Free Flow of Data and the Geography of Policies: Detecting Data Localisation Restrictions with Forward-Compatible Ontologies

Meem Arafat Manab, Víctor Rodríguez-Doncel, Elena Montiel-Ponsoda

Universidad Politécnica de Madrid

Fifth AI4LEGS Workshop collocated with 38th JURIX,

08.12.2025

Harnessing AI and Data-Intensive Technologies
HARNESS https://harness-network.eu/

The Problem

- EU Regulation 2018/1807 (Free Flow of Data Regulation): Prohibits unjustified data localisation within EU
  - Data should flow freely like goods, services, and people
- Policies increasingly use machine-readable languages (e.g. ODRL)
  - Manual compliance checking impractical at scale
  - Geographic references are heterogeneous (URIs, ISO codes, string literals)
- **November 2025: Digital Omnibus package announced to consolidate/replace this regulation – compliance tools must survive regulatory evolution**

Our Three Contributions

1. FFDRO Ontology
   o Free Flow of Data Restrictions Ontology
   o Captures key concepts of Regulation 2018/1807

2. HARNESS-FF Tool
   o Automated compliance verification
   o Systematically identifies geographic restrictions in ODRL policies

3. Forward-Compatible Design
   o Decouples stable legal concepts from regulatory provisions
   o Survives regulatory changes (e.g., Digital Omnibus package)

Theoretical Foundation - Legal Isomorphism

- Bench-Capon & Coenen (1992): Computational structures should faithfully reflect legal texts

- Our approach:
  - Semantic understanding (not just pattern matching)
  - Models the phenomenon of data localisation (conceptually stable)
  - Not just encoding specific regulatory articles
- Enables traceable compliance assessment
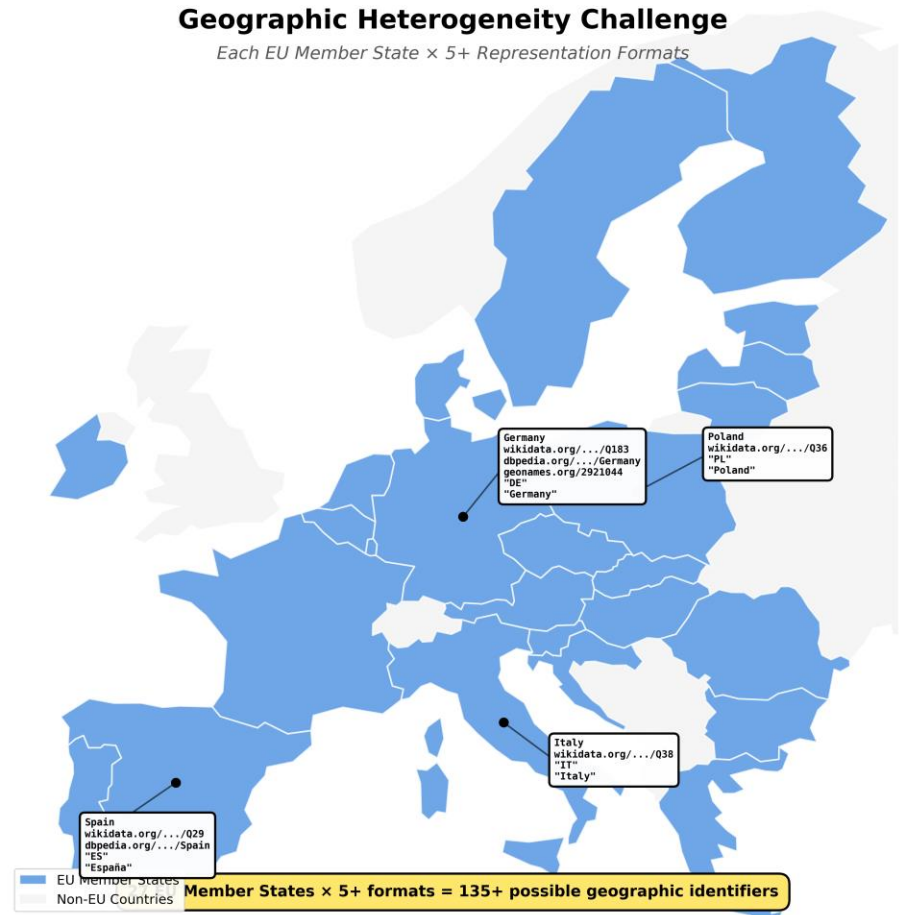
Learning from History - DPD to GDPR

