future internet

MDPI

Article

A Perfect Match: Converging and Automating Privacy and Security Impact Assessment On-the-Fly

Dimitrios Papamartzivanos [1,*], Sofia Anna Menesidou [1], Panagiotis Gouvas [1] and Thanassis Giannetsos [2]

[1] R&D Department, Ubitech Ltd., 11632 Athens, Greece; smenesidou@ubitech.eu (S.A.M.); pgouvas@ubitech.eu (P.G.)
[2] DTU Compute, Department of Applied Mathematics and Computer Science, Technical University of Denmark, 2800 Lyngby, Denmark; atgi@dtu.dk
* Correspondence: dpapamartz@ubitech.eu

Abstract: As the upsurge of information and communication technologies has become the foundation of all modern application domains, fueled by the unprecedented amount of data being processed and exchanged, besides security concerns, there are also pressing privacy considerations that come into play. Compounding this issue, there is currently a documented gap between the cybersecurity and privacy risk assessment (RA) avenues, which are treated as distinct management processes and capitalise on rather rigid and make-like approaches. In this paper, we aim to combine the best of both worlds by proposing the APSIA (Automated Privacy and Security Impact Assessment) methodology, which stands for Automated Privacy and Security Impact Assessment. APSIA is powered by the use of interdependency graph models and data processing flows used to create a digital reflection of the cyber-physical environment of an organisation. Along with this model, we present a novel and extensible privacy risk scoring system for quantifying the privacy impact triggered by the identified vulnerabilities of the ICT infrastructure of an organisation. We provide a prototype implementation and demonstrate its applicability and efficacy through a specific case study in the context of a heavily regulated sector (i.e., assistive healthcare domain) where strict security and privacy considerations are not only expected but mandated so as to better showcase the beneficial characteristics of APSIA. Our approach can complement any existing security-based RA tool and provide the means to conduct an enhanced, dynamic and generic assessment as an integral part of an iterative and unified risk assessment process on-the-fly. Based on our findings, we posit open issues and challenges, and discuss possible ways to address them, so that such holistic security and privacy mechanisms can reach their full potential towards solving this conundrum.

Keywords: Privacy Impact Assessment; General Data Protection Regulation; privacy scoring system; privacy quantification; healthcare data privacy

Citation: Papamartzivanos, D.; Menesidou, S.A.; Gouvas, P.; Giannetsos, T. A Perfect Match: Converging and Automating Privacy & Security Impact Assessment On-the-Fly. *Future Internet* **2021**, *13*, 30. https://doi.org/10.3390/fi13020030

Academic Editor: Claude Chaudet
Received: 31 December 2020
Accepted: 21 January 2021
Published: 27 January 2021

1. Introduction

Six decades since the start of the computer revolution, four decades since the invention of the micro-processor, and two decades into the rise of the modern Internet, all of the technology required to transform industries through software has finally matured and can be widely delivered at a global scale. Moreover, with the advent of the Internet of Things (IoT), the world just begun reaping the benefits of this evolution. However, this evolution brings several new challenges (or makes existing unsolved challenges urgent to be tackled) with security, interoperability, integrability, and composability being some of the major concerns at both logical extremes of a network. Currently, such challenges are addressed by next-generation approaches including model-based standards, ontology, Business Process Model Life-Cycle Management (BPM LCM), context of business process, and a host of other transport protocols [1–3].

Security, on the other hand, is of paramount importance and is addressed by conducting risk management as a first step. According to the European Union Agency for

Cybersecurity (ENISA), in the face of an increasing attack landscape (Figure 1), it is imperative to ensure the correct and safe operation of all safety-critical business processes because, by their very nature, the internal physical and cyber (data and computing) assets—of an ecosystem or application domain—may not always be in trusted custody. Towards this direction, organizations must perform risk management so that they can identify and assess risks in order to keep them at acceptable levels. It serves as the foundation on which organizations can start building a well-rounded cybersecurity strategy.

In this context, the most fundamental component of risk management is risk assessment. Risk assessment targets the identification of threats and of respective risk levels, thus, allowing for the overall risk management to keep cybersecurity risks under an acceptable threshold [4]. According to the Risk Management Framework (RMF) proposed by the National Institute of Standards and Technology (NIST), risk assessment can be performed in three tiers: the organizational; business process; and Information Systems tier [3]. In the era where "service is everything and everything is a service", this risk compartmentalization enables the emerging trend of the intelligent edge computing and the backend information service provider to operate in tandem, so as to provide flexible design choices that best meet business and operational goals. However, as the upsurge of such information and communication technologies has become the nervous system of all modern economies, fueled by the unprecedented amount of personal data being processed and exchanged [5], besides security concerns, there are also pressing privacy considerations that come into play and need to be taken into consideration.

The rising number of privacy related incidents has, therefore, mandated much stricter data protection and privacy principles, especially in sectors where massive amounts of sensitive data are processed; i.e., such as healthcare, Industrial IoT, etc. [6–8]. Privacy refers to the protection of personal data or Personally Identifiable Information (PII) and is regarded as a human right in the digital age [9]. Privacy is challenging not only because it is an all-encompassing concept that helps to safeguard important values, such as human autonomy and dignity, but also because there is a wide spectrum of threats against it, and at the same time, the means for achieving it can vary [10,11]. Towards this direction, privacy risk management approaches have emerged in order to identify and mitigate privacy risks raised during data processing activities. The paramount importance of privacy preservation, along with the establishment of the General Data Protection Regulation (GDPR) that governs the European territory, has rendered the privacy risk assessment a rapidly changing field, due to notable standardisation actions and legal establishments. However, a common language and a practical methodology that is flexible enough to address diverse privacy needs is still missing [10], while the community has identified the need for further research towards the implementation of tools to support Privacy Impact Assessment (PIA) conduction [12].

Even though there are several frameworks that set the principles for the conduction of privacy risk assessment, PIA remains a challenging and maze-like process mainly due to the multiple aspects that an assessor needs to consider; sometimes by having limited view to the details of the processing activities of interest and the supporting assets of the ICT infrastructure. In this line of research, intensive research efforts have converged to the proposal of a wide gamut of PIA complementary tools: While they can prove valuable during the assessment process, it appears that most of them perform the assessment without considering the cybersecurity status of the ICT infrastructure. This is rather contradictory considering the NIST proposed guidelines [10], according to which, the data protection lies in the intersection of the cybersecurity and privacy risks. Therefore, the lack of sufficient tools and methodologies that can offer the PIA a level of automation [12], so that risk impact assessment can be extracted in a timely manner, but most importantly the lack of proper metrics that can steer the decisions of the assessor in identifying, prioritising, anticipating and, finally, mitigating the risks, could raise questions on the effectiveness of the tools and the thoroughness of the assessments. Indeed, the diversity and complexity of proposed privacy properties makes an informed choice of metrics rather challenging [13] and sets

the challenge ahead on how to efficiently converge the usually contradicting security and privacy requirements, of a business ecosystem, towards a better risk assessment and estimation.

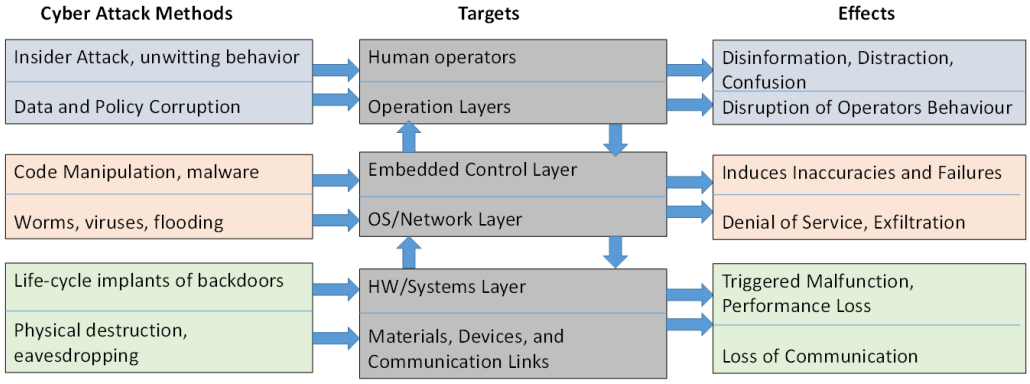

Figure 1. Potential cyber-attack methods, targets and effects.

Contributions: We meet these challenges with APSIA (Automated Privacy and Security Impact Assessment); a novel systematic approach towards assessing the privacy impact levels of organisations while also considering their cybersecurity status and threat landscape, as those are formed by the chains of interdependent ICT assets used to realise the data processing. APSIA is capable not only to support and complement PIA procedures, but also to enhance their assessments with dynamic asset inventory and vulnerability discovery techniques, extending beyond the initial setup and deployment of a business ecosystem to also consider the entire operational life-cycle of an organization. More specifically, APSIA: (i) is generic and applicable to any type of business ecosystem or application domain, (ii) leverages a novel and extensible privacy risk scoring system for quantifying the privacy risks triggered by the identified vulnerabilities of the ICT infrastructure of an organisation, (iii) provides a dynamic model for mapping core GDPR entities and requirements with tangible (e.g., databases, servers) and intangible assets (e.g., data records, PII) of the ICT infrastructures, so that to enable the risk assessors to effectively define data processing activities, and (iv) enables the risk assessor to keep track of all the needed information and assess the degree of compliance of the organisation including in the assessment the existence of risk mitigation actions. We provide a prototype implementation of APSIA and demonstrate its applicability and efficacy through a specific case study in a healthcare setting. While one could argue that APSIA is limited only to the GDPR-related privacy considerations, the core methodology is generic enough to facilitate any ecosystem with "privacy-by-design" properties. Since such properties are independent of the system or the application domain itself, putting forth the methodology to be able to precisely model them in the presence of strong adversaries is a prerequisite for any legal framework. Overall, our approach can complement any existing security-based risk assessment tool, closing the existing gap and enhancing the privacy posture of all involved actors against powerful adversaries.

The rest of this paper is organized as follows: In Section 2, we offer an overview of the related endeavors and frameworks in the PIA field. Section 3 presents the APSIA methodology and elaborates on the systematic approach of regulating and conducting the cybersecurity and privacy risk assessment. We then provide a detailed description of the APSIA's workflow of actions and internal components (Section 4), followed by a showcase of its mode of operation in a healthcare environment (Section 5). Section 6 puts forth open issues and challenges that still need to be considered towards even more holistic risk assessment services capable of capturing the intricacies of emerging "Systems-of-Systems" settings. Finally, Section 7 concludes the paper.

2. The Emergence of Privacy Impact Assessment

PIA is a risk management approach which aims at the evaluation of potential effects that systems may have on privacy, due to processing actions on personal data [11]. Via this systematic approach, organisations can anticipate the risks of their initiatives during their life-cycle, starting from the design phase—enabling a "privacy-by-design" approach—but also during their operational life-cycle by performing iterative assessment. Data Protection Authorities (DPAs) and standardisation bodies have established legal frameworks and guidelines which mandate the conduction of PIA. In an effort to gain the European citizens' trust for digital services, the General Data Protection Regulation (GDPR) refers to the obligation of the data controller to conduct an impact assessment and to document it prior to starting the intended data processing (Art. 35). The International Organization for Standardization (ISO) released a PIA guidelines standard, namely ISO/IEC 29134:2017, for standardising the PIA per se, and the reporting process and format [14].

In what follows, we provide an overview of relevant risk assessment standards and PIA tools in order highlight their benefits but also challenges in comparison to APSIA. The intuition is to showcase that besides the NIST privacy framework [10], the aforementioned privacy concerns are highly overlooked in today's standards while a common quantification formula for calculating the privacy risks together with the existing cybersecurity metrics is still an open and challenging task.

2.1. Methodologies, Standards and GDPR Guidelines

There are several privacy data protection standards, such as BS 10012:2017 [15], the ISO/IEC 29151:2017 [16] and the ISO/IEC 27018:2014 [17], where the Privacy Impact Assessment (PIA) included as a mandatory step towards conducting cyber risk assessment. Unfortunately, there is no explicit methodology that performs a PIA consolidated with a risk assessment process. The vast majority treats the PIA independently of the cyber risk assessment [18,19], even though according to the NIST privacy framework [10] the data protection lies in the intersection between the cyber security and privacy risks. In addition, there is a documented gap for automated tools that can support the conduction of a PIA [4]. A comprehensive guidance for carrying out a PIA presented in ISO/IEC 29134:2017 [20], however, it solely describes the the basic concepts for the impact analysis while the provided information for the risk assessor is inadequate [19]. Moreover, several privacy metrics have been documented in the literature by now, however these generally utilise criteria of privacy-enhancing technologies (PETs), such as the the quantification of leaked information or the number of indistinguishable users, instead of the privacy impact [21]. More recently, the NIST proposed a privacy framework in the form of a solid documentation and a methodology to manage the privacy risks of an organization by prioritizing privacy protection activities through enterprise risk management [10].

The GDPR commands controllers to perform a risk oriented approach for the personal data, the Data Protection Impact Assessment (DPIA) [20]. However, GDPR does not dictate a special assessment method, while at the same time mandates a good overview of the PIIs, since any inappropriate management of PIIs can possibly violate the GDPR [21]. Such an overview is a challenging task especially for complex systems designed before the GDPR era. To handle this issue, it is necessary to identify and document all the activities related to processing the PIIs [21].

Numerous national regulators have circulated guidelines for DPIA, among them are the French CNIL [22], the British ICO [20] and Canada's Privacy Act [23]. Such guides has been amended to support DPIAs and to provide comprehensive guidelines about their regulatory requirements and processes. These guidelines follow different approaches and propose diverse steps for conducting a PIA, while are abstract or imprecise, making difficult to conduct such methodologies [24]. Thus, the adoption of a single methodology becomes a difficult task for an organisation and this results to inadequate support for conducting a PIA [4]. A more recent research work, reviewed the most known DPIA methods based on seventeen questions derived from the literature in order to highlight

the absence of completeness among DPIA methods [4]. While there are differences in the aforementioned approaches, they are equally suitable for conducting a DPIA and produce largely similar results.

