Grant number:	883286	Horizon 2020: Secure societies
Project duration:	Sep 2020 – Aug 2022	SU-INFRA02-2019
Project Coordinator:	Joe Gorman, SINTEF	Security for smart and safe cities, including for public spaces
		Project Type: Innovation Action

IMPETUS

http://www.impetus-project.eu

IMPETUS Project Deliverable: D5.2

Initial mechanisms to preserve privacy in the secure smart city

Dissemination Status: Public

Editor: Nesrine Kaaniche (IMT)

Authors: Ravishankar Borgaonkar, Per Håkon Meland (SINTEF), Claudio Agostino Ardagna, Alessandro Balestrucci, Chiara Braghin (CINI), Joaquin Luzon (INS), Rafal Rynkiewicz (THA), Joachim Levy (CINEDIT)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Grant number:	883286
Project duration:	Sep 2020 – Aug 2022
Project Coordinator:	Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

About IMPETUS

IMPETUS (Intelligent Management of Processes, Ethics and Technology for Urban Safety) is a Horizon 2020 Research and Innovation project that provides city authorities with new means to improve the security of public spaces in smart cities, and so help protect citizens. It delivers an advanced, technology-based solution that helps operational personnel, based on data gathered from multiple sources, to work closely with each other and with state-of-the art tools to detect threats and make well-informed decisions about how to deal with them.

IMPETUS provides a solution that brings together:

- *Technology*: leverage the power of Internet of Things, Artificial Intelligence and Big Data to provide powerful tools that help operational personnel manage physical and cyber security in smart cities.
- *Ethics*: Balance potentially conflicting needs to collect, transform and share large amounts of data with the imperative of ensuring protection of data privacy and respect for other ethical concerns - all in the context of ensuring benefits to society.
- *Processes*: Define the steps that operational personnel must take, and the assessments they need to make, for effective decision making and coordination - fully aligned with their individual context and the powerful support offered by the technology.

Technological results are complemented by a set of practitioner's guides providing guidelines, documentation and training materials in the areas of operations, ethical/legal issues and cybersecurity.

IMPETUS places great emphasis on taking full and proper account of ethical and legal issues. This is reflected in the way project work is carried out, the nature of the project's results and the restrictions imposed on their use, and the inclusion of external advisors on these issues in project management.

The cities of Oslo (Norway) and Padova (Italy) have been selected as the site of practical trials of the IMPETUS solution during the project lifetime, but the longer-term goal is to achieve adoption much more widely.

The work is carried out by a consortium of 17 partners from 11 different EU Member States and Associated Countries. It brings together 5 research institutions, 7 specialist industrial and SME companies, 3 NGOs and 2 local government authorities (the trial sites). The consortium is complemented by the Community of Safe and Secure Cities (COSSEC) – a group established by the project to provide feedback on the IMPETUS solution as it is being developed and tested.

The project started in September 2020 with a planned duration of 30 months.

For more information

Project Coordinator:	Joe Gorman, SINTEF:	joe.gorman@sintef.no
Dissemination Manager:	Snježana Knezić, TIEMS:	snjezana.knezic@gmail.com

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Executive Summary

This document describes the different privacy challenges within a smart city, presents main privacy enhancing technologies and points-out their suitability to the different tools developed in IMPETUS. In particular, the document describes:

- *Privacy challenges with respect to IMPETUS scenarios.*

Along with the main use cases considered within the project, i.e., festival and protest events, several focus groups have identified a number of privacy issues that are summarized and categorized with respect to collection, storage and processing activities. Indeed, with respect to legislation, issues related to the massive collection of personal information, challenges of setting-up an informed consent agreement, the combination of different data sources that may lead to specific profiling and also challenges with the storage and outsourcing are discussed.

- *Privacy Enhancing Technologies.*

PET include all techniques and mechanisms that enforce the *privacy by default* principle. A panorama of these techniques is presented with respect to their best applicability to different data layers, i.e., collection, storage and processing.

- *Privacy mechanisms in IMPETUS*

The privacy overview describes privacy mechanisms adopted by different IMPETUS tools. A specific focus is given to Social Media Detection (SMD), Physical Threat Intelligence (PTI), Human Computer Interaction (HCI) and Weapon Detection (WD) tools as they are considered as potentially dealing with sensitive information, as discussed in Deliverable D11.3. The chronological description of the deployed privacy mechanisms at different data layers is discussed, namely how data are collected, what mechanisms are used for anonymization/pseudonymization, how stored data is protected and what data are used for processing.

Table of Contents

Table of Figures

List of Tables

List of Abbreviations

Table 1: List of Abbreviations

Abbreviation	Explanation
ABE	Attribute Based Encryption
AC	Anonymous Credentials
AI	Artificial Intelligence
API	Application Processing Interface
CIA	Confidentiality, Integrity, Availability
CSP	Cloud Service Provider
GDPR	General Data Privacy Regulation
HCI	Human Computer Interaction
ICT	Information and Communication Technology
IOI	Item of Interest
IOT	Internet of Things
LBS	Location Based Services
ML	Machine Learning
NIST	National Institute of Standards and Technologies
PET	Privacy Enhancing Technologies
PTI	Physical Threat Intelligence
SDC	Statistical Data Disclosure
SLA	Service Level Agreement
SMD	Social Media Detection
SMC	Secure Multiparty Computation
WD	Weapon Detection

List of Definitions

Table 2: List of Definitions

Term	Definition/explanation
Anonymization	*"An anonymisation process of personal data is an irreversible process that aims at making impossible to identify individuals within data sets"*CNIL
Privacy by Design	The privacy by design is an approach that ensures considering the privacy preservation and data protection issues at the design phase of any system, service or product and then throughout the lifecycle.
Pseudonymization	*"A pseudonymization process is a compromise between retaining raw data and producing anonymized datasets. It consists of replacing directly identifying data (surname, first name, etc.) in a dataset with indirectly identifying data (alias, random number in a filing system, etc.). They may result from a cryptographic hash of the data of individuals, such as their IP address, user ID, e-mail address."* CNIL

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1 About this deliverable

1.1 Why would I want to read this deliverable?

This document contains a description of the initial privacy-preserving mechanisms implemented to comply with GDPR and local legislation. The main objective is to ensure that IMPETUS solutions provide full compliance with data protection legislation in exploiting large amounts of data from the smart city.

1.2 Intended readership/users

The primary intended readership consists of people with an interest in preserving privacy in public spaces, and who are:

- interested in using results of the IMPETUS project;
- have specific interests in the data privacy within IMPETUS applications and scenarios.

The primary readership involves two categories:

- people outside of the IMPETUS consortium, including adopters of the project results after the completion of the project.
- members of the IMPETUS consortium involved in the design, testing and evaluation of IMPETUS solutions.

The deliverable (or parts of it) may also be of interest to a wider group: anyone with an interest in data privacy and legal issues in the context of using advanced technological solutions, especially in the smart-city environment.

Specifically, users *outside* the IMPETUS consortium wishing to use some or all of the IMPETUS tools and platform, would read this deliverable to:

- learn about the general data privacy and legal issues, including GDPR that they ought to be aware of if they want to deploy the tools;
- learn about any specific issues that may apply in relation to individual tools;
- access practical guidance that will help when ensuring compliance with relevant principles, regulations and laws with regards to privacy.

People *inside* the consortium using the tools and/or responsible for developing some solutions as part of the project activities would read this deliverable to:

- gain the same benefits as people outside the consortium.
- get help for providing feedback to tool and platform developers in the project about privacy aspects of their solutions - that might need to be designed differently- to be able to comply with relevant principles, regulations and laws.
- help to design their solution in a way that helps users address/overcome potential data privacy and legal barriers, and so increase the market potential of the product.

1.3 Structure

This report is organized, in three main chapters, as follows:

- Chapter 3 gives an overview of data privacy threats and requirements in smart cities. It first starts by introducing the functional architecture and pillars infrastructure (Section 3.1), and details the threats models and the security and privacy challenges in the context of IMPETUS (Section 3.2). Then, it recalls the jurisdictional privacy context, I.E., GDPR (Section 3.3) and both the NIST security and privacy frameworks (Section 3.4), before enumerating four main privacy requirements (Section 3.5) with respect to the identified requirements in D1.2.
- Chapter 4 reviews privacy preserving mechanisms and their deployment in existing smart cities. It relies on a taxonomy that involves 3 groups, based on the entity that will endorse the responsibility of maintaining a privacy-preserving setting. It details the privacy technologies that are mainly deployed at the service provider's servers and supported by the communication channels. This Chapter will first

present privacy preserving authentication (Section 4.1), privacy preserving computation (Section 4.2), Statistical Data Disclosure (SDC) (Section 4.3) and privacy preserving communication (Section 4.4). Then, it presents an overview of deployed mechanisms in 14 different cities (Section 4.5), and discusses their effectiveness and robustness (Section 4.6).

- Chapter 5 reviews the privacy challenges allied with four of IMPETUS tools, reported as dealing with sensitive and/or personal data, and discusses candidate privacy technologies in the context of IMPETUS scenarios. It recalls the selected IMPETUS tools and the deployed privacy mechanisms (Section 5.1). Relying on the analysis of the aforementionned tools, this chapter then details main building blocks and identifies suitable technologies to provide a full-stack privacy-preserving IMPETUS setting, ie., from the collection to the processing of data (Section 5.2).

1.4 Other deliverables that may be of interest

There is a relationship with “D11.3 – POPD Requirement 4” which describes the project’s approach to dealing with ethical issues and data privacy issues. It details the personal data collection and handling strategy to be followed by the project, and discusses the ethical considerations for different data collection methods.

There is a relationship with the different tools and results of the project. Indeed, the principles and privacy requirements point out *“constraints”* on the proposed tools, and privacy mechanisms provide an overview of the approaches to be considered to be compliant with EU, international and local legislation. There will be a relationship with:

- D1.2. “Requirements for public safety solutions”, providing requirements for privacy enhancement by different IMPETUS tools/
- D3.1. “Detection tool development report initial report”, providing a description of the functionalities of the algorithms of IMPETUS tools, that are considered as a starting point of a full-stack privacy preserving environment.
- D4.1. “Data analytics and ingestion-time access control initial report”, providing a description of the functionalities of Physical Threat Intelligence tool, and its interaction with the plateform.
- D5.1. “Initial ethical framework”, providing a comprehensive framework of ethical considerations by the different tools, mainly in terms of massive data collection.
- D6.1. “Initial concepts of operations”, developing strategies and operating guidelines, to support organizations in integrating the IMPETUS platform.

