

# 未来モビリティメーカー：対話と生成 AI による来場者参加型モビリティ共創体験

白井 公希<sup>1,a)</sup> 時田 聡実<sup>2,b)</sup> 内田 達弥<sup>4,c)</sup> 川上 悟<sup>1,d)</sup> 田中 五大<sup>1,e)</sup> 開 一夫<sup>3,f)</sup>  
石黒 祥生<sup>2,g)</sup> 大須賀 晋<sup>1,h)</sup>

**概要：**大規模展示会は、来場者が最新の技術や未来の社会像に触れる機会である一方で、展示内容と来場者自身の生活や価値観を結びつけるような「自分ごと化」された体験を提供することは難しい。とくに、来場者が自らの発想や関心に基づき、展示内容と結びつく新たなアイデアをその場で生成・可視化する機会は限られている。本研究では、AI エージェント「Saya」を用いて、来場者一人ひとりが自分の好みや生活スタイルに基づいた“未来のモビリティ”を生成し、その場で体験・共有できるシステム「未来モビリティメーカー」を提案・実装した。2025 年 10 月に開催された Japan Mobility Show において 12 日間にわたりフィールド評価を行った結果、来場者は日常的な趣味や行きたい場所を起点として未来のモビリティを構想しており、本システムが想像力のための足場として機能したことが示唆された。

## 1. はじめに

対話型エージェントや生成 AI の発展により、ユーザと協調してアイデアを生成する共創支援システムへの関心が高まっている。一方で、こうしたシステムの多くはラボ環境や小規模なデモにとどまり、数万人規模の来場者が訪れるような展示会で、スムーズかつパーソナルな体験を提供する例は少ない。また、展示会におけるガイドや案内ロボットは情報提供に重点が置かれることが多い。例えば、住宅、ファッション、食、モビリティなどの展示は、生活様式や価値観と密接に関わる領域にもかかわらず、自らの価値観を未来像に投影したり、新たなアイデアを考えたりする機会は限られている。

本研究で提案する「未来モビリティメーカー」は、フォトリアルな対話エージェント「Saya」が来場者と数ターンの

対話を行い、その中で得られた来場者の嗜好や興味をもとに、画像生成 AI によってオリジナルの未来モビリティ画像を生成し、来場者に提示・解説するシステムである。

2025 年秋、東京ビッグサイトで開催された一般社団法人日本自動車工業会（JAMA）主催の Japan Mobility Show 内の「Tokyo Future Tour 2035」において、本システムを 4 台設置し、12 日間にわたり一般来場者に公開した。対話ログをもとに、各フェーズへの到達率と離脱パターン、および来場者の発話内容を分析した。

本システムの貢献は以下の通りである。

- 未来のモビリティについて、単なる展示物ではなく、来場者の日常的な関心から未来像への接続を、構造化された対話によって支援した。
- 自らの言葉から即座に未来像を生成する体験を通じ、生成 AI との対話的な創作プロセスを提供した。
- 大規模展示会における対話エージェントの設計として、初期離脱を抑制し対話継続を促すためのシナリオ構成と質問設計に関する知見を示した。

## 2. 関連研究

### 2.1 公共空間における対話型ガイドシステム

公共空間や展示会における対話型エージェントやガイドロボットの研究では、通行人や来場者の注意を引き、会話を開始・維持するインタラクション設計が主要な課題とされてきた [1]。日本国内では、対話ロボット Robovie をショッピングセンターに設置し、来客に対して親しみやす

<sup>1</sup> 株式会社アイシン  
AISIN CORPORATION, Kariya, Aichi 448-0032, Japan  
<sup>2</sup> 東京大学大学院 情報学環・学際情報学府  
The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan  
<sup>3</sup> 東京大学大学院 総合文化研究科  
The University of Tokyo, Meguro, Tokyo 153-8902, Japan  
<sup>4</sup> Idein 株式会社  
Idein Inc., KandaJimbocho, Tokyo 101-0051, Japan  
a) koki.usui@aisin.co.jp  
b) tokidasatomi@g.ecc.u-tokyo.ac.jp  
c) t.uchida@idein.jp  
d) satoru.kawakami@aisin.co.jp  
e) godai.tanaka@aisin.co.jp  
f) khiraki@idea.c.u-tokyo.ac.jp  
g) ishiy@acm.org  
h) sohsuga@elec.aisin.co.jp

