Law, Science and Technology
MSCA ITN EJD n. 814177

RI E
Rights of Internet of Everything

Mirko Zichichi

Supervisors:
prof. **Stefano Ferretti** – UNIBO
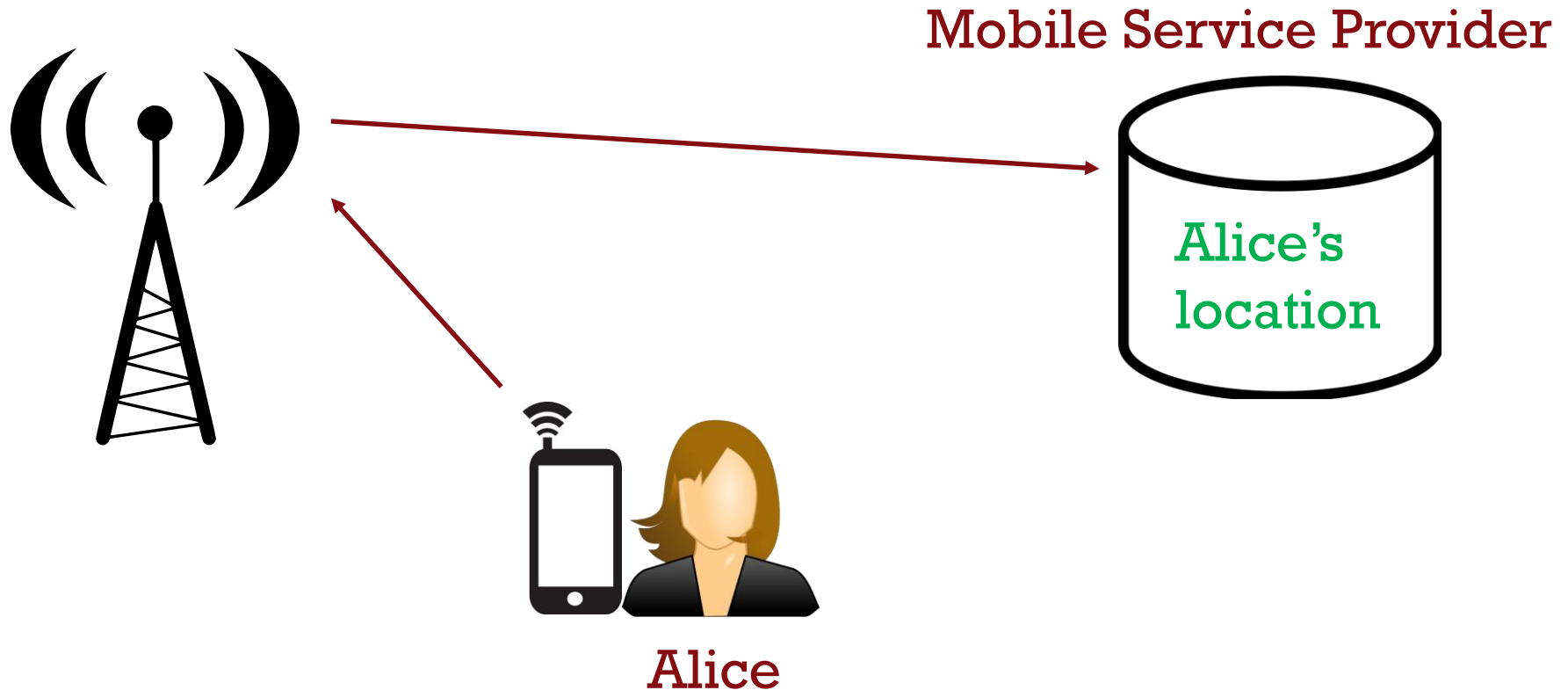prof. **Víctor Rodríguez Doncel** - UPM

