Distributed Ledger Technologies and Personal Data Protection

Mirko Zichichi

Supervisors:

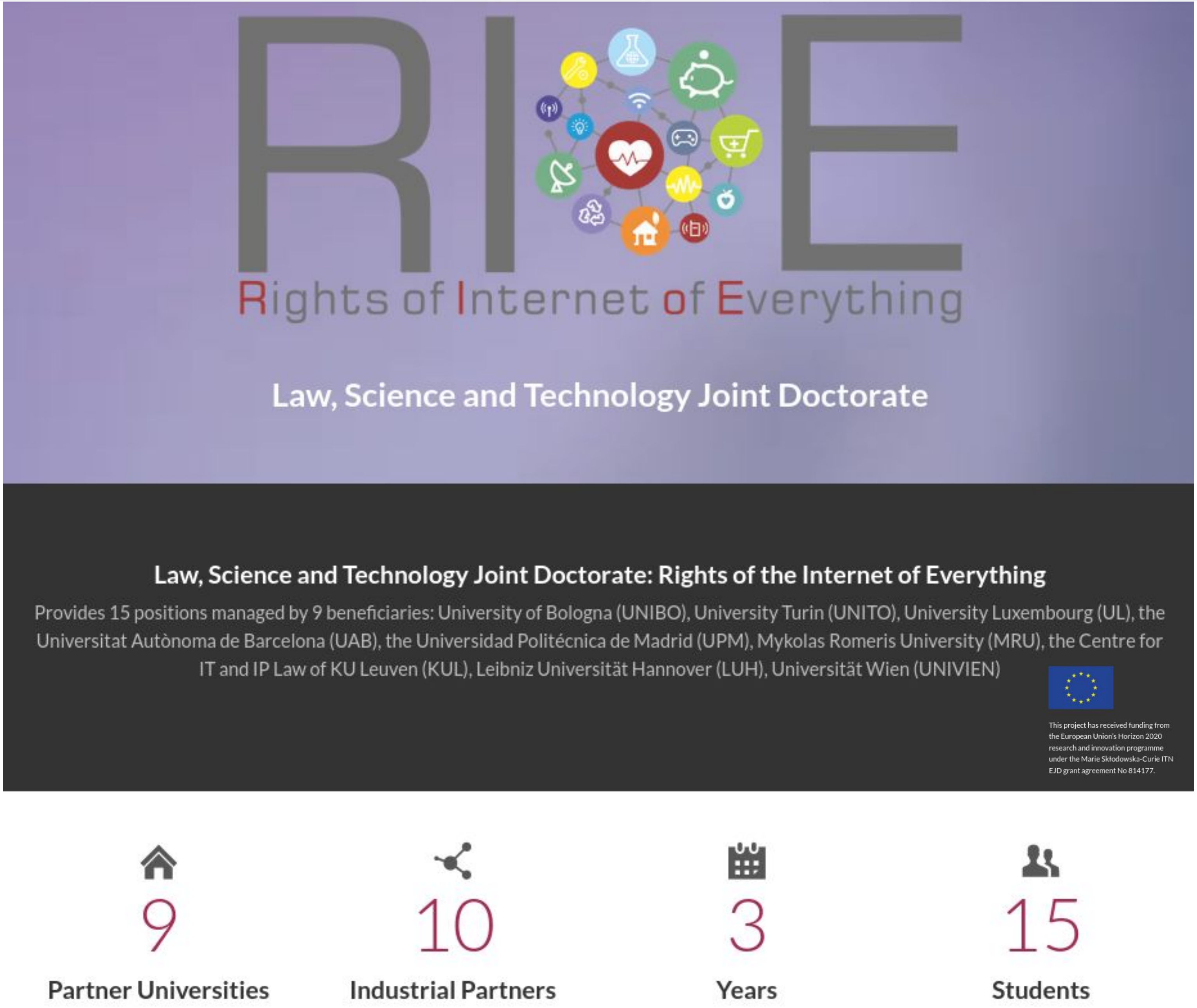
- **Víctor Rodríguez-Doncel**, UPM
- **Stefano Ferretti**, University of Bologna and University of Urbino, Italy

Mentor:
- **Massimo Durante**, University of Turin, Italy

https://mirkozichichi.me
✉ mirko.zichichi@upm.es

📅 03/06/2021

Law, Science and Technology Joint Doctorate



RIoE

Rights of Internet of Everything

Law, Science and Technology Joint Doctorate

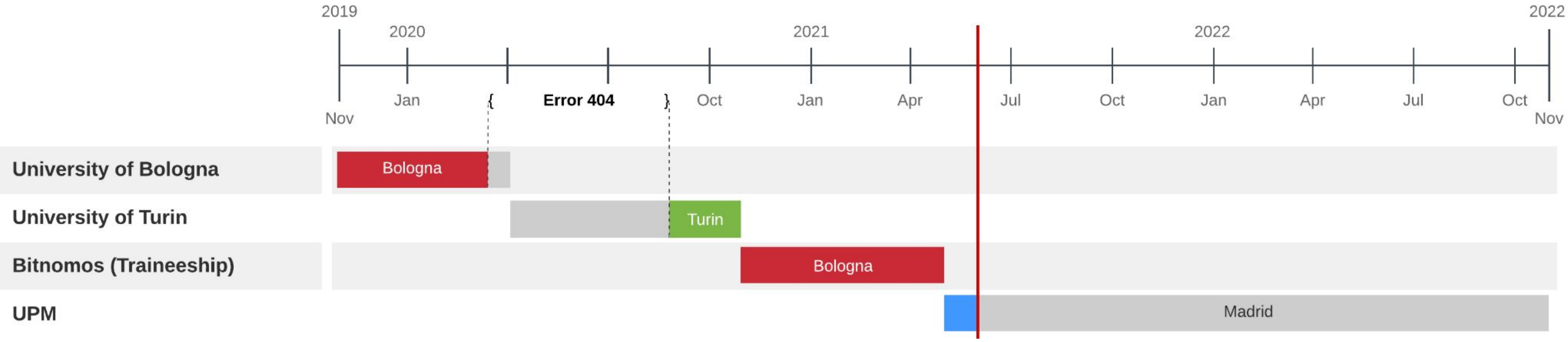
Law, Science and Technology Joint Doctorate: Rights of the Internet of Everything

Provides 15 positions managed by 9 beneficiaries: University of Bologna (UNIBO), University Turin (UNITO), University Luxembourg (UL), the Universitat Autònoma de Barcelona (UAB), the Universidad Politécnica de Madrid (UPM), Mykolas Romeris University (MRU), the Centre for IT and IP Law of KU Leuven (KUL), Leibniz Universität Hannover (LUH), Universität Wien (UNIVIEN)

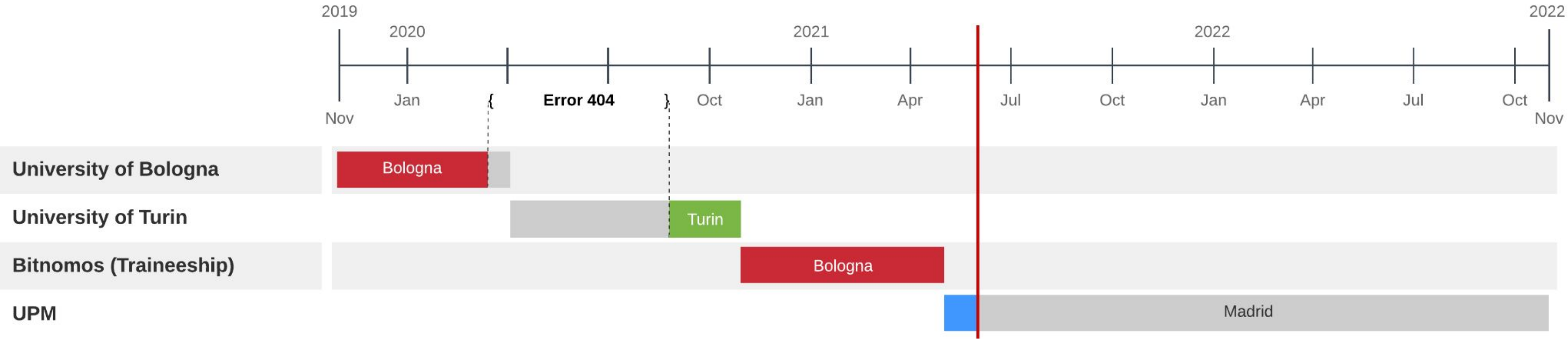
This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie ITN EJD grant agreement No 814177.

9	10	3	15
Partner Universities	Industrial Partners	Years	Students

Secondments and Labs



Secondments and Labs



Labs

- Analysis Of Networks And Simulation (AnaNSi) Research Group
- Legal Blockchain Lab
- BitNomos (traineeship)
- OEG

Background

Distributed Ledger Technologies (DLT)
and
Personal Data Protection

Why?

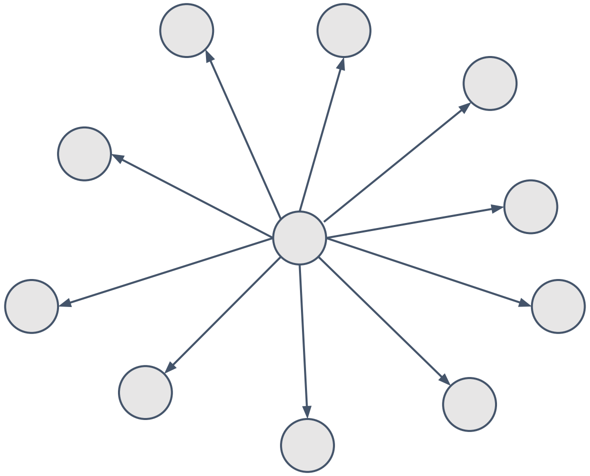
Background

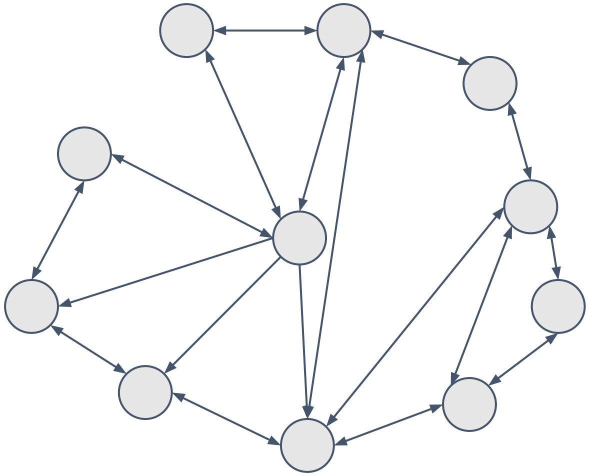
Distributed Ledger Technologies (DLT)

~~and~~

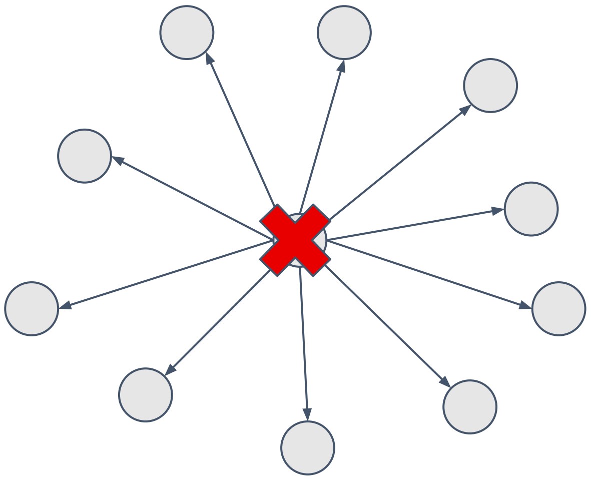
~~Personal Data Protection~~

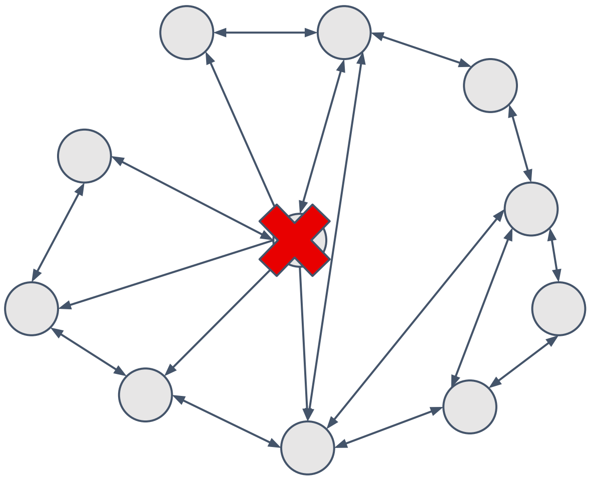
Why?



