PHOENIX

Electrical Power System's Shield against complex incidents and extensive cyber and privacy attacks

Deliverable D4.2
Privacy, Reputation and Mutual Auditability toolbox

Authors	Timotej Gale, Tomaž Bračič, Miha Smolnikar, David Carro Santome (CS), Francesco Bellesini (EMOT), Alessio Bianchini, Carmela Occhipinti, Elena Sartini, Luigi Briguglio (CEL)
Nature	Report
Dissemination	Public
Version	1.1
Status	Deliverable
Delivery Date (DoA)	31.10.2020
Actual Delivery Date	14.12.2020

Keywords	Privacy protection, advanced consent, reputation, mutual auditability, PRESS framework, compliance rules, governance policies, GDPR, data exchange, large-scale pilots.
Abstract	This deliverable provides the current state of the developments in the context of WP4. The deliverable presents a more detailed Privacy, Protection Enforcement component placement and an initial specification of the Privacy, Reputation and Mutual Auditability toolbox.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 832989

DISCLAIMER

	Participant organisation name	Short	Country
01	Capgemini Technology Services	CTS	France
02	THALES SIX GTS FRANCE SAS	TSG	France
03	THALES Research & Technology S.A.	TRT	France
04	SingularLogic S.A.	SiLO	Greece
05	DNV-GL AS	DNV	Norway
06	INTRASOFT International S.A.	INTRA	Luxemburg
07	Iskraemeco	ISKRA	Slovenia
08	Atos SPAIN SA [Terminated]	ATOS	Spain
09	ASM Terni	ASM	Italy
10	Studio Tecnico BFP srl	BFP	Italy
11	Emotion s.r.l.	EMOT	Italy
12	Elektro-Ljubljana	ELLJ	Slovenia
13	BTC	BTC	Slovenia
14	Public Power Corporation S.A.	PPC	Greece
15	E.ON Solutions Gmbh [Terminated]	EON	Germany
16	Delgaz Grid SA	DEGR	Romania
17	Transelectrica S.A.	TRANS	Romania
18	Teletrans S.A.	TELE	Romania
19	Centro Romania Energy	CRE	Romania
20	CyberEthics Lab	CEL	Italy
21	GridHound GmbH [Terminated]	GRD	Germany
22	Synelixis Solutions S.A.	SYN	Greece
23	ComSensus	CS	Slovenia
24	AALTO-KORKEAKOULUSAATIO	AALTO	Finland
25	Rheinisch-Westfälische Technische Hochschule Aachen	RWTH	Germany

26	Capgemini Consulting [Terminated]	CAP	France
27	ATOS IT Solutions and Services Iberia SL	ATOS IT	Spain
28	DNV GL NETHERLANDS B.V.	DNV-NL	Netherlands

ACKNOWLEDGEMENT

This document is a deliverable of PHOENIX project. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement Nº 832989.
The opinions expressed in this document reflect only the author's view and in no way reflect the European Commission's opinions. The European Commission is not responsible for any use that may be made of the information it contains.

Document History

Version	Date	Contributor(s)	Description
0.1	13.10.2020	CS	Initial ToC.
0.2	30.10.2020	CS	First draft.
0.3	24.11.2020	CS	Technology specification.
0.4	2.12.2020	CS	Interim internal review.
0.5	7.12.2020	CS	Insert missing content.
0.6	8.12.2020	CS, EMOT	LSP use cases.
0.7	9.12.2020	CEL	Interim review.
0.8	10.12.2020	CS	Finalized for review.
1.0	14.12.2020	CS	Address reviewers' comments and finalize for submission.
1.1	30.06.2021	SYN	Resubmission – based on review feedback, fixed the references

Document Reviewers

Date	Reviewer's name	Affiliation
11.12.2020	Jose-Ramon Martinez-Salio	Atos IT
11.12.2020	Francesco Bellesini	Emotion

Table of Contents

Figure List

Table List

Definitions, Acronyms and Abbreviations

AAM	Accountability and Access Management
API	Application Programming Interface
BC	Blockchain
BoEU	Block of Energy Unit
CA	Consortium Agreement
CoreDEX	Compliant Regulatory Data Exchange
CTI	Cyber Threat Intelligence
CybOX	Cyber Observable Expression
DACR	Data and Advanced Consent Registry
DC	Data Controller
DID	Decentralized Identifier
DLT	Distributed Ledger Technologies
DoA	Description of Action
DP	Data Processor
DR	Demand Response
DS	Data Subject
DSO	Distribution System Operator
EC-GA	European Commission Grant Agreement
EEA	European Economic Area
EPES	Electrical Power and Energy System
EU	European Union
EV	Electric Vehicle

EVM	Ethereum Virtual Machine
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
IBFT	Istanbul Byzantine Fault Tolerance
IMEC	Incidents Mitigation and Enforcement Countermeasures
IODEF	Incident Object Description Exchange Format
IT	Information Technology
LSP	Large-Scale Pilot
LV	Low Voltage
MV	Medium Voltage
OpenIOC	Open Incident of Compromise
PHOENIX	Electrical Power System's Shield against complex incidents and extensive cyber and privacy attacks
PK	Public Key
PPE	Privacy Protection Enforcement
PRESS	Privacy, Data Protection, Ethics, Security and Societal
PRIMULA	Privacy, Reputation and Mutual Auditability
RES	Renewable Energy Sources
SAPC	Situation Awareness, Perception and Comprehension
SC	Smart Contract
SCC	Secure Control Centre
SCADA	Supervisory Control and Data Acquisition
SK	Secret Key

SPC	Secure and Persistent Communications
STIX	Structured Threat Information Expression
TSO	Transmission System Operator
UC	Use Case
URL	Uniform Resource Locator
USG	Universal Secure Gateway
UUID	Universally Unique Identifier
WP	Work Package
XACML	Extensible Access Control Markup Language

Executive Summary

This document provides an overview of the privacy protection enforcement in the scope of the PHOENIX project. It defines the conceptual model and specifications related to privacy, reputation and mutual auditability with emphasis on Privacy, Data Protection, Ethics, Security and Societal (PRESS) Framework (see D4.1 [1]) and its compliance rules and governance policies. The deliverable positions the Privacy Protection Enforcement (PPE) component by identifying the inner components and interactions with external ones (i.e., rest of PHOENIX architecture). Privacy, Reputation and Mutual Auditability (from now on PRIMULA) toolbox is specified and presented along with design decisions and proposed technologies. This deliverable also illustrates the Large-Scale Pilot (LSP) scenarios in the context of the PPE component.

Chapter 1 provides a brief explanation on the objectives of the PHOENIX Project, the present deliverable and on the structure of the document.

Chapter 2 presents the privacy protection enforcement background. The transactions, associated with confidential and personal data handling by data controllers, are identified. PRESS framework (that includes GDPR) lays the foundation for building PPE components that ensure privacy protection. This deliverable introduces the PPE data model, and then focuses on main concepts such as Privacy, reputation, mutual auditability and advanced consent.

Chapter 3 describes the PPE component and places it in the context of PHOENIX framework. The interactions and actions of PPE and its subcomponents are specified.

Chapter 4 provides an overview of the PRIMULA toolbox, the design decisions and the proposed technologies.

Chapter 5 lists the LSP specific requirements and presents the PRIMULA/PPE use cases associated to LSP1 and LSP3, that are the first LSPs identified during the preparation of this document. Other PPE instances in LSPs will be evaluated during the rest of project lifecycle.

1. Introduction

The primary objective of the PHOENIX project is the protection of the European Electrical Power and Energy System (EPES) assets and networks against cyber-attacks. An important aspect of protection is information sharing, which – in PHOENIX scope – takes the form of Cyber Threat Intelligence (CTI). As exchanged CTIs and other EPES data may inhibit privacy and confidentiality constraints, appropriate mechanisms need to be developed to enable data exchange according to required governance policies and compliance rules, specified in the PRESS Framework (see D4.1 [1]).

This deliverable is produced within the Work Package (WP) 4. Aside from an overview of the privacy protection concepts, it provides the current state of development in WP4. The deliverable focuses on the positioning of the Privacy Protection Enforcement (PPE) component within the PHOENIX architecture and platform, as well as the specification of the Privacy, Reputation and Mutual Auditability (PRIMULA) toolbox. The document serves as a basis for further development within the WP4 and coordination among the other WPs.

1.1. Relation to Project Work

The general indications for the Project deployment have been defined in the European Commission Grant Agreement (EC-GA), the Description of Action (DoA) and the Consortium Agreement (CA). The present deliverable **D4.2 – Privacy, Reputation and Mutual Auditability toolbox** – as well as the other deliverables – does not replace any of these established agreements and Project deliverables, and Partners should abide by the order of precedence reported in Table 1.

Table 1: Order of precedence of PHOENIX agreements and deliverables.

1	European Commission Grant Agreement (EC-GA)
2	Commission Rules
3	Consortium Agreement (CA)
4	Project Handbook
5	D4.1: PRESS Framework

1.2. Document Structure

The document is divided into the following chapters (Table 2):

Table 2: Structure of D4.2 – Privacy, Reputation and Mutual Auditability toolbox.

	Chapter title	Summary
Chapter 1	Introduction	It provides a brief explanation on the objectives of the PHOENIX Project, the present deliverable and on the structure of the document.

Chapter 2	Privacy Protection Enforcement	It provides the privacy protection enforcement background.
Chapter 3	PPE Positioning in PHOENIX Framework	It describes the PPE component and places it in the context of PHOENIX framework.
Chapter 4	PRIMULA toolbox	It provides an overview of the PRIMULA toolbox, the design decisions and the proposed technologies.
Chapter 5	LSP Integration	It outlines the LSPs specific requirements and presents the PRIMULA/PPE use cases associated to LSP1 and LSP3.
Chapter 6	Conclusions	It provides the conclusions of this deliverable and the follow up of its outcomes.
Chapter 7	References	It provides the list of references used for the preparation of this deliverable.

2. Privacy Protection Enforcement

To ensure privacy and protection of personal and confidential data, privacy-preserving data handling mechanisms need to be developed. These mechanisms must comply with legal frameworks and regulations in force, as well as enforce the constituted principles. In practice, PPE's function is dual; privacy protection enforcement on local and global level is available by design whereas the component additionally provides important insights and serves as an aid for the operators (and possibly regulators) by granting accessible and comprehensive auditing capabilities.

Personal and confidential data goes through various lifecycle phases, such as collection, management and processing, with different constraints and requirements applying to each phase. Data controllers[1], as the main liable entity, must ensure proper data handling in each phase. To enable better analysis, development, integration and verification of various phases, the following Table 3 shows the identified transactions, associated with confidential and personal data handling by the data controllers.

