



decode



**Final version of the DECODE
architecture, documentation and
sustainability**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 732546



Project no. 732546

DECODE

DEcentralised Citizens Owned Data Ecosystem

D1.11 Final version of DECODE architecture, documentation and sustainability

Version Number: V1.0

Lead beneficiary: University College of London (UCL)

Due Date: 31 October 2019

Author(s): Ioannis Psaras, Sergi Rene (UCL)

Editors and reviewers: Oleguer Sagarra (Dribia), Denis Roio, Francesca Bria (Dyne), Jim Barrit (ThoughtWorks), Sam Mulube (Thingful).

Dissemination level:		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Approved by: Francesca Bria, DECODE Project Coordinator
Date: 15/12/2019

This report is currently awaiting approval from the EC and cannot be not considered to be a final version.

Contents

Figures.....	4
Tables	5
Abbreviations	6
1. Introduction.....	7
2. Goals	10
3. Architecture Principles	11
3.1. Free and open source	11
3.2. Modularity and interoperability	11
3.3. Reuse don't re-invent	11
3.4. Decentralisation and federation.....	12
3.5. Privacy by design	12
3.6. User friendliness	12
3.7. Determinism.....	13
4. Concept Foundations.....	14
5. DECODE Architecture	16
5.1. Zencode and Zenroom.....	17
5.2. Decode Ledger	18
Chainspace	18
Sawtooth.....	19
5.3. Coconut.....	20
5.4. Decode P2P Network.....	22
5.5. Decode OS.....	22
5.6. DECODE App	23
6. DECODE Pilots.....	24
6.1. Barcelona Pilots	24
Citizen Science and IoT Data Governance	25
Digital Democracy and Data Commons	28
6.2. Amsterdam Pilots.....	30
GebiedOnline (Neighbourhood online)	31
Claim verification (18+).....	33
7. Conclusion	35
Bibliography	36

Figures

Figure 1 - Digital Commons	- 6 -
Figure 2 – The value of DECODE	- 9 -
Figure 3 - DECODE Architecture	- 15 -
Figure 4 - What goes in the ledger?	- 17 -
Figure 5 - DECODE ledger	- 19 -
Figure 6 - Coconut architecture	- 20 -
Figure 7 - Schematic view of Barcelona pilot framework	- 26 -
Figure 8 – IoT Pilot framework	- 28 -
Figure 9 - Integration plan for the Digital Democracy and Data Commons pilot	- 31 -
Figure 10 - Amsterdam Pilots	- 32 -

Tables

Table 1 - DECODE elements

- 15 -

Abbreviations

ABC - Attribute Based Credentials

BFT - Byzantine Fault Tolerance

DECODE - DEcentralised Citizen-owned Data Ecosystems

EVM - Ethereum Virtual Machine

GDPR - General Data Protection Regulation

GNU - GNU's not Unix

GO - Gebiedonline

ICT - Information and Communications Technology

IP - Internet Protocol

IRMA - I Reveal My Attributes

IoT - Internet of Things

NGO - Non-governmental organization

NFC – Near-Field Communications

OCR - Optical Character Recognition

OS - Operating System

PbD - Privacy by Design

POSIX - Portable Operating System Interface

RBAC - Role Based Access Control

SDK - Software Development Kit

SBAC - Sharded Byzantine Atomic Commit

SNARK - Succinct Non-interactive ARGument of Knowledge

VM – Virtual Machine

ZK - Zero Knowledge

1. Introduction

DECODE (DEcentralised Citizen-owned Data Ecosystems) is a research project to enable practical alternatives to how we manage our personal data and interact on the Internet. DECODE developed decentralised technology that puts people in control of their personal data, giving them the ability to decide how it is shared [1]. The current models of data sharing enable service providers to collect citizens’ data in exchange for services. Given the high value of this data to service providers, a vast number of them appropriate data to extract value from it, for instance selling large aggregated datasets as well the possibility to target advertisement.

DECODE focuses research and development effort on novel notions of trust and privacy that can be operationalised in new governance frameworks, and innovative economic models (data commons) based on digital commons [2].

The digital commons are a form of commons involving the distribution and communal ownership of informational resources and technology [3]. Resources are typically designed to be used by the community by which they are created. In particular, the distinction between digital commons and other digital resources is that the community of people building them can intervene in the governing of their interaction processes and of their shared resources [4].

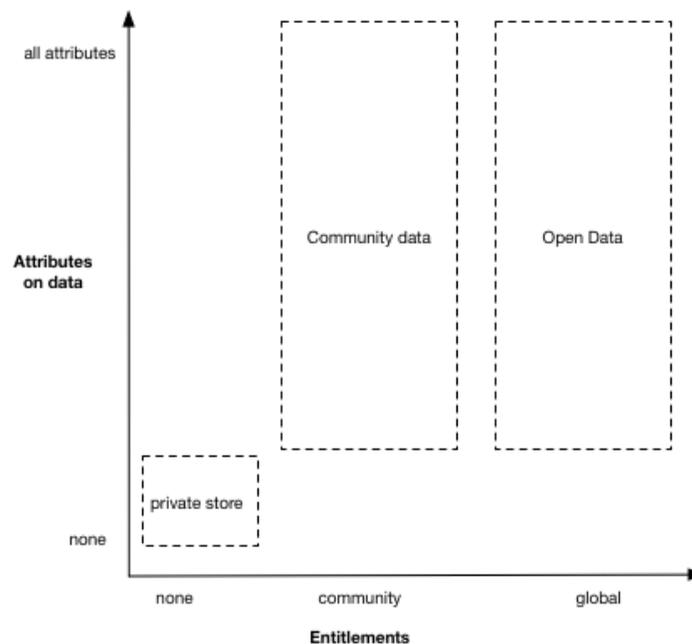


Figure 1 - Digital Commons

In DECODE, private data entitlements can be exhibited as authentication credentials in the public domain without revealing the contents of data. Novel concepts of data rights and

entitlements also apply to data being sent to or consumed by connected IoT objects in order to perform actions on the real world, making participants aware of who consumes the data gathered and how, while lowering the liability of data operators by implementing secure storage solutions that are accessible only with certain credentials and under certain conditions.

The project has researched, developed and tested free and open source software as the building block for distributed, privacy aware and trusted technology architectures, for decentralised data governance as well future identity management systems [1].

Identity on the internet has evolved from being implemented as centralised 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 centralised and federated models.

DECODE is an evolution of the concept of decentralised systems which leverages state of the art cryptographic techniques such as Distributed Ledgers, Attribute Based Credentials, Zero Knowledge Proofs, Threshold Credentials and Homomorphic Encryption to easily build systems to store data securely, give control and transparency over with whom and for what purpose data is shared and probably transact private information with other participants or organisations.

At a high level we can describe DECODE as being composed of the following:

- Set of specifications for distributed ledgers to support DECODE.
- 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)
- GNU/Linux based operating system for privacy by design networking (DECODE OS).
- Documentation needed for operators to write and deploy smart rules that request access to private data.
- Advanced cryptographic implementations to manipulate data using Zencode (Coconut scenario)

In the following Figure we list the points we aimed to develop in the DECODE project:

