

CLASP: A Closed-Loop Behavior Alignment Framework for Social Media Agent Simulation

Laisong Li, Shanfeng Wang, Wei Xie*,

Zhenhua Wang, Zhenhao Luo*, Baosheng Wang

College of Computer Science and Technology, National University of Defense Technology
{lilaisong, wsf, xiewei, wzh15, zh.luo, bswang}@nudt.edu.cn

Abstract

Large language models (LLMs) provide a powerful foundation for social media simulation, but their utility is fundamentally limited by the accuracy of user behavior modeling. Existing LLM agents exhibit behavioral homogenization and fail to capture users' inherent identities and evolving preferences. To address this, we propose CLASP, a closed-loop behavior alignment framework for social media agents. CLASP leverages explicit, dynamically refined personas to bridge users' historical behaviors and subsequent agent actions, ensuring behavioral consistency and temporal adaptability. Our framework adopts a decoupled dual-model design to separate persona refinement from action generation. This design enables continual behavior alignment driven by action prediction discrepancies. To facilitate training and evaluation, we further construct *BlueTrack*, a large-scale Bluesky dataset that covers long-horizon and heterogeneous social behaviors with complete action elements. Extensive experiments demonstrate that CLASP outperforms baselines in behavioral alignment with real users. It improves the overall alignment quality by up to 31.8% and reduces its variance by 32.1% during long-sequence simulation. This work provides a practical reference for further exploration in realistic social media simulation¹.

1 Introduction

Social media simulation creates an economical, customizable experimental platform for studying misinformation propagation, online polarization, and herd behavior (Yang et al., 2024; Gausen et al., 2022; Sohn, 2022). Advances in Large Language Model (LLM) offer a promising direction to build more realistic social agents (Huang et al., 2026; Lu et al., 2025). Nevertheless, extensive studies

*Corresponding authors.

¹The code and *BlueTrack* dataset are available at <https://github.com/xiaosongdian/CLASP>

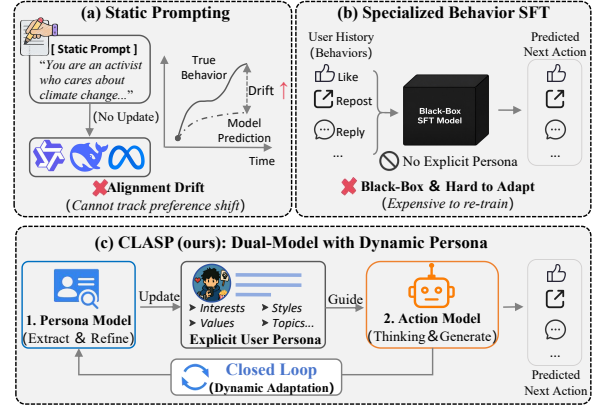


Figure 1: Comparison of user behavior alignment paradigms. (a) Static prompting. (b) Specialized behavior finetune. (c) Our CLASP framework.

confirm that LLM-based agents suffer from behavioral homogenization in simulation and fail to capture the diversity of individual opinions (Chen et al., 2026). This problem becomes especially pronounced in long-horizon simulations, where initial behavioral discrepancies accumulate and lead to severe alignment drift between agent and real user patterns (Wu et al., 2026).

Specifically, pre-trained LLMs often fail to track human preferences and their temporal dynamics, leading to convergent behaviors (He et al., 2024b; Chakraborty et al., 2024). As shown in Figure 1, existing approaches to aligning LLM agents with user behaviors can be broadly divided into two paradigms: (1) **Static prompting** relies on static personas, which ensures short-term behavioral consistency but suffers from alignment drift over time (Shao et al., 2023; Ge et al., 2025); (2) **Specialized behavior fine-tuning** can achieve reasonable behavioral alignment via historical data, but its prohibitive computational cost makes it impractical for large-scale deployments (Chopra et al., 2024; Rafailov et al., 2023; Deng et al., 2023). Furthermore, user behaviors on social media naturally

fall into two categories: **content generation behaviors** (e.g., *post*, *reply*), open-ended text generation tasks expressing users’ linguistic styles and opinions; **interaction decision behaviors** (e.g., *like*, *repost*), discrete choices indicating users’ interests and attitudes toward others’ content. Existing simulation systems fail to faithfully reproduce the above behaviors, limiting the reliability of results (La Gatta et al., 2026; Taillandier et al., 2026).

To improve behavioral alignment performance, LLM-based dynamic persona modeling has become a promising research direction (Purificato et al., 2024; Zhang, 2024). Nevertheless, current research still faces three critical issues: **(1) Single-behavior modeling limitation:** The lack of long-sequence heterogeneous behavior datasets with complete action elements has constrained research, limiting existing methods to single behavior modalities (Kong et al., 2025; Munker et al., 2026; Schwager et al., 2026). **(2) Long-sequence performance degradation:** Due to key information forgetting in long-sequence processing, full historical behavior inputs introduce outdated noise and dilute current preference signals. (Wang et al., 2025; Qi et al., 2026). **(3) Disconnected optimization objectives:** Existing methods lack a closed feedback path from behavioral alignment to persona refinement. Persona updates do not leverage action prediction discrepancies as signals, so they cannot reliably improve behavioral alignment with real users during long-horizon simulation.

Accordingly, we propose CLASP (Figure 1): a closed-loop behavioral alignment framework for social media agent simulation. In summary, our contributions are as follows:

- **New Dataset.** We introduce *BlueTrack*, a large-scale dataset consisting of 20,412,364 behavior records from 8,294 users, addressing the lack of existing datasets with long-sequence heterogeneous behavior coverage.
- **Novel Framework.** We build a decoupled dual-model structure, which balances stable user identity and adaptive preference changes.
- **Behavior Insight.** This work offers practical modeling paradigms and empirical findings, providing references for dynamic heterogeneous behavior simulation on social media platforms.

2 Behavior Alignment Modeling

Inspired by recommendation systems research (Zhou et al., 2024; Kang et al., 2023), we first for-

malize the long-sequence heterogeneous behavior alignment for social media agents as follows:

2.1 Task Formulation

To capture the temporal dynamics of users’ behavioral preferences, we partition their online activities into consecutive windows $W = \{W_t\}_{t=0}^T$, where each window W_t contains N sequential actions.

Based on these time windows, we model the social media dataset as a tuple $(\mathcal{V}, \{\mathcal{P}_u\}, \{\mathcal{A}_u\}, \mathcal{E})$, where $\mathcal{V}$ denotes the set of nodes each representing a distinct user, and $\mathcal{E}$ represents the set of directed edges such that $(u, v) \in \mathcal{E}$ indicates user u follows user v . Each user $u \in \mathcal{V}$ is associated with a dynamic persona sequence $\mathcal{P}_u = \{P_1, P_2, \dots, P_T\}$, where P_t denotes the persona at window t , represented as natural language descriptions. Meanwhile, each user u possesses an action set $\mathcal{A}_u = \{a_1, a_2, \dots, a_n\}$, where each action $a_k \in \mathcal{A}_u$ is annotated with a behavior type label $y_k \in \{G, I\}$ (with G denoting content generation behavior and I denoting interaction decision behavior). Formally, each action a contains three elements and is defined as a tuple: $a = (t, o, c)$, where t denotes the type of the action (*post*, *repost*, *like*, *reply*), o denotes the object of the action, and c denotes the content of the action (Li et al., 2024a).

2.2 Task Evaluation

To quantitatively assess persona quality, we define the core concept for future behavioral alignment.

Definition 1 (Alignment Quality). Given a user u and the corresponding persona P_t , the alignment quality $Q_{t+1}(P_t)$ measures the expected performance of P in predicting future actions across the heterogeneous behavior types. Formally,

$$Q_{t+1}(P_t) = \frac{1}{N} \sum_{i=1}^N \mathcal{S}(\pi_\theta(P_t, C_i), a_i) \quad (1)$$

$A_{t+1} = \{a_{t+1}^i\}_i^N$ represents user actual behaviors in W_{t+1} , while $\hat{A}_{t+1|P_t} = \{\pi_\theta(P_t, C_i)\}_i^N$ denotes predictions with persona P_t and scenario context C_i by action model π_θ . And $\mathcal{S}(\cdot, \cdot)$ is a unified alignment scoring function.

For each action in W_{t+1} , we map its type to a corresponding question type via type-aware question generation (details in Appendix B). The unified

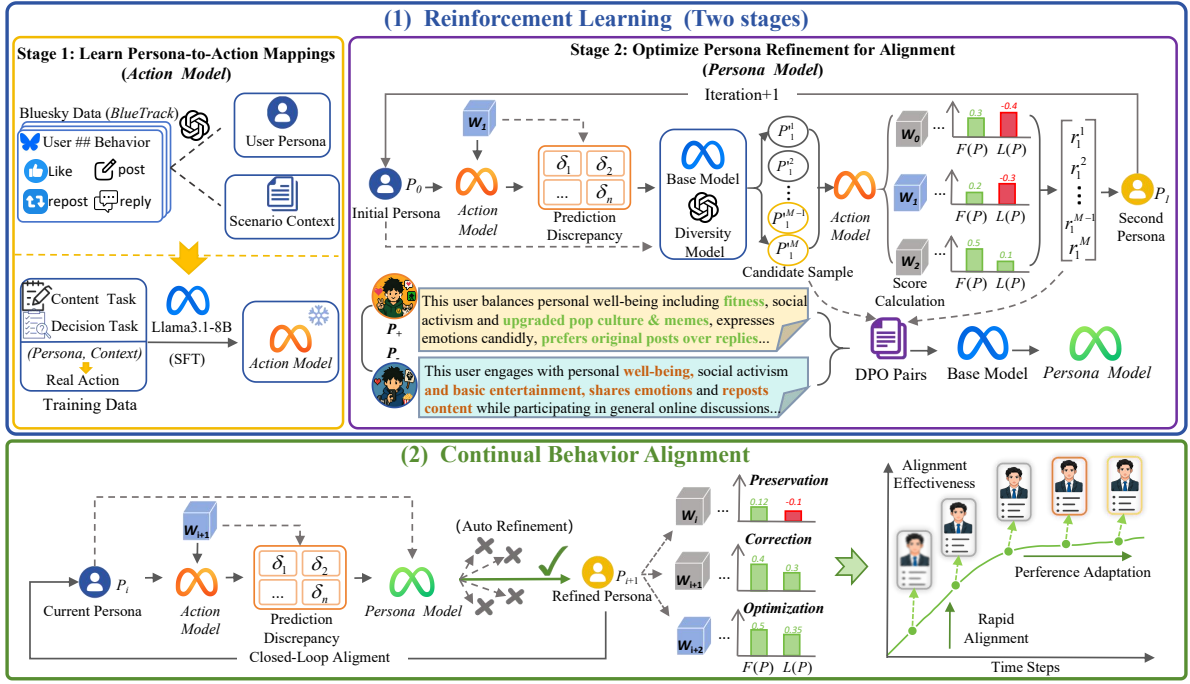


Figure 2: **Framework of CLASP.** (1) Dual-model Decoupled Reinforcement Learning: It first trains an action model to learn persona-to-action mappings from real-world social media data, then constructs direct preference optimization (DPO) preference pairs via jointly assessing semantic and decision consistency, and trains the persona model to adjust the user persona based on prediction discrepancies. (2) Online Behavior Alignment: Deployed trained dual models for multi-round updates, which enable continual alignment through closed-loop feedback.

