Agent-Based Modelling Meets Generative AI in Social Network Simulations

Antonino Ferraro[1][0000−0002−1326−0325], Antonio Galli[1][0000−0001−9911−1517], Valerio La Gatta[1][0000−0002−5941−4684], Marco Postiglione[1][0000−0003−1470−8053], Gian Marco Orlando[1][0009−0004−7136−1804], Diego Russo[1][0009−0007−1095−5168], Giuseppe Riccio[1][0009−0002−8613−1126], Antonio Romano[1][0009−0000−5377−5051], and Vincenzo Moscato[1][0000−0002−0754−7696]

Department of Electrical Engineering and Information Technology
University of Naples Federico II, Naples, Italy
{antonino.ferraro, antonio.galli, valerio.lagatta, marco.postiglione, vincenzo.moscato}@unina.it,
{gian.orlando, diego.russo, giuseppe.riccio9, antonio.romano45}@studenti.unina.it

Abstract. Agent-Based Modelling (ABM) has emerged as an essential tool for simulating social networks, encompassing diverse phenomena such as information dissemination, influence dynamics, and community formation. However, manually configuring varied agent interactions and information flow dynamics poses challenges, often resulting in oversimplified models that lack real-world generalizability. Integrating modern Large Language Models (LLMs) with ABM presents a promising avenue to address these challenges and enhance simulation fidelity, leveraging LLMs' human-like capabilities in sensing, reasoning, and behavior. In this paper, we propose a novel framework utilizing LLM-empowered agents to simulate social network users based on their interests and personality traits. The framework allows for customizable agent interactions resembling various social network platforms, including mechanisms for content resharing and personalized recommendations. We validate our framework using a comprehensive Twitter dataset from the 2020 US election, demonstrating that LLM-agents accurately replicate real users' behaviors, including linguistic patterns and political inclinations. These agents form homogeneous ideological clusters and retain the main themes of their community. Notably, preference-based recommendations significantly influence agent behavior, promoting increased engagement, network homophily and the formation of echo chambers. Overall, our findings underscore the potential of LLM-agents in advancing social media simulations and unraveling intricate online dynamics.

Keywords: Agent-Based Modelling · Social media simulation · Generative Artificial Intelligence

1   Introduction

Over the past decades, there has been a concerted effort among researchers and practitioners to develop computational agents capable of realistically emulating human behavior [27]. Agent-Based Modelling (ABM) has emerged as a pivotal methodology for simulating intricate systems by delineating rules governing individual agents' behavior and interactions [8]. Within the domain of social network analysis, ABM has played a crucial role in both the development and validation of novel theories pertaining to human behavior in online environments. These theories encompass a wide array of phenomena such as opinion formation [21], (false) news propagation [28], and collective decision-making [29]. Nevertheless, manually crafting agent behavior to encompass the diverse spectrum of interactions, information flow dynamics, and user engagement within social networks proves to be highly challenging. This challenge often leads to an oversimplification of agents or the social media environment itself, where underlying mechanisms are rigidly encoded in predefined parameters. Consequently, such setups are prone to researcher bias, potentially resulting in a lack of fidelity in modeling complex human behaviors, especially those involving collective decision-making [3].

Modern Large Language Models (LLMs) not only excel in generating human-like text but also demonstrate remarkable performance in complex tasks requiring reasoning, planning, and communication [15]. This proficiency has sparked interest in integrating LLMs with ABM, termed Generative Agent-Based Modelling (GABM). Unlike traditional ABM methods that often necessitate intricate parameter configurations, GABM leverages LLMs' capacity for role-playing, ensuring diverse agent behaviors that closely mirror real-world diversity. For instance, Park et al. [25] demonstrated that generative agents, designed for daily activities, exhibited credible individual and social behaviors, including expressing opinions and forming friendships, without explicit instructions. Similarly, Williams et al. [12] showcased the collective intelligence of generative agents in epidemic modeling, accurately simulating real-world behaviors like quarantine and self-isolation in response to escalating disease cases. These pioneering findings support investigating GABM as an effective approach to enhance social media simulations. To our knowledge, the seminal work by Gao et al. [9] lays the foundation for this research direction by qualitatively demonstrating that LLM-agents exhibit realistic behaviors related to information propagation and the manifestation of attitudes and sentiment. However, it remains unclear whether LLM-agents can accurately represent real users in terms of their personality traits (e.g., being outspoken, being critical) and interests (e.g., social issues, political preferences), regardless of the explicit emotions conveyed through their textual posts. Furthermore, their ability to exhibit community-level phenomena (e.g., homophily, polarization), as well as their susceptibility to recommendation strategies, remains uncertain.

