

テキストから推定される筆者の性格特性情報の活用の試みと考察

那須川 哲哉* 上條 浩一* 榎 美紀* 鈴木 祥子*

山下 紗苗** 上 泰** 権藤 恭之***

北村 英哉**** 尾崎 由佳****

*日本アイ・ビー・エム株式会社 東京基礎研究所

明石工業高等専門学校 *大阪大学 人間科学部 ****東洋大学 社会学部

表1: 寿命と関連付けられた性格特性

	説1[9]	説2 [10]	説3 [11]	説4 [12]	説5 [13]
Openness		高		高	
Conscientiousness	高	高	高	高	高
Extraversion	(低)		高	高	
Agreeableness		高		高	
Neuroticism		低	低	低	

1. はじめに

心理学の分野では、Big Five Model [1, 2] をはじめとして、人の性格特性を数値化して測定することにより、性格に関する様々な研究成果が生まれてきた。しかし、これまで性格特性を測定するには、調査対象者本人が、多数の質問文に対して回答する必要があった。例えば代表的な質問紙であるNEO-PI-Rでは、「私は心配性ではない。」、「大勢の人と一緒にいるのが好きだ。」といった240項目の質問文に対して、当てはまる度合いを5段階で回答する必要がある。NEOには60項目の短縮版、近年では、TIPIのように10項目で妥当性が確認されている性格検査も開発されているが、自分自身で質問項目に回答しなければならない点は変わらず、このような自己報告式の性格検査への依存が人の性格特性に関する研究を進める上での障壁となってきた。

例えば、人の性格特性と健康アウトカムや寿命の関係性を調べるためには、大規模なサンプリングをおこない、調査参加の同意を得、性格検査のほかにも社会経済変数を含めた多数の調査項目への回答を求め、さらに同一人物を長年に渡って追跡することがある。そのため、結果が出るまでに、数十年という年月が必要であり、仮説の検証が容易でない。さらに、そうして得られた調査結果の信頼性を確認するためには、同じような調査を何度も行う必要がある。

性格特性の測定が容易になれば、人の性格特性に関する研究の効率を向上させたり、その結果得られた知見を活用し易くなったりすると考えられる。

性格特性に関する研究成果の一つに、「書いたり話したりする言葉に人の性格特性が反映される」という知見 [3] がある。この知見に基づき、テキストから筆者の性格特性を推定するシステムが開発されるようになった [4]。筆者らは日本語のテキストから筆者の性格特性を推定するシステムを構築 [5] しネット上¹で公開している。このシステムを活用すれば、多数の質問への回答を依頼することなく、本人が書いた何らかのテキストさえあれば、性格特性の数値の推定が可能になる。

本稿では、このシステムを利用した、4つの研究事例から、その応用可能性と考慮すべき点を議論する。

2. 性格特性と推定システムの概要

筆者らが日本語版を開発したIBM Watson™ Personality Insights (以降PIと略) では、入力テキ

¹ <https://personality-insights-demo.ng.bluemix.net/>

ストに対し、その筆者の性格特性として、Big Five Modelとその構成要素のファセットに加え、Kevin FordのUniversal Needs Map に沿ったNeeds (欲求) 分析 [6] とSchwartzの価値概説 (Schwartz Value Survey) に沿ったValues (価値観) 分析 [7] の結果を出力する。

本稿で主に議論するBig Five Modelでは、性格の基本的な次元を下記の5つとし、個人の性格特性として、各次元に0から1の間の値を割り当てる。

- Openness to experience (知的的好奇心)
独創的・好奇心が強い vs. 着実・警戒心が強い
- Conscientiousness (几帳面さ・誠実性)
手際が良い・まめな人 vs. 楽天的・不注意
- Extraversion (外向性)
社交的・エネルギー豊富 vs. 孤独志向・控えめ
- Agreeableness (協調性)
人当たりが良い・温情あり vs. 冷たい・不親切
- Neuroticism (感情起伏)
繊細・神経質 vs. 情緒安定・自信家

PIでは、日本語テキストの場合、LIWC [8] に対応する代名詞や否定表現などに加え、助詞など日本語特有の表現の出現分布の特徴を分析し、質問回答結果から得られた性格特性とテキストが紐付いた学習データにおける出現分布情報との回帰分析結果をベースに性格特性を推定する。日本語版向けの千数百名分の学習データに対し、10-分割交差検証により、実測値と推定値のMean Absolute Error (MAE) を測定した結果は、各特性に対し概ね0.1程度の値である。このシステムを用いて、約10万人のツイッターユーザのテキストを分析し、その結果得られる性格特性が正規分布に従うように補正した結果も出力できるようにしており、本稿中の各調査における分析では、この補正された値を用いている。

