Novelty Assessment Report

Paper: Social Agents: Collective Intelligence Improves LLM Predictions
PDF URL: https://openreview.net/pdf?id=73J3hsato3
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-07

Abstract

In human society, collective decision making has often outperformed the judgment of individuals. Classic examples range from estimating livestock weights to predicting elections and financial markets, where averaging many independent guesses often yields results more accurate than experts. These successes arise because groups bring together diverse perspectives, independent voices, and distributed knowledge, combining them in ways that cancel individual biases. This principle, known as the Wisdom of Crowds, underpins practices in forecasting, marketing, and preference modeling. Large Language Models (LLMs), however, typically produce a single definitive answer. While effective in many settings, this uniformity overlooks the diversity of human judgments shaping responses to ads, videos, and webpages. Inspired by how societies benefit from diverse opinions, we ask whether LLM predictions can be improved by simulating not one answer but many. We introduce Social Agents, a multi-agent framework that instantiates a synthetic society of human-like personas with diverse demographic (e.g., age, gender) and psychographic (e.g., values, interests) attributes. Each persona independently appraises a stimulus such as an advertisement, video, or webpage, offering both a quantitative score (e.g., click-through likelihood, recall score, likability) and a qualitative rationale. Aggregating these opinions produces a distribution of preferences that more closely mirrors real human crowds. Across eleven behavioral prediction tasks, Social Agents outperforms single-LLM baselines by up to 67.45% on simple judgments (e.g. webpage likability) and 9.88% on complex interpretive reasoning (e.g. video memorability). Social Agents' individual persona predictions also align with human judgments, reaching Pearson correlations up to 0.71. These results position computational crowd simulation as a scalable, interpretable tool for improving behavioral prediction and supporting societal decision making.

Core Task Landscape

This paper addresses: **Improving Behavioral Prediction Using Multi-Agent Collective Intelligence**

A total of **50 papers** were analyzed and organized into a taxonomy with **18 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Multi-Agent Trajectory and Motion Prediction**
- **Multi-Agent Reinforcement Learning and Decision-Making**
- **LLM-Based Multi-Agent Systems and Collaboration**
- **Collective Behavior and Emergent Phenomena**
- **Explainability and Interpretability in Multi-Agent Systems**
- **Domain Applications and Cross-Disciplinary Studies**

Complete Taxonomy Tree

- Improving Behavioral Prediction Using Multi-Agent Collective Intelligence Survey Taxonomy
- Multi-Agent Trajectory and Motion Prediction
  - Interaction-Aware Trajectory Forecasting
  - Attention and Transformer-Based Approaches (3 papers)
    - [2] VAIN: Attentional Multi-agent Predictive Modeling (Hoshen, 2022) View paper
    - [5] AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting (Yuan Ye, 2021) View paper
    - [15] MotionLM: Multi-Agent Motion Forecasting as Language Modeling (Ari Seff, 2023) View paper
  - Graph and Relational Network Methods (3 papers)
    - [16] Mart: Multiscale relational transformer networks for multi-agent trajectory prediction (Seongju Lee, 2024) View paper
    - [21] Online Multi-Agent Forecasting With Interpretable Collaborative Graph Neural Networks (Maosen Li, 2022) View paper
    - [38] Evolvegraph: Multi-agent trajectory prediction with dynamic relational reasoning (Jiachen Li, 2020) View paper
  - Diffusion and Generative Models (2 papers)
    - [20] MotionDiffuser: Controllable Multi-Agent Motion Prediction Using Diffusion (Jiang, 2023) View paper
    - [49] Continual Multi-Agent Interaction Behavior Prediction With Conditional Generative Memory (Hengbo Ma, 2021) View paper
  - Domain-Specific Trajectory Prediction
  - Autonomous Driving and Traffic Scenarios (5 papers)
    - [1] Improving Multi-Agent Trajectory Prediction Using Traffic States on Interactive Driving Scenarios (Chalavadi Vishnu, 2023) View paper
    - [10] Correction to "Improving Multi-Agent Trajectory Prediction Using Traffic States on Interactive Driving Scenarios" (Chalavadi Vishnu, 2023) View paper
    - [13] Cmp: Cooperative motion prediction with multi-agent communication (Zehao Wang, 2025) View paper
    - [17] Reasoning multi-agent behavioral topology for interactive autonomous driving (Li Chen, 2024) View paper
    - [25] A hierarchical hybrid learning framework for multi-agent trajectory prediction (Yujun Jiao, 2024) View paper
  - Human Trajectory Prediction (2 papers)
    - [22] Sopermodel: Leveraging social perception for multi-agent trajectory prediction (Heming Yang, 2025) View paper

