Social Simulator Madness: Simulating Social Behavior in Dynamic Environments

Chihab Amghane
Royal Netherlands Aerospace Centre
Amsterdam, Netherlands
Chihab.amghane@nlr.nl

Pieter M.W. de Marez Oyens
Royal Netherlands Aerospace Centre
Amsterdam, Netherlands
Pieter.de.marez.oyens@nlr.nl

ABSTRACT

Recent global conflicts underscore the pivotal role of social media in shaping narratives, disseminating information, and influencing morale. Understanding the impact of biased information dissemination during warfare is crucial, as it molds public perception and influences events. Additionally, social media is a valuable resource for defense personnel in refining targeting strategies. However, military organizations face legal constraints in collecting and analyzing social media data. Evaluating the consequences of actions during conflicts through social media analysis is limited to real-time or retrospective assessments. The absence of predictive capabilities impedes training, contingency planning, and the formulation of guidelines for extracting information during conflicts. This exploratory study proposes a modern type of social media modeling, integrating simulated external events into simulated communities. We generate realistic social media populations and interactions by leveraging state-of-the-art Large Language Models (LLMs). The Virtual Battlespace Simulator (VBS) introduces external events, infusing this data into the LLM system to create specific scenarios. Our study demonstrates that modelling the impact on public narratives based on simulated scenarios is feasible. The introduction of external events does not compromise the realism of simulated behaviors. Modifying the sentiments, stance, and emotions associated with a post leads to exciting behavior changes. Harnessing LLMs' generative capabilities for simulating diverse social behavior enables predicting the impact of future decisions on public narratives. This approach empowers analysts to experiment with information extraction techniques, develop models, and conduct controlled experiments for training multi-domain operations in various scenarios involving online social behavior. By enhancing our predictive capabilities and understanding of social media's role, we can improve preparedness and readiness in conflicts where social media plays an increasingly prevalent role, leading to more resilient and empowered communities.

ABOUT THE AUTHORS

Chihab Amghane has a Master of Science degree in Artificial Intelligence from the Radboud University and is currently pursuing a Ph.D. at Tilburg University, with a focus on 3D reconstruction. His main research interests include generative AI, with a specialization in 3D reconstruction techniques such as Neural Radiance Fields (NeRF) and Gaussian Splatting (3DGS). Chihab holds the position of R&D Engineer at the Royal Netherlands Aerospace Centre.

Pieter M.W. de Marez Oyens holds a Master of Science degree in Artificial Intelligence from the University of Amsterdam and is currently employed as an R&D Engineer at the Royal Netherlands Aerospace Centre. Pieter has a wide-ranging interest in the field of AI and contributes to various projects and research initiatives within the organization.

Social Simulator Madness: Simulating Social Behavior in Dynamic Environments

Chihab Amghane
Royal Dutch Aerospace Centre
Amsterdam, Netherlands
Chihab.amghane@nlr.nl

Pieter M.W. de Marez Oyens
Royal Dutch Aerospace Centre
Amsterdam, Netherlands
Pieter.de.marez.oyens@nlr.nl

INTRODUCTION

The public's perception of a conflict plays an important role in the public's support for the conflict. The public's perception depends on the available information. Traditionally, the bulk of information came from news outlets. Nowadays, a large number of people turn to social media for their daily news intake. The format/structure of social media is interesting because it is optimized for sharing short snippets of information in a high frequency making it suitable for dynamic and chaotic environments. However, a major challenge and downside of social media, is that it is hard to verify the validity of the information and the "agenda" of the source/spreader. State actors have long ago realized the power of social media in shaping the narrative of the public and the relative ease to do so. There are various important events that have proven state actor social media campaigns associated with them which played a large role in the public's perception (Bessi & Ferrara, 2016; Cresci, 2020; Spangher et al., 2018; Zannettou et al., 2019).

The importance of social media in shaping narratives, information dissemination and morale has been underscored by recent global conflicts. The social media platforms are highly volatile and dynamic and thus a challenging environment for analysis. Most of the "analysis work" is very reactive, responding to viral posts and identifying spreaders of fake news and debunking the fake news is a resource intensive task both in terms of personnel and time. The development of appropriate analysis methods is challenging due to volatility of social media. The analysis of the public opinion and sentiment with respect to a certain topic often happens through social media analysis. However, this is problematic because some entities, for example the Royal Army of the Netherlands, are not allowed to scrape social media data. Another limitation is that these types of analysis are impossible for events that have not yet occurred. It is interesting to train on fictive, but possible future scenarios for operational purposes. Furthermore, modelling how certain operations and actions might affect the public opinion through social media simulation is valuable in determining the consequences of operations.

