

A semantic approach to digital-ready policymaking: the Legislative Financial and Digital Statement

Cecile Guasch

European Commission, DG DIGIT.B2,
Brussels, Belgium

cecile.guasch@ext.ec.europa.eu

Alessio Nardin

European Commission, DG DIGIT.B2,
Brussels, Belgium

alessio.nardin@ext.ec.europa.eu

Davide Audrito

European Commission, DG DIGIT.B2, Brussels, Belgium

davide.audrito@ext.ec.europa.eu

Abstract

As of January 2025, the main instrument of digital-ready policymaking in the European Commission is the digital chapter of the Legislative Financial and Digital Statement, which compiles essential information on the digital aspects of policy initiatives. To fully leverage the richness of such information, which is initially provided in plain text, and to enable further use cases, adopting a semantic approach is key. Accordingly, the paper focuses on presenting how the Legislative Financial and Digital Statement Vocabulary was developed, outlining the main design rationales, the use cases supported by the data model, and its relationships with other existing semantic specifications in the legal domain. Additionally, the article presents intermediate results obtained during the construction of this reference data, as well as the manual and semi-automated approaches used to reduce the burden on the policy officers responsible for completing the Legislative Financial and Digital Statement.

Keywords: legal semantics, artificial intelligence, digital-ready policymaking, Legislative Financial and Digital Statement

Introduction

As of January 2025, the European Commission has adopted the Digital-Ready Policymaking (DRPM) framework to comply with Article 3 of EU Regulation No. 2024/903, the so-called the Interoperable Europe Act. The Act, in force since 11. April 2024 establishes a comprehensive framework for cross-border interoperability of network and information systems used to provide or manage public services in the European Union (EU). Most prominently, Article 3 requires EU institutions and public sector bodies to conduct interoperability assessments when creating new or significantly modifying existing digital systems that enable cross-border public services, meaning cross-border digital public services in the EU. This marks a shift from voluntary cooperation to a mandatory, systematic evaluation of the interoperability implications of policy proposals.

Within this regulatory context, the European Commission has introduced the Legislative Financial and Digital Statement (LFDS)¹, which now includes a dedicated digital dimensions

¹ As its name suggests, the Legislative Financial and Digital Statement outlines the financial and digital aspects of a legislative proposal to support decision-makers and those involved in policy implementation to thrive for simplification, efficiency and burden reduction. In the context of digital-ready policymaking, any mention of the acronym LFDS refers specifically to its digital dimensions, as outlined in the fourth chapter.

chapter accompanying legislative proposals. The chapter consolidates key information on the digital aspects of policy initiatives and reinforces the European Commission obligation to conduct interoperability assessments under the Interoperable Europe Act. Besides the LFDS, which applies to legislative acts, the European Commission has also created the Digital Statement (DS), which is tailored to autonomous acts. This article focuses on the vocabulary derived from the information contained in the LFDS digital chapter, which covers both the LFDS and the DS. However, the model will include certain properties specifically designed to capture a few minor differences in the information required by each instrument. Together, the LFDS and DS provide the formal backbone for capturing the information required by the new interoperability assessment obligation in the European Commission.

The Digital-ready Policymaking Framework (DRPM)

The DRPM framework is the European Commission's governance and process architecture designed to ensure that digital and interoperability considerations are identified and addressed early and consistently throughout the policy lifecycle. It builds a path from initial screening to structured assessment, so that the digital implications of proposed policy initiatives in scope are assessed consistently. For the purposes of this article, the most relevant instruments within the framework are the digital dimensions chapter of the **Legislative Financial and Digital Statement (LFDS)** accompanying legislative proposals and the **Digital Statement (DS)** annexed to autonomous acts. The DS is a tailored version of the LFDS, designed to produce interoperability assessments specifically for autonomous acts, which are primarily implementing and delegated acts. The DS shall accompany these acts whenever an interoperability assessment is required under Article 3 of the Interoperable Europe Act. This requirement applies when the proposed act includes binding provisions that impact the design, development, or implementation of digital public services, as well as the networks and information systems supporting their provision, particularly where cross-border interoperability is affected.

