

D7.1 Guidelines and requirements, initial set of smart rules and related ontology - November 2022

GENTE

Distributed Governance for green ENergy communiTiEs



Source: ©GENTE, 2022

Impressum

Internal Reference

Deliverable No.	D 7.1 (2022)
Deliverable Name	Guidelines and requirements, initial set of smart rules and related ontology
Lead Participant	Prosume
Work Package No.	7
Task No. & Name	T7.1 Decentralised models for data and identity management, and privacy design strategies
Document (File)	GENTE-D7.1-Smart Rules and Related Ontology-C-R_A
Issue (Save) Date	15/02/2023

Document status

	Date	Person(s)	Organisation
Authors	2022-11-22	Miquel de la Mano Alex D'Elia	Prosume
Verification by	2022-12-17	Dogukan Ayci	REENGEN
Approval by	2023-02-15	Benjamin Bowler	HSLU

Document sensitivity

<input type="checkbox"/>	Not Sensitive	Contains only factual or background information; contains no new or additional analysis, recommendations or policy-relevant statements
<input type="checkbox"/>	Moderately Sensitive	Contains some analysis or interpretation of results; contains no recommendations or policy-relevant statements
x	Sensitive	Contains analysis or interpretation of results with policy-relevance and/or recommendations or policy-relevant statements
<input type="checkbox"/>	Highly Sensitive Confidential	Contains significant analysis or interpretation of results with major policy-relevance or implications, contains extensive recommendations or policy-relevant statements, and/or contain policy-prescriptive statements. This sensitivity requires SB decision.

Disclaimer

The content and views expressed in this material are those of the authors and do not necessarily reflect the views or opinion of the ERA-Net SES initiative. Any reference given does not necessarily imply the endorsement by ERA-Net SES.

About ERA-Net Smart Energy Systems

ERA-Net Smart Energy Systems (ERA-Net SES) is a transnational joint programming platform of 30 national and regional funding partners for initiating co-creation and promoting energy system innovation. The network of owners and managers of national and regional public funding programs along the innovation chain provides a sustainable and service oriented joint programming platform to finance projects in thematic areas like Smart Power Grids, Regional and Local Energy Systems, Heating and Cooling Networks, Digital Energy and Smart Services, etc.

Co-creating with partners that help to understand the needs of relevant stakeholders, we team up with intermediaries to provide an innovation ecosystem supporting consortia for research, innovation, technical development, piloting and demonstration activities. These co-operations pave the way towards implementation in real-life environments and market introduction.

Beyond that, ERA-Net SES provides a Knowledge Community, involving key demo projects and experts from all over Europe, to facilitate learning between projects and programs from the local level up to the European level.

www.eranet-smartenergysystems.eu



Date: 22 November 2022

Location: Barcelona

Co-financing:



Authors:

Miquel de la Mano, Prosume Energy,
miquel@prosume.io

The authors bear the entire responsibility for the content of this document and for the conclusions drawn therefrom.

Index

1. Introduction
2. Guidelines and requirements
 - a. Distributed ledger
 - b. Modularity and interoperability
 - c. Determinism
 - d. Decentralization of trust and federation
 - e. Privacy-by-design
 - f. Zero Knowledge Proof
3. Smart Rules for personal data sharing
 - a. Tools and information management
 - b. Different kind of datasets
 - c. The GDPR software compliance
6. Privacy-by-design system architecture
 - a. Zencode and Zenroom
 - b. The wallet
7. GENTE related ontology: exploring the Decentralized Identifiers (DIDs) framework
 - a. Overview
 - b. DID syntax and decentralized network adaptation
 - c. DID vs. Global-unique Identifiers
 - d. DID Methods
8. Order processing: initial approach

1. Introduction

The GENTE project is researching, developing and testing free and open source software as the building block for distributed, privacy aware and trusted technology architectures, for decentralized data governance as well future identity management systems. Identity on the internet has evolved from being implemented as centralized silos to federated identity models. Federated identity enables single sign-on available across several large service provider platforms. Service providers continue to be data controllers, in both centralized and federated models.