2 Introduction

Smart cities are emerging to dynamically optimize and efficiently use resources in traditional cities, provide better facilities and improve the quality of life for people. They involve a variety of components and include ubiquitous sensing devices, heterogeneous networks and large-scale databases. In this context, data is continuously collected, stored, transferred and processed among several actors, which significantly increases the surface of potential attacks. Indeed, data security and privacy preservation mechanisms are considered as main building blocks for the different tools proposed in IMPETUS.

While the data security is always considered as critical, the privacy is a recent *interdisciplinary* requirement that is rooted in a data governance strategy and enforced by national, European and international legislation, like the EU General Data Protection Regulation (GDPR) or the US California Consumer Privacy Act (CCPA). Thus, the term *"Privacy by Design"* has emerged as a development method for privacy-friendly systems and services, so going beyond conventional technical solutions, complying with legislation and addressing organizational procedures and business models as well.

This report will focus on the different mechanisms to enforce the privacy by design principle. One important element in this document are the main technical mechanisms, generally called Privacy Enhancing Technologies (PETs), such as malleable signatures, anonymous certification, attribute-based credentials, multiparty computation, private processing and data perturbation. The effectiveness of such mechanisms has been studied and demonstrated by researchers and with various pilot implementations. However, they are still not well perceived by many operators mainly because PETs are reputed to have an impact on the utility.

This report contributes to bridging the gap between the legal framework and the available technological implementation measures by providing an inventory of existing approaches, privacy design strategies, and technical building blocks of various degrees of maturity from research and development. Starting from the privacy principles of the legislation, important elements are presented as a first step towards a design process for privacy-friendly systems and services.

The report sketches a method to map legal obligations to design strategies, which allows the system designer to select appropriate techniques for implementing the identified privacy requirements, based on representative IMPETUS tools. Furthermore, the report reflects limitations of the approach, mainly with respect to the practical deployment of PETs for specific real-world scenarios. It distinguishes inherent constraints from those which are induced by the current state of the art.

3 Data privacy threats and requirements

Connected smart cities are built upon a variety of services adapted to specific needs and citizens' expectations. Mainly designed with respect to personalization techniques, these services and applications rely on massive collection and analysis of gathered data.

Indeed, a power imbalance between data processing entities, which determine what and how data is processed, and the individuals whose data is at stake, i.e., who might be influenced by decisions based on automated data analysis, or by failures to adequately protect private information, might be observed. To enforce privacy, the European Commission (EU) adopted, in 2018, the General Data Protection Regulation (GDPR) that sets up a legal context for the personal data collection, storage and processing. And, Privacy-Enhancing Technologies (PETs) [1, 2] have become a field that studies enabling techniques, investigates the level of data leakage, mitigates identification and traceability attacks and implements privacy-preserving processing.

By implementing PETs, the risks for the users' privacy would decrease and the legal data protection obligations of the entities responsible for the data processing would be fulfilled more easily. From this perspective, it is of utmost importance to first understand the main requirements and different security models to evaluate the relevant building blocks from the design phase. This chapter gives an overview of the main actors and settings in smart cities. It describes the core components, pillars, and applications. It also details the legal and ethical framework regarding data collection and dissemination, and discusses the security and privacy issues and requirements of the smart cities.

3.1 Background

Connected smart cities implement services to improve the quality of life of citizens. From this perspective, IMPETUS aims to provide cities with new intelligent means to address security and safety issues in public spaces. Using data gathered from multiple sources, the project aims to facilitate the detection of threats and help human operators dealing with threats to make better informed decisions, while enhancing the privacy of citizens. Hereafter, we present the generic functional architecture and main interactions between data layers.

3.1.1 Functional architecture

For our analysis, we consider a functional architecture which consists of four layers with specific responsibilities and components, defined as follows:

- *Sensing and data collection layer:* consists of various equipment that collect data from the surrounding physical environment and share it with the data collection layer. For instance, different types of sensors, actuators or CCTV cameras are considered as main components of the sensing layer.

- *Data transmission layer*: deals with the transmission of gathered data using reliable wired or wireless communication to the local or remote databases. This layer mainly involves the communication protocols and services, from a networking point of view.

  Let us emphasize that storing a big volume of data incurs high storage overhead on the existing databases of this layer. Thus, many applications would rely on remote edge or cloud servers to address this issue and remove the burden of maintaining large infrastructures, especially when data involves real-stream data. Referring to the GDPR nomenclature, these databases may be maintained by either data processors or data controllers.

- *Data storage and processing layer:* performs all "*pre*"-processing techniques, i.e., storing and analyzing large amounts of data and maintaining infrastructures, with regards to different applications and services' requirements.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

- *Application layer:* is responsible for exchanging data between operators (e.g., citizens and stakeholders) and smart applications. At this layer, the data can be either stored raw, aggregated or processed via accurate analysis and visualization algorithms.

3.1.2 Pillars infrastructure

In the context of IMPETUS' use cases, i.e., dealing with public security and safety, we identify four pillars' infrastructures, namely institutional, social, physical and economic, introduced as follows:

- *Institutional infrastructure:* deploys the fundamental activities, e.g., management, governance and planning of events. It relies on the citizens' collected and produced data in decision-making processes.

- *Physical infrastructure:* involves IT, communication and hardware components that yield to support a physical environment for urban safety and mobility. For instance, the urban mobility focuses on quality of cycling, pollution indicators, and smart transportation systems in cities, while public safety considers safe walking and secure gathering places and holding social events.

- *Social infrastructure:* involves diverse mechanisms to promote and develop human and social capital, and provide intelligent and straightforward connected infrastructure for addressing different social needs and services of citizens, such as environment, and inclusive planning.

- *Economic infrastructure:* refers to the basic services that help to promote the process of production and distribution of economic activities and develop proper infrastructure to generate employment opportunities and attract investments. Although, this type of activities may not directly be represented by IMPETUS outputs, the different tools may have an impact to the economic development and the attractiveness of a city.

3.2 Security and privacy challenges in the context of IMPETUS

The IMPETUS project has conducted a study with different focus groups whereas challenges related to privacy have been identified. These results have been compared with other surveys from the literature as documented in Deliverable 1.2. Below, we organized a summary of the different raised challenges.

- Maintaining privacy during regular Internet browsing in connected cities is considered as "*difficult*". The first main concern is the widespread deployment of artificially intelligent processing algorithms that can be used in combination with the collected personal information to deduce involuntary correlations, leading to specific identification (other people, web pages, organizations, etc.). This refers to data inference and reconstruction attacks which are detailed in Section 3.2.2.

- The second privacy concern is related the tracking of spatial mobility, e.g., in relation to pedestrians, consumers and vehicles. Tracking is already a legitimate part of smart city technologies, as per ensuring safety in the public space, but there is a fear of misuse, e.g., related to unwanted surveillance.

- The third reported challenge consists of properly informing citizens about what the information is being used for, obtaining and maintaining informed consent in a practical manner. This challenge becomes even harder when combining different data sources. The focus groups expressed that aggregation of data may lead to profiling, discrimination, and political manipulation.

- There is a general concern from many cities on unnecessary or unwanted use of personal data for purposes such as marketing campaign, telephone directory, contact-lists of some companies, etc. When it comes to personal data from video surveillance footage, the studies show a varying degree of concern from the citizens in different cities. Hence, we have to acknowledge that there are cultural differences between European countries on what is considered to be a privacy issue.

In order to understand the aforementioned challenges, we hereafter present threat models and main challenges, namely related to extracting information about data owners, and also referred to as citizens and users, or organizations, service providers and governmental institutions. Taking the example of the smart mobility, it is imperative that not only the privacy of the collected and analyzed data be preserved but also the running algorithms (usually considered as sensitive and proprietary). Regardless of the goal, the attacks and defenses relate to exposing or preventing the exposure of analysis algorithms (processes) and collected data.

3.2.1 Threat models

Privacy and security risks are mainly related to environments, technologies and involved parties. Indeed, as pointed out in a recent report of the European Network and Information Security Agency (ENISA) [1, 2], understanding privacy concerns from a technical point of a view, leads to identify:

a. Collected and processed data that are released and may be considered as sensitive, personal and identifying data,
b. Data that may be used to identify and/or revoke the anonymity of a user,
c. Potential adversaries (i.e., actors that may gain access to personal identifying information) which can rely on:
    - data being transferred and processed that the adversary has access to,
    - external and background knowledge of the adversary- possible collusion with other entities.

Recall that adversaries may be passive or active, and considered under either semi-trusted or untrusted environments [1], presented as follows:

- Passive attacks: the adversary passively observes the data and performs inference or concludes connections, e.g., without changing anything in the process. These attacks are usually considered against privacy requirements, namely anonymity, unlinkability and unobservability.
- Active attacks: the adversary actively changes the data or processes. These attacks are usually considered against security requirements, namely integrity and availability.

This categorization follows the common security literature distinguishing between honest-but-curious and fully malicious adversaries. It will be considered, afterwards, for our privacy analysis of the deployed tools and the whole IMPETUS platform.

3.2.2 Privacy and security issues

As discussed earlier, IMPETUS identified several challenges with respect to the enforcement of privacy technical requirements and the compliance with the legislation. These challenges are tightly related to technologies, the functional architecture and involved actors, defined in Section 3.1.
In the following, we point out the different vulnerabilities and issues that may arise from a technologies' point of view, while referring to the architectural layers considered in IMPETUS (c.f., Table 3).

Table 3 Summary of privacy challenges with respect to IMPETUS functional architecture layers

	IoT	***Cloud***	***AI***
Sensing and data collection layer	Data overcollection		-Data overcollection -Data poisoning -Backdoor injection
Data transmission layer	Lack of standardized secure short-band communication protocols		

Data storage and processing and layer	Limited computation resources for advanced secure (cryptographic) algorithms	-Loss of data and computation control -Lack of knowledge about effective SLA enforcement -Multi-tenancy	-Inference attacks -Model theft
Application layer	Open and insecure APIs		

3.2.2.1 *IoT-based challenges*

The Internet of Things (IoT) consists of interrelated, internet-connected (smart) objects that are able to collect and transfer data over a wireless network without human intervention. There are various privacy issues associated with smart devices that are mainly due to the massive collection of data, focused on the sensing and communications layers.

Indeed, the connected devices have the capability to be used as a mediator storage or a fog node to perform a small computation in the network. These sensing and pre-processing capabilities make them vulnerable end-points for collecting the exchanged data and enriching adversarial databases, thus conducting specific correlation and inference attacks. While a huge number of applications are continuously proposed to provide various benefits for citizens, the majority of these applications gain access to private information of users -*without acquiring explicit informed consent*- and may transfer the collected data to *unauthorized* third parties.

Finally, the sensing capabilities of the smart devices *facilitate* the bypass of the data minimization principle. Most applications usually collect more data than the necessities of original functions which are known as data overcollection.

