Literature Review: Multimodal Situational Awareness for Crisis Response

Georgi Tsvetanski

University of Baltimore

IDIA 612: Interaction Design

Professor Janisa Simmons

March 9, 2026

Abstract

Decision-making during a crisis is very often paralyzed by information fragmentation. This literature review evaluates the design requirements for the MUMOSA (MUlti-MOdal Situation Awareness) interactive dashboard, a project for the Army Research Laboratory. By synthesizing recent academic sources, this paper identifies the psychological and physiological needs of investigators and responders operating under high cognitive load. It outlines critical design heuristics for adaptive, vision-based multimodal interfaces and attention-preserving AI systems. The review compares existing spatial-temporal incident tools with emerging Large Language Model (LLM) document intelligence. The findings suggest a design path focused on

the "Resolve" phase of emergency management, utilizing schema graphs and 3D reconstruction to bridge the gap between digital evidence and real-world understanding.

I. Introduction

When a major crisis or disaster occurs, the people responsible for managing the response are often hit with a massive, unorganized wave of data. Investigators and responders find it incredibly difficult to "determine and assess what happened when" because the evidence is spread across text reports, 2D photos, and 3D spatial reconstructions (Lukin et al., 2024, p. 1). In the immediate aftermath of an event, the problem is rarely a lack of information; rather, it is the inability to connect these different pieces of data into a story that makes sense.

To address this "fragmented data" problem, the Army Research Laboratory (ARL) developed the **MUMOSA (MUlti-MOdal Situation Awareness)** interactive dashboard. This system uses Artificial Intelligence—specifically Large Language Models (LLMs) and Vision-Language Models (VLMs)—to "provide a comprehensive platform for event situational awareness" (Lukin et al., 2024, p. 1). By taking text reports and images and turning them into a unified "schema graph," the system allows users to see how different sub-events are connected over time.

This literature review explores the best way to design these high-stress dashboards. It analyzes the primary user groups, the psychological rules (heuristics) required to keep users from being overwhelmed, and what other researchers have built in this space. Ultimately, this review serves as a design guide for the MUMOSA system, ensuring it moves from a collection of data to a clear, evidence-based understanding of complex situations.

II. User Groups and Context of Use

The MUMOSA dashboard is intended for three primary groups: investigators, analysts, and emergency responders. These users do not work in standard office settings; they operate in "VUCA" contexts—Volatile, Uncertain, Complex, and Ambiguous. In these environments, the interface is not just a tool, but a cognitive lifeline.

Physiological and Psychological Stressors

Research into first responder training reveals that these users deal with "high anticipatory anxiety and cognitive load" (Giaume et al., 2024, p. 1). This stress is not just a feeling; it is a physical reality that changes how humans process information on a screen. When responders are tasked with identifying hazards or searching for victims, their "mental battery" or cognitive capacity is drained.

Giaume found that the stress levels of responders are measurable and impactful. The psychological, cognitive, and physiological impact of hazards casualties' trainings on first responders... proves emergency responders experience high anticipatory anxiety and cognitive load during crises.

For an interaction designer, this means that every extra button or unnecessary alert on the screen is a potential failure point. If the user is already at their mental limit, a cluttered dashboard can lead to "subpar performance" and cause "confusion or disorientation" (Zhou et al., 2023, p. 1).

Information Overload and Capacity Limits

While responders need speed, analysts and investigators need to process massive volumes of media to build accurate reports. However, even these specialized teams are reaching a breaking point. A recent design case study on Computer Emergency Response Teams (CERTs) found that analysts often report: "We do not have the capacity to monitor all media" (Kaufhold et al., 2024, p. 1).

There is a documented "gap between information supply and the capacity to monitor it" (Kaufhold et al., 2024, p. 1). This confirms that the MUMOSA dashboard cannot just be a "data dump." It must act as a filter that prioritizes the most important evidence based on the user's specific role at that moment.

III. Heuristics for High-Stress Dashboards

To support users under extreme pressure, we must follow design heuristics that prioritize attention, trust, and mental energy.

1. Attention-Preserving Heuristics

Standard software often uses "attention-capture" designs—like red notification badges or pop-ups—that can be dangerous in a crisis. Instead, we need the opposite. Monge Roffarello et al. (2025) propose a set of "digital attention heuristics" designed to "preserve and respect user attention by design" (p. 1).

