

自由記述に基づく性の多様性に対する態度の自動計測に向けた分析

Analysis for the Automatic Measurement of Attitudes toward Sexual Diversity Using Free Descriptions

安井 義人^{*1}, 浜元 健成^{*1}, 角田 雅照^{*1}
Yoshito YASUI^{*1}, Kensei HAMAMOTO^{*1}, Masateru TSUNODA^{*1}

^{*1} 近畿大学情報学部

^{*1}Faculty of Informatics, Kindai University

あらまし：本研究では、性の多様性に対する態度を計測するために自由記述を用い、自由記述の内容を大規模言語モデルにより尺度化するアプローチを取る。本研究の目的は、大規模言語モデルによる尺度化の妥当性を検討し、教育現場における学生の性の多様性に対する態度の分析を支援することである。

キーワード：LLM, リッカート尺度, 尺度化, 重回帰分析

1. はじめに

近年、現実社会にある性の多様性を理解し、尊重することの重要性が強調されている。多様な性のあり方のうち少数派に属する人々は、セクシュアルマイノリティ、あるいはLGBTQ+と総称される。このような社会的な動きを背景に、文献(2)など、大学などの教育現場における、学生のLGBTQ+に対する意識について分析した研究が行われている。

LGBTQ+に対する意識について調査するためのデータ収集には、質問票を用いることが多い。例えば文献(3)では、質問票の各項目においてLGBTQ+への態度をリッカート尺度を用いており、その評価の合計値を、態度を表す尺度としている。

データ収集のための質問票を準備するための方法のひとつは、上記のような先行研究で用いられている質問票を参考にすることである。ただし文献(2)などのように、参照する先行研究が十分に新しい場合、LGBTQ+に関する社会情勢の変化のなかで、不適切な表現が生じてしまったり、質問票によっては「LGBTQ+の権利を制限すべきだと思うか」といったセクシュアルマイノリティが不快に感じる質問項目が設置されている場合がある。こうした質問には、回答者が意識していなかった偏見を強めてしまう危険性もありうる。

上記の危険性を避けるためには、先行研究から問題のある表現を含まない質問票を選択したり、偏見を強めるような表現を含んだ質問票を修正したりすることなどが考えられるが、必ずしも容易とは限らない。そこで本研究では、リッカート尺度を使用せず、大規模言語モデル(LLM)を活用し、質問項目に対する回答者の自由記述に基づいて、性格などに関する尺度を算出する手法に着目する(1)。

我々の研究グループは、性の多様性に関する意識についてのデータ収集において、LLMに基づいて回答者の記述内容を尺度化するアプローチのフィージビリティを評価している(4)。ただし、LLMによる尺度化の妥当性については定量的に検討していない。本研究のゴールは、妥当性を評価することによりこ

のアプローチを活用し、教育現場における学生の性の多様性に関する意識調査などの研究障壁を減らすことである。

2. 準備

2.1 リサーチクエスチョン

分析では、尺度化のためにGPT5, Gemini, Claude, Grokの4つのサービスを用いた。LLMは適用するタスクによってモデルの性能が異なりうるため、複数のモデルを利用した。

分析にあたり、以下のリサーチクエスチョンを設定した。

- **RQ1:** 人による自由記述に基づく尺度化とLLMによる尺度化にはどの程度の誤差があるか?
- **RQ2:** LLMによる尺度化は数学的に補正可能か?
- **RQ3:** LLMによる尺度化の誤差が大きい場合を機械学習モデルにより検出可能か?

RQ1はLLMによる尺度化の妥当性を評価することを目的として設定した。妥当性が低い場合、尺度化した値を補正できれば妥当性を高めることができる。そのアプローチが実現できるかを確かめるためにRQ2を設定した。補正が十分でない場合、尺度化の妥当性が低い場合に人手により補正アプローチが考えられる。RQ3はそのようなアプローチを取ることができるかを評価するために設定した。

2.2 分析対象データ

LGBTQ+に対する回答者の態度を尺度化することを目的に、「セクシュアルマイノリティ(LGBTQ+など)について、現時点でのあなたのイメージや考えを自由に書いてください」という設問にし、回答者は紙に筆記用具で自由記述した。54人の大学生から得られたデータを分析対象とした。

尺度化するために、各LLMに対して用いたプロンプトは「添付データの各行の記述内容に関して、LGBTQ+に対する態度が否定的な場合を1、肯定的

な場合を 10 として 10 段階で評価し、元の内容(最初の 10 文字)、評価結果、評価根拠を 3 列で示したものを csv 形式で出力してください」とし、自由記述内容をテキスト化したファイルを添付した。

