THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : MSTII - Mathématiques, Sciences et technologies de l'information, Informatique
Spécialité : Informatique
Unité de recherche : Laboratoire d'Informatique de Grenoble

FORT: une théorie modulaire des relations ontologiques fondamentales pour représenter et raisonner sur la composition d'entités tangibles - observations du patrimoine culturel.

FORT: a modular Foundational Ontological Relations Theory for representing and reasoning over the composition of tangible entities - observations from cultural heritage.

Présentée par :

Fatme DANASH

Direction de thèse :

Danielle ZIEBELIN — Directrice de thèse
Professeure des universités, Université Grenoble Alpes

Emilie CHALMIN — Co-encadrante de thèse
Maitresse de conférence, Université de Savoie Mont-Blanc

Rapporteurs :

Nathalie AUSSENAC-GILLES
DIRECTRICE DE RECHERCHE, CNRS

Cesar GONZALEZ-PEREZ
SCIENTIST, Spanish National Research Council (CSIC)

Thèse soutenue publiquement le **28 septembre 2023**, devant le jury composé de :

Danielle ZIEBELIN — Directrice de thèse
PROFESSEUR DES UNIVERSITES, Université Grenoble Alpes

Béatrice MARKHOFF — Examinatrice
MAITRE DE CONFERENCES, Université de Tours

Jérome EUZENAT — Examinateur
DIRECTEUR DE RECHERCHE, INRIA Grenoble Rhône-Alpes

Nathalie AUSSENAC-GILLES — Rapporteure
DIRECTRICE DE RECHERCHE, CNRS

Cesar GONZALEZ-PEREZ — Rapporteur
SCIENTIST, Spanish National Research Council (CSIC)

Laurence CIAVALDINI-RIVIERE — Examinatrice
PROFESSEURE DES UNIVERSITES, Université Grenoble Alpes

Emilio SANFILIPPO — Examinateur
SCIENTIST, CNR Italian National Research Council,

Invités :

Emilie Chalmin Aljanabi
MAITRESSE DE CONFERENCES, Université Savoie Mont Blanc, Edytem

Dedication

To Kassem Danash & Layla Makki,
For making my dreams so vast
and the world within reach.

To Fadel, MohammadAli, and Zainab Danash,
For rendering distances so near,
even when continents apart.

All I am, I owe to you.

Acknowledgements

This thesis has marked a profound journey of intellectual and personal growth, during which several individuals have had great roles and contributed to his growth in multiple facets.

First and foremost, I express my sincere gratitude to Mrs. Nathalie Aussenac-Gilles and Mr. Cesar Gonzalez-Perez for accepting to review my work. Thank you for the time, expertise, effort and constructive feedback devoted into evaluating this work.

I would also like to express my deepest gratitude to Mr. Jérôme Euzenat, Mrs. Béatrice Markhoff, Mr. Emlio Sanfilippo, and Mrs. Laurence Ciavaldini-Riviere, the examiners of the thesis. I am particularly grateful to Mr. Jérôme, who had a pivotal role in my research journey ever since I embarked during my Master 2 internship. With your guidance, I learned the essence of research, how to conduct it diligently, and developed a genuine passion for this direction, Thank you.
Furthermore, I would like to give special recognition to Mr. Emilio (and Mr. Claudio Masolo from the Laboratory of Applied Ontology), with whom I had enlightening discussions during my second year. Your valuable insights helped me refine the problematic of my thesis and have positively influenced my approach.
Lastly, I extend my special thanks to Mrs. Béatrice for her role and dedicated time being the expert in my CSIs (Committé Suivi de Thèse) over the past years.
To all the reviewers and experts mentioned, thank you for contributing to the completion of my thesis.

Additionally, my heartfelt appreciation goes to my supervisors, Mrs. Danielle Ziébelin and Mrs. Emilie Chalmin-Aljanabi, for their support and guidance throughout this journey.

Danielle, I am truly grateful for your tremendous support, not only in research, but in all aspects. Over the course of these four years, you have consistently believed in my work, even during moments of self-doubt. Your attention to detail, reading and corrections, disciplined approach, and scientific dedication have been invaluable. I am especially thankful for your efforts in handling the parts of this thesis that I deviated away from, while you ventured into understanding between disciplines.
Emilie, I want to extend my gratitude for the time and effort you dedicated to introducing me to the field of cultural heritage at the beginning of the thesis, and for taking steps forward to understand modelisation-related concepts in computer science.

It has been a true pleasure working with both of you.

If I expand my thanks now, this definitely goes to all members of the Steamer team who have welcomed me since the beginning, starting by team's leader Mrs. Marlène Villanova-Oliver. This includes all the researchers in the team : Mrs. Camille Bernard (thank you for the multiple discussions we had at the beginning of this thesis), Mrs. Paul-Annick Davoine (thank you for all-time smile !), Mr. Philippe Genoud (thank you Philippe for the time you dedicated in meetings for the Patrimalp project, for your ski hints, and for your heartfelt support !), Mr. Sylvain Bouveret (thank you for having reviewed one of my writings during this thesis !), Mr. Jérôme Gensel (thank you for your feedback on my baking and cooking recipes, and for the encouragement to go further with them !), Mrs. Carole Adam, and Mrs. Julie Dugdale.

Thank you to Mr. German Vega also, for the meetings and discussions that consistently brought out the best in me.

As I reflect on this journey, I cannot overlook the international bonds that became an integral part of my thesis experience. Through the corridors of the laboratory and our coffee breaks, we were brought together crossing our cultural and linguistic differences. To Rouba Iskandar and Pier Castillo-Mora, you were family ! Daniela Milo-Flores, your warm hugs were always there. Rami Younes, your sense of humor sparkled everywhere. To Manon Predhumeau, Philip Scales, Thibaut Lopez, Samandar Ibragimov, Laura Alonzo, and Sina Ali Samir, thank you all.

This also expands to past Steamer members including Matthieu Viry (thank you for all the moments we shared in the office chatting and sharing our concerns !), Mathew Screeves, Maeva Seffar, Rabih Fares, Armen Inants, Ophélie Jobert, Anis Kadri, Badis Bouakaz, Etienne, Karam Slika, and Alice Jacquire, thank you !

I would also like to give a moment in the memory of Line Van-Den-Berg, a friend, accomplished researcher, and mountain enthusiast, whose recent passing has left me deeply saddened.

My acknowledgments continue to recognize certain special people.

Soha El Sayed, the person with whom I have shared countless memories throughout this entire journey. You were more than a friend ; you were family. You were my girl-next-door, my partner in crime during moments of sheer madness, and my go-to person in times of need. The moments we shared, traveling across the globe and exploring cuisines, cultures, and people will forever hold a special place in my heart. I am grateful for having had you by my side.

Khadija Al Khatib and Hanin Naser, the two who hold a special place in my heart as my childhood pals. Despite the distances that have put us across countries, you have always been by my side. Our friendship has been a constant source of support and joy, and especially through the hard times of my thesis.

Finally, I arrive here overwhelmed, with emotions rendering me speechless. My words fail to capture my deep gratitude to you. To my father (Kassem), my mother (Layla), my brothers (Fadel and Mohammad-Ali), my sister (Zainab), and my love (Hassan), I can never thank you enough ! It is with each of you that I embarked on this thesis journey. Alongside you, I have traversed its ups and downs. And to you, I dedicate this work.

My most profound gratitude is to God. Thank you for making me believe in me. And thank you for being present, throughout this journey, during the days and nights that I though would never pass, but passed !

Fatima Danash
Grenoble, France
July, 2023

Abstract

The materiality of tangible cultural heritage entities has garnered interest in heritage sciences due to its role in maintaining the unique identity of the entity across various interpretations as a boundary object. In multidisciplinary studies on tangible cultural heritage, an interdisciplinary approach is essential for developing a comprehensive understanding of human culture. To achieve an interdisciplinary approach within a multidisciplinary field such as cultural heritage, centered around a cross-disciplinary entity like the boundary object, a shared goal is necessary. By considering the *materiality* of the entity as the common point across multiple disciplines, and aiming for *a better understanding and representation of this materiality* as the shared goal, an interdisciplinary approach within cultural heritage can be fostered.

To effectively represent the materiality and composition of tangible cultural heritage entities, the use of an ontological model of *structural and spatial relations* is indispensable. Additionally, for a successful interdisciplinary integration, a meta-ontology approach is vital to overcome the challenges posed by the heterogeneity of multiple disciplines and promote interoperability across models of various domains.

In this thesis, we address the objective of representing and modeling the composition of any tangible entity using structural and spatial ontological relations, drawing insights from cultural heritage. For this purpose, we propose "**FORT** : a **F**oundational **O**ntological **R**elations **T**heory" within an applied ontological approach. FORT is designed with the following characteristics : (a) modular, i.e. composed of interlinked and intralinked relation modules ; (b) a meta-ontology i.e. specifying a meta-conceptualization of top-level abstractions and using a meta-modeling language of generic modeling primitives ; and (c) exclusively addressing relations and rule constraints.

To formalize FORT and illustrate its employment, we construct and adhere to an ontology engineering methodology. This methodology addresses various specification choices for FORT, namely expressivity and decidability, resulting in two versions : the *FORT reference ontology* and the *FORT lightweight ontology*. Furthermore, the methodology formalizes each specification, the reference and lightweight ontologies, at multiple levels, namely theoretical and empirical. Thus, FORT is formally expressed in a First-Order Logic (FOL) formalization with a Common Logic Interchange Format (CLIF) serialization for the reference ontology, and a decidable Description Logic (DL) formalization using the SROIQ fragment with an OWL2 implementation for the lightweight ontology. Moreover, the methodology bridges the two specifications through a systematic translation from the reference FOL theory to the lightweight SROIQ fragment.

Therefore, our approach contributes in the following ways. Firstly, we propose an expressive and well-founded language of exclusive relations and rule constraints through the FOL formalization of FORT. Secondly, we demonstrate the novelty and consistency of our proposed relations language through the CLIF serialization of FORT. Thirdly, we establish a decidable lightweight formalization of our relations language through a generic and systematic translation process, for the SROIQ formalization of FORT. Lastly, we provide this language as an OWL ontology and present different methods (direct and indirect) for its employment to support its practical use.

Résumé

La matérialité des entités du patrimoine culturel tangible a suscité l'intérêt des sciences du patrimoine en raison de son rôle dans le maintien de l'identité unique de l'entité à travers diverses interprétations en tant qu'objet frontière. Dans les études multidisciplinaires sur le patrimoine culturel matériel, une approche interdisciplinaire est essentielle pour développer une compréhension globale de la culture. Pour parvenir à une approche interdisciplinaire dans un domaine multidisciplinaire tel que le patrimoine culturel, centrée sur une entité cross-disciplinaire telle que l'objet frontière, un objectif commun est nécessaire. En considérant la *matérialité* de l'entité comme le point commun à plusieurs disciplines, et en visant *une meilleure compréhension et représentation de cette matérialité* comme objectif commun, une approche interdisciplinaire au sein du patrimoine culturel peut être encouragée.

Pour représenter efficacement la matérialité et la composition des entités tangibles du patrimoine culturel, l'utilisation d'un modèle ontologique des *relations structurelles et spatiales* est indispensable. Pour une intégration interdisciplinaire réussie, une approche *méta-ontologique* est essentielle afin de surmonter les défis posés par l'hétérogénéité de multiples disciplines et de promouvoir l'interopérabilité entre les modèles des différents domaines.

Dans cette thèse, nous abordons l'objectif de représenter et de modéliser la composition de toute entité tangible à l'aide de relations ontologiques structurelles et spatiales, en nous inspirant d'exemples issus du patrimoine culturel. Dans le cadre d'une approche ontologique appliquée, nous proposons "**FORT** : a **F**oundational **O**ntological **R**elations **T**heory". FORT est conçue avec les caractéristiques suivantes : (a) modulaire, i.e. composée de modules de relations interliées et intraliées ; (b) approche méta-ontologique, i.e. spécifiant une méta-conceptualisation d'abstractions de haut niveau et utilisant un langage de méta-modélisation de primitives de modélisation génériques ; et (c) exclusive, i.e. seules les relations et leurs contraintes de règles sont considérées.

Pour formaliser FORT et illustrer son utilisation, nous construisons et adhérons à une méthodologie d'ingénierie ontologique. Cette méthodologie aborde différents choix de spécification pour FORT, notamment l'expressivité et la décidabilité, et aboutit à deux versions : *FORT l'ontologie de référence* et *FORT l'ontologie légère*. De plus, la méthodologie formalise chaque spécification, l'ontologie de référence et l'ontologie légère, à plusieurs niveaux : théorique et empirique. Ainsi, FORT est formalisé en Logique du Premier Ordre (First-Order Logic - FOL) avec une sérialisation CLIF (Common Logic Interchange Format), pour l'ontologie de référence, et est formalisé en utilisant SROIQ le fragment décidable le plus expressif des Logiques de Description (Description Logics - DL) avec une implémentation OWL2, pour l'ontologie légère. En outre, la méthodologie fait le lien entre les deux spécifications grâce à une traduction systématique de la théorie FOL de référence vers le fragment SROIQ léger.

Les contributions de notre approche sont les suivantes : premièrement, nous proposons un langage expressif et bien fondé des relations et des contraintes de règles de FORT à travers une formalisation FOL ; deuxièmement, nous démontrons la nouveauté de FORT, ainsi que sa consistance de à travers une sérialisation CLIF ; troisièmement, nous établissons une formalisation légère et décidable de FORT par le biais d'un processus de traduction générique et systématique vers le formalisme SROIQ ; finalement, nous fournissons une implémentation de FORT en OWL et présentons différentes méthodes (directes et indirectes) pour l'utilisation pratique de cette ontologie.

Table des matières

1 Introduction

In this Chapter, we provide a general introduction of the thesis, followed by a Preliminaries section offering the essential background information.

1.1 Context

Heritage science, also called Patrimonial science referring to the original French term "Science du Patrimoine", is a multidisciplinary scientific field concerning cultural and natural heritage. It aims to improve the understanding, care, sustainable use and better management of heritage entities, so that it can enrich people's lives, now and in the future.

Cultural heritage (CH) is the legacy of the tangible artifacts (aka material entities) and the intangible attributes (aka immaterial entities) inherited from past generations [Logan2007]. It includes a wide range of entities such as artifacts, architecture, practices, customs, beliefs, knowledge, and expressions that reflect the cultural, historical, and social values of a particular group or region. These entities have had and continue to have an important place in our society, to which we attribute values (symbolic, commercial, cultural, social, scientific...) and significance [Sullivan2015]. CH involves the study, preservation, and interpretation of various aspects of human culture as the tangible and intangible aspects of a society or community's identity, history, and traditions that are passed down from generation to generation.

The difference between tangible (material) and intangible (immaterial) CH entities lies in their nature and characteristics [Munjeri2004]. Material CH entities refer to physical objects or artifacts that are considered important from a cultural or historical perspective, such as artworks, archaeological artifacts, historic buildings, manuscripts, traditional crafts, and ethnographic objects. They are tangible and can be preserved, protected, and displayed in museums, archives, or other cultural institutions. Material CH entities are often objects that are created, used, and appreciated by a culture or community, and they provide tangible evidence of a culture's history, identity, and artistic achievements. Whereas, immaterial CH entities refer to intangible aspects of a culture that are passed down from generation to generation including language, music, dance, oral traditions, rituals, festivals, and traditional knowledge. These are typically practices, expressions, knowledge, and traditions that are transmitted orally, through performances, or through other intangible means. Such intangible elements are often deeply embedded in a culture's social practices, beliefs, and values, and they contribute to the intangible cultural heritage that shapes a community's identity and way of life.

Both material and immaterial CH entities are important and interrelated, and together they contribute to the richness and diversity of a culture's heritage [Bouchenaki2003]. Within the scope of this thesis, we consider (1) material entities to which we refer as tangible entities, and (2) their significance and value to which we refer as the intangible aspects of tangible entities, throughout this document.

Moreover, a cultural heritage entity is considered as a "boundary object", a term adopted by heritage historians referring to the french concept "Objet Frontière" which initiated in the late 1980s by Star and Griesemer. "L'objet frontière est un objet suffisamment flexible pour s'adapter aux besoins et aux nécessités spécifiques des différents acteurs qui les utilisent et qui sont suffisamment robustes pour maintenir une identité commune" [1] [Star1989].

The term boundary object is meant to describe an object or artifact that is able to bridge different social worlds or domains of knowledge (the actors using the object), without necessarily being defined or interpreted in the same way by all of them (the robustness of the object). Howe-

1. It's corresponding translation : "A boundary object is one that is flexible enough to adapt to the specific needs and requirements of the different actors who use it, yet robust enough to maintain a common identity".

ver, despite the flexibility of the boundary object, it actually maintains a stable core, as a shared identity or common set of characteristics, that remains constant and recognizable to all parties involved (the common identity of the object). This shared identity allows the different actors to communicate and collaborate effectively, even if they have different backgrounds, perspectives, or goals [Akkerman2011]. In other words, the common identity which represents the common ground that unites the actors, acts as an interface between the multiple disciplines and allows for the exchange of knowledge, information, and ideas across disciplinary boundaries. This facilitates the collaboration and communication among actors from different disciplines with different perspectives, terminologies, or backgrounds [Bowker2000, Caccamo2022].

CH entities act as boundary objects [Leigh Star2010] because they are often multi-dimensional and can be interpreted in different ways, at different times, by different communities, while retaining a recognizable identity and significance. For example, a traditional dance or song may be performed and appreciated differently by different communities of different generations, but still recognized as a part of their shared cultural heritage. Similarly, a historical monument or artwork may hold different meanings and associations for different groups of people, but still serve as a point of connection and shared reference. This topic is examined in [Kopytoff1986] (not specifically about cultural heritage) where objects are argued to have a "cultural biography" that reflects their changing meanings and values over time as they are produced, exchanged, used, and discarded within different cultural contexts. Applying this approach to the study of cultural heritage tangible entities considers how their material qualities and historical contexts shape their significance and meaning within different cultural communities.

In this sense, CH entities are not just static artifacts or objects, but living and dynamic entities that are shaped and reinterpreted by the communities and contexts in which they exist [of Europe2009]. As boundary objects, CH entities provide a shared point of reference that facilitates the communication and collaboration across different disciplinary boundaries within a CH community, enabling the exchange of ideas, values, and practices.

In its role as a boundary object, a cultural heritage entity is not only considered of ***multi-disciplinary***, but also of ***cross-disciplinary*** nature. In order to study, interpret, and preserve cultural heritage tangible entities within diverse aspects of human culture (as mentioned earlier), it is mandatory to integrate the knowledge, methods, results, and perspectives from multiple disciplines [Harrison2013]. On the one hand, it is only with the expertise from multiple domains that insights can be drawn about the CH entity (multidisciplinary aspect) [Graham2012]. For example, the study of a cultural heritage entity such as a painting, may require expertise from art history, conservation science, and materials science, among others, to understand its historical context, artistic techniques, and material composition. On the other hand, in the context of boundary objects, a CH entity also serves as a common grounding that bridges disciplinary boundaries (the cross-disciplinary aspect) [Lowenthal1996]. For example, an archaeological site is a boundary object that brings together experts from various fields such as archaeology, history, anthropology, art history, conservation science, environmental science and geomorphology, to study and interpret the site's cultural significance and develop strategies for its preservation and management.

And according to both, the robust nature and the unique identity of a CH entity as a boundary object, it serves as an ideal tool for an interdisciplinary integration of knowledge and methods. Thus, cultural heritage is also an ***interdisciplinary*** field requiring in the collaboration and communication among experts from different fields to develop a comprehensive understanding of cultural heritage tangible entities [Smith2006]. An interdisciplinary approach to cultural heritage allows for a holistic comprehension and appreciation of human culture and its diversity. For that, a common goal or objective is necessary to ensure that each discipline's contribution is aligned and directed towards a shared purpose, to result in integrated and effective efforts to achieving an interdisciplinary approach. In [Lach2014], the authors argue that a shared goal is essential for interdisciplinary collaboration, as it helps to align the efforts of various disciplines towards a common purpose, as in the multi and cross-disciplinary field of environmental sustainability, involving fields such as ecology, economics, and policy. Without a common goal, each discipline may focus on its own objectives, resulting in conflicting strategies that do not effectively address environmental challenges. By establishing a shared goal, research teams can work together to develop holistic solutions that address their interdisciplinary challenges.

In heritage sciences, several interests arise for developing strategies for the preservation, restoration, management, and interpretation of CH entities. Among which is the goal of studying and constructing the tangible discourse of a CH entity and transmitting it over generations. A tangible discourse of a CH entity refers to the material characteristics of the entity that can be observed, analyzed, and interpreted. Material characteristics encompass the physical elements of the entity such as its form, material composition, construction techniques, decorative elements, and other physical attributes. For example, if the cultural heritage entity is a piece of artwork, its tangible discourse may involve its medium, technique, style, iconography, and condition. If it is an archaeological artifact, the tangible discourse may encompass its material composition, manufacturing techniques, and physical features.

In [Pearce2017], the author explores the ways in which museums and their collections shape our understanding of cultural heritage. In fact, objects and collections in museums provide tangible evidence of cultural heritage, and their materiality contributes to the discourse and interpretation of cultural heritage. The author suggests that the study of objects and collections can provide insights into the social, cultural, and historical contexts in which they were produced and used, and how they are perceived and valued by different communities. Similarly, [Dudley2012] explores the ways in which museum objects can be used as a means of accessing and interpreting cultural heritage, and highlights the importance of a tangible discourse for understanding the physical properties and material culture of these objects. Additionally, in [Miller2005], the authors discuss the importance of materiality in the study of culture, and how objects and artifacts provide tangible evidence of cultural practices and beliefs. It is argued that material culture can be understood as a "tangible discourse" that reflects the intangible aspects (e.g. the values, beliefs, and practices of a culture) by studying the tangible ones.
Indeed, several other works [Henare2007, Miller1997, Gell1998] highlight the significance of materiality in the communication and interpretation of cultural meanings and values, in a variety of disciplines, including anthropology, art history, and material culture studies.

Studying, analyzing and representing CH entities is necessary to construct their tangible discourse for which there is no written record. It provides important clues and evidence that can help researchers, conservators, and other stakeholders understand the history, cultural significance, and authenticity of the entity. Indeed, the politics and cultural implications of the UNESCO World Heritage Convention program examined in [Meskell2018] emphasizes the importance of a tangible discourse for understanding the material aspects of cultural heritage. By examining the material aspects of a CH entity, experts can gain insights into its creation, use, function, and meaning within its cultural context, and contribute to a deeper understanding of the ways in which physical objects and spaces shape learning and knowledge production [Sørensen2009].

The Patrimalp project, which this thesis is part of, is an interdisciplinary research project aiming to **develop an integrated and interdisciplinary approach for a better understanding of**

material (tangible) cultural heritage based on its cross-disciplinarity. The project is funded by the French National Research Agency [1], in the framework of the "Investissements d'Avenir" program, and divided into work packages (WP) involving several partners :

— (WP1) working on raw material resources and contextualization involving the EDYTEM lab from *Université Savoie Mont Blanc* (USMB), and the PACTE lab from *Université Grenoble Alpes* (UGA),

— (WP2) working on materials, manufacturing processes, and alteration involving the Néel institute from the *Centre National de la Recherche Scientifique* (CNRS), and the *European Synchroton Radiation Facility* (ESRF),

— (WP3) working on history, life, and trajectories of the artifacts involving the ARC-Nucléart lab from the *French Alternative Energies and Atomic Energy Commission* (CEA) Grenoble and LUCHIE lab from UGA, and

— (WP4) working on the modelisation and interconnected visualization of cultural heritage science involving the LIG lab (STeamer team, in which this thesis was conducted), and the LJK lab from UGA.

Patrimalp is part of the policy promoted by the European Council and the European Parliament to enhance the value of cultural heritage by establishing for the first time in 2018, a European Year of Cultural Heritage which draws attention to the importance of digital switchover, among others, on heritage elements.

The scientific approach of Patrimalp consists in a cross-disciplinary study of tangible cultural heritage entities dated from the Neolithic period to the pre-industrial period. These entities are preserved in the West Alps and the Rhone Corridor, and include parietal paintings in neolithic cave sites [2] and applied brocades on polychrome sculptures. Across the first three work packages, these entities incorporate an intrinsic analysis, allowing for the reconstruction of all the stages that enabled their achievement (materials, processes, models, etc.), and an extrinsic analysis allowing for the placement of these entities in their historical context of creation (natural, cultural, symbolic, etc.). And within the context of the fourth work package, only an integrated and interdisciplinary approach allows to reach such a level of results and understanding.

In this thesis, we particularly consider two complex examples of tangible cultural heritage entities as two case-studies ; rock art sites and their parietal paintings, and polychrome sculptures and their applied brocades. An example of each entity type is shown in figures 1.1 and 1.2 showing respectively the parietal/schematic paintings found in sites of the western Alps and southern of France such as those in "Rocher du Château" in the valley of Bessans, and the applied brocades found on the "Vierge de Pitié" and "Saint évêque Claude" sculptures.

Rock art sites are by definition complex CH tangible entities. The parietal paintings found are the expression of an ideal (intangible) universe, of collective representations, specific to prehistoric societies [David2002, Defrasne2023]. To study them, researchers in patrimonial sciences study the material remains of this ideal universe, the tangible manifestations of the ideas that animated prehistoric human groups. Researchers motivate that by stating that "If we will never be able to find their meaning, we can study their structure and nature through the different components that

1. ANR-15-IDEX-02 : CDP Patrimalp - Development of an Integrated and Interdisciplinary Heritage Science

2. Eighteen sites have been already identified and located within the context of Patrimalp in France and Italy. Some examples : Le Trou de la Fclaz, Saint-Jean-d'Arvey, Savoie ; Gias des Peintures, tende, Alpes Maritime ; West Abri n°2 de Pierre Rousse, Beauregard-Baret, Drôme ; Grotte du Loup, St-Laurent-sous-coiron, Ardèche ; Rocca di Cavour, Cavour ; Balma dei Cervi, Valle Antigorio ; le Rocher du Château, Bessans, Savoie ; Faravel, Fressinières, Les Hautes Alpes.

compose them" [David2018,Defrasne2019]. These parietal paintings were produced (i) by a cultural group at a time that is to specified ; (ii) on a site chosen from the environment for characteristics that is be to found and whose long-term evolution must be studied ; (iii) in interaction with other social practices that are to be highlighted ; (iv) in a particular cognitive register depending on the purpose of the graphic act (schematic forms and figurative representations probably used for communication purposes) ; (v) with coloring materials whose nature, origin, preparation and evolution are to be analyzed to shed light on cultural choices and social practices and the uses of territories. Thus, understanding the "rock art site" entity requires understanding each of the preceding elements, and their links, within an integrated approach. Using physio-chemical tools, it is possible to precisely identify the compounds and components of matter [Chalmin2019]. From geology, it is possible to correlate the use of geological materials with a source of supply [Salomon2022, de Kergommeaux2021], and so on.

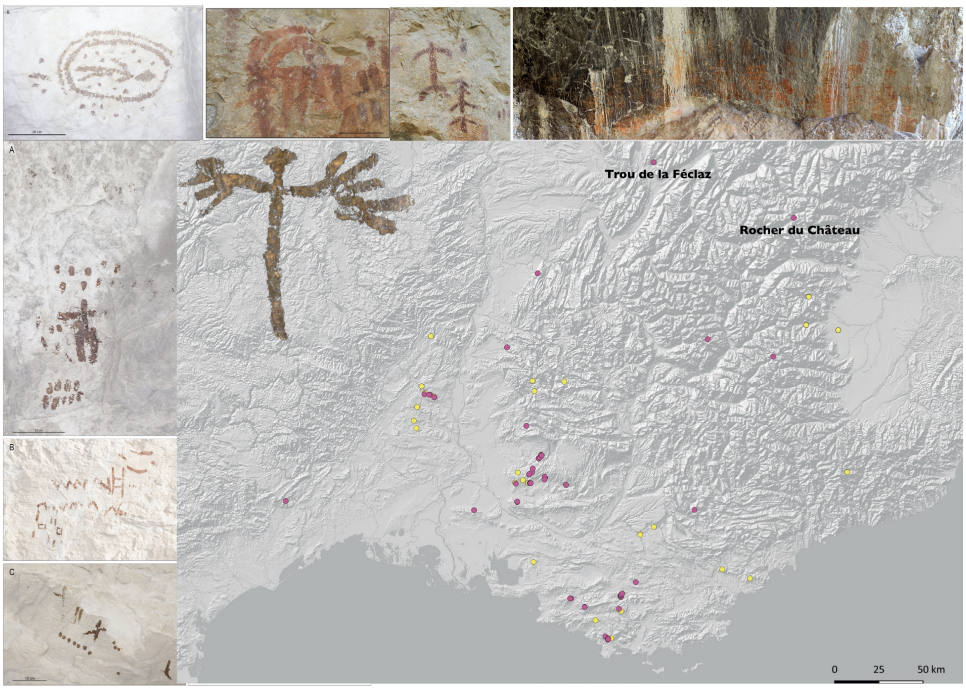


FIGURE 1.1 – *The distribution of sites with schematic paintings and engravings in southern France and the western Alps, with examples of their graphic representations. Location of study sites : Trou de la Féclaz (Massif des Bauges, St Jean d'Arvey) and Rocher du Château (Haute Maurienne, Bessans). In purple are mentioned the sites surveyed since 2014. Photo credit : Claudia Defrasne®.*

For polychrome sculptures, a corpus of Savoyard sculptures related to the western part of the duchy (Val d'Aoste in Italy and departments of Savoie and Haute-Savoie in France) present "applied brocades" patterns. "Applied brocade" is a decorative technique in relief based on molded and applied tin foil. The technique of pewter reliefs (cast relief) has existed since the 13th century [Brickhouse2016]. The decorations of "applied brocade" are intended to imitate, in volume, fashioned silk fabrics (velvet, lampas, damask), comprising motifs made of gold or silver threads in an edging weft (threads that run on the entire width of the woven piece) or stitched (which only run the width of the woven pattern) [Pinto2022]. These fabrics have been commonly described since the 19th century under the generic name of "brocades", and this term has been used since the middle of the 15th century in a non-specific way to designate silks woven with metallic threads. [1]

1. The first examples of "applied brocades" appeared in Northern Europe, around Hamburg, Berlin, Cologne, Brus-

The sculptures of the Savoyard corpus present two types of applied brocades defined by the shape of the cut motif and its location. The first type is called "juxtaposed", "covering" or "continuous", where the molded tin sheets, in a square or rectangular shape, are applied to the sculpture in such a way as to cover the entire surface of a garment. The pattern is repeated in a regular manner, horizontally or vertically. In the Savoyard corpus, this type is mainly present on clothes such as dresses, tunics or doublets but also on the bonnet of "Saint Crépin de Chambéry" [Pinto2021]. The evocation of this type of textile makes it possible to visually emphasize a character and its symbolic importance within a sculpted group by associating it with the most expensive precious fabrics [Van Duijn2012]. The second type of applied brocades is referred to as "dotted", "isolated" or "local", where the tin foil is cut so that the pattern takes on a round or polygonal shape and is placed fairly evenly on a painted surface of the sculpture. This type appears mainly on capes and more rarely on tunics [Pinto2021].
Applied brocades and the polychrome sculptures on which they are found are studied as complex cultural heritage tangible entities. Indeed, they represent the skills, techniques, and designs that have been passed down through generations, and reflect the history, values, and aesthetics of their respective cultures.
Researchers from patrimonial sciences in general, and in Patrimalp in particular, study these entities in an cross-disciplinary manner [Lelong2021]. From textile sciences, the physical and chemical properties of brocade fabrics, including their weave structures, fiber composition, and dyeing techniques can be studied e.g. [Dancause2018]. Researchers from art history study the historical contexts and cultural meanings associated with applied brocades, which involves examining historical texts, artwork, and other artifacts related to the production and use of brocades, as well as conducting ethnographic research with contemporary textile producers and users e.g. [Schoeser2007]. From the historical archives, it is possible to access receipts or trading acts, and so on. Also, material scientists use a variety of analytical techniques to investigate the physical and chemical properties of the materials used in brocades e.g. fiber, dyes, metals, and other decorative elements [Bordet2021], and so on.

For both preceding examples of CH tangible entities, which resemble our two case studies in this thesis, **the materiality of the entity** plays a vast role in its understanding and representation allowing for an interdisciplinary dialogue. It is certainly with the materiality of the entity that insights concerning its significance and value (immateriality) can be drawn for constructing eventually its tangible discourse.
And to establish an interdisciplinary dialogue around the shared goal that is **the understanding and representation of a tangible (material) cultural heritage entity**, defined as a boundary object, it is necessary to take up the challenge of understanding between disciplines.

Indeed, these disciplines need to understand each other, communicate their information, share their results, and interpret with one another their studies on the same patrimonial object. Yet, they use distinct terminologies, acquire diverse contextual backgrounds, and describe different viewpoints of the same entity, yielding in both syntactic (vocabulary-based) and semantic (content-based) heterogeneity. The former refers to differences in the way terms and concepts are used and represented across disciplines, whereas the latter concerns the differences in the meaning of the used terms and concepts [Klein2010, Klein2008].
Both create barriers to an interdisciplinary collaboration and limit the potential for knowledge

sels, the Netherlands, and France, on several sculptures between the end of the 14th and the beginning of the 15th century. Then the technique spread and grew in the rest of Europe and reached its peak between 1430 and 1530 in France, Switzerland, Austria, Bohemia, Lombardy (Italy), in the north of Spain, Portugal and in a much more fragmented way in Sweden, England and Wales. This complex decorative technique did not last more than a century and was quickly supplanted by simpler decorative techniques such as engraving, pastiglia, or sgraffito.

FIGURE 1.2 – *(a) Location of the juxtaposed "applied brocades" of the "Vierge de Pitié" sculpture group. In green : "applied brocade" juxtaposed with the tunic of Saint John. In yellow : "applied brocade" juxtaposed with the dress of Saint Mary Magdalene. In orange zoomed : details of the "applied brocade" of Saint Mary Magdalene.*
(b) Location of the "applied brocades" of "Saint évêque Claude" sculpture. In yellow : "applied brocades" juxtaposed with the bishop's dalmatic, in red "applied brocades" dotted with the bishop's cope. In orange zoomed : detail of an "applied brocade" dotted with the tread.
Photo credit : Florian Bouquet, Ariane Pinto, and ARC-Nucléart®.

integration causing (a) poor communication between researchers from different fields and between their organizations (people-level), (b) difficulties in identifying the common requirements which define the necessary specification for a possible common shared system (systems-level), and (c) lack of interoperability between the disparate modelings approaches of each organization, which yields in limiting the potential re-use and sharing of information across them (models-level) [Uschold1996].

To overcome syntactic and semantic heterogeneity, and achieve an interdisciplinary approach, we need to arrive to a shared, common and formal understanding that allows the bridging of disciplinary boundaries [Pohl2007]. A shared understanding ensures that the multiple disciplines agree on a consensus regarding the intersection of their interests, in view of their common shared goal. It is the foundation of a shared model that provides a common language (concepts, relationships, and rules) and a conceptual framework that allows the researchers from different disciplines to communicate and collaborate effectively [Klein1996]. Resolving the semantics of an interdisciplinary field via a shared model allows for semantic interoperability [Noy1997] and a semantic based data integration of the different knowledge sources via the common conceptual framework.

Ontologies address the need for a shared model in interdisciplinary fields by providing a formal and standardized way of representing and communicating knowledge, such as for Data Integration [Lenzerini2002] and the Semantic Web [Heflin2001]. Vice-versa, in [Guarino1995a, Guarino1997a], Guarino stresses the importance of an interdisciplinary approach in the practice of ontological engineering, underlying in particular the role played by formal ontology. **An ontology**

is a formal specification of a shared conceptualization i.e. it is defined in terms of a conceptual specification and a logical rendering [Guarino1995a]. [1] It includes a set of concepts, their definitions, and relationships between them. Such a presentation of an ontology term corresponds to its use in the *Applied Ontology* field. Together, with Knowledge representation and Reasoning, and the Semantic Web, they form the three fields that this thesis falls into. This shared conceptualization represented by a formal specification can be used to describe a common interest (goal), around a cross-disciplinary entity, across multiple disciplines, in an interdisciplinary approach.

It is possible to build an ontology as a meta-ontology that is not tied to any specific discipline i.e. not a domain-ontology. This is highly adopted in interdisciplinary approaches as a way of using the philosophical and linguistic aspects of the term ontology (within Artificial Intelligence) to analyze the structure of a given reality at a high level of generality and to formulate a clear and rigorous vocabulary [Guarino1998a]. A meta-ontology represents a shared meta-conceptualization using a shared meta-modeling language [Guizzardi2007]. This allows for the development of a generic and extensible framework for knowledge representation and facilitates knowledge integration across different domains and disciplines, which makes it a particularly useful for interdisciplinary approaches [Gruninger1995].

As for the employment of an ontology in interdisciplinary approaches, there are different ways of capturing implicit knowledge across heterogeneous data sources and creating semantic interoperability between them. As a result, a semantic- based data integration system is created using ontologies. Depending on the number of ontologies used, and their levels of abstraction, their are different approaches of *Ontology-Based Data Integration* (OBDI), based on [Wache2001].

Thus, adopting a (meta) ontology based approach for building a shared model in an interdisciplinary approach contributes to : (a) enhancing interoperability between the different systems through the provided common language (framework) [Berners-Lee2001], (b) resolving the semantic conflicts between the heterogeneous data resources and enabling the semantic integration and analysis of their data [Cruz2005], (c) advocating consistency and coherence between the multiple disciplines by providing a shared conceptualization upon which they agree and share an interest [Guarino1995a], (d) supporting knowledge management and reuse by providing a structured representation (hierarchy of concepts and relationships) of knowledge that can be shared and reused across different domains [Noy2001], and (e) facilitating communication and collaboration among researchers supporting more comprehensive and holistic analyses [Hastings2014].

As part of the Patrimalp project, to achieve an ***interdisciplinary*** approach, the ***multiple disciplines*** are brought together in order to access the materiality of these ***cross-disciplinary*** heritage entities. It is not a question of juxtaposing the elements brought by each discipline concerning the nature of the materials used. It is rather a problem of building a shared model, the meta-ontology, in which the common interest (the better understanding and representation of the materiality of tangible cultural heritage entities) shared by the different disciplines, is addressed.

For understanding it, a tangible CH entity that is a sensitive, authentic, concrete, material, visible, and touchable, must be characterized by a set of criteria, thanks to which each heritage object is unique. These include its spatial location at a given time, shape, dimensions, colors, famous author, material nature, production date, etc. Moreover, CH tangible entities that share a certain number of criteria (one or more) can come together to form a collection, which in turn has characteristics that distinguish it from another collection. For instance, in the case of sites with

1. This definition is based on Gruber's initial introduction of an ontology as an "explicit specification of a conceptualization" [Gruber1993]. Note that a detailed illustration of an ontology in terms of the notions of a conceptual specification and logical rendering is provided in the Preliminary Remarks section.

schematic paintings, these characteristics include : the type of location in the environment and that of the site (rock shelters or deep cavities, panoramic view, valley bottom, etc.), the mobilization of a particular cognitive register (schematic figures or figurative representations), similar time periods, associations of identical iconographic themes, dyestuff of the same nature or identical methods of preparation, etc.

In fact, the criteria mentioned not only apply to tangible CH entities, but they also characterize the structural and spatial aspects of any tangible entity. Therefore, the objective of this thesis expands beyond the creation of a model solely focused on the materiality of tangible CH entities. Instead, it aims to develop a *meta-ontological model that enables the understanding and representation of the materiality of any tangible entity by investigating its structural and spatial characteristics*, with a specific inspiration from CH as future goal for an application domain.

1.2 Problematic

Based on what preceded, we seek to build a meta ontological model for representing the materiality of tangible entities in general, and CH tangible entities in particular. Thus, in this thesis, we aim to answer the following research questions :

— How to model the materiality of tangible entities within a meta-ontological approach ? what are the requirements for this ontology to achieve a representation of materiality of the tangible entity concerning its structural and spatial settings ?

— How will this meta-ontology be employed in practice in general, and within an interdisciplinary approach in the presence of multiple disciplines in particular ?

— How is this meta-ontology approach convenient for applications in general, and the Patrimalp project's application in particular ?

The first inquiry pertains to a matter of **modelisation**, and our objective is to address it through the study of ontological relations within an Applied Ontology approach. We believe that it is useful and necessary to focus on ontological, structural and spatial, relations between entities and within entities to explicate the semantics of its materiality regarding its structural and spatial settings. To do so, only an ontological model of composition relations can adequately address this fundamental question.

The second inquiry revolves around the possible **employment approaches** of the proposed meta-ontology, with focusing on the matter in interdisciplinary approaches. Our proposed course of action to tackle this matter depicts two methods for employing the meta-ontology based on the setting of the application. In case of a single ontology application, the employment method is "direct". In case of multiple ontologies as is the case in interdisciplinary integration approaches, the employment is "indirect" accompanied by an ontology based data integration (OBDI) approach. For an OBDI paradigm, we believe the Global-as-View (GaV) approach adequately addresses interdisciplinary requirements.

Moreover, the third inquiry pertains to demonstrating the applicability and convenience of the proposed ontology. Within the scope of the Patrimalp project, the Cultural Heritage field is taken as the interdisciplinary application to show ontology in practice.

Therefore, we identify the fundamental objective of thesis as follows. Then, we clarify the approaches versus the fields of study that we address in view of our fundamental objective.

Fundamental objective :

Representing and modeling the composition of a tangible entity in general, and a Cultural Heritage tangible entity in particular, **using ontological structural and spatial relations**, within a Applied Ontology approach .

For the two matters of interest that we address in this thesis (modelisation and employment phases), we present some preliminary remarks in a Preliminary Remarks section focusing on *"ontologies"* and *"ontology-based data integration"* paradigms. This is to guide the reader with our followed approach throughout the manuscript in general, and Chapter 4 in particular upon which our contributions (in Chapters 5, 6, 7, 8) are based.

Regarding the state-of-the-art, two main fields necessitate consideration of the existing research conducted within each field. It is important to clarify that the state-of-the-art section serves as a general overview of the well-established literature that is relevant to our thesis, rather than a comprehensive review of the related topics addressed in this thesis. A more specific examination of related work is provided at each chapter level in each of the four main contribution chapters (5, 6, 7, 8).

The first field is motivated by modeling the composition of a tangible entity as a complex structure to enable the understanding and representation of its materiality, with the focus on CH entities [motivation-I]. This requires *investigating the extant ontological models of composition within the literature on cultural heritage model*, which is is carried out in Chapter 2.

The second field is motivated by acquiring a number of ontological, structural and spatial, relations that enable representing the composition of a tangible entity [motivation-II]. And this requires *investigating the extant literature on foundational ontological relations, focusing on structural and spatial relations, within the applied ontology field*, which is is carried out in Chapter 3.

After conducting a general review of the relevant literature pertaining to the two motivating factors, we have identified (in Chapter 4) three key challenges that need to be addressed in our thesis, and which we later contribute to.

- Challenge A : The need for a well-formalized language of a minimal set of ontological relations including rule constraints and excluding categories.
- Challenge B : The need for a meta-ontology that understands, represents, and models the structural and spatial constraints of a tangible entity.
- Challenge C : The need for a mapping and query pattern, according to specific employment method(s), to navigate and exploit the proposed ontology, and infer information relevant to the underlying questions concerning the materiality of the tangible entity.

Positioning of the thesis

The work of this thesis is positioned in the field of (a) Applied Ontology (AO), addressing in particular foundational ontological relations, using (b) Knowledge Representation and Reasoning (KRR) languages. The application of this work is carried out with a particular interest in the (c) Cultural Heritage (CH) field from which insights are drawn. For a CH interdisciplinary application, (d) an Ontology-Based Data Integration (OBDI) approach shall be used. Figure 1.3 depicts the four fields of study (AO, KRR, CH, and OBDI) and the corresponding positioning of our thesis's two phases, modelisation and employment, at the intersection between (AO and KRR), and (CH and OBDI) respectively.

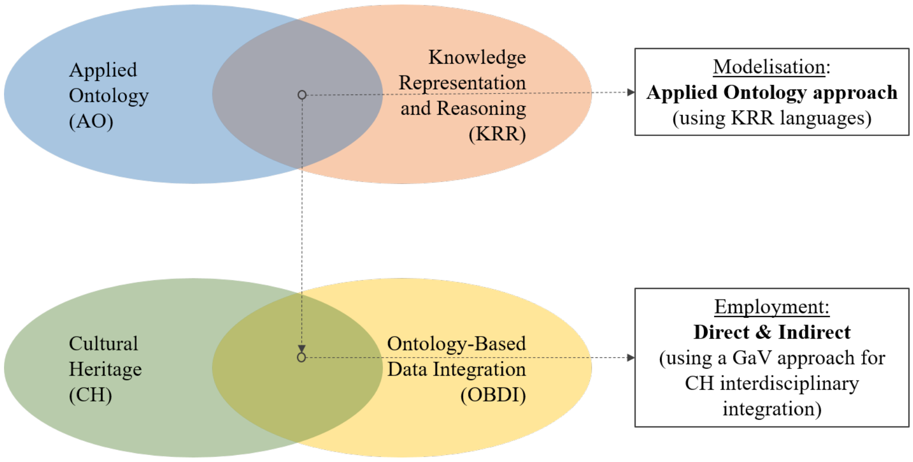


FIGURE 1.3 – *The four fields of study (AO, KRR, CH, and OBDI) and the corresponding positioning of our thesis's two phases, modelisation and employment, at the intersection between (AO and KRR), and (CH and OBDI).*

1.3 Contributions

In order to address the aforementioned challenges, we have established an applied ontological approach and an ontology engineering methodology, which serve as the fundamental pillars of this thesis.

From a modelisation standpoint, we propose *a meta-ontology at a metalevel defining a meta-conceptualization using a meta-modeling language*. This meta-ontology follows the formalization principles of meta-ontologies in the field of applied ontology [Guarino1998a, Guizzardi2007]. Our interpretation of this concept is as follows : a meta-conceptualization represents high-level abstractions of the world, encompassing the identification of concepts, relationships, objects, and more. On the other hand, a meta-modeling language employs generic vocabularies as modeling primitives to represent specific real-world examples.

For the employment method, we recommend *two possible eployment methods based on the application setting : direct and indirect*. In the case of a single ontology application, the employment method is referred to as "direct". Conversely, in situations involving multiple ontologies, as commonly encountered in interdisciplinary integration approaches, the employment method is deemed "indirect" and accompanied by GaV paradigm for ontology-based data integration (OBDI) to facilitate interdisciplinary integration based on the analysis in [Ekaputra2017].

To formalize the meta-ontology and design the proposed employment approach, we construct and adhere to a comprehensive ontology engineering methodology consisting of six distinct steps. Each step is designed to accomplish specific micro-objectives and ultimately overcome the challenges at hand. Our methodology encompasses four primary contributions, all of which revolve around a key deliverable presented by our approach, namely the **F**oundational **O**ntological **R**elations **T**heory (**FORT**).

As a first contribution, we propose the FORT reference ontology of relations at a theoretical level. This is by specifying and formalizing FORT using First-order logic (FOL) ; the expressive logic for constructing comprehensive theories. FORT introduces a minimal set of foundational ontological relations, namely parthood, dependence, location, membership, and constitution, based on the formalization of relations in existing literature. Each of these relations covers an intended representation for modeling the eventual composition of tangible entities. Thus, acting as a tool fostering semantically rich and well-formalized ontological relations.
FORT is designed with the following characteristics. It is a (1) modular ontology, composed of multiple ontology relation modules. These modules are internally linked using a set of definitions, axioms, and theorems, while being interconnected with each other through axioms. (2) It functions as a meta-ontology, both in terms of the conceptualization it defines and the modeling language it employs. The ontology specifies a meta-conceptualization that encompasses high-level abstractions, and employs a meta-modeling language consisting of generic vocabularies. (3) FORT exclusively focuses on relations and rule constraints, omitting the inclusion of entity types (ontological categories).
The objective is *to provide an expressive language that exclusively deals with relations and rule constraints*. This facilitates the understanding, representation, and reasoning of the materiality of any tangible entity as a complex structure.

In the second contribution, we undertake the analysis and validation of the FORT reference ontology at an empirical level. For the analysis, we interpret the construction of FORT in view of extant literature, more precisely : meta-ontologies that offer foundational ontological relations. To achieve this, we present three key arguments supporting the development of FORT. Additionally, we position FORT in relation to these ontologies and conduct a detailed comparison of relations, highlighting both the similarities and differences between them.
For the validation, this involves an additional serialization of FORT using Common Logic (CL) and specifically the CLIF format. By employing this serialization, we are able to conduct consistency checks, translate FORT into alternative serializations, and perform automated theorem proofs. Our objective is *to demonstrate the novelty and consistency of our proposed language for relations*.

For the third contribution, we establish the formalization of the FORT lightweight ontology at a theoretical level, employing the SROIQ Description Logic as a decidable language for knowledge representation and reasoning. To achieve this, we develop a generic procedure that facilitates the translation of First-order logic (FOL) theories into SROIQ-Tboxes. This translation process consist of multiple steps. It extracts a decidable fragment while preserving the highest level of expressivity. Subsequently, we apply this translation procedure to FORT, resulting in its representation within a decidable lite fragment, hence the term "lightweight ontology".
The objective is *to obtain a formalization of our proposed language of relations that is both decidable and capable of supporting rigorous reasoning*.

In our fourth contribution (proposal), we provide the FORT lightweight ontology at an empirical level, by implementing the translated SROIQ-Tbox into the OWL2DL web ontology language. This is followed by demonstrating the possible employment methods of the FORT lightweight ontology based on the application's settings and objective yielding in : direct and indirect methods. It is important to highlight that due to time limitations and limited inputs, we do not illustrate a complete application of FORT, the reason for which we refer to this contribution with "proposal". The objective is *to support the practice of FORT and demonstrate its applicability and convenience for real world applications*.

1.4 Thesis Outline

In the following Preliminary Remarks section, we present the foundations of ontology as a shared conceptualization and a modeling language, precisely within Applied Ontology. We also demonstrate the different ontology-based data integration approaches for interdisciplinary applications.

In Chapter 2, we conduct an investigation into the existing ontological models in the cultural heritage literature, with a specific emphasis on the representation of tangible entity compositions. We systematically classify these models and conduct a comprehensive analysis of three relevant models, examining their treatment of composition relations in terms of both structure and spatial aspects.

In Chapter 3, we clarify the notion of foundational ontological relations. Then, we categorize well-known structural and spatial relations from the literature based on different aspects of study. After that we illustrate some taxonomies and theories, those revolving around structural and spatial part-whole representations within and of entities, including mereological, mereotopological, and meronymic studies among other.

In Chapter 4, we interpret our Applied Ontological approach, as a meta-ontology of relations in terms of a meta-shared conceptualization and a meta-modeling language. Additionally, we establish an ontology engineering methodology for the formalization of the ontology and establishing its employment within a Global-as-View paradigm required for interdisciplinary applications.

In Chapter 5, we introduce our proposal of the FORT reference ontology, which focuses on selected foundational relations at both the microlevel and macrolevel. At the microlevel, we illustrate the micro-theories of FORT as multiple intralinked relation ontologies. At the macrolevel, FORT is presented as a modular macro-theory that interlinks the multiple relation micro-theories.

In Chapter 6, we analyze the FORT reference ontology, demonstrating its novelty in relation to existing literature on foundational relations. Additionally, we serialize FORT as a CL ontology in the CLIF format to perform operations and showcase its consistency.

In Chapter 7, we propose a procedure for translating First-order logic (FOL) theories into SROIQ-Tboxes, delineating a sequential series of steps. We analyze the properties of this procedure and apply it to the translation of FORT, resulting in a decidable fragment of FORT expressed in SROIQ, as the FORT lightweight ontology.

In Chapter 8, we implement the FORT lightweight ontology in OWL2DL, and demonstrate proposals for its employment oferring several methods based on the application's setting and objectives.

In Chapter 9, we recapitulate the motivations behind our thesis and outline our main contributions. We also discuss the limitations of our work, particularly the atemporal assumptions made in FORT, which provide potential avenues for future extensions of the approach. Furthermore, we provide perspectives for future research directions at various levels.

Preliminaries -
Ontologies : their modelisation and employment approaches

Tracing the roots of the term "ontology"

The term "ontology", having its roots in philosophy, was firstly studied by the Greek philosopher Aristotle, as the most fundamental branch of metaphysics referring to *the science of being qua being*. It deals with entities and relations found across categories, such as scientific disciplines, in addition to those recognized by common sense.

In contrary to scientific fields that aim at discovering and representing reality under a certain view, "Ontology" (with a capital 'O') is a discipline established by the German philosopher Husserl for developing scientific theories of reality, as a *science of essences*. It focuses on the nature and structure of entities independent of their actual existence and perspective. This field contributes to an analogy between *Formal Ontology* and *Formal Logic*. The former deals with formal ontological structures, as aspects of objects, independent of their specific kinds, types and instantiations, identity, unity, etc. While the latter deals with formal logical structures e.g. truth, validity, consistency, etc. The first ontology, namely "the set of theories of Substance and Accidents", was developed by Aristotle in his *Methaphysics* and *Categories* [Cohen2021].

An original clarification on the way the term "ontology" has been used in computer science is presented in [Guarino1995b].

On the one hand, in Information Systems (IS), according to Smith [Smith2012], the term 'ontology' firstly appeared in 1967 by S.H. Mealy, in his work "On the foundations of data modeling". Mealy argues that the existence of things in the world can be distinguished in the field of data processing into three distinct realms : (a) *"the real world itself"*, (b) *"ideas about it existing in the minds of men"*, and (c) *"symbols on the paper or some other storage domain"*. It basically discusses the existence of things in the world regardless of their (possible) multiple representations, and concludes by introducing the term "ontology" under the statement *"This is an issue of ontology, or the questions of what exists"* including some of Quine's claims in [Quine1948]. As such, an ontology as a particular system of categories accounting for a certain vision of the world. This view of "ontology" by the Information systems community conforms to its definition in philosophy as independent of a language i.e. an ontology is the same whether it is represented in a language as natural as English or as a formal as First-Order Logic.

On the other hand, in Artificial Intelligence (AI), the term "ontology" was introduced in [Hayes1979] to develop a physics ontology for liquids [Hayes1985], and later other ontologies started to appear in different domains as *domain ontologies*. As such, an ontology refers to an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words [Guarino1998a]. And a domain ontology is simply one that is focused on a particular domain of knowledge, such as medicine [Pisanelli2004, Gangemi1998], mechanical engineering [Borst1997], natural language translation [Knight1994, Mahesh1996], geographic information systems [Casati1998], and other fields. It is a formal model of the concepts, relationships, and rules that define the domain, and it provides a common vocabulary and conceptual framework that can be used to facilitate communication and collaboration among the stakeholders in the domain. The view of this community, including areas of computer science such as semantic web communities, corresponds thus to a concrete artifact designed for a specific purpose, and represented in a specific language.

In this manuscript, our reading of the term ontology refers to the AI reading, following the clarification in [Guarino1998a] and [Guizzardi2008]. Under this interpretation, an *ontology* uses the word *conceptualization* to refer to the philosophical representation, while being dependent on the language used i.e. two ontologies can be different in the vocabulary they use (English, Greek, First-order Logic, etc.) while sharing the same conceptualization.

Defining "ontology" and its variants

In terms of its definition, several efforts have surfaced in the literature e.g. [Swartout1999] or [Hendler2001]. Among which, as mentioned in [Guarino2009b], the one from Gruber in [Gruber1993] seems to be the most prevalent and most cited as an *"explicit specification of a conceptualization"*. Although the word ontology can be used to refer to an implicit representation i.e. existing only in someone's head or embodied in a piece of software. However, in the scope of this work and following its popular use AI, the notion of ontology and its respective conceptualization must be made accessible to users through an explicit and formal description.

