# HIERARCHICAL GENERATIVE AGENTS FOR SIMULATING SEQUENTIAL HUMAN BEHAVIOR

**Maria G. Mendoza*** **Lucas Waldburger*** **Jin Lee** **Shankar Sastry**
University of California, Berkeley
{maria_mendoza, lwaldburger, jinlee}@berkeley.edu
sastry@coe.berkeley.edu

## ABSTRACT

Complex cognitive, emotional, and social processes shape human evacuations during natural disasters. Accurate modeling and understanding of human behavior in disasters or emergencies can greatly impact the evacuation process by informing more effective planning and resource allocation. However, collecting human data in these situations is very difficult, and existing computational evacuation models assume rational, homogeneous behavior, leading to unrealistic, overly optimistic predictions. To address this gap, we present a simulation framework of sequential human decision-making during an evacuation scenario, introducing cognitively grounded, persona-driven agents. Our framework models evacuation behavior in a grid-based urban environment that evolves over time, capturing fire and other hazards. Human agents are modeled as personas that make sequential decisions in response to environmental stimuli with cognition structured in three levels: high-level evacuation goals, mid-level route reasoning, and low-level navigation. Decision-making is driven by large language models (LLMs) coupled with a cognitive module and calibrated with empirical human evacuation data. We propose a dynamic, stimulus-driven disaster simulation framework that models human evacuation decision-making using persona-conditioned LLM agents and a cognitive hierarchy.

## 1 INTRODUCTION

Inefficient or delayed evacuations during natural disasters result in preventable fatalities. Humans often fail to evacuate promptly due to confusion, misjudgment, or a lack of clear guidance. Disaster response is not purely driven by infrastructure or disaster severity; instead, human behavior is a key determinant of successful evacuation. Decades of research show that humans rarely behave as rational, utility-maximizing agents (1), especially in high-stress situations, such as disaster scenarios (2; 3). Empirical evacuation studies document behavioral delays, gathering of belongings, assisting neighbors or family members, strong attachments to property and pets, and the misinterpretation of environmental cues (4; 5). These behaviors are further influenced by emotional responses, such as panic or stress, and cognitive biases, such as familiarity seeking in route decision planning or mistrust in alarms (6). Moreover, vulnerable populations, including the elderly, disabled, and those living with dependents, are disproportionately affected, yet are often underrepresented or homogenized in computational models.

Despite extensive study of human behavior in disaster scenarios, existing simulations assume idealized behavioral patterns (2). These assumptions lead to suboptimal results when incorporated into search-and-rescue technologies, such as drones or other robotic systems, leading to failures when deployed in real-world scenarios (7; 8). Consequently, there is a pressing need for simulation frameworks that model heterogeneous social and cognitive decision-making as it evolves over time.

Recent advances in large language models (LLMs) offer new opportunities to emulate diverse human behavior. However, existing work using heuristics, rule-based systems, or learning-based methods typically focus on one-shot output. As a result, they fail to capture the sequential, time-varying

*Equal contribution.

nature of evacuation behavior where decisions unfold over minutes, hours, or days, influenced by both social context and dynamic environmental conditions.

In this work, we introduce a simulation framework for modeling human behavior during disaster evacuation using persona-conditioned LLM agents embedded in a dynamic urban environment.[1] The framework enables realistic, long-horizon evacuation modeling by coupling cognitively informed LLM decision-making with environments driven by external stimuli such as hazard progression and limited observability. Beyond behavior modeling, the framework supports applications including emergency resource allocation, evacuation strategy evaluation, alarm system design, urban planning, worst-case scenario analysis, and the testing of human–robot interaction strategies for rescue robotics. Our contributions are as follows:

1. A persona decision-making framework that incorporates empirical evacuation factors informed by survey data (9) and traditional Protection Action Decision Model (PADM) frameworks, implemented through LLM agents whose decisions better reflect realistic human tendencies (See Appendix A).
2. A lightweight, easy-to-deploy simulator that captures agent mobility and sequential decision-making in an urban environment with dynamic external stimuli.
3. A cognitively inspired modeling strategy that leverages LLM reasoning through API calls and multi-level planning to enable long-horizon simulations of diverse, evolving human behavioral responses.

## 2 Related Work

### 2.1 Human Behavior in Disasters

Human behavior during disasters is irrational, heterogeneous, and difficult to predict, yet it plays a decisive role in the outcome of evacuation and rescue operations (4). Factors such as risk perception, social influence, emotional responses (e.g., panic, stress, hesitation), and resource or physical constraints all shape how and when people choose to act (2). These behaviors can lead to delayed or inefficient evacuations, sometimes with catastrophic consequences (3). Accurate simulation of evacuation is therefore essential for planning and technology design, but current modeling approaches and data are limited.

Some evacuation models assume rational actors or focus narrowly on static route optimization and ignore assumptions such as (1) evolving environments, (2) cognitive and emotional variability in response to group dynamics and external stimulus, and (3) partial observability of evacuees. To advance evacuation planning, there is a need for models that capture how dynamic environments and social interactions influence human decision-making and generate a range of possible scenarios.

### 2.2 Traditional Rule-Based Methods

Agent-Based Modeling (ABM) is a widely used approach for simulating multi-agent systems governed by predefined rules, enabling the study of emergent collective behavior, and has been extensively applied to evacuation modeling (10). In this context, ABMs have been used to represent agent navigation, local interactions, communication, social influence, and deviations from rational behavior such as panic (11; 12). These models enable quantitative analysis of evacuation outcomes, including congestion, bottlenecks, and evacuation times. Despite these strengths, most ABM approaches rely on fixed rule-based or parametric decision logic, which limits their ability to capture adaptive, sequential, and context-dependent human decision-making under rapidly evolving disaster conditions (13).

### 2.3 Data-Driven Methods

Prior work has employed data-driven approaches, including survey and GPS-based mobility data, to study evacuee decision-making and movement patterns during real-world fire events (14; 15; 16). These datasets enable high-level analysis of evacuation flows, particularly at the city scale, but

[1] Code is available at github.com/lucaswaldburger/hierarchical_LLM_agents.

individual-level route choices are often unavailable due to privacy constraints. As a result, fine-grained evaluation of evacuation behavior across diverse demographic profiles remains limited.

