

# Yanapay: A Simulation Toolkit for Autonomous Agents Reasoning about Human Behaviour in Emergency Evacuations

Demonstration Track

Alexandros Kangkelidis  
The Open University  
Milton Keynes, United Kingdom

Carlos Gavidia-Calderon  
The Alan Turing Institute, UK  
London, United Kingdom

Bashar Nuseibeh  
The Open University  
Milton Keynes, United Kingdom

Ayodeji Abioye  
The Open University  
Milton Keynes, United Kingdom

Amel Bennaceur  
The Open University  
Milton Keynes, United Kingdom

## ABSTRACT

Deploying autonomous agents in safety critical domains, such as emergency evacuations, requires them to reason about and predict the often uncertain behaviour of humans. This paper presents Yanapay, a simulation toolkit designed to evaluate the decision-making strategies of autonomous Search and Rescue (SAR) agents operating in a transportation hub. Yanapay applies social psychology principles to simulate complex human behaviours, such as group dynamics and social identity. The toolkit enables researchers to design and benchmark agent architectures that reason about these social factors and predict human intent, for instance, which civilians are likely to aid in evacuation efforts. Yanapay automates the simulation workflow, allowing researchers to systematically evaluate and compare the performance of different strategies across diverse scenarios. A video demonstrating the tool is available at <https://youtu.be/erLxsmFtHzs>.

## KEYWORDS

Agent-Based Simulation; Emergency Evacuation

### ACM Reference Format:

Alexandros Kangkelidis, Carlos Gavidia-Calderon, Bashar Nuseibeh, Ayodeji Abioye, and Amel Bennaceur. 2026. Yanapay: A Simulation Toolkit for Autonomous Agents Reasoning about Human Behaviour in Emergency Evacuations: Demonstration Track. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, Paphos, Cyprus, May 25 – 29, 2026, IFAAMAS, 3 pages. <https://doi.org/10.65109/YQLN3900>

## 1 INTRODUCTION

Emergency evacuations can be modelled as complex multi-agent environments where human behaviour is uncertain and difficult to predict. Autonomous systems have increasingly been used in these settings to assist in evacuation efforts [2]. By collaborating with humans on the scene, these systems can significantly enhance the speed and safety of evacuations [15]. However, the development of such systems presents challenges related to modelling and reasoning about collaboration with humans [7].

These systems must be rigorously tested across diverse scenarios while accounting for complex human behavioural factors. While existing simulation frameworks focus either on technical system validation or human behaviour modelling [3, 14], there is a gap in tools that support the systematic evaluation of socially-aware decision making. Agent-based simulations can help address the complexities of modelling human behaviour [11]. However, the process of setting up these simulations, running experiments across diverse conditions, and analysing the results can be time-consuming [12].

To address this gap, we present Yanapay<sup>1</sup>, an agent-based simulation toolkit<sup>2</sup> designed to support researchers in evaluating autonomous agents in emergency evacuations of transportation hubs. Yanapay enables researchers to design and benchmark socially-aware Search and Rescue (SAR) agents that can reason about human intent and can collaborate with them. For instance, agents can predict which civilians are likely to act as ‘zero-responders’ (survivors who mobilise to help others) and task them to assist fallen victims (incapacitated by injury). The toolkit automates the simulation workflow to systematically evaluate and compare strategies under different scenarios.

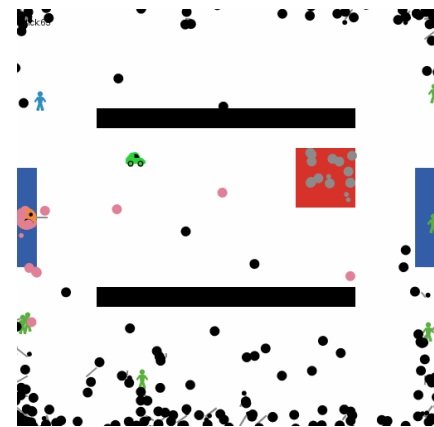



Figure 1: Snapshot of the simulation environment depicting a SAR robot (green) navigating a transportation hub during an evacuation. The robot identifies a fallen agent (orange) and recruits a zero-responder (blue) for assistance.

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

*Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026)*, C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems ([www.ifaaamas.org](http://www.ifaaamas.org)). <https://doi.org/10.65109/YQLN3900>

<sup>1</sup>Yanapay’ means ‘to help’ in Quechua, an indigenous language of Peru

<sup>2</sup>The Yanapay source code can be downloaded from the GitHub repository at <https://github.com/kangkelidis/robot-assisted-evacuation>

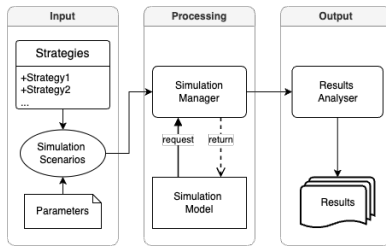


Figure 2: Diagram of the main entities of the Yanapay toolkit.

