

## 分担研究報告書

ウェブ検索クエリを用いた COVID-19 に対する市民の不安の定量的分析

研究分担者 若宮翔子(奈良先端科学技術大学院大学・先端科学技術研究科・准教授)

研究協力者 荒牧英治(奈良先端科学技術大学院大学・先端科学技術研究科・教授)

高 志偉(奈良先端科学技術大学院大学・先端科学技術研究科・大学院生)

### 研究要旨

マスクギャザリング時や新興・再興感染症の発生に備え、感染症に関する市民の反応をモニタリングするために、COVID-19を対象にウェブ検索行動を通じた市民の不安を調査する。具体的には、人々の不安の度合いを定量化した指標として、都道府県ごとの新規感染者あたりのCOVID-19に関するウェブ検索ボリュームLCIPP (Localized Concern Index by Patient Percentage) を定義し、時期別や年齢別に分析した。また、都市と地方における人々の不安の度合いの差異についても考察したところ、感染者が多い都市部よりも、感染者が少ない地方において人々のLCIPPが高い傾向にあることが示された。人々の過度な不安は風評被害やデマにもつながる可能性もあるため、定量的なモニタリングを継続することで、コロナ禍や今後の感染症発生時のリスクコミュニケーションの一助となることが期待される。

### A. 研究目的

マスクギャザリング時や新興・再興感染症の発生時の事例としてCOVID-19を題材に、一般市民の不安を定量化し、都市と地方における差異などについて調査する。具体的に、COVID-19の感染者数は東京都や大阪府などの都市部で多いが、1人の新規感染者に対する人々の不安の度合いという観点では、都市部よりもむしろ高齢者が多い地方の方が高い可能性がある。

本研究では、このような現象を分析するためのアプローチとして、ウェブ検索データを用いて都道府県ごとの人々の不安の度合いを定量化し、分析する。

### B. 研究方法

#### 【B-1. 検索クエリの選定と検索ボリュームの取得】

COVID-19に対する不安や心配を抱えている一般市民は、COVID-19に関する情報をウェブ検索などで積極的に取得していると考えられる。そこで、Yahoo!Japan検索エンジンにおけるCOVID-19前（2019年4月から5月）とCOVID-19中（2020年4月から5月）の全てのクエリの検索ボリュームを取得して比較した。この結果、COVID-19中に急増した上位100クエリのうち76クエリがCOVID-19に関連するクエリ

（“novel\_coronavirus”，“シャープ マスク”，“新型コロナウイルス”，“コロナ 感染者数”，“東京都 コロナウイルス感染者”など）であることがわかった。さらに、そのうち33ク

エリが都道府県名と“コロナウイルス感染者”というパターンであった。このようなクエリは、居住地域でのCOVID-19の感染状況を気に掛けている一般市民の心配や不安を反映していると考えられる。そこで、本研究では、都道府県名と“コロナウイルス感染者”からなるクエリ（ターゲットクエリと呼ぶ）の検索ボリュームを取得して分析する。なお、都道府県ごとの検索者数のばらつきの影響を抑えるために、都道府県ごとに常時検索されているベースラインクエリ（“東京23区”、“東京 天気 過去”など）とその検索ボリュームを取得した。

## 【B2. 検索ボリュームに基づく不安指標の算出】

B1. で取得した検索ボリュームを用いて、COVID-19に対する人々の不安や心配を定量化する。まず、検索ボリュームのみに基づく指標として、ベースラインクエリの検索ボリュームに対するターゲットクエリの検索ボリュームの割合をLCI (Localized Concern Index) として定義した。しかし、各都道府県のCOVID-19感染者数が多ければ、検索ボリュームも大きくなると考えられるため、新規感染者の割合に対するLCIをLCIPP (Localized Concern Index by Patient Percentage) として求めた。

(倫理面への配慮)

本研究で収集し分析するウェブ検索クエリデータは、Yahoo! Japanと奈良先端科学技術大学院大学ソーシャル・コンピューティング研究室との共同研究のもとで提供されている。個人を特定しないウェブ検索クエリの統計量として扱っており、データのプライバシーに配慮している。

