# YouTube's Silent Echo

**Team Members:** Simon Li, Vincent Tran, Isaac Ibidun
**GitHub Repo:** https://github.com/mayowaibi/EECS4461-Project/

## §1. Phenomenon Overview

### 1.1 Phenomenon Statement

Echo chambers are a common issue on social media platforms like YouTube, where recommendation algorithms and social bots reinforce people's existing beliefs while also filtering out different perspectives. This process, as described in the Brookings Report (Brown et al., 2023), creates mild ideological echo chambers, which means users gradually see more content that aligns with their views, and opposing viewpoints become less visible over time. One of the main drivers of this effect is YouTube's recommendation algorithm, which prioritizes content based on engagement. Social bots take advantage of this by boosting certain videos (using likes, comments, and shares) to make content seem more popular than it actually is. This tricks the algorithm into recommending the same kinds of videos repeatedly, making it harder for users to come across different viewpoints. Over time, this reinforces selective exposure, pushing people into tighter ideological bubbles (Brown et al., 2023). Our project simulates how social bots and YouTube's recommendation system interact to reinforce echo chambers. Using agent-based modelling, we will analyze how bots, algorithms, and users contribute to the formation of ideological clusters, which is similar to what happens in real life.

### 1.2 Problem Statement

It is important to understand how AI-driven recommendations and bots influence user exposure because social media shapes public opinion, contributes to polarization, and spreads misinformation. Research by Brown et al. (2023) found that while YouTube does not consistently push users into extreme rabbit holes, it creates mild ideological echo chambers where people are subtly guided/pushed into narrower content bubbles over time. This is especially concerning in areas like political content and misinformation, where engagement-driven algorithms can reinforce biases rather than challenge them. The issue is worsened by social bots, which manipulate engagement metrics by acting as fake users who inflate content popularity. Since YouTube's algorithm prioritizes engagement, these bots increase content visibility in an artificial way, which deepens ideological isolation. However, we should also take into account the existence of counterforces such as algorithmic interventions that promote content diversity. These factors introduce some variety in content exposure, but effectiveness in terms of disrupting echo chambers remains uncertain.

Our Project looks to answer:

- How does YouTube's recommendation algorithm and social bots influence what users see?
- To what extent do AI-driven interactions reinforce ideological bubbles?

- Are there ways to reduce ideological reinforcement and improve content diversity?

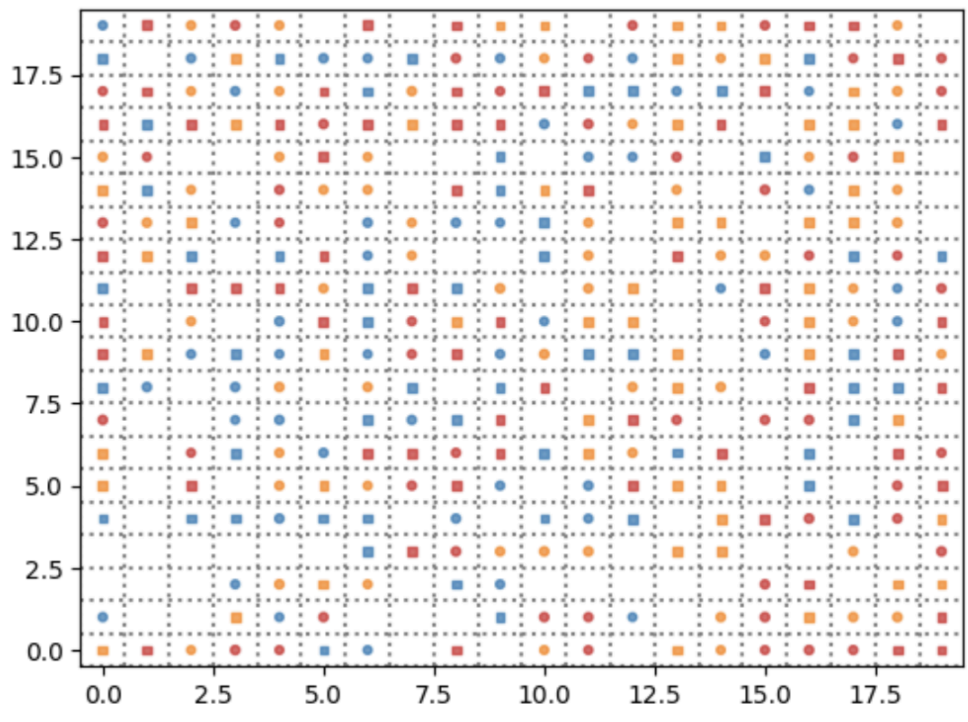
## 1.3 Why ABM?

