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

Project start: 1 February 2019
Project duration: 42 months

Cyber Security for Europe

D3.13

Updated version of enablers and components

Document Identification	
Due date	31 October 2021
Submission date	31 October 2021
Revision	3

Related WP	WP3	Dissemination Level	Public
Lead Participant	C3P	Lead Author	João Resende
Contributing Beneficiaries	C3P, UMU, UMA, CNR, AIT, UNILU, CNR, DTU, UM, VTT, ATOS	Related Deliverables	D3.2, D3.11

Abstract: This document presents D3.13 - "Updated Version of Enablers and Components". It is a supplement to D3.2 – "Cross Sectoral Cybersecurity Building Blocks": while D3.2 gives an overview of all CyberSec4Europe's enablers (also known as "assets" within the project), D3.13 focuses on making a demonstrator to show these assets working in a joint scenario and it is an updated version of the assets.

The demonstrator is based on a conceptual smart-campus platform, and all the assets are presented as a way of improving the quality and security of a smart campus. It consists of three scenarios for demonstrators:

(i) Video surveillances use case to manage the feeds, as it is a common scenario in smart buildings that allows a response team to continuously monitor the campus, such as monitoring fires or accidents.
(ii) Services with different access controls - Identity management for user authentication in various heterogeneous smart-campus services, including services for direct interaction with the University (e.g., enrollment in courses or activities). For this, it is necessary to have the guarantee of different access controls for different services. For example, some specific academic information should be restricted to users who are university students or enrolled in a particular course or year.
(iii) A smart-campus geolocation service is used to detect trends in the movement of people. This, in turn, can help with public transport planning, urban planning, creating safer and friendlier areas for residents.

This document includes the description of the use case scenario, and each asset describes an overview, the research challenges addressed and the demonstration example.

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 document presents D3.13 – “Updated version of enablers and components”, a product of the joint work of the members of task T3.2 – “Research and Integration on Cybersecurity Enablers and underlying Technologies”. As anticipated by its title, this deliverable focuses on privacy tools and the impact that they can have in a smart-campus infrastructure beginning with a short description of the general architecture, then the Scenarios of the smart-campus are introduced and then for each privacy asset we describe in detail the main contribution of each of them individually.

Document information

Contributors

Name	Partner
João Resende	C3P
Patricia Sousa	C3P
Rolando Martins	C3P
Inês Martins	C3P
Luís Antunes	C3P
Jorge Bernal	UMU
Rafa Torres	UMU
Jesus Garcia-Rodriguez	UMU
Ruben Rios	UMA
Rodrigo Roman	UMA
Javier Lopez	UMA
Eda Marchetti	CNR
Said Daoudagh	CNR
Christos Xenakis	UPRC
Eleni Veroni	UPRC
Dan Bogdanov	CYBER

Liina Kamm	CYBER
Armin Kisand	CYBER
Niko Lehto	VTT
Outi-Marja Latvala	VTT
Alireza Esfahani	UNILU
Alberto Lluch Lafuente	DTU
Stephan Krenn	AIT
Thomas Lorünser	AIT
Claudio Soriente	NEC
Alessandro Sforzin	NEC
Juan Carlos Pérez Baún	ATOS
Miryam Villegas Jimenez	ATOS
Marko Kompara	UM
Tamara Bubnjar	UM

Reviewers

Name	**Partner**
Sunil Chaudhary	NTNU
Liliana Pasquale	UCD

History

Version	Date	Authors	Comment
0.01	2019-03-28	João Resende	1st Draft
0.02	2021-01-07	Jorge Bernal Bernabe	Smart-Campus description
0.03	2021-02-21	CNR	Initial GENERAL_D description
0.04	2021-02-25	CNR	Specification of Test case for GENERAL_D validation
0.05	2021-03-10	CYBER	Added
0.06	2021-03-24	VTT	Added
0.07	2021-04-07	C3P	Update the description and structure
0.08	2021-04-28	AIT, NEC, CNR, DTU, Atos...	Updated descriptions of various assets
0.09	2021-05-20	C3P	Update structure
0.10	2021-09-01	CNR	Revision of section 4.12 (GENERAL_D)
2	2021-09-15	C3P	Revise the document
3	2021-09-15	C3P	Revise the document
3	2021-10-30	GUF	Final check & Submission

Table of Contents

List of Figures

List of Acronyms

2	**2FA**	Two-Factor Authentication
A	**ABC**	Attribute-based credential
	AES	Advanced Encryption Standard
C	**CA**	Certification authority
	CCTV	Closed-circuit television
D	**DLT**	Distributed Ledger Technology
	DPIA	Data Protection Impact Assessment
	DPO	Data Protection officer
	DHE	Diffie-Hellman Ephemeral
E	**eID**	Electronic identity
	ECDSA	Elliptic Curve Digital Signature Algorithm
	ECIES	Elliptic Curve Integrated
	ECDHE	Elliptic-curve Diffie–Hellman
G	**GDPR**	General Data Protection Regulation
	GPS	Global Positioning System
I	**IdM**	Identity Manager
	IdP	Identity Provider
	IoT	Internet of Things

N	**NGS**	Next Generation Sequencing
P	**p-ABC**	Privacy-preserving Attribute-Based Credential
	PET	Privacy Enhancing Technologies
S	**SAML**	Security Assertion Markup Language
T	**TEE**	Trust Execution Environment
U	**UAF**	Universal Authentication Framework
V	**vIdP**	Virtual Identity Provider
X	**XACML**	eXtensible Access Control Markup Language

1. Introduction

This document presents an updated version of the assets as an extension of D3.11 (“Definition of Privacy by Design and Privacy Preserving Enablers”)

This document presents a common scenario based on “Smart University Campus”, where different CS4E partners (mainly in T3.2) can engage and integrate their assets for evaluation and demonstration. This scenario will describe the storyline, processes, and test-cases employed to validate and evaluate the feasibility, novelty, soundness, accuracy, and performance of the assets devised and implemented in task T3.2.

Unlike in WP5, the focus of the T3.2 demonstrator should be validation and evaluation of the research aspects of the T3.2 investigations and associated assets, rather than an evaluation of mature or fully integrated software-implementation with the final users (as it is the focus of WP5 Pilots).

This scenario is broad enough to embrace different assets such as privacy-preserving Identity Manager (IdM) also involving IoT scenarios, TEE, and blockchain systems (main research topics of T3.2). The integration of the implementations of all different T3.2 assets within the Smart Campus testbed is not required since many assets are deployed in WP5, this is a demonstration on how the tools can work in this context. It will depend on the specific assets and effort per partner assigned in the task.

This common scenario for demonstration will be adopted in D3.13 and is expected to be used also as a baseline for the final T3.2 demonstrator, including the latest versions and research outcomes from partners.

1.1. Our Approach

To develop this deliverable, we focused on finding a scenario where most assets might be of use from a security and privacy perspective. The partners involved chose a smart-campus scenario, as it encompasses several users, services, and devices communicating with each other. Smart-campus aggregates much private information about students, professors, and other staff. Especially for the students, sensitive data is being transferred and stored between many services. Thus, it is necessary to guarantee that data is not tampered with and arrives as supposed. For this, it is necessary to have adequate identity management and access controls and guarantee the integrity, privacy, and security of systems and data. Also, it is necessary to ensure that the devices are secured.

1.2. Document Structure

This document is a follow-up of D3.11 (“Definition of Privacy by Design and Privacy-Preserving Enablers”), with an updated version of the assets. Thus, it is included a use case scenario to make the demonstration of each asset.

The document is structured as follows:

- Section 2 describes the T3.2 general architecture.
- Section 3 gives an overview of the smart-campus scenario description, including the three sub-scenarios for demonstrations:
    - CCTV,
    - Identity Management, and

- Geolocation service.

- Section 4 has a description of all assets. For each asset, the discussion is structured as follows:
  - An introduction to the asset and its modus operandi,
  - The research challenges addressed,
  - The demonstrations example.

Section 5 concludes the document.

2. T3.2 Architecture

The CyberSec4Europe Privacy-Preserving Architecture consists of several building blocks that expand over several intertwined domains, including the user domain, the web domain, and the IoT domain, as shown in Figure 1. A detailed description can be found in D3.2 “Cross-Sectoral Cybersecurity Building Blocks.” We also map all the assets used in this deliverable with the correspondent block of the framework. Some assets of the D3.11 deliverable were replaced by new privacy tools or rebranded under a new name and this information is also available. The total number of assets in this task are 19, divided by “Services Plane”, “User domain”, “Administration Plane”, “Intelligence Plane”, “Control and management plane”, “Blockchain Plane”, “IoT domain” and “Web domain”. Table 1 categorizes the assets accordingly with the plane, it also gives information regarding the difference regarding the previous deliverable and collaboration with WP5.

The building blocks are defined for different purposes which range from 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.

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 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, the 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 the 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 a 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)1 standards and specifications, the architecture features the building blocks needed for the creation, resolution, and discovery of decentralized identifiers (DID identifiers2) and names in heterogeneous blockchains through resolvers. In addition, the Identity Hubs keep secure, encrypted, privacy-preserving personal data storage and computation of data. Where the resolver services link user's DID's employed in blockchain with Identity Hubs. The blockchain Identity services provide the 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 a 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 **Error! Reference source not found.** can be also instantiated through processes related to the Web domain (e.g., eCommerce) in the CyberSec4Europe privacy-preserving architecture.

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

In this case, the Web domain is comprised of a 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 a 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.

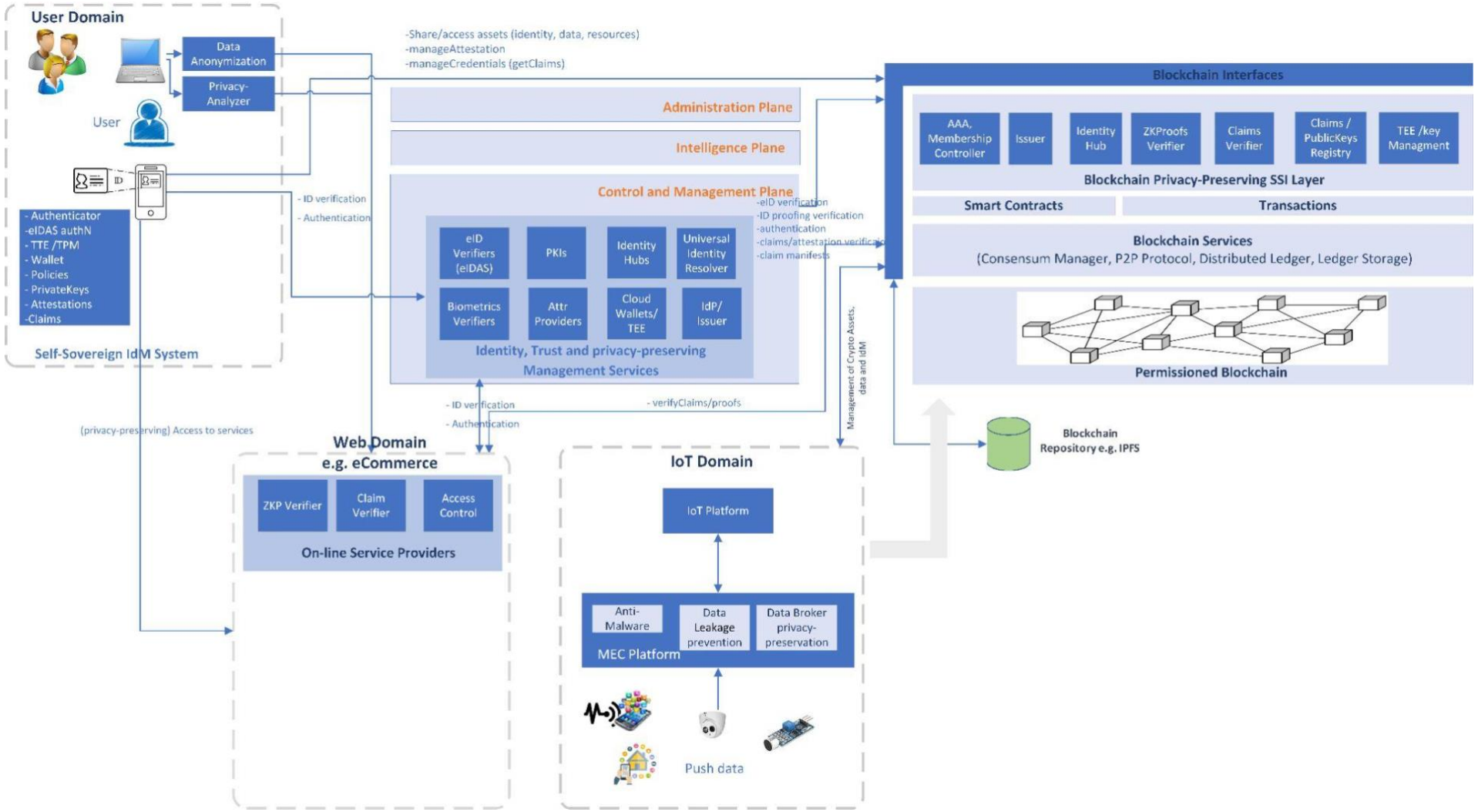


Figure 1: CyberSec4Europe Privacy-Preserving Functional Architecture

Asset	Partner	Services Plane	User domain	Administration Plane	Intelligence Plane	Control and management plane	Blockchain Plane	IoT domain	Web domain	Covered
Issuer-Hiding Attribute-Based Credentials	AIT		x			x				In this deliverable and in WP5
Cloud-Based Anonymous Credential Systems (eABCs)	AIT		x			x				In this deliverable and in WP5
FlexProd	AIT					x				In this deliverable
ArchiStar	AIT		x							In this deliverable (as submodule of FlexProd)
SS-PP-IdM	UMU	x	x							In this deliverable and merged with Mobile p-ABC and eIDAS browser
Password-less authentication	UPRC		x			x			x	In this deliverable
Edge-Privacy	UMA							x		In this deliverable
Privacy-Aware Aggregate Programming	DTU			x						In this deliverable, replacing the effort with AntibIoTic
DANS	ATOS		x			x				In this deliverable and evaluated in WP5 – including the previous affords of SPeIDI
Cryptovault	VTT		x							In this deliverable

Elastic Deployment of TEE-based applications in the cloud	NEC		x							In this deliverable
Backdoor-resistant TEEs	NEC							x		In this deliverable
Blockchain Platform	NEC						x			In this deliverable and also in WP5.
Sharemind	CYBER		x							In this deliverable, including PLEAK differential privacy analyzers
Privacy-Preserving for Genomic data	UNILU					x				In this deliverable with a specific genomic use case
GENERAL_D	CNR		x			x		x	x	In this deliverable and also in WP5
PTASC	C3P							x		In this deliverable and also in WP5
ARGUS	C3P					x				In this deliverable and also in WP5
GDPR compliant user experience	UM			x	x					In this deliverable and also in WP5.
Interoperability and cross-border compliance	UM			x						In this deliverable.

Table 1: Mapping of assets available and the CyberSec4Europe Privacy-Preserving Functional Architecture

3. Description of the Smart Campus Scenario

The Internet of Things (IoT) allows everyday objects (equipped with computational and communication capabilities) to connect to the Internet. The "things" can exchange data with each other and with the Internet, making decisions automatically, even without human interaction. IoT applications demand platforms that aim to facilitate their development process, which may involve integrating a diversity of heterogeneous devices with varying capacities, means of data transmission, and different communication protocols. The literature presents several middleware platforms that serve as the underlying infrastructure for the development of IoT applications [mineraud2016gap, ngu2016iot].

A smart campus uses technology solutions to manage services and users' life experiences. Sensors, networks, and applications are used to collect relevant data, such as the number of rooms used, energy use, and air quality. This data can be used to improve the smart campus services [sari2017study]. An example that improves the smart campus is the management of identity of users when interacting with the smart campus where heterogeneous services will be available to potential users, with services for direct interaction with the University (e.g., enrolment in courses or activities). These services may have widely varying requirements for their usage. A public transport service may be available for any user that interacts with the platform, while some specific academic information should be restricted to users that are students from the university or even enrolled in a specific degree/course.

Cloud-based IoT applications receive, analyze, and manage data in real-time to help institutions in the smart-campus, businesses, and citizens make decisions that improve the quality of life. Students engage with smart campus ecosystems in a variety of ways, using smartphones and mobile devices, as well as connected transportation and homes. Pairing devices and data with the infrastructure and physical services can reduce costs and improve sustainability.

The smart campus allows a set of known interactions Services such as (An Overview is represented in Figure 2:

- Public transport
- Parking availability
- Campus information
- Academic information
- Mobility data analysis (privacy-preserving)

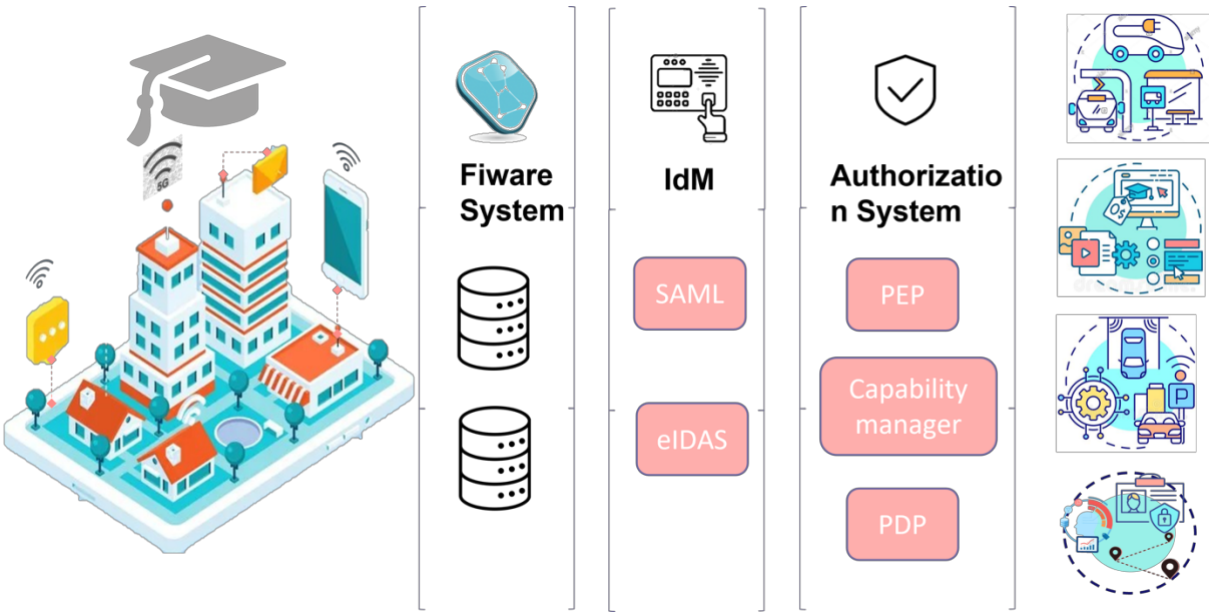


Figure 2: Smart Campus

A smart campus is composed of multiple infrastructure and components, such as:

- Smart-Building
- 5G - Base stations, WIFI Access Points (Edge Nodes)
- Smart-Campus IoT platform:
  - Fiware System (Orion Context Broker). Data Base that stores IoT data obtained from sensors (devices, users and services) in the smart campus.
  - IdM (Fiware Keyrock). Identity management system used for basic identity management and bridge to eIDAS (i.e., handles SAML communication flow with eIDAS node to obtain certified attributes).
  - Authorization System based on capability access control tokens. It includes (Policy Enforcement Point –PEP--, Capability Manager and Policy Decision Point --PDP)
    - **PEP:** Controls access to the services, checking that the request includes a valid capability token (i.e., the request is authorized).
    - **Capability Manager:** Generates capability tokens that bestow authorization to use specific services. Relies on the PDP for the decision (using XACML).
    - **PDP:** Checks if an authorization request should be conceded
  - User Mobile app: allow users to access Smart-campus services.

Given the multiple domains addressed in the smart campus and according to the partners' feedback we decided to set up a smart-campus high-level scenario from which we derive multiple scenarios that can be created use cases. In this context, we will focus the next subsection on 3 main sub-scenarios from the general smart-campus scenario: CCTV Surveillance in the Smart-Campus; Identity management and service usage in Smart Campus; Geolocation Service in the Smart Campus. The option for these scenarios focuses on the challenges addressed by the assets and available on the UMU smart campus. Table 2 identifies the main partners/assets integrated into each scenario. The only partner who does not have assets directly integrated into the smart campus scenarios is UNILU since the asset focuses on privacy-preserving of Genomic data, which is unrelated to the smart campus. However, we still show this asset in Table 2.

Use case	Partners	Assets
CCTV Surveillance in the Smart-Campus	C3P, DTU, CNR, UPRC, UM	ARGUS, PTASC, Password-less authentication, General_D, Interoperability and cross-border compliance, Password-less authentication
Identity management and service usage in Smart Campus	UMU, AIT, UM, VTT, NEC, UPRC, CNR	Blockchain Platform, SS-PP-IdM, Cloud-Based Credentials, FlexProd, ArchiStar, Issuer-Hiding Anonymous Credentials, Cloud-Based Credentials, GDPR compliant user experience, Password-less authentication, Cryptovault, GENERAL_D
Geolocation Service in Smart Campus	CYBER, UMA, ATOS, NEC	Elastic-TEE, Sharemind, Edge-Privacy, Back-door resistant TEE
Genomic data	UNILU	Privacy-Preserving for Genomic data

Table 2: Assets/partners mapped by use case

3.1. CCTV Surveillance in the Smart-Campus

This section presents an application use case as a motivation for the demonstrator. CCTV is a common scenario in smart buildings, which allows an incident CCTVs allow a response team to continuously monitor the campus for accidents, fires, and parking spots, to name a few. In case of emergency, police may be dispatched to the incident. However, the police must always have access to the video surveillance to gather information about incidents/unwanted situations, such as:

- A robbery is happening, and the robber is moving in a specific direction.
- A child is missing and the police need to analyze the CCTV feeds for possible locations of the child.
- A lady lost a bag, and the police officer needs to identify a possible suspect.

With the GDPR, sharing this type of data becomes even harder given the limitations introduced by the regulations with the protection of users. However, depending on the authorization level, users can have different qualities of the CCTV recordings. A normal user can only be authorized to track traffic lights or traffic jams but not extra information such as person faces, clothes, license plates, or any other information that can re-identify a person. However, this can be done using ML algorithms and face defacement techniques. Then, suppose it is an operator (or police, or other members with special authorization level) from the university tracking the same information, in that case, they should be able to see the information with more quality, to allow tracking cars from the municipalities. However, in this case, the faces must always be anonymized[3]. In case of emergency, we have policies that enable to track in real-time a specific user or license plate of the car but the police may not have sufficient permissions to access the specific CCTV feed and the request for higher permissions can be slow and inhibit the viewing of crucial images. To solve this latency, we propose using a system that uses a break the glass, where the police will have access to the required information immediately. Break the glass is

[3] https://www.theverge.com/2020/6/11/21280293/anonymize-blur-faces-photos-videos-camera-app-ios

generally used to do something in an emergency, especially in a medical or fire context, and refers to a quick means for a person who does not have access privileges to certain information to gain access when necessary. In this case, a pop-up will be shown that notifies the user (either a police officer or a user with specific access to the information), informing that their action will be registered on the system logs and, later on, it will be verified by a police chief department[4] to validate the reason for accessing that specific information. This will allow a faster response in cases of emergency and allow users to access information available from the CCTV feeds, which otherwise will only be accessed by an operator in a control room from the university. Also, the operators from the university will be capable of accessing the CCTV feeds on their smartphones depending on the policies of the institution.

3.1.1. Architecture

Our architecture includes users, a video surveillance feed, an authentication and access control management, and an offline verifier. The general architecture describes a scenario where users use an application to access surveillance videos. The users must have permission to access the videos, and this decision is made by the authentication management systems. The authentication depends on the type of authorization the person has, that is, users can have access to the videos in high quality, without quality, or with the faces blurred, creating the possibility of having different granularity levels. Still, there is also a registration mechanism for emergency access called Break the Glass. However, as it is an emergency case, it must be registered to be identified by an offline verifier if it is improper accessed. This register can be done using an immutable database, such as blockchain, in order to prevent the manipulation of data.

From the challenges from D3.11, we focus mainly in:

- DP-06 by providing mechanisms for control how the information is disseminated and control the private data on the communication.
- IDP-05 by providing a mechanism for guaranteeing proper identity management of things for authentication end-to-end.
- DP-07 by providing anonymization mechanism for control on how the information is stored and control the private data on the communication.
- DP-05 “Lack of mechanisms for controlling and limiting access to the data collected from numerous and geographically disperse IoT device
- Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP) (IDP-04).

Figure 3 represents the architecture with the identification of the steps to implement the CCTV scenario.

[4] This person in qualified to judge any misleading access to the information of the CCTV scenario.

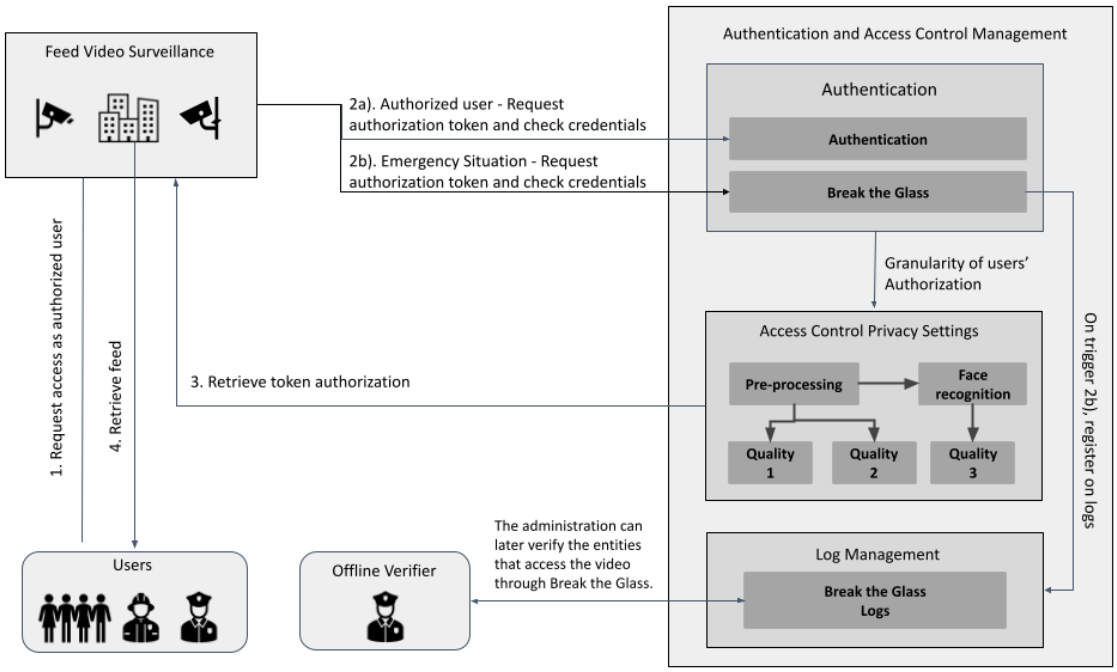


Figure 3: CCTV Scenario Overview

Video Surveillance Feed

The video surveillance feed is composed of a live stream of surveillance cameras, authenticated by login with username and password. The feeds are available depending on the logged users' authorization. For different users, it provides different video qualities and filters to comply with the data minimization principle.

This interface has an administration system that is accessible and protected by a username and password and which should only be accessed by the person in charge of the entities for managing video surveillance and security. In this interface, the admin can create or eliminate users and groups and set policies for them that define the different access controls for each entity.

Within the administration panel, it is possible to access records and logs (of those who accessed the system) for control and auditing purposes of the entity's internal platform, set the cameras that users and groups can access (e.g., by sectors), enable/disable filters for specific groups and users, remove non-accurate data and limit/define the period of conservation/retention of videos according to the GDPR.

Storage

Videos are stored for a limited time (defined by the DPO in the administration panel), but users who access these videos do not need to access all the information. Following the principle of data minimization, information should be reduced according to purpose and limited to what is strictly necessary. Thus, personal data must be hidden (for example, use filters to hide faces and license plates) to minimize exposure of personal data.

By default, the system automatically provides privacy filters and low-quality videos for users with general permissions.

The responsibility of the processing of information within the conditions defined in Article 9 from GDPR is from the administrator.

This storage is encrypted so that there is no access without permissions. Data access permissions are defined by the competent person (for example, DPO). Also, there are data retention policies because the lack of discipline around data lifecycle management means that organizations often hold onto data that add little value but lots of risks.

Authentication and Access Control

There are two types of authentications: Users and Groups. A group can have multiple users with the same policies. Individually, each user can have specific policies.

There are different authorization levels for different live feed streams. The first level presents the feed in high quality - the raw as the video camera produces it. The second level provides the image in low quality so that the image in high quality is not noticeable, for scenarios when it is not necessary to see all the details of the information. Lastly, the strictest level has privacy filters with face detection and concealment and detection of other types of personal data, such as license plates. For this type of personal data, it must be applied the appropriate filters, with the same type of policies also being applied to video storage. Our system only includes face detection, but the system is modular to include other filters.

Break the Glass

This system includes a recording mechanism for emergency access, called glass breakage. Breaking glass is often used to do something in an emergency, especially in a medical or fire context, and refers to a quick way for a person who does not have access privileges to information to gain access when needed.

This access breaks a security and privacy mechanism, especially concerning GDPR. The action of breaking the access control protection is a high-risk action, so it must be registered and identified and cannot be tampered with so that it can be audited later.

When there is an access attempt the user must agree to record that intention and write a description to define the purpose of the access. Thus, it is possible to define the user's purpose limitation to the data.

An offline verifier - who must be registered and authenticated - must be promptly notified and verify the intent data. Furthermore, it can revoke user access (in real-time) to the data.

The access token given to the user in "break the glass" mode has a limited time and can be revoked during its use. Time is short (default 5 min) and can be adjusted by the responsible person. To access it again, users must accept the conditions, and the system records the identity and the timestamp.

The following sections describe the specific enablers based on Figure 1, which represents the CyberSec4Europe Privacy-Preserving Functional Architecture [D3.2] and address the challenges enumerated previously.

3.1.2. Administration Plane

Techniques for providing password-less authentication often use some form of biometric data for the purpose of identifying users (primarily because of its convenience). The use of biometric data is especially popular with mobile devices – e.g., smartphones as proposed in the Password-less Authentication demonstrator. Biometric data, which is considered as a special category of personal data in the GDPR, when it is used for the purpose of uniquely identifying a natural person. Processing of such data is typically prohibited unless one of the exceptions (e.g., explicit consent) defined in the second paragraph of the GDPR's Article 9 applies. Member States can also introduce their own conditions and/or limitations. Before implementing any such authentication solutions, it is therefore good to know whether local, national legislations extend the GDPR and when the use of such biometric data is permitted (this tool can be used generically for other types of data).

3.1.3. Control and Management Plane

The storage of data can be maintained on the public Cloud so that we can maintain the costs of this type of system very low.

The logs and video storage collected in this context must be kept privately and securely stored on a system that ensures privacy and data loss protection.

This can be achieved using ARGUS (Section 4.2), as a privacy brokerage system to enhance trust in public cloud storage systems, guarantee data confidentiality, and improve the information's availability, a comparative study is provided by João et al.[argusprivacy]. The focus of ARGUS in this scenario is to store the relevant information using a decentralized approach, where the public cloud providers cannot access to the information stored, this is accomplished by using erasure coding techniques. Only ARGUS is capable of reconstructing all the bits of information to produce the original stream of video or logs. This also protects the privacy of the information since all the information is maintained in a secure public cloud.

It is important to ensure the integrity of data because it is important not to tamper with the logs, modify the original identification of the person that accesses through the break the glass mechanism to the video surveillance. Thus, it is essential to have a tamper-resistant and time-stamped database.

Moreover, the security of the authentication and authorization processes is crucial in this scenario. It is known that passwords are targets of multiple attacks, as they can be leaked, key-logged, replayed, eavesdropped on, brute-force decoded and phished. Therefore, the password-less authentication asset will be integrated with the Identity and Access Management system to enhance the authentication process. With the deployment of password-less authentication, the following aspects will be accomplished:

- Deployment of advanced authentication methods that combine strong cryptographic functions (e.g., FIDO protocol) with something the user knows (e.g., PIN), something the user has (e.g., USB key), or something the user is (e.g., biometrics).
- Utilizing a two factor2-Factor authentication (2FA) mechanism. Depending on the authorization that each user has, one of the authentication methods described above will be implemented.

3.1.4. User Domain

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 compromise the security of their accounts and the privacy of their data. For example, users 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 use a password manager. Furthermore, the password overload problem significantly affects the usability of an application. Password-less authentication not only will increase the security, but also will provide a more user-friendly authentication process. With its implementation the users will not have to use complex passwords to login into the application, instead, they will deploy the authentication method they prefer from a variety of available options (e.g., USB keys, biometrics, PIN, etc.).

In this process, particular attention will be devoted to developing a user-friendly mechanism for the management of personal data requirements (such as consent and purpose), and to automatically deriving GDPR-based access control policies useful for enforcing the user’s preferences.

3.1.5. IoT domain

The Edge is considered a security and privacy enabler in this architecture, 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 PTASC an end-to-end identity provider (Section 4.1). This privacy component is a new mechanism for authentication using YubiKey that allows extra control over the devices and has already been deployed in the smart-cities context.

In this case, the YubiKey will be used in the CCTV cameras' main categories and will allow two main privacy protection in the system: end-to-end communication with the central server; and privacy guarantees on the network to ensure that attackers do not try to connect to the system or connect to unauthorized remote servers.

Additionally, a GDPR-based Access Manager can manage data access in compliance with the GDPR and the data subject's consent.

3.2. Identity management and service usage in Smart Campus

This is another application case that justifies the selected demonstrator. In a Smart Campus, multiple heterogeneous services will be available to potential users, with services for direct interaction with the University (e.g., enrolment in courses or activities). These services may have widely varying requirements for their usage. A public transport service may be available for any user that interacts with the platform, while some specific academic information should be restricted to users that are students from the university or even enrolled in a specific degree/course.

Thus, the Smart Campus infrastructure needs to provide the means to perform authentication and authorization. Privacy concerns and related regulations like GDPR (which points to minimal disclosure, consent, etc.) must be taken into account. Indeed, users need to be empowered in their control over their identity, and over how much identifying information is shared in their interactions with the platform. In this context, we propose a capability-based authorization framework for the Smart Campus platform. To earn authorizations, users must authenticate using some properties of their identity, e.g., revealing an attribute or showing that it fulfils a condition (like being born in a specific period).

Within CyberSec4Europe, multiple related approaches to this challenge have been pursued, partially based on assets that were already developed before the start of the project, and also inspired by different requirements from the demonstration cases defined in WP5. Based on the concrete requirements, any of the assets below may be used.

- SS-PP-IdM ensures privacy by design in authentication processes, leveraging p-ABC technologies. The usage of eIDAS for attribute population and DLTs for auditability and distributing public parameters gives strong trust assurances to the solution.
- While not focusing on auditability, eABCs are rather focusing on the scenario of highly resource-constrained devices such as smart cards on the end user side. eABCs allow one to outsource a large fraction of the computations to a largely untrusted cloud-provider without impacting the end-user's privacy.
- Finally, issuer-hiding ABCs consider the case where the precise issuer of a certificate may be revealed. This might be interesting in the case of exchange students, where the issuer of the certificate might already uniquely identify the student within a university campus.

During student enrolment, the university gets to collect and process an extensive collection of personal data of enrolled individuals. In the scenario of Smart Campus, Data Protection Impact Assessment (DPIA) would be required on the basis that the processing involves the use of new technologies, processing affects a large number of data subjects, can include monitoring of publicly accessible areas on a large scale, and/or systematic monitoring. After the assessment shows that the use of personal data does not cause a high risk to the rights and freedoms of data owners (i.e., students), the University/Smart Campus can start using this data to provide their services. To help perform the assessment, we leverage the DPIA template we have designed in CyberSec4Europe.

Once students enroll in the university, the user's identity is augmented with extra information, like a university-specific identifier, an e-mail or other second factor authenticator. With the capabilities of our ABC-based assets, users can obtain credentials that attest with strong trust their personal and student information, while ensuring that their privacy is kept when interacting with services. A typical example showing the advantages of this kind of system over existing approaches is the obtaining of student discounts in different services. Showing a student identification card (in "physical" scenarios, like buying a movie ticket) or sending university registration information or documents (requested in streaming services like Spotify) reveal more information than necessary, like full name, date of birth, etc. Instead, with a couple of clicks, users could leverage our privacy-preserving identity management solutions, so they learn, and control exactly which information is being shared. This would enable privacy principles like minimal disclosure (e.g., just proving you are a university student, but not revealing any extra information) or informed consent (i.e., the cryptographic mechanisms ensure that only the agreed data is revealed). The whole framework of the IdM ecosystem and the soundness of the involved proofs would lead to strong trust by service providers even in this privacy-preserving environment.

3.2.1. Architecture

From the general architecture, this scenario focuses on most of the identity management related components, as well as the service usage and the required authentication (highlighted in Figure 3). A more detailed vision of the internal architecture and interactions between the components can be seen in Figure 15. There, specifically, the SS-PP-IDM is shown to rely on the blockchain instance and a distributed identity provider, which issues p-ABC credentials. Users then take advantage of those credentials to carry out the authentication processes through zero-knowledge proofs. The high-level architecture for the other two assets is similar yet does not include the distribute-ledger components.

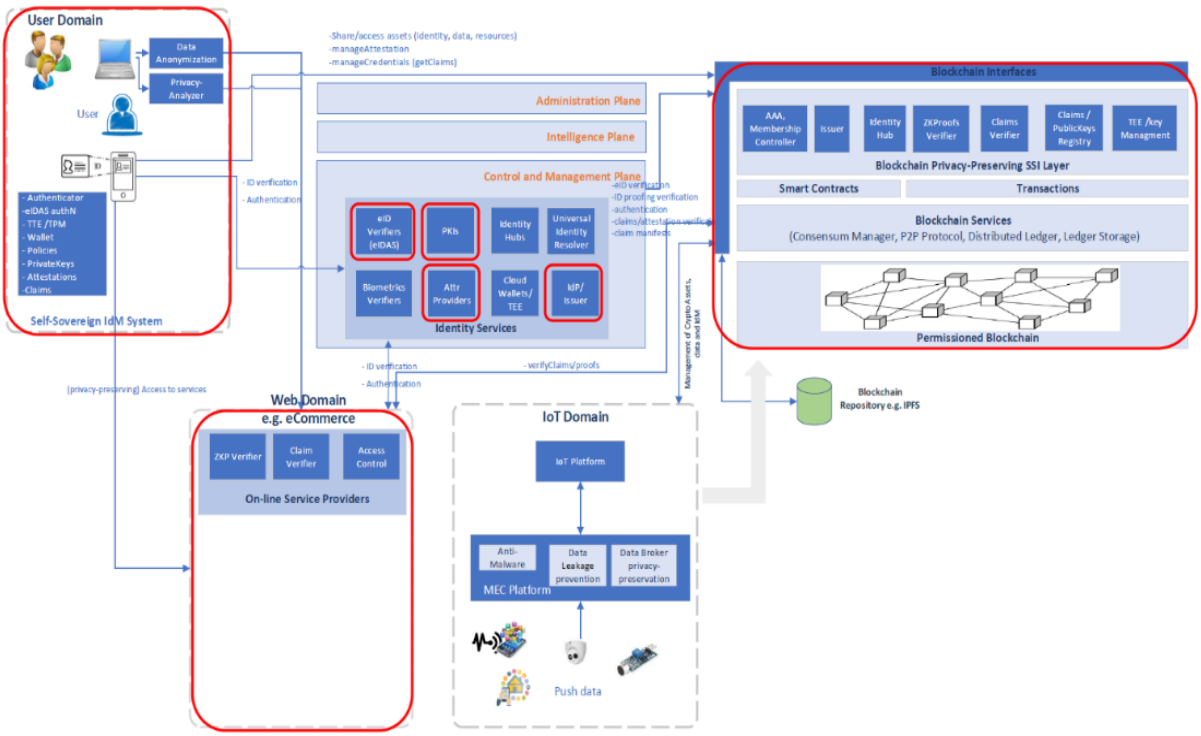


Figure 4: IdM scenario architecture overview and instantiation

From the challenges described in D3.11 this scenario addresses mainly the following:

- *IDP-02: Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks*.

- *IDP-03: User's privacy-preservation of transactions in distributed and immutable systems (e.g., blockchains).*
- *IDP-04: Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP).*
- DP-08 *When uploading information to the cloud the user partially loses control over the data.*

3.2.2. Intelligence Plane

Before a service using personal data can be designed and implemented, it is important to comply with relevant regulations. Recent security and privacy relevant examples include eIDAS[5], GDPR and the upcoming ePrivacy[6]. One important step prescribed in the GDPR is the Data Protection Impact Assessment (DPIA) which should be done any time processing of personal data is likely to result in a high risk to the rights and freedoms of natural persons. The assessment serves as a University's/Campus' evaluation of compliance to the GDPR requirements and assessment of the impact the risks of the envisaged processing operations will have on the protection of personal data used in providing Smart Campus services.

