HORIZON 2020
Information and Communication Technologies
Integrating experiments and facilities in FIRE+

Deliverable D1.2
Privacy by design report

Grant Agreement number:	**687884**
Project acronym:	**F-Interop**
Project title:	**FIRE+ online interoperability and performance test tools to support emerging technologies from research to standardization and market launch** **The standards and innovations accelerating tool**
Type of action:	**Research and Innovation Action (RIA)**
Project website address:	**www.finterop.eu**
Due date of deliverable:	**M9**
Dissemination level:	**PU**

This deliverable has been written in the context of the Horizon 2020 European research project F-Interop, which is supported by the European Commission and the Swiss State Secretariat for Education, Research and Innovation. The opinions expressed and arguments employed do not engage the supporting parties.

Document properties

Responsible partner	**University of Luxembourg**
Author(s)/editor(s)	**Luca Lamorte & Maria Rita Palattella (UL)** Eunah Kim (DG) Brecht Vermeulen (iMind) Sebastien Ziegler (Mandat)
Version	**v1.0**
Keywords	**Privacy, Security, Personal Data Protection, Risk**

Abstract

The deliverable **D1.2** – **Privacy by Design Report** describes how security and privacy have been integrated in the design of the F-Interop Architecture, to make the F-Interop platform secure, and trustable by users and contributors. The latter will be authenticated by the F-Interop server, and then authorized to access the testbeds for performing tests on their IUTs, and local implementation. The execution of the tests will imply the exchange of messages and information, specific for the type of test, and protocol implemented. To the final aim of protecting personal or sensitive data in accordance to the EU GDPR, all the data flows exchanged during an F-Interop session have been identified. Guidelines on how to secure personal data and sensitive one (i.e., test results) saved into resource and test repository have been provided as well.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ABC	Attribute Based Credential
CA	Consortium Agreement
CoAP	Constrained Application Protocol
ComSoc	Communications Society
DESCA	Development of a Simplified Consortium Agreement
DHCP	Dynamic Host Configuration Protocol
DHT	Distributed Hash Tables
DNS	Domain Name System
DNSSec	Domain Name System Security Extensions
DPA	Data Protection Authorities
DPO	Data Protection Officer
EC	European Commission
ENISA	European Union Agency for Network and Information Security
ETSI	European Telecommunications Standards Institute
EU	European Union
FP7	Seventh Framework Programme
GA	Grand Agreement
GA	General Assembly
GPS	Global Positioning System
HTTPS	Hypertext Transfer Protocol Secure
ICT	Information and Communication Technologies
ID	Identifier
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IERC	European Research Cluster on the Internet of Things
IETF	Internet Engineering Task Force
IoT	Internet of Things
IP	Internet Protocol
IPC	Intellectual Property Committee
IPM	IPR Monitoring and Exploitation Manager
IPR	Intellectual Property Rights
IPSEC	Internet Protocol Security
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISO	International Standards Organization
ISP	Internet Service Provider
IT	Information Technology
ITU	International Telecommunication Union
KPI	Key Performance Indicator
LSPI	Legal, Security and Privacy Issues
MAC	Media Access Control
MSc	Master of Science
M2M	Machine to Machine
OASIS	Organization for the Advancement of Structured Information Standards
OECD	Organization for Economic Cooperation and Development
OS	Operating System
OSN	Online Social Network

PC	Project Coordinator
PCP	Partner Contact Person
PDPO	Personal Data Protection Officer
PERT	Program Evaluation Review Technique
PhD	Doctor of Philosophy
PM	Person Month
PMB	Project Management Board
PPR	Periodic Progress Report
PRAAT	Privacy Risk Area Assessment Tool
P&T	Post & Telecom
QoS	Quality of Service
RAND	Reasonable and Non Discriminatory
RFC	Request For Comments
R&D	Research & Development
SME	Small Medium Enterprise
SMS	Short Message Service
SOTA (or SoA)	State Of the Art
SSL	Secure Sockets Layer
TC	Technical Coordinator
TCP	Transmission Control Protocol
TL	Task Leader
TLS	Transport Layer Security
Tor	The Onion Router
TRL	Technology Readiness Level
UK	United Kingdoms
UN	United Nations
UNCTAD	United Nations Conference on Trade and Development
UPRAAT	Universal Privacy Risk Area Assessment Tool
URL	Uniform Resource Locator
US	United States
VoIP	Voice over Internet Protocol
WES	Women's Engineering Society
WiTEC	Women in science, Engineering and Technology
WoT	Web of Trust
WP	Work Package
WPL	Work Package Leader
W3C	World Wide Web Consortium
XML	Extensible Markup Language

1 Introduction

1.1 About F-Interop

F-Interop is a Horizon 2020 European Research project, which proposes to extend the European research infrastructure (FIRE+) with online and remote interoperability and performance test tools supporting emerging technologies from research to standardization and to market launch. The outcome will be a set of tools enabling:

- Standardization communities to save time and resources, to be more inclusive with partners who cannot afford travelling, and to accelerate standardization processes;
- SMEs and companies to develop standards-based interoperable products with a shorter time-to-market and significantly lowered engineering and financial overhead.

F-Interop intends to position FIRE+ as an accelerator for new standards and innovations.

1.2 Deliverable Objectives

1.2.1 Work package Objectives

The overall objectives of this work package **WP1 – "Requirements and Architecture design"** are to:

- Analyze and specify the online testing tools requirements,
- Analyze and specify personal data protection and security requirements
- Design and specify the F-Interop architecture

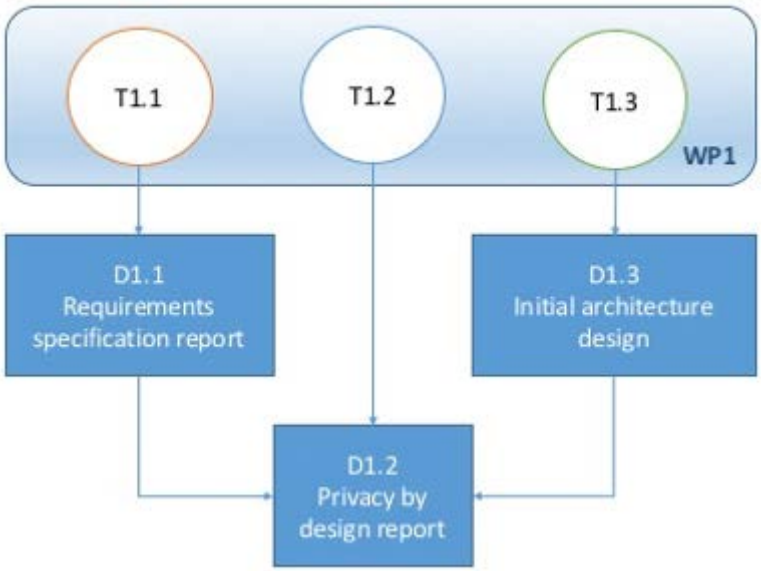

Figure 1 - WP1 Overview

Therefore, WP1 is articulated in three tasks which are closely interconnected. As shown in Figure 1, this deliverable is the main output of Task 1.2.

1.2.2 Task Objectives

Task T1.2 "**Privacy and Security by design**" aims to provide recommendations on the methods to adopt security and privacy in the design of the F-Interop Architecture, and testing tools. This is of foremost importance to make sure F-Interop users/contributors accept to perform their tests on the F-Interop platform. To this end, confidentiality of the tests (messages exchanged during the tests, test tools/scripts), as well as privacy of the test results (passed/failed) should be at least guaranteed. Finally a robust authentication method should be adopted to make sure F-Interop users/contributors are authenticated and authorized, before accessing the F-Interop services.

1.2.3 Deliverable Objectives and Methodology

1.2.3.1 Deliverable Objective

The **Deliverable D1.2 – Privacy by Design** describes how security and privacy have been integrated in the design of the F-Interop Architecture and the testing tools.

1.2.3.2 Methodology

Figure 2 summarizes the Methodology adopted to meet the task's and deliverable's objectives.

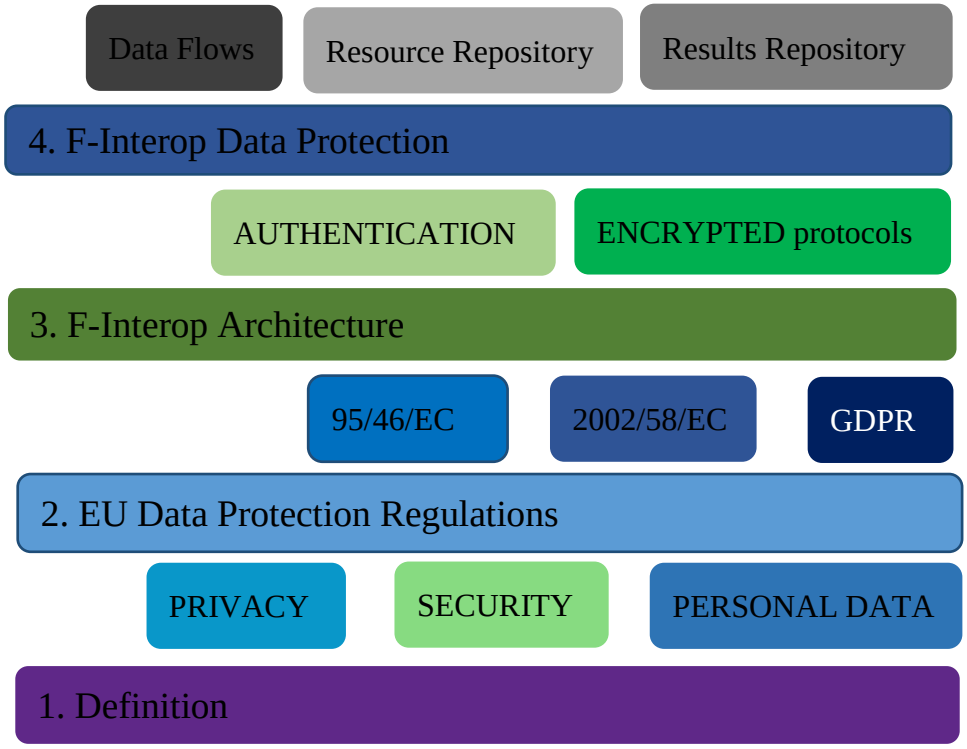

Figure 2 - Main steps of Methodology toward Privacy and Security by Design

Nowadays Internet is widely used for exchanging data, and some of the shared information may be confidential or sensitive for the owner (e.g., individual, company, organization). Therefore, there is a need to protect the data, and provide an acceptable level of *privacy*. Different users may have different interpretations of the meaning of privacy and security.

