

# LLMを用いた対話型芸術鑑賞システムの運用実験と対話ログ分析 ～ 三つの異なる鑑賞環境を対象として ～

鳥垣 耀平<sup>†</sup> 松田 裕貴<sup>†,††</sup>

<sup>†</sup> 岡山大学 〒700-0082 岡山県岡山市北区津島中3丁目1-1

<sup>††</sup> 理化学研究所革新知能統合研究センター 〒103-0027 東京都中央区日本橋1-4-1

E-mail: <sup>†</sup>yohei.torigaki@cocolab.jp, yukimat@okayama-u.ac.jp

あらまし 美術館や博物館における鑑賞体験を向上させるためには、鑑賞者の関心や理解に応じて解説内容を調整する対話的な鑑賞支援が重要である。しかし、学芸員やミュージアムエデュケーターの人員不足により、人手に依存した支援には限界がある。本研究では、大規模言語モデル（LLM）を組み込んだ対話型鑑賞システムを用い、三つの異なる鑑賞環境における運用実験を行った。鑑賞者とAIの会話ログを分析することで、対話型鑑賞支援が鑑賞体験に与える影響と、本手法の可能性と課題について考察する。

キーワード 美術館、大規模言語モデル、AI、音声対話システム、対話ログ分析

## Field Experiments and Dialogue Log Analysis of an LLM-based Interactive Art Appreciation System: A Case Study in Three Viewing Environments

Yohei TORIGAKI<sup>†</sup> and Yuki MATSUDA<sup>†,††</sup>

<sup>†</sup> Okayama University Tsushima-naka 3-1-1, Kita-ku, Okayama, Okayama 700-0082, Japan

<sup>††</sup> RIKEN Center for Advanced Intelligence Project Nihonbashi 1-4-1, Chuo-ku, Tokyo 103-0027, Japan

E-mail: <sup>†</sup>yohei.torigaki@cocolab.jp, yukimat@okayama-u.ac.jp

**Abstract** To enhance the quality of art appreciation experiences in museums, interactive appreciation support that adapts explanations according to viewers' interests and levels of understanding is essential. However, due to the shortage of curators and museum educators, appreciation support that relies on human resources faces inherent limitations. In this study, we conduct operational experiments using an LLM-based interactive art appreciation system across three different art viewing environments. By analyzing dialogue logs between viewers and the AI system, we examine the impact of interactive appreciation support on viewers' experiences and discuss the potential and challenges of the proposed approach.

**Key words** Museums, Large Language Models, AI, Spoken Dialogue System, Dialogue Log Analysis

### 1. はじめに

美術館や博物館における鑑賞体験の質は、鑑賞者の関心や理解の深まりに大きく依存しており、鑑賞者の反応に応じて解説内容を柔軟に調整する対話的な鑑賞支援が重要である。こうした役割を担うミュージアムエデュケーターは、鑑賞者との対話や非言語的反応を手がかりに心理状態を推定し、適切な問いかけや解説を行うことで鑑賞体験を支援している。しかし、学芸員・ミュージアムエデュケーターの人員は依然として不足しており、2008年から2018年にかけて学芸員数の増加は約20%にとどまる（1館あたりの学芸員数は4人に満たず、さらに相当博物館・類似博物館を含めると約1.5人に過ぎない）など、人的資源に依存した鑑賞支援には限界がある[1]。一方、既存の音声

ガイドシステムは、鑑賞者の状態を考慮しない一方的な情報提示が中心であり、人による対話的な鑑賞支援を十分に代替できていない。このように、鑑賞者の心理状態や行動に応じて動的に情報提示を行う鑑賞支援の仕組みは未だ確立されておらず、対話的かつ状況適応的な鑑賞支援を実現する技術的アプローチが求められている。

本研究では、鑑賞者の心理状態や視線といった非言語的コミュニケーションの活用を視野に入れつつ、その情報に基づいて対話内容を動的に調整することで、鑑賞体験を向上させるための対話型鑑賞システムの研究に取り組んでいる。本稿ではその実現に向け、大規模言語モデル（LLM）を組み込んだ対話型鑑賞システムを用い、三つの異なる鑑賞環境（展示会における

絵画鑑賞、歴史文化施設における現代アート鑑賞、大学講義における PC 画面上での絵画鑑賞)においてシステム運用実験を行った。実験中に取得した鑑賞者と AI との会話ログデータを分析することで、LLM を用いた対話型鑑賞システムが鑑賞体験に与える影響を検証し、本手法の可能性を考察する。

