

AI 要約において人格的評価語が付与される条件に関する観察的分析

松本賢一

有限会社ネットクレードル

k.matsumoto@netcradle.net

概要

本研究は、AI 検索・要約サービスにおいて、対象が自ら人格や価値観を明示していないにもかかわらず、特定の人格的評価語が付与される現象に着目する。情報発信量や一次情報の厚みが大きく異なる二つの企業事例を比較した結果、いずれにおいても同一の人格語が要約結果として生成されていた。本研究では、この現象を人格推定や再現の問題としてではなく、評価情報が要約過程において、複数の具体的表現をもとに、一般化された人格的評価語へとまとめ上げられる過程として記述的に整理する。本発表は、AI 要約における評価語選択の特性について、人格的評価語が要約過程において抽象化されるという観点を提示する。

1 はじめに

近年、大規模言語モデル (LLM) を用いた要約や検索支援が普及し、企業や個人に関する情報が「どのような存在か」という形で要約・整理される場面が増えている。こうした要約結果には、事実情報の整理にとどまらず、先行研究においても、自動要約において評価的・主観的表現が生成される可能性や、テキスト情報から対象の印象や評価が付与されることが指摘されている。本研究で扱う人格的評価語の付与も、こうした知見と接続可能な現象と位置づけられる。

先行研究では、対象が自ら属性や人格的特徴を明示していない場合であっても、語彙選択や表現の傾向といった言語的特徴から、読み手や分析側が特定の印象や評価を付与しうることが指摘されている[6]。また、自動要約は、既存の文を抽出する手法にとどまらず、文脈理解に基づいて情報を言い換え・再構成し、原文の要点を凝縮した新たな表現を生成する抽象型の手法を含むものとして整理されている[5]。

しかし、それらの評価語が、どのような条件のもとで生成されているのかについては、必ずしも明ら

かになっていない。従来の研究では、感情や性格特性を推定・分類する試みや、特定人物の「らしさ」を再現する研究が多く行われてきた。一方で、対象が自ら人格的特徴や価値観を明示していない場合であっても、AI による要約の中で人格語が自発的に付与される現象については、十分に検討されていない。

特に、情報量が限定され、理念や思想を語る発信がほとんど存在しない対象に対して、なぜ特定の人格語が選択されるのかという点は、AI による認識や要約のあり方を考える上で重要な問いである。

本研究は、このような状況下において人格的評価語が付与される現象に着目し、AI 要約において評価語が選択される背景や要因を明らかにすることを目的とする。本研究では、ブラックボックスとしての LLM を前提とし、その出力結果に観測される人格的評価語の出現を、事例に基づいて記述的に分析する立場を取る。

具体的には、情報発信量や発信内容が限定された対象を取り上げ、どのような情報配置や構成のもとで人格的評価語が付与されるのかを観察し、その特徴を整理することを試みる。なお、本研究は、AI が人物や組織の人格的特徴を正確に推定できるかを評価するものではない。また、人格語が生成される因果関係を、モデル内部の挙動から説明することも目的としない。

2 関連研究

AI による人物理解や評価に関する研究は、これまで主に二つの方向で進められてきた。

2.1 感情・性格特性の推定に関する研究

SNS 投稿やコメント、レビュー文などの大量テキストを対象に、書き手や対象の感情状態、性格特性、印象評価を推定する研究が多数報告されている[2][3][4]。また、架空キャラクターを対象として、人格特性を心理学的枠組みに基づいてあらかじめ定義し、語彙資源や言語特徴を用いて人格を推定・プロファイリングする研究も行われている[7]。これらの

研究では、感情カテゴリや性格特性をあらかじめ定義し、教師データや評価指標を用いて分類精度や推定性能を評価する手法が一般的である。

このアプローチは、大規模データを用いた統計的分析や再現性の確保に強みを持つ一方で、評価語がどのような文脈や情報配置から選択されたのかという生成過程そのものには、必ずしも焦点を当てていない。

