UNIVERSITAT POLITÈCNICA DE VALÈNCIA

A Framework for Conceptual Characterization of Ontologies and its Application in the Cybersecurity Domain

Beatriz Franco Martins

Advisor: Prof. Óscar Pastor López
Co-advisor: Prof. José Fabián Reyes Román

March – 2024

The thesis was submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Science by the *Universitat Politècnica de València*. Thesis defense on April 12th, 2024.

Author:
Beatriz Franco Martins

Advisor:
Prof. Óscar Pastor López

Co-advisor:
Prof. José Fabián Reyes Román

External reviewers:
Prof. Peri Loucopolus, Loughborough University, United Kingdom
Prof. João Paulo Almeida, Federal University of Espírito Santo, Brazil
Prof. Luiz Olavo Boninho, University of Twente, Netherlands

Examination committee:
President Prof. Peri Loucopolus, Loughborough University, United Kingdom
Secretary Prof. Maribel Yasmina Santos, University of Minho, Portugal
Speaker Prof. María Victoria Torres, University Politecnic of Valencia, Spain

This Doctoral Thesis has been accomplished thanks to the *Generalitat Valencia* for granting pre-doctoral scholarships *Santiago Grisolía* program (GRISOLIA C11273) and *Valencian Graduate School and Research Network of Artificial Intelligence* (ValgrAI 2022–2023), as well to the Accenture Tel Aviv Labs for the research collaboration.

Benedícam Dóminum, qui tríbuit mihi intelléctum: providébam Deum in conspéctu meo sempre: quóniam a destris est mihi, ne commóvear (Ps. 15, 7 et 8).

Tn Memoriam and honor of the author's beloved father
Engr. Hélio Brandão Martins M.Sc., who passed away during
the research and publication of this thesis.

Dedicated to him.

Acknowledgments

With new opportunities come great challenges. This thesis was a great personal challenge for me and my family. We have changed house, country, and continent, building a new life full of challenges and new things. Along the way, we have left many dear people behind. But with a lot of study, perseverance, and the help of many people, we have made it. I would like to express my sincere thanks to all of them.

I begin by expressing my gratitude to my advisor, Óscar Pastor López, who guided me along this path of study and research. It was years of great learning with him, I thank him for sharing his immense knowledge with me and for his support, patience, and dedication. More than a professor, he is an example to be followed, with a big sense of humor; always focused, and communicative, he supported me on this path. Thank you for accepting me on this journey as my advisor, but more for being my friend.

Thanks to José Fabián Reyes Román, my co-advisor who also helped me during all the research and for reviewing the preliminary version of this thesis. I am grateful to my colleagues from the Accenture Israel Cyber R&D Lab (Moshe Hadad, Dan Klein, Gal Engelberg, and Ethan Hadar) who helped with their immense knowledge in cybersecurity. Thanks to my dear old friends and Prof. Renata Guizzardi and Prof. Giancarlo Guizzardi for their valuable advice and support during the course of this research work. Also, thanks to my colleagues Bruno Duarte Borlini and Cristine Griffo for the countless philosophical and ontological discussions that opened me to new possibilities. Thanks to Vicente Javier Julian Inglada of DSIC, Ana Cidad Vila of ValgrAI,

Prof. José Ignacio Panach Navarrete, and the other PROS@VRAIN team members and colleagues. For all of you, who believed in me, especially when I was weakest and thought that the end never to come.

Besides my advisors and research team, I thank my thesis committee Prof. Peri Loucopolus, Prof. Maribel Yasmina Santos, and Prof. María Victoria Torres, the reviewers Prof. João Paulo Almeida and Prof. Luiz Olavo Boninho for agreeing to be a part of the court. It has been a great honor to have your participation in the last stage of the thesis.

In this research journey I made new friends, and I'm very grateful to them for the time we spent together. During our lunches at *La Vella* we formed a group of doctoral students who relaxed while running to catch up with our research hours. I am grateful for the companionship and moments of relaxation and joy we spent together as *The Hours Cartel*. Thanks to Ana, Ana Leon, Ángel, Carlos, Lenin, Nana, Pris, Jairo, and Julio. Special thanks to my personal friends Rocco, Cynthia, and Rafael, who stood up with me when I needed it most.

My greatest thanks to my parents, Hélio Brandão Martins (*in memoriam*) and Thelma Franco Martins, who are my examples of life; to my husband, Paulo Sérgio Pizoni, for being a steady partner; to my beloved son, Paulo Martins Pizoni, the greatest and best gift that God gave me, and to my entire family who directly or indirectly participated with me in this journey.

Thank you for all!

Abstract

Ontologies are computational artifacts with a wide range of applications. They represent knowledge as accurately as possible and provide humans with a framework for knowledge representation and clarification. Additionally, ontologies can be implemented and processed by adding semantics to data that needs to be exchanged between systems. In systems, data is the carrier of information and needs to comply with the *FAIR Principles* to fulfill its purpose. However, knowledge domains can be vast, complex, and sensitive, making interoperability challenging. Moreover, ontology design and development are not easy tasks; they must follow methodologies and standards and comply with a set of requirements. Indeed, ontologies have been used to provide data *FAIRness* due to their characteristics, applications, and semantic competencies.

With the growing need to interoperate data came the need to interoperate ontologies to guarantee the correct transmission and exchange of information. To meet the need to interoperate ontologies and at the same time conceptualize complex and vast domains, *Ontology Networks* emerged. Moreover, ontologies began to carry out conceptualizations, fragmenting knowledge in different ways depending on requirements, such as the ontology scope, purpose, whether it is processable or for human use, its context, and several other formal aspects, making *Ontology Engineering* also a complex domain. The problem is that in the *Ontology Engineering Process*, stakeholders take different perspectives of the conceptualizations, and this causes ontologies to have biases that are sometimes more ontological and sometimes more related

to the domain. These problems result in ontologies that lack grounding and ontology implementations without a previous reference model.

We propose a (meta)ontology grounded over the *Unified Foundational Ontology* (UFO) and supported by well-known ontological classification standards, guides, and FAIR Principles to address this problem of lack of consensual conceptualization. The **Ontology for Ontological Analysis** (O4OA) considers ontological-related and domain-related perspectives, knowledge, characteristics, and commitment that are needed to facilitate the process of *Ontological Analysis*, including the analysis of ontologies composing an ontology network. Using O4OA we propose the **Framework for Ontology Characterization** (F4OC) to provide guidelines and best practices in the light of O4OA for stakeholders. The F4OC fosters a stable and uniform environment for ontological analysis, integrating stakeholder perspectives. Moreover, we applied O4OA and F4OC to several case studies in the *Cybersecurity Domain*, which is intricate, highly regulated, and sensitive to causing harm to people and organizations.

The main objective of this doctoral thesis is to provide a systematic and reproducible environment for *ontology engineers* and *domain specialists* responsible for ensuring ontologies developed according to the FAIR Principles. We aspire that O4OA and F4OC be valuable contributions to the conceptual modeling community as well as the additional outcomes for the cybersecurity community through the ontological analysis in our case studies.

Resumen

Las ontologías son artefactos computacionales con una amplia gama de aplicaciones. Estos artefactos representan el conocimiento con la mayor precisión posible y brindan a los humanos un marco para representar y aclarar el conocimiento. Además, las ontologías se pueden implementar y procesar agregando semántica a los datos que deben intercambiarse entre sistemas. En los sistemas, los datos transportan información y deben seguir los *Principios FAIR* para cumplir su propósito. Sin embargo, los dominios del conocimiento pueden ser vastos, complejos y sensibles, lo que hace que la interoperabilidad sea un desafío. Además, el diseño y desarrollo de ontologías no es una tarea sencilla, y debe seguir metodologías y estándares, además de cumplir una serie de requisitos. De hecho, las ontologías se han utilizado para producir *FAIRness* de datos debido a sus características, aplicaciones y competencias semánticas.

Con la creciente necesidad de interoperar datos surgió la necesidad de interoperar ontologías para garantizar la correcta transmisión e intercambio de información. Para satisfacer esta demanda de ontologías interoperativas y, al mismo tiempo, conceptualizar dominios amplios y complejos, surgieron las *Redes de Ontologías*. Además, las ontologías comenzaron a presentar conceptualizaciones a través de la fragmentación del conocimiento de diferentes maneras, dependiendo de requisitos como el alcance de la ontología, su propósito, si es procesable o para uso humano, su contexto, entre otros aspectos formales, haciendo que la *Ingeniería Ontológica* sea también un dominio complejo. El problema es que en el *Proceso de Ingeniería de Ontologías*, las personas responsables toman diferentes perspectivas sobre las

conceptualizaciones, provocando que las ontologías tengan sesgos a veces más ontológicos y otras más relacionados con el dominio. Estos problemas dan como resultado ontologías que carecen de fundamento o bien implementaciones de ontologías sin un modelo de referencia previo.

Proponemos una (meta)ontología basada en la Ontología Fundacional Unificada (UFO, del inglés, *Unified Foundational Ontology*) y respaldada por estándares de clasificación ontológica reconocidos, guías y principios FAIR para resolver este problema de falta de consenso en las conceptualizaciones. La **Ontología para el Análisis Ontológico** (O4OA, del inglés, **Ontology for Ontological Analysis**) considera perspectivas, conocimientos, características y compromisos, que son necesarios para que la ontología y el dominio faciliten el proceso de *Análisis Ontológico*, incluyendo el análisis de las ontologías que conforman una red de ontologías. Utilizando O4OA, proponemos el **Marco para la Caracterización Ontológica** (F4OC, del inglés, **Framework for Ontology Characterization**) para proporcionar pautas y mejores prácticas a los responsables, a la luz de O4OA. F4OC proporciona un entorno estable y homogéneo para facilitar el análisis ontológico, abordando simultáneamente las perspectivas ontológicas y de dominio de los involucrados. Además, aplicamos O4OA y F4OC a varios estudios de casos en el *Dominio de Ciberseguridad*, el cual es complejo, extremadamente regulado y sensible, y propenso a dañar a personas y organizaciones.

El principal objetivo de esta tesis doctoral es proporcionar un entorno sistemático y reproducible para *ingenieros en ontologías* y *expertos en dominios*, responsables de garantizar ontologías desarrolladas de acuerdo con los Principios FAIR. Aspiramos a que O4OA y F4OC sean contribuciones valiosas para la comunidad de modelado conceptual, así como resultados adicionales para la comunidad de ciberseguridad a través del análisis ontológico de nuestros estudios de caso.

Resum

Les ontologies són artefactes computacionals amb una àmplia gamma d'aplicacions. Aquests artefactes representen el coneixement amb la major precisió possible i brinden als humans un marc per a representar i aclarir el coneixement. A més, les ontologies es poden implementar i processar agregant semàntica a les dades que han d'intercanviar-se entre sistemes. En els sistemes, les dades transporten informació i han de seguir els *Principis FAIR* per a complir el seu propòsit. No obstant això, els dominis del coneixement poden ser vastos, complexos i sensibles, la qual cosa fa que la interoperabilitat siga un desafiament. A més, el disseny i desenvolupament d'ontologies no és una tasca senzilla, i ha de seguir metodologies i estàndards, a més de complir una sèrie de requisits. De fet, les ontologies s'han utilitzat per a produir *FAIRness* de dades a causa de les seues característiques, aplicacions i competències semàntiques.

Amb la creixent necessitat de inter operar dades va sorgir la necessitat de inter operar ontologies per a garantir la correcta transmissió i intercanvi d'informació. Per a satisfer aquesta demanda d'ontologies inter operatives i, al mateix temps, conceptualitzar dominis amplis i complexos, van sorgir *Xarxes d'Ontologies*. A més, les ontologies van començar a presentar conceptualitzacions a través de la fragmentació del coneixement de diferents maneres, depenent de requisits com l'abast de l'ontologia, el seu propòsit, si és procesable o per a ús humà, el seu context i diversos altres aspectes formals, fent que el *Enginyeria Ontològica* també és un domini complex. El problema és que en *Procés d'Enginyeria d'Ontologies*, les persones responsables prenen diferents perspectives sobre les conceptualitzacions, provocant que les

ontologies tinguen biaixos a vegades més ontològics i altres més relacionats amb el domini. Aquests problemes donen com a resultat ontologies que manquen de fonament i implementacions d'ontologies sense un model de referència previ.

Proposem una (meta)ontologia basada en la Ontologia Fundacional Unificada (UFO, de le inglés, *Unified Foundational Ontology*) i recolzada per coneguts estàndard de classificació ontològica, guies i principis FAIR per a resoldre aquest problema de falta de consens en les conceptualitzacions. La **Ontologia per a l'Anàlisi Ontològica** (O4OA, de le inglés, **Ontology for Ontological Analysis**) considera perspectives, coneixements, característiques i compromisos, que són necessaris perquè l'ontologia i el domini faciliten el procés de *Anàlisi Ontològica*, incloent-hi l'anàlisi de les ontologies que conformen una xarxa d'ontologies. Utilitzant O4OA, proposem el **Marco per a la Caracterització Ontològica** (F4OC, de le inglés, **Framework for Ontology Characterization**) per a proporcionar pautes i millors pràctiques als responsables, a la llum d'O4OA. F4OC proporciona un entorn estable i homogeni per a facilitar l'anàlisi ontològica, abordant simultàniament les perspectives ontològiques i de domini dels involucrades. A més, apliquem O4OA i F4OC a diversos estudis de casos en el *Domini de Seguretat Cibernètica*, que és complex, extremadament regulat i sensible, i propens a danyar a persones i organitzacions.

L'objectiu principal d'aquesta tesi és proporcionar un entorn sistemàtic, reproduïble i escalable per a *engineers en ontologies* i *experts in dominis* encarregats de garantir les ontologies desenvolupades d'acord amb els Principis FAIR. Aspirem a fer que O4OA i F4OC aportin valuoses contribucions a la comunitat de modelització conceptual, així com resultats addicionals per a la comunitat de ciberseguretat mitjançant l'anàlisi ontològica dels nostres estudis de cas.

Resumo

Ontologias são artefatos computacionais com uma ampla gama de aplicações. Esses artefatos representam o conhecimento com a maior precisão possível e fornecem aos humanos uma estrutura para representação e esclarecimento do conhecimento. Ademais, ontologias podem ser implementadas e processadas adicionando semântica aos dados que precisam ser intercambiados entre sistemas. Em sistemas, os dados são portadores de informações e precisam seguir os *Princípios FAIR* para cumprir com sua finalidade. No entanto, os domínios de conhecimento podem ser vastos, complexos e sensíveis, tornando a interoperabilidade um desafio. Além disso, o design e o desenvolvimento de ontologias não são tarefas fáceis, devendo seguir metodologias e padrões, bem como cumprir uma série de requisitos. Na verdade, ontologias têm sido utilizadas para produzir *FAIRness* de dados devido às suas características, aplicações e competências semânticas.

Com a crescente necessidade de interoperar dados surgiu a necessidade de interoperar ontologias a fim de garantir a correta transmissão e troca de informações. Para atender a essa demanda de interoperar ontologias e, ao mesmo tempo, conceituar domínios amplos e complexos, surgiram as *Redes de Ontologias*. Ademais, as ontologias passaram a apresentar conceituações por meio da fragmentação do conhecimento de diferentes formas, dependendo de requisitos tais como o escopo da ontologia, sua finalidade, se são processáveis ou para uso humano, seu contexto e vários outros aspectos formais, tornando a *Engenharia de Ontologias* também um domínio complexo. O problema é que no *Processo de Engenharia de Ontologias*, as pessoas responsáveis assumem diferentes perspectivas das conceituações, fazendo com que as

ontologias tenham vieses que são às vezes mais ontológicos e às vezes mais relacionados ao domínio. Esses problemas resultam em ontologias que carecem de fundamentação e implementações de ontologias sem um modelo de referência prévio.

Propomos uma (meta)ontologia fundamentada na Ontologia Fundacional Unificada (UFO, do inglês, *Unified Foundational Ontology*) e apoiada por padrões de classificação ontológica reconhecidos, guias e Princípios FAIR para resolver este problema de falta de consenso em conceituações. A **Ontologia para Análise Ontológica** (O4OA, do inglês, **Ontology for Ontological Analysis**) considera perspectivas, conhecimentos, características e compromissos, que são necessários à ontologia e ao domínio para facilitar o processo de *Análise Ontológica*, incluindo a análise de ontologias que compõem uma rede de ontologias. Usando O4OA, propomos o **Estrutura para Caracterização de Ontologias** (F4OC, do inglês, **Framework for Ontology Characterization**) para fornecer diretrizes e melhores práticas aos responsáveis, à luz da O4OA. O F4OC fornece um ambiente estável e homogêneo para facilitar a análise ontológica, lidando simultaneamente com as perspectivas ontológicas e de domínio dos responsáveis. Além disso, aplicamos a O4OA e o F4OC a vários estudos de caso no *Domínio de Segurança Cibernética*, o qual é complexo, extremamente regulamentado e sensível, sendo passível de danos a pessoas e organizações.

O principal objetivo desta tese de doutorado é fornecer um ambiente sistemático e reproduzível para *engenheiros de ontologia* e *especialistas de domínio*, responsáveis por garantir ontologias desenvolvidas conforme os Princípios FAIR. Aspiramos que a O4OA e o F4OC sejam contribuições valiosas para a comunidade de modelagem conceitual, bem como resultados adicionais para a comunidade de segurança cibernética através da análise ontológica de nossos estudos de caso.

Contents

List of Figures

List of Tables

Part I

INTRODUCTION

Chapter 1

Introduction

Quasi nanos gigantum humeris insidentes,
Bernardo de Chartres (1117 – 1124 d.C.).

Computer Science is a modern science that flourished from the emergence of great technological and mathematical discoveries. From the first transistors to the current Tensor Processing Unit (TPU) [1, 217], and from the Turing Machine [241] to Machine Learning (ML) [257], Artificial Intelligence (AI) [255] algorithms, and *Large Language Models* (LLMs) [35], much has been done. This process produced abundant information, data, resources, and especially challenges. Not only to deal with everything already produced by humanity so far, but also to manage everything derived from this process. However, none of this would be possible without the human ability to create conceptual models. Indeed, the notion that cognitive processes are based on our mental models is not new, and Philosophy already in its foundations supports this statement [82]. Aristotle was the first philosopher to categorize *"things"* [1] to deal with his question of *"what is a being qua-being?"*, inaugurating what we call *Ontology* as a philosophical branch [2].

[1] Of things said without combination, each signifies either: (i) a substance (*"ousia"*); (ii) a quantity; (iii) a quality; (iv) a relative; (v) where; (vi) when; (vii) being in a position; (viii) having; (ix) acting upon; or (x) a being affected [3].

[2] Note that there is a clear difference among an ontology as an ontological artifact and Ontology as a branch of study in Philosophy [83].

From the philosophical perception that Ontology is the *study of being*, Computer Science borrowed its foundations to deal with different kinds of models, usually called *ontologies*. This term began to take on different meanings, depending on the community that dealt with these models [83]. In parallel, the definition of what ontology is has evolved, disclosing the multidisciplinary aspect of ontologies. Gruber defines an ontology as *"an explicit specification of a shared conceptualization"* [72, p. 1]. Borst defines as *"a formal specification of a shared conceptualization"* [30, p. 12]. Studer et. al. defines as *"a formal, explicit specification of a shared conceptualization"* [230, p. 184], which is a definition well accepted by communities.

The usefulness of ontologies is vast, such as providing conceptual support for data architectures (such as data mesh, data lake structuring, and big data solutions), schematizing knowledge graphs by providing knowledge representation and facilitating human-computer interaction through well-founded conceptual models [75, 83]. Indeed, ontologies can translate mental models [3] into conceptual models [158] and within the conceptual modeling community, they are a key support for AI and ML, prominent research branches [240]. Thus, a whole branch of study emerged for the ontology design and development process, which is *Ontology Engineering* [175].

Several methodologies and standards were proposed to control and manage the *Ontology Engineering Process*, clarifying the role of each stakeholder involved. Among them, *Ontology Engineers* and *Domain Specialists*, their roles and interactions are protagonists in this thesis. Briefly, ontology engineers must capture the domain notions provided by the domain experts, returning them with conceptualization solutions through well-founded ontological artifacts (e.g., documents, models, and implementations) to support managing their data, systems, and applications [167]. Indeed, these stakeholders' relationships, roles, and active participation are fundamental in guaranteeing that ontologies meet their technological objectives, being artifacts that may be reusable and interoperable.

[3] Mental models, considering the notion proposed in [51, 143].

1.1 Motivation

Ontologies as computational artifacts due to their characteristics, applications, and semantic competencies have been seen as a solution to reach data management and stewardship of computational assets towards the FAIR Principles [254]. Incidentally, all uses of ontologies must comply with the requirement of interoperating and reusing conceptualizations and data. This commonality requires that the semantics used and represented as ontologies be clear and consensual. Indeed, a domain conceptualization should use concise, complete, and unambiguous language to achieve its purposes [93]. Moreover, this demand must remain throughout the life cycle of the systems that use them [168]. Ontology engineers use philosophical principles to provide the semantic basis necessary for these requirements to be met by the ontologies produced, usually defining a *conceptualization grounding*.

As one of the ontologies uses, the conceptualization grounding comes from the notion that the meaning of each concept in a conceptualization is constrained formally to provide a better conceptual approximation in describing a domain in reality. For this to be possible, conceptualizations must have an *Ontological Level* [76, 77]. The ontological level reflects a specific *Ontological Commitment* regarding a particular axiomatization choice (in a language of representation). Indeed, within the ontology engineering process, the language choice can facilitate or hinder the representation (and implementation) of concepts, their relationships, and properties depending on the *Ontological Commitment* adopted [79, 102, 93].

Languages are essentially made by symbols (graphical or textual) that express certain knowledge (or *mental moment* [4]). The way these symbols are combined defines the syntax of such language. Modeling languages usually use graphical representations to represent models and require defining rules and primitives that compose their abstract syntax. As expected, ontological commitment is the key to making it feasible for languages to clearly express the desired meaning to provide an intelligible conceptualization through their constructs. Furthermore, ontologies are artifacts that, besides using modeling languages as a representation tool, serve to drive languages, constraining them. Thus, through the cognitive process, languages are constrained by ontologies, whether explicitly and formally (ontology-driven languages) or implicitly. However, stakeholders tend to have divergent interpretations of the conceptualization, either due to limitations of the representation or due to a lack of knowledge of the ontological commitment adopted.

[4] A mental moment in the philosophical sense, same adopted in [101].

Regardless of whether the commitment is explicit and formal, each mental moment represented by modeling language constructs intends to commit to a specific notion of reality. In other words, *"For an information structure to represent a conceptualization, it must commit to the existence of the entities constituting that conceptualization"* [99, p. 182]. Indeed, in this work, the author discusses the critical role of ontologies as promoters of FAIR Principles, specifically exploring the notions of *I*nteroperability and *R*euse. Thus, achieving FAIRness for data is an important use of ontologies, likewise attaining FAIRness for ontologies. In other words, ontologies are a pathway to guarantee FAIR data; hence, ontology engineers must guarantee these ontologies are FAIRness conceptualizations. Therefore, ensuring FAIRness of ontologies is required to guarantee that data from these ontologies can be successfully interoperated and reused. Moreover, this research area needs plenty of study and investigation, motivating us to bring together the stakeholders' views and guarantee their views follow FAIR Principles, i.e., through FAIRness ontologies.

1.2 Problem Statement

Despite all efforts in the ontology engineering community to provide methods, guidelines, metrics, standards, and maturity models for the ontology engineering process, ontology engineers and domain experts still develop unclear ontologies due to their unbalanced views and biases [165, 195]. The problem is that stakeholders take different conceptual perspectives, and this causes ontologies to have biases. Furthermore, the cognitive process enveloping semantic agreement depends on the ontological commitment [76, 77, 84] stakeholders adopt. Thus, their viewpoints introduce misinterpretations even about the same concept. Indeed, the way domain experts and ontology engineers seek to achieve consensus lacks a more robust semantic bond. As a result, ontology engineers provide well-grounded reference ontologies, which are rarely implemented. At the same time, domain experts produce operational ontologies storing large amounts of valid data but with naive ontological support or even without any [169, p. 106].

Metrics and criteria for evaluating the quality of produced ontologies are also focused only on the explicit characteristics of ontological artifacts, even though these artifacts carry their meta-characteristics plus those originating from the domain they represent. The work [146] identifies that their metadata is not treated adequately, in addition to the large number of ontologies produced. This is evidence that these ontologies do not follow the FAIR Principles,

i.e., there is a lack of FAIRness ontologies despite the amount produced. Metadata is commonly used to store information that implicitly characterizes computational artifacts, being able to deal with characteristics of ontological artifacts such as methodology of development, domain of application, version, granularity, application, and formalization, among others. Some ontology characteristics can be seen as intrinsic [5] properties (version, name, language, etc.) while others arise because of design decisions during the ontology engineering process; therefore, they demonstrate a relational character and/or some dependence. Those characteristics can be seen as meta-characteristics because relational aspects and dependence are at a higher ontological level [63]. For instance, the ontology formalization depends on the language used because languages impose limits on representation (or implementation).

In the same direction as our research, the work [173] discusses and describes the pros and cons of the ontology evaluation classes. The authors identify the challenges ontology engineers have when selecting ontologies to reuse and integrate due to the number of ontologies created to deal with a body of knowledge. The work explores the lack of homogeneous criteria to characterize ontologies and proposes an ontology classification considering an evaluation terminology. They also consider that these ontologies usually must deal with complex data, demonstrating this is a challenging problem, mainly because the knowledge representation carries its intrinsic issues.

Stakeholders' disagreements and challenges of the ontology engineering process are greater the larger the knowledge domain to be represented [167]. Complex domains greatly expand the cloud of concepts in the context of representation and the possible interpretations regarding those concepts. Domains that produce large volumes of data continuously, which are in constant development, in which the processed information requires a high level of detail, security, or risk (including life and death), are subject to producing catastrophic outcomes caused by problems of semantic interpretation. Cybersecurity, Human Genome, Medicine, Medical Science, and Government Intelligence and Defense are examples of complex domains.

[5] Intrinsic aspects in the sense of UFO [88]

1.3 Objectives and Research Questions

In this thesis, *Cybersecurity* is our case of study domain, and we are presenting throughout this document our proposal to tackle these presented problem statements. It is a domain that encompasses the most common and difficult problems that perfuse ontologies as a tool for semantic clarification. The cybersecurity domain is vast, complex, meaningful, and it is frequently used in practice. The number of published proposals for ontologies covering this domain and surrounding domains is large in terms of state-of-the-art and state-of-practice. Moreover, ontology-related data on cybersecurity is available and accessible. Additionally, in this type of research context is not feasible deal with more than one domain choice due to the enormous amount of knowledge and resources required for a cross-domain analysis. Therefore, we chose to focus on cybersecurity, as we know that the most important requirements are present.

We formulate some research questions based on the problem raised and the already identified context and stakeholders. We are answering these research questions throughout this dissertation to achieve the objectives defined in our investigation. We propose three main goals:

1st To clarify and homogenize the necessary meta-ontological requirements, data, and characteristics to help stakeholders achieve awareness and common sense about conceptualizations (ontologies).

2nd To provide a clear baseline for characterizing and comparing ontologies to facilitate the interoperability and reusability of ontologies through their *Ontological Analysis*.

3rd To validate the contributions of this research.

To address the proposed goals, we propose one knowledge question (KG) and some research questions (RQs) as part of the methodology [252] that we adopt in this work. We are answering these questions throughout this work.

KQ1 How to conceptually characterize ontologies?

RQ1 Which (meta)characteristics help it conceptually characterize ontologies?

RQ2 How do managing these (meta)characteristics aid ontological analysis to provide FAIRness?

RQ3 How should a framework be, to provide a precise and reproducible basis capable of favoring semantic agreement, and meeting stakeholders' perspectives?

RQ4 Do the stakeholders believe that this framework is useful to facilitate a semantic agreement between them?

1.4 Research Methodology

In our research, we apply the *Design Science Methodology* [252], which is defined as *"the design and investigation of artifacts in context"*. In this regard, a DSM *Design Cycle* has to target a *Design Problem* by providing some artifact to meet requirements. Moreover, stakeholders involved in the problem must be identified, and they are the information providers for elicitation of these requirements. Our final goal in this research is to provide a solution (the artifact) capable of facilitating ontologies' creation, management, analysis, and integration. Figure 1.1 shows this research *Engineering Cycle* we use according to the DSM.

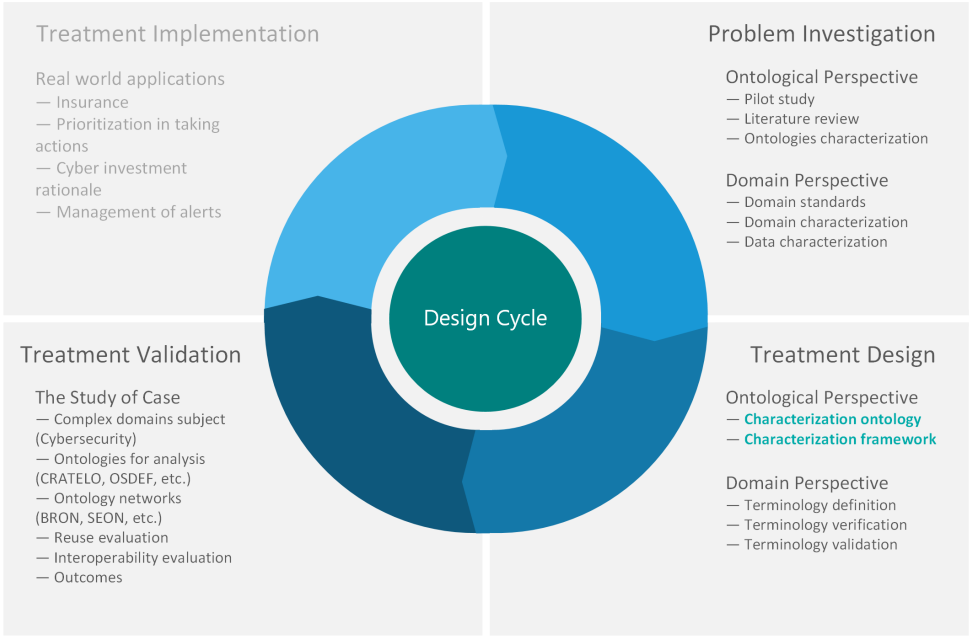


Figure 1.1: Engineering cycle of our research.

The three first steps of the Engineering Cycle compose the Design Cycle, and the last step is to validate the solution using it in real-world scenarios. The Design Cycle presented in Figure 1.1 has the above steps:

Problem Investigation: The first step to achieving our final target is to know state-of-the-art. In our pilot study [165], we searched ontologies covering the Cybersecurity domain to find their implementations and what are their technical approaches. Then, we search for ontology characterizations in general proposals to determine how they deal with ontology meta-characteristics, especially concerning FAIR Principles [254]. In parallel, we investigate the semantic support elements (standards, norms, regulations, etc.) used within the cybersecurity domain to understand the domain specialists' viewpoint and context of work. This helped us identify domain specialists' perspectives and possible biases in domain characterization. Finally, we have thoroughly reviewed the pilot study data and identified areas where we can align ontological and domain perspectives to ensure balanced perspectives on the conceptualization. This approach will help minimize any potential biases among stakeholders.

Treatment Design: Then, we focus on the treatment of the data we obtained in the first step, including defining a clear, traceable way to run ontological analysis while guaranteeing FAIRness for the data and their ontological model. We targeted mainly complex domains because the expense of ontological analysis grows proportionally as the domain gets vaster and more complex, as the cybersecurity domain. In doing so, we propose the **Ontology for Ontological Analysis** (O4OA) [169] and the **Framework for Ontologies Characterization** (F4OC) [167, 168]. In parallel, we propose a terminological verification and validation process of the cybersecurity domain definitions whose main outcome is forming a cloud of concepts about this domain safeguarded by domain meta-characteristics, i.e., with FAIRness.

Treatment Validation: We validate the reference version of the O4OA proposal through processes of model instantiation to explore possible issues or unexpected possibilities scenarios (branches or worlds). From the O4OA established model, we generate two implemented versions of it; the first called *gO4OA*, and the second is called *MongoO4OA*. We used the cybersecurity domain data and ontologies obtained in our state-of-the-art research to fill the O4OA knowledge bases. We used these data to produce ontological analysis results as study cases in this research.

Completing the *Engineering Cycle*, there is the Treatment Implementation step; however, it is important to point out that it is out of this thesis scope.

Treatment Implementation: We intend to use our approach by applying it in real-world scenarios, with the support of our industrial partners in future works.

1.5 Outline

This thesis is structured in 5 parts, comprising 8 chapters. Each part regards the phases described by the research methodology:

Problem Investigation Chapter 2 introduces the problem investigation. In this chapter, we analyze the characteristics and meta-characteristics of ontologies, the background that permeates them, and the roles and participation of stakeholders in the ontology engineering process. Chapter 3 introduces the Cybersecurity domain as our application context, describing stakeholders' challenges in producing findable, accessible, reusable, and interoperable ontologies with clear and reliable semantics. Chapter 4 presents a brief history of existing works focused on classifying or characterizing ontologies as tools to guarantee FAIR data that also must be FAIR themselves. Finally, we present an analysis of the meta-characteristics present in the domain to be represented (conceptualizations) and in ontologies as artifacts for representation, providing FAIRness to both. We structure the presented problems according to the stakeholders' perspectives and FAIR Principles.

Treatment Design Chapters 5 and 6 present the treatment design. In Chapter 5, we present the first part of the proposed solution to the problems already described. We address this problem of lack of consensual conceptualization by proposing a reference conceptual model (O4OA) that considers ontological and domain-related perspectives, knowledge, and commitment necessary to facilitate the process of Ontological Analysis, including the analysis of ontologies composing ontology networks. We propose the O4OA, a well-grounded meta-ontology supported by well-known ontological classification standards, guides, and FAIR principles. In Chapter 6, we present the second part of the solution proposed. Using the O4OA conceptualization as ontological support, we developed a framework for characterizing ontologies that provides a stable and homogeneous environment to facilitate ontological analysis by simultaneously dealing with ontological and domain stakeholders'

perspectives. The F4OA demonstrates the potential to facilitate identifying and looking at ontology metadata beyond the greater specific notions of them. The holistic view approximates with precision the metadata of the ontological angle (accessibility, availability, sharing, factors of modeling and implementation, and many others.) with the ones of the domain (cloud of concepts, area structuring, granularity, and so forth.), also present in ontologies and their relations in ontology networks.

Treatment Validation Chapter 7 focuses on the validation of the treatment. We present the O4OA implementations and the API solution we propose to manage the Cybersecurity domain real data we raise with the F4OA execution. We demonstrate the O4OA and the F4OA capabilities in favor of ontological analysis by showing the cybersecurity characterization cases we worked on. We also present the results from the ontological analysis made over proposed ontologies (and ontology networks) for the Cybersecurity domain. Additionally, we compare concepts used in these ontologies, clarifying their semantic misunderstandings. These outcomes are reviewed with the aid of a collection of professionals in Cybersecurity supported via a crew of Ontology Engineers for you to permit us to prove the validity of the proposal and the success in achieving our goal.

This thesis ends with conclusions that summarize the primary contributions of this work to the Conceptual Modeling and Cybersecurity communities, in addition to a dialogue of future lines of research (Chapter 8). Additionally, we provide complementary data through the appendices.

Part II

PROBLEM INVESTIGATION

Chapter 2

From Conceptualizations to Ontologies

Auribus teneo lupum, Terêncio (185 – 159 a.C.).

Taking the well-accepted definition that an ontology is ***"a formal, explicit specification of a shared conceptualization"*** [230, p. 184], we must deal with ontologies being computational artifacts specifically made to express knowledge according to certain commitments. As such, we must separate intrinsic characteristics of ontologies as computational artifacts from those of the conceptualizations themselves. The ontology engineering community has been discussing the multidisciplinary aspect of knowledge expression through computational and ontological artifacts. Indeed, this matter has already been better clarified through the notions of *Ontological Level* [76, 77] and language appropriateness [93, 102, 98]. From these studies, we extract relevant characteristics and meta-characteristics that make ontologies subject to misinterpretations.

2.1 Ontology Engineers and Domain Specialists

Ontology engineers are essentially knowledge modelers qualified to deal with the computational aspects that permeate the design and development of ontologies. As modelers, they consider the models produced (ontologies) understandable because they have already incorporated a series of restrictions and conditions previously introduced into their cognitive process. In other words, in the concern to producing the best possible representation of specific knowledge (Reference Ontologies), modelers abstract from computational aspects; meanwhile, they need to give up conceptual details in favor of the implementation (Operational Ontologies) because they need to be concerned about computational aspects at this moment [7]. However, design decisions can be lost throughout the life cycle of a computational artifact because they are sometimes implicit, or they are already part of the computational solutions, or even for economic or political reasons. This is particularly significant in the context of ontologies, as their primary objective is to elucidate, safeguard, and disseminate knowledge. Furthermore, this means that experts in the represented domain (designers or even users) may have a partial perception of the meaning contained (semantics), or worse, they may believe they know but ignore that they are unaware of it.

In essence, *"ontology engineers must capture the domain notions provided by the domain specialists, and give back to them with conceptualization solutions through well-founded ontological artifacts (e.g., documents, models, and implementations) to support managing their data"* [168, p. 106]. Figure 2.1 shows the `i*` [52] diagram expressing the distinct ontology engineers' and domain specialists' viewpoints.

Figure 2.1 shows a general view of the relation (roles and goals) between ontology engineers and domain specialists as stakeholders in the ontology engineering process [1]. Stakeholders of the ontology engineering process must have clear, homogeneous, and unambiguous conceptualizations in which they work. However, they have different concerns and viewpoints, i.e. different perspectives on the conceptualization. The goal of ontology engineers is to produce models to achieve the best possible approximation of a real-world domain and then implement those models. At the same time, domain specialists want conceptualizations to produce practical results given a certain set of requirements [168].

[1] It is important to note that the figure does not detail the ontology engineering process as a whole but illustrates the stakeholders' roles and their more general objectives.

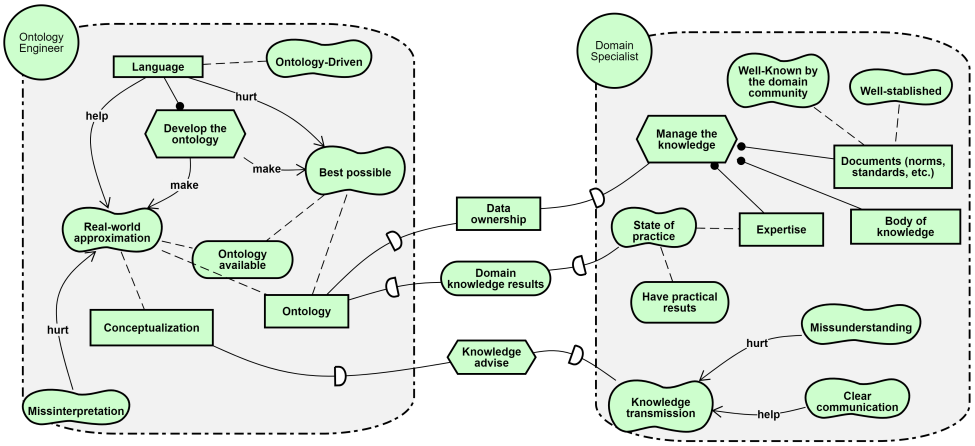


Figure 2.1: Ontology Engineers and Domain Specialists viewpoints `i*` diagram.

2.2 Ontologies Characterization Challenges

In the ontology engineering process, stakeholders must consider various technical aspects. Similar to software development, design decisions must be made when developing ontologies [7]. The implementation platform, the volume of data and its sources, and the modeling or implementation language influence these design decisions, whereby axiomatization aspects are often put aside in favor of the ability to draw logical conclusions. Therefore, we study and analyze these aspects in depth in ontological terms. Besides, it is fair enough that we can use ontological principles to analyze ontologies.

2.2.1 Ontological Level and Ontological Commitment

The notion of *Ontological Commitment* is defined as *"a mapping between a language and something which can be called an ontology"* [79, p. 560]. The authors advocate that the ontological commitment must be formal to give ontology engineers the linguistic tools necessary to model. Besides, models must be capable of reflecting the desired reality explicitly. Then, the authors focus on ontological commitment formalization, demonstrating that modal logic (in its neutrality) can express ontological constraints, presenting notions such as identity, counting conditions, and rigidity. Their approach abides by the meta-level categories that ontologies must distinguish, defining what is

called *Ontological Level* [76, 77]. The ontological level makes stakeholders take a formal ontological account of some representation.

Following the same philosophical line, the work [93] clarifies the relations between a thing in reality, its conceptualization, and a symbolic representation, and establishes the difference among the relations between *Conceptualization*, *Abstraction*, *Modeling Language*, and *Model*, as depicted in Figure 2.2.