        - [28] Recent Advances in Multi-Agent Human Trajectory Prediction: A Comprehensive Review (C'eline Finet, 2025) View paper
    - Uncertainty and Robustness in Prediction (3 papers)
    - [14] Collaborative Uncertainty in Multi-Agent Trajectory Forecasting (Tang, 2022) View paper
    - [30] Collaborative Uncertainty Benefits Multi-Agent Multi-Modal Trajectory Forecasting (Bohan Tang, 2023) View paper
    - [47] Multi-Agent Reachability Calibration with Conformal Prediction (Anish Muthali, 2023) View paper
    - Adaptive and Transfer Learning for Prediction (2 papers)
    - [42] ADAPT: Efficient Multi-Agent Trajectory Prediction with Adaptation (GÃ¶rkay Aydemir, 2023) View paper
    - [44] Multi-agent driving behavior prediction across different scenarios with self-supervised domain knowledge (Hengbo Ma, 2021) View paper
- Multi-Agent Reinforcement Learning and Decision-Making
    - Cooperative and Collaborative MARL (3 papers)
    - [11] Towards collaborative intelligence: Propagating intentions and reasoning for multi-agent coordination with large language models (Wang Haoyu, 2024) View paper
    - [23] Distributed multi-agent reinforcement learning via interactive relationship construction and behavior prediction for confrontation scenario (Jian Xiao, 2026) View paper
    - [29] Estimation of collective maneuvers through cooperative multi-agent planning (Jens Schulz, 2017) View paper
    - Competitive and Game-Theoretic MARL (3 papers)
    - [35] Cooperative Game-based Intelligent Actions Making for Constrained Multi-agent System (Xiaoyue Jin, 2024) View paper
    - [36] Behavior Reasoning for Opponent Agents in Multi-Agent Learning Systems (Yaqing Hou, 2022) View paper
    - [41] Multi-Agent cooperation based on opponent's behavior forecast (Chaoyang Ji, 2009) View paper
    - Large-Scale and Platform-Based MARL (2 papers)
    - [6] Magent: A many-agent reinforcement learning platform for artificial collective intelligence (Lianmin Zheng, 2018) View paper
    - [31] Generating intelligent agent behaviors in multi-agent game AI using deep reinforcement learning algorithm (Rosalina Rosalina, 2023) View paper
    - Application-Specific MARL (2 papers)
    - [7] Impact of collective behaviors of autonomous vehicles on urban traffic dynamics: A multi-agent reinforcement learning approach (JamrÃ³z, 2025) View paper
    - [40] Cognitive multi-agent empowering mobile edge computing for resource caching and collaboration (Rui Wang, 2020) View paper
- LLM-Based Multi-Agent Systems and Collaboration
    - LLM-Driven Collaborative Frameworks (2 papers)
    - [3] AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors. (Chen Weize, 2024) View paper
    - [4] AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors in Agents (Chen Weize, 2023) View paper
    - LLM-Based Behavioral Prediction ★ (3 papers)
    - [0] Social Agents: Collective Intelligence Improves LLM Predictions (Anon et al., 2026) View paper
    - [19] Position: Simulating Society Requires Simulating Thought (Wu Jiayi, 2025) View paper
    - [27] ElliottAgents: A natural language-driven multi-agent system for stock market analysis and prediction (JarosÅaw A. Chudziak, 2024) View paper
- Collective Behavior and Emergent Phenomena
    - Theoretical and Analytical Models of Collective Behavior (2 papers)
    - [9] Collective behavior of discrete time multi-agent systems with dynamical opinions (Han Guo, 2024) View paper
    - [39] Collective Behaviors of Discrete-Time Multi-Agent Systems Over Signed Digraphs (Zhen-Hua Zhu, 2021) View paper
    - Simulation and Modeling of Collective Behavior (5 papers)
    - [8] Designing emergent behaviors: From local interactions to collective intelligence (Maja J. MatariÄ, 1993) View paper
    - [18] TrafficSim: Learning to Simulate Realistic Multi-Agent Behaviors (Simon Suo, 2021) View paper
    - [33] Modeling human behavior during emergency evacuation using intelligent agents: A multi-agent simulation approach (Sharad Sharma, 2018) View paper
    - [43] Swarm intelligence, social force and multi-agent modeling of heroic altruism behaviors under collective risks (Peng LÃ¼, 2021) View paper
    - [48] Collective behavior simulation based on agent with artificial emotion (Yongbin Zhu, 2019) View paper
    - Collective Intelligence and Crowd Simulation (4 papers)
    - [24] Collective intelligence in teams: Contextualizing collective intelligent behavior over time (Margo Janssens, 2022) View paper
    - [26] Human-machine teamwork: an exploration of multi-agent systems, team cognition, and collective intelligence (Canonico, 2019) View paper
    - [37] Agent-based computational modeling of emergent collective intelligence (V. Singh, 2009) View paper
    - [45] Leveraging collective intelligence for behavioral prediction in signed social networks through evolutionary approach (Gaganmeet Kaur Awal, 2019) View paper
- Explainability and Interpretability in Multi-Agent Systems (2 papers)
    - [32] MACIE: Multi-Agent Causal Intelligence Explainer for Collective Behavior Understanding (Weinberg, 2025) View paper
    - [34] Multi-Agent Hybrid Prediction in Autonomous Driving (Yau, 2025) View paper
- Domain Applications and Cross-Disciplinary Studies (3 papers)
    - [12] Exploring the power of artificial intelligence in supply chain management: a literature review on the artificial intelligence applications and tools used in supply chains â¦ (MM Harrir, 2025) View paper
    - [46] Multi-agent systems in predicting global financial crises (Patrick, 2024) View paper
    - [50] Advances in Behavior Prediction Techniques for Intelligent Systems (Bighashdel, 2024) View paper

