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

Project start: February 1, 2019
Project duration: 42 months

Cyber Security for Europe

D3.2

Cross Sectoral Cybersecurity Building Blocks

Document Identification	
Due date	1 January 2020
Submission date	30 April 2020
Revision	2.0 (30 April 2020)

Related WP	WP3	Dissemination Level	PU
Lead Participant	AIT	Lead Author	Stephan Krenn
Contributing Beneficiaries	GUF, UMU, NEC, UNITN, AIT, ATOS, CYBER, CNR, VTT, DTU, UCY, UM, POLITO, UMA, UNILU, C3P, UPRC	Related Deliverables	D3.1, D3.13, D3.20

Abstract: A main objective of CyberSec4Europe is to develop core innovative cybersecurity building blocks, providing pioneering technologies on top of innovative tools to enhance the security and privacy of services. This task also includes identity management and authentication solutions over multiple non-federated providers, security and privacy services to deploy a basic Edge Computing platform, identify technologies to reduce the system attack surface, design security mechanisms based on Trusted Execution Environments (TEE) and design a framework for TEE-based cloud data processing, IoT Privacy Preserving Middleware Platform, improve integrated Security & Privacy by Design approaches, decentralized evidence-based authorization and distributed access control using blockchain, addressing applications in IoT and investigate approaches that achieve extreme privacy- and integrity-preserving storage and processing of critical data with long-term protection requirements.

In this document we give a summary over the relevant cybersecurity building blocks, assets, and expertise currently existing within the CyberSec4Europe consortium.

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

Executive Summary

This deliverable is the first outcome of task “T3.2 - Research and Integration on Cybersecurity Enablers and underlying Technologies”, which aims at identifying and improving generic and cross-sectoral enablers for privacy and cyber security, as well as the identification of research challenges for common technologies in these domains.

In this regard, this deliverable describes an initial overview of cross-sectoral cyber security building blocks. The document focuses on more than 20 building blocks and assets that are already existing within the CyberSec4Europe consortium, and which might be of relevance for “WP5 - Demonstration Cases”, where different verticals are showcased. Besides describing these building blocks, the document also identifies already known research challenges that need to be addressed within the project, thereby also linking to “WP4 - Research and Development Roadmap”.

The document also specifies a generic privacy and security architecture that instantiates the T3.2-related components of the overall CyberSec4Europe Global Architecture presented in “D3.1 - Common Framework Handbook 1”.

Document information

Contributors

Name	Partner
Valerio Cini	AIT
Stephan Krenn	AIT
Thomas Lorünser	AIT
Juan Carlos Perez Baun	ATOS
Salvador Pérez Franco	ATOS
Rolando Martins	C3P
João Soares	C3P
Eda Marchetti	CNR
Said Daoudagh	CNR
Liina Kamm	CYBER
Peeter Laud	CYBER
Nicola Dragoni	DTU
Ahad Niknia	GUF
Kai Rannenberg	GUF
Claudio Soriente	NEC
Alessandro Sforzin	NEC
Antonio Lioy	POLITO
Marko Kompara	UM
Marko Hölbl	UM
Javier Lopez	UMA
Ruben Rios	UMA
Jorge Bernal Bernabé	UMU
Alireza Esfahani	UNILU
Panagiotis Bountakas	UPRC
Niko Lehto	VTT

Reviewers

Name	Partner
Ahad Niknia	GUF
Marko Hölbl	UM
Ivan Pashchenko	UNITN

History

0.01	2019-09-10	Stephan Krenn	Initial ToC
0.02	2019-10-06	Stephan Krenn Thomas Lorünser	Merged building blocks for anonymous credentials, secret sharing, data analytics, and rewriting of blockchains
0.03	2019-10-07	Juan Carlos Perez Baun	Added SPeIDI and DANS
0.04	2019-10-08	Nicola Dragoni	Added AntibIoTics building block
0.05	2019-10-09	Ruben Rios	Added Privacy Monitor for the Edge
0.06	2019-10-19	Marko Kompara Marko Hölbl	Added GDPR compliant user experience and interoperability and cross-border compliance
0.07	2019-10-21	Peeter Laud	Flexible metrics and analyses for differential privacy
0.08	2019-11-02	Said Daoudagh Eda Marchetti	GDPR-based user stories in the access control perspective
0.09	2019-11-05	Liina Kamm	Provided CYBER's inputs on ShareMind
0.10	2019-11-06	Jorge Bernal Antonio Skarmeta	Added Self-Sovereign privacy-preserving IdM
0.11	2019-11-06	João Soares	Added UP's building blocks on Argus and pTASK
0.12	2019-11-20	Ruben Rios	First draft of Section 2 introducing the CyberSec4Europe Privacy-preserving architecture
0.13	2019-11-26	Jorge Bernal	Improved and extended Section 2 CyberSec4Europe Privacy-preserving Architecture.
0.14	2019-11-27	Stephan Krenn	Finalized document for first internal review cycle
0.15	2019-11-29	Liina Kamm	Added research challenges in Section 10
0.16	2019-11-30	Ruben Rios	Added research challenges in Section 4
0.17	2019-12-09	Stephan Krenn	Addressed reviewer comments
0.18	2019-12-12	Stephan Krenn Jorge Bernal Niko Lehto Claudio Soriente Ruben Rios Juan Carlos Perez Baun	Added missing research challenges, improved wording, etc.
0.19	2019-12-13	Stephan Krenn	Prepared final version for submission to Coordinator
1.0	**2019-12-20**	**Kai Rannenberg**	**Final alignments by Coordinator** **Version submitted to the European Commission**
1.01	2020-02-21	Nicola Dragoni	Updates on future work on AntibIoTic
1.02	2020-02-27	Claudio Soriente	Revised sections by NEC based on reviewer's comments
1.03	2020-03-06	Marko Kompara	Revised section on GDPR compliant user experience
1.04	2020-03-09	Eda Marchetti	Revised CNR contributions
1.05	2020-03-11	Juan Carlos Perez Baun	Revised SPeIDI and DANS based on reviewers' comments
1.06	2020-03-26	Stephan Krenn	Added metadata tables for all building blocks
1.07	2020-03-27	Valerio Cini Stephan Krenn Thomas Lorünser	Updated AIT-related building blocks regarding mapping to the architecture, related work, … as requested by the reviewers
1.08	2020-03-30	Juan Carlos Perez Baun	Refinements of ATOS components
1.09	2020-03-30	Ruben Rios	Added state of the art for Edge Privacy

1.10	2020-03-30	Jorge Bernal Bernabé	Extensions and refinements of UMU components; added eIDAS Browser App
1.11	2020-04-01	Claudio Soriente Alessandro Sforzin	Added research topics to Sections 6 and 9, and added related work on Sections 9.2 and 9.3
1.12	2020-04-07	Stephan Krenn Thomas Lorünser	Copy-editing, preparing document for internal review
1.13	2020-04-13	João Soares	Extended research challenges for Section 7
1.14	2020-04-15	Marko Kompara	Updated Section 10.6
1.15	2020-04-16	Panagiotis Bountakas	Added password-less authentication building block
1.16	2020-04-18	Jorge Bernal Stephan Krenn	Updated figures according to reviewers' comments
1.17	2020-04-20	Stephan Krenn	Final editing
1.18	2020-04-29	Ahad Niknia	High-level review
1.19	2020-04-29	Stephan Krenn	Incorporated reviewer comments
2.0	2020-04-30	Ahad Niknia	Final check and preparation to submission
2.0	**2020-04-30**	**Kai Rannenberg**	**Final alignments by coordinator** **Version submitted to the European Commission**

Table of Contents

List of Figures

List of Tables

List of Acronyms

ABCs	Attribute-based credentials
ACP	Access Control Policy
AI	Artificial Intelligence
API	Application programming interface
ARX	Data Anonymization Tool
ASD	Agile Software Development
BFT	Byzantine Fault Tolerance
CA	Certificate Authority

DANS	Data Anonymization Service
EC	European Commission
DH	Diffie-Hellman
DID	Decentralized identifier
DLT	Distributed ledger technology
EC	European Commission
eID	Electronic Identification
eIDAS	electronic IDentification, Authentication and trust Services
FIDO	Fast Identity Online
GDPR	General Data Protection Regulation
HMAC	Hash-based message authentication code
HTTPS	HyperText Transfer Protocol Secure
IdP	Identity Provider
IoT	Internet of Things
I/O	Input/output
LLC	Limited location channel
LoA	Level of assurance
MitM	Man-in-the-middle
MPC	Multi-party computation
NIZK	Non-interactive zero-knowledge proof
PAKE	Password-authenticated key exchange
PET	Privacy-enhancing technology
PKI	Public-key infrastructure
QR	Quick response
SAS	Short authentication string
SNARK	Succinct non-interactive argument of knowledge
SP	Service provider
SPeIDI	Service Provider eIDAS Integrator
SSI	Self-sovereign identity
TEE	Trusted Execution Environment
TRL	Technology Readiness Level
VM	Virtual Machine
VoIP	Voice Over Internet Protocol
ZKP	Zero-knowledge proof

1 Introduction

CyberSec4Europe aims to meet the EU and Member States' next generation cybersecurity challenges through strengthening research and innovation competence and cybersecurity capacities both at the national as well as at the European level. Thus, the project is conducting cybersecurity research and innovation through technology advancements supporting both the autonomy of the Digital Single Market as well as addressing the security of the European citizen, European industry, the European economy and society as a whole.

To this aim, CyberSec4Europe as part of WP3 is developing and implementing Cyber Security Enablers. One main focus is thereby put on the following eight different domains:

- Identity management and authentication solutions over multiple non-federated providers, with a special focus on user privacy, while still giving high authenticity guarantees to the relying party;
- security and privacy services for edge computing platforms;
- technologies to reduce the system attack surface;
- security mechanisms based on Trusted Execution Environments (TEE) and frameworks for TEE-based cloud data processing;
- Privacy-preserving middleware for the Internet of Things;
- Security & Privacy by Design approaches
- Decentralized, evidence-based authorization and distributed access control using blockchain technologies;
- Long-term privacy- and integrity-preserving storage and processing of critical data in potentially untrusted environments.

The ambition of this document is to give an overview of horizontal cybersecurity building blocks that already exist (partially as results from previous, e.g., H2020 projects), or that are currently under development, within the consortium. The document, therefore, collects a variety of cross-domains tools and technologies that solve specific cybersecurity challenges that occur in different applications scenarios and that are flexible enough to be adapted for different needs. While many of the components have already been developed as part of earlier research initiatives, this document serves as a baseline for the technologies available to the demonstrators that will be developed within our project. How the specific needs from WP5 have been incorporated, how the building blocks have been extended, and which new building blocks have been introduced, will be presented in future revisions of this document, especially as part of D3.13 and D3.20.

For each of the identified cybersecurity building blocks, an overview of the functionality and the related state-of-the-art is given. Furthermore, each component is put into relation to the CyberSec4Europe project using a metadata table of the following unified format:

Full name	**Acronym**	**Lead partner**	**TRL**
Addresses requirements from	**Addressed requirements**	**Further information**	

Table 1: Metadata template

In the following we briefly explain the semantics of the different fields:

- **Full name.** Specifies the full name of the building blocks
- **Acronym.** If available, acronyms or short names are given here
- **Lead partner.** Beneficiary responsible for the building block within CyberSec4Europe
- **TRL.** Technology Readiness Level[1] on a scale from 1 ("basic principles observed") to 9 ("actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space)"), following the definition of the European Commission.
- **Addresses requirements from.** Tasks/demonstrators from WP5, which identified requirements that might be addressed using the building block. Note that this does not necessarily mean that the demonstrator will actually deploy the building block, which will be discussed and clarified during the demonstrator's design and implementation in joint work with WP5.
- **Addressed requirements.** Specific requirements that could be addressed using this building block. For detailed descriptions of the requirements, please refer to D5.1.
- **Further information.** Further information, such as web links, publications, etc.

The document needs to be seen in the larger context of the overall CyberSec4Europe project. On the one hand, it provides input to the vertical demonstrator cases developed in WP5 to allow them for properly addressing the cybersecurity challenges encountered in their specific areas. On the other hand, WP5 already provided first feedback back to WP3, identifying necessary research gaps that need to be addressed for further advancements of their pilots. These inputs are included in the presentation of the building blocks. By doing so, this deliverable also links to WP4, where research and development roadmaps for the cybersecurity of critical sectors of Europe are developed.

1.1 Document Structure

This document is structured as follows:

- Section 2 recaps the CyberSec4Europe global security architecture and then presents in detail the privacy-preserving components of this architecture.
- Sections 3 to 10 then list the existing cybersecurity building blocks in the categories listed above, map them to the architecture, and identify already known research challenges for the different domains.
- Finally, we briefly conclude in Section 11.

[1] https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf

2 CyberSec4Europe Privacy-Preserving Architecture

2.1 CyberSec4Europe Global Architecture and Privacy-Preserving Building Blocks

The CybserSec4Europe Privacy-Preserving Architecture is part of the general CyberSec4Europe architecture defined in Deliverable D3.1, which consists of several planes.

In the global architecture shown in Figure 1, the blocks related to Privacy-preservation and Identity management are highlighted in yellow.

To cope with the privacy-preserving and identity management and authentication solutions, as can be seen in (highlighted in yellow), the CyberSec4Europe architecture features a set of functional components that are located in the User domain, above all in the user-smartphone. It allows realizing the self-sovereign and privacy-preserving identity management model. The user domain also includes security and privacy enablers needed to protect users' privacy in communications and data.

The Identity, Trust and privacy-preserving Management Services functional component of the architecture belonging to the Control and Management Plane, includes identity management services, identity providers, attribute providers, claims verifiers, PKIs, biometric verifiers, privacy-enhancing technologies (PETs) managers, and enablers for trusted execution environments (TEE), creation and monitoring (such as remote attestation).

In the Managed domain, the security and privacy-preservation tools will instantiate a set of middleware tools in the IoT-Edge domain, and enablers such as anti-malware, data leakage prevention, data anonymization, data broker privacy preservation, and in general tools to reduce the system attack surface. In other web-oriented domains (e.g. eCommerce use cases) the managed domain will embrace the enablers (including PET oriented technologies) required by service providers to verify proofs, and manage access control to services, based on claims and assertions obtained from IdPs.

In the Blockchain Privacy-Preserving SSI Layer there will be a set of functional components, similar to those allocated in the Identity, Trust and privacy-preserving Management Services but adapted to be deployed in the blockchain. It encompasses services for ID proofing, verification-authentication, claims/attestation verifications, TEEs. Thus, users will be able to share/access assets (identity, data, resources), manage attestation, manage credentials through blockchain in a privacy-preserving way. The functional components in blockchain deal with evidence-based authorization and distributed access control using blockchain, addressing applications in IoT.

2.2 CyberSec4Europe Privacy-Preserving Functional Architecture

The CyberSec4Europe Privacy-Preserving Architecture consist of a number of building blocks which expand over several intertwined domains, including the user domain, the web domain and the IoT domain, as shown in Figure 2. The building blocks are defined for different purposes which range from the compliance with current legal frameworks such as eIDAS and GDPR to mechanisms related to hardware-based solutions for managing keys and applications securely. Next we give an overview of the different building blocks that are being proposed.

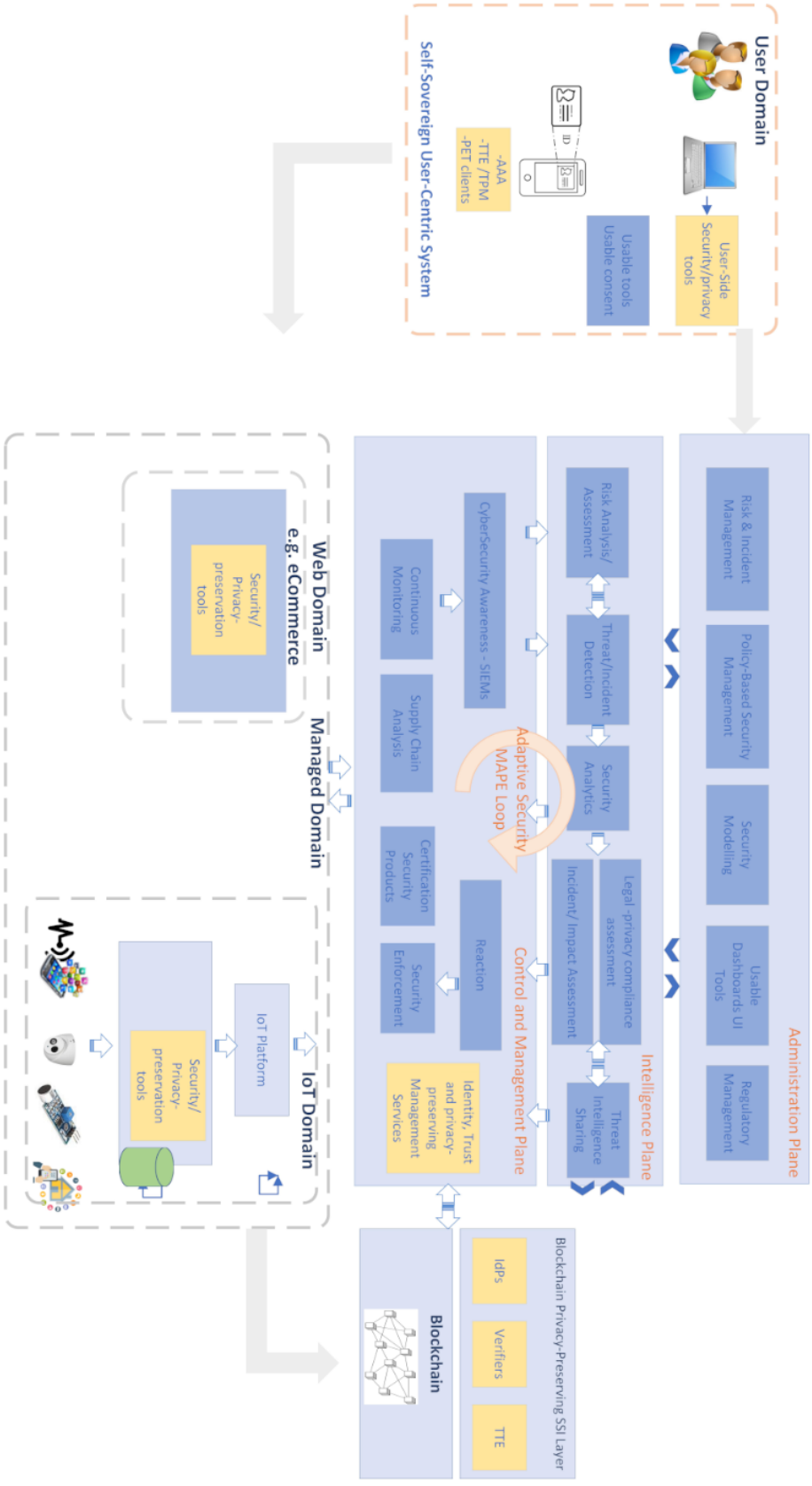

Figure 1: CyberSec4Europe Global Architecture

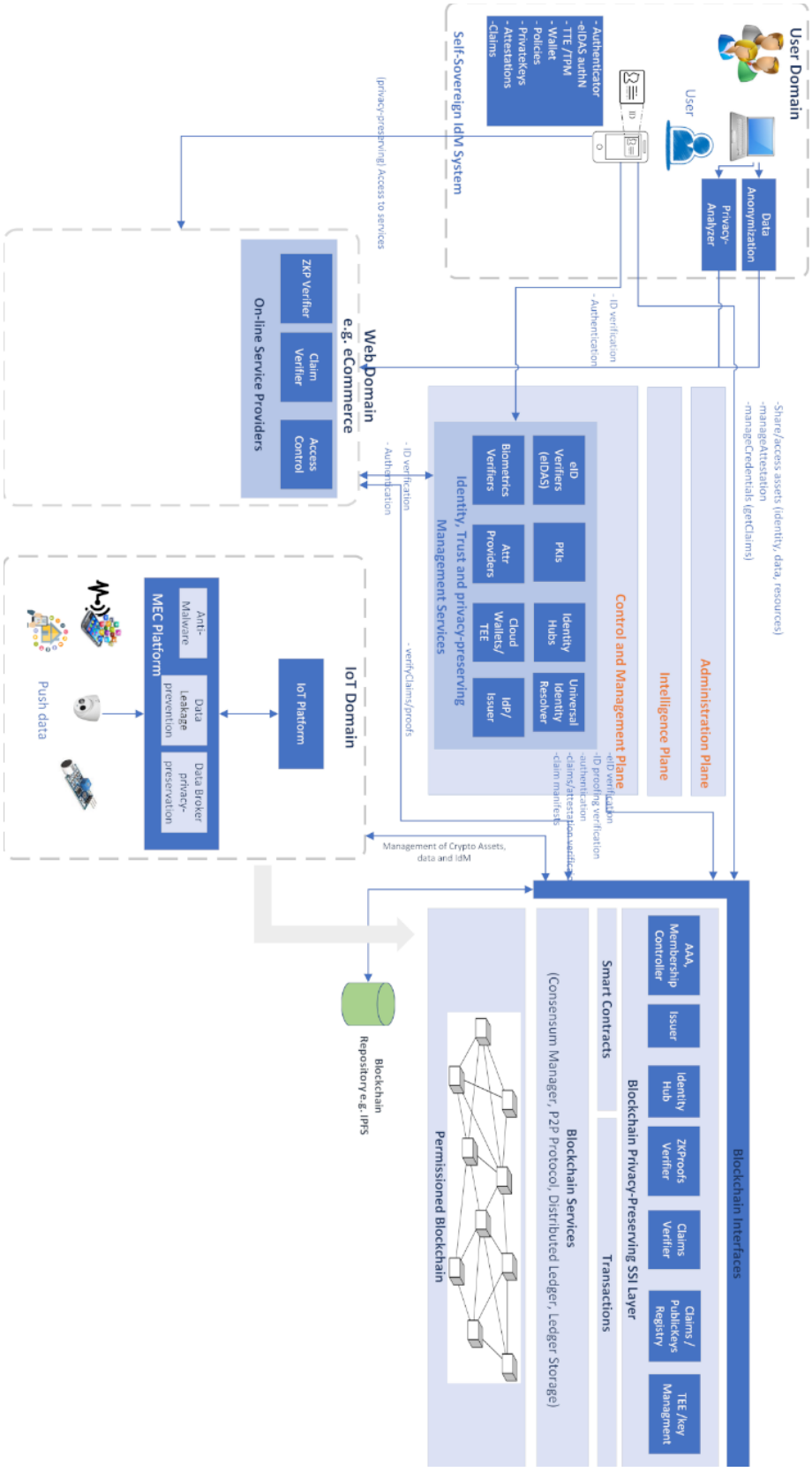

Figure 2: CyberSec4Europe Privacy-Preserving Functional Architecture

In the Control and Management plane of the CyberSec4Europe architecture, the **Identity and Privacy-preservation Services** plane includes the building blocks considered in the CyberSec4Europe Privacy-Preserving Architecture devoted to enabling privacy-respectful authentication based on the provision of anonymous credential systems and privacy-preserving identity management services, some of which rely on the use of secure distributed ledger technologies such as a Blockchain to provide a self-sovereign identity (SSI) model. The Identity and privacy-preservation Services also includes mechanisms for privacy-preserving computation technologies to reduce information leakage during the computations in the managed domain, thereby verifying that the systems comply with the users' privacy policies. Those privacy-preservation services can be run in the Cloud so that the architecture includes confidentiality-preserving and end-to-end secure sharing of sensitive data in the cloud among stakeholders using, for instance, secret sharing technologies. Besides, the architecture considers the privacy brokerage aiming at enhancing user trust in public cloud storage systems, guaranteeing data confidentiality and improving availability. The Privacy-preserving architecture includes functional building blocks for confidential and privacy-preserving storage that can employ techniques such as secret sharing to anonymize personal information during data analysis processes. Similarly, it also embraces privacy-preserving mechanisms for analyzing data from potentially different stakeholders in a way that gives high authenticity guarantees on the computation's result, while protecting the confidentiality and privacy of the input data, and ensuring data integrity.