Therefore, first of all, we provide definitions for the fundamental concepts of privacy, security, personal data and sensitive information.

Then, in order to build F-Interop privacy and security in accordance to the EU Data Protection Regulations, we overview the main ones, such as 95/46/EC, 2002/58/EC and the last defined General Data Protection Regulation (GDPR) that is currently under adoption and will be finally applied in 2018.

As third step, we focus on the F-Interop Architecture and the measurements taken for making it secure, and trustable by users and contributors. The F-Interop server will be authenticated as a user of the various federated testbeds. By doing so, the F-Interop users/contributors can be authenticated as regular users by each federation of the testbeds (Fed4Fire, OneLab, and IoT-Lab), using the AAA (Authentication, Authorization, Access) scheme they implement. Regardless the use of a AAA scheme, F-Interop users/contributors could misuse the federated testbeds, their resources, and the test tools. Therefore, practical management issues, including how to alert and handle such misuses will be defined in the project. Moreover, encrypted Protocols are used as well for making communication channels secure. In detail, both the data messages exchanged between the users and platform (encapsulated into HyperText Transfer Protocol, HTTP, packets), and the signaling messages exchanged during the test execution (encapsulated into Advanced Message Queuing Protocol, AMQP, packets), are encrypted with Transport Layer Security (TLS).

To the final aim of protecting personal or sensitive data, we have finally identified all the data flows exchanged during an F-Interop session. From our analysis it emerges that no sensitive data will be exposed to third party. Regardless this, potential privacy risks, and information leakages will be further evaluated with the privacy tools, developed in WP3, task T3.2.
Resources (e.g., users account, test tool scripts, etc.) as well as test results will be saved in repositories which will be properly protected. Test Results will be saved only for a reasonable period of time, with the final aim of producing statistics, or for allowing end users to compare the test results obtained with different versions of their implementation. F-Interop Users should be free to choose if they do (not) want to save their results, and for how long.

2 Definitions and Fundamental principles

The understanding of principles concerning **privacy,** security, and **data protection** has evolved over the years at the international, European and national level. Privacy and security are two different concepts, even though strongly interconnected.
In this Chapter, we provide a definition for the basic concepts of privacy, security, data protection, sensitive and personal data, which are fundamental for understanding the related EU regulations in this field, which should be followed and respected in the F-Interop Architecture.

2.1 Definitions

2.1.1 Privacy

The term privacy was conceived when people declared to be owners of some properties, and they wanted to defend them from intruders. In computer networks, the term privacy refers to the user's (data owner's) ability to choose which data he/she wants to expose to others, and which data he/she wants to keep from others. Thus, privacy gives users the *control of their data disclosure*. In Internet and computer networks, privacy relies on technical tools for data confidentiality and data integrity. While security gives users the ability to protect the data, and preserve their confidentiality.

The right to privacy was stated in the **Article 7** of the Charter of Fundamental Rights [1] of the European Union, which establish that:

> "*Everyone has the right to respect for his or her private and family life, home and* ***communications****"*

The European Convention of Human Rights [2] enforced this statement with the **Article 8**, which specifies also some protection right constraints.

> "*1. Everyone has the right to respect for his private and family life, his home and his correspondence*". "*2. There shall be no interference by a public authority with the exercise of this right except such as is* ***in accordance with the law*** *and is* ***necessary in a democratic society*** *in the interests of national security, public safety or the economic wellbeing of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others*".

However, over time, the European Union has replaced the concept of "Privacy" by the notion of "Personal Data Protection", whose scope and definition differs.

The newly adopted General Data Protection Regulation, Regulation (EU) 2016/679 [3] of the European Parliament and of the Council adopted on 27 April 2016, defines personal data as:

> *"personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person"*

2.1.2 Security

In information technology, Security is the protection of information assets through the use of technology, processes, and training. According to Article 4 of the Directive 2002/58/EC [4],

> *The provider of a publicly available electronic communications service must take appropriate technical and organizational measures to safeguard security of its services, if necessary in conjunction with the provider of the public communications network with respect to network security. Having regard to the state of the art and the cost of their implementation, these measures shall ensure a level of security appropriate to the risk presented. In case of a particular risk of a breach of the security of the network, the provider of a publicly available electronic communications service must inform the subscribers concerning such risk and, where the risk lies outside the scope of the measures to be taken by the service provider, of any possible remedies, including an indication of the likely costs involved.*

2.1.3 Nature of the data

Any kind of information can be personal data provided that it can be linked to an identified or identifiable natural person. Personal data may cover information pertaining to the private life of a person as well as information about his or her professional or public life. Data relate to persons also if the content of the information indirectly reveals data about a person. In some cases, where there is a close link between an object or an event – e.g. a mobile phone, a car, an accident – on the one hand, and a person – e.g. as its owner, user, victim – on the other, information about an object or about an event ought also to be considered personal data.

2.1.4 Personal Data

According to **Article 4**, section 1) of the GDPR [3]

The statement contains four main building blocks (Article 29 Opinion 4/2007)

a) Any information design a broad concept of personal data

Point of view of the information	
the *nature*	It covers "**objective**" information as well as "**subjective**" information, opinions or assessments
the *content*	This covers personal information considered to be "**sensitive data**" in Article 8 of the directive because of its particularly risky nature, but also more general kinds of information
the *format*	The medium on which that information is contained and whatever form, be it alphabetical, numerical, graphical, photographical or acoustic.

b) Information can be considered to "**relate**" to an individual when it is about that individual.

Elements	
Content	information can be linked to an identified or identifiable natural person through reasonable means, regardless of any purpose on the side of the data controller or of a third party, or the impact of that information on the data subject
Purpose	the data are used or are likely to be used, taking into account all the circumstances surrounding the precise case, with the purpose to evaluate, treat in a certain way or influence the status or behavior of an individual.
Result	the use of the data is likely to have an impact on a certain person's rights and interests, taking into account all the circumstances surrounding the precise case.

c) A natural person can be considered as "***identified***" when, within a group of persons, he or she is "distinguished" from all other members of the group. Accordingly, the natural person is "identifiable" when, although the person has not been identified yet, it is possible to do it.

How	
Directly	the name of the person is indeed the most common identifier
Indirectly	relates to the phenomenon of "unique combinations", whether small or large in size

d) The concept of **natural person** is referred to in Article 6 of the Universal Declaration of Human Rights, according to which *"Everyone has the right to recognition everywhere as a person before the law"*. Personal data are therefore data relating to identified or identifiable living individuals in principle

2.1.5 Special categories of personal data

Under EU law as well as CoE (Council of Europe) law, there are special categories of personal data which, by their nature, may pose a risk to the data subjects, when processed, and need enhanced protection. The processing of these special categories of data **(**better known as **'sensitive data'**) must therefore be allowed only with specific safeguards. A list of what should be considered as Sensitive Data, according to Article 10 of Regulation 45/2001, and Article 8 of Directive 95/46/EC [5], follows:

- National, racial or ethnic origin
- Political opinions or membership
- Religious beliefs or affiliations
- Health information and genetic
- Sexual preferences
- Labor/Trae union membership
- Criminal Records

2.2 Privacy and Security fundamental principles

Privacy involves much more than ensure secure access to data. It is all about control, enabling individuals to keep the control over their personally identifiable information with respect of its collection use and disclosure. Privacy laws are struggling to keep up with the ever-changing landscape lead by the rapid technological change. But this challenge is important to maintain the trust and confidence of the customers of each organization. Technology is not hindered by privacy, but rather, made far better by it.
Information security seeks to enable and protect the activities and assets of both people and enterprises. Its primarily scope is to implement proprietary processes and technologies as a defensive mechanism (basic protection) to protect them.

An approach to Privacy and Security can be defined in 7 fundamental principles [6], summarized in the following table.

Table 1 – Privacy and Security fundamental principles

Foundational Principles	Privacy: Respect and protect personal information	Security: Enable and protect activities and assets of both people and enterprises
Proactive not Reactive; Preventative not Remedial	Anticipate and prevent privacy-invasive events before they happen. Do not wait for privacy risks to materialize.	Leverage enterprise architecture methods to guide the proactive implementation of security.
Default Setting	Build privacy measures directly into any given ICT system or business practice, by default.	Implement "Secure by Default" policies, including least privilege, need-to-know, least trust, mandatory access control and separation of duties.
Embedded into Design	Embed privacy into the design and architecture of ICT systems and business practices. Do not bolt it on after the fact.	Apply Software Security Assurance practices. Use hardware solutions such as Trusted Platform Module
Positive-Sum	Accommodate all legitimate interests and objectives in a positive-sum "win-win" manner, not through a zero-sum approach involving unnecessary trade-offs.	Accommodate all stakeholders. Resolve conflicts to seek win-win.
End-to-End Security	Ensure cradle-to-grave, secure lifecycle management of information, end-to-end.	Ensure confidentiality, integrity and availability of all information for all stakeholders.
Visibility and Transparency	Keep component parts of IT systems and operations of business practices visible and transparent, to users and providers alike.	Strengthen security through open standards, well-known processes and external validation.
Respect for the User	Respect and protect interests of the individual, above all. Keep it user-centric.	Respect and protect the interests of all information owners. Security must accommodate both individual and enterprise interests.