Narrative

Core task: Improving behavioral prediction using multi-agent collective intelligence. The field encompasses diverse approaches to understanding and forecasting how multiple agents—whether autonomous vehicles, software agents, or simulated populations—interact and evolve over time. The taxonomy reveals six major branches: Multi-Agent Trajectory and Motion Prediction focuses on spatial forecasting in domains like autonomous driving, often leveraging graph-based architectures such as AgentFormer[5] and collaborative frameworks like Cooperative Motion Prediction[13]. Multi-Agent Reinforcement Learning and Decision-Making addresses sequential decision problems where agents learn policies through interaction, exemplified by platforms like Magent Platform[6]. LLM-Based Multi-Agent Systems and Collaboration explores how large language models enable richer agent communication and coordination, as seen in AgentVerse Collaboration[3] and AgentVerse Emergent[4]. Collective Behavior and Emergent Phenomena investigates how simple rules

yield complex group dynamics, from early work like Designing Emergent Behaviors[8] to recent studies on Emergent Collective Intelligence[37]. Explainability and Interpretability in Multi-Agent Systems tackles the challenge of making agent decisions transparent, while Domain Applications and Cross-Disciplinary Studies apply these methods to finance, supply chains, and social simulation.

Recent activity highlights contrasting philosophies: trajectory prediction methods emphasize geometric reasoning and uncertainty quantification, whereas LLM-based approaches prioritize semantic understanding and flexible collaboration. Social Agents[0] sits squarely within the LLM-Based Behavioral Prediction cluster, leveraging language models to simulate nuanced social interactions and predict human-like behaviors in multi-agent settings. This positions it alongside works like Simulating Society[19] and ElliottAgents[27], which similarly use LLMs to model complex social phenomena, but Social Agents[0] emphasizes collective intelligence mechanisms that emerge from agent interactions rather than purely individual reasoning. The tension between data-driven motion forecasting (e.g., MotionLM[15]) and knowledge-grounded social simulation remains a central open question, with hybrid approaches like Hybrid Prediction[34] attempting to bridge these paradigms by combining neural trajectory models with symbolic reasoning about agent intentions.

Related Works in Same Category

The following **2 sibling papers** share the same taxonomy leaf node with the original paper:

1. Position: Simulating Society Requires Simulating Thought

Authors: Wu Jiayi, Chance Jiajie Li, Mo Zhenze, Jiayi Wu, Qu, et al. (30 authors total) | **Year/Venue**: 2025 | **URL**: View paper

Abstract

Simulating society with large language models (LLMs), we argue, requires more than generating plausible behavior; it demands cognitively grounded reasoning that is structured, revisable, and traceable. LLM-based agents are increasingly used to emulate individual and group behavior, primarily through prompting and supervised fine-tuning. Yet current simulations remain grounded in a behaviorist "demographics in, behavior out" paradigm, focusing on surface-level plausibility. As a result, they ofte...

