Proposal No. 830929
Call H2020-SU-ICT-03-2018

Project start: 1 February 2019
Project duration: 42 months

Cyber Security for Europe

D3.23

Cybersecurity Outlook 2

Document Identification	
Due date	31 Jun 2022
Submission date	14 June 2022
Revision	1.0

Related WP	WP3	Dissemination Level	PU
Lead Participant	POLITO	Lead Author	Daniele Canavese (POLITO)
Contributing Beneficiaries	POLITO, ATOS, CNR, GUF, UMU, UNITN, UPS-IRIT, NTNU, SINTEF	Related Deliverables	D3.10

Abstract: This document presents the deliverable "D3.2 – Cybersecurity Outlook 2". It presents the outcome of the task T3.9 Continuous Scouting, whose goal is to constantly analyze the current state-of-the-art, new trends, and emerging technologies in the cybersecurity field. This document collects our findings, conclusions, research, and recommendations in various security-related areas, from highly technical ones (e.g., software and network security) to law and human-centric ones (e.g., governance and cybersecurity awareness).

This document is issued within the CyberSec4Europe project. This project has received funding from the European Union's Horizon 2020 Programme under grant agreement no. 830929.

Executive Summary

This deliverable reports the outcomes and findings of task T3.9 Continuous Scouting. It is a report detailing our discoveries and results in bleeding-edge technologies, new emerging trends, and discussions on how the cybersecurity field has recently impacted our lives. This document is a continuation of some topics described in D3.10 Cybersecurity Outlook 1 and an extension in several new areas.

All the investigations reported in this document are accompanied by a bibliography containing the most interesting and recent work in several cybersecurity-related areas, primarily scientific papers published in the last few years.

All our studies are grouped in several macro-categories for ease of organization and clarity. The first chapters focus on the most technical aspects of the cybersecurity world, such as software and network security, how to preserve privacy with the most recent technologies, and a discussion about 6G, the next-generation standard for wireless communications. Finally, we conclude with an entire chapter devoted to less technical matters as discussing the most recent evolution of laws concerning some cybersecurity issues and how the security and privacy of our data are affecting society.

This deliverable is of interest to everybody looking for the current directions and trends in the cybersecurity world from both a technical and non-technical point of view. Aside from research scientists, this document can also be interesting for industrial partners and companies as food for thought since it may provide a glimpse into the near future and evolution of cybersecurity.

Document information

Contributors

Name	Partner
Daniele Canavese	POLITO
Antonio Lioy	POLITO
Aljosa Pasic	ATOS
Angelica Liguori	CNR
Massimo Guarascio	CNR
Francesco Sergio Pisani	CNR
Giuseppe Manco	CNR
Manuel Cheminod	CNR
Luca Durante	CNR
Narges Arastouie	GUF
Dirk Müllmann	GUF
Christina von Wintzingerode	GUF
Indra Spiecker gen. Döhmann	GUF
Pablo Fernandez	UMU
Juan Francisco Martinez	UMU
Agustin Marin	UMU
Jorge Bernal	UMU
Antonio Skarmeta	UMU
Carlos E. Budde	UNITN
Fabio Massacci	UNITN
Silvia Vidor	UNITN
Afonso Ferreira	UPS-IRIT
Abdelmalek Benzekri	UPS-IRIT

Pierre-Henri Cros	UPS-IRIT
Elvire Prochilo	UP-SIRIT
Sunil Chaudhary	NTNU
Vasileios Gkioulos	NTNU
Shukun Tokas	SINTEF
Gencer Erdogan	SINTEF

Reviewers

Name	**Partner**
Renee Undrits	CYBER
João Resende	C3P

History

Version	**Date**	**Authors**	**Comment**
0.01	2022-03-29	Daniele Canavese, Antonio Lioy	Initial ToC
0.02	2022-04-22	Daniele Canavese, Angelica Liguori, Massimo Guarascio, Francesco Sergio Pisani, Giuseppe Manco, Afonso Ferreira, Dirk Müllmann, Christina von Wintzingerode, Abdelmalek Benzekri, Pierre-Henri Cros, Indra Spiecker gen. Döhmann, Elvire Prochilo, Juan Francisco Martinez, Agustin Marin, Jorge Bernal, Antonio Skarmeta, Carlos E. Budde, Fabio Massacci, Silvia Vidor, Sunil Chaudhary, Vasileios Gkioulos	Added Sections 2, 3.2, 4.1, 6.1, 6.2, 6.3

0.03	2022-05-04	Narges Arastouie	Added Section 5
0.04	2022-05-05	Aljosa Pasic	Added Section 6.4
0.05	2022-05-06	Shukun Tokas, Gencer Erdogan	Added Section 3.1
0.06	2022-05-09	Manuel Cheminod, Luca Durante	Adde Section 4.2
0.07	2022-05-09	Daniele Canavese	Added the abstract
0.08	2022-05-10	Daniele Canavese	Added the executive summary, introduction, and conclusion
0.09	2022-05-10	Daniele Canavese	Fixed the history and some affiliations
0.10	2022-05-10	Daniele Canavese	Fixed the formatting of several paragraphs and images
0.11	2022-05-11	Daniele Canavese	Merged Section 3 and 4 and various fixes in the bibliography
0.12	2022-05-23	Shukun Tokas, Gencer Erdogan	Fixed some typos in Section 3.1
0.13	2022-05-24	Daniele Canavese	Addressed most of the reviewers' comments
0.14	2022-05-25	Daniele Canavese	Shortened some titles, extended the introduction and the first paragraphs of Section 5, and grammar clean-up
0.15	2022-05-26	Narges Arastouie	Addressed the reviewers' comments in Section 4
0.16	2022-05-29	Daniele Canavese	Minor grammar clean-up
1.0	2022-06-14	Ahad Niknia	Final check, preparation and submission

Table of Contents

List of Figures

List of Tables

List of Acronyms

#	**3GPP**	3rd Generation Partnership Project
A	**AE**	AutoEncoder
	AHP	Analytic Hierarchy Process
	AI	Artificial Intelligence
	ANN	Artificial Neural Network
	API	Application Programming Interface
	AR	Augmented Reality
	ARN	Adversarial Reconstruction Network
B	**BGP**	Border Gateway Protocol
C	**CC**	Competence Community
	CCW	Certain Conventional Weapon
	CD-IM	Code Domain-Index Modulation
	CERN	Conseil Européen pour la Recherche Nucléaire
	CHECK	Community Hubs of Expertise in Cybersecurity Knowledge
	CI	Confidence Interval
	CI/CD	Continuous Integration/Continuous Delivery
	CLAW	Cyber Lethal Autonomous Weapon
	CMC	Crude Monte Carlo
	CNN	Convolutional Neural Network
	CNO	Collaborative Networked Organisation
	CoMP	Coordinated Multi-Point transmission
	CR	Cognitive Radio
	CR-BCI	Cognitive Radio Brain-Computer Interaction

	CRN	Cognitive Radio Network
	CSA	CyberSecurity Awareness
	CTI	Cyber-Threat Intelligence
D	**DDoS**	Distributed Denial-of-Service
	DEI	Digitising European Industry
	DIH	Digital Innovation Hub
	DL	Deep Learning / Distributed Ledger
	DLT	Distributed Ledger Technology
	DNNs	Deep Neural Networks
	DP	Differential Privacy
	DRL	Deep Reinforcement Learning
E	**ECCC**	European Cybersecurity Competence Centre
	ECSO	European Cyber Security Organisation
	EEA	European Economic Area
	eMBB	enhanced Mobile BroadBand
	EI	Edge Intelligence
	ENI	Experiential Networked Intelligence
	ENISA	European Network and Information Security Agency
	EP3R	European Public-Private Partnership for Resilience
	ETSI	European Telecommunications Standards Institute
	EU	European Union
F	**FDD**	Frequency Division Duplex
	FIG	Finite Improbability Generator
	FL	Federated Learning

	FT	Fault Tree
G	**GANs**	Generative Adversarial Networks
	GDPR	General Data Protection Regulation
H	**HAP**	High-Altitude Platform
	HTC	Holographic-Type Communication
I	**IBFD**	In-Band Full Duplex
	ICRC	International Committee of the Red Cross
	ICT	Information and Communications Technology
	IdM	Identify Management
	IDS	Intrusion Detection System
	IEC	International Electrotechnical Commission
	IEEE	Institute of Electrical and Electronics Engineers
	IETF	Internet Engineering Task Force
	IF	Isolation Forest
	IHL	International Humanitarian Law
	ILS	Iterated Local Search
	IM	Index Modulation
	IoE	Internet of Everything
	IOPS	Input/Output Operations per Second
	IOSA	Input/Output variant of Stochastic Automata
	IoT	Internet of Things
	IPFS	InterPlanetary File System
	IRS	Intelligent Reflecting Surface
	ISAC	Information Sharing and Analysis Centre
	ISG	Industry Specification Group

	ISPLIT	Importance SPLItting
	ISTN	Integrated Space and Terrestrial Network
	IT	Information Technology
	ITU-T	International Telecommunication Union – Telecommunication standardization bureau
K	**KPI**	Key Performance Indicator
F	**FG-ML5G**	Focus Group on Machine Learning for Future Networks
	FL	Federated Learning
	FW	FireWall
L	**LAW**	Lethal Autonomous Weapon
	LEO	Low Earth Orbit
	LLVM	Low Level Virtual Machine
	LSTM	Long-Short Term Memory
	LTE	Long-Term Evolution
M	**MBA**	Mixed-Boolean Arithmetic
	MC	Molecular Communication
	MIMO	Multiple-In, Multiple-Out
	MISP	Malware Information Sharing Project
	ML	Machine Learning
	MR	Mixed Reality
N	**NAT**	Network Address Translation
	NCC	Network of National Coordination Centre

	NG-RAN	Next Generation Radio Access Network
	NIB	Network In Box
	NIS	Network and Information Security
	NIST	National Institute of Standards and Technology
	NGMN	Next Generation Mobile Networks
	NFV	Network Function Virtualization
	NLP	Natural Language Processing
	NoC	Network on a Chip
	NoE	Network of Excellence
	NR	New Radio
	NS	Network Slicing
O	**OBSIDIAN**	Open Banking Sensitive Data Sharing Network for Europe
	OC-SVM	One-Class Support Vector Machines
	OFDM	Orthogonal Frequency Division Multiplexing
	OT	Operational Technology / Oblivious Transfer
	OWC	Optical Wireless Communication
P	**PET**	Privacy Enhancing Technology
	PCTL	Probabilistic Computation Tree Logic
	PLS	Physical Layer Security
	PPP	Point-to-Point Protocol / Public-Private Partnership
Q	**QKD**	Quantum Key Distribution
R	**RACI**	Responsible, Accountable, Consulted, Informed
	RAN	Radio Access Network
	R&D	Research & Development

	R&D&I	Research & Development & Innovation
	RES	Rare Event Simulation
	RIS	Reconfigurable Intelligent Surface
	RF	Radio Frequency
S	**SACM**	Security Automation and Continuous Monitoring
	SAE	Sparse Autoencoder
	SBA	Service Based Architecture
	SCDG	System Call Dependency Graph
	SD-IM	Space Domain-Index Modulation
	SDN	Software Defined Network
	SM	Space Modulation
	SMC	Statistical Model Checking
	SME	Small-to-Medium Enterprise
	SN	Social Network
	STA	Stochastic Timed Automata
T	**TATIS**	Trustworthy APIs for Threat Intelligence Sharing
	TDD	Time Division Duplex
	TEE	Trusted Execution Environment
	TRL	Technology Readiness Level
U	**UAV**	Unmanned Aerial Vehicle
	umMTC	ultra-massive Machine Type Communications
	uRLLC	ultra-Reliable Low Latency Communications
V	**VAE**	Variational Autoencoder

	VLC	Visible Light Communication
	VR	Virtual Reality
W	**WALDO**	Wasserstein Autoencoder for Learning the Distribution of Outliers
X	**XR**	eXtended Reality
Z	**ZSM**	Zero-touch network & Service Management

1 Introduction

CyberSec4Europe is an ambitious project whose goal is to strengthen the research and innovation in the cybersecurity area at national and international levels. Aside from technical work and designing various security-related tools, task T.39 Continuous Scouting aims to investigate new trends and technologies. The goal of these studies is not only to be kept updated on bleeding edge technologies and how laws and society are reacting to the most recent developments but also to offer interesting thoughts and ideas to academics and companies.

This document detail our findings and recommendations in various cybersecurity-related areas. All the discussions are accompanied by a collection of references published (mostly) in the last few years. The first document produced by T3.9 was deliverable D3.10 Cybersecurity Outlook 1, and this report can be thought both as a continuation of D3.10, since we reprise some arguments and update them with new developments, but also as an extension, since we expanded our investigations in several new directions, especially the less technical ones.

This deliverable collects a variety of cybersecurity-related topics ranging from highly technical ones such as threat intelligence via machine-learning and privacy management technologies, but also themes about how new trends and technologies are transforming our society.

Most of the topics reported in this deliverable were presented and discussed during the Privacy Symposium[1], held in Venice (Italy) on April 6th, 2022. This document, however, greatly expands what we discussed at the conference and contains new material and several additional topics.

1.1 Structure of the document

For the sake of organization, all the topics are grouped into several categories. The first chapters (from Section 2 to Section 4) examine the most technical aspects of cybersecurity. In contrast, Section 5 is entirely devoted to discussing a more human-centric view of how laws and our society is changing with respect to the world of cybersecurity. Every section contains a conclusion sub-section summarizing our findings and, occasionally, giving some recommendations in a specific area.

This document is structured as follows:

- Section 2 presents a discussion on some recent developments in the software security field;
- Section 3 describes how to increase the security in networks and discusses how to share data in a privacy-preserving manner;
- Section 4 is a dissertation about 6G, the next-generation wireless standard;
- Section 5 consists of a variety of topics related to the interaction between cybersecurity and our society;
- Section 6 contains the conclusions.

[1] https://privacysymposium.org/

2 Software security

Due to its ubiquitous nature, software dominates several aspects of our daily lives. Writing secure software that is bug-free and resilient to attacks has become a necessity that not all industries have welcomed. First, writing high-quality, secure software requires highly trained and expert developers. Second, this process frequently needs several assessment rounds, thus significantly increasing a product's time to market. To help human beings in this process, researchers have started to investigate several automatic or semi-automatic approaches in the last years.

2.1 Reverse engineering protected software

Software protections are widely used in a variety of commercial applications and videogames. They are used for various purposes, such as to avoid stolen intellectual properties (e.g., proprietary algorithms) or private data (e.g., credit card numbers), fighting piracy to safeguard license checks, or simply preventing users from cheating in video games.

The most common protection techniques are obfuscations, a family of methods whose basic idea is to increase the code complexity to hinder the code understanding by an attacker. By applying enough obfuscations, the time to successfully reverse engineer an application is so high that it becomes (economically) unworthy.

Reverse engineering a piece of protected software is the act of removing or bypassing protection applied to a piece of code. This operation is usually performed on binaries by looking at the assembly code level. Although attackers carry out reverse engineering for piracy purposes, it can also be used for several defensive reasons. Trying to reverse engineer commercial software can help assess the strength of its protections. Several companies worldwide have embraced this approach before releasing the product. On the other hand, several recent malware (e.g., viruses and worms) use protection techniques (e.g., packing, where the code is stored compressed, and decompressed on the fly when needed) to thwart the detection by anti-viruses. The analysis of these techniques can then be successfully used to increase the identification accuracy of anti-viruses and malware detectors.

The traditional approach is to reverse engineer a binary manually or with minimal automatic support in the form of ad-hoc scripts. However, a new trend has emerged slowly: trying to automate most operations using machine learning or symbolic analysis techniques.

2.1.1 Protected region identification

When reverse engineering a binary file, the first step is to identify the protected areas since this is a strong indication that they can be the assets the attacker is looking for (e.g., the license check function to crack). This operation is traditionally performed by hand with the help of disassemblers and debuggers, thus making it a slow and laborious task. However, some recent developments have shown that AI-based methods can be successfully used to quickly analyze a binary and detect the protected regions. NLP (Natural Language Processing) techniques seem particularly promising in this context. Based on deep learning, NLP was born for dealing with human languages (e.g., English or Italian). However, programming languages are still languages, so NLP can be used to reason on assembly instructions. Current state-of-the-art shows that these techniques can successfully detect patterns that are strong indicators of a protection presence.

The work of Kim et al. [1] uses a simple combinatorial neural network for detecting obfuscated functions in binaries for the Intel platform protected with Obfuscator-LLVM[2]. The approach presented by the authors

[2] https://github.com/obfuscator-llvm/obfuscator

is to analyze a binary by first disassembling it, then counting the occurrence of some specific opcodes that are then sent to a fully connected neural network for protection detection. This work shows an accuracy of 91% when a single obfuscation technique is used and an accuracy of 85% when two protections are applied to the same function.

Zhao et al. [2] implemented a mixed-NLP approach for identifying functions protected with Obfuscator-LLVM or Tigress[3] on Intel architectures. Their methodology consists of two major processing phases. First, after disassembling the binary, they create a custom embedding for the instructions that are grouped into basic blocks and then analyzed by a CNN (Convolutional Neural Network). Second, the neural network's output is then fed to an LSTM (Long-Short Term Memory) recurrent neural network for the final classification of the function. The tests showed that this approach could detect a single protected function with an accuracy of 91% and snippets safeguarded by two obfuscations with an average accuracy of 88%.

Canavese, Regano, and Basile [3] presented a patent for detecting functions protected with Obfuscator-LLVM, Tigress[4] , and Diablo on ARM and Intel architectures. The proposed approach leverages the power of the Radare2[5] disassembler and debugger to create a custom encoding for the assembly instructions. It then uses state-of-the-art NLP neural networks to analyze binary code sequences and classify them. Tests show that the accuracy for identifying a function obfuscated with a single protection technique is about 97%. In comparison, the accuracy slightly drops to 92% when the proposed system is used to identify functions protected with two or three obfuscations.

2.1.2 Code understanding

Once a potential asset or region of interest has been found, the attacker must perform some form of code understanding before attempting to remove the protection or bypass it somehow. This phase is usually completed by disassembling or decompiling the code. Hence this task is very time-consuming and requires significant skills and experience to be successful.

Symbolic execution is an interesting approach that can be used to analyze a piece of code quickly. The core idea of this technique is to simulate the code execution without actually running the application under test. To achieve this, a graph-like state model of the code is built, and then it is explored, thus emulating its execution. Though symbolic execution is not a new technique, it has started to gain popularity in the last few years for several reasons. On the one hand, advancements in computational power have opened the door for analyzing symbolically bigger applications and chunks of code. On the other hand, potentially dangerous applications (e.g., malware) can be investigated safely without running them. To reduce this technique's computational and memory consumption, concolic execution is sometimes used instead. The term concolic stems from the fusion of concrete (debugging) and symbolic execution. Launching an application with an attached debugger and analyzing its traces symbolically significantly decreases the model space, allowing a faster analysis.

Borisov and Kosolapov [4] tackle the problem of assessing the quality and security of obfuscating a piece of code by automatically computing some quantitative metrics by applying a symbolic execution approach. The authors propose to build a feature vector that depicts both the static and dynamic characteristics of an obfuscated binary by analyzing symbolically its original (non-protected) version and its obfuscated

[3] https://tigress.wtf/
[4] https://github.com/csl-ugent/diablo
[5] https://rada.re/

variation. Once these two vectors are obtained, a similarity score can be computed, effectively quantifying the strength of the applied obfuscation.

Sebastio et a. [5] proposed a system for automatically and quickly analyzing (obfuscated) malware via symbolic and concolic analysis. In this novel approach, the authors use the angr[6] binary analysis framework to perform the initial symbolic/concolic analysis of the binary, tuning its parameters to decrease the analysis time. Then, an SCDG (System Call Dependency Graph) structure is built. This data structure contains all the system calls invoked by the malware and their code correlations, thus, it is an excellent indicator of the behavior of a program. Subsequently, this graph is analyzed using gSpan[7], a simple yet effective software for graph mining to classify the malware and detect its type (e.g., worm, ransomware, crypto-miner). The experimental results showed that they could obtain an accuracy close to 97%.

Namani and Khan [6] also used a symbolic execution technique to process protected and unprotected malware with a mixed static and dynamic approach. A binary is disassembled, and its header and static library call sites are identified. Then, a symbolic/concolic execution is performed where the API calls are extracted. This mix of static and dynamic data is transformed into a feature vector, eventually performing feature hashing to convert variable-length features (e.g., strings) into fixed values. These vectors are then used to train various machine-learning classifiers to detect if a binary is malware or benign. The authors tested three different types of classifiers: decision trees, random forests, and fully connected neural networks. All the models showed similar results, achieving an accuracy ranging from 92% to 97%.

2.1.3 Protection removal/bypass

Once the code region of interest has been located and a sufficient code understanding phase has been performed, the attacker might try to remove or bypass the protection. This step requires a high level of expertise and patience since it usually involves writing and rewriting large chunks of assembly code by hand or with minimal automatic support by ad-hoc scripts. In the last few years, some work has started to appear proposing how to perform this code rewriting process in a completely automated fashion. Most of the work in this context focuses on obfuscations, the most common protection techniques used. Although automatic deobfuscation is still in its infancy, it has already shown potential, especially when coupled with AI-based approaches.

Menguy et al. [7] proposed Xyntia, an AI-based black-box deobfuscation framework for simplifying code protected with several. The core idea of Xyntia is to treat the deobfuscation as an optimization problem where the goal is to find the most straightforward code that is semantically equivalent to the protected region. The authors then used a variation of ILS (Iterated Local Search), an optimization technique known to work well in odd-shaped search spaces. The authors then showed that Xyntia could reconstruct several obfuscation techniques, such as Mixed-Boolean Arithmetic (MBA), a protection technique that transforms arithmetic expressions into more complex equivalent ones, and opaque predicates, an obfuscation approach that adds bogus branches.

David, Coniglio, and Ceccato [8] instead used concolic execution to automatically undo several different obfuscation techniques such as MBA, data-encoding, and virtualization (a protection technique where the assembly code is translated into bytecode that is then executed by a custom embedded virtual machine). Their approach works in two distinct phases. First, the application is run via concolic execution. Then, an abstract syntax tree is constructed by looking at the symbolically analyzed traces, and a simplified version is extracted using a top-down breadth-width search.

[6] https://angr.io/
[7] https://sites.cs.ucsb.edu/~xyan/software/gSpan.htm

Kochberger et al. [9] performed a systematic literature review of various approaches used to remove virtualization obfuscated functions, a common protection technique used by malware to hide the malicious code from anti-viruses. They experimentally tested 15 automatic deobfuscators on various protected applications with mixed results and analyzed their ability to understand the virtualized bytecode. In addition, they proposed a novel taxonomy for the deobfuscation techniques by looking at the artifacts used by the deobfuscators (e.g., traces, control flow graphs), the analysis type (e.g., static, dynamic, or hybrid), and the degree of automation (fully automatic or partially automatized).

2.1.4 Conclusion

Reverse engineering a piece of protected code is a task that requires expertise, time, and patience. Commonly performed by attackers for cracking commercial software, reverse engineering is also interesting for defensive purposes such as evaluating the security of an application before its release or helping anti-viruses detect new forms of malware.

The traditional approaches to reverse engineering employ commonly used tools such as debuggers, decompilers, and disassemblers with minimal automatized support. However, several attempts have been made to automatize most of these operations in the recent few years. Machine-learning and optimization techniques seem very promising in helping a person in this context. Successfully automatizing the reverse engineering process will considerably impact the future of software protection in many different regards: the quality assessment of commercial software can be significantly sped up, and the detection ability of anti-viruses can be drastically increased for the new types of obfuscated malware.

2.2 Using statistical model checking for cybersecurity analysis

Taken-for-granted technologies in today's digital societies include personal health-care appliances, assisted-driving cars, conversation robots, nightly build CI/CDs, etc. Distributed data storage with a central access point stands among these feats; In fact, from the technologies mentioned, cloud storage is arguably the one with the biggest immediate impact in our lives. Having your account stolen on social media apps is a twenty-first century bonfire story, not to mention the more serious implications of this happening with accounts that contain, e.g., your banking data.

Cybersecurity is usually portrayed as the guardian angel in these modern dreadful scenarios. In particular, the central access point that makes cloud technologies so useful for everyone's everyday life, is often what also makes them prone to attacks by malicious parties. Thus, good practices for access security—e.g., strong keys and multi-factor authentication—are enforced rather than expected from users. But not every practice can be enforced, since high-security and ease-of-use typically stand on opposite sides of the user-experience spectrum[8]. Therefore, enforceable user-sided security is at best of limited reach.

