

LLM Generated Persona is a Promise with a Catch

Anonymous authors

Paper under double-blind review

Abstract

The use of large language models (LLMs) to simulate human behavior has gained significant attention, particularly through personas that approximate individual characteristics. Persona-based simulations hold promise for transforming disciplines that rely on population-level feedback, including social science, economic analysis, marketing research, and business operations. Traditional methods to collect realistic persona data face significant challenges: they are prohibitively expensive and logistically challenging due to privacy constraints, and often fail to capture multi-dimensional attributes, particularly subjective qualities. Consequently, synthetic persona generation with LLMs offers a scalable, cost-effective alternative. However, current approaches rely on ad hoc and heuristic generation techniques that do not guarantee methodological rigor or simulation precision, resulting in systematic biases in downstream tasks. Through extensive large-scale experiments including presidential election forecasts and general opinion surveys of the U.S. population, we reveal that these biases can lead to significant deviations from real-world outcomes. Based on the experimental results, we argue that **a rigorous and systematic science of persona generation is needed to ensure the reliability of LLM-driven simulations of human behavior**. We call for not only methodological innovations and empirical foundations but also interdisciplinary organizational and institutional support for the development of this field. To support further research and development in this area, we have open-sourced approximately one million generated personas, available for public access and analysis.

1 Introduction

Recent advances in LLM-driven simulations have enabled the creation of synthetic agents that approximate real-world populations' behaviors at scale, potentially transforming various fields such as social science (Manning et al., 2024), political science (Argyle et al., 2023), economics (Horton, 2023), marketing research (Sarstedt et al., 2024), clinical psychology (Wang et al., 2024), entertainment (Shao et al., 2023), personalization (Zollo et al.), and business applications such as recommendation systems and web testing (Tseng et al., 2024).

A central component of these simulations is the use of personas as shown in Figure 1: digital representations of individuals characterized by demographic, psychographic, and behavioral attributes. When provided with persona information, LLMs can generate responses

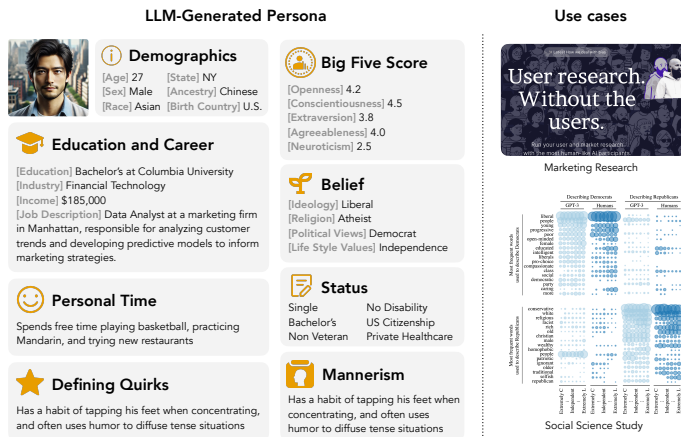


Figure 1: Left: An example of a LLM generated persona. Right: Applications of personas in the real world.

and simulate decision-making processes that aim to reflect those of real individuals (Park et al., 2024), serving as “silicon samples” to complement “human samples” for opinion testing (Argyle et al., 2023).

While “silicon samples” hold the potential to revolutionize societal-scale experimentation, a critical challenge remains: the systematic and reliable generation of persona profiles for targeted populations. Although large-scale collection of human data can maximize downstream simulation fidelity (Park et al., 2024), gathering such comprehensive data is far from trivial. It requires resource-intensive surveys and coordinated efforts that demand substantial financial investment, time, and operational support, while adhering to strict privacy regulations and ethical guidelines. On the other hand, existing large-scale real-world datasets (e.g., the U.S. Census (U.S. Census Bureau, 2024)) primarily contain marginal demographic information without capturing the joint distribution of multi-dimensional attributes. This fragmentation makes it impossible to reconstruct realistic, integrated personas that accurately mirror the complexity of real-world populations. Furthermore, these datasets often omit subjective attributes, such as lifestyle preferences or nuanced belief systems, due to privacy concerns or data collection limitations, despite their critical role in shaping individual opinions and values.

Meanwhile, LLM itself presents a viable solution to the above challenges by generating persona profiles directly in a cost-effective, efficient, and seemingly realistic manner. This nascent direction has received much attention recently from both academia and industry (Fröhling et al., 2024; Castricato et al., 2024; Ge et al., 2024; Schuller et al., 2024; Synthetic Users; CivicSync), where LLM-generated personas have been used to conduct surveys, marketing research, or even societal-scale simulations. Prior work has primarily focused on methodologies for scaling up the number of diverse personas, but there are no rigorous evaluations of their performance in different downstream applications or whether they faithfully capture a specific population’s opinion at scale.

Given the huge potential and interests of LLM-based persona generation and the lack of a scientific community that studies this problem, we argue that **a science of persona generation needs to be developed to fully realize the potential of LLM persona simulation**. Specifically, we observe that the current scalable persona generating methods are significantly biased and non-representative of the real-world distribution. *One example is that in a 2024 presidential election simulation, results generated by a specific type of LLM-based synthetic personas predict a Democratic sweep across all U.S. states* (see Figure 2). Since these generated personas can be widely used in applications ranging from opinion simulation to product testing, *their inherent biases can lead to harmful consequences, including skewed public decision-making, reinforcement of discrimination and stereotypes, and potential harm to minority groups*.

While collecting better persona data is vital and nontrivial, we believe that both data collection and persona generation and simulation methods are complementary rather than mutually exclusive. Regardless of the dataset’s size, generation methods are necessary to capture the complex, subjective aspects that traditional surveys often miss.

Therefore, we see an urgent need and call for a concerted effort from the AI community and interdisciplinary collaborations across social sciences to develop reliable methodologies, robust large-scale benchmarks, and broad community support to establish persona generation and silicon-sample simulation as a rigorous field. This paper is organized as follows:

1. *Systematization of Persona Generation Methods*. In Section 2, we categorize and systematize existing persona generation approaches, as summarized in Figure 3. This systematization lays the foundation for subsequent experimental evaluation and provides a structured framework for advancing the methodological rigor of persona generation.

2. *Current Persona Generation Practice Falls Short*: In Section 3, we present an extensive experimental evaluation of current persona generation practices. Using three commonly utilized persona types, we generate around 1,000,000 personas with six open-source LLMs and assess them across 500+ questions spanning diverse domains. The results demonstrate a substantial bias, amplified by the LLM-generated persona content. These findings critically undermine the validity of using LLMs as “silicon samples” and raise serious concerns about

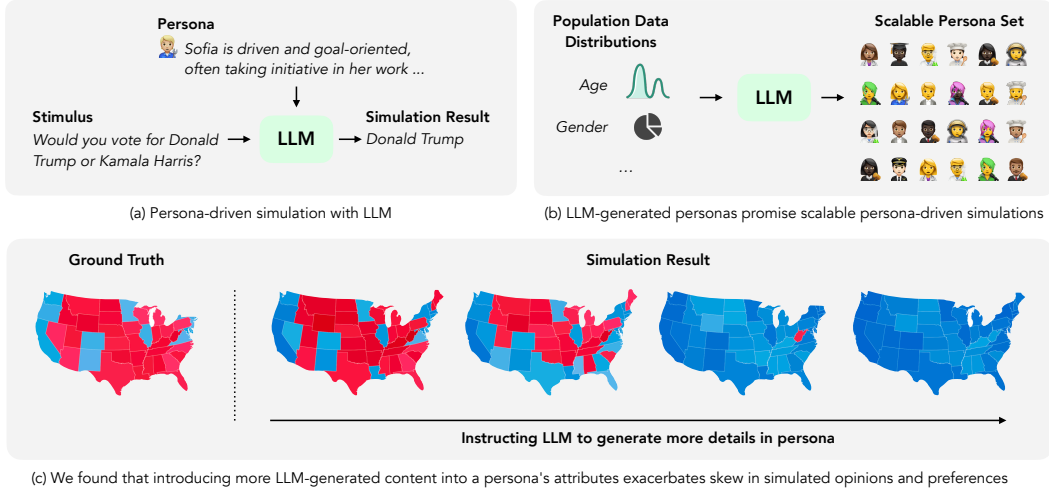


Figure 2: (a) Persona-driven simulation enables simulating human behaviors with LLMs. (b) Using LLM to generate personas promises scalable simulation of diverse population’s behaviors. (c) We caution that improper usage of LLMs as persona generators may lead to homogeneous results.