3.2.3. Control and Management Plane

The Identity Provider must ensure that users are actually in possession of the attributes that they claim form their identity. To this end, it must collaborate/rely on external trustworthy attribute providers (which may be supported by a PKI).

In this sense, eIDAS nodes are a great source of reputable identity information. Other attribute providers may be useful in this scenario. For example, the Smart Campus platform itself may provide an identifier for the user, while the university framework can provide other attributes like the user being enrolled in activities or courses.

The main function of the Identity Provider is issuing credentials to users so they can be used for authentication. In that task, it is supported by the DLT infrastructure for distributing the public parameters (e.g., verification keys) to all the entities that need them. However, it must assure that third-party elements cannot exploit the issuance process to obtain falsified/stolen credentials.

The Identity Provider will also need to fulfil all the account management operations contemplated in regulations like GDPR and in most use cases. Namely, it must allow the user to manage the attributes associated to his/her account (add/remove) and delete the account altogether, among others (managing metadata like password if necessary…). Lastly, depending on the use case it may be necessary for the management plane to have the means to suspend or remove accounts that misuse their credentials.

3.2.4. User Domain

Users access to and interact with services using the available interfaces (e.g., a mobile application). To do that, they need to comply with the access requirements and policies. Identity management tools are necessary for the involved verifications, so they must be included in user applications. While most details about the underlying identity management solutions do not need to be apparent to users (e.g.,

[5] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2014.257.01.0073.01.ENG
[6] https://digital-strategy.ec.europa.eu/en/policies/eprivacy-regulation

how complex is the distributed identity provider, how the underlying cryptography works…), users will have some direct interactions related to identity management:

- During enrolment phases, users may need to do actions specific to the attribute providers. For example, using their eID for identification in eIDAS.
- For the issuance process, users have to identify against the identity provider, e.g., through traditional username and password and optionally two-factor authentication.
- For service usage, users will be prompted to consent to access policies, revealing information about their identity.

3.2.5. Blockchain Interfaces

When using any blockchain technology, the security of the private key is paramount. Typically, one would use a wallet application to interact with the blockchain: wallets are used to store the private keys, generate signatures, and encode the transactions on behalf of the user. Online wallets require a lot of trust in a third-party service provider, while mobile and desktop wallets place the onus on the user to choose how to keep their keys safe, balancing delicately between convenience and security. Hardware wallets are dedicated devices that host the keys. They are very secure the user is fully in control of the keys, but on the other hand they are not as convenient as the other wallet types and losing or breaking the device means that the keys are forever gone.

The technical know-how of the user is important to consider when choosing a wallet type; in the case of the Smart Campus, IT administrators can be considered technically savvy, and we can assume that in a professional context they need to prioritize security and control of the keys. Therefore, a hardware wallet would be a likely choice. The security of the private key is not only dependent on the secrecy (or confidentiality) of the key, but also on its availability and integrity: what if the key is somehow lost? It is important to have a reliable and secure backup mechanism. CryptoVault is used to demonstrate a secure hardware wallet and backup method for the private keys of blockchain applications.

3.3. Geolocation Service in the Smart Campus

As in smart cities, geolocation in the smart campus can be used to detect trends in people's movement. This in turn can help with planning public transport, urban planning, creating safer and more resident-friendly areas. However, geolocation data has a lot of sensitive information that can be used adversely. In this scenario, we look at outdoor geolocation based on GPS or cell tower signaling, however, the assets we demonstrate are independent of the kinds of data collected and can be used for other data analysis as well.

In this scenario, we provide a means for analyzing the data and detecting patterns in a privacy-preserving way, so that the benefits of the data can be made use of without seeing any individual's records.

3.3.1. Architecture

From an architectural point of view, this scenario focuses on the User and IoT domains, as depicted in Figure 4.

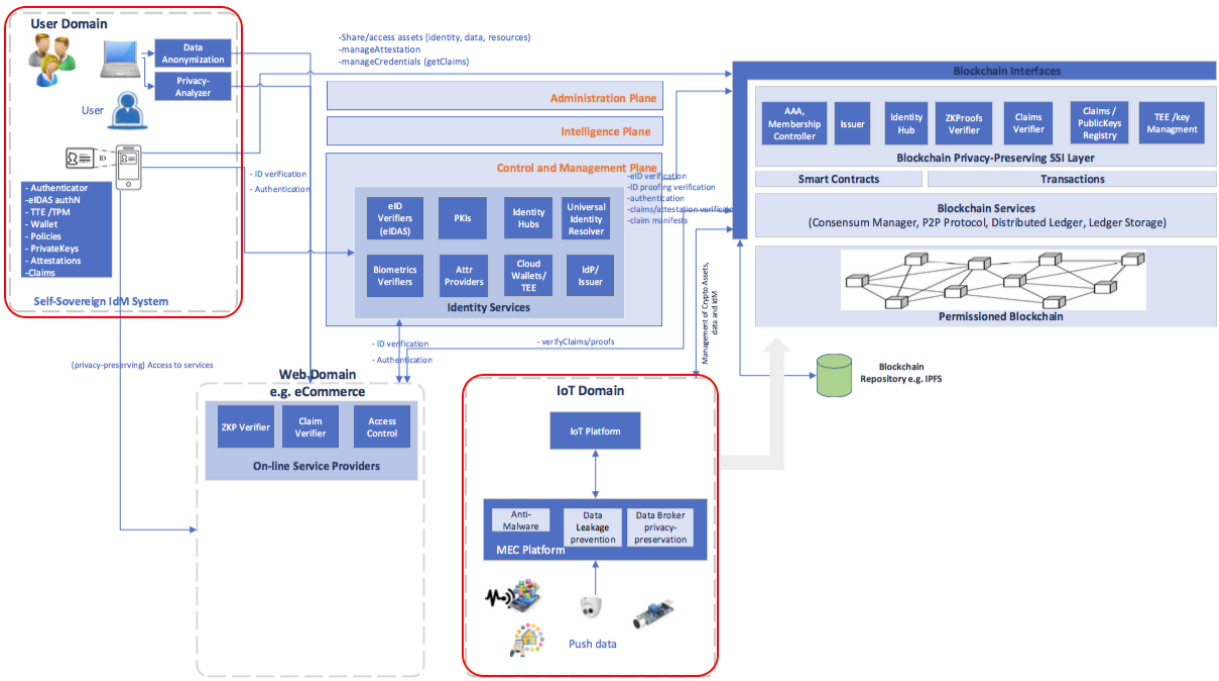


Figure 5: Mapping of the Geolocation Service in the General Architecture.

Data collection. We can have multiple sources for location data. First, sensor-based location can come from mobile phones and tablets with global positioning service (GPS) capability. These devices can determine their own location, combine it with a timestamp and encrypt the data. The data is uploaded then to the mobility analysis service, but without sharing the decryption key with the untrusted service provider.

Alternatively, location data can be collected from telecommunications operators who have bulk location data available from cell tower signaling (approximate triangulation of phone users' locations). This data is collected by the operators who can encrypt it in bulk and share it with the secure computing system. These two deployment options can be seen in Figure 5. Alternatively, the data set collected by the telecommunication operators could be anonymized to preserve the user's personal data privacy and facilitate the data analytics later on.

Data analytics. To calculate mobility statistics from the encrypted or anonymized data, the service uses a secure computing platform, e.g., a trusted execution environment (TEE). The exact method depends on the chosen technology.

Result presentation. Once the data preparation and analytics are completed, the results are stored in encrypted form on the platform. Authorized visualization services can then download the resulting mobility statistics and visualize them for the user. Different users can be allowed to have access to data using different views, as long as privacy is guaranteed. Remote attestation can be implemented to ensure that the correct code is being executed.

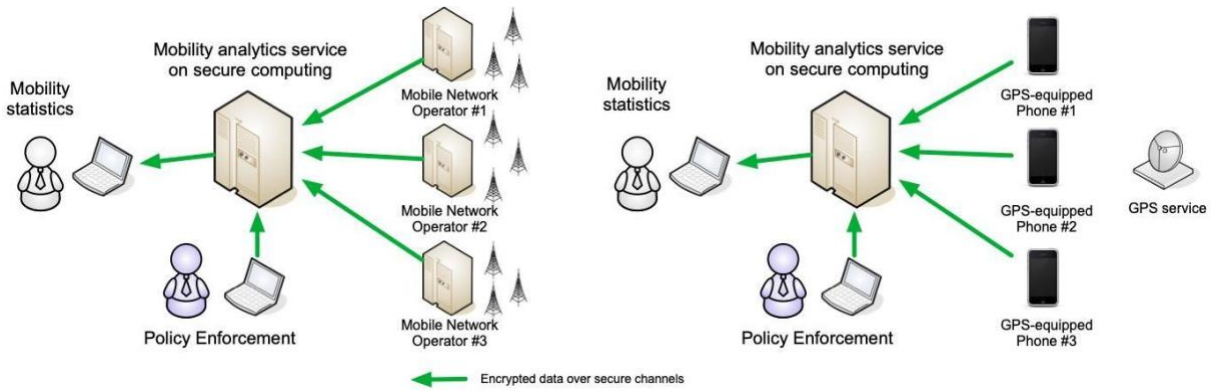


Figure 6: Two deployment options for secure computing

Research challenges. The main research challenges that the privacy-preserving data analysis of geolocation data addresses, is privacy by design. By using secure computing technologies like TEEs or secure multiparty computation, the privacy is built in at the start.

The more specific challenges are the following.

1. Data privacy challenge DP-02 (when using secure computing to analyze data, analysts are not able to see the individual data values).
2. Moving the internal state of an enclave to another platform is at odds with the main security provisions of SGX – the state of an enclave is private to the enclave itself.
3. TEEs need a subversion-resilient attestation mechanism that requires no secret (or secret share) on the hosting platform.
4. The privacy manager must be able to support the mobility of IoT devices without affecting data availability.

3.3.2. Control and Management Plane

This scenario uses privacy preservation tools to perform the data analysis so that the movement patterns of individuals are not visible to the data analyst. We propose using a trusted execution environment (TEE) so that the encryption key is stored inside a programmable hardware security module that loads the encrypted location data from storage and processes it within the same environment. At all other times, data remains encrypted, and the encryption keys do not leave the hardware security module. This way the data analysis can be carried out in a privacy-preserving manner so that even system administrators do not have access to individual values. In the case of anonymized data, the data analysis can be done without compromising the user privacy as the anonymized data are no longer personal data.

Sharemind HI can be used as the TEE, SR-EPID could be used as a drop-in replacement of the attestation protocol in use to guarantee security even in case the trusted execution environment is subverted, whereas the Elastic TEE asset can ensure that computing resources scale with the workloads.

3.3.3. User Domain

The application provides a PET client which is familiar to a data analyst (a language similar to R) that can easily be used to carry out data analysis or machine learning. The analyst will not be able to see individual records, just the aggregated results. The analyst does not need to know the details of how the underlying system works, the application provides them with a familiar workflow.

The Privacy Manager from the Edge-Privacy asset can be applied as means for the controlled release of information collected from devices belonging to Smart Campus students. The students' smartphones can be configured to upload their data to a Privacy Manager under the control of the user, which has been configured to encrypt location information before storing them in the Edge. The Privacy Manager can then receive queries from third parties and after checking, they are entitled to get access to the location information of the user, which can then be entered into the central system running the TEE.

3.3.4. IoT Domain

The location information necessary for the realization of this scenario may be obtained from IoT devices belonging to members of the university, including students and staff. Wearables such as smartwatches or smartphones can provide location information since they are fitted with positioning technologies such as GPS. However, not all these devices will be able to provide data in encrypted form.

The data collected from user related IoT devices can be uploaded to the Edge Platform where the Privacy Manager resides. Upon the reception of location information, the Privacy Manager will process the data so that it is securely stored by means of encryption. The Privacy Manager also serves as an interface to control who has access to these data.

4. Asset Demonstration

In this section, we describe the tools for the demonstrators in more detail. For each tool we first overview the components/architecture of the tool and include diagrams/screenshots to illustrate how the tool handles the challenges highlighted in the previous chapter, then, for each asset we address a simple example on the possible integration on the smart campus integration demonstrators.

For each asset we also provide a video explain the main concept and integration in the demonstrator[7].

4.1. PTASC

4.1.1. Overview

PTASC presents a decentralized, secure device-to-device communications solution in which device provisioning is focused on improving usability while providing security by default. The solution focuses on using a PKI where the CA is represented by a manager device that can be switched on/off to reduce single point of failure (SPOF) problems. The solution combines public-key cryptography and symmetric keys with the One Time Password (OTP) concept using a secure token. Device identity is guaranteed by physical access to this physical token. In addition to generating an OTP, the physical token also stores a public key to be transmitted to target devices only, eliminating attacks such as impersonation or man-in-the-middle. It also improves usability as we exclude configuration errors and difficulty choosing the right settings while provisioning the device. Although there is manual interaction to use the secure token, the process itself is as simple as finding the device to be provisioned and plugging in the security token. Along with the authentication and secure communications, PTASC encompasses a middleware layer solution for the IoT devices, which allows the control of the data generated by the device by its owner.

Manager Setup Phase
The manager device represents the CA system that plays an essential role in a certification system by signing public keys (or certificates) (Figure 7). This device should be assumed to be trusted and controlled only by trusted persons (such as the network owner). All certificates signed by the device will be implicitly trusted. Currently, systems that manage a PKI require a high degree of security and are installed on an isolated machine. In this proposed system, the PKI is installed on the manager device that is hybrid, meaning it may be offline from the network when not in use to prevent the possibility of the private key being stolen in a possible network intrusion. For added security, the manager device can use Intel SGX to secure all the cryptographic assets in a Trusted Execution Environment (TEE). The

[7] https://www.youtube.com/channel/UCSAJ78frZjdUTooAC4t6Wuw

manager begins by setting a CA using 256-bit Elliptic Curve Digital Signature Algorithm (ECDSA) and a 256-bit Elliptic Curve Integrated Encryption Scheme (ECIES) key pair to generate a shared key without the need for Diffie-Hellman exchange.

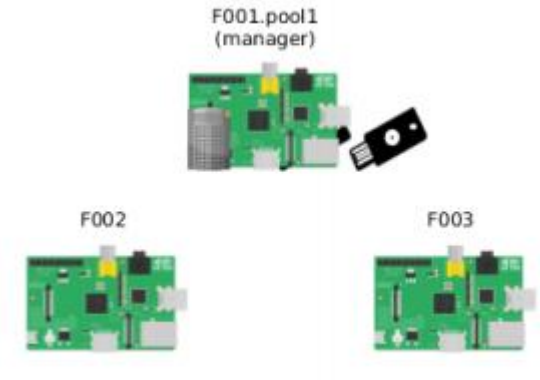


Figure 7: Manager Setup Phase

Device authentication

The authentication between a new device and the manager is essential for ensuring that it is added to the trusted device pool, Figure 8. To do this, the owner inserts the secure token into the target device, and then the new device is added to the pool (two cameras or one camera with another device).

As previously described, the secure token also has the manager's device public key. The new device starts by sending the Certificate Signing Request (CSR) of itself (which contains only the public key, not the private key, so the private key has not been compromised). When the new device sends the CSR to the manager, the latter will produce a signed x509 Certificate. Furthermore, it also sends OTP to verify that the new device is in physical presence with the security token and is therefore the correct device to authenticate. All this information is sent encrypted with the shared key (ECDHE) generated for both parties (client and manager) to encrypt a message that the manager can only decrypt. After the authentication is successful, the manager device sends back the signed certificate to certify that the client is a secure device added to the trusted device pool. These certificates are used to establish trust between client devices and provide decentralized, secure communication between them without the manager device's intervention.

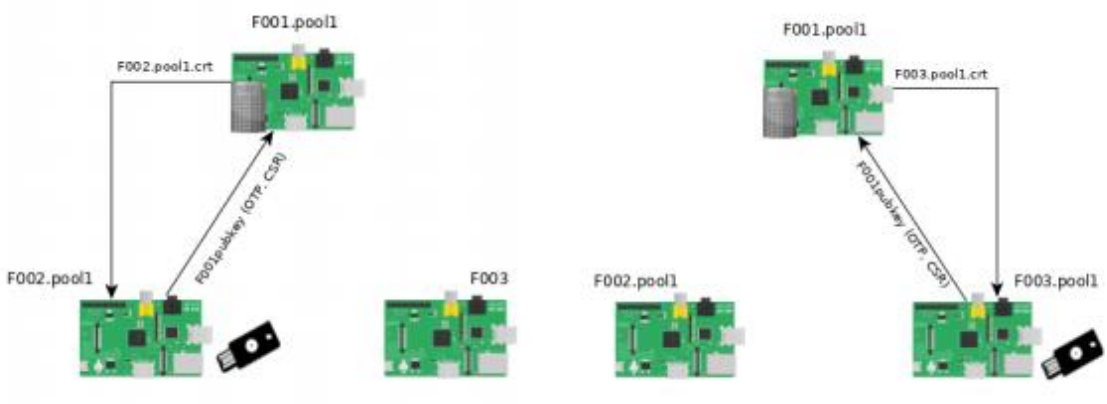


Figure 8: Device authentication

Decentralized Secure End-to-End Communications

After the discovery process, devices need to authenticate with each other. For mutual trust, both devices must exchange the manager's signed certificates. After verifying certificates' authenticity, a symmetric

key is generated between both devices to establish secure communication. Symmetric key generation needs authentication so that the nodes know each other. ECDSA was chosen for signing and verification, and ECIES was chosen for encryption. Then, Diffie-Hellman Ephemeral (DHE) or Elliptic-curve Diffie-Hellman Ephemeral (ECDHE) is used for key exchange. Ephemeral mode is important because if the key pair is used for more than a few hours, it must be stored somewhere. After all, devices can be turned off. There is always some risk that a stored key pair may be compromised, although a wide variety of methods can be and are used to mitigate this issue. This mode avoids this type of attack by not storing key pairs and generating a new key pair every millisecond, thus ensuring Perfect Forward Secrecy. After establishing a shared secret using ECDHE, the devices can exchange data with symmetric encryption using the secure cipher AES256 to encrypt messages. When all devices are provisioned, the manager can be turned off, Figure 9, until a new device needs to be added to the pool or there is a change on the device pool, such as certificate revocation or renewal.

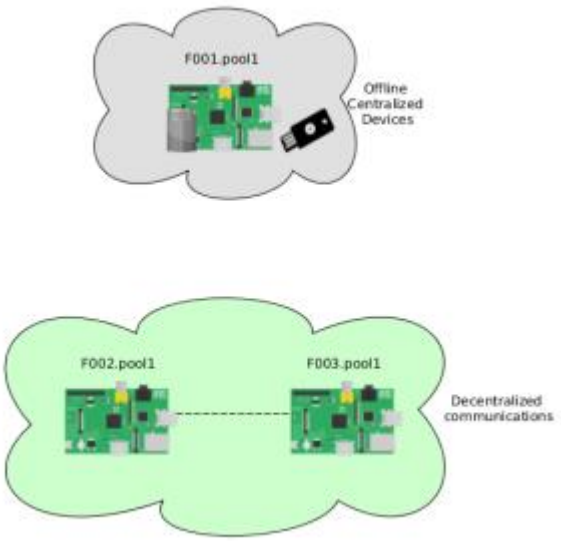


Figure 9: Decentralized Secure End-to-End Communications

Merge Two Trusted Devices Pools

An identity and authentication system must be flexible and highly scalable enough to handle billions of device infrastructures in multiple environments such as smart homes and smart cities in general. This system must support different environments, given the heterogeneity of applicability in IoT scenarios. For greater scalability, there needs to be a useful way to integrate different device pools to make the system more practical as it would not be feasible to re-provision devices already provisioned with another manager so that devices from different pools can communicate with each other. To address this issue, the system replicates the traditional mechanisms of having multiple CAs supported by a client. It is essential to ensure two points to deploy this in a real-world configuration: Use a secure token authentication scheme to enable enrollment and trust between different managers; and information dissemination on new pools among all new devices.

The authentication between managers (Figure 10 connection 1) uses the same mechanism to allow two pools to connect. After both managers perform mutual authentication, the next step is to spread the information across devices among different pools. To do this, the manager must send the signed and encrypted information to the devices. This allows any of the devices to read the information and verify the manager's signature. Figure 10 (connection 2) represents the agreement between managers and the corresponding spread of information from managers to their peers when they begin to trust each other

and announce on the network that others should move to include these new trusted colleagues in their trusted network.

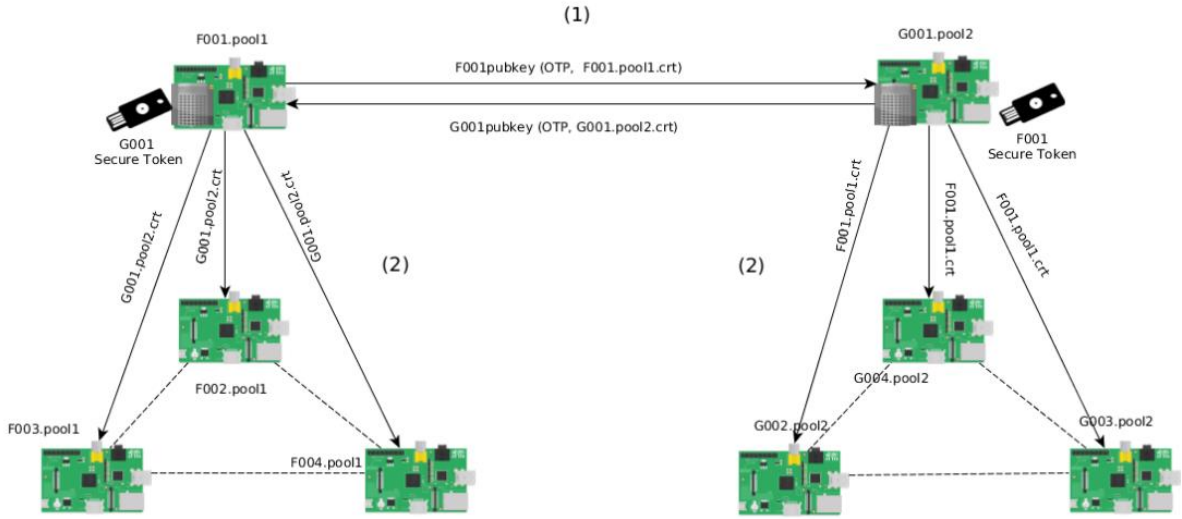


Figure 10: Merge Two Trusted Devices Pools

Privacy data controller

In PTASC, users can decide the data and traffic exchanged according to their preferences. Users have the option to block all traffic by default and to make exceptions for some specific domains. Therefore, users can block marketing/advertising sites and only communicate with the manufacturer's domains. Note that if users choose to block all communications from their devices to the Internet, some of the features may stop working, as some of these devices will not work in offline mode. Finally, users will have to choose between usability and privacy.

The middleware allows users to monitor incoming and outgoing connections to verify that the device is running an untrusted program and block or disable updates to specific resources, allowing them to block those connections "on the fly".

Along with this network traffic monitoring and, depending on the manufacturer's device firmware, users can store data offline on their home router for future reference. As users can connect with multiple routers (at home or work, for example), they can choose different permissions for their data depending on the device context, managing the data's life cycle.

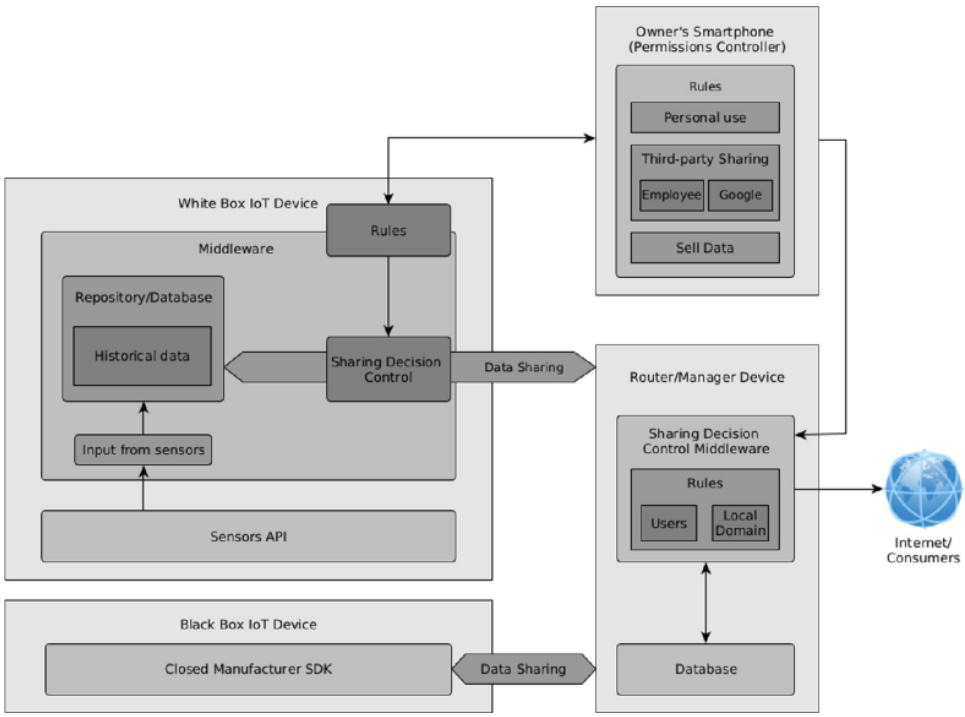


Figure 11: Privacy data controller

4.1.2. Research challenges addressed

Authentication and authorization of devices is a crucial aspect in the IoT ecosystem. PTASC addresses the current limitation of device pairing without using a PKI infrastructure or SSH keys. The main contribution of PTASC can be identified: Device Identity - A solution that relies on the combination of a secure token (capable of generating OTP and storage of a PKI) with cryptographic algorithms to provide an identity to devices. Managers can authenticate the trusted devices in their pool, giving them an identity; Devices' Pools - After device provisioning, the system can provide a decentralized architecture where the trusted devices can communicate end-to-end between each other (if they are in the same or trusted pools). Different pools can be trusted between each other if both managers agree on that (for this reason, we consider that the scalability is better than other systems that need to authenticate all devices between each other).

We can map this general discussion about challenges to the ones defined in D3.11. Concretely, the asset addresses the following challenges:

- DP-06 by providing mechanisms for control how the information is disseminated and control the private data on the communication.
- IDP-05 by providing a mechanism for guaranteeing proper identity management of things for authentication end-to-end.

4.1.3. Demonstrations Example

This work applies to video-surveillance cameras around the campus for authentication between them, forming islands of connected devices.

We can imagine a certification chain where the city is the root CA, and each manager can authenticate with the city manager, becoming a sub-CA, which means that when devices are authenticated with the

manager, they are also automatically authenticated with the city. In this way, we can have different independent pools, where all devices have certificates in which the city is part of the certification chain.

Before connecting to the city, the manager is a root CA for his devices but becomes an intermediate CA when he connects to the city. With this, the user's device certificates will have the city in the certification chain, which means that all the city devices can communicate. Thus, the city trusts the managers and their devices. By using Yubikey, we ensure that the city cannot accept provisions for non-manager devices.

Regarding privacy features, the administrator can control the data exchanged by the video-surveillance cameras, mainly by controlling whether the feed is being sent to other parties, such as third-party entities. Authorization is based on attribute/role-based access control, and we can configure different access for different device islands, which could map the hierarchical structure of organizations (different departments can be different islands, for example). In Europe, GDPR introduces a duty of a Data Protection Officer (DPO), who would be the preferred person to manage this structure. The officer would define different groups/classes and apply different policies regarding access to live streams. These policies must determine the video quality, the filters applied, and the time to access it. Thus, admins can change the permissions of feeds, which is an important layer to obey GDPR as the streams must not be shared with anyone other than those authorized as it contains sensitive information (both images and sound). Thus, admins have total control over who is accessing the video and to whom information about the camera is being sent (online connections).

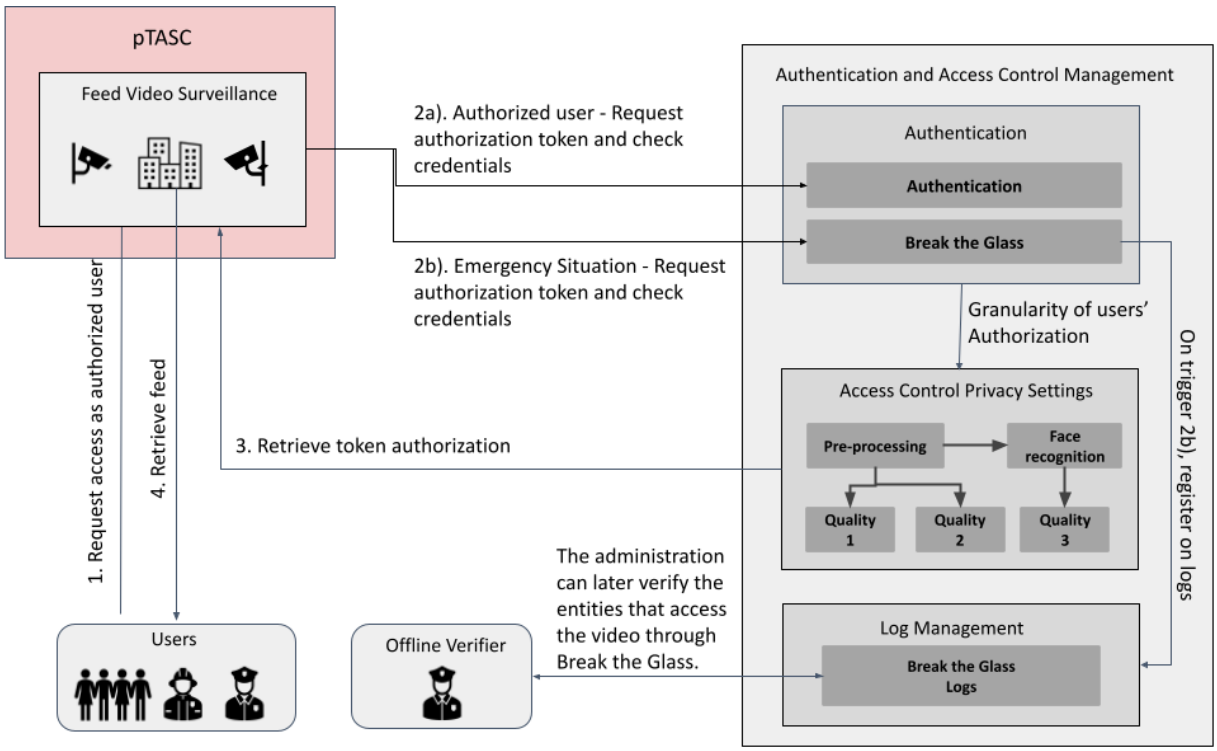


Figure 12: PTASC placement in the CCTV scenario

To simplify the process, in this demonstration, we will show only the provisioning phase, but further details are provided by Sousa et al. [sousa2021provisioning]. The provisioning phase is one of the most important phases of authentication because it guarantees an identity to each device.

Figure 12 represents the devices' provisioning process using PTASC, along with the construction of the trust anchors and certification chain.

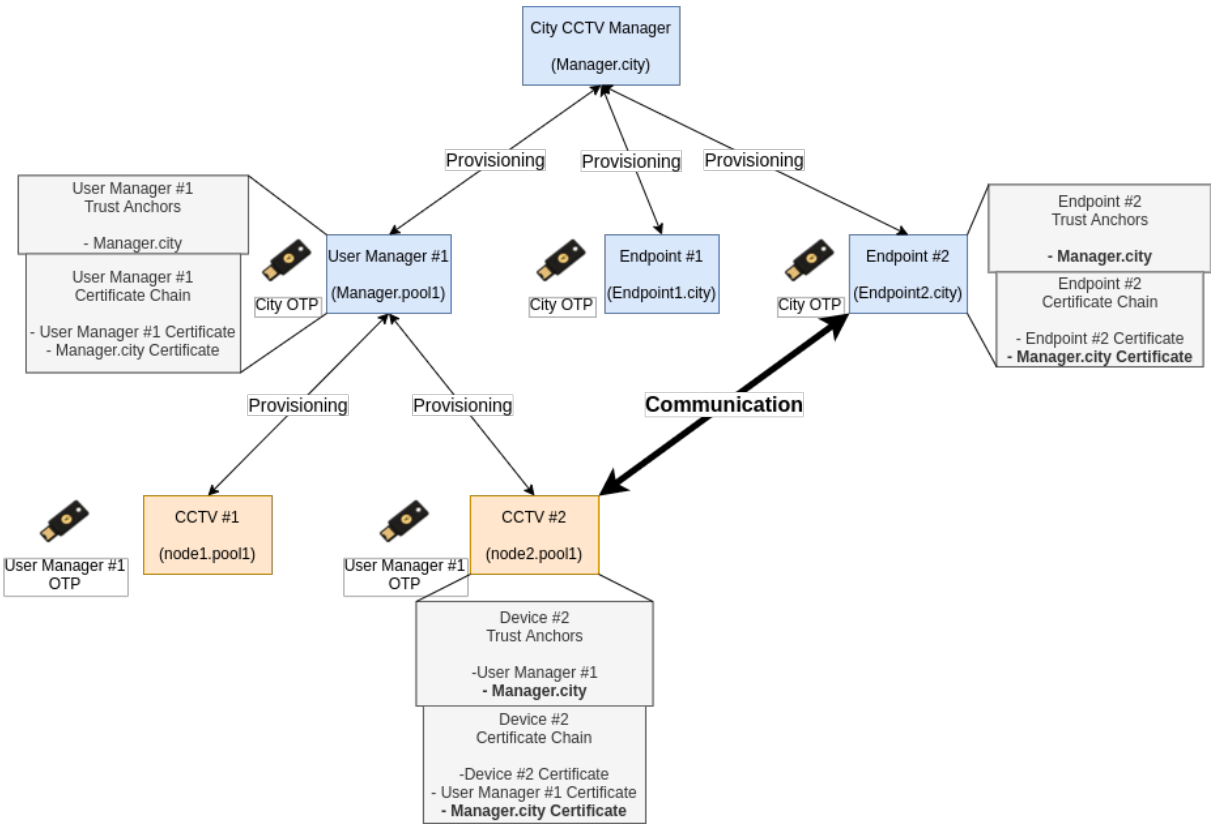


Figure 13: devices' provisioning process using PTASC

The certificate chain is created by adding the device certificate on top of the parent certification chain. The certification chain of the rootCA is its certificate. Developing the certification chains and trust anchors give advantages regarding the communication with all the nodes on the PKI tree, including nodes belonging to other users/company managers and nodes belonging to the city. The CCTV components described in Figure 13 are placed internally within the smart city infrastructure and are only accessible by endpoint nodes and the city manager (rootCA), not being connected to the PKI tree generated with PTASC.

The managers authenticated with the city manager are controlled by the secure token. The secure token combines public-key cryptography algorithms with OTP. The secure token acts as offline storage for private keys, allowing access to cryptographic operations to be kept offline without access to the network, contrary to the original solution. Each manager act as an OTP Server that can be switched off during the provisioning phase to reduce a SPOF problem. Due to the need to have a provisioning server, and since it is not necessary to be always on because not every day new devices are provisioned, the method works without it, having mutually authenticated devices that can communicate. This way, device identity is guaranteed by physical access to this physical token.

Decentralized provisioning and authentication in a video surveillance system encompass multiple independently operating cameras connected to a network. This decentralized approach to video surveillance ensures continuity in the case of a catastrophic network disruption. A vast number of authentication and security systems presented in the literature are centralized and frequently need to rely on some secure and trustful third party with communications. This, in turn, increases the time required for authentication and communication and decreases throughput due to known overheads, only all transactions are shared without centralized control. Decentralization provides the trust and safety of the data.

After provisioning between the manager and endpoint in the provisioning phase, the manager configures the user domain and policies for authorization grants on access to feeds according to the scenario policies. Note that the managers can authenticate other managers, creating the hierarchical structure for the city, organizations, and departments.

In brief, organization managers are the rootCA for their own devices until they are authenticated with the city manager. Then, when a manager connects to the city, it becomes an intermediate CA. With this, the user's devices' certificates will have the city in the certification chain and are authenticated to communicate with the endpoints for Orion.

4.1.4. Future Work

For future work that involves this asset (which may be realized during this project or not), we plan to improve the implementation including the usage of context-aware techniques to prevent attacks of impersonations and develop a disk image to allow to deploy directly when booting a new device.

4.2. ARGUS

Cloud storage allows users to remotely store their data, giving access anywhere and to anyone with an Internet connection. The accessibility, lack of local data maintenance and absence of local storage hardware are the main advantages of this type of storage. The adoption of this type of storage is being driven by its accessibility. However, one of the main barriers to its widespread adoption is the sovereignty issues originated by lack of trust in storing private and sensitive information in such a medium. Recent attacks to cloud-based storage show that current solutions do not provide adequate levels of security and subsequently fail to protect users' privacy. Usually, users rely solely on the security supplied by the storage providers, which in the presence of a security breach will ultimately lead to data leakage. We implemented a broker (ARGUS) that acts as a proxy to the existing public cloud infrastructures by performing all the necessary authentication, cryptography, and erasure coding. ARGUS uses erasure code to provide efficient redundancy (opposite to standard replication) while adding an extra layer to data protection in which data is broken into fragments, expanded, and encoded with redundant data pieces that are stored across a set of different storage providers (public or private). The key characteristics of ARGUS are confidentiality, integrity and availability of data stored in public cloud systems.

4.2.1. Overview

The main components of ARGUS can be seen as:

Privacy Broker:

We implement a broker that acts as a proxy to the public cloud infrastructures by performing all the necessary authentication, cryptography, and erasure coding. By doing so, it offloads the computational workloads from clients. Our approach ensures confidentiality, integrity, and availability of the data in the public cloud systems.

Dual Option File Encryption

In order to ensure the security and privacy of user data, the user can opt to encrypt everything locally and be responsible for the key management. In this way, the user does not need to rely on a third party. In addition, we have the option of delegating the encryption to a third party, which has the benefit of reducing the cost of execution on a limited device.

Confidentiality, Integrity and High-Availability

ARGUS maintains the integrity of the data as it stores an HMAC of all files. The confidentiality of the data is ensured as the data are encrypted, and the user can save their private key locally. ARGUS provides high availability through the redundancy that is assigned in the different cloud providers, that is, information is redundant on the three public clouds.

Intel SGX

The system uses Intel's CPU SGX extensions to cipher user credentials (access tokens that give access to the user's public cloud storage). This is an improvement over current implementations in systems that use the Google Drive API because the credentials are stored locally in the file system.

The general architecture of the ARGUS is represented in Figure 15, with all the components described in a detailed version.

4.2.2. Research challenges addressed

Storage of information in the cloud privately is a topic with years of development, but it is crucial to use the public cloud, for this many solutions exist, such as implementing encryption on the client side, split over multiple files. But current solutions lack on the privacy when sharing information or managing private data. In this context ARGUS in this project is extended to include a Run-time Adjustable Privacy Schemes (RAPS) an adjustable privacy mechanism that enables us to tweak the anonymization, storage location, and persistence parameters, allowing them to have more control over the processing that their data might suffer. This allows the system to detect anonymization patterns, using the parameters established on the RAPS, for possible privacy leakage or any other identifiable. This step is essential before sharing information with other parties, as it enhances anonymization and privacy. Also, we extended the work on the area with a Secure sharing using Machine Learning using MPC.

We can map this general discussion about challenges to the ones defined in D3.11. Concretely, the asset addresses the following challenges:

- DP-07 by providing anonymization mechanism for control on how the information is stored and control the private data on the communication.
- DP-08 by introducing the possibility of the user storing the files locally or in a private cloud depending on the user requirements.
- IDP-06 by ensuring that the encryption keys of the HTTPS protocol are manipulated in clear text only inside a trust zone (negotiating all the cryptographic material only inside the enclave)

4.2.3. Demonstrations Example

Since ARGUS is a cloud storage solution, the main advantage in the CCTV demonstration is the persistence storage, since the public cloud is the best option to retain the information over a small or long period allowing to scale as required. In ARGUS we take advantage of the known advantages of public cloud providers and use them to store information privately. In this example we can, for example as shown in Figure 14, manage the log management of the feeds accessed by each user.