At the entry point of the framework sits the lightweight **Digital check (DC)** conducted during policy planning with the support of multidisciplinary teams. It consists of a set of four simple questions that are aimed at conducting an early assessment of the potential digital aspects of a policy initiative. Its objective is to help policy officers assess the presence of potential digital dimensions (data exchanges, use of digital solutions, automation of stakeholder processes or impacts to digital public services) to be elaborated further during policy design.

The LFDS or DS is completed once a stable draft of the legal text exists, so that its contents can be considered in inter-service consultations (ISCs). Within the DRPM framework, the European Commission also provides **Digital-ready guidance** materials that support digital-ready policy design across the entire lifecycle. The guidance is further categorized in *DRPM essentials*, which mostly contain step-by-step instructions and for being digital-ready and templates for implementing DRPM procedures, *Training materials*, including self-paced modules to build DRPM knowledge and skills, and *Practical insights & examples*, including examples on how the DRPM is applied in specific policy initiatives. The guidance complements the DC and the LFDS or DS, ensuring methodological consistency and comparability of interoperability assessments.

Use Cases

The Legislative Financial and Digital Statement Vocabulary (LFDSV) (European Commission, 2025) is designed to operationalize the digital-ready policymaking framework by turning both the LFDS and DS into consistent, “queryable”, and machine-readable data. Concretely, the vocabulary aims to support the following use cases:

1) UC1: Complete the digital dimensions chapter of LFDS in a semi-automated way

Establishing precise semantic definitions for the concepts needed in the digital dimensions chapter of the LFDS, along with defining the interconnections, streamlines manual data entry. This foundation also supports the semi-automation of its completion by identifying relevant legal text snippets that align with these concept definitions and uncovering relationships with other concepts. Additionally, generating LFDS instances as “queryable” knowledge graphs – validated with SHACL and enriched with semantic property chains – enables assisted extraction and linking of relevant snippets, and automated checks for completeness and consistency.

2) UC2: Draft requirements of digital relevance in legal texts using a streamlined and digital-ready approach

When drafting legal texts, policy officers can leverage LFDS data to explore the digital dimensions of the EU legal acquis. This offers a comprehensive perspective on requirements of digital relevance and the digital assets involved, providing a user-centric view of their impact on stakeholders. The resulting LFDS knowledge graphs can then be queried to discover existing requirements, datasets, digital solutions, and components of digital public services, helping drafters assess stakeholder impacts, identify reusable patterns and dependencies, and align new provisions with established standards to solve potential interoperability gaps.

3) UC3: Convert the LFDS digital dimensions’ chapter into an interoperability assessment report

By leveraging the LFDS knowledge graphs, the digital dimensions’ chapter can be transformed into a comprehensive interoperability assessment report.

The LFDS

Completing the LFDS is mandatory for legislative proposals – EU regulations, directives, and decisions – having digital impact, and comprises the following five sections on digital implications, interoperability, and implementation measures:

1. **Requirements of digital relevance** captures provisions in the legal text that apply to the collection, processing, generation, exchange or sharing of data, the automation or digitalization of processes, the use of digital solutions, or digital public services. Once completed, each legal requirement is referenced and briefly described, the actors and processes affected are identified, and the requirements are categorized to ensure subsequent sections can trace back to the specific legal basis.
2. **Data** focuses on the requirements linked to the data dimension. A high-level description of the data covered by each requirement is described, and any applicable standards or technical specifications are specified. The LFDS requires to explain how the legal provisions under examination align with the European Data Strategy, confirm adherence to the once-only principle, and, for data exchanges, identify the actors involved and the triggers for exchange, including temporal aspects where relevant.
3. **Digital solutions** cover requirements for digital solutions – services, platforms, systems, applications, tools or infrastructure designed to automate, manage or improve processes through digital technologies. The LFDS allows for describing the

solution's functionalities as set out in the legal text, indicates the body responsible for the solution, and assesses compliance with the principles of reusability² and accessibility for persons with disabilities, besides reporting on the potential use of AI technologies. The section also reflects the alignment of the digital solution with EU digital policies, including the AI Act and the EU cybersecurity framework, to ensure coherence and implementability.