To mitigate these data limitations, Snopkova et al. introduced a controlled virtual-reality dataset capturing individual evacuation route choices, decision times, and route decisions in intersections (9). Complementary work has leveraged mixed-reality environments to collect fine-grained human evacuation behavior, including detailed mobility trajectories and routing decisions under simulated fire conditions (17). While the generalizability of such datasets to real-world environments requires further validation, they provide valuable empirical insight into human decision-making processes during evacuation. However, collecting high-fidelity human behavioral data remains costly and difficult to scale.

### 2.4 Hybrid Approaches

Previous studies have explored how LLMs can empower the agency of ABM. Park et al. explored emergent, interactive human behaviors using LLMs (18). Chopra et al. proposed LLM archetypes to integrate with ABM for large-scale simulations (19). Dai et al. use a persona-aware vision–language model with chain-of-thought reasoning to generate explainable cyclist safety and comfort assessments from street-view imagery (20). Although LLM-integrated agents are known to diversify human decision-making, their effectiveness in emergencies remains uncertain, as disasters often elicit irrational behaviors that are difficult to observe or ethically study. Chen et al. addressed some of these challenges by combining LLMs, survey data, and reinforcement learning to reduce the mismatch between behavior theory and LLM predictions (21), but their framework models evacuation as a binary outcome: deciding to evacuate or not. However, effective disaster response requires understanding not only *whether* individuals evacuate, but *when* and *how* they do so, as evacuation unfolds sequentially over hours or days, with highly variable departure times and congestion effects (14). This motivates a long-horizon modeling approach that captures a sequence of decisions rather than a single-point choice.

## 3 Methods

### 3.1 Simulation Environment

We model an urban environment as a discrete-time, 2D grid world and simulate an evacuation scenario using MiniGrid (22) as shown in Figure 1. The world contains static elements such as buildings or homes, roads, and parks. Each element can be accessible or visible to an agent depending on its public or private designation. For example, an agent's home is only accessible and visible to that specific agent, whereas parks and streets are public and therefore accessible to all agents. We also model spatio-temporal objects such as traffic, fire spreading stochastically, and smoke, which reduces an agent's visibility and may trigger an emotional response.

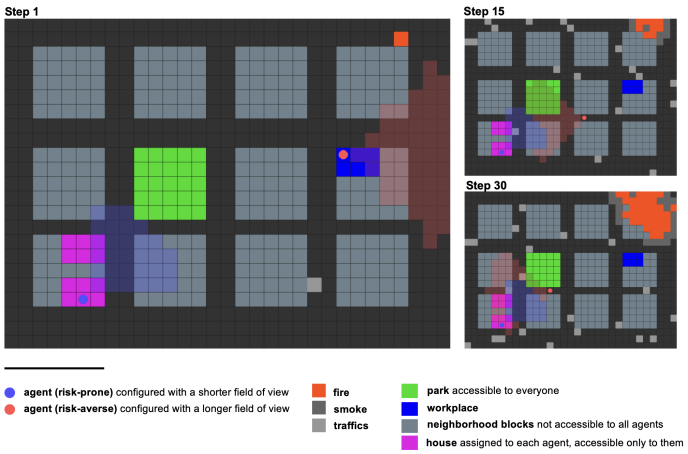


Figure 1: The MiniGrid world with indexed static elements

We formalize the environment as a finite 2D discrete grid $\mathcal{X} = \{1, \ldots, W-1\} \times \{1, \ldots, H-1\} \subset \mathbb{Z}^2$, where each cell $x \in \mathcal{X}$, is assigned a semantic type following a labeling function $\ell : \mathcal{X} \to \mathcal{L}$, The dynamic environment state at time step $k$ is $\mathcal{G}_k = \langle \mathcal{G}_{\text{static}},\ F_k,\ S_k,\ \mathcal{T}_k,\ e_k \rangle$, capturing the static

and dynamic elements of the environment, such as fire spread, smoke, traffic, and external stimuli. A detailed formulation is provided in Appendix B.1

### 3.2 PERSONAS

Each agent is modeled as a persona with demographic, cognitive, social, and spatial attributes, which are passed as conditioning variables to the LLM. These traits influence how agents perceive risk, interpret environmental and social cues, experience delays, prioritize goals, and navigate the environment. Persona attributes are informed by empirical evacuation literature and the PADM, capturing heterogeneity in risk perception, trust in alerts, social influence, and pre-evacuation behavior.

**Demographic Attributes:** Individual characteristics such as name, gender, age, occupation, and workplace location; and social ties such as friendships, dependents, and propensity to help others.

**Cognitive and Behavioral Characteristics:** Attributes that shape decision-making during disasters, informed by the PADM. These include risk perception, threat assessment levels, and sensitivity to environmental cues (e.g., alarms, smoke, traffic). In addition, we model susceptibility to social influence (e.g., messages from coworkers or neighbors), likelihood of delaying, ignoring, or acting on alerts. Lastly, we include personal motivations and priorities (e.g., retrieving a pet or belongings) such as housing type (renting vs. owning) (16).

**Spatial Knowledge and Memory:** Familiar routes, known locations, and place-based preferences that influence navigation.

Persona traits modify every component of the cognitive architecture. They shape the thresholds for interpreting risk, the urgency of updating goals, trust in alarms, willingness to help others, and preference for familiar or unfamiliar routes. These traits are fed into LLM prompts so that each agent responds differently to identical stimuli based on its cognitive and social profile.

Formally, each agent $i \in \mathcal{M}$ is characterized by a persona $\psi^i = \left(\psi^i_{\text{att}},\ \psi^i_{\text{cog}},\ \psi^i_{\text{soc}},\ \psi^i_{\text{spat}}\right)$, encoding demographic, social, and spatial attributes. The cognitive parameters $\psi^i_{\text{cog}} = (\rho^i, \tau^i, \theta^i)$ are grounded in the PADM, capturing risk perception, trust in alerts, and threat assessment sensitivity (see Appendix B.2.1 for the full specification).

