ACIG
APPLIED CYBERSECURITY & INTERNET GOVERNANCE

NASK

A Cybersecurity Digital Twin Architecture for Modelling Threats in Interconnected Systems

Matteo Repetto | Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes," Consiglio Nazionale delle Ricerche, Italy | ORCID: 0000-0001-8478-2633

Daniele Canavese | Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes," Consiglio Nazionale delle Ricerche, Italy | ORCID: 0000-0002-4265-7743

Abstract

The growing digitalisation of industrial and business processes is creating large deployments where Information Technology (IT) and Operational Technology (OT) are tightly interconnected and interdependent. While cybersecurity risks for IT are mostly related to data breaches and service unavailability, OT is far more critical, since it handles physical equipment. In fact, intrusions into cloud infrastructure may directly or indirectly affect critical processes for autonomous driving or energy operations, with high-impact consequences for people's lives. Additionally, the prevailing interconnection between providers across business and technological value chains creates further tight, recursive inter-domain security dependencies, which are challenging to address due to the fragmentation of cybersecurity operations. Secure and reliable operation of the whole chain requires each provider to improve the security posture of its suppliers. However, the current practice primarily relies on human intervention to disclose vulnerabilities, raise alerts, and suggest remediations, which has been proven to be largely ineffective and risky. In this position paper, we discuss how digital twins can help address the security implications of large, multi-ownership, interconnected systems. We start from the

Received: 17.11.2025

Accepted: 04.05.2026

Published: 05.06.2026

Cite this article as:
M. Repetto, D. Canavese, "A cybersecurity digital twin architecture for modelling threats in interconnected systems," ACIG, vol. 5, no. 2, 2026, doi: 10.60097/ACIG/221493.

Corresponding author:
Matteo Repetto, Istituto di Matematica Applicata e Tecnologie Informatiche "E. Magenes," Consiglio Nazionale delle Ricerche, Via de Marini 6, 16149, Genova, Italy; E-mail: matteo.repetto@cnr.it
0000-0001-8478-2633

concept of cybersecurity digital twin, a live threat model that combines digital assets and cyber threats, and then extend the scope to hybrid digital inter-twins, which brings physical devices into the abstraction. While the former can be used for threat hunting, lateral movement detection, and attack eradication, the latter also models cascading effects and hazards in critical infrastructure. We discuss the two models for two use cases, namely Smart City and Smart Grid.

Keywords

cibersecurity digital twin, hybrid digital twin, service chains, threat hunting, interconnected systems

1. Introduction

Today, digital and physical infrastructures are more interconnected than ever. The interconnection of software, devices, data, and infrastructures entails a large number of entwined, recursive, and often hidden supplier–consumer relationships. We will refer to these 'systems of systems' as digital service chains (DSCs) to distinguish them from 'meshed' supply models and 'linear' networks of services.

Attacks against the supply chain are an incremental threat to enterprises. The growing number of digital interconnections is expected to exacerbate the problem, as security controls across domains become looser, thereby exposing them to lateral movements between digital service providers (DSPs), which account for 25% of all attacks [1]. Even if most examples of DSCs are currently limited to a few providers (e.g. public cloud services and sensor networks), several factors, such as high bandwidth, massive connectivity, the Internet of Things (IoT), virtualisation, and multi-tenancy, turn 5G/6G verticals into fertile ground for attacks [2]. Similarly, the current interest in data spaces (e.g. the GAIAX[1] ecosystem) creates significant dependencies on data from external suppliers and on the infrastructures/services that generate it. Just considering smart city services, which are among the fastest-growing sectors for DSCs, cyberattacks have the potential to affect millions of citizens worldwide (including both residential population and tourist flows) [3].

1 https://gaia-x.eu/

The interconnectedness of physical and digital systems facilitates the propagation of threats because of loose security controls between the involved providers. As a matter of fact, supply-chain

attacks are among the top eight cybersecurity threats in 2022 and beyond, and DSPs are the second most-affected sector [4]. The massive integration of globally interconnected Information Technology (IT)/Operational Technology (OT) systems has challenged the continuous, seamless, and safe operation. Indeed, the combination of different technology value chains (e.g. electrical equipment, robot, vehicles, IoT devices, wide-area networks, cloud services, and software) creates tight inter-dependencies between them, hence security of the whole energy system is not the plain sum of security of each chain because failures and breaches can easily propagate across them and result in cascading effects [5]. For instance, even if power grids are designed to be resilient to the failure of single components, the cyberspace offers virtually unlimited attack paths that can be exploited in parallel to make the hardware redundancy ineffective, so systematic and coordinated attacks against two or more components likely lead to power outage and/or significant damage [5].

The prevailing fragmentation of cybersecurity operations in such multi-ownership systems hinders visibility across domains, thereby jeopardising coordinated, timely detection and response to attacks. Indeed, many organisations rely on the protection measures put in place by their suppliers and do not implement their own controls to address the remaining vulnerabilities; as a result, market surveys indicate that 98% of respondents have been negatively impacted by a cybersecurity breach in their supply chain [6]. In this scenario, even if each standalone provider has good detection and prediction capabilities, they usually resist cyberattacks without fighting back because they lack the technical means to eradicate attacks originating from any point in the chain. Moreover, without the ability to sanitise the root cause, attacks persist and may find other vulnerable points to move laterally along the chain [7]. This makes supply chains an attractive pathway for cyberattacks.

Cyber threats continue to evolve, and reactive defences remain ineffective against zero-day attacks, advanced persistent threats, and fast-moving cyber–physical impacts. Ensuring reliable operation of cyber–physical systems requires anticipating attacks, prioritising risks, and preparing mitigation strategies in advance. This demands a proactive, collaborative approach that identifies interdependencies across technology and business value chains, detects threats early in the cyber kill chain, and predicts cascading effects [8]. However, assessing suppliers' security posture remains difficult, as providers lack visibility into emerging issues and cannot reliably verify remediation efforts [6].

The use of a digital twin (DT) is not merely a matter of visualisation or simulation in this work. It is required because DSCs are dynamic, multi-domain, and partially opaque systems in which security-relevant dependencies cannot be captured by static inventories, isolated security information and event management (SIEM) deployments, or one-shot risk assessments. A cybersecurity digital twin (CDT) provides a continuously synchronised security abstraction of assets, vulnerabilities, trust relationships, and attack paths, while the hybrid digital inter-twin (HDIT) extends this abstraction to cyber–physical consequences. This enables non-invasive 'what-if' analysis, attack-path anticipation, response verification, and cross-domain threat hunting before mitigation actions are applied to operational infrastructure. The proposed model acts as a decision-support and risk-reduction mechanism for environments where direct experimentation on the real system would be unsafe or infeasible.

In this position paper, we discuss our vision of how DTs can be used for security purposes beyond their trivial use for penetration testing and cyber ranges already proposed in the literature. The main novelty of our work is the architecture for a CDT that models threats and attack paths for service chains. In our vision, CDT represents a new form of threat modelling, linking the knowledge of the physical context with related vulnerabilities and threats. Additionally, we discuss the concept and architecture of an HDIT, which integrates digital and physical aspects and models the mutual relationship between cyberattacks and safety/operational risks.

The rest of the paper is organised as follows. Section 2 reviews the existing architectures and usage of DTs for cybersecurity purposes. Section 3 explains the concept of DSC and provides two use cases for smart cities and smart grid. Section 4 introduces the concept and purpose of a CDT, and its architecture is discussed in Section 5. We further extend the discussion to federation issues for interconnected multi-ownership domains in Section 6. We then move one step forward in Section 7 with the concept by discussing how CDTs and DTs can be combined, with a practical focus on the smart grid scenario. Finally, we give our conclusions and discuss the road ahead in Section 8.

2. Related Work

The concept of DT has been around for many years. Hence, there are already good surveys that discuss its definition and implications, enabling technologies, the existing frameworks, current

trends, case studies, market forces, open challenges, prospects of DTs, and their significance in various domains [9], including the interconnection of multiple models [10]. The scope often entails the management of industrial applications and cyber–physical systems (CPSs). Still, DTs are used in almost any field, for example networking [11], critical infrastructure [12], smart cities [13], and healthcare [14].

A typical DT architecture spans physical and virtual layers (e.g. sensing, synchronisation, modelling, and visualisation) across IoT, edge, network, and cloud domains. In cybersecurity, DTs can replicate attack scenarios to collect data, identify vulnerabilities, and strengthen defences. However, as noted by Pokhrel et al. [15], practical implementations remain limited, and the use of DTs for cybersecurity is still immature, lacking standardised definitions and architectures.

Several works employ virtual replicas of physical processes for safe security analysis. Eckhart and Ekelhart [16] derive DTs from CPS specifications to support intrusion, detection, testing, and penetration analysis, while Empl and Pernul [12] highlight DTs' role in IoT security analytics. National Institute of Standards and Technology (NIST) [17] emphasises lifecycle security integration in CPS frameworks. Bitton et al. [18] propose a cost-effective Industrial Control System (ICS) twin balancing fidelity and budget constraints. Saad et al. [19] developed analytic models for power systems and IoT shadows using control theory for anomaly detection, though their approach remained largely agnostic to real-world attack patterns and domain-specific protocols.