3 European Personal Data Protection Normative Framework

In this chapter we overview the legislation that covers informational privacy of individuals and protection of their personal and contextual data. This study is fundamental to identify afterward the data flows which should be protected in the F-Interop Architecture. The adopted Privacy by Design approach will allow F-Interop Users/Contributors to run their tests without exposing personal data, as well as results of their tests.

3.1 Current Personal Data Protection Framework

Under EU law, personal data can only be gathered legally under strict conditions, for legitimate purpose. Furthermore, persons and organizations, which collect and manage personal information must protect it from misuse and must respect certain rights of the data owners which are guaranteed by EU law. Common EU rules have been established to ensure that personal data gets a sufficient high standard of protection. In what follows, we describe a brief history of the current data protection framework, and we highlight the main points of each directive/article.

3.1.1 Legislation History

In the 1990s, protection of personal data was regulated by non-harmonized laws in the Member States. All of them were following the same basic principles defined by the Council of Europe (CoE) in Convention for the *Protection of Individuals with regard to the Automatic Processing of Personal Data* - **Convention n. 108** (1985) [7].

To ensure a better functioning of EU's internal market a lot of effort was put toward a harmonized environment. Moreover, the explosion of the ICT-field requested a common set of data protection rules. This lead to the adoption, in 1995 of **Directive 95/46/EC** that regulated, for the first time in Europe, the protection of data.

Data Protection Directive 95/46/EC, defined by the European Parliament and the Council, establishes the protection of individuals with regard to the *processing of personal data* and on the free movement of such data. This framework was intended to guarantee the secure and free movement of personal data across the national borders of the EU member countries and to set a baseline of security around personal information wherever it is stored, transmitted or processed.

Since this directive could address only EU Members States, a new legal instrument was necessary to protect the processing of personal data by institution and bodies in EU. This gap was covered by **Regulation EC 45/2001 (EU Institution Data Protection Regulation)** on the protection of individuals with regards to the processing of personal data by the institution and bodies of the Community and the free movement of such data. This was done by adopting the Article 286 of the EC Treaty, in the Regulation.

More detailed data protection provisions, sometimes overlapping Directive 95/46/EC, were introduced to achieve the necessary clarity in balancing other legitimate interests. Two important examples are:

- **Directive 2002/58/EC on privacy and electronic communications** regulates areas which were not sufficiently covered by Directive 95/46/EC, such as confidentiality,

billing and traffic data, rules on spam, etc. The e-privacy directive was amended through **Directive 2009/136/EC**, which became effective in December 2009.

- **Directive 2006/24/EC** on **retention of data** generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC. The directive was invalidated by the Court of Justice on 8th of April 2008 for violating fundamental rights. (Article 7,8 EU Charter of the Fundamental Rights). [8]

Data Protection Directive (95/46/EC), E-Privay Directive (2002/58/EC) and Institution Data Regulation (EC 45/2001) have created a general and technology neutral system of data protection in all EU Member States.

3.1.2 The Data Protection Directive 95/46/EC

The Data Protection Directive 95/46/EC defines the basics elements of data protection that member states must transpose into national law. Each state manages the regulation of data protection and its enforcement within its jurisdiction, and data protection commissioners from the EU states participate in a working group at the community level (according to **Article 29** of the Directive). It sets up a regulatory framework which seeks to strike a balance between a high level of protection for the privacy of individuals and the free movement of personal data within the European Union. To do so, the Directive sets strict limits on the collection and use of personal data and demands that each Member State set up an independent national body responsible for the supervision of any activity linked to the processing of personal data. The key point of the directive are summarized hereafter:

A) The quality of the Data is a set of principles that defines how Personal Data must be (Article 6):
- *Processed* fairly and lawfully
- *Collected* for specified, explicit and legitimate purpose. Further processing for historical, statistical or scientific purposes shall not be considered incompatible provided that appropriate safeguards have been provided by the controller;
- *Adequate*, relevant and not excessive in relation to the purposes for which they are collected and/or further processed.
- *accurate* and where necessary kept up to date
- *kept* in a form which permits identification of data subjects for no longer than is necessary for the purposes for which they were collected or further processed. If data are stored for longer periods for historical, statistical or scientific use, they should be kept either in anonymous form only or, if not possible, only with the identity of the data subjects encrypted.

B) The legitimacy of the processing Personal Data is valid only when one of the following circumstances is valid (Article 7):
- If the *'data subject'* has unambiguously given his or her consent, after being adequately informed;
- if it is needed for a contract, for example billing
- if it is required by a legal obligation
- if it is necessary in order to protect the vital interest of the data subject, for example, processing of medical data of a victim of a car accident;

- if it is necessary to perform tasks of public interests or tasks carried out by government, tax authorities, the police or other public bodies;
- if the *'data controller'* or a third party has a legitimate interest in doing so, as long as this interest does not affect the interests of the data subject, or infringe on his or her fundamental rights, in particular the right to privacy.

C) The processing of special categories of data, needs enhanced protection states due to the categories of personal data ('sensitive data') which, by their nature, may pose a risk to the data subjects, when processed. (Article 8)

D) Information to be given to the data subject must be provided by the controller[1] or his representative with at least the following information: identity of the controller, purpose of the processing, category of data concerned, the recipients, the right of accessing and/or rectify those data. (Article 10,11)

E) The data subject's right of access to data without constraint at reasonable intervals and without excessive delay or expense by the controller (Article 12).

F) The right to object to the processing of data on legitimate grounds, to the processing of data relating to him/her (Article 14)

G) The confidentiality and security of processing : any person acting under the authority of the controller or of the processor[2], including the processor himself, who has access to personal data must not process them except on instructions from the controller, unless he is required to do so by law. A Controller must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing. (Article 16, 17).

H) The notification of the processing to a supervisory authority defines that a controller must notify the supervisory authority … before carrying out any wholly or partly automatic processing operation or set of such operations intended to serve a single purpose or several related purposes.(Article 18)

The Data Protection Directive provides a mechanism by which transfers of personal data outside the territory of the EU have to meet a level of processing "adequate" to the one prescribed by the directive's provisions.

[1] The controller is the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data; where the purposes and means of processing are determined by national or Community laws or regulations, the controller or the specific criteria for his nomination may be designated by national or Community law.

[2] The Processor is a natural or legal person, public authority, agency or any other body which processes personal data on behalf of the controller.

3.1.3 Directive 2002/58/EC

The Privacy and Electronic Communications Directive concerns the processing of personal data and the protection of privacy in the electronic communications sector. It is usually referred to as the "E-privacy Directive" and it was amendment by Directive 2009/136/EC.
It covers processing of personal data and the protection of privacy including provisions on:

- the security of networks and services
- the confidentiality of communications
- access to stored data
- processing of traffic and location data
- calling line identification
- public subscriber directories
- unsolicited commercial communications ("spam").

The main changes to the 2002 Directive include a rule requiring the notification of data breaches (for instance someone whose personal data are lost, modified or accessed unlawfully while being treated by its electronic communications provider should be notified if this breach is likely to affect him/her negatively) and an extension of the Directive to also cover various electronic tags, strengthened enforcement rules, etc.

3.2 New Legal Framework

In January 2012 the European Commission proposed a comprehensive reform of data protection rules in the EU. The aim of the new European Data Protection Regulation is to harmonize the current data protection laws in place across the EU member states, strengthen online privacy rights and boots Europe's digital Economy.

Figure 3 - GDPR validation process

Figure 3 shows the path that the General Data Protection Regulation (GDPR) has been following.

- On 25th January 2012 the European Commission presented a draft.
- In March 2014 the European Parliament adopted an amended version
- In June 2015, the Council adopted the amended release as well.
- In December 2015, the three members of the European Parliament released the final version of the Regulation after six months of negotiation.
- At the beginning of 2016, the formal approval was followed by the official publication of the Regulation.
- 2016-2018: A transaction period of two years will allow organizations and governments to adjust to the new requirements and procedures.
- Following the end of this transitional period, the Regulation will be directly applicable throughout the EU, without requiring implementation by the EU Member States through national law.

In fact, being a "Regulation" instead of a "Directive", it will be directly applicable to all EU member states without a need for national implementing legislation. This should increase legal certainty, reduce the administrative burden and cost of compliance for organizations that are active in multiple EU Member States, and enhance consumer confidence in the single digital marketplace.

3.2.1 Impact of the EU Data Protection Regulation

EU Data Protection contains strong requirements that the *data processor* must be compliant in order to improve consumer rights, but also to enhance transparency and accountability.