## 2. 関連研究

### 2.1 美術館におけるシステムの研究

美術館来館者の理解を助けるためのツールとして、近年では貸出型端末やスマートフォンアプリでのナビゲーション(音声ガイド)が国内外問わず広く用いられているが、広く普及しているシステムの多くは静的なガイドランスにとどまっている。一方で、美術鑑賞支援を鑑賞者に合わせて柔軟に変化させる重要性も指摘されてきた。Sun らは、個人の学習スタイルに合わせて美術館ガイドの情報提示方法を変更することで、認知コストの軽減や学習効果が高まることを明らかにしている[2]など、個人適応的な鑑賞支援の有効性を示している。

近年の LLM の急速な進化・普及に伴い、美術館鑑賞においても LLM を活用し、個人に合わせたガイドランスを対話的に提供する試みに注目が集まっている。Gao らは、スマートフォン上で動作する生成 AI チャットボットにより、作品検出に基づく解説提示や来館者の質問応答を通じて鑑賞体験を支援する手法を提案、既存アプリと比較してチャットボットが作品への関与や情報性を高める可能性を報告している[3]。Trichopoulos らは、GPT-4 を用いた美術館向け対話型ガイドシステムにより、来館者に対してパーソナライズされた解説や質疑応答を提供する手法を提案している[4]。また、Vasic らは、3D バーチャル美術館ツアーに LLM を組み込み、鑑賞者の自然言語による興味・希望に応じて鑑賞ルートや説明内容を動的にパーソナライズする手法を提案している[5]。Lewis らは、ChatGPT を教師モデルとした知識蒸留により、小規模モデルでありつつ応答品質を維持した仮想美術館ツアーガイド用の対話エージェントを構築している[6]。Wang らは、LLM を用いたマルチモーダル対話により、仮想美術館ツアーにおけるガイドランス支援システムを提案しており、利用者の状況に応じた対話的支援が没入感や理解促進に寄与することを示している[7]。さらに Rachidi らは、芸術分野の会話エージェント学習に向けて、文脈認識と個人特性(年齢・性別・民族性・知識レベル等)を考慮した合成対話データセットを設計・生成・評価し、作品情報の前処理などによるハルシネーション低減も図っている[8]。Ozaki らは大規模視覚言語モデル(LVLM)における芸術作品の多言語説明のためのデータセット構築と説明生成能力の評価を行っている[9]。

このような鑑賞支援システムの有効性を検証するためには、美術館における鑑賞体験を定量的に取得し評価できることが望ましい。Liu らは、美術館の評価基準を確立し、その評価基準を用いて美術館の面白さと来場者の満足度の関係性を導出する研究を行っている[10]。また著者らは、美術館における鑑賞者の感情や満足度などの心理状態推定に向けた鑑賞中の仕草や音声のデータセット構築を行っている[11]ほか、構築したデータ

セットを用いた美術鑑賞中の心理状態推定モデルを構築してきた[12],[13]。

### 2.2 本研究の立ち位置

これまで、美術館における美術作品鑑賞システムとして、鑑賞者の自然言語による文脈や、アンケートなどによる鑑賞者のパーソナライズにより、鑑賞者の興味・関心に応じて動的にガイドランス手法を変更する手法が提案されている。近年では大規模言語モデル(LLM)の発展により、鑑賞者に対して柔軟に解説を変更する手法についても提案されている。しかし、これらの手法においては、基本的に言語情報を基にしており、鑑賞中に生じる非言語的なコミュニケーションについては十分に考慮されていない。そこで本稿では、複数の鑑賞環境において、LLM を活用した対話型鑑賞システムを用いた運用実験と対話ログ解析を通じて、これらの可能性や課題について考察する。

## 3. 対話型鑑賞システムの運用実験

著者らは 2025 年 7 月～12 月にかけて、以下に示す三つの異なる実験設定(場所・対象物・対象者・シナリオ)において大規模言語モデル(LLM)を基軸とした対話型鑑賞システムを用いた鑑賞実験を実施してきた。なお、各実験における LLM の役割設定は実験環境に応じて調整している。

- (1) 美術展覧会における絵画の対話型鑑賞
- (2) 歴史文化施設における現代アートの対話型鑑賞
- (3) 大学内における PC を用いた絵画の対話型鑑賞

本稿では、各実験で用いたシステムや実験設定について述べるとともに、対話型鑑賞システムと鑑賞者の対話ログデータの分析を行う。なお、本研究は岡山大学「人を対象とする研究倫理審査委員会」の承認を受けて実施した(承認番号:自然2024-13)。