On top of that, the Privacy-Preserving Architecture includes several mechanisms that use Trusted Execution Environments (TEE) for different purposes that range from securely storing and managing secret keys to remote anonymous attestation even in the presence of compromised hardware. The building blocks can be used on the virtualized applications in the Cloud or directly installed in the user domain.

In the **User Domain**, the privacy-preserving architecture encompasses the wallets and TEE needed to maintain securely protected the credentials and manage key material obtained during the issuance and enrollment in diverse identity providers. The user domain is exemplified either with user mobiles, or software for desktop browsers. It contains the client-side software needed to perform authentication against service providers, eIDs-based authentication, and run protocols for proving privacy-Attribute Based credentials and claims (including zero-knowledge proofs).

Therefore, the user domain plays the role of *Recipient* and *Prover* in the privacy-ABC model. To this aim, user domain interacts with diverse online identity services (including IdPs, Attribute providers, PKIs, biometric verifiers, eID verifiers) placed in the *Control and Management Domain* of CyberSec4Europe architecture. In addition to credentials, the user domain needs to manage the attestations obtained from diverse attributes and identity providers, and short tokens obtained from IdPs (for single sign-on in Service Providers). The user-domain might also include ID-Proofing mechanisms, with client-side biometrics software needed to authenticate in biometric servers as second authentication factor.

Furthermore, the user-domain considers the data anonymization building blocks to share in a privacy-preserving way data in transactions online and between organizations using diverse different privacy models (e.g., the k-anonymity, k-Map, Average risk model, among others). In addition, in the user-domain, the privacy-analyzer allows reducing the attack surface preventing privacy breaches when sensitive personal data are managed.

Decentralized authorization, privacy-preservation and distributed access control are also important features considered in this architecture. In the **Blockchain privacy-preserving SSI Layer**, this is achieved by means of building blocks that are aimed at making blockchain technologies and consensus mechanisms more

scalable, efficient, guarantying on-chain transactional privacy. Besides it includes building blocks for modifying transactions (fine-granular rewriting)already present in the blockchain in a limited and traceable manner, which may be important for legal reasons.

The architecture considers privacy-preservation of identities and personal data in blockchains. To that aim, and following the *identity.foundation* (DIF)[2] standards and specifications, the architecture features the building blocks needed for the creation, resolution, and discovery of decentralized identifiers (DID identifiers [3]) and names in heterogeneous blockchains through resolvers. In addition, the Identity hubs keeps secure, encrypted, privacy-preserving personal data storage and computation of data. Where the resolver services links user's DID's employed in blockchain with Identity Hubs. The blockchain Identity services provide means to create, exchange, and verify crypto credentials and claims in a decentralized identity ecosystem with the User, following a Self-sovereign identity management model. Besides, the blockchain identity services might rely on authentication protocols open standards and cryptographic protocols, including DIDs and DID Documents.

Another group of solutions is intended to enable privacy preservation in Cloud computing environments as well as its extension towards the user side with **Edge computing**. The Privacy-Preserving architecture provides building blocks for secure data storage and processing in public clouds. In particular, it considers distributed data storage and privacy-preserving analytics as well as mechanisms for compliance with the provisions of GDPR regarding interoperability and cross-border data transfers.

The Edge is considered in this architecture as a security and privacy enabler especially for the **IoT domain**, where devices are typically extremely resource-constrained and may be subject to compromise or interference. In this respect, the proposed architecture includes data broker for both handling sensitive data according to a set of privacy policies as well as tools for monitoring and sanitizing IoT devices for reducing the attack surface in this domain. Likewise, the privacy-preserving architecture considers the privacy-preserving middleware and software for the IoT domain aimed to ensure secure and authenticated communication channels between IoT devices. The managed domain in the global IoT architecture of figure 1 can be also instantiated through processes related to **Web domain** (e.g. eCommerce) in the CyberSec4Europe privacy-preserving architecture. In this case the Web domain is comprised of set of functional components needed for the Service providers to authenticate their users, verify claims and privacy-preserving crypto-proofs (e.g. Zero-knowledge proofs). These service providers play the role of *Verifier* in the privacy-ABC model.

Finally, our privacy architecture also considers the application of security and privacy by design mechanisms by introducing components for GDPR-compliant software development as well as analyzing the information leakage produced by some particular privacy solutions.

[2] DIF Identity Foundation. https://identity.foundation.org
[3] Decentralized Identifiers (DIDs) v1.0. W3C. November 2019. https://w3c.github.io/did-core/

3 Identity Management and Authentication Solutions

Identity management, also referred to as identity and access management, is a framework for ensuring that the proper people have access to the proper technology resources in a network or system. Secure, scalable, and usable identity management solutions are, therefore, a requirement in many environments. Furthermore, as legal regulations as well as business requirements became more demanding, also the complexity and criticality of identity management systems increase.

In the following we present a variety of advanced identity and access management systems, with a special focus on self-sovereignty, privacy, and user-centricity.

3.1 Cloud-Based Anonymous Credential Systems

Anonymous (attribute-based) credential systems (also known as ABCs), first envisioned by Chaum [1,2], have gained significant attention in recent years. In such a system, a *user* receives a *credential* on her attributes from an *issuer* (e.g., a public authority or a cloud provider) certifying personal information such as name, date of birth, nationality, or similar. The user can then *present* this credential to *service providers* (or *relying parties*, e.g., a cloud service) in a way that allows her to control, on a fine granular level, which information to disclose. For instance, the user may decide to reveal her name but not her nationality, and prove that she is at least 18 years old, without disclosing her precise birth date. By doing so, the user receives high privacy guarantees, also against colluding issuers and service providers, as even in this case authentication sessions originating from the same user cannot be linked. On the other hand, the relying party receives very high authenticity guarantees, in the sense that it can be ensured that the received information was indeed certified by a given issuer.

Attribute-based credentials have been successfully proven useful in many scenarios and are successfully deployed in different real-world scenarios, cf., e.g., IRMA[4]. Furthermore, open-source implementations have been developed in different research and development project such as ABC4Trust[5].

However, one drawback of classical anonymous credential systems is that they are computationally expensive on the end-user's side, rendering them too inefficient for many embedded devices such as smart cards or in the Internet of Things context.

This drawback was recently addressed by Krenn et al [3,4], where an ABC system allowing for outsourcing the overwhelming part of the computations to a semi-trusted cloud provider (referred to as *Wallet*) has been presented. More precisely, in this system, the logic of the presentation flow is changed in the sense that the user merely uses her device (e.g., laptop, smart card, etc.) to prove that she is actively participating in the protocol. All other computations can securely be outsourced to a cloud service, while still guaranteeing privacy of the user also against this service. Using advanced encryption mechanisms dubbed *proxy re-encryption* [5] and flexible signature schemes called *redactable signatures* [6], the user only needs to store encrypted versions of her credentials on the cloud service. Now, when authenticating to a relying party, the user generates an ephemeral *re-encryption key* that can be used by the Wallet to translate the stored

[4] https://irma.app/
[5] https://abc4trust.eu/

encryptions into ciphertexts for the relying party - without itself ever learning information about the underlying attributes. Simultaneously, using the redaction functionality, the Wallet blanks out all the attributes that the user wishes to keep private. By following clear key-separation principles, it can then be guaranteed that not even in the case that the Wallet colludes with issuers or relying parties, the user's data-privacy can be broken, but at most unlinkability of authentication sessions could be affected.

The efficiency and usability of cloud-based anonymous credential systems have been prototypically demonstrated; yet, the service cannot yet be used "out-of-the-box" using publicly available libraries or implementations.

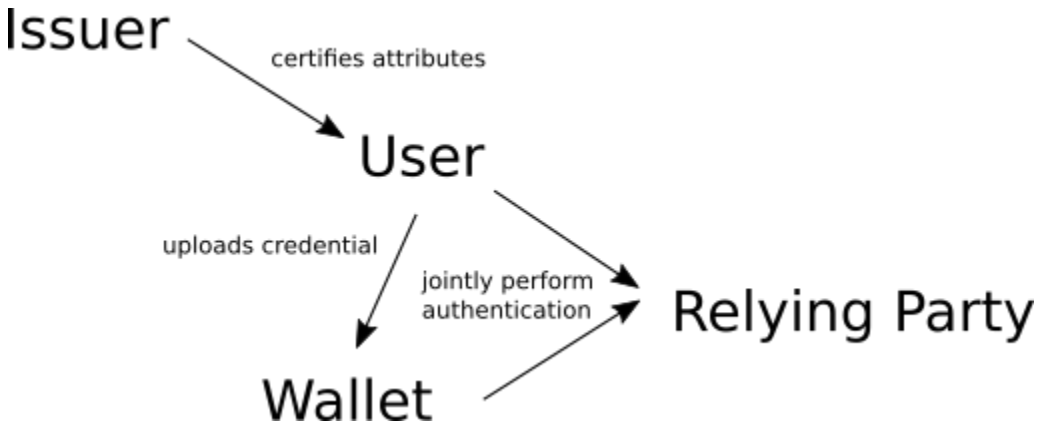

Figure 3: Parties and processes of cloud-based ABC systems

Figure 3 gives an overview of the actors and processes in a basic cloud-based attribute-based anonymous credential system.

Besides these basic functionalities, additional features like revocation or multi-show presentations can be added analogously to traditional ABC systems.

Anonymous credential systems have been invented and developed within a series of European funded projects, including the FP6 project *PRIME*[6], the FP7 projects *PrimeLife*[7] and *ABC4Trust*, as well as the H2020 project *CREDENTIAL*[8], where especially in the latter the concept of secure cloud-based ABCs has been introduced. This concept has recently been refined in a follow-up work within CyberSec4Europe [4] and will serve as a basis for further investigations regarding efficiency, proving of predicates (e.g., proving claims like “older than 18” without revealing the birth date), and metadata privacy.

[6] https://cordis.europa.eu/project/rcn/71383/factsheet/en
[7] https://cordis.europa.eu/project/rcn/85453/en
[8] https://credential.eu/

Full name	Acronym	Lead partner	TRL
Cloud-Based Anonymous Credential Systems	eABCs	AIT	6
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.3	IDM-SP02, IDM-SP05, IDM-SP06, IDM-SP07, IDM-SP08	https://credential.eu https://eprint.iacr.org/2019/1061.pdf	

Table 2: Metadata: Cloud-based Anonymous Credentials

3.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Attribute-based credentials allow for privacy-preserving authentication and access control, e.g., towards online service providers. Furthermore, they allow for cryptographically strong user authentication.

In the privacy-preserving functional architecture presented in Figure 2, cloud-based anonymous credential systems therefore belong to the following components: the user's device is located in the user domain, as part of the self-sovereign IdM systems component; the issuer as well as the Wallet are part of the IdP/issuer component in the identity and privacy-preserving services component of the control and management plane; finally, the relying party belongs to the access control and claim verifier components in the web domain.

3.2 Self-Sovereign & Privacy-preserving SS-PP-IdM

Emerging privacy-preserving proposals for blockchain [7,8,9,10], and platforms, such as uPort[9] or Sovrin[10], propose enhanced decentralized ledgers that empower users with mechanisms preserve their privacy in their digital transactions.

The management of user's related information in permissioned blockchains is being characterized by its privacy-preserving nature. With the rise of blockchain, Identity Management (IdM) systems are switching from traditional web-centric approach or identity federation approaches, towards the self-sovereign identity (SSI) paradigm [11]. Self-sovereign identities allow citizens to take control of their data in any-time in any online situation. Under this approach, user personal data is no longer kept in raw in third-parties services, neither in Service Providers or Identity Providers, and information regarding transactions and interactions of users in services can be anonymized. It avoids that third-parties can leak personal data, and, in the worst case, become a potential source of other, more important, risks, such as identity-related cybercrimes (e.g. identity-theft).

Identity Management based on Self Sovereign Identities (SSI) focuses on providing a privacy-respectful solution, enabling users with full control and management of their personal identity data without needing a third-party centralized authority taking over the identity management operations. Thus, citizens are not anymore data subjects, instead, they become the data controller of their own identity. This is, they can determine the purposes, and ways in which personal data is processed, as they manage directly their personal data during their online transactions.

[9] https://www.uport.me/
[10] https://sovrin.org/

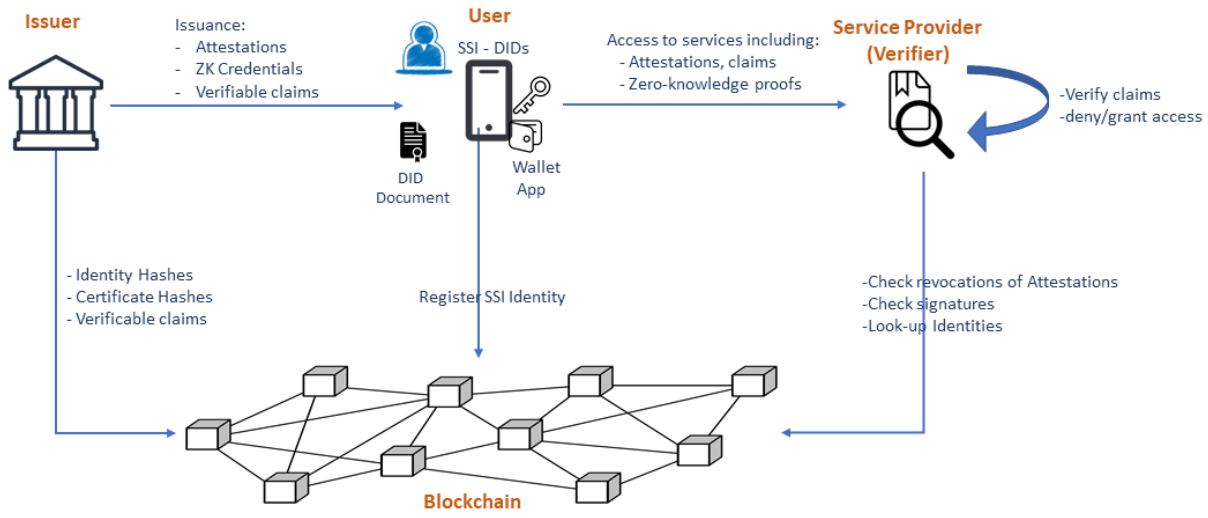

Figure 4: Representation of the SS-PP-IdM asset

SSI has been brought forward, as it is being materialized through blockchain, which facilitates the governance of the SSI system, increasing the performance to Internet scale and enabling the accessibility of identities to everyone. Blockchain enables sovereignty as users can be endowed with means to transfer digital assets, including user decentralized identifiers (DID) [12], DID documents, identity attributes, verifiable claims and proofs of identity [13] (including ZKPs), to anyone privately, without rules in behind, which ultimately increases the global democracy in the world. In this sense, latest blockchain solutions like uPort or Sovrin make use of DLTs, along with user-centric and mobile-centric approaches, and therefore, empowering users to maintain securely protected (in their mobile wallet) the needed crypto-credentials. In this scenario, the blockchain acts as distributed and reliable identity verifier, providing provenance and verifiability of identities. Thus, the ledger provides a cryptographic root of trust, which facilitates identity management without external authorities. In this sense, Wagner et al [14] have recently described the main SSI concepts on blockchain and the road ahead.

These SSI concepts, their main processes and associated entities are depicted in Figure 4 as envisioned in [15]. As it can be seen, a User (holder) might have DIDs and obtain verifiable claims and credentials from the Issuer authority, in a user-centric way, using his smartphone whereby the private-keys are kept securely protected in the wallet. To increase the privacy-preserving capabilities in the SSI model, the user can be empowered with means to present Zero-Knowledge crypto proofs against a Service Provider acting as verifier that checks in the blockchain the attestations and signatures. Indeed, SSI systems in permissioned blockchains can be leveraged with additional privacy-preserving capabilities, by using oblivious and distributed privacy-preserving crypto-solutions (using threshold cryptography) [16] where the IdP role is split up into several authorities, so that a single entity is not able to impersonate or trace user behaviors.

Full name	Acronym	Lead partner	TRL
Self-Sovereign & Privacy-preserving Identity Management	SS-PP-IdM	UMU	2
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.3, T5.6	IDM-SP03, IDM-SP05, IDM-SP06, MD-SP04, MD-OP03	n/a	

Table 3: Metadata: Self-Sovereign & Privacy-preserving Identity Management

3.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The presented technology is a fundamental building block for realizing the ambition of a privacy-preserving blockchain-based SSI layer. The software components of the SS-PP-IdM asset maps directly to building blocks to the CyberSec4Europe architecture. Thus, the user in the asset maps to the user domain in the CyberSec4Europe architecture, whereas the Verifier services map to Services providers in the CyberSec4Europe Web Domain. Besides, the Issuer in the SS-PP-IdM maps either to the identity services in the control and management plane or to the Issuer placed in blockchain plane.

3.3 SPeIDI – Service Provider eID Integration

In the context of Digital Single Market[11] promoted by the EC, the eIDAS regulation[12] (entered into force on 29 September 2018) is boosting and facilitating the use of electronic ID to European citizens for cross-border authentication. This situation allows citizens, businesses and organizations securely access, not only to online services provided by public administration, but also to those provided by the private sector.

The eIDAS network is the infrastructure developed for connecting the different national eID schemas. This network enables both public and private service providers connecting their services, allowing cross-border transactions.

With the aim to facilitate the integration of the digital services provided by the private sector the SPeIDI asset (Service Provider eID Integration) is developed by Atos. SPeIDI derives from LEPS[13] project and it aims at integrating online services with eIDAS infrastructure to European eID use for authentication scenarios when a user strong authentication is needed for securing the access to those services. This connectivity eIDAS-based solution is intended to provide a hub or proxy service between the private SP domain and the European country eIDAS nodes for securely accessing to the e-services using the eID issued by any European country. Based on the eID building block[14] provided by CEF following the eIDAS technical specifications, including signing, encryption and the SAML 2.0 standard. SP connection is based on a simple API based on JWT.

The European eIDAS Regulation[15] timeline establishes that from 29 September 2018citizens and companies, equipped with a notified eID mean, will be able to access digital public services in every MS. Regarding the private domains such as the financial sector the use of eID means is not widely spread enough

[11] https://ec.europa.eu/digital-single-market/en
[12] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2014.257.01.0073.01.ENG
[13] http://www.leps-project.eu/
[14] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eID
[15] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG

in most of the countries, even though the EC is trying to increase the use of eID for accessing these online services. A recent study published by the EC indicates that the adoption of eIDAS solutions is still low by the private sector, and namely by the SMEs[16]. The Connection European Facility (CEF) is offering the eID building block[17] (among others) for easing the use of eID by the different service providers

Some European projects (e.g. LEPS, ESMO) have demonstrated how a strong cross-border authentication, based on eIDAS network, increased not only the security and trustworthiness of the users, but the online service provider as well. Postal, stock market and educational digital services have leveraged this kind of authentication. The SPeIDI eIDAS connector was originally created for integrating private online services with the eIDAS network through the Spanish eIDAS node (following a SAML protocol). In this demonstrator a further step is made by allowing the integration with other country eIDAS node such as the French node by using OpenID Connect protocol. The following innovations will be achieved to address not only the specific requirements of the health stakeholders but of stakeholders from different domains, such as banking, educational or financial sector:

- Include new access protocols (e.g. OpenID Connect) to the eIDAS tool supporting the different stakeholders;
- Connect eIDAS tool with relevant attribute providers for several domain stakeholders;
- Extend authentication not only to a natural person but legal person when is supported by the country eIDAS infrastructure.

Once the use of eID schemas as secure authentication mean is established by the public sector, the industry sector has the opportunity for adopting the use of secure eID for accessing cross-border business services reduces ID theft and fraud. The benefits are for both, the user and the SP. The user access to a genuine and trusted website and the business is offering a secure and trusted service which can lead to enlarging user base (foreign European citizens are accepted as users). Businesses will reduce costs and increase time savings.

The use of eIDAS tool will help to control the access to those services requiring a high level of assurance, such as financial services or access to personal and sensitive data.

The modular design of this Java-based solution provides a simple and reusable component, which is SP infrastructure and client programming language independent, able to connect with different SP services in the same or from different domains. eIDAS authentication could be used by online services requiring to prove of user identity using strong authentication mechanisms (country eID card) and different LoA for accessing e-services.

Security is assured by using secure protocols to communicate with the SPs and the eIDAS network.

The main functionalities provided by SPeIDI are the following:

- Easy access to the eIDAS network from the SPs using standard protocols.
- Strong cross-border authentication using eID cards issued by European MS.

[16] https://op.europa.eu/en/publication-detail/-/publication/0627f219-5044-11e9-a8ed-01aa75ed71a1/language-en
[17] https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/eID

- HTTPS connection between SP and eIDAS solution and secure communication using JWT technology for user data transmission.
- Mapping/Translator service between the SP and the eIDAS network, translate SAML to common standards (JSON) and the way around.
- Support multiple e-Services in one domain.
- Docker deployment.
- Compliant with eIDAS specifications and the GDPR regulation.

Figure 5 shows the basic modules the SPeIDI service comprises and the main interactions with third parties (SPs and eIDAS network).

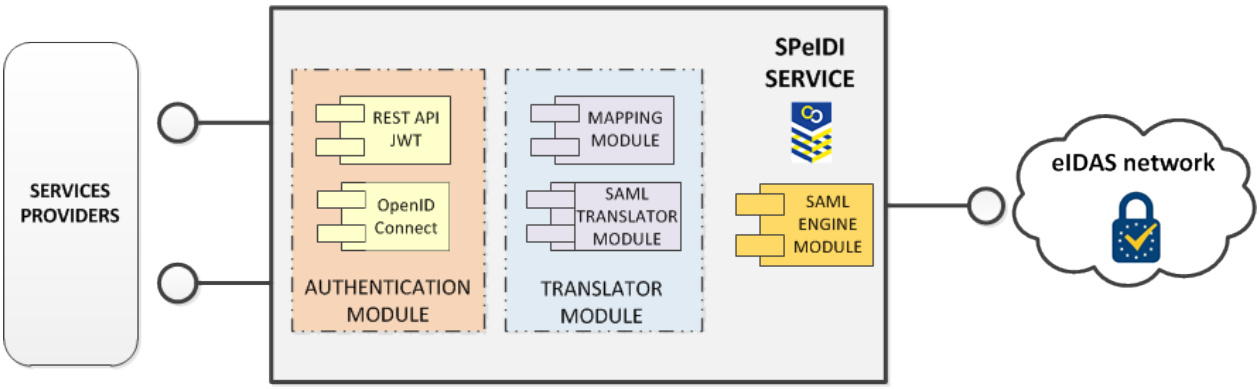

Figure 5: SPeIDI main modules and connections with third parties

SPeIDI asset has been developed within the Atos Blockchain, Identity & Privacy Unit and demonstrated in the private sector such as the financial market, the stock exchange market and postal services (LEPS CEF project[1813]), and in government and academic domains in previous projects (FIDES EIT project, and FutureID and Strategic FP7 projects). Hence, ATOS aims to elevate the underlying technical foundation of the described component from Technology Readiness Level 6 (technology demonstrated in relevant environment) to TRL 7 (system prototype demonstration in operational environment) at the end of the project.

The following innovations will be achieved to address not only the specific requirements of the health stakeholders but of stakeholders from different domains, such as banking, educational or financial sector:

- Include new access protocols (e.g. OpenID Connect) to the eIDAS tool supporting the different stakeholders;
- Connect eIDAS tool with relevant attribute providers for several domain stakeholders;
- Extend authentication not only to a natural person but legal person when is supported by the country eIDAS infrastructure.

