An Ontology for Data Regulation

Guillaume Delorme[a], Guilaine Talens and Eric Disson
Institute Laboratoire Magelan, UR Magellan, University Jean Moulin Lyon 3,
Iaelyon School of Management, Lyon, France

Keywords: Ontology, Compliance, Knowledge-based, Privacy.

Abstract: The recent upsurge enactment of regulations seeking to regulate data processing induces a complexification of compliance management for regulated firms. Firms wishing to implement efficient, cost effective and compliant information security and risk management require an increased comprehension of regulatory requirements. Following a previous paper defining Data Regulation Risk, this paper describes an ontology to apprehend the business and operational impacts of regulatory requirements. The ontology is structured to handle various firms' legal context while remaining agnostic of risk management methodologies.

1 INTRODUCTION

Over the past decades, the upsurge enactment of regulations seeking to reinforce the protection of individuals' rights and privacy, economic interests and national security has led to the appearance of a new class of risk called Data Regulation Risk (DRR) (Delorme et al., 2020). We defined Data Regulation as a norm governing data processing and/or ICT governances and processes and/or information technologies and services. Despite addressing similar concepts, such regulations are yet often demanding divergent or particular controls. As an example, the European Dual Use regulation specifies that access control to regulated data must be based solely on individual location and borders (COUNCIL REGULATION (EC) No 428/2009, 2009) while the U.S. Export Arm Regulations (EAR) also includes citizenship based controls (Export Administration Regulation, 2022). Efficient DR management then requires an in depth analysis involving a broad set of skills to apprehend the business and operational impacts of regulatory requirements. This article aims to furnish the necessary information to facilitate DR management.

Different Risk Management (RM) methodologies have already been developed. While classic RM methodologies tend to lack adaptability and turned out to be less effective than initially foreseen (Suh & Han, 2003), RM methodologies evolved to become domain and context specific (Tixier et al., 2002). Regardless of the chosen methodology, accurately identifying risks is crucial as it lays down the foundation for their management. Firms must therefore consider the granularity, exhaustiveness and accuracy of the risk factors' identification (Jallow et al., 2007).

Several authors pointed out the need for ontology in the security domain (Donner, 2003), (Tsoumas & Gritzalis, 2006). While the number of ontology and compliance management articles has been continuously growing over the past years and despite the variety of domain specific ontologies in the different branches of information security, they also tend to apply to only very limited scope. Similarly several conceptualizations of the legal domain have been presented and studies and comparison of legal ontologies can also be found (Larmande et al., 2013), (Visser & Bench-Capon, 1998). While the efforts have shown significant advancements in the field, it is still at an early stage with areas of study left to be thoroughly explored.

Despite important contributions supporting organizations in assessing and managing their risks, there is a need for methodologies and models to identify multi-disciplinary risks like DR management. We seek to address DR by building an ontology which facilitates its management. Ontologies are designed to facilitate the sharing, use and re-use of knowledge (Jones et al., 1998). Defined as explicit conceptualization of a domain (Gruber, 1992), they enable its modulization with the desired level of abstraction depending on the initial objective.

In order to support firms facing multi-disciplinary risks we develop an ontology following the Enterprise

[a] https://orcid.org/0000-0003-0307-1077

Delorme, G., Talens, G. and Disson, E.
An Ontology for Data Regulation.
DOI: 10.5220/0011548400003335
In *Proceedings of the 14th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2022) - Volume 2: KEOD*, pages 66-74
ISBN: 978-989-758-614-9; ISSN: 2184-3228

Model Approach (Uschold & King, 1995). With the ambition of facilitating the apprehension of business and operational impacts of regulatory requirements, the ontology is structured to handle various firms' legal context while remaining agnostic of risk management methodologies. The ontology focuses on regulatory controls while leaving the option of mapping the controls with additional threats for a broader or multi-disciplinary risk management. Our approach is an attempt to design a system capable of precisely and intelligibly representing the various deontic modalities a firm is confronted to while attempting to comply with DR. To reach our target, we use to the extent possible the terminologies of the WordNet database developed by Princeton University (Princeton Wordnet, 2022) as well as concepts present in existing ontologies.

This contribution is structured as follows. Section 2 discusses the core ontology and its purpose. Section 3 presents the building of the ontology. Section 4 sheds light on the usability of the ontology in practice. Finally, section 5 draws conclusions and discusses some further research directions.

