TESTUDO

AUTONOMOUS SWARM OF HETEROGENEOUS RESOURCES IN
INFRASTRUCTURE PROTECTION VIA THREAT PREDICTION AND PREVENTION

2nd Policy Brief

"Operational Readiness in Critical Infrastructure Protection: Early Insights from TESTUDO Pilots''

***Target audience**: EU-level policy officers working on critical infrastructure protection, cybersecurity, resilience, and AI governance.*

Highlights

- *TESTUDO strengthens critical infrastructure protection through an integrated, real-time platform combining AI-enabled detection, predictive analysis, multi-sensor fusion, and coordinated robotic assets to enhance operational readiness.*
- *UC2 pilot validation confirmed high-accuracy visual detection above 80% mean Average Precision (mAP), predictive accuracy of 80% or higher, stable low-latency performance under concurrent sensor feeds, and smooth system integration across components.*
- *User feedback identified usability constraints, including alert overload and limited multi-sensor correlation, highlighting the need for improved prioritisation and clearer visualisation.*
- *While aligned with the key EU frameworks, TESTUDO pilots underscored the absence of harmonised EU validation and certification pathways for AI- and autonomy-enabled critical infrastructure protection tools.*

TESTUDO is a project funded by the European Commission under the Horizon Europe Programme (HORIZON-CL3-2022-INFRA-01) under Grant Agreement No. 101121258.

1. Context and Objective

This second policy brief focuses on operational readiness and pilot validation, assessing how TESTUDO technologies perform in realistic conditions, what multi-sector partners have learned through joint deployment, and how early evidence can translate into an actionable EU policy. Building on the first brief's conceptual framing, this brief focuses on tangible progress, implementation constraints, and pathways to uptake across critical infrastructure ecosystems.

Problem definition: Current protection and resilience approaches remain hampered by fragmented threat intelligence, limited cyber–physical integration, and detection-response cycles that are often too slow for fast-evolving hybrid incidents. Interoperability gaps further constrain coordinated action across operators, sectors and borders.

Objective of this brief: To synthesise initial operational lessons from TESTUDO pilots, covering performance, usability, interoperability and validation, and to outline how these insights can inform ongoing EU policy development, capability building and regulatory preparedness for AI-enabled systems supporting critical infrastructure protection and resilience.

2. Key Achievements and Findings

2.1. Technological Progress

TESTUDO has delivered measurable progress towards an integrated, real-time operational platform that combines automated identification of road hazards and congestion with dashboard-level decision support tailored for critical infrastructure (CI) operators. Static rule sets have been implemented to filter signal from noise, ensuring only actionable information is surfaced to operators, while AI-driven alerting supports human-in-the-loop oversight rather than opaque automation.

In UC2, a Visual Detection model was validated with high accuracy for CI-relevant objects (above 80% mAP) and real-time performance (at least 25 fps), and it was successfully integrated into the wider platform stack, providing synchronised detections, Kafka-based communication, and robust live video stream handling. Thermal video analytics were also integrated and validated, enabling data ingestion from both legacy and embedded cameras and generating real-time alerts, indicating the presence of fire incidents or individuals in the area.

Beyond detection, TESTUDO integrated a Predictive Analysis module that ingests streamed sensor data (including upstream analytics outputs) to support continuous severity assessment, with predictive accuracy validated at 80% or higher.

In addition, the platform supports semi-automated UAV-based 3D reconstruction (with geolocalisedvisualisation), real-time mission planning via an Autonomous Fleet Coordinator across heterogeneous robotic assets (three platforms integrated in UC2), and an enhanced airborne chemical detection capability (T4i DOVER) with built-in self-calibration and real-time alarm transmission.

2.2. Pilot and User Insights

UC2 simulated a tunnel incident in which sensor fusion, threat assessment and decision support must operate under time pressure and degraded visibility, with events detected via CCTV

(visual/thermal/multispectral), escalated to the monitoring centre, and complemented by chemical sensing and response orchestration.

Technical debrief inputs confirm robust system performance: no crashes occurred, remained low even under concurrent CCTV and thermal feeds, and overall system behaviour remained stable as the scenario escalated, with only minor timing mismatches between thermal and visual detections.

Integration across components was smooth; the CCTV provided initial detections; the CBRN sensor captured chemical readings (including acetone); the mission planning incorporated inputs from the monitoring centre; the UGVs autonomously navigated low-visibility and GPS-denied areas to transmit telemetry data to the operators.

At the same time, early testing revealed a clear usability constraint: multi-sensor fusion did not reliably correlate events, leading to an alert overload that undermined prioritisation. This highlighted the need for better clustering of simultaneous alerts, clearer multi-hazard prioritisation (e.g., chemical and fire), and more intuitive overlays of robot paths on the UI.