[18] http://www.leps-project.eu/

Full name	Acronym	Lead partner	TRL
Service Provider eID Integration	SPeIDI	ATOS	6
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.4, T5.5, T5.6, T5.7	OB-SP01, OB-SP02, OB-SP09, OB-SP10, OB-SP26, OB-SP27, OB-SP28, OB-U03, OB-LR09, SCH-SP01, SCH-SP02, SCH-LR01, IDM-SP06, IDM-SP10, IDM-MP01, IDM-LR01, IDM-LR03, IR-SP01, IR-LR03, MT-SP01, MD-SP07, MD-OP02, SMC-F01, SMC-F03, SMC-SP01, SMC-SP03, SMC-SP10, SMC-SP16, SMC-SP21, SMC-LR02	http://www.leps-project.eu/node/345	

Table 4: Metadata: Service Provider eID Integration

3.3.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The SPeIDI component is an asset which allows secure and strong user authentication to digital services, belongs to the groups of Identity & Trust Management services embedded in the Control and Management layer on the CyberSec4Europe architecture. The SPeIDI service eases European citizens the use of his/her eID against their origin country IdPs for cross-border authentication purposes. leveraging the eIDAS network infrastructure. In the context of the T5.6 Medical Data Exchange demonstrator, SPeIDI will provide a strong cross-border user authentication mechanism for facilitating and securing access to the Dawex data exchange platform.

3.4 Mobile Privacy-Attribute Based Credentials (Mobile p-ABC)

In the context of p-ABC systems, mobile compatibility is important. Based on Idemix Anonymous Credential System and the implementation in the ABC4Trust project, Mobile p-ABC asset offers a minimal disclosure of personal information, through the use of zero knowledge proofs, for Android devices. This allows users to use their Android smartphones to present those ZK-Proofs against identity providers. This asset is already available in Y1 of CyberSec4Europe.

Thanks to the Android implementation of p-ABC systems, whenever users want to make transactions online with them, the solution provides an advanced mechanism of authentication and management of their personal data in a privacy-preserving way. The solution requires cooperation between service providers, identity providers and users.

In the user side, the software deals with the interaction with the service that the user wants to access and with the credential management. The issuance process is a two-round interaction where the User submits a request containing his attributes and the Issuer certifies the fact that the User has the claimed attributes by returning the credential that is stored in the user device (wallet).

Once the User has the Idemix credential stored in his wallet, it can use such a credential to derive partial identities for a privacy-preserving authentication mechanism when accessing to services. The wallet is a secure storage which encrypt, protect and operate with the credential. Protection methods such us PIN, patterns or fingerprinting are available.

Once the identity/ies is/are created, the user can select one to be used during the login in, for example, in a eCommerce scenario. The web page of the eCommerce will communicate the access policy to the client app, for example, with a QR code in the screen to be scanned by the client. Then the user should short press in the wanted identity and scan de QR code. After reading the QR, the app will display to the user the different attributes included in the identity selected previously. Finally, the user should click on each attribute she/he wants to share with the service provider of the eCommerce.

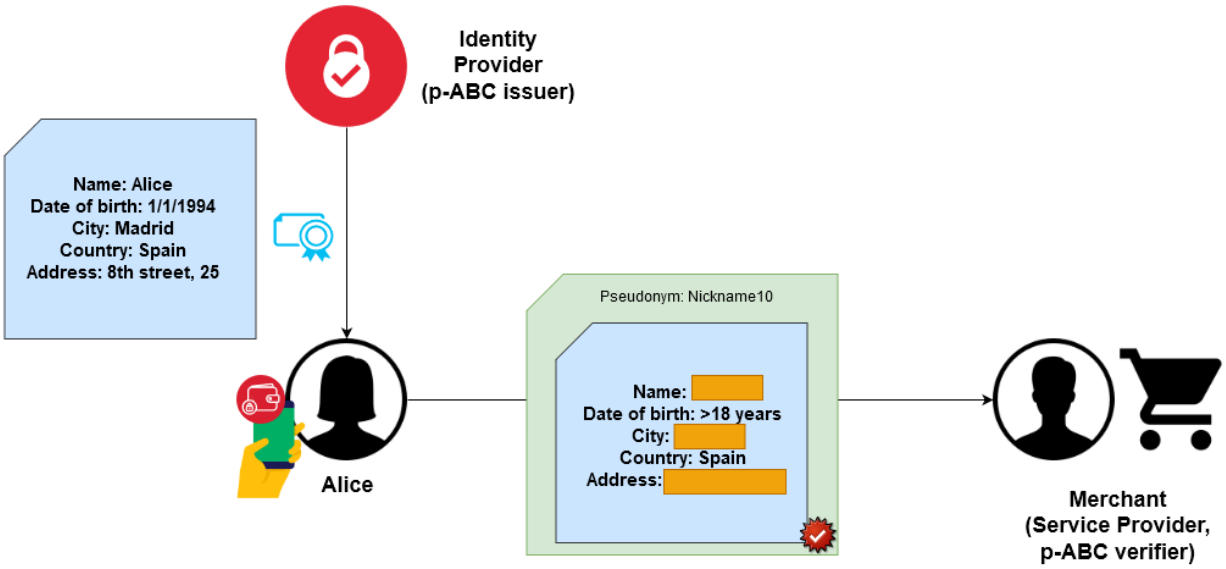

Figure 6: Mobile p-ABC

Projects such as ARIES[19] [17] have addressed the use of mobile p-ABC in the face of existing limitations such as the absence of standards and excessive complexity when it comes to deployment.

Full name	Acronym	Lead partner	TRL
Mobile Privacy-Attribute Based Credentials	Mobile p-ABC	UMU	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.3	IDM-SP02, IDM-SP05, IDM-SP06, IDM-SP07, IDM-SP08	https://doi.org/10.1016/j.future.2019.08.017	

Table 5: Metadata: Mobile Privacy-Attribute Based Credentials

3.4.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

In the privacy-preserving architecture, the mobile p-ABC asset maps to the component of the user domain that runs in the user's smartphone. The mobile wallet performs internally the crypto operations for obtain and keep securely protected the ABC credentials retrieved from the Issuer that is located online, in the

[19] https://www.aries-project.eu/

control and management plane of the architecture. The service provider (e.g. a Merchant), acts as p-ABC verifier and is located in the Web-domain.

3.5 eIDAS Browser App

The eIDAS browser app, is an Android application that empowers users with means to authenticate through eIDAS using their official European physical eIDs. The app connects with the eIDAS infrastructure and can be integrated with ATOS's SPeIDI asset. Namely, eIDAS browser allows any Spanish user to employ their national identity card, also known as DNIe, with NFC capabilities with their mobile phone to access any eIDAS and Spanish's CLAVE enabled service. This app was developed in the scope of LEPs European research project.

The app is indeed a browser that allows the user to surf internet and replaces the standard certificate selection mechanism provided with the operating system with one that includes the NFC DNIe mechanism for eIDAS. At that point the user can associate his physical DNIe card with the app by means of the CAN number. This CAN number is unique and hard coded/printed in the card itself. Then each time an authentication is requested the password for the user in addition with the card is needed. The user would then put card and phone together, insert the password and authentication will occur.

The app can access any web page but for a good user experience pages with adapted content are recommended. To simplify the use of the system during the project, a welcome page with production services urls associated with buttons has been added, avoiding the need of introducing urls by hand, although possible. The app requires an android device with NFC as well as Spanish DNIe with NFC support. It has been tested Nexus 5 terminal with Android 6.

Full name	Acronym	Lead partner	TRL
eIDAS Browser App	eIDAS browser	UMU	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.3	IDM-SP03, IDM-SP04, IDM-SP10	https://gitlab.atica.um.es/emtg.um.es/eu_leps_eIDASbrowser	

Table 6: Metadata: eIDAS Browser App

3.5.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

In the privacy-preserving architecture, the eIDAS browser asset maps to the component of the user domain that runs in the user's smartphone, and interfaces with the eIDAS services (such as the eIDAS verifier) that are running online, in the Identity & Trust Management services, which are in turn, scoped in the management plane of the global CS4EU architecture. The app connects could be integrated with SPeIDI asset described earlier. The eIDAS browser app could be exploited in the "Privacy-preserving identity management" CS4EU Pilot being developed in Task 5.3.

3.6 Password-less authentication system

Most web applications have authentication process that rely on the password paradigm. It is evident that a password can be considered secure when it contains 20 characters or more, is complex (is comprised of

alphanumeric characters, symbols and non-dictionary words), is only stored in the brain of the user, is used only in one application and is changed frequently. As the number of accounts each user maintains has greatly increased in the last few years, users are having a hard time memorizing and managing all these passwords. To solve this password overload problem, users have come up with solutions that directly affect the security of their accounts and the privacy of their data; they either simplify their passwords to be easy to remember, or reuse the same password on different services, or store their passwords in a "secure" place, on paper or using a password manager. At the same time, passwords are targets of multiple attacks, as they can be leaked, key-logged, replayed, eavesdropped, brute-force decoded and phished.

In this context, we develop a secure and user-friendly password-less authentication solution that integrates device centric authentication methods (e.g. biometrics). The solution relies on the users' biometrics to perform the authentication. Particularly, the password-less authentication system will contain a *FIDO client application* and a *FIDO server*. The *FIDO client application* is an android application that communicates with the *FIDO server* in order to authenticate the user in a service. The password-less authentication system performs the following actions:

- Registration, in which it is performed the authentication key agreement between the *FIDO server* and *FIDO client application*.
- Authentication, in which the user uses his biometric to unlock the private key and start the authentication procedure. The authentication process is presented in the next figure.
- Deregistration, in which the user's account and any related material is deleted from both *FIDO client application* and *FIDO server*.

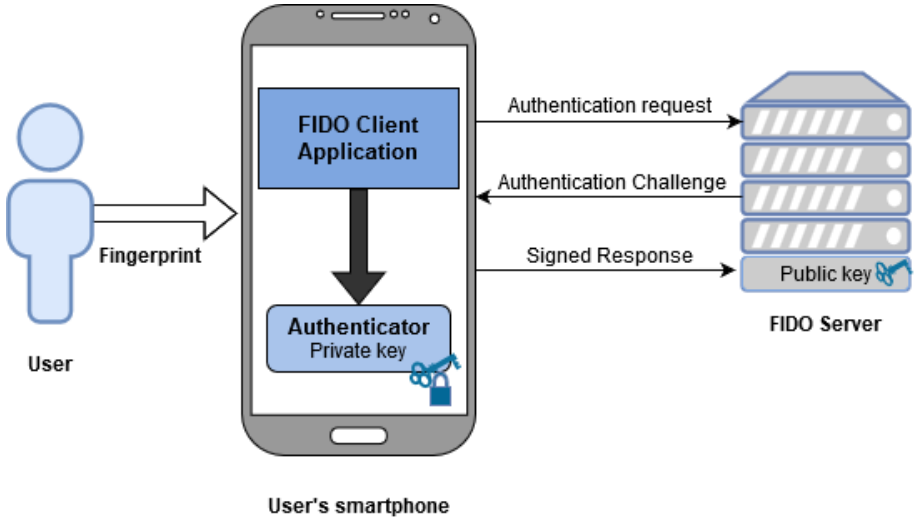

Figure 7: Password-less authentication

The communication between the *FIDO client application* and the *FIDO server* is based on the FIDO protocols. The user's fingerprint is used to authenticate him/her in the *client application* and then, the application sends a token to the server.

The presented building block resulted from the H2020 project ReCRED[20], and will be adopted towards the needs of CyberSec4Europe depending on the needs especially of T5.3, most likely during the second demonstration phase.

Full name	Acronym	Lead partner	TRL
Password-less authentication	Password-less authentication	UPRC	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.3	IDM-SP-02, IDM-SP-04, IDM-LF-01, IDM-LF-02, IDM-MP-01	https://www.recred.eu/ https://fidoalliance.org/	

Table 7: Metadata: Password-less Authentication

3.6.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The *FIDO client application* asset maps to the component of the user domain that runs in the user's smartphone. The authentication with the user's biometrics is performed internally in the smartphone and is not exchanged with the server. The *FIDO server* is located online, in the control and management plane of the CyberSec4Europe architecture.

3.7 Research Challenges Identified in WP5

So far, the following research potential for the listed cybersecurity building blocks has been identified within WP5:

- For cloud-based ABC systems, it is still an open challenge to prove predicates about attributes, i.e., it is currently only possible to reveal or redact an attribute, but not to prove that a certain relation (e.g., "older than 18 years") is satisfied. Such a feature would necessarily require joint computations by the Wallet and the user; however, it is unclear how this can be done in a way that minimizes the computation on the user's side.
- Also for cloud-based ABCs, it would be interesting to leverage blockchain technologies to increase trust in key material, etc. While it is clear how this could be done from a scientific point of view, there are potential challenges on the engineering side. Regarding privacy-preservation in distributed ledgers, these systems are subject to different privacy issues such as transaction linkability, on-chain data privacy, or compliance with privacy regulations (e.g. GDPR). There is a need to integrate and adapt privacy-preserving solutions like Anonymous Credentials Systems in distributed DLTs, following a Self-sovereign identity management approach. For instance, endow blockchains with Non-Interactive Zero-Knowledge Proofs (NI-ZKP) to protect privacy in blockchain transactions. In addition, it would be interesting to increase privacy-preserving capabilities against IdPs in permissioned blockchains, using novel crypto-solutions (e.g. using threshold cryptography) to split up the role of IdPs so that they cannot trace their users.
- Although the adoption of eIDAS by the private sector is growing during the recent years, and several implementations have been developed, the wide uptake of trust services is one of the main challenges the eIDAS framework is facing. As indicated above, leveraging disruptive technologies like blockchain could be helpful for increasing trustworthiness.

[20] https://www.recred.eu

4 Security and Privacy in Edge Computing

This section introduces tools devised for protecting information and assets in the Internet of Things (IoT) domain by taking advantage of Edge Computing technologies. The first tool described in this section is a data broker, called Privacy Monitor, which acts as intermediary between IoT devices and data consumers thereby enhancing data protection by enforcing privacy policies. The second tool is an intrusion detection and prevention system for the IoT, called AntibIoTic, which monitors the behavior of IoT devices in order to detect and sanitize corrupted devices.

4.1 Edge-Privacy

The Internet of Things (IoT) has seen in the Cloud a great ally for realizing real-world deployments as its storage capabilities and computing power allows us to hold great amounts of data and extract knowledge from these data by performing big data analytics. However, as the IoT matures and more devices are connected to the Internet, the amount of data that needs to be stored and processed grows significantly. This not only implies that the capacity of the servers in the Cloud needs to grow but also that more data travels from the IoT plane to the Cloud plane, wasting a great amount of bandwidth on the communication channels that connect both end-points. Not only that, the data are required to travel to a relatively distant location, possibly thousands of kilometers away from the source, to be processed and then act upon the IoT plane. This mode of operation introduces a significant delay that renders the deployment of some application scenarios with real-time requirements impossible. The Edge Computing paradigm emerges to address the aforementioned limitations [18].

In Edge Computing, data are processed at the edge of the network, close to the area where they are generated, by geographically distributed mini-clouds called edge devices. Since edge devices are more hardware-constrained than cloud servers, only current data is processed and stored in edge devices while historical and aggregated data are usually sent to the Cloud. However, the Edge is assumed to be relatively autonomous and not completely dependent on the presence of an online Cloud. In this paradigm shift, security services are to be revisited. One specific and important service strongly connected with data generation, storage and computation is *privacy*. Users and regulators demand that privacy is taken into consideration when deploying services that may use and process sensitive data [19].

Despite being a novel technology, there has been research on different aspects of privacy in Edge computing. Most privacy research has focused on the development of privacy-preserving authentication schemes [20] and access control [21] but also on topics related to data processing (i.e., data aggregation [22] and data analytics [23]) as well as data and task offloading [24]. However, little work has been on the definition of mechanisms for allowing the users to define their own policies for controlling the access to the data collected at the edge.

As such, we are currently in the process of designing a *Privacy Monitor* for the IoT based on edge computing technologies. The Privacy Monitor is a service to be deployed in edge devices which acts as a broker for all entities querying, pulling or pushing data from or to the edge infrastructure. The operation of this service will be defined by a set of rules determined by the privacy policies. The idea is to allow owners of IoT devices to define the privacy policies for controlling access to their data with a fine grain granularity. and possibly dependent on contextual information. These policies will determine how data is handled from the moment it is received at the Edge. In particular, we consider the possibility of having policies for controlling the amount and level of detail of the data to be stored by edge devices as well as policies for controlling the

entities that may access these data and with which level of detail. The filtering functionality (e.g., k-anonymity, homomorphic encryption, attribute-based encryption, etc.) is not expected to be implemented by the Privacy Monitor but instead borrowed from other elements of the CyberSec4Europe Privacy-Preserving Architecture, for example from the Data Leakage Prevention asset. A very high-level description of the expected functionality of the Privacy Monitor component is depicted in Figure 7.

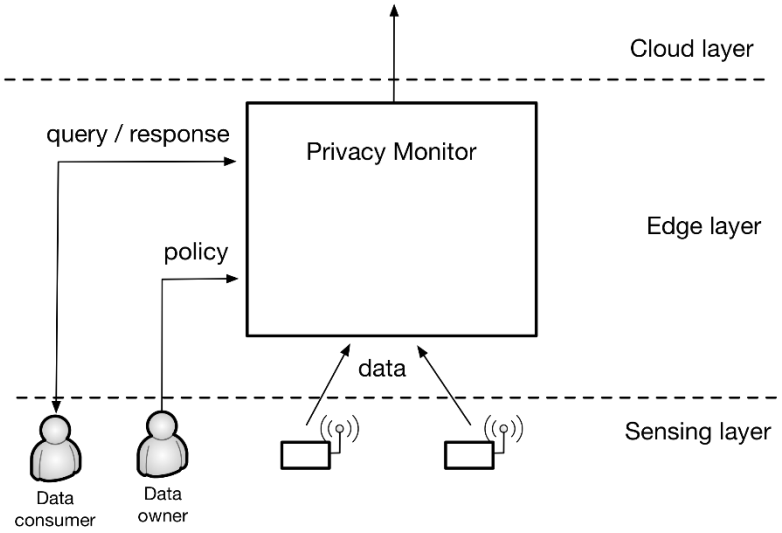

Figure 8: Functionality of the Privacy Monitor for Edge Computing

For the design of the Privacy Monitor, we will consider standard policy languages and architectures, such as XACML[21] for attribute-based access control. Nonetheless, it may be necessary to extend existing solutions or devise new ones for incorporating the particular features and requirements of edge computing platforms and the Internet of Things. Since we are dealing with highly distributed and dynamic environments, where data may move both horizontally, from edge device to edge device, and vertically from the edge to the Cloud, we may need to have a Privacy Monitor that is also distributed across different elements of the Cloud-Edge continuum. We will also explore the possibility of allowing the Privacy Monitor or some of its components to migrate to a new location as the IoT devices whose data are being protected move. This opens the door to figuring out intelligent or pre-emptive re-allocation of the Privacy Monitor. Moreover, we will analyze various virtualization technologies available for implementing such a service in existing Edge Computing platforms.

Finally, we envision that the Privacy Monitor can be useful for data protection in some of the pilots considered in WP5, where data from IoT devices are to be stored, processed and/or queried for by various entities. In particular, we consider that the functionality offered by the Privacy Monitor could support some of the requirements in pilots “Supply Chain Security Assurance”, “Maritime Transport” and “Smart Cities”.

[21] https://www.oasis-open.org/standards/#xacmlv3.0

Full name	Acronym	Lead partner	TRL
Edge-Privacy		UMA	2
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.2, T5.5, T5.7	SCH-SP08, MT- SP19, SMC-SP10, SMC-SP11, SMC-SP17	n/a	

Table 8: Metadata: Edge-Privacy

4.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The Privacy Monitor is a Security- and Privacy-Preservation tool that fits into the CyberSec4Europe Architecture in the Managed Domain and more precisely within the MEC Platform. It will serve as a privacy-preserving data broker between the IoT devices and anyone interested in accessing or pushing data. The Privacy Monitor depends on the functionality provided by other elements of the architecture such as attribute providers and identity resolvers for mapping access rights to data as well as data leakage prevention components for filtering data according to the policies defined in the system. Finally, the Privacy Monitor may also take advantage of Blockchain services for traceability and accountability on data accesses.

4.2 AntibIoTic

AntibIoTic 2.0 [25] is an anti-malware for the Internet of Things that relies on Fog computing to protect Industrial IoT devices. AntibIoTic 2.0 overcomes central limitations of its predecessor AntibIoTic (1.0) [26], in particular with respect to legal and ethical issues, e.g., related to Article 3 (Illegal access to information systems) and Article 5 (Illegal data interference) in the EU directive on attacks against information systems[22].

The idea behind AntibIoTic 2.0 is to use a Fog node (or a federation of Fog nodes) to monitor and sanitize the devices connected to it, allowing only safe ones to access the Internet. To this aim, the Fog node uploads on each IoT device an agent (lately addressed as AntibIoTic Bot) that works as an anti-malware sanitizing and securing them and reporting live information back to the Fog node. Then, depending on the information received from each IoT device, as well as the operation mode set for AntibIoTic 2.0, the Fog node decides if the host is allowed to connect to the Internet.

[22] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0040&from=EN

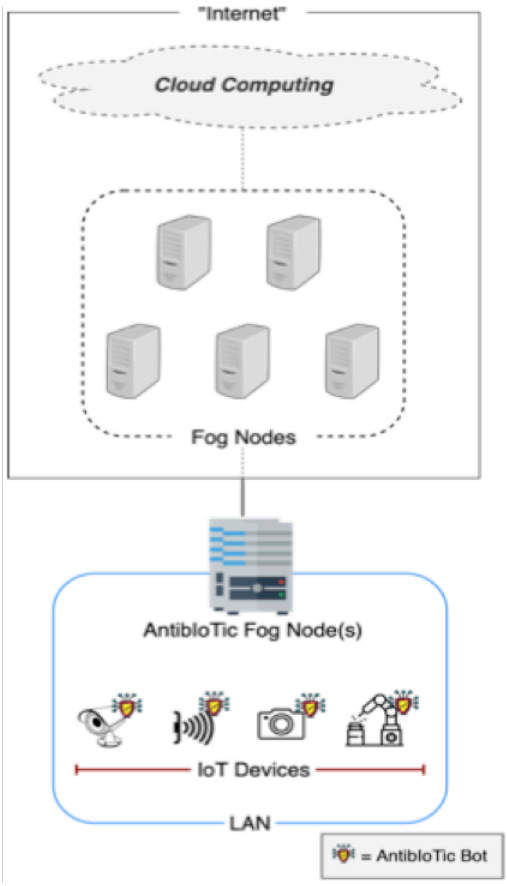

Figure 9: AntibIoTic 2.0 architecture

The main features characteristics of AntibIoTic 2.0 can be summarized as follows.

- *Easy to install & transparent to use*. After the first configuration, the system starts securing the IoT network in a transparent way.
- *Collect and process relevant data*. Automatically collects relevant data that can be later used to generate statistics and to improve the system.
- *Sanitize and secure IoT devices*. First, it cleans up the device and secures its perimeter against further intrusions; then, identifies security vulnerabilities of the device and takes action against them.
- *Persistent protection*. If a device is rebooted or temporary disconnected, it will be automatically protected again when available.
- *Versatile and scalable*. It is possible to scale horizontally or vertically and easily increase the intelligence of the system.

The features listed above are only the high-level summary of the basic AntibIoTic 2.0 functionalities, aimed at giving an idea of the system. Further extension and improvements are foreseen, which in particular will consist of several steps aimed at improving the AntibIoTic solution both in terms of implementation and design:

- *Implementation*: the aim is to expand the proof-of-concept to provide an implementation that includes as many functionalities as possible, making it closer to a final product. Some desired functionalities to be implemented are listed below:
  - expand the AntibIoTic Bot to generate a report about the security status of the hosting device;
  - expand the AntibIoTic Bot to generate keep-alive messages and short updates about the security status of the device;
  - expand the AntibIoTic Fog node implementation to be configurable with different operation modes;
  - expand the AntibIoTic Fog node implementation to parse the report sent from the Bot and take decisions based on that and on the operation mode;
- *Scalability*: the aim is to run and test AntibIoTic on a wider range of IoT devices, targeting different IoT architectures and scaling up in terms of number of devices.
- *Refine/Relax Security Assumptions*: both the Fog node and the interaction Bot-Fog~node within the LAN have been assumed secure. However, these represent important points for the full adoption of AntibIoTic 2. We are planning on working on this aspect.