## C. 研究結果

図1にターゲットクエリの検索ボリューム割合と新規感染者の割合に基づくLCIPPの値を地図上に可視化した結果を示す。全期間は2020年1月から9月までの結果を示す。1月から3月においては、最初にクラスターが発生した北海道のLCIPPの値が高くなっている。緊急事態宣言が発令された期間を含む4月から6月において

は、前後の期間に比べて全国的にLCIPPの値が高くなっていることが分かる。特に、新潟県、茨城県、長野県、岡山県、福井県など、地方におけるLCIPPの値が高くなっている。さらに、都市と地方で差があるかを考察するために、既存の統計データ（農業普及率、交通アクセス、人口変化や健康アウトカムなど）との比較を行った。図2に結果を示す。図2より、LCIPPと農業普及率に関する指標の一部（農家数  $p < .01$ ）、米の生産量 ( $p < .05$ ) についてやや相関があることを確認した。

図3に、年齢層別のLCIPPの結果を示す。年齢層は25歳から44歳、45歳から64歳、65歳以上の3区分とした。この結果、45 - 64歳が最も関心が高く、25 - 44歳が次いで高いことがわかった。一方、65歳以上は最も関心が低い結果となった。

## D. 考察

新規感染者に対する不安の度合いは都市部よりも地方でやや高い傾向があることが明らかになった。人々に過度な不安や心配を与えないように、地域ごとに情報発信方法や内容を工夫することなどが、今後のリスクコミュニケーションの一助になると期待される。図3の年齢層別の結果においては、45 - 64歳が最もLCIPPの値が高く、25 - 44歳が次いで高い結果となったが、これは自分自身に対する心配だけでなく、高齢の親などに対して心配する人が多いことが反映されている可能性があるとし唆される。一方、65歳以上の高齢者は最もLCIPPの値が低い結果となったが、そもそも65歳以上のネットユーザー数の少なさが影響している可能性があると考えられる。これはウェブ検索クエリデータを含む、ソーシャルメディアデータを利用する際に生じる課題である。

## E. 結論

COVID-19に対する不安を定量化するために、人々のウェブ検索クエリログを取得・分析した。新規感染者に対する不安の度合いは都市部よりも地方でやや高い傾向があるという結果を観察

することができた。本研究で提案した不安度合いを定量化した指標や、ウェブ検索クエリデータを分析して得られた知見は、今後のリスクコミュニケーションの一助になると期待される。

#### F. 研究発表 (発表雑誌名巻号・頁・発行年等も記入)

##### 1. 論文発表

- Shohei Hisada, Taichi Murayama, Kota Tsubouchi, Sumio Fujita, Shuntaro Yada, Shoko Wakamiya, and Eiji Aramaki: Surveillance of early stage COVID-19 clusters using search query logs and mobile device-based location information, Scientific Reports, 10, 18680 (2020).

<https://www.nature.com/articles/s41598-020-75771-6>

- Murayama T, Shimizu N, Fujita S, Wakamiya S, Aramaki E (2020) Robust two-stage influenza prediction model considering regular and irregular trends. PLoS ONE 15(5):e0233126.

<https://doi.org/10.1371/journal.pone.0233126>

##### 2. 学会発表

(査読なし)

- Zwiwei Gao, Kongmeng Liew, Taichi Murayama, Shuntaro Yada, Shoko Wakamiya, Eiji Aramaki, Nouyuki Shimizu, Sumio Fujita: Public Concern About COVID-19 Through Search Queries, 第13回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021(2021/3/2, オンライン)
- 村山太一、若宮翔子、荒牧英治: 検索クエリを活用した複数国におけるインフルエンザ流行予測、第13回データ工学と情報マネジメントに関するフォーラム (DEIM2021), 2021(2021/3/2, オンライン)

