Data-CASE:

Grounding Data Regulations for Compliant Data Processing Systems

Vishal Chakraborty, Stacy Ann-Elvy, Sharad Mehrotra, Faisal Nawab, Mohammad Sadoghi,
Shantanu Sharma, Nalini Venkatasubramanian, Farhan Saeed

March 28, 2024

EDBT '24 : Session 8

Paestum, Italy

Data Regulations Timeline

Enactment/Effective Dates

2018

General Data Protection Regulation

2019

Nevada SB220

CA Consumer Privacy Act

Personal Data Protection Bill

2020

Brazilian DPR

CA Privacy Rights Act

EU Data Governance Act

2021

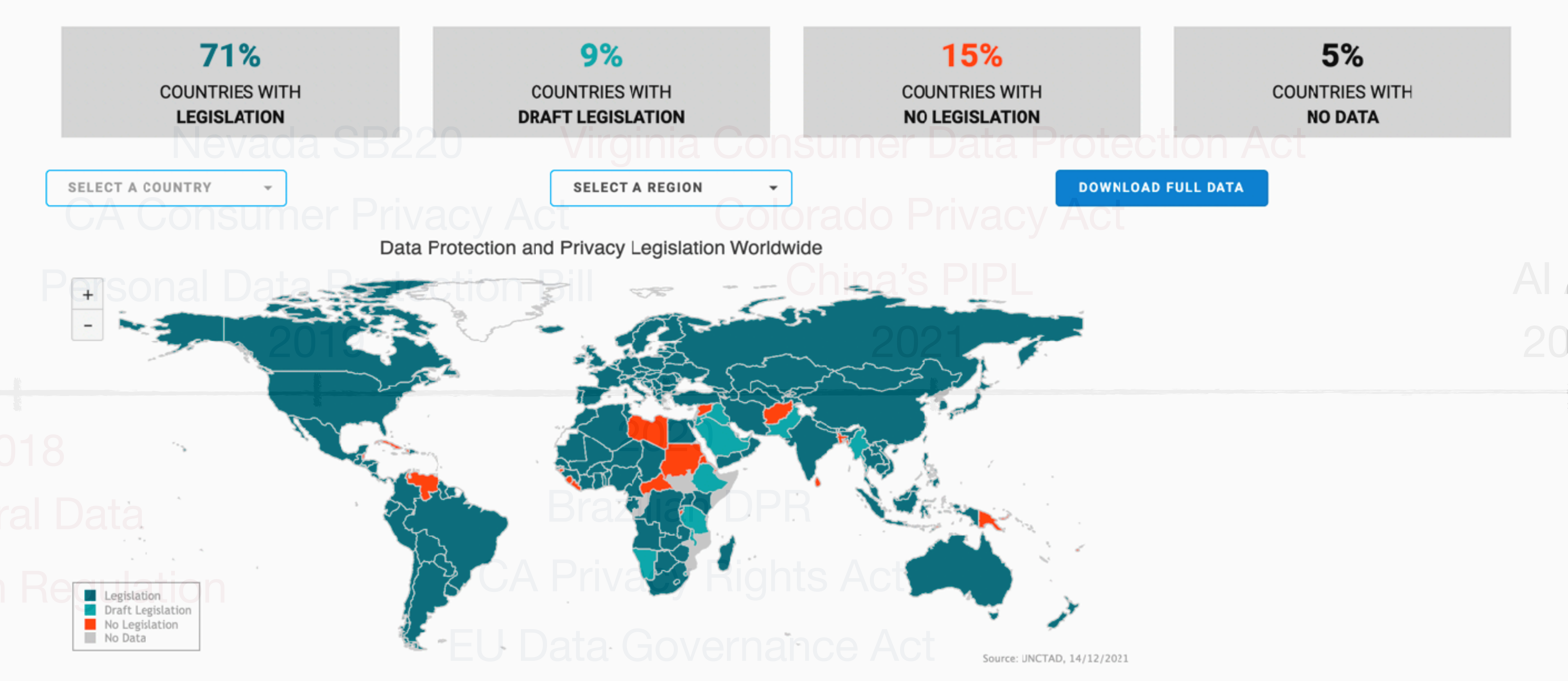
Virginia Consumer Data Protection Act

Colorado Privacy Act

China's PIPL

AI Act

2023



71%	9%	15%	5%
COUNTRIES WITH **LEGISLATION**	COUNTRIES WITH **DRAFT LEGISLATION**	COUNTRIES WITH **NO LEGISLATION**	COUNTRIES WITH **NO DATA**

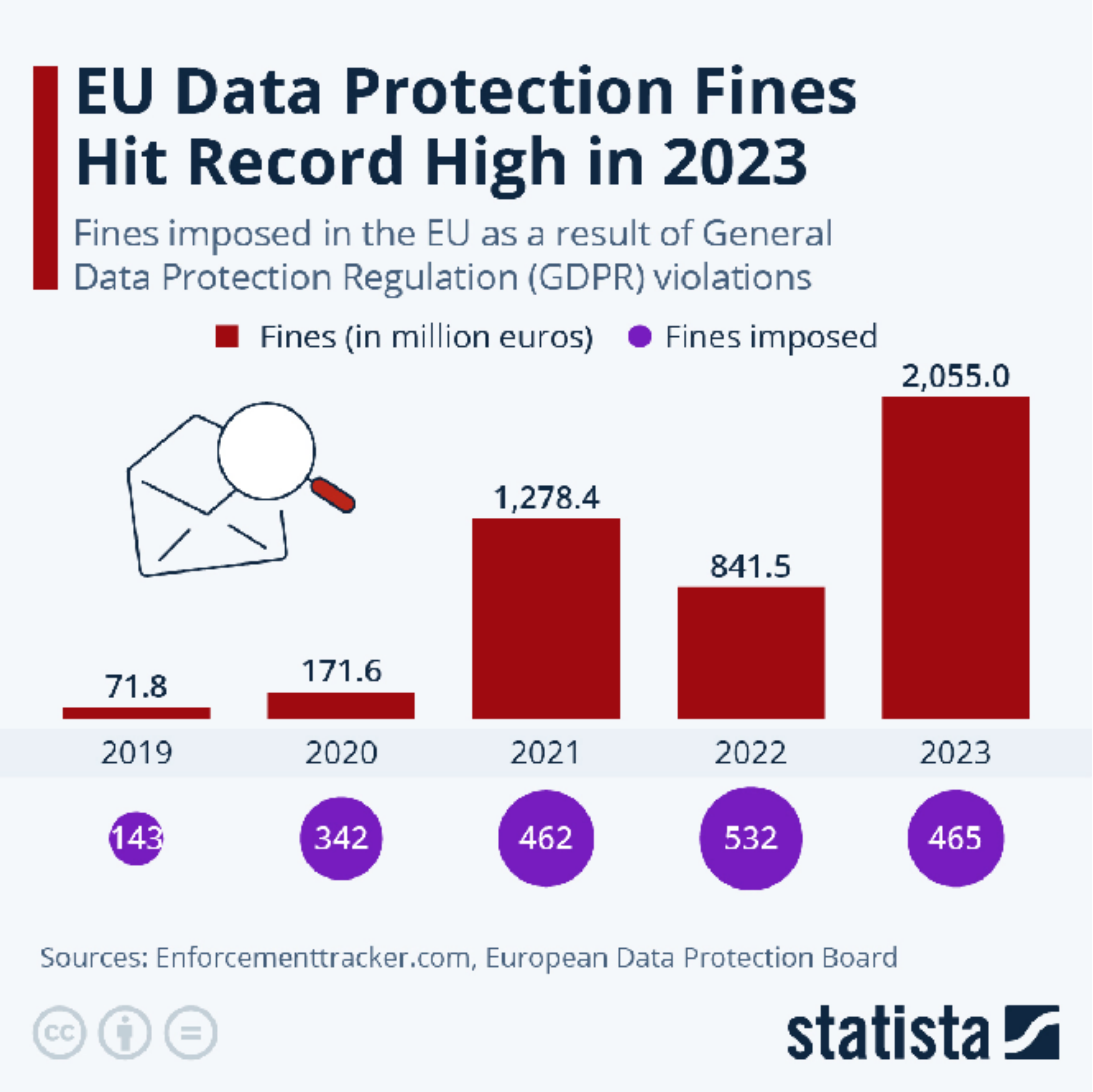
Data Protection and Privacy Legislation Worldwide

