Multimodal Situational Awareness for Civil-Military Coordination in Port Disruption Scenarios

Sabine Janzen*
German Research Center for Artificial Intelligence (DFKI)
email

Rishika Kumari
German Research Center for Artificial Intelligence (DFKI)
email

Wolfgang Maass
German Research Center for Artificial Intelligence (DFKI)
email

ABSTRACT

Disruptions in port access and transport infrastructures pose significant challenges for crisis management and civil–military coordination, particularly in dual-use environments. Decision-makers must cope with heterogeneous, incomplete, and partially conflicting information under time pressure. This work-in-progress paper introduces ARGUS, a generic multimodal decision-support architecture designed to consolidate diverse information sources into structured, geo-referenced incidents with explicit confidence values and traceable evidence. To explore the feasibility of core architectural concepts, a proof-of-concept instantiation is implemented for the port region of Wilhelmshaven, integrating a limited set of representative modalities, including emergency communication audio, social media reports, and hydrological sensor data. Exploratory results suggest that multimodal fusion can combine multiple raw alerts into fewer, more coherent incident views for specific port-access routes. At the same time, the proof of concept reveals challenges in merging nearby events too aggressively, interpreting confidence scores, and managing information in environments where civilian and military actors operate together.

Keywords

Crisis management, Multimodal decision support, Situational awareness, Information fusion, Dual-use infrastructure, Port operations, Explainable AI.

INTRODUCTION

Modern port operations are increasingly challenged by the combined effects of extreme weather, hybrid threats, and information overload in control centers. Storm surges, rapidly evolving traffic disruptions, and deliberate interference such as Global Navigation Satellite System (GNSS) jamming can disrupt both civilian supply chains and defense logistics (Braun et al., 2025; Bunker, 2020). This is particularly critical in Host Nation Support (HNS) settings and Seaports of Debarkation (SPOD) operations,[1] where convoys, hazardous materials, and time-critical movements place additional stress on infrastructure and coordination capacity.[2] Operational decision-making during such disruptions relies on diverse, high-velocity information streams, including radio communications, hydrological and weather sensors, structured alerts, CCTV/UAV imagery, social media, and local news reporting. Yet many

---

*corresponding author

[1] See NATO AJP-4.3 (Allied Joint Doctrine for the Execution of Support to Operations) for doctrinal context on logistics support and HNS-related coordination; public availability depends on edition and release: https://www.coemed.org/files/stanags/01_AJP/AJP-4.3_EDA_V1_E_2234.pdf.

[2] See Bundeswehr Host Nation Support (HNS) strategy and briefing documents for the German implementation context; publicly available versions vary.

command-and-control and crisis management systems remain modality-specific or pre-integrated around a limited set of sensor feeds. As a result, operators must reconcile fragmented and sometimes contradictory inputs under severe time pressure, increasing cognitive load and slowing response (Kaufhold et al., 2020; Steen-Tveit and Erik Munkvold, 2021; Van de Walle et al., 2016).

Prior research in crisis management and situational awareness has emphasized the need for systems that integrate heterogeneous data, support robust decision-making under uncertainty, and fit established operational workflows (Steen-Tveit and Erik Munkvold, 2021; Van de Walle et al., 2016). In crisis informatics, social media and other unstructured sources can provide early signals but also create noise, duplication and verification challenges that must be managed (Kaufhold et al., 2020; Luna and Pennock, 2018). Recent surveys further highlight both the opportunities and current limitations of multimodal and deep-learning approaches for disaster response, including issues of robustness, explainability and operational integration (Algiriyage et al., 2022; Ghaffarian et al., 2023; Zhao et al., 2024). To address this gap, we present ARGUS (cf. Figure 1), a modular multimodal AI system designed to support port access and movement coordination under disruptive conditions. ARGUS ingests text, image, audio, sensor, and infrastructure-graph inputs; performs anomaly and intent detection; and fuses signals into georeferenced incidents with confidence values and evidence trails. It produces prioritized response suggestions in interoperable formats (e.g., JSON/GeoJSON) intended to integrate with existing dashboards and routing or movement-planning tools. The system is designed to reduce operator burden by shifting attention from raw alerts toward synthesized, evidence-linked situational insights.

This work-in-progress paper reports on the design of ARGUS and a proof-of-concept (PoC) implementation in the dual-use port context of Wilhelmshaven in Germany, focusing on JadeWeserPort and its adjacent military logistics setting. In this paper, dual-use refers to port and transport infrastructures that serve both civilian and military operations and therefore require coordination across actors with different responsibilities, access regimes, and decision logics. The PoC tests simulated storm-surge and hybrid-threat scenarios and integrates representative open sources such as water level alerts (e.g., PegelOnline) and weather data (e.g., DWD), social media signals, and audio transcripts of German emergency service radio communications (BOS; Behoerden und Organisationen mit Sicherheitsaufgaben). The current PoC is based on simulated scenarios and simulated multimodal inputs derived from scenario-based ground truth. Accordingly, the presented results demonstrate the feasibility of the fusion architecture under controlled conditions rather than real-world field performance. The contribution of this work is twofold: (1) we introduce a domain-adapted multimodal fusion pipeline that bridges heterogeneous inputs into structured situational understanding; and (2) we provide scenario-based prototype evidence of real-time incident fusion and prioritized response generation. Our work is guided by three research questions:

