Exploring the Role of Enterprise Architecture Models in the Modularization of an Ontology Network: A Case in the Public Security Domain

Archimedes A. Detoni[1,2], Gabriel M. Miranda[1], Laylla D. C. Renault[1],
Ricardo A. Falbo[1], João Paulo A. Almeida[1], Giancarlo Guizzardi[1,3], Monalessa P. Barcellos[1]
[1]Ontology and Conceptual Modeling Research Group (NEMO)
Federal University of Espírito Santo (UFES), Vitória (ES), Brazil
[2]Federal Institute of Espírito Santo (IFES), Santa Teresa (ES), Brazil
[3]Free University of Bozen-Bolzano, Italy
archimedes@ifes.edu.br; gabrielmartinsmiranda@gmail.com; laylladuarte@gmail.com;
{jpalmeida, falbo, monalessa}@inf.ufes.br; giancarlo.guizzardi@unibz.it

Abstract — **Ontologies play a key role in semantic interoperability projects, capturing the conceptualization underlying the various systems to be integrated. In the case of large information systems landscapes, a single monolithic ontology often becomes hard to design and maintain. In this setting, it is common to divide the ontological model in consistently interlinked modular ontologies, forming an ontology network. This paper explores the role of Enterprise Architecture (EA) models in defining the ontologies in an ontology network. We report on an exploratory study in the scope of an e-Government interoperability project in the area of public security. In the reported study, an EA model provides guidance in the modularization of the ontology network.**

Keywords — Enterprise Architecture; Conceptual Modeling; Ontology Network; Networked Ontologies; Semantic Interoperability; Ontology Modularization; Exploratory Study; E-Government; Public Security.

I. Introduction

A common challenge in large organizations concerns the interoperability of Information Systems (ISs) that were designed independently. These ISs are decentralized, heterogeneous, independent and non-integrated. A common response to this challenge is the use of an ontology as "interlingua" to allow different ISs to interoperate [1]. In this approach, an ontology is specified to capture the conceptualization underlying the various systems to be integrated, thereby facilitating semantic interoperability.

Ontology-based integration strategies have been employed successfully in a number of settings, with significant attention to ontology-based techniques in the scope of e-Government (e-Gov) initiatives [2][3]. In the case of e-Gov, systems often are: (i) commissioned and maintained by different Public Administration (PA) agencies; (ii) designed to address different tasks; and (iii) positioned to support different business processes. These features have important implications to integration efforts.

First of all, in such a scenario, a solution based on a single monolithic ontology becomes hard to manipulate, use and maintain. According to [1][4], it is more appropriate to divide the ontological model in consistently interlinked modular ontologies, which tend to be easier to design and reuse. In this context, an Ontology Network (ON) is seen as an adequate approach to the development, integration, maintenance and management of a set of interrelated ontologies. Networked ontologies are ontologies that share concepts and relations in the context of a network [4].

Second, given the complexity of e-Gov initiatives, an overview of the various processes and ISs is often required, and integration efforts can thus profit from the practices of the Enterprise Architecture (EA) discipline. Among the many benefits of EA, there is a potential to enhance information sharing through interoperability [5].

Motivated by the potential benefits of both ONs and EA to e-Gov efforts, we have been investigating their synergy in the context of a public security ISs interoperability project. In early steps of this project, we have identified a number of open questions: How to detect interoperability problems among e-Gov ISs? What are the shared concepts between the various e-Gov ISs to be integrated? How to elicit knowledge for an ON that is intended to be used in an integration project? How to define the scope and depth of the ON for this purpose? How to modularize an ON, defining the focus of each networked ontology? The exploratory case study reported in this paper focuses on the last three questions pointed out above, concerning the initial steps in the design of an ON.

An exploratory case study aims to clarify the comprehension of an issue or situation, providing insights into and helping to gather preliminary information to define problems and to suggest hypotheses [6]. According to [7][8], this kind of study is suitable whenever there is no earlier model as a basis of study and when the available literature provides incipient conceptual framework within the concerned scope. Rather than testing a pre-formulated hypothesis, the research aims to develop general principles to account for the previous observations. The idea is to let questions emerge from the situation itself [7], and requires exposure to contextualized situations.

Our study began with the development of an as-is EA business process model for analyzing and understanding the Violent Crime Process (VCP) and related concepts. This strategy helped us to identify the VCP subprocesses (Police Incident Handling, Police Investigation, Preliminary Police Accusation, Indictment, Acceptance of Prosecution, Criminal Trial, and Imprisonment), IS infrastructure that supports these processes and information flow through the processes. Additionally, it provided us with a diagnosis of possible causes for the existing interoperability problems.

Inspired by [9], we decided to face the interoperability problems by using ontologies organized as a layered ON, which we call VCP-ON (Violent Crime Process Ontology Network). Moreover, to support knowledge acquisition for building VCP-ON, we used the EA business process model previously developed as a non-ontological resource [4]. We believe this approach has proven to be helpful because the domain of interest is process-oriented. In this context, the EA model was used for: (i) helping to delimitate the scope of and to point out the main concepts and relations of the overall ON; (ii) guiding the subject domain partition in more specific domain ontologies of the network; (iii) revealing possible relations between the networked domain ontologies; (iv) pointing out common concepts and relations that are present in the networked domain ontologies and must be used to interlink such ontologies.

The remainder of this paper is structured as follows: Section II introduces some basic notions underlying our approach. Section III presents an overview about the context in which we have developed this exploratory study. In this section, we present the developed EA business process model, discuss how this model is used for identifying interoperability issues and modularizing the VCP-ON, and at last present a fragment of the proposed ON. Section IV discusses related work. Finally, Section V discusses some hypotheses we have formulated as a result of this exploratory study along with an agenda for further investigation.

II. BACKGROUND

Ontologies are currently the main approach to overcome semantic conflicts, by making explicit and precise the meaning of the information to be interchanged across ISs that are intended to interoperate [1][3][10][11][12]. However, in large domains, it is prohibitive, and even undesirable, to provide a large monolithic ontology. This is the case, for instance, of e-Gov. When interoperating e-Gov ISs, there is a large number of involved PA agencies, processes and applications. This makes e-Gov a large and complex subject domain, even if we focus on a smaller context, as in the case of the public security ISs interoperability project addressed here. This project, albeit focused on a specific e-Gov subdomain, involves various PA agencies and a variety of ISs that support their administrative and operational processes.

