

Applying ontology engineering in public administration

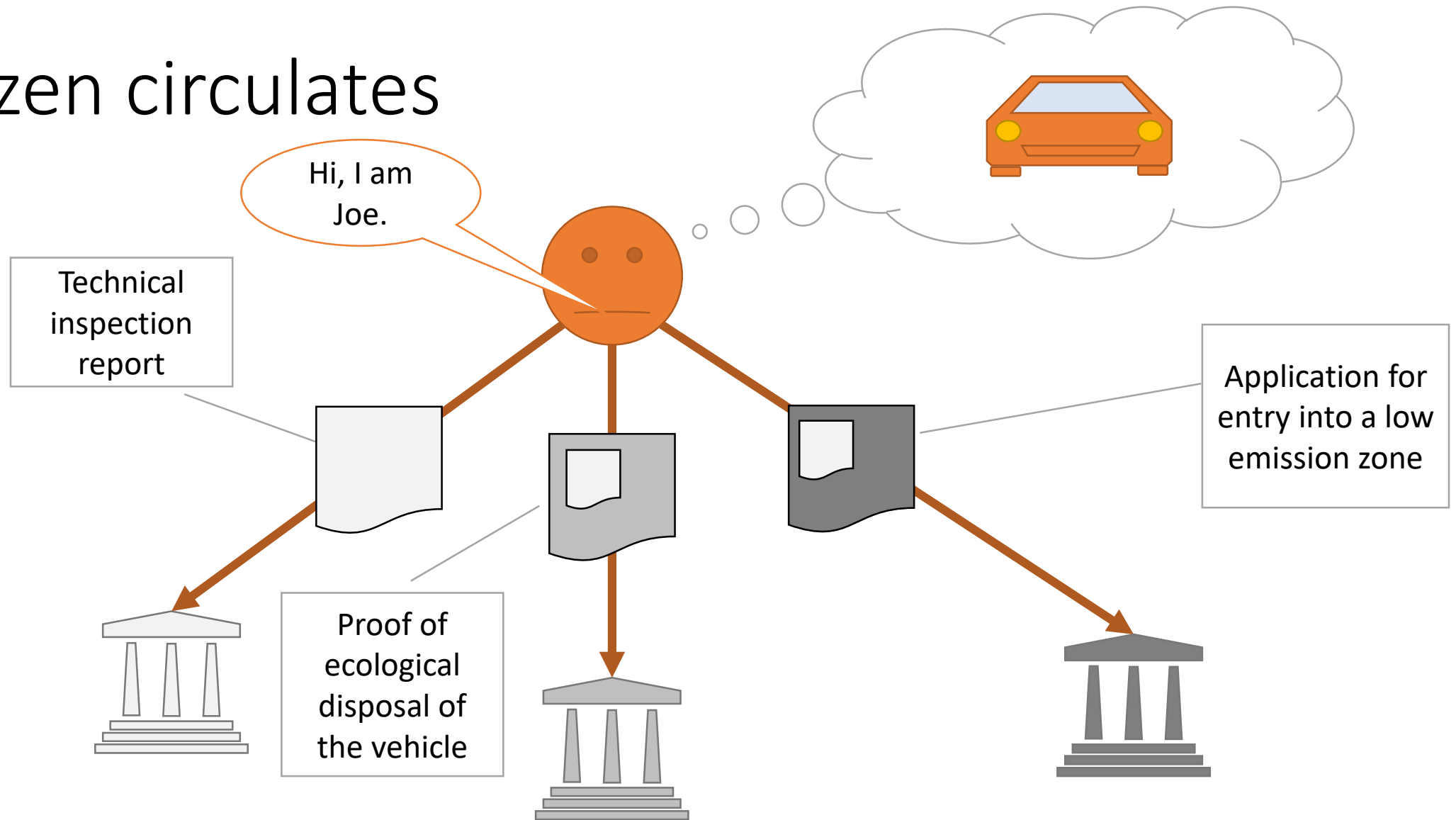
Martin Nečaský

martin.necasky@matfyz.cuni.cz

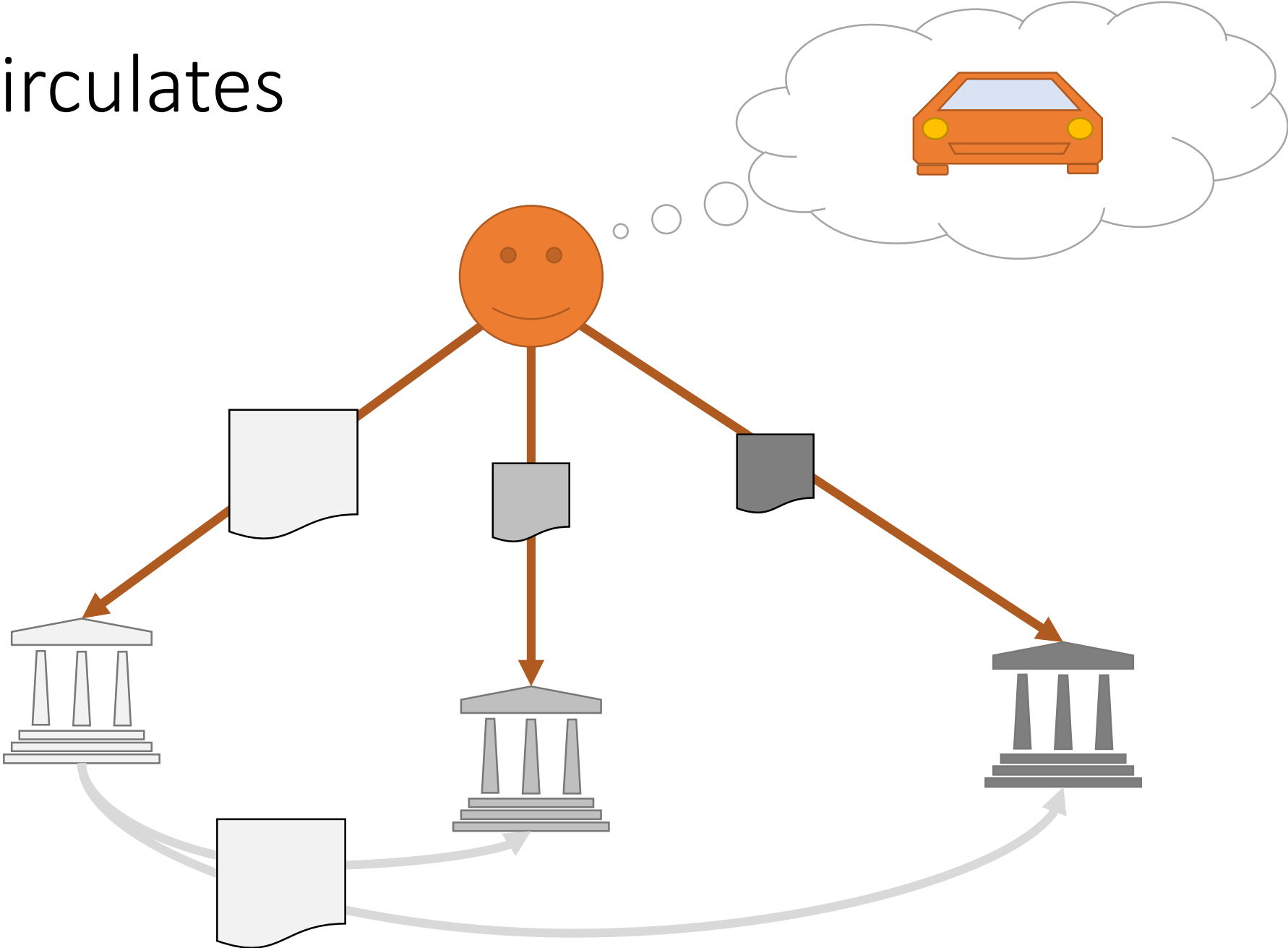
Department of software engineering
Faculty of mathematics and physics
Charles University