RQ1. To what extent can multimodal fusion consolidate heterogeneous port-disruption signals into consistent, georeferenced incidents and thereby reduce duplicate or conflicting alerts?
RQ2. How well do confidence values and evidence trails correspond to scenario ground truth, and how effectively do they communicate uncertainty when modalities disagree?
RQ3. What integration, interoperability (e.g., JSON/GeoJSON), and governance requirements arise for multimodal decision support in a dual-use port environment?

The PoC provides scenario-based evidence for these questions through injected incidents with known ground truth, system logs (fusion decisions and confidence), and analysis of the resulting incident objects and recommendations.

BACKGROUND WITH SCENARIOS

Port operations in dual-use environments increasingly face disruptions driven by extreme weather, infrastructure bottlenecks, and hybrid threat activity. In such settings, civil protection and military logistics actors must coordinate under time pressure while monitoring heterogeneous information streams that differ in reliability, latency, and format. When information arrives as fragmented sensor readings, short radio messages, and unstructured public reports, control centers risk alert overload, duplicated incident reporting, and delayed understanding of what is happening where and with what operational impact (Kaufhold et al., 2020; Steen-Tveit and Erik Munkvold, 2021; Van de Walle et al., 2016). ARGUS was developed in response to these conditions, with a focus on (i) consolidating heterogeneous signals into a small number of georeferenced, evidence-linked incidents and (ii) providing interoperable outputs that can be integrated into existing dashboards and planning tools. To illustrate the operational gap and to ground the proof-of-concept evaluation, we describe two representative scenarios derived from disrupted port-access conditions. These scenarios are designed to reflect the types of multimodal signals that can realistically be obtained during an incident (e.g., public warnings, sensor thresholds, radio transcripts, and open-source reporting) and the corresponding need for fusion and prioritization (Luna and Pennock, 2018).

Scenario 1: Storm surge and flooding affecting access corridors. In the first scenario, a storm surge causes localized flooding that threatens critical road segments and underpasses on the approach to JadeWeserPort near

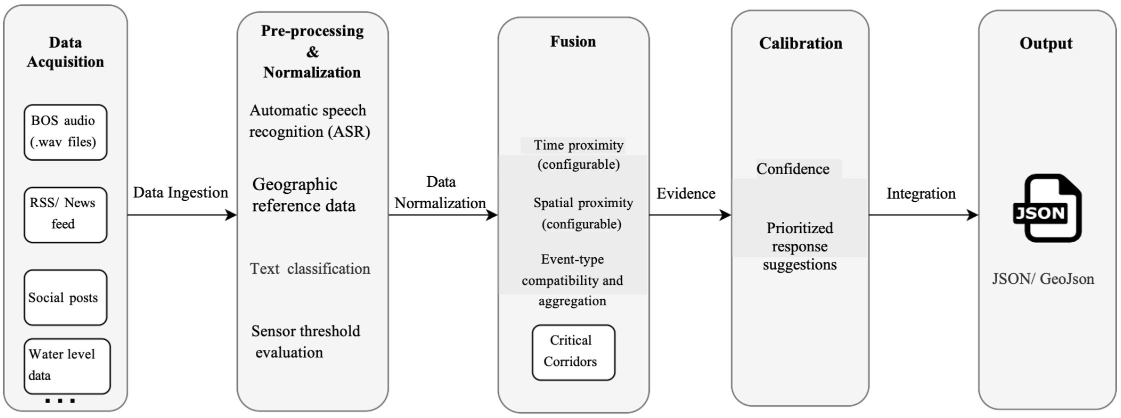


Figure 1. Generic architecture of ARGUS illustrating the conceptual processing stages from multimodal signal ingestion and normalization to evidence-based incident fusion, confidence calibration, and interoperable output. Specific data sources shown reflect an exemplary proof-of-concept instantiation and are configurable or replaceable in other deployments.

Wilhelmshaven (Green et al., 2025). Early indications may appear in structured sources (e.g., water-level thresholds from PegelOnline and weather warnings from DWD), but operationally relevant impacts are often first reported through unstructured channels such as BOS radio messages ("road is under water"), public posts, and short local news updates. Individually, these inputs are noisy and may be incomplete; combined, they can confirm both the event type and the affected corridor. In the PoC, ARGUS fuses these signals into a single flooding incident with an explicit evidence trail, assigns confidence based on cross-source agreement, and exposes the affected corridor as GeoJSON together with prioritized response suggestions (e.g., rerouting triggers, notification to civil protection units).

