CrisisGuideGPT: An Agentic-AI for Crisis Response Planning

Saeed H. Al-Shamrani
Dept. of Computer Science
University of Bisha, Saudi Arabia
University of Alabama at Birmingham
Birmingham, AL 35294-1241
shalsham@uab.edu

Ragib Hasan
Dept. of Computer Science
University of Alabama at Birmingham
Birmingham, AL 35294-1241
ragib@uab.edu

***Abstract*—Natural hazards are continuously posing threats to people's lives. When a crisis happens, people seek reliable and actionable information to ensure their safety. With many sources online, some information can be static and generalized, while others have incomplete or contradictory information. This paper introduces *CrisisGuideGPT*, an agentic Artificial Intelligence (AI) system designed to generate personalized, location-specific crisis response plans. By integrating real-time data from the Federal Emergency Management Agency (FEMA) and the National Oceanic and Atmospheric Administration (NOAA) with established emergency guidelines via Google's Gemini large language model (LLM), CrisisGuideGPT addresses the limitations of static emergency plans. The system employs a multi-stage prompting strategy, including knowledge base grounding, self-correction, and interactive refinement, to produce structured, context-aware guidance for individuals facing crises like hurricanes, floods, wildfires, or earthquakes. We present the system's architecture, methodology, and prompting techniques, and discuss its strengths (e.g., personalization and real-time integration), limitations (e.g., LLM unpredictability and API dependencies), and implications for transparency and accountability in high-stakes AI applications.**

***Index Terms*—Public Safety, Emergency Response, Communication Systems, Disaster Management**

I. Introduction

Natural hazards and emergencies are continuously posing threats to people's lives. In 2023 alone, more than 380 natural disasters occurred globally that affected 153 million people and caused more than $250 billion in damages [8]. Factors such as climate change are increasing the severity and frequency of these disasters around the globe [2]. It is important to always be prepared and have an effective crisis response plan to mitigate the impact of these hazards. Planning can also help in reducing casualties and facilitating recovery. Timely access to accurate and actionable information is therefore critical during emergencies [6].

Traditional crisis response planning is mostly based on static guidelines, predefined protocols, or coming from a single source. While these approaches form a solid foundation in emergency management, they often fail to address the dynamic nature of crises and the specific needs of individuals in different contexts [9]. Several key limitations hinder the effectiveness of conventional emergency planning:

- **Lack of adaptability:** Static plans and predefined guidelines can fail due to the rapid and evolving nature of emergencies. Pre-written guides cannot adapt to changing conditions and emerging threats. Crises are also subject to unexpected developments [6].
- **Information overload and inconsistency:** There is also an information overload during emergencies, which creates a significant barrier to effective response. Individuals, officials, and responders usually obtain information from multiple sources. Some of these multiple sources can have incomplete, vague, or even contradictory information, which hinders the response decisions and possibly leads to misguided actions [9].
- **Absence of personalization:** The lack of personalization in crisis guidance can also reduce its relevance and applicability. Generic recommendations can be brief and straightforward, however, they fail to account for personal circumstances, local conditions, or specific vulnerabilities [5].
- **Limited real-time integration:** Lastly, obtaining real-time data with established guidelines remains a significant technical challenge which exceeds traditional emergency management systems [13].

To address these challenges, we introduce CrisisGuideGPT, an agentic AI system [1] designed to generate personalized, location-aware crisis response plans, by integrating Large Language Models (LLMs), real-time data from authorized sources, and established emergency management guidelines.

The core concept of CrisisGuideGPT is to leverage agentic artificial intelligence to process and synthesize information from multiple sources and provide tailored response plans based on crisis type, location, and individual circumstance [1], [6].

The primary goals of our work include:

- Providing timely, accurate, and personalized crisis response guidelines based on users' location and crisis type.
- Integrating real-time data from authorized sources such as NOAA, FEMA, with their established guidelines.
- Structuring information in a clear, actionable format that facilitates following emergency management best

practices.
- Involving user feedback and clarification to adapt recommendations.
- Maintaining transparency through source attribution and structured output.