### 3.3 COGNITIVE ARCHITECTURE

We designed a cognitive architecture that enables each agent to interact with the environment and make human-centered decisions during disasters. Our approach integrates empirical behavioral data with LLM reasoning to diversify agent responses and model realistic variation across personas. To achieve this, we construct a cognitive module that simulates the decision-making processes most relevant to evacuation scenarios, focusing on the aspects of cognition that meaningfully influence behavior under time pressure and uncertainty. Details of these data sources and how we use the PADM to inform agent behavior are provided in Appendix A. While individual human responses vary widely, we approximate this diversity by parameterizing personas according to the characteristics described in Section 3.2.

**Perception:** Each agent has a finite field of view (FOV) representing how far they can observe in the environment. This determines which hazards, obstacles, social cues, and environmental changes they can detect. Perception also determines awareness of other agents, traffic, fire, and smoke.

**Memory:** The memory module stores key information about the environment and the agent's past behavior. This includes spatial memory (e.g., familiar routes, accessible areas in the environment), their current and past high-level goals, important decisions previously taken, and key hazard observations. Memory allows agents to maintain continuity in their plan, revise their strategies based on previous failures, and support behaviors such as returning to dependents or avoiding previously dangerous areas.

**Planning:** As described in Section 3.4, we implement a multi-level planning structure that allows agents to make sequential decisions over time. High-level planning governs intent formation in response to external stimuli; mid-level planning determines route choices conditioned on risk perception and persona traits; and low-level planning performs short-horizon navigation based on local observations. This hierarchy reflects how humans adjust plans dynamically as conditions evolve.

**Reflection:** Reflection evaluates urgency, weighs competing priorities, and updates goals. With LLM reasoning, agents consider factors such as whether to trust an alarm, whether to retrieve belongings, or whether to stop and help another agent. Reflection serves as the cognitive "check" that allows goals to be revised when conditions meaningfully change.

**Cognitive Loop and Module Interaction:** At each timestep, agents operate through a sequential loop: *Perception* → *Memory* → *Reflection* → *Planning* → *Action*, linking environmental observation, state, urgency assessment, hierarchical decision-making, and execution.

This cycle repeats continuously, enabling agents to adapt their behavior as the disaster evolves. The full agent state and cognitive module specification are given in Appendix B.2.

### 3.4 Multi-Level Planning System

Our framework uses a multi-level planning system, as in Figure 2, that shows how and when the LLM is prompted, and at which steps the agents make decisions.

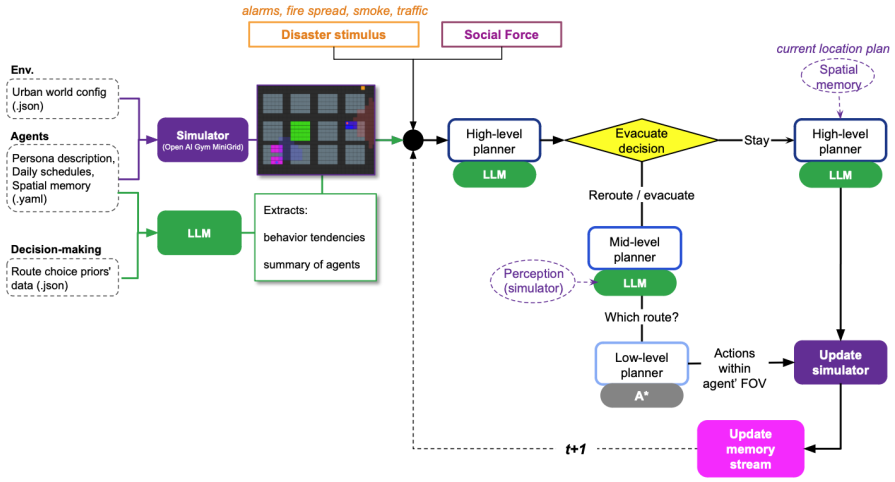


Figure 2: Multi-level decision-making architecture illustrating planner hierarchy, information flow, and event-triggered decision points.

The high-level planner engages the LLM when an agent experiences an external stimulus. External stimuli include alarms, official messages, visible smoke thresholds, fire spread, social messages, and environmental changes such as congestion. These triggers prompt high-level reassessment and may override routine plans and are typically broadcast to many or all agents. At this level, the LLM outputs high-level intentions such as ignoring the alert, preparing to evacuate, returning home, or gathering dependents. These decisions are constrained by spatial feasibility, memory, and reflection-based urgency assessments. Alternatively, the LLM may output a decision to evacuate immediately or travel to another location, such as going home to retrieve a pet or prepare belongings. Once a high-level intention is set, the mid-level planner translates it into route decisions, which are then executed step-by-step by the low-level navigator. The agent's policy decomposes as $\pi^i = (\pi^i_H, \pi^i_M, \pi^i_L)$, where $\pi^i_H$ produces a high-level goal $g^i_k \in \mathcal{G}_H = \{\texttt{stay}, \texttt{evacuate}, \ldots\}$, $\pi^i_M$ selects a route $r^i_k \in \mathcal{R}^i_k$ from feasible alternatives, and $\pi^i_L$ executes short-horizon navigation via A* search within the agent's field of view. The full hierarchical planner formulation is in Appendix B.4.1.

Mid-level planning captures the agent's route-selection behavior as it moves through the environment, guided by perceived risk, threat level, and traditional PADM principles. We condition the LLM using publicly available behavioral data from (9), enabling the agents to generate actions that reflect empirically grounded human decision patterns given their observations and persona (see Section A). At this stage, the LLM outputs a decision on which street or route to follow; for instance, some agents favor familiar paths, others respond more strongly to cues such as visible smoke or congestion, and some exhibit higher risk tolerance that delays protective action. By delegating these decisions to the LLM, the mid-level planner avoids assuming perfectly rational navigation and instead incorporates persona-specific behavioral variability. Once a direction is chosen, the low-level

planner executes a short sequence of actions, which is reasonable because over a short horizon, humans tend to follow efficient local movement until new observations prompt reconsideration.