A data-driven definition of DTs is provided by Bécue et al. [20], who describe a DT as an evolving digital profile built from cumulative, real-time measurements, and propose its use in cyber ranges for risk anticipation and impact prediction. Yigit et al. [21] employ DT-like agents to detect denial-of-service (DoS) attacks in core networks, although their approach resembles a digital shadow; collaboration among 'elementary' DTs enables improved detection of contextual and collective anomalies. De Benedictis et al. [22] model a railway signaling system using Petri nets and combine monitoring data with offline simulation and supervised machine learning (ML) for anomaly detection. While not explicitly focused on cybersecurity, their work demonstrates a concrete DT implementation that can be extended to cyberattack detection.

Atalay and Angin [23] propose executing threat-feed-based cyberattacks on a smart grid DT derived from system specifications,

emphasising continuous assessment and rapid model updates following technological or security changes. Recent work further integrates advanced measures, such as Zero Trust Architecture, into DTs to strengthen resilience and enforce strict access control. Dietz et al. [24] employ a DT during the design phase to analyse protocol and device vulnerabilities, though their approach is limited to simulation in the absence of a physical counterpart. Salvi et al. [25] envision DTs as sandbox and cyber-range environments for critical infrastructures, highlighting their integration into broader prevention and response frameworks while acknowledging trust concerns related to sharing sensitive security information.

De Benedictis et al. [26] extend the DT architecture with security functions across physical, virtual, and communication layers, introducing a dedicated layer for services such as anomaly detection, though without a concrete implementation. A common limitation of such approaches is the DTs' limited capability to accurately model protocol-level vulnerabilities, as most CPS-oriented DTs emphasise control and business logic rather than interfaces and software components.

Hadar et al. [27] propose a CDT that models attacker movements to identify exploit paths and derive necessary security controls. Their framework automatically discovers assets and vulnerabilities and maps attack tactics to controls. However, it lacks capabilities for detecting ongoing attacks and predicting the impact of system or threat changes, highlighting the need for continuous improvement and integrated predictive analytics in DT-based cybersecurity.

Although DTs can solve many problems and be very useful for testing and verifying physical twins, some security issues must be addressed before deploying and using them in cybersecurity. Such issues can concern the access that users or other systems have to the twin's data and functionality, or the integrity of the data received and stored [28]. Hence, some challenges and limitations exist in using DTs for cybersecurity, such as vulnerabilities in the underlying hardware, software, and communication networks.

Digital twins also pose communication, security, and scalability challenges. They require low-latency synchronisation with their physical counterparts, protection of large volumes of sensitive data throughout their lifecycle, robust authentication and access control, and scalable management when many twins and scenarios operate concurrently.

3. Digital Service Chains

Digital service chains are more challenging to set up and operate than other forms of supply chains because they usually follow 'meshed' operational models instead of more conventional 'linear' patterns [8].

A linear value chain incrementally refines goods and services at each stage, also combining resources from different providers. This is the typical case for industrial processes, and it also occurs in the software industry, where libraries and codes are progressively developed and integrated, culminating in delivery with deployment scripts and configurations, as shown in Figure 1. Instead, in a meshed model, applications interact continuously with external devices, services, infrastructures, and data at run-time, which typically yields complex, nonlinear, unclear, recursive, and dynamic interdependencies.

A rather meaningful scenario is represented by smart cities, where it is unlikely that a single provider owns all the necessary devices, infrastructures, services, and data to create the expected services (e.g. infomobility, parking management, and air and ground monitoring). Figure 1(B) shows a realistic example of a smart city where IoT and personal devices, metropolitan and cellular networks, cloud infrastructures, and applications from different providers are combined to create and deliver advanced services to citizens. The various shapes and colours of the placeholders highlight the heterogeneity and multi-ownership of digital resources involved in the service chain.

Virtualisation, cloud models, standard interfaces, and software-oriented architectures facilitate the deployment of DSCs across

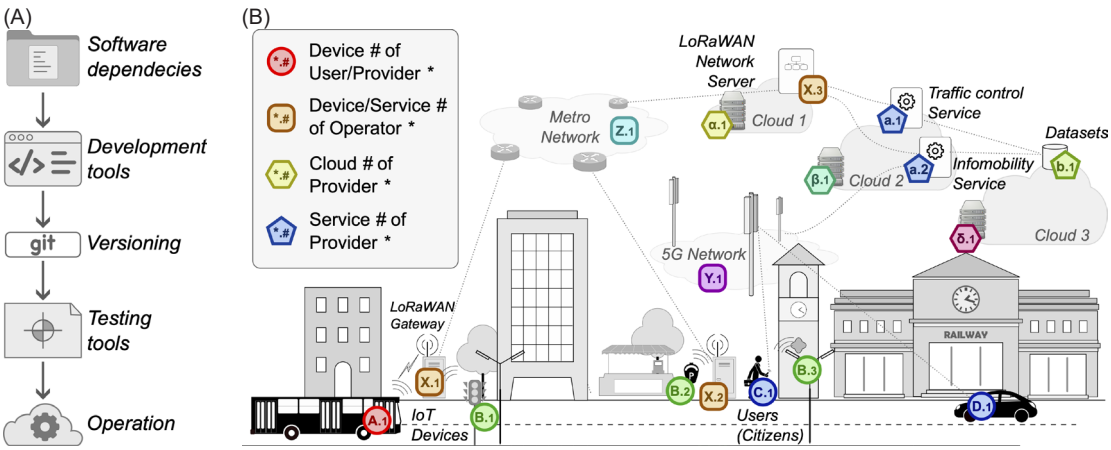

Figure 1. Linear versus meshed supply chain models.

multiple providers and make supplier replacement technically easier than in traditional industries. However, cybersecurity processes remain poorly integrated across domains and mostly rely on slow, partial, and often inaccurate exchanges of incident reports and threat intelligence (TI), which typically arrive only after malicious activity has already propagated.

The existing fragmentation of cybersecurity operations prevents a common and coherent strategy for the entire chain and leaves many open issues:

- Multi-ownership, which inhibits mitigation and response to attacks originating in other domains in the absence of collaboration from the owner.
- Dynamic, partially unknown, and opaque topologies, which hinder (i) a complete and holistic assessment of vulnerabilities as well as the establishment of strong trust relationships; (ii) the prediction of the impact of changes to configurations and/or alternative deployments; and (iii) the location of data and the tracking of their propagation across services.
- Scarce or no visibility and control over services and infrastructures operated by third parties.
- Lateral movements between services, which exploit weak security controls due to business relationships in place.
- Broad attack surface due to weak links in the chain that miss strong security policies.

A key missing innovation is the use of the existing technical interactions to achieve full situational awareness and coordinated response in multi-ownership environments. This would enable risk prioritisation based on business context, identification of root causes and responsible providers, and faster remediation. A CDT can support this by dynamically modelling service composition and operations while preserving confidentiality. However, anomaly detection and attack traceability in hyper-connected, data-intensive environments remain significant research challenges.

3.1. Use Case #1: Smart city

Smart cities integrate physical infrastructure, people, processes, and data through the information and communication technology (ICT)-enabled cyber–physical systems. They support mobility, energy efficiency, environmental sustainability, and improved public services by exploiting interconnectivity, data availability, and contextual awareness.

Boosted by significant investments, the scale and complexity of digital infrastructure and processes are increasing every year in cities around the world. As more social, medical, industrial, energy, and environmental services rely on ICT, digital infrastructures and technologies will become critical assets for smart cities. Core infrastructure components, such as sensor networks, cloud services, databases, and public mobile networks, need to be reliable and robust against cyber threats and service interruptions from both external attacks and internal misuse.

Figure 2 shows a typical service chain for a smart city scenario, where data is collected from multiple sources, including IoT sensors and data stores, and feeds urban intelligence platforms. The latter aggregate, enrich, process, and distill data, which specific services can then use for traffic management, info mobility services, etc. At the bottom of the chain, IoT devices generate data to feed the urban intelligence platform. Additionally, other data sources are available, including data sets, websites, etc. In the middle of the chain, wide-area networking services (like LoRaWAN and 5G/6G networks) deliver data to the urban intelligence platform. Smart city services then use the latter. All these services are typically hosted in cloud infrastructures. Finally, smart city services are made available

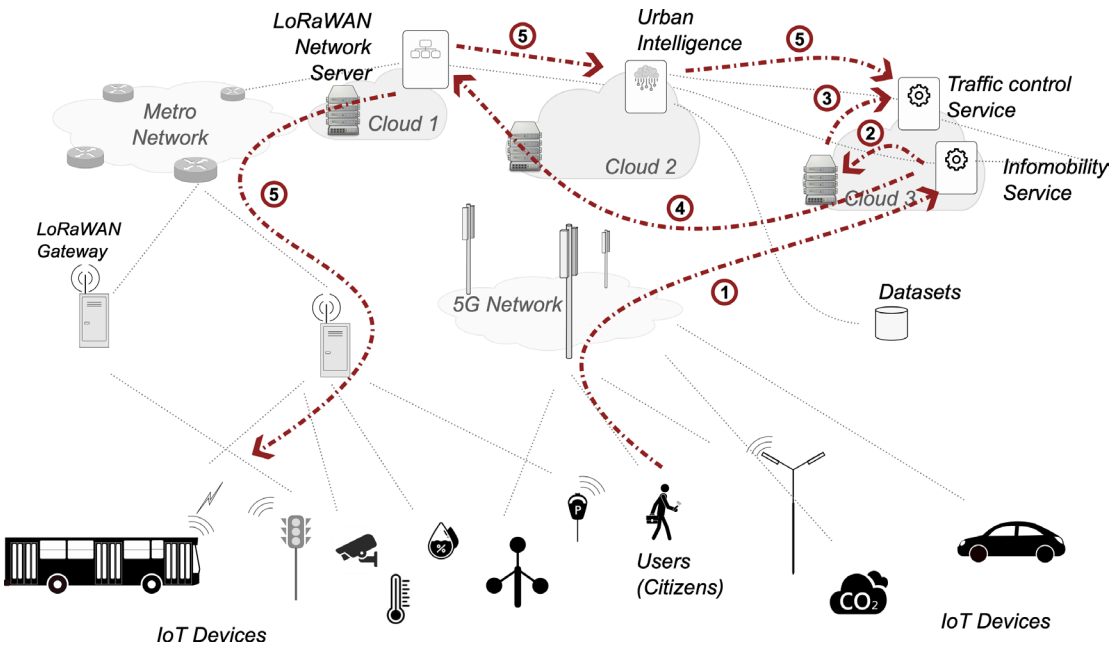

Figure 2. The smart city service chain and potential attack paths.

to citizens via web portals or mobile apps, again using public networks for device connectivity.

