

DGI Compendium

2026



University
of Exeter

Exeter Defence,
Security and Resilience

DDRC
DEFENCE DATA RESEARCH CENTRE



Contents

Foreward Zena Wood	3
Leveraging Federated Learning for Secure Data Fusion and Decision Advantage Michael Mattarock	7
Data-Driven Digital Twins for Maritime Infrastructure: Building Trustworthy Intelligence for Resilient and Sustainable Ports Mingyu Zhu, Carlos Calderon, Alastair Ford, Craig Robinson, Jiayi Jin	11
From Multidomain Warfare to Cognitive Resilience: Toward an Integrated NATO Decision Architecture Eugenia Hernández	16
Project NeuroGuards: Quantifying Human Susceptibility to Misinformation Through Neuro-Cognitive Evidence Hamoon Khelghat-Doost, Scott Kidd, Patrick Finnegan	20
Co-Evolutionary Optimisation for Decision Advantage in Adversarial Environments Jack Richings, Kureha Yamaguchi, Victoria Nockles	24
Interoperability Issues at the Strategic Level Frances Tammer	30

Foreward

Zena Wood

Achieving and maintaining information advantage requires data to be gathered and integrated from multiple sources at the speed of relevance. The fused data can then be analysed for improved situational awareness and decision advantage. However, data fusion for information advantage is an increasingly complex task. Multiple sources are available, all of which must be assessed for credibility and authenticity. The rapidly increasing number of data sources increases the likelihood of duplicate data be created and the risk of circular reporting. With the volumes of data rapidly expanding, AI could allow the data to be harnessed, to enable effective and improved decision-making. Although AI can bring many opportunities, it also brings many challenges (e.g., bias, validity). There are also concerns about the reliance on AI, particularly when it comes to critical decision-making.

This compendium contains six articles that address: what successful data fusion solutions exist that allow data to be fused from multiple sources to improve situational awareness and understanding with the goal of achieving decision advantage?

Authors were asked to consider at

least one of the following topics:

- How can such solutions detect and/or mitigate vulnerability to misinformation/disinformation?
- How can such solutions be used in both the classified and unclassified domains?
- How can such solutions prevent circular reporting and ensure decisions are made on good quality data?
- What is the role of open-source data?

The articles included in this compendium can be considered as covering three different topics: architectures and models, solution design and strategy.

Architectures and Models

Dr Michael Mattarock highlights that instead of dealing with data scarcity, the new challenge is how to fuse heterogeneous, and often untrusted data sources, to allow the provision of actionable insights at the speed of relevance. The article proposes the use of Federated Learning, an approach that allows raw data to remain at its source with only the model centrally aggregated. Such approaches have been used in Cyber Security for intrusion detection and the fusing of multi-source data for decision-making. The article discusses the benefits of the approach particularly around the practical data fusion challenges relating to circular reporting and propagation of misinformation.

Dr Mingyu Zhu and his colleagues focus on delivering a Digital Twin solution for dry-bulk terminals in maritime ports, a crucial part of our infrastructure and a backbone of global trade. Digital Twins allow physical infrastructures to be linked with data-driven models that can then be used for simulation, optimisation and prediction. Unlike other solutions, the presented Digital Twin models the independencies with environmental conditions, building energy systems and human operations through a unified data architecture. The model also recognises that accurate high-resolution real-time data streams are always going to be continuously available. Cross-validation and synthetic data auditing are utilised to

reduce the risk of circular reporting, duplication and misinformation. As the authors note, data-driven transformation requires using existing data more intelligently and transparently, not solely on collecting more data.

Professor Eugenia Hernández highlights the move away from a period of relative predictability when considering international alliances and the necessity to develop joint capabilities. A core question to this approach is 'how can command and control systems integrate such complexity and still enable timely, coherent, and ethical decision-making'? Professor Hernández believes that the answer lies at the intersection of the human factor and technology. The article proposes a NATO-governed fusion platform, anchored in existing NATO structures, that would offer collective validation of data, shared ethical oversight, doctrinal coherence and interoperability by design. Such a platform would strengthen resilience of NATO and its strategic edge in future conflicts. Instead of replacing national capabilities, it is designed to 'provide a common operational layer' that would enable coherence in decisions, trust and synchronization.

Solution Design

Dr Hamoon Khelghat-Doost and colleagues present their project, NeuroGuards, a brain-to-eye AI powered system that can identify when an audience becomes more

receptive to manipulative messages. AI-driven analysis of neuro-cognitive and behaviour signals during content exposure is used to measure receptivity. The work considers the conditions under which an audience's epistemic guard is lowered instead of trying to make a judgement on truthfulness or labelling individuals. Applications where the system could be used include testing public interest, designing proportionate counter-messaging strategies, or building uncertainty literacy among analysts.

able national level intelligence architectures. However, interoperability is still a challenge. Although technical challenges exist (e.g., cyber proof firewalls and data management capabilities), much wider systematic and operational enabling factors must be considered. The article outlines the interoperability issues that exist at a strategic level, calling for better private/public sector collaborations as well as the continued need to work with international partners.

Sensor networks will play an increasingly important role in information advantage. However, these networks must be resilient to the adaptive behaviour of intelligent adversaries. Effective data fusion requires quality inputs. Dr Jack Richings and his co-authors present a framework designed to optimise heterogeneous sensor network topologies by treating sensor placement as a dynamic, adversarial game. The article highlights the adaptability and scalability of the solution and its ability to ensure continued high-fidelity inputs with improved decision rates.

Strategy

Professor Frances Tammer focuses on the need to utilise the rapidly growing amount of OSINT data and leverage generative AI/human machine learning. OSINT data will play a crucial role in fully interoper-



Leveraging Federated Learning for Secure Data Fusion and Decision Advantage

Michael Mattarock
Carnegie Mellon University

Achieving and sustaining information advantage demands not only the gathering of data from multiple sources, but the rapid integration and fusion of that data to support improved situational awareness and decision advantage. In an era of proliferating sensors, intelligence streams, open-source feeds, and partner-data collaborations, the central challenge has shifted: it is no longer scarcity of data, but rather how to fuse heterogeneous, rapidly arriving, and often untrusted data sources in a way that yields reliable, actionable insight and does so at the speed of relevance.