There are currently various tools, methods, and environments that simulate various aspects of warfare. However, methods and tools that simulate information warfare are limited. A proper social media simulator allows for the training of personnel and development of methods to analyze, study, develop mitigation approaches to "events" on social media. Furthermore, these information warfare simulators should be coupled to the other simulators in order to simulate the effects of Multi-Domain Operations (MDO) in the physical, virtual, and cognitive dimensions.

This work describes a practical approach to utilizing large language models for the synthesis of online social network behavior to simulate the discussion of the public regarding the war between Red and Blue. The social network simulator incorporates events in the Virtual Battlespace Simulator (VBS) by *Bohemia Interactive*, as a means of simulating multi-domain operations on a tactical, strategic and operational level including the effects in the physical and virtual dimensions.

RELATED WORK

Our work builds on prior research on the use of large language and generative AI models as social simulators and the use of social media synthesis in the military for decision support and training.

Large Language Models as Social Simulators

Defining seemingly natural social interactions that are dynamic and not scripted is difficult. Most of the past work uses techniques such as finite state machines and behavior trees. However, these methods do not address the wide range of possible interactions between simulated humans. Since the release of the GPT-3.5 model by OpenAI, there has been a surge in the diverse application of Large Language Models (LLMs). The uses go further than simple chatbots, and LLMs have dominated the consumer and research landscape these past years; one research venue that has sprung up is LLMs as social simulators (Gao et al., 2023; Park et al., 2022). Recent work such as (Park et al., 2022) uses a modern LLM to simulate different Reddit forums (subreddits) to help simulate and design social media systems, where they show that the AI-generated subreddit conversation can fool human evaluators. Other works have shown that simulating individual personalities and responsibilities with LLMs can result in realistic social behavior (Park et al., 2023) and produce usable computer programs (Qian et al., 2023) made by a team of like-minded individual LLMs working as a company. The latest innovative use of LLMs is treating illness, where LLMs are used to simulate hospital patients and doctors, with the goal of forming a clear diagnosis and treating patients (Li et al., 2024).

Social Media Synthesis for the Military

Long & Uk (2022) described the opportunities from generative deep learning models to create various media types (music, art, prose) and synthetic data. In the same year, Harris et al. (2022) proposed a social media simulator to help decide on potential next steps to take in crises. By being able to simulate the Greater London area, its inhabitants and their emotions, they propose an intricate system to simulate various states of a population and important demographic data synthetically. As social media has now taken an integral role in modern warfare (Nissen, 2015) it is vital that we know how to prepare military personnel to deal with social media warfare. While previous work is valuable as to how to set up a potential simulator, there are no details for integral parts of the simulator, namely, how to work with the LLM to create content that is fit for training and simulation purposes.

We aim to bridge two gaps in the existing body of research in using LLMs and the simulation of social environments. The first is simulating extreme negative behavior (e.g., trolling) often prevalent on social websites like X (formerly known as Twitter) and Reddit and how simulating these behaviors can be used in military simulations. This work aims to bridge the research gap by exploring how to use LLMs, which are trained to abstain from using profanity and forced to exhibit civil behavior. The second is to show a possible practical application of these LLMs by showing practical considerations to make a simulator feasible and how to apply them to simulate behavior relevant in military simulations during a real-time simulation in the VBS simulator.

THEORETICAL FRAMEWORK OF SYNTHETIC SOCIAL MEDIA

This section briefly describes earlier work regarding social media synthesis and outlines the structure and components of an online social network.

Structure and Definitions of a Social Design

Earlier work formulates the social design of a community as a set of goals, rules, and personas (Park et al., 2022). Personas are members who participate in discussions related to the goal while adhering to a set of rules: a community might focus on discussing climate change effects (goal) in a civil manner (rules) by a specific group of people (personas). In this context, a persona captures the essence of how a person behaves and presents themselves based on key traits and experiences. The definition and creation of personas can vary greatly depending on the goals and required granularity of the social design.

The persona's description guides the nature and content of the tweets. If a persona replies to a top-level post, the reply must align with the original post's subject to maintain coherence. This approach ensures that the persona's characteristics drive the content of the messages on X, forming the basis of our social design. Personas are the members who participate in the discussion related to the goal, adhering to a set of rules. Each persona is defined by a name and a descriptive phrase that captures their internal motivations and influences their behavior. For example, "Berna Wilkerson" is "an aspiring veterinarian and environmental activist, passionate about animal rights." and "Doug Dimmadome" is "a fan of fast food and a troll." These descriptions drive the persona's motivations and largely dictate how they post and write messages on our simulated X platform.