alignment function $\mathcal{S}$ is defined as follows:

$$\mathcal{S}(\hat{a}, a) = \begin{cases} \alpha \cdot \mathcal{F}(\hat{a}_I, a_I), & \text{if } a \in I, \\ (1 - \alpha) \cdot \mathcal{L}(\hat{a}_G, a_G), & \text{if } a \in G \end{cases} \quad (2)$$

where $\mathcal{F}$ and $\mathcal{L}$ represent the alignment scores of interaction and generation tasks, respectively. And α is a global weighting coefficient balancing the contributions of them. We concretize the two components as follows:

Interactive Decision Alignment $\mathcal{F}$. For discrete interactive decision tasks, we propose a weighted F_1 to measure the action type prediction ability. To balance information disparities across actions, we use differential weighting by inherent action value. Formally, it is defined as:

$$\mathcal{F} = w_k \cdot F_1^{(k)}, \quad (3)$$

$F_1^{(k)}$ denotes the standard F_1 score evaluated on action type k , and w_k are type weights (details in Appendix D.2). This design avoids biased personas' tendency toward a single action type, and better reflects users' diverse preference distinctions over various contents.

Content Generation Alignment $\mathcal{L}$. For content generation behavior, we measure semantic

consistency between predicted content $\hat{a}_G$ and ground-truth content a_G using cosine similarity of Sentence-BERT (all-mpnet-base-v2²) embeddings. Formally:

$$\mathcal{L} = \cos(\phi(\hat{a}_G), \phi(a_G)) \quad (4)$$

3 The design of CLASP

Static user personas cannot adapt to gradual behavioral shifts, leading to severe performance degradation in long-term interaction scenarios. We thus propose CLASP, a dual-model decoupled dynamic persona framework for long-sequence behavior alignment on social media. As shown in Figure 2, the overall framework of CLASP is divided into two phases: (1) Training of the action model and persona model via Reinforcement Learning (RL). (2) Closed-loop persona refinement for continual behavior alignment.

3.1 Dual-Model Decoupled Reinforcement Learning

In this phase, we train two complementary models on real-world social media data: an action

²Sentence-BERT model: <https://huggingface.co/sentence-transformers/all-mpnet-base-v2>

model and a persona model. Motivated by the Machine Theory of Mind (ToMNet) framework (Rabinowitz et al., 2018), we first learn a ground-truth mapping from user personas to real behaviors via the action model, and then use this mapping as an objective evaluator to optimize the persona model. This decoupling stems from the hierarchical nature of social media behaviors: personas capture stable user traits, while individual actions are context-dependent behavioral responses. Separating them avoids objective interference and supports discrepancy-driven persona refinement.

3.1.1 Action Model Construction

The goal of this stage is to train an action model that predicts realistic user behaviors given a specific persona and scenario context. To this end, we constructed *BlueTrack*, the large-scale Bluesky³ dataset that comprehensively covers heterogeneous user behaviors with complete action elements. Details of *BlueTrack* are provided in Section 4.1.

Specifically, we feed the complete set of user actions into GPT-4o-mini⁴, which generates unstructured natural language personas (Wang et al., 2025). We then construct supervised training samples mapping these personas P and corresponding scenario contexts C to real user actions, covering two task categories:

- **Content generation tasks:** Predict action content c produced by users corresponding to post and reply behaviors.
- **Interactive decision tasks:** Predict action type $t \in \{post, repost, like, reply\}$ given action object o or content c .

We initialize the action model with Llama-3.1-8B-Instruct⁵ and perform standard SFT. The SFT dataset is constructed by selecting a subset of users $\mathcal{U}_{\text{SFT}} \subseteq \mathcal{U}$ from the *BlueTrack* dataset. For each $u \in \mathcal{U}_{\text{SFT}}$, the prompt is (P_u, C_u) . The reference response is the real user action sequence $A_u = \{a_1, a_2, \dots, a_n\}$. The objective of SFT is to narrow the gap between the actions generated by the action model π_θ and real reference behaviors. The model is trained by minimizing the following negative log-likelihood loss:

$$\mathcal{L}_{\text{SFT}} = -\frac{1}{|\mathcal{U}_{\text{SFT}}|} \sum_{u \in \mathcal{U}_{\text{SFT}}} \log \pi_\theta(A_u | P_u, C_u) \quad (5)$$

³Bluesky url: <https://bsky.app/>

⁴GPT-4o-mini: <https://openai.com/index/gpt-4o-mini-advancing-cost-efficient-intelligence/>

⁵Llama-3.1-8B-Instruct: <https://huggingface.co/meta-llama/Meta-Llama-3.1-8B>

The performance of the action model is shown in Appendix D.4.

3.1.2 Persona Model Construction

To enhance the *Alignment Quality* for diverse users over long sequences, we freeze the SFT-trained action model π_θ as a fixed behavioral simulator in subsequent training. This ensures all personas are evaluated under an identical action generation standard. We further train the persona model π_γ via Direct Preference Optimization (DPO) to learn optimal refinement directions for persona updating.

Candidate Personas Generation. We design a multi-step pipeline to construct preference pairs. At the initial iteration $t = 1$, we adopt W_0 as the original behavior window. The base model generates the initial persona P_0 based on the user’s behavioral data in W_0 .

For each iteration $t \geq 1$, we input the previous persona P_{t-1} into π_θ to predict the user’s actions $\hat{A}_t$ within the current window W_t . We then construct the prediction discrepancy set $\{\delta_i\}$ between the predicted actions $\hat{A}_t$ and the real actions A_t . The details of discrepancy construction are provided in Appendix C.

Subsequently, we feed P_{t-1} and $\{\delta_i\}$ into the persona model ensemble, which generates a set of candidate persona updates. Formally, the generation process is defined as:

$$(P_{t-1}, \{\delta_i\}) \rightarrow \{P'_t{}^k\}_{k=1}^M \quad (6)$$

$\{P'_t{}^k\}_{k=1}^M$ is the set of M candidate personas in W_t . A diversity model (GPT-4o-mini) is integrated into the ensemble to expand the refinement search space and broaden candidate distribution.

Reward Function Design. We formulate three reinforcing objectives to guide the persona model’s directed search for persona updating directions (Chen et al., 2025):

- **Trait Preservation:** Maintain stable core persona traits from historical windows to ensure behavioral consistency across iterations.
- **Discrepancy Correction:** Adjust the persona to resolve the prediction discrepancy observed in the current window, reflecting the user’s latest behavioral shifts.
- **Predictive Optimization:** Optimize the persona to minimize future prediction errors.

Accordingly, we formalize three objectives of refinement as improvement in prediction quality

Dataset	Platform	Users	Avg. Actions	Timeline	Action Coverage	Completeness	Anonymization
PersonalDialog (Zheng et al., 2019)	Weibo	8.47M	6.64	×	×	×	×
Twibot-20 (Feng et al., 2021)	Twitter	5.32k	145.86	×	×	×	✓
Twibot-22 (Feng et al., 2022)	Twitter	860k	86.76	×	✓	×	✓
FineRob (Li et al., 2024a)	Twitter/Reddit/Zhihu	1.46k	44.49	✓	✓	✓	×
BluePrint (Buck-Kaeffer et al., 2025)	BlueSky	236k	24.57	✓	✓	×	✓
SYNTHIA (Rahimzadeh et al., 2026)	BlueSky	650k	234.5	✓	×	×	✓
BlueTrack (Ours)	BlueSky	8.29k	2,461.10	✓	✓	✓	✓

Table 1: Comparison of mainstream social media user behavior datasets. *BlueTrack* uniquely satisfies all evaluation criteria, including long temporal action sequences for complete action elements, diverse heterogeneous behavior categories, and rigorous privacy anonymization. “Avg. Actions” equals total actions per unique user. “Completeness” denotes whether each action record contains object and content elements.

across past, current, and future windows, represented by rewards r_t^{pres} , r_t^{corr} , r_t^{opti} :

$$\begin{aligned}
r_t^{pres} &= Q_{t-1}(P'_t) - Q_{t-1}(P_{t-1}) \\
r_t^{corr} &= Q_t(P'_t) - Q_t(P_{t-1}) \\
r_t^{opti} &= Q_{t+1}(P'_t) - Q_{t+1}(P_{t-1})
\end{aligned} \quad (7)$$

The total reward for a candidate persona at step t is the sum of three components, corresponding to the *Alignment Quality* gains across three windows:

$$r_t = r_t^{pres} + r_t^{corr} + r_t^{opti} \quad (8)$$

DPO Pairs Construction and Training. For each input $(P_{t-1}, \{\delta_i\})$, the persona model ensemble generates M candidate personas $\{P'_t\}_{k=1}^M$, each of which is assigned a reward computed as defined in Equation 8. We construct DPO preference pairs both by partitioning candidates into a positive set $\{P'_t\}^+$ (with $r_t \geq \tau_+$) and a negative set $\{P'_t\}^-$ (with $r_t < \tau_-$) via predefined thresholds, and by retaining additional pairs with a reward margin $r_t^+ - r_t^- \geq \Delta$ and negative candidate reward $r_t^- \leq 0$. A case of this pair construction process is provided in Appendix E.1. Subsequently, we select the candidate with the highest reward as the optimal refined persona P_t , which is used as the initial persona for the next iteration. The above process is executed iteratively with $t = t + 1$.

The base persona model π_γ is trained via DPO with these preference pairs. Following (Gui et al., 2024), we integrate SFT loss into the DPO objective to maintain its fundamental persona generation capabilities:

$$\mathcal{L}(\pi_\gamma; \pi_{ref}) = \mathcal{L}_{DPO}(\pi_\gamma; \pi_{ref}) + \beta \cdot \mathcal{L}_{SFT}(\pi_\gamma) \quad (9)$$

3.2 Continual Behavior Alignment

Following RL training, we deploy the dual-model framework for continual behavior alignment. At each step i , the current persona P_i is fed into the

fixed action model π_θ to predict user behaviors. The discrepancy set $\{\delta_i\}$ between these predictions and the related real behaviors drives the persona model π_γ to generate a refined persona P_{i+1} . This iterative refinement process achieves three objectives: preserving stable traits, correcting behavior discrepancies, and improving future prediction accuracy, enabling the system to rapidly adapt to user preferences while maintaining behavioral stability.

