Planning AI Assistant for Emergency Decision-Making (PlanAID): Framing Planning Problems and Assessing Plans with Large Language Models

Jacqueline Lee, Michelle Cantu, Joel Korb, Eva Meth, John Griffith, Joanna Korman, Anna Yuen, Peter Schwartz, Abigail Gertner
The MITRE Corporation

Abstract

This paper proposes the use of agentic artificial intelligence (AI) to enhance emergency operations planning in response to the escalating frequency and complexity of health emergency response. We present Planning AI Assistant for Emergency Decision-Making (PlanAID), an AI planning assistant that combines a large language model (LLM) with a symbolic planner to improve public health preparedness. Like other recent planning tools, ours uses an LLM to translate a planning problem from natural language into a Planning Domain Definition Language (PDDL) specification to be solved by a symbolic planner. We extend this approach by using an LLM with a chat interface to actively help the user 1) frame the relevant components of the planning problem, 2) assess the plan provided by the symbolic planner, and 3) reframe the problem based on issues identified during the plan assessment. This integration allows for the generation of contextually appropriate plans by leveraging the strengths of both LLMs and symbolic planners. Our tool is built on a highly configurable backend that allows an administrator to tailor it for a specific team or incident type by specifying the relevant documents, data sources, plan components, and prompts before an emergency occurs. We also posit evaluation metrics focused on the system's ability to produce effective, resource-optimized plans and its adaptability to complex domains. Our research demonstrates the potential of AI technologies in emergency operations planning, offering a robust solution for improving public health resilience. Future work will aim to further validate the PDDL outputs and assess the system's practical utility in real-world scenarios, contributing to more effective human-machine teaming and planning processes.

Introduction

The landscape of public health preparedness is evolving rapidly, driven by the increasing frequency and severity of health emergencies from pandemics to natural disasters (Samet and Brownson 2024, Rudolph, et al. 2018). In recent years, the field of public health has faced increasingly complex challenges, necessitating innovative approaches to emergency planning and response. One notable example is the COVID-19 pandemic, which highlighted many difficulties facing the public health workforce (i.e., increasing and additional demands and staffing limitations) which significantly limit health response efforts (Schoch-Spana, et al. 2018, Leider, et al. 2023, Ravenhall, et al. 2021).

For emergencies such as this, one of the most significant pieces of planning guidance is the Federal Emergency Management Agency's (FEMA) Comprehensive Preparedness Guide (CPG) 101: Developing and Maintaining Emergency Operations Plans (Fagel, Mathews and Murphy 2022). Alongside its key principles for emergency operations, the CPG 101 emphasizes that plan creation, implementation and improvement require both following best practice doctrine and adapting to different stakeholders and unexpected real-life complications. For instance, while coordination is a best practice, the fragmented nature of the American emergency management system makes any "whole of government" planning especially challenging (Wolf-Fordham 2020).

As emergency operations planning remains difficult to standardize, models and guidance have historically fallen short, with lengthy, cumbersome emergency operations plans (EOPs) that are misaligned from the work as actually done, indicating a gap between decision makers and decision implementers (Keim, 2013, de Carvalho, et al. 2018, Son, et al. 2020). Given the numerous complex requirements for this field, many have long considered new technologies to improve the speed and efficacy of emergency response, including but not limited to the Internet of Things (Damaševicius, Bacanin, and Misra, 2023), cloud-based collaborative platforms (Gupta, et al. 2022), and AI (Sun, Bocchini and Davison 2020). With such optimism on technology integration, there remains a critical call to keep humans at the center of making these nuanced decisions with effective human-machine teaming.

In this paper, we present our progress in developing PlanAID, an agentic AI assistant that combines the benefits of LLMs and symbolic planners for emergency planning. Public health preparedness coordinators will be supported and guided by the assistant as they frame, formulate, solve, and assess emergency planning problems. With this flexible workflow and planning framework, we aim to improve both incident-specific plans, broader public health preparedness planning and overall public health resilience.