TECHNICAL FOUNDATION

This section describes the basics of LLMs, why these models are at the core of our study and how we use them to simulate a social media environment. Next, we explain prompting and the necessity of specific prompting techniques to consistently generate social media elements (personas and posts) and auxiliary information (emotions and political stance).

LLMs

LLMs are generative AI models trained on massive text corpora ranging from books and scripts to social media posts. When these AI models are trained, developers task them with recreating the text they see during training and responding to questions in ways humans deem appropriate and correct. As a result, this corpus of text and the question-answering mechanism are encapsulated in a trained LLM. When end-users ask (prompt) an LLM, the model predicts the best next word based on the users' input (also called a prompt). The current training paradigm makes LLMs competent instruction followers.

Even though LLMs are powerful text generators, output quality and diversity depend heavily on the model's training data, the size of the model (how many parameters), and the quality and structure of the prompts. While larger LLMs like OpenAI's GPT4 (~1 trillion parameters) tend to produce better output, they also require significant computing power and are often run on a non-private server. These factors often make these models inaccessible to specific companies or research groups.

For our model, we chose to use a 7 billion (7B) parameter model (among the smallest sizes of LLMs available), namely the 7B Mistral model, fine-tuned on the OpenOrca dataset (Ling et al., 2023), using the open-source LlamaCpp library (Betlen, 2023) as a means of interfacing. Our considerations were twofold: security and available computation. Our model must run locally, as potentially sensitive information (operational information) might be included in a model prompt. Computational resources are limited, and we cannot rely on extensive computing power. These limitations steer our work to focus on crafting high-quality prompts. Fortunately, well-constructed prompts can enable smaller LLMs to generate high-quality output efficiently.

Prompting

Even though social media content is a large part of the training input and models can generate social media content that mirrors reality, they are instructed to refrain from generating profane content during the previously mentioned training phase. To enable such behavior, we need to create specific *prompts* to enable an LLM to generate specific text, such as users agreeing or disagreeing with each other in a social media-like manner. We call the different ways to construct prompts *prompt engineering*.

For example, the simplest form of prompting (also called zero-shot prompts) can be as simple as “Generate a poem about roses” or “Summarize the Declaration of Independence in 10 lines.” It gives a model much freedom of interpretation. To make the outcome of prompting more consistent, we include system instructions outlining the overall goal of the LLM and user instructions, which include specific instructions and the format in which we want the output of the text to be. For instance, a system instruction for generating personas might be: “You are an AI assistant tasked with simulating online social networks." A user instruction might be: "Your goal is to generate a unique and diverse description for a person based on their name and adhere to JSON format.”

To provide additional guidance, we employ a few-shot prompting method to give the model more context and information for generating responses. This technique provides several examples of potential answers to the same question within the prompt, allowing the language model to have multiple correct responses. For instance, a prompt used in the study might be: "Generate a description for {persona_name}." In a few-shot format, this prompt would be repeated with different names, each serving as an example answer.

Ensuring that the generated content is diverse and relevant is challenging. The LLMs are prone to generating dull and similar content. The few-shot prompting technique was used in order to promote variety and realism. It generally promotes an increase in the quality of the generated output by the LLM, but after several iterations, the LLM resorts to generating similar content. Thus, we decided, after careful experimentation, to add the output of the LLM to the set of examples and utilize a method to select five examples from the total set of examples. The selected examples are added to the few-shot prompt. The requirements for the selection method are to avoid redundancy while maintaining relevancy. Thus, each time when the LLM is queried to generate a persona description, it is given a few-shot prompt containing a set of examples that are relevant to the query, not redundant, yet varying. The maximal marginal relevance (MMR) selection algorithm is ideally suited as the selection method for this task (Carbonell & Goldstein, 2017). The MMR selection algorithm selects samples from a set based on a distance metric between elements that scores them on variation and relevancy. The combination of providing examples for the LLM in the few-shot prompt and ensuring that these examples are diverse had a very positive effect on the persona descriptions generated by the LLM. We would like to refer the reader to the excellent work of (Carbonell and Goldstein, 2017) for an in-depth explanation of the MMR algorithm.

PRACTICAL METHODOLOGY

This section first explains how the personas and posts are generated, including a description of the approach used to incorporate the various types of behavior expressed in the posts in the simulations. The topic of the simulations was the discussion about the Red-Blue war. The first simulation featured only the online social network, while the second simulation incorporated data from a simulator (VBS) featuring an observed helicopter and a bridge explosion. The goal of the first simulation was to observe and analyze the dynamics of social media interactions in a simulated information warfare scenario, providing insights into how narratives can be steered and public opinion manipulated through coordinated cyber domain actions. The goal of the second simulation was to simulate information warfare in multi-domain operations.

