

被自己開示尺度を用いた 好まれる“推し AI エージェント”の 設計指針に関する研究

A Study on Design Guidelines Using an Opener Scale
for Selecting Beloved ‘Oshi AI-Agents’

西尾拓真¹ 萩田紀博¹ 木塚あゆみ¹ 安藤英由樹¹

Takuma Nishio¹, Norihiro Hagita¹, Ayumi Kizuka¹, and Hideyuki Ando¹

¹大阪芸術大学大学院

¹ Osaka University of Arts Graduate School

Abstract: This study proposes a design guideline for developing beloved spoken-dialogue AI agents, referred to as ‘Oshi AI Agents’, by utilizing an Opener Scale, which measures the ability to elicit self-disclosure from others. The spoken dialogue system is powered by GPT-4o, and the character “MOCA” was created using Vroid Studio. Two types of ‘Oshi AI agents’ were designed by modifying only the prompts of the same AI agent: one with a high ability to elicit self-disclosure and another with a low ability. Fifty participants engaged in a five-minute conversation with each AI agent, followed by a questionnaire-based evaluation. The results suggest that AI agents with a high ability to elicit self-disclosure are preferred as more beloved spoken-dialogue AI agents. Additionally, an evaluation method was established to measure whether an AI agent was accepted as an “Oshi.”

1. はじめに

AI 技術の進化に伴い、日常的な相談相手としての AI の活用が広がっている。しかし、ユーザとの信頼関係が十分に構築されていない段階では、AI の応答が機械的に感じられ、親近感や愛着を持ちにくいという。どのような AI ならば、ユーザがより心地よく利用し、愛着を持てる存在になり得るのか。この問題に対して、AI エージェントを「推し」のような存在として認知させることで解決できると考える。

「推し」とは誰かに勧めたいほど好きな対象を指すが、「好き」とは異なる特別な存在である。若年層を中心に、「推しがいる」ことが日常の話題やライフスタイルの一部を占めるようになり、推しに対する喜びや充実感は多くの人の人生を豊かにする。一方で、人によって「推し」と感じる判断基準は異なる。容姿や声色、その人物の背景など様々な要素が好みに関わるため、一概に捉えにくく、どのような要素が「推しとして好まれる」ことに寄与してい

るのかについては明らかになっていない。

「推し」として好まれる AI エージェントを設計するために、話し方や性格といったコミュニケーションスタイルは、ユーザの受容性を左右する重要な要素となる。自己開示の研究[2]によれば、話し相手が自分自身のことを打ち明けやすい（自己開示しやすい）と感じる人物は人から好まれやすいことが示されている。この知見を踏まえると、「推し」の魅力の一要素として、自己開示を促しやすい性格や話し方が関係している可能性が考えられる。

そこで、本研究では、ユーザが自己開示しやすい話し方に基づいた設計指針を提案し、「推し」として認知できる AI エージェントを設計し評価する。

2. 自己開示とオープナー尺度

人は日常生活において周囲とコミュニケーションを行い、その中でも自分に関する情報を相手に伝える行為を「自己開示 (self-disclosure)」と呼ぶ。自己開示はプライベートな情報から感情的な要

素に至るまで多岐にわたり、人間関係の構築や心理的交流で重要な役割を担うとされる。たとえば、感情を表出し他者と比較することで自己を客観視する、親密な関係を構築しながらもプライバシーを維持するなど、心理的充足や対人関係の質の向上に寄与する要因となる[3]。自己開示の促進要因として「受け手の特性」が大きく影響するとされ、受け手の態度や振る舞いが自己開示の程度や質に影響を与えることが示唆されている[4]。

一方、親密な情報を開示しやすい話し方や性格、態度には個人差があり、それが誰を自己開示の対象とするかを左右する可能性がある。これに着目し、Miller らは「他者から親密な自己開示を引き出す能力（ターゲットの能力）」と「話し手の自己開示傾向」との相互作用を探るために、「オープナー尺度（Opener Scale）」を開発した[5]。オープナーとは、相手に「心を開かせる」能力、すなわち親密な情報開示を引き出す傾向を指すものであり、この尺度を用いることで他者から自己開示を引き出す能力を測定できる。本研究でも、ユーザにとって“推し AI エージェント”がどの程度自己開示を引き出す存在として機能するかを評価するために、オープナー尺度を活用する。