„Data circulates, not the citizen.“

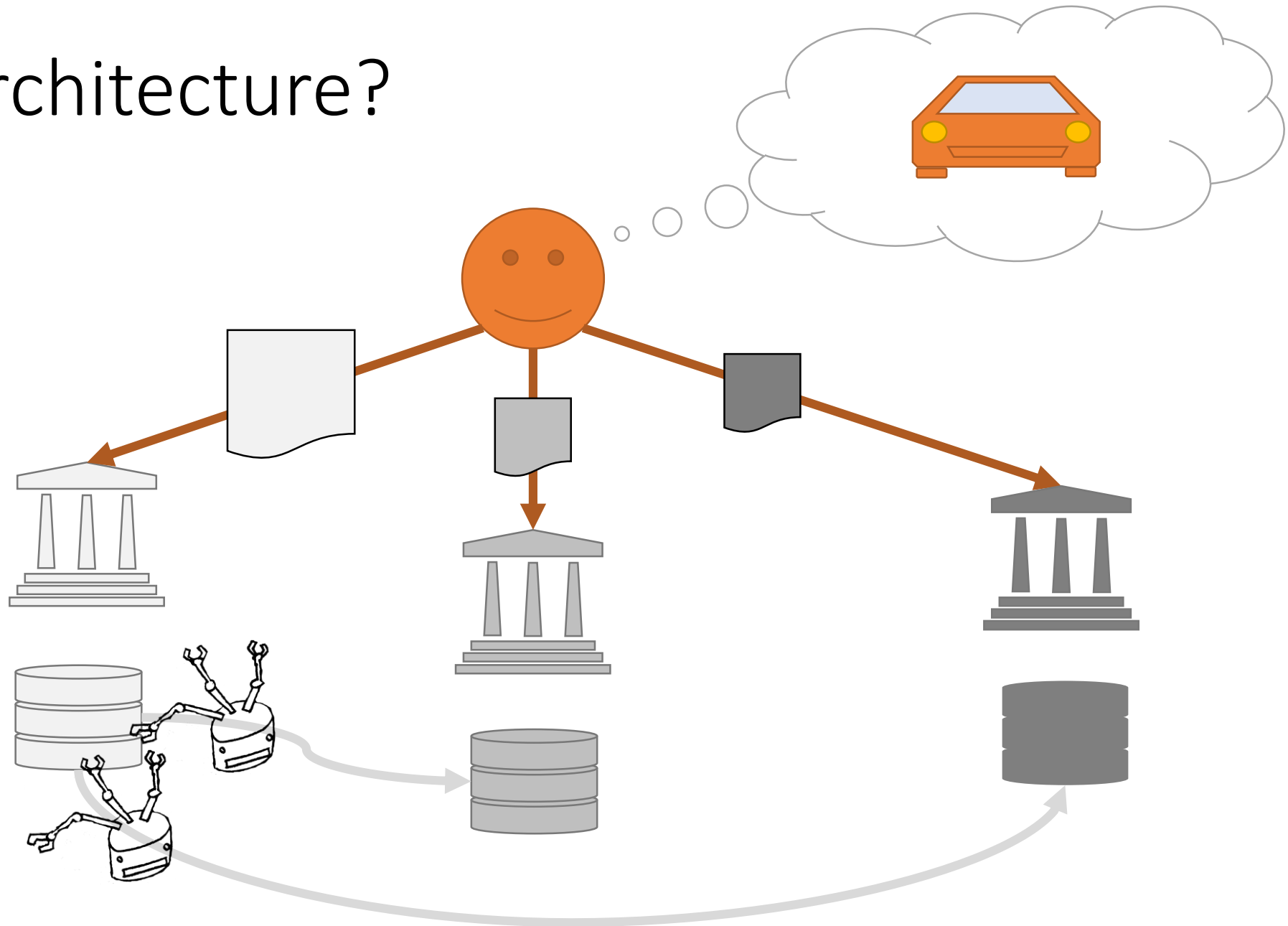
Citizen circulates



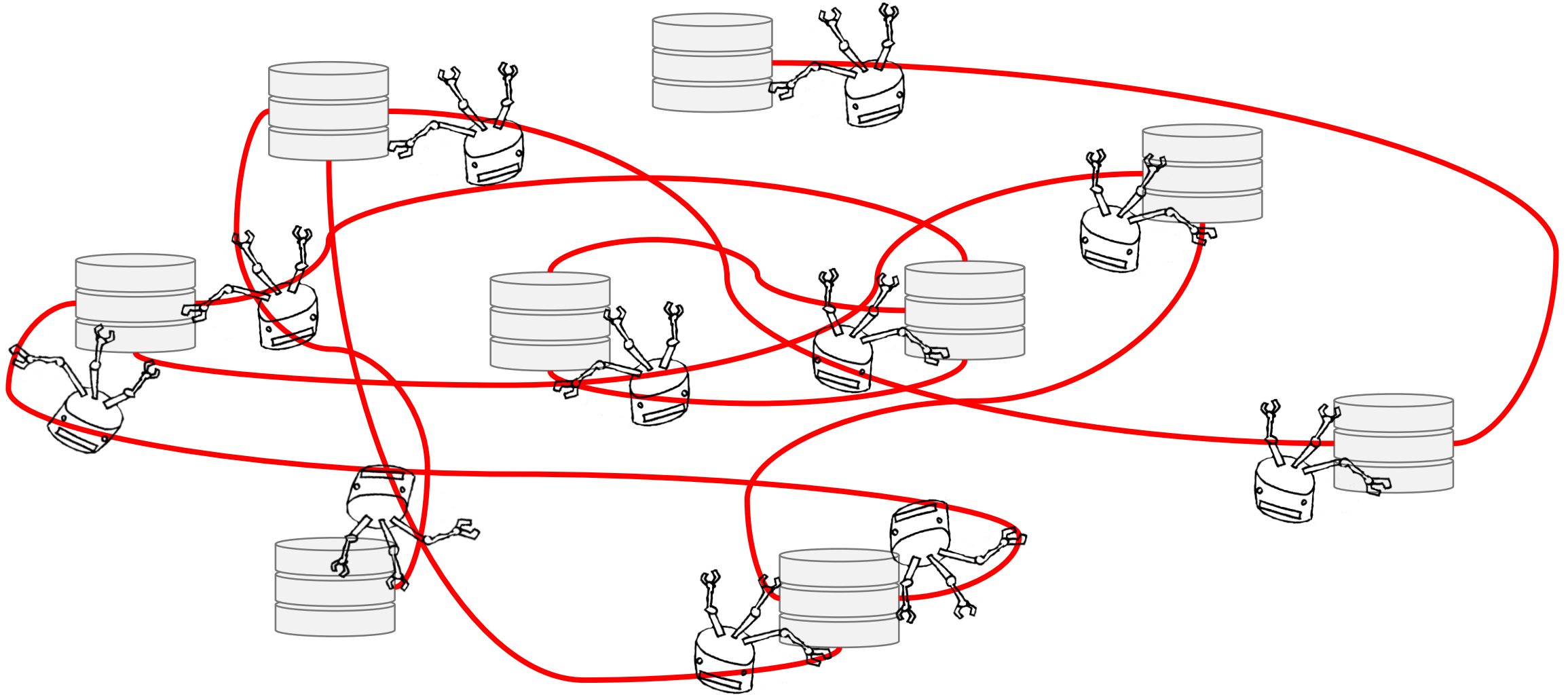
Data circulates



Bad architecture?



Hairball/big ball of mud is bad architecture!

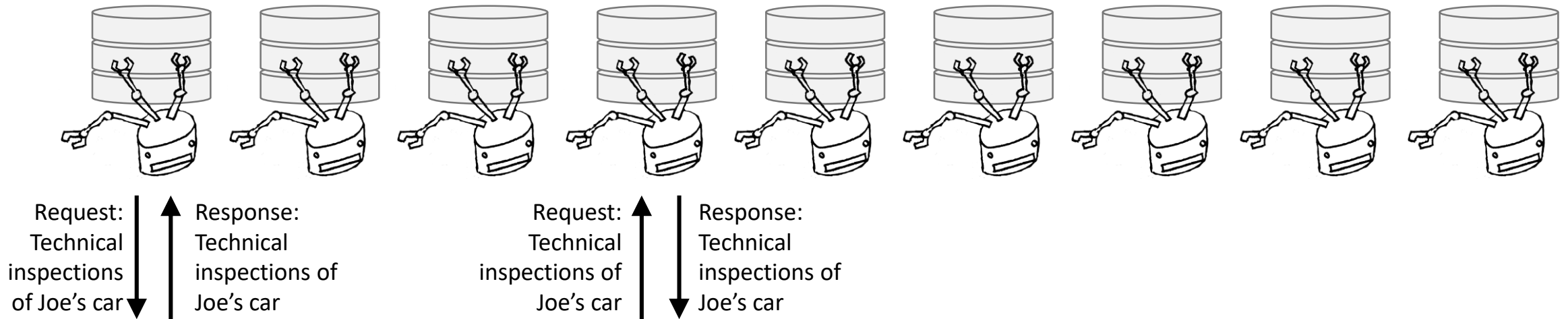


Possible solution?



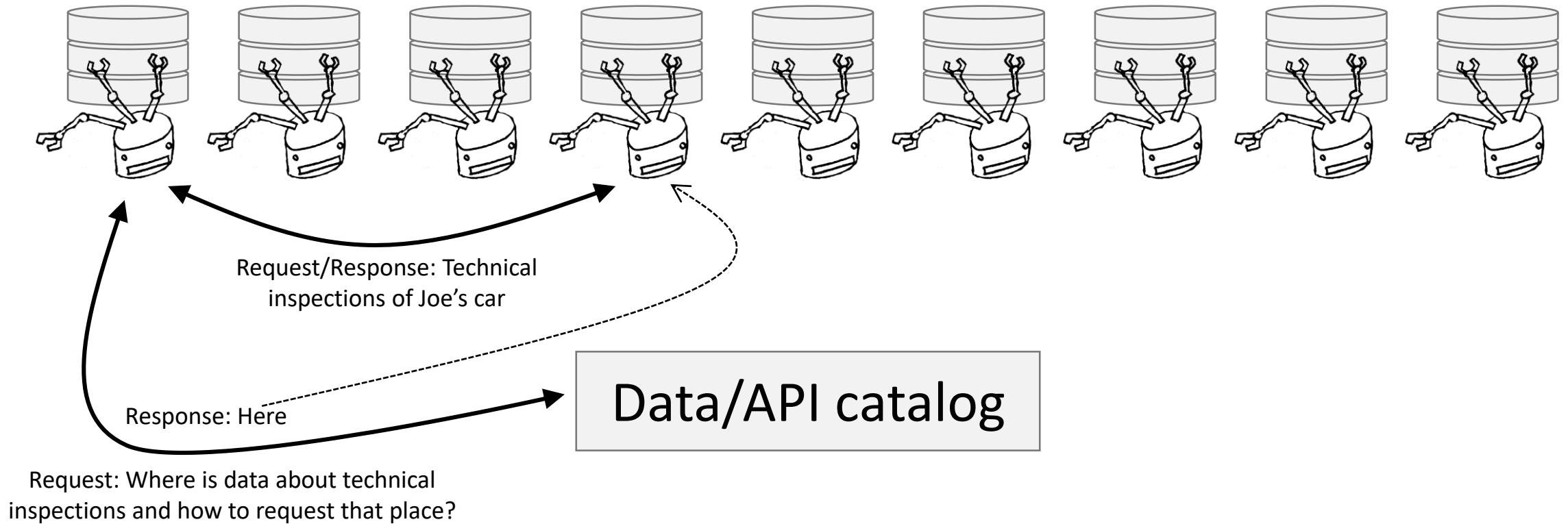
No point-to-point data exchange!!!

Possible solution – centralization



Enterprise service bus / API gateway

Possible solution – decentralization



Is such infrastructure enough?