2.2 特定人物の「らしさ」を再現する研究

一方、著名人や歴史的人物を対象に、その人物特有の言葉遣いや思想、文体を再現する研究も行われている[1]。これらの研究では、対象人物に関する十分なテキスト資料が存在することを前提とし、「らしさ」を語彙選択や文体的特徴として定義し、生成結果がどの程度それに近づいたかを評価する。

この系統の研究は、再現性や評価基準が比較的明確である反面、あらかじめ「正解となる人物像」が設定されている点に特徴がある。

2.3 本研究との位置づけ

本研究は、上記のいずれの立場とも異なる。感情や性格特性を数値的に推定・分類することや、特定人物の「らしさ」にどれだけ近づけるかを評価することを目的とはしない。

むしろ、対象が自らの価値観や人格的特徴を明示していない場合においても、AIによる要約や整理の過程で、特定の人格語が自発的に付与される現象そのものに着目する。

この点において、本研究は、人格的評価が生成される条件を観察的に捉えようとする試みとして、既存研究を補完する位置づけにある。特に本稿では、人格語が推定や再現の結果としてではなく、AIによる要約過程における抽象化の結果として現れる現象に着目する。

3 研究方法

3.1 観察対象

本研究では、AI 検索・要約サービスにおいて人格的評価語が付与された事例を観察対象とする。対象は、公式ウェブサイトを持有するものの、経営理念や価値観を体系的に表明するページを持たない企業である。

事例 A では、施工実績を中心とした事実情報の掲

載を主とし、ブログ記事や動画コンテンツは一定量存在するものの、人格や価値観に関する情報は断片的かつ暗黙的に散在しており、体系的に言語化された形では提示されていない。

一方、事例 B では、公式ウェブサイトは存在するものの、ブログ記事や動画コンテンツといった継続的な情報発信はほとんど行われておらず、人格や価値観に関する言語情報は極めて限定的である。

本研究では、このように、

- (1) 情報発信量は一定程度存在するが、人格的特徴が非構造的・暗黙的にしか示されていない事例と、
- (2) 情報発信量そのものが限定的で、人格的特徴を示す言語情報がほとんど存在しない事例

という異なる条件を備えた対象を比較することで、人格的特徴が明示的に言語化されていない状況において、AI 検索・要約サービスがいかなる条件下で人格的評価語を付与するのかを観察する。

3.2 分析手順

分析は以下の手順で行った。

まず、対象企業に関して公開されているウェブページ構成、掲載内容の種類、および情報発信の形態を確認し、情報量や情報配置の特徴を整理した。

次に、AI 検索・要約サービスによって生成された要約結果を収集し、そこに含まれる人格的評価語の有無および内容を確認した。

その上で、対象が自ら表明していない人格語が、どのような文脈で要約結果に現れているかを中心に、記述的な比較・整理を行った。

3.3 研究上の制約

本研究は、特定の人格語が生成される因果関係を明らかにすることを目的としない。また、本研究は、特定の AI 検索・要約サービスにおける一時点での観測結果に基づくものである。さらに、LLM 内部の推論過程や学習データの詳細についても立ち入らない。

4 事例分析

本章では、AI 検索・要約サービスにおいて同一の人格的評価語「誠実」が付与された二つの事例を取り上げ、それぞれの情報発信量、情報構成、および第三者評価の状況を整理する。ここでは評価や因果の推定は行わず、観測された事実の記述にとどめる。

4.1 事例 A：岡山県倉敷市 K 社

情報発信の概要

K 社は、公式ウェブサイト、YouTube 動画、FAQ、施工事例ページなどを通じて、継続的かつ大量の情報発信を行っている。公式ウェブサイトに掲載されているテキスト量は、ヘッダー・フッターを含めて約 20 万文字を超えており、以下のような構成を持つ。

- 施工事例ページ：116 ページ
- FAQ ページ：56 ページ
- お客様の声：7 ページ
- 代表者コラム/会社の姿勢を語る記事：複数
- YouTube 動画：ショート動画含め 800 本超

これらのコンテンツでは、施工手順や判断理由、過去の経験、価値観に関する語りが、代表者自身の言葉によって詳細に記述されている。一方で、「誠実」という人格語は、主として「お客様の声」内で限定的に使用されており、代表者自身が自らを「誠実」と表現する記述は確認されなかった。