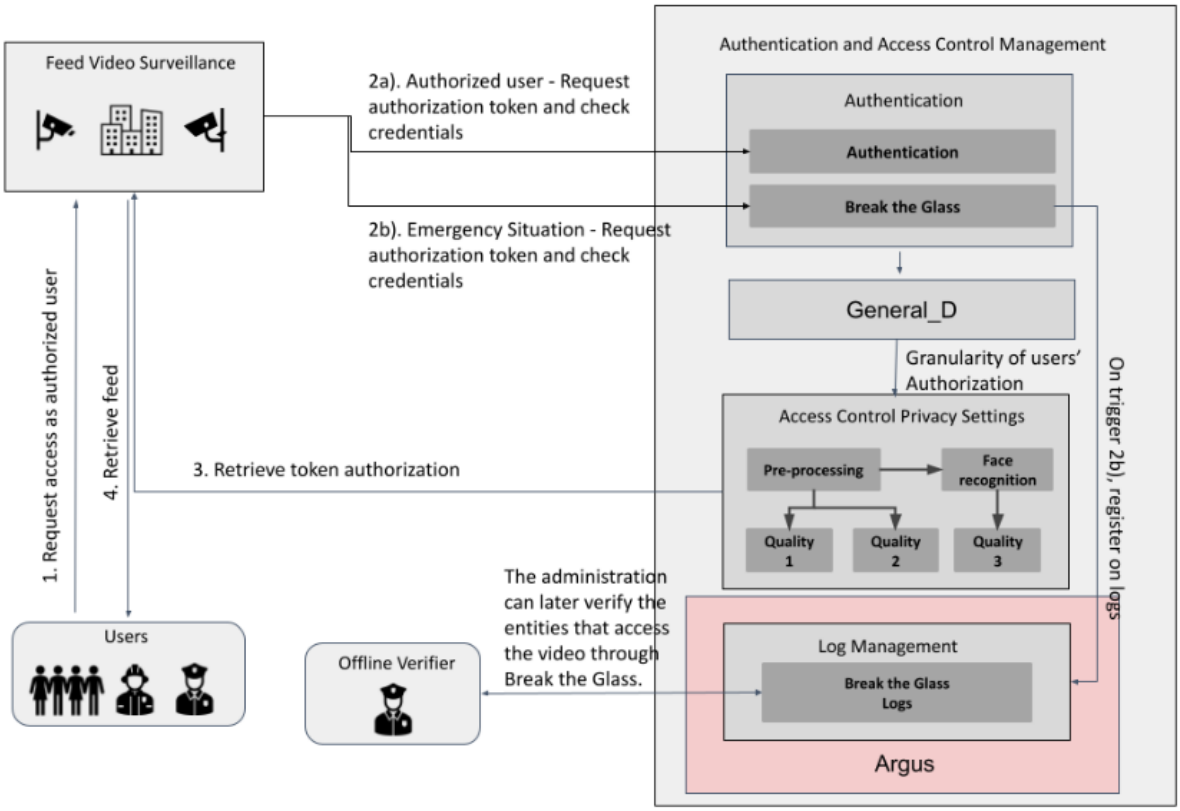


Figure 14: Argus location in the CCTV demonstrator

In our architecture, the user or CCTV software needs to be authenticated with the ARGUS server to upload the log file/stream content and communicate via HTTPS; otherwise, the connection is refused by the Authentication and SGX module (1). When the user uploads, Figure 15 on the left, the content is analyzed according to the specifications of the admin. If personal data are detected, it splits the information according to the anonymization algorithm that applies to the received file type. To cope with this, we allow the user to send the file together with a list of portions of the file content that contain personal information. The anonymization uses that information to redact the personal data from the received file, storing the private information on the private cloud (2). After an anonymization process, two files are generated: one stored on the private cloud, containing the personal data, and the other stored on the public cloud, containing the remaining data. Data must be encrypted and split (3) before uploading it to the different public cloud providers configured (4). A hash is generated for the split data chunks used to verify the integrity of the data on the download process. The hash, the encryption keys, and additional file metadata are stored in an index, allowing the on-premises server to know which files are stored and how to retrieve them. The upload process can be seen as a sequential process that starts in U1 up to the storage in U5. Contrarily to the upload of a file, the download is more complicated. It consists of the same processes, but the order is different. First, the user is authenticated in the system (D1), then the information is collected from the cloud-of-clouds and private storage (D2), then the file recovers the original format and is checked against the HMAC (D3), and finally, the final step is the analysis of the personal identifiers (D4) to recreate the final file to return to the client (D5). This architecture in the next deliverable may include a cache mechanism to maintain the most recent feeds on the premise of the infrastructure reducing the overall latency when using the CCTV feeds.

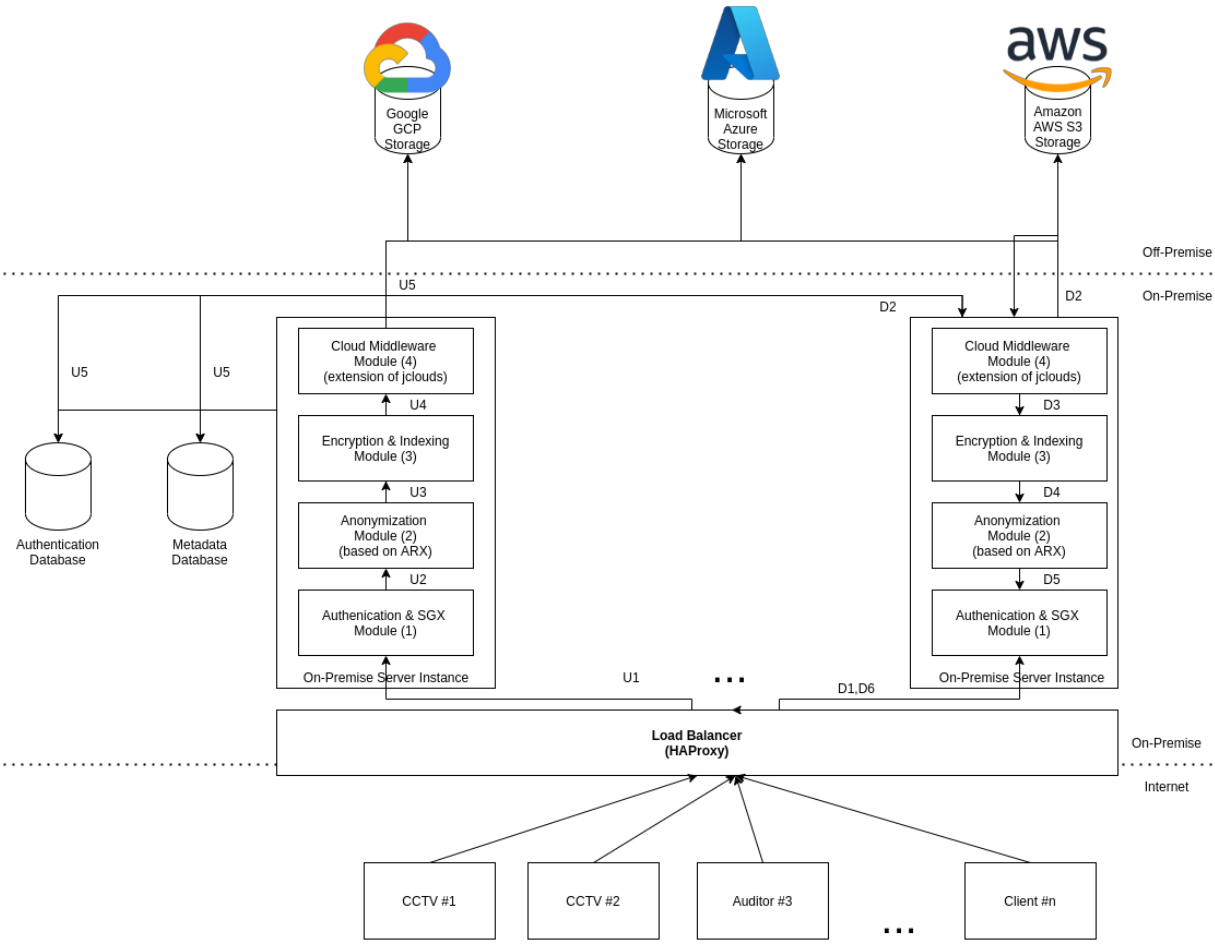


Figure 15: Argus architecture

4.2.4. Future Work

For future work that involves this asset (which may be realized during this project or not), we plan to improve the implementation including the usage of searchable encryption techniques. Also, we would like to implement a cache mechanism to improve the overall performance of the system.

4.3. Self-Sovereign Privacy-Preserving-IdM (SS-PP-IdM)

This asset leverages the OLYMPUS virtual identity provider, which is comprised of multiple individual IdPs, to manage user identities and authentication. It relies on distributed p-ABCs to offer privacy-preserving (minimal disclosure and unlinkability) and authentication (presentation of attributes) linked to eIDAS. Moreover, the asset proposes a trust framework based on Blockchain to complement the usage of credentials.

4.3.1. Overview

The smart campus scenario shows a full integration of OLYMPUS together with a blockchain infrastructure, which allows to increase the trust of the architecture, and an eIDAS node that allows the inclusion of certified attributes. The main components of the SS-PP-IdM are:

- **OLYMPUS vIdP**: OLYMPUS virtual identity provider comprised of multiple individual IdPs. Leverages distributed p-ABCs to offer privacy-preserving (minimal disclosure and unlinkability) authentication (presentation of attributes).

- **Blockchain platform**: Stores public information about the OLYMPUS infrastructure (endpoints, vIdP and cryptographic parameters) in a trusted way. This information can be consulted at any time.

In this case, we will rely on attributes strongly linked to user's identities through the use of eIDAS. Because of that, two extra components are involved in the identity management platform:

- **Keyrock**: Used as a bridge to eIDAS (i.e., handles SAML communication flow with eIDAS node to obtain certified attributes).
- **eIDAS node**: It handles authentication (of a natural person in the first pilot) with an electronic certificate or national eID following the eIDAS specification.

Lastly, the integration into the Smart Campus scenario means that there will be a close relationship between the identity management system and the smart platform's authorization framework:

- **Smart-campus platform**: Smart campus platform that offers services for Murcia city
  - **Services**: Public transport, parking availability…
  - **PEP**: Controls access to the services, checking that the request includes a valid capability token (i.e., the request is authorized).
  - **Capability Manager**: Generates capability tokens that bestow authorization to use specific services. Relies on the PDP for the decision (using XACML).
  - **PDP**: Checks if an authorization request should be conceded, using the OLYMPUS verification library to validate the presentation token against the policy.

In any usage of our identity management solution there are three key processes:

- **User Enrolment (a)**: A user account is created and populated with the attributes that will comprise her identity in the scenario. This will only happen if the user correctly proves possession of said attributes, with assertions being verified by the vIdP. Note that the sub-process of managing (adding, removing…) attributes to the user account can also be executed at other times.
- **Credential Issuance (b):** After they have enrolled, users can authenticate against the vIdP to obtain a credential. The resulting p-ABCs will hold all the user attributes and can be (securely) stored for later use.
- **Privacy-preserving service access (c):** Accessing services can require that a user fulfils some attribute-based policy. With the credentials generated in process *b* the user can prove the necessary assertions while ensuring minimal disclosure and unlikability between requests.

4.3.2. Research challenges addressed

Authentication and authorization of users is a crucial aspect in all online interactions, and an especially complex process in a large, heterogeneous and dynamic environment like the smart campus scenario for this demonstrator. Protecting the smart platform's services and data is not a simple task, and further challenges arise when privacy and security are pivotal objectives.
In that respect, being able to perform access control by authenticating users while ensuring privacy principles are respected is a key challenge. Attribute-based access control has been proposed for authorization in smart platforms, and more specifically, solutions like XACML framework [OASIS13] offer capabilities for fine-grained access control. However, these solutions have been applied with traditional federated identity management systems like OAuth [Hardt12], which have glaring issues in terms of privacy and security, such as user tracking or breaches of minimal disclosure. The SS-PP-IdM asset is based on distributed p-ABCs used in the OLYMPUS project [MBGFSMSPS20], allowing users to control which information is revealed when interacting with the capability-based access control and avoiding the IdP as a single point of failure. What is more, it tackles the lack of homogeneity in representing p-ABCs, one of the issues that jeopardized adoption of previous p-ABC systems, by integrating with the emerging W3C Verifiable Credential standard for serialization [GMBS21].

However, the original OLYMPUS IdM, like other similar proposals about privacy-preserving identity management [BDMCBS20], does not address the challenge of establishing a complete ecosystem where the different components can interact and ensure trust. The SS-PP-IdM asset defines and implements a complete environment with the addition of blockchain for auditing, traceability, and trusted public information (i.e., public keys, parameters, and support as verifiable data registry for W3C Verifiable Credential's elements). Thanks to this, users can safely discover and interact with identity and service providers [MGBS21].
Lastly, and also related to the challenge of security and trust in authentication, the asset addresses the challenge of establishing strong links to physical identities through the integration of the eIDAS [Dumortier17] system into the solution.

We can map this general discussion about challenges to the ones defined in D3.11. Concretely, the asset addresses the following challenges:

- *IDP-02: Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks*. The pp-IdM used in the scenario enables minimal disclosure through the application of p-ABCs. Also, it takes usability into account, offering simple authentication mechanisms to users easing their interaction with access policies (more work on this on T3.6).
- *IDP-03: User's privacy-preservation of transactions in distributed and immutable systems (e.g., blockchains).* This asset tackles the challenge by relying on zero-knowledge proofs and smart contracts (based on blockchain) that allow anonymity in interactions without storing sensitive data in the ledger.
- *DP-04: Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP)*. Our oblivious distributed pp-IdM avoids tracking by IdPs and the p-ABC scheme used provides unlinkability and minimal disclosure, avoiding tracking through the authentication mechanism by service providers (unless users consciously decide to reveal information that identifies them).

4.3.3. Demonstrations Example

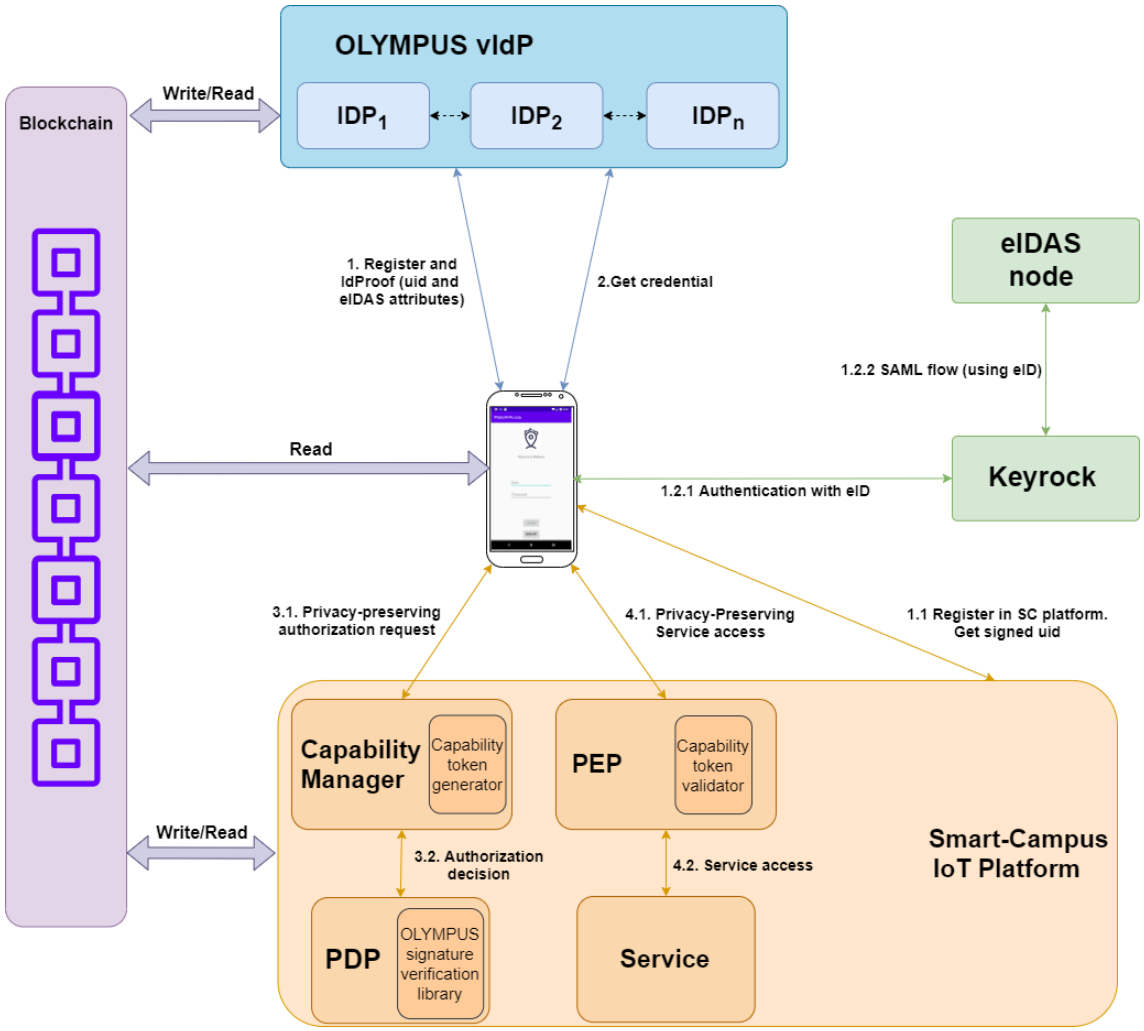


Figure 16: Example application architecture flow

The asset will be **validated** using two subcases. The first will demonstrate the application of the SS-PP-IdM to the Smart Campus scenario to perform the three processes (*a,b,c*) mentioned in the previous section, accessing a Smart Campus service. The second test case involves using the SS-PP-IdM to perform user enrolment in the university. The user can then show that she is a student using again the tools provided by the IdM, while not revealing any other data.

Figure 16 shows the architecture of the demonstrator, as well as an example flow involving the instantiation of the processes described in the previous section to this scenario. As we can see, the instantiation involves multiple steps specific to this use case, where the user relies on her mobile application to perform the different operations.

During **user enrolment**, the application redirects to the Keyrock platform so the user can obtain a set of attributes certified by an eIDAS node (using her eID). Also, the smart platform generates an id, so the user can be linked to an account in the platform. The application presents both assertions to the vIdP, and a successful verification will result in her OLYMPUS account having the information needed to generate credentials.

The **credential issuance** process, conversely, is equivalent to the one in any other application scenario. The user inputs her OLYMPUS account username and password to perform a login operation. If authentication is successful, the distributed issuance process will result in the user having a credential with her attributes certified by the IdM, which can be securely stored.

In this case, the **privacy-preserving service access** involves interaction with the smart platform Figure 17 shows the services that appear on the application's homepage. It may be conceptually divided into two sub-processes, authorization and actual service access. During the former, the user must prove that she fulfils some attribute-based policy (e.g., the one shown in Figure 18) to obtain the platform's authorization (in the form of a capability token issued by the capability manager). Later, the PEP checks that the user request includes a valid capability token to allow (or not, if the verification fails) service usage.

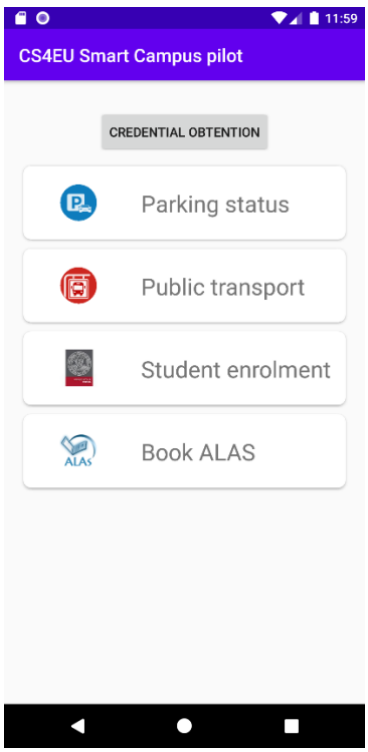


Figure 17: Application Homepage of the Smart Campus

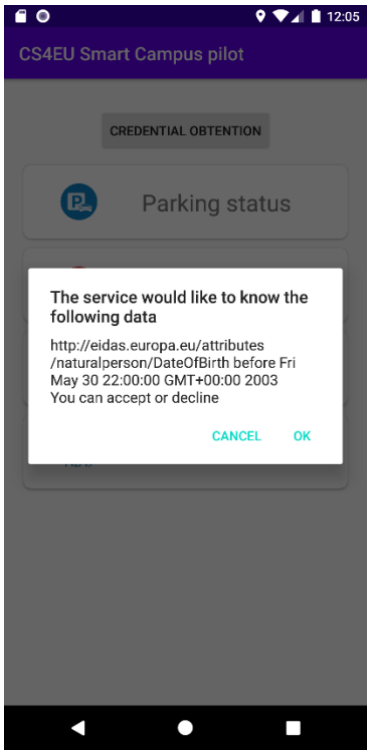


Figure 18: Using some attribute-based policy

Throughout the process, the involved actors will use the ledger to store and retrieve trusted public information. For example, vIdPs will register public information (endpoints, public keys…) in the ledger. The user application (and the verifier installed in the smart platform) can retrieve that information to perform the necessary configurations and ensure that the entity it is interacting with is trustworthy.

For this demonstration, we are deploying the necessary components (OLYMPUS partial IdPs, Keyrock, Smart platform's authorization components) in an instance (Ubuntu server 18.04, 8 GB RAM) deployed in an OpenStack platform within the University of Murcia. The ledger deployment (Hyperledger Fabric v2) also relies on OpenStack virtualization. The Fabric infrastructure consists of 2 organizations with two peers each, a certification authority and an Orderer node. Every Hyperledger machine is running Docker v19, and Docker composes v6.14. Also, we are implementing a user application in Android, controlling processes like enrolment, authentication and interaction with the smart platform.

We already have performed tests for the different needed functionalities with the current implementation, including (but not limited to): initialization with public information in the ledger, enrolment with Spanish eID, credential issuance, and authorization (and service usage) process. We

generated Verifiable Credential and Verifiable Presentation generated during a test execution of the application (accessing parking availability service, which requires being over 18). We even have some preliminary performance evaluation, through execution time measurements (using a Poco X3 NFC device) for the computations of different operations (note that the verifiable presentation requested in the process involved only revealing 2 attributes). The results shown in figure 19 include, on one hand, the enrolment and credential issuance performance, as well as the time required by different sub-processes (authorization through cap-tokens and p-abc credential evaluation) to handle the access request, as shown in the second chart of figure 19.

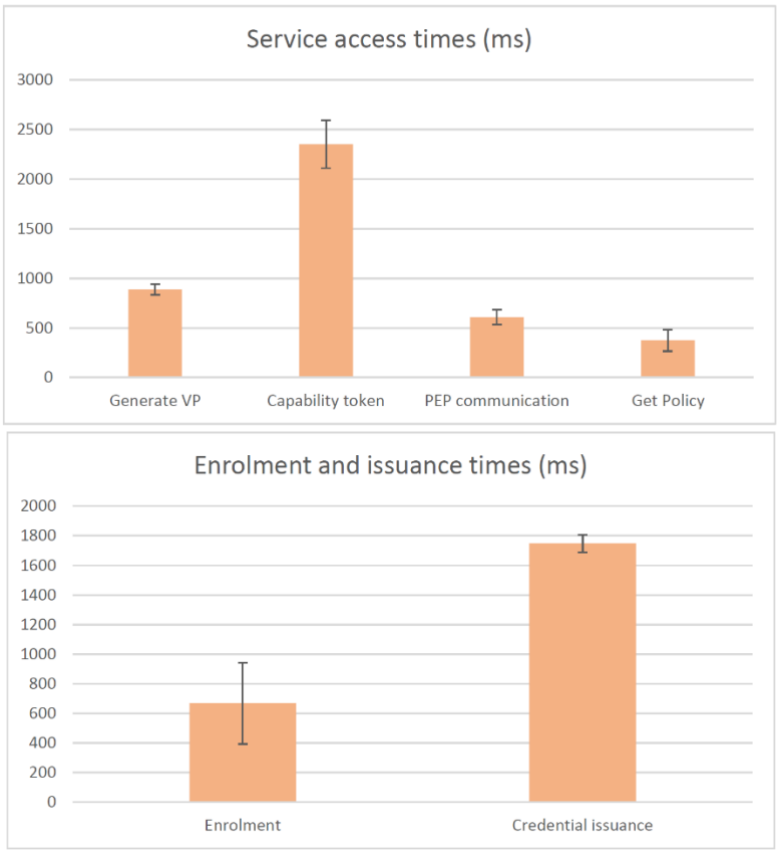


Figure 19: Preliminary execution time measurements for some of the operations

4.3.4. Future Work

For future work that involves this asset (which may be realized during this project or not), we plan to improve the ppIdM system with extra functionalities like inspection, revocation, set-membership proofs, or recovery functionalities. Also, we will explore the application of the p-ABC approach for identity management of constrained IoT devices. This might be accompanied by research on solutions for whole system integrations where identity management schemes coexist with other solutions to accommodate the different ranges of devices in terms of power (e.g., delegation for the most constrained, full capabilities for non IoT devices involved, reduced functionality for IoT devices “in the middle” etc). Lastly, we will explore architectures (and components that realize them) to provide pp-IdM applicable in distributed zero-trust scenarios, with dynamic participation, monitoring, dynamic trust, tackling the relationship between privacy and trust.

4.4. Password-less authentication

4.4.1. Overview

The password-less authentication asset is based on the FIDO standards[8]. It provides a device-centric authentication that implements a) a challenge-response scheme in which the user is authenticated locally (i.e., on the device that it is deployed to access the service) using alternative authentication methods, such as PIN, USB keys, and biometrics and b) public key cryptography to authenticate the device in the service. During the FIDO authentication, when a user (in our case a student) is authenticated in its device (for instance, using a USB key), it unlocks its private key, which subsequently is deployed to sign the challenge and the service deploys the user's public key, to decode the challenge. An overview of the FIDO authentication process is depicted in Figure 19.

Figure 20: FIDO Authentication Concept

The password-less authentication asset aims to replace or enhance the traditional username-password paradigm offering a more user-friendly and secure authentication procedure. Via this asset the students will deploy their smartphones, which are widely used in everyday life as well as in the campus premises, to be authenticated on numerous services. For this purpose, the FIDO protocol will be used, which will be responsible for the registration, authentication, and de-registration of the students.

A) Registration

The registration process allows the Server to verify the authenticity of the Authenticator and register it along with the user's account. The authenticator is represented by the device that the user will deploy to authenticate in the smart-campus platform, which in our case is the smartphone. Once an authenticator has been validated, the FIDO server can assign a unique identifier number (aaid) to the authenticator that can be used in future communication between the two parties. Specific policies are utilized in the FIDO server for the acceptance or not of the authenticator (for instance, some smartphones that have weak facial recognition can be omitted). The registration steps are described below:

1. The client application initiates the registration process.
2. The server sends a registration request.
3. The user enrolls in the client application and the key pairs (private and public) are created.
4. A registration response is sent to the server.

[8] https://fidoalliance.org/

5. The server validates the response

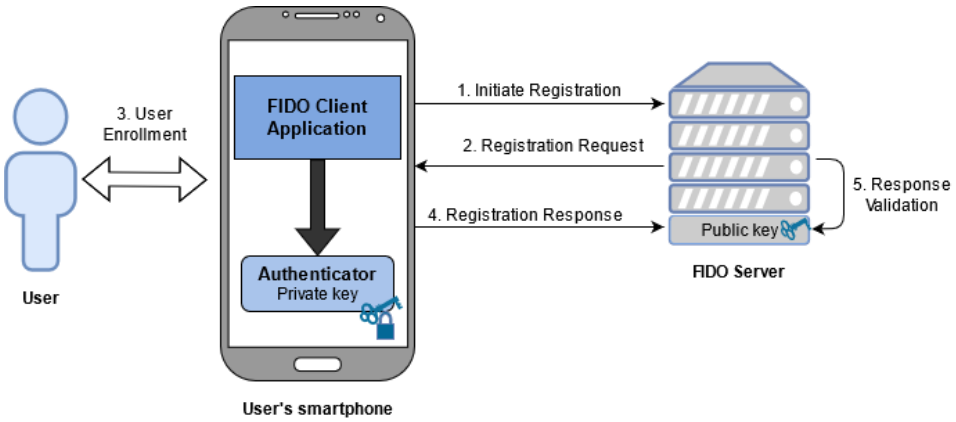


Figure 21: Registration

B) User Authentication

The user authentication process is based on a cryptographic challenge-response scheme in which the user is prompted by the server to be verified by the authenticator that was used in the registration process. The authentication is initiated by the client application, where the user chooses its preferred authentication method (i.e., fingerprint, PIN, pattern). The steps of the authentication process are:

1. The client application initiates the authentication process.
2. The server sends an authentication request along with a challenge.
3. The user is authenticated to the device using its preferred authentication method to unlock its private key.
4. The client application sends to the server the challenge signed by the user's private key.
5. The server validates the challenge using the user's public key.

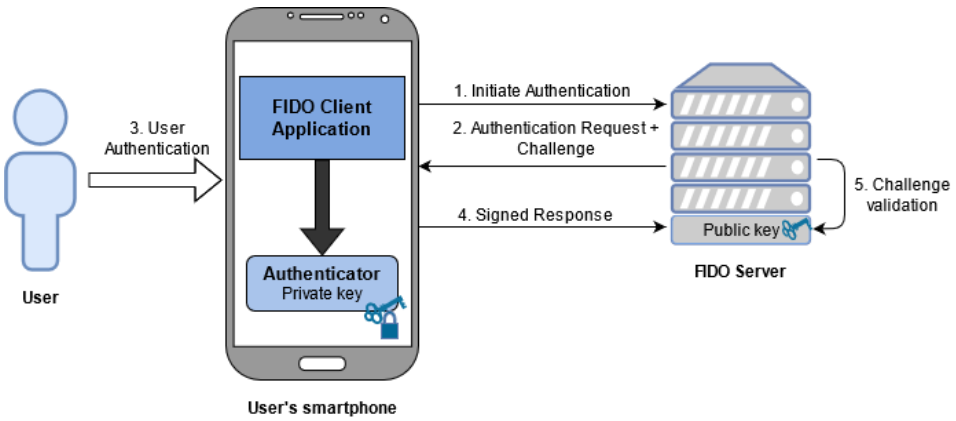


Figure 22: Password-less Authentication

C) De-registration

De-registration is required when the user account is removed from the Server. The server can trigger the Deregistration by asking the Authenticator to delete the associated FIDO Credentials that are bound to the user's account. The de-registration steps are:

1. The client application initiates the de-registration process.
2. The server sends a de-registration request.
3. The server deletes the user's account and the public key from the DB.

4. The client application deletes the information from the device.

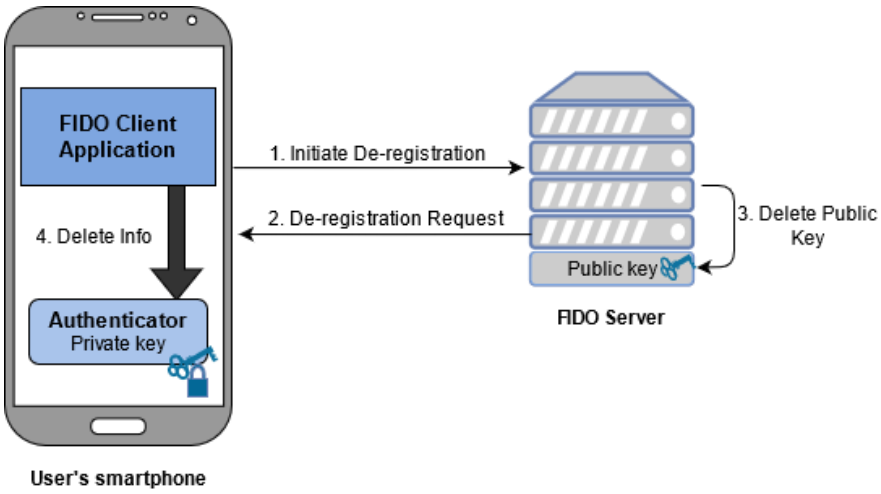


Figure 23: De-registration Process

In demonstration scenarios of CyberSec4Europe, the password-less authentication asset will provide an additional level of security in the authentication procedure, which will be based on the criticality of the service that the user wants to access.

4.4.2. Research challenges addressed

The authentication process of most web applications relies on the password paradigm. It is evident that a password can be considered secure when it contains 15 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[9]. The massive number of online accounts has led to a password overload problem that directly impacts the security and privacy of users' data, since they try to deal with this problem by simplifying their passwords or reusing the same password on different accounts or keeping their passwords unprotected, on paper or password managers. At the same time, passwords are targets of multiple attacks, as they can be leaked, key-logged, replayed, eavesdropped, brute-force decoded and phished. All the aforementioned reasons have rendered the traditional username/password authentication solution unreliable and made user authentication an open research challenge.

The password-less authentication asset aims to contribute to the challenging task of authentication by providing a solution that is both secure and user-friendly. Particularly, it addresses the following research challenges [vasile2021web, Angelo2021how, panos2017security]:

- The mitigation of significant security gap that comes from the extended usage of the username/password scheme by providing a password-less authentication solution that deploys secure methods to authenticate a user that merge sophisticated cryptographic algorithms (e.g., elliptic curves) with *something the user knows* (e.g., PIN), *something the user has* (e.g., USB key), or *something the user is* (e.g., biometrics).
- The enhancement of the authentication process by utilizing a two-factor authentication (2FA) mechanism. With 2FA an additional layer of security is added on the authentication process, where the user after its initial authentication with username/password is requested to provide an additional piece of information (namely, *something the user knows*, *something the user has*, or *something the user is*).

[9] https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/authentication-methods

- The limitation of password-related attacks, such as phishing, key-loggers, eavesdrop, and brute force.
- Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks (IDP-02).
- Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP) (IDP-04).

4.4.3. Demonstrations Example

The password-less authentication asset will be deployed in CCTV Surveillance in the Smart Campus and in Identity Management and service usage in Smart Campus scenarios. The asset's contribution will be similar in both scenarios and its aim will be to enhance the authentication process in a user-friendly way.

4.4.3.1. CCTV Surveillance in the Smart Campus

In the CCTV Surveillance, the password-less authentication asset based on FIDO 2 will be deployed to strengthen the user authentication process by implementing a two-factor authentication approach. Particularly, FIDO 2 will be integrated in the Identity Management system (i.e., Keycloak IdM), to add an extra authentication layer, namely the user will be authenticated by the first level using her username and password and by the second level using a password-less method, such as USB key, PIN, fingerprint, etc.

In Figure 23 one can notice the architecture of the demonstrator, as well as the integration of the password-less authentication asset at a high level. The password-less authentication communicates directly with the Authentication and Access Control Management component to exchange user information regarding the registration, user authentication, and de-registration processes as described in the previous section.

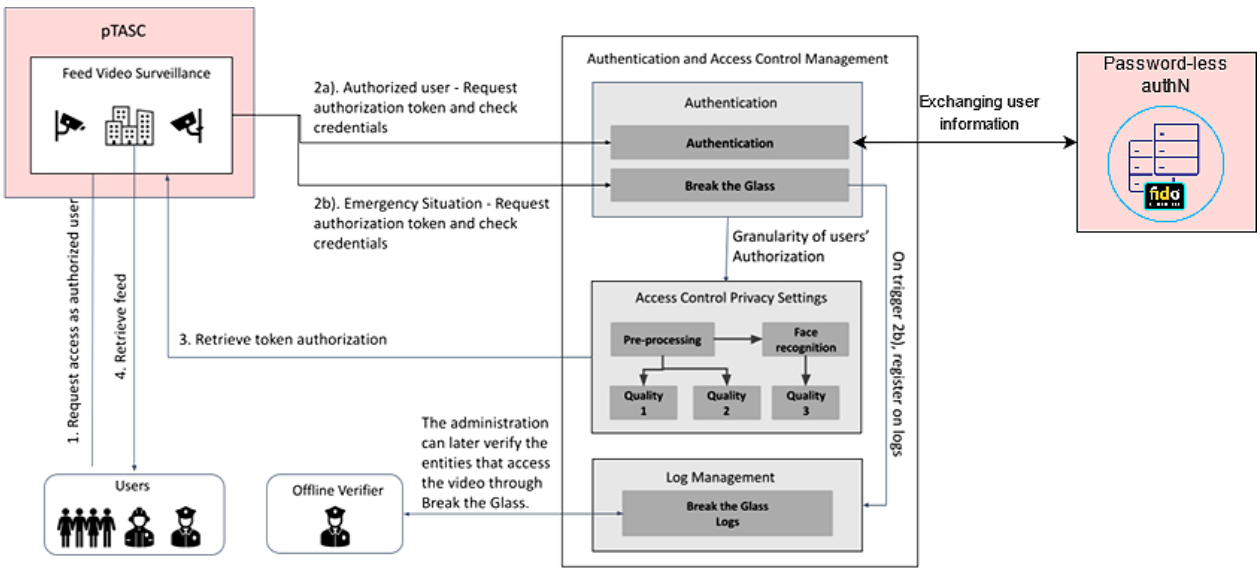


Figure 24: Password-less authN in the CCTV scenario

We argue that user-friendliness should always be taken into account, especially in the case of authentication solutions, thus we paid a lot of attention to providing a solution for two-factor authentication that will be attractive for the user. In the rest of the section, we will elaborate on the asset's main functionalities in the CCTV Scenario.

In the registration process, the user accesses its profile information in the account that it already has in the Keycloak IdM and in the Account Security>Signing In>Two-Factor Authentication>Security Key>Set Up Security Key field it registers the security key of its preference. To register a security key, the user will first have to connect its authenticator device on their machine (e.g., plug in its security USB token) or use the device's internal authenticator (e.g., Windows Hello). Then he will be able to initiate the registration process from the web interface and interact with his authenticator device to prove its presence (e.g., click a button on the USB token, scan his face). Right after the authenticator will send a newly generated public key bind to the user's account for this service.

In the example presented in the Figure below, the user already registered its machine's Windows Hello authenticator. Depending on the machine's hardware this user would be able to use the machine's internal authenticator and authenticate to the service using his Face, Fingerprint or just a PIN.

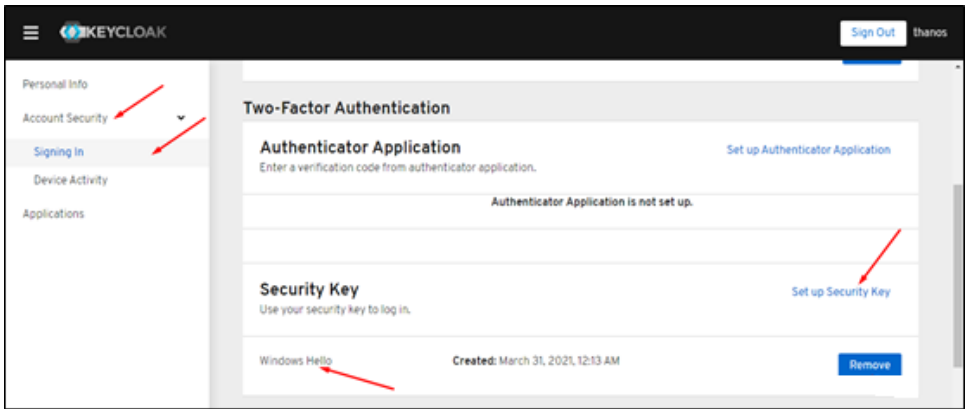


Figure 25: Password-less authN Registration

At this point it is important to mention that depending on the organization's security requirements, we can limit the authenticators that can be used with the service only to trusted ones that meet the organization's accepted security standards. For instance, the organization may configure the service to only accept specific security USB tokens (by adding the authenticator's GUID on a whitelist) and thus avoid the user of weak authenticators (e.g., face recognition authenticators with high false positive rates). This and other security policy configurations can be configured.

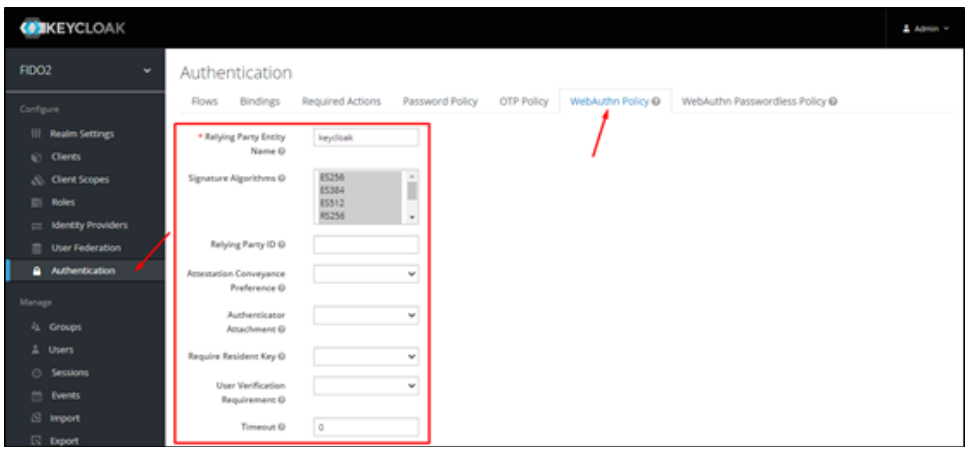


Figure 26: Configuring FIDO2 policy on Keycloak IdM

After the registration process, in the **authentication** process, the user will be asked to deploy its username and password to login into the service, as shown in Figure 26. However, after the successful authentication in the first layer, the user will be asked to connect its authenticator device that was registered to its account and prove its identity, otherwise, the authentication will fail. During this process, once again, the user will have to connect its authenticator (e.g., a secure USB token) or use the internal authenticator of its machine (e.g., Windows Hello).