At a high level we can describe GENTE privacy and identity management toolbox as being composed of the following:

- Set of specifications for distributed ledgers to support GENTE
- Free and open source reference implementation of a distributed ledger
- Human readable language for data transformation (Zencode)
- Lightweight, portable and secure virtual-machine to execute Zencode (Zenroom)
- Documentation needed for operators to write and deploy smart rules that request access to private data.

The GENTE toolbox is composed of both hardware and software components: each component adheres to the core architectural principles described below. The underlying philosophy adopted is that of the UNIX philosophy following key principles of modularity, clarity, composition, separation, simplicity, parsimony, transparency, robustness, representation and least surprise. The openness of the platform enables innovation and citizen participation.

The idea is to build a modularized system composed of:

- Software binaries to install on the meter devices, sub-meter, IoT, routers, etc. to provide flexibility and integration on the infrastructure layers
- Softwares modules like Zenroom to build the business logics in human understandable self programmable language, and to maintain interoperability
- Self Sovereign Identity and privacy modules, to build awareness and reduce the risk of algorithmic black holes

Energy Infrastructure is following the same path that telecommunication infrastructure has taken in the last 20 years, with the main difference that telecommunication became mobile while energy provisioning is fixed, but end users are in search of a flexible and portable management of their Energy Service Level.

To achieve these transformation objectives we have to align front/middle/back-office operations, separate the Identity and Privacy layers from the Smart Contract modules while providing strong cryptographic procedures and interoperability, enabling to improve customer experience and margin from the bottom-line profit. This process would establish a framework for every industry partner and participant making it easy to contract distributed resources, exactly the same way the wholesale markets work today.

Social digital innovation is led by new paradigms of immaterial value transfer, but when dealing with energy transition scenarios it is clear that we still live in a reality of physical authenticated systems. Parties have to share private information while participating in a public service and/or public infrastructure. Innovative research in cryptography has converged in our development and has helped to solve problems and envision even more opportunities for developments ahead.

Interoperability is a key factor, not only to avoid specific lock-in effects commonly caused by the oligopolies of Service Providers, but also to re-use legacy infrastructure. It is fundamental to support existing hardware which does not always provide a Linux Operating System, and therefore develop sensor components that can be retro-fitted to legacy infrastructure.

2. Guidelines and requirements

a. Distributed ledger

A distributed ledger with decentralized governance provides a public, resilient, tamper-resistant and censorship resistant record which allows any party to be able to verify some “fact” recorded within it. This verification is demonstrable through the use of cryptography.

b. Modularity and interoperability

GENTE develops modular privacy-aware tools and libraries that integrate with the operating system backed by a state of the art blockchain infrastructure supporting smart contracts and privacy protections. GENTE adopts a layered architecture, with components that build on top of each other. As opposed to building privacy aware applications solely in the application layer (layer 7) of the Operating System, privacy is built into the lower layers as well, such as transport, network and data-link layers.

c. Determinism

In a properly designed cryptographic flow, even for a simple encryption/decryption, each transformation needs to occur “end to end” (where different ends can be very different devices) while being deterministic and lead to the same results, otherwise the flow will break. The severity issue is directly proportional to the number of different software platforms and components present in the cryptographic flow: the implementation of algorithms can differ between multiple applications and libraries, or even executing the same library function on a different OS can produce a non-deterministic output, making the data unusable. The development of a virtual machine (VM) in GENTE aims to provide a platform-independent programming environment that abstracts away details of the underlying hardware or operating system and allows a program to execute in the same way on any platform.

d. Decentralization of trust and federation

The current era in technology has seen a shift from large monolithic systems to distributed energy systems. This is to meet the requirements of systems, scaling, resilience and fault tolerance but also provides for decentralized governance models. GENTE builds upon

decentralized models for data and identity management. This means that as much as possible, the control over the system should not be in the hands of a small number of entities over whom the participants of the system have no influence or recourse.