Agent-Based Modelling is used because it allows us to simulate individual user behaviours and observe how small interactions can create large-scale echo chamber effects. Our model will simulate user engagement, model social bots that manipulate engagement metrics to boost content, and observe how ideological clusters form based on user interactions or recommendation patterns. We're using Mesa because it allows us to simulate network-based ecosystems, track agent interactions, and generate visual outputs that help us illustrate ideological segregation.

## 1.4 Phenomenon Illustration

Initial State:

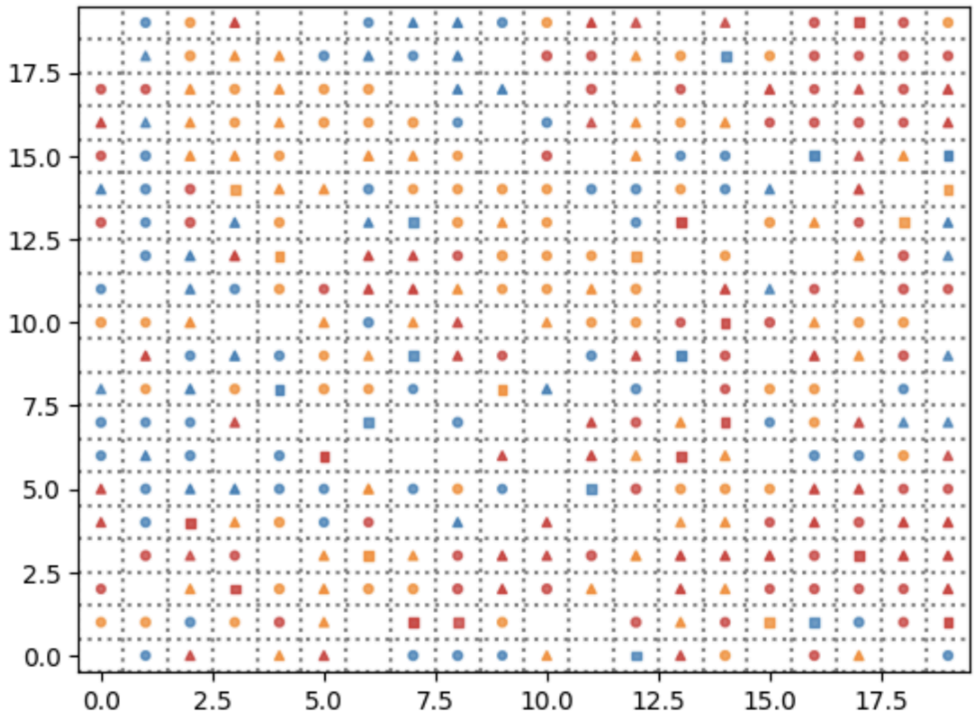
- The initial state of the simulation displays a randomized distribution of human users (circles) and social bots (squares) on a grid, with each agent assigned a colour for content preference.
- Similar to YouTube, human users start with no predefined algorithm and are exposed to a diverse range of content from various creators.