Relationship Analysis

Both papers belong to the LLM-Based Behavioral Prediction category, using LLMs to simulate human behaviors and judgments through persona-based agents. They overlap in their use of multi-agent systems with diverse personas to improve behavioral predictions, with both drawing on collective intelligence principles. However, the original paper (Social Agents) focuses on aggregating diverse persona predictions to improve accuracy across specific behavioral tasks (e.g., CTR, memorability), while the candidate paper (GenMinds) emphasizes cognitively grounded reasoning structures and causal belief formation, arguing for structured internal representations rather than surface-level behavioral mimicry.

---

2. ElliottAgents: A natural language-driven multi-agent system for stock market analysis and prediction

Authors: JarosÅ□aw A. Chudziak, MichaÅ□ Wawer | **Year/Venue**: 2024 | **URL**: View paper

Abstract

This paper presents ElliottAgents, a multi-agent system leveraging natural language processing (NLP) and large language models (LLMs) to analyze complex stock market data. The system combines AI-driven analysis with the Elliott Wave Principle to generate human-comprehensible predictions and explanations. A key feature is the natural language dialogue between agents, enabling collaborative analysis refinement. The LLM-enhanced architecture facilitates advanced language understanding, reasoning, a...

Relationship Analysis

Both papers belong to the LLM-Based Behavioral Prediction category, using LLMs to simulate human-like agents for prediction tasks. They overlap in employing multi-agent systems where diverse LLM-based personas independently generate predictions that are aggregated to improve accuracy. However, the original paper (Social Agents) focuses on predicting human behavioral responses to digital content (ads, webpages, videos) across diverse social contexts, while the candidate paper (ElliottAgents) applies multi-agent LLMs specifically to financial market analysis and stock price prediction using Elliott Wave Principle, representing a domain-specific application rather than general behavioral prediction.

Contributions Analysis

Overall novelty summary. The paper introduces Social Agents, a multi-agent framework that simulates diverse human personas to improve behavioral prediction through collective intelligence. It resides in the LLM-Based Behavioral Prediction leaf, which contains only three papers total, including this work and two siblings (Simulating Society and ElliottAgents). This represents a relatively sparse research direction within the broader taxonomy of fifty papers, suggesting the application of LLM-driven multi-agent systems specifically to behavioral forecasting remains an emerging area compared to more established branches like trajectory prediction or reinforcement learning.

The taxonomy reveals that LLM-Based Behavioral Prediction sits within the larger LLM-Based Multi-Agent Systems and Collaboration branch, which also includes LLM-Driven Collaborative Frameworks focused on task coordination rather than prediction. Neighboring branches pursue fundamentally different approaches: Multi-Agent Trajectory and Motion Prediction emphasizes spatial forecasting using graph networks and diffusion models, while Collective Behavior and Emergent Phenomena studies group dynamics through simulation and theoretical models. The scope note explicitly excludes trajectory prediction, positioning this work as complementary to geometric reasoning methods while sharing conceptual ground with Collective Intelligence and Crowd Simulation studies.

Among twenty-six candidates examined across three contributions, none were identified as clearly refuting the work. The Social Agents framework examined nine candidates with zero refutable overlaps, the empirical evaluation across eleven tasks examined seven candidates with similar results, and the synthetic dataset contribution examined ten candidates without finding substantial prior work. This suggests that within the limited search scope—primarily top-K semantic matches and citation expansion—the specific combination of LLM-based persona simulation, collective aggregation mechanisms, and systematic evaluation across diverse behavioral tasks appears relatively unexplored, though the analysis does not claim exhaustive coverage of all potentially relevant literature.

The limited search scope and sparse taxonomy leaf indicate the work occupies a nascent research direction where LLM capabilities meet collective intelligence principles. The absence of refutable candidates among twenty-six examined papers suggests novelty within the sampled literature, though the small sibling count and focused search strategy mean substantial related work may exist outside the top-K semantic neighborhood or in adjacent application domains not captured by this taxonomy structure.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Social Agents multi-agent framework

Description: The authors propose Social Agents, a framework that operationalizes the wisdom of crowds principle by creating ensembles of LLM-based persona agents with diverse demographic and psychographic characteristics. Each persona independently evaluates stimuli and provides quantitative predictions with qualitative rationales, which are then aggregated to produce collective judgments that mirror real human crowds.