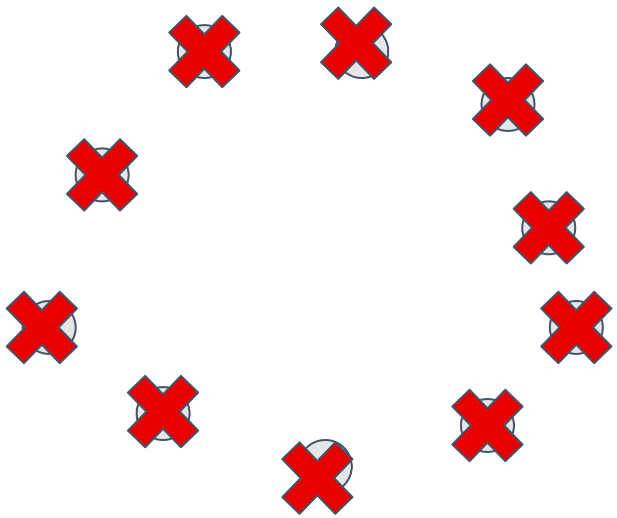


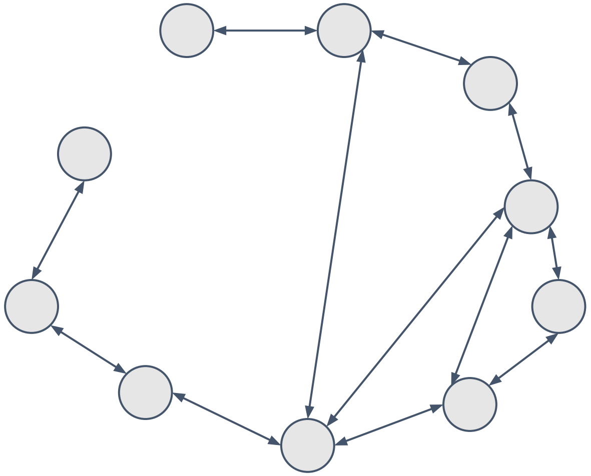
Single Point of Failure





Single Point of Failure





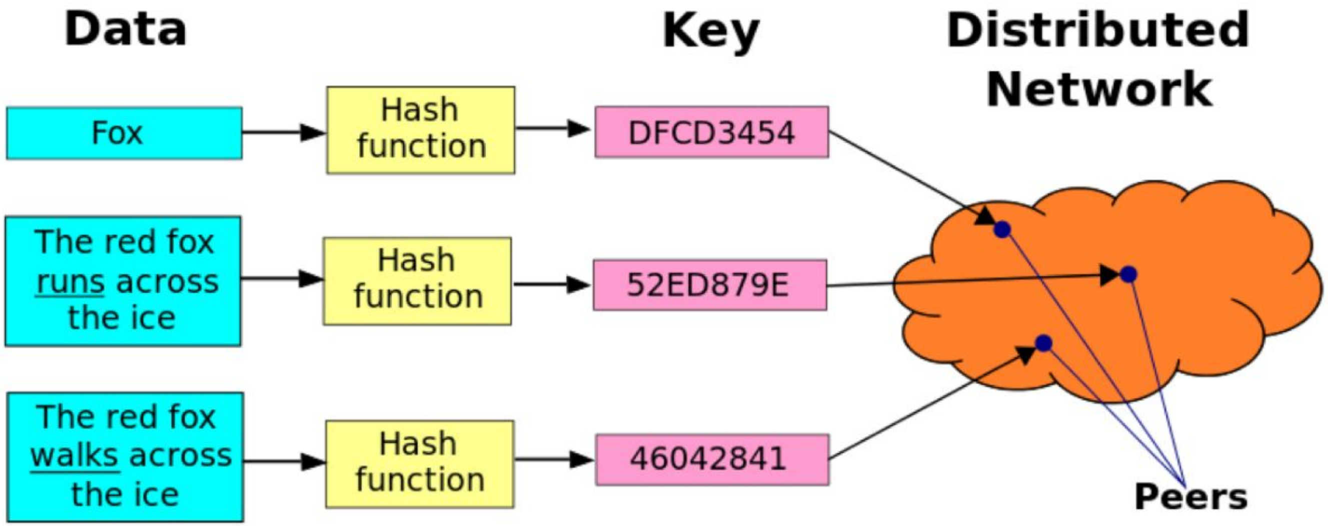
Failure?

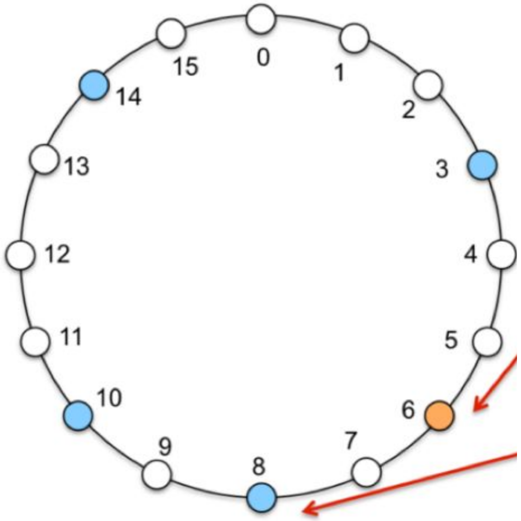
- 2016 Dyn cyberattack (Airbnb, Amazon, BBC, CNN, eBay, Netflix and Twitter affected)
- *2017 British Airways failure: 'accidentally switched off an uninterruptable power supply at a key data centre'*
- Amazon Web Services, Google Cloud, ...

https://www.techradar.com/news/5-of-the-worlds-biggest-network-outages

Distributed Hash Tables

Distributed Hash Tables where used for information sharing already **20 years ago**

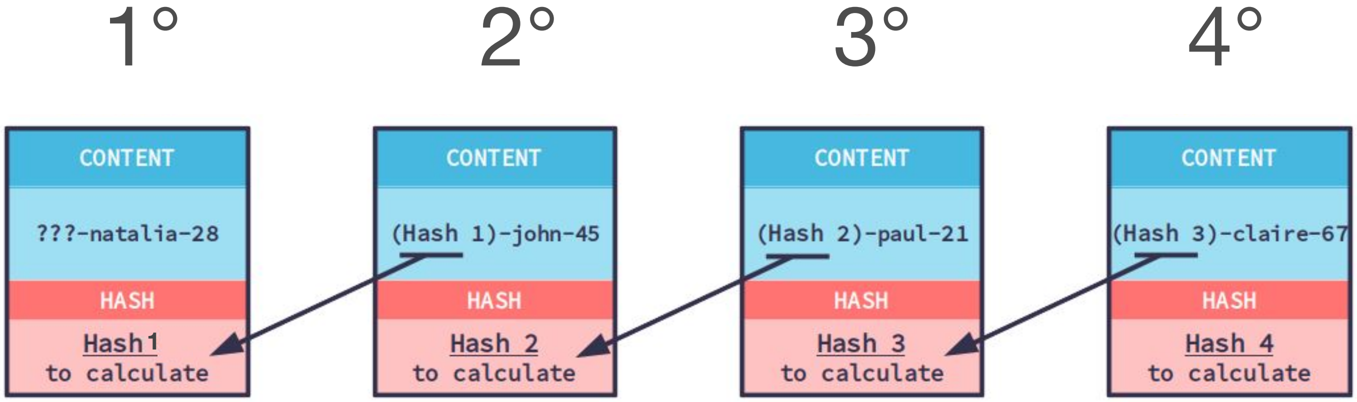




Distributed Ledger Technologies

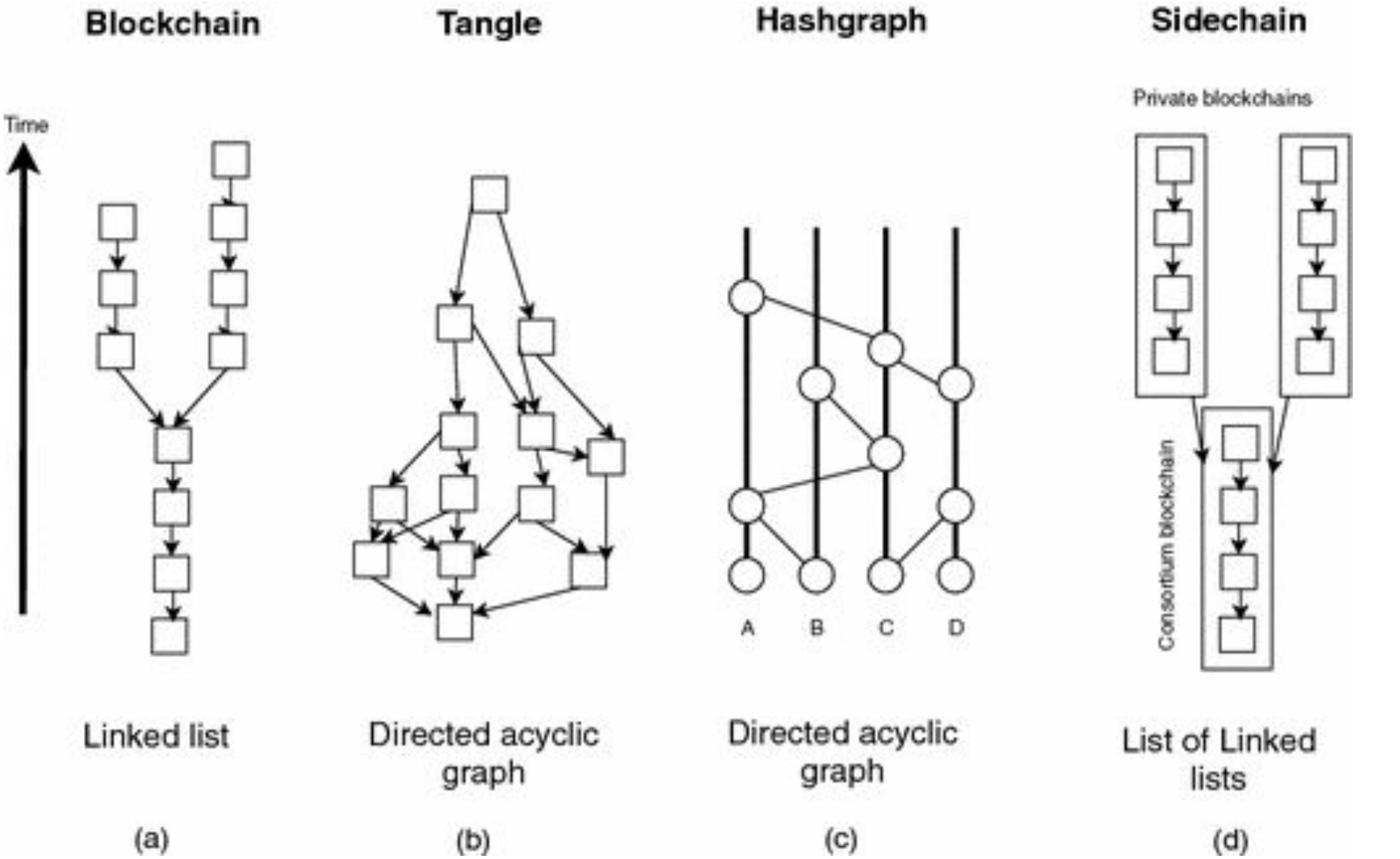
DLTs add “only”:
- persistence
- immutability
- **order**

DLTs add “only”:
- persistence
- immutability
- **order**



confident decentralized execution, that *sometimes* becomes “trustless”

Distributed Ledger Technologies



El Ioini, Nabil, and Claus Pahl. "A review of distributed ledger technologies." OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, Cham, 2018.