The complexity of the situation increases as servers, distributed worldwide, run on heterogeneous hardware and firmware [33]. While the user-experience challenge is to offer everyone a homogeneous and equally-responsive interface, cybersecurity strives to achieve a common minimal degree of protection. Ideally, system design can achieve this in a manner resilient to software heterogeneity and evolution; but developers seldom have the chance to design a system from scratch. More often than not, companies must work with legacy software, so the choices available to developers—already constrained w.r.t. enforceable user-sided

[8] https://cybercompetencenetwork.eu/

security—is restricted even more. In fact, it is not uncommon that the only feasible choice is which new library version to update to [26].

These are practical reasons why cybersecurity cannot be black or white, e.g., a social media platform is not either cyber-safe or -unsafe. The literature agrees that any software is ultimately susceptible to cyberattacks, which motivates concepts like cyber resilience [26, 23, 30]. Therefore and as usual, for companies, it boils down to investment strategies: how to spend resources in a manner that reduces the risk of cyberharm, viz. the degree of protection mentioned above.

Such investments range from buying specialised hardware and software, or implementing security policies, up to training attack-response teams, all the way through hiring more/specialised personnel to keep the system updated. The ultimate goal is to lower the risk of enduring or recovering from impactful attacks: the more vulnerable the system, the higher the investment needed. Thus, estimating the degree of software vulnerability has become an increasingly hot research topic [35, 23, 28].

Technically, however, vulnerability estimation comes with many theoretical and practical hurdles. One of the hardest to overcome is the sheer unpredictability of future technologies. In fact, most endeavors take an ostrich approach here and focus on known vulnerabilities, typically zero-day attacks, avoiding to speculate on issues to come. This line of action is mainly chosen due to the complexity of the field, where a seemingly innocuous code fragment can be exploited, but only when accessed via a specific browser with certain plugin installed.

In the face of such complexity not much can be projected accurately from existent code. Instead, this work discusses how an abstraction step can be taken to analyze a formalized model of the system's security. The added value of such an approach is that it can be used to estimate, with an arbitrary degree of accuracy, the likelihood and time window between the exploitation of software vulnerabilities.

2.2.1 Model-checking and cybersecurity

Formal system modeling and analysis are based on mathematical descriptions of systems whose properties are queried using (typically) temporal logic formulae. The field is vast, and a large part of it deals with model checking due to its attractive push-button approach, where formal checks can be fully automated [10].

Model-checking. The fundamental steps for model checking are:

1. defining a model M that describes the system to be analyzed;
2. defining a property φ that describes the query to perform;
3. checking whether (or the degree to which) the model satisfies the property, which is typically denoted $M \vDash \varphi$.

The subsequent formal guarantees on the automatically computed answer have resulted in many success stories of model checking applied to safety analysis—a trend that continues to this day [19, 21, 25, 29].

For example, in [29], the query from step 3 becomes "what is the probability that the power supply noise of my NoC surpasses the safety threshold". It is compelling to see the resemblance of such queries to measuring the degree of (cyber-) security resilience. This has, in fact, started to be noticed as researchers begin to apply model checking for general cybersecurity studies [160, 34].

There is, however, a zeroth step that precedes modeling: the selection of the formalism in which M and φ will be given semantics. This is crucial as it determines the type of questions that can be asked. For example, the semantics on which the model M is interpreted must contain time (namely, allow for a continuous state space) in order to query about the duration of events. Also, and quite to the point of attack-resilience, the chosen formalism must allow for probability measures in order to query about the likelihood of an event occurring. We now discuss these matters for cybersecurity analysis.

Semantic basis. In automata theory, many mathematical formalisms can express either time or probabilities [20]. Arbitrary combinations of these aspects are less common, and the complexity of the resulting models quickly reaches un- decidability even for reachability properties. That is, if the semantics are chosen to be too expressive, there may be no algorithm that can compute, e.g., whether a vulnerable situation is reachable. Computational efficiency is also a factor to consider: the more flexible the model, the more computation steps (and runtime) it will take for an algorithm to find the answer to a query.

Therefore, the modeling formalism must be chosen as expressive as needed and as simple as possible. For the case of software vulnerabilities, we are interested in two types of questions:

- "what is the probability of an attack in a defined time window?",
- "what is the expected time between independent attacks?".

Both questions are stochastic in nature, and the second one requires the estimation of potentially continuous time intervals. The simplest formalism from the literature that can cope with both continuous probability measures and time are Stochastic Automata, a subset of STA [159, 20].

A Stochastic Automata model M can encode the occurrence of attacks, according to empirical probability distributions adjusted from real-world cases. Then, a PCTL-like property φ can query the time-bounded probability of observing one such relevant event in the foreseeable future.

This regards steps 1 and 2 from the model checking procedure. However, step 3 encounters the extra requirement of verifying arbitrary distributions, which come from approximations of the empirical attack probabilities observed in the real world. Such potentially non-Markovian behavior rules out traditional (probabilistic) model checking and calls for simulation-based variants, in what is usually called statistical model checking (SMC [36]).

2.2.2 Statistical model checking for cybersecurity

SMC integrates Monte Carlo simulation with formal methods. Via discrete event simulation, it generates traces of states, which are samples of the states that the stochastic model M can visit. Via the generation and analysis of these stochastic samples, SMC estimates the degree to which M satisfies different properties.

Modeling considerations. For cybersecurity, the states S of the model M could encode (a) the first-level libraries that a main project depends on, (b) the number of known vulnerabilities for the own codebase and also for these libraries, (c) the criticality of these vulnerabilities, (d) the time since their publication, and (e) whether any of these codebases is currently under attack.

Note that, since the Markovian property can be dropped, values such as the number and criticality of vulnerabilities need not multiply the number of states. Instead, they can be kept as rewards that decorate the states during a simulation.

Item (e) above is speculative and defines the goal of the simulations. More precisely, a temporal logic property φ can query the probability of transitioning from the current safe state to a state in which one or more of the codebases is under attack, before T days have passed. Then SMC estimates this value (the step $M = \varphi$) by generating several samples, and computing the proportion of them that suffered an attack before T days[9]. Here, φ is said to characterize a subset of the states $S_\varphi \subset S$, whose reachability we are estimating.

[9] Transitions among states are governed by stochastic distributions, that describe the jump probabilities from past evidence. Stochastic Automata encode this via "clocks".

Thus, from M and φ, an SMC analysis yields an estimate $\hat{\gamma} \in [0,1]$ of the actual probability γ with which the model satisfies φ, i.e., the likelihood of an attack. Besides producing $\hat{\gamma}$, SMC can quantify the statistical error incurred via two numbers, $\delta \in (0,1)$ and $\varepsilon > 0$, such that $\hat{\gamma} \in [\gamma - \varepsilon, \gamma + \varepsilon]$ with probability δ. Thus, if $n \in \mathbb{N}$ traces are sampled, the full SMC outcome is the tuple $(n, \hat{\gamma}, \delta, \varepsilon)$.

This statistical quantification is usually returned as a confidence interval (CI) around $\hat{\gamma}$, and coveys an idea of the quality of the estimation. The usual approach is to fix the confidence δ prior to experimentation: then higher quality means smaller ε and thus a narrower CI, achieved by drawing more samples.

Computation considerations. Although flexible and automatic, the SMC approach is hindered by rare events. That is, if there is a very low probability γ to satisfy φ, then most traces sampled by SMC will not visit Sφ. The result is then either an incorrect estimate $\hat{\gamma} = 0$ or, if a few traces do visit S_φ, the confidence interval computed is very wide and hence uninformative.

This can affect cybersecurity analyses, since the likelihood of observing an exploit of a vulnerability is quite low in practice. To counter such phenomena, the number of samples n must increase as γ decreases. Unfortunately, for the sample mean, this causes n to increase in inverse proportion to the square of γ, which quickly results in unacceptably long run times. To tackle this issue, rare event simulation (RES) methods have emerged in many scientific disciplines [31].

Rare Event Simulation. Roughly speaking, RES can be divided into importance sampling and importance splitting (ISPLIT). The former modifies the stochastic transitions of the model, in a way that can later be undone when computing the estimate $\hat{\gamma}$. This is not clearly feasible for cybersecurity, where the transition distributions are arbitrary as they come from empirical data.

In contrast, ISPLIT methods are not directly affected by such matters, which makes them more attractive to our purposes. A caveat is that ISPLIT traditionally requires expert knowledge to split the state space S of the model M. This dampens the use of SMC as an automatic approach for cybersecurity analysis, since it necessitates user input beyond the definition of M.

However, novel theories are emerging to finally automate this step [11, 12, 24]. Next, and to conclude this work, a statistical model checker that implements automatic RES is briefly presented.

2.2.3 The FIG tool

The Finite Improbability Generator (FIG) is an SMC tool publicly available[10]. It uses the formal definitions of M and φ to derive the so-called importance function f and thresholds $\{\ell_i\}_{i=1}^{M}$ [13, 14]. These are the core components needed by ISPLIT to speed up the statistical convergence for the computation of the estimate $\hat{\gamma}$ [22].

For this, FIG runs a breadth-first search from S_φ, on the (inverted) transitions of M. This computes the number-of-transitions distance from each state to S_φ,. The heuristic importance function of FIG, f^*, is the inverse of this distance, stored as an array the size of S. To avoid the state explosion FIG works on a modular formalism called IOSA (a subset of Stochastic Automata), deriving a local f^* for each M_i whose parallel composition forms M. f^* is an aggregation of these functions, which in its most basic form adds the local f^* of every M_i whose variables appear explicitly in φ [12, 15].

Function f^* is solely based on the distance measured in number of transitions of M. All stochastic behavior that is omitted by f^*, such as probabilistic weights in the transitions, is captured in the thresholds ℓ_i. To choose these thresholds automatically FIG runs dynamic analyses, using either Expected Success [15] or a

[10] https://git.cs.famaf.unc.edu.ar/dsg/fig

variant of the Sequential Monte Carlo algorithm [16]. In both cases, finite-life simulations start from S_0, to estimate roughly the probability to reach states with higher importance via lightweight statistical analyses.

Demonstration. Finally, the capabilities of FIG are shown to study rare-event properties in two small examples from its test suite. The first is a triple tandem queue with Erlang service times: the IOSA model file is publicly available on the official website of FIG in the following path: `tests/models/3tandem_queue.sa`.

We compare crude Monte Carlo (CMC) and two RES strategies with the monolithic importance function, i.e. f^* built on the composition of all IOSA modules. The first strategy uses all of FIG default parameters, and the second one requests Expected Success to build thresholds and the restart engine with level-2 prolongations. The corresponding commands are:

```
fig --stop-time 5m 3tandem_queue.sa --cmc
fig --stop-time 5m 3tandem_queue.sa -amono
fig --stop-time 5m 3tandem_queue.sa --amono -t es -e restart2
```

We estimated the property $\varphi = S(q3 = 7)$, which asks the proportion of time that the third queue contains more than seven elements. Comparisons were done for a fixed simulation budged, namely a wall-clock time of 5 minutes. When the time is due, simulations stop, and CIs are reported: the estimation that achieves the narrowest CI for a fixed confidence level is the most efficient one.

Running these experiments in an Intel(R) Xeon(R) E-2124G CPU @ 3.40GHz (Linux kernel 5.14.8-arch1-1) resulted in the following 95% CIs: [3.81e-6, 4.52e-6] for CMC, [4.15e-6, 4.36e-6] for FIG defaults, and [4.25e-6, 4.40e-6] for the custom command. The widths of these intervals are 7.13e-7, 2.12e-7, and 1.53e-7 resp.

All CIs overlap and contain the expected value of 4.25e-6. However and as expected, RES can achieve tighter estimates for the same simulation budget. We highlight that the default FIG command is as bare as crude Monte Carlo, yet it produced an estimate more than three times more precise.

Finally, we experiment with a second model: a small repairable Fault Tree with non-Markovian failure and repair times (`FT.sa`), also available on the website of FIG as `tests/models/resampling_tiny_FT.sa`. The distribution families include exponential, Erlang, normal, and lognormal.

The case is quite interesting since ISPLIT has limited applications in FT analysis. Importance functions such as f^*, that only observe failures and repairs of components, resulting in efficient RES applications if and only if the dominant failure can be layered, e.g., as the result of the conjunctive failure of many subcomponents. To exploit this we have developed heuristics that automatically derive a composition strategy from the FT structure. A similar approach is envisioned for cybersecurity studies, using the closely related theory of attack tree analysis.

For this case, we estimated the time-bounded probability of observing a system failure before 150 time units. Again we compare CMV and two RES strategies: FIG with the `--ft` switch, Expected Success thresholds, Fixed Effort simulation engine, and (a) the default compositional importance function, and (b) the heuristic FT-structure importance function. The commands are:

```
fig --stop-time 5m FT.sa --cmc
fig --stop-time 5m FT.sa --ft -t es -e sfe --acomp +
fig --stop-time 5m FT.sa --ft -t es -e sfe --acomp \
```

```
'BE_0+max(BE_1,BE_2)+BE_4'
```

These experiments resulted in the 95% CIs: [1.93e-4, 3.02e-4], [2.28e-4, 3.12e-4], and [2.39e-4, 2.70e-4], whose widths are 1.09e-4, 8.41e-5, and 3.12e-5. As before, all CIs contain the expected value (2.65e-4), and RES achieved the tightest intervals for the same simulation budget. In this case, however, the difference between CMC and the default compositional strategy of FIG is much less pronounced than in the previous example. This is expected given the low redundancy required to cause a system failure (three components must simultaneously fail).

Yet, despite this, the heuristic composition strategy performed significantly better, producing a CI almost an order of magnitude narrower than CMC. Perhaps the most appealing feature of this strategy is that it is automatic: it is computed from the FT structure from which also the IOSA modules were created. In other words, this is effectively a fully-automatic deployment of RES. In subsequent research, we intend to apply analogous approaches to study the properties of models that encode cybersecurity problems.

2.2.4 Conclusion

SMC is a formal approach to model analysis via Monte Carlo simulation. Stochastic Automata semantics are proposed as underlying formalism: they combine continuous-time and probabilities, as required to estimate the likelihood and time of occurrence of future attacks. The need for rare event simulation is identified to achieve efficient computations, and the academic tool FIG is presented, which can deploy it automatically.

3 Network security and privacy

Since the rise of the Internet in the 90s, our world has become even more digitalized. Thanks to the revolution of cloud technologies and IoT devices in the last few years, we are more interconnected than ever. We use the Internet every day, and every day we transmit precious and sensitive data through third-party servers and devices. In this context, securely sharing this massive load of information requires protection against a variety of increasingly complex distributed attacks and to be treated in a privacy-preserving manner to avoid disclosing personal data.

3.1 A need for privacy assistive technology in notice and consent paradigm

The Internet of Things (IoT) describes the network of physical objects embedded with sensors, software, and other technologies to exchange data with other devices and systems over the Internet. These objects deployed in public and private spaces enable use cases that enhance productivity and quality of life. There is a strong incentive for IoT devices to be cheap, which leads to poor security and privacy. Several research studies and surveys reveal that privacy concerns are at an all-time high, as the collection and use of data in IoT are happening with very little or no control, and organizations collecting data are most of the time unknown to data subjects. For example, a Norwegian population survey [37] reveals that more than 50% of participants feel uncomfortable about commercial actors collecting information about them. CISCO's value/trust paradox report [38] reveals the divide between IoT value and trust: 53% of participants feel IoT makes their life more convenient, while only 9% trust that their data collected and shared through IoT is secure. Despite the trust deficit and perceived risk, 42% say IoT is too integrated into their lives to disconnect from IoT services. This growing trend of lack of transparency and absence of support for data subjects to control the collection and processing of their data in IoT may heavily affect many areas of our lives and even constitute a long-term danger for democracy.

In response to the emerging privacy concerns, the European Parliament has approved the General Data Protection Regulation (GDPR) [39] to strengthen and impose data protection across the European Union (EU) and the European Economic Area (EEA). Several studies have investigated the impact of GDPR on consumers. For instance, a survey carried out by Cisco [40] confirms that 55% of respondents view GDPR very favorably, and Godinho and Adjerid [41] found that only 6.2% of participants gave opt-in consent to the personal data collection (e.g., location data). In this article, we focus on a crucial privacy concept, i.e., consent. According to the GDPR, there are six lawful bases for processing personal data: contractual necessity, consent, legal obligation, vital interests, public interests, and legitimate interests. The service providers of data-enabled technologies or smart infrastructure must carry out the processing of personal data within the limit of the applicable processing grounds. Consent is one of the most discussed bases of processing, and is also a core principle of data protection as "it relates to the exercise of fundamental rights of autonomy and self-determination" [42]. Consent is the lawful ground that reflects a data subject's agreement and provides the data controller with permission to process a subject's personal data for specific purposes. Arguably, consent is often the most exploited legal ground for processing personal data. Most of us have had these experiences of giving consent in different situations where stakes can be low, e.g., small financial transactions, browsing news on the Internet, or even where stakes can be high, e.g., medical procedures, legal transactions, continuous monitoring through wearable technology. Consumers accept all the risks detailed in privacy notices, without even reading them. For example, an increasing number of consumers are using sophisticated fitness trackers, capable of sensing bodily states with precision, with very little awareness of the privacy risks of the collection and processing of fine-grained data. Privacy notices at large seem to be more about preventing litigation for data controllers rather than serving their real purpose, i.e., to aid in making informed privacy choices.

The paradigm of notice and consent, widely known as 'notice and choice', is based on a presupposition that consumers will adequately manage their privacy, if provided sufficient information about data collection and processing [43]. In fact, the GDPR resulted in more detailed and longer privacy notices. As computing becomes ubiquitous, the continued reliance on consumers to read several dozen of such notices and make an informed privacy decision every day is a practical problem. In addition to complicated privacy notices, and due to other prevalent malpractices, such as dark patterns, hidden default settings, notification overload, and low privacy notice comprehension, privacy notices fail their purpose of protecting privacy. This leads to the provocation of paradoxical behavior [44], i.e., despite being concerned about their privacy, consumers do not take necessary actions to protect their personal data, and privacy resignation [45], i.e., data subjects give up managing their privacy settings. Acquisti et al. [46], in their analysis of surveys, field studies, and experiments in privacy literature, conclude that privacy management that relies purely on consumer responsibilization has failed.

Research has shown that comprehending privacy notices imposes a high cognitive and time burden on data subjects [47]. In order to address this widespread issue of uninformed consent, research efforts are needed to advance the state-of-the-art by utilizing the existing research (e.g., [45, 48– 50]) to devise privacy notices more intuitive to consumers so that these notices are more likely to be read and understood. Even if the readability challenge is addressed and notices are made more comprehensible, there will remain the challenge of attending to numerous privacy notices on a regular basis. Privacy recommendations can assist IoT users in making meaningful privacy decisions. Privacy-preference recommenders can generate privacy recommendations based on users' behaviors with similar privacy tendencies, such as level of privacy awareness, sensitivity towards privacy, and understanding of utility-privacy trade-offs. The diverse privacy preferences of users need to be captured and modeled across different data collection and usage scenarios. However, it is important to make such recommendations in a privacy-preserving manner because privacy preferences can also be used to make inferences about the user and thus needs to be treated as personal data.

3.2 Privacy-preserving CTI sharing for enhanced intrusion detection in the financial sector

Over the last few years, financial entities have been improving the digitalization of their critical processes, thus amplifying the attack surface being exposed. Meanwhile, cyberattacks against the Financial Sector, such as those based on DDoS, Phishing, Ransomware, and Identity theft, are evolving quickly and becoming more refined. In the current state, attackers may hurt several entities without even changing the attack vector due to the lack of resources to provide fast reaction and response actions. In this context, Threat Intelligence Platforms (TIPs) enable organizations to exchange critical threat information in order to prevent further similar attacks to those that have already occurred. However, the sensitive nature of data managed in the financial sector may cause entities to refuse to share certain information mainly because of privacy concerns.

Given these circumstances, there is a need to propose a privacy-preserving and trust-minimized network that enables financial organizations to share critical information related to specific attacks or fraud events they may encounter, providing strong mechanisms to ensure at least: access control to the exchanged information, empowering organizations to keep control over what they want to share and with whom, the privacy of certain data being shared, enabling the obfuscation of sensitive attributes related both to stakeholders and organizations in the sharing network, and auditing resources, allowing involved actors to assess the provenance, integrity, and immutability of shared information during its whole lifecycle.

In recent years, several works have been proposed in this line. For instance, [51] proposed a solution based on a combination of Hyperledger Fabric and InterPlanetary File System (IPFS), along with real-world use cases extracted from the MITRE ATT&CK framework. The result is an ecosystem that can share and store threat files in a secure and trusted way. However, the work does not cover the application of anonymization techniques to sensitive attributes, so the receiver of the data, although trusted, still can access sensitive data, such as the identification number of a bank account in a financial environment.

On the other hand, [52] presented a privacy-preserving protocol for threat intelligence sharing based on the collaborative training of a decision tree classifier and exchanging training data between organizations in a private way using homomorphic encryption. Anyhow, this approach can be replaced by Federated Learning [53], which is a paradigm with a much broader background, so the need to design an alternative method for collaborative learning is removed. Moreover, this approach only provides a final machine learning model, but there may be scenarios in which it is necessary to access the original data for provenance and inspection purposes, e.g., in the event of cybercrime.

That way, a system that covers the shortcomings of most state-of-the-art proposals is presented on the basis of the TATIS asset [54], enhancing it with the use of Blockchain for auditing purposes, Privacy-Enhancing Technologies (PETs) for anonymization, and Federated Learning techniques to come up with a complementary workflow in which machine learning models are exchanged instead of data.

3.2.1 Framework

The framework involves several components that, in conjunction, create a trusted privacy-preserving network whereby financial entities can exchange threat information with other entities in the federation. It is designed to perform in conjunction with OBSIDIAN, which is an alternative network to exchange critical information about effective frauds, leveraging the latest online open banking services.

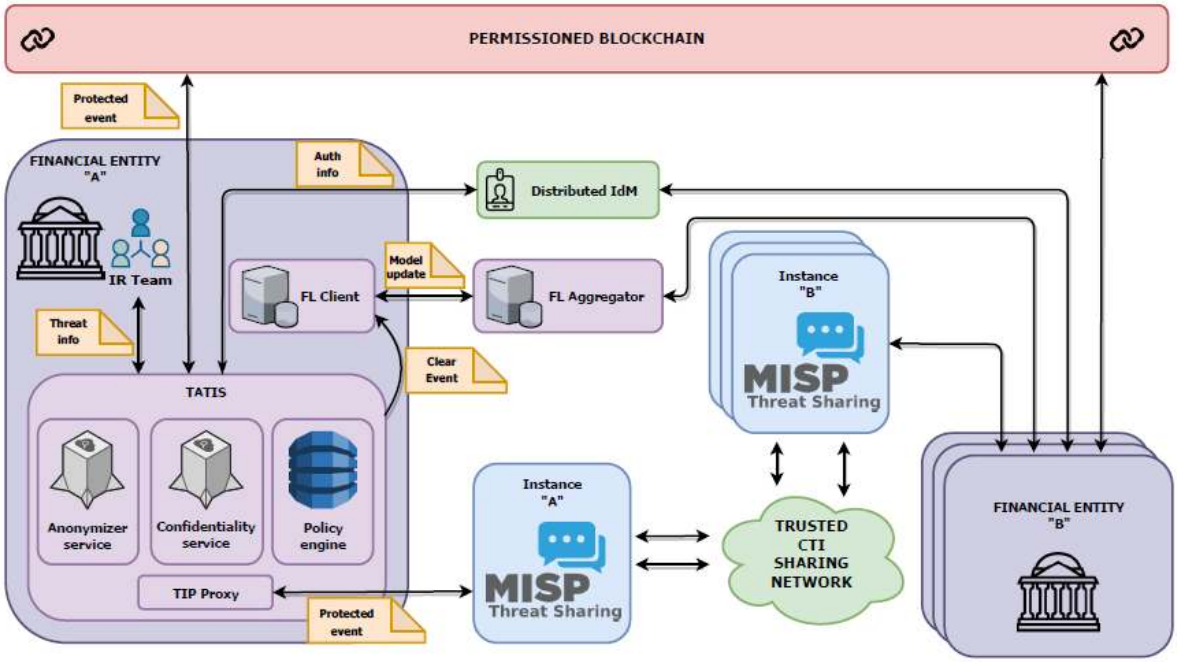


Figure 1 - CYTILIS architecture

The architecture of the solution can be consulted in **Error! Reference source not found.**. This scenario depicts the overall mode of operation of the whole system. As it should be noted, firstly, in addition to the encryption techniques for access control purposes offered in TATIS, the asset has been extended in order to provide an anonymizer service that is able to obfuscate certain sensitive attributes of an event also based on policies specified by the user.

This module supports several anonymization algorithms, such as K-anonymity [55], L-diversity [56], and T-closeness [57], which are previously selected by the user based on the specified policies. These policies are defined through extensible and inter-operable policy models that follow a specific design, that in turn allows determining adaptively which technique should be applied for each type of object and for each specific attribute of each object. In the end, it enables the anonymization of events with several different kinds of objects.