The implementation of AntibIoTic 2.0 hides some technical challenges, mainly residing in the great variety of IoT devices and in the use of a paradigm, Fog computing, not fully established yet. However, we have developed a proof-of-concept of the solution that proves its feasibility by implementing some of the basic functionalities of AntiIoTic 2.0.

Full name	**Acronym**	**Lead partner**	**TRL**
AntibIoTic		DTU	6
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.7	SMC-SP01, SMC-SP02, SMC-SP8, SMC-SP10, SMC-OP2	https://ieeexplore.ieee.org/abstract/document/8802381/	

Table 9: Metadata: AntibIoTic

4.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

AntibIoTic is an anti-malware which belongs to the set of user-side security and privacy tools in the CyberSec4Europe architecture presented in Figure 1, namely those supported by the MEC platform in the IoT-Edge domain.

4.3 Research Challenges Identified in WP5

We have so far identified some other research issues that may demand further attention:

- Protecting data privacy post-release is an extremely challenging task. Data may be released together with some protection mechanism that remains attached to the data (cf., sticky policies [27]) and at the same time monitors data usage, however doing so may introduce new privacy issues on the data consumer side. It is certainly unclear how this can be done in a way that protects the privacy of both data owners and data consumers. On top of that, data consumers may abuse the system once access to data has been granted. While trusted execution environments may help prevent abuse, this would require any data consumer to have access to such type of hardware which is unlikely especially for IoT devices.

- Technical challenges arise from the interaction of Fog and Edge computing platforms with hardware-constrained devices belonging to the Internet of Things plane. Dealing and protecting all sorts of heterogeneous devices introduces a number of technical issues.

Moreover, there are also some inherent challenges which are primarily due to the very nature of edge and fog computing platforms. These paradigms are still in an early stage of development and need to mature before being introduced to the general public.

5 Mechanisms Reducing the Attack Surface

The struggle between "attackers" and tools protecting software environments is and will be a never-ending story. The more software and new technologies appear, the greater the risk of potentially vulnerable points. Therefore, attackers and hackers will always have room for action.

Attacks are affecting hardware and software with the aim of obtaining restricted data. Cyber-security systems are developed and deployed avoiding the access to private information. In order to improve the security, it is worth to diminish the attack surface. In cases where vulnerabilities provoking a data breach are found, or even where some data are disclosed intentionally but privacy of user data must be guaranteed, it is necessary to use data privacy-preserving technologies. Moreover, the cyber-security systems can, unintentionally or not, learn from the stored or shared data. Thus, a trade-off between security and privacy must be achieved, and making hardy the clear information (the visible surface) the user privacy is preserved, and user identification and even future analytics are not compromised.

De-identification techniques such as the anonymization technology and especially *k*-anonymization methods can play an important role facilitating the analyst work but preserving user data privacy. An example of data privacy-preserving technology is the anonymization asset presented in the following section.

5.1 Data Anonymization Service DANS

Privacy principles are basic when personal and sensitive data are shared between organizations. With the aim to preserve data user privacy and minimize data disclosure risk, the use of anonymization tools for sanitizing the data before disclosing can help to limit the risks [28].

Data Anonymization Service (DANS) is an anonymization service, created by Atos, based on the data anonymization Java open-source tool (ARX[23]) that provides different privacy models (e.g., the k-anonymity, k-Map, Average risk model, among others) to enable the application of certain privacy criteria over a specific dataset. ARX is under Apache 2.0 license.

DANS allows preventing privacy breaches when sensitive personal data are managed. Towards this end, DANS will provide different privacy models, which enable the application of certain privacy criteria over a specific dataset. Particularly, this tool supports the k-anonymity model, which we will use to preserve privacy of users' sensitive health data. This model has been used on recent studies for protecting biomedical data against data disclosure [29], and strongly guarantee user data privacy [30]. The main purpose of this model is to protect dataset from identity disclosure, considering to this end different generalization hierarchies associated with the attributes of data registers. In this sense, DANS classifies attributes in four types, as follows:

- Identifying attributes: these attributes will be removed from the dataset.
- Quasi-identifying attributes: they will be transformed by using the previously defined generalization hierarchies.

[23] https://arx.deidentifier.org/

- Sensitive attributes: these attributes will not be modified.
- Insensitive attributes: as sensitive attributes, they will not be modified.

Therefore, a dataset is k-anonymous if, considering just the quasi-identifiers in a dataset, it is impossible to identify a specific data register from, at least, k-1 other ones.

Finally, it should be pointed out that DANS considers other privacy models, such as l-diversity, t-closeness or d-presence, so they could be also employed to guarantee users' privacy-preserving.

Based on the ARX high-level architecture view [31], Figure 9 depicts the DANS main components participating in the anonymization process:

- Core modules:
  - I/O Interfaces: services for input/export data
  - Data Encoding: transforms data
  - Data Management: orchestrator module
- Anonymization algorithms: algorithms to be used
- Privacy criteria: set privacy criteria
- Public API: access to anonymization too

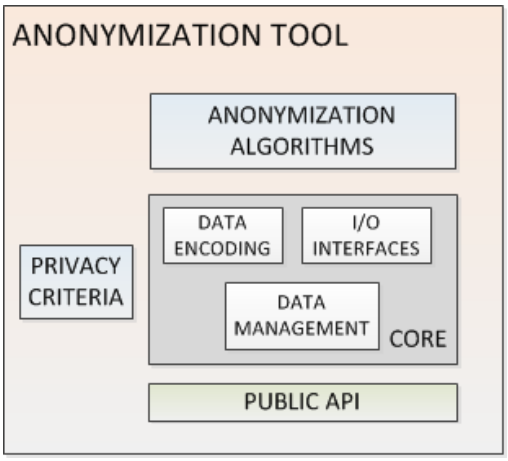

Figure 10: DANS main components

Due to the huge amount of shared health data and their use for different stakeholders such as research institutions, hospitals, companies, purposes, guarantee data privacy is a major goal, According to recent reviews on anonymization in the health domain [32]. Additionally, recent research shows how to integrate anonymization tools into extract, transform and load processes [33], in a similar approach as the suggested with this asset.

The Technology Readiness Level of DANS is 4 (technology validated in lab). DANS will be used in the medical data exchange demonstrator.

With the aim to address the privacy requirements of the medical data sharing stakeholders, but not limited to them, the following innovations are planned to be developed:

- Provide the anonymization service to the data providers for preserving user data privacy in different ways (e.g. as a library or as a REST service), assuring user data privacy when the data are processed avoiding user identification;
- Provide a tool easy to be used, improving the user experience of the data provider stakeholders.

Full name	Acronym	Lead partner	TRL
Data Anonymization Service	DANS	ATOS	4
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.5, T5.6, T5.7	OB-SP05, SCH-SP08, IDM-SP06, IDM-SP07, IDM-U02, MT-SP19, MD-SP02, MD-SPL01, SMC-SP17	https://arx.deidentifier.org/	

Table 10: Metadata: Data Anonymization Services

5.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

DANS component is part of the security and privacy tools the CyberSec4Europe architecture are including for preserving the user privacy data. Aligned with this objective the DANS asset will be used in the T5.6 Medical Data Exchange demonstrator offering this anonymization service to data providers for user data anonymization, preserving privacy on the created dataset..

5.2 Research Challenges Identified in WP5

Some potential research has been identified within WP5:

- Data anonymization tools are being used when health data are managed. Due to the increase of the number of shared data in the health domain, some de-identification process could be not sufficient for preserving personal privacy. Thus, additional research on stronger anonymize data tools which including more efficient and effective algorithms is needed.
- There exists a lack of standardization mechanisms when data are shared between different stakeholders. In this way the adoption in the health domain of specific standards (such as HL7-FHIR[24]) which facilitates the anonymization process will be investigated during the project;
- In spite of anonymization techniques are a good strategy for mitigating the risk of identifying the data subject, further research must be done on how the anonymization tools are used when health data are anonymized for creating anonymous data set[25], This process must assure that the GDPR regulation is fully complied, preventing identification when the anonymized data are processed.

[24] https://www.hl7.org/fhir/
[25] https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

6 Mechanisms Based on Trusted Execution Environments

A trusted execution environment (or TEE) is a secure area of processor that guarantees the protection of code and data with respect to confidentiality and integrity. These isolated execution environments offer an execution space that provides higher security guarantees than a full operation system, while also supporting more functionality than a secure element like a TPM. In the following we describe different assets for, and based upon, TEEs.

6.1 Cryptovault

An important property of blockchain networks, which are based on decentralized trust, is parties' authority to their own key-records. The initial problem is how to protect and maintain the availability of the key record when device containing the key will break.

Cryptovault, firstly, introduces an implementation that generates and maintains private keys in Intel SGX[26] enclave enabling usage of the private keys in a process isolated from all other processes running on the same system. Secondly, based on Shamir's secret sharing scheme, the method provides the ability to back-up key as an independent part to external records, such as remote servers, over the end-to-end secured connection.

The implementation is a client-side application, which can generate a key, initialize backup to remote server, restore backup and sign transaction. To accomplish functionalities program should be able to generate secp256k1-keypair, derive Ethereum address from public key, divide secret into shares, reconstruct secret, generate RSA key pair, encrypt and decrypt messages with RSA, and sign transactions in SGX enclave. Same functions can be used to implement remote-server functionality.

Signing transactions in accordance with Ethereum specifications have not yet been implemented. Currently, Cryptovault is only available for Intel SGX. The method itself can be implemented also into other trusted execution environments. Technology readiness level is 4 (technology validated in lab). Contribution now on will focus on finishing the missing functionalities.

Full name	Acronym	Lead partner	TRL
Cryptovault		VTT	4
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.6, T5.7	OB-SP13, MD-SP07, SMC-SP09	n/a	

Table 11: Metadata: Cryptovault

6.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

In task 3.2 architecture, Cryptovault belongs to the user domain. Such a method could be used to secure cryptographic keys on various devices such as mobile phones, autonomous cars, IoT- or Edge-devices.

6.2 Elastic Deployment of TEE-based applications in the cloud

[26] https://software.intel.com/en-us/sgx

Trusted Execution Environments like Intel SGX or ARM TrustZone are security architecture available on commodity hardware that enables computation on confidential data and public auditing of the data processing pipeline [34]. However, TEEs like Intel SGX arguably the most popular TEE on desktop and server machines, has been mainly designed to run on end-user machines so that software vendors can control the environment where their software is executed. However, the security provisions of SGX are also attractive for cloud platforms with different cloud providers like Amazon or Microsoft that have recently started to include SGX in their offerings. Nevertheless, SGX has been designed to run centralized applications and its deployment model is in sharp contrast with key features of cloud applications such as resource aggregation and elasticity. We, therefore, plan to design and implement security frameworks for SGX-based cloud application that could keep the security properties of SGX but, at the same time, would allow cloud providers to dynamically manage customer's VMs.

Full name	Acronym	Lead partner	TRL
Elastic Deployment of TEE-based applications in the cloud		NEC	3
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2	OB-UC1, OB-UC4, SCH-UC1, SCH-UC2	n/a	

Table 12: Metadata: Elastic Deployment of TEE-based applications in the cloud

6.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

This asset belongs to the User Domain of the CyberSec4Europe architecture. It enables data processing in the cloud while guaranteeing the privacy of user data and the integrity of the code that processes such data.

6.3 Backdoor-resistant TEEs

Anonymous attestation is a key feature of TEEs that allows a verifier to authenticate a platform as a member of a trusted set, while keeping the platform itself anonymous (within that set) [35]. This functionality is realized by using a privacy-enhanced flavor of group signatures in which signatures cannot be traced, not even by the group manager. The security of anonymous attestation schemes is grounded on the trustworthiness of the signer. In particular, anonymity and unforgeability definitions assume that the signer the SGX subsystem is trusted and does not exfiltrate any information via its signatures. Yet, in most applications, the signer is a small piece of hardware with closed-source firmware to which a user has only black-box access. In such a scenario, trusting the hardware to behave honestly may be too strong of an assumption for mainly two reasons. First, having only black-box access to a piece of hardware makes it virtually impossible to verify whether the hardware provides the claimed guarantees of security and privacy. Second, recent news on state-level adversaries corrupting security services have shown that subverted hardware is a realistic threat. In the context of anonymous attestation, if the hardware gets subverted (e.g., via firmware bugs or backdoors), it may output valid, innocent-looking signatures that, in reality, covertly encode identifying information (e.g., using special nonces). Such signatures may allow a remote adversary to trace the signer, thereby breaking anonymity. Using a similar channel, a subverted signer could also exfiltrate its secret key, and this would enable an external adversary to frame an honest signer, for example by signing bogus messages on its behalf. We plan to design backdoor-resistant anonymous attestation schemes tailored for TEEs such as SGX. The goal is to design a primitive with the same security and privacy properties of current attestation scheme while tolerating subverted hardware.

Full name	Acronym	Lead partner	TRL
Backdoor-resistant TEEs		NEC	2
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2	OB-UC1, OB-UC4, SCH-UC1, SCH-UC2	n/a	

Table 13: Metadata: Backdoor-resistant TEEs

6.3.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Backdoor-resistant TEE belongs to the IoT Domain of the CyberSec4Europe architecture. It enables IoT devices to leverage TEEs while, at the same time, guaranteeing meaningful security and privacy properties in case of hardware subversion.

6.4 Research Challenges Identified in WP5

So far, the following areas for further research and improvement of the presented building blocks have been identified:

- In order to also benefit from SGX-like offerings in cloud scenarios, we plan to design and realize a security framework for SGX-based cloud applications that keep the security of SGX, but at the same time allow cloud providers to dynamically manage their customers' VMs.
- We plan to design backdoor-resistant anonymous attestation schemes tailored for TEEs such as SGX. The goal is to design a primitive with the same security and privacy properties of current attestation scheme while tolerating subverted hardware. In this scope, we will provide formal definitions for subversion-resilient attestation schemes for SGX by starting from the current definitions of [36]. Then, we will revise the state of the art of subversion-resilient cryptography and design an attestation scheme that can tolerate subversion.
- In a network without central authority each node is responsible to their own key records. It is thus important to improve techniques to ensure the availability and security of the key record in decentralized networks.

7 Privacy-Preserving Middleware for IoT

Middleware subsumes all software components beyond operating systems that provide services to software applications, thereby making it easier for developers to focus on the main purpose of their applications.

In the following we present a privacy-preserving middleware component specifically designed for the Internet of Things.

7.1 pTASC Privacy Preserving Middleware

pTASC ensures secure and authenticated communication channels between Internet of Things (IoT) devices. It relies on Diffie-Hellman (DH) key exchange allowing it to overcome the complexity of PKI based systems and avoiding the use of trusted third parties for authentication purposes, making it a decentralized solution. However, DH solutions alone cannot ensure authentication and do not prevent Man-in-the-Middle (MitM) attacks [37].

pTASC provides authentication by adapting the approach of ZRTP [38], which allows the detection of MitM attacks by displaying Short Authentication Strings (SAS) for users to read and verbally compare over the phone in Voice Over Internet Protocol (VoIP) communication. For this, an extra communication channel is used for exchanging the SAS securely. It is then compared by both peers for detecting conflicts (i.e., non-identical values), preventing MitM attacks by dropping the connection.

The extra channel uses an adaptation of the Limited-Location Channel (LLC) concept, that takes advantage of proximity characteristics typical of devices in ad-hoc networks, where a secure pre-authentication channel can be established by visual or physical contact between the communicating devices [39]. This secure pre-authentication channel enables devices to exchange keys directly with each other without the need for public-key and CA certificates. The use of LLC is only needed for the first iteration. Remaining iterations rely on key continuity and forward secrecy capabilities.

Figure 11 presents the communication scheme between two devices (A and B) through pTASC. The communications scheme starts with A and B exchanging HELLO and HELLOack (i.e., acknowledgement of receipt of the HELLO message) messages, shown as steps F1 to F4. These messages include the identification of both devices. The identification is generated using a Pseudo-Random Number Generator (PRNG) and this identification is validated in this phase.

After this first exchange of messages, the key agreement exchange can begin with a Commit message in step F5 from the device B to the device A. The scheme as shown assumes that the device B is the initiator. There are two approaches that can now be carried out in order to agree a key between B and A:

- DH mode: pTASC endpoints exchange a new shared secret through the DH exchange;
- Pre-shared mode: In this mode, the DH calculation is omitted by the endpoints, as it is assumed that there exists a known shared secret from a previous session. However, DHPart1 and DHPart2 messages 1 are still exchanged to determine which shared keys should be used. Instead of DH values (hvi and pvr), the end-points use nonces, along with the retained secret keys, to derive the key material [40].

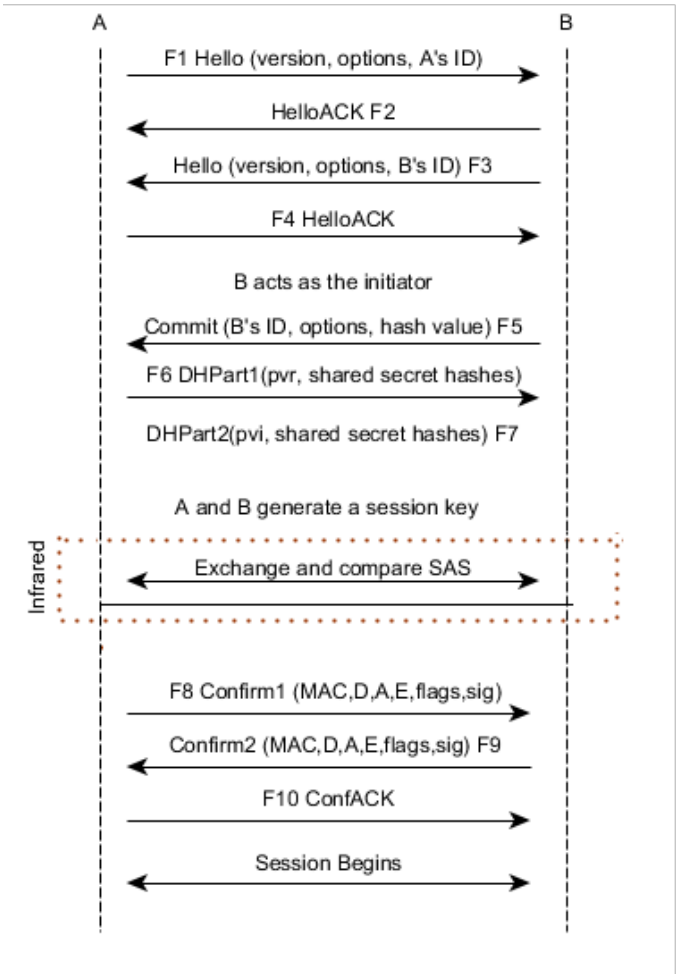

Figure 11: pTASC protocol handshake

After this phase, the LLC is used for privately exchange the SAS, with the comparison being made locally by both peers (current implementation uses Infrared connectivity for supporting LLC). If the comparison proves valid, a Confirm message is sent from both devices indicating that both have accepted and verified the SAS, and the protocol is ready to start sending data in this secure session. In case the validation process fails, a MitM attack is assumed and the connection is dropped.

pTASC takes advantage of some additional security and privacy properties, as the forward secrecy, in order to guarantee confidential communication in the future. This means that when devices A and B perform multiple connections at different times with each other (i.e., different communication sessions), the protocol rotates the keys between each ones of the sessions so that the same key is not re-used in a later session, thus, all the communications keys are different. An identifier file is used by each device for caching symmetric key material used to compute new secret session keys, with these values changing with each session.

If an attacker gets access to the local seed, the attacker will not be capable of reproducing any of the information exchanged in previous sessions. I.e., if a single communication is compromised (say the session key is "leaked"), it does not compromise the confidentiality of all communications prior to it. This way, pTASC is able to guarantee additional protection in communication because, even if the users do not bother with SAS, there is still fairly decent authentication against a MitM attack, based on a form of key continuity.

The main purpose of pTASC is to reduce human effort during device provisioning and configuration. This reduces the probability of miss-configuration that can lead to compromised privacy and security. Additionally, the two main characteristics of pTASC are usability; and decentralization.

Usability

While secure proposals exist, in order to apply them many procedures are necessary, and many of them require previous key changes. Note that the exchange of keys by unsecured means (email for example) compromises all of the security of the process. The system proposed in this work is easy to use. In the case of J-PAKE [41], for example, its use in Firefox Sync is less easy to use, as it requires that the SAS be written in both devices to provision these two devices. Also, in ZRTP, the protocol needs to exchange the SAS verbally and then, compare it on the phone. With LLC, we can exchange the SAS securely and privately between the peers and compare it locally in both sides.

Decentralized

Limitations of IoT devices (such as low power and processing capabilities) are well known. This makes PKI based solutions inefficient to use in such an environment. The protocol presented in this section differs from PKI models because it is decentralized – it does not rely solely and exclusively on a Certification Authority (CA), but rather encompasses a number of reliable, independent elements. Thus, our approach is more suitable for low-resource devices, without any need for PKI or CAs, key certification, trust models, etc., which encompasses inherent complexity.

Related Work

Device Pairing Using Short Authentication Strings. Device Pairing Using Short Authentication Strings (SAS) [42] is a two-device pairing mechanism based on the agreement and validation of a secret's authenticity using a SAS. The protocol consists of three phases: discovery, agreement, and authentication. When the pairing service starts, the server starts by publishing the chosen instance name. The client will discover that name and the corresponding connection parameters. After the server is discovered, the client and server use a TLS session which allows them to agree on a shared secret using a cryptographic protocol that produces a SAS. After this, there is an authentication phase, used to validate the pairing through a SAS. In this phase, the comparison of the SAS is made through manual verification, i.e., a user has to verify that both devices display the same string. If, instead, the server and client support Quick Response (QR) codes, then the server displays a QR code with the encoding of the SAS, and the client is capable of scanning the value of the SAS and comparing it to the locally computed value.

ZRTP. ZRTP is a key agreement protocol used by Voice over Internet Protocol (VoIP). It is based on Diffie-Hellman (DH) keys (also called shared secret keys) used for generating a secret master key that is later used for establishing Secure Real-time Transport Protocol (SRTP) Cryptographic Contexts or streams. DH alone does not provide authentication, thus are vulnerable against Man-in-the-Middle (MitM) attacks. In order to authenticate both peers, ZRTP uses Short Authentication Strings (SAS) during the key negotiation. Only peers that share SAS will be able to reach equal shared secrets. If the SAS is the same, the communication could be classified as secure, but, for this to happen, the SAS needs to be validated on both devices by a user. When a communication is classified as secure, the shared secrets are saved and used to generate new secure sessions in the future, thus decreasing the computational effort and skipping the user intervention on the comparison of SAS.

PAKE and J-PAKE. Password-Authenticated Key Agreement (PAKE) protocols are a recent addition to the cryptography literature. Currently, the most sophisticated algorithms do not perform key exchanges based on public-key cryptography, allowing low-entropy passwords to be used. Lancrenon et al [43], discuss three state-of-the-art PAKE protocols are discussed. J-PAKE is based on a shared password, which does not require PKI or third-party entities for establishing secure communication between two parties. It uses an elliptic curve DH for the key agreement and a Schnorr Non-Interactive Zero-Knowledge (NIZK) signatures [44] proof mechanism that authenticate two peers and establish a shared secret between them based on a passphrase. There are some services that use J-PAKE, such as the Pale Moon Web-Browser, the lightweight API in Bouncycastle[27] (1.48 and onwards), and the Thread (IoT wireless network protocol). This protocol has also been supported by FireFox Sync, OpenSSL and OpenSSH. However, it was removed after 2014 due to several known J-PAKE issues, already published by Mohsen Toorani [45], including vulnerable to password compromise, impersonation attack, and other shortcomings with respect to replay and Unknown Key-Share (UKS) attacks.