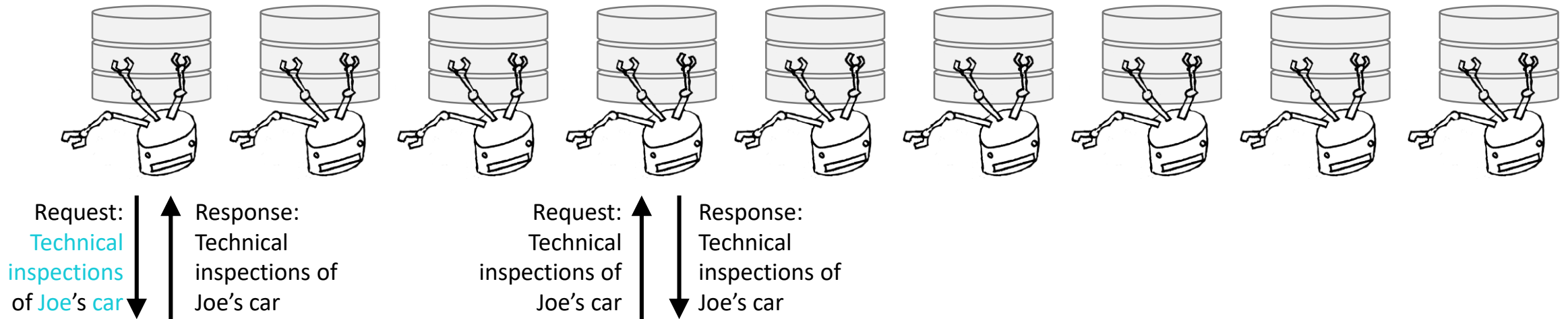
Know what data you have, where it is, how to access it, who is responsible, and who uses it for what.

Is such infrastructure enough?



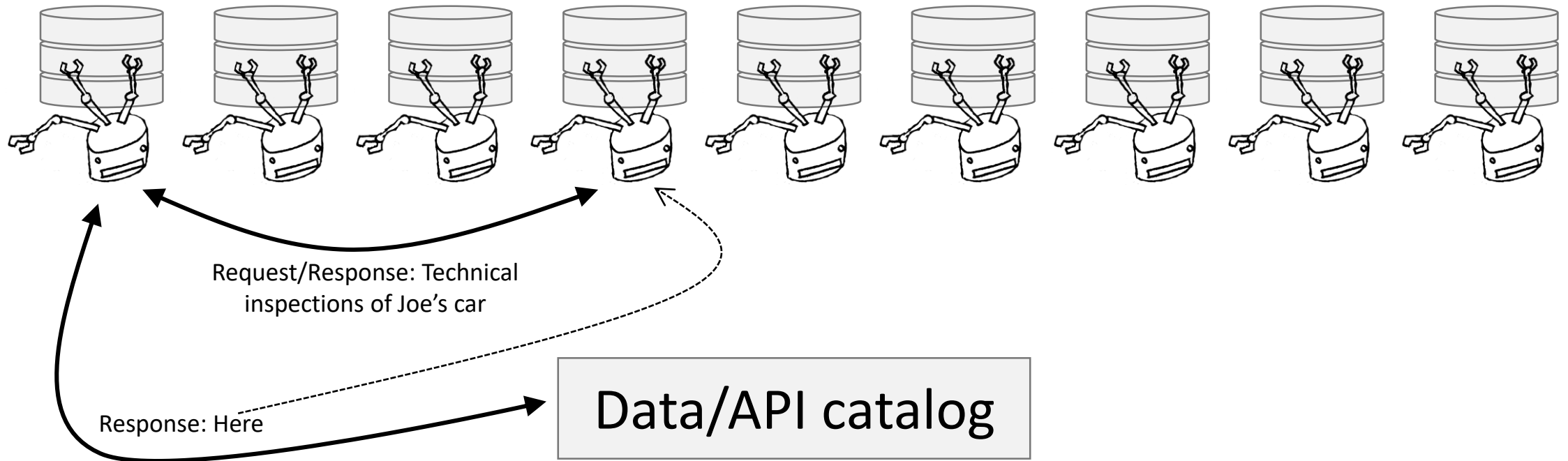
Know **what data you have**, where it is, how to access it, who is responsible, and who uses it for what.

Know what data you have



Enterprise service bus / API gateway

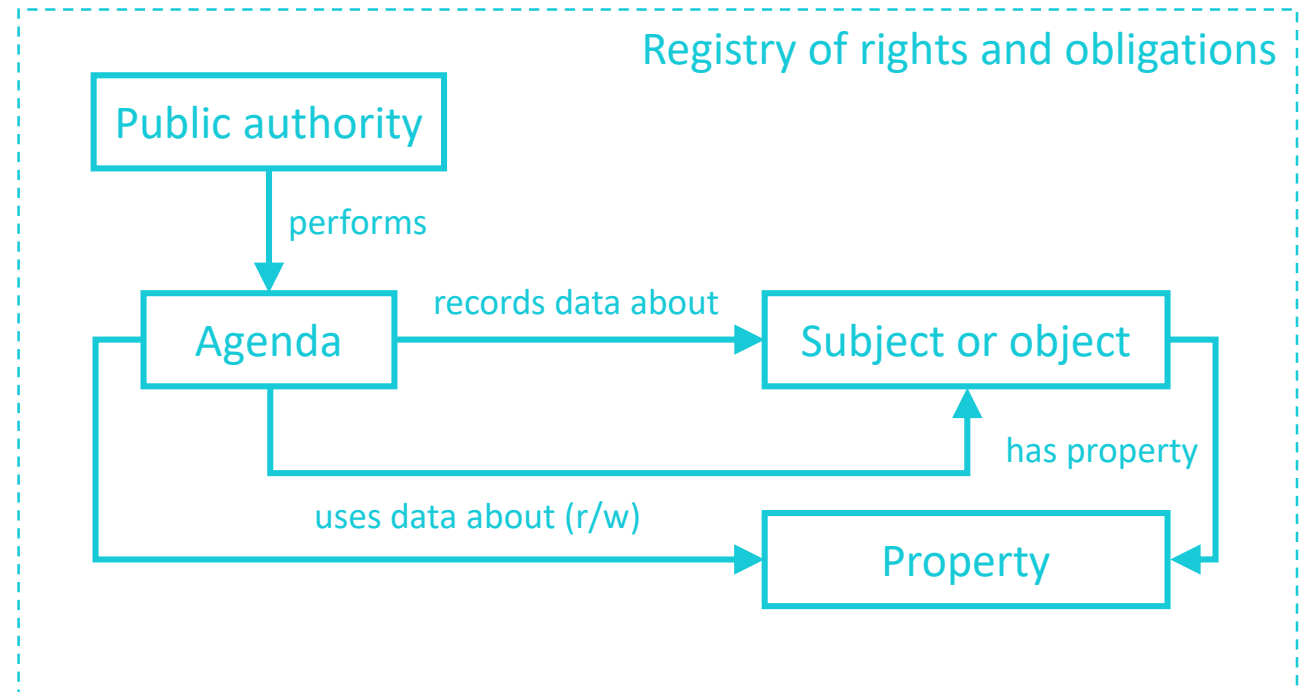
Know what data you have



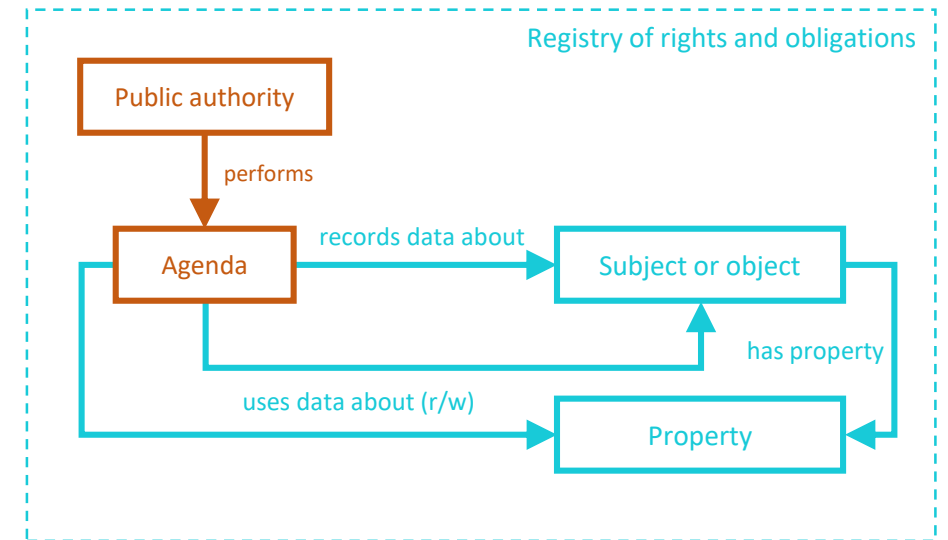
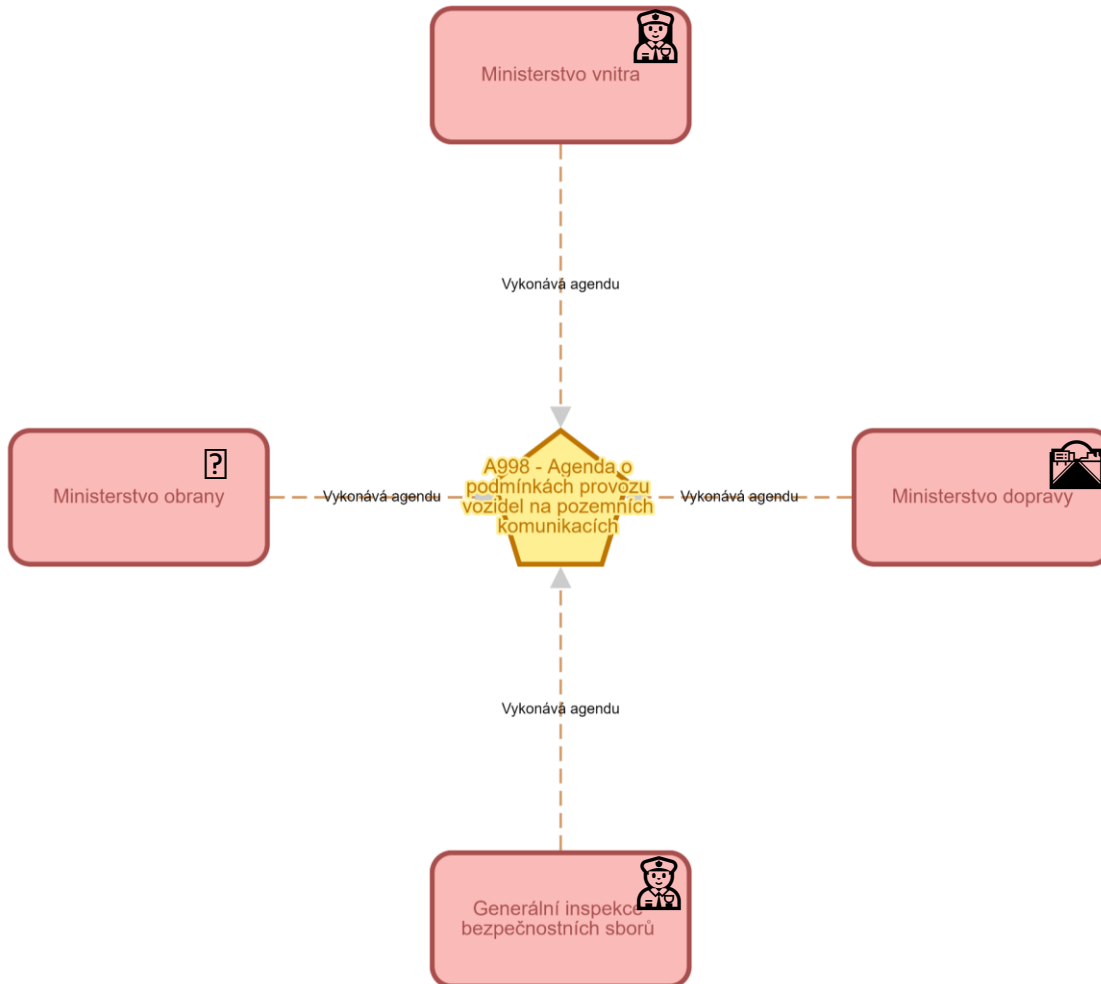
Request: Where is data about **technical inspections** of cars and car **owners** and how to request that place?