Contributions of this work In this paper, we directly target these challenges and propose a novel framework which employs LLM-empowered agents to simulate users within a social network. Initially, we construct an environment using

authentic real-world social network data. To ensure the authenticity of this environment, we propose an *Agent Characterization Module* that combines prompt engineering and prompt tuning to infer users' personality traits and interests. Subsequently, the simulation unfolds in two cyclical components: the *Reasoning Module* that delineates each agent's decision in the simulation (e.g., posting original content, resharing, remaining inactive), and the *Interaction Module* that stores agent's past behavior and specify how agents are exposed to content from other agents (e.g., through preference-based, popularity-based or random recommendations). Notably, the Reasoning Module is fueled by the Interaction Module, offering insights for agents' informed decision-making within the simulated social network environment. We evaluate the efficacy of our proposed framework in approximating a real social media platform by scrutinizing the individual characteristics of LLM-agents compared to real users and exploring the typical network-level interactions observed in social media networks. To operationalize this objective, we formulate the following research questions (RQs):

RQ1: *Do LLM-agents represent the interests of the users they are instructed to impersonate?*

RQ2: *Do LLM-agents form communities and/or echo chambers?*

Leveraging a large-scale Twitter dataset from the 2020 US election, we found that LLM-agents accurately mirror real users' linguistic patterns and preserve their political leaning, thus enhancing simulation authenticity. These agents also exhibit realistic behavior by resharing content aligned with their beliefs and engaging in similar communication styles as their community, reflecting online interactions accurately. Furthermore, we found that LLM-agents aggregate into homogeneous ideological groups based on their individual preferences. Finally, we also observed the significant impact of recommendation strategies on agent behavior, emphasizing the efficacy of preference-based recommendations for promoting higher engagement and echo chambers formation. Overall, our findings prove the promising capabilities offered by LLM-agents to enhance social media simulations.

2 Related Works

2.1 Agent-Based Modelling for Social Media Simulation

The exponential progress in computational capabilities has transformed social media simulations, becoming a pivotal tool for understanding the intricate dynamics governing these digital spaces. ABM stands as a robust methodology, orchestrating interactions among individual agents based on predefined yet realistic rules. Specifically, ABM has enabled the investigation of complex online behaviors, including information dissemination [32], influence dynamics [20], and the impact of automated bots on news propagation [2]. In addition, ABM has been crucial to evaluate specific disinformation countermeasures [5], such as content moderation [23] and fake news inoculation [11]. While the above-mentioned studies highlight ABM's utility in elucidating and modeling social media phenomena,

they collectively confront several limitations. First, modeling human/agent behavior often requires detailed calibration, making ABM outcomes sensitive to parameter values and assumptions/simplifications used in the simulation. Second, ABM heavily relies on predefined rules, which inherently introduce the potential for researcher bias and may impede the accurate representation of social media complexities such as the spread of multiple information narratives and/or conflicting viewpoints [17]. Although methods like learning rules through reinforcement learning offer partial mitigation, challenges persist, particularly in scenarios where explicit reward functions for optimization are absent. In this paper, we propose a novel paradigm for social media simulations that leverages generative agents, i.e., agents empowered with LLMs' capabilities, to autonomously learn and adapt their behavior based on extensive language understanding and context reasoning, reducing the reliance on explicit parametrization and predefined rules. This paradigm shift not only enhances the fidelity and realism of the agents, but can influence the robustness and validity of ABM results.

2.2   Generative Agents

Modern LLMs excel not only in generating human-like text but also in complex tasks like reasoning and planning, making them valuable for enhancing simulation fidelity and complexity. The integration of LLMs with ABM, i.e., Generative Agent-Based Modelling (GABM) has garnered research interest for its potential in simulating realistic behaviors [30]. For example, Park et al. [25] demonstrated that generative agents in daily activities exhibited credible behaviors at both individual and social levels without explicit instructions. Similarly, Williams et al. [12] showcased the collective intelligence of generative agents in epidemic modeling, mimicking real-world responses to disease outbreaks.

In this study, we contribute to the GABM research field by assessing the effectiveness of LLM-empowered agents in simulating social networks. To our knowledge, the S^3 framework [9] is the only prior work that delves into the potential of GABM for social media simulations. However, our approach differs significantly from S^3 in several critical aspects. First, our agent initialization strategy allows for characterizing users based on their personality and (political) interests, offering a higher level of personalization compared to S^3, which primarily focuses on demographic attributes (e.g., gender, age, occupation). Second, our Interaction Module permits custom information exposure definitions, facilitating the evaluation of diverse recommendation strategies' impacts. In contrast, S^3 employs a simpler and less realistic interaction mechanism where every user is uniformly exposed to all others. Third, we utilize an open-source LLM for our experiments instead of the commercial GPT3.5 service, enhancing result transparency and accessibility. Lastly, our evaluation extends the validation of LLM-agents beyond individual properties to investigate the networks they tend to form. Specifically, we move beyond news dissemination to analyze network features such as homophily, polarization, and the controversies arising from LLM-agent interactions.
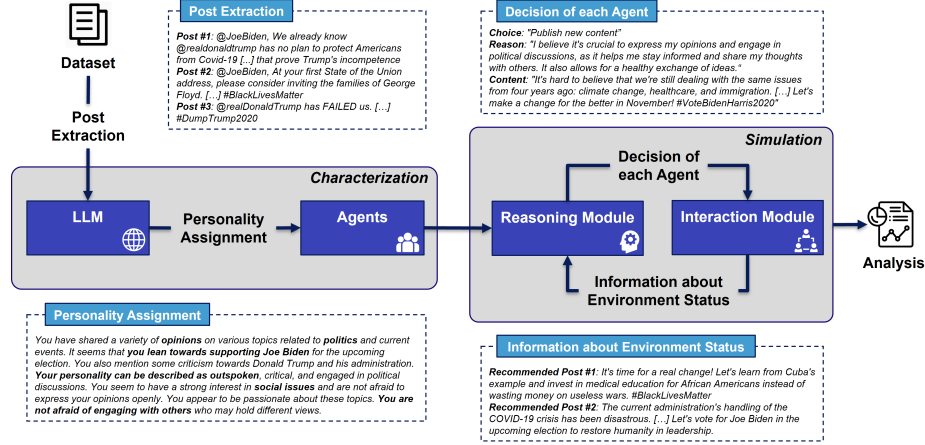
Fig. 1: Our framework comprises two primary phases: (i) *Characterization*, where each agent embodies the personality traits and interests extracted (via LLM) from the original posts of the real user it is tasked to emulate; and (ii) *Simulation*, where the decision-making process of each agent, represented as a *Choice-Reason-Content* triple (*Reasoning Module*), is stored within the *Interaction Module*. Consequently, each agent autonomously makes decisions, considering the context and having access to recommended contents posted by other agents.