e. Privacy-by-design

GENTE aims to develop a privacy preserving data distribution platform to foster energy-based sharing economy models, where citizens own and control their data. This asks for a privacy by design based approach, for which the concept of privacy design strategies have recently been developed.

f. Zero Knowledge Proof

The public nature of the ledger is in tension with a desire to maintain the privacy of the participants of the network. GENTE applies the concept of ZKPs to allow the cryptographic proof of a transaction to be recorded in the ledger without needing to publicly record the data within the transaction itself.

3. Smart Rules for personal data sharing

Despite the fact that the first good practice when dealing with personal data and DLTs is not to store personal data on the DLT (at least in permissionless DLT, until technology development, interpretation of the GDPR and standard legal solutions are not definitely able to allow legal compliance and rights safeguard), GENTE offers, nonetheless, the possibility to share personal data exploiting the architecture of DLTs, cryptography, and privacy design strategies combined, without any storage of personal data on the distributed ledger itself.

It is certainly possible to build digital data spaces made of personal data adopting free licenses: a good example of this is Wikipedia, that is formed by contributions made by many persons (tracked by their name, email and/or IP address) and thus formed with voluntarily contributed personal data.

Taking advantage of the DLT, particularly SR, there are new ways to foster data sharing: SRs allow to make declarations and commitments, to conclude agreements, and to perform specific actions in order to comply with defined obligations obligations

Such declarations, commitments, and undertakings do not need to be included within the licenses text: on the one hand, some free licenses do not accept any additional or different terms or conditions, on the other hand, free licenses normally do not deal with, and therefore do not overlap with, privacy rights. In short, it is neither useful nor efficient to design new licenses to foster the building of digital data commons including personal data.

It is instead a convenient way to be law abiding the adoption of SRs (providing information, receiving consent, making it easier for the data subject to exercise her rights, etc.), and the adoption of declarations and commitments that could strengthen the power of the data. For these reasons we propose in Annex A an initial GDPR compliance Terms and Conditions.

a. Tools and information management

The Linux Foundation supports the SPDX standard¹ that provides a common format for information about free software licenses and copyrights (SPDX Tools that provide translation, comparison, and verification functionality are also available). The Linux Foundation also supports the OpenChain Project², a project that, among others, provides the OpenChain Specification, a set of requirements for compliance programs. FOSSology³ is a free software license compliance software system and toolkit that allows to run license and copyright scans.

Information about free software licenses is easily accessible from different sources and good points to start with are:

- the GNU project website that lists licenses that comply with the free software definition, provides FAQ about the GNU licenses⁴ and other useful information;
- the Open Source Initiative website that lists licenses⁵ that comply with the Open Source Definition and provides other information;
- the Wikipedia website that provides information about most of the free software licenses (e.g., https://en.wikipedia.org/wiki/Apache_License) including comparison of free and open source software licenses⁶;
- the Choose a License⁷ website, the tldrLegal⁸ website and the Joinup Licensing Assistant (JLA)⁹ website that provide information about some of the most well known free licenses and the obligations to be complied with according to each of them.

For datasets, it is worth mentioning the licensing assistant made available by the European Data Portal¹⁰.

b. Different kind of datasets

Different kinds of datasets are processed throughout the GENTE Project.

Firstly, GENTE processes:

1. Data necessary to the running of the DLT services; this is the case of energy or technical data that is independent of data related to each specific service (or pilot) based on GENTE and data provided by users. Differently from the former, those latter are discretionary with respect to the functioning of the GENTE DLT.