4. **Interoperability assessment** addresses requirements for digital public services with cross-border relevance across the EU. The section demands completing an assessment of the impact of each requirement on cross-border interoperability, listing any Interoperable Europe solutions or other interoperability solutions that can be reused to implement the service efficiently, and identifying any remaining interoperability barriers that may require further considerations.
5. **Measures to support digital implementation** covers the policy initiative's planned digital implementation measures, such as implementing or delegated acts, policy implementation pilots, procurement procedures, and sandboxes. Each measure is described by outlining the European Commission's role, the actors involved, and the expected timeline, so that implementation is coordinated and traceable alongside the policy proposal.

Central to the digital dimensions chapter of the LFDS is the concept of requirements for digital relevance, which are crucial for assessing the digital implications of any policy initiative. This concept closely links the various sections of the digital dimensions chapter, ensuring a consistent approach throughout.

The LFDSV

The development of legal ontologies as semantic artefacts is well-established in AI and law research. It ranges from foundational ontologies such as DOLCE (Gangemi, Guarino, Masolo, Oltramari, & Schneider, 2002) and domain-specific frameworks aimed at modelling the General Data Protection Regulation (GDPR) like the Pronto ontology (Palmirani, Martoni, Rossi, Bartolini, & Robaldo, 2018), to more recent efforts that model reporting requirements (Palmirani M. , Nuzzolese, Poggi, Nicholson, & Estevez, 2024), classify legal definitions (Audrito, Sulis, Humphreys, & Di Caro, 2023), and employ hybrid methodologies that integrate legal concept mapping with Large Language Models (LLM) to analyze legal harmonization within the EU (Audrito, Spada, Mignone, Sulis, & Di Caro, 2025).

The development of the LFDSV adheres to established best practices in ontology engineering, with a focus on reusing, wherever possible, classes and properties from existing ontologies. These include the Reporting Requirement Metadata Vocabulary (RRMV) (Palmirani, Nuzzolese, & Generoso, 2025; European Commission, 2025), which serves as a foundational framework for representing reporting requirements; the European Legislation Identifier (ELI) ontology (European Commission, 2022), which provides a standardized method for identifying legal documents and representing their metadata; the Asset Description Metadata Schema (ADMS) (W3C, 2013), designed for describing and sharing interoperability assets; the Core Standards and Specifications Vocabulary (CSSV) (European Commission, 2025), which facilitates the interoperable description, cataloguing, and sharing of standards and specifications; and the Data Catalog Vocabulary (DCAT) (W3C, 2024), an RDF vocabulary developed by W3C to enable publishers to describe datasets and data

² According to the principle of reusability, data and metadata should be easy to find for humans and machines. Persistent identifiers, accurate metadata, and index in searchable resources should be prioritized. For more information, see <https://www.nature.com/articles/sdata201618>.

services in catalogues using a standardized model and vocabulary, thereby supporting metadata aggregation across multiple catalogues.

This research builds on these robust foundations, addressing a key gap in the automated generation of interoperability assessments. It combines the explicit semantics provided by formal ontologies with the analytical capabilities of LLMs, aiming to meet the requirements of digital-ready policymaking within the European context.

The primary goal of the LFDSV is to represent the legal and semantic content of the LFDS digital dimensions chapter in a standardized, machine-readable format, according to OWL. The chapter was conceived from the outset with semantic web modeling in mind. As previously outlined, policy officers conduct their assessments primarily in documents formatted as tables, where data is organized into rows and columns. This structure, in turn, facilitated the development of the LFDSV, whose graphical representation is illustrated in Figure 1 below.