い対話を行う実証実験が行われてきた [2]。また、近年では JR や私鉄駅構内で、タッチレス操作可能な AI サイネージ [3] やアバター駅員による案内システム [4] の導入も進んでいる。これらのシステムは、主に施設内の案内や情報提供を目的とし、ユーザの意図や状態に応じて適切な情報を提示するという設計思想に基づいている。本研究は、情報提供を主目的とするこれらの従来システムとは異なり、来場者との対話を通じてパーソナライズされたコンテンツを生成し、来場者自身が創造的プロセスに参加する体験を提供することを目指す。

## 2.2 AI 共創インタラクション

生成 AI の進展に伴い、人と AI が協調的にアイデアや創作物を生み出す共創支援システムへの関心が高まっている。単なる自動生成ではなく、ユーザの入力に応じた応答・補完・拡張といった対話的インタラクションにより、人間の想像力を喚起する役割を果たす [5]。日本国内でも、生成 AI を活用して未来の生活をビジュアル化する「わたしスタイル in 20XX」展（日本科学未来館）[6] や、生成 AI を介して未来の社会課題解決アイデアを共創する「ミライのタネ」（住友館／大阪・関西万博）[7] など、来場者の創造的参加を促す展示が始まっている。これらは、生成 AI を用いて来場者の関心や選択に応じた未来像を提示する点に特徴がある一方で、主に Web インタフェース上での入力や選択式操作を通じたインタラクションが中心であり、音声による対話的なやりとりを主要な入力手段とする構成は確認されていない。これに対し、本研究で提案する「未来モビリティメーカー」は、来場者との音声対話を通じて、趣味や関心といった個人的な語りを引き出し、その自由記述的な発話内容に基づいて未来のモビリティ像を動的に生成する。これにより、生成 AI を定型的な生成装置ではなく、来場者の語りに対応しながら未来像を協調的に構築するパートナーとして機能させる枠組みを提示する。これは、生成 AI の特性を最大限に活かしながら、パブリックな展示環境においても、短時間かつ無人でパーソナルな創造的体験を提供可能であることを示す試みである。

## 3. 未来モビリティメーカー

本研究で提案する「未来モビリティメーカー」は、フォトリアルな外見を持つ対話エージェント Saya がイベントの来場者と数ターンの対話を行い、その中で得られた来場者の嗜好や興味をもとに、画像生成 AI (DALL-E3) を用いて未来モビリティの画像を生成し、来場者に提示・解説するシステムである。

未来モビリティメーカーは4つの特徴を備えている。

- 共創的対話デザイン：Saya は4つの質問（好きなこと、行きたい場所、聞こえる音、仕上げたい雰囲気）を通じて来場者のパーソナルな文脈を引き出し、画像



図 1 対話シナリオと体験の様子。Saya との対話例 (左) と来場者が Saya と対話し、生成されたモビリティ画像を閲覧している様子 (右)。

生成プロンプトを動的に生成することで、来場者ごとにカスタマイズされたデザインの画像を生成する。

- リアルタイム画像生成：会話の内容をもとに、その場で画像生成 AI (DALL-E3) が個別の未来モビリティ画像を生成する。
- 自動説明・共有：生成された未来モビリティ画像は、同時に生成するモビリティの名前と説明文とともに来場者に提示され、Saya がそれを読み上げることで解説する。モビリティ画像は QR コードを通じて来場者がダウンロードすることが可能である。
- 多様な来場者対応：画像処理によって子どもを、音声認識によって英語話者を自動的に識別し、対応するモードに切り替えることで、幅広い属性に配慮した設計となっている。