We also note that a networking service might also build on third parties' infrastructure. For instance, the LoRaWAN network needs a transport network to interconnect its gateways with the network server (it may be a metropolitan network or a public Internet), and the network server itself is often deployed as a multi-container application in the cloud.

With reference to Figure 2, let us suppose a criminal wants to disrupt traffic control operations for the entire city, hence targeting the traffic control service. Of course, such a service is protected by access controls and firewalls, but the presence of shared infrastructure and connected services might create indirect paths for adversaries. For instance, the attacker could exploit the public-facing web application of the info mobility service to open a reverse shell within the system (step 1, CVE-202017530). Even if he/she is inside a container/virtual machine (VM), the attacker can now look for instance metadata, giving direct access to the cloud management interface (step 2). If such credentials are not the same as those used for the target service, privilege escalation might still be possible due to policy misconfiguration (step 3). Alternatively, the attacker might try a cross-site scripting attack towards the LoRaWAN server (step 4, CVE-2020-7656) to generate fake messages addressed to the traffic control service or the field devices, for example traffic lights, or report wrong data to the urban intelligence platform to invalidate traffic controls (step 5).

3.2. Use Case #2: Smart grid

The electrical grid is evolving from a largely stable design towards a more flexible, ICT-enabled infrastructure driven by efficiency, renewable integration, new market models, and bidirectional energy flows. This transformation relies on smart devices, communication networks, computing infrastructure, digital services, and automated processes.

The transition to a 'smart' grid brings more intelligence, flexibility, and automation in all control and management processes, from generation to transmission, distribution, and consumption; however, the presence of ICT infrastructure also exposes the grid to many cyber threats, which exploit vulnerabilities in intelligent electronic devices (IEDs), communication protocols, and network topology. Attacks against the smart grid not only impact the safety and

continuity of operations but also pose a threat to the large amount of customers' data that is usually owned and managed by distribution system operators (DSOs).

Boosted by the introduction of intermittent generation from renewables, electricity trading in the energy market, and the need for greater efficiency, the electrical grid has evolved towards more flexible production models than in the past. The concept of a smart grid requires significant infrastructure to be deployed throughout the entire supply chain, from production sites downward to transmission system operators (TSOs), DSOs, and up to users/prosumers. The ICT infrastructures include IoT devices for metering the entire grid, telecommunication networks that provide ubiquitous coverage, and computing infrastructures to run the multiple monitoring, management, control, and business applications of all involved players.

Figure 3 shows a simplified scenario for the operation of a smart grid. Smart metres are deployed at customer locations that measure the electricity consumption of both industrial and residential users. This data is typically collected through power line communication (PLC)

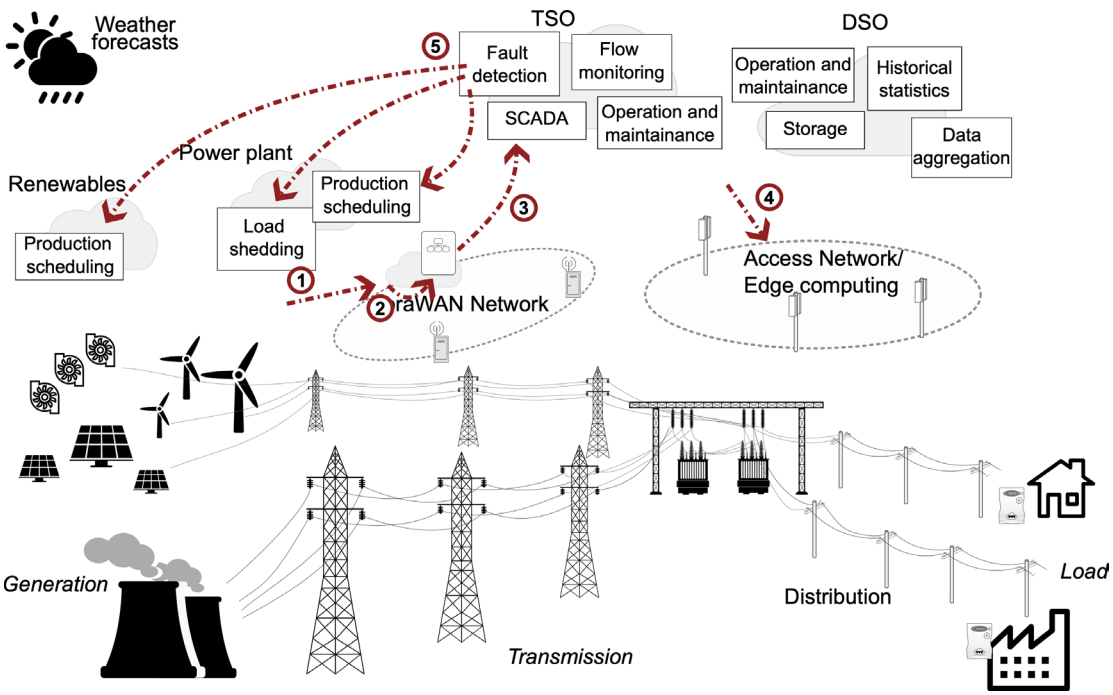

Figure 3. The smart grid service chain and potential attack paths.

up to the distribution substation; the latter aggregates the data and sends it via wired or wireless networks to the DSO facility. This data is mainly used for billing, but it can also be processed to create historical trends and statistics to predict load on a weekly, daily, or even hourly timescale. TSOs transport electricity from generation sites up to the substations of DSOs. They are responsible for buying electricity from producers and selling it to DSOs. They operate the critical backbone of the grid and must constantly monitor it to identify anomalies and faults. To this end, TSOs collect status information and measurements from their grid equipment (e.g. transformers, energy flows, phasors, and circuit breakers) using many sensors via a public network (e.g. LoRaWAN). For what concerns producers, they must know in advance the amount of electricity to generate in the next period (hour, day, or longer), which must fulfil the expected load from consumers. Renewables also need to estimate their production based on weather forecasts and historical data.

The overall objective of an electricity grid is to maintain a balance between the power injected by generation plants and the current load; this balance is necessary to avoid power outages while making the most efficient use of renewables and fossil fuels. Operation of the electrical grid requires a constant (and often bidirectional) flow of information among all stakeholders; it is no surprise that the electrical grid is one of the most critical infrastructures today, because in case one component gets compromised, it may drive the system away the equilibrium point, with catastrophic consequences for the safety of the electrical grid and the attached users.

Reliable and safe operation requires high service availability and trusted data. In the smart grid service chain, several vulnerable paths can be exploited to disrupt the electricity system.

With reference to Figure 3, let's suppose the attacker can compromise a VM running in the same cloud as the LoRaWAN server (step 1, for instance, by sending a phishing email or by server-side scripting). Now the attacker exploits vulnerability in the cloud software and can compromise the VM where the LoRaWAN network server runs (step 2, e.g. by escaping its VM or its virtual network). The attacker's purpose is now twofold: first, it can collect supervisory control and data acquisition (SCADA) data for several days to understand the expected behaviour of the transmission network. Second, at some point, it starts injecting fake data to deceive the TSO's energy flow monitoring (step 3). At the same time, the attacker initiates a denial of service (DoS) attack on the public mobile network, disrupting communication between smart metres and the DSO's

operation and maintenance centre (step 4) to deny consumers up-to-date information about the blackout. Fake data injected into the SCADA of the TSO will cause production sites to run ahead of the load balance (step 5); this will trigger protection devices, such as fuses and circuit breakers, resulting in a power outage.

4. A Cybersecurity Digital Twin for Digital Service Chains

The concept of CDT has been used recently to denote a live model for the security properties of ICT systems, which can be used to perform security assessments and to emulate cyberattacks and defence scenarios without disrupting its operational counterpart [29].

However, even the most recent literature fails to commonly and unambiguously define the concrete implementation of the general concept and its usage for the intended purposes. We argue that plain DTs of physical assets help detect anomalies and illicit usages, but the overall concept of CDT should entail many more capabilities in terms of understanding and predicting cyberattacks [30]. As a matter of fact, in the context of DSCs, it enables a better understanding of the relationships among individual domains within the same chain and the potential cascading effects of attacks, configuration changes, and mitigation actions.

4.1 Concept and Purpose

Specularly to the intended usage of DTs for physical systems, a CDT should support security operations (i.e. monitoring, detection, investigation, and response) by providing modelling and prediction capabilities of the evolution of cyberattacks and the risk that potential threats materialise in the current or a hypothetical context in which the system operates.