Background

Recent advances in AI present a transformative opportunity to enhance public health preparedness and resilience with enhanced planning, optimized resource allocation, and faster decision-making in the face of emergencies such as pandemics or natural disasters. Even so, to date, few public health departments have reported incorporating AI into their practice, suggesting a perceived lack of reliability of data and model outputs (Patel, et al. 2024). While some AI use cases in public health have been documented, especially during the COVID-19 response (Bharel, et al. 2024), few studies have shown the intersection of AI, public health preparedness, and emergency operations improvement planning. Alongside reliability concerns, this may also be due to technical challenges such as resource constraints that prevent the wider development, adoption, and sustainment of AI systems in this field. However, there is now much greater interest in leveraging these technologies, particularly with the recent surge in popularity and impressive progress of LLMs and their ability to support natural language interaction and agentic AI as general-purpose tools.

Symbolic Planning

Within the field of AI, symbolic planning develops plans based on a symbolic model of the world and actions that can be taken to change the state to achieve goals (Behnke 2024). Symbolic planners have long been used as a systematic, verifiable approach for robotic agents operating in specific environments with impressive results in continuous and discretized spaces for both low-level and high-level planning tasks (Garrett, Lozana-Perez and Kaelbling 2020, Konidaris, Kaelbling and Lozano-Perez 2018). Traditionally, the field of symbolic planning has focused on classical planning for simpler, deterministic problems, whereas application support for probabilistic planning and partial observability is limited (Nunez-Molina, Mesejo and Fernandez-Olivares 2024). This is due in no small part to the need for expertise in both symbolic planning and the operational domain in order to design an adequate planning domain description.

While symbolic planners use heuristics to accelerate state space search, they can still take a long time to produce a plan. This and the lack of assurance on the completeness and correctness of the problem specification significantly limit the adoption of these solvers in real-life applications (Nunez-Molina, Mesejo and Fernandez-Olivares 2024, Kambhampati 2007, Gragera and Pozanco 2023). Even if the planning domain is appropriate and the plan is formally correct, the resulting plan could still fail during execution if the real world does not behave as expected (Liu, et al. 2023).

Large Language Model Planning

Recent research has started to explore the use of LLMs to assist with planning (Valmeekam, et al. 2023, Liu, et al. 2023, Song, et al. 2023). Initial attempts to simply ask an LLM to produce a plan demonstrated that the plans were often invalid, either not achieving the desired goal or impossible to execute (Valmeekam, Marquez, Sreedharan, & Kambhampati, 2023). Researchers have found success by combining LLMs with symbolic planners at the outset of plan creation (Liu, et al. 2023, Song, et al. 2023) and for error handling (Chen, et al. 2024). In the former case, given a natural language description of a planning problem, an LLM can translate it into a formal planning representation such as Planning Domain Definition Language (PDDL) (McDermott, et al. 1998). A symbolic planner can then solve the planning problem, and the LLM can translate the output back into natural language.

Several human-in-the-loop planning frameworks integrating LLMs and symbolic planners have been recently proposed. Themes such as iterative refinement and intermediate representations have emerged, yielding promising results; however, these works use pre-determined PDDL domains (Agarwal and Sreepathy 2024, Zhou, et al. 2024). Initial attempts to create PDDL domain specifications using LLMs have thus far been largely limited to simple toy problems (Gestrin, Kuhlmann and Seipp 2024). A more complex domain involving emergencies is seen in the work done by Tang, Ni and Zhou (2020) and Yang and Liang (2023), but their applications are slightly different. The former work focuses on storyline generation after a disaster while the latter aims to solve scheduling problems for emergency material delivery. In both, little is proposed on evaluating outputted PDDL outside of manual review.

Though the combination of an LLM and a symbolic planner is promising, the technology does not yet provide the assurances required for many government applications. LLMs can misinterpret the user's description of the situation or choose the wrong details to include in the formal planning domain. While some researchers have started to investigate the use of LLMs to inform mission planning, these methods still rely on humans to generate and validate the final plans (Jensen and Tadross 2023).

Agentic AI

Agentic AI is an emerging concept that allows LLMs to perform complex, multi-step tasks through their use of tools, reflection, planning, and multi-agent collaboration (Chawla, et al. 2024). Reasoning in LLMs is typically managed by prompting the model to reflect on its output using a structured prompting procedure (Yao, et al. 2023, Wei, et al.

2024, Shin, et al. 2024) . The ability to reason is particularly useful in a human-machine teaming scenario where the LLM has information that the human does not about how to accomplish parts of the shared task.