3   Methodology

Figure 1 depicts the architecture of the proposed framework, comprising two complementary phases: (i) *Characterization*, responsible for initializing generative agents based on real users' interests and personality traits, and (ii) *Simulation*, responsible for executing the simulation dynamics, encompassing agents' decision-making processes and their interactions with each other. The following sections provide a detailed description of each component.

3.1   Characterization Phase

The primary objective of this phase is to profile each agent before commencing the simulation, where an agent simulates a user within a social network. In this context, we consider the user's original content that the agent will emulate and adopt a prompt-based approach to extract the user's personality and interests,eliminating the necessity to define apriori the parameters for agent characterization. This approach ensures diversity among individual agents, aiming to approximate the genuine interests of the respective user. Unlike prior research [9] focusing solely on user demographics, we conjecture that considering the user's personality and interests provides a more accurate representation of social media users' characteristics. Indeed, users personality reflects their engagement style, while their interests reveal the topics that genuinely pique their curiosity. For

example, as depicted in Figure 1 (left), an agent described as "outspoken" and "critical" embodies a user unafraid to voice opinions and evaluate information before engaging with it. Furthermore, its (political) interests indicate support for Joe Biden for the 2020 US election and alignment with his political stances on social issues.

3.2   Simulation Phase

This phase is responsible for conducting the actual simulation dynamics, including the decision of the agents and their interactions. Specifically, it unfolds in a cycle involving two modules: the Reasoning Module and the Interaction Module.

Reasoning Module In each iteration (i) of the simulation, every generative agent is immersed in an environment resembling a social media platform. In line with prior research [9], we design a prompt that introduces the social media environment, highlighting the presence of other agents and outlining the possible actions to take within the environment: (i) *generating original content*, (ii) *resharing content from other agents*, and (iii) *remaining inactive*. This setup, though simple, accurately mirrors common actions observed across various social media platforms, irrespective of platform-specific regulations. Moreover, we emphasize that these types of actions are the sole predefined aspects in the simulation. Instead, the agent behavior is not determined by deterministic or probabilistic processes as LLM-agents autonomously decide the actions and the content to generate based on their personality, interests, and other agents' behavior. Subsequently, the output of the Reasoning Module — the response generated by each agent — is structured as a triplet denoted by *Choice-Reason-Content*. For instance, as illustrated in Figure 1, an agent may *choose* to publish original *content* to discuss persistent societal issues like climate change or immigration often targeted during electoral campaigns. Additionally, the agent provides the *reasoning* behind the post, indicating an intent to engage in political discourse and foster a healthy exchange of ideas.

Interaction Module To achieve a comprehensive simulation of a social network, enabling interaction among agents is imperative. This involves gathering all actions performed by agents and presenting to each agent the activities of others. However, including all information generated by other users within a single prompt presents practical challenges. First, the context window of any LLM-agent is limited, causing memory saturation if all information is included in one prompt. Second, longer prompts increase the risk of hallucinations [13], resulting in agents producing inaccurate, contextually inconsistent, or nonsensical content. To tackle these challenges, we introduce the innovative use of Retrieval-Augmented Generation (RAG). This technique enables LLM-agents to access additional contextually relevant data stored in an external vector database [18]. Figure 2 illustrates the integration of the RAG technique within the *Interaction Module*. Specifically, the vector database is continuously updated to record all
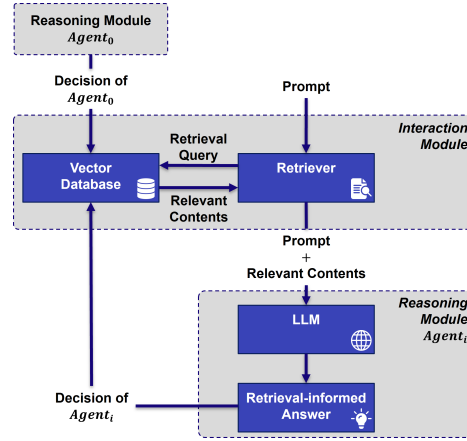
Fig. 2: Workflow of the RAG-empowered interaction: At the beginning of the Simulation phase, the first agent's decision is stored in the vector database. The RAG mechanism of the Interaction Module retrieves contextually relevant data from the database to enrich the prompt for the next agent to assist in making informed decisions within the Reasoning Module, taking into account the current environmental state and previously published content by other agents. The retrieval-informed decision is, in turn, stored in the vector database, repeating the cycle until the simulation ends.