3.2.2.2 *Cloud based issues*

In order to cope with the shortcomings of smart devices, i.e., processing and storage capacities, battery constraints, etc. various applications delegate the data and processing management to *external* cloud providers. While outsourcing data and processing has various economic advantages, several security and privacy challenges are identified in [3]. In the following, we summarize common challenges raised by cloud infrastructures, platforms and applications.

○ *Data and computation outsourcing:* by outsourcing the data to remote servers, data management is delegated to a third-party provider, usually considered as semi-trusted or honest-but-curious entity. This raises privacy concerns, such as the anonymity of data owners.

○ *Physical location of data:* the lack of knowledge about the physical location of data in cloud services may have an impact on the data security, quality of services and might harm users' privacy. This latter is utmost importance as data legislation regarding the collection and processing of data is different between different countries and regions, and can be more intrusive compared to the EU regulations.

○ *Lack of knowledge about Service Level Agreements (SLAs):* SLA is a contract signed between the client and the service provider including functional and non-functional requirement. It considers obligations, service pricing, and penalties in case of agreement violations. However, due to the abstract nature of clouds, SLA violation with regards to data involve data retention, privacy leakage.

○ *Multitenancy:* this cloud feature means that the cloud infrastructure is shared and used by multiple users. In a nutshell, data belonging to different users may be located on the same physical machine, based on a specific resource allocation policy. Due to the multi-tenancy's economic efficiency, providers usually select this feature as an essential block for the cloud environment design. However, it generates new threats, such that, malicious users may exploit this co-residence issue to perform privacy (inference) attacks.

3.2.2.3 *AI-based attacks*

Recent progress in Artificial Intelligence (AI) in general, and Machine Learning (ML) in particular, is continuously encouraging many sectors to integrate AI-based algorithms in different processes. AI is a key enabler of smart cities, where the size and complexity of smart cities' systems are key challenges. The ability of efficiently and process gathered data and monitor in real time the state of critical infrastructures increasingly become an added value and a practical need. Unfortunately, they are generally considered as data-hungry tools and their benefits are often accompanied by a mostly black-box character and high complexity of the final algorithms in use, rendering conventional methods for safety assurance insufficient or inapplicable. Hence, the need of enforcing the privacy by design [4,5].

As presented above, the massive collection of data from the different devices, a.k.a., referring to the sensing and data collection layers, constitute first threat vectors to attack intelligent systems due to their multitude and their limitations in terms of resources and security features. For instance, by poisoning smart city's data, adversaries can try to fake the models, impending them to learn the correct correlation between data and the state of a critical system (modifying the model boundaries), or they can push the model in taking decisions that are hampering city's infrastructure and population. In this context, there is the need to confirm the common assumption on the effectiveness of the employed ML models, adopting suitable privacy and security techniques. These techniques aim to counteract adversaries trying to deceive ML model at different layer of the ML pipeline: i) data collection and ingestion, ii) training, iii) inference. The state of the art in the domain show that while the performance of ML and neural network architectures had a boost in the last years, their robustness to adversarial settings is asking a step ahead [6,7,8,9]. It is clear that the strong link between sensor data and ML models, as well as the intrinsic weakness of the sensors themselves, introduce new risks and attacks that cannot be ignored [10,11,12]:

- *Backdoor injection:* the first one aims to manipulate data to attack the learning phase. In this case, the attacker crafts and distributes corrupted data, which are used by ML algorithms to build an inaccurate model of the system behavior. These attacks are referred Machine Learning Poisoning [13]. They produce a poisoned ML model that learns a wrong correlation between data and the state of the monitored infrastructure. Smart attacks based on careful data manipulation can open the door to stealth attacks on the infrastructure, providing adversaries with the ability to introduce *"backdoors"* in the model [14]. These backdoors induce erroneous classification of inputs, with possibly disastrous consequences on the working of the whole system. For instance, if an anomalous detection model is trained, the machine learning poisoning can introduce a backdoor that impedes the model to classify an anomaly, identifying it as a safe behavior. Let us consider a ML model that is monitoring the quality of air pollution, as considered within IMPETUS. An adversary can fake the model in believing that the presence of some chemicals in the air is innocuous, while they could be dangerous for the population.

- *Data Poisoning:* the second one aims to inject specific data, generally carefully selected, to fake an existing model into taking decisions that decrease performance and increase risks of city's infrastructure and population. This category of attacks, called adversarial examples [6], builds on sensor inputs that can trick a deployed model, trained on benign data, into making a wrong decision. The most difficult aspect here is that adversarial example attacks are difficult to counteract since poisoned data are generally indistinguishable from a normal input for humans. This represents a fundamental problem in the domain of neural network security [15]. For instance, let us consider a monitoring service in a smart city, via CCTV cameras. Adversarial examples can be used by an attacker interacting with different cameras to let the model believe that certain areas of the city are congested. This would force the model to reroute the traffic towards busy areas, as well as changing the traffic light timing, creating a gridlock. This would have disastrous consequences for instance in the case of a terroristic attack.

- *Model theft:* the third one is a mix of the previous two and is employed in scenarios, and mainly considered a security threat where ML models are either i) retrained over time or ii) alternative models have been trained and can be deployed on the basis of contextual information.

- *Inference attacks:* the fourth category of attacks involves two main attacks, namely (a) inference about members of the population and (b) inference about members in the training set. For the first case (a), an adversary can use the model's output to infer the values of sensitive attributes used as input to the model.

Note that it may not be possible to prevent this if the model is based on statistical facts about the population: for example, suppose that training the model has uncovered a high correlation between a person's externally observable phenotype features and their genetic predisposition to a certain disease; this correlation is now a publicly known fact that allows anyone to infer information about the person's genome after observing that person. For the second case (b), the focus is on privacy of the individuals whose data was used to train the model. For instance, given a model and an exact data point, the adversary infers whether this point was used to train the model or not. The adversary may also try to extract properties. In fact, training data may not be identically distributed across different users whose records are in the training set; unlike model inversion, the adversary tries to infer properties that are true of a subset of the training inputs but not of the class as a whole.

3.3 Jurisdictional privacy context

This section recalls data privacy legislation and data dissemination guidelines with a focus on GDPR. The jurisdictional privacy context provides a legal framework for technologists to efficiently analyze various aspects of security and privacy, such as involuntary visual and audio capture of personal property, massive collection of personal data, involuntary surveillance, etc.

In an effort to provide an analysis of various aspects of privacy, security, and surveillance concerning the involuntary visual and audio capture of personal property, access to personal data, involuntary surveillance, storage, and security of data collected, this section recalls data privacy legislation and data dissemination guidelines with a focus on GDPR.

In 2018, the General Data Protection Regulation (GDPR) came into force for effectively ensuring the protection of the data subject's personal data. In particular, the regulation clarifies the conditions under which it is compulsory to obtain the consent of the data subject before processing his/her personal data, especially for sensitive personal data and data relating to minors. The GDPR also introduces the new obligation of accountability for organizations (i.e., data processors and data controllers). Indeed, each entity processing personal data must be able to demonstrate at any times that it is complying with the obligations laid down by the GDPR. For more details regarding GDPR, please refer to Deliverable D1.2. "Requirements for public safety solutions".

According to the GDPR, the data subject's consent is given for specific purposes with which both the data controller and the data processor must comply. In this context, three main roles are defined. The data subject who gives his consent to a data controller (i.e., organization,) for the processing of his personal data, with the possibility to forward them to a data processor (i.e., organization) that may process data on behalf of the data controller. Here data controllers are responsible for:

(i) specifying to the data subject the purpose of data collection,
(ii) obtaining the data subject's consent, and
(iii) processing personal data according to the consented purposes, and not beyond.

For ease of presentation, we, in the following, refer to the data subject as the data owner and to both the data controller as well as the data processor as the service provider.

From a data owner perspective, there is a need for new security mechanisms that support data accountability and provenance auditing. In a nutshell, these solutions have to ensure that personal data were accessed by data controllers and/or forwarded to data processors. Indeed, it is important to conceive a secure and transparent solution that permits data owners to

(i) check that data controllers and processors are correctly using their personal data with respect to the consented purposes,
(ii) verify whether data were accessed, processed, or forwarded without their consent, and
(iii) withdraw their consent.

From a data controller or processor perspective, there is a need to design a trusted and transparent accountability solution that enables them to obtain a proof of the data owner's given consent prior to gathering, accessing, processing, or storing his personal data.

3.4 NIST cybersecurity and privacy frameworks

This section will review both the US National Institute of Standards and Technologies (NIST) cybersecurity framework and privacy framework and summarizes the main guidelines and best practices.

3.4.1 NIST cybersecurity framework (NIST-CSF)

The National Institute of Standards and Technology Cybersecurity Framework (NIST-CSF for short) is designed to help organizations manage their security risks. It is a set of guidelines and recommendations that combine industry standards and best practices to help organizations manage their cybersecurity risks. It consists of a framework of policies that describe how an organization can improve its ability to detect, respond to and prevent a cyber-attack. This framework provides a comprehensive system of methods for detecting and managing cyber risks.

NIST-CSF is designed under five main functions that aim at identifying risks and assets, protect services and infrastructures, detect attacks, and provide suitable responses and recovery strategies:

- **Identify:** the first step consists of identifying the possible vulnerabilities in the system and the risks associated with loopholes. When these points are clearly pointed out, it is important to prioritize the cybersecurity tasks according to business requirements.
- **Secure:** the second step consists of implement the right mechanisms and operatuons with respect to the identified vulnerabilities. This includes training the employees regarding cybersecurity risks, limiting access to critical systems and data, and having the right cybersecurity procedures and policies in place.
- **Detect:** this step stands with the development of monitoring solutions and processes to identify the occurrence of a cybersecurity event. To do so, all information systems need to be monitored and processes have to be tested regularly to detect unusual activity.
- **Respond:** this step consists on defining a strategy, once an event is detected. This includes coordinating and communicating with stakeholders and law enforcement agencies, controlling the events' flows, and reupdating the processes based on collected data.
- **Recover:** this step aims at developing a strategy that ensures the resilence of the whole system. The goal is to efficiently and quickly recover after an incident.