- 🟦 Legislation
- 🟦 Draft Legislation
- 🟥 No Legislation
- ⬜ No Data

Source: UNCTAD, 14/12/2021

Keeping-up with The Data Regulations

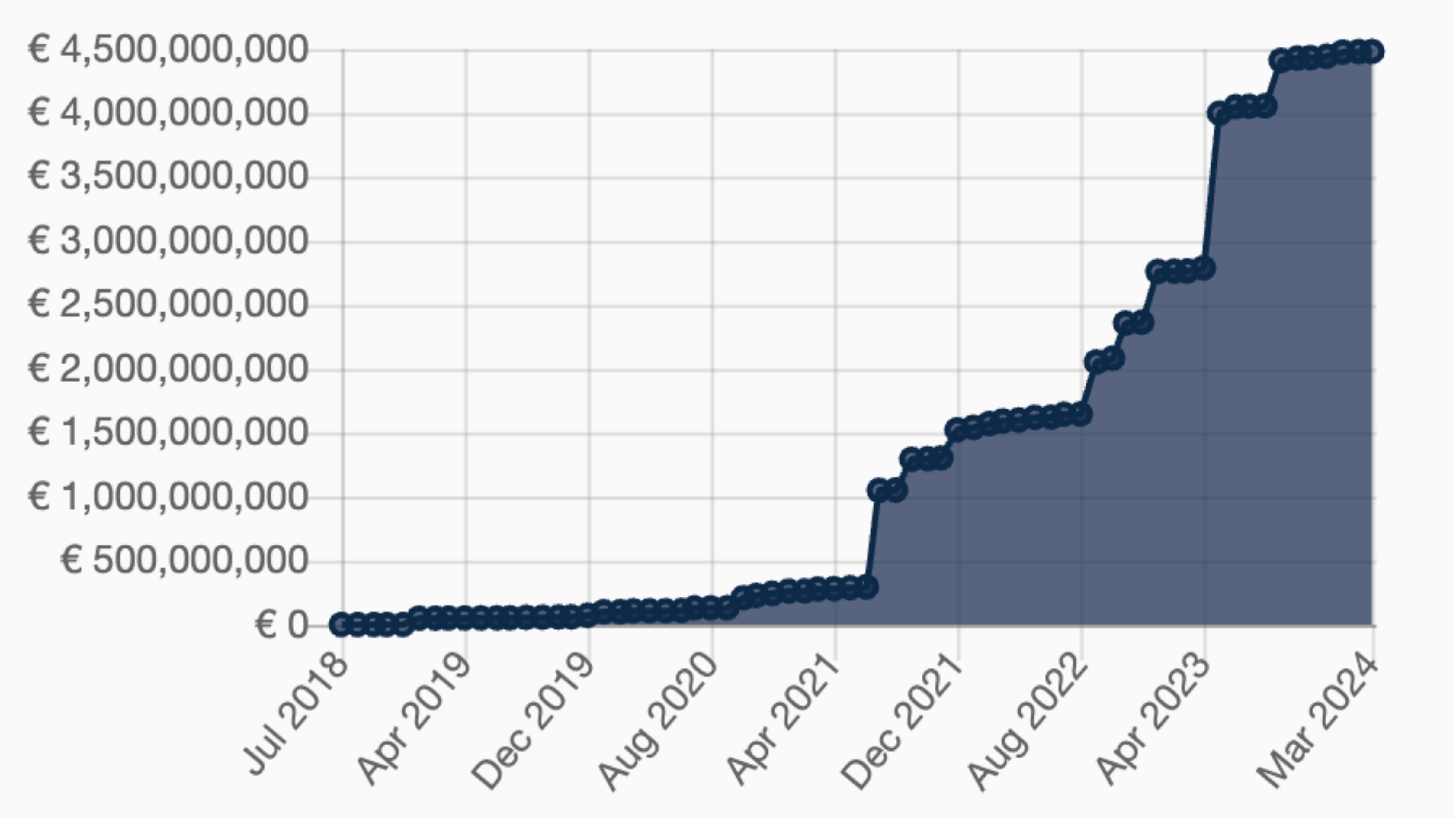
Violations at A Glance



EU Data Protection Fines Hit Record High in 2023

Fines imposed in the EU as a result of General Data Protection Regulation (GDPR) violations

Sources: Enforcementtracker.com, European Data Protection Board



https://www.enforcementtracker.com/?insights (March 27, 2024)