Federated Learning as an Enabler of Secure Fusion

A compelling alternative arises in the form of Federated Learning (FL) where the raw data remains at respective sources (or enclaves), while model training occurs locally and only model updates (not the raw data) are aggregated centrally. This preserves data locality and supports collaboration across data silos, including across classification boundaries. In cybersecurity and distributed systems contexts, FL has been applied to anomaly detection, intrusion detection across the industrial IoT, and

fusion of multi-source data for decision-making, especially given its ability to train and update significantly faster than traditional centralized learning.

In the context of information advantage, it enables fusion of data without centralizing raw streams, reducing exposure, respecting data ownership/security domains, and enabling faster integration across multiple sources. For example, different agencies might train local models on their own data, then participate in a federated aggregation to build a global model. That global model effectively fuses insights from across the network without exposing raw, sensitive data and allowing each enclave to maintain its independence. This was demonstrated through a recent intrusion detection study where smart-grid operators deployed FL to train anomaly-detection models across multiple utility substations without sharing raw telemetry (ICCK 2025). Similarly, Google's Gboard Keyboard uses FL to update predictive-text and malware-detection across billions of devices without collecting raw user data.

Mitigating Misinformation, Circular Reporting, and Poor-Quality Data

One of the practical challenges in data fusion is the risk of circular reporting (i.e., the same event being reported through multiple channels and then re-ingested) and the propagation of misinformation. FL can assist in mitigating these risks in several ways:

- **Weighted source aggregation:** During federated aggregation the contribution of each node (a participating data source that trains its own local version of the model) can be weighted based on assessed data-quality, provenance trust, historical reliability or statistical consistency. This reduces the influence of low-credibility sources. An example could be a sensor platform, like a UAV streaming imagery.
- **Audit and provenance metadata:** Because only model updates are shared, provenance of the underlying data remains at the edge (where the data lives), such as the case of a forward-deployed sensor. Nodes can attach metadata about source type, duplication indicators, confidence levels, and dataset overlap enabling the fusion system to detect possible circular reporting and duplicate input when training.

- **Hybrid human-in-the-loop oversight:** Federated systems must not be fully automated without oversight, especially in critical decision-making. Outputs of the fused model can be flagged for manual review where confidence is low or inconsistency is detected.
- **Robustness to adversarial manipulation:** The decentralized nature of FL means that adversarial data poisoning or model update manipulation must be addressed through secure aggregation, anomaly detection on updates, and robust aggregator algorithms.

Thus, FL becomes not only a mechanism for decentralized model training but a tool to embed data-quality and trust mechanisms into the fusion process helping to guard against poor-quality data, circular reporting, and misinformation. FL's architecture is also well suited to hybrid domains where some data sources are unclassified, others are classified, and cross-domain fusion is desired while respecting data security constraints.

Achieving Decision Advantage: A Synthesis

FL offers a promising paradigm for advancing data fusion toward decision advantage in cybersecurity and intelligence domains. By enabling decentralized training and secure aggregation, FL supports multi-source fusion

across heterogeneous, distributed data-owners while preserving data sovereignty and reducing exposure risk. When combined with mechanisms for provenance, weighting, anomaly detection and human oversight, FL can mitigate the twin risks of misinformation and circular reporting. This can support detecting circular reporting across text, imagery, and geospatial data, given the current absence of a unified multimodal classifier. Labelled data for circular reporting are more common in text than in imagery datasets, where availability and specificity vary by domain. In intelligence contexts, geolocation is often precise and useful for identifying duplicate reports, whereas open-source streams may lack reliable location data. Overall, federated learning can integrate these heterogeneous sources to improve circular-report detection without requiring centralized access to sensitive raw data.

References

Manzoor, H.U., Shabbir, A., Chen, A., Flynn, D., & Zoha, A. (2024). "A Survey of Security Strategies in Federated Learning: Defending Models, Data, and Privacy." *Future Internet*, 16(10):374. MDPI

Liu, P., Xu, X., & Wang, W. (2022). "Threats, attacks and defenses to federated learning: Issues, taxonomy and perspectives." *Cybersecurity*, 5:4. SpringerOpen

Tang, Y., & Liu, J. (2025). "A Survey on Federated Learning in Heterogeneous Data Environments." *Journal of Cybersecurity*, 3(2):2-11. *Journal of Cybersec*

Chaudhary, D., Rajasegarar, S., & Pokhrel, S. R. (2025). "Towards Adapting Federated & Quantum Machine Learning for Network Intrusion Detection: A Survey."

"Multi-Source Information Fusion for Anomaly Detection in Smart Grids Using Federated Learning." (2025). *ICCK Journal*. ICCK

"Federated Learning for Cybersecurity: A Privacy-Preserving Approach." (2025). *Applied Sciences*, 15(12):6878.



Data-Driven Digital Twins for Maritime Infrastructure: Building Trustworthy Intelligence for Resilient and Sustainable Ports

Mingyu Zhu*, Carlos Calderon**, Alastair Ford**, Craig Robinson**, Jiayi Jin***
*University of Glasgow, **Newcastle University, ***Northumbria University

Maritime ports form the backbone of global trade, essential for social and economic stability and security. Balancing resilience, efficiency, and sustainability has therefore become a defining challenge for port authorities and governments worldwide. Digital Twin offers a pathway to address this challenge by linking physical infrastructure with data-driven models that simulate, predict, and optimise port operations.

However, most existing Digital Twin applications for ports are narrowly scoped and heavily dependent on complete, high-resolution real-time data streams. They often model only single components, while overlooking interdependencies with building energy systems, environmental conditions, and human operations. Moreover, the assumption that continuous, accurate data is readily available does not reflect operational realities in many ports, where sensors may be limited, datasets incomplete, and data-sharing restricted for commercial or security reasons.

This study develops an integrated Digital Twin framework for dry-bulk terminals, designed to function under data uncertainty and partial ob-

servability. Centred on the Port of Tyne in the United Kingdom, the framework provides a holistic platform that connects terminal logistics, energy consumption, and environmental impact through a unified data architecture. It evaluates resilience and sustainability indicators under varying operational scenarios. It is funded by the Defence Science and Technology Laboratory (Dstl). An open-access publication is available in which no sensitive information is disclosed [1]).

Integrating Data

The framework draws on a diverse data ecosystem, including vessel tracking, climate, trading records, and energy usage, enabling a realistic, yet ethically and legally compliant, depiction of port systems (Figure 1). Those open-source data support transparency and replicability, allowing wider validation and collaborative development of the model. At the same time, sensitive datasets can be abstracted into synthetic or probabilistic equivalents to maintain confidentiality while preserving analytical value.