The authentication process flow can be seen on the following Figures using a Username-Password and a machine's Windows Hello authenticator already registered to the account, configured to work with a PIN.

Figure 27: Password-less authN Authentication Level-1

Figure 28: Password-less authN Authentication Level-2

The **de-registration** process can be used to remove an authenticator from an account. For instance, if the user wants to remove an old authenticator (e.g., the Windows Hello authenticator of its old laptop), it can deploy the de-registration process to achieve it. In the de-registration process the user has to follow the first steps of the registration process *Account Security>Signing In>Two-Factor Authentication>Security Key* and select the *Remove* option to remove the authenticator of his preference.

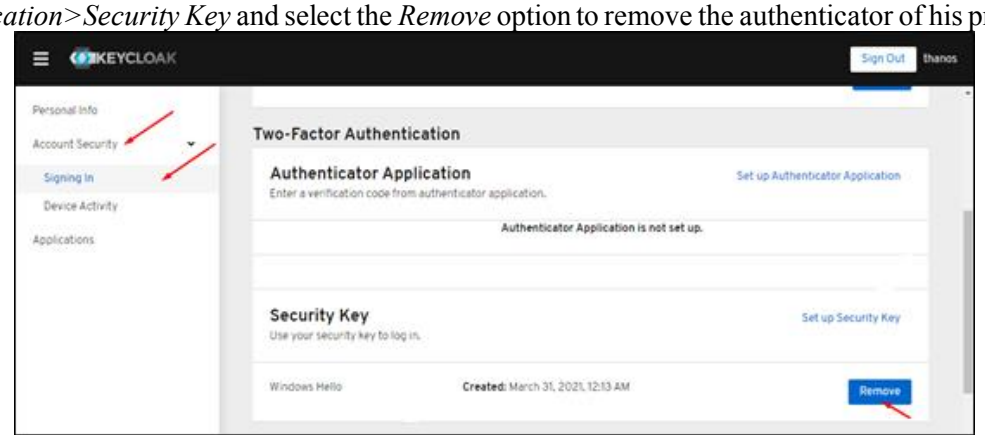


Figure 29: Password-less authN De-registration

4.4.3.2. Identity Management and service usage in Smart Campus

Moreover, in the Identity Management and service usage scenario, the password-less authentication asset will provide a user-friendly alternative to the eIDAS authentication, where the users will not have to deploy usernames and passwords, instead they can use their smartphone to access the campus services. The integration of the password-less authentication asset in the Identity Management and service usage architecture at the high-level can be seen in the below Figure 29.

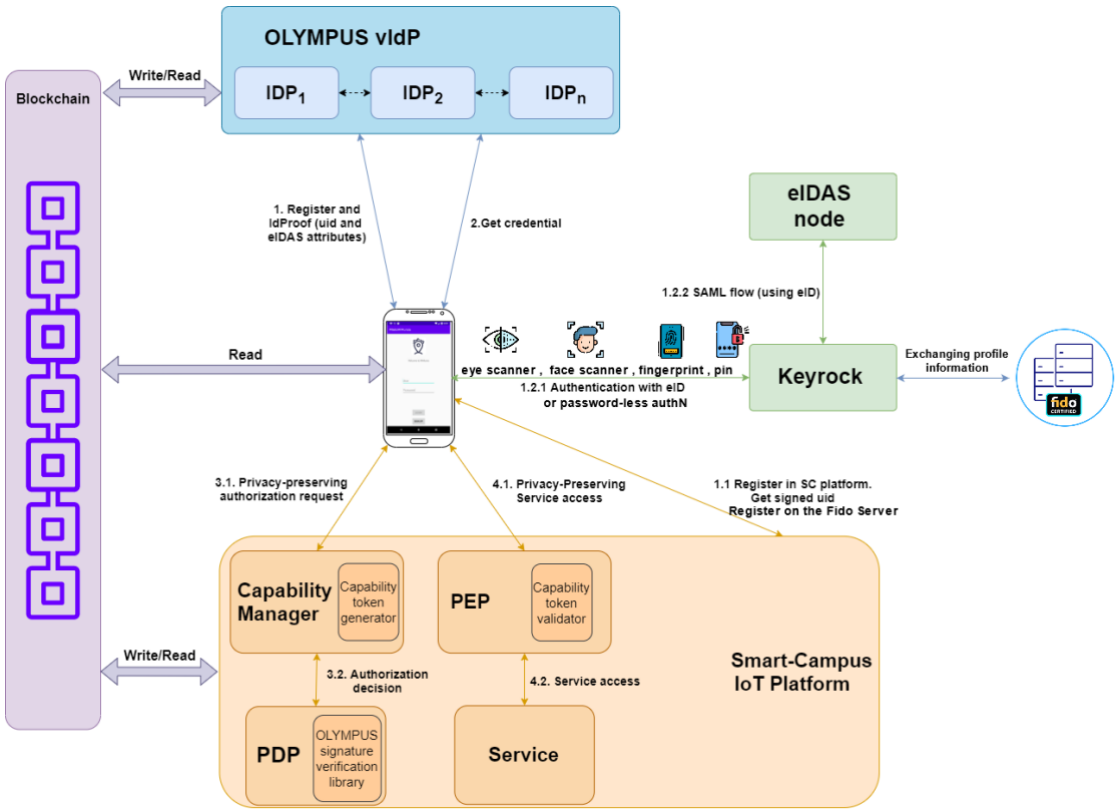


Figure 30: Password-less AuthN in the IdM and service usage scenario

The Identity and Access Management component is directly connected with the Password-less Authentication that FIDO UAF Protocol provides, in order to exchange information regarding the Registration, User Authentication and De-Registration processes.

In the Registration process, after launching the custom IdM Android Application that is based on Keycloak, the user must fill in the necessary personal information and a password (as a backup solution in case of login problems) as shown in the below Figure 30. This step can be skipped, if the admin console of the IdM has already access to the user's personal information.

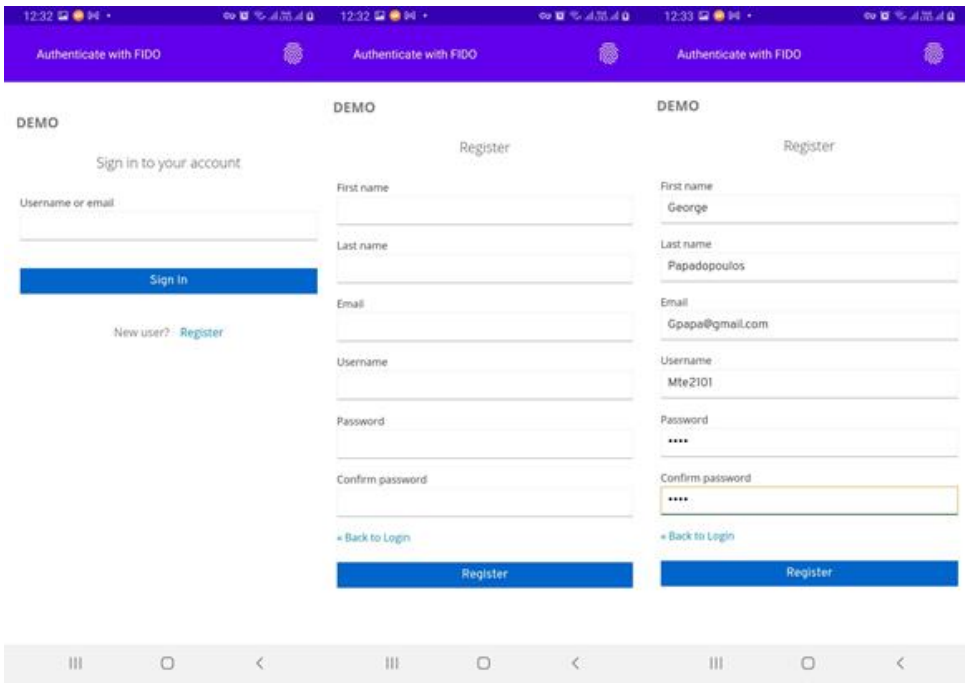


Figure 31: Password-less AuthN User Registration in the IdM App

The Registration process will be completed when the user is redirected to the FIDO UAF Android Application, where he/she will be asked to fill in the username, set the dedicated FIDO Server and tap on the Register button. Then he/she is prompted to confirm his/her identity using Biometrics (or even PIN /Pattern). When the Registration process is completed, the user is provided with a token generated from FIDO and IdM communication, which is necessary for the User's Authentication process.

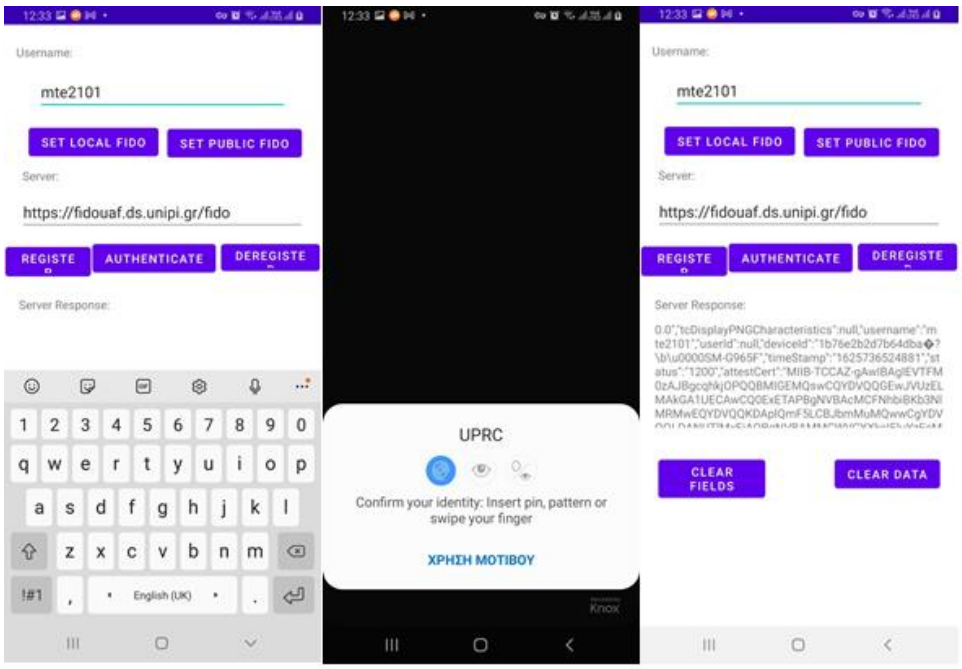


Figure 32: Password-less AuthN User Registration in FIDO App

In the Authentication process, the user only has to press the Authenticate button and then he/she will be asked once again to authenticate using the method chosen in the registration process.

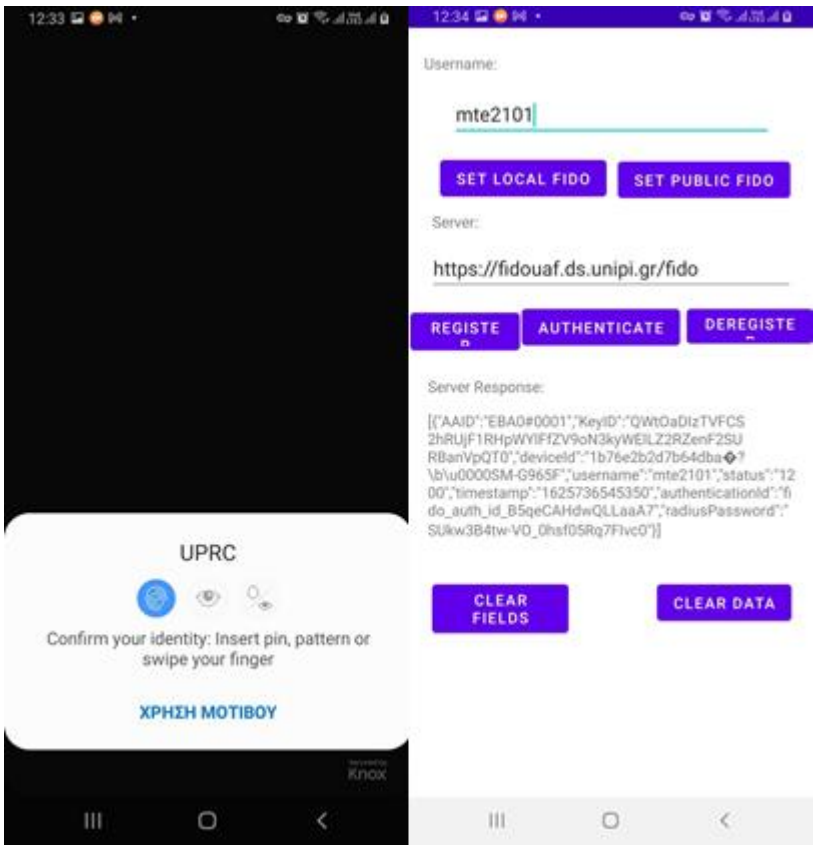


Figure 33: Password-less AuthN User Authentication

After the successful authentication, the user is redirected to the IdM Android Application and is asked to fill in the username or email and press the sign in button to access the service. The token issued in the registration process is used to validate the user's identity.

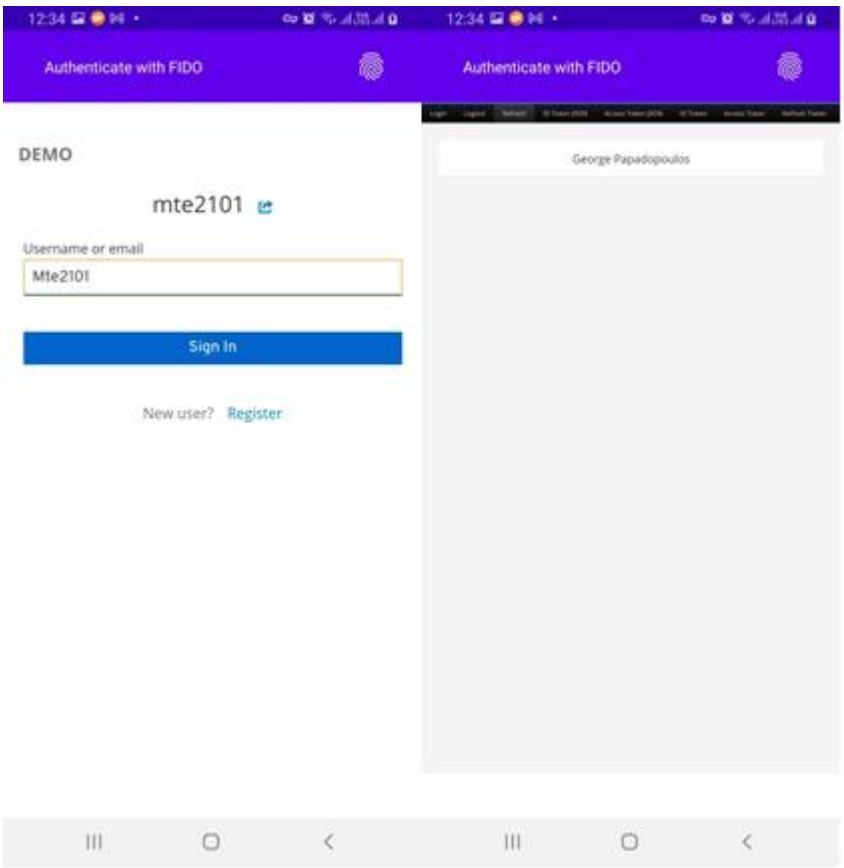


Figure 34: Password-less AuthN User Login

For the De-Registration process the user simply has to: a) fill his/her username, b) set the dedicated FIDO Server, and c) press on the Deregister button. At this point, the user will not have permission to sign in, unless he/she performs the re-registration process again.

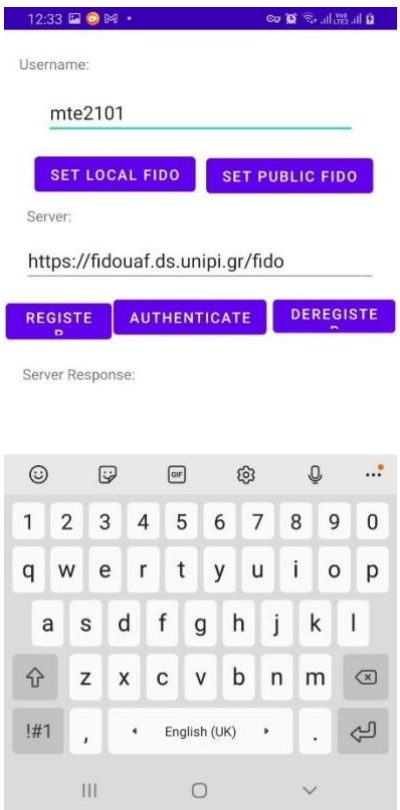


Figure 35: Password-less AuthN De-registration

4.4.4. Future Work

We plan to do further research activities regarding the FIDO protocol which is the backbone of the password-less authentication asset. Furthermore, we intent to employ the asset on several real-life demonstration scenarios in order to enhance the authentication process as well as to extend the applicability of the password-less authentication asset.

4.5. Edge-Privacy

4.5.1. Overview

As the number of IoT devices grows, more people wear or carry devices (e.g., smartwatches) with them and these devices are capable of collecting data about the environment or the users themselves. The Privacy Manager based on Edge Computing is aimed at helping users retain control of the data collected by their IoT devices.

The Privacy Manager [rios2021personal] is devised as a virtualized service that can be deployed in edge-ready devices. It allows data owners to decide under what circumstances the data collected from IoT devices are released to third parties and the level of detail at which these data are shared. The privacy preferences of the data owner are encoded as a set of rules defined in privacy policy files. In essence, the Edge Privacy Manager operates as a privacy proxy between IoT devices and data consumers willing to access their data.

The aforementioned characteristics have led to the architectural design of the Privacy Manager for IoT data, which is depicted in Figure 35. In this architecture, we have several components that enable the collection of IoT data (IoT interface), which will be later stored in a (Data) database, according to the privacy preferences of the data owner (Policies). There is also a component (AuthN) devoted to the authentication and authorization of entities. After successful authentication, a user can query for data, which will be released filtered (Data Filtering) according to the privacy policies of the data owner (Policies). A northbound (Cloud) interface is also included in the architecture to allow the secure storage of historical and backup data in the Cloud. Another component (PMEC interface) is used for the discovery, interaction and negotiation with other privacy manager entities deployed elsewhere controlling access to data from other IoT devices belonging to the user.

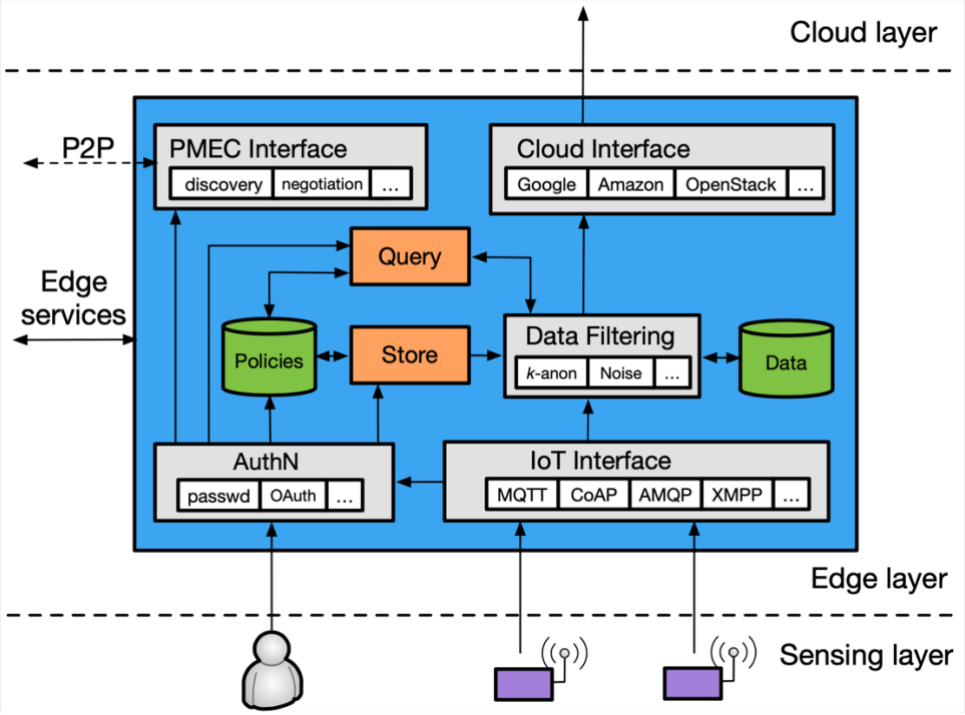


Figure 36: Architecture of the Privacy Manager for IoT data.

4.5.2. Research Challenges addressed

Controlling and limiting access to personal data in a hyperconnected society is an extremely difficult task especially in the context of the Internet of Things, where users own a number of devices that collect and share different types of personally sensitive information.

Developing privacy managers capable of fulfilling such a task is not trivial and several challenges need to be overcome. One of the challenges is related to the availability of the privacy manager, which needs to be accessible at any time in order for data requesters to retrieve information from them. Otherwise, data collected from IoT devices cannot be accessed even if they are functioning properly. The ubiquity of edge infrastructures helps to overcome this challenge.

Another relevant challenge related to the previous has to do with the mobility of IoT devices. The IoT devices a user owns may be geographically distributed across different locations. Moreover, some devices may be carried by the user or have the ability to move by themselves. Therefore, the privacy manager must be able to support the mobility of IoT devices without affecting data availability.

Finally, the privacy manager faces a challenge related to the storage of information. Privacy managers store information to be able to respond to data queries, but when the storage capacity is limited, some of the allocated data needs to be outsourced to an external repository. This implies that the data is distributed among different locations and must still be available to data requesters. When real-time constraints are in place this becomes even more challenging. Our privacy manager takes advantage of the tiered organization of edge infrastructures to store data according to its freshness thus reducing the impact of this challenge.

As a recap from the D3.11 deliverable, we focused on DP-05, the lack of mechanisms for controlling and limiting access to the data collected from numerous and geographically dispersed IoT devices. And DP-08 when uploading information to the cloud the user partially loses control over the data

4.5.3. Demonstration Example

The Privacy Manager can be applied as means for the controlled release of information collected from devices belonging to Smart Campus students, faculty members and staff. One relevant piece of information is the location of these entities while on the Campus.

Smartphones fitted with GPS or other positioning technologies can be used as location data sources. These devices can be configured to upload their data to a Privacy Manager under the control of the user, which has been configured to encrypt location information before storing them in the Edge.

The Privacy Manager can then receive queries from third parties and after checking they are entitled to get access to the location information of the user, they receive it without further processing of the data. The encrypted locations can be later used for computing statistics on mobility patterns, occupation levels, and so on by using secure computing tools like Sharemind (see Section 4.14).

The operation flow of the Privacy Manager is exemplified in Figure 36 and consists of the phases that are defined next.

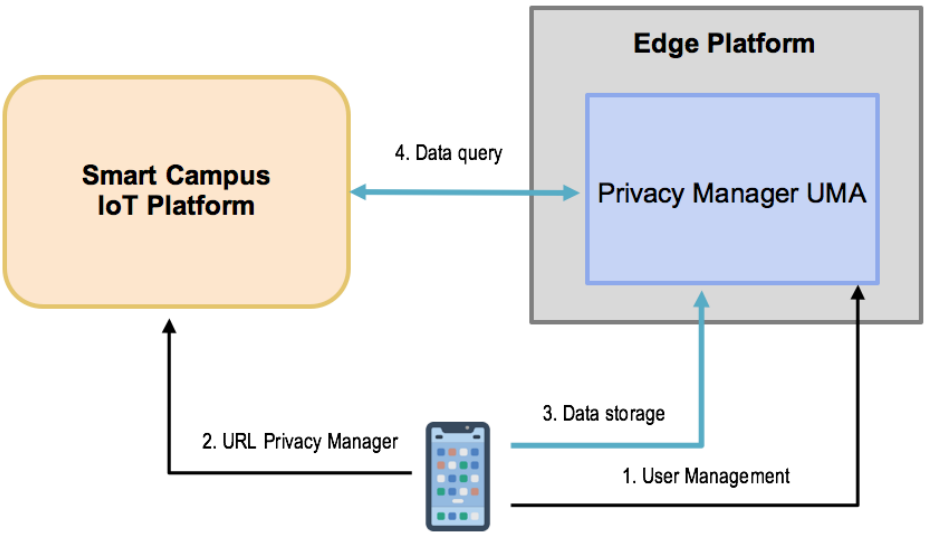


Figure 37: Example Privacy Manager flow

4.5.3.1. Deployment and User Management

Prior to the deployment of the Privacy Manager container into an edge-ready device, it must be pre-loaded with the credentials of the data owner, which will allow him or her to configure the privacy service once deployed. After deployment, the data owner can access the service to generate the structure of the data store and all necessary credentials for allowing IoT devices to store data within the Privacy Manager. The data owner also defines the privacy preferences at this point.

4.5.3.2. Privacy Manager Announcement

Once the Privacy Manager is deployed and properly configured, it is necessary to let interested entities learn how to reach it. Since DNS services geared to IoT networks, such as CoRE or DNS-SD, are not readily available, we opt here for contacting or configuring interested parties. In the case of the IoT devices belonging to the data owner, these are to be configured with the URL where the Privacy Manager will listen for data inputs. In addition to that, the user must first register the type of data his/her devices could provide to the Smart Campus by, for example, filling in an online form stating the URL where the Privacy Manager will respond to queries.

4.5.3.3. IoT Data Storage

After the configuration of IoT devices, they can start pushing data to the Privacy Manager. When data from a device is received, the Privacy Manager checks whether the IoT device is entitled to upload data and whether these data should be processed before being stored into the internal database. The pre-processing of the data will depend on the rules contained in the privacy policy files previously defined by the data owner. These rules can, for example, establish that a particular type of data will be stored in encrypted form.

4.5.3.4. Data Query

The Privacy Manager offers a RESTful API at the previously announced URL for third parties to query for data. When submitting a query, the entity must also present a credential that allows the Privacy Manager to identify the query source and check whether it is entitled to access these data. On top of that, it checks whether the privacy policy establishes the need for further processing of the requested data before it is delivered. For example, if the query asks for highly precise data but the data owner is only willing to share the data after applying some noise.

4.5.4. Future work

We plan to continue researching Edge Privacy in order to improve the stability and usability of our solution. We would also like to model the integration into the asset of new mechanisms for obfuscating data. Additionally, we would like to explore the ability of privacy manager instances to cooperate with one another while reducing the level of trust in these instances.

4.6. Privacy-Aware Aggregate Programming

4.6.1. Overview

Privacy-aware aggregate programming is a programming model centered around privacy protection and aggregate programming. The key concern is the trade-off between data utility and privacy protection. To illustrate this paradigm, consider one of the classical use-cases of aggregate programming for computing a so-called proximity field. The problem in this use-case is a 2D map where several agents, obstacles and locations of interest are spread. The agents want to collaborate in helping each other to find the locations of interest but they do not want to share their actual position with each other or with any central system. The aggregate programming solution to this problem is to build a proximity field: each agent in the map will indicate to closely located agents a notion of proximity. To calculate this field the agents, execute a simple aggregation-based program: iteratively aggregate the neighbors' proximity with the "min" operation, adding an estimate of the neighbor's distance.

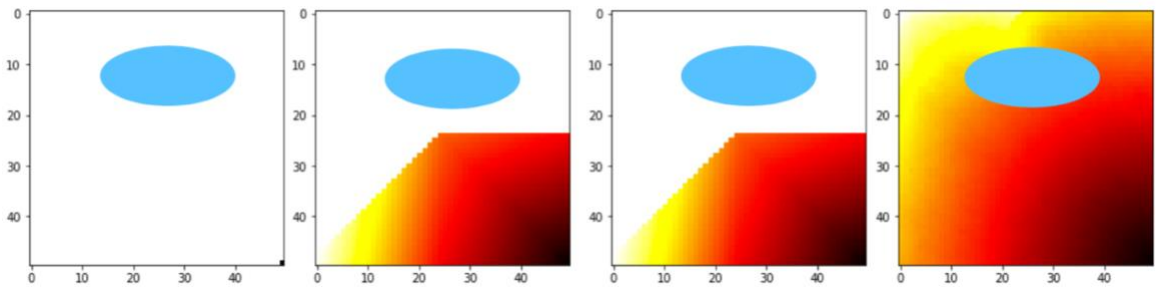


Figure 38: Incremental construction of a proximity field

In Figure 37 we see how the distributed computation of the proximity field evolves as the agents perform more and more rounds of the aggregation. In the Figure, there is only one point of interest (located at the bottom-right corner), one big oval (blue) obstacle, and agents in each of the rest of the coordinates

in the map. The proximity field is indicated with a “heat” gradient that goes from “dark red” (very close) to “yellow” (far away), with a blank indicating unknown proximity. Such a field can then be used by the agents to navigate the map and reach the location of interest.

This asset is in a preliminary proof-of-concept stage and, as such, there is still not a proper tool supporting the approach. The current status is based on a prototype implementation that allows for simulation of proximity field constructions with distinct noise-based privacy protecting techniques (à la differential privacy). The following sections provide more details on this and on possible future work to construct an effective tool supporting this research asset.

4.6.2. Research challenges addressed

Aggregate programming is a programming model that has “aggregation” as one of its key primitives. Aggregate programming methods are crucial in several areas, from traditional areas such as data analytics to more recent areas such as edge computing [roberto2021Eng]. A well-known example of aggregate programming languages is database query languages, which provide data aggregation operations (sum, count, etc.) as one of their main programming primitives. More sophisticated examples of aggregate programming can be found in the areas of multi-agent systems, for example in bio-inspired approaches to decentralized coordination based on so-called computational fields with related programming languages such as ScaFi (https://scafi.github.io), a Scala-based incarnation of the Field Calculus [viroli2013advance].

Aggregation is also a popular technique for providing specific privacy guarantees by putting together individual pieces of information and hence making it more difficult to distinguish their individual contribution and, ultimately, minimizing privacy leaks. Basic aggregation operations, however, do not offer sufficient privacy guarantees against sophisticated attacks privacy based on statistics. This requires the use of privacy-protection techniques like anonymization and noise-addition techniques such as differential privacy [dwork2013alg]. Privacy-aware aggregate programming aims at enriching the aggregate programming paradigm with privacy aspects so to facilitate the design, development and analysis of privacy-respecting software systems. A prior work in this area was presented in [kaminskas2018agg] where the authors advocated for data items to be shared with aggregation policies related to privacy protection so to empower individual data providers with the ability to control how their data is aggregated with other data and enforce the use of privacy protecting techniques. Here, one of our aims is to address privacy challenges identified in D3.11, in particular DP-05 “Lack of mechanisms for controlling and limiting access to the data collected from numerous and geographically disperse IoT devices”, by promoting privacy-aware aggregation in a distributed way and tools to assess impacts in utility and privacy risks.

4.6.3. Demonstrations Example

In this demonstrator, we illustrated how aggregate programming techniques can be used to provide privacy-protecting services to users and campus administrators in a way that we can also simulate and tune the effect of privacy-protecting measures (e.g., based on noise addition) so as to find the desired compromises between service utility and user privacy.

We consider a scenario where aggregate programming is used to handle the emergency on the campus. We consider a scenario where an event (possibly related to the emergency) causes a disruption in the CCTV system, leaving the users and their smartphones as the only available (decentralized) platform to detect, find and deal with the emergency.

In this scenario, two conflicting goals come into play. Since proximity does not provide full privacy guarantees (knowing the actual distance from an agent to a location of interest can be used to reveal its

actual placement and perhaps other sensitive information) users may want to introduce disturbances in the shared proximity information in a similar fashion as done for example by Apple privacy-protecting services or by CyberSec4Europe assets such as SOBEK. On the other hand, users wish to collectively build a field that is as accurate as possible to be able to efficiently deal with the emergency.

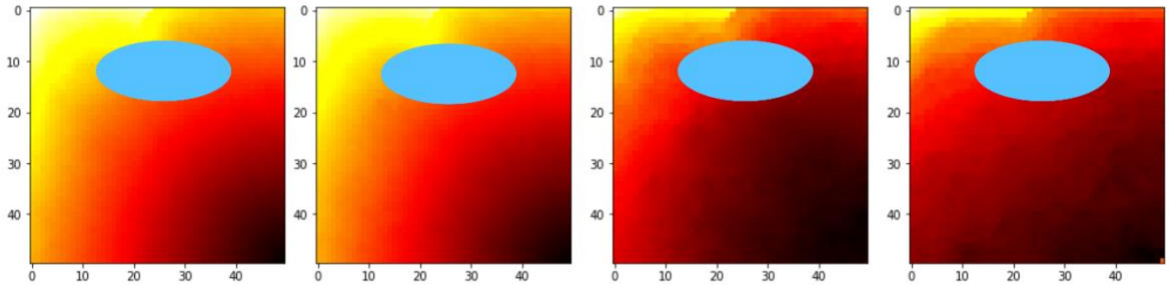


Figure 39: Proximity fields with increasing amount of privacy-protecting noise

Figure 38 shows the effect of noise addition in the field as in differential privacy techniques (more concretely we are adding Gaussian noise for varying mean deviation), while Figure 39 shows the result of simulating an agent (say a police agent) navigating the map towards the place where the emergency needs to be addressed. More noisy privacy-protecting fields require more steps to reach the location of the emergency.

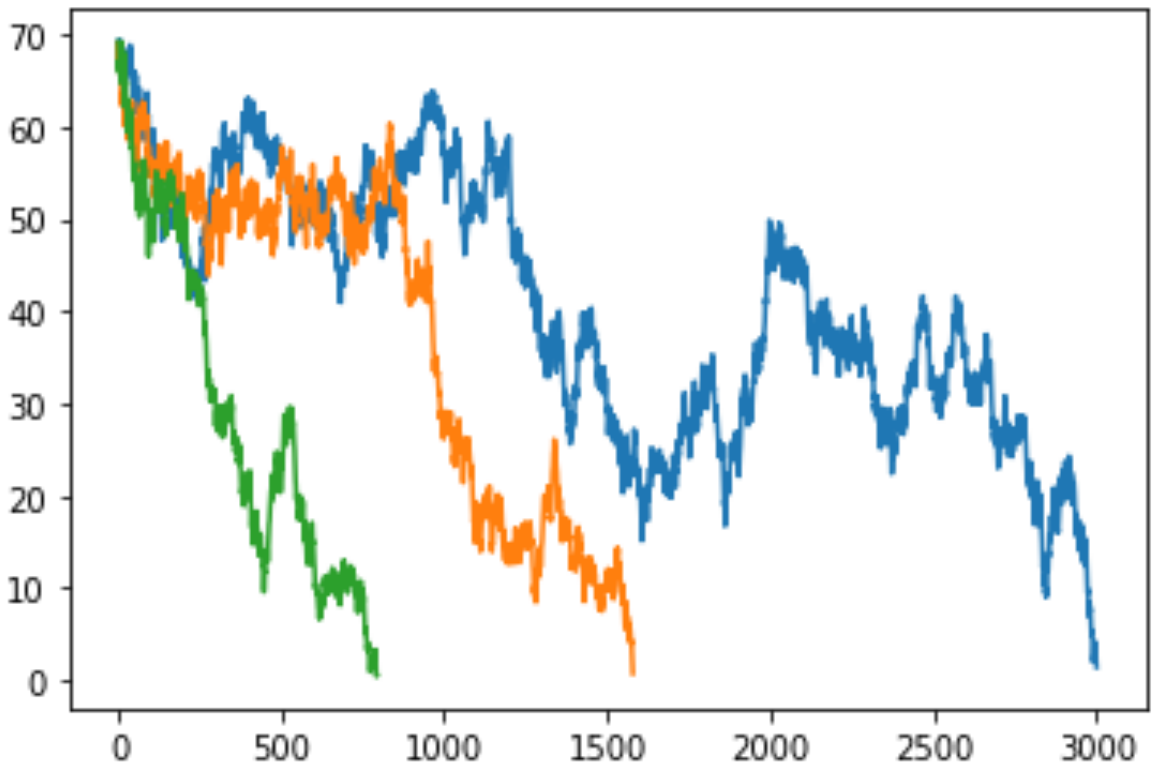


Figure 40: Time steps (x axes) to reach the emergency location. The y axis indicates proximity to the emergency location. The colored plots correspond to proximity fields with varying amounts of privacy-protecting noise.

4.6.4. Future Work

A possible next step could be the development of an effective tool to support the approach described in this section. The tool could allow for assessing the impact of privacy-protecting techniques in the tradeoff between utility and privacy risks. This would require first a deeper investigation of privacy risks, e.g., related to inference of actual locations based on proximities, and on utility measures for

proximity fields. Another long-term future work would be to extend the ideas to applications beyond proximity fields.

4.7.DANS

4.7.1. Overview

The main aim of the Data Anonymization Service (DANS) tool is data protection, namely preserving personal data privacy. Considering regulatory aspects anonymized data are excluded from GDPR regulation because anonymized data is no longer "personal data". In this way, the DANS asset is an anonymization tool that avoids user tracking and user re-identification by the use of privacy and risk models which prevents privacy threats when data are managed. As perfect anonymization is not possible it is necessary to balance between privacy and data accuracy for analytics. The DANS is managing the user attributes in different ways, to be delivered to the data consumers:

- Identifying attributes which will be removed from the dataset.
- Quasi-identifying attributes which will be transformed accordingly with the specified transformation procedure.
- Sensitive attributes will be kept as-is, but they can be protected using privacy models, such as t-closeness or l-diversity.
- Insensitive attributes which will be kept unmodified.

In order to preserve user data privacy, the DANS tool provides a K-anonymity privacy model for quasi-identifying attributes (as this model is the most used for health data) and also provides an L-diversity model for sensitive attributes.

This anonymization tool is provided in two flavors:

- As a service to be deployed on the data provider premises or a third party.
- As a java library to be integrated into the data provider system.

4.7.1.1. Integration in the Smart-Campus scenario

According to the scenarios provided in section 4, the DANS tool could be integrated into the Geolocation scenario. As geolocation data can be considered as quasi-identified attributes DANS would anonymize this kind of data in order to preserve the user privacy. This asset is integrated into the Common T.3.2 Architecture (Section 2) in the User domain and the IoT domain.

4.7.1.2. Asset Evaluation results

The DANS asset has been evaluated through the Medical Data Exchange demonstration in the context of Task 5.6 [Sforzin A. CyberSec4Europe. D5.3 Validation of Demonstration Case Phase 1. 2021 https://cybersec4europe.eu/wp-content/uploads/2021/02/D5.3-Validation-Demonstration-Case-Phase-1-v1.0-submitted.pdf]. The evaluation of this asset in the context of the Smart-Campus scenario could be done with the same approach. Instead of anonymizing COVID-19 data used in the Medical Data Exchange demonstrator, the geolocation data provided by the wearables or smartphones of university staff would be used.

4.7.2. Research Challenges addressed

Electronic health records gathered by hospitals and health organizations contain sensitive information. The aggregation of these data is an invaluable aid for preventing diseases, take accurate medical decisions and research purposes. These sensitive raw datasets cannot be directly shared and need to be protected in order to preserve the individual's data privacy. With the aim to avoid a breach of data privacy when these data are delivered outside the medical institutions, anonymization techniques have

been usually applied. De-identification of health records by applying the k-anonymity privacy model and using generalization techniques. With these techniques, identifiers such as name or unique identifiers are removed; quasi-identifiers such as sex, postcode or age are generalized; the sensitive data such as diagnosis are not modified for applying following analytics.

The application of these techniques provokes a loss of information, but it is necessary that the resulting anonymized dataset keeps the utility for subsequent analysis to be performed by the recipients of the data.

The anonymization approach developed by the DANS tool aims to provide a trade-off between privacy-preserving and utility.
We can map this general discussion about challenges to the ones defined in D3.11. Concretely, the asset addresses the following challenges:

- DP-07 by providing anonymization methods that detect and protect personal and sensitive identifiable information when data are managed out of the user's control.
- IDP-01 by providing models that inform about the re-identification risks.
- IDP-04 by transforming medical data the ability to obtain information beyond the necessary is limited.
- LDP-01 is partially covered, as DANS is providing a statistical report regarding the anonymization process.

4.7.3. Demonstrations Example

Once the geolocation data from the devices of university staff are collected and uploaded to the Edge Platform, the Privacy Manager could anonymize the geolocation data in two ways depending on how the DANS asset is used by the Edge platform either as a service or as an embedded library. The DANS asset can be deployed as a service providing the anonymization functionality on the service provider premises, in this case, the Edge Privacy Manager will communicate with the DANS service in order to anonymize the data location, as depicted in Figure 40 (a). The anonymization functionalities can also be embedded into the Privacy Manager component by using the DANS java library for anonymizing data location, as shown in Figure 40. (b). In both cases, the anonymized data can be used for analytics.