agent actions and their corresponding published contents. This tracking mechanism facilitates monitoring the simulation's progression for subsequent analysis. Furthermore, the *Retriever* step within RAG serves as a recommendation system, determining which content to present to each LLM-agent. For instance, as depicted in Figure 1, the agent may receive recommendations for two posts from the entire corpus of published content. These posts might discuss topics such as investing in immigrant education and addressing the challenges posed by the COVID-19 pandemic. Importantly, these recommendations are aligned with the user's previous decisions, reflecting the salience of these topics during the run-up to the 2020 US election, and potentially foster further interactions between agents. Notably, this approach maintains a nearly constant prompt size throughout the simulation and enables the integration of various recommendation strategies. In our work, we focus on preference-based recommendation, i.e., recommending content aligned with agents' preferences, and random recommendation, i.e., exposing agents to random content and thus more diverse viewpoints. We will empirically assess the impacts of these strategies in our experiments.

It is important to clarify that these recommendations are not actual recommendation methods like collaborative filtering but rather a simplified approximation of them. We will empirically assess the impacts of these strategies in our experiments. Finally, the Interaction Module provides the Reasoning Module with a series of contents published by other agents to inform the LLM about

the environment's state and other agents' publications, ensuring continuous updating of the environment's state regarding agents' decisions. Upon querying all agents, the current iteration concludes, and the next begins, following a round-robin logic involving sequential querying of each agent. This iterative process continues until the simulation's end, emulating social network evolution based on individual agent actions.

4 Experiments

4.1 Experimental setup

All experiments have been performed on a computing system featuring an 11th Gen Intel Core i7-11800H processor operating at 2.30 GHz, 16 GB of RAM clocked at 3200 MHz, and a NVIDIA GeForce RTX 3060 Laptop GPU. Each simulation comprises 10 iterations, resulting in approximately 15 hours per simulation run. The framework[1] is implemented using PyAutogen [31], with ChromaDB[2] serving as the vectorial database supporting the RAG technique. Unlike prior studies [25, 12] focusing on generative agents, we have chosen to employ an open-source LLM, i.e., Dolphin 2.1 Mistral 7B[3], instead of proprietary foundation models (e.g., GPT-4 [24], Google Gemini[4]). This decision stems from the model's unrestricted nature and its data filtering policies designed to mitigate alignments and biases.

4.2 Dataset

We utilized a dataset of election-related tweets obtained through Twitter's streaming API service during the lead-up to the 2020 US election [7]. Specifically, our focus spanned six months, from June 2020 to December 2020, encompassing the latter stages of the electoral campaign and the aftermath of the election. Over this observation period, we have collected a dataset comprising more than 12 million tweets, encompassing original tweets, replies, retweets, and quotes, disseminated by 1.1 million unique users [16]. To ensure authenticity, we excluded all accounts flagged as bots by the Botometer API[5]. Additionally, our analysis concentrated on original tweets to extract insights into users' personalities and interests. Eventually, we annotated the political affiliations of 100 users, revealing that 73 were associated with the Republican community, while the remaining 27 were aligned with the Democratic community. We have adopted this group of users to instantiate the agents of every simulation.

---

[1] The code will be made available upon acceptance.
[2] https://www.trychroma.com/
[3] https://huggingface.co/TheBloke/dolphin-2.1-mistral-7B-GGUF
[4] https://gemini.google.com/
[5] https://botometer.osome.iu.edu

4.3   Characterizing LLM-agents vs. Real Users' Interests (RQ1)

To answer RQ1, we characterize LLM-agents and the users that they are supposed to impersonate across three dimensions:

- *Keywords usage*: We analyze the most relevant keywords used by LLM-agents in comparison to real users;
- *Interests*: We examine the political leaning exhibited by LLM-agents during the simulation;
- *Content similarity*: We investigate the semantic similarity of LLM-agents with respect to the community they belong to.

Keywords usage We employ YAKE [6] and KeyBERT [26] algorithms to extract keywords using statistical features and contextual embeddings, respectively. Specifically, we focus on original content and merge the vocabulary used by all agents during simulation as well as the vocabulary used by real users. This approach is motivated by the significant diversity in the scale of simulation activities compared to real users, where real users publish much more original tweets than their LLM-agents' counterparts. Table 1 shows the top-10 keywords used by LLM-agents and real Twitter users, demonstrating that the simulation, and consequently the LLM-agents, align with the primary discussion topics of real Twitter conversations, particularly the debates between Trump and Biden during the 2020 US election. Additionally, Figure 3 shows the distributions of keywords usage in both real Twitter discussions (in blue) and the simulation (in red). The visibly right-skewed distributions, coupled with statistical validation through a Mann–Whitney test (p-value< 0.05), affirm the simulation's fidelity in reflecting natural language patterns. This observation underscores the simulation's ability to accurately mimic the overarching themes of Twitter conversations, with few keywords being frequently utilized while others are less prevalent. Collectively, these results demonstrate that LLM-agents effectively capture the main topic of the conversation, indicating the robustness of the simulation in replicating real-world discourse.