102 applying LLM-generated personas in social science research and business decision-making
 103 without proper caution. Section C provides a semantic analysis of the LLM-generated
 104 persona profiles.

105 3. *A Path Towards Rigorous Persona Generation:* In Section 4, we outline key directions for
 106 establishing a more scientific and systematic approach to persona generation, including: (i)
 107 developing refined frameworks for identifying essential persona attributes, (ii) building
 108 theoretical and empirical foundations for persona profile calibration, (iii) creating large-
 109 scale datasets and benchmarks to validate and refine persona generation methods, and (iv)
 110 advocating for interdisciplinary research and careful application-oriented consideration of
 111 risks and biases in persona generation and simulation.

112 Overall, we advocate for increased attention from the broader research community to realize
 113 the transformative potential of widespread LLM silicon sample adoption, enabling rapid,
 114 cost-effective, safe, and human-centric tests and innovation. As a first step toward this
 115 goal, we open-source all generated personas in our experiments to support and accelerate
 116 community efforts in this direction.

117 **A study of LLM persona generation bias rather than traditional LLM bias.** While biases
 118 in LLMs have been extensively studied in the literature (Gallegos et al., 2024; Ferrara,
 119 2023; Feng et al., 2023; Parrish et al., 2021; Bai et al., 2024; Lu et al., 2020; Li et al., 2020a),
 120 most prior work focuses on biases arising from *LLM simulation*. In contrast, we highlight
 121 a striking and underexplored source of bias introduced by *LLM persona generation* itself.
 122 Specifically, prior work in LLM bias has largely taken the persona generation process for
 123 granted, focusing instead on simulation outputs. In contrast, our study explicitly dissects
 124 the persona generation step, revealing that even if the simulation appears unbiased, the
 125 underlying generated personas may carry significant, unexamined biases that undermine
 126 their validity. Our findings indicate that this area requires significantly more attention
 127 to ensure fairness and representativeness in generated personas. For a comprehensive
 128 discussion of related work, please refer to Appendix A.

129 2 Persona-Driven Opinion Simulation

130 To highlight the shortcomings of current persona generation methods, we first establish the
 131 foundation for large-scale simulation experiments by implementing various commonly used

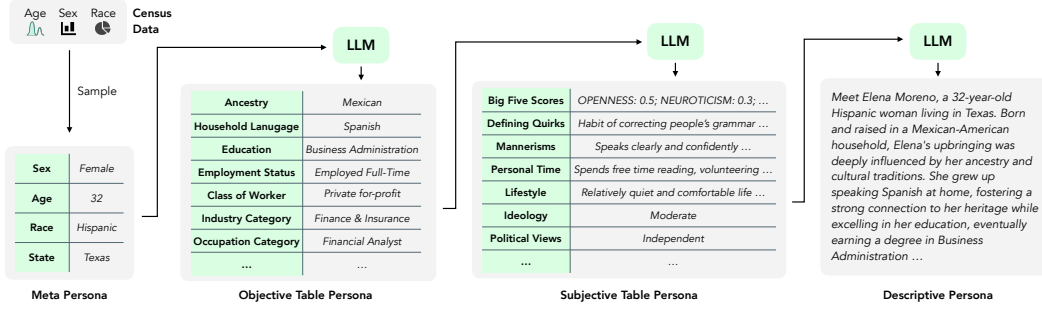


Figure 3: We categorize existing persona generation approaches into four tiers. Each tier adds more information generated by LLMs in generated personas to the previous tier.

132 persona generation strategies and simulation methods, drawing from both prior literature
133 and community practices.

134 2.1 Persona Generation

135 We build on prior literature to construct three distinct types of personas: **Meta Personas**,
136 **Tabular Personas**, and **Descriptive Personas**.

137 **Meta Persona.** Census data provides a foundational source for generating realistic per-
138 sonas due to its rich demographic information that reflects real-world population distri-
139 butions. However, census data typically presents marginal distributions (e.g., percentages
140 of different age groups and genders in New York), which limits its capacity for capturing
141 joint probabilities. This limitation can lead to incongruous combinations of attributes (e.g.,
142 a person earning \$500k annually but unable to afford health insurance). To address this,
143 we adopt methods from [Chang et al. \(2024\)](#) and [Argyle et al. \(2023\)](#), leveraging data with
144 joint distributions over key attributes, such as Age, Sex, Race, and State, provided by the
145 U.S. Census Bureau. We term these structured, sampled personas **Meta Personas** (No LLM
146 is involved in generating them). While these personas may lack diversity, they serve as a
147 robust foundation for generating more varied and realistic personas.

148 **Tabular Personas.** To enhance diversity and better capture nuanced differences for opinion
149 simulation, we extend the information encoded in **Meta Personas** using structured templates.
150 Following the proposed method of [Castricato et al. \(2024\)](#), we employ LLMs to fill in
151 additional attributes within these templates, conditioned on the generated meta personas.
152 We propose two variants:

153 1. *Objective Tabular Personas:* These personas expand **Meta Personas** by adding attributes
154 that can be objectively measured (e.g., Income, Education Level, Occupation, and Industry).
155 We align these entries with the Census Bureau’s predefined categories (detailed in the
156 appendix), which ensures consistency and mitigates bias. For example, occupations are
157 drawn from five predefined categories rather than generated freely by the LLM.

158 2. *Subjective Tabular Personas:* These personas extend *Objective Tabular Personas* by adding
159 subjective or less standardized attributes (e.g., Political Affiliation, Leisure Preferences).
160 Since these attributes lack comprehensive predefined categories, we prompt the LLM to
161 generate open-ended responses, guided by templates designed to maximize diversity while
162 preserving realism.

163 **Descriptive Personas.** The most flexible approach, **Descriptive Personas**, involves directly
164 prompting the LLM to create freeform personas conditioned on **Meta Personas**. This method
165 allows the generation of highly detailed and unconstrained personas by asking the LLM to
166 expand beyond structured attributes into narrative-style descriptions. While this approach
167 provides maximum freedom, it requires careful prompting and post-processing to ensure
168 the plausibility and coherence of the generated personas.

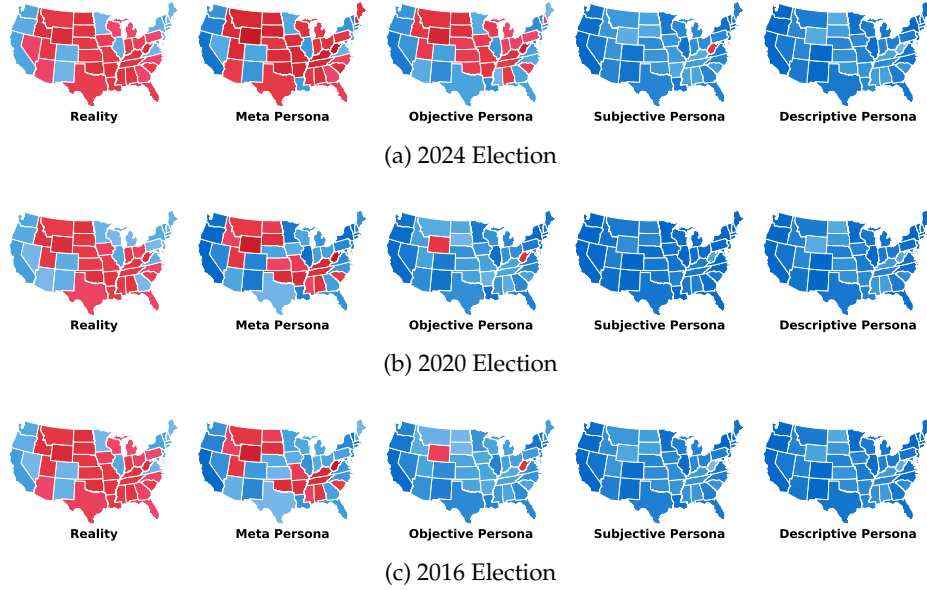


Figure 4: Persona-based simulations of elections 2016, 2020, and 2024.