Secondly, each GENTE pilot (or service that will be based on GENTE technology) manages datasets from different sources:

¹ <https://www.spdx.org/>

² <https://www.openchainproject.org/>

³ <https://www.fossology.org/>

⁴ <https://www.gnu.org/licenses/>

⁵ <https://opensource.org/licenses>

⁶ https://en.wikipedia.org/wiki/Comparison_of_free_and_opensource_software_licenses

⁷ <https://choosealicense.com/>

⁸ <https://tldrlegal.com/>

⁹ <https://joinup.ec.europa.eu/collection/eupl/joinuplicensingassistantjla>

¹⁰ <https://www.europeandataportal.eu/en/content/showlicense>

2. Datasets shared autonomously by users;
3. Datasets collected and released by public sector administrations or companies (we are referring more precisely to public sector bodies and public undertakings), that are freely and openly available to the public;
4. Datasets (belonging to users and/or public sector administrations) gathered automatically by IoT devices connected with the GENTE platform.

Thirdly, each GENTE pilot (or service) will produce new datasets:

5. Datasets generated by analytics based on other datasets, or datasets derived from a combination or a selection of datasets.

c. The GDPR software compliance

The General Data Protection Regulation (GDPR), as well as other data protection or privacy protection laws and regulations, define data protection in legal terms. These terms are soft, open to interpretation, and highly dependent on context. Because of this inherent vagueness, engineers find such legal requirements hard to understand and interpret. The GDPR also mandates privacy by design, without describing clearly what this means exactly, let alone giving concrete guidelines on how to go about implementing privacy by design when actually designing a system. Intuitively, privacy design means addressing privacy concerns throughout the system development life cycle, from the conception of a system, through its design and implementation, proceeding through its deployment all the way to the decommissioning of the system many years later. In terms of software engineering, privacy is a quality attribute, like security, or performance. To make privacy by design concrete, the soft legal norms need to be translated into more concrete design requirements that engineers understand. This is achieved using privacy design strategies.

Software can however enable or hinder an organization in achieving GDPR compliance. GENTE provides transparency for participants about exactly where their data is and with whom it has been shared which also enables GDPR compliance. Further, many of the privacy by design principles correlates with needs of GDPR compliance, for example right to be forgotten.

6. Privacy-by-design system architecture

a. Zencode and Zenroom

Zencode is a software project inspired by the discourse on data commons and technological sovereignty, consolidated in the Zencode Whitepaper¹¹ as a living document that will continue to be updated beyond the span of the GENTE project. The established goal is that of improving people's awareness of how their data is processed by algorithms, as well as facilitating the work of developers to create applications that follow privacy by design principles. The main use case taken in consideration is that of distributed computing capable of processing untrusted code and executing advanced cryptographic functions, for instance it can be used with any distributed ledger (blockchain) implementation as an interpreter of smart contracts.

¹¹ https://files.dyne.org/zenroom/Zenroom_Whitepaper.pdf

The Zencode domain specific language (DSL) makes it easy and less error-prone to write portable scripts implementing end-to-end encryption with operations executed inside an Zenroom isolated environment (the VM) that can be easily ported to any platform, embedded in any language and made inter-operable with any blockchain. Its interpreter, the Zenroom VM¹², supports secure isolation and protects its hosts from errors; it has no access to the calling process, the network, underlying operating system or filesystem. Zenroom VM is a process virtual machine: a restricted execution environment designed to process safely any Zencode instruction. Upon any failure during phases of interpretation of code, validation of data or execution of operations Zenroom aborts returning meaningful error messages that help programmers assess what problem had occurred. Zencode language scenarios are written following a declarative approach and provide functional tools to manipulate efficiently even complex data structures.

b. The wallet

The wallet is the minimum component a person requires to interact with the GENTE toolbox. Every participant has their own wallet. The wallet has several core functions:

- Store securely cryptographic material (e.g. private keys)
- Securely store Attribute based credentials, linked to private keys
- Execute GENTE transactions (via Smart Rules) and submit them to the Ledger for verification
- Store, encrypted the participant's attributes
- Provide the participant with a graphical user interface that allows them to manage their attributes, entitlements and applications.