Middle State:

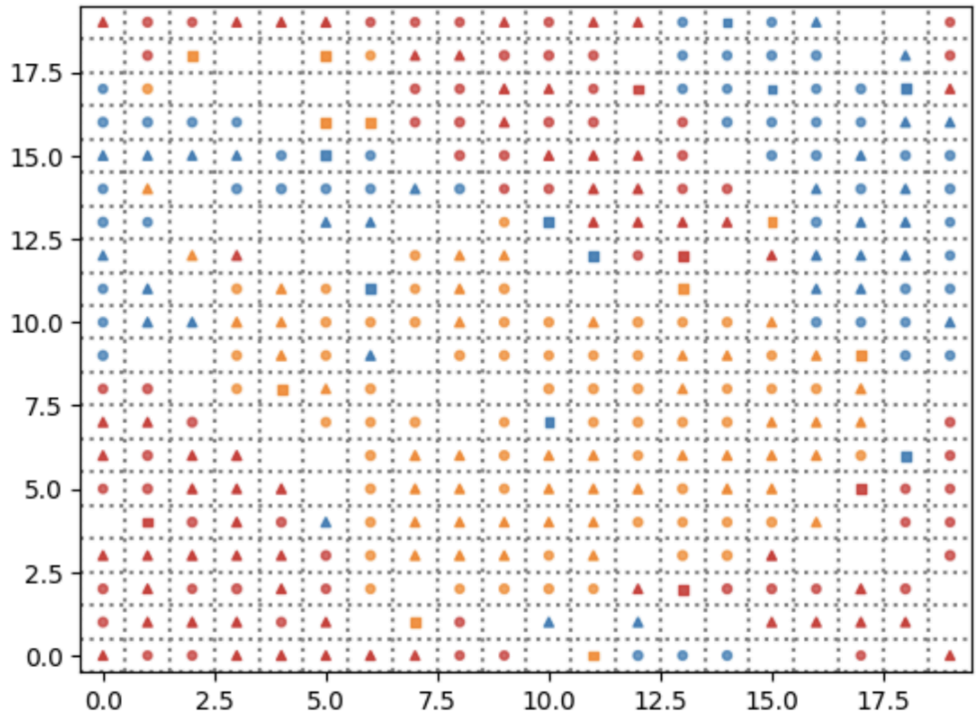
- As the simulation progresses, disctinct clusters begin to form, with human users becoming increasingly surrounded by others who share similar content preferences. Social bots actively engage with and amplify content, reinforcing echo chambers. As these bots become part of an echo chamber, they evolve from squares to triangles, signifying their increased influence in shaping the content ecosystem.

- Similar to YouTube, the recommendation algorithm begins adapting to user behaviour. As users engage more with specific content types, the algorithm prioritizes similar content.




Final State:

- By the final stage of the simulation, clear and isolated echo chambers have fully emerged, with most users interacting with content that aligns with their initial preferences. The influence of social bots within the echo chambers is evident as many have transformed into triangles.
- Similar to YouTube, the recommendation system at this stage has fully adapted to user preferences and they now see content almost exclusively from their preference.




## §2. Simulation Design & Implementation

### 2.1 System Overview

Our simulation models the formation of echo chambers on YouTube by representing human users, social bots, and YouTube's recommendation algorithm as agents in an agent-based modelling framework using the Mesa library. The system is designed to explore how algorithmic recommendations and bot interactions contribute to ideological clustering and selective exposure. The model is inspired by Schelling's Segregation Model, where agents move toward like-minded neighbours. Similarly, in our simulation, users are gradually exposed to content aligning with their preferences, which reinforces mild ideological echo chambers. Social bots further manipulate this system by amplifying engagement which influences content visibility. This reflects real-world algorithmic behaviour on platforms like YouTube, where engagement-driven recommendations can narrow people's exposure to diverse viewpoints.

Our system implements an interaction model where agents possess multiple attributes that influence their behaviour and impact on the environment. Human agents and social bots are distinguished not only by their type (0 for humans and 1 for bots) but also by their content preferences (represented as distinct categories). Each agent then maintains dynamic metrics, such as base and currency homophily levels, engagement rates and echo chamber strength (base metrics can be changed by the user before the start of the simulation). Social bots feature an additional complexity through their cluster formation mechanics, where their amplification power (capped at 3.0) increases based on cluster size and engagement levels.

The simulation then tracks these interactions through two key metrics visualized in real time: the AI cluster percentage and happy agents plots.