第三者評価の特徴

Google レビューは 7 件と件数自体は多くないものの、そのうち 4 件に「丁寧」という行為的評価語が含まれていた。また、公式サイトに掲載された直筆の「お客様の声」においても、「丁寧」という表現が 7 件中 4 件で確認された。

AI 要約結果

AI 検索・要約サービスにおいて、K 社は「誠実」という人格的評価語を用いて要約・紹介されていた。

4.2 事例 B：岡山県倉敷市 X 社

情報発信の概要

事例 B である X 社は、公式ウェブサイトを有しているものの、情報発信量は事例 A と比較して限定的である。公式ウェブサイト全体のテキスト量は、ヘッダー・フッターを含めて約 3 万文字弱であり、構成は以下の通りである。

- 施工実績ページ：19 ページ
- FAQ：1 ページ (6 件)
- お客様の声：1 ページ (8 件)
- 経営理念や価値観を明示的に語るページ：なし
- ブログ記事：1 ページ (1 件)
- YouTube 動画：1 件

掲載内容の多くは施工実績や事実情報であり、代表者自身による価値観や判断理由を詳細に語る記述はほとんど確認されなかった。

第三者評価の特徴

Google レビューは 8 件であり、そのうち 6 件に「丁寧」という行為的評価語が含まれていた。また、外部サービス (外壁塗装の窓口、ヌリカエ、リショップナビ等) に掲載された第三者評価においても、「丁寧」という表現が複数確認された。一方で、「誠実」という人格語は、公式サイトのお客様の声に 1 件、Google レビューに 1 件確認されるにとどまっていた。

AI 要約結果

AI 検索・要約サービスでは、X 社に対しても「誠実」という人格的評価語が用いられていた。

4.3 小結

以上の二事例はいずれも、AI 検索・要約サービスにおいて同一の人格的評価語「誠実」が付与されていた。しかし、情報発信量、情報構成、一次情報の厚みには大きな差が存在していた。

事例 A では、代表者自身による大量の一次情報と具体的な判断・語りが蓄積されているのに対し、事例 B では、第三者による行為的評価語を中心とした限定的な情報が主な材料となっていた。

本章では、これらの差異を事実として整理するとともに、人格語が生成された理由や因果関係についての解釈は、次章において検討する。

5 考察

5.1 人格的評価語が生成される条件について

本研究で取り上げた二つの事例では、情報発信量や一次情報の厚みに大きな差があるにもかかわらず、AI 検索・要約サービスにおいて同一の人格的評価語「誠実」が付与されていた。

この結果は、人格的評価語の生成が、単純な情報量や特定語彙の出現頻度のみに基づくものではない可能性を示唆している。特に、両事例に共通して確認された点として、第三者評価における行為的評価語「丁寧」が一定数存在していたことが挙げられる。

本事例において付与された人格的評価語は、公式サイト上で頻出している語ではなく、限定的な第三者評価や外部情報、ならびにコンテンツ全体の文脈

を踏まえた要約結果として生成されたものと考えられる。

ただし、「丁寧」という行為的評価語が直接「誠実」という人格語へと変換されたとは因果的に結論づけることはできない。本研究で観測されたのは、あくまで第三者評価における評価の方向性が一貫している状況下で、AI 要約結果として人格的評価語が付与されていたという事実である。

なお、同一企業に対する評価表現において、「丁寧」「誠実」といった信頼関連語が同時に出現する場合と、応答ごとにいずれか一方のみが強調される場合が観測された。本稿では主として人格的評価語が生成される条件の整理に焦点を当てるが、これら評価語の組み合わせや出現様式の差異が、意味的にどのような特徴を持つのかについては、今後の検討課題としたい。

5.2 下位評価語から人格語への要約的ジャンプ

人間が第三者評価を読む場合、「丁寧」「説明が分かりやすい」「対応が良い」といった行為的評価語に注目し、それらを個別に理解することが多い。一方で、本研究で観測された AI 要約では、これらの下位評価語にとどまらず、より抽象度の高い人格語によって全体がまとめられていた。