These heuristics are grounded in three fundamental psychological needs:

1. **Autonomy:** The user must feel in control of what they are looking at.
2. **Competence:** The user must feel effective in their task.
3. **Relatedness:** The user must feel connected to the rest of the response team (Monge

> Roffarello et al., 2025, p. 1).

In the context of the MUMOSA dashboard, this means the system should never interrupt an investigator who is deep in a 3D reconstruction unless the new data is life-critical.

2. Cognition-Driven Adaptive Systems

Because a user's stress level fluctuates, the interface should be "auto-adaptive." The **CogDNA** (Cognition-driven navigation assistive) framework proves that interfaces tailored to real-time cognitive load improve performance.

Zhou et al. (2023) state that:

> Adaptive interfaces tailored to real-time cognitive load reduce disorientation and improve performance... for emergency wayfinding where responders need to process a large amount of information while performing search and hazard identification tasks. (p. 1)

For the Army Research Lab project, this suggests that if the system detects the user is overwhelmed, it should automatically simplify the UI, hiding secondary data layers and highlighting the primary narrative timeline.

3. AI Trust and "Explainable" Heuristics

MUMOSA uses AI to "summarize" complex event data. However, if the user doesn't trust the AI, the tool is useless. Designers must use specific "design stimuli" to make AI interactions feel more natural and trustworthy. Jin et al. (2021) identify several key categories for AI ideas, including "Identifying," "Forecasting," and "Adaptive" signals (p. 9).

The best way to build this trust is through transparency. The current MUMOSA prototype handles this by including "Show Source" buttons next to every AI-generated event extraction (Lukin et al., 2024, p. 14). This allows an analyst to click a text summary and instantly see the original photo or text snippet it came from, proving the AI isn't just making things up.

IV. Prior Work: What Others Have Done

Understanding the design space requires looking at how others have tried to solve the situational awareness problem.

Spatial-Temporal Incident Dashboards

Vanderbilt University developed an "incident management and analysis dashboard for fire departments" (Pettet et al., 2019, p. 1). This system is highly effective at showing "where" and "when" incidents happen on a map. By analyzing survival data and incident distributions, it helps dispatchers send the closest responder (Pettet et al., 2019).

While these spatial-temporal tools are great for logistics, they have a major limitation: they don't tell the "story." They are great at showing a dot on a map, but they don't explain the "complex sequence of events" that led to that dot (Lukin et al., 2024, p. 1). This is where MUMOSA provides a unique advantage.

LLMs in Document Intelligence

On the other side of the design space are systems built purely for reading reports. Recent advances in **Large Language Models (LLMs)** have "dramatically transformed" document intelligence, allowing for much more accurate processing of reports (Ke et al., 2025, p. 1).

New methods like **Retrieval-Augmented Generation (RAG)** allow an AI to search through thousands of field reports to find a specific fact without getting confused by the "long context" (Ke et al., 2025). The gap in this space, however, is that these systems are usually text-only. They lack the "visual-language" integration needed to understand a photo of a disaster site.

Vision-Based Multimodal Interfaces (VMIs)

The "state-of-the-art" is currently moving toward **Vision-based Multimodal Interfaces (VMIs)**. Hu et al. (2025) conducted a survey of these interfaces and found that they are "essential for integrating multimodal data" (p. 1). By combining visual data with text, VMIs allow for a "more precise interpretation of user intentions" (Hu et al., 2025). This research confirms that MUMOSA's path of combining VLMs and LLMs is the correct academic direction for high-fidelity crisis management.

V. Discussion and Design Path Forward

The research consistently shows that our current mental model of crisis response is too simple. Most procedural guidelines say there are only two phases: "response" (the incident is happening) and "recovery" (the incident is over).

The Three-Phase "Resolve" Model

Research into multi-team emergency response has found that this two-phase model doesn't match how people actually behave. Instead, there is a middle ground called the "Resolve" phase.