Distributed Ledger Technologies - Use Cases

Cryptocurrencies

Distributed Ledger Technologies - Use Cases

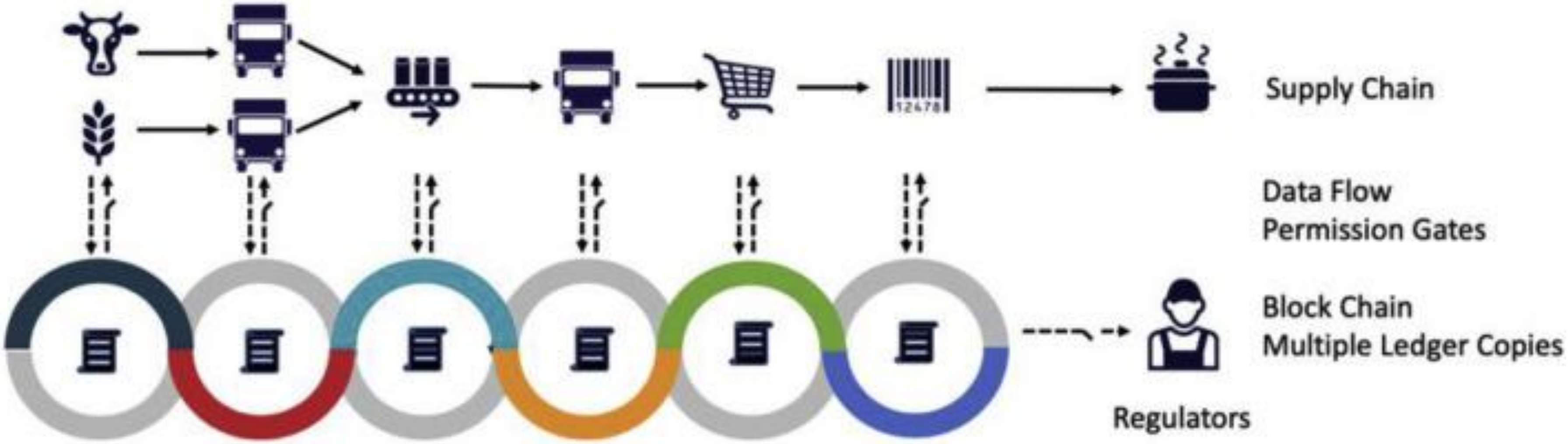
Cryptocurrencies

Distributed Ledger Technologies - Use Cases

Cryptocurrencies

Distributed Ledger Technologies - Use Cases

Traceability



Pearson, Simon, et al. "Are Distributed Ledger Technologies the panacea for food traceability?." Global food security 20 (2019)

Use Case - Personal Data Protection

Distributed Ledger Technologies (DLT)
and
Personal Data Protection

Distributed Ledger Technologies Features

DLTs enable the exchange of data or assets without the need to rely on a human intermediary:

- **transparency**, i.e. the guarantee for the auditability of transactions and data accesses
- **security**, i.e. the shifting of the trust that is normally placed to intermediaries (ISPs) towards a distributed consensus mechanism
- **decentralization**, i.e. the ability of direct user-to-user interactions and agreements, without intermediaries

(single point of failure comes up again)

Distributed Ledger Technologies Features

DLTs enable the exchange of data or assets without the need to rely on a human intermediary:

- **Smart Contracts**:
  An immutable set of instructions whose execution is computed deterministically by all peers in the DLT network. Each node executing the instructions receives the same inputs and produces the same outputs, thanks to a shared protocol, and act as verifier.
  It does not embody the same features of a legal contract.

Moving Data Sovereignty Towards Users

1. Systems that store and transfer personal data in a **transparent and non-centralized** manner

2. Support the right of individuals to the protection of personal data while at the same favoring **portability, social good and economic exploitation**

3. Represent and reason with policies in a **distributed execution to govern the access** to personal data

Moving Data Sovereignty Towards Users

1. Systems that store and transfer personal data in a **transparent and non-centralized** manner

Consumers privacy paradox for data disclosure
- *attitude*: profess their need for privacy (general)
- *behavior*: remain user of the tech that track and share their data (contextual)

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007).The privacy paradox: Personal information disclosure intentions versus behaviors.
Laufer, R. S. and Wolfe, M. (1977).Privacy as a concept and a social issue: A multidimensional developmental theory.
Acquisti, A., Taylor, C., and Wagman, L. (2016).The economics of privacy

Moving Data Sovereignty Towards Users

1. Systems that store and transfer personal data in a **transparent and non-centralized** manner

Consumers privacy paradox for data disclosure
- *attitude*: profess their need for privacy (general)
- *behavior*: remain user of the tech that track and share their data (contextual)

“Privacy Calculus”:
perceived value of disclosure utility= *privacy risk / benefits*

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007).The privacy paradox: Personal information disclosure intentions versus behaviors.
Laufer, R. S. and Wolfe, M. (1977).Privacy as a concept and a social issue: A multidimensional developmental theory.
Acquisti, A., Taylor, C., and Wagman, L. (2016).The economics of privacy

Moving Data Sovereignty Towards Users

1. Systems that store and transfer personal data in a **transparent and non-centralized** manner

Consumers privacy paradox for data disclosure
- *attitude*: profess their need for privacy (general)
- *behavior*: remain user of the tech that track and share their data (contextual)

“Privacy Calculus”:
perceived value of disclosure utility= *privacy risk* / *benefits*

Correct estimation?
undermined by asymmetric information or unawareness of possible alternative solutions.

Norberg, P. A., Horne, D. R., and Horne, D. A. (2007).The privacy paradox: Personal information disclosure intentions versus behaviors.
Laufer, R. S. and Wolfe, M. (1977).Privacy as a concept and a social issue: A multidimensional developmental theory.
Acquisti, A., Taylor, C., and Wagman, L. (2016).The economics of privacy

Moving Data Sovereignty Towards Users

2. Support the right of individuals to the protection of personal data while at the same favoring **portability, social good and economic exploitation**

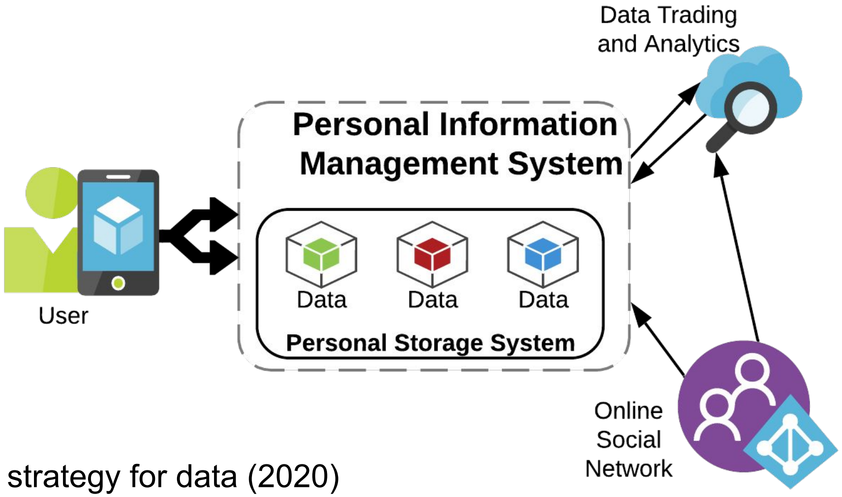
Internet of Persons (IoP) is emerging as a paradigm that places individuals and their personal devices at the heart of the data management design.

European Commission. A european strategy for data (2020)
Isabelle, L., Pelics, G., Binctin, N., and Pez-Pérard, V. (2018).
My data are mine: Why we should have ownership rights on our personal data.
Bock, S. (2018).My data is mine-users' handling of personal data in everyday life.SICHERHEIT 2018.

Moving Data Sovereignty Towards Users

2. Support the right of individuals to the protection of personal data while at the same favoring **portability, social good and economic exploitation**

Internet of Persons (IoP) is emerging as a paradigm that places individuals and their personal devices at the heart of the data management design. There is the need for **Personal Information Management Systems (PIMS)**


Data Trading and Analytics
Personal Information Management System
User
Data
Data
Data
Personal Storage System
Online Social Network

European Commission. A european strategy for data (2020)
Isabelle, L., Pelics, G., Binctin, N., and Pez-Pérard, V. (2018).
My data are mine: Why we should have ownership rights on our personal data.
Bock, S. (2018).My data is mine-users' handling of personal data in everyday life.SICHERHEIT 2018.

Moving Data Sovereignty Towards Users

3. Represent and reason with policies in a **distributed execution to govern the access** to personal data

The use of smart contracts may be crucial to regulate user's data flow **automatically**, since these can:

- enable the user to set data access rules in order to disclose data
- determine if a policy satisfies the legal requirements
- determine if a data request can be satisfied according to the individual's preferences

Cervone, Luca, Monica Palmirani, and Fabio Vitali. "The Intelligible Contract." HICSS. 2020.

Dynamic Personal Data

Dynamic Personal Data

‘dynamic data’ means documents in a digital form, subject to frequent or real-time updates, in particular because of their volatility or rapid obsolescence; data generated by sensors are typically considered to be dynamic data;

DIRECTIVE (EU) 2019/1024 on open data and the re-use of public sector information

Location Data

Location data (or geodata):

- is substantially different from the rest of personal data (Thesis 1)

Keßler, Carsten, and Grant McKenzie. "**A geoprivacy manifesto**." Transactions in GIS 22.1 (2018): 3-19.

Location Data

Location data (or geodata):

- is substantially different from the rest of personal data (Thesis 1)
- is easy to capture thanks to mobile devices and easy-to use APIs (Thesis 2)

Keßler, Carsten, and Grant McKenzie. "**A geoprivacy manifesto**." Transactions in GIS 22.1 (2018): 3-19.

Location Data

Location data (or geodata):

- is substantially different from the rest of personal data (Thesis 1)
- is easy to capture thanks to mobile devices and easy-to use APIs (Thesis 2)
- is more **useful when shared**, since it can improve the quality of services (Thesis 3)

Keßler, Carsten, and Grant McKenzie. "**A geoprivacy manifesto**." Transactions in GIS 22.1 (2018): 3-19.

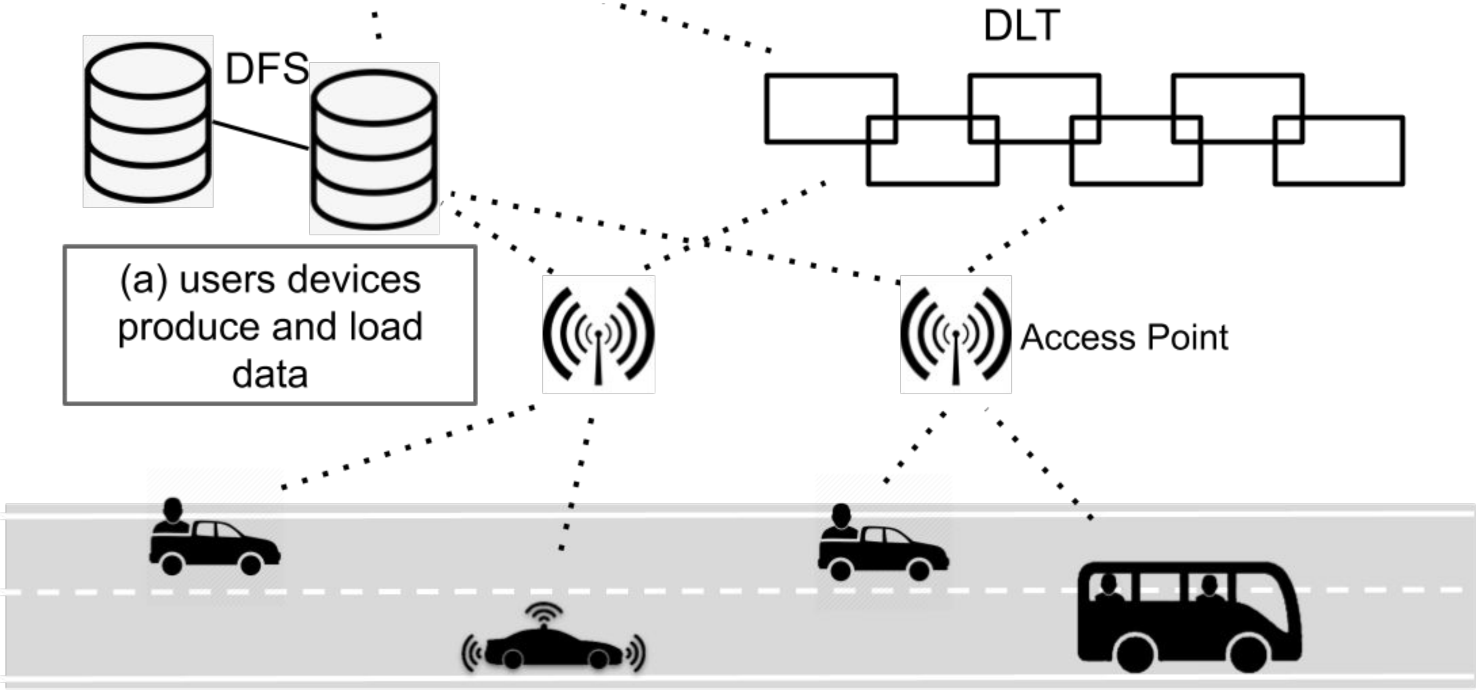
Location Data

Location data (or geodata):

- is substantially different from the rest of personal data (Thesis 1)
- is easy to capture thanks to mobile devices and easy-to use APIs (Thesis 2)
- is more **useful when shared**, since it can improve the quality of services (Thesis 3)
- allows to infer individuals activities (Thesis 5) that they never intended or agreed to share with a service (Thesis 6)

Keßler, Carsten, and Grant McKenzie. "**A geoprivacy manifesto**." Transactions in GIS 22.1 (2018): 3-19.

