LegIOn-IDS: Legal Interoperability Ontology for International Data Spaces⋆

Victor Oliveira[1,*], Patrício Silva[1] and João Moreira[2]

[1]*Federal Rural University of Semiarid Region - UFERSA, Mossoró, Brazil.*
[2]*University of Twente, Enschede, Netherlands.*

Abstract

The International Data Spaces (IDS) initiative aims to leverage secure and data-sovereign data spaces, allowing each data owner to provide a data usage policy. The European Interoperability Framework guides the IDS initiative, hence, they ought to comply with the four fundamental layers of interoperability, i.e., Legal, Organizational, Semantic, and Technical Layer. Based on a performed Systematic Literature Review (SLR), from 6 different sources, and 40 papers with an in-depth analysis, we identified the lack of research regarding the legal interoperability layer, even though, it is the fundamental layer. In this paper, we propose the development of the Legal Interoperability Ontology for IDS (LegIOn-IDS), a domain ontology encompassing the core participants, and the basic contractual flow, defining the legal aspects that allow the legal interoperability within. We followed an ontology engineering methodology (SABiO) for developing LegIOn-IDS ontology, producing a reference ontology in OntoUML, based on the Unified Foundational Ontology (UFO), with further transformation to an operational OWL ontology. This ontology covers specific definitions related to IDS, i.e., the Open Digital Rights Language and Reference Architecture Model. Finally, we provide an ontological connection with the Service Contract Ontology and the Information Model ontology, developed by the International Data Spaces Association. We designed the reference ontology from a set of competency questions and validated the ontology by answering these questions through informal (natural language) and formal (SPARQL queries) ways.

Keywords

Domain Ontology, International Data Spaces, Legal Interoperability, Ontology, European Interoperability Framework

1. Introduction

Along with Industry 4.0 arising, data has become one of the most valuable assets for companies, leading to data-driven business, hence, inducing concerns regarding its sharing. While acknowledging the advantageous trade-off of committing to data exchange, such as growth and competitiveness, they cannot restrict their usage to intern or open-access data sources. From predictive maintenance when sharing sensor data, to following market tendencies [1], data exchange leads to several benefits, such as allowing collaborative innovation, and value co-creation. Henceforth, the need for a data space ecosystem has led to the creation of the so-called Industrial Data Spaces [2]. The Industrial Data Spaces had as their main foundation data sovereignty, which stands for the power held by the data owner to delimit its usage, by defining who, when, and for how long data users may use it [3]. One year later, in 2016, the IDS initiative took shape as an enhanced version of Industrial Data Spaces, with international collaboration capabilities [4]. Adjacent to data sovereignty, IDS focuses on trust within its participants, relying on a thorough certification process for each of its participants.

The International Data Spaces Association (IDSA) is responsible for managing and providing a foundation for the IDS architecture. Throughout the years, it has developed the so-called Reference Architecture Model (RAM) [5], which establishes the framework for implementing, using, and main-

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands.

[*]Corresponding author.

victor.oliveira53295@alunos.ufersa.edu.br (V. Oliveira); patricio.alencar@ufersa.edu.br (P. Silva); j.luizrebelomoreira@utwente.nl (J. Moreira)
0009-0000-7026-6929 (V. Oliveira); 0000-0001-6827-1024 (P. Silva); 0000-0002-4547-7000 (J. Moreira)

taining IDS, grounding the complete engineering process. As an important domain instituted by the RAM, the interoperability within IDS is grounded by the European Interoperability Framework (EIF) [1] which provides the concept of interoperability as a result of four layers of interoperability, i.e., legal interoperability, organizational interoperability, semantic interoperability, and technical interoperability hierarchically. Henceforth, we may not achieve organizational interoperability, without providing a legal interoperability foundation. Leading to a sequence of constraints, on which legal interoperability is the precursor. The EIF defines legal interoperability as the capability of companies with different legal frameworks, policies, and strategies to work together, by dealing with their differences through negotiation. Furthermore, the EIF has recently appended two extra layers for interoperability, leveraging the collaboration of public services within the industry, i.e., interoperability governance and integrated public service governance. Nevertheless, for the scope of this paper, we comply with the four *fundamental* layer of interoperability, which ground the RAM.

To properly understand the domain of legal interoperability within IDS, and its similar data ecosystems, we proposed an SLR, which employed 6 different online databases and thoroughly analyzed 40 papers, with a detailed screening and systematic methodology. Through this SLR, we were able to provide several open issues regarding the legal aspects within IDS, especially the lack of legal interoperability and a single cohesive machine-readable language to represent service contracts. As an example of the few efforts toward legal aspects, the IDSA Dataspace Protocol[2] specifies schemas and protocols required from entities to publish data and negotiate data usage policy agreements. However, it lacks explicit guidance on enforcing legal restrictions and compliance in an IDS-based business ecosystem. All supplementary material of the proposed protocol (e.g., the performed SLR) is available in an open-access GitHub repository[3].

In addition to the RAM, IDSA provided the IM [4], which consists of an RDFS/OWL ontology encompassing fundamental concepts for describing actors in a data space. Although it lacks depth, it provides a road map towards its enhancement by IDSA and through community integration (each request should endeavor a systematic evaluation process). Based on that depth lack, the possibility of continuous integration, and key concepts retrieved by the SLR, through this work, we propose the Legal Interoperability Ontology for International Data Spaces (LegIOn-IDS), which encompasses the gap in the literature regarding legal aspects of IDS, leading to legal interoperability. Hence, we propose a domain ontology that better describes the legal aspects regarding International Data Space transactions, data exchange, and contract negotiation in a machine-readable way. We may present the main contributions of this paper as follows:

- Providing an ontological representation of the legal aspects domain regarding the IDS infrastructure.
- Providing an ontological relationship among the grounding blocks, yet regarded as distinct ideas i.e., IM, ODRL[4], and European Interoperability Framework. Aligning the idea of service contract provided by the Service Contract Ontology.
- Providing an unambiguous machine-readable language that can feasibly describe a service contract in IDS.
- Providing a clear concept of Legal Interoperability in IDS, and how to achieve it.