Registry of rights and obligations

- Central registry
- Map of
 - public authorities
 - their performed activities (called agendas)
 - data they collect and share



Registry of rights and obligations

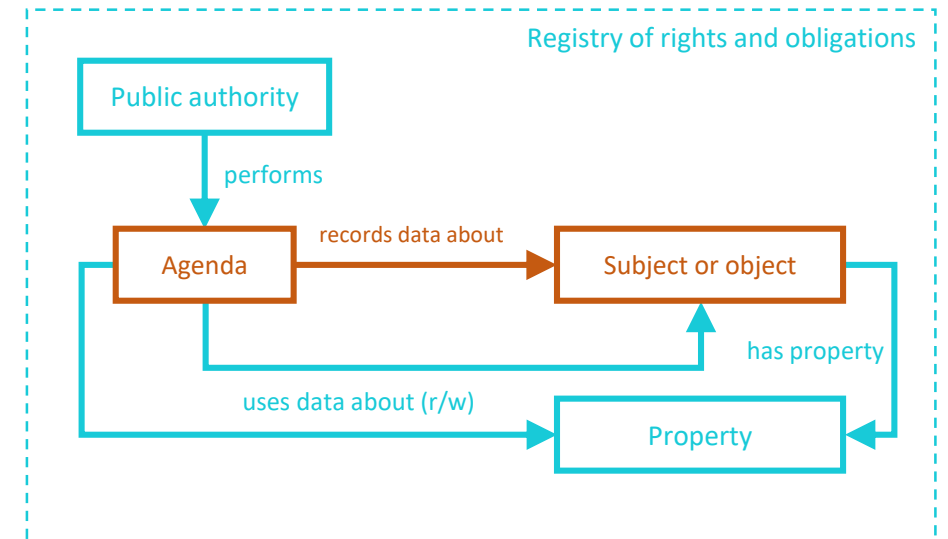


Registry of rights and obligations

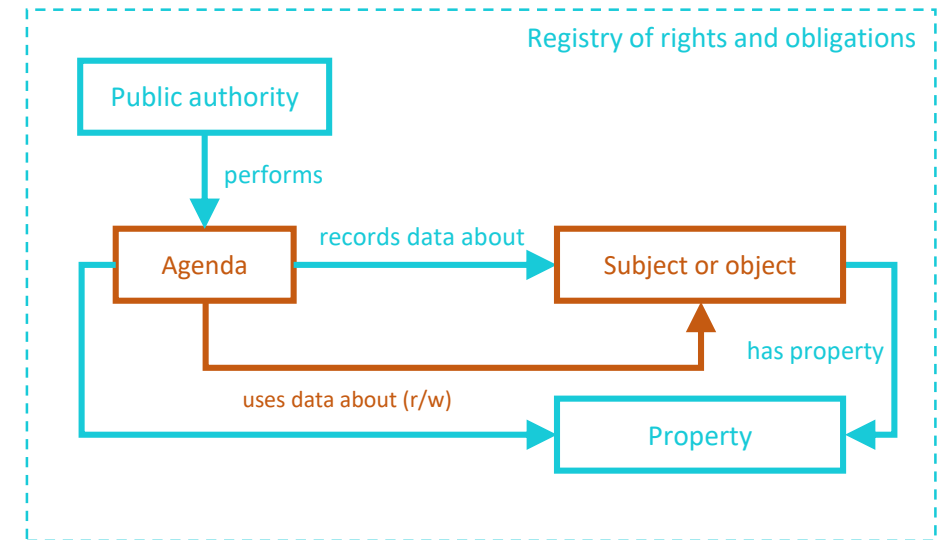
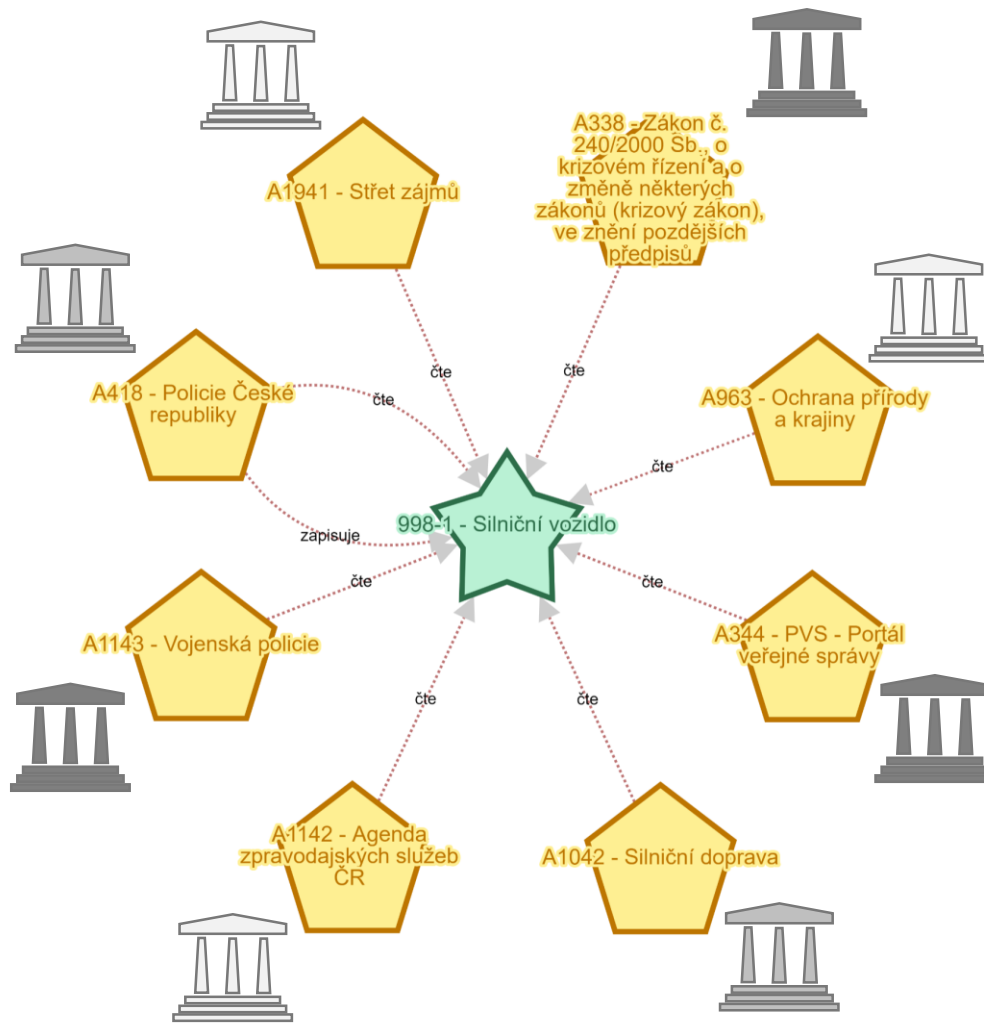
A998 - Agenda o podmínkách provozu vozidel na pozemních komunikacích

