Towards new models of information sharing

Camille DUBEDOUT [1], **Manuel MUNIER** [2]

[1] Université Grenoble Alpes
CESICE, Grenoble, France
Centre d'Études sur la Sécurité Internationale et les Coopérations Européennes

[2] Université de Pau et des Pays de l'Adour, E2S UPPA
LIUPPA, Mont-de-Marsan, France
Laboratoire d'Informatique de l'Université de Pau et des Pays de l'Adour

[CRiSIS] 2024 Aix-en-Provence (France)

Wednesday, November 27, 2024

Motivation

- **The rise of data exchange**
  - 21st century = era of the « information society »
  - ICTs promote the circulation of large amounts of data
    - IoT, smart cities, industry 4.0,…
    - connected information systems, systems of systems, cloud services,…
    - ➔ new services : transport, energy, waste management, traffic flow, smart agriculture,…
    - ➔ new risks to information security

Motivation

The rise of data exchange

- 21st century = era of the « information society »
- ICTs promote the circulation of large amounts of data
  - IoT, smart cities, industry 4.0,…
  - connected information systems, systems of systems, cloud services,…
  - ➔ new services : transport, energy, waste management, traffic flow, smart agriculture,…
  - ➔ new risks to information security
- However…
  - most of these data capture and processing technos are designed and managed by private entities
  - these private actors set the rules for data sharing, often under the guise of business secrecy
  - ➔ this can limit access to and use of data by external parties (e.g. citizens, local authorities)
  - ➔ many data flows are exploited for private purposes ➔ public interest in dissemination ?

Motivation

- **The challenges of data exchange**
  - ICTs are generally focused to « routing data from point A to point B »
    - networking : 4G/5G, FTTH, network security,…
    - cryptography (encryption, signature, watermarking,…), blockchain (distributed ledger, sync)
    - security policy : access control, Service Level Agreement, traceability & accountability,…
    - risk management mainly internal to the company

Motivation

The challenges of data exchange

- ICTs are generally focused to « routing data from point A to point B »
  - networking : 4G/5G, FTTH, network security,…
  - cryptography (encryption, signature, watermarking,…), blockchain (distributed ledger, sync)
  - security policy : access control, Service Level Agreement, traceability & accountability,…
  - risk management mainly internal to the company

- What is missing, from our point of view :
  - keeping control over the data we share ➔ users need trust
  - not imposing a single, overly strict tool
  - **decoupling the producer and the consumer** ➔ they don't necessarily know each other
  - ➔ a question arises: the responsibilities of each party
  - ➔ data governance concept

About this work

- « Convergences du Droit et du Numérique » (CDN) 2020
  - convergence of law and digital technology
    - https://cdn.u-bordeaux.fr/
  - round #1 ➔ **workshop**
    - each participant presents themes he or she would like to address
  - **creation of pairs** (1 lawyer, 1 IT specialist) & **brainstorming**
  - round #2 ➔ **symposium**
    - each pair presents their thoughts
    - the computer scientist presents the legal details; the lawyer presents the IT details
    - ➔ need to popularize the speech
    - discussion

About this work

- **C.Dubedout / M.Munier pair on responsibilities in data sharing**
  - C.Dubedout (law) ➔ new risks in cities, cybersecurity, data & privacy
    - employed at ANSSI (French Cybersecurity Agency)
  - M.Munier (computer science) ➔ usage control, risk management, cybersecurity, data & privacy
- **Content of our article**
  - (quick) overview of current and future regulations
  - proposal of an IT vision for data governance
  - paradigm shift to management by the Commons

Regulations

- **« Loi pour une République Numérique » (LRN)**
  - French Law for a Digital Republic – October 7, **2016**
  - encourages administrations to open and share public data, in order to facilitate its reuse by businesses and citizens
  - this law **mandates** that all state administrations, local authorities with more than 3,500 inhabitants, public establishments, and private organizations in charge of public services publish their source codes and public databases online ➔ **Open Data** (among other things…)

Regulations

- **« Loi pour une République Numérique » (LRN)**
  - French Law for a Digital Republic – October 7, **2016**
  - encourages administrations to open and share public data, in order to facilitate its reuse by businesses and citizens
  - this law **mandates** that all state administrations, local authorities with more than 3,500 inhabitants, public establishments, and private organizations in charge of public services publish their source codes and public databases online ➔ **Open Data** (among other things…)