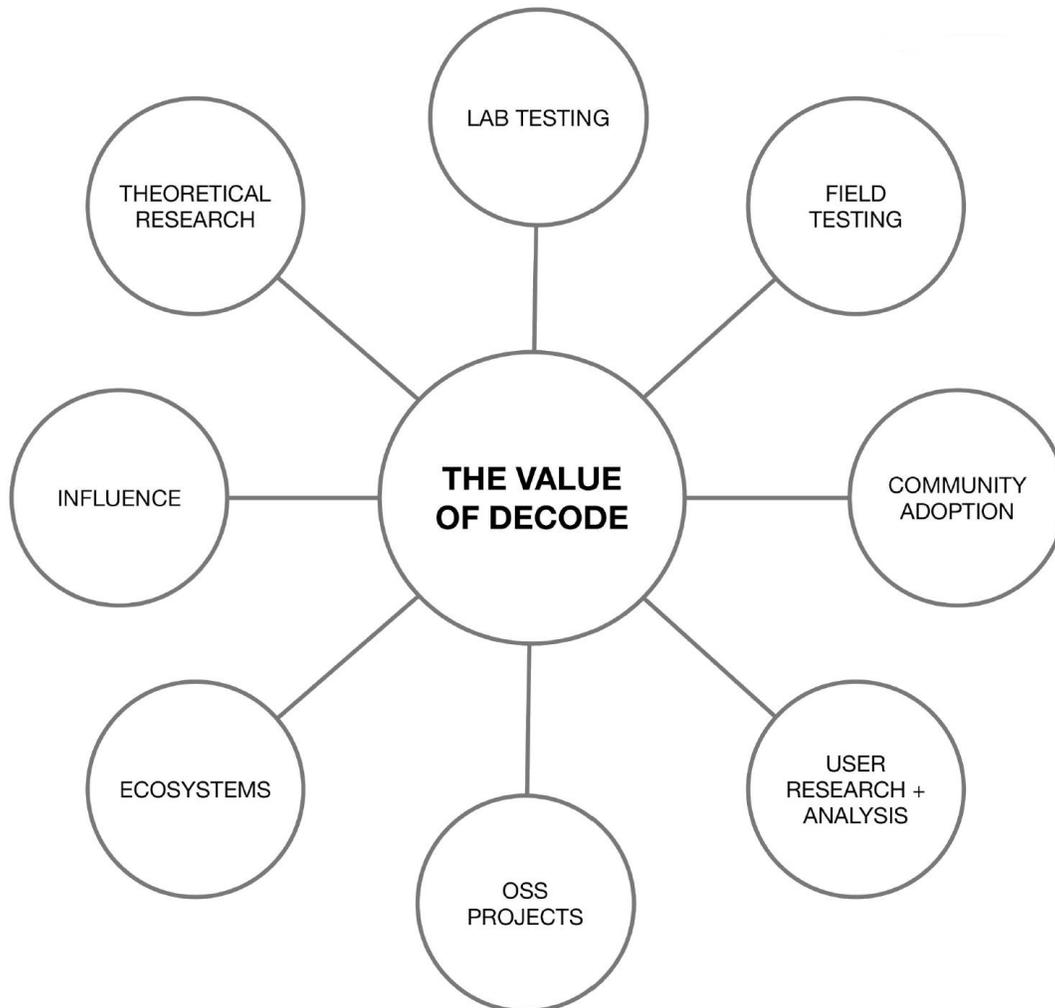


Figure 2 – The value of DECODE

The rest of the document is structured as follows: Section 2 describes the overall goals of the DECODE project. Section 3 goes through the architecture principles and Section 4 goes through the concept foundations of the project. Section 5 described the different elements of the DECODE architecture. Section 6 introduces the different pilots developed during the project. Section 7 concludes the document.

2. Goals

We identify the following key goals for DECODE:

- Ease the management of *private data, credential based authentication and secure storage across distributed networks*.
- Allow non-technical *operators* to write and review *smart rules* running on any device for end-to-end encryption.
- allow *smart rules* to access *private data* based on *entitlements* and matched *attributes*
- allow everyone to record *entitlements* on a *distributed ledger* whose integrity is resilient and verifiable

3. Architecture Principles

DECODE 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 [6] following key principles of modularity, clarity, composition, separation, simplicity, parsimony, transparency, robustness, representation and least surprise.

3.1. Free and open source

All work produced by DECODE is published as free and open source according to licenses approved by the Free Software Foundation Europe and emerging open hardware standards. The openness of the platform enables innovation and citizen participation.

DECODE adopts free software, however free software projects cannot be entirely considered as digital commons. As a matter of fact, writing a code and publishing it with a free license are not sufficient conditions in order to realize a free software. There are other necessary conditions, among them:

- the reputation inside the community
- the adoption of the good practices diffused in the community (for instance public repository, - continuous free upgrading, an efficient system of bugtrack and feedback the quality of the code, including its documentation to allow understanding of the code;
- the software's coverage, as the presence of automatic tests for evaluating the absence of bug on high percentages in the written code [3].

3.2. Modularity and interoperability

Adopting the key UNIX principle of modularity (simple parts connected by clean interface), enables building independent components which can be reused and combined to form a flexible eco-system of software products. DECODE 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.

DECODE 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.

3.3. Reuse don't re-invent

DECODE aims to be built upon the solid foundations of existing well proven software wherever appropriate. For example, "DECODE OS" is based on a well-known and solid OS, Devuan GNU+Linux.

3.4. Decentralisation and federation

The current era in technology has seen a shift from large monolithic systems to distributed decentralised systems, this is to meet the requirements of system - scaling, resilience and fault tolerance but also provides for decentralised governance models. DECODE builds upon decentralised models for data and identity management.

3.5. Privacy by design

DECODE aims to develop a privacy preserving data distribution platform to foster commons-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 [7].

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 lifecycle, 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 organisation in achieving GDPR compliance. As DECODE is designed with privacy in mind from the ground up it naturally affords a good foundation DECODE 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.

3.6. User friendliness

Building user-friendly tools and applications for end-users, and app developers for easy adoption is a core principle for DECODE. Using an outside in lean approach, where requirements from users' are researched and analysed, and prototypes are tested on target community groups allows DECODE to develop open, interactive and user friendly interfaces.

3.7. Determinism

One of the biggest challenges in managing distributed platforms built on diverse services and apps written in different languages is running on different devices (clients, servers, web browsers and mobile) while granting data integrity and the consistence of cryptographic transformations.

All data that undergoes a cryptographical transformation, will later need be transformed again, be transformed back or be matched against other data. 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.

Despite a good level of homogeneity among cryptographic algorithms and implementations, determinism can hardly be taken granted in this domain. 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 DECODE 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.

4. Concept Foundations

This section describes the foundational concepts that are combined to achieve the purpose of DECODE. DECODE is built upon several foundations:

- **Decentralisation of trust:** A key concept behind DECODE is to build a system which aims towards a model of *decentralised trust*. 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. For example, even though Bitcoin began life as a model for decentralisation, the current state is that the hash power is controlled by a few large mining pools. DECODE seeks to explore alternative, decentralised models and economic incentives (See section on Distributed Ledgers).
- **A distributed ledger:** A distributed ledger with decentralised 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.
- **Zero knowledge proofs:** The public nature of the ledger is in tension with a desire to maintain the privacy of the participants of the network. DECODE applies the concept of Zero knowledge proofs 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.
- **Attribute Based Cryptography:** In addition to the participant’s transactional data, a key element of privacy is to allow a strong control over data which is directly related to *Identity*. In DECODE all data is represented as Attributes. DECODE takes an approach to identity which states that what a participant discloses about their identity should only be related to the minimum transfer of information required for a particular interaction. Further, this transfer of information should occur through a privacy preserving mechanism. The cryptographic implementation of this mechanism is Attribute Based Credentials. This mechanism allows a participant to prove something about themselves without transferring any other identifying information to the *relying party* (The entity that requires proof). Underlying this selective disclosure mechanism is often a Zero knowledge protocol to allow for multi-show unlinkability. For instance, a participant can cryptographically prove their residency of a particular city without exposing sensitive information such as data of birth, national Id number, or the actual address. This mechanism can also be used to provide strong guarantees about authenticity of an interaction whilst preserving a level of individual anonymity, particularly relevant for scenarios of participatory democracy.
- **Cryptographically verifiable entitlements:** Building on its verifiable public record and privacy preserving cryptography and design, DECODE adds a mechanism for participants to declare and enforce agreements about how their data is consumed. DECODE refers to

this mechanism as the entitlements a participant agrees to over their data. These entitlements are cryptographically verifiable and can be extended to be cryptographically enforceable through Attribute Based Encryption. We can extend this data sharing capability to datasets for wide sharing - for example we can consider individual contributions to an aggregate dataset. This is often called the Digital Commons (3; 3).

- **A “Smart Rules” language to express governance of participants data:** Based on the principle of User Friendliness, giving participants access to a means to express and understand these entitlements is a further aim of DECODE, which is expressed through the development of a Smart Rules language and user interface.
- **A highly verifiable and controlled execution environment:** Finally, DECODE considers the full technology stack within which all of the above foundations executes. This includes the underlying hardware platforms Hub and the Operating System on which the software executes. DECODE terms this the *controlled execution environment* and explores ways to provide assurance, transparency and reproducibility of the execution environment.

Each of these conceptual building blocks are explored in detail in [8]. Each of the topics is a deep area of study in its own right so we provide references to allow further exploration.

5. DECODE Architecture

Figure 3 - DECODE Architecture shows a high-level view of how the DECODE components work together to create a network of validating nodes to which transactions can be submitted to the ledger. The DECODE P2P Network is formed by the Tor Dam nodes¹ communicating with each other. These run on DECODE OS which can be run on hardware “hubs” as explored in deliverable D4.4 [9]. This forms the foundation of the network between DECODE validating nodes. We call these validating nodes because they are running ledger nodes and therefore constitute the distributed ledger described within this document.