Location privacy and inference in online social networks

20/11/2019

Outline

- Introduction
  - Location -> Personal Data
  - Problem
  - GDPR
- State of the Art
  - Semantic Web
  - Solid *by Tim Berners Lee*
  - Distributed Ledger Technologies (DLTs)
- Objectives
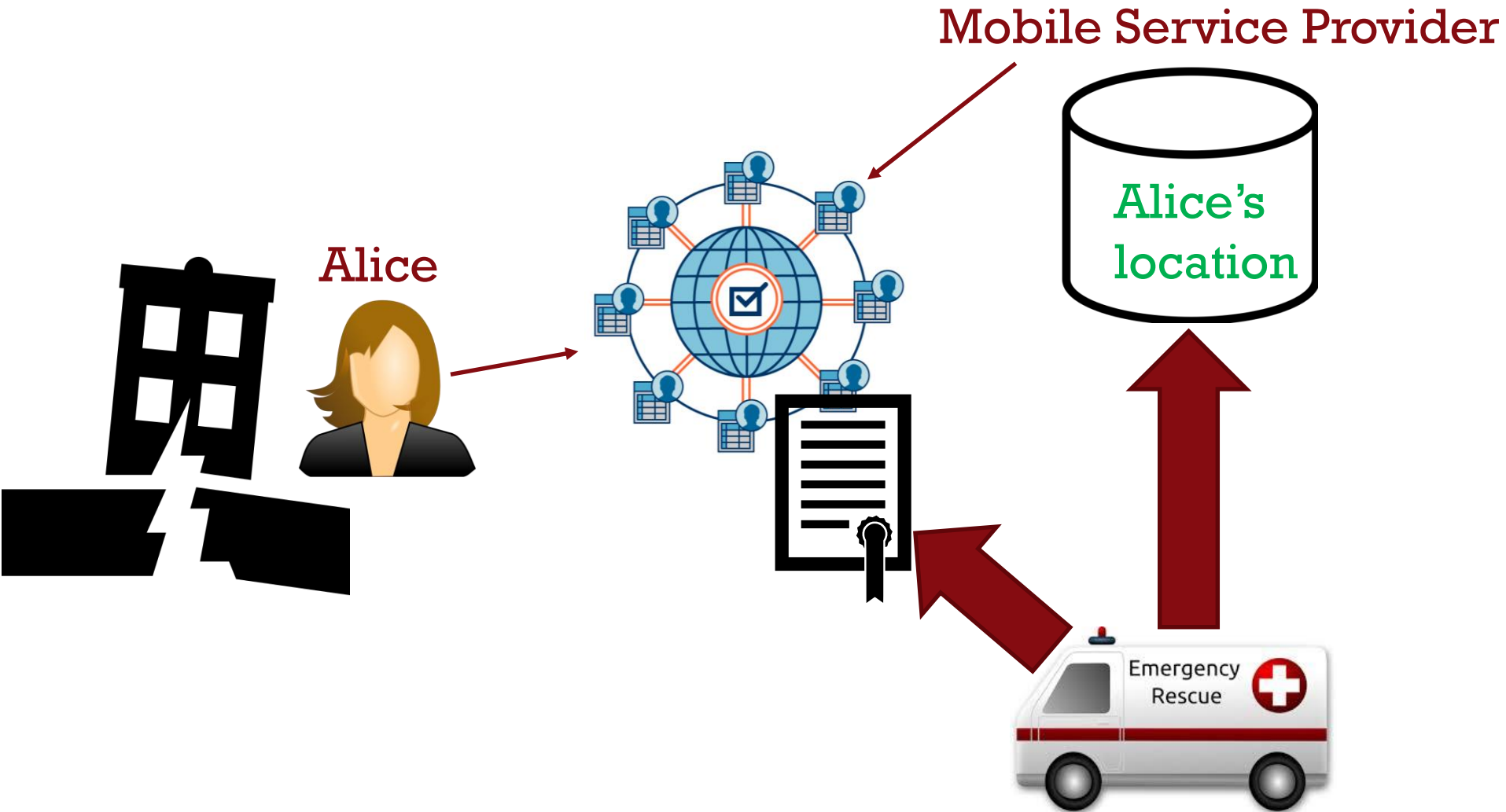- Hypoteses
- Research Questions
- Methodology
- Research Plan

RI E
Rights of Internet of Everything

Scenario (1/2)

Individual's location data generated by a provider

Mobile Service Provider

Alice's location

Alice

Individual's location data generated by a provider

Mobile Service Provider

Alice

Alice's location

Emergency Rescue

20/11/2019

+
Personal Data

- Any piece of information that can **identify** or be identifiable to a natural person

- Generated by the interaction of a user with a software or a hardware in form of:

    numbers, characters, symbols, images, sounds, electromagnetic waves, bits, etc. [1]

- Collected to improve the **safety and security** in citizens surveillance

- But also for a "not so new" **data-driven economy**

+ Problem

Abuse of personal information (Cambridge Analytica 2018)

- Personal data is sometimes **concentrated in few points** (e.g. online social networks) and transacted in **opaque transfers** without the individual's control or even knowledge

- Data is stored differently through several **data silos**, maintained by entities to which it is convenient **hampering** data exchange and its economical exploitation

- Individuals are not capable of determining the **fate** of their personal data, whereas they may be good willing to offer it for the **social good** (e.g. better policy making, research) or they want to make direct **profit** from it.

General Data Protection Regulation (GDPR)

GDPR [2] has empowered data privacy of citizens by radically changing operations carried out by data providers

Requires data providers to **release** to their users the complete dataset they collected on them, when requested.