3.4.2 NIST privacy framework (NIST-PF)

The NIST-PF is a voluntary tool developed in collaboration with stakeholders intended to help organizations identify and manage privacy risk to build innovative products and services while protecting users' privacy. Referring to NIST-PF, the privacy properties that have to be enforced by the designed services, are summarized as follows:

- *Anonymity:* it means the ability of the user to access a resource or service without disclosing his identity to third parties. That is, the anonymity of a user means that he is not identifiable within a set of subjects, known as the anonymity set. Several levels of anonymity have been defined in the literature, ranging from complete anonymity (i.e., no one can reveal the identity of the user) to pseudo- anonymity (i.e., the identity is generally not known but can be disclosed if necessary) to pseudonymity (i.e., multiple virtual identities can be created and used in different settings). ·
- *Data minimization:* it is a fundamental feature of privacy preservation. It requires that service providers collect and process the minimum amount of information, needed for appropriate execution of a service or a particular transaction. The goal is to minimize the amount of collected personal information by service providers, for instance, to reduce the risk of profiling and tracking users.
- *Unlinkability:* this property is essential for user privacy support and is closely related to the anonymity property. Unlinkability of two or more Items of Interest (IoIs, e.g., users, messages, actions, ⋯) from an attacker's perspective means that within the system (comprising these and possibly other items), the

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

attacker cannot sufficiently distinguish whether these IOIs are related or not. Unlinkability is divided into two properties: issue-show unlinkability and multi-show unlinkability, as follows:

- the issue-show unlinkability ensures that any information gathered during users credentials' issuing cannot be used later to link the proof of identity to the original credential, during the authentication process,
- the multi-show unlinkability guarantees that several presentation tokens derived from the same credential and transmitted over several sessions cannot be linked by the service provider.

- *Unobservability:* this property means the undetectability of a user against all users uninvolved in an Item of Interest (IoI) and its anonymity even against the other user(s) involved in that IoI. That is, a user can use a resource or a service, without being noticed by others. Unobservability also requires that third parties cannot determine if an operation is running.

3.5 Security and privacy requirements

This section summarizes the main security and privacy requirements with respect to the NIST framework and while considering IMPETUS scenarios and identified privacy challenges. These main requirements have to be taken into account at the design stage to enhance the privacy by design principle, as presented in Table 4. Each requirement is mapped with the respective ID, as presented in *D1.2. "Requirements for public safety solutions"*. For a complete view of technical and functional requirements, please refer to D1.2. that provides the requirements for the development of IMPETUS' results (platform, specific solutions, frameworks).

Table 4 Summary of security and privacy requirements w.r.t. IMPETUS identified challenges

	IMPETUS privacy challenges			
	Data overcollection	*Inference (correlation) attacks*	*Data injection attacks*	*Loss of data and processing control*
Sensing layer	-Data minimization (Rq. ID 15) -Fine-grained access control (Rq. ID 3)	-Data minimization (Rq. ID 15) -Fine-grained access control (Rq. N 3)	-Authentication (Rq. ID 79) -Data minimization (Rq. ID 15)	-Authentication (Rq. ID 79) -Data minimization (Rq. ID 15) -Fine-grained access control (Rq. ID 3)
Collection layer	Secure communication protocols			
Processing layer	-Lightweight secure (cryptographic) embedded algorithms (Rq. ID 93)	-Data minimization (Rq. ID15)	-Accountability and verifiability (Rq ID 103)	-Secure (encrypted) storage and processing of data (Rq. ID 93) -Accountability and verifiability (Rq ID 103) -Secure monitoring (Rq. N 16)
Application layer	Secure monitoring and dependable trustworthy monitoring (Rq. ID 16, Rq ID 103)			

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.5.1 Authentication and access control

Authentication and authorization are key requirements to enhance the security of gathered and exchanged data. While they are critical and mainly rely on the identification techniques in order to allow access with respect to granted privileges, they may constitute vulnerable asset as per collection of sensitive identifying data.

Thus, authentication and authorization mechanisms have to fulfill the data minimization principle, through the enforcement of the anonymity and unlinkability properties and fine-grained access control. These requirements need to be implemented with respect to data sensitivity, in order to be compliant with legislations, such as, user-consent.

Note that the implementation of privacy-preserving authentication and mechanisms has to support the accountability properties, as required by legislation. In fact, an authorized authority or legal actor should be able to remove the anonymity of the user (data owner) in case of disputes or suspicious behaviors.

3.5.2 Secure data storage and processing

Data storage and processing protection are mainly tight to scenarios and applications. Indeed, two main privacy requirements have to be fulfilled by the deployed mechanisms, with respect to the data sensitivity, namely: the data minimization and fine-grained multi-level access control. For instance, storing and processing e-health data records require the implementation strong privacy mechanisms, such that multi-level and fine-grained access to pseudonymized data, and processing over obfuscated and encrypted data.

3.5.3 Secure communications

By secure communications, we refer to the network protocols which are considered an essential component of the smart city architectures to join different components of the smart city for collecting, sharing, and transferring data throughout the cities. The selected communication protocol has to fulfill the agreed Quality of Service (QOS) while ensuring an acceptable security level for the different interconnected devices, i.e., devices, edge nodes, cloud servers, etc. The security level refers to the implementation of secure client-to-server and/or end-to-end protocols that can be support by each functional architectural layer.

3.5.4 Secure monitoring and trustworthy dependable control

It is of utmost importance to set up a monitoring strategy which is an indispensable requisite for all systems to control the surrounding environment and detect the active attacks and anonymous behavior. The automated response systems must have access to adequate information about attacks and automatic detection of suspicious behavior. Indeed, IMPETUS systems will require appropriate strategies to disclosure the potential vulnerabilities with the aim of migrating and updating all affected parties in a timely manner.

3.6 Summary

This chapter provides an overview of privacy issues and identified vulnerabilities with respect to the deployed technologies, e.g., IOT, cloud, and AI. It also points out main security and privacy requirements that have to be considered during the design of various IMPETUS tools. The following chapter gives an overview of PETs and discusses their suitability to the different applications.

4 Privacy-preserving mechanisms

This chapter gives an overview of privacy enabling techniques and details their suitability to different IMPETUS scenarios and identified challenges. Figure 1 classifies PET into three different groups and eight categories, according to which entity is mostly involved in the privacy protection decision, i.e., which entity is supporting the main cost for privacy and whether the channel between the client and the server is affected [1].

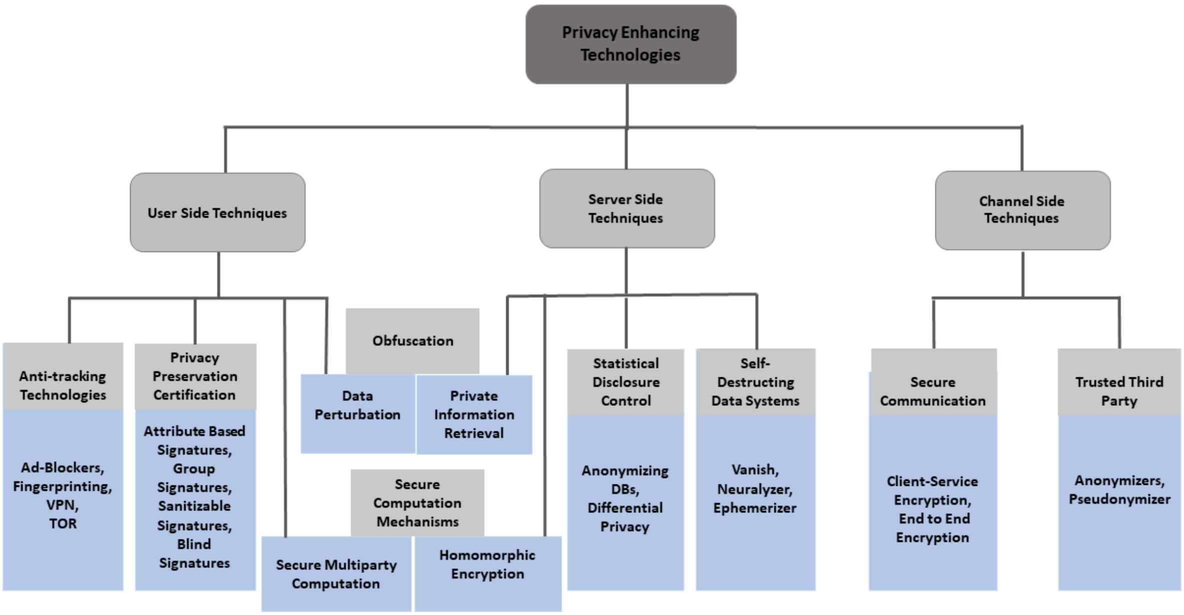
Privacy Enhancing Technologies
User Side Techniques
Server Side Techniques
Channel Side Techniques
Anti-tracking Technologies
Ad-Blockers, Fingerprinting, VPN, TOR
Privacy Preservation Certification
Attribute Based Signatures, Group Signatures, Sanitizable Signatures, Blind Signatures
Obfuscation
Data Perturbation
Private Information Retrieval
Secure Computation Mechanisms
Secure Multiparty Computation
Homomorphic Encryption
Statistical Disclosure Control
Anonymizing DBs, Differential Privacy
Self-Destructing Data Systems
Vanish, Neuralyzer, Ephemerizer
Secure Communication
Client-Service Encryption, End to End Encryption
Trusted Third Party
Anonymizers, Pseudonymizer

Figure 1 PETs Categorization [1]

As presented in Figure 1, the first category, called user-side techniques, requires the full involvement of the end-user in order to protect his privacy. User-side methods incorporate two fundamental PETs classes specifically, anti-tracking technologies, e.g., ad-blockers and anti-finger-printings; and privacy preserving certification. It also involves two sub-categories, called data perturbation and Secure Multi-Party Computation (SMC), under the obfuscation and secure computation mechanisms, respectively.

The second group, referred to as *server-side techniques*, requires the server to be firmly engaged with the privacy protection process either by anonymizing data sets for information sharing or valorization, or by performing substantial calculation over perturbated information while collaborating with end users. Server-side methods contain two classes: the Statistical Disclosure Control (SDC) and self-destructing data systems, and two sub-categories, to be specific Private Information Retrieval (PIR) procedures and homomorphic encryption algorithms, under the obfuscation and the secure processing mechanisms, respectively. It is worth noting that the obfuscation and secure computation techniques include both user side and server-side privacy-preserving procedures, and are implemented with respect to the framework's identified objectives.

The third group, named as *channel-side techniques*, specifies the nature of the channel between the client and the server - regardless of whether it is enciphered, encapsulated or encoded - or the nature of the exchanged information which can be intentionally corrupted. Channel-side procedures incorporate secure communications and Trusted Third Party such as anonymizers and pseudonymizers.

In the following, we will detail techniques that are deployed by data processors and data controllers to ensure users' privacy and be compliant with legislation. For this purpose, some of the aforementioned techniques are omitted, e.g., anti-tracking.

4.1 Privacy preserving authentication

Privacy preserving authentication, likewise known by privacy preserving certification or Attribute based certifications (AC), are cryptographic systems that permit clients to acquire certified credentials associated with their attributes from trusted authorities, and later determine presentation tokens that reveal just required data fulfilling service providers (SP)' predicates.