Contributions: CrisisGuideGPT is a practical implementation of Agentic AI for crisis response planning. It shows how Large Language Models can effectively generate reliable and structured outputs in high-stakes domains [9]. It is also a novel approach to use real-time data integration, combining current alerts and historical disaster data with established guidelines to create context-aware response plans [13].

Our framework implements an interactive refinement process that allows for personalization based on user-specific factors, overcoming the limitations of static guides and moving beyond one-size-fits-all approaches [5]. It provides a blueprint for similar AI-assisted systems in various high-stakes domains where data synthesis, personalization, and transparency are critical requirements [6].

Organization: This paper is organized as follows: Section II contains background information about public safety, crisis management, Information systems in emergency management, agentic AI, LLMs, and their applications in emergency management. Section III reviews some of the related works. In Section IV, we explain the methodology of our model. We demonstrate the results in Section V. Section VI provides a brief discussion about our framework and its strengths, weaknesses, and use cases. We elaborate on the limitations of our work in Section VII and conclude in Section VIII.

II. Background

A. Public Safety Crises

A crisis is an event that poses a threat to public safety [2]. Crises can be either natural or human-made. In this work, we only focus on natural hazards. A natural hazard is an environmental phenomenon that has the potential to cause significant harm to human life, property, or the environment. These phenomena can range in severity from no threat at all to a complete disaster that severely impacts individuals and communities [2].

We have identified 17 types of natural hazards according to authoritative sources such as FEMA and NOAA. The 17 types of natural hazards are illustrated in Table I.

Table I
Natural Hazards

Hazard
Earthquake
Tsunami
Volcanic Eruption
Hurricane
Tornado
Flood
Wildfire
Drought
Landslide
Avalanche
Winter Storm
Heatwave
Storm Surge
Hailstorm
Lightning Strike
Dust Storm
Coastal Erosion

B. Crisis and Emergency Management

Traditionally, crisis and emergency management are seen as a cyclic process consisting of several connected phases: preparedness, response, and recovery [6].

The first phase, preparedness, contains activities undertaken before the beginning of the crisis to build a capacity for effective response. The response phase follows the preparedness phase immediately, and it focuses on urgent and life-saving activities. The last phase, recovery, aims to restore affected communities to functioning normally or to build back with resilience [9].

In the United States, FEMA serves as the lead federal agency for emergency management, while NOAA, which is a national weather service that plays a critical role in monitoring, forecasting, and issuing warnings for natural hazards.

Information is the main driver of effective emergency management. Timely, accurate, and actionable information enables decision-makers at all levels to allocate the right resources efficiently, prioritize response activities, and protect vulnerable populations. Research has demonstrated that information quality directly impacts emergency outcomes [6].

C. Information Systems in Emergency Management

Traditional Decision Support Systems (DSS) in emergency management have evolved for the years from simple database applications to become sophisticated platforms incorporating data from multiple sources with better analytical capabilities and connectivity. Despite these advantages, traditional systems still have serious limitations, including limited real-time data integration, rigid information structure, and lack of personalization [5].

The advances in artificial intelligence and machine learning created impressive possibilities for emergency management information systems. AI applications in this domain include predictive analysis for disaster forecasting, natural language processing for monitoring social media during crises, computer vision for damage assessment, and decision support algorithms for resource allocation [9].

D. Agentic AI and LLMs and their Applications

Agentic AI refers to employing advanced artificial intelligence systems, such as Large Language Models (LLMs), to serve a certain purpose. It is an AI system that processes and analyzes information to achieve a specific goal. AI agents can make decisions and adapt to dynamic environments with minimal human oversight. Unlike other AI applications, agents demonstrate pro-active and goal-driven behavior [7], [12].

Large Language Models (LLMs) represent a significant advancement in artificial intelligence, particularly in natural language processing. These models were trained on large corpora of text data which allows them to generate coherent and contextually relevant texts based on users' prompts. Recent LLMs such as GPT, Claude, and Google's Gemini show remarkable capabilities in understanding context, following instructions and generating structured outputs [6].