Scenario 2: SPOD access under hybrid disruption signals. In the second scenario, port access for a scheduled convoy movement is affected by ambiguous, low-level signals consistent with a hybrid disruption (e.g., unauthorized vehicle reports near a gate, suspicious object messaging, or simulated GNSS instability along a route segment) (Braun et al., 2025). Such signals may emerge across modalities and vary in credibility: BOS radio reports, public posts about checkpoint activity, local media items, and route-level anomalies from navigation. The operational challenge is to avoid both underreaction (missing an emerging threat) and overreaction (triggering disruptive measures based on a single unreliable source) (Kaufhold et al., 2020). In the PoC, ARGUS consolidates these inputs into a georeferenced security-disruption incident linked to a specific access corridor, produces a confidence estimate that reflects both evidence diversity and contradiction, and generates a small set of response suggestions (e.g., delay movement, dispatch security, enable backup corridor), again exposed via JSON/GeoJSON for integration with existing tools.

Together, these scenarios capture the core challenge addressed in this work; that is transforming fragmented, cross-modal signals into structured, confidence-scored incidents that can be acted upon under time pressure. They also motivate the PoC evaluation design, where injected incidents with known ground truth and system logs enable assessment of fusion consolidation (RQ1), confidence behavior under agreement and disagreement (RQ2), and integration requirements for interoperable incident products (RQ3).

RELATED WORK

Research on crisis management information systems shows that responders must develop shared understanding under severe time pressure while information is fragmented, rapidly changing, and often inconsistent. Empirical and conceptual work on situation awareness, common operational pictures (COP), coordination, and sensemaking highlights both the value of enriched information and the risk of cognitive overload and delayed action when information environments are poorly integrated (Comes et al., 2011; Kaufhold et al., 2020; Nespeca et al., 2020; Steen-Tveit and Erik Munkvold, 2021; Van de Walle et al., 2016). In parallel, decision-support and geospatial literature identifies recurring challenges in operational integration, workflow fit, and interoperability of map-based products for emergency response (Newman et al., 2017; Opach et al., 2023; Schätter et al., 2019).

A second line of work in crisis informatics exploits unstructured streams—particularly social media—for early warning and rapid mapping. While such sources can provide timely local signals, they introduce noise, duplication, verification burdens, and geolocation uncertainty (Luna and Pennock, 2018; Reuter et al., 2018, 2020). Systems

Input Modality	**Normalization Role**	**Fusion-Relevant Processing**	**Contribution to Incident Evidence**
Social media text and news feeds	Early, low-latency disruption signals	Event-type classification and constrained geocoding (Transformer- based zero-shot classifier (e.g., German_Zeroshot), OSM-based gazetteer, fuzzy string matching	Weak-to-moderate evidence for incident emergence and spatial hypotheses
Hydrological sensor data, i.e., water level data	Structured environmental condition monitoring	Threshold-based normalization and criticality labeling (rules)	High-reliability evidence for flooding-related incidents
Emergency communication audio, i.e., BOS audio	Operational situation reporting	Automatic speech recognition and event-type classification (e.g., Whisper ASR, Transformer-based zero-shot classifier (e.g., German_Zeroshot))	Authoritative, context-rich evidence from coordination channels

Table 1. Functional role of heterogeneous input modalities within the ARGUS fusion pipeline, highlighting how different exemplary sources contribute varying evidential strength to incident construction.

such as E2mC demonstrate near-real-time filtering and geolocation pipelines for emergency management, and related work addresses social-media image processing and context-aware geolocation of emergency-related posts (Alam et al., 2018; Havas et al., 2017; Scalia et al., 2022).

Recent surveys highlight the growing potential of multimodal and deep-learning approaches for disaster response, while emphasizing robustness, explainability, and deployment constraints as key gaps for real-world adoption (Algiriyage et al., 2022; Ghaffarian et al., 2023; Zhao et al., 2024). ARGUS targets this intersection by fusing structured and unstructured signals across modalities (audio, text, sensors, and geospatial context) into unified, georeferenced incidents with confidence values and evidence trails, exposed via interoperable outputs (JSON/GeoJSON) designed to integrate with existing control-center tools. As ongoing work, we plan to extend ARGUS with retrieval-augmented generation to produce evidence-linked summaries and rationales grounded in retrieved guidance and incident context (Lewis et al., 2020).

SYSTEM APPROACH

ARGUS is a modular, multimodal decision-support architecture for situational awareness and action prioritization in disrupted transport and port-access environments. It can be instantiated for different regions, infrastructures, and operational contexts, including civil protection, host nation support, or military logistics. The core objective of ARGUS is to transform heterogeneous, high-volume information streams into a limited number of structured, geo-referenced incidents with explicit confidence values and evidence links, thereby supporting decision-making under uncertainty. The architecture is modality-agnostic and extensible. ARGUS does not assume that all data sources are available in every deployment. Instead, it is designed to operate under partial observability, tolerating missing, delayed, or noisy inputs while explicitly reflecting evidential uncertainty. This design principle addresses realities of crisis management and dual-use environments, where data availability, quality, and governance constraints vary across actors and situations. Conceptually, ARGUS follows a four-stage processing model: **(1) multimodal signal ingestion, (2) normalization and semantic interpretation, (3) evidence-based incident fusion,** and **(4) prioritization with confidence-supported recommendations** (cf. Figure 1).