## 2.2 Simulation Environment

Our simulation operates within a grid-based environment implemented using the Mesa agent-based modelling framework. Each agent in the simulation (humans, social bots, and later, recommendation algorithms) occupies a cell in a toroidal grid, meaning agents that move past one edge of the grid reappear on the opposite edge. This prevents artificial boundary effects and allows for continuous interactions between agents. Agents interact with neighbours within a defined radius based on homophily, engagement levels, and AI influence. The gird's density parameter determines the population concentration, affecting how frequently agents encounter each other and form clusters. Movement in the grid is then governed by the agent's happiness calculations, which combine content preference matching bot influence and engagement-based modifications. The homophily principle is used to model selective exposure, which means that humans prefer to interact with others who share their content preferences. Prior research done by Hussein et al. (2020) has shown that YouTube's recommendation system can lead to a filter bubble effect, where users who engage with misinformative content are subsequently recommended more of the same content. Our simulation reflects this mechanism by modelling how social bots and AI-driven engagement influence content exposure, leading to the gradual formation of echo chambers. This environment structure allows us to observe the emergence of ideological clusters, replicating how online users interact with algorithmic recommendations and bots in a networked media ecosystem.

## 2.3 Agent Design

<u>Human Users:</u> Starts with a random content preference and moves towards like-minded neighbours, influenced by both homophily and AI engagement.

<u>Social Bots:</u> Amplify content engagement, which reinforces echo chambers by increasing exposure to preferred topics.

<u>YouTube Recommendation Algorithm (for DEL 4):</u> In the next deliverable, the recommendation algorithm will dynamically adjust content promotion based on engagement.

## 2.4 Interaction Dynamics

Agents update their behaviour in step-based iterations, moving or reinforcing content exposure based on similarity, engagement, and AI influence. Humans determine whether they are "happy" with their neighbours and relocate if necessary. Social bots enhance content reinforcement and attract human agents to specific ideological clusters. AI-to-AI interactions further increase amplification effects. As Grusauskaite et al. (2024) suggests, real-world echo chambers are often self-sustaining because users actively engage with preferred content while rejecting counter-narratives. Our model reflects this by allowing users to cluster together and resist exposure to contrasting viewpoints, simulating the dynamics of selective exposure seen in real YouTube communities.

### 2.5 Data Collection & Visualization:

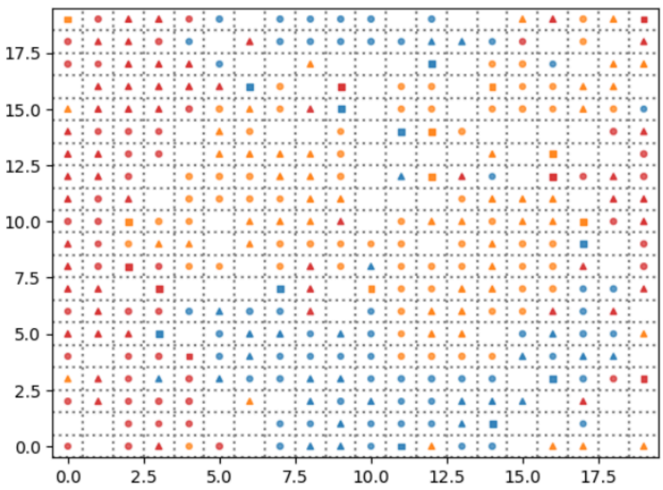
To track echo chamber formation, we collect data on agent clustering, AI influence, and happiness levels. The simulation includes a grid visualization showing agent distribution, along with graphs tracking AI cluster formation, and happiness trends over time. These metrics help analyze how engagement reinforcement and AI-driven interactions shape content exposure dynamics.