Progressing from **Meta Personas** to **Tabular Personas** to **Descriptive Personas** entails a growing content generated by LLMs for diversity. This means that the persona types from left to right in Figure 3 gradually introduce richer, more varied attributes and open-ended descriptions, which theoretically can yield broader and more diverse opinion distributions. However, as we show later, **increasing the amount of LLM-generated persona content can exacerbate biases in the simulation results in various domains.**

To approximate opinion distributions and capture diversity at the U.S. national level, we first sample 1,000 personas per state using the corresponding joint distribution. We then generate 1,000 augmented personas for each of the persona types (**Meta**, **Tabular**, and **Descriptive**). During the opinion simulation, we present the language model with a short survey: a question accompanied by multiple choices, a persona description, and a template prompting the model to select the most suitable response based on the persona’s attributes. We explore several prompt variants that encourage neutrality and mitigate unwanted biases or stereotypes, but observe no substantial performance differences across these variants. The specific prompts and templates can be found in Appendix D

2.2 Simulating Opinions with LLMs

We treat LLMs as impartial world models and simulators. Building on recent work such as Park et al. (2024), we hypothesize that, when conditioned on realistic and adequate information, LLMs can truthfully simulate a persona’s opinion, effectively serving as digital twins. To simulate each persona’s opinion, we first provide high-level instructions guiding the LLM to remain unbiased and objective while faithfully reflecting the persona’s attributes. Next, we input a persona along with a multiple-choice question, instructing the LLM to directly select one of the provided answers. The final prompt, includes the simulation instruction, persona description, and the simulation question, is fed once into the language model. Additional details can be found in Appendix B.

3 Can Personas Simulate the Society?

We now present empirical findings from two main simulation scenarios: (1) U.S. elections from 2016, 2020, and 2024, and (2) general opinion surveys drawn from OpinionQA (Santurkar et al., 2023). Our analysis focuses on how relying on different persona types (with varying amounts of LLM-generated content) affects alignment with real-world data.

3.1 A Case Study with US Presidential Election Voting

Motivated by prior work in political and social opinion simulations [Argyle et al. \(2023\)](#), we explore how our persona-based approach can serve as a viable methodology for modeling society-level electoral outcomes. Given that most large language models (LLMs) are trained on data collected prior to the 2024 U.S. election, we first test on the 2024 presidential election to avoid potential confounds from post-2024 information.

We plot the map of each state in blue and red with different color scales to indicate the support rate. In Figure 4a, we show one particular experimentation where the persona model and the simulation model are both Llama 3.1 70B. Our results reveal that with increasing LLM-generated persona attributes, the simulation results tend to deviate more from real-world outcomes. Specifically, **Meta Personas** yield simulated voting distributions closest to reality data, whereas **Generative Personas** produce results most divergent from reality. Moreover, we observe that as the level of LLM-generated persona information rises, the simulated electorate shifts progressively toward left-leaning stances, ultimately culminating in a surprising scenario where every state appears to vote for the Democratic candidate.

To investigate whether historical election data—presumably included in the training sets—reduces model bias, we conduct similar simulations for the 2016 and 2020 election cycles. One might expect that since these events are part of the training data, the models would “remember” historical facts and produce more accurate electoral distributions. Contrary to this expectation, we observe a similar pattern of leftward drift, suggesting that awareness of past elections alone does not sufficiently counterbalance the influence of LLM-generated content.

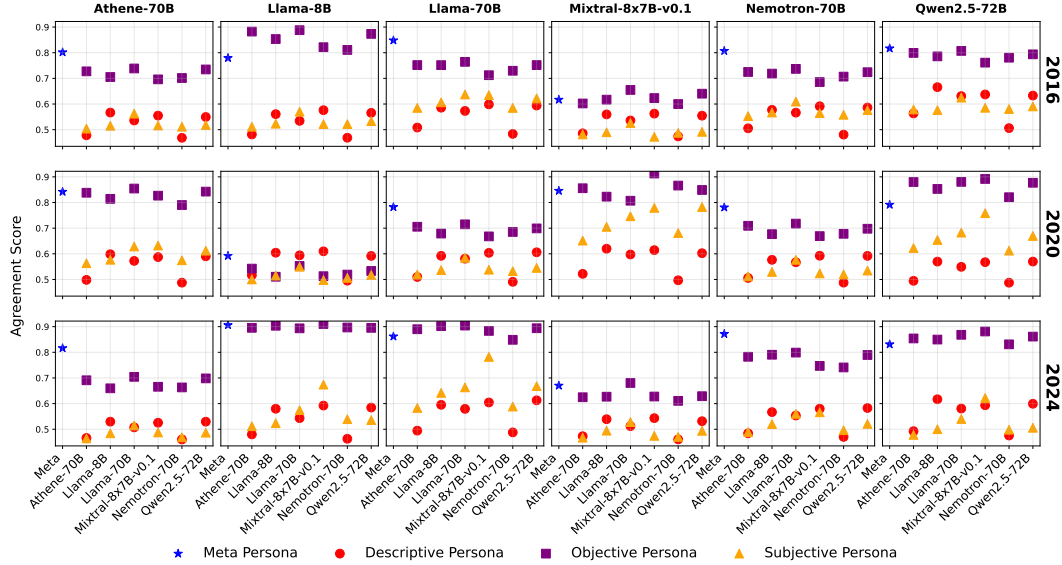


Figure 5: Alignment scores for cross-model simulation. Each column represents a simulation model, while the x-axis within each column corresponds to the persona generation model. The “meta” point is singular as it relies on sampling rather than generation.

Quantitative evaluation. To more rigorously compare our simulated results with ground truth, we adopt the alignment metric from [Santurkar et al. \(2023\)](#), using $1 - W(\hat{p}, p)$ as the alignment score. Here, $W(\hat{p}, p)$ denotes the Wasserstein distance between the simulated support rate (\hat{p}) and the actual voting rate (p), with higher alignment scores indicating results closer to real-world outcomes. Additionally, we perform *cross-simulation*, in which each LLM-generated persona set is simulated across all other language models. Our comprehensive assessment Figure 5 shows that this type of bias is not model-specific but rather universal in persona-driven simulations.

3.2 Exploring Bias Across Domains

We extended our analysis to assess whether the observed biases are limited to the political domain or span broader contexts.