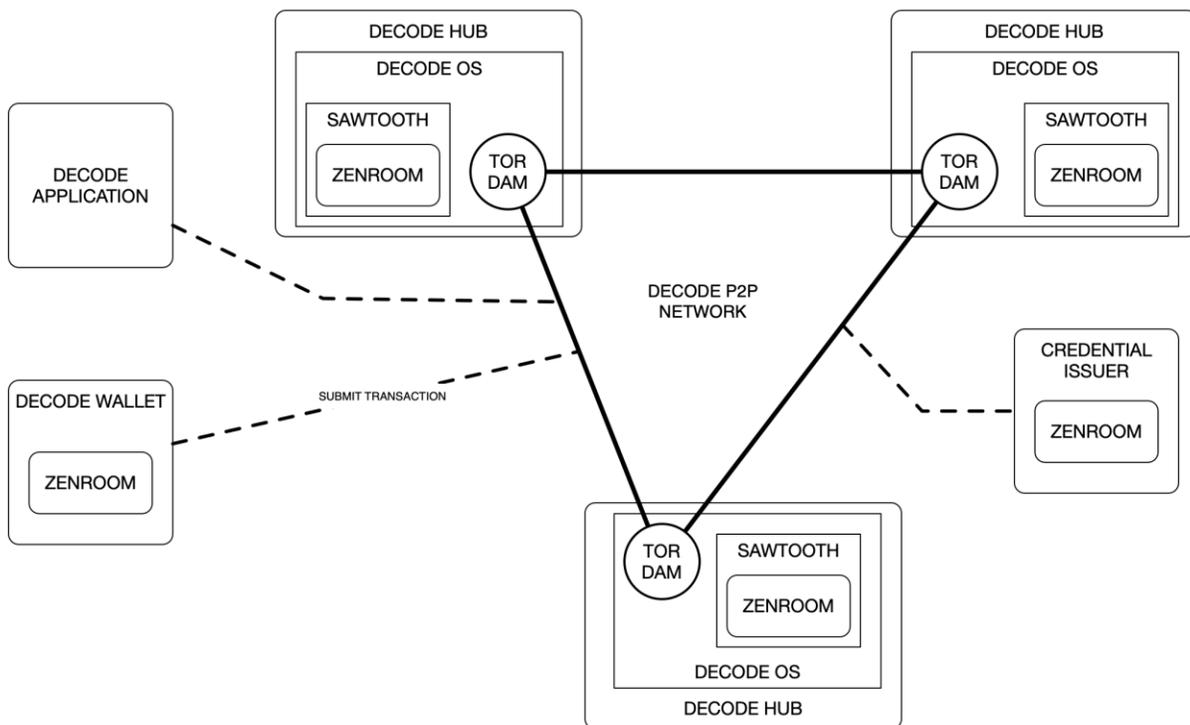


Figure 3 - DECODE Architecture

In the following table we list all the components developed in the DECODE project that are part of the architecture and are described in the subsections below:

¹<https://github.com/DECODEproject/tor-dam>

Table 1 - DECODE elements

DECODE Element	Functionality	Deliverable	Repository
Zencode/ Zenroom	Execution	D4.4	https://github.com/DECODEproject/Zenroom
Chainspace/ Sawtooth	Ledger	D1.4	https://github.com/DECODEproject/chainspace https://github.com/DECODEproject/Sawroom
Coconut	ABC encryption	D1.5	https://github.com/DECODEproject/
Decode P2P Network	Validation	D1.5	https://github.com/DECODEproject/
Decode OS	Execution	D4.4	https://github.com/DECODEproject/decode-os
Decode App	Key storage	D4.10	https://github.com/DECODEproject/bcnow

5.1. Zencode and Zenroom

Zencode is a software project inspired by the discourse on data commons and technological sovereignty, consolidated in the Zencode Whitepaper² as a living document that will continue to be updated beyond the span of the DECODE project. The established goal is that of improving people's awareness of how their data is processed by algorithms, as well facilitate 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. 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 isolated environment (the Zenroom³ VM) that can be easily ported to any platform, embedded in any language and made inter-operable with any blockchain. The Zencode implementation is inspired by modern research in language-theoretical security, it adopts Lua as direct-syntax parser to build a non-Turing complete DSL enforcing coarse-grained computations and recognition of data before processing. Its interpreter, the Zenroom VM, 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

² https://files.dyne.org/zenroom/Zenroom_Whitepaper.pdf

³ <https://zenroom.dyne.org/>

problem had occurred. Zencode language scenarios are written following a declarative approach and provide functional tools to manipulate efficiently even complex data structures.

5.2. Decode Ledger

DECODE has benefited from research and development made on the Chainspace software implementation [10], which constituted an early lab test-bed (in-vitro). To fulfil the goals outlined by this document it has been necessary to adopt a working DLT implementation that substitutes this component in DECODE, mostly due to operational shortcomings related to its capacity to scale, its reliability and the maintainability of its code.

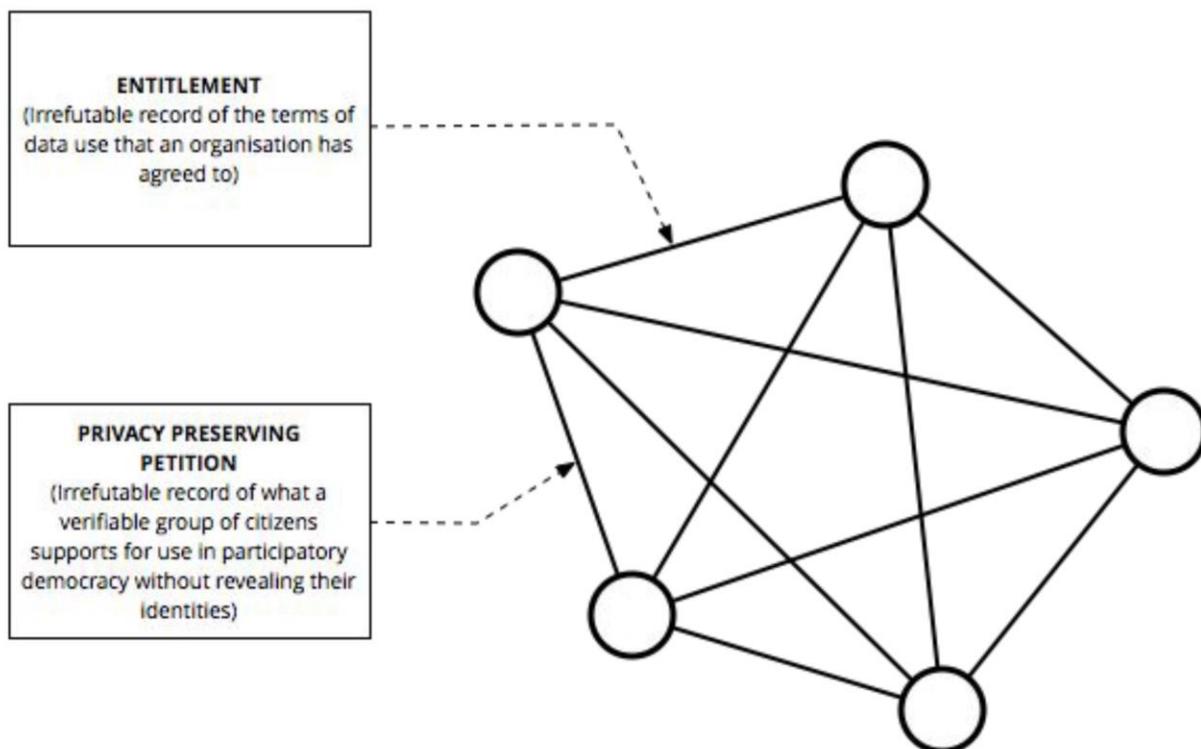


Figure 4 - What goes in the ledger?

Chainspace

Chainspace has been specifically designed for the needs of the project: to create and execute smart-contracts based on the Coconut⁴ crypto model. Chainspace aimed to become a decentralized infrastructure, known as a distributed ledger, to support user defined smart

⁴ Coconut: Threshold Issuance Selective Disclosure Credentials with Applications to Distributed Ledgers <https://arxiv.org/abs/1802.07344>

contracts and execute user-supplied transactions on their objects. The correct execution of smart contract transactions on any permissionless blockchain is verifiable by all. The system experimented with sharding state and execution of transactions, adopting Sharded Byzantine Atomic Commit (SBAC) protocol, a distributed commit protocol that aimed to guarantee consistency. To secure against subsets of nodes trying to compromise its integrity or availability properties, Chainspace adopted the consensus model of Byzantine Fault Tolerance (BFT). The full paper is available at [10].

Unlike other scalable but 'permissioned' smart contract platforms, such as Hyperledger Fabric [13] or BigChainDB [14], Chainspace aimed to be an 'open' system to allow anyone to author a smart contract, anyone to provide infrastructure on which smart contract code and state runs, and any user to access calls to smart contracts. Further, it aimed to provide ecosystem features, by allowing composition of smart contracts from different authors.

However, the security model of Chainspace, is different from traditional unpermissioned blockchains, that rely on proof-of-work and global replication of state, such as Ethereum. In Chainspace smart contract authors designate the parts of the infrastructure that are trusted to maintain the integrity of their contract---and only depend on their correctness, as well as the correctness of contract sub-calls. This provides fine grained control of which part of the infrastructure need to be trusted on a per-contract basis, and also allows for horizontal scalability.

Chainspace was a useful laboratory test, but had a number of areas that were not sufficiently developed to be satisfactory for production deployment, principally the implementation of Sharding. These were detailed during a "Production Readiness" assessment conducted by the DECODE team. Its development was however valuable for its experimentation value, to inform the development of Zencode smart-contracts. In early 2019 Facebook "acqui-hired" the Chainspace team meaning that development ceased and the technical issues were unlikely to be resolved. Consequently the decision was taken by the DECODE team to adopt a modular blockchain that is production-ready and has ways to plug different consensus and virtual-machines.