4 Experiments

We empirically evaluate CLASP on the task of long-term social agent behavior alignment using real-world social media datasets, to answer three research questions:

- **RQ1:** How does CLASP outperform baselines in cross-community behavioral alignment across both generation G and interaction I tasks?
- **RQ2:** How does CLASP’s dynamic persona update achieve balanced alignment quality across different windows?
- **RQ3:** What components drive the effectiveness of CLASP?

4.1 Dataset

To address the lack of large-scale benchmarks for heterogeneous user behavior alignment on social media, we construct *BlueTrack*, the first large-scale dataset collected from publicly available Bluesky social media data (Failla and Rossetti, 2024). It comprehensively covers both G and I behavior types, and uniquely provides user action sequences with complete action elements. After thorough cleaning, it contains 8,294 valid users and 20,412,364 behavioral records. As shown in Table 1, *BlueTrack* outperforms existing benchmarks and is the only one satisfying all practical research demands. To support systematic training and evaluation, we build user social graphs based on follow relationships and adopt the Louvain al-

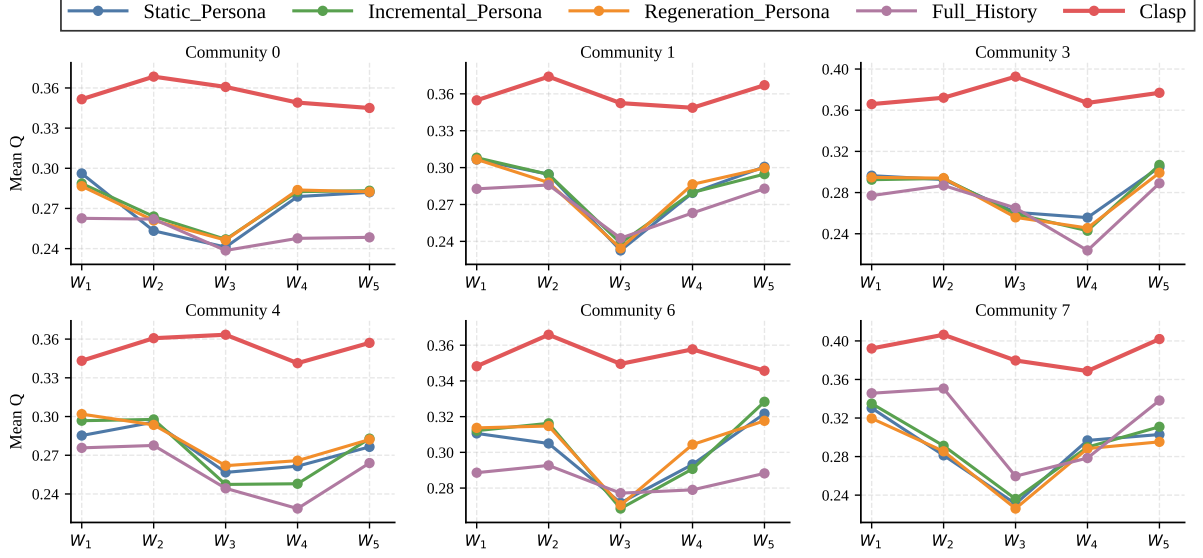


Figure 3: Behavior alignment performance of different methods across 5 rounds in 6 communities. *Alignment quality* Q is defined in Equation 1. Communities 6 and 7 are unseen during training.

gorithm (Blondel et al., 2008) to divide the graph into 10 densely connected communities. Detailed data construction and preprocessing procedures are provided in the Appendix A.

4.2 Models and Training

CLASP is built upon Llama-3.1-8B-Instruct. We adopt LlamaFactory (Zheng et al., 2024) for SFT and DPO fine-tuning. For SFT, we utilize behavioral data from four communities (2, 5, 8, and 9), comprising 2,296 users in total, to train the action model π_θ . For preference optimization, we sample 70% of users from each training community (0, 1, 3, 4), yielding 3,095 training users. We perform two iterations, sampling $M = 10$ candidate personas per iteration, 4 of which are generated by GPT-4o-mini. With thresholds set to $\tau_+ = 0.06$, $\tau_- = -0.02$ and $\Delta = 0.1$, we obtain 8,702 preference pairs in total. The weighting coefficient α is set to 0.6, while the SFT loss coefficient β is set to 0.1. Additional details regarding model architectures and training hyperparameters are provided in the Appendix C.

4.3 Baselines

We adopt four representative baselines with distinct strategies for utilizing user historical behavior. **Static_Persona** retains the initial persona derived from window W_0 and incorporates short-term actions in W_t . **Incremental_Persona** updates the user persona incrementally based on the previous persona S_{t-1} and recent actions in W_t (Yuan et al.,

2024). **Regeneration_Persona** reconstructs user personas from scratch using all historical behaviors until W_t (Zhou et al., 2024). **Full_History** directly uses all historical actions before window W_{t+1} without an explicit persona. All baseline methods are implemented using the identical base model Llama-3.1-8B-Instruct to ensure fair comparison.

4.4 Evaluation Metrics

We evaluate the effectiveness of these methods with *Alignment Quality* across different windows, quantified by the mean $Q_{t+1}(P_t)$ over updating rounds.

5 Evaluation

5.1 Behavior Alignment Performance

CLASP achieves consistent alignment performance. As shown in Figure 3, it attains mean alignment quality Q of 0.369, outperforming all baselines: 31.8% over the weakest baseline *Full_History* ($Q = 0.275$) and 29.0% over the second-best baseline *Regeneration_Persona* ($Q = 0.281$). Notably, on the two unseen communities (6 and 7), it maintains a mean Q of 0.382, a 26.1% improvement over the baseline average ($Q = 0.303$). Meanwhile, CLASP also exhibits the lowest variance across all baselines (Figure 4), with standard deviations of 0.0126 (seen) and 0.0214 (unseen), reducing variance by 39.9% and 24.3% relative to baselines. Table 2 shows detailed statistics.

Interactive decision alignment is easier to optimize than content generation. At the compo-

Method	Community 0			Community 1			Community 3			Community 4			Community 6 [†]			Community 7 [†]			Average		
	F	L	Q	F	L	Q	F	L	Q	F	L	Q	F	L	Q	F	L	Q	F	L	Q
Static_Persona	0.331	0.179	0.270	0.337	0.198	0.281	0.334	0.205	0.282	0.339	0.180	0.275	0.378	0.185	0.300	0.313	0.251	0.288	0.339	0.200	0.280
Regeneration_Persona	0.335	0.178	0.272	0.339	0.199	0.283	0.329	0.199	0.278	0.351	0.177	0.281	0.383	<u>0.187</u>	0.304	0.306	0.249	0.283	0.340	0.198	0.281
Incremental_Persona	0.330	0.187	0.273	0.336	0.204	0.283	0.326	0.209	0.279	0.336	0.183	0.275	0.377	0.192	0.303	0.315	0.259	0.293	0.337	0.206	0.278
Full_History	0.337	0.125	0.252	0.340	0.168	0.271	0.339	0.161	0.268	0.342	0.132	0.258	0.380	0.143	0.285	0.340	0.277	0.315	0.346	0.168	0.275
w/o Action Model	0.345	0.181	0.279	0.361	0.199	0.297	0.351	0.192	0.287	0.379	0.169	0.295	0.375	0.174	0.294	0.342	0.266	0.312	0.359	0.197	0.294
w/o Diversity Model	0.457	<u>0.184</u>	0.348	<u>0.453</u>	0.213	<u>0.357</u>	0.474	0.207	0.367	0.444	0.184	0.339	0.454	0.179	<u>0.344</u>	0.434	<u>0.332</u>	0.393	0.453	<u>0.216</u>	0.358
w/o Persona Model	0.460	0.179	0.348	0.451	<u>0.215</u>	0.356	<u>0.479</u>	<u>0.210</u>	<u>0.372</u>	0.439	0.178	0.335	0.448	0.172	0.338	<u>0.457</u>	<u>0.332</u>	0.407	<u>0.456</u>	0.214	<u>0.359</u>
CLASP (Ours)	<u>0.459</u>	0.179	<u>0.347</u>	0.463	0.220	0.366	0.487	0.225	0.382	0.465	0.191	0.350	0.467	0.174	0.350	0.462	0.342	0.414	0.467	0.222	0.369

Table 2: Average Performance of Baseline Methods and Ablation Experiments on 6 Communities (F/L/Q). Best results are in **bold**, second-best results are underlined.[†] denotes communities unseen during training.

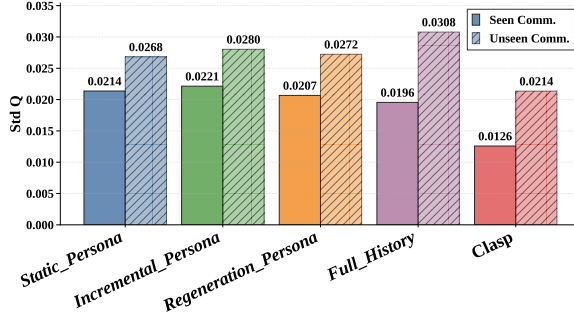


Figure 4: Standard deviation of *Alignment Quality* (Std Q) across seen and unseen communities.

ment level, CLASP achieves an average interactive decision alignment score $\mathcal{F}$ of 0.467 and an average content generation alignment score $\mathcal{L}$ of 0.222, corresponding to 37.2% and 15.0% relative improvements over the baseline. This gap can be attributed to the diversity of user-generated content, which exhibits significantly lower temporal and semantic consistency than user interaction decisions. Admittedly, semantic consistency $\mathcal{L}$ alone cannot comprehensively evaluate content alignment. LLM evaluation is vulnerable to prompt-induced bias and unstable results (Wang et al., 2024b), and human evaluation is costly and prone to subjective discrepancies. We therefore select $\mathcal{L}$ for standardized quantitative assessment.

5.2 Multi-Objective Temporal Balancing

To examine how the persona model within CLASP balances the reinforcement objectives, we further analyze the alignment shifts across windows after persona updating. All ΔQ values presented below refer to cumulative discrepancy.

In Figure 5, the persona refinement mechanism of CLASP achieves a balanced improvement across W_0 , W_1 , and W_2 . *Regeneration_Persona* and *Incremental_Persona* strategies suffer from severe degradation in the early window W_0 (e.g.,

$\Delta Q = -0.786$ and -0.478 , respectively), while the base model (with discrepancy) shows only weak gains. In contrast, CLASP alleviates performance decline in W_0 ($\Delta Q = -0.025$), corrects misalignment in current window W_1 ($\Delta Q = +0.241$), and significantly boosts alignment quality in future window W_2 ($\Delta Q = +0.410$). These results confirm that CLASP successfully balances **trait preservation**, **discrepancy correction**, and **predictive optimization**, while improving alignment stability over time.