The applications of LLMs in emergency management are an emerging field with high potential. These models can process, analyze, and synthesize large volumes of information. Such capabilities allowed them to generate personalized guides and quickly adapt to changing circumstances [9].

III. Related Work

The applications of artificial intelligence (AI) and natural language processing (NLP) in emergency management have seen significant growth, aiming to improve various aspects of the disaster management cycle [9].

Gridach *et al.* highlighted agentic AI ability to handle complex tasks, perform multi-step reasoning, and adapt to dynamic environments [4]. These characteristics are necessary in emergency management. Furthermore, Velev *et al.* explored how these systems can enhance disaster response by analyzing real-time data and generating adaptive response plans [10].

Faiaz *et al.* emphasized the importance of AI in emergency management and proposed an AI-driven disaster warning system [3]. Their model integrated predictive data with an LLM to generate personalized guidelines for crisis response. Their work demonstrated that LLMs can make better decisions, perform efficient real-time analysis, and communicate effectively. Nevertheless, certain challenges, such as data scarcity, unreliability, and inconsistency, must be addressed first.

Xia *et al.* proposed a machine learning framework to predict natural disasters and generate early warnings [13]. These kinds of systems leverage historical data, sensor networks, and meteorological models to forecast events like floods, earthquakes, and wildfires with increased accuracy. While these systems excel at prediction, they typically do not extend to generating a personalized response plan for users based on predicted events.

Stewart *et al.* utilized natural language processing (NLP) to analyze unstructured data from social media and news websites during crises [9]. NLP techniques like information extraction, topic modeling, and classification are employed to gain situational awareness, detect emerging events, understand public sentiment, and identify needs or damage reports. While valuable for situational awareness, these applications can lack the objective of generating structured, actionable response plans.

Based on the works presented above, it is evident that most existing research highlighted the potential for agentic AI and LLMs in generating effective emergency response plans, however, there has been no prior work specifically dedicated to developing an AI agent for multi-stage crisis response plan generation. Therefore, our study is the first to introduce an agentic AI approach tailored for crisis guidance with real-time data integration from authoritative sources.

IV. Methodology

Our system follows a modular architecture designed to handle user input. fetch real-time data, consult a knowledge base, interact with an LLM, and manage the overall workflow. The key components of CrisisGuideGPT are listed below, and presented in Figure 1:

- **Input Processing and Validation:** Handles user input for city, state, and crisis type. The function normalizes state names/codes and ensures the crisis type is within a predefined valid list.
- **API Handlers:** Manages interaction with external APIs (NOAA National Weather Service API for active alerts and FEMA OpenFEMA API for historical disaster declarations).
- **Knowledge Base:** A static text file was chosen for simplicity, which serves as the primary knowledge base and contains extracted crisis guidelines from FEMA and NOAA for the 17 types of crises presented in Table I.
- **LLM Interaction Module:** The system interacts with Google's Gemini API through the google-generativeai library, managing a chat session with the model.

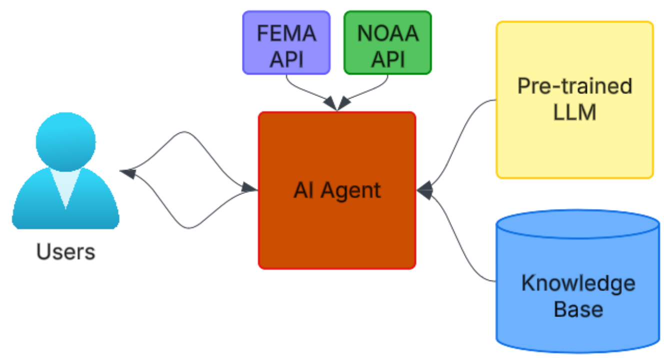

Figure 1. CrisisGuideGPT System Workflow

The knowledge base plays a crucial role in shaping the agent's responses and ensuring that the generated plans align with authoritative guidelines. The database contains pre-formatted text details about FEMA and NOAA guidelines for 17 crisis types illustrated in Table I. The agent follows a consistent structure with general guidelines and specific actions based on location, considering people with certain circumstances through user clarifications.