Authentication Models for ad-hoc Authentication

Nidal Aboudagga et al. have documented a taxonomy and research issues in the authentication protocols for ad-hoc networks. We describe in more detail some similar to our approach. Asokan and Ginzboorg [46] presented a key agreement protocol in ad-hoc networks which is based on PAKE and a location-based key agreement to authenticate through a limited channel such as Bluetooth. However, this protocol is not focused on IoT and does not have key continuity feature that is a good application in such context as it allows devices to move away after the first pairing and continuously have secure communications on the following connections.

In [39] the authors present new schemes for peer-to-peer authentication in ad-hoc wireless networks. The authors also describe how to use demonstrative identification to perform pre-authentication over LLCs. Amir Spahić et al. present an authentication mechanism (pre-authentication phase) which uses context information through LLC using Infrared. Serge Vaudenay [47] presents a concept that authenticates a short string, the SAS, through an extra insecure channel. This concept is similar to pTASC, however, this is based on the use of a narrow-band authentication channel. Another proposal [48] is based on a new LLC using biometrics. The protocol efficiently calculates a shared secret key from biometric data using quantization and cryptanalysis. The authors use grip pattern-based biometrics as a location limited channel to achieve pre-authentication in a protocol that sets up a secure channel between two handheld devices.

[27] https://bouncycastle.org/

Full name	Acronym	Lead partner	TRL
Privacy Preserving Middleware	pTASC	C3P	2
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.2, T5.3, T5.6, T5.7	SCH-SP01, SCH-SP02, SCH-SP03, SCH-SP04, SCH-SP05, SCH-SP06, SCH-SP07, SCH-SP08, SCH-SP10, SCH-SC01, IDM-SP01, IDM-SP02, IDM-SP03, IDM-SP04, IDM-SP07, IDM-SP11, MD-SP01, MD-SP02, MD-SP03, MD-SP04, MD-SP05, MD-SP06, MD-LR01, MD-SPL01, SMC-SP01, SMC-SP02, SMC-SP03, SMC-SP17, SMC-LR02	https://repositorio.inesctec.pt/bitstream/123456789/6936/1/P-00N-571.pdf	

Table 14: Metadata: Privacy Preserving Middleware

7.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

This asset maps with MEC Platform from the IoT Domain. Specifically, pTASK maps with *Data Leakage* Prevention and *Data Broker Privacy-Preservation*.

7.2 Research Challenges Identified in WP5

Up to date, the following directions for future work have been identified:

- Device provisioning is usually an arduous task that encompasses device configuration, including identity and key provisioning. Most solutions require human efforts to setup devices. This is a costly approach to scale and can lead to compromised security due to misconfiguration. Given the potential large number of devices, this process must be scalable and semi-autonomous.
- Device authentication remains without a viable solution, especially when considering a resilient decentralized approach that is the most suitable for this scenario, as it avoids some issues related to centralization, e.g., censorship, data leakage or profit from corporations.
- Given the number of things present in daily lives, data and consent acquisition should be a scalable process. In the presence of a data request, the middleware must evaluate the request on behalf of the data owner performing some kind of a preliminary filtering. Thus, (semi)autonomous negotiation strategies are needed for implementing this, taking into consideration such factors as data owner preferences, user profiling and rewards schemes, and combine different privacy-preserving techniques to support end-to-end privacy and GDPR compliance.

8 Security and Privacy by Design

Security and privacy by design describes an approach to systems engineering and software development, where security and privacy are already taken into account throughout the whole development process from the very initial phases onwards. In particular privacy by design has been incorporated into the European General Data Protection Regulation. In the following, we present three assets for supporting the implementation of this approach.

8.1 Privacy-Preserving for Genomic Data

The recent research shows that next-generation sequencing technologies have evolved to produce biological data faster. Those technologies are now well developed for short reads, however, a new generation producing longer reads has already entered the market [49]. In practice, those reads are often aligned to a reference genome to obtain their location in that reference, and then infer genomic insights [50].

On the other hand, processing, storing and sharing genomic data raises new privacy challenges, and risks to privacy are still an obstacle to sharing [51]. Indeed, several privacy attacks have been described in the literature. These genomic privacy attacks alerted the research community for the need to replace the conventional procedures by privacy-preserving frameworks. In the last years, several protocols have been designed to protect the sensitive information contained in genomic data. However, their slow performance fails to match the throughput of current sequencing machines at manageable cost.

Future genomics research would benefit from software solutions capable of protecting sensitive information throughout its whole life-cycle. That is, from its initial acquisition during the sequencing phase, to the day it is not sensitive anymore, including its storage and use by researchers or physicians.

Full name	Acronym	Lead partner	TRL
Privacy-Preserving for Genomic Data		UNILU	3
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.6, T5.7	MD-SP05, SMC-F02	n/a	

Table 15: Metadata: Privacy-Preserving for Genomic Data

8.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Privacy-preservation for genomic data is required in the biometrics verifiers module of the identity and privacy-preservation services, covered by the control and management plane of the presented architecture.

8.2 Flexible metrics and analyses for differential privacy

We make use of the data by performing computations on it, and then studying the outputs. There exist privacy-preserving computation technologies to reduce information leakage during the computations. To reduce the leakage through the results of the computation, one has to apply perturbation techniques, either during or after the computations. There are several metrics to measure the amount of leakage through the outputs, with *differential privacy* [52] having perhaps the most desirable composition properties.

Differential privacy characterizes, how much a unit change in the input of some computation can change the probability of getting a certain answer. For composition, we also need to know the *sensitivity* of computations - the possible amount of change in the output for the unit change in the input. Given such

characterization for the component computations, these can be aggregated into an upper bound on the differential privacy of the composed computation [53].

The component computations can be specified through a wide variety of languages. In case of data-oriented business processes, SQL is a commonly used language. Determining the sensitivity of a SQL-statement is non-trivial, and often the sensitivity may be infinite if the statement contains joins. For certain join-statements, though, the sensitivity may still be finite due to the filtering conditions in that SQL-statement. We have proposed a model-checking based method to determine the sensitivity of SQL queries, which can also accurately analyze statements containing join-clauses [54]. The method differs from other approaches by its accuracy and the lack of needed instrumentation of the SQL-statement about the maximum possible number of rows that can match a foreign key.

The guarantees offered by differential privacy depend not only on the numerical value of the "ε" parameter that comes with it, but also on the nature of what constitutes a unit change on the inputs. If the inputs are databases, then the distance on the inputs is usually defined as the number of changed rows. But not all changes are created equal, and the data owner may want to have a more fine-grained control over which changes should be hidden (the distance between databases differing only by such change should be small) and which are allowed to be reflected in the output of the computation (the distance between databases differing only by such change is allowed to be large). Such more generalized distances have been studied before, e.g. in the context of location privacy [55]. We have gone further in this direction and come up with a whole framework for defining the distances between databases; this framework comes together with the analysis of sensitivity of SQL queries with respect to these distances [56].

Full name	**Acronym**	**Lead partner**	**TRL**
Flexible metrics and analyses for differential privacy		CYBER	8
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.3, T5.4, T5.5, T5.6, T5.7	OB-SP13, OB-SP16, OB-SP27, OB-LR01, OB-LR08, SCH-SP06, SCH-SP08, SCH-SP09, SCH-LR01, SCH-LR04, IDM-SP05, IDM-SP06, IDM-SP07, IDM-LR01, IDM-LR02, IR-LR02, MT-SP19, MD-SP01, MD-SPL01, MD-SPL02, MD-SPL03, MD-LR01, SMC-SP10, SMC-SP11, SMC-SP16, SMC-LR02	https://pleak.io/home https://github.com/pleak-tools	

Table 16: Metadata: Flexible metrics and analyses for differential privacy

8.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Our privacy metrics are very much a mechanism to specify privacy policies. We see them as security/privacy tools, both at the user side and at the managed domain. At the user side, we need tools that allow the users to understand their privacy choices and trade-offs. For the managed domain, we need to verify whether systems comply with the users' privacy policies.

8.3 GDPR-Based User Stories in the Access Control Perspective

Nowadays, the Information Technology (IT) domain is moving towards systems with growing complexity, where digitalization, Artificial Intelligence (AI), interconnection and mobility are some key factors. Indeed, in their multidisciplinary nature, they require an extensive deployment of advanced Information and Communication Technologies (ICTs), as well as the adoption of effective measures for strengthening security, trust, dependability and privacy. These aspects have to be considered over the whole Software Development Life Cycle (SDLC), from gathering of the requirements to the deployment and subsequent maintenance of the system.

Over the last decade, especially for small and medium enterprises, Agile Software Development (ASD), first introduced in the Agile Manifest [57], and its subsequent evolution such as eXtreme Programming (XP) and Scrum [58] are becoming commonly adopted software development processes. Basically, ASD is an iterative approach that focuses on incremental specification, design and implementation, while requiring full integration of testing and development.

In this development process, a common means of capturing the user needs and describing the value that the user would get from a specific functionality is the so-called User Story (US). From a practical point of view, a US focuses on a requirement written according to a specific format and guidelines on how to implement it. Usually, depending on the granularity of the story, different names can be used for defining its contents: large ones may be known ad Epics, and small ones as Features, User Stories, and Tasks [59]. However, small organizations and software development groups could not expend the effort (in terms of budget and time) needed to collect and implement all the required User Stories prior to release. When the missing stories refer to privacy and data protection requirements, the side effect is to release software with high privacy risks [60].

With the entering into the force of the General Data Protection Regulation (GDPR) this situation is not affordable anymore, because it is changing how Personal Data should be processed. Indeed, the GDPR imposes several duties on the Controller and Processor, i.e., the data managers, concerning the processing of Personal Data, i.e., any information related to an identified or identifiable natural person called the Data Subject. Additionally, the GDPR defines a system of fines to induce controllers and processors to be compliant with its provisions. Thus, the controller and processor need to demonstrate compliance with the GDPR as required by the Accountability principle (Art. 5.2). However, this is not a trivial problem as it involves the definition of specifics purposes, the management of the consent given by the Data Subject whose personal data is referring to, and the need to demonstrate the compliance with the GDPR's provisions.

Within the Agile development, among the proposed solutions to tackle the security issues and vulnerabilities in an efficient way, one that is currently taking place is the possibility of using Security Backlogs to drive the software development work. In the Agile context, a backlog represents a prioritized features list describing the functionalities to be included in the final product. These backlog items are often in the form of User Stories, e.g., Asthana et al. The set of security backlogs is, therefore, a ready-made specification of the security items (requirements and task descriptions) useful for the implementation. An example of a security User Story is as follows (cf. Asthana et al.):

> As *[an information security manager]* I want *[that it is clearly defined which user account are authorised for which activities]* so that *[the effectiveness and correctness of access controls can be checked]*.

Following this tendency, the contribution of this section consists of: (1) introducing the concept of Data Protection Backlog that contains User Stories based on the GDPR requirements; (2) defining a methodology for mapping specific provisions of the GDPR to User Stories; (3) providing a ready to use set of User Stories each one having the corresponding implementation guidance so as to assure a GDPR-compliance design and (4) defining a systematic development process for implementing access control systems and their policies compliant-by-design with the GDPR.

This assures that the leveraged access control systems can protect personal data (security perspective) and process them lawfully (legal perspective),-Thus in this section we depict the process for minimizing errors and issues in the GDPR enforcement and providing the process for developing a consolidated, verified and predefined structure of Access Control Policies (ACPs) [61]. Accordingly, and in line with this tendency, for each identified User Story, we provide a GDPR-based Access Control Policy (ACP) template for each provision related to access control. Indeed, the templates represent meaningful, concrete and predefined blocks for ACP specification, which can be adopted and refined for the different scenarios, so as to overcome possible misinterpretations and reduce security and privacy risks.

Consequently, as schematized in Figure 13, we are providing a process for implementing access control systems and their policies compliant-by-design with the GDPR.

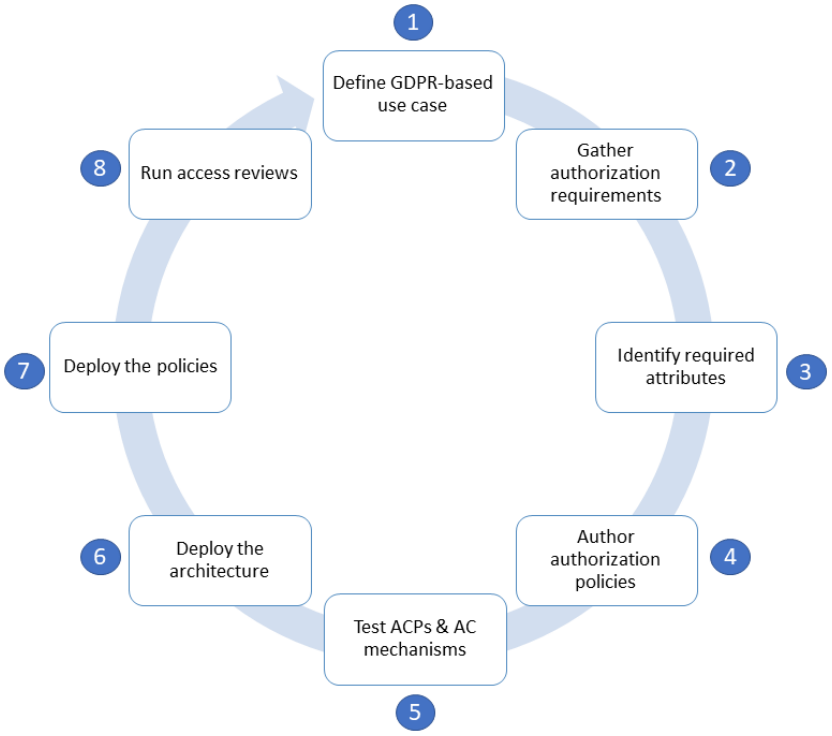

Figure 12: Overview of the The Authorization Policy Life Cycle

The life cycle is inspired by Brossard [62], which is a systematic approach to implementing authorization systems within enterprise. Figure 11 defines a specific, integrated GDPR focused process development life cycle for the specification, deployment and testing of adequate fine-grained authorization mechanisms able to take into account legal requirements. The result is an Agile ADLC (Agile Development Live Cycle),

which is profoundly rooted in the GDPR's "Data Protection by Design" approach (Art. 25) and the "Confidentiality and Integrity" principle defined in Art. 5.1(f).

As depicted in Figure 11, the resulted Agile ADLC is composed of eight phases: 1) Define GDPR-based use case; 2) Gather authorization requirements; 3) Identify required attributes; 4) Author the authorization policies; 5) Test ACPs & AC mechanisms; 6) Deploy the architecture; 7) Deploy the policies; and 8) Run access reviews.

Specifically Step 1 and 2 aim to define the context and an achievable scope so as to establish a common base to discuss with different stakeholders (e.g., Data Subject, Controller and Data Protection Officer (DPO)).

The established Use Cases are conceived according to GDPR implementation challenges. For this aim, we relies on the concept of Data Protection Backlog that contains User Stories based on the GDPR demands. These allow to gather authorization requirements and the sources they come from (step 2). In our case, the primary source is the GDPR regulation, therefore, authorization requirements are defined in terms of statements or natural language authorization policies. The systematic approach for mapping the specific provisions of the GDPR into to User Stories will be detailed more in the rest of this section.

Steps 3 and 4 involves the identification of the required attributes used in the selected requirements and their origin so as to make easier requirement reviews and authoring the authorization policies that are enforceable and directly deployable into the ACS. The outcome of these steps is a set of AC rules each related to a specific User Story and the derived ACPs take into account real data. Guidelines on how to proceed for obtaining such enforceable policies are briefly described in the remaining of this section.

Step 5 ensures that the implemented XACML policies meet the GDPR requirements. State-of-the-art and specifically conceived testing techniques should be used according to the different purposes. This step involves also the evaluation of the adequacy of the current AC mechanisms in the context of the GDPR.. The focus will be on three different testing scenarios: (1) Test Strategy Assessments; (2) Testing GDPR-based ACPs expressed in XACML 3.0; and (3) Evaluation the adequacy of AC mechanisms in the context of the GDPR.

The last three phases of the proposed development life cycle aim:

(1) to define of the contact points within the existing systems in order to make the different applications able to interact with authorization system (step 6);

(2) to deploying the authored XACML policies according to the selected (production) environment (step 7). This step is usually business dependent.

Finally, (3) to analyse the policies against a set of attributes to determine what these attributes grant (step 7).

As depicted in Figure 11, the asset proposed in this section relies on different outcomes. One of them is the ***Data Protection Backlogs***, which are lists of User Stories about the GDPR provisions told as technical requirements (outcome of step 2 of Figure 11). For each story thecorresponding Access Control Policy (ACP), so enabling the implementation of GDPR compliant Access Control (AC) systems (outcome of step 4 of Figure 11).

To better schematize the process for mapping specific provisions of the GDPR, first into User Stories and then into the corresponding implementation guidelines we report in Figure 12 the general schema.

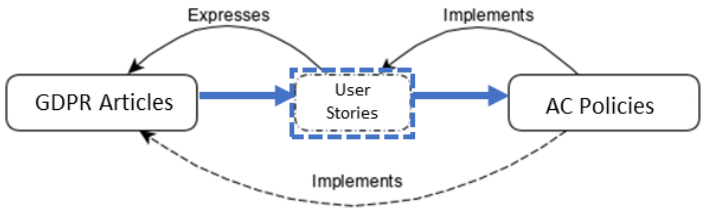

Figure 13: Overview of the conceptual model of GDPR-focused user stories

As a result, the set of User Stories, associated with the proper ACP templates, would be a valid starting point for privacy requirements specification, and generic guidance for who are facing to the problem of GDPR implementation. The provided ***Data Protection Backlogs*** can bu used by developers when a new ACS should be implemented. In this case the the developer can pick up the most suitable predefined User Story and easily derive the ACP implementing it. A more detailed conceptual model is however reported in [63].

In detail the process to define the set of User Stories, related to the provisions of the GDPR and the AC ruled, is composed of three main steps (see Figure 14): (1) GDPR Articles Selection; (2) User Stories Definition; (3) GDPR AC Rules Definition.

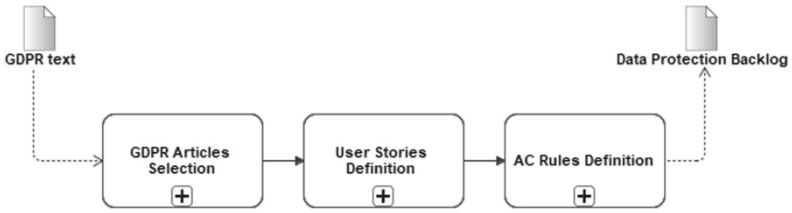

Figure 14: GDPR-focused User Stories Definition Process

GDPR Articles Selection. The input of the process if the GDPR text. Firstly, we selected only the mandatory part of the GDPR which consists of ninety-nine articles; for each article, we decided whether is related to AC concept, i.e., AC language or AC mechanism, and consequently we create and Epic associated with the current article. The result of this step was the selection of forty-one Epics (i.e., GDPR articles) related to AC. Specifically, three of them were concerning only AC mechanism; eight were referring only ACPs, and thirty articles related to ACPs and AC mechanism. For more details about this step we refer to our work [64].

User Stories Definition. For each article identified in the previous step, we extracted one or more technical requirements and defined a specific User Story for each of them. Thus, the User Stories were added to the Epic associated with the current article. In order to trace the covered GDPR's articles during the Agile development process, we defined for each Epic an identifier (name EpicID) able to find the GDPR's article the Epic is referring to. Similarly, we defined an identifier for each User Story (called UserStoryID) with the purpose to identify the specific part of the GDPR's article the User Story related to (e.g., the paragraph or the letter of the article).

GDPR AC Rules Definition. The final step deals with the translation of the technical requirements associated with the AC language, and consequently, we defined an AC rule for each User Story conceived in the previous step. In literature there exist different proposal for the derivation of ACPs from the natural language [65,66] or controlled natural language [67]. In our previous work we defined a systematic approach for deriving ACPs directly from the GDPR, and we refer to it for more details about this step.

As in Figure 8.1, the result of this process is *Data Protection Backlog*, i.e., a Privacy Backlog containing a set of AC rules organized in User Stories, Epics and Theme. This is a ready solution to be used during the Agile development of AC systems aligned with the GDPR requirements.

The *Data Protection Backlog* is the asset we are finalizing and we provide an extract of such asset in Table 17. The User Stories are reported from the perspective of both the Data Subject and the Controller. The table is composed of three columns: column *Article* (first column) contains the GDPR's articles; column *User Story* contains the GDPR-based User Stories defined; finally, *AC Rule* column reports the AC Rules related to the User Stories.

Article	User Story	AC Rule
Art. 6.1(a)	As a [Controller] I want [to process Personal Data only if Data Subject has given the Consent for one or more specific Purpose], so that [the processing shall be lawful]	[Controller] can [Process] [Personal Data] if [PersonalData.purpose = Process.purpose AND PersonalData.purpose.consent = TRUE]
Art. 7.3	As a [Data Subject] I want [to withdraw my Consent], so that [I can exercise my right as stated in Art. 7.3]	[Data Subject] can [Withdraw] [PersonalData.purpose.consent] if [PersonalData.owner = DataSubject] AND [PersonalData.purpose.consent = TRUE]
Art 15.1	As a [Data Subject] I want [to access my Personal Data], so that [I can be aware about my privacy]	[Data Subject] can [Access] [PersonalData] AND [Resource = PersonalData.purposes] AND [Resource = PersonalData.categories] if [PersonalData.owner = Data Subject]
Art. 15.3	As a [Data Subject] I want [to download a copy of my Personal Data], so that [I can check their correction]	[Data Subject] can [Download] [Personal Data] if [PersonalData.owner = Data Subject]

Table 17: GDPR-focused User Stories: Controller and Data Subject Perspectives

Full name	Acronym	Lead partner	TRL
GDPR-Based User Stories in the Access Control Perspective		CNR	4
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.3, T5.4, T5.6, T5.7	OB-SP05, OB-SP27, OB-LF04, OB-LR04, OB-LR05, OB-LR09, SCH-LR01, IDM-LR01, IR-LR02, MD-LR01, SCM-F02, SMC-F10, SCM-F11, SCM-SP06, SCM-SP10, SCM-SP11, SCM-LR02	n/a	

Table 18: Metadata: GDPR-Based User Stories in the Access Control Perspective

8.3.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The outcomes of the proposed asset can be used in the Regulatory Management component for different purposes. Indeed, the proposed methodology for extracting, implementing and testing the data protection regulation can be included in the Regulatory Management for translating any Legal Text (in our case the GDPR), analysing the articles related to ACs and creating a Data Protection Backlog containing a set of User Stories. Additionally the testing facilities can be used for validating the derived GDPR-based ACPs expressed in XACML 3.0 and evaluating the adequacy of the adopted AC mechanisms in the context of the GDPR.

Finally, the Data Protection Backlog can be used to conduct Data Protection Impact Assessment (DPIA) which is mandatory in the GDPR (Art. 35), and therefore the asset is related also with the Risk & Incident Management

8.1 Research Challenges Identified in WP5

The following challenges have been identified so far:

- Regarding the protection of sensitive genomic data, the consortium will contribute to biomedical data protection in the concept of security & privacy by design in Task 3.2. More precisely, it is planned to investigate mechanisms and techniques to reconcile sharing with privacy and integrity of high-risk data (e.g. genomics; highly sensitive EHRs, as well as other sectors). Furthermore, we plan to address this level of threat drawing from building on recent research on powerful and innovative automatic security and dependability techniques, like fault and intrusion tolerance or Byzantine fault tolerance (BFT), trusted computing and architectural hybridisation, secret sharing and secure multiparty computation, self-healing and diversity, or post-compromise security.
- The Supply Chain Security demonstration case has identified "Transparency" (SCH-LR04) as one of their requirements, meaning that adequate information related to the operations of the supply chain has to be given to monitoring organizations and agencies, as well as to the general public, in order to assure them of the compliance properties of the chain. When releasing information as such, care must be taken to balance the privacy loss and utility gain. In WP3 we can study, which information should be released to obtain the optimal trade-off. For this, we have to come up with

utility measures for these agencies, as well as direct the organizations in the supply chain to propose privacy policies and metrics for their data. The same policies and similar measures are also needed in addressing the requirement "Privacy-preserving analytics" (SCH-SP09).