This contribution was assessed against **9 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Comapoi: A collaborative multi-agent framework for next poi prediction bridging the gap between trajectory and language

URL: View paper

Brief Assessment

Comapoi[56] focuses on next POI prediction using three specialized agents (profiler, forecaster, predictor) for spatiotemporal tasks, not on simulating diverse human personas for collective behavioral predictions as in the original paper.

2. Internet of agents: Weaving a web of heterogeneous agents for collaborative intelligence

URL: View paper

Brief Assessment

Internet of Agents[54] focuses on integrating heterogeneous third-party agents with different tools and architectures across distributed environments, not on simulating diverse human personas for collective predictions. The frameworks serve fundamentally different purposes.

3. Drugagent: Multi-agent large language model-based reasoning for drug-target interaction prediction

URL: View paper

Brief Assessment

Drugagent[58] focuses on drug-target interaction prediction using specialized domain agents (AI agent, KG agent, search agent) with coordinator-based architecture, not on simulating diverse human personas for behavioral predictions. The frameworks serve fundamentally different purposes and domains.

4. MAGIS: LLM-Based Multi-Agent Framework for GitHub Issue Resolution

URL: View paper

Brief Assessment

MAGIS[51] focuses on GitHub issue resolution using specialized software development agents (Manager, Repository Custodian, Developer, QA Engineer), not on simulating diverse human personas for behavioral predictions or collective intelligence applications.

5. FinVision: A Multi-Agent Framework for Stock Market Prediction

URL: View paper

Brief Assessment

FinVision[53] focuses on financial trading with specialized agents processing financial data modalities (news, charts, trading signals), while the original paper addresses behavioral prediction tasks using diverse demographic/psychographic personas to simulate crowd wisdom. The application domains and agent design principles differ fundamentally.

6. Why Should Next-Gen LLM Multi-Agent Systems Move Beyond Fixed Architectures to Dynamic, Input-Driven Graphs?

URL: View paper

Brief Assessment

Dynamic Agent Graphs[59] focuses on dynamic, input-driven graph architectures for multi-agent systems, whereas Social Agents implements a fixed ensemble of diverse persona agents for behavioral prediction tasks. The candidate does not demonstrate prior work on wisdom-of-crowds-based persona ensembles for collective intelligence.

7. Large Language Model-Empowered Interactive Load Forecasting

URL: View paper

Brief Assessment

Interactive Load Forecasting[60] focuses on human-operator collaboration in power system load forecasting using LLM agents for a specific domain task, not on creating diverse persona agents for collective behavioral predictions across general tasks.

8. LLM-Powered Multi-Agent System for Automated Crypto Portfolio Management

URL: View paper

Brief Assessment

Crypto Portfolio Management[57] focuses on cryptocurrency investment using specialized agents for financial tasks (data analysis, portfolio management), not on simulating diverse human personas for behavioral predictions across general domains like ads, webpages, and videos.

9. Large language model based multi-agents: A survey of progress and challenges

URL: View paper

Brief Assessment

LLM Multi Agents[52] surveys multi-agent frameworks broadly but does not present a specific framework using diverse personas for collective predictions. The candidate focuses on architectural patterns, communication structures, and application domains rather than the wisdom-of-crowds aggregation mechanism central to Social Agents.

Contribution 2: Empirical evaluation across eleven behavioral prediction tasks

Description: The authors conduct a comprehensive evaluation of Social Agents across eleven diverse tasks spanning low-, medium-, and high-level construals based on Construal Level Theory. The framework demonstrates consistent improvements over single-LLM baselines and often exceeds task-specific trained models, showing that collective intelligence can improve LLM predictions across different cognitive domains.

This contribution was assessed against **7 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Towards Better Forecasting by Fusing Near and Distant Future Visions

URL: View paper

Brief Assessment

Fusing Future Visions[63] focuses on multivariate time series forecasting using construal level theory to fuse near and distant future predictions, not on behavioral prediction tasks involving human judgments across different cognitive domains.

---

2. Abstract thinking facilitates aggregation of information.

URL: View paper

Brief Assessment

Abstract Thinking Aggregation[64] focuses on numerical averaging tasks and emotional face processing under construal level manipulations, not on behavioral prediction tasks across diverse domains using LLM-based multi-agent frameworks.