The agent is trained using the database, and the initial prompt instructs the model to use this data and fetch real-time data from authoritative APIs. This approach aims to provide the agent with a foundational set of verified information and reduce its reliance on internal training data [9].

CrisisGuideGPT integrates real-time data and historical data from FEMA and NOAA to provide context-specific information:

- **FEMA API:** We use the OpenFEMA API using filters for the specified state and crisis type. The response carries a list of past disaster declarations.
- **NOAA API:** We use NOAA NWS API for active alerts relevant to the state and crisis type. The response is a filtered list based on keywords associated with the crisis type.

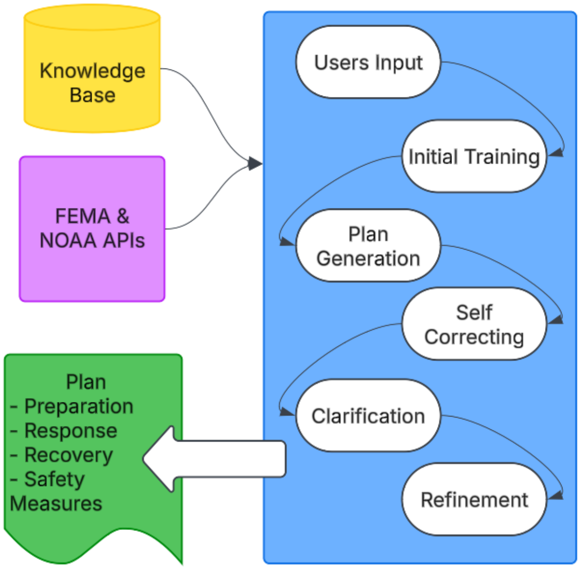

Figure 2. CrisisGuideGPT System Architecture

The information retrieved from both APIs is formatted into concise strings directly embedded within the prompt sent to the agent during initial plan generation. This step is necessary to ensure the agent is aware that there is both real-time and historical data.

A. Agent Prompting Strategy

Our model follows a multi-stage prompting strategy to guide Gemini's API towards generating accurate, structured, and personalized crisis response plans. The prompting strategy ans the system's architecture is explained below and illustrated in Figure 2:

- **Initial Training:** The first interaction involves sending a large prompt explaining the role definition and incorporating the knowledge from the database. The instructions are explicitly telling the agent to use these data along with API responses for future responses.
- **Initial Plan Generation Prompt:** This prompt provides the core request for plan generation. It includes the goal, location, and crisis type. It also contains the fetched FEMA and NOAA information with detailed instructions on the desired output.
- **Self-Correcting Prompt:** After receiving the initial plan, the system sends it back to the agent with a refining prompt. The agent is instructed to review the plan for consistency and completeness.
- **Clarification Prompt:** This prompt instructs the agent to analyze the plan and identify critical safety measures, evacuation routes, unclear or missing resources, and generate clarifying questions for the user to answer.
- **Plan Refinement Prompt:** If the user answers the clarifying questions, the refinement prompt tells the agent to refine and adjust the original plan based on the user's clarification.

V. Results

This section presents the results of CrisisGuideGPT. We evaluated the system through case studies and performance metrics. Results demonstrates CrisisGuideGPT's ability to integrate active alerts from NOAA with FEMA historical data and declarations. Results also shows CrisisGuideGPT's ability to produce structured, context-aware plans, leveraging Gemini's language capabilities and user inputs.

A. Case Studies

We tested CrisisGuide with two scenarios: a hurricane in Miami, FL, and a wildfire in San Francisco, CA. These cases highlight personalization through clarification questions and API integration.