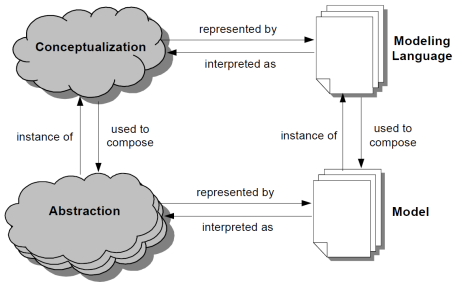


Figure 2.2: Relations between Conceptualization, Abstraction, Modeling Language, and Model according to [93, p. 21].

Ontologies such as DOLCE [27] and UFO [96] are based on these philosophical primitives. These ontologies are concerned with establishing criteria to be met by languages, which must be done through constraints at the level of their metamodels [93]. Non-compliance with a formal ontological commitment implies language whose specifications are more permissive. This permissiveness does not guarantee that the ontological principles (*Soundness*, *Completeness*, *Lucidity*, and *Laconicity*) [102] are fulfilled.

2.2.2 Ontology Identity and Rigidity

The *"Identity Principle is the feature that makes it possible for one to count and distinguish individuals"* [210, p. 251]. Counting enlightens the differentiation between individuals within a whole [89]. For example, among a group of programmers, Ada Lovelace is one of those. Besides, she was the same Ada girl before becoming the first programmer. Becoming a programmer did not change what is essential about Ada, i.e., what makes Ada to be Ada even after her death. *Rigidity* is the notion that explains how Ada's identity remains over time. Likewise Ada, ontologies as artifacts also have identity and rigidity.

Computational artifacts (software, operating systems, ontologies, etc.) are countable and have identity. Indeed, the concept of *Artifact* can be defined

as *"a collection of things created by agents where an Agent-Generic is a being that has desires or intentions and the ability to act on those desires or intentions"*[28, p. 282]. In this sense, should a software system in all its versions be counted as an individual, or should we count the various implementations of each, considering that the first version differs greatly from the last? These versions of the same system can perform the same function and keep historical records. Therefore, it is reasonable to think that they are the same software, just as Ada has always been Ada – from child to programmer, and so on. However, a newer version cannot run on a machine with an old CPU, nor can the older version be fully functional (discarding emulators) on a machine with current CPU generation `i#` because they are very different artifacts. Thus, unlike people who keep their essence no matter when or their characteristics (her essence is not apart from her humanity), software systems require a separation between their essence and artifacts. In other words, computational artifacts can change so much over time that these changes can mischaracterize the artifact. This can occur in such a way that they cannot be recognized as the same artifact anymore, only maintaining the same name for commercial reasons. Therefore, while we can say Ada is and was always Ada no matter what, we can recognize Windows as the same during its life cycle, but Windows 11 is a completely different computational artifact from Windows 3.1, for instance.

Likewise, in our example, an ontology (artifact) can change so much over time, including its properties, such as purpose, granularity, language used, etc. For instance, the Pizza Reference Ontology is a different artifact of the Pizza Operational Ontology. Indeed, several operational versions (implementations) may exist created from the pizza knowledge provided by the Pizza Reference Ontology. For example, one of those Pizza Operational Ontologies may be represented in gUFO [6]; then, we can think about two distinct artifacts, the canonical gUFO Pizza one and the fulfilled gUFO Pizza. Then, we can count several data-fulfilled gUFO Pizza implementations, each with its own different data. Thus, ontologies as artifacts have identity and rigidity, but we need to work with the multiple levels of instantiation and the context involved. This discussion regarding the identity of computational artifacts is relevant to the cybersecurity domain because computational artifacts are significant in this matter and are treated as assets. Moreover, we deal with ontologies and their characteristics and meta-characteristics in this work; therefore, we should treat ontologies as artifacts.

2.2.3 Ontology Structure, Limitations, and Formalization

Considering that ontologies as artifacts have identity and rigidity, we can work with the structural characteristics present in ontologies that make them able to represent some real-world part of a domain. Uschold and Gruninger [244, p. 10] show that the graphical representation interferes with the ontologies' level of formalization. In this sense, the authors are not dealing with ontologies in *particular* [2], but dealing with similar structural characteristics that certain *types of ontologies* [3] present. For instance, some ontologies present a structure that makes them vocabularies; others present a structure that makes them *Description Logics* [11] ontologies, and so on.

Structures of types of ontology determine the possible formalization range that ontologies (instances of this type) can reach. Thus, whether an ontology is at the minimum or maximum of this formalization range, it is always limited by what its structure can provide at a formalization level. Therefore, we identify that *types of ontologies* present the following structural characteristics:

Graphy: Symbolic characteristics present in certain types of ontologies. The works [156, 66] offer some common types of this characteristic, such as terms, ordinary glossaries, user classifications, web directories, data dictionaries, thesauri, structured glossaries, informal taxonomies, DB schemas, XML schemas, formal taxonomies, frames, formal lightweight ontologies, and models.

Limitations: Structural elements present in certain types of ontologies that impose limits on the ontologies represented. The works [244, 156, 66] show the degree of formalization that ontologies (instances of a certain type) can reach.

Appropriateness: the relational characteristics that types of ontologies have because of their Graphy and Limitations, making certain types of ontologies more appropriate than others to represent (or implement) certain knowledge. The works [102] clarify this notion.

Therefore, the notions of Graphy, Limitations, and Appropriateness are characteristics of types of ontologies that interfere with how ontologies (instances of types of ontologies) can be formalized. In other words, these are elements of characterization of ontology types that denote the meta-characteristics present in ontologies (instances of ontology types).

[2] Particular ontologies regarding the philosophical notion of *Individuals*
[3] Types of ontologies regarding the philosophical notion of *Universals*

2.2.4 Ontology Representation, Behavior, and Semantic Patterns or Anti-Patterns

Ontologies are artifacts made using some language; indeed, it is through the language used those ontologies express knowledge. According to [218, p. 10–16], a *"language is a convention that arises from the human capacity to construct a communication channel, i.e., a system of distinct signs (concrete things) corresponding to distinct ideas"*, and those *"signs are associations bearing the stamp of collective approval and which added together constitute language – are realities that have their seat in the brain"*. Meanwhile, ontologies are artifacts with different applications; some aim to facilitate human knowledge about a real-world domain, while others are made to be processable (implementations) [93]. This characteristic of a particular ontology being processable depends on the (representation or implementation) language used. High-axiomatization produces ontologies with processing limitations, while low-axiomatization makes it difficult to represent the desired semantic [66]. Therefore, ontology engineers usually consider the ontology's purpose and application (human or computational use) as justification in favor of a certain (representation or implementation) language instead of another, i.e., this is a design decision.

There are several languages used in conceptual modeling to produce ontologies. However, *"the first and foremost problem that can appear associated to a modeling language is its lack of adequacy for the specific application domain that is to be modeled"* [207, p. 271]. Characteristics of *lucidity*, *soundness*, *laconicity*, and *completeness* determine the degree of *isomorphism* of a language concerning some domain to be represented [102]. Thus, while appropriateness regards how much structural meta-characteristics interfere in ontologies, *isomorphism* regards how much structural meta-characteristics interfere in languages.

Languages use symbolism (a set of signs) to represent knowledge by assuming an ontological commitment [75, 79]. The symbolism used characterizes if some language is graphical or textual. Besides, it characterizes if an ontology uses one or other language (through its specification). For instance, an OWL representation is supported by the OWL language specification and must follow the rules engraved in its metamodel. In this set of rules, it is usual for design patterns [4] and design anti-patterns [5] to be established. Note that the OWL language specification is the thing that makes a representation presents

[4] A pattern is an abstraction from a concrete form that keeps recurring in specific, non-arbitrary contexts [213].

[5] An anti-pattern is a recurrent error-prone modeling decision [152]

the characteristics that make it an OWL, making this in different levels of abstraction; therefore, an OWL representation is a meta-characteristic present in ontologies (instances of ontology types). Likewise, languages also use the behavior of symbolism to represent knowledge, depicting Semantic Patterns and Anti-patterns for types of ontologies [102]. Thus, we identify that types of ontologies present the following additional characteristics:

Representation: Symbolic characteristics present in certain types of ontologies that make them an artifact for representation or implementation. They are used to represent a domain description or to produce an ontology schema.

Behavior: Behavioral characteristics present in certain types of ontologies make them artifacts to express/describe behaviors. They are used to represent a domain description behavior or to produce behavior in an ontology schema.

Semantic Patterns: Patterns are structures that produce recurrent and syntactically valid conceptual models. When an ontology is built, the (representation or implementation) language used induces the stakeholders to construct conceptualizations via the combination of existing ontologically motivated semantic patterns. These patterns constitute modeling primitives of a higher granularity when compared to usual language primitives. Besides, these higher-granularity modeling elements can only be combined in a restricted set of ways [103].

Semantic Anti-patterns: The anti-patterns, namely, model structures that, albeit producing syntactically valid conceptual models, are prone to result in unintended domain representations, i.e., Semantic Anti-Patterns. They are configurations that, when used in a model, will typically cause the set of valid (possible) instances of that model to differ from the set of instances representing the intended state of affairs in that domain [103].

2.2.5 Ontology Properties

Ontology engineering considers properties that ontologies have and need to be managed; most of those are associated with design decisions. For instance, competence questions are the pathway to define the ontology scope and provide its evaluation capabilities, complying with the stakeholder's expectations and requirements [73, 74]; the ontology purpose, which is important to determine the ontology level of generality [75] among other details; the version which must be managed, among others. Ontology engineering methodologies such

as *SABiO* [7] and *Methontology* [61] usually deal with this kind of aspect. Additionally, properties such as the ones related to the ontology and their data *Findability* and *Availability* are covered by the notions present in the FAIR Principles [254]. Well-controlled ontology catalogs [14], ontologies such as [22], data models such as [203], and proposals such as [50, 211] are good initiatives in this direction.

2.2.6 Contextualization and Commitment

Agreement and clear communication among stakeholders is fundamental during the ontology engineering process and must be maintained throughout the entire life cycle of the ontologies. Additionally, the defined and adopted ontological commitment in the engineering process must be traceable, reproducible, and available. Therefore, we searched within state-of-the-art ontology engineering to find the characteristics and meta-characteristics to characterize an ontology concerning the ontology engineers' perspective. We found vast literature within the context of ontology classification, such as the works [245, 230, 114, 58, 175, 147, 223, 156, 69] and FAIR Principles [254, 140, 155, 24, 99, 253, 177, 195], which we discuss in detail in Chapter 4. Thus, on one side, ontologies can be classified according to using many levels of abstraction, for instance, according to the degree of formalization and/or axiomatization of ontologies, their applicability, generality, structure, and development, among others. On the other hand, ontology engineers must pursue FAIR ontologies before using these ontologies to promote FAIR data.

We establish the requirements that must be accomplished to guarantee FAIRness for ontologies and help stakeholders achieve awareness and common sense about conceptualizations [168, 169]. This is in line with the notions of Ontological Commitment [79, 102, 93] and Ontological Level [76, 77], which are the keys to ensuring ontologies are correctly contextualized and have precise semantics. Therefore, we defined three pillars of support for this work, they are:

Classification: Classifies ontologies according to well-known classifications (level of applicability, level of generality, level of formalization, and level of axiomatization) and established within ontology engineering [165, 167].

Characterization: Characterizes ontologies by establishing relationships between ontological (meta)characteristics such as their language, representation, purpose, accessibility, copyright, reuse, and implementation, among others [168].

Discrimination: Provides a cloud of concepts that goes beyond the concepts adopted inside an ontology because it brings to light different standardizations and policies concerning the domain specialists' perspective. Moreover, this enlightens possible related ontologies, opening the domains' boxes and their ontologies, in which conceptualizations compound ontology networks [165].

2.3 Conclusions

Based on the main theoretical and philosophical bases that support the use and applications of ontologies to represent conceptualizations, we elicit the characteristics and meta-characteristics necessary for these artifacts to fulfill their main purpose of semantic clarification. This topic has been studied within ontology engineering, conceptual modeling, philosophical principles that involve ontology, and methodologies. However, the multidisciplinary characteristics that encompass this matter interfere with the viewpoints and biases of stakeholders. Although these aspects of ontologies are issues that have already been studied, they are often hidden in the concerns of ontology engineers and not apparent to domain specialists. Indeed, these are issues that usually add to other problems related to the representation of vast, complex, or sensitive domains, such as the cybersecurity domain, which also potentializes these problems. It is also important to note that in this work, we do not go into issues related to the politics and economics of the ontology engineering process.

Chapter 3

Towards Cybersecurity Ontologies

Dubitum sapientiae inittium, Descartes (1596 – 1650 d.C.).

Cybersecurity contributes to the primary protection of confidentiality, integrity, and availability of records through norms such as ISO/IEC 27032 [122], NIST 800-37 [144], among others. These norms relate to actions that deal with data protection, application safety, and network security. However, cybersecurity requires constant evolution, adoption of new technologies, and bringing substantial concerns to organizations, mainly regarding *Enterprise Architecture*. Besides, even the best technological aspects are no longer enough because the weakest link in the chain to guarantee assets' security is, in the end, a human issue. These range from the behavior of customers and attackers to how every stakeholder inside this environment participates and is aware of the principles and relationships they may be inserted into.

The sorts of problems organizations need to address, and the complexity of the cybersecurity domain induce misinterpretations and misunderstandings about the concepts. Indeed, these issues arise when it's necessary to ensure effective conversation among stakeholders, between human beings and systems, or promote systems reuse and interoperability [149]. Moreover, every stakeholder usually handles the facts according to their perception depending on the function he/she performs within the project, which interferes with the strategies followed. For example, notions such as *"Risk"* or *"Vulnerability"* for an enterprise architecture may be controversial, being hidden under presumptions of awareness that stakeholders have about these concepts. A supervisor can have a general-grained conception of these terms, while a cybersecurity engineer can think about the same concepts but with a specific belief. Both stakeholders assume they are speaking about the same idea, but this is not the truth because they have biases influenced by their roles. Indeed, usually, each one has strong argumentation that endorses their biases, most supported by well-established standards.

Throughout this chapter, we present the consequences of this problem and the challenges that lie behind it. To this end, we identify the domain specialists and their roles and perspectives. Then we surveyed and studied state-of-the-art cybersecurity ontologies with a view to the terminology adopted and their sources of support. Meanwhile, we also observed how these elements interfere with the characteristics that these ontologies present.

3.1 Identifying the Cybersecurity Experts Roles

Experts in the cybersecurity domain need to master an extremely vast body of knowledge because the domain is itself complex, and is coupled with several other domains, such as software engineering, requirements, analysis and social engineering, psychology, and human behavior, among others. This multidisciplinary characteristic makes these professionals subject to multiple perspectives and biases. Stakeholders that participate in these teams must present multiple abilities; so much so that companies usually identify teams by a color pattern, so that some teams antagonize others. Table 3.1 most frequent cybersecurity teams.

Table 3.1: Cybersecurity teams.

	Team	Role
●	Blue	The defenders, who are specialists in defensive tactics and techniques.
●	Red	The attackers, who are specialists in multiple kinds of offensive actions and strategies.
●	Yellow	The developers, who must project and codify security software.
●	Green	The mediators, who make the bridge between Blue and Yellow teams.
●	Purple	The renovator, who work on maximizing the performance by instigating competition between the Blue and Red teams.
●	Orange	The facilitators, who make the bridges between Red and Yellow teams.
○	White	The supervisors, who manage and solve any disputes among antagonist teams.

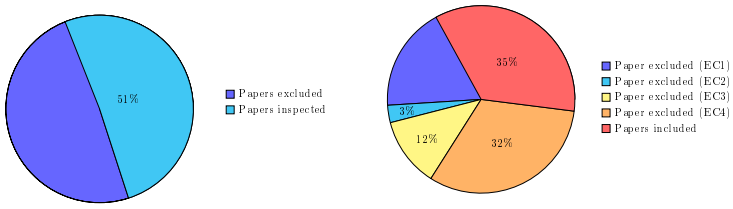
Although the Blue and Red teams are increasingly present in organizations, the rest of the palette often appears diluted among other organizational roles. For example, developers are not always aware that they are part of the Yellow team, the same as supervisors as part of the White team. Also, mediators, renovator, and facilitators can see themselves as members of one of the Red or Blue teams, depending on their personal profile and knowledge. However, among this palette of functions and roles, no one brings together a broader vision capable of filtering and transmitting knowledge, a *Black Team* for instance. Specialists to make a bridge among the palette teams with cybersecurity outsider stakeholders such as ontology engineers. Throughout this project, we received advice from specialists with this profile, which was extremely important.

Complex domains tend to be composed of subdomains. Consequently, there exist specialists who solely focus on these subdomains and those who possess a broader scope of expertise or who mediate relations among others. However, the roles of these stakeholders are not always treated as systematically as in the case of cybersecurity; even in this domain, we notice the absence of a Black team. This is a factor that interferes with the ontology engineering process.

3.2 Searching for Cybersecurity Ontologies

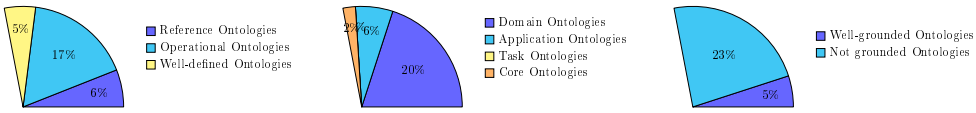
To identify proposals in the cross-field of cybersecurity and Ontologies, we conducted a Targeted Literature Review (TLR). This approach only keeps the significant references to maximize rigorousness while minimizing selection bias. We searched for *Cybersecurity Ontologies* and its related surroundings [165, 167]. We used the outcomes of this TLR to support our knowledge about this domain and as part of the solution we propose. The goal is to know the state-of-the-art ontologies covering the cybersecurity domain, whether they provide implementations and their technical approaches. Figure 3.1 summarizes the TLR results [1].

Figure 3.1: TLR selection process results.



In fifty-one papers included, we found thirty-five presenting ontology proposals: six works present only reference ontologies (not implemented), and only one is well-grounded. In comparison, twenty-seven works present operational ontologies (five papers present implementations supported by a reference ontology). Since some works refer to the same ontology; for instance, the works [237, 236, 238, 235] present the Ontology of Cybersecurity Operational Information which (a reference ontology) but only the works [237, 238] present an implementation (an operational version) suggesting that the implementation is partial concerning its reference model. Figure 3.2 shows the ontology classification statistics.

Figure 3.2: TLR ontology classification statistics.



[1] For more details about the TLR and absolute values, see Appendix A

3.3 Cybersecurity Conceptualization Challenges

Based on these outcomes, the most sizeable information we extracted from the TLR is the lack of foundational grounding within the cybersecurity ontologies. Only four papers point out an ontological foundation, and all of them are related to the CRATELO [197, 200, 198, 18] proposal. The ontologies found fail in their ontological foundation because they do not directly base their concepts on foundational ontologies nor use ontology-driven languages for their representations. The importance of a conceptual basis is noticeable because the aid of a foundational ontology avoids semantic interoperability issues on domain ontologies [87], even when the representation language is the bearer of this support [249]. Indeed, this is a common problem in other domains, as is evidenced in the works [239, 195, 194, 62].

It is also important to note that there is a relation between the lack of grounding and methodological problems. Indeed, the proposal that is based on a foundational ontology, CRATELO[193, 192], is exactly the one most dedicated to following an ontology engineering methodology, in this case, the *Methontology* [61]. The Common Ontology of Value and Risk (COoVR) [216], is another work that is well-grounded and uses a well-known methodology, SABiO [7]; however, it has not yet been implemented. Indeed, disagreements among stakeholders about the conceptualization can produce serious consequences because they may think they have good communication and agreement without actually having it. Moreover, in complex and sensible domains, prejudice can be immeasurable.

Besides the lack of grounding we detect, most papers bringing up operational ontologies had been implemented without earlier reference ontology. In contrast, most of the proposals of reference ontologies aren't applied, and no justification was supplied. Indeed, only OVM [251], CoCoa [201], and the Ontology of Cybersecurity Operational Information [237, 238] proposals offer an operational ontology supported via a previous reference ontology. The affirmation that operational ontologies require the support of a prior reference ontology is well-established in [93].

Another important issue is that obtaining further details about the implementations is difficult. For example, it is arduous for users external to the ontology project to determine which concepts present in the reference ontologies were implemented nor to track the losses of expressiveness that may occur because of the implementation choices. Moreover, these issues question the reliability of information available through these ontologies because reliability depends on the semantics, context, and ontological commitment

adopted. Therein lies the importance of taking care of the meta-characteristics of ontologies, which is why the FAIR Principles apply to ontologies [173].

The terminology used in Cybersecurity Ontologies is raised in several standards, norms, glossaries, recommendations, and other guides that support the cybersecurity domain [2]. These sources usually present information that is consistent with each other; however, their applicability context may create misinterpretations. Usually, *Ontological Analysis* [75] has been used to address problems like misinterpretations, misunderstanding, and disambiguation, including contextual and cognitive issues [167, 168]. However, the ontological analysis process in complex domains is not trivial because the volume of information to be studied and analyzed is enormous. Besides, ontology engineers must deal with the design process and the applicable standards. Thus, stakeholders take different conceptual perspectives, and this causes ontologies to have biases that are sometimes more ontological and sometimes more domain in nature.

Ontological analysis also helps stakeholders evaluate whether or not an ontology can be reused or how the ontology is interoperable in their projects. Indeed, ontologies built to describe vast or complex domains should not be overly large or be used in isolation. Therefore, the referred *Ontology Networks* (ON) arose, in which ontologies protecting subdomains of complex domains are co-related or interrelated [231]. Initiatives such BRON [116, 115] and OdTM [32] have been bringing a pragmatic view of the use of ON with the adoption of analytics from different cybersecurity data sources, and covering subdomains such as vulnerabilities [161] and weaknesses [164]. CVE (for the vulnerability concepts) and CWE (for the weakness concepts) are well-established sources contemplated in the TLR.

Thus, complex domains impose aggravating factors on the ontology engineering process. We summarize these factors:

1. *Methodological and ontology engineering process issues end up resulting in ontologies that are not well-grounded or not well-defined.*

2. *Stakeholders must deal with the natural (or technical) language in these sources, leaving room for more diverse interpretations.*

[2] Appendix B depicts the terminology and the sources that define the terms raised in the TLR

4. *Even well-established standards may provide conflicting definitions for the same term, depending on their use, goal, and applicability.*

5. *Complex domains tend to have sources with different granularity, which are created to cover their subdomains, being inevitable source overlaps or overload the resulting based-in conceptualizations.*

6. *Stakeholders tend to present difficulties in defining the correct granularity that ontologies require because the number of sources and the vocabulary is vast.*

7. *Complex domains require interoperable and reusable ontologies, making ontological analysis fundamental.*

8. *Usually, the stakeholders' roles and interests make ontologies reflect their biases, especially for intricate domains.*

3.4 Conclusions

From the presented in this chapter, we depict the problems faced by stakeholders involved in the ontology project. Although our work context is the cybersecurity domain, these problems are repeated in other domains that present greater complexity, are equally vast, or involve sensitive areas. Thus, knowing, controlling, and managing ontologies meta-characteristics in these scenarios is a key tool under the solution space. However, it is important to mention that there are limitations in making these meta-characteristics into useful, processable, and accessible metadata of ontologies.

Metadata can help by presenting ontology classification data and intrinsic properties (version, copyrights, representation or implementation languages, URL, among others). Still, it is unfeasible to aggregate in the ontology metadata covering aspects of the cloud of concepts that the conceptualization encompasses. It is generally at the discretion of ontology engineers to produce documentation that specifies the terminology, sources, and adopted definitions. However, the ontology documentation tends to be quite technical and hard to understand, primarily because constraints are mainly represented formally. Besides, the documentation may not even be available.

Chapter 4

State of The Art

Carpe viam et susceptum perfice munus,
Publius Virgilius (70 – 19 a.C.).

Due to the challenges that complex knowledge domains present when they need to be represented (modeled), several initiatives have emerged to address this issue. These initiatives emerged both within the Conceptual Modeling community with a model-agnostic viewpoint and from domain-specific proposals from their respective domain experts. Nonetheless, they are not convergent because agnostic proposals are too comprehensive to deal with the particularities of the domains. Specific proposals have high coupling, which increases the complexity of the data too much, depreciating the already complex domain. Thus, we seek state-of-the-art proposals to identify initiatives with potential for convergence, low coupling, and scalability. These proposals have similarities and differences compared to our proposal, which we will also show below.

4.1 FAIR Principles

The FAIR Principles proposed in [254] arise to clarify data management and stewardship, providing a set of best-practice indicators to allow these processes to be effective. The biotechnology community was among the first to benefit from the FAIR principles, as it deals with a knowledge domain that stores and distributes sensitive information through a vast quantity and variety of data sources [222]. This community has realized how beneficial it is to use

FAIR principles in managing their scientific data. They have acknowledged that humans cannot efficiently handle large amounts of data, while machines have limitations in processing information semantically. Works such as [188, 153, 214] are among these initial initiatives. The cybersecurity community has also recently adopted the FAIR Principles because they have dealt in a domain with sensitive and complex information stored in scale data [221].

The acronym `FAIR` stands for *F*indable, *A*ccessible, *I*nteroperable, and *R*eusable. These notions are described in [177, 24] as *"`F`"* states that *"datasets should be described, identified and registered or indexed clearly and unequivocally"*; *"`A`"* affirms that *"datasets should be accessible through a clearly defined access procedure, ideally using automated means – metadata should always remain accessible"*; *"`I`"* prescribes that *"data and metadata are conceptualized, expressed and structured using common, published standards"*; and *"`R`"* states that *"characteristics of data and their provenance are described in detail according to domain-relevant community standards, with clear and accessible conditions for use"*. Figure 4.1 shows the FAIR Principles, detailing the sub-principles proposed in [254].

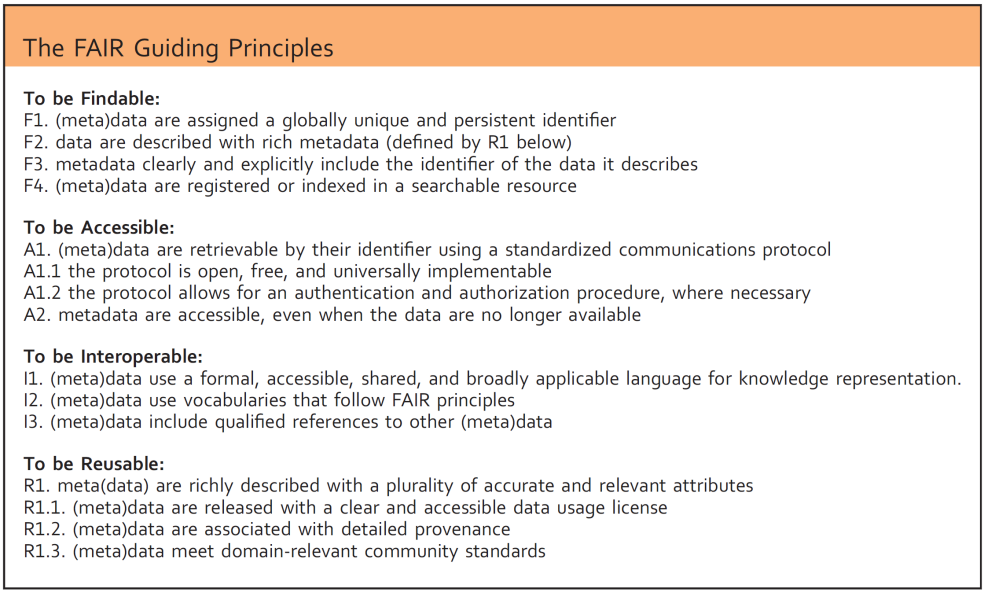


Figure 4.1: FAIR Principles detail [254, p. 4].

According to Wilkinson in [253, p. 1] *"the FAIR Principles are aspirational in that they do not strictly define how to achieve a state of FAIRness"*. Still, rather they describe a continuum of features, attributes, and behaviors that should move a digital resource closer to that goal. The GO FAIR [1] initiative defines that FAIRness in the context of the FAIR principles refers to the degree to which digital assets, data, and other research outputs adhere to the principles of being findable, accessible, interoperable, and reusable. GO FAIR proposes an ontology [2] already agreeing that ontologies act towards FAIRness. But are ontologies themselves FAIR?

In one direction, the community domain specialists search for answers focusing on domain-specific perspectives. Several communities have been persecuting archive FAIR Principles for their assets (models and data). For instance, the work [8] deals with metric and maturity indicators to provide an assessment workflow for data FAIRness in the life sciences. Domains such as the Internet of Things (IoT) and Web of Things (WoT) have initiatives under the *Ontology-Driven Interoperability* (ODI) matter, such as [180] that discuss interoperability concerning the ontology development process in IoT, and [13] that goes in line with the notion of FAIRness by adhering semantic principles in IoT ontologies reuse. Several ontologies conceptualize and regulate IoT and WoT domains; for instance, SSN [10], oneM2M [205], and SAREF [53] are ontologies proposed and adopted as W3C [3] standards. IoT and WoT are complex domains under which our work is applicable; however, these initiatives differ from ours, and several are the motives:

First, these are domain-specific ontologies dealing with the core characteristics in the IoT/WoT domain. In contrast, our proposal deals with (meta)characteristics present in any kind of ontological artifact created to represent any domain.

Second, our proposal rationalizes the notion of FAIRness over ontological analysis processes, while such ontologies rationalize ODI into their domain.

[1] `https://www.go-fair.org/fair-principles/`
[2] `https://github.com/peta-pico/FAIR-nanopubs/tree/master`
[3] `https://www.w3.org/`

> *Third, although these initiatives are examples of the reuse of ontologies, they do not deal with the notion of a broad cloud of concepts (and their details) nor relations among ontologies in any network. Indeed, they are data interoperability providers for IoT/WoT, while our proposal is an interoperability provider for any ontologies.*
>
> *Fourth, IoT/WoT ontologies have the same issues we detected in the cybersecurity ontologies, detailed in [165]; notably, lack of a grounding, making them require adaptations to interoperate or have proper reuse, with no assuring semantic. In other words, these works lack of a foundational ontology support. The work [17] runs ODI by making an ontological analysis and goes in line with the notion of FAIRness (like our proposal) under the ODI viewpoint (ontological perspective).*
>
> *Lastly, there is no mention of important domain-dependent aspects, i.e., domain (meta)characteristics (domain perspective). Instead, our approach is domain-agnostic but not domain-indifferent since the purpose of performing an ontological analysis is to elicit knowledge in a consensual, reproducible, traceable, and formal way. Indeed, ODI is among many uses where ontological analysis is a key contributor.*

In another direction, the ontology engineering community searches for answers, focusing on an ontological perspective and taking different approaches. The European Commission expert group on FAIR data addresses what is needed to implement FAIR Principles [45]. Their report discusses the interdisciplinary aspects of frameworks for FAIR. The work [211] summarizes some works that pursue FAIRness towards best practices and guidelines. The work [146] presents a vast study of the metadata of ontologies. This study shows that some works were more relevant in terms of being available to describe ontological metadata, such as Dublin Core, Ontology Metadata Vocabulary, and VoID, among others. Besides, the study compared these works and their implementations. The work [99] discusses the ontological principles that ontologies must comply with to be FAIR; moreover, to be FAIR, ontologies must be well-grounded. In the same foundational direction, the work [195] provides a systematic mapping study looking for FAIRness aspects in ontologies for the domain of Security.

In addition to these initiatives, the FAIR Principles research community started to develop ontologies to deal with the FAIR Principles themselves, i.e., provide ontologies to address the FAIR Principles contained knowledge. The Terms4FAIRskills (T4FS) [176] provides a terminology encompassing the skills required to provide and maintain FAIR data. The T4FS ontology [4] is under development through a collaboration of several institutions [5]. T4FS is an operational ontology implemented in OWL under the OBO Foundry [224] bases.

Going in the same direction, the work [22] provides an OntoUML reference ontology to describe FAIR Principles, clarify its definitions, terminology, and provide proper ontological analysis of this domain. Additionally, it clarifies the FAIR implementation branches into FAIRness assessment, FAIR tooling, and FAIR service support. Such initiatives deal with this targeted FAIRness terminology, while our proposal deals with the terminologies of domains whose data and conceptualization require FAIRness. Indeed, this distinction highlights an important standpoint; as well as the ontologies classification process helps us to target essential characteristics of ontologies, the process of classifying FAIR Principles terminology helps us to target essential characteristics that the domain perspective must consider achieving FAIRness in their assets. In other words, ontologies classification and FAIR principles conceptualizations are blueprints for the (meta)characterization of FAIRness, respectively, considering ontological and domain perspectives. Therefore, the proposal presented in this thesis considers both.

The Ontology Metadata Vocabulary (OMV) [112] is a proposal for describing ontologies and related entities, being the most similar approach to ours that we could find in the state of the art. The proposal has demonstrated usefulness in initiatives such as [54]. The approach distinguishes between an ontology base (a conceptualization) and an ontology document (a realization of a conceptualization – an implementation). The ontology covers metadata that is part of the FAIRness discussion, such as language, licensing, and quantitative data (number of classes, properties, and axioms). OVM precedes FAIR Principles adoption; however, it already presents ontologies classification as a key elicitation process for ontologies characterization, focusing on metadata of ontologies and intending to be the standard covering this domain, such

[4] `https://github.com/terms4fairskills/FAIRterminology/`

[5] CODATA, ELIXIR-EMBL-EBI, ELIXIR-FR, ELIXIR-NL, ELIXIR-UK, EOSC-Life, FAIRsFAIR, FAIRsharing, GO-FAIR, the Digital Curation Centre, the Dutch Centre for Life Science, DANS, Royal Holloway, Leiden University Libraries, The British Library, Oxford University, European University Association, VU Amsterdam, SURF, European Bioinformatics Institute, Australian Research Data Commons and TU Delft. See `https://terms4fairskills.github.io/`

as Guarino's classification [75]. In this respect, OVM is similar but lighter than our proposal; however, as an ontology, OVM in itself is not FAIR. Besides, it does not have the support of a prior reference ontology; indeed, it is an ontology implemented in OWL without using any foundation ontology for grounding. Conversely, our proposal is grounded on UFO [96], has a well-defined reference model written in OntoUML [19], and is implemented; besides, it supports our framework proposal following a solid methodological approach, SABiO [7].

4.2 Ontology Classification

Ontology classification is a criterion for observing and analyzing (meta)characteristics of ontologies. Since the beginning of the ontology research area in computer science, and more specifically in the conceptual modeling community, researchers observed the need to look at the building blocks of an ontology as an artifact of conceptualization. These classification proposals are strictly sustained in philosophical and mathematical foundations, observing essential properties of artifacts when they encompass the principles presented in Section 2.2. In other words, these classifications have only the key purpose of producing ontological artifacts (models) capable of representing a real-world phenomenon in the best possible way. Indeed, this philosophical vision at the origin of ontologies brings them closer to the FAIR Principles. Therefore, these classification proposals aim to establish a comparison base among ontologies using their (meta)characteristics to distinguish them.

In its logical foundation's work, Husserl in [117] already considers the notion of *Formal ontologies* versus *Informal ontologies*, perceiving those informal ontologies as actually non-ontologies. This inaugurated the classification criteria according to the level of formalization of ontologies. Then and adopting the notion of ontological commitment, the work [245] classifies ontologies according to the formality of the language used in the representation, consisting of *Highly Informal Ontologies*, *Informally Structured Ontologies*, *Semi-formal Ontologies*, and *Rigorously Formal Ontologies*. This classification is consistent with the approach outlined by Guarino in [84]. However, classifying ontologies based on their formalization level is not the only important aspect of applying ontologies. The very definition of what an ontology is discloses the multidisciplinary aspect of ontologies. Likewise, the Ontology Engineering community deals with the multidisciplinary aspect of knowledge expression through computational, logical, and linguistic aspects of ontological artifacts. This is better clarified by Guarino in [76, 77] and

Guizzardi in [93, 102, 98]. Thus, different dimensions of ontology classification emerged from this perception that ontologies transcend one perspective.

The classification based on the level of generality of the ontology (sometimes called knowledge kind) refers to a level of dependence on a specific viewpoint. Mizoguchi and Ikeda [175] classify ontologies into *Workplace Ontologies*, *Domain Ontologies*, *Task Ontologies*, or *General/Common Ontologies*, looking specifically from the knowledge reuse viewpoint. The work [230] provides a classification according to the generality level based on the works [2, 113, 31], already considering the notion that *Core Ontologies* as the ones between the General and Domain Ontologies, not so general as the first either so specific as the latter [113, 247]. In the same direction, Valente et al. define *Core Ontologies* as *"a very general ontology of a certain application domain, e.g., medicine. This core ontology should contain several of generic concepts and method-independent definitions, characteristics that would give high reusability to the elements of this library"* [246, p. 33[1]]. These approaches result in the following classification criteria.

Van Heijst et al. in [247] propose a classification according to two dimensions. First, according to the amount and type of structure of the conceptualization. Second, according to the subject of the conceptualization, as depicted in Table 4.1:

Table 4.1: Classification criteria proposed in [247].

Structure of the conceptualization	Subject of the conceptualization
• Terminological ontologies. • Information ontologies. • Knowledge representation ontologies.	• Application ontologies. • Domain ontologies. • Generic ontologies. • Representational ontologies.

** Core Ontologies are ontologies between foundational and domain ontologies, not as general as the firsts nor as specific as the latter [113, 246].*

Studer et al. in [230] propose a classification according to the generality level:

- Generic ontologies (core ontologies [247] or super theories [29]);
- Domain ontologies;
- Task ontologies;
- Representation ontologies.

Also working in the dimension of the structure, the work [114] proposes to classify ontologies as *High-level Ontologies*, *Domain Ontologies*, or *Task Ontologies*. The work [58] proposes to classify ontologies according to the generality level as *Generic or Common-sense Ontologies*, *Representational Ontologies*, *Domain Ontologies*, or *Method and Task Ontologies*. However, the most accepted classification of ontologies is based on their dependence on a particular task or viewpoint is the proposal of Guarino in [75], which complements the proposal of Mizoguchi and Ikeda [175], represented in Figure 4.2.

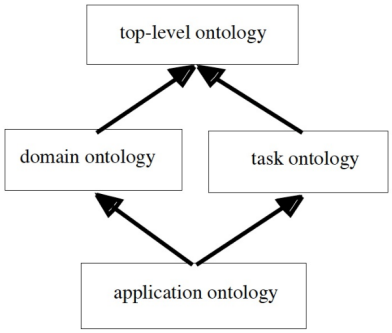


Figure 4.2: Graphical representation of the Guarino's proposal [75, p. 10].

In detail, these are the kinds of ontologies presented in Figure 4.2.

- *Top-level ontologies* describe very general concepts like space, time, matter, object, event, action, etc., which are independent of a particular problem or domain: it seems therefore reasonable, at least in theory, to have unified top-level ontologies for large communities of users.
- *Domain ontologies and task ontologies* describe, respectively, the vocabulary related to a generic domain (like medicine or automobiles) or a generic task or activity (like diagnosing or selling) by specializing the terms introduced in the top-level ontology.
- *Application ontologies* describe concepts depending on a particular domain and task, which are often specializations of both the related ontologies. These concepts often correspond to domain entities' roles while performing a certain activity, like replaceable unit or spare component [75, p. 10-11].

Additionally, Uschold and Gruninger in [244] propose to classify ontologies according to the knowledge kind as *Domain Ontologies*, *Task Ontologies*, or *Representation Ontologies*, already studying considering how the representation language used interference, as depicted in Figure 4.3.

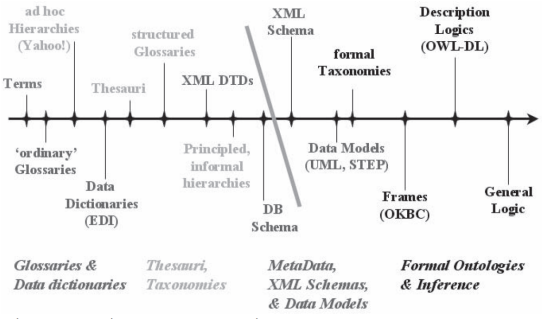


Figure 4.3: Graphical representation of the Uschold and Gruninger's proposal [244, p. 10].

Another important dimension of the ontologies classification concerns the purpose. In this dimension, authors may target real-world issues such as proposed in [147], classifying ontologies as *Static Ontologies*, *Dynamic Ontologies*, *Intentional Ontologies*, or *Social Ontologies*. Jasper et al. in [141] classify ontologies according to their application as *Neutral Authoring Ontologies*, *Ontology as Specification*, or *Common Access to Information Ontologies*. Regarding the classification based on the level of applicability, Guizzardi's classification allows us to differentiate when an ontology is an *"explicit and formal representation of a portion of reality for knowledge sharing"* or an *"implementation of this representation for knowledge computational management"* to differentiate a *Reference Ontology* from *Operational Ontology* [93, p. 24]. This work demonstrates greater acceptance by the community.