2 THE CORE ONTOLOGY ARCHITECTURE AND ITS KEY CONCEPTS

The creation of an ontology requires to determine what entities should be considered and studied. Our model must by default be designed to integrate the fast evolution of the regulatory landscape, the divergent or particular controls as well as being able to focus on a firm specific information systems' environment. Additionally, our model must furnish the necessary information to facilitate decision making and the overall corporate risk management.

2.1 Methodology

Building an ontology is an exciting yet complex endeavor. Few methodologies or guidelines explaining how to build an ontology have been developed over the years resulting in significant differences and disparity among existing ontologies. The variety remains even in ontologies constructed for similar purposes (Visser & Bench-Capon, 1998).

With the ambition of easing methodology building, (Jones et al., 1998) surveyed different ontology building methodologies such as TOVE (Fox et al., 1993), Enterprise Model Approach (Uschold & King, 1995), Methontology (Fernández-López et al., 1997) and Ontolingua (Gruber, 1992). Additional methodologies have been developed to focus on specific domains such as Methodology for building Legal Ontology (Palmirani et al., 2018). As no methodology seems to stand out and all of them have their pros and cons (Visser & Bench-Capon, 1998), we decided to adopt the Enterprise Model Approach which is a stage-based approach, widely spread, providing sufficient freedom of representation which is appropriate to a cross disciplinary ontology such as ours (Pinto & Martins, 2004). Articulated around four main stages, it consists of a skeletal methodology which includes: identify purpose, building the ontology, evaluation and documentation. The second stage incorporates the ontology capture, ontology coding and the integration of existing ontologies.

As opposed to the classic bottom-up and top-down approach to identify the main terms of our ontology, we opt for the middle out approach presented in (Uschold & King, 1995). This approach allows one to identify the primary concepts of the ontology before moving on to specialize or generalize terms, only if they are necessary (Fernández-López et al., 1997). The middle out approach implicitly leads to more stable concepts, less re-work and effort while 'increasing the clarity of the document especially for the non-technical portion of the intended audience'.

In regards to clarity, which is the foundation of the usability and reusability of an ontology, we need a world known, easily accessible, proven and accepted terminology database. Suggested Upper Merged Ontology (SUMO) (Niles & Pease, 2003) is a formal public ontology providing definitions for general purpose terms and is intended as a unifying framework for more specific domain level ontologies. As SUMO is designed as an upper ontology, it provides generic terms and therefore fails to address the needs of more specific domains ontologies (Boer et al., 2009)

We then decide to use when possible the terminologies of the WordNet database developed by Princeton University (Princeton Wordnet, 2022). WordNet "is a large lexical database of nouns, verbs, adjectives and adverbs grouped into sets of cognitive synonyms (Synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations." Each concept, relation or attribute in our ontology is mapped with a unique Synset using the Synset ID. For instance, the relation Govern may refer to "exercise authority over; as of nations" (Synset ID: 202586619), "direct or strongly influence the behavior of" (Synset ID: 202442205) or even "bring into conformity with rules or principles or usage; impose regulations" (Synset ID: 202511551). By specifying the Synset ID, the risk for misinterpretation is therefore greatly reduced while preserving semantic interoperability. For example,

we will use the Synset ID: 202511551 for the relation Govern in our ontology.

2.2 Purpose

As mentioned in (Uschold & King, 1995), being clear about why the ontology is being built and what its intended uses are, is the fundamental step towards developing an ontology. Defining an ontology purpose is accordingly presented as a key element of the specification activities which also happens to consist of the first step in the Methontology methodology (Fernández-López et al., 1997).

2.2.1 Intended Use and Scope

Our approach is an attempt to design a system capable of precisely and intelligibly representing the various deontic modalities or legal modalities (ought, ought not, may, or can) a firm is confronted to while attempting to comply with Data Regulation. DR management complexity resides in the necessity of translating the regulatory constraints and requirements into technical, organizational and operational terms. In other words, our model must be able to deliver pragmatic and concrete information for the users based on generalist and sometimes abstract body of laws. This system must then by default be designed to integrate the fast evolution of the regulatory landscape, the divergent or particular controls as well as being able to focus on a firm specific information systems' environment. Finally, this system must furnish the necessary information to apprehend the business and operational impacts of regulatory requirements.

Our ontology does not seek to assess the effective compliance nor the threat landscape of a company. Our work is solely to express the requirements and constraints based on the deontic models of the laws. In other words, it focuses on regulatory controls and DR while remaining agnostic of risk management methodologies and leaving the option of mapping the controls with existing threats or compliance assessment solutions for a broader or multi-disciplinary risk management.