### 3.1 物理環境の構成

本システムは、単一のデスクトップ PC にマイク、カメラ、スピーカ、ディスプレイを接続し、それらを筐体に格納した構成である。マイクには Fairy Devices T-03C、スピーカにはサンワサプライ 400-SP088、カメラには Logicool BRIO C1000eR、ディスプレイには Saya の存在感を高めるために等身大で高精細に表示可能な 65 インチで 4K 解像度の Panasonic TH-65EQ2J を採用した。筐体はマイク、カメラ、スピーカを視認できないよう設計し、来場者が機器を意識することなく自然に対話できる環境を目指した。

図 2 に示すように、システムは3つの主要モジュールで構成されている（認識、対話管理、出力生成）。これらのモジュールは協調して、リアルタイムでの応答生成と、来場者の回答内容に応じた対話進行を実現する。

### 3.2 認識

本システムは、画像および音声入力をリアルタイムに処理するマルチモーダルな認識機能を備える。

画像認識では、顔検出によりユーザの有無を判定し、対

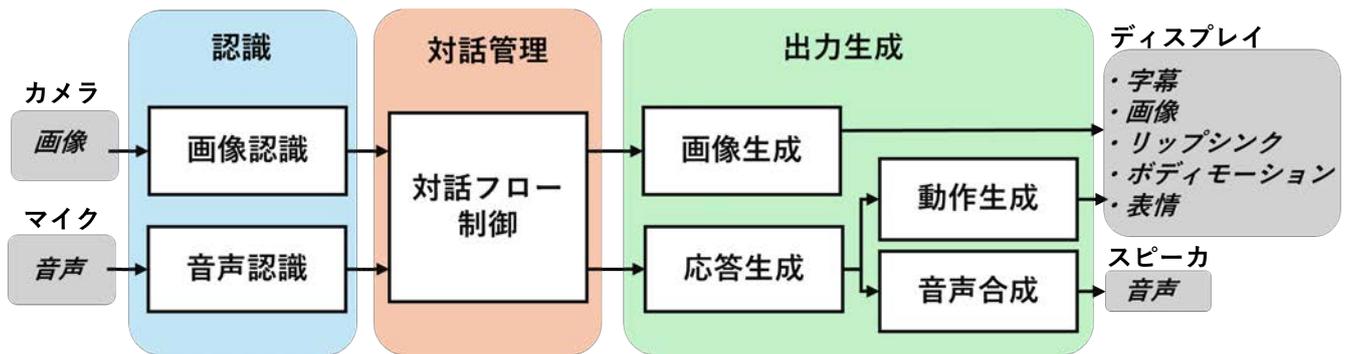


図 2 システム概要. Saya システムは、7つの構成要素 (画像認識、音声認識、対話フロー制御、画像生成、応答生成、動作生成、音声合成) で構成される。

話開始・終了のトリガーとして利用する。検出されたユーザに対して、年齢推定と顔検出時の位置情報に基づく子供判定を実施し、子供対応モードを実現する。子供対応モードでは、Sayaの姿勢を「かがみ姿勢」に切り替えることで子供との視線を合わせ、発話内容は敬語を排除し、子供でも理解可能な言葉遣いとする。また、口形状の時間的変化を用いた解析による発話判定を行い、発話区間抽出に利用することで、目の前のユーザのみの音声を選択的に受け付ける。さらに、ユーザ交代を検知し、対話をリセットすることで、高頻度な入れ替わりが発生しうる環境においても円滑な対話を維持する。

音声認識は、音量に基づくノイズゲート処理と Silero VAD (Voice Activity Detection: 音声区間検出)、画像認識による発話判定で音声区間を抽出し、YY System<sup>\*1</sup> (多言語音声認識システム) により音声認識を行う。それらの過程・結果は対話フロー制御と出力生成に利用される。

