

ソーシャルメディアデータとクラウドソーシングデータを用いた事例調査

研究分担者 若宮 翔子 奈良先端科学技術大学院大学 研究推進機構 特任助教

研究要旨

マスギャザリング時や新興・再興感染症の発生に備え、感染症に関する一般市民の反応をモニタリングするために、具体的な事例（エボラ出血熱感染疑いと風しん）を題材として、ソーシャルメディアやクラウドソーシングサービスを通して収集したデータをそれぞれ収集・分析した。ソーシャルメディアデータ（ツイート）の分析では、エボラ出血熱感染疑い事例発生時の空気感染に関するデマ情報の拡散について分析し、実際にはデマ情報は少ないにも関わらずデマ訂正情報が拡散されるという興味深い現象を観察できた。この結果は、今後のマスギャザリング時や新興・再興感染症の発生時を見据えた事例検証になったといえる。また、クラウドソーシングの分析では、風疹ワクチンに対する一般市民の印象や知識の調査を行った。この結果、回答者の偏りや回答の信頼性に関する議論もあるが、短時間で大規模なアンケート調査を実施し、ある種の傾向を捉えることが可能であることが示された。

A. 研究目的

マスギャザリング時や新興・再興感染症の発生に備え、感染症に対する一般市民の反応や意識をモニタリングするために、ソーシャルメディアやクラウドソーシングなどを通して得られたデータを調査・分析する。

B. 研究方法

(1) ソーシャルメディアデータの分析

2019年8月4日エボラ出血熱疑いに関する事例をもとに、ソーシャルメディア上での一般市民のツイートを分析する。具体的には、2019年8月4日にTwitterに投稿された「エボラ」を含むツイートを取得し、ツイート数の推移やツイートの内容を公式情報の発信タイミングと比較しながら調査する。また、風評被害を引き起こしうるデマ情報の一例として「空気感染」に着目し、ツイート数の推移やツイートの内容を調査する。

(倫理面への配慮)

本研究で収集し解析するソーシャルメディアデータは、原則としてTwitterが公開している公式APIを用いて取得可能な、インターネット上

に公開されているデータのみを対象としている。

(2) クラウドソーシングによるアンケート調査

感染症やそれに対する行政施策に関する一般市民の意識について調査するために、クラウドソーシングを用いたインターネット上でのアンケート調査を実施する。なお、クラウドソーシング(crowdsourcing)とは、インターネット上の不特定多数の人々(クラウド; 群衆)に様々なタスクを委託するプロセスである。今回はYahooクラウドソーシングサービスのプラットフォームを用いて、アンケートへの回答タスクを委託したものである。タスクを完了した回答者には報酬としてタスクオーナー(業務委託元)が指定したポイントが付与される。2019年に流行した風しんやそのワクチン摂取を題材とし、アンケートを作成して回答を得る。

(倫理面への配慮)

クラウドソーシングサービスでのアンケートでは、回答者に提示されるタスク詳細画面にて、回答内容の取り扱いについてのポリシーを提示し、回答の送信を持って、調査の趣旨に同意したもの

とみなす旨を明記し、回答を得ている。図1にタスク詳細画面の掲載イメージを示す。

C. 研究結果

(1) ソーシャルメディアデータの分析

Twitterから「エボラ」というキーワードを含むツイートを収集したところ、2019年の1年間で77,196ツイートを取得した。8月4日のツイート数の推移を図2に示す。図2よりツイート数は急増したものの、その日のうちにほぼ収束していることがわかる。また、ピーク時のツイートを、デマやフェイクニュースなど真偽の如何に関わらず十分な根拠がない情報（デマ情報と呼ぶ）と、デマ情報を訂正したりデマ情報であることを指摘する発言（デマ訂正情報と呼ぶ）に、人手で分類した。その結果、8月4日はデマ情報やデマ訂正情報を含む大量の情報があふれる状態となっていたものの、深刻なデマ情報は少なく、多くはデマ情報を正そうとする善意に基づくデマ訂正情報であった。さらに、デマ訂正情報の多さからデマ情報が実際よりも多いと想起してしまう誤謬があることが分かった。

(2) クラウドソーシングによるアンケート調査

2019年11月半ばに、風しんの第5期定期接種対象者100名に対するアンケートを実施して回答を得た。対象者の半数以上が第5期定期接種について知らないと回答し、抗体検査やワクチン摂取を行っていない理由として、クーポン券が届いていないことが最も多く、次いで、検査や摂取に行く時間が取れないこと、予防接種に不安があることを挙げた。また、受診場所あるいは受診希望場所として、病院・診療所が最も多かった。他にも、回答者を限定せず、1,000名に対して風しんワクチン摂取に対する印象、風しんについての知識や経験（自由文による回答）や風しんに関する情報を得る際の検索行動（どのようなWebページを閲覧してどのような情報を得たか。自由文による回答）などについても調査した。1,000名分の回答は20時間以内で得られた。2019年11月の調査では、ワクチン摂取に対する印象は、ポジティブが21%、どちらかというポジティブが50%、どちらかというネガティブが23%、ネガティブが6%という

結果であった。

D. 考察

(1) ソーシャルメディアデータの分析

今回の事例において、正確な情報源を発信していれば、誤った情報を流す人に対して訂正を求める傾向が見られた。具体的には、「空気感染」は一見デマの流布が行われているように見えたが、実際には、政府の公式アカウントが発信した情報を広める動きが多く観察された。このように公式機関が一般の人々の目に触れやすいソーシャルメディア上で迅速に情報発信することによって、より正しい情報認識を人々に広めることができるといえる。

