

ソーシャルメディアを利用した新型コロナ禍における スポット別性格値推移に関する調査

伊藤 桃[†] 榎 美紀^{††} 小口 正人[†]

[†] お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

^{††} IBM Research - Tokyo 〒103-8510 東京都中央区日本橋箱崎町 19-21

E-mail: [†]momo-i@ogl.is.ocha.ac.jp, enomiki@jp.ibm.com, oguchi@is.ocha.ac.jp

あらまし 昨今、中国の武漢市にて発生した新型コロナウイルスが世界中に多大な影響を与え続けている。日本においても生活様式が変化してきており、例えば屋外で過ごせる場所を重視したり、外食を自粛して家の中で過ごす時間を充実させるなど、人々の行動は従来と異なるものになってきている。それにより、従来のスポットへの訪問目的が変わったり、訪れるユーザ層にも変化が生じている可能性がある。本論文では、ソーシャルメディアを利用して、日本の各スポットに対するコロナ禍でのユーザの性格値の推移を分析することにより、スポットを利用するユーザにどのような変化が起きているのかを検証する。

キーワード ソーシャルメディア, Personality Insights, 性格値

1 はじめに

昨今、中国の武漢市にて発生した新型コロナウイルスが世界中に多大な影響を与え続けている。世界の累計感染者数は 2020 年 12 月 22 日現在において 7700 万人を超えた。この驚異的な感染力により未だにウイルスの収束も見えないことから、感染対策を意識した日常生活に送っている人が多いだろう。日本においても生活様式が変化してきており、例えば屋外で過ごせる場所を重視したり、外食を自粛して家の中で過ごす時間を充実させるなど、人々の行動は従来と異なるものになってきている。それにより、従来のスポットへの訪問目的が変わったり、訪れるユーザ層にも変化が生じている可能性がある。これまでの従来研究では、ソーシャルメディアから得られるデータから、美術館群、神社群など様々なスポット群ごとに訪れるユーザの性格値を算出し、各スポット群ごとにその性格値に傾向が見られる事がわかった [1]。本研究では、引き続き後述する Personality Insights(以下 PI) [2] と呼ばれるツールを用い、昨今の社会情勢による各スポットの変化をユーザの性格値から分析をしたいと考える。

本論文の構成は以下のとおりである。2 章で関連研究について述べ、3 章では本研究で使用する PI について具体的に説明する。4 章で各スポットにおける月ごと性格値推移の検証、最後に、5 章で本稿をまとめる。

2 関連研究

本研究では、Twitter にて情報収集を行いそこで得たテキスト情報からユーザの性格値を判断する。Jalal ら [3] や, Gou ら [4] が Twitter のテキスト情報から性格を判断するその正確性について説いている。Jalal らは、空港の待ち時間などその場にはないと分からないローカルな情報を Twitter のユーザから

収集する qCrowd というシステムを構築したいと考え、Twitter の情報からローカル情報を提供してくれるユーザか否かを判断するモデルを作成した。結果としては、65%以上の確率で正しくユーザをモデルが判別した。また、Gou らも、同じように Twitter の情報から最も主流な心理学的な性格の指標である Big Five Personality に基づいてユーザの性格値を判定するモデルを作成しており、正確性を立証している。

また本研究では、テキスト情報から性格値を判断するのに独自のモデルではなく、PI という IBM のサービスを用いる。このサービスは Twitter などのユーザが書いたテキスト情報をインプットとして、性格の特性を、Big five(個性), Needs(要求), Value(価値感) の 3 つの次元に分割して出力する。PI を用いた結果を検証している研究論文として、富永 [5] の論文を挙げる。この論文では、Twitter のユーザの時間とともに変化する人格を、PI を用いて分析している。結果として、人格特性の変化が感覚的に理解のできる要因によって変化していて、PI の正確性を裏付けるような結果であると解釈している。先述した従来の研究から、Twitter から取得したデータから正確性の比較的高い性格値の判断ができることが言える。

また、Twitter から収集したデータからコロナウイルス禍における人々の感情分析を行なった研究は既に複数存在するので紹介したい。例えば鳥海 [6] らはコロナ関連のキーワードを含んだツイートを長期的に収集し、感情表現辞書を用いて時系列的にどのような感情の変化があるかを分析している。その結果、コロナに関するニュースや報道を受けて、恐怖の感情の上下が確認された。四方田も [7] 同様に、コロナウイルスに関連する単語を複数グルーピングし、どの要因からコロナウイルスに関するツイートを行なっているのかを長期的に分析している。本研究と従来研究の相違点としては、PI を用いて分析を行っている点やスポットごとのユーザの違いを見ている点である。