Table 3: Identified transactions for personal data handling.

Category	Transaction	Description
Data management	Data registration	Data must be annotated as personal/confidential while possibly additional metadata may be provided (e.g., involved data subjects).
	Data update	In case of request from data subject, existing personal data has to be updated, or new data could be added.
	Data deletion	In case of request from data subject, personal data has to be deleted (art.17 GDPR).
	Data exchange	Data may be transferred among different data controllers/processors, that could be distributed in different locations according to and complying with specified permissions/consents.
	Data processing	Data controller may act as a data processor according to and complying with specified permissions/consents.
Permissions definition	Permissions granting	Data controller may grant permissions for specific data processing to data processors, in case of request from data subject.

[1] https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/controller-processor/what-data-controller-or-data-processor_en

	Permissions revocation	Data controller may revoke permissions for data processing to data processors, in case of request from data subject.
	Permissions update	Data controller may update data processing permissions for data processors, in case of request from data subject.
Consent management	Consent granting	Data subject may grant consent for data processing.
	Consent revocation	Data subject may revoke consent for data processing.
	Consent update	Data subject may update consent for data processing.
	Consent request	Data controller may request consent from data subjects.
Rules and policies	Rules and policies specification	Rules and policies for data and data operations may be specified.
	Rules and policies enforcement	Rules and policies must be enforced on data or data operations.
Auxiliary	Auditing	All operations must be recorded.
	Violation detection	Data controller must detect any data processing abnormalities.
	Notification	Data subject must be notified about any data processing as well as violations (e.g., data breaches).

In the following subsections is provided an overview of the PRESS framework (that includes GDPR) and its implications for PHOENIX privacy protection enforcement. The main concepts of privacy, reputation, mutual auditability and advanced consent are defined. A privacy-based data model and the related PPE envelope are introduced.

2.1. PRESS Framework and GDPR

PRESS framework aims to provide a comprehensive conceptual framework to the PHOENIX project. Delivered in D4.1 [1], the PRESS framework – focusing on privacy and data protection as fundamental rights/requirements, ethics, security, and social concerns – is composed of three pillars:

- Privacy and data protection requirements;
- Ethics and Social requirements;
- Security requirements.

Based on these pillars, a practical set of compliance rules and governance policies are provided to the PHOENIX project to be considered when developing PHOENIX architecture and components. As part of PRESS analysis, General Data Protection Regulation (GDPR) was investigated. GDPR is a law regulation in European Union (EU) and European Economic Area (EEA) governing data protection, privacy and transfer of personal data.

PRESS Framework is indeed relevant and serves as a basis for all the PHOENIX components, including the PPE, as its main function is to ensure the compliancy with rules and policies. PPE virtually serves as an authority and a mediator between entities with assigned GDPR roles in PHOENIX (i.e., data subjects, data controllers and data processors) (Figure 1).

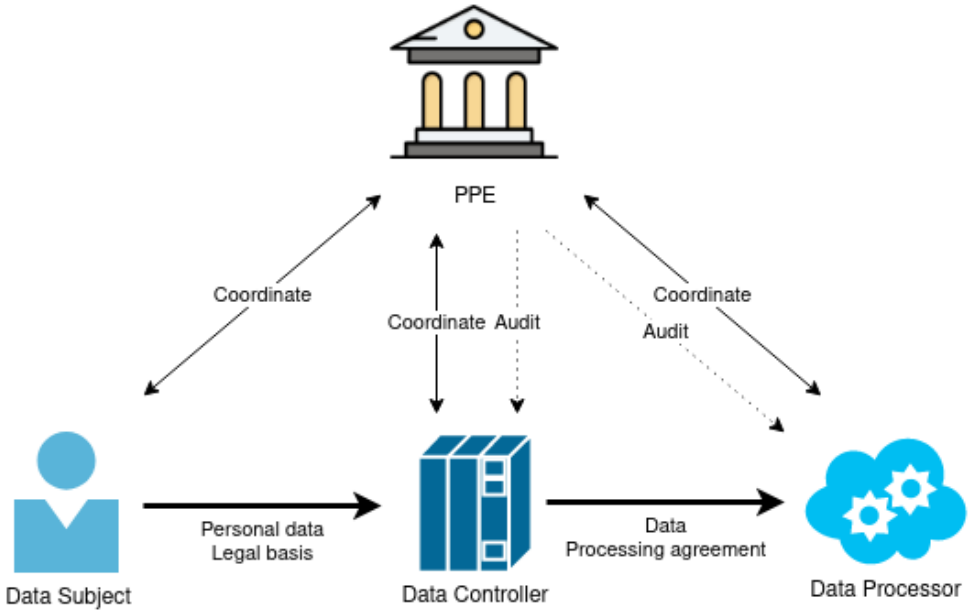

Figure 1: PPE as a GDPR authority and mediator.

In D2.2 [2], a template for test and validation report of the Secure and Persistent Communications (SPC) components was provided in order to comply with the PRESS framework. Since all PHOENIX components should adopt the PRESS framework, tools that administer the inclusion of PRESS recommendations, suggested mechanisms and techniques into components' specification are needed. Upon finalizing the specifications and based on conclusions emerging from deployment experiences, an extended checklist template for test and validation report of the PPE components that better captures the specific requirements and peculiarities of handling privacy-sensitive data will be provided in consequent deliverables (D4.3/D4.4: Cross-GDPR sensitive data exchange toolbox).

2.1.1. Advanced Consent

Generally, according to the PRESS framework, unless explicitly allowed by law or consented to by the involved data subject, the processing of personal data is prohibited. Consent is one of the legal bases for data processing, introduced in Article 6 of the GDPR. A data subject must provide consent prior (in advance) to any personal data collection and processing. Consent of the data subject, as defined in Article 4(11) of GDPR, is "any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her" [3]. Consent conditions are additionally formulated in Article 7 of GDPR [3]:

1. "Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data" [3].
2. "If the data subject's consent is given in the context of a written declaration which also concerns other matters, the request for consent shall be presented in a manner which is clearly distinguishable from the other matters, in an intelligible and easily accessible form, using clear and plain language. Any part of such a declaration which constitutes an infringement of this Regulation shall not be binding" [3].
3. "The data subject shall have the right to withdraw his or her consent at any time. The withdrawal of consent shall not affect the lawfulness of processing based on consent before its withdrawal. Prior to giving consent, the data subject shall be informed thereof. It shall be as easy to withdraw as to give consent" [3].
4. "When assessing whether consent is freely given, utmost account shall be taken of whether, inter alia, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract" [3].

Consent must be freely given, specific, informed, explicit and unambiguous. To be freely given, consent must be given on a voluntary basis by the data subject without pressure or influence. Another requirement is specificity; it must be apparent what data processing activities will be performed. The data subject must also be informed about the controller's identity, the purpose(s) of the data processing, the processing activities to be carried out and the data subject's right to withdraw the consent at any time. When requesting consent, clear and plain language must be used. Consent cannot be implied as a clear affirmative act is required for an unambiguous indication of consent. The GDPR does not specify a time limit for consent, however, a data subject has the right to revoke the consent at any time. Revoking consent must be as straightforward as giving consent. Whereas the GDPR does not restrict a third party from giving consent on behalf of a data subject, the third party's authority must be demonstrated.

2.2. Privacy, Reputation and Mutual Auditability

According to D4.1 [1], the term privacy can be considered as "providing for a general prohibition of interference in the private life of an individual". Privacy is represented by the ability to hide information and disseminate information by choice, such information being, at times, sensitive. Security, a directly related term, incorporates approaches of proper use and protection of information. In information technology (IT) domain, the term privacy is mainly associated with data. This concept determines the ability to regulate which data in a computer system is shared, with whom and under which conditions and procedures. Whereas data privacy dictates the means of data collection, sharing and usage, the data security deals with protection from data compromise.

Any data handling is associated with a certain degree of risk of data leakage. Data can be exposed while in transit (i.e., moving from one location to another) or while at rest (i.e., stored data). In order to prevent unauthorized data access, adequate security measures must be developed. Whereas data access policies may be used to define the responsibilities and roles of actors who are granted access to data, a very effective method for data protection – in transit, as well as at rest – is data encryption. Encryption is the process of encoding information, where plaintext data is converted into a ciphertext that is intelligible to third parties [4]. The encryption is usually performed by algorithms based on pseudo-random keys. A

key is also used by the recipient to decipher the message and access the original plaintext. In modern cryptography, the concepts of asymmetric encryption, where one key is used for encryption and a different key is used for decryption, and symmetric encryption, where one key is used for both encrypting and decrypting data, have been popularized. The decryption without a key is in theory possible, but due to the computation cost practically unfeasible.

In some cases, where total privacy is not required, a large computation overhead is not desired, or where data must be shared and at least some level of privacy is necessary, various privacy-preserving data transformations may be used. An example of such transformation is differential privacy [5]. In essence, this approach utilizes injection of noise (e.g., Gaussian or Laplace) into the data in order to increase privacy. The amount of added noise presents a trade-off between protection and data utility.

Privacy protection may also be achieved through anonymization, a technique by which identifying and quasi-identifying attributes are removed from the data. An alternative procedure to achieve near-anonymity is by adopting pseudonymity. Since permanent unique identifiers, especially identifiers that are used in multiple contexts, still pose a privacy risk as devices or persons may potentially be tracked and characterized, changeable identifiers have been previously considered. A possible approach for implementing changeable identifiers is called decentralized identifiers (DIDs) and has been proposed in [6]. DIDs have no central controlling authority and are managed by identity owners or their appointed trusted entity. This concept is known as self-sovereign identity [7]. Identifiers have also been problematized in the context of blockchain (BC) technologies. With the arrival of GDPR and emergence of distributed ledger technologies (DLTs), a critical problem surfaced; as blockchains are immutable, the right to be forgotten could not be assured. Several solutions have been identified to tackle this problem: private key destruction [8], block pruning [9], blockchain forking [10], … In [11], an alternative approach based on burnable pseudo-identities is considered. This approach implements a mechanism of continuous refresh and rotation of identities.

Reputation, a measure for facilitating trust among entities based on global perception of behavior, is paramount as more people and services with no prior direct relationship interact [12]. The fundamental goal of leveraging the notion of reputation is formation of trust/distrust among unfamiliar entities, which is directly proportional to the perceived risks. Reputation systems enable assessment of trustworthiness based on past transactions and are societal corrective, since good/bad behaviors are incentivized by positive/negative reputation in the long term. With trust and reputation management systems based on centralized approaches exhibiting several issues, the most prominent being the need to trust a central authority, distributed trust and reputation management systems emerged. Some of these systems also adapt the BC technology. A comprehensive overview is provided in [13].