### 3.3 対話管理

対話フロー制御は、事前に定義された対話シナリオ (対話開始、4つの質問、画像生成、画像説明、対話終了という一連の流れ) に基づき、次のアクションを決定する。まず、ユーザの顔検出をトリガーとして対話を開始し、4つの質問を順次提示する。この際、各質問に対するユーザの応答が適切かどうかを GPT-4o-mini で判定し、不十分な場合は回答のヒントを提示した上で、同じ質問を再度繰り返す。ターンテイクングは、発話区間抽出、音声認識、発話交替の適切なタイミング判定 (TRP; Transition Relevance Place)、出力生成の過程および結果を統合的に利用することで、ユーザの発話終了を検知し適切なタイミングで応答を開始する。さらに、ユーザの顔検出が2秒間途切れた場合、対話シナリオの途中であっても対話を終了する。この仕組みにより、ユーザの状態に応じた柔軟な対話制御を実現している。



図 3 Saya のボディモーションと表情の例。

### 3.4 出力生成

出力生成モジュールは、Sayaの応答を生成し、それに基づいて動作生成と音声合成を行う。また、画像生成 AI を用いてリアルタイムの画像生成を行う。

応答生成では、Sayaの発話内容のテキストが生成される。選択されたプロンプトと過去の音声認識結果、Sayaの過去の発話内容を GPT-4o-mini に入力し、文脈およびキャラクター特性に適合した発話を生成する。プライバシーの観点から、音声データおよびテキストデータは音声認識システムおよび ChatGPT のモデル学習には使用されない設定とした。また、音声データはローカルおよび外部サーバー上に保存しなかった。テキストデータはローカルに保存されたが、特定のユーザと紐づかず、実験結果の分析のみに使用された。

動作生成では、生成された発話内容に基づき、Sayaのボディモーションと表情を生成する (図 3)。ボディモーションと表情は発話の感情的ニュアンスと対話の文脈を考慮し、発話内容に付与されたセマンティックなタグに基づいて選択する。タグは7種類に分類され、応答生成時に必ず1文につき1つ付与する: [0] 普通, [1] 合意, [2] 否定, [3] 簡単 (即答できる場合), [4] 質問, [5] 対面時挨拶, [6] 別

<sup>\*1</sup> <https://yysystem.com>



図 4 来場者との対話により生成された未来モビリティの例。来場者の回答(好きなこと, 行きたい場所, 聞こえる音, 仕上げたい雰囲気)をもとに DALL-E3 が生成したオリジナル画像。

れの挨拶。例えば、適切にフォーマットされた応答の出力は以下のパターンに従う。

“ AIデザイナーの、サヤです。 [3] あなたが思い描く、未来のモビリティを、一緒にデザインしましょう。 [0] ”

この構造化された出力により、各文単位で発話内容に対応した非言語表現を割り当てる。リップシンクは、生成された発話内容に基づき、音素に対応する口形遷移を計算し、音声合成と同期させる。音声合成には高品質な音声合成システムである Azure の Speech Studio を用い、Saya の声質を保持しつつ、明瞭で理解しやすい音声を生成する。

画像生成では、ベースとなるプロンプトに加え、来場者の回答から抽出・整理した趣味や嗜好の情報を統合し、画像生成 AI (DALL-E3) を用いて未来モビリティの画像を生成する。その後、生成された表情やボディモーション、シナリオに応じて画像をディスプレイ左部に描画し、合成音声を適切なタイミングで再生する。システムは 4K 解像度の局所的なブレンダリング映像をリアルタイムで滑らかに接続・統合することで、途切れのない動作遷移を実現し、Saya の表情やジェスチャーの細部まで鮮明に表示する。音声出力については、イベント会場の音響環境を考慮した音量調整を組み込み、来場者にとって自然で快適な聴取体験を提供する。また、音声認識による来場者発話内容と Saya の応答を字幕としてディスプレイ右部に表示することで、視覚的補助を加え、より正確で円滑なコミュニケーションを実現している。