3 PI から分かる性格値

PI が推定する性格値は表 1 に示されるような性格の基本的な次元が 5 つであるという Big Five Model の特性項目と、Kevin Ford の Universal Needs Map に沿った Needs (欲求) 分析, Schwartz の価値概説 (Schwartz Value Survey) に沿った Values (価値観) 分析の特性項目からなる。

表 1 Personality traits estimated by PI

Big Five(個性)	数値の大小による意味
協調性	人当たりの良い・温情のある vs 冷たい・不親切
誠実性	勤勉・マメな人 vs 楽観的・不注意
外向性	外交的・エネルギー vs 孤独を好む・控えめ
感情起伏	繊細・神経質 vs 情緒安定な・自信家の
知的好奇心	好奇心が強い・独創的 vs 着実・警戒心が強い

4 各スポットにおける月ごと性格値推移の検証

本論文では、まず、各スポットに対して呟いているユーザの性格値を数ヶ月収集し、月間のそれぞれの性格値の差分の合計をみる。差分が大きければ、コロナウイルス禍にそのスポットに興味を持ったユーザ層に変化があったことを意味するので、その結果をもとにいくつか推察をしたいと考える。

4.1 検証用データセット概要

まず、Twitter の公式 API [8] を用いて、キーワードに複数の関東圏の観光スポットを設定しデータを収集した。キーワードに設定した各スポットについてツイートしているユーザを、そのスポットに興味を持つユーザと仮定し、設定したキーワードで見つけた日本人ユーザの過去の全てのツイートを PI を通して表 1 の各特性の数値を検証に用いる。数値は、具体的には 0 から 1 の連続値である。なお、PI は 100 単語以上からなるテキストからしか性格値を算出しない制限がある。よって 100 単語以上のツイートをしているユーザに限定している。また、取得スポットは、公益財団法人東京観光財団の運営する「GO TOKYO」[9] に記載されている離島などを除く都内のスポットを参照している。また、データ取得期間中一定以上のツイート数を得られたスポットを選び、正しく性格値を判断できないと思われる箇所 (URL 表記など) は適宜ツイートを削除するなどして加工している。データ取得期間は、2020 年 5 月から同年 8 月まで計 4 回収集している。

4.2 各月間差分合計算出

月間のそれぞれの性格値の差分の合計をみる。差分が大きければ、コロナウイルス禍にそのスポットに興味を持つユーザ層に変化があったことを意味する。以下に合計差分の大きかった上位 5 つのスポットを表 2 を示す

また同様に、合計差分の小さかった下位 5 つのスポットを表 3 を示す。

飛び抜けて差分合計の数値が小さい、あるいは大きいと言っ

表 2 各月間差分合計 (上位 5 抜粋)

	スポット名	各月差分合計
1	銀座ソニーパーク	11.7
2	駒形橋	11.1
3	谷中銀座商店街	10.98
4	赤坂 ACT シアター	10.71
5	ハーモニカ横丁	10.59

表 3 各月間差分合計 (下位 5 抜粋)

	スポット名	各月差分合計
1	荒川河川敷	4.26
2	アニメイト秋葉原	4.02
3	隅田川	3.99
4	日本橋三越	3.84
5	国立科学博物館	3.45

たスポットは見られないが、上位と下位で比較すると数値に開きが見られる。後の章にて、詳しく上位と下位のスポットのコロナウイルス禍における変化やツイートの内容と共に詳しく考察をしていきたい。

4.3 上位スポット考察

先ほどの結果から差分が大きかった銀座ソニーパークから見ていく。視覚的にどの月に特に変化が見られるかを観察するため、各月ごとに収集各性格値を x 軸にとり、y 軸にユーザの性格値の平均をとったグラフを図 1 に示す。