The remainder of this paper is organized as follows: Section 2 provides an overview tackling the ontological artifacts that grounded the development and design of LegIOn-IDS. Section 3 showcases the development of the ontology, providing an in-depth explanation of the proposed methodology, and thorough detailing of the ontology engineering process (reference ontology and operational ontology). Furthermore, Section 4 provides the evaluation of the ontology (verification through SPARQL queries),

[1] https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/european-interoperability-framework-detail

[2] https://docs.internationaldataspaces.org/dataspace-protocol/

[3] https://github.com/VictorBenoiston/towards_legal_interoperability_IDS_archive

[4] https://www.w3.org/TR/odrl-model/

and validation, through its real-world representation. Finally, we draft the conclusions, and present future works in Section 5.

2. Ontological Foundations

Overall, we must follow the framework provided by the IM [4]. As an addition to the RAM, the IM works as a semantic framework (knowledge base) that represents the domain of IDS, and its sole goal is to provide a foundation of the main concepts, allowing the specific growth of each environment. The RAM is aligned with the concept of usage control and a formal description of permissions and obligations. The IM insubstantially tackles this approach, by implementing the Open Digital Rights Language (ODRL), which provides the terms and concepts for these statements. The ODRL provides a vocabulary to express policies in a flexible and interoperable model, using policies to represent allowed and denied actions over certain assets, as well as obligations and constraints. The IM employs the ODRL to ground the proposed IDS usage control language [5], which is the adopted language to represent the IDS contracts. It provides an architecture of an IDS contract as the contract metadata and the usage control rules, which could be permission, obligation, or prohibition (as a minimal part of ODRL proposes). However, this IDS contract comprises 21 policy classes and is represented as 21 subclasses on the ontology, with no relationship or nested concepts. We propose further detailing of these constructs and defining their implication with a keen goal of enhancing the solely descriptive vocabulary to legally bind and enforce statements, and the results of the SLR incorporate this observation.

Furthermore, another important grounding ontology of LegIOn-IDS is the Service Contract Ontology (SCO), which targets the legal approaches regarding the awareness and compliance of imposed rules and explains the legal positions of participants in a service relation, clarifying their roles and actions [6]. The ontology is founded by the Unified Foundational Ontology (UFO), especially the high-level concepts available in UFO-L [7], which is a legal core ontology, and UFO-S [7], which is the ontology of services. Primordial concepts are recovered from SCO, such as legal moments, legal agreements, service provider burdens and entitlements, and, similarly, service customer burdens and entitlements, along with their respective claims and commitments. Such concepts are not yet available in IDS, even though it is presented by the RAM (currently version 3.0).

Moreover, the Unified Foundational Ontology (UFO) is an axiomatic theory (conceptual model) developed by joining several theories of formal ontologies i.e., philosophy, cognitive science, linguistics, and philosophical logic. Essentially, it was firstly organized into three main fragments, UFO-A, which is the main fragment, is called the ontology of *endurants* i.e., individuals that exist in time with all their parts, having accidental and essential properties that may qualitatively change while maintaining their numerical identity through time (e.g. a cat might be a kitten in a given time, but fully-grown in another). Along with the primordial concepts such as particular, universals, and moments, UFO-A also approaches the definition of concepts as relators and relations, all important concepts used as foundation. UFO-B, the ontology of *perdurants* i.e., individuals that accumulate temporal parts, only existing partially in the present, hence, at different time instants, their current properties may vary, the main concept retrieved from UFO-B is event. An event is essentially a transformation from pre-state to post-state situations, ontologically dependent on its participants. Finally, UFO-C, the ontology of social entities, is based on the latter two fragments and sets important definitions such as agents, objects, and normative descriptions. Normative descriptions are an important concept in our ontology once they define one or more rules recognized by at least one social agent (such as legal norms, data acts, governmental frameworks, etc.). From economics [8] to biology [9], it has been proven that UFO is a solid foundational ontology, leading to a better understanding of the proposed domain.

As one of our prior goals, we must provide an association among the IM precise description of IDS foundation (based on the RAM), SCO as the representation of a service contract domain, and ODRL as the proposed machine-readable policy language, combining different atomic nuances and allowing

[5] https://docs.internationaldataspaces.org/ids-knowledgebase/v/ids-g/UsageControl/Contract

the further enhancement and enforcement of Legal Interoperability in IDS, providing a grounding foundation of ontology reuse [10].

3. LegIOn-IDS: Legal Interoperability Ontology for IDS

This section provides the overall ontology engineering, providing (1) the methodological framework, (2) the requirements elicitation, (3) the reference ontology overview, (4) the operational ontology overview, and finally (5) a short presentation of the instances populated in the ontology. For the core foundation of this LegIOn-IDS, we employed the SABiO methodology, which is composed of five systematic steps i.e., requirements elicitation, ontology capture and formalization, design, implementation, and testing. Furthermore, it also proposes 5 support processes, enforcing the iteration of such steps. Moreover, we also partially applied its recently developed version, SABiOx [11], which proposes a modularization of the ontology, for agile development methodology, and easier maintenance.

3.1. Ontology Requirements

(SUÁREZ-FIGUEROA et al., 2012)[12] propose a systematic approach to document and set goals, granularity, and vocabulary for the proposed ontology. The lack of direction and well-documented goals and foundations for ontologies may lead to *construct overload* or *construct deficit* [7]. We may propose the so-called Ontology Requirements Specification Document (ORSD) through the systematization of such a document. The ORSD is summarized as follows:

- Purpose: Provide a legal interoperability domain description to foster an unambiguous presentation of policies, i.e., service contracts, within IDS architecture.
- Scope: A knowledge-based approach ranging from a semi-formal to a semi-formal degree of formality, based on a middle-out architecture. The proposed ontology is strictly a domain ontology, with foundational integration of top-level ontologies, such as UFO.
- Language: OntoUML with further translation to OWL.
- Intended Uses: Unambiguous Policy representation, and Legal Interoperability Endeavor, leading to secondary uses, such as Contract Automation.