Methodology

To support complex real-world domains like emergency operations planning, we have begun to investigate an agentic AI approach that combines an LLM with a symbolic planner. The system architecture implemented in PlanAID is shown in Figure 1. Prior to an emergency incident, users can configure the tool to access relevant documents and data sources. When an incident occurs, the agentic workflow facilitates an iterative process that includes four phases. First, the user and the LLM work together to properly frame the planning problem, identifying the relevant stakeholders, issues, and goals. The user can prompt the LLM to request relevant information from the documents it can access, but the LLM can also prompt the user to request necessary information that is missing from those documents. The LLM then formulates a PDDL specification based on this problem framing. Next, a symbolic planner solves the planning problem to produce a valid and complete plan. Finally, the user and the LLM work together to assess the plan based on factors such as correctness, risks, and level of detail. The user can optionally repeat this process by reframing the planning process based on the outcomes of previous iterations.

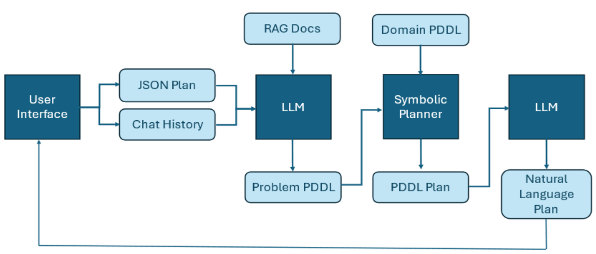

Figure 1: System Architecture Diagram

System Architecture

To maintain flexibility within PlanAID, we employ a configurable architecture. A domain-specific JSON schema file defines the essential components of the plan, such as resources, issues, and stakeholders, as well as their structure. This ensures that the planning process is tailored to the specific needs and constraints of the domain. Furthermore, the system allows for the upload of domain-specific documents, which serve two purposes: they provide contextual information for the LLM to reference via Retrieval-Augmented Generation (RAG) for plan generation and translation (Gao, et al., 2023), and they are available for users to consult during planning.

We use a simple database to manage plans in multiple formats. Initially, plans are stored in a structured, configurable JSON format, captured in the components field. This format not only specifies the required components of a plan for a given domain but also provides the user interface with a flexible framework to guide users through each component step by step. The LLM leverages this JSON format, along with a prompt, to generate a problem PDDL to send to the symbolic planner. The planner's output, which is stored in the PDDL field, is translated by the LLM into a final natural language plan, recorded in the final plan field, which serves as the primary deliverable of the system. Additionally, the database archives the complete chat history between the user and the LLM in the chat messages field, preserving context to enhance the LLM's understanding of user intent.

PDDL Domain and Problem Specification

Given the plan components, the LLM can begin to generate the proper inputs for the symbolic planner – namely the PDDL domain and problem specification files. This is done with an ordered sequence of prompts designed to extract key parts of the PDDL domain, which we have limited for now to object types, predicates, and actions. First, the relevant object types are generated based on the described emergency so that the basic types and subtypes are made explicit. With these types as additional input, the LLM reasons what predicates may be important to describe the relevant state of the world during the emergency. Finally, considering these object types and predicates, the LLM can produce a set of actions that affect the world state, making it possible to move from the initial state to a goal state. Given the complex nature of how actions in emergency operations planning may have various short- and long-term effects, the LLM is additionally advised to constrain the number of effects and to break down higher-level actions.

In this manner, the object types, predicates, and actions can be combined into a cohesive PDDL domain file. With this domain and its understanding of the emergency as learned from its chat interactions with the user and the verified plan components from the Planner Pane, the LLM is finally prompted to generate a problem specification file. At this point, the symbolic planner can be used to solve the PDDL problem and find a contextually appropriate and relevant plan for the emergency.

Graphical User Interface (GUI) and Workflow

A user interacts with PlanAID via its graphical user interface (GUI), which enables them to easily and intuitively interact with the LLM to generate an EOP. Its three panes – as shown in Appendix Figures A1 and A2 – allow the user to manage multiple incidents, chat with the LLM, and determine plan components. The Navigation Bar on the left allows the user to add, remove, and rename incidents. The Chat Pane in the center allows the user to chat directly with the LLM or select from prompts suggested by the tool. Lastly, in the Planner Pane on the right, the user can upload relevant documents, review and modify information the LLM has collected (i.e., plan components), and review plans suggested by the tool.