#### G. 知的財産権の出願・登録状況(予定を含む)

1. 特許取得  
なし
2. 実用新案登録  
なし
3. その他  
なし

#### H. 参考文献

なし

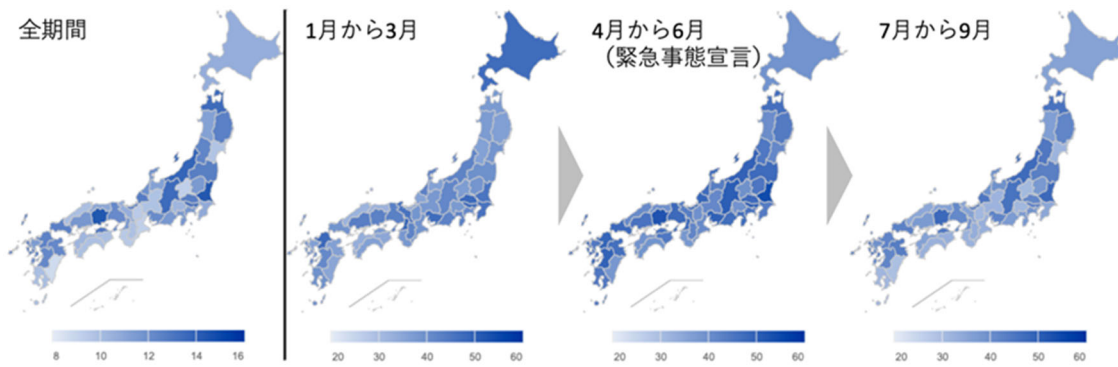


図1. 都道府県ごとのLCIPPの結果

比較指標		LCI	LCIPP
No. of farm households 農家数	Pearson's r	0.147	0.403
	p-value	0.323	0.005
Rate of population change 人口変化率	Pearson's r	0.522	0.022
	p-value	<.001	0.886
Farmland percent 農地割合	Pearson's r	0.235	0.232
	p-value	0.111	0.116
Rice production 米生産量	Pearson's r	0.020	0.309
	p-value	0.895	0.034
Reachable area within one hour 1時間以内到達可能面積	Pearson's r	0.585	0.258
	p-value	<.001	0.080
Travel time from Tokyo to major stations in each prefecture 東京から主要駅までの時間	Pearson's r	-0.264	-0.164
	p-value	0.073	0.270
Reported symptoms 有病者率	Pearson's r	-0.053	-0.160
	p-value	0.725	0.289
Daily outpatients 1日あたりの入院患者数	Pearson's r	-0.146	-0.019
	p-value	0.326	0.897

図2. 既存指標（農家数、人口変化率など8つの指標）との比較

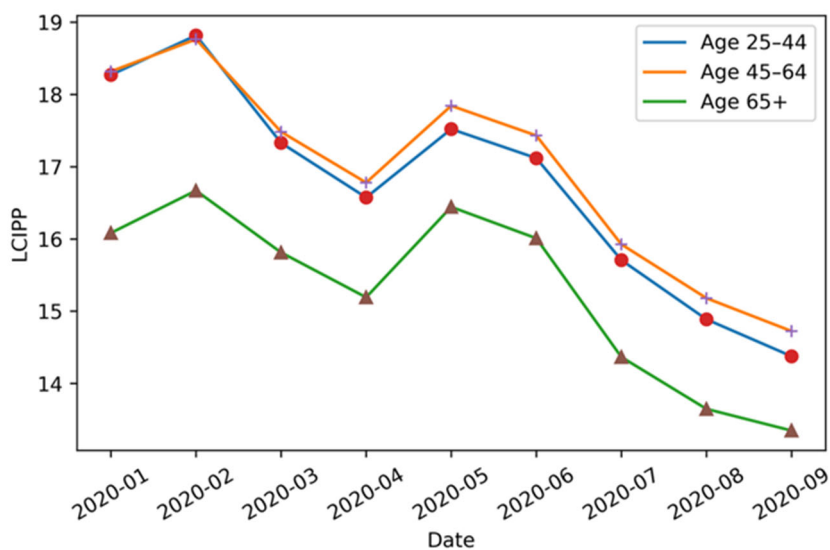


図3. 年齢層別のLCIPPの比較