Furthermore, the requirements are twofold presented as Functional Requirements (FR) and Non-Functional Requirements (NFR). The FRs must provide boundaries to the ontology's purpose, whereas the NFRs delimit efficiency, design, and performance aspects. For LegIOn-IDS, we propose the following NFRs: (Design) Support for natural language (English); (adaptation) Addresses the EIF; (Design/Performance) Follow the FAIR principles - Findable: It must have open access, Accessible: It shall possess unique URIs, Interoperable: Use a formal and broadly applicable set of concepts and languages for representation, finally, Reusable: Data meet domain-relevant accepted standards (we provide a complete ISO dictionary of terms in the GitHub repository), and holds a clear and accessible usage license (MIT); (design) UFO and SCO as foundational ontologies; (Design) Available at an open GitHub Repository. Finally, the FRs are translated into Informal Competency Questions (natural language) and Formal Competency Questions (SPARQL queries). Table 1 shows the Informal Competency Questions.

Finally, the Formal CQs - represented in SPARQL queries - and all supplementary material are available in an open-source Github Repository[6].

3.2. Reference Ontology

After identifying the purpose, elicit the requirements, and capture and formalize the reference ontology (1st and 2nd steps). The support process of knowledge acquisition ought to be employed as well. For reference ontology stage of LegIOn-IDS, we follow the design assumption proposed in the NFRs and ground the ontology with the presented foundational ontologies. Furthermore, the reference ontology

[6] https://github.com/VictorBenoiston/legal_interoperability_IDS_ontology

Table 1
Functional Requirements - Informal Competency Questions

Informal Competency Questions Group 1: SCO and EIF Related	
ID	**Informal Competency Question**
CQ1	What are the legal entitlements of the service provider 1?
CQ2	What are the legal burdens/lacks of the service provider 1?
CQ3	What are the legal entitlements of the service customer 1?
CQ4	What are the legal burdens/lacks of the service customer 1?
CQ5	What are the interoperability barriers in service contract 1?
CQ6	Which contracts represent joint controllership?
Informal Competency Questions Group 2: IM and RAM related	
CQ7	What are the permissions and duties of data user 2?
CQ8	Which service contracts are characterized as data rent?
CQ9	Which service contracts are characterized as data purchase?

has been developed using the OntoUML language, which is provided through a plugin for the modeling tool Visual Paradigm[7]. Using such a tool aligns with the first and second steps of our grounding methodology, which allows the informal axiom definition (classes, referred to on UFO as *types*) (e.g., kinds, subkinds, categories, mixins) - and formal axioms (e.g., characterization, mediation) were modeled using UFO stereotypes. The reference ontology comprises five views (granular topologies), alluding to the SABiOx modularization principles. The views are presented in a further section.

3.2.1. Main View

The main view encapsulates the main concepts tracing a parallel between the IDS core participants, and a service contract. For this view, we ought to consider the assumption that *in most cases, the Data Owner acts as the Data Provider and the Data Consumer acts as the Data User*, nonetheless, this view is henceforth documented. **IDS Core Participant**, as a *kind*, holds the identity of the fundamental participants every time a data exchange o. The IDS Core Participant may play three different *roles* (based on our assumption), i.e., *Data Owner*, *Data Consumer*, and *App Provider.* As proposed by UFO, the *role* stereotype is dependent on its *bearer*, and has a dynamic nature, allowing the same individual to play different *roles* at the same time, or in different slots of time, however, this liability may be blocked by a disjointness axiom (further exploited in the operational ontology).In ontoUML, generalization means correspondence, e,g., every *Data Owner* is an *IDS Core Participant*, but not every *IDS Core Participant* is an *Data Owner.* Furthermore, a *Data Owner* acts as a *Data Provider*, and a *Data Consumer* as a *Data User.* Each *Data User* is equivalent to a *Service Customer*, whereas *Data Provider* is equivalent to *Service Provider.* These two possible roles of a *Contractual Party* hold the essence of disjointness, not allowing simultaneous acting. A **Contractual Party** holds the stereotype of *kind*, once it holds the identity of actors involved in a *Service Contract* [7].

Moreover, each *Service Provider* and *Service Customer* has its own *Governing Law* [13], and it is composed of a *Data Protection Law* and a *Competition Law.* When two companies under different jurisdictions (own *Governing Law*) shall collaborate, the EIF proposes the comparison among them, allowing its *compliance*, which is classified as a *relator*, once it holds a *truth-maker* identity, i.e., one must exist in order to two or more individuals to be connected. On its own, the *Compliance* will be guaranteed after policy negotiation. The process has two possible beginnings, by a *Contract Request*

[7] https://www.visual-paradigm.com/

performed by the *Service Customer* or a *Contract Offer*, performed by the *Service Provider*, regardless of the beginning, the policies must be represented in the *Service Contract*, which is also a relator, providing a policy description to provide the aforementioned *compliance* materialized by the Service Customer and service provider. The Main View is available in Figure 1.