- There are similar issues in the Medical Data Exchange demonstration case, where the privacy of certain parties has to be weighed against the utility requirements of other parties. We have identified the requirements MD-SP01 about privacy-preservation during data sharing, as well as MD-SPL02 and MD-SPL03 about privacy and utility in certain computations as the targets of the research to be performed in WP3.

9 Decentralized Authorization and Distributed Access Control

A blockchain can be thought of as a growing database which is resistant to any subsequent modifications. While initially designed as a permission-less structure (cf. Bitcoin[28]), but soon a variety of applications requiring permissioned blockchains, and different types of consensus protocols, i.e., mechanisms to decide how which blocks to append to an existing blockchain, were introduced. The approaches presented in the following can be used to further increase the possible applications of blockchains, by increasing their scalability and efficiency, and allowing for fine-granular modifications of blocks, e.g., because of legal requirements.

9.1 Fine-Granular Rewriting on Blockchains

Blockchain technologies have recently gained considerable attention. While at the beginning the focus was mainly on the use of blockchains in crypto currencies (e.g. Bitcoin), today there are a large number of other application scenarios. Blockchains have the fundamental property that an object (e.g. a transaction) becomes unchangeable once it has been registered and included in the blockchain.

Although this feature is of central importance in many applications, it is sometimes - for example for legal reasons - necessary to be able to lever out this immutability in a limited and possibly traceable form.

Recently, Ateniese et al [68] proposed a block-level solution based on the use of Chameleon hashes [69]. In contrast to standard cryptographic hash functions, where it is computationally infeasible to find a collision and thereby change the message contained in a block of the blockchain, Chameleon hashes come with a dedicated trapdoor known to the party generating its parameters which allows this party to find arbitrary collisions, thereby becoming able to arbitrarily rewrite the blockchain.

However, unfortunately, in many applications, this approach is not fine-granular enough, as the whole block can be changed in an all-or-nothing manner. That is, the entity knowing the trapdoor can compute arbitrary collisions on a block level, but it is not possible for a party adding an entry to the blockchain to decide who is actually allowed to modify the included information (e.g., transaction).

Using the technology presented by Derler et al [70], it becomes possible to modify single transactions on a very fine-granular level, by specifying predicates that need to be satisfied by a user in order to become able to edit a given transaction. Furthermore, due to the decentralized setting where every entity should be able to play the role of an attribute authority and tag other users with attributes, it is no longer necessary to trust a central entity, but each user verifying a transaction can, at her own discretion, decide whether or not she wants to trust the attribute authority.

The realization of the scheme relies on a careful combination of chameleon hashes with ephemeral trapdoors [71] and attribute-based encryption schemes [72], and the efficiency of the solution has been proven on a prototype level, i.e., efficiency benchmarks have been performed for the resulting cryptographic primitive, but it has not yet been integrated and tested for usability on a large scale in real-world applications.

[28] https://en.wikipedia.org/wiki/Bitcoin

This technology has been developed within the H2020 ECSEL project *SECREDAS*[29]. Within CyberSec4Europe, further investigations to broaden the applicability of the concept are foreseen, in particular by continuing research on the underlying cryptographic primitive of Chameleon hashes. For instance, aspects like long-term security (especially against quantum-attacks), decentralization (e.g., to distribute the right to sanitize/edit a block), or restriction of modifications (to not allow for arbitrary, but only well-defined modifications) are under consideration. The precise focus of further research will be guided by continuous discussions with, and feedback received from, WP5, in order to guarantee for addressing real-world challenges.

Full name	Acronym	Lead partner	TRL
Fine-Granular Rewriting on Blockchains		AIT	3
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2	OB-LR01, SCH-LR01	https://eprint.iacr.org/2019/406.pdf	

Table 19: Metadata: Fine-Granular Rewriting on Blockchains

9.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The above technology is essential to achieve compliance with data-protection regulation in blockchain-based protocols and schemes, as it allows for enabling, e.g., the "right to be forgotten". Regarding the privacy-preserving functional architecture defined in Figure 2, it belongs to the "Blockchain services" component.

9.2 Scalable and Private Permissioned Blockchain

It is thanks to the rising interest in, and intense scrutiny of, Bitcoin that the world made the acquaintance of Bitcoin's underlying technology: the blockchain. It is a technology that allows information to be securely shared, stored, and verified without trust and without the need of a centralized entity.

In the past few years, research efforts and investments in this technology have only increased, and nowadays financial institutions are investigating how to exploit it for their benefits, and countless blockchain-based projects and startups are promoting new ways to employ it. Ethereum [73] and Hyperledger Fabric [74] are the most famous and widely used examples of the potential of the blockchain.

Nevertheless, research has shown that there are a number of important issues to be addressed to make the technology suitable for the fintech world [75]. Namely:

- **Privacy**: existing blockchain architectures work with the assumption that transactions and their order of execution are public. However, the both corporate world and private citizens value their privacy, therefore they restrict data sharing with only a limited number of intended parties. Encrypting the contents of transactions is a good, but insufficient, measure, as it still allows everyone in the network to know when a particular transaction occurred.
- **Scalability**: existing permissionless blockchains (e.g., Bitcoin and Ethereum) scalability is excellent, but is possible at the expense of the network's throughput (e.g., Bitcoin achieves a mere

[29] https://secredas.eu/

7 transactions per second [76]). On the other hand, permissioned blockchain are at the other end of the spectrum, that is, they can achieve higher throughput at the expense of scalability. However, a blockchain architecture needs both scalability and throughput to meet acceptable quality standards that would make it useful for both organizations and private citizens.

- **Lack of Governance**: the distributed and trustless nature of the blockchain is one of its main strong points. However, in real deployments, organizations and service providers want to retain a certain degree of control over their networks in order to enforce business logics and policies, and to be able to selectively grant or deny access to their services.

This asset is a blockchain platform capable of solving the aforementioned issues. Namely, satellite chains are a unique feature of this blockchain architecture that guarantee scalability and transactional privacy. They are small, independent sub-blockchains of a larger network with their own ledger, set of smart contracts, consensus algorithm, and set of participants. If needed, they can transfer assets to each other. The advantage of satellite chains is that they involve only the intended parties, thus limiting the sharing of knowledge to who matters. Transactions exchanged within a satellite chain are visible only to its members. See also Figure 13 for an illustration.

The first immediate consequence of this approach is efficiency: members do not have to receive and store transactions that are not relevant to them. The second consequence is scalability: by allowing an unbounded number of satellite chain to work in parallel, this architecture can easily handle a much greater number of active participants exchanging transactions privately. This is turn, gives this blockchain architecture a higher throughput than existing deployments.

Finally, organizations can enforce business logics and policies via smart contacts, Turing complete programs that can be uploaded to the blockchain.

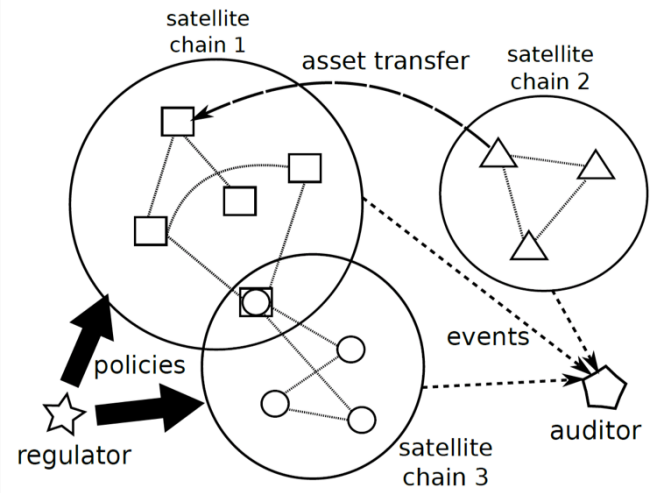

Figure 15: Blockchain platform architecture showing three independent satellite chains in action

Full name	Acronym	Lead partner	TRL
Scalable and Private Permissioned Blockchain		NEC	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2	OB-SP01, OB-SP02, OB-SP03, OB-SP05, OB-LF01, OB-OP01, OB-OP03, OB-LR01, OB-LR02, SCH-SP01, SCH-SP02, SCH-SP03, SCH-SP04, SCH-SP05, SCH-SP07, SCH-SP09, SCH-SP10, SCH-LF01, SCH-OP01, SCH-OP02, SCH-OP03, SCH-OP04, SCH-LR01, SCH-LR02, SCH-LR03	n/a	

Table 20: Metadata: Scalable and Private Permissioned Blockchain

9.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

This asset can act as the permissioned blockchain platform of choice to implement CyberSec4Europe architecture's blockchain services.

9.3 Scalable and Efficient Consensus Algorithms

Permissioned blockchains achieve high throughput at the expense of scalability. The cause of this imbalance is the consensus algorithm that these blockchain use. Namely, permissioned blockchains leverage Byzantine Fault Tolerant (BFT) protocols to reach consensus between nodes. Unfortunately, state of the art BFT protocols cannot scale to a large number of nodes, and as such can only realistically support the construction of networks amongst a few tens of nodes [77].

Recently, Proof-of-Stake (PoS) algorithms have emerged as a possible scalable alternative to traditional BFT protocols [78,79,80,81,82,83], but they were still proved to be vulnerable to a number of attacks [84,85,86].

Within CyberSec4Europe, we will look for answers to the scalability question, and design efficient and scalable consensus algorithms. The desirable outcome would be protocols able to reduce the number of participants required to resist faults in the system, and reach consensus with fewer communication rounds. These two features would tremendously increase the overall system performance.

Additionally, we plan to use such novel consensus protocols in the blockchain platform introduced in section 9.2. The combination of the two technologies would allow for a system architecture able to scale to thousands of nodes without sacrificing throughput.

Full name	Acronym	Lead partner	TRL
Scalable and Efficient Consensus Algorithms		NEC	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2	OB-SP03, OB-SP04, OB-OP01, OB-OP02, OB-OP03, SCH-SP03, SCH-SP04, SCH-OP01, SCH-OP02, SCH-OP03	n/a	

Table 21: Metadata: Scalable and Efficient Consensus Algorithms

9.3.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The results of these research efforts will be leveraged by CyberSec4Europe architecture's blockchain services as the consensus algorithm(s) of choice of the permissioned blockchain.

9.4 Research Challenges Identified in WP5

The following research questions have been identified so far:

- Regarding the editing of blockchains, multiple challenges may improve the applicability and versatility of the tool. For instance, restricting the way of possible modifications (e.g., solely to redactions) might be sufficient for realizing the right to be forgotten, but no further modifications might be desirable.
- Efficiently distributing the rewriting capabilities, e.g., using approaches from threshold cryptography, might be desirable in order to avoid a single point of failure.
- Long-term security guarantees managing the rights to edit entries in the blockchain are highly desirable for long-term protection of its integrity. On the one hand this requires post-quantum instantiations to also protect against the advent of large-scale quantum attackers. On the other hand, it should be possible to withdraw the right to edit certain entries if it is no longer needed, in an efficient and compact way, e.g., leveraging approaches from puncturable cryptography.
- As mentioned earlier, improving the scalability of consensus mechanisms is of key relevance for large-scale applications. We must also investigate novel methods to preserve blockchain users' data privacy. An important characteristic of the blockchain is its immutability: once uploaded to the ledger, records cannot be modified anymore. However, recent EU regulations – such as GDPR – say that users must give their consent for their data to be used. Furthermore, users can request their data to be deleted at any time. Therefore, EU regulation compliance should be a key research aspect for the project.

10 Privacy- and Integrity-Preserving Storage and Processing

Despite the many advantages of outsourcing data storage and processing, such as scalability and availability, there are also relevant drawbacks regarding privacy and integrity of the stored data, as full control over the data is given to a third party. While for plain storage, traditional cryptographic mechanisms like encryption and signatures can be employed to solve most related issues, things become more complicated if also computations on the outsourced data should be performed by the cloud providers.

It is therefore of paramount importance to develop mechanisms, tools, and guidelines to ensure the long-term integrity and privacy of outsourced data against cloud providers in case of data leaks, hacks, or intentional misbehavior.

10.1 ArchiStar and SECOSTOR Secure Distributed Storage

ArchiStar and SECOSTOR [87] are two versions of a tool for secure, distributed storage of data in the cloud. The data is fragmented and stored in different locations. The use of secret sharing technologies [88] ensures high redundancy and availability on the one hand and ultimate security guarantees on the other.

More precisely, secret sharing schemes allow one to decompose input data into *n* fragments (or shares) such that the original data can be recovered from any *t* shares. However, any collection of at most *t-1* shares does not contain any information about the original data in an information-theoretic sense. That is, the data remains secure even against powerful future adversaries that may have access, e.g., to quantum computers, as long as less than *t* storage locations are corrupted. Additionally, it must be noted that the confidentiality and availability guarantees in such schemes are achieved in a fully keyless manner, i.e., no key management issues arise on the end-user's side.

The advantages of ArchiStar and SECOSTOR are multifold, including the following:

- **Confidentiality and availability** increase equally in these systems: on the one hand, data remains secret as long as less than *t* storage nodes are corrupted. On the other hand, availability is guaranteed as long as at least *t* storage nodes are available. By choosing, e.g., *n~2t*, increasing the threshold simultaneously provides benefits for both properties.
- The amount of data to be stored per storage node is equal to the size of the original data. However, this **storage overhead** can be decreased significantly if one is willing to give up the information theoretic confidentiality guarantees and rely on the security of (symmetric) encryption schemes such as, e.g., AES.
- ArchiStar and SECOSTOR leverage existing and publicly available technologies. They do not depend on any specific storage offerings, and all used data formats are fully open and documented. By doing so, the tools contribute to **prevent vendor locking** and contribute to an open environment for secure outsourcing of data.
- By using encoding mechanisms instead of encryption, the entire storage solution provided by ArchiStar and SECOSTOR is **fully keyless**. Security is governed by non-collusion of storage providers and not by private keys. By doing so, no complex key management procedures need to be put in place. Nevertheless, by solely relying on non-collusion assumptions, the storage locations need to be carefully chosen when setting up the system.

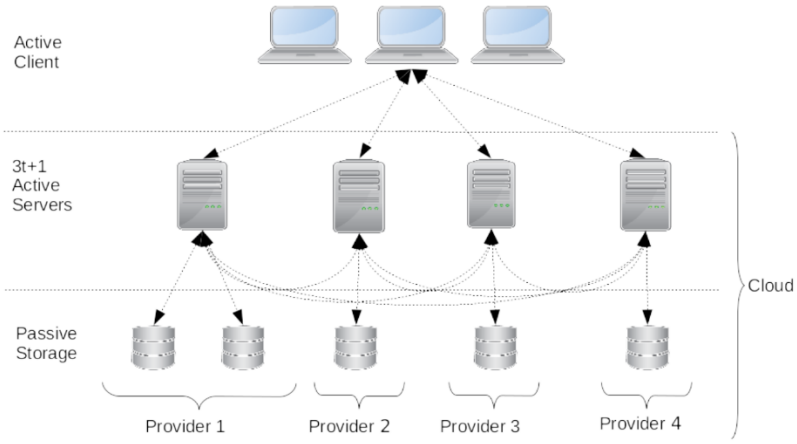

Figure 16: SECOSTOR architecture

Figure 16 gives an overview of the SECOSTOR architecture. One or more clients communicate with active components on the storage servers' side. By choosing *n=3t+1* and deploying Byzantine fault tolerant (BFT) [89] systems, concurrent (write) access to the data can be supported, without risking inconsistency of the stored data fragments. The active components then store the data fragments on passive storage nodes.

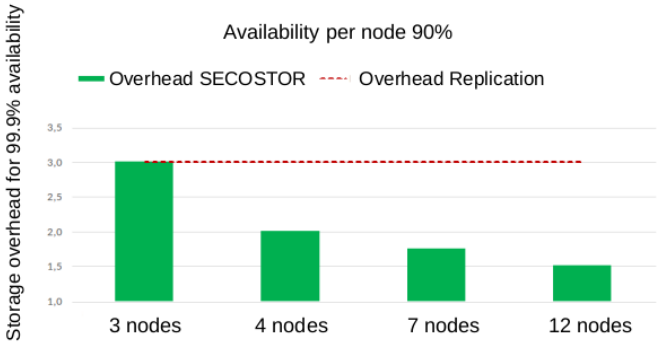

Figure 17: Storage overhead for different *(n,t)* for fixed availability

Besides the BFT components, SECOSTOR, in particular, deploys the ArchiStar secret sharing libraries and components, which can also be used directly if concurrent write access is not needed. In this case, choosing *n=2t+1* is optimal in many situations in order to balance security and availability.

Finally, Figure 17 compares the storage overhead of computationally secure variant of SECOSTOR with plain redundant storage, when aiming for a predefined availability level using storage nodes with lower availability guarantees.

The ArchiStar and SECOSTOR platforms have been partially developed within the FFG project *ArchiStar*[30] and the H2020 project *PRISMACLOUD*[31]. While the existing code base is already relatively stable and also being commercialized together with a local SME[32], additional features (e.g., integration with consumer

[30] http://archistar.at/
[31] https://prismacloud.eu/
[32] https://www.fragmentix.com/

frontends like OwnCloud), efficiency improvements, and alternative schemes are currently under consideration and might be further developed within the context of CyberSec4Europe.

Full name	Acronym	Lead partner	TRL
Secure Distributed Storage	Archistar/SECOSTOR	AIT	7
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.5, T5.6, T5.7	OB-SP12, OB-SP16, OB-LF02, SCH-SP06, SCH-LR01, MT-SP19, MD-SP01, MD-SPL01, SMC-SP16	http://archistar.at/ https://github.com/archistar	

Table 22: Metadata: Secure Distributed Storage

10.1.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Archistar and SECOSTOR are one approach to allow for confidentiality-preserving and end-to-end secure sharing of sensitive data among stakeholders. Furthermore, the storage model is compliant with the data analytics approaches presented in Sections 10.2 and 10.3.

With respect to the architectures presented in Figure 1 and Figure 2, Archistar belongs to the user domain. It can be considered as a user-side security/privacy tool, and, in particular when being used in combination with one of the following building blocks, as a PET client.

10.2 Sharemind MPC – Privacy-preserving data analysis

There is a strong push towards regulating data sharing (GDPR, eIDAS). Sharemind[33] MPC is a data storage and analysis platform that uses secret sharing in order to hide individual values from the data analyst while allowing data analysis to be performed.

In this solution, the data owners (input parties) use additive secret sharing to distribute their data among several, e.g., three computing parties. The computing parties are unable to recover individual values. The result parties send queries to the computing parties, who use secure computing protocols to perform the requested analyses and return the results. The result parties can then use the result recovery algorithm and learn the result of the computation.

Sharemind applications seek to achieve the following four security goals [90,91,92].

1. Cryptographic privacy – no computing party shall learn a private value held by an input party.
2. Source privacy – no computing party shall be able to relate the results of a computation to the inputs of a certain input party.
3. Query restrictions – no result party or an unauthorised set of computing parties shall be able to successfully initiate and complete a query that has not been authorised by the computing parties.
4. Output privacy – no result party shall learn a private value held by an input party from the results of queries.

[33] https://sharemind.cyber.ee/

Sharemind MPC is a distributed computing system where all parties run a part of the Sharemind MPC platform. Input parties and result parties run client applications that connect to Sharemind MPC, provide data, request computations and return results. These applications are responsible for automatically secret sharing the inputs and recombining the outputs of computations and communicating with the application server. The application server is a distributed system that upon receiving queries client applications, runs the cryptographic protocols to complete the requested computing tasks. After receiving a request, the servers authenticate the client and apply access control restrictions. If these checks pass, the servers synchronously start a privacy-preserving computing process based on a program description stored at each server and retrieved by name according to the query. This feature helps achieve the query restrictions security goal, as the agreement of all computing parties is needed to perform a particular query. After completion, the application server returns the results to the client application.

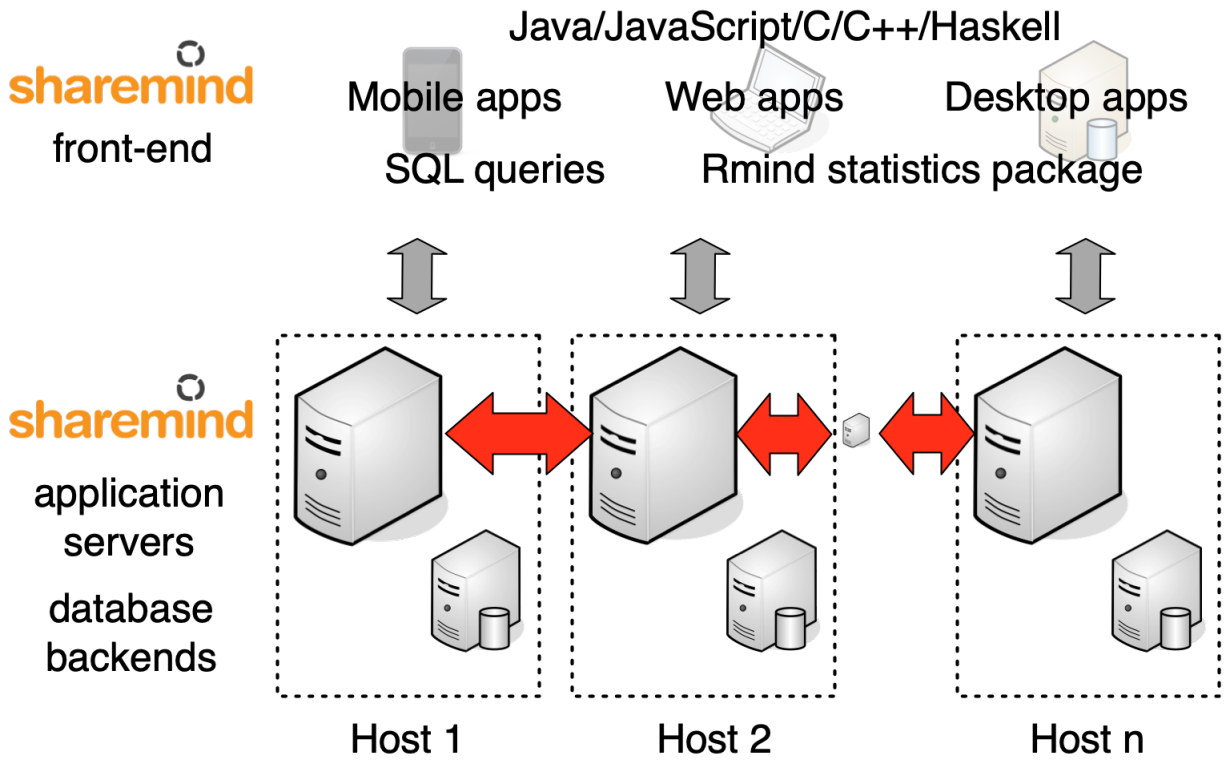

Figure 18: Sharemind platform deployment

Sharemind MPC provides central services for randomness generation, secure channels and access control for its applications. These are modular and can be replaced with newer implementations as required. While the Sharemind MPC platform supports various secure computing techniques like homomorphic encryption and trusted execution environments, we look at secure multi-party computation (MPC).

Sharemind has a two-level programming model. Privacy-preserving business logic is implemented using the SecreC 2 programming language [93]. SecreC is a high-level procedural language for writing privacy-preserving algorithms without any knowledge of the underlying cryptographic protocol.

Secure instructions in SecreC are designed to be composable sequentially, to form programs, and, in parallel, to enable SIMD operations [94]. SecreC provides different data types (integers, Booleans, floating-point numbers, fixed-point numbers, dates, text) and operations.