### 3.1 実験(1) 美術展覧会における絵画の対話型鑑賞

奈良県立美術館で行われた美術展覧会「わたしたちのびじゅつかん～きて・みて・はなして→たいけんする美術展」<sup>(注1)</sup>において、独自開発したシステムを利用した絵画鑑賞実験を行った。本実験は 2025 年 7 月 27 日(日)に同展覧会にて 10 名の参加者を対象に行った。実験対象作品として、絹谷幸二による《チェスキーニ氏の肖像》<sup>(注2)</sup>と田中敦子による《90E》<sup>(注3)</sup>を利用した。実験の様子は図 1 に示すとおりである。

本実験では、著者らが独自に開発した大規模言語モデル(LLM)を基軸とした対話型鑑賞システムを用いた。システム構成を図 2 に示す。主に、Apple の Speech Framework<sup>(注4)</sup>にて提供される Speech-To-Text (STT) と Text-To-Speech (TTS) の機能を用いた音声による入出力を行い、バックエンドで OpenAI 社の GPT API (gpt-4o)<sup>(注5)</sup>による言語生成を行う構成とした。音声に基づく言語的な情報に加えて、鑑賞者のリアルタイムな「感

(注1): <https://www.pref.nara.jp/69196.htm>

(注2): <https://kinutani.jp/artwork/1986/01/post-429.html>

(注3): <https://artplatform.go.jp/ja/collections/W1154466>

(注4): <https://developer.apple.com/documentation/speech>

(注5): <https://platform.openai.com/docs/models/gpt-4o>



図 1: 実験の様子 (実験 1)

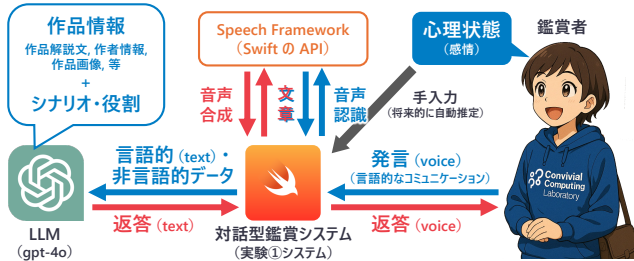


図 2: 対話型鑑賞システム (実験 1) のシステム構成

情」を実験実施者が手動入力できる構成としており、非言語的な情報を考慮した上で応答するプロンプトを与えた。感情推定に関しては別途検討を進めており [13]～[15]、将来的には本システムとの統合を想定している。なお、LLM には「鑑賞者に対して対話型鑑賞を提供する学芸員」という設定と作品情報（画像や解説文、メタデータ）をプロンプトとして与えた。対話型鑑賞システムの利用（1 セッション）の流れは以下の通りである。

- (1) 利用者が作品の前に移動
- (2) システム初期化、相互に自己紹介した後、鑑賞を開始する
- (3) 作品について一定時間システムとの対話を行う（3～5 分）
- (4) 時間をトリガにシステムが作品解説し、会話を終了する
- (5) 鑑賞を終了する

### 3.2 実験 (2) 歴史文化施設における現代アートの対話型鑑賞

岡山芸術交流<sup>(注6)</sup>において、ブリッジウェル社の開発する ToyTalk<sup>(注7)</sup>を利用した現代アート鑑賞実験を行った。本実験は 2025 年 9 月 26 日（金）～11 月 24 日（月）の岡山芸術交流会期を通して実施し、実験協力を希望した来訪者を対象に行った。なお、約 2 ヶ月間の会期中に 149 人のシステム利用があった。

本実験では、ブリッジウェル社の開発する対話型 AI システムである ToyTalk<sup>(注7)</sup>を用いた。画面例を図 3 に示す。本システムも、音声の入出力を受け付ける LLM を利用した対話型鑑賞システムとなっている。本実験に特有な点として、システムを一般公開し不特定多数の一般人を対象としたため、鑑賞者がどの作品を鑑賞しているのかが利用開始時にはわからないという制約がある。そこで、LLM には「仮想的に視覚情報を持たない設定」をプロンプトで与えることによって、どのような作品を見ていてどう感じたのか、といった状況の共有から会話を始めるシナリオを採用することとした。対話型鑑賞システムの利用（1 セッション）の流れは以下の通りである。

(注6) : <https://www.okayamaartsummit.jp/>

(注7) : <https://toytalk.ai/>



図 3: 対話型鑑賞システム (実験 2) の画面例, ToyTalk<sup>(注7)</sup>を利用

- (1) 利用者が特定の作品の前に移動する
- (2) ToyTalk にアクセスし鑑賞を開始する
- (3) 作品について一定時間システムとの対話を行う（3～5 分）
- (4) 時間をトリガに会話を閉じる
- (5) 鑑賞を終了する