The permissioned blockchain is used to securely and traceably store a series of data related to the provenance of the sharing event, such as the publisher, user, the anonymization processes, hashes of the event, and hashes of the privacy policies applied for each obfuscated attribute. Thus, entitled users will be able to query and verify in real-time not only a published event, associated privacy policies, and techniques applied but also, in case an inspection is needed, to request a comparison of the original event with the anonymized one. This would help, for instance, to uncover an attacker who is delivering malicious obfuscated events.

Moreover, the FL scenario is set up with a single aggregator and multiple FL clients, ideally one per entity or domain, receiving training data (threat events) from TATIS in real-time. It may be the case that some entities do not implement this functionality, but nevertheless, they could still receive the final model as it will be uploaded and broadcasted to the MISP network once the federated training has finished, based on [58].

Furthermore, the scenario will also be endowed with a distributed identity management system (IdM) in order to ensure that users are authenticated and have permission to share or receive events

3.2.2 Conclusion

A first framework implementation shows the feasibility of the proposal to anonymize exchanged CTI data and increase the overall anomaly detection accuracy in the Federated Learning setting compared to a centralized deployment. The proposed scenario works in a fully automated and autonomous mode, following a ”zero-touch” security management approach, thereby requiring the user to only define the privacy-preserving policies and other initial parameters, such as FL-specific configurations. In this way, the solution will be provided as a service, so final users will be able to access it through an API containing all the described functionality ready to use.

3.3 Generative methods for out-of-distribution prediction and applications to threat intelligence

Anomalous data and shifts in the data distribution can greatly affect the performances of Machine Learning (ML) models. It mainly depends on how these models are learned. Indeed in real application scenarios, the closed-world assumption, i.e., training data are a representative sample of the test ones, is not realistic. From a cybersecurity perspective, these perturbations on data can concern both adversarial and defense strategies. Therefore, learning robust models against these changes is a challenging and relevant research topic.

Outlier detection and generation techniques represent a precious tool for identifying malicious behaviors and strengthening the Artificial Intelligence (AI) based security systems. While outlier detection is a well-known research area in ML research [59] and has been widely adopted for identifying anomalies in several cybersecurity scenarios (e.g., Network and Host Intrusion Detection Systems, Malware Detection, Fraud Detection, etc.), outlier generation is a relatively new research line and focuses on devising algorithms for generating realistic examples so to enrich the ground truth for the model learning. Recently, with the rise of deep generative models, there has been a growing interest in these solutions [60], [61], [62] since they represent an effective method for generating realistic cyberattacks.

In this section, we survey some recent Deep Learning (DL) based methods and techniques for outlier detection and generation, and we provide an overview of the main application scenarios in the cybersecurity domain in which these approaches can be effectively adopted.

3.3.1 Deep learning models for outlier detection

Class imbalance, i.e., the overwhelming of a class w.r.t. the other ones, represents the main issue in devising outlier detection techniques. In many cybersecurity scenarios, a large part of the available data contains information on normal/expected behaviors. Therefore, unsupervised and semi-supervised approaches are the most frequently used methods for addressing this problem. In particular, one-class classification (e.g., One-Class Support Vector Machines (OC-SVM) [63], distance-metrics, e.g., Isolation Forest (IF) [64], or nearest neighbor algorithms [65] are considered state-of-the-art techniques in this research field. In the last years, with the rise of Deep Learning, there has been a growing interest in anomaly detection models exploiting Deep Neural Networks (DNNs) [66]. This is mainly due to the capability of DNNs to directly learn from raw data in incremental fashion without the necessity of a manual feature engineering phase. In this context, solutions based on the usage of autoencoders (AE) [67] have increasingly gained attention among researchers and practitioners [68], [69], [70], [71], [72].

An autoencoder is a neural architecture trained to reproduce as output a duplicate of its input. It is composed of two components: (i) the encoder, a neural network whose goal is to map the input to a low dimensional latent space, (ii) the decoder, a further neural network that, given the latent representation, aims at producing an output as close as possible to the original input. The main capability of the AE is to yield an encoding that ignores the noise. Therefore, the adoption of these architectures for anomaly detection tasks represents a natural choice since high reconstruction errors on new data are likely related to anomalies (e.g., cyberattacks). Basically, the reconstruction error estimates the outlierness score associated with a given instance. In more detail, AE is trained only against normal samples. Therefore, they learn to model normal behaviors. When a deviant example is provided as input, the output will diverge from the expected one, and an alarm is raised. However, the threshold τ, required to establish if a score is anomalous, plays a crucial role in this framework since it greatly affects the performance of the model. Some preliminary studies tried to address this issue, e.g., [73] introduced a (weakly) supervised anomaly detection model able to discover anomalies without setting τ. Basically, the proposed AE integrates two decoders, respectively named inlier decoder (D_in) and outlier decoder (D_out). These sub-networks are trained in a competitive fashion and allow modeling inliers and outliers' data distributions. Unlabelled data are provided as input to both the subnetworks, which are responsible for labeling them as outliers or inliers on the basis of the reconstruction errors. Specifically, the strategy relies on comparing the scores yielded by two decoders: if the inlier reconstruction error is smaller than the outlier reconstruction error, the inlier label is assigned to the instance, the second one otherwise.

A further emerging research line in outlier detection focuses on the usage of Generative Adversarial Networks (GANs) [74]. The main idea is to define a single comprehensive generative framework in which the detection process is directly embedded. Basically, two components, respectively generator G and discriminator D, are trained through an adversarial process [62]. The generator attempts to yield examples able to fool the discriminator and undermine its predictive performances, whereas the discriminator acts in the role of estimating the probability that a given example belongs to the training data or generated ones by G. The generative distribution over the data is estimated by learning the mapping from an initial noise distribution (used as prior) to the data space.

AnoGAN [75] represents the first attempt to adopt a GAN framework for outlier detection. The main idea consists in using an iterative process for learning the mapping from the latent space to realistic normal examples. The final result of this process is a point z in the latent space such that the generated sample is as similar as possible to the input one. The similarity between generated and input data is used as an anomaly score and calculated as a combination of residual and matching loss [76]. Low values of anomaly scores are associated with examples seen in the training phase (i.e., the normal data), while a high outlierness is likely

associated with an anomaly. The main issue of this approach relies on the adopted backpropagation procedure, which reduces the scalability of the approach.

Some subsequent studies tried to address this issue by extending AnoGAN framework [77], [78], [79]. As an example, ALAD [79] improves the model by introducing an encoder to refine the discriminator detection capabilities.

An encoding-decoding framework is adopted for combining latent representation with adversarial learning in GANomaly [80]. The main difference w.r.t. other GAN-based approaches relies on the generator that includes three components. An encoder-decoder architecture is devoted to mapping the input data into a latent space and vice versa. A different encoder is employed to learn a further representation from the reconstructed data space to a latent one. In this case, the outlierness is computed as the distance between the two learned latent representations. This model has been extended in [81], in which the usage of skip connections improves the convergence of the model and the detection performances.

In ADAE [82], an autoencoder architecture is used to model both the discriminator and generator. Here, the idea is to use the reconstruction error of the discriminator autoencoder to estimate the anomaly score. Finally, recent approaches propose to combine ensemble learning with autoencoders [83] or GANs [84]. Ensembles are well-known ML methods, where the output of several models trained against different data samples or using different algorithms are then combined according to a given strategy for classifying new unseen instances. In the GAN framework, they are typically used by exploiting combinations of different generators and discriminators.

3.3.2 Anomaly generation via deep learning

Artificial outlier generation represents an effective solution for addressing the class unbalance problem and can be exploited for both evaluating the prediction quality of the outlier detection models and improving their performances by feeding them with these data.

The DL-based architectures discussed in the following exploit a general approach that consists in adapting probabilistic generative models so to yield anomalous examples by sampling from low-density regions.

For example, a VAE-based approach for generating synthetic time series, including also outliers, is proposed in [85]. Basically, the solution consists of sampling anomalies from the outlier region of the latent space. The main issue in using this approach depends on the training data distribution, i.e., complex manifolds in original data can lead to learning overgeneralized models exhibiting weak performances.

WALDO - Wasserstein Autoencoder for Learning the Distribution of Outliers – defined in [86] is a supervised framework integrating into a single solution, detection, and generation capabilities. In more detail, it integrates the Wasserstein autoencoders [87] with the technique proposed in [73]. WALDO architecture is composed of two decoders, respectively, for inliers and outliers, stacked on a common encoder that is trained by using the Wasserstein distance. The objective is to learn how to minimize the Wasserstein Distance among inliers and outliers.

A different approach is used to learn FenceGAN (Ngo, et al., 2019). The standard GAN framework is extended on the basis of the following observation, i.e., adversarial samples tend to overlap the true distribution. Then, the main idea consists in devising a generator able to yield samples distributed within the boundaries of the data distribution. As a result, the learned discriminator is robust against difficult outliers, i.e., able to correctly classify these challenging examples.

ARN (Adversarial Reconstruction Network) [88] combines Variational Autoencoders (VAE) and Generative Adversarial Networks for the generation and the identification of outliers. The main goal is to generate realistic outliers for the learning phase of an outlier detector. Specifically, ARN exploits normal data to generate their abnormal counterpart, i.e., a synthetic but realistic outlier that is similar to the normal data but differs from it for minimal but substantial differences. The generation of realistic outliers enables

the learning of an outlier detector that is able to identify real outliers without the need to be fed with explicit information about them.

3.3.3 Applications in the cybersecurity domain

Since the main aim of cybersecurity is to protect systems, organizations, and users against attacks (in many cases, anomalous and rare events) perpetrated by black hats, the usage of outlier detection techniques in the cybersecurity domain has become more popular in recent years. As highlighted in the previous sections, many different types of malicious behaviors can be considered attacks, e.g., network intrusions, insurance, and credit card frauds, diffusion of fake information, and image manipulation. In principle, all these deviant behaviors can be discovered by exploiting anomaly detection techniques that are able to efficiently process large amounts of data and also identify slight changes w.r.t. an expected behavior. In the following, we illustrate some of the most relevant cybersecurity scenarios where these techniques can be adopted.

Intrusion Detection Systems (IDSs) are hardware and software systems able to identify malicious behaviors and are widely reckoned as an important tool for timely recognizing security breaches and attacks. The problem of detecting malicious activities in a host or network environment [89] is currently studied, and many approaches have been proposed in the literature. In particular, since the intrusions are typically rare events exhibiting slight differences in terms of traffic flow statistics, outlier detection techniques based on GANs can represent an effective solution to discover these anomalies, as shown in [90]. A BiGAN-inspired model combined with a custom loss function is adopted for identifying intrusions in computer networks.

Fraud detection represents a challenging issue in several application scenarios, and the automatic discovery of these fraudulent behaviors represents a very important task with great impact in many real-life situations. The main goal consists in detecting criminal activities and preventing unauthorized operations. In this domain, outlier detection techniques can be adopted to detect (i) fraudulent credit card applications or usage, (ii) misused usage of a mobile phone account, or (iii) insurance fraud claims in which an attacker manipulates the claim processing system for unauthorized and illegal claims [91]. As an example, [68] proposes a DL-based model to address the fraud detection problem. The framework exploits a combination of Sparse Autoencoder (SAE) and GAN to detect whether a transaction is a fraud or not.

The main aim of the Image Manipulation Detection is to discover any unauthorized modification in an image. Splicing, copy-move, and removal are some of the most common techniques used in this field to manipulate the content of an image. Manipulated images may seem normal to the human eye, but typically they can exhibit artifacts that makes them anomalous. A different type of manipulation relies on the usage of an information hiding technique. As an example, by using steganography techniques, malware or sensitive data can be embedded into an image and sent/exfiltrated. Once again, the human eye is not able to perceive the differences w.r.t. the original image (i.e., the image without the hidden information). Therefore, AI-based tools can represent an effective solution for analyzing the image and discovering tamperings, in particular when they are synthetically generated via advanced manipulation tools. For example, [92] proposes a deep forgery discriminator (DeepFD) based on the usage of contrastive loss to detect computer-generated images.

Timely detection of fake news spreading over well-known social networks (SN) represents a relevant research topic since fake/errata news can grow in popularity, becoming a real fact for the entities belonging to such SN communities. Fake news can be regarded as low-quality news with intentionally false information [93]. With the aim to fight the rise and spreading of fake news, different automatic detection methods based on AI and ML have been proposed in the literature. In particular, the recent results obtained by DL techniques in complex natural language processing tasks make them a promising solution for effectively detecting fake news. From this perspective, although in literature, Deep Learning approaches have already been adopted to detect fake news, the usage of GAN-based models is still unexplored. In

particular, GAN-based models have been mainly employed to address similar tasks, such as fake review detection. As an example, [94] demonstrated as FakeGAN, an architecture based on the simultaneous usage of two discriminators, can be effectively used to detect deceptive reviews in a real scenario.

3.3.4 Conclusion

The usage of Machine Learning techniques for automizing and improving the detection of cyberattacks is rapidly gaining importance. Unfortunately, AI-Based tools may exhibit poor performances when deployed in real environments since the training data used in the learning phase could greatly differ from the test ones. The development of new methods and techniques for learning models robust against out-of-distribution represents an emerging and challenging research line. In this section, we described some approaches based on innovative Deep Learning architectures moving a step in this direction and analyzed their application in different cybersecurity scenarios.

3.4 Distributed firewall reconfiguration in industrial networks

Modern interconnected systems are constantly facing numerous cyber security threats coming from heterogeneous sources. This is more and more true and alarming in industrial networks and, in general, in critical infrastructures. Such systems have strong connections with the physical world, and security-related accidents can quickly escalate to physical damage and harm to persons, together with costly production stops.

Industrial systems have experienced a slow but steady evolution toward the modern principles of Industry 4.0. There is, in fact, a convergence in technologies and architectures between the OT (Operational Technology) and the standard IT (Information Technology) worlds. This convergence is required to enable a more flexible and dynamic approach to production and also to optimize management costs. The widespread trend is to avoid a strict dependence on proprietary technologies for either devices or communications, as this does not facilitate high-level integrations among different parts of production systems. Instead, the convergence of technologies and communication protocols does encourage the creation of dynamic systems that can quickly adapt to market needs and high-level management decisions.

However, the other side of the coin is that modern industrial systems have an ever-increasing number of interconnections between different parts of the plants and with the external world. Moreover, many industrial plants are nowadays geographically distributed while belonging to the same communication network, which is possible through Internet-based channels. To add further complexity, we have to take into account the ever-increasing need for remote connections (e.g., remote maintenance, smart working, and so on) that open further communication channels. Given this scenario, there is a constantly rising awareness of the critical aspect of cyber security in these systems [95].

In this described context, we have to consider the possible sources of threats that are no more limited to the "Internet" or the "external world". Internal threats are more and more critical in such scenarios, including accidental misuses by legitimate users (operators, employees), software threats (malware, compromised devices), and malicious users (disgruntled employees). In this context, a specific aspect to consider is the security of the communication network and, in particular, the effective protection against unwanted or malicious traffic [96].

A fundamental element in this regard is the availability of filtering devices, that is firewalls. These devices have the specific purpose of checking and evaluating the network packets flowing in the system, allowing only those that belong to acceptable flows, and blocking the others. A correct deployment and configuration of such devices implement a network security policy. In the considered scenario of large distributed systems, it is reasonable to expect the deployment of multiple firewalls distributed throughout the network [97]. In industrial systems, this is very desirable to obtain the "defense in depth" suggested by many standards and best practices (e.g., IEC 62443 [98]), that is, to protect the different logical and physical layers in which we

can partition the system. However, two main problems arise when dealing with firewall deployment: the number and location of firewalls and their configuration. Here, we specifically address the second problem, and in particular, we consider the case of anomalous traffic in the network that can overload one or more firewalls. This kind of problem is particularly critical in industrial networks as overloaded firewalls cause excessive delays in network packet delivery to the point the packets can be lost. In both cases, these effects can involve critical communication flows in the production plant, thus breaking the correct execution of production processes and causing damage. The main idea here is to recognize that one or more firewalls are entering this critical condition and act to relax their load, preventing the severe consequences of lost and delayed packets. As the performance of many firewalls depends on the number of rules in their configurations, a viable approach to reduce the problem is to re-configure these firewalls and distribute the rules throughout available devices.

Our proposed approach to the problem is to leverage the presence of multiple firewalls along the critical communication paths. In particular, we identify, in any given topology, a hierarchy of firewalls depending on the overloaded traffic flow direction.

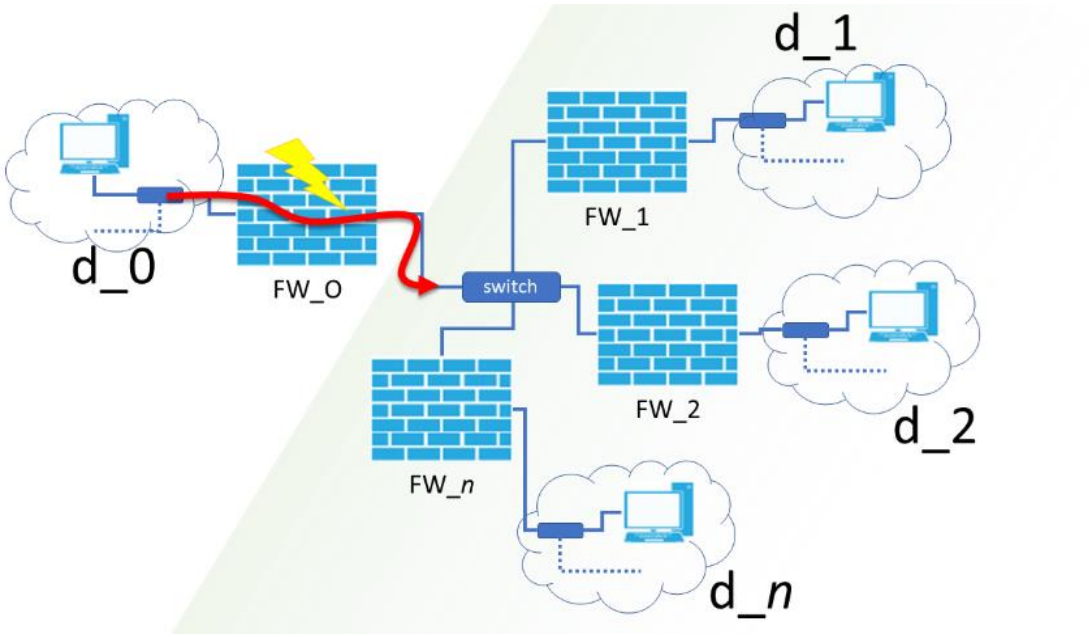


Figure 2 - reference topology

This situation is depicted in Figure 2, where the anomalous traffic is highlighted in red and originates from the domain d_O. This traffic crosses the overloaded firewall FW_O and reaches a protected domain on the other side (highlighted in green in Figure 2). We can further identify several sub-domains (e.g., d_1, d_2, …, d_n) in the protected area that may be protected by internal firewalls (FW_1, FW_2, …, FW_n). It is worth noting that domains d_1, d_2, and so on, represent sub-networks that can be further detailed with their internal nodes and topology. The main idea is to decompose the configuration on the overloaded firewall into several rule partitions that can be distributed to the internal firewalls. Reducing the number of rules deployed on FW_O effectively reduces the cost of managing the incoming packets, allowing the firewall to avoid the saturated condition. The partitioning of the rules is performed by analyzing the source and destination fields of all the rules. If a rule source matches the d_O domain and the destination matches only a specific sub-domain d_i, then the rule is flagged as belonging to the partition P_i. If the last condition does not hold, the rule is flagged as "multi-domain". When the analysis ends, each partition is considered

and all the rules belonging to a partition P_i are moved downstream to the firewall internal FW_i. The rules that are flagged as “multi-domain”, instead, are kept on the FW_O. The number of rules on firewall FW_O is reduced and, for each sub-domain d_i, a specific new "by-pass" rule is added at the beginning of FW_O configuration. Each "by-pass" rule r_b is written to accept all the packets with the sub-domain d_b as destination, since these packets will be checked by the firewall FW_b, which has now all the rules concerning specifically the subdomain d_b. This “by-pass” approach further reduces the load on FW_O as the related packets are quickly matched by one of the first rules, thus avoiding the need to further evaluate the other rules. It is worth noting that this approach has also the effect of allowing unwanted traffic to flow into the green area of Figure 1. In fact, packets that can be blocked by FW_O are now accepted and then blocked by one of the internal firewalls, such as FW_1. Moreover, while the main objective of this approach is to improve the performance of the filtering system, of course, we have also to ensure that the overall security policy is maintained and that any changes in the configuration do not affect the security properties of the system.

A first validation of the rule distribution algorithm can be found in [99]. The underlying firewall model is a simple but general one, where a firewall configuration is represented by a list of filtering rules, where each packet can only be accepted or rejected, and where only a "goto" instruction can modify the flow of rule evaluation in the firewall configuration. A first experimental evaluation is provided as well, where the effectiveness of the proposed algorithm is measured in a simulated environment. After these first results, two further steps are required to fully validate the approach. On the one hand, we need to consider more complex firewall models, adapting the algorithm to the specific features of these further models. On the other hand, we need to confirm the results found in a simulated environment in a real one, measuring the performances of real deployed firewalls in physical topology. About the first aspect, a good starting point is the iptables software firewall, widely used in Linux-based devices and environments. The structure of iptables configuration is based on groups of rules organized in so-called tables and chains that are traversed in a specific order, that can, up to some point, be modified through "goto" and "jump" operations. A preliminary model and modified distribution algorithm have been developed for this type of firewall, however, some specific features of this software firewall (such as the network address translation, NAT) require further consideration and analysis. Nonetheless, the performances of iptables firewall and the rule migration algorithm have been validated with a physical testbed [100] where two physical firewalls have been deployed in a subset of the topology of Figure 2. The results of this experimental validation confirm the effectiveness of the approach.

3.4.1 Conclusion

The security of industrial networks and critical infrastructures is a priority. One of the utmost important aspects to consider is the security and the availability of the underlying communication network. Surges in traffic loads in these systems can quickly result in severe damage. One possible approach to reduce this risk is to adapt the system to the changed situation by re-distributing the filtering rules throughout the available firewalls. This approach has been demonstrated to be feasible [99] and to effectively improve the network performance of the systems [100] while keeping the designed security policy unaffected.

4 6G technologies: key features, challenges, security issues, and potential solutions

By the end of 2025, more than half the world's population is estimated to have access to 5G networks [101,102]. Whereas these networks are still deployed across the world, research and development on 6G networks have already started, as 5G networks do not have sufficient bandwidth to accommodate new technologies such as holographic apps, and the increasing popularity of wearables and mobile video services strains bandwidth even further [103-107] It is expected that 6G which integrated with 5G with satellite networks for global coverage will be a radical departure from the traditional wireless mobile communication [101], allowing for the maximizing the synergy between AI and mobile networks. Furthermore, the introduction of 6G will mark the shift to a radio-optical system taking advantage of both electronic and photonic technologies. It has fast broadband signals through the air at high-speed optical fiber lines to transmit the secured information from transmitters to destinations. Also, it is expected that 6G with cheap and fast internet technology, zero distance connectivity between people, and incredible transmission speed in Terabit ranges will complete the wireless network through no limitation and maximizes the data throughput and IOPS (Input-Output Operations per Second) [102,106,108]. The wireless evolution from "connected things" to "connected intelligence" will be dramatically reshaped by the 6G system.