In [Guarino1995b], the authors carry out a detailed discussion for clarifying the term "ontology" across its bearable interpretations in AI, driven by the original analysis of Gruber.
In 1997, Borst defined an ontology by additionally requiring that the conceptualization should express a shared view between several parties, a consensus rather than an individual view, as a *formal specification of a shared conceptualization* [Borst1999]. This means that such a conceptualization should not only be explicit, but also expressed in a (formal) machine readable format. As for the "shared" property, it refers to approximations of minimal consensual conceptualizations that can be made based on a set of examples. And it is thanks to the "shared" property that ontologies offer interoperability at large-scale applications.
So in 1998, Studer et al. [Studer1998] merged these two definitions stating that an ontology is *"a formal and explicit specification of a shared conceptualization"*. All these definitions are grounded on the same terms ; *"shared conceptualization"* and *"explicit/formal specification"*.

Later, in [Guarino2009b], a precise and concise formalization was provided, focusing on the three major aspects of the definition by Studer et al. "conceptualization", "formal, explicit specification", and the "shared" property. In addition, a formal characterization of the notions of "ontology", "conceptualization" and "metamodel", as well as on the relations between these notions, is presented in [Guizzardi2007]. In the latter, the authors present these notions at different levels ; domain-level and meta-level. We explain these different notions in view of figure 1.4, based on [Guarino2009b] and [Guizzardi2007].

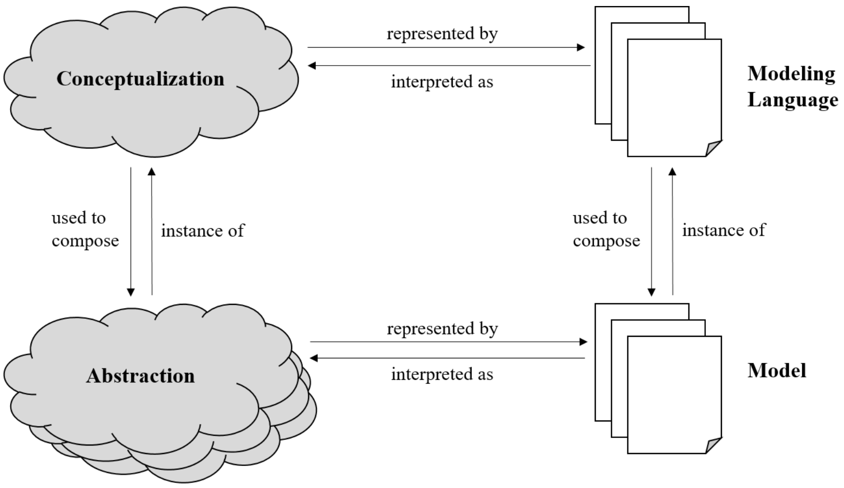


FIGURE 1.4 – *Figure taken from [Guizzardi2007] showing the relations between Conceptualization, Abstraction, Modeling Language and Model.*

A Conceptualization and a Modeling Language

Conceptualizations and abstractions are entities that exist only in the mind of a community using some language. A conceptualization refers to the abstract representations of knowledge, which can be used to reason, make decisions, or solve problems [Gruber1993, Gruber1995]. This is by identifying and defining "the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them". These elements constituting a conceptualization are used to articulate abstractions of certain state of affairs in reality i.e. a conceptualization is *used-to-compose* abstractions of reality. This refers to the *used-to-compose* relation in figure 1.4 between the *conceptualization* and an *abstraction*.

In order for the conceptualization to be documented, communicated and analyzed, a language is necessary to capture it. Thus, to concretize a conceptualization, concepts and their relationships are *represented-by* vocabularies (names) of a certain modeling language [Buccell2005]. This refers to the *represented-by* relation in figure 1.4 between the *conceptualization* and the *modeling language*. These vocabularies makeup the primitives of the modeling language that can directly express the conceptualization. In [Quine1969], Quine refers to the real world entities that the primitives of the modeling language commits to the existence of, as the ontological commitments of this language i.e. its commitment to the representation of reality that the domain conceptualization abstracts. In [Guarino1998a], this is referred to as the ontological commitment of the language L to a conceptualization C, where the authors formally specify it in terms of L and C.

Moreover, the primitives of a modeling language are used to form rules as valid combinations of these symbols, called models. These models that can be made with the primitives are based on some interpretations of the language. This refers to the *used-to-compose* relation in figure 1.4 between the *modeling language* and a *model*. The models also called (language models) resemble the context conditions that constraint the use of the language's syntax. Language models are enriched by context-conditions given in some constraint description language, such as the higher order logic named first-order logic (FOL) [Smullyan1995].

And finally, models are used to represent the abstractions of reality which are instances of the conceptualization that the modeling language's syntax represent. This refers to the *used-to-compose* relation in figure 1.4 between the *modeling language* and a *model*.

The preceding elements can be summarized following [Guizzardi2007] as : a model M can represent an abstraction A, only if the modeling primitives of the modeling language L used to produce M, represent the conceptualization C, which is used to articulate A.

An ontology : a conceptual specification within a logical rendering

Now what is an ontology in terms of what preceded ? An ontology is a model that can adequately specify the conceptualization using the modeling language. In fact, the language models which we mentioned earlier as the result of interpretations of the modeling language, resemble all the possible models that can be built using the modeling language. An ontology is a model, from the possible language's models, whose interpretation is approximated to target the *intended models* which are to represent the intended abstractions.
In [Guarino2009b], the interpretation under which the ontology model is built is called the *ontological commitment* of the model (ontology) towards the primitives of the modeling language. Figure 1.5, shows how an ontology is basically a model within the possible models of the language's models. And according to its specification, and the intended models, an ontology is approximated to be good or bad depending on its coverage of the intended models.

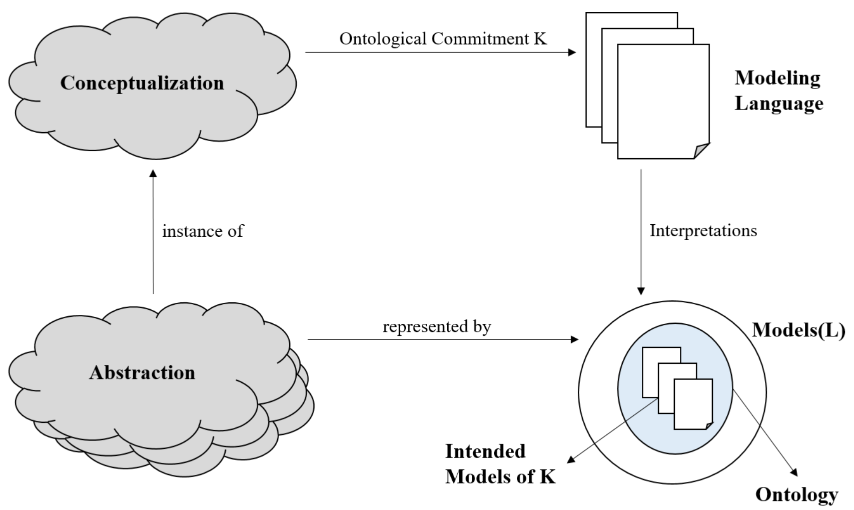


FIGURE 1.5 – *Figure adapted from [Guarino2009b] showing the relationships between the abstractions, their abstracted conceptualization, the language used to talk about such conceptualization, its intended models, and an ontology.*

Consider the conceptualization C, and the vocabulary V used to intentionally represent the universe of discourse of C using the modeling language L. An ontology O is defined in terms of a conceptual specification X, as a *concrete representation* of the universe of discourse which C abstracts in terms of V, and a logical rendering T_x, as the *logical theory* that explicates the specification X in terms of a formal language.

The conceptual specification X of the ontology describes knowledge about a domain in a manner that is independent of epistemic state of affairs [Guizzardi2007]. It intends to constraint the possible interpretations of a language's vocabulary so that its logical models approximate the set of intended world structures of a conceptualization C of that domain.

The logical rendering of the ontology is the theory that is the formal logical description, e.g. in first-order logic, of its specification X. The value of the logical theory in representing the specification depends thus on both : firstly the modeling language i.e. the choice of the modeling primitives, and secondly the formal logical language i.e. the formal logic used for rendering the vocabulary. The former ensures the suitability of the language and the appropriateness of the vocabulary regarding the intended shared conceptualization. The latter, however, constraints, based on its syntax and semantics, the description of the specification. For instance, adopting first-order logic as the logical language offers wide expressivity yet yields to a loss of some computation services. Whereas adopting a decidable description logic guarantees computational services, over a limited expressivity.

The meta-ontology level

The concept of levels of ontologies was first introduced by Guarino in [Guarino1998a] proposing a three-level ontology architecture consisting of an application ontology, a domain ontology, and a top-level ontology, according to their level of generality following a detailed discussion in [Guarino1997b, Guarino1997a]. Top-level ontologies describe very general concepts (e.g. space, time, matter, object, etc.) independent of a particular domain. Domain and task ontologies describe the vocabulary related to a generic domain (e.g. medicine, culture, etc.) or task (e.g. selling, paying, etc.), in which it specializes concepts of top-level ontologies. And application ones describe concepts depending on both, a particular domain and a particular task, as a specialization of both domain and task ontologies.

Later, other intermediate ontology levels were proposed such as core ontologies and mid-level ontologies. [1] Core ontologies are ones that define multiple domains and consist of the minimal concepts required to understand concepts across these domains [Falquet2011]. It is linked to a particular multidisciplinary domain and provides several view points relating to the different disciplines. An example of a core ontology is the CIDOC-CRM [ISO211272014] providing an ontology for concepts and information in cultural heritage and museum documentation. Whereas mid-level ontologies are ones that aim to bridge the gap between top-level generic terms and domain-level specific terms [Ceusters2015] such as the Common Core Ontologies (CCO) [2012019], and CIDOC CRM too.

A meta-level aims at providing a way to describe and analyze reality at a higher level of abstraction, allowing for more interoperability and use of models across different domains and applications. For a meta-ontology level, a conceptualization is a meta one that is encompasses formal ontological categories and relations of *top-level abstractions*. Similarly, its modeling language is a meta one whose syntax uses *generic primitives* i.e. domain-independent vocabularies.

Indeed in [Guizzardi2007], the authors distinguish between two levels : the level of material domains (such as genomics, archaeology, etc.) and the corresponding domain-specific modeling languages, and the meta-level of domain-independent meta-conceptualizations with a general modeling language (also called general ontology representation language or general ontology modeling language). This is depicted in figure 1.6 which shows the relations between the different notions at both levels. The meta-conceptualization defines a level of abstraction that is top-level, and can instantiate the set of all domain conceptualizations. And the general ontology representation language is a language whose primitives specify theories that can be used for the formal

1. A detailed illustration of the different levels is presented in [Cummings2017].

characterization of domain-specific languages and the production of domain ontologies.
For that, the primitives of the meta-modeling language representing the world space of the meta-conceptualization shall define the general laws that describe reality. This falls under the field of *formal ontology* in philosophy.

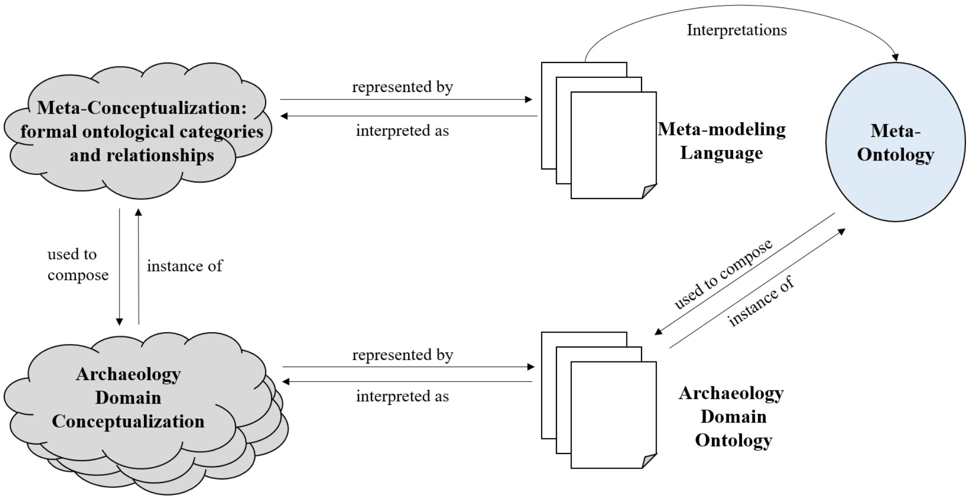


FIGURE 1.6 – *The relations between Domain Conceptualization, Domain Ontologies, Meta-modeling languages, their Meta-conceptualization, and a Meta-Ontology.*

Moreover, some meta-ontologies that are developed using the theories from *formal ontology*, are referred to as *foundational ontologies* (sometimes top-level ontologies or upper-ontologies) in the literature. These are : DOLCE (A Descriptive Ontology for Linguistic and Cognitive Engineering) [Masolo2003, Borgo2022a], BFO (the Basic Formal Ontology) [Smith2002], UFO (the Unified Foundational Ontology) [Guizzardi2015,Guizzardi2022], GFO (the General Formal Ontology) [Herre2010], GUM (the Generalized Upper Model) [Bateman2010, Bateman1995], SUMO (the Suggested Upper Merged Ontolog) [Niles2001], and YAMATO (the Yet Another More Advanced Top-level Ontology) [Mizoguchi2022] (as the next version of YATO [Mizoguchi2009]. Foundational ontologies are comprehensive meta-level (or top-level) ontological theories of categories (aka concepts) and relations.

Meta-level ontological theories provide a common framework not only for representing knowledge from different domains, but also for integrating their corresponding data based on the semantics of the shared common ontology. For example, in [Gangemi2002], the authors conclude how foundational ontologies (which are meta-ontologies) act not only as a reference for users providing a set of formal guidelines for domain modeling, but also as a tool for making heterogeneous ontologies interoperate or merge. Also, [Gómez-Pérez2006] discusses the challenges of integrating knowledge from different domains, and demonstrate how meta-ontologies can help to overcome these challenges by providing a shared framework for representing and exchanging knowledge.

Figure 1.7 show the different ontology levels which we have described. For our work, due to generalizing our approach to be applicable to an entity across any domain (domain-independent), we are interested in meta-ontologies. And following [Guizzardi2007], in this manuscript, we prefer the use of the notion ***"meta-ontology"*** to refer to an *"ontological theory at a meta-level i.e. specifying a meta-conceptualization using a meta-modeling language"*. We believe that the term "meta-ontology" or "top-level ontology" is more generic than "foundational ontology". This is because ontological theories can be comprehensive in terms of categories, relationships and their rules, as is the case of foundational ontologies, or not. The latter includes ontological theories

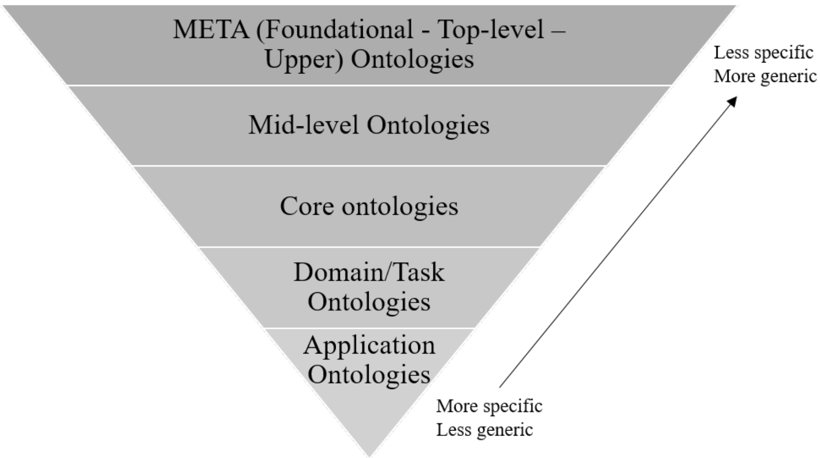


FIGURE 1.7 – *The different ontology levels starting from specific to generic levels as follows : Application, Domain/Task, Core, Mid-level, and Meta (Foundational - Top-level - Upper) ontologies.*

of sole relationships such as mereology [1] [Varzi2003]. Both cases are ontological theories at a meta-level (i.e. meta-ontologies), while presenting different approaches.

Applied ontology, and it's links to other fields

The study and development of ontological theories at a meta-level as meta-ontologies has some roots in philosophy, and is referred to as the *Applied Ontology* field in AI. Applied Ontology builds on philosophy, cognitive science, linguistics and logic with the purpose of understanding, clarifying, making explicit and communicating people's assumptions about the nature and structure of the world [Guarino2012]. This orientation towards helping people understanding each other distinguishes applied ontology from philosophical ontology, and motivates its unavoidable interdisciplinary nature. Whereas the study of the content (of these assumptions) as such independently of their representation is the *Ontological analysis* [Guarino2008], upon which, an interpretation (ontological commitment) can be built.

Of course, building meta-ontologies is not exclusively the business of Applied Ontology in AI. Other fields intersect within the context of ontologies, among which we are name *Knowledge Representation and Reasoning* (KRR) and the *Semantic Web* (SW).

For KRR, ontologies provide a structured and formal way to represent, integrate, and share knowledge about a domain, and across different domains. Thus, ontologies are used to represent and reason with complex knowledge in a structured manner (i.e. Ontologies' impact on KR). But also, ontologies use formal logical languages as machine-understandable languages, to explicate the conceptual specification of an ontology. These are indeed knowledge representation formalisms within the field of KRR [Grimm2009] (i.e KR's impact on ontologies). In [Guarino2009a], Guarino introduces the notion of "ontological level" to the four extant knowledge representation

1. Mereology is the study of parts, wholes, and part-whole relations in logic, philosophy, and formal ontology. A comprehensive exploration of this term will be undertaken in Chapter 3, where a more detailed analysis will be provided.

levels (logical, epistemological, conceptual, and linguistic). Examples of languages in the logical level include the KL-ONE [Brachman1989] language ; the frame language which had an influential aspect in defining concepts formally such as introducing the $Is-Arelation$ [Brachman1978]. Other knowledge representation languages, such as the Web Ontology Language (OWL), are specifically designed for representing ontologies. OWL is a language for defining and sharing ontologies on the web, and it provides a rich set of constructs for representing concepts, relationships, and rules in a structured and machine-readable format.

For the SW, and motivated by the growing interest in Linked Open Data (LOD), the technologies existing are a family of knowledge-based approaches that rely on a formal and shared model i.e. ontology [Gruber1993]. Ontologies are a fundamental building block for the SW because they provide a common vocabulary and understanding of concepts that can be used to annotate and describe web resources. Ontologies enable machines to understand the meaning of information on the web, which is critical for tasks such as information retrieval, data integration, and automated reasoning [Horrocks2008]. With ontologies, implicit knowledge can be captured in the SW across heterogeneous data sources, and semantic interoperability can be achieved [Wache2001]. Furthermore, the SW provides a set of standards and technologies for representing and sharing data on the web in a machine-readable format. These technologies include the Resource Description Framework (RDF), the Web Ontology Language (OWL), and the SPARQL query language, among others. By using these technologies, the SW allows different systems to exchange and integrate data in a way that is meaningful and understandable to machines.

Indeed, the underpinnings of key SW standards, such as RDF [Brickley2014] [1] and OWL [Bechhofer2004], are explicitly logical underlying Description logics. For instance, the OWL2 language [Krötzsch2012a] which is the latest version of OWL, is based on the SROIQ logic [Horrocks2006], which is the the most expressive Description Logic [Baader2003,Robinson2001]. This reflects that Semantic Web applications often rely on schema/ontology qualities as their structure (also called backbone), and allows for deductive reasoning [Hitzler2020].

Nevertheless, ontological theories (e.g. [Heller2004, Chisholm1996, Bunge1977]) have been successfully applied to the evaluation of conceptual modeling languages and frameworks such as UML [ISO/IEC195012005], ORM [Halpin2005], ER [Chen1976], and GOL [Degen2001]. Additionally, they have been used for the development of engineering tools such methodological guidelines and modeling design patterns that contribute to the theory and practice of their disciplines [Guizzardi2008]. It is important to see how the different fields interfere and use one another, in order to specify, later, in which field(s) the scope of this thesis contributions fall.

Ontology employment in interdisciplinary approaches

As we have mentioned in the Introduction chapter 1, data integration (sometimes called information or knowledge integration) is a major burden in interdisciplinary applications. For these applications, to make a cooperative work in which experts from each domain can share parts of their data, the multiple data sources must be linked and integrated. This is known as *information interoperability* i.e. the capacity of different information systems, applications, and services to communicate, access, share and interchange data, information, and knowledge in an effective

1. Additional information about RDF and OWL can be founded on `https://www.w3.org/RDF/` and `https://www.w3.org/2001/sw/wiki/OWL` respectively.

way [Wache2001].

Data integration raises several semantic heterogeneity problems, which consider the content of an information item and its intended meaning. Since most heterogeneity problems in information integration are caused by semantic conflicts, then richer semantics of data are needed to resolve this problem.

For achieving interoperability, the creation of a global schema has been proposed and studied as schema integration, also known as schema alignment, or schema matching [Rahm2001]. And to address the need of richer data semantics for resolving conflicts from heterogeneous data sources, researchers have proposed the use of ontologies [Bergamaschi2001,Hakimpour2001,Wache2001]. Compared to schemas, ontologies extract the implicit semantics of knowledge. In [Shvaiko2005], differences and commonalities between ontologies and schema have been made clear e.g. the semantics are not really not considered within a schema's specification. Using an ontology as a shared common model for integrating the multiple disciplines according to a shared understanding is a key to achieving semantic interoperability [Uschold1996,Cruz2005].

In the following we first present the two main data integration approaches in general (schema integration), followed by the three basic ontology-based data integration in particular. Then we show how the specific approaches overlap the general ones, and how this created a fourth ontology-based approach which we employ in this thesis.

Data integration approaches

Data integration is the problem of combining data residing at different sources, and providing the user with a unified view of these data [Halevy2001, Hull1997, Ullman1997]. Most data integration systems focus on architectures based a global schema and a set of sources [Lenzerini2002] in which the sources contain the data while the global schema provides a reconciled, integrated, and virtual view of the underlying sources. And based on *the specification of the mapping* (correspondences) made between the data of the sources and the data of the global schema, two basic approaches have appeared in the literature as systems of data integration : *Local-as-View* (LaV) and *Global-as-View* (GaV). A mapping between two schemas is constituted by a set of assertions of the form $q_1 \rightsquigarrow q_2$, where q_1 and q_2 are two queries of the same arity, expressed in a query language over a schema. [1]

The two approaches (LaV and GaV) are differentiated based on the mapping between the global schema G and the source schema S.

LaV approaches are based on characterizing each source in terms of a query (also called view) over the global schema yielding in one assertion for each element $s \in S$ of source, in the form $s \rightsquigarrow q_G$. This permits changes to source schemas without affecting the global schema, since the local schemas are defined as views over the global schema, but the query processing can be complex. Additionally, from a modeling point of view, the designer concentrates on declaratively specifying the content of the source in terms of the global schema [Lenzerini2002].

Whereas GaV approaches are based on associating each element $g \in G$ in the global schema with a view over the source local schema, in the form $g \rightsquigarrow q_S$. Therefore querying strategies are simple, but the evolution of the local source schemas is not easily supported. and in contrary to LaV approaches, the designer concentrates to specify how to get the data of the global schema by means of queries over the sources.

In each of the approaches, to better characterize each element of the source/global schema in terms of the global schema/sources, some sophisticated assertions (in some formal logical

1. The notion of query is a crucial for capturing a mapping between the global and local schemas. For further reading, we suggest the folowing paper [Calvanese2001].

languages) have been proposed to express complex mappings such as [Abiteboul1998, Calvanese2000] for both approaches.

Ontology-based data integration approaches

Moreover, data integration system can be based on an ontology as the global schema in both LaV approaches [Gruninger2002] and GaV ones [Calì2001].
Using ontologies as shared conceptualizations for data integration approaches is referred to as ontology-based data (or inormation) integration (OBDI). Sometimes the term "ontology-based data integration" is confused with the "ontology mediation" where the distinction between these two terms can be blurry, and different authors may use the terms interchangeably or with slightly different meanings. Basically, the difference lies in the following. Ontology mediation involves using an ontology to mediate between different data sources by resolving semantic heterogeneities and integrating data on the fly [De Bruijn2006]. It is a broad term that encompasses ontology mapping, ontology merging, and ontology alignment. While ontology-based data integration is method exclusively focused on integrating data from different sources possibly using a common ontology as a basis depending on the approach followed. During the process of data integration, the mapping task (mentioned above) between the global and local schemas becomes an ontology mapping task when the schemas employed are ontologies. The integrated data can be used for various purposes, including analysis, querying, and decision making.

Within OBDI systems, different approaches are introduced based on a criteria other than of data integration in general (which is highlighted above based on the specification of the mapping). This criteria is based on *the number and level of abstraction of the ontologies used within an integration system*. This was investigated in [Wache2001], in which the authors distinguish three variants reflecting the number and type of ontologies used for data integration : *single ontology approaches*, *multiple ontology approaches*, and *hybrid approaches*. These variants were identified by studying available OBDI systems examples from various domains in 2001. Figure 1.8 shows the three possible ways of using ontologies within an integration system :

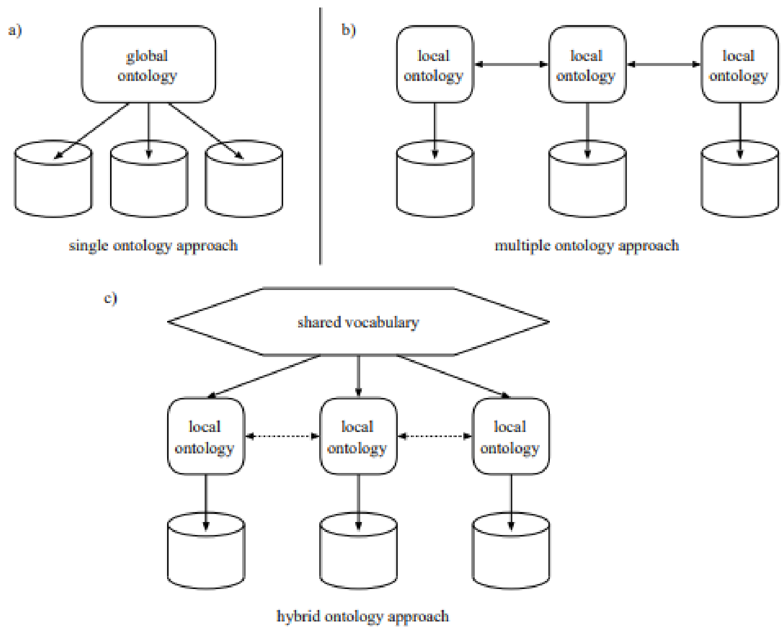


FIGURE 1.8 – *Figure taken from [Wache2001] showing the three possible ways for using ontologies for content explication.*

— *Single Ontology approaches :* these use one global ontology which provides a shared vocabulary for the specification of the semantics, to which all information sources (e.g. databases) are related, such as the OntoBroker frame-based representation system [Decker1999]. This is by relating the objects of each information source to those in the global ontology. These approaches are employed in cases where all information sources share nearly a same view on a domain, rather than different views in which developing a minimal ontology becomes difficult [Gruber1995]. In this approach, domain experts are needed to define and evaluate the semantics of all data sources in terms of the global ontology.

— *Multiple Ontology approaches :* these use multiple ontologies, each describing an information source, without sharing the same vocabulary, e.g. the OBSERVER. This requires inter-ontology mappings which identifies semantically corresponding terms across ontologies considering the different. Having no common and minimal ontology makes it difficult to compare the different source ontologies and to map the semantically heterogeneous views which might acquire different granularity levels.

— *Hybrid Ontology approaches :* these overcome the difficulties of single and multiple ontology approaches, in which the semantics of each information source is described at its own, while building one global shared vocabulary (not ontology) [Goh1997,Wache1999], e.g. the MECOTA/BUSTER [Vögele2003]. In these approaches, the terms of the source ontologies are described in terms of the primitive terms of the shared vocabulary.

In general, each of the three approaches possesses distinct benefits and limitations, and their utilization shall be determined based on the integration requirements within a given system.

In the context of interdisciplinary approaches, the utilization of ontologies has become increasingly prevalent. This is evident through two key observations : (1) each discipline within the multiple disciplines involved tends to possess its own schemas, often in the form of ontologies, which reflect their localized perspectives on the subject matter, and (2) there exists a consensus among these disciplines regarding a common goal and shared understanding pertaining to the (partial) common entity they collectively address.

Consequently, an additional approach appeared necessary to facilitate interdisciplinary integration in fields such as environmental sciences. It serves as an intermediary between the single and hybrid OBDIs approaches. In the subsequent discussion, we will examine the intersections between OBDI and data integration approaches, followed by the fourth variant of OBDI that best aligns with the requirements of our interdisciplinary cultural heritage field.

From "data integration" to "Ontology-based data integration approach"

The two aforementioned criteria for approaches operate at different levels : one is of a general nature, encompassing data integration regardless of the schema type employed, while the other pertains to the specific context of systems utilizing ontologies as their schemas. The former is based on the mapping between the two schemas leading to LaV and GaV data integration approaches. Whereas the latter is based on the number, and level of abstraction, of the used ontologies resulting in single, multiple, and hybrid OBDI approaches.

These two levels overlap as follows. Both single and hybrid ontology approaches are appropriate for cases where building a central data integration system via a global schema is beneficial [Cruz2005]. More precisely, single approaches, being focused on the view of the global ontology, seem more appropriate for a GaV data integration system, whereas hybrid approaches seem to be appropriate for LaV ones. As for multiple ontology approaches, it can be best used to construct "peer-to-peer" data integration systems in which no global schema is built i.e. is neither for LaV nor GaV approaches.

For example, [De Giacomo2018] describes a general framework for *Ontology-Based Data Access* (OBDA) employing the single-ontology approach within a GaV data integration system, where the mappings translate the operations on the global ontology in terms of queries on local ontologies.

Afterwards, other works started to appear and presented GaV approaches using both : a global ontology and local ontologies. This means a GaV approach as view within an OBDI that provides an intermediate approach between the single and hybrid ontology approaches regarding the number and abstraction-level criteria) e.g. [Gagnon2007, Moser2009, Moser2016, Ekaputra2017, Modoni2017]. This is referred to as the "*Global-as-View*" (GaV) ontology-based data integration (stressing on the term "ontology-based" which distinguishes GaV ontology-based data integration from the general GaV approach for data integration presented above), which we present below.

The GaV Ontology-based data integration approach

In [Gagnon2007], the authors propose an ontology-based information integration system with a local ontology for each data source and a global ontology. The system aims at exploiting the global ontology and its integration with local ones via a local to global ontology mapping which falls basically under the GaV data integration approaches, as shown in figure 1.9. The mediator in the figure maps the requests and answers between the global ontology and the local ontologies of the data source elements. Local ontologies capture the information from local data sources (e.g. databases), which may not be captured by the global ontology. Then, mappings are defined in terms of correspondences between the global ontology and local ones. An ontology is defined by a pair $O = (C,A)$ where C is a set of concepts and A is a set of axioms describing the interpretation of the concepts in a given domain. A total mapping from $O_1 = (C_1,A_1)$ to $O_2 = (C_2,A_2)$ can be expressed as a morphism $f : C1 \rightarrow C2$ to semantically relate concept C_1 to C_2, such that, $A_2 \pm f(A_1)$, i.e., all interpretations that satisfy O_2's axioms also satisfy O_1's translated axioms. And a partial ontology mapping form $O_1 = (C_1,A_1)$ to $O_2 = (C_2,A_2)$ if there exists a sub-ontology $O'_1 = (C'_1,A'_1)$ with $(C'_1 \subseteq C_1$ and $A'_1 \subseteq A_1)$ such that there is a total mapping from $O'_1 to O_2$. The approach stresses of the establishment of both, the global ontology and the local ones. The mappings can be expressed by means of various languages, depending on the ontology representation language used within the integration system e.g. OWL [Bechhofer2004].

In [Ekaputra2017], the GaV approach was explained as a fourth OBDI variant and compared to the other three OBDI variants explained above, as shown in figure 1.10. It is based on the GaV approach from the relational databases [Doan2012]. The authors aim at systematically reviewing the literature with a survey reflecting on OBDI applications in the context of multidisciplinary application. Based on their analyses and comparison of 23 OBDI applications in an environmental multidisciplinary field, they categorize the four OBDI variants and compare their strengths and limitations.

The GaV approach requires the definition of one local ontology per data source, similar to the multiple and hybrid approaches, in an independent manner from the global ontology which is defined afterwards. Then interoperability is achieved by defining independent mappings between the local and the global ontologies.

Based on the key characteristics of data integration in their environmental multidisciplinary domain application, they proceed with evaluating each of the four OBDI variants. These characteristics are : semantic heterogeneity, data access, mapping complexity, data source dynamic, and ontology implementation effort, as shown in Table 1.1.

In general, the GaV approach's strength points lie in : (1) supporting various levels of heterogeneity thanks to its global ontology, which is similar to the hybrid approach case although it

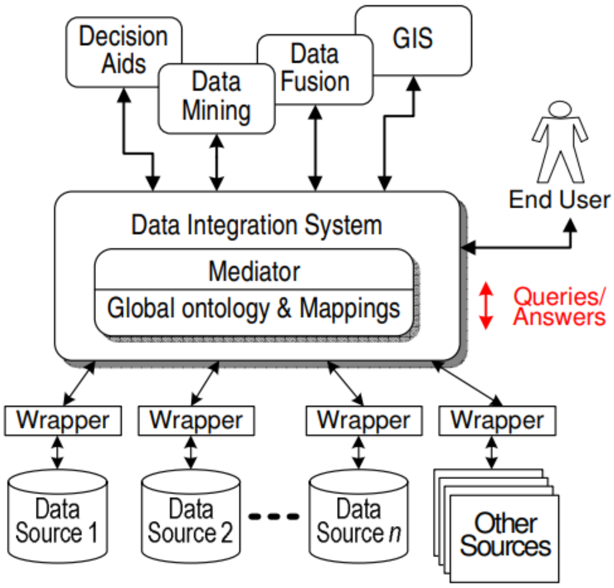


FIGURE 1.9 – *Figure taken from [Gagnon2007] of an ontology-based data integration architecture.*

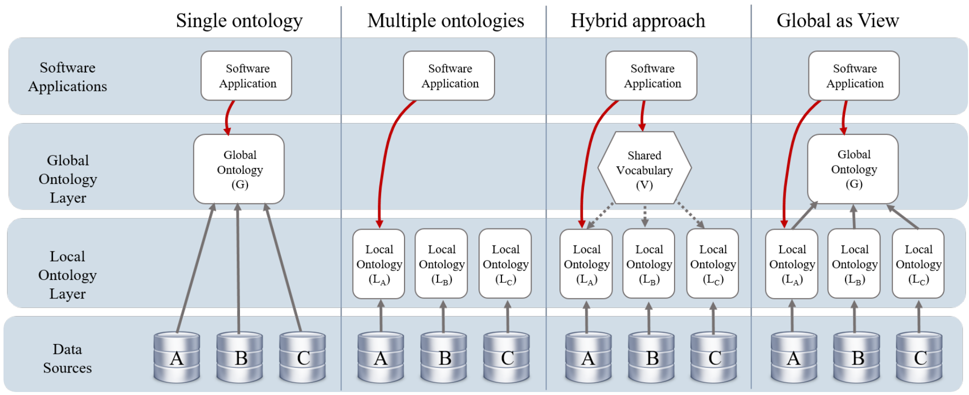


FIGURE 1.10 – *Figure adapted from [Ekaputra2017] showing the three variants of OBDI : (1) single-ontology, (2) multiple-ontology, (3) hybrid, and an additional OBDI variant (4) Global-as-View (GaV).*

does not encompass an ontology but a vocabulary, (2) providing access to data at both local and global ontology levels, and (3) supporting simple and complex mappings using queries which is probably the most important characteristic in an integration system. Concerning the dynamics of data sources, GaV approaches are noted to acquire a slight limitation of an additional ontology level : the definition of local ontologies. This leads in costly implementation efforts (at several levels) which appear to be a major limitation, especially in cases where all ontologies are made from scratch.

However, the GaV approach for OBDIs has demonstrated its utility and has been successfully applied and implemented in notable studies [Dubinin2014, Ekaputra2016, Lin2007].

	Single-ontology	Multiple-Ontology	Hybrid-Ontology	GaV OBDI
Semantic Heterogeneity	best applied for data sources with similar view of a domain	support heterogeneous views	support heterogeneous views	supports heterogeneous views
Data Access	only allows access to global data	allows access to each local ontology and to the aggregated local ontologies	allows access to each local ontology and to the global ontology	allows access to each local ontology and to the global ontology
Data Source dynamics (& data source types)	requires modifications in the global ontology	needs to provide a new local ontology & map it to other local ontologies	only requires a local ontology based on the shared vocabulary	requires a new local ontology and mappings to the global ontology
Mapping Complexity	N/A	supports simple mappings (semantic relations)	supports simple mappings (vocabulary refinement)	supports simple and complex mappings (queries and rules)
Ontology Implementation effort	straightforward	costly	reasonable	Rather costly

TABLEAU 1.1 – *Table from [Ekaputra2017] showing the characteristics, strengths and limitations of OBDI variants. (Green : strengths, yellow : slight limitations ; red : limitations).*

Conclusion

The content presented in this chapter has provided us with a comprehensive understanding of ontologies, encompassing both their role in the modelisation phase and their application in data integration. In terms of modeling, we have examined the fundamental aspects of ontologies, i.e. the shared conceptualization and the modeling language, as well as the "meta"-abstraction level. Regarding data integration, we have explored various approaches that account for ontologies within ontology-based data integration systems, with favoring the Global-as-View (GaV) approach as the most suitable choice for interdisciplinary applications like cultural heritage. This preliminary chapter equips us with the necessary knowledge to make informed decisions concerning the elements that wll be discussed in Section B, and more precisely in Chapter 4.

The next Section A is dedicated to a general state of the art and its analysis : the ontology modeling approaches that have been proposed in the Cultural Heritage field (Chapter 2), and the foundational ontological relations (structural and spatial) that have been studied in the Applied Ontology literature (Chapter 3).

SECTION A :
***
STATE-OF-THE-ART

The following section provides an essential overview of the state of the art divided into two parts : cultural heritage models and foundational ontological relations, preparing and guiding the reader for the subsequent section.

2

Ontology models for Cultural Heritage entities

This Chapter explores the ontological models constructed within the cultural heritage field and reviews the pertinent existing models.

Sommaire

2.1 Introduction

In Chapter 1, we have discussed the interest of building and utilizing an ontology in general, and a meta-ontology in particular, as a suitable modeling tool for fostering a shared understanding across multiple disciplines. Thereby, facilitating an interdisciplinary approach centered around a cross-disciplinary entity. Emphasis has been placed on the significance of establishing a shared goal as the foundation for constructing the ontology. In this regard, we have identified Patrimalp's objective of understanding and representing the material aspects of a tangible cultural heritage entity. Based on that, we have generalized and established the fundamental objective of the thesis as representing and modeling the structural and spatial settings of a tangible entity.

Furthermore, we have discussed the importance of selecting an appropriate employment approach for the ontology within an integration system, in the context of an ontology-based data integration system aligned with the requirements of an interdisciplinary approach. Specifically, our focus has been on systems that achieve semantic integration of their ontologies and data by adhering to a meta-ontology within a global view.

Later in the Preliminary Remarks section, we illustrated the foundations of ontology as a shared conceptualization and a modeling language. We also delved into various paradigms of ontology-based data integration to leverage the employment of a meta-ontology.

Now refocusing our attention on the field of Cultural Heritage (CH), and guided by the utilization of ontologies as a tool for modeling and fostering an interdisciplinary approach, we will delve into the various ontology modeling approaches that have been put forth in the extensive literature. As a result, we clarify below the [motivation-I] behind this chapter.

Motivation-I :
Modeling the composition of a tangible entity as a complex structure (i.e. an object, a place, a collection) in a manner that enables the understanding, constructing, and navigating into its tangible discourse (i.e. representation), and learning its intangible aspects (i.e. significance).

First, we present the overall challenges and needs for a interdisciplinary modeling approach in the CH field, highlight the issue of data heterogeneity and the need for interoperability as the key for a semantic based data integration (Section 2.2). Then, we shift to the management and organization of CH data (Section 2.3), in which we (a) present organizations, institutions, and information systems which manage CH data (Section 2.3.1), and (b) differentiate them from knowledge organization systems which structure and organize CH data (Section 2.3.2). After that, we narrow down to our focus on **ontology modeling approaches** as structures for modeling CH data and entities (Section 2.4) in which we systematically classify ontology models based on a set of criteria (Section 2.4.1), followed by excerpting and reviewing a three main ontology approaches that are most relevant with respect to our objective (Section 2.4.2). In conclusion, we synthesize the discussed topics and shed light on the requirement for further progression (Section 2.5).

2.2 Overall challenges and needs

In this section, we discuss the challenges and needs associated with managing and organizing CH data, emphasizing the importance of an interdisciplinary and semantic approach to address the heterogeneity of the data and achieve a semantic-based data/schema integration system.

Information about CH data is characterized by a wide range of highly specialized and overlapping systems, leading to challenges related to mutual accessibility, interchange, and integration of information. Two main challenges arise in this context : (a) the heterogeneity of CH data based on its semantics and interdisciplinary nature, and (b) the problem of integrating data and schemas across these diverse systems.

2.2.1 Data heterogeneity and integration problems

The heterogeneity of CH data poses a significant challenge to researchers, as it encompasses "diverse data types", including (a) textual such as descriptions, transcripts, and translations of historical documents, (b) visual such as images, photographs, and videos of artifacts or archaeological sites, (c) audio such as recordings of music, oral traditions, or interviews with local communities, and (d) spatial data such as geographic information system (GIS) data, 3D models, analytical data, and maps. Additionally, cultural heritage data can encompass metadata, annotations, and contextual information that provide additional layers of information about the cultural artifacts or practices.

Moreover, the "variability of data formats and structures", such as databases, spreadsheets, XML, RDF, or JSON, adds complexity to the integration process, as different systems and repositories may employ their own data management practices, standards, and protocols.

This heterogeneous nature of CH data restricts the ability for researchers to combine, compare, and contribute to consensual findings. Indeed, the potentials of this data when combined, supports the generation of knowledge relative to any period of times, geographic location, and aspect of past human activity [Bruseker2017].

Moreover, given the cross-disciplinary nature of CH research (as illustrated in Chapter 1), it is crucial to aggregate information, knowledge, expertise, and efforts within the CH field. This aggregation is essential for fostering a sense of shared purpose among the CH community, characterized by a collective commitment to the scientific analysis and presentation of the human past through empirical evidence [Doerr2009]. Consequently, the integration of data and schemas becomes imperative in order to address this need for unity and collaboration across disciplines within the CH field.

The CH field faces significant hurdles when it comes to recognizing the importance of shared search terms. Despite ongoing efforts, there remains a need for substantial work to be done in this area. One of the primary difficulties lies in achieving consensus on a common terminology [Doerr2009]. Dealing with diverse aspects of CH necessitates the utilization of extensive and nuanced language, often specific to particular communities or individual fields. Consequently, establishing an agreement on standardized terminology proves challenging due to the absence of equivalent terms in different languages. This lack of shared vocabulary, in turn, complicates the development of a unified conceptualization, making it to establish a common language structured as a hierarchical system.

The integration problem extends to metadata integration, where different institutions use various metadata schemas, controlled vocabularies, and descriptive standards. Indeed, the complexity of the CH domain has led to the adoption of different metadata schemas and models and to the use of large number of different value vocabularies i.e. thesauri, authority files and controlled lists (as will be illustrated in Section 2.3). This yields in greatly complicating the identification of a model for the integration of metadata collections and it making it difficult to develop aggregations during the integration of the different metadata sources [Peroni2013]. With this subject of metadata aggregation being important the focus of of the digital library community is highlighting the needs of **syntactic and semantic aspects of interoperability** of heterogeneous digital libraries e.g. [Candela2007] and to the creation of of aggregation frameworks that give access to heterogeneous metadata collections and expose them as integrated datasets [Brogan2005].

In [Doerr2009], it is suggested that CH terminology can be divided into an upper level that remains stable for search purposes and a lower level that is more flexible for hypothesis building. This division is based on insights gained from developing the ISO21127 CIDOC CRM [ISO211272014] and various information systems, as well as the experience with the Art & Architecture Thesaurus (AAT), which is the largest and most stable thesaurus in the field. [1] To address these challenges, the concept of shared core ontologies (explained in Section Preliminary Remarks) has been proposed to facilitate interoperability among different domain ontologies [Guarino1998a]. Ongoing projects, such as the British STAR project [Binding2010b], aim to investigate cross-search capabilities using the CIDOC CRM core ontology as an integrating framework.

2.2.2 The need for interoperability

To achieve interoperability and address the challenges of heterogeneity and data/metadata/-schema integration, a semantic approach is essential. Semantic interoperability ensures that information systems communicate data consistently with its intended meaning, enabling effective data integration, interoperability, and knowledge sharing. Syntactic interoperability focuses on compatibility between encoding and access protocols, while semantic interoperability ensures the consistent interpretation of data by considering its structure, terminology, and identifiers. Resolving semantic heterogeneity is crucial for achieving interoperability and enabling seamless communication and integration among different systems [Patel2005].

And as presented in the Introduction Chapter 1, ontologies have gained widespread acceptance as a means to ensure data interoperability and facilitate the efficient discovery and sharing of domain knowledge among interconnected sources [Moraitou2019]. This goes back to the shared consensus upon which ontologies are based providing a shared conceptualization of a knowledge domain. By utilizing ontologies as structures that organize knowledge, CH data can be effectively managed, categorized, and aggregated, enabling researchers to derive meaningful insights and knowledge from the integrated data.

In conclusion, the interdisciplinary nature of CH data and its inherent heterogeneity necessitate an interoperable and semantic approach for effective management, organization, and integration. By leveraging ontologies and adopting semantic interoperability, researchers can overcome the challenges associated with data heterogeneity and achieve meaningful integration and exchange of CH data, enabling new insights and discoveries in the field of Cultural Heritage.

1. Both of which will be presented in Section 2.3.

2.3 Managing and organizing CH data

In this section, our initial objective is to elucidate several distinct institutions responsible for the management of CH data. Subsequently, we proceed to outline various knowledge organization systems (KOS) that enable the organization of data and examine the approaches that have been designed and implemented in the context of the CH field. Ultimately we refine our analysis to concentrate specifically on ontology models, for the upcoming sections.

2.3.1 Managing CH data : organizations, institutions, and systems

To address the challenges of the heterogeneity of CH data, data standards, best practices, and guidelines have been developed by various organizations and initiatives. These efforts aim to promote data interoperability in the CH field, and facilitate the exchange, integration, and reuse of CH data.

Memory institutions :

Cultural heritage data are records about CH entities stored as content that is packaged in different object types (e.g. books, artifacts, videos, music, etc.) and managed by organizations of different types (e.g. libraries, museums, media companies, archives, etc.) [Mäkelä2012]. The preservation of this content is carried out by institutions called *memory institutions* i.e. museums which hold primary evidence for establishing and furthering knowledge [1], Sites and Monuments Records (SMR) departments of a Ministry of Culture, which pursue similar goals as museums, but for immobile sites, archives and libraries which maintain large amounts of original material – mostly written and image content – in their historical order, such as administrative records, letters from VIPs, photographic collections [Doerr2009]. Their international umbrella organizations which maintain the specific documentation policies of cultural heritage content are : the International Council of Museums (ICOM) (several committees form the ICOM, one of which is the CIDOC committee), the International Federation of Library Associations (IFLA), and the International Council of Archives (ICA).

Organizational Institutions and Infrastructures :

Several other organizational institutions and infrastructures contribute significantly to the management, preservation, and accessibility of CH data, playing a pivotal role in advancing research and knowledge in the field. From which we highlight some notable examples :

— The ARIADNEplus, which is an extension of the previous ARIADNE Integrating Activity, successfully integrates archaeological data infrastructures in Europe, indexing around 2,000,000 datasets in its registry through the ARIADNE portal2. ARIADNEplus aims to further develop and support the research community established by the previous project, while strengthening relationships with key stakeholders, including European archaeological associations, researchers, heritage professionals, and national heritage agencies.

— The Getty Research Institute is a prominent organization that actively contributes to the development of various knowledge organization systems in the CH field. Their expertise and resources make them a significant player in advancing CH knowledge organization.

— The Bibliothèque Nationale de France (BNF) is the national library of France, serving as the central repository for all published materials in the country. It houses extensive historical collections and plays a pivotal role in preserving and providing access to France's cultural heritage. For more information, visit the BNF website.

1. https://icom.museum/en/resources/standards-guidelines/museum-definition/

— The Bibliothèques Nationales européennes (BBF) is a European national library that operates within concentric circles, spanning from the local to the international level, including the national and European contexts. Given the evolving global information society, BBF's interconnectedness with various levels of libraries enables them to effectively meet the information needs of diverse users. Explore more about BBF on their website.

— The Institut national de l'information géographique et forestière (IGN), which is formerly known as Institut Géographique National, is a French public state administrative establishment established in 1940. It is responsible for producing and maintaining geographical information for France and its overseas departments and territories. IGN plays a vital role in providing geospatial data and services that contribute to CH research and preservation. Learn more about IGN on their official website.

Information Systems and databases :

Based on [Doerr2009], CH information systems can be distinguished into four categories according to their major functions with respect to CH information. [1] The first category includes collection management systems, mostly relying on relational or hierarchical database systems, and support the technical management and administration of collections/sites/monuments (e.g. acquisition, exhibitions, protection zones, etc.) [Grant1994, Grant1995]. The second category includes those systems responsible for conservation information i.e. the scientific, material analysis of the objects, preventive measures and interventions which allow scientists, such as art conservators, to accumulate and exchange their knowledge [Viñas2002]. The third category revolves around research information systems which are responsible for the building of uniform descriptions (e.g. the Union List of Artist Names [Bower1994]), integrating information from several resources for specific purpose analysis (e.g. statistical), geographical reasoning for archaeological site prediction, and running automatic classification (see for instance, [Doerr2002, Hermon2002]). And the last category includes presentation systems which give access to CH information to the general public such as portals.

A popular example is the Ishtar project for managing archaeological data and documentation (including archaeological finds) from archaeological operations. It takes shape as a *free software for archaeological data management based on a database* [2] under AGPL [3] 3.0 or any later version license.
The user interface depicted in figure 2.1 offers different *functionalities* depending on the chosen *module* [4]. Although it is organized around a common core (the database), yet it is associated with multiple modules linked to specific professional needs : administration of operations and inventories, archaeological warehouses, treatments related to restoration laboratories, advanced stratigraphic analysis, QR-code labeling, etc. Multiple levels of users are possible, from a public access (or not) to access for researchers, operations managers, warehouses managers, GIS connection, etc., with precise reading and writing rights, defined by geographical area or at the scale of the archaeological excavation, or by module. As for its functionalities, Ishtar allows for the entry of excavation data, context records, and archaeological finds data, as well as media management (pictures, reports, etc.), among others.
The software does not impose a particular workflow and adapts to multiple purpose entities such as

1. One should distinguish between data, information and knowledge. According to D. Soergel [Soergel1985], data is the form and information is the content, whereas knowledge has structure that ties together and integrates individual pieces of an image of the state of affairs and is the basis for action.
2. Downloadable from their public GitLab repository : https://gitlab.com/iggdrasil/ishtar.
3. https://en.wikipedia.org/wiki/Affero_General_Public_License
4. https://ishtar.readthedocs.io/fr/main/index.html

research/preventive excavations (data management, data pooling, etc.), research/restoration laboratory (close management of treatments, traceable records), archaeological warehouses (management of archaeological finds inventories, operations, planning files, reports, etc.), students (access to a free standardized database), and other possible users and uses of the archaeological data.

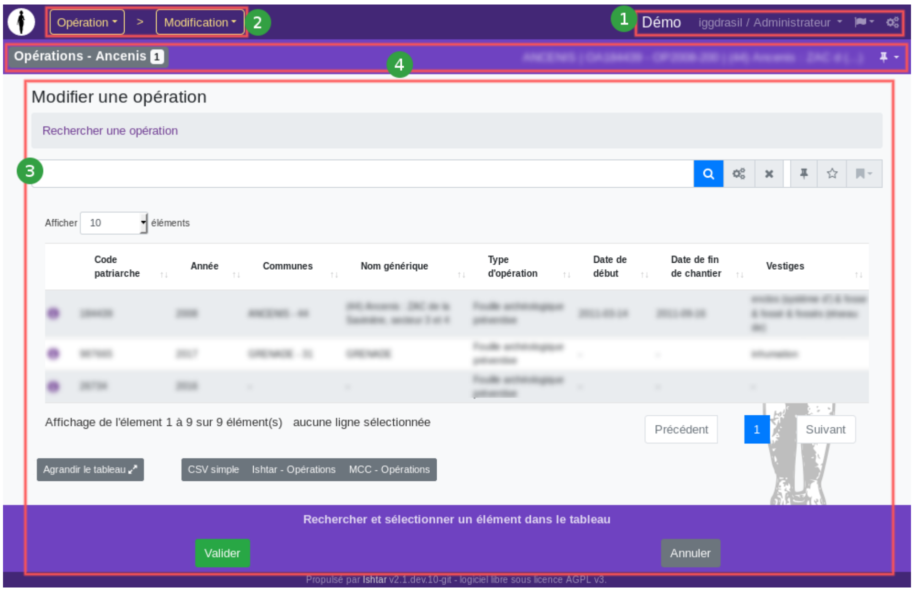


FIGURE 2.1 – *Figure taken from the documentation of the Ishtar software showing a screenshot of the "desktop" interface of the Ishtar software.*

Other online databases and online resources which handle CH (related) data include (a) Artefacts ; an online collaborative Encyclopedia of archaeological small finds, (b) Webminira ; a website representing a database for describing and presenting information about minerals and their composition, (c) the Romanian Database of Raman Spectroscopy (RDRS) ; a database for materials and composition, (d) the EarthCem ; an online tool that provides open data services to the geochemical, petrological, mineralogical, and related communities. Services include data preservation, discovery, access, and visualization, (e) the Enluminures database offering free consultation of more than 120,000 images, in the form of of vignettes and full screen, digital reproductions of illuminations and elements of decoration of more than 5,000 medieval manuscripts preserved in a hundred French municipal libraries, (f) Chelabs ; an open libratory innovative system for the CH safeguard and valorization promoting an approach based on open access and shared culture, (g) the Symogih project ; a modular historical information management system which developed a generic model for storing historical data allowing their interoperability and selective publication, and (h) Arches ; an open source data management platform for the heritage field, among others.

2.3.2 Organizing CH data : Knowledge Organization Systems

The increasing demand for using CH data stored in records has amplified the requirement for structured and controlled data entry and representation. Typically, this necessity is addressed by prioritizing the digitization of CH data, i.e. the transitioning from physical artifacts to digital format [Navarrete2013], organizing and structuring the data in a manner that enables the integration from different sources, and subsequently managing this data for an effective utilization. This is indeed the case of other interdisciplinary data records, such as medical data, where terminological systems are adopted [de Keizer2000]. The term "terminological systems" used in that context is wide and bears different types of *organization systems* as an umbrella including classification systems, thesauri, (controlled) vocabularies, and arriving to formal systems i.e. ontologies. [1] To examine organization systems in the CH field, we first distinguish their types.