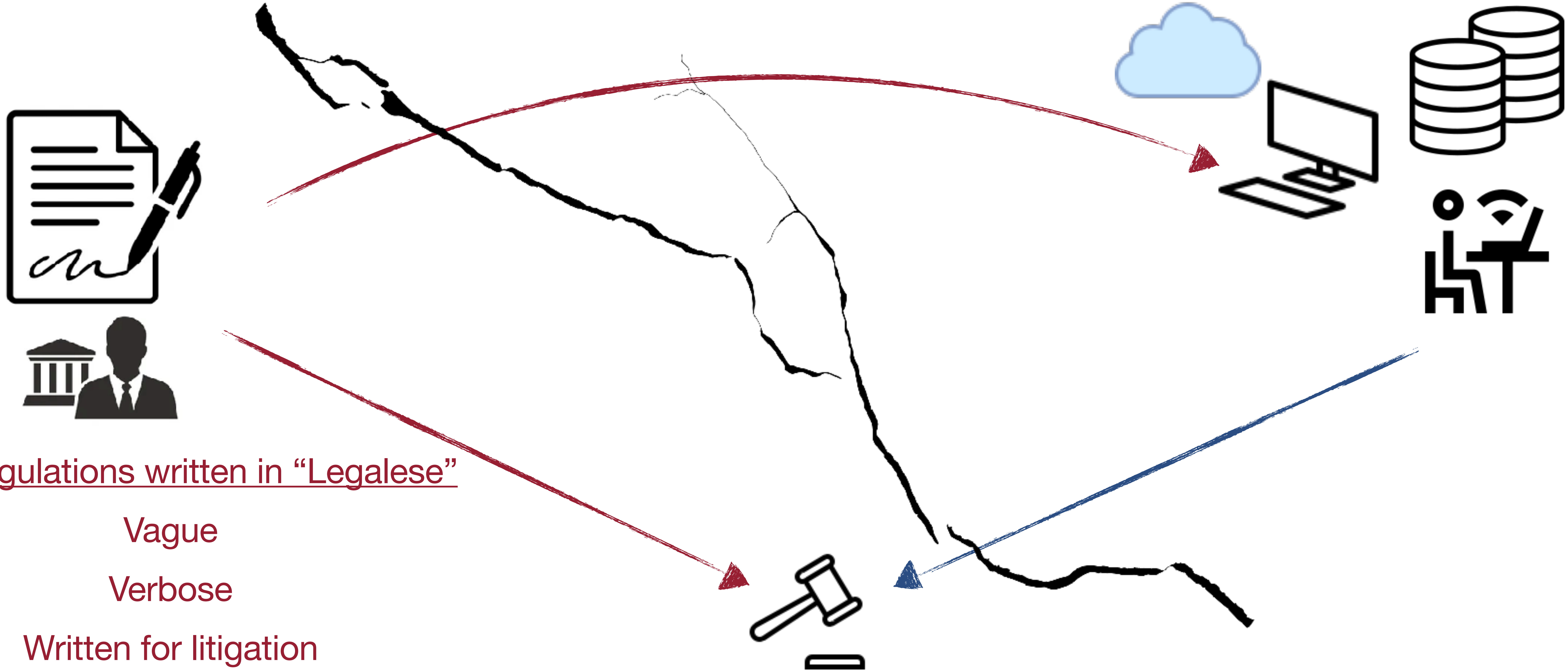
The Great Divide

System-actions and control-paths

Well-defined

Technical

Implement systems



Data-regulations written in “Legalese”

Vague

Verbose

Written for litigation

Example

Right to Erasure

"… shall have the right to obtain from the controller the erasure of personal data concerning him or her without undue delay and the controller shall have the obligation to erase personal data without undue delay… " Art. 17, GDPR

What is erasure?
Which data concerns the subject?
How much is undue delay?

Database Design Challenges

Data regulations are written for litigation

Data Regulations

Implement data- and control-paths

- Too many regulations with too may (varying) requirements
- Ambiguity [19]
- Article 29 Data Protection Working Party - GDPR [12]
- Recommendations have been unsound [19, 53]
- Pitted against industry practices [70, 71]
- Resource intensive [68]

Goal

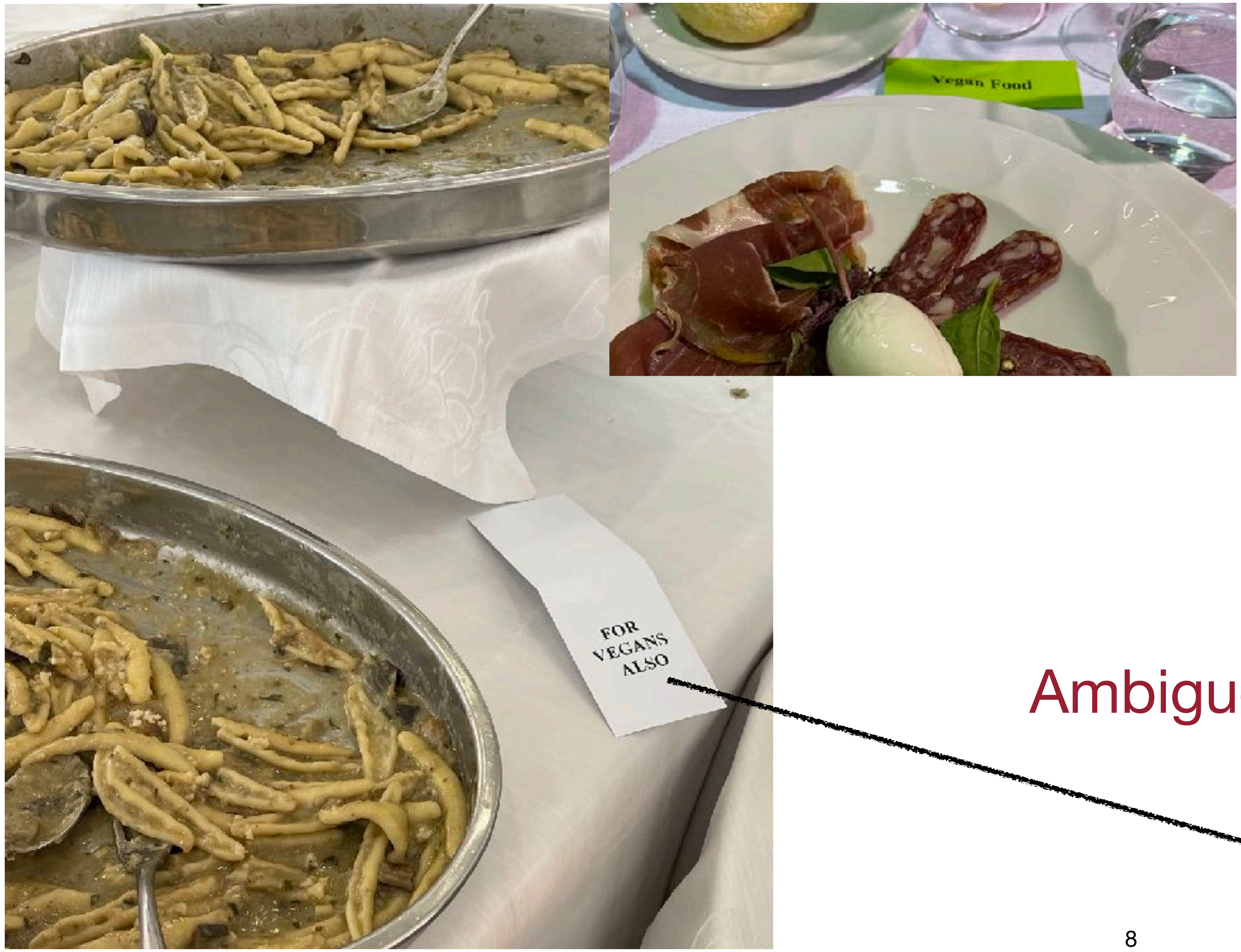
Vision

Ambiguous legal specifications

Data-CASE

Grounded (system-level)
technical specifications

High Level Idea - From dinner last night!