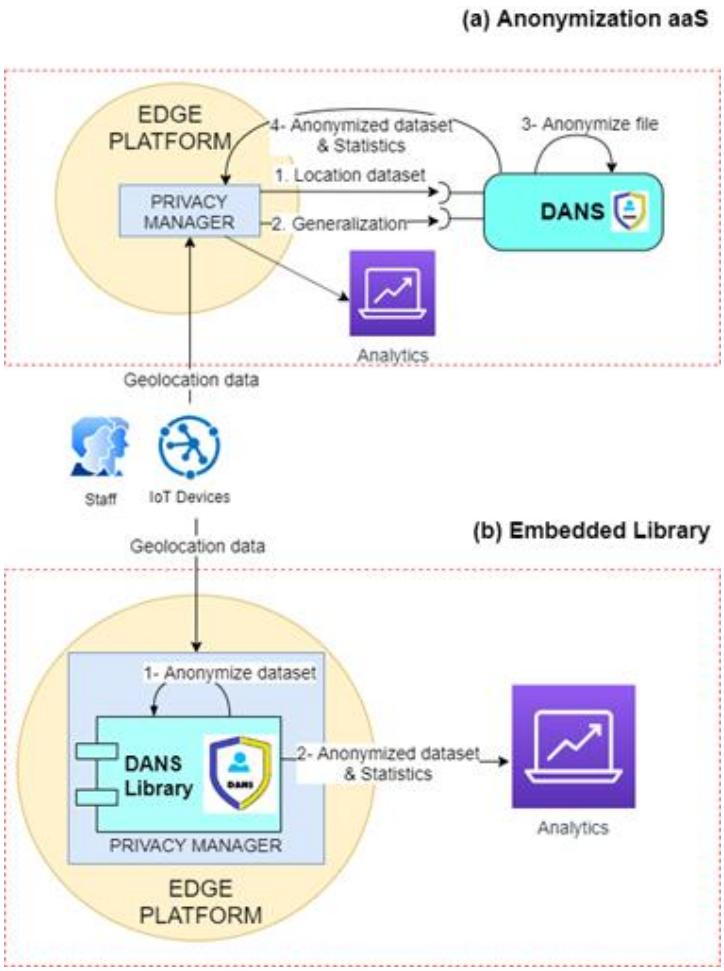


Figure 41: Flavours of DANS tool (a) Anonymisation as a Service, (b) Embedded library

4.7.4. Future Work

Within the context of this project, there will be no further development of DANS. The focus will be the integration in the WP5.

4.8. Cryptovault

4.8.1. Overview

CryptoVault is a system intended for users of different blockchain technologies. It comprises a hardware wallet that securely stores and manages the sensitive user keys, and a reliable method for backing up these keys as independent shares stored in multiple locations. The aim is to combine the best features of different wallet types while minimizing the risks related to these wallets. CryptoVault generates keys with high entropy, offers end-to-end protected key backup, signs transactions inside a trusted execution environment and can run key recovery without a single point of failure.

There are three security technologies used in the CryptoVault scheme: splitting keys into shares that are stored separately for secure backup purposes (on paper or on remote servers), isolating processing and storage from the main operating system by using the Intel SGX environment and using RSA to establish an end-to-end encrypted information channel between the SGX-enclave and remote servers when distributing the key shares.

In the identity management and service usage scenario, there are various options for using blockchain or smart contracts. CryptoVault can be integrated into this scenario as a tool for IT administrators to secure and backup their cryptographic keys.

4.8.2. Research Challenges Addressed

One of the main selling points of blockchain technologies is their decentralized nature; there is no need for a trusted third party to check transactions. The drawback of the blockchain approach is the hassle of key management. There is no "blockchain support personnel" that can reset your password or create new credentials in case you lose the originals. Often the solution is to use a wallet application, which can be either software, hardware, or a combination of the two. Wallets store the private keys, generate signatures, and encode the transactions on behalf of the user.

There are several types of wallet applications [suratkar2020crypto]: Online wallets can be used with any web browser; hence they are easy to use, but require an internet connection, a secure browsing setup and placing trust on a third party once again. Mobile wallets are similarly easy to use and enable storing the keys on your personal device, but they require good overall security for the device as well as careful use of that device from the user. Desktop wallets that operate on a personal computer are slightly less convenient to use, and not all wallets are available for all operating systems. The security of the desktop environment is integral, and the keys are likely to be hosted locally. The user probably uses the wallet in a safe physical environment, reducing the risk of shoulder surfing attacks, thefts, and lost or broken devices. Finally, there are hardware wallets, which are dedicated devices for hosting and operating with user keys. These are significantly less convenient to use, but the security is increased as the device is not used for other purposes and the user is in full control.

There is one issue that is common to all the different wallet types: the loss of keys or the destruction of the device containing the keys. If that happens, the user will be unable to access their cryptocurrencies or other blockchain services. Again, we come to the original problem of trusting a third party, e.g., a wallet operator, to have a secure key recovery protocol in place, which is what the blockchain was supposed to mitigate. The CryptoVault[niko2021sec] solution provides a key backup and recovery method that is under the user's control. The problem of a single point of failure with backing up a key to a different location, most probably controlled by a third party, is solved by using a secret sharing method and storing the shares separately. A highly secure but still accessible wallet is achieved by using a trusted execution environment on a regular, all-purpose computer.

By using the secret sharing method we can also address privacy challenge DP-08 from D3.11: When uploading information to the cloud the user partially loses control over the data. When using the CryptoVault system, the user is the only one who knows where all the key shares are stored, and the one share uploaded to a third-party cloud server is not enough to betray any information about the user's secret key.

4.8.3. Demonstrations Example

In the Smart Campus scenario, we assume that there is an IT administrator that uses a blockchain-based tool for their job, and they need to keep the private key secure. They cannot place the responsibility of controlling the key for a third party, so they choose to use either a desktop or a hardware wallet. Finally, they also need a way to restore their access to the blockchain application in case the private key is lost, e.g., due to hardware failure or a human mistake.

The CryptoVault implementation secures Ethereum keys. Ethereum uses the secp256k1 elliptic curve for digital signatures and in this implementation the secret key is a randomly selected positive integer below constant secp256k1n - 1. The corresponding public key is calculated by elliptic curve multiplication with secp256k1 generator point. Finally, the Ethereum address can be derived by calculating the keccak256-hash from the formatted public key. The secret key, public key and address are generated inside an enclave, and the address is stored in plain text into the file system. This sequence is illustrated in Figure 42 below.

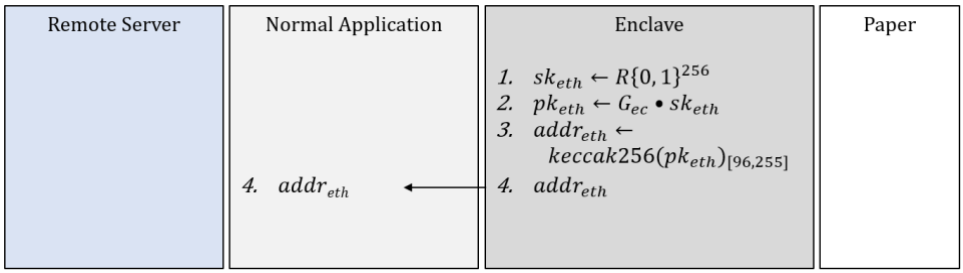


Figure 42: Key generation phase, where sk and pk are the secret and public keys and addr is the address.

The next step after calculating the Ethereum address is to make a backup of the secret key: the key can be divided into two or more shares with Shamir's Secret Sharing scheme. A user can use paper to save one share physically, and several remote servers can be used for the other shares. The user chooses a threshold value for how many shares out of the total number are needed to reconstruct the original key. We assume that the required minimum number of shares is available for recovery when needed, and that the number of malicious actors is smaller than the required minimum number of shares, thus ensuring that the secret cannot be recovered outside of the system.

A remote server needs to generate an RSA-2048 keypair. The public key is transferred into the Enclave and used to encrypt a key share. The encrypted share can now be transferred to the remote server. These steps are illustrated in Figure 43. The recovery process of the secret key is performed in reverse order. The share that was stored in the remote server is decrypted and then encrypted again with the public key retrieved from the enclave. The original secret key is recovered when Shamir's Secret Sharing Scheme is applied in reverse to the collected shares. This is depicted in Figure 44.

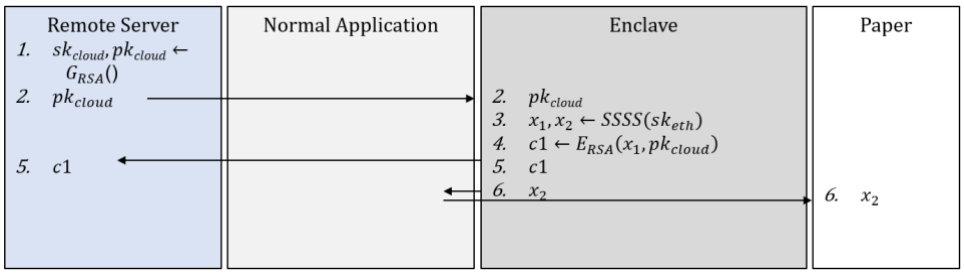


Figure 43: Key backup procedure. All participating servers need to generate (G) RSA key pairs for this scheme. The secret Ethereum key is split into shares (x) using Shamir's Secret Sharing Scheme (SSSS) and then remote shares are encrypted (E) with the server's

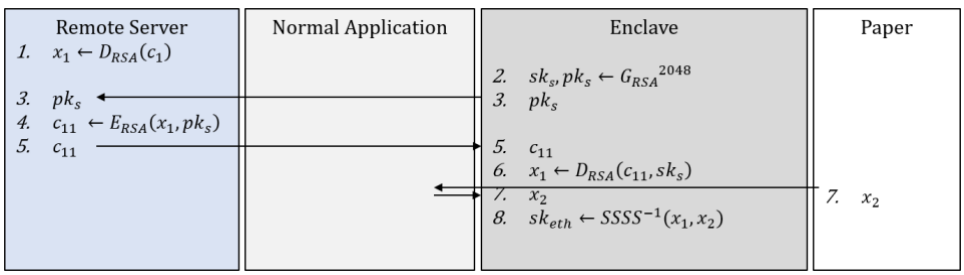


Figure 44: Recovering the key. The ciphertext (c) of the remote share needs to be decrypted (D) before encrypting it again with the enclave's public key for transfer. When enough shares are gathered, the enclave can reverse the secret sharing scheme to recover them.

For a practical demonstration, this asset can be run on a laptop that supports the Intel SGX environment. Signing an Ethereum transaction hash, creating a backup and recovery of a backup can be demonstrated.

4.8.4. Future Work

Within the context of this project, there will be no further development of CryptoVault. From a broader perspective, a new concept of social wallets has emerged as a way for users to back up their blockchain keys within their social circle. This is an interesting potential application for CryptoVault and for further research in general.

4.9. Elastic Deployment of TEE-based applications in the cloud

4.9.1. Overview

Processing a large amount of sensitive data requires secure and scalable solutions. For example, location data – as the one processed in the geolocation scenario - is considered as a privacy-sensitive data type so that the platform processing the data should be secured to minimize data leakage. At the same time, the platform should allow for dynamic resource allocations so to cope with different amounts of data to be processed. ReplicaTEE is a solution that enables dynamic enclave replication and de-commissioning for TEE-based applications in the cloud. In particular, ReplicaTEE is designed to enable replication of applications that use Intel SGX – arguably the most popular TEE for workstations – as the undelaying TEE. Given the current deployment model of Intel SGX enclaves, replication of an enclave across machines requires the application owner to either be always online so to provision secret material to newly deployed enclaves, or to trust the cloud provider with managing the enclave secrets. An always-online application owner reduces the benefit of outsourcing to the cloud; trusting the cloud provider with managing enclave secrets voids the advantages of using a TEE.

4.9.2. Research challenges addressed

Regarding enclave migration, moving the internal state of an enclave to another platform is at odds with the main security provisions of SGX – the state of an enclave is private to the enclave itself. Furthermore, when the state is persisted to disk, it can only be recovered by that same enclave running on that same platform.

The only proposal for enclave migration that we are aware of, is [gu2017secure, guerreiro2020Tee]. It leverages an SDK that augments enclaves with a thread dedicated to migration. This thread simply transfers the internal state of the source enclave to the matching thread of the destination enclave. The authors of [gu2017secure] also point out that some data structures that must be migrated are not available to the enclave. This is the case of the CSSA – a data structure that handles the nesting level of enclave exceptions. The solution proposed in [gu2017secure] is to infer this value by monitoring the behavior of the enclave and to rely on the untrusted OS to recreate the same conditions at the target platform. The authors show that their heuristics is indeed effective in few application scenarios. However, the effectiveness of this heuristic for general SGX applications remains to be assessed. The authors of [guerreiro2020Tee] propose an enclave migration framework that uses hardware security modules. The result is an increased trusted computing base, given the requirement to trust the SGX hardware and the hardware security module.

With respect to the challenges laid out in D3.11, this asset tackles challenge DP-03 and DP-05.

4.9.3. Demonstrations Example

ReplicaTEE leverages a distributed SGX-based service layer that interfaces with a Byzantine Fault-Tolerant (BFT) storage layer to orchestrate secure and dynamic enclave replication in the cloud. Enclave developers entrust application secrets to the service layer and can go offline. The service layer is a thin software layer that runs in SGX and handles the commissioning and de-commissioning of enclave replicas on behalf of the enclave owner. Application secrets are, therefore, shielded away from malware that penetrates the cloud, as they are securely transferred from the application owner to the service layer onto application enclaves. The service layer also controls the number of running replicas for a given application. To prevent attacks to the service layer itself, ReplicaTEE uses a distributed BFT storage layer that guarantees dependable storage despite the compromise of a fraction of its nodes.

ReplicaTEE is fully compliant with the existing Intel SGX SDK. A prototype implementation of ReplicaTEE increases the size of the Trusted Computing Base by approximately 800 Lines of Code

(LoC). Its deployment in a realistic cloud environment shows that ReplicaTEE does not add significant overhead to existing SGX-based applications.

ReplicaTEE augments the cloud software stack with a layer named Enclave Management Layer (EML), dedicated to elastic enclave provisioning. EML oversees provisioning and decommissioning enclaves on behalf of application owners. EML is designed to run entirely in SGX so that (i) application owners can verify its code, and (ii) sensitive data entrusted by application owners to EML is shielded by any other software running on the same host.

EML is distributed across enclaves and leverages a master-slave approach to ensure progress despite potential crashes. Since EML itself may be a victim of attacks, we couple it with a BFT Storage Layer (BFS) that provides consistent storage despite Byzantine faults of a fraction of its nodes. EML uses BFS to always maintain a consistent view of the requests to provision/remove enclaves and the progress it has made to handle such requests. This design allows us to prevent attacks on EML while, at the same time, keeping the codebase of the provisioning service small enough to be run entirely in an enclave. By coupling a lightweight management layer such as EML and a BFT storage layer such as BFS, we enable the cloud to dynamically provision enclaves to applications, while ensuring protection against forking attacks.

Our solution is depicted in Figure 44. In a nutshell, application owners entrust the cloud provider with the application code, and EML with the secret material that the application needs to run (e.g., a secret key).

When a new application enclave must be provisioned, EML acts on behalf of the application owner and ensures that (i) the deployment of the new enclave does not violate the policy defined by the application owner, (ii) the application code runs in an enclave on an SGX-enabled platform, and that (iii) the enclave is provisioned with the appropriate secret key, if required. When dealing with enclave decommissioning, we note that one cannot tell whether an enclave has been properly shut down or whether its messages are being blocked. To solve this issue, each application enclave is granted a lease upon provisioning. That is, when EML provisions an application enclave, it also provides an ``end-of-lease" timestamp. The application enclave should run until the lease expires unless the lease is otherwise renewed.

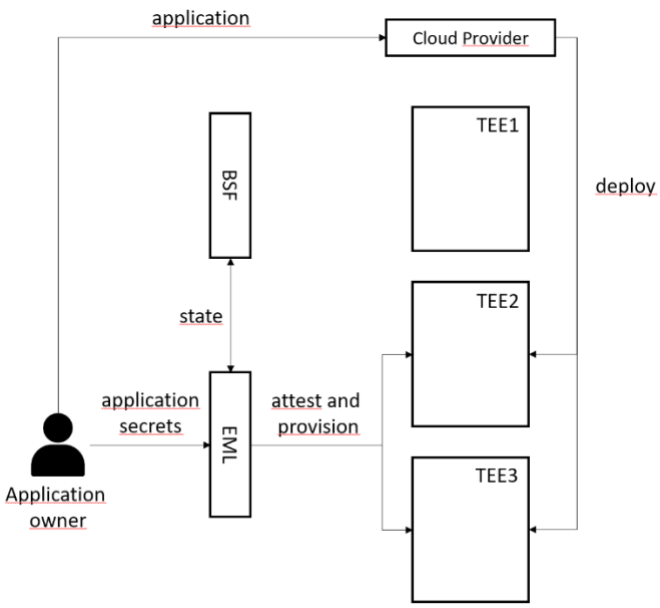


Figure 45: Backdoor-resistant setup

ReplicaTEE can be used to analyze privacy-sensitive data, for example, in the geolocation scenario. In particular, messages produced by IoTs, cameras and sensor devices may carry sensitive information and, as such, techniques are required to balance the privacy of the individuals providing the data and the utility of the applications harvesting such data. A Trusted Execution Environment can provide such a balance, but in the case of applications with dynamic workloads, a TEE may not provide adequate computing power as the computing platform must dynamically allocate and remove computing instances to cope with the current load. In this context, ReplicaTEE can provide a dynamic allocation of enclaves to process data in a privacy-preserving manner. The data processing logic is deployed in an enclave so that raw data is never exposed in the clear and only aggregate statistics are available. Further, ReplicaTEE enables the computing platform to adjust the computing resources to the current load so as to balance performance and operational costs.

We have built a prototype of ReplicaTEE including the EML and the storage later. We used our prototype to evaluate its performance in a close-to-realistic scenario. In particular, we deployed the storage service of ReplicaTEE on five identical servers with SGX supports. Each server is equipped with Intel Xeon E3-1240 V5 (8 vCores @3.50GHz) and 32 GiB RAM. The EML instances were deployed on a machine with Intel Core i5-6500 (4 Cores @3.20GHz) and 8 GiB RAM. All these machines are equipped with SGX to run enclaves and are connected with a 1 Gbps switch in a private LAN network. We argue that this setting emulates a realistic cloud deployment scenario where the compute servers and their corresponding storage servers communicate over the cloud's private LAN (e.g., Amazon AWS and S3).

Figure 45 and Figure 46 show throughput vs latency for the enclave provisioning process given different storage failure threshold f (i.e., the number of benign or malicious crashes that the storage system can tolerate). We see that when f=1 (3 storage servers), the system achieves a peak throughput of 85 op/s with a latency of 270 ms. On the other hand, when f=2 (5 storage servers), the latency remains almost the same, while the peak throughput is reduced to 75 op/s. Our findings suggest that the remote attestation process is the dominant factor in the operation latency. Notice that even if increasing the fault-tolerance threshold of the storage service reduces the peak throughput (since it requires more communication rounds), it has a limited impact on the witnessed latency.

In Figure 47, we further measure the constituent latencies incurred in the enclave provisioning process. In both cases when f=1 and f=2, we see that the time for remote attestation is around 260 ms while the state update only takes 10 ms without a noticeable difference in either case. Namely, the state update only comprises up to 3.7% of the whole provision process even when f=2.

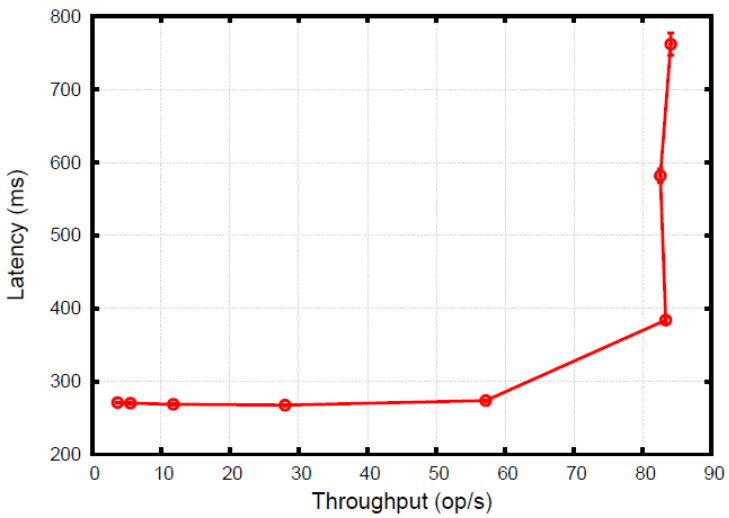
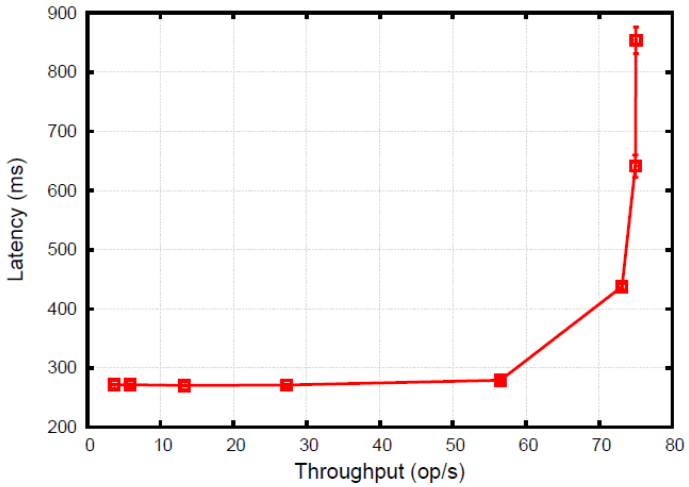


Figure 46: Performance of the solution



(b) Throughput vs. latency for enclave provisioning when $f = 2$.

Figure 47: Performance of the solution

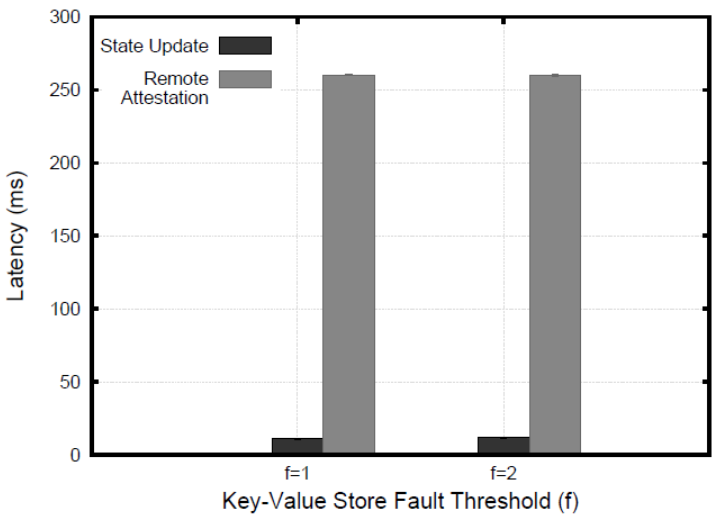


Figure 48: Fault Threshold

4.9.4. Future Work

We plan to complete most of the research and development work on this asset within the lifetime of the project. In the future, we plan to investigate the alternative mechanism for controlled and scalable distributed TEE application in the cloud.

4.10. Backdoor-resistant TEEs

4.10.1. Overview

SR-EPID is a subversion resilient version of Enhanced Privacy ID, the protocol used by Intel SGX for remote attestation. In the context of Trusted Execution Environments, remote attestation enables a party to establish trust in a TEE running on a remote platform. In the context of the demonstrator described earlier in this document, SR-EPID may be used in the geolocation scenario so that the integrity of the

devices where geolocation data is going to be processed, is verified before such data is uploaded. All the popular remote attestation protocols balance authenticity and privacy by using group signature schemes. The latter is a digital signature scheme that allows a verifier to verify a signature as issued by a member of a trusted set while keeping the signer itself anonymous (within that set). In the context of Intel SGX, remote attestation uses a special system enclave, named Quoting Enclave, that certifies the application enclaves running on the same platform. Certification is realized by signing a report with the identity of the application enclave being certified. The group signature scheme being used – EPID – allows a party to verify that the report was issued by a genuine Intel SGX platform, without revealing any other information about the issuing platform, i.e., keeping the platform anonymous. Given the embedded nature of TEEs and the increasing concern on state-level adversaries, the scientific community is designing cryptographic protocols that can withstand subverted parties. In the context of remote attestation, a subverted signer may exfiltrate, through innocent-looking signatures, identifying information or even the signing key. Thus, an adversary can either identify the signer, thereby breaking anonymity, or obtain the signing key, thereby breaking the unforgeability of signatures.

4.10.2. Research challenges addressed

Previous work [jan2017ano] has shown how to obtain "subversion-resilience" for Direct Anonymous Attestation – the remote attestation protocol used for the Trusted Platform Module (TPM). The subversion-resilient version of DDA uses a split signature scheme where the secret key is split between the TPM and the host. Intuitively, even if the TPM is subverted and its share of the secret known to the adversary, security is guaranteed by the fact that the host holds the remaining share. Our research challenge lies in designing a subversion-resilient attestation mechanism for SGX that requires no secret (or secret share) on the hosting platform.

With respect to the challenges laid out in D3.11, this asset tackles challenge IDP-04.

4.10.3. Demonstrations Example

Back-door resistant TEE could be used in the geolocation scenario to ensure that devices that process geolocation data have not been tampered with before data is uploaded. SR-EPID provides subversion resiliency to the remote attestation protocol used by Intel SGX. In order to counter subverted signers, our main idea is to enhance the EPID model by adding a "sanitizer" party whose goal is to ensure that no covert channel is established between a potentially subverted signer and external adversaries. The role of the sanitizer is shown in Figure 48. In practical application scenarios, the sanitizer could run on the same host of the signer (e.g., on a phone to sanitize signatures issued by the SIM card), or on a separate one (e.g., on a corporate firewall to sanitize signatures issued by local machines). Compared to a subversion-resilient anonymous attestation scheme that uses split-signatures, our approach comes with multiple benefits. First, signature generation is non-interactive, and the communication flow is unidirectional from the signer to the sanitizer, on to the verifier. Thus, our design decreases signing latency and provides more flexibility as the sanitization of a signature does not need to be done online. Another benefit of our design is the fact that the sanitizer holds no secret. This means that if a memory leak occurs on the sanitizer, one has nothing to recover but public information. Differently, in a split signature approach, security properties no longer hold if the TPM is subverted, and the key share of its host is leaked. Further, as sanitization is non-interactive and requires no secret, it may even be carried out by multiple parties in a cascade fashion so that covert channels are eradicated as long as one of the sanitizers has access to a good source of randomness – and such randomness is not available to the adversary. It is not clear how to achieve such "fault tolerance" with split signatures. One may split the signing key across several parties and design a multiparty signing protocol, but very likely this would lead to high latency for signature generation.

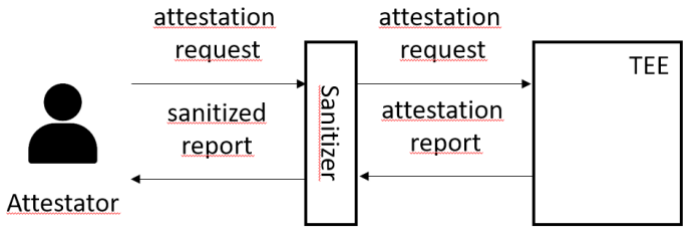


Figure 49: Backdoor-resistant TEEs

Our contributions include a new formalization of EPID to accommodate for potentially subverted signers and an instantiation of SR-EPID based on bilinear pairings. The resulting scheme provides the same functionality of EPID, tolerates subverted signers, and features signatures that are shorter than the ones in the original EPID proposal for reasonable sizes of the revocation list: ours have 28 + 2n group elements whereas EPID signatures have 8 + 5n, where n is the size of the revocation list (i.e., ours are shorter already for n≥7).

4.10.4. Future Work

Future work includes novel mechanisms for revocation of attestation keys so as to blacklist rouge TEEs. Revocation is one of the main challenges in attestation protocols and efficient, reliable revocation mechanisms are required.

4.11. Privacy-Preserving for Genomic Data (PP4Genomic)

In this asset, we presented a novel filtering approach to detect sensitive nucleotides in long reads, which are now produced by the newest sequencing technologies. In this context, our asset has focused on read filtering, i.e., the early detection of sensitive nucleotides in reads at the mouth of, and possibly inside, the sequencing machines. Requirements for an early read filtering method include i) to detect a maximum of nucleotide deemed sensitive; ii) not to be a throughput bottleneck (i.e., it must filter reads faster than the sequencing machine produces them); and iii) to be practical (i.e., it can be implemented in, or close to, a sequencing machine, with limited hardware resources). PP4Genomic demo can fulfil the three requirements that we initially set. More precisely, it is able to detect almost all of the sensitive nucleotides, the filter is faster than the sequencing speed of a modern next generation sequencing (NGS) and the hardware is as compact as a smartphone.

4.11.1. Overview

Context

The high throughput of Next Generation Sequencing (NGS) technologies coupled with lower genomic sequencing costs have enabled the creation of numerous genetic biobanks that collectively amount to gigantic genomic data volumes. Due to the interesting properties of genomic data, the scientific community has entered a new genomic era, which has been contributing to the flourishing of the biomedicine and -omics fields, and to the development of personalized medicine. More precisely, the rapid evolution of sequencing technologies promoted a variety of opportunities for genomic data, which have introduced significant performance and privacy challenges. PP4Genomic introduces the approach our research group has been following, in order to protect users' privacy at a low performance cost, with a special focus on the early stages of genomic data processing, which are routinely executed in public clouds. The keystone of our asset is a 'privacy filter' [j2018accurate], which detects sensitive information (i.e., variants) contained in unaligned sequencing data (reads).

Genomic Sequencing

As a matter of fact, the extraction of genetic material from biological tissue to the readout of the genetic code into a multimillion length sequence composed of a combination of letters “A”, “C”, “G” and “T”, called nucleotides, is a computationally demanding task. The typical pipeline encountered during the sequencing process involves first the generation of raw genomic reads by the NGS machine. The number of nucleotides that compose one read can vary from 30 to 1000 whereas the number of reads is in the millions or more. The second step consists in aligning the raw reads one by one against a reference genome. Due to the considerate amount of reads that need to be positioned to maximally match with the reference genome, the alignment process requires is computationally intensive and is oftentimes outsourced to cloud service providers.

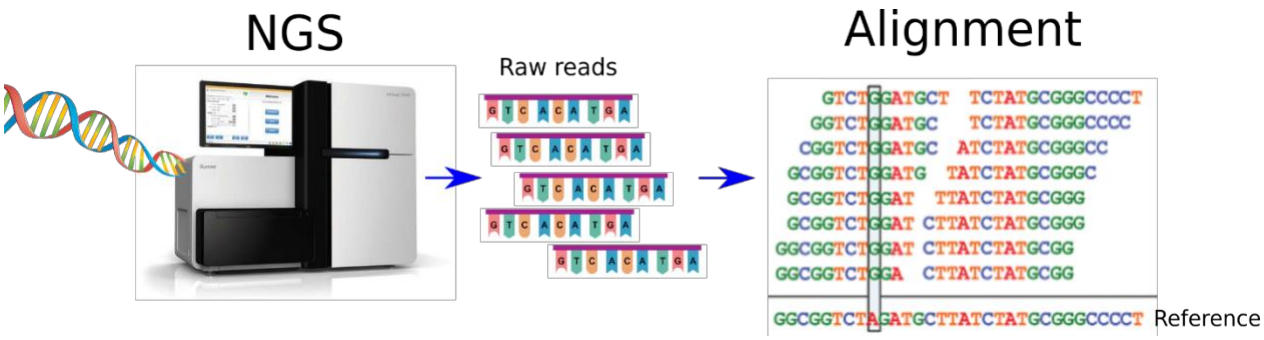


Figure 50: Genomic sequencing pipeline

Genomic Privacy

Clouds do not guarantee that the genomic raw data they host can be either accessed by the cloud service provider or be vulnerable to privacy attacks by an exterior adversary. Consequently, the increased availability of NGS technologies has come with a price, namely an increased vulnerability of personal biodata to privacy breaches. Indeed, an individual’s genome is privacy sensitive because it is uniquely associated to his identity. In addition, genetic variations are also correlated with private health information, such as disease predispositions. Several privacy attacks have been reported in the scientific literature, which pushed the research community to develop privacy-preserving solutions for secure genomic data analysis. However, few solutions have been proposed so far. One of them is data deidentification, where personal identifiers such as name, age and zip codes are removed. Data augmentation has also been proposed in which every data sample of a given individual is generalized so as to make it distinguishable from other data samples from different individuals. Lastly, cryptology-based methods have been put forward to encrypt genomic sequences and make them less vulnerable to adversarial attacks. Needless to say, each of these approaches has to face a trade-off between privacy and loss of utility. None of these methods focuses on attempting to protect the raw reads at the beginning of the sequencing pipeline, i.e., right at the mouth of sequencing machines, where it starts being vulnerable to privacy attacks.

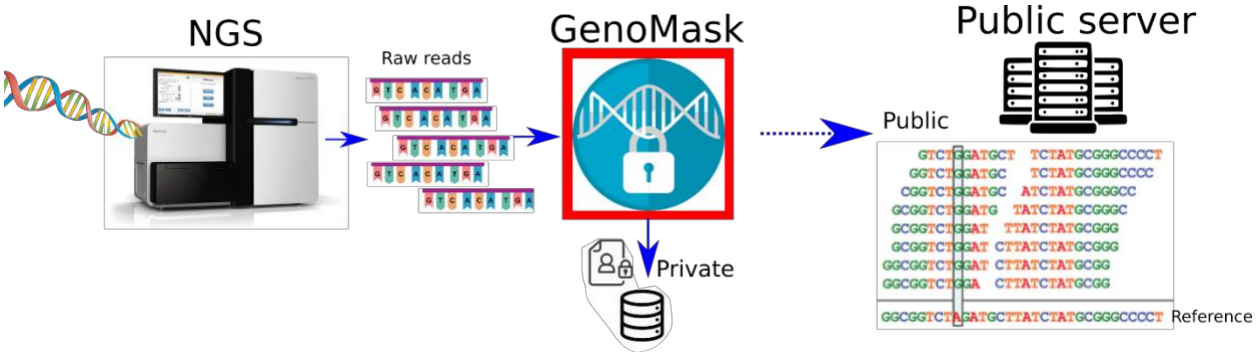


Figure 51: Privacy preserving genomic pipeline.

Nucleotide Sensitivity and Bloom Filters

This asset proposes an early read filtering solution, termed GenoMask. The requirement for an effective early read filtering technology is multi-fold:

1. Detect a maximum of nucleotides deemed sensitive

2. Filter throughput should be comparable to an NGS throughput

3. Filter should rely on limited hardware resources

The first requirement is achieved by constructing a dictionary of genomic variations that capture sequences and/or individual nucleotides that are rare inside a targeted population and could potentially be idiosyncratic to a given individual. The second condition is maintained by relying on a specific type of data structure called Bloom filters (BF). BFs are used to test whether a given element is part of a predefined set or for the purpose of genomic filtering, test whether a nucleotide is part of a pre-constructed dictionary, hence deemed sensitive. In order to be able to implement such a filtering device close to an NGS machine, the filtering process needs to be as computationally cheap as possible so as to minimize the amount of hardware, cost and consequently maximize its market viability.

Privacy Preserving Genomic Pipeline

In order to preserve the privacy of the genomic information the previously outlined genomic pipeline needs to be rethought. By design, GenoMask is intended to be installed right at the output of an NGS machine. In this setup, as the raw reads are generated by the sequencing machine and read by GenoMask, the latter will take care to separate the sensitive nucleotides from the non-sensitive ones. Two different files are created; a private file containing the reads where the non-sensitive reads have been masked but the sensitive ones are retained and a public file that contains exclusively the non-sensitive nucleotides. The first file is stored locally whereas the public one is transferred to an external machine where the alignment takes place.

4.11.2. Research Challenges Addressed

An increased availability of next-generation sequencing (NGS) technologies has come with a price, namely an increased vulnerability of personal biodata to privacy breaches. In fact, the increased speed and price decay of sequencing data production put high pressure on using high throughput alignment algorithms in cheap but unprotected environments (e.g., clouds). However, privacy risks due to the manipulation of data in plain text in the cloud, cast enormous shadows on this approach, and the recent adoption of GDPR measures will make this option untenable. However, the other extreme, like cryptographically strong alignment algorithms, provides high privacy protection, but has very limited performance, yielding delays that are unacceptable in a real-world production cycle. Thus, the current challenge, addressed by our results, consists in providing data privacy protection while taking advantage of the storage and computational power of clouds.

4.11.3. Demonstrations Example

Implementation and Performance

Since we didn't have access to an NGS machine to produce raw reads, we simulated the latter with software that ran on a desktop PC. GenoMask was built on a Raspberry Pi (model 4B) into which we plugged an external storage device. A second desktop PC served as a public server. The GenoMask box communicated via a TCP network connection with both desktop PC. We started the whole privacy-preserving genomic pipeline simulation by generating 1000000 raw reads of size 150. The GenoMask was able to detect 99.9 % of the nucleotides that were considered sensitive and stored them on the external hard drive. The unfiltered nucleotides were transmitted to the second PC. We measured the throughput i.e., the processing speed in terms of how many nucleotides the GenoMask box was able to filter per unit time. We observed a throughput speed approximately 3 times higher than the latest NGS state-of-the-art processing speeds.

The demo includes two entities as follows:

Client:

The client could be represented by any machine that pilots and controls a genomic sequencing machine (NGS). It receives raw data files from the NGS and converts (or demultiplexes in the case of Illumina machines) the raw files into raw genomic reads (FASTQ files). The FASTQ files are then transferred to the GenoMask box (server) via TCP connection for sensitive nucleotide filtering.

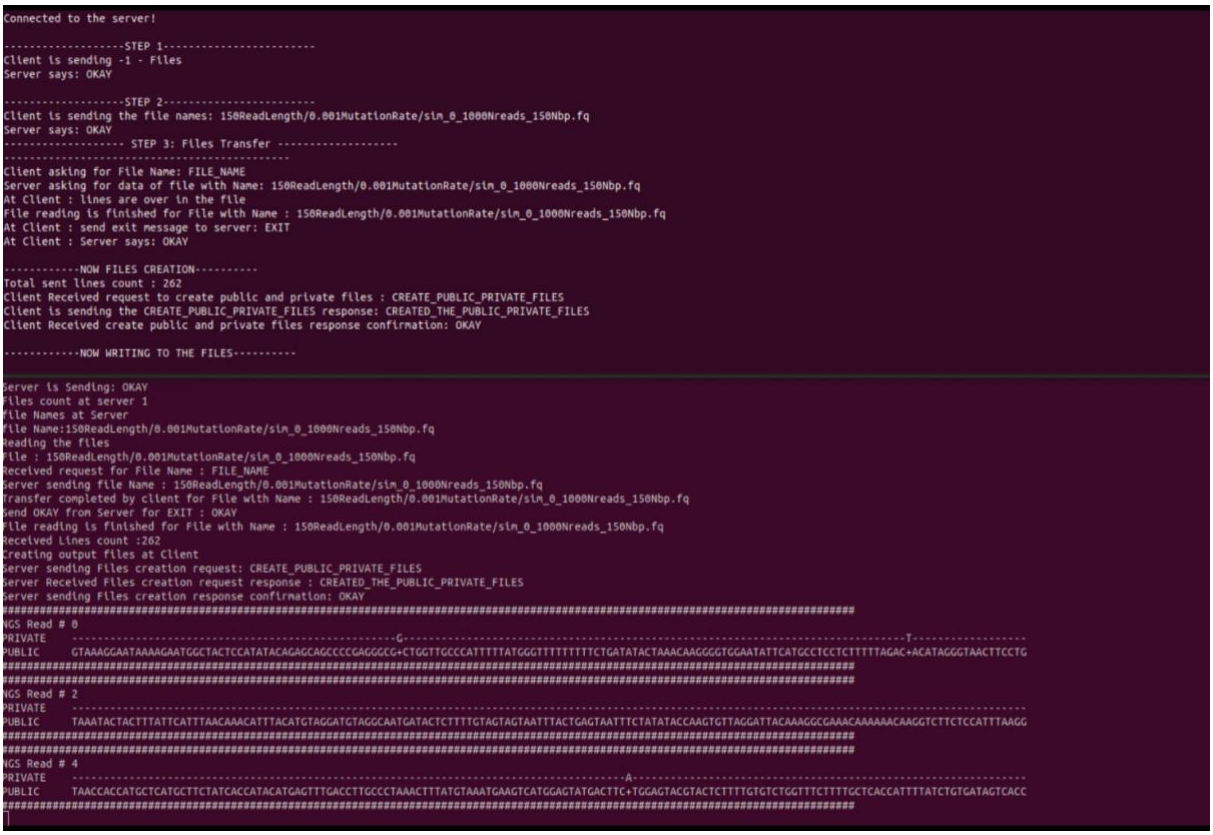


Figure 52: Sending a file to the server.