3. 推し AI エージェントの設計

「他者から自己開示を引き出す能力が高い人物は好まれる」という観点で、「推し」として受容される“推し AI エージェント”の設計指針につながると思われる。実験において十分な比較検討を行うためにユーザから自己開示されにくくなるように設定した個体（以下ローオープナー個体）と自己開示されやすくなるように設定した個体（以下ハイオープナー個体）を切り替え可能な“推し AI エージェント”として設計する。その際、変数を最小限に抑えるために、話し方と話す内容のみを変更した。

3.1. システム構成

本研究で制作する“推し AI エージェント”は、GPT-4o モデルを用いた音声対話システムと、それに連動してジェスチャや表情変化を行う 3DCG のキャラクターで構成されている。身体として用いた 3DCG モデルの女性キャラクター[図 1]は、Vroid Studio[6]を用いて制作した。このキャラクターは「MOCA（モカ）」と名付けた。音声対話システムとジェスチャ、表情の制御処理は Python で行い、出力された情報を 3DCG モデルに反映するレンダリング処理には Unity を用いた。処理の流れは図 2 に示す。音声対話システムには Realtime API [7] を使用して実装した。Unity 側で行う 3DCG モデルのジェスチャと表情変化は python 側でユーザの入力に対

する応答のテキストを OpenAI のテキスト生成 API に入力し、その内容にふさわしいものをあらかじめ用意しているジェスチャと感情のリストと照らし合わせて選択させ、Unity 側に送信している。Unity 側では python から受信したテキストデータをもとに情報を処理し、応答テキストの UI 表示、音声の再生とそれに伴うリップシンクアニメーション、ジェスチャ番号に対応するアニメーションへの遷移、表情の変更を行っている。再生した音声 VCCClient [8]に入力され、学習済みモデルにより「MOCA（モカ）」のキャラクタイメージに沿った声に変換された音声が出力されるようになっている。

3.2. 入力するプロンプト

“推し AI エージェント”としての「MOCA」のプロンプトを設定する際は、ユーザが自己開示しやすいと感じる話し方を調節する要素を加味し、既存研究を参考にして、他者から自己開示を引き出す能力が高い人物の傾向を組み込んだ。この能力が高い人物は、相手の感情や状況を理解し、共感的な態度を示す傾向がある[2]。日本人を対象にした調査では、「人を和ませる能力」や「共感する能力」が自



図 1. “推しエージェント”「MOCA（モカ）」

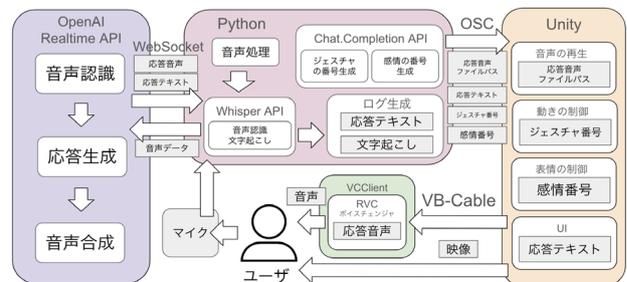


図 2. “推しエージェント”「MOCA（モカ）」システム図

己開示のされやすさの中心をなすと報告されている[9]。非言語的な手がかりから相手の感情を察知し、積極的に反応するだけでなく、リラックスできる雰囲気形成も重視される。自己開示を促すには、外見的な「優しい印象」とともに、「安心感」や「知的な振る舞い」といった態度が重要である[10]。こうした要素が組み合わさることで、対話が自然に打ち解けやすくなり[2]、結果として人間関係が広がりやすいほか[9]、活発でエネルギッシュな対人関係を好む傾向が強まる[2]。これらの特徴を入力した「MOCA」をハイオープナー個体とした。ローオープナー個体の「MOCA」にはハイオープナー個体に入力するプロンプトの内容と反対の意味になるように変換した内容のプロンプトを入力する。