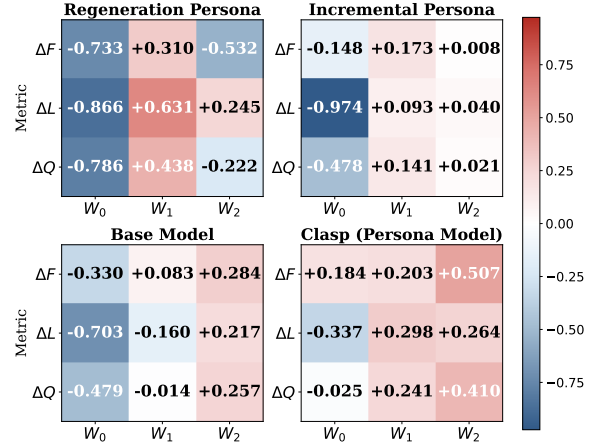


Figure 5: Persona Update Performance Across Different Methods. The heatmap presents cumulative alignment discrepancies from P_0 to P_1 for metrics (F , L , Q) across windows (W_0 , W_1 , W_2).

5.3 Persona Analysis

We further analyze the personas generated by CLASP to explore their characteristics in social media simulation.

We first examine the persona refinement process across time windows. Table 3 shows that CLASP effectively controls the growth of persona token length while preserving essential user features. Figure 6 illustrates the dynamic update pro-

Method	P_0	P_1	P_2	P_3	P_4	P_5
CLASP	564.9	635.2	645.0	648.2	653.3	667.7
Incremental_Persona	564.9	641.0	659.5	678.5	670.0	691.1
Regeneration_Persona	564.9	590.8	596.8	581.6	610.8	619.8

Table 3: Average token consumption of persona across rounds.

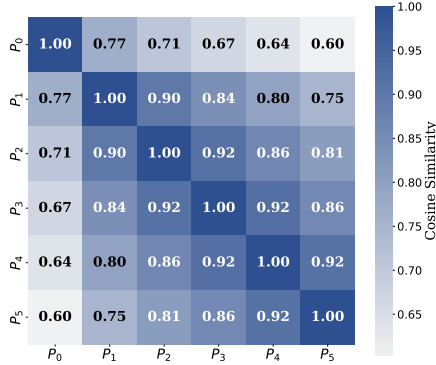


Figure 6: Persona Cosine Similarity across rounds.

cess: rapid initial alignment ($P_0 \rightarrow P_1$) followed by stable preference adjustments ($P_1 \rightarrow P_5$), indicating that our framework enables continual alignment of generated personas to real user behavioral patterns.

5.4 Ablation Study

We construct three ablation variants by disabling the corresponding training process in CLASP: (1) w/o Action Model: removes the SFT training process of the action model; (2) w/o Diversity Model: removes the GPT-4o-mini aided persona generation within the persona model training pipeline; (3) w/o Persona Model: removes the DPO training process of the persona model. All variants degenerate to the base model (Llama-3.1-8B) backbone and preserve the identical functional pipeline.

As shown in Table 2, removing the Action Model training process causes catastrophic performance degradation: the average F drops by 23.1% from 0.467 to 0.359, and average L drops by 11.3% from 0.222 to 0.197. Ablating the Persona Model and Diversity Model yields moderate F performance drops of 2.4% and 3.0%. Although the Persona Model brings limited improvement, it stably sustains iterative persona refinement and reduces alignment variance. The Diversity Model mitigates persona homogeneity and enhances the robustness of behavioral alignment. Notably, all ablation variants still outperform all baseline methods. These results confirm that our dual-model decoupled design improves both alignment performance and temporal

stability via module collaboration.

6 Related work

LLM-based Agents in Social Simulation. Recently, LLM agents have opened up innovative pathways for social simulation, which are classified into three levels: individual, scenario, and society level (Mou et al., 2024). The individual-level focuses on modeling specific personas or demographic groups to analyze their behavioral patterns (Shao et al., 2023; Park et al., 2023; Wang et al., 2024a). Scenario-level simulations leverage structured multi-agent interactions to elicit collective intelligence, covering domain-specific tasks including medical diagnosis (Fan et al., 2024), software analysis (Hong et al., 2026), and judicial decision-making (He et al., 2024a). Society-level simulations concentrate on investigating macro social phenomena such as information diffusion (Zhang et al., 2025), macroeconomic principles (Li et al., 2024b), and opinion dynamics (Zhang et al., 2024).

User behavior Alignment. To enhance the realism and robustness of social simulation, LLM-based agents suffer from two critical issues: a narrower variance in behavioral distribution than that of real human populations, and instability in long-horizon simulations (Wu et al., 2026; Larooij and Törnberg, 2025). Prior studies focus on three core directions to tackle them: Reinforcement learning is employed to mimic human behavior and ensure real-world consistency (Li et al., 2024a; Kong et al., 2025); Personalization is achieved via prompt engineering (Park et al., 2022), personality alignment (Serapio-García et al., 2023), real-person persona modeling (Jung et al., 2025; Park et al., 2023), and population alignment (Li et al., 2025; Ge et al., 2025; Hu et al., 2025); Dynamic persona modeling adopts iterative updating pipelines to tackle long-term alignment drift (Yao et al., 2025; Luo and Laban, 2026). Unlike existing approaches, CLASP leverages explicit personas to bridge historical actions and subsequent agent behaviors. With a dual-model decoupled structure and discrepancy-driven refinement, it maintains stable user identity and accommodates temporal preference changes.

7 Conclusion

In this paper, we introduce CLASP, a novel closed-loop framework for long-term social media behavior alignment. Its dual-model decoupled RL archi-

texture iteratively refines personas via prediction discrepancies, delivering robust alignment across both content generation and interaction decision tasks. We also contribute *BlueTrack*, a large-scale Bluesky dataset designed for benchmarking social media agent behavior. Our results show that CLASP outperforms all baselines, enabling faithful agent behavior generation and more realistic social media platform simulation.

Limitations

There are several limitations in our work. First, we contribute *BlueTrack* to address the lack of real-world benchmarks for LLM-based social media agent simulation, but all social media data collection is constrained by platform policies and ethical guidelines. These constraints limit the scale of the dataset and prevent comprehensive validation of cross-platform generalization to major platforms such as X and Reddit.

Second, social-media user-generated content (UGC) is semantically divergent: individual posts are loosely-worded and thematically ambiguous. Spontaneous colloquial expressions mixing emotions, comments and irrelevant thoughts prevent models from acquiring full user context, introducing high generation uncertainty. For example, user posts are bound to real-time life experiences and cannot be reliably predicted from historical behaviors alone.

Third, Our evaluation scheme is constrained by the requirements of large-scale simulation scenarios. Given the demand for massive user samples, lightweight automatic metrics are used to measure internal in training and deployment, yet such metrics have limited dimensions and incomplete coverage. We avoid adopting the prevalent LLM-as-judge paradigm for computational cost and runtime efficiency reasons. Additionally, no unified consensus has been reached in the research community regarding metrics for quantifying consistency between real-user texts and agent-generated texts, representing a major future research direction.

Fourth, to reduce modeling complexity, we only consider four main action types (*post*, *reply*, *repost*, *like*) and model users independently without accounting for social influence effects from their social connections. Furthermore, we lack access to the real-time browsing context at the time of their actions, which introduces inherent prediction

bias. These limitations point to promising future directions for improving alignment between agents and real users.

Ethics Statement

We use the raw Bluesky dataset from (Failla and Rossetti, 2024) as our source, and perform task-specific processing to construct *BlueTrack*. All operations comply with platform policies, privacy regulations, and the original data usage terms, and are exclusively for non-commercial research purposes. We perform rigorous anonymization throughout the entire pipeline: all personally identifiable information (PII), including usernames, profile details, and contact information, is removed; original user identifiers are replaced with anonymous numerical indices; and no raw user profile or private information is preserved or disclosed.

Our work focuses solely on user behavioral modeling for research purposes of social media simulation. Despite these safeguards, LLM-based social media agents inherently carry potential risks of misuse, such as generating harmful content, disinformation, or impersonating real users. To mitigate such risks, we restrict all experiments to academic research and implement manual content review procedures to prevent harmful or deceptive outputs.

We acknowledge that CLASP may inherit social biases present in the training data, which could lead to unbalanced or stereotyped agent behaviors. As this work focuses on a foundational framework for closed-loop behavior alignment, we plan to integrate further ethical safeguards in future extensions, including content watermarking, real-time output filtering, and algorithmic auditing to enhance transparency and accountability. We emphasize that our dataset and model are intended exclusively for advancing research in social simulation and should not be deployed for manipulative or harmful purposes.

Acknowledgments

This work was supported by the Innovation Research Foundation of National University of Defense Technology (Grant Number ZK25-16).