3. 性格特性と寿命の関係の調査

3.1. 従来手法による調査

前述の通り性格特性と寿命の関係の調査には長い年月がかかるが、2000年代になって長年の成果が実り、様々な調査結果が報告されるようになってきた。その結果をまとめたのが、表1である。

表 2: 入力テキストから推定された筆者の性格特性 (「確度」は入力テキスト長に基づく推定精度の確度)

名前	寿命	入力テキスト [ソース]	確度	推定執筆時年齢	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism
樋口一葉	24	たけくらべ (1~4章) [A]	非高	24	0.94	0.32	0.03	0.82	0.85
石川啄木	26	一握の砂 (冒頭約2千5百文字) [A]	適正	24	0.99	0.37	0.36	0.93	0.9
中原中也	30	汚れちまった悲しみに [B]	低	23	0.72	0.5	0.53	0.47	0.33
芥川龍之介	35	名言集 [C]	低	31-35	1	0.55	0.52	0.37	0.36
夏目漱石	49	道草 [D]	非高	48	0.82	0.63	0.81	0.68	0.78
森鷗外	60	舞姫 (冒頭約6千文字) [A]	高	28	1	0.55	0.48	0.94	0.95
ガンジー	79	Speech [D]	非高	52	1	0.74	0.03	0.01	0.6
渋沢栄一	91	Speech [C]	低	76	0.97	0.87	0.82	0.52	0.47
松下幸之助	94	Speech [E]	低	73	0.95	0.95	0.75	0.53	0.29

[A] 青空文庫 (<https://www.aozora.gr.jp/>)

[B] 中原中也・全詩アーカイブ (<http://nakahara.air-nifty.com/blog/>)

[C] 偉人の名言・格言 (<https://iyashitour.com/meigen/greatman>)

[D] PIデモ・サンプルテキスト (<https://personality-insights-demo.ng.bluemix.net/>)

[E] 松下幸之助 (http://www.q-bic.net/biker_blog/2008/12/post-696.html)

表 3: 表 2 と同じ入力テキストから推定された筆者の Conscientiousness ファセットの性格特性

名前	寿命	死因	orderliness (秩序性)	self efficacy (自己効力感)	dutifulness (忠実さ)	achievement striving (達成努力)	cautiousness (注意深さ)	self discipline (自制力)
樋口一葉	24	肺結核	0.91	0.91	0.66	0.39	0.34	0.33
石川啄木	26	肺結核	0.8	0.82	0.25	0.65	0.31	0.4
中原中也	30	結核性脳膜炎	0.59	0.7	0.48	0.7	0.29	0.68
芥川龍之介	35	服毒自殺	0.66	0.84	0.08	0.44	0.33	0.53
夏目漱石	49	胃潰瘍	0.91	0.94	0.07	0.89	0.12	0.38
森鷗外	60	萎縮腎	0.92	0.95	0.45	0.56	0.35	0.32
ガンジー	79	銃による暗殺	0.75	0.66	0.7	0.63	0.99	0.44
渋沢栄一	91	直腸がん	0.61	0.86	0.4	0.8	0.3	0.64
松下幸之助	94	肺がん	0.98	1	0.89	1	0.23	1

調査により多少の違いがあるが、Conscientiousness が高いほど長寿であるという傾向がどの調査でも示されている。このように繰り返し得られた傾向に基づき、性格特性における Conscientiousness の高さが、寿命と正の相関関係にあることが定説になってきている。

3.2. テキスト分析を活用した調査

筆者らは、PIを用い、各種テキストから筆者の性格特性を推定し、その結果と筆者の寿命との関連を調査した。その結果を表2に示す。さらに、横軸に死亡年齢を、縦軸に性格特性の値を取って、プロットした結果を図1に示す。ここからConscientiousnessが高いほど寿命が長い傾向が読み取れる。表2から相関係数を求めると、Conscientiousnessが0.95 (P値: 1.23×10^{-9}) と実際に非常に高い値であり、対象者十名弱の小規模かつ簡易的な調査であるが、従来手法で長年をかけて導き出された結果と同じ傾向の結果が得られた。

さらに、Conscientiousnessの構成要素となるファセットレベルの分析を行った結果を表3に示す。この結果から得られた寿命と各ファセットとの相関係数を求めると、achievement strivingの値が最も強く0.62 (P値: 0.025)であった。self disciplineの値が次に強く0.53 (P値: 0.082)となり、他の値はP値が0.1を超える結果となった。従来の疫学研究では、ファセットレベルまで評価できるフルバージョンではなく、5因子のみを評価できる短縮版が使われている研究が殆どである。