In computer science, these systems are referred to as "Knowledge Organization Systems" (KOS) and encompass all types of schemes for organizing information and promoting knowledge management. The term was explored by Hodge in [Hodge2000a] in the context of digital libraries, in which the author examines the importance of these systems in improving information retrieval and navigation within digital library environments, and provides insights into the design, implementation, and evaluation of KOS in the digital context. In his later note "Taxonomy of Knowledge Organization Sources/Systems" [Hodge2000b], Hodge distinguishes four mutually non-exclusive groups of KOS, based on the complexity of their *structures* and major *functions* [Hill2004]. This understanding was adopted and standardized by *Networked knowledge Organization Systems/Services/Structures* group (NKOS), and comprehensively presented by Zeng in [Zeng2008]. Thus, based on their structure, KOS are organized from simpler to more complicated structures as :

— Term Lists : these involve lists of terms with (a) some sequential order i.e. pick lists, (b) alphabetical order and their definitions i.e. dictionaries/glossaries, or (c) equivalences i.e. synonym rings.

— Metadata-like Models : These include lists of terms representing (a) names and their associated contact information i.e. directories, (b) variant names for an entity or the domain value for a particular field i.e. authority files, or (c) named/typed places as geospatial dictionaries i.e. gazetteers.

— Classification and Categorization : These emphasize the creation of subject sets as schemes (a) providing a set of controlled terms to represent subjects of items in a collection and the set of rules for combining terms into compound headings i.e. categorization schemes, (b) dividing items into ordered groups based on a particular characteristic i.e. taxonomies, and (c) arranging numerical or alphabetical notations in a hierarchical or faceted manner to represent broad topics i.e classification schemes. The formalization and comprehension of the relation holding between categorical statements with the analysis and definition of categories and their shared properties, represented by tree-like structures [Rosch1978]. They provide controlled vocabularies and hierarchical frameworks for classifying and categorizing CH data, enabling consistent and standardized representation.

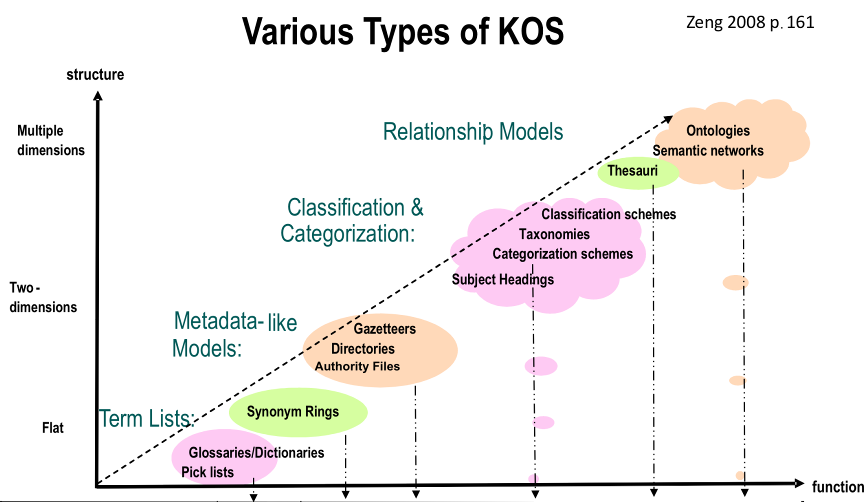
— Relationship Models : These encompass the connections between terms and concepts as (a) thesauri which are sets of terms representing concepts and the hierarchical, equivalence, and/or associative relationships, (b) semantic networks which are sets of terms representing concepts modeled as nodes with variable relationship types, and (c) ontologies which are formal models of concepts and relationships, as well as rules and axioms.

As for their functions, a KOS structure can be designed to fulfill multiple functions such as eliminating ambiguity, controlling synonyms, establishing relationships (hierarchical and associative) and/or presenting properties. Hierarchical relationships are based on degrees of superordination and subordination [NISO2005, Iyer1995]. And according to [Zeng2008], they cover three

1. Extensive literature has addressed the terminological confusion surrounding information problems, particularly in relation to the distinctions between thesauri, taxonomies, and ontologies [Gilchrist2003]. Scholars have also explored the relationship between thesauri and ontologies [Arano2005], clarified the concepts of classification systems and metadata taxonomies in relation to ontologies [Madsen2009], and compared ontologies to terminologies based on principles, methodology, formality, and complexity perspectives [Zemmouchi-Ghomari2012]. In this manuscript, we adopt the perspective that thesauri, classification systems, controlled vocabularies, metadata models, ontologies, and other terminological systems share a common underlying structure referred to as a terminology, with varying degrees of complexity, formality, and additional characteristics, as discussed in [de Keizer2000].

primary logically different and mutually exclusive conditions : generic relationships identifying the link between a class (category) and other categories that fall under it as sub sub-species using the notion "*IsA*", instance relationships identifying the link between a class (category) and its instances (entities that are classified as belonging to this category) using the "*IntanceOf*" notion, and the whole-part relationships covering the context in which one category is inherently included in another and thus linking them logically.

Figure 2.2 shows an overview of the structures (from simple to more complex : flat, two-dimensional, and multiple-dimensional) and functions (ambiguity elimination, controlling synonyms, hierarchical relationships, associative relationships, and properties presentations) of KOS. It visualizes a summary of [Hodge2000b] adopted by the NKOS group based on Hodge's article on KOS [Hodge2000a].



Major functions						
eliminating ambiguity	xxx		xxx	xx	xxxx	xx
controlling synonyms		xxxx	xxx	xx	xxxx	xx
establishing relationships: hierarchical			x	xxxx	xxx	xxx
establishing relationships: associative					xxxx	xxxxx
presenting properties						xxxxx

FIGURE 2.2 – *Figure taken from [Hodge2000b] after [Zeng2008] showing the types of knowledge organization systems (KOS), arranged according to the degree of controls introduced (from natural language to controlled language) and the strength of their semantic structure (from weakly structured to strongly structured), corresponding to the major functions of KOS.*

Since KOS are mechanisms for organizing information, they are at the heart of every library, museum, and archive. Their primary objective is to either create or extract vocabularies in scenarios where they are absent, or merge and map existing vocabularies within diverse systems that differ in terms of their structure, domain, language, or granularity [Lei Zeng2004].

Table 2.1 showcases a selection of prominent KOS tools used by various memory and information institutions. This compilation is based on the comprehensive study conducted by Bruseker et al. [Bruseker2017], which extensively addresses the challenges associated with their implementation. The highlighted tools encompass a range of traditional KOS, including controlled vocabularies, taxonomies, thesauri, metadata, and data schemas, along with ontologies.

In the following section, we present a selection of KOS structures derived from the literature [Zeng2008, Hodge2000b, Hodge2000a]. These structures are further illustrated through examples

	Library	Museum	Archives
Controlled Vocabulary	Library of Congress Name Authority File, Authority List for Journal Titles	The Revised Nomenclature for Museum Cataloging, Gazetteer of British Place Names	A Glossary of Archival and Records Terminology
Taxonomy	Dewey Decimal Classification	Traditional Biological Taxonomy	
Thesaurus	LCSH	AAT	UKAT
Metadata and Data Schema	Dublin Core, UniMARC, METS	Core Data Index to Historic Buildings and Monuments of Architectural Heritage, MIDAS Heritage, CDWA	EAD

TABLEAU 2.1 – *An illustration of well-known examples of different types of KOS used in memory institutions*

relevant to the CH field, incorporating insights from existing literature [Patel2005, Bruseker2017]. The examination of these KOS structures takes into account their inherent structural characteristics and aligns with the relevant scholarly discourse on the subject matter.

Authority files :

Authority files are considered quite simple structures that serve as metadata models. Libraries and information services have a history of creating authority files to establish forms of names (for persons, places, meetings, and organizations), titles, and subjects used in bibliographic records [Zeng2008].

A name authority file is a type that acts as intermediate files linking names to bibliographic information by controlling the variants of personal names and serve as tools for catalogers and indexers. These ensure that the proper form of the name, rather than an unapproved variant, is used and bring together all works by or about the entity whose name is represented. For instance, for a digital library of images of artists' works, name authority files contain integrated variant names which can be searched according to the name appearing in the digital library collection, and allows for displaying and providing a wide range of contextual material about the artist to the user [Hodge2000a].

VIAF. The Virtual International Authority File (VIAF) is an example of a directory that combines multiple name authority files into a single OCLC-hosted name authority service [1]. The goal of the service is to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web.

CONA. The Cultural Objects Name Authority (CONA) focuses on names and terminology related to cultural objects. It provides controlled vocabulary for describing and accessing cultural objects, including art objects, artifacts, archaeological finds, and historical items. CONA supports the cataloging and retrieval of cultural heritage data, ensuring consistency and precision in object descriptions.

1. OCLC, referring to Ohio College Library Center, is an nonprofit organization providing shared technology services, original research, and community programs.

ULAN. the Union List of Artist Names (ULAN) contains more than 293,000 names with biographical and bibliographic information on artists and creators across different disciplines, including painters, sculptors, architects, photographers, and performers. It provides standardized names, biographical information, and related metadata for individuals and corporate bodies associated with artistic production.

Both CONA and ULAN are developed by the Getty Research Institute managed by the Getty which is an international cultural and philanthropic organization dedicated to the preservation, study, and dissemination of art and cultural heritage. It encompasses several distinct entities, including the Getty Museum, Getty Research Institute, Getty Conservation Institute, Getty Foundation, and Getty Publications. Getty supports and promotes art historical research, critical analysis, and interdisciplinary studies through its library, archives, and research programs.

Classification and Categorization schemes :

Moving to a higher complexity of KOS structures, classification and categorization schemes aim at hierarchically structuring systems with the emphasis on the creation of subject sets. The terms taxonomy, classification, and categorization have been used interchangeably by different disciplines and professions. Although their differences are subtle, these types all provide ways to separate entities into broad topic level, without any explicit relationships (which are normally found in thesauri) [Hodge2000a]. However, the relationship upon which they structurally rely on is the subsumption relationship for ordering a diverse set of entities. It can be generic/individual type of relationship to express predication e.g. "Socrates is a man", or a generic/generic relationship to express sub-type e.g. "a car is a vehicle" [Brachman1983].

In libraries and information services, famous universal classification schemes include :

IconClass. The IconClass is a multilingual classification system developed for the classification of visual art and iconography. It provides a hierarchical structure for categorizing and indexing subjects and themes in art, including religious and mythological iconography.

UDC. The Universal Decimal Classification (UDC) : The Universal Decimal Classification is a comprehensive classification system used in libraries and information centers. It covers a wide range of subjects, including those related to cultural heritage. UDC combines both numeric and symbol notations to represent different subject areas.

DDC. The Dewey Decimal Classification (DDC) system is another widely used scheme for organizing library collections. It covers a broad range of subjects, including cultural heritage topics. The DDC system assigns a unique number to each subject area, allowing for easy browsing and retrieval of materials.

LCC. The Library of Congress Classification (LCC) system is widely used in libraries, including those in the cultural heritage sector, to organize and categorize books and other resources. It covers various subject areas, including art, history, archaeology, and literature.

Thesauri :

At a higher structure complexity level, thesauri play a crucial role in the CH field by standardizing controlled and structured vocabularies displaying hierarchical, associative, or equivalence relations among CH terms/concepts for the organization, classification, and retrieval of CH data. These relationships are generally represented by the notations "BT" (broader term), "NT" (narrower term), "SY" (synonym), and "RT" (associative or related term). More specific relationships can be developed as sub-relations e.g. more than 40 relationships have been defined as associative ones in the Unified Medical Language System (UMLS) from the National Library of Medicine. On the other hand, a thesaurus such as *the Roget's thesaurus* [1] In addition to these relations, thesauri often have a preferred term (called descriptor) and alternative terms (called synonyms) to improve search and retrieval.

In the context of CH data, thesauri have served as valuable tools for enhancing information discovery, interoperability, and knowledge representation [Patel2005]. Several thesauri have been developed by the Getty Research Institute providing vocabularies as Linked Open Data, XML, Relational tables, and through APIs [2].

AAT. The Art & Architecture Thesaurus (AAT) is one of the most widely used thesauri in the CH field. It encompasses a vast range of terms related to art, architecture, and material culture. AAT provides controlled vocabulary for describing and accessing art objects, architecture, visual materials, and conservation practices. Its hierarchical structure and relationships enable semantic navigation and improved data interoperability.

TGN. The Thesaurus of Geographic Names (TGN) thesaurus is also developed by Getty and focuses on geographic locations and features. It provides standardized terms for describing places, including historical, archaeological, and cultural heritage sites. The TGN facilitates the linking of cultural heritage data to specific geographic locations, supporting research, analysis, and visualization in a spatial context.

Although TGN is constructed in a thesaurus format, its is also considered as a gazetteer. Indeed, Gazetteers can be regarded as a special kind of authority file in the form of a spatial dictionary of named and typed places. With the development of digital libraries, digital gazetteers now have extended to become a service where relationships between places are represented inherently through geospatial representations as well as through explicitly stated relationships such as "IsPartOf ".

Heritage Data. another national cultural heritage thesaurus is the Heritage Data which provides vocabularies as Linked Data [3] used by national organizations and local authority Historic Environment Records.

Pactols. A thesaurus of the form of a reservoir of controlled keywords managed with the Opentheso software as a multilingual thesaurus manager developed by the Technological patform "Semantic Web & Thesauri" (WST). PACTOLS is a polyierarchical, multilingual, standardized, and interoperable thesaurus for archaeology, from prehistory to the contemporary period and the sciences of Antiquity. It consists of almost 60,000 concepts organized into 6 micro-thesauri (Peoples and cultures, Anthroponyms, Chronology, Toponyms -list not managed here-, Works, Places, Subjects).

Semantic Networks :

In addition to thesauri, **semantic networks** are other types of relationship models which focus on structuring concepts and terms as a network or web rather than as hierarchies. Their relation-

1. According to the Merriam webster https://www.merriam-webster.com/words-at-play/rogets-thesaurus, the Roget's thesaurus was intended by Peter Marl to be a classification of all knowledge.
2. The Getty vocabulary data are available on the Getty Vocabulary page online at https://www.getty.edu/research/tools/vocabularies/index.html, through online tools e.g. a SPARQL end point http://vocab.getty.edu/, in LOD formats at https://www.getty.edu/research/tools/vocabularies/lod/index.html, in addition to other formats e.g. XML and relational tables at the download center : https://www.getty.edu/research/tools/vocabularies/obtain/download.html
3. The use of the term "Linked Data" here refers to that within the context of the Semantic Web, as introduced by Tim Berners-Lee in 2006 as the means by which data can be represented as resources (using URIs) on the web, related to other linked data, and accessible.

ships go beyond the ones in thesauri to include other specific types e.g. whole-part, cause-effect, or parent-child relationships. The most famous semantic network is Princeton University's WordNet, used widely in natural language processing and information retrieval applications. It organizes words into synsets (sets of synonyms) and represents relationships between words, such as hyponymy (is-a) i.e. the generic hierarchical relationship, hypernymy (is-a-kind-of) i.e. the instanceship hierarchical relationship, and meronymy (part-of) i.e. the whole-part relationship.

Ontology models :

Ontologies represent the newest label attached to KOS as a relationship model type, placed at the extremity of the scatter plot shown in Figure 2.2, i.e. resembling a multi-dimensional structure and acquiring the most of functionalities out of KOS types. It is characterized by their formal nature, semantic richness, and intricate structure, enabling a broad spectrum of functionalities. To differentiate ontologies from semantic networks, Hodge provides valuable insight by stating, "*The knowledge-management community is developing ontologies as specific concept models. They can represent complex relationships among objects, and include the rules and axioms missing from semantic networks.*" [Hodge2000a]. As an ontology is a specification of a shared conceptualization for a shared domain of disocurse, including definitions of classes, relations, functions, and other objects [Gruber1993, Studer1998] i.e. they do not only represent complex relationships between objects, but also rules and axioms.

In comparison to other models such as taxonomies and thesauri, ontologies embrace the same classificatory structure with additionally presenting the properties of each class within each classificatory structure. This unique characteristic enables ontologies to serve as both a conceptual vocabulary and a practical framework for storing, searching, and reasoning.

Over the past decade, the field of CH has gradually adopted knowledge representation methods and semantic web tools for building ontological models that structure and manage CH data [Bruseker2017]. Some ontologies are influenced from extant traditional KOS like thesauri, or existing databases that have previously managed and organized data within respective institutions. In this manner, ontology development renders the existing structure and terminology to facilitate ontology development and integration.

A notable example is the Getty Art and Architecture Thesaurus (AAT), which serves as a foundational resource for constructing ontologies specific to art artifacts. In [Wielinga2001], the authors reported a project building an ontology prototype based on the existing AAT and Visual Resource Association's (VRA) Core Categories metadata element set, with the purpose of creating a knowledge-rich description of art objects. The ontology contained a taxonomy of a subset of the art-object descriptions in AAT, those related to antique furniture and a template showing the properties of class "furniture". This template includes the 17 VRA Core metadata elements and eight additional elements defined by the project. Indeed the authors argue that annotation of large amounts of information sources with knowledge rich meta-data should be based on a rich metadata structure in connection with an ontology. And since thesauri have been built in such a large CH domain, thesauri can be a basis of such construction under satisfying certain criteria.

The subsequent presents an exhaustive list of the predominant ontology models in the CH domain. Each model is succinctly introduced, outlining its roots, current status, scope, and intended objectives. However, for the examination of each approach regarding our objective, a detailed analysis is provided in Section 2.4.

CIDOC CRM. The CIDOC [1] Conceptual Reference Model (CIDOC CRM) [Doerr2003a] is a formal ontology intended to facilitate the integration, mediation and interchange of heterogeneous CH information. [2] The development of CIDOC CRM followed a bottom-up approach, involving the reengineering and integration of semantic contents derived from diverse database schemata, documentation structures, and resources across various museum disciplines, archives, and, more recently, libraries.

It has been recognized as an ISO standard [ISO211272014] designed to provide (a) high level information retrieval, as well as (b) the formulation and documentation of very specific data points and questions. While this latest standard dates to 2014, the latest stable version of its specification is the 7.1.1 [Velios2021] dating to 2021, along with an under development ISO standard referring to the latest official version, as of the time of writing this document.

The CIDOC CRM achieves its goals (a and b mentioned above) by providing definitions and a formal structure for describing the implicit and explicit concepts and relationships used in CH documentation and for the querying and exploration of such data, as formal ontologies. These formal descriptions are intended to to promote a shared understanding of CH information allowing for a common and extensible framework for the integration of CH data from multiple sources in a software and schema agnostic fashion. This is by serving as common language to be used by domain experts and conceptual modelers to formulate requirements for their information systems, and thus serving as a guide for their conceptual modeling tasks. In this way, it can capture specific CH domain domains and provide the intermediate ontological layer needed to mediate between different sources of CH information.

The ontology provides a comprehensive and formalized framework for describing and representing CH objects, events, actors, and their relationships, including their temporal, spatial and event-based aspects, making it suitable for complex and rich CH data.

While CIDOC CRM provides the basic classes and relations that represent the various CH disciplines, it is also extended by eleven modular models which cover documentation requirements of specific disciplines of the CH domain (FRBRoo, PRESSoo, CRMinf, CRMarchaeo, CRMsci, CRMgeo, CRMdig, CRMba,CRMtex, CRMsoc, and CRMact). The different versions of the both the CIDOC CRM and the CRM official extensions mentioned are available online [3].

EDM. Europeana is the European Commission's digital platform for CH maintained through a digital platform for cultural heritage collections called the Europeana portal (europeana pro). The main aim of the project is to collect *metadata* about CH entities from european CH institutions, and to enable the search and discovery of these items through a unified representation within a data model.

Due to the considerable digitization efforts undertaken by numerous institutions and data providers across Europe, a diverse array of CH objects have been made accessible in digital

1. CIDOC referring to the "Comité International pour la Documentation" is one of the committees that form part of ICOM - the International Council for Museums. The *Documentation Standards working group* - formed originally from the fusion of the data modeling and terminology working groups - took the decision in 1996 to embark on the development of a detailed conceptual model of the domain of cultural heritage information, known as the Conceptual Reference Model (CRM).

2. CRM was intended initially to extend and finally to replace the existing CIDOC relational data model [Reed1995], with the initial scope being restricted to that of the International Guidelines for Museum Object Information : The CIDOC Information Categories, published in June 1995 [Grant1995].

3. `https://CIDOCCRM.org/versions-of-the-CIDOCCRM`

form. However, the challenge lies in unifying this data, as each institution and provider adheres to different metadata standards.

Consequently, there is a pressing need to present this data in a cohesive manner within a cross-disciplinary framework, and thus the Europeana project emerged. It started in [Aloia2011] and proposed the development of the Europeana Data Model (EDM) which encompasses a set of classes of properties necessary to describe the CH objects in Europeana [1].The metadata aggregation is based on the mappings between the institutions' data and the EDM.

Nowadays, Europana stands as the principal European Digital Library, while EDM serves as the conceptual model for mapping metadata from various repositories into Europeana [Peroni2013]. The documentation of EDM comprises a comprehensive set of resources : The documentation of EDM comprises a comprehensive set of documents and online resources : (a) the EDM Definition providing the formal specification of the classes and properties [2], (b) the EDM Primer [Isaac2013] (the companion of the Definition document) providing the story of EDM and explaining the manner of usage of the classes and properties together to models data, (c) the EDM mapping guidelines [3] as a guide for providers on how to map their data to EDM, (d) the EDM profiles and mappings [4] gathering works works by Europeana and partners which yield in adapting EDM to the needs of the specific domains and applications and making domain systems interoperable, (e) the EDM Object Templates [5] providing templates for data providers to show the application of properties with respect to class and data type choices, (f) the EDM roadmap for an overview of the plan of development of the EDM, and (g) additional resources such as the EDM XML Schema [6] and the EDM validation guideline [7].

The EDM model serves as a metadata aggregator, combining multiple metadata schemas into a unified framework to enrich metadata and enhance granularity. The specifications of EDM emphasize the inclusion of vocabularies, models, and ontologies within its data model, enabling the representation and access of metadata pertaining to cultural heritage objects. Thus, operating within the context of data aggregation, EDM accommodates complex objects and acknowledges the possibility of diverse perspectives from various data providers.

LIDO. The Lightweight Information Describing Objects (LIDO) has emerged as a collaborative effort among museum stakeholders, aiming to establish a unified solution for contributing their CH content to portals and other repositories of aggregated resources [Coburn2010], such as the previously mentioned Europeana portal.

In the context of extracting valuable metadata from collection management environments, which hold immense potential for offering meaningful insights to users, the ATHENA pro-

1. EDM is in fact an improvement of the Europeana Semantic Elements (ESE) which aimed to express the datasets of european CH institutions using the Dublin Core (DC) standard as the lowest common denominator for object metadata, forcing interoperability, converting datasets to a flat data representation yielding in the loss of the semantics of the original data [Doerr2010].
2. The latest standardized version of the EDM Definition is downloadable on the permanent link https://pro.europeana.eu/page/edm-documentation.
3. https://europeana.atlassian.net/wiki/spaces/EF/pages/987791389/EDM+-+Mapping+guidelines
4. https://pro.europeana.eu/page/edm-profiles
5. https://github.com/europeana/corelib/wiki/EDMObjectTemplatesProviders
6. https://www.europeana.eu/schemas/edm/EDM.xsd
7. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM%20Schematron%20validation%20in%20Oxygen%20XML%20editor.pdf

ject undertook the task of determining the most suitable metadata schema to employ. And while the Dublin Core (DC) schema does not fully encompass the perspective of museum contents, such as differentiating between the creator and finder of a museum entity, resulting in the loss of significant data and relationships, the initiative for developing an alternative solution was handled by the ATHENA project by developing LIDO schema. It serves the purpose of assisting museums in effectively providing their object data to Europeana. By adopting the LIDO standard, museums can streamline the aggregation, transformation, and delivery of their data to Europeana and other CH repositories, simplifying the process and enhancing accessibility to their valuable collections.

The LIDO schema is conveniently available in XML format [1], and it incorporates various standards and existing formats to ensure compatibility and interoperability. These are the *CDWA Lite* which is an XML schema for encoding core records for arts and material works [ARTstor2006], the *museumdat* which is an XML schema which is a reconfiguration of CDWA Lite to account an event-oriented multi-disciplinary approach of the CIDOC CRM, the SPECTRUM XML schema which provides formats for exchanging object records between different collections management systems and aggregating data, and the CIDOC CRM ISO 21127 for providing a formal structure for describing the implicit and explicit concepts and relationships used in CH documentations.

In terms of functionality, LIDO offers a straightforward and flexible schema that enables the representation of essential descriptive information about CH objects. This includes object identification, classification, relation, and other pertinent elements. LIDO is commonly employed as a means of providing metadata for museum collections and exhibitions, contributing to effective documentation and information dissemination in the CH domain.

ABC. The ABC model was conceptualized as part of the Harmony international digital library project, which aims to investigate methods and models for effectively describing the diverse range of rich content that is becoming increasingly prevalent on the Web and digital libraries. While the earlier version of the model was enhanced through collaborations with various entities, such as the Dublin Core Metadata Initiative and the IFLA Functional Requirements for Bibliographic Records [Plassard1998], the current updated version [Lagoze2002] has further benefited from collaborations specifically focused on the museum community. These latter involve the CIDOC CRM and the CIMI consortium, contributing valuable insights and expertise to the model's development.

The primary objective of the ABC metadata model and ontology is to establish a conceptual foundation for comprehending and analyzing existing metadata vocabularies and instances. Additionally, it strives to offer guidance to communities embarking on the examination and development of descriptive vocabularies, while also serving as a basis for facilitating automated mapping among different metadata vocabularies.

In this context, the model encompasses a range of fundamental entities and relationships that are commonly found in various metadata vocabularies. These elements serve as the building blocks for developing domain-specific, role-specific, or community-specific vocabularies. In particular, the primitive category of the ABC ontology is an *entity* representing the description of a certain object/record, in addition to three other main categories covering different aspects around the entity : *temporality category*, *actuality category*, and the

1. `https://www.lido-schema.org/schema/v1.0/lido-v1.0.xsd`

abstraction category. The temporality aspect in ABC, represented by the *Temporality* class is based on theoretical foundations such as Situational Calculus [McCarthy1959]. It allows to express the states in which object properties exist using the *State* class, the transitions that define those states using the *Event* class, and the actions and agencies that participate in these events using the *Action* class. The actuality category, represented by the *Actuality* class, encompasses sensible entities -can be heard, seen, smelled, or touched, and is used to describe the *entity* using the *inState* and *hasInstance* properties to express time-dependent and time-independent facets of the same entity. And the abstraction category, represented by the *Abstraction* class, makes it possible to express insensible entities that never have a *State* and cannot exists by themselves, but are rather linked to some Actuality using *hasRealization* property such as the *Work* class.

As such, the ABC model provides an abstract, syntax-neutral conceptual framework for modeling metadata. It is available as an RDF schema using the RDF/XML syntax along with a search interface which is capable of more sophisticated queries than less-expressive, object-centric metadata models will allow.

Furthermore, the authors of [Doerr2003b] present a mediation process that establishes a harmonization between the CIDOC CRM ontology and the ABC metadata model. This harmonization serves as a valuable foundation for integrating knowledge from both the CH and museum community (target users of CIDOC CRM) and the digital library community (target users of ABC). It facilitates the seamless integration of these two domains, enabling effective collaboration and knowledge exchange between the CH and museum community and the digital library community. This was followed by proposing ABC as a core ontology to provide a common understanding of the basic enetities and relationships for achieving semantic interoperability and to enable information integration from diverse sources in multimedia [Hunter2003].

FinnONTO. The Finnish National Semantic Web Ontology project, known as FinnONTO, is an ambitious undertaking in Finland aimed at developing a semantic web infrastructure on a national level [Hyvönen2006a, Hyvönen2008, Hyvönen2010]. The project has yielded a diverse range of scientific outcomes, specifications, services, demonstrations, and applications demonstrating its dedication to advancing the field of semantic web ontology and its practical applications.

First is establishing national metadata standards across various application domains e.g. [Suominen2007]. Second, is going beyond traditional thesauri and developing core ontologies that can be used by both machines and humans [1]. In this sense, a number of core ontologies have been developed e.g. the cultural research ontology (KULO), the ontology for museum domain (MAO), in addition to the Finnish General Upper Ontology (YSO) which is based on the widely used Finnish General thesauri. Third is providing public ontology services for using and integrating ontologies with a client application. this is done through an ontology library and web service framework, called ONKI to enable ontology use in ontology, content indexing, and information retrieval [Viljanen2008]. Fourth is extending the project to pilot applications by implementing semantic web portals in different field such as CultureSampo [Hyvönen2006b, Hyvönen2009a] and MuseumFinland [Hyvönen2005] for eCulture, HealthFinland [Hyvönen2007] for eHealth, etc.

Thus, the CultureSampo portal represents an application of the FinnONTO infrastructure of ontologies within the CH field in Finland. It incorporates the FinnONTO ontologies in-

1. https://seco.cs.aalto.fi/ontologies/

frastructure, a range of metadata schemas and tools, including ONKI, SAHA, POKA, and VERA, and a user-interface, all integrated into a publication system based on semantic web technologies.

As such,the FinnONTO infrastructure's ontologies play a crucial role in the CultureSampo system, complementing generic semantic web recommendations like RDF with domain-specific concept descriptions in various domains. Many of the core ontologies were developed by transforming thesauri into lightweight ontologies, using a combination of automatic and manual approaches. Automatic methods, as described in [van Assem2004], were employed in some cases, while manual refinement of relations in subsumption hierarchies required human intervention [Hyvönen2009b].

The core ontologies were aligned with YSO using equivalence and subclass relationships, resulting in a system of interconnected and harmonized ontologies known as KOKO shown in figure . The alignment process was carried out using Protege, a popular ontology editor, ensuring consistency and compatibility between the ontologies.

The integration of the core ontologies within the KOKO system, along with the use of YSO as a foundational ontology, fosters semantic coherence and facilitates the exchange of information across different domains within the CultureSampo platform.

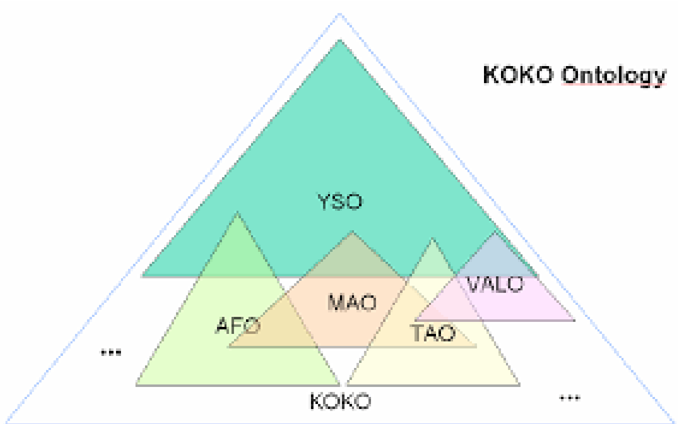


FIGURE 2.3 – *Figure taken from [Hyvönen2009b] showing the KOKO system of mutually mapped cross-domain ontologies.*

By leveraging these ontologies, the portal addresses the challenges of interoperability posed by multiple ontologies from different domains and the integration of diverse metadata schemas and cross-domain content into a unified semantic portal [Mäkelä2012].

One of the notable contributions of the CultureSampo system is its novel approach to facilitating collaboration and distributed ontology and content development among various memory organizations and citizens. It proposes and demonstrates a content creation process that promotes active participation and contributions from multiple stakeholders.

The system is designed to be multilingual, supporting Finnish, Swedish, and English languages, and offers CH content to end-users through nine thematic perspectives [Hyvönen2009b]. These perspectives include map views, relational search, search and organize functionalities, collection views, and Finnish history views, among others. Figure 2.4 provides a snapshot of the "Finnish History" thematic perspective, showcasing a timeline with corresponding event resources spanning the years 1853 to 1895 in Finland.

Inspire. The Infrastructure for Spatial Information in Europe (INSPIRE) is a framework initiated

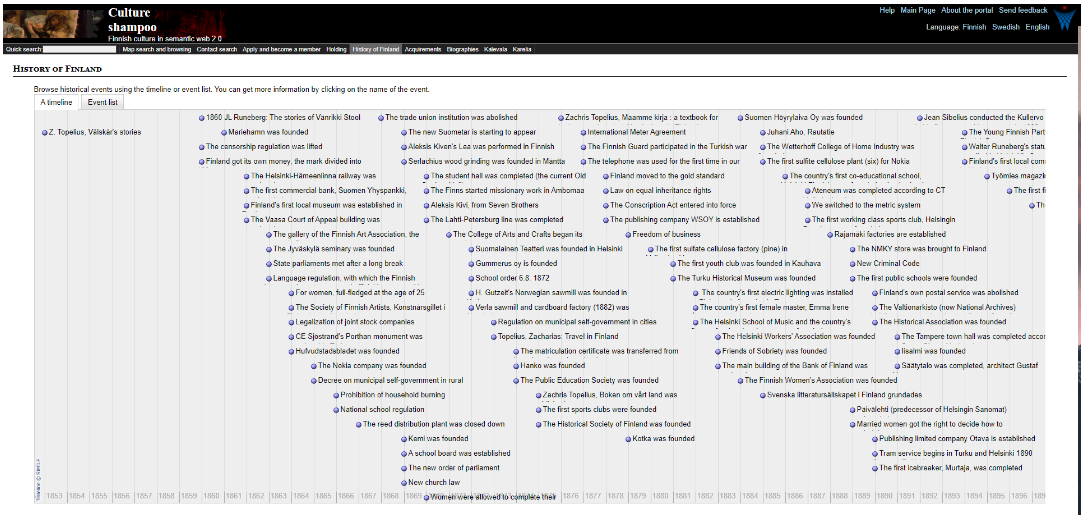


FIGURE 2.4 – *A snapshot of the "Finnish History" thematic perspective of the CultureSampo portal, showcasing a timeline with corresponding event resources spanning the years 1853 to 1895 in Finland.*

by the European Union to enhance the accessibility and interoperability of spatial/geographical information [Parcero-Oubiña2013]. Integrating CH into the INSPIRE framework is crucial for incorporating heritage information into the broader domain of spatial data. The spatial component of heritage elements plays a significant role in their characterization and subjective value as cultural assets.

While CH is included in INSPIRE to some extent, its treatment within the framework is not extensively developed. Therefore, there was a need to further develop a concrete data model for CH within INSPIRE to promote the development of CH Spatial Data Infrastructures (SDIs) within an interoperable framework. This would acknowledge the spatial nature of CH data and enhance their role in territorial governance, protection management, and public accessibility.

INSPIRE primarily focuses on natural protected sites, and only a specific subset of CH related to geographically protected locations is included. To address this limitation, a specific schema for protected CH sites was developed to incorporate this data into INSPIRE. This schema complements the description and spatial characterization of protected sites by linking them to specific themes listed in Annex III of INSPIRE [INSPIRE2013].

The integration of CH into INSPIRE aims to make a diverse range of geographic information accessible to the public, manage the protection of CH, and bring it closer to the general public. This necessitates studying the spatial components of CH, including protected natural sites and other elements.

The abstract data model created for integrating cultural heritage into INSPIRE consists of three main parts : *legal*, *cultural*, and *documentary*. The cultural entity encompasses various types of non-material and material entities, such as *HumanMadeObjects*, *HumanMadeFeatures*, and *NaturalFeatures*, as long as they have a spatial reference. The model allows for extensions based on the specific implementation requirements.

However, the INSPIRE model does not define specific kinds of objects or features ; instead, it focuses on the specificities of natural protected sites. This delimits the domain that an INSPIRE-derived data model must consider, but it provides a foundation for incorporating cultural heritage into the framework.

CHARM. The Cultural Heritage Abstract Reference Model (Charm) [Gonzalez-Perez2018] is an ontology that encompasses various aspects of cultural heritage, including tangible entities, agents, performative entities, valorizations, representations, locations, and occurrences.

The model's purpose is to represent entities that hold value within cultural heritage, along with their associated valorizations and representations. This includes entities that receive cultural heritage value as well as those necessary for their description and understanding. Thus the term "Cultural Heritage". As for the term "Reference Model", it signifies that Charm is designed to be used by diverse organizations and individuals to establish a common understanding. The "Abstract" characteristic implies that it provides an independent and extendable view that can accommodate additional specificities based on the requirements or perspectives of each organization or individual.
Charm serves as an infrastructure to assist domain experts, such as those in the cultural heritage community, in expressing their own conceptualizations of cultural heritage phenomena. It can be utilized for exploring, documenting, and communicating various aspects of archaeological and anthropological entities, among other applications.

Starting from the basis that valuable entities and valorizations are different, Charm is constructed around three basic pillars : *valuable entities* resembling entities that have received, currently receive, or may receive CH value, *valorizations* representing values that are granted to entities, and *representations* which are accounts or portrayals of other things, including both preceding pillars. Delving into *valuable entities*, figure 2.5 shows the two kinds *PrimaryEntity* and *DerivedEntity* distinguishing between those entities that do not need an explicit interpretive process to associate value to them (i.e. primary) e.g. a sculpture in a museum can be classified by anyone as a cultural heritage entity, versus those that require external explicit explanation of their value (i.e. derived) such as an archaeological site that required an achaeologist to classify it as such.

Charm is structured on the fundamental understanding that valuable entities and valorizations are distinct concepts. It encompasses three core elements : *valuable entities*, which encompass entities that have, currently possess, or may acquire cultural heritage (CH) value ; *valorizations*, which represent the values attributed to these entities ; and *representations*, which encompass accounts or depictions of other entities, encompassing both the preceding pillars.
Within the realm of *valuable entities*, Figure 2.5 provides an overview of the two categories : *PrimaryEntity* and *DerivedEntity*. The distinction lies in the fact that primary entities do not necessitate an explicit interpretive process to ascribe value to them, for example, a sculpture housed in a museum can be classified as suchy by anyone. Conversely, derived entities rely on external explicit explanations of their value, such as an archaeological site requiring an archaeologist to classify it as such.

GVP. The Getty Vocabulary Program (GVP) Ontology [1] includes various classes, properties and individuals (values) used in the GVP LOD [Alexiev2015]. It is designed to provide a standardized framework for representing the concepts, terms, and relationships found within the Getty vocabularies [2], which include AAT, ULAN, TGN, and CONA (all of which are mentioned earlier). The GVP ontology serves as a semantic model for expressing the structure

1. http://vocab.getty.edu/ontology
2. http://vocab.getty.edu/

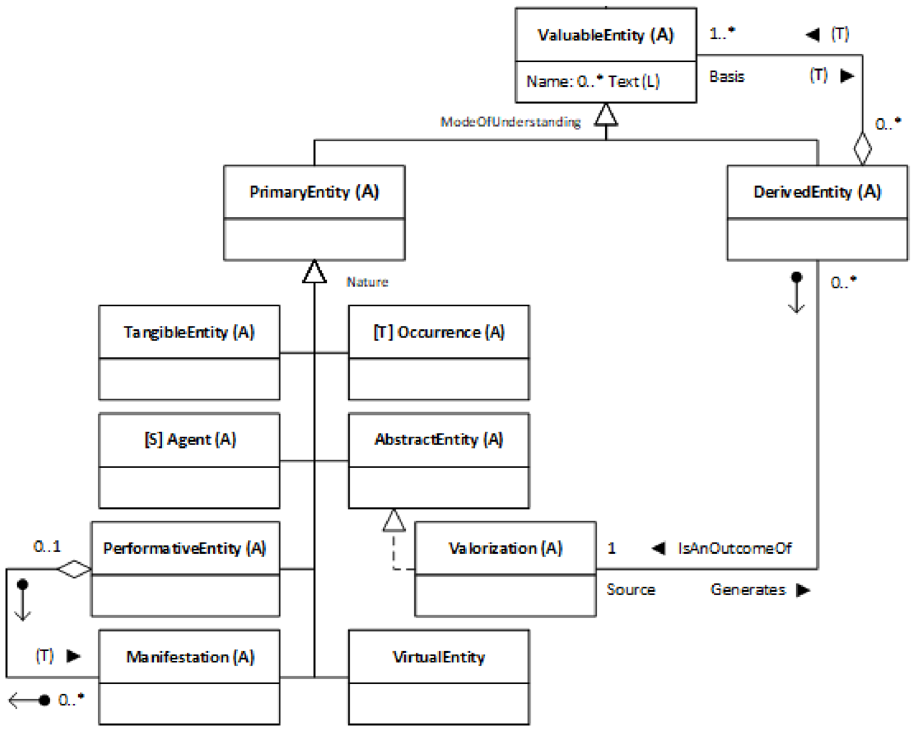


FIGURE 2.5 – *CHARM's topmost view showing the most abstract concepts.*

and semantics of the concepts and relationships within these vocabularies. It is complete regarding the AAT, TGN and ULAN vocabularies and will be extended in time with more elements needed for other GVP vocabularies (CONA). It also uses popular namespaces of extant ontologies such as SKOS, SKOS-XL, ISO 25964, DC, DCT, BIBO, FOAF and PROV.

The GVP ontology enables users to explore and navigate the rich semantic relationships within the Getty vocabularies, facilitating enhanced search, discovery, and analysis of CH resources. It supports the description and classification of artistic works, artists, geographic locations, and cultural objects, among other aspects of CH information.

OAIS. The Open Archival Information System (OAIS) [1] [BOOK2002] model, developed by CCSDS and ISO, focuses on digital information preservation, encompassing primary forms of information and supporting data for both digital and physical archives. The OAIS model facilitates consensus on archival requirements across disciplines and plays a significant role in the ISO 16363 standard [2], particularly concerning metrics related to Digital Object Management that rely heavily on OAIS concepts.

LRM. The Library Reference Model (LRM) is an conceptual model that serves as a reference model for library-related information, encompassing bibliographic records, authorities, and cataloging rules [Riva2018]. Although primarily used within library settings, LRM is used to represent relevant information about resources resembling CH entities such as books, manuscripts, cultural place, and archival materials.

1. http://www.oais.info/
2. http://www.iso16363.org/

The landscape of ontology models in the CH field is extensive, encompassing numerous models that serve various purposes. Each of these ontologies has its own strengths and weaknesses, and their suitability for modeling CH entities depends on the specific use case, domain, and requirements of the application. Thus, within an application, it is important to carefully evaluate and choose the appropriate ontology based on the specific modeling scope and requirements of your cultural heritage data project.

2.3.3 A synthesis of the preceding systems

All of the aforementioned examples of KOS, which exhibit varying levels of structural complexity and functionality, serve the purpose of organizing and enhancing access to knowledge concerning CH data within digital libraries or any other institution. Figure 2.6 illustrates the collective representation of KOS approaches in the field of CH, organized according to the classification presented above.

In general, KOS capture the underlying semantic structure of a specific domain and offer semantics, navigation, and translation through labels, definitions, typing, relationships, and properties associated with concepts [Hill2004, Koch2004]. Numerous systems of different structures have been employed to tackle data heterogeneity, and enhance interoperability [Tudhope2004a] as presented in Section 2.2. These systems, often embodied as (Web) services, play a crucial role in facilitating resource discovery and retrieval. Acting as semantic roadmaps, they provide a common reference point for indexers and future users, both human and machine, enabling effective navigation and exploration of digital resources [Tudhope2004b].

In particular, ontologies offer robust organization systems with advanced functionalities and a multidimensional structure. They go beyond modeling concepts and relations by also incorporating properties of concepts and additional rules. As a result, the use of ontologies in the CH domain has experienced significant growth [Moraitou2022].

Initially, ontologies emerged as a solution to address interoperability challenges associated with diverse and fragmented cultural data. By providing a unified framework for collecting, managing, and exchanging data across different CH institutions, ontologies have played a crucial role in achieving data harmonization and integration [Moraitou2019]. Their adoption within the CH community has been particularly prominent to represent the complex and multidimensional nature of cultural artifacts, practices, and contexts. They provide a structured way to not only describe and organize CH data, but also transform data into semantics enabling semantic interoperability, data integration, and knowledge discovery.

Thus, in the upcoming sections of this chapter (Section 2.4), our focus will primarily be on ontology KOS structures within the CH domain. Having presented the prominent ontology modeling approaches used in the CH field in the previous section, we will proceed to systematically classify the existing approaches based on specific criteria. This systematic classification will be followed by a detailed review of the most relevant approaches.

2.4 Modeling CH data (entities) using Ontologies

In this section, we first carry out in Section 2.4.1 a systematic classification of the extant ontology approaches that we have illustrated earlier based on specific criteria. Through this process,

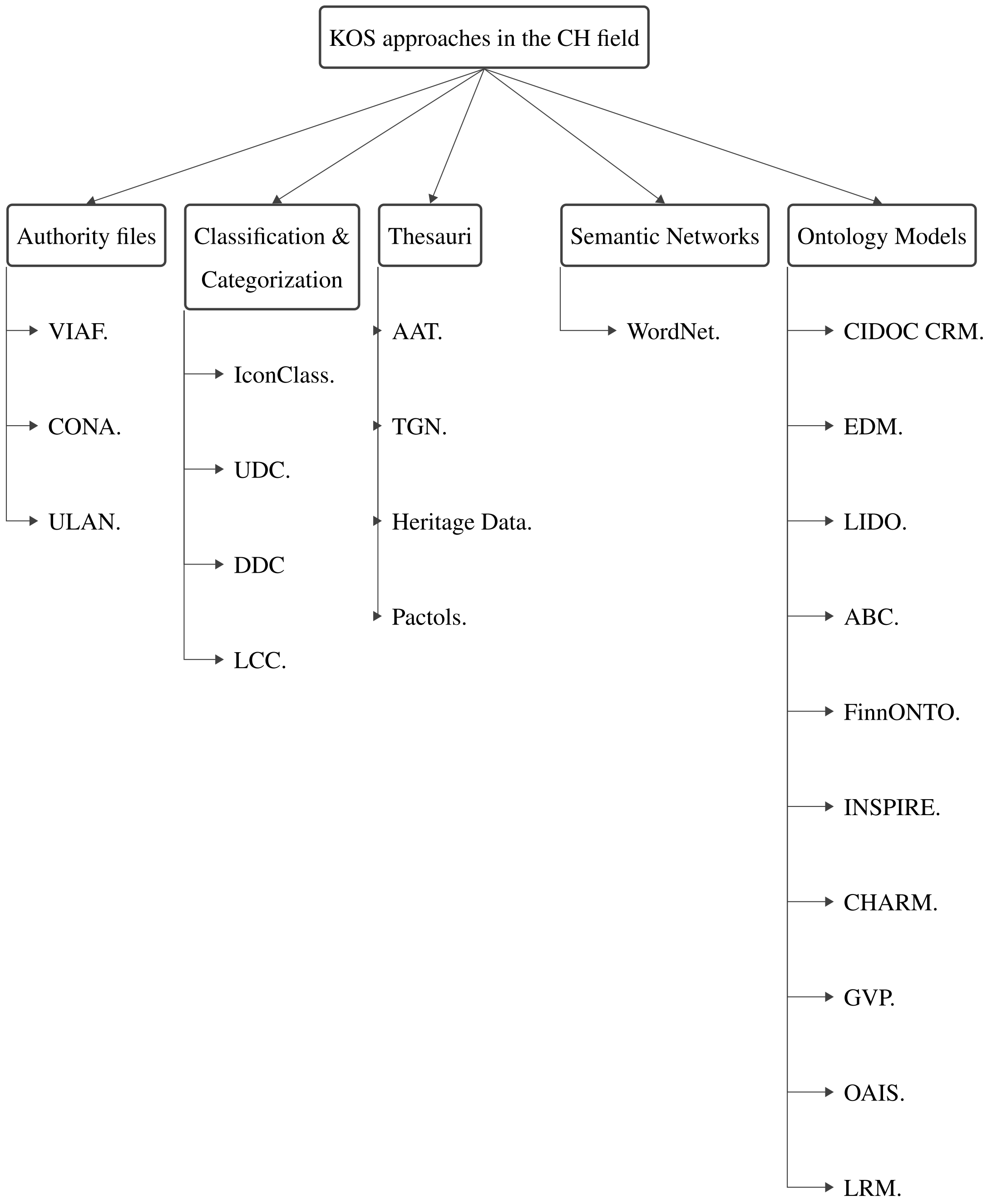


FIGURE 2.6 – *Some examples of KOS systems that have been proposed in the CH field classified according to their structure into five KOS types ; authority files, classification and categorization schemes, thesauri, semantic networks, and ontology models.*

we will carefully evaluate and select the most relevant approaches that align with our fundamental objective and [motivation-I] in particular.

Following the classification, we conduct in Section 2.4.2 a comprehensive examination and analysis of two selected relevant ontology models, considering the objectives outlined in our thesis. This analysis will shed the light on the extent to which these chosen approaches do not cover our intended representation a CH entity, but also examine how they can contribute to our research goals in the future.

2.4.1 A systematic classification of extant approaches

In the previous sections, we have observed a wide range of KOS approaches, ontology modeling approaches in particular, that have been proposed in the literature to address the challenge of integrating diverse and heterogeneous CH content.

These approaches exhibit significant variations in terms of several factors. For instance, some operate at the European international level, such as Europeana, while others are developed at the national level, like FinnONTO. Additionally, the approaches employ different levels of formalization, ranging from complex theoretical formalizations like CIDOC CRM with a first-order logic (FOL) formalization, to simpler empirical formalizations like CHARM in ConML. As such, each ontology modeling approach has been designed with a specific purpose and scope, tailored to the requirements of its respective project or intended application.

When there is a need to model CH data/entities within an institution, experts or ontologists will evaluate existing ontology approaches and choose a model that aligns with their specific needs. In cases where no suitable model is available, they may decide to develop their own relevant model.

In our case, we aim to identify the most relevant approaches that can effectively represent the composition of a CH entity, particularly focusing on its materiality. To achieve this, we establish three criteria, each with various potential options, inspired by our literature review. We then classify the approaches into a unique category for each criterion. This systematic classification enables us to select the relevant approaches that align with our modeling objectives and meet our chosen criteria.

The purpose of this classification is to identify and extract the approaches that are most pertinent to our objective. By doing so, we can conduct a detailed review of these approaches to assess their suitability for modeling our entities.

The classification of extant approaches :

For each criterion, we provide a concise description and explain its significance in evaluating and categorizing the modeling approaches. Then, we present the options (aka categories) of each while discussing how approaches align to each category. It's important to note that the categories of each criterion are not mutually exclusive and can overlap i.e. different ontological models may have characteristics that fall into multiple categories.

⋙ **The Geographical scale :**
This criterion considers the spatial boundaries or scopes within which CH data is considered for integration. It emphasizes the extent to which the models are designed to capture and represent CH information at different geographical scales, ranging from local to national and international contexts. This criterion recognizes that CH is often inherently tied to specific geographic regions, communities, or nations. It acknowledges that modeling approaches occasionally account for the diverse characteristics, contexts, and needs of CH data based on its geographical origin or relevance. And according to this geographical scale, modeling approaches select the CH entities whose data are to be integrated within their models.

The criterion of geographical scale offers two distinct option (categories) which we explain below while classifying the ontology approaches into each according to Table 2.2 :

— National-scale modeling approaches encompass modeling efforts that span an entire country or nation. These approaches aim to capture the diverse CH elements of a specific country, considering its unique heritage assets, history, and cultural practices, involving collaboration among CH institutions at the national level. In this category falls the FinnONTO approach developed at the national Finnish scale.

— International-scale modeling approaches are designed to facilitate interoperability and information exchange across multiple countries or regions. These models have a broader scope, typically covering transnational or pan-European CH initiatives. Thus, enabling the integration of diverse cultural heritage data from different countries or regions, fostering collaboration, harmonization, and cross-cultural research. For example, Inspire and Europeana aim to promote interoperability and access to CH data at the European level. Also, CIDOC CRM is classified as international-scale approach for that it provides a global framework for representing CH information. The ABC ontology is part of the Harmony project which is international and targets CH content across all the web and digital libraries. Similarly, LIDO addresses the conceptualizing of metadata at a wide range with focusing on delivering metadata schemes to Europeana i.e. a European scale. In addition, both OAIS and LRM account for international standards for modeling archives and and bibliographic records. As for CHARM, the ontology does not limit its application to single scale, instead it conceptualizes entities and relationships at a CH general level. Thus, we classify it as international, rather than limited to a certain geographical scale.

Geographical Scale	National	International
Ontology Model	FinnONTO	CIDOC CRM, Europeana, LIDO, ABC, Inspire, CHARM, GVP, OAIS, LRM

TABLEAU 2.2 – *Classifying some extant ontology approaches into the two geographical scales ; national and international.*

⋙ The Semantics and Formality level :

This criterion refers to the degree of emphasis on the semantic representation and the level of formality in the models used to capture CH information. It emphasizes the nature and depth of the semantic modeling employed in representing CH entities, relationships, and concepts, as well as the level of formalization and expressivity. This criterion shows how and to which extent a modeling approach captures the CH domain. It gives insights on its intended role in facilitating data interoperability, knowledge sharing, and the intended understanding of CH information.

We categorize this criterion into three distinct options (categories) for classifying the modeling approaches according to Table 2.3 [1] :

— Metadata-based models : This category includes modeling approaches that primarily focus on the modeling of metadata, such as their identification, classification, and contextual details. These models provide structured and highly interoperable frameworks for capturing and organizing the extant information related to CH objects, enabling efficient retrieval and management of CH data. Out of the previously illustrated ontology models, examples of metadata-based ones are EDM as it aims for mapping metadata from various CH repositories into a structured model (organizing data about data), LIDO whose goal is to provide a unified reference model for these repositories to organize their data according to it and map to EDM accordingly, and the ABC model which addresses the description of the heterogeneous CH data distributed in various repsitoes into a common conceptual foundation.

1. Although we refer to the first two categories as "metadata-based" and "thesauri-based," it is important to note that these categories resembles ontologies and not metadata models or thesauri. The terms used are intended to highlight the specific focus or inspiration behind these ontologies where the former emphasizes ontologies that model data about data, and the latter signifies ontologies that draw inspiration from thesauri.

— Thesauri-based models : This category encompasses models that are derived from or inspired by existing thesauri, which serve as controlled vocabularies for organizing and retrieving information. These models use thesauri concepts and relationships to represent and link CH entities and concepts. They enable semantic enrichment and provide a high expressivity of domain-specific terms. The GVP ontology is a typical example of a thesauri-based model whose classes, properties, and individuals are those present in the Getty Vocabulary Program of LOD. In addition, the core ontologies of the FinnONTO infrastructure are developed based on traditional thesauri, maintaining the concepts and relations expressed within Finnish vocabularies, with the goal of replacing these thesauri. Even YSO, the most central and top-level ontology in the FinnONTO infrastructure, is based on the general Finnish keyword thesaurus YSA maintained by the National Library of Finland.

— Formal ontologies/Conceptual models : This category involves the approaches that concentrate on modeling CH as a domain itself, rather than organizing data in the domain or structuring the vocabularies that used in the domain. These models aim at offering a shared understanding, a conceptualization, yielding in a new semantic level level with high formality and expressivity in representing CH objects. They often adhere to established standards and ontological principles, providing a well-defined and standardized framework for representing complex domain knowledge. They allow for precise and structured representation of entities, relationships, and concepts, enabling advanced reasoning and inference capabilities. In this category falls the CIDOC CRM, CHARM, and Inspire which provide different formal representations of CH entities based on a formal understanding. Although EDM is considered a metadata-based model, however within the EDM specification exists elements that serve as consensus of the different data representations across repositories. These elements are EDM classes and properties, some of which map to CIDOC CRM elements, providing top-level representation under which the different resources classify. And last, both the OAIS model and LRM classify as reference models offering frameworks for the understanding of concepts and the description of (CH) entities in archival and library systems, respectively.