Scenario: Intelligent Transportation Systems



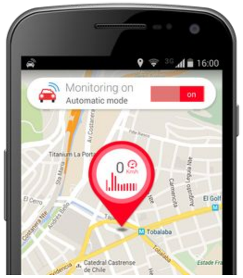
DFS

DLT

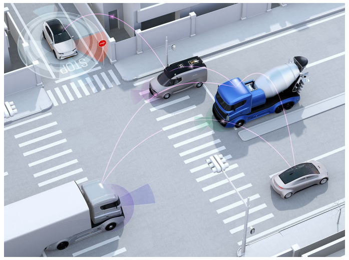
(a) users devices produce and load data

Access Point

crash alert



intersection safety



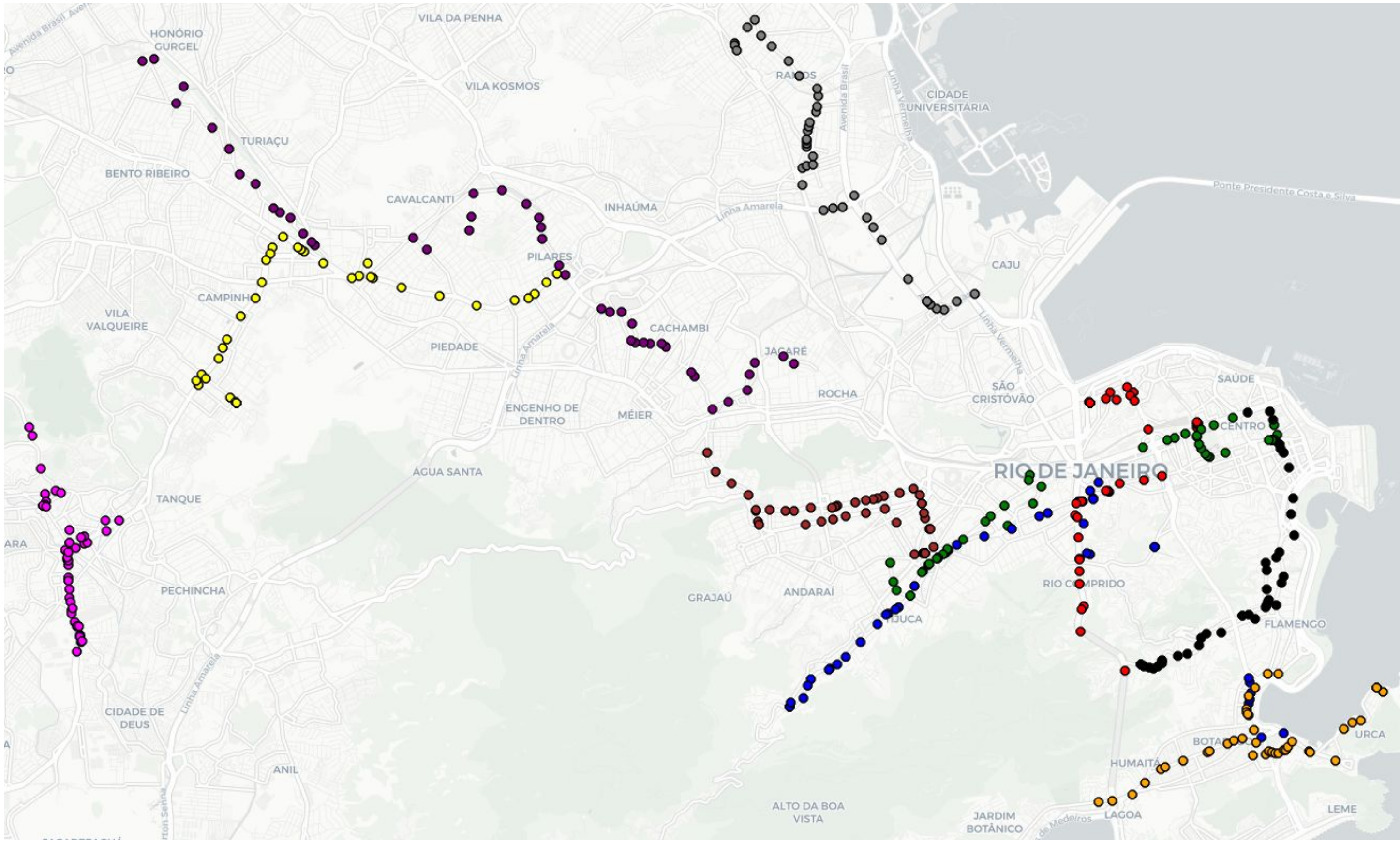
combat wrong way driving

road weather condition alerts

Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "A distributed ledger based infrastructure for smart transportation system and social good." 2020 IEEE 17th Annual Consumer Communications & Networking Conference (CCNC). IEEE, 2020.

Scenario: Intelligent Transportation Systems

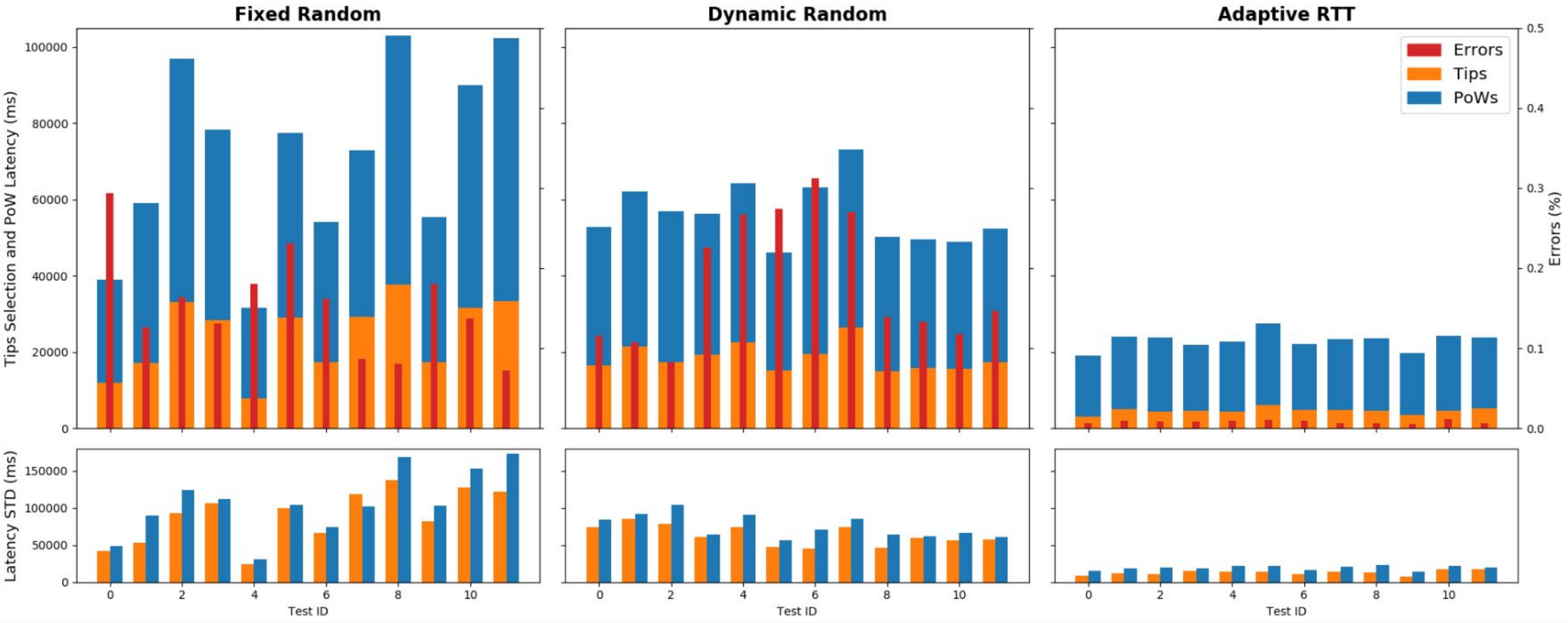
Dataset of real mobility traces of buses in Rio de Janeiro (Brasil)



We simulated up to 240 users issuing (geolocation) data to the **IOTA DLT** in real time

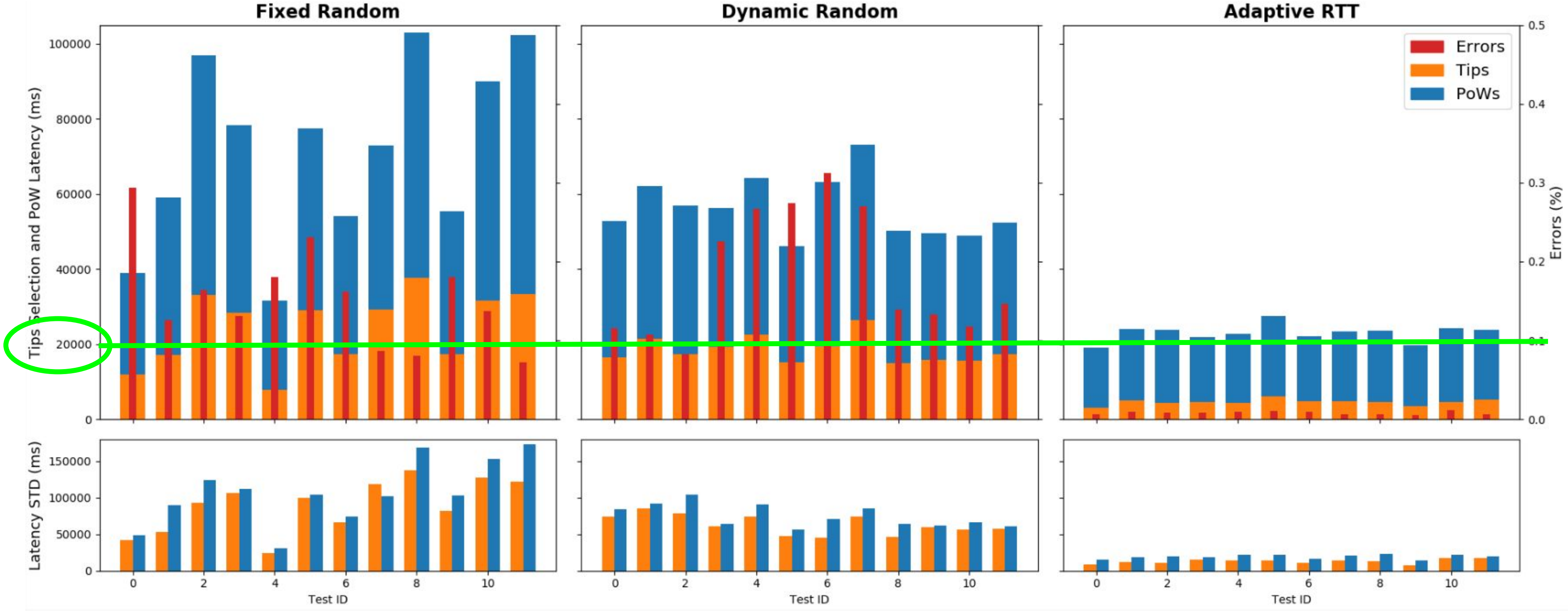
Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "Are distributed ledger technologies ready for intelligent transportation systems?." Proceedings of the 3rd Workshop on Cryptocurrencies and Blockchains for Distributed Systems. 2020.

Scenario: Intelligent Transportation Systems



Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "Are distributed ledger technologies ready for intelligent transportation systems?." Proceedings of the 3rd Workshop on Cryptocurrencies and Blockchains for Distributed Systems. 2020.

Scenario: Intelligent Transportation Systems



Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "Are distributed ledger technologies ready for intelligent transportation systems?." Proceedings of the 3rd Workshop on Cryptocurrencies and Blockchains for Distributed Systems. 2020.