Our vision is illustrated in Figure 4. Federated bidirectional models abstract the composition, topology, and security properties of interconnected systems and are continuously synchronised with real-world events and vulnerability data. Combined with shared TI, these models enable attack modelling and prediction, forming the CDT that enhances detection, analysis, protection, and response processes, with enforcement actions applied through the same framework.

In our vision, a CDT should map TI to the real system to generate hypotheses, perform analysis, and predict what could happen in

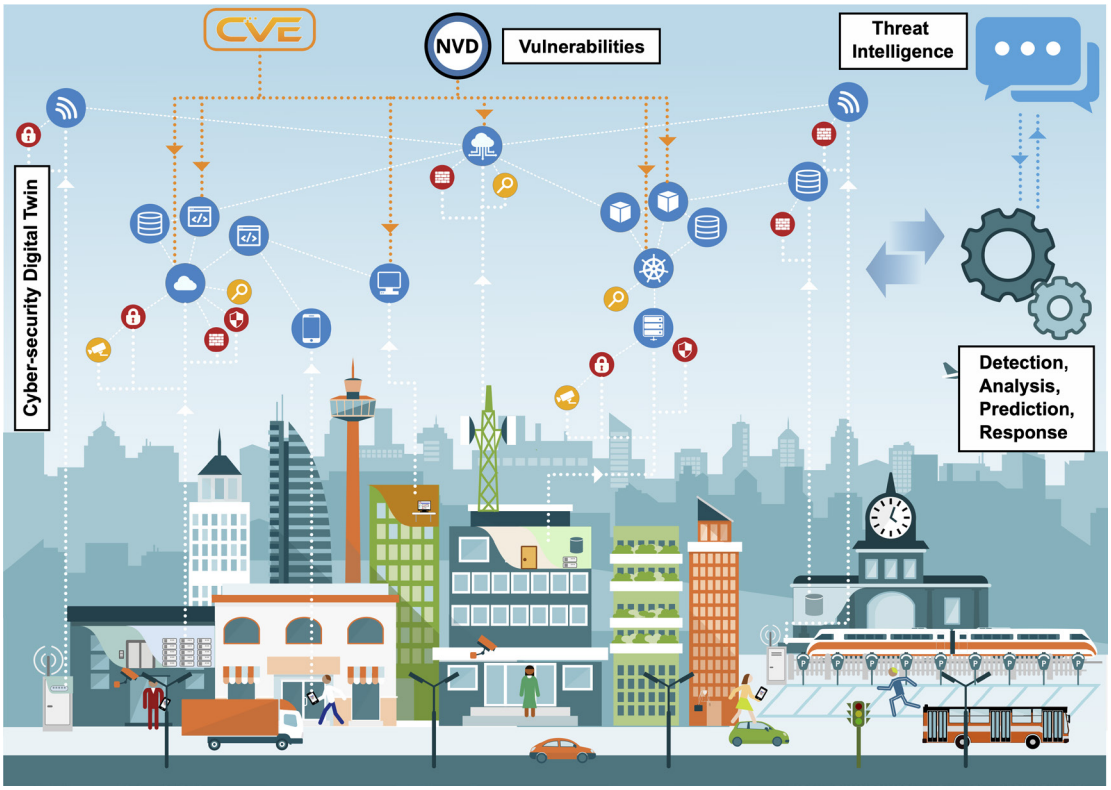

Figure 4. The scope of a CDT framework in the representative example of smart city services.

the real system in a proactive yet non-invasive way. It should leverage a bidirectional flow of information with security agents, including both monitoring and enforcement capabilities, which keeps the states of the two twins synchronised. In some way, the CDT could be seen as an extension of the IoT shadow provided by Saad et al [19]. A CDT for DSCs must accordingly be designed to be holistic, adaptive, agile, zero-trust, opaque, and, most of all, safe. A holistic CDT must account for the broad heterogeneity of digital resources, including physical devices; computing, networking, and storage infrastructures; software applications; data; and processes. The latter do not have a direct cyber–physical counterpart but correspond to human tasks (e.g. sending email, inserting data, and locking doors).

A major challenge for a CDT is adaptability to evolving operational and threat contexts. This requires not only bidirectional monitoring and control but also automated discovery of system changes and relationships. Scalability and agility are essential to handle large service chains, variable loads, and long-term trends. A zero-trust

architecture must regulate dynamic inter-provider interactions, ensuring context-aware data opacity based on runtime trust levels. Finally, security and privacy-by-design principles are necessary to prevent the CDT itself from becoming an attack vector.

A CDT should not remain a standalone modelling component but should be integrated into a more general framework that includes cybersecurity processes on top of it. The scope encompasses monitoring, detection, analysis, investigation, and response activities that improve legacy operations by leveraging modelling/prediction capabilities, as previously shown in Figure 4.

4.2 Expected Capabilities

A CDT is expected to model the aspects of its real counterpart relevant to security analysis. This requires capturing the main properties of all interconnected systems as well as the business and technical relationships that may enable lateral movement between them. Hence, the design of a CDT that models attacks and threats on DSCs should be based on two main capabilities:

1. Context abstraction, which captures technical and functional properties of all interconnected systems as well as the business and operational relationships that allow lateral movements between them.
2. Attack modelling, which is used for the prediction and emulation of attacks.

Context abstraction can rely on service context graphs (SCGs) [8], which capture: (i) service identity and ownership; (ii) execution environments and configurations; (iii) operational relationships and communication patterns; (iv) known vulnerabilities and threats; and (v) available cybersecurity functions (CSFs) and their capabilities. Figure 5 illustrates an SCG derived from the smart city scenario, where digital resources are linked through application and infrastructure relationships. Applications interact via network flows or application programming interfaces (APIs), while infrastructures host them through virtualisation mechanisms. These relationships constitute potential attack vectors depending on their nature.

We assume the availability of CSFs in each domain, namely log collectors, L3/L7 firewalls, Intrusion Detection System (IDS)/Intrusion Prevention System (IPS), antivirus, encryption, network scanners, network telemetry, trusted computing platforms, etc. Additionally, we assume that a specific management endpoint is available to

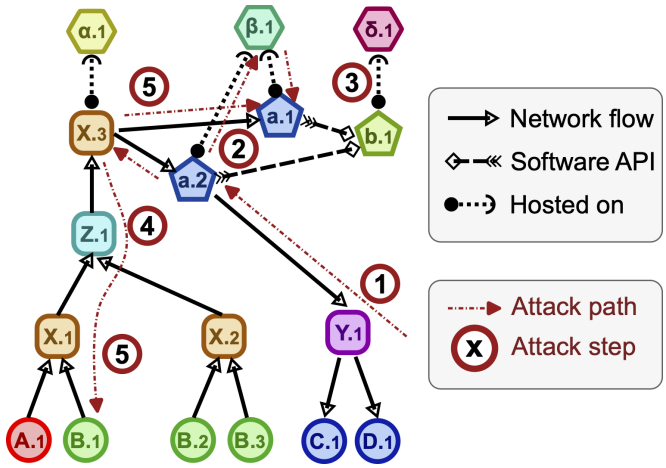

Figure 5. SCG model for the smart city example depicted in Figure 1(B).

retrieve the description of the execution environment and the list of CSFs; collectively, this information provides the context that will be abstracted in SCG. This assumption does not constitute a hard constraint for a specific scenario of DSCs, as it is commonly provided for any kind of self-provisioning interface, including cloud services, Software-Defined Networking (SDN), Network Function Virtualisation (NFV) controllers, etc.

Attack modelling should concern logical rules that describe how an attacker advances within the service chain (see Figure 5). These logical rules enable adversarial lateral movement, which a defender must eliminate and nullify [27]. The scope entails all types of models that can be fed by typical security data (logs, system metrics, measures), including analytical formalisms (e.g. analytical attack graphs (AAGs), petri networks [PNs]), probabilistic the ories [e.g. Bayesian methods], and data-driven models [e.g. ML/artificial intelligence {AI}]). The modelling and prediction capabilities of a CDT underpin the implementation of advanced cybersecurity operations on the whole service chain. They include monitoring, detection, hunting, and response, which are made agile, automatic, and adaptive. The main objective is to recognise the ongoing multi-step attacks, to predict potential attack paths based on the current service configuration and vulnerabilities, to identify the final target, and to tailor detection and mitigation/response actions to the current configurations (e.g. IP addresses and topologies).

5. CDT Architecture

Our proposed conceptual architecture, shown in Figure 6, reflects the notions and assumptions described so far. It is

organised into three layers, namely service, twinning, and security operations.

The *service* layer contains the physical/virtualised infrastructure, data, applications, and devices. Here, several CSFs are deployed for both monitoring (e.g. to collect events, logs, system metrics, network measures, and packet and software traces) and enforcement tasks (e.g. packet filtering rules, and access control rules). The connector is the management endpoint that is responsible for describing the overall execution environment and tracking the available CSFs. It extends the management agents already required by SIEM tools (e.g. Wazuh agents). Given the multi-domain/multi-ownership scope, identity management and access control are inescapable requirements for this component.