Considering this aspect, developing a set of ontologies to model parts of that domain seems to be more adequate than providing a large monolithic ontology. Modular ontologies tend to be more easily designed (considering a well-known divide-and-conquer approach) and reused (a helpful characteristic for integration purpose) [1][4]. As argued in [4], "*monolithic ontologies are hard to manipulate, use, and maintain. Modular ontologies on the contrary divide the ontological model in self-contained, interlinked components, which can be considered independently, while at the same time participate to the definition of a specific aspect of an ontology. Therefore, modules share the relation that they are common components of a larger ontology, and often include dependencies and alignments to other modules*".

The development of modular ontologies requires an approach that supports the development, integration, maintenance and management of this set of interrelated ontologies. According to [4], an appropriate approach to address this challenge is by means of an ON, i.e. a set of ontologies related together through various relationships, such as alignment and dependency. A networked ontology, in turn, is an ontology included in such a network, sharing concepts and relations with other ontologies.

Ruy et al. [9] argue in favor of organizing ONs in layers. In the background, a foundational ontology should be used to provide the general grounding knowledge for classifying concepts and relations in the ON, as an upper level of abstraction. A foundational ontology spans across many fields and models the very basic and general concepts and relations that make up the world, such as object, event, intrinsic and relational properties, parthood relations, etc. [13][14]. Core ontologies should be used in the ON to represent the domain knowledge in an intermediate level of abstraction. Core ontologies provide precise definitions for concepts in a specific field that spans across different application domains within that field. Core ontologies are built based on foundational ontologies and refine the latter by adding detailed concepts and relations in their respective specific field [18]. Finally, in a yet more refined manner, domain ontologies are used to describe more specific domain knowledge that is not accommodated in foundational and core ontologies.

As shown in Figure 1, in this work, we follow such layered organization of the ON using the Unified Foundational Ontology – UFO [14] as foundational ontology.

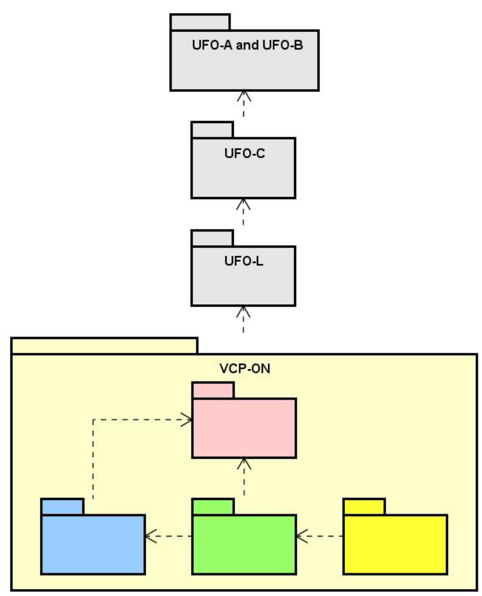

Figure 1. Organization of the VCP-ON in abstraction layers

UFO-A and UFO-B are in the more abstract layer in Figure 1. They are modules of UFO concerned with providing respectively categories of endurants (objects) [14] and perdurants (events) [15]. UFO-C [16] is an ontology of social entities built upon UFO-A and UFO-B, addressing social aspects (e.g., social agent, action, intention, participation, commitments and claims). All these modules are used for offering the ontological foundation of UFO-L [17], a core ontology of legal aspects, which represents essential legal concepts based on a theory of fundamental rights. Finally, the VCP-ON encompasses interlinked domain ontologies, which are grounded in the core and foundational ontologies of the upper layers.

We use OntoUML for representing the ontologies in the ON. OntoUML is an ontologically well-founded profile for UML 2.0 class diagrams, grounded in UFO [14]. OntoUML introduces in class diagrams a set of stereotypes that represents the UFO categorization, enabling the creation of ontologies that are consistent and aligned with that foundational ontology. Table I shows the subset of OntoUML stereotypes used in this paper. A relevant OntoUML feature explored in the ON presented here is the full-fledged mechanism for relationship objectification through the use of *relators* [19]. A *relator* is an objectified relational context, capturing in a structural model how entities relate to each other by playing certain *roles*.

TABLE I. ONTOUML STEREOTYPES SUBSET

Stereotypes	Main Features
<<kind>>	Rigid types that provides a uniform principle of identity for their instances (e.g. Person, Car). A type *T* is rigid iff for all instance *x* of *T*, *x* is necessarily an instance of *T*, i.e., if *x* instantiates *T* in a given world, then *x* must instantiate *T* in all other possible worlds in which *x* exists.
<<collective>>	Rigid types whose instances are collections (e.g., Working Group, Deck of Cards).
<<role>>	Anti-rigid types instantiated within the scope of a relational context (e.g. Student, Spouse). In the case of an anti-rigid type *T*, if an instance *x* instantiates *T* in a given world, there is a possible world in which *x* does not instantiate *T*.
<<relator>>	Types that objectify a material relational context (e.g., a Marriage is a *relator* that associates people playing the role of Spouses).
<<mode>>	Rigid types that capture (potentially complex) objectified intrinsic properties. Their instances are existentially dependent of exactly one other entity (e.g. Skill is a *mode* that characterizes a Person and so is the Intention of an Agent).
<<mediation>>	Formal relationships between a *relator* type and the roles related through that *relator* (e.g., the *mediation* relationship between Marriage and Spouse).
<<characterization>>	Formal relationships between a *mode* and the type that *mode* characterizes (e.g. the *characterization* relationship between a *mode* Belief and a *kind* Person).
<<2ndOT>>	Second-order types whose instances are other types, not individuals (often represented with the power type pattern) (e.g. "Crime Type" is a type whose instances are other types, "Kidnapping", "Homicide")

Besides defining the foundation and the language used to represent the ontologies in the network, we needed guidelines for developing the networked ontologies. In this sense, we have adopted an approach considering the main aspects proposed in the NeOn Methodology [4], which is a method for ON design. NeOn defines detailed processes, guidelines and different scenarios for building networked ontologies. It prescribes an "*ontology requirements specification activity*" as the starting point in the ontology development process. This activity is responsible for defining "*the purpose, the scope, and the implementation language of the ontology network, the target group, and the intended uses of* the *ontology network, as well as the set of requirements that the ontology network should fulfill, mainly in the form of competency questions (CQs)*" [4]. CQs are questions that the ontology should be able to answer. They refine the ontology scope and guide its development, and can also be used as basis for ontology evaluation.