Scenario: Intelligent Transportation Systems

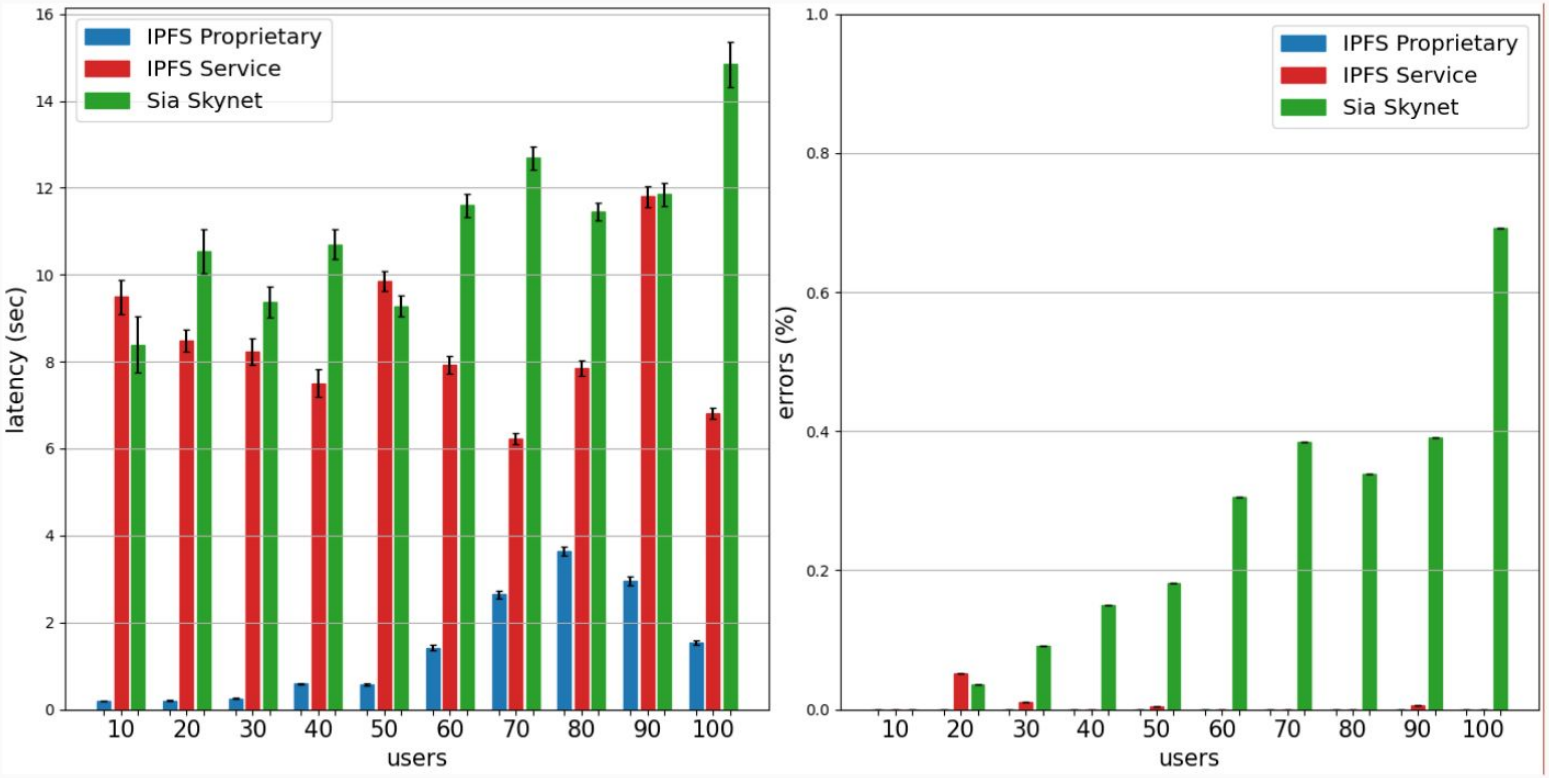


Figure 1: Latencies and errors sending geolocation. Black line → confidence interval (95%)

Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "On the efficiency of decentralized file storage for personal information management systems." 2020 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2020.

Scenario: Intelligent Transportation Systems

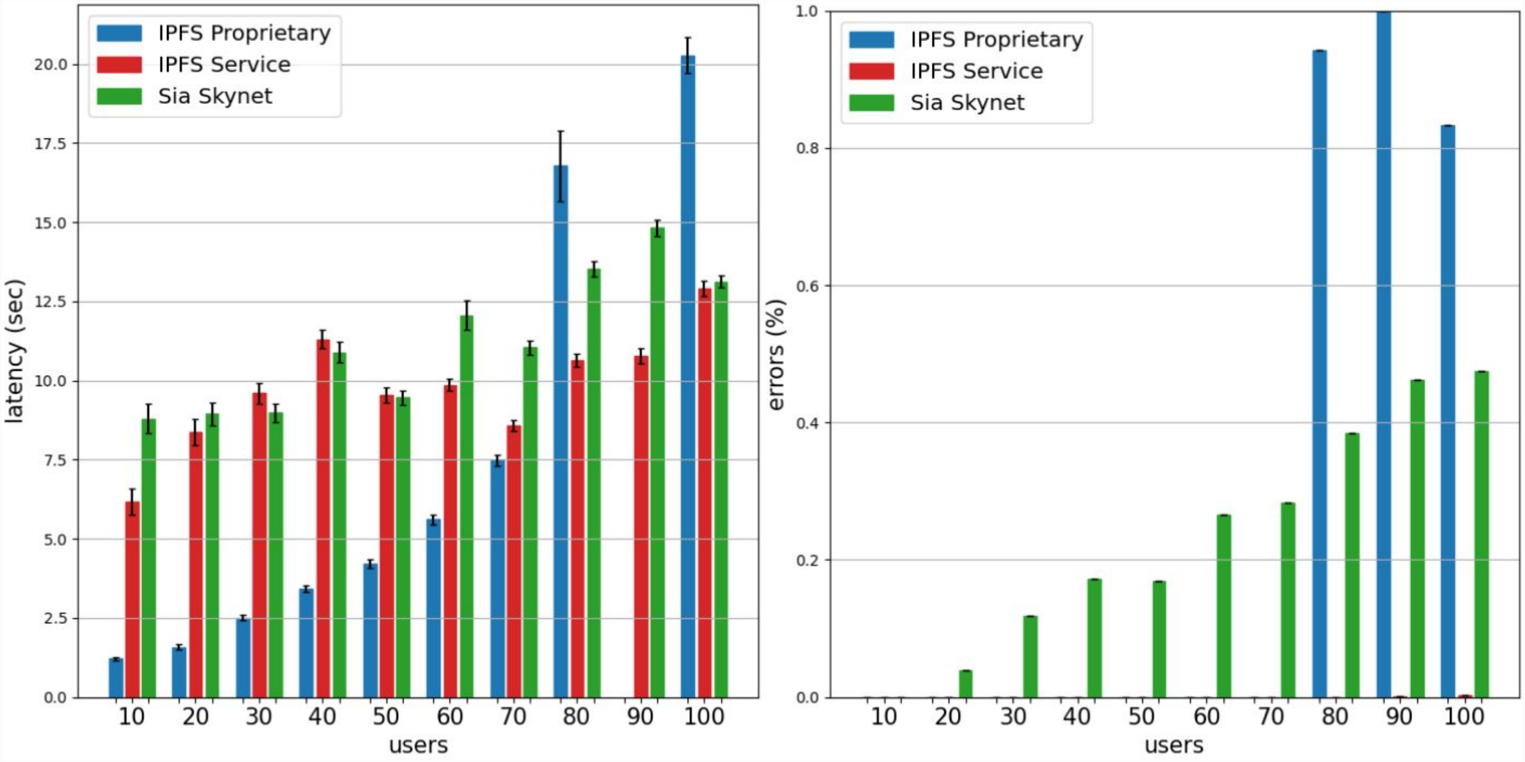


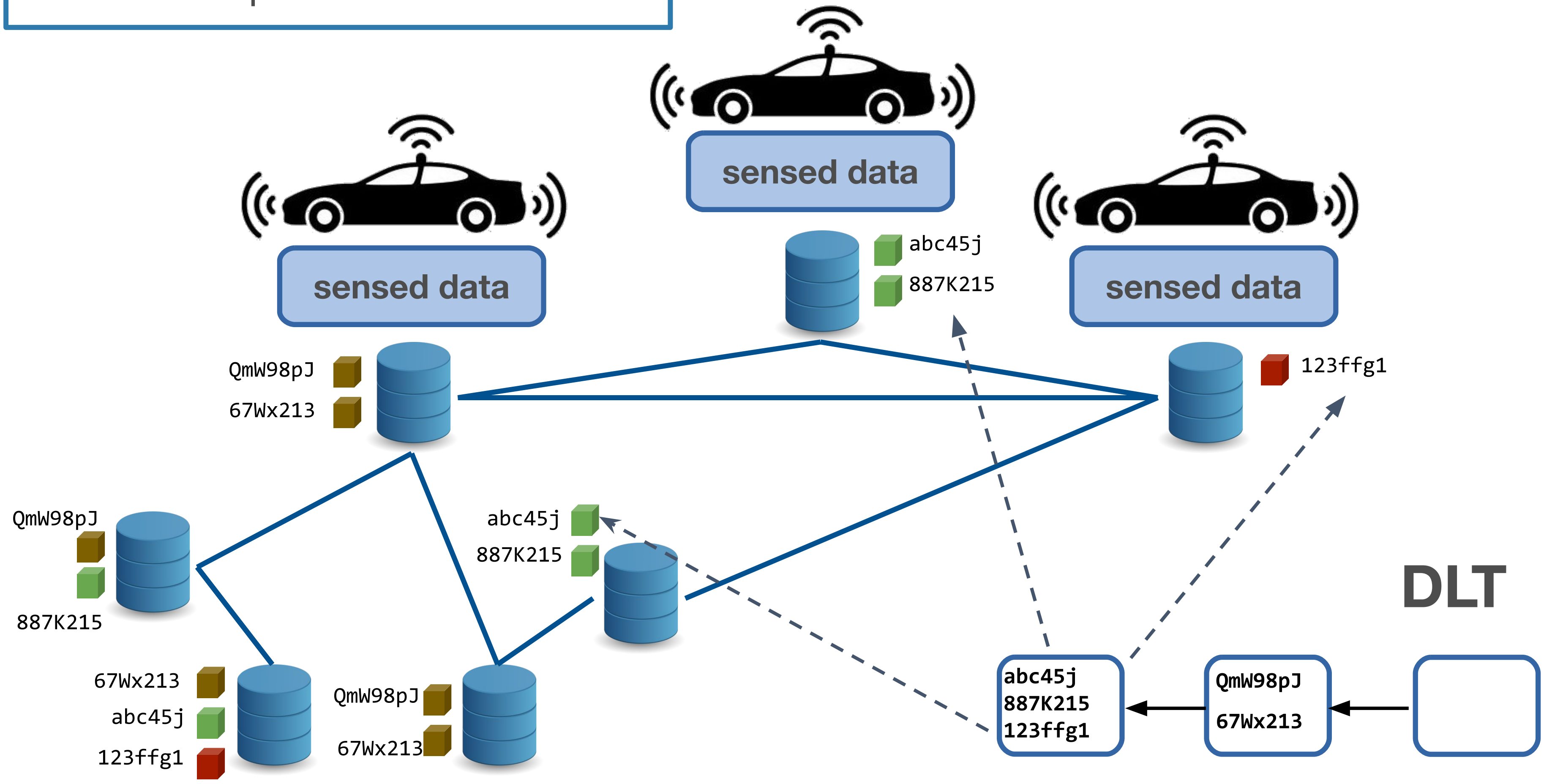
Figure 2: Latencies and errors sending photos (1 MB). Black line → confidence interval (95%)

Zichichi, Mirko, Stefano Ferretti, and Gabriele D'Angelo. "On the efficiency of decentralized file storage for personal information management systems." 2020 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2020.