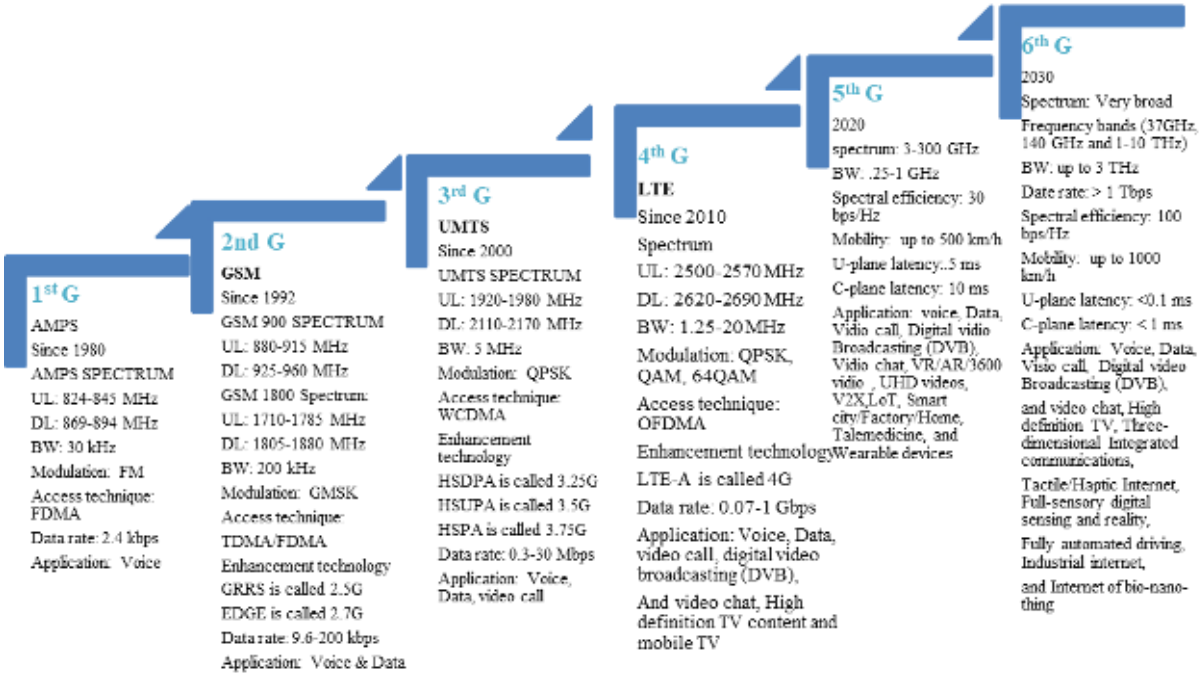


Figure 3 - different generations of communications

Since the beginning of 2020, there has been an increase in research papers concerning 6G. Major companies, like Facebook and Verizon, as well as political and academic institutions, like the Chinese Ministry of Science and Technology and the European Commission, are discussing ways in which the mobile networks of the future can be realized. Not only do new technologies such as Holographic-type Communication (HTC) and Tactile Internet, with their higher bandwidth demands, need to be accommodated by the new networks, which also applies to the ongoing rise of cloud computing [106], [102]. Ethical questions on topics such as privacy preservation need to be addressed as well, as the risk of personal data abuse and loss increases with the interconnectedness of the physical and digital worlds [109]. On top of that, the

development of 6G networks must aim for an intelligent, green, sustainable, and secure system [106]. In Figure 3, the different generation of communications is demonstrated. The paper provides a broad overview of 6G networks, and the goal is to discuss the advantages of 6G in mobile communication, as well as challenges and potential solutions that arise from the new networks and their key technologies [101], [110], and [111].

4.1 Key features and use-cases for future 6G networks

The first 6G network is expected to be deployed in 2030, possibly earlier. It will need to meet remarkably high demands in terms of latency [102], reliability, mobility [105], and security. The new 6G technology is characterized by a shift to a radio-optical system taking advantage of both electronic and photonic technologies, as well as maximized synergy between AI and mobile networks. Further characteristics that distinguish 6G from other network technologies are the ubiquitous 3D coverage of the surface of Earth, a smart compute-connect entity, as well as its higher intelligence and sustainability.

New KPIs measured in 6G network performance are reliability, signal bandwidth, positioning accuracy, coverage, timeliness, security, and privacy, as well as capital and operational expenditure. Some of the KPIs used to evaluate 5G networks will continue to be used, such as peak data rate, user-experienced data rate, and latency [103,106,112], PLS, network information security, and AI/ML-related security are all factors to consider for characterizing security [112].

In short, the Advantage of 6G are as follows:

- Protect your mobile and secure your data
- Provide a real online gaming experience and reduce the lag while gaming
- The extremely fast data transmission
- Super-fast streaming without buffering and high efficiency
- Record calls and forwards them to email also forward calls to other numbers
- Provide Intelligent batteries and Improved storage capacity
- High mobile-TV resolution
- Increase availability of social network
- Control natural disaster
- Satellite to satellite communication
- Create smart homes, cities, and village

Performance compared to 5G and possess high-performance matrices. For instance, the peak data rate of 6G is anticipated at 1-10 Tbps with optical frequency band and THz assistance, whereas this data rate in 5G is 20 Gbps.

It is also important to mention that conventional Cognitive Radio (CR) users' devices such as laptops, mobile phones, and taking health-related intelligent devices could be under strict restrictions [113]. 5G targets uRLLC, umMTC, and eMBB, whereas 6G CRNs will widely improve and also spread the application situations [114]. Some applications and use-case are as follows:

- Reality Extended by a combination of augmented reality (AR), virtual reality (VR), and mixed realities (MR).
- Smart Housing Societies to push life quality improvement, automation, and surrounding monitoring with the use of artificial intelligence-based machines.
- Remote Information transfer of five Human Senses (sight, touch, hearing, smell, and taste) to experience the world's surroundings.
- CR Brain-Computer Interaction (CR-BCI in innovative societies, pointedly medical systems and home used appliances.

- Haptic Communication, such as system applications and implementation, are predicted to maintain the higher features of 6G networks.
- Internet of Everything (IoE) stands for autonomous coordination and unified integration among many elements (computing elements); people, objects, devices or sensors, and the internet using data and processing.
- AI-based manufacturing and industrial intelligence automation [113].
- Digital Twin technology involves the creation of comprehensive and detailed virtual copies of physical objects.
- Intelligent healthcare systems by innovation in VR, XR, MR, AR, mobile edge computing, telepresence, holographic, and artificial intelligence [103], [113].

4.1.1 Key features for future 6G

The future 6G networks will be defined by key technologies such as AI, terahertz (THz), Blockchain, three-dimensional networking, wireless optical communication, and other potential technology, which we discuss in detail.

4.1.1.1 AI technology

AI and ML play a key role in 6G networks [106,115,116]. AI technologies can be utilized to enable 6G wireless systems to be in Autonomy [116,117]. Edge Intelligence (EI) is also partly based on AI technology and can be applied to use cases such as the automation of the management and orchestration tasks of the virtual resources in NG-RAN architecture [106,118]. It is crucial to offer AI technology to end-users in an AI-as-a-service paradigm [106]. AI provides intelligence for wireless networks by simulating some pointing at specific human processes and intelligent behaviors. In addition, autonomous applications such as autonomous aerial vehicles and autonomous robots are used in 6G [107,119]. AI Applications in 6G are capable of handling both the physical layer and network layers with unsupervised learning algorithms. These algorithms can be used in routing, traffic control, parameter prediction, resource allocations, etc. To put it in a nutshell, we have

- Supervised learning to train the machine model using labeled training data.
- Unsupervised learning to be leveraged to look for invisible patterns without using labels.
- Model-driven approach to be used in Artificial Neural Networks (ANN) with prior information based on professional knowledge.
- Deep reinforcement learning (DRL) for Markov decision models to select the next action based on the state transition models
- Explainable AI to build trust between humans and machines [107].
- Federated Learning (FL) to develop a machine learning model with training data remaining distributed to clients to protect data owners' privacy [107,116,120].

4.1.1.2 Terahertz communications

The THz spectrum is expected to provide extra high bandwidth [102,113,121]. Not only do they work with smaller antenna sizes, but they also are not limited by atmospheric effects either. Applications of Terahertz limit the use of THz for short-distance transmission such as indoor communications [107].

4.1.1.3 Blockchain-based networks

Blockchain technologies have rapidly grown both in the industry and academic institutions because of decentralized transparency, security, and privacy. Blockchain can provide a more flexible, secure, and efficient information infrastructure in 6G networks.

4.1.1.4 Satellite communication

By integrating with satellite communication (i.e., earth stations communicate with each other via satellites), 6G can provide localization services, broadcast, Internet connectivity, and weather information to cellular users [102, 106]. Therefore, enablers with new spectrum technologies will play a crucial role in supporting 6G key; such as mmWave [101], THz communications, and also softwarization and virtualization. Other enablers are new air interface technologies such as massive MIMO, IRS, and CoMP, as well as new architecture like 3D coverage using integrating large-scale satellite constellations and new technologies [106].

4.1.1.5 Index modulation

Index Modulation (IM) has high spectral- and power efficiency due to its idea of sending extra information through the indexed resource entities. IM used in 6G Networks are as follows:

- Time Division Duplex (TDD)
- The Orthogonal Frequency Division Multiplexing (OFDM), due to its high spectral efficiency
- SD-IM Technique or Spatial Modulation (SM), without inter-antenna synchronization and interference and with low complexity in the receiver. The spare information bits can trigger the transmit antennas.
- CD-IM Technique, by changing the property of the radio frequency (RF) environment due to employing RF mirrors or electronic switches [107].

4.1.1.6 Full-duplex and In-band Full-duplex (IBFD)

As a full-duplex and in-band full-duplex (IBFD) technologies improve the communication efficiency by considering Frequency-(FDD) and Time-Division-Duplex (TDD) in scheduling algorithms. Therefore, devices can transmit and receive a signal in the same frequency band.

4.1.1.7 Holographic radio

To develop holographic radio, one of the most promising interference-exploiting technologies, unwanted signals can be useful.

4.1.1.8 The NIB technique or Network in Box

The NIB technique or Network in box needs to be taken care of due to a device that can provide seamless connectivity between different services [107].

4.2 Challenges and potential solutions

The 6G, the next-generation advanced mobile communications system, is expected to integrate mobile and AI networks and work as a huge supercomputer. Different aspects such as distributed communication, computing, and storage will be combined in this system [106]. Since the networks will need to meet higher demands than their predecessors, particularly in terms of bandwidth, privacy preservation, and sustainability, questions on factors such as safety and energy efficiency need to be solved [122].

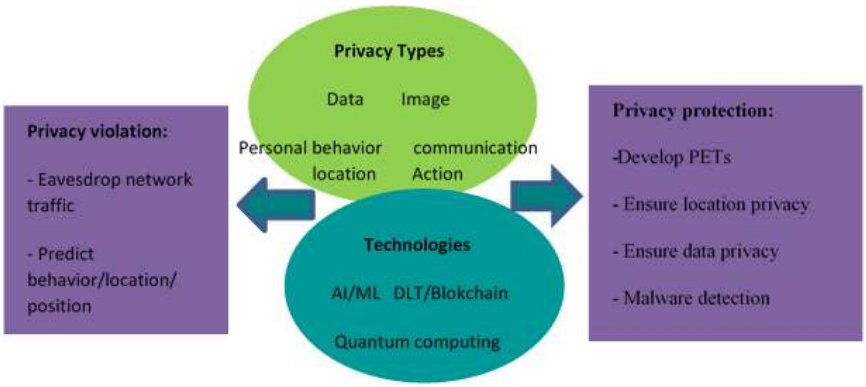
4.2.1 Challenge 1: security and safety

In the evolution of the mobile security landscape from 4G towards 6G, LTE, LTE advanced technology used in 4G makes MAC layer threats and attacks. Similarly, in 5G with NR, SDN, NFV, and NS technologies, we can see cyberware, critical infrastructure threats, SDN/NFV threats, and cloud computing-related threats. In 2030 at 6G networks with new technologies (i.e., AI/ML, blockchain, VLC, THz, quantum computing), there would be AI/Ml based intelligent attacks, zero-day attacks, quantum attacks, PHR layer attacks for VLC, THz, etc. The privacy types in 6G networks [112] (like Data, image, personal

behavior, communication, location, and actions) have to be protected, as described in Figure 4. We will discuss security in different dimensions; Platform and architecture, applications, and technologies.

Figure 4 - summary of 6G privacy

4.2.1.1 Security impact on 6G platforms and architectures



Due to the interlink of physical and digital worlds in 6G, it is necessary to address safety and security issues, especially in platforms and architectures. Since it is expected to be linked to mobile technology (stronger mobile than now), not only private security but also national security play a key role in insufficient trust models, clear practices, and rules. The increasingly widespread use of IoT technology, in particular, raises the demand for more network access points, network capacity, and service capabilities, which in turn creates a need for reliable trust models [102]. This need is amplified by the rise of cloud and edge computing. Also, a new routing solution must be developed because Border Gateway Protocol (BDP more efficient than currently used [109].

In order to achieve user trust, it is mandatory to aim for a holistic security approach. Classic cryptographic methods must be enhanced, and end-to-end security must be provided. Physical layer security is one method that can complement classic cryptographic methods [109], [123]. Another promising approach is using AI for end-to-end security. Other ways to achieve this include the combination and harmonization of technologies such as SDN and NFV [109], [115].

A new alternative for trust verification is Distributed Ledgers (DL). Current DL-related research challenges include improvement of DL privacy, trust management for the wide-area across multiple domains, and a specific DL for the 6G mobile network [109], [123]. Execution offloading is another promising method for enhancing privacy preservation. However, it might increase the risk of data abuse and data loss. In order to counter this, the following methods could be used: authentication of remote endpoints, certification of the platforms, support of remote attestation, and support of secure properties in spite of insecure execution platforms [109]. As far as ML is concerned, Federated Learning technology can be used in order to pre-process the data before distribution, thus omitting any information that cannot be distributed legally [106].

To summarize, the development of huge connections in 6G architecture and platform raises security and privacy issues. These issues are defined in different categories; orchestration, intelligence network management, edge intelligence, cloudification, specialized 6g networks, intelligence radio [112], RAN-core convergence, and end-users (terminals and users). Table 1 shows different categories and the corresponding security issues.

6G Architecture	Potential Security Issues
Orchestration	• Open API security threats • AI/Ml attacks • Security threats with closed-loop network automation
Intelligence network management	
Edge intelligence	• Data privacy threats • AI/ML attacks • Security threat on cloud • Security threat edge
Cloudification	
Specialized subnetworks	• Trust violations • AI/ML attacks • PHY security threats
Intelligence radio	• AI/ML attacks • PHY layer security threats
RAN core convergence	
Devices	• DDoS attacks on devices • User privacy threats
Consumer or End-user	

Table 1 - key security aspect of 6G platform and architecture

Physical Layer Security (PLS) techniques have a very important role in 6G architecture and are defined in four techniques. PLS techniques to improve confidentiality and conduct lightweight authentication and key exchange depend on the unique physical features of the random and noisy wireless channels,

- *Terahertz technology*: Data transmission exposure, eavesdropping, and access control assaults are all possible with THz communications. The potential solution for the challenge is to characterize the channel's backscatter to discover some eavesdroppers to improve information security. In order to investigate THz propagation multipath and path-loss, a device fingerprint in a THz time-domain spectroscopy setup may be developed.
- *Visible Light Communication (VLC) technology*: Eavesdropping attacks are common on VLC systems. Therefore designing PLS mechanisms play a key role since VLC has fast data rates, a lot of available spectrum, and robustness against interference. Designing the linear precoding in terms of the reachable secrecy rate enriches the secrecy performance of a multiple-input multiple-output (MIMO) VLC system. The transmitted signal is subjected to a peak-power limit, and only discrete input signaling schemes are employed. To improve the secrecy of a VLC system, a jamming receiver with the spread spectrum watermarking scheme (a watermark-based blind PLS) was combined.
- *Reconfigurable Intelligent Surface (RIS)*: RIS is a software-controlled meta-surface that can dynamically regulate its reflective coefficients, allowing them to control the amplitude or phase shift of reflected signals and thus improve wireless propagation performance. The reflected signals can be added coherently at the intended receiver to improve the quality of the received signal or destructively at a non-desired receiver to improve security by intelligently managing the phase shifts of RIS [113], [124].
- *Molecular Communication (MC)*: Molecules in an aqueous environment or chemical signals are a communicating way in healthcare 6G-applications such as wearable body sensors and telemedicine.

Security and privacy issues in the communication, authentication and encryption processes in sensitive data have to be addressed at the beginning of its actual development by calculating the secrecy capacity, i.e., the number of secure symbols that may be sent via a diffusion-based channel [113].

4.2.1.2 Security impact on 6G applications

The identified important applications and use-cases in 6G (summarized in **Error! Reference source not found.**) [112], [215] have varying security requirements and implantation issues in three levels of requirement/impact: low (L), medium (M), and high (H) (L) (see **Error! Reference source not found.**, L in green color, M in yellow, H in red).

4.2.1.3 Security impact on 6G technologies

The new technologies used in 6G (such as distributed ledger technology (DLT), distributed and scalable AI/ML and quantum computing, and some PLS-related topics (THz, VLC, RIS, MC) mitigate security threats [112].

DLT and 6G are expected to operate together. Blockchain technology, as the most attractive among DLTs, can integrate various advantages in 6G technology with privacy and security [113], such as non-reputation, disintermediation, proof of provenance, immutability, pseudonymity, and integrity. Blockchain can handle trust elements (e.g., immutable records for AI data integrity and distributed trust across different stakeholders [112], [115] and privacy by combining the communication process, access control, and verification. However, Blockchain may have unintended security impacts on 6G networks, as described in **Error! Reference source not found.**. As potential solutions for the security issues are:

- Assuring the smart contract's accuracy
- Testing functionality before deploying a smart contract across thousands of Blockchain nodes
- using suitable access control and authentication techniques to identify malicious bots and AI-agent based Blockchain nodes
- incorporating additional privacy preservation mechanisms to mitigate privacy leakages (as privacy by design), and TEE can be into Blockchain-based 6G services
- selecting the appropriate Blockchain/DLT type for the 6G application and services to mitigate certain attacks' impact

Quantum computing is expected to be used in 6G communication networks for security vulnerability detection, mitigation, and prevention [101], [112], and [126]. Post-quantum safe cryptography, quantum-resistant networking hardware, quantum key distribution (QKD), and quantum-based attacks will be the advantage of quantum computing in 6G, especially in IoT networks and devices (as shown in Figure 5). Incorporating post-quantum crypto solutions that are resistant to quantum-based attacks into IoT devices is a hot challenge, in which IoT needs light-weight cryptographic solutions. The Oblivious Transfer (OT) is incapable of retaining quantum information since any leakage might jeopardize the entire two-party communication system. Quantum computers feature a no-cloning characteristic that makes maintaining an exact copy of a quantum state difficult [115].

6G Application	Requirements
Holographic Telepresence	• Reduce operational cost • Diversity of devices • High Bandwidth • High Privacy
Digital Twin	• Secure Communication • High scalability • IoT data security • AI security
Industry 5.0	• Reduced operational cost • Interoperability • Real-time Operation • High Scalability • IoT data security
Connected Autonomous Vehicle (CAV)	• Reduced operational cost • Diversity of devices • Interoperability • Real-time Operation • High Scalability
UAV based mobility	
Intelligent Healthcare	• Ethical AI security • Interoperability • Real-time Operation • High Privacy • Scalable IoT data security
Smart Grid 2	• Terrorist attacks • Scalable IoT security • Physical Tampering • Intermittent Connectivity
Extended Reality	• Reduced operational cost • Diversity of devices • Limited Resources • High Privacy

Table 2 - key security requirement of prominent 6G applications

Potential 6G Applications: Security Requirement (SR) and Implementation Challenges (IC)		UAV based mobility	Holographic Telepresence	Extended Reality	Connected Autonomous Vehicles	Smart Grid 2.0	Industry 5.0	Hyper-Intelligent Healthcare	Digital Twin
SR	Ultra-lightweight security	M	M	H	L	H	M	H	M
	Zero-touch security	H	L	M	H	M	H	M	H
	High privacy	L	H	H	M	M	L	H	L
	Proactive security	M	L	L	H	H	H	M	L
	Security via Edge	H	M	H	H	L	H	H	M
	Domain-specific security	L	L	L	H	H	H	H	M
IC	Limited resources	H	H	H	L	H	H	H	L
	Diversity of devices	M	M	M	M	L	H	H	M
	High Mobility	H	L	M	H	L	H	M	L
	Physical Tampering	M	M	H	M	H	M	M	L
	Terrorist Attacks	H	L	L	H	H	L	L	H
	Intermittent Connectivity	L	L	L	L	H	M	M	M
	Localized environment	L	M	L	L	L	L	H	L
	Lack of security standards	L	M	H	L	L	H	M	L
	E2E security orchestration	H	H	H	H	L	M	H	H
	Energy Efficiency	H	H	H	M	M	H	H	M

Table 3 - 6G applications: security requirement and possible challenges

Quantum-resistant technologies and encryption solutions have previously been investigated by researchers. Lattice computational issues perform better in IoT devices in the current environment. They fit better in 32-bit architecture due to the smaller key length. However, due to performance and memory restrictions, as well as communication capabilities, these categories are still being developed and are suggested for IoT devices.

Potential Security Issue	Description
Majority attack	A group of malicious users could capture the 51% or more nodes and take over the control of the Blockchain
Double spending attack	A user spends a single token multiple time
Re-entrancy attack	A smart contract invokes another iteratively, and the invocation of the secondary contract is malicious
Sybil attacks	An attacker attempts to take over the peer network by conceiving fake identities explicitly
Broken authentication and access control	potential vulnerabilities and issues in the implementation of authentication and access control mechanisms
Security misconfiguration	use of insecure security configurations or outdated configurations that make the system vulnerable to attack
Privacy leakages	vulnerable to leakage privacy of transaction data, smart contract logic, and user privacy
Other Vulnerabilities	Other security threats such as destroyable contracts, exception disorder, call stack vulnerability, bad randomness, underflow/overflow errors, and unbounded computational power-intensive operation.

Table 4 - key security issues of blockchain in 6G services

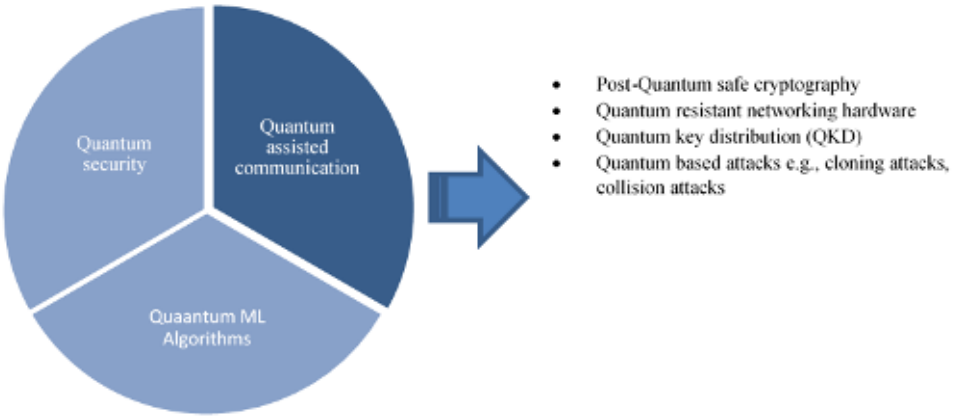


Figure 5 - role of quantum computing in 6G

The relation between security and AI is shown in Figure 6. The security and data privacy issues in 6G will be more challenging when the number of smart devices is increasing and tracking every move of a person with a lack of transparency about what is exactly collected. Therefore, Zero-touch Network & Service Management (ZSM), defined as autonomous networks performing Self-X activities without the need for human involvement in 6G networks., is very important. On the one hand, Differential Privacy (DP), which considers privacy in the context of statistical and machine learning analysis, is another emerging privacy-preserving technology that will likely feature in future 6G wireless applications [112], [115]. On the other hand, the factors such as the ML components' trustworthiness, visibility, AI Ethics and -Liability, scalability and feasibility, model- and data resilience are the AI/ML challenges in 6G networks. AI and machine learning will make the 6G intelligence network management system vulnerable to AI/ML-related attacks.

Compromise of AI frameworks to exploit flaws in such artifacts or traditional attack vectors against their software, firmware, and hardware parts is a serious concern at the AI middleware layer. API-based attacks are another form of attack in which an adversary requests and attacks an API of an ML model to get predictions on input feature vectors that have model inversion, model extraction, and membership inference attacks as a result of it. The potential solution for these challenges are as follows:

- Adversarial training and defensive distillation
- Input validation and control of the information (Provided by ML APIs to the algorithms against evasion attacks and adversarial attacks.)
- Protection of data integrity and authentication of the data origin
- Input validation and moving target defense against poisoning attacks.
- Control the information provided by ML APIs to the algorithms (i.e., add noise to ML prediction to prevent model inversion attacks)

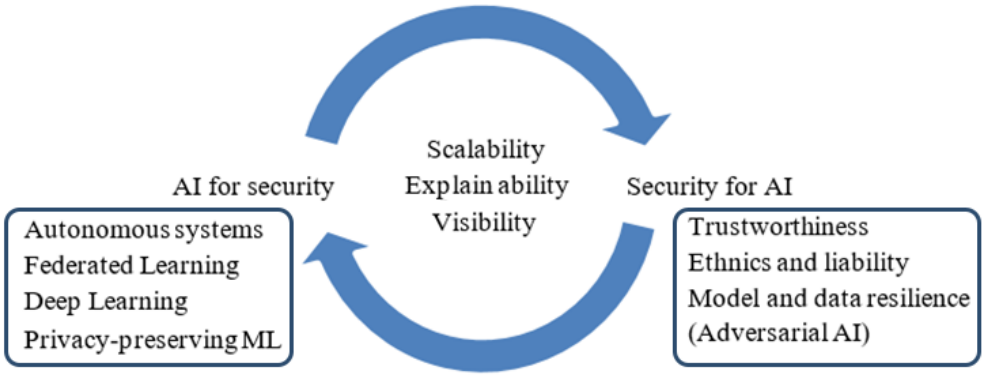


Figure 6 - 6G and AI in the security aspect

4.2.2 Challenge 2: standardization in security

The security sector has an active standardization as a vital part of next-generation networks. The standardization divers for 6G security are shown in Figure 7.