Sawtooth

Sawtooth is a free software blockchain implementation part of the Hyperledger consortium, a world-wide effort in industrial-grade blockchain development coordinated by the Linux Foundation. At the demise of Chainspace, the choice of Sawtooth was motivated by the fact that this blockchain has been engineered with simplicity and modularity in mind since the very beginning. Sawtooth does not impose a specific VM or consensus algorithm, allows building "permission-less" as well permissioned blockchain clusters and provides an extensively documented Application protocol interface (API) and a Software Development Kit (SDK) to plug custom made components. This approach adheres to a fundamental aim in DECODE: that of

developing technologies that are blockchain agnostic and portable across different DLT platforms.

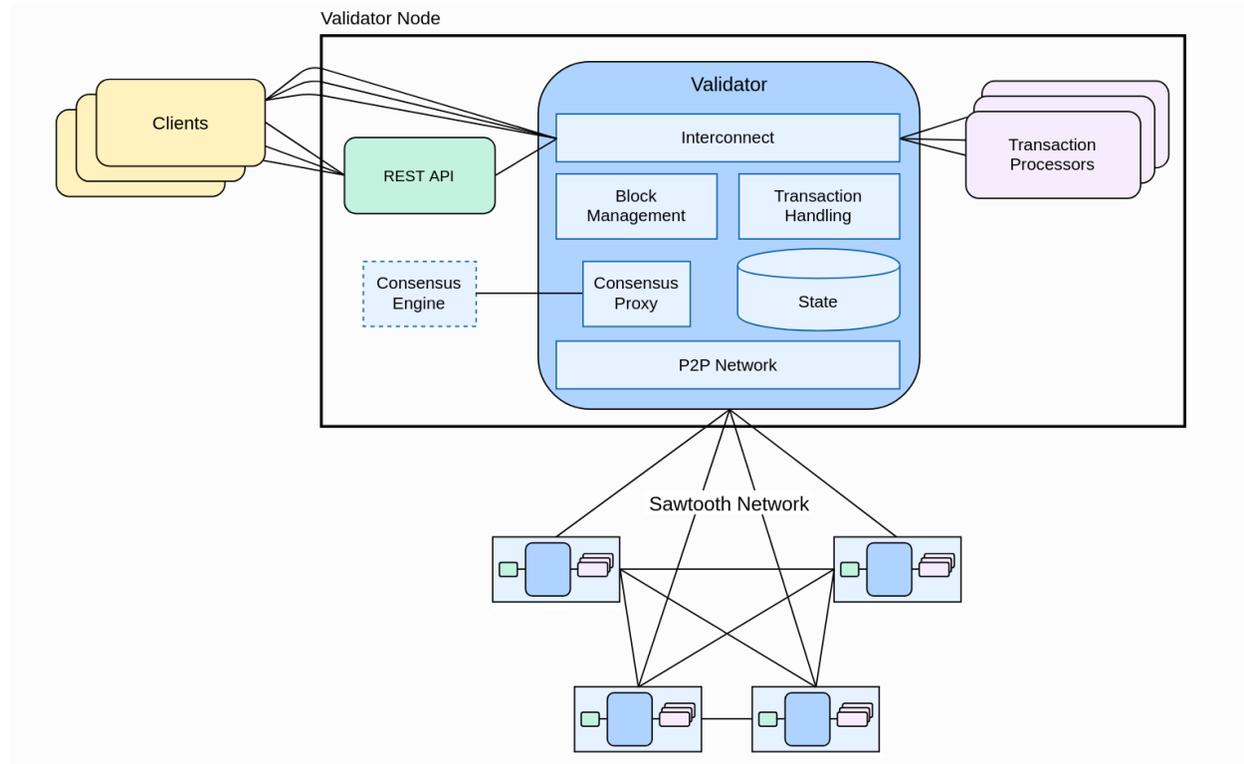


Figure 5 - DECODE ledger

Within the current DLT panorama then Sawtooth proves itself to be well beyond infancy by providing a production-ready setup that can clearly demonstrate the functioning of our VM and the Coconut crypto contracts in a deterministic environment.

5.3. Coconut

Selective disclosure credentials allow the issuance of a credential to a user, and the subsequent unlinkable revelation (or ‘showing’) of some of the attributes it encodes to a verifier for the purposes of authentication, authorisation or to implement electronic cash. While a number of schemes have been proposed, these have limitations, particularly when it comes to issuing fully functional selective disclosure credentials without sacrificing desirable distributed trust assumptions. Some entrust a single issuer with the credential signature key, allowing a malicious issuer to forge any credential or electronic coin. Other schemes do not provide the necessary re-randomisation or blind issuing properties necessary to implement modern

selective disclosure credentials. No existing scheme provides all of threshold distributed issuance, private attributes, re-randomisation, and unlinkable multi-show selective disclosure.

Coconut is a novel scheme that supports distributed threshold issuance, public and private attributes, re-randomization, and multiple unlinkable selective attribute revelations. Coconut allows a subset of decentralised mutually distrustful authorities to jointly issue credentials, on public or private attributes. These credentials cannot be forged by users, or any small subset of potentially corrupt authorities. Credentials can be re-randomised before selected attributes being shown to a verifier, protecting privacy even in the case all authorities and verifiers collude.

The lack of full-featured selective disclosure credentials impacts platforms that support ‘smart contracts’, such as Ethereum, Hyperledger and Chainspace. They all share the limitation that verifiable smart contracts may only perform operations recorded on a public blockchain. Moreover, the security models of these systems generally assume that integrity should hold in the presence of a threshold number of dishonest or faulty nodes (Byzantine fault tolerance). It is desirable for similar assumptions to hold for multiple credential issuers (threshold aggregability). Issuing credentials through smart contracts would be very useful. A smart contract could conditionally issue user credentials depending on the state of the blockchain, or attest some claim about a user operating through the contract – such as their identity, attributes, or even the balance of their wallet. As Coconut is based on a threshold issuance signature scheme, that allows partial claims of maintaining blockchain, or a side credential, it allows collections of authorities in charge of maintaining a blockchain, or a side chain based on a federated peg, to jointly issue selective disclosure credentials.

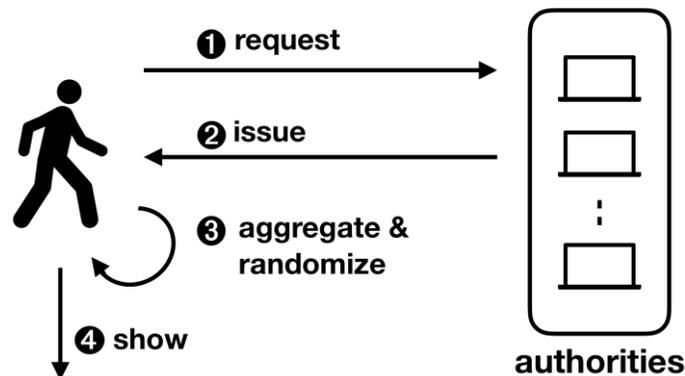


Figure 6 - Coconut architecture

Coconut is a fully featured selective disclosure credential system, supporting threshold credential issuance of public and private attributes, re-randomisation of credentials to support multiple unlinkable revelations, and the ability to selectively disclose a subset of attributes. It is embedded into a smart contract library, that can be called from other contracts to issue credentials. The Coconut architecture is illustrated in Figure 6 - Coconut architecture. Any Coconut user may send a Coconut request command to a set of Coconut signing authorities;

this command specifies a set of public or encrypted private attributes to be certified into the credential (1). Then, each authority answers with an issue command delivering a partial credentials (2). Any user can collect a threshold number of shares, aggregate them to form a consolidated credential, and re-randomize it (3). The use of the credential for authentication is however restricted to a user who knows the private attributes embedded in the credential—such as a private key. The user who owns the credentials can then execute the show protocol to selectively disclose attributes or statements about them (4). The showing protocol is publicly verifiable and may be publicly recorded. Coconut is detailed in deliverable D1.5 [16] and is implemented inside Zenroom, its credential and petition schemes are operable via Zencode language that is understandable by humans.

5.4. Decode P2P Network

The integrity and resilience of the network is provided through a distributed network of validating nodes. One of the key architectural features of DECODE is that it separates *execution* of logic (*contracts*) from the *verification* of that logic, which allows for privacy aware execution.

The validating nodes are key to providing the integrity and availability of the DECODE network. Therefore, we build them from the ground up with a strong emphasis on verifiability by basing them on the DECODE OS. Each node also contains the distributed ledger node and any other libraries and software that is required to participate in the DECODE network. This includes cryptographic functionality and P2P networking capabilities to allow dynamic and evolving P2P networks to be formed over the Tor hidden service protocol, granting an additional level of privacy for connections which is "by design" lowering the liability of their Internet Service Providers (ISP)

It is important to note that DECODE does not mandate that every participant host a validating node. The minimum software you need access to in order to participate is a DECODE App.