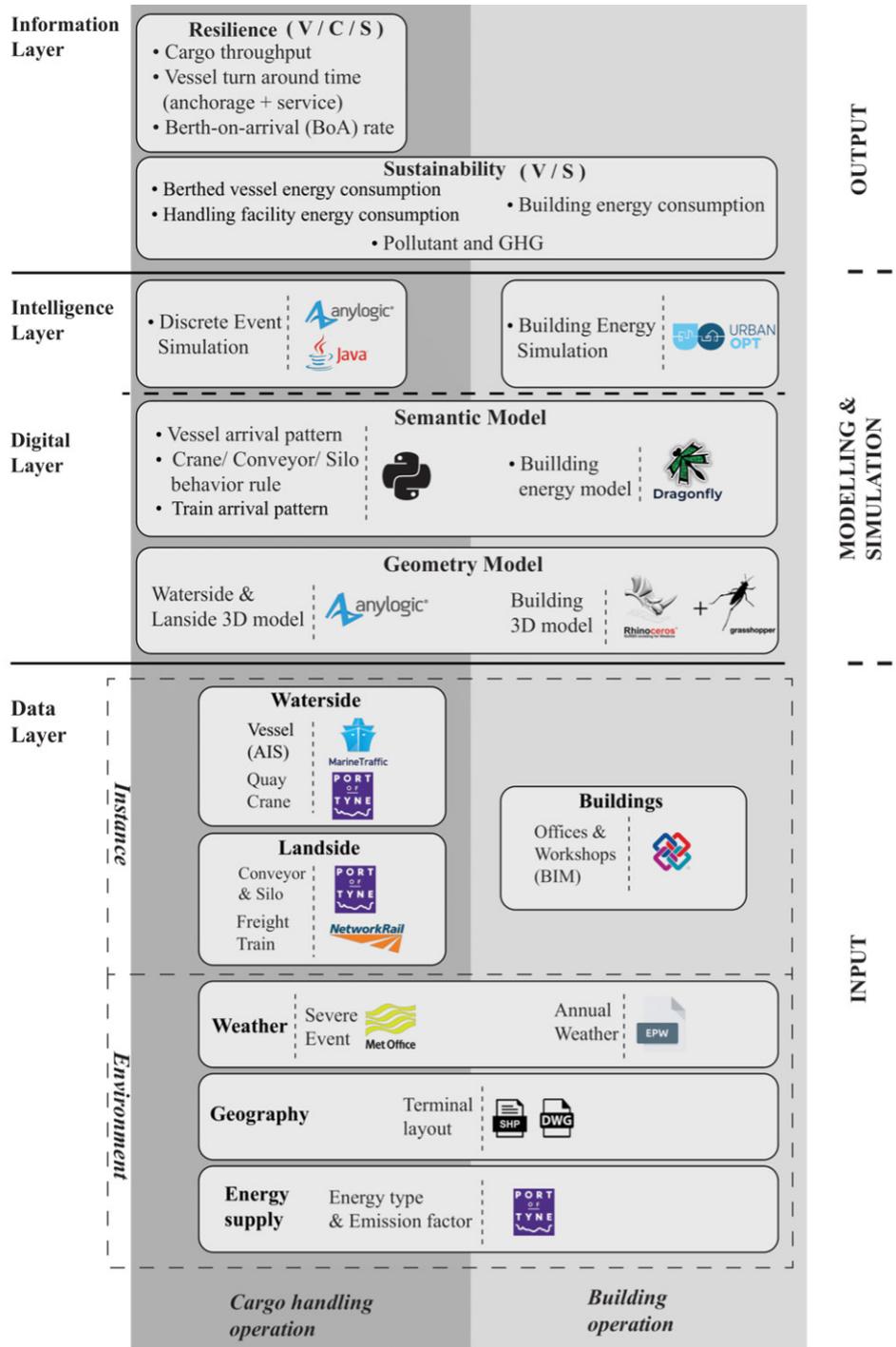


Figure 1 Digital Twin Framework

Mitigating misinformation and ensuring data credibility

In environments where data streams originate from multiple actors and sources, the risk of misinformation, duplication, and circular reporting becomes significant. Two methods were adopted to address this issue:

1. Cross-validation, comparing independent datasets. For example, verifying vessel calls against AIS movements and energy usage to identify inconsistencies or false signals.
2. Synthetic data auditing, where modelled or imputed data is explicitly flagged and periodically reviewed against new empirical evidence.

Through these measures, the Digital Twin environment becomes a trusted analytical space, capable of highlighting data vulnerabilities and reducing over-reliance on uncertain information.

Explainable decision support at multiple scale

Port operations involve decisions ranging from daily scheduling to long-term planning. The Digital Twin provides explainable decision support by linking detailed operational data with higher-level performance indicators.

At the operational scale, each simulation records its data sources and assumptions, ensuring that outcomes, such as energy use or equipment efficiency, are traceable and reproducible. This multi-scale transparency prevents data recycling and builds confidence that decisions are based on verifiable, high-quality evidence.

Data fusion and resilience under uncertainty

Recognising that perfect data availability is unattainable in most operational contexts, the research introduces a hybrid calibration method that combines heuristic search algorithms with synthetic data generation. When real-time data are missing or incomplete, the model infers plausible operational states using probabilistic rules derived from historical or open-source datasets. This adaptive calibration enables the Digital Twin to function in low-data environments, main-

taining analytical continuity without relying on fragile sensor networks.

The simulation results illustrate how alternative operational strategies, such as crane scheduling, electrification of yard vehicles, or retrofitting of port buildings, affect both emissions and throughput. These scenario analyses reveal complex trade-offs between resilience and sustainability, providing actionable insights for port operators seeking to decarbonise without compromising efficiency.