7. GENTE related ontology: exploring the Decentralized Identifiers (DIDs) framework

Decentralized Identifiers (DIDs) v1.0 are a proposed standard by the World Wide Web Consortium (W3C) for creating and managing unique identifiers that are independent of any centralized authority. DIDs enable individuals, organizations, and devices to create and own their own identifiers, which can be used to verify their digital identity, establish trust, and enable secure data exchange without relying on third-party intermediaries. DIDs are designed to be flexible, privacy-preserving, and interoperable across different systems and networks. The proposed standard includes a data model, a syntax, and a set of security and privacy considerations.

a. Overview

The origin of Decentralized Identifiers (DIDs) can be traced back to the need for a more secure, flexible, and privacy-preserving way of creating and managing digital identities on the web. Traditional identity systems rely on centralized authorities to issue and verify identities, which can lead to a range of issues, including vendor lock-in, identity theft, and censorship.

¹² <https://zenroom.dyne.org/>

To address these issues, a group of experts from various fields came together to explore the idea of decentralized identities based on distributed ledger technology. This led to the creation of the Decentralized Identity Foundation (DIF) in 2017, which aimed to promote the development of open standards for decentralized identity systems.

One of the key outputs of the DIF was the Decentralized Identifiers (DIDs) specification, which was proposed to the World Wide Web Consortium (W3C) in 2018. The aim of the specification was to create a standardized way of creating and managing DIDs that would be interoperable across different systems and networks.

The DIDs v1.0 specification went through a rigorous process of review, feedback, and refinement by a wide range of stakeholders, including experts from the W3C, DIF, and other organizations. The final specification was published as a W3C recommendation in November 2020, providing a solid foundation for the development of decentralized identity systems that can help to address the challenges of the traditional identity landscape.

At a superficial level, a DID is simply a new type of globally unique identifier. But at a deeper level, DIDs are the core component of an entirely new layer of decentralized digital identity and Public Key Infrastructure (PKI) for the Internet. This *decentralized public key infrastructure* (DPKI)¹³ could have as much impact on global cybersecurity and cyberprivacy as the development of the *SSL/TLS protocol*¹⁴ for encrypted Web traffic (now the largest PKI in the world).

This description is designed to give newcomers to DID architecture the background they need to understand not just the DID specification, but the overall architecture for decentralized identity represented by the family of DID-related specifications currently under development. It covers:

- Background on the origin of DIDs and the DID specification.
- How DIDs differ from other globally-unique identifiers.
- How the syntax of DIDs can be adapted to work with decentralized networks.
- How DIDs resolve to DID Documents containing public keys and service endpoints.
- The key role that DID Methods play in the implementation of DID infrastructure.
- Privacy considerations for the use of DIDs.
- How DID infrastructure lays the foundation for verifiable credentials.

b. DID syntax and decentralized network adaptation

The syntax of Decentralized Identifiers (DIDs) can be adapted to work with decentralized networks by leveraging existing standards and protocols for decentralized systems. The basic syntax of a DID is "did:method:specific-id", where "method" is a string that identifies the method used to create and manage the DID, and "specific-id" is a string that uniquely identifies the DID within that method.

¹³

<https://github.com/WebOfTrustInfo/rebooting-the-web-of-trust/blob/master/final-documents/dpki.pdf>

¹⁴ https://en.wikipedia.org/wiki/Transport_Layer_Security

To work with decentralized networks, different methods can be used to create and manage DIDs. For example, a method could use a public blockchain to store and manage the DIDs, with each DID represented as a unique address on the blockchain. Alternatively, a method could use a distributed hash table (DHT) to store and manage the DIDs, with each DID represented as a unique key in the DHT.

To support these different methods, the syntax of DIDs can be extended to include additional parameters that are specific to each method. For example, a method that uses a public blockchain might include a parameter that specifies the blockchain network and address format to use, while a method that uses a DHT might include a parameter that specifies the DHT implementation and configuration.