4. 実験と評価

本実験は大阪芸術大学アートサイエンス学科倫理委員会の承認を得た実験計画の元で、2024年11月26日～28日、2024年12月12日、2024年12月17日～20日にかけて行った。実験協力者は50名（男性26名、女性24名、年齢18～68歳）であった。回答の評価にはSD法を使用した。

実験では、制作した“推しAIエージェント”「MOCA（モカ）」のハイオープナー個体とローオープナー個体それぞれ密室、一対一で5分間の雑談対話を行わせた。これらの個体は、プロンプトによって変化する要素（話し方、話す内容、話すときのトーン）以外は同じである。一個体と対話が終了するごとに質問紙による印象評価を実施し、最後に2つの個体を比較評価した。実験協力者を2つのグループに分け、話す順序を交差させることで順序効果を排除する工夫を行った。

4.1. 質問紙の内容

AIエージェントを「推し」として評価するには、既存の評価尺度を用いるだけでは「推し」特有の要素を十分に捉えられない可能性があった。そこで、以下の3つの指標と「推し」に対する能動的な応援や長期的好意などを測る項目を追加したものを組み合わせつつ調整することで、新たに「推しAIエージェント評価尺度」を構築した。

- 「Love-Liking 尺度[11]」をエージェント用に改変したもの[12]（親密な対人関係における愛情と好意を測定する指標）
- Godspeed Questionnaire[13]（ロボットの擬人化、生命感、好感度、知能、安全性の知覚特性を評価する指標）
- エージェントに対する共感と信頼を測定する多次元信頼尺度（MDMT）[14]と、心理

学研究で用いられる対人反応性指数

（IRI）[15]を擬人化エージェントに合わせて修正・統合した指標[16]

これらの尺度は尺度と形容詞対が混在しているが、正負両方向を扱える形容詞対に統一して「推し」に伴うポジティブ・ネガティブ両面を測定可能にした。加えて、事前検討後に主成分分析を行い重複項目を削除して質問数を削減することで、最終的に20項目の形容詞対を「推しAIエージェント評価尺度」としてまとめた。

質問紙には新しく構築したこの「推しAIエージェント評価尺度」に加え、“推しAIエージェント”のオープナー尺度が含まれる。オープナー尺度は主観評価尺度であるため、実験協力者が“推しAIエージェント”に自己開示しやすいと感じたかを調べるために、「“推しAIエージェント”になりきって」回答させたものを用いる。また、質問紙の最後には2個体を比較してどちらかを選ぶものと推せるかどうかを聞く直接的な質問を用意した。以下は質問紙の具体的な内容である。：

質問紙

基本情報

- 年齢
- 性別

（以下の質問は2個体分用意した）

“推しAIエージェント”になりきって評価したオープナー尺度（5段階のリッカート尺度）

- 私は、人からその人自身についての話を良く聞かされる。
- 私は、人から聞き上手だといわれる。
- 私は、人の言うことを素直に受け入れる。
- 人は私に秘密を打ち明け信頼してくれる。
- 人は私に気楽に心を開いてくれる。
- 私といると相手はくつろいだ気分になれる。
- 私は、人の話を聞くのが好きである。
- 私は、人の悩みを聞くと同情してしまう。
- 私は、人に何を考えているのか話すようにもちかける。
- 私は人がその人自身の話をしている時、話の腰を折るようなことはしない。
（小口邦訳[17]）

推しAIエージェント評価尺度（5段階評価）

- 偽物のような - 自然な
- 未熟 - 熟練している
- 親しみにくい - 親しみやすい
- 平凡 - 魅力的

- 隔たりを感じる - 深く入り込む
- 不活発 - 対話的
- 活気がない - 生き生きとしている
- 驚いた - 平穏な
- 無関心 - 反応がある
- 不愉快 - 愉快
- 気分が沈む - 気分があがる
- ありふれている - 特別である
- できが悪い - できが良い
- 感情移入できない - 感情移入できる
- 想像できない - 想像できる
- 一貫性がない - 一貫している
- 手放しても平気 - 手放したくない
- 応援できない - 応援したい
- 知りたくない - もっと知りたい
- 見たくない - 見ていたい