Sdružuje údaje vedené nebo vytvářené v rámci agendy

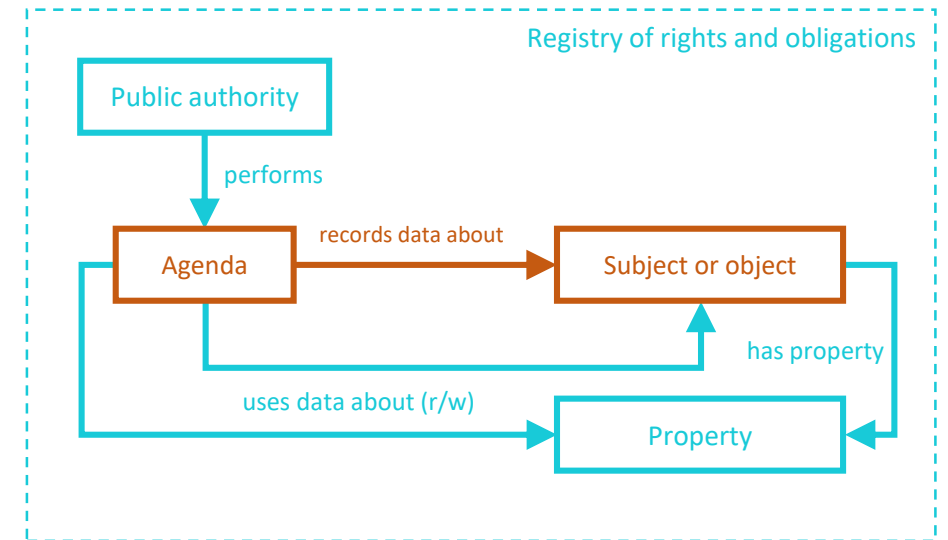
998-1 - Silniční vozidlo



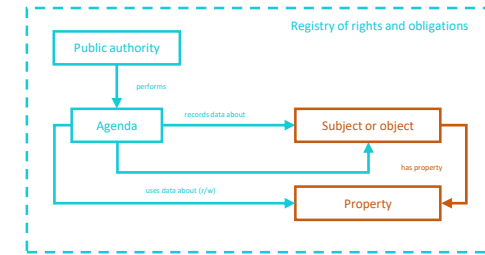
Registry of rights and obligations



Registry of rights and obligations

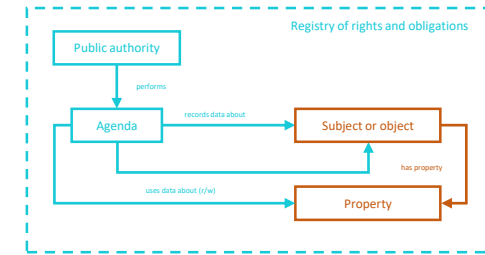


998 – Road traffic: 998-1 Road vehicle (properties)



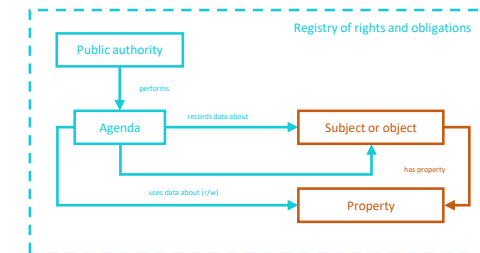
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-1	Údaje o vlastníkovi a provozovateli, není-li totožný s vlastníkem
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-10	Vyřazení silničního vozidla z provozu a zániku silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-11	Údaje o vyvezení silničního vozidla do jiného státu
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-12	Údaje o provedení záznamu v registru
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-13	Údaje o pojištění odpovědnosti
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-14	Údaj o tom, že vozidlo je historicky původní
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-15	Datum převodu vozidla z registru silničních vozidel do registru historických a sportovních vozidel
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-16	Účel užívání vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-17	Druh a kategorie silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-18	Značka, název a typ vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-19	Identifikační číslo vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-20	Údaje o výrobcu silničního vozidla, podvozku, motoru a karoserie silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-21	Typ, výkon a zdvihový objem motoru a palivo
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-22	Typ, výrobní číslo a barva karoserie, počet míst k sezení a stání, popřípadě lůžek, rozměry ložné ...
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-23	Údaj o celkových rozměrech silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-24	Největší přípustná, povolená a provozní hmotnost vozidla a největší přípustná a povolená hmotnost ...
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-27	Údaje o technických prohlídkách a technické způsobilosti silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-4	Registrační značka
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-5	Technický průkaz
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-6	Číslo osvědčení o registraci silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-7	Datum první registrace silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-8	Stát poslední registrace silničního vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/998-1-9	Odnětí osvědčení o registraci vozidla a technického průkazu vozidla

998 – Road traffic: 998-2 Registration of vehicles (properties)



https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-1	Evidence silničních vozidel, vlastníků a provozovatelů těchto vozidel
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-2	Evidence silničních vozidel členů diplomatické mise
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-3	Evidence ztracených, odcizených, poškozených a zničených osvědčení o registraci silničního vozidla, technických průkazů silničního vozidla a tabulek s přidělenou státní poznávací značkou (dále jen „registrační značka“)
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-4	Evidence přidělených zvláštních registračních značek a ztracených, odcizených, poškozených a zničených tabulek s přidělenou zvláštní registrační značkou
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-5	Evidence vyrobených formulářů osvědčení o registraci silničního vozidla a technického průkazu silničního vozidla opatřených identifikačními znaky a vyrobených a nevydaných tabulek registrační značky vozidla
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/údaj/998-2-6	Evidence technických údajů schválených typů silničních vozidel

1086 – Health services: 1086-1 Provider (properties)



https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/1086-1-1	Údaje provozovatele fyzická osoba	Osobní údaje: jméno, popřípadě jména, příjmení, rodné číslo, kontaktní údaje
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/1086-1-2	Údaje provozovatele fyzická osoba podnikatel	jméno, popřípadě jména, příjmení, IČO, kontaktní údaje
https://rpp-opendata.egon.gov.cz/odrpp/zdroj/udaj/1086-1-3	Údaje provozovatele právnická osoba	název, IČO, kontaktní údaje



Registry of rights and obligations - quantification

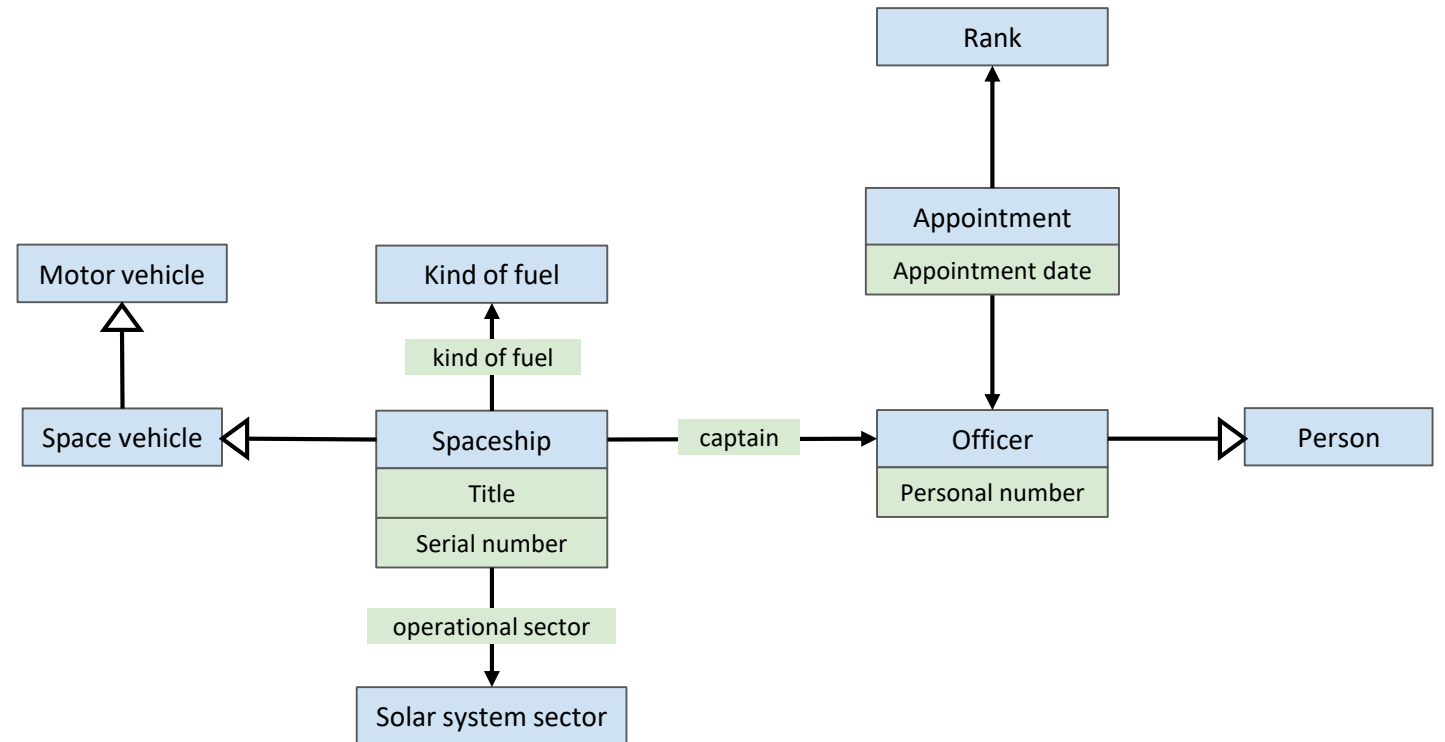
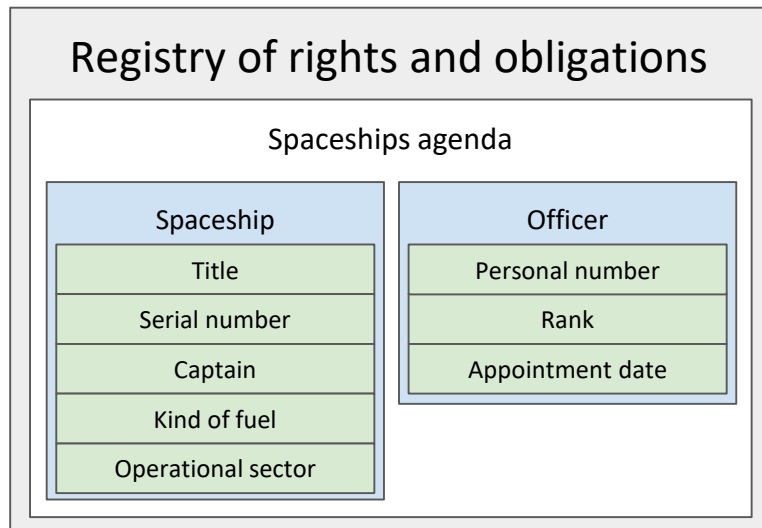
On 27.09.2022