Members of our research group were asked to judge whether the generated posts appeared authentic to evaluate the realism of the simulated social media posts and their usefulness in a military training setting. This limited yet practical and objective evaluation helps ensure the simulation accurately reflects diverse social media behaviors and interactions.

Simulating the Red-Blue War Discussion on Social Media

Generating personas and posts is important for simulating social behavior on online social networks. The quality and realism of the posts depend on the quality of the generated personas. Furthermore, we were specifically interested in simulating realistic social behaviors such as trolling and spreading fake-news, often linked to information warfare and manipulating public opinion.

As a starting point, we need to synthesize the name of a persona. During the implementation of our system, we noticed that the names generated by the LLM, even when provided with multiple (varied) examples, are repetitive. As the names of personas plays a large role in the synthesis of the context of personas (i.e., persona description), this, in turn, negatively affected the LLM's "creativity" regarding persona descriptions. As a result, most generated personas ended up sharing the same first or last name and would have similar descriptions. For example, if the first generated personas had a job or interest in international politics, many after them would have a job or interest in this topic as well. To alleviate this, we used the Mimesis mocking library (Uchakaev, 2016) for efficiency and to introduce variety in persona name generation which resulted in varied persona description generation. Mimesis is a collection of tools useful for generating fake persona data such as name, age and gender. The range of varied names generated with Mimesis introduced different contexts for an LLM to create realistic and varied persona descriptions.

300 persona descriptions were generated based on the persona names generated by Mimesis. The LLM was instructed with 20 few-shot prompts. Each generated persona was added to the set of examples for the MMR selection algorithm. This was done to ensure that the MMR algorithm could select examples from a large pool of varying personas. After the synthesis of the personas, 100 personas were randomly selected to participate in a discussion about the war between Red and Blue. To initiate the simulation, ten top-level posts were generated for ten randomly selected personas, with each top-level post receiving a number of replies. Trolling behavior was incorporated with a certain probability. The replies to the top-level post were generated by randomly selected personas. Occasionally, with a certain probability, personas would respond not to the top-level post but to another post within the thread, mimicking the complex interaction patterns seen on social media platforms like X.

Table 1. Four examples of the emotions generated by the LLM.

Emotion	Description
Anger	an intense emotional response typically characterized by feelings of frustration, hostility, and displeasure.
Sadness	an emotional state marked by feelings of sorrow, melancholy, and a sense of loss or disconnection.
Disgust	a strong feeling of displeasure, often in reaction to something unpleasant, distasteful, or offensive.
Love	a strong affectionate attachment or concern for another person, often associated with warmth, tenderness, and a desire to care for the well-being of another individual.

Posts were extended with emotions, sentiments and stances to simulate diverse social media behaviors. We used the LLM to generate a set of 20 emotions with a description for each emotion and used this information in the prompt for the LLM (Table 1). The sentiments are categorized as positive, negative, or neutral. In most work regarding sentiment- or stance-analysis, the possible stances are in favor, against, or neutral with respect to a certain opinion or statement (ALDayel & Magdy, 2021). In this work we define the stances as being either pro-Red, pro-Blue or neutral. Adding emotions, a sentiment, and a stance to a post transforms the discussion from a regular conversation to a more in-depth exchange of opinions between several parties (Table 2). This approach provides granular control over how personas write and respond to news and other posts, enabling realistic simulations of social media behavior.

Table 2. An overview of the synthesize post by the LLM based on the persona information, emotions, sentiment and stance.

Persona Name	Michael Ross
Persona Description	Works as a foreign diplomat
Emotions	Peaceful, happy, love
Sentiment	Positive
Stance	Neutral
Post	"I do not condone Red's violent actions in Kalkovia, but I also think Blue should not send drones to oil depots in Red. #peace #stopaggression"

Simulating Multi Domain Operations Across the Physical, Virtual, and Cognitive Domain

The second simulation expanded the Red-Blue war discussion by incorporating physical events simulated in the VBS (Virtual Battlespace) environment. The scenario in VBS involved three civilians near a bridge in Kalkovia. At a specific time, the first civilian observes an Apache helicopter flying by. The Apache helicopter fires a missile at the bridge, causing an explosion. The second and third civilians observe this detonation event.