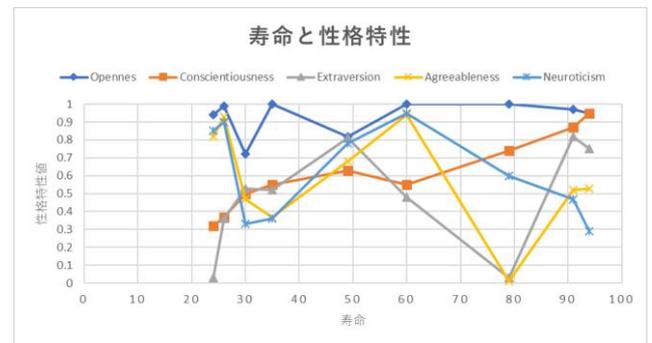


図1: 著名人のテキストから推定された性格特性と寿命の関係

性格特性と寿命の関係の調査を通じ、特定の性格特性が長寿に結び付くことが明らかになれば、その性格特性に特徴的な行動習慣などを調べることで、長寿につながる行動の解明につながる可能性がある。調査時間に制約がない本手法のアプローチでは、ファセットレベルのより詳細な側面が評価でき、これまで想定されていない、長寿の因果関係を含め、よりアウトカムと関連のある性格特性の詳細を検証できる可能性がある。

4. 生活環境と性格特性推定値の関係の調査

性格特性は、基本的には個人の特徴として一貫性があることが期待される。しかし、Big Five Model の各特性値は、

その判断基準が「人生を楽しんでいる」かどうかといった本人の感覚に依存していることから、個人の置かれている状況に応じて、多少なりとも変動する。

4.1. 被災状況における推定状態の変化

筆者らは、2016年に発生した熊本地震の前後でのTwitterデータをPIで分析した[14]。ツイート内容から被災者と認められるTwitterユーザのツイートを収集し、PIに入力したところ、震災発生前のテキストと被災中(地震発生から2週間)のテキストから推定される結果の一部(例えば、Neuroticismのファセットのvulnerability(低ストレス耐性)やConscientiousnessのファセットのcautiousness(注意深さ)など)の値に変化が見られた。多くのユーザにおいては、地震発生2週間以降のテキストから推定される結果は被災前のテキストから推定される性格特性とほぼ同じ値に戻っており、大震災の被災状況においては、一部の値の状態が一時的に変化するという傾向が見られた。こういった状態の変化を捉えることで、より良い被災者支援の実現に繋がれば、それは意義があると考えられる。

こうした調査は、質問紙を利用する従来型の手法では、被災者への負担をかけてしまうといった観点から、非常に困難である。Twitterデータの分析による調査は、Twitterユーザの特徴というバイアスが結果に反映されている可能性があり、簡易的な調査に過ぎない。しかし、その結果得られた仮説を考慮して本格的な調査に取り組むことで、調査の効率を上げることができる可能性がある。

4.2. 入院状況における推定状態の変化

今回、新たな取り組みとして、入院に関するつぶやきをしたTwitterユーザの分析を行った。

まず、ツイート内容から一度入院し既に退院したと認められるTwitterユーザ54名分のツイートを収集し、その中で、病気などにより、前もって入院が予定されていた39ユーザ(予定入院群)と、事故などで急遽入院することになった15ユーザ(突然入院群)を区別した。

各群のユーザに関し、入院前、入院中、退院後のテキストをPIに入力したところ、例えば、Agreeablenessの値の変化が、予定入院群では図2(a)のように、突然入院群では図2(b)のようになった。折れ線グラフの各線が一人分の変化を示している。予定入院群では、下がったり上がったりしている傾向に共通性が見出し難いが、突然入院群では、入院前のAgreeablenessの値が0から0.8程度までの値のユーザはことごとく入院中に値を上げているという共通性が見出された。得られた推定値の変化の特徴に共通性が見られたケースを表4に示す。