The *twinning* layer provides modelling and prediction capabilities of cybersecurity properties. It has two interfaces for interacting with external components. Synchronisation with the real counterpart happens on the southbound interface, which implements both data and control API towards local *connectors* and CSFs. The control API will leverage open standards, namely OpenC2, for the discovery

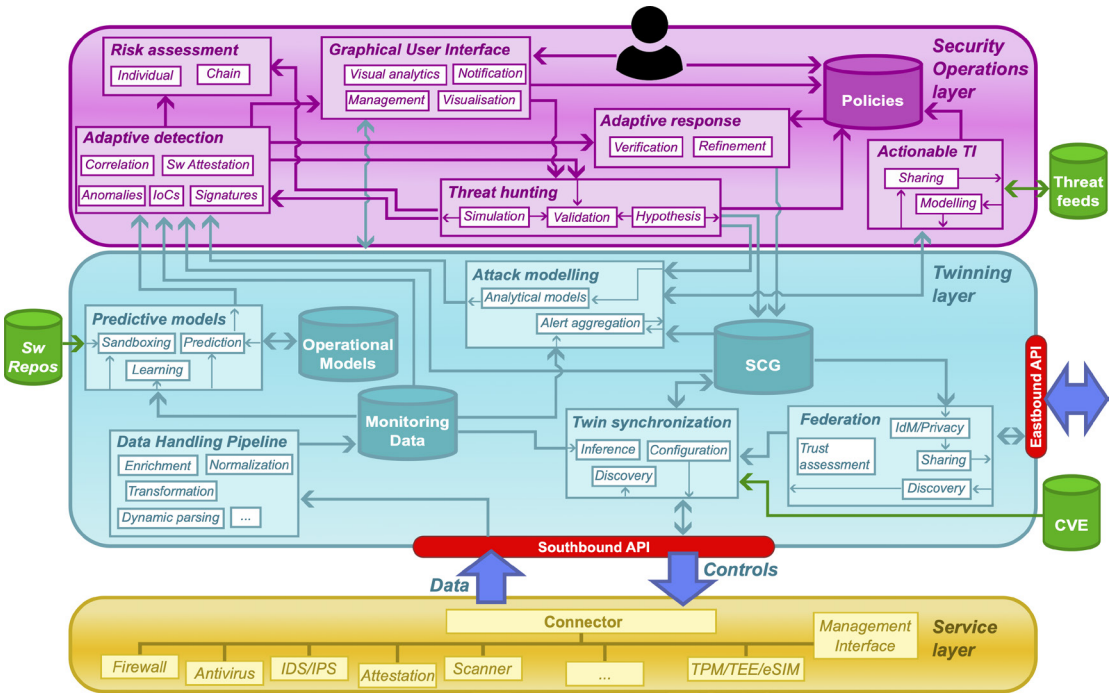

Figure 6. Four-layer architecture of a CDT framework.

and configuration of the CSFs. Data collection will build on open-source protocols and pipelines (e.g. the Elastic Stack, Prometheous, and Graphana).

The *eastbound* interface is conceived to federate with the CDTs of other providers. This will build on the ongoing work on DTs in other domains (e.g. IETF DT Networks, ISO 23247, and IEEE P2806), also investigating the possibility of using OpenC2[2] for this purpose as well. The scope of this interface is many-fold: (i) to retrieve the description of other parts of the supply chain; (ii) to notify alerts about threats and vulnerabilities to suppliers; and (iii) to delegate prediction tasks to remote models.

2 https://openc2.org/

Internally to the *twinning* layer, the *twin synchronisation* module is responsible for keeping the real system and its internal abstraction, namely SCGs, in sync. This module has a bidirectional scope, since changes in the composition and configuration of the execution environments are reflected in the internal models, and changes in the internal models are applied as configurations to the real system. It is also fed from the federation module, external feeds, and internal TI sources. *Twin synchronisation* has also inference capabilities, for example to derive technical and business relationships between services from monitored data. For instance, it detects open network ports from the configuration and communication traffic patterns from network probes.

The data handling pipeline (DHP) collects, enriches, aggregates, transforms, normalises, and indexes security-related data, so as to maintain a real-time and historical track of system status. The *predictive models* module implements different methods for emulating the behaviour of real systems. They include: (i) traditional ML and trusted federated learning (FL) that save models in an internal secure repository (*operational models*), and use them for short-term predictions (e.g. volumes or patterns of network traffic, software control flows, API syntax); sandboxing capabilities and network simulation (e.g. mininet) that retrieve individual software images or entire templates and run them with real or fuzzy inputs in a safe environment for both deriving trusted control flow graphs and detonating suspicious code.

The *attack modelling* module comprises analytical methods for investigating cyberattacks. On the one hand, it takes the logical rules that describe how attackers move between services and contextualises the kill chain to the current SCG (i.e. it maps different stages of the attack to the potential targets). The implementation

will consider the most appropriate formalism among AAG, attack vectors, attack trees, Petri nets, etc.

Finally, the *federation* module has similar objectives as the *twin synchronisation*, but it retrieves/exports data from/to the CDTs of other providers. The *federation* module also makes internal data opaque to external CDT by applying access control, anonymisation, and shadowing mechanisms based on estimated trust levels. To this purpose, *trust assessment* builds dynamic trust relationships based on the availability of hardware-based roots-of-trust (RoTs) and identities. The computation will consider (i) the existence of direct business relationships and reputation, (ii) privacy-preserving and accountable zero-trust authentication and attestation services (e.g. attributed-based direct anonymous attestation [DAA]), including integrity verification of both configurations and the software execution. The latter mechanisms will establish the chain of trust between services for evidence collection and derive the trust scores used in risk assessment and FL and transfer learning (TL) algorithms.

The *security operation* layer makes typical monitoring, detection, investigation, and response processes more agile, adaptive, and automated by leveraging the unique capabilities of a CDT. Adaptivity is achieved mainly by combining monitoring data with SCGs, which provides better opportunities to correlate data based on the existing relationships and potential attack paths.

The actionable threat intelligence module is responsible for generating, sharing, and using data that is distilled, contextual, and real-time. It will therefore prefer tactics, techniques, and procedures (TTPs) formalisms over descriptive reports, as the former are more suitable for machine processing. The scope includes the transformation of TTP descriptions into (i) analytical models for the *attack modelling* module, (ii) tailored indicator of compromise (IoC) for the *adaptive detection* module (e.g. IP addresses, port numbers, and URLs), (iii) high-level protection and enforcement policies; (iv) parsing of Open Source Intelligence (OS-INT) sources. This module will also generate new TI by enriching new findings from the *threat hunting* module with necessary context (relationships, configurations, etc.) from SCGs.

The adaptive detection module detects IoCs and anomalies, and maps them to attack models, so to effectively reconstruct the entire kill chain and trigger the adaptive response module before the attack reaches the final target. Remote attestation will provide evidence of untrusted or compromised components. The *adaptive*

response module will be responsible for refining high-level end-to-end policies into intra-domain policies and low-level configurations/playbooks based on the specific set, capability, and position of security agents at any time, hence making them agile and adaptive to the evolving context. *Adaptive response* will also provide formal verification of the correctness of security controls with respect to high-level policies, thereby identifying missing security controls (e.g. packet/application firewall rules, access control rules, etc.).

The *risk assessment* module will include the evaluation of the required trust level for each component in the system, hence supporting the operation of *federation* and trusted FL algorithms. It will estimate the risk that threats materialise downstream or upstream in the chain. Finally, the Graphical User Interface (GUI) will provide a user-friendly, intuitive, and robust environment for humans. This includes the possibility to perform *visual analytics* with the intuition and creativity that distinguish humans, to receive alert notifications, and to supervise the operation of the whole system.

A practical implementation of the proposed CDT can be organised as an incremental deployment composed of five stages. First, each provider deploys a local connector that extracts asset inventories, service dependencies, communication flows, software versions, security controls, and available management interfaces. These data populate the SCG and provide the minimum abstraction required by the CDT. Second, the DHP normalises telemetry from logs, network flows, IDS/IPS alerts, vulnerability scanners, cloud APIs, and orchestration platforms, such as Kubernetes. Third, vulnerabilities and TI are mapped to the graph by associating software components, exposed interfaces, credentials, privileges, and communication edges with known CVEs), TTPs, and attack preconditions. Fourth, the attack modelling module instantiates possible attack paths using analytical models, such as attack graphs or logic-based reasoning. Automated attack-graph generation is appropriate here because manual construction becomes error-prone and impractical at large scale, as shown in classical attack-graph research. Fifth, the adaptive response module translates high-level mitigation decisions into actuator-specific commands, for example, by using OpenC2.

6. Federation and Trust Management

In a multi-ownership environment, it is challenging to assume that all parties will perform their tasks honestly. Various reasons, such as financial benefits or reputation, may influence their behaviour. Privacy-aware, zero-trust architectures are

therefore necessary to provide innovative mechanisms that dynamically and continuously verify the identity and assess the trustworthiness of providers and their services, thereby preventing misuse and external attacks.

Federations and zero-trust schemes must be conveniently combined for this purpose to enable reliable, dynamic interactions that balance confidentiality with visibility. On the one hand, granting third parties visibility into internal assets and data is not acceptable for most service providers. However, the latter often occurs when security operations are externalised; in this case, the delegation involves only a single trusted entity. On the other hand, blindly trusting data and alerts from unknown upstream and downstream providers may be risky.