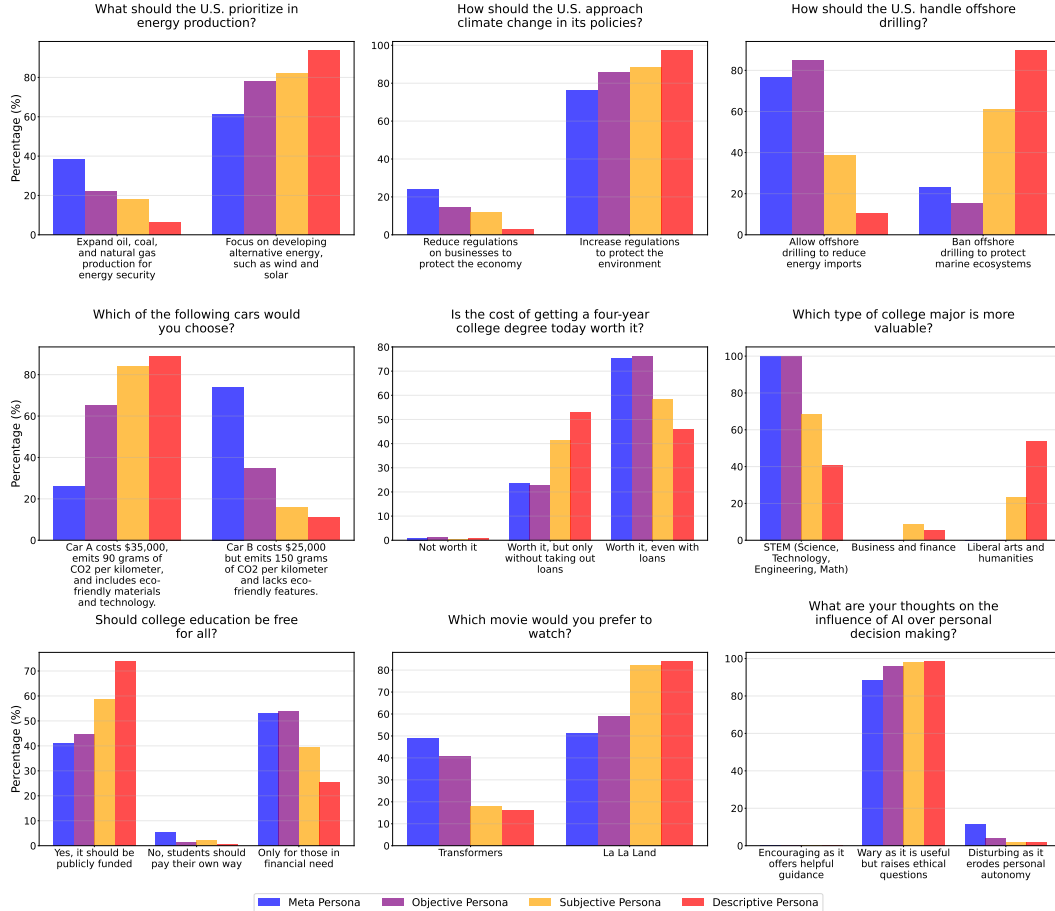


Figure 6: A preview of some specific topics that incurs interesting phenomenons

To this end, we designed 9 questions in Figure 6 covering five distinct domains: climate, consumer choices, education, entertainment, and technology.

For these questions, we do not have ground truth responses to measure alignment scores as in the political example. However, the plots provide a qualitative understanding of the general trend across different persona types' perspectives. Specifically, as persona types become more LLM-generated—from Meta to Descriptive—their perspectives shift from traditional to more progressive views. For example, LLM-generated personas increasingly prefer environmentally friendly cars over cheaper options, value liberal arts over STEM majors, and choose movies like "La La Land" over "Transformers."

While our experiments are not exhaustive, they reveal significant trends in how different persona models respond to diverse societal and individual decision-making scenarios. This raises concerns about deploying these "silicon samples" for social science and marketing research. For instance, the hidden preferences of silicon samples, such as favoring electric cars over gas-powered vehicles, could bias the decision-making of a car business owner. Similarly, an entertainment company might be misled for the market trend if these samples show a bias for musical or romantic movies over action films.

OpinionQA. To gain a more quantitative understanding, we evaluate various personas on the OpinionQA dataset (Santurkar et al., 2023), which provides a comprehensive evaluation of model behavior across a wide range of sensitive and broader topics.

The OpinionQA dataset collects questions and real US population’s responses from Pew research surveys, covering a diverse range of topics, including sensitive areas such as race, misinformation, and gender, as well as broader subjects like science, biomedical research, and automation. This allows us to examine how different persona generation methods influence simulated opinions on various social, economic, and personal issues. We use the first 500 questions in OpinionQA, covering 15 distinct topics (see Figure 7 for the list of topics). For each question, we simulate responses using each of the four persona types and calculate the alignment score between the simulated distribution of answers and the ground truth distribution provided in the OpinionQA dataset. We can clearly see that persona opinions diverge more on controversial topics.

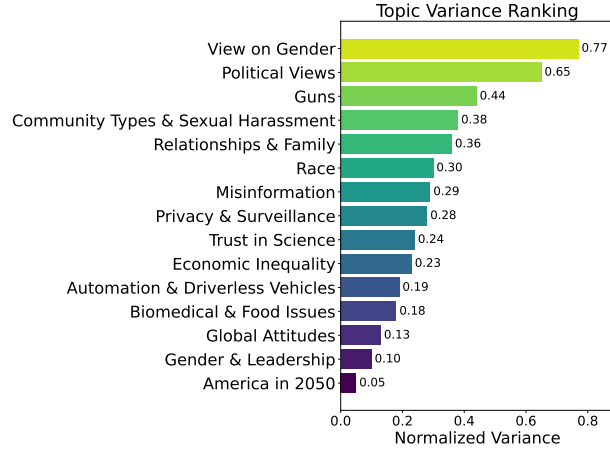


Figure 7: Rank of variance in alignment score among four persona types for the first 500 questions in OpinionQA based on topics.

4 Paths Forward: Toward a Scientific Approach to Persona

Our investigation reveals a critical limitation of current persona generation practices: none of the existing methods reliably produce realistic societal-level opinions. This underscores the urgent need for a rigorous and scientific approach to persona generation, one that moves beyond ad-hoc techniques and is borne out of principled methodologies. We put an initial analysis of persona in Appendix C. We further list a few potential directions and challenges that the community could work on.

Identifying essential information needed in a persona. A foundational challenge in persona-based simulation is identifying the essential information required for effective persona generation and how that information should be represented. The goal is to move beyond simply listing attributes to understanding what truly drives realistic simulation outcomes. Existing research offers conflicting evidence. While Argyle et al. (2023); Park et al. (2024); Salewski et al. (2023); Toubia et al. (2025) demonstrated that well-crafted conditioning can enable LLMs to simulate opinions aligned with real human responses, other studies such as Hu and Collier (2024); Gupta et al. (2023); Zheng et al. (2024); Beck et al. (2024) and our own experiments raise concerns about the efficacy and potential pitfalls of persona-based simulations. This discrepancy underscores the need to identify the crucial elements for effective persona-driven simulations, including which attributes are most important, such as demographic, psychographic (e.g., personality traits, values, attitudes, interests, lifestyles), behavioral (e.g., online activity), or contextual (e.g., social environment, current events), and the optimal format and prompting strategies for presenting that information to the LLM.

Calibrating LLM-generated personas towards real population. A parallel direction lies in accurately reconstructing realistic joint distributions of persona attributes from fragmented data sources, and subsequently calibrating these distributions to match a specific target population (Valliant et al., 2013). Even if we identify the crucial attributes for a given simulation, generating a population of personas requires sampling from the correct distributions. Existing datasets, such as the U.S. Census, often provide only marginal distributions of individual attributes (e.g., age, income, education level). This makes it impossible to sample from the true joint distribution. While Castricato et al. (2024) offers a first step by using

LLMs to filter out invalid attribute combinations sampled from marginal distributions, their method does not fully generalize to calibrate real-world joint distributions. Therefore, a crucial research direction is the development of robust sampling and calibration methods that can combine fragmented data and LLMs to accurately recover any target population.

Open-source benchmark and datasets. To accelerate progress in scientific understanding of the science of persona, we propose the creation of a large-scale, open-source benchmark dataset of rich and realistic persona profiles. Analogous to ImageNet [Deng et al. \(2009\)](#), a large-scale and open benchmark for persona generation would serve as a crucial resource for the research community. Specifically, this dataset would serve the following purposes: (i) a benchmark for evaluating the performance of different LLM-based persona generation methods; (ii) a training dataset such as [\(Toubia et al., 2025\)](#) for developing and testing new persona generation methods; and (iii) a high-quality profile library of diverse, realistic population-level personas suitable for direct use in “silicon sample” simulations. Constructing such a comprehensive dataset necessitates addressing data privacy concerns and requires a substantial investment of time and resources. However, we believe the potential benefits outweigh the effort.

Interdisciplinary research and broad community collaborations. Persona-driven simulation has the potential to revolutionize numerous fields, making it necessary for interdisciplinary collaboration between AI researchers and domain experts. Specifically, the ability to conduct rapid, cost-effective human experiments could be transformative for both academic fields, such as social science, economics, and political science (where human studies are crucial for advancement), and industry applications, like software and web design (where A/B testing is prevalent). We encourage broader exploration of silicon sample applications to better understand their potential and limitations, thereby advancing the field through collective community effort.

5 Discussions

While our paper highlights the current flaws and potential of LLM-generated personas for various applications, it is important to acknowledge that other approaches, such as personas derived from real-world data only, can achieve significantly higher prediction accuracy at the population level compared to synthetic ones. For example, [Park et al. \(2024\)](#) simulated over 1000 people based on their respective survey results and obtained 85% accuracy for prediction; similarly, [Toubia et al. \(2025\)](#) collected responses from over 2,000 participants across 500 questions, achieving 88% accuracy on held-out evaluation data. While we acknowledge the superior performance of real data, we have to point out that this approach fails to adequately address cost and ethical concerns. For instance, [Park et al. \(2024\)](#) could not release their dataset publicly, while [Toubia et al. \(2025\)](#) compensated thousands of participants with substantial payments, investing around \$100,000 total. In addition to that, a variety of papers have supported the role of synthetic data in model training and improving simulation ([Drechsler and Haensch, 2023](#); [Eckman et al., 2024](#)). Though not directly related to simulating human behaviors with LLMs, they provide foundational support for our position.