### 3.3 実験 (3) 大学内における PC を用いた絵画の対話型鑑賞

中京大学現代社会学部の協力の下、独自開発したシステムを利用し PC を用いた絵画鑑賞実験を行った。本実験は 12 月 17 日（水）に実施し、実験参加に同意した 17 名を対象とした。

本実験では、実験 1 のシステムを改良し著者らが独自に開発した LLM を用いる対話型鑑賞システム「MuseMate」を使用した。システム構成を図 4 に示す。主に、OpenAI 社の提供する Realtime API (gpt-realtime)<sup>(注8)</sup>を中核とした対話生成部、視線・心理状態の入力部、それらをつなぎ合わせるインタフェース部から構成される。画面例を図 5 に示す。鑑賞する作品がシステム画面上に表示されており、視線（注目点）については鑑賞者が自身で気になる箇所をクリックすることで LLM に注目地点を共有できるようにした。心理状態については、実験 1 と同様に手入力できるようにした。なお、視線・心理状態推定ともに、今後は推定モデルと連結し自動化を想定している。LLM には「美術館の対話型鑑賞ナビゲータ」という設定と作品情報（画像や解説文、メタデータ）をプロンプトとして与えた。また、視線や心理状態の情報を考慮して、対話内容に反映したうえで応答を生成するように指定している。対話型鑑賞システムの利用（1 セッション）の流れは以下の通りである。

- (1) MuseMate にアクセスし鑑賞する作品を選択し、画面に表示された作品について鑑賞を開始する
- (2) 作品について一定時間システムとの対話を行う（3～5 分）
- (3) 会話終了または実験時間超過に基づき会話を終了する
- (4) 鑑賞を終了する

(注8) : <https://platform.openai.com/docs/models/gpt-realtime>

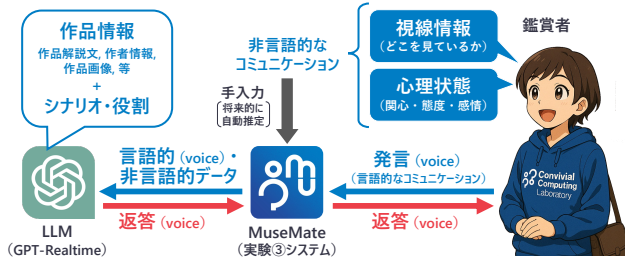


図4: 対話型鑑賞システム（実験3）のシステム構成

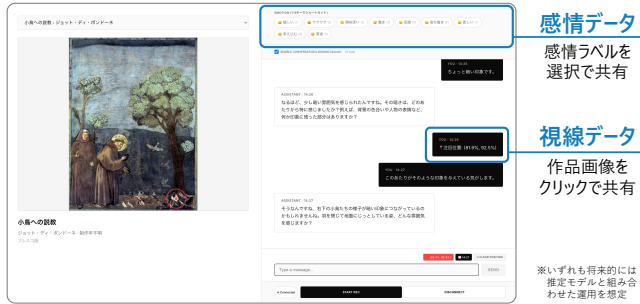


図5: 対話型鑑賞システム（実験3）の画面例

## 4. 実験結果の分析・考察

### 4.1 反応遅延時間と発話文字数の関係分析

実験1, 実験2, および実験3におけるユーザの応答遅延時間（Response Latency）と、それに対するユーザの平均発話文字数（Utterance Length）の関係について分析結果を述べる。その結果をそれぞれ図6, 図7, 図8に示す。

まず、AIの問いかけに対してユーザが応答するまでの時間である応答遅延時間について比較を行う。実験1および実験3は、中央値でそれぞれ約1.5秒、約1.7秒、平均値においても4秒台前半（実験1: 4.1秒, 実験3: 4.3秒）となっていた。これに対し、実験2は中央値で約13.2秒、平均値では18.1秒を要しており、他の二つの実験条件と比較して顕著な遅延が見られた。

次にユーザの発話文字数に着目すると、最も応答速度の速い実験1において、ユーザの平均発話文字数は26.8文字（中央値18.5文字）と最大値を記録した。次いで実験3が平均22.6文字（中央値17.5文字）となり、応答遅延が最も大きい実験2では、平均18.0文字（中央値13.0文字）と最も短い結果となった。

全体を通して、多くのユーザの発言は応答遅延時間が短く、発言量が少ない傾向にある。すなわち、AIの問いかけに対して「即座に」「端的に」回答している場合が多い。一方で、応答遅延時間が長いパターンや、発言量が多いパターンも一定数存在しており、発話文字数が多い場合には応答遅延時間が短くなる傾向、発話文字数が少ない場合には応答までに時間を要する傾向が確認された。これらは必ずしも矛盾するものではなく、鑑賞者の認知状態や発話方略の違いを反映した結果と考えられる。