To address the first issue, a federation of CDTs could be created to exchange models and run predictions in an 'opaque' way. Here, opacity refers to a mechanism for selectively sharing visibility and controls based on trust levels as well as to current privacy-preservation techniques [31]. The concept is shown in Figure 7, where service provider A's CDT federates with those of directly interconnected systems, and recursively discovers additional providers in the same service supply chain. Within the federation, each CDT determines what information can be shared with other peers by applying fine-grained access control rules. The level of detail exposed by each CDT depends on trust in the requesting entity, resulting in progressively more opaque models, up to complete darkness when no trust is present or the CDT is not implemented.

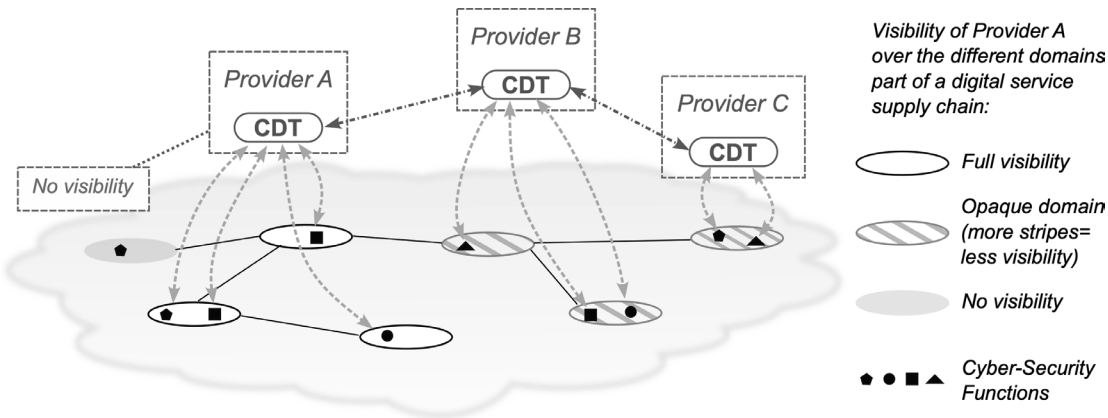

Figure 7. A federated model to share visibility and controls in multi-ownership digital service chains.

Concerning the second issue, establishing trust relationships between providers requires extending existing zero-trust models. The new frontier in this domain is attestation of the configuration and behavioural integrity of entities while addressing privacy concerns in a decentralised architecture, which goes far beyond mere verification of knowledge of secrets. However, there is a non-negligible risk in sharing attestation evidence with potentially vulnerable or malicious verifiers. The challenge is therefore the creation of strong 'chains-of-trust' that take into account both security and privacy requirements posed by the presence of multiple domains.

The combination of federation and privacy-aware authentication mechanisms will enable unprecedented levels of adaptivity and autonomy in managing technical and business relationships between DSPs, thereby boosting recursive schemes. As a matter of fact, if, for instance, service provider A does not have a direct trust relationship with service provider C, it may ask service provider B to act as an intermediary, and to attest to the identity and the integrity of the configurations and software execution paths of services and devices in domain C. This way, each domain can obtain evidence of the integrity and trustworthiness of all other domains without divulging unnecessary confidential information, including their identities.

7. Towards a Hybrid Digital Inter-Twin

7.1. Concept

While the concept of CDT offers new ways to predict, anticipate, and detect cyber threats, it cannot assess its impact on physical systems. This is particularly critical for systems that are also interconnected at the physical layer (e.g. the power grid), because systematic and coordinated attacks against two or more digital components are likely to trigger cascading effects on the underlying physical layer (e.g. power outages, brownouts, overvoltage).

For this reason, we advance our architecture towards a combined HDIT that captures mutual cross-dependencies between physical interaction and IT/OT operations. To be more concrete, we contextualise our architecture in a specific scenario, namely the smart grid discussed in Section 3.2. Here, an HDIT should progress beyond plain and standalone replicas of energy and IT/OT systems used to run legacy detection, testing, training, and investigation [19, 22, 32] (e.g. Eclipse Ditto[3]), and combine analytical models, simulators, and hardware/software-in-the-loop emulation with a CDT.

3 https://eclipse.dev/ditto/

For what concerns the 'physical' facet, the HDIT should be used to (i) compare real measures from grid sensors (e.g. phasor measurement units [PMUs]) and data from system telemetry (e.g. logs and packet traces) from the physical environment with their corresponding values in the twin, (ii) provide visibility over hidden states and components that are not observable in the physical system [33], and (iii) predict the evolution of grid and IT/OT behaviour (including cascading failures and blackouts) under alternative deployments, configurations, and inputs. Concerning the 'digital' facet, the CDT will provide topology and relationships among IT/OT devices. Finally, the two facets will be linked by identifying the impact that OT devices have on physical equipment.

Overall, the whole framework will include three layers (see Figure 8):

1. Context discovery for each provider, which creates proper abstractions of physical and cyber resources, their operational and security properties, and their relationships (blue/green graphs), including dependencies between different technology value chains (red dashed lines). The essential and distinctive characteristic is the adaptive level of granularity and aggregation, which allows for 'blur' or 'opacify' the abstraction based on trust concerns.

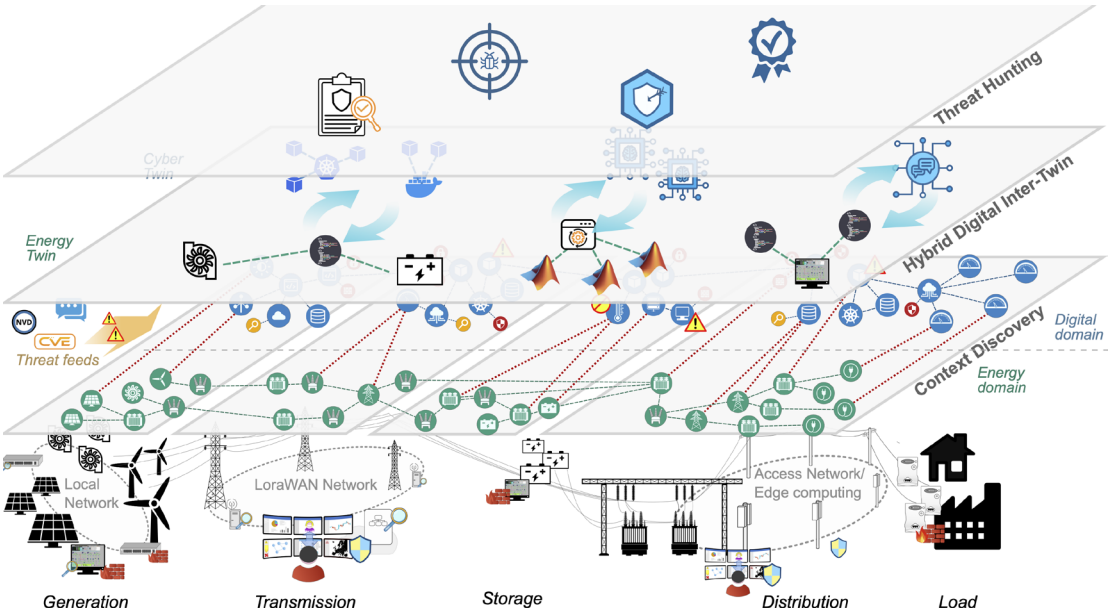

Figure 8. Hybrid digital inter-twin framework.

2. HDIT for the whole chain, which combines cyber and energy twins (including HW/SW-in-the-loop emulation) and feeds them with context and data from individual providers. The identification of mutual relationships (e.g. data communication patterns under different operating conditions of grid elements and subsystems, operational states, and energy flows in response to commands sent over data networks) is a key element for improving visibility and understanding cascading effects across the whole value chain.
3. Threat hunting over the whole system, which leverages the rich set of modelling and emulation capabilities of the HDIT to carry out proactive investigations like (i) penetration testing and dynamic/runtime application security testing (Dynamic Application Security Testing [DAST]/Run-time Application Security Testing [RAST]), (ii) compliance verification, (iii) identification of the existing threats, (iv) anticipation of new attack paths, and (v) detection of stealthy attacks and advanced persistent threats.

7.2. Reference Architecture

The architecture of the HDIT is organised in the four typical layers of DTs [28]. The two lower layers (leftmost in Figure 9) are implemented by each provider in the chain and create the common abstraction that feeds the two upper layers (rightmost in Figure 9).

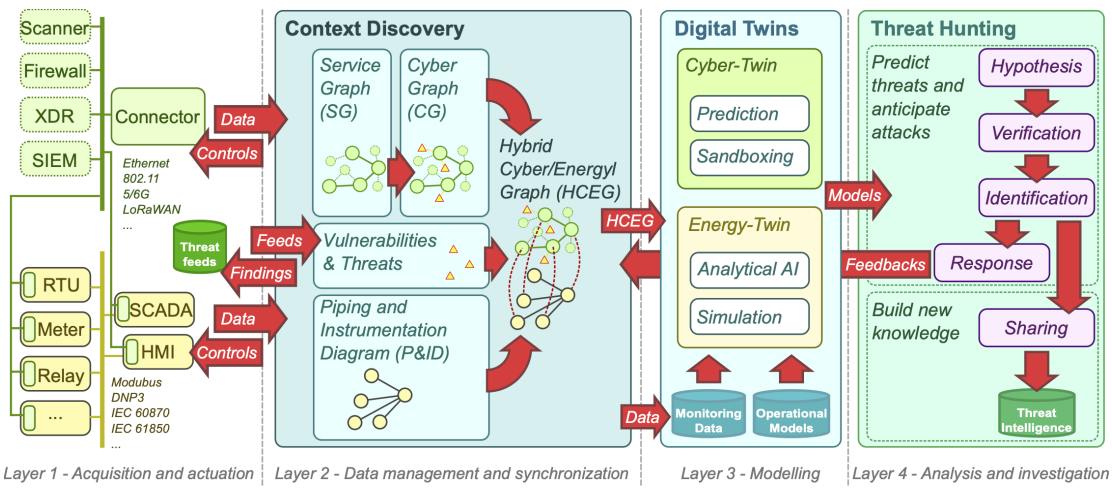

Figure 9. Hybrid digital inter-twin architecture.