A framework based on DLT for data management and services in ITS

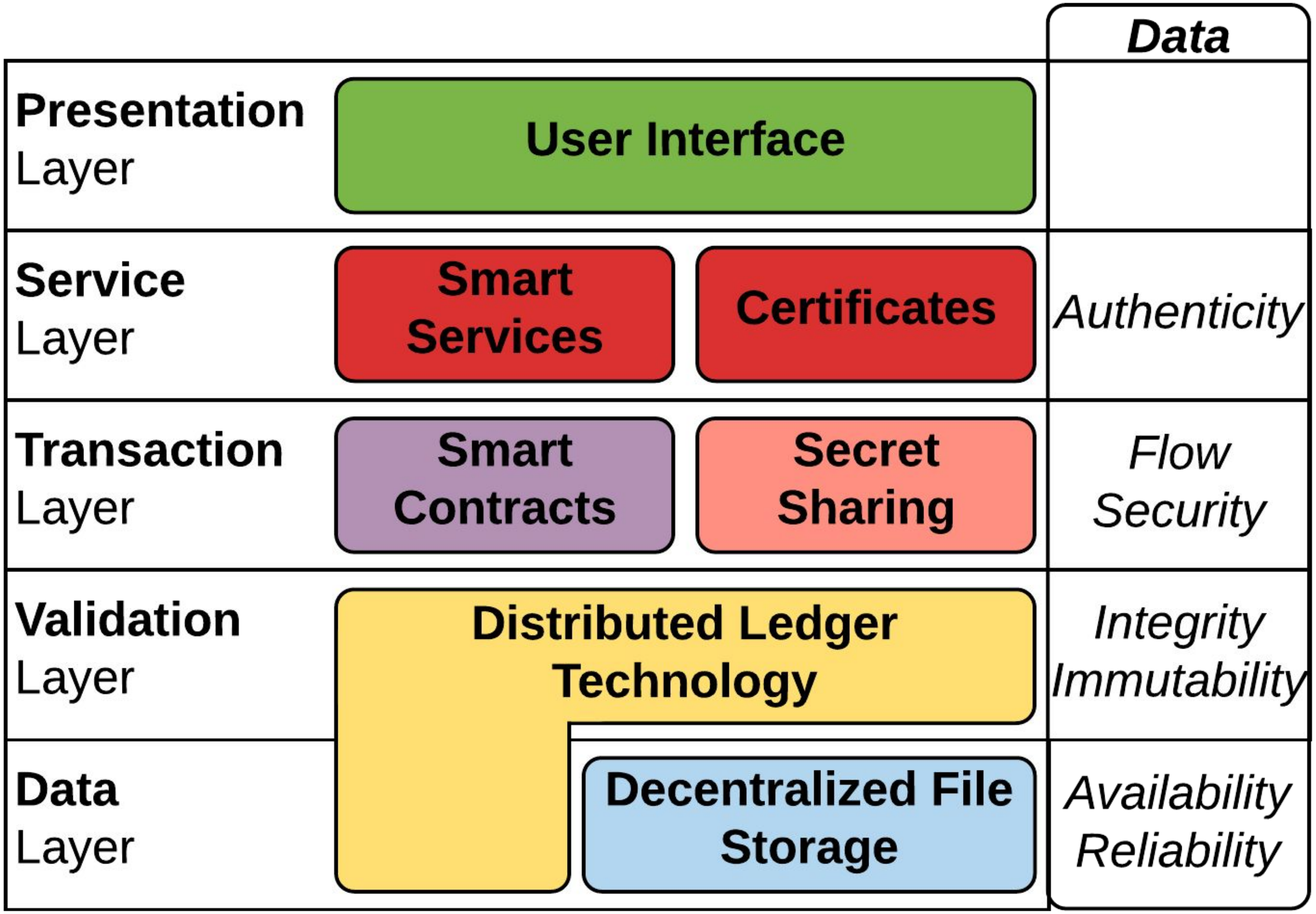
Content based addressing
(instead of location based)

Use of hash pointers → DLTs



Zichichi, Mirko, Stefano Ferretti, and Gabriele D’angelo. "A framework based on distributed ledger technologies for data management and services in intelligent transportation systems." IEEE Access 8 (2020): 100384-100402.

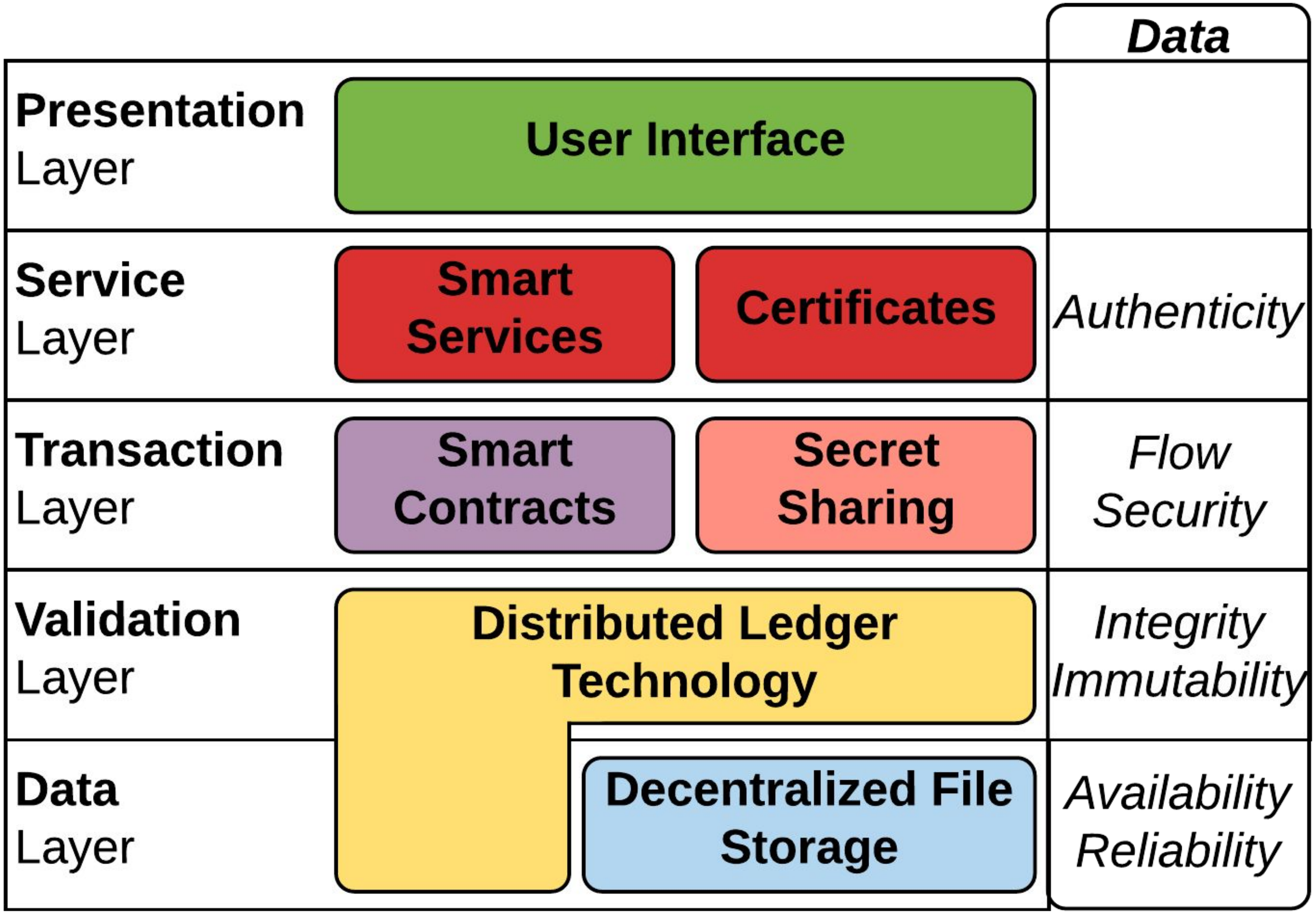
A framework based on DLT for data management and services in ITS



			Data
Presentation Layer	User Interface		
Service Layer	Smart Services	Certificates	*Authenticity*
Transaction Layer	Smart Contracts	Secret Sharing	*Flow Security*
Validation Layer	Distributed Ledger Technology		*Integrity Immutability*
Data Layer		Decentralized File Storage	*Availability Reliability*

Zichichi, Mirko, Stefano Ferretti, and Gabriele D’angelo. "A framework based on distributed ledger technologies for data management and services in intelligent transportation systems." IEEE Access 8 (2020): 100384-100402.

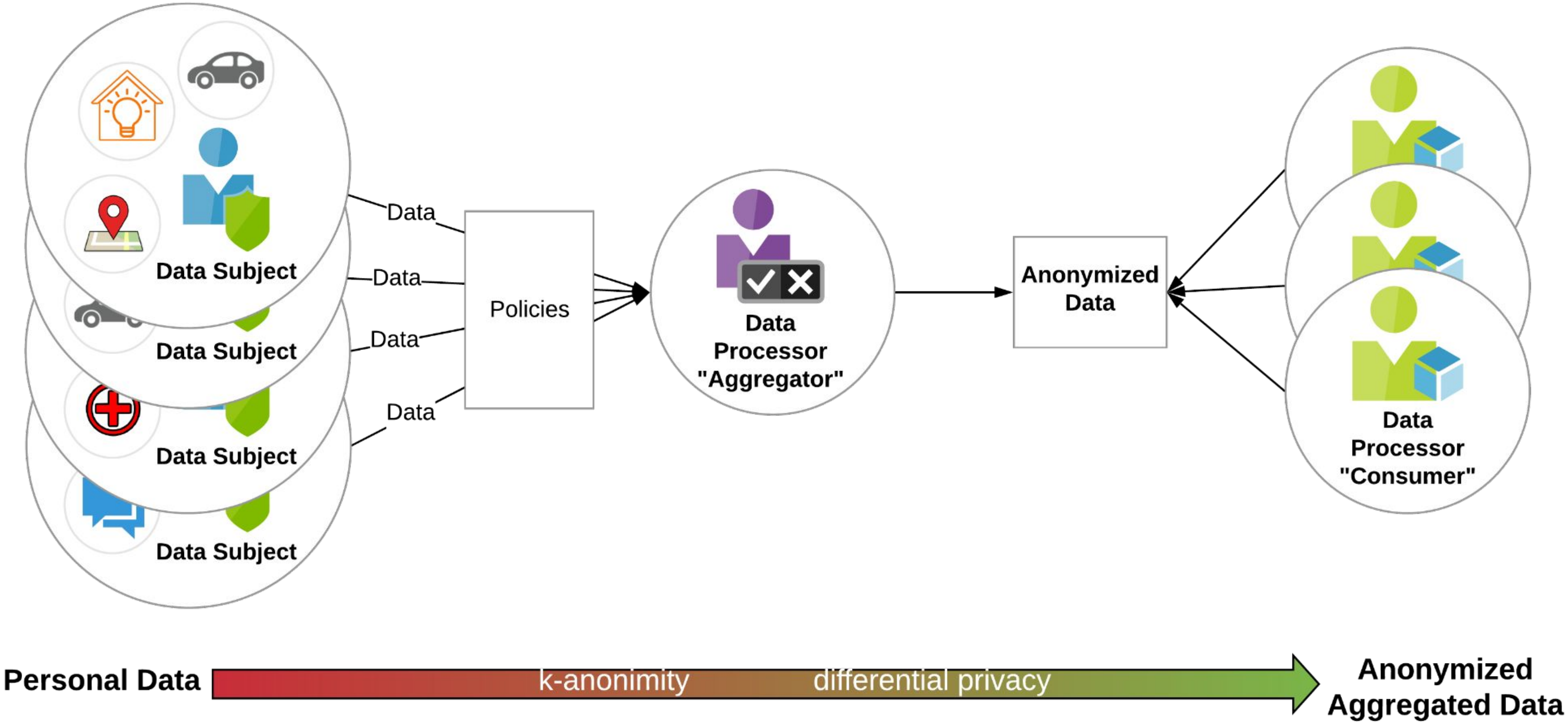
A framework based on DLT for data management and services in ITS



			Data
Presentation Layer	**User Interface**		
Service Layer	**Smart Services**	**Certificates**	*Authenticity*
Transaction Layer	**Smart Contracts**	**Secret Sharing**	*Flow Security*
Validation Layer	**Distributed Ledger Technology**		*Integrity Immutability*
Data Layer		**Decentralized File Storage**	*Availability Reliability*