この結果から、ユーザが即時的に回答できる質問に対しては発言量が多くなり、一方でユーザが即時的に回答することが難しい質問に対しては発言量が少なくなる傾向にあると言える。即時的に量の多い回答をしている場合は鑑賞への高い関与が示

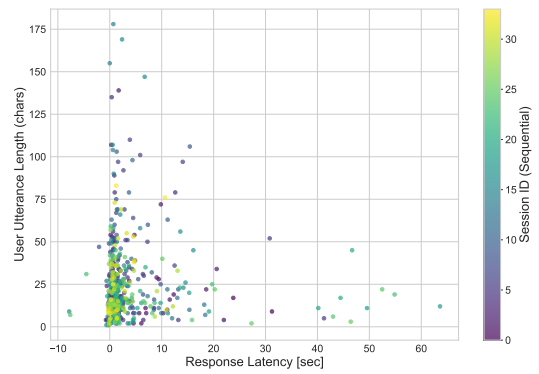


図6: 応答遅延時間と発言量（実験1, 奈良県立美術館）

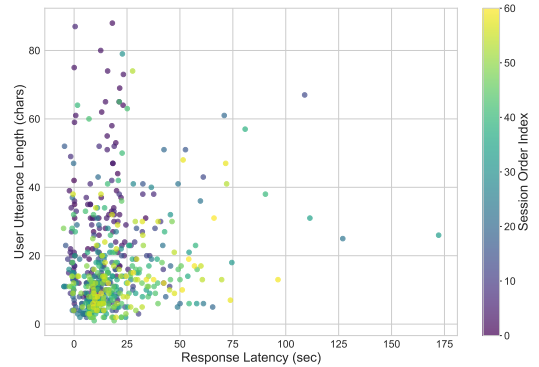


図7: 応答遅延時間と発言量（実験2, 岡山芸術交流）

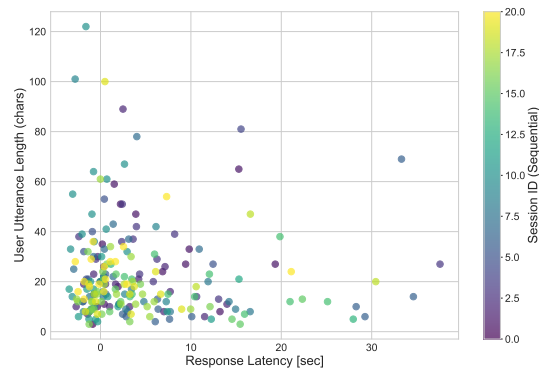


図8: 応答遅延時間と発言量（実験3, 大学内）

唆され、逆に長考した上で端的に回答している場合は鑑賞への関与の在り方が異なる可能性があると考えられる。これらの傾向から、美術鑑賞者の興味・関心の推定にこれら二つの指標を取り入れることによって推定性能の向上が期待できる。

### 4.2 会話ログの感情分析

3.章に記載した三つのシステム利用実験の会話ログについて感情分析を行った。会話ログの感情分析手法として、OpenAI社のGPT API (gpt-4o) <sup>(注5)</sup>を活用した。鑑賞者 (User) と AI (Assistant) の発話内容のテキストを Positive, Negative, Unclassified の三つに分類した。結果を図9, 図10, 図11に示す。

まず、感情表出の割合について、実験3におけるユーザの発言においては、全体の46.3%がPositiveまたはNegativeな感情を示しており、約66~67%が中立的となった他の実験で用いたシステムよりも感情的なやり取りが多かったと考えられる。ユーザからのポジティブな反応は実験3, 実験1, 実験2の順に多くなっている。一方でユーザからのネガティブな反応は実



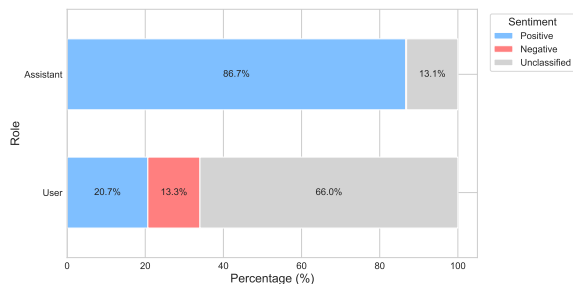


図 9: AI と鑑賞者の発話感情分類 (実験 1, 奈良県立美術館)