- **Data Protection Directive (1995) → GDPR (2018)**
- Problem: Early DPD ontology attempts became obsolete
- Required complete rebuilding, not adaptation
- Root cause: Tight coupling between legal concepts and regulatory implementation

- Our lesson: Separate conceptual stability from regulatory contingency
- GDPR introduced fundamental structural changes: new rights (portability, erasure), expanded definitions, direct processor regulation
- **Successful exceptions (ELI, DPV)** survived through institutional mandate and *modular stratification* – not just good design
- **FFDRO applies this pattern: stable legal phenomena + implementation agnostic classifications + regulation-specific provisions (smallest, most replaceable component)**

Technical Challenge - Geographic Heterogeneity
How is "Germany" represented in ODRL policies?

- **Wikidata URI:** https://www.wikidata.org/resource/Q183

- **DBpedia**: https://dbpedia.org/resource/Germany

- **GeoNames**: https://www.geonames.org/2921044/

- **ISO 3166**: "DE"

- **String literal**: "Germany"

- × 27 EU Member States!

Geographic Heterogeneity Challenge
Each EU Member State × 5+ Representation Formats

Germany
wikidata.org/.../Q183
dbpedia.org/.../Germany
geonames.org/2921044
"DE"
"Germany"

Poland
wikidata.org/.../Q36
"PL"
"Poland"

Italy
wikidata.org/.../Q38
"IT"
"Italy"

Spain
wikidata.org/.../Q29
dbpedia.org/.../Spain
"ES"
"España"

EU Member States
Non-EU Countries

Member States × 5+ formats = 135+ possible geographic identifiers

Why Not Use LLMs?

LLMs are unsuitable for regulatory compliance:

- **Non-deterministic**: Same input → different outputs
- **Hallucinate**: Generate plausible but incorrect information
- **Cannot be traced**: No audit trail to authoritative sourcesxx
- **Scale problems**: Latency and cost prohibitive

SPARQL-based approach provides:

- **Deterministic, auditable results**
- **Formal verification**
- **Millisecond execution** at negligible cost

HARNESS-FF Architecture
Three-stage analysis pipeline

Parse

- ODRL policy is parsed into RDF graph
- SPARQL queries for geographic constraints

Normalize

- Resolve references to canonical EU member states
- Record detection method & confidence score

Evaluate

- Generate structured compliance report (FFDRO vocabulary)
- Compliance status + actionable recommendations

FFDRO Ontology - Three-Module Design

Core Module (Stable concepts)

- Localisation Restriction, Compliance Evaluation, Policy Violation
- EU Member State, Regulatory Article