When an ontology is being developed, it is essential to elicit the knowledge that it should contain. After the ontology requirements specification activity, NeOn recommends to carry out a search for candidate knowledge resources – existing ontologies, ontology design patterns, and non-ontological resources (NORs) – that can be reused. Although NeOn includes an ontology support activity for knowledge acquisition (elicitation) [4], it lacks more prescriptive guidelines for performing these tasks.

In this context, we decided to explore the use of EA process-related models as NORs. These models are very useful in process-rich social domains, such as e-Gov, and can be employed as valuable sources of knowledge. In particular, EA models provide mechanisms to organize and represent enterprise systems at subsystem and module levels [20][21], helping to deal with issues related to ontology modularization. Hence, our purpose in adopting an EA approach is to elicit knowledge of the subject domain. These models provide not only a mechanism to systematic structuring knowledge from the subject domain, but also provide tools to analyze and understand the institutional processes and the alignment between the institutional structure and the information technology (IT) architecture.

Here, we adopt ArchiMate as the framework for modeling and structuring EA models. ArchiMate is an EA language that provides a uniform representation for diagrams that describe EA artifacts [22].

III. EXPLORATORY CASE STUDY

According to the Brazilian Health Ministry, between 1996 and 2010, there were almost 1.9 million violent deaths in Brazil, including 710 thousand homicides; and 174 thousand deaths whose basic cause of the event could not be determined by the State. That is, it was not possible to define the basic reason for death in 9.6% of violent events. In developed countries, these incidents of undetermined cause represent a residue less than 1% of all violent cases [23]. As Cerqueira [23] noted, this indicator reveals a huge problem of data quality about violent deaths, which brings serious implications for the State and society. He points out that the main damaging causes for the availability of

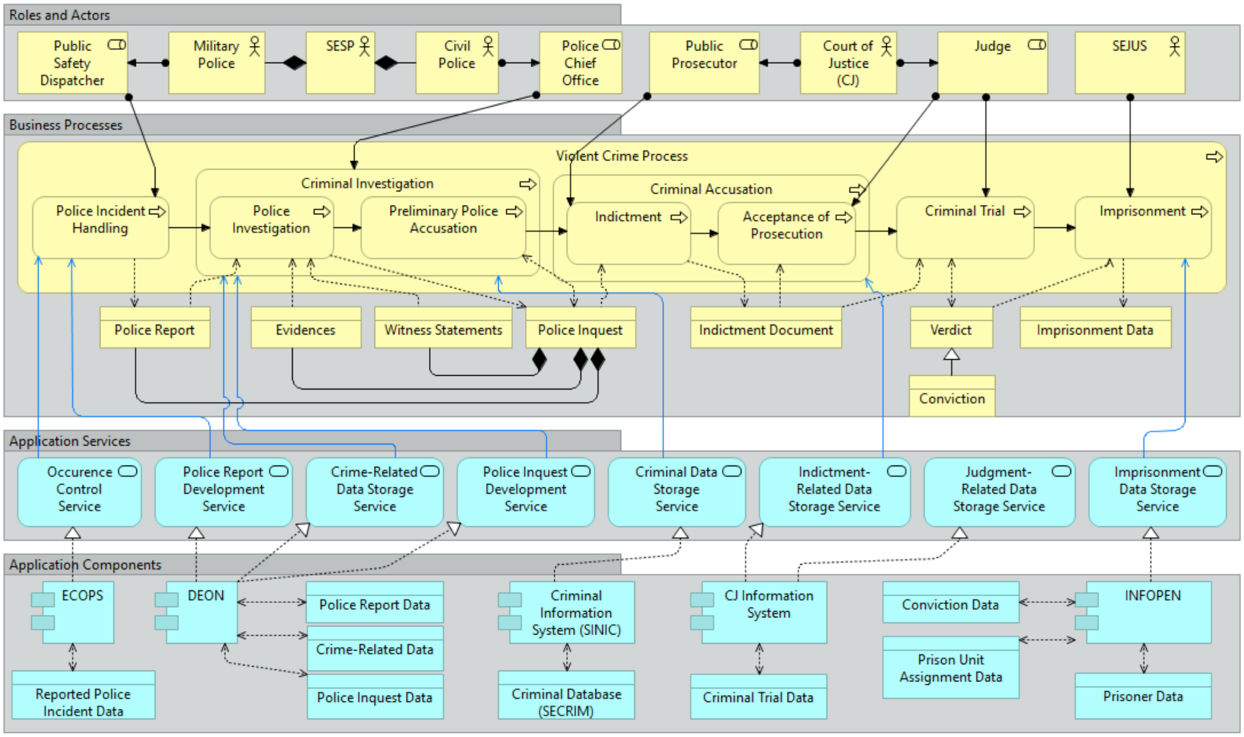

Figure 2. Public Security Agencies As-Is EA Model

consistent and qualified information on crimes and violent deaths in Brazil are problems related to sharing and dissemination of consistent and updated knowledge and information between PA agencies involved in the public security sector. In this context, interoperability is a key element to allow the cooperation and exchange of information between PA agencies. This motivated the project in which this study has been conducted.

The project "Semantic Interoperability of Information in Public Security – SIIPS" aims to improve the quality of criminal processes information in order to support strategic decisions about public security. An ontology-based framework is being developed to allow such improvement, and facilitate the interoperation of the ISs that support the operational and administrative processes performed by the involved PA agencies. For achieving this goal, the following activities are planned: (i) Understand the current processes followed by PA agencies to deal with violent crimes, and ISs infrastructure that supports them; (ii) Identify interoperability gaps in the current approach; (iii) Develop an ontology network (VCP-ON) for dealing with the identified gaps; (iv) Identify scenarios for integrating data from different systems, trying to answer questions that cannot be answered by standalone ISs; (v) Provide integration solutions for these scenarios; (vi) Develop an approach for governance of public security information, based on the results obtained in the previous steps.

We have already accomplished the first two activities, and now we are working on activities (iii) and (iv). To support activities (i) and (ii), we have developed EA models. This paper focuses on the use of the process EA model as basis for developing VCP-ON (activity (iii)). In particular, we discuss how we used the process model to define the networked ontologies that are part of VCP-ON, and to identify ontology relations in the network. In the next subsection, we present the developed EA process model. In the sequel, we discuss how we used this model to define VCP-ON.

A. *EA Process Model*

The first step we conducted was the development of the EA model shown in Figure 2. It is intended to represent the current aspects (as-is model) of the PA agencies involved in the public security sector. This is done by means of: their roles in the so-called *violent crime processes*, the subprocesses performed by each PA agency, the IS infrastructure that supports these processes and the information flow.