Towards a trustworthy digital ecosystem for ports

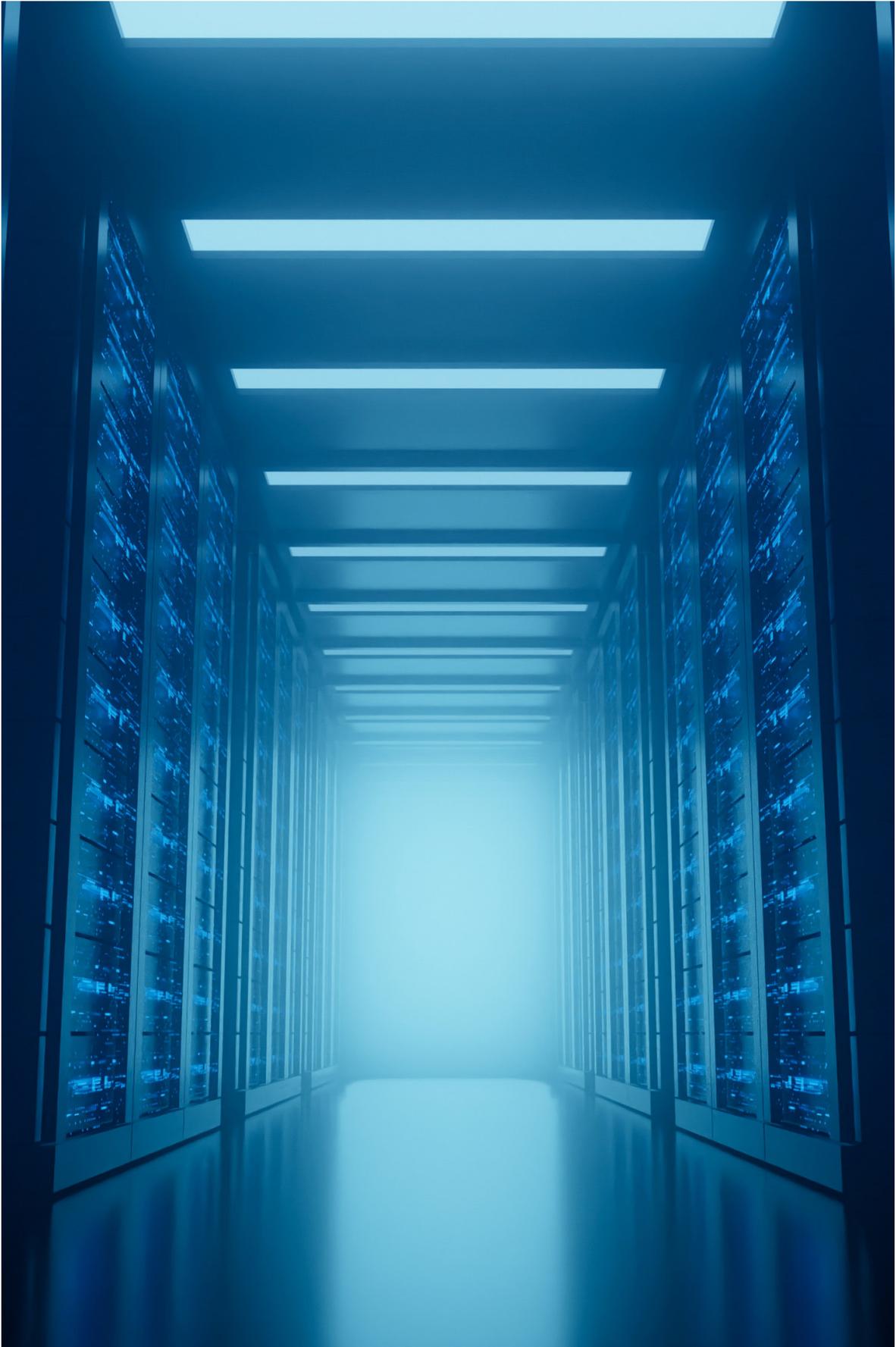
The study ultimately demonstrates that data-driven transformation does not depend solely on acquiring more data, but on using existing data more intelligently and transparently. By embedding cross-validation and interdisciplinary integration within the Digital Twin architecture, ports can build trustworthy intelligence systems that withstand misinformation, operate across classified and unclassified contexts, and avoid the pitfalls of circular reporting.

In an age where ports sit at the intersection of economic resilience, environmental responsibility, and national security, such trustworthy digital in-

frastructures are indispensable. The framework developed here provides a replicable model for critical infrastructure operators worldwide, showing how Digital Twins can become engines not only of operational optimisation but also of data integrity, transparency, and informed decision-making.

Reference:

- [1] Zhu, M., Calderon, C., Ford, A., Robson, C., & Jin, J. (2025). Digital Twin for resilience and sustainability assessment of port facility. *Sustainable and Resilient Infrastructure*, 1-34.



From Multidomain Warfare to Cognitive Resilience: Toward an Integrated NATO Decision Architecture

Eugenia Hernández

Intelligence Analysis Unit (UNINT), Autonomous University of Madrid (UAM)

For more than seven decades, NATO has been more than a military alliance. It has functioned as a community of friends and allies, a convergence of democratic societies, and a cumulative project of political, military, technological, and academic cooperation. Its resilience has not relied solely on material superiority, but on shared norms, trust, interoperability, and the capacity to translate innovation into collective security. This ecosystem flourished under an international order governed by commonly accepted, if minimal, rules that framed sovereignty, deterrence, and the legitimate use of force.

That order is now eroding. After fifty years of relative predictability, the international system is reverting toward historical patterns of power politics, spheres of influence, and strategic rivalry. The return of great power competition obliges the Alliance to reassess its assumptions. In this context, the development of conventional military power and credible joint capabilities is no longer optional. It is a strategic necessity, not in opposition to NATO's values, but as a prerequisite for their protection.

Developing Capability

Joint capability development must be understood as an ethical imperative. Deterrence, and the responsible use of force under Article 5 of the Washington Treaty, requires not only military mass but integrated, interoperable, and trustworthy systems. The challenge is no longer limited to force generation; it lies in how capabilities are connected, governed, and employed in complex operational environments.

If contemporary conflict is understood as multidomain, analytical rigor demands that we begin with the traditional physical domains—land, sea, and air—upon which professional armed forces were built. These domains remain foundational. However, over the last two decades, two additional domains have decisively reshaped the battlespace: cyberspace and outer space. Cyber operations now directly affect superiority across the physical domains, while space—particularly through Low Earth Orbit satellite constellations—has expanded the air domain into a vertically integrated operational layer.

The result is an unprecedented density of interacting capabilities: combat aircraft, drones and drone swarms, satellites and mini-satellites, cyber effects, electromagnetic operations, and space-based sensors. This interaction does not generate efficiency; it generates friction. NATO forces must operate in environments that are contested, degraded, and denied, where communications are disrupted, sensors are unreliable, and decision cycles are compressed.