Regarding logical (meta)characteristics in conceptualizations, ontologies have other classification dimensions. Gómez-Pérez and Corcho propose that *"Lightweight and Heavyweight refer to two different kinds of ontologies: those ontologies where concepts (described by their attributes and are organized in taxonomies using only the subclass-of relationship), relations and functions, and possibly instances are the only components that are represented, and those ontologies that also contain axioms"* [68, p. 58]. Lassila and MacGuinness propose a bi-dimensional classification, considering the richness of internal structure and formalization [156]. At the same time, the bi-dimensional classification proposed by Gómez-Pérez and Corcho considers the richness of internal structure and the subject of conceptualization [70]. Oberle's tri-dimensional classification considers the purpose, expressiveness, and specificity of ontologies [191]. These approaches result in the following classification criteria.

According to the richness of internal structure and formalization [156] as depicted in Figure 4.4.

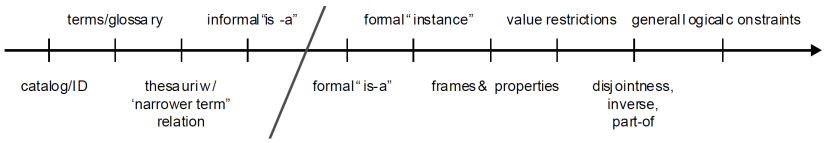


Figure 4.4: Graphical representation of the Lassila and MacGuinness' proposal [156, p. 907].

Table 4.2 depicts the kinds of ontologies presented in Figure 4.4.

Table 4.2: Classification criteria proposed in [156].

Internal structure	Formalization
• Controlled vocabularies. • Terms/Glossaries. • Thesauruses (narrower term, relation). • Informal hierarchies.	• Formal hierarchies. • Formal instance. • Frames (properties). • Value Restriction. • General Logical constraints. • Inverse Disjoint (Part-of).

According to the richness of internal structure and the subject of conceptualization [70] as depicted in Figure 4.5.

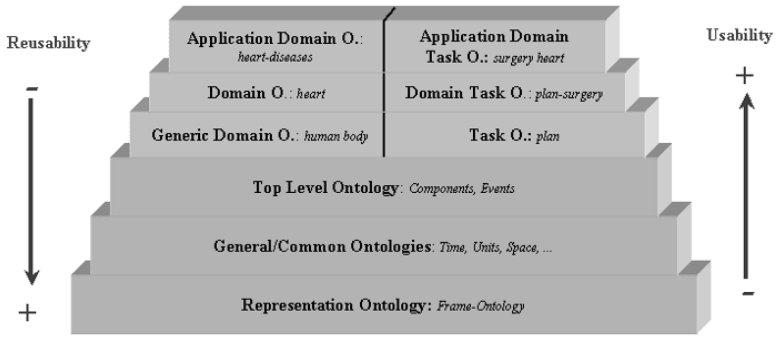


Figure 4.5: Graphical representation of the Gómez-Pérez and Corcho's proposal [70, p. 35].

Table 4.3 depicts the kinds of ontologies presented in Figure 4.5.

Table 4.3: Classification criteria proposed in [70].

Richness of internal structure	Subject of conceptualization
• Controlled vocabularies. • Glossaries. • Thesauruses. • Informal hierarchies. • Formal hierarchies. • Frames. • Ontologies with value constraints. • Ontologies with generic logical constraints.	• Knowledge representation ontologies. • Common or generic ontologies. • High-level ontologies. • Domain ontologies. • Task ontologies. • Domain task ontologies. • Method ontologies. • Application ontologies.

The tri-dimensional (purpose, expressiveness, and specificity) classification [191]as depicted in Figure 4.6.

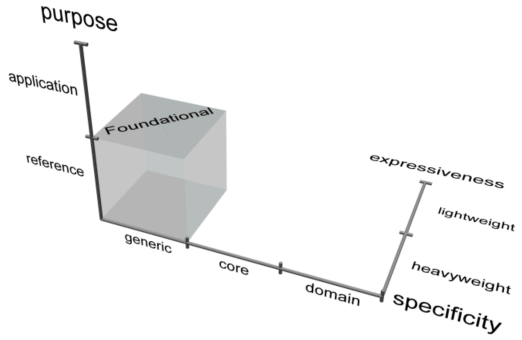


Figure 4.6: Graphical representation of the Oberle's proposal [191, p. 49].

Table 4.4 depicts the kinds of ontologies presented in Figure 4.6.

Table 4.4: Classification criteria proposed in [191].

Purpose	Expressiveness	Specificity
• Application Ontology. • Reference Ontology.	• Heavyweight Ontology. • Lightweight Ontology.	• Generic Ontology. • Core Ontology. • Domain Ontology.

Taking the works [244, 68, 156] as a reference, Giunchiglia and Zaihrayeu propose a multidimensional classification [66] as shown in Figure 4.7.

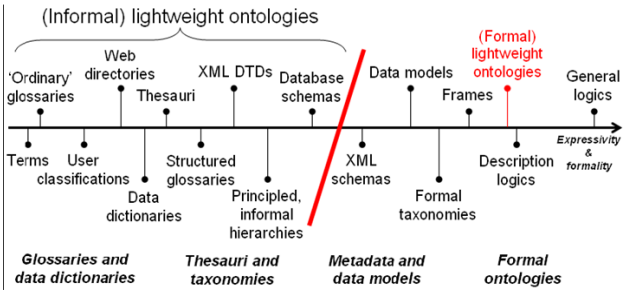


Figure 4.7: Graphical representation of the Giunchiglia and Zaihrayeu's proposal [66, p. 10].

4.3 Conclusions

Several other ontological characteristics became valuable elements to characterize them from these initial classification and characterization efforts. The work [173] presents state-of-the-art research on the evaluation criteria of ontologies. The authors group the ontologies' characterization efforts into five main lines: *domain/task fit*, *error checking*, *libraries*, *metrics*, and *modularization*. Already the *OntoUML/UFO Catalog* [6] initiative [14] provides an open solution for ontology metadata availability. All those efforts target ontology as artifacts, pointing to the ontology engineers' eyes and how they perceive the conceptualizations. There is high value in this auto-evaluation process, indeed, the O4OA adopts the auto-evaluation metrics summarized in this work [7] since these metrics are directly or indirectly encompassed in the classification criteria we use.

Even the efforts under the *domain/task fit* group are limited to the ontology engineers' viewpoint. This happens because these works are generally tested with ontologies that have little use (sometimes not even implemented) by experts in the conceptualized domain. The most used ontologies usually have

[6] `https://github.com/OntoUML/ontouml-models`.

[7] Adaptability, Alignment, Clarity, Cohesion, Completeness, Conciseness, Correctness, Coupling, Craftsmanship, Deployability, Domain Fit, Essence, Expandability, Expressiveness, Extensibility, Fidelity, Fitness, Intelligibility, Interoperability, Linked Data, Vocabularies , Pragmatics , Pruning, Reusability, Richness, Rigidity, Semantics, Semantic Interoperability, Semantic Web, Semiotics, Sensitiveness, Social Quality, Syntactic, Task Fit, and Upper Ontology [173].

a strong agreement among domain experts; however, they are naive concerning ontology engineering best practices. In our proposal, we postulate that this divergence arises from the difficulty in identifying the characteristics and the meta-characteristics of the conceptualization, both from the perspectives of ontology engineers and domain experts. Therefore, the FAIR Principles add value in identifying which characteristics are meta and which are not when we evaluate ontologies as artifacts that express domains of knowledge in their best possible expression. However, the ontologies that conceptualize the FAIR Principles domain do not achieve FAIRness and present the same problems presented in works [165, 195].

The work [146] demonstrates the lack of grounding as an issue, which confirms conclusions presented in [195, 167]. In this work, some works were more relevant in terms of being available to describe ontological metadata, such as Dublin Core, Ontology Metadata Vocabulary, VoID, etc. The study compared these works and their implementations, demonstrating the lack of foundational grounding. All in all, these works are, among others, examples of how domain ontologies lack grounding, problems in using development methodologies, erroneous design decisions, implementation issues, and difficulties holding on to data FAIRness. Consequently, to achieve FAIRness, it is not enough to use ontologies. It is mainly necessary that these ontologies be well-founded, well-defined, and FAIRness due to their characteristics and meta-characteristics, both concerning the representation of reality and through the data contained therein.

Part III

TREATMENT DESIGN

Chapter 5

The Meta-Ontology to describe Ontologies

Animus hominis quicquid sibi imperat obtinet,
Plauto (254 – 159 a.C.).

The **Ontology for Ontological Analysis** (**O4OA**) describes the characteristics and meta-characteristics of an ontology as an artifact that expresses a conceptualization can have. The goal (purpose) of O4OA is to clarify and homogenize the necessary (meta)ontological requirements, data, and characteristics to help stakeholders (*Ontology Engineers* and *Domain Specialists*) achieve awareness and common sense about conceptualizations (ontologies). O4OA models the foundational and domain-related concepts and the relations necessary to facilitate the process of *Ontological Analysis*.

In this context, we deal with the ontological perspective supported by well-known classification approaches and FAIR Principles by considering ontologies as artifacts expressed through some language. Besides, we also deal with the domain perspective, which deals with a semantic consensual agreement about the terms and their definitions (concepts) present in ontologies and their sources of information (norms, standards, etc.). O4OA intends to correct misalignment and miscommunication between stakeholders, establishing a clear, reproducible, and homogeneous standpoint to conceptually characterize ontologies, including addressing their possible relationships in networks. These are requirements that make ontologies, and their data comply with FAIR Principles.

5.1 Methodology, Stakeholders and Research Questions

Ontology engineering best practices strongly recommend the adoption of a known methodology to guide the ontology design; therefore, we adopt the SABiO methodology [7]. Accordingly, we first identify the ontology stakeholders and their roles, as presented in Part II. Then, we define the purpose of the ontology. In this case, the goal of O4OA is to clarify and homogenize the necessary (meta)ontological requirements, data, and characteristics to help stakeholders achieve awareness and common sense about conceptualizations (ontologies). Next, we elicit the characteristics and meta-characteristics of ontologies as artifacts through best-practices ontology engineering, philosophical grounding, and FAIR Principles available in the literature. Finally, we identify two important stakeholder responsibilities that our proposal must cover to make ontologies accomplish their purpose concerning the ontology development process: While *domain specialists* are concerned with identifying the relevant aspects of knowledge that are part of a conceptualization, *ontology engineers* aim to represent this ontology in such a way that it expresses this knowledge with real-world semantics that can be unambiguously interpreted, either by humans or by computational systems.

Still following SABiO, we propose the O4OA in a partnership project that brings together a research consortium to develop sound cybersecurity knowledge graphs (operational ontologies) through a comprehensive solution. The project involves teams from several academic institutions working with a private company. After many discussions among the project participant stakeholders, this team reached a consensual agreement. The team has included multidisciplinary participation, providing different contributions [1], composed of ontology domain specialists, UFO/OntoUML specialists, cybersecurity domain specialists, cybersecurity data researchers, literature review specialists, the project manager, the project advisor, and this thesis author. From the team discussions, we define the *Competence Questions* (CQs) that are the pathway to define the ontology scope and provide its evaluation capabilities, complying with the stakeholder's expectations and requirements [74, 73]. Readers may find the complete description of O4OA competence questions in Appendix C.

The CQs contemplate a cross-perspective of ontological and domain-related, extending them to consider the relationships among ontology when they need to interoperate in networks. Thus, from the defined scope,

[1] Part of the elicitation process happened during the COVID-2019 pandemic, so the remote strategy was mandatory.

purpose, commitment, and competence questions and knowing the involved stakeholders, we proceed with the engineering process according to SABiO. Due to the O4OA model characteristics (size and complexity) and design decisions, we adopt agile development. Figure 5.1 summarizes the development cycle we adopt.

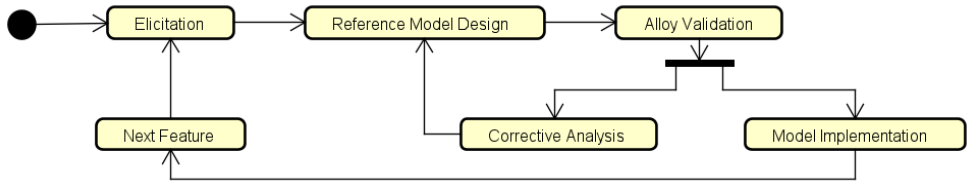


Figure 5.1: O4OA development cycle.

After the elicitation, we represent the O4OA reference model using the OntoUML [19] language, which provides grounding over UFO [96]. Next, we adopted the Alloy analyzer tool [139], applying the OntoUML notions present in the work [20] to proceed with the validation. The tool helps us run the instantiation of each model package in an individual and modular way. In this validation process, we elicit the additional constraints required (in addition to those already present in OntoUML), and we also check model cardinalities to ensure correct semantics. Thus, if we find any error, issue, or other problem, we proceed with a process of corrective analysis and restart the cycle. If we reach the stakeholders' agreement, we proceed with the feature implementation through the operational version of the ontology.

To answer the O4OA questions, we partition the ontology into three main packages. The first is the `Domain` package that covers the domain specialists' perspective. The second is `Linguistic`, which covers the representational aspects of conceptualizations. And, the third is `Ontological`, which covers the ontology engineering perspective. Each of these packages was divided into sub-packages to deal with the features of the ontology. In Appendix C, we present the complete package structure.

5.2 Conceptual Characterization of Ontologies

We present the ontology engineers' perspective by using four levels of abstraction to classify ontologies according to the ontology application, the level of generality, and the level of formalization and/or axiomatization of ontologies. This selection encompasses the most relevant and comprehensive classification criteria as the referential base for O4OA, being supported by the works [93, 75, 247, 66, 68]. These dimensions encompass a systematic ontology classification approach to guarantee the FAIRness of ontological artifacts.

The classification based on the *level of applicability* proposed in [93] allows stakeholders to differentiate when an ontology is an *"explicit and formal representation of a portion of reality for knowledge sharing"* or an *"implementation of this representation for knowledge computational management"*, i.e., and if it is a Reference or an Operational Ontology.

The classification based on the *level of generality* (sometimes called knowledge kind) of ontologies refers to dependence on a specific viewpoint. We opt for Guarino's [75] proposal since it is the most accepted by the community and complements the proposal of Mizoguchi and Ikeda [175]. This classification characterizes conceptualizations as *Foundational Ontologies* and, *Non-Foundational Ontologies* (i.e., *Domain Ontologies*, *Task Ontologies*, or *Application Ontologies*). We added in this classification another widely accepted proposal; the *Core Ontologies* [247].

Figure 5.2 shows the classification approach adopted in O4OA, in which we describe the classification levels using the <<*Subkind Pattern*>> [97], considering the aforementioned classification describes types of ontologies.

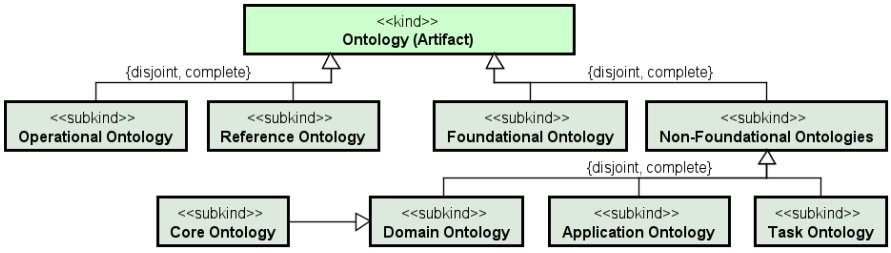


Figure 5.2: Fragment of the O4OA (meta)ontology – Classifications according to [93, 75, 247].

We also consider the classification based on the axiomatization level (and the language's limitations) provided by Gomés-Peréz and Corcho [68]. Thus, stakeholders can identify ontologies' computational limitations when a conceptualization becomes an implemented ontology (i.e., an operational ontology). This classification positions ontologies linearly according to the expressiveness of the language used. Already, the bi-dimensional classification [66], based on [244] and [68], provides a link between the axiomatization and formal levels, focusing on the approach and expressiveness of the language. These classification dimensions comprise the meta-characteristics of ontologies, i.e., they are aspects existentially dependent on types of ontologies (`Ontology Type`) appearing in their instances (`Ontologies`) as a consequence of the ontological commitment adopted.

Figure 5.3 presents the notion of ontologies as artifacts adopted in O4OA, in which we describe the classification levels considering the aforementioned meta-characteristics and dimensions. The notions of expressiveness, formalization, and axiomatization levels are categories of modes (<<*category*>>) of ontologies (instances of `Ontology Type`).

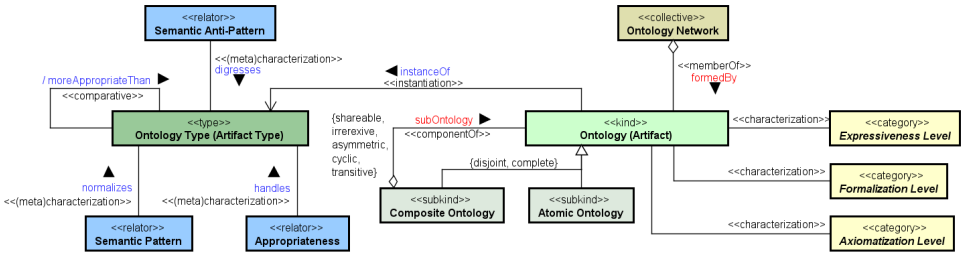


Figure 5.3: Fragment of the O4OA (meta)ontology – Types of Ontologies.

Note, that Figure 5.3 also depicts two custom characterization relations. The relation `normalizes` between `Ontology Type` and `Semantic Pattern`; then the relation `handles` between `Ontology Type` and `Appropriateness`. We define the stereotype <<*(meta)characterization*>> for these relations because we want to highlight their nature as meta-characteristics of ontologies (as instances of `Ontology Type`). Thus, these relations are typically characterizations (defined with the stereotype <<*characterization*>> in OntoUML) established at the type level and concerning the context of O4OA. We adopted this approach even though this stereotype is not present in OntoUML.

The notion of appropriateness [102] arises from the relationship between the language symbolism (`Graphy`) and its limitations (`Limitation`) in representing some domain using this language. `Graphy` and `Limitation` are modes part of (<<*commitment*>>) language specifications (more specifically, the concepts of `Language Specification` and `Ontology-Driven Language Specification`, presented in Section 5.4). Figure 5.4 presents the notion of appropriateness by using the *Relator Pattern* [97], considering the notions of *types of types* [63, 64].

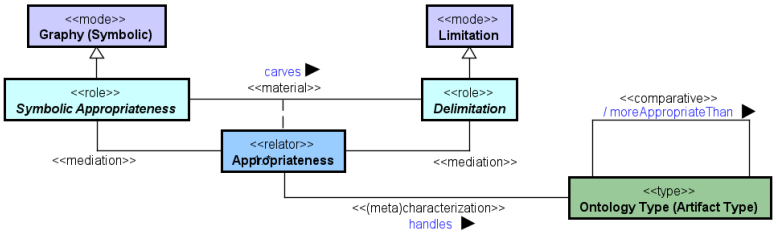


Figure 5.4: Fragment of the O4OA (meta)ontology – Appropriateness.

Language types are delimited (`Delimitation`) to be more suitable to model phenomena in a given domain when the limited structural characteristics of ontologies (`Limitation`) are present due to the symbolism (`Symbolic Appropriateness`) used. Therefore, `Appropriateness` characterizes language types (`Language Type`); meanwhile, `Limitation` and `Graphy` interfere with ontology characteristics (formalization level, expressiveness level, and axiomatization level). Similarly, the semantic patterns and anti-patterns [103] notions follow the same ontological foundation. Section 5.4 details how languages and ontologies are related and the relational aspects that rely on these notions.

5.3 Domain Cloud of Concepts in Conceptual Characterizations

We present the domain specialists' perspective through the concepts belonging to a conceptualization that must be represented and described. For this to be possible, we clarify the philosophical grounding of what encompasses a conceptualization and precisely distinguish what a concept is (as an abstraction) and what is the concept representation (as an artifact) [169, p. 111]. Thus, O4OA uses `Term` as a syntactical artifact (an *Object Kind*) used to describe the notion of a concept (as *Trope Kind*).

Sources provide relevant ontological information and are represented through the OntoUML *Rolemixin Pattern* [97]. Besides these notions, roles are relational-dependent [105]. Figure 5.5 presents the relation between concepts (`Concept`) and documentation sources (`Source`). Note that we use the *OntoUML notion of Part/Whole* through the relation `componentOf` to represent the definitions that compose each domain description.

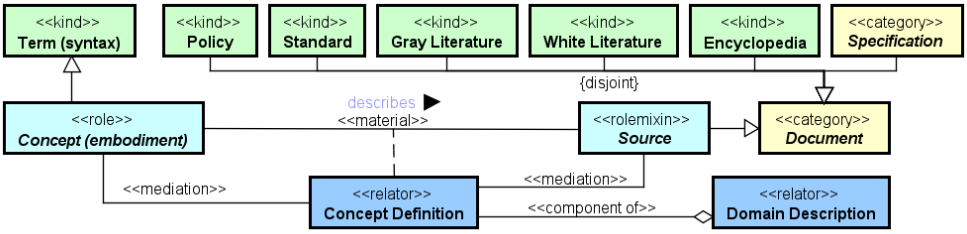


Figure 5.5: Fragment of the O4OA (meta)ontology – Domain definitions.

According to [93], a concept is an ability (or a *domain abstraction*), a *Trope*. It relies on the notion that agents possess moments (mental and cognitive) that regard the capacity of some properties of specific individuals to refer to possible situations of reality [101]. Already, a concept, as an embodiment (an *Object Kind*), refers to the symbolism these agents assign to a concept as an assignation of ability (a *Trope Kind*). In this view, concepts are the meanings [209] – (containing) modes of presentation, in Fregean's sense [256]. Note that a concept (*Trope* – ability) is externally dependent on a Source (*<<rolemixin>>*) to provide a `Concept Definition`; moreover, it has an existential dependence on a concept (*Object Kind* - embodiment). This aligns with the Relator Pattern that bears the notion of Relational Moments of UFO [64].

Thus, O4OA names a concept as a syntactic object that provides an embodiment because a definition is associated with a term to refer to this concept as an ability (an abstraction). Thus, a term takes on the role of embedding an abstract concept when it is used as a syntactic element in a definition (`Concept Definition`). In other words, a `Concept` is syntactic (non-abstract) and refers to the embodiment of an abstraction in an ontology (`Domain Description`). A `Domain Description` is a building block used to clarify grammatically (terminologically) the notion of a *Concept* (as *Trope Kind*) according to some source of information.

5.4 Linguistics in Conceptual Characterizations

Stakeholders' perspectives must be connected to reach a semantic agreement capable of sustaining an ontological commitment [93, 76]. From the classification of ontologies according to their applicability level, O4OA derives the notions of domain descriptions (reference ontologies) as the ontological support to ontology schemas (operational ontologies) [169, p. 113]. Indeed, there should not be any operational ontology without a previous reference ontology in which concepts and their relationships are *well-defined*. This affirmation promotes FAIRness for ontologies because it helps ontology engineers deal with implementation language limitations by knowing which ontological aspects can (or cannot) be implemented. Indeed, stakeholders can make better design decisions to guarantee that ontologies comply with the requirements. This approach facilitates identifying whether an ontology is a conceptual model or an implementation. Figure 5.6 presents reference and operational ontologies, their roles, and their relation.

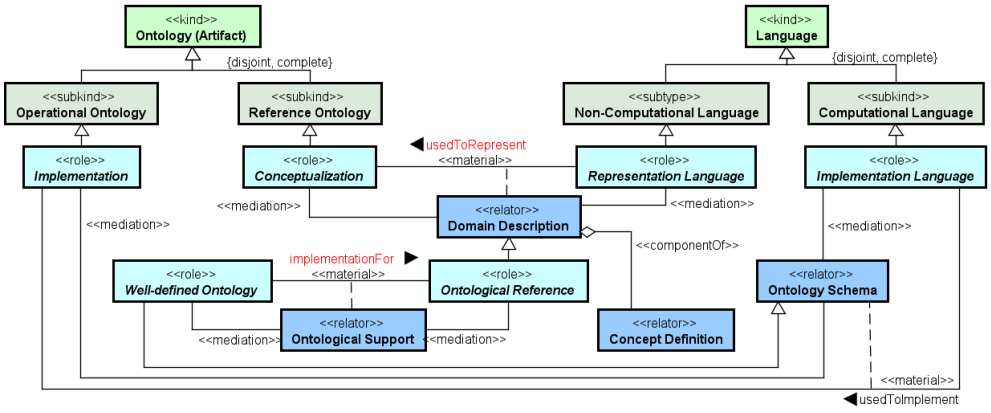


Figure 5.6: Fragment of the O4OA (meta)ontology – Applicability Level.

Figure 5.6 shows that `Conceptualization` is a reference ontology that is represented through a `Modeling Language`. Likewise, an `Implementation` is an operational ontology that works through an `Implementation Language`. We use the *Relator Pattern* [97] to represent these relational aspects of ontologies. According to its Applicability Level, they are `Domain Description` and the `Ontology Schema`. Thus, the notion that an ontology is or has an implemented version (`implementationFor`, a *<<material>>* relation) derives from the fact that reference ontologies provide ontological support

for ontological schemes. Indeed, `Ontology Schema` instances are the possible ontology implementations, such as they have relational dependence on the language used for their implementations. When a schema has the ontological support of a reference ontology, it is considered a well-defined ontology (<<*role*>>).

Incidentally, ontologies (domain descriptions or ontology schemas) are thought of in some languages specified by metamodels. Metamodels play the role of `Abstract Syntax` to specify languages playing the role of `Concrete Syntax`; this relation denotes the notion of what are language specifications. These roles, the external and non-descriptive associations of characterization connecting mode types `Graphy` and `Limitation` with the role `Abstract Language` plus the mode types `Representation` and `Behavior` with the role `Concrete Language` [64]. The package `Ontological` previously defined these aspects (see Section 5.2). This relation clarifies the external dependence and the commitment embedded in language specifications. Figure 5.7 describes these notions.

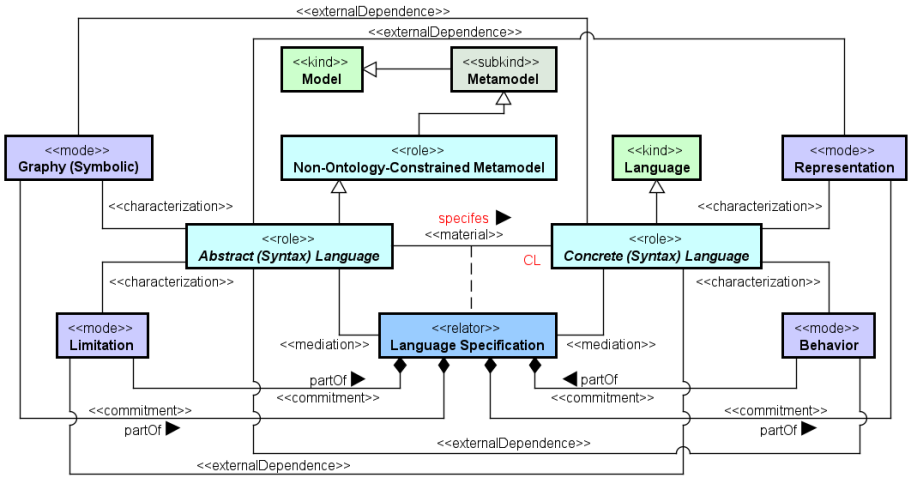


Figure 5.7: Fragment of the O4OA (meta)ontology – Language Specification

O4OA considers metamodels (`Metamodel`) as particular models (<<*subkind*>>) that can carry the available modeling primitives forms the lexical layer of some language. Besides, metamodels are made available to be used as a language specification by its mediation as abstract syntax (<<*role*>>). This approach is in line with the metamodel definition present in [93, 36]. Note that in [36] the authors consider that abstract

languages (`Abstract (Syntax) Language`) already contain all the syntactic rules and constraints **implicit** in the metamodel. Indeed, we extract these different relational aspects, making the relations explicit from the role played by metamodels, i.e., the `Abstract (syntax) Language` and `Abstract Ontology-Driven Language`. It only makes sense to deal with abstract language when these relationships are present since the metamodel is a model; in other words, a metamodel only differs (from models) because it has *"Meta"* potential, which occurs based on these relations. Figure 5.8 expands the abstract syntax notion through the metamodel relations.

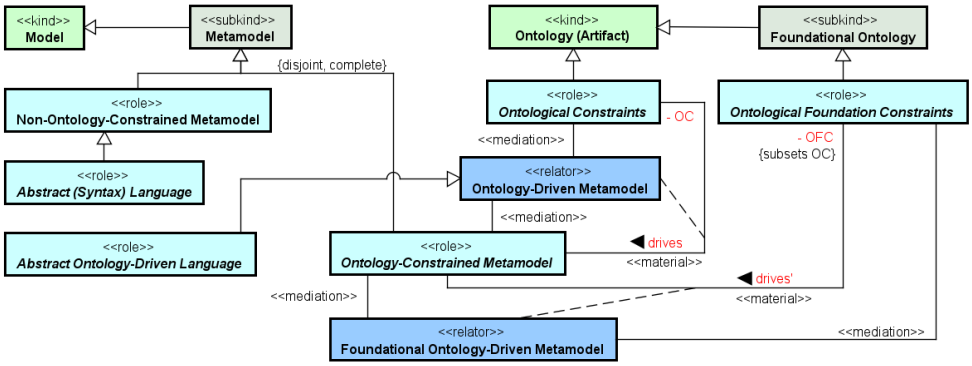


Figure 5.8: Fragment of the O4OA (meta)ontology – Abstract Language

Furthermore, modeling languages (typically ODML) play the role of representation languages in domain specifications so that ontology-driven metamodels specify ODML. Indeed, ontologies drive ODML and philosophically constrain their metamodels by enlightening the `specifys` (<<*material*>>) relation and define the notion of `Ontology-driven Language Specification` as shown in Figure 5.9.

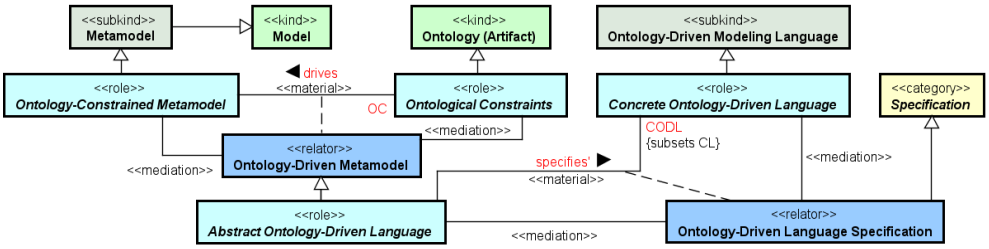


Figure 5.9: Fragment of the O4OA as a (meta)ontology – Ontology-Driven Modeling Languages.

An example of an ODML is the OASIS language, an *Ontology-Driven Domain Specific Language* driven by O^3 [208, 206], an ontology for object-oriented programming based on BWW [250]. Similarly, OntoUML is a *Foundational Ontology-Driven Language* based on UFO. In both cases, we are dealing with *Foundational Ontologies* that drive *Domain-Specific Languages*; i.e., these languages are the carriers of the `Ontological Foundation` for conceptualizations that are made through them. Thus, an indirect foundation is created by the `Ontology-driven Language Specification`, as shown in Figure 5.10. Note that the *association-end* (`CODL` subset of `CL`) constrains out the notion of `Concrete Ontology-Driven Language`.

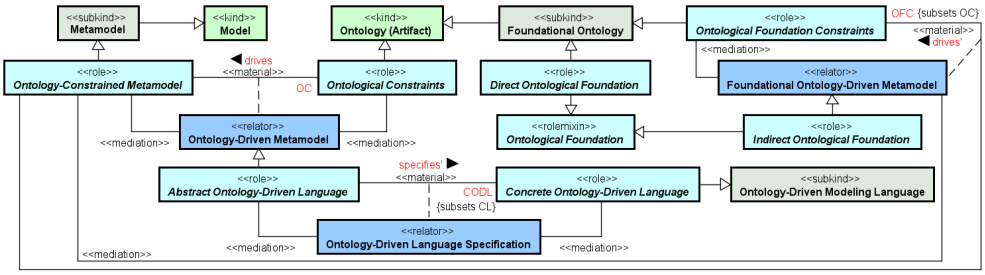


Figure 5.10: Fragment of the O4OA (meta)ontology – Ontologies driving languages.

Foundational ontologies avoid semantic interoperability problems in more specific ontologies [87] in two ways. In other words, directly or indirectly, they can produce *Well-grounded* ontologies. Directly, when more specific notions of a conceptualization can be inherited from the more general notions of foundational ontologies, a kind of approach is much used in the ontological analysis process. Indirectly, languages (ODML) are the carriers of the most general notions when these languages are used to produce domain, task, application, and core ontologies. Therefore, *"ontologies must be evaluated according to their grounding, separating ontologies that are driven by foundational ontologies (i.e., well-grounded) from ontologies without this support (i.e., not grounded)"* [169, p. 114].

5.5 Relations Among Ontologies and Ontology Networks

The notion of well-grounded ontologies makes it feasible for stakeholders to understand the importance of grounding when they have to interoperate concepts among ontologies with these characteristics and, at the same time, guarantee FAIRness. Jointly, the classification according to generality level helps them properly position ontologies, focusing on the relationship between a foundational ontology and its grounded ontologies. The O4OA covers grounding ontologies relationship by defining the <<*material*>> relation `groundedOver` established through the Foundational Ontologies roles. Thus, stakeholders can make solid semantic considerations about conceptualizations from more general conceptual (philosophical) notions of a foundational ontology. Figure 5.11 shows the (`groundedOver`) relation using the *Relator Pattern* to describe how Foundational Ontologies ground the non-foundational ontologies.

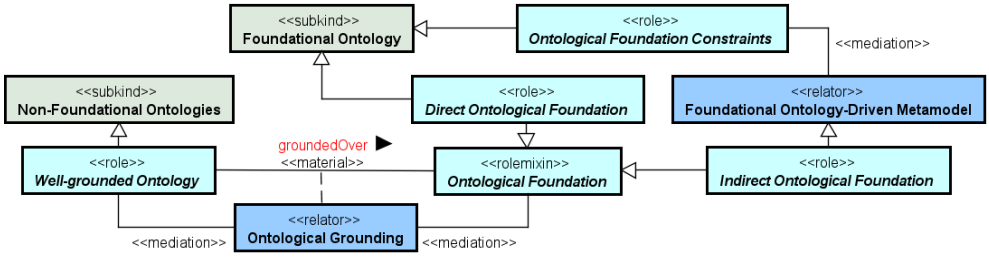


Figure 5.11: Fragment of the O4OA as a (meta)ontology – Well-grounded ontologies.

Still chasing the FAIRness for ontologies, O4OA characterizes them according to their generality level and facilitates their reuse, the *"R"* principle of FAIR. Ontologies can be reusable in several ways, depending on how the *Reuser Ontology* lays hold of and uses concepts of the *Reused Ontology*. The most usual kinds of reuse are:

- non-foundational ontologies reusing other non-foundational ontologies by specialization to attend requirements and provide stakeholders' agreement;
- non-foundational ontologies reusing other foundational ontologies by specializing general notions (thought a `groundingOver` relation);
- not-grounded ontologies gathering ontological grounding by incorporating concepts and notions from well-grounded ontologies (ontological analysis);

- by establishing relationships between concepts from different ontologies, i.e., adding relations;
- by establishing types of types abstractions, i.e., refining or thickening the level of abstraction (by using a multi-level theory such as [63]).

Figure 5.12 depicts the reuse of ontologies and how ontologies can be composed by other ontologies (sub-ontologies). The reuse of ontologies depicts an *Intersection* among ontologies. In contrast, *whole/part* follows the *Weak Supplementation Pattern*, which states that every whole must be composed of at least two parts [92, 100].

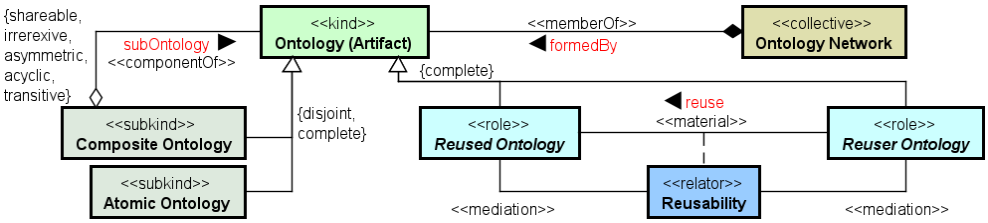


Figure 5.12: Fragment of the O4OA (meta)ontology – Reuse a Whole/Part.

The difference between the relations of reuse among ontologies and the notion that a larger ontology *"uses"* smaller ontologies. This means that, in essence, the reuse of ontologies denotes an *Intersection* among ontologies while sub-ontologies reveal a *Whole/Part* relationship. Stakeholders make strenuous efforts to recognize the role of each ontology within a network of ontologies where these relationships generally co-occur because these kinds of relationships become confusing the more complex the domain. For example, UFO is composed of UFO-A, UFO-B, and UFO-C sub-ontologies, and simultaneously, UFO-B and UFO-C reuse UFO-A. Indeed, in our search for state-of-the-art cybersecurity ontologies, we came across widely reused ontologies, such as CVE and CWE, even though these are cases of grounding lack. In none of those cases, these relations are clear.

5.6 Conclusions

We presented **the Ontology for Ontological Analysis**, a meta-ontology that classifies and characterizes ontologies from their characteristics and meta-characteristics. In developing this ontology, we had the support of a multidisciplinary team and followed the SABiO methodology to accomplish best practices in ontology engineering. We used the OntoUML language to represent the O4OA conceptualization to guarantee ontological grounding through the UFO. Besides, the proposal is supported by a series of well-established ontology classifications, principles, and best practices supported by FAIR Principles. Thus, O4OA, besides answering the proposed CQs (see Table C.4), also complies with the requirements imposed therein, being FAIRness. For readers who want to delve deeper or know the details, visit the ontology repository [2].

O4OA establishes a clear and systematic process for characterizing ontologies using the most recognized works and with the best coverage of philosophical principles regarding ontological analysis and ontology engineering process. Although the proposal is domain-agnostic, it is not domain-indifferent. It is a reference model to study, manage, and maintain ontologies that describe real-world complex, highly regulated, and data-sensible domains since, in these cases, semantic problems tend to be aggravated. Moreover, O4OA establishes a joint, stable, and systematic environment for improving communication among ontology engineers and domain experts, avoiding misinterpretations, misunderstandings, structural issues in ontologies, and communication problems that interfere with FAIRness. Indeed, the proposal can allow semi-automated management and clarification of cloud-of-concepts in ontologies. We present these results in Chapter 7.

Going intimately into the relationships between concepts belonging to ontologies is not part of the scope of O4OA because foundational ontologies and multi-level theories already deal with this. Nor is the objective to manage all possible characteristics involving FAIRness, but their effect on ontologies through well-established ontological principles. Furthermore, ontologies that manage FAIR Principles can be used in conjunction with O4OA to add these details to it; an example is the work [22], which sounds compatible and promising.

[2] `https://bfmartins.gitlab.io/o4oa`

Chapter 6

The Framework for Ontology Characterization

Labor omnia vicit improbus, et duris urgens in rebus egestas,
Publius Vergilius Maro (70 – 19 a.C.).

Domain specialists seek ontologies as tools to achieve FAIRness because they need data to produce relevant and necessary information, i.e., results. They comprehend that achieving desired outcomes transcends the confines of merely handling and manipulating data within a system but also requires systems interaction and information analysis and exchange. Observe that data and information are distinct things [16]. Thus, when specialists seek to clarify how a particular concept is treated in an ontology, they want to know if that concept is the same semantically in similar ontologies, if it is possible to reuse that concept in their proposal or ontologies, and whether it is possible to interoperate data supported by the definition of this concept between different ontologies.

Ontological analysis helps ontology engineers address these questions; however, this is not simple, as explained in this book. Chapter 5 presents O4OA that helps stakeholders in the make ontological analysis. In this chapter, we offer *The Framework for Ontology Characterization* (F4OC), which presents a systematic procedure to characterize ontologies using the features of O4OA. The objective of the framework is to provide a homogeneous, clear, and well-established base to compare ontologies [167, p. 337].

6.1 Framework Description

After clarifying the characteristics and meta-characteristics of an ontology as an artifact expressing a conceptualization, we must ensure that this acquired knowledge is used appropriately. Therefore, we propose the F4OC is an ontology-driven framework that is in line with the FAIR principles (mainly *"I"* and *"R"*) through the ontological support of the O4OA.