- It is mandatory to clarify and potentially expand the concept of **Personal data**
  - for example, various numerical online identifiers that can be connected to a person(s) would be considered personal data.
- Introduce tighter controls on consent as a ground for lawfully processing personal data
  - consent must be 'explicit' with onus on the controller to prove it was properly given.
- Introduce a new "transparency" principle defining types of information which have to be given. The data provided to the consumers has to be in a simple and understandable form. Privacy notes have also to be explicit in order to provide information, for instance, the frequency of data collection, the usage, etc.
- Make data controllers more accountable, and increasing responsibilities of data processors
  - e.g. with regards to security, via obligations for privacy by design and default, or requirements to prove that they have obtained consent, etc.
- Increase sanctions for data protection violations and giving organizations
  - (e.g. digital rights and consumer organizations) the right to complain or take collective action for data breaches on behalf of individuals.
- Data breach notification - and prescriptive high fines for breach.
- Privacy impact assessments for processing that is considered risky. This is also incidentally contained in the EU Recommendation.

4 F-Interop Privacy and Security Approach

Security attacks have been increasing in frequency and sophistications during the years. For this reason, a *proactive* and *preventative* approach should be pursued (rather than a reactive one) during the design of IT Platforms, such as F-Interop. As suggested by Gartner in the discipline of an Enterprise Architecture (EA) [9], a strategic security view should be foreseen *at design time* rather than responding to threats as they arise just with tactical actions. This mind-set follow the first principle of Privacy and Security by Design: "Proactive not Reactive; Preventative not Remedial" (see Table 1, Sec. 2.2).
For the same principle, privacy cannot be added on an ICT system after-the-fact, e.g. by adding a "compliance layer" on top of its core functionality to address relevant privacy legislation. Rather privacy must be proactively embedded into the design and architecture of an ICT system.
In this chapter we describe how security and privacy have been taken into account since the first design of the preliminary F-Interop Architecture.

4.1 Security of F-Interop Architecture

F-Interop aims to integrate and extend several European testbeds (in particular Fed4Fire, OneLab and IoT-Lab), federating them with a shared "Testbed as a Service" platform.

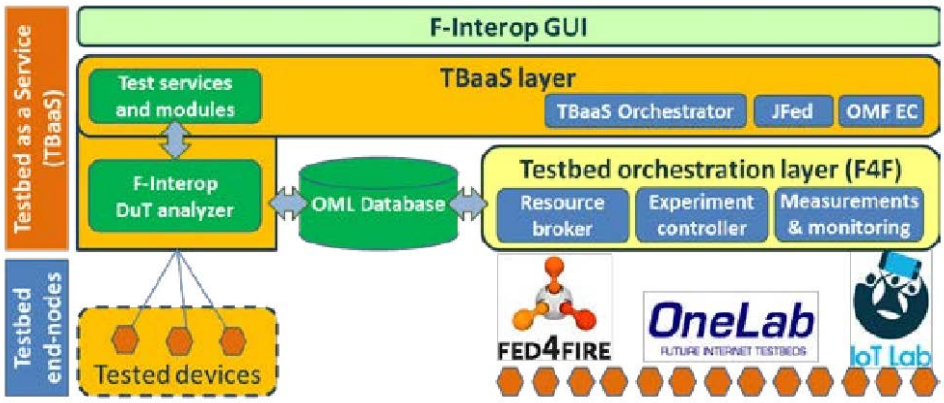

Figure 4 - Initial envisage F-Interop Architecture

F-Interop will build on the security and privacy by design strategies already adopted in the federated testbeds. All F-Interop modules and especially those that expose functionalities outside the architecture boundaries will be protected at least with the same security measurements, or with more solid ones, as needed.

The initial envisaged architecture (shown in Figure 4) has been organized in three main modular blocks, illustrated in Figure 5:

1. F-Interop Central Services
2. F-Interop Orchestrator and Testbeds
3. User premises

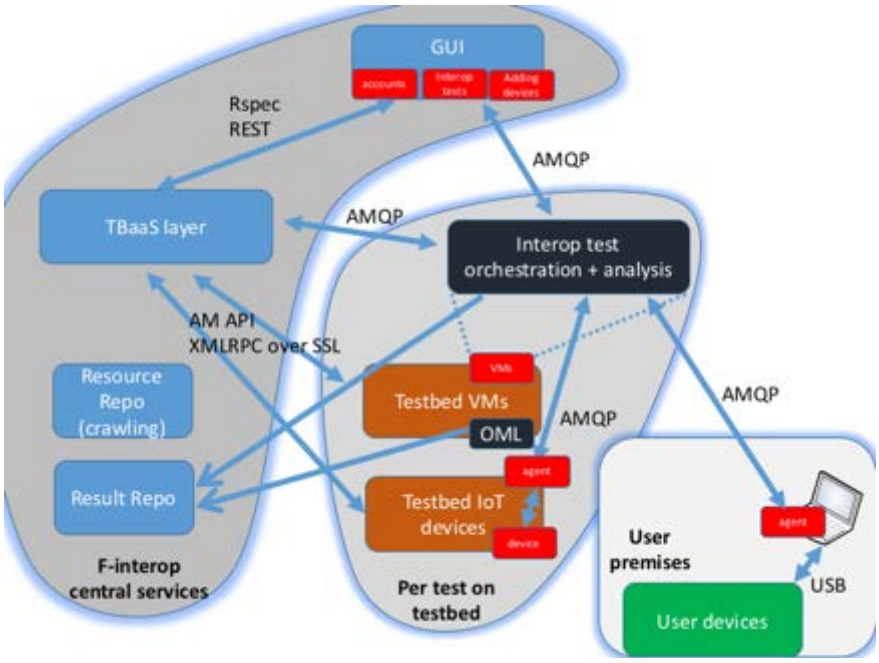

Figure 5 - Architecture intra-module communication overview

The *F-Interop central services* contain the GUI to access the Testbed as a Service (i.e., TBaaS layer), as well as the repositories where resources and test results will be saved. As containers of confidential information (such as users credentials/accounts, and test results), the repositories should be properly protected from potential attacks, and any data leakage which could expose user personal data.

The *F-Interop Orchestrator and Testbeds* block is responsible for the allocation and instantiations of the different resources and tools required to perform a remote test. Several data streams might be exchanged outside this block, if one or more user(s)/device(s) are localized outside the testbeds. Hence, a secure channel should be established for the data flows. Beside that all the users will be authenticated before being authorized to access the F-Interop resources and test tools. .

The *F-Interop User premises* block simply consists in "the User laptop", which the end user will use to connect from any location with any IUT device, and to perform remotely conformance and performance tests on the FI-Interop Platform.

Each block can be executed in a different physical location, and secure communication channels should be established to protect data exchanges among them Figure 5 shows (with blue arrows) all the communication among the modules, as well as the different protocols used for communicating each other. Being the scope of this deliverable the (proactive) privacy and security by design Approach of the F-Interop Architecture, hereafter we focus on the cross-border communication (i.e., communication among different blocks), being these, the ones more expose to potential security risks. Table 2 summarizes the main cross-border communications among the different blocks, together with their level of risk.

Table 2 – Cross border communications

End-Point A	End-Point B	SCOPE	PROTOCOL	RISK
User Device (Web)	FI-GUI (MySlice)	Authentication, Experiment Configuration and Execution, Result Consultation, Sharing resources	HTTPs REST /SSH	Low
User Agent	Interop Test Orchestration	Experiment Protocol packets, Experiment Control messages	AMQP	High
TBaaS Layer	Interop Test Orchestration	Experiments lifecycle management (Initiation, Execution, Termination)	AMQP	Low
TBaaS Layer	Testbed VM/ IoT	Reserving and instantiating remote resources (testbed),	XMLRPC over HTTPS	Low
Interop Test Orchestration	Result Repository	Store partial and final experiments session and result	HTTPs REST/ SQL	High
Testbed VMs OML	Result Repo	Store partial and final experiments session and result	HTTPs REST/ SQL	High

F-Interop modules are basically using three application protocols to vehicle different type of messages:

- HTTPs REST API
- Advanced Message Queuing Protocol (AMQP).
- XMLRPC over HTTPS

In what follows, we briefly describe both of them in more details.

4.2 HTTPS REST API

REST or REpresentational State Transfer is a kind of architecture design largely used to build services on top of the Web. In the Rest approach everything is considered as a resource that is exposed by a Rest Server and accessed by a Rest Client though simplified URLs.
REST is also a lightweight alternative to mechanisms like RPC (Remote Procedure Calls) and Web Services (SOAP, WSDL, et al.).

The privacy and the integrity of data exchanged between the REST client and the REST server MUST be obtained through the use of HTTP within a connection encrypted by the Transport Layer Security (TLS) or its predecessor, the Secure Sockets Layer (SSL). It provides a bidirectional channel encryption that protects the typical Eavesdropping, Tampering and Man In the Middle (MIM) attacks.

There are well known best practices to define REST APIs. Some of them are addressing security aspects. For example, a very simple but effective rule is: Do not use "physical" URLs. A physical URL points at something physical. Much more important to bear in mind is that a RESTful API **should be** stateless. This means that request authentication should not depend

on cookies or sessions. Instead, each request should come with some sort of authentication credentials.

There are several options for authenticating with the API:

Method	**Description**	**Pros**	**Cons**
Basic Auth	Username & Password passed in the API request	Simple to implement	every message contains credentials No way to logout the user
OAuth	A delegation Protocol useful for conveying authorization decision across a network.	Use a large supported mechanism.	More complex to obtain the access.
Access Token (AT)	A valid AT should be provided in the API to prove that the user has granted permission	AT can expires and revoked. Different grant access to resources can be associated with AT	A complex infrastructure is required to generate and control the generation and validation of the AT

Table 3 – Authentication methods

4.3 Advanced Message Queuing Protocol (AMQP)

The Advanced Message Queuing Protocol (AMQP) is an open standard for passing business messages between applications or organizations.