This reality raises a central operational question: how can command and control systems integrate such complexity and still enable timely, coherent, and ethical decision-making? The answer lies at the intersection of technology and the human factor. Data fusion, artificial intelligence, and resilient digital architectures are essential to achieving decision superiority, but they must remain anchored in meaningful human control and ethical rules of engagement.

This approach underpins initiatives such as The Future of Decision Making, led by Stanford University and the Defence Data Research Centre at the University of Exeter. These efforts seek to develop AI-enabled decision-support systems designed not to replace commanders, but to augment judgment, accelerate sense-making, and enable an integrated Mission Command model grounded in human responsibility.

Not Just Technology

Multidomain warfare cannot be fully understood through physical and technical layers alone. The war in Ukraine has demonstrated that when secure military communications are degraded, civilians become inadvertent intelligence actors. Open platforms such as Telegram and other social networks have provided real-time visibility on troop movements, logistics, and battlefield effects. This phenomenon reveals the operational centrality of the cognitive domain.

The cyber vector connects directly with cognition. Smartphones, social platforms, and open digital ecosystems function simultaneously as sensors, communication channels, and attack surfaces. Influence operations, psychological pressure, fear management, and resilience-building unfold at scale. In this domain, citizens are not passive observers; they are part of the operational environment.

Ukraine has also reinforced a hard truth: while infrastructure—energy, water, transport, airspace control, communications, and spectrum—remains critical, so do the minds and morale of both populations and combatants. Cognitive effects can shape strategic outcomes as decisively as kinetic action.

NATO must therefore develop cognitive capabilities aligned with its moral and legal framework. This requires interoperable fusion platforms capable of integrating traditional ISR with social, emotional, and narrative indicators to support informed decision-making. While market-based solutions remain indispensable, without an Alliance-level integration layer they risk reinforcing fragmentation.

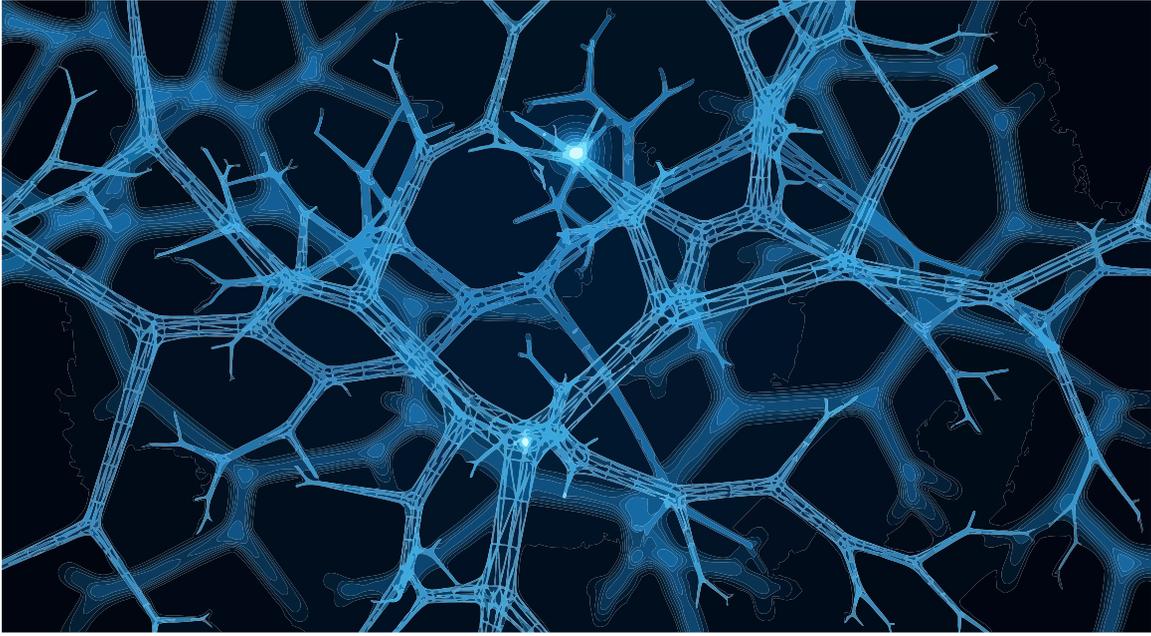
A Fusion Platform

A NATO-governed fusion platform would offer interoperability by design, shared ethical oversight, collective validation of data, and doctrinal coherence. Leveraging Allied industry, innovative SMEs, and academic research networks, such a platform would transform information chaos into structured, ethical, and actionable insight—strengthening NATO’s resilience and preserving its strategic edge in future conflicts.

From a doctrinal perspective, this evolution requires NATO to move beyond ad hoc integration toward a structured, Alliance-wide decision architecture. Multidomain and cognitive operations demand shared frameworks for data fusion, command support, and ethical governance that are interoperable by design and resilient under degraded conditions. A NATO-governed

fusion platform would not replace national capabilities or market-driven innovation, but would provide a common operational layer enabling trust, synchronization, and decision coherence across the Alliance. Anchored in existing NATO structures and supported by academic research, industry, and Centres of Excellence, such an architecture would strengthen collective deterrence while preserving democratic accountability.

In an era defined by strategic competition and cognitive confrontation, NATO’s ability to integrate technology, human judgment, and shared values into its decision-making processes will be decisive in ensuring credible deterrence, operational effectiveness, and long-term resilience.



Project NeuroGuards: Quantifying Human Susceptibility to Misinformation Through Neuro-Cognitive Evidence

Hamoon Khelghat-Doost, Scott Kidd, Patrick Finnegan
University of Lincoln

Scientific Framework and Methodology

NeuroGuards detects and mitigates vulnerability to mis- and disinformation by treating susceptibility as context dependent receptivity to manipulative political content. It measures receptivity during exposure and turns the results into proportionate, auditable decision support for training. Instead of judging truthfulness or labelling individuals, it identifies conditions under which audiences' epistemic guard is lowered — when identity cues, affect or social norm signals heighten predicted uptake — providing insight without judging beliefs or intentions.

Technically, NeuroGuards aligns eye movement dynamics recorded during naturalistic media exposure with neural markers of valuation and self relevance. These markers are established through simultaneous fMRI and eye tracking, then distilled into an eye only surrogate that infers latent receptivity without imaging. The pipeline yields calibrated read outs of predicted acceptance versus resistance during message encoding; an empirically grounded proxy for susceptibility.