While the PlanAID system is flexible to what users may find to be most helpful for their planning task, we present a sample user workflow for demonstration purposes. A user begins the process by creating a new incident in the Navigation Bar and importing relevant documents and data sources in the Documents tab of the Planner Pane. The user can then provide PlanAID with additional context about the incident through the chat interface in the Chat Pane. When the user determines that the agent has enough information, the user can ask it to suggest key components of the plan (e.g., stakeholders, issues, and goals). The user will then review and copy that information into a form in the Planner tab of the Planner Pane, editing as necessary. The user can return to chatting with the agent and ask for clarifications or edits. When the user confirms that all information is complete and correct, the user can click a button to generate the plan.

The LLM will work with the symbolic planner to produce a plan based on the plan components as specified in the Planner Pane planner form and user-provided context extracted from the chat conversation. The user will then read through the plan and if edits are needed, they can ask the agent to refine the plan or edit the plan directly themselves. The agent can also review the plan and suggest potential issues or risks to the user. Once satisfied, the user can export the EOP document for use outside of PlanAID.

Measures of Effectiveness and Performance

This prototype tool has not been fully tested yet on real plans, but the overall effectiveness of PlanAID will be assessed by its ability to produce EOPs that are more successful than traditional methods and that improve public health preparedness and resiliency. Desired outcomes include plans that enhance public safety, confidence, and well-being, are easier and faster to implement, and require fewer resources. Human expert evaluations will compare AI-generated plans to traditional ones, focusing on novel steps, plan brevity and level of granularity, and the identification of dependencies between steps. Additionally, the plans' generalizability to a broader range of incidents and their ability to improve coordination and communication with other agencies and the public will be key indicators of effectiveness.

Performance will be measured through key performance indicators such as the time taken to generate plans, resource optimization, and the system's adaptability to complex domains. The user experience with the GUI and the accuracy of the LLM's outputs will also be evaluated. For this evaluation, experiments involving different forms of the PlanAID system will be conducted as described in Table 1.

Planner	Description
Single-shot	Plan produced in a single iteration of workflow
Iterative	User repeatedly interacts with PlanAID to refine plan
Ablation	Use PlanAID as above without abilities such as: • Reasoning on user's mental state or context • Acting proactively (i.e. limit PDDL actions' individual long-term effects) • Suggesting future action(s) (i.e., explicitly limit length of sequences of actions)

Table 1: Evaluation Experiments

Discussion and Future Research

With the increasing frequency of major health emergencies and the current state of public health preparedness, there is an urgent need for assistive tools to support emergency operations planning. Operators must balance between following general planning guidelines and rapidly adapting to situational realities. LLMs and agentic AI present a user-friendly way of interacting with powerful technologies, which would otherwise require a higher level of technical expertise and more time to customize with important details of the emergency. Our planning tool, PlanAID, uses an LLM to help a user frame the important details of an emergency incident, both being prompted by and offering prompts to the user. It then formulates the situation as a formal planning problem in PDDL so it can be solved by a symbolic planner. Finally, the LLM actively helps the user assess the resulting plan so they can reframe the problem if needed.

Looking forward, we plan to ensure further validation of the generated PDDL domain and problem specification beyond manual review. Additionally, though there has been some initial research on the translation of natural language problems to PDDL, to our knowledge there has been no research verifying the possibility of doing the reverse – converting a PDDL solution to a natural language plan. This remains an interesting question for future research. Lastly, while having readable planning documents would be a significant first step, it is imperative to evaluate the utility of our system for actual users in the emergency planning domain. For this, we would like to conduct the evaluation experiments introduced in the previous section for a real-world emergency problem. We expect that PlanAID's combination of an LLM and a symbolic planner working together in an interactive, configurable agentic AI architecture will significantly improve planning in complex domains like public health preparedness and emergency operations.