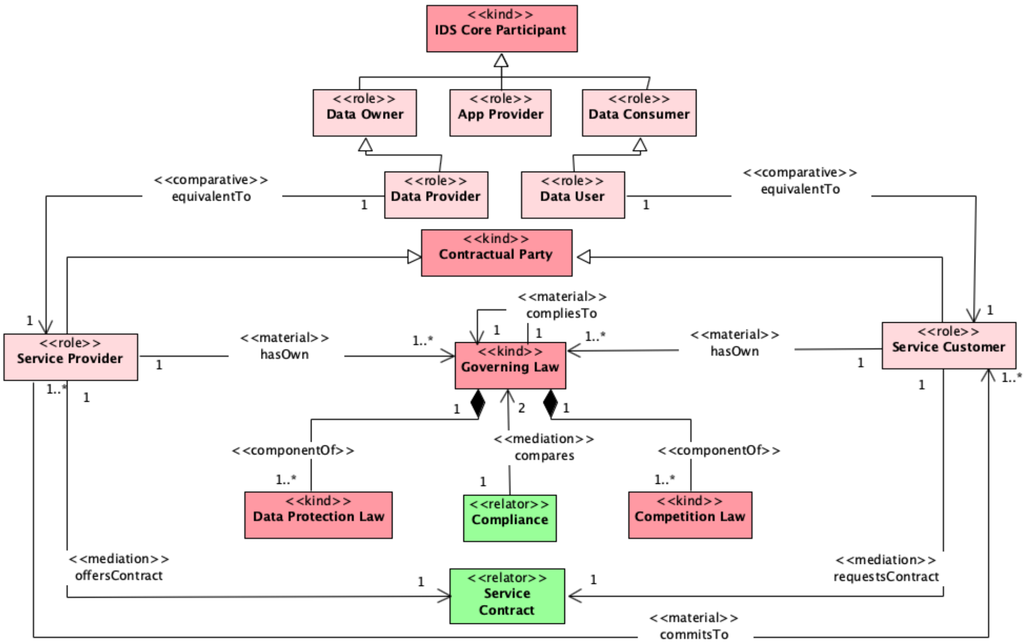


Figure 1: Reference Ontology (Main View)

3.2.2. Service Contract View

This view comprises the aforementioned *Service Contract*. It classifies it as a relator, once it binds the mediation of a *Service Provider* and a *Service Customer* through a non-empty set of *Policies* [13], changing the status of the beares, allowing an intrinsic material relation of commitment, and being existentially dependent. Both *Contractual Parties* must provide a *Usage Consent*, which will further characterize the *Contractual Agreement*. A *Service Contracts* has two phases, the *Contract Negotiation*, which occurs by definition (characterization) of the *Type of Contract*, and *Contractual Agreement*, guaranteed by the agreement of both parties (clear link to the UFO-S [14] lifecycle principles). IDSA proposes two contract models, *Data As a Service* - which allows the *Usage Right* of *Asset* to a Service Customer with a rent purpose, hence, characterizing a temporary transfer. Whereas *Data Purchase* contracts characterize a perpetual transfer, allowing the purchase right of the *Asset*. Those types of contracts are *subkinds* of a *service contract*, once they share the same identity functions, which will be exploited in the operational ontology (e.g., contractual mode, duration, usage rights, licensing terms, etc. [13]). Although we point to those predefined sets of characteristics (service contracts), the RAM defines an IDS contract as an open issue, hence, it is up to the parties to provide a set of policies that enables the IDS architecture. Another negotiated aspect in a *Service Contract* is the *Processing Purpose*, which defines how data will be processed and allows further value creation. The processing purpose may be its *Own Purpose* or *Joint Controllership*. The former refers to the *Service Customer*'s purpose of data usage and further value creation. In contrast, the latter alludes to the joint interest of *Service Customer* and *Service Provider* to create value over data (*Asset*). The overview of this subset is available in Figure 2.

3.2.3. Policy View

As mentioned, *LegIOn-IDS* is grounded by a few perspectives, among which, is the ODRL. This view emphasizes the concept of *Policy*, which is composed of a non-empty set of *Rules*, fostering the unam-

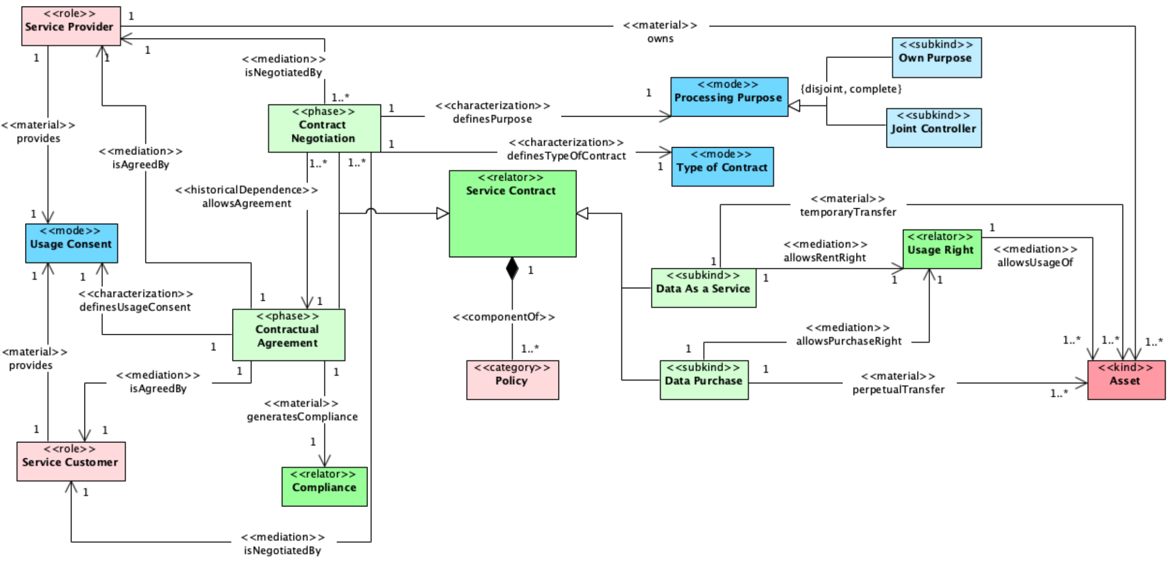


Figure 2: Reference Ontology (Service Contract View)

