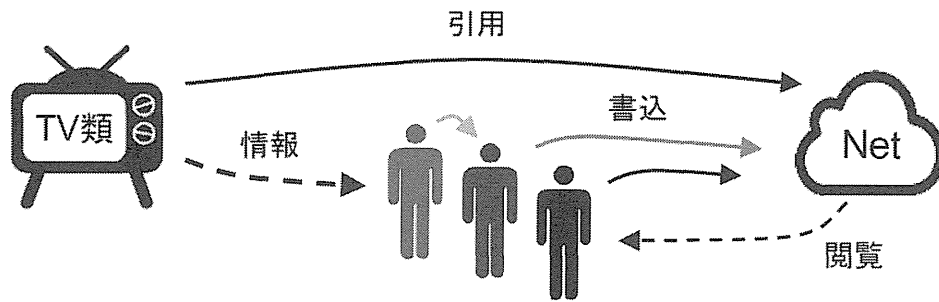


図 1. 意見の表出における既存メディアとネットの関係 (模式図)

このような「リスクコミュニケーション」は、今までにも、環境汚染による風評被害などの際に実践されてきた。しかしながら、従来のリスクコミュニケーションは、マスメディアを通じた一方通行な情報伝達か、専門家と一般人の顔の見える交流かのいずれかを前提としており、近年のようにネットを介して公的機関や専門家に加えて一般人までもが独自に情報発信をするような状況を想定していなかった。そのため、今後の公的機関による情報伝達においては、ネット上での反応を考慮せざるをえないが、ネットには膨大な情報が存在するため、社会心理学や統計学にのみに基づく従来の社会調査手法では状況の全体像を把握することすら困難となっている。

こうして、ネット時代のリスクコミュニケーションを検討していく上で、多量の文書を効果的に処理しうる「自然言語処理技術」が期待されることになる。その際、もっとも重要となるのは、ネット上に表出されている意見と、ネットを利用しない者も含めた社会全体の意見との乖離である。実際、ネット上の世論調査と実社会の世論調査に相違が生じることは繰り返し指摘されている。ネット上の意見は、自由に複製が可能であるうえ、PC やスマートフォンなどの情報機器を利用できる年齢層に偏っており、「いいね」や「リツイート」等の消極的な形で示される意見もある。このように多様な意見すべてを正確に把握することは

困難であるため、ネット上での意見について多いか少ないかという定量的な判断を行うことは容易ではない。

そこで、本研究分担は、社会調査により得られる意見とネット解析により得られる意見の比較を行い、インターネットにおける意見解析手法の妥当性についての検討を進めてきた。今年度は、昨年度の社会調査により収集した意見を元にして、より解析の精度向上に向けた研究を試みた。

B. 研究方法

インターネットに表明されている意見は、発信者の素性が特定できず、また、発信者が複数発言したり一つの意見が複製される可能性があることから、有効な定量的分析が難しい。一方で、テレビやラジオなどのマスメディアにおける言説は、現代では、直ちにネット上に複製され、賛否双方からのさまざまな意見が表明されることになる。また、ネットを利用していない者の意見も、伝聞情報としてネットに記載されうる(図 1)。したがって、意見の多様性に関しては、ネット上の意見は、社会調査により収集される意見を凌駕する可能性がある。そのため、低コストで大規模化するネットにおける意見の解析は、意見の網羅性や意見間の関係に関する定性的解析に関しては、社会調査以上に社会を代表しうる可能性がある。

Q27. 食品の放射能汚染について、お伺いします。

- a) 食品の安全性について、どのような情報を求めていますか。
- b) その情報は、国が提供すべきだと思いますか、思いませんか。どうしてそのように思うかもあわせて、ご意見をお聞かせ下さい。
- c) 国が、どのような情報を、どんな形で提供してくれればいいと思いますか。

Q28. 仮に食品の買い物の際欲しかった品物に「福島産」と書かれていたら、どうしますか。

Q30. 放射能は、あなたの暮らしにどのように影響を与える、または与えていると思いますか。自由に考えをお聞かせ下さい。

図 2. 放射能汚染に関する質問一覧

この仮説を検証するため、昨年度は、社会調査により放射能汚染に関する意見を収集したうえで、それぞれの意見がネット上に存在するかを確認し、ネット上の意見が社会を代表する網羅性を備えるかを検証した。具体的には、東日本大震災の影響が少なかった地域から大都市部と地方都市として東京、秋田、高知を選び、それぞれ 70 件のインタビュー調査を行った。インタビューでは、放射能汚染に関するデマ等に関するアンケートに加えて、放射能汚染に関する意見を自由回答して頂いた。社会調査による意見の抽出に際しては、賛成や反対などで回答しうる単純かつ構造的な調査も可能ではあるが、こうした調査は効率的である一方で、回答の自由度を制限することで独自性の高い意見を収集し損ねる懸念がある。

そこで、昨年度は、今後の調査を設計するための大まかな傾向の把握を目的として、あえてより多彩な意見を収集しうる自由回答を採用した。上述お通り、3 地点で各 70 名、合計 210 名を対象として、放射能汚染に関する 5 種の質問への回答を収集した。このアンケートに利用した質問票の抜粋を図 2 に示す。これらの質問について自由な意見を求めることで、210 名からの各 5 件、合計 1050 件の意見が集まった。昨年度は、これらの意見を手作業により分類し、それぞれがネット上に存在するか否