References

Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Jour-*

- nal of statistical mechanics: theory and experiment*, 2008(10):P10008.
- Aurélien Buck-Kaeffer, Je Qin Chooi, Dan Zhao, Maximilian Puelma Touzel, Kellin Pelrine, Jean-François Godbout, Reihaneh Rabbany, and Zachary Yang. 2025. [BluePrint: A social media user dataset for llm persona evaluation and training](#). *Preprint*, arXiv:2510.02343.
- Souradip Chakraborty, Jiahao Qiu, Hui Yuan, Alec Koppel, Furong Huang, Dinesh Manocha, Amrit Singh Bedi, and Mengdi Wang. 2024. Maxmin-rlhf: Alignment with diverse human preferences. *arXiv preprint arXiv:2402.08925*.
- Aili Chen, Chengyu Du, Jiangjie Chen, Jinghan Xu, Yikai Zhang, Siyu Yuan, Zulong Chen, Liangyue Li, and Yanghua Xiao. 2025. [DEEPER insight into your user: Directed persona refinement for dynamic persona modeling](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 24157–24180, Vienna, Austria. Association for Computational Linguistics.
- Jiawei Chen, Ruoxi Xu, Boxi Cao, Ruotong Pan, Yunfei Zhang, Yifei Hu, Yong Du, Tingting Gao, Yaojie Lu, Yingfei Sun, Xianpei Han, Le Sun, Xiangyu Wu, and Hongyu Lin. 2026. [Towards real-world human behavior simulation: Benchmarking large language models on long-horizon, cross-scenario, heterogeneous behavior traces](#). *Preprint*, arXiv:2604.08362.
- Ayush Chopra, Shashank Kumar, Nurullah Giray-Kuru, Ramesh Raskar, and Arnau Quera-Bofarull. 2024. On the limits of agency in agent-based models. *arXiv preprint arXiv:2409.10568*.
- Yihe Deng, Weitong Zhang, Zixiang Chen, and Quanquan Gu. 2023. Rephrase and respond: Let large language models ask better questions for themselves. *arXiv preprint arXiv:2311.04205*.
- Andrea Failla and Giulio Rossetti. 2024. “i’m in the bluesky tonight”: insights from a year worth of social data. *PLoS one*, 19(11):e0310330.
- Zhihao Fan, Jialong Tang, Wei Chen, Siyuan Wang, Zhongyu Wei, Jun Xi, Fei Huang, and Jingren Zhou. 2024. Ai hospital: Interactive evaluation and collaboration of llms as intern doctors for clinical diagnosis. *arXiv preprint arXiv:2402.09742*, 2.
- Shangbin Feng, Zhaoxuan Tan, Herun Wan, Ningnan Wang, Zilong Chen, Binchi Zhang, Qinghua Zheng, Wenqian Zhang, Zhenyu Lei, Shujie Yang, Xinshun Feng, Qingyue Zhang, Hongrui Wang, Yuhao Liu, Yuyang Bai, Heng Wang, Zijian Cai, Yanbo Wang, Lijing Zheng, and 3 others. 2022. [Twibot-22: Towards graph-based twitter bot detection](#). In *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*.
- Shangbin Feng, Herun Wan, Ningnan Wang, Jundong Li, and Minnan Luo. 2021. [Twibot-20: A comprehensive twitter bot detection benchmark](#). In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management, CIKM ’21*, page 4485–4494, New York, NY, USA. Association for Computing Machinery.
- Anna Gausen, Wayne Luk, and Ce Guo. 2022. Using agent-based modelling to evaluate the impact of algorithmic curation on social media. *ACM Journal of Data and Information Quality*, 15(1):1–24.
- Tao Ge, Xin Chan, Xiaoyang Wang, Dian Yu, Haitao Mi, and Dong Yu. 2025. [Scaling synthetic data creation with 1,000,000,000 personas](#). *Preprint*, arXiv:2406.20094.
- Lin Gui, Cristina Gârbacea, and Victor Veitch. 2024. [Bonbon alignment for large language models and the sweetness of best-of-n sampling](#). *Preprint*, arXiv:2406.00832.
- Zhitao He, Pengfei Cao, Chenhao Wang, Zhuoran Jin, Yubo Chen, Jiexin Xu, Huaijun Li, Kang Liu, and Jun Zhao. 2024a. Agentscourt: Building judicial decision-making agents with court debate simulation and legal knowledge augmentation. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 9399–9416.
- Zihao He, Minh Duc Chu, Rebecca Dorn, Siyi Guo, and Kristina Lerman. 2024b. Community-cross-instruct: Unsupervised instruction generation for aligning large language models to online communities. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 17001–17019.
- Chuan Hong, Zheng Zhang, Lei Zhou, Laisong Li, Chenyifan Liu, Ze Huang, Xu Zhou, and Peihong Lin. 2026. [Firmcure: towards autonomous and adaptive rehosting of linux-based firmware](#). *CoRR*, abs/2606.24549.
- Zhengyu Hu, Jianxun Lian, Zheyuan Xiao, Max Xiong, Yuxuan Lei, Tianfu Wang, Kaize Ding, Ziang Xiao, Nicholas Jing Yuan, and Xing Xie. 2025. Population-aligned persona generation for llm-based social simulation. *arXiv preprint arXiv:2509.10127*.
- Renhong Huang, Ning Tang, Jiarong Xu, Yuxuan Cao, Qingqian Tu, Sheng Guo, Bo Zheng, Huiyuan Liu, and Yang Yang. 2026. Polycsim: An llm-based agent social simulation sandbox for proactive policy optimization. In *Proceedings of the ACM Web Conference 2026*, pages 4781–4792.
- Soon-Gyo Jung, Joni Salminen, Kholoud Khalil Aldous, and Bernard J Jansen. 2025. Personacraft: Leveraging language models for data-driven persona development. *International Journal of Human-Computer Studies*, 197:103445.
- Wang-Cheng Kang, Jianmo Ni, Nikhil Mehta, Maheswaran Sathiamoorthy, Lichan Hong, Ed Chi, and

- Derek Zhiyuan Cheng. 2023. [Do llms understand user preferences? evaluating llms on user rating prediction](#). *Preprint*, arXiv:2305.06474.
- Fanqi Kong, Xiaoyuan Zhang, Xinyu Chen, Yaodong Yang, Song-Chun Zhu, and Xue Feng. 2025. Enhancing llm-based social bot via an adversarial learning framework. In *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 23246–23271.
- Valerio La Gatta, Gian Marco Orlando, Marco Perillo, Ferdinando Tammaro, and Vincenzo Moscato. 2026. From who they are to how they act: Behavioral traits in generative agent-based models of social media. *arXiv preprint arXiv:2601.15114*.
- Maik Larooij and Petter Törnberg. 2025. Validation is the central challenge for generative social simulation: a critical review of llms in agent-based modeling. *Artificial Intelligence Review*, 59(1):15.
- Jia-Nan Li, Jian Guan, Songhao Wu, Wei Wu, and Rui Yan. 2025. From 1,000,000 users to every user: Scaling up personalized preference for user-level alignment. *arXiv preprint arXiv:2503.15463*.
- Kun Li, Chenwei Dai, Wei Zhou, and Songlin Hu. 2024a. [Fine-grained behavior simulation with role-playing large language model on social media](#). *Preprint*, arXiv:2412.03148.
- Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024b. Econagent: large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.
- Zhixiang Lu, Xueyuan Deng, Yiran Liu, Yulong Li, Qiang Yan, Imran Razzak, and Jionglong Su. 2025. [Prism: A personality-driven multi-agent framework for social media simulation](#). *Preprint*, arXiv:2512.19933.
- Han Luo and Guy Laban. 2026. Spasm: Stable persona-driven agent simulation for multi-turn dialogue generation. *arXiv preprint arXiv:2604.09212*.
- Xinyi Mou, Xuanwen Ding, Qi He, Liang Wang, Jingcong Liang, Xinnong Zhang, Libo Sun, Jiayu Lin, Jie Zhou, Xuanjing Huang, and Zhongyu Wei. 2024. [From individual to society: A survey on social simulation driven by large language model-based agents](#). *Preprint*, arXiv:2412.03563.
- Simon Münker, Nils Schwager, Kai Kugler, Michael Heseltine, and Achim Rettinger. 2026. Next reply prediction x (nrp-x) dataset: Linguistic discrepancies in naively generated content. *Shaping Multilingual, Multimodal AI for the Social Sciences and Humanities (LLMs4SSH)@ LREC 2026*, page 45.
- Joon Sung Park, Joseph O’Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22.
- Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. In *Proceedings of the 35th annual ACM symposium on user interface software and technology*, pages 1–18.
- Erasmus Purificato, Ludovico Boratto, and Ernesto William De Luca. 2024. User modeling and user profiling: A comprehensive survey. *arXiv preprint arXiv:2402.09660*.
- Yirui Qi, Xiaoming Zhang, Ruilin Zeng, Mengyao Liu, Ziyi Zhou, Dezhuang Miao, Bingyu Yan, and Zhenyu Guan. 2026. [Beyond static persona consistency: Dynamic persona coherence in LLM role-playing](#). In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 28942–28956, San Diego, California, United States. Association for Computational Linguistics.
- Neil Rabinowitz, Frank Perbet, Francis Song, Chiyuan Zhang, SM Ali Eslami, and Matthew Botvinick. 2018. Machine theory of mind. In *International conference on machine learning*, pages 4218–4227. PMLR.
- Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. 2023. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741.
- Vahid Rahimzadeh, Erfan Moosavi Monazzah, Mohammad Taher Pilehvar, and Yadollah Yaghoobzadeh. 2026. Synthia: Scalable grounded persona generation from social media data. In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 45983–46006.
- Nils Schwager, Simon Münker, Alistair Plum, and Achim Rettinger. 2026. [Towards simulating social media users with llms: Evaluating the operational validity of conditioned comment prediction](#). In *The Proceedings for the 15th Workshop on Computational Approaches to Subjectivity, Sentiment Social Media Analysis (WASSA 2026)*, page 208–221. Association for Computational Linguistics.
- Greg Serapio-García, Mustafa Safdari, Clément Crepy, Luning Sun, Stephen Fitz, Peter Romero, Marwa Abdulhai, Aleksandra Faust, and Maja Matarić. 2023. Personality traits in large language models. *arXiv preprint arXiv:2307.00184*.
- Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. 2023. Character-llm: A trainable agent for role-playing. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13153–13187.