2.2.2 Scenario of Use

In order to facilitate strategic decision making by providing easy access to information, our system is designed to provide and retrieve the necessary information required to answer the following competency questions :

- What data processing is falling under what regulations ?
- What are the deontic modalities involved ?
- Where is the data processing occurring
- Where can assets or data be located ?
- Who is involved in the data processing ?
- When is the data processing occurring ?
- Under what condition and what context is the data processing occurring ?
- Why are the deontic modalities applied to the data processing ?

For the sake of clarity, we will refer to data processing as operations performed on data, both by manual or automated means. The processing includes the collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction of data (Directive 95/46/EC General Data Protection Regulation, 2016) As for deontic modalities, they refer to the legal modality (ought, ought not, may, or can) (Van Kralingen, 1997). In our scenario of use, they may take the form of mandatory security measures, restriction in data location, restriction in data access, etc.

2.2.3 End Users

As previously mentioned, the complexity of DR management resides in the necessity of translating the regulatory constraints and requirements into technical, organizational and operational terms. Our work is driven by the will to ease the comprehension of the requirements a firm has to comply with by the different domain experts. Not to mention that DR is context specific and depends on one organization's markets, geographical presence and jurisdictions, it therefore requires an in-depth analysis involving a broad set of skills which are usually fragmented across the organization's departments such as legal, IS security, IT operational, Human Resources, etc. In other words, one specific legal requirement may not only require a legal interpretation but also the involvement of security practitioners, IT managers, governance and business experts.

We then identified three main types of users which are: IT managers, security practitioners and compliance managers. All three of them require different pieces of information extracted from the laws in order to perform their duties while ensuring business continuity and their company compliance. For example, the IT manager will need the deontic modalities and regulatory requirements to build and manage the overall information systems while the security practitioners will focus on the mandatory security controls that need to be implemented. Finally, the compliance manager will require an

overview of the regulatory landscape and the global information system scope falling under compliance to perform his tasks.

3 ONTOLOGY BUILDING

3.1 Reused Ontology

While the number of ontology and compliance management articles has been continuously growing over the past years, to the best of our knowledge there is still no existing work that perfectly matches our needs. However, during our search we were able to distinguish two main areas of work that are closely related to ours and that we could reuse to some extent: information security management ontologies and compliance or legal ontologies.

3.1.1 Information Security & Management Ontologies

As show in (Blanco et al., 2008), security ontologies can be sorted by: general security ontology, security ontology applied to a specific domain and theoretical work. They also conclude that seeking to formalize all security concepts is impossible and requires always evolving ontologies. This work was later reused by (Souag et al., 2012) who extended the classification to eight categories, namely: beginning security ontologies, security taxonomies, general security ontologies, specific security ontologies, web oriented security ontologies, risk based security ontologies, ontologies for security requirements and security modeling ontologies. They reached the conclusion that the existing security ontologies vary a lot in the way they cover security aspects and no ontology covers all of the aspects and subdomains of the security domain.

A strong basis for information security domain knowledge may be found in (Fenz & Ekelhart, 2009) Their Information Security Ontology is composed of three sub-ontologies (security, enterprise and location) and is based on established documentation, industry best practices and controls. In their previous work, (Ekelhart et al., 2006) also proposed a security ontology as a basis for a low cost risk management solution as well as an ontology focusing on threats to corporate assets. The ontology consists of five sub-ontologies (threat, attribute, infrastructure, role and person). The role and person concepts enable the ontology to map natural persons while providing a certain degree of liberty by allowing the addition of specific roles if needed.

Other works introduce ontologies specific to vulnerability analysis and management (Wang & Guo, 2009), risk assessment (Tsoumas & Gritzalis, 2006), security annotations of agents and web services (Denker et al., 2003), dependability requirements that include security (Dobson & Sawyer, 2006), secure development (Karyda et al., 2006). Despite the variety of domain specific ontologies in the different branches of information security, they tend to apply to only very limited scope which prevent us from reusing most of them. We will nonetheless reuse the role and person concepts found in (Fenz & Ekelhart, 2009) as much as possible.

3.1.2 Compliance & Legal Ontologies

One of the first international works on legal ontology dates from 1997 (Larmande et al., 2013). Since then, several conceptualizations of the legal domain have been presented and studies and comparison of legal ontologies can also be found (Larmande et al., 2013), (Visser & Bench-Capon, 1998). For instance, the McCarty's Language for Legal Discourse (McCarty, 1989) is semi-formal conceptualization with the ambition of creating a general language for legal domain knowledge. In particular, it enables the expression of certainty in relations and logic rules connecting the certainty. By dividing the domain in three: norm, act and concept description, the issue of reusability of legal ontologies is presented in (Van Kralingen, 1997). The three concepts are designed to be sufficient to conceptualize the subdomains of the legal domain.