かに関して、検索エンジンを用いて手動での類似意見検索を試みた。その結果、大まかな傾向として、ネット上の意見は社会調査により得られる意見を包含するという調査結果を得たが、これはあくまで手作業による調査であり、精度に限界がある。そこで今年度は、昨年収集した意見に対して自然言語処理を用いた類似意見検索の自動化を試み、より精緻な検証を目指した。

ネット上からの類似意見検索に際しては、まず、ネット上の膨大な情報の中から、類似性の高い意見をいかに効率的に抽出するかという課題がある。このスクリーニングに際しては、今年度は、検索対象をブログ記事へと限定するものとするこゝで対応した。これにより、広告等のノイズを減らし、ネットユーザーにより表明された意見のみへと検索範囲を限定することが出来る。そのうえで、とある話題についてウェブ上に意見が表明されるのは何らかの事件や事故によりその話題が広く喚起された場合であると仮定し、検索する意見についてウェブ上で活発にブログ記事が掲出された期間を推定し、その期間に投稿された記事のみを対象に意見抽出を行うこととした。これらの手法により、探索効率を高めるとともに、常時特定の話題について投稿しているユーザの偏りを軽減することが出来る。その上で、類似意見の確認作業を行った。

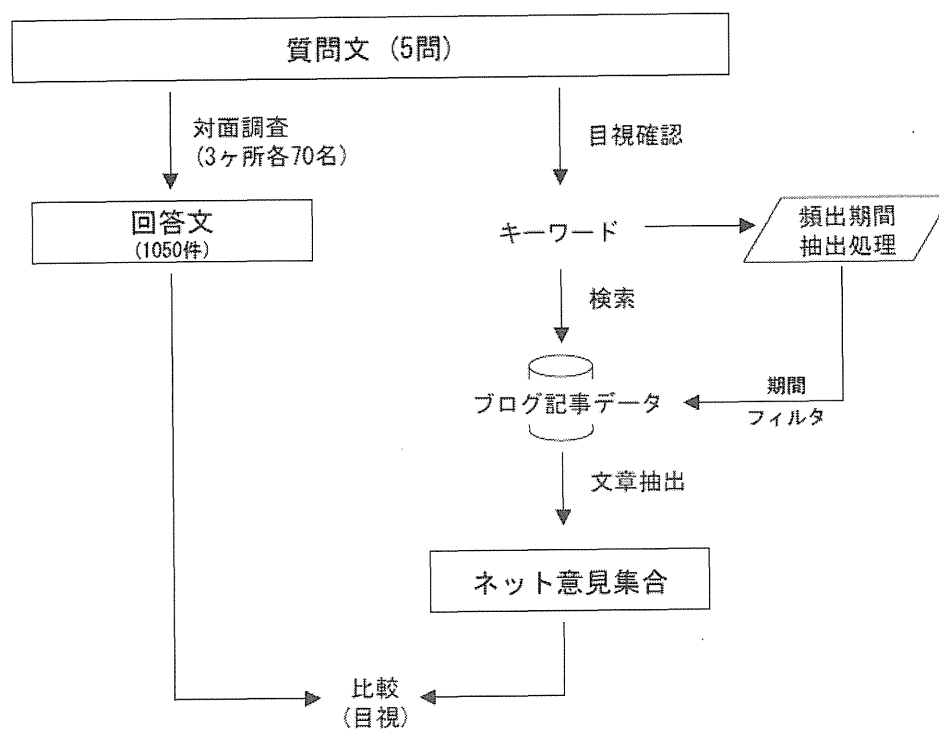


図 3. データ処理概要

(倫理面への配慮)

本アンケート調査は、「震災後のデマ情報の情報源と、その拡散、終息に寄与する要因の分析」として国立保健医療科学院倫理審査委員会より承認されている調査(NIPH-IBRA#12037)の一部である。

C. 研究結果

社会調査側の意見収集に際しては、ネットを日常的に利用している層ではなく、出来る限りネットへの接触が少ない相手を対象とした調査が望ましい。そこで、昨年度の社会調査においては、調査地域の住宅を5世帯おきに訪問したうえで、対面方式のアンケート調査への協力を要請するランダムウォーク法を用いた。合計210件の協力者のうち、年齢構成は20代4.8%、30代

12.9%、40代16.2%、50代18.6%、60代以上47.6%と高齢者が多く、平均年齢は56.6歳となった。また、男性36.2%、女性63.8%と、回答者の約2/3が女性となった。学歴別では、中学卒13.3%、高校・専門学校・短大卒65.7%、大学・大学院卒20.5%と、8割が高等教育を受けていない。インターネットの利用に関しては、「毎日利用している」31.4%、「週1回以上」12.9%、「週1回未満」9.0%、「利用しない」46.7%と、インターネットを日常的には利用していない者の割合が7割近くを占め、計画通りインターネットを日常的に利用しない層を抽出することが出来た。

また、類似意見を検索する対象のネット上の意見としては、放射能汚染に関する話題を抽出するために、2006年以降に日本語圏の3573万6960ユーザにより投稿された30億3462万3833件のブログ記事のなかから、2011年3月～2012年12月に投稿