予定入院では大きな変化が見られなかったが、突然入院群では、一部の推定値において、入院中の変化に共通する傾向が見られた。

5. イベントにおける性格特性情報活用の試み

筆者らは、社内イベントの活動において、性格特性を考慮したメンバー構成を行うことで、協業の促進など、より良い成果に結び付ける可能性を調査した。

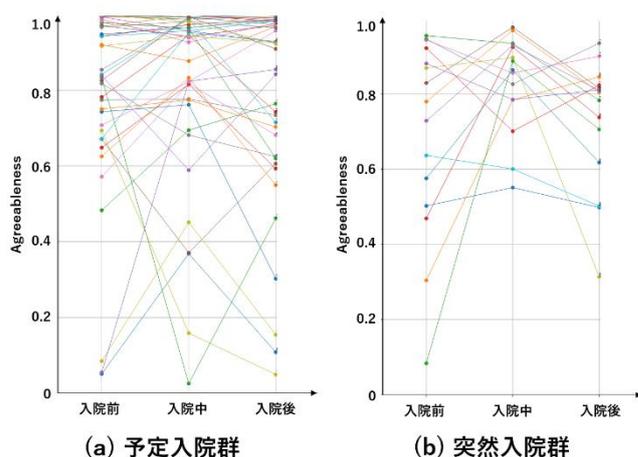


図2: 入院したTwitterユーザのテキストから推定したAgreeablenessの推定値の変化

表4: 入院中の変化が大きかった推定値の変化の平均値(変化具合=(入院中の値-入院前の値)+(入院中の値-退院後の値))

性格特性	変化具合		入院前後の差	
	突然入院	予定入院	突然入院	予定入院
Agreeableness	0.31	0.03	-0.01	-0.01
[Conscientiousness] dutifulness (忠実さ)	0.28	0.02	0.02	0
[Agreeableness] altruism (利他主義)	0.28	0	0.01	0
[Agreeableness] sympathy (共感度)	0.23	-0.01	-0.01	-0.02
[Needs] practicality (実用主義)	-0.29	-0.06	-0.01	-0.02
[Value] self enhancement (自己増進)	-0.39	-0.03	0.03	0.02

5.1. 親睦イベントでの活用

IBM 東京基礎研究所では、自然言語処理関連の研究を中心として、毎年夏に高尾山へハイキングに行き、中腹にある飲食施設で懇親会を行っている。2019年は8月23日に30名²で開催した。

開催に先立ち、参加者中の成人22名にPIを利用して得られた性格特性の値を提出してもらい、各提出者に対し、PIが推定する全特性値およびBig Five Modelの特性値のcosine類似度が一番近い人と一番遠い人、ValueおよびNeedsのcosine類似度が一番近い人という6つの基準で選んだ6名のリストを(誰がどの基準で選ばれたか知らせず)提供し、ハイキング中に積極的に話をするように依頼した。

また、2時間制飲み放題・食べ放題の懇親会を前半と後半に分け、前半は同じテーブルのメンバーのValueのcosine類似度なるべく近くなるような組み合わせの5グループのリストを、後半は同じテーブルのメンバーのBig Five Modelのcosine類似度なるべく近くなるような組み合わせの4グループ³のリストを用意し、各リストに従ったテーブ

² 参加者には、インターンの学生や研究所以外の部門の社員およびその家族といった多くのゲストが含まれており、大半の参加者にとって半数以上の参加者が初対面であった

³ 一部の参加者が途中で抜けるため後半のグループ数を減らした

ルに着いて飲食し歓談してもらった。

そして、イベント後に各参加者にアンケートを依頼し、「話をする機会ができて良かった」「話が盛り上がった」と感じた相手などを回答してもらった。

具体的名前をあげなかった参加者も多く、絶対数が少ないため、参考情報にすぎないが、Big Five Model の cosine 類似度が一番近い人の名前が一番多く回答されるという傾向が見られた。懇親会に関しては、これも数が少ないものの、後半の Big Five Model の類似度が近くなる組み合わせの方に、お互い話が盛り上がったと報告された組み合わせがより多く含まれていた。

5.2. 社内ワークショップでの活用

日本 IBM には COSMOS と名付けられた女性技術者のネットワークコミュニティが存在し、技術分野やイノベーション分野における女性の活躍を支援するために活動している。2019 年 11 月 5 日にはその一環として、グループディスカッションを中心とする 2 時間弱のワークショップが開催された。このワークショップ参加者に対しても、予め PI を使って得られた性格特性値を提出してもらい、その結果に基づき、各 4 人から構成されるグループの編成を行った。

上記親睦イベントの分析結果から、Big Five Model の有効性が比較的高いという感触が得られていたため、ここでは、Big Five Model の類似度が高い人が集まるようにした 7 グループ(類似群)と、逆に Big Five Model の各次元の値の分散がなるべく大きくなるようにした 8 グループ(多様性群)を編成した。さらに、性格特性値の事前提出をしていないメンバーから構成されるグループ(ランダム群)が 7 つ編成された。グループディスカッション終了後に「グループが盛り上がったと思う」かどうかなどのアンケートに回答してもらいその結果を分析した。その結果、盛り上がったという回答の度合いが最も高かったのは「多様性群」であった。