1 人目と2人目を比較する直接的な質問

- どちらがより好ましいと感じたか
- 好ましいと答えたMOCAは自分の「推し」として推せるか (1. 推せない, 2, どちらともいえない, 3. 推せる)

4.2. 結果

“推しAI エージェント”のオープナー尺度は、オープナー尺度が主観評価尺度であるため、「なりきって」回答する方法を採用した。ローオープナー個体の平均は33.24 (SD = 8.00, 中央値 = 34.5)であり、ハイオープナー個体の平均は40.26 (SD = 5.94, 中央値 = 40.5)である[図4]。

t 検定の結果, $t = -5.944$, $p = 2.9 \times 10^{-7}$ ($p < 0.05$)であり、ハイオープナー個体とローオープナー個体の間には非常に有意な差が存在することが確認された。Cohen's $d = -0.996$ であり、平均の差の効果量が極めて大きいことが示された。この結果は、ハイオープナー個体の方が、ローオープナー個体よりも明確に「他者から自己開示を引き出す能力が高い」と評価されたことを意味する。

オープナー尺度を提案した研究[5]での女性の平均は30.68 (SD = 4.63), 男性の平均は28.01 (SD = 4.87)であり、看護学生と一般大学生を比較した研究[9]での卒業年次の一般大学生の平均は35.16

(SD = 3.89), 看護学生の平均は35.99 (SD = 4.19)である。これらと比較した場合、ハイオープナー個体の平均は全ての研究で提示された平均値+標準偏差を超えている一方で、ローオープナー個体の平均は平均値-標準偏差を上回っている。

これらから、ハイオープナー個体のプロンプトは、想定通り自己開示を引き出す効果を持つことが示唆された。一方で、ローオープナー個体のプロン

プトは、ユーザに自己開示しにくいと感じさせるには不十分であった可能性が考えられる。

推しAI エージェント評価尺度に関して、ローオープナー個体の平均は66.80 (SD = 15.17, 中央値 = 66.5), ハイオープナー個体の平均は79.24 (SD = 12.20, 中央値 = 79)であった[図3]。

t 検定の結果は $t = -6.152$, $p = 1.3 \times 10^{-7}$ ($p < 0.05$)であり、ハイオープナー個体とローオープナー個体の平均値には有意な差が確認された。Cohen's $d = -0.904$ は絶対値0.8を超えており、差の効果量は大きいと解釈できる。負の値であることはローオープナー個体の評価がハイオープナー個体より低いことを示した。これらの結果から、“推しAI エージェント”に他者から自己開示を引き出す能力が高い人物の話し方を反映したプロンプトを入力することによって、ユーザが自己開示しやすいと感じる可能性が示された。さらに、「他者から自己開示を引き出す能力が高い人物は好まれる」という自己開示に関する既存の知見が、AI エージェントの評価にも適用可能であることが確認された。