The *field operation* layer perceives and controls both energy and cyber dimensions of physical systems by collecting data and measurements and by uniformly applying controls. This may leverage both consolidated and emerging protocols for this purpose (e.g. IEC 61850, IEC 61970, IEC 60870, and OpenC2) by implementing such interfaces in a zero-trust framework, targeting multi-ownership scenarios and interoperability at network interfaces. The *connector* will act as a proxy when a native interface is unavailable or inconvenient to develop.

The *context discovery* layer builds an abstraction of both physical and cyber properties that is agnostic to the specific model, serialisation, and representation, and keeps them synchronised with real system(s) (*context/controls*). This also includes known vulnerabilities and threats that may affect the deployed assets (*feeds*). This layer collects context from energy devices (SCADA, advanced metering infrastructure [AMI], remote terminal unit [RTU]), etc.), software-defined infrastructures (OpenStack, Kubernetes, etc.), and cybersecurity appliances (firewall, SIEM, eXtended detection and response [XDR]), etc.), and applies back any change made to the internal abstraction (*controls*). Beyond plain data, this layer is designed to discover both intra- and inter-domain relationships, enabling inter-twinning models in the upper layer. For this purpose, it may combine device (UPnP), service (SSDP and WS-DD), functional (MUD), model (IEC 61850 MMS), and security (OpenC2 CTXD) discovery protocols to automatically build and continuously update the system topology. The context will be represented by direct acyclic graphs denoted as service graphs (SGs) for IT/OT resources and Piping and Instrumentation Diagram (P&IDs) for energy devices, which will then be combined together and enriched with information from threat feeds (CVE, MITRE ATT&CK, and OWASP) to create hybrid cyber-energy graphs (HCEGs) for each domain. Finally, the *federation* module 'blurs' context and data before exporting to external providers in order to preserve confidentiality and privacy; similarly, it receives domain-wide response strategies as *feedbacks* and verifies their impact before turning them into local controls.

The *HDIT layer* combines *cyber* and *energy twins* together, fed by the context and real-time data from multiple providers in the chain (e.g. Eclipse Ditto) to make predictions and investigate alternative configurations and scenarios from *operational models*. The scope for the *energy twin* includes steady and dynamic energy/power flows, distributed energy resources (renewables and storage), and energy management and optimisation. The scope for the *cyber twin*

includes logs and network flows, attack strategies, and protocol operations. Both data-driven, simulation, and emulation mechanisms should be included to support the broadest set of hunting methodologies.

Several methods will be used in the *prediction* modules, including analytical (e.g. first-order logic rules), ML (e.g. graph neural networks), and large language models (e.g. multi-scale self attention networks – MSANets, transform graph [34]) for learning and anticipating common operational patterns. The *simulation* module will leverage conventional simulators as Simulink, OpenPLC4,[4] ScadaBR,[5] modRSsim2,[6] and EasyModbusTCP.[7] Finally, emulation will leverage cloud technologies (Kubernetes, Docker, and OpenStack) to *sandboxing* real-word applications and services as well as model- and hardware-in-the-loop solutions like the RTMS[8] simulator.

The *threat hunting* layer hunts for potential and ongoing threats in the system and shares new findings with the relevant community. Hunting is a four-step process that: (i) automatically generates *hypotheses* about attack paths and next steps based on the current system status, known techniques and vulnerabilities, and real-time data; (ii) carries out *verification* through HDIT; (iii) uncovers new patterns and TTPs through an identification phase that merge and reorder sequences and conditions from hypotheses which turned to be true; and (iv4) elaborates a *response* strategy and verifies its effectiveness and safety by running the HDIT under different conditions and configurations. The Knowledge process turns new findings into TTP descriptions and enriches them with the necessary context and IoCs before *sharing threat intelligence*.

7.3. Findings and Open Challenges

Our proposed framework advances from fragmented, domain-specific monitoring towards coordinated, model-driven cyber defence across DSCs. By integrating digital, cyber, and physical dimensions, it improves visibility, predictive capability, and coordinated response in multi-ownership environments. This integration introduces trade-offs in computational overhead, architectural complexity, trust management, and operational safety.

Security of massive IoT deployments: IoT and edge devices operate beyond traditional perimeters and are exposed to tampering, insecure configurations, and faulty software. Hardware-based roots of trust and zero-trust attestation mechanisms are therefore essential. Approaches based on attribute-based DAA, zero-knowledge proofs,

4 https://autonomylogic.com/
5 https://www.scadabr.com.br/
6 https://github.com/Cavaler/ModRSsim2/tree/main
7 https://sourceforge.net/projects/easymodbustcpserver/
8 https://www.rtds.com/

decentralised identifiers, and AI-assisted control-flow attestation enable continuous and privacy-preserving verification of device integrity, supporting reliable trust assessment at scale.

Interoperability: Our CDT promotes interoperability through standardised security control interfaces, with OpenC2 as a primary candidate. Nevertheless, limited maturity in open-source and practical validation constrains its demonstrated effectiveness. Constraining OpenC2 profiles to shared capabilities while allowing vendor-specific extensions mitigates integration issues but reduces full interoperability.

Threat intelligence and predictive capabilities: Automated cyber-threat intelligence (CTI) requires integrated processes for detection, contextualisation, enrichment, and attribution, leveraging Natural Language Processing (NLP) and data-driven analysis. Within our framework, the SCG concept provides a comprehensive network perspective. Predictive components – such as attack graph instantiation and sandbox-based analysis – introduce computational overhead, especially in large-scale environments. Selective model activation and adaptive twin scoping help balance prediction accuracy and resource consumption.

Cloud and edge visibility: Cloud- and edge-based environments limit traditional security monitoring. Technologies, such as eBPF[9] and DynamoRIO,[10] enable lightweight, high-speed observation of containers and IoT devices. Virtualisation further supports sandboxing and continuous vulnerability assessment, while unified configuration models enhance cross-domain situational awareness and support automated threat hunting across both infrastructure-as-code (IaC)-managed and legacy systems.

Operational security: Although HDIT enables adaptive response, misconfigured, or compromised twins may propagate unsafe control actions, zero-trust principles, attestation, fine-grained access control, and graduated response strategies mitigate these risks, reinforcing HDIT's role as a decision-support system. Maintaining synchronisation between physical systems and twins remains challenging in dynamic, federated environments. Future research will address the scalability of attack modelling, quantitative evaluation of trust mechanisms, long-term system stability, and fail-safe designs for operation under degraded trust conditions.

Implementation challenges and future directions: Maintaining consistency between physical systems and their DTs remains challenging in dynamic and heterogeneous environments, where synchronisation

9 ——— https://ebpf.io/
10 ——— https://dynamorio.org/

delays and inconsistent data can degrade attack prediction and risk assessment, particularly in federated settings. Preliminary results show that OpenC2 introduces negligible latency (a few milliseconds per command [35]). Future work will address scalable attack modelling, quantitative evaluation of trust mechanisms, long-term stability, validation across heterogeneous infrastructures, and fail-safe designs for operation under degraded trust conditions.

8. Conclusions

In this paper, we have elaborated on the use of DTs to improve cybersecurity processes. Our work has shown that the current literature primarily focuses on the use of DTs for penetration testing and vulnerability analysis but largely fails to model threats at the system- and chain-wide levels. We have proposed two novel concepts, the CDT and HDIT, that bring together digital, cyber, and physical facets within a single model. This approach has the potential to enhance the understanding of technical and operational relationships among components, and to support the anticipation of failures and cascading effects on critical systems.

Our work sheds light on a topic that has not been sufficiently investigated until now. By elaborating on how complementary components can cooperate within a common, coherent framework, it fosters new research directions to be explored with a precise target application in mind. Indeed, we are already working on the implementation of the CDT concept in the MIRANDA project.[11] Our specific contribution focuses on automating context discovery and threat hunting, boosting a proactive approach that anticipates attacks before they materialise.

11 https://www.mirandaproject.eu/

Acknowledgements

This project received funding from the European Union's Horizon Europe Research and Innovation Programme under grant agreement No. 101168144 (MIRANDA). The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Cybersecurity Competence Centre. Neither the European Union nor the granting authority can be held responsible.

References

[1] VMware, *Global incident response threat report*, 2022. [Online]. Available: https://www.vmware.com/content/microsites/learn/en/1553238_REG.html. [Accessed: Dec. 20, 2022].

[2] B. Coffin, "6G cybersecurity will transform how we deal with cyberattacks." *Access Newswire*. Oct. 24, 2022. [Online]. Available: https://www.accesswire.com/721805/6g-cybersecurity-will-transform-how-we-deal-with-cyberattacks. [Accessed: Dec. 19, 2022].

[3] T. Thibodeaux, *Smart cities are going to be a security nightmare*, *Harvard Business Review*, Apr. 28, 2017. [Online]. Available: https://hbr.org/2017/04/smart-cities-are-going-to-be-a-security-nightmare. [Accessed: Feb. 11, 2023].