Methodologically, the programme links social scientific constructs (framing, group norms, legitimacy cues) to concurrent measures of attention (fixations, dwell time, transition entropy, pupillary dynamics) and latent neural valuation signals. Initial multimodal studies identify cross validated signatures of acceptance and rejection that train predictive models mapping ocular dynamics to latent neural states. The diagnostic flags which elements (for example, grievance laden frames or identity affirming cues) are most likely to elicit uncritical uptake in given audiences, enabling more precise intervention design than blunt exposure metrics.

Mitigation, Safeguards and Cross-National Validation

Mitigation is delivered through governed decision support rather than automated decision making. Because the system pinpoints segments and contexts that raise predicted receptivity, it is used to: (i) pre test and stress test public interest communications; (ii) design proportionate prebunking or counter messaging at the junctures where manipulative frames are likely to take hold; and (iii) build uncertainty

literacy by visualising calibration, confidence and out of distribution warnings alongside interpretable artefacts. Deployment includes model cards, standard operating procedures and an uncertainty first interface; the tool is scoped for training and reflective analysis in prevention and resilience settings, not for real time profiling or automated individual targeting. Safeguards including data minimisation, pseudonymisation (including MRI defacing), fairness diagnostics and immutable audit operationalise Responsible and Trustworthy AI while enabling concrete workflows.

foundations. In unclassified settings such as academia, civil society, local authorities and police training, the eye only diagnostic is deployed at TRL 5–6 as a beta decision support tool for scenario based training, content pre testing and reflective analysis. Trainees examine how changes to framing, norm cues or narrative pacing shift predicted receptivity and practise designing proportionate pre-bunking or corrective messages. All uses rely on model cards, uncertainty visualisations and purpose limitation constraints, with open governance artefacts (e.g., DPIA templates) to support oversight and reproducibility.

Cross national validation across European sites supports culturally sensitive mitigation. Stimulus sets are translated and back translated, and portability is tested so that guidance remains proportionate and locally legitimate. This base helps practitioners anticipate which frames may prove misleading or manipulative in particular media ecologies, reducing the risk that counter measures inadvertently amplify polarisation.

In classified environments such as secure government or law enforcement analysis cells, the same capability is containerised for use in enclaves with access control, immutable audit and network separation. It supports red teaming and structured wargaming of influence threats: analysts evaluate which hostile narratives or presentation strategies are most likely to exploit known vulnerabilities in specified contexts, and test the protective value of alternative public information strategies before release. The tool remains analyst in the loop: it surfaces calibrated indicators and interpretable segment level diagnostics to inform judgement; it never produces automated individual classifications. These design choices implement purpose limitation and proportionality while enabling sensitive threat assessment and preparedness.

Dual-Use Deployment and Governance

NeuroGuards is designed for dual use across classified and unclassified domains through separation of functions, governance and data handling, while maintaining identical scientific

Across domains, governance is harmonised but tailored. The research phase employs layered consent, site level REC/IRB approvals, pseudonymisation at source and compliant cross border transfer; the operational phase enforces purpose specific SOPs, role based access and comprehensive provenance capture. Fairness and robustness are monitored via pre specified subgroup analyses and calibration metrics, with corrective actions gated by promotion criteria embedded in the engineering pipeline. This end to end design ensures that any use—classified or unclassified—remains auditable and legally compliant, while delivering practical value for prevention, training and resilience against mis and disinformation.

Taken together, NeuroGuards offers a single, scientifically grounded solution that detects vulnerability as heightened receptivity under particular message designs and contexts, mitigates that vulnerability through proportionate prebunking, counter messaging and analyst training, and does so within a governance envelope compatible with both unclassified and classified practice, without collapsing into automated individual profiling.

Co-Evolutionary Optimisation for Decision Advantage in Adversarial Environments

Jack Richings, Kureha Yamaguchi, Victoria Nockles

Defence AI Research Centre (DARe), The Alan Turing Institute

Abstract

Achieving information advantage requires sensor networks that remain effective against the adaptive behaviour of intelligent adversaries. We outline a co-evolutionary double oracle framework to optimise the topology of a heterogeneous sensor network. By modelling the problem as a zero-sum game between a sensor placement optimiser and a path-planning adversary, we iteratively construct a defence strategy that ensures data quality for downstream fusion. We utilise genetic algorithms (GAs) as efficient response oracles to explore the parameter space of adversarial trajectories, offering a computationally efficient alternative to deep reinforcement learning (RL) for securing situational awareness.

1. Introduction

Data fusion for decision advantage is strictly bounded by the validity of raw input data. While AI allows data to be harnessed at the speed of relevance, it also introduces risks if inputs are biased or compromised. In security applications, the environment is adversarial [1]. Attackers actively exploit sensor limitations to minimise detection, effectively corrupting the “ground truth” available to the fusion process. Consequently, configurations optimised against historical tactics yield poor quality data against adaptive threats, causing situational awareness failure.

To address this, we propose a co-evolutionary approach that treats sensor placement as a game between

a defender (placing sensors) and an attacker (planning trajectories). The goal is to identify a configuration that ensures decisions are made on good quality data, minimising the escape rate against an ensemble of adaptive tactics.

2. Methods

The Double Oracle Framework

We employ a double oracle framework [2]. An oracle is a subroutine that identifies the optimal counter-strategy against a fixed opponent. We implement these using genetic algorithms (GAs) [3]. GAs are selected to avoid local minima in this non-smooth search space. The process alternates between two routines:

1. The Defender optimises sensor locations to maximise detection against a fixed ensemble of known attacker tactics. 2. The Attacker finds paths minimising detection against the current sensor configuration.

New successful attacks are added to a cumulative ensemble. Subsequently, the defender optimises against the full attack history. This memory mechanism prevents “Rock-Paper-Scissor” cycling common in co-evolutionary systems [4], ensuring the defence increases in robustness monotonically. Our approach is visualised in Figure 1.

Simulation Environment

We utilise a complex grid world where obstacles block movement and line-of-sight. Heterogeneous sensors have variable radii, simulating mixed modalities (e.g., thermal, optical, radar) typical in fusion hubs. We prevent placement near adversary spawn and end points. The Defender GA maximises the fraction of known adversary paths detected. Adversaries are parametrised paths (moving from top to bottom) connecting flexible waypoints. The Attacker GA maximises binary success (undetected arrival) and efficiency (path length). To ensure convergence, we enforce diversity by requiring new attacks to be sufficiently different from the archive.