The framework comprises a reproducible set of layers of components required to do ontological analysis and build FAIRness ontologies. Part of the proposed components are of a technical or documentary nature, while part of them comprises processes to be followed. The proposal provides a predefined structure, rules, and guidelines that developers can build upon, reducing the need to start from scratch and enabling more efficient development. The F4OC unveils characteristics and meta-characteristics of ontologies and contextualizes their cloud of concepts (in line with the ontological commitment notion) to provide interoperability and reuse of ontologies effectively [168, p. 188]. Figure 6.1 presents the framework.

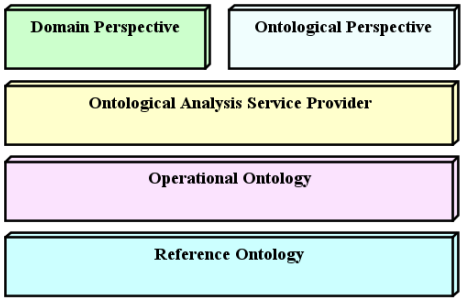


Figure 6.1: The Framework for Ontology Characterization.

Therefore, the F4OC composes:

Reference Ontology: O4OA, a meta-ontology to describe ontologies to guarantee FAIRness in ontologies and data.

Operational Ontology: implementations of O4OA to make it feasible to characterize ontologies, their characteristics, and meta-characteristics.

Ontological Analysis Service Provider: tools used to manipulate and manage O4OA implementations to promote easy information access to stakeholders and provide automation to the framework process.

Stackeholders' Perspectives: the stable and homogeneous environment in which the framework procedure is applied, facilitating the ontological analysis process by dealing simultaneously with ontological and domain perspectives.

Regarding the domain perspective, the framework process focuses on identifying the cloud of concepts that ontologies encompass, especially when ontologies compose networks. Besides, the strategy links to this cloud with the sources (standards, norms, etc.) established within the knowledge domain, embracing the domain context described in Chapter 3. This *Terminological Verification and Validation* process was firstly presented in the works [165, 168].

The framework domain perspective comprises five steps, as depicted in Figure 6.3 [1].

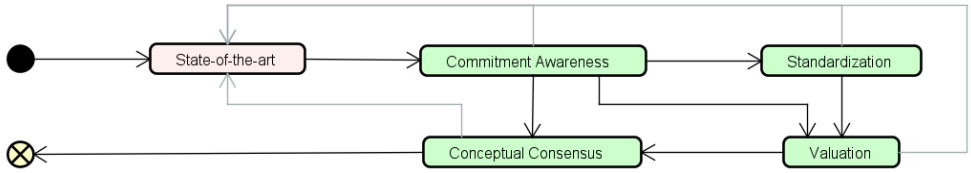


Figure 6.2: Framework for Classifying Ontologies – Domain Perspective.

From the ontological perspective, the framework process classifies ontologies through the ontological support of O4OA. Characterizing an ontology based on these classification levels is orthogonal since each classification can be carried out in an encapsulated form. However, there is a correlation between them. Thus, each classification level looks at ontologies with a separation of concerns. Still, there are important aspects [2] and relationships that ground these concerns, and they are grounded over O4OA. Indeed, we demonstrate in Chapter 2 the importance of a homogeneous basis of comparison to face the involved challenges in the ontology engineering process, and in Chapter 4 how this basis of comparison can help in terms of ontological analysis for FAIRness. This ontological perspective of the framework was first presented in the work [167].

[1] Throughout this chapter, we are using the UML Activity Diagram notation to present the F4OA details.

[2] Aspects in an ontological sense (essential properties).

The framework ontological perspective also comprises five steps, as depicted in Figure 6.2.

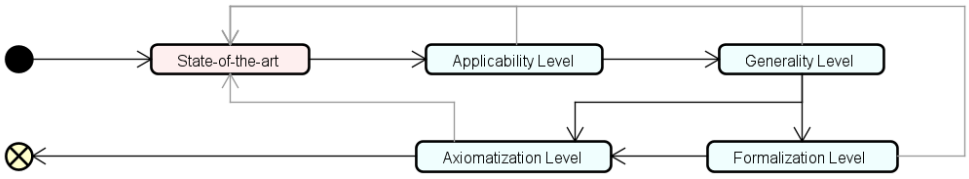


Figure 6.3: Framework for Classifying Ontologies – Ontological Perspective.

In Figures 6.3 and 6.2 the gray arrows indicate that it is possible to return to step 1 to do additional state-of-the-art research for further details or fulfill any lack of information if needed.

6.2 State of The Art

The F4OC step of *State of the art* is present in both perspectives as a starting point. Stakeholders can perform a single step uniting domain and ontological perspectives or separate the step, focusing mainly on each perspective separately. Stakeholders must search for relevant information concerning the state-of-the-art domain scope and ontologies describing this domain. The granularity of the search depends on the domain coverage. Therefore, it may refer to an ontology covering the entire domain composed of sub-ontologies, each covering domain parts, more specific ontologies, or network ontologies. This step process can be performed through direct research with specialists, a survey, a literature mapping, or even a systematic literature review when reproducibility is required. The information sources (documents, norms, standards, etc.) must be traceable and accessible. This cyclical process must be repeated until the largest information is obtained, and we made this step for our case study (see Chapter 3). Figure 6.4 presents the process in this step. Note that we highlight in red the next classification steps, in this case, the *Applicability Level* and/or the *Commitment Awareness*.

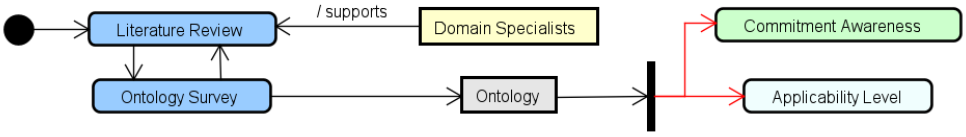


Figure 6.4: Framework for classifying Ontologies – State-of-the-art step.

6.3 Domain Perspective

As each domain has particular characteristics that are not always present in all domains, it is arduous (or impossible) to adopt a classification system such as that adopted from the ontological perspective. Generally speaking, the conceptual modeling community has targeted the solution by addressing domain characteristics through the FAIR Principles. Proposals that use a higher level of abstraction are useful in this sense, and the [22] proposal is a compatible work with O4OA in this direction. Therefore, the framework describes how this type of solution should be used and fed, pointing out the most important landmarks and the stakeholders' roles. Indeed, several people can play several roles, and each role can be played by several people. Therefore, stakeholders participating in one step may not be the same ones participating in another.

After the state-of-the-art step, the second step regarding the domain perspective is the *Commitment Awareness*. At this moment, ontology engineers are introduced to participating domain experts and the knowledge domain. They must identify and be aware of the scope, granularity, purpose, competence questions, and other information relevant to design ontologies in the domain.

The third step is the *Standardization*, in which stakeholders search for the best available and consolidated documents that are sources of information for the various definitions of the terms in the conceptualization. Some examples are classical books, international standards, glossaries, lexicons, classification schemes, reference models, and other ontologies. Each term and its definitions are confronted with the one used in the ontology; the objective is to identify other possible biases for the terms. Likewise, each definition present in the ontology can appear in sources to define other terms than those used in the ontology. The objective is to identify synonyms and terminology mismatches using the confronting of the clouds of concepts of ontologies.

Each term and its definitions are confronted with the terminology used by the domain specialists. The aim is to identify other possible biases for the terms. Similarly, any definition present in the ontology may appear in sources that define concepts other than those used in the conceptualization. The aim is to identify synonyms and terminological inconsistencies by comparing the term clouds of the ontologies. It is important to clarify that not every concept present in the cloud will be used by the ontology that describes such conceptualization.

The next step is the *Valuation*, in which stakeholders must compile conceptualization information in search of modeling anti-patterns, misinterpretations, and flaws that are potential points of divergence or causes of errors. At this moment, the under-development ontology competence questions must be already defined because they state the ontology requirements [7] and interfere in design decisions [74, 73]. The objective is to guarantee that the obtained for the ontology answers all these questions and accomplishes requirements. Moreover, stakeholders must agree by comparing the expected conceptualization with the one obtained for the ontology, preferentially a reference ontology, before its implementation.

The last step is the *Conceptual Consensus*, in which stakeholders reach a semantic consensus regarding the conceptualization and its functional requirements [3] The consensus must preserve the principles of representation defended in [76, 93] as well as the FAIR Principles. Finally, this agreement must (preferably) be accompanied by some contractual formalization so that future changes in understanding are necessarily well justified and tracked.

We also note that the already consolidated umbrella of ontological analysis activities within the Ontology Engineering community supports these steps. Besides, they align with the *Support Process* of the SABiO methodology [7]. The domain perspective of the framework strongly agrees with this methodology and provides two additional contributions with bias in ontology networks [168].

[3] *"Functional requirements refer to the knowledge to be represented"* [7].

6.4 Ontological Perspective

Likewise, in the domain perspective, the ontological perspective goes forward in the second step, classifying ontologies according to the *Applicability Level* [93]. This classification determines if the ontology documentation provides a *Reference Ontology*, an *Operational Ontology*, or both. The presence (or absence) of a reference ontology before its implementation depends on the choice of design methodology used. Indeed, the design methodology used by the ontologies developers is an important aspect of analyzing the applicability level, but not a classification itself. After classifying the ontologies, stakeholders identify which ontologies are well-defined and which are not, documenting and fulfilling all information in an O4OA implementation. Figure 6.5 presents the classification process in this step.

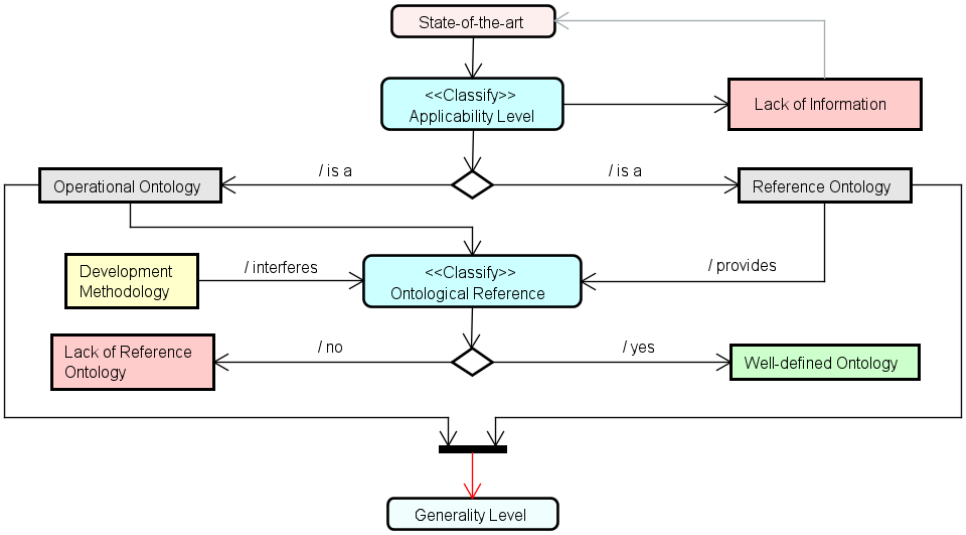


Figure 6.5: Framework for classifying Ontologies – Applicability level.

The third step uses the *Generality Level* classification [75, 247], in which stakeholders must consider several ontology aspects to position the ontologies correctly. Aspects such as competence questions, purpose, scope, and granularity are indirectly related to this classification and help to identify the generality level of ontologies. However, the key aspect stakeholders must consider is whether ontologies are well-grounded, and which are not. Therefore, they must verify whether the ontologies have any ontological

grounding through some *Foundational Ontology*. Similarly to what was made in the prior step, stakeholders must document and fulfill all information in an O4OA implementation. Figure 6.6 presents the classification process in this step.

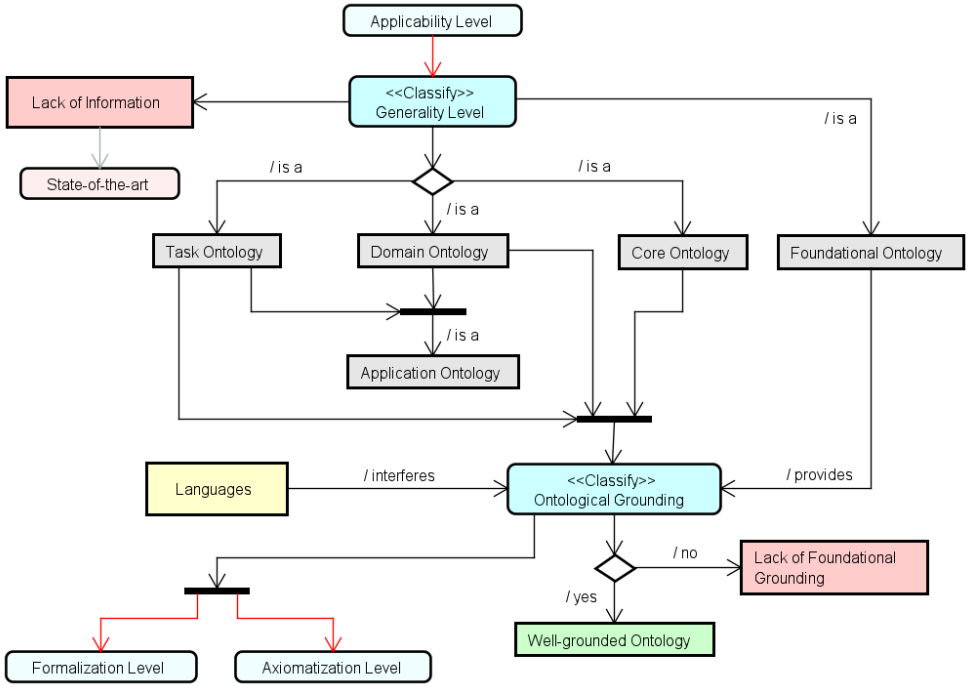


Figure 6.6: Framework for classifying Ontologies – Generality level.

In the fourth step, stakeholders can classify ontologies based on their formalization following the bi-dimensional approach of [66]. This evaluation depends on the implementation (if it exists) and other ontology characteristics. Additionally, several meta-characteristics imply a greater or lower formalization, such as the type of ontologies appropriateness, semantic patterns, and anti-patterns. Moreover, language isomorphism implies that a greater or lower axiomatization allows conceptualizations to be represented. We explain in Chapter 3 and clarify in O4OA how these characteristics and meta-characteristics interact. In the framework, we establish a linear dimension from *Informal Ontologies* to *Formal Ontologies* to quantify the formalization, delimiting ranges depending on the ontologies' characteristics

and meta-characteristics. We also establish a second linear dimension regarding the classification proposed in [156, 68], where the *Heavyweight Ontologies* correspond only to the ones from *Logic programming* to *General Logic* (this includes *First-order Logic*, *Higher-order Logic*, *Modal Logic*). Figure 6.7 presents the process in this step.

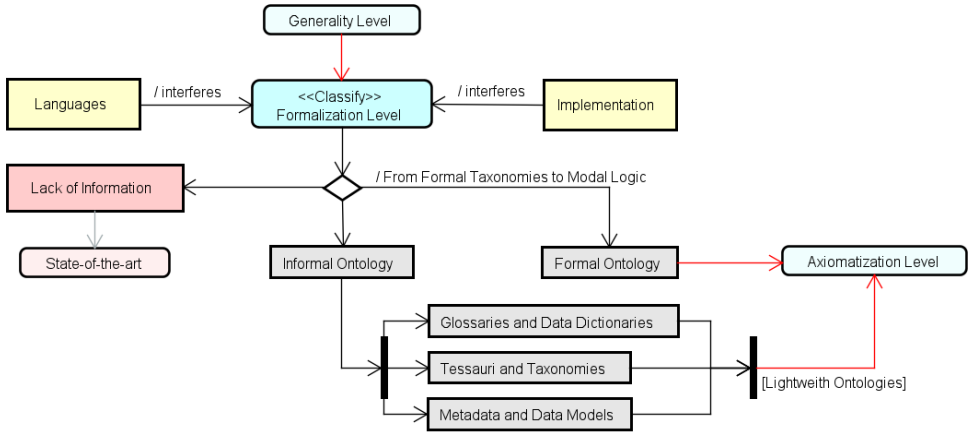


Figure 6.7: Framework for classifying Ontologies – Formality level.

The classification according to the *Axiomatization Level* is related to the previous (*Formalization Level*) because it is subject to the influence of the same characteristics and meta-characteristics. Therefore, we also establish a linear dimension regarding ontologies axiomatization [68], from *Lightweight Ontologies* to *Heavyweight Ontologies* based on the number of axioms (this value may be estimated). In this case, stakeholders can use greater granularity to evaluate the axiomatization, which is useful when there is a lack of information. The axiomatization can also be derived from the analysis already carried out in the previous step and can even be done with some automatism. Using greater granularity at this stage indicates that additional state-of-the-art research is probably necessary to reach further details. Likewise, in the other steps, stakeholders must document and fulfill all information about formalization and axiomatization in an O4OA implementation. Figure 6.8 presents the process in this step.

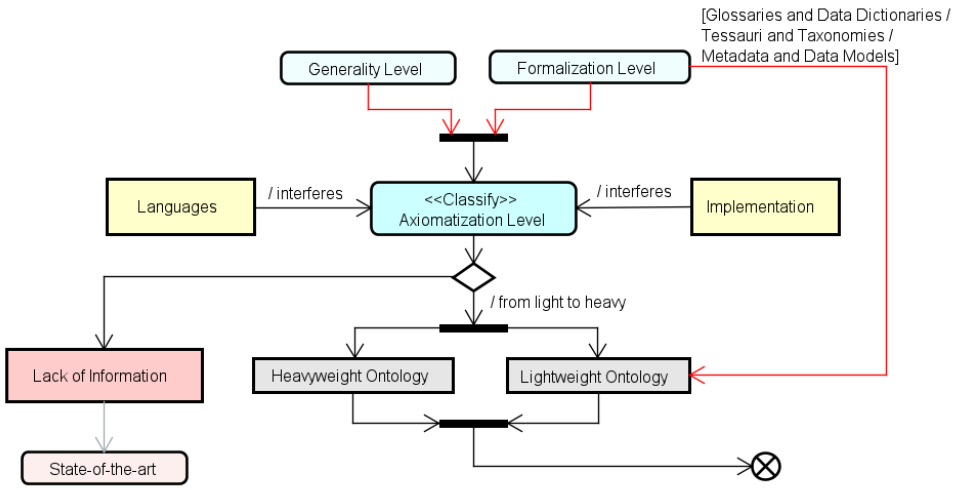


Figure 6.8: Framework for classifying Ontologies – Axiomatization level.

6.5 Conclusions

From its ontological support, the F4OC drives stakeholders in the ontological analysis, approximating their distinct perspectives (domain and ontological). Besides, it provides a robust and systematic comparison base to support conceptualization interoperability. These are key features for FAIRness ontologies that are tools to provide FAIRness data, making it feasible to deal with network ontologies, especially when the goal is to interoperate information among domains. The F4OC facilitates the cross-analysis of ontologies, i.e., provides a comparative ontological analysis among conceptualizations [167]. This approach aims to explore and make available ontology characteristics and meta-characteristics, providing the necessary basis for the interoperability process instead of analyzing each ontology singly. By using O4OA and following the F4OC guidelines, stakeholders can design a solution that ensures compliance with their requirements in an environment where they can apply their metrics. This allows stakeholders to find ontological problems and helps them fill concept misinterpretation gaps. Chapter 7 explores the outcomes we have in this regard.

Part IV

TREATMENT VALIDATION

Chapter 7

Applications of O4OA and F4OC in the Cybersecurity Domain

Experientia docet, Cornelius Tacitus (56 – 118 d.C.).

Considering the characteristics of the cybersecurity domain already presented in Chapter 2 and the issues that interfere with the ontological analysis process of domains like this (presented in Chapter 3), we proposed O4OA (presented in Chapter 5) and F4OC (presented in Chapter 6) targeting these issues. To validate our proposal, we opted for a semi-automatic approach to deal with the cybersecurity domain. Figure 7.1 depicts the general of the framework, showing its application in the cybersecurity domain (as our case study). Indeed, this is a F4OC application (among many possible) made over the general view presented in Figure 6.1 of the Chapter 6. Note that some of the proposed components (presented in white color) are not fully implemented (`gO4OA` and `Frontend`) because the design approach or are not in the scope of this thesis (`AI Solution`).

As a reference ontology, O4OA is the foundation for several implementation versions. We chose an implementation in *MongoDB* [1] and called it MongoO4OA in our case study. Also, we strongly recommend that an OWL version be implemented, in this case using gUFO, given its compatibility

[1] `https://www.mongodb.com/`

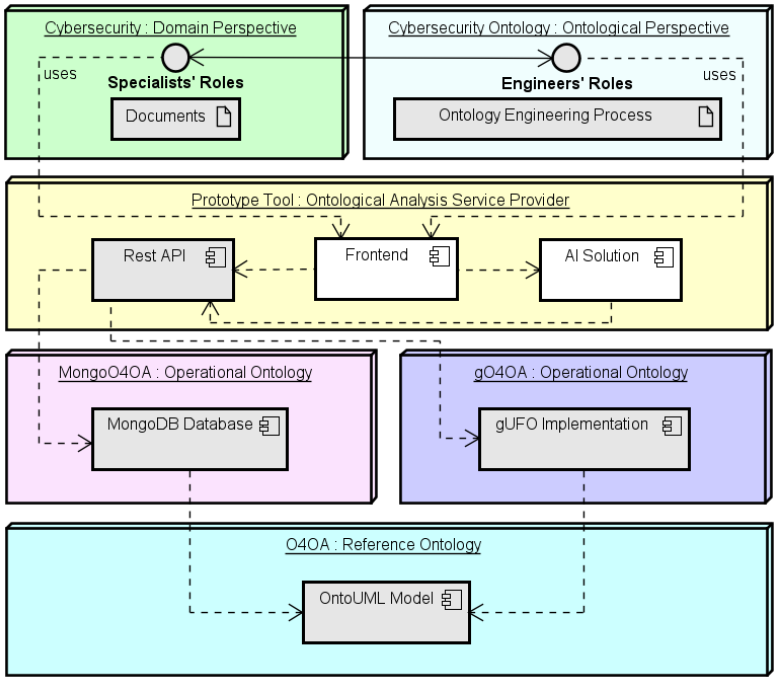


Figure 7.1: The Framework for Ontology Characterization – cybersecurity case study.

with the OntoUML reference model, accessibility, and dissemination in the ontology engineering community. Above this structure, the development of an Ontological Analysis Service Provider is suggested since complex and vast domains must manage large amounts of data that require some level of automation. Finally, we must deal with the stakeholders' perspectives (domain and ontological): the *Ontological Perspective* as a conceptual modeling approach and the *Domain Perspective* as a domain of knowledge specialists' viewpoint. The former looks at the semantic foundation, while the latter deals with the knowledge domain itself.

7.1 An Operational Version of O4OA

According to SABiO, developing an ontology version from a reference model (reference ontology) requires that this model to be analyzed through processes of model instantiation to explore possible issues or unexpected possibilities scenarios (branches or worlds). Thus, we use the validation notions present in [20] to validate O4OA by performing it concurrently with the development of the operational version of the ontology. We fragment the analysis, running an Alloy Analyzer [139] instantiation of each model package in an individual and modular way due to the O4OA model characteristics (size and complexity) [2].

The instances model helps elicit additional constraints (in addition to those already present in OntoUML) required, and we also check model cardinalities to ensure correct semantics [7]. For example, when analyzing the instantiation of the contents of the `Reuse` package, because the reused and reuser ontology roles are not *disjoint*, we had to add a constraint to avoid cyclic reuses, i.e., a *Transitive Closure* predicate for the relations `reuses`. Note that some required constraints must be implemented directly in the persistence while others must be in the API. More details are available in the ontology repository.

Given each part of the O4OA reference model, we implemented it (operational version) in data storage with MongoDB. We chose this data storage because it composes a modern open-source suite able to work in *Microservices Architectures*, provides a cloud solution compatible with AI solutions, and deals with data in scale. It is a document-oriented database in which data documents are stored in structures called *Collections*, which can have several free fields; however, it is possible to add validation rules to constrain the data. Figure 7.2 shows the MongoO4OA actual version visualized thought the *Studio3T* [3] GUI tool[4].

We manually added the data collected from the TLR (see Chapter 3) about the ontologies belonging to our case studies [165]. Up to now, we have assessed 161 concepts in the cybersecurity domain from the TLR and many others obtained from the associated foundational and domain-correlated ontologies studied. Associated with these concepts (in the cloud of concepts), we registered 73 reliable sources, providing a burst of possible usage definitions in ontologies of this and its related domains. For instance, taking the concept of *Risk*, we found 18 definitions of *what it is a Risk*. Besides, we also found many other risk-related definitions, the ones for concepts such as *Level of Risk*, *Residual*

[2] See the O4OA packages in Appendix C.3

[3] `https://studio3t.com/`

[4] The figure's hidden parts (gray dash) are sensitive information.

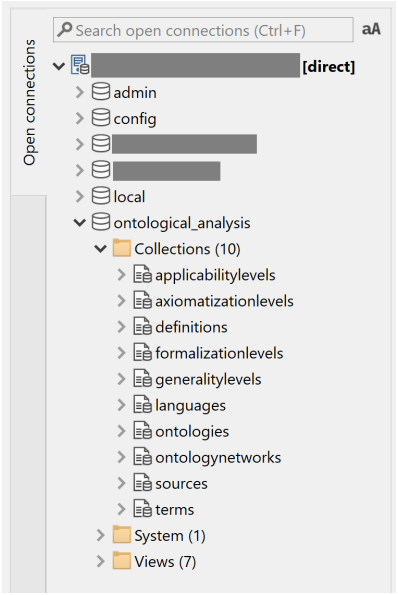


Figure 7.2: MongoO4OA version – partial implemented O4OA version.

Risk, *Risk Criteria*, and other 11 associated ones. See in Appendix D.2 some query results from the database.

7.2 A Semi-automatic Support for Ontological Analysis

In order to manage and analyze the data stored in the MongoO4OA we developed a backend solution composed by *REST-API* [166] made with *NodeJS* [5] and *ExpressJS* framework [6]. The solution uses a distributed and compartmentalized environment through *Docker* containers [7]. Besides, we develop a frontend prototype solution to provide easy, responsive, and multiplatform access to the data by using *Angular Material* UI component library [8] and several additional open-source libraries to perform complementary features. Figure 7.3 shows a prototype tool screenshot with (the first page) the list of ontologies that belong to one of the O4OA case studies.

[5] `https://nodejs.org/en`
[6] `https://expressjs.com/`
[7] `https://www.docker.com/`
[8] `https://material.angular.io/`

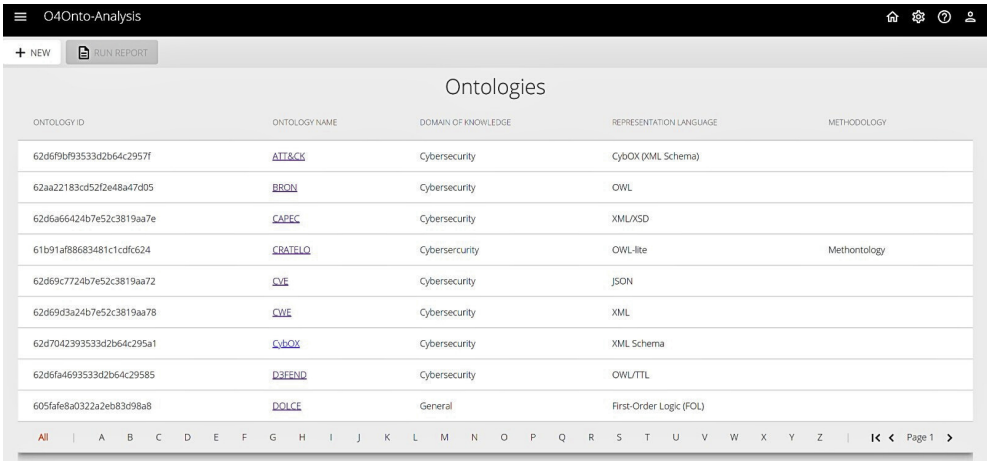


Figure 7.3: Screenshot of the O4OA prototype tool – list of ontologies.

Figure 7.4 shows a prototype tool screenshot with the classification of CWE, one of the ontologies belonging our case study.

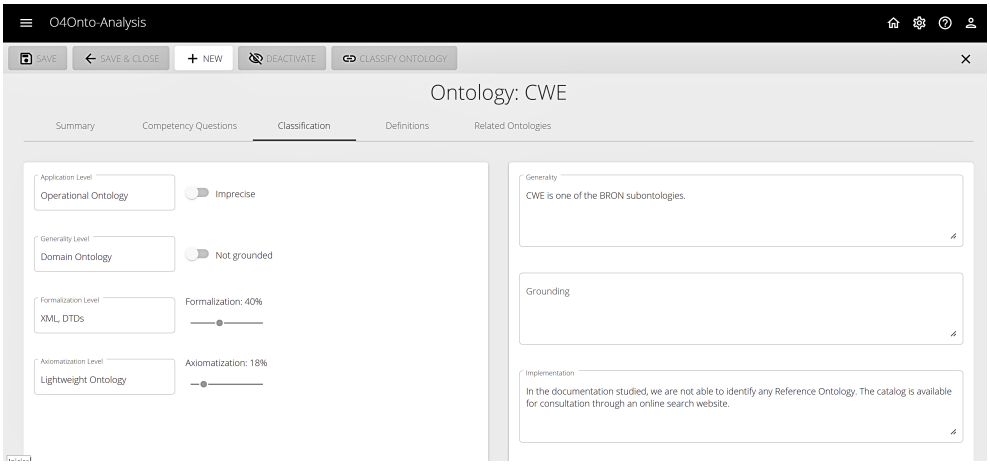


Figure 7.4: Screenshot of the O4OA prototype tool – CWE classification.

Figure 7.5 shows the concept definitions that are linked with CWE.

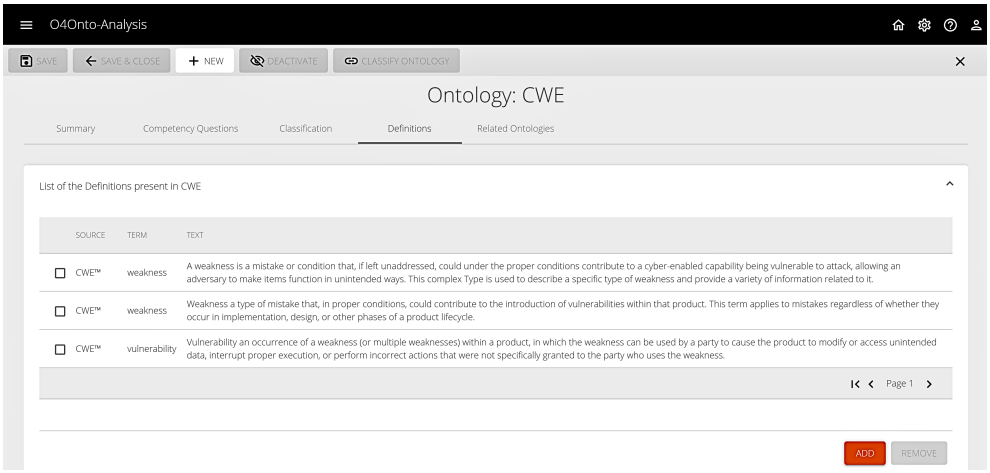


Figure 7.5: Screenshot of the O4OA prototype tool – CWE definitions.

Figure 7.6 shows a prototype tool screenshot with (the first page) the cybersecurity cloud of concepts belonging our case studies.

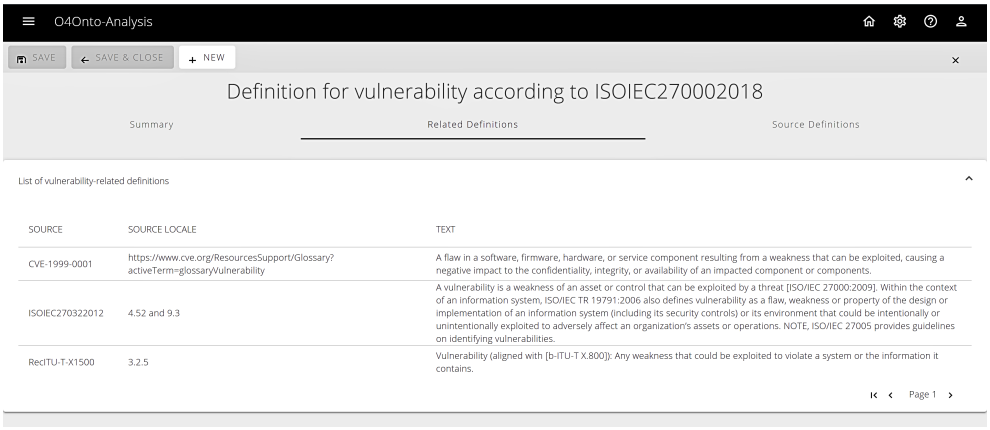


Figure 7.6: Screenshot of the O4OA prototype tool – Cloud of concepts.

We consider this prototype tool as an *alpha* version because it is still under development. However, the implemented parts are already sufficient to explore the potential of the proposal since we adopt the strategy of *Agile Development* [9]. Therefore, despite this limitation, we can trace the concepts from their definitions until we reach the ontologies, use them, and classify them according to the lights of O4OA.

7.3 Applying FO4C on the Cybersecurity Domain

During the progress of the project, the team members (stakeholders) participated in the work stages of the proposed framework. They were able to take several interesting outcomes from the ontologies analyzed. We focus on some key concepts of the domain instead of producing a complete ontological analysis of each ontology. The main reasons that led us to this choice were:

Firstly, as the number of concepts in the domain is enormous and there is a central set of concepts that repeatedly appear in the various ontologies raised [165]. Thus, having a more specific look at these concepts were considered more relevant.

Second, we aim not to criticize a particular ontology or be extremely precise in correcting it as a whole but to point out the benefits of applying the framework to detect semantic problems in concepts.

Third, dealing with a few concepts in different ontologies opens the possibility of finding semantic variations covering interoperability and reuse of ontologies.

Fourth, by demonstrating the usefulness of the framework in applying it to central concepts, it appears to be an indicative sample of the good scalability, traceability, and formalization of the proposal.

Finally, the analysis process can be extended to other concepts and even complete ontologies if necessary for the project in future opportunities.

[9] It is important to point out that we adopted an *Agile Development* approach to provide fast initial results meanwhile being scaled.

The first concept we studied was the concept of *Risk*, and the second was *Vulnerability*, extending both to the related conceptualization and neighboring concepts.

7.3.1 What is Risk?

The classification and characterization of ontologies help stakeholders in the ontological analysis process, but there are many other outcomes F4OC can provide. For instance, trace the relationships among the ontologies. In our case study, stakeholders can identify if one ontology is a sub-ontology of another if one reference ontology provides one or more (different) implementation versions (operational ontologies), which ontologies use the same (or similar) foundational ontology, which are the ontologies overlapping the domain (or domain parts); as well as others relationship results. Besides, the cloud of concepts provides the standardization support and context applied to the ontologies that allow knowing how the domain specialists use the conceptualizations in-depth. Furthermore, by bringing these perspectives together, we can produce outcomes like those presented in [167] by answering stakeholders' questions. In this case, these are the questions made during the ontological analysis process comparing the notion of *Risk* used in SECCO (a sub-ontology of CRATELO [197]) with other ontologies (COoVR [216], Ontology of ISO/IEC 27005 [4], and Mulval [202]).

- Is the *Risk* concept interoperable between SECCO and COoVR?
- Is the *Risk* concept interoperable between SECCO and the Ontology of ISO/IEC 27005?
- Is the *Risk* concept interoperable between SECCO and Mulval?
- Is the *Risk* concept interoperable between SECCO and other ontologies?

Answering these ontological analysis competence questions, stakeholders aim to verify if the *Risk* concept is the same at the ontological level, identify the concepts and relations in which the *Risk* is applicable (*"the thing at risk"*), and verify if the context of use for both *"the thing at risk"* and the *Risk* itself. Figure 7.7 summarizes the concept of *Risk* cross-analysis (see the running results in Appendix D.3).

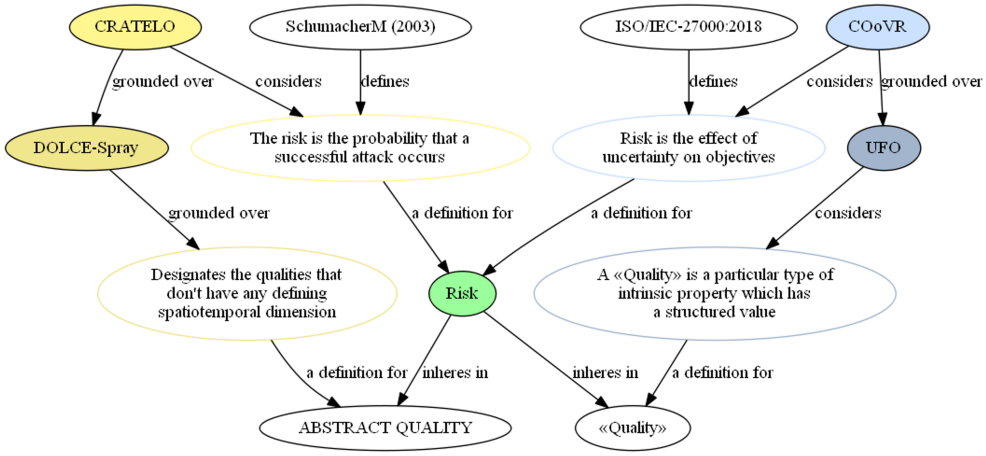


Figure 7.7: The concept of *Risk* concept cross-analysis as an outcome of the ontology characterization framework.

In the comparative study considering the concept of *Risk* and its surrounding conceptualization, stakeholders deal with definitions from diverse contexts, all of which are well-supported by known cybersecurity standards. Besides, each ontology studied involves different contexts, and the stakeholders' biases are present. Indeed, the authors of one of the proposals studied mention that *"at the same time, neither practitioners nor ontologists pay comparable attention to the concepts traditionally associated with risk, such as probability or likelihood of an adverse event, and the cost of consequences or impact of the event. Such concepts, which are canonical in most definitions inspired by traditional definitions of risk, are mentioned very infrequently in discourses of practitioners and with only moderate frequency by ontologists"* [196, p. 64]. In summary, conceptual ambiguities caused by the lack of an ontological foundation combined with the complexity of the domain itself are the main cause of misinterpretations and problems found, emphasizing the importance of analyzing possible intersections and unions of the definitions taken and according to their contexts (ontological commitment and foundation). F4OA proved to be fundamental in clarifying these outcomes, considering the large volume of information that needs to be analyzed and considering that stakeholders analyze a single concept in this case.

7.3.2 What is Vulnerability?

The notion of *Vulnerability* is another representative concept for the cybersecurity domain. Therefore, we formulate similar competence questions to those made for *Risk*. In this case, the objective targets clarify this concept by comparing it in CVE and CWE, which are widely used by domain specialists (cybersecurity specialists). The competence questions are:

- Is the *Vulnerability* concept interoperable between CVE and CWE?
- Is the *Vulnerability* concept interoperable between CVE and other ontologies?
- Is the *Vulnerability* concept interoperable between CWE and other ontologies?

Figure 7.8 summarizes the concept of *Vulnerability* cross-analysis between CVE and CWE (see the running results in Appendix D.2).

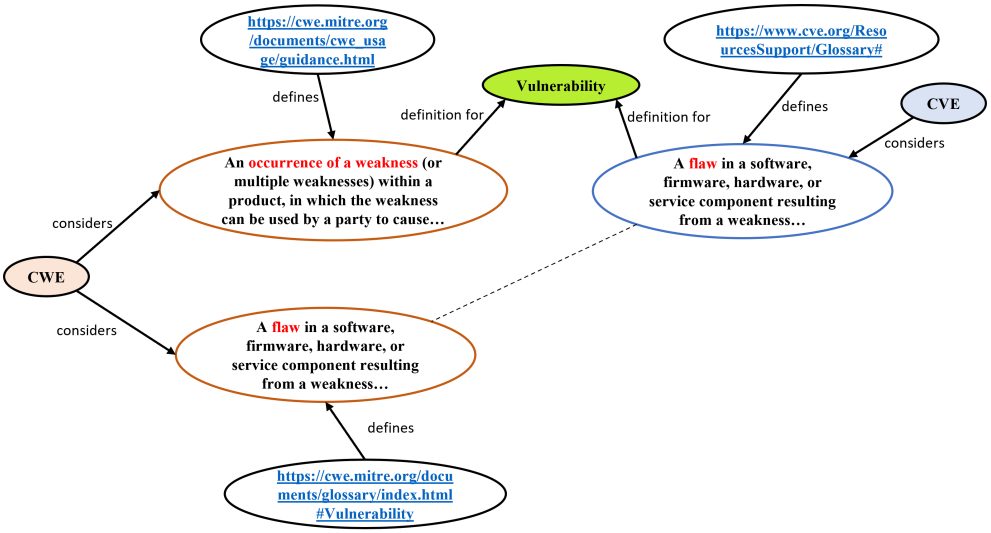


Figure 7.8: The concept of *Vulnerability* concept cross-analysis CVE and CWE as an outcome of the ontology characterization framework.