5.5. Decode OS

The DECODE OS is the base operating system and cloud container running all server-side software designed, developed and deployed for the DECODE project. This operating system is based on the renown Devuan GNU+Linux distribution, a fork of the now 20 years old Debian distribution, maintained by the Dyne.org foundation and an open community of volunteers. Devuan forked Debian to preserve the security, simplicity and minimalism of UNIX systems, still running modern software applications and inheriting the security patches from Debian.

DECODE OS is capable of building a P2P network automatically from a few published seed hosts and to plug a consensus mechanism that authorises the connection and propagation of node lists. It does so leveraging Tor as an underlying network protocol, communicating via socks5 channels and providing to inner application the whole network mapping via a Redis server. Its

operation requires only 128MB of RAM and runs smoothly on the hardware ARM devices designed to be the minimum sized nodes in DECODE.

DECODE OS also makes it possible to deploy very easily nodes, bypassing firewalls and lowering the liability of ISPs hosting the nodes, which are agnostic to the content vehicled. Entry nodes can be programmed to offer connectivity to the DECODE App through simple and documented APIs.

5.6. DECODE App

In the context of the DECODE pilots described in the next section, a smartphone app has been developed in order to have a user interface to provide DECODE functionalities to end-users. Within the app (DECODE App) a wallet functionality is included. The wallet is the minimum component a person requires to interact with DECODE. 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 DECODE 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.

Optionally a participant can push the cryptographic functions of the wallet onto a hardware device, similar to Trezor⁵, Ledger Nano⁶ from the blockchain world.

The app is the primary interface between DECODE and the participant, the other being the *application* themselves. The app interface has been developed through a user centric design process which provides a state-of-the-art experience for participants focused on transparency of who they have shared their data with. The app is also be the point at which a participant interacts with smart rules.

⁵<https://trezor.io/>

⁶<https://www.ledger.com/>

6. DECODE Pilots

The DECODE pilots are a fundamental part of the project as they must serve to engage large number of citizens and local communities to demonstrate the use of the different privacy-enhancing, decentralized and rights-preserving digital tools we are developing, while also integrating the legal, economic and social aspects of the data commons development. 4 pilots have been developed during the project: 2 pilots in Barcelona and 2 pilots in Amsterdam.

6.1. Barcelona Pilots

The Barcelona pilots have one clear goal in addition to those shared with DECODE: Use the city as a laboratory to develop, enable and sustain city data commons. This means, that the pilots must aim at testing the use of DECODE technology with real users, in order to enable real data sovereignty for citizens and offer communities the possibility of commonly sharing data to enhance the public good. By combining private and public sources of data city problems can be addressed by the crowd, that is, solving them collectively, while also preserving citizen's digital rights such as privacy and the right to information self-determination. The technopolitical details on what data commons is detailed in deliverable D2.5 [18]. The data commons vision within the context of DECODE was also outlined in an article by DECODE Project Coordinator Francesca Bria for the Guardian⁷: "Our goal is to create "data commons" from data produced by people, sensors and devices. A data commons is a shared resource that enables citizens to contribute, access and use the data – for instance about air quality, mobility or health – as a common good, without intellectual property rights restrictions".

To that end, and given the extended exploration performed in Barcelona as described in D1.1 [17], we have tried to build a unified common framework for pilot development that is depicted in Figure 7 - Schematic view of Barcelona pilot framework. Specifically, what we aim to realize is:

- Citizen Science and IoT Data Governance pilot: Testing the concept of granular data sharing permissions - data entitlements- at different levels (individual, community and public) with privacy enhancing technologies (PETs) and IoT devices. This has the aim to crowdsource data and to that end we must raise awareness with interested communities and work with them to define policies for the use of crowd-sourced aggregated data. The secondary objective of the pilot is to test privacy-enhancing (PETs) technology with low-risk personal data in order to consider its later expansion to more sensitive domains such as health and others.

⁷<https://www.theguardian.com/commentisfree/2018/apr/05/data-valuable-citizens-silicon-valley-barcelona>

- Digital Democracy and Data Commons pilots: Test PETs technology to protect political opinions of citizens in digital democracy platforms, while enabling them to generate anonymized datasets that collectively serve to detect city problems. We also aim to apply state of the art distributed ledger and cryptography technology to the open democracy platform Decidim (decidim.org) to discuss collectively the data governance policies and use of the aggregated generated datasets for public good. Such use should allow for transparent, auditable, yet privacy aware management of political initiatives support. This open democracy governance experiment directly informs the city of Barcelona ethical and responsible data strategy.

While we do not consider it as a pilot by itself, it is relevant to mention at this stage the role of the Barcelona Now⁸ tool developed during this first project stage. In order to demonstrate and integrate the technology tested in the two pilots, the consortium has built this data commons analytics and visualization tool, which in conjunction with the DECODE wallet (described in this document), is the main interaction point for users to visually interact with the datasets and policies being discussed and created.

Barcelona Pilots: How they fit together

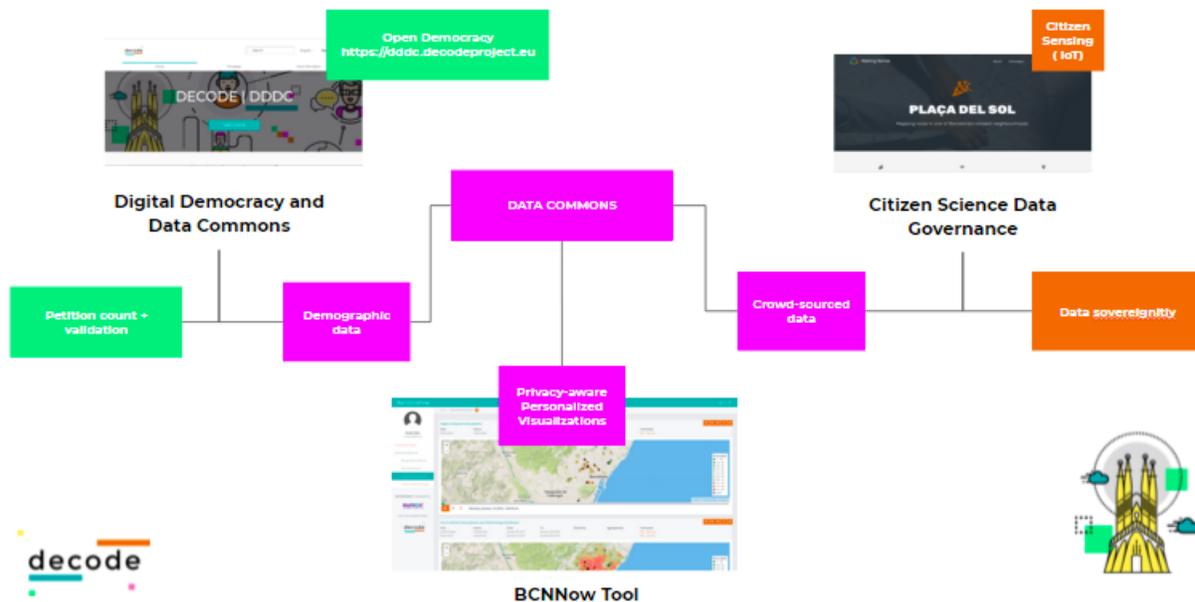


Figure 7 - Schematic view of Barcelona pilot framework

Additional information about the Barcelona pilots can be found in D5.6 [21] and D5.9.

Citizen Science and IoT Data Governance

Earlier work within DECODE involved the creation of a suite of cooperating microservices that attempted to provide a robust architecture for IoT devices but one in which the normal

⁸<https://github.com/DECODEproject/bcnnow>

structures of control were inverted. Rather than the provider of the device or the IoT infrastructure being in control of the storage and transmission of data, we attempted to give that control to the users generating the data.

The mechanism used for this inversion was based on applying strong encryption techniques to the data generated by devices and giving control of that encryption to the user rather than the relying on the infrastructure. This "cryptographic control" allows DECODE to sidestep the philosophical question of whether data can ever be owned, by instead framing the problem as being "who controls the keys used to encrypt the data?". With this framing it does not matter who gains access to the raw data being emitted from the device as only the specific entities for whom the data is encrypted will be able to decode and make sense of it.

More detail on this scheme can be found in deliverable D3.7 [19] and D3.9 [24] which attempts to document the previous work. An implementation of the DECODE app described in Section 5.6 was developed for the related DDDC [20] pilot. The consortium decided to improve this mobile app to incorporate a range of extended functionalities and make it fully modular, extensible and ready for deployment by third parties, however it was realised that this work would take too long in relation to the schedule already planned for the IoT pilot workshops.

To bridge this gap a webapp was developed by Thingful which replicated the core functionality of the DECODE app in terms of being able to participate in the Coconut entitlement flow, but with additional functionality required for the IoT pilot. This application was intended to be the primary interface by which pilot participants could configure and add IoT devices to DECODE, and control the flow of data from these devices to the Barcelona Now visualization platform. To do this the application was required to provide functionality to implement the following two primary user journeys:

1. To onboard new IoT devices and allow the user to control where and how the data from this device is transmitted and encrypted by adding their device to one or more communities.
2. To allow the user to prove their membership of a community using the Coconut to securely share an attribute with the Barcelona Now visualization tool.