Server:

The server side represents the GenoMask box (Rasberry Pi), which receives the FASTQ files via TCP connection from the NGS (client). It processes each FASTQ sequentially and outputs the performance in terms of throughput, processing time per file, and how many nucleotides have been filtered. It then creates a private file (.PRI) and public files (.PUB) where the first contains the filtered nucleotides, deemed sensitive, and the latter contains the non-sensitive nucleotides. The PRI files are kept locally by the GenoMask box whereas the PUB files are sent back to the client which can then transfer it to a public server (non-secure) for genome alignment without disclosing the sensitive part of the sequence's genome.

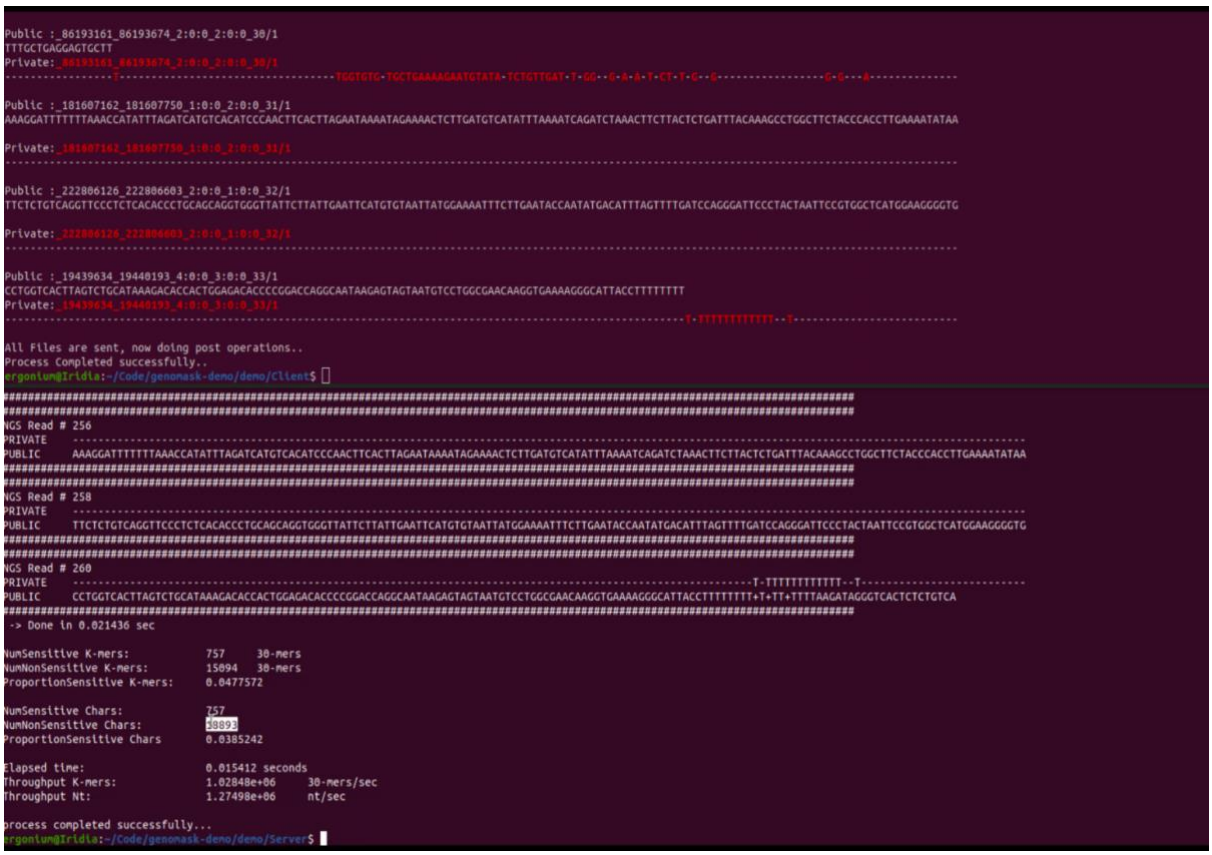


Figure 53: The process completed successfully.

4.11.4. Future Work

As future work, our intuition is to explore Intel Software Guard Extensions (SGX) enclaves which could provide new alternatives for the privacy-preserving analysis of genomic data. Furthermore, we aim to extend our approach in order to tolerate malicious adversaries and considering memory-oblivious algorithms.

4.12. GENERAL_D

4.12.1. Overview

As for any other requirement, a fundamental step for any organization (e.g., a Small and Medium-sized Enterprise (SME)) is to guarantee the compliant realization of the GDPR requirements by design. This means the integration of the data protection concepts into the overall software life cycle: from gathering the requirements to deployment and subsequent maintenance of the system. In particular, research attention has been devoted to authorization systems because they are recognized, by scientific communities and private companies, as the successful elements for the development of privacy-by-design solutions in compliance with the GDPR. However, to the best of our knowledge, most of the available proposals tend to target just a single aspect of authorization system development, and no integrated solutions for the GDPR-by-design compliant development through the entire life cycle are provided. Therefore, GENERAL_D asset (or enabler) has the following objectives:

> **OBJ 1:** defining a GDPR-based Life Cycle for authorization systems. This means to define a specific and integrated process development life cycle for the specification, deployment, and testing of adequate fine-grained authorization mechanisms able to take into account legal requirements.

4.12.1.1. **OBJ 2:** providing an integrated environment for automatically enforcing the data protection or privacy regulations. Indeed, we define an integrated environment where some of the available solutions are combined for: specifying the privacy requirements, controlling personal data, processing them, and demonstrating compliance with the GDPR in collecting, using, storing, disclosing, and/or disposing of the personal data. GENERAL_D Life cycle

We refer to by referring to Deliverable D3.2 - "Cross-Sectoral Cybersecurity Building Blocks" and D3.11 - "Definition of Privacy-by-Design and Privacy-Preserving Enablers" and the references from [generaDref1] to [generlaDref18] for a detailed description of the final release of GENERAL_D Life Cycle (Figure 53).

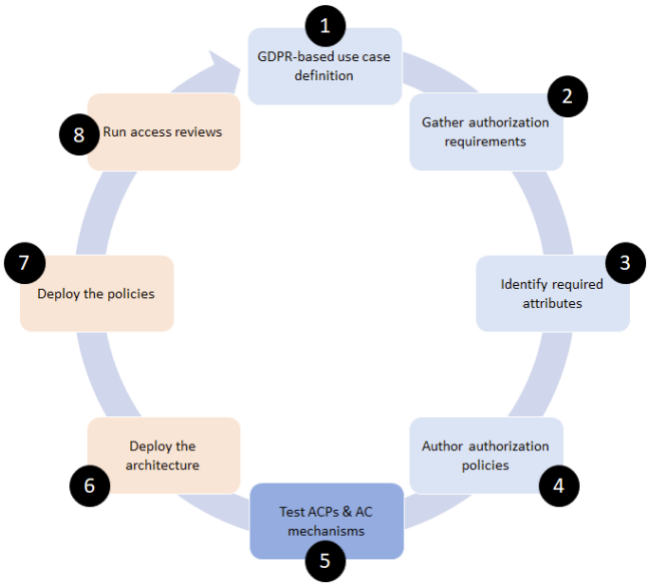


Figure 54: The Authorization Policy Life Cycle.

4.12.2. Research Challenges Addressed

The GENERAL_D solution [generalDref1] is based on the definition of AC systems that guarantees compliance with the GDPR. This required the definition of Access Control Policies (ACPs) able to express requirements aligned with GDPR's provisions and features for automatically translating the natural language requirements of the law into technical ones.

In leveraging AC systems, as a technical means for protecting "personal data by-design", and gaining legal compliance with the GDPR, the GENERAL_D solution contributed to the definition of a GDPR-based Life Cycle for authorization systems and its reference architecture, enabling data protection by-design.

It also contributes to the definition of a GDPR-based AC ontology useful for building ACPs in reference to the GDPR and to the description of a GDPR profile for a standardized AC.

Finally, the systematic approach for gathering and developing ACPs compliant-by-design with the regulation lets the definition of a comprehensive testing framework for validating both GDPR-based and traditional ACSs useful for promoting the application of ACSs in different contexts.

Considering the challenges reported D3.11 - "Definition of Privacy-by-Design and Privacy-Preserving Enablers, GENERAL_D targets the following challenges:

- DP-01 by providing the data protection backlog containing GDPR-based user stories each one connected with a specific article of the regulation [generalDref1, generalDref18].
- DP-05 by providing automatic facilities for assessing and testing access control systems that regulate/limit access to personal data [generalDref1, generalDref4, generalDref5, generalDref6, generalDref8, generalDref10, generalDref12, generalDref13, generalDref14, generalDref15].
- LDP-03 by providing an agile based authorization life cycle for the development of access control systems which is rooted in the data protection by design principles [generalDref1, generalDref2, generalDref7, generalDref9, generalDref17].
- LDP-04 by providing a set of tools supporting the overall development life cycle. They enable the controller to assess and demonstrate the compliance with the GDPR [generalDref1, generalDref1, generalDref4, generalDref2, generalDref7, generalDref9].

4.12.3. Demonstration Example 1: CCTV Surveillance

This scenario focuses on a situation in which a response team in the Smart Campus wants to continuously monitor the campus for accidents, fires, and parking spots. In particular, it wants to leverage its access control system to manage the access to the video surveillance, in normal and emergency conditions, according to the GDPR demands. According to this scenario, the preconditions are:

- The CCTV users have been already registered to the Smart Campus and provided the required consents.
- The response team already uses an access control system to regulate access to its resources and data (assets)
- The response team (i.e., controller) needs to manage Personal Data in compliance with the regulation in case of an emergency situation

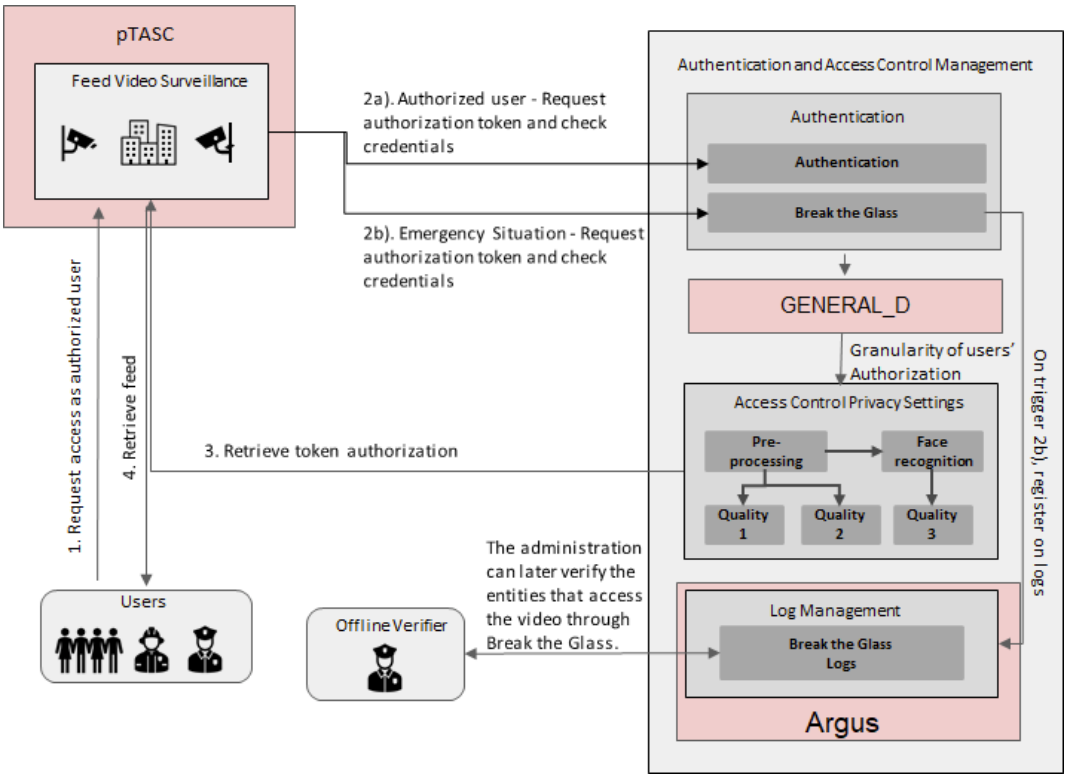


Figure 55: Integration of GENERAL_D within the CCTV Use-Case Scenario.

GENERAL_D can contribute to the specification of standard and emergency access control policies, so as to allow the management of Authorized User (step 2a in Figure 54) and Emergency Situation (step 2b in Figure 54). In this scenario, GENERAL_D provides also valid means for the access control policies enforcement in a lawful way, i.e., in compliance with the GDPR.

The first five phases of the proposed Life Cycle can help the response team to obtain a GDPR-based access control system able to automatically manage personal data. Indeed, the process can help the controller to focus on the most suitable set of GDPR obligations and to define their relations with Access Control (AC) rules. This allows to:

1 monitor the evolution of the compliance over time.
2 lawfully managing both the normal and emergency situations.
3 traceback which obligations are already covered, i.e., enabling the definition of a traceability function or feature; and
4 easily perform the review process, when necessary.

4.12.3.1. GDPR-based use case 0 (step 1)

The first step is the definition of its typical response team use-case scenarios. The aim is to help the response team in the definition and collection of the following information: Data Subjects (in terms of the GDPR), the Personal Data (i.e., the service required and the purpose the processing of the collected data). During this activity, the interactions between the end-users and the system and the possible actions should be also envisaged (e.g., Read, Write and Delete).

As a final result, each of the identified use-cases will be focused on a specific obligation so as to make the GDPR compliance process iterative and controllable. Considering the CCTV Surveillance Scenario description in Section 2 the following use-cases were identified:

1. **GENERAL_D_CCTV_UC_1 consent management use case:** for each of the involved users an explicit consent is required for a specific purpose for allowing the lawfulness of processing of personal data. The consent management behavior is identified as the derivation of the specific access control policy taking into account the collected consent and under the GDPR requirements.
2. **GENERAL_D_CCTV_UC_2 Normal use case:** The normal use case has as precondition the execution of GENERAL_D_CCTV_UC_1 and its behavior is identified as: authorized users CCTV Surveillance Scenario "*can only be authorized to track traffic lights, traffic jams but not extra information such as person faces, clothes, license plates, or any other information that can re-identify a person*"
3. **GENERAL_D_CCTV_UC_3 Special use case:** The Special use case has as precondition the execution of GENERAL_D_CCTV_UC_1 and its behavior is identified as: "a*n operator (or police, or other members with special authorization level) from the university*" .... "*Should be capable of seeing the information"* of the normal behavior "*with more quality, to allow to track cars from the municipalities or the motivation behind the agglomerate of people. However, in this case, the faces must always be anonymized.* "
4. **GENERAL_D_CCTV_UC_4 Municipality use case:** The Municipality use case has as precondition the execution of GENERAL_D_CCTV_UC_1 and its behavior is identified as: an operator in a control room from the municipality can access to the video information in high quality without anonymization. Additionally, "*the municipality operators will be capable of access the CCTV feeds on their smartphones with the information required*".
5. **GENERAL_D_CCTV_UC_5 Emergency use case:** The Emergency use case behavior is identified as: "*In case of emergency*" an operator needs "*to track a specific user or license plate of the car" in real-time. Therefore, he/she*:
   a. Make an informed emergency access request (i.e., using break the glass mode)

i. Provides the consent to the emergency situation. "*In this case, it will be shown a pop-up that advertises the user, claiming that their action will be registered on the system logs, and any later on will be verified by a chief department to validate the reason for access to that specific information"*. If the consent is not provided the emergency request will be deleted and rights of the GENERAL_D_CCTV_UC_3 Special use case
ii. Send the informed emergency access request ()

b. Authenticate the break the glass mode, where the operator "*will have access to the required information immediately*."
c. Authorize the breaking the glass by activating breaking the glass access policy
d. The break the glass log system is activated, where the operator's activities are registered.
e. A specific access control policy of the operator asking the emergency situation is generated by including the GENERAL_D_CCTV_UC_1.

GENERAL_D_CCTV_UC_6 Post Emergency use case: Post Emergency use case has as precondition the execution of GENERAL_D_CCTV_UC_1 and its behavior is identified as: an operator in a control room from the municipality verifies that the registered emergency actions on the system logs and validates the reasons for access to that specific information.

4.12.3.2. Gather authorization requirements (step 2)

For each of the use cases defined in the previous step, one or more authorization requirements are specified and represented according to the following form: [Subject] can [Action] [Resource] if [Condition].

Successively, the GDPR terms of Controller, Processor, Personal Data (and their categories), can be associated with the collected elements. Without loss of generality and for the aim of simplicity we detail in this section only the gathered authorization requirements for the *GENERAL_D_CCTV_UC_1*, *GENERAL_D_CCTV_UC_5*, and *GENERAL_D_CCTV_UC_6*.

Considering the **GENERAL_D_CCTV_UC_1 consent management use case:**

- **UC_1_R1:** Controller/processor can process users' (Data Subject) personal data if explicit consent is given for a specific purpose
- **UC_1_R2:** Users (Data Subject) can access their Personal Data anytime

GENERAL_D_CCTV_UC_5 Emergency use case:

- **UC_5_R1:** Operator can access video surveillance having full quality video in an emergency situation (breaking the glass access policy).
- **UC_5_R2:** (Controller/processor) can register the operator's (Data Subject) activities on the break the glass log if explicit consent is given for being verified by a chief department.
- **UC_5_R3:** Operator (Data Subject) can access her/his activities on the break the glass log anytime

GENERAL_D_CCTV_UC_6 Post Emergency use case:

- **UC_6_R1:** Municipality Operator can access the operator's registered activities on the break the glass log for verification purposes.

4.12.3.3. Identify required attributes (step 3)

Considering each identified use case and the collected data, this step is composed of three main sub-steps:

1 identify the GDPR's concepts involved in the authorization requirements.
2 identify the concrete attributes defined in the reference scenario; and
3 classify the identified attributes into the commonly used entities (or categories) of the AC specification: namely, Subject, Resource, Action, and Environment. As a result, a precise mapping

between the GDPR's concepts, the concrete entities involved in the reference use case, and the access control attributes are identified.

In the following, we report as an example of the application of this step, the results in a tabular form related ***UC_1_R1, UC_1_R2, UC_5_R1 and UC_6_R1.***

GENERAL_D_CCTV_UC_1 consent management use case:

UC_1_R1. Controller/processor can process users' (Data Subject) personal data if explicit consent is given for a specific purpose

Since the use case is given in generic terms, by referring to realistic data, we consider a requirement.

Smart-Campus can process operator's (Operator1 and Operators2) personal data (name, birth date, address) if explicit consent is given for the administrative purpose

Use-case Attributes	Concrete Values	GDPR attributes	Authorization Attributes
Controller (generic)	{Smart campus, CCTV team}	Controller	Subject
Processor (generic)	{Municipality Operator}	Processor	Subject
Process (generic)	{Read, Write, Modify, Analyze}	Processing Activity	Action
Personal Data (generic)	{name, birth date, address}	Personal Data	Resource
User (generic)	{Operator1, Operator2}	Data Subject	Subject
Explicit Consent	{YES, NO}	Consent	Environment
Specific Purpose (generic)	{Administrative}	Purpose	Environment

UC_1_R2: Users (Data Subject) can access their Personal Data anytime

Use-case Attribute	Concrete Values	GDPR attribute	Authorization Attribute
Users	{Operator1, Operator2}	Data Subject	Subject
Access	{Read}	Access	Action
Personal Data	{name, birth date, address}	Personal Data	Resource

GENERAL_D_CCTV_UC_5 Emergency use case:

UC_5_R1: Operator can access video surveillance having full quality video in an emergency situation (breaking the glass access policy).

Use-case Attribute	Concrete Values	GDPR attribute	Authorization Attribute
Operator	{Operator1}	-	Subject
Access	{Visualize}	-	Action
Video Surveillance	{Full Quality Video}	-	Resource
Emergency	{Emergency Condition}	-	Environment

GENERAL_D_CCTV_UC_6 Post Emergency use case

UC_6_R1: Municipality Operator can access the operator's registered activities on the break the glass log for verification purposes.

Use-case Attribute	Concrete Values	GDPR attribute	Authorization Attribute
Municipality Operator	{MunOperator1}	Processor	Subject
Access	{Validate}	Processing Activity	Action
Operator	{Operator1}	Data Subject	Resource Owner
Registered Activities	{Activity1, Activity2}	Personal Data	Resource
Break the glass log	{BreakTheGlassLog}	-	Resource
Verification purpose	{Verification}	Purpose	Resource (Attribute)/ Environment

4.12.3.4. Author authorization policies (step 4)

An access control policy is defined for each use case. This can be performed through the following activities:

1) define a set of abstract access control policies, each related to a specific access control requirement identified in step (2), by using the mapping results obtained in step 3.
2) combine the obtained rules into a single abstract access control policy by (i) defining the order in which the rules will be evaluated, (ii) adding a default rule, and (iii) selecting the rule conflict resolution algorithm. As a result, an abstract policy (i.e., ACP not directly enforceable by the AC system), expressed in terms of GDPR's concepts can be derived.

3) replace each GDPR's concept in the abstract policy with the corresponding one (i.e., the concrete attribute gathered from the reference scenario) according to the precise mapping result in the previous step

Among the obtained access control policies, in Figure 55 we report the one associated with GENERAL_D_CCTV_UC_5 Emergency use case, by highlighting the part related to requirement UC_5_R1, i.e., breaking the glass access policy.

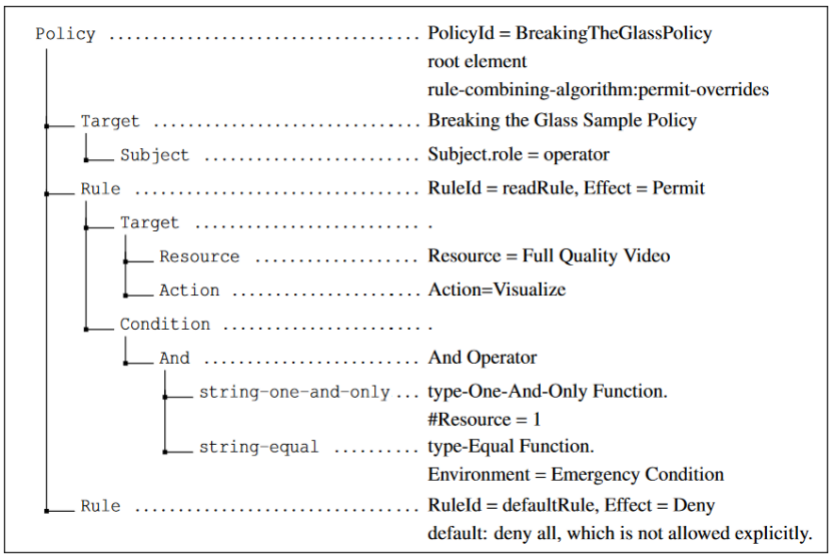


Figure 56: An XACML-like Breaking the Glass Sample Policy.

4.12.4. Demonstration Example 2: Identity Management and Service Usage

In this section, we illustrate how GENERAL_D has been integrated within the Identity management and service usage scenario from the architectural point of view. In this instantiation, GENERAL_D has been enhanced with consent management capabilities, so as to allow the management of the consent given by the different data subjects (e.g., Students) and the definition of specific purposes defined by controllers (e.g., University or Parking Service Provider).

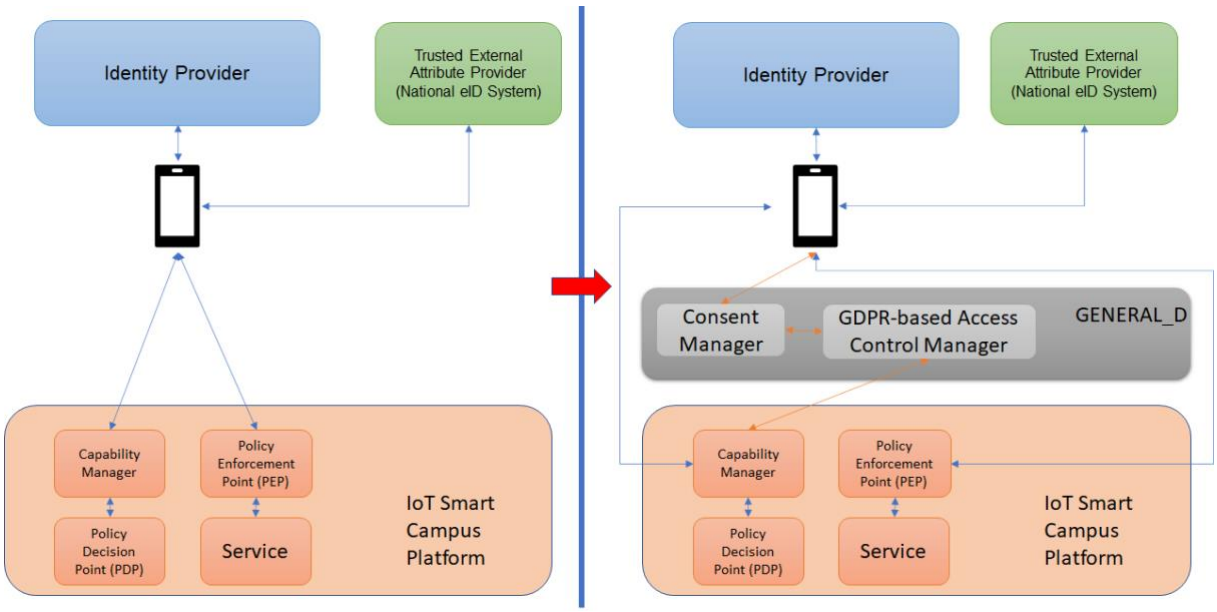


Figure 57: GENERAL_D in the Context of Identity Management and Service Usage.

For this, in Figure 56, we depict on the left side the standard simplified architecture of the Smart Campus and on the right side of the Figure our enhanced view with GENERAL_D.

On the left side of Figure 56 we depict a simplified version of the Smart Campus IoT Platform provided by the University of Murcia described in section 4.3 Figure 16, and in its enhancement with GENERAL_D so as to allow left side,

The proposal of this demonstration example is to leverage extra components that form an additional layer for actions related to GDPR like user consent, as reported in Figure 56 right side. GENERAL_D is in charge of modeling and enforcing the GDPR legal framework and includes two specific components: Consent Manager that translates the textual consent into structured representation, and Access Control Manager that provides enforceable access control policies.

Consent Manager

The aim of the Consent Manager is to manage and control personal data during the interaction among Data Subjects and public and private services that act as Data Controller and Processors (e.g., University or research centers). It provides facilities for lawful data sharing processes, with the ability to grant and withdraw consent to third parties for accessing personal data.

The aim of the consent manager is to perform steps 1 and 2 of the proposed life cycles. For this, it interacts with the Smart-campus platform app by preparing the consents to be subscribed by the users. Therefore, after receiving the Consent Record representing the given consent by the end-user of the system (i.e., the Data Subject), the Consent Manager extracts the useful personal data from the signed consents and stores them into the personal data DB.

Additionally, the Consent Manager should guarantee by-design the compliance with the GDPR's demands, such as data minimization and purpose limitation principles

GDPR-Access Control Manager

The aim of this component is to perform steps 3 and 4 of the life cycles by creating Access Control Policies that are compliant by-design with the GDPR. This component works in collaboration with the

Consent Manager by receiving, as input, the machine-readable specification of services definitions and the related Data Subjects' consents. Basically, the Access Control Manager component uses: Personal Data related to Data Subject classified in categories as required by the GDPR; information about the Controller of each service and the defined purposes; the consent given by the Data Subject in terms of a relation between Personal Data and Purposes. Based on that information, the Access Control Manager is able to create specific Access Control Policies, each related to a specific article of the GDPR. The peculiarities of the Access Control Manager are the possibility to (a) be integrated with different Consent Managers, and (b) collaborate with different Access Control systems.

4.12.4.1. GENERAL_D Customization

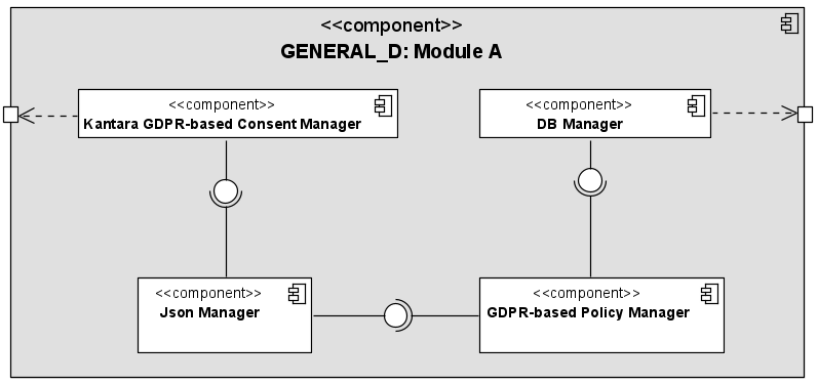


Figure 58: Customization of GENERAL_D and Consent Manager Integration.

For this demonstration, a customization of GENERAL_D has been used for obtaining GDPR-based access control policies directly from the consent provided by the Kantara-based Consent Manager, so as to allow a lawful processing of personal data managed by Smart-Campus Platform. As depicted in Figure 57, the customization of GENERAL_D involves only Module A of the reference architecture (see Figure 58), and it includes four main components: Kantara Consent Manager, Json Manager; GDPR-based Policy Manager; and DBs Manager.

Kantara GDPR-based Consent Manager: In the current implementation, among the consent format available in both industry and academia, we rely on “Consent Receipt Specification” proposed by the Kantara initiative, and more precisely, we refer to its draft GDPR extension version named “GDPR Explicit Consent Record & Receipt Extension for Kantara CISWG: Consent Receipt”, which is under active development. The specification proposes a JSON schema for a consent receipt and it contains all the required GDPR concepts useful for authoring ACPs in compliance with the regulation. This component is therefore responsible for managing the interaction with the end-users; it allows controller to define specific purposes and data subjects to give an informed consent for processing their personal data.

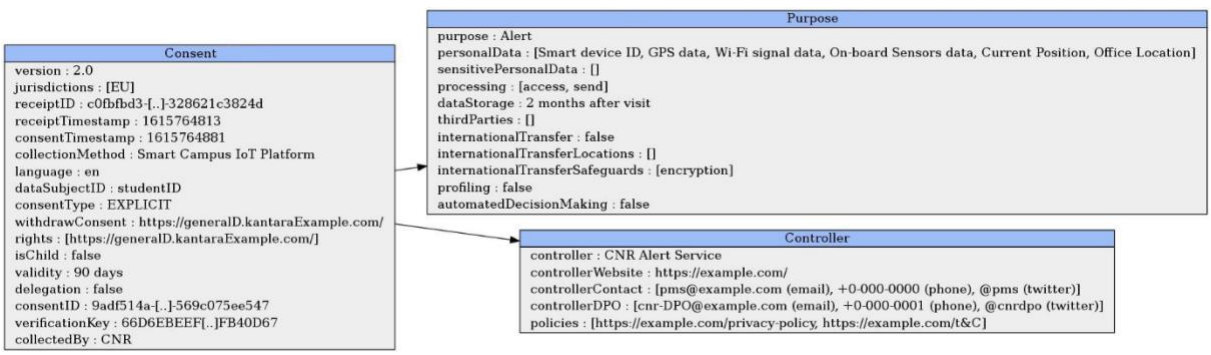


Figure 59: Example of Kanatara Consent Receipt of CNR Alert Service.

An Example of a concrete receipt is depicted in Figure 58. The receipt contains:

- Consent element, which reports information about:
  - the Data Subject, e.g., dataSubjectID, isChild attributes. In our example, Student is an adult person identified by ID studentID.
  - the consent itself, e.g., the consentTimestamp, the validity of the consent, the collection method and who performed it (i.e., the CNR Alert Service in our case).
- Purpose element, which is related to the explicit given consent, contains the following attributes among others:
  - the purpose name and the allowed actions (i.e., processing) to achieve it.
  - the set of personal data the data subject is given to controller and their storage validity.
  - the involved third parties in the processing, and whether personal data are using for profiling or for automated decision making.
- Controller element, instead, contains both its contact information and the DPO's contact.

JSON Manager has the responsibility to interact directly with Kantara GDPR-based Consent Manager; it receives the consent in JSON format, and it parses that consent so as to extract the relevant information for the ACP generation purpose. Such information includes, among others, the Consent ID and the defined purposes of processing, the allowed operations, and Personal Data provided by the Data Subject.

GDPR-based Policy Manager: has the responsibility of creating enforceable ACPs encoded in the XACML language. It interacts with 1) JSON Manager for retrieving the data to be processed; 2) DB Manager to retrieve the GDPR-based ACPs templates and store the obtained policies.

DB Manager: offers databases supporting functionalities to the GDPR-based Policy Manager (e.g., create/modify/delete a database, and insert/modify/delete specific entries in the available tables).

A preliminary performance evaluation of the customized version of GENERAL_D framework has been conducted including but not limited to the parsing of the Kantara consent receipt and the generation of executable access control policies. We, therefore, focused on the process for converting the consent provided by the Kantara GDPR-based Consent Manager into an executable GDPR-based access control policy expressed in the XACML formalism. The measurements were carried out using a Dell Laptop (Intel(R) Core (TM) i9-9980HK CPU @ 2.40GHz and 16 GB RAM), and each final time value was obtained as the average of 10 executions. Figure 59 reports the results associated with two main phases: parsing the consent and XACML policy generation.

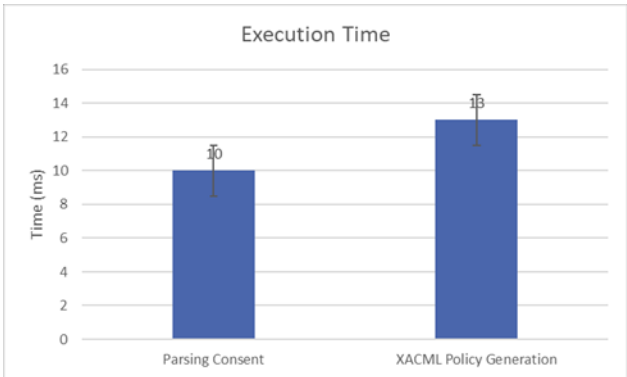


Figure 60: GENERAL_D Execution Time.

For the evaluation, different consent structures have been considered either coming from real case studies or obtained by operating specific modifications or extensions to the existing one. They differentiate each other in the complexity and in the number of personal data and purposes considered so as to cover the most realistic scenarios. For the aim of clarity, we report in Figure 60 a more complex

example of Kantara consent receipt which contains two purposes: Show Personalised Ads and Manage Delivery of Goods. The consent allows also the controller to share personal data with two different third parties: Delivery playing the role of partner and Ad Generator which is a processor.

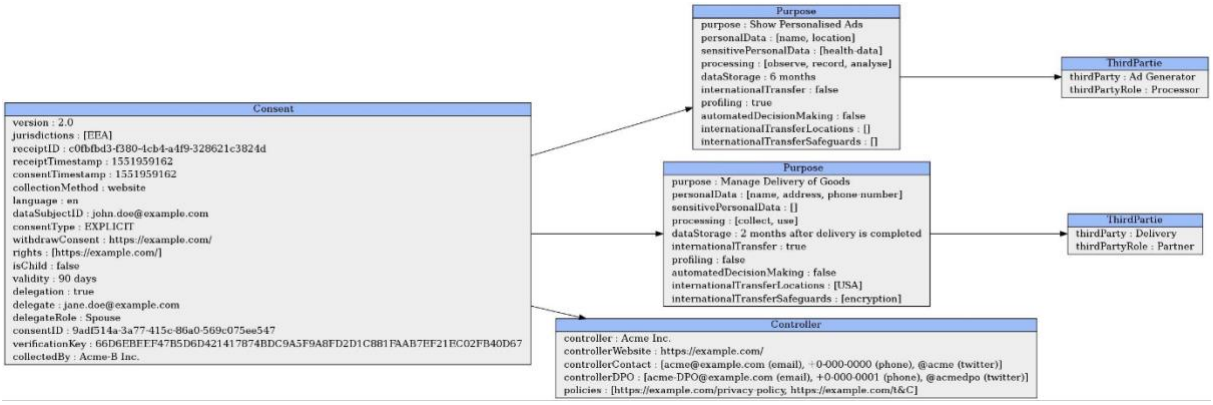


Figure 61: Complext Example of Kanatara Consent Receipt.

4.12.5. Future Work

Despite the accuracy devoted in investigating the main challenges reported in this deliverable, future research activities will include Standardization of the XACML GDPR Policy Profile, Validation of the GENERAL_D proposal through additional case studies, discussion and validation of the proposal with legal experts so as to confirm the compliance with the GDPR; releasing a reference architecture of GENERAL_D and provide the user stories templates suitable for other legal frameworks.

4.13. Blockchain Platform

4.13.1. Overview

This asset provides a blockchain-as-a-service platform but with improvements over state-of-the-art solutions (e.g., Hyperledger Fabric) that address a number of important issues the technology suitable for the fintech world[10]. Namely:

- Privacy: existing blockchain platforms assume work by broadcasting transactions to the entire network. However, both the corporate world and private citizens value their privacy, wanting to restrict data sharing to a number of parties of their choice. Exchanging encrypted transactions is a good practice, but it still lets the network know when a particular transaction occurred.
- Scalability: permissionless blockchains (e.g., Bitcoin, Ethereum) scalability is excellent, but sacrifices throughput. Permissioned blockchain can achieve higher throughput but sacrifice scalability. Today's standards require a scalable blockchain architecture that does not sacrifice throughput, making it useful for both organizations and private citizens.
- Lack of Governance: blockchains offer a distributed platform that does not need a trusted third-party. However, in real deployments, organizations and service providers want to be in charge of their networks in order to enforce business logics and policies of their choice and to be able to grant or deny access to their services as they see fit.

[10] Li, W., Sforzin, A., Fedorov, S. and Karame, G.O., 2017, April. Towards scalable and private industrial blockchains. In *Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts* (pp. 9-14).

The key feature of the platform is its architecture comprising "satellite chains": small, independent blockchains with their own ledger, applications, consensus algorithm, and participants. Note that this model does not prevent the transfer of assets from one satellite chain to another, because all satellite chains are part of a larger "parent" network usually managed by the network's owner (see "lack of governance" point above). The parent keeps track of the active satellite chains but does not interfere with their normal operations.

The advantage is that information is exchanged only between the intended parties, thus limiting the sharing of knowledge to who matters. The only requirement is instantiating a satellite chain that they all join. Transactions exchanged within that satellite chain are visible only to its members. See also Figure 61 for an illustration. In the example, the "circles" satellite chain does not that the "triangles" satellite chain exists because they are independent. Similarly, the "squares" satellite chain and the "circles" satellite chain are not aware of each other, even if one entity is a member of both. Only that member knows the existence of both. There can be as many satellite chains as the use case needs, there is no upper bound.

This design gives two benefits: efficiency, because participants do not have to store transactions that are not relevant to them; scalability, because by allowing an unbounded number of satellite chain to work in parallel, the platform achieves a higher throughput than existing deployments.

Note that smart contracts are available as well: organizations can thus implement their business logics and policies via a program uploaded to the blockchain. In particular, the organization managing the network can assume the role of a "regulator" that implements policies that apply to all satellite chains active in the network. For the second cycle of the project, the asset named "Consensus Research"[11] has been integrated into this one.

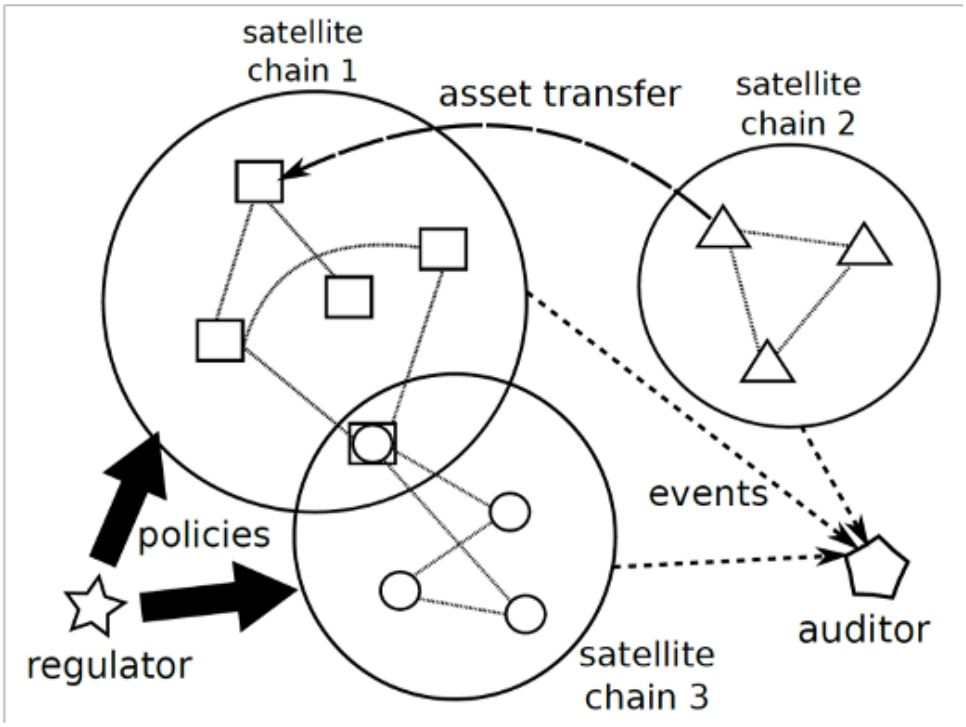


Figure 62: Blockchain platform architecture showing three independent satellite chains in action

[11] See D3.2 - *Cross Sectoral Cybersecurity Building Blocks*

Supporting these features is a Byzantine Fault Tolerant (BFT) consensus protocol - named FastBFT - NEC has perfectioned over the course of the project[12]. The recent researchers' interest in BFT protocols comes after years of the blockchain being in the spotlight. For example, Bitcoin relies on proof-of-work (PoW) to agree on the order and correctness of transactions. However, pow has been proven to be inefficient and too slow to be useful in the fintech world. This is why research institutions and private organizations are investing in designing new BFT protocols that could finally allow organizations, (e.g., financial institutions and supply chains) to leverage the blockchain for their business.