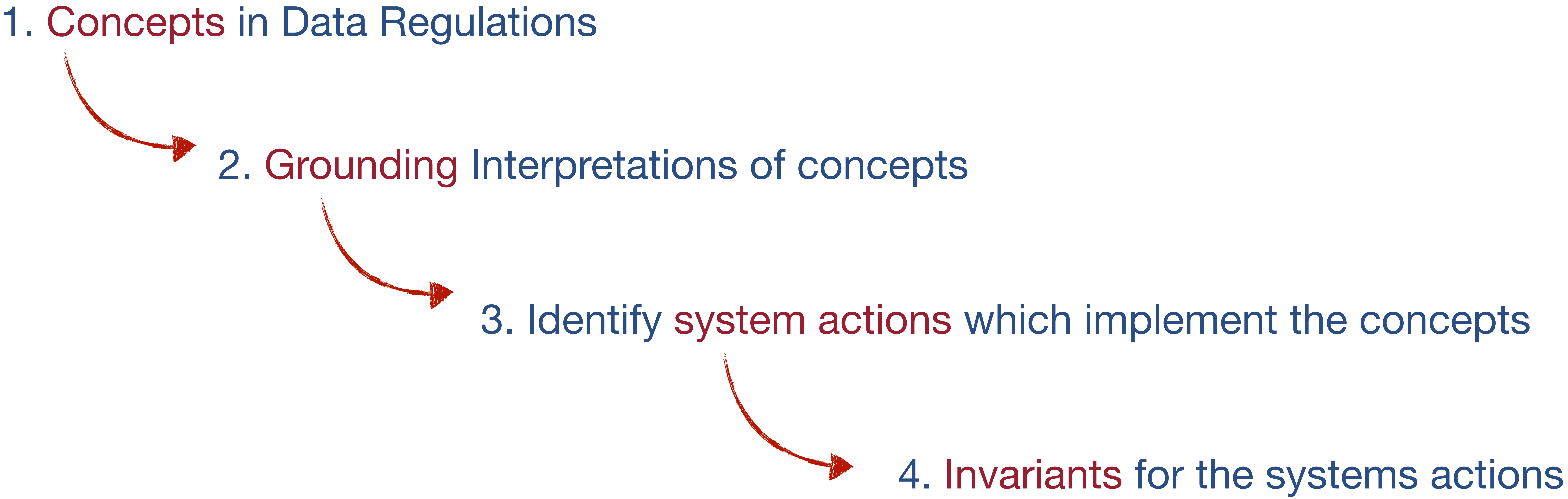
- Vegan
    - No animal products/derived
- Vegetarian
    - No meat
    - Includes eggs, dairy
    - Includes fish(?)

Ambiguous. Use simple, well defined concepts!

Contains: Eggs, Dairy, …

Steps In Data-CASE

Process



1. Concepts in Data Regulations

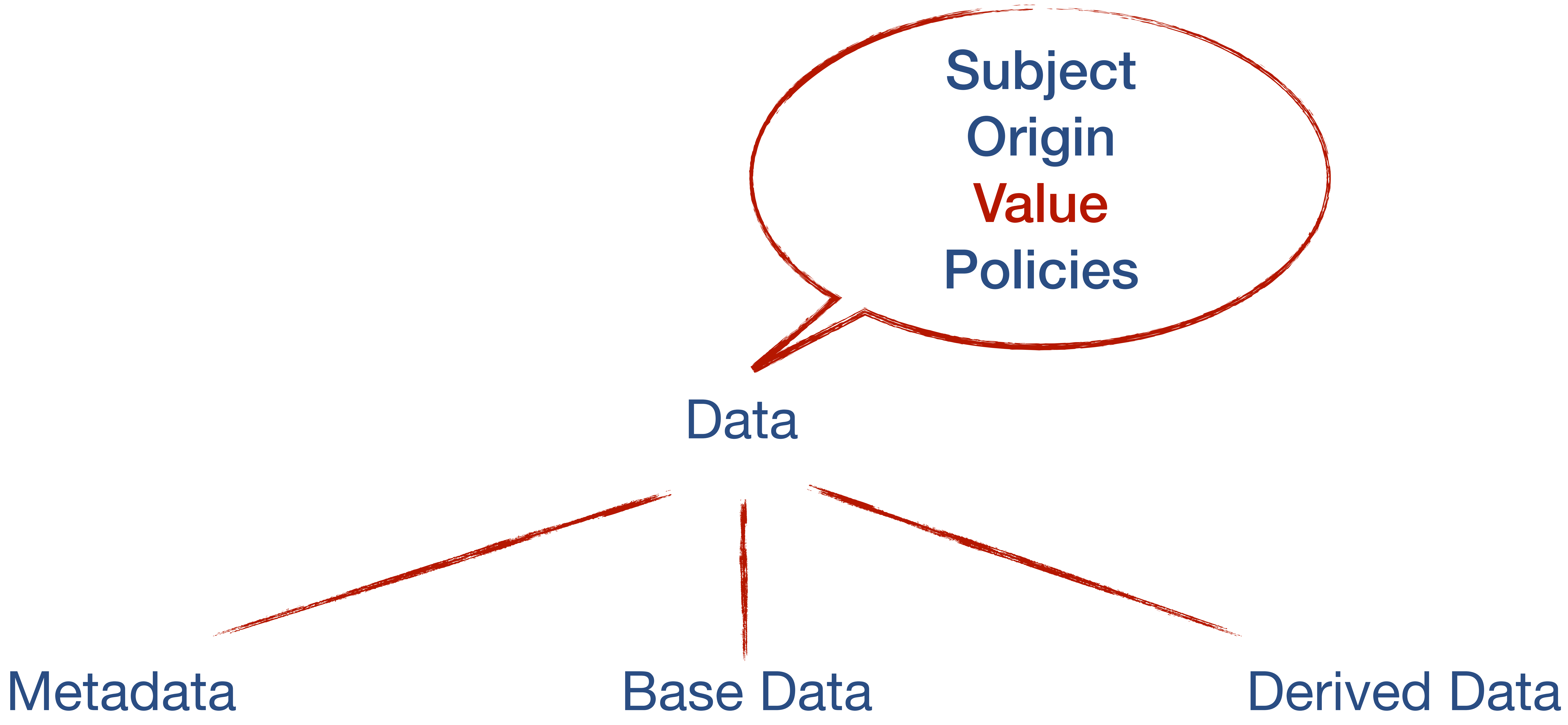
2. Grounding Interpretations of concepts