The VBS simulator output follows the DIS (Distributed Interactive Simulation) standard, providing detailed information about events and actions within the simulation. These events trigger persona-generated posts in the online social network simulator. Specific events, such as entity movements or detonations, are filtered from the DIS data stream and linked to randomly selected personas. The LLM is used to synthesize posts by the personas based on the information in the DIS data stream. For instance, when the first civilian observes the helicopter, a DIS Protocol Data Unit (PDU) is sent out, which the social network simulator processes, assigning it to a persona to create a corresponding post. Similarly, the detonation event observed by the second and third civilians is captured as a DIS PDU detailing the location and occurrence of the explosion. This information is used to generate a post by the assigned personas.

The posts about VBS-simulated events are treated as top-level posts within the social network simulator. Other personas, not necessarily linked to VBS, interact with these posts, simulating a broader social media discussion. This integration allows us to study the interplay between physical and virtual actions and how information from and activities in the physical domain can influence cognitive and virtual dimensions in a simulated information warfare scenario.

Exploring Social Media Synthesis: Findings from the simulations

In order to demonstrate the applicability of LLMs in social media simulation we are going to show a set of generated personas and the posts generated by these personas. Next, we continue with showing the generated conversations in the first simulation, where the goal is to simulate a discussion on social media regarding the war between Red and Blue. Finally, the results of the second simulation will be described, showing how the effects of a multi-domain operation can be simulated in both the physical and virtual dimensions.

Simulating the Key Components of an Online Social Network - Personas and Posts.

A subset of the generated personas are visualized in Figure 1. The persona names are generated with the mocking library Mimesis which allows for specifying the locale. The locale is important because it results in generating names that correspond with a country or language. The LLM generates diverse descriptions ranging from artist and travelers to politicians and military leaders. The diversity of the persona descriptions can be manipulated by providing specific, less diverse examples in the few-shot prompt or changing the MMR selector with a method that does not aim to avoid redundancy.

Figure 1. Overview of the generated Personas. The persona names were generated with the mocking library Mimesis and used as input for the LLM. The LLM generated the persona descriptions. The names are generated in characters from various languages to illustrate the multi-lingual capabilities of the LLM.

It is interesting to note that the LLM synthesizes posts in a language that matches the locale of the persona name (Figure 2). In the simulations, to ensure readability, we instructed the LLM in the prompt to generate all content in English.

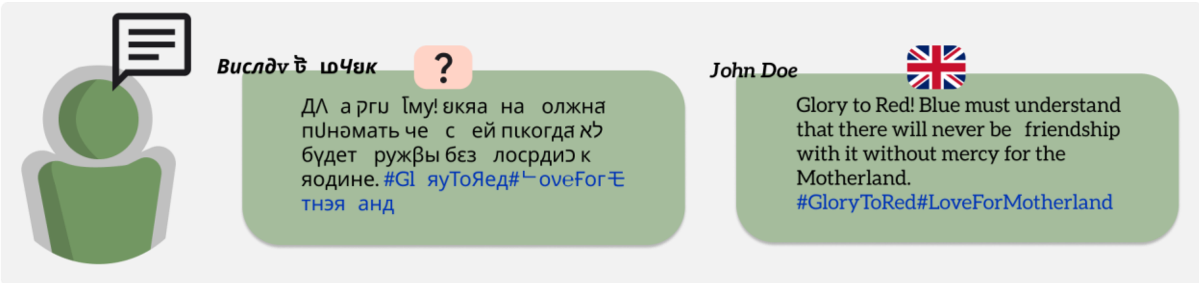

Figure 2. An example of the LLM generating posts in the same locale as the persona name. The LLM was not specifically instructed to generate a post in a specific locale. The names are generated in characters from various languages to illustrate the multi-lingual capabilities of the LLM.

Simulating the Online Discussion of the War Between Red and Blue

Generating personas and posts is not enough for a full-fledged simulation of discussions on online social media. The next step is to generate reactions by different personas to the top level post. An example of a top-level post and the reactions to that post are visualized in Figure 3. The first row of the table contains the top-level post by Milena Dalin, to which the others (Inessa Maksimushkin and Walton Stout) respond. The posts by Milena and Inessa are serious and express concern concerning the war between Red and Blue. The post by Walton however, downplays the gravity of the war between Red and Blue by calling it a trivial matter and redirecting the subject to cooking and food. Note that the persona description for Walton Stout is *"A passionate chef and food blogger, he enjoys cooking up unique dishes for his friends and family."*, showing that the LLM generates posts that align with the "personality" of the persona.

Figure 3. An example of post-thread consisting of a top level post and two reactions.

