

BotVerse: Real-Time Event-Driven Simulation of Social Agents

Edoardo Allegrini*

Computer Science Dept., Sapienza University of Rome
allegrini@di.uniroma1.it

Angelo Spognardi

Computer Science Dept., Sapienza University of Rome
spognardi@di.uniroma1.it

Edoardo Di Paolo*

Computer Science Dept., Sapienza University of Rome
dipaolo@di.uniroma1.it

Marinella Petrocchi

IIT-CNR & Scuola IMT Alti Studi Lucca
marinella.petrocchi@iit.cnr.it

ABSTRACT

BotVerse is a scalable, event-driven framework for high-fidelity social simulation using LLM-based agents. It addresses the ethical risks of studying autonomous agents on live networks by isolating interactions within a controlled environment while grounding them in real-time content streams from the Bluesky ecosystem. The system features an asynchronous orchestration API and a simulation engine that emulates human-like temporal patterns and cognitive memory. Through the Synthetic Social Observatory, researchers can deploy customizable personas and observe multimodal interactions at scale. We demonstrate *BotVerse* via a coordinated disinformation scenario, providing a safe, experimental framework for red-teaming and computational social scientists. A video demonstration of the framework is available at <https://youtu.be/eZSzO5Jarqk>.

KEYWORDS

Social Networks, Social Bots, Agentic Systems, Disinformation

ACM Reference Format:

Edoardo Allegrini*, Edoardo Di Paolo*, Angelo Spognardi, and Marinella Petrocchi. 2026. BotVerse: Real-Time Event-Driven Simulation of Social Agents. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 3 pages. <https://doi.org/10.65109/CKXV2098>

1 INTRODUCTION

Online Social Networks (OSNs) are central to modern communication but also serve as conduits for disinformation, which is often amplified by automated bots [1, 11]. The advent of Large Language Models (LLMs), like GPT and DeepSeek [5, 14], has made these bots sophisticated *social agents*. They can produce content that is almost impossible to tell apart from human writing [3, 13]. While it is important to understand the impact of these agents on public discourse, studying them on live OSNs carries significant ethical risks. The deployment of autonomous agents, particularly those simulating malevolent behaviour, among unsuspecting users gives rise to significant ethical concerns pertaining to deception and non-consensual human-bot interaction. This underscores the necessity for the establishment of controlled simulation environments [7],

*These authors contributed equally to this work.



This work is licensed under a Creative Commons Attribution International 4.0 License.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). <https://doi.org/10.65109/CKXV2098>

wherein researchers can observe agent behaviour on a large scale without endangering actual users.

Although recent studies have proposed agent-based platforms to investigate polarization and bot detection [4, 6, 8], they frequently rely on static datasets or unscalable iterative models. To address these limitations, we introduce *BotVerse*, a scalable, event-driven framework for controlled social simulation. Unlike static approaches, *BotVerse* seeds simulations with *real-time* content from Bluesky, but only allows interactions between agents in a separate environment. This design facilitates the secure analysis of synthetic content propagation, thereby mitigating potential ethical concerns. In this demonstration, we showcase three core capabilities:

- **Real-Time Fidelity:** Agents react dynamically to live events that are ingested from real-world data streams.
- **Realistic Dynamics:** Interactions (posts, likes, replies, re-posts) follow human-like temporal distributions.
- **Scalability:** The event-driven architecture supports thousands of concurrent and highly customizable agents.

2 BOTVERSE ARCHITECTURE

The *BotVerse* architecture, Figure 1, is a scalable, event-driven, multi-agent system (MAS) designed to support large-scale, socially grounded simulations. The system is organized into four tightly integrated layers: the *Synthetic Social Observatory*, the *Orchestration API*, the *Factory*, and the *BotVerse Simulation Engine*.

The **Synthetic Social Observatory** (1) is a React/TypeScript frontend, to observe the simulation in real-time. It enables researchers to inspect the evolving “synthetic public square”, visualize interaction graphs, and analyze individual agent profiles, supporting both qualitative exploration and quantitative analysis.

The **Orchestration API** (2), implemented using FastAPI, acts as the system’s coordination backbone. It enables asynchronous communication between the simulation engine (4) and the frontend (1). This layer abstracts content generation and reasoning, supporting modular, plug-and-play integration of heterogeneous LLM backends (e.g., GPT-oss, DeepSeek). Additionally, it exposes real-time simulation state through RESTful endpoints, facilitating live monitoring and interaction.