## 2 RELATED WORK

Simulations are extensively used as a research tool in software engineering [4]. Due to the nature of emergencies, using simulations allows us to explore multiple configurations and scenarios safely, ethically, and in a cost-effective way [13]. On the one hand, simulation tools for emergencies using autonomous systems, e.g., Dronology [3] focus on simulating the software system itself or a single human participant rather than group behaviour. On the other hand, social psychology researchers use agent-based models to capture people’s behaviour [14] but without representing autonomous systems. Mass casualty incident simulators primarily address medical or operational aspects, such as triage [5], rather than focusing on Human-AI collaboration. Yanapay combines these streams by introducing a platform that integrates the socio-cultural dynamics of human crowds with the decision-making strategies of SAR robots, specifically, modelling how robots can recruit civilian zero-responders to assist in the evacuation of a transportation hub.

## 3 THE YANAPAY TOOLKIT

Yanapay builds upon the IMPACT+ model, an extension of the IMPACT model [14], which was engineered by domain experts in social psychology and validated against data from real-world evacuation drills to simulate human behaviour in a transportation hub during an emergency evacuation [8]. Yanapay models agents with internal cognitive states derived from social psychology. Agents maintain a belief state representing their perception of danger; environmental or social cues trigger a desire to evacuate. If they do not decide to evacuate, agents move randomly through the space. Agents are assigned demographic attributes, specifically age, gender, cultural group, familiarity, and compliance. Decisions are driven by group leaders who guide their members; once a leader decides to evacuate, the entire group follows. This creates realistic bottlenecks as members seek each other before exiting. High agent concentration increases the likelihood of accidents, causing agents to fall. This state persists for a set duration, which is shortened if assistance is provided by distinct agent types: group leaders (who guide agents), first-responders (who trigger evacuations), or zero-responders (civilians who assist victims based on shared social identity). The SAR robot aims to leverage this behaviour to speed up evacuation times.

Figure 1 depicts the SAR locating and assisting fallen agents. Operating distinctly from civilian group leaders, it must decide whether to request help from a zero-responder or from a first-responder. This allows researchers to plug in and compare different prediction models without modifying the underlying simulation

engine. For instance, researchers can implement strategies based on social psychology, where the robot evaluates shared identity markers, subject to privacy safeguards in real-world deployments, to predict if a civilian is likely to comply with a request for help [6].

As seen in Figure 2, Yanapay automates the generation and parallel execution of simulations based on user-defined scenarios (e.g., varying crowd and robot density) and SAR robot strategies. The Simulation Manager orchestrates experiments, using PyNetLogo [10] and Flask [9] to bridge the Simulation Model (implemented in NetLogo [1]) with Python-based strategy controllers. This design allows researchers to define SAR behaviours in Python while using NetLogo’s capabilities. Upon termination, the Results Analyser processes the results of the experiments to generate statistical comparisons and visualisations of the strategies’ performance, as illustrated by the evacuation time distributions in Figure 3.

The demonstration presents a scenario where a configurable number of SAR robots are deployed to assist scarce first-responders, aiming to speed up the evacuation of a transportation hub. Yanapay allows researchers to test various decision-making strategies to optimise the autonomous agent’s response to unpredictable human behaviour. If the robot incorrectly tasks a civilian who refuses to help, the evacuation is delayed due to the wasted attempt and the subsequent delay in alerting a first-responder. Conversely, if the robot relies too heavily on first-responders, the evacuation time increases because professionals are a scarce resource and the potential of zero-responders remains underutilised.

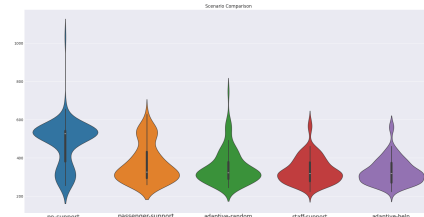


Figure 3: Predefined scenarios showing evacuation times. Rank by median values, from left to right: no SAR robot, always asking for help, random decision, always calling for help, predicts the chance of a passenger accepting to help.

## 4 CONCLUSION AND FUTURE WORK

In this paper, we introduced Yanapay, a toolkit providing a standardised environment to benchmark autonomous decision-making in emergency evacuations of transportation hubs. Yanapay addresses the challenge of simulating human behaviour in environments where SAR robots must reason about social dynamics to minimise evacuation times. The toolkit automates the workflow for setting up, executing, and analysing simulations, enabling researchers to systematically benchmark strategies for the SAR robots.

The current implementation is constrained by the simplifications of human behaviour inherent in the IMPACT model. Additionally, the robot’s current decision logic considers only the nearest candidate; future iterations will implement a prioritised ranking system. We also aim to refine the model to simulate the complex crowd dynamics where fallen agents act as obstacles that may reduce congestion at exits, affecting total evacuation times.