The main element of this as-is EA model is the "Violent Crime Process" (VCP). The VCP is a complex business process composed by other business processes, as follows:

- Police Incident Handling: Based on a request received, the military police[1] is triggered to verify the occurrence and to perform all necessary procedures. After these procedures are performed, the public safety dispatcher records information about the police incident (e.g., location, time, possible victim, among others) in

[1] The Military Police is the State police charged with maintaining order. It patrols the streets and imprisons suspects of criminal activity. It is a "militarized" institution (gendarmerie) because it is based on military principles of hierarchy, uniform, discipline, and ceremony.

the police report, which will be used as the basis for establishing the police inquest.

- Police Investigation: At this point, there is already suspicion about the authorship of the crime, but evidence still needs to be raised to confirm the authorship. In this way, the civil police[2] establishes an investigation to determine authorship of the alleged crime. The police chief officer requests testimony from potential witnesses. Police investigators go to the crime scene to search for information. The autopsy and death reports, when ready, are sent by the Scientific Police. All these documents are attached to the police inquest.
- Preliminary Police Accusation: The accusation is based on the Police Inquest, developed in the previous process. After the police investigation ends, if the evidences gathered are sufficient to declare who is the offender (i.e., it is confirmed that it was indeed the preliminary police suspect), an accusation is made.
- Indictment: The police chief officer requests a public prosecutor to offer an indictment (formal complaint). The public prosecutor analyzes the police inquest and defines whether to offer the indictment or not.
- Acceptance of Prosecution: The indictment is sent to a judge, who can decide on its acceptance. If accepted, the police inquest, now turned in an indictment (represented by an indictment document), becomes the criminal procedure that will continue to the Judiciary.
- Criminal Trial: The criminal trial begins when the judicial act is established. The judicial act is a request with which the indictment is manifested, accompanied by an exposition of the fact and the law. A judicial act is usually accompanied by a court hearing, where the grand jury hears the parties through themselves or their lawyers. Parties of a legal proceeding are the defendant (the person against whom the legal action is opposed), victim (who has suffered the offense) and witness (who has seen or heard something and is called to testify. At the end of a criminal trial, the defendant is sentenced and can be subject to a penalty (the defendant receives the conviction) or acquitted.
- Imprisonment: After conviction, the guilty party has to comply with the sentence determined by the judge after conclusion of the criminal trial. Imprisonment here does not necessarily imply confinement in jail, but also encompasses any kind of lawful restraint of a person's liberty imposed after criminal trial.

The VCP subprocesses are performed by the various PA involved in violent crime response: "SESP" (which encompasses the civil and military police), "Court of Justice (CJ)" and "SEJUS" (which encompasses the penitentiary system).

The business objects "Police Report", "Evidences", "Witness Statements", "Police Inquest", "Indictment Document", "Verdict" and "Imprisonment Data" represent the information accessed (or altered) by business processes in the business layer.

As shown in the application layer in Figure 2, there are five ISs (represented by application components) related with the public authorities, named: "ECOPS", "DEON", "Criminal Information System", "CJ Information System" and "INFOPEN". These ISs support the business layer through application services[3] (e.g., "Occurrence Control Service", "Police Report Development Service", "Crime-Related Data Storage Service") and manage the information from the business layer using data objects (e.g., "Reported Police Incident Data", "Police Report Data", "Crime-Related Data").

The access relations between business processes and business objects represent the information shared among business processes. While some business processes create information (e.g., "Police Investigation" creates the "Police Inquest"), others consume information during their execution (e.g., "Indictment" reads the "Police Inquest").

B. *Interoperability Issues*

At a first glance, the model in Figure 2 suggests that information exchange occurs seamlessly between the various subprocesses in the VCP (business layer). However, careful analysis performed in tandem of the business layer and application layer reveals interoperability issues. This analysis was conducted with the following general steps: (i) starting by analyzing the business layer, we can identify that a given business process "BP1" exchanges information with a different business process "BP2", employing business object "BO1"; (ii) this business object "BO1" represents an information asset that should be realized in the application layer by data objects, e.g., the business process "BP1" is supported by an application component "AC1" which access a data object "DO1" and, that in turn realizes the business object "BO1"; (iii) the business process "BP2", which shares the business object "BO1" with the business process "BP1", is supported by another application component, called "AC2". It occurs that, in the application layer, the application component "AC2" is not connected to the application component "AC1" (that supports the business process "BP1"), as well as the data object "DO1" is not shared between the application components "AC1" and "AC2". In short, although *in the business layer* the business process "BP1" (supported by "AC1") exchanges the information asset carried by "BO1" with a different business process "BP2" (supported by "AC2"), *in the application layer*, the data object "DO1", which realizes "BO1", is not shared between the application services "AC1" and "AC2", i.e. there is no ISs integration to support information exchange between business processes.

By using the elements represented in the EA model of Figure 2, we can instantiate a real example of ISs

[2] The Civil Police is the State police with criminal law enforcement duties. It has the function of investigating crimes committed in violation of Brazilian criminal law. It does not patrol the streets.

[3] ArchiMate [22] defines that an "*application service exposes the functionality of components to their environment*". Thus, an application service represents the functionalities of an application component that enable it to support the business process.

interoperability issue in the SIIPS project: (i) the "Police Investigation" process needs to exchange the "Police Inquest" business object with the "Indictment" process; (ii) the "Police Inquest" information is stored in the "Police Inquest Data" database, which is accessed by the "DEON" application, which in turn supports the "Police Investigation" process; (iii) the "Indictment" process, which uses the "Police Inquest" information recorded during the "Police Investigation" process, is supported by the "Criminal Information System" application. However, the applications "DEON" and "Criminal Information System" are not connected, as well as the "Police Inquest Data" database is not shared among these applications, revealing thus an integration issue. The same kind of integration problem occurs in other parts of the VCP.

C. *VCP Ontology Network*

From the diagnosis of the public security IS interoperability problem, considering the complexity and size of the problem, we decided to employ an ontology-based solution, adopting a layered ON architecture as proposed by [9]. The VCP-ON architecture encompass: (i) domain ontologies that capture the conceptualizations of specific subdomains of the public security area; (ii) core ontologies, representing the general domain knowledge that spans across the different subdomains in this field; (iii) a foundational ontology (UFO in our case) that provides a well-grounded classification for the concepts and relations in the networked ontologies. Having a common foundational grounding simplifies the integration of the core and domain ontologies at hand [4][9].