[4] European Parliament , *Cybersecurity: Main and emerging threat*, 2025. [Online]. Available: https://www.europarl.europa.eu/pdfs/news/expert/2022/1/story/20220120STO21428/20220120STO21428_en.pdf. [Accessed: Jun. 7, 2024].

[5] T. Schleker, *Study on clean energy R&I opportunities to ensure European energy security by targeting challenges of distinct energy value chains for 2030 and beyond*. Final report. Luxembourg: EU Publications Office, 2024.

[6] BlueVoyant Government Solutions, *The state of supply chain defense – annual global insights report 2022*. Blueoyant Report, 2022. [Online]. Available: https://www.bluevoyant.com/resources/the-state-of-supply-chain-defense-2022. [Accessed: Apr. 4, 2023].

[7] Z. Zhou, X. Kuang, L. Sun, L. Zhong, C. Xu, "Endogenous security defense against deductive attack: When artificial intelligence meets active defense for online service," *IEEE Communications Magazine*, vol. 58, no. 6, pp. 58–64, 2020, doi: 10.1109/MCOM.001.1900367.

[8] M. Repetto, "Adaptive monitoring, detection, and response for agile digital service chains," *Computers & Security*, vol. 132, Art. no. 103343, 2023, doi: 10.1016/j.cose.2023.103343.

[9] S. Mihai, M. Yaqoob, D.V. Hung, W. Davis, P. Towakel, et al., "Digital twins: A survey on enabling technologies, challenges, trends and future prospects," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 4, pp. 2255–2291, 2022, doi: 10.1109/COMST.2022.3208773.

[10] Y. Wu, K. Zhang, Y. Zhang, "Digital twin networks: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13789–13804, 2021, doi: 10.1109/JIOT.2021.3079510.

[11] C. Zhou, D. Lopez, M. Boucadair, H. Yang, C. Jacquenet, et al., *Digital twin network: Concepts and reference architecture*, IRTF Internet-Draft, Working document of the Internet Engineering Task Force (IETF). Fremont, CA: IETF, 2023.

[12] P. Empl, G. Pernul, "Digital-twin-based security analytics for the internet of things," *Information*, vol. 14, no. 2, Art. no. 95, 2023, doi: 10.3390/info14020095.

[13] E. Glaessgen, D. Stargel, "The digital twin paradigm for future NASA and US Air Force Vehicle," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference, 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, Art. no. 1818, Reston, VA: American Institute of Aeronautics and Astronautics, 2012, pp. 1–14, doi: 10.2514/6.2012-1818.

[14] M. Batty, "Digital twins," *Environment and Planning B: Urban Analytics and City Science*, vol. 45, no. 5, pp. 817–820, 2018, doi: 10.1177/2399808318796416.

[15] A. Pokhrel, V. Katta, R. Colomo-Palacios, "Digital twin for cybersecurity incident prediction: A multivocal literature review," in *Proceedings of IEEE/ACM*

42nd International Conference on Software Engineering Workshops (ICSEW'20), Gregg Rothermel, Doo-Hwan Bae Seoul, Eds. 2020, pp. 671–678, doi: 10.1145/3387940.3392199.

[16] M. Eckhart, A. Ekelhart, "A specification-based state replication approach for digital twins," in *Proceedings of the 2018 Workshop on Cyber–Physical Systems Security and Privacy (CPS-SPC '18)*, Toronto, 2018, pp. 36–47, David Lie, Mohammad Mannan, Eds. New York: ACM, doi: 10.1145/3264888.3264892.

[17] E. Griffor, C. Greer, D. Wollman, M. Burns. (2017). *Framework for Cyber-Physical Systems: Volume 1, Overview*, National Institute of Standards and Technology, 2017. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-201.pdf. [Accessed: May 13, 2026].

[18] R. Bitton, T. Gluck, O. Stan, M. Inokuchi, Y. Ohta, Y. Yamada, et al., "Deriving a cost-effective digital twin of an ics to facilitate security evaluation," in *23rd European symposium on research in computer security, ESORICS 2018*, Ser. LNCS, J. Lopez, J. Zhou, M. Soriano, Eds., vol. 11098, Part I. Barcelona: Springer, 2018, pp. 533–554.

[19] A. Saad, S. Faddel, T. Youssef, O.A. Mohammed, "On the implementation of IoT-based digital twin for networked microgrids resiliency against cyberattacks," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5138–5150, 2020, doi: 10.1109/TSG.2020.3000958.

[20] A. Bécue, Y. Fourastier, I. Praça, A. Savarit, C. Baron, et al., "CyberFactory#1 — Securing the industry 4.0 with cyber-ranges and digital twins," in *14th IEEE International Workshop on Factory Communication Systems (WFCS)*, Imperia, 2018, IEEE (Piscataway, USA), pp. 1–4. doi: 10.1109/WFCS.2018.8402377.

[21] Y. Yigit, B. Bal, A. Karameseoglu, T.Q. Duong, B. Canberk, "Digital twin enabled intelligent DDoS detection mechanism for autonomous core networks," *IEEE Communications Standards Magazine*, vol. 6, no. 3, pp. 38–44, 2022, doi: 10.1109/MCOMSTD.0001.2100022.

[22] A. De Benedictis, F. Flammini, N. Mazzocca, A. Somma, F. Vitale, "Digital twins for anomaly detection in the industrial Internet of Things: Conceptual architecture and proof-of-concept," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 12, pp. 11553–11563, 2023, doi: 10.1109/TII.2023.3246983.

[23] M. Atalay, P. Angin, "A digital twins approach to smart grid security testing and standardization," in *IEEE International Workshop on Metrology for Industry 4.0 & IoT*, Roma, 2020, IEEE (Piscataway, USA), pp. 435–440, doi: 10.1109/MetroInd4.0IoT48571.2020.9138264.

[24] M. Dietz, L. Hageman, C. von Hornung, G. Pernul, "Employing digital twins for security-by-design system testing," in *Proceedings of the 2022 ACM Workshop on Secure and Trustworthy Cyber–Physical Systems (Sat-CPS '22)*, Baltimore, MD, 2022, pp. 97–106, M. Gupta, S. Khorsandroo, M. Abdelsalam, , Eds. New York: ACM, doi: 10.1145/3510547.3517929.

[25] A. Salvi, P. Spagnoletti, N. Saad Noori, "Cyber-resilience of critical cyber infrastructures: Integrating digital twins in the electric power ecosystem," *Computers & Security*, vol. 112, Art. no. 102507, pp. 1–11, 2022, doi: 10.1016/j.cose.2021.102507.

[26] A. De Benedictis, C. Esposito, A. Somma, "Toward the adoption of secure cyber digital twins to enhance cyber-physical systems security," in *Quality of information*

and communications technology, A. Vallecillo, J. Visser, R. Pérez-Castillo, Eds. Cham: Springer, 2022, pp. 307–321, doi: 10.1007/978-3-031-14179-9_21.

[27] E. Hadar, D. Kravchenko, A. Basovskiy, "Cyber digital twin simulator for automatic gathering and prioritization of security controls' requirements," in *IEEE 28th International Requirements Engineering Conference (RE)*, Zurich, 2020, pp. 250–259, IEEE (Piscataway, USA), doi: 10.1109/RE48521.2020.00035.

[28] C. Alcaraz, J. Lopez, "Digital twin: A comprehensive survey of security threats," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 3, pp. 1475–1503, 2022, doi: 10.1109/COMST.2022.3171465.

[29] D. Holmes, M. Papathanasaki, L. Maglaras, M.A. Ferrag, S. Nepal, et al., "Digital twins and cyber security – Solution or challenge?" in *6th South-East Europe Design Automation, Computer Engineering, Computer Networks and Social Media Conference (SEEDA-CECNSM)*, Preveza, 2021, IEEE (Piscataway, USA), doi: 10.1109/SEEDA-CECNSM53056.2021.9566277.

[30] *Input to the horizon Europe programme 2021–2027: Priorities for the definition of a strategic research and innovation agenda in cybersecurity*. Brussels: ECSO Publications, 2020.

[31] Q. Liu, P. Li, W. Zhao, W. Cai, S. Yu, V.C.M. Leung, "A survey on security threats and defensive techniques of machine learning: A data driven view," *IEEE Access*, vol. 6, pp. 12103–12117, 2018, doi: 10.1109/ACCESS.2018.2805680.

[32] C. Gehrmann, M. Gunnarsson, "A digital twin based industrial automation and control system security architecture," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 1, pp. 669–680, 2020, doi: 10.1109/TII.2019.2938885.

[33] S. García, M. Fresia, J. Mora-Merchán, A. Carrasco, E. Personal, C. León, "A data-driven topology identification method for low-voltage distribution networks based on the wavelet transform," *Electric Power Systems Research*, vol. 243, 2025, doi: 10.1016/j.epsr.2025.111517.

[34] Q. Zhang, J. Chen, G. Xiao, S. He, K. Deng, "TransformGraph: A novel short-term electricity net load forecasting model," *Energy Reports*, vol. 9, pp. 2705–2717, 2023, doi: 10.1016/j.egyr.2023.01.050.

[35] M. Repetto, "Otupy: A flexible, portable, and extensible framework for remote control of security functions," *Computers & Security*, vol. 158, Art. no. 104597, pp. 1–20, 2025, doi: 10.1016/j.cose.2025.104597.