By complying to the AMQP standard, middleware products written for different platforms and in different languages can send messages to one another. AMQP addresses the problem of transporting value-bearing messages across and between organizations in a secure manner.

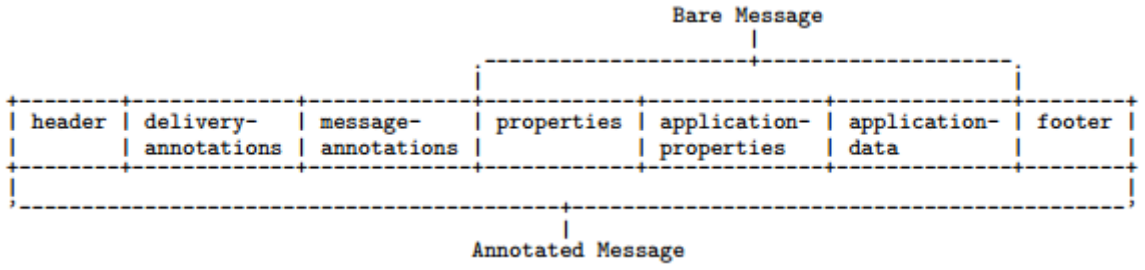

Figure 6 - AMQP Message

F-Interop is going to use RabbitMQ [10], an Open Source message broker software, to distribute AMQP messages around some of the architecture components. It is written in

Erlang [3]and it is distributed under the Mozilla Public License. RabbitMQ has inbuilt support for Transport Layer Security (TLS). This means that AMQP messages exchanged by the peers are secured, because the client and the server, or better the publisher and the consumer of the messages are using TLS to authenticate each other and then use it to encrypt messages between the authenticated parties. TLS protocol is based on public key cryptography.

RabbitMQ has pluggable support for various Simple Authentication and Security Level (SASL) authentication mechanism. The server has three built in mechanisms:

- PLAIN: (default) a simple clear text password mechanism.
- AMQPLAIN: non-standard version of PLAIN defined by AMQP 0-8 spec [11].
- RABBIT-CR-DEMO: challenge-response authentication. Security equivalent to PLAIN

4.3.1 Overview of SSL/TLS Encryption

The main purpose of this protocol is to encrypt data sent over the network. The SSL/TLS security protocol is layered between the application protocol layer and the TPC/IP layer, where it can secure and then send application data to the transport layer. Due of its position, it can support multiple application layer protocol.

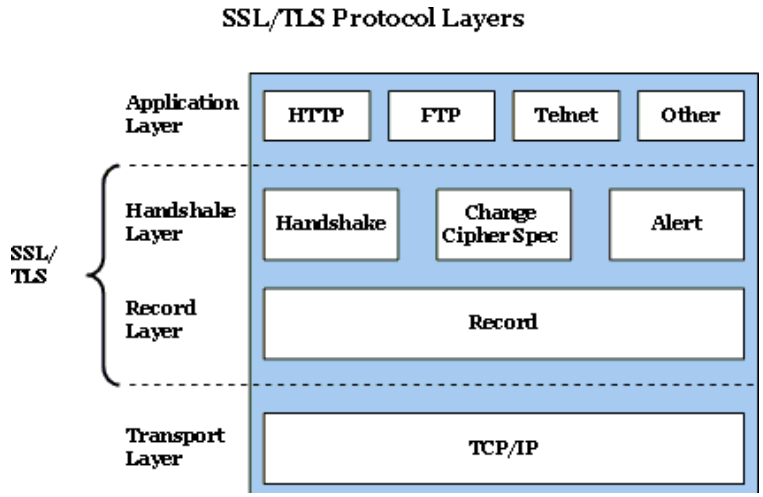

Figure 7 - SSL/TLS protocol layers.

The SSL/TLS protocol can be divided in two layers:

- Handshake Layer: negotiate the session, key encryption, change status messages.
- Record Protocol Layer: receives and delivers encrypted data.

In the Handshake sub-protocol provides a number of very important security function. It is responsible for the authentication and negotiation of the encryption phase.
For authentication purposes, it uses an X.509 certificate to provide strong evidence to a second party that helps to prove the identity of the party that hold the certificate and the corresponding private key.

[3] Erlang is a general-purpose, concurrent functional programming languages. – www.erlang.org

A certificate is a digital form of identification that is usually issued by a Certification Authority (CA) and contains identification information, a validity period, a public key, a serial number, and a digital signature of the user.

4.3.2 XMLRPC over HTTPS

The base of this protocol is based on the same SSL/TLS protocol as discussed before. XMLRPC is a remote procedure call protocol which uses XML to encode its calls and arguments. Security and privacy wise HTTPS is the main protocol here.

In the AM API which is used to talk to the testbeds, also client based authentication is used to set up the SSL connection. This means that a client needs to use a signed certificate and accompanying private key to initiate the connection. The server side verifies the certificate against a set of trusted root certificates[4].

4.4 Security Modules Requirements

Authentication is the first security measurement taken by the F-Interop Platform. To be able to run a remote online test on one of the federated testbed, F-Interop users/contributors MUST first of all being authenticated by the F-Interop Server.

Authorization is the second level of the data security. In fact, authenticated users can be authorized to have access to all the F-Interop modules needed for running their own experiments, i.e., associated resources, repositories, test logs and results. As a general rule, a given user MUST have access ONLY to these resources, and not to those associated to other F-Interop users' tests. In particular test results will not be open to public, and each user should be able to access only the results of his/her own experiments.

Rules or filter mechanisms MUST be implemented by F-Interop to avoid data leakage and data manipulation from not authorized entities.

F-Interop should also take into consideration practical management policies when discovering abuses of the use of shared resources (IUT, Testbed, Tools). A notification mechanism has to be implemented to inform the F-Interop user about the issue raised. Based on the severity of the action, strong decision should be made like blocking or banning temporally that user from the platform.

[4] See also https://fed4fire-testbeds.ilabt.iminds.be/asciidoc/federation-am-api.html#_xml_rpc_over_https_with_client_authentication

5 Data Flow in F-Interop Session

The execution of any test in F-Interop Platform has been articulated in several steps, which together constitute the so called F-Interop Session, described in detail in Deliverable D1.1. These steps represent the lifecycle of an experiment executed at least by one remote user running a specific test on the F-Interop platform. Not all those steps require a direct interaction of the user. In some cases, internal modules interact as a consequence of the user configuration or his/her previous actions.

At each step of the Session, data flows are exchanged among the different modules of the F-Interop Architecture (see Sec. 4.1).

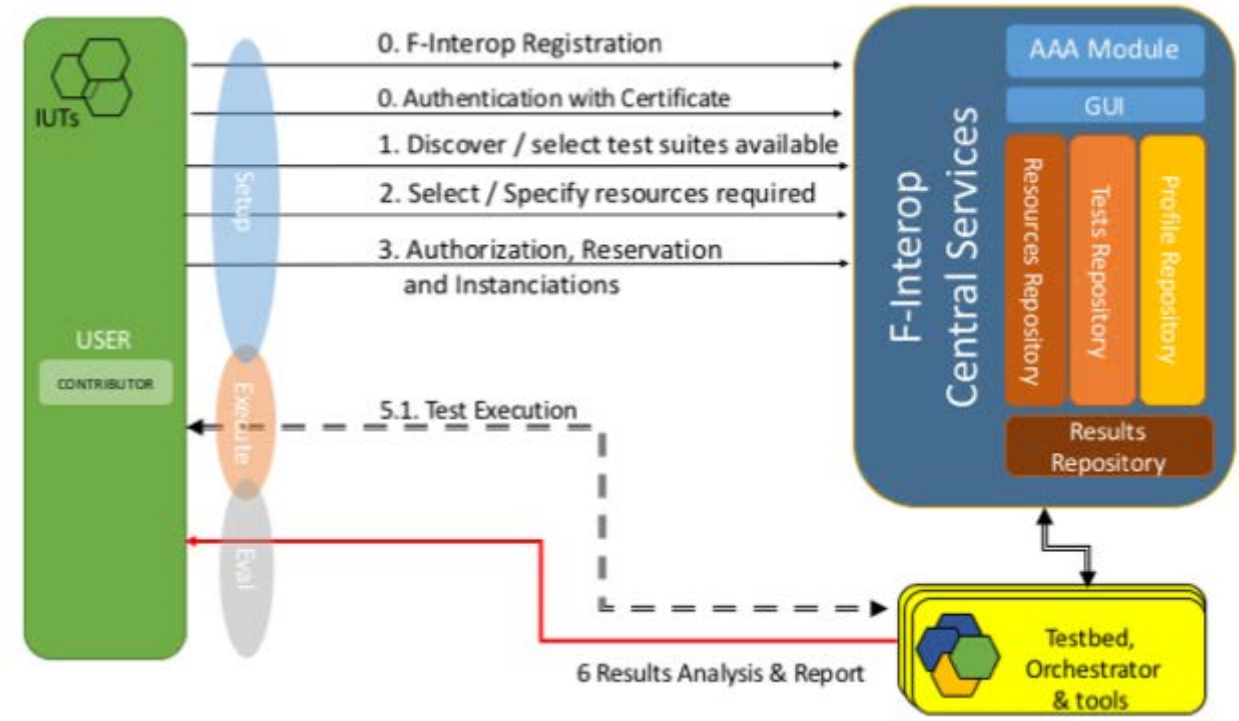

Figure 8 - F-Interop Session

Figure 8 shows the actions performed by the user against F-Interop Modules.

During a session, at least two different Data Flows will be established:

- Control Flow
- Data Flow

Control flow refers to all commands (i.e., control messages) exchanged to orchestrate and manage an experiment, together with any other information needed to execute remote operations on F-Interop platform.