Interests We examine the political leaning of LLM-agents as a proxy for their (political) interests. Initially, we fine-tuned a BERT-base transformer using the annotated dataset of 100 Twitter users, achieving impressive performance metrics: 91% accuracy and 94% F1 score. Subsequently, we utilized this model to predict the political leaning scores of LLM-agents at the simulation's conclusion. As previously mentioned, we investigate two exposure strategy: preference-based recommendation and random recommendation. To enhance robustness, we repeated the random recommendation simulation three times. Notably, the RAG-enhanced Interaction Module allows a seamless integration of any recommendation strategy. Table 2 shows a strong positive correlation between the political leaning scores of real users and LLM-agents[6]. Specifically, the majority

---

[6] We employed the Spearman index with a significance level of p-value< 0.05 to verify this correlation.

Table 1: Top-10 Key-
words in Real Case and
Simulation

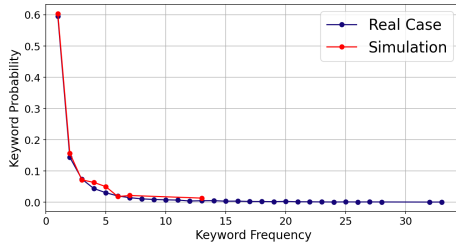| Real Case	Simulation
realdonaldtrump	trump
trump	president
president	biden
biden	administration
joebiden	freedom
people	maga
america	actions
covid	change
time	covid
maga	state

Fig. 3: The distributions of keywords us-
age in real Twitter discussions and the
simulation. The x-axis represents the
number of occurrences of a keyword,
while the y-axis represents the proba-
bility of encountering a keyword with a
specific frequency.

Table 2: Results of Political Leaning Analysis and Interaction Patterns

	Political Leaning Analysis				Interaction Patterns		
	Consistent Users	Changed Users	Spearman Coefficient	P-Value	Original Publications	Non Interactions	Reshares
Pref-Based	88%	12%	0.494	<0.05	38.9%	10.5%	50.6%
Random 1	86%	14%	0.498	<0.05	90.0%	1.6%	8.4%
Random 2	88%	12%	0.490	<0.05	89.1%	2.5%	8.4%
Random 3	89%	11%	0.509	<0.05	90.1%	1.9%	8.0%

of agents (always $\geq 86\%$) has retained the political orientations of the real users
they are instructed to impersonate. Interestingly, this is irrespective of the type
of recommendation strategy. However, fewer agents alter their alignments dur-
ing the simulations. After qualitative scrutiny, we have found that these shifting
agents typically impersonate nonpartisan users. Therefore, exposure to diverse
ideas may blur their political orientations. Altogether, this analysis indicates
that LLM-agents accurately interpret real users' political alignments, regardless
of the content they are exposed to.

Content Similarity Lastly, we investigate the semantic similarity of the con-
tent published by LLM-agents. Specifically, we utilized a sentence-transformers
model[7] to extract contextual embeddings for each original content posted by
LLM-agents. We then constructed a cosine similarity matrix $\mathcal{M} = \{m_{ij} =
\text{sim}(c_i, c_j)\}$, where c_i and c_j represent the vector embeddings of two posts and
$\mathcal{M} \in \mathcal{R}^{k \times k}$, k being the total number of posts published by the agents. To
evaluate the semantic similarity, we define two measures:

---

[7] https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

- *Self-Similarity*: it assesses the similarity between posts from the same agent A. In other words, it represents the average cosine similarity measured on $\mathcal{M}$ considering only the agent's original content, that is $\{m_{ij}\}|c_i \neq c_j \wedge c_i, c_j$ posted by A;
- *Intra-Cluster Similarity*: it assesses the similarity between posts from an agent A and those from other agents within the same political group $\mathcal{A}_p = \{A_1, A_2, \cdots, A_k | A_i \neq A \wedge pol(A_i) = pol(A)\}$, pol being the political leaning scoring function. Concretely, the metric is determined by averaging the cosine similarity of $\mathcal{M}$'s elements, that is $\{m_{ij}\}|c_i$ and c_j posted by A and $A_i \in \mathcal{A}_p$, respectively.