- 222 types of subjects/objects
 - 1643 properties
- for 63 agendas out of 446 in total

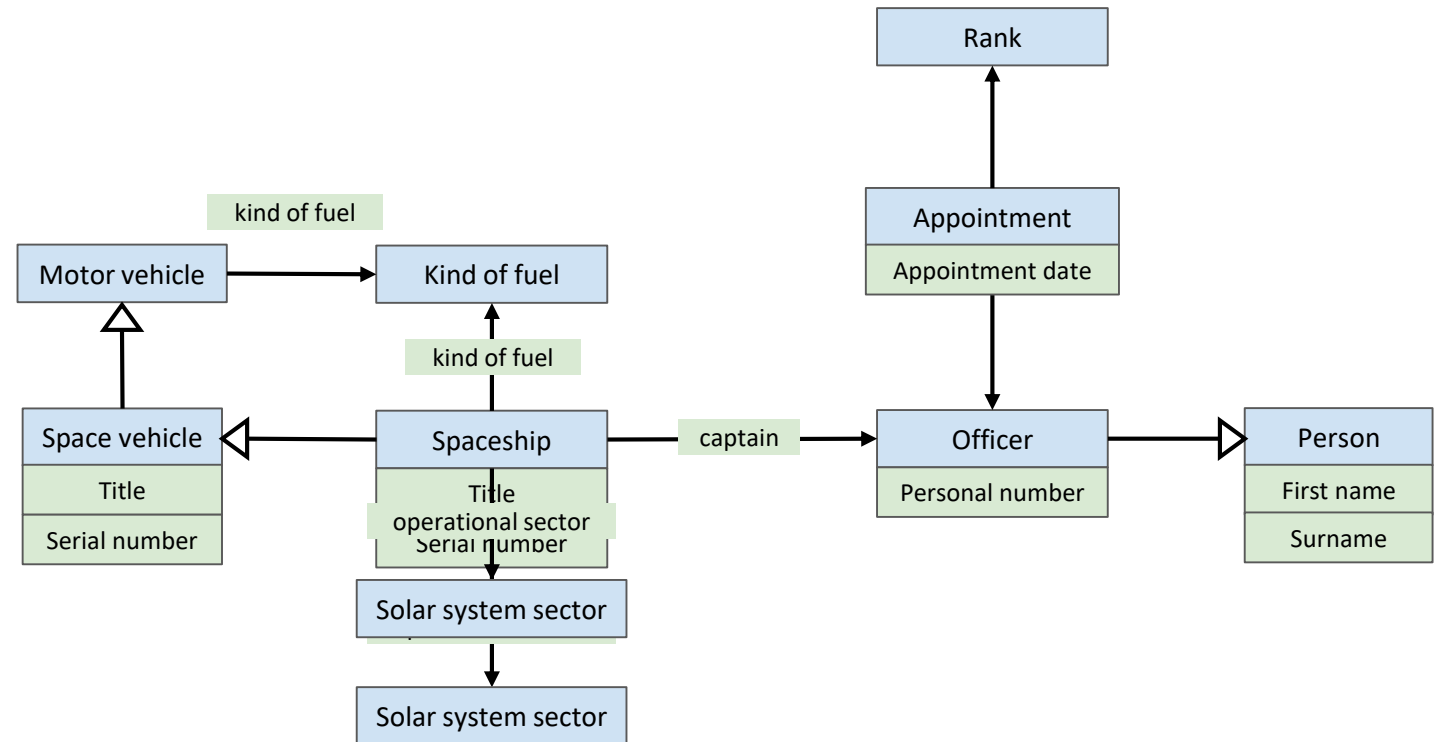
Observations

- Registry of rights and obligations is a data model.
 - It describes data processed and stored by public authorities in their information systems.
- Registry of rights and obligations is a data model with elements of a conceptual model.
 - It is expressed at the conceptual level, but something is missing.

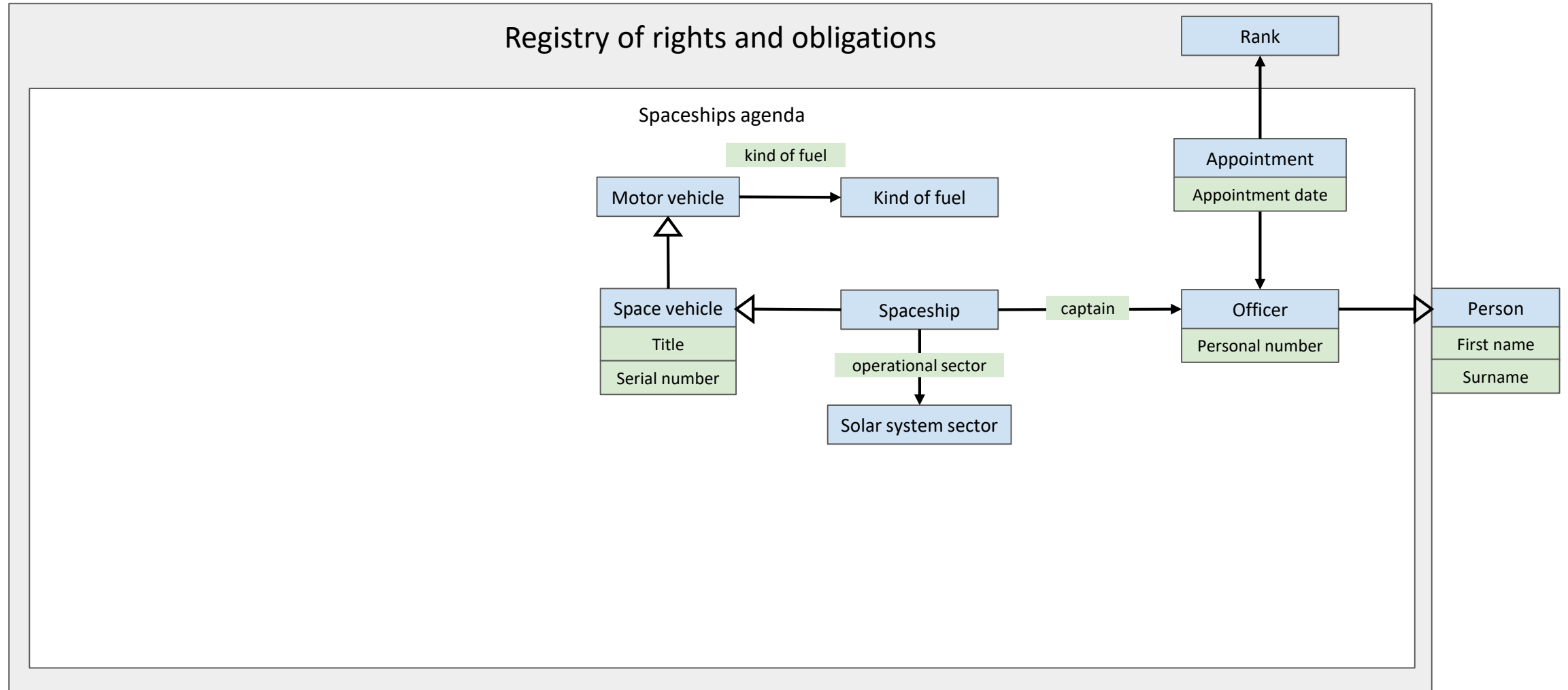
What is missing?



What is missing?



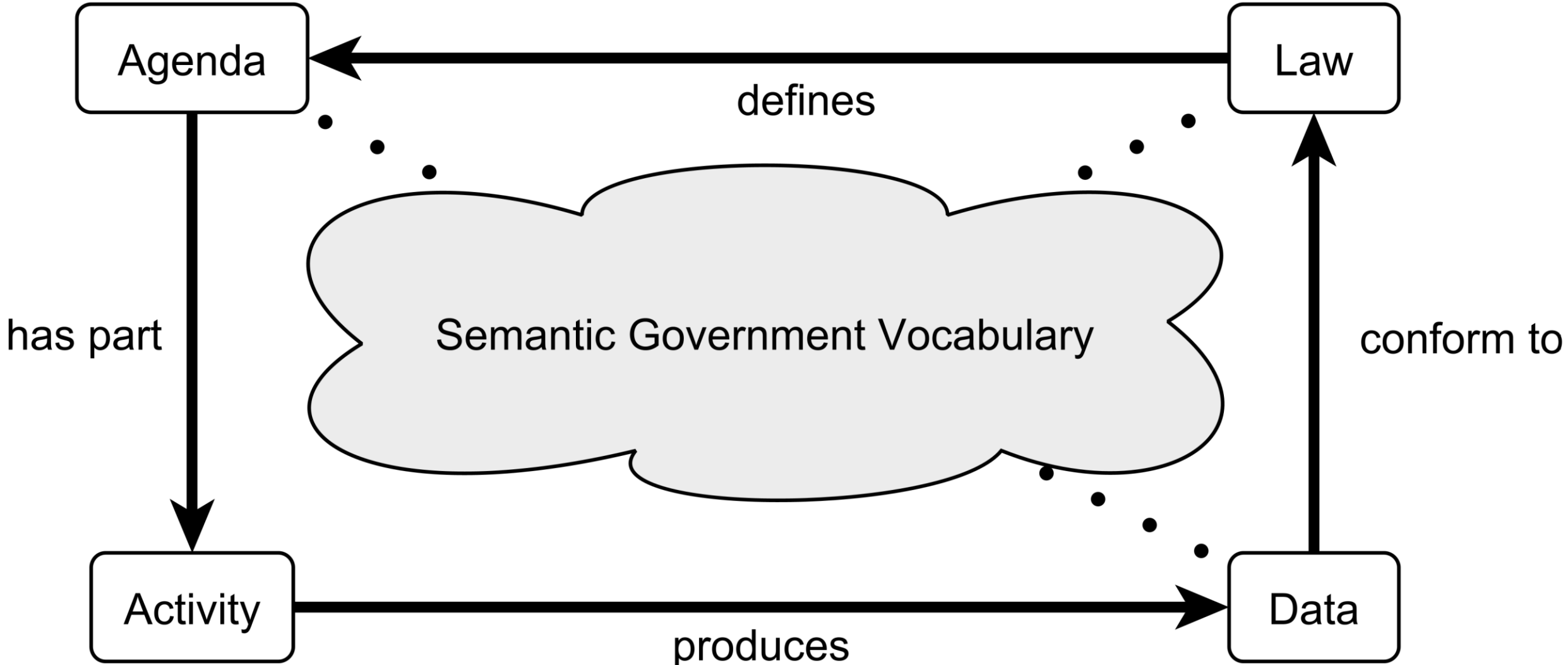
Conceptual model



Formalized and shared



Semantic Government Vocabulary (SGOV)