6. 考察

以上、テキストから推定される筆者の性格特性情報を活用する試みを示した。

蓄積されているテキストデータを活用することで、性格検査の負担をかけずに性格特性を推定できるため、従来型の手法では難しかった性格特性に関する調査が可能になり、仮説設定などで心理学の発展への貢献が期待できるものと考えられる。また、性格特性を考慮したグループ形成を試みたところ、小規模な実験ではあり、結論付けるには、再現実験を繰り返す必要があるものの、Big Five Model の性格特性に関して、懇親会では近い方が盛り上がり、グループディスカッションでは、多様性が大きい方が盛り上がるという結果が得られた。

性格特性が人の行動に影響を与えるのであれば、それを考慮することで、より良い社会の実現に近づける可能性が出てくると考えられる。例えば、新人に対して性格特性の近いメンターを付けることで新環境に溶け込み易くしたり、あるいは、メンバーの性格特性が多様化するよう組織を構成しダイバーシティを高めることで議論が活性化し易い

ようにしたりといった工夫が考えられる。

但し、こういった性格特性の情報を活用する上で留意すべき点が、その可変性である。本稿で示した通り、被災や入院といった環境の変化で性格特性の推定値が多少なりとも変化する場合がある。そのため、こうした性格特性の推定値を、採用試験などで、足切りのために使うのは不適切である。人材採用のような目的で性格特性情報を使うのであれば、従来型の基準で対象外になった候補者の中に、性格特性の観点で有望性を見出せる候補が存在しないかを再検討するような形で活用するのが望ましいと考える。

IBM Watson は International Business Machines Corporation の米国およびその他の国における商標。

参考文献

- [1] Goldberg, Lewis R. An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology* 59.6: 1216, 1990.
- [2] McCrae, R. R. and John, O.P. "An introduction to the five-factor model and its applications." *Journal of Personality*, 60(2), 175-215, 1992.
- [3] Mairesse, F., Walker, M.A., Mehl, M.R., and Moore, R.K., Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. In *Journal of Artificial Intelligence Research*, 30: 457-500, 2007.
- [4] Fabio Celli, Fabio Pianesi, David Stillwell, and Michal Kosinski, "Workshop on Computational Personality Recognition: Shared Task", AAAI Technical Report WS-13-01, 2013.
- [5] 那須川哲哉, 上條浩一, 日本語における筆者の性格推定の取組み, 言語処理学会第 23 回年次大会予稿集, pp.807-810, 2017.
- [6] Yang, H., and Li, Y. "Identifying user needs from social media." *IBM Research Report*, RJ10513 (ALM1309-013), 2013.
- [7] Ryan L. Boyd1, Steven R.Wilson, James W. Pennebaker1, Michal Kosinski, David J. Stillwell, and Rada Mihalcea, "Values in Words: Using Language to Evaluate and Understand Personal Values." In 9th International AAAI Conference on Web and Social Media, p.31-40, 2015.
- [8] Pennebaker, James W., Martha E. Francis, and Roger J. Booth. "Linguistic inquiry and word count: LIWC 2001." *Mahway: Lawrence Erlbaum Associates* 71 (2001): 2001.
- [9] Friedman, H.S. and Martin, L.R., 2011. The longevity project: surprising discoveries for health and long life from the landmark eight decade study. Hay House, Inc.
- [10] Chapman, B.P., Huang, A., Horner, E., Peters, K., Sempeles, E., Roberts, B. and Lapham, S., 2019. High school personality traits and 48-year all-cause mortality risk: results from a national sample of 26 845 baby boomers. *J Epidemiol Community Health*, 73(2), pp.106-110.
- [11] Spears, S.K., Montgomery-Downs, H.E., Steinman, S.A., Duggan, K.A. and Turiano, N.A., 2019. Sleep: A pathway linking personality to mortality risk. *Journal of Research in Personality*, 81, pp.11-24.
- [12] Kelly, E.L., 1955. Consistency of the adult personality. *American Psychologist*, 10(11), p.659.
- [13] Kern, M.L. and Friedman, H.S., 2008. Do conscientious individuals live longer? A quantitative review. *Health psychology*, 27(5), p.505.
- [14] 青木竜哉, 那須川哲哉, 吉川克正, 高村大也 and 奥村学, 2017. 震災被災者のソーシャルメディア分析を通じたメンタルケアの可能性 (言語理解とコミュニケーション)-(第 10 回テキストマイニング・シンポジウム). 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, 116(451), pp.65-70.