There are also ontologies focusing on a single regulation or a type of regulation such as privacy ontologies. For instance, PrivOnto is a semantic framework to represent annotated privacy policies (Oltramari et al., 2018). The solution is oriented to provide a linguistic instrument for the privacy domain as it is based on an ontology which represents legal issues such as data practices in privacy policies. Another example is GDPRtEXT (Pandit et al., 2018) which is a list of concepts present in the General Data Protection Regulation (GDPR). Based on the European Legislation Identifier ontology, the GDPR text extensions expose the GDPR as linked data. Its goal is to provide a way to refer to the concepts and terms found in the GDPR without providing an interpretation of compliance obligations. Similarly, PrOnto, a privacy ontology models the GDPR main conceptual cores such as data types and documents, agents and roles, processing purposes and legal bases (Palmirani et al., 2018). As explained by the authors, "the explicit goal of PrOnto is to support legal reasoning and compliance checking by employing defeasible logic theory".

Similarly to the Frame-Based Ontology, the PrOnto ontology manages to model legal norms through its conceptualization of deontic operators as

right, obligation, permission and prohibition. We will reuse and follow as much as possible these design patterns in order to support DR reasoning.

Finally, LKIF (Boer et al., 2009) is a legal core ontology presented as a knowledge representation formalism that enables the translation between different legal bases. Comparably to the role and person concepts found in (Fenz & Ekelhart, 2009), LKIF presents the organization, role and person concepts which we will be reusing.

3.2 Ontology Capture

The preceding sections presented the requirements for our ontology and some concepts we reuse from existing ontologies. The ontology capture includes the identifications of the key concepts and relationships in our domains. In other words, the ontology capture represents the scoping phase of building our ontology.

3.2.1 Key Concepts

Capturing our ontology implies the findings of precise unambiguous text definitions and terms' identification for the different concepts and relationships (Uschold & King, 1995). We group the top level concepts of our ontology in four subontologies: enterprise, security, legal and location. We used various data sources for the ontology development such as established documentation (National Institute of Standards and Technology, 2022) or industry best practices (International Organization for Standardization, 2022), existing ontologies and regulations.

We reuse the top concepts Individual and Role from (Boer et al., 2009) and (Fenz & Ekelhart, 2009). The concept Individual (Synset ID: 100007846), (ent: Individual $\sqsubseteq \top$) is used to represent an identifiable natural person. The concept Role (Synset ID: 100722061), (ent: Role $\sqsubseteq \top$) and its corresponding subconcepts are used to represent the normal or customary activity of a person in a particular social setting. Every individual has one or more roles which enables a flexible handling of the concepts in complex scenarios such as an external administrator having specific rights on the information system.

The creation of the subontologies enterprise, security and location is derived from (Fenz & Ekelhart, 2009). While the whole subontologies do not fit the needs of ours, we reuse and adapt their concepts Control, Asset, Organization, and Data and Location to create respectively Security_Measure (Synset ID: 100823316), (sec: Security_Measure $\sqsubseteq \top$), Information_System (Synset ID: 103164344), (ent: Information_System $\sqsubseteq \top$), Legal_Entity (Synset ID: 100001740), (ent: Legal_Entity $\sqsubseteq \top$), Technological_Data (Synset ID: 105816622), (ent: Technological_Data $\sqsubseteq \top$) and Country (Synset ID: 108544813).

The concept Technological_Data and its corresponding subconcepts are used to represent data in digital format. For this ontology, we model the subconcepts: Business_Data and Personnal_Data. The former (ent: Business_Data $\sqsubseteq$ Technological_Data) corresponds to data involved in the course of conduct of activities of a Legal_Entity while the latter (ent: Personnal_Data $\sqsubseteq$ Technological_Data) are any information relating to an identified or identifiable natural person.

The concept Legal_Entity and its corresponding subconcepts represent a natural or legal person, a public authority body which carries out an activity whatever its legal form. The following subconcepts modeled so far are: Business_Organization, Independant_Organization and Regulatory_Agency.

We use the concept Information_System to describe an organized set of resources (hardware, software, individual, data and procedure) which makes it possible to process data. In practice, an Information_System may require to involve IT_System that are not the property of the studied organization in the case of cloud services such as Software as a Service. The capacity of representing external ownership within a firm specific information systems' environment is crucial to express the requirements and constraints based on the deontic models of the laws.