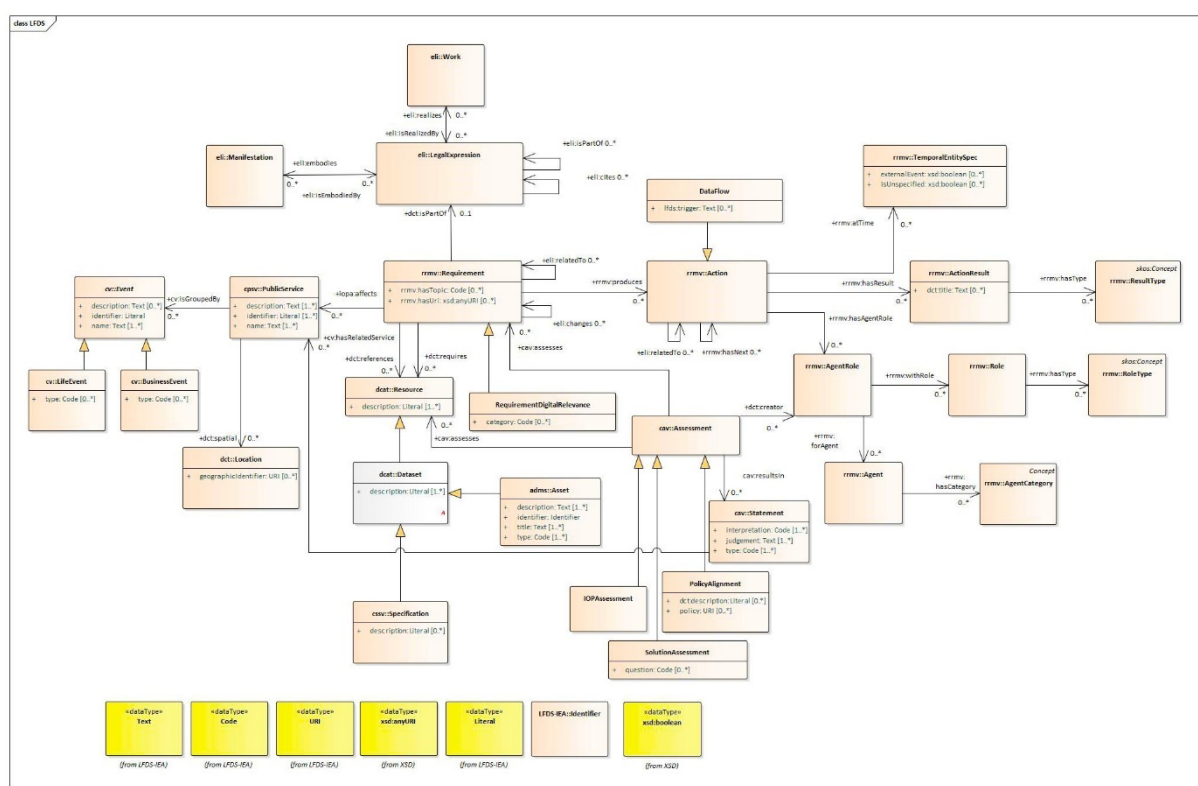


Figure 1 – Legislative Financial and Digital Statement Vocabulary (LFDSV, release 0.2.0)

The LFDSV reuses established EU vocabularies, anchored by the `eli:LegalExpression` class from the European Legislation Identifier (ELI) ontology. This serves as the class connecting all versions of requirements from a specific legislative act under assessment, with relationships to `eli:Work` abstractions and concrete `eli:Manifestation` instances through standard ELI properties like `eli:realizes`, `eli:embodies`, and `eli:cites`.

The LFDS section 4.1 devoted to collecting requirements of digital relevance is formalized by building upon the RRMV with the core `rrmv:Requirement` class. The model introduces `lfds:RequirementDigitalRelevance` as a specialized subclass expressing requirements of digital relevance that are categorized via the `lfds:RequirementCategory` property, preserving the LFDS categorization logic for data, process digitalization and automation, digital solutions, and interoperability assessment while enabling formal SKOS-based reasoning.

The LFDSV transforms procedural descriptions into an action-centered model through the `rrmv:Action` class with temporal specifications through the reused `rrmv:TemporalEntitySpec`, an abstract superclass linked from `rrmv:Action` through the `rrmv:atTime` property. Legislative requirements are indeed often connotated by one or more actions such as reporting data, adopting organizational measures or implementing components and functionalities of online platforms.