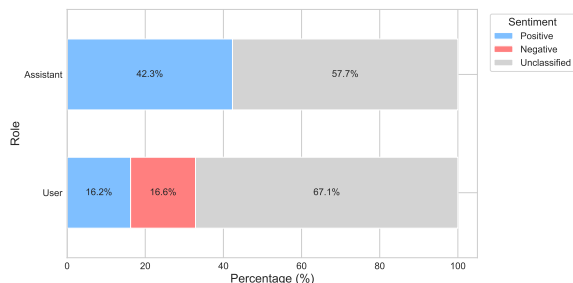


図 10: AI と鑑賞者の発話感情分類 (実験 2, 岡山芸術交流)

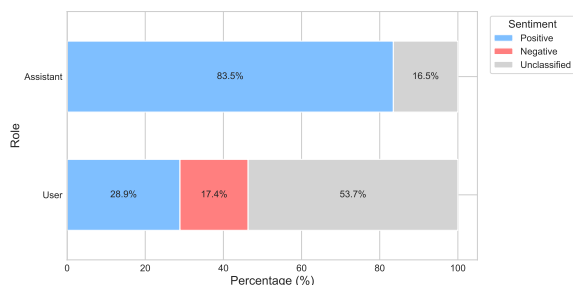


図 11: AI と鑑賞者の発話感情分類 (実験 3, 大学内)

験 3, 実験 2, 実験 1 の順に多くなっている。

AI の振る舞いの違いとして、Positive の割合が実験 1, 実験 3 のシステムでは 80% を超えており、ユーザに対して励ましや共感を示す振る舞いが多く、受容的・肯定的な対話スタイルをとっていることが分かる。一方で実験 2 のシステムでは Positive の割合が 42.3% にとどまっており、他のシステムと対話スタイルが異なっていることが分かる。一方で Negative の割合はどのシステムでも極めて低かった。

これらの結果から、AI が好意的に振る舞った場合、ユーザの発言が Positive および Negative の発話が相対的に増加する可能性があることが示唆された。

#### 4.3 ユーザーフィードバックの分析

本研究での最新システムである「MuseMate」を用いた実験 3 において、実験参加者に対して事後アンケートを実施した。対話型鑑賞システムを用いた鑑賞体験に関する質問への回答集計結果は図 12 に示すとおりとなった。

特に「4: ややそう思う」と「5: そう思う」の割合が高い項目は以下の 4 項目である。

- 作品を鑑賞している間、集中できましたか？
- 作品や展示の細部や特徴に気づくことができましたか？
- アプリの利用により、鑑賞体験の没入感は増しましたか？
- 対話を通じて新しい視点や発見を得られましたか？

この結果から、本研究におけるシステムでは作品鑑賞への集中力を高めることが、鑑賞体験における没入感を高めることにつながり、作品の細部や特徴など、美術作品鑑賞に関して新たな視点の発見に役立てることが期待できる。

一方で、以下の項目については肯定的な回答が少なかった。

- アプリとの対話が鑑賞に自然に溶け込んでいたと感じましたか？
- 鑑賞中、感情的な体験（驚き・感動・楽しさなど）がありましたか？

この結果から、鑑賞者の感情面に訴えかける情緒的な鑑賞体験を提供することにおいては課題があると考えられる。自由回答を確認したところ、対話型鑑賞を行うというシナリオを与えたことで、AI が鑑賞者に対して類似する問いかけを繰り返し行う傾向があり、「話題が堂々巡りをしている」という感覚を鑑賞者が感じていたことが明らかとなった。このことから、適切なタイミングでの話題転換についても改善の必要性がある。

#### 4.3.1 考察

以上の結果から、LLM を用いた対話型鑑賞システムは、鑑賞者の認知的関与や没入度を反映する行動指標を通じて、鑑賞体験を動的に支援できる可能性を有することが示唆された。特に、応答遅延時間と発話量の関係は、鑑賞者が即時的に思考・言語化できている状態と、熟考を要する状態の違いを捉える指標として有効であると考えられる。また、感情分析の結果から、AI が肯定的・共感的に振る舞うことで、鑑賞者の感情表出が促進される傾向が確認された。ユーザアンケートでは LLM との対話を通じて細部への発見や新しい視点を鑑賞者に提供することができた一方で、情緒的体験や対話の自然さに関する評価が相対的に低く、言語情報のみに基づく対話には限界があることも明らかとなった。

これらの課題に対し、視線や感情といった非言語的情報を統合することで、鑑賞者の言語化されない関心や内的状態を推定し、それに即した問いかけや応答を生成できる可能性がある。非言語情報を取り入れた LLM 主導の対話は、鑑賞者の没入感や情緒的体験を高め、より自然で適応的な鑑賞支援へとつながり、今回確認された対話の自然さや美術鑑賞の情緒的な部分へアプローチといった課題の解決が可能であると考えられる。