By comparing the approach used in CVE and CWE, stakeholders found inconsistencies in the definitions used in these ontologies. In this case, CWE presents two different definitions for this concept, one whose source is the

same as that used in CVE and another with a different source. From this divergence, stakeholders add other questions to be answered in this analysis:

Domain specialists' question: Does the *Vulnerability* registered is correct?

Ontology engineers' question: Does *Vulnerability* is a flaw or an occurrence of weakness?

To answer the first question, we firstly resort to the source of the *Vulnerability* definitions in the MongoO4OA, in this case the *CVE Glossary*, in which this concept means: *"A flaw in a software, firmware, hardware, or service component resulting from a weakness that can be exploited, causing a negative impact to the confidentiality, integrity, or availability of an impacted component or components"* [10]. However, we also found references to the other definition in several works; some examples are [243, 55, 170]. As these publications are before the glossary update, we resort to the *Web Archive* looking for prior versions of the glossary. Indeed, the prior definition is: *"Vulnerability, an occurrence of a weakness (or multiple weaknesses) within a product, in which the weakness can be used by a party to cause the product to modify or access unintended data, interrupt proper execution, or perform incorrect actions that were not specifically granted to the party who uses the weakness"* [11]. Then, both definitions are valid, depending on the glossary version used (See pieces of evidence in Appendix D.5). By the way, CWE also considers *Vulnerability* as a *Flaw* just like the definition most recently adopted in CVE.

To answer the second question, stakeholders searched in the cloud of concepts for other possible definitions of the concepts of *Flaw*, *Weakness*, and synonyms, such as *Failure*. This way, it was possible to find the ontology OSDEF [56], which is part of the SEON [26], an ontology network that is part of another ongoing case study already registered and not directly related. Even so, stakeholders proceeded with the ontological analysis based on the data provided by this ontology because it is classified as well-grounded (grounded over UFO). Furthermore, OSDEF is an ontology applicable to the software engineering domain; therefore, it is a related area. Figure 7.9 summarizes the concept of *Vulnerability* cross-analysis between CWE and OSDEF.

[10] `https://cwe.mitre.org/documents/glossary/index.html#Vulnerability` with last update on November, 16th of 2022

[11] `https://web.archive.org/web/20221003210527/https://cwe.mitre.org/documents/glossary/index.html#Vulnerability` with last update on October, 10th of 2021

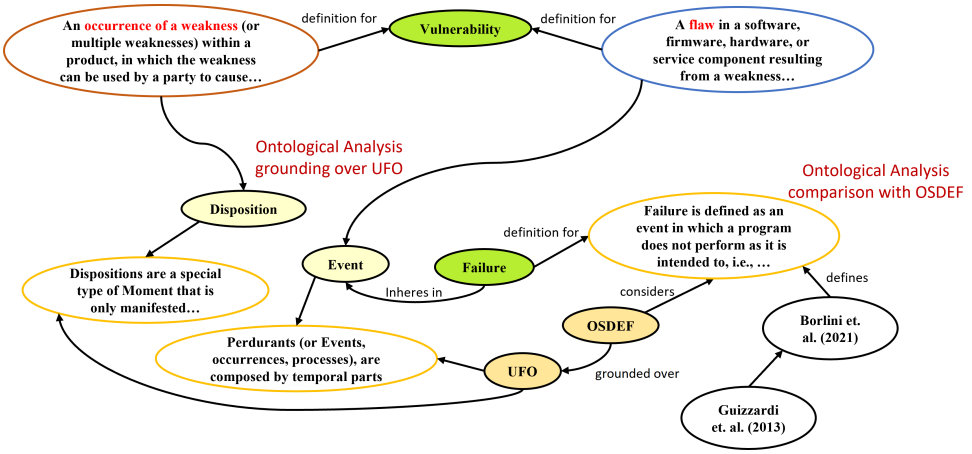


Figure 7.9: The concept of *Vulnerability* concept cross-analysis CWE and OSDEF as an outcome of the ontology characterization framework.

In this case, OSDEF defines *Vulnerability* (i.e., Disposition), which is the manifestation of a *Failure* (i.e., Event), i.e., these are different concepts [168]. This means that those CVE *Vulnerability* definitions in the light of a UFO ontological analysis imply different grounding concepts. In detail, the CVE Glossary updated definition takes *Vulnerability* in the ontological sense of *Event* [12], while the prior definition takes *Vulnerability* in the ontological sense of *Disposition* [13]. Therefore, we are dealing with different valid semantics of the same concept in the same ontology in distinct time moments. Then, stakeholders made another question:

- Considering that the *Vulnerability* definition was updated, does the data also was updated?
- Does the data context refer to events or dispositions?

Answering these questions, we identify that CVE controllers properly control the updating of data contained therein (*ID* and *Records*), but we could not find records about the semantic adjustments of their model. Upon consultation, we found no evidence that the related data underwent changes due to the update in the definition of *Vulnerability*. Changes in agreement, understanding, evolutions, and corrections are part of the ontology life cycle; however, they must be properly controlled, recorded, and published. Indeed, this is part of

[12] UFO notion of *Perdurant* [96], and as also described in OSDEF [56].
[13] UFO notion of *Moment*, such that dispositions manifest as events [101].

the best-practices ontology engineering process, aiming to ensure ontologies FAIRness. Moreover, this helps ontology users remain properly updated.

With the support of our domain specialists' team, we verify that CVE (and CWE) vulnerability data refers to **vulnerabilities as types** (*Universals*). For instance, the `CVE-1999-0067` vulnerability details [14] where any affected *"PHF CGI program allows remote command execution through shell metacharacters"* [15]. Therefore, any CGI BIN program (types of programs of the subtype CGI BIN, an *Endurant Universal*) that presents certain characteristics (included with NCSA httpd, and Apache 1.0.3, a *Mode Universal*) suffers this type of vulnerability (as a *Disposition*). This is different from the CGI BIN program (in the server named x_1), which is vulnerable because it belongs to this subtype of programs that are CGI BIN vulnerable (*Individual*).

A similar (meta)characteristic notion can be seen in magnets (types of things that are magnets – *Endurant Universal*) have the characteristic of magnetism (*Mode Universal*) having the capacity to attract ferromagnetic objects (a *Disposition*). Indeed, object dispositions are the cause of events (*Event* instances) to occur when a particular magnet (*Individual*) attracts a coin, for instance. This is why a *Disposition* is existentially dependent on its bearers [101]. In summary, the concept of *Vulnerability* in OSDEF represents *instances of* the concept of *Vulnerability* in CWE and CVE, although the term used is the same. Furthermore, for both ontologies, CWE and CVE, the previous definition used in CVE is more adequate than the updated one to indicate dispositions. Therefore, CWE and CVE can interoperate with OSDEF through the concept of *Vulnerability*. However, for this to be possible, it is necessary to consider *Vulnerabilities* as dispositions and that the bearers of these dispositions (vulnerabilities) must relate to each other through an `instance of` relationship.

[14] See `https://www.cvedetails.com/cve/CVE-1999-0067/`

[15] *"There exists a vulnerability in the sample cgi bin program, PHF, which is included with NCSA https, and Apache 1.0.3, an NCSA derivative. By supplying certain characters with special meaning to the shell, arbitrary commands can be executed by remote users. In case of a successful attack, a remote user may retrieve any world-readable files, execute arbitrary commands, and create files on the server with the privileges of the httpd process, which answers HTTP requests. This may be used to compromise the http server and, under certain configurations, gain privileged access"*, a source at `https://advisories.checkpoint.com/advisory/cpai-2003-47/`

7.3.3 How do BRON initiative ontologies interact?

BRON [116, 115] is an open-source initiative [16] from ALFA group at CSAIL, MIT [17]. It is an ontology network that links together offensive, defensive, and vulnerability concepts to analyze potential attacks and their potential countermeasures [115]. This is a work of interest to our project team, which is why it was added to this work as a case study. Furthermore, this case study adds a more general view to our validation because it deals with the relationships between ontologies instead of dealing with specific concepts. Therefore, we apply F4OA to BRON and all its ontologies. We use the registered concept cloud (in MongoO4OA) to work with the concepts involved, as it already comprises the necessary elements.

BRON implements its ontological approach as a knowledge graph using OWL and RDF formats [18]. The implementation of BRON links together *ATT&CK* [228, 154, 33] [19] and *CAPEC* [20] (for the offensive concepts), *D3FEND* [148] and *Engage* [215] [21] (for the defensive concepts), and *CWE* [163, 162] [22] and *CVE* [161] [23] (for the vulnerability concepts).

Observe CVE and CWE compose BRON and already are in our *Vulnerability* concept case study. Therefore, stakeholders previously know they are interoperable themselves, and the semantics of their data is mostly under the context of *Universals*; likewise, the *Vulnerability* concept has already been analyzed. They also know how this concept relates to OSDEF (an ontology external to BRON). Therefore, stakeholders formulated two competency questions for this case study.

Domain specialists' question: How do the ontologies that makeup BRON relate to each other?

Ontology engineers' question: What benefits for the domain that interoperates OSDEF with BRON can bring?

To answer the domain specialists' question, we classified each ontology of BRON based on the information sources we obtained searching the state-of-the-art. We have the advisor of our team specialists to clarify our

[16] `http://bron.alfa.csail.mit.edu/info.html`
[17] `http://alfagroup.csail.mit.edu/`
[18] `https://github.com/ALFA-group/BRON`
[19] `https://attack.mitre.org/`
[20] `https://capec.mitre.org/index.html`
[21] `https://github.com/mitre/engage`
[22] `https://cwe.mitre.org/`
[23] `https://www.cve.org/`

doubts about the domain specificities. Then, we carry out all the steps proposed in F4OA for BRON itself and incorporate the information collected about each of its related ontologies. In this case, we identify the following relations:

ATT&CK is a sub-ontology of BRON;

D3FEND is a sub-ontology of BRON;

DAO is a sub-ontology of D3FEND (and BRON);

DAO is a reuse of ATT&CK, mapping concepts from ATT&CK to D3FEND [148];

ATT&CK reuses CAPEC, CVE, and CWE, such that its concept of *Technique* links to an existing instance of CAPEC (type of type relations);

CAPEC, CVE, and CWE are sub-ontologies of BRON.

Using the framework, stakeholders can understand the absence of ontological patterns (semantic consensus) in those reuse relations adopted in BRON. Indeed, the *Vulnerability* concept analysis presented in Subsection 7.3.2 is one evidence of misinterpretations or at least the use of additional programming to resolve this issue, especially comparing the definition used to conceptualize what is a *Vulnerability* in CVE and CWE, that is reused by ATT&CK. Other competency questions arise from this situation, but we do not detail the analysis further to avoid overloading the text repetitively. In summary, the integration of BRON (its ontologies) with an ontology such as OSDEF, aiming to track (or reason about) vulnerabilities to predict and avoid (or mitigate) weaknesses in software systems, becomes more intricate, given the analysis presented through the F4OC. This happens because BRON deals with types of vulnerabilities; in contrast, OSDEF deals with particular system vulnerabilities.

Figure 7.10 visually represents the MongoO4OA outcomes about BRON used in this F4OA case study. Note there is a difference between BRON as an ontology part of the BRON initiative network, and both are elements present in MongoO4OA (see running results in Appendix D.4).

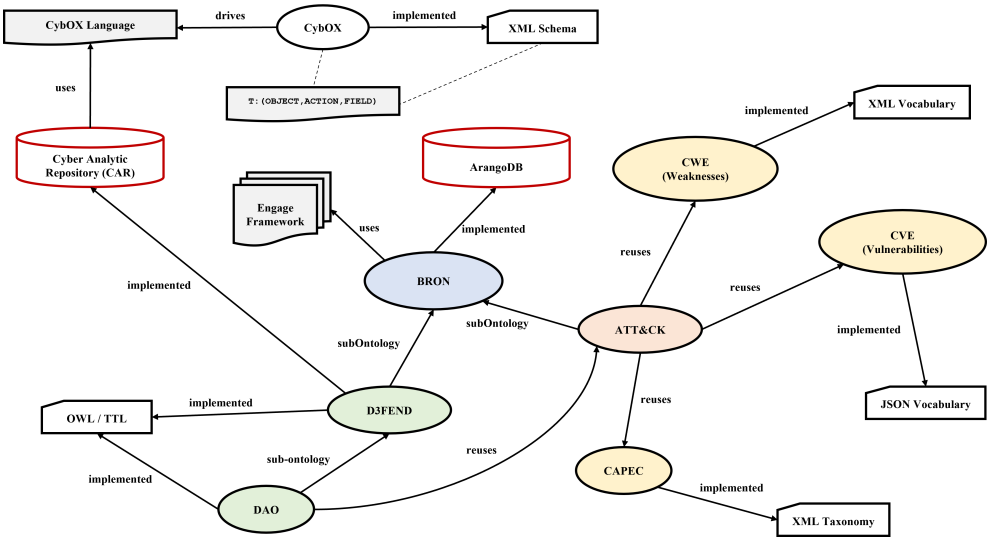


Figure 7.10: Framework outcome showing relational aspects present BRON.

Answering the ontology engineers' question, Cybersecurity specialists usually play distinct roles; the so-called *Red Team* act as if they were attackers, while the *Blue Team* is an integral part of the defense actions. Both use public catalogs, such as CVE and CWE, to identify vulnerabilities and patterns of attack. However, not surprisingly, real attackers often use these sources of information about vulnerability patterns and attacks as inspiration for their actions, and they are very creative. Furthermore, real-world attacks typically aim to achieve specific and concrete objectives; they are not launched fortuitously. Therefore, the blue team must focus on defending the systems individually, tracking where these general attack patterns are present.

Like real attackers, the red team uses patterns to achieve specific goals, and they can test different options repeatedly until they succeed. The most common is using new forms of attack, combining types of vulnerabilities creatively and often innovatively. Instead, the blue team starts from concrete systems that can have vulnerabilities (including an unexpected mix of them) that are yet unknown. Therefore, identifying known types of vulnerabilities and patterns of attack helps, but only when corrective actions are applicable, hardly as preventive actions. Besides, the many possibilities to explore vulnerabilities in a specific system favor the attacker. Worse, a predictive defense does not have a second chance to avoid catastrophic consequences. Solutions based on proposals such as OSDEF can effectively track in-depth

vulnerabilities and predict specific system failures with requirements and coding traceable, making interoperability between OSDEF and BRON (its ontologies) an interesting approach.

7.4 Conclusions

The F4OC facilitates the process of ontological analysis by elaborating its characteristics and meta-characteristics that go beyond the explicit notions of ontologies. The framework includes an approximation of the key factors of the ontological perspective (accessibility, availability, sharing, aspects of modeling and implementation, etc.) to those of the domain perspective (cloud of concepts, domain structuring, granularity, etc.). In addition, F4OC can work with individual concepts, ontologies, or ontology networks, ensuring traceability via a stable basis for comparison. The project members validate the usefulness of F4OC. The team is composed of a multidisciplinary group of stakeholders, including cybersecurity specialists, ontology engineers, academic researchers, and managers from various institutions. The team validates the capabilities of the framework through a series of case studies that match the interests of the project partners, as we demonstrate in Sections 7.3.1, 7.3.2, and 7.3.3. The validation process for the work was conducted with our partners, especially our industrial contributors. We held weekly meetings [24] with stakeholders in which we were able to interact with the red and blue teams, analyzing their requirements and receiving advice and feedback from them regarding our work.

We implemented the central elements to make feasible proof of the F4OC usefulness, and the frontend does have limitations on the visual representation of the API results. We intend to improve it, adding graph presentation to allow dynamic navigation in the cloud of concepts as well as in ontology (meta)characteristics [25]. Therefore, to illustrate the presented case studies e made manually, Figures 7.7, 7.8, 7.9, and 7.10 according to API results (in JSON) because the prototype does not have an effective GUI to present these results as visual graphs yet.

In practical terms, ontological analysis supported by reference ontologies has already demonstrated its benefits in managing and sharing knowledge in the human-machine context. However, when it is necessary to apply

[24] Our meetings were duly scheduled and documented with minutes. But because of a contract, we can't share the meeting details or the internal data we use during the work.

[25] It is important to point out that we adopted an *Agile Development* approach to provide fast initial results meanwhile being scaled.

semantic capabilities in the machine-to-machine scope through systems supported by operational ontologies, there is still a lot to be done. New proposals within the spectrum of AI and ML can be a way for machines to deal with semantics, as operational ontologies can enrich mathematical models, such as the ones used in LLM, for instance. Following our approach with F4OA, our industrial partners are already working on a commercial solution. Their objective is the automatic and concrete identification of countermeasures based on vulnerability descriptions. They aim to mitigate or even prevent cyberattacks by enabling professionals to apply cybersecurity recommendations to real-world use cases [151].

Part V

CONCLUSIONS AND FUTURE WORKS

Chapter 8

Conclusions and Future Works

Bis vincit qui se vincit in victoria,
Publilius Syrus (85 – 43 a.C.).

We begin this research by formulating questions that pursue three main goals: to clarify and homogenize the necessary meta-ontological requirements, data, and characteristics to help stakeholders achieve awareness and common sense about conceptualizations; to provide a clear baseline for characterizing and comparing ontologies to facilitate the interoperability and reusability of ontologies through their *Ontological Analysis*; and to validate the contributions of this research.

From the introduction, the problems associated with developing ontologies for representing and processing knowledge are from vast, largely regulated, or sensitive domains, such as the cybersecurity domain. We propose the *Ontology For Ontological Analysis* (O4OA), a meta-ontology that classifies and characterizes ontologies from their (meta)characteristics. Then, above the O4OA knowledge, we propose the *Framework for Ontologies Characterization* (F4OC). O4OA is made with OntoUML (UFO) and was developed through the SABiO methodology. Considering domain specialists' and ontology engineers' perspectives and roles in the ontology engineering process and managing a cloud of concepts, the proposed meta-ontology investigates ontologies (conceptualization artifacts) based on FAIR Principles; meanwhile, it is trapped by complying with these principles to be subject to FAIRness itself. F4OC depicts how to deal with the O4OA knowledge by providing a

homogeneous, reproducible, and traceable environment to facilitate ontological analysis. This ensures FAIRness to ontologies and their data.

Finally, we tested the validity of the proposal by implementing O4OA and the necessary tools for applying F4OA to the cybersecurity domain. The tools support the application of the framework in three case studies where we were able to deliver several results to stakeholders. Throughout this chapter, we present these results, answering the research questions (8.1) and presenting the impact of the work based on the accepted publications, as well as contributions to the funding project and others in the research group (8.2). Ultimately, we will present future research directions based on these contributions (8.3).

8.1 Answers to Research Questions

Throughout this research, we have worked towards answers to the research questions associated with the objectives described in the first chapter.

8.1.1 First Objective Outcomes

Along Chapters 2 and 3, we answer the knowledge question associated with the first objective of this thesis: Along Chapters 2 and 3, we answer the knowledge question associated with the first objective of this thesis: To clarify and homogenize the meta-ontological elements that interfere in stakeholders' semantic awareness about a complex domain, as well as, that elements surrounding their common sense about these conceptualizations (and ontologies).

KQ1: *How to conceptually characterize ontologies?*

The evolution of conceptual modeling concerning ontologies as tools has allowed the modeling process to move from representations based on neutral language patterns to conceptualizations with a philosophical foundation. The notion of *Ontological Level* [76, 77] brought this contribution, just as *Ontological Commitment* [79] added context to the modeling process. From this base, in Chapter 2, we explore how characteristics and meta-characteristics of ontologies affect the knowledge representation. We did this using an approach from the perspective of the ontology engineering process. Already, to treat the characteristics and meta-characteristics related to the knowledge domains to be represented, in Chapter 3, we explore how the characteristics and meta-characteristics of the knowledge to be represented affect the

representation and consequently, its processing. Using a TLR and with the support of a team of experts, we elicit the problems that complex domains, such as the cybersecurity domain, present when they need to be represented through ontologies. The team of experts who worked on this project has specialized training and extensive practical experience, as they are research professionals from a large consulting organization, working specifically in the organization's cybersecurity research laboratory. We answer this question of competence by bringing to light both perspectives, domain and ontological.

8.1.2 Second Objective Outcomes

From the baseline established in Chapter 4, we developed a solution proposal throughout Part III targeting to answer the questions associated with the second objective of this research: To provide a clear baseline for characterizing and compare ontologies to facilitate the interoperability and reusability of ontologies through their *Ontological Analysis*.

RQ1: *Which (meta)characteristics help it conceptually characterize ontologies?*

To bring domain and ontological perspectives together and to answer this question, we look at the state of the art of what work exists that focuses on the classification and characterization of ontologies as tools for ensuring FAIR data, which in turn must also be FAIR. In section 4.1, we focus on the characteristics and meta-characteristics that make FAIRness data by introducing and exploring the FAIR Principles approach. In section 4.2 we focus on the characteristics and meta-characteristics of ontologies based on established classification criteria. Thus, we select the better and more comprehensive criteria to contrast the domain and ontological perspectives.

RQ2: *How do managing these (meta)characteristics aid ontological analysis to provide FAIRness?*

Ontologies (artifacts) are modeling tools with wide applications, fundamental in clarifying concepts via ontological analyses and conceptualizing real-world domains as close to reality as possible. In this sense, using this tool to clarify your characteristics and meta-characteristics is completely reasonable. We then proposed O4OA, a meta-ontology, i.e., an ontology that clarifies what ontologies are and, at the same time, guarantees FAIRness to the conceptualization contained therein. Chapter 5 presents O4OA.

RQ3 *How should a framework be a clear and reproducible basis capable of favoring semantic agreement meeting stakeholders' perspectives?*

From the ontological basis of O4OA and using the data about cybersecurity ontologies and their sources of information, we proposed an O4OA implementation solution and an ontological analysis service provider prototype to deal with the domain and ontological perspectives. These components compose the Framework for Ontological Characterization (F4OC). In Section 6.1, we present the F4OC; then, we start with the state-of-the-art step, which aims to support the ontological analysis processes (Section 6.2). In Section 6.3, we describe the F4OA treatment applicable to the domain perspective, while in Section 6.4, we describe the treatment applicable to the ontological perspective. Thus, we answer this question and present the framework components and all the required steps to perform ontological analysis.

8.1.3 Third Objective Outcomes

Chapter 7 is dedicated to answering the research questions associated with the third objective of this thesis by presenting the case studies we conducted and stakeholders' impressions about it.

RQ4: *Do the stakeholders believe that our framework is useful to facilitate the semantic agreement between them?*

To answer this research question, we present how data relating to cybersecurity ontologies (its characteristics and meta-characteristics) and the information sources of these conceptualizations were collected, recorded, and processed. After performing the framework required steps, we selected two critical concepts – *Risk* and *Vulnerability* – within the cybersecurity domain as initial case studies for validating the proposed framework, F4OC, and the ontology that supports it, O4OA. These concepts frequently appear in cybersecurity ontologies, and despite this, they are still subject to controversies and misunderstandings. Subsection 7.3 shows the details. In addition, we conducted an additional case study. Cybersecurity specialists from the organization participating in the project asked us to analyze the BRON ontology network to identify its effectiveness for a possible integration solution with source code tracking system ontologies. After reviewing the results, stakeholders found the framework helpful in the conceptualization clarification, showing the context and the ontological level in which the involved concepts must be treated. Indeed, they invited us to proceed with

another ontology network proposal analysis for future comparison with BRON (work in progress). Subsection 7.3.3 details the BRON case study.

8.2 Thesis Impact

This research has been validated by publishing the results in relevant international forums. Apart from this, some academic papers were developed as complementary works and participation in research projects. This section summarizes these contributions and provides some general considerations about the research.

8.2.1 Publications

The outcomes obtained during this research have been published and presented in major forums in the field of conceptual modeling and information systems, all with international impact:

Martins, B. F. *and Guizzardi, R. and Reyes Román, J. F. and Hadad, M. and Pastor, O.; The Ontology for Conceptual Characterization of Ontologies. In: Conceptual Modeling: 42th International Conference (ER), Lisbon, Portugal, November 06 – 09, 2018, Proceedings 42. Ed. by João Paulo A. Almeida et al. Cham: Springer Nature Switzerland, 2023, pp. 105 – 124. ISBN: 978-3-031-47262-6. DOI:* `https://doi.org/10.1007/978-3-031-47262-6_6`.

Martins, B. F. *and Reyes Román, J. F. and Pastor, O. and Hadad, M.; Improving Conceptual Domain Characterization in Ontology Networks. In: Research Challenges in Information Science: 17th International Conference (RCIS), Corfu, Greece, May 23 – 26, 2023, Proceedings. Springer International Publishing, 2023, pp. 187 – 202. ISBN: 978-3-031-33080-3_12. DOI:* `http://dx.doi.org/10.1007/978-3-031-33080-3_12`.

Martins, B. F. *and Serrano Gil, L. J. and Reyes Román, J. F. and Panach, J. I. and Pastor, O. and Hadad, M. and Rochwerger, B.; A framework for conceptual characterization of ontologies and its application in the cybersecurity domain. In: Software and Systems Modeling 21.4, 2022, pp. 1437 – 1464. DOI:* `https://doi.org/10.1007/s10270-022-01013-0`.

Martins, B. F. *and Serrano Gil, L. J. and Reyes Román, J. F. and Panach, J. I. and Pastor, O.; Towards the Consolidation of Cybersecurity Standardized Definitions: A Tool for Ontological Analysis. In: Proceedings of the XXIV Iberoamerican Conference on Software Engineering (CIbSE), San José, Costa Rica, 2021. Ed. by Tayana Conte et al. Curran Associates, 2021, pp. 290 – 303. ISBN: 978-1-7138-3944-6.*

Martins, B. F. *and Serrano Gil, L. J. and Reyes Román, J. F. and Panach, J. I. and Pastor, O. and Rochwerger, B.; Conceptual Characterization of Cybersecurity Ontologies. In: The 13th IFIP WG 8.1 working conference on the Practice of Enterprise Modelling (PoeM), Riga, Latvia, 2020, pp. 323–338. DOI:* `https://doi.org/10.1007/978-3-030-63479-7_22`.

Additional contributions:

Besides the leading publications, this research assisted as a contribution to works that were written in parallel, such as:

Martins, B. F. *and Guizzardi, R. and Pastor, O.; O4OA Specification. In: RiuNet Technical Report, Universidad Politecnica de Valencia, 2023,* `http://hdl.handle.net/10251/196721`.

Martins, B. F. *and Serrano Gil, L. J. and Reyes Román, J. F. and Panach, J. I. and Pastor, O.; Towards the Consolidation of Cybersecurity Standardized Definitions. In: RiuNet Technical Report, Universidad Politecnica de Valencia, 2021,* `http://hdl.handle.net/10251/163895`.

Serrano Gil, L. J. and ***Martins, B. F.*** *and Reyes Román, J. F. and Panach, J. I. and Pastor, O.; Una encuesta acerca de la Definición de Conceptos de Ciberseguridad. In: RiuNet Technical Report, Universidad Politecnica de Valencia, 2021,* `https://riunet.upv.es/handle/10251/174756`.

Verdonck, M. and Gailly, F. and Pergl, R. and Guizzardi, G. and ***Martins, B. F.*** *and Pastor, O.; Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. In: Information Systems 81, 2019, pp. 92 – 103. DOI:* `http://dx.doi.org/10.1016/j.is.2018.11.009`.

Martins, B. F.*; The OntoOO-Method: An Ontology-Driven Conceptual Modeling Approach for Evolving the OO-Method. In: Conceptual Modeling: 38th International Conference (ER), Salvador, Bahia, Brazil, November 04 – 07, 2019. In: Guizzardi, G., Gailly, F., Suzana Pitangueira Maciel, R. (eds) Advances in Conceptual Modeling. Lecture Notes in Computer Science, vol 11787. Springer, Cham. DOI:* `https://doi.org/10.1007/978-3-030-34146-6_23`.

8.2.2 Academic Work

The research performed throughout this thesis work has accomplished its objectives. Additionally, it has enabled the student involved in the research to obtain a master's degree through her internship with our research team. The academic work produced as a result of the research has made valuable contributions towards the advancement of knowledge in the relevant field:

> *Information security assessment model for Bring Your Own Device in the South African healthcare sector. Master's Thesis in Informatics at Tshwane University of Technology. Cathy B. Moeketsi. Advisor: Prof. Adeyelure Tope Samuel D.Sc. Co-advisor: Prof. Mmatshuene Anna Segooa D.Sc.*

8.2.3 Industrial Work

The investigation carried out throughout this work also provided means for members of the professional team of the contributing company to initiate their solution at the industrial level. As a result of the research, this solution can bring valuable contributions to the advancement of knowledge in the company's various areas of activity, in addition to the Cybersecurity domain:

> *Klein, D. and Engelberg, G.; Cyber Attack Countermeasures Generation With LLMs & Knowledge Graphs. In Connections: a Neo4J virtual event – Neo4j Video Channel. Accenture Labs Israel. Live stream on December 12nd, 2023.*

8.2.4 Research Projects

In addition to the contributions mentioned above, I have worked on the following research projects:

> *Modelado Conceptual, Inteligencia Artificial, Ontologías y Ingeniería de Ontologías. L'Escola de Postgrau i Xarxa de Centres d'Investigació en Intelligència Artificial (ValGRAI). From December 25th, 2022 to January 25th, 2024. Ref. VALGRAI/22/1.*
>
> *Plataforma de Computación Intensiva Mediante Aceleradores Gráficos (GPUS) para su Aplicación en Medicina Personalizada. Generalitat Valenciana (IDIFEDER). From January 1st, 2019 to December 31th, 2019. Ref. IDIFEDER/2018/032.*
>
> *Um Método de Producción de Software Dirigido por Modelos para el Desarrollo de Aplicaciones Big Data. Agencia Estatal de Investigación. From January 1st, 2017 to September 30th, 2019. Ref. TIN2016-80811-P.*

8.2.5 General Considerations

Throughout this research, we adopted some strategies, such as focusing on two main groups of stakeholders (ontology engineers and domain specialists) and their technical perceptions. By avoiding delving into administrative aspects of the relationship between and within these groups, we could focus more calmly on the problems addressed in this thesis. Therefore, we focused on the essential elements of ontologies, searching in the literature which of their elements (meta-characteristics) are relevant for their classification.

However, throughout this study, we found that it is also necessary to consider the essential elements of the domain represented in these ontologies. The FAIR Principles were added to our work because we considered the domain represented, its information, and its representation in the form of ontologies and ontology networks. The observation of both stakeholders' perspectives was fundamental in this regard. We also opted for a non-domain-specific approach, allowing both the framework (F4OC) and its base ontology (O4OA) to be applied to any knowledge domain, despite our primary research focusing

on the cybersecurity domain. Furthermore, we observe that the results provided through the application of the framework in this domain can be used to provide the ontological support necessary for the implementation and management of modern architectures based on knowledge graphs. We implemented a prototype using well-known data and information sources in this domain to illustrate our approach. In summary, our work already presents promising results in the industrial scope and indicates that we made the right choices during this research.

8.3 Future Work

We formulate future research directions for O4OA and F4OC proposals according to the results we achieve, validation performed, and considering the research team's interests in continuing the project.

In this research, we applied the F4OC in case studies, proving its reliability and usefulness. The work focuses on a qualitative approach strongly linked to the FAIR Principles and ontology engineering best practices; therefore, quantitative measurements are outside the scope. However, we recognize this is an important step forward for future improvements. Indeed, the semantic problems and misunderstanding in representing conceptualizations as ontologies can mean tremendous losses for organizations and provide measurements about an opportunity for quantification. Besides, as F4OC is reproducible and traceable, it provides the required baseline for precise qualitative measurements. This opens opportunities for future research in qualitative assessments under the umbrella of the ontology engineering process and FAIRness.

The F4OC proposal was validated by applying it to answer the research questions considering the cybersecurity domain, which is a domain of interest of the project stakeholders. In addition, based on the outcomes, the project financier intends to use the approach in other domains of interest, for instance, Car Insurance and Business Mapping. Also, our research group is interested in applying the hanging framework in their Genomic Domain research.

The framework works by building a cloud of concepts that so far are based on syntactic elements (terms and their definitions) and, despite this, has already presented relevant results. However, it is in the interest of the funding organization to extend the project to add AI to the prototype. Given the recent progress in LLMs, it is proposed to use the syntactic elements already in use in conjunction with these models. This research direction seems

promising and could bring F4OA to a semantic level (or close to it) and thus further support the ontological analysis process of vast domains.

To summarize, the proposed framework and its ontological basis represent a step toward ontology engineering. However, we know that more research needs to be done, and more cases need to be analyzed to verify the real potential of the solution. We will continue this work by developing new research projects and academic papers in the hope that we can continue to show the positive results we are already achieving.

Bibliography

[1] Martin Abadi et al. "TensorFlow: A system for large-scale machine learning". In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*. 2016, pp. 265–283 (cit. on p. 3).

[2] Manfred Aben. "Formally specifying reusable knowledge model components". In: *Knowledge Acquisition* 5.2 (1993), pp. 119–141 (cit. on p. 39).

[3] J. L. Ackrill. "Aristotle's Categories". In: *Aristotle: A Collection of Critical Essays*. Ed. by J. M. E. Moravcsik. London: Palgrave Macmillan UK, 1967, pp. 90–124. DOI: 10.1007/978-1-349-15267-4_6 (cit. on p. 3).

[4] Vivek Agrawal. "Towards the Ontology of ISO/IEC 27005:2011 Risk Management Standard". In: *HAISA*. 2016 (cit. on pp. 82, 143).

[5] Antognoni Albuquerque and Giancarlo Guizzardi. "An ontological foundation for conceptual modeling datatypes based on semantic reference spaces". In: *IEEE 7th International Conference on Research Challenges in Information Science (RCIS)*. IEEE. 2013, pp. 1–12 (cit. on p. 151).

[6] J. P. A. Almeida et al. *gUFO: a lightweight implementation of the Unified Foundational Ontology (UFO)*. 2019 (cit. on pp. 19, 151).

[7] Ricardo de Almeida Falbo. "SABiO: Systematic Approach for Building Ontologies." In: *Onto. Com/odise@ Fois*. 2014 (cit. on pp. 16, 17, 23, 29, 38, 50, 68, 77).

[8] Ammar Ammar et al. "A semi-automated workflow for FAIR maturity indicators in the life sciences". In: *Nanomaterials* 10.10 (2020), p. 2068 (cit. on p. 35).

[9] Paul A. Grassi And, Michael E. Garcia And, and James L. Fenton. *Digital Identity Guidelines*. Tech. rep. NIST, 2017. DOI: `https://doi.org/10.6028/NIST.SP.800-63-3` (cit. on p. 150).

[10] Rob Atkinson et al. *Semantic Sensor Network Ontology*. Tech. rep. OGC 16-079. World Wide Web Consortium, 2017 (cit. on p. 35).

[11] Franz Baader, Ian Horrocks, and Ulrike Sattler. "Description logics". In: *Foundations of Artificial Intelligence* 3 (2008), pp. 135–179 (cit. on p. 20).

[12] Radu F. Babiceanu and Remzi Seker. "Cybersecurity and resilience modelling for software-defined networks-based manufacturing applications". In: *Studies in Computational Intelligence* 694 (2017), pp. 167–176. ISSN: 1860949X. DOI: `10.1007/978-3-319-51100-9_15` (cit. on p. 142).

[13] Hamza Baqa et al. *Semantic IoT Solutions-A Developer Perspective*. 2019 (cit. on p. 35).

[14] Pedro Paulo F. Barcelos et al. "A FAIR model catalog for ontology-driven conceptual modeling research". In: *International Conference on Conceptual Modeling*. Springer. 2022, pp. 3–17 (cit. on pp. 23, 44).

[15] Sean Barnum. "Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIX ™)". In: *Mitre Corporation* 11.Version 1.1, Revision 1 (2014), pp. 1–22 (cit. on p. 151).

[16] Carlo Batini and Monica Scannapieco. "Data and information quality". In: *Cham, Switzerland: Springer International Publishing* (2016) (cit. on p. 63).

[17] Martin Bauer et al. "Towards semantic interoperability standards based on ontologies". In: *AIOTI White paper* (2019) (cit. on p. 36).

[18] Noam Ben-Asher et al. "Ontology-based Adaptive Systems of Cyber Defense". In: *STIDS*. 2015, pp. 34–41 (cit. on pp. 29, 142).

[19] Alessander Botti Benevides and Giancarlo Guizzardi. "A model-based tool for conceptual modeling and domain ontology engineering in OntoUML". In: *Enterprise Information Systems* (2009), pp. 528–538 (cit. on pp. 38, 51).

[20] Alessander Botti Benevides et al. "Validating Modal Aspects of OntoUML Conceptual Models Using Automatically Generated Visual World Structures." In: *J. Univers. Comput. Sci.* 16.20 (2010), pp. 2904–2933 (cit. on pp. 51, 77).

[21] Sandra Bergner and Ulrike Lechner. "Cybersecurity Ontology for Critical Infrastructures." In: *KEOD*. 2017, pp. 80–85 (cit. on p. 142).

[22] Anna Bernasconi et al. "Ontological representation of FAIR principles: A blueprint for FAIRer data sources". In: *International Conference on Advanced Information Systems Engineering*. Springer. 2023, pp. 261–277 (cit. on pp. 23, 37, 62, 67).

[23] CVE Editorial Board. *Common Vulnerabilities and Exposures – CVE downloads data last generated: 2020-06-23*. `https://cve.mitre.org/data/downloads/index.html`. Online; accessed 23-Jun-2020. 2006 (cit. on p. 150).

[24] Martin Boeckhout, Gerhard A. Zielhuis, and Annelien L. Bredenoord. "The FAIR guiding principles for data stewardship: fair enough?" In: *European Journal of Human Genetics* 26.7 (2018), pp. 931–936 (cit. on pp. 23, 34).

[25] H. Booth and C. Turner. "Vulnerability Description Ontology (VDO)". In: *A Framework for Characterizing Vulnerabilities. NIST* (2016) (cit. on p. 143).

[26] Fabiano Borges Ruy et al. "SEON: A software engineering ontology network". In: *Knowledge Engineering and Knowledge Management: 20th International Conference, EKAW 2016, Bologna, Italy, November 19-23, 2016, Proceedings 20*. Springer. 2016, pp. 527–542 (cit. on p. 85).

[27] Stefano Borgo and Claudio Masolo. "Ontological Foundations of DOLCE". In: *Theory and Applications of Ontology: Computer Applications*. Dordrecht: Springer Netherlands, 2010, pp. 279–295. ISBN: 978-90-481-8847-5. DOI: 10.1007/978-90-481-8847-5_13 (cit. on pp. 18, 152).

[28] Stefano Borgo and Laure Vieu. "Artefacts in formal ontology". In: *Philosophy of technology and engineering sciences*. Elsevier, 2009, pp. 273–307 (cit. on p. 19).

[29] Pim Borst, Hans Akkermans, and Jan Top. "Engineering ontologies". In: *International journal of human-computer studies* 46.2-3 (1997), pp. 365–406 (cit. on p. 39).

[30] Willem Nico Borst. "Construction of engineering ontologies for knowledge sharing and reuse." In: (1999) (cit. on pp. 4, 151).

[31] WN Borst. "Construction of engineering ontologies (PhD thesis, University of Twente, Enschede)". In: (1997) (cit. on p. 39).

[32] Andrei Brazhuk. "Towards automation of threat modeling based on a semantic model of attack patterns and weaknesses". In: *arXiv preprint arXiv:2112.04231* (2021) (cit. on p. 30).

[33] Andrei Brazhuk. "Towards automation of threat modeling based on a semantic model of attack patterns and weaknesses". In: *CoRR* abs/2112.04231 (2021) (cit. on p. 88).

[34] Bret Jordan, Rich Piazza, and Trey Darley, ed. *OASIS – STIX™ Version 2.1*. OASIS Committee Specification 01, 2020 (cit. on p. 150).

[35] Nicholas Carlini et al. "Extracting training data from large language models". In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021, pp. 2633–2650 (cit. on p. 3).

[36] Victorio A. de Carvalho, João Paulo A. Almeida, and Giancarlo Guizzardi. "Using reference domain ontologies to define the real-world semantics of domain-specific languages". In: *Advanced Information Systems Engineering: 26th International Conference, CAiSE 2014, Thessaloniki, Greece, June 16-20, 2014. Proceedings 26*. Springer. 2014, pp. 488–502 (cit. on p. 57).

