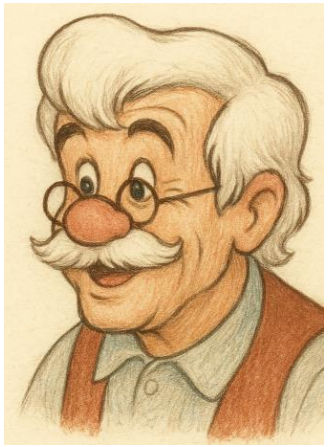
Formalizing Institutional Data Custody: Multi-Tier Data Governance in European Data Spaces

Meem Arafat Manab

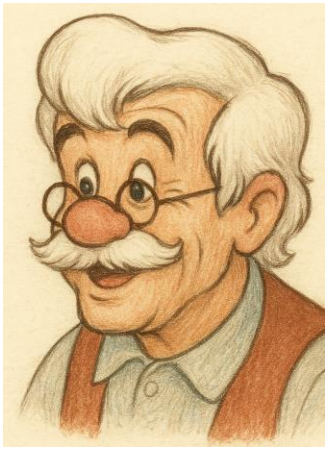
Ontology Engineering Group
Universidad Politécnica de Madrid, Spain

Meet Giuseppe



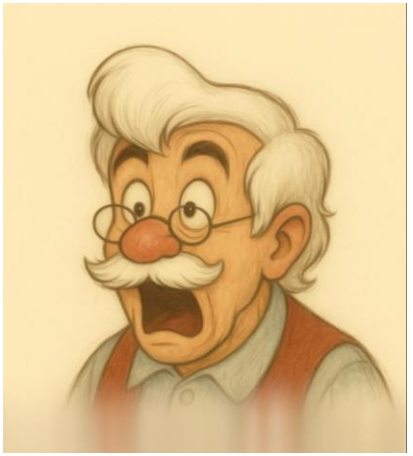
This is Giuseppe

Giuseppe's Technical Skills



Giuseppe lives in Zaragoza. He's "very good with computers" because he knows how to restart his router.

One week, three institutions come knocking...



The **Municipality**
wants vehicle emissions
data for climate policy

The **National Water Authority**
needs consumption patterns
for drought planning

The **Regional Waste Consortium**
seeks disposal habits to
optimize collection routes

Giuseppe confidently opens his "data pod settings"...

...then immediately closes his laptop!

The Core Problem

Giuseppe trusts these institutions serve legitimate purposes:

- Cleaner air
- Secure water
- Efficient waste collection
- Better life for all Zaragozans

But managing:

- separate pods *(personal online data stores),*
- granular permissions,
- AND access audits?

That requires technical knowledge!!

HARNESS

And It Gets More Complicated

DS4SSCC: Data Spaces for Smart and Sustainable Cities

When Zaragoza collaborates with neighboring cities on waste optimization:

- Who ensures Giuseppe's water data stays within drought planning?
- What happens when Aragon shares algorithms/data with French Occitania?
- How do we handle cross-border data governance?

Municipal institutional custody lacks formal standards

Why Current Solutions Fall Short

Solid Protocol	IDSA/Gaia-X	Legal Ontologies
User-controlled datastores	Technical infrastructure ready	GDPR formalization (PrOnto, GConsent, DPV)
Emergency access scenarios	High-level guidance only	Focus on controller/processor roles
Most citizens lack technical capability	No machine-readable legal specifications	Don't model custody relationships

Our Research Questions

RQ1: Extending Legal Standards

How can Akoma Ntoso and LegalRuleML formalize multi-stakeholder data governance agreements?

- New document types for data governance agreements (parties, data flows, purposes, custody chains)
- Validate by encoding DS4SSCC municipal agreements

RQ2: Cross-Border Interoperability

What framework enables semantic interoperability across different jurisdictions?

- Core Custody Vocabulary (following EU SEMIC methodology)
- Semantic mapping: Spanish municipalities ↔ Slovenian municipalities
- 6-month research stay at University of Bologna (CIRSFID-AlmaAI)