された記事を対象として検索を行った。抽出に際しては、質問文に関連するキーワードの出現統計を用いて、話題が盛り上がった期間をさらに限定したうえで、当該の話題に関連する意見を抽出した。質問文のキーワード抽出に際しては、たとえば、「仮に食品の買い物の際欲しかった品物に「福島産」と書かれていたら、どうしますか？」という質問に対して「食品」「福島産」というキーワードを抽出し、検索に用いた。さらに、検索結果の各記事に対して、キーワードが出現する前後 150 文字以内の文章を抽出した。以上の処理の全体図を図 3 に示す。

こうした前処理を行ったうえで、「実社会には存在するもののネット上には存在しない意見」の存在を検証するために、アンケートにより収集した放射能に関する意見について、それぞれがネット上に存在するか逐次検証を行った。この社会調査の各回答と類似した意見を検索する作業においては、当初、類似意見検索技術の精度向上を目指して自動処理を試みていた。しかしながら、試みた手法では実用的な性能を出すことが出来なかったことから、最終段階での意見抽出作業については手動で代替した。なお、この抽出に際しては、意見中の各文を独立した意見と見なして処理を行った。たとえば、「自分自身で米作りを行っており、以前に比べて基準値が厳しくなった。風評被害のとばっちりを受けているように感じているので、産地別の米の放射能濃度を食品に表示して公表してほしい。」という回答文があれば、これは 2 つの意見と見なして処理した。

最終的に、各質問文のそれぞれ頻度上位 8 件の回答文全てについて、ブログ記事から類似意見を確認することができた。このことから、昨年度と同様に、ネット上の意見の定性的分析は、社会を知るうえで程度信頼に足る手段であると考えられた。

D. 考察

今回、社会調査によって得られた意見について、ブログ記事より類似意見を抽出することができた。しかしながら、抽出途中の記事には、多くのノイズが混入していることが明らかとなった。とりわけ、ブログには新聞記事の引用が多数含まれている傾向がある。今後、新聞記事に良く含まれる表現や意見に良く含まれる表現をヒントに、これらのノイズを除去する工夫が望まれる。

また、今回は、質問文から抽出したキーワードを用いてスクリーニングを行ったが、ブログ上の意見に該当するキーワードが欠落していた場合、本手法では検知することが出来ない。したがって、社会調査によって得られた意見側から特徴的なキーワード抽出を行い、このキーワードを用いたスクリーニングを行う等の工夫が求められる。

また、今回の手法では、各回答文毎に確認できた類似意見の数や質にばらつきがあった。たとえば、「～欲しい」「～すべき」という要求型の意見もあれば、「～と思う」「～と考える」のように感想型の意見もある。こうした特長に着目し、回答の種類毎に類似意見の抽出に際した戦略を変える工夫も求められるかも知れない。また、前述の通り収集できる意見に制約が生じる可能性があるものの、より検索に適した回答が得られるよう質問文に対して工夫を行う方向性についても検討の余地がある。

最後に、抽出した意見文と回答文の類似性判定に関しては、より大規模で客観的な処理を実現する自然言語処理による自動化が望ましい。一方で、ソーシャルメディアに見られる口語的な文書は表記揺れが大きく、厳密な類似性の判定が容易ではない。今後、機械的な類似性の判別に向け、辞書の整備等により単語や表現の曖昧性を解消する必要があると考えられる。

E. 結論

本研究分担は、社会調査により得られる意見とネット解析により得られる意見の比較を通じて、インターネットにおける意見解析手法の妥当性に関して検討した。その結果、社会調査により得られる意見のそれぞれについて、インターネット上のブログにおいても類似する意見を大まかに確認することが出来た。このことから、インターネットにおいて表明されている意見は、社会調査と比較して、同等以上の多様性を有するものと考えることが出来る。このことは、社会調査における情報源として、ネット情報の定性的な解析が一定の妥当性を有しうることを示唆している。

この妥当性のさらなる検証に際しては、大規模化と客観性の担保に向けて、現在目視判定に頼っている類似意見判定の自動化が望ましい。今後、検証の網羅性、並びに、精度の向上に向け、ノイズの除去、スクリーニング手法の効率化、意見の類型化、曖昧性の除去等、要素技術の発展が求められる。

F. 研究発表

1. 論文発表

なし

2. 学会発表

なし

厚生労働科学研究費補助金（食品の安全確保推進研究事業）
分担研究報告書

大規模 Web 情報分析システムを用いたリスク情報分析と
リスクコミュニケーションへの応用

研究分担者 鳥澤健太郎・大竹清敬・川田拓也
(独立行政法人情報通信研究機構(NICT) 情報分析研究室)