The limitations of real-world data further highlight the potential of synthetic personas, but only if they are constructed properly. This reinforces the need for a dedicated scientific community focused on understanding the capabilities and limitations of LLMs in persona generation and simulation, in order to fully unlock the potential of LLM-based persona modeling. This community must develop calibration methods for reducing biases, establish best practices for persona creation, and create benchmarks for assessing simulation results. Through interdisciplinary collaboration, we can envision a world where LLM-generated personas not only revolutionize various social science domains, but also transform our society with high-fidelity digital twins.

References

- Altera AL, Andrew Ahn, Nic Becker, Stephanie Carroll, Nico Christie, Manuel Cortes, Arda Demirci, Melissa Du, Frankie Li, Shuying Luo, et al. Project sid: Many-agent simulations toward ai civilization. *arXiv preprint arXiv:2411.00114*, 2024.
- Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023.
- Xuechunzi Bai, Angelina Wang, Ilia Sucholutsky, and Thomas L Griffiths. Measuring implicit bias in explicitly unbiased large language models. *arXiv preprint arXiv:2402.04105*, 2024.
- Tilman Beck, Hendrik Schuff, Anne Lauscher, and Iryna Gurevych. Sensitivity, performance, robustness: Deconstructing the effect of sociodemographic prompting. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2589–2615, 2024.
- Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. Language (technology) is power: A critical survey of “bias” in nlp. *arXiv preprint arXiv:2005.14050*, 2020.
- Graham Caron and Shashank Srivastava. Identifying and manipulating the personality traits of language models. *arXiv preprint arXiv:2212.10276*, 2022.
- Louis Castricato, Nathan Lile, Rafael Rafailov, Jan-Philipp Fränken, and Chelsea Finn. Persona: A reproducible testbed for pluralistic alignment, 2024. URL <https://arxiv.org/abs/2407.17387>.
- Serina Chang, Alicja Chaszczewicz, Emma Wang, Maya Josifovska, Emma Pierson, and Jure Leskovec. Llms generate structurally realistic social networks but overestimate political homophily, 2024. URL <https://arxiv.org/abs/2408.16629>.
- Kevin Chapuis, Patrick Taillandier, and Alexis Drogoul. Generation of synthetic populations in social simulations: a review of methods and practices. *Journal of Artificial Societies and Social Simulation*, 25(2), 2022.
- Jiangjie Chen, Xintao Wang, Rui Xu, Siyu Yuan, Yikai Zhang, Wei Shi, Jian Xie, Shuang Li, Ruihan Yang, Tinghui Zhu, et al. From persona to personalization: A survey on role-playing language agents. *arXiv preprint arXiv:2404.18231*, 2024.
- Yun-Shiuan Chuang, Nikunj Harlalka, Siddharth Suresh, Agam Goyal, Robert Hawkins, Sijia Yang, Dhavan Shah, Junjie Hu, and Timothy T Rogers. The wisdom of partisan crowds: Comparing collective intelligence in humans and llm-based agents. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 46, 2024.
- CivicSync. Civicsync. <https://civicsync.com/>. Accessed: 2025-01-30.
- Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.
- Jörg Drechsler and Anna-Carolina Haensch. 30 years of synthetic data. *arXiv preprint arXiv:2304.02107*, 2023.
- Stephanie Eckman, Barbara Plank, and Frauke Kreuter. Position: Insights from survey methodology can improve training data. *arXiv preprint arXiv:2403.01208*, 2024.
- Shangbin Feng, Chan Young Park, Yuhuan Liu, and Yulia Tsvetkov. From pretraining data to language models to downstream tasks: Tracking the trails of political biases leading to unfair nlp models. *arXiv preprint arXiv:2305.08283*, 2023.
- Emilio Ferrara. Should chatgpt be biased? challenges and risks of bias in large language models. *arXiv preprint arXiv:2304.03738*, 2023.

- 401 Leon Fröhling, Gianluca Demartini, and Dennis Assenmacher. Personas with attitudes:
402 Controlling llms for diverse data annotation. *arXiv preprint arXiv:2410.11745*, 2024.
- 403 Isabel O Gallegos, Ryan A Rossi, Joe Barrow, Md Mehrab Tanjim, Sungchul Kim, Franck
404 Dernoncourt, Tong Yu, Ruiyi Zhang, and Nesreen K Ahmed. Bias and fairness in large
405 language models: A survey. *Computational Linguistics*, pages 1–79, 2024.
- 406 Tao Ge, Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. Scaling synthetic data
407 creation with 1,000,000,000 personas, 2024. URL <https://arxiv.org/abs/2406.20094>.
- 408 Shashank Gupta, Vaishnavi Shrivastava, Ameet Deshpande, Ashwin Kalyan, Peter Clark,
409 Ashish Sabharwal, and Tushar Khot. Bias runs deep: Implicit reasoning biases in persona-
410 assigned llms. *arXiv preprint arXiv:2311.04892*, 2023.
- 411 John J Horton. Large language models as simulated economic agents: What can we learn
412 from homo silicus? Technical report, National Bureau of Economic Research, 2023.
- 413 Tiancheng Hu and Nigel Collier. Quantifying the persona effect in llm simulations. *arXiv
414 preprint arXiv:2402.10811*, 2024.
- 415 Jen-tse Huang, Wenxuan Wang, Eric John Li, Man Ho Lam, Shujie Ren, Youliang Yuan,
416 Wenxiang Jiao, Zhaopeng Tu, and Michael Lyu. On the humanity of conversational ai:
417 Evaluating the psychological portrayal of llms. In *The Twelfth International Conference on
418 Learning Representations*, 2023.
- 419 Bernard J Jansen, Joni Salminen, Soon-gyo Jung, and Kathleen Guan. *Data-driven personas*.
420 Springer Nature, 2022.
- 421 Kyuwon Lee, Simone Paci, Jeongmin Park, Hye Young You, and Sylvan Zheng. Ap-
422 plications of gpt in political science research. May. [https://hyeyoungyou.com/wp-
423 content/uploads/2024/05/gpt-polisci.pdf](https://hyeyoungyou.com/wp-content/uploads/2024/05/gpt-polisci.pdf), 2024.
- 424 Tao Li, Tushar Khot, Daniel Khashabi, Ashish Sabharwal, and Vivek Srikumar. Unqovering
425 stereotyping biases via underspecified questions. *arXiv preprint arXiv:2010.02428*, 2020a.
- 426 Zheng Li, Yue Zhao, and Jialin Fu. Sync: A copula based framework for generating synthetic
427 data from aggregated sources. In *2020 International Conference on Data Mining Workshops
428 (ICDMW)*, pages 571–578. IEEE, 2020b.
- 429 Steven Loria. textblob documentation. *Release 0.15, 2*, 2018.
- 430 Kaiji Lu, Piotr Mardziel, Fangjing Wu, Preetam Amancharla, and Anupam Datta. Gender
431 bias in neural natural language processing. *Logic, language, and security: essays dedicated to
432 Andre Scedrov on the occasion of his 65th birthday*, pages 189–202, 2020.
- 433 Bolei Ma, Xinpeng Wang, Tiancheng Hu, Anna-Carolina Haensch, Michael A Hedderich,
434 Barbara Plank, and Frauke Kreuter. The potential and challenges of evaluating attitudes,
435 opinions, and values in large language models. *arXiv preprint arXiv:2406.11096*, 2024.
- 436 Benjamin S Manning, Kehang Zhu, and John J Horton. Automated social science: Language
437 models as scientist and subjects. Technical report, National Bureau of Economic Research,
438 2024.
- 439 Xianghe Pang, Shuo Tang, Rui Ye, Yuxin Xiong, Bolun Zhang, Yanfeng Wang, and Siheng
440 Chen. Self-alignment of large language models via monopolylogue-based social scene
441 simulation. *arXiv preprint arXiv:2402.05699*, 2024.
- 442 Joon Sung Park, Joseph O’Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and
443 Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. In
444 *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages
445 1–22, 2023.
- 446 Joon Sung Park, Carolyn Q Zou, Aaron Shaw, Benjamin Mako Hill, Carrie Cai, Meredith
447 Ringel Morris, Robb Willer, Percy Liang, and Michael S Bernstein. Generative agent
448 simulations of 1,000 people. *arXiv preprint arXiv:2411.10109*, 2024.