2.2. Understanding the Differences between PIAs and the GDPR's DPIAs

In addition to the aforementioned regulatory efforts, amendments on the DPIA processes has also been proposed from the academia. These efforts include making the DPIA process more systematic and structured by using formal modeling techniques for privacy threats [19]. LINDDUN is a privacy threat analysis framework that can support analysts in systematically eliciting and mitigating privacy threats and consists of six main steps [25]. The acronym stands for Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness and Non-compliance. LINDDUN and the CNIL method [26] have common principles. LINDDUN, however, has the functionality to visualise data flow diagrams and privacy threat tree patterns [27] compared to the CNIL. However, LINDDUN is missing assessment steps from a legal perspective and also is not integrated with a risk assessment process [21]. The work in [28] presents a methodology, firstly introduced in [24], to reinforce the privacy enhancement of a system design model, since there is a luck of a common a methodology concerning the design of IT systems [24]. More precisely, a systematic model-based cost estimation methodology is proposed that takes into consideration a range of privacy controls, including privacy-design strategies, patterns, and PETs as well as the interrelations and dependencies among them. More recently, the authors in [29] propose a detailed methodology for identifying and quantifying data privacy risks, while the risk values are calculated at two different levels for helping the senior management and the operational personnel to assess and mitigate privacy risks. In addition, the work in [19] present a systematic privacy-related information security risk assessment (pISRA) model, which combines both a privacy impact analysis and a risk assessment. In [30], the authors present an empirically evaluated privacy risk assessment framework, the DPIA Data Wheel. This framework considers the contextual integrity that practitioners can use to take decision around the privacy risks of Cyber Physical Systems (CPS). Unfortunately, most of the research efforts do not implement their proposed method/model.

2.3. The Current Landscape in PIA Frameworks & Tools

When it comes to existing PIA tools, these mainly fall under the efforts that have been conducted by various standardization bodies which have resulted to the following schemes: the ENISA tool [31], the GS1 EPC/RFID PIA Tool [32], the SPIA Tool [33] and the CNIL tool [34]. Most tools on the market have a rather narrow scope of application, with a single use case being the norm. There exist a considerable number of tools supporting the documentation of data processing practices, the formulation of consent templates, or the documentation of privacy and data protection policies. However, the cybersecurity status of the organisation in which the impact analysis is performed is largely neglected.

ENISA tool: ENISA offers an online privacy tool for evaluating the risk level of the personal data processing operations [31]. This tool builds on existing works [26,35,36] in the field and aims to provide guidance to SMEs and support data controllers/processors. The adopted approach consists of six steps towards offering a simplified approach and guide the SMEs to a data processing operation and assist them evaluate the privacy-related security risks. Through the proposed steps, the assessor defines the context of the processing operation and determines manually how the fundamental rights and freedoms of the individuals could be harmed from the possible loss of security of the personal data. Four levels of impact are supported ranging from Low to Very High. In addition the assessor manually documents both the external and internal threats of the environment and assess their threat occurrence probability. After the impact evaluation of the personal data processing operation and the corresponding threat probability, the final evaluation of

risk is delivered. Based on the result, the tool assists the process of adopting new privacy security measures.

GS1 EPC/RFID PIA Tool: The GS1 EPC/RFID PIA Tool [32] focuses on EPC/RFID applications in the context of large corporations and small and medium enterprises (SMEs). The tool assists in the conduction of the assessment of privacy risks of RFID implementations and contributes to the identification of privacy controls to be considered during the development of the applications. The tool is aligned with the European Commission's RFID Recommendations and with EPC Privacy Guidelines. The tool is an MS excel file which assists in the definition of the risk level scores based on the formula $Rik = Impact \times Likelihood - Controls$. The score considers the control effectiveness to determine the residual risk. Through the process the assessor answers specified questions/consideration and can define the arbitrary controls and their effectiveness in the scale of [1–5]. The tool does not focus on identifying technical aspects of the implementations to shed light on privacy risks that can be raised due to actual threat vectors targeting the deployment. In addition, the criteria for scoring are rather generalized and not specific for privacy risks [37], while the assessment is limited to the technology field of EPC/RFID applications.

SPIA Tool: The Security and Privacy Impact Assessment (SPIA) Tool [33] aims to help organisations to conduct PIA, by identify areas of risks and select the suitable strategies and timeframes for the risk mitigation. SPIA focuses on both security and privacy for protecting data with a focus on safeguards and is a tool developed by the University of Pennsylvania. The first version of the tool is an MS excel file, while the second version, SPIA 2.0 Tool shifted to a web-based application. SPIA enables organizations to take probability rankings and threat consequences and automatically score risk into categories of High, Significant, Moderate and Low [38]. The calculated risk score is actually the product of probability and consequence, while the level of probability and consequence are entered in the tool manually. Last but not least, the SPIA Tool is also an adaptable and versatile tool that supports additional threats both security and privacy.

CNIL PIA method and tool: The CNIL tool [34] aims to assist data controllers to perform DPIA based on the methodology published by CNIL in [22,26]. According to CNIL's methodology a PIA is based on two ain aspects:

- fundamental rights and principles, which are "non-negotiable", mandated by law and which must be respected, regardless of the risk nature.
- management of data subjects' privacy risks, which determines the appropriate technical and organisational controls to protect personal data.

The PIA practitioners need to carry out the following necessary steps:

- Define and document the context of the data processing action under consideration.
- Analysis of controls that can protect fundamental principles.
- Assessment and management of privacy risks related to data.
- Formal documentation and validation of the PIA.

The PIA tool assist the practitioners in fulfilling the above-mentioned steps. The result of the assessment is represented through a heat map, where the risks are positioned based on their criticality and the risk likelihood. CNIL tool supports four levels of severity scales; Negligible, Limited, Significant, Maximum. However, defining the processing activities with the engaged actors and data, and the underlined threats, is a manual process which requires considerable effort by the tool user and deep understanding for the current data processing actions of the organisation. As a result, the tool lacks automation that can increase the awareness of the risk assessor, while the cyber security status, i.e., the vulnerabilities of the ICT assets are not reflected in the final scores.

Overall, ENISA's on-line tool [31], consists of six steps for the calculation of the privacy risk. The assessment of risks is the first step towards the adoption of appropriate security measures to protect the personal data. This tool engages the user in a manual process of data entry and does not provide any level of automation. The CNIL's PIA tool [34]

considers data controllers that are familiar with the PIA process. This tool lacks automation, in terms of ICT asset inventory and privacy threat and vulnerability detection, that can increase the awareness of the risk assessor, while the resulted risk levels do not consider the cyber security status of the organization. Moreover CNIL's PIA tool does not use visual representation of the information flows, which is a corner stone of DPIA.

The scoring criteria of SPIA Tool [33] are difficult to be extracted from the available online sources. According to the tool designers, SPIA is able to conduct both cybersecurity and privacy assessment. However, considering that the the tool is offered in a web environment, it is not destined to interact with the assessed infrastructure and support asset inventory and vulnerability detection functionalities. The scoring criteria of S1 EPC/RFID PIA Tool [32] are rather generalized [37]. Moreover, S1 EPC/RFID PIA Tool is simpler than SPIA Tool, but is incompatible with older computers [37]. Last but not least, S1 EPC/RFID PIA Tool has not been well-received or widely-used and the user community has identified the absence of technology-specific guidance regarding both risks and controls [39]. In addition, it has to be noted that S1 EPC/RFID PIA Tool is a domain specific tool that cannot be considered usable to other application domains. Table 1 offers an overview of the key characteristics of the aforementioned tools and highlights the additional aspects that makes APSIA to excel.

Table 1. Comparison of tools.

		Features							
		Asset Interdependencies	**Data Processing Flow visualisation**	**Automation & Dynamicity**	**Cyber Risk Consideration**	**Mitigation Controls**	**GDPR PIA Support**	**Automated Privacy Impact Scoring**	**Legal Framework**
PIA Tools	ENISA Tool	-	-	-	-	✓	✓	-	GDPR
	GS1 EPC/RFID PIA Tool	-	-	-	-	✓	-	-	N/A
	SPIA Tool	-	-	-	✓	✓	-	-	HIPAA
	CNIL Tool	-	-	✓ [1]	-	✓	✓	-	GDPR
	APSIA	✓	✓	✓	✓	✓	✓	✓	GDPR

[1] Offers a level of automation but not dynamicity when it comes to ICT asset management and vulnerabilities detection.

In this context, as aforementioned, APSIA aims to bridge the gap between the cyber- and privacy-risk assessment, which are treated as distinct management processes [18,19]. It tackles one of the current field's shortfalls, which is the lack of a risk scoring system that adequately considers the context and the primary goal, i.e., the privacy preservation, of the environment for identified vulnerabilities, as this can lead the organisations to improperly prioritise their mitigation efforts. To this end, our approach for supporting a PIA, goes beyond a mere documentation of an organisation's data and procedures, but offers a direct connection between the data processing flows and the existing cybersecurity status of the organisation in order to identify the magnitude of privacy risks. The developed scoring system complements the systematic approach for identifying data processing flows based on Inderdependency Graphs, by annotating them with the risk scores of the supporting assets. APSIA is assisted by asset inventory to offer a great level of automation in the process, while the vulnerability detection capabilities contribute to the dynamicity of the tool in detecting new threats. One could argue that, APSIA is only applicable in the context of GDPR privacy impact assessment. Indeed, our method and tool have instantiated based

on GDPR. Nonetheless, APSIA is based on a generic methodology and technical artifacts that can be adjusted to comply with other legal frameworks and standards and we aim to extend our tool in the future towards this direction. Overall, the GDPR requirements, along with the cybersecurity status and the privacy levels of the organisation are offered in a unified manner under the umbrella of the methodology presented in the following section.

3. Towards a Hybrid Risk Assessment Methodology: Unified Security & Privacy IAs

In this section, we elaborate on the methodology followed to treat the cybersecurity and privacy risk assessment in a unified manner and integrate the PIA as part of a dynamic iterative process of infrastructure monitoring and risk assessment. Our approach capitalises on a typical risk assessment workflow, but considers the interdependence of cybersecurity and privacy risks by leveraging the following technology offerings:

- The introduction of interdependency graphs, as a modelling technique and enabler, for capturing and visualising the data flows and supporting assets' interdependecies in the context of an organisation.
- A novel and extensible privacy risk scoring system aiming to quantify the privacy risks imposed by identified vulnerabilities on ICT assets.
- A dynamic and extensible system model that maps core GDPR entities and requirements for assisting the decision makers in keeping track of all needed information and assessing the degree of compliance upon the occurrence of threats.

In what follows, we break down the methodology's steps for regulating the hybrid assessment process of APSIA, while in the following section we elaborate on the aforementioned technology offerings used to realise the methodology. Figure 2 provides an overview of the methodology which consists of widely accepted steps, but aims to facilitate the conduction of the cyber and privacy risk assessment in tandem:

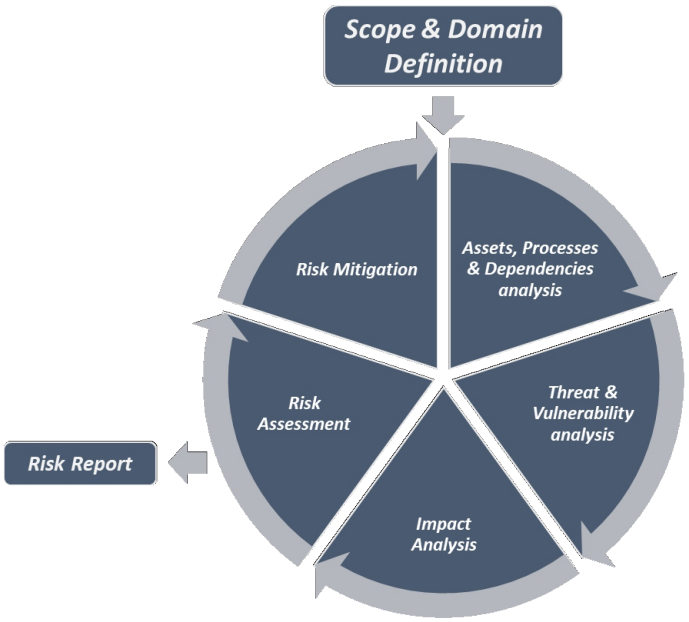

Figure 2. Structural view of the APSIA (Automated Privacy and Security Impact Assessment) methodology.

Step 1: Scope and Domain definition. The focal point of this step is the definition of the type of the assessment and the fragmentation of the specific organisation and its domain (e.g., Healthcare, Energy sectors, etc.). During this task, the definition of the regulatory framework, the standards and guidelines that drive the operation of the domain must be defined. In fact, this action point engages the principal actors of the organisation to

document the identified requirements. Based on the above, the outcome of the analysis of the domain shall be the identification and definition of all the processing activities, the corresponding assets and personnel which will be involved in the assessment. Any organisation and its ICT infrastructure can be seen as a set of linked assets, resources, processes which belong into the sphere of influence of diverse actors who may have different access and control rights. In this direction, it is vital to form a common perception of the environment, where a risk assessor aims to evaluate the cybersecurity and privacy risks considering those multiple dependencies. Thus, this preliminary step aims to clearly define the goals, the scope, and the envisioned outcome of the assessment.

- Scope: Identification of all legacy ICT and domain-specific assets, actors, data subjects, and processes that fall into the scope of the assessment methodology. All the aforementioned entities are treated as tangible and intangible assets in the scope of the APSIA method and tool and define the scope and the boundaries of the assessment.
- Goal: Identification, analysis and assessment of the threats, vulnerabilities, fundamental rights and risks, which are associated to the assets that fall into the scope of the risk assessment.
- Outcome: Evaluation of the cybersecurity and privacy risks, as a consequence of the multiple asset interdependencies, that can lead to proper mitigation actions.