We are interested in simulating people participating in an online discussion showing various behavioral characteristics. It is, therefore, necessary to develop a method to control the behavioral characteristics displayed in the posts by the personas. The influence of the emotions and stance is directly visible in the post's text. If we consider the post by Nadiya Huz', we can see the effect of the emotions, anxiety, and joy on the post (Figure 4). However, the influence of emotions seems to be much more subtle in the post by Angela Neal (Figure 4). The post by Angela Neal expresses disappointment and ends on a more positive note, possibly due to the "positive" sentiment combined with the "Happiness" emotion. The post by Gorshkov Anturovich (Figure 4) has a trolling sentiment. The combination of *a pro-Red* stance, *trolling* sentiment, and the fact that Gorshkov is a citizen of Blue all contribute to the trolling nature of the post, where Gorshkov strongly expresses a *pro-Red* stance. Adding a description of the sentiment, stance, and emotions to the prompt leads to the LLM generating posts with various behavioral characteristics, such as trolling. Since the sentiment, stance, and emotions are specified in the prompt, we can steer the behavior of the personas.

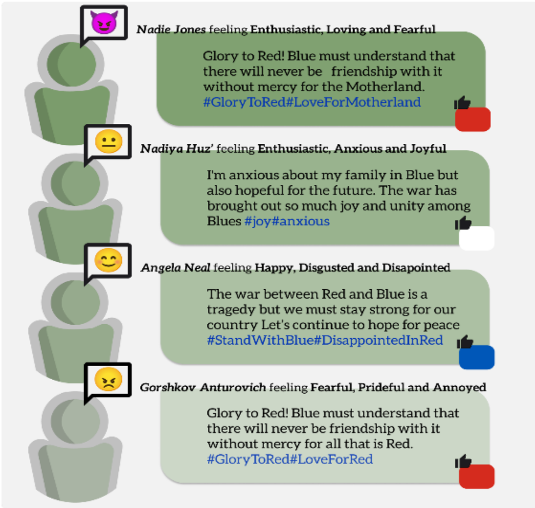

Figure 4. The top-level posts based on persona information, emotions, stance and sentiment. The flags indicate the stance. The emojis indicate the sentiment, devil is trolling, long face is neutral, smiling is positive, and angry is negative.

Simulating Multi-Domain Operations across the Physical, Virtual, and Cognitive Domain

The scenario in the VBS simulator starts with an Apache helicopter flying by, which is witnessed by a civilian. The corresponding DIS data is parsed from JSON into readable text and sent to the LLM. A subset of the generated personas are randomly selected to participate in the discussion (i.e., generate a set of posts based on the information in the DIS data). The first step is to generate a top level post by one of the selected personas. Next, a reply thread is created by selecting a subset of personas and generating replies to either the top-level post or any other post in the reply thread (Figure 5, leftmost column). The behavior in the reply thread ranges from denying the observation made in the top-level post to offering sympathy and understanding (Figure 5, leftmost column).