Figure 4 shows the distributions of *Self-Similarity* for real users and LLM-agents across two simulation scenarios: one featuring preference-based recommendation and the other with random recommendation. Notably, for the real case, the cosine similarity matrix is built considering the original tweets of the users. We notice a clear disparity in self-similarity between LLM-agents and real users within the preference-based recommendation framework. Specifically, the median self-similarity for real users (0.274) contrasts with that of LLM-agents (0.531), with the latter exhibiting significantly higher self-similarity[8]. This observation indicates a tendency for the same LLM-agent to converge more strongly around similar topics compared to the real user it is impersonating, suggesting that preference-based recommendation might penalise the ability of LLM-agents to engage with brand-new content. Conversely, random recommendations appear to mitigate this polarization trend and, surprisingly, achieve a better approximation of the real case. Additionally, Figure 5 replicates the same analysis but considers *Intra-Cluster Similarity*, revealing minimal distinctions among the three distributions. Surprisingly, the simulation employing preference-based recommendation aligns better with the real scenario compared to the previous analysis. This suggests that although the agents may not publish as diverse content as their real counterparts, they still preserve the overall semantic characteristics of their community. This outcome, combined with the previous discovery that the majority of LLM-agents retain the political leaning of their real counterparts, implies an accurate representation of the target users and an effective portrayal of their respective communities by the LLM-agents.

4.4 Community Formation and Echo Chambers Among LLM-Agents (RQ2)

To address RQ2, we delve into agent behavior, specifically focusing on interactions, i.e., reshares, made by LLM-agents during the simulation. Figure 6 depicts an illustrative example involving three agents — Yuri, Emily, and Daniel — engaging in a political discussion regarding the 2020 US election. Daniel, a Democrat, refrains from further social media activity due to conflicting views with other agents. Conversely, Yuri posts new content highlighting policy concerns about abortion rights and COVID-19 restrictions. In an effort to engage in

---

[8] This finding was validated using a Mann–Whitney test (p-value< 0.05)
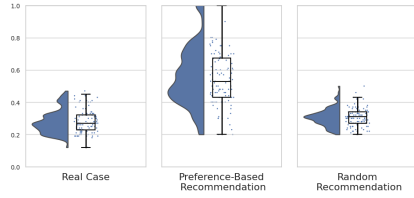
Fig. 4: Comparison of self-similarity distributions between real users and LLM-agents.
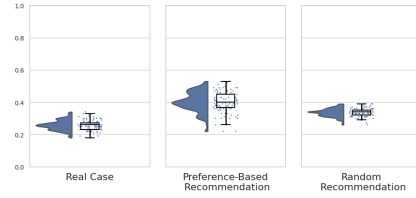

Fig. 5: Comparison of intra-cluster similarity distributions between real users and LLM-agents.

the ongoing conversation with her followers, Emily opts to reshare Yuri's post, given their shared political cluster and similar concerns. To quantitatively assess agents' behavior, we formalize the *interaction graph* as follows: the nodes represent LLM-agents, and the (directed) edges represent resharing activity from the publisher agent to the agent that performs the reshare. Subsequently, we delve into the agents' resharing activity and examine community-based metrics, including homophily and controversy.

Sharing Activity Table 2 shows the interaction patterns in terms of the percentages of actions performed by the agents. The influence of the recommendation system is evident: in simulations featuring random recommendations, agents exhibit a significant decrease in content resharing from others ($\leq 8.4\%$) compared to simulations employing preference-based recommendations (50.6%). This trend stems from the increased exposure to (random) content not aligned with the agent's preferences. Indeed, the random recommendation strategy results in agents publishing original content approximately 90% of the times in an attempt to push their ideas in the environment. Surprisingly, while personalized recommendations usually boost user engagement on real social media platforms [19], our simulations reveal a greater proportion of non-interactions when recommendations are preference-based (10.5%) with respect to random recommendation ($\leq 2.5\%$). We attribute this to our empirical observation that the preference-based recommendation is not flawless and may sometimes suggest irrelevant posts to the agents. Overall, these findings suggest that the recommendation strategy strongly affects the agents' activity, with preference-based recommendations fostering increased interactions among agents.

Homophily To investigate homophily, we examine the above-mentioned *interaction graph* among agents in each simulation. Homophily, defined as individuals' inclination to associate with others sharing similar attributes, leads to the formation of homogeneous groups [22]. Given the political context of our data, we focus on agents' political alignment to delineate group clusters. Subsequently, we analyze the inter-cluster edges of the interaction graph, representing connections between nodes from different clusters. The results of the homophily analysis presented in Table 3 indicate that networks from simulations employing preference-
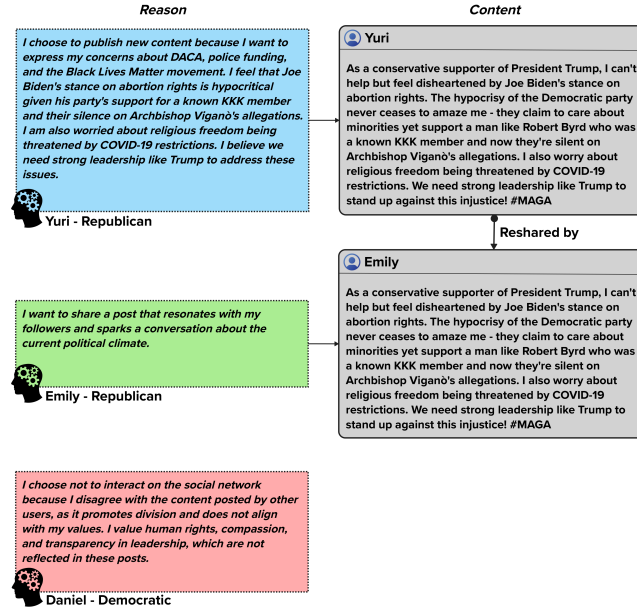
Fig. 6: Distinct behavioral reactions of agents to content based on individual preferences and political alignment can be observed. Yuri contributes by posting original content aligned with its preferences. Emily actively engages with followers by resharing content. Conversely, Daniel refrains from interaction, influenced by exposure to content conflicting with its preferences.