Accordingly, we create the concept IT_System (Synset ID: 104377057), (ent: IT_System $\sqsubseteq \top$) and its corresponding subconcepts to represent a combination of interacting elements (resources) organized to achieve one or more desired objectives. We introduce this concept to provide an agile ontological structure according to the granularity of regulations. For this ontology, the following subconcepts have been modeled so far: Data Center, Network, Physical Server, Virtualization, Operating System, Database, Application, Device. To illustrate data processing with sufficient granularity, we add the concept Action (Synset ID: 100037396), (ent: Action $\sqsubseteq \top$) and its corresponding subconcepts to represent something done (i.e. action or processing of data).

We then need to create the concept Functionnal_Process (SynSet ID: 101023820), (ent: Functionnal_Process $\sqsubseteq \top$) to describe a set of interrelated or interacting activities that uses inputs to produce an intended result. As an example, an instance of a Functionnal_Process would be the payroll process within an organization.

As mentioned above, our model uses the concept Security_Measure and its corresponding subconcepts to represent a prescribed countermeasure for an

information system or organization designed to protect the confidentiality, integrity, and availability of its information and to meet a set of defined security requirements. They are usually gathered within different classes of documents. We then create the concept Documentation (Synset ID: 106588326), (sec: Documentation ⊑ ⊤) to represent the set of texts from a Legal Entity. Documentation can take the form of policies, guidelines, procedures or frameworks. The documentation concept is also useful to illustrate external documentation such as NIST standards or ISO frameworks which are often cited in regulations and used for compliance.

Finally, we need the concept Norm (Synset ID: 106532330), (leg: Norm ⊑ ⊤) to describe texts of laws that seek to regulate data processing and/or ICT governances and processes and/or information technologies and services. To show an action that is governed by a regulation through deontic

modalities or legal modalities, we will use the concept Act (Synset ID: 100030358), (leg: Act ⊑ ⊤) as introduced by [23]. Accordingly, a Norm governs an Act which itself governs Individual, Technological_Data, Legal_Entity and Security_Measure.

3.2.2 Key Relationships

Once the primary concepts identified and clearly defined, our next task focuses on determining the relationships between them. Our model consists of two types of relationships: characteristic relationships and action relationships. In the first type, the patient concept does not perform an action directly on the agent concept but makes it possible to specify the attributes of the latter. Characteristic relationships are used to represent the links between the different concepts of the model. On the other hand, action relationships are used when an agent concept performs a direct action on a patient concept. Our model is composed of 11 characteristic relationships (govern, has_a, isLocatedIn, belong, involve, protect, define, manage, isOwnedBy, isComposedOf and create) and 3 action relationships (process, isUsedBy, perform).

The following section will further explain the diverse relations and concepts by the formalization of pieces of regulations. The recent upsurge of regulations seeking to regulate data processing materializes by more than 120 countries already engaged in some form of international privacy laws (Thales Group, 2022). Similarly, firms are coping with an increasing number of export control regulations or other types such as Sarbanes Oxley. As our ontology purpose is not to exclusively provide a theoretical understanding of the legal domain, but to retrieve the necessary information to apprehend the business and operational impacts of regulatory requirements, we will formalize parts of two regulations: the U.S. Export Arm Regulations and the General Data Protection Regulation. Opting for one privacy and one export control regulation from different regulators allows us to cover every concept and relation presented above while ensuring that our model remains agnostic of the type of regulations.

4 THE ONTOLOGY IN PRACTICE

4.1 Formalization of the EAR Supplement No. 18 to part 734

We will take the Export Arm Regulations (EAR), EARNorm is_a (leg: Norm ⊑ ⊤), as our first example. This norm is issued by the United States Department of Commerce, Bureau of Industry and Security (Export Administration Regulation, 2022). There are numerous other legal authorities underlying the EAR listed in the Federal Register documents promulgating it. To illustrate the different primary concepts of our model, we will formalize the EAR Supplement No. 18 to part 734 which states the following :

§ 734.18 ACTIVITIES THAT ARE NOT EXPORTS, REEXPORTS, OR TRANSFERS

Transmitting or otherwise transferring "technology" or "software" to a person in the United States who is not a foreign person from another person in the United States.

Using the concept Act, Supplement No. 18 to part 734 is therefore represented as : EAR734.18Act is_a (leg: Act ⊑ ⊤). Finally, using the govern relation: EARNorm governs EAR734.18Act. Data regulated by the EAR Supplement No. 18 to part 734 then correspond to: EARBusiness_Data is_a (ent: Business_Data ⊑ Technological_Data).