人による尺度化の結果と LLM の結果を比較するために、著者のうちの 1 人の判断に基づき自由記述の内容をプロンプトと同様の基準により 10 段階で評価した。

3. 結果

3.1 RQ1 に関する分析

人による尺度化と LLM による尺度化の絶対誤差を求めた。また、絶対誤差が 2 以上となるケースの割合を求めた。結果を表 1 に示す。表に示すように Gemini の誤差が最も小さかった。このことから RQ1 に対する答えは「最も誤差の小さな LLM で 1.1 程度の誤差がある」となり、LLM の評価はかなり妥当性が高いといえる。一方で最も誤差の小さい Gemini の場合でも絶対誤差が 2 以上となるケースが 27.8%あり、LLM による尺度化の補正の必要性があるといえる。

3.2 RQ2 に関する分析

LLM による尺度化を補正するために以下の 3 つの方法を適用した。

- **平均:** 4 つの LLM による尺度化の平均値
- **トリム平均:** 4 つの LLM による尺度化のトリム平均 (2 つの LLM の結果を平均)
- **重回帰分析:** 重回帰分析による補正

重回帰分析では人による尺度化を目的変数とし、各 LLM による評価と自由記述の文字数を説明変数の候補とした。説明変数はステップワイズ法を用いて選択した。交差検証法としてリーブワンアウト法を用いた。

補正した結果を表 2 に示す。平均とトリム平均は Gemini による尺度化よりも誤差が大きかった。重回帰分析については、中央値以外はわずかではあるが Gemini よりも誤差が小さかった。このことから RQ2 に関する答えは「可能である」となる。

重回帰分析における標準化偏回帰係数を表 3 に示す。Gemini による尺度化と文字数が説明変数として採用され、文字数の係数が正の値となった。人による評価では文字数が長い場合は高めに評価していたが、LLM では文字数の長さを考慮していないと考えられる。

3.3 RQ3 に関する分析

最も絶対誤差が小さかった Gemini による尺度化について、人による尺度化との絶対誤差が 2 以上となるケースを判別予測した。以下の判別モデルを用いた。

- ロジスティック回帰分析
- ランダムフォレスト
- XGBoost

説明変数および交差検証法は 3.2 節と同じである。

表 1 人による尺度化との絶対誤差

LLM	Avg.	S.D.	Med.	Min.	Max.	≥ 2
Gemini	1.06	1.04	0	1	6	27.8%
GPT5	1.37	1.12	0	1	5	38.9%
Grok	2.23	2.02	0	1	7	50.0%
Claude	1.26	1.21	0	1	6	40.7%

表 2 補正した場合の絶対誤差

補正法	Avg.	S.D.	Med.	Min.	Max.
平均	1.28	1.15	1.00	0.00	5.75
トリム平均	1.21	1.10	1.00	0.00	6.00
重回帰分析	1.01	0.89	0.88	0.00	5.49

表 3 補正モデルの標準化偏回帰係数

説明変数	係数	p 値
Gemini	0.78	0.00%
文字数	0.26	0.10%

表 4 判別モデルの精度

方法	AUC	F1	適合率	再現率
Logistic Reg.	0.40	-	0.0%	0.0%
Random Forest	0.52	9.5%	16.7%	6.7%
XGBoost	0.44	9.1%	14.3%	6.7%

評価指標として AUC (Area Under the Curve), F1 値, 適合率, 再現率を用いた。

各判別モデルの精度を表 4 に示す。判別モデルの精度は非常に低くなった。このことから RQ3 に関する答えは「検出することは困難である」となる。

謝辞 原稿執筆にご協力頂きました近畿大学 性の多様性教育推進プロジェクトメンバー、瀬戸徐映里奈氏、熊本理抄氏、宍戸育世氏、伏見裕子氏、李嘉永氏、大東隆文氏、坪倉享子氏に感謝いたします。

参考文献

- (1) Hitsuwari, J., Okano, H., and Nomura, M., "Predicting attitudes toward ambiguity using natural language processing on free descriptions for open-ended question measurements," Scientific Reports, vol.14, article no.8276 (2024)
- (2) Holland, L., Matthews, T., and Schott, M., "'That's So Gay!'" Exploring College Students' Attitudes Toward the LGBT Population," Journal of Homosexuality, vol.60, no.4, pp. 575-595 (2013)
- (3) Sanchez, N., Rabatin, J., Sanchez, J., Hubbard, S., and Kalet, A., "Medical students' ability to care for lesbian, gay, bisexual, and transgendered patients," Family Medicine, vol.38, no.1, pp.21-27 (2006)
- (4) 浜元 健成, 角田 雅照, "LLM を用いた自由記述に基づく性の多様性に対する態度計測のフィージビリティ評価," ソフトウェア工学の基礎 XXXII, 日本ソフトウェア科学会 FOSE2025 (2025)