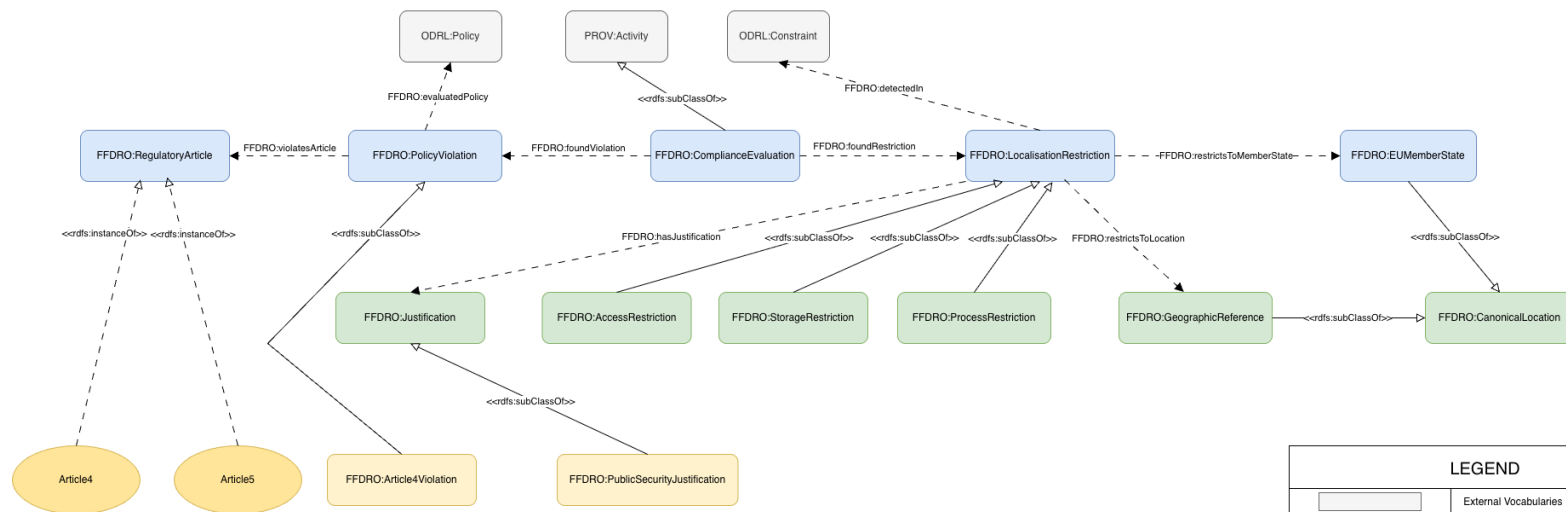**Extended Module (Technical classifications)**

- Geographic Reference, Canonical Location
- Storage Restriction, Processing Restriction, Access Restriction

Integration Module (Regulation-specific)

- Article4 Violation, Article5
- **This module changes when regulations evolve**

FFDRO Architecture Diagram

FFDRO Vocabulary Architecture: Three-Module Design for Data Localisation Compliance

Forward Compatibility in Action

Validation of our approach:
- November 2025: Digital Omnibus package announced
- Proposes consolidating Regulation 2018/1807 with other instruments
- **Our design responds:**
  - Core & Extended modules remain unchanged
  - Only Integration Module needs updating
  - No wholesale redesign required

Example Detection Policy

```
Restriction Instance:
- Type: AccessRestriction
- Location: Germany (detected from "DE")
- Detection method: iso-code
- Confidence: 1.0

Compliance Evaluation:
- Status: non-compliant
- Violation: Article 4
- Justification required: Yes (public security
grounds)
```

Integration with Data Spaces

Where HARNESS-FF fits:

- GAIA-X: Policy Decision Point (PDP) → Policy Enforcement Point (PEP)
- IDSA: Dataspace Connector's PEP
- Ensures data usage adheres to legal requirements before access granted
- ODRL selected as core technology for European data spaces

Impact & Future

Current state:
- ODRL adoption still limited today

Future importance:
- European data spaces initiative (health, agriculture, mobility, energy)
- National governments, universities, research institutions need compliance
- HARNESS-FF transforms compliance from ad-hoc manual process to systematic, scalable practice

Design pattern for legal ontologies:
- Three-tier modular stratification
- Conceptual (stable) / Technical (classification) / Regulatory(specific)

Conclusion

Key takeaways:

- Data flows are as important as movement of goods and people
- Digital borders can be invisible but enforceable
- HARNESS-FF makes compliance automated and systematic
- Forward-compatible design survives regulatory evolution
- Both tool and ontology publicly available online

"Freedom of movement requires not just absence of walls, but presence of open roads ..."
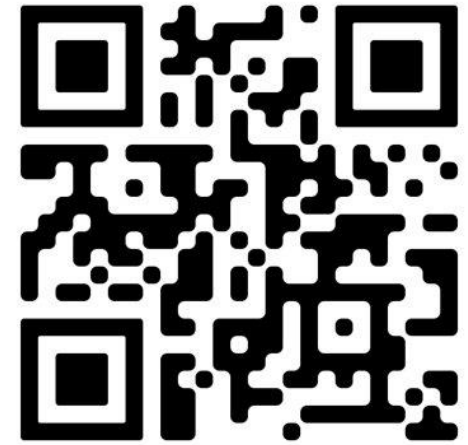
Thank you!

Please share research ideas with
Meem Arafat Manab | meem.manab@upm.es
HARNESS Doctoral Network (MSCA Grant 101169409)