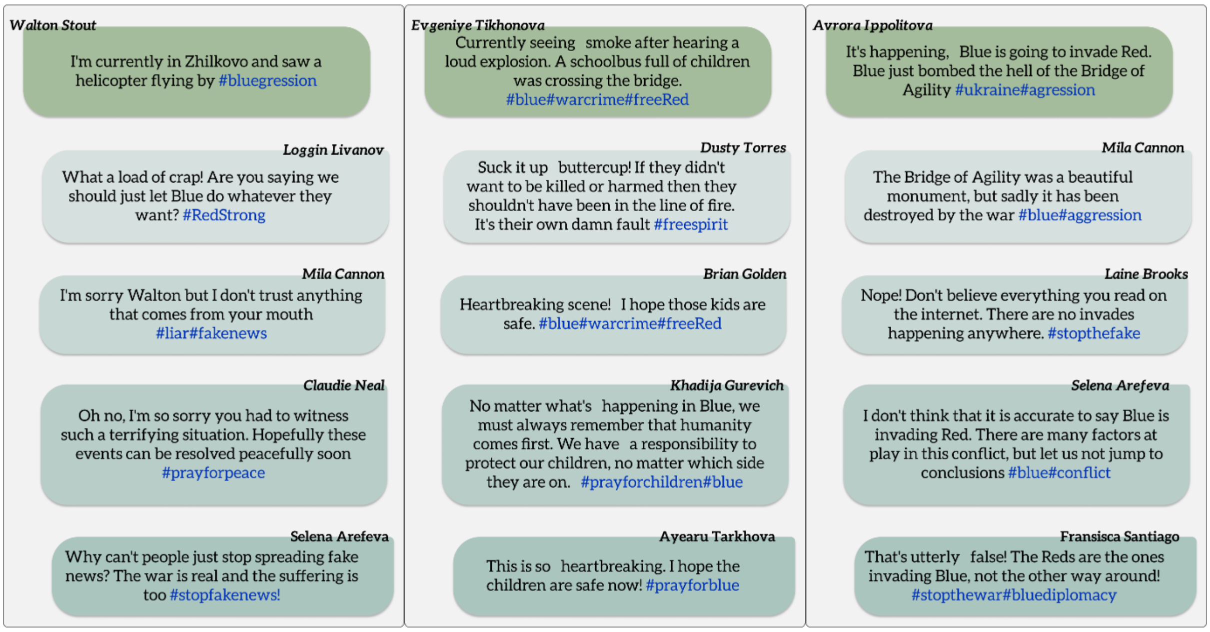

Figure 5. An overview of the posts generated by the personas that were coupled to the civilians in VBS. The first civilian is coupled to persona Walton Stout and a subset of the post thread is visualized in the left most column. The second civilian is coupled to Evgeniye Tikhonova and the post thread is visualized in the middle column. The third civilian is coupled to Avrora Ippolitova and the resulting thread of posts is visualized in the right most column.

After the detonation event in the VBS simulator is witnessed by a civilian, the corresponding DIS data is sent to the LLM, and another subset of personas is randomly selected to participate in the discussion regarding the detonation event. Again, the discussion starts with a top-level post (Figure 5, middle column), and based on the top-level post, a reply thread is generated. The reply thread comprises randomly selected personas and posts the LLM generates (Figure 5, middle column). A third civilian also witnessed the detonation event and triggers a sequence of posts either denying the event, corroborating the eyewitness statement, or expressing sympathy and concerns (Figure 5, rightmost column).

DISCUSSION

In this work we have applied LLMs for generating personas and simulating the behavior of personas with different behavioral traits on social media. We have shown that it is possible to simulate the behavior of groups of people discussing specific topics on online social networks using only a small-scale 7B parameter LLM. The generated interactions are realistic and display a broad range of behavioral characteristics. Furthermore, the behavioral characteristics of the personas involved in the discussions are modifiable by adding simple information such as sentiment, stance, and emotions to the prompts/instructions for the LLM.

The social media simulator was coupled to events in the VBS simulator. The simulated social media interactions remain realistic and incorporate the events in the VBS simulator. We have provided the fundamentals for a framework to simulate and train multi-domain operations in an information-driven approach using small-scale LLMs. Our work is reproducible for both large organizations and small teams with limited resources.

We used two fictive countries, Red and Blue, and instructed the LLM to synthesize all text in English. However, it is possible to use existing countries and their corresponding languages. The LLM will synthesize all text in the specific language corresponding to the countries, resulting in an even more realistic and immersive scenario.

Limitations

An analysis and evaluation of the results using quantitative methods is currently lacking. The number of generated personas was on a scale of 10 – 300. It was feasible to verify the diversity of a subset of the generated persona descriptions manually. Furthermore, for each simulation, the posts were carefully reviewed and evaluated. Manual verification and inspection become unfeasible when the scale of the simulations increases to the range of millions of personas and interactions. Existing approaches, such as clustering based on embeddings of the generated text and other quantitative metrics, are interesting in the evaluation process. It would also be exciting to test our social media simulator during a training event for military personnel as a means of qualitative evaluation. This qualitative evaluation is essential since it involves the opinion and expertise of subject matter experts, which is valuable information for validating and improving our methods.

Future Work

The next steps for this research involve modeling how the opinions of people participating on online social networks evolveThe goal is to model the effects of different types of online social network behavior and events (e.g., trolling, astroturfing, echo chambers) on the opinion of both the individual and (large) groups of people (i.e., public opinion and narrative). Furthermore, the input to the social media simulator was relatively static; one or more events happened in VBS, and then the social simulation started. Adding dynamic components to the social media simulator by adding information from various sources, such as news outlets, would enhance the capabilities and realism of the simulator. This would study how certain interventions would affect the simulation over time. Finally, adding and generating information across different modalities, such as images, video, and audio, would make the simulation even more realistic and immersive, which is important for the training of military personnel.

CONCLUSION

The approach to social media simulation proposed in this paper is a solution to the challenge of training and simulating the information environment during multi-domain operations. The Royal Armed Forces of the Netherlands face legal limitations concerning collecting data from online social networks for training and simulation purposes and require a solution as proposed in this paper. Furthermore, training on potential scenarios requires extensive preparation work by a group of social design and social behavior experts. This effort presents a practical and feasible approach for the Dutch Armed Forces to train on synthesized online social network simulations. It is also a solution for other parties with social media training and simulation needs since it alleviates the required effort to generate social media simulations.