研究要旨

本研究は、総務省所管の独立行政法人情報通信研究機構情報分析研究室が運営費交付金により開発を行っている大規模 Web 情報分析システム WISDOM X を一例とし、自然言語処理技術を応用した Web 情報分析システムのリスクコミュニケーションへの応用可能性を検討した。特に、食品の放射線被害や風評被害、森林破壊問題に対して、億単位の Web ページを分析する WISDOM X を用いて実際に分析し、その結果がリスク情報発信者やそれを受け取る一般の人々に対して有用な情報が提供できるか検討を行った。その結果、WISDOM X を用いることで、ユーザは必要な情報を効率的に得られるだけでなく、意見分析によって多様な視点からの情報を、また「なぜ型質問」回答機能によってある情報の明確な根拠となり得る情報を得ることができ、さらに、因果関係分析により意外な情報を提供することでユーザに新たな気づきをも与えることが可能であった。インターネット情報の深い意味解析を行う Web 情報分析システムは今後のリスクコミュニケーションに対して、大きな転換をもたらし、将来的に幅広く活用されてゆけば、リスク情報発信者とそれを受け取る一般の人々に対して有用なツールとなることが期待される。また、行政側も、WISDOM X を用いて効果的に風評等誤ったリスク情報を収集し、その誤ったリスク情報自体やネット上に存在するその情報の根拠に対して迅速に反論することで風評に歯止めをかけることできる。このように行政による WISDOM X を利用した能動的かつ機動的な情報提供が今後のリスクコミュニケーションの一つのモデルになると考えられる。

A. 研究目的

2011 年 3 月に発生した東日本大震災および福島第一原子力発電所事故により、福島県及び周辺地域の農水産物に通常時以上の放射線量が検出される事態が生じた。それに伴い、検出された線量が基準値以下であっても全国的に福島県及び周辺地域の農水産物が敬遠されるなど風評被害も生じた。その背景にはインターネットの発達が挙げられよう。インターネットは誰もが自由に

情報発信できるという特性から、玉石混淆の情報が蔓延し、風評被害を助長する可能性がある。一方で、インターネット上には信頼性の高い有用な情報もあり、有効に活用すれば、風評に惑わされない意思決定も可能となる。しかしながら、インターネット上からユーザが必要な情報を的確に探し出すためにはある程度の情報リテラシーが要求される。さらに、インターネットから得られた情報の断片を取捨選択し、それらを元に推論し、自分なりの仮説や予測を立

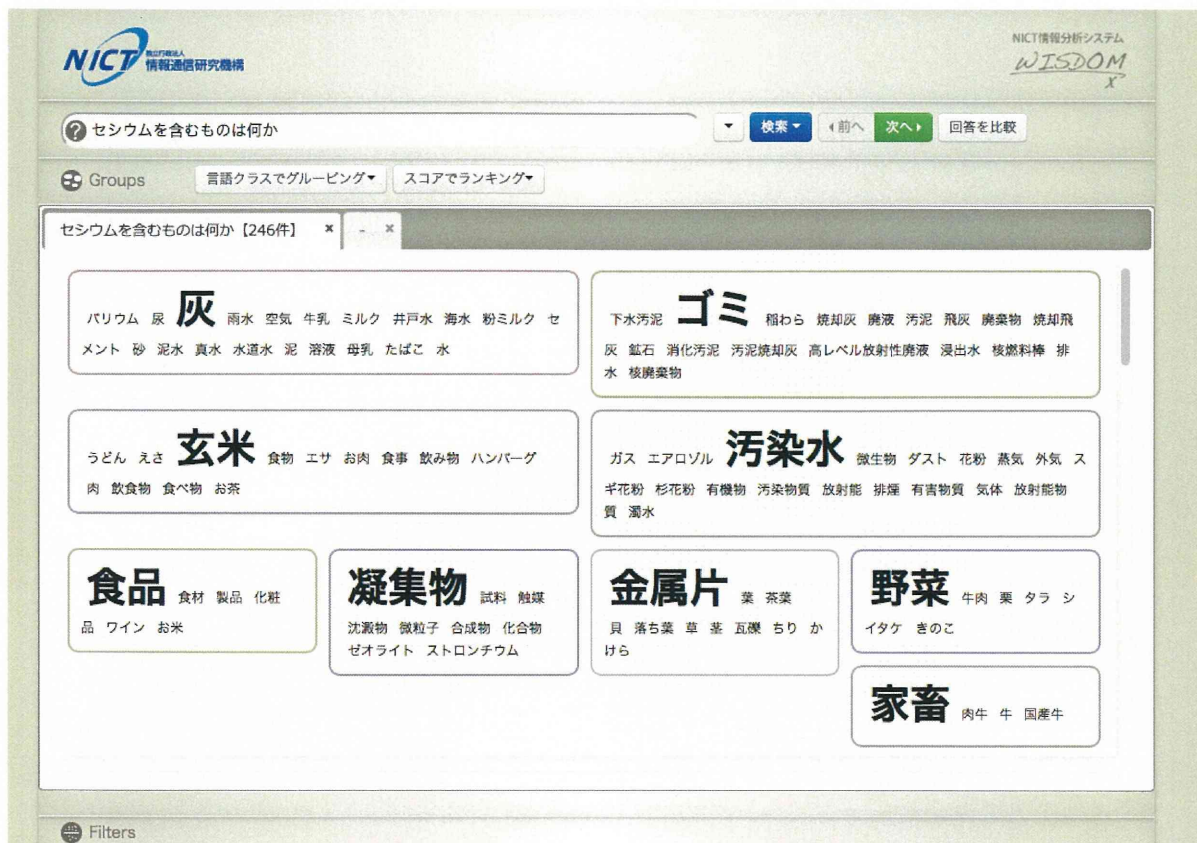


図 1 WISDOM X 検索結果 (Factoid 型質問)