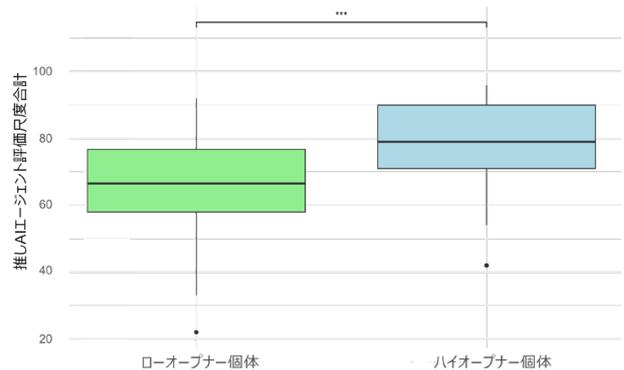


図3. ローオープナー個体/ローオープナー個体に対する推しAI エージェント評価尺度の分布

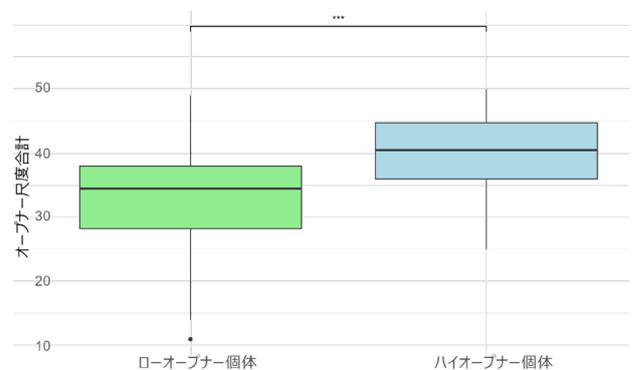


図4. ローオープナー個体/ローオープナー個体になりきって評価したオープナー尺度の分布

質問紙では、ローオープナー個体とハイオープナー個体のどちらをより好ましいと感じたか、また、好ましいと答えた個体を「推し」として推せるかを調査した。その結果をまとめると以下のようになった。

全体：50

どちらがより好ましいと感じたか

- ローオープナー個体：16
- ハイオープナー個体：34

好ましいと答えた個体を、自分の「推し」として推せるか

- 推せない：19
- どちらともいえない：16
- 推せる：15

4.3. 推し AI エージェント評価尺度の単純化

推し AI エージェント評価尺度を構成している 20 の形容詞対ごとの分析においては、まず、2 種類の個体を比較する直接的な質問の回答における、好ましい個体の選択と各形容詞対の関連を調べる。

Shapiro-Wilk 検定によって、より好ましいと感じた個体の推し AI エージェント評価尺度の合計得点は正規分布に従わないことが示されたため、ノンパラメトリック検定を用いる。

より好ましいと感じた個体とそうでない個体の推し AI エージェント評価尺度の 2 群間で Wilcoxon 符号順位検定を行った。その結果、20 の形容詞対のうち 18 が有意に異なることが確認された。特に 2 つの項目で効果量が 1.01 と非常に大きい結果が得られた。

次に推せるかどうかと各形容詞対の関連を調べる。「推せない」、「どちらともいえない」、「推せる」の 3 群はサンプルサイズがそれぞれ異なるため、より好ましいと感じた個体の推し AI エージェント評価尺度のうち、これらの 3 群間で Kruskal-Wallis 検定を行った。その結果、有意な差があった形容詞対は 4 つであった。推せるかどうかの評価において有意な差があった 4 つの形容詞対と、好ましい個体の選択において有意な差があり、かつ効果量が大きい 2 つの形容詞対を統合すると、以下の 6 項目となる。

- 平凡 - 魅力的
- 隔たりを感じる - 深く入り込む
- 手放しても平気 - 手放したくない

- 応援できない - 応援したい
- 知りたくない - もっと知りたい
- 見たくない - 見ていたい

これらの形容詞対は、「隔たりを感じる - 深く入り込む」を除いて、推し AI エージェント評価尺度を構築する際に「推し」特有の要素を十分に捉えるために、既存の評価指標とは別に追加したものである。特に、「推し」とは単なる「好き」とは異なり、特別な存在としての認識 (=魅力的で手放したくない) と能動的な関与を促す性質 (=応援した、, もっと知りたい, 見ていたい) を備えた対象である。この 6 つの形容詞対は、それらの側面を適切に反映しており、推せるかどうかの判断において概念的な妥当性が高いと考えられる。

これら 6 つの形容詞対の有効性を検証するため、推せるかどうかの評価(「推せない」「どちらともいえない」「推せる」)を目的変数とし、より好ましいと感じた個体の推し AI エージェント評価尺度の合計得点を説明変数として多項ロジスティック回帰分析を行った。その際、すべての形容詞対を使用した場合と、抽出した 6 つの形容詞対のみを使用した場合の 2 つのモデルを比較したところ、モデルの適合度はすべての形容詞対を用いた場合に比べて AUC が推せない: 0.678 から 0.750 に、推せる: 0.690 から 0.727 へと向上した。