Step 2: Assets, Processes & Dependencies analysis. After defining the scope and breaking down the domain to perform the assessment, the first step of the iterative lifecycle of the assessment is the analysis of the Assets, Processes, along with their interdependencies. The fundamental goal of this step is the identification and modelling of the main cyber or/and physical (controlled/monitored by a cyber system) processes that comprise the organisation, the detection of assets which are involved in, and their dependencies. In this context, the risk assessor should perform the following activities:

- Business Processes and processing activities identification: All cyber processes of interest, along with their data sources, must be defined and recorded in order to be part of the evaluation process.
- Actors Association: The identified actors are linked to the defined Business Processes and the corresponding assets.
- Assets Identification: All assets involved in the identified data processing activities that may be part of the provision of services and their risk level needs to be evaluated, must be identified and reported. It should be noted that the cyber assets can be identified either manually or by using automated tools, e.g., network scanners.
- Assets Interdependencies Identification: Specification and Illustration of the interconnections that exist among the entities and the assets comprising the investigated organisation and domain.

Note that, the actions of this step aim to provide an accurate reflection of the current status of the operational field and the ICT infrastructure of the organisation in order to provide a solid basis to the next steps of the assessment. In order to maintain this accurate reflection, the aforementioned actions shall be executed every time a new assessment cycle should be initiated. The initiation of this process could be triggered periodically or upon the detection of an event like, among others, the detection of a new asset entry, the definition of a new procession activity.

Step 3: Threat & Vulnerability analysis. This step aims to analyse and document the threats and vulnerabilities that may occur against the underlined infrastructure. The probability of the occurrence of a threat is a factor which may vary based on the several factors, such as the nature of the infrastructure per se, the accessibility provided to the targeted assets, among others. Hence, the definition of a threat probability is a subjective matter which is usually undertaken by the security administrator of the infrastructure. In this regard, it is vital to define the threat landscape which is a product of the technologies used, potential vulnerabilities, and the historic log of cyber incidents. Hence, the aim of this step is to identify the set of the applicable individual cyber threats per asset. However,

towards conducting the assessment in a dynamic manner, our appoach capitalises on vulnerability discovery tools that periodically probe the network for identifying vulnerabilities on the assets of the infrastructure. Thus, this step is crucial for the next steps of the utilised methodology, which aim to define the impact and risks based on the dynamically monitored cybersecurity conditions of the infrastructure.

The cyber threat analysis can be based on the following possible sources:

- Domain knowledge of Security or Data Protection officer: These entities usually have the experience to identify applicable and relevant threats to the involved assets. Their insights may also base on historic records of cyber incidents that can reveal weak point of the assets and specific attack patterns against them.
- Existing repositories and threat intelligence platforms: Those repositories (such as NVD [40], CVE [41]) can offer details on existing and new threats found and documented by organisations and companies active in the cyber defence field. In addition, intelligent threat exchange platforms enable the participants to investigate emerging threats in the wild and quickly identify if their endpoints have been compromised in major cyber attacks.

Step 4: Impact Analysis. Following the traditional cyber risk assessment methodology, a threat or a vulnerability exploitation can affect the three security qualities, namely the Confidentiality (C), Integrity (I) and Availability (A) of a targeted asset (also known as the CIA triad).

The impact notion in the field of the cybersecurity risk assessment is a well-known and well-defined notion. However, a crucial research and practical question is how, and to which extend, a vulnerability exploitation may have a cascading impact to the privacy realm. In this direction, our work aims to bridge the gap between the cyber and privacy risk assessment, which are treated as distinct management processes [18,19], and address the cybersecurity and privacy impact assessment under a unified step in the context of our methodology and technology offering. As it will be explained in detail in Sections 4.1 and 4.3 our method capitalises on the interdependency graph model and to an expansible privacy scoring systems for quantifying the privacy impact triggered by identified vulnerabilities of the ICT infrastructure in a dynamic manner.

Step 5: Risk Assessment. This step capitalizes on the outcomes of the previous steps in order to proceed to the final risk estimation. After documenting the assets, the actors, the business processes, the data processing activities, the vulnerabilities, the threats and the associated impact, the risk assessment methodology combines this information to evaluate the overall risk imposed to the infrastructure and trigger the necessary reporting phase. During this phase, especially for the privacy risk assessment, our approach evaluates all the interdependencies identified during the previous phases of the methodology to acquire the dataflows of all the involved personal data (e.g., PII) through the processing activities of the organisation, along with the imposed risk due to the engagement of vulnerable supporting assets.

Step 6: Risk Mitigation. Given the results of the risk assessment processes, Mitigation Controls (MC) can be applied as an action for regulating the risk to the desired level. In this Step, the risk level values are evaluated by the risk assessor and are compared to predefined thresholds, which have been set as requirements during the initial steps on the methodology. In cases where a risk exceeds the desired threshold the infrastructure administrators may proceed to mitigation actions. These actions may vary based on the nature of the assets, the vulnerabilities, and other requirements, such as the cost and complexity of a control adoption. The adoption decision of mitigation controls is undertaken by the security administrator, who has a well-established knowledge of the infrastructure and the domain knowledge.

However, in order to support the decision making of the experts, there have been a plethora of guidelines and standards that define mitigation actions. Several controls that refer to procedural and technical remediation actions can be taken into consideration. In the context of our methodology, since we aim to define the dependence between a cyber

and privacy risk, by extrapolating the former in the privacy field using the privacy scoring system, we mainly focus on mitigation actions that can mitigate the cyber risk that triggers the privacy one. The arsenal of mitigation measures depends on the adopted standards, guideline in the context of the organisation, but they also depend on the domain and infrastructure knowledge of the engaged professional in the process. Our tool utilises as a baseline the CIS controls [42], as a valuable source for identifying mitigation actions that fit to the needs of several business domains. In the context of the risk assessment process, once appropriate mitigation controls are applied for mitigating a vulnerability or a threat on an asset (or a set of assets), the assessment shall be triggered again in order to calculate the residual risk and thus, evaluate the efficacy of the adopted measures.

Note that, the selection of optimal mitigation actions is out of the scope of the proposed methodology and tool. Our aim is to support the decision makers for taking informed decision during the risk mitigation lifecircle.

4. APSIA Conceptual Architecture and Building Blocks

The iterative steps of the APSIA methodology aim to keep the enhanced assessment up-to-date and consider any changes that may (dynamically) occur in the Cyber, Physical and Operational environment of the target ecosystem. APSIA is capable to support and complement PIA procedures and to enhance the assessment with dynamic asset inventory and vulnerability discovery techniques in order to feed the assessment steps dynamically. Figure 3 presents the high-level architecture of APSIA, which highlights the separation between the physical environment and the APSIA operational environment, while intuitively reflects the interconnections that enable the flow of information to the building blocks of APSIA.

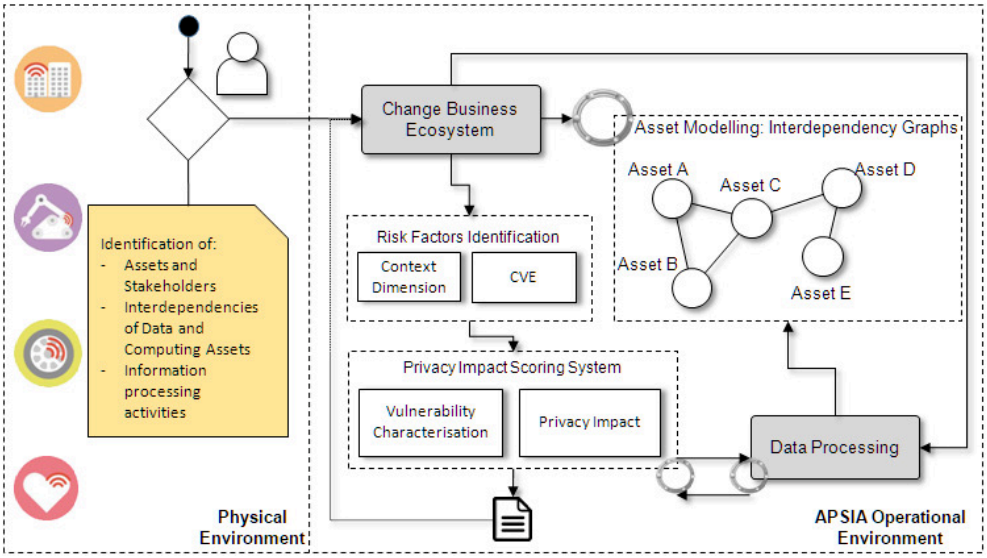

Figure 3. APSIA Conceptual Architecture.

Physical Environment: APSIA takes advantage of information regarding the tangible and intangible assets (along with their interdependencies), the different actors, threats and vulnerabilities, and the organisation's processing activities. The discovery of ICT assets and vulnerabilities is achieved through the use of inventory tools (OpenVAS), while the documentation of the processing activities is part of the initial configuration of APSIA by the CISO or DPO of an organisation, who is aware of the data and processes that should be in the scope of the assessment.

APSIA Operational Environment: The APSIA environment incorporates the necessary building blocks for performing the assessment. More specifically, the internal modeling components of APSIA generate the interdependency graph and identify the chain of assets included in the processing activities. The interdependency graphs assist the security analyst,

to identify potential privacy risks based on a cartography of assets, which encapsulate their vulnerabilities and the potential privacy threats posed against them. This graph is updated every time a new assessment is conducted in order to include possible updates from the assessed ecosystem. The interdependency graphs are also used for constructing the data processing flows, which are formed through the asset paths and data sources defined by the assessor. After defining the aforementioned dependencies and having an updated reflection of the assessed environment, the Privacy scoring system undertakes the quantification of the privacy threats. To do so, the scoring system is based on risk factors of existing repositories, such as CVE, but can be extended by other context dimensions.

The following sections elaborate on each of the APSIA building blocks by providing more details on their structure and functional behavior.

4.1. Interdependency Graphs for Data and Asset Modelling

In order to meet the requirement of a cyber and privacy risk assessment methodology, our work capitalises on a graphical representation, namely the interdependency graphs [43,44]. This particular representation offers the necessary flexibility to define relations among large number of objects of arbitrary types, and thus, provides a model that can be adapted to the needs of the risk assessor of virtually any kind of organisation. This graphical representation model is a cornerstone in our work, as it works as the "glue" that holds together ICT assets, data entries, threats and vulnerabilities in order to identify risky data processing activities of an organization.

Our densely interconnected world is based on the provision of services that may form rather complex flows of data processing activities, which engage actors and supporting assets which may be distributed, not only on a network basis, but also across different physical locations. Thus, keeping track of these workflows implies the need of a modelling techniques that can meet the challenges introduced in modern infrastructures that store, process and require the exchange of data. This mesh of interrelations results in a very sensitive network of critical entities, where the various types of interdependencies have to be identified in the context an assessment.

The interdependency graph model allows physical, cyber, and human elements to be combined, including combinations of legacy systems and new technologies, and data. More precisely, in order to be able to have a global view of the infrastructure and the workflows of data processing and be in position to detect possible cascading and escalating effects of threats against users' privacy, it is crucial to maintain a cartography of all the dependent assets. In the model, Nodes are used to represent the individual assets and edges to represent the interdependencies amongst them. In [45] four general types of interdependencies have been identified: physical, cyber, geographical, and logical. Depending on the granularity of the performed analysis, some of these types can be omitted. In our work we use the interdependency types *IsConnectedTo*, *IsUsedBy*, *IsProcessedBy*, *isLocatedIn*, *isStoredOn* and *IsInstalledOn* to annotate the relation among assets. These relations are not only used to denote connections among tangible ICT assets, but also intangible ones, such as data, health records and PIIs. The IsConnectedTo and IsInstalledOn represents network and systemic inter-dependencies, the IsUsedBy and isLocatedIn represent physical inter-dependencies, the IsProcessedBy and isStoredOn represent logical inter-dependencies. Each type/dependency is represented by a different edge arrow.

Overall, by utilizing the interdependency graphs, the risk assessor is in position to identify on-the-fly potential privacy risks based on a cartography of assets, which encapsulate their vulnerabilities and the potential privacy threats posed against them. In this way, the graph model contributes, not only to the uncovering of privacy risky individual assets, but crucially, it ease in highlighting privacy risky paths which are formed by chains of assets included in a specific processing activity. An example of an interdependency graph is illustrated in Figure 4.

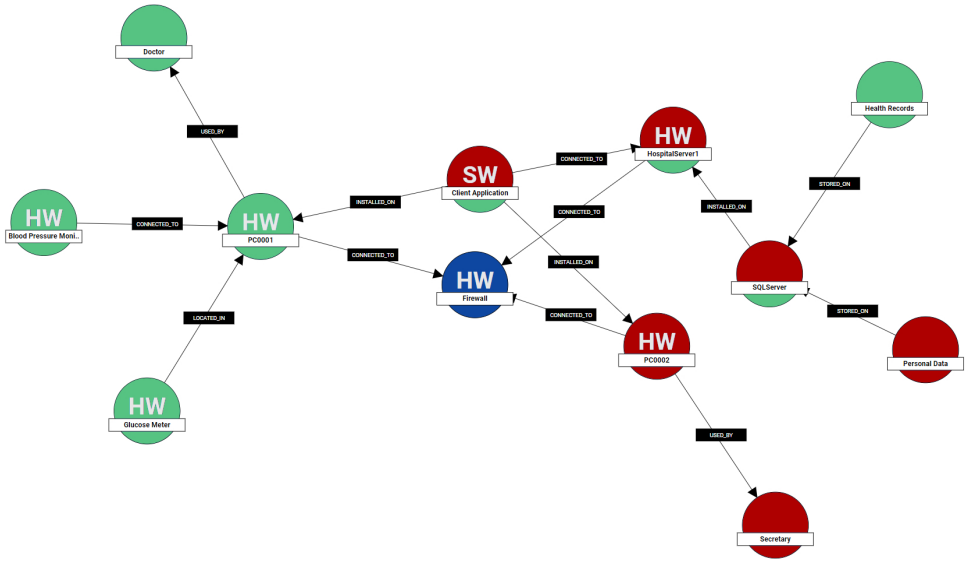

Figure 4. Asset interdependency graph with the processing activities highlighted.