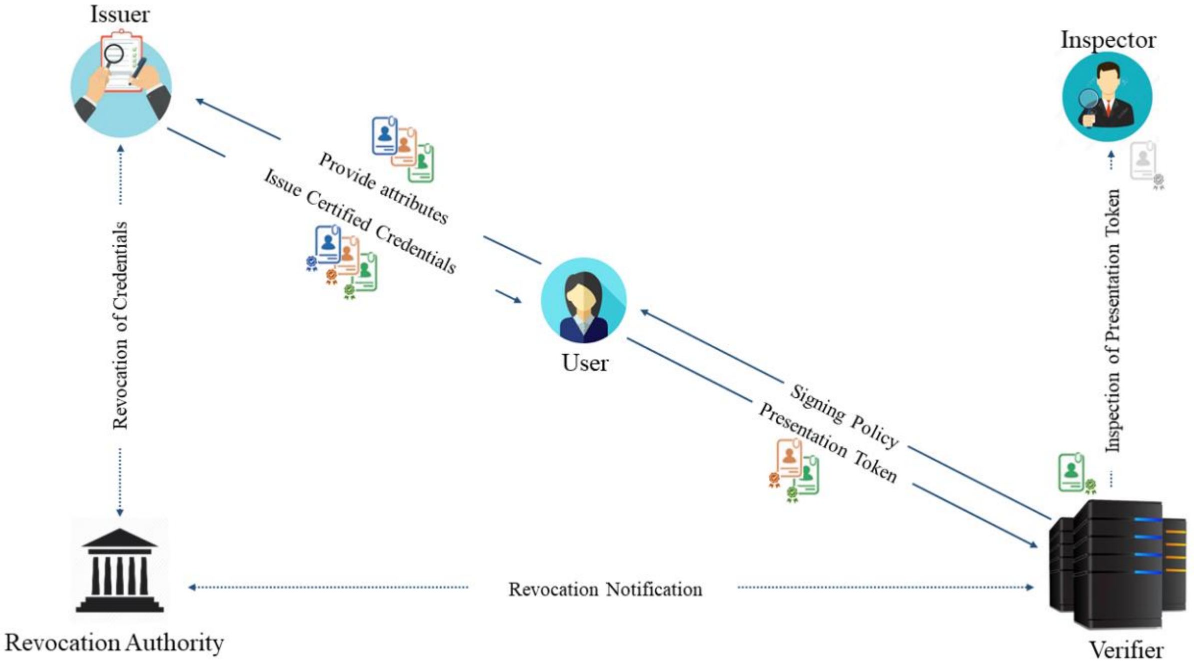

Figure 2 Privacy preserving identity management systems

Diverse substantial developments have been proposed and considered as fundamental building blocks in identity management systems. Indeed, depending on AC, every client can demonstrate to a service provider, that he holds validated properties, referred to likewise as credentials, obtained from issuing authorities. In addition, AC techniques prevent SPs to -trace- users' activities based on successive authentication sessions. That is, the user derives a proof associated to each different access request, such that the SP is not able to link a single received proof to another or to any information relative to its owner, even in case of collusion between providers and with the credential issuer, as depicted in Figure 2.

AC methods attract a lot of interest and complete consideration from industries and academia, thanks to their capacity to help the data minimization basic component The design of a privacy preserving certification scheme strongly relies on the use of malleable signature schemes that provide several interesting properties, such as the selective disclosure feature and the unforgeability property. In fact, the selective disclosure property refers to the ability provided to the user to present to the verifier partial information extracted or derived from his credential, for instance, to prove he is older than 18 to purchase liquors, while not revealing his birth date. The unforgeability property ensures that unless a user possesses a legitimate and certified credential, i.e., secret key, he is not able to generate a valid authentication proof, i.e., user's signature over the SP's access policy.

Two main industrial solutions emerged, namely IBM-Idemix and Microsoft-U-Prove. An Idemix credential relies on a variant of group signatures, generated by the issuer over the user's secret key and the attribute values to transform a credential into a presentation token, the user creates a zero-knowledge proof showing that he knows a valid signature on a committed value. On the other side, U-Prove is based on Brands signature which is a variant of blind signatures [1].

According to this viewpoint, malleable signatures are also considered as key building blocks to build privacy preserving yet authenticated access, specifically ABS [17], sanitizable signatures and group signatures [16], supporting the data minimization guideline. For example, ABS enable a client to sign a message with respect to a particular access structure defined over attributes. Every client, holding a bunch of properties, needs to acquire a private key related with his attributes from an issuing entity. Accordingly, he can sign a message w.r.t. a predicate fulfilled by any subset of his certified attributes (c.f., Figure 3). The verifier cannot deduce more than the correctness of received signature, i.e., he cannot then guess which attributes have been used.

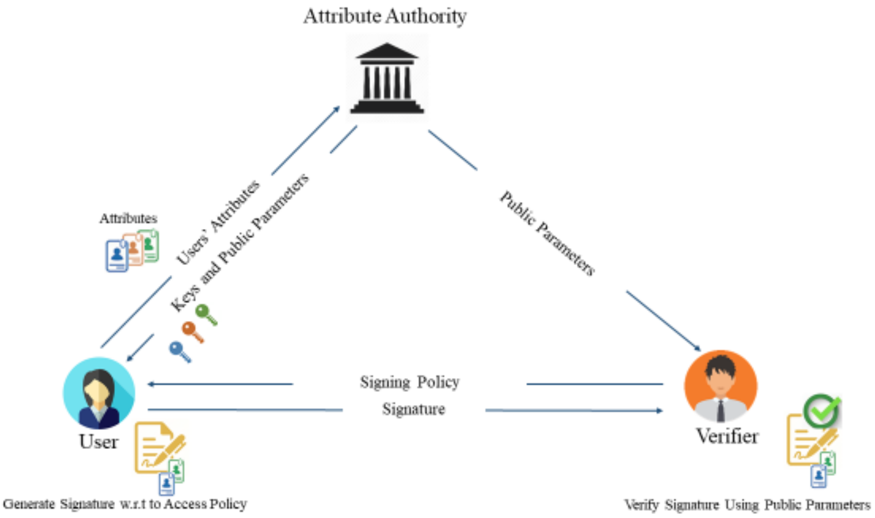
Attribute Authority
Attributes
Users' Attributes
Keys and Public Parameters
Public Parameters
User
Verifier
Signing Policy
Signature
Generate Signature w.r.t to Access Policy
Verify Signature Using Public Parameters

Figure 3 Attribute based Signatures

In the same vein, group signatures allow members of a group to generate signatures on behalf of the group they are belonging to (cf., Figure 4). As such, any verifying entity can be convinced that the generated signature has been produced by a legitimate group member, without inferring his real identity or being able to identify him. In fact, a user has first to interact with the group manager to join the group and receive related secret information. Once registered, each group member can sign messages on behalf of the group. The generated signature can then be verified by any entity, referred to as verifier that is able to check the correctness of a group signature without identifying the actual signer. In case of dispute, a designated entity is able, when needed, to identify the actual signer. Thus, group signatures permit to ensure two main privacy properties, namely, users' anonymity and accountability.

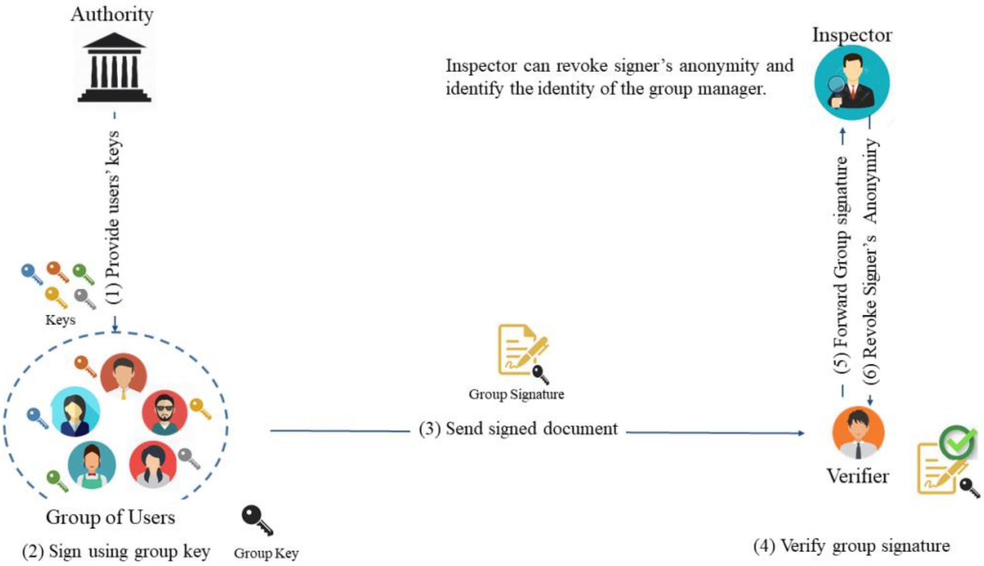
Authority
Inspector
Inspector can revoke signer's anonymity and identify the identity of the group manager.
(1) Provide users' keys
Keys
(5) Forward Group signature
(6) Revoke Signer's Anonymity
Group Signature
(3) Send signed document
Verifier
Group of Users
(2) Sign using group key
Group Key
(4) Verify group signature

Figure 4 Group signatures

4.2 Privacy preserving computations

This section describes privacy-preserving processing techniques that can be applied at both the collection and preprocessing layers.

4.2.1 Data perturbation

As displayed in Figure 1, obfuscation techniques form a subcategory of the data perturbation techniques. They aim at intentionally making information difficult to understand or perceive for security and privacy reasons. In fact, the speed of dissemination of information, the technical progress and the global nature of Internet make it difficult to delete data that may be too personal, embarrassing or confidential. Thus, obfuscation consists mainly in publishing large amounts of information that are false, imprecise, irrelevant and/or organized in such a way that the information that one wishes to protect is hidden, i.e., embedded in a large volume of data.

Data perturbation techniques are used for enhancing privacy in various querying services. In order to protect queries, one idea consists of generating dummy queries that will be sent to the central server along with the real query. The main issues of these techniques are the privacy utility trade-offs induced by the suppression technique, remove some records or details [16].

4.2.2 Secure Multiparty Computation

Privacy preserving computation techniques aim at protecting users' privacy and the secrecy of data contents during processing over these data. The goal of SMC techniques is to enable distributed computing tasks among participating entities in a secure manner. That is, SMC considers that a group of participants wants to carry a joint computation of a given function while keeping secret the input data of each party. SMC has been used to solve several privacy-preserving problems such as private database queries, secret voting, privacy preserving data mining and privacy preserving intrusion detection tools and mechanisms.