Data flow is the real stream packets generated by an experiment, usually associated to a specific protocol under test (e.g., CoAP, 6TiSCH, etc.). Source or destination of these streams are generally those IUTs involved in the experiment. In particular, such data flow is

exchanged between the IUT that the user wants to test, and another IUT belonging to the F-Interop federated testbeds, or to another end user (depending on the scenario).

To the final aim of providing privacy by design, and thus, data protection, in the following sections, we first identify the type of data exchanged among the modules involved for each step of the F-Interop Session, and then we provide guidelines on the security/privacy measurements that should be taken in order to avoid data leakages and platform breaches.

5.1 User Authentication and Authorization

All the F-Interop users/contributors will be first of all authenticated before being able to access the F-Interop Platform, and perform any test. Authentication is the procedure with which a person is able to prove his/her identity. Authentication is performed asking the user for information that should be known only to him/her, such as a personal identification number (PIN) or a password. If the provided information is correct, then the user can be authenticated, and he/she receives a session key or token. The authenticated user can then have access to the resources, according to a given policy. This process is known as Authorization.
Following the same approach of the federated testbeds (Fed4Fire, IoT-Lab, and OneLab), the F-Interop Platform will use a *Public Key Infrastructure (PKI) with X.509 Certificates* [12] to authenticate and authorize users. Authenticated and authorized users/contributors will be able to perform remote online experiments, and they will have access to resources as well as their own test results, stored respectively in the resource and results repositories.

5.1.1 Public key cryptography

As discussed in Chapter 4, to provide security by design, each communication exchanged among the different F-Interop modules has to be encrypted. But encryption without authentication does not protect data in transit. For connections to be safe, first each party needs to prove his identity to the other. One methodology that enables this is HTTPS's Public Key Infrastructure (PKI). PKI is based on digital certificates and public key cryptography. A Certificate lets a user prove his/her identity. Practically speaking, a certificate is a file with some identity information about the owner, a public key, and a signature from a Certificate Authority (CA). Each public key has an associated private key, which is kept securely under the certificate owner's control. The private key can be used to create digital signatures that can be verified by the associated public key.

A certificate is a powerful tool for proving the user identity online, therefore it is a perfect approach for remote test sessions. The owner of a certificate can digitally sign data, and a verifier can use the public key from the certificate to verify it. The fact that the certificate is itself digitally signed by a third party CA means that if the verifier trusts the third party, they have assurances that the certificate is legitimate. The CA can give a certificate certain rights, such as a period of time in which the identity of the certificate should be trusted.

Sometimes certificates are signed by what's called an intermediate CA, which is itself signed by a different CA. In this case, a certificate verifier can follow the chain until they find a certificate that they trust, the root. This chain of trust model can be very useful for the CA. It allows the root certificate's private key to be kept offline and only used for signing intermediate certificates. Intermediate CA certificates can be shorter lived and be used to sign

endpoint certificates on demand. Shorter-lived online intermediates are easier to manage and revoke if compromised.

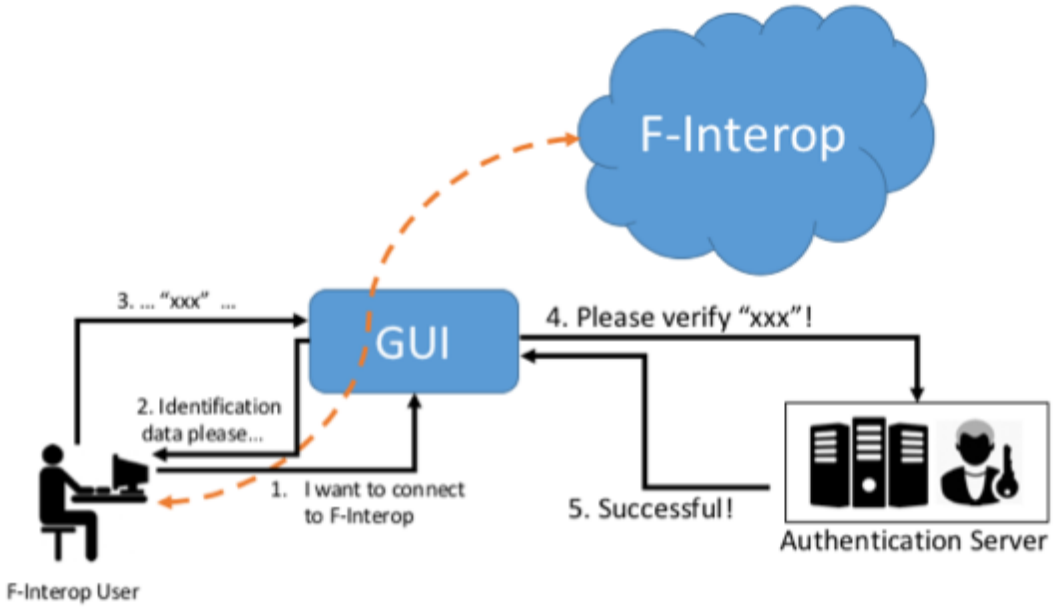

Figure 9 - User Authentication

5.1.2 F-Interop User Registration and Authentication

Each federation of testbeds (Fed4Fire, OneLab, IoT-Lab) which is part of F-Interop Platform, is currently associated to a specific distinct Certificate Authority (CA) that can grant access to users on a specific testbed. This means that currently User Profiles Repositories are separated and, isolated. Thus, a user registered on one testbed cannot access another one, using the same certificate. This raise the need of a new CA that creates a common certificate for any user that want to access F-Interop, and a chain connection with the other CAs.
By doing so, a registered and authorized F-Interop user will be able to access with the same credential not only the test tools, but also all the resources, regardless the federation of testbeds to which they belong.

The authentication form to create a new F-Interop User, will require the following minimum set of identity information:

Information	Category	Risk	Comment
Name	Personal Data	High	Identify a specific user
Surname	Personal Data	High	Identify a specific user
Email	Personal Data	High	Identify a specific user
Institute or company	Personal Data	High	Identify a specific user
Username	Personal Data	High	Identify a specific user
Country	Location	Low	Identify a location
State	Location	Medium	Identify a location
City	Location	High	Identify a location
Password	Security	High	Strictly confidential
SSh Public key	Security	Low	Used in PKI

Table 4 – Registration Form

The Certificate Authority (CA) will use such information to generate a personal and unique certificate that will uniquely identify the user inside F-Interop. With it, F-Interop will grant access to resources, tools and testbeds.

The user before providing such information has to explicitly agree on the terms of conditions, acceptable and permitted inside F-Interop. Therefore, the registration page MUST contain a link that guides the user to the Terms of Conditions page in order to make him aware of that. As Controller, the CA MUST provide information regarding the processing of personal data in a concise, transparent, intelligible and easily accessible form, as specified in GDPR.

The information provided by the user MUST be stored in a secure way in a Repository under the responsibility of the Certificate Authority. F-Interop will be tightly coupled with it during the validation process to authenticate the User and let him access all the available resources.

5.2 Discover and Select an Experiment

Once the user is authenticated, and logged into the F-Interop Platform, he/she can select the tests/experiments he/she would like to perform. To each test F-Interop associate a *Slice* [13], the unit of isolation of an experiment. Only users that are part of a slice can make changes to experiments in that slice. A Slice is also the container for resources (IUTs).

Following the same approach adopted in OneLab testbed, in F-Interop an experiment will be managed by an improved version of MySlice, an user-centric tool that support user's interaction with the federation of testbeds. The software offers both a web interfaces and a programmable API through which users can manage their slices.

If a new experiment is created, the first think to define is the type of test to perform in the slice. A set of pre-configured test-case will be provided by F-Interop through the GUI, but each user will be able to extend them or creating a new one from scratch.
A test case defines the set of events and responses that IUTs involved in the experiment should implements. The full behavior is described in a file called Test Spec to be stored in a specific repository inside the F-Interop Central Service.

In the repository a set of correlated information SHOULD be stored, like the following:

- **Ownership**: the user ID that has created the TestSpec and has the admin role.
- **Slices**: the list of slices that are using that TestSpec
- **Visibility**: define who can have access to that file
    - i.e. private, protected, public, read-only
- **Timestamp**: the date when it was created
- **Last-update**: the date of the last modification
- **Version**: a number to define a version of the TestSpec
- **Description**: a short description to help the user to understand the general behavior
- **Tags**: a list of keywords that will help a future search mechanism.

The experiments and their information MUST be organized in the repository providing the right access control and the security check. The My-Slice extension which will allow the user to select an experiment, SHOULD control which TestSpec a user can see and what kind of right he/she has on it. Moreover, it SHOULD be the only component of the platform that can access the Repository in order to keep secure that data and avoid malicious behavior.

5.3 Select/Specify Resource required

Once a user has created the experiment or joined an existing one he/she can then add resources in the associated slice. As mentioned before the slice is the container for the users sharing the slice, and the resources they are using for running the test together. A given experiment can only use the resources in the associated slice.

In general, a resource can be broadly defined as:

- A physical resource (Server, IoT)
- A virtual resource (VM, SW, Library)

In F-Interop a resource can also be distinct based on its locations:

- Local: that means in the same location of the user
- Remote: connected in a remote location
    - in one of the federated testbed connected
    - in the location of another remote user

In the selection process this should be transparent for the user, and there should not be differences in term of reachability, configurability and access. The GUI and the other F-Interop software infrastructure have to provide the right level of abstraction for researching a desired resource, based on its capabilities.