[37] CCDB, ed. *CC and CEM addenda Exact Conformance , Selection-Based SFRs , Optional SFRs*. Vol. V0.5. Geneva – Switzerland: CCDB, May 2017 (cit. on p. 149).

[38] CCITT & ITU-T, ed. *DATA COMMUNICATION NETWORKS: OPEN SYSTEMS INTERCONNECTION (OSI); SECURITY, STRUCTURE AND APPLICATION – SECURITY ARCHITECTURE FOR OPEN SYSTEMS INTERCONNECTION FOR CCITT APPLICATIONS*. Geneva – Switzerland: CCITT & ITU-T, 1991 (cit. on p. 150).

[39] CCMB, ed. *Common Criteria for Information Technology Security Evaluation Part 1 : Introduction and general model*. Revision 5. Vol. Version3.1. CCDB, 2017 (cit. on p. 149).

[40] CCMB, ed. *Common Criteria for Information Technology Security Evaluation Part 2 : Security functional components*. Revision 5. Vol. Version3.1. CCDB, 2017 (cit. on p. 149).

[41] CCMB, ed. *Common Criteria for Information Technology Security Evaluation Part 3 : Security assurance components*. Revision 5. Vol. Version3.1. CCDB, 2017 (cit. on p. 149).

[42] CCMB, ed. *Common Methodology for Information Technology Security Evaluation Evaluation methodology*. Revision 5. Vol. Version3.1. CCDB, 2017 (cit. on p. 149).

[43] CCRA. *Common Criteria Portal*. `https://www.commoncriteriaportal.org/cc/`. Online; accessed 23-Jun-2020. 2017 (cit. on p. 150).

[44] Mark Chaplin et al. "The 2011 Standard of Good Practice Principal authors Review and quality assurance". In: June (2011) (cit. on p. 151).

[45] European Commission, Directorate-General for Research, and Innovation. *Turning FAIR into reality – Final report and action plan from the European Commission expert group on FAIR data*. Publications Office, 2018. DOI: doi/10.2777/1524 (cit. on p. 36).

[46] MITRE Corporation. *Digital Artifact Ontology*. https://d3fend.mitre.org/dao/. Online; accessed 23-Jun-2020 (cit. on p. 150).

[47] MITRE Corporation. *The Common Weakness Enumeration (CWE) Initiative*. http://cwe.mitre.org/. Online; accessed 20-Jun-2020 (cit. on p. 150).

[48] MITRE Corporation. *Trusted Automated eXchange of Indicator Information — TAXII™ Enabling Cyber Threat Information Exchange*. Tech. rep. MITRE Corporation (cit. on p. 151).

[49] Dolors Costal, Cristina Gómez, and Giancarlo Guizzardi. "Formal Semantics and Ontological Analysis for Understanding Subsetting, Specialization and Redefinition of Associations in UML". In: *Conceptual Modeling – ER 2011*. Ed. by Manfred Jeusfeld, Lois Delcambre, and Tok-Wang Ling. Springer, 2011, pp. 189–203. ISBN: 978-3-642-24606-7 (cit. on p. 151).

[50] G. Cota. "Best practices for implementing fair vocabularies and ontologies on the web". In: *Applications and practices in ontology design, extraction, and reasoning* 49 (2020), p. 39 (cit. on p. 23).

[51] Kenneth James Williams Craik. *The nature of explanation*. Vol. 445. CUP Archive, 1967 (cit. on p. 4).

[52] Fabiano Dalpiaz, Xavier Franch, and Jennifer Horkoff. "istar 2.0 language guide". In: *arXiv preprint arXiv:1605.07767* (2016) (cit. on p. 16).

[53] Laura Daniele et al. *Smart Applications REFerence Ontology (SAREF)*. Last accessed August 2023. 2019 (cit. on p. 35).

[54] Mathieu d'Aquin et al. *Watson: A gateway for the semantic web.* 2007 (cit. on p. 37).

[55] Vladimir Dimitrov. "Chapter Two: CVE Annotation". In: *Information Security in Education and Practice* (2020), pp. 19–28 (cit. on p. 85).

[56] Bruno Borlini Duarte et al. "Towards an ontology of software defects, errors and failures". In: *International Conference on Conceptual Modeling.* Springer. 2018, pp. 349–362 (cit. on pp. 85, 86, 152).

[57] Sam Adam Elnagdy, Meikang Qiu, and Keke Gai. "Cyber incident classifications using ontology-based knowledge representation for cybersecurity insurance in financial industry". In: *2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud).* IEEE. 2016, pp. 301–306 (cit. on p. 143).

[58] Dieter Fensel. "Ontologies". In: *Ontologies.* Springer, 2001, pp. 11–18 (cit. on pp. 23, 40).

[59] James L. Fenton et al. *Digital Identity Guidelines: Authentication and Lifecycle Management.* Tech. rep. NIST, 2017. DOI: `https://doi.org/10.6028/NIST.SP.800-63b` (cit. on p. 150).

[60] James L. Fenton et al. *Digital Identity Guidelines: Enrollment and Identity Proofing.* Tech. rep. NIST, 2017. DOI: `https://doi.org/10.6028/NIST.SP.800-63a` (cit. on p. 150).

[61] Mariano Fernández-López, Asunción Gómez-Pérez, and Natalia Juristo. "Methontology: from ontological art towards ontological engineering". In: (1997) (cit. on pp. 23, 29).

[62] Roberta Ferrario, Nicola Guarino, and Meritxell Fernández-Barrera. "Towards an ontological foundation for services science: The legal perspective". In: *Approaches to Legal Ontologies: Theories, Domains, Methodologies.* Springer, 2010, pp. 235–258 (cit. on p. 29).

[63] Claudenir M. Fonseca et al. "Incorporating Types of Types in Ontology-Driven Conceptual Modeling". In: *International Conference*

on Conceptual Modeling. Springer. 2022, pp. 18–34 (cit. on pp. 7, 54, 61).

[64] Claudenir M. Fonseca et al. "Relations in ontology-driven conceptual modeling". In: *Conceptual Modeling: 38th International Conference, ER 2019, Salvador, Brazil, November 4–7, 2019, Proceedings 38*. Springer. 2019, pp. 28–42 (cit. on pp. 54, 55, 57, 152).

[65] H. Gasmi, J. Laval, and A. Bouras. "Cold-start cybersecurity ontology population using information extraction with LSTM". In: *2019 International Conference on Cyber Security for Emerging Technologies (CSET)*. 2019, pp. 1–6. DOI: 10.1109/CSET.2019.8904905 (cit. on p. 142).

[66] Fausto Giunchiglia and Ilya Zaihrayeu. *Lightweight ontologies*. Tech. rep. University of Trento, 2007 (cit. on pp. 20, 21, 44, 52, 53, 70, 151, 184).

[67] *Glosario de términos de ciberseguridad – Una guía de aproximación para el empresario*. Instituto Nacional de Ciberseguridad – Spanish National Cybersecurity Institute, 2017 (cit. on p. 150).

[68] Asunción Gómez-Pérez and Oscar Corcho. "Ontology languages for the semantic web". In: *IEEE Intelligent systems* 17.1 (2002), pp. 54–60 (cit. on pp. 41, 44, 52, 53, 71, 152).

[69] Asuncion Gomez-Perez, Mariano Fernández-López, and Oscar Corcho. *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web*. Springer Verlag, Jan. 2004 (cit. on p. 23).

[70] Asuncion Gomez-Perez, Mariano Fernández-López, and Oscar Corcho. *Ontological Engineering: With Examples from the Areas of Knowledge Management, E-Commerce and the Semantic Web*. Jan. 2004 (cit. on pp. 41–43).

[71] André Grégio et al. "Ontology for malware behavior: A core model proposal". In: *2014 IEEE 23rd International WETICE Conference*. IEEE. 2014, pp. 453–458 (cit. on p. 142).

[72] Thomas R. Gruber. "A translation approach to portable ontology specifications". In: *Knowledge acquisition* 5.2 (1993), pp. 199–220 (cit. on pp. 4, 151).

[73] Michael Gruninger. "Designing and evaluating generic ontologies". In: *12th European Conference of Artificial Intelligence*. Vol. 1. Citeseer. 1996, pp. 53–64 (cit. on pp. 22, 50, 68).

[74] Michael Gruninger. "Methodology for the Design and Evaluation of Ontologies". In: *International Joint Conference on Artificial Intelligence*. 1995 (cit. on pp. 22, 50, 68).

[75] N. Guarino. "Formal Ontology in Information Systems". In: *Proceedings of the first international conference (FOIS'98)*. Trento, Italy: IOS Press, June 1998, pp. 19–28. ISBN: 9051993994 (cit. on pp. 4, 21, 22, 30, 38, 40, 52, 69, 141, 151).

[76] Nicola Guarino. "The ontological level". In: *Philosophy and the Cognitive Sciences* (1994) (cit. on pp. 5, 6, 15, 18, 23, 38, 56, 68, 96, 142, 143, 184).

[77] Nicola Guarino. "The ontological level: Revisiting 30 years of knowledge representation". In: *Conceptual Modeling: Foundations and applications* (2009), pp. 52–67 (cit. on pp. 5, 6, 15, 18, 23, 38, 96).

[78] Nicola Guarino, Riccardo Baratella, and Giancarlo Guizzardi. "Events, their names, and their synchronic structure". In: *Applied Ontology* 17 (Jan. 2022), pp. 1–35. DOI: 10.3233/AO-220261 (cit. on p. 151).

[79] Nicola Guarino, Massimiliano Carrara, and Pierdaniele Giaretta. "Formalizing ontological commitment". In: *AAAI*. Vol. 94. 1994, pp. 560–567 (cit. on pp. 5, 17, 21, 23, 96).

[80] Nicola Guarino and Giancarlo Guizzardi. "Relationships and events: towards a general theory of reification and truthmaking". In: *Conference of the Italian Association for Artificial Intelligence*. Springer. 2016, pp. 237–249 (cit. on p. 152).

[81] Nicola Guarino and Giancarlo Guizzardi. ""We need to discuss the Relationship": Revisiting Relationships as Modeling Constructs". In: *International Conference on Advanced Information Systems Engineering*. Springer. 2015, pp. 279–294 (cit. on p. 152).

[82] Nicola Guarino, Giancarlo Guizzardi, and John Mylopoulos. "On the philosophical foundations of conceptual models". In: *Information Modelling and Knowledge Bases* 31.321 (2020), p. 1 (cit. on p. 3).

[83] Nicola Guarino, Daniel Oberle, and Steffen Staab. "What is an ontology?" In: *Handbook on ontologies* (2009), pp. 1–17 (cit. on pp. 3, 4).

[84] Nicola Guarino and Roberto Poli. "The role of formal ontology in the information technnology". In: *International journal of human-computer studies* 43.5-6 (1995), pp. 623–965 (cit. on pp. 6, 38).

[85] Nicola Guarino, Tiago Prince Sales, and Giancarlo Guizzardi. "Reification and Truthmaking Patterns". In: *International Conference on Conceptual Modeling*. Springer. 2018, pp. 151–165 (cit. on p. 152).

[86] *GUÍA DE SEGURIDAD (CCN-STIC-401) GLOSARIO Y ABREVIATURAS*. 2015 (cit. on p. 150).

[87] G. Guizzardi. "The Role of Foundational Ontology for Conceptual Modeling and Domain Ontology Representation, Keynote Paper". In: *7th International Baltic Conference on Databases and Information Systems (DB&IS), Vilnius, IEEE Press*. 2006 (cit. on pp. 29, 59).

[88] G. Guizzardi, C. Masolo, and S. Borgo. "In the Defense of a Trope-Based Ontology for Conceptual Modeling: An Example with the Foundations of Attributes, Weak Entities and Datatypes, 25th Intl". In: *Conf. on Conceptual Modeling, Berlin*. 2006 (cit. on pp. 7, 151).

[89] Giancarlo Guizzardi. "Agent roles, qua individuals and the counting problem". In: *International Workshop on Software Engineering for Large-Scale Multi-agent Systems*. Springer. 2005, pp. 143–160 (cit. on p. 18).

[90] Giancarlo Guizzardi. "Agent Roles, Qua Individuals and the Counting Problem". In: *Software Engineering for Multi-Agent Systems IV*. Ed. by Alessandro Garcia et al. Springer, 2006, pp. 143–160. ISBN: 978-3-540-33583-2 (cit. on p. 151).

[91] Giancarlo Guizzardi. "Logical, ontological and cognitive aspects of object types and cross-world identity with applications to the theory of conceptual spaces". In: *Applications of Conceptual Spaces*. Springer, 2015, pp. 165–186 (cit. on p. 151).

[92] Giancarlo Guizzardi. "Modal Aspects of Object Types and Part-Whole Relations and the de re/de dicto Distinction". In: *International Conference on Advanced Information Systems Engineering*. Springer. 2007, pp. 5–20 (cit. on pp. 61, 151).

[93] Giancarlo Guizzardi. "On ontology, ontologies, conceptualizations, modeling languages, and (meta) models". In: *Frontiers in artificial intelligence and applications* 155 (2007), p. 18 (cit. on pp. 5, 15, 18, 21, 23, 29, 39, 41, 52, 55–57, 68, 69, 141, 151).

[94] Giancarlo Guizzardi. "On the Representation of Quantities and their Parts in Conceptual Modeling". In: *FOIS*. 2010, pp. 103–116 (cit. on p. 152).

[95] Giancarlo Guizzardi. "Ontological Foundations for Conceptual Part-Whole Relations: The Case of Collectives and Their Parts". In: *Advanced Information Systems Engineering*. Ed. by Haralambos Mouratidis and Colette Rolland. Springer, 2011, pp. 138–153. ISBN: 978-3-642-21640-4 (cit. on p. 152).

[96] Giancarlo Guizzardi. *Ontological Foundations for Structural Conceptual Models*. CTIT, Centre for Telematics and Information Technology, 2005 (cit. on pp. 18, 38, 51, 86, 152).

[97] Giancarlo Guizzardi. "Ontological patterns, anti-patterns and pattern languages for next-generation conceptual modeling". In: *Conceptual Modeling: 33rd International Conference, ER 2014, Atlanta, GA, USA, October 27-29, 2014. Proceedings 33*. Springer. 2014, pp. 13–27 (cit. on pp. 52, 54–56).

[98] Giancarlo Guizzardi. "Ontology-based evaluation and design of visual conceptual modeling languages". In: *Domain Engineering: Product Lines, Languages, and Conceptual Models* (2013), pp. 317–347 (cit. on pp. 15, 39).

[99] Giancarlo Guizzardi. "Ontology, ontologies and the "I" of FAIR". In: *Data Intelligence* 2 (2020), pp. 181–191. DOI: 10.1162/dint_a_00040 (cit. on pp. 6, 23, 36).

[100] Giancarlo Guizzardi. "The problem of transitivity of part-whole relations in conceptual modeling revisited". In: *International Conference on Advanced Information Systems Engineering*. Springer. 2009, pp. 94–109 (cit. on pp. 61, 152).

[101] Giancarlo Guizzardi, Ricardo de Almeida Falbo, and Renata SS Guizzardi. "Grounding Software Domain Ontologies in the Unified Foundational Ontology (UFO): The case of the ODE Software Process Ontology." In: *CIbSE*. 2008, pp. 127–140 (cit. on pp. 5, 55, 86, 87, 151).

[102] Giancarlo Guizzardi, Luís Ferreira Pires, and Marten Van Sinderen. "An ontology-based approach for evaluating the domain appropriateness and comprehensibility appropriateness of modeling languages". In: *International Conference on Model Driven Engineering Languages and Systems*. Springer. 2005, pp. 691–705 (cit. on pp. 5, 15, 18, 20–23, 39, 54).

[103] Giancarlo Guizzardi and Tiago Prince Sales. "Detection, simulation and elimination of semantic anti-patterns in ontology-driven conceptual models". In: *Conceptual Modeling: 33rd International Conference, ER 2014, Atlanta, GA, USA, October 27-29, 2014. Proceedings 33*. Springer. 2014, pp. 363–376 (cit. on pp. 22, 54).

[104] Giancarlo Guizzardi and Gerd Wagner. "Conceptual simulation modeling with Onto-UML". In: *Proceedings of the Winter Simulation Conference*. Winter Simulation Conference. 2012, p. 5 (cit. on p. 151).

[105] Giancarlo Guizzardi and Gerd Wagner. "What's in a relationship: an ontological analysis". In: *International Conference on Conceptual Modeling*. Springer. 2008, pp. 83–97 (cit. on pp. 55, 152).

[106] Giancarlo Guizzardi, Gerd Wagner, and Heinrich Herre. "On the Foundations of UML as an Ontology Representation Language". In: *Engineering Knowledge in the Age of the Semantic Web*. Ed. by Enrico Motta et al. Springer, 2004, pp. 47–62. ISBN: 978-3-540-30202-5 (cit. on p. 151).

[107] Giancarlo Guizzardi and Veruska Zamborlini. "Using a trope-based foundational ontology for bridging different areas of concern in ontology-driven conceptual modeling". In: *Science of Computer Programming* 96 (2014), pp. 417–443 (cit. on p. 152).

[108] Giancarlo Guizzardi et al. "An ontologically well-founded profile for UML conceptual models". In: *International Conference on Advanced Information Systems Engineering*. Springer. 2004, pp. 112–126 (cit. on p. 151).

[109] Giancarlo Guizzardi et al. "Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story". In: *Applied ontology* 10.3-4 (2015), pp. 259–271 (cit. on p. 152).

[110] Giancarlo Guizzardi et al. "Towards ontological foundations for the conceptual modeling of events". In: *International Conference on Conceptual Modeling*. Springer. 2013, pp. 327–341 (cit. on p. 152).

[111] Renata S. S. Guizzardi and Giancarlo Guizzardi. "Ontology-based transformation framework from TROPOS to AORML". In: *Social modeling for requirements engineering* (2010), pp. 547–570 (cit. on p. 152).

[112] Jens Hartmann et al. "OMV–ontology metadata vocabulary". In: *ISWC*. Vol. 3729. 2005 (cit. on p. 37).

[113] VAN GERTJAN Heijst. "The Role of Ontologies in Knowledge Engineering". In: *Dr. Thesis, University of Amsterdam* (1995) (cit. on p. 39).

[114] HAAV Hele-Mai and LUBI Tanel-Lauri. "A survey of concept-based information retrieval tools on the web". In: *Proceedings of the 5thEast-European Conference AD BIS*. 2001, pp. 29–41 (cit. on pp. 23, 40).

[115] Erik Hemberg and Una-May O'Reilly. "Using a Collated Cybersecurity Dataset for Machine Learning and Artificial Intelligence". In: *arXiv preprint arXiv:2108.02618* (2021) (cit. on pp. 30, 88).

[116] Erik Hemberg et al. *Linking Threat Tactics, Techniques, and Patterns with Defensive Weaknesses, Vulnerabilities and Affected Platform Configurations for Cyber Hunting*. 2021. arXiv: 2010.00533 [cs.CR] (cit. on pp. 30, 88).

[117] Edmund Husserl. *Formal and transcendental logic*. Springer Science & Business Media, 1969 (cit. on pp. 38, 151).

[118] Chadni Islam, M. Ali Babar, and Surya Nepal. *Automated Interpretation and Integration of Security Tools Using Semantic Knowledge*. Springer International Publishing, 2019, pp. 513–528. ISBN: 978-3-030-21290-2. DOI: 10.1007/978-3-030-21290-2_32 (cit. on p. 143).

[119] ISO Central Secretary. *Information technology Security techniques — Evaluation criteria for IT security — Part 3: Security assurance components*. en. Standard ISO/IEC 154083:2008. Geneva: International Organization for Standardization, 2008 (cit. on p. 150).

[120] ISO Central Secretary. *Information technology — Security techniques — Code of practice for information security controls*. en. Standard ISO/IEC 27002:2013. Geneva: International Organization for Standardization, 2013 (cit. on p. 150).

[121] ISO Central Secretary. *Information technology — Security techniques — Evaluation criteria for IT security — Part 2: Security functional components*. en. Standard ISO/IEC 154082:2008. Geneva: International Organization for Standardization, 2008 (cit. on p. 150).

[122] ISO Central Secretary. *Information technology — Security techniques — Guidelines for cybersecurity*. en. Standard ISO/IEC 27032:2012. Geneva: International Organization for Standardization, 2012 (cit. on pp. 25, 150).

[123] ISO Central Secretary. *Information technology — Security techniques — Information security management systems — Overview and vocabulary*.

en. Standard ISO/IEC 27000:2018-02. Geneva: International Organization for Standardization, 2018 (cit. on p. 150).

[124] ISO Central Secretary. *Information technology — Security techniques — Information security risk management*. en. Standard ISO/IEC 27005:2018. Geneva: International Organization for Standardization, 2018 (cit. on p. 150).

[125] ISO Central Secretary. *Irmation technology – Security techniques – Evaluation criteria for IT – Part 1: Introduction and general model Technologies*. en. Standard ISO/IEC 154081:2009. Geneva: International Organization for Standardization, 2009 (cit. on p. 150).

[126] ISO Central Secretary. *SEVOCAB*. en. Standard **ISO/IEC/IEEE 24765:2010**. Geneva: International Organization for Standardization, 2010 (cit. on p. 150).

[127] ITU-T, ed. *DATA NETWORKS AND OPEN SYSTEM COMMUNICATION SECURITY – INFORMATION TECHNOLOGY – OPEN SSTEMS INTERCONNECTION – SECURITY FRAMEWORKS FOR OPEN SYSTEMS: INTEGRITY FRAMEWORKS*. Vol. 11/95. Geneva – Switzerland: ITU-T, 1995 (cit. on p. 150).

[128] ITU-T, ed. *Data Networks and Open System Communications – Security – Information Technology – Open Systems Interconnection – Security Frameworks for Open Systems: Overview*. Vol. 11/95. ITU-T, 1996 (cit. on p. 150).

[129] ITU-T, ed. *DATA NETWORKS AND OPEN SYSTEM COMMUNICATIONS – SECURITY – INFORMATION TECHNOLOGY – OPEN SYSTEMS INTERCONNECTION – SECURITY FRAMEWORKS FOR OPEN SYSTEMS: SECURITY AUDIT AND ALLARMS FRAMEWORK*. Vol. 11/95. Geneva – Switzerland: ITU-T, 1996 (cit. on p. 150).

[130] ITU-T, ed. *DATA NETWORKS AND OPEN SYSTEM COMMUNICATIONS – SECURITY – INFORMATION TECHONOLOGY – OPEN SYSTEMS INTERCONNECTION – SECURITY FRAMEWORKS FOR OPEN SYSTEMS: AUTHENTICATION FRAMEWORK*. Vol. 04/95. Geneva – Switzerland: ITU-T, 1996 (cit. on p. 150).

[131] ITU-T, ed. *DATA NETWORKS AND OPEN SYSTEM COMMUNICATIONS – SECURITY – INFORMATION TECHONOLOGY – OPEN SYSTEMS INTERCONNECTION – SECURITY FRAMEWORKS FOR OPEN SYSTEMS: CONFIDENTIALITY FRAMEWORK*. Vol. 11/95. Geneva – Switzerland: ITU-T, 1995 (cit. on p. 150).

[132] ITU-T, ed. *DATA NETWORKS AND OPEN SYSTEM COMMUNICATIONS SECURITY – INFORMATION TECHNOLOGY – OPEN SYSTEMS INTERCONNECTION – SECURITY FRAMEWORKS FOR OPEN SYSTEMS : ACCESS CONTROL*. Vol. 11/95. Geneva – Switzerland: ITU-T, 1996 (cit. on p. 150).

[133] ITU-T, ed. *SERIES X: DATA NETWORKS AND OPEN SYSTEM COMMUNICATION – Security – Information techology – Open Systems Interconnection – Security Frameworks in open systems: Non-repudiation framework*. Vol. 10/96. ITU-T, 1997 (cit. on p. 150).

[134] ITU-T, ed. *SERIES X: DATA NETWORKS AND OPEN SYSTEM COMMUNICATIONS Security – Security architecture for systems providing end-to-end communications*. Vol. 10/2003. ITU-T, 2003 (cit. on p. 150).

[135] ITU-T, ed. *SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY – Cybersecurity information exchange – Overview of cybersecurity – Overview of cybersecurity information exchange*. 1.0. Vol. 04/2011. ITU-T, 2012 (cit. on p. 150).

[136] ITU-T, ed. *SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY Cyberspace security – Cybersecurity – Capabilities and their context scenarios for cybersecurity information sharing and exchange*. 1.0. Vol. 12/2010. ITU-T, 2010 (cit. on p. 150).

[137] ITU-T, ed. *SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY Cyberspace security – Cybersecurity – Design considerations for improved end-user perception of trustworthiness indicators*. 1.0. Vol. 03/2017. ITU-T, 2017 (cit. on p. 150).

[138] ITU-T, ed. *SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY Telecommunication security – Overview of cybersecurity*. Vol. 04/2008. ITU-T, 2008 (cit. on p. 150).

[139] Daniel Jackson. *Software Abstractions: logic, language, and analysis*. MIT press, 2012 (cit. on pp. 51, 77).

[140] Annika Jacobsen et al. "FAIR Principles: Interpretations and Implementation Considerations". In: *Data Intell* 2.1-2 (2020), pp. 10–29. DOI: 10.1162/dint_r_00024 (cit. on p. 23).

[141] Robert Jasper and Mike Uschold. "A framework for understanding and classifying ontology applications". In: *Proceedings 12th Int. Workshop on Knowledge Acquisition, Modelling, and Management KAW*. Vol. 99. Citeseer. 1999, pp. 16–21 (cit. on p. 41).

[142] Yan Jia et al. "A practical approach to constructing a knowledge graph for cybersecurity". In: *Engineering* 4.1 (2018), pp. 53–60 (cit. on p. 142).

[143] Philip Nicholas Johnson-Laird. *Mental models: Towards a cognitive science of language, inference, and consciousness*. 6. Harvard University Press, 1983 (cit. on p. 4).

[144] JOINT TASK FORCE. *Risk Management Framework for Information Systems and Organizations – A System Life Cycle Approach for Security and Privacy*. Tech. rep. NIST, 2018. DOI: https://doi.org/10.6028/NIST.SP.800-37r2 (cit. on pp. 25, 150).

[145] JOINT TASK FORCE TRANSFORMATION INITIATIVE. *Security and Privacy Controls for Federal Information Systems and Organizations Security and Privacy Controls for Federal Information Systems and Organizations*. Tech. rep. NIST, 2013. DOI: http://dx.doi.org/10.6028/NIST.SP.800-53r4 (cit. on p. 150).

[146] Clement Jonquet et al. "Harnessing the power of unified metadata in an ontology repository: the case of AgroPortal". In: *Journal on Data Semantics* 7.4 (2018), pp. 191–221 (cit. on pp. 6, 36, 45).

[147] Igor Jurisica, John Mylopoulos, and Eric Yu. "Using ontologies for knowledge management: An information systems perspective". In: *Proceedings of the Annual Meeting-American Society For Information Science*. Vol. 36. Information Today; 1998. 1999, pp. 482–496 (cit. on pp. 23, 41).

[148] Peter E. Kaloroumakis and Michael J. Smith. "Toward a knowledge graph of cybersecurity countermeasures". In: *Corporation, Editor* (2021) (cit. on pp. 88, 89, 151).

[149] Dongwoo Kang et al. "An ontology-based enterprise architecture". In: *Expert Systems with Applications* 37.2 (2010), pp. 1456–1464 (cit. on p. 26).

[150] Elmar Kiesling et al. *The SEPSES Knowledge Graph: An Integrated Resource for Cybersecurity*. Vol. 11779 LNCS. Springer International Publishing, 2019, pp. 198–214. ISBN: 9783030307950. DOI: 10.1007/978-3-030-30796-7_13 (cit. on p. 142).

[151] Dan Klein and Gal Engelberg. *Cyber Attack Countermeasures Generation With LLMs & Knowledge Graphs*. https://neo4j.com/video/connections/generative-ai-and-knowledge-graphs-unveiling-the-future-of-knowledge-retrieval/. Oct. 2023 (cit. on p. 92).

[152] Andrew Koenig. "Patterns and Antipatterns". In: *The Patterns Handbooks: Techniques, Strategies, and Applications*. USA: Cambridge University Press, 1998, 383–389. ISBN: 0521648181 (cit. on p. 21).

[153] Pius Krütli et al. "How to fairly allocate scarce medical resources: ethical argumentation under scrutiny by health professionals and lay people". In: *PloS one* 11.7 (2016), e0159086 (cit. on p. 34).

[154] Kabul Kurniawan, Andreas Ekelhart, and Elmar Kiesling. "An ATT&CK-KG for Linking Cybersecurity Attacks to Adversary Tactics and Techniques". In: *In Proceedings of the International Semantic Web Conference (ISWC)*. 2021 (cit. on p. 88).

[155] Anna-Lena Lamprecht et al. "Towards FAIR principles for research software". In: *Data Science* 3.1 (2020), pp. 37–59 (cit. on p. 23).

[156] Ora Lassila and Deborah McGuinness. "The role of frame-based representation on the semantic web". In: *Linköping electronic articles in computer and information science* 6.5 (2001), p. 2001 (cit. on pp. 20, 23, 41, 42, 44, 71).

[157] Kun Li et al. *CSKB: A Cyber Security Knowledge Base Based on Knowledge Graph*. Vol. 1268 CCIS. Springer Singapore, 2020, pp. 100–113. ISBN: 9789811591280. DOI: 10.1007/978-981-15-9129-7_8 (cit. on p. 142).

[158] Wolfgang Maass, Veda C Storey, and Roman Lukyanenko. "From mental models to machine learning models via conceptual models". In: *International Conference on Business Process Modeling, Development and Support*. Springer. 2021, pp. 293–300 (cit. on p. 4).

[159] *MAEC™ Specification – Core Concepts*. Malware Attribute Enumeration and Characterization (MAEC™). 2017 (cit. on p. 150).

[160] *MAEC™ Specification – Vocabularies*. Malware Attribute Enumeration and Characterization (MAEC™). 2017 (cit. on p. 150).

[161] David E. Mann and Steven M. Christey. "Towards a common enumeration of vulnerabilities". In: *2nd Workshop on Research with Security Vulnerability Databases, Purdue University, West Lafayette, Indiana*. 1999 (cit. on pp. 30, 88).

[162] Bob Martin et al. "CWE". In: *SANS top* 25 (2011) (cit. on p. 88).

[163] Robert A. Martin and Sean Barnum. "A status update: The common weaknesses enumeration". In: *Proc. of the Static Analysis Summit (NIST Special Publication 500-262)* (2006), pp. 62–64 (cit. on p. 88).

[164] Robert A Martin and Sean Barnum. "Common weakness enumeration (CWE) status update". In: *ACM SIGAda Ada Letters* 28.1 (2008), pp. 88–91 (cit. on pp. 30, 150).

[165] Beatriz F. Martins et al. "Conceptual Characterization of Cybersecurity Ontologies". In: *13th IFIP WG 8.1 working conference on the Practice*

of Enterprise Modelling. Springer. 2020, pp. 323–338 (cit. on pp. 6, 10, 23, 24, 28, 36, 45, 65, 77, 81, 139–142).

[166] Beatriz F. Martins et al. "Towards the Consolidation of Cybersecurity Standardized Definitions: a Tool for Ontological Analysis". In: *Proceedings of the XXIV Iberoamerican Conference on Software Engineering, CIbSE 2021, San José, Costa Rica, 2021*. Ed. by Tayana Conte et al. Curran Associates, 2021, pp. 290–303. ISBN: 978-1-7138-3944-6 (cit. on p. 78).

[167] Beatriz Franco Martins et al. "A framework for conceptual characterization of ontologies and its application in the cybersecurity domain". In: *Software and Systems Modeling* 21.4 (2022), pp. 1437–1464 (cit. on pp. 4, 7, 10, 23, 28, 30, 45, 63, 65, 72, 82, 139–142).

[168] Beatriz Franco Martins et al. "Improving Conceptual Domain Characterization in Ontology Networks". In: *Research Challenges in Information Science: 17th International Conference, RCIS 2023, Corfu, Greece, May 23–26, 2023, Proceedings*. Springer International Publishing, 2023, pp. 187–202. ISBN: 978-3-031-33080-3_12 (cit. on pp. 5, 10, 16, 23, 30, 64, 65, 68, 86, 184).

[169] Beatriz Franco Martins et al. "The Ontology for Conceptual Characterization of Ontologies". In: *Conceptual Modeling: 42th International Conference, ER 2023, Lisbon, Portugal, November 06–09, 2018, Proceedings 42*. Ed. by João Paulo A. Almeida et al. Cham: Springer Nature Switzerland, 2023, pp. 105–124. ISBN: 978-3-031-47262-6 (cit. on pp. 6, 10, 23, 54, 56, 59).

[170] Jenny Martinsson. "The use of vulnerability data for risk assessment". In: (2021) (cit. on p. 85).

[171] Claudio Masolo et al. "Relational roles and qua-individuals". In: *AAAI Fall Symposium on Roles, an interdisciplinary perspective*. AAAI PRESS-MIT PRESS. 2005, pp. 103–112 (cit. on p. 152).

[172] Claudio Masolo et al. "Wonderweb deliverable d17". In: *Science Direct Working Paper No S1574-034X (04)* (2002), pp. 70214–8 (cit. on p. 152).

[173] Melinda McDaniel and Veda C. Storey. "Evaluating domain ontologies: clarification, classification, and challenges". In: *ACM Computing Surveys (CSUR)* 52.4 (2019), pp. 1–44 (cit. on pp. 7, 30, 44).

[174] Mitre Corporation. *Science of Cyber-Security*. Tech. rep. McLean, Virginia: The MITRE Corporation, 2010 (cit. on p. 150).

[175] Riichiro Mizoguchi and Mitsuru Ikeda. "Towards ontology engineering". In: *Journal-Japanese Society for Artificial Intelligence* 13 (1998), pp. 9–10 (cit. on pp. 4, 23, 39, 40, 52).

[176] Laura Molloy, Peter McQuilton, and Yann Le Franc. *EOSC Co-creation funded project 074: Delivery of a proof of concept for terms4FAIRskills: Technical report*. Tech. rep. Version 1.0.0. May 2021. DOI: 10.5281/zenodo.4772741 (cit. on p. 37).

[177] Barend Mons et al. "Cloudy, increasingly FAIR; revisiting the FAIR Data guiding principles for the European Open Science Cloud". In: *Information services & use* 37.1 (2017), pp. 49–56 (cit. on pp. 23, 34).

[178] Bruno Augusti Mozzaquatro et al. "An ontology-based cybersecurity framework for the internet of things". In: *Sensors* 18.9 (2018), p. 3053 (cit. on p. 142).

[179] David A. Mundie et al. "An Incident Management Ontology". In: *STIDS*. 2014, pp. 62–71 (cit. on p. 142).

[180] Paul Murdock et al. "Semantic interoperability for the Web of Things". PhD thesis. Dépt. Réseaux et Service Multimédia Mobiles (Institut Mines-Télécom-Télécom ..., 2016 (cit. on p. 35).

[181] Sandeep Narayanan et al. "Cognitive Techniques for Early Detection of Cybersecurity Events". In: *arXiv preprint arXiv:1808.00116* (2018) (cit. on p. 142).

[182] National Institute of Standards and Technology. *Framework for Improving Critical Infrastructure Cybersecurity*. Tech. rep. National Institute of Standards and Technology, 2014 (cit. on p. 150).

[183] National Institute of Standards and Technology. *Framework for Improving Critical Infrastructure Cybersecurity*. Tech. rep. National Institute of Standards and Technology, 2018. DOI: https://doi.org/10.6028/NIST.CSWP.04162018 (cit. on p. 150).

[184] National Institute of Standards and Technology. *Security Self-Assessment Guide for Information Technology Systems*. Tech. rep. Gaithersburg, M. D.: National Institute of Standards and Technology, 2001 (cit. on p. 150).

[185] NERC. *CIPC Control Systems Security Working Group*. Tech. rep. NERC, 2014 (cit. on p. 150).

[186] NERC. *Glossary of Terms Used in NERC Reliability Standards*. Tech. rep. NERC, 2020 (cit. on p. 150).

[187] William Newhouse et al. *National Initiative for Cybersecurity Education (NICE) Cybersecurity Workforce Framework*. Tech. rep. NIST, 2017. DOI: https://doi.org/10.6028/NIST.SP.800-181 (cit. on p. 150).

[188] B. Nolan Nichols et al. "Linked Data in Neuroscience: Applications, Benefits, and Challenges". In: *bioRxiv* (2016), p. 053934 (cit. on p. 34).

[189] Michael Nieles, Kelley Dempsey, and Victoria Yan Pillitteri. *An Introduction to Information Security An Introduction to Information Security*. Tech. rep. NIST, 2017. DOI: https://doi.org/10.6028/NIST.SP.800-12r1 (cit. on p. 150).

[190] NIST, ed. *Generally accepted principles and practices for securing information technology systems*. NIST, 2018. ISBN: 3019753058 (cit. on p. 150).

[191] Daniel Oberle. *Semantic management of middleware*. Vol. 1. Springer Science & Business Media, 2006 (cit. on pp. 41, 43).

[192] Leo Obrst. "Ontological architectures". In: *Theory and applications of ontology: computer applications*. Springer, 2010, pp. 27–66 (cit. on p. 29).

[193] Leo Obrst, Penny Chase, and Richard Markeloff. "Developing an Ontology of the Cyber Security Domain". In: *STIDS*. 2012, pp. 49–56 (cit. on pp. 29, 142).

[194] Ítalo Oliveira et al. "Boosting D3FEND: Ontological Analysis and Recommendations". In: *Formal Ontology in Information Systems*. Ed. by IOS Press. Nieuwe Hemweg, The Netherlands: IOS Press, 2023, pp. – (cit. on p. 29).

[195] Ítalo Oliveira et al. "How FAIR are security core ontologies? A systematic mapping study". In: *Research Challenges in Information Science: 15th International Conference, RCIS 2021, Limassol, Cyprus, May 11–14, 2021, Proceedings*. Springer. 2021, pp. 107–123 (cit. on pp. 6, 23, 29, 36, 45).

[196] Alessandro Oltramari and Alexander Kott. "Towards a reconceptualisation of cyber risk: An empirical and ontological study". In: *Journal of Information Warfare* 17.1 (2018), pp. 49–73 (cit. on p. 83).

[197] Alessandro Oltramari et al. "Building an Ontology of Cyber Security". In: *STIDS*. Citeseer. 2014, pp. 54–61 (cit. on pp. 29, 82, 142).

[198] Alessandro Oltramari et al. "Computational ontology of network operations". In: *MILCOM 2015-2015 IEEE Military Communications Conference*. IEEE. 2015, pp. 318–323 (cit. on pp. 29, 142).

[199] Alessandro Oltramari et al. "Senso Comune". In: *LREC*. Citeseer. 2010 (cit. on p. 152).

[200] Alessandro Oltramari et al. "Towards a Human Factors Ontology for Cyber Security." In: *STIDS*. 2015, pp. 26–33 (cit. on pp. 29, 142).

[201] C. Onwubiko. "CoCoa: An Ontology for Cybersecurity Operations Centre Analysis Process". In: *2018 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA)*. 2018, pp. 1–8 (cit. on pp. 29, 142).

[202] Xinming Ou, Sudhakar Govindavajhala, and Andrew W. Appel. "MulVAL: A Logic-based Network Security Analyzer". In: *USENIX*

security symposium. Vol. 8. Baltimore. 2005, pp. 113–128 (cit. on pp. 82, 142).

[203] Hercules Panoutsopoulos, Christopher Brewster, and Spyros Fountas. "A Semantic Data Model for a FAIR Digital Repository of Heterogeneous Agricultural Digital Objects". In: *2nd Integrated Food Ontology Workshop*. Bolzano, Italy. 2021 (cit. on p. 23).

[204] Mary C. Parmelee. "Toward an Ontology Architecture for Cyber-Security Standards". In: *STIDS* 713 (2010), pp. 116–123 (cit. on p. 142).

[205] oneM2M Partners. *ONEM2M TECHNICAL SPECIFICATION*. Tech. rep. TS-0012-V3.7.3. Type 1 (ARIB, ATIS, CCSA, ETSI, TIA, TSDSI, TTA, TTC). oneM2M, 2019 (cit. on p. 35).

[206] Oscar Pastor. "Diseño y Desarrollo de un Entorno de Producción Automática de Software basado en el modelo orientado a Objetos". PhD thesis. Tesis doctoral dirigida por Isidro Ramos, DSIC, Universitat Politècnica de València, 1992 (cit. on p. 59).

[207] Oscar Pastor et al. "Model-driven development". In: *Informatik-Spektrum* 31 (2008), pp. 394–407 (cit. on p. 21).