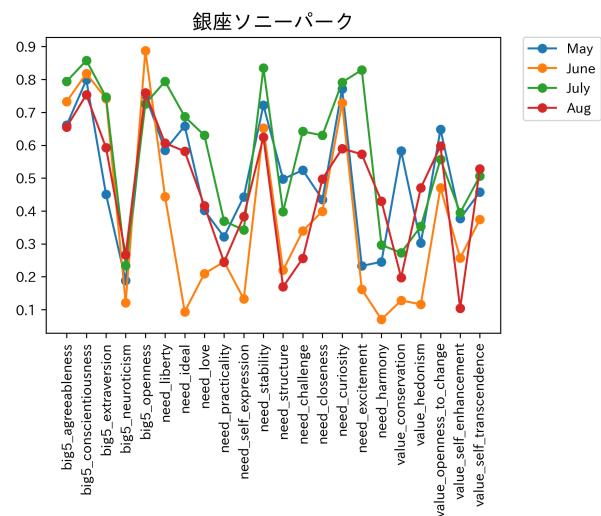


図 1 銀座ソニーパーク 月毎性格値推移

図を見ると、二つ考察できる事がある。一つには 6 月に大きな性格値の変動があるという事である。6 月に 2021 年 9 月末まで今のソニーパークを開園延長する発表があったのでその影響により、普段銀座ソニーパークに強く関心があるユーザ層ではないユーザも興味を示しツイートをしたので、それ以外の月の性格値と違う傾向を示したのでないか、と考える。また、二つ目に、7 月のグラフを見ると他の月と比べ、”need_liberty”(自由の求めやすさ) や”need_excitement”(熱意や意欲をそそる経験の追求しやすさ) といった値が大きい。これは 7 月に銀座ソ

ニーパークはオンライン上での配信ライブを行っていた事が関係しているのではないかと考える。実際にツイートの中身を見てみても、配信ライブをみた感想を呟いているユーザが他の月より多くみられている。

二番目に大きかった駒形橋のグラフは図2に示す。スカイツリーと共に見られる橋として有名である。グラフをみると全体的に性格値の値の上昇が見られる。PIにおいて、大半の性格値は上昇をするとポジティブな意味を持つことから、ポジティブなツイートをするとユーザ層が増えたのではと考えた。収集時期に関連するニュースをみると、駒形橋はコロナウイルス禍においては医療従事者を応援するためのライトアップをしていた。実際に「COVID-19 感染症に対応する医療関係者に感謝を表す為に欄干部をブルーにライトアップ! (日没後~23時) で、行ってきた」のように、ライトアップを見に行っていると思われるユーザも存在し、このようにライトアップに興味を示し、ツイートをしたユーザ層によって大きな差分に繋がったのではないかと考える。

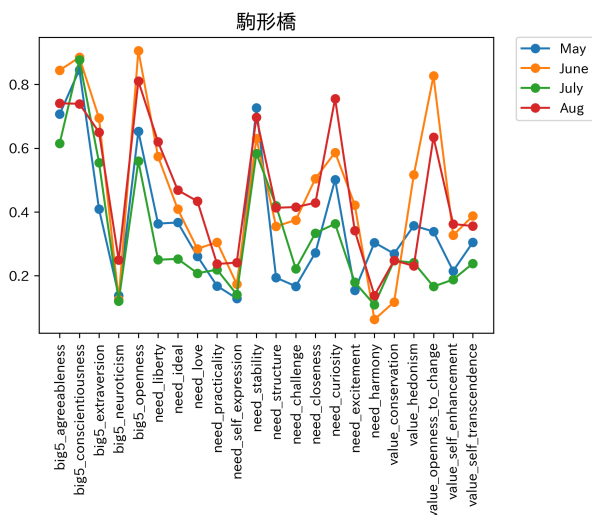


図2 駒形橋 月毎性格値推移

三番目に大きかった谷中銀座商店街は図3に示す。中身のツイートを確認してみると、谷中銀座についてツイートしているユーザに二つのパターンが見られた。一つ目のパターンとしては、コロナ禍の商店街の検査体制に不満を募らせているパターンである。実際、谷中銀座商店街を含む複数の具体的な商店街名を挙げ、「この辺りの商店街を検査対象にしないと不公平」などと呟いているユーザが見られた。二つ目のパターンとしては、谷中銀座商店街に最近行ってきた、あるいは行きたいと呟いているパターンである。例えば、「谷中銀座商店街! メンチカツ食べながら歩いた」などの実際に直近訪れていたであろうユーザの呟きを確認した。ツイート両者のユーザ層に違いがあり、またそのような様々なユーザ層が性格値の差分につながったと考える。

四番目の上位スポットであった赤坂 ACT シアターに関しては、図4に示す通り、各月を追うごとに数値が下がっていている。考えられる原因としては、ACT シアターの閉館がある。