5.3.1 Local resources

When a user wants to use personal resources (local), it is necessary to register first them in the F-Interop Platform. The registration makes them available in a Session and selectable by the user for a testing session. The resource registration might require the following information:

- Unique ID: probably generated during the registration
- Serial Number
- Resource Owner
- Resource Manufacturer
- Resource Name
- MAC addresses
- IP Addresses
- Protocols implemented (with version)
- Operating System and SW version
- Reference Links
- Visibility (Private, Protected, Public)
- Location (GPS coordinate, Country, …)
- Availability (used to reserve a shared resource)

All the data listed above (and others that will be properly added for the IUT registration) SHOULD be securely stored in the Resource Repository located in F-Interop Central Service. Among the different field, the resource Owner field is the only one representing personal data, given that it allows to identify the user owning the IUT. Therefore, such information should be properly protected by F-Interop, and anonymized.

Once registered (and thus, part of the F-Interop resources), the device MUST be reachable by the other F-Interop devices (and users) that share the same session and slice. The

communication channel between the new resource and the other IoT devices, used both for control and data flow, MUST be secured.

5.3.2 Remote Resources

When a resource is set as discoverable, any other F-Interop user involved in the same session can see it and use it. To make the resource discoverable, the owner has to explicitly set the visibility flag as not private. Private SHOULD be the default value to prevent unexpected resource sharing.
In F-Interop a registered user that is part of a session MUST be able to select among a list of registered remote (and not private) resources. The searching process should hide some confidential information like the precise location, and the resource's owner. But it has to simply inform about the kind of capabilities and protocols it implements and its availability, without disclosing any information related to the end user, owing the shared resource (IUTs).

5.4 Reservation and Resource Instantiation

In this phase of the session, there is no interaction between the user and the platform. All the necessary resources required for the testing session are registered and selected in the previous step. The orchestration component has to perform the following tasks:

- Check the availability of the resources
- If resource are available, the orchestration module can instantiate them. Depending on the type of the resources, an Operating System and the relevant tools SHOULD be installed in order to set up the test and run it in the following execution phase.
- If a resource is used by in another slice, the orchestrator will require it as soon as it becomes available.

From the security point of view, F-Interop has to ensure in this step that all resources required for an experiment in a slice should be accessed only by users involved in it. Moreover, even though most of resource's details should be hidden by the F-Interop GUI, likes location and resource ownership, further measurements should be taken to avoid any data leakage, and unauthorized access to resources and information.

5.5 Test Execution

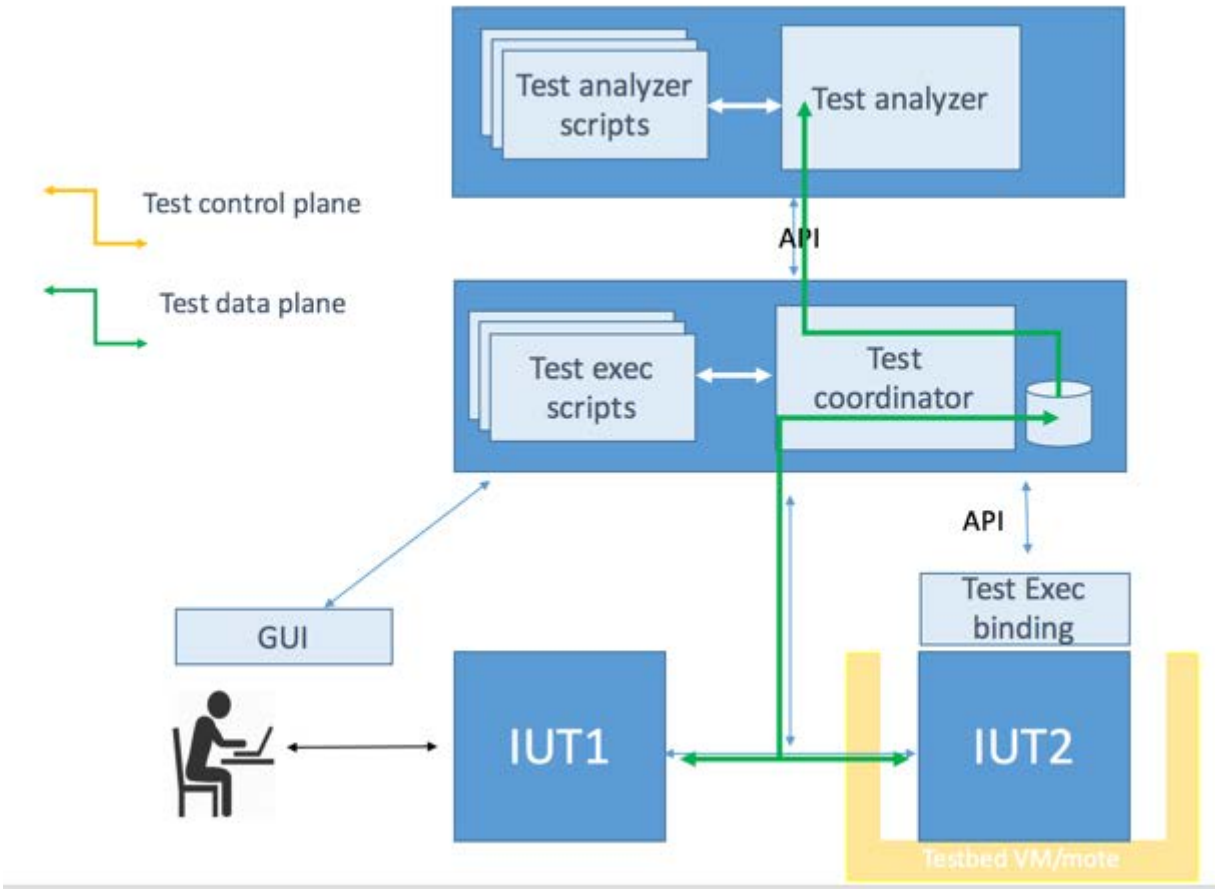

Figure 10 - Test Execution

Test execution is the real interactive part of the session where several elements of the platform are involved. In detail, as shown in Figure 10, the main actors (which can play an active or passive role) are the following:

- User(s)
- IUT(s)
- The Testing Coordinator (TC)
- A bunch of agents (mainly designed to root packets)
- Repositories (Profile, Experiment, Resources, Results)

User is the primary actor in the execution of the test. An F-Interop web interface, (GUI) will be provided as support to control the full test lifecycle, and check the correct execution of all the intermediate steps. All the instructions are translated by the GUI into control messages exchanged between the GUI and the Testing coordinator. These packets are sent through a secure channel (the control pipe), by agents. As mentioned in Sec. 4.2, all packets are encapsulated in AMQP messages, encrypted using TLS protocol. These packets are pure signaling messages, necessary to orchestrate the whole test lifecycle (start, next, pause, resume, stop, quit). From the security point of view, there is no critical information exchanged on these communication channels. Considering the nature of the messages the only information that can be inferred by analyzing those packets is the beginning and the end of the testing session.

IUT is the main target of the test. The primary goal of conformance/interoperability tests is to check if the IUT behaves as expected, according to the implemented protocol, and also if it is able to interoperate with other devices (belonging to the F-Interop testbeds, or to another end user). Packet flows are generated among all IUTs involved. To secure any data exchanges, the nearest **agent,** responsible to route F-Interop messages among components, encapsulates and encrypt those packets in AMQP messages. Before reaching the final destination an AMQP message is stored in a specific queue (called exchange), used by **the Test Coordinator (TC)** to perform mid-term evaluation and finally store results in the repository. Those data exchanges, which can imply flow crossing F-Interop boundaries (see Figure 5) can be an interesting target for malicious observers, trying to extrapolate sensitive / private information.

The result of the testing session and all the other correlated information MUST be protected and made accessible only to the users directly involved in the session.

The Result **Repository**, located in the F-Interop Central Service, is the place where all that information is stored. To reduce risks, repositories SHOULD not be directly accessed by the end user, but only by a limited list of components, like the Testing Coordinator. The latter is responsible of generating a final report, by reading raw packets stored in the repository during test execution.

5.6 Result Analysis

At execution time, most of the packets exchanged among IUTs are collected and stored in F-Interop Result Repository. This collection of raw and partial data is necessary in order to perform a complete benchmark of the experiment and produce a final report about the execution. The Testing Analyzer tool is the platform component responsible of generating those verdicts (pass/fail), one for each step of the testing case performed. Once generated they are sent back to the user, visualized in a simple and readable way using the visualization tool, that is part of the GUI.

In particular, the information collected in the repository are:

- Packets dump
- (Partial) verdicts
- System logs

Packets dump represents the full history of the execution of a test session. In general, they are composed by the collection of packets exchanged during the test, and they are related to a specific protocol (CoAP, 6TiSCH, TCP, UDP). Usually such information is extracted using a sniffer.

The **Packets dump** will be saved in clear. To provide confidentiality of the information they carry, they may be encrypted before being saved in the Repository. While increasing the level of data protection, this will increase complexity for the Test Tool Analyzer, in the computation of the test verdict.

For performing test statistics and for allowing a F-Interop user to repeat a test and to compare the results obtained with different versions of his/her implementation, packets dumps are kept in the database also after the test verdict has been sent to the user(s). When users start an experiment, they should select through GUI whether they want to save the results on the F-

Interop Result Repository or not, and if stores, for how long. According to the limited retention of data principle, a clear policy MUST be defined in F-Interop to erase old data or anonymized it.

Verdict is the output of the Testing Analyzer. It is generated at the end of the test execution by elaborating all traces stored in the repository. When ready, the verdict is sent to the user in a JSON format, and then translated as a visual report in the GUI. The JSON value should be stored in the repository.