Operational stakeholders have also pointed to practical enablers (such a well-prepared control room setup supporting coordination) and areas for improvement, including enhanced incident visualisation for decision-makers and more realistic trials enabled by lower traffic and longer durations.

2.3. Policy and Regulatory Alignment

TESTUDO is positioned to support EU resilience and cybersecurity priorities and is being developed with explicit attention to compliance and trustworthy AI. The project aligns its approach with the CER Directive, NIS2, the AI Act, the Data Act, the GDPR, the Free Flow of Non-Personal Data Regulation, and the Cybersecurity Act, while recognising a key systemic barrier: the absence of harmonised EU validation pathways and certification frameworks for AI- and autonomy-enabled CI protection tools.

In operational terms, the compliance mechanisms were embedded through legal mapping and guidance, Data Management Plans, and structured ethics and robustness activities (including an AI ethics workshop and an AI robustness and reliability assessment template). No regulatory barriers were reported during the pilot phase.

3. Evidence and Methodology

Use Case 2 (UC2) required a multi-tier evaluation to demonstrate performance across the full trial stack, individual modules, platform integration, operator workflows, and multi-agency coordination, rather than relying on a single, tool-specific assessment. This aligns with TESTUDO's broader user-guided, iterative development and evaluation logic, in which operational testing and evaluation cycles guide subsequent technical refinement.

The first methodological step was to identify and engage role-holders with direct operational accountability for the trial. In UC2, interviews targeted three central roles: the Trial Owner (overall command, operational realism, and alignment with real emergency management practice), the Technical Coordinator (technical setup, integration, and performance), and the Practitioner Coordinator (coordination of first responders and practitioners). These interviews were designed to complement

questionnaire data by capturing structured qualitative insights into operational, technical, and coordination dimensions, including usability, preparedness, communication, and overall performance. To preserve accuracy, while limiting the respondent's burden, the interviews were kept to no more than 20 minutes and conducted within 1–24 hours after the trial. Findings were analysed qualitatively and triangulated against questionnaires, structured observations, and system logs.

In parallel, evidence was gathered from a broader stakeholder group, including end users (e.g., firefighters and police officers), consortium partners responsible for tool development, and other relevant observers with an interest in the validated solutions. For technical validation, module developers supported the co-definition of KPIs and the collection of performance evidence during the UC2 execution to verify that target metrics were met (e.g., reliability, latency, throughput and operational output quality). Finally, data collection was governed by informed consent and secure handling practices consistent with GDPR principles, including controlled access to recordings, notes and transcripts.

TESTUDO in Numbers

4. Policy Recommendations

Short-term (0–1 year): The Commission and relevant EU agencies should accelerate the establishment of EU-wide testbeds that enable CI operators and public authorities to trial AI-enabled, cyber-physical protection tools in controlled yet realistic environments. These testbeds should prioritise cross-domain (cyber and physical) scenarios and enable rapid iteration on usability, interoperability and operational procedures. In parallel, Horizon Europe should reinforce the structured exchange of pilot data and lessons learned through clustering mechanisms, ensuring that validation evidence, integration patterns and operational insights are shared across projects and sectors.

Medium-term (1–3 years): Member States should embed predictive AI and digital twin capabilities into national transpositions and implementation plans under the CER Directive, supported by guidance that links technology deployment to operational continuity, contingency planning, and multi-agency

coordination. At EU level, there is a clear need to develop common validation standards for autonomous and AI-based CI protection tools, including minimum requirements for performance metrics, human-in-the-loop oversight, auditability, stress testing and cross-sector interoperability.

Long-term (3+ years): The EU should support permanent cross-sector cooperation platforms that bring together competent authorities, CI operators, and technology providers to sustain joint preparedness, conduct common exercises, and develop capabilities. Finally, AI certification under the AI Act should be operationally aligned with CI resilience frameworks, so that conformity assessment and assurance mechanisms reflect the real-world safety, reliability and continuity requirements.

Key Message to Take Away from the Brief

Europe's critical infrastructure is entering a risk environment where climate disruption, hybrid interference, and cyber-physical attacks converge, and where traditional, siloed protection models are no longer sufficient.

Early TESTUDO pilot evidence shows that AI-enabled, multi-agent capabilities can measurably strengthen operational readiness by improving detection speed, situational awareness, and cross-domain interoperability, while keeping humans in control of critical decisions.

At the same time, TESTUDO's pilots underscore that the policy uptake depends on practical enablers: usable operator interfaces, clarity on accountability for autonomous functions, and, most urgently, harmonised EU validation and certification pathways.

The clear priority for EU policy is to move from the isolated demonstrations to a scalable deployment by investing in EU-wide testbeds, embedding predictive tools into the CER implementation, and establishing common standards that make trustworthy AI for critical infrastructure both operationally viable and regulatorily clear.