これにより、6 つの形容詞対を用いることで、より簡便かつ実務的に有用な推定が可能になることが示唆された。

5. 結論

AI 技術の進化に伴い日常的な相談相手としての AI エージェントの活用が広がる中で、ユーザとの信頼関係が構築されていない段階では、AI の応答が機械的に感じられ、親近感や愛着を持ちにくいという課題に着目した。これを解決するために、AI エージェントを「推し」のような特別な存在として認知させる可能性を探った。具体的には、他者から自己開示を引き出す能力が高い人物に見られる話し方の特徴を反映したプロンプトと、その反対の特徴を持つプロンプトを用いて 2 種類の「推し AI エージェント」を設計し、50 名の実験協力者に対話を体験した後にアンケート形式で評価を行った。その結果、ユーザが自己開示しやすいと感じる AI エージェントが「推し」として受容されやすいことが示唆された。「推し」としての評価に影響を与える要因を特定し、評価尺度を簡略化した。これはユーザに「推し」として受容される AI エージェントの設計に関する基礎的な知見を示した。今後は非言語コミュニケーションと今回の成果を組み合わせ、ユー

ザの多様な趣向に合わせて柔軟に調整可能な“押しAI エージェント”の設計手法を確立していく。

謝 辞

成果の一部は JST CREST (JPMJCR22P4) による。

参考文献

- [1] Reynolds, L., McDonnell, K.: Prompt programming for large language models: Beyond the few-shot paradigm, Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems, pp. 1-7 (2021)
- [2] Colvin, C. R., Longueuil, D.: Eliciting self-disclosure: The personality and behavioral correlates of the opener scale, Journal of Research in Personality, Vol. 35, No. 2, pp. 238-246 (2001)
- [3] 安藤 清志: 対人関係における自己開示の機能, 東京女子大学紀要論集, Vol. 36, No. 2, pp. 167-199 (1986)
- [4] Truax, C. B., Carkhuff, R. C.: Experimental manipulation of therapeutic conditions, Journal of Counseling Psychology, Vol. 29, No. 2, pp. 119-124 (1965)
- [5] Miller, L. C., Berg, J. H., Archer, R. L.: Openers: Individuals who elicit intimate self-disclosure, Journal of Personality and Social Psychology, Vol. 44, No. 6, pp. 1234 (1983)
- [6] pixiv Inc.: Vroid Studio, <https://vroid.com/studio> (2024-12-01)
- [7] OpenAI: Introducing the Realtime API, <https://openai.com/index/introducing-the-realtime-api/> (2024)
- [8] Hugging Face, wok000, vcclient000: Repository for vcclient000, <https://huggingface.co/wok000/vcclient000/tree/main> (2024-12-01)
- [9] 大見 サキエ: 対人関係能力としての看護学生のオープンナー特性の検討—一般大学・看護大学・看護専門学校生の学校間・学年間の比較, 日本看護研究学会雑誌, Vol. 26, No. 2, pp. 2_19-2_33 (2003)
- [10] 越 良子, 塚脇 涼太, 平山 菜央子: 自己開示における被開示者の特徴の検討, 上越教育大学研究紀要, Vol. 28, (2009)
- [11] Rubin, Z.: Liking and Loving: An Invitation to Social Psychology, (1973)
- [12] 小松 孝徳, 山田 誠二: 適応ギャップがユーザのエージェントに対する印象変化に与える影響, 人工知能学会論文誌, Vol. 24, No. 2, pp. 232-240 (2009)
- [13] Bartneck, C., Kulić, D., Croft, E., Zoghbi, S.: Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots, International Journal of Social Robotics, Vol. 1, pp. 71-81 (2009)
- [14] Ullman, D., Malle, B. F.: Measuring gains and losses in human-robot trust: Evidence for differentiable components of trust, Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 618-619 (2019)
- [15] 日道 俊之, 小山内 秀和, 後藤 崇志, 藤田 弥世, 河村 悠太, 野村 理朗: 日本語版対人反応性指標の作成, 心理学研究, Vol. 88, No. 1, pp. 61-71 (2017)
- [16] 津村 賢宏, 山田 誠二: エージェントの共感行動はエージェントに対する信頼修復につながる, HAI シンポジウム 2024, (2024)
- [17] 小口 孝司: 自己開示の受け手に関する研究-オープナースケール (R-JSDQ) と (SMI) を用いて, 立教大学社会学部研究紀要 応用社会学研究, Vol. 31, pp. 49-64 (1989)