The `rrmv:TemporalEntitySpec` provides the broad range of temporal relations starts, ends, before, after, in, meets, and overlaps, semantic typing via `hasFunction` as an attribute of `rrmv:TemporalFunction`, recurrence via `hasFrequency` as an attribute of `rrmv:Frequency`, and flags `externalEvent` and `isUnspecified`. The `externalEvent` property indicates whether a temporal entity specification depends on events outside the direct control of the regulatory framework, while `isUnspecified` signals whether a temporal aspect exists but is undefined. Specializations are `rrmv:Instant` expressing a point-in-time, `rrmv:Interval` indicating a period bounded by start and end instants via `hasStart` and `hasEnd`, and `rrmv:Duration` specifying a relative length via the `hasDuration` property. In practice, this pattern allows users to type a temporal specification as a deadline or validity period, anchor it with an instant or compose an interval, add recurrence (e.g., yearly) with frequency, and relate with starts, ends, before, or after. When timing is driven by external triggers or not yet defined, the boolean flags can be leveraged to mark that explicitly.

The `rrmv:AgentRole` pattern connects `rrmv:Agent` entities with `rrmv:Role` classifications through `rrmv:forAgent` and `rrmv:withRole` properties, replacing the LFDS actor descriptions with structured semantic encoding. Action outcomes are captured through `rrmv:ActionResult` instances with `rrmv:ResultType` classifications, supporting automated tracking of implementation progress. This action-result modeling directly supports the LFDS emphasis on responsible bodies and expected timelines through formal property chains.

Section 4.2 of the LFDS, which covers data analysis, is reorganized around the `lfds:DataFlow` class that captures dynamic exchange patterns, while the `lfds:trigger` property supports multilingual descriptions. The model leverages DCAT vocabulary through `dcat:Dataset` and `dcat:Resource` classes for static assets, while maintaining requirement traceability through properties like `dcterms:references` and `dcterms:requires`. Digital assets, including digital solutions outlined and described in section 4.3 of the LFDS, are unified under `adms:Asset` with mandatory `adms:Identifier` instances and structured metadata using Dublin Core properties. The once-only principle and the European Data Strategy alignment are incorporated through the `lfds:PolicyAlignment` class. This class is designed to be prefilled, reflecting the policies and principles that guide the assessment process. Meanwhile, the `cav:Statement` class contains the actual text of the assessment.

Section 4.4 of the LFDS, which addresses interoperability assessment, leverages the Core Assessment Vocabulary (CAV) (European Commission, 2025) with `cav:Assessment` as the foundational class. Specialized sub-classes `lfds:IOPAssessment` and `lfds:SolutionAssessment` (also applicable to digital solutions) provide domain-specific evaluation mechanisms, with each assessment producing `cav:Statement` instances containing mandatory `cav:judgement` properties and `cav:interpretation` values from SKOS concept schemes. Digital public services are modeled through Core Public Service Vocabulary (CPSV) (European Commission, 2024), possibly with spatial characteristics (`dcterms:Location`) and event relationships that include `m8g:Event`, `m8g:BusinessEvent`, and `m8g:LifeEvent` classes. This enables cross-border interoperability modeling.

Policy compliance from the LFDS various sections is formalized through the `lfds:PolicyAlignment` class linking specific policies via `lfds:policy` properties with descriptive

assessments. This approach enables tracking alignment with digital policies such as the AI Act, eIDAS, and cybersecurity frameworks across multiple legislative initiatives through formal semantic relationships. Technical standards integration utilizes the Core Standard and Specification Vocabulary (CSSV) for capturing compliance requirements.

The ontological transformation fundamentally changes the LFDS from document-centric to machine-processable knowledge resource. SHACL validation shapes ensure data quality through cardinality constraints (sh:minCount, sh:maxCount), datatype validation (sh:datatype), and class restrictions (sh:class), while semantic property chains enable automated reasoning about requirement dependencies, action sequences, and compliance gaps. This semantic foundation supports systematic cross-legislative analysis, automated interoperability gap detection, and evidence-based policy impact assessment, transforming LFDSs into queryable and interoperable knowledge graphs suitable for EU-wide digital governance requirements.