Appendix

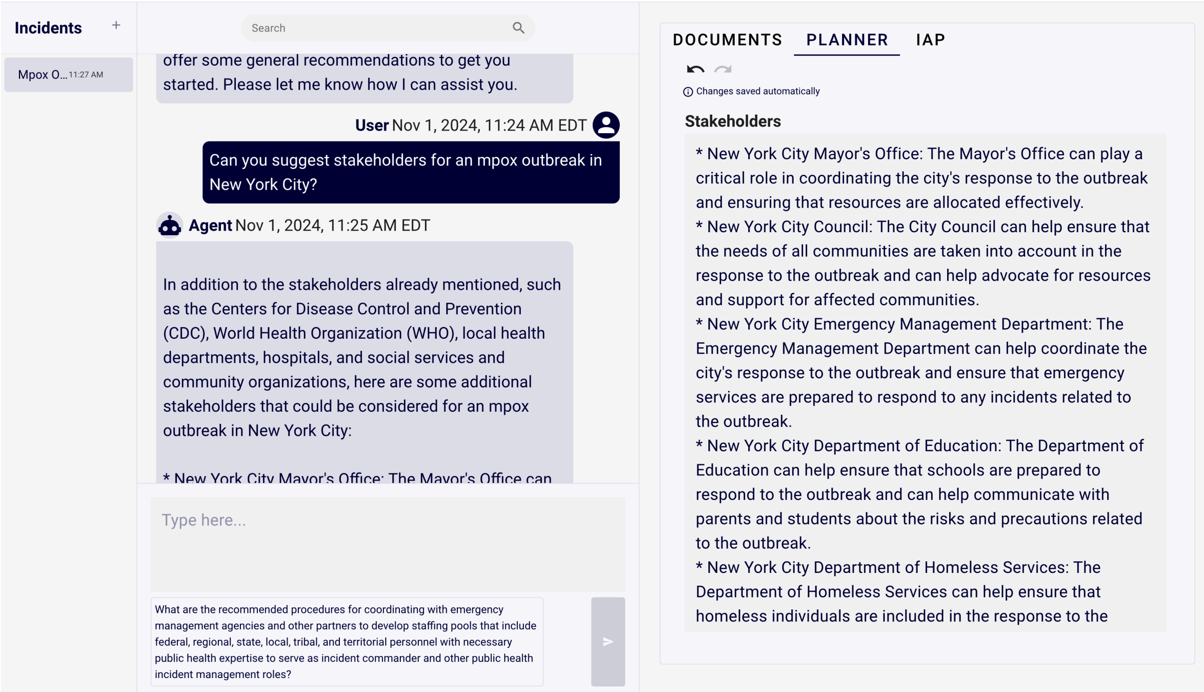

Figure A1: GUI with inputted stakeholders from LLM recommendations

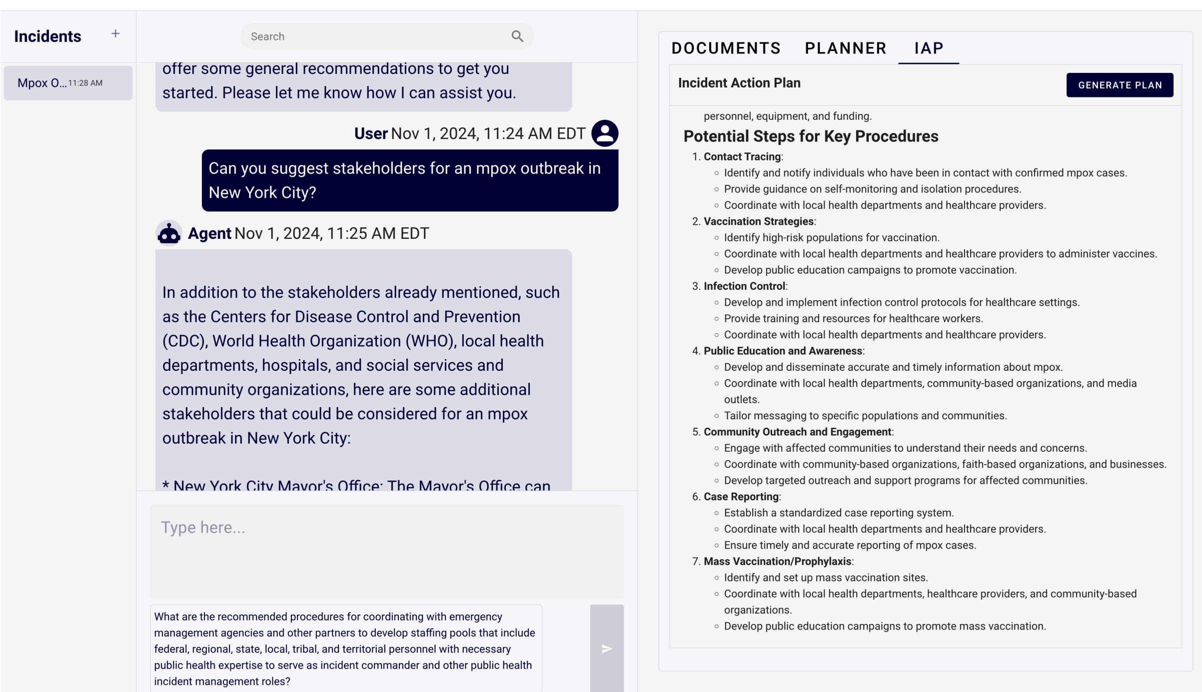

Figure A2: GUI with notional example output incident action plan (IAP)

References

Agarwal, Sudhir, and Anu Sreepathy. 2024. "TIC: Translate-Infer-Compile for Accurate "Text to Plan" Using LLMs and Logical Representations." *Neural-Symbolic Learning and Reasoning: 18th International Conference, NeSy 2024, Barcelona, Spain, September 9–12, 2024, Proceedings, Part II.* 222-244. doi.org/10.48550/arXiv.2402.06608

Behnke, Gregor. 2024. "Symbolic Reasoning Methods for AI Planning." *Proceedings of the AAAI Conference on Artificial Intelligence.* Vancouver, Canada: AAAI Press. doi.org/10.1609/aaai.v38i20.30277

Bharel, Monica, John Auerback, Von Nguyen, and Karen DeSalvo. 2024. "Transforming Public Health Practice With Generative Artificial Intelligence." *Health Affairs* 776-782. doi.org/10.1377/hlthaff.2024.00050

Chawla, Chhavi, Siddharth Chatterjee, Sanketh Gadadinni, Pulkit Verma, and Sourav Banerjee. 2024. "Agentic AI: The building blocks of sophisticated AI business applications." *Journal of AI, Robotics & Workplace Automation.* doi.org/10.69554/XEHZ1946

Chen, Guanqi, Lei Yang, Ruixing Jia, Zhe Hu, Yizhou Chen, Wei Zhang, Wenping Wang, and Jia Pan. 2024. "Language-Augmented Symbolic Planner for Open-World Task Planning." *Proceedings from Robotics: Science and Systems 2024.* Delft, Netherlands. doi.org/10.48550/arXiv.2407.09792