We then need to create a first person using the concept Individual: PersonReceiveingEARData is_a (ent: Individual ⊑ ⊤). Then, this individual must be physically located in the United States (US is_a (loc: Country ⊑ ⊤)) and be an US citizen (USCitizenship is_a (loc: Citizenship ⊑ ⊤)). PersonReceiveingEARData isLocatedIn US and has_a USCitizenship. We can proceed to create our second individual residing in the US who is the sender of the data: PersonSendingEARData is_a (ent: Individual ⊑ ⊤) and PersonSendingEARData isLocatedIn US.

Translating the term Release as mentioned in Supplement No. 18 to part 734 into practical and technological terms would result in granting or

receiving access allowing the consultation of EAR controlled data. To encapsulate this in our model, the concept Action will be used to represent the transfer of controlled data: TransferEARData is_a (ent: Transfer ⊑ Action ⊑ ⊤). To add an extra layer of granularity, we can come up with additional subconcepts such as granting access and its reverse, receiving access: GrantAccessEARData is_a (ent: GrantAccess ⊑ Transfer) and ReceiveAccessEARData is_a (ent: ReceiveAccess ⊑Transfer). In the end, Transmitting or otherwise transferring would be : An individual that uses an EARSystem is_a (ent:IT_Sysem ⊑ T) to perform the action TransferEARData that process EARBusiness_Data.

To conclude, we can formalize the EAR Supplement No. 18 to part 734:

EAR734.18Act ⊑ governs ((PersonReceiveingEARData ⊓ isLocatedIn.US ⊓ has_a.USCitizenship) and (EARSystem ⊓ perform.ReceiveAccessEARData ⊓ process.EARBusiness_Data))

EAR734.18Act ⊑ governs ((PersonSendingEARData ⊓ isLocatedIn.US) and (EARSystem ⊓ perform.GrantAccessEARData ⊓ process.EARBusiness_Data))

4.2 Formalization of Article 32 of the GDPR

The previous example introduced several primary concepts (Norm, Act, Individual, Country, Citizenship, IT_System and Action). We will rely on parts of article 32 of the GDPR as a way of explanation for the concepts Legal_Entity, Functionnal_Process, Information_System, Role, Security_Measure and Documentation.

While section 2 of GDPR sets the rules for personal data security, Article 32 focuses on security of processing by stating the following:

1. […] the controller and the processor shall implement appropriate technical and organisational measures to ensure a level of security appropriate to the risk, including […] :

(b)the ability to ensure the ongoing confidentiality, integrity, availability and resilience of processing systems and services;

[…]

4. The controller and processor shall take steps to ensure that any natural person acting under the authority of the controller or the processor who has access to personal data does not process them except on instructions from the controller, unless he or she is required to do so by Union or Member State law.

A controller is defined by the GDPR as a natural or legal person, public authority, agency or other body which determines the purposes and means of the processing of personal data while a processor designates the entity which processes personal data on behalf of the controller (Directive 95/46/EC General Data Protection Regulation, 2016). We therefore use the concept Legal_Entity to illustrate a processor and a controller: EnterpriseProcessor is_a (ent:Processor ⊑ Business_Organization ⊑ Legal_Entity) and similarly: EnterpriseController is_a (ent:Controller ⊑ Business_Organization ⊑ Legal_Entity)

The legal entities subject to GDPR must therefore create SecurityPolicy is_a (sec: Policy ⊑ Documentation) which defines SecurityControl is_a (sec: Security_Measure ⊑ ⊤).

The security measures must protect the firm's information system, process and IT systems leads us to create: EnterpriseIS is_a (ent: Information_System ⊑ ⊤); GDPRProcess is_a (ent: Functionnal_Process ⊑ ⊤) and GDPRApplication is_a (ent: Application ⊑ IT_System).

Considering the regulated data as GDPRData is_a (ent: Personnal_Data ⊑ Technological_Data) and AccessGDPRData is_a (ent: Action ⊑ ⊤), we can determine the following: SecurityControl protects (EnterpriseIS ⊓ isComposedOf.GDPRProcess ⊓ involve.GDPRApplication ⊓ perform.AccessGDPRData ⊓ process.GDPRData)

We then need to create the GDPRUserRole is_a (ent: Role ⊑ ⊤) and an Individual: Individual1 has_a GDPRUserRole and uses GDPRApplication.