Semantics & Formality Level	Metadata-based models	Thesauri-based models	Formal ontologies/ Conceptual models
Ontology Model	EDM, LIDO, ABC	FinnONTO, GVP	CIDOC CRM, CHARM, Inspire, EDM, OAIS, LRM

TABLEAU 2.3 – *Classifying some extant ontology approaches into the three semantics and formality levels ; metadata-based models, thesauri-based models, and formal ontologies or conceptual models.*

⋙ The Modeling scope :

The criterion refers to the focus and specific domain coverage of the ontologies, i.e. the extent and boundaries of the subject matter, within the CH domain, that the ontology aims to represent and capture. The significance of this criterion lies in assessing the applicability and suitability of the modeling approaches to different aspects within the CH field. It allows for the identification of models that are designed to capture specific types of data, spatial information, or comprehensive representations of CH entities. By considering the modeling scope, we can better understand the intended use and strengths of the ontologies in representing and integrating CH information.

Each ontology model possesses its distinct modeling scope. However, within the group of ontology models under consideration in this Section, we have identified three distinct categories of modeling scopes. Accordingly, the modeling approaches can be classified as follows and as shown in Table 2.4 :

— Data-centric Modeling : This category encompasses ontology models that primarily focus on modeling bibliographic records, library data, archival systems, and information objects. These models are designed to capture and represent data-related aspects of CH, such as descriptive metadata, classification schemes, and information organization. In this category falls the ABC ontology, LRM, LIDO, and the OAIS.

— Spatial-centric Modeling : This category includes models that specifically focus on modeling the spatial and geographic properties of CH entities. They are tailored to capture and represent information related to natural and archaeological sites, geographic features, and their spatial relationships. Examples within this category include Inspire which resembles the infrastructure of spatial information in Europe, whose application to the CH field considers natural protected sites as spatial entities of CH value. In addition, the CRMarchaeo which i a CIDOC CRM extension specifically designed for archaeological data modeling.

— Entity-centric Modeling : This category covers models that aim to represent CH entities in general, including all types of entities such as physical objects and spatial (archaeological) sites. These models provide comprehensive representations of CH entities, capturing their attributes, relationships, and contextual information, and thus the CH domain as an overall. The CIDOC CRM, EDM, FinnONTO and CHARM models are well-suited for this category, as they primarily focus on modeling CH entities rather than CH data. These approaches do not limit themselves to a specific type of CH entity but encompass all types of CH entities in a comprehensive manner.

Modeling Scope	Data-centric models	Spatial-centric models	Entity-centric models
Ontology Model	LIDO, ABC, LRM, OAIS	Inspire, CRMarchaeo	CIDOC CRM, EDM, FinnONTO, CHARM

TABLEAU 2.4 – *Classifying some extant ontology approaches into three modeling scopes ; data-centric models, spatial-centric models, and entity-centric models.*

Other possible criteria : It should be noted that the classification criteria and their corresponding categories that we have established may not be universally applicable to all systematic classification systems or specific application needs. Indeed, other criteria could be considered which can be more significant in certain contexts. For instance, the criterion of interoperability and reusability of an ontology, which classifies models based on their adherence to interoperability standards and scopes, is of considerable importance. Also the temporal representation is another critical criterion for classifying ontologies, as it addresses the temporal constraints inherent in CH entities and forms a basic element for reasoning about them.

However, the chosen criteria have proven to be important in guiding our approach and contributing to our research endeavors. Another criterion of significance to us, although not explicitly discussed here, is the abstraction level of ontologies [1]. This criterion encompasses application, domain/task, core ontologies, mid-level, and top-level ontologies. While our primary focus lies on top-level ontologies as our approach is to be generic one, the classification of the selected ontology approaches according to this criterion did not significantly impact the final outcome of our system, given that the relevant models extracted are predominantly core or mid-level ontologies.

1. This term is presented in the Preliminary Remarks Section

Overall, these criteria have shaped our direction and align with our research goals, and they will further contribute to the development of our system.

A criteria-based selection of relevant approaches :

In this section, we first revisit our [fundamental objective] : "*Representing and modeling the composition of a tangible entity in general, and a Cultural Heritage tangible entity in particular, using ontological structural and spatial relations, within an Applied Ontology approach*".

We also recall the particular [motivation-I] behind this Chapter : "*Modeling the composition of a tangible entity as a complex structure (i.e. an object, a place, a collection) in a manner that enables the understanding, constructing, and navigating into its tangible discourse (i.e. representation), and learning its intangible aspects (i.e. significance)*".

Taking into account our interdisciplinary approach, we analyze how our intended approach aligns with each criterion and position it within the relevant category. This positioning helps us establish the requirements we seek in an ontology model that can effectively represent the composition of a tangible entity. By defining these requirements, we are able to identify a subset of the aforementioned ontology approaches that are *relevant* to our objective. In the following section (Section 2.4.2), we conduct a comprehensive review of these selected approaches, evaluating their suitability and effectiveness in relation to our objective.

In the context of geographical scale, our objective is to adopt a model that goes beyond specific geographical boundaries and encompasses CH entities irrespective of their location, cultural context, or nomenclature. This requires a model that focuses on the broader scope of CH rather than being confined to specific regions or countries. Among the ontology models considered in Table 2.5, CIDOC CRM, Europeana, LIDO, ABC, Inspire, CHARM, GVP, OAIS, and LRM fulfill this requirement by offering a perspective that is independent of any geographical boundary.

Geographical Scale	National	International
Ontology Model	FinnONTO	CIDOC CRM, Europeana, LIDO, ABC, Inspire, CHARM, GVP, OAIS, LRM

TABLEAU 2.5 – *Selecting relevant ontology approaches according to their geographical scale.*

Regarding the semantics and formality level, our aim is to utilize a formal ontology that surpasses the confines of thesauri-based models and instead builds upon a shared understanding of CH, in the sense presented in Section Preliminary Remarks. We emphasize the importance of semantic interoperability among various sub-fields involved in CH, enabling them to overcome differences in terminology and arrive to a common representation that is based upon the shared understanding. While both conceptual models and formal ontologies contribute to the conceptualization of a subject matter, we prioritize the use of formal ontologies due to their logical rigidity and their establishment as formal logical theories. In line with this, the selected ontology models from are those highlighted in Table 2.6, including CIDOC CRM, CHARM, Inspire, EDM, OAIS, and LRM, which align with our requirement of utilizing formal ontologies, but also include conceptual models in general.

And lastly, concerning the modeling scope criterion, our objective is to represent the composition and materiality of tangible CH entities. Thus, we seek an entity-centric model that places significant emphasis on the entities themselves, encompassing their structural and spatial aspects, rather than organizing data about CH entities. An ideal ontology model for our purpose would

Semantics & Formality Level	Metadata-based models	Thesauri-based models	Formal ontologies/ Conceptual models
Ontology Model	EDM, LIDO, ABC	FinnONTO, GVP	CIDOC CRM, CHARM, Inspire, EDM, OAIS, LRM

TABLEAU 2.6 – *Selecting relevant ontology approaches according to their semantics and formality levels.*

offer rich semantic relationships that allow for the representation and reasoning of entity composition. Based on this consideration and Table 2.7, we identify the models from the fourth row, CIDOC CRM, FinnONTO, EDM, and CHARM, as potential candidates to fulfill our intended representations.

Modeling Scope	Data-centric models	Spatial-centric models	Entity-centric models
Ontology Model	LIDO, ABC, LRM, OAIS	Inspire, CRMarchaeo	CIDOC CRM, EDM, FinnONTO, CHARM

TABLEAU 2.7 – *Selecting relevant ontology approaches according to their modeling scope.*

Consequently, based on the intersection of these three criteria-based requirements, the ontology models that best meet our objectives are CIDOC CRM, EDM, and CHARM. These models encompass a broad geographical scale, adhere to the formality of a formal ontology, and offer the necessary capabilities to represent and reason about the composition of tangible CH entities.

2.4.2 Representing CH entities as complex structures : an analysis

This section analyses the selected relevant ontology models -CIDOC CRM, EDM, and CHARM- in view of our [fundamental objective] and [motivation-I], focusing in particular on the models' potentials in expressing the composition of a tangible entity, using structural and spatial relations.

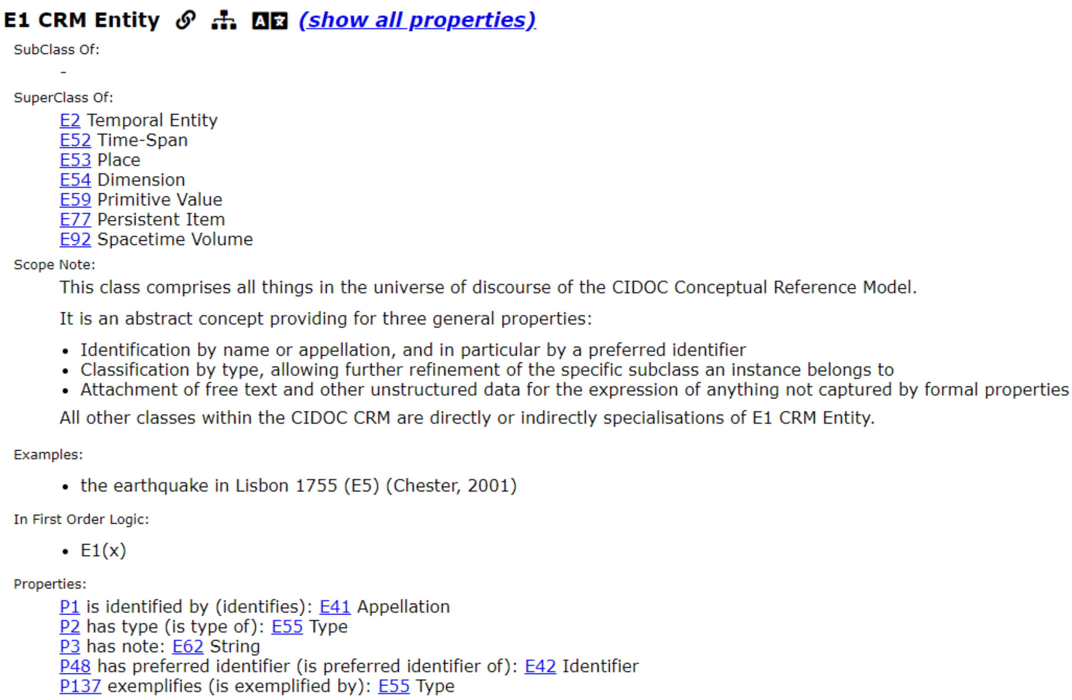
For each model, we (1) start by its logical formalization, if any, (2) go through its different serializations as an ontological model using (or not) semantic web technologies, (3) and examine its forte points and the application domains in which it has been employed. Finally, in view of our objective and the criteria-based elimination/inclusion process that we performed in Section 2.4.1, we (4) address its compositional (structural and spatial) relations, presented as object properties, which are at the core of our requirements.

The CIDOC Conceptual Reference Model (CRM) :

The CIDOC CRM is a formal ontology in the sense introduced by Guarino in [Guarino1998a] i.e. a specification of a set of named concepts used to describe and approximate a part of reality, plus a first-order logic theory to narrow down the intended meaning of the named concepts. As such, it is expressed in terms of FOL as a logic-based knowledge representation language by means of logical axioms. For each definition of a class/property, their corresponding FOL axioms are stated using unary/binary/ternary predicate symbols and basic symbolic FOL operators. Thus, focusing on semantic precision and expressiveness.

In addition to FOL, the ontology is expressed in an object-oriented semantic model understandable by experts and information scientists i.e. documentation in natural language format, as well as in machine-readable formats using RDFS, OWL, XML, and JSON LD, among others. This is by defining classes (resembling the unary predicates), properties (resembling the binary predicates), properties of properties (i.e. reification of properties resembling the ternary predicates), and axioms in a formal and machine-readable interoperable manner. The additional extension models of CIDOC CRM are provided in RDF format to support their use in the SW too [Cyganiak2014] such as that of CRMarchaeo [1].

Throughout the CIDOC CRM, classes are identified by numbers preceded by the letter *E* (historically classes were sometimes referred to as "Entities") e.g. the class *E63 Beginning of Existence*, while properties are identified by numbers preceded by the letter *P* e.g. the property *P126 employed*. The CIDOC CRM defines 81 classes and 160 unique properties. A tool is available online [2] providing an interface for users to navigate through the classes, properties declarations, and translations and versioning information, as well as visualizing the class/property hierarchy graph with all the corresponding outgoing and incoming properties of each class of the CIDOC CRM ontology. Figure 2.7 shows the user-interface of the web page after navigating through the *E1 CRM Entity* class, while figure 2.8 shows the same entity's hierarchy graph.



E1 CRM Entity *(show all properties)*

SubClass Of:

-

SuperClass Of:

E2 Temporal Entity
E52 Time-Span
E53 Place
E54 Dimension
E59 Primitive Value
E77 Persistent Item
E92 Spacetime Volume

Scope Note:

This class comprises all things in the universe of discourse of the CIDOC Conceptual Reference Model.

It is an abstract concept providing for three general properties:

- Identification by name or appellation, and in particular by a preferred identifier
- Classification by type, allowing further refinement of the specific subclass an instance belongs to
- Attachment of free text and other unstructured data for the expression of anything not captured by formal properties

All other classes within the CIDOC CRM are directly or indirectly specialisations of E1 CRM Entity.

Examples:

- the earthquake in Lisbon 1755 (E5) (Chester, 2001)

In First Order Logic:

- E1(x)

Properties:

P1 is identified by (identifies): E41 Appellation
P2 has type (is type of): E55 Type
P3 has note: E62 String
P48 has preferred identifier (is preferred identifier of): E42 Identifier
P137 exemplifies (is exemplified by): E55 Type

FIGURE 2.7 – *Figure showing the user interface of the online navigation tool after selecting the E1 CRM Entity class.*

CIDOC CRM was developed following rigorous principles that selectively incorporated concepts serving the purpose of global information integration [Doerr2003a]. These principles have resulted in its successful application as it remains compact without compromising its adequacy [Smith2004a].

Regarded as the most comprehensive ontology for integrating CH information, CIDOC CRM has gained increasing prominence in real-world integrated information environments for CH systems [Doerr2009]. Numerous use cases have demonstrated its effectiveness, and a compilation

1. https://cidoc-crm.org/crmarchaeo/fm_releases
2. https://cidoc-crm.org/html/cidoc_crm_v7.1.1.html

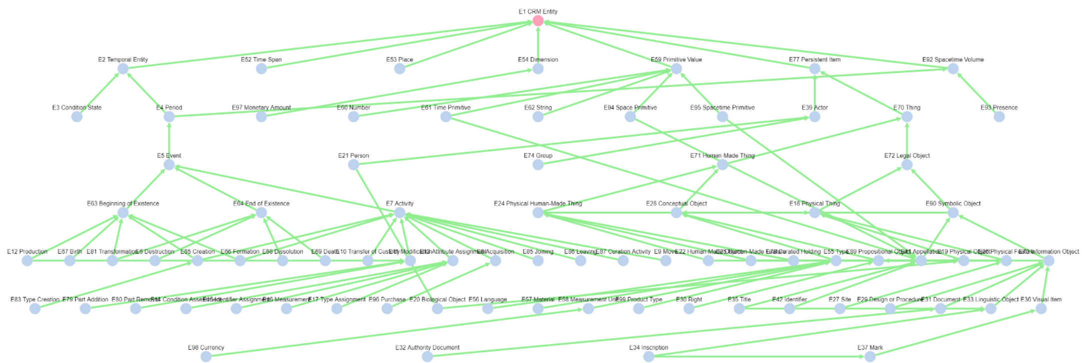


FIGURE 2.8 – *Figure showing the user interface of the online navigation tool after selecting the E1 CRM Entity class's hierarchy graph.*

of references highlighting its application can be found in the dedicated use-cases section of their site [1]. Moreover, the CRM's coverage of CH data and CH data structures has been extensively validated through mappings from various sector-specific data structures to the CRM, including the mapping from Dublin Core to the CRM [Doerr2000, Kakali2007]. This observation demonstrates the seamless integration of the well-defined common semantics of CIDOC CRM with prevalent metadata formats. As in terms of the implemented RDFS ontology, the mappings can be seen as a specification for Local as View (LaV) schema integration, as referred by Doerr in [Doerr2009].

And as discussed in Section 2.3.2, certain CH ontologies align with established traditional KOS, including CIDOC CRM. This ontology's data structures, i.e. the classes and properties, are compliant with renowned online thesauri such as the Getty AAT and TGN. In the following, we illustrate this synergy :

— an *E32 Authority Document* is a subclass of *E31 Document* and it comprises encyclopedia, thesauri, authority lists ad other documents that define terminology or conceptual systems for consistent use. Examples of instances of *E32* are Webster's Dictionary, the Getty AAT [Petersen1990], and the CIDOC CRM [Gergatsoulis2010].

— an *Information Object* which comprises immaterial items, such as texts and multimedia objects, having an objectively recognizable structure could possibly refer to the Getty AAT pieces of data published as Linked Open Data.

— the property *P71 lists* which takes place between an *E32 Authority Document* (domain) and a *E1 CRM Entity* could refer to the AAT as a thesauri containing lists of CRM entities e.g. the list of *alcazars* [2].

— the property *P189 approximates* which takes place between two places i.e. *E53 Place*, uses the TGN thesaurus, in addition to other online databases such as the GeoNames geographical database [3], to approximate the declarative places with point shapes. For example, $[4031'00.1"N2116'00.1"E]$ which is a *E53* instance *approximates* Kastoria, Greece, which is another *E53* entity, based on the TGN *ID* : 7010880 [4].

Moreover, in [Navigli2006a, Navigli2006b], the authors present a pattern-based method to automatically enrich a core ontology with the definitions of a domain glossary with an application on

1. https://cidoc-crm.org/useCasesPage
2. https://www.getty.edu/vow/AATFullDisplay?find=&logic=AND&note=&subjectid=300006897
3. GeoNames is a well-known geospatial dataset providing geographical data and metadata of around 7 million unique named places from all over the world collected from several sources.
4. https://www.getty.edu/vow/TGNFullDisplay?find=&place=&nation=&english=Y&subjectid=7010880

the enriching the CIDOC CRM core ontology using the resources in the AAT glossary, Wordnet, and the Dmoz taxonomy for named entities and other types of entities (based on the examples in CIDOC's specification).

In addition to the aforementioned aspects, CIDOC CRM encompasses a wide range of object properties that cover various types of relationships. These properties enable the identification of items through names and identifiers, describe the participation of items in temporal entities, specify the spatial and temporal location of entities, capture observation and assessment relationships, represent part-decomposition and structural properties, account for the influence of things and experiences on human activities, and facilitate the reference of information objects to other entities.

When it comes to temporal and event-based relationships, CIDOC CRM is particularly noteworthy for its comprehensive representation which enables the capturing of the historical context of CH entities e.g. [Binding2010a]. This capability allows for a more nuanced understanding of the temporal dimensions associated with CH data and facilitates the analysis of complex historical narratives and dependencies. Consider the historical event of *the signing of the Declaration of Independence in 1776*. The CIDOC CRM employs the concept of temporal entities and relationships to capture the temporal aspects of events. In this example, the event itself can be represented using the CRM *E4 Period* which represents a cultural manifestations occurring in time and space, with the appropriate temporal attributes. Additional properties can be used, e.g. *P7 took place at*, to make links to the other entities, e.g. the specific location in which this event took place with the corresponding range being an instance of the CRM *E53 Place*. Also, the historical context can be modeled by incorporating relationships between events and other entities that contextualize them. For instance, one can establish a *P9 consists of* relationship between the event of signing the Declaration of Independence and the specific document that was signed, such as *E31 Document*.

Furthermore, the CRMgeo model [Hiebel2017] is a global schema consistent with CIDOC CRM for integrating spatiotemporal properties of temporal entities and persistent items using the conceptualizations, formal definitions, encoding standards and topological relations defined by the Open Geospatial Consortium (OGC) [1]. Thus, linking the CIDOC CRM to the OGC standard of GeoSPARQL through a spatiotemporal refinement [Doerr2013]. Figure 2.9 shows the property hierarchy of the CRMgeo including classes and properties ; specific to the CRMgeo model, also referred to as "Spatiotemporal Model", with the prefix *"SP"* and *"Q"* ; those referring to CIDOC CRM entities and relationships with the prefix *"E"* ; and some classes of the GeoSPARQL ontology e.g. *SpatialObject* and *Feature*.

Besides CRMgeo, the CRMarchaeo model is a schema intended o provide all necessary tools to manage and integrate existing documentation in order to formalize knowledge extracted from observations made by archaeologists, recorded in various ways and adopting different standards [Doerr2018]. Figure 2.10 shows the property hierarchy of the CRMarchaeo which encompasses classes and properties ; specific to the description of archaeological excavations processes and entities, also referred to as "Excavation Model", with the prefix *"A"* ; those referring to CIDOC CRM entities and relationships, with the prefix *"E"* ; and those referring to the Scientific Observation Model entities and relationships with the prefixes *"S"* and *"O"*.

As mentioned earlier, all CIDOC CRM compatible models are provided as RDF schemas available online, including the CRMgeo [2] and CRMarchaeo [3].

1. OGC is a consortium of experts committed to improving access to geospatial or location information by producing standards and specific means to describe at high level geographic data in the the Semantic Web such as W3C Geo Vocabulary and GeoSPARQL [Query2012].
2. `https://cidoc-crm.org/crmgeo/sites/default/files/CRMgeo_v1_2.rdfs`
3. `https://cidoc-crm.org/crmarchaeo/sites/default/files/CRMarchaeo_v1.4.1.rdfs`

P. id	Property Name	Entity – Domain	Entity - Range
Q1	occupied	E4 Period	SP1 Phenomenal Spacetime Volume
Q2	occupied	E18 Physical Thing	SP1 Phenomenal Spacetime Volume
Q3	has temporal projection	SP1 Phenomenal Spacetime Volume	SP13 Phenomenal Time-Span
Q4	has spatial projection	SP1 Phenomenal Spacetime Volume	SP2 Phenomenal Place
Q5	defined in	E53 Place	SP3 Reference Space
Q6	is at rest in relation to	SP3 Reference Space	E18 Physical Thing
Q7	describes	SP4 Spatial Coordinate Reference System	SP3 Reference Space
Q8	is fixed on	SP4 Spatial Coordinate Reference System	E26 Physical Feature
Q9	is expressed in terms of	E94 Space Primitive	SP4 Spatial Coordinate Reference System
Q10	defines place	E94 Space Primitive	SP6 Declarative Place
Q11	approximates	SP6 Declarative Place	E53 Place
Q12	approximates	SP7 Declarative Spacetime Volume	E92 Spacetime Volume
Q13	approximates	SP10 Declarative Time-Span	E52 Time-Span
Q14	defines time	SP14 Time Expression	SP10 Declarative Time-Span
Q15	is expressed in terms of	SP14 Time Expression	SP11 Temporal Reference System
Q16	defines spacetime volume	SP12 Spacetime Volume Expression	SP7 Declarative Spacetime Volume
Q17	is expressed in terms of	SP12 Spacetime Volume Expression	SP11 Temporal Reference System
Q18	is expressed in terms of	SP12 Spacetime Volume Expression	SP4 Spatial Coordinate Reference System
Q19	has reference event	SP11 Temporal Reference System	E5 Event

FIGURE 2.9 – *The property hierarchy of the CRMgeo model.*

Property id	Property Name	Entity – Domain	Entity-Range
AP1	produced (was produced by	A1 Excavation Process Unit	S11 Amount of Matter
AP2	discarded into (was discarded by)	A1 Excavation Process Unit	S11 Amount of Matter
AP3	excavated (was excavated by)	A1 Excavation Process Unit	E53 Place
AP4	produced surface (was surface produced by)	A1 Excavation Process Unit	S20 Physical Feature
AP5	removed part or all of (was partially or totally removed by)	A1 Excavation Process Unit	A8 Stratigraphic Unit
AP6	intended to approximate (was approximate)	A1 Excavation Process Unit	A3 Stratigraphic Interface
AP7	produced (was produced by)	A4 Stratigraphic Genesis	A8 Stratigraphic Unit
AP8	disturbed (was disturbed by)	A5 Stratigraphic Modification	A8 Stratigraphic Unit
AP9	took matter from (provided matter to)	A4 Stratigraphic Genesis	S10 Material Substantial
AP10	destroyed (was destroyed by)	A1 Excavation Process Unit	S22 Segment of Matter
AP11	has physical relation (is physical relation of)	A1 Stratigraphic Unit	A8 Stratigraphic Unit
AP12	confines (is confined by)	A1 Stratigraphic Interface	A2 Stratigraphic Volume Unit
AP13	has stratigraphic relation (is stratigraphic relation of)	A1 Stratigraphic Modification	A5 Stratigraphic Modification
AP14	justified by	AP13 has stratigraphic relation	AP11 has physical relation
AP15	is or contains remains of (is or has remains contained in)	A8 Stratigraphic Unit	E18 Physical Thing
AP16	assigned attribute to (was attributed by)	A6 Group Declaration Event	A8 Stratigraphic Unit
AP17	is found by (found)	E7 Embedding	S19 Encounter Event
AP18	is embedding of (is embedded)	E7 Embedding	E18 Physical Thing
AP19	is embedding in (contains embedding)	E7 Embedding	A2 Stratigraphic Volume Unit
AP20	is embedding at (contains)	E7 Embedding	E53 Place

FIGURE 2.10 – *The property hierarchy of the CRMarchaeo model.*

During our examination of CIDOC CRM, we have identified two aspects that warrant further consideration regarding its composition relationships (object properties).
First, for structural relationships, it's modeling capabilities are not highly detailed when it comes to representing composition relations Indeed, the relations that it encompasses which can cover the structural composition between physical entities are (merely) *parthood* and *constitution* shown in Table 2.8, and explained below :

— *P45 consists of* : This property resembles ***constitution*** i.e. identifies the instances of *E57 Materials* of which an instance of *E18 Physical Thing* is composed, allowing the different materials to be recorded e.g. a silver cup (*E22 Human-Made Object* subclass of *E18*) consists of silver (*E57*).

— *P46 is composed of* : This property resembles ***parthood between physical entities*** i.e. associates an instance of *E18 Physical Thing* with another instance of physical thing that forms part of it, while the spatial extent of the composing part is included in the spatial extent of the whole e.g. the Royal carriage (*E22*) forms part of the Royal train (*E22*). If a component is not part of a whole from the beginning of its existence or until the end of its existence, the classes *E79 Part Addition* and *E90 Part Removal* can be used to document when a component became part of a particular whole and/or when it stopped being a part of it. *P46 is composed of* is transitive and asymmetric.

However, CIDOC CRM does not explicitly incorporate two significant structural relations : membership and dependence. These relations play a crucial role in representing the connections between parts and their (integral) wholes.

ObjectProperty	domain	range
P45 consists of (is incorporated in)	E18 Physical Thing	E57 Material
P46 composed of	E18 Physical thing	E18 Physical Thing

TABLEAU 2.8 – *The compositional structural relationships in CIDOC CRM.*

For ***membership***, representations such an artwork being part of a specific exhibition, or a collection of artifacts belonging to a particular CH category, are not supported. In general, expressing the inclusion of a physical entity within a larger group/collection is missing such as grouping of a number of entities under a physical characteristic e.g. the brocades that have vertical juxtaposed patterns on the sculpture.

For ***dependence***, the ontology does not provide explicit support for representing dependencies between entities. Dependencies capture the reliance or interconnectedness between different elements in a system. For instance, in the context of CH, a manuscript being dependent on its pages or an archaeological artifact being dependent on one part. The absence of these structural relations in CIDOC CRM limits its ability to fully capture and represent certain types of relationships between entities in the cultural heritage domain.

It should be noted that our analysis specifically focuses on relationships between endurant entity types, specifically physical entities, and does not consider relationships involving perdurant entities. Therefore, compositional structural relations such as *P5 consists of (forms part of)*, *P9 consists of (forms part of)*, *P106 is composed of (forms part of)*, *P148 has component (is component of)*, *P69 has association with (is associated with)*, *P150 defines typical parts of (defines typical wholes for E55 Type entities which resemble abstract perdurants)*, *P165 incorporates (is incorporated in)*, and others, are not within the scope of our analysis. We are specifically interested in relationships that hold between tangible entities.

Second, regarding the spatial compositional relations, CIDOC CRM defines relations identifying the location of an instance of *E18 Physical Thing* using an instance of *E53 Place* at different period/intervals of time : (a) former or current location using *P53 has former or current location* e.g. the silver cup (*E22*) has former or current location Display Case 4, Room 23, Museum of Oxford, (b) current permanent location using *P54 has current permanent location* e.g. the cup has current permanent location Shelf 3.1, Store 2, Museum of Oxford, and (c) current location at the time of validity of the record using *P55 has current location* e.g. the cup has current location Display Cabinet 23, Room 4, British Museum.

In addition, the ontology also provides the *P59 has section* property to link an instance of *E18 Physical Thing* to the instance of *E53 Place* in the which it is found e.g. HMS Victory (*E22*) has section HMS Victory section B347.6 (*E53*). In the preceding example, the instance of *E53*, B347.6 is expressed in terms of a coordinate system and takes the shape of the respective HMS Victory.

And the last spatial relationship is the *P89 falls within* property which identifies two instances of *E53 Place* upon which one falls wholly within the extent of the other another i.e. spatial containment without implying any relationship between things or phenomena occupying these places, e.g. the area covered by the World Heritage Site of Stonehenge (*E53*) falls within the area of Salisbury Plain (*E53*).

While considering these spatial relationships and their diverse temporal representations within CIDOC CRM, we believe that a specific representation of spatial relations is missing. This particular relationship occurs between two distinct *E18 Physical Thing* that are not parts of each other, but their respective parts share a common spatial location. In such cases, it would be a relationship between two instances of *E18 Physical Thing*. However, CIDOC CRM does not currently support this specific relation.

ObjectProperty	doamin	range
P53 has former or current location (is former or current location of)	E18 Physical Thing	E53 Place
P54 has current permanent location (is current permanent location of)		
P55 has current location (currently holds)		
P59 has section (is located on or within)		
P89 falls within (contains)	E53 Place	E53 Place

TABLEAU 2.9 – *The compositional spatial relationships in CIDOC CRM.*

To illustrate the need for this relationship, let's consider the example of a sculpture exhibition in a museum. The exhibition consists of multiple individual sculptures, each considered as separate Physical entities within CIDOC CRM. While the sculptures are not parts of each other, they share a common spatial location within the exhibition hall. This shared spatial location creates a meaningful relationship between the sculptures, indicating their coexistence within the same physical space i.e. the museum.

Similarly to our focus in structural relations, we also limit our analysis in spatial relations to those between endurant entities within CIDOC CRM. Therefore, spatial relations involving perdurant entities, such as *P10 falls within (contains)* and *P86 falls within (contains)* are not within the scope of our examination.

While the CIDOC CRM ontology does not fully encompass the desired representations regarding composition relations, it does extensively cover the scope of representing the spatiotemporal aspects of a CH entity. Combining the comprehensive coverage of spatiotemporal aspects provided by CIDOC CRM with the composition aspects that are sought after would result in a robust model significant for such a representation.

The Europeana Data Model (EDM) :

The basis of the metadata description is RDF statements. An XML Schema has been defined for describing classes and properties.

In order to facilitate the process of metadata aggregation, EDM defines a comprehensive set of elements, encompassing classes and properties, with the aim of incorporating as many elements as feasible from the descriptions provided by a content provider. Within this set, certain elements are re-used from other namspaces such as the Resource Description Framework (RDF) and the RDF Schema (RDFS) namespaces, the Object Reuse and Exchange (ORE) namespace, the Simple Knowledge Organization System (SKOS) namespace, the Dublin Core namespaces for properties (DCMI), the W3C Data Catalog Vocabulary (DCAT), the Creative Commons (CC) namespace, and the SIOC Services Ontology Module namespace (services).

Furthermore, other elements are specifically introduced within the EDM specification as EDM elements that align with predicates commonly employed in prevalent ontologies. In cases where this applies to certain classes and properties, the EDM specification provides explicit mappings to the equivalent classes and properties in these other ontologies, such as CIDOC-CRM and FRBR. For instance the *edm :Agent* class is equivalent to CIDOC CRM's *E39 Actor*. These mappings serve to establish connections and facilitate interoperability between different models.

EDM has been highly influenced by some formulated requirements and design principles, out of which is the requirement to distinguish between the "provided object" and it digital representation in order for metadata values to be associated properly [Isaac2013]. Upon this distinction, the representation of the CH object is presented, as shown in figure 2.11. The figure presents the three EDM core classes used to represent the core CH object : *ore :aggregation*, referring to a set of related resources that are grouped together as an aggregation under the ORE's namespace, and the two EDM classes *edm :WebResource* and *edm :ProvidedCHO* referring to the information resource having a URI, and to the CH object about which Europeana collects descriptions, respectively.

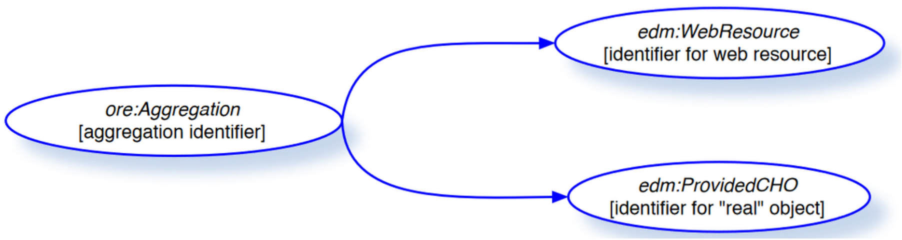


FIGURE 2.11 – *Figure taken from documentation of the EDM Mapping Guidelines showing the three core classes : an Aggregation with a Provided CHO and a Web Resource.*

In addition to classes and properties, Europeana is defining also controlled vocabularies useful for CHO interoperability (such as AAT, DDC, DBpedia, Iconclass). The main aim of Europeana is to work on Linked Data both exposing record sets [Haslhofer2011] and using Linked Data resources [Haslhofer2010] in order to augment Europeana content.

Ultimately, Europeana fulfills its purpose by serving as a platform for metadata aggregation. Its primary objective is not to serve as a model for directly representing the entirety of a CH object itself, but rather to collect and consolidate data pertaining to the object in a centralized repository. Indeed this focus holds significant importance and is one of the primary goals in the CH domain, e.g. using EDM as a top-level metadata model and mapping cross-domain metadata to it [Charles2013] and proposing new aspects for richer aggregation of data using new concepts [Noor2019].

However, if we are to use the EDM for representing the composition of a CH entity (as is the fundumental objective of this thesis), EDM will not fulfill that requirement adequately. Basically, the relationships that EDM encompasses which allow for covering part of our representation are the following :

— *ore :aggregates* : an ORE re-used property, expressing that an object resource is member of the set of aggregated resources of the subject i.e. hold between an Aggregation and its Aggregated resources

— *dcterms :hasPart* (*dcterms :isPartOf*) : a DC terms re-used property, expressing a related resource that is physically or logically included in the described resource.

— *edm :currentLocation*, and EDM specific property that is a sub property of the *dcterms :spatial*, exressing the geographic location, i.e. the range that is an instance of the EDM class *edm :Place*, of the object whose boundaries include the resource being described, i.e. the domain that is an instance of the EDM class *edm :ProvidedCHO*. It is equivalent to CIDOC CRM's *P55 has current location*

It is worth highlighting that EDM does encompass relationships other that enable the description of the composition of entities such as *edm :incorporates*, however this is not between tangible entities. We can use EDM as a complementary data model for describing CH specific properties such as *edm :isRelatdTo*, *edm :isSimilarTo*, *edm :isDerivativeOf*

Despite the fact that EDM may not fulfill our specific objectives in terms of representing the composition of a tangible entity using structural and spatial relationships, it can still be utilized as a complementary data model for describing other types of representations that center around objects as descriptive properties. For instance, EDM offers CH-specific properties such as *edm :isRelatdTo*, *edm :isSimilarTo*, and *edm :isDerivativeOf* which can be employed to capture and convey relationships around the CH entity as well.

The Cultural Heritage Abstract Reference Model (CHARM) :

The CHARM ontology is expressed in ConML ; a conceptual modeling language [Gonzalez-Perez2012] developed for humanities and social sciences. It is designed to be affordable for end users with no previous exposure to information technologies, and is based n object-oriented paradigm. The language is composed of a "types" package allowing for the creation of type models to represent the world such as classes (i.e. categories), attributes (i.e. properties of categories), and associations (i.e. relationships), and an "instances" package allwing for the creation of instance models represnting real world entities in terms of the types package.

Both, CHARM and ConML, are research outcomes of the Incipit lab of Spanish National Research Council (CSCI) [1]. While CHARM is expressed in ConML, ConML is also used to facilitate the extension mechanisms that allow users to extend CHARM into particular models that best suit their particular needs.

As presented in Section 2.3.2 and in figure 2.5, CHARM distinguishes between a *PrimaryEntity* and a *DerivedEntity*. Among the sub-types of primary entities is the *TangibleEntity* category, on which we focus. Tangible entities are defined as primary entities that are *fundamentally perceived in a direct fashion and through their materiality*, i.e. composed of matter and can be touched. Figure 2.12 further explores the different types of tangible entities in CHARM including places e.g. a valley, structures e.g. a building, objects e.g. a pebble, stratigraphies and samples. Further browsing through the list of classes and enumerated types is available at the CHARM website [2].

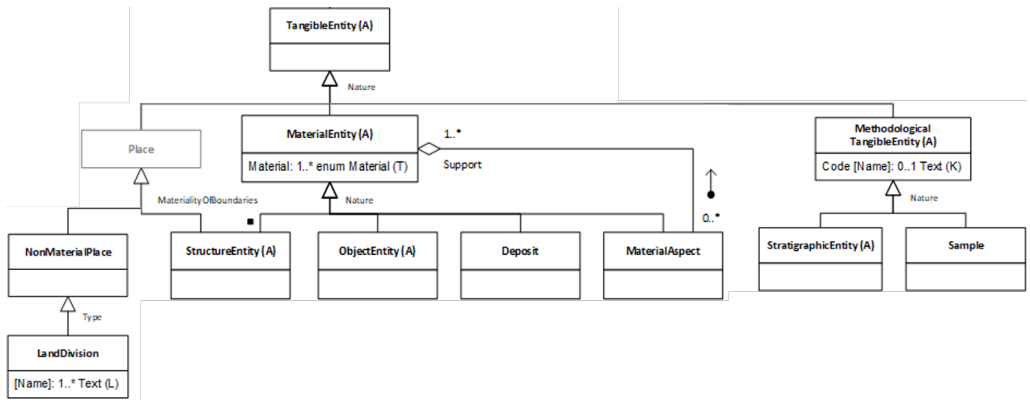


FIGURE 2.12 – *A snapshot of CHARM's hierarchy of tangible entities including Places, Material Entities, and Methodological Entities with some of their sub-classes.*

Regarding relationships, within the framework of CHARM in ConML, multiple types are provided that encompass different semantics, with many bearing similarities to UML relationships observed between entities.

The first type is the subsumption relationship, represented in a conceptual model through a generalization/specialization relationship. An illustration of subsumption can be observed in Figure 2.13, where a *chair* is a *sub-type* of *PieceOfFurniture*.

1. https://www.incipit.csic.es/
2. http://charminfo.org/Reference/Browse.aspx

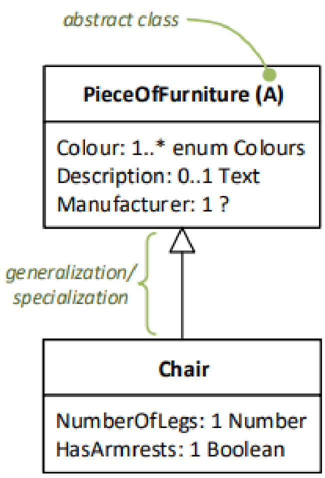


FIGURE 2.13 – *A snapshot of ConML's (version 1.5) Basic Graphical Notation Summary showing the generalization/specialization relationship between two classes.*

The second type is the classification relationship, depicted by means of an instantiation relationship in ConML, specifically avoiding the term *IsA*, as it may cause confusion between classification and subsumption, as pointed out by Guarino [Guarino1998b]. Figure 2.14 exemplifies a classification relationship, where an individual *ch1 :Chair* is classified as an instance of the class *Chair*.

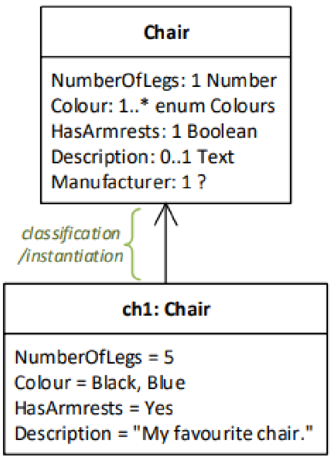


FIGURE 2.14 – *A snapshot of ConML's (version 1.5) Basic Graphical Notation Summary depicting the instantiation relationship between an individual and a class.*

The third type encompasses any intricate relational characteristic of classes, which is represented through association relationships, serving as semantic links connecting classes. An association comprises two semi-associations that are inverse of each other. Each semi-association belongs to a class referred to as the *participant class*, and connects to another class known as the *opposite class*. Each semi-association possesses a name and a cardinality constraint that describes the number of instances of the opposite class allowed for each instance of the participant class. Figure 2.15 demonstrates an example of an association named *IsLocatedIn* between the participant class *Chair* and the opposite class *Room*, with a cardinality of *0..**, indicating that 0 or any number of *Chair* instances can be located within a *Room* instance.

Lastly, a special type of association in ConML, extensively employed in CHARM, is the aggregation relationship, utilized to express the semantics of a part-whole relationship. Within the

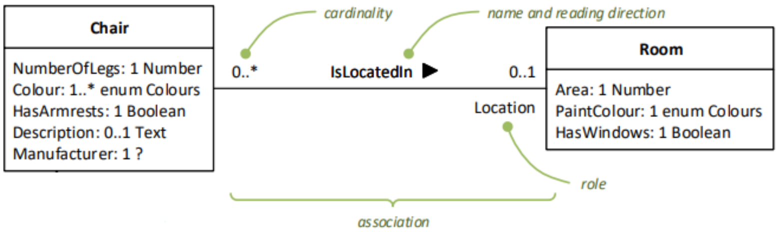


FIGURE 2.15 – *A snapshot of ConML's (version 1.5) Basic Graphical Notation Summary showing an example of an association between two classes.*

context of ConML, an aggregation relationship between two classes implies that an instance of class *A*, representing the *whole*, is composed of, consists of, or contains several instances of class *B*, signified by the star cardinality annotation $*$.

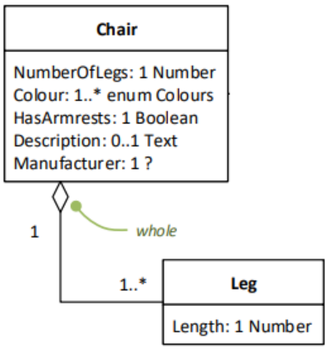


FIGURE 2.16 – *A snapshot of ConML's (version 1.5) Basic Graphical Notation Summary showing the aggregation relationship between two classes.*

As we are interested in representing the composition of tangible entities (following our objective), we proceeded with inspecting the aggregation associations between the tangible entities in CHARM, those of Figure 2.12.

And thus, we found around 14 aggregation relationships in CHARM between the following tangible entities [1], among which are the following between :

- *SubPlace* between a *Place* and itself, as a self-association, resembling spatial containment of a place entity in another place entity i.e. a place can be subdivided into subplaces, e.g. areas and regions contained in continents.
- *SubDivision* between a *LandDivision* and itself, as a self-association inherited from places (based on Figure 2.12, we can see that a *LandDivision* is a sub-class of *Place* and thus it inherits relations from it) with redefined cardinalities, e.g. states, provinces, and municipalities.
- *Support* between a *MaterialEntity* and a *MaterialAspect*, referring to a constitution relation. The class *MaterialEntity* is a tangible entity that is defined by its materiality through its association with the matter which consititutes it. In CHARM, this is represented through the class *MaterialAspect*, i.e. the matter, in its aggregation with the entity which it constitutes such as ceramic, bone, wood, etc.

1. For additional information on the class hierarchy in CHARM and the corresponding aggregations holding between them, please refer to the book [Gonzalez-Perez2018], and/or to the ConML diagrams available at `http://charminfo.org/Reference/Browse.aspx`.

— *Content* between a *StructureEntity* and a *MaterialEntity*, referring to the containment of material entities in a structure, e.g. a house being inside an enclosure.

— *Content* between an *ObjectEntity* and itself, as a self association to refer to the containment of an object inside another object, e.g. a basket containing tools.

— *Fragment* between a *CompleteStructure* and a *StructureFragment*, referring t a part-whole relation between a structure having altered its material integrity, e.g. broken, and a fragment of it, e.g. an ashlar block from a wall.
Similarly, a *Fragment* relation holds between *CompleteObject* and an *ObjectFragment*, e.g. a shard from a broken clay beaker.

— *SubStructure* between a *CompleteStructure* and itself as a self-association, referring to the containment of structures in larger frame structures, e.g. a farm containing a house, a trench, and a barn as as complete structures too.
Similarly, a *SubObject* relationship holds between a *CompleteObject* and itself.

— *Element* between a *ConstructedStructure* and a *ConstructiveElement* as a material part/whole relationship e.g. the columns of a house being part of the house

Thus, one of the strong points we identified in the CHARM model is its capture of a wide range of structural and spatial relations through the use of aggregations. This comprehensive approach ensures that various aspects of CH entities can represented effectively.

However, the semantics of its associations are not formally specified. Although these associations serve to represent connections between classes, the precise interpretations of these connections are not explicitly defined. Instead, they rely on the names, directions assigned to them, and natural language.

This aspect can be attributed to the nature of the ConML conceptual modeling language itself as a UML class metamodel in which its graphical notation is primarily described through natural language and diagrams. As a result, no formal logical specification is provided to support reasoning capabilities or the inclusion of additional rules for expressing rich semantics of relations.

Both, the CHARM model and ConML language, are user-friendly tools that enable domain experts to construct domain-specific ontologies. These tools provide a starting point for capturing the semantics of various fields. Additionally, ConML diagrams can be translated into more formal ontologies like OWL, enhancing the expressiveness and interoperability. CHARM can also be directly employed by domain institutions to represent the semantics of their systems. Overall, these tools facilitate interdisciplinary collaboration and support the representation of CH entities in a structured and meaningful way.

2.5 Final considerations

Having defined the motivation behind this Chapter ([motivation-I]), we have specified the modeling scope of the intended interdisciplinary modelisation. This is to model the composition of a tangible entity as a complex structure (i.e. an object, a place, a collection) in a manner that enables the understanding, constructing, and navigating into its tangible discourse (i.e. representation), and learning its intangible aspects (i.e. significance).

In Section 2.2, we have seen the basic challenges present in the CH field focusing around the heterogeneity of data (both syntactic and semantic) as well as the integration problem. Integration is problematic not only at the level of data, but also metadata and schemas. Furthermore, we assured the importance of achieving interoperability for solving both problems. This is by arriving to a common conceptualization based on shared goal around the cross-disciplinary CH entity.

Then, in Section 2.3 we have presented the memory and organization institutions, as well as some information systems and databases that manage CH data using best practices and standards. This was followed by outlining various knowledge organization systems that enable the organization of data in the CH field, ending by refining our analysis to concentrate specifically on ontology models as knowledge organization structures.

After that, in Section 2.4, we performed a systematic classification of the extant ontology modeling approaches based on the predefined criteria : the geographical scale, the semantics and formality level, an the modeling scope. We have categorized each criteria and classified the approaches according to these categories. Based on this classification system, we were able to select the relevant ontology models that we believe could (maybe) do the intended representations required by the specified [fundamental objective] in general and [motivation-I] in particular. After excerpting the relevant models, CIDOC CRM, EDM, and CHARM, we analyzed each model in details in terms of its formalization, available resources, and classes and properties hierarchy focusing on the spatial and structural relationships conveyed in each.

The analysis showed the following. In the case of CIDOC CRM, while a pretty good number of relationships is provided, still there is not a complete coverage of the intended representation with respect to our goal such as the memberhsip, dependence, and representing spatial inclusion between entities. With EDM, the case is more difficult since it is more likely a data-centric approach than an entity-centric approach, where a limited number of compositional (structural and spatial) relations is present. As for CHARM, the issue is more likely with the formality of the model and the use of ConML modeling language which is very useful for non-ontologists, yet not rich in terms of semantics.

In conclusion, we recognized that there is a need for a semantically robust and well-formalized set of foundational ontological relations that can capture the composition of a CH entity and enable an accurate representation of its materiality. This emphasis on composition relations forms the core objective of this thesis. Furthermore, integrating such composition relations into existing models, such as CIDOC CRM as a core ontology and CHARM as a domain-specific ontology for certain CH domains, would provide a comprehensive framework for the presentation of CH entities, filling the existing need in current models.

As such, in the next chapter 3, we delve into foundational ontological relations, those that allow the representation of the structural and spatial composition of a tangible entity, namely some foundational ontological relations.

3

Foundational ontological relations

This Chapter investigates some pertinent structural and spatial foundational ontological relations from the applied ontology literature.

Sommaire

3.1 Introduction

In Chapter 2, we conducted an investigation into the existing ontological models found in the literature concerning cultural heritage. Our focus was specifically on models that enable the representation of the composition of tangible entities. We performed a comprehensive analysis of three relevant ontology models, specifically examining their treatment of composition relations in terms of structure and spatial aspects. The analysis revealed a clear requirement for a more semantically comprehensive theory of ontological relations that can effectively capture and model the composition of tangible entities within cultural heritage.

In this chapter, aligned with our fundamental objective, we introduce a complementary motivation ([motivation-II]) to further strengthen our primary motivation ([motivation-I]) by emphasizing the specific focus required to develop a comprehensive model of composition. This additional point underscores the necessity for a language that encompasses a range of ontological relations pertaining to both structural and spatial composition within a unified theory. Consequently, we delve into an examination of the existing literature on foundational ontological relations, specifically targeting the domain of structural and spatial relations within the field of applied ontology.

Motivation-II :
Acquiring a number of foundational ontological, structural and spatial, relations that enable representing the composition of a tangible entity, within a theory.

First, in Section 3.2, we provide an introduction to foundational ontological relations, encompassing their definition, fields of study, and various types. Our focus narrows down to specifically examine structural and spatial relations. Subsequently, in Section 3.3, we explore several studies that delve into the realm of structural part-whole relations, encompassing mereological, meronymic, and other related approaches. Shifting our attention in Section 3.4, we delve into the presentation of spatial relations within the context of representing the spatial aspects of tangible entities, including mereotopology and location relations. Finally, we discuss and conclude in Section 3.5.

3.2 The study of foundational ontological relations : a categorization

"Foundational ontological relations are formal and (often) primitive relations that pay a fundamental role in the foundations of ontological analysis and modeling". Our interpretation of this definition is based on the following aspects.

- Firstly, these relations are considered "**ontological**" as they are studied and represented within ontologies to capture the relationships between entities.
- Secondly, the term "**foundational**" highlights their fundamental significance in the field of applied ontology.
- Thirdly, we define "foundational ontological relations" as possessing two key characteristics : **formality** and **primitiveness**.
  - The formality aspect relates to the universality and specification of relations. In [Gangemi2001], formal relations are explained as relations that involve entities in all "material spheres", so that they are understandable as universal notions. This is based on Smith's reflection, in [Smith1998], on formal ontology as dealing with with properties

of objects which are formal in the sense that they can be exemplified, in principle, by objects in all material spheres or domains of reality. This is a formality in view of the philosophical literature referring to being domain neutral ; what we also like to refer as generic in the sense of general and independent from a specific domain. Additionally, formality extends the neutrality of the domain to that of the framework in which the relations are specified within a formal framework using formal logical languages.

— The primitiveness characteristic generally pertains to notions (such as concepts, properties, or relationships) that are not defined in terms of previously-defined notions [1]. Typically, the introduction of primitives is motivated informally, drawing on intuition and everyday experiences. In axiomatic theories, primitives are established through a group of axioms that characterize the notion and provide restrictions on its usage.

Where are foundational ontological relations studied ?

Although formal relations hold a central position in classical philosophical investigations (e.g. [Edmund1970, Fine1983, Fine1999]) within the domain of "formal ontology," as discussed in [Smith1998], their formalization in computer science, through ontologies, has been explored extensively in the applied ontology literature as investigating fundamental relationships between entities or entities and properties, at a meta-level. Given that applied ontology focuses on studying and developing ontological theories at a meta-level, as outlined in the Preliminary Remarks section, the examination of foundational ontological relations primarily resides within this domain. Examples of studies encompass a wide range of relationships including identity [Guarino2000a], parthood [Simons1987], causation [Salmon1984], dependence [Edmund1970], etc.

How are foundational ontological relations studied ?

The study of foundational ontological relations has given rise to various contributions, encompassing approaches that offer *taxonomies as hierarchical structures of relations*, and others that provide *theoretical frameworks for formalizing relations and establishing interconnections*.

A taxonomy of relations refers to a systematic hierarchical classification system that organizes relations based on their inherent properties and hierarchical levels. Within the context of foundational relations, taxonomies present a structured arrangement of relations, showcasing their subsumption or specialization relationships, which facilitates a clear understanding of their interrelationships and hierarchical structure. The development of taxonomies of relationships has been mostly deployed in cognitive sciences studies such as a taxonomy of part-whole relations [Winston1987] in linguistics.

Whereas a theory involves the development of formal frameworks often accompanied by logical systems designed to capture and represent the essence of foundational relations. For instance, in [Gerstl1996] a conceptual theory of part-whole relation in common-sense reasoning was presented without necessarily formalizing the relationships in terms of a formal logical language. Other theories provide a set of axioms (algebraic such as transitivity, symmetry, etc.) and rules governing the behavior and properties of relations, enabling their formalization and reasoning. In such formal theories e.g. [Bittner2005], the emphasis is on defining a logical structure that enables rigorous representation and inference.

Both approaches play crucial roles in enhancing our understanding of entity relationships and providing a systematic framework for their application in ontological and conceptual modeling tasks. Notably, there has been a particular emphasis on ontology-driven conceptual modeling, exemplified by the development of the Unified Foundational Ontology (UFO) [Guizzardi2005] and

1. https://en.wikipedia.org/wiki/Primitive_notion

the adoption of the UFO-based ontology-driven conceptual modeling language, OntoUML, in various domains, including the business sector with applications in value, risk, service, and contract modeling [Verdonck2016]. In a related work [Fonseca2019], the authors revisit UFO's theory of relations based on empirical feedback from different experiences, proposing a new theory that subsequently informs the design of a new metamodel for OntoUML. This development enhances the utilization of relations in ontology-driven conceptual modeling tasks.

In general, formalizing foundational ontological relations/categories through ontological analysis serves as a valuable addition that benefits modeling tasks, whether in ontology modeling or conceptual modeling. It provides a reference meta-model that enables the validation of domain-specific models in accordance with the formal semantics of the meta-level.

What are examples of studied foundational ontological relations ?

Several examples demonstrate the significance of certain foundational ontological relations across different domains. For instance, the parthood relation has proven crucial in aligning and correlating ontologies within the bioinformatics domain [Bittner2004a]. Location and topological relations have been instrumental in disambiguating spatial information in biomedical ontologies, thereby enhancing automatic reasoning capabilities [Donnelly2006]. In the realms of cognitive sciences, linguistics, ontology, and conceptual modeling, contextualizing parthood typologies based on the types of participating entities has been a focus of research [Pribbenow2002, Winston1987, Bittner2004c, Guizzardi2009].

Furthermore, other studies, such as the work by Smith et al. [Smith2004b], have addressed the disambiguation of similarities between relations that can lead to problematic inconsistencies, specifically between class subsumption (the is-A relation) and partonomic inclusion (the part-of relation).

These examples highlight the importance of understanding and disambiguating relations within various domains, as they play a significant role in aligning ontologies, enhancing reasoning capabilities, contextualizing typologies, and resolving inconsistencies that can arise in relation-based modeling.