Hence, our modelling techniques is a cornerstone that supports the methodology adopted in this work, while the key functionalities that this component is offering are:

- The asset representation, along with the interdependencies that connect both tangible (e.g., a Database) and intangible assets (e.g., health records).
- Constitutes the steppingstone for defining the Processing Activities, since each Processing Activity is represented as a chain of supporting assets.
- Gives the ability to detect, in how many processing activities a vulnerable asset is engaged in order to assist the privacy scoring system (see Section 4.3) to quantify the imposed privacy impact score.
- Last but not least, works as the connection point between the privacy impact scoring system and the Data processing flows (see Section 4.2), as the GDPR data processing activities inherit the privacy risk levels of their supporting ICT assets.

4.2. Data Processing Flows in the Frame of GDPR

The Processing Activity is a principal aspect of the GDPR and aggregates all the GDPR-related information. To form the processing activities, we utilise a dynamic and extensible model which is able to maintain a map among the core GDPR entities and requirements, and the processed data. As mentioned in the previous section, the formation of the Processing Activities is based on the asset modelling, which maintains a representation of the tangible and intangible asset chains. The exact GDPR entities and the formation of the processing activities is part of the initial configuration, which is undertaken by the security analysts or the risk assessor (e.g., CISO and/or DPO) of the organisation.

Hence, the Processing Activity is the key aspect of the GDPR modelling that enables the assessor to understand how data flow among the organisation's processes and which entities are engaged in it. The main information that a Processing Activity includes can be divided in three parts: (a) the processing purpose along with the involved entities, (b) all the processed data assigned to specific Subjects, and (c) the asset chain that is involved in the processing activity. The IsProcessedBy interdependency introduced in the previous section enables the functionality of forming the Processing Activities as a chain of assets. Processing Activities are an important part for the overall functionality, as the privacy risk calculation formula considers the scope of impact factor (see Section 4.3.2) to calculate the privacy score.

Indicatively, in a Processing Activity, among other attributes, we define the data subject, data controller, the data processor, the legal grounds, the data recipients, the processing country, processing purpose and lawfulness of processing, all the involved personal data (e.g., PII such as the name, surname, etc.), and all the assets that are included

in this processing activity. Those attributes are combined together through the use of interdependency graphs.

4.3. Privacy Impact Scoring System

As mentioned in the introduction of this work, the current privacy risk assessment methods and tools perform the risk assessment by ignoring the cybersecurity state of the underlined infrastructure. In fact, the reported tools in Section 2 base the assessment on documenting the organisation's procedures and the definition of the final risk depends completely on the perception of the administrator. In fact, one of the current scoring systems shortfalls is the lack of a risk scoring system that adequately considers the context of the environment for identified vulnerabilities [46]. This can lead the organisations to improperly prioritise their mitigation efforts. In the frame of our method and tool, the context is the data subject's privacy and our aim is to provide a formula which considers the peculiarities of the organisation's data processing procedures along with its cyber security state.

Current methods for analysing identified cyber vulnerabilities of traditional information technology systems tend to focus on the impact to systems' CIA triad to discern end-user risks. Although this approach is sufficient for evaluating traditional information technology systems, it fails to consider the operational ramifications for complex systems-of-systems. Thus, there is a need for a Risk Scoring System that provides the means to characterize identified vulnerabilities and numerically score the effect that a potential exploitation of a vulnerability may have on users' privacy.

To address this limitation, approach quantifies the privacy risk based on a generic and extensible scoring system, which takes into consideration, not only the vulnerability score per se, but also the potential impact to patient's privacy. This scoring system will be used to measure the privacy impact of the data processing activities of an organisation, given the vulnerabilities of the supporting assets. Given that GDPR compliance is necessary for organisations, this scoring system can contribute towards measuring the degree of GDPR compliance.

The developed scoring system incorporates two primary factors, which are used to calculate a Total Score: (i) Vulnerability Characterization, and (ii) Privacy Impact. On the one hand, the vulnerability characterization refers to the details and the severity of the vulnerability, which in turn may threaten a user's privacy. On the other, the Privacy impact is based on three contributing factors, namely:

a the level of impact on the fundamental rights and freedoms of the individuals
b the scope of impact to the data processing activities
c the type (i.e., sensitivity) of the processed data

By combining these two elements, the scoring system provides the mean for reflecting the severity of an identified vulnerability in the context of users' privacy.

In the context of our work, the vulnerability score is based on the Common Vulnerability Scoring System (CVSS) [47], where a value from 0 to 10 is assigned to the identified vulnerability, following the system's scoring formulas. However, the determination of the privacy impact is based on the aforementioned contributing factors, which will be explained in detail in Section 4.3.2.

The final score and the contributing factors follow the structure depicted in Figure 5. While the scoring algorithm combines these two factors, it treats the impact as the leading factor for the final assessment. Indeed, it uses a weighted scale to focus on the impact to users' privacy, while incorporating the vulnerability score. The scores are evaluated on a 0 to 10 scale, with higher numbers indicating more severe ratings. It must be stated that the exact value of the weights is a parameter that can be adjusted accordingly, given the preferences and the domain knowledge of experts of the organisation. The weighted scale formula is given in Equation (1).

$$Privacy\ Score = (Vulnerability\ Characterization + 2 \times Privacy\ Impact)/3 \quad (1)$$

4.3.1. Vulnerability Characterisation

A vulnerability may appear either as a technical limitation of a system, or a gap in the security procedures and practices of an organisation. We base our method on dynamically detecting vulnerabilities by deploying the OpenVAS Vulnerability Assessment Scanner [48]. In this way, we create a dependence between the infrastructure's asset and their vulnerabilities. The most prominent way of quantifying the severity and the impact that a vulnerability may have on a targeting system, is the adoption of a common and vendor-independent scoring system enables the security society, and especially the risk assessors, to form a common understanding for the criticality level of vulnerabilities. In the context of distributed organisations, there is a need for privacy assessment tool to be deployed on different sites, and thus, the use of an open, independent and globally accepted scoring system is of significant importance in forming a common risk perception among the involved parties. The utilisation of CVSS contributes to the design of a unified risk assessment methodology that leaves aside the diversity of the engaged assets. In fact, depending on the domain, an organisation may utilise a wide range of legacy ICT and domain-specific devices (e.g., IoT, Industrial, Energy, healthcare devices) and applications. Inevitably, such a setup results to an extended attack surface (more vulnerabilities) against a significant amount of data processing activities of Personally Identifiable Information (PII), and thus, an increased possibility of data leakage.

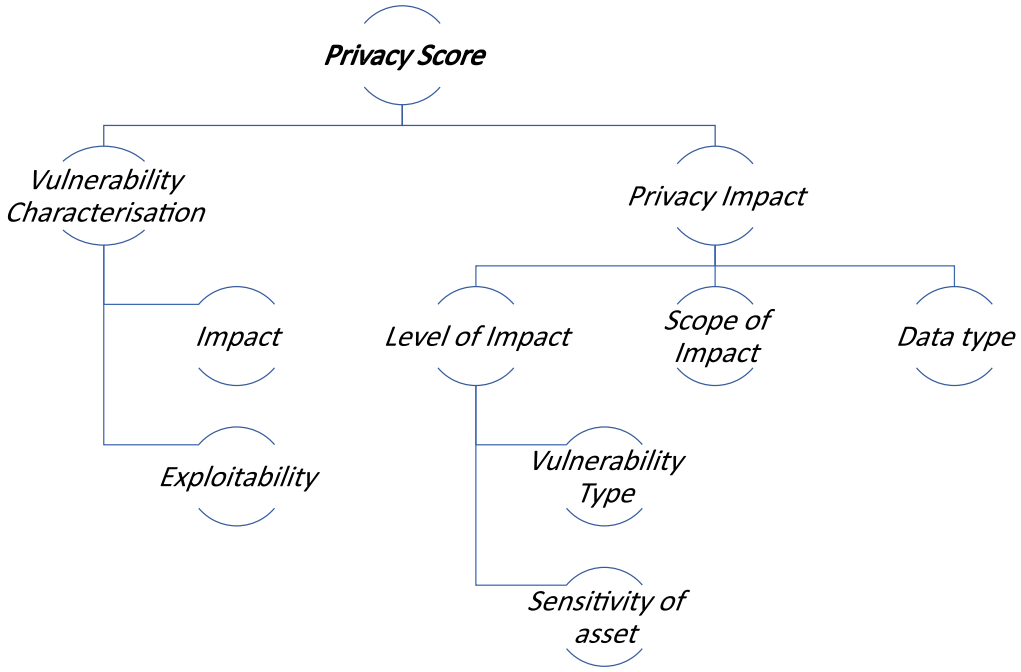

Figure 5. Structural view of Privacy score with the contributing factors.

Last but not least, the adoption of CVSS enables the risk calculations can be compatible both for vulnerabilities of old and legacy ICT assets, as for new ones, and are in-line with global vulnerability repositories. It has to be stated that this is a requirement that must be satisfied by the developed privacy assessment methodology. Crucially, the adoption of CVSS guarantees that all the developed components and automated tools used for threat and vulnerability detection have a common reference point that enhances their interoperability towards the dynamic conduction of the privacy impact assessment.

Given the above, the vulnerability characterization aims to measure the severity of the vulnerability under the traditional risk assessment perspective. Note that the

vulnerability characterization is one of the contributing factors of the Score Equation (1) and is responsible for reflecting the cyber security status of the assets in the organisation.

$$Vulnerability\ Characterisation = (0.6 \times Impact + 0.4 \times Exploitability - 1.5) \times f(impact) \quad (2)$$

$$Impact = 10.41 \times (1 - (1 - ConfImpact) \times (1 - IntegImpact) \times (1 - AvailImpcat))$$

$$Exploitability = 20 \times AccessComplexity \times Authentication \times AccessVector$$

$$f(impact) = \begin{cases} 0 & \text{impact = 0} \\ 1.176 & \text{otherwise} \end{cases}$$

4.3.2. Privacy Impact

The Privacy impact reflects the consequences an exploited vulnerability may have on a data subject's privacy. As have been already stated, one of the current scoring systems shortfalls is the lack of a risk scoring system that adequately considers the context of the environment for identified vulnerabilities. Thus, the utilisation of a scoring system that focuses exclusively on the impact on CIA metrics, fails to consider the operational ramifications imposed to the affected organisation. To address this issue, the proposed scoring system incorporates the privacy impact as contributing factor, to extrapolate the vulnerability exploitation impact to the privacy dimension. The privacy impact itself consists three components, namely, (a) the level of impact on the fundamental rights and freedoms of the individuals, (b) the scope of impact to the data processing activities, and (c) the type (i.e., sensitivity) of the processed data. The aforementioned components and their contribution in the scoring formula are described in the following sections.

$$Privacy\ Impact = Level\ of\ Impact + Scope\ of\ Impact + Data\ type \quad (3)$$

A. Level of Impact

The level of impact aims to assess the impact on the fundamental rights and freedoms of the individuals, resulting from the possible loss of security of the personal data. Four levels of impact are considered (Low, Medium, High, Very High) as shown in Table 2, following the taxonomy proposed by ENISA for the personal data processing [49] considering also the case where there is no impact to the rights and freedoms of the individual. As can be inferred by the taxonomy proposed by ENISA, the lowest level of impact considers minor consequences on individuals, while at the highest level, the affected individuals may suffer significant of even irreversible consequences. Although this taxonomy has been adopted on ENISA's tool [31] for evaluating the level of risk of personal data processing operations, the determination of the appropriate value relies explicitly on the situational awareness and the domain experience of the assessor. In the context of the scoring formula of this work, the exact level of impact is determined based on a systematic approach which considers the type of the vulnerability which targets an asset, as well as the importance of the latter in data processing activities of the organisation. In this way, the scoring system considers the nature of the vulnerability which targets an asset and the privacy-oriented business value of it.

Vulnerability type: In order to identify the type of the vulnerability, we adopted the taxonomy used by CVEdetails [50] online vulnerability database. CVEdetails provides an easy to access interface to CVE vulnerability data. Vulnerabilities are categorised based on vendors, products, and versions. CVE vulnerability data are taken from National Vulnerability Database (NVD) [40] xml feeds provided by NIST. Additional data from several sources like exploits from exploit-db [51], vendor statements and additional vendor supplied data, Metasploit modules are also published in addition to NVD CVE data. Hence, CVEdetails is a database which enriches the basic CVEs with additional metadata and offers a classification that reveals 13 type of the vulnerabilities, as can be seen below.

Table 2. Impact categories on the fundamental rights and freedoms of the individuals, resulting from the possible loss of security of the personal data according to ENISA (European Union Agency for Cybersecurity).

Level of Impact	Description	Value
None	Individual will not encounter inconveniences or consequences	0.0
Low	Individuals may encounter a few minor inconveniences, which they will overcome without any problem (time spent re-entering information, annoyances, irritations, etc.).	1.0
Medium	Individuals may encounter significant inconveniences, which they will be able to overcome despite a few difficulties (extra costs, denial of access to business services, fear, lack of understanding, stress, minor physical ailments, etc.).	2.5
High	Individuals may encounter significant consequences, which they should be able to overcome albeit with serious difficulties (misappropriation of funds, blacklisting by financial institutions, property damage, loss of employment, subpoena, worsening of health, etc.).	4.5
Very High	Individuals which may encounter significant, or even irreversible consequences, which they may not overcome (inability to work, long-term psychological or physical ailments, death, etc.).	7.0

The type of a vulnerability is a parameter of significant importance that enhances the situational awareness of a security defender when it comes to the prioritisation of actions for mitigating cyber risks. The same applies in the case where the defender must make decisions considering the data protection and the user privacy. For instance, vulnerabilities of the Gain Information or the SQL Injection categories, can have a greater impact on users' privacy in contrast to a Denial of Service, which mainly affect the availability of a source. Given the above, the proposed scoring system takes into consideration the type of the vulnerability in order to convey this information to the final score, which will be used by the decision maker in the privacy risk mitigation actions.