These extensions can be defined and managed through the DID method specification, which provides a standardized way for creating and managing DIDs using different methods. This enables interoperability across different decentralized networks and systems, while still allowing for the flexibility and customization needed to meet specific use cases and requirements.

c. DID vs. Global-unique Identifiers

Decentralized Identifiers (DIDs) differ from other globally-unique identifiers in several key ways:

1. **Decentralization:** DIDs are designed to be independent of any centralized authority, unlike traditional identifiers such as domain names or social security numbers, which are typically issued and controlled by centralized organizations or governments.
2. **Self-Sovereignty:** DIDs enable individuals, organizations, and devices to create and own their own identifiers, rather than relying on third-party intermediaries to manage their identity.
3. **Security:** DIDs provide a high degree of security through the use of public-key cryptography, which enables users to prove ownership of their identifier without revealing any sensitive information.
4. **Flexibility:** DIDs are designed to be flexible and extensible, allowing for the creation of new identifier types that can be tailored to specific use cases and requirements.
5. **Interoperability:** DIDs are designed to be interoperable across different systems and networks, enabling users to use the same identifier across multiple platforms and services.
6. **Privacy:** DIDs provide a high degree of privacy by allowing users to control the information that is associated with their identifier, and by enabling selective disclosure of information to different parties as needed.

These characteristics make DIDs a powerful tool for enabling secure, flexible, and privacy-preserving digital identity systems that can help to address the challenges of the traditional identity landscape.

d. DID Methods

There are currently many Decentralized Identifier (DID) methods that have been proposed and developed, each with its own unique set of rules and procedures for creating and managing DIDs. Some of the most widely used and well-known DID methods include:

1. **DID Key:** This is a simple DID method that is based on public key cryptography. DIDs are derived from a public key, and the associated DID document contains the public key that can be used to verify digital signatures.
2. **DID Web:** This method is designed to work with the traditional web infrastructure and uses domain names to create and manage DIDs. The DID document contains information about the domain name and other web-related information.
3. **DID BCR:** This method uses the Bitcoin blockchain to create and manage DIDs. The DID document is stored on the blockchain, and each DID is associated with a Bitcoin private key.
4. **DID Sovrin:** This method is designed for use in self-sovereign identity systems and uses a public permissioned blockchain to create and manage DIDs. The DID document contains information about the DID controller and verification methods.
5. **DID Sidetree:** This method uses a decentralized hash table (DHT) to create and manage DIDs. The DID document is stored in the DHT, and each DID is associated with a unique public key.

Each DID method has its own unique set of features and characteristics, but they all share the common goal of providing a standardized way to create and manage DIDs in a decentralized and interoperable way. The choice of DID method will depend on the specific use case and requirements of the application or system that is being developed.

8. Order processing: initial approach

By adopting the GENTE toolbox, participants are offered the freedom to choose how their energy is managed. Participants can also benefit from flexibility management of energy production and accumulation, for instance by deciding to store excess energy and place it on the market for additional revenues according to custom conditions defined by smart contracts.

We have to build an online distributed marketplace where “Smart Contracts” manage offers and bids, automatically matched by the platform based on participant’s preferences. Members of a certain area have complete transparency to monitor with whom, when, and at what price their energy was traded: transparency is contextual to local environments as defined by “Community Bindings” (PRESETS contracts).

Following standard Order Management Systems we provide participants with the possibility to establish market and limit orders, further recording booked orders and exchanges as immutable records in the Distributed Ledger. To establish a sort of “digital single market” the execution of Smart Contracts takes place in the distributed system based on given specific attributes:

- **UID:** type of asset and name, status and trade date/time
- **BUY/SELL:** value, price, quantity

- **Seller:** Sales, Shipper, Provider
- **Counterparty:** UID/name/address/date/payment/settlement/delivery_type (as for the Seller it requires specific Attributes)

Given the specific Attributes, it is important to have the Order Booking logic separated from the DLT system and the Accounting systems who might inherit attributes from other systems.