Brown et al. (2021) found that:

> The current response/recovery classification does not fit the nuanced context of an emergency. Instead, a three-phase structure of 'response/resolve/recovery' is more reflective of behaviour. (p. 591)

MUMOSA is a "Resolve-phase" tool. It is built for that moment when the immediate danger has passed, but the team is still working to "coordinate across event phases" and verify what happened (Brown et al., 2021).

The Path for the MUMOSA Prototype

Based on the heuristics and prior work analyzed in this review, the design path for the MUMOSA prototype should focus on three specific elements:

1. **Mission-Aware Filtering:** Using the attention heuristics from Monge Roffarello et al. (2025), the dashboard must hide low-priority data when it detects high cognitive load. It should prioritize "Mission Awareness" over general data browsing.
2. **Cross-Modal Verification:** Every AI extraction must follow the "Show Source" heuristic (Lukin et al., 2024). This bridges the gap between the LLM's text summaries and the visual evidence.
3. **Phase Transition Logic:** The UI should physically shift its layout as the user moves from the "Response" phase (speed/alerts) to the "Resolve" phase (deep analysis/schema graphs).

By following these heuristics and grounding the design in the three-phase "Resolve" model, we can ensure the MUMOSA dashboard helps investigators and responders turn "fragmented data" into a "clear, evidence-based understanding" (Design Brief).

VI. References

Brown, O., Power, N., & Conchie, S. M. (2021). Communication and coordination across event phases: A multi-team system emergency response. *Journal of Occupational and Organizational Psychology*, *94*(3), 591–615. https://doi.org/10.1111/joop.12349

Giaume, L., Le Roy, B., Daniel, Y., Lauga Cami, H., Jost, D., Travers, S., & Trousselard, M. (2024). Psychological, cognitive, and physiological impact of hazards casualties' trainings on first responders: The example of a chemical and radiological training. An exploratory study. *Frontiers in Psychology*, *15*, 1336701. https://doi.org/10.3389/fpsyg.2024.1336701

Hu, Y. O., Tang, J., Gong, X., Zhou, Z., Zhang, S., Elvitigala, D. S., Mueller, F. F., Hu, W., & Quigley, A. J. (2025). Vision-based multimodal interfaces: A survey and taxonomy for enhanced context-aware system design. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, Article 758, 1–31. https://doi.org/10.1145/3706598.3714161

Jin, X., Evans, M., Dong, H., & Yao, A. (2021). Design heuristics for artificial intelligence: Inspirational design stimuli for supporting UX designers in generating AI-powered ideas. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, Article 219, 1–8. https://doi.org/10.1145/3411763.3451727

Kaufhold, M.-A., Riebe, T., Bayer, M., & Reuter, C. (2024). 'We do not have the capacity to monitor all media': A design case study on cyber situational awareness in computer emergency response teams. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, Article 580, 1–16. https://doi.org/10.1145/3613904.3642368

Ke, W., Zheng, Y., Li, Y., Xu, H., Nie, D., Wang, P., & He, Y. (2025). Large language models in document intelligence: A comprehensive survey, recent advances, challenges, and future trends. *ACM Transactions on Information Systems*, *44*(1), Article 18, 1–64. https://doi.org/10.1145/3768156

Lukin, S. M., Bowser, S., Suchocki, R., Summers-Stay, D., Micher, J., Matuszek, C., Francis, F., & Voss, C. R. (2024). MUMOSA, interactive dashboard for multi-modal situation awareness. *Proceedings of the Workshop on the Future of Event Detection (FuturED)*, 32–47.

Monge Roffarello, A., De Russis, L., & Lukoff, K. (2025). The digital attention heuristics: Supporting the user's attention by design. *ACM Transactions on Computer-Human Interaction*, *32*(4), Article 38, 1–41. https://doi.org/10.1145/3725215

Pettet, G., Mukhopadhyay, A., Samal, C., Dubey, A., & Vorobeychik, Y. (2019). Incident management and analysis dashboard for fire departments: ICCPS demo. *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, 336–337. https://doi.org/10.1145/3302509.3313329

Zhou, T., Xia, P., Zhu, Q., & Du, J. (2023). Cognition-driven navigation assistive system for emergency indoor wayfinding (CogDNA): Proof of concept and evidence. *Safety Science*, *162*, 106100. https://doi.org/10.1016/j.ssci.2023.106100