---

3. People watching: Social perception and the ensemble coding of bodies

URL: View paper

Brief Assessment

Ensemble Coding Bodies[62] examines ensemble perception of physical body characteristics (BMI, waist-to-hip ratio) and social attributes (emotion, gender) from groups of bodies in visual perception tasks. The original paper evaluates LLM-based social agents across behavioral prediction tasks spanning construal levels. These are fundamentally different research domains with no overlap in methodology or objectives.

---

4. Limits of a Deductive Construal of the Function of Scientific Theories A Comment

URL: View paper

Brief Assessment

Deductive Construal Limits[67] discusses theoretical limitations of deductive construal in scientific theories and unique event ensembles, not empirical evaluation of behavioral prediction tasks using ensemble methods or LLMs.

---

5. â¦ COGNITIVE MEMORY SYSTEMS: A UNIFIED FRAMEWORK FOR SEQUEN-TIAL INFORMATION PROCESSING AND LONG-TERM BEHAVIORAL PREDICTION

URL: View paper

Brief Assessment

Cognitive Memory Systems[66] focuses on temporal reasoning in financial forecasting and e-commerce recommendation using bio-inspired memory architectures, not on ensemble methods for behavioral prediction across construal level tasks as in the original paper.

---

6. How does the construal level affect consumers' intention to adopt product ratings and individual reviews?

URL: View paper

Brief Assessment

Construal Level Adoption[61] focuses on consumer adoption of product ratings versus individual reviews in online shopping contexts, not on ensemble methods for behavioral prediction across diverse tasks. The candidate applies construal level theory to understand consumer decision-making between aggregate metrics and individual reviews, which is a fundamentally different research question than evaluating multi-agent LLM frameworks across behavioral prediction tasks.

---

7. The effect of construal level on predictions of task duration

URL: View paper

Brief Assessment

Construal Task Duration[65] focuses on how construal level affects predictions of task duration in a controlled experimental setting, not on ensemble methods for behavioral prediction across diverse tasks.

Contribution 3: Synthetic dataset of persona-conditioned predictions

Description: The authors release a dataset containing persona-conditioned predictions, definitions, and rationales generated by Social Agents across all eleven behavioral tasks. This dataset captures how diverse personas interact with and evaluate digital content, providing a resource for understanding collective intelligence in synthetic crowds.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

---

1. Faithful Persona-based Conversational Dataset Generation with Large Language Models

URL: View paper

Brief Assessment

Faithful Persona Dataset[74] focuses on generating persona-based conversational dialogues between users, not on creating datasets of persona-conditioned predictions for evaluating digital content (ads, webpages, videos) as in the original paper.

---

2. LLM Generated Persona is a Promise with a Catch

URL: View paper

Brief Assessment

LLM Persona Promise[68] focuses on generating persona profiles themselves and analyzing biases in persona generation methods, not on releasing datasets of persona-conditioned predictions for behavioral tasks. The candidate's contribution is the open-sourcing of generated personas for validation, not behavioral prediction datasets.

---

3. Personabench: Evaluating ai models on understanding personal information through accessing (synthetic) private user data

URL: View paper

Brief Assessment

Personabench[69] focuses on generating synthetic user profiles and private documents to evaluate AI models' ability to extract personal information through RAG pipelines. The original paper releases persona-conditioned predictions for behavioral tasks like ad CTR and webpage likability, while Personabench[69] creates personal Q&A datasets from synthetic user documents. These are distinct evaluation paradigms with different objectives.

---

4. MCPDial: A Minecraft Persona-driven Dialogue Dataset

URL: View paper

Brief Assessment

MCPDial[76] focuses on persona-driven dialogue generation for Minecraft gameplay interactions, not on persona-conditioned predictions for evaluating digital content like ads, videos, or webpages. The datasets serve fundamentally different purposes.

---

5. Virtual personas for language models via an anthology of backstories

URL: View paper

Brief Assessment

Virtual Personas Backstories[75] focuses on generating backstories to condition LLMs for survey approximation, not on creating datasets of persona-conditioned predictions for digital content evaluation tasks like ads, videos, and webpages as in the original paper.

---

6. PersonaGym: Evaluating Persona Agents and LLMs

URL: View paper

Brief Assessment

PersonaGym[71] focuses on evaluating persona agents through dynamic question generation and scoring mechanisms, not on releasing datasets of persona-conditioned predictions for digital content evaluation. The candidate's contribution is an evaluation framework, not a dataset release.