4.13.2. Research Challenges Addressed

Our platform is based on Hyperledger Fabric and improves upon it. The satellite chains feature provides transaction privacy and system scalability that is unachievable with another existing platform. Another popular blockchain platform is Corda [4]. Corda leverages so-called flows to provide transactional privacy; they establish point-to-point connections between nodes that want to exchange transactions. By doing this, transactions will be only visible to the communicating parties. The consensus protocol in Corda, a cluster of nodes (notaries) that receive and verify all the transactions. In practice, notaries need to inspect transactions to validate them, therefore, when traversing the transaction graph to verify a newly received transaction, nodes learn about other transactions issued by other nodes, thus partially breaking transactional privacy. Finally, given the lack of scalability of existing blockchains, recent works have proposed to use shading. However, shading protocols split the load of only transaction processing. All nodes still need to receive and store all confirmed transactions and blocks, and validator nodes still have to store the complete blockchain history.
Our consensus algorithm FastBFT leverages a novel message aggregation technique combining hardware-based trusted execution environments (e.g., Intel SGX) with lightweight secret sharing. Message aggregation allows FastBFT to have a message complexity of O(n). Furthermore, unlike existing schemes, our message aggregation protocol does not require any public-key operations (e.g., multi-signatures), thus lowering computation and communication overhead. Typically, byzantine protocols performance decrease as the number of nodes communicating increases. Our experiments show that, as the number of nodes increases, FastBFT's throughput deteriorates considerably slower than other BFT protocols. This makes our protocol well-suited for next generation blockchain systems. For example, assuming 1 MB blocks and 250 bytes transaction records (as in Bitcoin), FastBFT can process over 100,000 transactions per second.

With respect to the challenges laid out in D3.11 [D3.11], this asset tackles challenge IDP-03.

4.13.3. Demonstrations Example

This asset is integrated in two WP5 demonstrators: *Open Banking* (T5.1) and *Supply Chain Security Assurance* (T5.2).

In the context of WP3, this asset could be integrated into the *Smart Campus* demonstrator, to serve as its DLT backbone. The combination of satellite chains and smart contracts can take care of the heterogeneous services available to users. For example, different services can have different dedicated satellite chains accessible only to those users that can access them. Complex authorization operations, such as registration and ID verification could be carried out by smart contracts.

The heart of the platform is FastBFT, a novel consensus protocol comprising a message aggregation technique that leverages hardware-based trusted execution environments (e.g., Intel SGX), allowing it to lower its message complexity from $O(n^2)$ to O(n). Further, the protocol works in optimistic BFT mode, that is, it only requires a subset of nodes to run the protocol.

[12] Liu, J., Li, W., Karame, G.O. and Asokan, N., 2018. Scalable byzantine consensus via hardware-assisted secret sharing. *IEEE Transactions on Computers*, *68*(1), pp.139-151.

Regarding performance, as the number of nodes increases, the protocol's throughput in terms of transactions processed per second decreases slower than other state-of-the-art BFT protocols, reaching, in some scenarios, over 100,000 transactions per second. This makes it the ideal consensus algorithm for scalable and performant blockchains able to cater to the industry's needs. Figures 62, 63, and 64 show its performance compared with other popular consensus protocols.

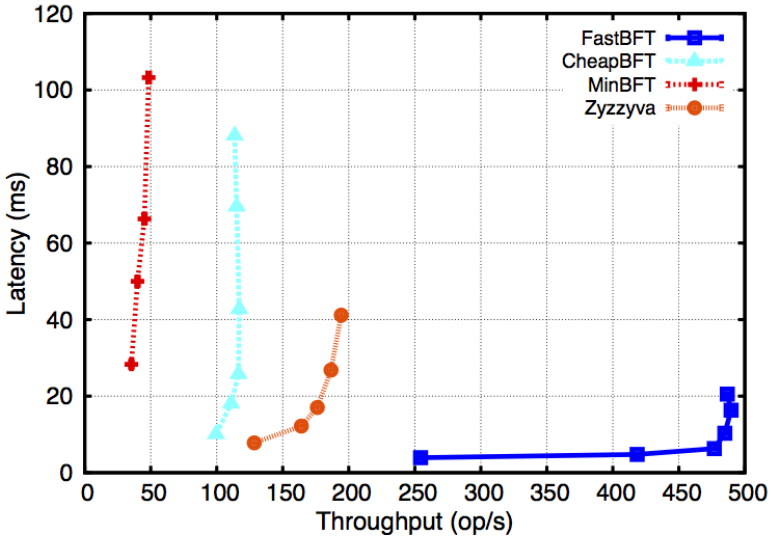


Figure 63: Protocol latency as the number of operations per seconds increases.

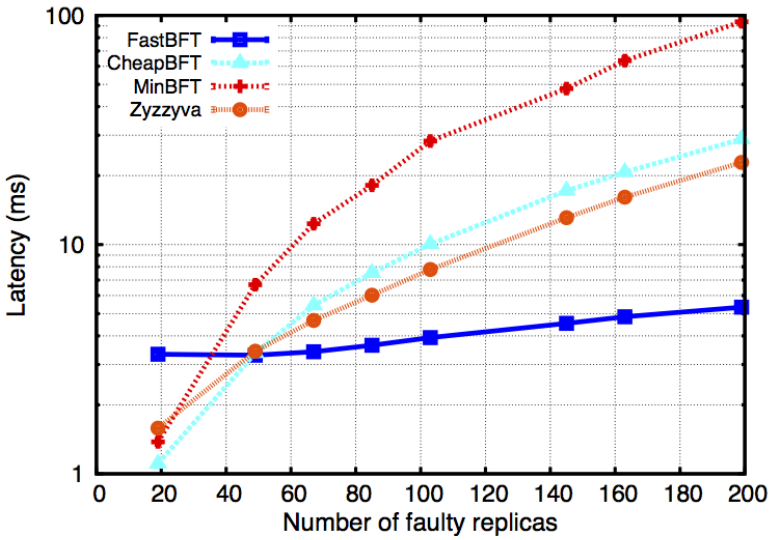


Figure 64: Protocol latency as the number of nodes increases.

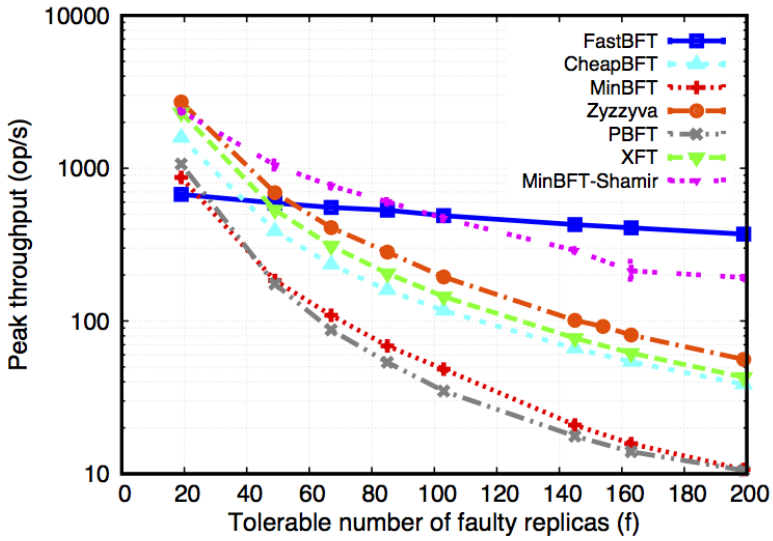


Figure 65: Protocol throughput as the number of nodes increases.

As the Figure shows, our protocol has a steadier performance when compared to other well-known protocols. In particular, the size of the network has a significantly lower impact on the protocol's latency and throughput than its competitors.

4.13.4. Future Work

We plan to complete most of the research and development work on this asset within the lifetime of the project. In the future, we plan to investigate new use cases that this asset can implement, with the additional goal of using it as a Blockchain-as-a-Service (BaaS) infrastructure.

4.14. Sharemind

4.14.1. Overview

Sharemind is a secure computing platform that consists of Sharemind MPC (based on secret sharing) and Sharemind HI (based on trusted execution environments). Here we will describe Sharemind HI, which is a development platform for the confidential analysis of data from multiple parties on a centralized server with full control overexposing the data and results to others.

Sharemind HI was created to reduce the risk of a privacy breach when processing confidential data. The data is encrypted at the source, by the data owner, and only then sent to the Sharemind HI service. The host of the service will not have access to the unencrypted data nor the encryption keys. Sharemind HI does not remove the data protections even while processing it, the data will remain protected by cryptographic means during the whole analysis.
Sharemind HI relies on a trusted execution environment (TEE) technology to provide security guarantees. A TEE isolates the security sensitive parts of an application from the rest of the system with the help of some trusted hardware. The TEE technology used in Sharemind HI to implement the privacy-preserving data processing is Intel® Software Guard Extensions (SGX) which is available in modern Intel® processors.
Sharemind HI is built as a client-server service. The solution is based on tasks that run inside SGX. Each task resides in a separate SGX enclave. The client is an application that calls operations on the server, encrypts data and performs remote attestation on the server. The Sharemind HI server does the bulk of the work and is responsible for the following:
• checking if a user has the right to access the system (authentication),
• checking if a user has proper roles and data access permissions to perform an operation (authorization),
• managing the encryption keys of the data,

• managing the secure data transport in the solution between tasks and external stakeholders including data upload and download,
• storing a log of the operations performed in the server,
• scheduling the solution tasks to run.
The security model of Sharemind HI relies on the security guarantees provided by SGX. The data encryption model of Sharemind HI is illustrated in Figure 65. The input data, shown in light blue, is encrypted at the client side and sent to the server. The input data encryption keys of the data are securely transferred to the SGX protected enclaves. Likewise, the output data, shown in green, is encrypted inside the enclave and stored on the server. When requested, the enclave securely transfers the output data encryption keys to the authorized clients.

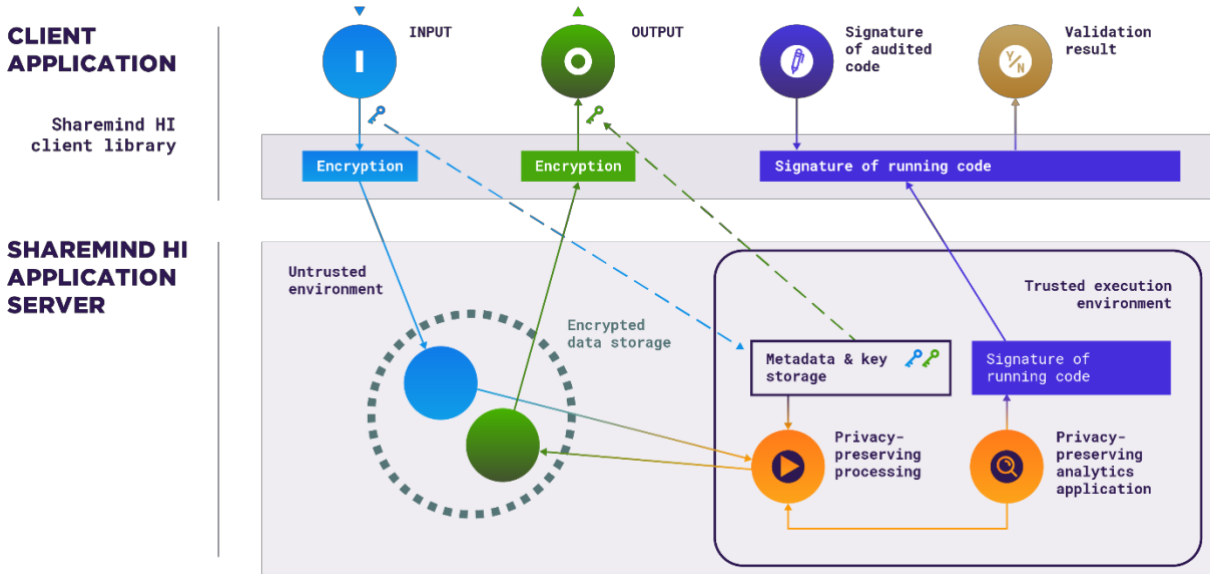


Figure 66: Sharemind HI security model

It is the obligation of the enforcers to verify that a task is configured as expected. Input providers and output consumers specify which enforcers they trust with this task. Sharemind HI ensures that they can only upload data to and download data from tasks which have been approved by their trusted enforcers. This link of trust prevents clients from sending data to or receiving data from a wrong task.
At any point during the deployment, a client can request cryptographic proof of what analysis code is running in the server, shown in dark blue in Figure 65. This proof can be compared against a previously generated proof by an auditor who has validated the code to be secure.

For each Sharemind HI deployment, a deployment coordinator has to generate a deployment specific private key and public key certificate. This private key is used to sign all the client keys that want to communicate with the Sharemind HI server. The signed deployment certificate is loaded into the server at startup and is used to authenticate clients in the remote attestation. The root CA certificate is embedded into the server and verifies the validity of the deployment certificate. This process is needed to establish a root of trust for the server.

4.14.2. Research challenges addressed

Using secure computing, whether multi-party computation or trusted execution environments, adheres to the principle of privacy-by-design as these privacy-enhancing technologies ensure that the privacy of individual values is not an afterthought. In this solution, it is possible to use either trusted execution environments or multi-party computation as the backend. No individual values can be viewed by the data analyst or the decisionmaker, however, aggregation results are clearly useful and usable. This

addresses the data privacy challenge DP-02 (when using secure multi party computation (MPC) to analyze data, analysts are not able to see the individual data values) from Deliverable D3.11.

Publication: Ostrak, A.; Randmets, J.; Sokk, V.; Laur, S.; Kamm, L. Implementing Privacy-Preserving Genotype Analysis with Consideration for Population Stratification. Cryptography 2021, 5, 21. https://doi.org/10.3390/cryptography5030021

4.14.3. Demonstration example

Statistical mobile phone location data analysis helps smart cities and campuses adapt to their visitors and provide services. For example, the countries of origin affect the languages of signs and services, or their times of arrival affect event planning. During the COVID-19 pandemic, mobility data was also used to understand the movement of people during lockdowns.

Mobile phone operator data is obtained by triangulation via cell towers. It is unbiased data - a user does not need to own a smartphone or install a special app on the phone to be included in the dataset. This data includes all mobile phone users, and it is therefore not possible to seek consent from each of them. We need a solution providing the best possible anonymity of the users. Secure computing with trusted execution is one technology that can provide this kind of privacy assurance. We demonstrate a solution that ensures fully privacy-preserving aggregation of mobile phone location data.

Location data are encrypted at the source with a key that is only available within the secure computing environment and never to its host. All aggregation of data takes place within the Sharemind environment where confidentiality and integrity are enforced by hardware. This ensures that statistical reidentification and linking with auxiliary data are infeasible attacks.

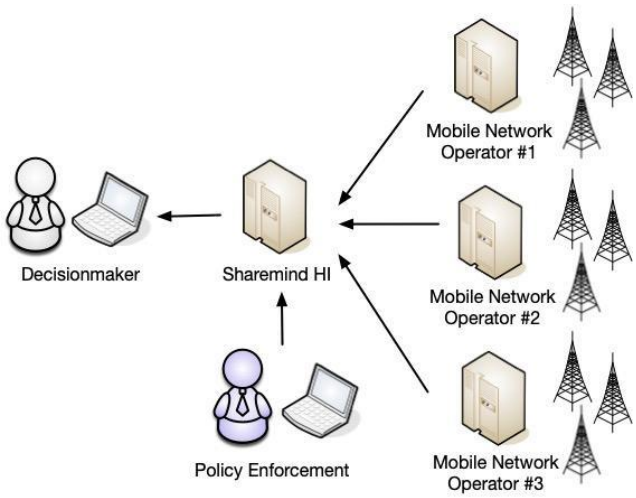


Figure 67: Sharemind deployment scheme for the smart campus scenario

Using mobility analytics insights from our partner Positium (https://positium.com), we built a demonstrator that analyses information from roaming phones from foreign countries and allows this data to be analyzed visually. The data is stored and analyzed in the trusted execution environment of Sharemind HI, meaning that all of the data is aggregated without having access to individual values. Sharemind HI is a backend service (the server with the trusted execution environment) and as such, it is difficult to demonstrate. Hence, the demonstrator itself focuses more on the front end and shows how the data can be efficiently used without seeing individual values. The mobility data in the demonstrator is from Estonia, so the mobility analysis is also performed with respect to Estonia.

Such an application offers a good dynamic overview to the decision-maker. It is possible to see where people are coming from. The data can be filtered for example based on country, nature of the visit, duration of the visit, country of origin (Figure 67).

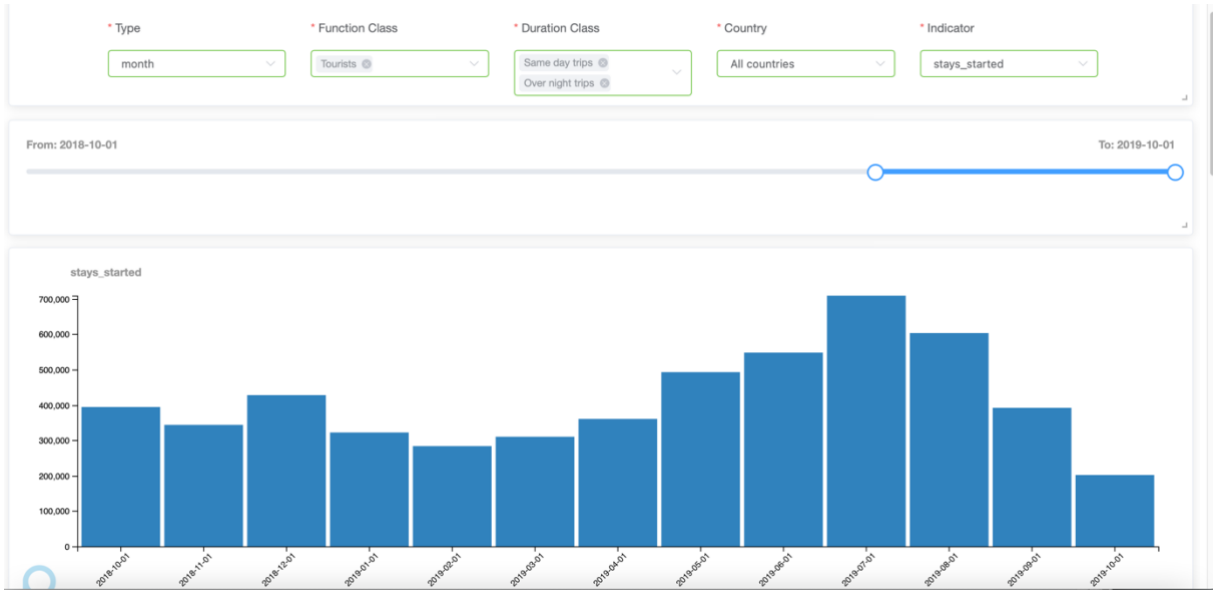


Figure 68: Filtering the data

It is possible to easily view which months are more popular for visiting and whether seasonality affects travel from different countries (Figure 68).

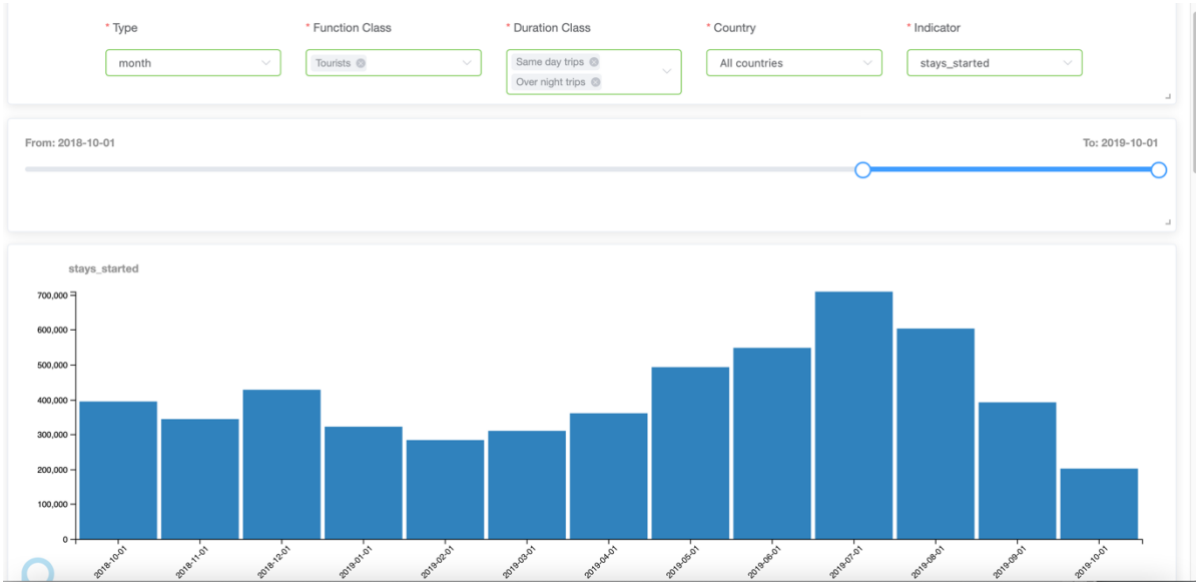


Figure 69: Visit seasonality comparison for Spain (ES, above) and Finland (FI, below)

For more advanced analysis, it is also possible to view statistics based on the classification of people - whether tourists, transit travelers, or cross-border workers. These sorts of in-depth analyses are helpful in other users like national statistics offices, central banks, and national tourism ministries.

4.14.4. Future Work

For Sharemind, we plan to implement new machine learning algorithms, so the data analysis that can be done with secure computing will be even more extensive. We are investigating ways of making the implementation more developer friendly, so more people would be able to implement applications for Sharemind, thus making the adoption more accessible.

We also plan more work on the PLEAK differential privacy analyzers, improving the tool and determining how to best choose epsilon for different cases.

4.15. Cloud-Based Credentials

Cloud-based credential systems provide a mechanism for privacy-friendly identity management. They enhance the state of the art by not only offering means for selective disclosure and data minimization (cf. also asset SS-PP-IdM (Section 4.3)) but by additionally focusing on resource-constrained devices. This is achieved by outsourcing all computationally heavy operations to the cloud without putting a user's privacy at risk.

4.15.1. Overview

Cloud-based anonymous credentials were first introduced within the H2020 CREDENTIAL project[13]. The main components of cloud-based anonymous credential systems are as follows:

- **Identity provider**: This is the identity provider (or issuer) in a cloud-based credential environment. In our implementation, this is given by a web service where users can register and receive certificates on their sensitive data.
- **User application**: This is the user's local application required for performing privacy-preserving authentications. In our implementation, this is realized through an Android mobile

[13] https://credential.eu/

app. However, the corresponding cryptographic libraries require only a minimum computational capacity and could easily be carried out also, e.g., on a student ID card.

- **CREDENTIAL Wallet**: This is a cloud-based service where users can upload encrypted versions of their credentials. To authenticate, the user application triggers the necessary computations in the Wallet, which, depending on the concrete implementation, directly computes with the relying service, or routes all communications through the user's device.
- **Relying party**: Due to the cryptographic specificalities of cloud-based credentials, relying parties (e.g., cloud services) need to integrate this end point into their system.

4.15.2. Research challenges addressed

While attribute-based credentials already provide high privacy guarantees, they are often too expensive to be used on low-cost or embedded devices, including IoT devices. This has also been identified as an open challenge in D4.4 (Section 5.5.14). The presented solution, building on results from previous projects[14], minimizes the computation costs on the embedded devices, making attribute-based credentials available also in resource-constraint environments.

Regarding the challenges identified in D3.11, cloud-based ABCs also address the following challenges:

- *IDP-02: Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks*. Cloud based credential system give the user full control over their data, in the sense that they can selectively decide which information to disclose and which information to keep confidential. In combination with intuitive user interfaces, the risk of unnecessary over-identification is directly reduced.
- *IDP-04: Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP).* Our design of cloud-based ABCs directly addresses potentially malicious Wallet providers by guaranteeing that they never receive access to any information in the plain. By routing the data flow through the user's device, also metadata could be kept inaccessible to them.

4.15.3. Demonstrations Example

In the following, we describe in detail how cloud-based credentials can be used in the context of the smart-campus scenario. The underlying implementations of the components introduced above have been carried out within CREDENTIAL, while CyberSec4Europe mainly focused on the advancement of the cryptographic libraries, which are application-agnostic.

In a first step, upon **user enrolment** at the university, the student obtains access to the university's identity management system and also opens an account in the CREDENTIAL Wallet. Furthermore, the student installs the user application on her smartphone. In order to link her smartphone application to the identity management system, the student logs into the identity provider and scans a QR code containing an ephemeral one-time key.

[14] https://credential.eu/

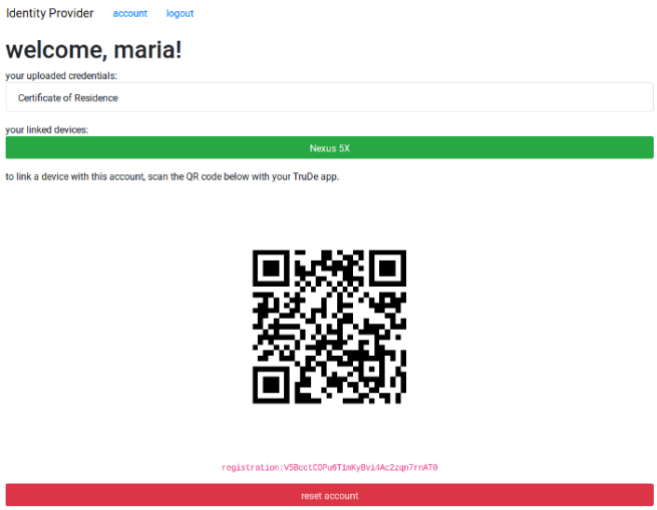


Figure 70: Linking a device to the IdP in cloud-based credential systems

Once having linked a mobile phone to the student's account at the university issuer, the user can receive anonymous credentials from this issuer in the **credential issuance** process. To do so, the user displays the corresponding entry in her IdP account. Again, scanning a specific QR code, the user application initiates the issuance process of the credential system, during which the anonymous credential is generated in an interactive protocol. The resulting credential is then directly uploaded to the user's account in the CREDENTIAL Wallet, and only the user's master key (a 256-bit random integer) remains in the user application. (Note here that in our current scenario we are only considering a single instance of the CREDENTIAL Wallet. While multiple independent instances would be conceivable, in which case an advanced linking process would be necessary to link the student's accounts and mobile application, this would pose additional privacy challenges as the current design of cloud-based credentials would potentially reveal the instance being used.)

Now, the **privacy-preserving service access**, the user once again scans a QR code displayed by the service provider. This QR contains a detailed specification of the presentation policy, in particular defining the attributes required for being granted access. This presentation policy is then presented to the user in a human accessible format, as shown in Figure 70:

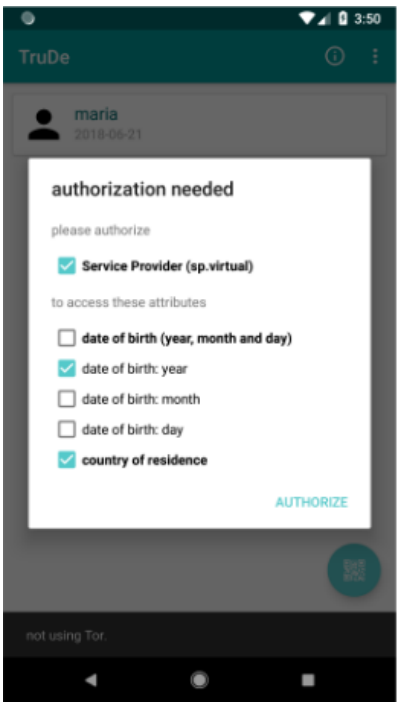


Figure 71: Users need to explicitly authorize service providers to access their attributes

Upon the student's consent, the CREDENTIAL Wallet, jointly with the user application, now computes a presentation token for the selected anonymous credential, only disclosing the displayed attributes. This token is then forwarded to the relying party.

In the current implementation, the presentation token is transmitted directly from the CREDENTIAL Wallet to the relying party, thereby disclosing the identity of the relying party to the Wallet. A possible refinement of the implementation could be to route the presentation token through the user application in order to disguise the information for which service provider a specific token is computed by the Wallet. A high-level overview of the authentication process is given in Figure 71. For details, we refer to the original literature [KLSS17,K18, HK19].

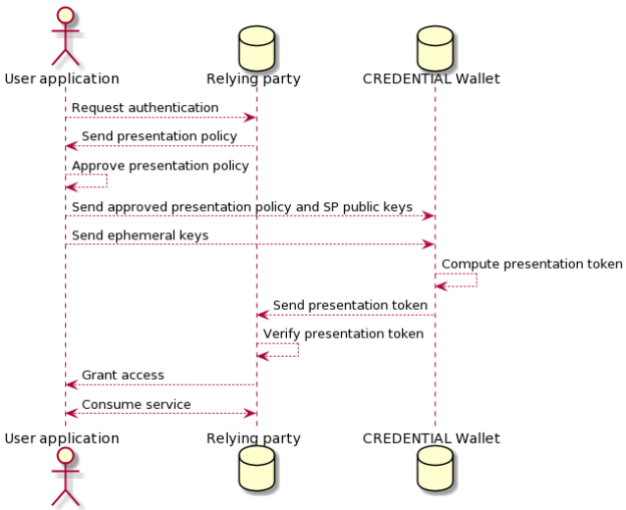


Figure 72: High-level data flow of cloud-based anonymous credentials.

An initial performance analysis showed that all computations of the refined cryptographic libraries can still be carried out in near real-time, causing no usability drawback compared to the initial version.

4.15.4. Future Work

Open research challenges are in particular related to two aspects. Firstly, revocation is a non-trivial challenge because to the best of our knowledge existing revocation approaches could only be efficiently implemented in a cloud-based setting if the revocation handle was revealed to the Wallet as clear text. While this is not an immediate privacy challenge, it might however give the Wallet the possibility to revoke a credential without the user's consent. Secondly, in order to prove predicates over attributes (e.g., "older than 18") requires access to the plain text attribute as well, which poses an immediate privacy risk. A possible way forward could be to perform parts of the computation on the user's device. It would therefore be interesting to develop solutions that allow for proving such predicates over attributes while minimizing the required computations on the user's side.

4.16. Issuer-Hiding Anonymous Credentials

Issuer-hiding anonymous credential systems provide a mechanism for privacy-friendly identity management. They enhance over the state of the art by not only offering means for selective disclosure and data minimization (cf. also asset SS-PP-IdM (Section 4.3) and cloud-based credentials (Section 4.15)), but also allow for hiding the issuer of a certain credential.

4.16.1. Overview

An issuer-hiding anonymous credential system consists of the three main actors:

- **Identity provider**: The identity provider certifies users' attributes and issues anonymous credentials on them.
- **User application**: This is the user's local application required for performing privacy-preserving authentications.
- **Relying party**: Due to the cryptographic specificalities of issuer-hiding credentials, relying parties (e.g., cloud services) need to integrate this end point into their system.

The main difference to other attribute-based credential systems (E.g., SS-PP-IdM or cloud-based credentials) is that all of them reveal the issuer of the credential. This may seem like a natural property upon first glance. after all, the relying party must be able to decide whether it is willing to trust attributes certified by this issuer. However, this is too restrictive in many interesting scenarios. For instance, a national electronic identity might be used to participate in a European wide opinion poll: a priori, there is no need to reveal the citizenship of a participant. Similarly, students might want to use their student identities in different contexts: when requesting access to the university campus, it might be necessary to prove that they are currently enrolled at this university; however, when authenticating towards a cloud service, it might be sufficient to prove that they are enrolled at *some* university in order to get a student discount, but there is no need to reveal the precise university. In issuer-hiding anonymous credentials, the relying party can, in an ad-hoc fashion, decide which issuers it is willing to accept, and the user then shows that she owns a credential that was issued by one of these accepted issuers, without revealing which one.

By following this approach, issuer-hiding credentials immediately also address the problem of revocation of issuers upon corruption. In traditional settings, it would be required to invalidate all certificates that have ever been issued by the given issuer, leading to significant scalability issues in case of, e.g., millions of potentially affected users. With issuer-hiding ABCs, this challenge is solved by allowing for federated systems, where only a limited number of credentials is issued under each specific key, without compromising the users' privacy.

4.16.2. Research challenges addressed

While attribute-based credentials already provide high privacy guarantees, our solution introduces an additional level of metadata privacy and unlikability, which has been identified as one of the key

challenges in the area of privacy-preserving identity management (cf. D4.4, Section 5.5.9, and D3.11, challenge IDP-02). The technique has proved practically efficient in the first basic benchmarks. Furthermore, our approach solves the scalability problem related to the revocation of compromised key material on the issuer side, as it directly enables federated systems, where, e.g., local administrations can use different issuer keys (compared to using a single key nation-wide) without affecting the privacy guarantees for users. For additional use cases, we also refer to Bobolz et al. [BEK+21].

Regarding the challenges identified in D3.11, issuer-hiding ABCs also address the following challenge:

IDP-02: Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks. Issuer-hiding ABC systems reduce the risk of unnecessary information disclosure even one step further than, e.g, cloud-based ABCs. Specifically, besides letting the user selectively decide which information to disclose, they also let the user hide the issuer of her credential, reducing the risk to be re-identified based on this information (cf. the example given before).

4.16.3. Demonstrations Example

Issuer-hiding credentials are a new concept that was first instantiated recently by Bobolz et al. [BEK+21], and consequently, no full-fledged implementation is yet available. We thus will only briefly discuss the relevant processes in the following.

- **User enrolment**: This process would work similar to, e.g., to SS-PP-IdM.
- **Credential issuance**: After having verified the user's attributes, the identity provider would sign the user's attributes using a digital signature scheme. The signature constitutes the credential.
- **Privacy-preserving service access**: In issuer-hiding credential systems, service providers can define so-called *issuer policies*, defining a set of accepted issuers. This is done by signing the public keys of the issuers using an ephemeral digital signature key. These policies can be re-used and can be publicly accessed by all users of the system. In order to authenticate, the user now receives from the service provider which attributes she has to reveal, and which issuer policy she needs to fulfil. Upon consenting, the user's application would compute cryptographic proof that the user owns a credential on the disclosed attributes, which was issued by an issuer whose public key was signed in the defined issuer policy. A high-level flow is depicted below:

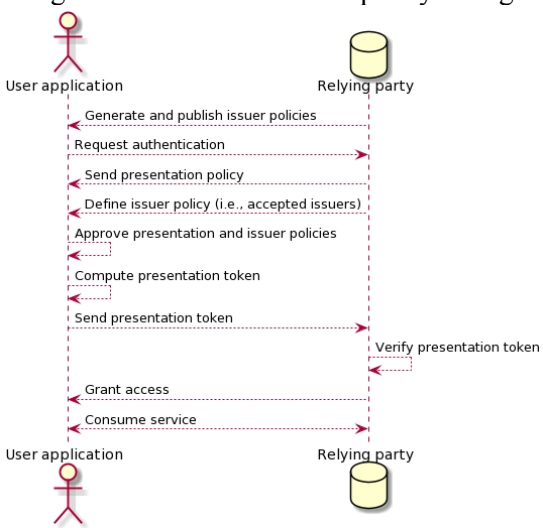


Figure 73: High-level data flow of issuer-hiding anonymous credentials.

Given the novelty of this technology, only standalone implementations and benchmarks could be carried out so far, leading to verification times in the area of 1-2ms on standard computers. The design of the schemes in [BEK+21] is such that the number of computations carried out by the user are independent of the number of accepted issuers, except for the verification of the issuer policy. As policies are re-

usable, they do not need to be verified every time, and we thus consider the amortized costs as negligible. The actual session-dependent computations can all be carried out in well below one second, proving the practicality of our approach.

4.16.4. Future Work

In the current setting, policies defining the set of accepted issuers are linear in the number of issuers. An open question is to overcome this limitation and achieve sublinear policy sizes. This might be achievable by using an accumulator-based approach; however, we are currently not aware of building blocks allowing for efficient instantiations of the relevant zero-knowledge proofs. Another interesting question would be to carry over this concept more generally to the self-sovereign identity (SSI) setting.

4.17. FlexProd and ArchiStar

The goal of the FlexProd platform is to enable computations on sensitive data without compromising the privacy of the data sources, while at the same time giving high integrity guarantees. Besides working on the functionality and efficiency of the existing platform prototype, research performed within CyberSec4Europe during the last months mainly focused on possible extensions to achieve end-to-end integrity and authenticity. The envisioned features are still under analysis and no implementation supporting these features is available for demonstration purposes yet.

4.17.1. Overview

FlexProd is an integrity- and privacy-preserving platform for distributed computations on potentially sensitive data, using ArchiStar libraries as subcomponents. Like Sharemind (cf. Section 4.14), FlexProd is based on secure multi-party computation, which allows multiple computation nodes to jointly perform computations on their respective inputs, without the other nodes learning any information about the other nodes' input. Users can now share their data using the ArchiStar secret sharing libraries, and store one share for each compute node, which can then perform, e.g., statistical analytics on data from potentially numerous users.
The following actors are involved in such a computation:

- **Data sources**: These are individuals or devices contributing data.
- **Compute nodes**: These are the entities performing the secure multiparty computations.
- **Data consumers**: These are the receivers of the computations of the result.

In an ongoing development and research effort, the asset is enhanced to also give the end-to-end integrity guarantees, by letting users sign their data before splitting it for the different nodes. The nodes then jointly perform the computation and furthermore will generate a zero-knowledge proof (potentially a so-called zk-SNARK for efficiency reasons) that all computations were performed correctly, and that the input data was equipped with valid signatures.
For a more detailed description of these assets, we also refer to the respective sections in D3.2 and D3.11.

4.17.2. Research challenges addressed

While a variety of secure multi-party computations exists (cf. Hastings et al. [HHNZ19] for an overview), existing frameworks do not directly cover the end-to-end authenticity and integrity of the results provided by the MPC framework. That is, the privacy and integrity of the computation are always based on the assumption that a sufficiently high fraction of the MPC nodes behaves honestly. With this asset, we aim at overcoming this limitation in the case that authentic (i.e., signed) data is processed by the MPC nodes. In this case, we aim at jointly generating a cryptographic proof (a zk-SNARK) that will allow the receiver to verify that all computations were performed correctly, where even a fully malicious MPC network would not be able to forge a proof. By doing so, all integrity guarantees can be based, e.g., on a physical trust anchor located in the users' devices collecting and generating the data. This

approach could thus be seen as one possible approach towards addressing the integrity and privacy of supply chain information assets, cf. D4.4 (Section 4.5.10).

Regarding the challenges identified in D3.11, FlexProd also addresses the following challenges:

- *DP-02: When using secure multi-party computation (MPC) to analyze data, analysts are not able to see the individual data values.* By design and definition, MPC aims at not revealing the input data to any entity, including the receiver. However, one ambition of this asset is to remove any trust assumptions that need to be put into the MPC network by the data analyst. We believe that by doing so, and by aiming for formal integrity and authenticity guarantees, data analysts will be more willing to rely on the results provided by the MPC network.
- *DP-08: When uploading information to the cloud the user partially loses control over the data.* In collaboration with other projects (e.g., H2020 KRAKEN), the goal is to let the user, on a fine granular basis, specify for which computations her data may be used. Before triggering any computation, the MPC network would then verify that these policies are satisfied, thereby minimizing the risk of data abuse by malicious parties.