1) Hurricane in Miami, FL

The system processed inputs: city = "Miami", state = "FL", crisis type = "Hurricane". After the inputs were validated, the NOAA API returned no active hurricane alerts, and the FEMA API reported historical 1,000 hurricane-related disaster declarations. The initial plan, generated in approximately 18.95 seconds, followed the structured format (Preparation, Response, Recovery, Safety Measures), incorporating FEMA and NOAA guidelines. Table II summarizes key plan components.

Table II
SUMMARY OF CRISISGUIDEGPT PLAN FOR HURRICANE IN MIAMI, FL (NO ACTIVE NOAA ALERTS, AND 1000 HISTORICAL FEMA DECLARATIONS)

Section	Key Recommendations
Preparation	preparing supplies, securing property, knowing evacuation routes, reviewing insurance, and creating a family communication.
Response	Stay updated, evacuate if told, use safe routes, shelter wisely, turn off utilities if needed, and seek higher ground if flooding
Recovery	Return only when safe, inspect for damage, report losses, clean up carefully, seek aid, follow disposal rules, and plan to prevent future risks.
Safety Measures	Practice safety drills, install detectors, reinforce your home, get proper insurance, stay informed, and keep your emergency plan up to date.

With no active NOAA alerts, the clarification module generated 4 non-urgent questions presented in Figure 3:

The refined hurricane response plan, after answering the clarifying questions, adds a backup communication method like a satellite phone and emphasizes practicing evacuation routes. It updates the emergency kit to include medication storage solutions for power outages and communication aids for people with disabilities. Finally, it confirms that the insurance coverage is not only reviewed but also sufficient to cover major hurricane-related losses.

2) Wildfire in San Francisco, CA

The second test used inputs: city = "San Francisco", state = "CA", crisis type = "Wildfire". The NOAA API returned no active alerts, and the FEMA API reported no historical wildfire

```
1. Have you identified your specific evacuation zone in Miami-Dade County and practiced your evacuation route?
Your answer:  Yes
2.  Beyond the basic hurricane kit, have you considered specialized needs like medication storage for power outages,
or communication methods for those with visual or auditory impairments?
Your answer:  No
3. Given Miami's history of significant hurricane damage (1000 FEMA declarations), have you reviewed and updated you
r homeowners and flood insurance policies to ensure adequate coverage for potential losses?
Your answer:  Yes
4.  Does your family communication plan include a backup method (e.g., satellite phone) in case cell service is disr
upted during a hurricane?
Your answer:  No
```

Figure 3. Clarifying Questions in CrisisGuideGPT

declarations. The initial plan relied on the knowledge base and included:

- **Preparation:** Take proactive steps to plan evacuation, gather supplies, and reduce fire risks around your home.
- **Response:** Act quickly during a wildfire by following alerts, evacuating safely, and bringing essential items.
- **Recovery:** Safely return, assess damage, and seek help to restore your home and well-being after the wildfire.
- **Safety Measures:** Follow official guidance, stay informed, and protect yourself from environmental hazards during a wildfire.

B. Performance Metrics

CrisisGuide's performance was assessed on a standard laptop (Intel i7, 16GB RAM) running the Jupyter Notebook implementation, focusing on plan generation time.

- **Plan Generation Time**: Across 17 test runs for all available scenarios, the average time to generate an initial plan, including API calls and Gemini processing, was 20.67 seconds. Figure 4 below presents the generation time in seconds for all hazards.