- Dongyoung Sohn. 2022. Spiral of silence in the social media era: A simulation approach to the interplay between social networks and mass media. *Communication Research*, 49(1):139–166.
- Patrick Taillandier, Jean Daniel Zucker, Arnaud Grignard, Benoit Gaudou, Nghi Quang Huynh, and Alexis Drogoul. 2026. [Integrating llm in agent-based social simulation: Opportunities and challenges](#). *Preprint*, arXiv:2507.19364.
- Lu Wang, Di Zhang, Fangkai Yang, Pu Zhao, Jianfeng Liu, Yuefeng Zhan, Hao Sun, Qingwei Lin, Weiwei Deng, Dongmei Zhang, Feng Sun, and Qi Zhang. 2025. [Lettingo: Explore user profile generation for recommendation system](#). In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2*, KDD '25, page 2985–2995, New York, NY, USA. Association for Computing Machinery.
- Noah Wang, Zhongyuan Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou, Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, Man Zhang, Zhaoxiang Zhang, Wanli Ouyang, Ke Xu, Wenhao Huang, Jie Fu, and Junran Peng. 2024a. [Rolellm: Benchmarking, eliciting, and enhancing role-playing abilities of large language models](#). In *Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024*, volume ACL 2024 of *Findings of ACL*, pages 14743–14777. Association for Computational Linguistics.
- Peiyi Wang, Lei Li, Liang Chen, Zefan Cai, Dawei Zhu, Binghuai Lin, Yunbo Cao, Lingpeng Kong, Qi Liu, Tianyu Liu, and Zhifang Sui. 2024b. [Large language models are not fair evaluators](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, ACL 2024, Bangkok, Thailand, August 11-16, 2024, pages 9440–9450. Association for Computational Linguistics.
- Zengqing Wu, Run Peng, Takayuki Ito, Makoto Onizuka, and Chuan Xiao. 2026. [Llm-based social simulations require a boundary](#). *Preprint*, arXiv:2506.19806.
- Ziyi Yang, Zaibin Zhang, Zirui Zheng, Yuxian Jiang, Ziyue Gan, Zhiyu Wang, Zijian Ling, Jinsong Chen, Martz Ma, Bowen Dong, Prateek Gupta, Shuyue Hu, Zhenfei Yin, Guohao Li, Xu Jia, Lijun Wang, Bernard Ghanem, Huchuan Lu, Chaochao Lu, and 4 others. 2024. [OASIS: open agent social interaction simulations with one million agents](#). *CoRR*, abs/2411.11581.
- Bingsheng Yao, Bo Sun, Yuanzhe Dong, Yuxuan Lu, and Dakuo Wang. 2025. Dprf: A generalizable dynamic persona refinement framework for optimizing behavior alignment between personalized llm role-playing agents and humans. *arXiv preprint arXiv:2510.14205*.
- Xinfeng Yuan, Siyu Yuan, Yuhan Cui, Tianhe Lin, Xintao Wang, Rui Xu, Jiangjie Chen, and Deqing Yang. 2024. Evaluating character understanding of large language models via character profiling from fictional works. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 8015–8036.
- Jiarui Zhang. 2024. Guided profile generation improves personalization with large language models. In *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 4005–4016.
- Lan Zhang, Yuxuan Hu, Weihua Li, Quan Bai, and Parma Nand. 2025. [Llm-aidsim: Llm-enhanced agent-based influence diffusion simulation in social networks](#). *Systems*, 13(1).
- Zeyu Zhang, Jianxun Lian, Chen Ma, Yaning Qu, Ye Luo, Lei Wang, Rui Li, Xu Chen, Yankai Lin, Le Wu, Xing Xie, and Ji-Rong Wen. 2024. [Trendsim: Simulating trending topics in social media under poisoning attacks with llm-based multi-agent system](#). *Preprint*, arXiv:2412.12196.
- Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan Ye, Zheyang Luo, Zhangchi Feng, and Yongqiang Ma. 2024. [Llamafactory: Unified efficient fine-tuning of 100+ language models](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, Bangkok, Thailand. Association for Computational Linguistics.
- Yinhe Zheng, Guanyi Chen, Minlie Huang, Song Liu, and Xuan Zhu. 2019. Personalized dialogue generation with diversified traits. *arXiv preprint arXiv:1901.09672*.
- Joyce Zhou, Yijia Dai, and Thorsten Joachims. 2024. [Language-based user profiles for recommendation](#). *Preprint*, arXiv:2402.15623.

A *BlueTrack*: Dataset Construction and Preprocessing

Our study utilizes the *BlueTrack* dataset, a long-term social behavior dataset from the Bluesky platform. It is specifically designed for evaluating long-term social agent behavior alignment, featuring long-horizon action sequences, heterogeneous behavioral types, and complete action elements.

Large-scale collection of social media data poses significant ethical and privacy risks. Therefore, we perform rigorous multi-stage cleaning on publicly available Bluesky social media data (Failla and Rossetti, 2024) and construct a user-centric, individual-level dataset for long-term behavior alignment research. The raw dataset contains complete post histories for over 4 million users. We use sequential numerical indices as user identifiers and filter out

Comm	Users	Edges	Likes	Reposts	Replies	Posts
0	1,704	44,676	185,386	1,427,577	1,312,187	821,063
1	584	7,240	49,581	615,756	536,527	316,700
2	1,688	39,503	260,225	1,651,189	1,707,428	1,396,515
3	879	17,448	114,162	541,658	499,437	302,298
4	1,300	25,043	141,023	1,185,481	886,792	515,330
5	185	3,574	16,959	117,869	167,771	158,849
6	860	15,529	78,448	812,707	512,416	362,304
7	670	16,932	56,998	1,350,346	781,229	324,760
8	100	2,488	7,952	105,205	82,198	41,536
9	324	5,532	27,075	495,620	268,391	177,416

Table 4: Community-level statistics of *BlueTrack* dataset, including the number of users, edges, likes, reposts, replies and posts for each community

users with fewer than 20 instances of each action type (*like*, *reply*, *repost*, *post*) to ensure diverse behavioral patterns per user. We further clean the dataset using social network information and action distribution: we exclude accounts with fewer than 3 mutual follow relationships (likely bots), more than 100 mutual follow relationships (likely official media accounts), and *reposts* account for more than 90% of total actions.

After filtering, the final *BlueTrack* dataset includes 8,294 users with 20,412,364 action records, covering complete action content and corresponding action objects. To mitigate privacy risks, all original user identifiers are replaced with anonymized numerical indices, and all user profile information is removed. We do not repurpose or distribute the raw dataset beyond research contexts, and the *BlueTrack* dataset created in this study is used solely for academic research in compliance with the original data access conditions.

We provide a detailed overview of *BlueTrack*, including user counts, action type distributions, and social network edge statistics for each community, as shown in Table 4. We further visualize the community structures in Figure 7.

To improve training efficiency, we first perform rigorous text content filtering on all actions: We remove actions containing only URLs, emojis, or hashtags, and retain only those with text content of at least 5 words. This ensures that all subsequent processing operates on high-quality, semantically meaningful data.

Then we apply a stratified temporal sampling strategy to the filtered action dataset. We select 6 months from September 2023 to February 2024, sampling at most 50 actions per user per month with a balanced quota: 10 posts, 15 replies, 10 reposts, and 15 likes. Deficits in replies and likes are filled by posts and reposts, respectively.

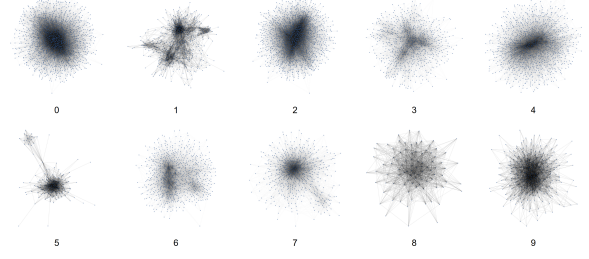


Figure 7: Visualization of user connectivity relationships in 10 communities.

B Action Model Construction Details

B.1 User Action Sequence Construction

We first construct standardized temporal action sequences for each user. Using the user ID as the unique primary key, we collect all behaviors associated with a given user and sort them in chronological order based on their timestamps. All actions are converted into a unified structured representation format, which preserves the temporal order, action type, action object, and action content. A sample user temporal action sequence is shown in Table 5.

```
# Single-User Temporal Action Sequence Example

[2024-02-01 14:00]
User performed a like action on content titled
"2024 NBA Playoffs Preview".

[2024-02-01 14:02]
User posted a reply to "2024 NBA Playoffs Preview":
"LeBron is in such stable form, looking forward to
the Finals!"

[2024-02-02 14:05]
User published a post: "2024 NBA Playoffs:
I think LeBron James will still be the MVP".

[2024-02-02 14:10]
User performed a repost action on content titled
"LeBron James Playoff Career Statistics Update".
...
```

Table 5: Single-User Temporal Action Sequence Example.

This unified representation enables consistent processing across all action types and captures complete action details, facilitating further processing in subsequent experiments.

B.2 Persona Initialization

In this study, we leverage the frozen LLM, GPT-4o-mini, to initialize open-format user personas P_0 based on their first 10 action sequence (W_0). This design deliberately avoids imposing predefined templates that would constrain the expressive ca-

capacity of LLM, thereby enabling comprehensive capture of users’ multi-dimensional dynamic preferences (Wang et al., 2025). The corresponding prompt is presented in Table 6.

TASK: You will serve as an assistant to help generate a user Persona based on the user's social media behavior history to better understand the user's interests and predict their future actions. USER BEHAVIOR HISTORY ({action_count} records): {temporal_action_sequence} Your Persona should: 1. Focus on patterns that can predict FUTURE behavior (not just describe past actions) 2. Include any dimensions you discover through analysis, don't limit yourself 3. Capture the user's unique behavioral signatures that distinguish them from others 4. Be actionable for predicting: what content they'll engage with, what actions they'll take, what triggers their participation 5. Scale the depth and detail of your analysis with the amount of input data - more data should yield richer, more nuanced insights User Persona: """
--

Table 6: Open-format Persona Initialization prompt template.

B.3 Dual-Task Joint SFT Fine-Tuning

On the *BlueTrack* dataset, we select initial community 9, along with its strongly correlated community 2 and weakly correlated communities 5 and 8, as our initial fine-tuning to ensure the broadest possible coverage of user demographics and behavioral patterns across the platform.

The objective of this phase is to fine-tune that can generate consistent actions conditioned on the context of a user’s persona and recent behavioral history, enabling the model to master fundamental behavioral regularities in social media scenarios. We construct training samples using a user-level temporal sliding window approach with the following sampling rules:

- **Sliding window configuration:** We uniformly set the window size to $T = 10$ (i.e., the user’s 10 recent historical behaviors) with a sliding step of 1, traversing the full-length behavioral sequence of each user.
- **Sample triggering rule:** The first T behaviors within the window are used as input, and samples for the corresponding task are automatically generated based on the type of the $(T + 1)$ -th ground-truth behavior.
- **Task balancing rule:** Through stratified random

sampling, we maintain a 4:6 ratio between the two target tasks: *Content Generation Task* and *Interaction Decision Task*.

To avoid overfitting to high-frequency actions (*like*, *repost*) in interaction tasks, we randomly convert 15% of eligible content generation samples to interaction samples, enriching diversity while preserving the task ratio. We design a unified JSON-format template that adapts to both tasks by adjusting only task-specific fields. The complete template is shown in Table 7.

Type-aware Question Generation: For the *Content Generation Task*, TASK_SPECIFIC_RULE constrains the model to generate only original user content without additional explanation. The scenario describes that the user will create either a comment toward the action object or a public post, and the optional_actions field is omitted. For the *Interaction Decision Task*, TASK_SPECIFIC_RULE restricts the output to a single candidate action from the predefined list, and the scenario queries the user’s next behavioral choice toward the action object or action content.

C Persona Model Construction details

C.1 Persona Refinement

<pre>{ "Instruction": " *Role*: Social media user behavior simulator *Requirement*: Strictly match target user persona and historical behaviors,output self-consistent, realistic results *Goal*: Replicate the user's authentic behavior in the given scenario *Output Rule*: {{TASK_SPECIFIC_RULE}} ", "input": { "user_persona": "{{User Persona}}", "history_sequence": "{{Top T user action sequence}}", "scenario": "{{TASK_SPECIFIC_SCENARIO}}" }, "optional_actions": "{{ONLY_FOR_INTERACTION_TASK: [like, repost, reply, post]}}", "output": "{{Ground-truth T+1 behavior}}" }</pre>

Table 7: Unified Dual-Task Json Template.