てて意思決定を行うことも可能だが、相当のスキルとコストが必要となる。すなわち、インターネットは潜在的には、風評に惑わされない食品の安全情報の取得とその後の意思決定に有効である可能性があるものの、必ずしも一般的に活用できてはいない。

現在、総務省所管の独立行政法人情報通信研究機構情報分析研究室では、運営費交付金により、“WISDOM X”と呼ばれる Web 情報分析システムを開発している。これは自然言語処理技術を用いて、インターネット上にある大量のテキストの意味を解析し、ユーザの様々な質問に対して、多様な視点からの回答を提示するシステムである。これまではユーザが検索エンジンを用いて経験と勘を頼りに繰り返し試行錯誤しながら必要とする情報を探索しなければなら

が、WISDOM X を利用することで、必要とする情報を質問として送信するだけで、回答を得ることができる。さらに解析した情報の含意・因果関係を認識する事で、情報同士を組み合わせ、Web 上には書かれていない意外な仮説を構築することも可能である。本研究では、WISDOM X を一例に、Web 情報分析システムの活用事例としてリスクコミュニケーションへの応用可能性を検討する。

Web 情報分析システムを用いることで、これまで大量の情報に埋もれてしまい十分に活用しきれなかったインターネット上の情報を効率的に取得し、食品の放射線被害や健康被害、風評被害に対する有用性を検証する。



図 2 WISDOM X 検索結果(意見分析機能)

B. 研究方法

本研究は、Web 情報分析システムのリスクコミュニケーションへの応用可能性を検討し、インターネット情報の解析技術がリスクコミュニケーションに活用できる事例の収集と課題の整理を行うため、我々が開発している WISDOM X を利用して、リスクコミュニケーションに関連する事例を実際に分析して、その結果を定性的に評価するという方法を採用することとする。

そこで、まず、WISDOM X の概要を紹介する。図 1~3 に WISDOM X のスクリーンショットを示す。前述したとおり、

WISDOM X は質問応答システムで、ユーザからの様々な質問に対して回答を提示するシステムである。WISDOM X はユーザが問い合わせた質問の種類に応じて、その質問に適した分析結果を提示する。WISDOM X の機能は主として以下に挙げる 4 点である。

- Factoid 型質問への回答
- 意見分析機能
- なぜ型質問への回答
- どうなる型質問への回答

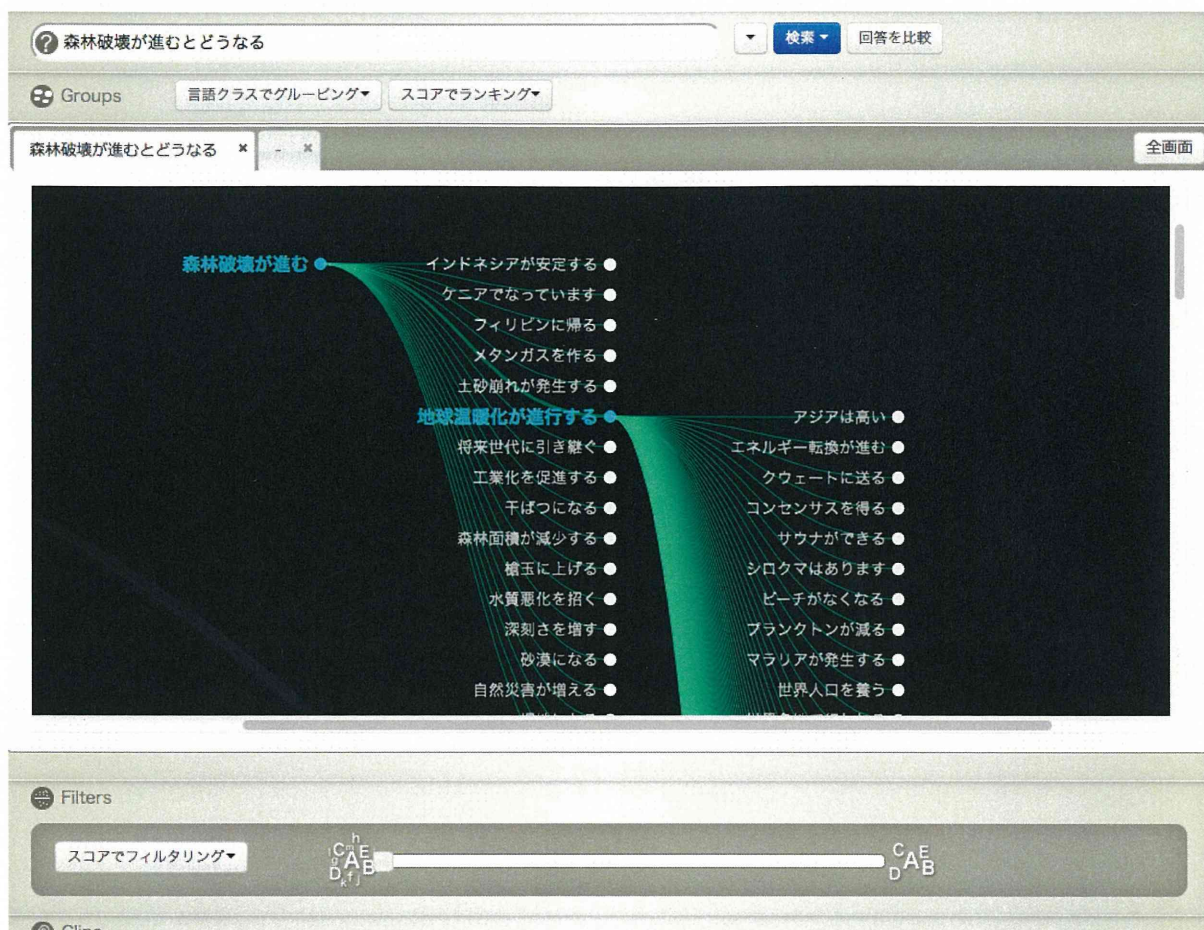


図 3 WISDOM X 検索結果(どうなる型質問)