By adopting an ON, we intend to take the advantages of that architecture, which enables a progressive development of the domain ontologies, extending the network, and reusing the concepts already existing in other networked ontologies (favoring the integration and semantic alignment among the conceptualizations captured by the networked ontologies).

However, to develop the networked ontologies, we need guidelines to answer the following key questions: (i) Which are the core and domain ontologies that compose the network? (ii) How should these networked ontologies be organized (i.e., according to which ontology modularization principles)? (iii) Which concepts and relations must compose each domain ontology, and which concepts span the various (sub) domains and can be generalized belonging to a core ontology?

The purpose of VCP-ON is to establish a consensual conceptual model for the domain of violent crimes. The intended use of VCP-ON is to make possible to interoperate the several ISs that support particular subprocesses of the violent crime process. The scope of VCP-ON must encompass all the VCP subprocesses, the PA agencies and person roles involved in them, and the most important information spread along these subprocesses.

The network as a whole should be able to answer competency questions that cannot be answered by a single networked ontology, such as:

CQ1 – What formal suspects investigated for homicide crime have been convicted?

CQ2 – Which police investigations conducted by an investigator led to the effective conviction of the formal suspects?

By using the EA process model to analyze important information on the VCP, such as business process architecture, we can perceive possible partitions for the represented domain guiding the ON modularization, as illustrated in Figure 3.

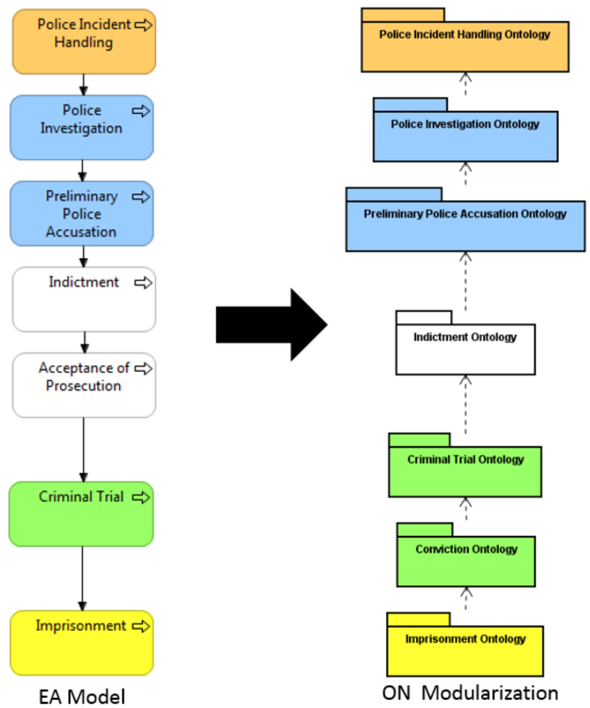

Figure 3. VCP-ON Process-Oriented Modularization

Each color represents a set of VCP subprocesses that gives rise to a module in VCP-ON in Figure 3. We should highlight that a given subprocess can yield to more than one module. This is the case of the "Criminal Trial" subprocess, which yields two modules ("Criminal Trial Ontology" and "Conviction Ontology"). We decided to isolate aspects related to the conviction given its intrinsic complexity. Further, it is also possible for two subprocesses to yield a single module. This is the case of the two subprocesses related to criminal investigation ("Police Investigation" and "Preliminary Police Accusation"), which correspond to a single ontology ("Criminal Investigation Ontology"). The domain ontologies are linked by dependency relations, which reflect historical dependencies. This shows that there are concepts in the domain ontologies that are interlinked, reinforcing the idea of "network". With this 'process-oriented' modularity principle presented in Figure 3, we can revise Figure 1 detailing the VCP-ON domain ontologies and the dependency relations among them and the foundational ontologies as shown in Figure 4.

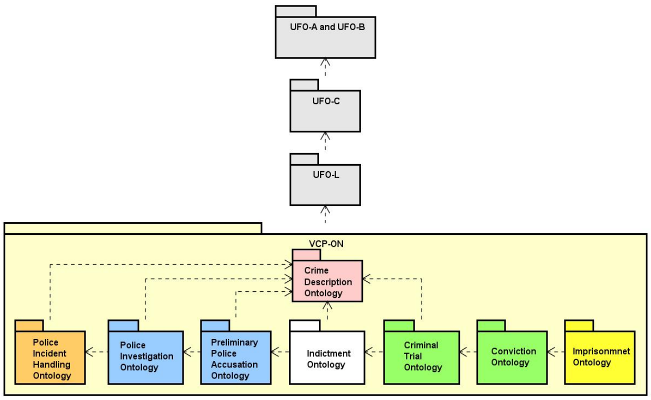

Figure 4. Detailing of the VCP-ON Domain Ontologies

Considering the VCP-ON modularization, we notice that all the VCP subprocesses need to deal with criminal descriptions. Thus, we created the "Crime Description" domain ontology. Figure 5 shows the "Crime Description Ontology" jointly with a fragment of the "Police Investigation Ontology" and the "Preliminary Police Accusation Ontology". We can understand "Crime Description" as the description of the occurred fact, which is composed by other descriptions, such as "Alleged Victim Description", "Alleged Weapon Description" and "Alleged Location Description", among others. The "Crime Description" represents information inherent to concepts of other networked ontologies, exemplifying the interrelation among the VCP-ON domain ontologies, as illustrated in Figure 4 through the dependency relations.

Furthermore, in Figure 5 the "Police Investigation" is a *relator* connecting the "Investigator" (the *role* played by a "Police Chief Officer" when performing an investigation) to the "Preliminary Police Suspects" (the *role* played by a "Person" being investigated). The "Police Investigation" is characterized by an investigation content, which refers to a "Crime Description". This description grounds the indication of some participant as "Preliminary Police Suspect", justifying the relation "refers to" holding between "Alleged Participant Description" and "Preliminary Police Suspect".

Based on the investigation, the "Police Chief Officer" (playing the *role* of "Accuser") can accuse the "Preliminary Police Suspect", which is now the "Formal Suspect" of a "Preliminary Police Accusation". Like a "Police Investigation", a "Preliminary Police Accusation" is characterized by a police accusation content, which refers to a "Crime Description".