- **General Data Protection Regulation (GDPR)**
  - European Union – effective May 25, **2018**
  - only applies to **personal data**
  - does not require specific tools ➔ principles, rights and obligations ; guidelines ; compliance « controls »

Regulations

Data Governance Act (DGA)

- European Union – entered into force on June 23, 2022 – applicable since September **2023**

- the initiative aims to make more data available and facilitate data sharing across sectors and EU countries in order to leverage the potential of data for the benefit of European citizens and businesses

- promotes « free » data sharing across the Union by encouraging the establishment of common data-sharing spaces and **data intermediaries** (between **data holders** and **data users**)

- how ? ➔ the DGA establishes a European governance framework in order to :
  - build trust in voluntary data sharing
  - stimulate mechanisms aimed at increasing data availability
  - remove technical barriers to data reuse

Regulations

- **Data Act**
  - European Union – entered into force on January 11, **2024**
  - its main objective is to make Europe a leader in the data economy by harnessing the potential of the ever-increasing amount of industrial data, in order to benefit the European economy and society

Regulations

- **Data Act**
  - European Union – entered into force on January 11, **2024**
  - its main objective is to make Europe a leader in the data economy by harnessing the potential of the ever-increasing amount of industrial data, in order to benefit the European economy and society

- **Artificial Intelligence Act (AI Act)**
  - European Union – published in the Official Journal (OJ) of the European Union on July 12, **2024**
  - it says that AI systems that can be used in different applications are analysed and classified according to the risk they pose to users ➔ the different risk levels will mean more or less regulation
  - transparency requirements
    - disclosing that the content was generated by AI
    - designing the model to prevent it from generating illegal content
    - publishing summaries of copyrighted data used for training

So what ?

- **Europe and its citizens are becoming aware of the value of data**
  - data (or information) is a wealth and must be protected
  - data should not become the property of private actors (to which access fees must be paid)
  - ➔ on the contrary, their sharing should be encouraged
  - laws aim to give rights but also impose obligations ➔ responsibilities of the actors in data sharing

So what ?

- **Europe and its citizens are becoming aware of the value of data**
  - data (or information) is a wealth and must be protected
  - data should not become the property of private actors (to which access fees must be paid)
  - ➔ on the contrary, their sharing should be encouraged
  - laws aim to give rights but also impose obligations ➔ responsibilities of the actors in data sharing

- **How could ICT help?**
  - in information security risk management (eg. ISO 27005) data are **primary assets** for organizations
  - need to « keep control » of information exchanged between systems / actors
  - ➔ tools to define and apply **data governance**
  - how ?
    - allow data providers to define the usage rules they would like to be applied ➔ license
    - allow data consumers to verify whether their intended use complies with this license

Relying on formal tools

- **How to represent the concepts ?**
  - stakeholder identification
  - identification of assets and their properties
  - governance and participation rules
    - usage control
    - quality / traceability / transparency obligation
  - identifying the type of commitment ➔ license
  - determining responsibilities and penalties

Relying on formal tools

- **How to represent the concepts ?**
  - stakeholder identification
  - identification of assets and their properties

  } semantic web ➔ OWL for ontologies

  - governance and participation rules
    - usage control
    - quality / traceability / transparency obligation

  } SWRL for rules
  ➔ usage control rules expressed in OrBAC then translated into SWRL

  - identifying the type of commitment ➔ license
  - determining responsibilities and penalties

Relying on formal tools

- **How to represent the concepts ?**
  - stakeholder identification
  - identification of assets and their properties
  - governance and participation rules
    - usage control
    - quality / traceability / transparency obligation
  - identifying the type of commitment ➔ license
  - determining responsibilities and penalties

semantic web ➔ OWL for ontologies

SWRL for rules
➔ usage control rules expressed in OrBAC then translated into SWRL

- **Use of standard semantic web tools**
  - OWLAPI and SWRLAPI programming APIs
  - SWRL rules usable with semantic web reasoners (e.g. Drools, HermiT, Pellet)