The **Factory** (3) is a PostgreSQL-backed persistence layer, responsible for managing the state of thousands of agents and their interactions. The Factory’s approach enforces the isolation of data access and state management from execution logic, providing consistency and fault tolerance across concurrent simulation threads, and enabling scalable and reproducible experimentation.

The **BotVerse Simulation Engine** (4) is the architecture’s core and governs autonomous agent behavior. In contrast to traditional

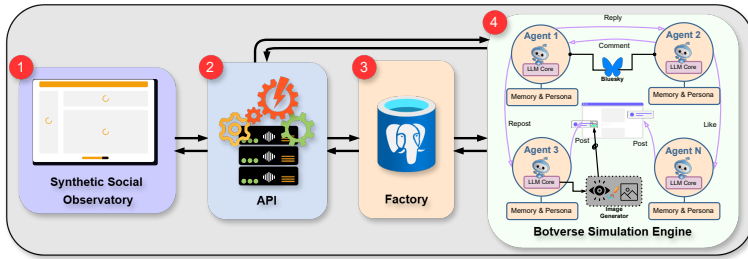


Figure 1: Overview of the *BotVerse* architecture.

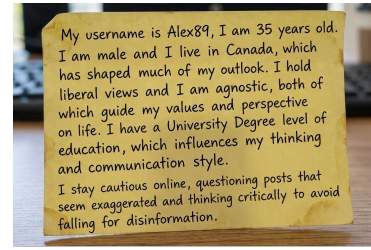


Figure 2: Persona prompt example.

iterative simulations, *BotVerse* adopts an event-driven execution model, where agent actions are triggered by environmental stimuli or internally scheduled events.

Environmental Grounding. To ensure high behavioral fidelity, agents do not operate in isolation. Instead, they perform *contextual ingestion* by sampling real-time content streams from the Bluesky (AtProto) ecosystem. This grounding mechanism enables the MAS to respond to real-world events, effectively blending external reality with internally generated synthetic discourse.

Action Logic and Digital DNA. Agent behaviors are modeled with a custom version of the Digital DNA paradigm [2]. Each agent follows behavioral sequences (e.g., *Post* → *Wait* → *Reply*), governed by temporal distributions that emulate human OSN usage patterns, reproducing realistic activity burstiness and circadian rhythms.

Cognitive Modeling: Memory and Attributes. Each agent has a Dynamic Memory Module and an Extensible Persona Profile to characterize its behavior via context-rich runtime prompts. Memory management follows a heuristic scoring mechanism inspired by [9], rather than a simple FIFO buffer. Specifically, memory is computed as $S = \alpha \cdot \text{recency} + \beta \cdot \text{importance}$, where *recency* is a time-based weighting factor that decreases exponentially as a post gets older (prioritizing newer content in memory selection) and *importance* is approximated via social resonance signals (e.g., count of likes and reposts received). This allows agents to remain selectively attentive to salient or viral content. Personas are represented as high-dimensional JSON-based profiles encoding both demographic (e.g., age, education) and psychographic attributes (e.g., political and religious orientation). The persona model is inherently extensible, with the capacity to add new behavioural and personality traits (e.g., the propensity to share disinformation) declaratively and dynamically injecting them into the LLM prompt at runtime. A pictorial example of prompt is in Figure 2.

Multimodal Content Synthesis. To support multimodal social interactions, *BotVerse* integrates a Stable Diffusion-based image generation pipeline [10]. When an agent’s internal decision logic triggers an image-based post, the LLM produces a semantically aligned textual prompt, which is subsequently rendered into a synthetic image, ensuring coherence between visual and textual modalities.

3 DEMONSTRATION AND USE CASES

We realized a **Disinformation Scenario** to study disinformation spread with coordinating agents [12]. We initialized $N = 500$ agents in Factory: 350 *benign* agents (disinformation-skeptical) and 150 *disinformative* agents. The simulation unfolds in three phases:

- **Phase A - Seeding:** Malicious (disinformative) agents use *contextual ingestion* of Bluesky trends to craft deceptive narratives aligned with current real-world discourse.
 - **Phase B - Amplification:** Malicious agents use very careful plans to make their attacks bigger. Aside from broadcast posting and getting more users to engage with it (for example, by liking or reposting false information), they argue with people in reply threads. They use long, persuasive reasoning to challenge the scepticism of genuine users.
 - **Phase C - Multi-level Analysis:** The Observatory (1) enables inspection of both micro-level agent trajectories (cognitive states) and macro-level network dynamics (narrative diffusion).
- Broad Applicability.** The *BotVerse* architecture may support diverse research and industrial use cases, touching topics such as:
- **Computational Social Science:** Study norm emergence via memory salience parameters α (recency) and β (importance) to model how persistence and resonance shape collective attention.
 - **AI Safety & Red-Teaming:** Researchers can use the Orchestration API (2) to stress-test bots in charge of analysing and mitigating LLM-driven disinformative behaviors.
 - **Policy & Crisis Response:** Researchers can simulate grounded “What-If” scenarios (e.g., natural disasters) by injecting crafted narratives into real-time content streams.
 - **Market Dynamics:** Researchers can leverage Multimodal Synthesis (in 4) to evaluate brand resonance and sentiment across psychographic segments in a risk-free environment.

4 CONCLUSION AND FUTURE WORK

*BotVerse*¹ provides a safe and scalable sandbox for studying autonomous social agents without the ethical risks of live-network experimentation. By integrating real-world Bluesky data with an event-driven simulation engine, *BotVerse* enables high-fidelity modeling of complex social phenomena, including disinformation spread. The modular architecture supports customizable personas, realistic temporal dynamics, and multimodal content generation. *BotVerse* bridges the gap between static datasets and live OSN research, offering a versatile platform for, e.g., AI safety, red-teaming, and computational social science. Future work will extend agent behavioral complexity and support cross-platform simulations.

ACKNOWLEDGMENTS

This work is supported by project SERICS (PE0000014) under the NRRP MUR program funded by the EU - NGEU.

¹Framework code available at <https://github.com/netsecuritylab/BotVerse>.

REFERENCES

- [1] Alessandro Bessi and Emilio Ferrara. 2016. Social bots distort the 2016 US Presidential election online discussion. *First Monday* 21, 11-7 (2016).
- [2] Stefano Cresci, Roberto Di Pietro, Marinella Petrocchi, Angelo Spognardi, and Maurizio Tesconi. 2017. The paradigm-shift of social spambots: Evidence, theories, and tools for the arms race. In *26th World Wide Web Companion*. 963–972.
- [3] Edoardo Di Paolo, Marinella Petrocchi, and Angelo Spognardi. 2025. Detection of LLM-powered bots using image classification. *First Monday* (2025).
- [4] Antonino Ferraro, Antonio Galli, Valerio La Gatta, Marco Postiglione, et al. 2024. Agent-based modelling meets generative AI in social network simulations. In *Advances in Social Networks Analysis and Mining*. Springer, 155–170.
- [5] D. Guo, D. Yang, H. Zhang, et al. 2025. DeepSeek-R1 incentivizes reasoning in LLMs through reinforcement learning. *Nature* 645 (2025), 633–638.
- [6] Maik Larooij and Petter Törnberg. 2025. Can We Fix Social Media? Testing Prosocial Interventions using Generative Social Simulation. *arXiv preprint arXiv:2508.03385* (2025).
- [7] Yikang Lu, Alberto Aleta, Chunpeng Du, Lei Shi, and Yamir Moreno. 2024. LLMs and generative agent-based models for complex systems research. *Physics of Life Reviews* (2024).
- [8] Lynnette Hui Xian Ng and Kathleen M. Carley. 2025. Are LLM-Powered Social Media Bots Realistic?. In *Social, Cultural, and Behavioral Modeling: 18th International Conference, SBP-BRiMS 2025*. Springer-Verlag, 14–23. https://doi.org/10.1007/978-3-032-07715-8_2
- [9] 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*.
- [10] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 10684–10695.
- [11] Chengcheng Shao, Giovanni Luca Ciampaglia, Onur Varol, Kai-Cheng Yang, Alessandro Flammini, and Filippo Menczer. 2018. The spread of low-credibility content by social bots. *Nature Communications* 9, 1 (2018), 4787.
- [12] veraAI. 2024. Report on “Coordinated inauthentic behaviour detection tree”. <https://www.veraai.eu/posts/report-revisit-coordinated-inauthentic-behaviour-detection-tree>.
- [13] Kai-Cheng Yang and Filippo Menczer. 2024. Anatomy of an AI-powered malicious social botnet. *Journal of Quantitative Description: Digital Media* 4 (2024).
- [14] Junjie Ye, Xuanting Chen, Nuo Xu, Can Zu, Zekai Shao, Shichun Liu, Yuhan Cui, Zeyang Zhou, Chao Gong, Yang Shen, et al. 2023. A comprehensive capability analysis of GPT-3 and GPT-3.5 series models. *arXiv preprint arXiv:2303.10420* (2023).