## §3. Preliminary Observations & Results

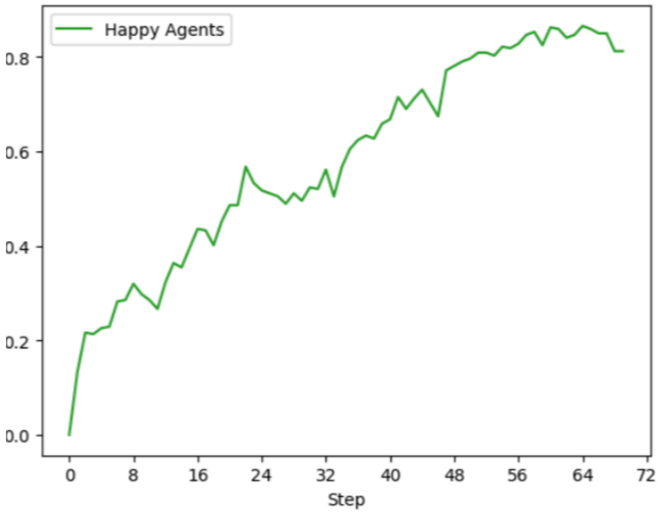
### 3.1 Early Simulation Results

Our early simulation demonstrates the formation and dynamics of echo chambers on Youtube by showcasing the interaction between human users and social bots. Using Mesa's agent-based modeling along with Solara visualization, we are able to observe how content preferences and social bot influence leads to the emergence of distinct content bubbles. At the start of the simulation, Human agents (circles) and social bots (squares) each with a colour that represents their content preference on Youtube are distributed randomly on a grid. Then as the simulation progresses, we see these agents move across the grid based on their content preferences and homophily levels. This visualization then reveals four key behaviours that emerge from our echo chamber simulation: 1) agents with similar content preferences cluster together, forming distinct colour groupings that represent content bubbles, 2) social bots transform from squares to triangles when they cluster, indicating coordinated influence, 3) agent sizes increase within clusters due to higher engagement levels, and 4) clustered bots become more opaque, representing their increased amplification power.

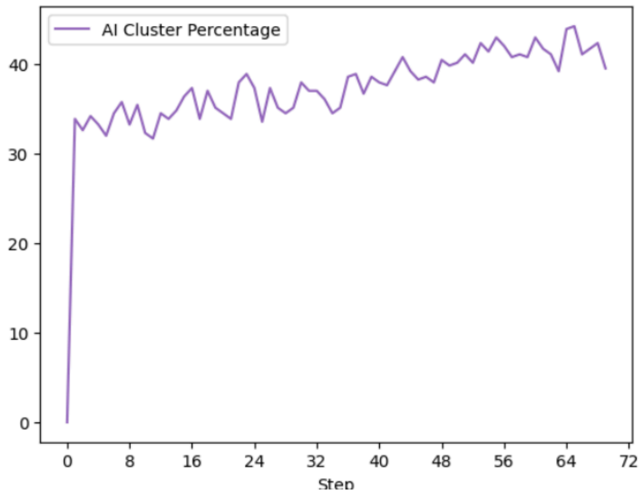
The simulation tracks these dynamics through two key metrics: the AI cluster percentage plot and the happy agents plot. The AI cluster percentage plot showcases how social bots with content preferences form clusters over time, while the happy agents plot demonstrates the overall satisfaction level of agents as they settle into these content-based communities. AI cluster percentage rises as bots form effective groups, while the happy agents percentage increases as humans settle into bot-influenced content bubbles. Both plots as a result reflect the pattern that when bots cluster together, shown as triangles that are more opaque, it leads to a strong influence on nearby human agents, shown as larger agents. The colour coding of agents based on content preferences then makes it visually apparent when echo chambers, as agents with similar content preferences cluster together, with bot clusters often at the center of these formations. This pattern mirrors Youtube's phenomenon where social bot to social bot interactions can create and reinforce content echo chambers.



Agent Grid: visually showcases the above mentioned four key behaviours




Happy Agents Graph: metrically showcases overtime that agent satisfaction levels increase over time as they settle into content-based communities




AI Cluster Percentage Graph: metrically showcases overtime that social bot to social bot interactions overtime form AI clusters over time

### 3.2 Unexpected Behaviors

During early runs there were several unexpected patterns that emerged from the agent interactions. First of which was that bot clusters would occasionally form “super-nodes” with amplification powers approaching the maximum cap of 3.0, creating unexpectedly strong zones of influence. This emerged from the compounding effects of cluster size bonuses (set to 10% per bot) and engagement bonuses in the “_amplify_bot_power” method in agents.py. A second unexpected pattern was that we observed a “cascade effect” where human agents rapidly shifted their positions once bot clusters reached a critical mass, that is bot_cluster > 3. This behaviour seems to stem from the combined impact of bot influence (set to 30% weight) and the additional cluster bonus (up to 15% weight) in the happiness calculation.