For a full explanation of the system architecture and terms used above please refer to the following deliverables:

- D5.6 Deployment of Pilots in Barcelona [21]
- D3.7 Control and entitlement system for data owners [19]
- D3.9 IoT privacy-enhancing data sharing: integration with pilot infrastructures [24]

However, we also include a copy of the system architecture here in Figure 6 – IoT Pilot framework.

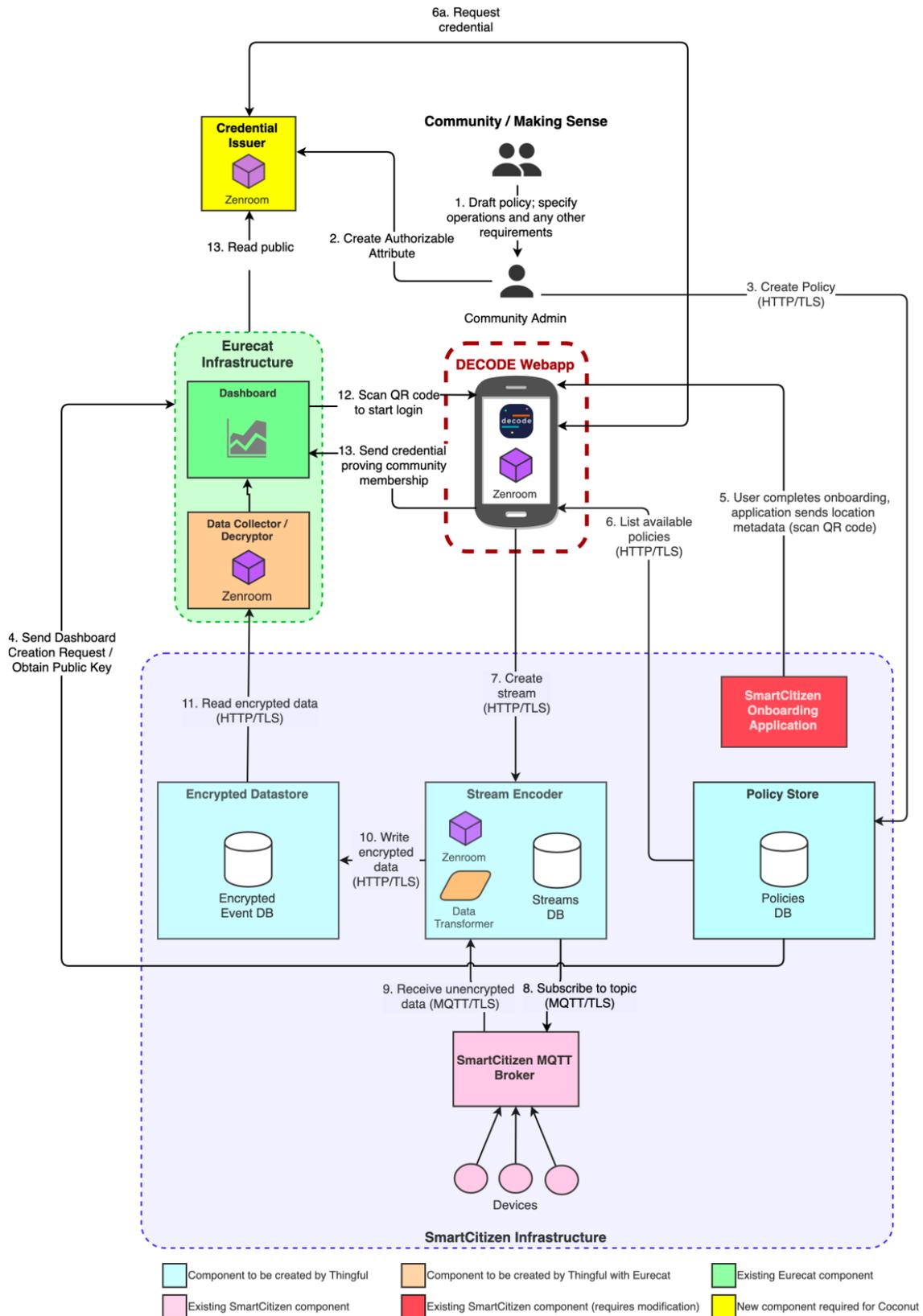


Figure 8 – IoT Pilot framework

Digital Democracy and Data Commons

The Digital Democracy and Data Commons pilot experimentally implemented the DECODE mission⁹. The core of the pilot is a technologically-enabled (via DECODE and Decidim technologies) participatory process for experts, citizens and city representatives to: 1-test the new DECODE-Decidim system (from now on DecidimCODE) for strongly secure, private, transparent and data enriched democratic decision making; 2-deliberate upon data politics and economics, at the local level and beyond; and 3-constitute an experimental digital data commons, whose shape was defined by the ideas and practices coming from the Digital Democracy and Data Commons participatory process itself and later linked to the data commons framework defined with the City of Barcelona¹⁰.

The Digital Democracy and Data Commons (from now on DDDC) pilot is oriented to experimentally implement the DECODE mission. The primary DECODE mission is to develop technologies and tools that enable sociotechnical systems (be they political or economic in character) that give people more individual and collective (democratic) control over their data, while enabling uses that provide more collective benefits from it.

In the pilot, this mission is advanced by the convergence of two technological systems: DECODE and Decidim. As a result, the DDDC pilot has two threads, the Digital Democracy thread and the Data Commons thread. The Digital Democracy thread, the primary one, has a central aim: testing the DecidimCODE system. Ultimately, this thread speaks to the potential of DECODE technology to push forward Decidim's technology and vision of participatory democracy. The Data Commons thread, which is a complementary (and somehow secondary) thread, is oriented to cover the other two general aims of the pilot, namely, to collectively deliberate upon data policies and experiment with data commons. Ultimately, this thread speaks to the potential of Decidim to advance DECODE's vision of alternative forms of data governance and digital economy. The pilot orientation has a theoretical background behind and can be broken down into a series of more detailed objectives, listed below:

- Test and improve DECODE technology;
- Integrate DECODE technology with Decidim;
- Develop and test DECODE legal tools;
- Test toolkit;
- High quality and quantity participation;
- Awareness raising;
- Uptake;
- Policy and social innovation;

⁹DECODE high-level vision outlined in this article by Francesca Bria: <https://www.citymetric.com/horizons/people-should-control-their-digital-identity-barcelona-s-chief-technology-officer-decode>

¹⁰The data commons policy of the city of Barcelona can be found here: https://ajuntament.barcelona.cat/digital/sites/default/files/2018_mesuradegovern_en.pdf

- Test concepts and frameworks.

The two key sets of technologies to be deployed in the Digital Democracy and Data Commons (DDDC) pilot are DECODE technologies and Decidim. The DECODE technology is a complex ensemble of digital tools and infrastructures composed by a wallet, a distributed ledger, a network of nodes, and a dashboard. We describe each of them in turn. Decidim is a free/open, digital platform for participatory democracy. Its software is fully open and available at decidim.org. Furthermore, Decidim is a common's free and open project and infrastructure involving not only code but also documentation, design, training courses, a legal framework, collaborative interfaces, user and facilitation communities, and a common vision.

The Decidim technology plays three key roles in the DDDC pilot.

- Digital Democracy and Data Commons (DDDC). Decidim software is used to set up a web that has three key purposes:
 - facilitate the test of the DECODE technology;
 - enable the DDDC participatory process;
 - enable the experimental constitution and democratic governance of the DDDC.
- Decidim.barcelona. The Decidim.barcelona instance serves to publicly announce the DDDC participatory process.
- Decidim.org. As a potential result of the pilot, DECODE technology may be integrated in the decidim software, thereby ensuring compatibility and promoting its use. In the longer term, DECODE legal tools and concepts may be including in Decidim's social contract, license, and data governance, more broadly.

The connection of the various DECODE and Decidim components can be seen in Figure 7 - integration plan for the Digital Democracy and Data Commons pilot, and were achieved through an integrated flow that were at work during the final petition signing step of the participatory process, detailed in D2.5 [18].