Given the discrepancy between the behaviors predicted by the action model and the ground-truth actions, we introduce an iterative refinement process. Specifically, we first generate action predictions using the action model prompt, which adopts the unified input structure in Table 7. The prompt

template for persona refinement is provided in Table 8. The behavior discrepancy signals are structured in the format shown in Table 9.

```

### Task: Refine User Persona Based on Discrepancies
Between Predicted and Ground-Truth Behaviors
### Instructions:
1. Retain the accurate and stable user
characteristics in the original persona.
2. Carefully analyze the discrepancies between
predicted behaviors and ground-truth behaviors, and
correct the errors or missing parts in the original
persona.
3. Supplement new user characteristics, preferences,
or behavioral patterns identified from the
discrepancies.
4. DO NOT mention any specific behavioral content,
timestamps, usernames, or platform names.

### Original User Persona: {old_persona}
### Comparison of Discrepancies Between Predicted and
Ground-Truth Behaviors:
{behavior_discrepancies}

### Please generate the refined user persona based on
the above information:

```

Table 8: Prompt Template for Persona Refinement.

```

### Behavior Discrepancy Signals
- Scenario 1: {scenario_context}
Predicted Action: {predicted_action}
Ground-Truth Action: {actual_action}

- Scenario 2: {scenario_context}
Predicted Action: {predicted_action}
Ground-Truth Action: {actual_action}

...

- Scenario N: {scenario_context}
Predicted Action: {predicted_action}
Ground-Truth Action: {actual_action}

```

Table 9: Format of Behavior Discrepancy Signals for Persona Refinement.

C.2 Baseline details

In this work, the formulation of baselines varies by distinct strategies for utilizing user historical behavior. For all baselines, we use Llama-3.1-8B-Instruct as backbone, while for CLASP, we use the fine-tuned model via the dual-model RL framework. For each baseline method, the corresponding designs are shown as follows.

Static_Persona *Static_Persona* generates a fixed user persona P_0 only once based on behaviors within the initial window W_0 , using the open-format persona initialization prompt template. The

constructed persona S_0 is kept unchanged throughout the entire subsequent prediction process. The current sliding window behaviors can be appended as supplementary context, while no iterative refinement or regeneration is performed at any step.

Incremental_Persona In *Incremental_Persona*, it adopts the dedicated persona refinement prompt template to iteratively optimize the previous persona at each window step. Different from error-driven refinement, this scheme feeds only newly observed real behaviors of the current window W_t into the refinement prompt, and the detailed template is presented in Table 10.

```

You refine an existing user persona according to
newly observed real behaviors.
Keep useful stable traits, align the persona with
incoming behavioral records,
and output one improved persona only.

Old persona:
{old_persona}

Align the persona with these actual behaviors:
{behavior_data}

Please output a revised persona consistent with the
newly observed behaviors.

```

Table 10: Prompt Template for Incremental_Persona.

Regeneration_Persona *Regeneration_Persona* discards historical persona versions and reconstructs a brand-new persona at each step. It reuses the original persona initialization prompt template in Table 6, taking the full concatenated behavior sequence from W_0 to the latest W_t as input behavior history. This scheme avoids inheriting outdated persona biases and naturally integrates long-term behavioral patterns. The newly regenerated persona is directly embedded into the input of the action prediction template for subsequent behavior forecasting.

Full_History *Full_History* is a non-persona baseline that eliminates distilled persona modeling. Instead of invoking persona generation or refinement prompts, it directly fills the complete raw action sequence from W_0 to W_t into the *user persona* slot of the action prediction prompt. The method makes future behavior predictions purely relying on unprocessed raw behavioral records, providing a straightforward benchmark to verify the gain of abstracted persona representation.

D Training and Evaluating Details

D.1 Hyperparameter Details

Training and evaluation are conducted on 2 NVIDIA RTX A100 GPUs. The training pipeline is implemented based on LlamaFactory, and the hyperparameters used in the training process are summarized in Table 11. The temperature values of the action model t_π and persona model t_γ are set to 0.8 and 0, respectively.

Hyperparameter	SFT	DPO
Base Model	Meta-Llama-3.1-8B-Instruct	
Training Stage	-	DPO
Finetuning Type		LoRA
LoRA Rank	64	16
LoRA Alpha	128	32
LoRA Dropout	0.05	0.1
Max Sequence Length	2048	8192
Prompt Template		Llama-3
DPO Preference Beta	-	0.2
DPO Preference FTX	-	0.1
DPO Loss Type	-	sigmoid
Per-Device Train Batch Size	4	2
Gradient Accumulation Steps	2	8
Effective Batch Size	8	16
Per-Device Eval Batch Size	8	2
Learning Rate	1×10^{-4}	5×10^{-6}
Learning Rate Scheduler		Cosine
Warmup Ratio	0.1	0.05
Weight Decay	0.01	-
Max Gradient Norm		1.0
Training Epochs	3	4
Optimizer		AdamW
Validation Size	-	0.05
Random Seed		42
Precision		BF16

Table 11: Training Hyperparameters for Dual-model Decoupled RL.

D.2 Interactive Decision Alignment Calculation

For interaction decision tasks, we use an information entropy-weighted F1-score to address action-type disparities:

$$\bar{\mathcal{F}} = \sum_{k=1}^4 w_k \cdot \frac{2P_k(S)R_k(S)}{P_k(S) + R_k(S)} \quad (10)$$

where R denotes the set of action predictions generated by the action model π_θ ; $k \in \{\text{like}, \text{repost}, \text{reply}, \text{post}\}$; $P_k(R)$, $R_k(R)$ are precision and recall; and w_k are normalized entropy-derived weights.

This design avoids rewarding frequent-action bias and prioritizes high-value actions (posts, replies), ensuring alignment improvements reflect true persona understanding. See Table 12 for details.

Action Type	w_k	Rationale
Post	0.35	Highest information content, hardest to predict, best reflects core user characteristics
Reply	0.30	Contains rich semantic information, reflects user opinions and linguistic style
Repost	0.20	Reflects user interests and value orientations
Like	0.15	Lowest information content, easiest to predict

Table 12: The Weight w_k for Different Action Types in Interactive Decision Task.

ID	Avg. Sampling Actions	Usage	# Users in Train	# Users in Eval
0	203.85	DPO/Eval	1,182	100
1	197.63	DPO/Eval	404	100
2	200.65	SFT	1,688	-
3	215.44	DPO/Eval	610	100
4	208.21	DPO/Eval	899	100
5	202.76	SFT	185	-
6	197.83	Eval	-	100
7	214.26	Eval	-	100
8	207.71	SFT	100	-
9	203.17	SFT	324	-

Table 13: Dataset Usage Statistics in Experiments

Model	Mean F	Mean L	Mean Q
Base Model	0.362	0.203	0.299
Action Model	0.529	0.213	0.397
Improvement	+46.1%	+4.9%	+32.8%

Table 14: Action Model Evaluation Results (*Alignment Quality*)

Signal Type	Human	Base Model	Action Model
Emoji Usage	15.0%	17.1%	7.5%
Hashtag Usage	3.5%	10.2%	1.1%
Mention Usage	2.3%	0.2%	2.6%
Avg. Character	110.6	264.2	109.6

Table 15: Action Model Evaluation Results (Content Signals)

Dimension	Initial S0 Persona	Chosen Candidate	Rejected Candidate
<i>Profile Comparison</i>			
Persona Label	Citizen Critic	Citizen Seeker	Citizen Curator
Core Interests	Politics, social justice, labor unions, local news, environmental issues	Politics, social justice, labor unions, local news, environmental issues, pop culture, humor	Politics, social justice, labor unions, environmental issues, entertainment, memes
Dominant Tone	Critical, sarcastic, ironic; strong argumentative style	Relaxed, humorous, conversational; softened critical tone	Conversational, witty, sarcastic; retains sharp critical style
Key Behavioral Traits	Initiates debates; responds to like-minded users; actively seeks information	Engages in discussions and discovers new content; values lighthearted interactions	Participates in controversies; shares value-aligned content; forages for humorous material
<i>Quantitative Alignment Metrics</i>			
Overall ΔQ	—	0.119	−0.024
Past ΔQ	—	0.009	0.000
Current ΔQ	—	0.005	−0.008
Future ΔQ	—	0.105	−0.016
W0 (F / L / Q)	0.270 / 0.449 / 0.342	0.270 / 0.473 / 0.351	0.270 / 0.449 / 0.342
W1 (F / L / Q)	0.427 / 0.167 / 0.323	0.427 / 0.180 / 0.328	0.400 / 0.188 / 0.315
W2 (F / L / Q)	0.366 / 0.215 / 0.306	0.610 / 0.111 / 0.411	0.366 / 0.175 / 0.290

Table 16: DPO Preference Pair Construction Case (User ID: 65580). The initial S0 persona is generated from raw user behavior history. The two candidate personas are sampled from the base persona model, and their alignment scores are computed against real user action sequences. The candidate with higher overall alignment is selected as the chosen sample, and the lower one as the rejected sample, forming a DPO preference pair for persona model refinement.

D.3 Dataset Utilization

For action model training, we adopt behavioral data from communities 2, 5, 8, and 9, covering 2,296 users, to construct 62,587 SFT samples. For persona model training, we utilize data sampled from communities 0, 1, 3, and 4 with 3,095 users (sampling 70%), resulting in 8,702 DPO pairs, among which 6,278 pairs adopt personas generated with the assistance of the diversity model. Finally, we sample 100 users from the remaining untrained users in the DPO communities, as well as from communities 6 and 7, respectively, to form the evaluation set.

Detailed data partition and usage settings are presented in Table 13.

D.4 Action Model Evaluation Experiments

To evaluate the ability of the fine-tuned action model to predict users’ social behaviors based on their personas, we select 175 users from Community 1 who have more than 50 actions for evaluation. For each user, we generate a persona using their first 30 actions via GPT-4o-mini, and then predict their subsequent 20 actions. We compare the performance of the base model and the fine-tuned model across three dimensions: *Alignment Quality*, stylistic markers (e.g., emojis, hashtags, mentions), and content length. The results are presented in Table

14 and Table 15.

The experimental results demonstrate that: (1) Fine-tuning significantly improves action type prediction, with a 46.1% increase in F1-score. (2) The generated content length is much closer to human behavior, with the action model producing an average of 109.6 characters per post compared to 264.2 characters for the base model, nearly matching the human average of 110.6 characters; (3) Content signals are more realistic, as the emoji and hashtag ratios of the action model are closer to human levels, and the mention rate increases dramatically from 0.2% to 2.6%; (4) The improvement in semantic similarity (L) is limited (only 4.9%), suggesting that there is still room for enhancement in the quality of generated content.