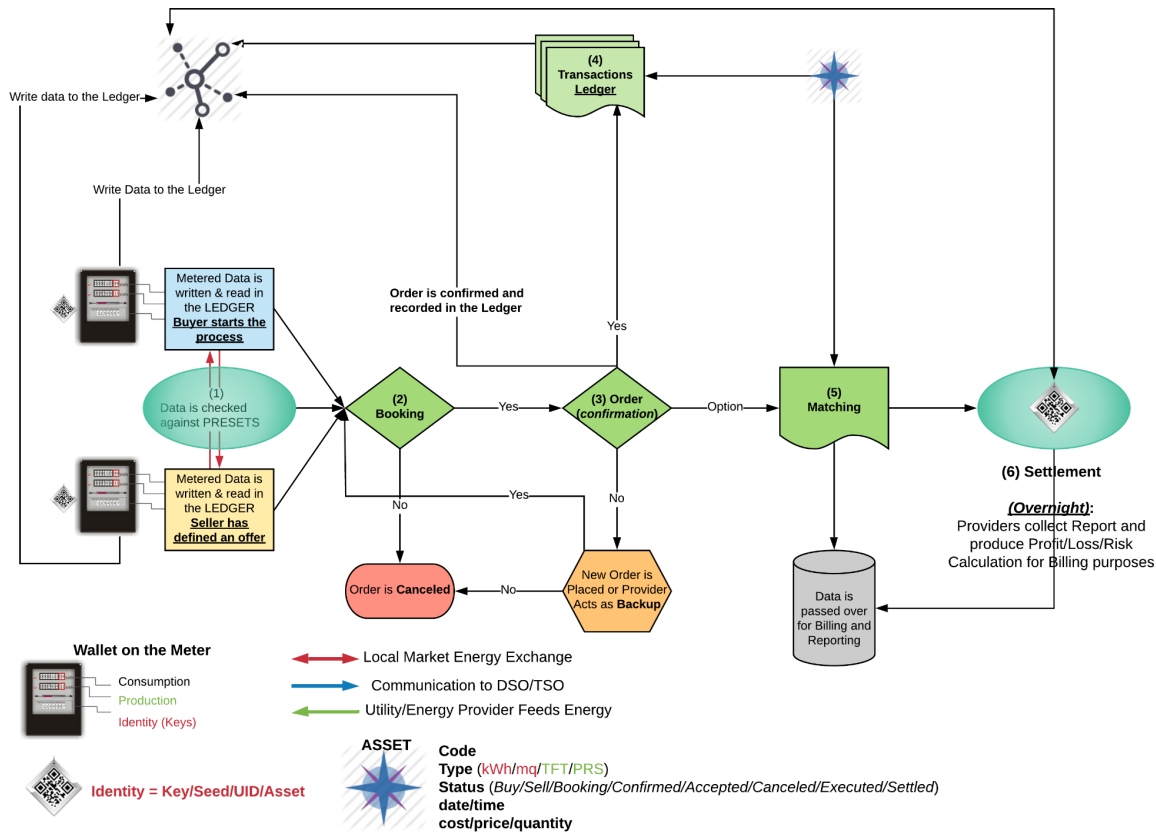


Figure 3: Order processing following the 6 steps of a settlement life-cycle.

1. Given that the Seller or Buyer have Proper ID, Credit (enough Asset), and a Specific Service Level, there can be an offer to buy or sell. This is the Pre-Execution step where data are checked against Attributes and metered data values which are registered in the Ledger by metering devices
2. Booking: Both parties agree on the characteristics and assets. Only then there can be an order between them
3. Order (Confirmation): Once agreed, there must be an Order Confirmation and at this step the order becomes a transaction and is written into the Ledger
4. Transaction: happening on the Ledger. Orders are matched and values are exchanged
5. Matching: after the transaction has happened, given that specific exchanges happened and values transacted, those values are registered as confirmed and settled for the Billing procedure. In case something goes wrong they have to be canceled
6. Settlement: at this step all the orders are registered on the Ledger and on the systems managing operations and providing then the reports generated for Overnight profit/loss calculation, risk and billing procedures and any different type of document might be requested by the accountant offices

This type of development would help us provide a flexible solution to relevant stakeholders' co-design enabling a greater autonomy in building new sets of business logic smart-contracts with high societal impact.