At the multimodal signal ingestion step, ARGUS acquires a broad range of potential data sources, including meteorological and hydrological services, traffic and closure feeds, port and gate systems, remote sensing platforms (e.g., UAV, SAR, optical imagery), emergency communication audio, news outlets, and public social media. All incoming signals are treated as observations associated with timestamps, spatial references (explicit or implicit), and modality-specific reliability characteristics.

Normalization and semantic interpretation transform raw signals into a shared internal representation. This process extracts event-type hypotheses, spatial references, temporal markers, and preliminary confidence values. Event-type normalization maps heterogeneous observations into a common operational taxonomy (e.g., flooding, access restriction, security-relevant disruption), enabling consistent downstream fusion across modalities. Spatial interpretation relies on explicit coordinates, matching against structured geographic and infrastructure reference datasets, infrastructure identifiers, or graph-based context. For text and audio-derived reports without explicit coordinates, spatial assignment is based on constrained geocoding against known place names, route identifiers, and predefined corridor polygons. If a report is spatially underspecified, the system anchors it at corridor level rather than claiming precise point-level localization.

Incident fusion represents the analytical core of ARGUS. Normalized signals are aggregated into candidate incidents based on configurable temporal, spatial, and semantic compatibility constraints. Fusion is evidence-driven rather than majority-based, i.e., each incident maintains an explicit evidence set that records contributing observations, their modalities, and their inferred relationships. Related event categories can be aggregated into a single incident based on explicit, predefined rules that specify event-type compatibility and incident-level labeling precedence, while all contributing observations remain preserved as traceable evidence. ARGUS supports the spatial contextualization of incidents by mapping them to operationally significant infrastructure structures such as critical transport corridors (e.g., Wilhelmshaven – Oldenburg rail corridor, Road Corridor A28), port access routes, gate zones, and military-relevant infrastructure segments. The framework incorporates a predefined corridor model representing transport infrastructure of both civil and defense relevance in the Jade Region (Wilhelmshaven, Germany). These structures are represented as polygons or graphs and allow the system to reason about impacts at the level of movement and access rather than isolated point events. This design aligns situational awareness outputs with routing, access-control, and scheduling workflows. Each fused incident is assigned a confidence estimate reflecting evidential strength, diversity, and consistency. Confidence calibration is designed to be interpretable and conservative, i.e., confirmation across modalities increases confidence, while contradictory observations reduce it without suppressing the incident itself. Confidence is treated as a communicative signal for human operators rather than as a hard decision criterion.

Finally, ARGUS generates a limited set of prioritized response suggestions intended to support operator deliberation rather than prescribe actions. These suggestions are derived from incident type, spatial context, and evidential patterns and are intended to support human decision-making. All outputs are deployed via interoperable interfaces, including structured JSON objects and geospatial layers (e.g., GeoJSON), enabling integration into existing control-center dashboards and decision-support tools.

PROOF-OF-CONCEPT AND FIRST RESULTS

The proof-of-concept (PoC) presented in this paper represents a partial instantiation of the generic ARGUS architecture for a specific geographic and operational context. It does not implement the full range of modalities, data sources, or functional components envisioned in the aforementioned system approach. Instead, the PoC focuses on a subset of capabilities that are particularly relevant for evaluating multimodal fusion behavior, confidence calibration, and corridor-based situational awareness under controlled conditions. The PoC is instantiated for the port region of Wilhelmshaven, Germany, and integrates a limited set of representative input modalities, including simulated emergency communication audio (BOS), simulated social media posts and news headlines, and simulated hydrological sensor signals (water levels). All incidents, modalities, and detections in the current PoC are simulated and generated from scenario-based ground truth using a seeded stochastic framework.

Scenarios

Across the two aforementioned scenarios ”storm surge and flooding” (S1) and ”SPOD access” (S2) 100 incidents were injected with known ground truth attributes. For both scenarios S1 and S2, in each case 50 injected incidents include event type (flooding, closure), spatial location (road corridors, urban access zones), time windows and expected modalities (water level, BOS audio reports, social media reports, news headlines). Across both scenarios, 198 raw detections were generated across all implemented modalities (e.g., BOS audio, social media, news, water levels). The simulated time horizon spans approximately 50 hours, beginning at 24 January 2026, 00:00 UTC, with each incident assigned a five-minute active window. All incidents were spatially distributed within the Wilhelmshaven transport network. Ground truth includes incident-level attributes and detection-level logs to support evaluation. For each incident, ground truth specifies its intended semantic category and spatial-temporal extent. While scenarios implicitly vary in modality coverage (e.g., single-modality vs. multimodality support), no explicit conflict labels are encoded in the current ground truth specification. This structure enables systematic evaluation of consolidation behavior, spatial-temporal matching accuracy, and confidence–correctness under varying modality combinations.