Constricting our scope to structural and spatial relations

Our research narrows down its focus to structural and spatial relations, aligning with the fundamental objective stated in Chapter 1 and the specified [motivation-II]. We aim to investigate and utilize a well-formalized theory that encompasses a comprehensive set of foundational ontological relations addressing the structural and spatial aspects of tangible entity composition, provided such a theory exists

Thus, we have conducted an extensive review of the literature to gather taxonomies and theories concerning structural and spatial foundational ontological relations.

For structural relations, our investigation centers on studies that primarily focus on part-whole relations and their typologies. It should be noted that our use of the term "typologies of part-whole relations" does not reflect our subjective opinion on the taxonomies/theories as definitive part-whole relation typologies. Our intention is to describe the nature of these studies, wherein some researchers designate their work as specializing in distinct "types of part-whole relations," while others do not label them as such. To differentiate between the specific term "part-whole relations" used in reference to studies that specifically denote their relations as part-whole ones, and our general use, we will denote the former by enclosing the term within double brackets, such as "((part-whole)) relations" when referring to the specific taxonomy/theory. This indicates the terminology used by the authors of a specific reference, without necessarily endorsing our particular viewpoint on these relation types.

Regarding spatial relations, we explore approaches that capture location and spatial part-whole relationships.

In Table 3.1, we categorize some popular extant approaches pertaining to both structural and spatial relations based on their context of studies : formal applied ontological studies (e.g. location theories, topology, and mereology), cognitive sciences studies (e.g. meronomy), and other approaches that combine both meronymic and mereological studies in single taxonomies/theories. For the latter, we distinguish approaches based on the language used for formalization of the theory into conceptual modeling languages (Unified Modeling Language (UML) [ISO/IEC195012005], Entity-Relationship (ER) language [Chen1976], Object-Role Modeling (ORM) language [Halpin2010]), and other knowledge representation and reasoning languages (Description Logics (DL) [Baader2003, Calvanese2003] and First-Order logic (FOL) [Smullyan1995]).

In the following sections, we delve into further exploration of the approaches wround sturtcural and spatial relations in Sections 3.3 and 3.4 respectively.

<table>
<tr><th colspan="2">Spatial relations</th><th colspan="6">Structural relations</th></tr>
<tr><th colspan="3">Formal Applied Ontological studies</th><th>Cognitive Sciences studies</th><th colspan="4">The meddling of both < mereology >and < meronomy >in formal taxonomies</th></tr>
<tr><th>Spatial Location</th><th>Spatial Connection</th><th>Formal Parthood</th><th>(Part-Whole) Relation Taxonomies</th><th colspan="2">Conceptual modeling languages</th><th colspan="2">Knowledge representation languages</th></tr>
<tr><th><Location></th><th><Topology></th><th><Mereology></th><th><Meronomy></th><th>UML</th><th>ER/ORM</th><th>DL</th><th>FOL</th></tr>
<tr><td rowspan="3">[Varzi1996]
[Casati1999]
[Varzi2007]</td><td>[Randell1992]
[Randell1989]
[Cui1993]</td><td>[Simons1987]
[Varzi2003]</td><td>[Iris1986]</td><td>[Odell1998]
[Opdahl2001]
[Motschnig-Pitrik1999]</td><td rowspan="3">[Keet2006b]</td><td>[Schulz2000]
[Sattler2000]</td><td>[Guizzardi2005]</td></tr>
<tr><td colspan="2"><Mereotopology></td><td>[Winston1987]</td><td rowspan="2">[Barbier2003]
[Shanks2004]
[Berardi2005]</td><td>[Artale1996b]
[Artale1996a]</td><td>[Keet2008]</td></tr>
<tr><td colspan="2">[Varzi1993] [Varzi1996]
[Casati1999] [Varzi2007]</td><td>[Gerstl1995]
[Gerstl1996]</td><td colspan="2">[Bittner2005]</td></tr>
</table>

TABLEAU 3.1 – *A categorization of some structural and spatial relations from the literature in formal applied ontology, cognitive sciences, and other common-sense reasoning approaches .*

3.3 Studies on structural (part-whole) relations

The part-whole relation (*part-of*) has gained significant attention in the field of knowledge representation and reasoning, serving as a fundamental ontological relation [Burkhardt1991]. Extensive research has been devoted to exploring various aspects of this relation, leading to the development of taxonomies and theories. [1]

From a formal ontological perspective, traditional accounts of the part-of relation are predominantly found in formal ontology and are considered widely accepted and universal. These theories, categorized under the term **mereology** [Leśniewski1991, Simons1987, Varzi2003], fall under the classification of "formal parthood" in Table 3.1. Further investigation into mereological theories will be conducted in Section 3.3.1.

Later in works such as [Lyons1977] and [Cruse1986], it was recognized that mereological relations alone cannot fully capture the complexities of part-whole relations, particularly in cases of intransitivity observed in natural language. Building upon this observation and the suggestions of

1. Please note that a significant portion of this review is based on the papers [Keet2006a] and [Fernández-López2008], which provide a comprehensive studies of part-whole relations from various perspectives.

Lyons, several authors explored the hypothesis of multiple part-whole relations to address cognitive tasks. Subsequently, several other authors, including Keet in [Keet2006a], distinguish between mereology and meronomy as distinct fields of investigation, with the former being ontological and the latter focusing on linguistics.

Based on that, the study of part-whole relations within cognitive science originated in linguistics, with works such as [Iris1986] and [Winston1987] differentiating various cognates of "part" (e.g., portion, element, member, fragment, component, constituent) based on their differing semantics. Subsequent research, including refinements proposed in works like [Gerstl1995] and [Vieu2007], delved further into the exploration of different types of part-whole relations. The distinction between types of part-whole relations is grouped under the term **meronomy**. Some meronymic studies, found under "part-whole relation taxonomies" in Table 3.1, will be investigated in Section 3.3.2.

Besides the aforementioned studies, various approaches have been developed for modeling part-whole aspects using different knowledge representation and conceptual modeling languages. Examples include extensions to UML to incorporate reasoning capabilities behind part-whole relations [Barbier2003], using ER modeling [Shanks2004], ORM-based proposals [Keet2006b], application of FOL [Guizzardi2005, Keet2008], and employment of DL [Artale1996b, Bittner2005]. These approaches, which combine elements of both mereology and meronomy in formal taxonomies, can be found in the "the meddling of both mereology and meronomy in formal taxonomies" section of Table 3.1. Further exploration of these approaches will be discussed in Section 3.3.3.

Additionally, in Section 3.3.4, we provide a brief overview of some approaches that have extended the study of part-whole relations to account for properties such as functionality, dependence, typologies of wholes, and granularity.

3.3.1 Formal theories of parts : Mereology

The research on mereology started with Lesniewski's seminal work "Foundatios of the general theory of sets" in the early 20th century (1901-2000). Since then, significant contributions to the field have been made by Peter Simons [Simons1982, Simons1987] and Achille Varzi [Varzi1996, Varzi2003], who have played prominent and influential roles in expanding the research on mereology from a philosophical standpoint.

Ground mereology **M** is the common core of any mereological theory presenting *formal parthood*. It is a denoted using the primitive $\mathrm{P(x,y)}$, standing for "x is part of y", as a partial order relation : reflexive (Pa1), antisymmetric (Pa2), and transitive (Pa3).

$$(\forall \mathrm{x})\mathrm{P(x,x)} \tag{Pa1}$$

$$(\forall \mathrm{x,y})(\mathrm{P(x,y)} \wedge \mathrm{P(y,x)}) \rightarrow \mathrm{x} = \mathrm{y} \tag{Pa2}$$

$$(\forall \mathrm{x,y,z})(\mathrm{P(x,y)} \wedge \mathrm{P(y,z)}) \rightarrow \mathrm{P(x,z)} \tag{Pa3}$$

Using P, other mereological predicates are built for a wider semantic range ; *proper-part* (PP) which is asymmetric and irreflexive [1], *equal* (EQ), *overlap* (O), *underlap* (U), *overcross* (OC), *undercross* (UC), *proper-overlap* (PO), and *proper-underlap* (PU). These predicates are depicted in Figure 3.1. [2]

1. A relation R is asymmetric iff ; if $R(x,y)$ then $\neg R(y,x)$, and irreflexive iff ; $\neg R(x,x)$.
2. The visual presentations regarding mereology presented in Figure 3.1 are taken from `http://journal.b-pro.org/article/the-ultimate-parts/`.

$$(\forall x,y)PP(x,y) \leftrightarrow (P(x,y) \wedge \neg P(y,x)) \quad \text{(Pd1)}$$
$$(\forall x,y)EQ(x,y) \leftrightarrow (P(x,y) \wedge P(y,x)) \quad \text{(Pd2)}$$
$$(\forall x,y)O(x,y) \leftrightarrow \exists z(P(z,x) \wedge P(z,y)) \quad \text{(Pd3)}$$
$$(\forall x,y)U(x,y) \leftrightarrow \exists z(P(x,z) \wedge P(y,z)) \quad \text{(Pd4)}$$
$$(\forall x,y)OC(x,y) \leftrightarrow (O(x,y) \wedge \neg P(x,y)) \quad \text{(Pd5)}$$
$$(\forall x,y)UC(x,y) \leftrightarrow (U(x,y) \wedge \neg P(y,x)) \quad \text{(Pd6)}$$
$$(\forall x,y)PO(x,y) \leftrightarrow (OC(x,y) \wedge OC(y,x)) \quad \text{(Pd7)}$$
$$(\forall x,y)PU(x,y) \leftrightarrow (UC(x,y) \wedge UC(y,x)) \quad \text{(Pd8)}$$

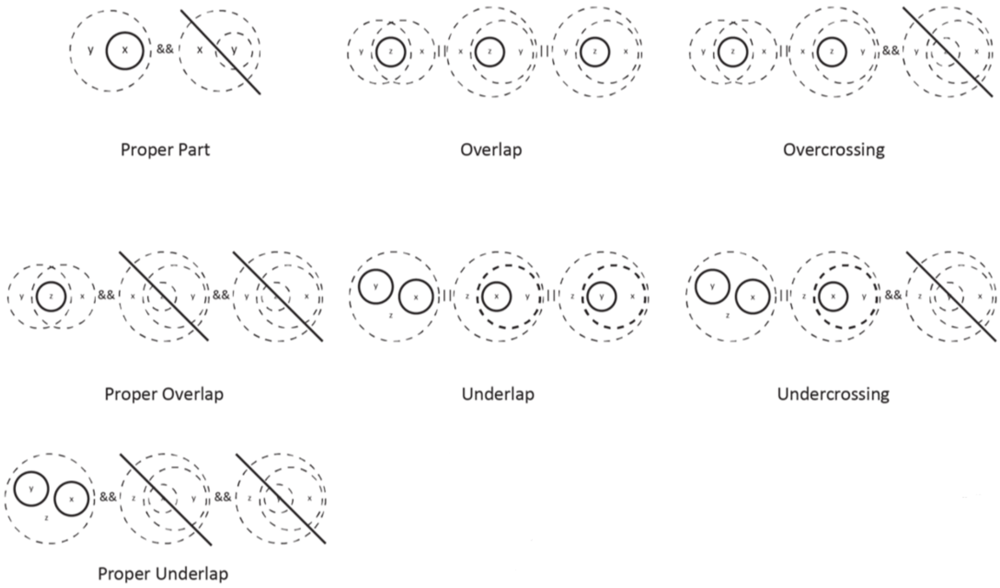


FIGURE 3.1 – *A visual representation of some mereological predicates, figure taken "The ultimate parts" blog artcile.*

Based on ground mereology, other mereological theories are generated by adding axioms to **M** to allow for finer grained relations or/and permit intransitivity in some cases. As illustrated in figure 3.2, starting from **M**, one can add assumptions of decomposition and/or composition principles. For each axiom, we clarify its meaning based on the explanations from [Fernández-López2008].

The principle of decomposition introduces the argument that if some y has a proper part x, then there should be a remainder because x is less than y. This can be added either using weak supplementation Pa4 resulting Minimal Mereology **MM**, or strong supplementation Pa5 resulting in Extensional Mereology **EM**.

— Pa4 : Every object y with a proper part x has another part z (different than x) that is disjoint of x. For example, given that Spain is a proper part of the Europe, then Europe has other parts that are disjoint of Spain : Portugal, France, Italy, etc.

— Pa5 : If y is not part of x, then there is a part of y that does not overlap with x. For example, given that Spain is not part of Africa, there is a part of Spain (e.g. Madrid) that is not part of Africa

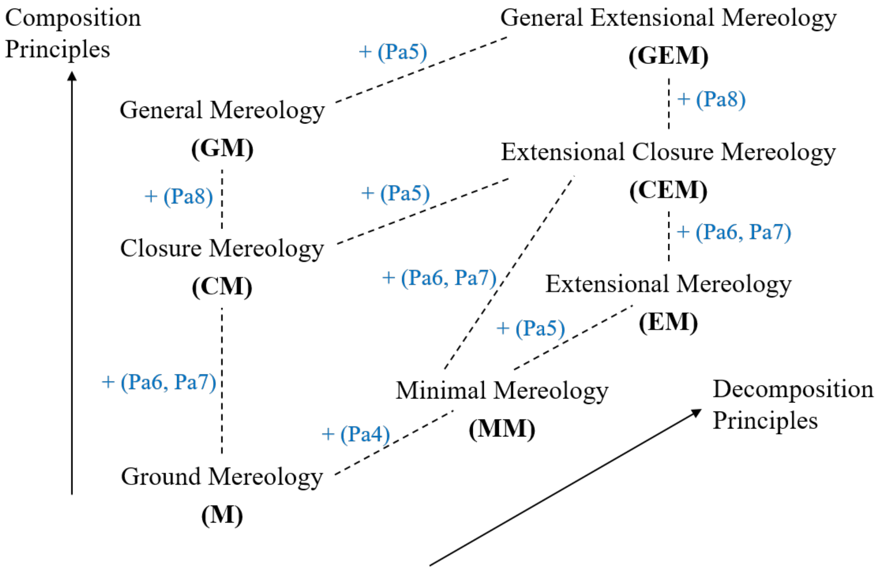


FIGURE 3.2 – *Hasse's diagram for mereological theories (from weaker to stronger).*

Note that in EM, the theorems Pt1 and Pt2 are implied upon the addition of the supplementation axioms, and therefore the theorem Pt3 [Varzi1996].

The principle of composition introduces the argument that a mereological domain shall be closed under various operations. This done either via finitary operations, i.e. finitary sum Pa6 and finitary product Pa7, resulting in Closure Mereology **CM**, or via unrestricted fusions Pa8 i.e. the fusion of all objects satisfying the formula ϕ in which x is a free variable, resulting in General Mereology **GM**.

— Pa6 : (Sum principle)If x and y underlap, then there is a z such that, for all w's, w overlaps z if and only if w overlaps x or w overlaps y. That is, if two objects underlap, then it may be assumed that there is a smallest object of which they are part (an object that exactly and completely exhausts both)

— Pa7 : (Product principle) If x overlaps y, then there is a z such that for all w's, w is part of z if and only if w is part of x and w is part of y. That is, if two objects overlap, then it may be assumed that there is the largest object that is part of both (the common part at their junction). For example, Spain and Africa overlap, and it may be assumed that there is the largest object overlapped by both : Canaries, Ceuta, Melilla, etc.

— Pa8 : (Unrestricted fusion principle) For every satisfied property or condition ϕ (if there exists an entity x that satisfies ϕ), then there is a z such that for all y's, y overlaps z if and only if there is an x such that it satisfies ϕ and overlaps y. That is, there is an entity consisting of all those things that satisfy ϕ. For example, suppose that ϕ means "country with more than 10 million inhabitants", then there is an object that consists of all the countries with more than 10 million inhabitants.

Similarly, the result of adding Pa6 and Pa7 to MM/EM yields in **CMM/CEM** i.e. Closure Minimal/Extensional Mereology. However, Pa7 implies Pa5 by means of Pa4, thus making **CMM** the same as **CEM** [Varzi2003]. Moreover, to ensure the unicity of the entities that are held by means of Pa6 (the sum entity) and Pa7 (the product entity) under the presence of extensionality, it is possible to add the axioms Pa9 and Pa10, respectively. These last two axioms Pa9 and Pa10 ensure both, the existence of the sum/product entity, and the unicity (uniqueness) of this entity.

$$(\forall x,y) PP(x,y) \rightarrow \exists z(P(z,y) \land \neg O(z,x)) \quad \text{(Pa4)}$$
$$(\forall x,y) \neg P(y,x) \rightarrow \exists z(P(z,y) \land \neg O(z,x)) \quad \text{(Pa5)}$$
$$PP(x,y) \rightarrow \exists z(PP(z,y) \land \neg O(z,x)) \quad \text{(Pt1)}$$
$$(\exists z(PP(z,x)) \land \forall z(PP(z,x) \rightarrow PP(z,y))) \rightarrow P(x,y) \quad \text{(Pt2)}$$
$$(\exists z PP(z,x) \land \forall (PP(z,x) \leftrightarrow PP(z,y))) \rightarrow x = y \quad \text{(Pt3)}$$
$$(\forall x,y) U(x,y) \rightarrow \exists z \forall w(O(w,z) \leftrightarrow (O(w,x) \lor O(w,y))) \quad \text{(Pa6)}$$
$$(\forall x,y) O(x,y) \rightarrow \exists z \forall w(P(w,z) \leftrightarrow (P(w,x) \lor P(w,y))) \quad \text{(Pa7)}$$
$$(\exists x)\phi(x) \rightarrow \exists z \forall y(O(y,z) \leftrightarrow \exists x(\phi(x) \land O(y,x))) \quad \text{(Pa8)}$$
$$(\forall x,y) \exists z(O(z,x) \land O(z,y)) \rightarrow \exists z(\forall w(O(w,z) \leftrightarrow (O(x,w) \lor O(y,w))) \quad \text{(Pa9)}$$
$$(\forall x,y) \exists z(P(z,x) \land P(z,y)) \rightarrow \exists z(\forall w(P(w,z) \leftrightarrow (P(x,w) \lor P(y,w))) \quad \text{(Pa10)}$$

Mereology offers a solid and formally grounded framework for analyzing and representing part-whole relations, providing valuable insights from both mathematical and philosophical perspectives. However, its application as a theory of parthood in conceptual and ontological modeling tasks presents challenges, as discussed in [Guizzardi2005] and supported by various authors (e.g., [Odell1998], [Opdahl2001], and [Pribbenow2002]). These challenges arise from either the theory being deemed too strong to capture the nuances of part-whole relations at the conceptual level, where it imposes constraints that may not universally apply, or too weak to adequately distinguish between different typologies of the part-whole relation. We clarify these issues based on the insights presented in [Guizzardi2005] and [Guizzardi2005].

— Firstly, concerning **M**, it is observed that transitivity (Pa3) is derived in cases where it may not be applicable. For instance, it falsely implies that my hand is part of a research group, based on the fact that my brain is part of myself, and I am part of the group. Guizzardi attributes this problem to mereology being a theory of parthood and highlights the need for an additional theory of wholes [Gangemi2001] to complement it in the context of conceptual modeling. This is because a theory of parthood alone does not account for the diverse roles that parts can play within a whole.

— Secondly, regarding **EM** and its extensions, these theories infer identity between entities that share the same proper parts (theorem Pt3, which is implied by the supplementation axiom Pa5), even in cases where it may be irrelevant. For instance, it may incorrectly identify a soccer team and an orchestra group consisting of the exact same members as identical entities. Authors such as [Guizzardi2005] and [Gerstl1995] argue that these inferences are unacceptable and that the property of extensionality in **EM** theory leads to incorrect inferences, failing to differentiate entities that are perceived as distinct and equating entities that should be considered different.

— Lastly, **GEM** introduces entities such as the sum (Pa9) and product (Pa10) entities with the inclusion of the unrestricted fusion axiom (Pa8). However, these entities may be deemed irrelevant in certain contexts. For example, considering the sum of my brain, my cat's leg, and my car as a meaningful entity in conceptual modeling systems unless these entities serve a specific role or represent a genuine universal (such as a class type or concept), as discussed in [Pribbenow2002] and [Guizzardi2005] respectively. Pribbenow argues that, in everyday understanding, we only accept the summation of entities if the resulting mereological sum plays a meaningful role in the intended conceptual model, such as the bottle and its cap forming an integral whole that is recognized by humans.

To address these challenges in the context of conceptual modeling tasks, alternative approaches have been proposed, such as complementary theories of wholes [Gangemi2001, Guizzardi2005], typologies of part-whole relations within meronomy [Gerstl1995], and typologies of universals [Armstrong2018].

3.3.2 Other meronoymic taxonomies

In order to illustrate some meronymic studies, it is important to clarify the distinction between "meronymy" and "meronomy". Meronymy refers to the semantic relation between a part (meronym) and a whole (holonym) [1], while meronomy pertains to the hierarchical organization of meronymic relations. Meronomies have been introduced as a means of studying part-whole relations in everyday cognition, where these relations are not necessarily transitive.

The exploration of meronomy began with the groundbreaking work of Winston, Chaffin, and Heramnn in the development of the WCH taxonomy [Winston1987]. This seminal research paved the way for subsequent investigations that aimed to model and build upon the WCH taxonomy [Gerstl1995, Artale1996b, Odell1998, Guizzardi2005]. Notably, Grestl and Pribbenow conducted a prominent study in 1995, wherein they proposed "a common sense theory of part-whole relations". While there have been other proposed approaches, these two studies have remained widely recognized in the field of meronymic research, which we present and analyze below.

The WCH taxonomy, 1987

The first investigation on meronymic relations was motivated by linguistics in the WCH taxonomy [Winston1987]. The authors distinguished between three types of inclusion; spatial, meronymic, and class inclusion. And within meronymic inclusion, they proposed a taxonomy of part-whole relations based on the type of the whole entity and its corresponding part. The whole can be a concrete physical object, a collection, a mass, an area, an assembly, a representational object, an abstract object, or an organization.

The authors distinguished three characteristic properties of relations : functional (F), homeomerous (H), and separable (S). For each property, two possible values exist, either true or false ($F/\neg F$, $H/\neg H$, $S/\neg S$). A part-whole relation is :

- functional, if parts are considered to be functional if they possess a specific restriction, according to their spatial or temporal location, towards their whole. A functional part, without the restriction of its spatial or temporal location, does not function as it is supposed to be.
- homeomerous, if parts are considered to be homeomerous if they are the same kind of their whole.
- separable, if parts are considered to be separable from their wholes if they can be separated by any means of separation e.g. methodological sampling.

Based on these properties, the taxonomy yielded in 6 meronymic relations which we list below, each with its value of the three preceding characteristic properties and a clarifying example from [Winston1987] :

1. Component/Integral Object (F, $\neg H$, S) e.g. pedal-bike

1. https://en.wikipedia.org/wiki/Meronymy_and_holonymy

2. Member/Collection ($\neg F$, $\neg H$, S) e.g. tree-forest
3. Portion/Mass ($\neg F$, H, S) e.g. slice-pie
4. Stuff/Object ($\neg F$, $\neg H$, $\neg S$) e.g. steel-car
5. Feature/Activity (F, $\neg H$, $\neg S$) e.g. paying-shopping
6. Place/Area ($\neg F$, H, $\neg S$) e.g. Everglades-Florida

This proposal significantly contributes to the recognition that there exist various ways in which parts can relate to each other and to the whole they form. However, it is important to note that the focus of the study is primarily on the linguistic term "part-of" and its related terms, without sufficient consideration of the ontological and conceptual adequacy of the proposed distinctions. We examine below two facets of concern regarding this approach.

The first facet pertains to the combinations of values for the proposed characteristic properties. Figure 3.3 illustrates the eight possible combinations of values for the three characteristic properties. Table 3.2 further presents these combinations and the corresponding relations associated with each combination of values, if any. Based on that, we not that two combinations, namely (F, S, H) and (F, $\neg S$, H), do not correspond to any identifiable part-whole relation. The reasons behind the absence of these relations were not discussed, whether due to the fact that no part-whole relation exhibits these specific combinations or that these combinations lack a common-sense semantics in the first place.

The second facet concerns the two characteristic properties : separability and functionality. While separability refers to the physical disconnection of a part from the whole, and functionality relates to the spatio-temporal position of the part with respect to the whole, neither property provides insights into the dependence between the part and the whole. [1] In other words, these properties do not capture the implications of separating a part from the whole on the persistence, identity, or overall functionality of the whole. For instance, consider the example of a pedal, which is a separable component of a bicycle. If the pedal is removed, the bicycle would no longer function properly, yet it would still retain its identity as a bicycle (i.e., the case of an assembled whole).

A common sense theory of part-whole relations 1995, 1996

In their attempt to enhance the original WCH taxonomy, Gerstl and Pribbenow [Gerstl1995, Gerstl1996] concentrate on the role of specific well-defined parts in contributing to the overall functionality of the whole, a concept that has been highlighted in [Cruse1979]. Their approach also seeks to complement mereology by considering the distinct roles played by two different entities in relation to an entity that they both constitute parts of.

In their theory of part-whole relations based on common-sense understanding, the authors propose a classification system that encompasses various ontological categories, including physical objects, temporal and spatial entities, and abstract entities. This classification distinguishes between part-whole relations that arise from the compositional structure of the whole entity (such as uniform, homogeneous, and heterogeneous) and those that are independent of the compositional structure (arising from intrinsic features like external partitioning or partitioning based on properties). The former category encompasses three part-whole relations : Collection/Element, Mass/Quantity, and Complex/Component, while the latter category includes two : segments of wholes and portions of wholes.

1. An examination of the notion of dependence will be provided in Section 3.3.4. Additionally, Chapter 5 offers a thorough re-examination of this concept.

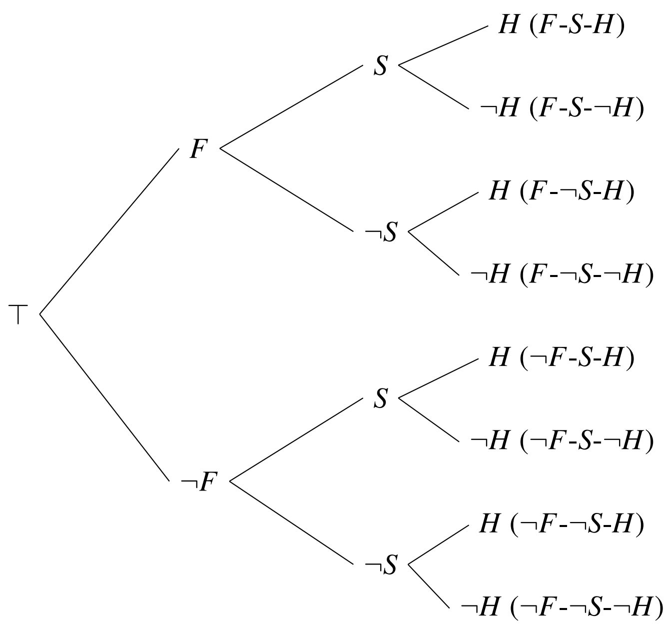


FIGURE 3.3 – *A binary-tree based criterion to present all the possible combinations of the values of the three characteristic properties.*

Combination	Corresponding Relation
F, S, H	—
$F, S, \neg H$	Component-Integral object
$F, \neg S, H$	—
$F, \neg S, \neg H$	Feature-Activity
$\neg F, S, H$	Portion-Mass
$\neg F, S, \neg H$	Member-Collection
$\neg F, \neg S, H$	Place-Area
$\neg F, \neg S, \neg H$	tuff-Object

TABLEAU 3.2 – *The 8 possible combinations of the values of the characteristic properties and the corresponding part-whole relations, if any.*

- Part-whole relations induced by the compositional structure of the whole (dependent on the compositional structure of the whole). Independent of its categorical status if it is a physical object, a situation, or an abstract entity, it is viewed as having parts by means of a potentially inherent compositional structure. Then depending on what aspects of this structure one focuses on, different types of compositional part-whole relations arise i.e. depending on the type of the whole. In order for proper distinction, it critically depends on the level of granularity which is assumed for classifying the entity. So here we have 3 types of compositional wholes and the corresponding 3 part-whole relations :
  - Uniform compositional structure of the whole (Collection-Element). Examples include ; two of the three apples in the basket, one of the three visits. The whole (collection) is an integral whole (i.e. the entities that form this mereological sum are considered to be compatible/belonging to each other). The part (element) can be ; (i) non-atomic, if its primary view is again another whole (collection) and corresponding to the set-theoretic notions of set inclusions ; or (ii) atomic, if its primary view is not another whole (collection) and corresponding to a membership view.
  - Heterogeneous compositional structure of the whole (Complex-Component) e.g. the engine of the car. If the parts are distinguished on the basis of their spatiotemporal arrangement with respect to the whole, then the entity can be viewed as a heterogeneously structured complex comprising different sorts of components. The compositional structure of the whole is based on different relations of the type complex-component depending on the contribution of the part in the function of the whole. This contribution may presuppose a specific spatiotemporal arrangement between the part and the whole.
  - Homogeneous compositional structure of the whole (Mass-Quantity) such as an amount of rice. The whole (mass) is assumed not to have any compositional structure. It may be separated into quantities by applying a certain kind of quantitative measure. The part (quantity) is characterized by a quantitative measure as an arbitrary piece. The relation can be represented by a pair (A, B) where B is the whole and A is a quantitative measure of B, which in turn is represented as a pair (D,N) consisting of a dimension D and a numerical value N (can be unspecified). The main difference between the mass/-quantity relation and the preceding two is that a mass accounts to amounts of matter e.g. amounts of water, sugar, sand, while a component or a member accounts for an object entity e.g. the engine of the car or an apple of the three.

- Part-whole relations that are independent of the compositional structure of the whole. Some partitions are processed independent of the compositional structure of the whole (whether being heterogeneous, uniformly structured or a homogeneous mass), they are instead induced by intrinsic features like external schemes or properties. This is the case where partitions are “segments” or “portions”, and no distinguishing between different kinds of wholes is made. Segments differ from portions by the extent of partitioning whether it is an inherent-to-the-whole partitioning or external partitioning. So here we have two types of parts, and their corresponding part-whole relations
  - External partitions called segments, based on external schemes i.e. spatial. For the whole, it presupposes certain attributes to its internal structure : to be of one-dimensional boundness. Note that this attribute is provided by some masses. For the segments, they may have vogue boundaries with respect to one or more dimensions. In such cases, these vague boundaries may conceptually coincide with some other “part” induced by the compositional structure of the whole. Example : the upper part of the house (segment) – the roof of the house (component of complex). In such cases, the boundary of the segment gets anchored in the boundary of the component that it coincides with.

Examples of external schemes : the topological scheme of segmenting an entity into exterior/boundary/interior parts. These external schemes are spatial and can only be applied on spatial entities or those that can be projected on a spatial dimension. They can be applied to any one-dimensional entity (e.g. street, rope, queue of people) or to an entity after being projected to a one-dimensional entity (e.g. story, cinema play, career. . .).

— Dimension partitions called portions, based on property dimensions to select parts. Portions are those parts of the whole that provide the requested value of relevant dimension property. Examples of property dimensions upon which partitioning into portions can be made : color dimension (the red parts of the painting), dimension of valuation (the scary parts of the movie), a combination of properties (people of the population who are females, workers, and over fifty). Portions are not necessarily connected even if their whole is connected. Also, as in segments, it is possible for portions to coincide with parts depending on the compositional structure of the whole.

In their later work [Gerstl1996], two applications of the theory are presented ; in natural language semantics and for modeling parts of physical objects.

In comparison to the WCH taxonomy, the authors propose a classification that extends beyond linguistic problems. They consider the relevance of partitioning into parts for various domains such as visual perception, object partonomies, and languages. The theory also encompasses the combination of relations, the inheritance of properties between parts and wholes, and goes beyond the transitivity of the relation.

An additional notable difference is that the classification of part-whole relations in the proposed approach is based on the study of the compositional structure of the whole, rather than relying on linguistic cognates of part-whole relations. In view of this goal, some points of critique can be raised :

— The distinction between the complex/component relation and the feature/activity relation may not be necessary. According to [Guizzardi2005], both relations involve a whole where the parts play a role based on their spatial or temporal positions in the internal structure. The functionality property applies to both relations in the WCH taxonomy, with the only difference being the type of entities involved : tangible entities (endurants) for complex/-component relations and intangible entities (perdurants) for feature/activity relations.

— The stuff/object relation is considered ambiguous as it combines two different relations : constitution and another part-of relation. In the WCH taxonomy, the example "a bike is partly steel" is classified as a stuff/object relationship. However, the authors argue that there are actually two distinct relations in this example : (1) a part-whole relation indicating the presence of the skeleton as a part of the bike, establishing a component/integral whole relationship, and (2) a constitution relation where steel constitutes the part (the skeleton).

— Similar to the WCH taxonomy, the approach is not ontological in which no logical formalization of the theory is considered.

3.3.3 Other approaches/theories formalizing PWR

As a compromise between the ontological aspect of parthood in mereology and the consideration of part-whole typologies in meronomy, other approaches have emerged that blend elements from both perspectives. These approaches offer taxonomies of ((part-whole)) relation types, allowing for intransitive in some cases [1], within formalized theories using either conceptual modeling languages (such as UML, ER, ORM) or knowledge representation languages (such as FOL, DL).

Given our focus on foundational ontological relations, we are particularly interested in the latter category, which includes knowledge representation and reasoning languages, as they are more suitable for ontology modeling. In the following sections, we provide a brief overview of some notable work within the realm of conceptual modeling, followed by a more detailed presentation of well-known works within knowledge representation and reasoning languages.

Using conceptual modeling languages :

In the context of conceptual modeling, the term "aggregation" is commonly used and often resembles a form of part-whole relation. However, it is important to note that aggregation differs from a typical part-whole relation, as discussed earlier in contrast to the strong supplementation axiom of mereology (Pa5).

Specifically, in UML [Force2010], aggregation is defined in two forms : composite and shared. A composite aggregation is a transitive asymmetric relationship, represented by a black filled diamond on the side of the whole in UML diagrams. It signifies a strong form of aggregation that requires the presence of at least one part instance for the existence of the whole. Additionally, if the composite (the whole) is deleted, the associated parts are also deleted or removed. For example, in a class diagram, the class "player" may be linked to the class "team" through a composite aggregation relation. However, based on the discussion in [Keet2006a], some ambiguity around composite aggregation arises from its binary nature, which means it can only represent a whole composed of one specific type of part. To represent a whole that can be built from different types of parts, multiple aggregation associations need to be used, resulting in potentially ambiguous semantics between the parts and the whole.

On the other hand, shared aggregation, depicted by an empty diamond on the side of the whole in UML diagrams, represents a more general form of part-whole relation without imposing constraints on the part and whole entities. Unlike composite aggregation, shared aggregation allows a part to be shared by multiple wholes simultaneously.

While no formal semantics are provided within UML, and with the aim to (a) clarify the ambiguities of the two aggregation relations, and (b) account for semantically richer part-whole relations, some researchers proposed extensions to the UML's aggregation e.g. [Motschnig-Pitrik1999, Barbier2003, Berardi2005, Shanks2004]. For instance, in [Barbier2003], the authors propose a formal definition for part-whole relation in UML i.e. for aggregation associations to be incorporated in the version 2.0 of UML, to incorporate reasoning capabilities behind part-whole relations. The formalization is expressed in the Object Constraint Language (OCL) [Warmer1998], a textual language that is part of UML 1.1 for expressing constraints that cannot be shown in UML diagrams.

Using knowledge representation languages :

Despite the fact that the part-whole relation has not been widely embraced as a fundamental modeling primitive in Semantic web languages, several authors have acknowledged its significance for reasoning in description logics. For instance, the relevance of the part-whole relation has

1. Please refer to [Cruse1979] and [Varzi2006] for discussions on the transitivity of the part-whole relation.

been highlighted in works such as [Artale1996a, Lambrix2000, Sattler2000, Bittner2005].

Bittner and Donnelly, 2005 : In [Bittner2005], the authors focus on the *proper-part-of* relation as the main parthood relation. They introduce a theory that encompasses parthood, componenthood, and containment relations, using the relations *proper-part-of*, *contained-in*, and *component-of*, respectively. The distinctions between these relations are based on specific properties, such as the ability to relate to a single part for containment and componenthood, while parthood lacks this property. However, all three relations share common algebraic properties like transitivity and asymmetry.

To explicate the semantics of these relations and their corresponding properties, the use R-structures $(\triangle, R)$: a structures consisting of a non-empty domain $(\triangle)$ resembling the entity types, and a binary relation (R) for denoting relationships and holding between entities whose types are indicated in $(\triangle)$. They present an ontological theory for the three relations in both FOL and DL. FOL demonstrates expressive power in distinguishing properties between the relations, while the DL language offers less expressive power for some properties but proves to be relatively easier to use and suitable for reasoning tools.

Based on their findings, the authors propose a computational ontology comprising two complementary parts. The DL-based ontology enables automatic reasoning and restricts the meaning in the most concise manner. On the other hand, the FOL-based ontology serves as a knowledge base for the relations and makes explicit the properties that cannot be expressed in DL. This work indeed highlighted the importance of two-folded complementary formalization supporting both expressiveness and decidability.

Guizzardi, 2005 : One notable contribution in the field of part-whole relations studies is the approach presented by Guizzardi in his thesis [Guizzardi2005]. This work can be seen as an advancement building upon several previous works, all from the perspective of conceptual modeling. These include :

- A comprehensive summary of Varzi's ontological study [Varzi2003] from the standpoint of conceptual modeling, providing a concise overview of mereology.
- An enhancement of the conceptual study conducted by Gerstl and Pribbenow in [Gerstl1995], which introduced "3 types of conceptual parthood".
- An extension of UML's treatment of the part-whole relation (referred to as aggregation), achieved through the proposal of a first-order logic (FOL) formalization and a graphical notation.

To address the issues of mereology, Guizzardi proposed an extension of the theory of parthood by introducing the concept of "Integral Wholes", as a complementary theory of wholes, in his thesis [Guizzardi2005] (Section 5.3). This extension is based on the ontological distinction made by Simons [Simons1987], which considers the existence conditions of entities. While mere sums exist whenever the parts exist, integral wholes require additional conditions, such as a unifying condition, to exist as a cohesive whole. Guizzardi defines a parthood relation, denoted as A-parthood, which signifies the acquisition of an integral whole and is defined in terms of a relationship that unifies its parts.

In order to formally characterize part-whole relations and their ontological distinctions, Guizzardi proposes axioms using modal logic. These axioms are based on the analysis of secondary characteristics of relations presented in Opdahl's work [Opdahl2001]. The examined properties include shareability and separability (presented in Section 5.4 of his thesis), which allow for the classification of parts into essential and mandatory parts based on their dependence properties. Essential parts exhibit specific dependence, while mandatory parts exhibit generic dependence. It is important to note that Guizzardi's notion of separability differs from the one discussed in the WCH taxonomy [Winston1987]. The former is based on ontological dependence, while the latter is based on physical disconnection (though not formalized).

Furthermore, Guizzardi proposes an ontological theory of conceptual part-whole relations that aims to provide formal characterization and common-sense semantics. This theory expands the literature on mereology by incorporating meronomy and distinguishes four types of relations based on the ontological entities involved : quantities (sub-quantities), collections (sub-collections), functional complexes (member-of), and integral wholes (parthood). This first dimension of distinction is based n the ontological entities. At a second dimension, meta-properties such as separability and shareability, along with their corresponding cognates like dependence, are introduced to create typologies of relations. It is important to note that the two dimensions "ontological types" and "meta-properties" are not completely orthogonal, as the ontological type of a part-whole relation implies certain meta-properties.

Guizzardi's noteworthy contribution presented a significant approach that incorporated a complementary theory of integral wholes in conjunction with the theory of the four basic relations. This work extended to the mere development of a formal ontology, as its primary objective was to support conceptual modeling tasks, particularly within the context of UML. By providing an ontological tool comprising relations, ontological categories, and roles, Guizzardi aimed to facilitate ontology-driven conceptual modeling. The resulting ontology offers a package of ontological choices for users to adopt in their conceptual modeling tasks rather than a theory of structural part-whole relations. Additionally, it is important to shed the light on some formalization choices. While modal logic provides a rigorous foundation, it may introduce complexities and make the theory less accessible to those who are not well-versed in modal logic. The reliance on such formalism could limit the practical applicability and comprehensibility of the theory for conceptual modelers.

Keet and Artale, 2008 : In their work, Keet and Artale [Keet2008] developed an approach that builds upon Keet's earlier work in ORM [Keet2006b] and aims to formalize it using FOL. The objective is to assist conceptual modelers in selecting the appropriate part-whole relations, resulting in a taxonomy of meronymic and mereological part-whole relations.

The taxonomy differentiates between mereological and meronymic part-of relations, based on the choice between transitive part-whole relations (mereology) and non-transitive relations that can be either transitive or intransitive (meronomy). This distinction is illustrated in Figure 3.4, where transitive mereological part-whole relations are denoted using *part-of*, and non-transitive meronymic part-whole relations are denoted using *mpartof*.

Furthermore, within the categories of *part-of* and *mpart-of*, the taxonomy specializes several types of part-whole relations based on the categories of the entities involved. This second level of distinction relies on concepts from the DOLCE foundational ontology [Masolo2003].

The resulting leaf relations in the taxonomy, such as *member-of*, *constitutes*, *sub-quantity-of*, *participates*, *involved-in*, *contained-in*, and *located-in*, are formally defined in terms of their parent relation (either *part-of* or *mpart-of*) and the types of domain and range entities, which are categorized according to DOLCE.

In a subsequent work [Keet2012], the taxonomy was extended to include some mereotopological relations (specifically, *PP* and *EQ*) within a DL-formalized ontology. This extension aimed to overcome limitations in expressing the semantics of these relations in OWL.

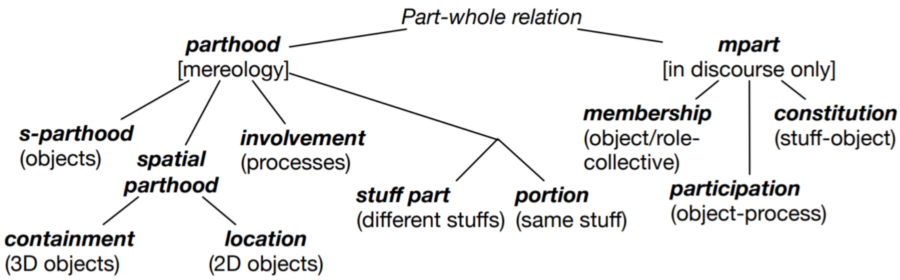


FIGURE 3.4 – *Keet and Artale's taxonomy of meronymic and mereological part-whole relations.*

While it is commendable to specialize ((part-whole)) relations to account for different types and consider intransitivity in certain cases, we have some concerns regarding Keet's taxonomy, particularly with its second level of distinction that utilizes DOLCE categories for further classification of sub-types of *part-of* and *mpart-of*. It should be noted that DOLCE is a foundational ontology that presents a formal and well-defined axiomatic system of categories and relations.

When DOLCE is employed to annotate the relata (i.e. the domain and range) of the relation, it essentially involves projecting the axioms of DOLCE's classes onto the domain and range entities of the relation. Consequently, the use of DOLCE's axiomatization to define these relations seems to rely heavily on leveraging the semantics already provided by axiomatized classes. As a result, apart from the naming of the relations, the taxonomy does not appear to offer significant additional value beyond the exploitation of existing semantic associations derived from the axiomatized classes of DOLCE.

3.3.4 Some properties of part-whole relations discussed in the literature

Granularity :

Some approaches have taken the steps to relation mereology to granularity ; a concept which involves organizing information a hierarchical manner, guided by specific criteria known as the granular perspective. It involves the differentiation of levels, where lower levels encompass more detailed knowledge or data, while higher levels simplify or generalize finer details. Each granular level, also referred to as grain size, consists of one or more entities or instances. Examples of these approaches include [Bittner2001, Bittner2003, Rector2006a, Keet2006c] in which authors use mereology as the theory for formalizing granularity, partition entities (wholes) based on granularity, and propose a taxonomy of types of granularity and discussing how entities in each granularity level relate.

Ontological dependence :

The exploration of the concept of dependence traces back to 1970 with Husserl in the realm of classical philosophy. In this context, a dependence relation is established between two individuals, wherein x depends on y if the existence of x (necessarily) implies the existence of y. The formalization of dependence has been approached in various ways, including as a primitive relation i.e. not defined in terms of any other relation, but through axioms, such as in [Fine1983], or as a non-primitive relation i.e. defined in terms of another primitive, often the existence or persistence primitive, such as in [Simons1987]. Moreover, modal logic has been used in some formalizations later to capture the notion of necessity in dependence for part-whole relations [Vieu2007, Masolo2003]. For a comprehensive overview of dependence and its formalization with/without modal logic, we refer the reader to [Tahko2020].

Furthermore, within the context of mereology, dependence introduces the concept of "mereological essentialism" ([Chisholm1975, Plantinga1975] among others) which allows for the definition of essential and mandatory parts. Formal ontology employs mereological essentialism to characterize and formalize the essentiality of parts to their wholes, using ontological dependence. Many authors make use of the distinction between essential (e.g., the brain or the heart as parts of a human) and non-essential (e.g., a single hair as a part of a human) to introduce the notions of essential parts and wholes. For instance, in [Guizzardi2005] as mentioned earlier, these terms are linked to separability, which is defined in terms of dependence.

Functionality :

The concept of functionality was initially introduced in [Winston1987] without a formal framework, as an inherent characteristic of binary part-whole relations. It is denoted by F/¬F, indicating whether parts are in a specific spatial/temporal arrangement with respect to each other, thus supporting their functional role within the whole. TIt applies to parts of complex objects, wherein the spatial/temporal positioning of these parts triggers the overall function of the whole, as stated in [Winston1987].

Subsequently, various authors have explored the notion of functionality within the context of part-whole relations. For instance, Vieu and Aurnague presented a theory of functional dependence in [Vieu2007], while Johansson and Garbacz provided formal frameworks for defining functional parthood in [Johansson2004, Garbacz2007]. In the following section, we provide a brief explanation of Vieu and Aurnague's approach, which served as the foundation for Guizzardi's work in identifying functional dependence through conceptual part-whole relations.

In [Vieu2007], these specific parts of complex wholes are referred to as functional components, such as the relationship between an organ and a body, or an engine and a car. Here, the functional link between a part and a whole in part-whole relations encompasses not only a functional property but also a dependence property, indicating the interdependency between the part and the whole. Consequently, the authors dedicated their work to analyzing, characterizing, and formalizing this functional link within part-whole relations, which they referred to as "component-integral whole."

In their research, functionality is treated as a generic term denoted by "F" to establish its semantics within the context of the component-integral whole relation. Additionally, functional dependence is recognized as a special type of generic dependence that occurs at two levels : generic and individual.
Generic functional dependence (*GFD(X,Y)*) represents the functional link between two entity types (X and Y) as lexical categories. It can manifest in both directions as *GFD(X,Y)* and/or *GFD(Y,X)*. Whereas individual functional dependence, in its direct form, pertains to the functional link between two specific entities x and y (*IFD(x,y)*), where x and y belong to types X and Y, respectively. It should be noted that this functional dependence does not necessarily align with the *GFD* of their class types. In other words, whenever x is functioning, y also participates in the functionality.
As for the indirect form individual functional dependence (*IIFD(x,y)*), it represents the functional link between x (e.g., a handle) and y (e.g., a door). In this case, for x to fulfill its function, it does not specifically require y as an instance, but rather any entity that possesses a role enabling the functionality of x (e.g., any object that can be manipulated or used by hand, such as a knife or a bag).

Based on these considerations, four scenarios of functional dependence between a part (x) and a whole (y) are identified : *IFD(x,y)*), *IIFD(x,y)*), *IFD(y,x)*), and *IIFD(y,x)*). Therefore, the authors advocate for a notion of functional dependence that is not tied solely to essential parts or wholes. In other words, parts and wholes can exist independently but still maintain a functional dependence, rather than an existential dependence.

However, the term "function" in computer science poses philosophical challenges, as it encompasses various aspects that are difficult to simplify. The exploration of functionality in artifacts originated from the analysis of "function" in engineering, leading to diverse definitions proposed in the literature. These definitions range from identifying function based on the nature of the artifact versus its behavior [De Kleer1984], to considering the intentions, decisions, and actions of the artifact's creator [Dipert1993], to differentiating between the notions of "function" and "creator's intentionality" (despite both categorizing artifacts) [Bloom1996], to linking the term to its contextual dependence in applications [Kumar1998] or its independence from context [Roy2001], among other perspectives.

Therefore, opting for a neutral and generic notion of function does not provide a definitive solution. Rather, what is needed is a clear identification of its semantics within formal ontology, enabling precise applications without the requirement of achieving a consensus on the term itself.

3.4 Studies on spatial (part-whole) relations

In the realm of spatial relations, a multitude of relations fall under the classification of spatial, including topology which broadly understood as a theory of qualitative spatial relations such as continuity and contiguity [Varzi2007]. The Region Connection Calculus (RCC) [Randell1992, Cui1993] is intended for qualitative spatial representation and reasoning, based on the "connection" relation, denoted as *C*. For instance, **RCC8** is a family of the RCC calculus serving as a spatial logic represent and reason about topological/spatial relationships among entities in spatial databases. It is based on the primitive *C* upon which it defines 8 topological relations.

However, as explicitly outlined in 3.2, our specific focus pertains to spatial relations that allow for the representation of spatial configurations within entities or of entities themselves. This specific criterion narrows down the scope of spatial relations investigated to those capable of performing part-whole representations, as well as location relations, which will be explored in detail in Section 3.4.1 and Section 3.4.2 respectively.

3.4.1 From mereology and topology, to mereotopology

The issues with a purely mereological theory in adequately addressing the properties of both parts and wholes have been discussed earlier. More specifically, several limitation have been addressed in [Varzi1996] regarding mereology's need for the global properties of wholeness. Examples such as the relationship between an entity and its surface or the proximity of one entity to another demonstrate the inadequacy of pure mereology in capturing fundamental spatial relations. Consequently, the incorporation of complementary topological analysis becomes necessary to characterize entities and the spatial relations that exist among them.

To address some tasks of spatial representation and reasoning, three main strategies have been discussed in [Varzi1993], aiming to combine mereology (the theory of parthood), with hololog (the theory of wholeness) which is provided by topology. The first strategy considers mereology and topology as independent theories, the second regards mereology as the overarching theory subsuming topology, and the third treats topology as the general theory subsuming mereology.

Through this discussion, the following outcome emerges : by employing P as the primitive predicate of mereology, an additional predicate "C" is introduced, intuitively understood as the relation of topological "connection," following the suggestion of certain authors who propose "C" as a join relation [Whitehead1925]. Thus, while having a mereological system based on parthood and a topological system based on connection, the question at hand is how to expand mereology into a more comprehensive part-whole theory. More precisely, it is to explore the interaction between the parthood-based mereological system and the connection-based topological system.

In this context, we begin by presenting Ground Topology (T) as the theory of connection, followed by the discussion regarding the integration of the two theories of mereology and topology (M+T). Subsequently, we illustrate mereotopology and some of its extensions (MT).

Ground Topology (T)

For a topological theory, the reflexive axiom (Ca1) and the symmetric axiom (Ca2) make up the proper/sufficient axioms of the connection relation C. The basic theory defined by the minimal axioms (Ca1) and (Ca2) is referred to as *Ground Topology* **(T)** [Varzi1996], in analogy to the theory of parthood **(M)**.

$$(\forall x, y) C(x, x) \tag{Ca1}$$

$$(\forall x, y)(C(x, y) \rightarrow C(y, x)) \tag{Ca2}$$

The integration of topology and mereotopology (M+T)

T is considered to be extremely weak, for that a model of **T** can be obtained simply by interpreting C as mereological overlap O (Pd3). And so for a combination of **(T)** and **(M)**, further principles should be added so as to distinguish C from O. Indeed, it is of no interest to simple add Ca1 and Ca2 of **(T)** to Pa1, Pa2 and Pa3 of **(M)** unless one also adds some new principle bridging **M** and **T** [Casati1999].

According to the explanation provided in [Varzi2007], most theories, if not all, adhere to a bridging principle that revolves around the fundamental notion that, regardless of the comprehensive characterization of P and C (both are fully characterized), they must be related in a manner that ensures a strong connection between a whole and its constituent parts. In order to capture this intuition, three distinct approaches have been proposed, each corresponding to a specific axiom : integrity, unity, and monotonicity.

$$(\forall x, y) P(x, y) \rightarrow C(x, y) \tag{Integrity axiom}$$

$$(\forall x, y) O(x, y) \rightarrow C(x, y) \tag{Unity axiom}$$

$$(\forall x, y) P(x, y) \rightarrow E(x, y) \tag{Monotonicity axiom}$$

$$(\forall x, y) E(x, y) \leftrightarrow \forall z (C(z, x) \rightarrow C(z, y)) \tag{Monotonicity}$$

The first principle (Integrity axiom) indicating that everything must be connected to its parts. However it is considered weak for that it does not even capture that if something is part of two things, then those two things are connected *because* of that common part (not *by* that common part).

The second principle (Unity axiom) is stronger than the first since parthood is a sub-relation of overlap (it implies overlap based on Pd3). But is still weak to capture the intuition identified above, for that it is true that it guarantees that overlapping a part is sufficient for being connected to the whole, it doesn't however secure that touching a part (without being actually sharing parts) is also sufficient. For example, something can touch a bottle just by touching its cap, without sharing parts of the bottle.

Hence, it is only with the third principle (Monotonicity axiom), that a plausible formulation of the basic idea is achieved. Indeed, if connection is to behave properly, it must be monotonic with respect to parthood, meaning if an entity x is part of another y, then whatever is connected to x is connected to y i.e. whatever is connected to a part is also connected to the whole. In addition, it is easily checked that the third principle (Monotonicity axiom) implies the second (Unity axiom), hence the first (Integrity axiom).

Mereotopology (MT)

Therefore, adding the monotonicity axiom (Ca3) [1] , bridges **M** and **T** yielding in Minimal mereoTopology (**MT**) (or ground mereotopology) as the unified theory combining mereology and the global properties of wholeness of topology [Varzi1996, **?**].

$$(\forall x,y)P(x,y) \rightarrow \forall z(C(z,x) \rightarrow C(z,y)) \tag{Ca3}$$

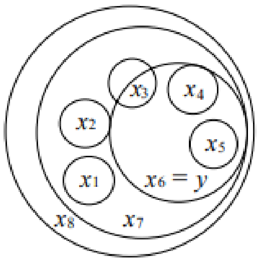
In the framework of **MT**, a diverse range of relations can be established, enabling the representation of connected entities that do not share parts, referred to as external connection (Cd1). Additionally, **MT** incorporates various spatial sub-relations within the scope of P, including tangential and interior parts (Cd2-Cd3), within PP such as tangential and interior overlap (Cd4-Cd5), within O such as tangential and interior overlap (Cd6-Cd7), and within U such as tangential and interior underlap (Cd8-Cd9). These basic merotopological relations contribute to a more comprehensive representation of the spatial relationships between entities, as depicted in Figure 3.5.

$$(\forall x,y)EC(x,y) \leftrightarrow C(x,y) \wedge \neg O(x,y) \tag{Cd1}$$
$$(\forall x,y)TP(x,y) \leftrightarrow P(x,y) \wedge \exists z(EC(z,x) \wedge EC(z,y)) \tag{Cd2}$$
$$(\forall x,y)IP(x,y) \leftrightarrow P(x,y) \wedge \neg TP(x,y) \tag{Cd3}$$
$$(\forall x,y)IPP(x,y) \leftrightarrow PP(x,y) \wedge \forall z(C(z,x) \rightarrow O(z,y)) \tag{Cd4}$$
$$(\forall x,y)TPP(x,y) \leftrightarrow PP(x,y) \wedge \neg IPP(x,y) \tag{Cd5}$$
$$(\forall x,y)IO(x,y) \leftrightarrow \exists z(IP(z,x) \wedge IP(z,y)) \tag{Cd6}$$
$$(\forall x,y)TO(x,y) \leftrightarrow O(x,y) \wedge \neg IO(x,y) \tag{Cd7}$$
$$(\forall x,y)IU(x,y) \leftrightarrow \exists z(IP(x,z) \wedge IP(y,z)) \tag{Cd8}$$
$$(\forall x,y)TU(x,y) \leftrightarrow U(x,y) \wedge \neg IU(x,y) \tag{Cd9}$$

Similar to the extensional theories of **M**, other extensional mereotopological theories can be constructed based on the foundation of ground mereotopology (**MT**), by incorporating specific axioms. The various extensions are summarized in Figure 3.6.