[208] Oscar Pastor et al. "The OO-Method approach for information systems modeling: from object-oriented conceptual modeling to automated programming". In: *Information Systems* 26.7 (2001), pp. 507–534. ISSN: 0306-4379. DOI: https://doi.org/10.1016/S0306-4379(01)00035-7 (cit. on p. 59).

[209] Christopher Peacocke. *A study of concepts*. The MIT Press, 1992 (cit. on p. 55).

[210] Robert Pergl, Tiago Prince Sales, and Zdeněk Rybola. "Towards OntoUML for software engineering: from domain ontology to implementation model". In: *Model and Data Engineering: Third International Conference, MEDI 2013, Amantea, Italy, September 25-27, 2013. Proceedings 3*. Springer. 2013, pp. 249–263 (cit. on p. 18).

[211] María Poveda-Villalón et al. "Coming to terms with FAIR ontologies". In: *International Conference on Knowledge Engineering and Knowledge Management*. Springer. 2020, pp. 255–270 (cit. on pp. 23, 36).

[212] Shengzhi Qin and K. P. Chow. "Automatic Analysis and Reasoning Based on Vulnerability Knowledge Graph". In: *Communications in Computer and Information Science*. Ed. by Huansheng Ning. Vol. 1137 CCIS. Singapore: Springer Singapore, 2019, pp. 3–19. ISBN: 9789811519215. DOI: 10.1007/978-981-15-1922-2_1 (cit. on p. 143).

[213] Dirk Riehle and Heinz Züllighoven. "Understanding and using patterns in software development". In: *Tapos* 2.1 (1996), pp. 3–13 (cit. on p. 21).

[214] Alejandro Rodríguez-Iglesias et al. "Publishing FAIR data: an exemplar methodology utilizing PHI-base". In: *Frontiers in plant science* 7 (2016), p. 641 (cit. on p. 34).

[215] Hy Rothstein and Barton Whaley. *The art and science of military deception*. Artech House, 2013 (cit. on p. 88).

[216] Tiago Prince Sales et al. "The common ontology of value and risk". In: *International Conference on Conceptual Modeling*. Springer. 2018, pp. 121–135 (cit. on pp. 29, 82, 142).

[217] Kaz Sato, Cliff Young, and David Patterson. "An in-depth look at Google's first Tensor Processing Unit (TPU)". In: *Google Cloud Big Data and Machine Learning Blog* 12 (2017) (cit. on p. 3).

[218] Ferdinand Mongin Saussure. *Course in general linguistics*. Ed. by Charles Bally and in collaboration with Albert Riedlinger Albert Sechehaye. New York, Toronto, and London: McGraw-Hill Book Company, 1966 (cit. on p. 21).

[219] Noemi Scarpato, Nicole Dalia Cilia, and Marco Romano. "Reachability matrix ontology: a cybersecurity ontology". In: *Applied Artificial Intelligence* 33.7 (2019), pp. 643–655 (cit. on p. 143).

[220] Markus Schumacher. "Toward a Security Core Ontology". In: *Security Engineering with Patterns: Origins, Theoretical Model, and New*

Applications. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 87–96. ISBN: 978-3-540-45180-8. DOI: 10.1007/978-3-540-45180-8_6 (cit. on p. 151).

[221] Luiz Olavo Bonino da Silva Santos et al. "FAIR Data Point: A FAIR-Oriented Approach for Metadata Publication". In: *Data Intelligence* 5.1 (Mar. 2023), pp. 163–183. ISSN: 2641-435X. DOI: 10.1162/dint_a_00160. eprint: https://direct.mit.edu/dint/article-pdf/5/1/163/2074301/dint_a_00160.pdf (cit. on p. 34).

[222] Luiz Olavo Bonino da Silva Santos et al. "FAIR data points supporting big data interoperability". In: *Enterprise Interoperability in the Digitized and Networked Factory of the Future. ISTE, London* (2016), pp. 270–279 (cit. on p. 33).

[223] Elena Simperl et al. "ONTOCOM: A reliable cost estimation method for ontology development projects". In: *Journal of Web Semantics* 16 (2012), pp. 1–16 (cit. on p. 23).

[224] Barry Smith et al. "The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration". In: *Nature biotechnology* 25.11 (2007), pp. 1251–1255 (cit. on p. 37).

[225] Sarah K. Squire et al. *Federation and Assertions*. Tech. rep. NIST, 2017. DOI: https://doi.org/10.6028/NIST.SP.800-63c (cit. on p. 150).

[226] *Standard 1300 — Cyber Security*. Standard 1300 – Quality Assurance and Improvement Program. 2004 (cit. on p. 150).

[227] Keith Stouffer et al. *Guide to Industrial Control Systems (ICS) Security NIST Special Publication 800-82 Guide to Industrial Control Systems (ICS) Security – Supervisory Control and Data Acquisition (SCADA) Systems, Distributed Control Systems (DCS), and Other Control Syst.* Tech. rep. NIST, 2015. DOI: http://dx.doi.org/10.6028/NIST.SP.800-82r2 (cit. on p. 150).

[228] Blake E. Strom et al. "Mitre att&ck: Design and philosophy". In: *Technical report*. The MITRE Corporation, 2018 (cit. on p. 88).

[229] Blake E. Strom et al. *MITRE ATT&CK(trademark): Design and Philosophy*. Tech. rep. MITRE CORP MCLEAN VA MCLEAN, 2018 (revised 2020) (cit. on p. 151).

[230] Rudi Studer, V. Richard Benjamins, and Dieter Fensel. "Knowledge engineering: principles and methods". In: *Data & knowledge engineering* 25.1-2 (1998), pp. 161–197 (cit. on pp. 4, 15, 23, 39, 151).

[231] Mari Carmen Suárez-Figueroa et al. *Introduction: Ontology engineering in a networked world*. Springer, 2012 (cit. on p. 30).

[232] Romilla Syed. "Cybersecurity vulnerability management: A conceptual ontology and cyber intelligence alert system". In: *Information & Management* 57.6 (2020), p. 103334. ISSN: 0378-7206. DOI: https://doi.org/10.1016/j.im.2020.103334 (cit. on p. 142).

[233] Romilla Syed and Haonan Zhong. "Cybersecurity vulnerability management: An ontology-based conceptual model". In: *AMCIS 2018 Proceedings*. 2018 (cit. on p. 142).

[234] Zareen Syed et al. "UCO: A unified cybersecurity ontology". In: *Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*. 2016 (cit. on p. 143).

[235] T. Takahashi and Y. Kadobayashi. "Reference Ontology for Cybersecurity Operational Information". In: *The Computer Journal* 58.10 (2015), pp. 2297–2312 (cit. on pp. 28, 143).

[236] Takeshi Takahashi, Hiroyuki Fujiwara, and Youki Kadobayashi. "Building ontology of cybersecurity operational information". In: *Proceedings of the Sixth Annual Workshop on Cyber Security and Information intelligence Research*. 2010, pp. 1–4 (cit. on pp. 28, 142).

[237] Takeshi Takahashi and Youki Kadobayashi. "3-5 cybersecurity information exchange techniques: Cybersecurity information ontology and cybex". In: *Journal of the National Institute of Information and Communications Technology Vol* 58.3/4 (2011) (cit. on pp. 28, 29, 142).

[238] Takeshi Takahashi, Youki Kadobayashi, and Hiroyuki Fujiwara. "Ontological approach toward cybersecurity in cloud computing". In: *Proceedings of the 3rd international conference on Security of information and networks*. 2010, pp. 100–109 (cit. on pp. 28, 29, 142).

[239] Cassia Trojahn et al. "Foundational ontologies meet ontology matching: A survey". In: *Semantic Web* 13.4 (2022), pp. 685–704 (cit. on p. 29).

[240] Juan Trujillo et al. *Conceptual modeling in the era of Big Data and Artificial Intelligence: Research topics and introduction to the special issue*. 2021 (cit. on p. 4).

[241] Alan Turing. "Intelligent machinery (1948)". In: *B. Jack Copeland* (2004), p. 395 (cit. on p. 3).

[242] Jeffrey Undercofer et al. "A target-centric ontology for intrusion detection". In: *Workshop on Ontologies in Distributed Systems, held at The 18th International Joint Conference on Artificial Intelligence*. 2003 (cit. on p. 142).

[243] Cybersecurity Unit. "A Framework for a Vulnerability Disclosure Program for Online Systems". In: *US Department of Justice* (2017) (cit. on p. 85).

[244] Michael Uschold and Michael Gruninger. "Ontologies and semantics for seamless connectivity". In: *ACM SIGMod Record* 33.4 (2004), pp. 58–64 (cit. on pp. 20, 40, 41, 44, 53, 152).

[245] Michael Uschold and Michael Gruninger. "Ontologies: Principles, methods and applications". In: *The knowledge engineering review* 11.2 (1996), pp. 93–136 (cit. on pp. 23, 38, 152).

[246] Andre Valente and Joost Breuker. "Towards principled core ontologies". In: *Proceedings of the International Knowledge Acquisition Workshop, Banff, Canada./ cited*. Citeseer. 1996 (cit. on p. 39).

[247] Gertjan Van Heijst, A. Th Schreiber, and Bob J. Wielinga. "Using explicit ontologies in KBS development". In: *International Journal of*

human-computer Studies (1997), pp. 183–292 (cit. on pp. 39, 52, 69, 142, 143, 152, 184).

[248] Bret Jordan Varner and Drew, eds. *OASIS – TAXII™ Version 2.1*. OASIS Committee Specification 01, 2020 (cit. on p. 151).

[249] Michaël Verdonck et al. "Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study". In: *Inf. Syst.* 81 (2019), pp. 92–103. DOI: 10.1016/j.is.2018.11.009 (cit. on p. 29).

[250] Yair Wand and Ron Weber. "On the deep structure of information systems". In: *Information Systems Journal* 5.3 (1995), pp. 203–223 (cit. on p. 59).

[251] Ju An Wang and Minzhe Guo. "OVM: an ontology for vulnerability management". In: *5th Annual Workshop on Cyber Security and Information Intelligence Research: Cyber Security and Information Intelligence Challenges and Strategies*. 2009, pp. 1–4 (cit. on pp. 29, 143).

[252] Roel Wieringa. *Design Science Methodology for Information Systems and Software Engineering*. Springer, 2014. ISBN: 978-3-662-43838-1. DOI: 10.1007/978-3-662-43839-8 (cit. on pp. 8, 9).

[253] Mark D. Wilkinson et al. "A design framework and exemplar metrics for FAIRness". In: *Scientific data* 5.1 (2018), pp. 1–4 (cit. on pp. 23, 35).

[254] Mark D. Wilkinson et al. "The FAIR Guiding Principles for scientific data management and stewardship". In: *Scientific data* 3.1 (2016), pp. 1–9 (cit. on pp. 5, 10, 23, 33, 34).

[255] Patrick Henry Winston. *Artificial intelligence*. Addison-Wesley Longman Publishing Co., Inc., 1984 (cit. on p. 3).

[256] Edward N Zalta. "Fregean senses, modes of presentation, and concepts". In: *Philosophical Perspectives* 15 (2001), pp. 335–359 (cit. on p. 55).

[257] Zhi-Hua Zhou. *Machine learning*. Springer Nature, 2021 (cit. on p. 3).

[258] Carson Zimmerman. *Ten Strategies of a World-Class Cybersecurity Operations Center*. Bedford, MA: The MITRE Corporation, 2014. ISBN: 9780692243107 (cit. on p. 151).

[259] Elisabetta Zuanelli. "The cybersecurity ontology platform: the POC solution". In: *e-AGE2017* (2017), p. 1 (cit. on p. 143).

APPENDIX

Appendix A

Cybersecurity Ontologies TLR

A.1 Selection Process

Table A.1: Cybersecurity Ontologies' TLR works selection process [165, 167].

	Selection criteria applied
SC1	By filtering the publication title
SC2	By the read of the publication abstract
SC3	By the read of the whole document
	Inclusion criteria applied
IC1	Papers that present cybersecurity ontologies
IC2	Papers that present parts of cybersecurity ontologies
	Exclusion criteria applied
EC1	Papers inaccessible for reading
EC2	Papers with low relevance by the number of citations
EC3	Papers that present parts of cybersecurity ontologies
EC4*	Papers already selected in the first round of search

* We added the fourth exclusion criterion in the second round of search.

At least two researchers carried out this work, one for each perspective: a domain specialist regarding Cybersecurity and an ontology engineer for the Ontological bias. The ontology engineer conducts all the search stages, and the domain specialist gives expert advice, typically on cybersecurity details and doubts clarification matters.

Search string used in the first round [165]:

```
//Accessed on April 2020 at ACM, IEEE, Scopus, and Google Academic.

// First round search string -- with wildcard characters
TITLE = ("Cybersecurity Ontolog*")
```

Search string used in the second round [167]:

```
//Accessed on January 2021 at ACM, IEEE, Scopus, and Google Academic.

// Second round search string -- with operator OR
TITLE = ("Cybersecurity Ontology" OR "Cybersecurity Ontologies")
```

A.2 Search Outcomes

Table A.2: Cybersecurity Ontologies' TLR works summary [165, 167].

Search	Number of Publications			
	April 2020	**January 2021**	**Manual add**	**Total**
Papers found	198	229 (31 additional)	2	231
Papers inspected	32	48 (17 additional)	2	51
Papers excluded (EC1)	3	0	0	3
Papers excluded (EC2)	0	0	0	0
Papers excluded (EC3)	4	9	0	12
Papers excluded (EC4)	–	32	0	32
Papers included	25	8	2	35
*Ontologies found**	*19*	*8 (1 sub-ontology)*	*2*	*28*

* Italic data refers to the total of ontology references found in the included papers

Note, that the total of ontologies that we consider does not correspond to the number of papers found, because some refer to the same ontology and others to multiple ontologies (or sub-ontologies).

Table A.3: Level of Applicability of the studied ontologies [165, 167].

Level of Applicability	**Number of Ontologies**			
	April 2020	**January 2021**	**Manual add**	**Total**
Reference Ontology	5	1 (0 additional)	1	6
Operational Ontology	20	8 (7 additional)	1	28
*Operational Ontology supported by a Reference Ontology**	*4*	*0*	*1*	*5*

* Italic data refers to the notion of well-defined ontologies presented in [93]

Table A.4: Level of Generality of the studied ontologies [165, 167].

Level of Generality	**Number of Ontologies**			
	April 2020	**January 2021**	**Manual add**	**Total**
Domain Ontology	11	6	2	19
Task Ontology	0	0	0	0
Application Ontology	5	1 (sub-ontology)	0	5
Core Ontology	2	0	0	2
*Ontology grounded over a Foundational Ontology**	*4*	*1*	*0*	*5*

* Italic data refers to the notion of well-grounded ontologies presented in [75]

Table A.5: Summary of Cybersecurity Ontology Characterization [165, 167].

Search	Proposed Ontology	Level of Applicability		Level of Generality	
		Reference Ontology	Operational Ontology	Well-grounded	According to [76, 247]
Apr/20	CCS [181]	No	Yes	No	Domain
Apr/20	CoCoa [201]	Yes	Yes	No	Application
Jan/21	Cold-start Cybersecurity Ontology [65]	No	Yes	No	Domain
Jan/21	CSR-Cybersecurity Ontology [12]	No	Yes	No	Domain
Manual	COoVR [216]	Yes	No	Yes	Domain
Apr/20	CVO – CIA System [233]	Yes	No	No	Domain
Jan/21	CVO & CIO – CIA System [232]	Yes	Yes	No	Application
Apr/20	CRATELO [197]	No	Yes	Yes	Application
Apr/20	CRATELO [200]	No	Yes	Yes	Domain
Apr/20	CRATELO [198]	No	Yes	Yes	Domain
Apr/20	CRATELO [18]	No	Yes	Yes	Domain
Apr/20	Cyber Ontology [204]	No	Yes	No	Application
Apr/20	Cybersecurity Ontology for Critical Infrastructures [21]	No	Yes	No	Domain
Jan/21	Cybersecurity Ontology for the CSKB [157]	No	Yes	No	Domain
Apr/20	IDS [242]	No	Yes	No	Core
Apr/20	IM [179]	No	Yes	No	Domain
Apr/20	IoTSec [178]	No	Yes	No	Domain
Apr/20	Cybersecurity Knowledge Base [142]	No	Yes	No	Domain
Apr/20	Malware Ontology [71]	Yes	No	No	Domain
Apr/20	MITRE Co approach [193]	No	Yes	No	Core
Apr/20	MulVAL [202]	No	Yes	No	Application
Jan/21	Ontology for the SEPSES KG Cybersecurity [150]	No	Yes	No	Domain
Apr/20	Ontology of Cybersecurity Operational Information [237]	Yes	Yes	No	Domain
Apr/20	Ontology of Cybersecurity Operational Information [236]	Yes	No	No	Domain
Apr/20	Ontology of Cybersecurity Operational Information [238]	Yes	Yes	No	Domain

– *continued on next page.*

Table A.5 continued from previous page.

Search	Proposed Ontology	Level of Applicability		Level of Generality	
		Reference Ontology	**Operational Ontology**	**Well-grounded**	**According to [76, 247]**
Apr/20	Ontology of Cybersecurity Operational Information [235]	Yes	No	No	Domain
Manual	Ontology of ISO/IEC 27005 [4]	No	Yes	No	Domain
Apr/20	OVM [251]	Yes	Yes	No	Domain
Apr/20	POC [259]	No	Yes	No	Application
Apr/20	SCIC [57]	No	Yes	No	Domain
Jan/21	RMO Ontology [219]	No	Yes	No	Domain
Jan/21	SecOrP [118]	No	Yes	No	Domain
Apr/20	UCO [234]	No	Yes	No	Domain
Apr/20	VDO [25]	Yes	No	No	Domain
Jan/21	VulKG [212]	No	Yes	No	Domain

Appendix B

Cybersecurity Conceptualization Sources

B.1 Cybersecurity Terminology

Table B.1: The initial list of terms present in the cybersecurity cloud of concepts.

Initial Terminology		
access	information	phishing
access control	information need	policy
application	information security	process*
asset	information system	provider
attack	integrity	reliability
authentication	internet	requirement
availability	likelihood	review
bot	malicious software	risk
competence	malware	risk assessment
confidentiality	measure	risk management
consequence	measurement	stakeholder
control	monitoring	threat
countermeasure	objective	trojan
event*	organization	trojan horse
indicator	performance	vulnerability

* Term also used in foundational ontologies

Table B.2: Additional list of terms present in the cybersecurity cloud of concepts.

Additional Terminology	
access control functions	interested party
access control mechanism	internal context
access control policies	internet crime
access control policy	internet safety
adware	internet security
application service provider	internet service provider
application services	internet services
application software	internet work
artifact*	level of risk
attack mechanisms	machine code
attack potential	malicious contents
attack vector	management system
audit	measurement function
audit scope	measurement method
authenticity	nonconformity
authorization	non-repudiation
avatar	organizational assets
base measure	outsource
blended attack	personal assets
botnet	physical asset
code	potentially unwanted software
conformity	program specification
consumer	program
continual improvement	programming language
control objective	residual risk
cookie	review object
correction	review objective
corrective action	risk acceptance
cyber-squatter	risk analysis
cybercrime	risk communication and consultation
cybersafty	risk criteria
cybersecurity	risk evaluation
cyberspace	risk identification
cyberspace application services	risk management process
cyberspace security	risk owner
deceptive software	risk treatment
defensive tactic	robot
defensive technique	scam
derived measure	security implementation standard
digital artifact	software system
digital artifactual value	source code
digital object	spam
documented information	spyware
drone	system specification
effectiveness	technique reference
external context	technique
governance of information security	threat agent

– continued on next page.

Table B.2 continued from previous page.	
Additional Terminology	
governing body	top management
hacking	trusted information communication entity
hacktivism	unsolicited email
information asset	virtual artifact
information processing facilities	virtual asset
information security continuity	virtual currency
information security event	virtual world
information security incident	weakness
information security incident management	zombie computer
information security management system professional	zombie
information sharing community	

* Term also used in foundational ontologies

Table B.3: List of terms present in the cloud of concepts from foundational ontologies.

Foundational Ontologies Terminology		
abstract	human state	proposition
abstract individual	individual	qua individual
abstract quality	institutional agent	quale
abstract region	intangible	quality
accomplishment	intention	quality kind
achievement	intentional moment	quality region
action	interaction	quality structure
action contribuition	internal closed commitment universal	quality universal
action universal	internal commitment universal	quality value
agent**	internal relation	quantity
agentive physical object	intrinsic moment	quantity universal
amount of matter	intrinsic moment universal	region
antirigid mixin	kind	relation universal
anti-rigid sortal	material relation	relational qua individual
appointment	measurable quality universal	relator
appointment goal	measurement quality dimension	relator universal

– continued on next page.

Table B.3 continued from previous page.

Foundational Ontologies Terminology

arbitrary sum	measurement quality domain	resource
atomic action	measurement quality region	resource participation
atomic action universal	measurement quality structure	rigid mixin
atomic event	mediation	rigid sortal
atomic measurement quality region	meets	role
basic measurement quality region	member	rolemixin
before	memberof	scientific measurement domain
belief	mental moment	self appointment
category	mental object	semirigid mixin
change	mental process	sensible quality
characterization	meronymic	set
circular measurement dimension	mixin	situation
closed appointment	mixin universal	social agent
closed commitment	mode	social appointment
closed commitment universal	mode kind	social claim
cognitive measurement domain	mode universal	social commitment
collective	moment	social moment
collective member	moment universal	social object
collective social agent	monadic universal	social relator
collective universal	natural object	social role
commitment	natural process	social rolemixin
commitment universal	nominal quality region	society
communicative act	nominal quality structure	sortal universal
complex	nominal quality universal	space region
complex action	non agentive physical object	spatial location
complex action universal	non physical endurant	spatial quality
complex closed appointment	non physical object	starts
complex closed commitment	non atomic measurement quality region	state
complex event	non boundary measurement dimension	stative
complex member	non perceivable quality universal	structuration

– *continued on next page.*

Table B.3 continued from previous page.

Foundational Ontologies Terminology		
componentof	non-rigid mixin	subcollectionof
composite measurement quality region	non-sortal universal	subkind
concrete individual	normative description	subquantityof
creation	object	substance sortal
delegation	object kind	substantial
delegatum	one boundary measurement dimension	substantial moment
derivation	open commitment	substantial universal
desire	overlaps	successful action
disposition	participation	tangible
domain formal relation	participation universal	temporal location
during	perceivable quality universal	temporal quality
ends	perdurant	temporal region
endurant	perdurant universal	temporal structure
endurant universal	phase	termination
entity	phased qua individual	thing
equal	phasemixin	time interval
event**	physical agent	time interval relation
event universal	physical endurant	time point
fact	physical object	trope kind
feature	physical quality	type
formal relation	physical region	unfulfiled intention
fulfilled intention	physical state	universal
function	plan description	usage
goal	process**	

** Term also used in non-foundational ontologies

B.2 Cybersecurity Sources

Table B.4: List of sources used in the cybersecurity (and surroundings) cloud of concepts.

Conceptualization Sources	
CCDB-2017-05-xxx [37]	Standard
CCMB-2017-04-001 [39]	Standard
CCMB-2017-04-002 [40]	Standard
CCMB-2017-04-003 [41]	Standard
CCMB-2017-04-004 [42]	Standard

– *continued on next page.*

Table B.4 continued from previous page.

Conceptualization Sources	
CCN-STIC-401 [86]	Standard
CCv31-Release 5 [43]	Standard
CVE-1999-0001 [23]	Standard
CWE [47, 164]	Standard
DAO Version 0.10.1-BETA-1 (2022-06-13) [46]	Electronic Document
Enforcement 2020 [186]	Standard
Incibe [67]	Standard
ISO/IEC 154081-2009 [125]	Standard
ISO/IEC 154082-2008 [121]	Standard
ISO/IEC 154083-2008 [119]	Standard
ISO/IEC 27000-2018 [123]	Standard
ISO/IEC 27002-2013 [120]	Standard
ISO/IEC 27002-2018 [124]	Standard
ISO/IEC 27032-2012 [122]	Standard
ISO/IEC/IEEE 24765-2010 (SEVOCAB) [126]	Standard
ITU-T-RecX.811 [130]	Standard
ITU-T-Rec-X805 [134]	Standard
ITU-T-Rec-X810 [128]	Standard
ITU-TRecX812 [132]	Standard
ITU-T-Rec-X813 [133]	Standard
ITU-T-RecX814 [131]	Standard
ITU-T-RecX815 [127]	Standard
ITU-T-RecX816 [129]	Standard
JSR-10-102 [174]	Standard
MAEC 5.0 Vocab [160]	Standard
MAEC 5.0 Spec [159]	Standard
NERC-CIPv3-v5 [185]	Standard
NIST.SP.800-63-3 [9]	Standard
NIST.SP.800-63a [60]	Standard
NIST.SP.800-63b [59]	Standard
NIST.SP.800-82r2 [227]	Standard
NIST-800-14 [190]	Standard
NIST-800-181 [187]	Standard
NIST-800-37 Revision 2 [144]	Standard
NIST-800-53 Rev 4 [145]	Standard
NIST-800-63c [225]	Standard
NIST-CSWP-04162014 [182]	Standard
NIST-CSWP-04162018 [183]	Standard
NIST-SP-800-12 Revision 1 [189]	Standard
RecITU-T-X1205 [138]	Standard
RecITU-T-X1209 [136]	Standard
RecITU-T-X1212 [137]	Standard
RecITU-T-X1500 [135]	Standard
Rec X800 [38]	Standard
Spec Publ 800-26 [184]	Standard
Standard 1300 [226]	Standard
STIX-v21-CS01 [34]	Standard

– *continued on next page.*

Table B.4 continued from previous page.	
Conceptualization Sources	
STIX and TAXII [15]	Standard
STROM 2020 [229]	White Literature
TAXII and STIX [48]	Standard
TAXII-v21-CS01 [248]	Standard
Ten Strategies of a World-Class Cybersecurity Operations Center [258]	White Literature
The 2011 Standard of Good Practice [44]	White Literature
Toward a knowledge graph of cybersecurity countermeasures [148]	White Literature
Toward a Security Core Ontology[220]	While Literature

Table B.5: List of sources used in the cloud of concepts grounding.

Grounding Sources	
A translation approach to portable ontology specifications [72]	While Literature
Agent Roles, Qua Individuals and the Counting Problem [90]	White Literature
An ontological foundation for conceptual modeling datatypes based on semantic reference spaces [5]	White Literature
An ontologically well-founded profile for UML conceptual models [108]	White Literature
Conceptual simulation modeling with Onto-UML [104]	White Literature
Construction of engineering ontologies for knowledge sharing and reuse [30]	While Literature
Events, their names, and their synchronic structure [78]	White Literature
Formal and transcendental logic [117]	While Literature
Formal Ontology in Information Systems [75]	While Literature
Formal Semantics and Ontological Analysis for Understanding Subsetting, Specialization and Redefinition of Associations in UML [49]	White Literature
gUFO: a lightweight implementation of the Unified Foundational Ontology (UFO) [6]	White Literature
Grounding Software Domain Ontologies in the Unified Foundational Ontology (UFO) [101] In the Defense of a Trope-Based Ontology for Conceptual Modeling: An Example with the Foundations of Attributes, Weak Entities and Datatypes, 25th Intl [88]	White Literature
Knowledge engineering: principles and methods [230]	While Literature
Lightweight ontologies [66]	While Literature
Logical, ontological and cognitive aspects of object types and cross-world identity with applications to the theory of conceptual spaces [91]	White Literature
Modal Aspects of Object Types and Part-Whole Relations and the de re/de dicto Distinction [92]	White Literature
On ontology, ontologies, conceptualizations, modeling languages, and (meta)models [93]	While Literature
On the Foundations of UML as an Ontology Representation Language [106]	White Literature

– *continued on next page.*

Table B.5 continued from previous page.

Grounding Sources	
On the Representation of Quantities and their Parts in Conceptual Modeling [94]	White Literature
Ontological Foundations for Conceptual Part-Whole Relations: The Case of Collectives and Their Parts [95]	White Literature
Ontological Foundations for Structural Conceptual Models [96]	White Literature
Ontological Foundations of DOLCE [27]	
Ontologies and semantics for seamless connectivity [244]	While Literature
Ontologies: Principles, methods and applications [245]	While Literature
Ontology languages for the semantic web [68]	While Literature
Ontology-based transformation framework from TROPOS to AORML [111]	White Literature
Reification and Truthmaking Patterns [85]	White Literature
Relational roles and qua-individuals [171]	White Literature
Relations in ontology-driven conceptual modeling [64] Relationships and events: towards a general theory of reification and truthmaking [80]	White Literature
Senso Comune [199]	White Literature
The problem of transitivity of part-whole relations in conceptual modeling revisited [100]	White Literature
Towards an ontology of software defects, errors and failures [56]	White Literature
Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story [109]	White Literature
Towards ontological foundations for the conceptual modeling of events [110]	White Literature
Using a trope-based foundational ontology for bridging different areas of concern in ontology-driven conceptual modeling [107]	White Literature
Using explicit ontologies in KBS development [247]	While Literature
"We need to discuss the Relationship": Revisiting Relationships as Modeling Constructs [81]	White Literature
What's in a relationship: an ontological analysis [105]	White Literature
Wonderweb deliverable d17 [172]	While Literature

Appendix C

O4OA Documentation Details

C.1 Related Ontologies

Table C.1 presents the ontologies related to O4OA reference model:

Table C.1: O4OA Related Ontologies.

Ontology	Relation	Compatibility
UFO	Ontological Grounding through OntoUML.	High
UFO-A	Structural aspects of grounding.	High
UFO-B	Dynamic aspects (Ontology Versions) of grounding.	High
UFO-C	Social aspects (Ontological Commitment) of grounding.	High
gUFO	gUFO is the light light version of UFO implemented in OWL	High
gO4OA	O4OA light implemented model in OWL and grounded over gUFO	High
MongoO4OA	O4OA implemented model in MongoDB	High

C.2 Competence Questions

Table C.2 shows the competence questions that O4OA must answer.

Table C.2: Competence Questions.

Ref.	Competence Question
CQ1	*How to conceptually characterize an ontology (as an artifact)?*
CQ1.1	What is the applicability level of an ontology?
CQ1.2	What is the generality level of an ontology?
CQ1.3	What is the formalization level of an ontology?
CQ1.4	What is the axiomatization level of an ontology?
CQ1.5	What is an ontology that is considered a well-grounded conceptualization?
CQ1.6	What is an ontology that is considered a well-defined conceptualization?
CQ1.7	Which (meta)characteristics of an ontology interfere with its characterization?
CQ1.8	How do the languages in which ontologies are represented or implemented interfere with their conceptualization?
CQ2	*How to conceptually characterize the domain cloud of concepts of an ontology?*
CQ2.1	What is a concept when it is represented in the context of one or more conceptualizations?
CQ2.2	Which information sources support the conceptualization of a domain?
CQ2.3	Which are the terms and their definitions (cloud of concepts) belonging to the conceptualization of a domain?
CQ2.4	Which sources provide the same (equal, equivalent, similar) definition for a particular term, in the face of one or more conceptualizations?
CQ2.5	Which are the terms (or synonyms) have the same definition, in the face of one or more conceptualizations?
CQ2.6	What are the different definitions for a term in the cloud of concepts, in the face of one or more conceptualizations?
CQ2.7	What are the different terms for a definition in the cloud of concepts, in the face of one or more conceptualizations?
CQ2.8	Which is the distribution of terms in the cloud of concepts, in the face of one or more conceptualizations?
CQ3	*How to conceptually characterize ontology networks (its ontologies as a whole)?*
CQ3.1	What is an ontology grounded over another ontology?
CQ3.2	What are sub-ontologies as parts of another ontology?
CQ3.3	What is the reuse of ontologies?
CQ3.4	What is a language that is ontology-driven?
CQ3.5	What is an ontology implemented?

C.3 Packages

Figure C.1 shows the O4OA packages.

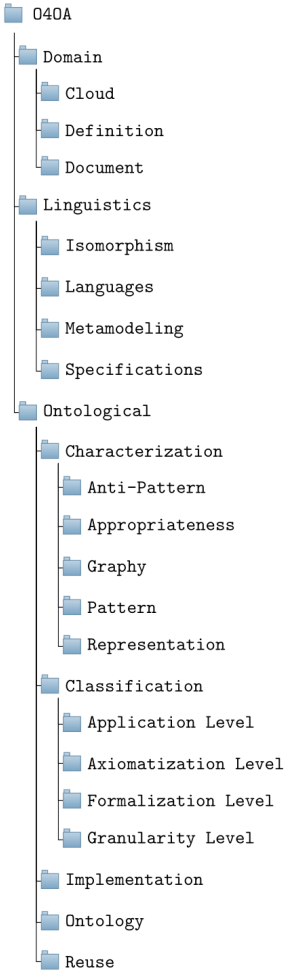


Figure C.1: O4OA packages organization.

C.4 Competence Questions Verification

Table C.3 shows the results of O4OA coverage regarding the proposed CQs.

Table C.3: Results of the O4OA verification.

Ref.	Concepts and Relations
CQ1	*How to conceptually characterize an ontology (as an artifact)?*
CQ1.1	The `Applicability Levels` compounds the intrinsic aspects (<<*mode*>>) that characterize Ontologies Types and make explicit the two disjoint ontology <<*subkind*>>, `Reference Ontology`, and `Operational Ontology`. These modes externally depend on ontology roles when a `Conceptualization` represents a `Domain Description` or an `Implementation` represents an `Ontology Schema` of a domain.
CQ1.2	The `Generality Levels` compounds the intrinsic aspects (<<*mode*>>) that characterize `Ontologies Type` and make explicit these disjoint ontology <<*subkind*>>: `Foundational Ontology`, `Domain Ontology` (including `Core Ontology`), `Task Ontology` and `Application Ontology`.
CQ1.3	The `Formalization Levels` of an ontology is a categorization of the possible aspects (<<*mode*>>) that characterize `Ontologies Type`. The `Expressiveness Levels` is another category of aspects that characterize types of ontologies meanwhile interfere in the `Formalization Levels`.
CQ1.4	The `Axiomatization Levels` of an ontology is a categorization of the possible aspects (<<*mode*>>) that characterize `Ontologies Type`. The `Expressiveness Levels` is another category of aspects that characterize types of ontologies meanwhile interfere in the `Axiomatization Levels`.
CQ1.5	`Well-grounded Ontology` represent ontologies that participate in a relation of grounding through some ontological grounding; i.e. they are grounded over a `Foundational Ontology`.
CQ1.6	`Well-defined Ontology` represents ontologies that participate in a relation of implementation through some ontological support; i.e. there is an `Ontology Schema` (implemented model) that has a correspondent `Domain Description` (reference model) that supports it.
CQ1.7	The intrinsic and relational aspects (characteristics) of ontology types (`Ontology Type`) interfere in ontology characterization, usually these are ontology (meta)characteristics.
CQ1.8	`Ontology Schema` (implemented models) and `Domain Description` (reference models) are mediated by languages, in this case, the roles of languages, `Implementation Language` and `Representation Language` respectively. Therefore, language (meta)characteristics interfere with conceptualizations represented or implemented.
CQ2	*How to conceptually characterize the domain cloud of concepts of an ontology?*
CQ2.1	`Concept Definition` is the <<*relator*>> and its roles (`Concept` and `Source`) using the *Relator Pattern* explain how sources describe concepts in conceptualizations, denoting the relation `describes`.
CQ2.2	`Document` represent the category of consolidated (well-known) bibliographic material (such as standards, policies, and etc.) that may be used as source of information to support the conceptualization.
CQ2.3	The `Cloud of Concepts` collection defines the set of concepts on which it is possible to infer the possible definitions that support one or more ontologies. The relation formed by express which concepts are members of (<<*memberOf*>>) a cloud of concepts collection.
CQ2.4	The <<*relator*>> `Concept Definition` provides all possibilities of sources for conceptualizations since an `Ontology Version` represent conceptualization that contains (<<*componentOf*>>) definitions. Thus, the same source can support multiple conceptualizations.

– *continued on next page*

Table C.3 continued from previous page

Ref.	Concepts and Relations
CQ2.5	The <<*relator*>> `Concept Definition` provides all possibilities of concepts (terms) for conceptualizations since an Ontology Version represent conceptualization that contains (<<*componentOf*>>) definitions. Thus, the same term can appear multiple conceptualizations.
CQ2.6	A `Concept Definition` may (or not) be a component of an `Ontology Version`; therefore, multiple definitions of the same source may be associated with conceptualizations (represented as `Ontology Version`). These definitions may (or may not) be similar, depending on the commitment adopted in each conceptualization.
CQ2.7	A `Concept Definition` may (or not) be a component of an `Ontology Version`; therefore, multiple definitions of the same `Term` (as `Concept`) may be associated with conceptualizations (represented as `Ontology Versions`). These definitions may (or may not) refer synonymous, depending on the commitment adopted in each conceptualization.
CQ2.8	The part-hood relation contains is between a whole conceptualization (represented as `Ontology Version`) and its parts (`Concept`) delimit the cloud of concepts for this conceptualization.
CQ3	*How to conceptually characterize ontology networks (its ontologies as a whole)?*
CQ3.1	`Ontological Grounding` is the relator and its roles (`Well-grounded Ontology` and `Ontological Foundation`) using the *Relator Pattern* explain how ontologies are grounded, denoting the relation `groundedOver`.
CQ3.2	`Composite Ontology` and `Atomic Ontology` are subkinds of `Ontology` and using the *Weak Supplementation Pattern* describe the notion of Whole/Part of ontologies, denoting the relation `componentOf`.
CQ3.3	`Reusability` is the relator and its roles (`Reused Ontology` and `Reuser Ontology`) using the *Relator Pattern* explain ontologies reuse, denoting the relation `reuses`.
CQ3.4	`Ontology-driven Modeling Language` subtypes of languages which have `Ontology-driven Language Specification`. The relator `Ontology-Driven Metamodel` and its roles (`Abstract Ontology-driven Language` and `Ontological Constraints`) using the *Relator Pattern* foundation explains how ontologies provide constraints that drive languages, denoting the relation `drives`.
CQ3.5	`Ontology Schema` is the relator and its roles (`Implementation Language` and `Implementation`) using the *Relator Pattern* explains how ontologies are implemented, denoting the relation `implementationFor`.