Outcome	N	Mean Confidence	Median Confidence
Correct	56	0.66	0.78
Partial	13	0.62	0.58
Incorrect	15	0.65	0.50

Table 2. Relationship between cluster correctness and incident-level confidence estimates. The table illustrates how heterogeneous detections are aggregated into incidents, highlighting consolidation effectiveness as well as over-fusion behavior.

Confidence Bin	N	% Correct	%Partial	%Incorrect
[0.0-0.3]	5	100.0%	0.0%	0.0%
[0.3-0.7]	37	54.1%	24.3%	21.6%
[0.7-1.0]	42	73.8%	9.5%	16.7%

Table 3. Relationship between incident-level confidence estimates and correctness in the proof-of-concept evaluation. Percentage of correct, partial and incorrect matches for confidence bins.

Data sources and modalities

The PoC integrates representative data sources spanning structured and unstructured modalities (cf. Table 1). The dataset was generated using a ground-truth-driven stochastic simulation framework with fixed random seeds for reproducibility. Incident-level ground truth defined event type, approximate location, active time window, and expected modality coverage. Individual modality-specific detections were then synthesized from this ground truth with controlled spatial jitter, temporal offsets, and wording variation to emulate realistic reporting uncertainty. Simulated water-level measurements provide time-stamped, structured sensor signals associated with flood risk. These signals serve as high-confidence, structured evidence that can reinforce or contextualize flooding-related incidents during fusion. Short text entries from simulated social media posts (e.g., "At JadeWeserPort: The underpass is completely flooded.") and news headlines (e.g., "Flooding near Wilhelmshaven (JadeWeserPort/ Naval Base): Traffic partially suspended") provide early, but potentially noisy, indicators of disruptions such as road flooding, temporary closures, or access disturbances. Simulated BOS radio messages (e.g., "Wilhelmshaven control center. A29 exit, completely flooded, vehicles turning around") represent short operational communications typical of emergency and security coordination. Audio inputs are transcribed using automatic speech recognition and subsequently classified into event categories. These example utterances are simulated operational-style messages. In the current PoC, geolocation is not intended as dispatch-grade point localization; instead, the system performs constrained geographic reference extraction and corridor-level assignment based on infrastructure context. Predefined GeoJSON corridor polygons encode critical urban access routes, gate zones, and port-related segments. During fusion, clustered incidents are spatially mapped to corridor groups, enabling corridor-level situational representation and operational interpretation.

Implementation

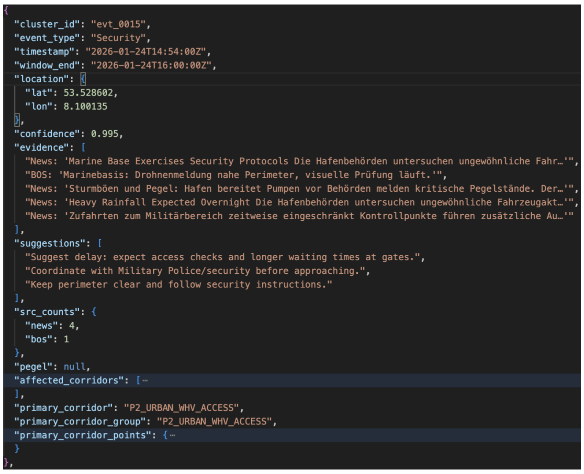
The PoC is implemented as a modular Python-based pipeline[3] that normalizes each input modality into a shared event schema comprising timestamp, event-type hypothesis, location estimate, source type, and source-level confidence. Normalized signals are clustered using fixed temporal (60 minutes) and spatial (3 km) thresholds combined with event-type compatibility rules. Fusion decisions are implicitly recorded through the produced cluster objects and evaluation outputs with assigned cluster ID, start and end time window, aggregated centroid location, source counts per modality, evidence list, and corridor assignment. Each fused incident includes a computed confidence value bounded to the interval [0,1]. Incident-level confidence values are computed deterministically from interpretable heuristic components based on average source-level confidence, cross-modal agreement, and number of supporting signals. Up to three rule-based response recommendations are generated per incident. Fused incidents are exported as structured JSON and GeoJSON objects. Each recommendation is stored together with the associated cluster ID, event type, and evidence trails that triggered the rule.