Relying on formal tools

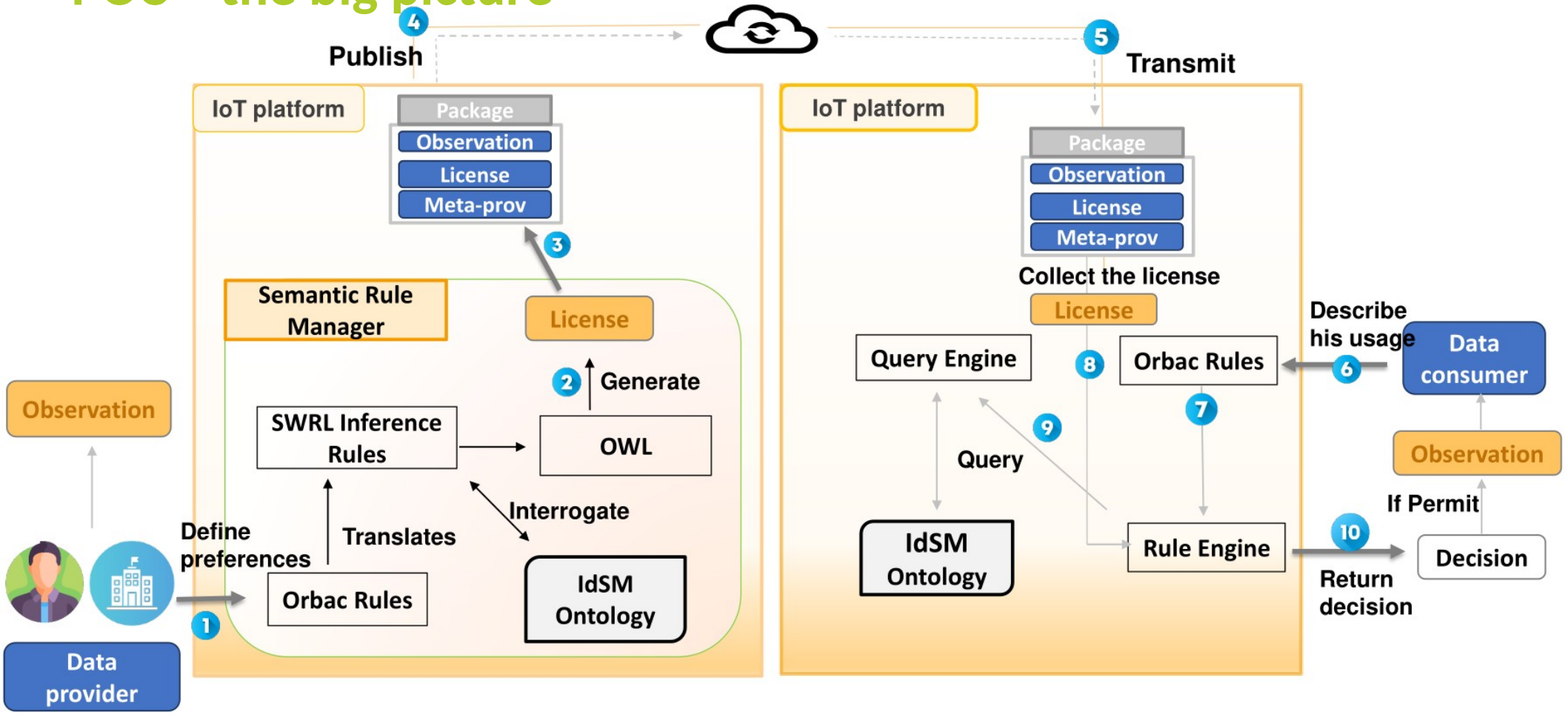
- **Benefits**
  - agreement on the terminology to be used (interoperability) ➔ as when drawing up "paper" contracts
  - rules used as such in inference engines
    - NB : some tools can provide **explanations** of inferred decisions
  - integration of existing ontologies
    - **SSN** (IoT), **PROV-O** (provenance), **OrBAC-O** (usage control policy)

Relying on formal tools

- **Benefits**
  - agreement on the terminology to be used (interoperability) ➔ as when drawing up "paper" contracts
  - rules used as such in inference engines
    - NB : some tools can provide **explanations** of inferred decisions
  - integration of existing ontologies
    - **SSN** (IoT), **PROV-O** (provenance), **OrBAC-O** (usage control policy)

- **First practical results**
  - **IdSM-O** ➔ an IoT Data Sharing Management Ontology for Data Governance
  - POC (proof of concept)
    - definition of rules, building of licenses, compliance of operations
    - wrapping of data published by legacy IoT platforms (with real sensors)
  - ➔ our "vision" of information sharing is feasible !