SecurityControl protects ((GDPRApplication ⊓ perform.AccessGDPRData ⊓ process.GDPRData) and (isUsedBy.Individual1 ⊓ has_a.GDPRUserRole))

To conclude, the formalization of the mentioned parts of the article 32:

GDPRNorm is_a (leg: Norm ⊑ ⊤)

GDPRArt32 is_a (leg: Act ⊑ ⊤)

GDPRNorm governs GDPRArt32

GDPRArt32 ⊑ governs (EnterpriseController ⊔ EnterpriseProcessor)

GDPRArt32 ⊑ governs (EnterpriseIS ⊓ isComposedOf.GDPRProcess ⊓ involve.GDPRApplication ⊓ isUsedBy.Individual1 ⊓ perform.AccessGDPRData ⊓ process.GDPRData)

5 CONCLUSION AND FURTHER RESEARCH DIRECTION

Based on various data sources such as established documentation or industry best practices, existing ontologies and regulations, we present an ontology able to formalize firms' legal context while enabling

the sharing and reuse of knowledge to support decision making.

Following the Enterprise Model Approach (Uschold & King, 1995), we build an ontology with 14 top level concepts grouped in four subontologies (enterprise, security, legal and location) and 14 relationships. We then successfully formalize pieces of different regulations to test our model. Our model is capable of precisely and intelligibly representing the various deontic modalities a firm is confronted to while attempting to comply with data regulations. With the ambition of facilitating the apprehension of business and operational impacts of regulatory requirements, our ontology is designed to by any type of firm. We are currently developing the ontology using Protégé and implementing it at Solvay (Solvay, 2022), a worldwide chemical company subject to over 30 privacy regulations, 20 export control regulations and additional data regulations.

To further develop the existing ontology, we also plan to integrate further existing information security and risk management ontologies. We believe that combining them will enable the model to further facilitate the role of security practitioners and compliance manager by providing a more holistic risk management with information knowledge from traditional information security threats. With a desire to optimize efforts, we also hope to lead to more efficient risk management by combining regulatory risk and information security risk.

REFERENCES

Delorme, G., Talens, G., Disson, E., Collard, G., & Gaget, E. (2020, December). On the Definition of Data Regulation Risk. In International Conference on Service-Oriented Computing (pp. 433-443). Springer, Cham.

COUNCIL REGULATION (EC) No 428/2009 of 5 May 2009 setting up a Community re-gime for the control of exports, transfer, brokering and transit of dual-use items (EU dual-use), last consolidated version 2018/15/12.

Export Administration Regulation (EAR), 15 C.F.R. § 730 et seq, https://www.bis.doc.gov/index.php/regulations/export-administration-regulations-ear, last accessed 2022/02/20.

Suh, B., & Han, I. (2003). The IS risk analysis based on a business model. Information & management, 41(2), 149-158.

Tixier, J., Dusserre, G., Salvi, O., & Gaston, D. (2002). Review of 62 risk analysis methodologies of industrial plants. Journal of Loss Prevention in the process industries, 15(4), 291-303.

Jallow, A. K., Majeed, B., Vergidis, K., Tiwari, A., & Roy, R. (2007). Operational risk analysis in business processes. BT Technology Journal, 25(1), 168-177.

Donner, M. (2003). Toward a security ontology. IEEE Security & Privacy, 1(03), 6-7.

Tsoumas, B., & Gritzalis, D. (2006, April). Towards an ontology-based security management. In 20th International Conference on Advanced Information Networking and Applications-Volume 1 (AINA'06) (Vol. 1, pp. 985-992). IEEE.

Larmande, P., Arnaud, E., Mougenot, I., Jonquet, C., Rouge, T. L., & Ruiz, M. (2013, May). Proceedings of the 1st International Workshop on Semantics for Biodiversity. In 1. International Workshop on Semantics for Biodiversity (pp. 001-131).

Visser, P. R., & Bench-Capon, T. J. (1998). A comparison of four ontologies for the design of legal knowledge systems. Artificial Intelligence and Law, 6(1), 27-57.

Jones, D., Bench-Capon, T., & Visser, P. (1998). Methodologies for ontology development.

Gruber, T. R. (1992). Ontolingua: A mechanism to support portable ontologies.

Uschold, M., & King, M. (1995). Towards a methodology for building ontologies (pp. 1-13). Edinburgh: Artificial Intelligence Applications Institute, University of Edinburgh.

Princeton WordNet , https://wordnet.princeton.edu/ last accessed 2022/02/20.