The low-level planner manages short-horizon navigation constrained by the agent's FOV. While the mid-level planner may designate a general direction (e.g., turn right, proceed along a given street), the agent cannot anticipate obstructions beyond what is locally observable. To capture this uncertainty, movement is restricted to positions within the FOV, and a simple A* search is used to plan only a few steps. This choice serves as a computational approximation of human local navigation behavior, where individuals typically follow the most direct visible path and adapt incrementally to obstacles. The mid and low-level planners aim to improve realism by capturing how decisions evolve with changing observations and computational efficiency by limiting long-horizon pathfinding.

Beyond environmental cues, agents are also influenced by social interactions, which we model through a dedicated layer. The social interaction layer captures interpersonal influences that shape evacuation timing and route choice. These interactions include brief information exchanges, such as a co-worker warning about smoke or encounters with neighbors, where an agent may decide to wait, offer help, or exchange additional information depending on their persona.

# 4 RESULTS

## 4.1 FIRE DISASTER SCENARIO

Our simulation consists of $n-$agents in the grid world, where they react to hazards like fire or traffic, and take actions in an evacuation setting. The fire spreads probabilistically based on the number of adjacent fire tiles (23). A tile is set on fire if any of its neighboring cells are burning with probability $p = 0.05$. Surrounding the fire is smoke, which reduces the FOV size of the agent.

In addition to the disaster, the agent must navigate through dynamic obstacles, such as traffic. Spatio-temporal density of traffic can be fine-tuned within specific areas in the environment. This is intended to simulate how traffic might be higher in evacuation zones during a disaster condition. Throughout the episode, traffic forces the agent to take detours from their original trajectory, creating congestion that affects route planning and evacuation timing. Together, the fire and traffic produce non-deterministic patterns that test the agent's decision-making and ability to navigate changing conditions.

Agents begin with different starting contexts (e.g., at home with family or at a workplace) and receive external events of varying severity, eliciting heterogeneous responses. For instance, an agent at work who is risk-prone or exhibits low trust in authorities may choose not to react to an initial text alert indicating the presence of a fire but advising that evacuation is not yet necessary. As hazard urgency increases, such as when smoke is close to the agent, the high-level planner triggers evacuation behavior. Interestingly, persona-based differences emerge naturally among the agents. The risk-averse agent chooses a familiar park as a safe zone, while the risk-prone agent tends to remain in place longer unless the danger becomes unmistakable. At the mid- and low-level, agents dynamically reroute when traffic or fires block their path, exhibiting a preference for familiar streets. Representative outputs are shown in Figure 3.

Other agents' behavior outputs shown in Figure 3 confirms alignment with the PADM analysis workflow (see Figure 6). The LLM output components, such as the Urgency score, Safety assessment, and Social/environmental cues, are mapped to the sequential steps of the PADM. For example, in Step 28, Isabella's log shows her perception ("heavy traffic near walkable") affecting the Urgency score (0.20), which triggers the decision to evacuate and communicate with Klaus. Isabella's resulting action to move right to avoid the perceived traffic hazard eventually serves as the Low-level goal. This demonstrates how our model captures the PADM steps from threat and risk perceptions to behavioral response.

## 4.2 BEHAVIOR VALIDATION

We evaluated how well agent characteristics align with observed environmental behavior using a minimal validation scenario. The setup consists of two agents, a risk-prone (blue) and a risk-averse (red) agent, and a static fire hazard (middle red square). The fire is kept static to reduce episode-to-

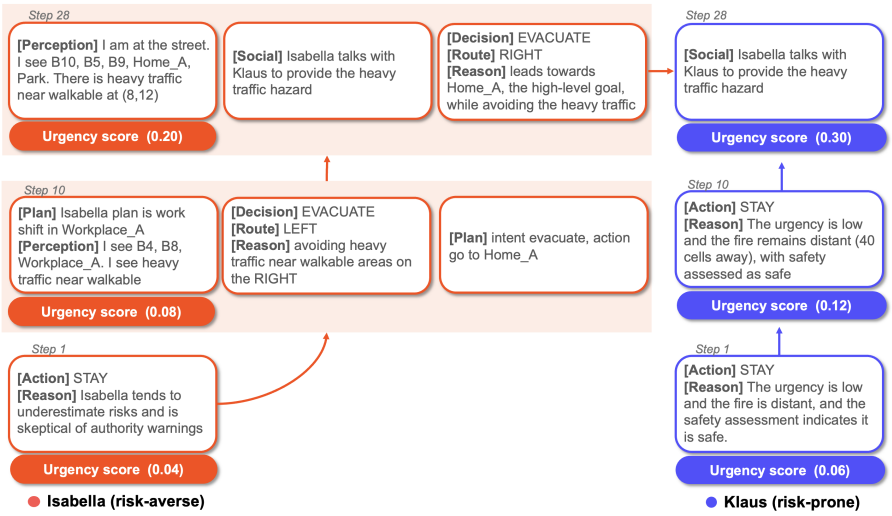


Figure 3: A route-choice tree for Isabella and Klaus based on high (decision), medium (route), and low-level (action) planners.

episode variance. Given the hazard, each agent must go to their corresponding home, indicated by pink tiles at the left and right ends of the environment. The risk-averse agent consistently maximizes its distance from the fire by navigating along the boundary of the simulation before returning home. In contrast, the risk-prone agent tolerates closer proximity to the fire compared to the risk-averse agent before rerouting toward a safe location. These results indicate that at a high-level, there is alignment between the agent's characteristics and their behavior within the environment, while there are minor specifications that need to be optimized, likely through improved prompt engineering.

Figure 4: Agent behavior validation. Left: The behavior validation environment where the risk-averse agent (red) and risk-prone agent (blue) must avoid the fire in the middle and go to either of the safe zones (pink). Right: Trends in mean fire proximity as a function of time for agents with different risk profiles.