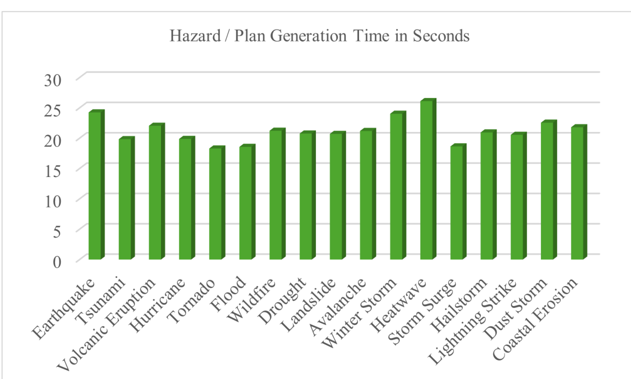

Figure 4. Plan Generation Time

C. Discussion of Results

The evaluation confirms that CrisisGuideGPT provided structured, context-aware emergency plans by integrating NOAA alerts, FEMA data, and user input. In both cases, the system generated relevant guidance even without active alerts by relying on historical declarations and the knowledge base. The clarification module played a crucial role in personalizing plans and refining recommendations based on specific users' needs, like medical concerns or communication barriers. Performance remained practical, with average generation time around 20 seconds. Compared to static FEMA documents, CrisisGuideGPT provides faster, adaptable, and user-specific planning.

VI. Discussion

Our framework demonstrates notable strengths in crisis response and emergency management planning. One of our key advantages is focusing on personalization. It was a necessary step to overcome the traditional static guidelines. We incorporated location, crisis type, and user-provided clarification in order for the agent to tailor a response plan effective for individual circumstances.

Real-time integration is another important feature in CrisisGuideGPT, which enables it to incorporate current conditions and active alerts into its plan generation. This integration is necessary to address one of the most significant limitations in traditional emergency management frameworks.

CrisisGuideGPT was also trained using a knowledge base from credible and authoritative sources: FEMA and NOAA. This allowed the agent to gain increased accuracy and reduce the risk of generating misleading or incorrect information. Furthermore, the refinement mechanism allowed the agent to enhance the plan's completeness and relevance while involving the user in the process.

Despite the advantages our model has, there are still some limitations. First, there is a room for error as Large Language Models can be unpredictable. This inherent drawback can make the output occasionally contain inaccuracies or inappropriate recommendations. Another limitation is the dependency on external APIs. While the integration can provide timely and official information, external APIs can experience some downtime or changes in configuration. In addition, LLMs work in a black box nature, which hinders transparency and accountability. The lack of explainability in LLMs makes it difficult to audit them or explain why certain details were included or omitted.

A. Use Cases

CrisisGuideGPT's primary goal and intended use are to enhance people's preparedness for various crisis types. By providing personalized, location-based guidance, the agent can help individuals and the community develop more effective emergency plans designed to adapt to their circumstances.

We believe CrisisGuideGPT can serve as a valuable tool for emergency responders and planners. It can be used to generate a draft response plan for various scenarios rapidly. First responders can also benefit from getting quick access to relevant protocols.

VII. Limitations

While CrisisGuideGPT demonstrates strong potential in enhancing crisis response planning, several limitations must be acknowledged. These limitations span technical, data-related, and ethical dimensions, and they highlight areas for future improvement.

Technical Limitations: Dependence on external APIs is one of the core functionalities of CrisisGuideGPT. Each API comes with its own considerations and constraints. While the current implementation includes basic error handling, a more robust implementation might require caching strategies, multiple fallback data sources, and local deployment options.

Another technical concern is the limitations of LLM technologies. Issues such as inaccuracy, hallucination, and inconsistency remain a significant challenge. Especially in a high-stakes domain such as public safety. The major concern here is hallucination, particularly in crisis contexts where made-up information can lead to harmful results.

Data Limitations: One of the major limitations is the fact that the knowledge database used in training the agent is static in nature. There are no update mechanisms or version information. This limitation is not exclusive to CrisisGuideGPT, but rather, an existing disadvantage for all Large Language Models [11].

The API information provides a valuable context, but it has limitations in terms of granularity and timeliness. NOAA alerts, for example, are filtered by state rather than the user's precise location or city. Furthermore, both APIs provide only snapshot data at the time of the query.

The agent doesn't incorporate hyperlocal data sources that could potentially enhance the relevance or accuracy of its guidance, such as local emergency management agencies, transportation departments, utility companies, and community organizations.

Transparency and Trustworthiness: While CrisisGuideGPT is designed to enhance transparency and trustworthiness by incorporating data from authoritative sources, the agent faces transparency limitations due to its reliance on a Large Language Model. LLMS are generally vague and lack transparency.