Sharemind MPC provides standardised tools for commonly occurring tasks such as the Rmind statistical analysis environment. Rmind is similar to the popular statistical analysis environment R (The R Project for Statistical Computing: https://www.r-project.org). Both are interactive environments where the statistician can enter commands to perform data manipulations and analytics, however, R works with data available locally, while Rmind connects to a remote Sharemind MPC instance. Rmind parses commands from the user and connects to Sharemind MPC using the client API to complete the queries on uploaded data tables. Rmind's SecreC implementations enforce the query restrictions configured by the computing parties to ensure that the user follows the study plan and does not break output privacy.

Rmind supports privacy-preserving calculation of new attributes from existing ones, date manipulation, filtering, descriptive statistics, outlier detection, statistical tests, multiple tests, linear regression, linear modelling, sorting, aggregation, database join and plotting results as graphs. Rmind is designed to support exploratory data analysis where the statistician has limited freedom to pick the next hypothesis to check.

Full name	Acronym	Lead partner	TRL
Privacy-preserving data analysis	Sharemind MPC	CYBER	9
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.3, T5.6	OB-SP05, OB-SP06, OB-SP12, OB-LF02, OB-LR01, OB-LR02, OB-LR03, OB-LR07, SCH-SP06, SCH-SP07, SCH-SP08, SCH-SP09, IDM-SP07, MD-SP01, MD-SP02, MD-SP06, MD-SPL01, MD-SPL02, MD-LR01	https://sharemind.cyber.ee https://dspace.ut.ee/handle/10062/29041 https://ieeexplore.ieee.org/document/7505613	

Table 23: Metadata: Privacy-preserving data analysis

10.2.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

The main purpose of the Sharemind MPC platform is to enable privacy-preserving data storage and analysis. As such it falls under Security/Privacy-preservation tools in the Managed domain. It also offers PET clients (User domain) for both input parties (tool for secret-sharing and uploading data) and output parties (tool for privacy-preserving data analysis).

10.3 FlexProd – Integrity-Preserving Data Analytics

Secure multi-party computation (MPC) allows parties to jointly evaluate a function while keeping the respective inputs private. This is leveraged by the FlexProd[34] platform to provide means to analyze data from potentially different stakeholders in a way that gives high authenticity guarantees on the computation's result, while protecting the confidentiality and privacy of the input data. This is achieved by combining secret sharing technologies [88] with zero-knowledge proofs of knowledge to guarantee correctness of the evaluations and blockchain technologies to guarantee integrity of the input data. On a high level, a zero-knowledge proof of knowledge allows one to prove that a certain statement is true (e.g., that one knows a

[34] https://flexprod.at/

secret key) without revealing more than the validity of the statement (e.g., without disclosing the secret key).

Conceptually, the FlexProd platform works as follows, where we assume that each party potentially holds a variety of potentially authenticated (i.e., signed) data items as inputs.

- In the first step, each party decomposes her data using secret sharing into $n=3$ shares using a threshold of $t=2$. Each of these shares is then encrypted for one of the three FlexProd computation nodes, using the public encryption key of the respective node. Furthermore, it encrypts the input data under her own public key.
- Next, the party computes a zero-knowledge proof of knowledge that shows that the decomposition and encryption were done correctly, i.e., that the plaintexts contained in the computation nodes' ciphertexts correspond to consistent shares of the plaintext encrypted for the user. Optionally, if the data was previously signed, the proof further shows that the original data was indeed authenticated by a legitimate party. All ciphertexts, as well as the zero-knowledge proof, are stored in a blockchain to avoid future changes on the input data.
- In a third step, each of the computation nodes verifies the zero-knowledge proofs for all inputs necessary to perform a given computation. The nodes then decrypt their own shares and engage in an MPC protocol for the given computation.
- The nodes then jointly compute a zero-knowledge proof of knowledge which shows that the output value has indeed been computed correctly. The output value and the resulting zero-knowledge proof are again stored in the blockchain, potentially encrypted under the key of the designated receiver of the result.
- Finally, the receiver can decrypt the result and verify the zero-knowledge proof to check the correctness of the computation.

Besides high privacy guarantees for the data owners, this sketched approach bears multiple advantages for all involved stakeholders. Firstly, the data providers do not need to communicate with each other but can register their data at their own convenience. Secondly, the same data can be re-used in many computations, i.e., each data provider needs to register their data in the blockchain only once. Finally, the receiver gets high authenticity guarantees as he can verify the validity of the input data by verifying the data sources' and the nodes' zero-knowledge proofs, thereby receiving end-to-end authenticity guarantees without having to rely on any trust assumptions to either entity in the system, while still allowing for computations across different data owners and data sources.

While the above approach works generically for any type of data analytics, its efficiency highly depends on the concrete computation to be performed. Up to date, the FlexProd platform has been implemented and validated for the purpose of confidentiality-preserving auctions, where the goal is to identify the maximum of a number of inputs (i.e., the highest bid). While an extension to simple analytics (e.g., mean value) is straightforward, enabling more complex computations is future work.

The current version of the above platform has been designed as part of the FFG project *FlexProd* and is planned to be further extended in close collaboration with CyberSec4Europe as well as the H2020 project *KRAKEN*[35].

Full name	Acronym	Lead partner	TRL
Integrity-Preserving Data Analytics	FlexProd	AIT	4
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.6	OB-SP06, OB-SP12, OB-LF02, OB-LR01, OB-LR02, OB-LR03, SCH-SP09, MD-SP01, MD-SP02, MD-SP06, MD-SPL02, MD-LR01	https://flexprod.at/	

Table 24: Metadata: Integrity-Preserving Data Analytics

10.3.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

Considering Figure 2, FlexProd is a platform for privacy- and integrity-preserving data analysis, and thus falls into the Security/Privacy-preservation tools in the Management domain, and is compatible with the storage solution presented in Section 10.1. Also beyond Task T3.2, the platform can contribute to the threat intelligence analysis and sharing components in the Intelligence plane presented in the overall WP3 architecture in Figure 1.

10.4 Argus - Enforcing Privacy and Security in Public Cloud Storage

ARGUS [95] is a privacy brokerage system aiming at enhancing user trust in public cloud storage systems, guaranteeing data confidentiality and improving availability.

Guaranteeing Confidentiality

ARGUS guarantees confidentiality in different ways. First, by cyphering user data before storing it on the Cloud. Secondly, by partitioning data and storing each data partition into different Cloud providers. The former guarantees confidentiality since only the user is able of accessing data his/her data due to privately stored cypher keys. The latter improves confidentiality, since it prevents a single provider from gaining access to all the necessary data, even if they are somehow able to get hold the encryption keys.

To guarantee the confidentially of user cloud credentials, ARGUS uses Intel SGX technology, using secure enclaves for cyphering access tokens, thus guaranteeing confidentiality in case of data breaches.

Improving Availability

Argus improves the availability over current Cloud storage providers using erasure coding. This partitions data and simultaneously creates parity blocks for recovery improving redundancy. Each partition and parity block is stored into different Cloud providers, thus, in a scenario where three different cloud providers are

[35] https://www.kraken-project.eu/

used, even if a single cloud provider is unavailable, data stored on the remaining is sufficient for ARGUS to guarantee user access, by recovering absent date from the recovery blocks.

Integrity verification

For data integrity, we store a Hash-based Message Authentication Code (HMAC) that allows ARGUS to ensure integrity of the stored or transmitted information over an unreliable channel. A new HMAC is produced for each file, before splitting the data, to ensure that the data was not tampered with while being hosted in the public cloud. For complete security against lost data, at least three cloud providers are required, so there is at least one parity block for every two data blocks.

Additional Features

ARGUS can be configured for addressing limitations of mobile devices, offloading computation from the clients to a privately configured server, making it viable for resource-constrained devices, such as smartphones.

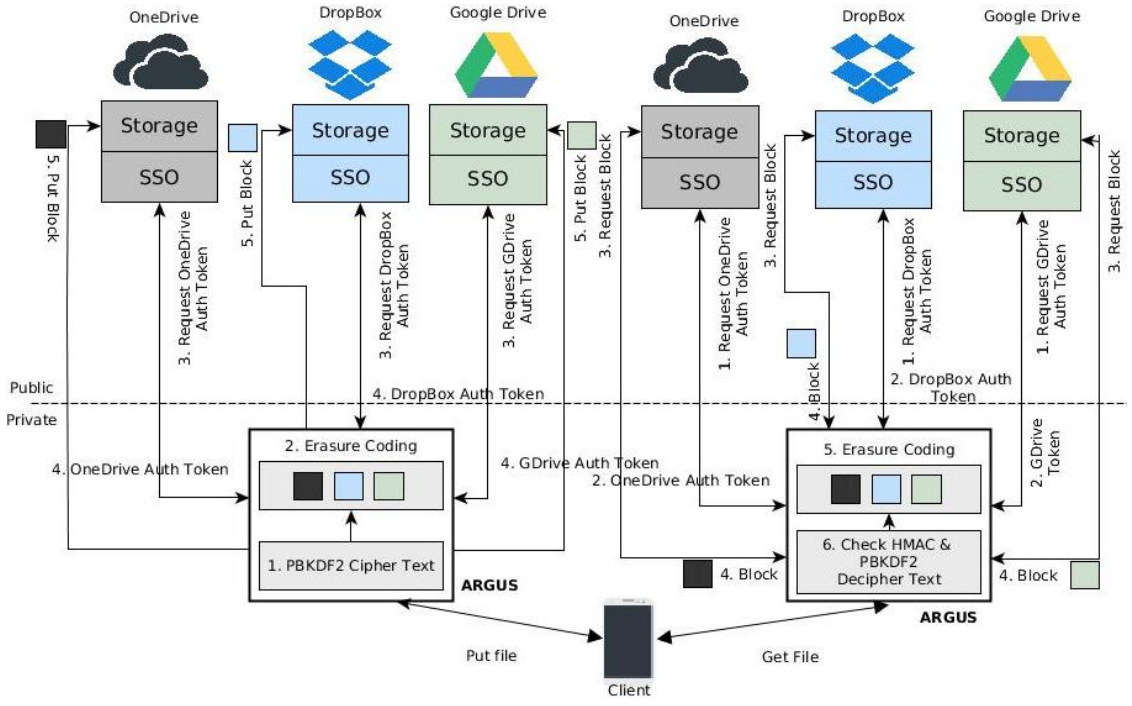

Figure 19: Argus architecture

Figure 19 depicts ARGUS architecture, presenting its behavior when storing and accessing data.

Storing data

When the broker receives a request to upload a file, it performs three different tasks. First, it decides if the received file is in clear text (i.e., its contents have yet to be cyphered) or if the contents have already been ciphered. If the contents have yet to be cyphered, ARGUS does so using a predefined symmetric key to

cypher the document ((1) in Figure 19). Second, it applies the erasure code algorithm, partitioning the file into two partitions, and creating an additional recovery block (parity data) ((2) in Figure 19). Third, it collects authorizations from the public clouds, i.e., the system must get the authentication of the user using OAuth2 or request credentials from Intel SGX, assigns each partition (parity block inclusive) to a cloud providers ((3) and (4) in Figure 14), and uploads each partition to the respective provider ((5) in Figure 14).

Accessing data

When the broker receives a request to access a file, it performs the following tasks. First, the system verifies its permissions to search the cloud providers for the corresponding partition of the file that is being requested by the user ((1) and (2) in Figure 14). Second, the system downloads the necessary partitions for reconstructing the original encrypted file and applies the erasure code ((4) in Figure 14). In order to recover the original cyphered file, at least two out of three partitions are required. When the system gets them, the cyphered file is reconstructed and the stored HMAC [96] is validated against the recovered file ((5) and (6) in Figure 14). Third, if previous steps are concluded successfully, the system sends the file to the user.

Related Work

	Resilio	AeroFs	Safecloud Photos	Boxcryptor	Whisply	NetApp	ARGUS
Share Files	X	X		X	X	X	X
Dropbox API			X	X	X	X	X
OneCloud API			X	X	X	X	X
Google Drive API			X	X	X		X
Priv. Cloud support		X		X		X	X
Encrypted Public				X	X		X
Erasure Coding							X
ZK Encryption	X	X		X		X	X
Collaborative						X	X
Computation Offload							X
Data Partitioning							X

Table 25: Comparison of ARGUS and related systems

Current cloud systems try to overcome issues regarding trust in the provider, limitation of sharing files, or bottleneck of having a local hybrid cloud. However, there are new limitations. Users nowadays use edge devices, such as mobile or IoT devices, to send information to the cloud. The traditional cloud systems do not address the limitations of these systems. An outline of the current state-of-the-art systems is depicted in

Table 25. For this comparison, we have included Resilio[36], AeroFs[37], Safecloud Photos, Boxcryptor[38], Whisply[39] and NetApp[40].

The literature also describes uses of NFS (Network File System), or other technologies that allow the users to communicate with the Cloud. An example can be the implementation of a Shared Cloud-backed File System (SCFS [97]), where the user needs to interact with a coordination service. The computation is completed locally with an extension of Depsky [98]. The use of coordination's systems is not the ideal for IoT devices because the computation is intensive and all the tasks are performed by the user device, the coordination service just grants permission to write in the cloud and solves duplication's and detection of collisions. An example of these systems is SCSS and SCFS. Systems such as the FUSE-based file system, backed by Amazon S3 (S3FS[41]), typically not solve the limitation of sharing a file with a third party. In ARGUS, we address all the limitations of current implementations, joining efforts of multiple projects in one, with users able to choose from public or private cloud providers along with the computation offload to guarantee data partitioning and security in the storage providers.

Full name	Acronym	Lead partner	TRL
Enforcing Privacy and Security in Public Cloud Storage	Argus	C3P	5
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.6	MD-SPL01	https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8514195	

Table 26: Metadata: Enforcing Privacy and Security in Public Cloud Storage

10.4.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

This asset can be used for improving privacy during access to On-Line Service Providers.

10.5 GDPR compliant user experience

Regulation (EU) 2016/679 of the European Parliament and of the Council or more commonly known as General Data Protection Regulation (GDPR) is a legal framework that sets guidelines for the collection and processing of personal information. The regulation applies across the entire European Union (EU) and European Economic Area (EEA) region. The regulation was designed to give the citizens of the EU and EEA greater control over their personal data and ensure that their information is being adequately protected. According to the GDPR, personal data is any information related to a person such as a name, a photo, an email address, a computer IP address etc. For any entity that processes personal data and does not comply with the regulation, the GDPR stipulates harsh fines. In turn, the companies processing personal data have put a lot of effort into compliance with the GDPR requirements. However, like any other legislation, and even more so as the GDPR is meant to be a framework, the exact meaning or the correct implementation of it, is often subject to interpretation. The final interpretation of the GDPR is within the jurisdiction of the

[36] https://www.resilio.com
[37] https://www.aerofs.com
[38] https://www.boxcryptor.com/en/
[39] https://whisp.ly/en
[40] https://www.netapp.com/us/index.aspx
[41] https://github.com/s3fs-fuse/s3fs-fuse

European Court of Justice. However, the opinions of the supervisory authorities are also very relevant as they are responsible for investigations of non-compliance, the corrective powers they hold and their knowledge of local (member) specific legislation (GDPR relevant and other).

For this reason, we would like to present just a few cases of frequently encountered situations that are often hard to implement without further interpretation than is provided by the GDPR itself.

One of the basic goals of the GDPR is the idea of data transparency, which gives the person the ability to know which data is being processed and for what purpose [99]. This information must be conveyed in a concise, transparent, intelligible and easily accessible form, using clear and plain language. These are tricky and subjective criteria, that are hard to measure and need constant revision as the perception and tendencies of users change. A separate issue is the data obtained indirectly, where and immediate notification of users is not possible and collection of data from minors.

GDPR further requires (Article 4) that the controller/processor reports personal data breaches to the regulator. Consequently, after the introduction of GDPR, some of the Information Commissioner's Offices were reporting a sharp increase in the number of reported security incident [100]. This was not a surprise, given the severe sanctions non-compliance with this requirement can attract under Article 83, and potentially under Article 82. However, many of the reports were unnecessary as the regulation only stipulates reporting of personal data breaches that are likely to result in a risk to the rights and freedoms of natural persons and not all forms of a data breach. The circumstances in which the breach must be reported to the regulator and possibly also to the natural persons are therefore clearly not defined well enough.

Some misconceptions regarding the need for consent and what are the best lawful bases for the collection and processing of any personal data have taken hold amongst organizations [101]. While there are specific circumstances where this is true, this is not a rule to blindly follow. Additionally, when consent is given, it must be informed, specific, freely given, granular, revocable, affirmative, and recorded. How to achieve all these conditions can again be challenging for organizations that also have to take into account that consent may be revoked at any time.

GDPR (Article 22) prohibits decision making based solely on automatic processing of personal data, that could have a legal or similarly significant effect for the data subject. When does an effect become significant for a data subject is not defined and can cause problems to the entities that process personal data for some decision making or profiling [102]. The banks are mostly taken as examples when determining creditworthiness using automated means, but whether it is rightly so is still to be determined.

GDPR (Article 35) has introduced the mandatory Data Protection Impact Assessments whenever the processing of personal data is "likely to result in a high risk to the rights and freedoms of natural persons". The regulation does not address in detail how and when the assessment itself should be conducted, therefore further guidance is needed [103].

Under the GDPR data protection by design and by default is a requirement (Article 25). However, in the regulation neither the concepts of protection by design nor of protection by default are well defined or explained, and consequently, organizations sometimes find it hard to know exactly how to implement the concept practically [104]. Those concepts are not yet fully included in current development methodologies and lifecycles. The protection by design and by default is specifically complicated for organizations because it requires technical and organizational measures (or controls) that are specific for each product, service or

system. Therefore, a list of requirements to achieve protection by design and by default could be constructed, with some general guidelines or more specific examples. Furthermore, the role of certification in relation to the process of complying with the protection by design and by default could also be better described. The results may also impact the development of ISO/PC 317 standard that is in its infancy.

GDPR compliant user experience is a solution (see deliverable D3.1, Section 5.6, asset *Guidelines for GDPR compliant user experience*) that collects important interpretations of the regulation, together with good implementation examples to meet the specified requirements. As we have shown there are many open questions regarding GDPR compliance and this solution provides a singular platform, where users can look for clarification of the specific GDPR requirements and legitimate ways the regulators have suggested of addressing them. As the primary principle behind the GDPR is that it views personal data as the property of the individual/natural person, it applies to all companies selling to and storing personal information about citizens in Europe, including companies on other continents. The created guidelines will, therefore, be useful not only to the EU and EEA states but to the wider audience.

GDPR compliance and the help to achieve it could be listed under a few sections that are a part of this document. We chose to primarily list it in "Privacy- and Integrity-Preserving Storage and Processing" as the GDPR is primarily focused on the privacy, storage and processing of personal data.

Full name	**Acronym**	**Lead partner**	**TRL**
GDPR compliant user experience		UM	6
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.3, T5.4, T5.5, T5.6, T5.7	OB-SP01, OB-SP05, OB-SP06, OB-SP12, OB-SP27, OB-LF02, OB-U01, OB-LR01, OB-LR03, OB-LR08, OB-LR09, SCH-SP01, SCH-SP08, SCH-U01, SCH-LR01, SCH-LR04, IDM-SP06, IDM-SP07, IDM-SP08, IDM-U03, IDM-LR01, IDM-LR02, IR-F01, IR-F06, IR-SP01, IR-LR02, MT-SP01, MT-SP19, MD-SP01, MD-SP02, MD-SP05, MD-SP06, MD-OP01, MD-SPL01, MD-SPL02, MD-LR01, SMC-F08, SMC-F10, SMC-F11, SMC-F12, SMC-SP10, SMC-SP11, SMC-SP13, SMC-SP14, SMC-SP16, SMC-SP17, SMC-SP19, SMC-SP21, SMC-OP05, SMC-MP02, SMC-MP03, SMC-LR02	https://cybersec4europe.eu/wp-content/uploads/2020/02/D3.6-Guidelines-for-GDPR-compliant-user-experience-Submitted.pdf	

Table 27: Metadata: GDPR compliant user experience

10.5.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

GDPR guidelines for compliant user experience can be used in the Administration plane to construct privacy-compliant governance and management practices throughout adaptive security lifecycle. Regulatory management impacts other key components in the Administration Plane:

Regulation is the basis for Risk and Incident Management as risk analysis needs to take into consideration regulation-based risks and risks of regulatory fines. Incident management needs to consider timeframes and triggers for mandatory notification.

Policy-Based Security Management needs to consider policy implications on privacy risks and should, therefore, be linked to data privacy impact assessments.

Security modelling may be found as a complementary tool to data privacy impact assessments and should consider privacy by design and privacy by default requirements.

Regulatory management indirectly impacts the Intelligence plane and the Control and Management Plane. In the Intelligence Plane, the assessment of compliance with legal requirements must be established and carried out. Assessment should be linked to data privacy impact assessment and risks that were identified in the Administration Plane. This is also the point where actual decisions will take place, e.g. decision whether an incident qualifies for mandatory reporting to the supervisory authority.

10.6 Interoperability and cross-border compliance

The purpose of this building block is to highlight some of the problems with interoperability and cross-border compliance, primarily with the Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market, also more commonly known as eIDAS. The mentioned trust services include electronic signatures, electronic seals, time stamps, electronic delivery service, and website authentication. Together with electronic identification, they allow for trust, security and legal certainty in electronic transactions. This regulation applies across the entire European Union (EU) and European Economic Area (EEA) region. The basic idea of the eIDAS regulation is for all the Member States that offer an online public service for which access is granted based on an electronic identification scheme, then they must also recognize the notified electronic identification of other Member States.

Each of the member states was required to implement the EU Electronic Signature Directive into their national law. This caused two undesirable outcomes. In some cases, the local legislation was not produced in time to support the rollout of the eIDAS. The freedom the regulation left the member states when designing their own systems, has also led to problems. Different member states have proposed and implemented different solutions that are not necessarily compatible between member states, in turn defeating the principal idea behind the eIDAS. The main objective of this work is to identify these discrepancies between member states.

Further, member states were left with the freedom to regulate own measures in other areas of electronic commerce. This is leading to the position where other regulations come into conflict with the eIDAS regulation. This is blocking further harmonization of the Single European Market.

For example, a member state may not accept certificates issued under eIDAS in some special cases, e.g. when opening a bank account. This is closing some markets to competitors and is completely defeating the purpose of the regulation. Such small details, unfortunately, stretch to other industries and have adverse effects on them. In the presented case IT industry offering online identification of natural persons cannot provide their services cross-border, resulting in market segmentation.

The regulation sets several requirements for advanced electronic signature. However, just their definition raises a number of questions:

- The signature must be uniquely linked to the signatory: However, digital signatures are linked to the private key, which are typically stored on a computer that is secured in some other way [105]. As a result, the digital signature is not directly linked to the user.
- The signatory must have sole control of the signature creation data, however, when a private key is used, a recipient will still not be able to confirm if the owner of the key was actually the one who used it. This problem is exacerbated in real-world where electronic signatures are now offered as cloud services. These services need to use highly secure environments (e.g. hardware security modules) to store the private key in order to keep eIDAS compliance. Though authentication of the users invoking cloud signature functionality is many times questionable (e.g. username and password-based authentication). Therefore, a significant gap exists between regulatory requirements and actual security resulting from those requirements.
- An advanced electronic signature must be constructed in a way that any change in the signed data is detectable. Digital signature based on asymmetric cryptography does satisfy this requirement, but even these types of signatures are not immune to attack. Therefore, the only technical solution that is completely compliant with this requirement is holding a complete copy of the signed document in the signed data itself, which is impractical as a single electronic signature would take hours to complete. If requirements are to be followed literally, then almost none of the technical solutions used today to comply with the regulation.
- Establishing an efficient and usable infrastructure of electronic identifications and trust services across the member states demands adaptation and integration of many systems and legislation of the members, that were originally established and run by different entities. Each of the 28 Member States of the EU was required to implement the EU Electronic Signature Directive into their national law. However, the Electronic Identification and Trust Services Regulation applies directly to every EU Member State. This means that many laws, if not every law, might need to be amended in due course. Further many businesses can't properly distinguish between trust levels and don't understand which one they should be using.