SGOV architecture

Basic Vocabulary
(Z-SGoV)

Vocabulary of Public Sector
(V-SGoV)

Vocabulary of Legislation
(L-SGoV)

Vocabulary of Agenda
(A-SGoV)

Vocabulary of Dataset
(D-SGoV)

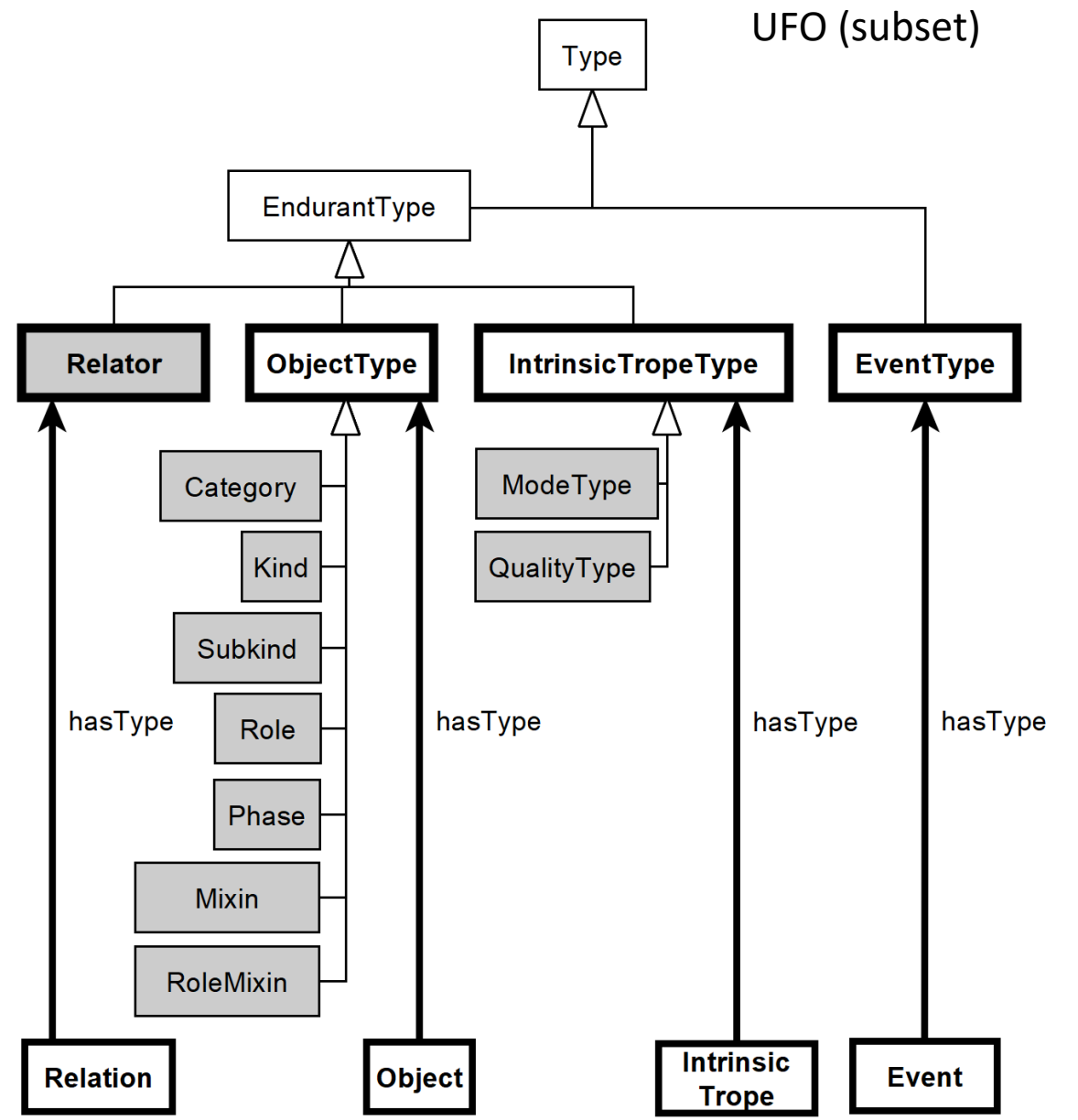
Basic Vocabulary
(Z-SGoV)

Vocabulary of Public Sector
(V-SGoV)

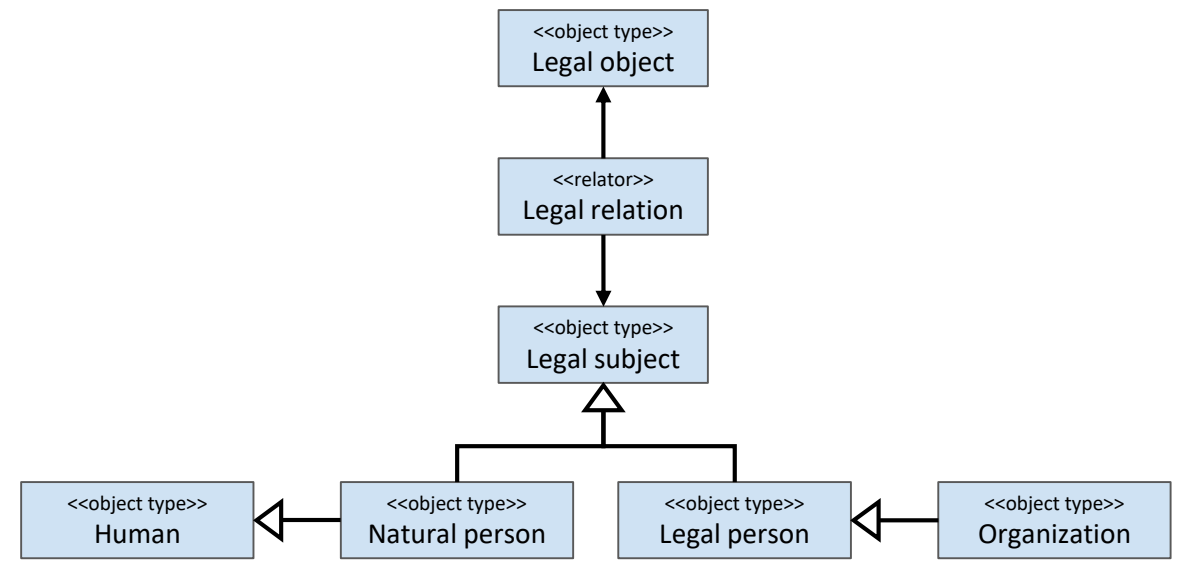
Vocabulary of Legislation
(L-SGoV)

Vocabulary of Agenda
(A-SGoV)

Vocabulary of Dataset
(D-SGoV)



- Basic Vocabulary (Z-SGoV)
- Vocabulary of Public Sector (V-SGoV)
- Vocabulary of Legislation (L-SGoV)
- Vocabulary of Agenda (A-SGoV)
- Vocabulary of Dataset (D-SGoV)



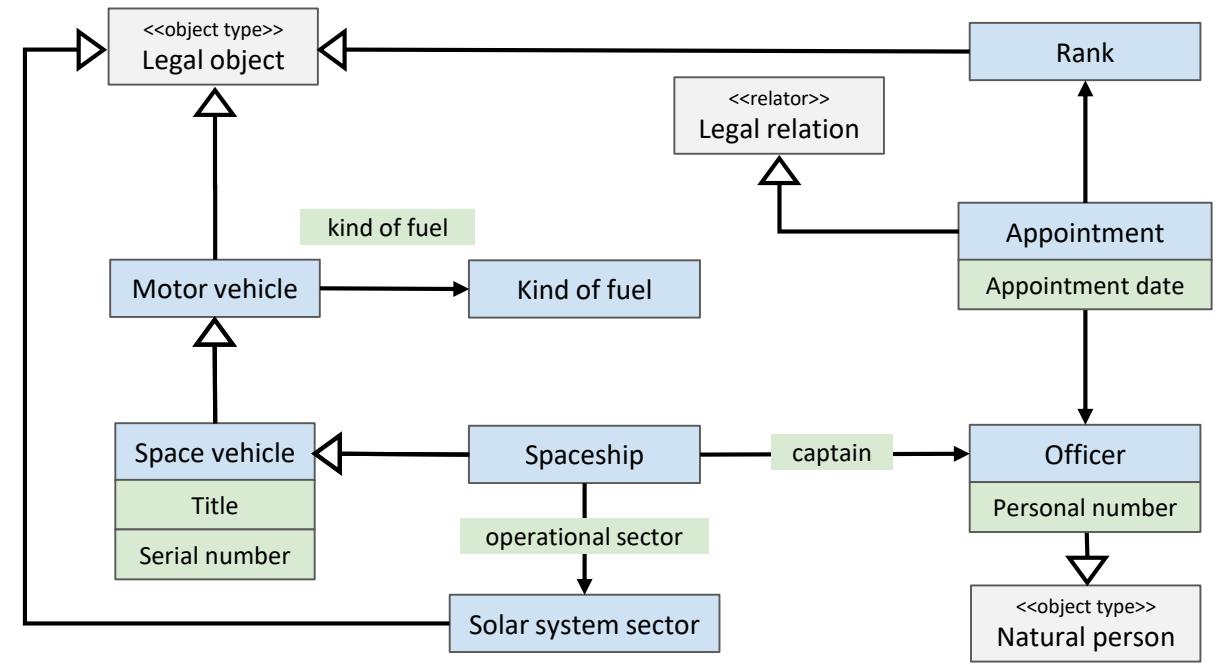
Basic Vocabulary
(Z-SGoV)