### 4.3 Scalability

We sought to quantify how well the current implementation scales as a function of agents in order to study emergent behaviors in multi-agent systems in disaster scenarios (Figure 5). Since each agent requires API calls for decision-making, it is important to consider how these computational costs might affect community-level simulation. We ran the fire disaster simulation over three episodes with one to five agents. The number of tokens used increases drastically as more agents are placed in the simulation. The token usage spikes at time steps 10 and 20 when there are three or more agents in the simulation. This likely coincides with time steps where several agents must plan due to an external stimulus and thereby make an API call. We also found that the difference in token usage

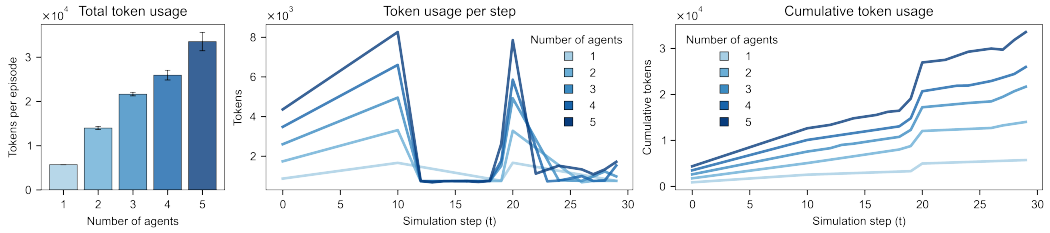


Figure 5: Scalability analysis of the multi-agent system. Left: Distribution of total token usage by the LLM as a function of the number of agents. Middle: Total number of API calls (completions) required per experiment, grouped by agent count. Right: Cumulative token usage over time steps, averaged across runs, illustrating how token consumption grows with the addition of more agents. Each experiment was repeated across 3 episodes.

becomes more volatile as a function of agents since the number of planning steps becomes more variable. These results indicate that the current implementation has limited scalability. Improved scalability can be achieved by using an agent archetype, as in Chopra et al. (19) and modifications to the cognitive architecture.

## 5 Discussion

Our framework introduces persona-conditioned LLM agents grounded in behavioral theory and empirical evacuation data for modeling sequential human decision-making during disasters. Through controlled experiments, we demonstrate that persona conditioning yields risk-sensitive behavior aligned with the PADM: risk-averse and risk-prone agents exhibit statistically distinguishable hazard proximity profiles (Section 4.2), and LLM output components (urgency scores, safety assessments, social cues) map onto the sequential PADM stages (Figure 6). A central challenge for quantitative evaluation of agent-level evacuation models is the absence of ground-truth sequential decision data; existing real-world datasets capture aggregate mobility flows (14) or post-hoc survey responses (15; 16), not the step-by-step decision traces our model produces. We plan to address this through distributional validation, comparing simulation outputs such as departure-time curves and congestion patterns against documented empirical signatures, and by running agents through the same T-intersection scenarios as human participants in (9) to compare route-choice distributions quantitatively. The framework is intentionally lightweight, prioritizing rapid iteration and compatibility with learning-based methods, enabling the study of system-level phenomena such as congestion formation, evacuation delays, and bottleneck dynamics.

## Acknowledgments

This work was supported in part by Provably Correct Design of Adaptive Hybrid Neuro-Symbolic Cyber Physical Systems, DAF Air Force Research Laboratory award number FA8750-23-C-0080.

## References

[1] S. Choi, S. Kariv, W. Müller, and D. Silverman, "Who is (more) rational?" *American Economic Review*, vol. 104, no. 6, p. 1518–50, June 2014. [Online]. Available: https://www.aeaweb.org/articles?id=10.1257/aer.104.6.1518

[2] E. Bakhshian and B. Martinez-Pastor, "Evaluating human behaviour during a disaster evacuation process: A literature review," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 10, no. 4, pp. 485–507, Aug. 2023. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S2095756423000740

[3] N. Elhami-Khorasani, M. Kinateder, V. Lemiale, S. L. Manzello, I. Marom, L. Marquez, S. Suzuki, M. Theodori, Y. Wang, and S. D. Wong, "Review of Research on Human Behavior in Large Outdoor Fires," *Fire Technology*, vol. 59, no. 4, pp. 1341–1377, Jul. 2023. [Online]. Available: https://link.springer.com/10.1007/s10694-023-01388-6

[4] Y. Xenidis and G. Kaltsidi, "Prediction of humans' behaviors during a disaster: The behavioral pattern during disaster indicator (bpdi)," *Safety Science*, vol. 152, p. 105773, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925753522001126

[5] N. Verdière, G. Cantin, D. Provitolo, V. Lanza, E. Dubos-Paillard, R. Charrier, M. Aziz-Alaoui, and C. Bertelle, "Understanding and simulation of human behaviors in areas affected by disasters: From the observation to the conception of a mathematical model," *Global Journal of Human-Social Science: H Interdisciplinary*, vol. 15, no. 10, 2015.

[6] M. K. Lindell and R. W. Perry, "The protective action decision model: Theoretical modifications and additional evidence," *Risk Analysis*, vol. 32, no. 4, pp. 616–632, 2012. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.2011.01647.x

[7] M. Nayyar and A. Wagner, "Effective robot evacuation strategies in emergencies," in *2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019*, ser. 2019 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019. United States: Institute of Electrical and Electronics Engineers Inc., Oct. 2019, 28th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2019 ; Conference date: 14-10-2019 Through 18-10-2019.

[8] S. Javdani, S. S. Srinivasa, and J. A. Bagnell, "Shared autonomy via hindsight optimization," in *Robotics: Science and Systems*, 2015.

[9] D. Snopková, M. Tancoš, L. Herman, and V. Juřík, "Predictors of evacuation behavior: dataset on respondents' route choice and web interaction," *Scientific Data*, vol. 12, no. 1, p. 116, Jan. 2025. [Online]. Available: https://www.nature.com/articles/s41597-025-04440-y

[10] G. P. Senanayake, M. Kieu, Y. Zou, and K. Dirks, "Agent-based simulation for pedestrian evacuation: A systematic literature review," *International Journal of Disaster Risk Reduction*, vol. 111, p. 104705, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212420924004679

[11] S. Sharma, K. Ogunlana, D. Scribner, and J. Grynovicki, "Modeling human behavior during emergency evacuation using intelligent agents: A multi-agent simulation approach," *Inf. Syst. Frontiers*, vol. 20, no. 4, pp. 741–757, 2018. [Online]. Available: https://doi.org/10.1007/s10796-017-9791-x

[12] A. Trivedi and S. Rao, "Agent-based modeling of emergency evacuations considering human panic behavior," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 277–288, 2018.