---

[3] The full implementation of the ARGUS fusion pipeline, including synthetic data generation, multimodal normalization, and evaluation scripts, is implemented in Python. The link to the code repository will be provided.

```
{
  "cluster_id": "evt_0006",
  "event_type": "Flooding",
  "timestamp": "2026-01-24T09:35:00Z",
  "window_end": "2026-01-24T09:38:00Z",
  "location": {
    "lat": 53.573766,
    "lon": 8.139293
  },
  "confidence": 0.77,
  "evidence": [
    "BOS: 'JadeWeserPort Terminal: Wasser läuft über Fahrbahn, LKW-Zufahrt eingeschränkt.'",
    "Social: 'Terminalzufahrt am JadeWeserPort: Wasser steht auf der Straße.'"
  ],
  "suggestions": [
    "Suggest re-route: avoid affected roads and follow detours."
  ],
  "src_counts": {
    "bos": 1,
    "social": 1
  },
  "pegel": {
    "max_m": null,
    "critical": false
  },
  "affected_corridors": [⋯
  ],
  "primary_corridor": "P2_URBAN_WHV_ACCESS",
  "primary_corridor_group": "P2_URBAN_WHV_ACCESS",
  "primary_corridor_points": {⋯
  }
},
```

Figure 2. Example of a fused flooding incident anchored to an access corridor, combining BOS audio and social media evidence into a single incident object with explicit confidence and traceable evidence links.



```
{
  "cluster_id": "evt_0015",
  "event_type": "Security",
  "timestamp": "2026-01-24T14:54:00Z",
  "window_end": "2026-01-24T16:00:00Z",
  "location": {
    "lat": 53.528602,
    "lon": 8.100135
  },
  "confidence": 0.995,
  "evidence": [
    "News: 'Marine Base Exercises Security Protocols Die Hafenbehörden untersuchen ungewöhnliche Fahr…'",
    "BOS: 'Marinebasis: Drohnenmeldung nahe Perimeter, visuelle Prüfung läuft.'",
    "News: 'Sturmböen und Pegel: Hafen bereitet Pumpen vor Behörden melden kritische Pegelstände. Der…'",
    "News: 'Heavy Rainfall Expected Overnight Die Hafenbehörden untersuchen ungewöhnliche Fahrzeugakt…'",
    "News: 'Zufahrten zum Militärbereich zeitweise eingeschränkt Kontrollpunkte führen zusätzliche Au…'"
  ],
  "suggestions": [
    "Suggest delay: expect access checks and longer waiting times at gates.",
    "Coordinate with Military Police/security before approaching.",
    "Keep perimeter clear and follow security instructions."
  ],
  "src_counts": {
    "news": 4,
    "bos": 1
  },
  "pegel": null,
  "affected_corridors": [⋯
  ],
  "primary_corridor": "P2_URBAN_WHV_ACCESS",
  "primary_corridor_group": "P2_URBAN_WHV_ACCESS",
  "primary_corridor_points": {⋯
  }
},
```

Figure 3. Example of a fused security-related incident integrating heterogeneous reports across modalities, illustrating high-confidence aggregation as well as corridor-level spatial anchoring for access control decisions.

Results

Across current scenario runs, the PoC ingested 198 raw signals and produced 84 fused incidents. Regarding RQ1 (*To what extent can multimodal fusion consolidate heterogeneous port-disruption signals into consistent, georeferenced incidents and thereby reduce duplicate or conflicting alerts?*), the PoC achieves a compression ratio of 2.36, indicating that more than two heterogeneous signals are consolidated into a single fused incident on average. The incident match rate against injected ground truth is 0.81, with exact matches accounting for 0.64 and partial matches for 0.17. This indicates that the fusion engine is able to recover the majority of injected incidents while maintaining reasonable semantic alignment. After fusion, we have 84 fused clusters and 35 unmatched clusters. At the same time, 51 over-fusion cases are observed, where multiple injected incidents are merged into a single cluster, suggesting that the current spatial and temporal thresholds are overly permissive in dense situations. This behavior also indicates the need to distinguish between spatial proximity and operational relevance, especially in dense transport environments where nearby disruptions may not affect the same access corridor. Figure 2 shows an example of a flooding incident fused from two evidence items spanning two modalities (BOS audio and social media) anchored to corridor segment ”P2 URBAN WHV ACCESS”. The exported GeoJSON description includes confidence 0.77 and links to the underlying evidence trail. Figure 3 illustrates a security incident also anchored in this corridor segment fused from five evidence items spanning two modalities (news and BOS audio) with confidence 0.995.

For RQ2 (*How well do confidence values and evidence trails correspond to scenario ground truth, and how effectively do they communicate uncertainty when modalities disagree?*), confidence values show a positive, though imperfect, association with correctness (cf. Table 2). Correct clusters exhibit higher mean and median confidence values than partial or incorrect clusters, and higher-confidence bins contain a larger proportion of correct incidents (cf. Table 3). In the high-confidence bin [0.7-1.0], 73.8% of incidents matched ground truth, compared to 54.1% in the low-confidence bin [0.3-0.7]. These results indicate a positive association between confidence scores and correctness, suggesting that the confidence estimation mechanism provides meaningful reliability. For each fused incident, evidence trail completeness was measured using number of evidence items and modality diversity, i.e., distinct contributing modalities. The average evidence items per cluster is 2.05 and the average of modality diversity is 1.57. Most clusters are supported by one or two modalities, with a smaller subset integrating three modalities (e.g., BOS + news + water level). When contradictory evidence patterns occur, confidence values are consistently reduced (penalty term), signaling uncertainty rather than suppressing incidents. The contradiction detection heuristic identified five clusters containing conflicting evidence patterns. All such clusters received low confidence scores (< 0.60), indicating that the contradiction penalty mechanism operates as intended by reducing confidence in ambiguous cases.