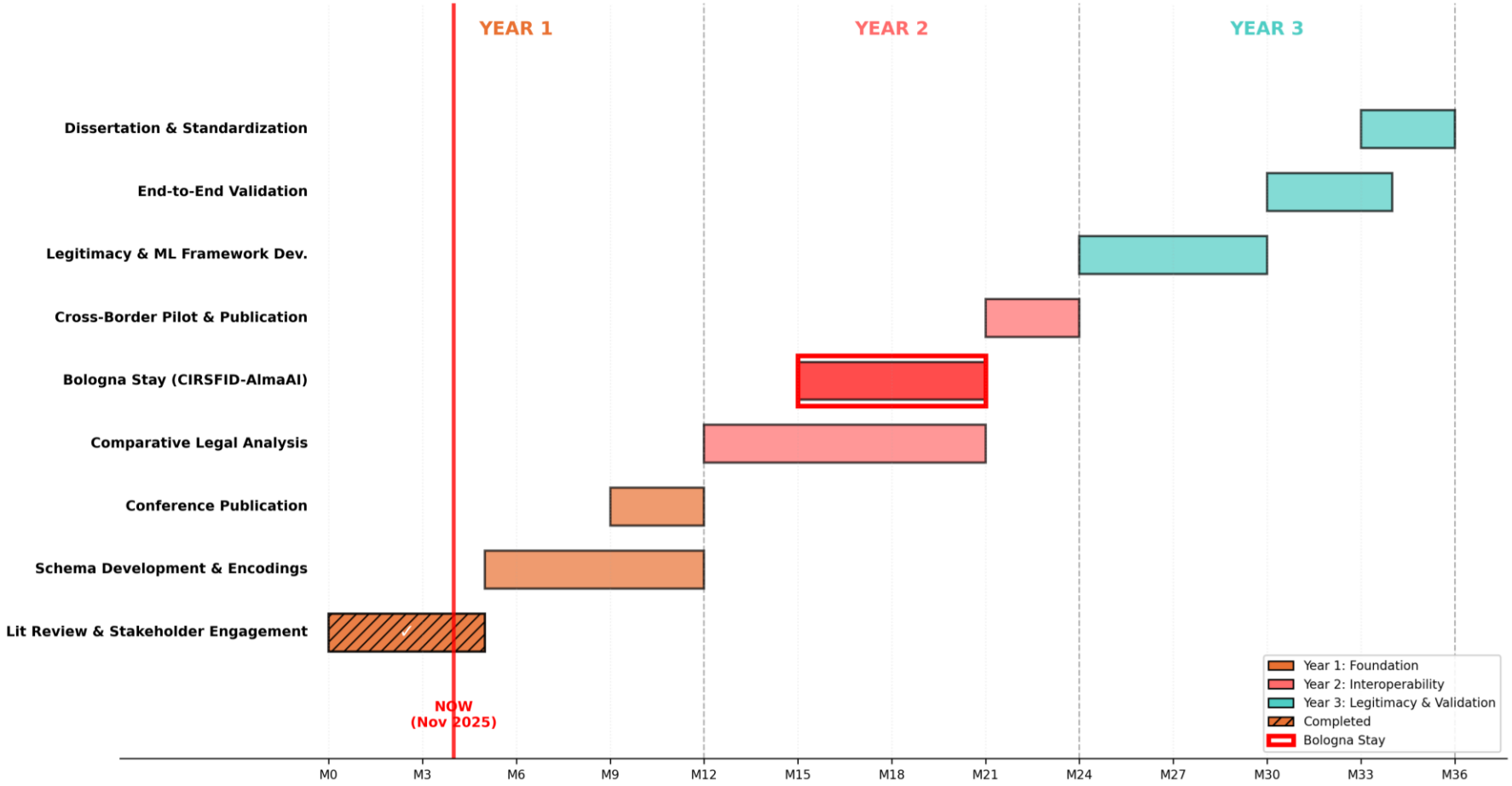
RQ3: Legitimacy & Monitoring

When is institutional custody legitimate, and can ML detect misuse?

- Operationalize fiduciary commons theory into machine-readable criteria
- ML-based anomaly detection with explainable AI (SHAP values)

HARNESS

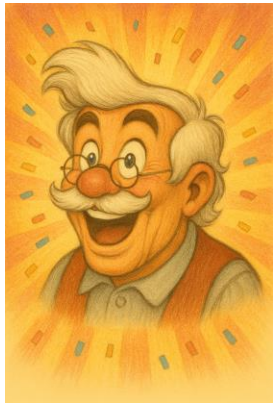
Research Timeline



What this will mean

For Giuseppe:

- Municipality manages data with formalized fiduciary duties
- No nephew required
- Trust in institutional custody
- Zero learning curve



For European Data Spaces:

- **DS4SSCC**: 10-12 pilots benefit from formal legal foundations
- **Standardization**: Contributing to established legal XML standards
- **Trust Infrastructure**: Machine-readable legal specifications for cross-border collaboration

For us:

- Extending Akoma Ntoso/LegalRuleML to data governance
- Semantic interoperability across jurisdictions
- Legitimacy framework with ML monitoring

HARNESS

THANK YOU!

Meem Arafat Manab | meem.manab@upm.es
HARNESS Doctoral Network (MSCA Grant 101169409)