It shows that the addition of the spatial enclosure axiom (Ca3) to **T** yields the theory of Minimal mereoTopology (**MT**). Further expansion is achieved by combining **MT** with **GEM** and incorporating additional axioms, namely the self-connected axiom (Ca4) i.e. a self-connected whole

1. Note that we continue to label the definitions and axioms within merotopology using the denotation "C", knowing that these definitions and axioms are not purely topological.



	Ux_ny	ECx_ny	Ox_ny	$TPPx_ny$	$IPPx_ny$	EQx_ny	$TPEx_ny$	$EPEx_ny$
$n = 1$	+	–	–	–	–	–	–	–
$n = 2$	+	+	–	–	–	–	–	–
$n = 3$	+	–	+	–	–	–	–	–
$n = 4$	+	–	+	+	–	–	–	–
$n = 5$	+	–	+	–	+	–	–	–
$n = 6$	+	–	+	–	–	+	–	–
$n = 7$	+	–	+	–	–	–	+	–
$n = 8$	+	–	+	–	–	–	–	+

FIGURE 3.5 – *Figure taken from [Varzi2007], showing the basic mereotopological relations, with shaded cells indicate connection and darker shading tands for parthood.*

such as a mug and its handle part in as opposed to a scattered entity, the converse-monotonicity axiom (Ca5) for allowing dependence between connection and parthood, the bridging connection with part axiom (Ca6), and the fusion axiom (Ca7). This amalgamation results in the theory of General Extensional Mereotopology (**GEMT**).

Moreover, **GEMT** enables the definition of the *interior-proper-part* relation $\mathsf{IPP}(x,y)$ (Cd10), which subsequently allows for the definition of the *tangential-proper-part* relation $\mathsf{TPP}(x,y)$ (Cd11).

$$(\forall x)SC(x) \leftrightarrow (\forall y,z)(x = y+z \rightarrow C(y,z)) \quad \text{(Ca4)}$$

$$(\forall x,y)(E(x,y) \rightarrow P(x,y)) \quad \text{(Ca5)}$$

$$(\forall x,y)\exists z(SC(z) \wedge O(z,x) \wedge O(z,y) \wedge \forall w(P(w,z) \rightarrow (O(w,z) \vee O(w,y)))) \rightarrow C(x,y) \quad \text{(Ca6)}$$

$$(\forall z)z = \Sigma x\phi x \rightarrow \forall y(C(y,z) \rightarrow \exists x(\phi x \wedge C(y,x))) \quad \text{(Ca7)}$$

$$(\forall x,y)IPP(x,y) =_{df} PP(x,y) \wedge \forall z(C(z,x) \rightarrow O(z,y)) \quad \text{(Cd10)}$$

$$(\forall x,y)TPP(x,y) =_{df} PP(x,y) \wedge \neg IPP(x,y) \quad \text{(Cd11)}$$

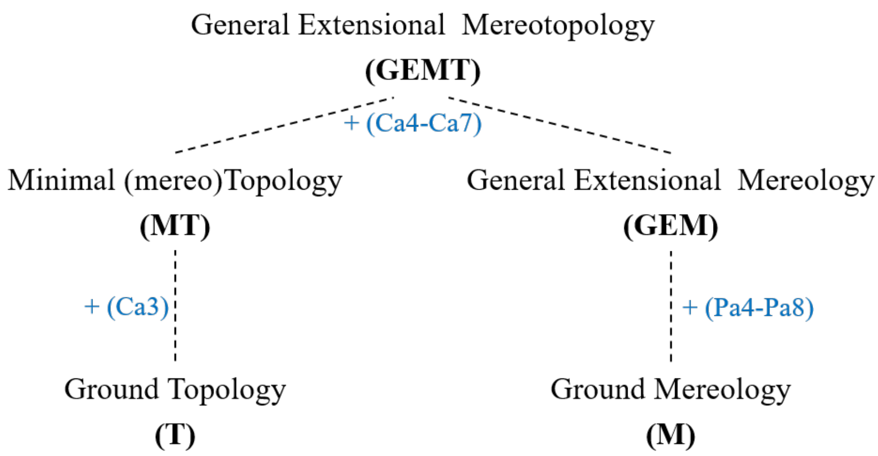


FIGURE 3.6 – *Hasse's diagram for mereotopological theories (from weaker to stronger).*

For further insights and comprehensive summaries of mereotopological theories, readers may refer to [Varzi2007]. Additionally, [Keet2017] presents a comprehensive exploration of the network of mereotopological theories across multiple formal frameworks.

3.4.2 Casati and Varzi's spatial location theory

The theory of location proper focuses specifically on the relationship between an entity and the spatial region it occupies [Gilmore2018]. This theory not only addresses the mereological aspects of an object but also considers the interconnection between the mereology of the object and its spatial location.

Furthermore, as highlighted in [Parsons2007], various perspectives emerge regarding the formulation of the location predicate and its potential extensions. These viewpoints encompass statements such as "*x* is exactly located at *y*," "*x* is weakly located at *y*," and "*x* is entirely located at *y*." These alternative perspectives prompt discussions on the precise framing of the location predicate and the nature of the relationship between an entity and its spatial location.

To establish a framework for defining the axioms governing the location relation and its interaction with parthood and other mereotopological relations, Casti and Varzi have proposed logics of location [Casati1999]. These logics aim to capture and represent the systematic connection between two fundamental aspects : (a) the mereological properties and relations among located entities and (b) the mereological properties and relations among the corresponding spatial locations of these entities. In essence, the goal is to capture the inherent relationship whereby aspect (a) must align with aspect (b).

In [Casati1999], the authors introduce a theory that treats the location relation as an independent primitive alongside parthood and connection. The formalization of this theory is designed to remain neutral regarding the ontological status of the entities involved in the relation. It does not aim to establish categorical types for the domain and range entities.

Using the primitive L, a location relation captures the intuition *being (exactly) in a place*. $L(x,y)$ takes place between an entity x that is not a spatial region and y that is the spatial region at which x is located. L is formalized as (1) conditionally reflexive (La1), and not reflexive, i.e. reflexivity hold only on the region entities, and (2) functional (La2) for ensuring that it represents exact location and not a any notion of minimal address location, in the sense that an entity can have one exact location at a specific time.

$$(\forall x,y)L(x,y) \rightarrow L(y,y) \tag{La1}$$
$$(\forall x,y,z)(L(x,y) \wedge L(x,z) \rightarrow y = z) \tag{La2}$$

By La2, it is implied that no distinct region can be exactly co-located (Lt1). And with (La1), it is guaranteed that L is antisymmetric (Lt2) and transitive (Lt3).

$$(\forall x,y,z,w)(L(x,y) \wedge L(z,w) \wedge L(y,w) \rightarrow y = w) \tag{Lt1}$$
$$(\forall x,y)(L(x,y) \wedge L(y,x) \rightarrow x = y) \tag{Lt2}$$
$$(\forall x,y,z)(L(x,y) \wedge L(y,z) \rightarrow L(x,z)) \tag{Lt3}$$

Moreover, using L, other locative relations can be built to include cases of in exact location : *partial location* (PL) and *whole location* (WL), defined using parthood in Ld1 and Ld4 respectively. Additional predicated of each can then be pictured with the help of mereotopology (namely the tangential Cd2 and interior Cd3 parts) such as *tangential partial location* (TPL) Ld2, *interior partial location* (IPL) Ld3, *tangential whole location* (TWL) Ld5, and *interior whole location* (IWL) Ld6. Upon the addition of these predicates, it is supposed that exact location L is a the case in which an entity x is both partially and wholly located in y La3.

$$(\forall x,y)PL(x,y) =_{df} \exists z(P(z,x) \wedge L(z,y)) \tag{Ld1}$$
$$(\forall x,y)TPL(x,y) =_{df} \exists z(TP(z,x) \wedge L(z,y)) \tag{Ld2}$$
$$(\forall x,y)IPL(x,y) =_{df} \exists z(IP(z,x) \wedge L(z,y)) \tag{Ld3}$$
$$(\forall x,y)WL(x,y) =_{df} \exists z(P(z,y) \wedge L(x,z)) \tag{Ld4}$$
$$(\forall x,y)TWL(x,y) =_{df} \exists z(TP(z,y) \wedge L(x,z)) \tag{Ld5}$$
$$(\forall x,y)IWL(x,y) =_{df} \exists z(IP(z,y) \wedge L(x,z)) \tag{Ld6}$$
$$(\forall x,y)L(x,y) \rightarrow PL(x,y) \wedge WL(x,y)) \tag{La3}$$

To guarantee the links with parthood and connection, axioms La4 and La5 are added implying that the exact locations of part and whole acquire themselves a part-whole relationship, and that connected entities have their exact locations connected too, respectively.

$$(\forall x,y,z,w)P(x,y) \wedge L(x,z) \wedge L(y,w) \rightarrow P(z,w) \tag{La4}$$
$$(\forall x,y,z,w)C(x,y) \wedge L(x,z) \wedge L(y,w) \rightarrow C(z,w) \tag{La5}$$

To support reasoning on the systematic links between location and mereoptology, the following are assumed as basic principles :

— (a) reasoning about the location of the mereotopological properties of entities x and y and the location z of y in Lt4, Lt5, and Lt6,

$$(\forall x,y,z)P(x,y) \wedge L(y,z) \rightarrow WL(x,z) \tag{Lt4}$$
$$(\forall x,y,z)TP(x,y) \wedge L(y,z) \rightarrow TWL(x,z) \tag{Lt5}$$
$$(\forall x,y,z)IP(x,y) \wedge L(y,z) \rightarrow IWL(x,z) \tag{Lt6}$$

— (b) reasoning about the location of an entity z, part of y, with respect to the mereotopological properties of of its whole's (y) location denoted x in Lt7, Lt8, and Lt9,

$$(\forall x,y,z)L(x,y) \wedge P(z,y) \rightarrow PL(x,z) \tag{Lt7}$$
$$(\forall x,y,z)L(x,y) \wedge IP(z,y) \rightarrow IPL(x,z) \tag{Lt8}$$
$$(\forall x,y,z)L(x,y) \wedge TP(z,y) \rightarrow TPL(x,z) \tag{Lt9}$$

— and (c) reasoning about the location of the mereotopological properties of an entity x with respect to the mereotopological properties of its location y in Lt10, Lt11, Lt12, Lt13, Lt14, and Lt15.

$$(\forall x,y,z)PL(x,y) \wedge P(z,y) \rightarrow PL(x,z) \tag{Lt10}$$
$$(\forall x,y,z)IPL(x,y) \wedge P(z,y) \rightarrow IPL(x,z) \tag{Lt11}$$
$$(\forall x,y,z)TPL(x,y) \wedge P(z,y) \rightarrow TPL(x,z) \tag{Lt12}$$
$$(\forall x,y,z)WL(x,y) \wedge P(z,y) \rightarrow WL(x,z) \tag{Lt13}$$
$$(\forall x,y,z)IWL(x,y) \wedge P(z,y) \rightarrow IWL(x,z) \tag{Lt14}$$
$$(\forall x,y,z)TWL(x,y) \wedge P(z,y) \rightarrow TWL(x,z) \tag{Lt15}$$

Several researchers have employed Casati and Varzi's location theory to represent and reason about spatiotemporal entities. For instance, in [Bittner2004b], the location relation is employed to characterize the spatiotemporal aspects of endurants and perdurants, examining the connection between a located entity and the region where it is situated, including whether they share common parts.

In other approaches, such as in [Donnelly2006] for anatomical reasoning, the location theory is expanded by introducing a region function R_x as the second argument. This function maps each individual x to the unique spatial region at which it is precisely located during a given time. The interpretation of the locative relation L is twofold : (a) it is linked to mereology, indicating that a locative relation holds between two entities if one is part of or overlaps with the other, and (b) it is independent of mereology, denoting locative relations between partially or wholly coinciding entities that do not share parts.

Furthermore, in [Gangemi2001], an alternative approach is presented, which extends the theory of Casati and Varzi to encompass the relationship between arbitrary entities and four-dimensional regions. However, the term "location" is not employed in this approach, as it does not distinguish between continuants and occurrents. Instead, the relation (the equivalent of Casati and Varzi's L) is renamed as "being extended in a (n-dimensional) region," using the primitive $E(x,y)$. This relation is also functional and conditionally reflexive.

3.5 Final considerations

Aligned with our fundamental objective, we have presented the motivation for this chapter ([motivation-II]), as a complement to [motivation-I]. This addition emphasizes the critical need for a language encompassing foundational ontological relations that can offer primitives for the representation and reasoning over the composition of tangible entities, while highlighting more specifically, structural and spatial relations.

In Section 3.2, we introduced the concept of "foundational ontological relations" and situated the field of study concerned with the characterization and formalization of these relations. We also discussed the various approaches used to investigate these relations, including taxonomies and theories, providing illustrative examples. Following this general overview of foundational ontological relations, we narrowed the focus, as guided by [motivation-II], to concentrate specifically on structural and spatial (part-whole) relations.

In Section 3.3, we have focused on structural ((part-whole)) relations, in which we first presented the study of part-of as a significant ontological relation under different aspects.

First are the formal studies of parthood which are conducted under the theory named "mereology". These studies have provided a rigid formal framework for representing and assessing parthood relations. However, mereology have been shown as either too weak to capture the distinctions that mark different types of ((part-whole)) relations in conceptual modeling and cognitive tasks, or too strong to hold as a generalization of a theory of part-whole relations at a conceptual level.

Second are the studies conducted in the field of cognitive sciences, particularly in linguistics, and fall under the umbrella term "meronomy." These studies have often been carried out within non-formal frameworks, making it challenging to apply the resulting theories effectively in practical assessments of relations.

And last are the efforts that have been made to integrate the strengths of both formal mereology and typologies of relations in meronomy within various formal frameworks, using conceptual modeling and knowledge representation languages. Notably, Guizzardi's work introduced a significant approach that encompasses a complementary theory of integral wholes alongside the theory of the four basic relations, among other aspects within an ontology. Although the ontology encompasses well-formalized and expressive set of relations, however the aim of the approach is focused to support ontology-driven conceptual modeling tasks (mostly within UML) by offering a foundational ontology of relations, ontological categories, and roles. As a result, the foundational ontology proposed is centered in the ontological category of endurants and endurants universals, which was then extended to the Unified Foundational Ontology (UFO), which will be inspected in more details in Chapter 6.

Additionally, we presented some other works that have studied some characteristic properties of these relations such as ontological dependence and functionality.

After that, in Section 3.4, our attention has been directed towards spatial relations, which enable the representation of spatial configurations within entities, specifically the spatial aspects of part-whole studies and the location relations of entities. As a result, we introduced mereotopology as a comprehensive theory that integrates mereology and topology, facilitating reasoning about the spatiotemporal aspects of tangible entities. Furthermore, we exemplified Casati and Varzi's location theory, elucidating the location relation and its connection to mereotopology.

In conclusion, we acknowledge the significance of the various approaches presented in different aspects of studies and emphasize the importance of their collective integration as complementary facets of research. Thus, we recognize the need for a unified theory that encompasses a foundational set of ontological relations, facilitating the structural and spatial representation of entities to describe their composition. These relations would include formal parthood, extension of parthood with additional properties to encompass a broader range of semantic relations, additional primitives beyond parthood that are not considered as cognates of the part-whole relation, as well as spatial relations to represent location and internal spatial aspects within entities. Therefore, our objective is to combine the relevant elements observed in the current state of the art within a well-formalized theory of relations.

An important question arises : if we seek a theory of foundational ontological relations, do we also require ontological categories to characterize and formalize these relations ? Are categories obligatory in this context ? We leave this question unanswered for now and explore it further in the subsequent chapters.

The subsequent Section B focuses on the contributions of this thesis, starting with Chapter 4, wherein we delineate our theoretical approach, outline the objectives we aim to achieve through this approach, and expound upon the methodology employed to fulfill these objectives.

SECTION B :
***
CONTRIBUTIONS AND RELATED WORK

This section presents the contributions of this thesis work in the form of four articles (Chapters 5, 6, 7, and 8), preceded by a introductory Chapter 4 presenting the thesis approach and methodology.

4

Theoretical Approach and Methodology

This Chapter synthesizes the state of the art section, and introduces our thesis approach and methodology guiding the reader throughout the contributions section.

Sommaire

4.1 Introduction

In chapter 1, we have established the [fundamental objective] which we recall here : representing and modeling the composition of a tangible entity in general, and CH tangible entity in particular, using ontological structural and spatial relations, within an Applied Ontology approach. In light of this objective, we specified two essential fields of investigation to be considered in the state-of-the-art Section A, driven by two motivating factors as follows.

In Chapter 2, we specified our modeling scope via [motivation-I] as modeling the composition of any tangible entity as a complex structure, in a way that enables navigating into the entity and understanding and representing its material structure.

To address it, we inspected the types of knowledge organization systems that have been proposed to model CH data, and focused on ontology models as types of systems. Then, we performed a systematic classification of extant ontology models based on three criteria, the geographical scale, the semantics and formality level, and the modeling scope. Based on this classification we were able to excerpt relevant ontology models ; those that we believed can fulfill the specified objective. The selected models (CIDOC CRM, EDM, and CHARM) were then analyzed in details focusing on their capabilities to represent the composition of tangible entities using ontological relations : structural and spatial ones.

The analysis revealed certain insights with respect to our modeling objective. For CIDOC CRM, the model's relations are insufficient when it comes structural relations, such as membership which is not represented, among other possible structural relations. EDM, on the other hand, demonstrated a more data-centric approach, which deviated from our intended entity-centric representation. In the case of CHARM, the conceptual modeling language (ConML) used for its formalization exhibited a degree of less pronounced formality with respect to the desired semantic richness in the relations we seek for.

Based on these insights, we recognized the need for a semantically robust and well-formalized set of foundational ontological relations in order to capture the intended representation of the composition.

After that in Chapter 3, we set [motivation-II], as complementary for the first, calling for acquiring a number of foundational ontological relations that enable understanding and representing the composition of a tangible entity, within a formalized theory.

To address it, we clarified the notion of foundational ontological relations upon which we identified and classified some well-known structural and spatial ones that have been proposed within the applied ontology literature under different aspects of studies. For structural foundational ontological relations, those resembling part-whole relations, we presented the formal studies grouped under mereology, some cognitive sciences studies grouped under meronomy, and other approaches that combine mereology and meronomy, as well as additional properties of relations. As for spatial relations, we illustrated those that enable representing spatial configurations within and among entities i.e. mereotopology and location theories.

Based on our examination, we recognized the need for a unified theory that collectively integrates the significance of various approaches of both structural and spatial relations.

In this Chapter, first we specify the three challenges that we aim to overcome in this thesis using six micro-objectives that we set (Section 4.2). Then, we describe our thesis proposal which builds upon an applied ontological approach and an ontology-engineering methodology as the foundations of this thesis (Section 4.3). After that, we synthesize the thesis's structure and make links to our proposals and their corresponding Chapters in which they are presented, with describing each proposal briefly (Section 4.4). Finally, we conclude (Section 4.5).

4.2 Challenges and Objectives

In the following, we recall the thesis's two motivations. For each, we pose the challenges that we need to overcome, and for each challenge, we set some micro-objectives to accomplish. We cover first motivation-II that responds to motivation-I's requirements.

By [motivation-II] :

In view of [motivation-II], and the state of the art explored on extant theories and taxonomies around selected foundational ontological relations in Sections 3.3 and 3.4, we raise the following challenge and set the necessary micro-objectives to overcome it.

Acquiring a number of ontological, structural and spatial, relations that enable representing the composition of a tangible entity, within a theory.

Challenge A : The need for a well-formalized language of a minimal set of ontological relations including rule constraints and excluding categories.

- ▷ **Micro-objective 1 :** To construct, characterize, and formalize a language specification of relations and rules.
- ▷ **Micro-objective 2 :** To demonstrate the novelty and consistency of the approach proposed in (1) by analyzing its micro-theories concerning extant theories and providing a language serialization that validates its consistency.

By [motivation-I] :

In view of [motivation-I], and the state of the art explored on the ontological models for cultural heritage entities representation in Section 2.4, similarly, we raise the following challenges and set the necessary micro-objectives to overcome them.

Modeling the composition of a tangible entity as a complex structure (i.e. an object, a place, a collection) in a manner that enables the understanding, constructing, and navigating into its tangible discourse (i.e. representation), and learning its intangible aspects (i.e. significance).

Challenge B : The need for a generic ontology that understands, represents, and models the structural and spatial constraints of a tangible entity.

- ▷ **Micro-objective 3 :** To establish a decidable lite formalization of the language built in (1).
- ▷ **Micro-objective 4 :** To provide a lightweight ontological model as a tool offering the language re-formalized in (3).

Challenge C : The need for a mapping and query pattern, according to specific employment method(s), to navigate and exploit the proposed ontology, and infer information relevant to the underlying questions concerning the materiality of the tangible entity.

- ▷ **Micro-objective 5 :** To demonstrate the applicability of the model provided in (4) in real life applications i.e. its employment in practice.
- ▷ **Micro-objective 6 :** To demonstrate the convenience of the model provided in (4) in answering queries and inferencing new triplets.

4.3 The FORT approach and methodology

To meet these micro-objectives, and contemplate the [fundamental objective] (reinforced by [motivation-I] and [motivation-II]), we demonstrate in this section our thesis's approach and the methodology built for its implementation.

Our thesis builds upon (1) an applied ontological approach (Section 4.3.1) and (2) an ontology engineering methodology (Section 4.3.2). Together, the ontology and the methodology, make the two foundations of the thesis.

4.3.1 An Applied Ontological approach

a) In order to represent the composition, ontological relations are required to represent the links between entities, between an entity and the entities that compose it, and those that locate it.

b) In order to build a meta-ontology, the relations in the ontology shall be syntactically and semantically domain-independent, yet applicable to any domain-specific entities.

c) In order to achieve an interdisciplinary approach on a cross-disciplinary entity, the heterogeneous views of each discipline regarding it own models shall be maintained, i.e. preserving domain-specific models, yet enriching them with the requirements mentioned in (a) and (b).

Requirements of the ontology :

Thus, we shall contribute with a list of selected foundational ontological relations, offered in a ***modular*** ontology (also called theory) of relation modules. Each relation module addresses a specific relation, i.e. its characterization and formalization, and allows for certain representation(s) following (a).

This ontology shall be a ***meta***-ontology in terms of both, the conceptualization which it specifies and the modeling language which it uses (following the presentation of an ontology in Preliminary Remarks), following (b). This former corresponds to specifying a meta-conceptualization representing top-level abstractions i.e. the composition of any tangible entity. The latter refers to using a meta-language of generic (domain-independent) vocabulary.

And for the ontology to be imported (by ontology engineers, users, domain experts, etc.) into domain models with the precise goal of enriching the semantics of representing the composition of their entities, then the practices in terms of its employment shall be straightforwardly and ***exclusively addressing relations and rule constraints***, without overburdening this employment with categories, following (c).

After having identified the requirements of the meta-ontology, we explain its modelisation within an ***Applied Ontological*** context. Following the presentation in the Preliminary Remarks section of an ontology in the applied ontological field, we interpret first the notions of our intended *shared conceptualization* in [A] and the intended *vocabulary of the modeling language* in [B]. Then, we show how to specify and formalize the intended *ontology* in [C].

A. The shared meta-conceptualization :

A shared consensus conceptualization is to refer to an abstract model regarding a shared subject upon which the multiple disciplines agree on. The elements that constitute this conceptualization are used to articulate certain state of affairs in reality, called abstractions, as explained in Preliminary Remarks. These abstractions are mainly template examples inspired from reality. Since our intended application domain is cultural heritage, we seek for examples of cultural heritage tangible artifacts, as templates of tangible entities (the abstractions) whose composition is to be modeled. This task of abstracting examples allows for understanding the requirements of representing the structure of tangible entities, and in particular cultural heritage entities, with the interest of studying their spatial and structural constraints to have the necessary information for their preservation and restoration. [1] Such information can only be acquired from domain experts, i.e. researchers within the Patrimalp project e.g. material sciences, historians, archaeologists, geologists, and physical-chemists, whom share a common interest : studying the structure of their cultural heritage patrimonial artifacts in order to preserve/restore them.

In order to guide interactions with domain experts, we interviewed them using a pre-prepared questionnaire, as experiments. To do so, we adhered to the experimental protocol [Mandran2018] designed within the framework of the THEDRE method [Mandran2022] of research for supporting experimental methods in information systems research. The scientific protocol aims at guiding the process of producing the intended data and specifying the measures and methods for analyzing it. Multiple guides and tools for prepare, lead, and assess experiments exist, from which we built our experimental protocol and questionnaire, attached in Annex A.

Examples of entities which participants focus on are archaeological and paritial entities such as archaeological sites, caves, rock shelters, settlements, theater, cathedral, ancient ruins, medieval town, castle, sculptures, statues, and collections of entities. The experiments confirmed the need for semantically rich ontological relations that describe the structural composition and the spatial circumstances of the entity, and in particular the links between (called representations) :

- (**R.1**) : an entity and its different parts e.g. the micro-sample taken from a rock art figure, the hematite powder producing the color of the red coloring matter,
- (**R.2**) : an entity as collective and the entities it groups under certain semantics e.g. a collection of figures sharing a common shape, collection of statues having the same brocade design and/or composition,
- (**R.3**) : an entity and its constituents e.g. a brocade's layer constituted of tin, the type of clay constituting the "Saint Jean" statue,
- (**R.4**) : an entity and the spatial region that locates it e.g. the Rocher-du-chateaux site situated in Bessans in the valley of Maurienne,
- (**R.5**) : and the spatial constraints among entities e.g. the schematic figures located on the rock art panel, the brocade located on the robe of the statue.

B. The meta-modeling language :

A modeling language is to provide a set of modeling primitives (vocabulary) that can directly express the shared conceptualization. For the ontology to be a *meta*-ontology (as expressed in the [Requirements] above), a general representation language of primitives is required i.e. a meta-language whose terms are generic, formal, and primitive. And for the ontology to address *exclusively relations*, the primitives of the language must contain terms to represent sole relations (not concepts). Moreover, the choice of the vocabularies of the modeling language plays a role in having an (appropriate) suitable representation of the conceptualization, and a (relatively) comprehensible vocabulary language i.e primitives to construct models that represent the abstractions of the conceptualization [Guizzardi2007].

To select the relations that are indispensable for expressing our shared conceptualization, we select those that resolve the representations mentioned above, **(R.1)** to **(R.5)**, which respond to our objective of representing and reasoning over the composition of tangible entities.

1. Please note that this is not a Knowledge Acquisition task i.e. we are not collecting data from different sources.

— For **(R.1)**, the part-whole relation, aka parthood, ($Part-of$) studies the link between an entity and its parts within different aspects as seen in Section 3.3 (formal logical aspect i.e. mereology, or common-sense reasoning aspect i.e. meronomy). Nevertheless, additional semantics possibly accompany parthood to represent some dependency properties between the part and the whole such as existential ($ExistentialDependence$).

— For **(R.2)**, the membership relation, often studied within parthood theories, addresses the link between collective whole and their particular parts as members ($Member-of$).

— For **(R.3)**, the link between an entity and its constituents is studied in the context of constitution, which, similar to membership, is frequently considered as parthood ($Constitutes$).

— For **(R.4)** and **(R.5)**, both representations refer to locative relations taking place between an entity to be located, and, either a region as the standard location relation ($Location$), or another entity ($Entity-Location$).

The relations that we identified form a list of modeling primitives (vocabularies) of our meta-modeling language. To formalize and characterize each relation, rules are required in the form of definitions, axioms, and theorems, to constraint the semantics of the relations. And because these rules (might) require additional vocabularies, we say that we have selected the basic list of vocabularies as the set of indispensable relations to achieve the required representations, rather than exclusively limiting the vocabulary choice to the selected relations. We refer to this set as a minimal one since it encompasses the basic and minimum number of relations needed to provide the required semantics.

As in defense of the use of the term "minimal", we provide two arguments. The first argues that the removal of a single relation of the selected list yields in a limitation in covering the required number of representations specified above, i.e. each relation is inevitable for achieving an intended representation. We defend this argument in a detailed manner in Chapter 8 in which we illustrate the employment of FORT in practice for two case studies within the cultural heritage domain. The second argument, discussed verbosely in Chapter 6, claims that these relations have been commonly and exhaustively addressed in the applied ontological literature on foundational relations such as mereology, and in other theories as foundational ontologies, as fundamental ontological relations. This ensures the significance of each relation as an ontological foundational relation, and the emphasis of adding each into a basic set of relations.

C. The meta-ontology in a nutshell :

In terms of our preceding interpretations, the shared conceptualization [A] and the modeling language [B], we propose here in [C] our meta-ontology. More precisely, in this Chapter we describe only briefly the ontology's specification and establish a methodology for formalizing the ontology in multiple renderings (formal languages), while the detailed illustration of both the specification and the original rendering is presented in Chapter 5.

As illustrated in Preliminary Remarks, an ontology is a conceptual specification within a logical rendering. The conceptual specification (in [i] below) describes how the primitives of the modeling language are to be used, under a certain interpretation, to formalize the shared conceptualization, as ***the concrete representation*** of the ontology written in natural language.
While the logical rendering (in [ii] below) is the description of this specification in terms of the primitives, as ***the logical theory*** of the ontology written in formal logic.

i. The conceptual specification of the ontology : Our ontology specifies a language of primitive relations and rule constraints that address the [fundamental objective] of the ontological representation-of and reasoning-over the structural and spatial constraints of a tangible entity in general, and a cultural heritage tangible entity in particular. The ontology is named **FORT**, abbreviating a **F**oundational **O**ntological **R**elations as a **T**heory. FORt characterizes and formalizes some selected foundational relations within an ontological context i.e. modelisation of relations only within the scope applied ontology and not pure philosophical ontology. [1] The relations, parthood, dependence, membership, constitution, and location are the indispensable relations forming the minimal set for achieving the intended representation (addressing [motivation-II]). [2] Such a representation allows to model the composition of tangible entities and to navigate its complex structure using generic foundational relations and thus to answer question of restoration and preservation (addressing [motivation-I]). Based on that, we specify the ontology requirements as follows :

Purpose. The purpose of building the ontology is to provide a consensual representation of the composition of any tangible entity.

Scope. The ontology is *modular*, *meta*, and focuses *exclusively* on ontological, structural and spatial, relations and their rule constraints, whose :

- **Abstraction.** level of abstraction is a top-level considering the representation of the composition of any entity (a meta-conceptualization), and represented using generic vocabulary (a meta-modeling language).
- **Users.** intended users encompass ontology engineers, researchers, or other domain experts aiming to model the composition of domain specific entities and requiring the importation of a relations ontology.
- **Uses.** intended uses comprise : representing and understanding the structure of the entity, navigating the structural parts of an entity, locating entities with respect to each other, spatially locating entities in spatial regions, inferring new data about possible links between entities according to their common structure and/or shared location, etc.

ii. The logical rendering of the ontology : To explicate the semantics of the ontology's specification, a logical rendering is crucial to formalize the ontology. It resembles the logical theory of the conceptual specification, using some formal logical language. As discussed previously, the choice of the formal logic used plays a role in achieving a good value ontology, based on its syntax and semantic properties which identify in turn the expressive and computational power of the logic. Thus, according to the desired properties of the ontology, a choice of a knowledge representation language is contemplated. For that, we developed an ontology engineering methodology for formalizing the ontology using multiple logical renderings. The methodology is illustrated verbosely in the next Section 4.3.2, as the complementary foundation of the thesis, besides the ontology.

4.3.2 An ontology engineering methodology

For the formalization of the ontology and establishing its employment within a Global-as-View paradigm required for interdisciplinary applications (e.g. cultural heritage), we construct and adhere to an ontology engineering methodology.

1. In fact, an ontology underlying such a level of abstraction and a general representation language of primitives is referred to as foundational ontology (FO) [Guizzardi2007]. However, FOs study the general laws that describe reality regarding both : foundational concepts and relations. And since our ontology is one of exclusive relations and rules, we do not claim proposing a FO, but a ontology of foundational relations instead.
2. The detailed illustration of the ontological choices made for the characterization of each relation are presented and explained in Chapter 5 i.e. the concrete interpretation and explicit representation of FORT.

It is important to interpret the term "*ontology engineering*" in view of our use of the term as a characterization of our methodology. Firstly, ontology engineering is a term encompasses a broad range of activities related to the development, maintenance, alignment, merging, and evaluation of ontologies. It offers a manner of solving interoperability problems derived from semantics, by providing a set of tasks related to the development of ontologies for a particular domain [Gómez-Pérez2019]. Secondly, a methodology is a "comprehensive, integrated series of techniques or methods creating a general systems theory of how a class of thought-intensive work ought be performed" [IEE1996] i.e. it is to provide generic guidelines for users on how to construct some systems theory. Thus, an ontology engineering methodology concerns the different processes of building and managing an ontology, whether its conceptualization e.g. [Uschold1996], or its reuse across different applications within technical domains e.g. KACTUS [Schreiber2003], or its engineering and development e.g. the Cyc methodology [Lenat1989] for building ontologies with the Cyc Knowledge Base and SENSUS [Knight1994, Knight1995], or even the whole life cycle of the ontology e.g. Methontology [Fernández-López1997] and OnToKnowledge [Sure2003, Sure2004]. Moreover, some ontology engineering methodologies were developed specifically for constructing ontologies within a certain domain such as in the domain of enterprise modeling e.g. TOVE [Gruninger1996] and the Enterprise Ontology [Uschold1995, Uschold1998]. Some extensive literature reviews on methodologies for ontology engineering can be found in [Fernández-López1999] and [Fernández-López2002].

In this methodology, we are particularly concerned with the (a) *formalization of the ontology* which falls under the modelisation phase, and *(b) the possible employment of the ontology using a Global-as-View paradigm for interdisciplinary application.* [1] Ontology formalization is an important step of ontology development which involves other steps as conceptual modeling, specification, implementation, and testing [Noy2001]. Thus, formalization is a crucial step in the ontology engineering ; the reason for which we refer to our methodology as an ontology engineering one. For our ontology, we are interested in providing multiple formalizations, addressing different specification choices, such as expressivity and decidability, using well-known formal logical languages, such as First-Order-Logic (FOL) and Description Logics (DLs). In addition, for each specification choice, it is crucial to offer both : a theoretical formalization and a language serialization that allows for its empirical validation. Moreover, we are also interested in providing the final practice of our ontology within the context of Semantic Web i.e. using Ontology Web Language (OWL) as a Semantic Web standard.

The methodology consists of six steps, depicted in Figure 4.1. Each step is designed to address a/several micro-objective/s from those listed in Section 4.2, as will be illustrated in the next Section 4.4 and throughout the thesis :

1. specify (conceptually) and formalize (logically) the relations of FORT in a highly-expressive formal language that is adequate for the formalization of foundational theories : a first-order-logic (FOL) formalization of the FORT reference ontology.
2. analyze the relations of FORT in the presence of other foundational theories that encompass foundational relations as a relation-based alignment, and validate FORT as a theory by serializing FORT in another formal language that validates the existence of models using consistency checks : a Common Logic (CLIF) serialization of the FORT reference ontology.

1. For the employment of the meta-ontology, we suggest the Global-as-View paradigm specifically for interdisciplinary applications such as cultural heritage. However, the possible employment methods of FORT are two-fold, and will be addressed in Chapter 8.

3. extract a secondary decidable fragment from the original formalization that guarantees desirable computational services, and translate the FOL-formalization into a decidable, yet expressive, knowledge representation and reasoning language : a SROIQ Description Logic formalization of the FORT lightweight ontology.
4. specify and implement the T-boxes of the SROIQ formalization into a semantic web ontological model : an OWL2-DL implementation of the FORT lightweight ontology.
5. import the ontology in practice according to the application setting, and link it to other ontologies based on an employment method : a proof of FORT's applicability.
6. populate the RDF graph using FORT's semantics, and navigate through different data sources to infer new data : a proof of FORT's convenience.

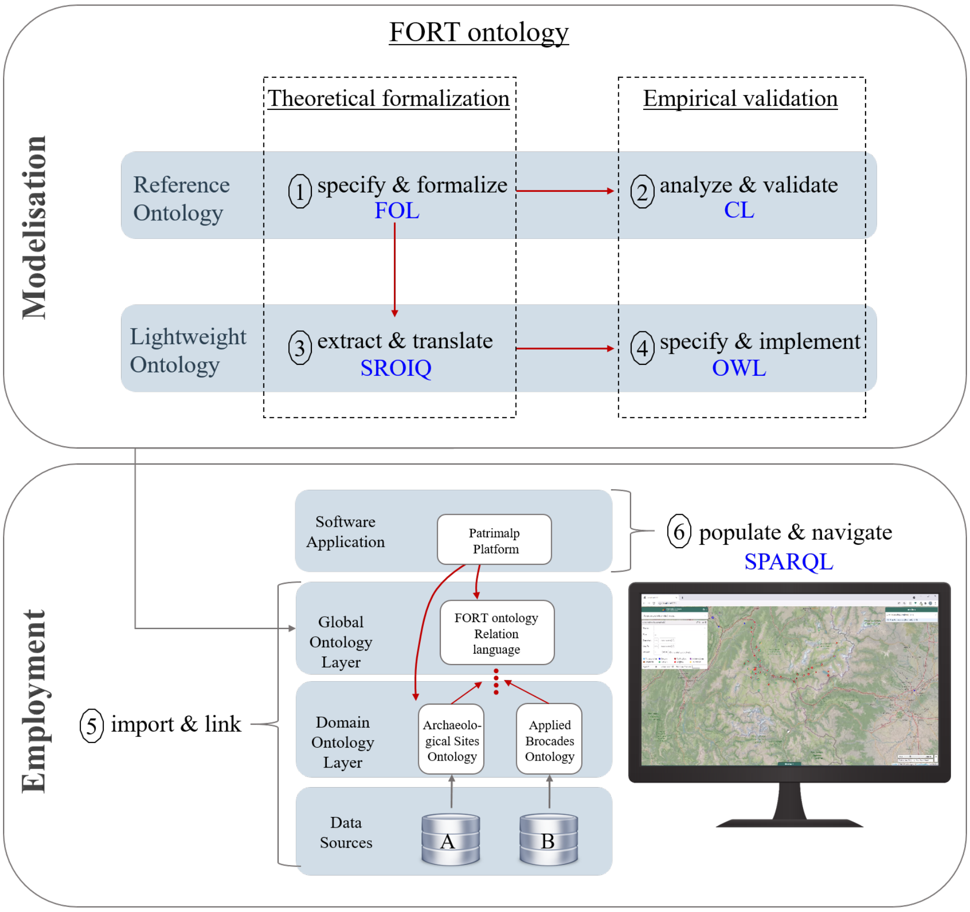


FIGURE 4.1 – *Our ontology engineering methodology designed for (a) formalizing FORT while addressing : two specification choices yielding in a Reference-ontology and Lightweight-ontology, and two formalization levels (theoretical and empirical) yielding in a formalization and implementation of each specification, and (b) achieving a Global-as-View ontology approach with FORT.*

The forte points of our methodology lie in providing a generic guideline for ontology engineers to built a meta-ontology and employ it in practice while robustly :

— addressing different specification choices : *expressivity and decidability*, by providing two specifications of the ontology. The original specification is the *reference-ontology* addressing expressivity using a highly-expressive formal language (FOL), and the secondary specification is the *lightweight-ontology* addressing decidability using a decidable formal language (SROIQ). This is inspired by the work in [Botti Benevides2019] (after being initiated in [Guizzardi2007] following [Guarino1998a]) which promotes to offering two formalizations of an ontology as a way of compromising the trade-off between expressivity and decidability.

— formalizing each specification, the reference and the lightweight ontologies, at multiple levels, *theoretically and empirically*, by providing firstly a theoretical formalization e.g. FOL, SROIQ, and secondly an empirical serialization that allows for validation e.g. CLIF, and supports its use in practice e.g. OWL.

— bridging the two specifications (the reference and the lightweight ontologies) by building a translation procedure which allows computing the lightweight SROIQ formalization (the secondary formalization), given the reference FOL formalization (the original formalization).

4.4 Thesis structure and Contributions

The structure of our thesis reflects the successive examination of the micro-objectives identified in Section 4.2, following the methodology steps specified in Section 4.3.2. Figure 4.2 sketches this structure starting from the thesis's fundamental objective and its two motivations, to the challenges revealed by the state-of-the-art, to the micro-objectives set to overcome the challenges, to the methodology steps managing the micro-objectives, reaching the proposals made (and their corresponding Chapters) to cover the methodology steps.

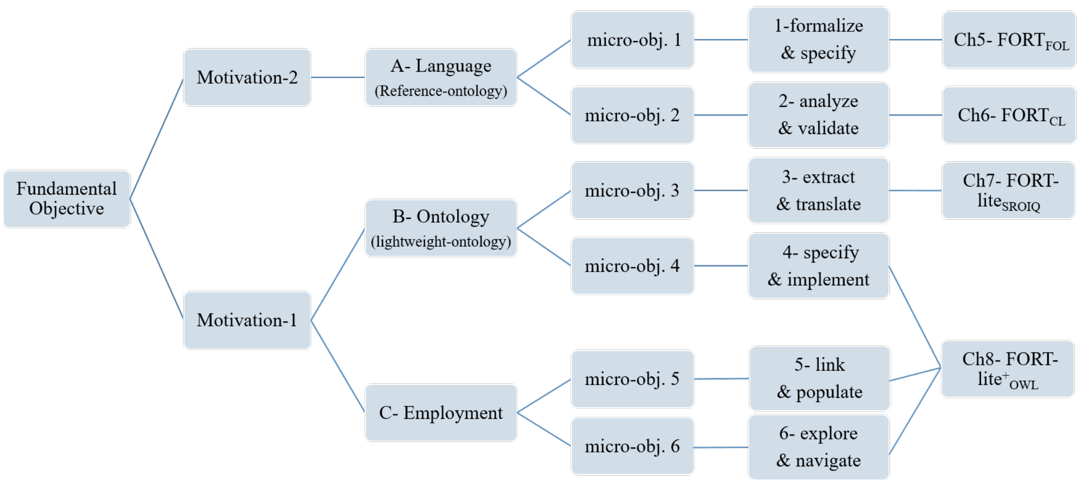


FIGURE 4.2 – *Overview of the thesis structure linking the fundamental objective of the thesis to the two emanated motivations and their corresponding challenges revealed by the state-of-the-art, the challenges to the micro-objectives that we set to overcome them, each micro-objective to the methodology step(s) in which we achieve it, and finally each methodology step and the corresponding Chapter in which we make a proposal to cover it.*

In Chapter 5. we propose and present the Foundational Ontological Relations Theory (FORT). It contributes to a minimal set of foundational ontological relations (parthood, dependence, membership, constitution, and location) formalized as relation modules in FOL. Hence, FORT is *modular*. Moreover, FORT is a *meta*-ontology : specified based on the representations interpreted in [A] (the meta-conceptualization), and using the primitive vocabulary of the modeling language interpreted in [B] (the meta-language), specified based on the representations interpreted in [A]. For the formalization of FORT, we re-use extant formalization of some theories (mereotopology and location), and (re-)formalize other relations (dependence, membership, constitution, and entity-location). This formalization addresses *exclusively* relations, without categories, by normalizing constraints on the relata of the relation without defining entity-types for their domain and range.

This Chapter covers the first step of the methodology by specifying and formalizing the FORT reference ontology, and thus attains the first micro-objective of constructing a language theory.

In Chapter 6. we defend our FORT ontology proposal in the presence of some comprehensive foundational ontologies in which foundational relations are studied and offered. To do so, we elaborate on three arguments for FORT, and elucidate an alignment between the relation-based content of foundational ontologies and the relation micro-theories in FORT. In addition to this analysis, we proceed with validating FORT by serializing the reference ontology in Common Logic that enables running consistency checks and automatic theorem proving, using the CLIF serialization.

This Chapter addresses the second step of the methodology by analyzing and validating the consistency of the FORT reference ontology, and thus realizes the second micro-objective of demonstrating the adequacy and consistency of proposed theory.

In Chapter 7. we develop a generic procedure for translating FOL theories into a SROIQ-formalization fragment. The procedure computes different logical formalisms at each step by performing operations such as rewriting formulas, syntactic/semantic checks, graph transformations, and rule-rolling techniques. We then use this procedure to apply it to FORT's translation into the SROIQ DL relevant for placing the theory in practice.

This Chapter tackles the third step of the methodology by extracting a decidable fragment of the FORT reference ontology (FOL) as a FORT lightweight ontology (SROIQ), and thus establishes the third micro-objective of translating the original expressive theory into a decidable secondary formalization for its use in practice, while contributing with a generic translation procedure.

In Chapter 8. we provide an OWL2DL implementation of the FORT lightweight ontology to support its employment in practice. Then, we discuss the different employment methods of FORT based on the application's setting and objectives showing : both direct use and indirect use (in interdisciplinary applications). After that, based on limited time and inputs from the application domain, we discuss requirements for a complete application and conclude our thesis.

This Chapter thus aims covering the last three steps of the methodology by providing FORT in practice with designing possible employment scenarios for its applicability and convenience proofs, thus achieving the fourth and fifth micro-objectives while laying ground proposals for future work regarding the sixth micro-objective.

4.5 Conclusion

In this Chapter, we have synthesized the requirements of the state-of-the-art section by recalling the fundamental objective of the thesis reinforced by the two motivations which have called for the two Chapters of the state-of-the-art (Section 4.1).

For each motivation, we have determined the challenge(s) to overcome based on the issues revealed in the both Chapters of literature, and established for each challenge the micro-objectives to cover it (Section 4.2).

Then, we introduced our Applied Ontological approach, and ontology engineering methodology (Section 4.3). For the approach, first we identified the requirements for the ontology as modular, meta, and of exclusive relations. Second, we have interpreted the ontology in terms of a meta-conceptualization and a meta-modeling language, followed by presenting the ontology's conceptual specification and logical rendering.

As for the ontology engineering methodology, it aims at guiding the steps for : (a) the formalization of the ontology (the logical rendering phase) addressing different specification choices (expressivity and decidability) and formalization levels (theoretical and empirical), and (b) the establishment of its employment within a Global-as-view paradigm for interdisciplinary approaches.

And last, we illustrated the thesis structure as a map showing our contributions, the micro-objectives they address, and the Chapters in which they are presented (Section 4.4) .

We present in the following Chapters 5, 6, 7, and 8, four consecutive principle contributions, referring to five conference papers, addressing sequentially the methodology steps indicated in Section 4.3 and accomplishing the corresponding micro-objectives specified in Section 4.2.

5

FORT : a modular Theory of Foundational Ontological Relations

This Chapter refers to our contribution "FORT : a minimal Foundational Ontological Relations Theory for Conceptual Modeling Tasks" published as a conference paper at the 41st International Conference on Conceptual Modeling (ER2022), October 17–20, 2022, Forum track.

Sommaire

5.1 Introduction

In the previous Chapter 4, we interpreted our Applied Ontological approach, as a meta-ontology of relations in terms of a meta-shared conceptualization and a meta-modeling language. Additionally, we established an ontology engineering methodology for the formalization of the ontology and achieving its employment in interdisciplinary applications.

In this Chapter, we present our first contribution of a language of exclusive relations and rule constraints [micro-objective-1]. This is by specifying and formalizing the relations of FORT through importing and (re-)formalizing foundational relations from the applied ontological literature, i.e. *formalizing the FORT reference ontology at the theoretical aspect* according to the specified [methodology-step-1].

We first position, in Section 5.2, our proposal based on the review of existing works in Chapter 3 while recalling our motivations. Second, we focus on the presentation of our proposed language at micro and macro levels in Sections 5.3 and 5.4 respectively. In the former, we illustrate the micro-theories of FORT as multiple relation ontologies, whereas the latter presents FORT as a macrotheory (the reference ontology) and specifies its characteristics and design choices. Finally we conclude in Section 5.5 .

5.2 Context : positioning FORT with respect to the literature

As discussed in Chapter 3, the inclusion of foundational relations, such as parthood and membership, along with their formal properties, plays a crucial role in providing a semantically comprehensive ontological analysis for representing a knowledge domain. These relations have been proven essential in the development of formal ontologies, such as DOLCE [Masolo2003], in the alignment of ontologies within the biomedical domain [Smith2004b, Bittner2004c], in disambiguating the semantics of relations such as *isA* (subsumption) and *partOf* (parthood), as well as in providing modeling templates for representing and reasoning over specific subject areas, among others.

Our thesis is particularly driven by the [fundamental objective] of modeling the composition of tangible entities using foundational ontological relations, both structural and spatial in nature. To achieve this objective, we propose to build a modular ontological theory that employs a set of primitive relations and rule constraints. This approach aligns with [motivation-II], allowing for an effective representation of the composition of tangible entities.

The selected relations for inclusion in this ontology are minimal, based on the intended representations (**(R.1)**, **(R.2)**, **(R.3)**, **(R.4)**, and **(R.5)**) specified in Section [A], and are identified in Section [B] as *dependence*, *parthood*, *location*, *membership*, and *constitution*. Furthermore, this ontology whose scope and purpose are set in Section [C] shall be designed to be modular i.e. consisting of multiple relation modules, meta i.e. specifying a meta-conceptualization for top-level abstractions using a meta-modeling language of generic vocabularies, and exclusively addressing relations i.e. not commiting to entity types, based on the specified [requirements].

Given the absence of an existing language that encompasses this minimal set of ontological relations within a modular, meta, and exclusive relations theory, our contribution lies in the development of such an ontology theory, which we refer to as the Foundational Ontological Relations Theory (FORT).

The positioning of FORT regarding the existing literature on foundational ontological relations in both philosophical and applied ontology domains is depicted in Figure 5.1. Within FORT, the utilization and (re-)formalization of relations from the literature occurs at a micro-level through its micro-theories. At a macro-level, FORT comprises a collection of connected micro-theories that collectively form the FORT reference ontology.

As such, it is important to point out the following. In this Chapter, we present each micro-theory as a relation module providing a concise overview of the relevant literature associated with each specific relation and adopting and/or re-formalizing certain analysis (Section5.3), followed by discussing the characteristics of FORT as a macrotheory (Section 5.4). In both, our focus remains on the *demonstration of FORT*, emphasizing key design choices and conducting ontology analyses in relation to existing works.

As for *the novelty of FORT*, it is demonstrated through a comparative analysis with other theories that offer foundational relations as macrotheories (as depicted in Figure 5.1) such as the set of relations considered in the DOLCE foundational ontology, rather than to individual micro-theories. To ensure a robust comparison, two key considerations are taken into account. Firstly, the investigation will solely encompass theories within the realm of ontologies, excluding other philosophical theories. Secondly, the comparison will be conducted at the holistic level of the theory itself. This assessment of novelty and comparison will be carried out in Chapter 6. Furthermore, regarding the applicability of the theory in general, and to the CH field in particular, this will be targeted in Chapter 8.

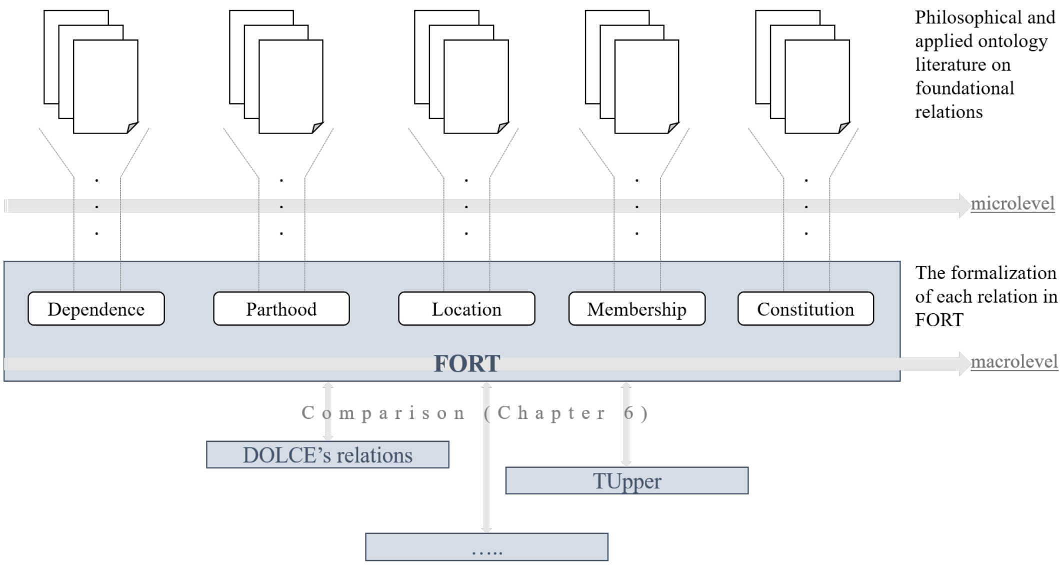


FIGURE 5.1 – *Positioning FORT with respect to the literature on some foundational relations and in the view of other theories offering foundational relations.*

5.3 Methodology step 1 : a FOL formalization [$FORT_{FOL}$] (microlevel)

In this Section, we characterize and formalize the relations in FORT, each as a micro-theory at a microlevel. The basic relation of each micro-theory is recognized as a (primitive) and generic relation, i.e. is not defined in terms of other relations and spans multiple application domains. Each relation is defined and formally characterized in terms of ground axioms that are algebraic properties and non-ground axioms that project some constraints on its use. Thus, the relations are **formal** and often **primitive**. Additionally, relations can be linked to other relations of other FORT micro-theories or extant imported theories, when possible.

Thus, with the relations selected being **foundational**, **ontological**, and possessing **formality** and **primitiveness** satisfying the characteristics we have identified in Section 3.2 in Chapter 3.

In the context of selecting, characterizing, and formalizing general ontological categories and relations, the methodology introduced in [Gangemi2001] holds significant importance. This methodology seeks to avoid making strong ontological commitments while establishing a minimal formal framework essential for understanding, comparing, and evaluating ontological choices. The methodology comprises several steps, where the authors first select relations from classical philosophical literature. Subsequently, they proceed to *select and adapt ground axioms (such as those exhibiting algebraic properties like transitivity)*, and introduce non-ground axioms. Through this process, they identify the formal properties that delineate the nature of these relations.

Given the challenging task of formally characterizing ontological relations, given the highly debated and extensive literature in the field, we are aware of the difficulty and do not claim presenting novel formalizations (as explicitly clarified in Section 5.2). Instead, we rely on the existing literature to inform our use and (re-)formalization of these relations, for a unified theory of foundational ontological relations.

5.3.1 Ontological dependence

As presented briefly in the Section 3.3.4 of Chapter 3, the original analysis on the dependence notion started with Husserl [Edmund1970] within a classical philosophical context, upon which later formalizations were proposed. In [Fine1983], Fine and Smith considered dependence as a quasi-mereological primitive relation introduced in terms of 4 axioms ; reflexive, transitive, and links to parthood. Later in [Simons1987], Simons criticizes the preceding axioms as resembling a sort of a topological relation, and presents dependence within a modal-logic approach, using the existence relation as the primitive. Several authors also addressed the notion of dependence such as ; Thomasson by introducing different kinds of temporal dependence ; constant and historical in [Thomasson1999] ; and Vieu and Aurnague by specializing kinds of generic dependence (functional dependence) both within modal logic formalizations in [Vieu2007]. Moreover, in foundational ontologies such as in BFO [Smith2002] ; it is studied between qualities, realizable entities (e.g. roles and functions), or processes (also process boundaries), and (in)dependent continuants or processes, and in DOLCE [Masolo2003] ; it is deeply axiomatized (ontological and spatial dependence notions) in a modal approach using the presence primitive relation.