- **No standards** for this requests
- There is the tendency to **hinder the progress** of these

GDPR **data portability** provides the right to have data directly transferred from one data provider to another, making a step towards user-centric platforms of interrelated services

- **Interoperability** [3]

Semantic Web

Extension of the World Wide Web through standards provided by the World Wide Web Consortium (W3C)

Semantic Web brings structure to the meaningful contents of the Web by promoting **common data formats and exchange protocols** [4] e.g.:

- **RDF** (Resource Description Framework) [5]
- **OWL** (Web Ontology Language) [6]

+ **Linked Data**: data published in a structured manner, in such a way that information can be found, gathered, classified, and enriched using annotation and query languages.
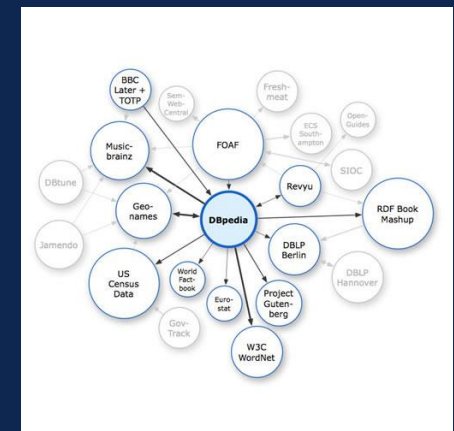
SOLID (Tim Berners Lee's project)

Involves the use of distributed technologies and Semantic Web integration in social networks. Born with the purpose of giving users their data sovereignty, letting them choose where their data resides and who is allowed to access and reuse it [7]

https://rubenverborgh.github.io/Solid-DeSemWeb-2018/
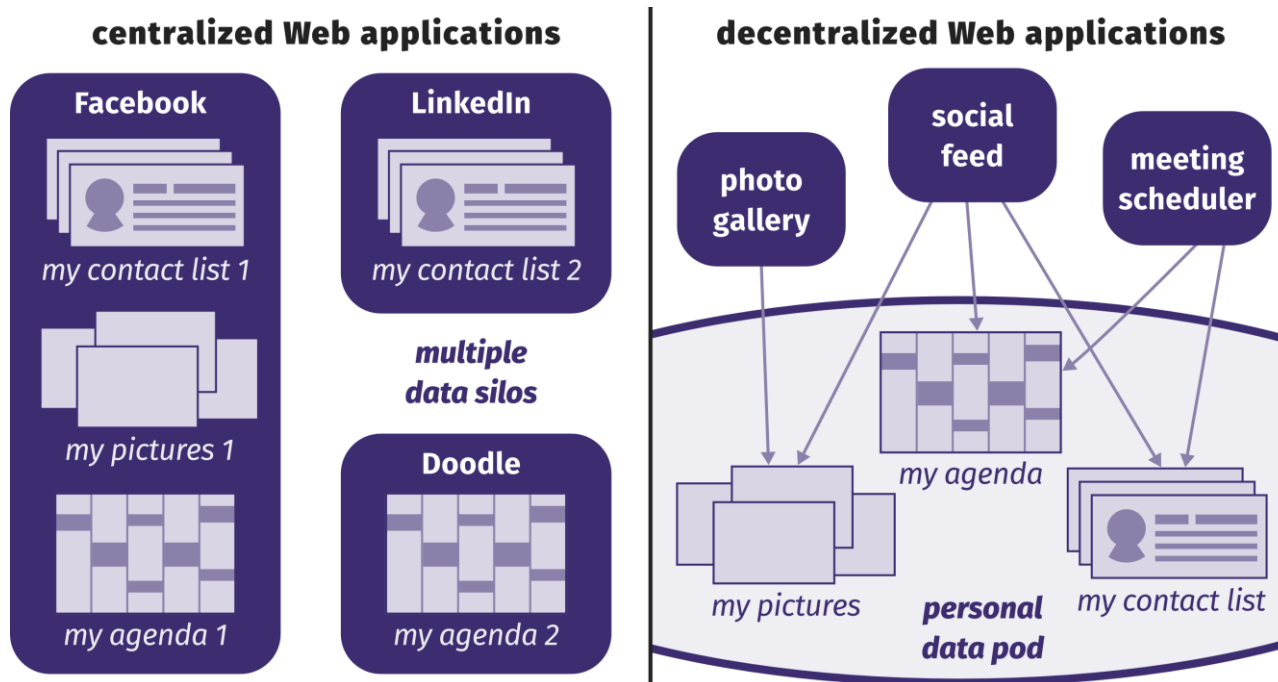
Distributed Ledger Technologies

- A software infrastructure maintained by a p2p network, where the network participants must reach a **consensus** on the states of transactions submitted to the distributed ledger

- A DLT brings trust when there are several parties that concur in handling some data in a **trustless** manner

- The Ethereum **Smart Contract** [8] is a new concept of contract that brought a second blockchain revolution

- SCs remove the technology bond with finance and provide a new paradigm where **unmodifiable instructions** are executed in an **unambiguous manner** during a transaction between two parts