Appendix D

F4OC Applied to the Cybersecurity Domain

D.1 API query in MongoO4OA Tracing Concepts

API code functions called trace the concepts from its definitions until the ontologies use them.

```
//...
// Get tree nodes with ontologies
function getOntologyDefinitionNodesByTerm(req,res){
  var definition = new Definition();
  definition.term = req.params.term;

  Definition.aggregate([
  { \$match: { term: definition.term } },
  { \$addFields: {
    node: "\$text",
    type: "definition" }
  },
  { \$project: { _id: 1, node: 1, type: 1 } },
  { \$unionWith: { coll: "terms", pipeline: [
    { \$match: { _id: definition.term } },
    { \$addFields: {
      node: "\$syntax",
      type: "term" }
```

```
19      },
20      { \$project: { _id: 1, node: 1, type: 1 } }
21      ] } },
22    { \$unionWith: { coll: "sources", pipeline: [
23      { \$lookup : {
24        from : "definitions",
25        localField : "_id",
26        foreignField : "source",
27        as : "definitions" }
28      },
29      { \$match: { definitions: { \$elemMatch: { term: definition.term } } } },
30      { \$addFields: {
31        node: "\$label",
32        type: "source" }
33      },
34      { \$project: { _id: 1, node: 1, type: 1 } }
35      ] } },
36    { \$unionWith: { coll: "ontologies", pipeline: [
37      { \$lookup : {
38        from : "definitions",
39        localField : "definitions",
40        foreignField : "_id",
41        as : "definitions" }
42      },
43      { \$match: { definitions: { \$elemMatch: { term: definition.term } } } },
44      { \$addFields: {
45        node: "\$name",
46        type: "ontology" }
47      },
48      { \$project: { _id: 1, node: 1, type: 1 } }
49      ] } }
50    ]).exec((err,diagramNodeData) => {
51      if(err) return res.status(500).send({message: 'Incorrect request.'});
52
53      return res.status(200).send({diagramNodeData});
54    });
55  }
56
57  // Get tree links with ontologies
58  function getOntologyDefinitionLinksByTerm(req,res){
59    var definition = new Definition();
60    definition.term = req.params.term;
61
62    Definition.aggregate([
63    { \$match: { term: definition.term } },
64    { \$addFields: {
65      from: "\$_id",
66      to: "\$term",
67      link: "definition for" }
68    },
69    { \$project: { _id: 0, from: 1, to: 1, link: 1 } },
70    { \$unionWith: { coll: "definitions", pipeline: [
71      { \$match: { term: definition.term } },
```

```
72      { \$addFields: {
73        from: "\$source",
74        to: "\$_id",
75        link: "defines" }
76      },
77      { \$project: { _id: 0, from: 1, to: 1, link: 1 } }
78      ] } },
79    { \$unionWith: { coll: "ontologies", pipeline: [
80      { \$lookup : {
81        from : "definitions",
82        localField : "definitions",
83        foreignField : "_id",
84        as : "definitions" }
85      },
86      { \$match: { definitions: { \$elemMatch: { term: definition.term } } } },
87      { \$unwind: "\$definitions" },
88      { \$addFields: {
89        from: "\$definitions._id",
90        to: "\$_id",
91        link: "concept for" }
92      },
93      { \$project: { _id: 0, from: 1, to: 1, link: 1 } }
94      ] } }
95    ]).exec((err,diagramLinkData) => {
96      if(err) return res.status(500).send({message: 'Incorrect request.'});
97
98      return res.status(200).send({diagramLinkData});
99    });
100 }
101 //...
```

D.2 API query in MongoO4OA Tracing the Concept of Vulnerability

The JSON result from the API call that traces the concept of *Vulnerability* from its definitions until the ontologies use them.

```
1  {
2    "diagramNodeData": [
3    {
4      "_id": "62eaa61726118d27d0b9b465",
5      "node": "A flaw in a software, firmware, hardware, or service component
6      resulting from a weakness that can be exploited, causing a negative
7      impact to the confidentiality, integrity, or availability of an impacted
8      component or components.",
9      "type": "definition"
10   },
11   {
12     "_id": "62eaadee26118d27d0b9b473",
13     "node": "Weakness of an asset or control (3.14) that can be exploited by
```

```
    one or more threats (3.74)",
    "type": "definition"
  },
  {
    "_id": "62eaaf1b26118d27d0b9b478",
    "node": "A vulnerability is a weakness of an asset or control that can be
    exploited by a threat [ISO/IEC 27000:2009]. Within the context of an
    information system, ISO/IEC TR 19791:2006 also defines vulnerability as a
    flaw, weakness or property of the design or implementation of an
    information system (including its security controls) or its environment
    that could be intentionally or unintentionally exploited to adversely
    affect an organization's assets or operations. NOTE, ISO/IEC 27005
    provides guidelines on identifying vulnerabilities.",
    "type": "definition"
  },
  {
    "_id": "62eab0ed26118d27d0b9b47f",
    "node": "Vulnerability (aligned with [b-ITU-T X.800]): Any weakness that
    could be exploited to violate a system or the information it contains.",
    "type": "definition"
  },
  {
    "_id": "62eab81e26118d27d0b9b4d5",
    "node": "A vulnerability is a weakness in a system, system security
    procedure, internal controls, or implementation that could be exploited
    by a threat source (Threat Source - The intent and method targeted at the
    intentional exploitation of a vulnerability or a situation and method
    that may accidentally exploit a vulnerability). Vulnerabilities leave
    systems susceptible to a multitude of activities that can result in
    significant and sometimes irreversible losses to an individual, group, or
    organization. These losses can range from a single damaged file on a
    laptop computer or mobile device to entire databases at an operations
    center being compromised. With the right tools and knowledge, an
    adversary can exploit system vulnerabilities and gain access to the
    information stored on them. The damage inflicted on compromised systems
    can vary depending on the threat
    source.",
    "type": "definition"
  },
  {
    "_id": "62f24f2f8f5fe43590d12e16",
    "node": "Vulnerability an occurrence of a weakness (or multiple
    weaknesses) within a product, in which the weakness can be used by a
    party to cause the product to modify or access unintended data, interrupt
    proper execution, or perform incorrect actions that were not specifically
    granted to the party who uses the weakness.",
    "type": "definition"
  },
  {
    "_id": "5eee523ad541e23b1e3855ce",
    "node": "vulnerability",
    "type": "term"
  },
```

```
  {
    "_id": "5ef0920097fc002808331b01",
    "node": "ISOIEC270322012",
    "type": "source"
  },
  {
    "_id": "5ef09d8197fc002808331b04",
    "node": "ISOIEC270002018",
    "type": "source"
  },
  {
    "_id": "5ef09eac97fc002808331b08",
    "node": "RecITU-T-X1500",
    "type": "source"
  },
  {
    "_id": "5ef10be571196615684c0a68",
    "node": "NIST-SP-800-12Revision1",
    "type": "source"
  },
  {
    "_id": "5ef1cb9cf5ff54170cfa1408",
    "node": "CVE-1999-0001",
    "type": "source"
  },
  {
    "_id": "63db7eba2cd3a82a72beca73",
    "node": "CWE",
    "type": "source"
  },
  {
    "_id": "62d69c7724b7e52c3819aa72",
    "node": "CVE",
    "type": "ontology"
  },
  {
    "_id": "62d69d3a24b7e52c3819aa78",
    "node": "CWE",
    "type": "ontology"
  } ]
}

{
  "diagramLinkData": [
  {
    "from": "62eaa61726118d27d0b9b465",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
  },
  {
    "from": "62eaadee26118d27d0b9b473",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
```

```
  },
  {
    "from": "62eaaf1b26118d27d0b9b478",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
  },
  {
    "from": "62eab0ed26118d27d0b9b47f",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
  },
  {
    "from": "62eab81e26118d27d0b9b4d5",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
  },
  {
    "from": "62f24f2f8f5fe43590d12e16",
    "to": "5eee523ad541e23b1e3855ce",
    "link": "definition for"
  },
  {
    "from": "5ef1cb9cf5ff54170cfa1408",
    "to": "62eaa61726118d27d0b9b465",
    "link": "defines"
  },
  {
    "from": "5ef09d8197fc002808331b04",
    "to": "62eaadee26118d27d0b9b473",
    "link": "defines"
  },
  {
    "from": "5ef0920097fc002808331b01",
    "to": "62eaaf1b26118d27d0b9b478",
    "link": "defines"
  },
  {
    "from": "5ef09eac97fc002808331b08",
    "to": "62eab0ed26118d27d0b9b47f",
    "link": "defines"
  },
  {
    "from": "5ef10be571196615684c0a68",
    "to": "62eab81e26118d27d0b9b4d5",
    "link": "defines"
  },
  {
    "from": "63db7eba2cd3a82a72beca73",
    "to": "62f24f2f8f5fe43590d12e16",
    "link": "defines"
  },
  {
    "from": "62eaa61726118d27d0b9b465",
```

```
    "to": "62d69c7724b7e52c3819aa72",
    "link": "concept for"
  },
  {
    "from": "62f24c3a8f5fe43590d12e0a",
    "to": "62d69d3a24b7e52c3819aa78",
    "link": "concept for"
  },
  {
    "from": "62f24ee18f5fe43590d12e11",
    "to": "62d69d3a24b7e52c3819aa78",
    "link": "concept for"
  },
  {
    "from": "62f24f2f8f5fe43590d12e16",
    "to": "62d69d3a24b7e52c3819aa78",
    "link": "concept for"
  } ]
}
```


D.3 API query in MongoO4OA Tracing the Concept of Risk

The JSON result from the API call that traces the concept of *Risk* from its definitions until the ontologies use them.


```
{
  "diagramNodeData": [
  {
    "_id": "600eeb4ffdaefa069006421f",
    "node": "RIESGO Efecto de la incertidumbre sobre la consecucion de los
    objetivos. [UNE-ISO GUIA 73:2010] NOTA 1 Un efecto es una desviacion,
    positiva y/o negativa, respecto a lo previsto. NOTA 2 La incertidumbre es
    el estado, incluso parcial, de deficiencia en la informacion relativa a
    la comprension o al conocimiento de un suceso, de sus consecuencias o de
    su probabilidad. NOTA 3 Con frecuencia, el riesgo se caracteriza por
    referencia a sucesos potenciales y a sus consecuencias o una combinacion
    de ambas NOTA 4 Con frecuencia, el riesgo se expresa en terminos de
    combinacion de las consecuencias de un suceso (incluyendo los cambios en
    las circunstancias) y de su probabilidad. NOTA 5: En el contexto de
    sistemas de gestion de la seguridad de la informacion, los riesgos de
    seguridad de la informacion se pueden expresar como el efecto de la
    incertidumbre sobre los objetivos de seguridad de la informacion. NOTA 6:
    El riesgo de seguridad de la informacion se relaciona con la posibilidad
    de que las amenazas exploten vulnerabilidades de un activo o grupo de
    activos de informacion y causen dano a una organizacion.
    [UNE-ISO/IEC 27000:2014] ",
    "type": "definition"
  },
  {
    "_id": "600ef803e6ae1b0b54bd6046",
    "node": "RIESGO Estimacion del grado de exposicion a que una amenaza se
```

```
    materialice sobre uno o mas activos causando danos o perjuicios a la
    organizacion. [UNE-71504:2008] ",
    "type": "definition"
  },
  {
    "_id": "600ef85fe6ae1b0b54bd6047",
    "node": "RIESGO Efecto de la incertidumbre sobre la consecucion de los
    objetivos. NOTA 4. Con frecuencia, el riesgo se expresa en terminos de
    combinacion de las consecuencias de un suceso (incluyendo los cambios
    en las circunstancias) y de su probabilidad. [ISO Guia 73:2010]",
    "type": "definition"
  },
  {
    "_id": "600ef8cce6ae1b0b54bd6048",
    "node": "RIESGO Un posible Evento que podria causar dano o perdidas, o
    afectar la habilidad de alcanzar Objetivos. Un Riesgo es medido por la
    probabilidad de una Amenaza, la Vulnerabilidad del Activo a esa Amenaza,
    y por el Impacto que tendria en caso que ocurriera. [ITIL:2007]",
    "type": "definition"
  },
  {
    "_id": "600ef923e6ae1b0b54bd6049",
    "node": "RIESGO Estimacion del grado de exposicion a que una amenaza se
    materialice sobre uno o mas activos causando danos o perjuicios a la
    Organizacion. [Magerit:2012]",
    "type": "definition"
  },
  {
    "_id": "600ef949e6ae1b0b54bd604a",
    "node": "RIESGO Probabilidad de que una amenaza se materialice
    aprovechando una vulnerabilidad causando dano (impacto) en un proceso o
    sistema. [CCN-STIC-401:2007]",
    "type": "definition"
  },
  {
    "_id": "600ef97de6ae1b0b54bd604b",
    "node": "RIESGO El potencial de que una amenaza especifica explote las
    debilidades de un activo o grupo de activos para ocasionar perdida y/o
    dano a los activos. Por lo general se mide por medio de una combinacion
    del impacto y la probabilidad de ocurrencia. [COBIT:2006]",
    "type": "definition"
  },
  {
    "_id": "600ef99fe6ae1b0b54bd604c",
    "node": "RIESGO Probabilidad de que una vulnerabilidad propia de un
    sistema de informacion sea explotada por las amenazas a dicho sistema,
    con el objetivo de penetrarlo. [CESID:1997]",
    "type": "definition"
  },
  {
    "_id": "600efa56e6ae1b0b54bd604e",
    "node": "Risk The potential business impact and likelihood of particular
    threats materialising - and the application of controls to mitigate risks
```

```
    to acceptable levels.",
    "type": "definition"
  },
  {
    "_id": "600efc9ce6ae1b0b54bd6053",
    "node": "risk effect of uncertainty on objectives (3.49) Note 1 to entry:
    An effect is a deviation from the expected - positive or negative. Note 2
    to entry: Uncertainty is the state, even partial, of deficiency of
    information related to, understanding or knowledge of, an event, its
    consequence, or likelihood. Note 3 to entry: Risk is often characterized
    by reference to potential ``events'' (as defined in ISO Guide 73:2009,
    3.5.1.3) and ``consequence'' (as defined in ISO Guide 73:2009, 3.6.1.3),
    or a combination of these. Note 4 to entry: Risk is often expressed in
    terms of a combination of the consequences of an event (including changes
    in circumstances) and the associated ``likelihood'' (as defined in ISO
    Guide 73:2009, 3.6.1.1) of occurrence. Note 5 to entry: In the context of
    information security management systems, information security risks can
    be expressed as effect of uncertainty on information security objectives.
    Note 6 to entry: Information security risk is associated with the
    potential that threats will exploit vulnerabilities of an information
    asset or group of information assets and thereby cause harm to an
    organization.",
    "type": "definition"
  },
  {
    "_id": "600efdc7e6ae1b0b54bd6055",
    "node": "Risk A measure of the extent to which an entity is threatened by
    a potential circumstance or event, and typically a function of: (i) the
    adverse impacts that would arise if the circumstance or event occurs; and
    (ii) the likelihood of occurrence.",
    "type": "definition"
  },
  {
    "_id": "600efe1ce6ae1b0b54bd6057",
    "node": "Risk A measure of the extent to which an entity is threatened by
    a potential circumstance or event, and typically a function of: (i) the
    adverse impacts that would arise if the circumstance or event occurs; and
    (ii) the likelihood of occurrence.",
    "type": "definition"
  },
  {
    "_id": "600effbbe6ae1b0b54bd6059",
    "node": "Risk A measure of the extent to which an entity is threatened by
    a potential circumstance or event, and typically a function of: (i) the
    adverse impacts that would arise if the circumstance or event occurs; and
    (ii) the likelihood of occurrence. [Note: System-related security risks
    are those risks that arise from the loss of confidentiality, integrity,
    or availability of information or systems and reflect the potential
    adverse impacts to organizational operations (including mission,
    functions, image, or reputation), organizational assets, individuals,
    other organizations, and the Nation. Adverse impacts to the Nation
    include, for example, compromises to systems that support critical
    infrastructure applications or are paramount to government continuity of
```

```
    operations as defined by the Department of Homeland Security.]
    SOURCE: SP 800-37",
    "type": "definition"
  },
  {
    "_id": "600f00abe6ae1b0b54bd605b",
    "node": "Risk A measure of the extent to which an entity is threatened by
    a potential circumstance or event, and typically is a function of: (i)
    the adverse impact, or magnitude of harm, that would arise if the
    circumstance or event occurs; and (ii) the likelihood of occurrence.",
    "type": "definition"
  },
  {
    "_id": "600f0192e6ae1b0b54bd605d",
    "node": "Risk [FIPS 200, Adapted] A measure of the extent to which an
    entity is threatened by a potential circumstance or event, and typically
    a function of: (i) the adverse impacts that would arise if the
    circumstance or event occurs; and (ii) the likelihood of occurrence.
    Information system-related security risks are those risks that arise from
    the loss of confidentiality, integrity, or availability of information or
    information systems and reflect the potential adverse impacts to
    organizational operations (including mission, functions, image, or
    reputation), organizational assets, individuals, other organizations, and
    the Nation.",
    "type": "definition"
  },
  {
    "_id": "600f024de6ae1b0b54bd605f",
    "node": "Risk The level of impact on agency operations (including
    mission, functions, image, or reputation), agency assets, or individuals
    resulting from the operation of an information system, given the
    potential impact of a threat and the likelihood of that threat occurring.
    SOURCE: NIST SP 800-30 [79]",
    "type": "definition"
  },
  {
    "_id": "600f02cae6ae1b0b54bd6061",
    "node": "Risk is the possibility of harm or loss to any software,
    information, hardware, administrative, physical, communications, or
    personnel resource within an automated information system or activity.",
    "type": "definition"
  },
  {
    "_id": "600f5a13d289480c60440184",
    "node": "(Risk). The risk is the probability that a successful attack
    occurs.",
    "type": "definition"
  },
  {
    "_id": "5eee523ad541e23b1e3855cb",
    "node": "risk",
    "type": "term"
  },
```

```
  {
    "_id": "5ef09d8197fc002808331b04",
    "node": "ISOIEC270002018",
    "type": "source"
  },
  {
    "_id": "5ef104b071196615684c0a50",
    "node": "NIST-CSWP-04162018",
    "type": "source"
  },
  {
    "_id": "5ef1053f71196615684c0a51",
    "node": "NIST-CSWP-04162014",
    "type": "source"
  },
  {
    "_id": "5ef108ed71196615684c0a5c",
    "node": "Chaplin2011",
    "type": "source"
  },
  {
    "_id": "5ef109f471196615684c0a5f",
    "node": "SpecPubl800-26",
    "type": "source"
  },
  {
    "_id": "5ef10be571196615684c0a68",
    "node": "NIST-SP-800-12Revision1",
    "type": "source"
  },
  {
    "_id": "5ef10c1a71196615684c0a69",
    "node": "NIST.SP.800-82r2",
    "type": "source"
  },
  {
    "_id": "5ef10c6671196615684c0a6a",
    "node": "NIST800-37Revision2",
    "type": "source"
  },
  {
    "_id": "5ef10ccd71196615684c0a6c",
    "node": "NIST800-53Rev4",
    "type": "source"
  },
  {
    "_id": "5ef10e2271196615684c0a70",
    "node": "CCN-STIC-401",
    "type": "source"
  },
  {
    "_id": "600f59aed289480c60440183",
    "node": "Schumacher2003",
```

```
    "type": "source"
  },
  {
    "_id": "61b91af88683481c1cdfc624",
    "node": "CRATELO",
    "type": "ontology"
  } ]
}

{
  "diagramLinkData": [
  {
    "from": "600eeb4ffdaefa069006421f",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef803e6ae1b0b54bd6046",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef85fe6ae1b0b54bd6047",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef8cce6ae1b0b54bd6048",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef923e6ae1b0b54bd6049",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef949e6ae1b0b54bd604a",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef97de6ae1b0b54bd604b",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600ef99fe6ae1b0b54bd604c",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600efa56e6ae1b0b54bd604e",
```

```
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600efc9ce6ae1b0b54bd6053",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600efdc7e6ae1b0b54bd6055",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600efe1ce6ae1b0b54bd6057",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600effbbe6ae1b0b54bd6059",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600f00abe6ae1b0b54bd605b",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600f0192e6ae1b0b54bd605d",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600f024de6ae1b0b54bd605f",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600f02cae6ae1b0b54bd6061",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "600f5a13d289480c60440184",
    "to": "5eee523ad541e23b1e3855cb",
    "link": "definition for"
  },
  {
    "from": "5ef10e2271196615684c0a70",
    "to": "600eeb4ffdaefa069006421f",
    "link": "defines"
  },
```

```
345   {
346     "from": "5ef10e2271196615684c0a70",
347     "to": "600ef803e6ae1b0b54bd6046",
348     "link": "defines"
349   },
350   {
351     "from": "5ef10e2271196615684c0a70",
352     "to": "600ef85fe6ae1b0b54bd6047",
353     "link": "defines"
354   },
355   {
356     "from": "5ef10e2271196615684c0a70",
357     "to": "600ef8cce6ae1b0b54bd6048",
358     "link": "defines"
359   },
360   {
361     "from": "5ef10e2271196615684c0a70",
362     "to": "600ef923e6ae1b0b54bd6049",
363     "link": "defines"
364   },
365   {
366     "from": "5ef10e2271196615684c0a70",
367     "to": "600ef949e6ae1b0b54bd604a",
368     "link": "defines"
369   },
370   {
371     "from": "5ef10e2271196615684c0a70",
372     "to": "600ef97de6ae1b0b54bd604b",
373     "link": "defines"
374   },
375   {
376     "from": "5ef10e2271196615684c0a70",
377     "to": "600ef99fe6ae1b0b54bd604c",
378     "link": "defines"
379   },
380   {
381     "from": "5ef108ed71196615684c0a5c",
382     "to": "600efa56e6ae1b0b54bd604e",
383     "link": "defines"
384   },
385   {
386     "from": "5ef09d8197fc002808331b04",
387     "to": "600efc9ce6ae1b0b54bd6053",
388     "link": "defines"
389   },
390   {
391     "from": "5ef1053f71196615684c0a51",
392     "to": "600efdc7e6ae1b0b54bd6055",
393     "link": "defines"
394   },
395   {
396     "from": "5ef104b071196615684c0a50",
397     "to": "600efe1ce6ae1b0b54bd6057",
```

```
398     "link": "defines"
399   },
400   {
401     "from": "5ef10be571196615684c0a68",
402     "to": "600effbbe6ae1b0b54bd6059",
403     "link": "defines"
404   },
405   {
406     "from": "5ef10c6671196615684c0a6a",
407     "to": "600f00abe6ae1b0b54bd605b",
408     "link": "defines"
409   },
410   {
411     "from": "5ef10ccd71196615684c0a6c",
412     "to": "600f0192e6ae1b0b54bd605d",
413     "link": "defines"
414   },
415   {
416     "from": "5ef10c1a71196615684c0a69",
417     "to": "600f024de6ae1b0b54bd605f",
418     "link": "defines"
419   },
420   {
421     "from": "5ef109f471196615684c0a5f",
422     "to": "600f02cae6ae1b0b54bd6061",
423     "link": "defines"
424   },
425   {
426     "from": "600f59aed289480c60440183",
427     "to": "600f5a13d289480c60440184",
428     "link": "defines"
429   },
430   {
431     "from": "600f5a13d289480c60440184",
432     "to": "61b91af88683481c1cdfc624",
433     "link": "concept for"
434   } ]
435 }
```

D.4 API query in MongoO4OA Tracing BRON Relations

```
1  ...
2  // Get the tree of sub-ontologies from an ontology from parent to children
3  function getSubOntologiesTreeTopDown(req,res){
4
5    if (req.params.id) {
6      var ontology = new Ontology();
7      ontology._id = req.params.id;
8
9      Ontology.aggregate([
10     { $match: { _id : ontology._id } },
```

```
11      { $graphLookup: {
12        from: "ontologies",
13        startWith: "$subOntologies",
14        connectFromField: "subOntologies",
15        connectToField: "_id",
16        as: "children" ,
17        depthField: "depth" } },
18      { $project: {
19        _id: 1,
20        name: 1,
21        domain: 1,
22        classification: 1,
23        depth: 1,
24        superOntologies: 1,
25        children: {
26          _id: 1,
27          name: 1,
28          domain: 1,
29          classification: 1,
30          depth: 1,
31          superOntologies: 1 } } },
32      { $unwind: { path: "$children", "preserveNullAndEmptyArrays": true } },
33      { $unwind: { path: "$children.superOntologies",
34          "preserveNullAndEmptyArrays": true } },
35      { $lookup: {
36        from: "applicationlevels",
37        localField: "classification.applicationLevel",
38        foreignField: "_id",
39        as: "classification.applicationLevel" }
40      },
41      { $unwind: { path: "$classification.applicationLevel",
42          preserveNullAndEmptyArrays: true } },
43      { $lookup: {
44        from: "generalitylevels",
45        localField: "classification.generalityLevel",
46        foreignField: "_id",
47        as: "classification.generalityLevel" }
48      },
49      { $unwind: { path: "$classification.generalityLevel",
50        preserveNullAndEmptyArrays: true } },
51      { $lookup: {
52        from: "formalizationlevels",
53        localField: "classification.formalizationLevel",
54        foreignField: "_id",
55        as: "classification.formalizationLevel" }
56      },
57      { $unwind: { path: "$classification.formalizationLevel",
58        preserveNullAndEmptyArrays: true } },
59      { $lookup: {
60        from: "axiomatizationlevels",
61        localField: "classification.axiomatizationLevel",
62        foreignField: "_id",
63        as: "classification.axiomatizationLevel" }
```

```
64      },
65      { $unwind: { path: "$classification.axiomatizationLevel",
66        preserveNullAndEmptyArrays: true } },
67      { $lookup: {
68        from: "applicationlevels",
69        localField: "children.classification.applicationLevel",
70        foreignField: "_id",
71        as: "children.classification.applicationLevel" }
72      },
73      { $unwind: { path: "$children.classification.applicationLevel",
74          preserveNullAndEmptyArrays: true } },
75      { $lookup: {
76        from: "generalitylevels",
77        localField: "children.classification.generalityLevel",
78        foreignField: "_id",
79        as: "children.classification.generalityLevel" }
80      },
81      { $unwind: { path: "$children.classification.generalityLevel",
82        preserveNullAndEmptyArrays: true } },
83      { $lookup: {
84        from: "formalizationlevels",
85        localField: "children.classification.formalizationLevel",
86        foreignField: "_id",
87        as: "children.classification.formalizationLevel" }
88      },
89      { $unwind: { path: "$children.classification.formalizationLevel",
90        preserveNullAndEmptyArrays: true } },
91      { $lookup: {
92        from: "axiomatizationlevels",
93        localField: "children.classification.axiomatizationLevel",
94        foreignField: "_id",
95        as: "children.classification.axiomatizationLevel" }
96      },
97      { $unwind: { path: "$children.classification.axiomatizationLevel",
98        preserveNullAndEmptyArrays: true } },
99      { $sort: { "children.depth": -1 } },
100     { $group: {
101       _id: "$_id",
102       superOntologies: { $first: "$superOntologies"},
103       name: { $first: "$name"},
104       children: { $push: "$children"} } },
105     { $addFields: { children: { $reduce: {
106       input: "$children",
107       initialValue: {
108         currentDepth: -1,
109         currentDepthSub: [],
110         previousDepthSub: [] },
111       in: { $let: { vars: {
112         prev: { $cond: [ {
113           $eq: [ "$$value.currentDepth", "$$this.depth"] },
114           "$$value.previousDepthSub",
115           "$$value.currentDepthSub"] },
116         current: { $cond: [ {
```

```
          $eq: [ "$$value.currentDepth", "$$this.depth"] },
          "$$value.currentDepthSub",
          [] ] } },
      in: {
        currentDepth: "$$this.depth",
        previousDepthSub: "$$prev",
        currentDepthSub: { $concatArrays: [
          "$$current",
          [ { $mergeObjects: [
            "$$this",
            { children: {
            $filter: { input: "$$prev", as: "e", cond: { $eq:
            [ "$$e.superOntologies", "$$this._id"] } } } } ] }
          ] ] } }
      } } } } } },
    { $addFields: { children: "$children.currentDepthSub"} }
    ]).exec((err,ontologies) => {
      if(err) return res.status(500).send({message: 'Incorrect request.'});
      if(!ontologies) return res.status(404).send(
      {message: 'No ontology found.'});

      return res.status(200).send({ontologies});
    });
  } else {
    return res.status(200).send(
      {message: 'Inform all mandatory fields of the request.'});
  }
}
...
```

```
{
  "ontologies": [
  {
    "_id": "62aa22183cd52f2e48a47d05",
    "superOntologies": null,
    "name": "BRON",
    "children": [
    {
      "_id": "62d6a66424b7e52c3819aa7e",
      "name": "CAPEC",
      "domain": "Cybersecurity",
      "classification": {
        "applicationLevel": {
          "_id": "605f92a275b2e96017664770",
          "name": "Operational Ontology",
          "children": [],
          "edge": {
            "label": ""
          },
          "parent": "605f92a275b2e96017664772"
        },
        "wellDefined": false,
        "generalityLevel": {
          "_id": "61099c21e18238d81d2f8280",
```

```
          "name": "Domain Ontology",
          "edge": {
            "label": ""
          },
          "children": [],
          "parent": "61099c21e18238d81d2f8283"
        },
        "wellGrounded": false,
        "formalizationLevel": {
          "_id": "60637d0810afc9096eeb81fb",
          "name": "XML, DTDs",
          "axiomatization": "606c78f12bce7070e3b08ab8",
          "children": [],
          "edge": {
            "label": "",
            "max": 43.75,
            "min": 37.51
          },
          "parent": "60637d0810afc9096eeb81fe"
        },
        "formalization": 42,
        "axiomatizationLevel": {
          "_id": "606c78f12bce7070e3b08ab8",
          "name": "Lightweight Ontology",
          "children": [],
          "edge": {
            "label": "",
            "max": 50,
            "min": 0
          },
          "parent": "606c78f12bce7070e3b08aba"
        },
        "axiomatization": 20
      },
      "depth": 0,
      "children": []
    },
    {
      "_id": "62d6f9bf93533d2b64c2957f",
      "name": "ATT&CK",
      "domain": "Cybersecurity",
      "classification": {
        "applicationLevel": {
          "_id": "605f92a275b2e96017664770",
          "name": "Operational Ontology",
          "children": [],
          "edge": {
            "label": ""
          },
          "parent": "605f92a275b2e96017664772"
        },
        "wellDefined": false,
        "generalityLevel": {
```

```
          "_id": "61099c21e18238d81d2f827f",
          "name": "Task Ontology",
          "edge": {
            "label": ""
          },
          "children": [],
          "parent": "61099c21e18238d81d2f8283"
        },
        "wellGrounded": false,
        "formalizationLevel": {
          "_id": "60637d0810afc9096eeb81f7",
          "name": "XML Schema",
          "axiomatization": "606c78f12bce7070e3b08ab8",
          "children": [],
          "edge": {
            "label": "",
            "max": 62.5,
            "min": 56.26
          },
          "parent": "60637d0810afc9096eeb81f8"
        },
        "formalization": 64,
        "axiomatizationLevel": {
          "_id": "606c78f12bce7070e3b08ab8",
          "name": "Lightweight Ontology",
          "children": [],
          "edge": {
            "label": "",
            "max": 50,
            "min": 0
          },
          "parent": "606c78f12bce7070e3b08aba"
        },
        "axiomatization": 28
      },
      "superOntologies": "62aa22183cd52f2e48a47d05",
      "depth": 0,
      "children": []
    },
    {
      "_id": "62d69c7724b7e52c3819aa72",
      "name": "CVE",
      "domain": "Cybersecurity",
      "classification": {
        "applicationLevel": {
          "_id": "605f92a275b2e96017664770",
          "name": "Operational Ontology",
          "children": [],
          "edge": {
            "label": ""
          },
          "parent": "605f92a275b2e96017664772"
        },
```

```
        "wellDefined": false,
        "generalityLevel": {
          "_id": "61099c21e18238d81d2f8280",
          "name": "Domain Ontology",
          "edge": {
            "label": ""
          },
          "children": [],
          "parent": "61099c21e18238d81d2f8283"
        },
        "wellGrounded": false,
        "formalizationLevel": {
          "_id": "60637d0810afc9096eeb81fb",
          "name": "XML, DTDs",
          "axiomatization": "606c78f12bce7070e3b08ab8",
          "children": [],
          "edge": {
            "label": "",
            "max": 43.75,
            "min": 37.51
          },
          "parent": "60637d0810afc9096eeb81fe"
        },
        "formalization": 40,
        "axiomatizationLevel": {
          "_id": "606c78f12bce7070e3b08ab8",
          "name": "Lightweight Ontology",
          "children": [],
          "edge": {
            "label": "",
            "max": 50,
            "min": 0
          },
          "parent": "606c78f12bce7070e3b08aba"
        },
        "axiomatization": 18
      },
      "depth": 0,
      "children": []
    },
    {
      "_id": "62d69d3a24b7e52c3819aa78",
      "name": "CWE",
      "domain": "Cybersecurity",
      "classification": {
        "applicationLevel": {
          "_id": "605f92a275b2e96017664770",
          "name": "Operational Ontology",
          "children": [],
          "edge": {
            "label": ""
          },
          "parent": "605f92a275b2e96017664772"
```

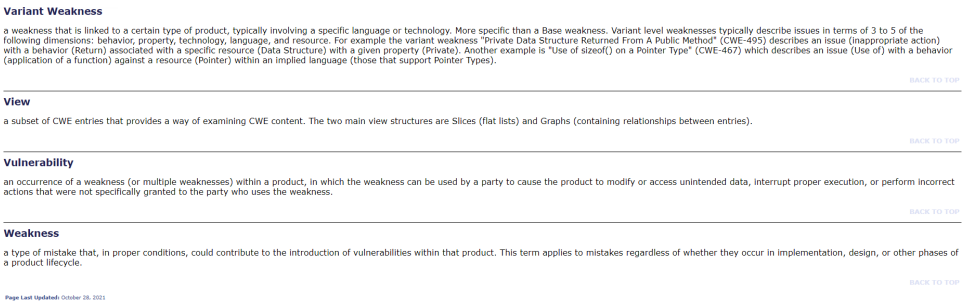
```
        },
        "wellDefined": false,
        "generalityLevel": {
          "_id": "61099c21e18238d81d2f8280",
          "name": "Domain Ontology",
          "edge": {
            "label": ""
          },
          "children": [],
          "parent": "61099c21e18238d81d2f8283"
        },
        "wellGrounded": false,
        "formalizationLevel": {
          "_id": "60637d0810afc9096eeb81fb",
          "name": "XML, DTDs",
          "axiomatization": "606c78f12bce7070e3b08ab8",
          "children": [],
          "edge": {
            "label": "",
            "max": 43.75,
            "min": 37.51
          },
          "parent": "60637d0810afc9096eeb81fe"
        },
        "formalization": 40,
        "axiomatizationLevel": {
          "_id": "606c78f12bce7070e3b08ab8",
          "name": "Lightweight Ontology",
          "children": [],
          "edge": {
            "label": "",
            "max": 50,
            "min": 0
          },
          "parent": "606c78f12bce7070e3b08aba"
        },
        "axiomatization": 18
      },
      "depth": 0,
      "children": []
    },
    {
      "_id": "62d6fa4693533d2b64c29585",
      "name": "D3FEND",
      "domain": "Cybersecurity",
      "classification": {
        "applicationLevel": {
          "_id": "605f92a275b2e96017664770",
          "name": "Operational Ontology",
          "children": [],
          "edge": {
            "label": ""
          },
```

```
          "parent": "605f92a275b2e96017664772"
        },
        "wellDefined": false,
        "generalityLevel": {
          "_id": "61099c21e18238d81d2f8280",
          "name": "Domain Ontology",
          "edge": {
            "label": ""
          },
          "children": [],
          "parent": "61099c21e18238d81d2f8283"
        },
        "wellGrounded": false,
        "formalizationLevel": {
          "_id": "60637d0810afc9096eeb81f3",
          "name": "Lightweight Ontology",
          "axiomatization": "606c78f12bce7070e3b08ab8",
          "children": [],
          "edge": {
            "label": "",
            "max": 81.25,
            "min": 75.01
          },
          "parent": "60637d0810afc9096eeb81f5"
        },
        "formalization": 76,
        "axiomatizationLevel": {
          "_id": "606c78f12bce7070e3b08ab8",
          "name": "Lightweight Ontology",
          "children": [],
          "edge": {
            "label": "",
            "max": 50,
            "min": 0
          },
          "parent": "606c78f12bce7070e3b08aba"
        },
        "axiomatization": 42
      },
      "superOntologies": "62aa22183cd52f2e48a47d05",
      "depth": 0,
      "children": [
      {
        "_id": "62f2c51e8f5fe43590d12eea",
        "name": "DAO",
        "domain": "Cybersecurity",
        "classification": {
          "applicationLevel": {
            "_id": "605f92a275b2e96017664770",
            "name": "Operational Ontology",
            "children": [],
            "edge": {
              "label": ""
```

```
            },
            "parent": "605f92a275b2e96017664772"
          },
          "wellDefined": false,
          "generalityLevel": {
            "_id": "61099c21e18238d81d2f8280",
            "name": "Domain Ontology",
            "edge": {
              "label": ""
            },
            "children": [],
            "parent": "61099c21e18238d81d2f8283"
          },
          "wellGrounded": false,
          "formalizationLevel": {
            "_id": "60637d0810afc9096eeb81f3",
            "name": "Lightweight Ontology",
            "axiomatization": "606c78f12bce7070e3b08ab8",
            "children": [],
            "edge": {
              "label": "",
              "max": 81.25,
              "min": 75.01
            },
            "parent": "60637d0810afc9096eeb81f5"
          },
          "formalization": 76,
          "axiomatizationLevel": {
            "_id": "606c78f12bce7070e3b08ab8",
            "name": "Lightweight Ontology",
            "children": [],
            "edge": {
              "label": "",
              "max": 50,
              "min": 0
            },
            "parent": "606c78f12bce7070e3b08aba"
          },
          "axiomatization": 42
        },
        "superOntologies": "62d6fa4693533d2b64c29585",
        "depth": 1,
        "children": []
      }]
    }]
  }]
}
```

D.5 CVE Glossary about the Concept of Vulnerability

Figure D.1 shows the definition of the *Vulnerability*



Variant Weakness

a weakness that is linked to a certain type of product, typically involving a specific language or technology. More specific than a Base weakness. Variant level weaknesses typically describe issues in terms of 3 to 5 of the following dimensions: behavior, property, technology, language, and resource. For example the variant weakness "Private Data Structure Returned From A Public Method" (CWE-495) describes an issue (inappropriate action) with a behavior (Return) associated with a specific resource (Data Structure) with a given property (Private). Another example is "Use of sizeof() on a Pointer Type" (CWE-467) which describes an issue (Use of) with a behavior (application of a function) against a resource (Pointer) within an implied language (those that support Pointer Types).

BACK TO TOP

View

a subset of CWE entries that provides a way of examining CWE content. The two main view structures are Slices (flat lists) and Graphs (containing relationships between entries).

BACK TO TOP

Vulnerability

an occurrence of a weakness (or multiple weaknesses) within a product, in which the weakness can be used by a party to cause the product to modify or access unintended data, interrupt proper execution, or perform incorrect actions that were not specifically granted to the party who uses the weakness.

BACK TO TOP

Weakness

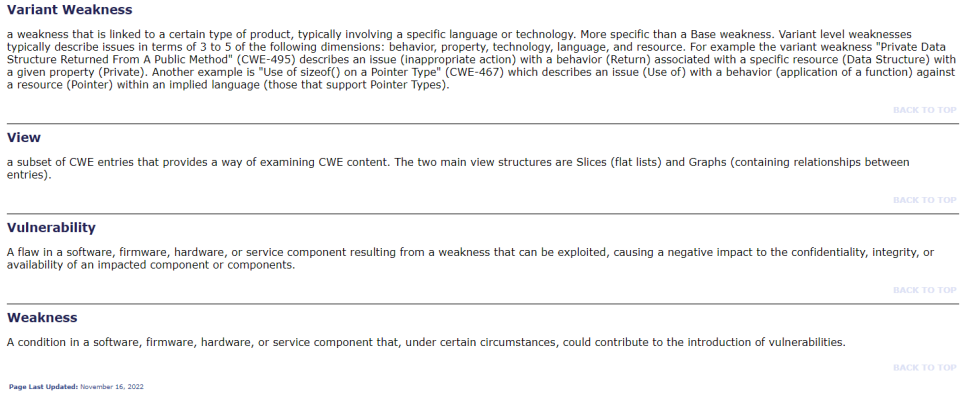
a type of mistake that, in proper conditions, could contribute to the introduction of vulnerabilities within that product. This term applies to mistakes regardless of whether they occur in implementation, design, or other phases of a product lifecycle.

BACK TO TOP

Page Last Updated: October 28, 2021

Figure D.1: The concept of *Vulnerability* in the CVE Glossary on October, 10$^{\text{th}}$ of 2021.

Figure D.2 shows the prior definition of the *Vulnerability*



Variant Weakness

a weakness that is linked to a certain type of product, typically involving a specific language or technology. More specific than a Base weakness. Variant level weaknesses typically describe issues in terms of 3 to 5 of the following dimensions: behavior, property, technology, language, and resource. For example the variant weakness "Private Data Structure Returned From A Public Method" (CWE-495) describes an issue (inappropriate action) with a behavior (Return) associated with a specific resource (Data Structure) with a given property (Private). Another example is "Use of sizeof() on a Pointer Type" (CWE-467) which describes an issue (Use of) with a behavior (application of a function) against a resource (Pointer) within an implied language (those that support Pointer Types).

BACK TO TOP

View

a subset of CWE entries that provides a way of examining CWE content. The two main view structures are Slices (flat lists) and Graphs (containing relationships between entries).

BACK TO TOP

Vulnerability

A flaw in a software, firmware, hardware, or service component resulting from a weakness that can be exploited, causing a negative impact to the confidentiality, integrity, or availability of an impacted component or components.

BACK TO TOP

Weakness

A condition in a software, firmware, hardware, or service component that, under certain circumstances, could contribute to the introduction of vulnerabilities.

BACK TO TOP

Page Last Updated: November 16, 2022

Figure D.2: The concept of *Vulnerability* in the CVE Glossary on November, 16$^{\text{th}}$ of 2022 – from Web Archive.

D.6 Summary of BRON Initiative Ontologies Characterization

Table D.1: Summary of BRON Ontologies Characterization [168].

	Applicability Level		**Generality Level**		**Bi-dimensional Classification [66]**	
Ontology	**Ref. Ontology**	**Oper. Ontology**	**Found. Ontology**	**Classif. [76, 247]**	**Formalization Level**	**Axiom. Level**
BRON	No	Yes	No	Application Ontology	ArangoDB	Lightweight
ATT&CK	No	Yes	No	Task Ontology	Formal Ontology (XMLS)	Lightweight
D3FEND	No	Yes	No	Domain Ontology	Formal Ontology (XMLS)	Lightweight
CWE	No	Yes	No	Domain Ontology	Informal Ontology/Tesauri and Taxonomies	Lightweight
CVE	No	Yes	No	Domain Ontology	Informal Ontology/Tesauri and Taxonomies	Lightweight
CAPEC	No	Yes	No	Domain Ontology	Informal Ontology/Tesauri and Taxonomies	Lightweight