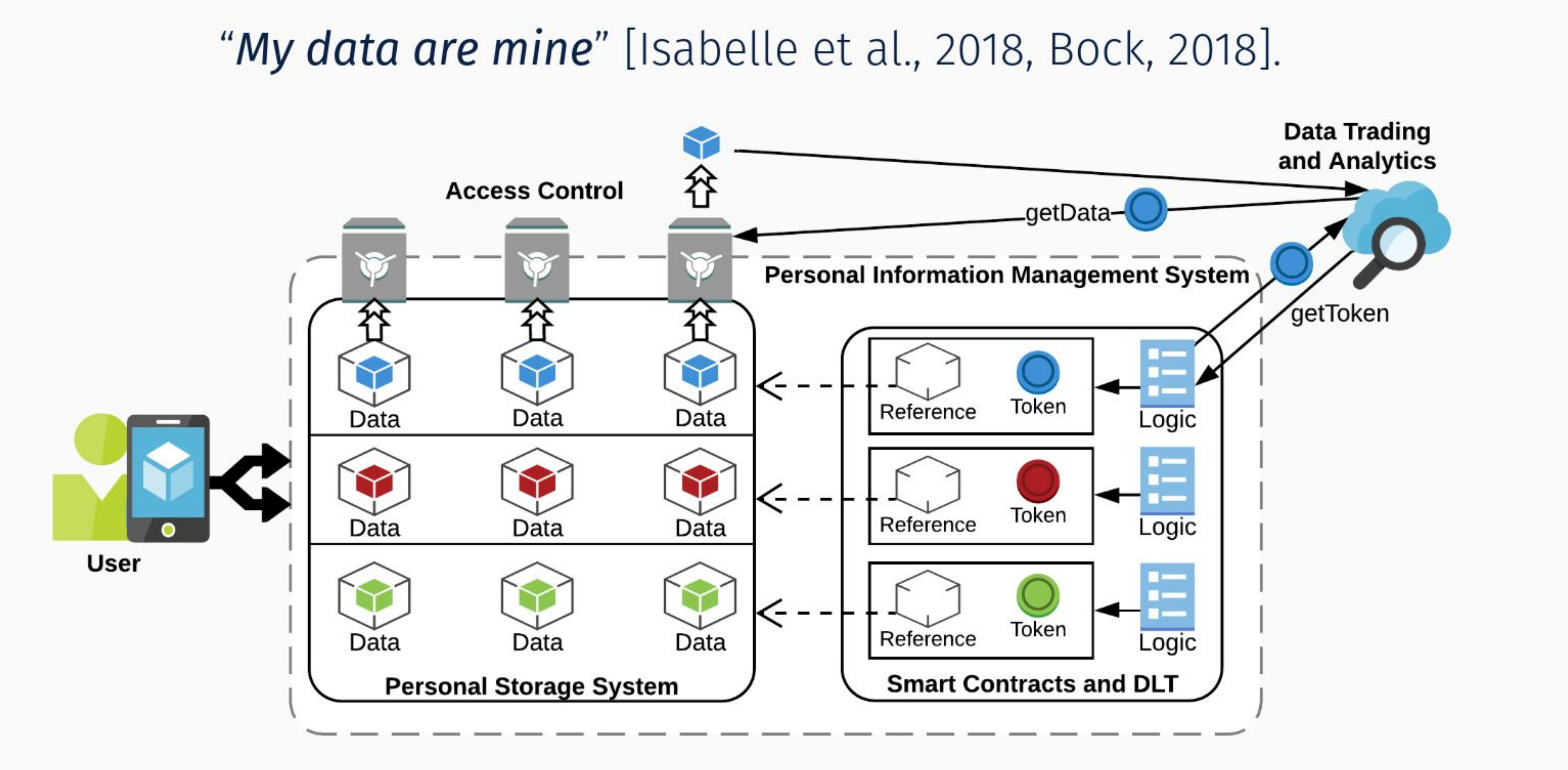
Zichichi, Mirko, Stefano Ferretti, and Gabriele D'angelo. "A framework based on distributed ledger technologies for data management and services in intelligent transportation systems." IEEE Access 8 (2020): 100384-100402.

Decentralized Personal Data Market



Zichichi Mirko, Contu Michele, Stefano Ferretti, and Rodríguez-Doncel Victor. "Ensuring personal data anonymity in data marketplaces through sensing-as-a-service and distributed ledger technologies." In 3rd Distributed Ledger Technology Workshop. ITASEC, 2020.

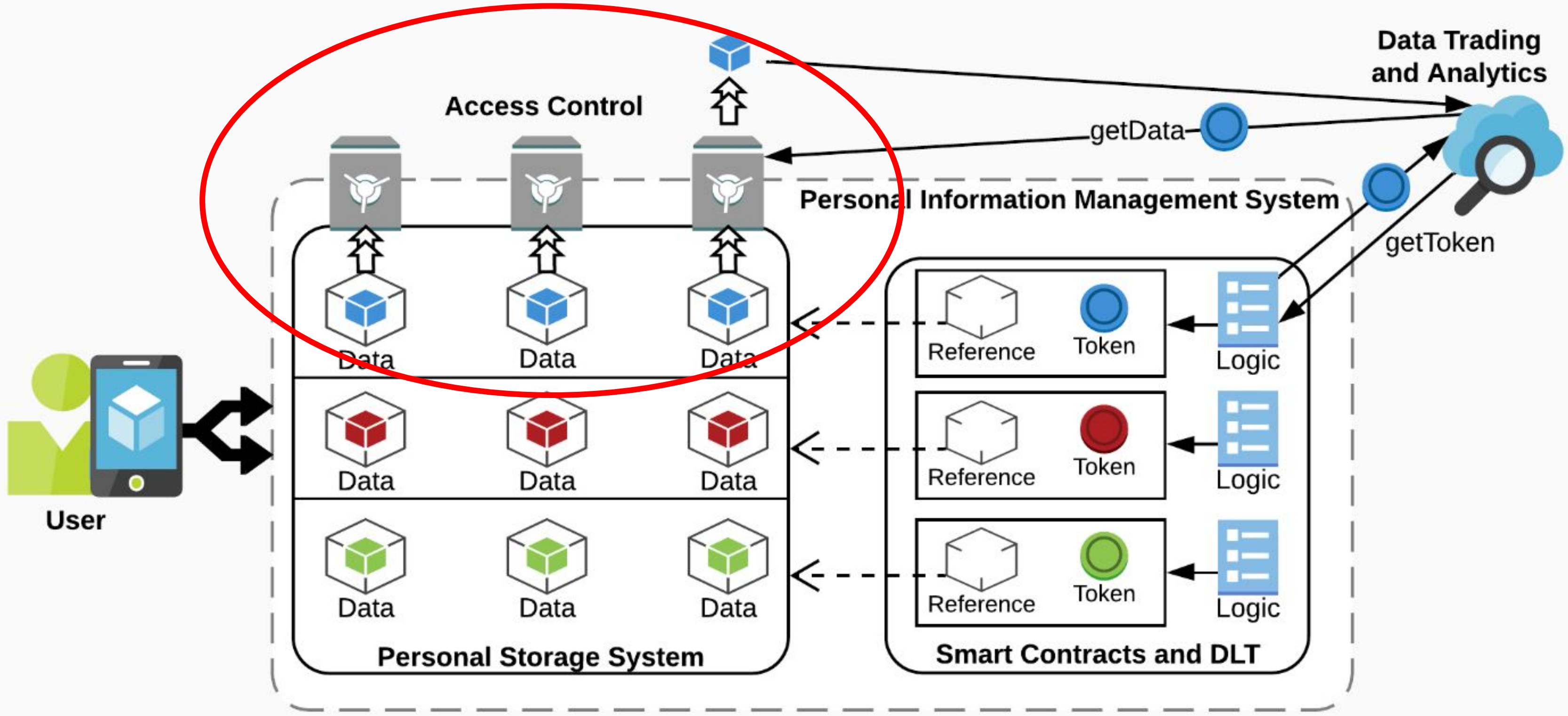
Decentralized Personal Data Market



Zichichi, M., Ferretti, S., D'Angelo, G., & Rodríguez-Doncel, V. (2020, November). Personal Data Access Control Through Distributed Authorization. In 2020 IEEE 19th International Symposium on Network Computing and Applications (NCA). IEEE.

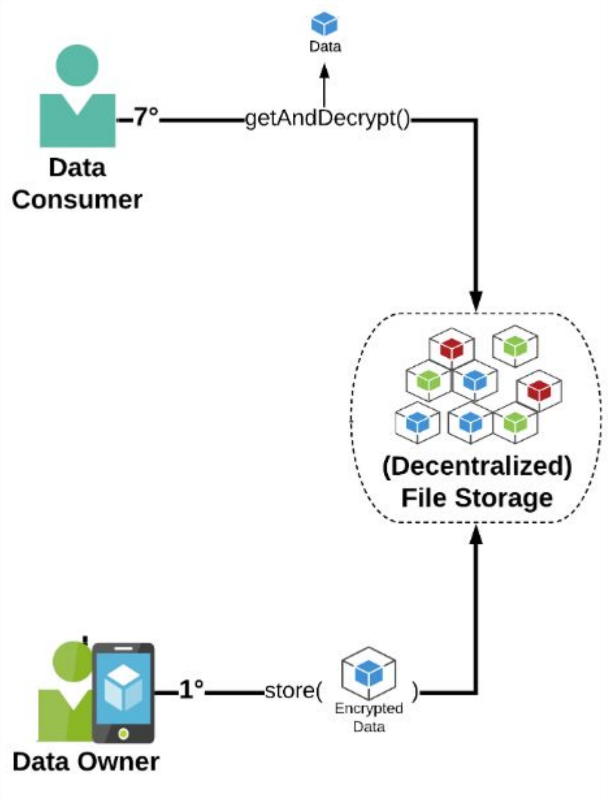
Decentralized Personal Data Market

“*My data are mine*” [Isabelle et al., 2018, Bock, 2018].



Zichichi, M., Ferretti, S., D'Angelo, G., & Rodríguez-Doncel, V. (2020, November). Personal Data Access Control Through Distributed Authorization. In 2020 IEEE 19th International Symposium on Network Computing and Applications (NCA). IEEE.

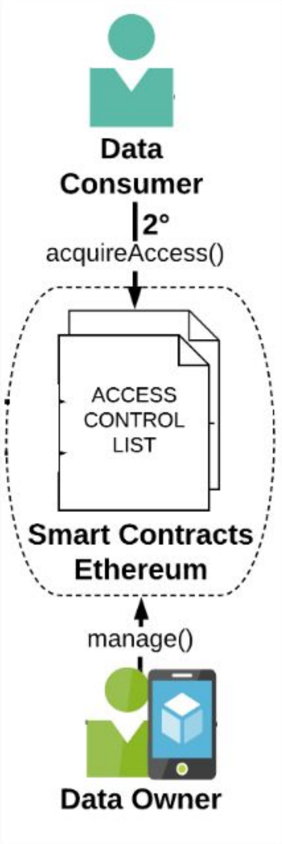
Decentralized Personal Data Market



- Personal data → *encrypted* + off-chain
- In a (possibly decentralized) **File Storage (FS)** and then referenced in a DLT
- **Digest** → verification of **integrity** + **pepper and salt**
- **GDPR** → requires the modification or deletion of data under certain circumstances

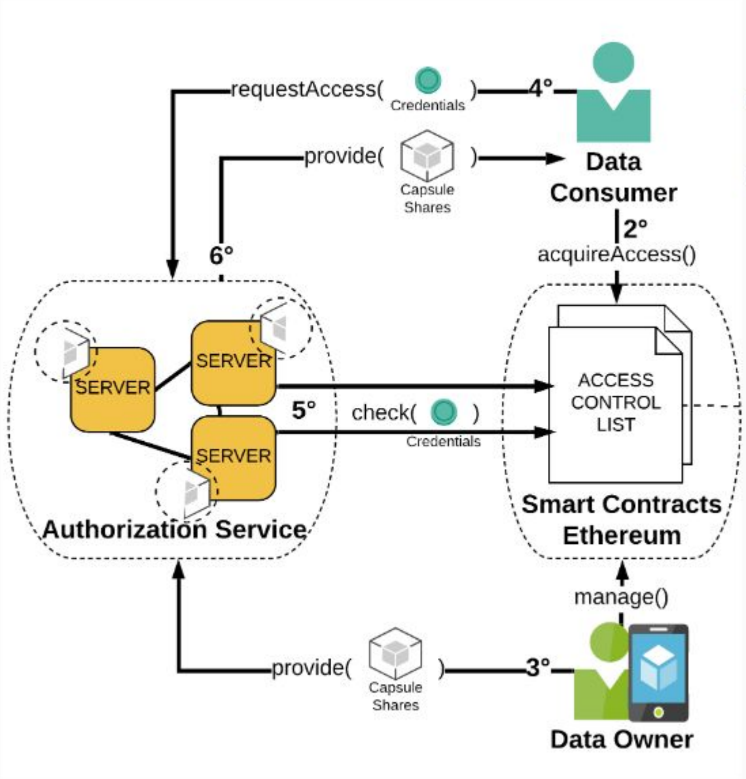
Zichichi, M., Ferretti, S., D'Angelo, G., & Rodríguez-Doncel, V. (2020, November). Personal Data Access Control Through Distributed Authorization. In 2020 IEEE 19th International Symposium on Network Computing and Applications (NCA). IEEE.