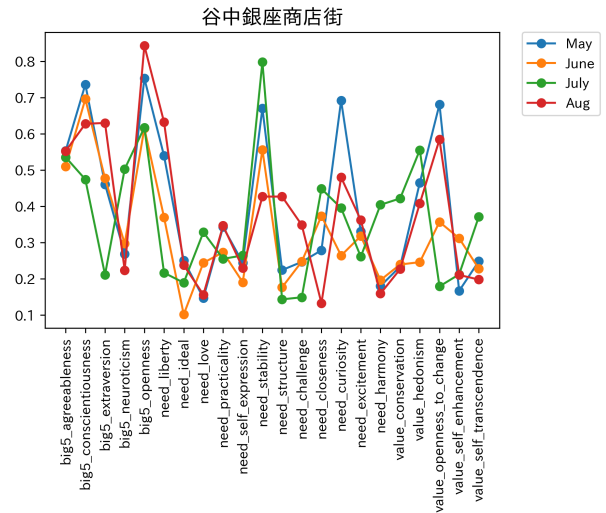


図3 谷中銀座商店街 月毎性格値推移

実は ACT シアターは 2021 年に閉館する事が決まっている。ツイートの中身を確認すると、閉館を惜しむ声が大きかった。前述したとおり、PI の性格値は値が高いほどポジティブな意味を示す性格項目がほとんどであることから、ネガティブな呟きは全体的な性格値を下げることに繋がらう。ACT シアターについて呟いているユーザは頻繁に ACT シアターの劇に関して呟いているような固定のファンが多くみられた。そのようなユーザ層のツイートが全体的にネガティブな呟きが増え、全体的な性格値を下げることは容易に想像できる。

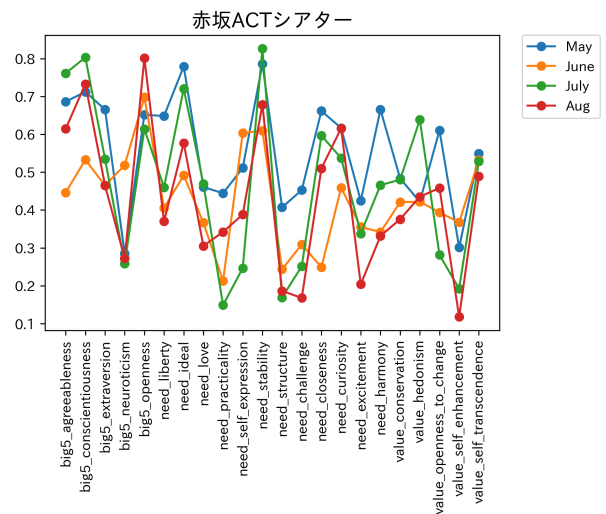


図4 赤坂 ACT シアター 月毎性格値推移

五番目であるハーモニカ横丁に関しては、以下の図5に示す。このような居酒屋の多い横丁に関してはコロナウイルス禍において批判的にあげられることが多く、実際にツイートの中身を見てみると「ハーモニカ横丁」を例に出して批判をするユーザと、実際に横丁に足を運んできた旨を呟くユーザと、横丁に行きたいと呟くユーザと大まかに三パターンの呟きを確認した。よって様々な性格のユーザによって、このような差分の大きさにつながったのではないかと考える。

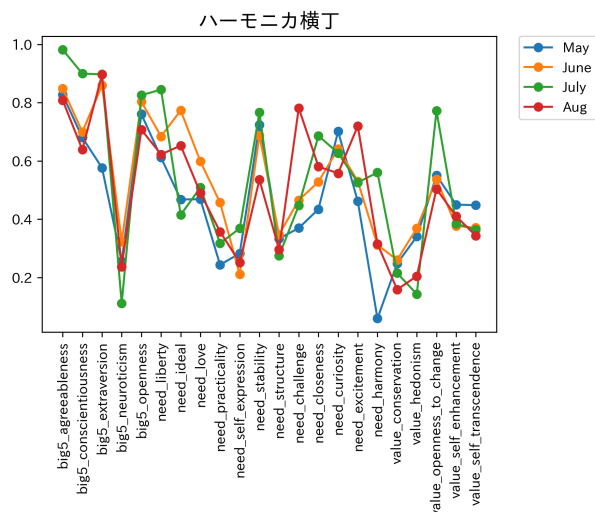


図5 ハーモニカ横丁 月毎性格値推移

4.4 下位スポット考察

先ほどの章と同様に差分が小さかった国立科学博物館をみていく。国立科学博物館より上位の下位スポットはコロナウイルス禍においても通常通りであるスポットが多かった。反対に上位のスポットはコロナウイルス禍の中で、閉館していたりといった事があり、その間のイベントやニュースによって性格値に大きな変化があった。しかしながら、国立科学博物館は6月まで閉館していたにも関わらず、先ほどの合計差分表からも下に示す図6からも性格の変動が見られない。