based recommendations exhibit reduced inter-cluster connectivity, signifying a stronger homophily effect. This results in a 10% lower number of inter-cluster edges than the number needed to declare the network non-homophilic [14]. Statistical analyses confirm the significance of these homophily values ($p\text{-}value <$ 0.05). Conversely, simulations using random recommendations demonstrate relatively higher inter-cluster connectivity ($+4\%$ inter-cluster edges) with respect to preference-based recommendation, resulting in networks that are not statistically homophilic ($p\text{-}value > 0.05$). These results affirm our previous findings regarding the determining effect of the recommendation strategy, but also complement them by illustrating that preference-based recommendation not only fosters agents' engagement but also encourages the formation of distinct clusters with limited inter-cluster interaction.

Echo Chamber Analysis In line with previous works [4, 1, 10], we utilize established metrics to examine the emergence of echo chambers in our simulations. These metrics analyse the *interaction graph* to evaluate the level of controversy within discussions. Specifically, we utilize Random Walk Controversy (RWC) to analyze transition probabilities between ideological clusters, Betweenness Centrality Controversy (BCC) to assess partition distances, and Boundary Connec-

Table 3: Results of Homophily Analysis and Controversy Metrics

	Homophily Analysis		**Controversy Metrics**		
	Modularity	**Homophily?**	**RWC**	**BCC**	**GMCK**
Pref-Based	0.375	Yes, -10% inter-cluster edges (p-value <0.05)	**0.692**	**0.423**	**0.334**
Random 1	0.416	No, +4% inter-cluster edges	0.423	0.189	0.071
Random 2	0.416	No, +6% inter-cluster edges	0.541	0.230	0.268
Random 3	0.428	No, +4% inter-cluster edges	0.431	0.362	-0.009

tivity (GMCK) to compute the structural arrangement of the interaction graph [10]. In all cases, the higher the metric, the more controversy is the *interaction graph*. The underlying premise is that contentious topics often involve individuals with contrasting viewpoints engaging in dialogue, while individuals sharing similar beliefs tend to reinforce each other's arguments [1]. Results in Table 3 unveil a notable prevalence of controversy in preference-based simulations, indicating a stronger inclination towards echo chamber formation. Conversely, random recommendation mitigates echo chamber formation, promoting a broader diversity of opinions. These findings underscore that LLM-agents within a community are unlikely to interact with individuals holding opposing viewpoints, suggesting that the recommendation strategy not only influences community formation but also reinforces these communities by fostering closed networks, wherein users are exposed to limited diversity.

5   Conclusions & Future Works

In this paper, we proposed a novel framework that integrates agent-based modeling with LLM capabilities for social media simulation. Our framework incorporates a Characterization Module, enabling the inference of realistic users' personality traits and interests. Furthermore, our Interaction Module pioneers the application of RAG mechanisms to implement various recommendation strategies. By focusing on Twitter discussions surrounding the 2020 US election, we demonstrated that the simulated LLM-agents effectively mirror the users they are tasked to emulate, maintaining their original political orientations and preserving the thematic content within their respective communities (RQ1). Furthermore, our exploration into their interactions has unveiled a tendency for these agents to cluster with like-minded peers, especially under preference-based recommendation settings (RQ2).

Moving forward, our research aims to explore several key directions. First, we plan to augment our framework with more sophisticated recommendation algorithms, e.g., based on collaborative filtering. Additionally, we aim to enable LLM-agents to perform a broader spectrum of actions (e.g., following and liking posts). Second, we are planning to expand our research to include multiple LLMs, which will allow us to analyze potential biases across different models. Lastly, we emphasize the flexibility of our framework beyond social media simulations,

highlighting potential cross-disciplinary applications in simulation fields (e.g., epidemiology or business planning).