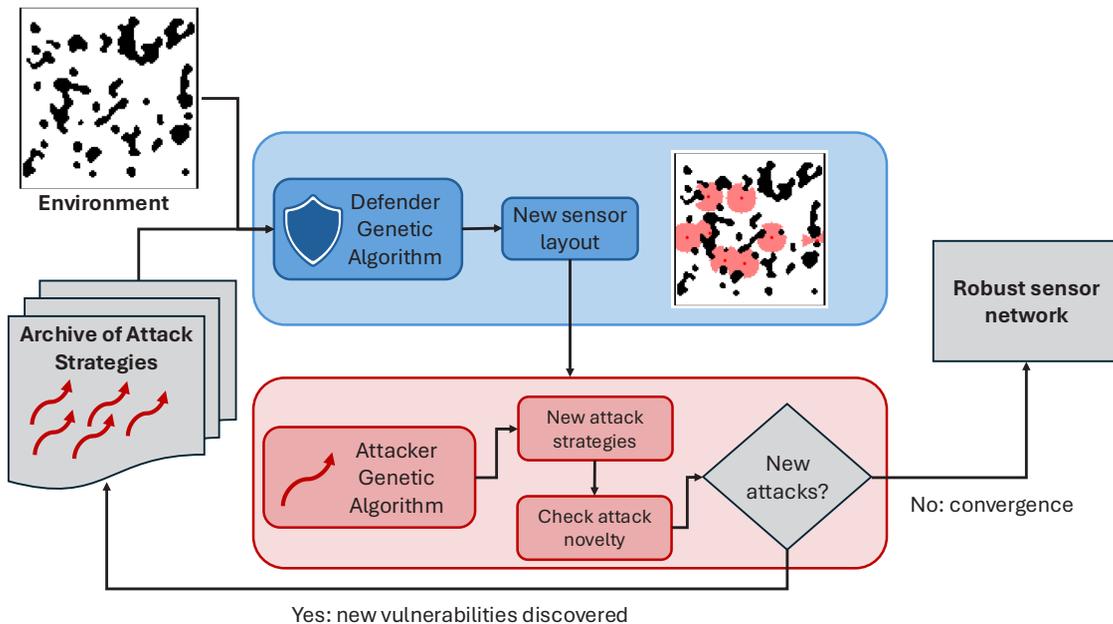
3. Results

Figure 2 compares the co-evolutionary method against a greedy baseline [5]. The co-evolutionary method outperforms the greedy algorithm (86.2% vs. 61.9% detection). Unlike the greedy approach, which places sensors in easily circumvented open areas, the co-evolutionary strategy leverages topological chokepoints. This adversarial process also exposed the leftmost corridor as a vulnerability. A standard GA optimising against historical data would likely overlook such edge cases, resulting in unrealistically high confidence in the network’s security and degraded situational awareness.

4. Limitations and Future Work

Unrealistic Behaviours

A common risk in adversarial optimisation is the generation of unrealistic behaviours relying on simulation artifacts (e.g., pixel-perfect skirting of sensor cones). Mitigating this involves weighting strategies that prioritise key avenues of approach over edge-case exploits. Alternatively, one can fix a subset of sensors (e.g. for critical infrastructure), removing them from the optimisation loop. Furthermore, the approach relies on high simulation fidelity to ensure evolved tactics are physically viable. Additionally, GAs cannot capture the highly reactive strategies



achievable with deep reinforcement learning (RL), although RL is often computationally excessive for static placement [6].

Scalability and Extensions

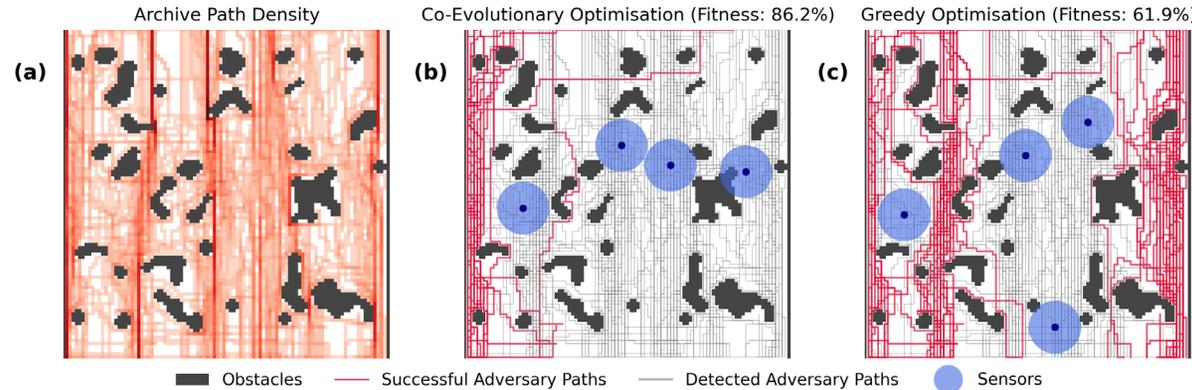
The framework is readily adaptable to diverse environments, sensor counts, and dynamic scenarios. Mobile patrols can be parametrised similarly to adversaries (waypoints and timestamps), allowing the GA to co-evolve patrol routes alongside static sensor placement. The approach is also scalable from a computational perspective. The principal

bottleneck is the fitness evaluation inside the GA loop. Evaluations of individual solutions are independent of one another, meaning that the system can be significantly accelerated with parallel processing. Crucially, the framework decouples optimisation logic from environmental topology, enabling rigorous testing on open-source data before deployment in proprietary environments.

5. Conclusion

Effective data fusion relies on quality inputs. This study demonstrates that sensor networks optimised

against static historical data are inherently fragile to adaptive threats. By shifting sensor placement from a static allocation to a dynamic, adversarial game, we account for intelligent exploitation of topological blind spots. This framework improves detection rates while ensuring continued high-fidelity inputs required for genuine decision advantage.



Acknowledgement

This report was funded by the UK government through the FCDO's Integrated Security Fund (ISF) supported Strategic Stability Programme and conducted in partnership with the Icelandic Ministry of Foreign Affairs and the Icelandic Coast Guard.