We should highlight the historical dependence relation between "Preliminary Police Accusation" and "Police Investigation", and the generalization relation between "Preliminary Police Suspect" and "Formal Suspect". These relations capture two important notions: the first captures the idea that a "Police Investigation" is required for a "Preliminary Police Accusation" to exist; the second shows that the "Formal Suspect" must be a "Preliminary Police Suspect" in the scope of a "Police Investigation". Similar patterns are manifested throughout the whole VCP process.

Figure 6 shows a view of the VCP-ON focusing on the chain of *relators* that reflect the corresponding chain of business process. In this view, the central concepts are the very same *relators* linked by historical dependence relations, according to the VCP subprocesses sequence presented in the EA model. Together with these *relators*, we have a chain of roles (at right side) accounting for the possible roles that an alleged crime suspect can perform along the VCP. The specialization relations reflect dependencies between the roles that are not explicitly represented in the EA model.

Note that the ontology network fragment in Figure 6 is key to answer competency questions that cross different subdomains, such as CQ1 and CQ2 presented before. The various ontologies each of which captures details of the interlinked subprocesses are related through a backbone of

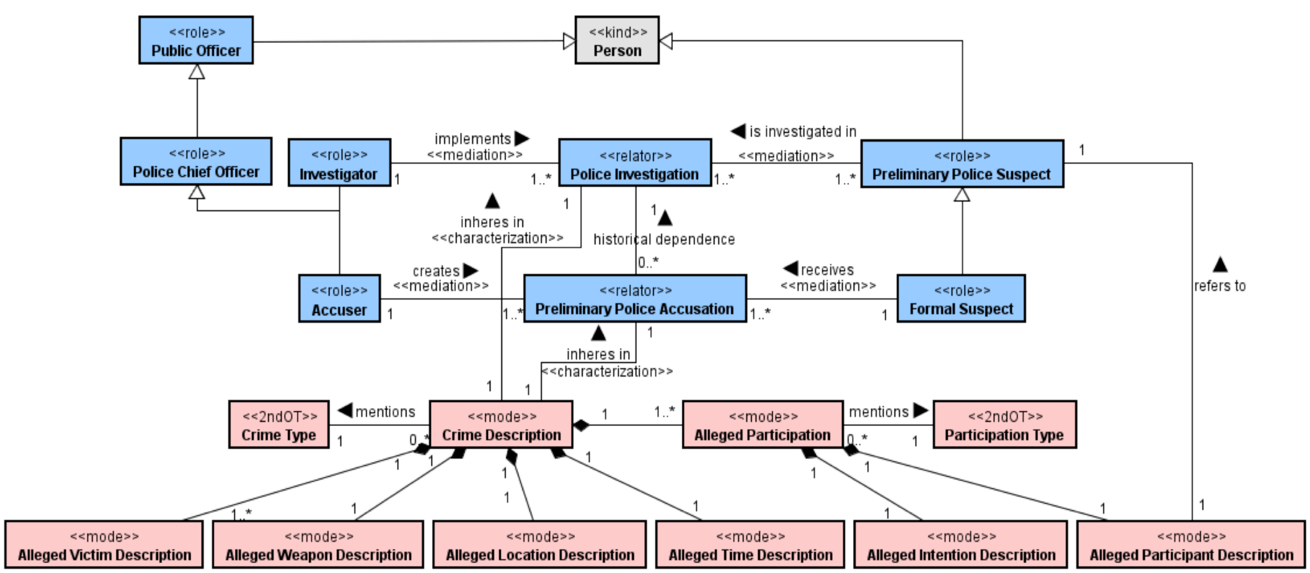

Figure 5. Example of Interrelation among the VCP-ON domain ontologies

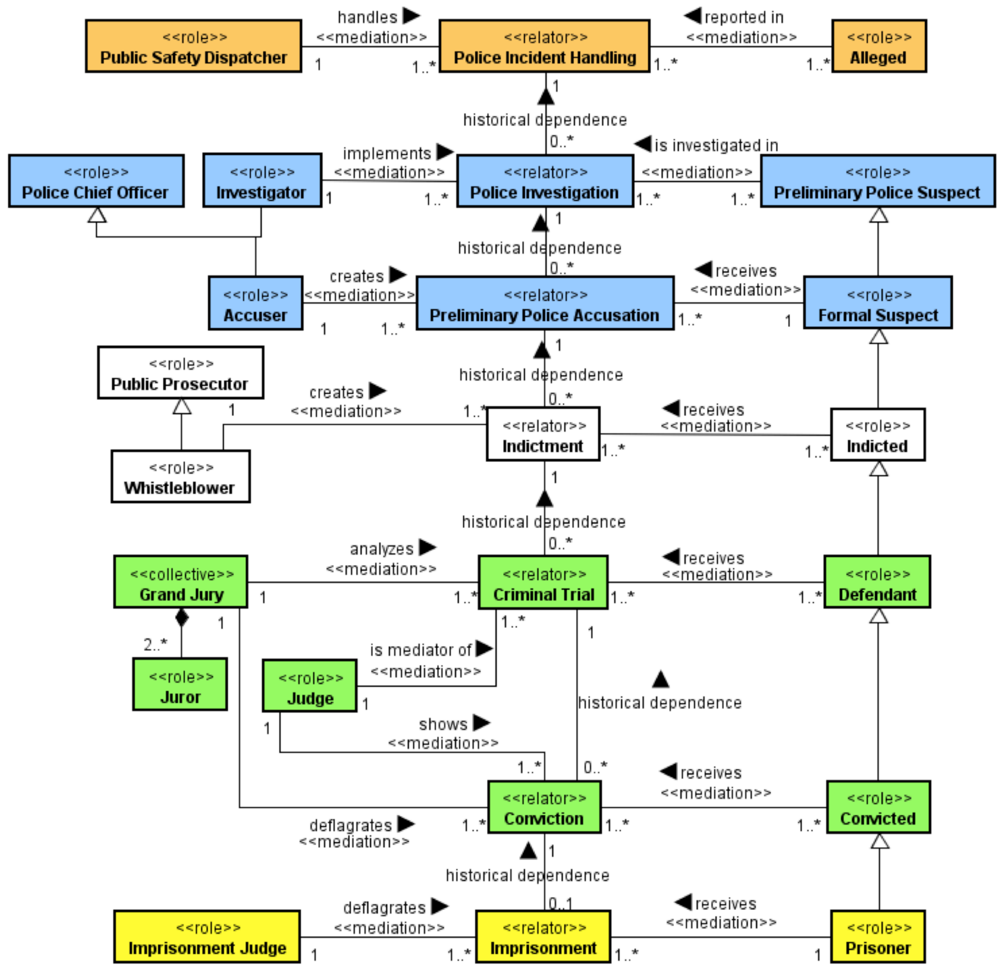

Figure 6. VCP-ON View on *Relators*

relators connected by historical dependence.