3. Identify system actions which implement the concepts

4. Invariants for the systems actions

1.Concepts in Data-CASE

Data



Subject
Origin
Value
Policies

Data

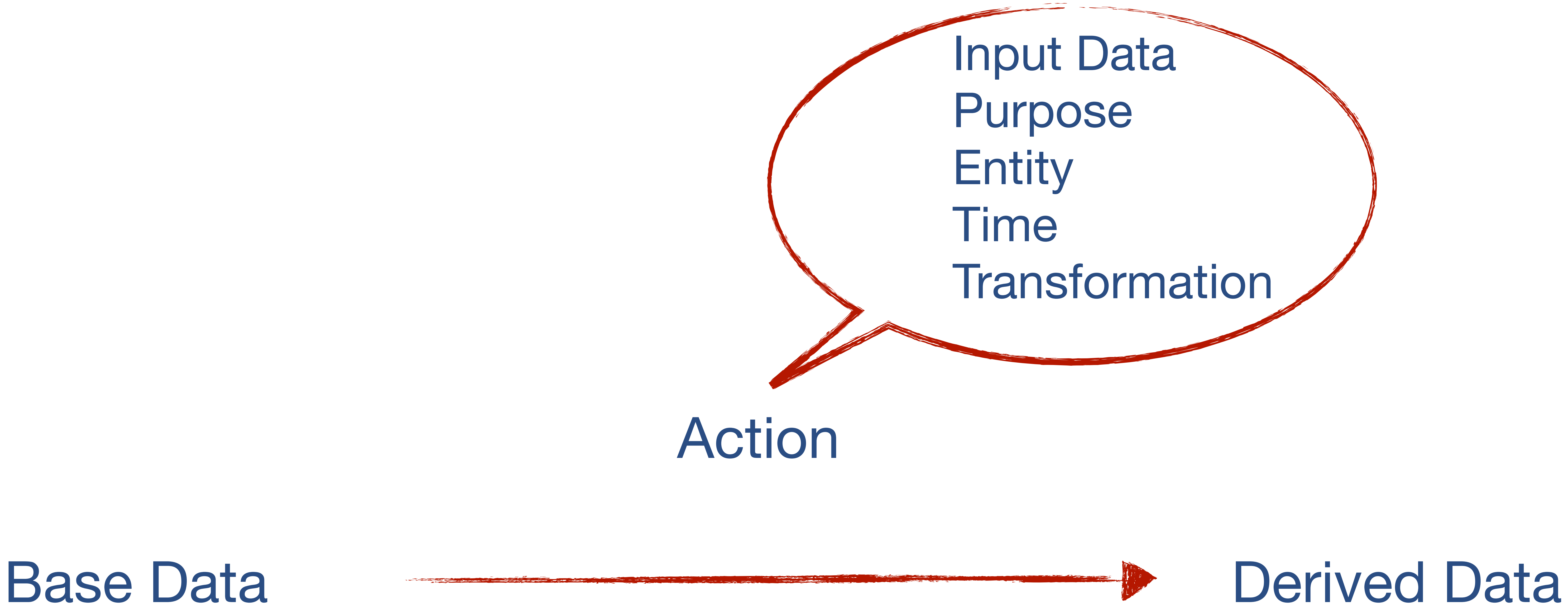
Metadata

Base Data

Derived Data

1.Concepts in Data-CASE

Actions and action-history



Input Data
Purpose
Entity
Time
Transformation

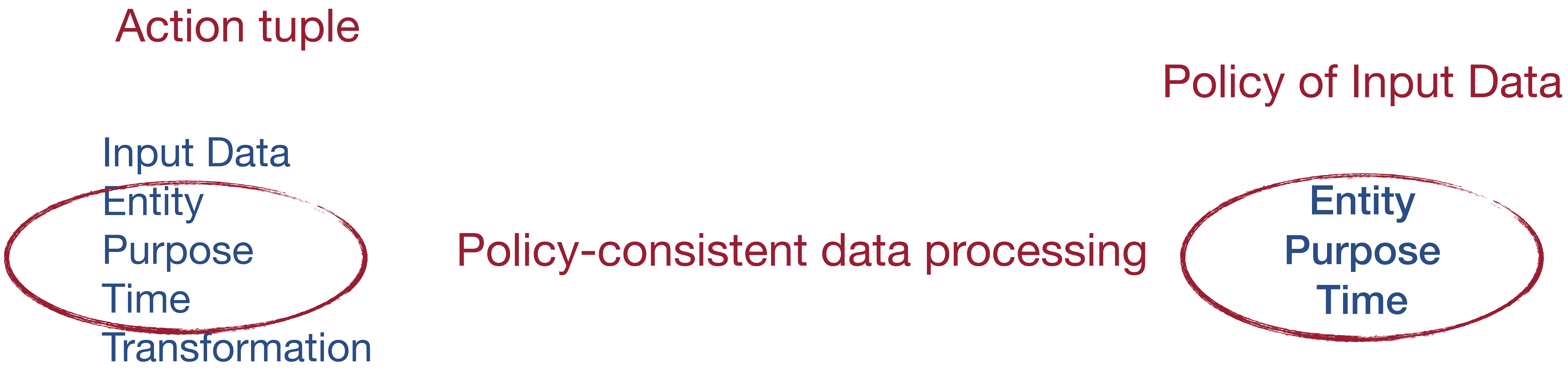
Action

Base Data

Derived Data

1.Concepts in Data-CASE

Consistent Data processing



Action tuple

Input Data
Entity
Purpose
Time
Transformation

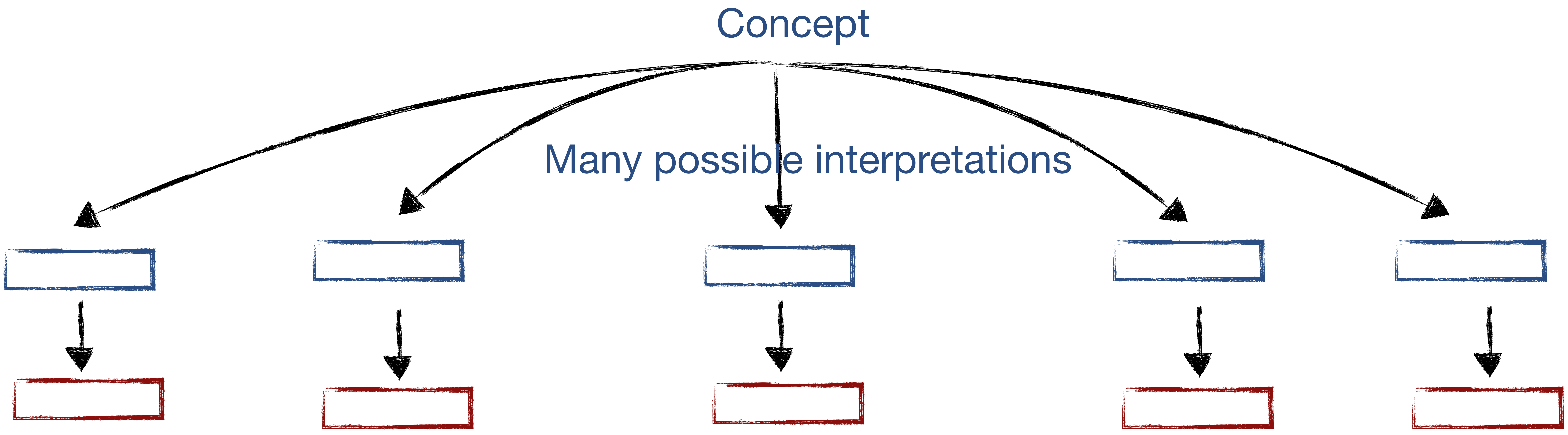
Policy-consistent data processing

Policy of Input Data

Entity
Purpose
Time

2. Grounding Concepts

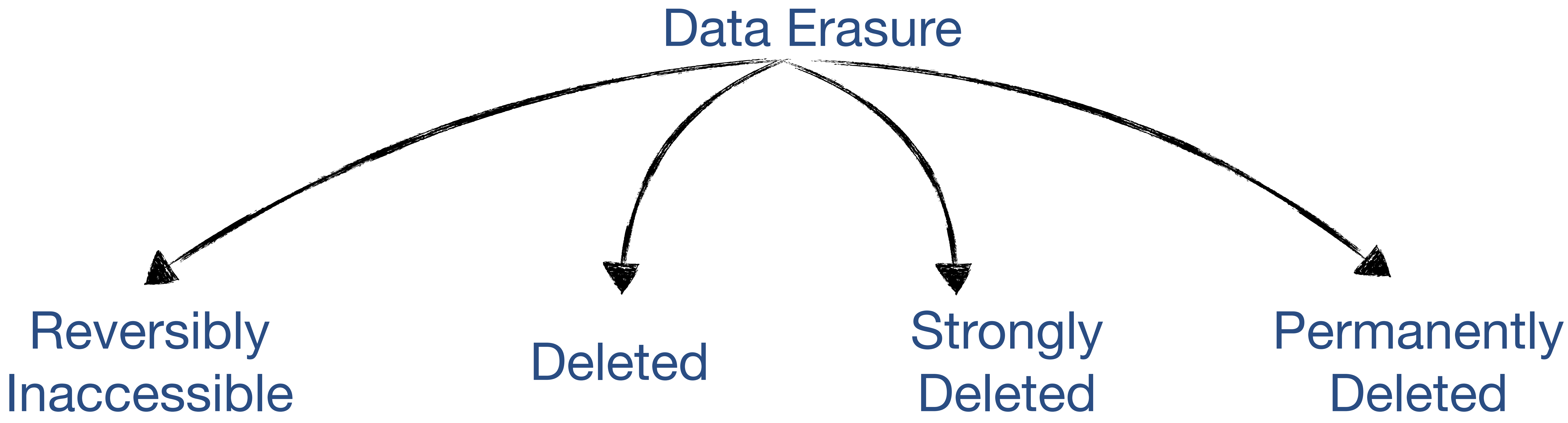
Fixing an interpretation



Grounded Concept

Technically sated.
Unambiguous interpretation.