Factoid 型質問は名詞一語で答えられるような質問を指す。図 1 は「セシウムが含まれるのは何か」という factoid 型質問の回答結果を表している。意見分析機能は、調べたい対象に関する、Web 上の肯定意見と否定意見を提示する機能で、図 2 は福島産の桃に関する意見を分析した結果を表している。意見分析機能を用いることで Web 上の人々の意見を比較することができ、多様な視点からの情報を得ることが可能となる。「なぜ型質問」とは、理由や原因を問う質問に対する回答を提示する機能で詳しくは事例分析で説明する。「どうなる型質問」はユーザが興味を持つ現象や出来事にまつわる因果関係の連鎖を推定し、その現象や

出来事が将来的にどうなるのか予測する機能で、図 3 は「森林破壊が進むとどうなるか」という質問に対する回答（シナリオ）の一部である。WISDOM X でさらに調べると「森林破壊が進む」→「地球温暖化が進行する」→「海水温度が上がる」→「腸炎ビブリオが多い」→「食中毒になる」という意外なシナリオが得られる。この意外な結果は荒唐無稽な因果関係とは必ずしも言えず、実際、地球温暖化が進んだ結果、腸炎ビブリオが増加したことを示した研究結果が報告されている¹。しかも WISDOM X が予測したシナリオは 2007 年の Web データが基になっているが実際に研究結果が

¹<http://www.nature.com/nclimate/journal/v3/n1/full/nclimate1628.html>



報告されたのは2012年である。また、「地球温暖化」と「腸炎ビブリオ」を含む文書は2007年のWebデータにはごくわずかしかな存在せず、それらを人手で精査した結果、いずれの文書においてもこれらのキーワードは独立のトピックとして記載されており、因果関係を表すシナリオとしての記載はなかった。すなわち、WISDOM Xは、分析時にはWeb上に書かれていない意外なシナリオを提示できる可能性が示唆される。

C. 研究結果

福島第一原子力発電所事故以後、連日のように、福島県および、その周辺地域の農水産物に対して放射線物質が検出されたという報道が流れたが、実際にどの農水産物に、何が検出されたのか報道後に改めて調べ、まとめることは容易ではない。

WISDOM Xを用いることで、セシウムが

検出されたものが何かを調べ、そのリストを得るという用途が考えられる。さらに検出されたものに対して、なぜその農水産物にセシウムが検出されたのか、その具体的な原因、根拠を調べることで、ある程度信頼性が担保された情報を得ることができる事が予想される。そこで「セシウムを含むものは何か」という質問をWISDOM Xに入力すると、図4のような結果が得られる。「牛」や「米」、「飲料水」といった回答の他、「ワカサギ」という回答が見られる。回答が抽出された元文をみると「ワカサギが640ベクレル/kgのセシウムを含んでいました」²と書かれたWebページが見つかる。同ページでは群馬県にある赤城大沼のワカサギからセシウムが検出された事は書かれているが、その理由については書かれていない。得られた回答の根拠を分析する手段として、WISDOM Xでは理由、原因を問

²http://tom.as76.net/social/umi_osen.php



う「なぜ型質問」に対する回答を提示する機能がある。たとえば、ここで、「なぜワカサギにセシウムが検出されるか」という質問を改めて WISDOM X に入力し、送信すると、図 5 で示したような結果が表示される³。回答の中には、ワカサギ特有の性質が関わっているとされる情報が得られ、例えば、回答の一つに次のようなパラグラフが Web ページから抽出されている。『なぜ赤城大沼のワカサギだけ高い濃度のセシウムが検出されたのか。地元関係者の多くは、主に二つの見方を挙げる。「一つは淡水魚特有の事情。海水を大量に飲み込んで吐き出す海水魚と異なり、淡水魚は、えさから取り入れたナトリウムやカリウムを体内で維持しようとし、次第に放射性物質の濃縮が進む」と漁業関係者は指摘。「もう一つは、赤城大沼特有の緩やかな水循環のスピ

ード。』⁴無論、ここで得られた根拠が真実であるかは保証できないが、こうした説得力のあるテキストを書くのに要するコスト等を考えると、根拠のない情報と根拠が与えられている情報とでは、後者の方がより確度は高いことが予想され、また、ユーザに判断材料が提供されるという意味では有用である。ただし、ここで得られた根拠から読み取れる情報は、ワカサギがセシウムを蓄積しやすい可能性に限られ、それが人体に影響するかどうかは程度の問題であり、ワカサギに蓄積したセシウムが人体に有害なレベルかどうかは不明である。仮に人体に無害なレベルであれば、このような根拠情報はかえって風評の原因にもなりかねない。そこで、行政側には、WISDOM X で発見した根拠情報を元に迅速かつ積極的な情報発信が期待される。例えば、ワカサギ