Building on this semantic foundation, the next LFDSV release will introduce an implementation layer to model the concrete measures and artefacts undertaken to operationalize legislative requirements. This layer will capture, among others, information transmission (e.g., as datasets), the development of digital solutions, and the creation or integration of components of digital public services. It will also enable explicit representation of external events that affect the status of a legislative requirement, allowing traceability of status changes, dependencies, and real-world implementation progress over time.

AI4DRPM: programmatically instantiate Digital Statements

AI4DRPM is a software microservice implemented in Python that automates the ingestion, transformation, annotation, classification and compilation of legislative documents to support evidence-informed policymaking. It ingests legal resources through federated retrieval (e.g., SPARQL-based queries to Cellar) or by accepting uploaded AKN4EU artifacts, converts them into intermediate structured representations, i.e. LegalJSON3 and AKN4EU), and materializes a provision-level schema in a relational database in which each provision is identified by the tuple (CELEX, eId)⁴. Looking ahead, the system is being designed to extend its capabilities by accommodating draft legal texts as additional data sources via the integration with EdIT, the upcoming legislative drafting solution of the EU. This planned feature will enable the ingestion and processing of preliminary legislative content, further enhancing its use in policymaking workflows. Pipeline orchestration and model invocation are configuration-driven, enabling the parametrisation of prompt templates, model endpoints and execution parameters via a JSON configuration file.

Core architecture

The software architecture is layered and modular. A REST API implemented with FastAPI exposes endpoints for resource ingestion, classification, analysis and compilation, while a persistence tier uses SQLAlchemy ORM models for core entities (legal resources, provisions, per-dimension analyses, classification artifacts and task records). Computational logic is factored into small service modules that orchestrate repositories and pipeline steps; legal provision classification and analysis are implemented as pluggable components.

Operational workflows follow a canonical sequence: resource collection and deterministic transformation, asynchronous execution of classification and analytical steps (enqueued via a

³ LegalJSON is a standard JSON with a schema based on AKN4EU attributes.

⁴ In the next release, ELI identifiers will replace eIDs.

task factory and executed by Celery workers), validation and persistence of structured outputs, and finally compilation into human- or machine-readable artifacts (e.g., DOCX or RDF graphs).

AI architecture

AI4DRPM employs a hybrid AI architecture that combines classical ML pipelines and prompt-driven large language models (LLMs). The design separates model invocation from orchestration: pluggable step classes encapsulate interaction logic for a given technique (for example, an LLM-based classification step or an ML pipeline for rule-based tagging), and service modules invoke those steps according to configuration. LLM calls are made to an OpenAI-compatible endpoint and are parameterised by prompt templates, temperature, and maximum token constraints provided in the configuration file.

LLM-driven steps follow a three-stage pattern: (1) prompt construction, in which contextual information (provision text, metadata, and schema constraints) is composed into the prompt; (2) model invocation; and (3) structured output extraction and validation. Prompt templates are kept external to code to enable versioned experimentation and reproducibility. The implementation anticipates non-deterministic or malformed LLM outputs by validating responses against expected schema before persistence; invalid items are either corrected with deterministic post-processing, flagged for human review, or skipped with an auditable log entry.

Although generative AI models are inherently non-explainable and prone to biases and hallucinations, the AI4DRPM architecture requires the model to generate an explanation for each automatically produced output, enabling policy officers responsible for the interoperability assessment to verify the information. While this step does not guarantee explainability or eliminate susceptibility to bias, it nonetheless moves in the direction of mitigating these issues by promoting human verification.

Classical ML pipelines are implemented for deterministic classification techniques where available. For reproducibility, model artifacts, hyperparameters and preprocessing steps are to be recorded in configuration or an external model registry; the codebase is structured so that ML steps conform to the same handler interface as LLM steps, making it straightforward to swap or augment techniques.

The AI4DRPM approach demonstrates a tangible reduction in the manual effort required to complete the LFDS, while generating structured outputs, both textual and RDF. The resulting legal knowledge graphs represent a critical step towards realizing a digital twin of legislation, enabling new forms of analysis, interoperability, and automation in support of digital-ready policymaking across Europe.