Auditability is defined as the auditor’s capacity to realize a comprehensive review of records. The rise of the BC technologies enabled improved accountability and audit capabilities. In addition to mechanisms for verifying accountability, the BC ensures authenticity of records and non-repudiation through the use of public key infrastructure and digital signatures. The immutability of the ledger on which BC transactions are saved, thus keeping an audit trail, provides integrity by preventing alteration of the records. In BC, mutual auditability (i.e., the ability of all parties to verify their actions and actions of other parties in the system) is available by design.

2.3. PRESS Data Model

Data originating from electricity power and energy systems (EPES) as well as related subsystems and sources may contain private information. This includes personal data or any data from which personal information can be derived from or may identify a physical person, or any data requiring a certain level of secrecy. For example, meter data – if not properly anonymized - has been previously identified as problematic since various patterns can be extracted from power consumption and be used to classify and monitor human behavior [14]. Furthermore, cyber threat information (CTI) may disclose important information about business processes or expose business secrets, consequently harming the organization, and a country (due to the fact that CTI contains information impacting critical infrastructures). To provide a common framework for handling various types of privacy-constrained data, we introduce the notion of *PRESS data* (see Figure 2). PRESS data is an abstract class for all categories of personal and confidential data.

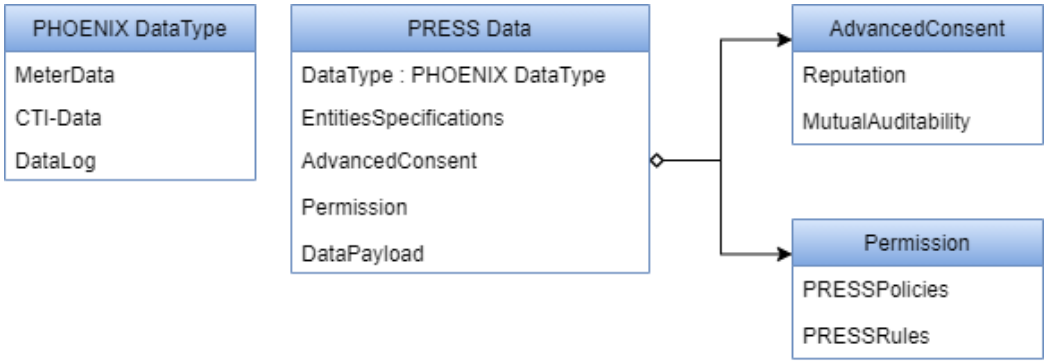

Figure 2: PRESS data model.

PRESS data, by definition, has various constraints and associated features that may be attributed to these data. We introduce the concept of *PPE envelope* (Figure 3) that encapsulates data's metadata alongside data. PPE envelope provides a logical view of privacy-sensitive data by capturing all associated privacy aspects, thus enabling straightforward data modeling, data flow construction and privacy enforcing system design. The following metadata may be in part or fully assigned to PRESS data:

- **Data type:** Characterization of data's structure. Standardized ontologies are preferred; however, an arbitrary data structure may be adopted. Data structure definition is necessary for advanced access control and advanced consent/constraints applications.
- **Entity specification:** Characterization of implicated entities (data subjects, data controllers, data processors).
- **Advanced consent:** Data subject's consent specification. Consent format is implementation-specific and is to be defined at the implementation stage but would normally take a form of a human-readable string or a machine-readable structure.
- **Permission:** Data access control policy and roles definition. Internal organizational authorization schemas in addition to interorganizational authorization schemas on data subject/controller/processor level formalization. Permissions are dependent on given consent

and additional (privacy) constraints applying to data, e.g., cross-border exchange limitations. Permissions are implementation-specific and are to be defined at the implementation stage but would normally take a form of a machine-readable structure or adopt existing authorization mechanisms (e.g., permissioned blockchain – see Section 4.2.1).

The PPE envelope format will be defined during the implementation step and reported in consequent deliverables (D4.3/D4.4).

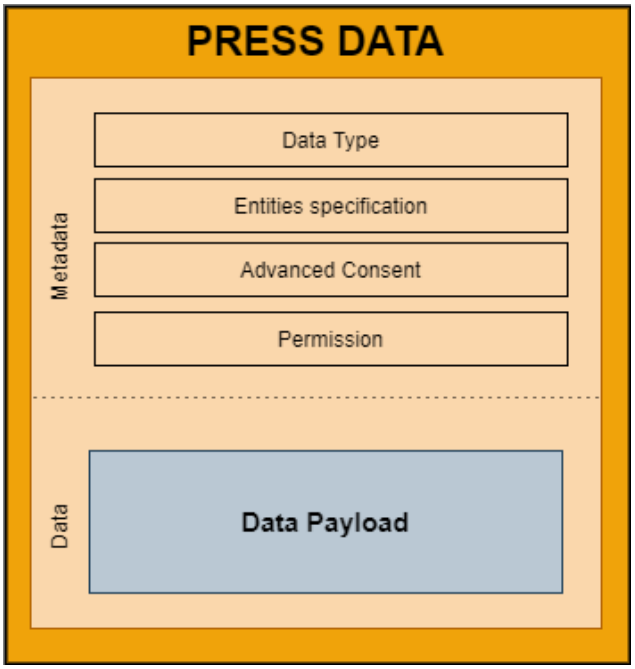

Figure 3: PPE envelope.

2.3.1. Data Structure

Data structure definition enhances interoperability and is necessary for advanced access control and advanced consent/constraints applications as understanding the data structure is a requirement for granular data policy specification. Standardized ontologies are preferred; however, an arbitrary data structure may be adopted. In cases where data structure is not known, the data may only be treated as a binary blob where any specified data policies apply to the whole data object.

Ontologies are a semantic technology enabling computer interpretation of data models using standardized data representations, knowledge can be acquired from represented information using various concepts and relationships between those concepts [15]. Ontologies represent abstract models that are independent from a certain language. Ontologies are normally domain-specific and facilitate knowledge sharing as well as knowledge consolidation.

PHOENIX project is mainly focused on cyber threat intelligence (CTI), several ontologies pertaining to CTI exist: Incident Object Description Exchange Format (IODEF), Open Incident of Compromise (OpenIOC),

Cyber Observable eXpression (CybOX), Structured Threat Information eXpression (STIX). STIX[2] has been previously identified as a suitable ontology for CTI sharing in PHOENIX project. STIX provides a standardized, community-driven, widely adopted and structured representation of threat information. Whereas the STIX format is human readable, it also provides a comprehensive CTI notation.

2.3.2. Data Access and Transmission

Data access policies are a fundamental element of data privacy. Proportionate and adequate security measures and appropriate data access authentication/authorization roles must be developed to prevent intrusion. Internal organizational authorization schemas, in addition to interorganizational authorization schemas, on data subject/controller/processor level must be specified. When personal data is considered, the access policy is decidedly dependent on given consent. Various tools for defining access policies exist, the most prominent being the eXtensible Access Control Markup Language (XACML) [16]. Several ontology-based access control solutions are available [17] [18].

XACML is a standard that defines an architecture, a processing model and a declarative attribute-based access control policy language. While providing a fine granularity of access rules, the standard also promotes common terminology and interoperability between access control implementations. The current version 3.0 was ratified by OASIS standards organization in January 2013. XACML provides functions for manipulation and attribute comparison in addition to time-based authorization. XACML defines the following policy elements:

- **Policy set, policy, rule:** Rules constitute a policy, a policy set can contain several policies or policy sets.
- **Subject, resource, action, environment:** Main entities in rules, policies and policy sets.
- **Attribute, category:** Subject, resource and action elements may have one or more attributes.
- **Target:** A set of conditions that must be met before applying a policy set, policy or rule.
- **Condition:** Conditions constitute rules and are used to compare attributes.
- **Obligation, advice:** A directive on (mandatory) actions before/after access is approved.

Many open-source as well as commercial XACML implementations exist: Balana, Axiomatics Policy Server, SunXACML, ...

An additional important aspect of data privacy is data form and data transmission. In addition to encryption, end-to-end encryption is encouraged to prevent unwanted access while in transit. In some cases, data must be transformed prior to transmission to ensure privacy or reduce complexity; by employing differential privacy, noise is added to the data in order to increase protection (see D2.2 [2]). When personal data are stored, applying data referencing using pseudonymization and encrypting data pointers is required. To store mappings of subjects and personal data references, a cryptographically authenticated data structure like Merkle-Patricia trees[3] should be used.

[2] https://oasis-open.github.io/cti-documentation/
[3] https://eth.wiki/en/fundamentals/patricia-tree

3. PPE Positioning in PHOENIX Framework

The Privacy Protection Enforcement (PPE) enables handling and management of PRESS data in compliance with EU Regulatory Framework based on the PPE envelope concept (see Section 2.3). It is built as a part and on top of a common PPE DLT. Moreover, PPE provides the PHOENIX development and operation team with compliance rules and governance policies (PRESS Framework) to be adopted by all the PHOENIX components and services. The PPE component focuses on three main aspects, each of them represented by its corresponding subcomponent:

- **Privacy, data Protection, Ethics, Security and Societal (PRESS) framework** provides compliance rules and governance policies.
- **Privacy, Reputation Mechanism and Mutual Auditability (PRIMULA) toolbox** enables consent-driven exchange of PRESS data through managing the advanced consent among parties (data subjects, data controllers, data processors) via smart contracts (see Section 4.2.2). PRIMULA provides consent auditing functionalities and reputation mechanisms to assess the party's reputation based on global perception of its behavior.
- **Compliant Regulatory Data Exchange (CoreDEX)** enforces compliance rules and governance policies from the PRESS framework by acting as a mediator for data exchange among parties. CoreDEX implements appropriate mechanisms/protocols for data transfer authorization to other PHOENIX components and mechanisms/protocols for direct data transfer between two parties without intermediary channels. The authorization is dependent on advanced consent and reputation mechanisms from PRIMULA.

Emanating from transactions identified in Chapter 2, the actions of the PPE component are depicted in Figure 4. On the left side are the input actions, that are triggered by other components. On right side are PPE's output actions, which affect other components. Some input actions have corresponding output actions; these pairs of actions are a result of action forwarding or immediate response operation. The actions are related to PHOENIX components, assigned to PPE inner components and further discussed in Section 3.2.