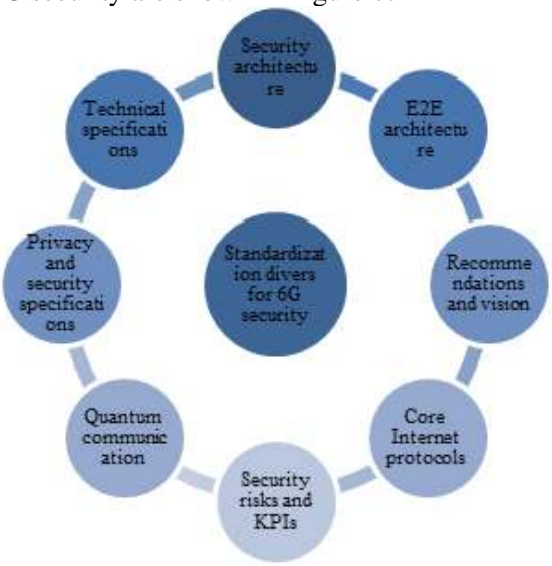


Figure 7 - 6G security standardization landscape

- ETSI: ETSI has created various Industry Specification Groups (ISG) to look into 5G component technologies, including NFV (ETSI NFV), AI, and network automation. Also, ETSI ISG ENI was established to develop a Cognitive Network Management architecture that would leverage AI methods and context-aware rules to adapt offered services
- ITU-T: The ITU-T Focus Group on Machine Learning for Future Networks (FG-ML5G) has been created by ITU to work on technical standards for machine learning for future networks, such as interfaces, and network topologies, protocols, algorithms, and data formats.
- 3GPP: By establishing the Network Data Analytics Function, the 3GPP has already addressed the application of AI/ML in the 5G Core Service Based Architecture (SBA). presently, 3GPP SA3 is working on a draft TR
- NIST: The National Institute of Standards and Technology is in charge of standardizing post-quantum cryptography methods.
- IETF: The IETF Security Automation and Continuous Monitoring (SACM) Architecture RFC specifies an architecture for a cooperative SACM ecosystem.
- 5G PPP: The 5G PPP Security Work Group was formed as a collaborative effort to address 5G security threats and issues, as well as provide insights on 5G security and how it should be addressed.
- NGMN: The NGMN 5G End-to-End Architecture Framework v4.3 (2020) lays out the requirements for end-to-end framework capabilities, including security, in terms of network entities and functions.
- IEEE: The IEEE P1915.1 Security Standard for Software Defined Networking, and Network Function Virtualization (SDN/NFV) aims to offer a framework for constructing and operating secure SDN/NFV environments [112].

4.2.3 Challenge 3: high bandwidth demands and accessibility

Since new technologies like holographic apps and the ongoing increase in connected things will place high demands on 6G networks, it is necessary to find ways to extend network coverage and increase efficiency.

This includes researching solutions to increase energy efficiencies, such as low energy utilization and elongated battery charge life duration. Cost efficiency is another vital factor, particularly in network deployment and expansion. This is especially crucial in covering more remote areas. On top of that, high intelligence applications can hurt user costs. Since new technologies like holographic apps and the ongoing increase in connected things will place high demands on 6G networks, it is necessary to find ways to extend network coverage and increase efficiency. This includes researching solutions to increase energy efficiencies, such as low energy utilization and elongated battery charge life duration. Cost efficiency is another vital factor, particularly in network deployment and expansion. This is especially crucial in covering more remote areas. On top of that, high intelligence applications can hurt user costs.

One way to improve the efficiency of wireless communication systems is the exploitation of the high-frequency spectrum. Bandwidth can be broadened with new spectrum technologies such as mmWave (millimeter wave) technologies [102], [106], [123]. However, mmWave can cause severe non-linear distortions and has a limited transmission range. Therefore, it must be used with caution and complemented by another technology. One possible addition to mmWave is THz, which has a broader transmission range but is prone to high path loss. Furthermore, network efficiency can be increased with the use of NG-RAN architecture. Whereas it is expected to support a massive amount of RAN slice subnets, further research is needed to create a virtualized and slicing-aware RAN for 6G mobile networks. Blockchain technology can also increase network efficiency.

As far as network coverage is concerned, non-terrestrial areas, in particular, 6G networks, have to be deployed as an ISTN (Integrated Space and Terrestrial Network). This network is predicted to carry with it the ground-based layer constructed by terrestrial base stations, the airborne layer empowered by HAP and UAV, and therefore the spaceborne layer implemented by satellites. So as to expand network reach, LEO satellites are used successfully. Additionally, the reusable rocket Falcon 9 developed by SpaceX helps decrease launching costs for LEO satellites, which may be helpful in increasing network coverage. Some alternatives to LEO satellites are HAP (high altitude aerial platform), which has lower maintenance costs and therefore the possibility of repairing, and UAV (unmanned aerial vehicles), which are more flexible than the satellites and offer the chance of re-planning the RAN dynamically [106].

4.3 Conclusion

Research on 6G technology has already started in the late 2010s. It is not only conducted by companies such as Huawei [127] and Samsung but also by governmental institutions such as China's Ministry of Industry and Information Technology and the EU [101]. Specialized research groups include Technologies for Network 2030 and the Next Generation Mobile Networks. Projects such as 5G-COMPLETE, 5G-CLARITY, and ARIADNE aim to explore the potential of the new beyond 5G networks. The key technologies, the advantages of the 6G technologies, THz, AI, green networks, use cases, ML, and VLC, make it widely accessible and increase its energy efficiency [106]. A holistic overview of security and privacy challenges and potential solutions is discussed. That way, the idea of 6G as a supercomputer-like network can be realized in the next years, which will not only benefit the industry but academia and private users as well. It is necessary to solve security and privacy issues, considering multidisciplinary challenges in fields such as politics, technology, and ethics. This leads to research challenges focusing on technologies such as trust networking, security architectures, PLS, and privacy solutions, including the development of frameworks to understand the trade-offs between privacy and trust in 6G systems. For these research challenges, legal factors, such as regulation and standardization of privacy and personal identity preservation, as well as workers' rights, must be taken into account.

5 Cybersecurity and society

When thinking about cybersecurity, our mind instinctively races to the technological world. However, the adoption of several cybersecurity-related technologies also has several profound repercussions on how we human beings interact with the world around us.

Information theft, ransomware, and social engineering constantly threaten industries and also our personal devices. Cybercrime has already started to invade our homes and our most private data years ago. This invasion has begun to shape how we think and how we can (legally) react to this new type of crime. Various countries have started to create ad-hoc laws to defend their citizens against cybercrime. Companies nowadays include security policies in their governance models and train their personnel accordingly. There is no doubt that the cybersecurity field is influencing our society deeper than from a mere technical perspective.

5.1 Building principles for lawful cyber lethal autonomous weapons

There is a broad consensus over the fact that the choice of means in international warfare is not unlimited. Over time, several agreements have limited the use of some weapons, from the Geneva Conventions to the United Nations' "Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons Which May Be Deemed to Be Excessively Injurious or to Have Indiscriminate Effects" [the Convention on Certain Conventional Weapons (CCW)]. Recently, a new generation of weapons is emerging: lethal autonomous weapons (LAWs) and applications of artificial intelligence to the military. Their lawful use is highly debated, but the discussion focuses only on their embodiment as "killer robots". We argue that the category should include cyberweapons: malware/exploits used by state actors for military or intelligence aims, generally toward other state actors, in what is known as cyberwarfare [128]. While cyberweapons ultimately target devices and software applications, they can also indirectly hurt or kill a person; for instance, a compromised industrial control system can lead to somebody's death. We would like to open the debate: What security design principles make a lawful cyber lethal autonomous weapon or CLAW?

5.1.1 Conventional and autonomous weapons

Conventional weapons are mostly regulated through the CCW, which includes five different protocols banning or limiting their use in international warfare. To do so, International Humanitarian Law (IHL) builds upon four fundamental principles, limiting the adverse effects of armed conflicts [129]: humanity, distinction, proportionality, and military necessity. One of the main consequences of these principles is that weapons that may disproportionately affect civilian targets in comparison with military targets or cause unnecessary suffering are supposed to be banned. Throughout the CCW, the reasons behind the ban of a weapon can be classified into three main categories: technical (the specific characteristics of a weapon violate IHL), target (the way a weapon affects its targets violates IHL), and damage (the type of damage provoked by the weapon violates IHL). Conventional weapons regulations, however, do not consider the possibility for the weapon to be autonomous.

According to the International Committee of the Red Cross (ICRC), a LAW should be able to perform the following actions:

- target selection, which includes research, detection, identification, tracking, and selection;
- target attack, which can include the use of force, neutralization, damage, or destruction.

The heart of the issue is the definition of autonomy: international law places responsibility on humans, but increasing levels of autonomy complicate the identification of the legally responsible human. So far, three hypothetical and apparently mutually exclusive scenarios for LAW's autonomy have been identified:

- Human-in-the-loop: the LAW performs activities based on human input or authorization.

- Human-on-the-loop: the LAW performs activities under the supervision of a human operator, who may override the system in case of necessity.
- Human-out-of-the-loop: the LAW performs activities independently, without the need for human input or oversight [130].

The human-in-the-loop scenario seems excluded by the very definition of LAWs: continuous human input would make it non-autonomous and, by definition, some other kind of weapon. The human-on-the-loop scenario is considered to be the most feasible. However, this model risks undermining a LAW's rapidity of action and reaction (its very military advantages) or being meaningless by reducing the human's possibility to override the system due to so-called automation bias [131], a phenomenon that is well known to operators of security operations centers [132, 133]. The human-out-of-the-loop scenario seems out of scope, given the current embodiment of LAWs as "killer robots", which are limited by the physical supply chain of energy or transportation. This, however, is not the case for a CLAW.

5.1.2 Cyber lethal autonomous weapons

Concerning the autonomy of CLAWs, we argue that all three scenarios – human-in-the-loop, human-out-of-the-loop, and human-on-the-loop (in that order) –will take place during the CLAW's lifecycle. Stuxnet's lifecycle is an illustrative case in point [134].

Some moments in a CLAW's working process require an input coming directly from the human operator: the decision to start the system, for example, or to retire it at the end or in case of malfunctioning. These actions fall under the human-in-the-loop scenario. Then, a considerable number of actions will likely be performed by CLAWs with no human intervention at all and would fall under the human-out-of-the-loop scenario. For example, Stuxnet's propagation happened without human supervision. This is typical of cyber. Finally, some particularly crucial phases (for example, launching or recalling an attack) could be programmed to require at least a minimal level of human supervision and could therefore be traced back to the human-on-the-loop scenario. The retention of some level of human control in the most critical phases is an essential criterion for the lawful use of LAWs in conflict; in the case of programs with essentially deterministic effects, the very fact that they are programmed might provide this notion of on-the-loop scenarios. From this perspective, Stuxnet, which has been programmed to act on some specific nuclear turbines and not just any turbines, would map to the human-on-the-loop scenario.

When discussing lethality, the key point is that it is just a kinetic, incendiary, or explosive consequence of a conventional weapon system. Modern weapons do not touch the ship they sink: the explosion of the torpedo creates a void space in the water under the hull, and it is the suction filling the void that ultimately cracks the hull. The same reasoning applies to the actions of a cyberattack on the targeted cyber-physical system. A simple example is a malware whose primary effect is overheating of a lithium-ion battery – used in Internet of Things systems and smartphones but also in military vehicles – which can lead to lethal consequences for those placed in the vicinity of the concerned systems and to cascading explosions [135].

5.1.3 Building security principles for lawful CLAWs

Of course, traditional building security principles do apply [136, 137], but what we want to do here is to propose building security principles for lawful CLAWs. This is way trickier than one thinks to avoid falling into the realm of the irrelevant.

Consider the principle of distinction: it establishes the inviolability of certain non-military organizations, such as the ICRC, but also of medical buildings and personnel. These organizations cannot be targeted by attacks and must carry a distinctive emblem (such as the red cross or crescent) that cannot be used by military

forces improperly. Physical armies of a country are also supposed to be always identifiable by wearing a uniform or a distinctive sign.

Property	Condition	Description	Purpose
Technical	Absence of disrupting functionless fragments	The CLAW should not disseminate fragments of code that do not provide functionalities to the weapon itself but can disrupt the execution of the target system if invoked by chance.	We are back to "unnecessary suffering". The target systems can be fully DoSsed or taken over, but the malicious code should be purposeful in the same way that ROP gadgets are.
Technical	Permanent self-identification	Each CLAW should have a fingerprint or signature recognizable by its own designer, preserved through obfuscation or mutations.	While detectability is not reasonable as a criterion for a CLAW, the designer should be able to recognize its own to remove it after the conflict.
Technical	Eventual self-deactivation	The CLAW should be capable of deactivating itself (e.g., after a timeout or by inserting a key) or through a command-and-control system.	Indefinitely operational presence and damage might be needed throughout the conflict, but the impossibility of stopping it would be against the principle of proportionality and unnecessary suffering.
Target	Deterministic target (or nontarget boundaries)	Deterministic target or nontarget boundaries for the actual deployment of a lethal payload should be controlled through algorithmic fingerprinting.	The ability to perform stealth but not disruptive propagation across cyberspace might be justified by military necessity, whereas indiscriminate payload unleashing would be against the principle of proportionality.
Target	Initial validated specification for learning	Learning algorithms should start from an initial target definition that has been validated before deployment.	Trial and error for target identification would be against the principle of proportionality and unnecessary suffering, so the initial definition of a target should be done offline.
Damage	Appropriate software stack position	The programmed type of damage should be at the appropriate point of the software/hardware stack to achieve the CLAW's aim.	This would make it possible to avoid collateral damage to components (and related cyber-physical systems) that is not a consequence of the failure of the attacked component.

Table 5 - building principles for lawful CLAWs

Unfortunately, having software presenting itself is a technical no-go. Cyberattacks work precisely by confusing the target program into thinking it is interacting with a legitimate client program rather than being attacked [138]. However, the idea of recognizing and fingerprinting the targets once control has been taken

over is actually possible and implemented by several malware authors: DarkSide, the malware behind the Colonial Pipeline hack, has a hard-coded do-not-install list of countries.

To decide what should be allowed and what should be prohibited by international treaties, we should start with the key functionality of a cyberweapon: eventually provide stealth control of an IT system for integrity attacks, or disable it for denial-of-service (DoS) attacks.

We tried to sketch positive principles in Table 1 based on the criteria that determined a weapon's ban within the framework of the CCW and according to the general principles of IHL. For example, launching a disk encryption attack on military systems as done by ransomware might be a lawful CLAW in a conflict, provided the attacker holds the key to unlock the system. Ransomware, where nobody has a key, would instead be unlawful.

There is one aspect we haven't tackled so far: the identification of the perpetrators of (unlawful) CLAW attacks [139]. Attributing cyberattacks to state actors even in peaceful times is hard. The absence of on-the-ground human personnel recognizable by identifying signs, the possibility of disguising one's location in cyberspace, and the presence of human-out-of-the loop propagation phases are fundamental complicating factors.

5.1.4 Conclusion

Cyberspace is not among the fields of warfare traditionally considered by international law, and it is unclear which codified, or customary norms would apply to cyberwarfare. There is uncertainty over the application of the principles determining the legality of conventional weapons to cyber (autonomous) weapons. With this white paper, we would like to kick-start a discussion before such weapons start being effectively used in conflict scenarios – without knowing if they are lawful or not.

5.2 Governance foundations for the European cybersecurity community

With Regulation EU 2021 / 887 of 20 May 2021 establishing the European Cybersecurity Industrial, Technology and Research Competence Centre and the Network of National Coordination Centres, the European legislator intends to end the fragmentation of efforts in the research and development of cybersecurity products in the EU.

However, in comparison to the rules regarding the establishment, structure, and tasks of the European Competence Centre and the national coordination centers, the establishment, governance structure, and tasks of what the regulation termed the "Community" are rather vaguely described in the Regulation.

It is against this background that the European Commission decided to fund four pilot projects to help build and strengthen cybersecurity capacities across the EU, as well as provide valuable input for the set-up of the Cybersecurity Competence Network with a European Cybersecurity Research and Competence Centre.

The pilot CyberSec4Europe is designing, testing, and demonstrating potential governance structures for a future European Cybersecurity Competence Network using best practice examples derived from well-proven concepts like CERN, as well as the expertise and experience of partners.

With a focus on community-building, the CyberSec4Europe pilot project proposes the installation of additional regional and cross-border networks at the Community level. As one element to achieve this goal, CyberSec4Europe envisions the introduction of Community Hubs of Expertise in Cybersecurity Knowledge (CHECKs) into a future form of the regulation. The network would be significantly strengthened and advanced into a true structure that would ensure efficient flows of information that are implemented swiftly and occur within the most efficient layers.

Accordingly, this paper contributes to the foundations of a framework facilitating the emergence of bottom-up communities of knowledgeable cybersecurity experts that would also integrate potential users and their needs, including from the civil society.

5.2.1 Legislative framework

With Regulation EU 2021 / 887 of 20 May 2021 establishing the European Cybersecurity Industrial, Technology and Research Competence Centre and the Network of National Coordination Centres, the European legislator intends to end the fragmentation of efforts in the research and development of cybersecurity products in the EU. The Regulation intends to transfer existing expertise, particularly in cybersecurity research, into practice in the form of products that are tailored to needs and marketable.

The European Cybersecurity Network appointed to implement these plans is intended to empower the relevant stakeholders, especially those working in the field of cybersecurity, to network the wealth of expertise and experience. This is intended to strengthen and further develop Europe's cybersecurity posture. At the same time, the network aims to improve the link between the players in the field and decision-merks to enhance and support their political decisions.

For the institutionalized implementation of these goals in the network, the Regulation provides for three categories of actors. The central role is envisaged for the European Cybersecurity Competence Centre, which will be responsible for coordinating cybersecurity-related activities at the European and national level, as well as for planning and implementing research and development tasks, and for advising and supporting the relevant actors in the EU.

The second category of actors on the level of the member states is the national coordination centers, whose task is to support and promote the competence center and the members of the cybersecurity community as a third group of actors. The national coordination centers act as focal points and provide assistance in the form of funding, expertise, and technical support.

The mission of the aforementioned cybersecurity competence community is to support the competence center, the network, and their projects, as well as to develop and disseminate expertise in the field of IT security. The circle of members is deliberately kept open and is intended to bring together the most important stakeholders. These are, on the one hand, institutions from industry, science, and research, civil society organizations, but also standardization organizations, public or private institutions involved in cybersecurity, as well as stakeholders from other areas that have an interest in cybersecurity or face related challenges. Community affiliation is regulated by the Competence Centre and the national coordination centers.

5.2.1.1 The Cybersecurity competence community

However, in comparison to the rules regarding the establishment, structure, and tasks of the European Competence Centre and the national coordination centers, the establishment, governance structure, and tasks of the Community are rather vaguely described in the Regulation. The broad definition of the Community in Art. 8 (2) of the Regulation involves a heterogeneous group of very different private and public actors. The inclusive approach of this broad definition is to be welcomed. However, it remains unclear how this Community will be established. The wording of the Regulation ("…shall consist of…") does not seem to assume an already existing community. This would also be contradictory given the fragmentation of industrial and research efforts, lacking alignment, and common mission identified by the European Commission, which was part of the reasons why the Regulation was adopted by the European legislator.

5.2.1.2 Possible role of the community

As regards the role and tasks of the Community, the Regulation is very concise, too. In summary, the Community’s role is a supportive one, and the Competence Centre and Network shall collaborate with the

Community only “as appropriate” (Art. 3 (2)). Besides contributing to the mission of the Competence Centre and the Network and enhancing, sharing, and disseminating cybersecurity expertise across the Union (Art. 8 (1)), the Community shall work closely with the Competence Centre and the national coordination centers, participate in activities and working groups and support the promotion of specific projects (Art. 9). The Regulation follows a top-down approach for the involvement of the Community in the working groups (Art. 13 (3)(n)) and the Strategic Advisory Group (Art. 13 (3)(o)) and does overall not stipulate an active role of the Community or its involvement in any decision- making processes.

5.2.1.3 Opportunities and barriers for community building

In order to activate and effectively use the wealth of expertise and experience in cybersecurity research, technology, and industrial development that exists in Europe, it is crucial to establish a true and agile Community as soon as possible and give it a governance structure that enables the Community’s full potential. This also includes means of cooperation, collaboration, knowledge exchange, and financial opportunities. Since the top-down approach chosen by the Regulation does not provide for such means and lacks benefits or incentives for joining and contributing, this could make the Community building otherwise difficult or even impossible.

Furthermore, despite the broad definition of the Community, the participation design in the Regulation is rather exclusive. The organizational efforts that come with the involvement of community members in working groups or in the Strategic Advisory Group are likely to exceed the organizational possibilities of smaller members, no matter how much expertise and experience they have.

5.2.2 Organizing the community

The fact that the Regulation says very little about the Community can be seen as an opportunity. The absence of strict rules leaves the opportunity for the development of a true bottom-up approach to the Community building. The power of bottom-up approaches results from their possibility to provide broad expertise and knowledge of industry, academia, and stakeholders in specific areas by organizing information gathering and distribution. It is, thus, an appropriate way of activating research and development capacities.

One way to establish and organize the Community could therefore be the introduction of hubs in which different stakeholders could join their efforts, accumulate special expertise, promote scientific exchange and facilitate research or development of solutions. These Community Hubs of Expertise in Cybersecurity Knowledge (CHECKs) would be low-level, easy-to-access points of accumulation of regional, sectoral or topical interests and information, and they can serve as an accelerator to demands and problem identification as well as solution mechanisms.

As a first step, it is necessary to further develop the idea of CHECKs. Their optimal design requires answering a variety of questions, such as the organization, composition, tasks, and funding of CHECKs, as well as their relationship to each other, the national coordination centers, and the European Competence Centre. The answers should be based on practical experience gained in the pilot projects, analysis of legal frameworks, expediency, and teleological considerations, and stakeholder feedback.

Therefore, the CyberSec4Europe pilot project proposes the installation of additional regional and cross-border networks at the Community level. As one element to achieve this goal, CyberSec4Europe envisions the introduction of CHECKs into a future form of the regulation. The network would be significantly strengthened and advanced into a true structure that would ensure efficient flows of information that are implemented swiftly and occur within the most efficient layers. It has to be noted that also, from the point of the Community membership, it can be advisable to establish different decision-making processes, which will not always include all partners on all issues.

5.2.2.1 Our use-case

Two types of CHECK emerged after the preliminary analyses, namely one that is an economic actor in the cybersecurity landscape and must be sustained by a sound business model, and another that is part of the public administration and financed as a public good. The case described here, namely the CHECK-T pilot, in Toulouse, France, that is used to validate a specific governance model, is an example of the former type.

In view of the implementation of CHECK-T, interviews were conducted with stakeholders in order to learn about their needs and requirements regarding CHECKs, e.g., which details make the concept of CHECKs attractive for them to participate and contribute to the cybersecurity Community. These results together with possible changes in the governance structures may constitute the basis for the improvement of the European cybersecurity governance in future revisions of the regulation.

The interview campaign was carried on in order to identify the main needs and expectations, types of financially sustainable activities and a multidisciplinary pool of actors that would be willing to participate in the creation and development of the CHECK-T, aiming at:

- Mobilise communities of actors with different but complementary challenges
- Project a common vision
- Identify a consensus on the expected missions within the consortium
- Highlight the benefits for each stakeholder by sharing, contributing, and financing in common

5.2.2.2 Needs and expectations

The interview campaign included a total of 40 stakeholders from four large community groups (cybersecurity end users, cybersecurity solutions providers, technology centres, and economic development accelerators) and six mission classes:

- Expertise Development: Guarantee the sharing of data, sensitive information, and technological research with other partners on all types of incidents and on the responses provided.
- Technological leadership: Sharing expertise and general know-how, infrastructure, and investment costs by obtaining R&D funding.
- Transfer of uses, pooling of R & D & I costs: implementing methodological processes transferable from one sector to another, at lower costs.
- Design driven by need: Eliminate barriers by studying use cases and demonstrating scientific and technological know-how before large-scale deployment towards industrial products.
- Confidence-building: Building a local and European base of trust, promoting cooperation and coopetition between members (ethical framework, protection of freedoms, dissemination of trust).
- ROI of companies: facilitate the obtaining of funding in Cybersecurity Innovation and accelerate the maturation of projects and products to the market and awareness- raising, co-innovation activities.