Example of Grounding: Erasure



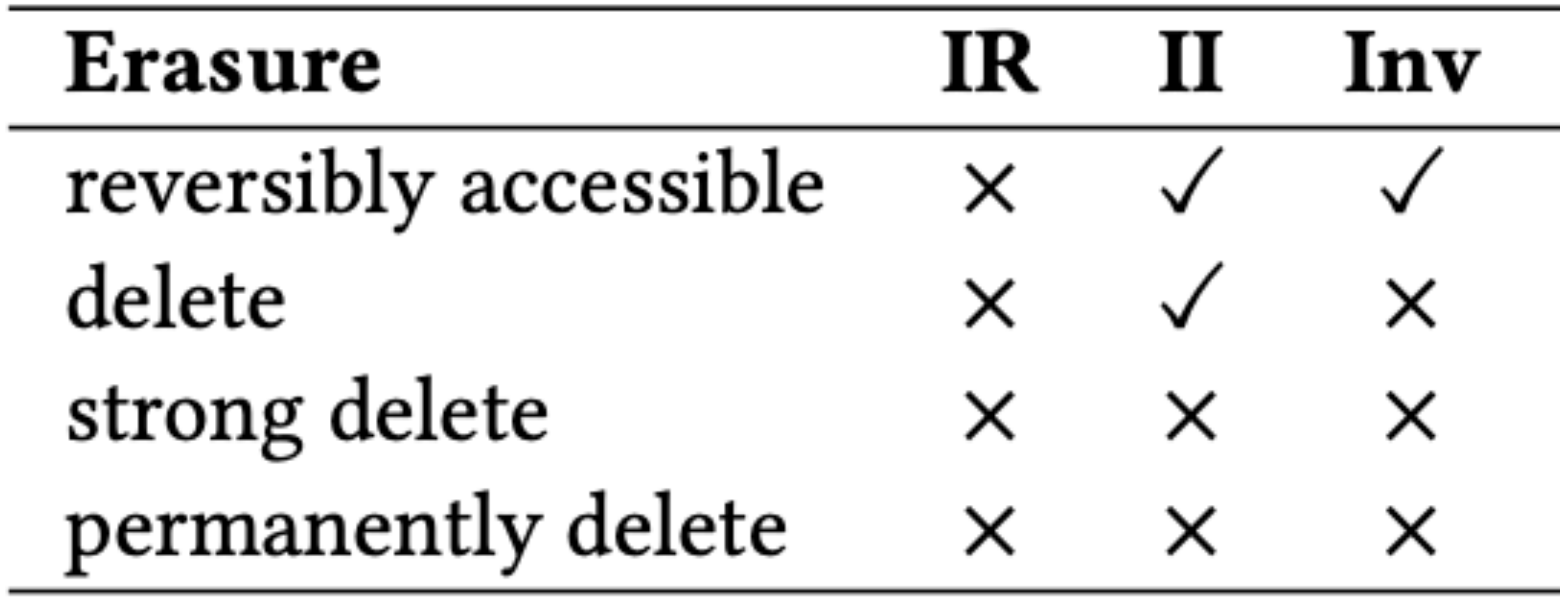
Data Erasure

Reversibly Inaccessible

Deleted

Strongly Deleted

Permanently Deleted

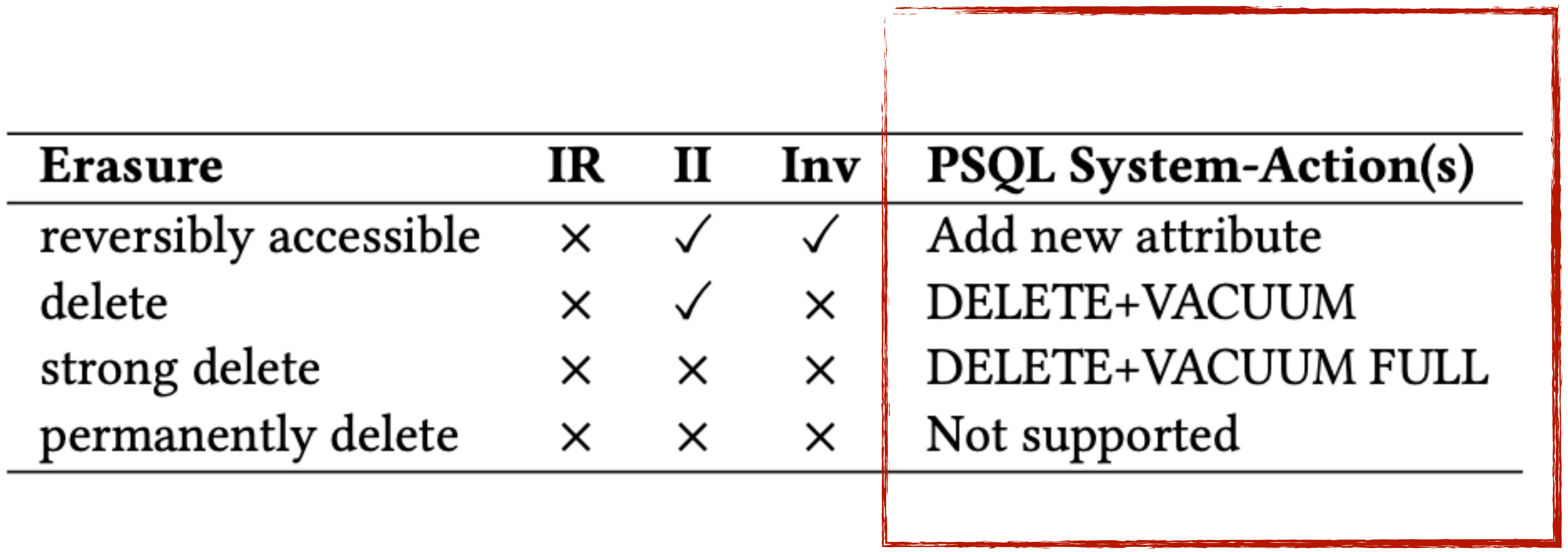


Erasure	IR	II	Inv
reversibly accessible	×	✓	✓