- 449 Alicia Parrish, Angelica Chen, Nikita Nangia, Vishakh Padmakumar, Jason Phang, Jana
450 Thompson, Phu Mon Htut, and Samuel R Bowman. Bbq: A hand-built bias benchmark
451 for question answering. *arXiv preprint arXiv:2110.08193*, 2021.
- 452 Weihong Qi, Hanjia Lyu, and Jiebo Luo. Representation bias in political sample simulations
453 with large language models. *arXiv preprint arXiv:2407.11409*, 2024.
- 454 Leonard Salewski, Stephan Alaniz, Isabel Rio-Torto, Eric Schulz, and Zeynep Akata. In-
455 context impersonation reveals large language models’ strengths and biases. *Advances in*
456 *neural information processing systems*, 36:72044–72057, 2023.
- 457 Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cino Lee, Percy Liang, and Tatsunori
458 Hashimoto. Whose opinions do language models reflect? In *International Conference on*
459 *Machine Learning*, pages 29971–30004. PMLR, 2023.
- 460 Marko Sarstedt, Susanne J Adler, Lea Rau, and Bernd Schmitt. Using large language
461 models to generate silicon samples in consumer and marketing research: Challenges,
462 opportunities, and guidelines. *Psychology & Marketing*, 41(6):1254–1270, 2024.
- 463 Andreas Schuller, Doris Janssen, Julian Blumenröther, Theresa Maria Probst, Michael
464 Schmidt, and Chandan Kumar. Generating personas using llms and assessing their
465 viability. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*,
466 pages 1–7, 2024.
- 467 Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. Character-llm: A trainable agent for
468 role-playing. *arXiv preprint arXiv:2310.10158*, 2023.
- 469 Emily Sheng, Kai-Wei Chang, Premkumar Natarajan, and Nanyun Peng. Societal biases in
470 language generation: Progress and challenges. *arXiv preprint arXiv:2105.04054*, 2021.
- 471 Synthetic Users. Synthetic users. <https://www.syntheticusers.com/>. Accessed: 2025-01-30.
- 472 Olivier Toubia, George Gui, Tianyi Peng, Daniel Merlau, Ang Li, and Haozhe Chen. Twin-
473 2k-500: A dataset for building digital twins of over 2,000 people based on their an-
474 swers to over 500 questions, May 2025. URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5265253. SSRN Working Paper.
- 476 Yu-Min Tseng, Yu-Chao Huang, Teng-Yun Hsiao, Yu-Ching Hsu, Jia-Yin Foo, Chao-Wei
477 Huang, and Yun-Nung Chen. Two tales of persona in llms: A survey of role-playing and
478 personalization. *arXiv preprint arXiv:2406.01171*, 2024.
- 479 U.S. Census Bureau, 2024. URL <https://www.census.gov/data/datasets.html>.
- 480 Richard Valliant, Jill A Dever, and Frauke Kreuter. *Practical tools for designing and weighting*
481 *survey samples*, volume 1. Springer, 2013.
- 482 Ruiyi Wang, Stephanie Milani, Jamie C Chiu, Jiayin Zhi, Shaun M Eack, Travis Labrum,
483 Samuel M Murphy, Nev Jones, Kate Hardy, Hong Shen, et al. Patient- $\{\Psi\}$: Using large
484 language models to simulate patients for training mental health professionals. *arXiv*
485 *preprint arXiv:2405.19660*, 2024.
- 486 Zekun Moore Wang, Zhongyuan Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou,
487 Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, et al. Rolellm: Bench-
488 marking, eliciting, and enhancing role-playing abilities of large language models. *arXiv*
489 *preprint arXiv:2310.00746*, 2023.
- 490 Benfeng Xu, An Yang, Junyang Lin, Quan Wang, Chang Zhou, Yongdong Zhang, and
491 Zhendong Mao. Expertprompting: Instructing large language models to be distinguished
492 experts. *arXiv preprint arXiv:2305.14688*, 2023.
- 493 Ziyi Yang, Zaibin Zhang, Zirui Zheng, Yuxian Jiang, Ziyue Gan, Zhiyu Wang, Zijian Ling,
494 Jinsong Chen, Martz Ma, Bowen Dong, et al. Oasis: Open agents social interaction
495 simulations on one million agents. *arXiv preprint arXiv:2411.11581*, 2024.

- 496 Jinman Zhao, Zifan Qian, Linbo Cao, Yining Wang, and Yitian Ding. Bias and toxicity in
497 role-play reasoning. *arXiv preprint arXiv:2409.13979*, 2024.
- 498 Mingqian Zheng, Jiaxin Pei, Lajanugen Logeswaran, Moontae Lee, and David Jurgens.
499 When “a helpful assistant” is not really helpful: Personas in system prompts do not
500 improve performances of large language models. In *Findings of the Association for Compu-*
501 *tational Linguistics: EMNLP 2024*, pages 15126–15154, 2024.
- 502 Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-
503 Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, et al. Sotopia: Interactive
504 evaluation for social intelligence in language agents. *arXiv preprint arXiv:2310.11667*, 2023.
- 505 Thomas P Zollo, Andrew Wei Tung Siah, Naimeng Ye, Ang Li, and Hongseok Namkoong.
506 Personallm: Tailoring llms to individual preferences. In *The Thirteenth International*
507 *Conference on Learning Representations*.

508 A Related Work

509 **Persona generation.** Synthetic persona generation (Tseng et al., 2024; Chen et al., 2024)
 510 involves creating artificial user profiles for a variety of downstream tasks such as political
 511 simulations Lee et al. (2024) and economic research Horton (2023). Traditional approaches
 512 often rely solely on real-world datasets, such as census and survey data, to ensure sta-
 513 tistical validity while addressing scalability and diversity challenges (Argyle et al., 2023;
 514 Chang et al., 2024; Huang et al., 2023; Xu et al., 2023). For example, agent-based modeling
 515 studies generate synthetic populations by sampling attributes (age, income, etc.) from
 516 census or survey distributions, ensuring the aggregate matches real demographics Chapuis
 517 et al. (2022). Common techniques include clustering user data, factor analysis, and matrix
 518 decomposition to identify archetypal personas from large datasets Jansen et al. (2022). Gen-
 519 erative personas, on the other hand, leverage LLMs to automatically scale up the number
 520 of distinct personas. Generative tabular personas attempt to balance demographic validity
 521 with enriched attributes by either employing LLMs (Castricato et al., 2024) or probabilistic
 522 methods (Li et al., 2020b) to fill in extra characteristics. Descriptive personas leverage LLMs
 523 to directly generate free-form personas conditioned on text input Ge et al. (2024). While
 524 lacking validity, descriptive personas serve as the most common approach for large-scale
 525 persona generation in the industry.

526 **LLM Driven Simulation.** LLMs have enabled new possibilities in simulating human-like
 527 interactions and societal behaviors. Early works demonstrated feasibility in diverse settings,
 528 from digital twins of real individuals (Park et al., 2024; 2023) and role-playing of famous
 529 characters (Wang et al., 2023; Shao et al., 2023; Chen et al., 2024) or specific personalities
 530 (Caron and Srivastava, 2022), to society simulations (Argyle et al., 2023; Chuang et al., 2024;
 531 Yang et al., 2024; AL et al., 2024; Zhou et al., 2023; Pang et al., 2024). However, though
 532 those works provide foundational support for our research, they primarily focus on demon-
 533 strating feasibility rather than conducting rigorous, large-scale simulations across diverse
 534 populations. In our work, instead, we systematically examined how diverse personas
 535 influence the outcomes of LLM-driven simulations. Prior studies, such as (Park et al., 2024;
 536 Argyle et al., 2023), have shown that conditioning simulations on well-crafted personas can
 537 yield results that closely approximate real-world behavior. However, while these works
 538 validate the feasibility of such simulations, none of them provide rigorous evidence that
 539 conditioning on the right persona guarantees accurate behavior. As an orthogonal direction,
 540 future research should rigorously refine simulation methodologies to improve both realism
 541 and scalability.