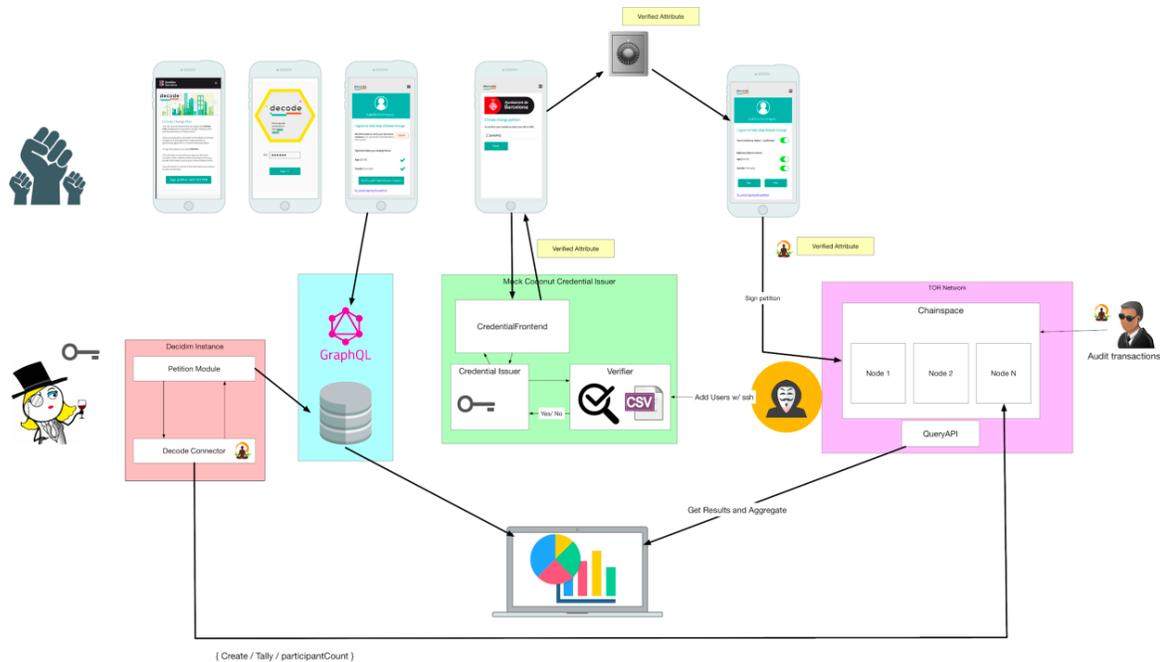


Figure 9 - Integration plan for the Digital Democracy and Data Commons pilot

6.2. Amsterdam Pilots

The Amsterdam pilots take an educational approach towards citizens and public administrations, which is guided by the following:

- Pilot partners aim to build technical solutions that demonstrate opportunities and possibilities in the realm of digital identity.
- The solutions should be useable by the general public and provide a clear path forward for further development.
- Amsterdam pilot partners utilize this technology as a framework to educate people about DECODE values such as privacy by design, attribute-based credentials, digital identity, data ownership, and data commons.
- Finally, the direct, applied, real-world effects of these pilots should clearly demonstrate the benefits of DECODE values to users.

This has led to two [2] current pilots in Amsterdam: Gebiedonline and Claim Verification (18+). Both focus on implementation of Attribute-Based Credentials. More information about the Amsterdam pilots can be found in D5.5 [22].

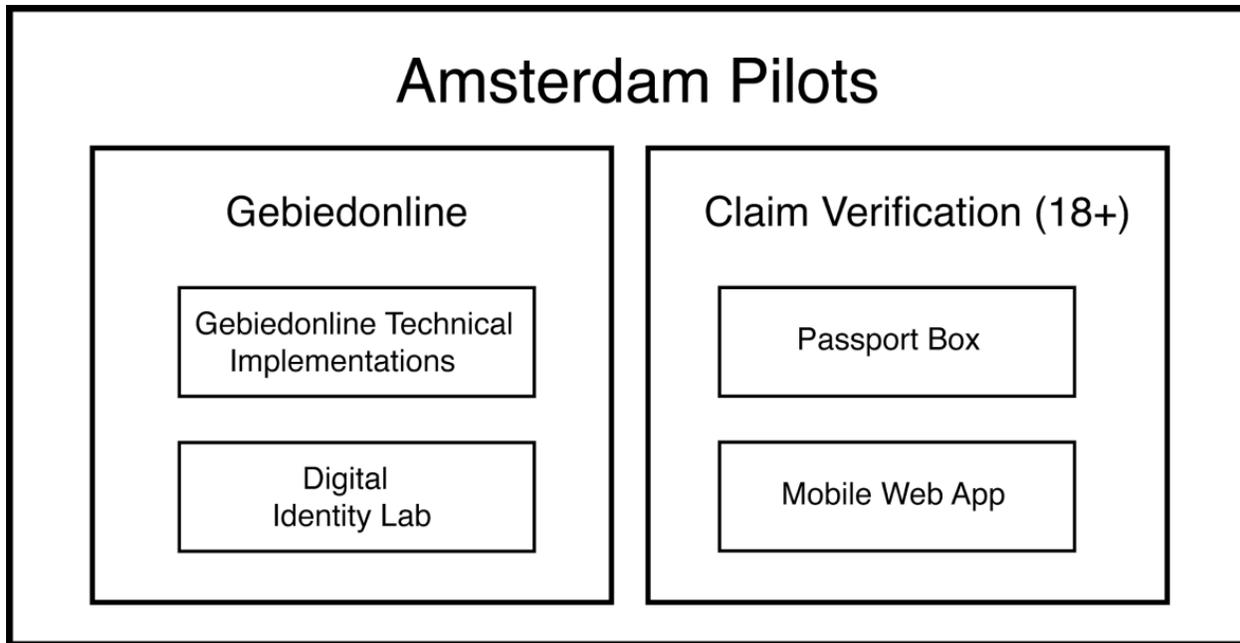


Figure 10 - Amsterdam Pilots

GebiedOnline (Neighbourhood online)

Gebiedonline (GO) is a pre-existing digital platform that enables local people, groups and organisations to view events taking place in their neighbourhood, share news, exchange and borrow products and services, and meet people. It is a community owned and operated member-based cooperative. The Gebiedonline board and community have a clear view: they want their platform to be open source, inclusive and to enable their users to have granular control over the data they share with the platform and with other users in the community. Gebiedonline has been a DECODE partner since having taken part in an open challenge organized by Waag in Spring 2017, which challenged communities to take part in DECODE.

Gebiedonline has strongly and successfully implemented the concept of a community-managed social network. It thus represents DECODE many criteria for pilot selection (giving people ownership of their personal data, decentralized IoT access, data shared for the public good). However, as noted in D1.1 [17], Gebiedonline is still “missing features on security and is not sufficiently privacy aware”. As such, the platform provides a great opportunity for a use case that technically implements security and verification improvements.

To these ends, the pilot proposed to build and pilot a set of features to use and access the platform using Attribute Based Credentials. With Gebiedonline, the Amsterdam pilot developed two main technical use cases utilizing the IRMA¹¹ platform:

¹¹<https://irma.app/docs/overview/>

- Integrating the IRMA API into an existing GO neighborhood/network. This would allow members of a (to-be-selected) neighborhood/network to share personal attributes with others in their community without sharing non-essential personal data. For example, a user could prove his or her place of residency based on census records, without ever having to share their address with Gebiedonline.
- Logging in to GO with IRMA, which provides a more secure login than is currently available on the site.

IRMA is at its core a set of software projects implementing the Idemix attribute-based credential scheme. An *attribute* is a statement or property about a person, such as "I am over 18 years old" or "my name is John Doe".

These attributes are grouped together in a *credential*. In attribute-based credential schemes such as Idemix, such a credential can be issued to a user by a trusted party called the *issuer*. This issuer creates a digital signature over the credential and its containing attributes using its *private key*. The user receives the credential as well as the issuer's signature in her [IRMA mobile app](#).

After that, the user can disclose these attributes to other parties, who are called *verifiers*, selectively showing some and hiding the other attributes from the credential. The verifier then receives the disclosed attributes, as well as a *proof of knowledge* which proves to the verifier that the user:

- knows the attributes from the credential which are not being disclosed
- owns a valid issuer signature over the disclosed attributes and hidden attributes.

The verifier can check the validity of this proof of knowledge using the issuer's *public key* that corresponds with the private key with which the issuer signed the attributes (thus, the verifier must know this public key). The verifier can tell from this that the user has at some point received the disclosed attributes from the trusted issuer. Therefore, it can trust the authenticity of the attributes. (This proof of knowledge does *not* include a full copy of the signature over the attributes, so that even if all attributes of the credential were disclosed simultaneously, the verifier can impossibly use the received attributes and proof of knowledge to disclose these attributes itself to others.)

Discussing specific options for further and future use cases. Integration of the IRMA API paves the way for peer-to-peer (P2P) identification, which would allow, for example, two verified community members to approve that a third person does indeed live in their neighbourhood.

During initial development for this pilot, it became clear that there was an additional value in the educational aspects of this process—discussing attribute-based credentials with people and exploring its potential applications and implications. Doing so with tangible examples on hand provided pilot partners with a unique and highly practical lenses through which to approach concepts such as digital identity, data ownership and how these are examples of the way in which technology embeds or obstructs shared ethics and public values.