4.17.3. Demonstrations Example

FlexProd is based on collaborations with other currently ongoing research and innovation projects and is in parts demonstrated there (e.g., for privacy-preserving auctions), depending on the specific requirements and contexts. Not having finally decided on a use case within CyberSec4Europe, we sketch a simple possible application scenario within the Smart Campus context in the following. We stress that this is a hypothetical example that could be realized using the envisioned asset, but due to the ongoing research effort for achieving end-to-end authenticity, the implementation has not yet been completed.

In order to increase the overall health and wellbeing of their students, the university announces a fitness challenge among the students, in which students can report their daily physical activities (e.g., number of steps) into a system. At the end of the month, the students with the highest activity during the last week can claim a reward.

To guarantee the privacy of their students, the students do not report their activities in the plain but use the ArchiStar secret sharing libraries to decompose the shares and securely transfer them to three servers of non-colluding entities (e.g., the servers could be hosted by a student council, the university, and a public cloud storage). On the other hand, to guarantee the correctness of the provided data, only data coming from certain fitness tokens capable of signing the data are accepted. At the end of the month, students can now access the FlexProd network and, after proving their identity (e.g., using one of the assets described above), can now check in which quantile among all the students they were. In the case that a student was among the most active students, she receives a verifiable certificate from the computation nodes, allowing her to claim her reward.

A possible data flow for this scenario is depicted below:

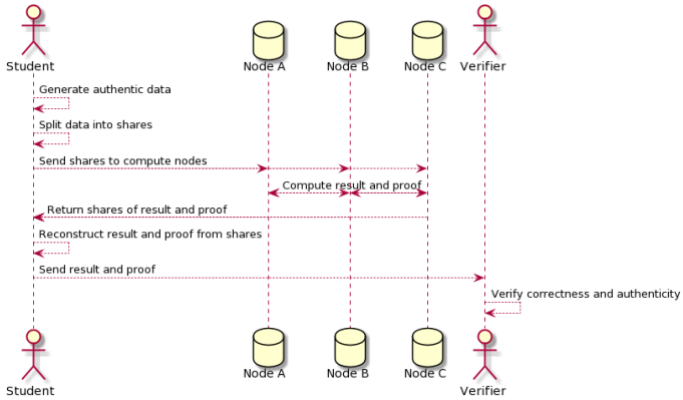


Figure 74: FlexProd high-level data flow.

4.17.4. Future Work

Given the early development stage of FlexProd, many research challenges are still open. The most central challenge is the design of zkSNARKs that would allow for efficient proof generations in a distributed setting. Furthermore, scalability and efficiency of the overall system are still to be analyzed.

4.18. GDPR compliant user experience

GDPR compliant user experience is combined from two sections. The first is the Guidelines for General Data Protection Regulation (GDPR) which present the regulation's requirements through the GDPR principles. The second part of the enabler is the Data Protection Impact Assessment (DPIA) template. As the name suggests, this part of the enabler can be used to help guide the user through the process of doing a DPIA and also serve as documentation for the performed analysis.

A DPIA could be considered in many areas of the Smart Campus scenario as many of the services will use the personal data of the attending students and personnel. There is also a lot of potential for processing of personal data that (these are some of the requirements that can cause the DPIA to be required):

- is likely to pose a high risk.
- involves the use of new technologies.
- involves systematic monitoring.
- involves sensitive personal information (e.g., biometric data).
- refers to a significant amount of personal data at the regional level.
- can affect a large number of data subjects.
- prevents individuals from using the service or contract.
- includes monitoring of publicly accessible areas on a large scale (e.g., public university).

A DPIA must be performed before any type of processing is carried out and is an ongoing process that has to be regularly reviewed and brought up to date. A finished and properly performed DPIA will also help an organization evaluate, document, and later show how they comply with all the personal data protection requirements. A DPIA template could therefore be implemented as a forerunner to many different scenarios within Smart Campus.

The purpose of the DPIA is to systematically analyze, identify and minimize the impact the identified risks could have on the privacy of the data subjects. We will therefore prepare a DPIA example; however, in it, we will limit ourselves to the data collected and later processed in a typical student enrolment process.

4.18.1. Overview

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 was designed to strengthen the rights of individuals across the EU and ensure uniform and coordinated action across the Member States. The GDPR has caused a significant amount of panic and confusion among businesses. This can be attributed to multiple factors, such as high fines, applying to all organizations (as long as they process personal data) equally regardless of size or amount of data they process, often vague or open to interpretation requirements, etc. Therefore, the goal was to create something to help, especially smaller organizations, understand the requirements GDPR demands from them and support them in carrying some of the demanded tasks out.

Guidelines for GDPR Compliant User Experience is a deliverable that was produced as D3.6 in the CyberSec4Europe project. As its name implies, it is a collection of guidelines, best practices and recommendations for achieving GDPR compliance. However, here we will focus on a specific section of the deliverable, which was designed to serve as a template for the process of performing a Data Protection Impact Assessment (DPIA). We will refer to it as the DPIA template. The template is like a to-do list with guidelines on how to perform specific tasks and some pre-prepared structures to support the user. DPIA template is a combination of a guide and pre-prepared content in the form of table templates that personal data controllers can use to perform the DPIA. This solution aims to be primarily of use to the smaller organizations having problems performing or having questions about the assessment's specific steps by giving them a starting point on which they can build.

DPIA is meant to identify and minimize personal data protection risks by systematically analyzing the processing of personal data. Unlike most other risk analyses , DPIA is concentrated on the prevention of harm to data subjects, individuals, and overall society rather than the risk to the organization itself. A DPIA is a legal requirement under the GDPR when the processing is likely to result in a high risk to natural persons' rights and freedoms. This is an excellent example of a condition set by the GDPR for which it is difficult to instinctively know whether it applies or not because there is no definition for "likely to result in a high risk" and the type of issue the enabler is meant to resolve.

The major elements of the DPIA template are presented in Figure 74. The DPIA template aids with the initial decision on the necessity of performing a DPIA. If the circumstances demand the organization to perform the assessment, then the template describes and provides guidelines for the DPIA steps. The "Conduct the self-assessment" (bottom left former in the Figure) is optional and the last step in the DPIA. Before the solution/process can be implemented in the organization, it is important to also make sure all other GDPR requirements are met, which is the purpose of the more broad GDPR compliant user experience enabler. DPIA template contains all the basic information about the assessment as well as many recommendations and good practices on how to perform it. Next, we briefly describe the DPIA template sections where the users can directly utilize the content to perform their own assessment (the template part of the deliverable).

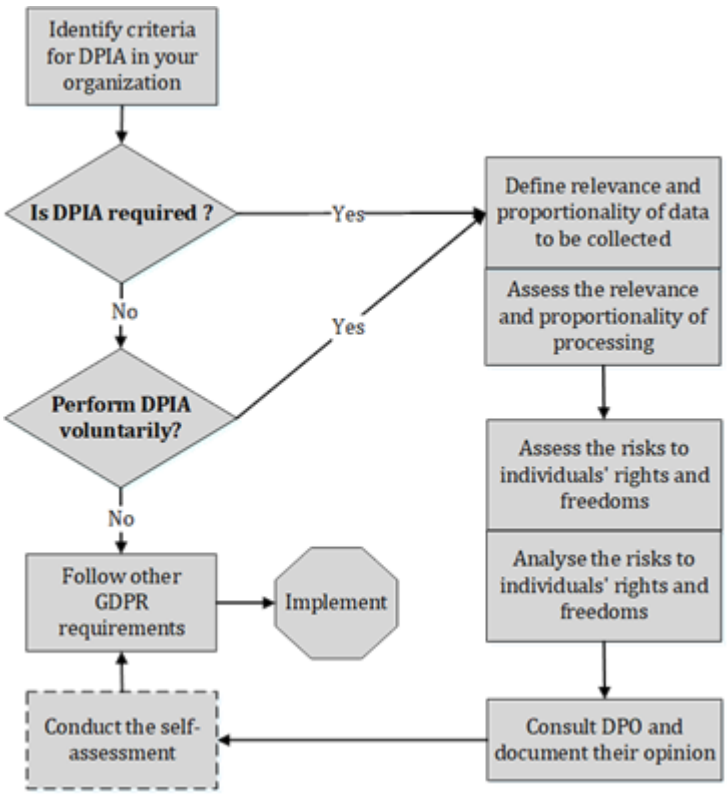


Figure 75: The main steps in the DPIA Template.

The first step of performing a DPIA is determining if a DPIA is required. The DPIA template provides a list of criteria based on the GDPR and recommendations given by the Article 29 Working Party and endorsed by the European Data Protection Board. The template requires the users to answer a few questions about the type of processing they intend to perform. Based on the answers, the template enables the user to make an easy judgment about the necessity of the DPIA, given the users' circumstances. This decision process can also be beneficial to show an organization has performed the DPIA voluntarily. Performing DPIA when not necessary can improve the trustworthiness and reputation of an organization, assist in ensuring that the best practices for data security and privacy are being utilized, and help minimize the organization's liability.

The next major component of the DPIA is focused on risks arising from the processing of personal data. The first step is to establish what type of personal data will be processed, whether data processing is proportional/necessary, for how long it will be stored, and on what legal basis it will be collected. From this information, compliance with the GDPR can be determined. The template helps establish compliance levels based on the collected information. The DPIA template provides instructions on how to measure risk based on the severity and probability of threats. The risk assessment methodology is aligned with the ISO 31000:2018 and ISO 31010:2011 and can be directly used to assign risk levels to all identified threats. The DPIA template provides the users with a template to fill this data into, but more importantly, it already includes a long list of personal data processing threats related to the GDPR requirements. Users can freely add other threats specific to their organizations, circumstances, used processes, etc. Finally, the risk to individuals' rights and freedoms are evaluated.

GDPR requests that the Data Protection Officer (DPO) provides an opinion on the assessment if one is appointed. The template suggests when a consultation with a DPO might be beneficial. These opinions should be documented.

The final part of the template provides a form for self-assessment. This step of the DPIA is optional and not required by the GDPR. The prepared self-assessment can help organizations track the work they have done and learn from it. Based on their performance, they can improve future work on DPIAs for other processes/services.

4.18.2. Research challenges addressed

The content of the assessment can vary depending on different circumstances, and the same structure is not always necessarily the best for everybody. That is why the template is fairly soft on the structure and encourages users to change, expand, and upgrade the given template to suit their own organization requirements and circumstances better. The simple structure designed for smaller organizations, together with the preprepared list of potential risks that have to be addressed in the assessment and the self-assessment form for the users to check and evaluate their work, are the main elements that distinguish this template from other similar tools designed to support the process of performing the DPIA. This work addresses the data privacy challenge DP-01 from the D3.11 (Section 2.1). We have discussed other similar tools in D3.11 (Section 3.1.1) and mentioned some additional ones in D4.4 (Section 5.5.1.6).

4.18.3. Demonstration Example

This asset is demonstrated in the example scenario of student enrolment for the smart campus. For this purpose, we have prepared a full DPIA using the DPIA template from the deliverable D3.6 GDPR guidelines for compliant user experience. The process of student enrolment and legal circumstances have been defined based on the Slovenian (i.e., our local) higher education institution model and national legislation. Specific information and/or justifications in the demonstration example may therefore differ or not be valid in the reader's country. In the DPIA we have limited the content to the typical data processed in the student enrolment and have not extended the example to include all the personal data and processing required by other assets demonstrated in the Smart Campus scenario. The demonstration is done on a legal basis, valid in Slovenia, where personal data collected by higher education institutions is prescribed. This can therefore serve as a basis for the collection of data. Other personal data that would have to be collected for the purpose of smart campus would therefore have to be centered on another basis - e.g. a contract between students and the University/Smart Campus or student consent (although this would not be the best choice as those not willing to give their consent or in case of withdrawn consent, the University/Smart Campus would still need to provide all the services it provides to all other students). The example DPIA is performed as if for a third party that will process the data on behalf of the University/Smart Campus.

The GDPR guidelines for compliant user experience assets provide the DPIA template and instructions on how to use it. The template allows for a more straightforward approach to performing and documenting the DPIA. However, the provided demonstrator of the DPIA documentation for the specific scenario of student enrolment can serve more as a blueprint on what anybody using the template should aim to produce using the enabler. It gives the users a better understanding of how the template is to be filled out and what they are trying to achieve. In addition to this example of a DPIA we will produce a few other demonstration examples of the template's application for different circumstances in other tasks. The result will be a diverse set of examples on how to use the DPIA template and how to perform the assessment for a wide range of circumstances, types of personal data, and types of processing.

The GDPR guidelines for compliant user experience enabler have been produced and finished under the CyberSec4Europe project and is currently not under any further serious development. As such, this enabler will not be demonstrated again in the final round of demonstrators for this task – D3.20 Final cybersecurity enablers and underlying technologies components.

The documented DPIA can be found in a separate file named "Addendum to D3.13 - Student Enrolment Data Protection Impact Assessment"[15].

The GDPR guidelines for compliant user experience enabler have been produced and finished under the CyberSec4Europe project and is currently not under any further serious development. As such, this

[15] https://cybersec4europe.um.si/Addendum%20to%20D3.13%20-%20Student%20Enrolment%20Data%20Protection%20Impact%20Assessment.pdf

enabler will not be demonstrated again in the final round of demonstrators for this task – D3.20 Final cybersecurity enablers and underlying technologies components.

4.18.4. Future work

The DPIA template demonstrated in section 4.18 is getting a small upgrade and an addition. As part of task T3.6, deliverable D3.16, we are extending the list of potential risks that is part of the template. In hopes of making the list of potential risks friendlier to use and reduce the amount of work assessing risks, we are also developing a system to identify implausible risks (dependent on the user's use case) from the list.

4.19. Interoperability and cross-border compliance

The Interoperability and cross-border compliance enabler address issues related to different eIDAS (electronic Identification, Authentication and trust Services) implementations and legislation differences in EU member states, ultimately hampering the idea of a Single European Market. To a lesser degree, the report also looks at the differences between Member states from the perspective of the GDPR.

This enabler is still in development, and therefore there are still some unknowns about its final form. For this reason, we will limit the results from this asset to the GDPR related aspects, which is the smaller part of the whole. Also, because at its core this asset is a report on the current situation regarding the (differences in) implementation of eIDAS and GDPR in the EU, and not a tool to be demonstrated, we will in this deliverable present the relevancy of gathered results for the given Smart Campus scenario.

Within the CCTV Surveillance in the Smart Campus scenario, partners have proposed password-less authentication, which includes the use of biometric data. Under the GDPR, biometric data is considered a special type of personal data when it is used for the purpose of uniquely identifying a natural person. Processing of such data is typically prohibited unless one of the exceptions defined in the second paragraph of the GDPR's Article 9 applies. Member States can also introduce their own conditions and/or limitations.

Before using password-less authentication with biometric data, it is, therefore, necessary to gather information to find out if such use of biometric data is allowed or is restricted in a way that would reduce the functionality of the solution. For such a solution, we look at the results from the asset to ascertain if such data can be used and if there is any additional legislation that governs their use. As mentioned before, such limitations and legislation can change from country to country, so we have, in this case, limited ourselves to the possibility of implementation in Spain.

4.19.1. Overview

The purpose of this enabler is to highlight some of the problems with interoperability and compliance, primarily with the eIDAS. The mentioned trust services include electronic signatures, electronic seals, time stamps, electronic delivery services, and website authentication. Together with electronic identification, they allow for trust, security and legal certainty in electronic transactions. This regulation applies across the entire EU and the 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 to recognize the electronic identification of other Member States as well.

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.

Given the discussed situation, this asset will identify differences between the implementations of eIDAS legislation in the different Member States. These inconsistencies will be shown in selected cases of deployed solutions. 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 main objective of this work is to identify the main characteristics of different implementations and discrepancies that emerge from them. However, this work is still ongoing and will not be demonstrated in this deliverable.

The second part of this asset which will be showcased in this deliverable is on the topic of GDPR employment differences between the Member States. GDPR allows for the Member States to define or change some parts of the regulation in the way they wish. The prime example of this is the consent age, which is in the GDPR set at 16 (persons aged 16 years and older do not require parental consent); however, the regulation allows individual countries to change this and go as low as 13 years old.

For this purpose, we have performed a survey, where we have asked national Supervisory Authorities from each Member State to fill in some information regarding current legislation in their own states. The information-gathering was centered around biometric data, but it did also include other topics such as the age of consent, legislation on storage of data, and whether the country has additional legislation extending the GDPR. In the survey, we have managed to so far receive feedback from 19 (Austria, Belgium, Croatia, Cyprus, Czechia, Denmark, Estonia, Finland, Greece, Germany, Hungary, Latvia, Luxembourg, Malta, Poland, Romania, Slovakia, Slovenia, and Spain) out of the 27 Member States. The results were published on the project's website: https://cybersec4europe.eu/heterogeneity-of-data-protection-legislation-in-the-eu/.

4.19.2. Research challenges addressed

Similar collections indexing GDPR related legislation of the Member States, and their differences have been done and are available. The more notable are [REF4.19.1], [REF4.19.2] and [REF4.19.3]. The first collected relevant legislation from all Member States, but there is no actual comparison, or rather there is still a lot of work on the user to extract the necessary information and compare countries. The second collection is similar to the first; however, it includes fewer Member States but is more specific on what exactly each legislation contains. The last research is the most similar to ours; however, the overlap between it and the data collected here is very small. Our solution, therefore, brings information that (at least to our knowledge) has not been collected and compared in such a way, as well as a very easy way to visually compare differences between the Member States.

4.19.3. Demonstration Example

While the collection of data on certain aspects of how each country adopts/changes GDPR and any additional relevant legislation they have was primarily collected to ascertain where the differences between the Member States are, this same information can also be used to get a rough idea on how different parts of the regulation are applied and if there is any other relevant legislation that should be paid attention to in an individual country. We assume the example of the Smart Campus and the included password-less authentication in the CCTV Surveillance in the Smart Campus scenario to be deployed in Spain; however, from the asset, you would be able to get the same types of information for all of the other Member States for which we have collected the data in the previously mentioned survey.

As we have already mentioned, the GDPR considers biometric data when used to uniquely identify a natural person, a special category of personal data. Processing of such data is prohibited unless one of the specific exceptions applies. Information collected from the Spanish Supervisory Authority (overview in Figure 75 and Figure 76) indicates that they have additional legislation on protection of citizen security[16], national identity document and its electronic signature certificates[17], which includes legislation on collecting of biometric data, photographs, fingerprints and signatures, and regulation on the issuance of the passport[18], which includes legislation on collecting digitized signatures and digitized

[16] https://boe.es/buscar/doc.php?id=BOE-A-1992-4252

17 https://boe.es/buscar/act.php?id=BOE-A-2005-21163

18 https://boe.es/buscar/act.php?id=BOE-A-2003-13978

photographs. The Spanish Data Protection Agency, which is also the Supervisory Authority, has published a document on mistakes in relation to biometric identification and authentication[19] and has made some rulings on the processing of biometric data (primarily video surveillance)[20]. Even though the last two are not legislation, they are both prime examples of how to process biometric data in Spain in a way that is agreeable with their Supervisory Authority. Finally, like most Member States (but not all), Spain allows the use of biometrics in the work environment (with the understanding that all other data protection regulations still apply).

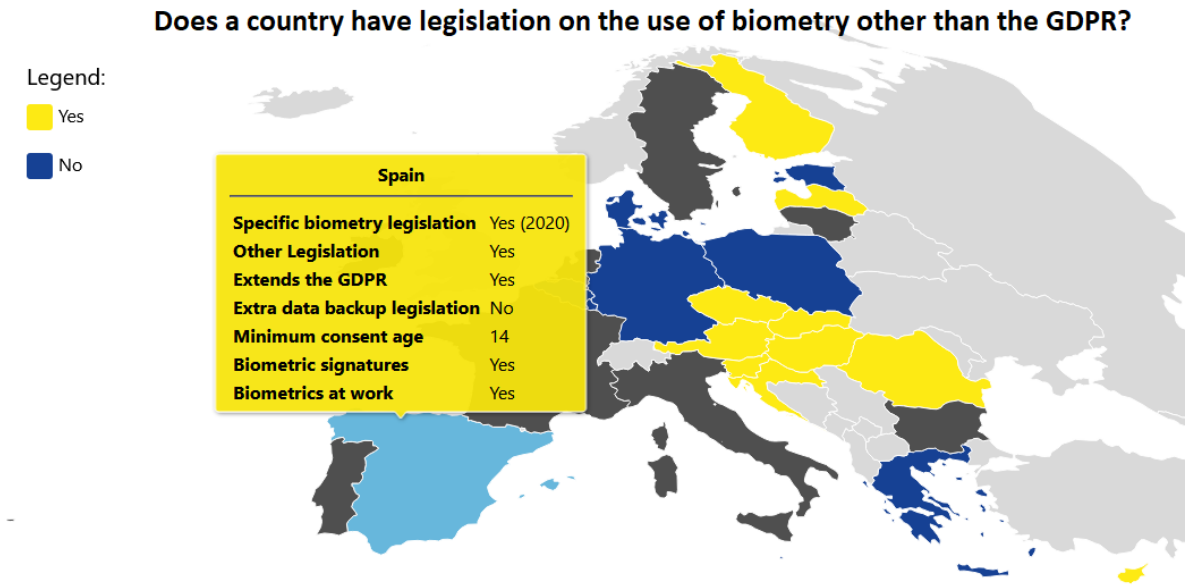


Figure 76: Overview of GDPR related legislation in Spain

Some other information specific to their application and specific legislation surrounding GDPR that could be useful in the context of a Smart Campus includes anonymization[2122], pseudonymization[23], minimum age consent (14 years old)[24] and more specific rules regarding the erasure of data. Spain also allows the use/collection of biometrics in the electronic acquisition of handwritten signatures.

19 https://www.aepd.es/sites/default/files/2020-06/nota-equivocos-biometria.pdf

20https://www.aepd.es/es/prensa-y-comunicacion/notas-de-prensa/AEPD-informe-sistemas-reconocimiento-facial-empresas-seguridad-privada

21https://www.aepd.es/sites/default/files/2019-09/guia-orientaciones-procedimientos-anonimizacion.pdf

22 https://www.aepd.es/sites/default/files/2019-09/nota-tecnica-kanonimidad.pdf

23 https://www.aepd.es/sites/default/files/2020-05/estudio-hash-anonimidad.pdf

24 https://boe.es/buscar/doc.php?id=BOE-A-2018-16673

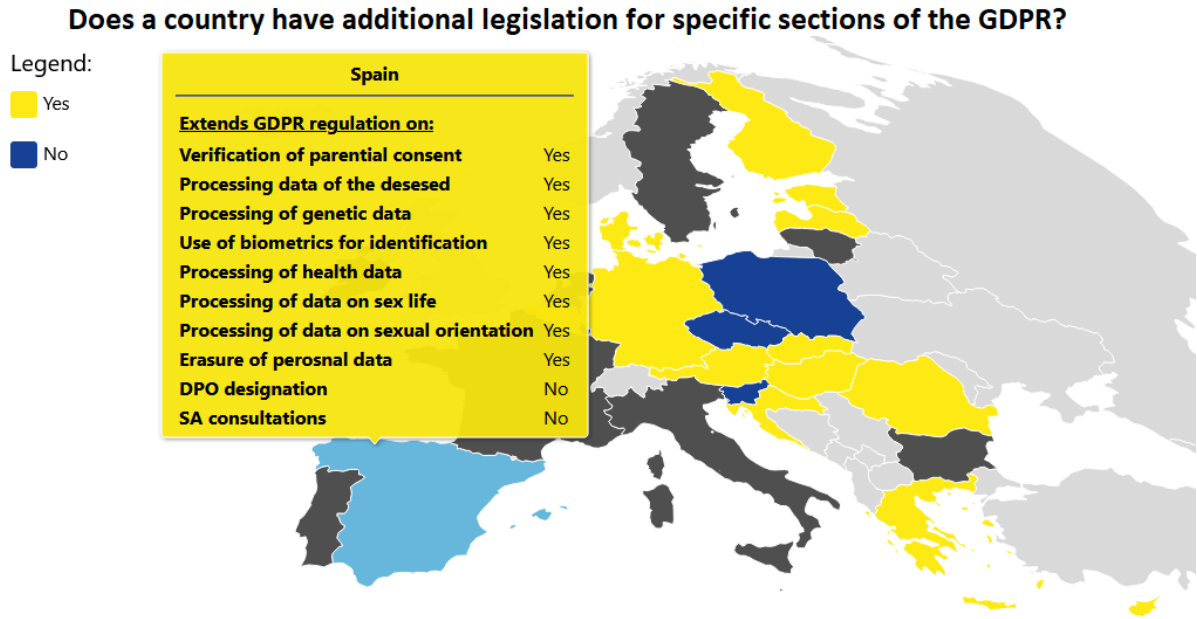


Figure 77: Legislation that extends specific sections of the GDPR in Spain.

4.19.4. Future work

The study on interoperability and cross-border compliance (section 4.19) is still ongoing. The GDPR section is practically finished (presented in this demonstration), but until the end of the project, we plan to add the research on interoperability and cross-border compliance for eIDAS. The results will be published in deliverable D3.18.

5. Conclusions

This document presented D3.13, which is entitled “Updated Version of Enablers and Components”. It focused on CyberSec4Europe’s assets, concerned with privacy and security functionalities, to show an updated version of all assets and the respective integrations between them and with a specific usage scenario. The purpose of this deliverable is to show the usefulness of each one so that readers can use them for future integrations in other usage scenarios as well.

For the integration of these assets, the smart-campus scenario was chosen. This document begins with a detailed description of this usage scenario and respective services where the security and privacy modules must be integrated, namely a video surveillance service, a smart-campus management service (classes, teachers, students), and a geolocation service for statistics.

Then, this document provides a detailed description of the assets that guarantee the security and privacy of the scenario, including the architecture, the research questions addressed, and the use case demonstration example.

This document serves to show a complete overview of the updates on CyberSec4Europe's security and privacy technologies from task 3.2 and their importance in the context of today's research challenges, namely on a concrete use case.

In this document, we mainly address the following challenges of D3.11:

- Data privacy challenge DP-02 (when using secure computing to analyze data, analysts are not able to see the individual data values).

- DP-05 “Lack of mechanisms for controlling and limiting access to the data collected from numerous and geographically disperse IoT device
- DP-06 by providing mechanisms for control how the information is disseminated and control the private data on the communication.
- DP-07 by providing anonymization mechanism for control on how the information is stored and control the private data on the communication.
- IDP-02: Unnecessary over-identification and information disclosure due to a lack of awareness and usability drawbacks.
- IDP-03: User's privacy-preservation of transactions in distributed and immutable systems (e.g., blockchains).
- IDP-04: Honest but curious Service Providers and Identity Providers that might be tracking users, IdP could also become a potential point of failure (identity/data leakage in the IdP).
- IDP-05 by providing a mechanism to guarantee proper identity management of things for authentication end-to-end. DP-08 When uploading information to the cloud the user partially loses control over the data.

In the next deliverable, we will address some open challenges addressed in each asset and also solve some more challenges of D3.11.

6. References

[sari2017study] Sari, Marti Widya, Prahenusa Wahyu Ciptadi, and R. Hafid Hardyanto. "Study of smart campus development using internet of things technology." IOP Conference Series: Materials Science and Engineering. Vol. 190. No. 1. IOP Publishing, 2017.

[mineraud2016gap] Mineraud, Julien, et al. "A gap analysis of Internet-of-Things platforms." Computer Communications 89 (2016): 5-16.

[ngu2016iot]Ngu, Anne H., et al. "IoT middleware: A survey on issues and enabling technologies." IEEE Internet of Things Journal 4.1 (2016): 1-20.

[D3.2] Stephan Krenn, D3.2 –"Cross Sectoral Cybersecurity Building Blocks"

[sousa2021provisioning] Sousa, P. R., Magalhães, L., Resende, J. S., Martins, R., & Antunes, L. (2021). Provisioning, Authentication and Secure Communications for IoT Devices on FIWARE. Sensors, 21(17), 5898.

[OASIS13] OASIS. eXtensible Access Control Markup Language (XACML) Version 3.0. http://docs.oasis-open.org/645xacml/3.0/xacml-3.0-core-spec-os-en.html, 2013

[Hardt12] D. Hardt et al., “The oauth 2.0 authorization framework,” tech.rep., RFC 6749, October 2012

[MBGFSMSPS20] R. T. Moreno, J. Bernal Bernabe, J. García Rodríguez, T. K. Frederiksen, M. Stausholm, N. Martínez, E. Sakkopoulos, N. Ponte, and A. Skarmeta, “The olympus architecture—oblivious identity management for private user-friendly services,” Sensors, vol. 20, no. 3, p. 945, 2020

[BDMCBS20] J. B. Bernabe, M. David, R. T. Moreno, J. P. Cordero, S. Bahloul, and A. Skarmeta, “Aries: Evaluation of a reliable and privacy-preserving european identity management framework,” Future Generation Computer Systems, vol. 102, pp. 409 425, 2020.

[MGBS21] R. T. Moreno, J. García-Rodríguez, J. B. Bernabé and A. Skarmeta, "A Trusted Approach for Decentralised and Privacy-Preserving Identity Management," in IEEE Access, vol. 9, pp. 105788-105804, 2021, doi: 10.1109/ACCESS.2021.3099837.

[Dumortier17] Dumortier, J. (2017). Regulation (EU) No 910/2014 on Electronic Identification and Trust Services for Electronic Transactions in the Internal Market (eIDAS Regulation). In EU Regulation of E-Commerce. Edward Elgar Publishing.

[rios2021personal] R. Rios, J. A. Onieva, R. Roman, and J. Lopez, "Personal IoT Privacy Control at the Edge", IEEE Security & Privacy, vol. 20, issue 1, Early Access. Doi: 10.1109/MSEC.2021.3101865

[roberto2021Eng] Roberto Casadei, Mirko Viroli, Giorgio Audrito, Danilo Pianini, Ferruccio Damiani: Engineering collective intelligence at the edge with aggregate processes. Eng. Appl. Artif. Intell. 97: 104081 (2021)

[viroli2013advance] Viroli M., Damiani F., Beal J. (2013) A Calculus of Computational Fields. In: Canal C., Villari M. (eds) Advances in Service-Oriented and Cloud Computing. ESOCC 2013. Communications in Computer and Information Science, vol 393. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-45364-9_11

[dwork2013alg] Dwork, C., Roth, A.: The algorithmic foundations of differential privacy. Found. Trends Theor. Comput. Sci. 9(3–4), 211–487 (2013)

[suratkar2020crypto] Suratkar, S., Shirole, M., & Bhirud, S. (2020, September). Cryptocurrency Wallet: A Review. In 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP) (pp. 1-7). IEEE.

[niko2021sec] N. Lehto, K. Halunen, O. -M. Latvala, A. Karinsalo and J. Salonen, "CryptoVault - A Secure Hardware Wallet for Decentralized Key Management," 2021 IEEE International Conference on Omni-Layer Intelligent Systems (COINS), 2021, pp. 1-4, doi: 10.1109/COINS51742.2021.9524133.

[gu2017secure] J. Gu, Z. Hua, Y. Xia, H. Chen, B. Zang, H. Guan, and J. Li, “Secure live migration of SGX enclaves on untrusted cloud,” in International Conference on Dependable Systems and Networks, DSN, 2017, pp. 225–236.

[guerreiro2020Tee] J. Guerreiro, R. Moura, and J. N. Silva, “TEEnder: Sgx enclave migration using HSMs”, Computers & Security, 2020

[jan2017ano] Jan Camenisch, Manu Drijvers, Anja Lehmann: Anonymous Attestation with Subverted TPMs. CRYPTO (3) 2017: 427-461

[j2018accurate] J. Decouchant, M. Fernandes, M. Volp ¨ et al., “Accurate filtering of privacy-sensitive information in raw genomic data,” Journal of Biomedical Informatics, vol. 82, pp. 1–12, 2018.

[HK19] Ulrich Haböck, Stephan Krenn: Breaking and Fixing Anonymous Credentials for the Cloud. CANS 2019: 249-269

[K18] Dominik Koehle: ABC for PrivacyImplementing a Demonstrator Model for Privacy-Preserving Authentication in the Cloud. MSc Thesis, 2018

[KLSS17] Stephan Krenn, Thomas Lorünser, Anja Salzer, Christoph Striecks: Towards Attribute-Based Credentials in the Cloud. CANS 2017: 179-202

[BEK+21] Jan Bobolz, Fabian Eidens, Stephan Krenn, Sebastian Ramacher, Kai Samelin: Issuer-Hiding Attribute-Based Credentials. Technical report, currently under submission. 2021

[HHNZ19] Marcella Hastings, Brett Hemenway, Daniel Noble and Steve Zdancewic, “SoK: General Purpose Compilers for Secure Multi-Party Computation,” 2019 IEEE Symposium on Security and Privacy (SP), San Francisco, CA, USA, 2019, pages 1220-1237

[REF4.19.1] Data protection comparison,” activeMind.legal. https://www.activemind.legal/law/

[REF4.19.2] “GDPR Tracker,” Bird & Bird. https://www.twobirds.com/en/in-focus/general-data-protection-regulation/gdpr-tracker

[REF4.19.3] GDPR Derogations Tracker,” Latham & Watkins, Apr. 2018. https://gdpr.lw.com/Home/Derogations

[generalDref1] Said Daoudagh: The GDPR Compliance Through Access Control Systems. PhD Thesis, 2021

[generalDref2] Paolo Barsocchi, Antonello Calabrò, Antonino Crivello, Said Daoudagh, Francesco Furfari, Michele Girolami and Eda Marchetti: COVID-19 & Privacy: Enhancing of Indoor Localization Architectures towards Effective Social Distancing. Array. https://doi.org/10.1016/j.array.2020.100051
[generalDref3] Said Daoudagh, Francesca Lonetti and Eda Marchetti: An automated framework for continuous development and testing of access control systems. J Softw EvolProc. 2020; e2306. https://doi.org/10.1002/smr.23063
[generalDref4] Said Daoudagh, Francesca Lonetti and Eda Marchetti: XACMET: XACML Testing& Modeling. Softw. Qual. J. 28(1): 249-282 (2020)
[generalDref5] Said Daoudagh and Eda Marchetti: GROOT: A GDPR-based Combinatorial Testing Approach. ICTSS 2021.
[generalDref6] Said Daoudagh and Eda Marchetti: GRADUATION: A GDPR-based Mutation Methodology. QUATIC 2021.
[generalDref7] Said Daoudagh, Eda Marchetti, Vincenzo Savarino, Roberto Di Bernardo and Marco Alessi: How to Improve the GDPR Compliance Through Consent Management and Access Control. ICISSP 2021.
[generalDref8] Said Daoudagh, Francesca Lonetti and Eda Marchetti: Continuous Development and Testing of Access and Usage Control: A Systematic Literature Review. ESSE2020
[generalDref9] Paolo Barsocchi, Antonello Calabrò, Antonino Crivello, Said Daoudagh, Francesco Furfari, Michele Girolami and Eda Marchetti: A Privacy-By-Design Architecture for Indoor Localization Systems. QUATIC 2020: 358-366
[generalDref10] Said Daoudagh, Francesca Lonetti and Eda Marchetti: Assessing Testing Strategies for Access Control Systems: A Controlled Experiment. ICISSP 2020: 107-11
[generalDref11] Said Daoudagh and Eda Marchetti: A Life Cycle for Authorization Systems Development in the GDPR Perspective. ITASEC 2020: 128-140
[generalDref12] Said Daoudagh and Eda Marchetti: Defining Controlled Experiments Inside the Access Control Environment. MODELSWARD 2020: 167-176
[generalDref13] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A Framework for the Validation of Access Control Systems. ETAA@ESORICS 2019: 35-51
[generalDref14] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A Decentralized Solution for Combinatorial Testing of Access Control Engine. ICISSP 2019: 126-135
[generalDref15] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A General Framework for De-centralized Combinatorial Testing of Access Control Engine: Examples of Application. ICISSP (Revised Selected Papers) 2019: 207-229
[generalDref16] Cesare Bartolini, Said Daoudagh, Gabriele Lenzini and Eda Marchetti: Towards a Lawful Authorized Access: A Preliminary GDPR-based Authorized Access. IC-SOFT 2019: 331-338
[generalDref17] Antonello Calabrò, Said Daoudagh and Eda Marchetti: Integrating Access Control and Business Process for GDPR Compliance: A Preliminary Study. ITASEC2019
[generalDref18] Cesare Bartolini, Said Daoudagh, Gabriele Lenzini and Eda Marchetti: GDPR-Based User Stories in the Access Control Perspective. QUATIC 2019: 3-17
[generalDref3] Said Daoudagh, Francesca Lonetti and Eda Marchetti: An automated framework for continuous development and testing of access control systems. J Softw EvolProc. 2020; e2306. https://doi.org/10.1002/smr.23063
[generalDref4] Said Daoudagh, Francesca Lonetti and Eda Marchetti: XACMET: XACML Testing& Modeling. Softw. Qual. J. 28(1): 249-282 (2020)
[generalDref5] Said Daoudagh and Eda Marchetti: GROOT: A GDPR-based Combinatorial Testing Approach. ICTSS 2021.
[generalDref6] Said Daoudagh and Eda Marchetti: GRADUATION: A GDPR-based Mutation Methodology. QUATIC 2021.
[generalDref7] Said Daoudagh, Eda Marchetti, Vincenzo Savarino, Roberto Di Bernardo and Marco Alessi: How to Improve the GDPR Compliance Through Consent Management and Access Control. ICISSP 2021.
[generalDref8] Said Daoudagh, Francesca Lonetti and Eda Marchetti: Continuous Development and Testing of Access and Usage Control: A Systematic Literature Review. ESSE2020

[generalDref9] Paolo Barsocchi, Antonello Calabrò, Antonino Crivello, Said Daoudagh, Francesco Furfari, Michele Girolami and Eda Marchetti: A Privacy-By-Design Architecture for Indoor Localization Systems. QUATIC 2020: 358-366
[generalDref10] Said Daoudagh, Francesca Lonetti and Eda Marchetti: Assessing Testing Strategies for Access Control Systems: A Controlled Experiment. ICISSP 2020: 107-11
[generalDref11] Said Daoudagh and Eda Marchetti: A Life Cycle for Authorization Systems Development in the GDPR Perspective. ITASEC 2020: 128-140
[generalDref12] Said Daoudagh and Eda Marchetti: Defining Controlled Experiments Inside the Access Control Environment. MODELSWARD 2020: 167-176
[generalDref13] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A Framework for the Validation of Access Control Systems. ETAA@ESORICS 2019: 35-51
[generalDref14] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A Decentralized Solution for Combinatorial Testing of Access Control Engine. ICISSP 2019: 126-135
[generalDref15] Said Daoudagh, Francesca Lonetti and Eda Marchetti: A General Framework for De-centralized Combinatorial Testing of Access Control Engine: Examples of Application. ICISSP (Revised Selected Papers) 2019: 207-229
[generalDref16] Cesare Bartolini, Said Daoudagh, Gabriele Lenzini and Eda Marchetti: Towards a Lawful Authorized Access: A Preliminary GDPR-based Authorized Access. IC-SOFT 2019: 331-338
[generalDref17] Antonello Calabrò, Said Daoudagh and Eda Marchetti: Integrating Access Control and Business Process for GDPR Compliance: A Preliminary Study. ITASEC2019
[generalDref18] Cesare Bartolini, Said Daoudagh, Gabriele Lenzini and Eda Marchetti: GDPR-Based User Stories in the Access Control Perspective. QUATIC 2019: 3-17
[GMBS21] Jesús García-Rodríguez, Rafael Torres Moreno, Jorge Bernal Bernabé, and Antonio Skarmeta. 2021. Towards a standardized model for privacy-preserving Verifiable Credentials. In The 16th International Conference on Availability, Reliability and Security (ARES 2021). Association for Computing Machinery, New York, NY, USA, Article 126, 1–6. DOI: https://doi.org/10.1145/3465481.3469204

[kaminskas2018agg] Kaminskas L., Lluch Lafuente A. (2018) Aggregation Policies for Tuple Spaces. In: Di Marzo Serugendo G., Loreti M. (eds) Coordination Models and Languages. COORDINATION 2018. Lecture Notes in Computer Science, vol 10852. Springer, Cham. https://doi.org/10.1007/978-3-319-92408-3_8
[argusprivacy]Resende, J. S., Magalhães, L., Brandão, A., Martins, R., & Antunes, L. (2021). Towards a Modular On-Premises Approach for Data Sharing. Sensors, 21(17), 5805.
[panos2017security]Panos, C., Malliaros, S., Ntantogian, C., Panou, A., & Xenakis, C. (2017, September). A security evaluation of FIDO's UAF protocol in mobile and embedded devices. In *International Tyrrhenian Workshop on Digital Communication* (pp. 127-142). Springer, Cham.
[Angelo2021how]Angelogianni, A., Politis, I., & Xenakis, C. (2021). How many FIDO protocols are needed? Surveying the design, security and market perspectives. *arXiv preprint arXiv:2107.00577*.

[vasile2021web]Vasileios Grammatopoulos, A., Politis, I., & Xenakis, C. (2021, August). A web tool for analyzing FIDO2/WebAuthn Requests and Responses. In *The 16th International Conference on Availability, Reliability and Security* (pp. 1-10).