The clarification step represents a form of human oversight that partially addresses the transparency and trustworthiness concern, however, this is limited in several ways, and a robust mechanism is left for future development.

VIII. Conclusion

In this study, we introduced CrisisGuideGPT, an Agentic AI system designed to generate personalized emergency response plans during crises using a few-shot learning approach with the Google Gemini API. By integrating real-time data and official guidelines from authoritative sources, CrisisGuideGPT demonstrated strong performance in producing structured, context-aware, accurate, and personalized response plans. Experimental results demonstrated our framework's ability to generate crisis response plans based on user location and specific circumstances. CrisisGuideGPT averaged around 20 seconds in plan generation, which is efficient, especially with the narrow window of action during crises.

While the model shows its great potential, there is room for future improvements. CrisisGuideGPT limitations included the reliance on external APIs, the possibility of generating inaccurate and inconsistent information due to LLMs' drawbacks, the static nature of its knowledge base, and the lack of interpretability.

Future work will focus on developing strategies to mitigate external APIs' failure and a mechanism to periodically update the knowledge base. This enhancement will increase the robustness and reliability of the framework and improve its adaptability against the rapid and evolving nature of crises.

Ultimately, CrisisGuideGPT represents a foundational step toward intelligent, adaptive, and user-centric crisis response systems. Its architecture and methodology offer a blueprint for future AI-driven tools in emergency management and other high-stakes domains where real-time decision support is essential.

References

[1] D. B. Acharya, K. Kuppan, and B. Divya. Agentic AI: Autonomous intelligence for complex goals—a comprehensive survey. *IEEE Access*, 13:18912–18936, 2025.

[2] M. T. Chaudhary and A. Piracha. Natural disasters—origins, impacts, management. *Encyclopedia*, 1(4):1101–1131, 2021.

[3] M. A. Faiaz and N. Nawar. Short paper: Ai-driven disaster warning system: Integrating predictive data with llm for contextualized guideline generation. In *Proceedings of the 11th International Conference on Networking, Systems, and Security*, pages 247–253, 2024.

[4] M. Gridach, J. Nanavati, K. Z. E. Abidine, L. Mendes, and C. Mack. Agentic ai for scientific discovery: A survey of progress, challenges, and future directions. *arXiv preprint arXiv:2503.08979*, 2025.

[5] L. A. Jardim Gonçalves, S. Ferreira, and P. J. Ribeiro. Assessing urban mobility resilience: An exploratory approach using hazard-based duration models. *Electronics*, 13(21):4220, 2024.

[6] O. A. Maher, L. Cegolon, and S. Bellizzi. Artificial intelligence as a tool for enhancing the performance of public health emergency operation centres (eoc). *BMJ Global Health*, 10(3), 2025.

[7] 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*, pages 1–22, 2023.

[8] Statista. Number of natural disasters events globally from 2000 to 2024, 2025. Accessed: May 3, 2025.

[9] J. Stewart, J. Lu, A. Goudie, G. Arendts, S. A. Meka, S. Freeman, K. Walker, P. Sprivulis, F. Sanfilippo, M. Bennamoun, et al. Applications of natural language processing at emergency department triage: A narrative review. *Plos one*, 18(12):e0279953, 2023.

[10] D. Velev and P. Zlateva. Challenges of artificial intelligence application for disaster risk management. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48:387–394, 2023.

[11] S. Wang, Y. Zhu, H. Liu, Z. Zheng, C. Chen, and J. Li. Knowledge editing for large language models: A survey. *ACM Computing Surveys*, 57(3):1–37, 2024.

[12] Z. Xi, W. Chen, X. Guo, W. He, Y. Ding, B. Hong, M. Zhang, J. Wang, S. Jin, E. Zhou, et al. The rise and potential of large language model based agents: A survey. *Science China Information Sciences*, 68(2):121101, 2025.

[13] Y. Xia and J. Lu. Short-term flood prediction model based on pre-training enhancement. *Electronics*, 13(11):2203, 2024.