[13] S. K. Kasereka, O. M. Kabwe, M. W. Kinyanta, A. L. Kasongo, G. W. Ilunga, K. Muhambya, T. Tashev, and K. Kyamakya, "Enhancing building safety: A brief review of agent-based modeling for fire evacuation simulation," *Procedia Comput. Sci.*, vol. 257, no. C, p. 668–675, Jan. 2025. [Online]. Available: https://doi.org/10.1016/j.procs.2025.03.086

[14] Q. Wang and J. E. Taylor, "Patterns and Limitations of Urban Human Mobility Resilience under the Influence of Multiple Types of Natural Disaster," *PLOS ONE*, vol. 11, no. 1, p. e0147299, Jan. 2016. [Online]. Available: https://dx.plos.org/10.1371/journal.pone.0147299

[15] E. D. Kuligowski, X. Zhao, R. Lovreglio, N. Xu, K. Yang, A. Westbury, D. Nilsson, and N. Brown, "Modeling evacuation decisions in the 2019 kincade fire in california," *Safety Science*, vol. 146, p. 105541, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925753521003842

[16] A. Forrister, E. D. Kuligowski, Y. Sun, X. Yan, R. Lovreglio, T. J. Cova, and X. Zhao, "Analyzing risk perception, evacuation decision and delay time: A case study of the 2021 marshall fire in colorado," *Travel Behaviour and Society*, vol. 35, p. 100729, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214367X23001801

[17] M. Chen, R. Yang, Z. Tao, and P. Zhang, "Mixed reality lvc simulation: A new approach to study pedestrian behaviour," *Building and Environment*, vol. 207, p. 108404, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0360132321008015

[18] J. S. Park, J. O'Brien, C. J. Cai, M. R. Morris, P. Liang, and M. S. Bernstein, "Generative Agents: Interactive Simulacra of Human Behavior," in *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. San Francisco CA USA: ACM, Oct. 2023, pp. 1–22. [Online]. Available: https://dl.acm.org/doi/10.1145/3586183.3606763

[19] A. Chopra, S. Kumar, N. Giray-Kuru, R. Raskar, and A. Quera-Bofarull, "On the limits of agency in agent-based models," Nov. 2024, arXiv:2409.10568. [Online]. Available: http://arxiv.org/abs/2409.10568

[20] Y. Dai, Z. Wang, C. Wang, K. Zhou, Y. Qian, S. Xu, and X. Yan, "Persona-aware and explainable bikeability assessment: A vision-language model approach," *https://arxiv.org/abs/2601.03534*, 2026.
[21] R. Chen, C. Wang, Y. Sun, X. Zhao, and S. Xu, "From perceptions to decisions: Wildfire evacuation decision prediction with behavioral theory-informed LLMs," *arXiv preprint arXiv:2502.17701*, 2025. [Online]. Available: https://arxiv.org/abs/2502.17701
[22] M. Chevalier-Boisvert, B. Dai, M. Towers, R. de Lazcano, L. Willems, S. Lahlou, S. Pal, P. S. Castro, and J. Terry, "Minigrid & miniworld: Modular & customizable reinforcement learning environments for goal-oriented tasks," *CoRR*, vol. abs/2306.13831, 2023.
[23] C. Hill, *Learning Scientific Programming with Python*, 2nd ed. Cambridge University Press, 2020.

## A Human Behavior in Disaster

We use the following data to build our framework:

- **Predictors of evacuation behavior (9):** provides empirical data on human route-choice decisions during fire evacuations in controlled virtual environments, capturing how environment cues like corridor width, length, and the presence of other transitioning routes influence behavior across different groups. Although the study focuses on human participants navigating simplified building layouts, we leverage this dataset to abstract key decision-making tendencies, rather than replicate individual behavior. In our work, these insights inform how agents with different demographic profiles and risk sensitivities respond to perceived congestion, spatial constraints, and local route preferences. This allows us to ground the LLM reasoning approach in empirically observed patterns while maintaining scalability and generalization to dynamic evacuation scenarios.
- **Human mobility patterns in wildfire (14):** offers the geotagged positional data, capturing human movement patterns before (steady state), during (perturbed state), and after (perturbed state) each wildfire event. This dataset allows for a comparison of mobility characteristics under normal versus disaster conditions. In our project, this dataset influenced how we modeled our disaster scenario, how far people move during the evacuation phase, and the length of the horizon for evacuation.
- **LLM-based method integrating behavioral theory Protective Action Decision Model (PADM) (21):** utilize data collected from local residents via surveys following wildfire events. Here, the PADM, which is a core conceptual framework in disaster psychology, is used to ground the LLM's reasoning. The final output is based on a binary classification (yes or no) for individual evacuation choices, which is simple but shows high predictive accuracy. We refer to their LLM's plan, such as how to integrate the PADM theory with the LLM. Considering PADM, a resident agent's decision-making process focuses on mental states such as risk perception or threat assessment (Figure 6), especially using prior data (9) and then selecting relevant variables that influence a resident's perceptions.

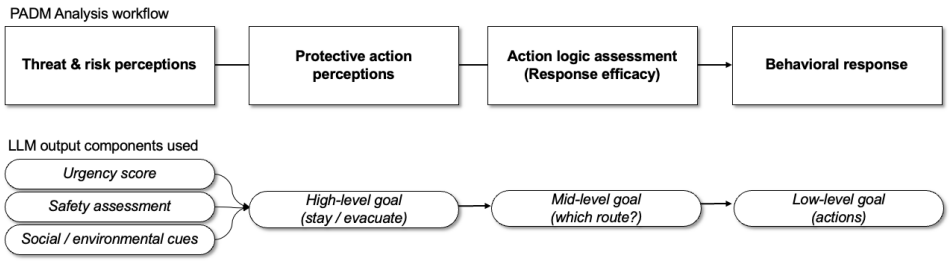


Figure 6: Correspondence between the PADM analysis and the LLM output components used.

# B FORMULATION

## B.1 ENVIRONMENT

We model the environment as a finite 2D discrete grid

$$\mathcal{X} = \{1, \dots, W-1\} \times \{1, \dots, H-1\} \subset \mathbb{Z}^2,$$

where each cell $x \in \mathcal{X}$ represents a spatial unit of the urban area, and $W, H \in \mathbb{Z}_{>0}$ are fixed dimensions.