Three different approaches are generally deployed to provide secure multiparty computation functionalities, namely oblivious transfer, homomorphic encryption, and secret sharing techniques. The oblivious transfer protocol generates high processing and communication overheads. The secret sharing approach gives better results in terms of computation cost, thanks to the usage of primitive operations. However, it requires the existence of secure channels between different participating entities, hence generating a high bandwidth consumption, due to the involved interactions between users. The homomorphic encryption does not require the existence of secure channels and assures high level of privacy. However, it necessitates several processing operations to ensure homomorphism properties, thus generating high computation complexity.

4.3 Privacy preserving statistical data disclosure

Statistical Disclosure Control (SDC) components are basically used to protect data within statistical databases. They permit to resolve the trade-off between data usability and users' privacy preservation, as revealed results, either the databases or a specific result over the database is do not permit to reveal information related to a specific user. SDC techniques include database anonymizing techniques and Differential Privacy mechanisms. Anonymization techniques are relevant for various use-cases, namely applications that do not require to learn the original user's identity, but only context information. Anonymization techniques mainly refer to database privacy preservation. Even so, for cooperative applications where the database belongs to several corporations, it comes to the privacy protection of the various collaborating entities.

4.3.1 Anonymizing databases

Main techniques for anonymizing databases w.r.t. respondent, owner and users' privacy include *k-anonymity, t-closeness* and *l-diversity*. Note that these techniques that are originally used over statistical databases have extended usage to dynamic data.

K-anonymity aims to prevent the conflict between information loss and disclosure risk. For defining the k-anonymity approach, we first distinguish three types of attributes, for a microdata set S:

- *identifiers:* attributes that exactly identify the respondent, such as, his social security number or tax number. Generally, it is assumed that during a pre-processing step, identifiers in $\mathcal{S}$have to be removed or encrypted.
- *key attributes:* attributes of $\mathcal{S}$ that are useful to the application and which combination with external information can serve to re- identify respondents of the database. Examples of these attributes are: gender, age, ZIP code, ⋯ Unlike identifiers, these attributes cannot be removed from S.
- *confidential outcome attributes or sensitive attributes:* Attributes which values are of high interest for the adversary. They usually include religion, salaries, etc.

In a nutshell, to implement k-anonymity, it is important to recognize which attributes are considered as key credits, called likewise semi-identifiers. In other words, k-anonymity can forestall character divulgence, i.e., a record in the k-anonymized set S, k cannot be planned back to the comparing record in the first S, subsequently, by guaranteeing that each record is indistinguishable by essentially other k − 1 records dependent on the worth of key credits.

However, k-anonymity is not resistant to attribute disclosure attacks, as illustrated by the following example: Let us suppose that a patient's health record is k-anonymized into a group of k-patients, while considering three different key attributes, namely *Age = 42, Height = 16 and Weight = 75*. Thus, if all patients share the same confidential attribute *Disease= Cancer*, k-anonymization may be useless as an attacker may link an external record with the above group of patients, based on the key attributes. Thus, the attacker can successfully perform an attribute disclosure attack by inferring that suffers from Cancer.

For Location-Based Services (LBS), an attacker, having access to users' location, may be able to identify the requesting user, relying on its spatio-temporal parameters. Consequently, several research works propose to expand the precise location of the user to involve several potential requesting issuers. This leads to generalizing several context-data to ensure anonymity, thus resulting in the context information released to the service provider being sometimes too large and imprecise to provide an acceptable quality for the service.

4.3.2 Differential Privacy

Differential protection (DP) is acquiring a growing interest, primarily to guarantee security saving information mining. In a nutshell, differential privacy ensures that the removal or addition of a single database item does not (substantially) affect the outcome of any analysis (i.e., the probability distribution of released items does not significantly change). This property is enforced by adding random noise, to the exact outcome. Moreover, note that differential privacy addresses data leakage attacks as even if a user has removed his data from the data set, no outputs would become significantly more or less likely.

When DP techniques are applied at the data owner side, without the need for a third party, it is called Local Differential Privacy (LDP). The main idea behind LDP is to allow users to locally perturb their input data. LDP algorithms have been applied in several contexts, and three main practical realizations are introduced, namely:

- *Google's solution:* it permits to identify popular visited web-sites (URLs) without revealing any individual user's browsing habits and interests. It combines Randomized Response with Bloom Filters to compactly encode massive sets.
- *Apple's DP technique:* it combines a set of algorithms and functions to ensure a perfect differentially private large data-sets. For instance, it relies on the Fourier transform to spread out signal information, and sketching techniques to reduce the dimensionality of the massive domain. This technique was introduced in 2016, at the keynote address of Apple's Worldwide Developers' Conference, where the company's senior vice president of software engineering Craig Federighi emphasized that Apple does not assemble user profiles. Afterwards, a patent of Apple's technique was filed.
- *Microsoft's collection:* it applies an efficient LDP algorithm adapted to repeated collection of counter data, such as daily applications' usage statistics. This solution relies on fixed random numbers to collect data over time.

4.4 Secure communications

In the context of connected cities, it is difficult to physically prevent pervasive surveillance over all physical connections. Indeed, users' generated data should be protected, namely personal information or sensitive user inputs. However, even access to public resources should be protected through security mechanisms to prevent an attacker from deducing users' patterns of browsing, profiling, service use or extracting identifiers that may be used for future tracking.

4.4.1 Client-service secure communications

To secure communications against pervasive surveillance, several service providers propose to deploy encrypted communication channels. It is important to emphasize that encrypted channels need to be implemented and configured correctly, to ensure a sufficient security level. Several technologies and protocols have been introduced, namely the well-known Transport Layer Security 1.2 protocol (TLS 1.2) and the Secure Shell (SSH) protocols). These technologies provide a confidential and conceivably authenticated channel between users and service providers. Indeed, deploying TLS 1.2, or an equivalent secure channel, for every network interaction should consider currently recommended cipher-suites. Both TLS and SSH rely on public key cryptography techniques. Consequently, users and servers are able to set up an encrypted channel with no need to share secrets. TLS 1.2 is based on a public key certificate infrastructure to ensure the authenticity of the server involved in the communication. Conversely, compromised authorities may lead to the security of channels being compromised. On the other side, SSH relies on manual user verification of the service provider's key, which is more resource-consuming for end-users. The SSH protocol also requires to ensure whether end-users are able to perform periodical verification. A number of technologies may be used for communication within an organization to ensure the security of transmitted data. For instance, the Internet Protocol Security (IPSec) permits to create secure communication tunnels between networked machines, or between networks connected by public network links. It is recommended that traffic internal to an organization (local-area network) is encrypted if it may contain user information, such as for performing back-ups or communications between application and database servers.

4.4.2 End-to-end secure communications:

End-to-end encrypted services refer to encrypted communications between end-users, meaning that the encryption layer is added at one end-user and is only cleared at the other end-user. Hence, transmitted data cannot be read by any third party including the service provider. Service providers usually need to assist users to authenticate them, in order to create an end-to-end encrypted channel. However, it is preferable that the keys used to subsequently ensure the confidentiality and integrity of data never be available to the service providers, but derived on the end-user devices. Meanwhile, several service providers may require having some visibility for either routing data contents to the correct destination, or providing value added services, i.e., w.r.t. users' experience. As such, the minimum amount of information should be exposed to service providers. Web-sites and applications can approximately identify the location of devices, for instance, based on users' IP addresses. Most IP addresses permit to report the city or metropolitan area, while others may even refer to more specific places. Although most tracking mechanisms are mainly deployed at the application layer, hiding the original IP address is a well-known solution to avoid the simple technique of IP addresses' tracking. The most common method for hiding the remote site's IP address is to use virtual private networks (VPNs) or TOR (The Onion Router).

4.5 Comparison between existing privacy-enhancing smart cities

Table 5 presents an overview of commonly deployed privacy mechanisms with respect to different enabling technologies, introduced in Section 3.2. It also describes different implemented solutions relying on the underlying PET. Table 5 shows that PETs are deployed in different contexts and various purposes, thus enabling the privacy by design for the whole data lifecycle, from the sensing to the processing and application levels.

However, it is important to emphasize that in order to ensure privacy properties and guarantee an acceptable level of privacy in a connected city, usually considered as open and complex environment, it is crucial to consider the interactions between all different actors. That is to say, it might be insufficient to implement PET to ensure the privacy requirement of a specific service or application, i.e., as a standalone tool with no consideration of the surrounding environment.

Table 6 summarizes the examples of real smart cities, as well as their applications, potential privacy issues, and what privacy protections and enabling technologies are in place. Please note that "NA" that no information has been found with respect to privacy protection.

Table 5 Summary of PETs with respect to enabling technologies

PET	Technology	Description
Anonymous Credentials	Internet of Things	Use short-lived pseudonyms for car-to-car communication
		Use of attribute-based encryption and signature for private service discovery, i.e., select peers
	Cloud	Deployment of pseudonymous architecture, based on AC
		Use of attribute-based encryption for access control
Secure Communications	Ubiquitous Connectivity	Secure public WiFi with WPA2
		Use anonymous communication to protect metadata, i.e., Tor
		Ensure correct usage of SSL/TLS with static analysis

Secure Multiparty Computation	Private Information Retrieval	Cloud	Process data with private inputs, e.g., genomic tests
			Hide access patterns to remote files and databases
	Homomorphic Encryption		Perform privacy-preserving data mining over distributed datasets Privately process data at third parties
		Internet of Things	Aggregate data over multiple users
Database Anonymization		Internet of Things	Ensure k-anonymity of sensor readings
			Use *l*-diversity or hierarchical map quantization to prevent location inference attacks against *k*-anonymity
			Cluster IoT data streams and only release clusters with at *at least k*- members
		Open Data	Release only data that satisfy *k*-anonymity, *l*-diversity, *m*-invariance, and *t*-closeness
		Ubiquitous Connectivity	Change device identifiers frequently to prevent fingerprinting, randomize browser fingerprints, insert cover traffic
Differential Privacy		Internet of Things	Apply noise to meter readings
		Open Data	Release noisy aggregates of data, e.g., public transport data or t-closeness

4.6 Summary and discussion

This chapter provides a review of most commonly deployed PET in the context of smart cities applications. It also discusses existing real-world implementations and compares the deployed solutions in various cities.