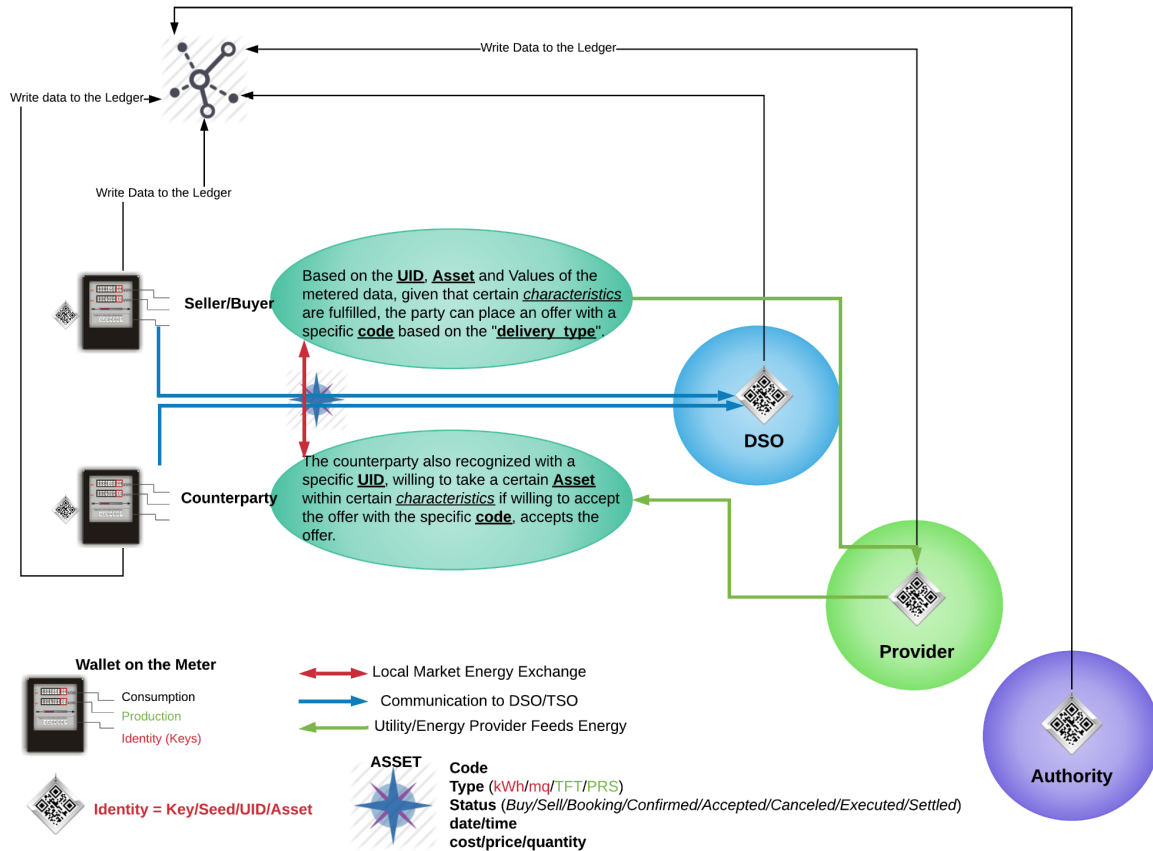


Figure 4: Access map of stakeholders for booking orders and trading operations.

IoT is especially relevant to the Smart Grid since it provides systems to gather and act on energy and power-related information in an automated fashion with the goal to improve the efficiency, reliability, economics, and sustainability of the production and distribution of electricity.