System Logs will be collected during the whole test execution. Each module can store useful information in a F-Interop log collector, in order to identity problems and isolate the responsible of a failure. F-Interop Service Manager MUST document the purposes for which logs are produces. For each purpose:

- Service managers MUST check whether Personal Information might be logged. If so, a justification must be given and steps to ensure compliance with the EC Directive.
- Service manager MUST ensure that retention periods are appropriate.

Examples of Persona Data in the a log traces are:

- Usernames: a clearly Personal Data as they directly identify individual users.
- IP Addresses: must almost always be processed as Personal Data
- MAC Addresses: as a unique identifier may be linked to individuals.

5.6.1 Limited retention of data principle

According to Article 6 of the Data Protection Directive, **Personal Data** has to be *"kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the data were collected or for which they are further processed"*.

As defined in Chapter 2, **Personal Data** is "any information relating to an identified or identifiable natural person, referred to as "**data subject**".[5]

GDPR to some extent clarifies and possibly expands the sometimes vague concept of personally identifiable information, defining a "data subject" as a natural person who can be identified "by means reasonably likely to be used by the controller or by any other natural or legal person," including by reference to "an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person" (Art. 4(1)). According to GDPR, which adds **location data** and **online identifiers** among the personal data, IP and MAC addresses of IUTs connected to the F-Interop platform, will represent personal data which can reveal the user identity.

In general, Data stored in the repository must therefore be erased when those purposes have been served. The time limitation for storing personal data applies, however, only to data kept in a form which permits identification of data subjects. Lawful storage of data which are no longer needed, could, therefore, be achieved by anonymization of the data or pseudonymisation (see Annex 8.1).

[5] **the data subject** is the person whose personal data are collected, held or processed. – from European Data Protection Glossary

6 Conclusion

In the present “Privacy By Design” Report we have described how Privacy and Security have been integrated into the design specifications, and Architecture of the F-Interop Platform. Taking a privacy by design approach (i.e., following a proactive approach) is essential to minimize privacy risks and build trust. This is particularly important for F-Interop, to make sure F-Interop Users / Contributors can trust the F-Interop Testbed as a Service (TaaS), and run their experiments online, without exposing any personal or sensitive data to third parties. The TaaS will be accessible only to authenticated and authorized users, which should use and not abuse of the platform. The Architecture will therefore implement appropriate management schemes, for blocking misbehaved users.

While running the tests, users will not exchange any sensitive data, but rather some personal data (e.g., location, usernames, IP addresses, etc.) that could allow potential attackers to identify them. Such data should be properly protected, according to the GDPR recommendations. Moreover, test results must be accessible only to the user(s) performing the tests, and be saved for the minimum needed period of time (for further elaboration: comparison with previous test implementation, statistics, etc.)

Being, at this stage of the project, the F-Interop Architecture still under development, the actual security and privacy measurements which will be finally implemented may slightly diverge from what currently foreseen in the present document. But general consideration and recommendation will still apply.

7 References

[1] "CHARTER OF FUNDAMENTAL RIGHTS OF THE EUROPEAN UNION," [Online]. Available: http://www.europarl.europa.eu/charter/pdf/text_en.pdf.

[2] "European Convention on Human Rights," [Online]. Available: http://www.echr.coe.int/Documents/Convention_ENG.pdf.

[3] "General Data Protection Regulation," [Online]. Available: http://ec.europa.eu/justice/data-protection/reform/files/regulation_oj_en.pdf.

[4] "Directive on privacy and electronic communications," 12 7 2002. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:en:HTML.

[5] "Directive 95/46/EC," 24 10 1995. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML.

[6] "The 7 Foundational Principles," [Online]. Available: https://www.ipc.on.ca/images/Resources/7foundationalprinciples.pdf.

[7] "Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data," [Online]. Available: http://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/108.

[8] "LEGISLATION," [Online]. Available: https://secure.edps.europa.eu/EDPSWEB/edps/site/mySite/Legislation.

[9] "Enterprise Architecture (EA)," [Online]. Available: http://www.gartner.com/it-glossary/enterprise-architecture-ea.

[10] "RabbitMQ," [Online]. Available: https://www.rabbitmq.com/.

[11] "Advanced Message Queuing Protocol Protocol Specification Version," [Online]. Available: https://www.rabbitmq.com/resources/specs/amqp0-8.pdf.

[12] "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," [Online]. Available: https://tools.ietf.org/html/rfc5280.

[13] "Key GENI Concepts," [Online]. Available: http://groups.geni.net/geni/wiki/GENIConcepts.

[14] [Online]. Available: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

[15] 18 12 2000. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:008:0001:0022:en:PDF.

[16] "ARTICLE 29 - DATA PROTECTION WORKING PARTY," [Online]. Available: http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

8 Annex

8.1 Anonymised and pseudonymised data

Personal Information (PI) contains identifiers, such as a name, date of birth, sex and address. For any organizations collecting PI and not protecting it, there are serious legal and financial issues. Moreover, as a consequence of the limited retention of data principle, Data Subject would have to be **anonymised** if a Controller wanted to store them after they were outdated and no longer served their initial purpose. Data is **anonymised** if all identifying elements have been eliminated from a set of Personal Data. No element may be left in the information which could, by exercising reasonable effort, serve to re-identify the person(s) concerned. Where data have been successfully anonymised, they are no longer personal data.

The processing step of anonymizing personal data is the last legal second that this data falls under the scope of EU data protection laws as personal data. Data Protection Working Party 29 (WP29) [14] considers several anonymization techniques:

- **Noise addition**. This means that an imprecision is added to the original data.
- **Substitution**. Information values of the original data are replaced with other parameters. Substitution is often combined with noise addition.
- **Aggregation**. In order not to be singled out, an individual is grouped with several other individuals that share some or all personal data, i.e. their place of residence and age.
- **Differential privacy**. This comes into play when a company gives a third party access to an anonymized data set. A copy of the original data remains with the company, and the third-party recipient only receives an anonymous data set. Additional techniques such as noise addition are applied prior to the data set transfer. Differential privacy is applied when an authorized third party is requesting data.

The GDPR introduces also a new concept in European Data Protection law that is called "**pseudonymization**". It is the separation of data from direct identifiers so that linkage to an identity is not possible without additional information that is held separately. Pseudonymization, therefore, may significantly reduce the risks associated with data processing, while also maintaining the data utility. To pseudonymize a data set, the "additional information" must be "kept separately and subject to technical and organizational measures to ensure non-attribution to an identified or identifiable person." This is a privacy-enhancing technique where directly identifying data is held separately and securely from processed data to ensure non-attribution.

The WP29 provides the following selected examples of pseudonymization techniques:

- **Hash functions**. Hashes are a popular tool because they can be computed quickly. They are used to map data of any size to codes of a fixed size.

- **Tokenization**. Tokenization is a process by which certain data components are substituted with a non-sensitive equivalent. That equivalent is called the token. The token has no exploitable value, but it serves as an identifier. It is a reference that traces back to the original data.

8.2 Personal Data breach notification standards

Unlike Directive 95/46/EC [5], which was silent on the issue of data breach, the GDPR [3] contains a definition of "Personal Data breach" and notification requirements to both the supervisory authority and affected data subjects. Under the GDPR, a "personal data breach" is

"a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorized disclosure of, or access to, personal data transmitted, stored or otherwise processed.".

This broad definition differs from that of most U.S. state data breach laws, for example, which typically are triggered only upon exposure of information that can lead to fraud or identity theft, such as financial account information.

In the event of a personal data breach, data controllers must notify the **Supervisory Authority**[6] "competent under Article 55" which is most likely (looking to Article 56(1)) the **Supervisory Authority** of the member state where the controller has its main establishment or only establishment, although this is not entirely clear. Notice must be provided "without undue delay and, where feasible, not later than 72 hours after having become aware of it." If notification is not made within 72 hours, the controller must provide a "reasoned justification" for the delay.

Article 33 (1) contains a key exception to the **Supervisory Authority** notification requirement:

Notice is not required if "the personal data breach is unlikely to result in a risk for the rights and freedoms of natural persons," a phrase that will no doubt offer data protection officers and their outside counsel opportunities to debate the necessity of notification.

A notification to the authority must "at least":

1. describe the nature of the personal data breach, including the number and categories of data subjects and personal data records affected;
2. provide the data protection officer's contact information;
3. "describe the likely consequences of the personal data breach";
4. describe how the controller proposes to address the breach, including any mitigation efforts. If not all information is available at once, it may be provided in phases.

When a data processor experiences a personal data breach, it must notify the controller but otherwise has no other notification or reporting obligation under the GDPR.

If the **Controller** has determined that the personal data breach "is likely to result in a high risk to the rights and freedoms of individuals," it must also communicate information regarding the personal data breach to the affected data subjects. Under Article 34, this must be done "without undue delay."

The GDPR provides exceptions to this additional requirement to notify data subjects in the following circumstances:

[6] Article 28 of Directive 95/46/EC [6]

- the controller has “implemented appropriate technical and organizational protection measures” that “render the data unintelligible to any person who is not authorized to access it, such as encryption”;

- the controller takes actions subsequent to the personal data breach to “ensure that the high risk for the rights and freedoms of data subjects” is unlikely to materialize; or

- when notification to each data subject would “involve disproportionate effort,” in which case alternative communication measures may be used.

Assuming the **Controller** has notified the appropriate supervisory authority of a personal data breach, its discretion to notify data subjects is limited by the DPA’s ability, under Article 34(4), to require notification or conversely to determine it is unnecessary under the circumstances.