一方、件数は少ないものの、エボラ出血は空気感染するかどうかについては情報が錯綜していた。このような場合には、感染症や疾病予防の観点からも、国民に正しい情報を伝える手段として、ソーシャルメディアプラットフォームを活用し、正しい情報と人々の疑問や不安に対する回答を、短く簡潔でわかりやすい文章で発信することが望まれる。

さらに、今回対象としたTwitterには、話題になっているキーワードを表示する機能（トレンド機能）があり、同日に「空気感染」というキーワードが表示されていたことがわかり、デマ情報が実際よりも多いと想起してしまう誤謬があったと考えられる。今後、同様の事例が生じた場合には、こういった情報を監視するとともに、単なるキーワード抽出ではなく、要約フレーズ生成技術の向上も必要になるだろうという知見が得られた。

(2) クラウドソーシングによるアンケート調査

クラウドソーシングを用いることにより、回答者の偏りや回答の信頼性に関する議論もあるが、短時間で大規模なアンケート調査を実施し、ある種の傾向を捉えることが可能であることが示された。調査では自由回答文も得ているため、これを自然言語処理により解析することで、感染症に対する誤情報や誤認識の有無の検出や施策などのイベント前後での人々の行動や考えの変容の観察への応用も期待される。また、定期

的にアンケートを実施してモニタリングを継続することで、風しんやそのワクチン摂取に対する意識や認知度の広まりについてさらに発展的に調査する予定である。

E. 結論

感染症に関する一般市民の反応をモニタリングするために、具体的な事例（エボラ出血熱感染疑いと風しん）を題材として、ソーシャルメディアやクラウドソーシングサービスを通して収集したデータを収集・分析した。特に、ソーシャルメディアデータの分析では、感染症発生時のデマ情報について対象にしており、今後のマスギャザリング時や新興・再興感染症の発生時を見据えた事例検証であり、デマ情報は少ないにも関わらずデマ訂正情報が拡散されるという興味深い現象を観察することができた。今後は、日本に対して言及しているツイートをモニタリングすることで、海外における人々の反応についても調査・解析し、東京オリンピック・パラリンピック期間前・中・後を対象にケーススタディを行う予定である。

F. 研究発表

1. 論文発表
なし
2. 学会発表
 - 1) 村山太一, 若宮翔子, 荒牧英治. 点過程を用いたフェイクニュース検出. 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020) 論文集, 2020年
 - 2) 村山太一, 若宮翔子, 荒牧英治. 訂正投稿を用いたフェイクニュース収集システムの開発. 第12回データ工学と情報マネジメントに関するフォーラム (DEIM2020) 論文集, 2020年

G. 知的財産権の出願・登録状況

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

募集中

風しんに関するアンケート 2

条件付きポイント進呈 ヒント

タスクの種類

アンケート

もらえるポイント

5

タスクにかかるめやす時間

15分00秒

残り件数

1,000件

(全1,000件)

タスク実施中のユーザー: 0人

0%完了

このタスクって、どんな内容?

【なにをするの?】

風しんに関する質問にお答えいただけます。

【なにに使うの?】

風しんの予防接種に対してみなさまがお持ちのイメージや知識、そして予防接種の実施状況などを調査することを目的としています。

【回答内容の取り扱いについて】

回答の送信をもって、この調査の趣旨にご同意いただけたものとさせていただきます。
 また、回答は無記名ですので、このタスクへの参加によって個人が特定されることはありません。
 回答内容は研究利用の範囲において第三者に公開することがあります。
 調査結果につきましては、報告書を作成したり、情報系・医療系の学会や専門誌などで発表する可能性があります。

【注意事項】

回答に個人情報を含めるのはお止め下さい。

図 1. クラウドソーシングのタスク詳細画面の掲載イメージ

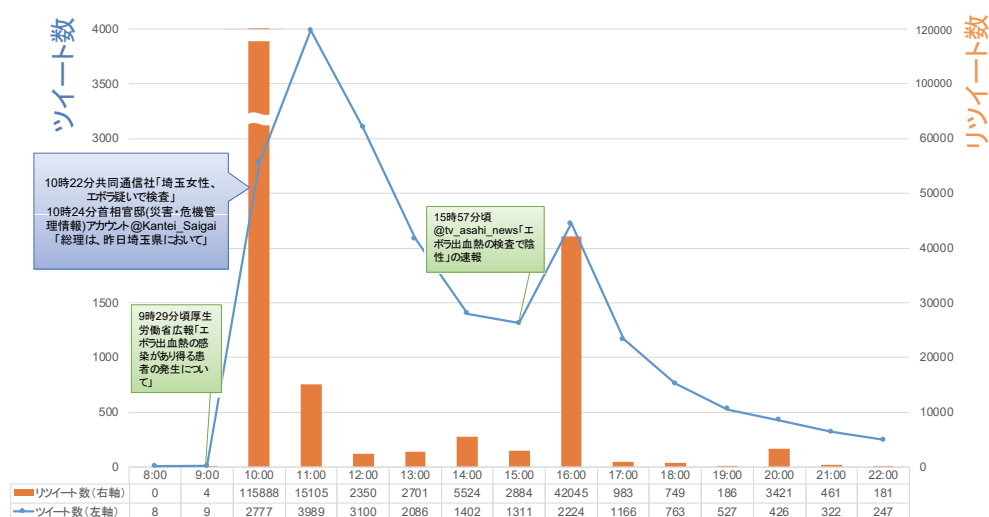


図 2. 「エボラ」に関するツイート数とリツイート数の推移 (8月4日1時間ごと)