- Denial of Service
- Bypass Something
- Execute Code
- Gain Information
- Overflow
- Gain Privilege
- SQL Injection
- XSS
- Directory Traversal
- File Inclusion
- Memory Corruption
- CSRF
- Http Response Splitting

Sensitivity of ICT asset: In order to define the level of impact in a more reliable way,the privacy sensitivity of an asset is considered. More specifically, as the infrastructure assets are engaged in the data processing activities of an organisation, undoubtedly some may have a more crucial role in the data processing contrary to others that simply support the activity. For example, a central database which is used to store the Personal or Sensitive information of a hospital's patients, is of greater importance -in terms of privacy- than an ICT network component. To materialise this, we use a 4-tier scale to categorise the assets into the Low, Medium, High, Very High tiers. Given the above, the scoring system takes into consideration the importance of the assets in order to convey this information to the final score, which will be used by the decision maker in the privacy risk mitigation actions. The definition of the level of impact factor is the product of the two above-mentioned notions namely, the type of the vulnerability and the sensitivity of the assets, following the mapping illustrated in Table 3. It must be stated that the mapping has been generated based on the domain knowledge of cybersecurity experts and penetration testers of our corporate environment, who have the necessary experience to perceive how a specific type of a vulnerability can affect privacy-sensitive assets. Although the proposed mapping conveys the domain knowledge of experts, it cannot be considered foolproof, as security experts of different expertise and different organisations may have a different point of view on the mapping. The exact definition of the map is part of the configuration steps of the tool and depend also on the cybersecurity posture and risk appetite of the assessor.

Table 3. Definition of the level of impact factor as a product ofthe type of a vulnerability and the sensitivity of an asset.

	Vulnerability Type												
Asset sensitivity		**DoS**	**Code Execution**	**Overflow**	**Memory Corruption**	**SQL Injection**	**XSS**	**Directory Traversal**	**Http Response**	**Bypass Something**	**Gain Information**	**Gain Privileges**	**CSRF**	**File Inclusion**
	Low	L	M	L	L	M	L	L	L	L	M	L	L	L
	Medium	L	M	M	M	M	M	L	L	M	M	M	M	M
	High	L	VH	H	M	VH	H	M	M	H	VH	VH	H	H
	Very High	L	VH	H	H	VH	VH	VH	VH	VH	VH	VH	VH	VH

B. Scope of impact

The scope of impact is used to reflect the number of data processing activities affected by an instance of exploiting the vulnerability. The data processing activities may consist of several supporting assets which are used to process, store, and visualise the data. However, those assets may have vulnerabilities which can impose a risk to the processing activities in which they are engaged. In this direction, the scoring algorithm considers the scope of impact factor in order to reflect the severity that the vulnerability exploitation may have to the dependent processing activities. Hence, three options have been identified based on the impact values, namely Single ($value \leq 0.5$), Multi ($0.5 < value \leq 1.0$), and All ($1.0 < value \leq 1.5$). Following this approach, the scoring system takes into consideration the dependence between the vulnerable assets and the data processing activities of the organisation and conveys this information to the final score, which will be used by the decision maker in the privacy risk mitigation actions.

C. Data types

Information systems may store and process a huge amount of data. However, the criticality of the data is not always the same. For instance, some processing activities may focus on publicly available data, others on financial data, and other personal or even sensitive data. This variation indicates the need to assign a different criticality levels to the aforementioned data types and treat personal and sensitive data, as data types that can clearly have a higher impact on the fundamental rights and freedoms of the individuals in case of data breaches [52]. To materialize this in the proposed method, the data that an organisation stores/processes, are classified in the following categories, following the classification proposed in [53].

- Sensitive personal data (medical data, legal documents, etc.)
- Personal data (data which uniquely identify a person, such as IDs, Social Security Number (SSN), personal or marital status, etc.)
- Financial data (data related to financial transactions, accounting entries, etc.)
- Operational data (data generated during the execution of a service, log files, etc.)
- Other data (data that cannot be classified in any of the above categories, and belong to a lower criticality level)

Thus, each data entry, which is part of a processing activity, falls into one these categories and the privacy scoring algorithm considers the criticality of the processed data. Table 4 provides the mapping between the data type and the corresponding quantitative value. Given the above, this property is reflected to the final score, which will be used by the decision maker in the privacy risk mitigation actions.

Table 4. Data Type values.

Data Type	Value
Sensitive Personal Data	1.5
Personal Data	1.0
Financial Data	0.75
Operational Data	0.5
Other Data	0

4.3.3. Privacy Scores of Processing Activities

Given the methodology steps described in Section 3 and the privacy impact scoring of Section 4.3, one can infer that based on the inderdependencies of the tangible and intangible assets of an organisation, a vulnerable asset may trigger a privacy threat. Thus, the privacy scoring formula given in Equation (1) is calculated per vulnerability and per asset. Hence, the scoring approach produces as many privacy scores as the existing vulnerabilities on the assets. This is reasonable, as each vulnerability may trigger a different privacy impact.

Having said that, overall we define 3 different types of scores:

- the asset-level privacy score (APS)
- the processing activity-level privacy score (PAPS)
- the organisation-level (global) privacy score (OPS)

All three scores range between 0 and 10. The corresponding qualitative value is mapped to a 5-tier scale values ranging from "Very Low" to "Very High", as can be seen in Table 5.

Table 5. Five-tier scale values for privacy score.

Privacy Score Range	Qualitative Value
0.0–2.0	Very Low
2.0–4.0	Low
4.0–6.0	Medium
6.0–8.0	High
8.0–10	Very High

Asset-level privacy privacy score (APS): An asset-level privacy score is assigned to each asset, associated with a vulnerability, following the scoring system presented in Section 4.3. It must be noted that in the case where an asset is affected by multiple vulnerabilities, the highest value (max) among the Asset-level privacy scores will be assigned to the asset. The privacy score is calculated following Equation (1).

Processing activity-level privacy score (PAPS): Each processing activity is represented as a chain of interrelated assets (supporting assets). Thus, the highest Asset-level privacy score among the assets of the processing activity represents the privacy risk of the activity. For instance, the privacy score assigned to the Processing Activity i that contains n assets in its asset chain is calculated based on the following formula:

$$PAPS_i = max(APS_1, APS_2, \ldots APS_n) \quad (4)$$

Organisation-level (Global) privacy score (OPS): The global privacy score is the highest Asset-level privacy score among all the ICT assets of the infrastructure of the organisation. Thus, the highest Asset-level privacy score among the assets of the organisation, and in turn among all the processing activities, represents the global privacy score of an organisation. For instance, the privacy score assigned to the organisation that contains k assets in its infrastructure is formalised as follows:

$$OPS = max(APS_1, APS_2, \ldots APS_k) \quad (5)$$

The selection of the highest score to be represented in the APS, PAPS and OPS scores aims to simplify the assessment process, especially in the cases where a great number of assets and processing activities is engaged in an organisation. In addition, this quick view can be beneficial in cases where the lifecycle of the methodology described in Section 3, is triggered periodically or upon the detection of events that can trigger the assessment process in a dynamic manner (e.g., detection of new vulnerability, new asset entry in a processing activity, new device in the topology).

However, in order not to miss the holistic view of the threats and their corresponding risks, in case where multiple risks exist for an assets, a histogram is generated to complement the analysis and represent the distributional characteristics of the risks for the assets. In this way, the histogram balances the narrow view of focusing solely to most severe privacy impact. Section 5 elaborates on the aforementioned points in the frame of a case study of a healthcare organisation.

5. Case Study for the Healthcare Sector

This section demonstrates the applicability of the APSIA methodology and tool in a case study of a hospital. The case study considers security and privacy concerns against the hospital's ICT infrastructure that threaten processing activities. The case study is based on actual requirements driven from a real use case in the context of the H2020 CUREX project [54]. More specifically, during the testing process of APSIA we approached a set of healthcare related stakeholders, ranging from the IT support team and the security administrator of a hospital to physicians, for acquiring information regarding the actual assets, medical devices and processes which shall be considered in a healthcare environment. Based on this information, we fleshed out the actual topology of ICT assets and applications along with their interdependencies, and we defined the data and the processing activities. The gathered information was used to create an accurate virtual topology in order to simulate in a lab environment the existence of vulnerable devices that support the hospital's data processing activities and may trigger privacy risks. APSIA deployed in a virtual machine to ensure network visibility and enable asset inventory and vulnerability detection. The generated setup of all the aforementioned aspects of the virtual environment and the visualisation results shown in Appendix A, were then ratified by the board of relevant professionals of the CUREX project, who identified the value of APSIA.

More specifically, the IT infrastructure of the hospital facilitates the deployment of Healthcare Point of Care (POC) technologies. POC have been widely used during the last decade to pave the way to the emergence of healthcare monitoring and management. POC technologies are hospital information systems that includes terminals or other devices for medical diagnostic testing at or near the site of patient care [55]. The advances of POC have enabled patients to receive better care. However, along with these advances, there are concerns regarding the connectivity of these devices to the Internet or to a Health Information System (HIS), since it may affect both the security and the privacy of a patient. Even though, the connected medical devices improve the quality of patients' care, they also expose a wide attack surface and introduce new and domain-specific vulnerabilities. In addition, the rising number of the processed personal data in conjunction with the increasing data breaches has led to an attention towards data protection and privacy. Hospitals and care centers need to address these challenges by efficiently assessing the privacy risks in tandem with GDPR that mandates such an assessment. To this end, in our scenario we consider a subset of common assets and devices that exist in a hospital's POC to demonstrate the APSIA methodology and tool.

5.1. Scenario Overview

Patients Patient1, Patient2 with diabetes want to perform a regular check-up. Patient1 is visiting the Hospital for the first time and, hence, a registration process is initiated for collecting personal information. After the registration phase, a Doctor monitors their blood pressure and glucose considering their clinical history of diabetes. Two common processing

activities are identified in the aforementioned scenario (a) Patient Registration and (b) Patient Monitoring. In the former processing activity, we assume patient's registration to the hospital, where the hospital's secretary Secretary1 inserts into the HIS details such as their full name, their contact details etc. In the latter processing activity, we assume the monitoring and storing of the patient's blood pressure and glucose measurements to the hospital's HIS.

Each processing activity includes a chain of assets considering the interdependencies among them. The hospital's CISO or DPO, periodically performs internal interviews, as a part of her role, in order to confirm the GDPR compliance levels of the organisation. In this sense, she is already aware of the processing activities that should be documented in order to keep track of how sensitive or personal data flow among the various entities and supporting assets. Overall, by utilising the interdependency graphs, a security analyst of the hospital, will be in position to identify potential privacy risks based on a cartography of assets, which encapsulate their vulnerabilities and the potential privacy threats posed against them.

5.2. Assessment Results

Interdependency Graph: Ten tangible assets and three intangible (e.g., data and personnel) are included with different connections among them. Figure 4 illustrates the interdependency graph of the scenario. In the Patient Registration processing activity the engaged assets are the Client Application, PC002, Secretary, HospitalServer1, the SQLServer and the Personal Data (red color), while in the Patient Monitoring the included assets are the Glucose Meter, Blood Pressure Monitor, Doctor, PC001, Client Application, HospitalServer1, SQLServer and Health Records (green color). As can be seen, the Health Records and Personal Data nodes isStoredOn on the SQLServer. The latter IsInstalledOn the HospitalServer1, while a ClientApplication IsConnectedTo to the same server. The Doctor interacts with the PC0001 based on the physical interaction IsUsedBy.

It becomes obvious that several interconnections can be defined as a result of the actual dependencies of cyber assets, data sources and actors. Especially for large scale or dynamic environments, our approach can be proved beneficial as it offers a cartography that can assist the assessor to understand how the interconnected assets facilitate the data flows of the procession activities. In this way, the interdependency graphs contribute, not only to the uncovering of privacy risks on individual assets, but crucially, they ease in highlighting privacy risky paths which are formed by chains of assets.

Privacy scoring system: As aforementioned, our proposed asset-centric scoring system incorporates the privacy impact as contributing factor for quantifying the impact that a vulnerability may have on user's privacy. Table 6, provides an overview of the vulnerabilities of core assets which are engaged in the processing activities of the organisation. As can be seen, independently of the vulnerability characterisation score, which reflects the cybersecurity criticality against the CIA triad, based on the introduced privacy scoring system the risk assessor can have the impact reflection of this cyber threat to the privacy dimension.

Considering the peculiarities of each case, i.e., the sensitivity of the affected asset, the vulnerability type, the type of processed data and the number of the affected processing activities (scope of impact), our scoring system calculates the Privacy Impact score and the Privacy score. In this way, in the context of a privacy impact assessment, where mitigation actions should be driven by setting the privacy preservation as the main goal to achieve, the assessor can take advantage of the privacy impact and score to clearly identify the potential privacy risks and prioritise the mitigation actions accordingly.

Table 6. Use case threats, assets and privacy scores association.

ID	Asset	Vulnerability	Vulnerab. Charact. [1]	Asset Sensit.	Vuln. Type	Level of Impact	Scope of Impact	Data Type	Privacy Impact [2]	Privacy Score [3]
ID-01	SQLServer	CVE-2020-11898	6.4	VH	Gain Info	VH (7.0)	All (1.5)	Sensitive (1.5)	VH (10)	VH (8.8)
ID-02	Glucose Meter	CVE-2019-10964	5.8	H	Bypass Auth.	H (4.5)	Single (0.5)	Sensitive (1.5)	H (6.5)	H (6.3)
ID-03	Blood P. Mon	CVE-2017-11579	4.8	H	Gain Info	VH (7.0)	Single (0.5)	Sensitive (1.5)	VH (9)	H (7.6)
ID-04	PC0001	CVE-2017-13993	9.3	H	Code Exec.	VH (7.0)	Single (0.5)	Sensitive (1.5)	VH (9)	VH (9.1)
ID-05	PC0001	CVE-2019-1343	7.1	H	DoS	L (1.0)	Single (0.5)	Sensitive (1.5)	L (3)	M (4.4)
ID-06	PC0002	CVE-2019-5831	6.8	H	DoS	L (1.0)	Single (0.5)	Personal (1)	L (2.5)	L (3.9)

[1] Following CVSS v2 and Equation (2). [2] Following Equation (3). [3] Following Equation (1).