Such analysis is summarised in

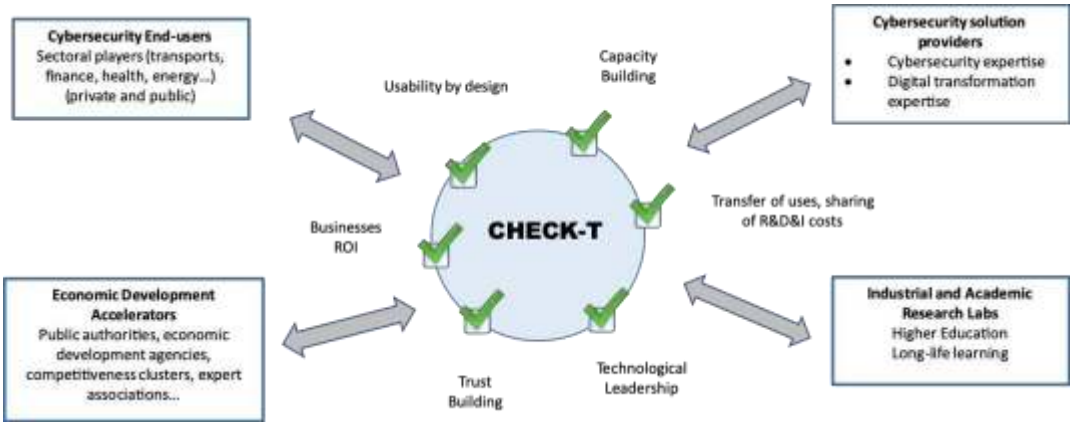


Figure 8 - community groups and mission classes for a CHECK

Figure 8.

Figure 8 - community groups and mission classes for a CHECK

5.2.2.3 Findings and recommendations

The possible interactions between the actors (from one community group to another and peer-to-peer by highlighting the concept of coopetition) were synthesized in Figure 9.

Four strategic application areas, which must be implemented in order for the stakeholders to take an interest in the creation of a CHECK, are described in Figure 10, along with the priority and evolving activities that have been identified.

Finally, the activities that needed to be implemented in order to increase the likelihood of success of a CHECK were chosen, as shown in Figure 11.

5.2.3 Conclusion

In view of the installation of additional European regional and cross-border networks at the Community level, CyberSec4Europe envisions the introduction of CHECKs into a future form of the regulation. For stakeholders to take an interest in CHECKs, four strategic application areas must be implemented, namely access to funding in R&D&I, capacity building, market access, and dedicated services. On the other hand, Member States should establish contractual Public-Private partnerships with the CHECKs in order to increase their attractivity in the eyes of stakeholders because the return on investment of joining one such CHECK will become more evident.

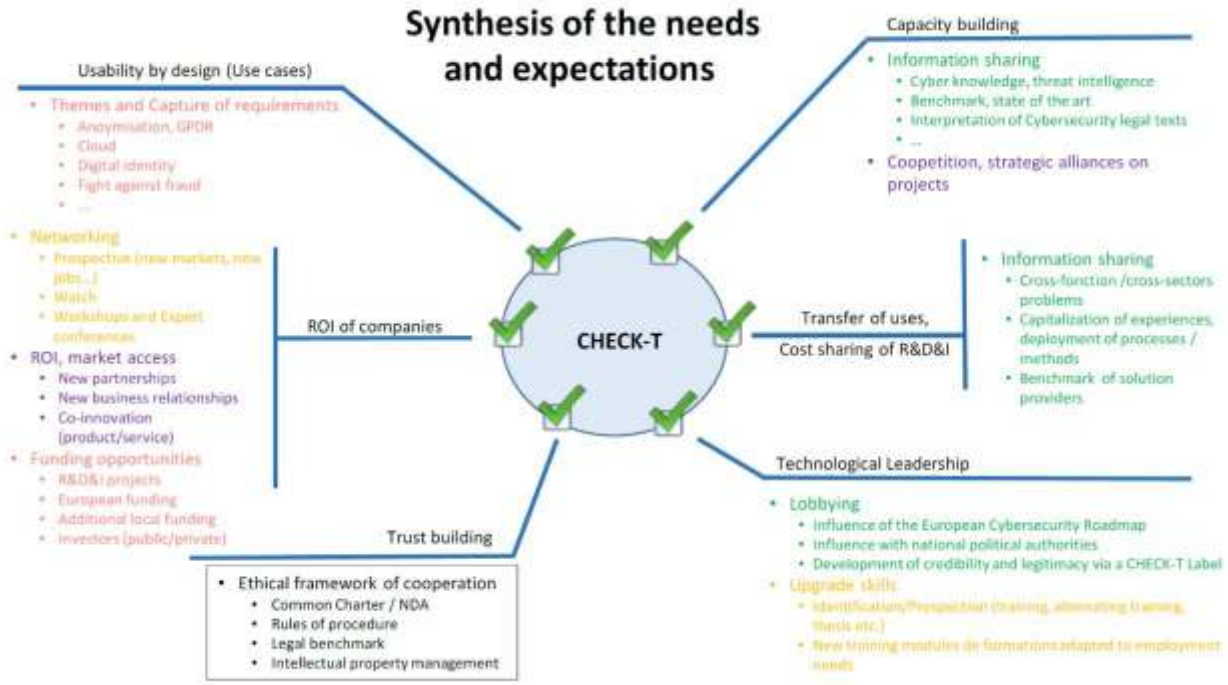


Figure 9 - synthesis of the needs and expectations

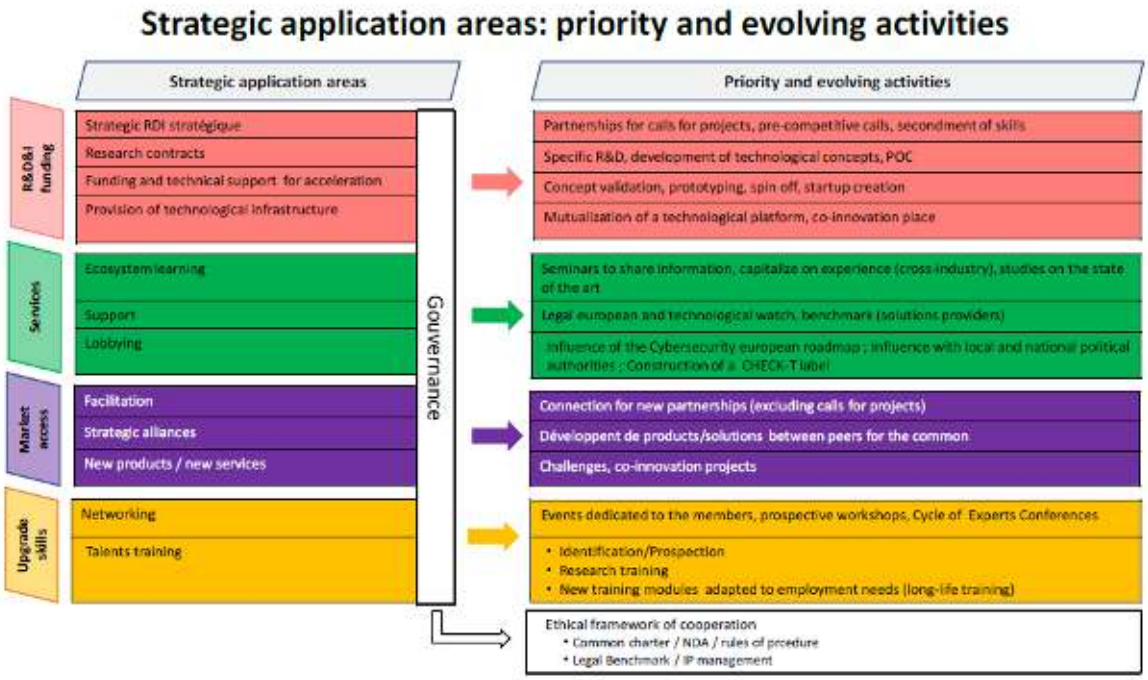


Figure 10 - the four strategic application areas emerging from the interview campaign

Figure 11 - list of activities selected to establish the CHECK in Toulouse

5.3 Cybersecurity awareness

Cybersecurity awareness (CSA) is about making people mindful of security risks and threats relevant to them, aiming to motivate them to adopt good security practices in their personal and professional lives. For a cybersecurity awareness program to achieve this objective, the communicated information, the medium of delivery, the timing, and the periodicity of it must be adjusted to the target audience [141]. In terms of content, the delivered information must be of sufficient depth to alert the audience of relevant security risks, enhancing their understanding of both the potential ramifications and the suitable adaptations of their security posture.

In recent times, almost every discipline is exploring artificial intelligence (AI) and machine learning (ML), as the means for increased efficiency and higher levels of automation. Cybersecurity is no exception and widely utilizes AI and ML methods to achieve various objectives, for example, security incident patterns extraction [142], intrusion detection [143], malware classification and detection [144], spam detection [145], and so on. Alarmingly, the usage of AI and ML methods has not been limited to constructive purposes, but cybercriminals and other malicious actors are also misusing and abusing them for ill gain. Deepfake, AI-supported password guessing, human impersonation on social networking platforms, AI-supported hacking, and so on are a few examples of the misuse and abuse of AI and ML methods by cybercriminals [145].

In the context of CSA, AI and ML are also finding potential uses, primarily through accelerating the transition from traditional computer-assisted CSA training to computer-based CSA training. In this computer-based CSA training, AI and ML methods can be applied to automate its various crucial activities so that a more personalized, customized, and optimized CSA training could be offered to mass easily and effectively. Some potential CSA activities in which AI and ML methods can be applied are mentioned next.

5.3.1 Identification of vulnerable audience group

Cyber risks change over time within an organization. Therefore, timely identification of the vulnerable employee group for the given cyber threats is crucial for effective CSA training. This information is important for CSA training planning and for determining security topics and focus groups that need to be prioritized. Analyzing relevant data such as the history of compromises, employee knowledge and expertise, and role-specific risks by using ML algorithms can help with tailoring the message and medium of delivery

to selected groups. Similarly, such analysis can assist with determining the probability of employees being targeted but also with inferring the probability of such an attack succeeding.

5.3.2 Personalization and customization of CSA training

According to current best practices, the content and procedures used for CSA training must be tailored to the needs and preferences of the audience. This has been proven to improve the impact of the training and extend the period of adherence to a more robust security posture. However, with limited resources available for CSA purposes, often this recommendation is not adhered to.

Every audience is diverse. In the same organization, employees can be different in their work nature, age group, education, work experience, national culture, and so on. Moreover, the employees can have different levels of security expertise and attitude towards security. These variances also determine the employee needs and expectations from a CSA program. For example, some employees may prefer short video training, while games with relatable characters may appeal to others, and some may thrive on task-based learning and simulations. In meeting the individual needs and expectations of employees from a CSA program, ML can play a vital role.

With the use of ML algorithms, it is possible to comprehend, and to some degree predict, the audience's requirements and expectations from CSA training. Furthermore, these algorithms can be applied to estimate what each audience knows, what they need to learn, and how they learn best or prefer to learn. This information can be utilized to design a more personalized and customized CSA content and training approach that best fits the audience's roles, responsibilities, and challenges. The training contents will be more relevant (up-to-date information on the security issues suitable to the audience), relatable (engaging characters and real-world scenarios which the audience can relate to), and fitting to the audience's expertise and knowledge level. Similarly, the delivery approach will be adjusted according to the learning curves (attention spans and retention) of the audience, and how s/he uses the training materials (time availability for the training).

5.3.3 Projection of future threats

It is preferable to have CSA initiatives that prepare the audience for both the existing and the evolving cyber risks and threats. But cyber risks and threats are dynamic in nature, and keeping pace with them is naturally difficult for any organization. Moreover, raising awareness of every evolving cyber risk and threat through the traditional approach can be resource taxing.

In such a situation, ML techniques can be applied to analyze the past cyber threats to find patterns in their evolvement and then apply the knowledge to project how the threats will change over time and generate potential future threats. These machine-generated threats can be used to teach the audience about the future course of attacks and how to stay protected from them.

5.3.4 Conclusion

AI and ML methods can become very effective means to automate various crucial activities of CSA training, for example, identification of vulnerable groups; projection of future threats; personalization, customization of training content and approach; and others. Automating these activities will help to get rid of the "one size fits all" approach rampant in CSA training [146] and rather provide a more optimized, customized, and personalized CSA training to the audience [141], [147]. In a nutshell, it can assist with delivering the right CSA content to the right audience at the right time in the right way.

5.4 Building European cybersecurity ecosystems: lessons from the past

The European Union has articulated its ambition to maintain its digital sovereignty and become a global leader in the digital economy, which is closely linked to challenges in the area of cybersecurity, such as the

lack of cooperation between the Member States, industries and academia, leading to fragmented efforts in research and development (R&D), insufficient investment in cybersecurity at EU level, increased demand for cybersecurity skills, know-how, and infrastructure, or (in)consistency in policies, legal frameworks, and actual practice. To meet these challenges, the European Commission proposed to set up the European Cybersecurity Competence Centre (ECCC), together with the Network of National Coordination Centres (NCCs) and Competence community (CC), which together can be considered a specific type of a multi-organizational structure or ecosystem aiming at strengthening of the capacities of the cybersecurity.

The current regulation[11] still leaves many uncertainties, and four pilot projects of the cybersecurity competence community (ECHO, SPARTA, CSEU, and CONCORDIA, see [27]) are continuously providing their feedback on these open issues. At this stage, for example, not many details about the governance of financial provisions are revealed about the ECCC, which will be located in Bucharest[12], or the other two layers (NCC, CC), so this paper is based on an analysis of lessons learned from the other similar value co-creation experiences, as well as the feedback collected in four pilot projects.

Cybersecurity4Europe project report [17], for example, proposes a governance structure, with an overview of the inputs on which it is based, while later deliverable [18] is bringing some experiences and experimentations with governance structures of cybersecurity ecosystems from various EU member states. In October 2021, ENISA, ECSO, and four pilot projects (CONCORDIA, SPARTA, CS4EU, ECHO) submitted their draft recommendations to the ECCC, following a consensus process about the future priorities. Besides these recommendations, representatives of these projects and institutions elaborated "concept on the way forward" with 11 strategic directions where the "competence community" could make a significant contribution. These strategic directions are not presented here, as they need to be further enhanced and maybe finetuned for specific target stakeholders. However, stimulating cost-efficient instruments for growth, transfer, or creation of value networks should also be considered in future revisions.

In order to have a cybersecurity ecosystem capable of addressing such a diversity of identified challenges, it is essential to take into account the stakeholder views and current best practices. Four pilot projects have already gathered views from more than 80 stakeholders via surveys, interviews, and workshops. In this paper, we additionally present some of the previous lessons learned and best practices in different kinds of cybersecurity initiatives, networked organizations, or "cybersecurity ecosystem", as it is called in the CONCORDIA project.

In this context, we define an ecosystem as "a system of people, practices, values, and technologies in a particular environment" [165]. The ecosystem includes roles, tasks, and relationships, which could be customized for different layers or even different member states [166]. Unlike the concept of a network, the ecosystem also brings also dynamicity since different alternatives need to be also considered from the economic perspective (e.g., reuse of software components, shared resources such as lab, scale-up of new solutions, etc.).

5.4.1 Related work

Network effects that increase the value are also studied in the context of internet platforms, which often enable a single company to take large market shares [149], with well-recognized internet services such as advertising, social networks, and search engines being the most prominent examples. As a result, so-called scaling benefits within "platform economy" or "multi-sided markets" arise. Depending on the type of

[11] https://eur-lex.europa.eu/eli/reg/2021/887/oj
[12] https://cybersecurity-centre.europa.eu/

application, the added value benefits users in different ways and depends on openness and centralization [150]. In the more open examples, the value should grow for everyone, although smaller entities have it more difficult. Although some elements could be applicable to the cybersecurity ecosystem and CC, for example, lessons from the multi-sided market for the e-identity services, these platforms are not directly related to the CC model.

Business ecosystems [147, 148] are structures where large companies can co-evolve their skills together with academic partners and smaller, more agile companies. Unlike CC, these are centered around one large company, although it also builds upon the idea of value creation by putting together different assets and skills. This process is often non-linear as in CC, but the configuration is not that complex, and stakeholders are mainly partnering from the supply side. These partners are expected to complement each other, while in CC, we will certainly have to deal with overlapping and competing supply-side stakeholders.

A similar notion but applied exclusively to digital technologies is termed digital business ecosystem (DBE), and it has been defined [153] as "a collaborative environment made up of different entities that co-create value through information and communication technologies (ICTs)". The concept of Digital Business Ecosystems was coined initially in the context of the implementation of the Europe 2002 action plan and the projects funded by the 6th Framework Programme of the European Commission, and clustered under the name "Technologies for Digital Ecosystems", presented their main research and empirical achievements in a book [156].

According to [154], DBE comprises two main tiers: digital and business, where this second tier refers to an economic community of stakeholders that operate outside their traditional industry boundaries. It does not, in comparison to the envisaged cybersecurity ecosystem or CC, consider academic research or regional economic development, for example. On the other hand, the similarity lies in the fact that DBE relies on the synergy between different stakeholders and values co-creation as an important driver.

The main characteristics of DBEs are [155] platform, symbiosis, co-evolution, and self-organization. Here we find further differences with ECCC/NCC/CC approach. In four pilot projects for the cybersecurity ecosystem and CC, the platform vaguely refers to a collection of tools, innovations, or services, but its focus differs from federated cyber-range platforms to research testbeds. Symbiosis and synergy between stakeholders in pilot projects are also common elements, but the instruments ad governance support is still lacking. Finally, co-evolution and self-organization are characteristics that might be desirable but impossible to evaluate and validate in four pilot projects. What seems inevitable is that different categories of relationships will co-exist in ECCC/NCC/CC, including organized and ad-hoc collaborations among stakeholders. Organized collaborations might include long-term strategic relationship networks, e.g., between ECCC and NCC, or inside of a regional hub, while informal and ad-hoc collaborations might be desirable on a cross-border level, for example, in R&D consortia.

The Digital Innovation Hubs (DIHs) is a kind of ecosystem that has existed with this name since 2016. Originally it was linked to the Digitising European Industry (DEI) initiative [161]. Afterward, the DIH concept was also evolving towards the European DIH and outside of the Industry sector and has also been suggested as a model for cybersecurity ecosystems. This also caused divergence from the original vision, and there was an attempt [162] to establish a shared common conceptual framework of a DIH within the European DIH community. This included five different building blocks as the backbone of the European DIH network, roughly described around competencies, services, economies and finance, and finally, collaborations and networks. One of the distinctive features of the "DIH-like ecosystem" is the focus on SMEs and Mid-caps that are faced with a "valley of death", a term that in innovation denotes the period between prototype and market-ready solution. Elevated economic risks and market failures are also common in cybersecurity, while "economies of scale", one of the main DIH assumptions, is also very important for the cybersecurity ecosystem.

Given that DIH also has a regional focus, synergy was established with a mechanism called "Smart Specialisation Platform". This European Commission initiative has existed since 2011 in order to facilitate mutual learning, networking opportunities, and other activities for regions. Thematic Smart Specialisation platforms have also been created, and while digital technologies are considered a transversal priority in most regional S3 across Europe, Digital Innovation Hubs (DIH) were started to be used as a policy instrument to boost these priorities [163]. Synthesis of empirical research and the analysis of the governance arrangements underpinning Smart Specialisation strategies were published in 2021 [164], but there is no specific analysis yet of DIH synergy and impact on economy and policy objectives.

Some experiences in what can be labeled as a "cybersecurity ecosystem" in Europe already exist, and those should also be taken into account when shaping governance and evolution of the forthcoming ECCC/NCC/CC ecosystem.

The EP3R (European Public-Private Partnership for Resilience) was established in 2009 and closed down in April 2013. In the ENISA report [167], it was described as "the very first attempt at the Pan-European level to use a Public-Private Partnership (PPP) to address cross-border Security and Resilience concerns in the Telecom Sector". This statement might be questionable since it was not a contractual PPP and also because different working groups with an information security focus already existed in several European initiatives ([168, 169, 170]), co-funded by the European Commission within the Sixth Framework Programme (2002-2006).

The EP3R had important support from ENISA that initiated, supported, and participated in many discussions. The PPP approach was judged to be particularly appropriate for addressing complex cooperation problems, and the model was even proposed for Information Sharing and Analysis Centres (ISACs) and similar initiatives.

Initially, the EP3R was facing challenges and value propositions such as team building, trust, joint objectives, action plan identification, and others, but stakeholders soon lost interest, and in a related survey, they mentioned a couple of shortcomings, such as the need for smaller working groups, focused and limited in time, need to improve motivation and incentives of demand-side stakeholders, simple but formal rules and governance, etc. In 2011 ENISA published a Good Practice Guide on Cooperative Models for Effective PPPs [171] and included some of these opinions in it.

Private-Public Partnerships (PPP) are a well-known instrument that has been used many times by the Commission that even published Guidelines for Successful Public-Private Partnerships[13] in 2003. The establishment of the network and information security (NIS) Public-Private Platform was announced in the Cybersecurity Strategy of the European Union in 2013 [157]. The NIS Platform was supposed to complement and underpin the proposed NIS Directive, while at the same time, its subsections (working groups) were addressing objectives such as input to the secure ICT Research & Innovation agenda at the national and EU level or assessment of Business Cases and Innovation Paths. The Commission called the first plenary meeting of the NIS platform on 17 June 2013, but this initiative, similar to EP3R, was also short-lived. It is also worth noticing that both EP3R and NIS platforms were "self-proclaimed" or designated PPP, as opposed to later "contractual" PPP (cPPP)[14] that emerged in the H2020 programme.

Nevertheless, the NIS Platform paved the way for the establishment of the first cPPP in cybersecurity, established in July 2016, where the EU was committed to investing €450 million in cPPP, whose private

[13] https://ec.europa.eu/regional_policy/sources/docgener/guides/ppp_en.pdf
[14] https://ec.europa.eu/programmes/horizon2020/en/contractual-public-private-partnerships

counterpart was represented by the European Cyber Security Organisation (ECSO). This cPPP was preceded by similar initiatives in the other digital technologies, such as photonics, robotics, high-performance computing, or 5G technologies in 2013 or big data in 2014. It should be noted that the "private" part includes many public stakeholders, such as academia and the public sector, and as such, it represents, in essence, also a kind of ecosystem. In the case of ECSO, many are confused by the differences between cPPP and ECSO proper, which explains on its web page[15] that within their governance, the Partnership Board is the actual communication channel between the European Commission and the ECSO. While at the time of writing this paper, the role of ECSO in the upcoming cybersecurity ecosystem is still unclear, there is no doubt that there are some useful lessons and elements that could be reused from ECSO and cPPP that were active during the H2020 framework programme.

Finally, we should also mention cybersecurity ecosystems with very focused objectives, such as data sharing. ENISA conducted a study on Cooperative Models for Public-Private Partnership (PPP) and Information Sharing and Analysis Centers (ISACs) [158], collating information on best practices and common approaches. As for the ecosystems focused on cybersecurity research, there are useful lessons in the EU 7th framework programme, where the "Network of Excellence" (NoE) was used as the funding scheme. It did so by supporting a "Joint programme of activities" implemented by various stakeholders, although mainly from the academic sector, with a possibility of longer-term cooperation. An example is the NESSOS project[16], which focused on secure software and services, or European Network of Excellence in Cryptology ECRYPT 2[17]. Both of these examples are limited in focus and type of stakeholders, and although they could be an interesting starting point for specific substructures, e.g., working groups within CC, reusability of their conclusions is limited.

5.4.2 Stakeholder analysis

Some stakeholders in the cybersecurity ecosystem, similar to what has already happened in four pilot projects of the cybersecurity competence community (ECHO, SPARTA, CSEU, and CONCORDIA, see [27]), will have both contributor and beneficiary roles, being present simultaneously on the supply and demand sides (for example telecommunication use case in CONCORDIA project[18]).

Stakeholder analysis in the cybersecurity ecosystem goes further than only looking at the supply and demand side or inclusion of "other" external stakeholders, such as policymakers, certification and standardization bodies, or legal organizations. Segmentation could and should take into account the current level of maturity, territorial approach, cultural differences, size of organizations, risk appetite, and many other parameters.

Collaboration and cooperation can be analyzed from several perspectives, from co-design (e.g., research project) of a solution to service co-delivery (e.g., coordinated response from security teams from different member states, see also work on Collaborative Automated Course of Action Operations (CACAO) Security Playbooks in CONCORDIA[19]). Less visible issues and challenges, such as SME networking, where supply-side SMEs could complement each other, should also be investigated from a wider economic angle.

Besides complementarities and cooperation, other value drivers for stakeholder collaboration and cooperation should be considered in the economic models for the cybersecurity ecosystem, including efficiency, avoiding vendor lock-in, or interplay between economics and digital sovereignty. Value network reconfiguration or government intervention through policy might be needed when addressing technology

[15] https://ecs-org.eu/cppp
[16] http://www.nessos-project.eu/
[17] https://cordis.europa.eu/project/id/216676
[18] https://www.concordia-h2020.eu/
[19] https://www.concordia-h2020.eu/blog-post/an-update-on-security-playbook-standardization/

acceptance or user adoption. The growth and evolution of the ecosystem (CONCORDIA consortium expands every year with new project partners) and the motivation for diverse stakeholders to collaborate and cooperate should also be analyzed.