Decentralized Personal Data Market



- **Ethereum smart contracts** → access to the data enabled by the owner or purchased
- The smart contract maintains an **Access Control List (ACL)** that represents the rights to access a bundle of data.
- Once a **consumer is listed in the ACL**, he can access data through an access key, which is provided by the authorization service.

Zichichi, M., Ferretti, S., D'Angelo, G., & Rodríguez-Doncel, V. (2020, November). Personal Data Access Control Through Distributed Authorization. In 2020 IEEE 19th International Symposium on Network Computing and Applications (NCA). IEEE.



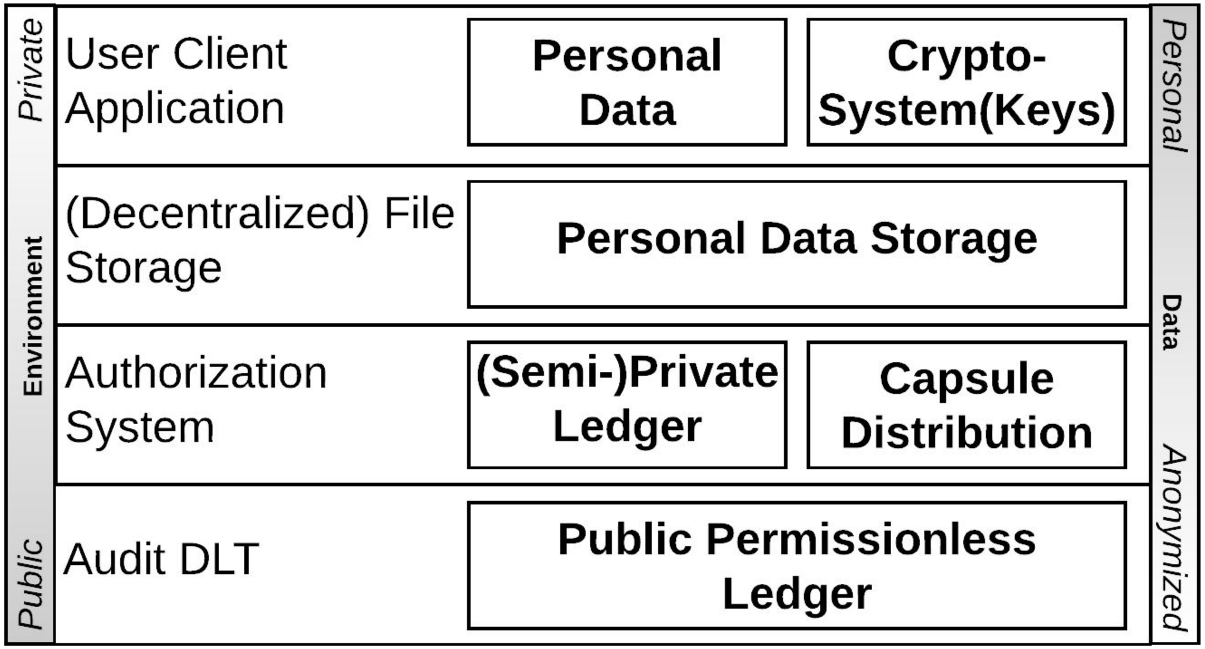
- Enforces the **access rights** specified in the ACLs
- A valid data consumer request enact the release of the "**capsule**" that holds the k_{DEM} secret key, needed to decrypt the desired data
  - **SS** - splits capsule in n shares, but only t shares $(t < n)$ are needed to "open" it
  - **Threshold PRE** - splits capsule in in n shares and uses a (t, n)-threshold scheme with t "re-encryption shares" to re-encrypt it for consumer

Secret Sharing (SS) & Threshold Proxy Re-Encryption (T-PRE)

Zichichi, M., Ferretti, S., D'Angelo, G., & Rodríguez-Doncel, V. (2020, November). Personal Data Access Control Through Distributed Authorization. In 2020 IEEE 19th International Symposium on Network Computing and Applications (NCA). IEEE.

Now?

1. **Data Governance Through a Multi-DLT Architecture in view of the GDPR**
Mirko Zichichi, Stefano Ferretti, Gabriele D’Angelo, Victor Rodriguez-Doncel
(journal paper being reviewed)



Private Environment Public

User Client Application

Personal Data

Crypto-System(Keys)

(Decentralized) File Storage

Personal Data Storage

Authorization System

(Semi-)Private Ledger

Capsule Distribution

Audit DLT

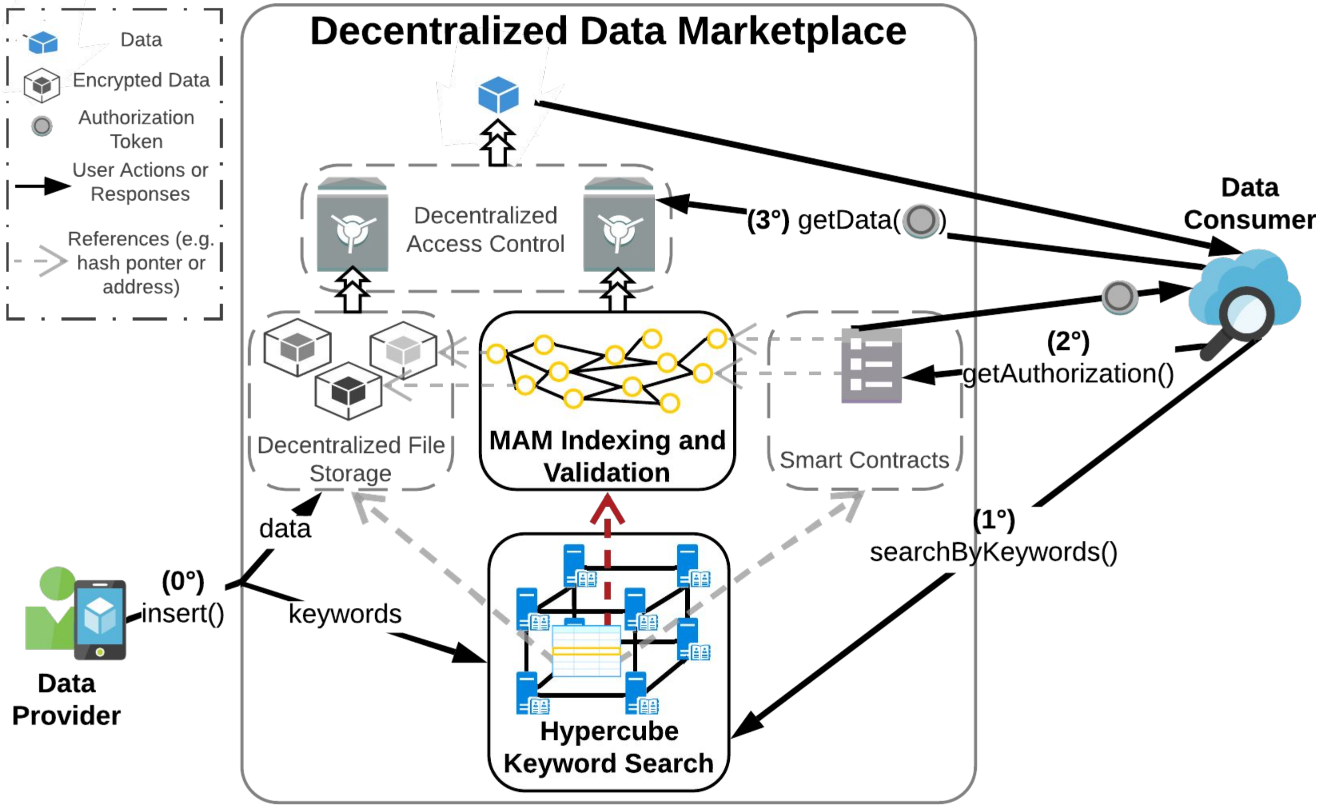
Public Permissionless Ledger

Personal Data Anonymized

2. **Towards Decentralized Complex Queries over Distributed Ledgers: a Data Marketplace Use-case**

Mirko Zichichi, Luca Serena, Stefano Ferretti, Gabriele D'Angelo

(IEEE ICCCN21 conference paper accepted)



3. **Intelligible Identities and Certificates**

(Bitnomos traineeship)

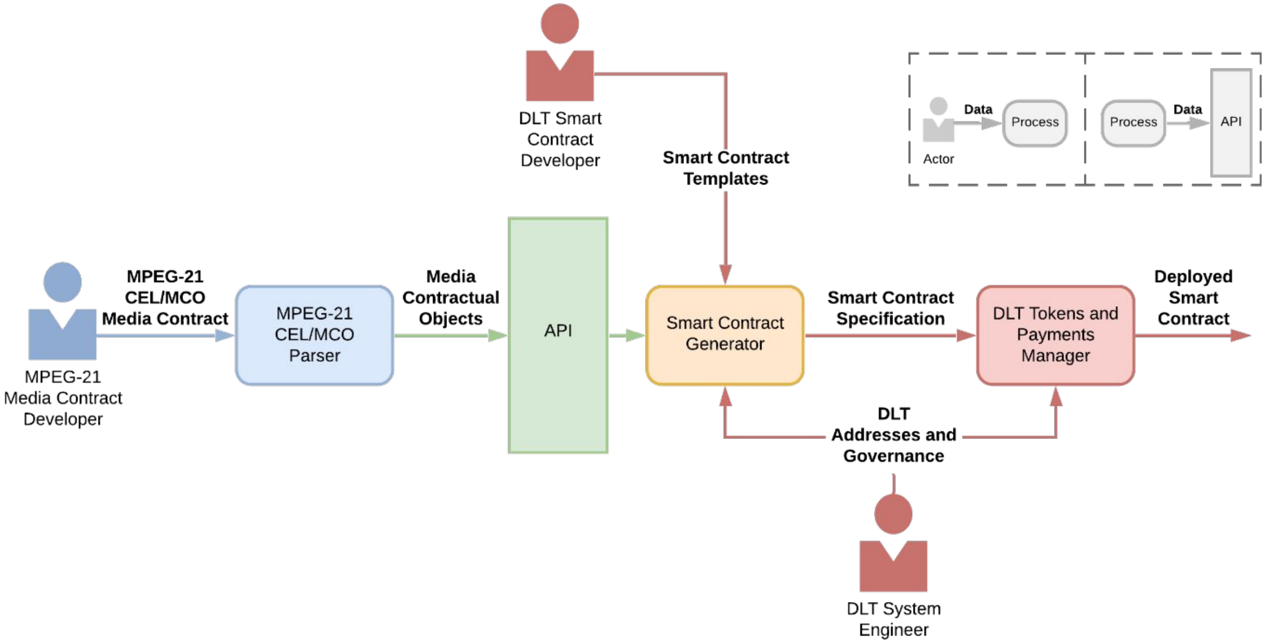
- Two models for managing intelligible, legal, and machine-processable certificates.
- Using standards for mark-up and identification of legal documents, in order to associate an easily verifiable legal and operational context https://demo.intelligible.io

4. **Richer policies expression for smart contracts access control**

(ODRL?)

MPEG ISO/IEC 21000-23 Smart Contracts for Media

A standard way to further translate MPEG-21 contracts to smart contracts ensuring users that the clauses of the smart contracts executed by a DLT correspond to the clauses of the MPEG-21 ontologies and languages (ISO/IEC 21000-21 Media Contracts Ontology & ISO/IEC 21000-20 Contract Expression Language)


DLT Smart Contract Developer
Smart Contract Templates
Data
Process
Actor
Process
Data
API
MPEG-21 CEL/MCO Media Contract
MPEG-21 CEL/MCO Parser
Media Contractual Objects
API
Smart Contract Generator
Smart Contract Specification
DLT Tokens and Payments Manager
Deployed Smart Contract
MPEG-21 Media Contract Developer
DLT Addresses and Governance
DLT System Engineer

Thank you!

https://mirkozichichi.me
https://github.com/miker83z/
https://scholar.google.com/citations?user=ikl56qoAAAAJ