542 **LLM Bias.** Bias in LLMs has been widely studied Gallegos et al. (2024); Ferrara (2023); Feng
 543 et al. (2023); Parrish et al. (2021); Bai et al. (2024); Lu et al. (2020); Li et al. (2020a); Blodgett
 544 et al. (2020); Sheng et al. (2021); Ma et al. (2024), including in LLM-driven simulations. Recent
 545 work shows that assigning personas to LLMs can expose hidden biases and harm reasoning
 546 Gupta et al. (2023), and that role-play increases both response quality and bias risk, especially
 547 with dynamic roles Zhao et al. (2024). In political simulations, LLMs overemphasize political
 548 homophily Chang et al. (2024) and favor English-speaking, democratic systems Qi et al.
 549 (2024). Distinct from prior work, our study investigates bias brought by persona generation,
 550 rather than the simulation itself.

551 B Additional Simulation Details

552 B.1 Language Model Selection

553 We evaluate six open-source language models: *Athene 70B*, *Llama 3.1 8B*, *Llama 3.1 70B*,
 554 *Mistral-8x7B Instruct V0.1*, *Nemotron 70B*, and *Qwen 2.5B*. Our selection is driven by two key
 555 considerations:

- 556 1. **Alignment strategy.** Some models, such as the *Llama* family, are purely instruction-
 557 tuned, while others (e.g., *Athene* and *Nemotron*) undergo additional refinement
 558 through Reinforcement Learning from Human Feedback (RLHF).

Qualitative word cloud analysis of the language patterns in descriptive personas supports these quantitative findings (Figure 9b). The prevalence of positively-valenced terms (“love”, “proud”) alongside community-oriented words (“family”, “community”) indicates LLMs systematically generate personas with optimistic life outlooks and strong social connections. Terms related to achievement and stability (“education”, “work”, “cultural”, “heritage”) further reinforce this positive framing. Notably absent are terms reflecting life challenges, social difficulties, or negative experiences, suggesting LLMs may be systematically avoiding less favorable characterizations. These findings help explain the previously observed biases in simulation outcomes: the combination of positive sentiment and highly emotional characterizations in the underlying personas naturally leads to skewed responses, particularly on social issues where emotional reasoning may play a stronger role.

C.1 Feedback Collection

Instead of computing token-level log probabilities for each choice, we aggregate the counts of each selected choice across the simulated population, leveraging the instruction-following capabilities of modern LLMs.

D Persona Generation and Simulation Prompts

D.1 Persona Generation System Prompts

You are an AI assistant specialized in detailed and unbiased persona generation for opinion simulations. Your task is to generate a specific, realistic, and diverse persona based on the provided demographic information and fill in a comprehensive JSON template.

D.2 Objective Tabular Persona

Persona Generation Instructions

```
### INSTRUCTIONS ###
1. You will be provided with a persona meta file that has the core
   demographic information of a person.
2. You will also be provided with a final persona template. Your task is
   to create a detailed, concrete persona that is fully consistent with
   ALL features in the given metadata by filling the template.
3. Elaborate on all metadata points, providing specific details that
   flesh out the persona while remaining true to the given information.
4. For all of the features in the metadata, you will be provided with a
   range of values in the VALUE RANGES AND CATEGORIES section below.
   Select one of the values for each of the features. DO NOT ADD EXTRA
   INFORMATION OR ELABORATION TO THE VALUES. DO NOT ADD EXTRA FEATURES
   TO THE TEMPLATE.
5. IMPORTANT: Place your entire response in the ### PERSONA GENERATION
   ### section below. Start your response with 'Persona:' and then
   provide only the persona description. Do not include any other
   prefixes, headers, or additional text.

### VALUE RANGES AND CATEGORIES ###

ANCESTRY: [
  "British",
  "Irish",
  "German",
  "Italian",
  "Polish",
  "French",
  "Norwegian",
  "Dutch",
```

```

632     "Swedish",
633     "Russian",
634     "Chinese",
635     "Filipino",
636     "Asian Indian",
637     "Vietnamese",
638     "Korean",
639     "Japanese",
640     "Mexican",
641     "Puerto Rican",
642     "Cuban",
643     "African American",
644     "West Indian",
645     "Arab",
646     "American Indian"
647 ]
648
649 HOUSEHOLD_LANGUAGE: [English, Spanish, Other Indo-European,
650     Asian/Pacific Islander languages, Other]
651
652 EDUCATION: [Less than HS, HS Graduate, Some College, Bachelor's,
653     Graduate Degree]
654
655 EMPLOYMENT_STATUS: [Employed, Unemployed, Not in Labor Force]
656
657 CLASS_OF_WORKER: [Private, Government, Self-employed, Unpaid family
658     worker]
659
660 INDUSTRY_CATEGORY: [
661     "Management, business, science, and arts occupations",
662     "Service occupations",
663     "Sales and office occupations",
664     "Natural resources, construction, and maintenance occupations",
665     "Production, transportation, and material moving occupations"
666 ]
667
668 OCCUPATION_CATEGORY: [
669     "Management, business, science, and arts occupations": [
670         "Management, business, and financial occupations",
671         "Computer, engineering, and science occupations",
672         "Education, legal, community service, arts, and media
673             occupations",
674         "Healthcare practitioner and technical occupations"
675     ],
676     "Service occupations": [
677         "Healthcare support occupations",
678         "Protective service occupations",
679         "Food preparation and serving related occupations",
680         "Building and grounds cleaning and maintenance occupations",
681         "Personal care and service occupations"
682     ],
683     "Sales and office occupations": [
684         "Sales and related occupations",
685         "Office and administrative support occupations"
686     ],
687     "Natural resources, construction, and maintenance occupations": [
688         "Farming, fishing, and forestry occupations",
689         "Construction and extraction occupations",
690         "Installation, maintenance, and repair occupations"
691     ],
692     "Production, transportation, and material moving occupations": [
693         "Production occupations",
694         "Transportation occupations",
695         "Material moving occupations"
696 ]

```

```

697 ]
698
699 INCOME: [Range $0-$1,000,000 annually]
700
701 MARITAL_STATUS: [Never Married, Married, Divorced, Widowed, Separated]
702
703 HOUSEHOLD_TYPE: [Family, Non-family]
704
705 PLACE_OF_BIRTH: [US State, Foreign Country]
706
707 VETERAN_STATUS: [Veteran, Non-veteran]
708
709 DISABILITY: [None, Physical, Mental, Both]
710
711 HEALTH_INSURANCE: [Private, Public, None]
712
713 ### RESPONSE FORMAT ###
714 Persona: [The completed FINAL PERSONA TEMPLATE]
715
716 ### PERSONA METADATA ###
717 {METADATA}
718
719 ### FINAL PERSONA TEMPLATE ###
720 {TEMPLATE}
721
722 ### PERSONA GENERATION ###
723

```

724 Template

```

725 "AGE": "",
726 "SEX": "",
727 "RACE": "",
728 "STATE": "",
729 "ANCESTRY": "",
730 "HOUSEHOLD_LANGUAGE": "",
731 "EDUCATION": "",
732 "EMPLOYMENT_STATUS": "",
733 "CLASS_OF_WORKER": "",
734 "INDUSTRY_CATEGORY": "",
735 "OCCUPATION_CATEGORY": "",
736 "INCOME": "",
737 "MARITAL_STATUS": "",
738 "HOUSEHOLD_TYPE": "",
739 "FAMILY_PRESENCE_AND_AGE": "",
740 "PLACE_OF_BIRTH": "",
741 "CITIZENSHIP": "",
742 "VETERAN_STATUS": "",
743 "DISABILITY": "",
744 "HEALTH_INSURANCE": ""
745

```

747 D.3 Subjective Tabular Persona

748 Persona Generation Instructions

```

749 ### INSTRUCTIONS ###
750
751 1. You will be provided with a persona meta file that has the core
752    demographic information of a person.
753 2. You will also be provided with a final persona template. Your task is
754    to create a detailed, concrete persona that is fully consistent with
755    ALL features in the given metadata by filling the template.
756 3. Elaborate on all metadata points, providing specific details that
757    flesh out the persona while remaining true to the given information.
758 4. For some of the features, you will be provided with a range of values
759    in the VALUE RANGES AND CATEGORIES section below. Select one of the