POC – the big picture



Paradigm shift...

- **Use of metadata**
  - from metadata to the obligation of data quality
    - the law could mandate that data producers and intermediaries publish their metadata
    - users could access more precise information, such as the production date of the data, any modifications made, and the identity of the organization or individual responsible for these changes
    - consider a certification process to be able to share quality data
  - from metadata to traceability and liability concerns
    - in case of litigation metadata can be also used to determine legal responsibility ("**liability**")
    - NB : liability differs from imputability and accountability
    - upward traceability ("where does the data come from ?") and downward traceability ("who uses my data ?")

Paradigm shift...

- **Another way to share information**
  - decoupling the provider and the consumer ➔ they don't necessarily know each other
  - no central authority to control data sharing ➔ stakeholders would group together into communities
  - data governance
    - usage control rules set by the data provider ➔ **informational self-determination**
    - metadata to increase data quality and liability

Paradigm shift to management by the Commons

Another way to share information

- decoupling the provider and the consumer ➔ they don't necessarily know each other
- no central authority to control data sharing ➔ stakeholders would group together into communities
- data governance
  - usage control rules set by the data provider ➔ **informational self-determination**
  - metadata to increase data quality and liability

Commons

- model for data management proposed by economist Elinor Ostrom (1977, 1990)
- information commons as « *resources governed by a system of* ***distributed rights*** *and a* ***governance structure*** *that ensures compliance with rights and obligations* » (Benjamin Coriat, 2015)
- « *a new economic paradigm that prioritizes* ***use value over ownership*** » (Jeremy Rifkin, 2014)

Conclusion

- **Towards better exchange practices**
  - information sharing can be a source of new vulnerabilities ➔ new risks
  - if we want users to share more, we need to provide them with tools for trust and **liability**
  - Q : who is responsible for managing data governance ? ➔ **governing Commons**

Conclusion

- **Towards better exchange practices**
  - information sharing can be a source of new vulnerabilities ➔ new risks
  - if we want users to share more, we need to provide them with tools for trust and **liability**
  - Q : who is responsible for managing data governance ? ➔ **governing Commons**
  - our reflection began a few years ago
    - about responsibilities in data sharing → **2020**
    - on an informational self-determination oriented approach → **2019** (PhD N.Laamech 2021-2024)
    - ➔ the EU was just beginning to present its draft « European strategy for data » (2020)
    - ➔ long before the Data Governance Act / Data Act / AI Act came into force (2023, 2024)

Conclusion

- **Towards better exchange practices**
  - information sharing can be a source of new vulnerabilities ➔ new risks
  - if we want users to share more, we need to provide them with tools for trust and **liability**
  - Q : who is responsible for managing data governance ? ➔ **governing Commons**
  - our reflection began a few years ago
    - about responsibilities in data sharing → **2020**
    - on an informational self-determination oriented approach → **2019** (PhD N.Laamech 2021-2024)
    - ➔ the EU was just beginning to present its draft « European strategy for data » (2020)
    - ➔ long before the Data Governance Act / Data Act / AI Act came into force (2023, 2024)
- **We are on the right way !**
  - higher-level language for expressing governance rules (rather than OrBAC or, worse, SWRL…)
  - transparency, explanation of inferred decisions, modeling of sanctions, etc.
  - more code, study data spaces, etc.

[CRiSIS] 2024

Thank you for your attention

CONTACT

Manuel MUNIER
Université de Pau et des Pays de l'Adour, E2S UPPA
LIUPPA, Mont-de-Marsan, France

Manuel.Munier@univ-pau.fr
http://munier.perso.univ-pau.fr/

UNIVERSITÉ DE PAU ET DES PAYS DE L'ADOUR | COLLÈGE STEE SCIENCES ET TECHNOLOGIES POUR L'ÉNERGIE ET L'ENVIRONNEMENT

e2s UPPA energy environment solutions

UGA Université Grenoble Alpes