## 4. フィールド評価

2025 年 10 月 29 日から 11 月 9 日までの 12 日間 (10 月 29 日はプレスデー)、東京ビッグサイトで開催された Japan Mobility Show にてフィールド評価を行った。会場内の Tokyo Future Tour 2035 エリアに未来モビリティメーカを 4 台設置し、一般来場者に公開した。会期中の Japan Mobility Show 全体の来場者は 101 万人であった。本システムはカメラによる人物検知を起点として自動的に対話を開始する仕様となっており、評価期間中に合計 11,907 回の対話開始が記録された。

### 4.1 対話シナリオ

未来モビリティメーカにおける Saya と来場者の対話は、5 つのフェーズで構成される (図 1)。対話は、来場者が Saya の前に立ちカメラに検知されることで開始される (対話開始フェーズ)。Saya は「ここでは私の質問に答えるだけで、あなたが思う未来の乗り物を描けます」と導入を行い、来場者の参加を促す。続く質問フェーズでは、Saya が 4 つの質問を順に投げかける。第 1 問では「あなたの好きなことやワクワクすることは何ですか?」と来場者の興味や関心を引き出す。例えば、来場者が「電車です。」のように回答すると、Saya は「電車に乗るのは楽しいですね!」と共感を示しながら、第 2 問として「行ってみたい場所はどこですか?」と行きたい場所に関する質問を続ける。第 3 問では「そこではどんな音がしますか?」と聴覚的なイメージを、第 4 問では「最後に、どんな雰囲気で仕上げたいですか?」と理想的な雰囲気を引き出す。これらの質問を通じて、来場者の趣味、行きたい場所、モビリティの出す音、雰囲気という多面的な情報が収集される。質問フェーズが終了すると、Saya は「これで質問は終了です。あなたの描くモビリティを描いてみますので、そのままお待ちください」と伝え、画像生成フェーズに移行する。画像生成が完了すると画像説明フェーズに移行し、Saya は生成されたモビリティの名称と特徴を読み上げる。例えば「電車型浮遊モビリティのガタリンです。ガタンゴトンという愛らしい音を響かせながら架空の街を浮遊する、丸みを帯びた電車型のモビリティです。」のように、来場者の回答を反映したパーソナライズされた説明が提示される。最後の退場フェーズでは、QR コードを通じて生成画像をダウンロードできることを案内し、「この後もモビリティショーをお楽しみください」と挨拶して対話を終了する。

### 4.2 評価指標

本評価では、定量的評価と定性的評価の 2 つのアプローチを採用した。

定量的評価では、対話の継続性を定量的に評価するため、



## 5. 考察

### 5.1 対話継続を促す設計要因

本評価では、対話開始から質問 1 への移行時に 51.1%の離脱が見られた一方、質問フェーズを進むにつれて離脱率は漸減し、画像生成以降は 10%未満となった。この結果は、公共空間における対話エージェント研究が指摘してきた初期関与の獲得の困難さを改めて示すとともに、一度対話に参加した来場者の参加意欲を段階的に高める設計の有効性を示唆している。従来の展示会ガイドシステム [3], [4] が情報提供を主目的としてきたのに対し、本システムでは来場者自身の語りを引き出し、それを可視化するという双方向的な構造を採用した。質問ごとの離脱率低下は、来場者が自分の回答の蓄積を認識し、最終的なアウトプットへの期待を高めていったためと考えられる。

本評価における定量データには、いくつかの未反映項目が存在する。カメラによる人物検知では、同一人物が一時的に検知範囲外に出て再検知される入れ替わり誤検知が発生しており、これらは実際には 1 回の対話であるにもかかわらず複数回としてカウントされている可能性がある。また、会場時間中のデバッグ動作に伴う発話回数も含まれている。これらをログデータから事後的に間引くことは困難であり、本研究の限界である。