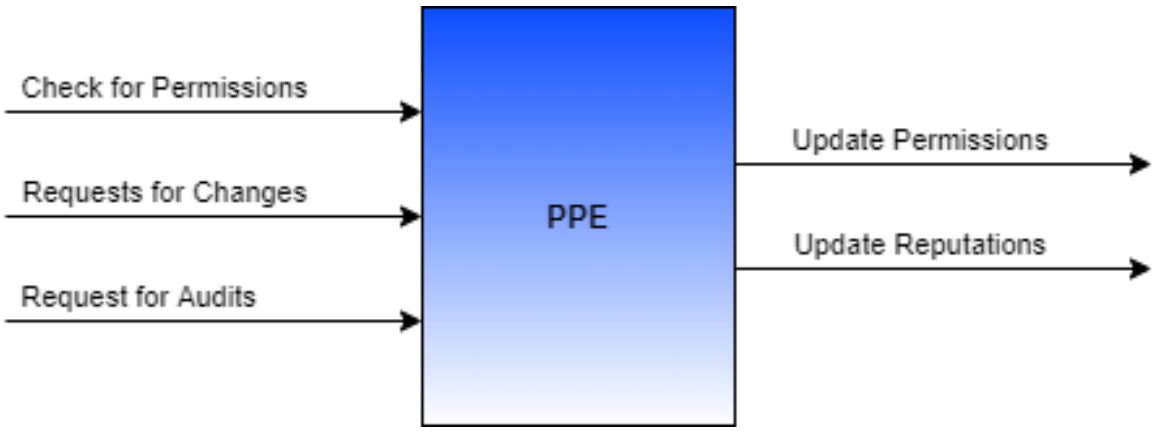

Figure 4: PPE Interactions with external components.

To place the PPE component in a broader PHOENIX framework scope, the envisioned solution based on initial requirements and PHOENIX components' interactions are presented in the following subsections.

3.1. PPE Requirements

This section lists the PPE security and functional requirements and outlines how these requirements will be reflected in the developed solution (see Table 4: PPE requirements and envisioned solution.Table 4).

Table 4: PPE requirements and envisioned solution.

Requirement ID	Requirement Name	Requirement Description	Envisioned Solution
PRI01	Confidentiality	The solution has to prevent unauthorized disclosure of personal or privacy-sensitive data between data owners and DPs.	CoreDEX will allow only authorized data exchanges as defined by consents managed by PRIMULA.
PRI02	Auditability	The solution has to enable mutual auditing and investigation of consent proofs in case of non-compliant activities.	All consent and permission changes as well as any related actions will be recorded in PPE DLT.
PRI03	Transparency	Allow for data owner to have a complete transparent view over the collected, shared and processed data.	All personal and privacy-sensitive data will be registered with PPE. All operations – including processing and sharing – will be recorded in PPE DLT. CoreDEX will provide access to data whereas PRIMULA will provide advanced consent through PHOENIX components or existing EPES infrastructure.
PRI04	Change requests	Allow for data owner to change permitted actions over its data: contract deletion, contract inactivation and modification of data usage policies	PRIMULA will enable advanced consent change through PHOENIX components or existing EPES infrastructure.
CRD01	Consent-based Exchange Rules	Data exchange will be allowed only on Consent-based Exchange Rules. This also includes checks on data confidentiality and data sensitivity.	CoreDEX will permit or prohibit data exchange according to advanced consent from PRIMULA as well as

			additional CoreDEX compliance rules and governance policies.
CRD02	Advanced Consent	Advanced Consent for data exchange is signed by data subject data controller (DS DC). Advanced Consent explicitly and clearly specifies what data is needed, its purpose, and how and who will process the data. Furthermore, it checks if the parties have the proper rights to access the data.	DS signs consent through PHOENIX components or existing EPES infrastructure. Consent and its specification (purpose, …) is recorded in PRIMULA. Upon data request, CoreDEX checks whether the requesting party has the rights to access data.
CRD03	Change Requests	DS has always the right to withdraw or request the change of the data exchange rules. Thus, the system has to provide the ability for managing requests for changes.	PRIMULA will allow withdrawal or change of consent. CoreDEX will allow change of exchange rules.
CRD04	Status Notification	Parties involved in the data exchange are always notified about the status of data exchange processes.	All data exchanges are recorded in PPE DLT, PRIMULA will notify the parties through PHOENIX components or existing EPES infrastructure.
CRD05	Right to be forgotten	DS has always the right to request personal data erasure (art. 17 GDPR). Thus, the storage has to allow the deletion of personal data.	PHOENIX components or existing EPES infrastructure will implement data erasure mechanisms.

3.2. PHOENIX Components Interoperability

An overview of the PPE component, its subcomponents and interactions are depicted in Figure 5. PRIMULA and CoreDEX are built on top of a common PPE DLT and provide privacy protection enforcement functionalities to several entities:

- **Secure Control Center (SCC) Dashboard (via GDPR & Privacy Service):** The SCC dashboard, built as a part of WP6, shows the reputation score of different entities. Such actionable information enables the analysis of various actors and allows carrying out appropriate response measures. The dashboard also serves as a configuration interface for the PPE component; data management, advanced consent options, permissions, rules and policies may be configured.
- **PHOENIX components:** PPE enables data exchange between PHOENIX components and also existing EPES infrastructure by providing a secure data exchange channel in addition to data exchange authorization. The components can register private or confidential data and may act as

either data controllers and/or data processors. They may also act on behalf of data subjects if such behavior is required.

- **Existing EPES infrastructure:** The role of existing EPES infrastructure is similar to the role of other PHOENIX components and SCC Dashboard. The rationale behind connecting EPES infrastructure directly to the PPE is two-fold; EPES infrastructure might already have a customer management system in place that is able to handle the required rules and policies related to data subjects with minor adaptations. Moreover, data exchange is sometimes required/desired also between the systems that may not be directly incorporated into the PHOENIX framework.

Since PPE is designed to be component agnostic, the above functionalities are extendable, and roles are interchangeable between components.

CoreDEX implements mechanisms for data exchange and data exchange authorization based on built-in compliance rules and governance policies in addition to PRIMULA's inputs: advanced consent and reputation score. CoreDEX provides data access control to SPC through data exchange authorization. PPE may be accessed through means of DLT node instantiation, interledger communication or via SPC. Data subjects interact with the PPE through the PHOENIX components or existing EPES infrastructure.

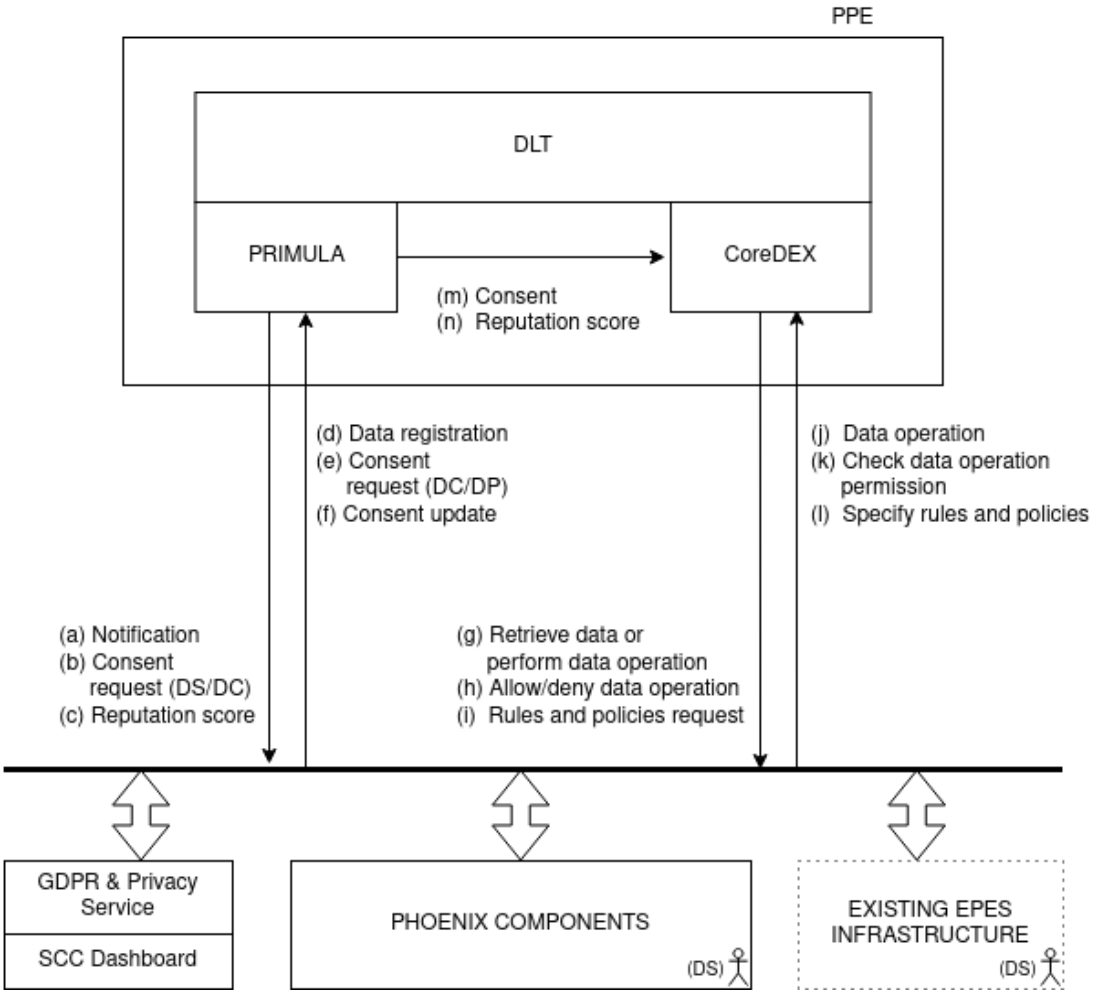

Figure 5: Interactions with PPE's inner components.

The actions are described in Table 5.

Table 5: PPE's inner components interactions.

Action	Description
(a) Notification	PRIMULA sends notifications regarding data processing etc.
(b) Consent request (DS/DC)	PRIMULA requests consent for data processing from DS/DC on behalf of requesting entities.
(c) Reputation score	Reputation score is communicated to entities.
(d) Data registration	Entities register personal or confidential data with PRIMULA.
(e) Consent request (DC/DP)	Entities (DC/DP) request consent for data processing.
(f) Consent update	DS/DC updates consent.

(g) Retrieve data or perform data operation	CoreDEX sends requested data or performs a data operation as a result of requested data operation (j).
(h) Allow/deny data operation	CoreDEX permits or denies a data request or data operation.
(i) Rules and policies request	CoreDEX requests a rules and policies operation from entities.
(j) Data operation	Entities request data or data operation.
(k) Check data operation permission	Entities check whether they have the rights to access data or perform an operation on data.
(l) Specify rules and policies	Entities specify rules and policies pertaining to their registered data.
(m) Consent	PRIMULA forwards the consent to CoreDEX.
(n) Reputation score	PRIMULA forwards the reputation score to CoreDEX.