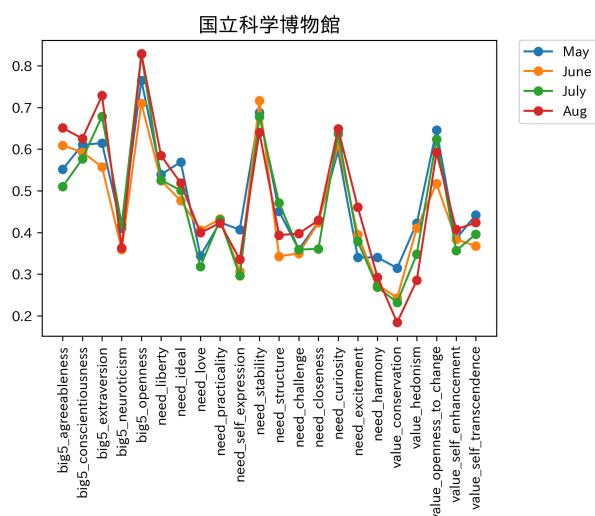


図6 国立科学博物館 月毎性格値推移

取得期間の科学博物館に関するニュースを見ると、閉館中でも自宅で博物館を楽しめるようにVRで館内を楽しめるサービスを開始していた。ツイートの中身を見てみても、自宅で楽しんでいる様子をツイートをしているユーザーが多く、実際に博物館に行っているような感覚があるので、月ごとに性格値の推移がなかったのではないと思われる。

5 まとめと今後の課題

今回は、コロナウイルス禍での変化を性格値の推移からスポットごとに分析をした。結果として、3ヶ月間を通して大きく性格値に変化のあるスポットはユーザ層に大きく変化はないが、スポット自体が何かしらのイベントを打ち出している、あるいは何かしらの影響によってユーザ層に変化があったスポットなどがあった。また、逆に性格値に変化はないが、コロナウイルス禍における特別なサービスによって通常通りのユーザ層とリアクションを獲得しているスポットも見受けられた。本研究によって、性格値によって変化を発見する事ができた。この研究の発展によって、例えばスポットの新規サービスに対する評判分析や新規ユーザに対するスポット推薦などに話を発展させる事ができるだろう。今後はそのような具体的なシステム構築に発展させたい。

文 献

- [1] 伊藤桃, 榎美紀, 小口正人. パーソナライズされた観光地推薦のための、性格特性を考慮することの効果検証. *DEIM Forum 2020*.
- [2] Personality Insights. <https://www.ibm.com/cloud/watson-personality-insights> (2020年12月22日閲覧).
- [3] Jalal Mahmud, Michelle X Zhou, Nimrod Megiddo, Jeffrey Nichols, and Clemens Drews. Recommending targeted strangers from whom to solicit information on social media. pp. 37–48, 2013.
- [4] Liang Gou, Michelle X Zhou, and Huahai Yang. Knowme and shareme: understanding automatically discovered personality traits from social media and user sharing preferences. pp. 955–964, 2014.
- [5] 富永登夢, 土方嘉徳. Twitter ユーザの受け取るフィードバックと人格特性の変化の関係に関する調査と分析. *知能と情報*, Vol. 31, No. 1, pp. 516–525, 2019.
- [6] 鳥海不二夫, 榎剛史, 吉田光男. ソーシャルメディアを用いた新型コロナウイルス禍における感情変化の分析. *人工知能学会論文誌*, Vol. 35, No. 4, pp. F–K45.1, 2020.
- [7] 四方田健二. 新型コロナウイルス感染拡大に伴う不安やストレスの実態: Twitter 投稿内容の計量テキスト分析から. *体育学研究*, Vol. 65, pp. 757–774, 2020.
- [8] Twitter Search API. <https://dev.twitter.com/rest/public/search>. (2020年12月22日閲覧).
- [9] GO TOKYO. <https://www.gotokyo.org/>. (2020年12月22日閲覧).