Conclusions and future work

The LFDSV represents an advance in transforming the LFDS digital chapter from a static document into a “dynamic”, queryable knowledge graph. By building upon well-established EU vocabularies, the model captures requirements of digital relevance together with affected actors, timelines, data flows, digital solutions, and interoperability assessments in a machine-readable format.

This semantic foundation enables the semi-automated completion of interoperability assessments, streamlines the legislative drafting of digital requirements, and enhances traceability from legal bases to concrete implementation measures. As a result, the LFDSV

supports not only compliance with the Interoperable Europe Act but also the broader ambition of digital-ready policymaking in the EU.

Knowledge graphs have the potential to improve the accessibility of legal knowledge, facilitate cross-legislative analysis, and enable new forms of evidence-based policymaking. AI-driven approaches will continue to play a key role in semi-automating policymaking workflows, enhancing data quality, and enabling advanced analytics. Together, these advances pave the way for a new generation of digital-ready, interoperable, and adaptive policymaking across Europe.

References

- Audrito, D., Spada, I., Mignone, R., Sulis, E., & Di Caro, L. (2025). Towards semi-automating European legislative harmonisation analysis: A harmonised glossary for LLM-based legal concept detection. *Computer Law & Security Review* 58 106171.
- Audrito, D., Sulis, E., Humphreys, L., & Di Caro, L. (2023). Analogical lightweight ontology of EU criminal procedural rights in judicial cooperation. *Artificial Intelligence and Law*, 31(3), 629-652.
- European Commission. (2022). *European Legislation Identifier (ELI) ontology v1.4*. Retrieved from <https://op.europa.eu/documents/3938058/11669184/ELI-note.pdf>
- European Commission. (2024). *Core Public Organization Vocabulary (CPOV)*. Retrieved from <https://semiceu.github.io/CPOV/releases/2.1.1/>
- European Commission. (2025). *Core Assessment Vocabulary (CAV) v1.3.0*. Retrieved from <https://github.com/isa-camss/CAV/releases>
- European Commission. (2025). *Core Standards and Specifications Vocabulary (CSSV) v1.4.1*. Retrieved from <https://interoperable-europe.ec.europa.eu/collection/common-assessment-method-standards-and-specifications-camss/solution/core-standards-and-specifications-vocabulary-cssv/distribution/cssv-v141-html>
- European Commission. (2025). *Regulatory Reporting Metadata Vocabulary (RRMV) v0.2.0*. Retrieved from <https://semiceu.github.io/RRMV/releases/0.2.0/>
- European Commission. (2025). *The Legislative Financial and Digital Statement (LFDSV) v0.2.0*. Retrieved from <https://semiceu.github.io/LFDS/releases/0.2.0/>
- Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., & Schneider, L. (2002). Sweetening ontologies with DOLCE. *International conference on knowledge engineering and knowledge management* (S. 166-181). Berlin, Heidelberg: Springer Berlin Heid.
- Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., & Robaldo, L. (2018). Pronto: Privacy ontology for legal reasoning. *International Conference on Electronic Government and the Information Systems Perspective* (pp. 139-152). Cham: Springer International Publishing.
- Palmirani, M., Nuzzolese, A. G., Poggi, F., Nicholson, J., & Estevez, J. (2024). Tracking WHO reporting requests using hybrid AI. *Proceedings of the 17th International Conference on Theory and Practice of Electronic Governance*, (S. 139-146).
- Palmirani, M., Nuzzolese, A., & Generoso, L. (2025). Reporting Requirements Ontology for European Legislation. *Electronic Government and the Information Systems Perspective: 14th International Conference, EGOVIS 2025* (S. 18). Bangkok, Thailand: Springer Nature.
- W3C. (2013). *Asset Description Metadata Schema (ADMS)*. Retrieved from <https://www.w3.org/TR/vocab-adms/>
- W3C. (2024). *Data Catalog Vocabulary (DCAT) v3*. Retrieved from <https://www.w3.org/TR/vocab-dcat-3/>