4. PRIMULA Toolbox

This chapter provides an overview of the PRIMULA toolbox, the design decisions and the proposed technologies. PRIMULA enables consent-driven exchange of personal and privacy-sensitive data (also referred to as PRESS data) through managing the advanced consent among parties (data subjects, data controllers and data processors) via smart contracts (SCs, see Section 4.2.2). PRIMULA also provides consent auditing functionalities and reputation mechanisms to assess the party's reputation based on global perception of its behavior.

PRIMULA handles the exchange of the SCs for the following entities:

- data subject (DS) who is a consumer of the provided service,
- data controller (DC) who collects and manages the data, and
- data processor (DP) which consumes data in order to provide the desired service.

It is remarked here that, in order to provide the desired genericity of PRIMULA, a data subject role may be assigned to any data owners (i.e., entities that own a data collection). In this case, data owners normally exhibit also the role of data controllers. The concepts and relations pertaining to traditional GDPR roles may thus be adapted and valid also for the cases when no physical persons are involved.

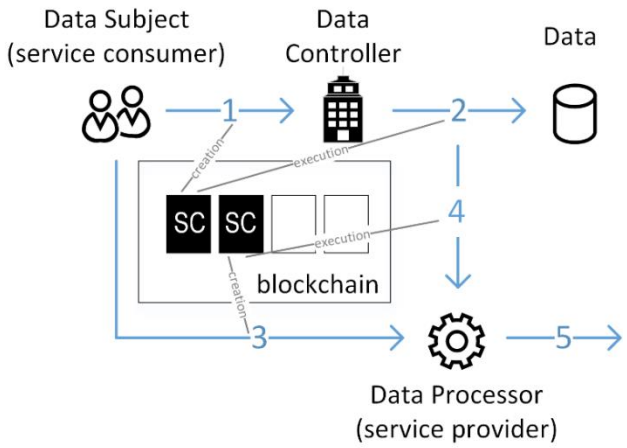

Figure 6: Exchange and utilization of SCs.

The exchange and utilization of SCs takes place as follows (see Figure 6):

1. DS and DC create a data usage SC. This SC governs the policies concerning the transferred and other common data.
2. DC has to comply with the SC during the collection and processing of DS's data.
3. DS and DP create a SC for governing policies for providing a service using DS's data collected by DC.
4. DC hands over DS's data to DP. The transaction is logged in the blockchain SC.
5. DP provides the service.

4.1. Design

PRIMULA is designed as a part and on top of a privacy-preserving blockchain for exchanging personal and confidential data. Built using a DLT, PRIMULA provides a set of ledgers and associated SCs for managing consents, assessing reputation of participating entities and logging any transactions related to the system. Additionally, a notification/request system, which serves as a messaging bus between various entities is provided. All entities in the system are defined by a pair of keys: private key (also secret key, sk) and public key (pk). The entities are identified using their pk. A registry of entities and their associated keys may be provided by the Accountability and Access Management (AAM) service, built as a part of WP6. In scenarios when DS does not directly interact with the DLT, the DS's keys are kept by a respective DC for interacting with the system on DS's behalf. In such cases, DS's keys should be encrypted using a secret key known only by the DS and decrypted when required.

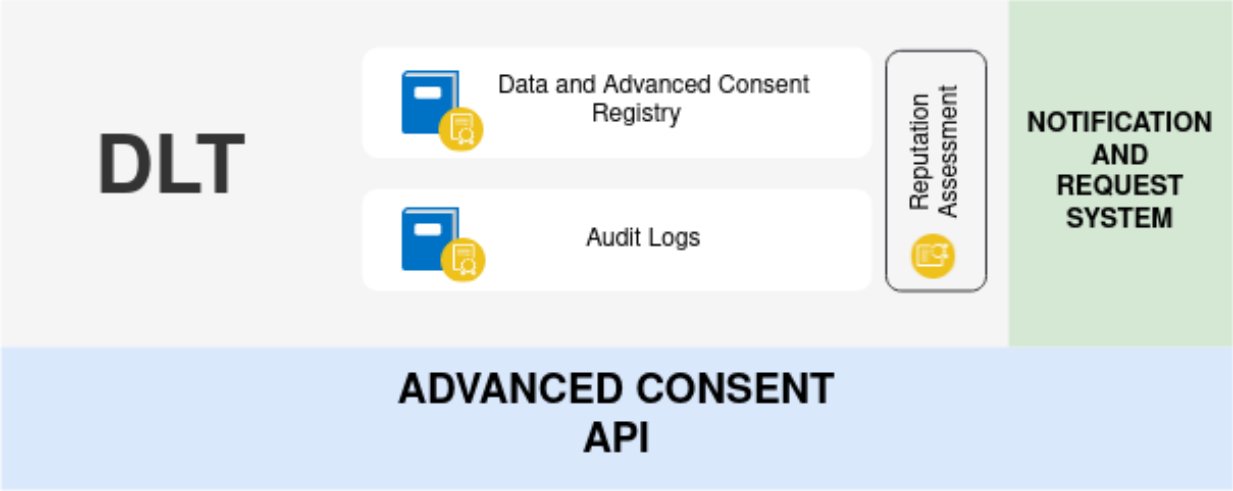

Figure 7: PRIMULA architecture.

PRIMULA's architecture is depicted in Figure 7. The architecture is composed of the following components:

- **DLT** is a blockchain on which PRIMULA is implemented.
  - **Data and Advanced Consent Registry** provides mechanisms for registering personal/confidential data with the blockchain and specification of consents.
  - **Audit Logs** ensures traceability by logging all interactions.
  - **Reputation Assessment** assesses the actors in the system and computes their reputation scores.
- **Notification and Request System** enables sending/retrieval of consent requests and various notifications (e.g., data breach notification).
- **Advanced Consent API** exposes PRIMULA's functionalities via defined application programming interfaces.

The components are further described in the following subsections.

4.1.1. Data and Advanced Consent Registry

Data and Advanced Consent Registry (DACR) enables secure registration of data and management via blockchain. Access management is provided through consent specification. DACR provides a distributed ledger and a corresponding SC that implements all actions.

Common ledgers come in the form of key-value pairs. The ledger keeps a full history of state transitions where a state denotes a snapshot of the ledger at any specific time. The state mutates according to transactions, which are used for creating, updating and deleting the key-value pairs. DACR's ledger keeps the following state:

```
{
      "key": {
            "ds": pk_DS,
            "dc": pk_DC
      },
      "value": {
            "timestamp": 1606672706,
            "identifier": "4899gu3lkduf93ufd...",
            "location": "v5909wjdjdfhepqss3...",
            "description": "b928fjaoqv84kfn39gjl...",
            "access": {pk_DS, pk_DC, pk_DP1, ...},
            "consent": {
                  pk_DP1: {"f93uflkdjf03f...", ...},
                  pk_DP2: {"f93uflkdjf03f...", ...}
            },
            "pk_d": "49jdlsuf93jfvmvvefi0..."
      }
}
```

The *identifier* field (a data hash or UUID) identifies the data whereas the *location* provides the data pointer (e.g., a connection string, URL, …). Access policy is specified using the *access* field, where the allowed public keys are listed. Human readable description of the data, used primarily for a better user experience in GUI application development, is provided in *description*. The consent for each DP is specified using the *consent* field. Consent content is DP-specific and is to be defined by each DP individually, but would normally take a form of a human-readable string or a machine-readable structure (i.e., XACML). Field pk_d denotes the public key used for encryption of data as well as fields. To protect confidential information, location, description and consents are encrypted using pk_d, which is a part of the key pair (pk_d, sk_d) that is shared between the entities using a secure channel.

Initially, a DS grants consent and access to DC (as a response to request), which in turn registers the data on DS's behalf by providing DS's signed consent. Direct data registration by a DS is possible. Upon data registration, a key pair (pk_d, sk_d) is generated, and a new record is appended to the ledger. DC is permitted to update the data location whereas the DS is allowed to change access rights and consent. Description and data identifier are immutable to ensure consistency. Changing any of those fields requires a new data registration.

Data registration can be carried out directly by a DS or, more commonly, indirectly by a DC. DS data registration is straightforward, a new record is simply appended to the ledger. The responsible DC can be initially set by a DS along with consents upon data registration. In a more realistic scenario, the data entry is added by a DC. In this case, a signed consent must be obtained from the DS first. In this scenario, the data registration takes place as follows:

1. The DC wants to start collecting some data in which a DS is implicated.
2. The DC requests initial consent for data collection/storing from the implicated DS via the Notification and request system or through other channels, e.g., existing infrastructure.
3. A consent object is formed by the DS and signed using DS's private key. Additionally, a set of keys for encrypting the data is created.
4. The signed consent and keys from step 3 are forwarded to the DC via the Notification and request system or through other channels (i.e., the existing infrastructure).
5. The DC requests data entry creation on BC by forwarding the signed consent.
6. The SC verifies the DS's signature and appends the record to the ledger.

Further updates to consent and other fields are carried out individually by each participating entity where the SC manages and authorizes the permitted updates of each entity. For instance, when a DS attempts to update the consent for a specific data entry, the SC checks whether the DS in question actually is the DS of that data entry. Normally, the updates are performed as a result of a request. Request mechanism is described in Section 4.1.3.

4.1.2. Audit Logs

DACR provides basic traceability in a tamper-resistant and transparent manner by design. However, the scope is limited to transactions associated with consent specification and data registration. To ensure complete transparency, mutual auditability and traceability, additional logging is required. For this reason, a second ledger is developed, which additionally documents all API calls and notifications/requests. This ledger records all interactions with the PPE and PRIMULA, namely:

- data registration and updating,
- consent checking, updating and retrieval,
- data access operations,
- request and notification submission and retrieval,
- system logs retrieval, and
- transactions associated with reputation.

Logs are appended to the BC as a tuple *{timestamp, pk, action}*, where *pk* is the public key (or part of the public key) of the actor and *action* object denotes the activity. The structure of the activity object will be

defined during the implementation step but should include all relevant information for comprehensively characterizing the individual interaction with the system.

4.1.3. Notification and Request System

Notifications are custom messages that arise from PRESS requirements. The aim of notifications is to provide comprehensive and timely information about various aspects, changes and irregularities in the system; in the context of GDPR, notifications are used by DCs/DPs to inform the DS regarding:

- data processing,
- data breaches, and
- other related events.

While these are the primary use cases, the usage of notifications is not limited to these scenarios. As notifications can be considered as a generic messaging platform, additional use cases – such as advertising of DP services – may be implemented using this system.