Another unexpected dynamic was the formation of “bridge agents”. These are human users who remained stable between different content clusters due to the balanced influence of multiple bot groups. This seems to have emerged from our weighted happiness calculation where content matching (50%), bot influence (30%) and engagement (20%) created stable positions at cluster boundaries.

## §4. Challenges & Next Steps

### 4.1 Development Challenges

One of the most difficult aspects of implementing the simulation was ensuring that agent behaviours accurately represented real-world engagement dynamics. Fine tuning AI influence also proved to be slightly challenging, as overly strong reinforcement reduces the organic formation of echo chambers.

Another major challenge involved working within the constraints of the Mesa library. Although Mesa's built-in visualization tools allowed us to display agent states, visualizing AI-driven recommendation effects was difficult. Implementing more detailed tracking to show how AI agents influence human content exposure over time required extensive modifications to the portrayal function.

Additionally, we had some troubles with our initial stopping condition which halted when all agents were “happy”. This was impractical because ideological clusters prevented some agents from ever reaching a fully satisfied state. To address this, we implemented a temporary step limit of 100 steps to ensure that the simulation doesn't continue indefinitely.

Engagement tracking also presented some challenges, as AI-driven amplification effects were difficult to quantify precisely in real-time. Due to this, we opted to remove the engagement graph for this version of the prototype. However, in the last deliverable, we plan to refine our engagement tracking metrics to capture how different agents types interact over time, making it possible to visualize gradual shifts in content exposure.

A significant implementation challenge we came across was modeling the complex feedback loop between bot clustering and their amplification effects. A change we had to make to the model was adding the method "_amplify_bot_power" which required calibration of cluster bonuses (10% per bot, capped at 50%) and engagement bonuses to prevent runaway amplification effects. Initially, bot clusters would become too influential too quickly, creating artificially echo chambers that didn't reflect realistic formation patterns seen on Youtube. To address this, we implemented a gradual influence scale and multiple influence caps: a cluster size cap of 3 for how far bots can influence humans, an amplification power cap of 3, and a 15% cap on bot cluster influence.

### 4.2 Refinements for the Final Report

As we move toward the final deliverable, there are a few aspects of the simulation that need further refinement and expansion.

One key area of development is the recommendation algorithm agent, which will simulate YouTube's engagement-driven content promotion. Unlike social bots, which act as static engagement amplifiers, the recommendation algorithm will dynamically adjust content visibility based on user trends. This feature will allow us to explore how algorithmic bias shapes long-term ideological exposure.

Additionally, we plan to enhance the data collection process by refining engagement tracking metrics. The engagement graph, which was removed in this version, will be reintroduced in DEL 4 to provide better insight into how human and AI engagement evolve over time. We aim to refine AI clustering analysis by tracking how social bots and recommendation algorithms interact to influence content diversity over multiple steps.

We will also revisit our stopping condition, as using a fixed step limit may not always provide the best representation of echo chamber formation. We aim to implement a more dynamic stopping criterion to offer a better indicator of when an echo chamber has fully formed.

To make the model easier to understand, we will add a legend with descriptions explaining what the different agent types, colours and shapes represent. This will provide clearer insight into what is happening at each step of the simulation.

Lastly, we will incorporate more academic literature on echo chambers, algorithmic bias and social media engagement patterns. By grounding our findings in existing research, we can better understand our results and discuss how our simulation aligns with our chosen phenomena.