Each cell is assigned a semantic land-use type via a static labeling function

$$\ell : \mathcal{X} \to \mathcal{L},$$

where $\mathcal{L}$ is a finite set of mutually exclusive categories. We partition $\mathcal{L}$ into *public* types, accessible to all agents, and *private* types, accessible only to designated agents:

$$\mathcal{L}_{\text{pub}} = \{\texttt{road}, \texttt{park}, \texttt{safe_zone}\}, \qquad \mathcal{L}_{\text{priv}} = \{\texttt{building}, \texttt{workplace}, \texttt{home}\}.$$

For each type $\ell \in \mathcal{L}$, we denote the corresponding cell set by $\mathcal{X}_\ell = \ell^{-1}(\ell)$.

**Accessibility and visibility.** For a population of agents $\mathcal{M} = \{1, \dots, M\}$, we define two binary predicates for each agent $i \in \mathcal{M}$ and cell $x \in \mathcal{X}$. Let $\text{owns}^i(x) = 1$ if cell $x$ is the private location assigned to agent $i$, and 0 otherwise.

- **Accessibility:** agent $i$ can physically occupy cell $x$:
$$\mathbf{access}^i(x) = 1\big[\ell(x) \in \mathcal{L}_{\text{pub}}\big] \;\vee\; 1\big[\ell(x) \in \mathcal{L}_{\text{priv}} \wedge \text{owns}^i(x) = 1\big].$$
- **Visibility:** agent $i$ can observe cell $x$:
$$\mathbf{vis}^i(x) = 1\big[\ell(x) \neq \texttt{building}\big] \;\vee\; 1\big[\text{owns}^i(x) = 1\big].$$

Roads, parks, and safe zones are both accessible and visible to all agents. Private cells (home, workplace) are accessible and visible only to their assigned agent. Buildings are neither accessible nor visible to any agent.

**Hazard dynamics.** For a cell $x \in \mathcal{X}$, let $\mathcal{N}(x) \subseteq \mathcal{X}$ denote the neighborhood of $x$ under a fixed adjacency relation (e.g., 4- or 8-connectivity). At each discrete time step $k \in \{0, 1, \dots\}$, let $F_k \subseteq \mathcal{X}$ denote the set of burning cells. Fire spreads stochastically: each cell $x \notin F_k$ ignites independently according to

$$\mathbb{P}(x \in F_{k+1} \mid F_k) = 1 - (1 - p_f)^{|\mathcal{N}(x) \cap F_k|},$$

where $p_f \in (0, 1)$ is the per-neighbor ignition probability. The model assumes fire is permanent: $F_k \subseteq F_{k+1}$. Smoke occupies the traversable cells adjacent to active fire,

$$S_k = \mathcal{N}(F_k) \cap \{x \in \mathcal{X} : \ell(x) \in \mathcal{L}_{\text{pub}}\},$$

where $\mathcal{N}(F_k) = \bigcup_{x \in F_k} \mathcal{N}(x)$.

**Traffic.** At each time step $k$, traffic is a time-varying occupancy on road cells:

$$\mathcal{T}_k = \mathcal{T}_k^{\text{exo}} \cup \mathcal{T}_k^{\text{agents}} \;\subseteq\; \mathcal{X}_{\texttt{road}},$$

where $\mathcal{T}_k^{\text{exo}}$ represents exogenous background vehicles and $\mathcal{T}_k^{\text{agents}}$ captures congestion from evacuating agents. In the current instantiation, $\mathcal{T}_k^{\text{exo}}$ is a time-varying exogenous process parameterized by a spawn rate and density, and $\mathcal{T}_k^{\text{agents}} = \{x \in \mathcal{X}_{\texttt{road}} : |\{j \in \mathcal{M} : c_k^j = x\}| \geq \kappa\}$ for a congestion threshold $\kappa \in \mathbb{Z}_{>0}$.

**Environment state.** The static environment collects the structural elements:

$$\mathcal{G}_{\text{static}} = \big\langle \mathcal{X},\ \ell,\ (\mathbf{access}^i)_{i \in \mathcal{M}},\ (\mathbf{vis}^i)_{i \in \mathcal{M}} \big\rangle.$$

The dynamic environment state at time $k$ is:

$$\mathcal{G}_k = \big\langle \mathcal{G}_{\text{static}},\ F_k,\ S_k,\ \mathcal{T}_k,\ e_k \big\rangle,$$

where $e_k$ is the set of external stimuli broadcast at time $k$, such as evacuation orders, alarms, and social warnings.

### B.2 AGENTS

#### B.2.1 PERSONA

Each agent $i \in \mathcal{M}$ is characterized by a **persona** $\psi^i$, a fixed attribute vector that conditions all aspects of perception, memory, and decision-making throughout the simulation:

$$\psi^i = \big(\psi^i_{\text{att}},\ \psi^i_{\text{cog}},\ \psi^i_{\text{soc}},\ \psi^i_{\text{spat}}\big).$$

- $\psi^i_{\text{att}}$: demographic attributes (age, occupation, housing type, number of dependents).
- $\psi^i_{\text{cog}} = (\rho^i, \tau^i, \theta^i)$: cognitive parameters: risk perception $\rho^i \in [0,1]$, trust in alerts $\tau^i \in [0,1]$, and threat assessment sensitivity $\theta^i \in [0,1]$, grounded in the Protective Action Decision Model (PADM).
- $\psi^i_{\text{soc}}$: social attributes, including ties to other agents (family, neighbors, coworkers) and susceptibility to social influence $\eta^i \in [0,1]$.
- $\psi^i_{\text{spat}}$: spatial knowledge, encoding familiar routes and salient locations within $\mathcal{X}$.

### B.3 OBSERVATION

At each time step $k$, agent $i$ does not have access to the full environment state $\mathcal{G}_k$. Instead, it receives a local observation determined by its current position $c^i_k \in \mathcal{X}$ and its field of view.