The second type of asynchronous interactions are requests, which mainly differ from notifications by their composition and response requirements. Whereas the notifications normally do not require a direct response and are intended as means of passing a piece of information, the requests, as indicated by the name, require a response. In PRIMULA, the requests provide a mechanism for service request and advanced consent specification throughout the complete system's lifecycle. In the context of PRIMULA, the requests are triggered by a DS to request a service from a DP or sent to a DC to initialize the data collection/handling. Conversely, a DP may request consent for data processing from a DS or, in an extended use case, a DP may request consent from a DC, should the DC already have received appropriate consent from a DS or no DS is implicated in the data processing.

Notification and request system is essentially a messaging system, an alternative dedicated persistent communication channel between entities that interact with PPE/PRIMULA. Internally, to ensure privacy and permit access only to the recipient, the content of the requests and notifications that are stored inside a secure persistent database, is encrypted using recipient's public key. A database entry is represented by a tuple $\{pk_{sender}, pk_{recipient}, content\}$. Message routing and delivery is based on $pk_{recipient}$. For verifying the authenticity and integrity of a message, the message is cryptographically signed by the sender. The message content, depending on the message type, may exhibit different internal structure. For instance, requests are normally bound to a specific set of data, therefore a data identifier specification is necessary. To satisfy the PRESS requirement of ensuring minimal necessary storage time, the notifications and requests are deleted from the system upon retrieval.

Due to the nature of the message flows and connected systems' modes of operation, both request/response and publish/subscribe approaches should be supported by the notification and request system. Frequently, a DS will not have direct access to PPE/PRIMULA and the related BC. In such cases, a DC can act as a mediator. Figure 8 shows a sequence diagram for a combination of publish/subscribe and request/response cases where the DC acts as a broker between the DS, the infrastructure and other entities. First, the DC subscribes to requests and notifications for respective DSs. When a DP publishes a notification/request, the messages are stored for DS's consumption at a later

time (e.g., when a DS logs in to a DC's application). Since all messages are encrypted, the DC does not have access to the message content.

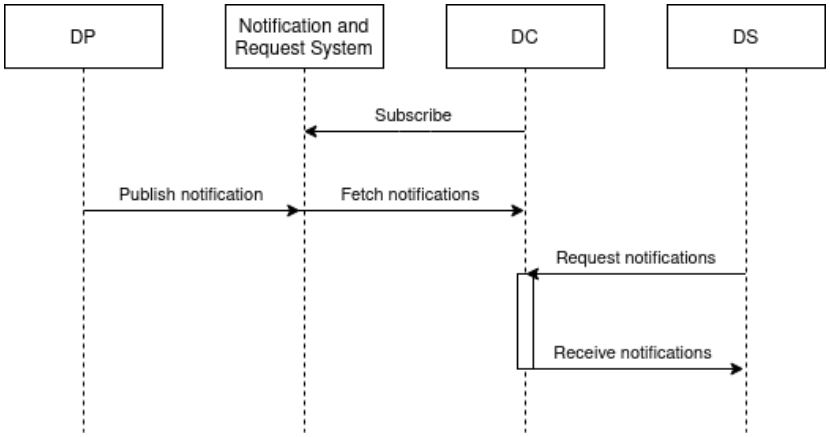

Figure 8: Publish/subscribe notification sequence.

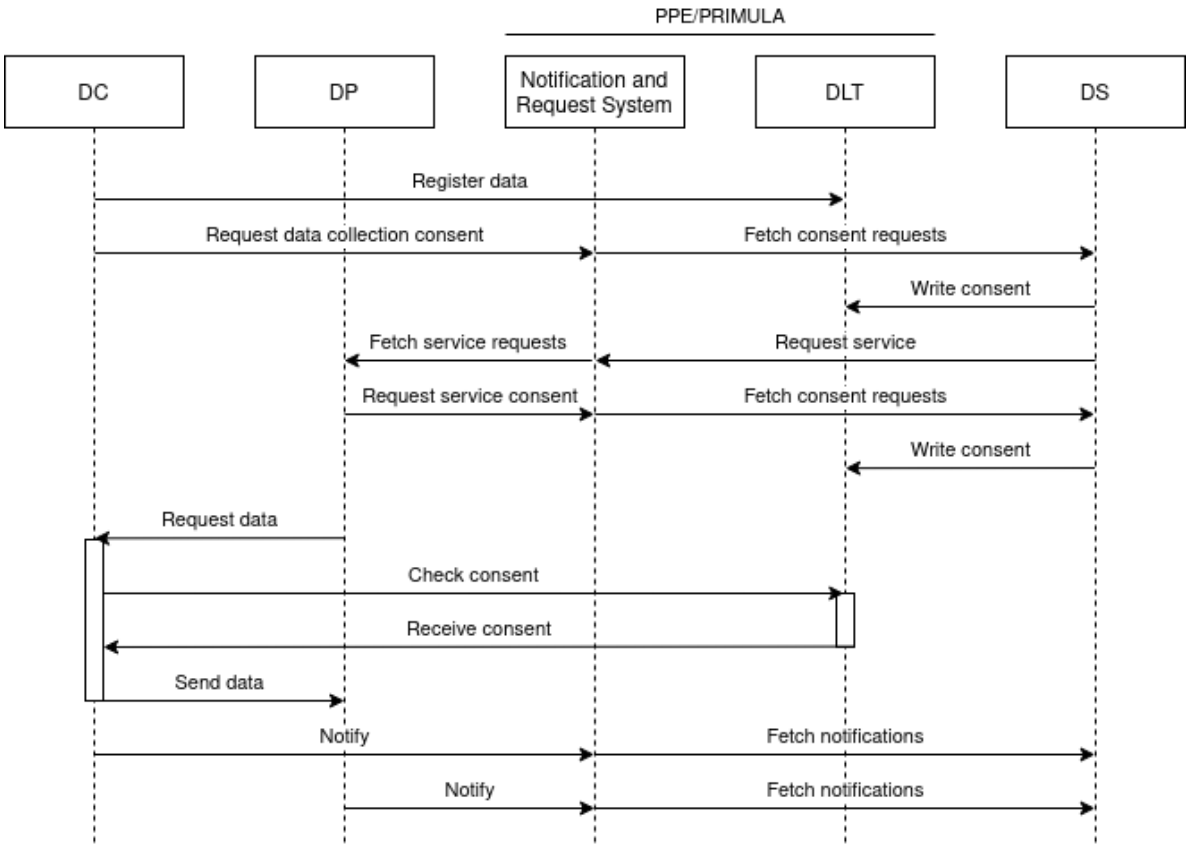

Figure 9: Data registration and exchange sequence.

Figure 9 depicts an illustrative sequence for a complete process of data registration, service request and consent exchange involving PPE/PRIMULA with focus on the notification and request system. Upon data registration by the DC, the initial consent for data collection is requested from the DS via the notification and request system. Then, the DS updates the consent for registered data on BC and requests a service from a DC. In turn, the DC requests consent for data processing that is part of the requested service. After the DS consents, the DP may request the data from the DC. The DC then checks the consent with the BC and consequently sends the data to the DP. Finally, the DS is notified about the data processing.

4.1.4. Reputation Assessment

Personal and confidential information is, due to its nature, prone to fraud. Since some entities participating in data transactions might be malevolent, a scoring mechanism for entities must be developed and the participation of dishonest parties should be prevented or appropriately limited. To enable such mechanisms, a reputation record that is calculated based on perceived behavior of all participating entities must be kept. This record should reflect the risks associated with each entity as well as their related trust level. In the context of PPE and DLT, two layers of reputation assessment are identified:

1. **BC level:** Reputation of each participating BC node/peer is assessed.
2. **PRIMULA level:** Reputation of actors in the PPE system (DS, DC, DP) is assessed.

Trust in the current ledger state among trustless nodes in a BC is primarily achieved using consensus algorithms and, in permissioned networks, also through authority. The dishonest and malicious peers may be identified by utilizing trust models and trust evaluation systems. Whereas the BC level reputation assessment focuses on BC infrastructure, in PRIMULA, the risk and reputation analysis targets consent/data operations.

PRIMULA reputation assessment module calculates the reputation score based on behavior monitoring of data controllers and data processors in the system to detect and evaluate possible risks and problems in personal/confidential data exchange. The reputation assessment module is implemented on a BC as a SC to enable decentralized computation and follows principles from log anomaly detection. It operates on data and transactions from the data and consent registry and audit logs, which also incorporates interactions associated with notifications and requests.

Reputation assessment consists of two steps:

1. Identification of anomalous behavior, and
2. scoring of actors based on anomalous behavior.

In relation to PPE/PRIMULA, the following anomalous behaviors are identified:

(a) Requesting data for which no consent was granted.
(b) Requesting consent for more data than usual.
(c) Requesting consent from more entities than usual.

For each of those three behaviors, a risk factor is derived as a ratio of anomalous events versus all events. The risk factors are denoted as r_a, r_b and r_c for behaviors (a), (b) and (c) respectively. The factor r_a is defined as

$$r_a = a_{total} \div (a_{total} + e_{distinct} - a_{distinct}),$$

where a_{total} is the total number of anomalous data requests where no consent was granted, $e_{distinct}$ is a number of distinct requested data sets and $a_{distinct}$ is a number of distinct requested data sets where no consent was granted (anomaly). The factor is normalized with respect to unique data requests to prevent scenarios where an entity would artificially try to lower the factor by repeatedly legitimately requesting data many times or requesting several different data sets.

For calculating r_b and r_c, two adjacent sliding windows are applied. The number of consent requests is summed up in each window and the windows are compared according to a threshold or a statistical test. The windows with a significant deviation are treated as anomalous. To obtain the risk factors, the anomalous windows and requests are added up and divided by the total number of windows/requests. The risk score R is then calculated as a weighted average $R = \alpha r_a + \beta r_b + \gamma r_c$, where $\alpha + \beta + \gamma = 1$. Finally, the reputation score S is computed as $S = 1 - R$.

Before the reputation of an entity can be calculated with sufficient confidence, ample information must be known about the entity. Essentially, the entity must perform a number of actions in the system. This phenomenon is also known as a cold start problem. For such cases, the use of global averages is proposed.

When a reputation score drops below a certain threshold, the implicated entities may be notified via the notification and request system, or automated response procedures may be executed (e.g., consent revocation). Individual anomalous behaviors in addition to the calculated reputation scores are sent to the SCC dashboard for further inspection.

The parameter analysis in addition to investigation of machine learning and stream processing approaches will be reported in subsequent deliverables (D4.3/D4.4).