The technology provider group, for example, could be expanded to other providers (external to the current participants of four pilot projects or the initial CC) that could replace one of the existing technologies or connect them to the different environments. Research groups (universities, institutes) might need to collaborate with members who provide consultancy services to enable the transfer of knowledge to the industry or the creation of new start-ups. Open-source and other related EU communities in digital technologies (e.g., GAIA-X, DAIRO, FIWARE, AIOTI, different DIHs, etc.) could have their role in the adoption and further development and the sustainability of the ecosystem. Standardization and certification bodies, individual investors, business angels, governmental organizations, incubators, accelerators, innovation centers, professional associations of cybersecurity practitioners, citizens, and others should all have clear roles and rules of engagement within the ecosystem.

This engagement may adopt various forms, depending on the context. Project or action types could include contracted work, transferring technology, smaller expert and consultancy services, permanent cooperation structures, industry-specific or sector-specific associations or sub-structures, spin-off and start-up companies supported by ecosystem incubators, etc. We might expect that the final ecosystem configuration will strongly depend on the choice of a governance model and the pool of funds and equipment available.

From the Pan-European perspective, solutions are not only supposed to be reused or used in a collaborative and cooperative manner but also evaluated by the “peers” from the same target audience, as well as by the stakeholders with different perspectives. Failure to synchronize activities across member states pertinent to the second tier of this ecosystem (NCC) would cause an excess of effort put into “reinventing the wheel”, too many overhead activities, additional challenges of benchmarking, interoperability, matching and reconfiguration, and the possibility to miss compatible value propositions and others.

5.4.3 Governance model

As mentioned before, the CC governance model is still under construction, but the four pilot projects, ECHO, SPARTA, CSEU, and CONCORDIA, have already made proposals in this regard, as well as regarding its strategic directions or set of principles. The ecosystem, for example, should monitor capability for scaling and capacity utilization, rapid identification of members that do not bring in value and can be excluded, reciprocal interdependence between members, members acting as mediators or “glue” for a value proposition, and it should be able to identify gaps and challenges in a collaborative and cooperative manner, taking into account multiple perspectives and multidisciplinary views. While ecosystem objectives are directly related to the long-term digital society and cybersecurity goals, specific stakeholder configurations in sub-structures, e.g. sector or technology dependant, should be feasible.

While all four projects addressed possible CC governance models in different ways, the Cybersecurity4Europe project went one step further by proposing a network of Community Hubs of Expertise in Cybersecurity Knowledge (CHECKs) [17], as a kind of the bottom-up approach to put forward a sub-structure of the cybersecurity ecosystem. There have been several categories taken into account for governance best practice analysis prior to the CHECK proposition. Besides stakeholder analysis and organizational structures, decision-making or coordination mechanism with external bodies have also been discussed. The model and its assumptions were validated in the real-life scenario, namely with cybersecurity stakeholders in Toulouse in an entity named Community Hubs of Expertise in Cybersecurity Knowledge (CHECK Toulouse) [18].

SPARTA project deliverables [152] describe the structures, processes, and activities that characterize the governance of the SPARTA and their adequacy for the ecosystem under the European Cybersecurity Competence Centre (ECCC). The project uses its own governance model as the case and concludes that there was strong utilization of some committees and processes (e.g., road mapping) while other governance bodies (notably the Certification Task Force, the Ethics Board, and the Advisory Committee were under-utilized, something to be attributed to the early stage of the project

ECHO project formulated the ecosystem as a Collaborative Networked Organisation (CNO) and described the development, assessment, and selection of the governance model alternative [26]. It used the Analytic Hierarchy Process (AHP) method to reach consensus among stakeholders, engaging European Cyber Security Organisation (ECSO) as well. The alternatives assessment was done by comparison of the governance model performance against each pre-defined criterion. The solution accepted by most stakeholders was to create one "umbrella" alternative, so-called Alternative 0 (A0), which would be based on best practices from the other four alternative models. In a later deliverable [151], the overall design framework was described through RACI (Responsible, Accountable, Consulted, Informed) matrix, while key process discovery (e.g., Strategic Planning Process; Partnership Development Process; Catalogue Management and Customer Relations Management process description; Innovation (R&D) Process) was made with the help of COBIT framework. There is also an assessment of organizational structure similarity of the ECHO Consortium and future ECHO CNO, with a conclusion that the consortium has similarities to ECHO CNO's Central Hub. The future roadmap was also considered with possibilities such as merging with ECSO and other three pilot projects in one organization (forming ECSCON in line with the joint Governance White paper agreed by the Commission and other pilot projects). Another option is the establishment of the ECHO Network as an NGO with Chapters and Central Hub to be the ECHO Cybersecurity Competence Centre after the end of the project.

5.4.4 Strategic directions, gaps, and challenges

The rationale for the "ecosystem" based model is also that it fits demand uncertainty & fragility. For example, there is little certainty regarding the window of opportunity, as cybersecurity technology moves and changes very fast. The length of time that will elapse between the prototype launch and the development of meaningful or scalable demand from EU users is shorter than it is for the other digital technologies. Early-stage users tend to be particularly fragile. Pioneer adopters do not have the testimonials, benchmarks, or references from the existing operational environment users or other types of evidence such as proof of value (PoV). Even if such evidence exists, a solution might not be easily transferred or reusable from one operational environment to another.

One idea coming from the joint meeting of four pilot projects and ENISA and related to ecosystem market-driven strategic directions is to use the CC marketplace as the single-entry point with the emphasis on early product and prototype visibility. The value proposition of such a marketplace needs to be worked out, but it could basically create an environment for the conversation and mapping between demand and supply-side in the EU cybersecurity market, eventually leading to better and more mature products. "Try before buy" or "test before invest" could also be reflected or linked in this marketplace, as well as a range of other free or subsidized services and business processes that could serve as a "hook" to sign up new customers, also outside EU.

Ecosystem-based trials or PoV projects might need innovative instruments involving the assumption of successful adoption from the demand side "pioneers" and later scaling-up. However, it is likely that some less profitable research or some cybersecurity segments will need to follow different paths, and this is where governance mechanism should include multiple tiers or substructures with parallel sets of rules or policies to keep the ecosystem aligned with the latest economic, business, but also breakthrough research objectives.

In the case of the CONCORDIA pilot project, for example, there are already some services offered for free to the CONCORDIA community, such as the possibility to test technologies or the catalog of online training

offerings. The project also addresses limitations and barriers when it comes to the adaptation of cybersecurity to the needs of SMEs and start-ups, such as the lack of expertise (CONCORDIA training), financial resources (CONCORDIA virtual lab), or optimization of the solutions to their needs and scale (CONCORDIA experiments). Ranking of exploitable results is done on an annual basis, having in mind the evolution of technology readiness level (TRL), but also market readiness level (MRL), innovation potential, and the importance of ecosystem support in its exploitation path.

Beyond the support for integrating specific cybersecurity technologies in the processes and products created by start-ups and SMEs, there is also significant business potential for European cybersecurity start-ups and SMEs to showcase their innovations and solutions and, in this way, attract investors, finding better geographical coverage and expanding besides national boundaries, as well as identifying partnerships in value networks, whether it is with larger cybersecurity suppliers or with the other start-up or SME that complements their solutions. The challenge here is to be able to implement assessment of highly innovative ideas through the community, especially for so-called "deep tech" start-ups, without disclosure of the idea itself.

The gap in the ecosystem might also appear between the top-down policy or market issues and bottom-up research and innovation. Finally, dynamics of relationships that have a direct impact on trust, and in consequence, on economics should also be addressed as a challenge for the ecosystem.

5.4.5 Conclusion

Cybersecurity ecosystems modeling is a topic of particular importance for the future EU cybersecurity policy, as it includes the establishment of the EU Cybersecurity Competence Centre (ECCC), Network of National Coordination Centres (NCCs), and Cybersecurity Community (CC). Four pilot projects of the cybersecurity competence community (ECHO, SPARTA, CSEU, and CONCORDIA) have started their work in 2019 and are continuously providing their feedback on a number of issues, including research roadmap, governance model, strategic directions, gaps, and challenges.

In this section, we tried to make an analysis of lessons learned from the other similar value co-creation experiences, cybersecurity communities, and public-private partnerships, as well as to present some work and the feedback collected in four pilot projects. Stakeholders in four pilot projects, which are also representative of the future ecosystem, might have multiple roles, while levels of membership or relationships between are likely to be very dynamic, in both formal and informal dimensions. EU cybersecurity market fragmentation is likely to be reduced, therefore increasing the economy of scale, but shaping instruments to enable mapping between demand and supply still needs to find a place. The overall benefit is expected to be very positive but also difficult to assess, as the economic impact attribution to "ecosystem existence" is hard to validate. Inclusion of "ecosystem role" parameters in the exploitation plan descriptions might help, for example, by addressing support for the testing, certification preparation, maintenance, data exchange, etc.

Although there is an aim to align industrial strategy and policy priorities with the generation of innovative ideas, there is still a need to improve policy-market-technology-society alignment, as well as embed economic issues in the ecosystem. Further gaps might appear in territorial coverage, capacity, and maturity of the cybersecurity landscape in different member states. The challenges related to the EU cybersecurity ecosystem and stakeholders outside of the EU are also yet to be addressed. Finally, the dynamics of relationships that have a direct impact on trust and, in consequence, economics is also be explored. Our next steps will be to focus on a concrete exploitable result that has heavy reliance on "ecosystem support", namely cybersecurity threat intelligence (CTI) sharing, and to analyze economic impact as a function of stakeholder involvement and support for it. We will also explore incentives and motivations for participants in CTI, as well as added value services that could be exploited by individual stakeholders.

6 Conclusion

This document, D3.23 Cybersecurity Outlook 2, updates and extends in new fields its previous version, D3.10 Cybersecurity Outlook 1. It reports and discusses new emerging technologies, current trends, and recent interactions between the cybersecurity worlds and human society.

This deliverable presented several new technologies and trends related to:

- software security and how an automatized assessment can improve the code quality and its protection level (Section 2);
- network security and the use of more intelligent security techniques to handle threats and share private data (Section 3);
- a discussion about the future and challenges of 6G, the next-generation wireless standard (Section 4);
- how cybersecurity awareness, laws, and our society are coping with the recent changes in the cybersecurity field (Section 5).

The content of this deliverable will hopefully provide some interesting ideas and food for thought about how the current cybersecurity technologies and developments are evolving and are starting to build a better technocentric world.

7 Bibliography

[1] J. Kim, S. Kang, E. S. Cho, J. Y. Paik, *"LOM: Lightweight Classifier for Obfuscation Methods"*, proceedings of WISA 2021: International Conference on Information Security Applications, pp. 3-15, Springer, 2021, https://doi.org/10.1007/978-3-030-89432-0_1

[2] Y. Zhao, Z. Tang, G. Ye, D. Peng, D. Fang, X. Chen, Z. Wang, *"Semantics-aware obfuscation scheme prediction for binary"*, Computers & Security, vol. 99, Elsevier, 2020, https://doi.org/10.1016/j.cose.2020.102072

[3] D. Canavese, L. Regano, C. Basile, *"Method for the identification of protected assets in software binaries"*, patent-pending, priority number 102021000012488, 2021, https://www.knowledge-share.eu/en/patent/method-for-the-identification-of-protected-assets-in-software-binaries/

[4] P. D. Borisov, Y. V. Kosolapov, *"On Characteristics of Symbolic Execution in the Problem of Assessing the Quality of Obfuscating Transformations"*, Modeling and Analysis of Information Systems, vol. 28, Âroslavl, 2021, https://doi.org/10.18255/1818-1015-2021-1-38-51

[5] S. Sebastio, E., F. Biondi, O. Decourbe, T. Given-Wilson, A. Legay, C. Puodzius, J. Quilbeuf, *"Optimizing symbolic execution for malware behavior classification"*, Computers & Security, vol. 93, Elsevier, 2020, https://doi.org/10.1016/j.cose.2020.101775

[6] N. Namani, A. Khan, *"Symbolic execution based feature extraction for detection of malware"*, proceedings of ICCCS 2020: International Conference on Computing, Communication and Security, pp. 1-6, IEEE, 2020, https://doi.org/10.1109/ICCCS49678.2020.9277493

[7] G. Menguy, S. Bardin, R. Bonichon, C. de Souza Lima, *"AI-based Blackbox Code Deobfuscation: Understand, Improve and Mitigate"*, 2021, https://arxiv.org/abs/2102.04805

[8] R. David, L. Coniglio, M. Ceccato, *"QSynth - A Program Synthesis based Approach for Binary Code Deobfuscation"*, proceedings of BAR 2020: workshop on Binary Analysis, pp. 1-12, IEEE, 2021, https://doi.org/10.14722/bar.2020.23009

[9] P. Kochberger, S. Schrittwieser, S. Schweighofer, P. Kieseberg, E. Weippl, *"SoK: Automatic Deobfuscation of Virtualization-protected Applications"*, proceedings of ARES 2021: International Conference on Availability, Reliability and Security, pp. 1-15, ACM, 2021, https://doi.org/10.1145/3465481.3465772

[10] C. Baier, J. P. Katoen, *"Principles of model checking"*, MIT Press, 2008, ISBN: 9780262026499

[11] B. Barbot, S. Haddad, C. Picaronny, *"Coupling and importance sampling for statistical model checking"*, procedings of TACAS 2012: Tools and Algorithms for the Construction and Analysis of Systems, pp. 331-346, Springer, 2012. https://doi.org/10.1007/978-3-642-28756-5_23

[12] C. E. Budde, *"Automation of Importance Splitting Techniques for Rare Event Simulation", Ph.D. thesis"*, Universidad Nacional de Córdoba, 2017, https://doi.org/10.1007/978-3-319-23267-6_18

[13] C. E. Budde, *"FIG: The finite improbability generator"*, procedings of TACAS 2020: Tools and Algorithms for the Construction and Analysis of Systems, pp. 483-491, Springer, 2020, https://doi.org/10.1007/978-3-030-45190-5_27

[14] C. E. Budde, *"FIG: the finite improbability generator"*, proceedings of TOSME 2021: Workshop on Tools for Stochastic Modelling and Evaluation, (to appear), 2021, https://www.performance2021.deib.polimi.it/tools-for-stochastic-modelling-and-evaluation/

[15] C. E. Budde, P. R. D'Argenio, A. Hartmanns, *"Automated compositional importance splitting"*, Science of Computer Programming, vol. 174, Elsevier, 2019, https://doi.org/10.1016/j.scico.2019.01.006

[16] F. Cérou, P. Del Moral, T. Furon, A. Guyader, *"Sequential Monte Carlo for rare event estimation"*, Statistics and Computing, vol. 22, Springer, 2012, https://doi.org/10.1007/s11222-011-9231-6

[17] P. R. D'Argenio, J. P. Katoen, *"A theory of stochastic systems part I: Stochastic automata"*, Information and Computation, vol. 203, Elsevier, 2005, https://doi.org/10.1016/j.ic.2005.07.001

[18] Z. Fang, H. Fu, T. Gu, Z. Qian, T. Jaeger, P. Hu, P. Mohapatra, *"A model checking-based security analysis framework for IoT systems"*, High-Confidence Computing, vol. 1, Elsevier, 2021, https://doi.org/10.1016/j.hcc.2021.100004

[19] R. Faqeh, C. Fetzer, H. Hermanns, J. Hoffmann, M. Klauck, M. A. Köhl, M. Steinmetz, C. Weidenbach, *"Towards dynamic dependable systems through evidence-based continuous certification"*, proceedings of ISoLA 2020: International Symposium on Leveraging Applications of Formal Methods, pp. 416–439, Springer, 2020, https://doi.org/10.1007/978-3-030-61470-6_25

[20] A. Hartmanns, *"On the analysis of stochastic timed systems"*, Ph.D. thesis, Saarland University, 2015, https://doi.org/10.22028/D291-26597

[21] S. Khan, J. P. Katoen, *"Synergising reliability modelling languages: BDMPs and repairable DFTs"*, proceedings of PRDC 2021: Pacific Rim International Symposium on Dependable Computing, pp. 113-122, IEEE, 2021, https://doi.org/10.1109/PRDC53464.2021.00023

[22] P. L'Ecuyer, F. Le Gland, P. Lezaud, B. Tuffin, *"Splitting techniques"*, Rare Event Simulation using Monte Carlo Methods, Wiley & Sons, 2009, ISBN: 9780470772690

[23] F. Massacci, I. Pashchenko, *"Technical leverage in a software ecosystem: Development opportunities and security risks"*, proceedings in ICSE 2015: International Conference on Software Engineering, pp. 1386-1397, IEEE, 2021, https://doi.org/10.1109/ICSE43902.2021.00125

[24] B. L. Mediouni, A. Nouri, M. Bozga, M. Dellabani, A. Legay, S. Bensalem, *"SBIP 2.0: Statistical model checking stochastic real-time systems"*, proceedings of ATVA 2018: International Symposium

on Automated Technology for Verification and Analysis, pp. 536-542, Springe, 2018, https://doi.org/10.1007/978-3-030-01090-4_33

[25] V. Menzel, J. L. Hurink, A. Remke, *"Securing SCADA networks for smart grids via a distributed evaluation of local sensor data"*, proceedings of SmartGridComm 2021: International Conference on Communications, Control, and Computing Technologies for Smart Grids, pp. 405-411, IEEE, 2021, https://doi.org/10.1109/SmartGridComm51999.2021.9632283

[26] I. Pashchenko, H. Plate, S. E. Ponta, A. Sabetta, F. Massacci, *"Vuln4Real: A methodology for counting actually vulnerable dependencies"*, Transaction on Software Engineering, IEEE, 2020, https://doi.org/10.1109/TSE.2020.3025443

[27] G. V. Post, A. Kagan, *"Evaluating information security tradeoffs: Restricting access can interfere with user tasks"*, Computers & Security, vol. 26, Elsevier, 2007, https://doi.org/10.1016/j.cose.2006.10.004

[28] K. Rindell, J. Ruohonen, J. Holvitie, S. Hyrynsalmi, V. Leppänen, *"Security in agile software development: A practitioner survey"*, Information and Software Technology, vol. 131, Elsevier, 2021, https://doi.org/10.1016/j.infsof.2020.106488

[29] R. Roberts, B. Lewis, A. Hartmanns, P. Basu, S. Roy, K. Chakraborty, Z. Zhang, *"Probabilistic verification for reliability of a two-by-two network-on-chip system"*, proceedings of FMICS 2021: International Conference on Formal Methods for Industrial Critical Systems, pp. 232-248, Springer, 2021, https://doi.org/10.1007/978-3-030-85248-1_16

[30] A. Z. Rose, N. Miller, *"Measurement of Cyber Resilience from an Economic Perspective"*, Applied Risk Analysis for Guiding Homeland Security Policy, Wiley & Sons, 2021, ISBN: 9781119287490

[31] G. Rubino, B. Tuffin, *"Introduction to rare event simulation"*, Rare Event Simulation using Monte Carlo Methods, Wiley & Sons, 2009, ISBN: 9780470772690

[32] G. Rubino, B. Tuffin, *"Rare Event Simulation using Monte Carlo Methods"*, Wiley & Sons, 2009, ISBN: 9780470772690

[33] T. W. B. Silva, D. C. Morais, H. G. Andrade, A. M. N. de Lima, E. U. Melcher, A. V. Brito, *"Environment for integration of distributed heterogeneous computing systems"*, Journal of Internet Services and Applications, vol. 9, Springer, 2018, https://doi.org/10.1186/s13174-017-0072-1

[34] M. Stoelinga, C. Kolb, S. M. Nicoletti, C. E. Budde, E. M. Hahn, *"The marriage between safety and cybersecurity: Still practicing"*, proceedings of SPIN 2021: International Symposium on Model Checking Software, pp. 3-21, Springer, 2021, https://doi.org/10.1007/978-3-030-84629-9_1

[35] C. Theisen, N. Munaiah, M. Al-Zyoud, J. C. Carver, A. Meneely, L. Williams, *"Attack surface definitions: A systematic literature review"*, Information and Software Technology, vol. 104, Elsevier 2018, https://doi.org/10.1016/j.infsof.2018.07.008

[36] H. I. S. Younes, R. G: Simmons, *"Probabilistic verification of discrete event systems using acceptance sampling"*, proceedings of CAV 2002: International Conference on Computer Aided Verification, pp. 223-235, Springer, 2002, https://doi.org/10.1007/3-540-45657-0_17

[37] Norwegian Consumer Council, *"Surveillance-based advertising: Consumer attitudes to surveillance-based advertising"*, 2021, accessed: 2021-12-21, https://fil.forbrukerradet.no/wp-content/uploads/2021/0 6/consumer-attitudes-to-surveillance-based-advertising.pdf

[38] CISCO, *"The IoT Value/Trust Paradox: Building Trust and Value in the Data Exchange Between People, Things and Providers"*, 2021, accessed: 2021-12-2, https://www.jasper.com/sites/default/files/cisco_ iot_survey_-_the_value_trust_paradox_final-1_2.pdf

[39] European Parliament - Council of the European Union, *"The General Data Protection Regulation (GDPR)"*, accessed: 2019-11-24, https://eur-lex.europa.eu/eli/reg/2016/679/oj

[40] CISCO, *"Consumer Privacy Survey: The growing imperative of getting data privacy right"*, accessed: 2022-01-06, https://www.cisco.com/c/dam/global/en_uk/products/collateral/security/cybersecurity-series-2019-cps.pdf

[41] M. Godinho de Matos, I. Adjerid, *"Consumer consent and firm targeting after GDPR: The case of a large telecom provider"*, Management Science, 2021, https://doi.org/10.1287/mnsc.2021.4054

[42] Privacy International, *"Grounds for Processing of Personal Data"*, accessed: 2022-03-17, https://privacyinternational.org/sites/default/files/2018-09/Part%205%20%20Grounds%20for%20Processing%20of%20Personal%20Data_0.pdf

[43] A. Ravichander, A. W. Black, T. Norton, S. Wilson, N. Sadeh, *"Breaking Down Walls of Text: How Can NLP Benefit Consumer Privacy?"*, proceedings of IJCNLP 2021: International Joint Conference on Natural Language Processing, vol. 1, ACL, 2021, http://dx.doi.org/10.18653/v1/2021.acl-long.319

[44] S. Barth, M. D. De Jong, *"The privacy paradox–Investigating discrepancies between expressed privacy concerns and actual online behavior–A systematic literature review"*, Telematics and informatics, vol. 34, Elsevier, 2017, https://doi.org/10.1016/j.tele.2017.04.013

[45] Y. Feng, Y. Yao, N. Sadeh, *"A Design Space for Privacy Choices: Towards Meaningful Privacy Control in the Internet of Things"*, proceedings of CHI 2021: Conference on Human Factors in Computing Systems, pp. 1–16, ACL, 2021, https://doi.org/10.1145/3411764.3445148

[46] A. Acquisti, L. Brandimarte, G. Loewenstein, *"Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age"*, Journal of Consumer Psychology, vol. 30, SSRN, 2020, https://doi.org/10.1002/jcpy.1191

[47] B. Fabian, T. Ermakova, T. Lentz, *"Large-scale readability analysis of privacy policies"*, proceedings of WI 2017: International Conference on Webintelligence, pp. 18–25, ACL, 2017, https://doi.org/10.1145/3106426.3106427

[48] V. Morel, M. Cunche, D. Le Métayer, *"A generic information and consent framework for the IoT"*, proceedings of TrustCom 2019: International Conference On Trust, Security And Privacy In Computing And Communications. pp. 366-373, IEEE, 2019, https://doi.org/10.1109/TrustCom/BigDataSE.2019.00056

[49] H. Habib, Y. Zou, Y. Yao, A. Acquisti, L. Cranor, J. Reidenberg, N. Sadeh, F. Schaub, *"Toggles, dollar signs, and triangles: How to (in) effectively convey privacy choices with icons and link texts"*, proceedings of CHI 2021: Conference on Human Factors in Computing Systems, pp. 1–25, ACL, 2021, https://doi.org/10.1145/3411764.3445387

[50] P. Emami-Naeini, J. Dheenadhayalan, Y. Agarwal, L. F. Cranor, *"Which Privacy and Security Attributes Most Impact Consumers' Risk Perception and Willingness to Purchase IoT Devices?"*, proceedings of SP 2021: Symposium on Security and Privacy, pp. 1937-1954, IEEE, 2021, https://doi.ieeecomputersociety.org/10.1109/SP40001.2021.00112