delete	×	✓	×
strong delete	×	×	×
permanently delete	×	×	×

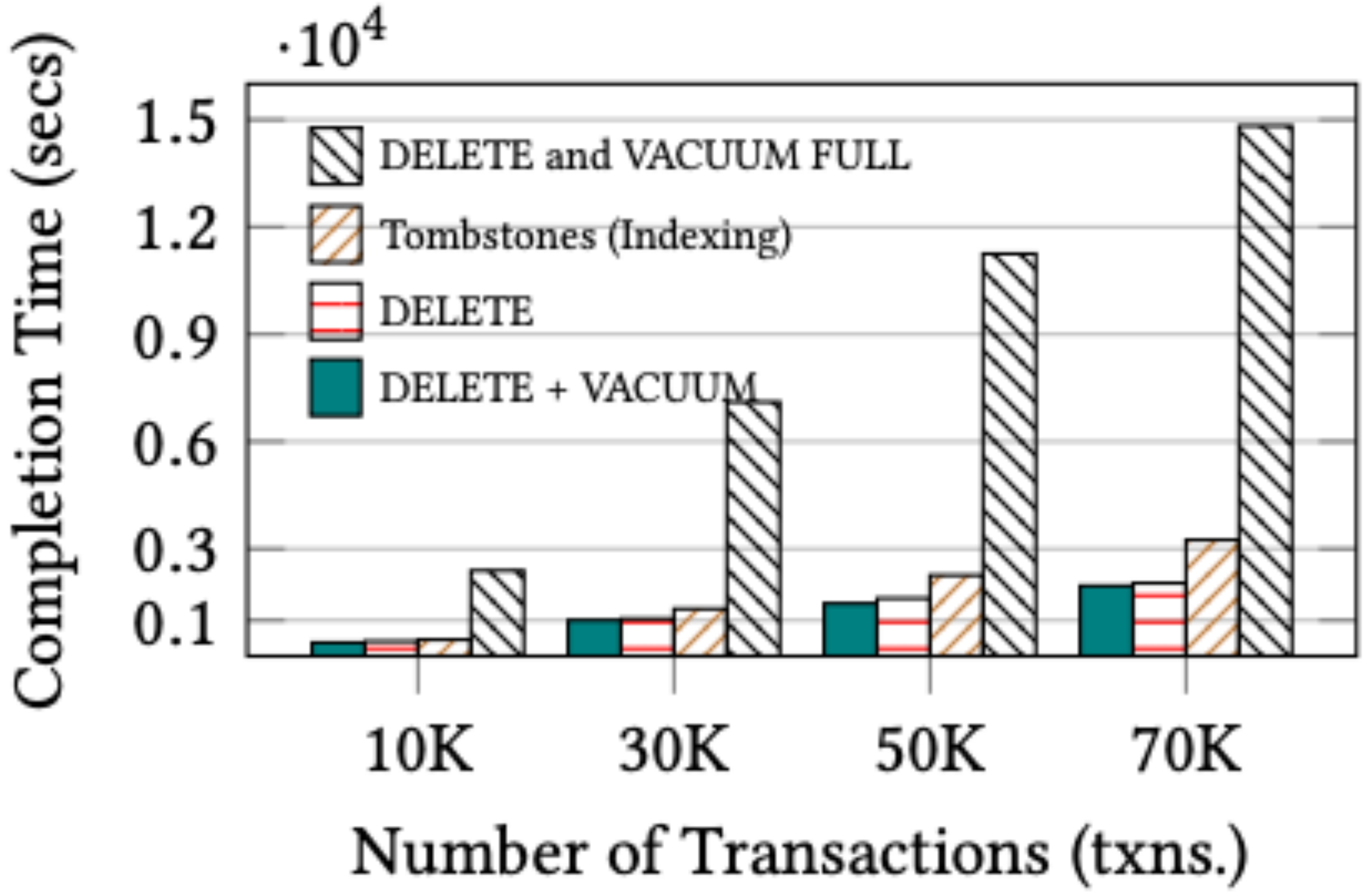
3. System Actions For Groundings

From grounded concepts to system actions

- System actions define the grounded concepts for a given system.