4.1.5. Advanced Consent API

The interactions between multiple software systems are regularly defined using computer interfaces called application programming interfaces (API). The advanced consent API abstracts all PRIMULA BC operations, specifies the possible communication endpoints, thus exposing the PRIMULA functionalities. Due to the asynchronous nature of the PPE and PRIMULA, both request/response and publish/subscribe approaches are required for internal and external interactions. Complying with the REST architectural constraints, a preliminary PRIMULA RESTful API specification is provided in the following. An authenticated user is assumed.

Data management APIs enable data registering and retrieving/updating registered data entries:

[GET] /data	
This endpoint retrieves all data entries available to the requesting entity.	
Parameters	/

Returns	• a list of data entries
	  ○ data identifier
	  ○ data description

	○ data location

[POST] /data	
This endpoint inserts a new data entry.	
Parameters	• data description • data location
Returns	• data identifier

[GET] /data/{data identifier}	
This endpoint retrieves all information for a particular data entry.	
Parameters	• data identifier
Returns	• data identifier • data description • data location

[PUT] /data/{data identifier}	
This endpoint updates information for a particular data entry.	
Parameters	• data identifier • data description • data location
Returns	/

[DELETE] /data/{data identifier}	
This endpoint removes a data entry.	
Parameters	• data identifier
Returns	/

Consent APIs enable consent updating and retrieval:

[GET] /data/{data identifier}/consents	
This endpoint retrieves consents for a particular data entry.	
Parameters	• data identifier

Returns	• a list of consents

[PATCH] /data/{data identifier}/consents	
This endpoint updates the consent of the current authenticated entity for a particular data entry.	
Parameters	• consent
Returns	/

Notification/request APIs enable retrieving/sending notifications consent requests:

[GET] /notifications	
This endpoint retrieves the notifications sent to the current authenticated entity.	
Parameters	/

Returns	• a list of notifications
	  ○ sending entity
	  ○ data identifier
	  ○ notification content

[GET] /requests	
This endpoint retrieves the consent requests sent to the current authenticated entity.	
Parameters	/

Returns	• a list of notifications
	  ○ sending entity
	  ○ data identifier
	  ○ request content

[POST] /notifications	
This endpoint sends a notification to concerned entities for a data entry.	
Parameters	• recipient entities • notification content
Returns	/

[POST] /requests	
This endpoint sends a consent request to concerned entities for a data entry.	
Parameters	• recipient entities

	• request content
Returns	/

Reputation APIs enable retrieving reputation scores of participating entities:

[GET] /score	
This endpoint retrieves reputation scores for all entities.	
Parameters	/
Returns	• a list of entities and their reputation scores ○ entity ○ reputation score

[GET] /score/{entity}	
This endpoint retrieves a reputation score for a specific entity.	
Parameters	• entity
Returns	• entity • reputation score

Auditing API provides auditing information:

[GET] /data/{data identifier}/logs	
This endpoint retrieves logs for a particular data entry.	
Parameters	• data identifier
Returns	• a list of audit logs

4.2. Technology

This section provides an overview of the decided technologies for the implementation of the PRIMULA component as part of the PPE component.

4.2.1. ConsenSys Quorum

ConsenSys Quorum [19] is an open-source protocol layer that enables development of Ethereum-based blockchain solutions. It is optimized for business applications built on private and public Ethereum networks. Quorum comprises a suite of configurable components and APIs, thus facilitating a high level of use case and production environment customization.

GoQuorum [20], a Go-based fork of the Ethereum client focused on R&D within the Ethereum[4] community, enables development of enterprise applications requiring secure and high-performance transaction processing in a private network. GoQuorum is established and has been thoroughly tested in production permissioned networks with live applications. It implements an Ethereum-based protocol that runs private permissioned networks and provides Proof-of-Authority (Raft, IBFT, and Clique) consensus mechanisms. Compared to geth[5], an official Go implementation of the Ethereum protocol, GoQuorum incorporates the following modifications [21]:

- Proof-of-Work consensus algorithm is replaced with Proof-of-Authority.
- Connections in P2P layer are permitted only between permissioned nodes.
- Global state root check is replaced by global public state root check in block generation logic.
- In block header, the global state root is replaced with global public state root in block validation logic.
- A public state Patricia trie (also called prefix tree) and a private state Patricia trie is provided.
- Handling of private transactions is added to the block validation logic.
- Transaction data can be replaced by encrypted hashes on transaction creation.

As part of the ConsenSys Quorum ecosystem, a Java-based open-source private transaction manager Tessera for GoQuorum is provided. Tessera manages processes related to private data. In addition to generation and storage of private/public key pairs, Tessera enables node discovery and provides APIs between GoQuorum and Tessera nodes. Tessera implements the following two concepts:

- **Transaction manager** handles private transactions. It forms a P2P network with other transaction managers, manages peers and database access, distributes private payloads, and interfaces with Quorum as a gateway as well as the enclave for encrypting/decrypting private payloads.
- **Enclave** is a secure software processing environment for protecting information from attacks. It enables key management and encryption/decryption operations.

PRIMULA, as one of the core services of the PPE component, will be based on Quorum and will – together with CoreDEX – implement permissioned-network-like functionalities. From the deployment standpoint, this can be either a standalone setup or installed as a part of a wider infrastructure in cloud-based or on-premises setups. PRIMULA, built using Quorum blockchain layer, will adopt Tessera as a transaction manager and smart contracts from Quorum. Smart contracts will enable DACR and reputation assessment development whereas the Quorum blockchain ledgers will record DACR transactions and audit logs.

4.2.2. Smart Contracts

Smart contracts (SCs) are programs or protocols, that are automatically executed based on predetermined terms and conditions according to a contract or an agreement[6]. Removing a need for a

[4] https://ethereum.org/en/
[5] https://geth.ethereum.org/
[6] https://ethereum.org/en/developers/docs/smart-contracts/

trusted third party, SCs are self-executed in a trusted computer network such as BC. Ethereum promoted SCs as a form of general-purpose computation on a BC or distributed ledger and provides a Turing-complete language for implementing SCs on its BC. SCs may thus store arbitrary state and execute arbitrary computations on a BC. Deployment and execution of SCs occurs by a BC transaction. An overview of the SC technology is provided in [22].

By design of BC technology, SCs are immutable, however, a certain degree of mutability is normally desired, especially when applying bug fixes and improvements according to iterative releases. To address this, a concept of upgradable smart contracts (also called dynamic smart contracts) was developed [23].

SCs are implemented using various programming languages, such as Solidity[7]. Solidity enables development of SCs using an object-oriented, statically typed, high level language. It supports complex user-defined types, inheritance and inclusion of libraries. The implemented SCs govern the behavior of accounts within the Ethereum state and other Ethereum-based networks. Development, deployment and management of SCs for Ethereum-like BCs is facilitated by Remix IDE prototyping tool. Additionally, Truffle suite provides a development environment, testing framework and asset pipeline for Ethereum Virtual Machine (EVM) BCs.

For the implementation of the PRIMULA functionalities with regard to SCs, Remix IDE will be used in addition to Truffle suite development environment. SCs will be developed with Solidity of version 0.7.x or higher. The concept of upgradable smart contracts will be applied within the PPE component.

[7] https://docs.soliditylang.org/en/v0.7.4/

5. LSP Integration

This section outlines the specific requirements and use cases related to Large-Scale Pilots (LSPs). As a baseline, the LSP1 and LSP3 are considered, since they are dealing with personal as well as confidential data exchange among users and EPES stakeholders, thus foreseeing the adoption of PHOENIX components aimed at tight integration of technical processes (e.g., Universal Secure Gateway, USG), and anticipating the integration of PPE component in legacy products (i.e., smart meter data management and EV charging flexibility management). The initially captured specific requirements of LSP1 and LSP3 will be generalized and complemented with those of other LSPs upon the exact specification of data to be exchanged either for the purpose of other components operation (e.g., Situation Awareness, Perception and Comprehension, SAPC) or the exchange of specific CTI data. These specifics will be reported in D7.3: LSP set-up & test specifications. An overview of the LSPs is provided in D2.1 [24].

5.1. LSP Specific Requirements

The data to be exchanged among the EPES stakeholders constituting the LSP, can be classified either as personal data subject to GDPR and managed on the EU level, or confidential data associated with business operation or cybersecurity that is subject to individual stakeholder and managed on the national level. The CTI data per se are a subset of confidential data and subject to intra- and cross-sector coordination procedures.

For the narratives of data exchange requirements provided by the following subsections, it is important that the PRIMULA/PPE components fulfil the requirements of remaining agnostic to the kind of data (i.e., respecting the rules and policies that apply to certain data) as well as not interfere with the actual data exchange, but rather put in place the management of specific permissions and advanced consent.

5.2. Use Case 1 (LSP 1)

Regarding LSP1 (partners ASM, EMOT and BFP), practical implementation will focus on next assumptions:

- **ASM's** role is DSO (Distribution Service Operator, MV/LV grid owner).
- **EMOT** as charging station/EV fleet manager providing energy flexibility to ASM.
- **BFP's** role is renewable energy plants owner.

ASM Terni S.p.A. is a Public Company fully owned by the local municipality (City of Terni). The operation activities of ASM are very essential public services in the City of Terni as:

- Production and distribution of Electric Energy (DSO),
- Gas and Water distribution,
- Environmental Health.

In particular, LSP1 will focus on ASM's district consisting of the following block of energy units:

- Two PV arrays (180 kWp and 60 kWp), connected to the LV network.
- 72 kWh 2nd life Li-ion battery energy storage is the Block of Energy Unit (BoEU) providing the electric power storage and supply services. It is the BoEU that plays an important role in providing

the district with the flexibility necessary to implement different services, especially ancillary services like Primary reserve, Dynamic reactive Power control and Reactive Power Compensation.

- ASM Terni buildings comprising a 4,050 m2 three-story office building, a 2,790 m2 single-story building consisting of technical offices, a computer center and an operation control center and a 1,350 m2 warehouse; usually the base load varies between 50 kW and 90 kW and peak load is between 120 kW and 170 kW, depending on seasonal factors.
- Three smart charging stations and six electric vehicles managed by EMOT.