de Carvalho, Paolo Victor R., Angela W. Righi, Gilbert J. Huber, Caio de F. Lemos, Alessandro Jatoba, and José Orlando Gomes. "Reflections on Work as Done (WAD) and Work as Imagined (WAI) in an Emergency Response Organization: A Study on Firefighters Training Exercises." *Applied Ergonomics* 68 (2018): 28-41. doi.org/10.1016/j.apergo.2017.10.016

Damaševičius, Robertas, Nebojsa Bacanin, and Sanjay Misra. 2023. "From Sensors to Safety: Internet of Emergency Services (IoES) for Emergency Response and Disaster Management" *Journal of Sensor and Actuator Networks* 12, no. 3: 41. doi.org/10.3390/jsan12030041

Fagel, Michael J, Rick C Mathews, and J Howard Murphy. 2022. *Principles of Emergency Management and Emergency Operations Centers (EOC).* New York: Taylor & Francis.

FEMA. 2021. "Developing and Maintaining Emergency Operations Plans: Comprehensive Preparedness Guide (CPG) 101." September. Accessed November 2024. https://www.fema.gov/sites/default/files/documents/fema_cpg-101-v3-developing-maintaining-eops.pdf.

FEMA. 2021. "Developing and Maintaining Emergency Operations Plans: Comprehensive Preparedness Guide (CPG) 101." September. Accessed November 2024. https://www.fema.gov/sites/default/files/documents/fema_cpg-101-v3-developing-maintaining-eops.pdf.