## §5. References

Brown, M. A., Nagler, J., Bisbee, J., Lai, A., & Tucker, J. A. (2023, October 26). *Echo chambers, rabbit holes, and ideological bias: How YouTube recommends content to real users.* Brookings Institute. https://www.brookings.edu/articles/echo-chambers-rabbit-holes-and-ideological-bias-how-youtube-recommends-content-to-real-users/

Grusauskaite, K., Carbone, L., Harambam, J., & Aupers, S. (2024). *Debating (in) echo chambers: How culture shapes communication in conspiracy theory networks on YouTube.* New Media & Society, 26(12), 7037–7057. https://doi.org/10.1177/14614448231162585

Hussein, Eslam, Prerna Juneja, and Tanushree Mitra. (2020). *"Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube."* Proceedings of the ACM on Human-Computer Interaction, vol. 4, no. CSCW1, May 2020, Article 48, 27 pages. https://doi.org/10.1145/3392854.

## §6. Attestation

### 6.1 Documentation and Overview of Contributions

Simon:

- Found references 1-3, which were used to back some of our decisions/findings and were cited in APA format (section 5)
- Tasked with writing the “Phenomenon Statement” and “Problem Statement”
    - Explained the impact of AI-driven recommendations and social bots on echo chambers
- Contributed alongside other group members for sections 2 & 4
    - Described the system overview, explaining how simulation models ideological clustering using agent-based modelling in Mesa
    - Outlined the simulation environment, detailing agent interactions, homophily, engagement influence, and the grid-based structure
    - Described the difficulties in fine-tuning AI influence and ensuring realistic engagement dynamics in the “Development Challenges” section

Vincent:

- Wrote section 3, preliminary observations and results
    - Described how our early simulation results illustrate the phenomenon of interest and described unexpected behaviors and/or emergent dynamic
- Contributed alongside other group members for sections 2 & 4
    - In section 2, contributed by adding details about our system overview (such as specific agent attributes and bot clustering mechanics) and by adding details about our simulation environment (such as grid density and its impact and weighted happiness calculations)

    - In section 4, contributed by adding additional challenges we faced such as bot clustering balance challenges and multiple influence cap implementation and how we addressed them accordingly

Isaac:

- Wrote the phenomenon illustration portion of section 1
    - Created detailed descriptions for the simulation in different stages
    - Provided relevant diagrams to illustrate how the simulation changes over time
    - Identified relationships between simulation and chosen phenomenon at each stage
- Contributed alongside other group members for section 4
    - Highlighted the struggles we faced when trying to work with the constraints of the Mesa library
    - Outlined our plans for the next deliverable to revisit our stopping condition, add more descriptions to the model to make it easier to understand, and incorporate more literature in our final report
- Revised and formatted the entire document to enhance clarity and ensure it meets all expectations as specified.

## 6.2 Assignment of CRediT Roles:

Simon:

- Data curation: lead // located and selected all academic references used in the report
- Writing - original draft: lead // responsible for writing the majority of Section 1
- Writing - review & editing: supporting // refined sections 2 & 4 with added details
- Methodology: support // helped describe the agent-based modelling approach and simulation environment
- Visualization: equal // contributed to explaining how the simulation visually represents ideological clustering and agent interactions

Vincent:

- Project administration: lead // assigned tasks to each team member for prototype & report
- Validation: equal // each group member was responsible for verifying what each member wrote
- Writing - original draft: lead // responsible for writing Section 3
- Formal analysis: lead // interpreted simulation results and described behaviours
- Writing – review & editing: supporting // refined sections 2 & 4 with added details
- Visualization: equal // provided visualizations of prototype in section 3

Isaac:

- Writing - original draft: support // created complete descriptions for portions of sections 1 and 4
- Writing - review & editing: lead // reviewed and formatted sections in the entire document
- Investigation: lead // investigated the relationships between our implemented simulation and chosen phenomenon

- Visualization: equal // provided visualizations of prototype in section 1

## 6.3 CRediT Acknowledgement

All authors have read and approved the final report.

Simon: Data curation (lead), Writing – original draft (lead: section 1), Writing – review & editing (supporting), Methodology (supporting), Visualization (equal)

Vincent: Project administration (lead), Validation (equal), Writing - original draft (lead: section 3), Formal analysis (lead: section 3), Writing – review & editing (supporting), Visualization (equal)

Isaac: Writing - original draft (support: sections 1 and 4), Writing - review & editing (lead), Investigation (lead), Visualization (equal)