[51] H. Ali, P. Papadopoulos, J. Ahmad, N. Pitropakis, Z. Jaroucheh, and W. Buchanan, *"Privacy-preserving and trusted threat intelligence sharing using distributed ledgers"*, 2021, https://arxiv.org/abs/2112.10092

[52] S. Badsha, I. Vakilinia, and S. Sengupta, *"Privacy preserving cyber threat information sharing and learning for cyber defense"*, proceedings of CCWC 2019: Computing and Communication Workshop and Conference, pp. 708-714, IEEE, 2019, https://doi.org/10.1109/CCWC.2019.8666477

[53] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Aguera y Arcas, *"Communication-efficient learning of deep networks from decentralized data"*, proceedings of AISTATS: Conference on Artificial Intelligence and Statistics, pp. 1273-1282, MLR, 2017, https://doi.org/10.48550/arXiv.1602.05629

[54] D. Preuveneers, W. Joosen, J. Bernal Bernabe, and A. Skarmeta, *"Distributed security framework for reliable threat intelligence sharing"*, Security and Communication Networks, vol. 2020, Hindawi, 2020, https://doi.org/10.1155/2020/8833765

[55] L. Sweeney, *"K-anonymity: A model for protecting privacy"*, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, vol. 10, ACL, 2002, https://doi.org/10.1142/S0218488502001648

[56] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "L-diversity: Privacy beyond k-anonymity", Transaction on Knowledge Discovery from Data, vol. 1, ACL, 2007, https://doi.org/10.1145/1217299.1217302

[57] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity", proceedings of ICDE 2007: International Conference on Data Engineering, pp. 106-115, IEEE, 2017, https://doi.org/10.1109/ICDE.2007.367856

[58] D. Preuveneers and W. Joosen, *"Sharing machine learning models as indicators of compromise for cyber threat intelligence"*, Journal of Cybersecurity and Privacy, pp. 140-163, 2021, https://doi.org/10.3390/jcp1010008

[59] L. Ruff. et al., *"A Unifying Review of Deep and Shallow Anomaly Detection"*, Proceedings of the IEEE, https://arxiv.org/abs/2009.11732

[60] D. P. Kingma, M. Welling, *"Auto-Encoding Variational Bayes"*, ICLR, 2014, https://arxiv.org/abs/1312.6114

[61] I. T. Bousquet, S. Gelly, B. Schoelkopf, *"Wasserstein Auto-Encoders"*, ICLR, 2019, https://arxiv.org/abs/1711.01558

[62] I. J. Goodfellow et al., *"Generative adversarial nets"*, NIPS, pp. 2672-2680, https://arxiv.org/abs/1406.2661

[63] B. Scholkopf, et al., *"Support vector method for novelty detection",* NIPS, 1999

[64] F. T. Liu, K. M. Ting, Z. H. Zhou, *"Isolation Forest",* ICDM, 2008

[65] S. Ramaswamy, R. Rastogi, K. Shim, *"Efficient Algorithms for Mining Outliers from Large Data Sets"*, SIGMOID, 2000, https://doi.org/10.1145/335191.335437

[66] G. Pang, C. Shen, L. Cao, A. V. D. Hengel, *"Deep Learning for Anomaly Detection: A Review"*, Computing Surveys, vol. 54, ACM, 2021, https://doi.org/10.1145/3439950

[67] D. Bank, N. Koenigstein, R. Giryes, *"Autoencoders"*, CoRR, 2020, https://arxiv.org/abs/2003.05991

[68] J. Chen, Y. Shen, R. Ali, 2018, *"Credit Card Fraud Detection Using Sparse Autoencoder and Generative Adversarial Network"*, proceedings of IEMCON 2018: Annual Information Technology, Electronics and Mobile Communication Conference, pp. 1054-1059, IEEE, 2018, https://doi.org/10.1109/IEMCON.2018.8614815

[69] J. An, S. Cho, *"Variational autoencoder based anomaly detection using reconstruction probability"*, 2015

[70] A. L. Alfeo et al., "*Using an autoencoder in the design of an anomaly detector for smart manufacturing"*, Pattern Recognition Letters, vol. 136, Elsevier, 2020, https://doi.org/10.1016/j.patrec.2020.06.008

[71] S. Hawkins, H. He, G. J. Williams, R. A. Baxter, *"Outlier detection using replicator neural net-works"*, DaWaK, 2002

[72] C. Zhou, R. C. Paffenroth, *"Anomaly detection with robust deep autoencoders"*, KDD, 2017

[73] K. Tian, S. Zhou, J. Fan, J. Guan, *"Learning competitive and discriminative reconstructions for anomaly detection"*, proceedings of AAI 2019: Symposium on Educational Advances in Artificial Intelligence, pp. 5167-5174, ACM, 2019, https://doi.org/10.1609/aaai.v33i01.33015167

[74] F. D. Mattia, P. Galeone, M. D. Simoni, E. Ghelfi, *"A survey on gans for anomaly detection"*, CoRR, https://arxiv.org/abs/1906.11632

[75] T. Schlegl et al., *"Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery"*, proceedings of IPMI 2017: Conference on Information Processing in Medical Imaging, pp. 146-157, Springer, 2017, https://doi.org/10.1007/978-3-319-59050-9_12

[76] T. Salimans et al., *"Improved Techniques for Training GANs"*, NIPS, 2016

[77] T. Schlegl et al., *"F-anogan: Fast unsupervised anomaly detection with generative adversarial networks"*, Medical Image Analysis, vol. 54, Elsevier, 2019, https://doi.org/10.1016/j.media.2019.01.010

[78] H. Zenati et al., *"Efficient gan-based anomaly detection"*, CoRR, 2019

[79] H. Zenati et al., *"Adversarially learned anomaly detection"*, ICDM, 2018

[80] S: Akcay, A. Atapour-Abarghouei, T. P. Breckon, *"GANomaly: Semi-Supervised Anomaly Detection via Adversarial Training"*, proceedings of ACCV 2018: Asian Conference on Computer Vision, pp. 622-637, ACM, 2018, https://doi.org/10.1007/978-3-030-20893-6_39

[81] S. Akçay, A. Atapour-Abarghouei, T. P. Breckon, *"Skip- ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection"*, 2019, https://arxiv.org/abs/1901.08954

[82] H. S. Vu. et al., *"Anomaly Detection with Adversarial Dual Autoencoders"*, CoRR, 2019

[83] J. Chen, S. Sathe, C. Aggarwal, D. Turaga, *"Outlier detection with autoencoder ensembles"*, proceedings of SIAM 2017: International Conference on Data Mining, pp. 90-98, SIAM, 2017, https://doi.org/10.1137/1.9781611974973.11

[84] X. Han, X. Chen, L. P. Liu, *"Gan ensemble for anomaly detection"*, proceedings of AAI 2021: Symposium on Educational Advances in Artificial Intelligence, pp. 4090-4097, ACM, 2021, https://arxiv.org/abs/2012.07988

[85] N. Laptev, *"Anogen: Deep anomaly generator"*, 2018

[86] S. G. Rizzo, L. Pang, Y. Chen, S. Chawla, *"Probabilistic outlier detection and generation"*, 2020, https://arxiv.org/abs/2012.12394

[87] I. Tolstikhin, O. Bousquet, S. Gelly, B. Schoelkopf, *"Wasserstein auto-encoders"*, ICLR, 2019, https://arxiv.org/abs/1711.01558

[88] A. Liguori, G. Manco, F. S. Pisani, E. Ritacco, *"Adversarial Regularized Reconstruction for Anomaly Detection and Generation"*, proceedings of ICDM 2021: International Conference on Data Mining, pp. 1204-1209, IEEE, 2021, https://doi.org/10.1109/ICDM51629.2021.00145

[89] A. Milenkoski et al., *"Evaluating Computer Intrusion Detection Systems: A Survey of Common Practices"*, Computing Surveys, vol. 48, ACM, 2015, https://doi.org/10.1145/2808691

[90] H. Chen, L. Jiang, *"Efficient GAN-based method for cyber-intrusion detection"*, 2019, https://arxiv.org/abs/1904.02426

[91] V. Chandola, A. Banerjee, V. Kumar, *"Anomaly Detection : A Survey"*, Computing Surveys, vol. 41, ACM, 2009, https://doi.org/10.1145/1541880.1541882

[92] C. C. Hsu, C. Y. Lee, Y. X. Zhuan, *"Learning to Detect Fake Face Images in the Wild"*, 2018, https://arxiv.org/abs/1809.08754

[93] K. Shu et al., *"Fake News Detection on Social Media: A Data Mining Perspective"*, 2017, https://arxiv.org/abs/1708.01967

[94] H. Aghakhani et al., *"Detecting Deceptive Reviews using Generative Adversarial Networks"*, 2018, https://arxiv.org/abs/1805.10364

[95] C. Zunino, A. Valenzano, R. Obermaisser, S. Petersen, *"Factory communications at the dawn of the fourth industrial revolution"*, Computer Standards & Interfaces, vol. 71, Elsevier, 2020, https://doi.org/10.1016/j.csi.2020.103433

[96] J. E. Rubio, C. Alcaraz, R. Roman, J. Lopez, *"Current cyber-defense trends in industrial control systems"*, Computers & Security, vol. 87, Elsevier, 2019, https://doi.org/10.1016/j.cose.2019.06.015

[97] W. Yuhong, H. Xiangdong, *"Industrial internet security protection based on an industrial firewall"*, proceedings of ICAICA: International Conference on Artificial Intelligence and Computer Applications, pp. 239-247, IEEE, 2021

[98] International Electrotechnical Commission, *"IEC 62443 Security for Industrial Automation and Control Systems"*

[99] L. Durante, L. Seno, and A. Valenzano, *"A formal model and technique to redistribute the packet filtering load in multiple firewall networks"*, Transactions on Information Forensics and Security, vol. 16, IEEE, 2021, https://doi.org/10.1109/TIFS.2021.3057552

[100] L. Seno, M. Cheminod, I. Cibrario Bertolotti, L. Durante and A. Valenzano, *"Improving performance and cyber-attack resilience in multi-firewall industrial networks"*, to appear on

proceedings of WFCS 2022: International Conference on Factory Communication Systems, Springer, 2022

[101] M. H. Alsharif, A. H. Kelechi, M. A. Albreem, S. A. Chaudhry, M. S. Zia, S. Kim, *"Sixth generation (6G) wireless networks: Vision, research activities, challenges and potential solutions"*, Symmetry, vol. 12, MDPI, 2020, https://doi.org/10.3390/sym12040676

[102] A. Shahraki, M. Abbasi, M. Piran, A. Taherkordi, *"A comprehensive survey on 6G networks: Applications, core services, enabling technologies, and future challenges"*, 2021, https://arxiv.org/abs/2101.12475

[103] I. F. Akyildiz, A. Kak, S. Nie, *"6G and beyond: The future of wireless communications systems"*, Access, vol. 8, IEEE, 2020, https://doi.org/10.1109/ACCESS.2020.3010896

[104] B. Han, W. Jiang, M. A. Habibi, H. D. Schotten, *"An abstracted survey on 6G: Drivers, requirements, efforts, and enablers"*, 2021, https://arxiv.org/abs/2101.01062

[105] M. Ikram, K. Sultan, M. F. Lateef, A. S. Alqadami, *"A Road towards 6G Communication—A Review of 5G Antennas, Arrays, and Wearable Devices"*, Electronics, vol. 11, MDPI, 2022, https://doi.org/10.3390/electronics11010169

[106] W. Jiang, B. Han, M. A. Habibi, H. D. Schotten, *"The road towards 6G: A comprehensive survey"*, Open Journal of the Communications Society, vol. 2, IEEE, 2021, https://doi.org/10.1109/OJCOMS.2021.3057679

[107] Y. Zhao, J. Zhao, W. Zhai, S. Sun, D. Niyato, K. Y. Lam, "A survey of 6G wireless communications: Emerging technologies", proceedings of FICC 2021: Future of Information and Communication Conference, pp. 150-170, Springer, 2021, https://doi.org/10.1007/978-3-030-73100-7_12

[108] J. Wang, X. Ling, Y. Le, Y. Huang, X. You, *"Blockchain-enabled wireless communications: a new paradigm towards 6G"*, National Science Review, vol. 8, Oxford Academic, 2021, https://doi.org/10.1093/nsr/nwab069

[109] M. Ylianttila et al., *"6G white paper: Research challenges for trust, security and privacy"*, 2020, https://arxiv.org/abs/2004.11665

[110] M. Alsabah et al., *"6G wireless communications networks: A comprehensive survey"*, Access, vol. 9, IEEE, 2021, https://doi.org/10.1109/ACCESS.2021.3124812

[111] A. A. Solyman, K. Yahya, *"Key performance requirement of future next wireless networks (6G)"*, Bulletin of Electrical Engineering and Informatics, vol. 10, 2021, https://doi.org/10.11591/eei.v10i6.3176

[112] P. Porambage, G. Gür, D. P. M. Osorio, M. Liyanage, A. Gurtov, M. Ylianttila, *"The roadmap to 6G security and privacy"*, Open Journal of the Communications Society, vol. 2, IEEE, 2021, https://doi.org/10.1109/OJCOMS.2021.3078081

[113] M. M. Aslam, L. Du, X. Zhang, Y. Chen, Z. Ahmed, B. Qureshi, *"Sixth Generation (6G) Cognitive Radio Network (CRN) Application, Requirements, Security Issues, and Key Challenges"*, Wireless Communications and Mobile Computing, vol. 2021, Hindawi, 2021, https://doi.org/10.1155/2021/1331428

[114] J. He, K. Yang, H. H. Chen, *"6G cellular networks and connected autonomous vehicles"*, Network, vol. 35, IEEE, 2020, https://doi.org/10.1109/MNET.011.2000541

[115] V. L. Nguyen, P. C. Lin, B. C. Cheng, R. H. Hwang, Y. D. Lin, *"Security and privacy for 6G: A survey on prospective technologies and challenges"*, Communications Surveys & Tutorials, vol. 23, IEEE, 2021, https://arxiv.org/abs/2108.11861

[116] Y. Siriwardhana, P. Porambage, M. Liyanage, M. Ylianttila, *"AI and 6G security: Opportunities and challenges"*, proceedings of EuCNC: European Conference on Networks and Communications, pp. 616-621, IEEE, 2021, https://doi.org/10.1109/EuCNC/6GSummit51104.2021.9482503

[117] V. Raj, A. CA, *"Understanding the Future Communication: 5G to 6G"*, International Research Journal on Advanced Science Hub, vol. 3, RSP, 2021, http://dx.doi.org/10.47392/irjash.2021.159

[118] B. Yang et al., *"Edge intelligence for autonomous driving in 6G wireless system: Design challenges and solutions"*, Wireless Communications, vol. 28, IEEEW, 2021, https://doi.org/10.1109/MWC.001.2000292

[119] D. C. Nguyen et al., *"6G Internet of Things: A comprehensive survey"*, Internet of Things Journal, vol. 9, IEEE, 2021, https://doi.org/10.1109/JIOT.2021.3103320

[120] Z. Yang, M. Chen, K. K. Wong, H. V. Poor, S. Cui, *"Federated learning for 6G: Applications, challenges, and opportunities"*, Engineering, vol. 8, Elsevier, 2021, https://doi.org/10.1016/j.eng.2021.12.002

[121] A. Slalmi, H. Chaibi, A. Chehri, R. Saadane, G. Jeon, *"Toward 6G: Understanding network requirements and key performance indicators"*, Transactions on Emerging Telecommunications Technologies, vol. 32, ACM, 2021, https://doi.org/10.1002/ett.4201

[122] M. Matinmikko-Blue, S. Yrjölä, P. Ahokangas, K. Ojutkangas, E. Rossi, *"6G and the UN SDGs: Where is the Connection?"*, Wireless Personal Communications, vol. 121, Springer, 2021, https://doi.org/10.1007/s11277-021-09058-y

[123] B. Ozpoyraz, A. T. Dogukan, Y. Gevez, U. Altun, E. Basar, *"Deep Learning-Aided 6G Wireless Networks: A Comprehensive Survey of Revolutionary PHY Architectures"*, 2022, https://arxiv.org/abs/2201.03866

[124] S. Basharat, S. A. Hassan, H. Pervaiz, A. Mahmood, Z. Ding, M. Gidlund, *"Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6G wireless networks"*, Wireless Communications, vol. 28, IEEE, 2021, https://doi.org/10.1109/MWC.011.2100016

[125] N. Arastouie, M. Sabaei, B. Bakhshi, "*Near-optimal online routing in opportunistic networks*", International Journal of Communication Systems, vol. 32, Wiley & Sons, 2019, https://doi.org/10.1002/dac.3863

[126] N. Arastouie, M. Sabaei, *"Self-adaptive risk-aware routing in opportunistic network"*, The Journal of Supercomputing, vol. 74, ACM, 2018, https://doi.org/10.1007/s11227-018-2264-2

[127] K. B. Letaief, Y. Shi, J. Lu, J. Lu, *"Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications"*, Journal on Selected Areas in Communications, vol. 40, IEEE, 2021, https://doi.org/10.1109/JSAC.2021.3126076

[128] A. P. Liff, *"Cyberwar: A new 'absolute weapon'? The proliferation of cyberwarfare capabilities and interstate war"*, Journal of Strategic Studies, vol. 35, 2012, https://doi.org/10.1080/01402390.2012.663252

[129] N. Melzer, *"International Humanitarian Law: A Comprehensive Introduction"*, Geneva, Switzerland: International Committee of the Red Cross, 2016

[130] P. D. Scharre, *"Where does the human belong in the loop?"*, accessed: 2021-07-02, https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2014)/Scharre_MX_LAWS_technical 2014.pdf

[131] M. Leese, *"Configuring warfare: Automation, control, agency"*, Technology and Agency in International Relations, M. Hoijtink and M. Leese, Eds., 2019

[132] S. Bhatt, P. K. Manadhata, L. Zomlot, *"The operational role of security information and event management systems"*, Security & Privacy, vol. 12, IEEE, 2014, https://doi.org/10.1109/MSP.2014.103

[133] F. B. Kokulu et al., *"Matched and mismatched SOCs: A qualitative study on security operations center issues"*, proceedings of SIGSAC 2019: Conference on Computational Communication Security, pp. 1955–1970, ACM, 2019, https://doi.org/10.1145/3319535.3354239

[134] R. Langner, *"Stuxnet: Dissecting a cyberwarfare weapon"*, Security & Privacy, vol. 9, IEEE, 2011, https://doi.org/10.1109/MSP.2011.67

[135] A. B. Lopez, K. Vatanparvar, A. P. D. Nath, S. Yang, S. Bhunia, M. A. A. Faruque, *"A security perspective on battery systems of the Internet of Things"*, Journal of Hardware System Security, vol. 1, 2017, https://doi.org/10.1007/s41635-017-0007-0

[136] J. H. Saltzer, M. D. Schroeder, *"The protection of information in computer systems"*, Proceedings of the IEEE, vol. 63, IEEE, 1975, https://doi.org/10.1109/PROC.1975.9939

[137] F. Massacci, *"Is 'deny access' a valid 'fail-safe default' principle for building security in cyberphysical systems?"*, Security & Privacy, vol. 17, IEEE, 2019, https://doi.org/10.1109/MSEC.2019.2918820

[138] W. Lee, *"Malware and Attack Technologies Knowledge Area Issue"*, 2021, https://www.cybok.org/media/downloads/CyBOK_v1.1.0.pdf

[139] T. Rid, B. Buchanan, *"Attributing cyber attacks"*, Journal of Strategic Studies, vol. 38, 2015, https://doi.org/10.1080/01402390.2014.977382

[140] S. Chaudhary, S. Pape, M. Kompara, G. Kavallieratos and V. Gkioulos, *"D3.19 Guidelines for Enhancement of Societal Security Awareness"*, CyberSec4Europe Deliverable, 2022

[141] I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, A. Ng, *"Cybersecurity data science: an overview from machine learning perspective"*, Journal of Big Data, vol. 7, Springer, 2020, https://doi.org/10.1186/s40537-020-00318-5

[142] M. Abdullahi, Y. Baashar, H. Alhussian, A. Alwadain, N. Aziz, L. F. Capretz, S. J. Abdulkadir, *"Detecting Cybersecurity Attacks in Internet of Things Using Artificial Intelligence Methods: A Systematic Literature Review"*, Electronics , vol. 11, MDPI, 2022, https://doi.org/10.3390/electronics11020198

[143] D. Gibert, C. Mateu, J. Planes, *"The rise of machine learning for detection and classification of malware: Research developments, trends and challenges"*, Journal of Network and Computer Applications, vol. 153, Elsevier, 2020, https://doi.org/10.1016/j.jnca.2019.102526

[144] M. Musser, A. Garriott, *"Machine Learning and Cybersecurity: Hype and Reality"*, Center for Security and Emerging Technology, 2021

[145] M. Stone, *"How to Optimize Security Awareness Training for Different Groups"*, Security Intelligence, accessed: 2019-09-26, https://securityintelligence.com/articles/how-to-optimize-security-awareness-training-for-different-groups/

[146] A. Alruwaili, *"A Review of the Impact of Training on Cybersecurity Awareness"*, International Journal of Advanced Research in Computer Science, vol. 10, MDPI, 2019

[147] E. Anggraeni, E. Hartigh, M. Zegveld, *"Business ecosystem as a perspective for studying the relations between firms and their business networks"*, proceedings of ECCON: International Economics Conference, 2007

[148] S. Wieninger, R. Götzen, G. Gudergan, K. Wenning, *"The strategic analysis of business ecosystems : New conception and practical application of a research approach"*, proceedings of ICE 2019: International Conference on Engineering, Technology and Innovation, pp. 1-8, Springer, 2019, https://doi.org/10.1109/ICE.2019.8792657

[149] T. Noe, G. Parker, *"Winner Take All: Competition, Strategy, and the Structure of Returns in the Internet Economy"*, Journal of Economics & Management Strategy, vol. 14, https://doi.org/10.1111/j.1430-9134.2005.00037.x

[150] J. Arkko, *"The influence of internet architecture on centralised versus distributed internet services"*, Journal of Cyber Policy, vol. 5, https://doi.org/10.1080/23738871.2020.1740753

[151] ECHO project, Deliverable D3.3, "Governance Model Description"

[152] SPARTA project, Deliverable D1.2, "Lessons learned from internally assessing a CCN pilot"

[153] F. Nachira, P. Dini, A. Nicolai, *"A network of digital business ecosystems for Europe: Roots, processes and perspectives"*, Digital business ecosystem, European Commission Information Society and Media, 2017

[154] J. Stanley, G. Briscoe, *"The ABC of digital business ecosystems"*, Journal of Computer, Media and Telecommunications Law, vol. 15, https://arxiv.org/abs/1005.1899

[155] P. K. Senyo, K. Liu, L. Sun, J. Effah, *"Evolution of norms in the emergence of digital business ecosystems"*, Socially aware organisations and technologies. Impact and challenges, https://dx.doi.org/10.1007/978-3-319-42102-5_9

[156] EC publications, *"Digital Business Ecosystems"*, 2007

[157] "*Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace"*, JOIN/2013/01 final, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52013JC0001

[158] ENISA, *"ENISA study on Cooperative Models for Public-Private Partnership (PPP) and Information Sharing and Analysis Centers (ISACs)"*, https://op.europa.eu/en/publication-detail/-/publication/597dee0f-2285-11e8-ac73-01aa75ed71a1

[159] Cybersecurity4Europe project, Deliverable D2.1, *"Governance Structure v1.0"*

[160] Cybersecurity4Europe project, Deliverable D2.2, *"Internal Validation of Governance Structure"*

[161] Digitising European Industry Imitative Report, Working Group 2, *"Digital Industrial Platforms"*, 2017

[162] DIHNET.EU project, Deliverable, *"Defining Digital Innovation Hubs as part of the European DIH network"*

[163] JRC technical reports, *"Digital Innovation Hubs in Smart Specialisation Strategies, Early lessons from European regions"*, 2018

[164] JRC Policy Insights, *"The Impact of Smart Specialisation on the Governance of Research and Innovation Policy Systems"*, 2021

[165] CONCORDIA project, Deliverable D6.3, *"Innovation management strategy"*

[166] ECHO project, Deliverable D3.2, *"Governance alternatives"*

[167] ENISA, *"EP3R 2010-2013, Four Years of Pan-European Public Private Cooperation"*, 2014

[168] A. Pasic, *"NESSI and ESFORS: Paving the way towards secure software services"*, European Critical Information Infrastructure Newsletter, 2006

[169] A. Pasic, "Building blocks for Future Internet of Services: Trust, Security, Privacy and Dependability", New Architectures for Future Internet, MIT press, 2009

[170] A. Pasic, *"Delivering Building Blocks for Internet of Services: Trust, Security, Privacy and Dependability"*, Studies in Computational Intelligence, vol. 297, Springer, 2010,

[171] ENISA report, *"Good Practice Guide on Cooperative Models for Effective PPPs"*, October 2011