Erasure	IR	II	Inv	PSQL System-Action(s)
reversibly accessible	×	✓	✓	Add new attribute
delete	×	✓	×	DELETE+VACUUM
strong delete	×	×	×	DELETE+VACUUM FULL
permanently delete	×	×	×	Not supported



4. Invariants

Formal properties

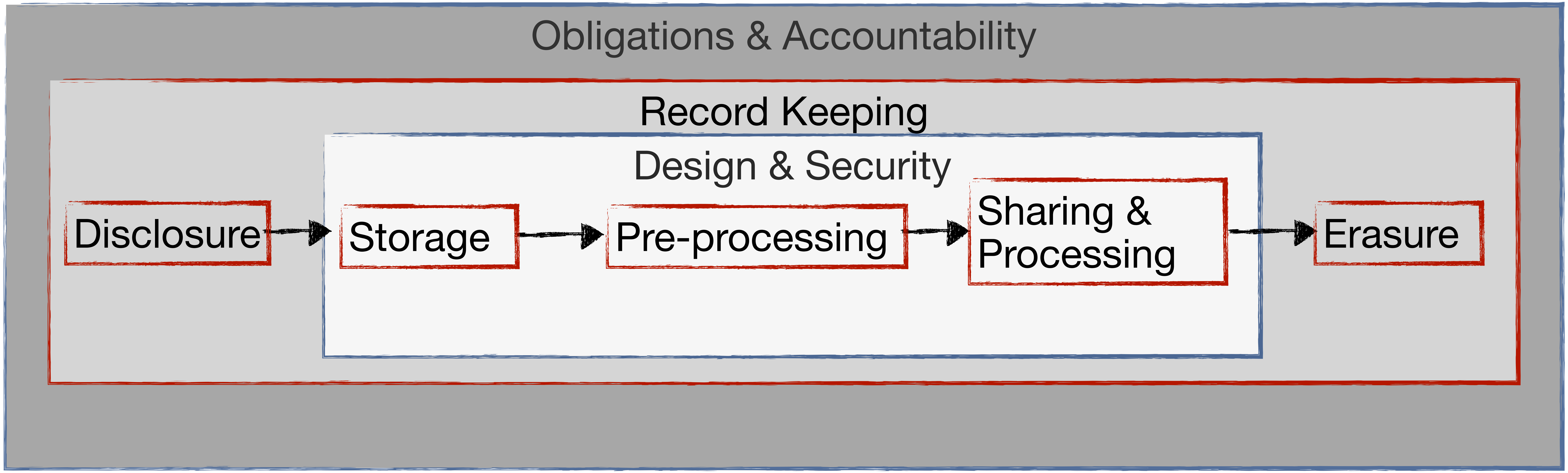
- Characterize system actions with formal invariants that must hold in the system.
  - Think: “When” and “how”?

$$\forall X \,.\, erasure\_req(subject_X, X, t) \implies erase(x, [t, t+\delta])$$

grounded and mapped to system actions

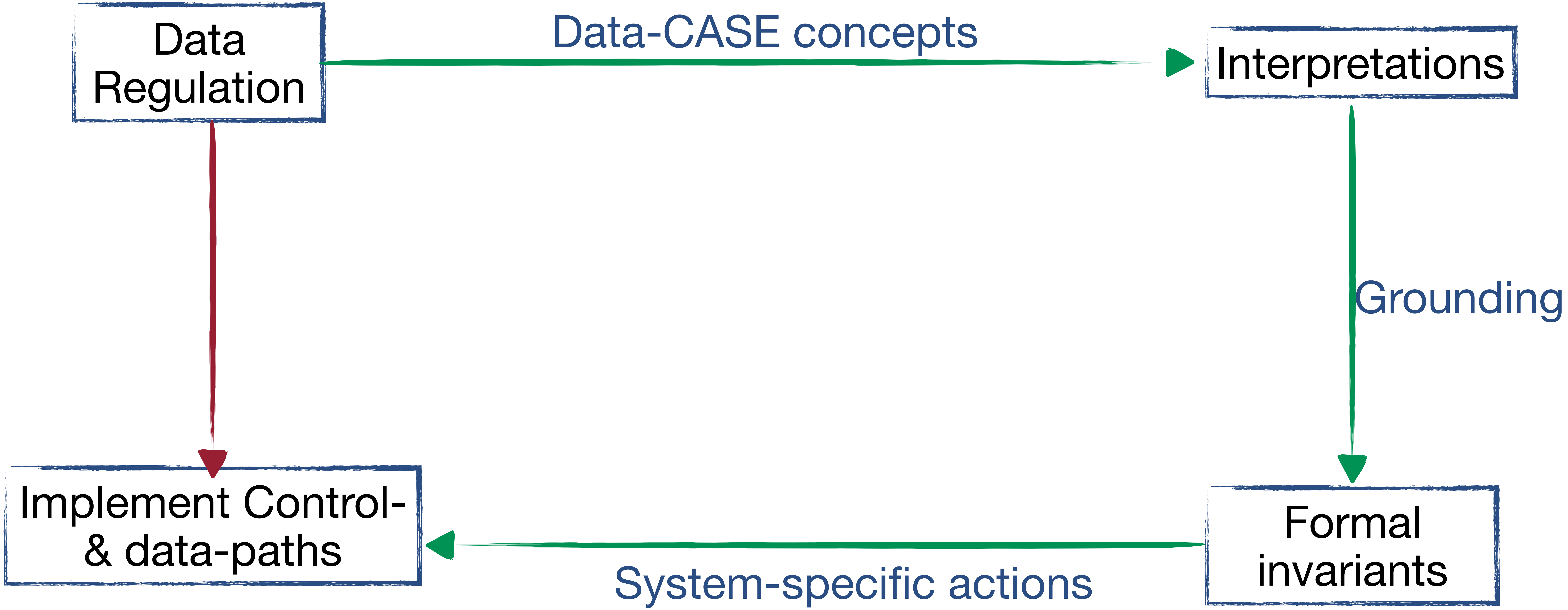
How To Come Up With Invariants?

Classification of Data Regulations



Overview

Data-CASE



Uses

Of Data-CASE

See the paper for case studies.

Data Collectors

Database Providers

Data Processors

Service Providers

App developers

Regulators

Regulatory Agencies

Multinational Orgs

Privacy Impact Assessments

Data-
Collection
Access
Storage
Erasure

- Data-CASE makes data regulations amenable for compliant system design
  - **Amenable**: capable of being acted upon in a particular way
- It doesn't determine what's legal and what's not

Questions?

Vishal Chakraborty
vi.c@uci.edu

Funding Sources

U.S. National Science Foundation

HPI Hasso Plattner Institut
Digital Engineering · Universität Potsdam

HPI @ UCI