# Starting document for preparation of a HEU grant proposal

Call: HORIZON-CL3-2024-INFRA-01-03: "Advanced real-time data analysis used for infrastructure resilience". Link: Funding & tenders (europa.eu).



**Background to the Call**

This project targets an increase in cyber resilience across (critical) public and private infrastructures. The Call originates from a strong increase in (attempted) manipulations of data held by or processed by such infrastructures, with potential of cascading effects that may undermine European economic or social structures. The aim of the Call is to develop a system that allows for real-time identification, analysis, response and continuous monitoring of anomalies in and across large data sets, with the possibility for stakeholders to share their findings.

**Expected outcomes of a funded project**

- Improved capabilities for risk and faulty events identification in infrastructure networks and smart cities through real-time analysis (including big data) by public and private actors where the collaboration follows clear legal and political frameworks.
- Tools and processes to identify, analyse, assess and continuously monitor risks and boost adaptive capacity to unexpected events risks in advance by allowing for the analysis of various data sources.
- Fast and continuous real-time identification, classification and tracking of hazardous agents, contaminants or anomalies in infrastructure networks and supply-chains.
- Interoperable interfaces and improved collaboration between infrastructure operation detection and response systems, national/EU risk management/coordination centres and first responder equipment.
- Increased cyber-resilience of industrial xG networks and cloud data covering specific infrastructure domains.
- Improved ability to map in real-time the source(s) of risk factors that could endanger the networked infrastructure supported by Earth Observation and geolocation data.

**Project end-levels**

This Call expects the results to be at TRL5-6, meaning: a demonstration and validation of the technology in a controlled environment.

**Problem description**

Global interconnected networks have allowed ubiquitous data availability. Data availability does not automatically equal awareness of the presence of that data, nor of the significance of its contents. In other words: the availability of big data in many organisations, in particular (critical) infrastructures, and the continuous data exchange means there a high potential for insertion of malicious data that may disrupt or distort normal services, or otherwise misrepresent a factual state. The EU is facing an exponential increase in cyber-related threats against its (critical) infrastructures, which could have serious harmful implications for the functioning of European democratic society. These threats (can) happen across all public and private sectors, and the interconnectivity between organisations increases the risk of cumulative effects exponentially. In the extreme, this could lead to a 'perfect storm' situation where economic and social coherence breaks down.

European regulators and responders invest a lot of time and energy in identifying cyber-security threats, but thus far responses tend to be (too) slow and analysis too granular to be effective or responsive. Moreover, data is not always shared and differences in interpretation (and thus prioritisation and response) result in fragmented intervention (if at all). Predictive risk analysis and cross-referencing of suspicious data activity/transactions are also not possible. At this time, there is no proven, transparent and coherent system on the market for real-time detection and analysis of anomalies in/across large datasets. There is also no adequate legal framework in place that is sufficiently flexible to allow responders to fight emerging types of cyber threats.

**The SHIELD project**

As far as we know, CERN is the only organisation that has a proven anomaly detection system in place for the detection of targeted particles in their LHC. This system, called ROOT, is able to cope with very high data volumes and can operate with high dimensional data. In addition, its complementary BioDynaMo software allows the simulation of interactions between many different data points and thus provide strong predictive capability.

ROOT and BioDynaMo form the methodological foundations of the SHIELD project. Both approaches are statistically sound and comprise full big data analysis tools and visualisation capabilities. Even more importantly, both are fully open source and web-based, which means that they allow local adaptations and wide collaboration. In a recent study, it was possible to transpose the ROOT methodology and tools onto financial market data to identify malicious trading behaviour in futures and options markets.

SHIELD will further strengthen the technical capabilities of ROOT and BioDynaMo with regard to the analysis and visualisation of anomalies, the inclusion of more data sources and the predictive capabilities. To respond to the needs of different types of critical infrastructures, we will further finetune the ROOT methodology with respect to its accuracy, response times, modelling capability, data structure, time correlations and computation cost. Based on this, the SHIELD consortium will create a common framework for data interpretation, specifically with regard to threat level analysis and prioritisation, visualisation, and impact measurement. That will result in a common analytical toolbox and a transparent process for data collection, data analysis and data sharing across stakeholders.

To prove the interoperability of the SHIELD system, 4 use cases will be created, whereby historical data from each is analysed and benchmarked against each partner's current analytical approach. The HighLo project will be used as a technical reference and template for each use case. White hacking will be used for technical validation of the methodology, the analysis toolbox and the interoperability of systems and data.

Lastly, we will assess the SHIELD methodology against the existing European legal frameworks (notably: Cyber Resilience Act and Cyber Security Act) so that modifications or new legislation can be proposed that will give regulators and responders the legal foundation to take appropriate response actions.

**Proposed SHIELD project structure (v.01)**

1. Project Management
2. Common methodology for data analysis
    a. Requirements
    b. Interpretation
3. Technical implementation
    a. Libraries & Toolbox
    b. Data source expansion
    c. Data and analysis interoperability & understanding at application side
4. Use cases
    a. Finance – stock market data
    b. Port – container movements
    c. Telecom – social media messaging
    d. Energy – energy trading
    e. Smart city – IoT mobility/water
5. Evaluation and validation of results
    a. Technical evaluation: anomaly identification/analysis/source identification/simulation
    b. Interoperability & data sharing
6. Recommendations
    a. National and European legal frameworks
    b. User acceptance
7. Dissemination & exploitation of project results
    a. Next steps toward technical upscaling and interoperability

b. Pathway toward regulator acceptance and approval
c. Business case development

**Consortium & partner search**

The consortium comprises CERN and WUR as initiating partners, and we are currently identifying further key organisations to join our effort. These public or private organisations should be either:

- demonstration partners able to locally implement the technical template toward a use case.
- a university with specialism in EU digital law and user acceptance strategies.

**Observers**

We are also looking for public regulators related to the identified sectors, technical standards agencies and responders (i.e. crime units, forensic experts, anti-terrorism agencies) to act as project observers and guide/advise the consortium in the further development of the SHIELD technology toward upscaling and market acceptance.