biguous representation of policies in a contract, leveraging its (semi)automatic negotiation. A *Rule* is an abstract concept that characterizes the permission, duty, or prohibition by some party to perform an *Action* over some *Asset*. It must be abstract, hence, unable to characterize (defined). The ODRL proposes the classification of *Rule* into *Permission Rule*, which allows some *Action*, *Duty Rule*, which obligates some *Action*, and *Prohibition Rule*, which denies some *Action*. The Permission may also have the duty property that expresses an agreed *Action* that must be exercised (as a pre-condition to be granted the Permission) [13]. Similar to a *Contract Negotiation*, a *Policy Negotiation* is the capability of dynamization of specific policies, and as a relator, mediates (performed by) the *Service Provider* and *Service Customer*. Moreover, an *Action* represents an operation on an *Asset*, which is a resource or a collection of resources that are the subject of a Rule and may be classified as *Personal* and *Non-Personal Data*. A *Rule* is seldom characterized by a mode of *constraint*, i.e., a boolean/logical expression that refines an *Action*, and *Asset* collection or the conditions applicable to a *Rule*. This view is available in Figure 3.

3.2.4. Legal Moments View

This view entails the concepts available in the Service Contract Ontology (SCO) [7]. A *Service Contract* is composed of a set of *Legal Moments*, which recovers the concept of *moment* defined by UFO and employs it to the legal positions described by [15]. As a *category*, it holds a general concept, with different identity principles, which may be a *Legal Entitlement*, which implies positivity, or *Legal Lack/Burden*, which implies negativity [7]. The *Legal Entitlements* may be derived into *Power*, *Permission*, *Right*, and *Immunity*, as*kinds*. Whereas the *Legal Burdens/Lacks* are divided into *Duty*, *NoRight*, *Subjection*, and *Disability*. A *Legal Moment* is compared to a *Rule*, which as seen before, may split into prohibition, duty, and permission. By tracing the compatibilities of SCO and ODRL, we might compare the *Legal Entitlements* to *Permission Rules* (once both rely on a positive endeavor), *Legal Burdens/Lacks* may be compared to *Prohibition Rules* (by carrying a negative payload), and finally, the *Duty Rule* may be a placeholder for both *Legal Moments*. This view is represented in Figure 4.

3.2.5. Interoperability View

The last view of LegIOn-IDS (reference ontology) consists of the *Interoperability Checks* proposed by the EIF. The EIF defines Legal Interoperability as *the capability of organizations operating under different legal frameworks, they manage to work together*. This is accomplished by aligning policies and strategies, requiring that current legislation does not block the proposed policies, generating clear

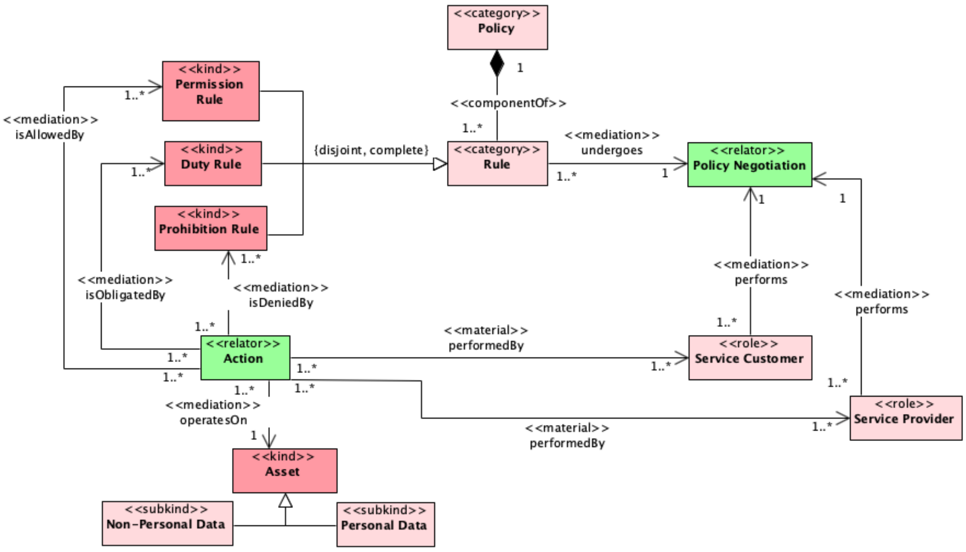


Figure 3: Reference Ontology (Policy View)

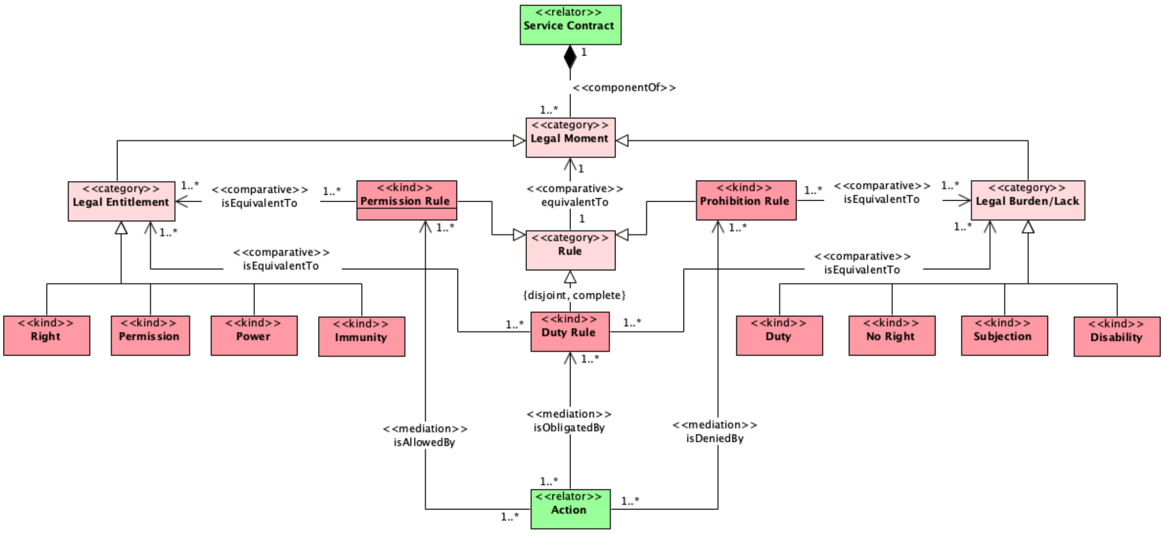


Figure 4: Reference Ontology (Legal Moments View)