Given the discussed situation, this asset will, as part of the interoperability and cross-border compliance, identify differences between the implementations of eIDAS legislation in different Member States as well as between different industry sectors. These inconsistencies will be demonstrated in selected cases of deployed solutions. An analysis of the security characteristics of selected use cases will also be included. In response to the findings, we also plan to try and identify possible solutions or ways in which problematic areas can be improved upon. The result will be a catalogue of present issues regarding interoperability and cross-border compliance in regard to the eIDAS.

Full name	Acronym	Lead partner	TRL
Interoperability and cross-border compliance		UM	4
Addresses requirements from	**Addressed requirements**	**Further information**	
T5.1, T5.2, T5.3, T5.4, T5.5, T5.6, T5.7	OB-SP01, OB-SP02, OB-SP08, OB-LR01, OB-LR08, OB-LR09, SCH-SP01, SCH-SP02, SCH-LR01, IDM-SP03, IDM-LR01, IDM-MP01, IDM-LR02, IDM-LR03, IR-SP01, IR-LR02, IR-LR03, MT-SP01, MT-SP05, MT-SP06, MT-SP07, MT-SP19, MT-SP22, MT-OP05, MT-OP06, MT-MP01, MT-MP02, MT-MP03, MD-SP07, MD-OP02, MD-LR01, SMC-F03, SMC-SP01, SMC-F10, SMC-SP03, SMC-SP10, SMC-LR02, SMC-MP01	n/a	

Table 28: Metadata: Interoperability and cross-border compliance

10.6.1 Relation to the CyberSec4Europe Privacy-Preserving Functional Architecture

eIDAS falls into the Regulatory Management by definition. Similar to GDPR and other requirements eIDAS needs to be included in the complete governance and management lifecycle to provide governance from strategic starting point, to implementation, operations and compliance monitoring process and activities.

10.7 Research Challenges Identified in WP5

The following research challenges have already been identified:

- Depending on the specific application domain of tools like FlexProd or Sharemind, it is an interesting research challenge to develop efficient ZKP technologies such as SNARKS or similar which allow for proving the correctness of the performed computation in a compact way as part of the output, in order to reduce the necessary trust assumptions into the MPC nodes. Namely, such techniques would allow the MPC nodes to add a compact cryptographic proof that the computation was done correctly, without the verifier needing to rely on non-collusion assumptions among the nodes. SNARKS and similar already exist for certain types of computations, however, it would be interesting to support a broader range of statements, as well as design them having the distributed setup of MPC in mind to optimize efficiency.
- As privacy-preserving analytics are needed in the different verticals in WP5, research into privacy- and integrity preserving data processing is very relevant, especially for open banking and medical data exchange. This is due to regulatory requirements such as GDPR and ePrivacy, but also due to privacy becoming more important in the marketplace as a differentiator. For otherwise equivalent services, privacy could be a tiebreaker. In open banking, we can use the technology for preventing the leakage of actors' sensitive information (OB-SP05), performing privacy-preserving analytics (OB-SP06), avoiding clear storage of sensitive data (OB-SP12, OB-LR03), and allowing data

sharing while complying with the bank secrecy principle (OB-LR07). While doing this, it would be possible to provide fraud protection (OB-LR02) while preserving the privacy of the users. Similarly, for medical data exchange, which also deals with sensitive personal information, privacy-preserving data analysis can be useful (MD-SP01). We can use secure multi-party computation to analyze data sharing by multiple parties while still maintaining anonymity of the data donor (MD-SPL02). The need to preserve the integrity of data is also a research challenge that we can research (MD-SP06). The GDPR requirements of both verticals (OP-LR01 and MD-LR01) are also easier to comply with when using privacy-preserving techniques. More specifically, the proposed technologies allow for better minimization than has been available until now. Minimization is an important principle in data protection and it also reduces the risks related to storing or processing confidential data. Reduced risks will also reduce costs on other cybersecurity measures and liabilities.

11 Conclusions

This document presents deliverable "D3.2 – Cross Sectoral Cybersecurity Building Blocks" of CyberSec4Europe.

The document details the privacy-preserving architecture of CyberSec4Europe and followed by an identification of more than 20 generic and cross-sector building blocks that are already available within the project consortium. For each of the building blocks, a functional description, as well as description of the current maturity level, has been presented. Each of these building blocks is mapped into the presented architecture, and indicates the demonstrator for which it could be used, and which requirements it would satsify. As such, this deliverable provides a basis for "WP5 – Demonstration Cases", by offering developers a rich set of cybersecurity building blocks to leverage in their scenarios..

Furthermore, for each category of building blocks, this document lists research challenges that were already identified together with WP5, and that will receive further attention by WP3 in the next phases of the project, where the focus will be put depending on the priorities and feedback received from WP5. In addition, this research outlooks were also communicated to „WP4 – Research and Development Roadmap" for consideration for their research agenda.

In summary, this document provides a baseline for the development of the demonstrators, and an important project-internal link between work packages WP3, WP4, and WP5. This living document will be further revised and updated to „D3.13 – Updated version of enablers and components" in month M30 and „D3.20 – Final cybersecurity enablers and underlying technologies components" in month M36.

References

[1] D. Chaum: *Untraceable electronic mail, return addresses, and digital pseudonyms*. Commun. ACM 24: 84–88 (1981)

[2] D. Chaum: *Security without identification: Transaction systems to make big brother obsolete.* Commun.ACM 28: 1030–1044 (1985)

[3] S. Krenn, T. Lorünser, A. Salzer, C. Striecks: *Towards Attribute-Based Credentials in the Cloud.* CANS 2017: 179-202

[4] U. Haböck, S. Krenn: *Breaking and Fixing Anonymous Credentials for the Cloud*. CANS 2019: 249-269

[5] M. Blaze, G. Bleumer, M. Strauss: *Divertible Protocols and Atomic Proxy Cryptography*. EUROCRYPT 1998: 127-144

[6] R. Steinfeld, L. Bull, Y. Zheng: *Content Extraction Signatures*. ICISC 2001: 285-304

[7] M. Crosby, P. Pattanayak, S. Verma, and V. Kalyanaraman: *Blockchaintechnology: Beyond Bitcoin.* Appl. Innov., vol. 2, 6–10 (2016)

[8] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou: *Hawk:The blockchain model of cryptography and privacy-preserving smartcontracts*. IEEE Symp. Secur. Privacy (SP) 2016, 839–858

[9] I. Miers, C. Garman, M. Green, and A. D. Rubin: *Zerocoin: Anonymous distributed E-cash from bitcoin*. IEEE Symp. Secur. Privacy (SP) 2013, 397–411

[10] S. Alboaie and D. Cosovan: *Private data system enabling self-sovereign storage managed by executable choreographies*. Distributed Applications and Interoperable Systems 2017, 83–98

[11] A. Tobin and D. Reed: *The Inevitable Rise of Self-Sovereign Identity*. Available online: https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf (2016)

[12] D. Reed, M. Sprony, D. Longley, C. Allen, R. Grant, M. Sabadello: *Decentralized identifiers (DIDs) v0. 11 data model and syntaxes fordecentralized identifiers (DIDs)*. W3C, Tech. Rep., 2018.

[13] M. Sporny and D. Longley: *Verifiable claims data model and representations*. W3C, Tech. Rep., 2017.

[14] K. Wagner, B. Némethi, E. Renieris, P. Lang, E. Brunet, E. Holst: *Self-sovereign identity: A position paper on blockchain enabled identity and the road ahead.* Blockchain Bundesverband, Tech. Rep., 2018.

[15] J. B. Bernabe, J. L. Canovas, J. L. Hernandez-Ramos, R. Torres Moreno, A. Skarmeta: *Privacy-preserving solutions for Blockchain: review and challenges*. IEEE Access, Vol 7, Issue 1, 164908-164940 (2019).

[16] R. Torres Moreno, J. B. Bernabe, J. García Rodríguez, T. K. Frederiksen, M. Stausholm, N. Martínez, E. Sakkopoulos, N. Ponte, A. Skarmeta: *The OLYMPUS Architecture—Oblivious Identity Management for Private User-Friendly Services*. Sensors 20(3), 945 (2020).

[17] J. Bernal, M. David, R. Torres Moreno, J. Presa Cordero, S. Bahloul, A. F. Skarmeta: *ARIES: Evaluation of a reliable and privacy-preserving European identity management framework*, Future Generation Computer Systems, Volume 102, January 2020, Pages 409-425

[18] W. Shi, S. Dustdar: *The Promise of Edge Computing.* Computer, vol. 49, no. 5, pp. 78-81, 2016.

[19] R. Rios, R. Roman, J. A. Onieva, J. López: *From Smog to Fog: A Security Perspective*. FMEC 2017, 56-61

[20] Y. Tian, J. Yuan, H. Song: *Efficient privacy-preserving authentication framework for edge-assisted Internet of Drones*. J. Inf. Secur. Appl. 48 (2019)

[21] T. Khalid, M. Abbas Khan Abbasi, M. Zuraiz, A. N. Khan, M. Alim R. W. Ahmad, J. J. P. C. Rodrigues, M. Aslam. *A survey on privacy and access control schemes in fog computing*. Int J Commun Syst. (2019)

[22] X. Li, S. Liu, F. Wu, S. Kumari, J. J. P. C. Rodrigues: *Privacy Preserving Data Aggregation Scheme for Mobile Edge Computing Assisted IoT Applications*. IEEE Internet of Things Journal 6(3): 4755-4763 (2019)

[23] C. Xu, J. Ren, D. Zhang, Y. Zhang: *Distilling at the Edge: A Local Differential Privacy Obfuscation Framework for IoT Data Analytics*. IEEE Communications Magazine 56(8): 20-25 (2018)

[24] X. Xu, Y. Xue, L. Qi, Y. Yuan, X. Zhang, T. Umer, S. Wan: *An edge computing-enabled computation offloading method with privacy preservation for internet of connected vehicles*. Future Gener. Comput. Syst. 96: 89-100 (2019)

[25] M. De Donno, J. Manuel Donaire Felipe, N, Dragoni: *ANTIBIOTIC 2.0: A Fog-based Anti-Malware for Internet of Things*. EuroS&P Workshops 2019: 11-20

[26] M. De Donno, N. Dragoni, A. Giaretta, M. Mazzara, P. Ciancarini, S. Litvinov, A. Messina, A. Sillitti, G. Succi: *AntibIoTic: Protecting IoT Devices Against DDoS Attacks*. Conference in Software Engineering for Defence Applications, 2018

[27] S. Pearson, M. Casassa-Mont: *Sticky Policies: An Approach for Managing Privacy across Multiple Parties*. Computer, vol. 44, no. 9, pp. 60-68, 2011

[28] G. Fisk, C. Ardi, N. Pickett, J. S. Heidemann, M. Fisk, C. Papadopoulos: *Privacy Principles for Sharing Cyber Security Data*. IEEE Symposium on Security and Privacy Workshops 2015: 193-197

[29] H. Spengler, F. Prasser: *Protecting Biomedical Data Against Attribute Disclosure*. GMDS 2019: 207-214

[30] R. Bild, K. A. Kuhn, F. Prasser: *SafePub: A Truthful Data Anonymization Algorithm With Strong Privacy Guarantees*. PoPETs 2018(1): 67-87 (2018)

[31] F. Prasser, F. Kohlmayer, R. R. Lautenschläger, K. A. Kuhn: *ARX - A Comprehensive Tool for Anonymizing Biomedical Data*. AMIA 2014

[32] R. Chevrier, V. Foufi, C. Gaudet-Blavignac, A. Robert, C. Lovis. *Electronic signatures. Westlaw UK. Use and Understanding of Anonymization and De-Identification in the Biomedical Literature: Scoping Review*. J Med Internet Res 2019;21(5)

[33] F. Prasser, H. Spengler, R. Bild, J. Eicher, K.A. Kuhn. *Privacy-enhancing ETL-processes for biomedical data*. International Journal of Medical Informatics 126 (2019) 72–81

[34] Victor Costan, Srinivas Devadas. *Intel SGX Explained*. IACR Cryptology ePrint Archive 2016: 86 (2016)

[35] E. F. Brickell, J. Camenisch, L. Chen: *Direct anonymous attestation*. ACM Conference on Computer and Communications Security 2004: 132-145

[36] E. Brickell, J. Li: *Enhanced privacy ID from bilinear pairing for hardware authentication and attestation*. IJIPSI 1(1): 3-33 (2011)

[37] Dong Hwi Seo and P Sweeney. *Simple authenticated key agreement algorithm*. Electronics Letters 35, 13, 1073–1074 (1999)

[38] P. Zimmermann, A. Johnston, J. Callas: *ZRTP: Media Path Key Agreement for Unicast Secure RTP*. RFC 6189: 1-115 (2011)

[39] D. Balfanz, D. K. Smetters, P. Stewart, H. Chi Wong: *Talking to Strangers: Authentication in Ad-Hoc Wireless Networks*. NDSS 2002

[40] P. Thermo, A. Takanen: *Securing VoIP Networks: Threats, Vulnerabilities, and Countermeasures*. Addison-Wesley Professional. (2007)

[41] F. Hao, P. Ryan: *J-PAKE: Authenticated Key Exchange without PKI.* Trans. Computational Science 11: 192-206 (2010)

[42] C. Huitema, D. Kaiser. *Device Pairing Using Short Authentication Strings*. (2016).

[43] J. Lancrenon, M. Skrobot, Q. Tang: *Two More Efficient Variants of the J-PAKE Protocol*. ACNS 2016: 58-76

[44] F. Hao: *Schnorr Non-interactive Zero-Knowledge Proof*. RFC 8235: 1-13 (2017)

[45] M. Toorani: *Security analysis of J-PAKE*. ISCC 2014: 1-6

[46] N. Asokan, P. Ginzboorg: *Key agreement in ad hoc networks*. Computer Communications 23(17): 1627-1637 (2000)

[47] S. Vaudenay: *Secure Communications over Insecure Channels Based on Short Authenticated Strings*. CRYPTO 2005: 309-326

[48] I. R. Buhan, J. M. Doumen, P. H. Hartel, R. N. J. Veldhuis: *Feeling is Believing: a location limited channel based on grip pattern biometrics and cryptanalysis*. Number 06-29 in CTIT technical reports series (2006)

[49] S. E. Levy, R. M. Myers: *Advancements in next-generation sequencing*. Annual review of genomics and human genetics 17: 95-115 (2016)

[50] H. Li, R. Durbin. *Fast and accurate short read alignment with Burrows Wheeler transform*. Bioinformatics 25.14: 1754-1760. (2009)

[51] E. Ayday, J. L. Raisaro, J.-P. Hubaux, J. Rougemont: *Protecting and evaluating genomic privacy in medical tests and personalized medicine*. WPES 2013: 95-106

[52] C. Dwork, F. McSherry, K. Nissim, A. D. Smith: *Calibrating Noise to Sensitivity in Private Data Analysis*. TCC 2006: 265-284

[53] M. Dumas, L. García-Bañuelos, P. Laud: *Differential Privacy Analysis of Data Processing Workflows*. GraMSec@CSF 2016: 62-79

[54] P. Laud, M. Pettai, J. Randmets: *Sensitivity Analysis of SQL Queries*. PLAS@CCS 2018: 2-12

[55] K. Chatzikokolakis, M. E. Andrés, N. Emilio Bordenabe: *Catuscia Palamidessi: Broadening the Scope of Differential Privacy Using Metrics*. Privacy Enhancing Technologies 2013: 82-102

[56] P. Laud, A. Pankova, M. Pettai: *Achieving Differential Privacy using Methods from Calculus*. To appear at PET Symposium 2020.

[57] M. Fowler, J. Highsmith: *The Agile Manifesto*. Softw. Dev. 9(8), 28-25 (2001)

[58] H. Kniberg: *Scrum and XP from Trenches* (2015)

[59] J. Ahola, C. Frühwirth, M. Helenius, L. Kutvonen, J. Myllylahti, T. Nyberg, A. Pietikäinen, P. Pietikäinen, J. Röning, S. Ruohomaa, C. Särs, T. Siiskonen, A. Vähä-Sipilä, V. Ylimannela: *Handbook of the Secure Agile Software Development Life Cycle*. University of Oulu (2014)

[60] V. Asthana, I. Tarandach, N. O'Donoghue, B. Sullivan, M. Saario: *Practical Security Stories and Security Tasks for Agile Development Environments*. (2012)

[61] S. Wachter: *Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR*. Comput. Law Secur. Rev.34(3), 436-449 (2018)

[62] D. Brossard, G. Gebel, M. Berg. *A systematic approach to implementing abac*. ACM ABAC 2017: 53–59

[63] C. Bartolini, S. Daoudagh, G. Lenzini, E. Marchetti: *GDPR-Based User Stories in the Access Control Perspective*. QUATIC 2019: 3-17 (2019)

[64] C. Bartolini, S. Daoudagh, G. Lenzini, E. Marchetti: *Towards a Lawful Authorized Access: A Preliminary GDPR-based Authorized Access*. ICSOFT 2019: 331-338 (2019)

[65] M. Alohaly, H. Takabi, E. Blanco: *Automated Extraction of Attributes from Natural Language Attribute-Based Access Control (ABAC) Policies*. Cybersecurity 2(1), 2 (2019)

[66] X. Xiao, A. Paradkar, S. Thummalapenta, T. Xie: *Automated Extraction of Security Policies from Natural-Language Software Documents*. ACM SIGSOFT FSE 2012, pp. 12:1-12:11 (2012)

[67] K. Fatema, C. Debruyn, D. Lewis, D. O'Sullivan, J. P. Morrison, A. Mazed: *A Semi-Automated Methodology for Extracting Access Control Rules from the European Data Protection Directive*. IEEE SPW, 25-32 (2016)

[68] G. Ateniese, B. Magri, D. Venturi, E. R. Andrade: *Redactable Blockchain - or - Rewriting History in Bitcoin and Friends*. EuroS&P 2017: 111-126

[69] H. Krawczyk, T. Rabin: *Chameleon Signatures*. NDSS 2000

[70] D. Derler, K. Samelin, D. Slamanig, C. Striecks: *Fine-Grained and Controlled Rewriting in Blockchains: Chameleon-Hashing Gone Attribute-Based*. NDSS 2019

[71] J. Camenisch, D. Derler, S. Krenn, H. C. Pöhls, K. Samelin, D. Slamanig: *Chameleon-Hashes with Ephemeral Trapdoors - And Applications to Invisible Sanitizable Signatures*. Public Key Cryptography (2) 2017: 152-182

[72] A. Sahai, B. Waters: *Fuzzy Identity-Based Encryption*. EUROCRYPT 2005: 457-473

[73] "Home." *Ethereum.org*, ethereum.org/.

[74] "Hyperledger Fabric." *Hyperledger*, https://www.hyperledger.org/projects/fabric

[75] W. Li, A. Sforzin, S. Fedorov, G. O. Karame: *Towards scalable and private industrial blokchains*. ACM Workshop on Blockchain, Cryptocurrencies and Contracts,9-14, 2017

[76] A. Gervais, G. O. Karame, K. Wüst, V. Glykantzis, H. Ritzdorf, S. Capkun: *On the Security and Performance of Proof of Work Blockchains*. ACM Conference on Computer and Communications Security 2016: 3-16

[77] M. Castro, B. Liskov: *Practical byzantine fault tolerance and proactive recovery*. ACM Trans. Comput. Syst. 20(4): 398-461 (2002)

[78] S. King, S. Nadal: Ppcoin: *Peer-to-peer crypto-currency with proof-of-stake*. self-published paper (2012).

[79] D. Pike, P. Nosker, D. Boehm, D. Grisham, S. Woods, J. Marston. *PoST White Paper*. Whitepaper, available online: https://cdn. vericonomy. com/documents/VeriCoin-Proof-of-Stake-Time-Whitepaper.pdf (2015)

[80] P. Vasin. *Blackcoin's proof-of-stake protocol v2*. Available online: https://blackcoin. co/blackcoin-pos-protocol-v2-whitepaper.pdf (2014).

[81] Novacoin-Project. *Novacoin-Project/Novacoin*. *GitHub*, github.com/novacoin-project/novacoin/wiki/Proof-of-stake.

[82] Nxt.*Nxt Whitepaper*. Available online: nxtdocs.jelurida.com/Nxt_Whitepaper. (2016)

[83] F. Schuh, D. Larimer: *Bitshares 2.0: general overview*. Available online: http://docs.bitshares.org/downloads/bitshares-general.pdf (2017).

[84] Ethereum: *Ethereum/Wiki*. Available online: github.com/ethereum/wiki/wiki/Proof-of-Stake-FAQ#what-is-the-nothing-at-stake-problem-and-how-can-it-be-fixed.

[85] Ethereum Foundation: *Long-Range Attacks: The Serious Problem With Adaptive Proof of Work*. Available online: blog.ethereum.org/2014/05/15/long-range-attacks-the-serious-problem-with-adaptive-proof-of-work/. (2014)

[86] Ethereum: *Ethereum/Wiki*. Available online: github.com/ethereum/wiki/wiki/Proof-of-Stake-FAQ#how-does-validator-selection-work-and-what-is-stake-grinding.

[87] T. Lorünser, An. Happe, B. Rainer, F. Wohner, C. Striecks, D. Demirel, G. Traverso: *Advanced architecture for distributed storage in dynamic environments (SECOSTOR Tool)*. PRISMACLOUD project deliverable D5.3. 2017.

[88] A. Shamir: *How to Share a Secret*. Commun. ACM 22(11): 612-613 (1979)

[89] M. Castro, B. Liskov: *Practical byzantine fault tolerance and proactive recovery*. ACM Trans. Comput. Syst. 20(4): 398-461 (2002)

[90] D. Bogdanov. *Sharemind: programmable secure computations with practical applications*. PhD thesis, University of Tartu, 2013

[91] D. W. Archer, D. Bogdanov, L. Kamm, Y. Lindell, K. Nielsen, J. Illeborg Pagter, N. P. Smart, R. N. Wright: *From Keys to Databases – Real-World Applications of Secure Multi-Party Computation*. The Computer Journal, vol. 61, Issue 12, Pages 1749–1771 (2018).

[92] D. Bogdanov, L. Kamm, S. Laur, V. Sokk. *Rmind: a tool for cryptographically secure statistical analysis*. IEEE Transactions on Dependable and Secure Computing, Pages (99):1–14, 2016.

[93] J. Randmets: *Programming Languages for Secure Multi-party Computation Application Development*. PhD thesis, University of Tartu, 2017.

[94] D. Bogdanov, P. Laud, S. Laur, and P. Pullonen: *From input private to universally composable secure multi-party computation primitives*. CSF 2014, 184–198

[95] J. S. Resende, R. Martins, L. Antunes: *Enforcing Privacy and Security in Public Cloud Storage*. PST 2018: 1-5

[96] H. Krawczyk, M. Bellare, R. Canetti: *HMAC: Keyed-Hashing for Message Authentication*. RFC 2104: 1-11 (1997)

[97] A. N. Bessani, R. Mendes, T. Oliveira, N. Ferreira Neves, M. Correia, M. Pasin, P. Veríssimo: *SCFS: A Shared Cloud-backed File System*. USENIX Annual Technical Conference 2014: 169-180

[98] A. Bessani, M. Correia, B. Quaresma, F. André, P. Sousa: *Depsky: dependable and secure storage in a cloud-of-clouds*. ACM Transactions on Storage, 9(4):12, 2013.

[99] K. Oastler. *Tricky transparency requirements and how to overcome them*. Westlaw UK.

[100] R. Corbet. *When to report your data breach*. Westlaw UK

[101] M. Watts. *Consent vs legitimate interest: Part 1*. Westlaw UK.

[102] K. Oastler. *GDPR series: automated decisions - what controllers need to know*. Westlaw UK.

[103] S. Tsakiridi. *A practical guide to conducting data protection impact assessments*. Westlaw UK.

[104] B. Treacy. *Data protection by design and by default: from theory to practice*. Westlaw UK.

[105] S. Mason. *Electronic signatures*. Westlaw UK.