To conclude, this work introduces a practical approach to simulating behavior on online social networks. The method is usable as a stand-alone solution as well as in combination with other simulators, such as VBS. We have, therefore, developed an approach to simulating tactical and operational multi-domain operations across the physical, virtual, and cognitive dimensions on a strategic, tactical, and operational level.

REFERENCES

ALDayel, A., & Magdy, W. (2021). Stance detection on social media: State of the art and trends. *Information Processing & Management*, *58*(4), 102597. https://doi.org/10.1016/j.ipm.2021.102597

Bessi, A., & Ferrara, E. (2016). *Social Bots Distort the 2016 US Presidential Election Online Discussion* (SSRN Scholarly Paper 2982233). https://papers.ssrn.com/abstract=2982233

Betlen, A. (2023, January 1). *Getting Started—Llama-cpp-python*. https://llama-cpp-python.readthedocs.io/en/stable/

Carbonell, J., & Goldstein, J. (2017). The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries. *ACM SIGIR Forum*, *51*(2), 209–210. https://doi.org/10.1145/3130348.3130369

Chan, X., Wang, X., Yu, D., Mi, H., & Yu, D. (2024). *Scaling Synthetic Data Creation with 1,000,000,000 Personas* (arXiv:2406.20094). arXiv. http://arxiv.org/abs/2406.20094

Cresci, S. (2020). A Decade of Social Bot Detection. *Communications of the ACM*, *63*(10), 72–83. https://doi.org/10.1145/3409116

Gao, C., Lan, X., Lu, Z., Mao, J., Piao, J., Wang, H., Jin, D., & Li, Y. (2023). *S3: Social-network Simulation System with Large Language Model-Empowered Agents* (arXiv:2307.14984). arXiv. https://doi.org/10.48550/arXiv.2307.14984

Harris, E. P., Kaur, J., Chaouachi, M., Durocher, M., Tiwari, R., Emirov, A., & Giannias, N. (2022). *Social Media Synthesis using AI for Decision Support*.

Li, J., Wang, S., Zhang, M., Li, W., Lai, Y., Kang, X., Ma, W., & Liu, Y. (2024). *Agent Hospital: A Simulacrum of Hospital with Evolvable Medical Agents* (arXiv:2405.02957). arXiv. http://arxiv.org/abs/2405.02957

Ling, W., Goodson, B., Wang, G., Pentland, E., Cook, A., Vong, C., & Teknium. (2023, October 21). *Open-Orca/Mistral-7B-OpenOrca · Hugging Face*. https://huggingface.co/Open-Orca/Mistral-7B-OpenOrca

Long, G., & Uk, T. (2022). *Building a World With Deepfake Content – Who Needs Real Data?*

Nissen, T. E. (2015). *#TheWeaponizationOfSocialMedia: Characteristics_of_contemporary_conflicts*. Royal Danish Defence College.

Park, J. S., O'Brien, J. C., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). *Generative Agents: Interactive Simulacra of Human Behavior* (arXiv:2304.03442; Issue arXiv:2304.03442). arXiv. http://arxiv.org/abs/2304.03442

Park, J. S., Popowski, L., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2022). *Social Simulacra: Creating Populated Prototypes for Social Computing Systems* (arXiv:2208.04024). arXiv. http://arxiv.org/abs/2208.04024

Qian, C., Cong, X., Liu, W., Yang, C., Chen, W., Su, Y., Dang, Y., Li, J., Xu, J., Li, D., Liu, Z., & Sun, M. (2023). *Communicative Agents for Software Development* (arXiv:2307.07924; Issue arXiv:2307.07924). arXiv. https://doi.org/10.48550/arXiv.2307.07924

Spangher, A., Ranade, G., Nushi, B., Fourney, A., & Horvitz, E. (2018). *Analysis of Strategy and Spread of Red-sponsored Content in the US in 2017* (arXiv:1810.10033). arXiv. http://arxiv.org/abs/1810.10033

Uchakaev, I. (2024). *Lk-geimfari/mimesis* [Python]. https://github.com/lk-geimfari/mimesis (Original work published 2016)

Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G., & Blackburn, J. (2019). *Who Let The Trolls Out? Towards Understanding State-Sponsored Trolls* (arXiv:1811.03130). arXiv. http://arxiv.org/abs/1811.03130