As such, the Gebiedonline technical use cases have provided the basis for further educational outreach about digital identity in this pilot, of which citizens and public administrations are

target groups. In particular, this outreach aims to inform people about ethical and technical issues and risks regarding attributes, identification data ownership and management, as well as to provide opportunities to address these issues through the uptake of technology. During dedicated events, pilot partners, citizens, and local public administrators co-created concepts of digital identity together; this went beyond collecting input, and allowed those present to create a broad perspective on the topic. This conversation has proven to be especially fruitful when, using decode principles, architecture and technology, specific actionable alternatives can be offered to the audience. This task has been carried out through various publications, events, and consultations which are described further in document D5.6

Claim verification (18+)

Digital identity is a sprawling, abstract subject that can be made more tangible. Rather than focus on complex organizational problems, pilot partners wanted to make a manageable user experience case study on how digital identity could look in a concrete form. More specifically, with this pilot, partners wanted:

- to develop something of use to the citizens of Amsterdam;
- a clear user experience, a full understanding of the entire customer journey from A to B;
- to give citizens access to their personal data as stored in the municipal database, and allow them to share these data in a different context, on- or offline;
- a pilot that could be approached iteratively, where initial steps could both have a clear impact and also leave room for further development and applications;
- to demonstrate that these aims are possible and can happen, to inspire further trust and belief in this approach towards citizen identity.

With the Passport Box, pilot participants were able to prove that they are over a certain age (for example, over 18) - without having to share their full identity, date of birth, or social security number - through the use of attribute-based credentials.

The pilot provides a station to verify the personal data contained inside a passport (name, nationality, birthdate, photo) so that an operator can decide to release a credential when a condition bound to such data verifies. Zenroom can then be used to provide a portable credential associated to the issuer and the specific attribute issued. The credential released can also be a zero-knowledge proof attribute whose use can be publicly verified without leaving a trace that is directly connected to the subject holding the attribute. In practice this realised an easy to setup station to make the 18+ pilot possible: releasing proofs that people are older than 18 years old and for instance is possible for them to use the credential issued to buy alcohol at a supermarket or in a club, without having to disclose their full identity.

A modular approach to implement the core features of DECODE's software stack has been developed by Dyne with particular emphasis on data encryption. It is based on the Zenroom VM which also features:

- advanced zero knowledge proof functionalities (Coconut crypto scheme)
- natural language execution for cryptographic transformations

- data manipulation (Zencode interpreted since version 0.8.1)
- modular bindings to Python and NodeJS

The proposed implementations focus on privacy and added benefits for users, but also on developer experience to facilitate the integration in bigger applications which can be developed ad-hoc for pilots.

The main pilot implementation realized in cooperation with the UX and development team of Amsterdam municipality includes:

- Hardware and software to scan a passport's digital contents via NFC
- Zenroom integration to encrypt and manipulate the data extracted
- 3D printed prototype of casing and totem for the passport reader

Further description of the pilot, including progress detail and results, is included in D5.5 [22].

7. Conclusion

DECODE provides tools that put individuals in control of whether they keep their personal data private or share it for the public good. DECODE is an evolution of the concept of decentralised systems which leverages state of the art cryptographic techniques such as Distributed Ledgers and Attribute Based Credentials and makes it easy to build modular and distributed systems that provide its *participants* the capability to store and exchange data securely, give control and transparency over with whom and for what purpose data is shared and transact with other participants or organisations.

In this document we described all the tools built during the project, that experimentally developed practical alternatives to how we use the internet today – 4 European pilots show the wider social value that comes with individuals being given the power to take control of their personal data and given the means to share their data differently. DECODE explored how to build a data-centric digital economy where data that is generated and gathered by citizens, the Internet of Things (IoT), and sensor networks is available for broader communal use, with appropriate privacy protections. As a result, innovators, startups, NGOs, cooperatives, and local communities can take advantage of that data to build apps and services that respond to their needs and those of the wider community.

All the software developed during this project is not only useful for the objectives of the DECODE project but will be also useful for future EU initiatives in open data and new governance methods since all software developed is fully available free and open source.

Given the sophistication, effort and academic specialisation that was required to create a protocol such as Coconut, a key contribution of DECODE has been the Zenroom VM which has taken these ideas, industrialised them and made them accessible to a much broader community, effectively democratising this knowledge and providing a platform to allow its benefits to be realised by wider society.

Bibliography

1. **Tom Symons, T. B.** *Me, my data and i: The future of the personal data economy [online]*. . 2017. Available from: <https://decodeproject.eu/publications/me-my-data-and-ithe-future-personal-data-economy>.
2. **Bria, F.** (2018) Our data is valuable. Here's how we can take that value back. Published in The Guardian: <https://www.theguardian.com/commentisfree/2018/apr/05/data-valuable-citizens-silicon-valley-barcelona>
3. **Eleonora Bassi, J. C. D. M., Marco Ciurcina.** *D1.8 legal frameworks for digital commons decode os and legal guidelines*. s.l. : DECODE Project, 2017.
4. **Buterin, V.** *Ethereum white paper [online]*. s.l. : Available from: <https://github.com/ethereum/wiki/wiki/White-Paper#history>.
5. **Drummond Reed, D. H., Jason Law.** *The technical foundations of Sovrin [online]*. s.l. : Available from: <https://sovrin.org/wp-content/uploads/2017/04/The-Technical-Foundations-of-Sovrin.pdf>, 2016.
6. **Raymond, E. S.** *The art of unix programming*. s.l. : Pearson Education, 2003.
7. **Jaap-Henk Hoepman, E. B., Shehar Bano.** *D1.2 privacy design strategies for the decode architecture*. s.l. : DECODE Project, 2017.
8. **George Danezis, Shehar Bano, Mustafa Al Bassam, Alberto Sonnino.** *D1.4 First Version of the DECODE Architecture*. s.l. : DECODE Project, October 2017.
9. **Jelinčić, Denis Roio and Ivan.** *D4.4 First release of the DECODE OS*. s.l. : DECODE Project, August 2017.
10. **M. Al-Bassam, A. Sonnino, S. Bano, D. Hrycyszyn and G. Danezis.** *Chainspace: A Sharded Smart Contracts Platform*. In Proceedings of the Network and Distributed System Security Symposium (NDSS), 2018.
11. **Bootle, j., Cerulli, a., Chaidos, p., and Groth, j.** *Efficient zero-knowledge proof systems*. In Foundations of Security Analysis and Design VIII. Springer, 2016.
12. **Danezis, g., Groth, j., Fournet, c., and Kohlweiss, m.** *Square span programs with applications to succinct nizk arguments*. Springer Berlin Heidelberg.
13. **Cachin, c.** *Architecture of the hyperledger blockchain fabric*. In Workshop on Distributed Cryptocurrencies and Consensus Ledgers, 2016.
14. **Mcconaghy, t., Marques, r., Müller, a., De jonghe, d., Mcconaghy, t., McMullen, g., Henderson, r., Bellemare, s., and Granzotto, a.** *Bigchaindb: a scalable blockchain database. white paper, BigChainDB*. 2016 : s.n.
15. **Olson, K., Bowman, M., Mitchell, J., Amundson, S., Middleton, D., Montgomery, C.** *Sawtooth: An Introduction*. s.l. : Linux Foundation, January 2018.
16. **Alberto Sonnino, Shehar Bano, George Danezis, Jim Barritt.** *D1.5 Intermediate Version of DECODE Architecture*. s.l. : DECODE Project, January 2019.
17. **Jill Irving, Jen Hughes, Jim Barritt, Pinar Wennerberg, Priya Samuel, Andrei Biasprozvanny, Camilla Serri Colombo, Ina Tsetsova.** *D1.1 DECODE Scenarios and requirements definition report*. s.l. : DECODE Project, 2017.

18. **Calleja-López, Antonio.** *D2.5 Technopolitical Democratization and Digital Commoning: the Case of the Digital Democracy and Data Commons (DDDC) pilot.* s.l. : DECODE Project, September 2018.
19. **Francesca Bria, Oleguer Sagarra, Javier Rodríguez, Pau Balcells.** *D5.6 Deployments of pilots in Barcelona.* October 2018.
20. **Mulube, S.** *D3.7 Control and entitlement system for sensor data owners.* s.l. : DECODE Project, 2018.
21. **DDDC Participatory Platform.** [Online] [Cited: 24th May 2019.]
<https://dddc.decodeproject.eu/>.
22. **Max Kortlander, Job Spierings, Tom Demeyer.** *D5.5 Deployment of Pilots in Amsterdam.* s.l. : DECODE Project, March 2019.
23. **Shehar Bano, Eleonora Bassi, Marco Ciurcina, Ana Freire, Sara Hajian, Jaap-Henk Hoepman.** *D1.2 Privacy Design Strategies for the DECODE Architecture.* s.l. : DECODE project, June 2017.
24. **Mulube, Sam.** *D3.9 IoT privacy-enhancing data sharing: integration with Pilot Infrastructures.* s.l. : DECODE Project, May 2019.
25. **Denis Roio, Puria Nafisi Azizi, Andrea D'Intino.** *D3.10 Implementation of Blockchain platform and ABC in DECODE pilots.* s.l. : DECODE Project, September 2019.
26. **Federico Bonelli, Taco van Dijk, Guy Samuel.** *D4.9 Design & implementation interface for smart rules.* s.l. : DECODE Project, December 2018.