Our consideration of ontological dependence is a property characterizing the persistence semantics between two entities (individuals) or entity types (categories). It plays an important role in the representation and reasoning over relations e.g. parthood or connection in particular. We investigate ontological dependence as based on a primitive existence relation (in contrast to Smith and Fine's primitivity) and at both ; instances and universals levels, as inspected in [Smith1997] and following to DOLCE's distinction between specific and generic constant dependence, within a non-modal formalization. For the primitive existence relation we use the binary predicate E, with the notation E(x,t) standing for "entity x exists at time t". The introduction of t as an instance of time in E does not put the approach in a temporal framework, but simplifies the representation of the framework at instants of time, when necessary.

$$\forall(x)\neg E(x,x) \tag{Da1}$$
$$\forall(x,y)E(x,y) \rightarrow \neg E(y,x) \tag{Da2}$$

An entity x is specifically existentially dependent entity y, denoted $SED(x,y)$, if ; at any time t, x cannot exist at t unless y exists at t ; & x and y are different entities ; & x exists at some t (Da3). For example, a person is be specifically existentially dependent on its brain.

$$\forall(x,y)SED(x,y) \rightarrow \forall t(E(x,t) \rightarrow E(y,t)) \wedge \neg(x = y) \wedge \exists t E(x,t) \tag{Da3}$$

By (Da3), it follows from $\neg(x = y)$ that SED is irreflexive and asymmetric ;

$$(\forall x)\neg SED(x,x) \tag{Dt1}$$
$$(\forall x,y)SED(x,y) \rightarrow \neg SED(y,x) \tag{Dt2}$$

We furthermore consider the transitivity of the SED as follows ;

$$(\forall x,y,z)SED(x,y) \wedge SED(y,z) \rightarrow SED(x,z) \tag{Da4}$$

An entity type ϕ is generically existentially dependent on entity type φ, denoted $GED(\phi,\varphi)$, if ; at any time t, for every x instance of ϕ, x cannot exist at t unless there exists some instance y of φ at$t$ and x and y are different entities ; & there exists time t such that there exists instance x of ϕ ; & ϕ and φ are disjoint (Da5). For example, a person might be generically constantly dependent on having a heart.

$$\forall(\phi,\varphi)GED(\phi,\varphi) \rightarrow \forall x,t((\phi(x) \wedge E(x,t)) \rightarrow \exists y(\varphi(y) \wedge E(y,t))) \wedge \exists x,t(\phi(x) \wedge E(x,t)) \wedge \neg\exists z(\phi(z) \wedge \varphi(z)) \tag{Da5}$$

By (Da5), it follows from $\neg\exists z(\phi(z) \wedge \varphi(z))$ that GED is irreflexive and asymmetric ;

$$(\forall\varphi)\neg GED(\varphi,\varphi) \tag{Dt3}$$
$$(\forall\phi,\varphi)GED(\varphi,\phi) \rightarrow \neg GED(\phi,\varphi) \tag{Dt4}$$

We furthermore consider the transitivity of the GED under the condition of ϕ and α being disjoint as follows ;

$$GED(\phi,\varphi) \wedge GED(\varphi,\alpha) \wedge \neg\exists x(\phi(x) \wedge \alpha(x)) \rightarrow GED(\phi,\alpha) \tag{Da6}$$

We have introduced both ontological dependence relations, specific and generic, as axioms i.e. only the necessary conditions of dependence hold ($SED \rightarrow$ ... and $GED \rightarrow$ ...), rather than a definitions. The choice of eliminating the sufficient condition of both ontological dependence relations (... $\rightarrow SED$ and ... $\rightarrow GED$) is to ensure that the inverse does not hold i.e. if two entities/categories x/X and yY coexist at all times of their persistence, it does not necessarily mean that x/X or y/Y is specifically/generically existentially dependent on y/Y or x/X, respectively. For example, two different persons, born on the same exact time, coexisted during their lifetimes, and having died on the same exact time, does not yield in them being existentially dependent on one another. This view of ontological dependence is not in-line with DOLCE. Indeed, we believe that for the sufficient condition of the formula to hold, a stronger predicate is needed more than just the existence relation E.

5.3.2 Parthood and dependence

In FORT, we adopt parthood P as satisfying the axioms of Closure Extensional Mereology, and thus importing **CEM** i.e. importing Pa1-Pa8 along with the corresponding theorems and parthood predicated illustrated in Section 3.3.1. Furthermore, to allow for additional semantic inferences on the persistence of wholes depending on the persistence of their parts, we provide an extension of parthood P using ontological dependence *SED* and *GED*, inline with DOLCE [Masolo2003], UFO [Guizzardi2015], and other foundational theories.

Relations grouped under this section provoke for part entities that are inseparable from their wholes. The notion of separability that we seek for is not that of physical connection/disconnection of a part/whole from its whole/parts by any means of physical (e.g. by hand, clippers, scissors) or chemical (e.g. chemical filtration process) separation which depends on some granularity level of separation. Instead, separability is elucidated by means of ontological dependence (specific and generic existential dependencies) which serves an important role in the reasoning over the persistence conditions of the parts and whole entities. Thus, adding semantically specialized parthood relations and benefiting the usage and representation of dependencies in conceptual modeling tasks. Using the two preceding dependency definitions of section 5.3.1, we introduce the notions of components and elements. We tag the definitions, axioms and theorems of this micro-theory as starting by *PD* referring to parthood and dependence.

Componenthood :

x is a $ComponentOf$ y iff ; x is a part of y & y is generically existentially dependent on x (PDd1). For example ; the engine is a component of the car, the heart is a component of the body of a living being.

$$(\forall x,y)ComponentOf(x,y) =_{df} P(x,y) \wedge GED(\phi(y), \varphi(x)) \quad \text{(PDd1)}$$

From the axioms of P and PDd1, $ComponentOf$ is a strict partial order relation (PDt1-PDt3).

$$(\forall x,y,z)ComponentOf(x,y) \wedge ComponentOf(y,z) \rightarrow ComponentOf(x,z) \quad \text{(PDt1)}$$

$$(\forall x)\neg ComponentOf(x,x) \quad \text{(PDt2)}$$

$$(\forall x,y)ComponentOf(x,y) \rightarrow \neg ComponentOf(y,x) \quad \text{(PDt3)}$$

Démonstration. From PDd1, consider ; $\exists a,b(P(a,b) \wedge GED(\phi(b), \varphi(a)))$ and $\exists c(P(b,c) \wedge GED(\alpha(c), \phi(b)))$, i.e. $ComponentOf(a,b) \wedge ComponentOf(b,c)$. By the transitivity of P (Pa3) and that of *GED* (Da6), it follows that $P(a,c)$, and $GED(\alpha(c), \varphi(a))$, i.e. $ComponentOf(a,c)$. □

Démonstration. Assume that $ComponentOf$ is reflexive. Then, $\forall x ComponentOf(x,x)$ i.e. $P(x,x)$ $\wedge GED(\phi(x), \phi(x))$. However, *GED* is irreflexive by Dt3. Thus, $ComponentOf$ is irreflexive. □

Démonstration. If we assume $ComponentOf$ is antisymmetric, this yields in $ComponentOf(x,x)$ and $GED(\phi(x), \phi(x))$ by PDd1 which does not hold according to the irreflexivity of $ComponentOf$ PDt2 by that of *GED* Dt3. On the other hand, assuming that $ComponentOf$ is symmetric, means $ComponentOf(x,y)$ and $ComponentOf(y,x)$, results in $P(x,y) \wedge P(y,x)$, thus $x = y$ by Pd2 yielding in a contradiction according to PDt2. Hence, $ComponentOf$ is asymmetric. □

$ComponentOf$ is thus a proper parthood satisfying weak supplementation axiom (PDt4).

$$(\forall x,y)ComponentOf(x,y) \rightarrow PP(x,y) \quad \text{(PDa1)}$$

$$(\forall x,y)ComponentOf(x,y) \rightarrow \exists z(P(z,y) \wedge \neg O(z,x)) \quad \text{(PDt4)}$$

Démonstration. From PDa1 and the weak supplementation axiom Pa4 satisfied by P.

□

Elementhood :

x is an $ElementOf$ y iff ; x is an part of y & y is specifically existentially dependent on x (PDd2). For instance ; the tin layer is an element of the brocade, the brain is an element of the human's body. Elements parts of a whole are those whose existence is elementary. For instance, the spatial existence of the tin layer in the brocade is (along the other layers) what makes up the identity of a brocade as a whole.

$$(\forall x,y)ElementOf(x,y) =_{df} P(x,y) \wedge SED(y,x) \tag{PDd2}$$

From the axioms of P and (PDd2), $ElementOf$ is a strict partial order relation (PDt5-PDt7).

$$(\forall x,y,z)ElementOf(x,y) \wedge ElementOf(y,z) \rightarrow ElementOf(x,z) \tag{PDt5}$$
$$(\forall x)\neg ElementOf(x,x) \tag{PDt6}$$
$$(\forall x,y)ElementOf(x,y) \rightarrow \neg ElementOf(y,x) \tag{PDt7}$$

Démonstration. From PDd2, consider ; $\exists a,b(P(a,b) \wedge SED(b,a))$ and $\exists c(P(b,c)SED(c,b))$, i.e. $ElementOf(a,b) \wedge ElementOf(b,c)$. By the transitivity of P Pa3 and that of SED Da4, it follows that ; $P(a,c)$, and $SED(c,a)$ respectively, i.e. $ElementOf(a,c)$. □

Démonstration. Assume that $ElementOf$ is reflexive. Then $\forall x ElementOf(x,x)$ i.e. $P(x,x) \wedge ED(x,x)$. By Dt1, SED is irreflexive. Thus $ElementOf$ is irreflexive. □

Démonstration. Similar to the proof of PDt3, on the one hand, assuming that $ElementOf$ is antisymmetric yields in $ElementOf(x,x)$, and in $SED(x,x)$ by PDd2 which does not hold according to PDt6. On the other hand, assuming that $ElementOf$ is symmetric results in having $ElementOf(x,y)$ and $ElementOf(y,x)$, i.e. $P(x,y) \wedge P(y,x)$, thus $x = y$ by Pa2 yielding in a contradiction. Hence, $ElementOf$ is asymmetric. □

Similar to (PDa1), it follows that $ElementOf$ is a proper parthood relation (PDa2).

$$(\forall x,y)ElementOf(x,y) \rightarrow PP(x,y) \tag{PDa2}$$

From PDa2 and the weak supplementation axiom satisfied by P, $ElementOf$ satisfies weak supplementation axiom (PDt8).

$$(\forall x,y)ElementOf(x,y) \rightarrow \exists z(P(z,y) \wedge \neg O(z,x)) \tag{PDt8}$$

Démonstration. From Pa2 and the weak supplementation axiom a4 satisfied by P. □

At this point, it is possible to assert predicates of $ComponentOf$ and $ElementOf$, using predicates of parthood P, such as $ComponentOverlap$, $ElementOverlap$, $ComponentUnderlap$, and $ElementUnderlap$ using the overlap and underlap relations O (Pd3) and U (Pd4). Such predicates allow for additional inferences about entities sharing a common element or component as being identical ; PDa3 and PDa4.

$$(\forall x,y)ComponentOverlap(x,y) =_{df} \exists z(ComponentOf(z,x) \wedge ComponentOf(z,y)) \tag{PDd3}$$
$$(\forall x,y)ElementOverlap(x,y) =_{df} \exists z(ElementOf(z,x) \wedge ElementOf(z,y)) \tag{PDd4}$$
$$(\forall x,y)ComponentOverlap(x,y) \rightarrow x = y \vee ComponentOf(x,y) \vee ComponentOf(y,x) \tag{PDa3}$$
$$(\forall x,y)ElementOverlap(x,y) \rightarrow x = y \vee ElementOf(x,y) \vee ElementOf(y,x) \tag{PDa4}$$

Following several works in the applied ontological literature, parts whose wholes are generically dependent on these parts are referred to as mandatory parts, which applies of our $ComponentOf$ relation in FORT. A popular example from the literature is the human being heart which is considered as a mandatory human body part, that can be replaced by another heart entity that is classified within the category as a heart, thus, a heart is a component of the human body. Whereas, parts whose wholes are specifically dependent on these parts are referred to as essential parts, which applies to our $ElementOf$ relation in FORT. This is the case of the human brain upon which a human's body is specifically dependent on that brain making it an essential body part i.e. the brain is an element of the human body.

5.3.3 Location

In FORT, we treat the location relation in a three-fold manner, as follows.

Region-to-region locative relations :

These are location relations expressing the spatial inclusion of entities of the same type which is regions. In such a case, the problem of spatial inclusion representation can be generalized into a simple inclusion representation using parthood P to show one entity is inside another entity i.e. part of it. For a wide range of expressive representations, the mereo(topo)logical theory of relations, including parthood, connection, and the fusion of both, illustrated in Section 3.4.1 of Chapter 3, is imported as a theory in FORT to express region-to-region locative relations without the commitment to formalizing a region entity type.

Entity-to-region locative relations :

These are locative relations resembling the address of an entity in a spatial space while remaining neutral with respect to the ultimate ontological status of the reference which located them, based on the location theory presented by Casati and Varzi [Casati1999]. Thus, for these locative relations, FORT import Casati's and Varzi's location theory (presented in Section 3.4.2 of Chapter 3) using the L primitive (denoting "exactly located at region") and borrowing the axioms (La1-La3), definitions (Ld1-Ld6) and theorems (Lt1-Lt3). The borrowed formalization of L further establishes ; links to parthood and connection (La4-La5) ; reasoning on the location of the mereotopological properties of an entity with respect to its location (Lt4-Lt6) ; reasoning on the location of an entity with respect to the mereotopological properties of its location (Lt7-Lt9) ; and reasoning about the location of the mereotopological properties of an entity with respect to the mereotopological properties of its location (Lt10-Lt15).

Entity-to-non-region locative relations :

Motivated by the work in [Donnelly2006], these are additional locative relations expressing the spatial link between entities that are temporarily or permanently located in spaces, but never share parts with each other or with these spaces. For example ; the painting is located on the wall, the basket is located at the top of the fridge. These examples hold even though the wall and the top of the fridge are not regions (i.e. L solely cannot hold), and neither the painting is part of the wall nor the basket is parts of the fridge (i.e. P cannot hold). In such cases, neither pure mereological relations, i.e. the case of region-to-region locative relations, nor solely region-location relations, i.e. entity-to-region locative relations, are useful to represent such links.

Since we believe that this aspect of a locative relation is essential, we introduce the entity-location relation using a primitive EL indicating "located at/on/in", with making links to parthood and connection. We assume that every entity is located in/at itself i.e. reflexive (ELa1), and that transitivity is guaranteed i.e. is x is located in y and y is located in z then x is located in z (ELa2).

As links to mereology, we assert the axiom (ELa3) ; if x is part of y then x is located in y, which directly implies (from ELa2) the theorems (ELt1) ; if x is part of y and y is located in z then x is located in z, and (ELt2) ; if x is located in y and y is part of then x is located in y. The preceding two theorems serve in the reasoning about the entity-location of entity x, with respect to the entity-location of its whole y (e.g. the red dear is part of the painting, and the painting is located at the wall, then the red dear is located at the wall), and the location of an entity x with respect to the whole of its location entity z (e.g. the painting is located at the rock wall, and the wall is part a of the Rocher du Château, then the painting is located in the Roche du Chateaux site). In both cases, the entity x is (not necessarily) part of the entity y, rather than a pure locative relation.

As links with the locative relation L, using (7.16) and (L.3), theorems (ELt3) ; if x is entity-located in y, and y is exactly located in z, then x is wholly located in z, and (ELt4) ; if x is entity-located in y, then x's spatial region if part of y's spatial region, are provable respectively.

$$(\forall x)EL(x,x) \tag{ELa1}$$
$$(\forall x,y,z)(EL(x,y) \wedge EL(y,z)) \rightarrow EL(x,z) \tag{ELa2}$$
$$(\forall x,y)(P(x,y) \rightarrow EL(x,y)) \tag{ELa3}$$
$$(\forall x,y,z)(P(x,y) \wedge EL(y,z) \rightarrow EL(x,y)) \tag{ELt1}$$
$$(\forall x,y,z)(EL(x,y) \wedge P(y,z) \rightarrow EL(x,y)) \tag{ELt2}$$
$$(\forall x,y,z)(EL(x,y) \wedge L(y,z) \rightarrow WL(x,z)) \tag{ELt3}$$
$$(\forall x,y,z,w)(EL(x,y) \wedge L(x,z) \wedge L(y,w) \rightarrow P(z,w)) \tag{ELt4}$$

Furthermore, one can define other entity-location locative predicates, such as tangential and interior entity-location relations, using mereotopological definitions of tangential and interior parts, as well as partial and whole entity locative relations using locative definitions of partial and whole locations.

5.3.4 Membership

Upon the characterization of the membership relation ; a debate arises two views on the relation. The first considers the relations as a part-whole relation typology i.e. formalized using parthood such as in formal ontological studies e.g. [Gangemi2001] following the analysis in [Guarino2000b], and in the meronymic literature on part-whole relations e.g. [Winston1987] and [Gerstl1996]. The second view analyzes membership as a primitive independent relation e.g. [Simons1987]. However in both, the members of a whole participating in a membership relation acquire a unifying relation (also referred to as a uniform structure in meronymic literature) that binds all the members together and a maximality constraint on the members with respect to this relation.

Furthermore, an imperative point in the formalization of membership derives from ranging the relation over collective wholes, also called aggregates, and proceeding with characterizing collectives/aggregates as a mandatory for characterizing membership. Characterizing aggregates can be performed by ; a relation holding among members [Fine1999] ; a common role played by all members [Guizzardi2011] ; specifying a single entity type or a least common subsumer type that all members are instantiated to as in [Rector2006b] and [Galton2010] respectively ; the uniqueness of the collective's decomposition into members in [Masolo2020].

In Fine [Fine1999], a collective structure could be represented by means of a relation holding among the members. Collectives can be then reduced to variable embodiments i.e. entities that at any time t at which they are present, are constituted by a rigid embodiment i.e. a sort of compound of the whole entity. In [Rector2006b] and [Guizzardi2011], the authors further expand by assigning a common role played by all the members of a collective, which in turn necessitates the formal characterization of such a role.

Moreover, in [Rector2006b] and [Galton2010], it is assumed that given a collective, one tends to specify the type of its members e.g. a collective of type forest has members only entities as type trees. In the former, authors thus state that all members of a collective are of a specified type, whereas in the latter, the authors does not require the type(s) of all the members to be unique, but to have a least common subsumer i.e. a minimal type that under which all the classes are subsumely closed.

In [Masolo2020], and in-line with the structural constitution view introduced by Harris for group agents [Harris2020], the authors propose 3 axiom forms (f1,f3, and then f5) to characterize collectives that decompose into members and distinguish them from composites that decompose into components, which were considered still quite weak to make the distinction. This enhanced two further (a) the uniqueness of the collective's decomposition into members (vs the multiple possible decomposition criterion of a composite into components) via an axioms (a9) that enforces the unique decomposition of collectives, and (b) the possibility of recursively decomposing composites but not collectives i.e. intransitivity of membership via axioms (a10-a11).

In FORT, we regard membership as a primitive relation distinct from parthood inline with that in [Simons1987]. More precisely, we follow Simons's approach in attaching a notion of unity to the range of the membership relation i.e. the aggregate, and adopt the two preceding aspects as follows. First, we formalize membership, notating "is member of", using the binary predicate memberOf as irreflexive (Ma1) and asymmetric (Ma2).

$$(\forall x)\neg memberOf(x,x) \tag{Ma1}$$

$$(\forall x,y) memberOf(x,y) \rightarrow \neg memberOf(y,x) \tag{Ma2}$$

Second, we proceed with characterizing the range by implanting a number of axioms. We employ axioms from BFO's axiomatization of an aggregate entity [Smith2002] ; an aggregate has more than one member at least one time (Ma3) ; all proper parts of an aggregate overlap some member (Ma4) ; and all members of an aggregate are disjoint proper parts (Ma5). Then we add the axioms (Ma6) stating that an aggregate is the exact sum of its members.

$$(\forall x,y) memberOf(y,x) \rightarrow \exists t, m_1, m_2 (m_1 \neq m_2 \wedge E(x,t) \wedge memberOf(m_1,x) \wedge memberOf(m_2,x)) \tag{Ma3}$$

$$(\forall x,p,y) memberOf(y,x) \wedge PP(p,x) \rightarrow \exists o (memberOf(o,x) \wedge O(o,p)) \tag{Ma4}$$

$$(\forall x,y) memberOf(x,y) \rightarrow (PP(x,y) \wedge \forall m (memberOf(m,y) \rightarrow x = m \vee \neg O(m,x))) \tag{Ma5}$$

$$(\forall x,y) memberOf(y,x) \rightarrow (\forall w (O(w,x) \leftrightarrow \exists m (memberOf(m,x) \wedge O(w,m)))) \tag{Ma6}$$

Third, we advance with characterizing the members of an aggregate by sharing a characteristic property : a unity according to a unifying relation. The dispute resides on the ground axioms of the unifying relation ; more precisely transitive in [Guarino2000b] and intransitive in [Gangemi2001]. In our theory, we endorse Gangemi's intransitivity and proceed with its formalization. Let $\mathfrak{R}_R$ denote a finite set of binary predicates that are unifying relations representing characteristic relations of entities, such that $\forall R_i \in \mathfrak{R}_R$, R_i is conditionally reflexive (Ra1) and symmetric (Ra2). Then we define the unification, denoted $\mathscr{U}_{R_i}(z)$ as ; an entity z is unified under R_i iff z is the sum of entities in the domain of R_i, and all entities that possess R_i and are parts of z are linked by R_i (Rd1).

$$(\forall x,y)R_i(x,y) \rightarrow R_i(x,x) \wedge R_i(y,y) \quad \text{(Ra1)}$$
$$(\forall x,y)R_i(x,y) \rightarrow R_i(y,x) \quad \text{(Ra2)}$$
$$(\forall z)\mathscr{U}_{R_i}(z) =_{df} \forall r(R_i(r,r) \rightarrow P(r,z)) \wedge \forall m(O(m,z) \leftrightarrow \exists r(R_i(r,r) \wedge O(m,r))) \wedge \forall a,b(R_i(a,a) \wedge R_i(b,b) \wedge P(a,z) \wedge P(b,z) \rightarrow R_i(a,b)) \quad \text{(Rd1)}$$

Some interpretation must be given about the choice of predicates in Rd1. The first statement, z is the sum of entities in the domain of R_i, indicates that each entity that is satisfied by R is a part of z. However, alone, it is not sufficient to assert that z is exactly the sum of the entities satisfying R_i and not more. For example; consider $p1$, $p2$, and $p3$ as the entities satisfying R (i.e. $R(p1,p1)$, $R(p2,p2)$, and $R(p3,p3)$), and consider z as the entity built by $p1$, $p2$, $p3$, and some other entity g. In this case, z satisfies the first statement, knowing that it does hold as the intended meaning.

Thus, a second statement is needed to strengthen the declaration of the sum of R's entities; for every entity m that is overlapping z; there must exist an r entity belonging to the domain of R, such that m overlaps r.

The third statement, for each pair of entities that are parts of z "and satisfying R_i", the entities must be linked with R_i too. The predicate "and satisfying R_i" is to ensure that parts of parts of z (that do not satisfy R) are not (necessarily) linked by R_i. For instance, the hand of a person of jury (who is part of a jury) should not be linked with the role of being a jury member, so that the jury is unified by the relation : being a jury member.

After defining $\mathscr{U}_{R_i}(z)$, we further axiomatize the range of $memberOf$ as a unified entity (Ma7); each aggregate has a unified relation R_i under which it is unified at all the times that it exists.

$$(\forall x,y)memberOf(y,x) \rightarrow \exists i(\mathscr{U}_{R_i}(x) \wedge (\forall t(E(x,t) \rightarrow \mathscr{U}_{R_i}(x))) \quad \text{(Ma7)}$$

In contrast to [Masolo2020] in characterizing collectives, we consider that the characteristic property unifying the whole of a membership relation (or as it is formalized being the plurality constituting the collective), holds at all times of the existence of the collective, and applies on any plurality that constitutes. In other words, even if this sum (plurality) changes with time e.g. a member ceases, then the whole of the membership relation (collective) still maintains its unification under R_i. While in their approach, they consider that for the plurality to be x to be characterized by a property F at t notated $F_t x$, it has to be wholly present at t, notated $\varepsilon_t^w x$. Our interpretation for our disagreement is that we consider the aggregate as unified by R_i at all the times that it exists even if some members change or decreased in number. This only meas that these members do not satisfy R_i anymore, while the aggregate is still unified by R_i.

Fourth, we use what preceded to infer the conditions under which two aggregates are considered identical in (Mt1) if they are unified by the same unification relation.

$$(\forall x,y,w,z)memberOf(x,y) \wedge memberOf(w,z) \wedge \exists i(\mathscr{U}_{R_i}(y) \wedge \mathscr{U}_{R_i}(z)) \rightarrow y = z \quad \text{(Mt1)}$$

5.3.5 Constitution

In the ontological literature on constitution, its formalization varies according upon two philosophical views of the world. A multiplicative-based view approach allows for different entities to be co-localized in the same space-time. Different entities signify incompatible essential properties, such as persistence properties, yet related. Whereas, a reductionist-based view approach presupposes that each space-time location contains at most one entity, and the incompatible essential properties are only unintended different interpretations of different perspectives that one can assume about spatio-temporal entities. Thus, a main difference between the two views regards the mode of existence of entities populating the world at a metaphysical level [Guarino2017]. Also, foundational ontologies adopt different philosophical views which develop highly in some foundational relations such as constitution. We demonstrate the difference through a popular example ; the vase and the clay which constitutes it.

For instance, DOLCE [Masolo2003] is a descriptive ontology (multiplicative view) adopting a cognitive based representation of the world underlying natural language and human commonsense. With constitution, DOLCE recognizes a vase, as constituted by an amount of clay, and clay, as an amount of matter. A vase and amount of clay are taken as two different types that are co-localized in the same space-time location. DOLCE supports the claim of constitution is not identity based on three arguments following [Thomson1998] and [Baker2000]. Firstly, the two entities have different histories ; clay can be present before the vase. Secondly, the two entities have different persistence conditions ; the clay can persists upon a change of change while the vase ceases to exist ; and the vase can undergo a replacement of a certain amount of clay by another amount, while a piece of clay cannot i.e if replaced, the piece of clay is not the same piece anymore. Thirdly, the two entities differ in their essential metaphysical relational properties i.e. the clay can exist without any artificial intervention while a vase needs an intended intervention to exist.

While BFO [Smith2002] is a realist ontology (reductionist view) capturing the world as (multiple) particular perspectives of reality i.e. a possibly multiple instantiations of the same particular individual. In contrast to DOLCE, BFO regards the entities participating in a constitution relation, e.g. the vase and the clay, as the same spatio-temporal individual that instantiates different universals at the same spacetime location.

Thus, a main difference between the two views regards the mode of existence of entities populating the world at a metaphysical level [Guarino2017]. The incompatible essential properties referred to in the multiplicative approach are only unintended different interpretations of different perspectives that one can assume about spatio-temporal entities, which, in the reductionist approach are perceived as identical.

In FORT, we adhere to the multiplicative view to formalize constitution. First, we build the primitive constitution relation using the binary predicate constitutes as a strict partial order relation (COa1-COa3), and link it to existence (COa4) and parthood (COa5) relations.

$$(\forall x)\neg constitutes(x,x) \tag{COa1}$$

$$(\forall x,y)constitutes(x,y) \rightarrow \neg constitutes(y,x) \tag{COa2}$$

$$(\forall x,y,z)constitutes(x,y) \wedge constitutes(y,z) \rightarrow constitutes(x,z) \tag{COa3}$$

$$(\forall x,y)constitutes(x,y) \rightarrow \exists t(E(x,t) \wedge E(y,t)) \tag{COa4}$$

$$(\forall x,y,z)constitutes(x,y) \wedge P(z,y) \rightarrow \exists x'(P(x',x) \wedge constitutes(x',z)) \tag{COa5}$$

Second, on the link with the dependence relations, following DOLCE's specific and generic constant constitution. We define two relations ; specific constitutional dependence *SCD* (COd1) and generic constitutional dependence *GCD* (COd2). Using the definitions and axioms of dependence in 5.3.1, theorems (COt1-COt4) are implied.

$$(\forall x,y)SCD(x,y) =_{df} \exists t E(x,t) \land \forall t(E(x,t) \rightarrow constitutes(y,x)) \quad \text{(COd1)}$$

$$(\forall \phi,\psi)GCD(\phi,\psi) =_{df} \neg\exists z(\phi(x) \land \psi(z)) \land \forall x(\phi(x) \rightarrow \exists t E(x,t)) \land \forall x,t(\phi(x) \land E(x,t) \rightarrow \exists y(\psi(y) \land constitutes(y,x)) \quad \text{(COd2)}$$

$$(\forall x,y)SCD(x,y) \rightarrow SED(x,y) \quad \text{(COt1)}$$

$$(\forall \phi,\psi)GCD(\phi,\psi) \rightarrow GED(\phi,\psi) \quad \text{(COt2)}$$

$$(\forall x,y,z)SCD(x,y) \land SCD(y,z) \rightarrow SCD(x,z) \quad \text{(COt3)}$$

$$(\forall \phi,\psi,\varphi)GCD(\phi,\psi) \land GCD(\psi,\varphi) \land \neg\exists z(\phi(z) \land \varphi(z)) \rightarrow GCD(\phi,\varphi) \quad \text{(COt4)}$$

Third, we adjoin constitution with a dependence between the relata types. The dependence that we seek for is not specific i.e. the existence of vase ($v1$) does not depend specifically and constantly on that of the instance of clay ($c1$). This is to say that ($c1$) can be replaced with another piece ($c2$) without violating the persistence of $v1$, hence no specific existential dependence of $v1$ on $c1$ ensues. Nevertheless, any other instance $v\#$ of the same type "vase", could not have been artificially created without the presence of some clay, any instance $c\#$ of the type clay. Hence the dependence regarded in constitution is general constitutional dependence (*GCD*) between classes. To represent *GCD*, they types shall be disjoint to ensure that the causal existential connection between instances of the classes comes to an end. While DOLCE asserts generic constant constitution between categories (e.g. $GK(NAPO, M)$ a generic constant dependence between a non-agential physical object and amount of matter), we permit the relation itself to apply *GCD* between the relata of the relation without the obligation of instantiating the types to categories that ensure constitutional dependence. This is done via the axiom (COa6) asserting *GCD* between the relata types and ensuring their disjointedness.

$$(\forall x,y,\phi,\psi)constitutes(x,y) \land \phi(x) \land \psi(y) \rightarrow GCD(\psi,\phi) \quad \text{(COa6)}$$

Fourth, we link constitution to parthood. Since the matter of the constitution relation is taken to be mereologically invariant [Masolo2003], i.e. it changes identity when some parts change, then parts of matter are considered to be essential ones (COa7).

$$(\forall x,y)constitutes(x,y) \rightarrow \forall z(P(z,x) \rightarrow ElementOf(z,x)) \quad \text{(COa7)}$$

5.4 FORT at a macrolevel

Building upon what preceded, the subsequent section introduces the FORT macro-theory, which encompasses the presented group of micro-theories pertaining to various relations. The primary objective is to construct an ontology comprising the selected foundational relations. This entails showcasing the FOL formalizations of each relation alongside their corresponding philosophical and applied ontological interpretations, thereby allowing for potential exploration in different directions.

At a macrolevel, FORT is a modular ontology offering a group of intralinked and interlinked relation micro-theories, also called ontology modules. Figure 5.2 shows FORT's relations, imported relations, and their interlinks as definitions (plain lines) and as axioms (dotted lines) participating in the formalization of one another. In the following, we clarify these characteristics of FORT with highlighting some modeling decisions made.

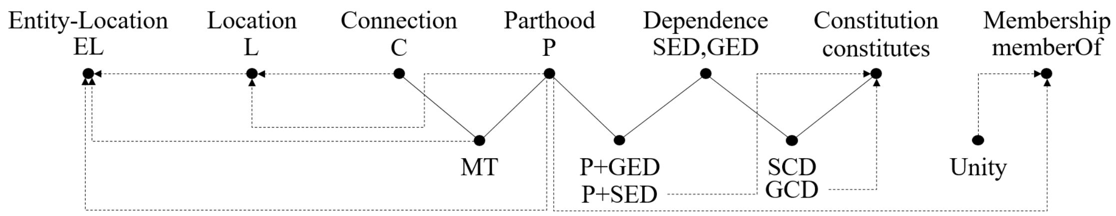


FIGURE 5.2 – *FORT's relations and their interlinks as definitions (plain lines) and as axioms (dotted lines) participating in the formalization of one another.*

First, the notion of intralinked relation micro-theories corresponds to a group of definitions, axioms, and theorems that characterize our view (ontological analysis) of a specific relation. Theoretically, each relation has been thoroughly discussed taking into account its literary ontological and philosophical work. Empirically then, the relation has been formalized by importing, reusing and adapting extant theories according to our ontological analysis and requirements.

Second, the notion of interlinked micro-theories corresponds to links drawn between the relations of each micro-theory using either axioms as is the case with the Entity-Location relation in which the Location relation is used to characterize additionally located entities, or definitions as is the case with the componenthood and elementhood relations which use both parthood and ontological dependence for their definition.

Third, FORT distinguishes the parthood relation from membership and constitution, rather than considering the two latter relations as part-whole relations typologies, in oppositiont to taxonomies of part-whole relations such as the ones presented Section 3.3.2 e.g. the WCH taxonomy [Winston1987], and in Section 3.3.3 e.g. Keet's taxonomy [Keet2008]. This demarcation is adopted to (a) refrain a philosophical debate on the consideration of these relations as part-whole typologies, (b) delimit the scopes upon which transitivity holds, and (c) advocate for the additional semantics that each relation acquires divergent from one another.

Thus, FORT is ***modular*** ([Requirements : 1]), refering to the practice of designing FORT as a set of intralinked and interlinked relations modules. Each module captures a specific relation as a single micro-theory, and together they are combined to create a more complex macro-theory. This manner of breaking down an ontology into more focused modules gives greater flexibility for managing and modifying the modules, and re-usability.

For the characterization of its relations, FORT ***is built exclusively upon relations and rule constraints*** ([Requirements : 3]) i.e. it is entity-type free. This corresponds to the theory not committing to specific kinds, aka categories, for the entities participating in the relation ; the domain and range of relations. Instead, FORT characterizes formally the relations by normalizing constraints on the relata of the relation upon their identification in practice. This is done by introducing axioms that project restrictions on the types that will be allocated for each relation while used in conceptual models, since the theory's utilization is intended as a. This makes the theory straightforward to integrate within extant theories, without obliging the compliance to a hierarchy of entity types.

For the formalization of its relations, the FORT reference ontology which is logically rendered in FOL, is built using abstract and generic modeling primitives, as identified in the modeling language [B] in Section 4.3.1. This yields in FORT being a ***meta***-ontology i.e. at a meta-level ([Requirements : 2]), conceptually specifying a meta-conceptualization using a meta-modeling language.

Lastly, it is possible for one to argue that FORT can be perceived as merely a repository of relations compiled together in a grab-bag ontology. While it is true that FORT incorporates existing theories, such as mereology, mereotopology, and location, and utilizes certain modeling choices from foundational ontologies, FORT goes beyond a mere aggregation.

In FORT, we (1) position FORT within the established body of literature on foundational relations, where the analysis and valorization of extant work have been conducted ; (2) analyze the requirements within the context of our motivation where the selection, re-use, and re-formalization of foundational relations have been rendered ; and (3) provide what preceded within a modular, meta, and entity-type free theory, which underscores the significance of FORT. We firmly believe that this approach distinguishes FORT as a theory rather than a mere repository.

In the work by Gruninger et al. [Grüninger2014], an upper ontology is perceived from a "sideways" perspective as a reducible ontology comprising a collection of generic ontologies, with each generic ontology automating a specific set of generic concepts and relations. This "sideways" view enhances the analysis and design of upper ontologies, simplifies ontology verification by understanding its scope as reduced to its generic modules, and facilitates constructing an ontology by combining multiple generic ontology modules.

Although FORT is not claimed as a foundational ontology, as discussed in Chapter 4, it is considered a meta-ontology in the sense of conceptualizing top-level abstractions of relations. Consequently, FORT aligns with the perspective presented by Gruninger et al. [Grüninger2014], as it focuses on acquiring foundational relations as the primary scope of each generic ontology module (the micro-theories) within a meta-ontology.

5.5 Conclusion

In this Chapter, we have contributed to the applied ontological field by proposing and building a modular foundational ontological relations theory (FORT), as the FORT reference ontology [methodology-step-1]. The theory is formalized using an expressive knowledge representation and reasoning language ; a First-order logic (FOL) formalization.

FORT consists of a minimal set of foundational ontological relations, structural and spatial, that are indispensable for the representation-of and reasoning-over the composition of tangible entities, in general, and a cultural heritage tangible entity in particular. For the composition of tangible entities, several representations are required (as specified in [A] and [B] in Chapter 4) to model the links between ; an entity as a whole and its different inseparable parts (i.e. parthood and dependence) ; an entity as collective whole and the entities that it groups under certain semantics (i.e. membership) ; an entity and its constituents (i.e. constitution) ; an entity and the spatial region that locates it (i.e. location) ; and the spatial constraints among entities (i.e. entity-location). Hence, we are concerned with parthood (and connection), location, membership, and constitution relations, with the formal properties that characterize them. This yielded in a minimal set of generic relations to express the composition of a tangible entity across any domain.

Thus, the FORT reference ontology provides a language of exclusive relations and rule constraints [micro-objective-1] serving as a tool to aid modelers who aim to use foundational relations in practice.

A major limitation of the work is that, up to now, the context is assumed in an atemporal framework, i.e. it captures reality as it exists at a single moment of time. Thus, we do not consider (yet) the behavior of these relations at different times. Though, we use the time predicate to axiomatize some notions that need the intrusion of time to be defined. However, we do not study the evolution of (the relata of these) relations with time neither their conservation of identity while undergoing changes, which is widely discussed subject [Thomson1983].

Nevertheless, the forte contribution points of FORT are twofold. First, FORT is *modular* and addresses *exclusively relations*, yet it normalizes constraints on the relata of the relation. This makes the theory straightforward to integrate within extant theories, without obliging the compliance to a hierarchy of entity types. In addition, it is a *meta*-ontoloy specifying a meta-conceptualization for top-level abstractions, using generic primitives of a meta-modeling language. Thus, FORT covers the basic characteristics specified as [Requirements] for the intended ontology.

Second, FORT offers a minimal set, yet inclusive, of foundational relations that is ample for representing the internal structure, spatial conditions, and interrelations between entities via the selected set of relations. The importing of concrete extant relation theories such as mereology and location derives in FORT being adequate at ontological level with the existing philosophical and formal literature. The selection and reformulation of other relation theories such as membership, constitution, entity-location, componenthood, and elementhood serves the goal of a minimal, yet comprehensive, set of foundational ontological relations.

In conclusion, having presented the FOL theory as an ontology in this chapter, our future direction will remain ontology-oriented. There are various potential paths to explore further with the FOL theory ; however, based on the explanations provided in Chapter 4, we will proceed as follows.

In the upcoming Chapter 6, our focus for comparing FORT with other literature works that offer foundational relations will be limited to those within the scope of ontologies, excluding other philosophical theories. This deliberate choice narrows down the scope of works that will be included in the subsequent chapter, ensuring a more focused and relevant comparison. We will demonstrate our second contribution, addressing the second step of the proposed methodology [methodology-step-2] to cover the second micro-objective [micro-objective-2].

Acknowledgments :

We are indebted to Claudio Masolo and Emilio Sanfilippo for their helpful comments and multiple feedback on earlier versions of this work.

6

On the Analysis and Validation of FORT

This Chapter refers to our contribution "On the Analysis of FORT : arguments, alignment to FOs, and CLIF validation" published as a workshop paper at The Joint Ontology Workshops (JOWO'2022); the 6th Workshop on Foundational Ontology (FOUST VI), August 15–19, 2022.

Sommaire

6.1 Introduction

In the previous Chapter 5, we contributed to the applied ontological domain with a reference ontology of selected foundational relations. The ontology has been presented as a modular macro-theory characterizing and formalizing five, intralinked and interlinked, relation micro-theories.

In this Chapter, we present our second contribution of demonstrating the novelty and consistency of our proposed relations language [micro-objective-2]. This is by analyzing FORT in view of extant ontology relation theories (i.e. a comparison with other theories at same level merely), and providing an additional serialization of FORT to perform consistency checks, i.e. *validating the FORT reference ontology at the empirical aspect* according to [methodology-step-2].

We first present in Section 6.2 the study of foundational relations in foundational ontologies as meta-ontologies at the same level as FORT, and emphasize the primary challenges associated with their use in the sake of foundational relations. Second, we analyze in Section 6.3 FORT in view of these ontologies to demonstrate its novelty. For this, we expound in Section 6.3.1 the arguments for building FORT, and analyze in Section 6.3.2 FORT in view of some foundational ontologies. Third, we proceed in Section 6.4 to associate FORT with a CLIF-serialization as a Common Logic ontology that validates the existence of models for FORT using consistency checks. Lastly, we conclude in Section 6.5.

6.2 Context : foundational relations in foundational ontologies

As demonstrated in Chapter 5, Section 5.3, FORT incorporates certain theories and adopts the perspectives of other meta-ontologies (some foundational ontologies) in making modeling decisions for its formalization. Furthermore, in Section 5.2, we have illustrated in Figure 5.1 the placement of FORT within the context of the existing literature on foundational relations. This positioning highlighted that FORT leverages established theories from philosophical/applied ontology at a micro-level, while functioning as a meta-ontology alongside other meta ontologies, such as foundational ontologies and other top-level ontologies. According to Figure 1.7, which we depicted the ontology abstraction levels, we situate FORT at the meta-level as a meta-ontology.

In order to conduct a comparison between FORT and other relevant literary works to demonstrate its novelty, it is necessary to consider theories operating at the same meta-level. This entails comparing FORT with other meta-ontologies that provide foundational relations, which primarily includes *foundational ontologies and top-level ontologies of relations*. As a preliminary introduction to the upcoming analysis section 6.3, we provide in this section an overview of foundational ontologies from the perspective of "meta-ontologies that offer foundational relations", followed by posing some questions that we aim to answer throughout this chapter.

Foundational ontologies

Foundational ontologies (FOs), aka upper/top-level ontologies, are comprehensive ontological theories comprising both foundational concepts and relations. As referred to in [Guizzardi2006a] following [Guizzardi2007], FOs are "meta-ontologies constructed using the theories developed formal ontology in philosophy" including the study of in areas such as descriptive metaphysics, philosophical logics, cognitive sciences, and linguistics. As mentioned in the Preliminaries section, these include DOLCE (A Descriptive Ontology for Linguistic and Cognitive Engineering) [Masolo2003, Borgo2022a], BFO (the Basic Formal Ontology) [Smith2002], UFO (the Unified Foundational Ontology) [Guizzardi2015, Guizzardi2022], GFO (the General Formal Ontology) [Herre2010], GUM (the Generalized Upper Model) [Bateman2010, Bateman1995], SUMO (the Suggested Upper Merged Ontology) [Niles2001], and YAMATO (the Yet Another More Advanced Top-level Ontology) [Mizoguchi2022] (as the next version of YATO [Mizoguchi2009].

The development of FOs, which deal with broad concepts applicable to diverse subject areas, has presented similar challenges to most applied ontologists. That is finding themselves addressing the same issues precisely as philosophers. The fields of Artificial Intelligence, Knowledge Representation, and Knowledge Engineering within Computer Science have significantly contributed to the advancement of applied ontologies. The exploration of foundational issues has been especially vital in the creation of top-level ontologies, which encompass fundamental aspects such as time [Allen1984] and space [Randell1992]. As presented in the Preliminary Remarks section, the research within the applied ontology domain often requires interdisciplinary e.g. philosophy, linguistics, cognitive science, and computational science. [1]

Most FOs are pretty similar in their fundamental representation aspects [Keet2022] while possibly differing at a philosophical level in their viewpoint to reality and its concrete representation in the ontology under certain modeling and design choices. These are called ontological commitments [Guarino1998a] and encompass philosophical views such as : eternalism, endurantism, actualism, multiplicative, realist, universals, individials, etc. Different FOs adopt different ontological commitments yielding in varying categories and relations hierarchies in the representation of real world entities. [2]

Moreover, FOs are regarded by ontologists from two distinct perspectives : as indispensable elements for ontology development and/or as impractical and burdensome complexities [Keet2011]. From the former viewpoint, FOs play a crucial role in advancing the development of specific ontologies, including core, domain-specific, and application-oriented ontologies. By adhering to a FO and its ontological commitments, the methodologies employed for ontology modeling and development are enhanced. These commitments ensure that models can be verified against an ontological analysis of a subject domain, thereby avoiding the need to reinvent basic categories and relations. Additionally, using FOs for constructing ontological models improves their quality and facilitates semantic interoperability across diverse models. This is particularly useful since FOs serve as meta-ontologies at a meta-level of abstraction, allowing thus the integration of different ontologies based on a common FO within interdisciplinary and multidisciplinary fields.
Conversely, the latter perspective considers FOs as highly abstract, expressive, and excessively comprehensive for relatively simple subject domain ontologies. Furthermore, due to their expressiveness, challenges arise when representing certain attributes, such as attributes and qualia, within the decidable fragments of these specifications, such as OWL implementations.

It is important to note that both of the aforementioned views can be valid in different contexts, depending on various factors. These factors include the element of time, which encompasses two aspects : (a) the time taken to comprehend, to some extent, the different existing FOs and their

1. Please refer to [Borgo2022b] for an overview on the historical perspectives on FOs, as well as the basic philosophical issues addressed by them.

2. Please, refer to [Guarino1998a] and the work in [Guizzardi2007] for the interpretation of the notion of ontological commitment, and it formalization in view of a conceptualization and a modeling language. In our manuscript, we refrain from extensively discussing the different ontological commitments of different FOs since it is not within our scope. However, we do acknowledge these differences and leverage them in Section 6.3 for conducting an analysis of FORT in view of the relations based content of FOs.

respective ontological commitments, which shall lead to the selection of a specific FO, and (b) the time required to become acquainted with the hierarchy of concepts and relations within the chosen FO. Additionally, other important factors appear, such as (c) the level of experience and knowledge of the ontologist or modeler undertaking the task, which significantly impacts the aforementioned time-related factors, and (d) the relevance of the mapping between the subject of the domain ontology and the abstraction provided by the FO. Considering these factors is crucial as they influence the suitability and practicality of employing a foundational ontology in a given context.

Both preceding views can be relevantly occasionally valid depending on different factors. These include the time-factor regarding (a) making a choice of employing a specific foundational ontology after spending some time on the comprehension -to some extent- of the different extant foundational ontologies and their different ontological commitments, and (b) getting familiar with the concepts and relations hierarchy of the chosen foundational ontology. In addition to other factors such as (c) the level of experience and degree of knowledge of the ontologist/modeler making the task which highly affects both preceding time-factors, and (d) the relevance of mapping between the subject of the domain ontology and the abstraction of the FO. [1]

It is noteworthy that certain tools have been developed to aid ontology developers in the selection process of the appropriate FO to employ. Examples of such tools include the ONSET FO selection tool [Khan2012], the BFO classifier [Emeruem2022], and comparative studies like the one conducted by Partridge et al. [Partridge2020].

However, it is important to emphasize that even with the existence of these tools, a minimum level of knowledge on FOs is still required. This includes understanding their ontological commitments and the philosophical distinctions between them. These tools serve as valuable aids, but they do not negate the need for a solid understanding of FOs by ontology developers.

On the use foundational ontologies for foundational relations

Furthermore, since FOs encompass comprehensive theories that encompass both relations and concepts, each FO provides and formalizes a distinct set of foundational relations such as parthood and constitution. While certain relations are commonly addressed across FOs, the specific set of relations varies based on the ontological commitments established by each FO. Aside from FORT, it has become widely prevalent to incorporate foundational relations in practical modeling tasks by using a FO that offers a selection of such relations. Consequently, adopting a FO necessitates adherence to the entire ontology as a complete package, which encompasses its metaphysical perspectives, as well as its hierarchical structure of classes and relationships. Therefore, it becomes feasible to pose inquiries such as the following, to which we intend to provide answers in this chapter.

(a) Why employing/not employing a FO for the use of its set foundational relations ?
(b) If the relations offered in an FO are not sufficient, why not extending a FO with the additional relations depending on the intended application ?
(c) how is FORT novel in view of its imported theories and other meta-ontologies ?
(d) What about the consistency and automatic theorem proofs in FORT as a theory of relations ?

1. On this subject, preliminary empirical experiments considering very basic factors on the effectiveness of the use/no-use of a FO in domain ontology development [Keet2011], and in conceptual data model development for database and software application design [Verdonck2019] have been performed. In [Keet2011], these concluded in better quality ontologies using FOs in comparison to not using any FO.

6.3 Methodology step 2.1 : an analysis

In this section, we analyze FORT in the presence of other meta ontologies offering foundational relations, i.e. we address **(a) FOs and (b) top-level ontologies of relations** as shown in Figure 6.1. First, we expound in Section 6.3.1 the arguments for building FORT. Then, we analyze in Section 6.3.2 FORT in view of extant literature on foundational relations by elucidating an alignment between FORT's micro-theories and the relation-based content of some FOs.

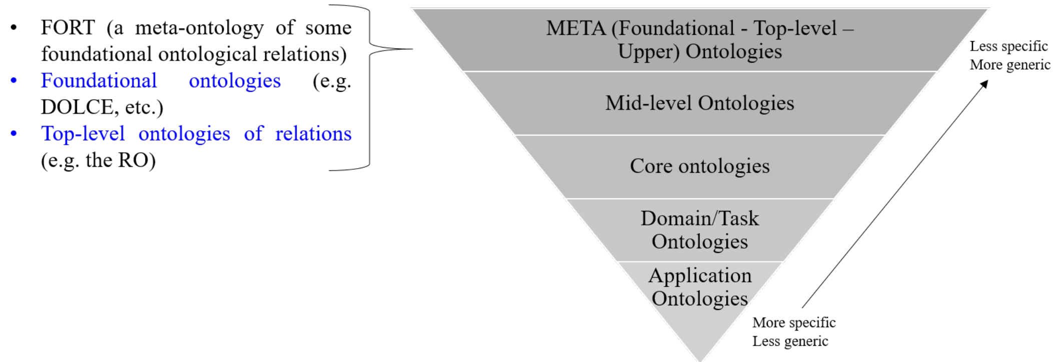


FIGURE 6.1 – *Positioning FORT and other meta-ontologies, including foundational ontologies and top-level relation ontologies, with respect to Figure 1.7 which depicts the different ontology abstraction levels.*

6.3.1 Arguments for FORT

The arguments derive from the scope of our meta-ontology specified in Section [C], and in view of what the literature provides as ontologies of foundational relations.

⋙Argument 1 : "No extant theory of exclusive relations"

This argument addresses two *top-level ontologies of relations* as meta-ontologies.

The Relation Ontology (RO) : is a popular ontology standardizing relationships [Smith2005]. It is one of ontologies used in the Open Biological and Biomedical Ontologies Foundry (OBO) [Smith2007], a community for the development of interoperable ontologies for the biological sciences. Among the several ontologies employed in OBO for sharing vocabularies within an upper integration level, we have : the Core Ontology for Biology and Biomedicine (COB) as the common biological upper layer abstracted under the foundational Basic Formal Ontology (BFO) [Smith2002], the OBO Metadata Ontology (OMO), and the Relation Ontology (RO).

The RO was initially proposed by Smith in [Smith2005] and has been formulated as part of an OWL licensed specification [1]. It serves as a comprehensive collection of relations with the aim of achieving standardization across ontologies within the OBO Foundry. It encompasses a range of relations, including top-level relations such as "part-of," as well as specific relationship types in the biological domain, such as "develops-from".

Although the RO is a relation ontology, it should be noted that it encompasses more than just relations. The relations within the RO are axiomatized in terms to categories, which are classes

1. https://oborel.github.io/.

derived from domain ontologies within the OBO Foundry. For example, the "expressed-In" relation is axiomatized based on the "gene" class as its domain, which is a class from the Sequence types and features domain Ontology (SO) within the OBO Foundry. The range of this relation is specified as the "material anatomical entity" class, which is a class from the Common Anatomy Reference Ontology (CARO) within the OBO Foundry.

Additionally, this triplet of subject-relation-object types is axiomatized (partially) in terms of a subject-relation-object from BFO. In other words, BFO further axiomatizes these relations. This topic is deeply investigated in a conference talk at the Ontology Forum [1] showcasing the structure of the OBO foundry in general, and that of the COB in particular.

TUpper : is another well-known modular ontology of relations [Grüninger2022]. It is a top level ontology with standards being the fourth part of ISO 21838 (Top Level Ontologies) [2]. TUpper offers an alternative approach to the conventional upper ontology paradigms considering an upper ontology to be a modular one composed of generic ontologies, designed based on the "sideways" view of [Grüninger2014] (as explained in Chapter 5 Section 5.4). This is by covering a collection of generic ontology modules incorporating classes relating to time, process, and space, rather than a single taxonomy-centered axiomatization. TUpper is therefore composed of a multiple generic ontologies, called ontology modules [Grüninger2022]. Each of these ontologies adheres to the ontological analysis of an existing standards and adopts it within its formalization. These standards include : ISO 18629 (Process Specification Language), ISO 19150 (Geographic Information Ontology), ISO 80000 (Quantities and Units), and the OWL-Time ontology.

As such, similar to the case of the RO, TUpper uses the classes formalized in these standards to designate categories and further characterize the relations in its ontology modules. For example, the classes *point* and *geometry* are used for formalizing relations such as *in* and *equal*. The CLIF formalization, as well as the OWL axiomatization, of TUpper are available onine as part of the COLORE repository [3].

COLORE, also known as the "Common Logic Repository," is an open repository for first-order ontologies. It has been in existence prior to the development of TUpper, and it serves as an alternative repository alongside TUpper for storing and accessing first-order ontologies. It started in the context of the COLORE project, to develop all the techniques and metatheoretic relationships between theories as discussed in [Grüninger2012].
COLORE serves as a testing ground for evaluating and integrating ontologies, as well as supporting their design, evaluation, and application using FOL. All ontologies within COLORE are specified using Common Logic (ISO 24707), a recently standardized logical language specifically designed for expressing first-order ontologies and knowledge bases.
The fundamental organizational principle in COLORE revolves around the concept of a hierarchy, which consists of a collection of ontologies sharing the same signature. [4] In other words, ontologies are treated as logical theories within this framework. They are used for the verification of upper ontologies, as well as the application of the generic ontologies within COLORE for the specification of mappings between upper ontologies and the design of new upper ontologies [Grüninger2009].

1. https://ontologforum.org/index.php/ConferenceCall_2023_01_25.
2. https://www.iso.org/standard/78928.html.
3. https://github.com/gruninger/colore/tree/master/ontologies/tupper
4. Examples of these ontologies including mathematical ontologies e.g. the arithmetic and algebra ontologies, FOs like some of DOLCE's modules e.g. constitution and participation, and other generic ontology modules that axiomatize generic domains such as time, process, space, and shape.

