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
Gener Protection

71%
COUNTRIES WITH LEGISLATION

9%
COUNTRIES WITH DRAFT LEGISLATION

15%
COUNTRIES WITH NO LEGISLATION

5%
COUNTRIES WITH NO DATA

SELECT A COUNTRY

SELECT A REGION

DOWNLOAD FULL DATA

Data Protection and Privacy Legislation Worldwide

- General Data Protection Regulation
- AI Act
- Brazilian DPR
- CA Privacy Rights Act
- Colorado Privacy Act
- Nevada SB220
- China's PIPL
- Virginia Consumer Data Protection Act
- EU Data Governance Act
- Personal Data Protection Bill

Act 23
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

- Fines (in million euros)
- Fines imposed

<table>
<thead>
<tr>
<th></th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
<th>2022</th>
<th>2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fines</td>
<td>71.8</td>
<td>171.6</td>
<td>1,278.4</td>
<td>841.5</td>
<td>2,055.0</td>
</tr>
</tbody>
</table>
| Sources: Enforcementtracker.com, European Data Protection Board

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

Data-regulations written in “Legalese”
  Vague
  Verbose
  Written for litigation

System-actions and control-paths
  Well-defined
  Technical
  Implement systems
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

- Too many regulations with too many (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]

Implement data- and control-paths
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

- Metadata
- Base Data
- Derived Data

Subject
Origin
Value
Policies
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 of Input Data
Policy-consistent data processing

Entity
Purpose
Time
2. Grounding Concepts
Fixing an interpretation

Concept

Many possible interpretations

Grounded Concept
Technically sated.
Unambiguous interpretation.
Example of Grounding: Erasure

Data Erasure

- Reversibly Inaccessible
- Deleted
- Strongly Deleted
- Permanently Deleted

<table>
<thead>
<tr>
<th>Erasure</th>
<th>IR</th>
<th>II</th>
<th>Inv</th>
</tr>
</thead>
<tbody>
<tr>
<td>reversibly accessible</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>delete</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>strong delete</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>permanently delete</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>
3. System Actions For Groundings

From grounded concepts to system actions

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

<table>
<thead>
<tr>
<th>Erasure</th>
<th>IR</th>
<th>II</th>
<th>Inv</th>
<th>PSQL System-Action(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>reversibly accessible</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
<td>Add new attribute</td>
</tr>
<tr>
<td>delete</td>
<td>×</td>
<td>✓</td>
<td>x</td>
<td>DELETE+VACUUM</td>
</tr>
<tr>
<td>strong delete</td>
<td>×</td>
<td>×</td>
<td>x</td>
<td>DELETE+VACUUM FULL</td>
</tr>
<tr>
<td>permanently delete</td>
<td>×</td>
<td>×</td>
<td>x</td>
<td>Not supported</td>
</tr>
</tbody>
</table>

Completion Time (secs)

- DELETE and VACUUM FULL
- Tombstones (Indexing)
- DELETE
- DELETE + VACUUM

Number of Transactions (txns.)
4. Invariants

Formal properties

• Characterize system actions with **formal invariants** that must hold in the system.

• Think: “When” and “how”?

\[ \forall X. \text{erasure}_\text{req}(\text{subject}_X, X, t) \implies \text{erase}(x, [t, t + \delta]) \]

grounded and mapped to system actions
How To Come Up With Invariants?
Classification of Data Regulations
Overview

Data-CASE

Data Regulation

Data-CASE concepts

Interpretations

Grounding

Implement Control- & data-paths

System-specific actions

Formal invariants
Uses
Of Data-CASE

- Data Collectors
- Database Providers
- Service Providers
- App developers
- Multinational Orgs
- Privacy Impact Assessments

See the paper for case studies.
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