First, it is important to emphasize that due to the diversity of smart applications, different privacy technologies need to be combined to ensure an acceptable level of privacy. Indeed, smart cities combine so many technological components that it is not enough to simply apply privacy technologies to each component. Instead, we advise that the interactions between technologies and data have to be considered to design *"joint privacy technologies."* This is especially important because applications start with isolated solutions that get integrated gradually. Thus, one approach to facilitate joint privacy protection is to focus on the interfaces between different systems, on their interactions and in particular on the data flow. For example, different components in a sensor-based application may all deploy independent differential privacy mechanisms before transferring data to the processing layer. Taking this into consideration will help to define appropriate privacy enabling mechanisms for the data storage and processing.

Second, it is crucial to consider the architecture patterns that define the system's components, responsibilities, and the relationships between them. There are two main architectural design. The first group contains variations of a simple centralized architecture that does not take into account the diversity of attackers and smart city applications. The second group relies on distributed settings that are tailored to specific application areas within the smart city and may induce communication overheads.

Both joint privacy mechanisms and privacy architectures aim to integrate isolated privacy protection mechanisms into more general solutions. In smart cities, this integration is complicated not only by a large number of subsystems, but also by a large number of stakeholders. To implement joint privacy mechanisms in a coherent privacy architecture, various stakeholders should collaborate on an operational level. However, this collaboration can entail privacy risks because it may enable stakeholders to combine data from several sources.

Table 6. Comparison between connected cities: applications and privacy measures

City	Application	Privacy measures	Possible privacy issues
Chicago	Aggregated health data on city map	*K*-anonymity	Sensitive health information about individuals
Sydney	Tap-on tap-off data	Differential privacy	Data about transport usage
Japan	Collaborative control: intersection collision warning	NA	Location tracking is possible
Copenhagen	WiFi hotspots for traffic flow and safety	Aggregation and data anonymization	Location tracking is possible
Oulu	App to track running data	User consent: tracking only enabled during runs	Location tracking
Bristol	Wireless mesh and sensors on street lamps	NA	Location tracking and audio surveillance
Zwolle	Smart grid	Data minimization, aggregation, and separation of knowledge	Profiling via energy consumption
Glasgow	Intelligent street lights	Sensors detect presence, But not individuals	Location tracking
	Operations center	Compliance with UK data protection law, no information about privacy technologies	Combination and automated analysis of CCTV footage
Rio de Janeiro	Operations center	NA	Combination of CCTC footage
Waseda, Japan	Teaching robot Pepper	NA	Profiling of children
California	Surveillance robot Knightscope K5	Wifi communications are encrypted	Profiling, video and audio surveillance
Estonia	Free WiFi access	Mandatory data retention	Browsing history and location tracking
Hong Kong	Free WiFi access	WPA2 encryption, user activity is recorded and available to authorities	Browsing history and location tracking
	Companion app HK GovWiFi	NA	Profiling through phone permissions: network access, location, phone identity, etc.
Malaysia	Compulsory citizen card MyKad	NA	Card number leaks information about user

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5 Preserving privacy in IMPETUS

This chapter analyzes the sensitivity of the collected information within IMPETUS pilots to determine the appropriate dissemination level according to EU legislation, the most prevalent one in this respect being the General Data Protection Regulation (GDPR). It gives an insight on different building blocks of privacy preserving mechanisms for IMPETUS tools to be compliant with GDPR namely to ensure data minimization, user consent and also accountability (verifiability).

For this purpose, we first recall the collected data by various IMPETUS tools and expose the deployed privacy techniques. Then, we analyze the data flow of standalone tools and briefly describe their interactions within the whole platform.

5.1 Collected data

D11.3. provided a detailed analysis of the types of data and anonymization/pseudonymization techniques adopted in each data collection activity of IMPETUS.

Table 7 presents four different tools that are considered as collecting identifying information and briefly describes the adopted processes for ensuring privacy, at the collection layer. Deployed in the context of a connected city to enhance security and safety of citizens, Table 7 reviews the Social Media Detection (SMD) tool, Weapon Detection tool (WD), Physical Threat Intelligence tool (PTI) and the Human Computer Interaction tool (HCI). They, collectively, with the connection of other integrated software, collect and analyze data at different levels, in order to assist cities on the detection of anomalies and threats with respect to particular settings and events.

Therefore, the data lifecycle should be reviewed in order to assess the suitability of the deployed PET and ensure reasonable privacy levels.

Table 7 Summary of collected data and anonymization techniques

Tool	Examples of collected data	Privacy techniques
Social Media Detection tool (SMD)	- Public social media data - Public local press data *Collected data may include information such as: name, pseudonym, geo-location, gender, date of birth, e-mail, website address(es), employer, occupation, phone number, hometown address.*	Data obfuscation – metadata and identifying attributes are removed from the analyzed database, and pseudonyms are attributed. A matching database is accessible based on granted privileges.
Weapon Detection tool (WD)	- Data sensed using CCTV camera devices, with respect to two different modes: 1. Calibration mode[1]: data include raw video sequences. Biometric data is not anonymized when raw video sequences are shared during the calibration period. 2. Red Alert mode[2]: data include specific snapshots and alerts' information: (a) jpeg snapshots with a visual bounding box of the	Data obfuscation: all biometric data -in the buffer, i.e., the collecting device- is hidden using either a black rectangle to fully mask people or large pixels.

[1] "Calibration mode" refers to the process where the tool is configured for use at a particular location with a specific type of camera.

[2] "Red alert mode" refers to the situation when the tool has recognised something in the image that appears to be a weapon.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

	anomaly, (b) video sequence of the red alert with a visual bounding box of the anomaly, (c) a raw video sequence of the alert and (d) GPS coordinates of the red alert.	
Physical Threat Intelligence tool (PTI)	- License plate numbers of vehicles, i.e., may be considered an "identifier" in the context of GDPR	Data obfuscation: data is removed or anonymized using SDC techniques.
Human Computer Interaction tool (HCI)	- Neuro-physiological signals *The sensors will encompass EEG (ElectroEncephaloGram) to capture brain activity and PPG (Photoplethysmogram) to measure heart rate.*	Data obfuscation: Participants is assigned with a random ID that is not linked to a specific subject.

5.1.1 SMD

The SMD tool is a human-in-the-loop AI tool that collects and analyzes massive amounts of websites and social media data. It relies on the deployment of AI methods based on a combination of computational linguistics and ML analysis to help security professionals to detect potential threats. SMD processes pseudonymised text and metadata from social media - Twitter, YouTube, TikTok - and the comments section from local news' websites in Padova and Oslo: mattinopadova.gelocal.it, document.no, reset.no, vg.no and dagbladet.no.

The output is provided via a friendly User Interface (UI) that is represented by a dashboard with graphs for user-friendly visualization of the results (bar plots, word clouds, tables…). The content analysis is shown with different perspectives, while obfuscating identifying data of each user.

5.1.2 PTI

The PTI tool exploits big data analytics algorithms to perform i) anomaly detection and ii) event classification. The tasks are performed via data-driven approaches to construct models that will be capable to identify i) anomalies in the data distributions or ii) specific threats from the sensor data for the city of Padova and Oslo. Anomaly detection is an unsupervised data-driven approach that aims to train a predictive model that is capable of catching anomalies from data. Starting from a predefined set of threats (e.g. fire, car accident, attack with guns…), similarly to the anomaly detection task, the event classifier aims to classify the current unclassified sensor data under analysis as a particular threat or as a normal case. Therefore, in addition to the anomaly detector, the event classifier would be able to indicate also the type of the threat under analysis.

PTI tool is built on an access control based data governance approach that works at ingestion time and supports perturbation and transformation of data prior to the storage in the data lake. This approach enables the support for compliance requirements at ingestion time. Access control policies are designed and enforced to guarantee that the data stored in the data lake are compliant with the policies defined for specific services/actors (possibly mandated by laws and regulations) and directly accessible with no delay when requested.

Three main phases are defined as follows (c.f., Figure 5)

- *Phase 1: Data ingestion*
  - INPUT: sensor data, annotation rules, access control policies
  - OUTPUT: sanitized data
  -
- *Phase 2: Training phase*
  - INPUT: sanitized data
  - OUTPUT: machine learning models

- *Phase 3: Inference phase*
    - INPUT: sanitized data
    - OUTPUT: anomalies, events

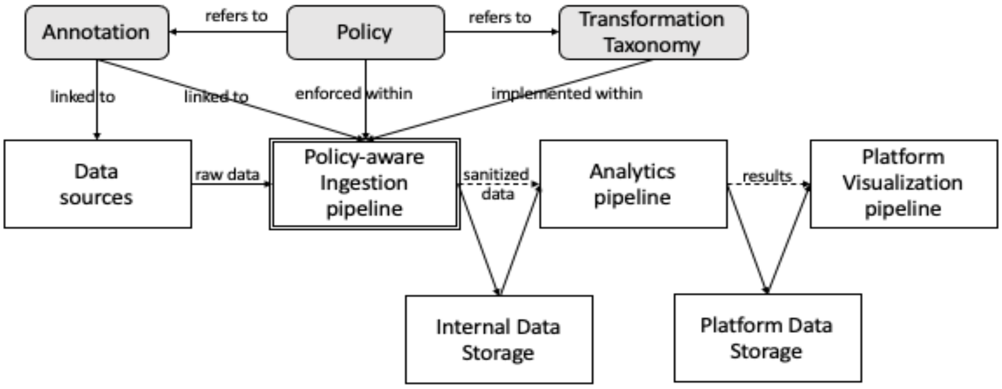

Figure 5 PTI data workflow

Figure 5 defines the data workflow. Raw data are first gathered from different data sources. This collection pipeline is made of a sequence of tasks such as data extraction, data transformation (in most of the cases format transformation) and data storing (to store data on a specific medium/storage service). Perturbed data are then processed in an analytics pipeline that is responsible to remove the anonymity of users and retrieve obfuscated data if needed (e.g., build a ML model, or apply a model for prediction).

5.1.3 WD

The WD tool aims at analyzing real-time collected data from installed CCTV in the different pilots, i.e., cities of Oslo and Padua. Two main collection modes are supported. The first mode consists of including all raw data sequences, where biometric data is shared and no privacy protection techniques are enforced. This mode is activated only when needed, for instance, in case of legitimate threat and identified risks. The second mode provides only specific snapshots and alerts' information, once a threat is detected. Data will include, jpeg snapshots with a visual bounding box of the anomaly, video sequence of the red alert with a visual bounding box of the anomaly, a raw video sequence of the alert and GPS coordinates of the red alert.

The main privacy techniques that are supported by supported by WDT are data perturbation, at the processing layer. It consists of hiding biometric data from unauthorized entities, through the application of nuances of grey pixels. The data collection and storage layer are out of scope the tool. Data will be transferred via an encrypted channel to the data controller, and stored in a protected volume, provided by the platform.