IV. RELATED WORK

Many works in the literature, in particular in the context of Ontology Engineering methods such as [4] and [24], have pointed out the importance of modularizing ontologies. However, prescriptive guidelines for modularization are still lacking, and ontology modularization is a very active research area.

Some authors have clearly pointed out that work in ontology modularization is incipient and that modularization is not as well understood in the context of ontologies as it is in software engineering [27]. Because of that, they have explored and evaluated different approaches to modularization, revealing a variety of criteria to evaluate the resulting modules (size, redundancy, connectedness, distance). They have also concluded that the suitability of modularization criteria is highly application dependent.

The work reported in [28] surveys a number of ontology modularization techniques. According to the authors of that study, ontology modularization involves a step of sub-ontology extraction. In [28], we have analyzed, as approaches for modularization, both *network partitioning* and *traversal approaches*. In network partitioning, the ontology is treated as a network of nodes connected by links, and the approaches focus in decomposing the ontology by creating clusters. In the traversal approach, modules are extracted by starting from one or several concepts of the ontology with the further inclusion of the concepts and relations that are linked to these elements. We should note that [27] and [28] investigated approaches which assume the existence of an original and complete ontology, from which all modules are created. Here, in contrast, we propose the identification of modules before the development of an ontology network.

Moreover, [29] performed a survey in the literature of the ontology modularization field, and presented an analysis of several studies on techniques for partitioning the ontology in modules. For example, both in [30] and [31], the authors present results on the characterization of modular ontologies based on structural criteria (e.g. connectedness, size, redundancy of representation, cohesion, and coupling), leading to patterns based on ontology imports from knowledge resources. On the other hand, [32] and [33] present logic-based semantic notions of modularization, which have a main focus on module inseparability, i.e., the module and the source ontology are deemed to be inseparable if they give the same answers to any query. Further, [34] proposed some goals for modularity, e.g., maintenance and scalability for reasoning, with the motivation that the way in which modularization is approached depends on such goals. There is also a list of strategies that are proposed for creating modules, e.g., semantic-driven and structure-driven strategies. In a study on the foundational goals of modularity, [35] classifies ontology modules into different types. There are several different types, e.g., isolating/developing branches of a taxonomy, collecting categories according to a domain, and isolating patterns.

Finally, in contrast with these consolidated works that have proposed several criteria that could be adopted for ontology modularization, in this paper, we have been specifically concerned with the use of EA models to guide decisions in 'process-oriented' ontology modularization activities. Differently from many of these aforementioned efforts, we have focused on the opportunities for modularization that arise in early phases of ON design, and which may be used to guide the definition of requirements for each network ontology to be further designed.

V. CONCLUSIONS

We have focused here on a particular aspect of the interplay between EA models and ontologies: the relation between process-oriented EA models and structural ontology-based conceptual models. Based on the study we have conducted, we hypothesize that the conceptual decomposition introduced naturally in process-oriented models supports the identification of a corresponding modularization principle in the ontology network. This *'process-oriented' modularization principle* manifests itself in a sequence of relations of historical dependencies between *relators*, as shown in Figure 6.

These observations corroborate the conceptual discussion in [19] [25], which characterizes the nature of *relators* and the intimate connection they bear to relational processes that are ubiquitous in social reality. As discussed in depth in those works, *relators* such as marriages, enrollments, employments, service contracts, presidential mandates are the endurantistic (object-like, structural) counterparts of their respective (typically homonymous) relational processes (i.e., the marriage process, the mandate process, etc.). For example, on one hand, the *marriage* real-world (relationally dependent) object (a stable bundle of commitments and claims) gives rise to a number of actions that constitute the marriage process. On the other hand, the marriage process is the life of the marriage object (i.e., the sum of everything that happens to and that can change the properties of the marriage object). For this reason, as discussed also in [26] and observed in this study, *relators* can serve as a natural connection point between structural and dynamic (processual) aspects of reality. Note that the focus here is not on "data objects" but rather on real-world objects (including social objects, commitments, and relationships) that are pervasive in the business world. Thus, representing these objects and their relations to these networked ontologies is key to capture business reality accurately [26].

Despite the correspondences, we have observed that the behavioral and structural models play different yet complementary roles. In particular, the structural ontology-based models provides an in-depth specification of the concepts involved in each activity of a business process and further characterize the business objects that are only identified opaquely in an EA model. This is key in the semantic interoperability tasks envisioned here. After all, the ultimate objective of the developed ON is to provide semantic grounding for concepts that crosscut the various activities, organizations and systems.

The modularization strategy discussed here supports not only the identification of the scope of each individual networked ontology, but also makes the relations between each networked ontology explicit. For example, we can 'trace' a "Conviction" to the "Police Investigation" through a chain of "Criminal Trial", "Indictment" and "Preliminary Police Accusation". This is key for the ontology to answer questions that are posed in the project. This is because these questions cannot be answered if singular ontology are considered in isolation.

Along the study, we have observed that the identified 'process-oriented' modularization strategy can be combined with other modularization strategies. In the studied ontology network, these strategies include layering (with domain ontologies specializing foundational and core ontologies) and the identification of cross-cutting reusable ontologies (e.g., the Crime Description Ontology). Further work is required to investigate the interplay between the different modularization strategies as well as the combination with other strategies identified in the literature.

Based on this study, we can hypothesize that EA models may play a supporting role in other tasks of the Ontology Engineering process. For example, we have observed that process-oriented EA models can support the elicitation of requirements in the form of competency questions. Moreover, we envision that EA motivation models may support the identification of ontology purpose, in light of its required alignment with organizational objectives. These other uses of EA models in the Ontology Engineering process will be reported in the near future.

Finally, we intend to investigate the application of these techniques outside the scope of administrative e-Gov processes. More specifically, we aim to apply these techniques in the scope of scientific workflows in the domain of aquatic toxicology research, in the context of which a second interoperability project is being conducted by our group. By approaching significantly different application domains, we aim to explore whether the modularization strategy we identify here is applicable in a broad range of scenarios.