+ Objectives

Design methods and systems to support the right of individuals to the **protection** of personal data, at the same favoring its **portability** and economic exploitation and fostering the social good

1. To design methods and systems that store and transfer personal data in a **controlled, transparent and non-centralized** manner

2. To identify **modeling and evaluation** methodologies for the analysis of decentralized and complex systems, e.g. to understand possible actors and manners to **infer** data

3. To specify languages and protocols that favour personal data **interoperability**

4. To specify the languages and algorithms necessary to **represent and reason with policies in smart contracts** to govern the access to personal data

+ Hypotheses

1. The use of DLTs for data management would grant: *data validation, access control, no central point of failure, immutability and traceability*

2. It is possible to use **decentralized file systems** for storage in order to allow continuous data availability.

3. Location privacy can be guaranteed through "suitable" cryptographic techniques (e.g. Zero Knowledge Proof)

4. Interoperability can be best achieved if data models adapt the **W3C specifications** for the semantic web.

5. By means of defeasible deontic logic in **smart contracts** individuals are able to state how their personal data is managed.

6. Operating with these technologies is **fast enough** to ensure the "correct" execution of processes that require individuals' personal data.

+
Research Questions

Are decentralized technologies and semantic web standards able to optimally support individuals' personal data protection and interoperability?

1. Is it possible, using these technologies, to handle large quantity of data **maintaining privacy** and **efficiency** in indexing and accessibility? And how can it be evaluated?

2. Is the current specification of smart contracts able to assure the **correct execution** of individuals intentions?

3. Which challenges to the **use and diffusion of semantic web technologies** do entities, that extract and/or process data from individuals, present?

+ Methodology

1. A **decentralized digital space** will be specified.
   This methodology is requirement-driven and empirically validated.

2. Standard system evaluation methods may not be sufficient in such environment, hence compliant methods must be studied
   (e.g. **complex networks analysis**).

3. A network of **ontologies** will be developed to model the personal data life-cycle and their actors.

4. The design of **Smart Contracts** will be focused towards **legal requirements** and **privacy preferences**, in compliance with GDPR

Research Plan

Tasks	2020													2021													2022							
	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44
	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11

Stays: UNIBO | UNITO | UPM | UNIBO

Board Meeting

Gain Focus! — Overall Goal Y1
- StateOfTheArt+Define Architecture
- Prototype Design
- Modifications
- First Year's Report
- Minimal Working Prototype

Go Semantics! — Overall Goal Y2
- Defined Solution

Evaluate! — Overall Goal Y3
- Complete Thesis

Spring/Summer Schools
- Blockchain Bologna
- Blockchain Padova
- Semantic Web Bertinoro

Conferences — Circa 5/6 conferences
- Jurix 2019
- Jurix 2020

20/11/2019

RI°E
Rights of Internet of Everything

Publications

- M. Zichichi, S. Ferretti, and G. D'Angelo, "**A distributed ledger based infrastructure for smart transportation system and social good**," in IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, USA, 10-13 January, 2020

- M. Zichichi, S. Ferretti, and G. D'Angelo, "**Are Distributed Ledger Technologies Ready for Smart Transportation Systems?**", *submitted to* IEEE International Conference on Communications (ICC), Dublin, Ireland, 7-11 June 2020

Not for the LAST-JD RIoE project, but related:

- M. Zichichi, M. Contu, S. Ferretti, and G. D'Angelo, "**Likestarter: a Smart-contract based social DAO for crowdfunding**," in Proc. of the 2st Workshop on Cryptocurrencies and Blockchains for Distributed Systems (CryBlock'19), Paris, France, 29 April, 2019

References

1. R. Kitchin, The data revolution: Big data, open data, data infrastructures and their consequences. Sage, 2014.

2. Council of European Union, "Regulation (eu) 2016/679 - directive 95/46," pp. 1–88

3. P. De Hert, V. Papakonstantinou, G. Malgieri, L. Beslay, and I. Sanchez, "The right to data portability inthe gdpr: Towards user-centric interoperability of digital services,"Computer Law & Security Review,vol. 34, no. 2, pp. 193–203, 2018

4. T. Berners-Lee, J. Hendler, O. Lassilaet al., "The semantic web,"Scientific american, vol. 284, no. 5,pp. 28–37, 2001

5. https://www.w3.org/TR/rdf-syntax-grammar/

6. https://www.w3.org/TR/owl-features/

7. A. V. Sambra, E. Mansour, S. Hawke, M. Zereba, N. Greco, A. Ghanem, D. Zagidulin, A. Aboulnaga,and T. Berners-Lee, "Solid : A platform for decentralized social applications based on linked data,"2016

8. V.Buterin et al.,"Ethereum whitepaper" 2013.[Online]. Available: https://github.com/ethereum/wiki/wiki/White-Paper