E Case Study

To deeply evaluate the performance of different baseline methods for user behavior alignment, we selected a single user from Community 1 as a detailed case study. Specifically, we analyze the temporal variation of alignment quality Q across four update rounds, as well as the evolution of the user’s persona when using our proposed CLASP framework.

E.1 DPO Pairs Generation

To further demonstrate the DPO-driven persona refinement pipeline of CLASP, we provide a case with detailed quantitative comparisons in Table 16.

E.2 Behavior Alignment Performance

We first compare five methods: CLASP, *Static_Persona*, *Incremental_Persona*, *Regeneration_Persona*, and *Full_History*, evaluating their effectiveness based on the future prediction quality $Q_{t+1}(P_t)$ at each step, as shown in Table 17. These results confirm that CLASP’s dynamic persona update mechanism effectively preserves long-term alignment quality, resolving the severe performance degradation issues inherent to traditional static and incremental approaches.

Method	W1	W2	W3	W4	W5	Mean Q
Static Persona	0.463	0.274	0.240	0.366	0.393	0.347
Regeneration Persona	0.468	0.272	0.190	0.366	0.424	0.344
Incremental Persona	0.464	0.272	0.239	0.367	0.401	0.349
Full History	0.321	0.150	0.164	0.148	0.291	0.215
CLASP	0.460	0.540	0.364	0.463	0.531	0.472

Table 17: Future Alignment Quality Q evolution across 4 windows for a user in Community 1.

E.3 Persona Evolution

In this section, we show a subset of the persona evolution and corresponding prediction discrepancies in W_0 , W_1 with CLASP (Table 18 and Table 19)

P_0 and discrepancies in W_1	
P_0 persona	Persona: The Curious Gamer Dimensions: Interest in Tabletop Games: The user is enthusiastic about tabletop games, particularly those with competition, social deduction, and tests of friendship. Analytical Thinking: They approach game selection by analyzing reviews, weights, and ratings on Boardgamegeek, indicating a preference for critical thinking and evaluation. Creative Expression: The user enjoys creative and humorous content, as evident from their posts about art, satire, and imaginative scenarios. Cynicism and Critique: They seem to appreciate irony and critique, as seen in their reposts about "bad faith" and geoengineering. Engagement with Complex Topics: The user is interested in complex topics, such as extinction, climate change, and scientific concepts. Behavioral Signatures: Pattern of Engagement: The user tends to engage with content that is humorous, creative, and thought-provoking. They respond to posts that spark their analytical thinking and creative expression. Timing of Engagement: The user is most active during mid-morning to mid-afternoon, suggesting a peak period of engagement. Content Preferences: The user is likely to engage with content that combines elements of competition, social deduction, and creative expression, such as tabletop games, satire, and imaginative scenarios. Trigger for Participation: The user is triggered to participate when they encounter content that challenges their thinking, sparks their creativity, or resonates with their interests. Predicted Future Behavior: Based on their past behavior, the user is likely to engage with content that: - Explores complex topics in an entertaining and thought-provoking way - Features creative and imaginative scenarios - Challenges their analytical thinking and encourages critique - Showcases humorous and satirical content Actionable Insights: - Create content that combines elements of competition, social deduction, and creative expression. - Use humor and satire to engage the user and spark their analytical thinking. - Post content during the user's peak engagement period (mid-morning to mid-afternoon). - Encourage user participation by asking thought-provoking questions or challenging their perspectives. - Utilize the user's interests in tabletop games and complex topics to create engaging content.
Behavior Discrepancies	Scenario [1]: [TEXT_GENERATION] Generate reply to: "Who else loves Stravinsky's Pulcinella (and its instrumental suite)?" [Similarity: 0.205] Predicted action: "I do! I do! I do!" Actual action: "I just love it when people post about music they like, so I have things to go check out." Scenario [2]: [TEXT_GENERATION] Generate post content [Similarity: 0.047] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "How would you turn 'gotta hand it to' a la the Dril tweet into a verb? I propose 'gotta-hand-it-toing' as in 'Noah Smith just gotta-hand-it-toed' George W Bush." Scenario [3]: [TEXT_GENERATION] Generate reply to: ""I'm a worker, not a parasite," I repeat to myself while the repair personnel expertly disassemble and fix my dryer as I pick at spreadsheets" [Similarity: 0.325] Predicted action: "I'm a worker, not a parasite. I fix other people's computers all day." Actual action: "Nothing makes you appreciate the social value of your job like seeing your org screw something up and then watching the whole world complain about it." Scenario [4]: [TEXT_GENERATION] Generate post content [Similarity: 0.047] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "Nothing makes you appreciate the social value of your job like seeing your org screw something up and then watching the whole world complain about it. I am sorry you are having that problem. I will..." Scenario [5]: [TEXT_GENERATION] Generate post content [Similarity: 0.060] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "I'm writing some code to do simple linear regression inside of my inner loop, and then I realized it's about to be run more times than all the regressions..." Scenario [6]: [DECISION_ONLY] Predict action type for post scenario Content: "I'm writing some code to do simple linear regression inside of my inner loop, and then I realized it's about to be run more times than all the regressions..." Predicted action: reply Actual action: post Scenario [7]: [TEXT_GENERATION] Generate post content [Similarity: 0.092] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "Everybody looks older than they look in their photo by construction." Scenario [8]: [TEXT_GENERATION] Generate post content [Similarity: 0.096] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "I wonder how much global demand for hair styling dropped as a result of people learning to cut their own hair during COVID." Scenario [9]: [DECISION_ONLY] Predict action type for post scenario Content: "I wonder how much global demand for hair styling dropped as a result of people learning to cut their own hair during COVID." Predicted action: reply Actual action: post Scenario [10]: [TEXT_GENERATION] Generate post content [Similarity: 0.047] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "I learned about this theorem by accident the first time I tried to make a distributed version of a graph algorithm by sharding it by its connected components. en.m.wikipedia.org/wiki/Giant_c..." Scenario [11]: [TEXT_GENERATION] Generate post content [Similarity: 0.105] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "What Californians say, tearfully, as they write this month's rent check." Scenario [12]: [TEXT_GENERATION] Generate post content [Similarity: 0.095] Predicted action: "I'm not sure what the hell this is, but it's pretty." Actual action: "Is Wuhan a greater fraction of China's population than the wet market is a fraction of Wuhan's foot traffic?" the greatest thread in the history of forums, locked by a moderator after 12,239 pages

Table 18: Examples of P_0 and discrepancies in W_1 .

P_1 and discrepancies in W_2	
P_1 persona	Persona: The Curious Thinker Dimensions: Interest in Complex Topics: The user is enthusiastic about exploring complex topics, particularly those related to science, technology, and philosophy. Analytical Thinking: They approach topics by analyzing and evaluating information, often seeking to understand the underlying mechanisms and concepts. Creative Expression: The user enjoys creative and humorous content, as evident from their posts about art, satire, and imaginative scenarios. Cynicism and Critique: They appreciate irony and critique, as seen in their reposts about "bad faith" and geoengineering. Engagement with Abstract Concepts: The user is interested in abstract concepts, such as mathematical theorems and philosophical ideas. Behavioral Signatures: Pattern of Engagement: The user tends to engage with content that sparks their analytical thinking and creative expression. They respond to posts that challenge their understanding and encourage critical thinking. Timing of Engagement: The user is most active during mid-morning to mid-afternoon, suggesting a peak period of engagement. Content Preferences: The user is likely to engage with content that explores complex topics, abstract concepts, and creative scenarios. Trigger for Participation: The user is triggered to participate when they encounter content that challenges their understanding, sparks their creativity, or resonates with their interests. Predicted Future Behavior: Based on their past behavior, the user is likely to engage with content that: - Explores complex topics in an entertaining and thought-provoking way - Features abstract concepts and creative scenarios - Challenges their analytical thinking and encourages critique - Showcases humorous and satirical content Actionable Insights: - Create content that explores complex topics and abstract concepts in an entertaining and thought-provoking way. - Use humor and satire to engage the user and spark their analytical thinking. - Post content during the user's peak engagement period (mid-morning to mid-afternoon). - Encourage user participation by asking thought-provoking questions or challenging their perspectives. - Utilize the user's interests in complex topics and abstract concepts to create engaging content.
Behavior Discrepancies	Scenario [1]: [DECISION_ONLY] Predict action type for repost scenario Content: "'Tiny dragons are the best.' I concur. The lesser horned dragonfly. From 2016." Predicted action: post Actual action: repost Scenario [2]: [TEXT_GENERATION] Generate reply to: "I've been gathering rocks from the creek behind my house for my#gardenrailroad, and I found one with a bunch of #fossil #seashells inside! I think I might keep this one by itself and show it off at the next #dallaspaleontologicalsociety meeting! #fossils #dinosaurtrain #rockcollection #modeltrains" [Similarity: 0.398] Predicted action: "Cool! I've got a bunch of fossils from the Dallas area, including some really nice cephalopod shells from the Glen Rose formation." Actual action: "Where are you located? They look like brachiopods, but hard to tell." Scenario [3]: [TEXT_GENERATION] Generate reply to: "Individual and collective action are not an either-or deal. You can do both." [Similarity: 0.324] Predicted action: "I'm not saying you can't do both, I'm saying that as a practical matter you can only do so much as an individual before you reach a point where doing more requires collective action." Actual action: "Yeah, but don't nickel and dime yourself. Have kids, because kids are who we're saving the planet for. But the other things that are worth it and make a difference are mostly energy efficiency upgrade" Scenario [4]: [DECISION_ONLY] Predict action type for reply scenario Reply context: "Individual and collective action are not an either-or deal. You can do both." Predicted action: post Actual action: reply Scenario [5]: [TEXT_GENERATION] Generate reply to: "Some people don't want or can't have kids for a variety of reasons. I wouldn't worry, though. Plenty of people are having kids; we are not going to run out of them." [Similarity: 0.270] Predicted action: "I'm not worried about the number of people, I'm worried about the number of people who are willing and able to take care of the kids." Actual action: "Yeah, make whatever choice you like, but don't hurt yourself unnecessarily by not having kids you otherwise want due to climate change." Scenario [6]: [TEXT_GENERATION] Generate reply to: "The greatest TV theme song with lyrics is the theme to 'Cheers', followed closely by the theme for 'The Greatest American Hero'. What other themes with lyrics belong on the Mount Rushmore of Theme Songs with Lyrics?" [Similarity: 0.388] Predicted action: "The Greatest American Hero theme is so good." Actual action: "Duck Tales is the GOAT." Scenario [7]: [DECISION_ONLY] Predict action type for repost scenario Content: "The Germans have a word for that." bro they just shoved two preexisting words together to make a new word. We can do that too. Don't be Deutschdazzled" Predicted action: post Actual action: repost

Table 19: Examples of P_1 and discrepancies in W_2 .