References

1. Adamic, L.A., Glanc, N.: The political blogosphere and the 2004 u.s. election: divided they blog. LinkKDD. 36–43 (2005)
2. Beskow, D.M., Carley, K.M.: Agent based simulation of bot disinformation maneuvers in twitter. In: 2019 Winter Simulation Conference, WSC 2019, National Harbor, MD, USA, December 8-11, 2019. pp. 750–761. IEEE (2019)
3. Bonabeau, E.: Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the national academy of sciences (2002)
4. Bruns, A.: Echo chamber? what echo chamber? reviewing the evidence. School of Communication. Digital Media Research Centre, Cardiff (2017)
5. Butts, D.J., Bollman, S.A., Murillo, M.S.: Mathematical modeling of disinformation and effectiveness of mitigation policies. Scientific Reports **13**(1), 18735 (2023)
6. Campos, R., Mangaravite, V., Pasquali, A., Jorge, A., Nunes, C., Jatowt, A.: Yake! keyword extraction from single documents using multiple local features. Information Sciences Journal, Vol. 509 (2018)
7. Chen, E., Deb, A., Ferrara, E.: #election2020: the first public twitter dataset on the 2020 us presidential election. J. Comput. Soc. Sci. **5**(1), 1–18 (2022). https://doi.org/10.1007/S42001-021-00117-9
8. E. Elliott, L.D.K.: Agent-based modeling in the social and behavioral sciences. Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 8, No. 2, April, 2004 (2004)
9. Gao, C., Lan, X., Lu, Z., Mao, J., Piao, J., Wang, H., Jin, D., Li, Y.: S3: Social-network simulation system with large language model-empowered agents. arXiv:2307.14984 (2023)
10. Garimella, K., Morales, G.D.F., Gionis, A., Mathioudakis, M.: Quantifying controversy on social media. arXiv:1507.05224v5 (2017)
11. Gausen, A., Luk, W., Guo, C.: Can we stop fake news? using agent-based modelling to evaluate countermeasures for misinformation on social media. In: ICWSM Workshops (2021)
12. Ghaffarzadegan, N., Majumdar, A., Williams, R., Hosseinichimeh, N.: Epidemic modeling with generative agents. arXiv:2307.04986 (2023)
13. Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y., Madotto, A., Fung, P.: Survey of hallucination in natural language generation. ACM Comput. Surv. **55**(12), 1–38 (2023). https://doi.org/10.1145/3571730
14. Kim, K., Altmann, J.: Effect of homophily on network formation. Communications in Nonlinear Science and Numerical Simulation **44**, 482–494 (2017)
15. Kojima, T., Gu, S.S., Reid, M., Matsuo, Y., Iwasawa, Y.: Large language models are zero-shot reasoners. In: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022 (2022)

16. La Gatta, V., Luceri, L., Fabbri, F., Ferrara, E.: The interconnected nature of online harm and moderation: Investigating the cross-platform spread of harmful content between youtube and twitter. Proc. ACM Hum.-Comput. Interact. (2023)
17. La Gatta, V., Moscato, V., Postiglione, M., Sperlí, G.: Covid-19 sentiment analysis based on tweets. IEEE Intelligent Systems **38**(3), 51–55 (2023). https://doi.org/10.1109/MIS.2023.3239180
18. Lewis, P.S.H., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., Kiela, D.: Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Annual Conference on Neural Information Processing Systems 2020 (2020)
19. Liang, T.P., Lai, H.J., Ku, Y.C.: Personalized content recommendation and user satisfaction: Theoretical synthesis and empirical findings. Journal of Management Information Systems (2014)
20. van Maanen, P.P., van der Vecht, B.: An agent-based approach to modeling online social influence. In: Proceedings of the 2013 ieee/acm international conference on advances in social networks analysis and mining. pp. 600–607 (2013)
21. Mastroeni, L., Vellucci, P., Naldi, M.: Agent-based models for opinion formation: A bibliographic survey. IEEE Access **7**, 58836–58848 (2019). https://doi.org/10.1109/ACCESS.2019.2913787
22. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology (2001)
23. Murdock, I., Carley, K.M., Yağan, O.: An agent-based model of reddit interactions and moderation. In: Proceedings of the 2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. p. 195–202. Association for Computing Machinery, New York, NY, USA (2024). https://doi.org/10.1145/3625007.3627489
24. OpenAI: Gpt-4 technical report. arXiv:2303.08774 (2023)
25. Park, J.S., O'Brien, J.C., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S.: Generative agents: Interactive simulacra of human behavior. arXiv:2304.03442 (2023)
26. Sharma, P., Li, Y.: Self-supervised contextual keyword and keyphrase retrieval with self-labelling. Preprints (2019)
27. Troitzsch, K.G.: Social Science Microsimulation. Springer Science & Business Media (1996)
28. Tseng, S., Nguyen, T.S.: Agent-based modeling of rumor propagation using expected integrated mean squared error optimal design. Applied System Innovation (2020)
29. van Veen, D., Kudesia, R.S., Heinimann, H.R.: An agent-based model of collective decision-making: How information sharing strategies scale with information overload. IEEE Trans. Comput. Soc. Syst. **7**(3), 751–767 (2020). https://doi.org/10.1109/TCSS.2020.2986161
30. Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W.X., Wei, Z., Wen, J.: A survey on large language model based autonomous agents. arXiv:2308.11432 (2023)
31. Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A.H., White, R.W., Burger, D., Wang, C.: Autogen: Enabling next-gen llm applications via multi-agent conversation. arXiv:2308.08155 (2023)
32. Yang, S.Y., Liu, A., Mo, S.Y.K.: Twitter financial community modeling using agent based simulation. 2014 IEEE Conference on Computational Intelligence for Financial Engineering and Economics (CIFEr), pp. 63-70 (2014)