agreements on how to deal with those differences across borders, and even allowing the inclusion of a new legislation. Our proposed SLR, however, defines Legal Interoperability as *the semantic capability of unambiguously representing policies and rules among companies while respecting the so-called state sovereignty*. This view treats Legal Interoperability as a consequence, achieved when aligning the proposed checks and alignments compliance. The *Interoperability Check* is a *relator* (event), which mediates (analyzes) the *Data Protection Law*, and *Data Competition Law* of the parties involved in the *Service Contract*. Furthermore, the *Interoperability Check* may find *Interoperability Barriers*, which are classified by the EIF, as *Restriction*, *Data License Model*, *Contradictory Requirement*, *Outdated Security*, *Data Protection Need*, and *Over Restrictive Obligation*. The application of *Interoperability Checks* leads to the definition of *Coherence*, which is the alignment of the analyzed governing laws, and further creates the so-called *Legal Availability*. *Legal Availability* refers to the dynamization potential of aligning the analyzed governing laws, based on their interoperability constraints. Figure **??** describes the proposed view.

3.3. Operational Ontology

The operational ontology should take an in-depth vision of the proposed domain, adding details to the ontology, such as disjointness, closing axioms, relationships, object properties, and data properties. The properties are the *predicate* in a semantic *triple*, which provides a structure of *Subject → Predicate → Object*. Moreover, once we are grounded by UFO, the class hierarchy ought to provide the stereotypes as the foundation. [10] proposes, that once we have a reference model ontology, we might translate it into an operational version, which computer applications could use. To achieve such operational ontology, we must design and implement it in a machine-readable ontology language, such as OWL. To design and develop LegIOn-IDS, we utilized desktop Protégé[8], which allows the employment of an automated reasoner. For our ontology, we proposed the usage of the automated reasoner Pellet [16]. Fostering the principles of Findability and Reuse, the complete operational ontology is available in the previously mentioned open-access GitHub repository.

One of the benefits of using an operational ontology to describe the legal nuances regarding IDS is the unambiguous representation of a service contract. Once we provide enough axioms to describe the identity of a service contract, the reasoning enables the deductive classification of it, as composed by the legal moments of its parties. The ontology is thoroughly documented and fully available in the previously mentioned GitHub repository, along with its glossary of terms[9] [17].

In this representation, we ground the service contract with a *Service Customer*(SC) and *service provider*(SP) (disjoint classes). Each SC and SP has its governing law, and hence, its own set of policies to rely on. Furthermore, as established by IDSA, the contract is composed of usage control rules. We propose a distinct set of policies (rules) for each contractual party. A *Rule* (R), must be either a Duty rule (DR), a Prohibition Rule (ProR), or a Permission Rule (PerR), as stated in Eq. 1. As for a different view, we define the *Legal Moments* (LM) as *Legal Entitlements* (LE) and *Legal Burdens/Lacks* (LBL). As one of the main observations of this document, we compare the concept of Rule to Legal Moment as shown in Eq. 2, specifically, ProR as a Legal Burden/Lack which is defined in Eq. 3, PerR as a Legal Entitlement showcased in Eq. 4, and DR has a twofold equivalent, as Entitlement and Lack/Burden, presented in Eq 5.

$$\forall R(R \rightarrow ProR \lor PerR \lor DR)) \tag{1}$$

$$\forall R(R \leftrightarrow LM) \tag{2}$$

$$\forall ProR(ProR \leftrightarrow LBL) \tag{3}$$

$$\forall PerR(PerR \leftrightarrow LE) \tag{4}$$

$$\forall DR(DR \leftrightarrow LE \lor LBL) \tag{5}$$

Furthermore, each governing law is composed of a data *protection law*, and a *competition law*, and based on the comparison of the existing governing laws, we may state the *interoperability barriers*. Moreover, each *service contract* is composed of a set of lacks and entitlements from the SC, and a different set for the SP as presented in Eq. 6, and each policy (moments) has its own description.

$$\forall SC(SC \rightarrow (\forall LM_{SC}(LM_{SC} \in SC)) \land (\forall LM_{SP}(LM_{SP} \in SC))) \tag{6}$$

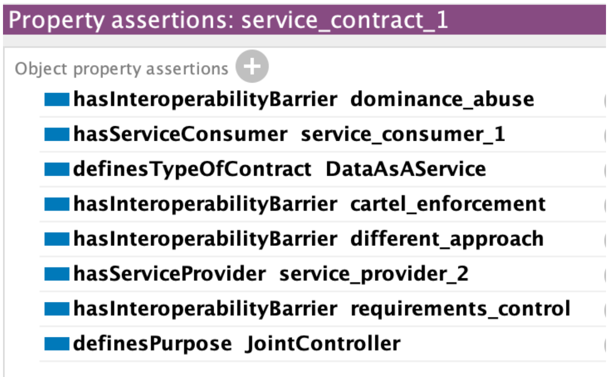
Additionally, we point out the types of service contracts (data as a service and data purchase) as defined classes, hence, they hold the closing axiom of their object properties i.e., (Contractual model, Duration, Usage Rights, Licensing term, Sublicensing, Complying and distribution, Sui generis right of database maker, and usage types, following the contracts matrix provided by [13]).