ACKNOWLEDGEMENTS

This research is funded by CNPq (grants number 311313/2014-0, 461777/2014-2 and 407235/2017-5), CAPES (23038.028816/2016-41), FAPES (69382549), as well as the Free University of Bozen-Bolzano (OCEAN Project). We would like to thank the anonymous EDOC PC members who provided useful comments for improvement.

REFERENCES

[1] M. Uschold, and M. Gruninger, "Ontologies: principles, methods and applications", The knowledge engineering review, vol. 11.02, pp. 93-136, 1996.

[2] H. J. Scholl, and R. Klischewski, "E-government integration and interoperability: framing the research agenda", International Journal of Public Administration, vol. 30.8-9, pp. 889-920, 2007.

[3] V. Peristeras, T. Konstantinos, and S. K. Goudos, "Model-driven eGovernment interoperability: a review of the state of the art", Computer Standards & Interfaces, vol. 31.4, pp.613-628, 2009.

[4] M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, and A. Gangemi, eds. “Ontology engineering in a networked world”, Springer Science & Business Media, 2012.

[5] N. Banaeianjahromi, and K. Smolander, “What do we know about the role of enterprise architecture in enterprise integration? A systematic mapping study”, Journal of Enterprise Information Management, vol. 29.1, pp. 140-164, 2016.

[6] P. Kotler and G. Armstrong, “Principles of marketing”, Addison Wesley Publishing Company, 11th edition, 2006.

[7] R. K. Schutt, “Investigating the social world: The process and practice of research”, Pine Forge Press, 7th edition, 2011.

[8] R. Yin, “Case study research: Design and methods”, Sage Publications Inc, 5th edition, 2013.

[9] F. B. Ruy, R. A. Falbo, M. P. Barcellos, S. D. Costa, and G. Guizzardi, “SEON: A software engineering ontology network”, in Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management. Bologna, Italy, pp. 527-542, 2016.

[10] F. B. Vernadat, “Interoperable enterprise systems: principles, concepts, and methods”, Annual Reviews in Control 31, 2007.

[11] S. Izza, “Integration of industrial information systems from syntactic to semantic integration approaches”, Enterprise Information Systems, Vol. 3, No. 1, 2009. 1-57.

[12] H. Watche et al. “Ontology-based information integration–a survey of existing approaches”, In: Proceedings of the Workshop Ontologies and Information Sharing, IJCAI. 2001.

[13] N. Guarino, “Formal ontology and information systems”, in Formal Ontology and Information Systems, Proceedings of the International Conference on Formal Ontology and Information Systems (FOIS), IOS Press, Amsterdam, pp. 3–15, 1998.

[14] G. Guizzardi, “Ontological foundations for structural conceptual models”, Fundamental Research Series, Centre for Telematics and Information Technology, Netherlands, 2005.

[15] G. Guizzardi, G. Wagner, R. de A. Falbo, R. S. S. Guizzardi, and J. P. A. Almeida, “Towards Ontological Foundations for the Conceptual Modeling of Events”, in 32th International Conference, ER 2013, pp. 327–341, 2013.

[16] R. S. S. Guizzardi, “Agent-Oriented Constructivist Knowledge Management” University of Twente, 2006.

[17] C. Griffo, J. P. A. Almeida, and G. Guizzardi, “Towards a Legal Core Ontology based on Alexy’s Theory of Fundamental Rights”, Multilingual Workshop on Artificial Intelligence and Law, ICAIL 2015.

[18] A. Scherp, C. Saathoff, T. Franz, and S. Staab, “Designing core ontologies”, Applied Ontology, v. 6, n.3, pp. 177–221, 2011.

[19] N. Guarino, and G. Guizzardi, “We need to discuss the relationship: Revisiting relationships as modeling constructs”, Advanced Information Systems Engineering, p. 279–294, 2015.

[20] M. Lankhorst, “Enterprise architecture at work: Modelling, communication and analysis”, in The Enterprise Engineering Series, 2nd ed. Springer: Germany, Berlin, 2009.

[21] D. Chen, G. Doumeingts, and F. Vernadat, “Architectures for enterprise integration and interoperability: Past, present and future”, Computers in industry, vol. 59.7, pp. 647-659, 2008.

[22] The Open Group, “ArchiMate 3.0 specification”, United Kingdom, 2016.

[23] D. Cerqueira; “Map of hidden homicides in Brazil” (text translated). Texto para Discussão, 1848, IPEA, Rio de Janeiro, 2011.

[24] R. A. Falbo, “SABiO: systematic approach for building ontologies”, *ONTO. COM/ODISE@ FOIS*, 2014.

[25] N. Guarino, G. Guizzardi, “Relationships and Events: Towards a General Theory of Reification and Truthmaking”, 15th International Conference of the Italian Association for Artificial Intelligence, Genova, Italy, 2016.

[26] G. Guizzardi, N. Guarino, and J. P. A. Almeida, “Ontological considerations about the representation of events and endurants in business models”, International Conference on Business Process Management, Springer International Publishing, 2016.

[27] M. d’Aquin, A. Schlicht, H. Stuckenschmidt, and M. Sabou, “Ontology modularization for knowledge selection: Experiments and evaluations”. In Database and Expert Systems Applications, pp. 874-883, Springer Berlin/Heidelberg, 2007.

[28] J. M. L. Aparicio, “Ontology view: a new sub-ontology extraction method”, 2015.

[29] Z. C. Khan, and C. M. Keet, “An empirically-based framework for ontology modularization”. Applied Ontology, pp. 171-195, 2015.

[30] S. B. Abbès, A. Scheuermann, T. Meilender, and M. d’Aquin, “Characterizing Modular Ontologies”. In The 6th International Workshop on Modular Ontologies. Graz, Austria, 2012.

[31] A. Schlicht, and H. Stuckenschmidt, “Towards structural criteria for ontology modularization”. In 1st International Workshop on Modular Ontologies. Georgia, USA, 2006.

[32] B. Konev, C. Lutz, D. Walther, and F. Wolter, “Formal Properties of Modularisation”. In Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization, pp. 25–66, Springer, 2009.

[33] F. Loebe, “Requirements for logical modules”. In The 1st International Workshop on Modular Ontologies. Georgia, USA, 2006.

[34] C. Parent, and S. Spaccapietra, “An Overview of Modularity”. In Modular Ontologies, pp. 5–23, Springer Berlin/Heidelberg, 2009.

[35] S. Borgo, “Goals of modularity: A voice from the foundational viewpoint”. In The Fifth International Workshop on Modular Ontologies, pp. 1–6. Ljubljana, Slovenia, 2011.