References:

- [1] N. Agmon, S. Kraus, G. A. Kaminka, Multi-robot adversarial patrolling: facing a full-knowledge opponent, *Journal of Artificial Intelligence Research* 42 (2011) 887–916.
- [2] H. B. McMahan, G. J. Gordon, A. Blum, Planning in the presence of cost functions controlled by an adversary, in: *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, 2003, pp. 536–543.
- [3] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*, University of Michigan Press, Ann Arbor, MI, 1975, second edition, 1992.
- [4] C. D. Rosin, R. K. Belew, New methods for competitive coevolution, *Evolutionary Computation* 5 (1) (1997) 1–29.
- [5] A. Krause, A. Singh, C. Guestrin, Near-optimal sensor placements in gaussian processes: Theory, efficient algorithms and empirical studies, *Journal of Machine Learning Research* 9 (Feb) (2008) 235–284.
- [6] T. Salimans, J. Ho, X. Chen, S. Sidor, I. Sutskever, Evolution strategies as a scalable alternative to reinforcement learning (2017). arXiv:1703.03864.



Interoperability Issues at the Strategic Level

Frances Tammer

University of Exeter

The Issue

Open-source intelligence (OSINT) plays an important role, as a single intelligence source, in supporting evidential-based national strategic level and international partner defence strategy, policy and warfighting decision making. However, many national level intelligence communities are not innovating sufficiently, constituent parts are not evolving at the same pace, and, in real time, firstly to utilise the explosion of OSINT, and secondly, to leverage human machine learning/generative AI to handle this exponential OSINT growth. Our adversaries are highly likely edging ahead. Interoperability across the whole intelligence enterprise is still a long way off. This has implications for the future expeditious use of geoint within this system, both as a single source and when used in a multi-source and/or all source environment.

AI-assisted intelligence analysis offers significant benefits in deriving insights from unstructured and disparate datasets. This is of even more relevance in the OSINT context, considering the importance of the web as a vector for intelligence datamining, and the magnitude of human

resource required for the burgeoning data discovery, collection and organization. Crucially, these datasets have differing levels of authenticity and accreditation. This has implications for actionable operational and decision making, at the highest level of government and warfighting.

Criteria Setting

To set the stage for immediate, mid-term and long-term interoperability, political will; economic policy; organisational, institutional and funding systems and processes; innovation and resources must all be aligned and sustained. These are indivisible. Pivotal, within this continuous capitalisation strategy, is the promotion of better private/public sector collaboration, as the private sector holds the key to technology innovation. This private industry innovation needs to be incentivised and nested within soundly-based national level fiscal and economic policies. Partnerships cannot be fostered, with as full interoperability as possible, unless procurement constraints and security barriers are managed adequately, with some enhanced level of risk undertaken. Underpinning all of this must be better cyber protection. This requires work-

ing with international partners, given the global nature of our collective adversaries. In addition, to encourage first-class academic research, with a triumvirate approach of academia/industry/government collaboration.

Best Practice

One development, being rolled out across the US Intelligence Community, is the Open AI GPT40-based generative AI system in its top-secret cloud environment. This tool will address security concerns related to large language models (LLMs), typically connected to the internet, by 'air-gapping' the tools to a cloud-based environment, so it cannot be accessed from the web. It will allow secure conversations with a chatbot similar to ChatGPT. Crucially, the AI platform has been developed so that it can read files, but not learn from them in any way that would impact its output. This is a major step towards having enhanced interoperability. It is unknown whether the ultimate objective is for other 5 Five EYES partners to host this system as well. Given that the US is still very much the leader within the 5 EYES Top Secret community, this development theoretically augurs well for its wider implementation. However, secure cloud-based hosting needs to be made available at all security levels – below Secret and above. is already available for Official Sensitive data, with all levels able to speak to each other.

Educating the Decision-makers/Operators

Significantly, training and education about the risks and benefits of AI-enriched intelligence should be communicated to senior strategic decision-makers and war-fighters in national security and defence, to ensure the principles of analytical rigour, transparency and reliability of intelligence reporting and assessments are upheld. They all need help in understanding the new uncertainties introduced by all aspects of AI-enriched intelligence.

Recommendations:

In essence, national level intelligence architectures need to ensure OSINT is fully hosted and integrated within their highly classified digital platforms, otherwise full interoperability will fail. Cyber proof firewalls will be critical as will data management. However, there are much wider systemic and operational enabling factors at stake, which coalesce around –:

- requisite skills and training of the workforce;
- safeguarding against disinformation, misinformation and hallucinations, ensuring that all meta-data is tagged accordingly at the strategic level;

- safeguarding against disinformation, misinformation and hallucinations, ensuring that all metadata is tagged accordingly at the individual agency, tactical and operational levels, which feed into national level defence all source strategic level assessments to mitigate against inherent corruptions;
- the elimination of cognitive bias, working out will this be human-led only or machine -learning or both?
- governance with legal and ethical frameworks in place and cultural mindset transformation.

About DGI

DGI is the longest-running international conference for geospatial intelligence professionals in the world, welcoming delegations from more than 60 nations in 2025. With a focus on providing innovative, interactive formats for learning and networking across three days DGI is a unique opportunity for nations to benchmark their national geospatial intelligence strategies, assess new technologies and forge stronger relationships to deliver collective security.

Editor: Zena Wood

Professor Zena Wood is an Associate Professor in Digital Economy at the University of Exeter and Director of the Defence Data Research Centre (DDRC). DDRC is a transdisciplinary research centre that aims to help the UK defence sector overcome their technical, cultural, and behavioural barriers to the use of data within AI and Data Science.

DDRC is part of the Exeter Defence, Security and Resilience, which brings together over 200 University of Exeter academics from across all disciplines, University Institutes and Research Centres. The network brings together academics from metamaterials, human performance, CBRN and AI through to strategy, Homeland security and policing working at the cutting edge of defence and security research.

DGI Compendium 2027

If you have a problem set connected to the continuous evolution of GEOINT that could form the basis of the next DGI Compendium please contact :

Prof. Zena Wood (z.m.wood2@exeter.ac.uk)

Tom Webber (Tom.Webber@wbr.co.uk)



DGI Compendium

2026



University
of Exeter

Exeter Defence,
Security and Resilience

DDRC
DEFENCE DATA RESEARCH CENTRE