We partially populated the ontology (Abox), with enough instances to provide clear and concise answers to the proposed CQs (partial completeness), hence, one of the main instances that enables the major part of the axioms, and provides the foundation for different *phases* and *relators* is the *service contract*. To properly verify, and pre-validate the ontology, we instantiated the ontology with generic

[8] https://protege.stanford.edu/

[9] https://legionids.netlify.app/

instances, provided by the SLR, and the contract models provided by IDSA (for data purchase and data rent). Figure 5 provides the overview of an instance of a *service contract*.



Property assertions: service_contract_1

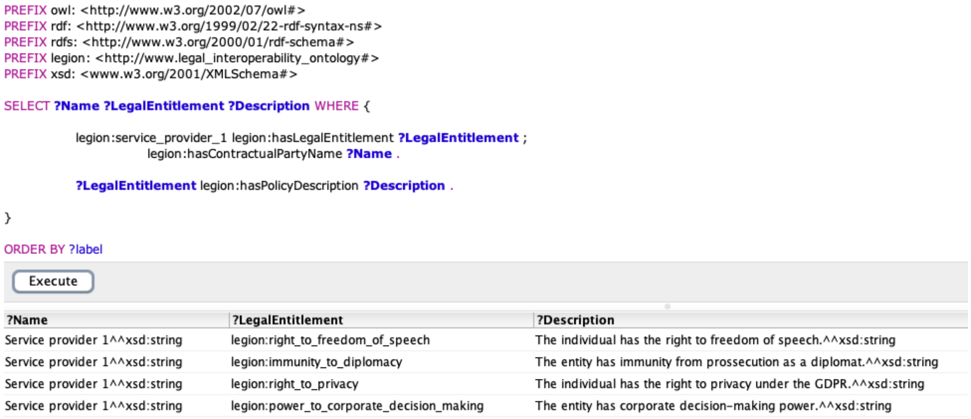
Object property assertions

hasInteroperabilityBarrier	dominance_abuse
hasServiceConsumer	service_consumer_1
definesTypeOfContract	DataAsAService
hasInteroperabilityBarrier	cartel_enforcement
hasInteroperabilityBarrier	different_approach
hasServiceProvider	service_provider_2
hasInteroperabilityBarrier	requirements_control
definesPurpose	JointController

Figure 5: Representation of a generic instance of *service contract*

4. Evaluation

As proposed by [17], the evaluation process is divided into the *verification* and *validation*. Although these works usually uphold semantic similarity, in this context, the former refers to the capacity of the ontology to answer the elicited formal competency question, and for such, we provided SPARQL queries that were able to retrieve knowledge from the instances, for instance, CQ1 is translated and provides the following outcome, showcased in Figure 6.



```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX legion: <http://www.legal_interoperability_ontology#>
PREFIX xsd: <www.w3.org/2001/XMLSchema#>

SELECT ?Name ?LegalEntitlement ?Description WHERE {

        legion:service_provider_1 legion:hasLegalEntitlement ?LegalEntitlement ;
                legion:hasContractualPartyName ?Name .

        ?LegalEntitlement legion:hasPolicyDescription ?Description .

}

ORDER BY ?label
```

?Name	?LegalEntitlement	?Description
Service provider 1^^xsd:string	legion:right_to_freedom_of_speech	The individual has the right to freedom of speech.^^xsd:string
Service provider 1^^xsd:string	legion:immunity_to_diplomacy	The entity has immunity from prossecution as a diplomat.^^xsd:string
Service provider 1^^xsd:string	legion:right_to_privacy	The individual has the right to privacy under the GDPR.^^xsd:string
Service provider 1^^xsd:string	legion:power_to_corporate_decision_making	The entity has corporate decision-making power.^^xsd:string

Figure 6: Formal Competency Question 1 (Translated to SPARQL query) and Outcome

The complete set of SPARQL queries and outcomes are available in the supplementary material. For validation purposes, we must map the real-world scenarios, with the domain established by the ontology, and trace for similarity of its representation. In order to achieve such a step, we provide a demonstration using instances retrieved by the SLR, and contractual models enabled by the IDSA. To properly validate this ontology, we must assess its representation capacity of the real-world domain. To achieve this, we instantiated the ontology using Web Protégé[10], which allows us to generate the instances' map. We provide a map referring to one instance of *Service Customer*, one for *Service Provider*, and the complete set, in a map for *Service Contract*. Those maps are available in the supplementary material.

Finally, in order to properly instantiate this ontology for complete validation and evaluation, we must employ real data, from empirical studies and use cases. For instance, we may populate the ontology

[10] https://webprotege.stanford.edu/

with contractual clauses from companies currently negotiating their clauses (entitlements) in a service contract of IDS.

5. Conclusions

The proposed domain ontology is a component of a research agenda regarding legal interoperability for IDS and acts as an instrument goal. Eventually, the Service Contract will be composed of the legal moments of its Service Customer, and service provider. By doing so, we follow the provided architecture by IDSA, which relates an IDS contract as the contract metadata (initially provided as generic metadata, such as the initiation and finalization time and date). We not only provided a framework that enables the alignment with several employed technologies, such as the framework provided by the IM, and the policy description language ODRL, but applied the equivalence of legal moments for describing such policies, leveraging the machine-readability of the policies, and allowing the comparison of different governing laws (once those are as well composed of policies). With this first step, the development of *LegIOn-IDS* provides a solid foundation for machine-readable contracts representation, fostering the automation of those contracts, and maybe providing the chance to apply machine learning models such as large language models (LLMs) to provide a sample contract based on the metadata provided by the parties, and the overall contractual data (Service Customer's and service provider's legal moments).