By taking a closer look at the scores of Table 6, one can notice that for ID-01, the SQLServer can be affected by a vulnerability that can lead to leak of information. Such a type of threat against a sensitive asset (in terms of privacy) can have a major impact to data subjects' privacy. However, if the assessor prioritise the mitigation actions following solely the CVSS score (Vulnerability characterisation), the ID-05 case, would look more severe, even thought the ID-01 case has a direct impact of users privacy via the information leakage. In fact, the same applies for the case of ID-06. In consideration of the potential impact that CVE-2020-11898 may have on patients' privacy, our risk scoring system considers the severity of the vulnerability, but via the weighted scale of Equation (1) the final score focuses on the privacy aspect. Hence, the scoring system addresses the need of considering the operational ramifications posed by the domain and quantifies the privacy risk based on an extensible scoring system. The same applies in the cases of ID-02 and ID-03, where the medical devices have been found vulnerable to Authentication Bypass and Gain Information attacks. Again in both cases, the final privacy score results to a higher score than the vulnerability characterisation and assists the assessor to notice the privacy risk from the pertinent point of view. In the case of ID-04, the code execution vulnerability of PC0001, leads to an almost equal privacy score as the vulnerability characterisation.

Data processing flows in the frame of GDPR: As we analysed thoroughly in Section 4.2, one of the main contributions of APSIA is the formation of the data processing flows in the context of an organisation following the principles of GDPR. Thus, given the use case scenario and the Patient Registration and Patient Monitoring processing activities, Figures A1–A3 are generated automatically by considering the interconnections of the interdependency graph of Figure 4.

More specifically, Figure A1 creates a flow that maps the PII, Data subjects, (i.e., Patient1, Patient2), Purpose of processing and the Lawfulness of Processing for revealing what type of data are processed in the organisation's processing activities and under which conditions and legal grounds. Note that, this view can be modified in order to provide a mapping among various GDPR entities and requirements in order to give the necessary flexibility to the assessor to keep track of particular flows of interest.

Moreover, Figures A2 and A3 present the data flows visualisation of the two identified processing activities. Each data flow consists of the PII, Data Subjects, Processing Activities, the supporting assets and the corresponding Privacy Risks. This type of figures enables the assessor to associate the identified privacy threats, based on the privacy scoring system, with the assets that trigger the threat, the data and the individuals which are being threaten. This graph enables the assessor not only to prioritise the mitigation actions for specific assets, but crucially, to keep track of the data of subjects that may face an impact on their fundamental rights and freedoms.

Taking a closer look at Figure A2, the Patient Registration activity and the corresponding data subjects can be affected by vulnerabilities identified in assets PC0002 and SQLServer. The total risk score assigned to the specific process is "Very High" and overshadows the "Low" risk of PC0002, following the approach described in Section 4.3.3. The same approach is illustrated in Figure A3 for the Patient Monitoring activity. In this case, more supporting assets bring threats against the data processing activity. In the case of the PC0001, which was found vulnerable to DoS and Code Execution attacks, the latter privacy risk overshadows the "Medium" risk of the DoS vulnerability, following the approach described in Section 4.3.3.

Apart from the data flow graphs, APSIA offers a set of adjustable visualisation tools that enhance the risk assessment process. Indicatively, Figure A4 presents the heat-maps and histograms with information regarding both the privacy and the cyber risks in order to provide an holistic view of the privacy and cybersecurity levels of the organisation. More specifically the Threat Probability–Vulnerability Impact heat-map provides an overview of the cybersecurity risk assessment based on the CVSS scores given in Table 6. The privacy impact assessment process is supported by the heat-map in Figure A4, where the correlation of the Privacy Impact–Vulnerability Characterisation is given. As can be seen, based on the

scores of Table 6, there are risk associations that reveal a "Very High" privacy impact even if the vulnerability characterisation is "Medium". By utilising this view, the risk assessor can properly prioritise the mitigation actions based on the heat of the privacy impact axis.

In addition, in an effort to provide an holistic view of the Privacy scoring results of the organisation, the histogram in Figure A4 presents the distributional characteristics of the privacy scores for all the processing activities of the organisation. This view complements the data flows in which some of the identified privacy scores may be overshadowed, as only the highest score is illustrated following the approach described in Section 4.3.3. Last but not least, APSIA is able to represent the evolution of risks among consecutive risk assessments by highlighting the differences in the histograms. Such differences may occur as a result of mitigation actions or the emergence of new cybersecurity and privacy threats.

Overall, APSIA offers a gamut of enablers to converge and automate the privacy and cybersecurity risk assessment by integrating asset inventory and vulnerability detection tools to support the assessment on a dynamic manner. The visualisation options based on the use of the interdependency graphs and the data processing flows constitute one of the competitive advantages of APSIA. In fact, as highlighted by the authors in [21], the CNIL method, which is one of the most prominent methods in the PIA field, does not use visual representations of the information flows. Such a feature is the foundation stone for the whole PIA and is facilitated by APSIA to a great extend.

6. Cybersecurity and Privacy Risk Assessment: The Road Ahead

While the area of risk assessment is rather mature, as aforementioned, the landscape of privacy impact assessment is fragmented into various families based on emerging research challenges. Undoubtedly, converging the usually contradicting security and privacy requirements towards a better risk assessment and estimation is a prominent challenge with a number of consequences, should it is not addressed appropriately. APSIA methodology can, therefore, be considered as the first step towards not only the development of a holistic framework that can alleviate this hurdle, but also as the basis for future research that will attempt to address the following emerging questions.

Cybersecurity and Privacy Risk Assessment for Supply chains: As aforementioned, nowadays, emerging application domains can be seen as "Systems-of-Systems" made up of heterogeneous cyber-physical systems, supplied by multiple vendors, that are increasingly connected to global information and management networks. Consequently, we must understand such ecosystems as federated safety critical systems designed, implemented, operated and owned by multiple tenants with different security and privacy goals, requirements, and priorities. Furthermore, security and privacy cannot be seen in an isolated way, but must be considered also in the face of the safety of the overall system. Therefore, it is necessary to understand what is semantically sensible for a component of a certain type to do and from this microscopic view expand to overall system analysis. Particularly with respect to security and privacy, components must be enabled to make and prove statements about their state and actions so that the other components can align their actions appropriately and an overall system state can be assessed. This goes substantially beyond simple authorization schemes telling who may access whom but will require understanding of semantics of requests and chains of effects throughout the system and an analysis both statically at configuration-time and dynamically during runtime. The latter will then allow to conduct dynamic risk assessment and decide at runtime if an entity is still safe to be used; even if some components are compromised and fail. Such a reactive, runtime risk assessment model, facilitating the real-time handling of threats and identified risks, is therefore needed for conducting a holistic threat assessment of such hyper-connected Systems-of-Systems. This will enable the dynamic assessment and forecast of individual, cumulative and propagated risks. Based on the representation of assets along with their dependencies, the associated threats and vulnerabilities and the potential cascading effects, thus, allowing for enhanced situation awareness adaptation of

the entire SoS-enabled ecosystem supporting policy adjustments and the compilation of updated mitigation strategies.

Connection between the Physical and Cyber world: Nowadays, ICT infrastructures can be seen as complex environments that integrate various technologies and create a mesh of interconnections among systems, processes and actors. In the context of the well-documented cybersecurity risk assessment field, there is a plethora of standardised frameworks and guidelines that aim to capture and document the cyber risks which are triggered as a result of vulnerabilities, attacks and technical flaws that emerge from the cyber field. In addition, there is a wide spectrum of works that focus on the analysis of cyber-physical systems, but they revolve around the cyber threats of those systems [56,57]. Hence, the connection between the two worlds, and the assessment of risks as a result of cascading actions, which are triggered from actors or events in the physical world and may have an impact to the cyber dimension, is a research domain with major challenges and room for innovation. Especially when aiming to privacy preservation, the actions of physical actors (e.g., data processor) could lead to security incidents that may put data subjects in position to encounter significant, or even irreversible consequences, which they may not overcome. In this context, there is an open research question on how to assess, and possibly quantify, those privacy related cascading risks. Physical-equivalent scoring systems like those developed for the quantification of the impact of technical vulnerabilities and weaknesses, like CVSS and CWE would contribute in this direction. In the context of APSIA method and tool, the interdependency graphs can be used for defining dependencies between the physical and cyber world using the isUsedBy and isLocatedIn relations. This feature could be used as an enabler for building future methods that consider this connection between the physical and cyber world.

Risk Mitigation through Optimal Countermeasure Selection: In light of the complex and demanding task of risk mitigation, one of the current hurdles that both the scientific community and ICT industry are trying to overcome pertains to the identification of appropriate mitigation actions towards increasing the robustness of cyber defense solutions. In this context, there have been several mechanisms proposed for supporting automated reaction; delivering a set of mitigation suggestions to the decision makers [58]. The vast majority of such frameworks capitalise on multi-objective optimisation techniques and constraint solvers in order to find an optimal mitigation action, or a set of actions, that can eliminate the identified risks. More, complex techniques consider game theoretic methods for emulating the engagement between defenders and aggressors [59]. This specific field has a number of challenges to tackle in order to bridge the gap from theory to practice. The scalability of optimisation solutions, in cases where the environment poses several objectives to be met, in conjunction with graph modeling techniques and live feeds of intrusion detection systems [60,61], create wide search areas for the optimal solution and has been reported as a major challenge to overcome. In addition, these optimisation processes do not consider any standard representation of remediation actions or a clear mapping between well-correlated sets of mitigation actions and threats in order to provide interoperable and effective solutions [62]. Notably, the aforementioned endeavors widely neglect the privacy preservation aspect in their objectives. This is mainly due to the fact that cybersecurity and privacy risk assessments are treated as distinct management processes, but also there is a lack of scoring systems that can capture and quantify the interdependence between these two aspects. Therefore, the instantiation of tools like APSIA can enable the model-based risk workflow, scoring and assessment for both the security and privacy assurance of mixed-criticality applications. Risks that have been identified in core assets (comprising the target application), can be weighted according to their criticality and impact degree, allowing the resolution of an optimization problem for the selection, deployment and placement of the best set of possible mitigation actions; ranging from the enactment of security solutions based on traditional cryptographic primitives (i.e., symmetric/asymmetric encryption, digital signatures [63], etc.) to more advanced trust assurance services [64,65] including network-based intrusion detection systems or (remote)

attestation controls. However, despite the clear benefits of such approaches, their applicability in environments with resource constrained devices, especially when it comes to privacy-related mitigation actions, is rather challenging [66].

For instance, strong cryptographic protocols can be used to increase trust, by not letting privacy risks be technically possible. Over the past years, a number of technologies have been developed to build Privacy Preserving Attribute-based Credentials (Privacy-ABCs) in a way that they can be trusted, like normal cryptographic certificates, while at the same time they protect the privacy of their holder [67]. Such Privacy-ABCs are issued just like ordinary cryptographic credentials using a digital secret signature key [63]. However, Privacy-ABCs allow their holder to transform them into a new token, in such a way that the privacy of the user is protected [68]. Bringing more control on the edge side, however, also gave birth to Direct Anonymous Attestation [8,69]; a platform authentication mechanism that enables the provision of privacy-preserving and accountable services. DAA is based on group signatures that allow remote attestation of a device associated to a Trusted Component (TC) while offering strong anonymity guarantees. Standardised by the Trusted Computing Group (TCG), DAA retains user anonymity, provides device-controlled unlinkability, and identifies signatures created by compromised devices. Given the aforementioned challenges, the privacy scoring system and methodology of APSIA could be used as an enabler for building such decision support systems that focus to the privacy realm.

7. Conclusions

In this work, we presented the APSIA methodology and tool in order to bridge the gap between the (automated) cybersecurity and privacy risk assessment conduction, which are widely treated as distinct management processes. The building blocks of APSIA, namely the Interdependency Graphs, the Privacy Scoring System, and the Data processing flows modeling based on GDPR are unified under the umbrella of an assessment methodology that enables the automated risk identification from data flows, the automatic creation of PIA reports, and the continuous execution of the risk assessment steps. The aforementioned offerings, were evaluated in the context of a heavily regulated sector (namely, the assistive healthcare domain) where strict security and privacy considerations are not only expected but mandated so as to better showcase the beneficial characteristics of APSIA. More specifically, the interdependency graphs were used to generate a detailed cartography of the engaged assets in the data processing activities of an emulated healthcare network. Based on this view, APSIA generated a number of flows that revealed the manner and and under what conditions patients' sensitive data are handled in the organisation. Crucially, the identified cyber security vulnerabilities of the core infrastructure assets were evaluated based on the novel privacy scoring system, in order to acquire a quantification of the impact of the vulnerability in the privacy dimension. This privacy-oriented score steers the decision maker in identifying, prioritising, anticipating and, finally, mitigating the risks by having privacy as the prominent quality that needs to be safeguarded.

Overall, all these APSIA innovations enable the presented tool to exceed the rigid approaches of privacy impact assessment and provide the means to conduct an enhanced, dynamic and generic assessment as an integral part of an iterative and unified risk assessment process on-the-fly. In this context, while one could argue that APSIA is limited only to the GDPR-related privacy considerations, the core methodology is generic enough to facilitate any ecosystem with "privacy-by-design" properties. Since such properties are independent of the system or the application domain itself, putting forth the methodology to be able to precisely model them in the presence of strong adversaries is a prerequisite for any legal framework. Given this remark, however, we aim to extend APSIA and evaluate it in scenarios with privacy principles that are documented by additional legal frameworks (i.e., NIST). Furthermore, as a future work, we aim to deploy and evaluate APSIA in a larger pilot and perform improvements on the usability of the tool according to the views of engaged stakeholders.