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., . . . Wang, H. (2023, December 18). Retrieval-Augmented Generation for Large Language Models: A Survey. Retrieved from https://doi.org/10.48550/arXiv.2312.10997

Garrett, Caelan Reed, Tomas Lozana-Perez, and Leslie Pack Kaelbling. 2020. "PDDLStream: Integrating Symbolic Planners and Blackbox Samplers via Optimistic Adaptive Planning." *International Conference on Automated Planning and Scheduling (ICAPS) 2020.* March 23.

Gestrin, Elliot, Marco Kuhlmann, and Jendrik Seipp. 2024. "NL2Plan: Robust LLM-Driven Planning from Minimal Text Descriptions." *ICAPS 2024 Workshop on Human-Aware and Explainable Planning.* May 7. doi.org/10.48550/arXiv.2405.04215

Gragera, Alba, and Alberto Pozanco. 2023. "Exploring the Limitations of using Large Language Models to Fix Planning Task." *International Conference on Automated Planning and Scheduling 2023.*

Gupta, Shivam, Sachin Modgil, Ajay Kumar, Uthayasankar Sivarajah, and Zahir Irani. 2022. "Artificial intelligence and cloud-based Collaborative Platforms for Managing Disaster, extreme weather and emergency operations ." *International Journal of Production Economics.* doi.org/10.1016/j.ijpe.2022.108642

Jensen, Benjamin, and Dan Tadross. 2023. "How Large-Language Models Can Revolutionize Military Planning." *War On The Rocks.* April 12. Accessed November 6, 2024. https://warontherocks.com/2023/04/how-large-language-models-can-revolutionize-military-planning/.

Kambhampati, Subbarao. 2007. "Model-lite Planning for the Web Age Masses: The Challenges of Planning with Incomplete and Evolving Domain Models." *AAAI'07: Proceedings of the 22nd national conference on Artificial intelligence - Volume 2.* 1601-1604.

Keim, Mark E. 2013. "An innovative approach to capability-based emergency operations planning." *Disaster Health* 53-62. doi.org/10.4161/dish.23480

Konidaris, George, Leslie Pack Kaelbling, and Tomas Lozano-Perez. 2018. "From Skills to Symbols: Learning Symbolic Representations for Abstract High-Level Planning." *Journal of Artificial Intelligence Research* 215-289. doi.org/10.1613/jair.5575

Leider, Jonathon, Mac McCullough, Simone Rauscher Singh, Annie Sieger, Moriah Robbins, Jessica Solomon Fisher, and Paul, Castrucci, Brian Kuehnert. 2023. "Staffing Up and Sustaining the Public Health Workforce." *Journal of Public Health Management & Practice* E100-E107. doi.org/10.1097/PHH.0000000000001614

Liu, Bo, Yuqian Jiang, Xiaohan Zhang, Qiang Liu, Shiqi Zhang, Joydeep Biswas, and Peter Stone. 2023. "LLM+P: Empowering Large Language Models with Optimal Planning Proficiency." September 27. doi.org/10.48550/arXiv.2304.11477

McDermott, Drew, Malik Ghallab, Adele Howe, Craig A Knoblock, Ashwin Ram, Manuela Veloso, Daniel Weld, and David Wilkins. 1998. *PDDL - The Planning Domain Definition Language.* Technical report, Yale Center for Computational Vision and Control.

Nunez-Molina, Carlos, Pablo Mesejo, and Juan Fernandez-Olivares. 2024. "A Review of Symbolic, Subsymbolic and Hybrid Methods for Sequential Decision Making." *ACM Computing Survey.* doi.org/10.1145/3663366

Patel, Krishna, Meghan Shea, Gelila Tamrat, Gwen Davis, David Okereke, Kellie Hall, Timothy McCall, and Aurimar Ayala. 2024. *2024 Public Health Informatics Profile.* Assessment, Washington, DC: National Association of County and City Health Officials.

Ravenhall, Sarah, Nicole Levy, Kathryn Simpson, Molly Fleming, Mayela Arana, Peggy DiManno, Yesenia Grijalva, and Marita Murrman. 2021. "New York State Local Health Department Preparedness for and Response to the COVID-19 Pandemic: An In-Progress Review." *Journal of Public Health Management & Practice* 240-245. doi.org/10.1097/PHH.0000000000001340

Rudolph, Linda, Catherine Harrison, Laura Buckley, and Savannah North. 2018. *Climate Change, Health, and Equity: A Guide for Local Health Departments.* Guide, Oakland, CA; Washington D.C.: Public Health Institute and American Public Health Association.