### 5.2 日常と未来をつなぐ共創体験

定性的評価の結果は、本システムが、来場者が未知の概念（未来のモビリティ）を構想する際に、既知の経験（趣味、行きたい場所）から出発できるよう導く足場として機能したことを示唆している。来場者は「ゲーム」「キャンプ」「サッカー」といった自身の趣味、「北海道」「沖縄」「宇宙」といった行きたい場所、「波の音」「静か」といった感覚的な要素を通じて、抽象的な「未来のモビリティ」を自身の生活や価値観に接続していた。このプロセスにより、来場者は受動的な展示の見学者から、未来を能動的に構想する創造者へと役割を転換することができたと考えられる。

対話ログには、生成されたモビリティに対する「未来を感じます」「かっこいいです」「すごく素敵です」といった肯定的な反応が多く記録された一方、「もうちょっとワイルドな方が良かったな」「もっと面白おかしい感じでも良かった」といった批判的なフィードバックも見られた。後者は否定的な体験を意味するのではなく、来場者が AI を単なるツールではなく創造的なパートナーとして捉え、自身の意図をより正確に反映した結果を期待しているとも解釈できる。来場者自身の言葉を反映したパーソナライゼーションが、生成結果に対する当事者意識を持った批判的関与を誘発した可能性がある。今後の共創型システムは、単なる新規性だけでなく、来場者が生成プロセスに介入できる自由度を提供することが求められる。

今後の展望としては、展示会以外の場面への展開が考えられる。例えば、教育現場における将来の職業や社会像の探索支援、地域活性化イベントにおける住民参加型の未来ビジョン策定などが考えられる。対話を通じて個人の関心を引き出し、生成 AI によって即座に可視化するという本システムの枠組みは、「自分ごと化」が求められる様々な文脈に応用可能である。

## 6. デモンストレーションの説明

インタラクション 2026 でのデモンストレーションでは、Japan Mobility Show 2025 で展示したものと同様の「未来モビリティメカ」システムを体験する。来場者は Saya の前に立ち、4 つの質問に回答することで、自分だけの未来モビリティ画像を生成できる。生成された画像は Saya による解説とともに提示され、QR コードを通じて持ち帰ることが可能である。

### 参考文献

- [1] Masahiro Shiomi, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. A larger audience, please! — encouraging people to listen to a guide robot. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 31–38, 2010.
- [2] 宮下善太, 神田崇行, 塩見昌裕, 石黒浩, 萩田紀博. 来客と顔見知りになる案内ロボット. *インタラクション 2008 論文集*, 2008.
- [3] 東日本旅客鉄道株式会社. 駅係員の案内を「非接触」でサポート～「案内 AI みんなで育てようプロジェクト」第 2 弾～, 2020. [https://www.jreast.co.jp/press/2020/20201118\\_ho01.pdf](https://www.jreast.co.jp/press/2020/20201118_ho01.pdf), 閲覧日: 2025 年 12 月 18 日.
- [4] 日立製作所. 【日立×京王電鉄】生成 ai で駅の“案内係”が変わる! 「ai 駅係員サービス」実証実験スタート, 2024. [https://digital-highlights.hitachi.co.jp/\\_ct/17750346](https://digital-highlights.hitachi.co.jp/_ct/17750346), 閲覧日: 2025 年 12 月 18 日.
- [5] Janin Koch, Andrés Lucero, Lena Hegemann, and Antti Oulasvirta. May ai? design ideation with cooperative contextual bandits. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–12, New York, NY, USA, 2019. Association for Computing Machinery.
- [6] 日本科学未来館. Ai とえがく? わたしスタイル in 20xx, 2025. [https://www.jreast.co.jp/press/2020/20201118\\_ho01.pdf](https://www.jreast.co.jp/press/2020/20201118_ho01.pdf), 閲覧日: 2025 年 12 月 18 日.
- [7] 住友館 (大阪・関西万博). Experience — ミライのタネ, 2025. <https://sumitomoexpo.com/experience/mirainotane>, 閲覧日: 2025 年 12 月 18 日.