---

7. Co-persona: Leveraging LLMs and Expert Collaboration to Understand User Personas through Social Media Data Analysis

URL: View paper

Brief Assessment

Co-persona[72] focuses on developing user personas from social media data for product design and marketing applications, not on generating synthetic datasets of persona-conditioned predictions for digital content evaluation tasks as described in the original paper.

---

8. Evaluating Persona Prompting for Question Answering Tasks

URL: View paper

Brief Assessment

Persona Prompting Evaluation[73] focuses on evaluating persona prompting effectiveness on question answering tasks, not on releasing datasets of persona-conditioned predictions for digital content evaluation across behavioral tasks.

---

9. BluePrint: A Social Media User Dataset for LLM Persona Evaluation and Training

URL: View paper

Brief Assessment

BluePrint[77] focuses on clustering real user behaviors into personas for social media simulation, not on generating synthetic persona-conditioned predictions across behavioral tasks like the original paper's dataset of ad CTR, webpage likability, and memorability predictions.

---

10. Higher-Order Binding of Language Model Virtual Personas: a Study on Approximating Political Partisan Misperceptions

URL: View paper

Brief Assessment

Political Partisan Misperceptions[70] focuses on generating backstories for political science studies (partisan misperceptions, democratic backsliding) rather than general behavioral predictions across digital content evaluation tasks. The dataset described contains political survey responses, not predictions about ads, videos, or webpages as in the original paper.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] Social Agents: Collective Intelligence Improves LLM Predictions View paper
- [1] Improving Multi-Agent Trajectory Prediction Using Traffic States on Interactive Driving Scenarios View paper
- [2] VAIN: Attentional Multi-agent Predictive Modeling View paper
- [3] AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors. View paper
- [4] AgentVerse: Facilitating Multi-Agent Collaboration and Exploring Emergent Behaviors in Agents View paper
- [5] AgentFormer: Agent-Aware Transformers for Socio-Temporal Multi-Agent Forecasting View paper
- [6] Magent: A many-agent reinforcement learning platform for artificial collective intelligence View paper
- [7] Impact of collective behaviors of autonomous vehicles on urban traffic dynamics: A multi-agent reinforcement learning approach View paper
- [8] Designing emergent behaviors: From local interactions to collective intelligence View paper
- [9] Collective behavior of discrete time multi-agent systems with dynamical opinions View paper
- [10] Correction to "Improving Multi-Agent Trajectory Prediction Using Traffic States on Interactive Driving Scenarios" View paper
- [11] Towards collaborative intelligence: Propagating intentions and reasoning for multi-agent coordination with large language models View paper
- [12] Exploring the power of artificial intelligence in supply chain management: a literature review on the artificial intelligence applications and tools used in supply chains â View paper
- [13] Cmp: Cooperative motion prediction with multi-agent communication View paper
- [14] Collaborative Uncertainty in Multi-Agent Trajectory Forecasting View paper
- [15] MotionLM: Multi-Agent Motion Forecasting as Language Modeling View paper
- [16] Mart: Multiscale relational transformer networks for multi-agent trajectory prediction View paper
- [17] Reasoning multi-agent behavioral topology for interactive autonomous driving View paper
- [18] TrafficSim: Learning to Simulate Realistic Multi-Agent Behaviors View paper
- [19] Position: Simulating Society Requires Simulating Thought View paper
- [20] MotionDiffuser: Controllable Multi-Agent Motion Prediction Using Diffusion View paper
- [21] Online Multi-Agent Forecasting With Interpretable Collaborative Graph Neural Networks View paper
- [22] Sopermodel: Leveraging social perception for multi-agent trajectory prediction View paper