## 5. おわりに

本研究では、大規模言語モデル (LLM) を用いた美術作品の対話型鑑賞システムを対象として、三つの異なる鑑賞環境における運用実験を実施し、鑑賞者と AI との対話ログ分析およびユーザーフィードバックに基づく評価を行った。その結果、LLM を活用した対話型鑑賞は、鑑賞者の思考や注意を促し、鑑賞体験への没入感や新たな視点の獲得に寄与することが示唆された。特に、応答遅延時間と発話量の関係分析からは、鑑賞者が即時的に言語化できている状態と、熟考を要する状態の違いを捉えられる可能性が示され、鑑賞中の認知的関与を推定する行動指標として有効であることが示唆された。また、会話ログの感情分析より、AI が肯定的・共感的な対話スタイルをとること

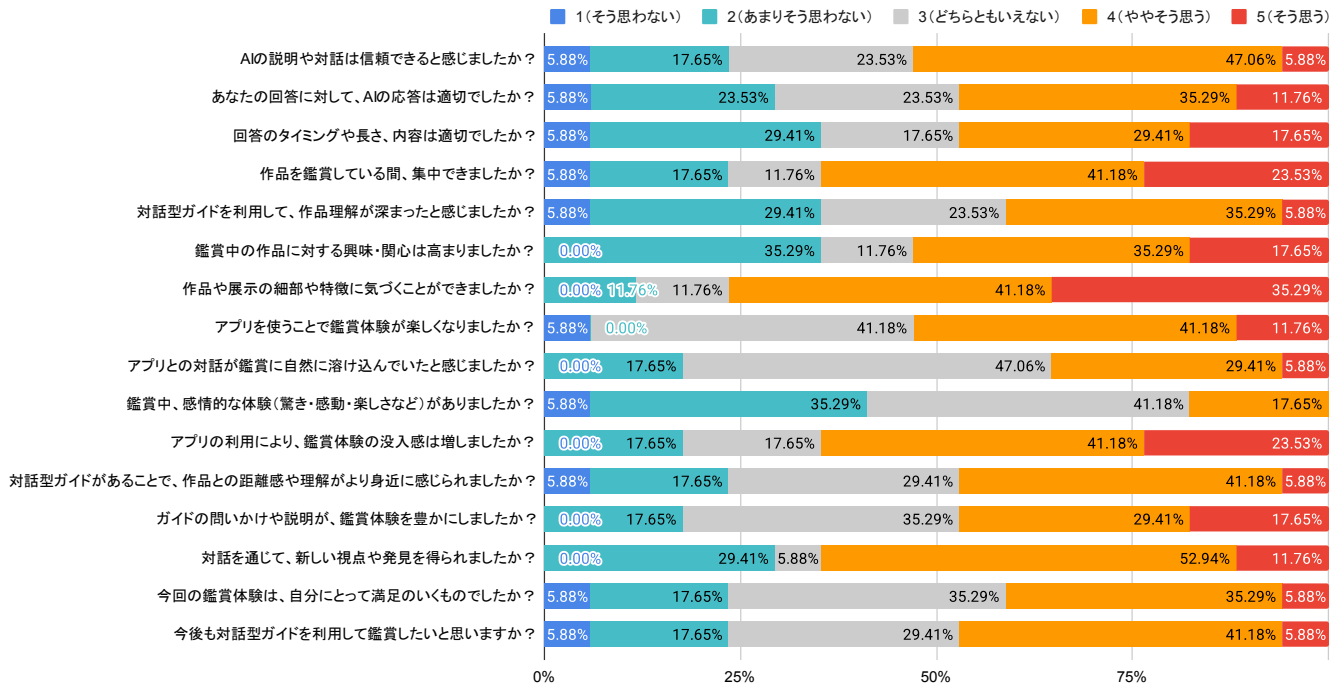


図 12: 実験 3 におけるユーザアンケートの集計結果

で、鑑賞者の感情表出が促進される傾向が確認された。一方で、ユーザアンケートの結果からは、対話の自然さや情緒的体験の喚起といった側面に課題が残ることも明らかとなった。これらは、言語情報のみに基づく対話型鑑賞支援の限界を示すものであり、鑑賞者の視線や感情といった非言語的情報を統合した支援の必要性を示唆している。

今後は、視線・表情・生体情報などから推定される鑑賞者の心理状態をリアルタイムに LLM へ統合し、鑑賞者の暗黙的な反応に即した問いかけや解説を生成することで、より自然で情緒的な鑑賞体験の実現を目指す。