Samet, Jonathan, and Ross Brownson. 2024. "Reimagining Public Health: Mapping A Path Forward." *Health Affairs* 750-758. doi.org/10.1377/hlthaff.2024.00007

Schoch-Spana, Monica, Jennifer, Ravi, Sanjana Nuzzo, Laura Biesiadecki, and Geoffrey Mwaungulu. 2018. "The local health department mandate and capacity for community engagement in emergency preparedness: a national view over time." *Journal of Public Health Management & Practice* 350-359. doi.org/10.1097/PHH.0000000000000680.

Shin, Hyungyu, Seulgi Choi, Ji Yong Cho, Sahar Admoni, Hyunseung Lim, Taewan Kim, Hwajung Hong, Moontae Lee, and Junho Kim. 2024. "Towards an Evaluation of LLM-Generated Inspiration by Developing and Validating Inspiration Scale." *1st HEAL Workshop at CHI Conference on Human Factors in Computing Systems.* Honolulu, HI, May 12.

Son, Changwon, Farzan Sasangohar, Timothy J. Neville, S. Camille Peres and Jukrin Moon. "Investigating Resilience in Emergency Management: An Integrative Review of Literature." *Applied Ergonomics* 87 (2020): 103114. doi.org/10.1016/j.apergo.2020.103114

Song, Chan Hee, Jiaman Wu, Clayton Washington, Brian M Sadler, Wei-Lun Chao, and Yu Su. 2023. "LLM-Planner: Few-Shot Grounded Planning for Embodied Agents with Large Language Models." *International Conference on Computer Vision.* doi.org/10.48550/arXiv.2212.04088

Sun, Wenjuan, Paolo Bocchini, and Brian D Davison. 2020. "Applications of artificial intelligence for disaster management." *Natural Hazards* 2631-2689. doi.org/ 10.1007/s11069-020-04124-3

Tang, Mingrong, Jinxin Ni, and Qifeng Zhou. 2020. "A PDDL based Disaster Storyline Generation Approach." *2020 39th Chinese Control Conference (CCC).* Shenyang, China: IEEE. 4273-4279. doi.org/10.23919/CCC50068.2020.9188462

Valmeekam, Karthik, Matthew Marquez, Sarath Sreedharan, and Subbarao Kambhampati. 2023. "On the Planning Abilities of Large Language Models (A Critical Investigation with a Proposed Benchmark)." *2023 Conference on Neural Information Processing Systems.* February 13. doi.org/10.48550/arXiv.2302.06706

Wei, Jason, Xuezhi Wang, Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H Chi, Quoc V Le, and Denny Zhou. 2024. "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models." *NIPS'22: Proceedings of the 36th International Conference on Neural Information Processing Systems.* 24824-24837. doi.org/10.48550/arXiv.2201.11903

Wolf-Fordham, Susan. (2020). Integrating Government Silos: Local Emergency Management and Public Health Department Collaboration for Emergency Planning and Response. *The American Review of Public Administration*, *50*(6-7), 560-567. doi.org/10.1177/0275074020943706

Yang, Liping, and Ruishi Liang. 2023. "An AI Planning Approach to Emergency Material Scheduling Using Numerical PDDL." *Proceedings of the 2022 International Conference on Artificial Intelligence, Internet and Digital Economy.* 47-54. doi.org/10.2991/978-94-6463-010-7_7

Yao, Shunyu, Dian Yu, Jeffrey Zhao, Izhak Shafran, Thomas L Griffiths, Yuan Cao, and Karthik Narasimhan. 2023. "Tree of Thoughts: Deliberate Problem Solving with Large Language Models." *37th Conference on Neural Information Processing Systems.* December 3. doi.org/10.48550/arXiv.2305.10601

Zhou, Zhehua, Jiayang Song, Kunpeng Yao, Zhan Shu, and Lei Ma. 2024. "ISR-LLM: Iterative Self-Refined Large Language Model for Long-Horizon Sequential Task Planning." *2024 IEEE International Conference on Robotics and Automation.* Yokohama, Japan, May. doi.org/10.48550/arXiv.2308.13724