With respect to RQ3 (*What integration, interoperability (e.g., JSON/GeoJSON), and governance requirements arise for multimodal decision support in a dual-use port environment?*), the PoC highlights several integration

Identified requirement	Operational rationale	Observed PoC evidence	Design challenge
Evidence-level provenance	Auditability and accountability in dual-use settings	Evidence trails store text snippets without structured identifiers	Design a provenance schema including source type, timestamps, and reliability classes
Corridor-based spatial representation	Operational relevance for routing and access control	Incidents anchored to GeoJSON corridors improved interpretability	Refine corridor definitions and integrate dynamic corridor availability
Confidence thresholds for surfacing incidents	Avoid alert overload and misprioritization	High-confidence but incorrect clusters observed	Define adaptive or context-aware confidence threshold policies
Recommendation provenance	Accountability for decision-support outputs	Recommendations lack explicit rule traceability	Attach recommendation rationales linking rules to triggering evidence
Data governance separation	Handling sensitive and open-source inputs	PoC mixes simulated operational and open-source data	Define access control, retention, and separation mechanisms for live deployment

Table 4. Integration, interoperability, and governance requirements derived from observed PoC behavior, illustrating how multimodal fusion characteristics translate into concrete design and research challenges for dual-use port environments.

and governance requirements for multimodal decision support in dual-use port environments. These requirements emerge directly from observed fusion behavior and are summarized in Table 4.

IMPLICATIONS, LIMITATIONS AND FUTURE WORK

The PoC results indicate that consolidating heterogeneous signals into corridor-anchored incident objects can reduce alert fragmentation while preserving access to underlying evidence (cf. Table 4). This supports prior work on crisis management systems and suggests that multimodal fusion can shift operator attention from raw alerts toward structured situational understanding. Confidence values in ARGUS function primarily as communicative cues rather than strict decision thresholds. Their positive association with correctness and their sensitivity to contradictory evidence suggest that confidence values can support operator judgment under uncertainty rather than replace it.

The PoC has several limitations. All data and scenarios are simulated, and no claims can be made regarding real-world performance, robustness, or latency. Fusion behavior is sensitive to manually defined spatial and temporal thresholds, leading to observable over-fusion. Confidence calibration is heuristic due to missing labeled historic data and does not adapt to context or operator feedback. Governance aspects such as provenance granularity, retention policies, and access control remain only partially addressed. However, over-fusion and high-confidence incorrect clusters are treated as diagnostic signals that motivate adaptive fusion strategies and operator-in-the-loop designs. The current evaluation focuses on recovery and consolidation of injected incidents, but it does not yet separately test the system's ability to reject geographically proximate yet operationally irrelevant events. In the port-access domain, this distinction is crucial and will be addressed in future evaluation. Future work will validate such scenarios against historical real-world disruption cases and, where available, archived or live operational data streams. Here, the focus will be on operator-centered evaluations, adaptive fusion, explicit false-positive and relevance-filtering tests, and integration with historical as well as live data feeds to assess end-to-end workflow impact.

CONCLUSION

This work-in-progress paper presented ARGUS, a generic multimodal decision-support architecture for incident-centric situational awareness in disrupted port-access and transport environments. ARGUS is designed to consolidate heterogeneous and partially unreliable information sources into structured, geo-referenced incidents with explicit confidence values and traceable evidence, supporting decision-making under uncertainty. A proof-of-concept instantiation for the port region of Wilhelmshaven demonstrated the feasibility of key architectural elements, including multimodal normalization, evidence-based incident fusion, confidence calibration, and corridor-level spatial anchoring. The exploratory results highlight both the potential benefits of multimodal fusion for reducing alert fragmentation and the challenges associated with over-fusion, confidence interpretation, and governance in dual-use contexts. The findings underline the importance of explainability, configurability, and governance-aware design for multimodal decision-support systems in crisis management settings.

REFERENCES

Alam, F., Ofli, F., & Imran, M. (2018). Processing social media images by combining human and machine computing during crises. *International Journal of Human–Computer Interaction*, *34*(4), 311–327. https://doi.org/10.1080/10447318.2018.1427831

Algiriyage, N., Prasanna, R., Stock, K., Doyle, E. E., & Johnston, D. (2022). Multi-source multimodal data and deep learning for disaster response: A systematic review. *SN Computer Science*, *3*(1), 92.