**Field of view.** The observable set of agent $i$ at time $k$ is

$$\mathcal{V}^i_k = \big\{\, x \in \mathcal{X} : \|x - c^i_k\|_\infty \le \phi^i_k \ \wedge\ \mathbf{vis}^i(x) = 1 \big\},$$

where $\phi^i_k \in \mathbb{Z}_{>0}$ is the effective field-of-view radius at time $k$. Smoke reduces observability: letting $\phi^i_{\text{base}}$ denote the baseline radius (a function of $\psi^i$) and $\Delta\phi > 0$ a fixed reduction,

$$\phi^i_k = \begin{cases} \phi^i_{\text{base}} - \Delta\phi & \text{if } c^i_k \in S_k, \\ \phi^i_{\text{base}} & \text{otherwise.} \end{cases}$$

**Local observation.** The local observation of agent $i$ at time $k$ is the tuple

$$o^i_k = \Big(c^i_k,\ \mathcal{G}_k\big|_{\mathcal{V}^i_k},\ e_k\Big),$$

where $\mathcal{G}_k|_{\mathcal{V}^i_k}$ denotes the restriction of the dynamic environment state to the observable cells $\mathcal{V}^i_k$, and $e_k$ is the set of stimuli broadcast globally at time $k$.

#### B.3.1 MEMORY

Agent $i$ maintains a memory state $m^i_k$ that is updated at each time step upon receiving a new observation:

$$m^i_k = \text{Update}(m^i_{k-1},\ o^i_k).$$

The memory state is a structured record comprising:

- **Spatial memory** $m^i_{\text{spat},k}$: the cumulative map of observed cells $\bigcup_{k' \le k} \mathcal{V}^i_{k'}$, annotated with cell types, known routes, and locations of salient destinations (home, dependents, safe zones).
- **Hazard history** $m^i_{\text{haz},k}$: the record of observed fire, smoke, and traffic configurations up to time $k$.
- **Goal history** $m^i_{\text{goal},k}$: the sequence of high-level goals pursued by agent $i$ up to time $k$.
- **Decision history** $m^i_{\text{dec},k}$: the sequence of actions taken and their observed outcomes.

### B.4 AGENT STATE

The full state of agent $i$ at time $k$ is:

$$s_k^i = \big(c_k^i,\ o_k^i,\ m_k^i,\ \psi^i\big).$$

Since $\psi^i$ is fixed and $o_k^i$ is determined by $c_k^i$ and $\mathcal{G}_k$, the agent state is fully characterized by the triple $(c_k^i, m_k^i, \psi^i)$. The agent's behavior at each time step is a deterministic or stochastic function of $s_k^i$.

#### B.4.1 COGNITIVE MODULE

The cognitive module governs how agent $i$ maps its current state $s_k^i$ to a decision. It is structured around two components: an **urgency assessment** and a **hierarchical planner**.

**Urgency.** At each time step $k$, the agent computes an urgency score $u_k^i \in [0, 1]$ that quantifies the perceived need for immediate protective action:

$$u_k^i = f_{\text{cog}}\big(o_k^i,\ m_k^i;\ \psi_{\text{cog}}^i\big),$$

where $f_{\text{cog}}$ integrates current observations (proximity to fire and smoke, traffic density, incoming stimuli $e_k$) with memory and persona-conditioned cognitive parameters $(\rho^i, \tau^i, \theta^i)$. Concretely, $u_k^i$ increases with hazard proximity, alarm severity weighted by $\tau^i$, and risk perception $\rho^i$, and is modulated by threat assessment sensitivity $\theta^i$. In our framework, $f_{\text{cog}}$ is implemented via LLM reasoning conditioned on $\psi^i$.

**Hierarchical planner.** The agent's decision-making (or policy) $\pi^i$ is decomposed in three levels of planning:

$$\pi^i = \big(\pi_H^i,\ \pi_M^i,\ \pi_L^i\big).$$

- **High-level planner** $\pi_H^i$: activated when $u_k^i \geq \theta_H^i$ or upon receipt of a high-severity stimulus $e_k$. Produces a high-level goal

  $$g_k^i = \pi_H^i\big(s_k^i,\ u_k^i\big) \in \mathcal{G}_H,$$

  where $\mathcal{G}_H = \{\texttt{stay}, \texttt{evacuate}, \texttt{go-home}, \texttt{gather-dependents}, \texttt{help-neighbor}, \ldots\}$ is the discrete set of high-level intents.
- **Mid-level planner** $\pi_M^i$: translates the active goal $g_k^i$ into a route decision

  $$r_k^i = \pi_M^i\big(g_k^i,\ o_k^i,\ m_k^i;\ \psi^i\big) \in \mathcal{R}_k^i,$$

  where $\mathcal{R}_k^i$ is the set of feasible routes given the agent's current spatial memory and accessibility constraints. Route selection reflects persona-specific behavioral tendencies such as familiarity bias and congestion sensitivity, grounded in empirical evacuation data.
- **Low-level planner** $\pi_L^i$: executes the current route $r_k^i$ over a short horizon $K \in \mathbb{Z}_{>0}$ within the observable set $\mathcal{V}_k^i$, via shortest-path navigation:

  $$(a_k^i, \ldots, a_{k+K}^i) = \pi_L^i\big(r_k^i,\ \mathcal{V}_k^i\big),$$

  where each $a_k^i \in \mathcal{A} = \{\texttt{N},\ \texttt{S},\ \texttt{E},\ \texttt{W},\ \texttt{stay},\ \texttt{interact}\}$ is a primitive action.

**Cognitive loop.** At each time step $k$, agent $i$ executes the following sequential cycle:

$$o_k^i \xrightarrow{\text{perceive}} m_k^i \xrightarrow{\text{assess}} u_k^i \xrightarrow{\text{plan}} (g_k^i, r_k^i) \xrightarrow{\text{navigate}} a_k^i \xrightarrow{\text{act}} c_{k+1}^i.$$

This cycle repeats at every time step, enabling continuous adaptation of behavior as the disaster evolves.

## C LLM USAGE STATEMENT

We used `gpt-4.1-mini` for the agent action calls through the OpenAI API. We also used `ChatGPT 5.1` to assist in generating the environment JSON files, structuring agent persona files, and grammar/sentence structure assistance in writing.