As for the scope of this ontology, we accomplished the pointed goals, and as an instrument design, the ontology must enable the description of service contracts by its policies, constraints, and metadata.

The implementation and further use of such ontology are twofold, *Bottom-up*, in which, a machine learning model will be able to provide the textual classification of the given contract, and results in its legal moments, unambiguously representing it, in a machine-readable way (grounded by the IDS architecture), or *Top-bottom*, in which lawyers could use a web interface to input the clauses of the contract, using the ontology as a fundamental schema of relations, and through a machine learning model of natural language processing, generate a model of contract (fostering the data as a service and data purchase kinds of contract, at first).

Holding to those features, and implementing those exploited models, we may propose the automation of the contract negotiation, and providing a base contract, that will act as a framework for performing changes, enhancing the human resources time consumption, and finally, allowing financial and time savings. As an ongoing work, the future works relate to the total instantiation of the ontology, with data from a use case, and the validation of the ontology by a group of specialists through a focal discussion group, or different expert opinions. Furthermore, we also propose implementing a tool that enables the usage of this ontology as the foundation (grounding an ontology-driven development) to provide a base contract that encompasses all the metadata from the instances provided by the ontology. Finally, utilizing machine learning models to comprehend and classify the text provided by the ontology, we may showcase a human-readable contract, that leverages an easier evaluation for law specialists and IDS representatives.

References

[1] H. Bendjenna, P.-J. Charrel, N. E. Zarour, Identifying and Modeling Non-Functional Concerns Relationships, Journal of Software Engineering and Applications 3 (2010) 820–826. URL: https://www.scirp.org/journal/paperinformation.aspx?paperid=2483. doi:10.4236/jsea.2010.38095, number: 8 Publisher: Scientific Research Publishing.

[2] R. U. Ayres, Creating industrial ecosystems: a viable management strategy?, International Journal of Technology Management 12 (1996) 608–624. URL: https://www.inderscienceonline.com/doi/abs/10.1504/IJTM.1996.025505. doi:10.1504/IJTM.1996.025505, publisher: Inderscience Publishers.

[3] M. Jarke, B. Otto, S. Ram, Data Sovereignty and Data Space Ecosystems, Business & Information Systems Engineering 61 (2019) 549–550. URL: https://doi.org/10.1007/s12599-019-00614-2. doi:10.1007/s12599-019-00614-2.

[4] C. Lange, J. Langkau, S. Bader, The IDS Information Model: A Semantic Vocabulary for Sovereign Data Exchange, in: B. Otto, M. ten Hompel, S. Wrobel (Eds.), Designing Data Spaces : The Ecosystem Approach to Competitive Advantage, Springer International Publishing, Cham, 2022, pp. 111–127. URL: https://doi.org/10.1007/978-3-030-93975-5_7. doi:10.1007/978-3-030-93975-5_7.

[5] B. Otto, S. Steinbuß, A. Teuscher, S. Lohmann, Reference architecture model—international data spaces (version 3.0), 2019.

[6] C. Griffo, G. Guizzardi, J. Almeida, Conceptual Modeling of Legal Relations, 2018.

[7] C. Griffo, J. P. A. Almeida, G. Guizzardi, J. C. Nardi, Service contract modeling in Enterprise Architecture: An ontology-based approach, Information Systems 101 (2021) 101454. URL: https://www.sciencedirect.com/science/article/pii/S030643791930506X. doi:10.1016/j.is.2019.101454.

[8] G. Guizzardi, G. Amaral, D. Porello, T. Prince Sales, A Core Ontology for Economic Exchanges, 2020, pp. 364–374.

[9] G. Guizzardi, A. Bernasconi, O. Pastor, V. Storey, Ontological Unpacking as Explanation: The Case of the Viral Conceptual Model, 2021, pp. 356–366. doi:10.1007/978-3-030-89022-3_28.

[10] R. de Almeida Falbo, Sabio: Systematic approach for building ontologies., Onto. Com/odise@ Fois 1301 (2014).

[11] C. Z. d. Aguiar, F. Zanetti, V. E. S. Souza, Source code interoperability based on ontology, in: Proceedings of the XVII Brazilian Symposium on Information Systems, 2021, pp. 1–8.

[12] M. C. Suárez-Figueroa, A. Gómez-Pérez, M. Fernández-López, The NeOn Methodology for Ontology Engineering, Springer, Berlin, Heidelberg, 2012, p. 9–34. URL: https://doi.org/10.1007/978-3-642-24794-1_2. doi:10.1007/978-3-642-24794-1_2.

[13] A. Duisberg, Legal Aspects of IDS: Data Sovereignty—What Does It Imply?, in: B. Otto, M. ten Hompel, S. Wrobel (Eds.), Designing Data Spaces : The Ecosystem Approach to Competitive Advantage, Springer International Publishing, Cham, 2022, pp. 61–90. URL: https://doi.org/10.1007/978-3-030-93975-5_5. doi:10.1007/978-3-030-93975-5_5.

[14] J. C. Nardi, R. d. A. Falbo, J. P. A. Almeida, G. Guizzardi, L. F. Pires, M. J. v. Sinderen, N. Guarino, C. M. Fonseca, A commitment-based reference ontology for services, Information Systems 54 (2015) 263–288. URL: https://www.sciencedirect.com/science/article/pii/S0306437915000228. doi:https://doi.org/10.1016/j.is.2015.01.012.

[15] R. Alexy, A Theory of Constitutional Rights, Oxford University Press UK, 2002.

[16] E. Sirin, B. Parsia, B. C. Grau, A. Kalyanpur, Y. Katz, Pellet: A practical owl-dl reasoner, Journal of Web Semantics 5 (2007) 51–53. doi:10.1016/j.websem.2007.03.004.

[17] M. Fernández-López, A. Gómez-Pérez, N. Juristo, Methontology: from ontological art towards ontological engineering (1997).