## ACKNOWLEDGMENTS

This work was supported by EPSRC grants EP/V026747/1 and EP/R013144/1.

## REFERENCES

- [1] [n.d.]. Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>, Northwestern University, Evanston, IL.
- [2] Andreas Birk and Stefano Carpin. 2006. Rescue robotics—a crucial milestone on the road to autonomous systems. *Advanced Robotics* 20, 5 (Jan. 2006), 595–605. <https://doi.org/10.1163/156855306776985577>
- [3] Jane Cleland-Huang, Michael Vierhauser, and Sean Bayley. 2018. Dronology: an incubator for cyber-physical systems research. In *ICSE*. ACM, 109–112.
- [4] Breno Bernard Nicolau de França and Nauman Bin Ali. 2020. The Role of Simulation-Based Studies in Software Engineering Research. In *Contemporary Empirical Methods in Soft. Eng.* Springer, 263–287.
- [5] Ruben De Rouck, Michel Debacker, Ives Hubloue, Selma Koghee, Filip Van Uterbeeck, and Erwin Dhondt. 2018. SIMEDIS 2.0 : ON THE ROAD TOWARD A COMPREHENSIVE MASS CASUALTY INCIDENT MEDICAL MANAGEMENT SIMULATOR. In *2018 Winter Simulation Conference (WSC)*. 2713–2724. <https://doi.org/10.1109/WSC.2018.8632369>
- [6] John Drury, Chris Cocking, and Steve Reicher. 2009. Everyone for themselves? A comparative study of crowd solidarity among emergency survivors. *British Journal of Social Psychology* 48, 3 (Sept. 2009), 487–506. <https://doi.org/10.1348/014466608X357893>
- [7] Carlos Gavidia-Calderon, Amel Bennaceur, Anastasia Kordoni, Mark Levine, and Bashar Nuseibeh. 2022. What do you want from me? adapting systems to the uncertainty of human preferences. In *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: New Ideas and Emerging Results* (Pittsburgh, Pennsylvania) (ICSE-NIER '22). Association for Computing Machinery, New York, NY, USA, 126–130. <https://doi.org/10.1145/3510455.3512791>
- [8] Carlos Gavidia-Calderon, Anastasia Kordoni, Amel Bennaceur, Mark Levine, and Bashar Nuseibeh. 2024. The IDEA of Us: An Identity-Aware Architecture for Autonomous Systems. *ACM Trans. on Soft. Eng. and Meth.* (March 2024), 3654439. <https://doi.org/10.1145/3654439>
- [9] Miguel Grinberg. 2018. *Flask web development: developing web applications with python*. " O'Reilly Media, Inc".
- [10] Marc Jaxa-Rozen and Jan H. Kwakkel. 2018. PyNetLogo: Linking NetLogo with Python. *Journal of Artificial Societies and Social Simulation* 21, 2 (2018), 4. <https://doi.org/10.18564/jasss.3668>
- [11] Charles Macal and Michael North. 2014. Introductory tutorial: Agent-based modeling and simulation. In *Proc. of the Winter Simulation Conf.* IEEE, Savannah, GA, USA, 6–20. <https://doi.org/10.1109/WSC.2014.7019874>
- [12] Tong Niu, Weihao Zhang, and Rong Zhao. 2024. Solution-oriented Agent-based Models Generation with Verifier-assisted Iterative In-context Learning. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems* (Auckland, New Zealand) (AAMAS '24). International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 1473–1481.
- [13] Salil Purandare, Urjoshi Sinha, Md Nafee Al Islam, Jane Cleland-Huang, and Myra B. Cohen. 2023. Self-Adaptive Mechanisms for Misconfigurations in Small Uncrewed Aerial Systems. In (SEAMS). 169–180. <https://doi.org/10.1109/SEAMS59076.2023.00030>
- [14] C. Natalie Van Der Wal, Daniel Formolo, Mark A. Robinson, Michael Minkov, and Tibor Bosse. 2017. Simulating Crowd Evacuation with Socio-Cultural, Cognitive, and Emotional Elements. In *Transactions on Computational Collective Intelligence XXVII*, Jacek Mercik (Ed.). Vol. 10480. Springer International Publishing, Cham, 139–177. [https://doi.org/10.1007/978-3-319-70647-4\\_11](https://doi.org/10.1007/978-3-319-70647-4_11) Series Title: Lecture Notes in Computer Science.
- [15] Christopher Van Tilburg. 2017. First Report of Using Portable Unmanned Aircraft Systems (Drones) for Search and Rescue. *Wilderness & Environmental Medicine* 28, 2 (June 2017), 116–118. <https://doi.org/10.1016/j.wem.2016.12.010>