³ 本文に言及されていない分析結果は著作権等の問題を考慮し、ぼかしを入れている。

⁴<http://aws-fish.com/bulletin+index.page+article+storyid+477.htm>



図6「牡蠣 ノロウイルス」の意見分析結果

にセシウムが蓄積しやすいことは事実であるものの、人体に影響しないレベルであれば、風評の元にもなりかねないこの根拠情報に対して、専門家も交えた検証を経て科学的データと共に反論情報を提示することで、風評の根元を絶つことが期待される。逆にある程度この情報が支持できるものであれば、行政のサイトでリンクするなど積極的に情報提示するという事も考えられるだろう。

このようにユーザ側から見れば、WISDOM X を用いることで、どの農水産

物にセシウムが含まれ、さらになぜそれにセシウムが含まれるのか、ある程度の根拠を得ることができ、また、行政サイドでは、ここで挙げたような民間から発信された根拠情報に対して専門家の検証を経て、行政のサイトにて引用する方法も、より多様な複数者がコミットした信頼のおける情報を提供するという意味で検討する必要があると思われる。こうした先進的な情報システムと専門家等の連携による地道な活動が説得力の高いリスクコミュニケーションにおいては重要であると考えられる。

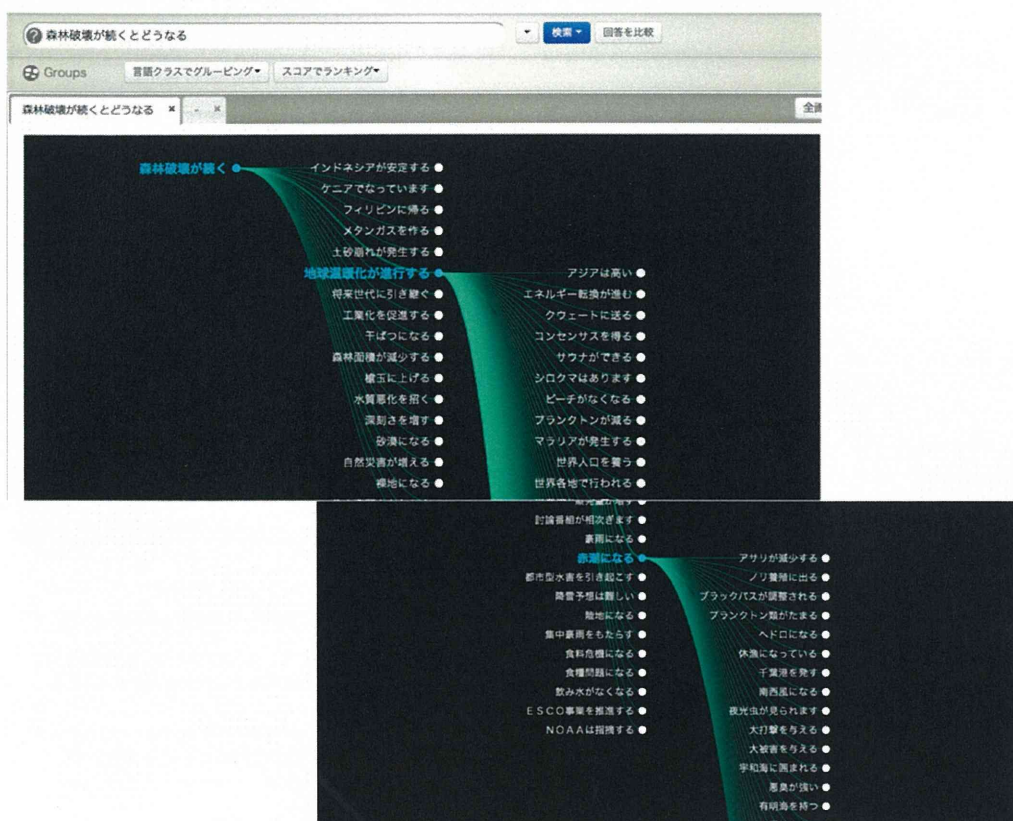


図7 「どうなる型質問」回答画面

次に風評被害について WISDOM X を用いて分析する。まず、factoid 型質問として「風評被害を受けたのは何か」という質問を入力してみる。すると、回答として「野菜」「牡蠣」「カイワレ」「農家」「福島」などが回答として表示される。その中で「牡蠣」について元文を見てみると「ノロウイルスの風評被害をもちに受けた「牡蠣」のことが、地方紙に載っていました」⁵という記事が見つかる。「意見分析機能」を使って、人々の意見を分析することで、当該トピックについて多様な視点からの深い分析をすることも可能である。例えば、「牡蠣 ノロウイルス」で意見分析をすると、図6のように、牡蠣のノロウイルスに関する意見が抽出される。否定意見の中には「風説被害

が深刻化している牡蠣(かき)についてだが、今までのところノロウイルスによる食中毒の原因食材として特定された事例はないという。」⁶という文が見られ、情報源のページを確認すると上記は YOMIURI ONLINE の記事の転載であることが分かる。このように、意見分析機能を使って、回答のさらなる分析が可能となる。