- [23] Distributed multi-agent reinforcement learning via interactive relationship construction and behavior prediction for confrontation scenario View paper
- [24] Collective intelligence in teams: Contextualizing collective intelligent behavior over time View paper
- [25] A hierarchical hybrid learning framework for multi-agent trajectory prediction View paper
- [26] Human-machine teamwork: an exploration of multi-agent systems, team cognition, and collective intelligence View paper
- [27] ElliottAgents: A natural language-driven multi-agent system for stock market analysis and prediction View paper
- [28] Recent Advances in Multi-Agent Human Trajectory Prediction: A Comprehensive Review View paper
- [29] Estimation of collective maneuvers through cooperative multi-agent planning View paper
- [30] Collaborative Uncertainty Benefits Multi-Agent Multi-Modal Trajectory Forecasting View paper
- [31] Generating intelligent agent behaviors in multi-agent game AI using deep reinforcement learning algorithm View paper
- [32] MACIE: Multi-Agent Causal Intelligence Explainer for Collective Behavior Understanding View paper
- [33] Modeling human behavior during emergency evacuation using intelligent agents: A multi-agent simulation approach View paper
- [34] Multi-Agent Hybrid Prediction in Autonomous Driving View paper
- [35] Cooperative Game-based Intelligent Actions Making for Constrained Multi-agent System View paper
- [36] Behavior Reasoning for Opponent Agents in Multi-Agent Learning Systems View paper
- [37] Agent-based computational modeling of emergent collective intelligence View paper
- [38] Evolvegraph: Multi-agent trajectory prediction with dynamic relational reasoning View paper
- [39] Collective Behaviors of Discrete-Time Multi-Agent Systems Over Signed Digraphs View paper
- [40] Cognitive multi-agent empowering mobile edge computing for resource caching and collaboration View paper
- [41] Multi-Agent cooperation based on opponent's behavior forecast View paper
- [42] ADAPT: Efficient Multi-Agent Trajectory Prediction with Adaptation View paper
- [43] Swarm intelligence, social force and multi-agent modeling of heroic altruism behaviors under collective risks View paper
- [44] Multi-agent driving behavior prediction across different scenarios with self-supervised domain knowledge View paper
- [45] Leveraging collective intelligence for behavioral prediction in signed social networks through evolutionary approach View paper
- [46] Multi-agent systems in predicting global financial crises View paper
- [47] Multi-Agent Reachability Calibration with Conformal Prediction View paper
- [48] Collective behavior simulation based on agent with artificial emotion View paper
- [49] Continual Multi-Agent Interaction Behavior Prediction With Conditional Generative Memory View paper
- [50] Advances in Behavior Prediction Techniques for Intelligent Systems View paper
- [51] MAGIS: LLM-Based Multi-Agent Framework for GitHub Issue Resolution View paper
- [52] Large language model based multi-agents: A survey of progress and challenges View paper
- [53] FinVision: A Multi-Agent Framework for Stock Market Prediction View paper
- [54] Internet of agents: Weaving a web of heterogeneous agents for collaborative intelligence View paper
- [55] A comprehensive survey on multi-agent cooperative decision-making: Scenarios, approaches, challenges and perspectives View paper
- [56] Comapoi: A collaborative multi-agent framework for next poi prediction bridging the gap between trajectory and language View paper
- [57] LLM-Powered Multi-Agent System for Automated Crypto Portfolio Management View paper
- [58] Drugagent: Multi-agent large language model-based reasoning for drug-target interaction prediction View paper
- [59] Why Should Next-Gen LLM Multi-Agent Systems Move Beyond Fixed Architectures to Dynamic, Input-Driven Graphs? View paper
- [60] Large Language Model-Empowered Interactive Load Forecasting View paper
- [61] How does the construal level affect consumers' intention to adopt product ratings and individual reviews? View paper
- [62] People watching: Social perception and the ensemble coding of bodies View paper
- [63] Towards Better Forecasting by Fusing Near and Distant Future Visions View paper
- [64] Abstract thinking facilitates aggregation of information. View paper
- [65] The effect of construal level on predictions of task duration View paper
- [66] â COGNITIVE MEMORY SYSTEMS: A UNIFIED FRAMEWORK FOR SEQUEN-TIAL INFORMATION PROCESSING AND LONG-TERM BEHAVIORAL PREDICTION View paper
- [67] Limits of a Deductive Construal of the Function of Scientific Theories A Comment View paper
- [68] LLM Generated Persona is a Promise with a Catch View paper
- [69] Personabench: Evaluating ai models on understanding personal information through accessing (synthetic) private user data View paper
- [70] Higher-Order Binding of Language Model Virtual Personas: a Study on Approximating Political Partisan Misperceptions View paper
- [71] PersonaGym: Evaluating Persona Agents and LLMs View paper
- [72] Co-persona: Leveraging LLMs and Expert Collaboration to Understand User Personas through Social Media Data Analysis View paper
- [73] Evaluating Persona Prompting for Question Answering Tasks View paper
- [74] Faithful Persona-based Conversational Dataset Generation with Large Language Models View paper
- [75] Virtual personas for language models via an anthology of backstories View paper
- [76] MCPDial: A Minecraft Persona-driven Dialogue Dataset View paper
- [77] BluePrint: A Social Media User Dataset for LLM Persona Evaluation and Training View paper