Use case 1 (UC1) analyzes a problem that arose after the advent of distributed generation from renewable energy plants. Before the integration of the renewable energy sources, the power flow was unidirectional: the power was generated in the large power plants from which the high voltage networks managed by the TSO (Transmission System Operator) departed, then the power was transformed into medium and low voltage in the networks managed by the DSO and finally reached the end users. After the integration of the renewable energy sources (RES), the power flow has become bi-directional. In the part of the electrical network where the end users are located (LV/MV grid) now there are production plants from renewable sources, such as photovoltaic or wind turbines, which generate intermittent energy and often their energy produced is not consumed instantaneously locally. As that part of the electricity network had been built with the idea that the end user was exclusively a consumer of energy and as therefore no energy storage facilities had been installed, the phenomenon of reverse power flow has arisen. Reverse power flow causes stability and safety problems, like voltage rise, frequency imbalance and fault equipment tripping; for this reason, different mechanism are being experimentally deployed, with the aim of avoiding this problem. The most advantageous mechanism for the DSO turns out to be the demand response (DR) mechanism, as it does not require the installation of devices for balancing the network, like storage systems, but it subsists in interacting with the end-users to ensure that they consume when the network has the flexibility need. A particular successful DR mechanism is the one in which the EV users are involved, since not only the capacity of the EV user is significantly greater than a home user but also because the flexibility can be provided both in the required time and in the required place. In UC1, therefore, the DSO (ASM) will identify the need for flexibility in the network, also taking advantage of the renewable production forecasts provided by BFP, and the Fleet Manager (EMOT) will make sure to provide the required flexibility stimulated by a discounted charging energy cost and the guarantee of charging with energy produced 100% from renewable sources.

In this context, the exchange of confidential data internal to stakeholders is a prerequisite, while particular attention should be devoted to respecting end-users' privacy and management of metadata exchanged between DSO, energy producer, fleet manager and EV users. This will be achieved through deployment of PRIMULA/PPE.

5.3. Use Case 2 (LSP 3)

In the context of LSP3 (partners BTC and Elektro Ljubljana), hands-on implementation will focus on next assumptions:

- **BTC** in the role of data service provider (legal entity as end user).
- **Elektro Ljubljana** in the role of utility, providing electricity to BTC.

There are other combinations possible, including data subjects (based on the GDPR terminology), which represents end users/consumers/data owners, though focus withing LSP3 will be on entities described above.

BTC as one of the biggest shopping centers in Slovenia with its 450+ shops, 175.000m2 shopping premises, 55.000m2 office spaces and 25.000m2 storage, runs a critical infrastructure, logistics, electrical infrastructure, demand/response, etc. It is crucial that all those entities and BTC as a whole run as smoothly and without interruptions as possible.

With all industrial systems in place, BTC runs various SCADA systems. Cyber Threats are something they are aware of. It is of most importance that systems run nonstop and provided information are relevant, tested and secure because of its sensitive nature.

The focus of privacy will target metadata exchanged between BTC and Elektro Ljubljana as the one "holding/storing" the data which is in fact owned by BTC – in this case smart meters owner.

Between BTC and Elektro Ljubljana, there is more than one layer of communication channels. Industrial one is equipped with SCADA systems, and is there for its main purpose: electrical distribution. Operational data from smart meters is using this layer, which is important from operational perspective and for billing needs. This data is stored by Elektro Ljubljana but is not owned by them.

Another layer of communication between BTC and Elektro Ljubljana is business oriented, where confidential business information is shared. This so-called metadata must be protected and therefore it is important that permissions on who can see/operate on this data are delegated by BTC.

Data owner needs to be in control of data. Data owner needs to have all the means to manage access, availability and restrictions on data. Solutions, which will provide management of permissions, its auditing, with highest level of trust into infrastructure beneath will base on DLT or more specifically on so called Permissioned Blockchain – Quorum, which will be implemented as a part of PRIMULA/PPE.

PRIMULA/PPE will have an important role in providing transparency in a way that the data owner will trust the actions above this layer to be temper proof. This layer will provide important functionalities as auditing to data owner, permissions delegations through smart contract functionality, and for potential data subjects also consent functionality. All these functionalities will be handled either by human/user faced GUI or APIs.

GUI will provide rich user experience for managing permissions, consent, in some cases even potential data and service registry. Based on the context of GUI, GUI will call appropriate smart contract on BC and finalize actions in audited way.

In our case, a 3rd party on electricity market will be interested in smart meter data (usage), metadata (confidential nature) and will need permissions from data owner (BTC). If data owner will be interested, they will be able to use the GUI and set permissions on specific data to be available to this 3rd party. In a way, BTC will consent that this 3rd party can use data owned by BTC and stored by Elektro Ljubljana for its further activities, which will be well defined before within the GUI/smart contract.

6. Conclusions

This deliverable provides an overview of the ongoing activities within WP4 together with the current development status. WP4 has developed the PRESS framework and aims to design as well as implement the PPE component and its constituent elements: PRIMULA and CoreDEX. This document's focus is on the former element.

This document described the main concepts related to privacy, reputation and mutual auditability. An extended overview of the GDPR targeting the different roles and their inclusion in PRIMULA was provided. The notion of advanced consent was presented. As confidential and personal data exhibit several constraints, the introduced concepts of PRESS data together with the PPE envelope provide means for better data analysis and system development for handling such data.

The PPE component was positioned in the scope of the PHOENIX project by defining interactions between components and specifying internal flows. An envisioned solution to PPE's requirements specified in D2.1 [24] was suggested. This deliverable further specified the PRIMULA toolbox. The different functionalities were presented through design decisions and envisioned technologies aiming to enable consent-driven exchange of personal and confidential data through managing the advanced consent among parties. A preliminary PRIMULA API was provided.

The LSPs as an integral part of showcasing the envisioned solution hold specific requirements. Such requirements were identified and consolidated considering the PRIMULA's and PPE's functionalities to form several use cases to be deployed in LSPs, namely LSP1 and LSP3.

The document better placed the PPE component and presented the first specification of PRIMULA. Further specification and implementation details will be reported in the subsequent deliverables (D4.3 and D4.4).

7. References

[1] E. Sartini, L. Briguglio, C. Occhipinti, A. Fiorentino and J. R. Martinez, “Deliverable D4.1: PRESS Framework Analysis,” H2020 PHOENIX Project, 2020.

[2] M. Pahlevan, L. Pasi, N. Pekka, W. Ben Jaballah, C. Piatte, N. Peiffer, V. Thouvenot, N. Wirtz, H. Flamme, C. Eze, C. Joglekar, Ö. Sen, J. Martinez, A. Palomares and D. Skias, “Deliverable D2.2: Secure and Persistent Communications Layer,” H2020 PHOENIX Project, 2020.

[3] GDPR.EU, “What are the GDPR consent requirements? - GDPR.eu,” [Online]. Available: https://gdpr.eu/gdpr-consent-requirements/. [Accessed 14 December 2020].

[4] G. C. Kessler, “Overview of Cryptography,” in *Handbook of Local Area Networks*, 1998.

[5] C. Dwork, “Differential privacy: A survey of results,” in *International conference on theory and applications of models of computation*, Springer, Berlin, Heidelberg, 2008.

[6] Y. Kortesniemi, D. Lagutin, T. Elo, N. Fotiou, D. Hill, J. Zhao and F. Luo, “Improving the privacy of IoT with decentralised identifiers (DIDs),” *Journal of Computer Networks and Communications 2019,* 2019.

[7] C. Allen, “The Path to Self-Sovereign Identity,” 25 April 2016. [Online]. Available: http://www.lifewithalacrity.com/2016/04/the-path-to-self-soverereign-identity.html. [Accessed 10 December 2020].

[8] F. Martin-Bariteau, “Blockchain and the European Union General Data Protection Regulation: The CNIL's Perspective,” *Blckchn. ca Working Paper Series 1,* 2018.

[9] G. Ateniese, B. Magri, D. Venturi and E. Andrade, “Redactable blockchain–or–rewriting history in bitcoin and friends,” in *2017 IEEE European Symposium on Security and Privacy (EuroS&P). IEEE*, 2017.

[10] M. Finck, “Blockchains and data protection in the European Union,” *European Data Protection Law Review,* vol. 4, no. 1, 2018.

[11] G. M. Riccio, A. Peduto, F. Iraci, L. Briguglio, E. Sartini, C. Occhipinti, I. Gutierrez and D. Natale, “The PoSeID-on Blockchain-based platform meets the ‘right to be forgotten’,” *Media Laws Journal,* vol. 2020, no. 2, 2020.

[12] H. Ferry, K. Bubendorfer and R. Chard, “Reputation systems: A survey and taxonomy,” *Journal of Parallel and Distributed Computing,* no. 75, pp. 184-197, 2015.

[13] E. Bellini, Y. Iraqi and E. Damiani, “Blockchain-based distributed trust and reputation management systems: a survey,” *IEEE Access,* vol. 8, pp. 21127-21151, 2020.

[14] W. Kong, Z. Y. Dong, J. Ma, D. Hill, J. Zhao and F. Luo, “An extensible approach for non-intrusive load disaggregation with smart meter data,” *IEEE Transactions on Smart Grid,* vol. 9, no. 4, pp. 3362-3372, 2016.

[15] I. Horrocks, “Ontologies and the semantic web,” *Communications of the ACM,* vol. 51, no. 12, pp. 58-67, 2008.

[16] M. Lorch, S. Proctor, R. Lepro, D. Kafura and S. Shah, “First experiences using XACML for access control in distributed systems,” in *Proceedings of the 2003 ACM workshop on XML security*, 2003.

[17] C. Choi, J. Choi and P. Kim , “Ontology-based access control model for security policy reasoning in cloud computing,” *The Journal of Supercomputing,* vol. 67, pp. 711-722, 2014.

[18] A. Masoumzadeh and J. Joshi, “Ontology-based access control for social network systems,” *International Journal of Information Privacy, Security and Integrity 1.1,* pp. 59-78, 2011.

[19] ConsenSys, “ConsenSys Quorum,” [Online]. Available: https://consensys.net/quorum/. [Accessed 14 December 2020].

[20] ConsenSys, “GoQuorum,” [Online]. Available: https://docs.goquorum.consensys.net/en/stable/. [Accessed 14 December 2020].

[21] ConsenSys, “Architecture - GoQuorum,” [Online]. Available: https://docs.goquorum.consensys.net/en/stable/Concepts/Architecture/. [Accessed 3 December 2020].

[22] T. Kerikmäe and A. Rull, The Future of Law and eTechnologies, Springer, 2016, pp. 133-147.

[23] OpenZeppelin, “Upgrading smart contracts,” [Online]. Available: https://docs.openzeppelin.com/learn/upgrading-smart-contracts. [Accessed 14 December 2020].

[24] W. Ben Jaballah, C. Piatte, N. Peiffer, K. Fotiadou, A. Voulkidis, H. Flamme, N. Wirtz, Ö. Sen, A. Lekidis, P. Nikander, M. Pahlevan, P. Lassila, D. Skias, L. Briguglio and Smolnika, “Deliverable D2.1: PHOENIX Platform Architecture Specification,” H2020 PHOENIX Project, 2020.