最後に「どうなる型質問」を用いた分析を行う。近年、世界的に森林破壊が問題となっており、このまま森林破壊が進むと、何が起こるのか、議論が続いている。専門家でも立場が異なれば、正反対の情報を発信する可能性があり、一般の人々が予測することは難しい。そのような状況で森林破壊が続くとどうなるか、様々な情報源から

⁵http://takeyama.weblogs.jp/takeyama/2007/02/post_58e0.html

⁶<http://www.gamenews.ne.jp/archives/2006/12/1159650.html>



図 8 「なぜ温暖化が赤潮を引き起こすか」

その予測を提示するシステムがあれば、一般の人々が判断する材料として有用である。WISDOM X は、億単位の Web データを基に、事象間の因果関係を推定し、ある事象が今後どうなるかを予測として提示することができる。例えば「森林破壊が続くとどうなる」という質問を入力すると、森林破壊が続いた結果、どのような事態が生じるか推定する。その結果の一部を図 7 に示す。この図が示すように森林破壊が進むと、「地球温暖化が進行する」ことが示唆されている。地球温暖化は様々な事象を引き起こすことが WISDOM X の結果から見えてくるが、ここでは「赤潮になる」という事態に注目した。森の環境破壊は、最終的には海の被害である「赤潮」へとつながり、その結果、「アサリが減少する」「漁業被害をもたらす」などの事態が生じる可能性が回答として提示される。

さらに生成されたシナリオをなぜ型質問回答機能で検証することもできる。例えば、「なぜ温暖化で赤潮が発生するか」とい

う質問を入力すると、「地上の温度が上昇すると海水の温度も上昇します。赤潮は、生活廃水や工業廃水により、窒素やリンが増加し、さらに水温が一定温度に達することにより大量に発生します。地球温暖化で水温が上昇し、赤潮が発生しやすい状態になっているとも考えられます。」⁷ という情報が見つかる(図 8)。このように WISDOM X によって因果関係の検証を行いつつ、シナリオを構築していくことが可能となる。これによって、生成された因果関係がある程度根拠づけられる。

D. 考察

本節では WISDOM X の分析を通して、Web 情報分析システムのリスクコミュニケーションへの応用可能性について考察を試みる。

⁷ <http://aksio.seek-pi.com> (現在はリンクが切れている)

東日本大震災および福島第一原子力発電所事故では、ツイッターを始めとする Web を利用した情報発信が盛んに行われ、その結果、真偽不明の情報も出回り、専門家の正確な情報が埋もれたり、不正確な形で伝播したりするなど、負の側面も明らかになった。誰もが容易に情報発信を行えるようになったことで、リスクコミュニケーションのあり方が改めて問われている。しかし一方で、多くの正確で有用な情報も Web 上には蓄積されている。これまで専門家と一般の人々の間でリスクコミュニケーションが行われる場は限られており、テレビなど既存のメディアでは能動的な情報収集は不向きで、幅広くリスク情報を得ること自体、容易なことではなかったことを考えると、インターネットを効果的に利用することができれば、一般の人が容易にリスク情報にアクセスできるようになり、ある程度能動的に情報収集できるという点から、Web 上の情報はこれまでのリスクコミュニケーションを補完するものとしての役割が期待される。

WISDOM X は億単位の Web 文書を基盤にユーザの質問を回答として抽出することで、玉石混淆の Web 情報からユーザの要求する情報を抽出することができる。さらにその回答について以下のように深い分析を行うことも可能である。

- 「意見分析」を用いることで、その回答に対する Web 上で言われている意見や態度を多様な視点から把握可能
- 「なぜ型質問応答」を用いることで、回答に関する根拠となる情報を得ることが可能
- 「どうなる質問」では Web 上に書かれていない予想外の帰結を導き出す仮説を生成することが可能

このように、多様な視点から情報を得られることにより、偏った情報で判断することを防止できる。また、回答の根拠となりうる情報が得られることで、(完全ではないものの) より確度の高い情報をもとにした意思決定が支援される。さらに今後起こりうる Web 上にはまだ書かれていない意外な情報が生成されることにより、ユーザに新たな気づきを与えることができるだろう。

これは、リスク情報の発信者とそれを受け取る一般の人々双方にとって有益であると思われる。すなわち、リスク情報の発信者は WISDOM X を用いて、分析することで、Web 上でどのような意見が優勢なのか、一般の人々が何を求め、Web 上で何を表明しているのか把握することができる。そのため WISDOM X はリスク情報発信者にとっては、効果的なリスク情報発信の指針となりうる。逆にリスク情報を受け取る一般の人々は、WISDOM X を用いることで、Web 上から必要な情報を幅広く得ることができるため、既存のメディアを通じてしか得られなかった、偏りのあるリスク情報が補完され、より正確な意思決定が可能となることが期待される。

前にも述べたが、行政サイドとしては、リスク情報を発信する側として、WISDOM X を用いて、民間から発信されている情報を効率的に収集し、誤った根拠に基づく風評があれば、専門家による検証を経て、それに対して素早く反論して情報公開するなど、機動的な対応が可能となる。このように、WISDOM X が備えるような膨大な情報を意味的に深く分析する機能を利用し、行政側に風評等に効果的に反論することはより良いリスクコミュニケーションの一つのモデルになると考えられる。