Finally, by taking into consideration the salient characteristics of risk assessment, as a building block, along with the requirements of the involved actors, we identified a number of open research challenges. It is our strong belief that if these challenges are tackled now while APSIA is still at an early stage, then this emerging security and privacy mechanism can reach its full potential.

Author Contributions: Conceptualization, D.P. and S.A.M.; methodology, D.P. and S.A.M.; software, P.G.; validation, T.G. and P.G.; investigation, D.P. and S.A.M.; resources, P.G.; writing—original draft preparation, D.P. and S.A.M.; writing—review and editing, D.P., S.A.M., T.G. and P.G.; funding acquisition, P.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the European Commission, under the CUREX and ASSURED projects; Grant Agreement No. 826404 and 952697, respectively.

Data Availability Statement: Not Applicable, the study does not report any data.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

APS	Asset-level Privacy Score
APSIA	Automated Privacy and Security Impact Assessment
BPM LCM	Business Process Model Life-Cycle Management
CISO	Chief Information Security Officer
CPS	Cyber Physical Systems
CVE	Common Vulnerabilities and Exposures
CVSS	Common Vulnerability Scoring System
DPA	Data Protection Authority
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
ENISA	European Union Agency for Cybersecurity
GDPR	General Data Protection Regulation
HIS	Health Information System
ICT	Information and Communication Technologies
IoT	Internet of Things
ISO	International Organization for Standardization
MC	Mitigation Controls
NIST	National Institute of Standards and Technology
NVD	National Vulnerability Database
OPS	Organisation-level Privacy Score
PAPS	Processing Activity-level Privacy Score
PET	Privacy-Enhancing Technologies
PIA	Privacy Impact Assessment
PII	Personally Identifiable Information
pISRA	Privacy-considered Information Security Risk Assessment
POC	Point of Care
RMF	Risk Management Framework
SMEs	Small and Medium Enterprises
SPIA	Security and Privacy Impact Assessment
SSN	Social Security Number

Appendix A. APSIA Graphical User Interface Components

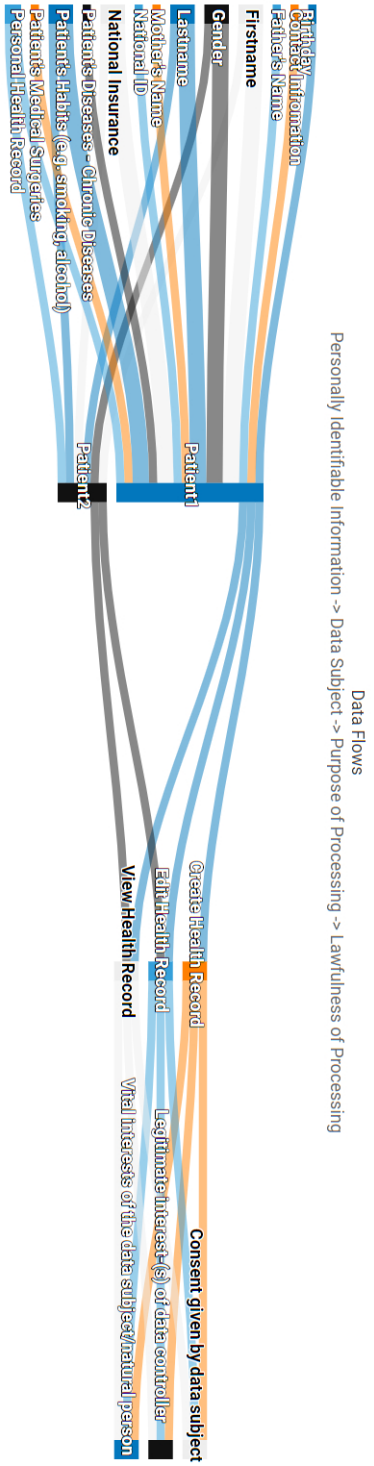

Figure A1. Data processing flows in the frame of GDPR.

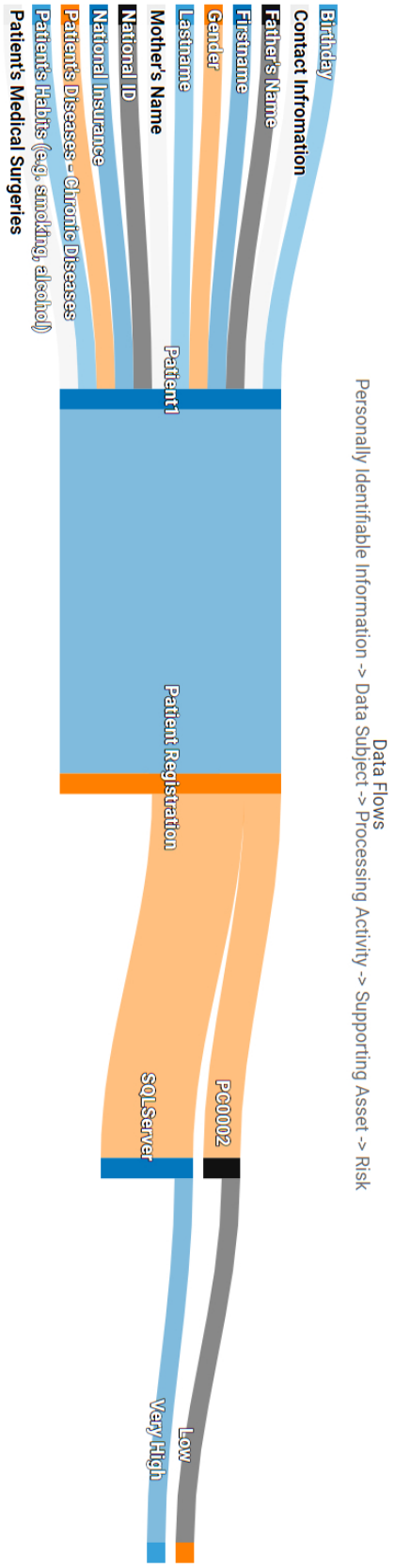

Figure A2. Patient Registration Risk Data Flows.

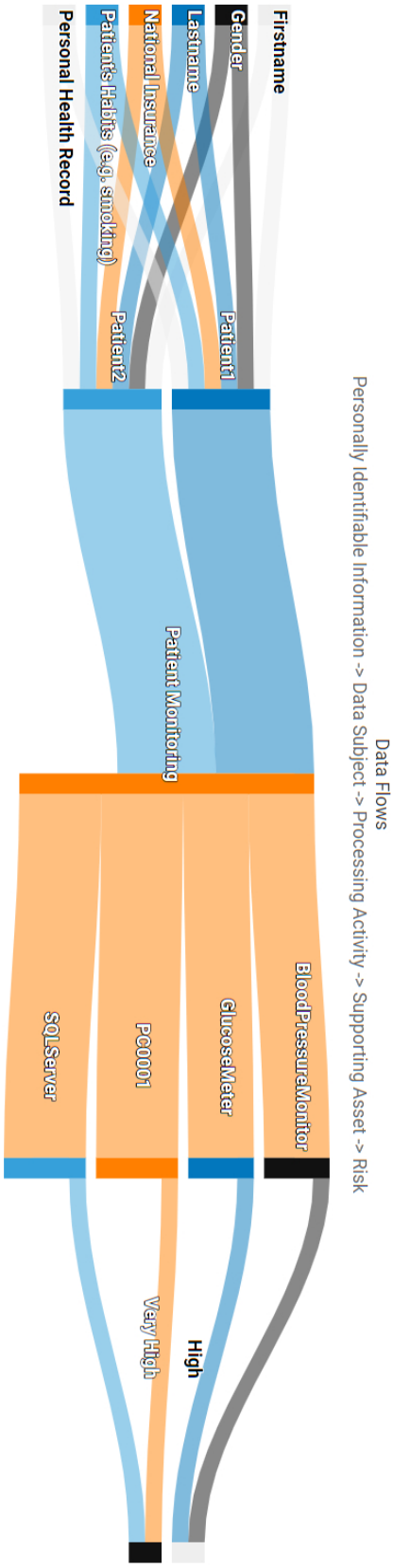

Figure A3. Patient Monitoring Risk Data Flows.

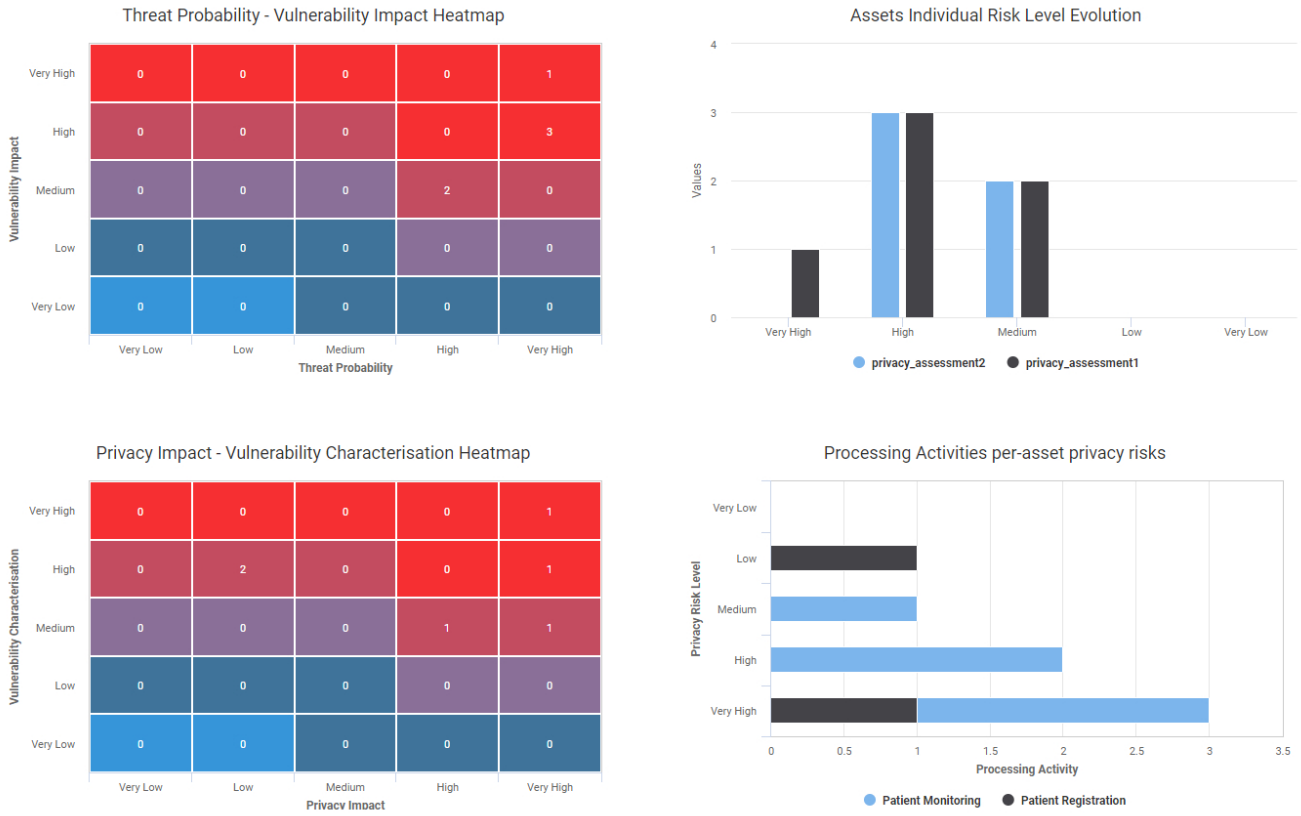

Figure A4. APSIA Dashboard heat-maps and histograms.

References

1. Jelisic, E.; Ivezic, N.; Kulvatunyou, B.; Anicic, N.; Marjanovic, Z. A Business-Context-Based Approach for Message Standards Use-A Validation Study. *Commun. Comput. Inf. Sci.* **2019**, *1064*, 337–349. [CrossRef]
2. Kulvatunyou, B.S.; Ivezic, N.; Srinivasan, V. On architecting and composing engineering information services to enable smart manufacturing. *J. Comput. Inf. Sci. Eng.* **2016**, *45–52*. [CrossRef] [PubMed]
3. Dimitriadis, A.; Flores, J.L.; Kulvatunyou, B.; Ivezic, N.; Mavridis, I. ARES: Automated Risk Estimation in Smart Sensor Environments. *Sensors* **2020**, *20*, 4617. [CrossRef]
4. Vemou, K.; Karyda, M. An Evaluation Framework for Privacy Impact Assessment Methods. In Proceedings of the MCIS 2018 Proceedings, Corfu, Greece, 28–30 September 2018.
5. National Institute of Standards and Technology. Risk Management Framework for Information Systems and Organizations: A System Life Cycle Approach for Security and Privacy. 2018. Available online: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-37r2.pdf (accessed on 26 January 2021).
6. Dimitriou, T.; Giannetsos, T.; Chen, L. REWARDS: Privacy-preserving rewarding and incentive schemes for the smart electricity grid and other loyalty systems. *Comput. Commun.* **2019**, *137*, 1–14. [CrossRef]
7. Giannetsos, T.; Dimitriou, T.; Prasad, N.R. People-centric sensing in assistive healthcare: Privacy challenges and directions. *Secur. Commun. Netw.* **2011**, *4*, 1295–1307. [CrossRef]
8. Whitefield, J.; Chen, L.; Giannetsos, T.; Schneider, S.; Treharne, H. Privacy-enhanced capabilities for VANETs using direct anonymous attestation. In Proceedings of the 2017 IEEE Vehicular Networking Conference (VNC), Torino, Italy, 27–29 November 2017; pp. 123–130. [CrossRef]
9. Pearson, S.; Yee, G. *Privacy and Security for Cloud Computing*; Springer Publishing Company: London, UK, 2014. [CrossRef]
10. National Institute of Standards and Technology (NIST). NIST Privacy Framework—A Tool to Help Organizations Improve Individuals' Privacy Through Enterprise Risk Management. 2020. Available online: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf (accessed on 26 January 2021).
11. Clarke, R. Privacy impact assessment: Its origins and development. *Comput. Law Secur. Rev.* **2009**, *25*, 123–135. [CrossRef]
12. Vemou, K.; Karyda, M. Evaluating privacy impact assessment methods: Guidelines and best practice. *Inf. Comput. Secur.* **2020**, *28*, 35–53. [CrossRef]
13. Wagner, I.; Eckhoff, D. Technical Privacy Metrics: A Systematic Survey. *ACM Comput. Surv.* **2018**, *51*.