Fox, M.S., Chionglo, J., Fadel, F. A Common-Sense Model of the Enterprise, Proceedings of the Industrial Engineering Research Conference 1993

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

Gruber, T. R. (1992). Ontolingua: A mechanism to support portable ontologies.

Palmirani, M., Martoni, M., Rossi, A., Bartolini, C., & Robaldo, L. (2018, October). Pronto: Privacy ontology for legal compliance. In Proc. 18th Eur. Conf. Digital Government (ECDG) (pp. 142-151).

Pinto, H. S., & Martins, J. P. (2004). Ontologies: How can they be built?. Knowledge and information systems, 6(4), 441-464.

Niles, I., & Pease, A. (2003). Mapping WordNet to the SUMO ontology. In Proceedings of the ieee international knowledge engineering conference (pp. 23-26).

Boer, A., Di Bello, M., Breuker, J. & Hoekstra, R. (2009). LKIF core: Principled ontology development for the legal domain. Law, ontologies and the semantic web: channelling the legal information flood, 188, 21.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Official Journal L,(2016).

Van Kralingen, R. (1997, June). A conceptual frame-based ontology for the law. In Proceedings of the first international workshop on legal ontologies (pp. 6-17).

Blanco, C., Lasheras, J., Valencia-García, R., Fernández-Medina, E., Toval, A., & Piattini, M. (2008, March). A systematic review and comparison of security

ontologies. In 2008 Third International Conference on Availability, Reliability and Security (pp. 813-820). Ieee.

Souag, A., Salinesi, C., & Comyn-Wattiau, I. (2012, June). Ontologies for security requirements: A literature survey and classification. In International conference on advanced information systems engineering (pp. 61-69). Springer, Berlin, Heidelberg.

Fenz, S., & Ekelhart, A. (2009, March). Formalizing information security knowledge. In Proceedings of the 4th international Symposium on information, Computer, and Communications Security (pp. 183-194).

Ekelhart, A., Fenz, S., Klemen, M. D., & Weippl, E. R. (2006, December). Security ontology: Simulating threats to corporate assets. In International Conference on Information Systems Security (pp. 249-259). Springer, Berlin, Heidelberg.

Wang, J. A., & Guo, M. (2009, April). OVM: an ontology for vulnerability management. In Proceedings of the 5th Annual Workshop on Cyber Security and Information Intelligence Research: Cyber Security and Information Intelligence Challenges and Strategies (pp. 1-4).

Denker, G., Kagal, L., Finin, T., Paolucci, M., & Sycara, K. (2003, October). Security for daml web services: Annotation and matchmaking. In International Semantic Web Conference (pp. 335-350). Springer, Berlin, Heidelberg.

Dobson, G., & Sawyer, P. (2006, November). Revisiting ontology-based requirements engineering in the age of the semantic web. In Proceedings of the International Seminar on Dependable Requirements Engineering of Computerised Systems at NPPs (pp. 27-29).

Karyda, M., Balopoulos, T., Dritsas, S., Gymnopoulos, L., Kokolakis, S., Lambrinoudakis, C., & Gritzalis, S. (2006, April). An ontology for secure e-government applications. In First International Conference on Availability, Reliability and Security (ARES'06) (pp. 5-pp). IEEE.

Larmande, P., Arnaud, E., Mougenot, I., Jonquet, C., Rouge, T. L., & Ruiz, M. (2013, May). Proceedings of the 1st International Workshop on Semantics for Biodiversity. In 1. International Workshop on Semantics for Biodiversity (pp. 001-131).

McCarty, L. T. (1989, May). A language for legal discourse i. basic features. In Proceedings of the 2nd international conference on Artificial intelligence and law (pp. 180-189).

Oltramari, A., Piraviperumal, D., Schaub, F., Wilson, S., Cherivirala, S., Norton, T. B., ... & Sadeh, N. (2018). PrivOnto: A semantic framework for the analysis of privacy policies. Semantic Web, 9(2), 185-203.

Pandit, H. J., Fatema, K., O'Sullivan, D., & Lewis, D. (2018, June). GDPRtEXT-GDPR as a linked data resource. In European Semantic Web Conference (pp. 481-495). Springer, Cham.

National Institute of Standards and Technology (NIST), https://www.nist.gov/, last accessed 2022/02/20.

International Organization for Standardization (ISO), https://www.iso.org/home.html, last accessed 2022/02/20.

Thales Group, https://www.thalesgroup.com/en/markets/digital-identity-and-security/government/magazine/beyond-gdpr-data-protection-around-world, last accessed 2022/02/20.

Solvay Group, https://www.solvay.fr/, last accessed 2022/02/20.