```

```

760     values for each of the features. DO NOT ADD EXTRA INFORMATION for
761     those features.
762 5. For the other features, fill in the values with a reasonable and
763     succinct description. Be as objective as possible.
764 6. IMPORTANT: Place your entire response in the ### PERSONA GENERATION
765     ### section below. Start your response with 'Persona:' and then
766     provide only the persona description. Do not include any other
767     prefixes, headers, or additional text.
768
769 ### VALUE RANGES AND CATEGORIES ###
770
771 ANCESTRY: [
772     "British",
773     "Irish",
774     "German",
775     "Italian",
776     "Polish",
777     "French",
778     "Norwegian",
779     "Dutch",
780     "Swedish",
781     "Russian",
782     "Chinese",
783     "Filipino",
784     "Asian Indian",
785     "Vietnamese",
786     "Korean",
787     "Japanese",
788     "Mexican",
789     "Puerto Rican",
790     "Cuban",
791     "African American",
792     "West Indian",
793     "Arab",
794     "American Indian"
795 ]
796
797 HOUSEHOLD_LANGUAGE: [English, Spanish, Other Indo-European,
798     Asian/Pacific Islander languages, Other]
799
800 EDUCATION: [Less than HS, HS Graduate, Some College, Bachelor's,
801     Graduate Degree]
802
803 EMPLOYMENT_STATUS: [Employed, Unemployed, Not in Labor Force]
804
805 CLASS_OF_WORKER: [Private, Government, Self-employed, Unpaid family
806     worker]
807
808 INDUSTRY_CATEGORY: [
809     "Management, business, science, and arts occupations",
810     "Service occupations",
811     "Sales and office occupations",
812     "Natural resources, construction, and maintenance occupations",
813     "Production, transportation, and material moving occupations"
814 ]
815
816 OCCUPATION_CATEGORY: [
817     "Management, business, science, and arts occupations": [
818         "Management, business, and financial occupations",
819         "Computer, engineering, and science occupations",
820         "Education, legal, community service, arts, and media
821             occupations",
822         "Healthcare practitioner and technical occupations"
823     ],
824     "Service occupations": [

```



```

825         "Healthcare support occupations",
826         "Protective service occupations",
827         "Food preparation and serving related occupations",
828         "Building and grounds cleaning and maintenance occupations",
829         "Personal care and service occupations"
830     ],
831     "Sales and office occupations": [
832         "Sales and related occupations",
833         "Office and administrative support occupations"
834     ],
835     "Natural resources, construction, and maintenance occupations": [
836         "Farming, fishing, and forestry occupations",
837         "Construction and extraction occupations",
838         "Installation, maintenance, and repair occupations"
839     ],
840     "Production, transportation, and material moving occupations": [
841         "Production occupations",
842         "Transportation occupations",
843         "Material moving occupations"
844     ]
845 ]
846
847 INCOME: [Range $0-$1,000,000 annually]
848
849 MARITAL_STATUS: [Never Married, Married, Divorced, Widowed, Separated]
850
851 HOUSEHOLD_TYPE: [Family, Non-family]
852
853 PLACE_OF_BIRTH: [US State, Foreign Country]
854
855 VETERAN_STATUS: [Veteran, Non-veteran]
856
857 DISABILITY: [None, Physical, Mental, Both]
858
859 HEALTH_INSURANCE: [Private, Public, None]
860
861 IDEOLOGY: [Very Liberal, Liberal, Moderate, Conservative, Very
862     Conservative]
863
864 POLITICAL_VIEWS: [Democrat, Republican, Independent, Other]
865
866 ### RESPONSE FORMAT ###
867 Persona: [The completed FINAL PERSONA TEMPLATE]
868
869 ### PERSONA METADATA ###
870 {METADATA}
871
872 ### FINAL PERSONA TEMPLATE ###
873 {TEMPLATE}
874
875 ### PERSONA GENERATION ###
876

```

877 Template

```

878 "AGE": "",
879 "SEX": "",
880 "RACE": "",
881 "STATE": "",
882 "ANCESTRY": "",
883 "HOUSEHOLD_LANGUAGE": "",
884 "EDUCATION": "",
885 "EMPLOYMENT_STATUS": "",
886 "CLASS_OF_WORKER": "",
887 "INDUSTRY_CATEGORY": "",
888 "OCCUPATION_CATEGORY": "",
889

```

```

890 "DETAILED_JOB_DESCRIPTION": "",
891 "INCOME": "",
892 "MARITAL_STATUS": "",
893 "HOUSEHOLD_TYPE": "",
894 "FAMILY_PRESENCE_AND_AGE": "",
895 "PLACE_OF_BIRTH": "",
896 "CITIZENSHIP": "",
897 "VETERAN_STATUS": "",
898 "DISABILITY": "",
899 "HEALTH_INSURANCE": "",
900 "BIG_FIVE_SCORES": {
901   "OPENNESS": "",
902   "CONSCIENTIOUSNESS": "",
903   "EXTRAVERSION": "",
904   "AGREEABLENESS": "",
905   "NEUROTICISM": ""
906 },
907 "DEFINING_QUIRKS": "",
908 "MANNERISMS": "",
909 "PERSONAL_TIME": "",
910 "LIFESTYLE": "",
911 "IDEOLOGY": "",
912 "POLITICAL_VIEWS": "",
913 "RELIGION": "",
914 "COGNITIVE_DIFFICULTY": "",
915 "ABILITY_TO_SPEAK_ENGLISH": "",
916 "VISION_DIFFICULTY": "",
917 "FERTILITY": "",
918 "HEARING_DIFFICULTY": ""
919

```

920 D.4 Descriptive Persona

921 Persona Generation Instructions

```

922 ### INSTRUCTIONS ###
923
924 1. You will be provided with a persona meta file that has the core
925    demographic information of a person.
926 2. Your task is to create a detailed, diverse, and vivid persona that is
927    fully consistent with ALL features in the given metadata.
928 3. Elaborate on all metadata points, providing specific details that
929    flesh out the persona while remaining true to the given information.
930 4. For any ranges or categories provided in the metadata, select and
931    specify exact values or details within those ranges/categories.
932 5. Ensure diversity in perspectives, backgrounds, and personality
933    traits. Provide enough specific details to make the persona feel
934    real and three-dimensional.
935 6. Maintain diversity by acknowledging various experiences within the
936    demographic group, but commit to specific details for this
937    individual persona.
938 7. IMPORTANT: Place your entire response in the ### PERSONA GENERATION
939    ### section below. Start your response with 'Persona:' and then
940    provide only the persona description. Do not include any other
941    prefixes, headers, or additional text.
942
943 ### RESPONSE FORMAT ###
944 Persona: [A detailed, vivid, and diverse description of a specific
945    individual. Ensure all details are consistent with and elaborate
946    upon the provided metadata.]
947
948 ### PERSONA METADATA ###
949 {METADATA}
950
951 ### PERSONA GENERATION ###
952

```

D.5 Opinion Simulation Prompts

```
You are an AI assistant tasked with generating realistic opinions based
on a given persona and a specific topic.

### TASK ###
You will simulate a persona answering a multiple-choice opinion
question. Select the answer that best matches your persona's
viewpoint and interests.

### GUIDELINES ###
1. Be Faithful to the Persona: Ensure your answer is consistent with the
persona's data.
2. Focus on Relevant Aspects: Center your reasoning on the relevant
factors that would influence the persona's opinion on that topic.
3. Be Objective: Avoid injecting personal bias or overly politically
correct views that may not align with the persona's standpoint.

### INSTRUCTIONS ###
- Choose ONE option (A, B, C, or D depending on the number of options)
that best fits the persona
- If multiple answers are possible, randomly select based on their
probability
- Always pick an option, even in unclear cases - treat it as a
forced-choice survey
- Output format: 'Answer: [Letter]' only, no explanation needed

### PERSONA ###
{PERSONA}

### QUESTION ###
{QUESTION}

### YOUR RESPONSE ###
```