## 謝 辞

本研究の一部は、SCAT 研究費助成および JSPS 科研費 (JP24K20763) の助成を受けて行われたものです。また運用実験実施に際して、奈良県立美術館、株式会社ブリッジウェル (ToyTalk)、中京大学現代社会学部真鍋講師の多大なる協力をいただきました。ここに感謝の意を申し上げます。

## 文 献

- [1] 文化庁, “博物館数・入館者数・学芸員数の推移,” [https://www.bunka.go.jp/seisaku/bijutsukan\\_hakubutsukan/shinko/suii/](https://www.bunka.go.jp/seisaku/bijutsukan_hakubutsukan/shinko/suii/). Accessed on 18 July, 2024.
- [2] J.C.-Y. Sun and S.-J. Yu, “Personalized Wearable Guides or Audio Guides: An Evaluation of Personalized Museum Guides for Improving Learning Achievement and Cognitive Load,” *International Journal of Human-Computer Interaction*, vol.35, no.4-5, pp.404–414, 2019.
- [3] H. Wang and A. Matvienko, “Experiencing Art Museum with a Generative Artificial Intelligence Chatbot,” *Proceedings of the 2025 ACM International Conference on Interactive Media Experiences*, pp.430–436, IMX ’25, 2025.
- [4] G. Trichopoulos, M. Konstantakis, G. Caridakis, A. Katifori, and M. Koukoulis, “Crafting a Museum Guide Using ChatGPT4,” *Big Data and Cognitive Computing*, vol.7, no.3, pp.1–15, 2023.
- [5] I. Vasic, H.-G. Fill, R. Quattrini, and R. Pierdicca, “LLM-Aided Museum Guide: Personalized Tours Based on User Preferences,” *Extended Reality*, pp.249–262, 2024.

- [6] A. Lewis and M. White, “Mitigating Harms of LLMs via Knowledge Distillation for a Virtual Museum Tour Guide,” *Proceedings of the 1st Workshop on Taming Large Language Models: Controllability in the Era of Interactive Assistants!*, pp.31–45, 2023.
- [7] Z. Wang, L.-P. Yuan, L. Wang, B. Jiang, and W. Zeng, “Virtuwander: Enhancing multi-modal interaction for virtual tour guidance through large language models,” *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, pp.1–20, CHI ’24, 2024.
- [8] I. Rachidi, A. Ezzakri, J. Bellver-Soler, and L.F. D’Haro, “Design, Generation and Evaluation of a Synthetic Dialogue Dataset for Contextually Aware Chatbots in Art Museums,” *Proceedings of the 15th International Workshop on Spoken Dialogue Systems Technology*, pp.20–28, 2025.
- [9] S. Ozaki, K. Hayashi, Y. Sakai, H. Kamigaito, K. Hayashi, and T. Watanabe, “Towards Cross-Lingual Explanation of Artwork in Large-scale Vision Language Models,” *Findings of the Association for Computational Linguistics: NAACL 2025*, pp.3773–3809, 2025.
- [10] S. Liu and J. Guo, “Smart museum and visitor satisfaction,” *Journal of Autonomous Intelligence*, vol.7, no.3, pp.1–13, 2024.
- [11] 近藤亮介, 松田裕貴, S.D. Terlizzi, D. Ahmetovic, 福光嘉伸, 平良繁幸, 諏訪博彦, S. Mascetti, 安本慶一, “美術館における鑑賞者の心理状態推定に向けたマルチモーダルデータセットの構築,” *社会情報学会関西支部研究会 (SSI)*, pp.1–10, 2024.
- [12] 鳥垣耀平, 諏訪博彦, 安本慶一, 松田裕貴, “美術館における鑑賞者の心理状態推定に向けた基礎分析,” *2024 年度 情報処理学会 関西支部 支部大会 講演論文集*, no.G-16, pp.1–6, 2024.
- [13] 鳥垣耀平, 松田裕貴, “表情や仕草データを用いた美術鑑賞者の心理状態推定手法の提案と評価,” *電子情報通信学会技術研究報告, センサネットワークとモバイルインテリジェンス研究会 (SeMI)*, 第 125 巻, pp.31–36, 2025.
- [14] Y. Matsuda, D. Fedotov, Y. Takahashi, Y. Arakawa, K. Yasumoto, and W. Minker, “Emotour: Estimating emotion and satisfaction of users based on behavioral cues and audiovisual data,” *Sensors*, vol.18, no.11, pp.1–19, 2018. <http://www.mdpi.com/1424-8220/18/11/3978>
- [15] N. Rach, Y. Matsuda, S. Ultes, W. Minker, and K. Yasumoto, “Estimating subjective argument quality aspects from social signals in argumentative dialogue systems,” *IEEE Access*, vol.9, pp.11610–11621, 2021.