5.1.4 HCI

The HCI tool ensures an analysis of human behaviors based on collected and processed biodata in order to detect anomalies. Considered as highly sensitive information, data controllers will collect explicit and informed consent with respect to various usage (i.e., storage and sharing) of the collected biodata and –other- personal information.

Note that for accountability purposes, the data controller makes and maintains a list, with real data owners' names and linked ID numbers. The data processor, i.e., entity collecting biodata associated with subjects, does not have access to this list and only refers to anonymized records. Collected data will be then shared with the data controllers to be processed and create the personal profile for each ID. These personal profiles are then stored on designated USB sticks, each marked with the ID of the person it was made for. Backup copies of the data stored on these USB disks are kept in storage at the data controller's premises.

The data collected at the sensing layer and used to generate the personal profiles, will be permanently destroyed upon the completion of creation of said profiles.

The personal profile data only contains the ID and the information needed to interpret the measured biodata in such a way, that the assessment based on this measured data will correctly reflect the individual's cognitive workload. From now on, it is the responsibility of the data owners and controllers to (securely) handle the USB sticks. This includes the possibility of owners or the controller making copies of the data stored on the USB sticks. In case of damage or loss of the USB stick or the data it contains, new copies can be made upon request (based on the backed-up data) and sent to the supervisor. Every event will be logged.

5.1.5 Summary

This section summarizes the three main PET that are enforced by different IMPETUS tool. They are focused at the server and channel's sides and defined as follows:

1- ***Anonymization techniques****:* are widely deployed by all the tools. They rely on a random identifier referring to each data owner, while removing quasi-identifiers. This technique may provide "standalone" anonymity if quasi-identifiers have been carefully removed. However, it is still vulnerable to databases' linkability and more generally, open data [18]. Indeed, de-identifying data has no guarantee of anonymity. Released data from other tools or publicly may always contain information that can lead to the identification of individuals [18, 19].

2- ***Data perturbation:*** is mainly deployed in order to hide specific points of interests from captured CCTV streams, and also by means of aggregation while processing collected data. Data perturbation is considered as efficient in large databases, as it can be implemented without degrading the utility of the algorithm.

3- ***Data obfuscation****:* is implemented in WD. It consists on 'totally' hiding the captured data subject. While other information are still available, such as the geo-location coordinates and the raw localization videos, these are shared upon explicit consent's agreement between the data controller and the data processor.

5.2 Privacy preserving building blocks

As discussed in Section 5.1, the efficient implementation of the privacy by design principle in different IMPETUS tools requires a full consideration of the data lifecycle, and the selection of appropriate PET at each data layer. From our first analysis, we identified (i) a plausible deployment of anonymization techniques, mainly at the data collection phase, and (ii) a mild usage of secure processing at the data processing and storage levels, usually topped up with strong consent agreements between processing entities and secure client-server communication channels.

We recall that the interactions between technologies and data have to be considered to design *"joint privacy technologies."* This is especially important because applications start with isolated solutions that get integrated gradually within the IMPETUS platform. Thus, one approach to facilitate joint privacy protection is to focus on the interfaces between different systems. For instance, instead of only relying on anonymization, different tools may all deploy independent differential privacy mechanisms before transferring data to the processing layer. Taking this into consideration will help to define appropriate privacy enabling mechanisms for the data storage and processing. To do so, three main privacy-preserving building blocks should be deployed in different tools, in order to enable "global" protection of users' privacy, introduced hereafter.

5.2.1 Privacy preserving data collection

Private data collection is a key element in the design of smart applications. Thus, a careful deployment of anonymized data records before processing will increase the privacy protection at all subsequent levels.

SDC techniques are usually implemented by the service provider at the storage layer, *i.e.*, after the data collection. to provide an anonymized data-lake for further analysis and processing. Associated with data perturbation techniques, SDC can be deployed at the end-user side, e.g., by implementing LDP. This will enable an end-to-end privacy-preserving collection of data.

SDC are reputed to have an impact on the utility of processes. Indeed, recent results showed that this degradation may be costly in terms of setting-up. However, it does very slightly degrade the performances of the whole system compared to most of privacy-preserving techniques. This makes them an effective candidate to all IMPETUS tools. For instance, DP will ensure that the removal or addition of a single database item does not (substantially) affect the outcome of any analysis (i.e., the probability distribution of released items does not significantly change). Recall that this property is enforced by adding random noise to the exact outcome.

5.2.2 Authenticated multi-level access

Protecting access to sensitive identifying data constitutes the second step before sharing with the data processing and storage layers, thus potentially to third parties. It is important to emphasize that in large complex system, data should be shared by different group of users, i.e., different data controllers and data processors that have different granted privileges. Users belonging to several groups may be revoked and enrolled dynamically during the whole lifetime of the application. This requires a dynamic enforcement of access control mechanisms that supports the multi-level feature and efficient revocability. Main techniques involve attribute-based primitives and sanitization techniques.

5.2.3 Secure data processing

Secure data storage and processing is the third building block in a privacy by design deployment. It aims at protecting users' privacy and the secrecy of data contents during processing over these data. That is to say, these techniques will operate on ciphered or uncomprehensive data, thus being beyond the GDPR legislation, and ensuring a high level of security and privacy.

Secure private processing is reputed to be inefficient and costly in terms of computation and communication costs, while secure storage, based on adapted encrypted volumes is a valuable approach for IMPETUS tools and the main platform.

5.3 Conclusion

This chapter analyzed the data anonymization techniques that are deployed by different IMPETUS tools, discussed main advantages and identified challenges in large complex system.

It also presented a first analysis of main privacy preserving building blocks that should be deployed in order to ensure "global" privacy for citizens and provide balance between safety (w.r.t. utility) and privacy protection.

6 Conclusion

When designing a new smart city application, a standardized design process should assist in ensuring appropriate privacy protection. This design process needs to integrate existing methods, e.g., privacy requirements engineering and privacy testing, into a holistic process. In the software engineering discipline, many design processes have been proposed. However, it is unclear how privacy design should be integrated into these. There are some proposals to incorporate privacy by design in agile environments but it is important to find out which is the best design process associated to each scenario. Especially with regard to the high level of interconnectivity and complexity in IMPETUS, safeguarding single applications with a privacy design process might not be sufficient. Instead, there exists a need for a general design process that guides how to make a city smart with the privacy requirements of its citizens.

Privacy-enhancing technologies are often not well perceived and adopted because of a fear that they will degrade data quality to a point where the quality of the provided service is affected. Even though we have discussed several utility-neutral privacy mechanisms, e.g., DP, that can be adopted by IMPETUS tools, it is important to formally and practically assess how privacy-enhancing technologies affect the utility of services, before their integration. This problem needs to be tackled from both sides: First, operators and service providers should define more specifically what data (and with which accuracy) is required for the proper functioning of an application. Data overcollection can only be reduced if it is clear what portion of data is essential for an application, and what portion is not. Second, it is important to keep the required level of utility in mind when developing PET to increase the likelihood of adoption.

7 References

[1] Kaaniche, N., Laurent, M. and Belguith, S., 2020. Privacy enhancing technologies for solving the privacy-personalization paradox: Taxonomy and survey. Journal of Network and Computer Applications, p.102807.

[2] Chaum, D.L., 1981. Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM, 24(2), pp.84-90.

[3] Kaaniche, N. and Laurent, M., 2017. Data security and privacy preservation in cloud storage environments based on cryptographic mechanisms. Computer Communications, 111, pp.120-141.

[4] L. Cui, G. Xie, Y. Qu, L. Gao and Y. Yang, "Security and Privacy in Smart Cities: Challenges and Opportunities," in IEEE Access, vol. 6, pp. 46134-46145, 2018, doi: 10.1109/ACCESS.2018.2853985

[5] A. Gharaibeh et al., "Smart Cities: A Survey on Data Management, Security, and Enabling Technologies," in IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2456-2501, Fourthquarter 2017, doi: 10.1109/COMST.2017.2736886

[6] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2014.

[7] IoT Security Foundation. Physical security best practices. https://www.iotsecurityfoundation.org/best-practice-guide-articles/physical-security/.

[8] N. Papernot, P. McDaniel, S. Jha, M. Fredrikson, Z. B. Celik, and A. Swami. The limitations of deep learning in adversarial settings. In 2016 IEEE European Symposium on Security and Privacy (EuroS P), 2016.

[9] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. Intriguing properties of neural networks, 2013.

[10] Anirban Chakraborty, Manaar Alam, Vishal Dey, Anupam Chattopadhyay, Debdeep Mukhopadhyay: Adversarial Attacks and Defences: A Survey] [https://pdfs.semanticscholar.org/57c5/2c98730c26290b2044ad45924e58cb2fb5cf.pdf

[11] Security Matters: A Survey on Adversarial Machine Learning Guofu Li, Pengjia Zhu, Jin Li, Zhemin Yang, Ning Cao, Zhiyi Chen

[12] V. Behzadan and A. Munir, "Adversarial Exploitation of Emergent Dynamics in Smart Cities," 2018 IEEE International Smart Cities Conference (ISC2), 2018, pp. 1-8, doi: 10.1109/ISC2.2018.8656789

[13] J. Zhang, J. Chen, D. Wu, B. Chen, and S. Yu. Poisoning attack in federated learning using generative adversarial nets. In 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), 2019.

[14] T. Gu, K. Liu, B. Dolan-Gavitt, and S. Garg. Badnets: Evaluating backdooring attacks on deep neural networks. IEEE Access, 2019.

[15] Anish Athalye, Nicholas Carlini, and David Wagner. Obfuscated gradients give a false sense of security: Circumventing defenses to adversarial examples, 2018.

[16] Marco Anisetti, Claudio Agostino Ardagna, Ernesto Damiani, Paolo G. Panero: A Methodology for Non-Functional Property Evaluation of Machine Learning Models. MEDES 2020: 38-45

[17] Kaaniche, N. and Laurent, M., 2016, September. Attribute-based signatures for supporting anonymous certification. In *European symposium on research in computer security* (pp. 279-300). Springer, Cham.

[18] Samarati, P. and Sweeney, L., 1998. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression.

[19] Culnane, C., Rubinstein, B.I. and Teague, V., 2017. Health data in an open world. *arXiv preprint arXiv:1712.05627*.

Members of the IMPETUS consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Şos. Bucureşti – Ploieşti, nr. 73 – 81, Sector 1, Bucureşti, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
XM CYBER	XM Cyber, Galgalei ha-Plada St 11, Herzliya, Israel https://www.xmcyber.com	Lior Barak lior.barak@xmcyber.com Menachem Shafran menachem.shafran@xmcyber.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
T S I M E	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
UniSMART Fondazione Università di Padova	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it