このことから、AI 要約過程においては、行為やプロセスに関する具体的な評価語を保持・列挙するよりも、評価の方向性を包括的に表現できる人格語が選択されやすい可能性が考えられる。これは、情報の省略や要約を目的とする AI の性質に起因するものと解釈できる。

ただし、このような要約的ジャンプが、対象の人格や人となりを正確に反映しているかについては、慎重な検討が必要である。

5.3 本研究の限界

本研究は、観測事例が二件に限られており、AI 検索・要約サービスとしても特定のサービスに依存している。また、モデルの更新や学習データの変更により、同様の結果が将来も再現される保証はない。

さらに、人格的評価語の生成過程については、AI 内部の推論や重み付けを直接検証したものではなく、あくまで出力結果に基づく記述的考察にとどまっている。これらの制約を踏まえ、本研究の結果は探索的・観察的な知見として位置づけられる。

本研究で観測された人格語付与の傾向が、実務的な情報設計にどのような影響を与えるかについては、今後の検討課題としたい。

6 おわりに

本研究では、AI 検索・要約サービスにおいて、対象が自ら人格的特徴や価値観を明示していない場合であっても、特定の人格的評価語が付与される現象に着目し、二つの事例をもとに観察的な分析を行った。分析の結果、情報発信量や一次情報の厚みに大きな差がある事例においても、AI 要約結果として同一の人格的評価語「誠実」が生成されていることが確認された。一方で、両事例に共通して見られたのは、第三者評価における行為的評価語「丁寧」を中心とした、評価の方向性の一貫性であった。

これらの結果は、人格的評価語の生成が、特定の語彙の頻出や自己申告的な表現によって直接制御されるものではなく、第三者評価や情報全体の配置・構造を要約する過程において生じる可能性を示唆している。ただし、本研究は因果関係の解明を目的とするものではなく、観測された現象の記述にとどまるものである。

本研究で観察されたように、AI 要約は対象の人格的特徴を表す語を付与する一方で、その背後にある判断や迷い、経験といった一次情報を十分に反映しない場合がある。

日本の古典芸能の文脈で用いられてきた「ニン（人・仁）」という概念は、行為や語りの積み重ねの中ににじみ出る人となりを指す言葉であるが、本研究の事例からは、AI 要約によって生成された人格語が、こうした「ニン」に相当するものを必ずしも捉えていない可能性が示唆される。

この点は、AI 要約を無批判に受け入れることの危険性を示すと同時に、人間理解を伴う情報提示における AI 利用の在り方を再考する必要性を示している。

本研究は、生成 AI による人格理解や評価の正確性を論じるものではなく、人格的評価語がどのような条件下で立ち上がるのかという現象を記述的に捉えた探索的研究である。今後は、対象事例の拡張や、異なる AI サービスとの比較を通じて、人格語生成の多様なパターンを検討することが課題として考えられる。

参考文献

- [1] 大西直,高岸智,鎌田理久,山西宏平,神崎雄介,小林拓,徐曉琳,菅原陸,杉浦いぶき,篠崎友悠,河村岳. 松下幸之助再現 AI システムの開発. 言語処理学会第31回年次大会, 2024.
- [2] 中川隼三郎,野元裕樹. ChatGPT が考える日本語ジョークの面白さ: 人間との比較. 言語処理学会第31回年次大会, 2024.
- [3] 菅野祐希,坂野遼. YouTube 動画コメントを用いた視聴者感情の推定と感情処理能力の比較. 言語処理学会第31回年次大会, 2024.
- [4] 藤川直也,伊藤和浩,若宮翔子,荒牧英治. 書き手の孤独感を予測できるか?. 言語処理学会第31回年次大会, 2024.
- [5] Ani Nenkova, Kathleen McKeown. A Survey of Text Summarization Techniques. *Mining Text Data*, Springer, 2012.
- [6] David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. Gender identity and lexical variation in social media. *Journal of Sociolinguistics*, Vol. 18, pp. 135-160, 2014.
- [7] Lucie Flekova and Iryna Gurevych. Personality profiling of fictional characters using sense-level links between lexical resources. *Proceedings of the Workshop on Computational Personality Recognition*, 2015.