14. International Organization for Standardization (ISO). ISO/IEC 29134:2017 Information Technology—Security Techniques—Guidelines for Privacy Impact Assessment. 2017. Available online: https://www.iso.org/obp/ui/#iso:std:iso-iec:29134:ed-1:v1:en (accessed on 26 January 2021).
15. BSI. Data Protection-Specification for a Personal Information Management System. 2017. Available online: https://www.bsigroup.com/en-GB/BS-10012-Personal-information-management/ (accessed on 30 December 2020).
16. International Organization for Standardization (ISO). Iso/Iec 29151:2017 Information Technology—Security Techniques—Code Of Practice For Personally Identifiable Information Protection. 2017. Available online: https://www.iso.org/obp/ui/#iso:std:iso-iec:29151:ed-1:v1:en (accessed on 26 January 2021).
17. International Organization for Standardization (ISO). Iso/Iec 27018:2014 Information Technology—Security Techniques—Code Of Practice For Protection Of Personally Identifiable Information (Pii) In Public Clouds Acting As Pii Processors. 2014. Available online: https://www.iso.org/obp/ui/#iso:std:iso-iec:27018:ed-1:v1:en (accessed on 26 January 2021).
18. Oetzel, M.C.; Spiekermann, S. A systematic methodology for privacy impact assessments: A design science approach. *Eur. J. Inf. Syst.* **2014**, *23*, 126–150. [CrossRef]
19. Wei, Y.C.; Wu, W.C.; Lai, G.H.; Chu, Y.C. pISRA: Privacy considered information security risk assessment model. *J. Supercomput.* **2020**, *76*, 1468–1481. [CrossRef]
20. Information Commissioner's Office. Data Protection Impact Assessments (DPIAs). 2018. Available online: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/ (accessed on 8 November 2020).
21. Bisztray, T.; Gruschka, N. Privacy Impact Assessment: Comparing Methodologies with a Focus on Practicality. In *Secure IT Systems*; Springer International Publishing: Cham, Switzerland, 2019; pp. 3–19. [CrossRef]
22. French Data Protection Authority (CNIL). *Privacy Impact Assessment (PIA) Methodology*; 2018. Available online: https://www.cnil.fr/sites/default/files/typo/document/CNIL-PIA-1-Methodology.pdf (accessed on 26 January 2021).
23. Treasury Board of Canada Secretariat. Directive of Privacy Impact Assessments. 2010. Available online: https://www.isc.upenn.edu/sites/default/files/introduction_to_spia_program.pdf (accessed on 29 December 2020).
24. Ahmadian, A.S.; Strüber, D.; Riediger, V.; Jürjens, J. Supporting Privacy Impact Assessment by Model-Based Privacy Analysis. In Proceedings of the 33rd Annual ACM Symposium on Applied Computing, Pau, France, 9–13 April 2018; pp. 1467–1474. [CrossRef]
25. Wuyts, K.; Joosen, W. LINDDUN Privacy Threat Modeling: A Tutorial. 2015. Available online: https://lirias.kuleuven.be/retrieve/331950 (accessed on 30 December 2020).
26. French Data Protection Authority (CNIL). Methodology for Privacy Risk Management—How to implement the Data Protection Act. 2012. Available online: https://www.cnil.fr/sites/default/files/typo/document/CNIL-ManagingPrivacyRisks-Methodology.pdf (accessed on 30 December 2020).
27. ENISA—European Union Agency for Cybersecurity. Privacy and Data Protection by Design—From Policy to Engineering. 2014. Available online: https://arxiv.org/ftp/arxiv/papers/1501/1501.03726.pdf (accessed on 30 December 2020).
28. Ahmadian, A.S.; Strüber, D.; Jürjens, J. Privacy-enhanced system design modeling based on privacy features. In Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing, SAC 2019, Limassol, Cyprus, 8–12 April 2019; pp. 1492–1499. [CrossRef]
29. Manna, A.; Sengupta, A.; Mazumdar, C. A Quantitative Methodology for Business Process-Based Data Privacy Risk Computation. *Adv. Comput. Syst. Secur.* **2020**, *10*, 17–33. [CrossRef]
30. Henriksen-Bulmer, J.; Faily, S.; Jeary, S. DPIA in Context: Applying DPIA to Assess Privacy Risks of Cyber Physical Systems. *Future Internet* **2020**, *12*, 93. [CrossRef]
31. ENISA—European Union Agency for Cybersecurity. On-Line Tool for the Security of Personal Data Processing. Available online: https://www.enisa.europa.eu/risk-level-tool/risk (accessed on 30 December 2020).
32. GS1. EPC/RFID Privacy Impact Assessment Tool. 2015. Available online: https://www.gs1.org/standards/epc-rfid/pia (accessed on 26 January 2021).
33. University of Pennsylvania. Introduction to the SPIA Program. 2016. Available online: https://www.isc.upenn.edu/sites/default/files/introduction_to_spia_program.pdf (accessed on 29 December 2020).
34. French Data Protection Authority (CNIL). Privacy Impact Assessment (PIA) Tool. 2015. Available online: https://www.cnil.fr/en/privacy-impact-assessment-pia (accessed on 30 December 2020).
35. ENISA Ad Hoc Working Group on Risk Assessment and Risk Management. Information Packages for Small and Medium Sized Enterprises (SMEs). In *Information Packages for SMEs, Deliverable 2, Final Version, Version 1.0*; 2006. Available online: https://www.enisa.europa.eu/publications/information-package-for-smes/at_download/fullReport (accessed on 26 January 2021).
36. Manson, C.; Gorniak, S. Recommendations for a Methodology of the Assessment of Severity of Personal Data Breaches. In *ENISA (European Union Agency for Network and Inform. Security) Working Document, v1.0*; 2013. Available online: https://www.enisa.europa.eu/publications/corporate/enisa-annual-report-2013 (accessed on 26 January 2021).
37. Agarwal, S., Developing a Structured Metric to Measure Privacy Risk in Privacy Impact Assessments. In *Privacy and Identity Management. Time for a Revolution*; Springer International Publishing: Edinburgh, UK, 2016; pp. 141–155. [CrossRef]
38. Wadhwa, K.; Rodrigues, R. Evaluating privacy impact assessments. *Innov. Eur. J. Soc. Sci. Res.* **2013**, *26*, 161–180. [CrossRef]

39. Piatkowska, E.; Bajraktari, A.; Chhajed, D.; Smith, P. Tool support for data protection impact assessment in the smart grid. *Elektrotechnik Inf.* **2017**, *134*, 26–29. [CrossRef]
40. National Institute of Standards (NIST). National Vulnerability Database (NVD). Available online: https://nvd.nist.gov/ (accessed on 26 January 2021).
41. The MITRE Corporation. Common Vulnerabilities and Exposures (CVE). Available online: https://cve.mitre.org/ (accessed on 26 January 2021).
42. Centre for Internet Security. CIS Controls v7.1. 2020. Available online: https://www.cisecurity.org/controls/ (accessed on 30 December 2020).
43. Polemi, N.; Kotzanikolaou, P. Medusa: A Supply Chain Risk Assessment Methodology. In *Cyber Security and Privacy*; Springer International Publishing: Brussels, Belgium, 2015; pp. 79–90. [CrossRef]
44. Kalogeraki, E.M.; Papastergiou, S.; Mouratidis, H.; Polemi, N. A Novel Risk Assessment Methodology for SCADA Maritime Logistics Environments. *Appl. Sci.* **2018**, *8*, 1477. [CrossRef]
45. Rinaldi, S.M.; Peerenboom, J.P.; Kelly, T.K. Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control. Syst. Mag.* **2001**, *21*, 11–25. [CrossRef]
46. QED Secure Solutions. Risk Scoring System for Medical Devices (RSS-MD)-Technical Specification Guide. Available online: https://www.riskscoringsystem.com/medical/techspecmedical.pdf (accessed on 8 November 2020).
47. FIRST. Common Vulnerability Scoring System (CVSS). Available online: https://www.first.org/cvss/ (accessed on 26 January 2021).
48. Greenbone Networks. Open Vulnerability Assessment Scanner (OpenVas). Available online: https://www.openvas.org/ (accessed on 26 January 2021).
49. ENISA—European Union Agency for Cybersecurity. Handbook on Security of Personal Data Processing. 2018. Available online: https://www.enisa.europa.eu/publications/handbook-on-security-of-personal-data-processing (accessed on 30 December 2020).
50. CVEdetails. Available online: https://www.cvedetails.com/ (accessed on 26 January 2021).
51. Offensive Security. Exploit Database-Exploits for Penetration Testers. Available online: https://www.exploit-db.com/ (accessed on 26 January 2021).
52. De Capitani di Vimercati, S.; Foresti, S.; Livraga, G.; Samarati, P. Data Privacy: Definitions and Techniques. *Int. J. Uncertain. Fuzziness Knowl. Based Syst.* **2012**, *20*, 793–818. [CrossRef]
53. Makri, E.L.; Georgiopoulou, Z.; Lambrinoudakis, C. A Proposed Privacy Impact Assessment Method Using Metrics Based on Organizational Characteristics. In *Computer Security*; Springer International Publishing: Luxembourg, 2020; pp. 122–139. [CrossRef]
54. Mohammadi, F.; Panou, A.; Ntantogian, C.; Karapistoli, E.; Panaousis, E.; Xenakis, C. CUREX: seCUre and pRivate hEalth data eXchange. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence, Thessaloniki, Greece, 14–17 October 2019; Volume 24800, pp. 263–268. [CrossRef]
55. Quesada-González, D.; Merkoçi, A. Nanomaterial-based devices for point-of-care diagnostic applications. *Chem. Soc. Rev.* **2018**, *47*, 4697–4709. [CrossRef]
56. Zhong, S.; Zhong, H.; Huang, X.; Yang, P.; Shi, J.; Xie, L.; Wang, K., Connecting Physical-World to Cyber-World: Security and Privacy Issues in Pervasive Sensing. In *Security and Privacy for Next-Generation Wireless Networks*; Springer International Publishing: Cham, Switzerland, 2019; pp. 49–63.
57. Humayed, A.; Lin, J.; Li, F.; Luo, B. Cyber-Physical Systems Security—A Survey. *IEEE Internet Things J.* **2017**, *4*, 1802–1831. [CrossRef]
58. Nespoli, P.; Papamartzivanos, D.; Gómez Mármol, F.; Kambourakis, G. Optimal Countermeasures Selection Against Cyber Attacks: A Comprehensive Survey on Reaction Frameworks. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 1361–1396. [CrossRef]
59. Fielder, A.; Panaousis, E.; Malacaria, P.; Hankin, C.; Smeraldi, F. Decision support approaches for cyber security investment. *Decis. Support Syst.* **2016**, *86*, 13–23. [CrossRef]
60. Papamartzivanos, D.; Gómez Mármol, F.; Kambourakis, G. Introducing Deep Learning Self-Adaptive Misuse Network Intrusion Detection Systems. *IEEE Access* **2019**, *7*, 13546–13560. [CrossRef]
61. Papamartzivanos, D.; Gómez Mármol, F.; Kambourakis, G. Dendron: Genetic trees driven rule induction for network intrusion detection systems. *Future Gener. Comput. Syst.* **2018**, *79*, 558–574. [CrossRef]
62. Nespoli, P.; Mármol, F.G.; Vidal, J.M. Battling against cyberattacks: Towards pre-standardization of countermeasures. *Clust. Comput.* **2020**, 1–25. [CrossRef]
63. Sanchez, J.L.C.; Bernal Bernabe, J.; Skarmeta, A.F. Integration of Anonymous Credential Systems in IoT Constrained Environments. *IEEE Access* **2018**, *6*, 4767–4778. [CrossRef]
64. Larsen, B.; Debes, H.B.; Giannetsos, T. CloudVaults: Integrating Trust Extensions into System Integrity Verification for Cloud-Based Environments. In *Computer Security. ESORICS 2020. Lecture Notes in Computer Science*; Springer International Publishing: Guildford, UK, 2020; Volume 12580, pp. 197–220.
65. Camenisch, J.; Drijvers, M.; Lehmann, A. Anonymous Attestation with Subverted TPMs. In *Advances in Cryptology-CRYPTO 2017*; Springer: Santa Barbara, CA, USA, 2017; pp. 427–461.

66. Saraiva, D.A.F.; Leithardt, V.R.Q.; de Paula, D.; Mendes, A.S.; Villarrubia-González, G.; Crocker, P. PRISEC: Comparison of Symmetric Key Algorithms for IoT Devices. *Sensors* **2019**, *19*, 4312. [CrossRef]
67. Sabouri, A.; Krontiris, I.; Rannenberg, K. Trust relationships in privacy-ABCs ecosystems. In *International Conference on Trust, Privacy and Security in Digital Business*; Springer International Publishing: Cham, Switzerland, 2014; pp. 13–23.
68. Gisdakis, S.; Giannetsos, T.; Papadimitratos, P. SPPEAR: Security & Privacy-preserving Architecture for Participatory-sensing Applications. In Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless, New York, NY, USA, 23–25 July 2014; pp. 39–50.
69. Brickell, E.F.; Camenisch, J.; Chen, L. Direct anonymous attestation. In Proceedings of the ACM Conference on Computer and Communications Security, CCS, Washington, DC, USA, 25–29 October 2004.