Braun, B., Montenbruck, O., Markgraf, M., Hrschgen-Eggers, M., & Kirchhartz, R. (2025). Gnss jamming observed on sounding rocket flights from northern scandinavia. *Engineering Proceedings*, *88*(1). https://doi.org/10.3390/engproc2025088055

Bunker, D. (2020). Who do you trust? the digital destruction of shared situational awareness and the covid-19 infodemic. *International Journal of Information Management*, *55*, 102201. https://doi.org/https://doi.org/10.1016/j.ijinfomgt.2020.102201

Comes, T., Hiete, M., Wijngaards, N., & Schultmann, F. (2011). Decision maps: A framework for multi-criteria decision support under severe uncertainty. *Decision Support Systems*, *52*(1), 108–118. https://doi.org/10.1016/j.dss.2011.05.008

Ghaffarian, S., Taghikhah, F. R., & Maier, H. R. (2023). Explainable artificial intelligence in disaster risk management: Achievements and prospective futures. *International Journal of Disaster Risk Reduction*, *98*, 104123. https://doi.org/https://doi.org/10.1016/j.ijdrr.2023.104123

Green, J., Haigh, I. D., Quinn, N., Neal, J., Wahl, T., Wood, M., Eilander, D., de Ruiter, M., Ward, P., & Camus, P. (2025). Review article: A comprehensive review of compound flooding literature with a focus on coastal and estuarine regions. *Natural Hazards and Earth System Sciences*, *25*, 747–816. https://doi.org/10.5194/nhess-25-747-2025

Havas, C., Resch, B., Francalanci, C., & Pernici, B. (2017). E2mc: Improving emergency management service practice through social media and crowdsourcing analysis in near real time. *Sensors*, *17*(12), 2766. https://doi.org/10.3390/s17122766

Kaufhold, M.-A., Rupp, N., Reuter, C., & Habdank, M. (2020). Mitigating information overload in social media during conflicts and crises: Design and evaluation of a cross-platform alerting system. *Behaviour & Information Technology*, *39*(3), 319–342. https://doi.org/10.1080/0144929X.2019.1620334

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive nlp tasks. *NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems*.

Luna, S., & Pennock, M. J. (2018). Social media applications and emergency management: A literature review and research agenda. *International Journal of Disaster Risk Reduction*, *28*, 565–577. https://doi.org/https://doi.org/10.1016/j.ijdrr.2018.01.006

Nespeca, V., Comes, T., Meesters, K., & Brazier, F. (2020). Towards coordinated self-organization: An actor-centered framework for the design of disaster management information systems. *International Journal of Disaster Risk Reduction*, *51*, 101887. https://doi.org/10.1016/j.ijdrr.2020.101887

Newman, J. P., Maier, H. R., Riddell, G. A., Zecchin, A. C., Daniell, J. E., Schaefer, A. M., van Delden, H., Khazai, B., O'Flaherty, M. J., & Newland, C. P. (2017). A review of decision support systems for natural hazard risk reduction: Current state and future directions. *Environmental Modelling & Software*, *96*, 378–409. https://doi.org/10.1016/j.envsoft.2017.06.042

Opach, T., Rød, J. K., & Munkvold, B. E. (2023). Map functions to facilitate situational awareness during emergency events. *Cartography and Geographic Information Science*, *50*, 546–561. https://doi.org/10.1080/15230406.2023.2264759

Reuter, C., Hughes, A. L., & Kaufhold, M.-A. (2018). Social media in crisis management: An evaluation and analysis of crisis informatics research. *International Journal of Human–Computer Interaction*, *34*(4), 280–294. https://doi.org/10.1080/10447318.2018.1427832

Reuter, C., Stieglitz, S., & Imran, M. (2020). Social media in conflicts and crises. *Behaviour & Information Technology*, *39*(3), 241–251. https://doi.org/10.1080/0144929X.2019.1629025

Scalia, G., Francalanci, C., & Pernici, B. (2022). Cime: Context-aware geolocation of emergency-related posts. *GeoInformatica*, *26*, 125–157. https://doi.org/10.1007/s10707-021-00446-x

Schätter, F., Hansen, O., Wiens, M., & Schultmann, F. (2019). A decision support methodology for a disaster-caused business continuity management. *Decision Support Systems*, *118*, 10–20. https://doi.org/10.1016/j.dss.2018.12.006

Steen-Tveit, K., & Erik Munkvold, B. (2021). From common operational picture to common situational understanding: An analysis based on practitioner perspectives. *Safety Science*, *142*, 105381. https://doi.org/https://doi.org/10.1016/j.ssci.2021.105381

Van de Walle, B., Brugghemans, B., & Comes, T. (2016). Improving situation awareness in crisis response teams: An experimental analysis of enriched information and centralized coordination. *International Journal of Human-Computer Studies*, *95*, 66–79. https://doi.org/https://doi.org/10.1016/j.ijhcs.2016.05.001

Zhao, F., Zhang, C., & Geng, B. (2024). Deep multimodal data fusion. *ACM Comput. Surv.*, *56*(9). https://doi.org/10.1145/3649447