Vocabulary of Public Sector
(V-SGoV)

Vocabulary of Legislation
(L-SGoV)

Vocabulary of Agenda
(A-SGoV)

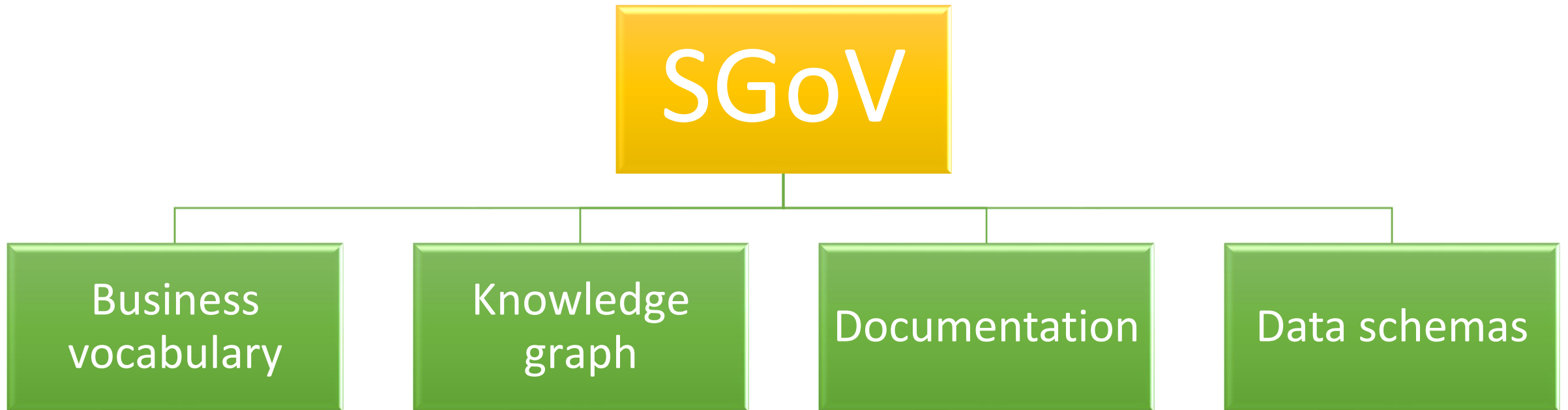
Vocabulary of Dataset
(D-SGoV)



Benefits

- Formally defined vocabulary of governmental business concepts from different domains
- Domain vocabularies (L-SGoV, A-SGoV) can be managed in a federated manner
- Domain vocabularies defined within a shared context of V-SGoV
- A domain vocabulary can be defined by reusing concepts from other vocabularies
- Concepts related to each other – for each concept related concepts are known (generalizations, specializations, relators, tropes, events, ...)

Benefits



Let's see the benefits live.

<https://slovník.gov.cz>

(Assembly line (TermIt, OntoGrapher), ShowIt, NODC, DataSpecer)

Is it a success?

- Working assembly line prototype for building SGoV in a federated manner (Termit + Ontographer)
- Follow-up tools working on top of SGoV (ShowIt, NODC, Dataspecer)
- Working public domain slovník.gov.cz
- Methodologies
- Training courses
- Several public organizations using the tools

So, yes, this seems like a success. But ...

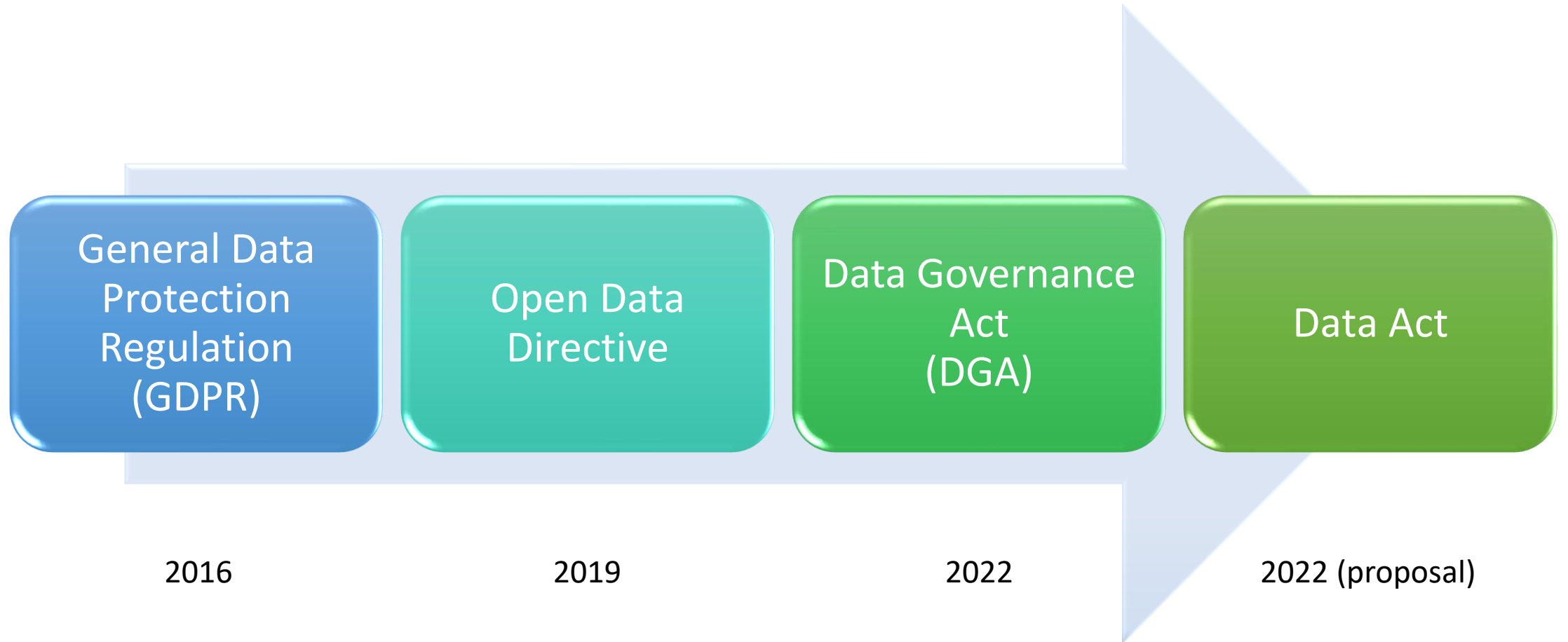
The general adoption of our approach is hard to achieve. Why?

- Of course, I could say that it costs money and human resources we do not have. But this would be too easy.
- We have money and hard-working people, but they are heading in the wrong direction!
 - People do not consider data as a first-class citizen.
 - Front-end features and services are what matter.
 - People naturally design data exchange as point-to-point but do not see the enormous technical debt they leave behind.
- These problems are blockers. They are not only engraved in the heads of foot soldiers but also captains and decision-makers.
- There is no focus on data in the current reorganization efforts.
- Suppliers do not help.

... and it is going to get worse



EU legislation and data



Self-reflection

What about us, teachers and academicians, do we teach so that new generations are ready to face these challenges?

Thank you!

Martin Nečaský

martin.necasky@matfyz.cuni.cz

Department of software engineering
Faculty of mathematics and physics
Charles University

Back-up slides

More on data-related EU legislation ...



General Data Protection Regulation (GDPR)

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

<http://data.europa.eu/eli/reg/2016/679/oj>



General Data Protection Regulation (GDPR)

- introduces individual's control over their data.
- provides a legal basis for processing personal data
- defines rights of the subjects of the data
 - right of access
 - right of erasure
- strict restrictions on how personal data can be shared



Open Data Directive

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information

<http://data.europa.eu/eli/dir/2019/1024/oj>



Open Data Directive

- provides common rules for publishing and sharing government-held documents (data) in the EU
- introduces the legal concepts of machine readability, open format, formal open standard, standard license, and Open Data
- encourages public sector bodies to produce and make available documents (data) in accordance with the principle “open by design and by default”



Open Data Directive

Public sector bodies shall make their documents (data) available in formats that are open, machine-readable, accessible, findable, and re-usable, together with their metadata.

Both the format and the metadata shall, where possible, comply with formal open standards



Open Data Directive

Public sector bodies shall make their documents (data) available in formats that are open, machine-readable, accessible, findable, and re-usable, together with their metadata.

Both the format and the metadata shall, where possible, comply with formal open standards.

F

A

I

R



Open Data Directive

- introduces high-value datasets in the following thematic categories
 - Geospatial, Earth observation and environment, Meteorological, Statistics, Companies and company ownership, Mobility
- high-value datasets must be published mandatorily by all Member States as Open Data
- high-value datasets are defined by European Commission
 - the first version is expected in the second half of October 2022



Data Governance Act (DGA)

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act)

<http://data.europa.eu/eli/reg/2022/868/oj>



Data Governance Act (DGA)

- aims to make more data available by regulating the re-use of
 - publicly held protected data (personal data and commercially confident data)
 - privately protected data provided voluntarily by individuals and companies (data altruism)
- introduces the concept of domain-specific common European data spaces
 - health, mobility, manufacturing, financial services, energy, agriculture, etc.
 - or combination of such areas, e.g., energy and climate as well as thematic areas such as the European Green Deal



Data Governance Act (DGA)

Common European data spaces should make data findable, accessible, interoperable, and reusable (the 'FAIR data principles') while ensuring a high level of cybersecurity.

Member States shall ensure that all relevant information is available and easily accessible through a single information point. The functions of the single information point may be automated.



Data Act

Proposal for a REGULATION OF THE
EUROPEAN PARLIAMENT AND OF THE
COUNCIL on harmonised rules on fair
access to and use of data (Data Act)

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN>



Data Act

- lays down a framework specifying who, other than the manufacturer or other data holder is entitled to access the data generated by products
- ensures that users of a product in the EU can access the data generated by the use of that product, including by sharing them with third parties of their choice
- does not explicitly enforce FAIR principles, but strongly enforces interoperability