Thus, both the RO and TUpper serve as comprehensive repositories for ontological relations, aiming to standardize the representation of relations within specific domains (RO for biology and TUpper for a broad range of conceptualizations including time and space). However, neither of these ontologies can be considered exclusive theories of relations, which is a requirement of our specified set of [Requirements]. Moreover, to the best of our knowledge, there are no other existing meta-ontologies of foundational ontological relations apart from TUpper and the RO. Based on our analysis, it can be concluded that there is currently no meta-ontology available that fully satisfies our specified set of requirements for foundational ontological relations.

⋙**Argument 2 : "No theory incorporates inclusively the specified set of relations"**

This argument addresses *FOs as meta-ontologies* providing foundational relations. It draws attention to a significant issue that has not (yet) been addressed in existing FOs : none of the current FOs encompass the complete set of intended foundational relations. In the following discussion, we elaborate on this claim by examining popular FOs such as BFO, DOLCE, and UFO. For each ontology, we emphasize the foundational relations that it does not address, but are of particular interest to us, as identified in [B], and addressed in FORT.

BFO does not (and cannot) express constitution due to its reductionist view. BFO is a realist ontology capturing the world as (multiple) particular perspectives of reality i.e. possibly multiple instantiations of the same particular individual. For constitution, BFO regards the entities participating in a constitution relation, e.g. the vase and the clay, as the same spatio-temporal individual that instantiates different universals at the same space-time location i.e. an identity relation instead. Such an argument is added to BFO's formalization using an axiom that prevents two material entities from occupying the same spatial region unless they coincide i.e. the relations collapse to an identity. Now if constitution is to be represented in BFO, then it will be a relationship taking place between an individual and itself as an instance of two different universals i.e. individual ID1 as an instance of a statue and the same ID1 as an instance of clay. This thus requires the fact that individual ID1 be instantiating two classes at the same time instance t. Let's say BFO succeeds in representing such restrictions on a constituted relation. What about the specificity vs generality of this constitution ? In other words, is it a constitution that is applied specifically to these to individuals ? Or is it a constitution that takes place between any instance of the category of the constituent entity (any instance of clay) and any of the constituted entity (any instance of statue) ? In the latter case, the representation of constitutional dependence (generic) between the constituent and the constituted entity, which is an important axiom in the characterization of the constitution relation, will still be unachievable. This is because constitutional dependence requires the two class types to be disjoint to ensure that the causal existential connection between instances of the classes comes to an end. Thus, constitution remains to be problematic for BFO.

DOLCE [Masolo2003], on the other hand, does not explicitly define a locative relation nor does it adopt a specific location theory such as that of Varzi. Instead it offers locative representations via qualities and quales as will be explained in section 6.3.2. In view of such a comprehensive treatment of relations, membership is not (yet) covered within DOLCE's specification. Although in the earlier analysis of building DOLCE, in [Gangemi2001] which presented a preliminary work for DOLCE as a systematic methodology for selecting and defining some general ontological categories and relations, the authors have already considered the treatment of the membership relation. The analysis has been induced in terms of parthood in the spirit of the analysis presented in [Guarino2000b], with an interpreting a unifying relation that binds all members of a whole (collective/aggregate) together and a maximality constraint on the members with respect to his relation. However, no consideration of membership has (yet) been shown in DOLCE.

Similar to DOLCE, with some differences in their treatment of regions, UFO does not integrate a theory of location relations. Instead it presents an entity-type based approach of *attributes* and *attribute value spaces* following [Guizzardi2006b]. Such an approach can cover the locative representation between regions, and between entities (i.e. material ones) and regions. However, what about representing locative relations between entities that are not spatial regions ? For example, consider two entities that occupy a shared spatial region. On the one side, these entities are not parts and do not share any parts i.e. mereology is insufficient. On the other side, these entities are not exactly located in the same spatial region i.e. the "being exactly located at" primitive of the Varzi' location theory is also insufficient. There is a need for a module other than mereology or basic region locations, such as containment [Bittner2004a] and inclusion theories [Donnelly2006], for clarifying and determining the spatial information embodied between entities in ontologies and enhancing their automatic reasoning.

Consequently, each of the examined FOs possesses a distinct and formalized set of relations. While certain relations may be commonly addressed across FOs, the specific set of relations varies based on the ontological commitments made by each ontology. Thus, considering the minimal set of relations necessary for our representation, none of the existing FOs inclusively incorporates the specified set of foundational relations. Additionally, this observation underscores that the relations in which we are interested have been sparsely and exhaustively addressed in the literature. This highlights the significance of each relation as an ontological foundational relation and emphasizes the need to include each relation in a fundamental set of relations, as discussed in Section 4.3.1.

⋙ At this point, one might raise the question of "why not simply extend an extant FO with the missing relation(s) from our targeted set of relations ?" Indeed, such a task might exceed the empirical and theoretical flexibility of a theory to comprise a specific relation. Earlier in this thesis, we tried to adopt an approach of extending a foundational ontology with a relations theory, taking BFO as a case-study to extend with the constitution relation mainly and some other typologies of parthood. For that, we have built a taxonomy of relations as part-whole typologies using BFO categories. However, this approach resulted in violating theoretically the ontological commitments made by BFO, and empirically, the user-intended representation and inferences where the constituent and the constituted entity are regarded as identical i.e. the same individual. It is possible now to claim this is a particular case with only BFO being problematic. But what about other FOs ? Why does DOLCE not account for a membership theory ? Why does UFO not consider a comprehensive theory of location ? The question here goes beyond "why not extend an extant FO" to "can extant FOs accommodate for additional relations needed in practice". If yes, then why is there not any consideration yet ?

⋙**Argument 3 : "The difficulties upon the adoption and employment of a FO"**

In this argument, we delve into the challenges that modelers encounter when adopting a FO *specifically for the use of foundational relations*. It is important to note that this argument does not seek to dispute the benefits of FOs in ontology development, as they have been shown to enhance semantic interoperability, guide ontology development, improve the quality of systems, etc. [Borgo2008, Keet2011, Keet2018]. However, we are solely addressing the difficulties that modelers face when employing a FO and the reasons why they may choose not to use a FO.

The first difficulty is the responsibility that a FO choice incorporates towards understanding the different ontological and philosophical assumptions made in each FO. With the numerous FOs available today, the modeler having to choose a FO for practice has to understand : what does each FO offer differently from one another ? what are the theoretical modeling decisions made in each FO, i.e. the ontological commitments that correspond to the world view it acquires ? how are these ontological commitments empirically translated into formalizations ? Performing such a comparison between FOs is a difficult task since most differences occur at a high meta-physical [1] level which requires deep philosophical understanding. After having this comparison, the modeler has to elect a FO for practice according to the needs of the application domain. This involves answering the following : how can the modeling dilemmas present in the application domain be represented in each FO ? how are these representations different, and what are the conveniences offered by each ? Only if the modeler can answer all these questions, which require massive work and time efforts, then the proper justification of the choice of FO can be made. [2]

The second difficulty is the entire commitment to the assumptions made in the chosen FO regarding its world-view. Upon employing a FO, the modeler is committed to its view on the elements of reality that is wholly interpreted in its formalization. This can be problematic for the modeler when the modeling dilemma ought to be modeled, cannot be fully represented by the chosen FO due to the fact that a specific assumption made by it, rejects this view in the dilemma. In case the assumption in the FO is ignored, the representation will yield in contradicting semantics. And in case all assumptions are complied to, then the representation of the dilemma is not fully achieved. Knowing that the list of ontological commitments varies from one FO to another, one might want to comply with one assumption from one FO e.g. BFO's [Smith2002] reductionist view (rejects for the co-existence of two entities in the same space-time location), and another assumption from another FO (that the first, BFO, rejects) e.g. DOLCE's [Masolo2003] parthood theory (general extensional mereology GEM) to solve the modeling dilemma he has. This representation is not plausible, and the modeler needs thus to presuppose the requirements of the modeling problem in order to make the best choice of FO with the least missing representations.

The third difficulty is the obligation to adhere to the categories-hierarchy of the chosen FO. This does not only oblige modelers to map the domain-specific types (aka kinds or categories) in their models to the top-level categories in the chosen FO, but also to comply to the constraints that each type acquires. For example, the category "Non-Agentive-Physical-Object" in DOLCE is disjoint from the category "Amount-of-Matter", and requires that any instance of the former be generically constantly constituted by an instance of the latter by the axioms characterizing both categories. Thus, all categories that are mapped to any of the preceding two, will acquire additional inherited axioms and semantics that the modeler shall comply with. Thus, several obstacles arrive with FOs in general, and with the commitment to category hierarchies in particular.

Therefore, when considering the use of a FO solely for its foundational relations, one may raise concerns regarding the difficulties encountered by modelers in employing these relations.

1. According to Merriam Webster, metaphysics is "*a division of philosophy that is concerned with the fundamental nature of reality and being and that includes ontology, cosmology, and often epistemology*".

2. An initial step towards the establishment of library of FOs is the Repository of Ontology for MULtiple USes (ROMULUS) [Khan2016], which specifically focuses on the OWL serializations of FOs [Khan2015]. Building upon ROMULUS, a tool called ONSET [Khan2012] has been developed to assist modelers in selecting the most appropriate FO. The selection process relies on the inputs provided by the developer, upon which the tool generates an automated explanation to support the proposed FO. However, the developer's inputs are responses to questions that are categorized into five distinct categories, one of which is "ontological commitments". This category encompasses all the philosophical choices that the developer must make, including decisions regarding universals or particulars, descriptive or realist approaches, multiplicative or reductionist perspectives, endurantism or perdurantism, and so on.

⋙**A synthesis of arguments**

As a result, the construction of FORT serves the purpose of fulfilling the requirement for the development of an ontology that is based on exclusive relations (argument 1). Furthermore, it highlights the importance of having a theory that specifically focuses on pure foundational relations, separate from large and complex foundational ontologies (argument 2). This aspect can be particularly beneficial for modelers who intend to utilize foundational relations without encountering the challenges associated with theories that only partially emphasize relations in addition to taxonomic category axioms (argument 3).
Our proposal, presented in Chapter 5, addresses this issue by introducing a modular theory that encompasses a minimal set of foundational ontological relations (FORT). This theory is significant for the representation and reasoning of the composition of tangible entities.

6.3.2 FORT in view of extant theories

After presenting the arguments that have led to the development of FORT, aligned with the objectives of our thesis and the specified requirements, it is essential to demonstrate the novelty of this approach in relation to other meta-ontologies for comparison. In theory, we shall include in this comparison : RO, TUpper, DOLCE, BFO, UFO, GFO, GUM, SUMO, and YAMATO. However, based on the following considerations, we have sifted only three main ontologies for comparison.

Regarding the RO, as discussed in Section 6.3.1, it forms an integral part of the highest integration level of the OBO Foundry, along with BFO and several core ontologies. Furthermore, BFO and RO are commonly used together and are viewed as complementary ontologies. The BFO repository on GitHub offers the second latest version of the ontology in the form of BFO-RO (and OWL file). This ontology version presents a hierarchy of classes and relations offered by BFO, along with additional complementary relations from the RO hierarchy. Importantly, the domains and ranges of these RO relations align with the classes defined in BFO. Notably, the latter set of relations specializes from the former [Arp2015]. Thus, RO is considered within BFO.

While TUpper [Grüninger2022] initially appeared to be a relevant candidate for this comparison, the generic ontology modules considered in TUpper are significantly similar to those found in DOLCE. In fact, most of these modules' axiomatizations correspond closely to those in DOLCE, upon which evidenced could be inspected in the COLORE repository.

As for some other foundational ontologies (DOLCE, BFO, UFO, GFO, GUM, SUMO, and YAMATO), as mentioned earlier in 6.2, most FOs are similar in their fundamental representation aspects but can differ at a philosophical level regarding their ontological commitments [1].

For instance, GFO [Herre2010] serves as a foundational framework for ontology development, upon which UFO is built. UFO expands the coverage of GFO by incorporating more specific ontological categories such as events, roles, and goals. Furthermore, UFO provides a more formal and logical foundation for the ontology while emphasizing the role of natural language in ontology development and utilization. Consequently, if UFO is included in the comparison (which is the case), GFO is implicitly considered as part of the comparison as the foundational framework for UFO. [2]

1. Please consult [Borgo2022b] for a comprehensive survey of existing FOs. This survey provides a concise introduction to each FO, including its historical background, and presents fundamental distinctions among FOs based on their ontological commitments, illustrated with relevant examples.
2. It is probably noteworthy and assuring that a mediation effort between GFO, BFO, and DOLCE has been conducted in [Khan2015] as to facilitate interchangeability, which illustrated the similarities among these three FOs. For an ad-

With DOLCE, BFO, and UFO being well-known FOs, and inherently including (part of the) views of TUpper, the RO, and GFO respectively, we have opted to include them as primary meta-ontologies in our comparison with FORT. This choice is particularly justified due to the notable resemblances between FORT and these three ontologies as will be illustrated. Thus, in the following, we elucidate a relation-based comparison between each micro-theory in FORT and the corresponding consideration of the relation made in BFO [Smith2002], DOLCE [Masolo2003], and UFO [Guizzardi2015].

Figure 6.2 portrays a visual representation where the first column, highlighted in blue, showcases the micro-theories imported by FORT, denoted by solid squares, as well as the partially reused micro-theories, represented by dotted squares.

Subsequent columns, highlighted in gray, present some FOs, each accompanied by their relation-based content, specifically those that align with the micro-theories within FORT, if applicable.
Regarding the imported micro-theories of FORT (solid cells in the first column), these encompass Classical Extensional Mereology (CEM) theory for parthood, ground Mereotopology (MT) for the link between parthood and connection, and a portion of Varzi's location theory for entity-to-region locative relations. Each micro-theory is explicitly labeled above its respective representation.

Regarding the partially reused micro-theories of FORT (dotted cells in the first column), these encompass dependence, membership, constitution, and dependence with constitution. For each of these micro-theories, within the same row, an additional dotted cell is present. This cell corresponds to a micro-theory from another meta-ontology, which shares significant similarities with the corresponding micro-theory in FORT.

		CEM		MT	Varzi				
FORT	Dependence	Parthood	Parthood + Dependence	Parthood + Connection	Entity-to-Region	Entity-to-Entity	Membership	Constitution	Constitution + Dependence
BFO	s-depends-on	own mereology	-	-	occupies-SR	located-in	member-of -U, + Aggregate	-	-
DOLCE	SD & GD …	GEM	-	-	(qualities)	-	-	k	SK & GK …
UFO	ed, ind & gfd	GEM	-	-	(attributes)	-	member-of -U, + Collection	constituted -by	GCD

FIGURE 6.2 – *A relation-based comparison between the micro-theories in FORT and the corresponding consideration of these relations in the some FOS : BFO, DOLCE, and UFO. Some micro-theories of FORT are highlighted using : solid lines denoting those that import theories, or dotted lines denoting those that re-use some of the relation's formalization in the highlighted FOs.*

Ontological dependence

The dependence relation is generally defined in terms of an existence primitive relation, also referred to as "being present". The existence relation is pretty similarly presented in all FOs : *existsAt* in BFO, *PRE* (being present at) in DOLCE as binary predicates between an entity and an instance of time, and *ex* in UFO as a unary predicate due to a single time slice formalization.

ditional analysis of the theoretical ontological alignment between BFO and DOLCE, readers are referred to [Seyed2009] and [Temal2010].

In BFO, ontological dependence is introduced in the form of entity types that are dependent on other entity types i.e. $SpecificallyDependentContinuants$ such as : qualities, functions, roles, and dispositions, and $GenericallyDependentContinuants$. This dependence is carried out via relation `s-depends_on` which has as a domain a $SpecificallyDependentContinuant$ and as a range either an $InependentContinuant$ or $Process$ in case of a *one-sided s-dependence*, or a $DependentContinuant$ in case of a *reciprocal s-dependence*.

This view however omits other continuant entities that are not qualities, dispositions, or roles, but still can be dependent in their existence on other entities within a certain context. For instance, it can be asserted that the presence of a brain is indispensable for the existence of a human body, given that no human body has ever come to birth without a brain. Depicting this scenario necessitates the consideration of both entities involved in the dependence relation as material entities. Within the framework of BFO, the representation of such a case becomes somewhat ambiguous and requires the inclusion of the brain's functional state as a material entity and the designation of the human body's state as being alive, followed by the assertion of dependency between the two states. It is convenient to have dependence as a relation between entities that are not explicitly defined within a particular type, encompassing qualities, functions, or other specific categorizations.

In UFO, in addition to existential dependence $ed(x,y)$ and independence $ind(x,y)$, functional dependence is also studied in its generic $gfd(\Phi,\Psi)$ and individual $ifd(x,\Phi,y,\Psi)$ -defined in terms of $gfd(\Phi,\Psi)$- forms, according to the treatment of functional parts by Guizzardi in [Guizzardi2009], as mentioned in Section 3.3.3 of Chapter 3.

The investigation of the relationship between functionality and dependence has been extensively studied, with a substantial body of literature dedicated to this subject. Notably, Vieu and Aurnague [Vieu2007] have examined functional concepts within certain typologies of the part-whole relation, resulting in characterizing the formally specified functional dependence.
However, the endeavor to characterize and formally specify functional dependence is subject to intense philosophical debate, particularly concerning the identification of functions. It should be noted that an entity can encompass multiple functions simultaneously. For instance, a car can serve various purposes such as movement, support (as a means of staying grounded), stopping, breaking, and more. The functionality associated with an entity can be based on its design, as identified by its creator, or even the most prevalent functionality at a given moment in time. Thus, determining the functional dependency between entity x and another entity y poses a challenging task.
Moreover, even if we are able to identify all the potential functionalities that an entity can possess at a specific point in time (e.g., entities capable of movement, entities that burn fuel, etc.), it remains a complex undertaking to (1) instantiate individual entities using these categories and (2) identify the dependencies that exist among these functionalities.

DOLCE introduces two types of dependence relations : specific between particulars $SD(x,y)$ or universals/properties $SD(\Phi,\Psi)$, and generic between universals/properties $GD(\Phi,\Psi)$. Using these two types, other forms are also defined e.g. one-sided and mutual dependencies.

Similar to DOLCE's account for dependence, FORT analysis two types of the relation, namely existential dependence in two forms, specific $SED(x,y)$ and generic $GED(x,y)$, however both treated as single-directed axioms within a non-modal approach, rather than definitions as in DOLCE. This is less similar to UFO's approach, and completely dissimilar from that of BFO which does not meet our interpretation as explained in section 6.3.1.

Parthood

FORT adopts CEM, while DOLCE and UFO both adopt GEM, and BFO constructs its own mereology using its "continuant-part-of" relation. FORT further accounts the combination of parthood with dependence to introduce the notion of inseparable parts (essential and mandatory) : elements and components. This is plausible in DOLCE since the primitive relations (parthood and specific/generic existential dependence) exist. In UFO, it is also feasible using the existential dependence relation to introduce essential parts (aka elements in FORT) but not mandatory parts (aka components in FORT). However, UFO does define dependent parts as components using the $componentOf(x,y)$ relation that is proper parthood $PP(x,y)$ accompanied with a restriction on the individual functional dependence of the whole on the part $ifd(x,x',y,y')$.

As for the mereotopological aspect, none of the considered FOs imports a topological theory to account for the connection relation, whereas FORT imports minimal (ground) mereotopology i.e. the primitive connection relation with its corresponding topological predicates.

Location

— region-to-region locative relations : Without committing to a region entity type, FORT suggests the use of the primitive parthood relation $part-of$ to express mereotopological representations between entity types that can be regions. This is similar to BFO's utilization of its own mereological relation $continuantPartOf$ between spatial regions, and DOLCE's and UFO's use of the primitive mereological relation $PartOf$.

— entity-to-region locative relations : In FORT, we import Varzi's location theory [Casati1999], which in turn links to the imported mereotopological theory. While BFO uses the $occupies-SpatialRegion$ relation to express the spatial region that an independent continuant entity is acquiring, both DOLCE and UFO adopt different views for location representations via qualities (in DOLCE) and attributes (in UFO). In DOLCE, in terms of quality types and quales, location can be described as a scenario encompassing : (a) a quality type, which in the case of expressing a location, is the spatial location of an entity e.g. $SL1$ as a instance of the class $SpatialLocation$ which is a subclass of $PhysicalQuality$, which is a subclass of $Quality$, (b) a quale i.e. the spatial region which the entity is covering e.g. $SR1$ which is an instance of the $SpatialRegion$ which is a subclass of $PhysicalRegion$, which is a subclass of $Region$, and (c) a relation Qlt that links both the quality type and it corresponding quale i.e. links $SR1$ to $SL1$, at a specific time. Whereas in UFO, a similar representation is done using attributes and attribute value spaces.

— entity-to-entity locative relations : these are expressed in FORT using the entity location EL primitive relation indicating "located at/on/in". BFO in turn uses a simple primitive relation $located-In$ between two independent continuants that are not spatial regions, while both DOLCE and UFO do not account for such a representation.

Membership

FORT examines membership $memberOf(x,y)$ as a primitive relation that is defined in terms of ground axioms with the characterization of its whole entity. Although FORT does not account for entity types, it requires restriction on the range of the membership relation to be unified $U_{R_i}(y)$, through its members, by a unification relation $R_i(x,x)$. The axiomatization provided in FORT is similar to that in BFO except that BFO uses a class type $AggregateEntity$ to characterize the whole while FORT does not, and BFO does not oblige the binding of all members according to a similarity constraint i.e. unifying relation. UFO also provides a membership primitive

$memberOf(x,y)$ that holds between an object and a collection following the preliminary analysis in [Guizzardi2005] and [Guizzardi2015]. However, for DOLCE, membership is not (yet) considered although the notion has been addressed in the preliminary studies for ontological distinctions as mentioned in section 6.3.1.

Constitution

FORT treats constitution in a very similar manner to that in DOLCE and UFO. Using a constitution primitive $constitutes(x,y)$ ($K(x,y)$ in DOLCE, and $constitutedBy(y,x)$ in UFO) along with defining specific and generic constitutional dependencies $SCD(x,y)/GCD(\Psi,\Phi)$. However, in contrary to DOLCE's concept $AmountOfMatter$, FORT does not restrict types but applies additional axiomatization on the relata of the relation. For BFO, as discussed in section 6.3.1, constitution is regarded to be identity.

6.4 Methodology step 2.2 : a CL-ontology [$FORT_{CL}$]

In this section, we demonstrate the consistency of the proposed FORT reference ontology by providing FORT within a logical rendering that allows running consistency checks. Indeed, for each FO, this has been done using various serializations like Common Logic [ISO/IEC247072007] (e.g. used for DOLCE, BFO, and TUpper), Alloy [1] [Jackson2012] (e.g. used for UFO), Prover9 [2] [McCune2007] (e.g. used for BFO) or TPTP syntax [3] [Sutcliffe2017] (e.g. used for UFO) serializations.

For FORT, we construct a Common Logic (CL) [ISO/IEC247072007] ontology that validates the existence of models using consistency checks. CL is a logical language based on FOL, with the purpose of standardizing syntax. It further extends FOL for that (1) any term can be used as function or predicate, and (2) sequence markers allow for talking about sequences of individuals directly, and in particular, provide a succinct way for axiomatising polyadic functions and predicates [Mossakowski2014]. CL is used in ontological theories as a formal language tool to prove the consistency of a theory by validating the existence of model(s) `M` for a theory `T`. For example, BFO [4] and some of DOLCE's modules [5] are encoded in CLIF. In addition to FOs, "`COLORE`", the

1. Alloy is an open-source language and analyzer designed for software modeling. It utilizes a simple structural modeling tool based on first-order logic and aims to create micro-models that can be automatically checked for correctness. The Alloy Analyzer is used to verify Alloy specifications. Unlike many other specification languages, Alloy allows the definition of infinite models, making it unique in its approach.
2. Prover9, a theorem prover developed by William McCune, is widely recognized for its capabilities in automating first-order and equational logic. Notably, Prover9 is distinguished for generating proofs that are relatively easy to comprehend and for its robust hints strategy [Phillips2008]. Additionally, Prover9 is intentionally complemented by Mace4, a companion tool that specializes in searching for finite models and counterexamples. The two tools can be executed concurrently from a shared input [Berghammer2010], with Prover9 focusing on finding proofs and Mace4 striving to identify counterexamples. The implementation of Prover9, Mace4, and several other tools leverages the LADR ("Library for Automated Deduction Research") library, which simplifies the development process.
3. The TPTP (Thousands of Problems for Theorem Provers) problem library is a widely recognized and standardized collection of test problems specifically designed for first-order automated theorem proving (ATP) systems. It serves as a standard benchmark for evaluating the performance of different ATP systems and facilitates effective communication among researchers in the field [?]. The TPTP language itself offers a convenient means of expressing both problems and solutions, making it a unique and valuable tool for ATP. The simplicity of its syntax further enhances its usability and effectiveness as an ATP language.
4. https://github.com/BFO-ontology/BFO-2020/tree/master/21838-2/common-logic
5. https://github.com/gruninger/colore/tree/master/ontologies

open-access repository, is implemented as CL-ontology modules [1].

To write the CL ontology, we employ the CLIF serialization ("`CLIF`", the Common Logic Interchange Format) which is specifically designed to represent ontologies and logical theories in a standardized and machine-readable manner. By selecting CLIF, we ensure compatibility with a wide range of reasoning systems and tools that support CL, which can facilitate consistency checks and automated theorem proving. CLIF provides a formal and unambiguous syntax, allowing for clear and precise representation of your theory. Its expressive power enables the encoding of complex relations and logical constructs, ensuring that the nuances of your ontology are accurately captured. Furthermore, CLIF supports a range of logical fragments and features, such as quantification, modalities, and nonmonotonic reasoning, which may be crucial for representing certain aspects of your theory.

6.4.1 A CLIF-serialization for FORT

Considering that FORT is a group of micro-theories, also called ontology modules, importing and reusing some extant theories from the literature, firstly we import those that are already encoded and available online at the "`Colore`" repository. These are : the CEM mereological theory [2], the MT mereotopological theory along with the basic connection topological theory [3], Varzi's Location theory [4], and their corresponding definitions. Note that, we modified some files e.g. the mereological definitions file to account for additional definitions e.g. overcross, undecross, etc, the location root file to remove region axioms, etc. We show below in Listing 6.1 the CLIF serialization of the primitive *existing* relation, and that of the imported CEM theory from the COLORE repository in Listing 6.2.

```
(cl-text https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-
    ontology/existing.CLIF

(cl-comment 'existing relation is asymmetric')
(cl-comment 'Identifier: FORT_')
(forall (x t)
  (if (existing x t)
    (not (existing t x) )
  )
)
)
```

Listing 6.1 – The "existing" relation in CLIF.

```
(cl-text https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-
    ontology/colore/mereology/cem_mereology.CLIF

(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/colore/mereology/cm_mereology.CLIF )

(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/colore/mereology/em_mereology.CLIF )

```

1. https://github.com/gruninger/colore/tree/master/ontologies
2. https://raw.githubusercontent.com/gruninger/colore/master/ontologies/mereology/cem_mereology.clif
3. https://raw.githubusercontent.com/gruninger/colore/master/ontologies/combined_mereotopology/mt.clif
4. https://raw.githubusercontent.com/gruninger/colore/master/ontologies/location_varzi/L_location.clif

```
)
```

Listing 6.2 – The imported CEM theory.

Secondly, we initiate the serialization of FORT's micro-theories in the sense of "what comes first", i.e. starting by the basic primitive relations that do not necessitate the use of other primitives, and moving on to relations that import other relations in their modules. The full CL formalization of the theory is available online on GitHub repository : FORT. We present below the code serialization of the $component - of$ (Listing 6.3) and the $member - of$ (Listing 6.4) relations, as examples.

```
(cl-text https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-
    ontology/componentOf_definition.CLIF

(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/colore/mereology/cem_mereology.CLIF )

(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/dependence_definitions.CLIF )

(cl-comment 'x is a component of y iff x is a part of y and y is
    generically existentially dependent on x')
(cl-comment 'Identifier: FORT_Pd1')
(forall (x y)
  (iff  (componentOf x y)
      (and (part x y)  (exists (PSI PHI) (and (PSI y) (PHI x) (GED (PSI y)
    (PHI x)) ) )
      )
  )
)
(cl-comment 'componentOf is a proper part of relation')
(cl-comment 'Identifier: FORT_Pa1')
(forall (x y)
  (if (componentOf x y)
      (ppart x y)
  )
)

)
```

Listing 6.3 – The CLIF serialization of the $component - of$ relation.

```
(cl-text https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-
    ontology/memberOf_root.CLIF

(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/colore/mereology/m_mereology.CLIF )
(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/colore/mereology/definitions/mereology_definitions.CLIF )
(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/unified_entity_definition.CLIF )
(cl-imports https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-
    CL-ontology/existing.CLIF )

(cl-comment 'Identifier: FORT_Ma1')
(forall (x)
  (not (memberOf x x) )
)
(cl-comment 'Identifier: FORT_Ma2')
(forall (x y)
  (if (memberOf x y)
    (not (memberOf y x) )
```

```
  )
)
(cl-comment 'Identifier: FORT_Ma3')
(forall (x y)
  (if (memberOf y x)
    (exists (t m n)
      (and (existing x t)
         (not(= m n))
         (memberOf m x)
         (memberOf n x)
      )
    )
  )
)
(cl-comment 'Identifier: FORT_Ma4')
(forall (x p y)
  (if (and (memberOf y x) (ppart p x) )
    (exists (o)
      (and (memberOf o x) (overlaps o p) )
    )
  )
)
(cl-comment 'Identifier: FORT_Ma5')
(forall (x y)
  (if (memberOf x y)
    (and (ppart x y)
       (forall (m)
         (if (memberOf m y)
           (or (= x d) (not(overlaps m x)) )
         )
       )
    )
  )
)
(cl-comment 'Identifier: FORT_Ma6')
(forall (x y)
  (if (memberOf y x)
    (forall (w)
      (iff (overlaps w x)
         (exists (m)
          (and (memberOf m x) (overlaps w m) )
         )
      )
    )
  )
)
(cl-comment 'Identifier: FORT_Ma7')
(forall (x y)
  (if (memberOf y x)
    (exists (R)
      (and (U (R x x))
         (forall (t) (if (existing x t) (U (R x x)) ) )
      )
    )
  )
)
)
```

Listing 6.4 – The CLIF serialization of the $member-of$ relation.

6.4.2 Automatic translations into TPTP, LADR, and CASL

Besides a CLIF serialization, we translated of FORT into the LADR syntax [1] which is the format required by Prover9 [McCune2007], the TPTP format [Sutcliffe2017], and the CASL syntax [2] [Astesiano2002,Bidoit2004], to offer multiple serializations of CL. For performing automatic translation, several tools exist such as the following.

The CLtools repository : It has been adopted for a theory's translation to LADR syntax [3]. However, no maintenance of the tool or resolving of its issues has been carried out in the last decade.

The Macleod environment : In [Katsumi2011], an ontology development environment in the Common Logic (CL) language called "Macleod" is proposed. The current implementation of Macleod is available as a collection of Python scripts. This tool offers several functionalities, including the translation of a CLIF file to TPTP and LADR formats, extraction of an OWL approximation of a CLIF ontology/module, verification of logical consistency, verification of non-trivial logical consistency, and theorem/lemma proving based on the ontology or module.

In order to use the Macleod tool, we made necessary modifications to configuration files and adapted some scripts. The modified scripts, along with the procedure followed, can be found online in the FORT GitHub repository, specifically in the Macleod and virtual environment folders [4].

During our usage of the Macleod tool, we encountered a bug in the current version. In certain ontology modules of FORT, quantification over unary/binary predicate names is employed, such as universal quantification on a unary predicate (concept) in the generic existential dependence axiom Da5, and existential quantification over a binary predicate (unifying relation) in the unification axiom for the range of the membership relation Ma7. While the CLIF syntax, which is a variant of CL, supports such quantification where properties can range over unary and binary predicates, Macleod was unable to parse these CLIF files containing such quantifications or perform the intended translations and consistency checks. However, for files without such quantifications, the Macleod tool functioned correctly, allowing us to generate TPTP and Prover9 translations for the corresponding ontology modules. Therefore, it is crucial to address and fix this issue in the current implementation of the tool to fully support the CLIF syntax.

The Hets tool : While a detailed evaluation of the tools supporting CL showed a major calling for coverage in [Mossakowski2014], the "The Heterogeneous tools set" (Hets) [Mossakowski2005, Mossakowski2007, Mossakowski2013], which existed before the survey, was extended with parsers for CLIF, enabling CL support.

Hets is a versatile tool that encompasses parsing, static analysis, and proof management functionalities, integrating various provers and specification languages. It includes logic-specific tools for parsing and static analysis of basic logical theories written in different logics, as well as a logic-independent parsing and static analysis tool for structured theories and theory relations. Hets offers automatic translations of theories from multiple input languages (e.g., DOL, CASL, OWL, CLID) to various serializations. Hets can be used via a web-based interface [5].

1. As mentioned earlier, LADR ("Library for Automated Deduction Research") is the syntax required by Prover9 and Mace4. An online forum for Prover9, Mace4, and LADR can be found at https://web.archive.org/web/20081230011330/http://forums.prover9.org/.
2. The CASL (Common Algebraic Specification Language) is an expressive specification language that has been designed to supersede many existing algebraic specification languages and provide a standard. CASL consists of several layers, including basic (unstructured) specifications, structured specifications and architectural specifications ; the latter are used to prescribe the modular structure of implementations.
3. http://www.cs.toronto.edu/~torsten/DCT-BCont/
4. https://github.com/DanashFatima/FORT
5. http://rest.hets.eu/

⋙ In comparison to other available tools, we have chosen to use Hets for several reasons. It offers extensive support for various specification languages and logics, incorporates of parsers for CLIF allowing us to seamlessly work with CL ontologies, provides automatic translations of theories from multiple input languages to different serializations, and features a web-based interface, offering a user-friendly and accessible environment for interacting with the tool especially for users who are not familiar with ATP tasks. Additionally, Hets benefits from a dedicated development team [1] that actively works on bug fixes, updates, and enhancements. This continuous support ensures that any reported issues are addressed promptly, and the tool remains up-to-date with the latest advancements in ontological reasoning and specification languages.

To use Hets, we first parse the CLIF file to get the development graph. Then, we navigate to the theory's node in the graph, which leads us to a translation page. Finally, we select the desired translation : the "`CLImp2CFOL`" translation for CASL and the "`CLImp2CFOL:CASL2TPTP_FOF`" translation for TPTP.

An example of the development graph of the *constitutes* relation module is shown in figure 6.3. Upon clicking on the url of the theory [2], we are directed into the translation page, where we translated into CASL (Listing 6.5 shown below) using the "`CLImp2CFOL`" translation.

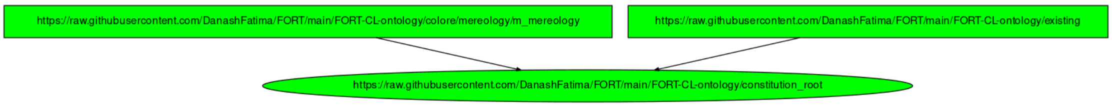


FIGURE 6.3 – *The development graph of the constitutes relation module, saved from the Hets web interface.*

```
sorts individual
op constitutes : individual
op exxisting : individual
op part : individual
op z : individual
pred rel : individual * individual * individual
. (forall xx : individual . not rel(constitutes, xx, xx))
     /\ (forall xx : individual; y : individual
                . rel(constitutes, xx, y) => not rel(constitutes, y, xx))
     /\ (forall xx : individual; y : individual; z : individual
                . rel(constitutes, xx, y)
                    /\ rel(constitutes, y, (var z : individual))
                    => rel(constitutes, xx, (var z : individual)))
     /\ (forall xx : individual; y : individual
                . rel(constitutes, xx, y)
                    => exists t : individual
                                    . rel(exxisting, xx, t) /\ rel(exxisting, y, t
   ))
     /\ forall xx : individual; y : individual
              . rel(constitutes, xx, y) /\ rel(part, z, y)
                  => exists w : individual
                                  . rel(part, w, xx) /\ rel(constitutes, w, z)
%(https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology
   /constitution_root.CLIF)%
forall xx : individual; t : individual
. rel(exxisting, xx, t) => not rel(exxisting, t, xx)
```

1. https://github.com/spechub/Hets/
2. https://github.com/DanashFatima/FORT/tree/main/FORT-CL-ontology for constitution.

```
%(https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology
    /existing.CLIF)%
. (forall xx : individual . rel(part, xx, xx))
     /\ (forall xx : individual; y : individual
             . rel(part, xx, y) /\ rel(part, y, xx) => xx = y)
     /\ forall xx : individual; y : individual; z : individual
           . rel(part, xx, y) /\ rel(part, y, (var z : individual))
               => rel(part, xx, (var z : individual))
%(https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology
    /colore/mereology/m_mereology.CLIF)%
```

Listing 6.5 – The automatic Hets translation of the constitution module from CLIF to CASL.

6.4.3 Running consistency checks

As for running consistency checks, we continued with Hets and selected the "darwin [1]" compiler as the consistency checker for each ontology module, followed by the full FORT theory. To ensure that no problems might be caused by importations, we created an additional file, where we merged all axioms and stripped out the comments. While using Hets for running consistency checks, we were able to contribute to some bug findings, and their corresponding fixes thanks to Fabian Neuhaus and Till Mossakowski : Issue#2101 [2] where "CLFull2CFOL" was not properly implemented, and Issue#2105 [3] where for impredicative CL, Hets uses "CLFull2CFOL" as the default translation, while it should use "CLImp2CFOL" instead, which does not require sequence markers.

The translation works properly with the following steps : (1) parse the theory using CLIF syntax, (2) select check consistency via tools, (3) select "`darwin`" ("`darwin-non-fc`", "`e-darwin`", "`e-prover`", or "`ekrh`") as a prover, and "`CLImp2CFOL:CASL2SoftFOL`" as the translation. It is also possible to do so by using the CASL or TPTP syntax of the theory. However, to check what the model of the theory looks like (i.e. the theory is proved to have a model), then the downloaded version of Hets shall be used, which allows to log it in log files.

The result showing the consistency of each micro-theory, and that of FORT macro-theory, is shown in Annex B, saved from Hets web interface. Note that consistency checking for FOL/CLIF theories will likely fail at some point when the ontology grows either in size or with the logical complexity. This is not a limitation of darwin (or Hets) but rather a consequence of the fact that there is no algorithm that is able to determine consistency for an arbitrary set of FOL formulas.

6.4.4 Automatic proofs using Hets/DOL

For automated theorem proving, Common Logic does not support logical consequences, relative theory interpretations, and other features related to the structuring and comparison of logical theories, as discussed in [Mossakowski2014]. Instead, the authors encourage to use the Distributed Ontology Language (DOL), which is supported by Hets. Thus, we employ DOL for running automatic theorem proofs.

DOL is a language specifically designed for expressing and integrating ontologies and logical theories written in different formalisms and logics [Mossakowski2015]. It provides a unified framework for representing diverse ontologies and their relationships, allowing for seamless interoperability and integration of heterogeneous knowledge resources. It does support the specification

1. http://combination.cs.uiowa.edu/Darwin/
2. https://github.com/spechub/Hets/issues/2101
3. https://github.com/spechub/Hets/issues/2105

of logical consequences, theory interpretations, translations between different logical languages, and various other features that facilitate the structuring and comparison of logical theories, allowing thus for running automatic theorem proofs.

It is widely used in the ontology engineering and research communities to address the challenges of integrating and aligning ontologies across different domains and formalisms. Indeed, DOL has gained significant traction within the foundational ontologies community, as evidenced by its extensive use in the COLORE repository. This reinforces the validity and practicality of adopting DOL as a robust and widely-accepted language for our theorem proving needs.

To form a DOL ontology file, it usually imports files written in specific logics such as Common Logic or OWL 2. Consider the CLIF file below (Listing 6.6) of the specific and generic dependence relations simplified i.e. comments and links stripped out. We construct a DOL ontology, shown below in Listing 6.7, by specifying the "`Logic`" as "`CommonLogic`", importing the CL ontology as the ontology assumption, setting variables as indiscourses, and attaching an (or several) ontology conjecture(s). The conjecture is basically the theorem that is to be proved.

To use the DOL to run automatic theorem proofs, first the file is to be parsed building the development graph as an assumption node i.e. the ontology, and the conjecture(s) node(s) i.e. the theorem(s), as shown in figure 6.4a. Second, we click on "commands" and choose "`auto`", upon which a formerly green node turns red (figure 6.4b) denoting a *local proof goal* . Third we click on "tools" and select "`automatic proofs`", which directs us to the prover page. For the prover selection, not all provers work, however, "`SPASS`" and "`Eprover`" (not "`eprover`") run properly. So, we select one the preceding two provers and specify the conjecture(s) to proof (or the option "`All`"), adjust the timeout property based on the size of the ontology and the complexity of the proof, and click "prove". The results of some proofs, showing theorems proved, are available at the FORT github repository. Returning back to the development graph, the red not turns green i.e. local implications are proved (figure 6.4c).

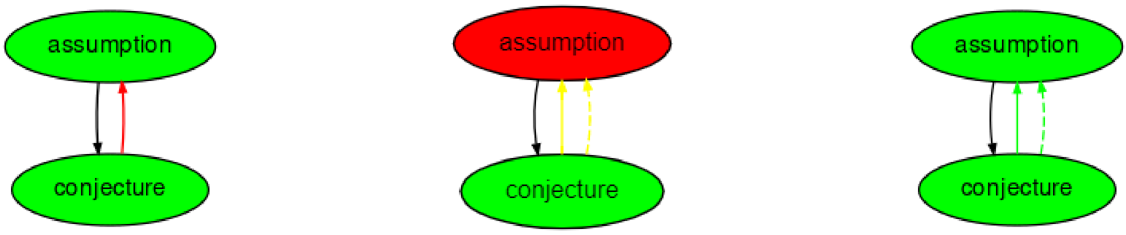
```
(forall (x t)
  (if (existing x t)
    (not (existing t x) )
  )
)
(forall (x y)
  (iff  (SED x y)
      (and  ( forall (t)
            (if (existing x t)
              (existing y t)
            )
          )
          ( not (= x y) )
          ( exists (t)
            (existing x t)
          )
      )
  )
)
(forall (PSI PHI)
  (iff  (GED (PSI x) (PHI y))
      (and  (forall (x t)
            (if (and (PSI x) (existing x t) )
               (exists (y)
                (and (PHI y) (existing y t) )
               )
            )
          )
```

```
          (exists (x t)
            (and (PSI x) (existing x t))
          )
          (not (exists (z) (and (PSI z) (PHI z) ) )
          )
      )
  )
)
```

Listing 6.6 – Example of the simplified dependence relations module.

```
logic CommonLogic
ontology  assumption =
(forall (x t)
  (if (existing x t)
    (not (existing t x) )
  )
)
(forall (x y)
  (iff  (SED x y)
      (and  ( forall (t)
            (if (existing x t)
              (existing y t)
            )
          )
          ( not (= x y) )
          ( exists (t)
            (existing x t)
          )
      )
  )
)
(forall (PSI PHI)
  (iff  (GED (PSI x) (PHI y))
      (and  (forall (x t)
            (if (and (PSI x) (existing x t) )
               (exists (y)
                (and (PHI y) (existing y t) )
               )
            )
          )
          (exists (x t)
            (and (PSI x) (existing x t))
          )
          (not (exists (z) (and (PSI z) (PHI z) ) )
          )
      )
  )
)
(indiscourse a)
(indiscourse now)
end
ontology conjecture = assumption then %implies
(if (existing a now) (not (existing now a )))
end
```

Listing 6.7 – The DOL ontology of the simplified dependence relation module: the theory and a conjecture.



(a) *(b)* *(c)*

FIGURE 6.4 – *The development graph of the simplified DOL ontology example in its three stages : (a) parsed, (b) assumption turned into a local proof goal, and (c) assumption proved.*

6.5 Conclusion

In this Chapter, we have analyzed FORT in the presence of extant literature and associated the theory with a CLIF serialization as the FORT-CL-ontology to support consistency checking [methodology-step-2].

For the analysis of FORT, we elaborated on the three arguments behind constructing FORT based on our observations of the extant literature (Section 6.3.1). These argue the absence of a modular theory of exclusive relations, the need for an inclusive incorporation of the intended minimal set of relations, and the difficulties that arise with employing FOs for the use of foundational relations. In addition, we positioned and aligned FORT to some FOs by elucidating a relation-based comparison between each module in FORT and its corresponding aligned relation in each of the selected FOs, to which FORT presents high similarities (Section 6.3.2). Thus, we have demonstrated the (need and) novelty of FORT with respect to what the literature offers on foundational relations.

As for the CL-ontology (Section 6.4), we associated FORT with a CLIF serialization as an additional formalization for running logical proofs. Then, we used the CLIF serialization to automatically translate FORT into other ontology serializations e.g. TPTP, CASL, and Prover9. The translation task was performed using some online tools, in which we hihlight Hets. After that, we validated the consistency of the FORT-CL ontology (as a single and simplified CLIF file), and constructed some DOL ontologies for running automatic theorem proofs, both using Hets. Thus, we have demonstrated the consistency of FORT by validating the existence of models for FORT.

Therefore, with both the novelty and consistency of FORT shown, we have achieved in this Chapter [micro-objective-2], and overcame, along with Chapter 5, the need for a well-formalized consistent language of an inclusive minimal set of exclusive ontological relations ([challenge A]).

We proceed in the following Chapter 7 towards our third contribution, addressing the third step of the proposed methodology [methodology-step-3] to cover the third micro-objective [micro-objective-3].

Acknowledgments :

We would like to thank Fabian Neuhaus for the discussion over the CLIF syntax and his helpful comments on the use of Hets.

7

On the translation of FORT into SROIQ - a translation procedure

This Chapter refers to our contribution "Translating FOL-theories into SROIQ-Tboxes" published as a short conference paper at The 38th ACM/SIGAPP Symposium on Applied Computing (SAC'2023), March 27-31, 2023.

Sommaire

7.1 Introduction

In the previous Chapter 6, we analyzed our proposed FORT reference ontology, in which we demonstrated the need and novelty of FORT in view of extant literature on foundational relations, in addition to showing its consistency, accomplishing thus the second step of the methodology.

In this chapter, we fulfill our third contribution by offering a decidable lite formalization of our proposed language of foundational ontological relations [micro-objective-3]. This is achieved by extracting a secondary specification from the original one, ensuring desirable computational capabilities through the translation of the FOL formalization into an SROIQ Description Logic formalization i.e. *theoretically formalizing the FORT lightweight ontology* [methodology-step-3].

We first begin in Section 7.2 by providing contextual information on the use of formal languages for theory specification, with a specific focus on First-order logic (FOL) and Description Logics (DL). We present the syntax and semantics of SROIQ DL and emphasize the importance of having multiple specifications that serve different purposes, highlighting the necessity for a translation service between these two languages. Second, we introduce in Section 7.3 the proposed translation procedure as a sequential series of steps, followed by a walk-through example. Third, in Section 7.4, we demonstrate the application of the proposed procedure by translating the original FOL formalization of FORT into a decidable SROIQ DL representation. Finally, Section 7.5 discusses the results and concludes with remarks.

7.2 Context and preliminaries : logical languages

Logical languages provide rigid formalisms for theories, standards, and knowledge domains. Several languages have been designed with varying expressive powers and scalable complexities. For the Semantic Web [Shadbolt2006], which is the next step in the evolution of the World Wide Web, the goal is to have a standard formal representation that is expressive enough to model any knowledge domain, yet decidable enough to be read, understood, and compiled by machines. Thus, the trade-off arises between the expressivity and the decidability properties of logical languages.

The Web Ontology Language [Bechhofer2004] (OWL), a World Wide Web Consortium (W3C) recommendation language for the SW, achieves a balance between these requirements as based on the *SHOIN* [Horrocks1999] logical language. *SHOIN* is a logic from the family called Description Logics [Baader2003] (DLs). DLs form a decidable fragment of the expressive First-Order Logic [Smullyan1995] (FOL), thus, a compromise between expressivity and scalability. DLs are significantly less powerful than FOL but (relatively) easily implemented on the computer [Baader2003]. Each DL formalism models simple descriptions (concepts, roles, and individuals), and differs from other DLs by the constructors used. Constructors offer complex descriptions that derive from simple ones, thus they identify the expressivity of each DL and the complexity of its algorithms and reasoning services. The more the constructors, the more the DL is expressive and complex, until arriving to an expressivity level where decidability is lost. As a semantic web standardization, the latest version of OWL ; OWL2DL [Group2009] is based on the *SROIQ* [Horrocks2006] logic which is the most expressive (yet decidable) commonly used DL.

7.2.1 Preliminaries : the SROIQ DL

The SROIQ DL [Horrocks2006] is a widely used DL in the field of knowledge representation and reasoning. It is known for its expressive power and ability to capture complex relationships and constraints within a domain. SROIQ DL offers several features, including role hierarchies, inverse roles, transitive roles, nominal individuals, and qualified number restrictions. One of the key advantages of SROIQ DL is its decidability, which means that it has effective algorithms for reasoning and consistency checking. This makes it suitable for practical applications where automated reasoning and inference are required. As mentioned earlier, it is the underlying logic used in popular behind the OWL2 DL ontology language, which is widely adopted in semantic web applications and ontology engineering. In the following, we demonstrate the syntax and semantics of the SROIQ DL.

Concepts and roles in *SROIQ*

In DLs, concepts and roles are represented using unary and binary predicates respectively. Let N_C and N_R be mutually disjoint sets of atomic concepts.

Definition 7.2.1 (Role)**.** Every role name R is a role description (atomic role), for $R \in N_R$. A role is either a role name R or the inverse R– of a role name R, where $R \in N_R$. U is a role name representing the Universal – role.

Definition 7.2.2 (Concept)**.** Every concept name C is a concept description (atomic concept), for $C \in N_C$. $\top$ and $\bot$ are concept names representing the Top – concept and the Bottom – concept respectively.
Complex concepts can be built by means of class constructors ; $\neg C$ (concept-complement) ; $C \sqcap D$ (concept-intersection) ; $C \sqcup D$ (concept-union) ; $\{a\}$ (nominals) ; $\exists R.C$ (existential restriction), $\forall R.C$ (universal restriction), $\exists R.SELF$ (local reflexivity), $\geq nR.C$ (at-least restriction), $\leq nR.C$ (at-most restriction), and $= nR.C$ (exact-value restriction) where C and D are concepts in N_C, R is a role in N_R, and n is a natural number in $\mathbb{N}$.

In DL, the terminological knwoledge is expressed by a *TBox*, which contains axioms refering to concepts and roles. The *SROIQ* DL allows for representing complex role inclusion axioms (the "*R*" constructor in *SROIQ*) along with role assertions and composition extensions, nominals (the "*O*" constructor), inverse properties (the "*I*" constructor), and qualified cardinality restrictions (the "*Q*" constructor). In FORT, nominals are not used, so it can be argued that the theory fits in a smaller logic *SRIQ* [Horrocks2005]. However, since the presentation of *SRIO* does not allow for asymmetry role assertions, we employ the *SROIQ* DL. In the following, we present the main syntax of *SROIQ* to which the result of our proposed procedure yields.

Definition 7.2.3 (General concept inclusion axiom)**.** Subsumptions between concepts can be represented by means of general concept inclusions (GCIs) using the operator $\sqsubseteq$. If C and D are concepts in N_C, then $C \sqsubseteq D$ is a GCI in N_C.

Definition 7.2.4 (Role composition axiom)**.** Compositions of binary relations can be represented by means of role compositions using the $\circ$ operator. If S role and τ is a role or role composition in N_R, then $S \circ \tau$ is a role composition in N_R.

Definition 7.2.5 (Role inclusion axiom)**.** Subsumptions between binary relations can be represented by means of role inclusion axioms (*RIAs*) using the operator $\sqsubseteq$. If S role and τ is a role or role composition in N_R, then $S \sqsubseteq \tau$ is a *RIA* in N_R.

Definition 7.2.6 (Role assertion axioms)**.** Constraints on roles can be represented by means of role assertions. If S and T are roles in N_R, then symmetry $Sym(R)$, asymmetry $Asy(R)$, transitivity $Tra(R)$, reflexivity $Ref(R)$, irreflixivity $Irr(R)$ and pairwise disjointness $Dis(R,S)$ are role assertions in N_R.

Definition 7.2.7 (TBox $\mathscr{T}$)**.** A TBox is a union of an RBox and a CBox. A role box (RBox) is a finite set of role assertions or *RIAs*. A concept box (CBox) is a finite set of GCIs.

Structural restrictions in *SROIQ*

To maintain the decidability of the TBox, which can easily be violated by the RBox, *SROIQ* imposes additional syntactic restrictions on the use of roles ; the *simplicity* and *regularity* constraints.

Definition 7.2.8 (Simplicity)**.** A role R is non-simple if (i) R subsumes a role composition τ i.e. a *RIA* of the form $\tau \sqsubseteq R$, or (ii) R appears in the role assertions $Tra(R)$ or $Tra(R^-)$, or (iii) R appears in *RIAs* of the form $S \sqsubseteq R$ and $R \sqsubseteq S$ where S is a non-simple role name. The simplicity constraint bans the use of non-simple roles in concepts definitions as $\exists R.SELF$ and $< | = | > nR.C$, and in role assertion of the form $Irr(R)$, $Asy(R)$, and $Dis(R,S)$.

Definition 7.2.9 (Regularity)**.** A *RIA* of the form $\tau \sqsubseteq R$ is $\prec$-regular if (i) $\tau = R \circ R$, or (ii) $\tau = R^-$, or (iii) $\tau = R \circ S_1 \circ .. \circ S_n$, or (iv) $\tau = S_1 \circ .. \circ S_n \circ R$, or (v) $\tau = S_1 \circ ... \circ S_n$ where $S_i \prec R \forall 1 \leqslant i \leqslant n$ and $S_i \prec R \Leftrightarrow S_i^- \prec R$. The regularity constraint forces a *set of RIAs* of the RBox to be regular i.e. to have a regular order $\prec$ such that each *RIA* in the set is $\prec$-regular.

7.2.2 Transitioning from FOL to the SROIQ DL

For the formalization of large comprehensive theories such as foundational ontologies (e.g. BFO [Smith2002], DOLCE [Masolo2003], UFO [Guizzardi2015], etc.) and relation theories (e.g. mereology [Varzi2003], location [Gilmore2018], etc.), FOL is employed. Since in these theories the purpose is to represent a knowledge domain in a way that rules out unintended models. Such a precision in capturing the domain of interest requires a wide expressive power. Moreover, the employment of such theories in practice is highly desirable where applications seek to guarantee decidability and tractability in their reasoning services. Thus, to support its application in the semantic web, a lightweight fragment of the theory is serialized within a decidable language, such as *SROIQ*. In ontology engineering, as discussed in Chapter 4 and proposed in Section 4.3.2, it is becoming more popular to offer a two-folded formalization of a knowledge domain ; a FOL-formalization as an initial *reference ontology*, and a DL-formalization as a secondary *lightweight ontology* [Guizzardi2007].

In our thesis, to overcome [challenge B], we aim to provide a *SROIQ* formalization of FORT (*FORT lightweight ontology*) to offer its practice in the semantic web as an OWL ontology that supports reasoning about relations for modeling the composition of tangible entities. The passage from the former FOL-formalization to a latter SROIQ-formalization is a crucial task that allows for the transition from expressive theories to building semantic web applications for the world wide web. To fulfill this task, a translation is required to guide a systematic and principled rewriting of the set of formulas (definitions, axioms and theorems) in the theory.

The investigation on extant literature concerning the transition from FOL to *SROIQ* reveals the following. In [Horrocks2004a], the approach of combining function-free Horn rules with expressive DLs , shows that determining the consistency of knowledge bases is undecidable. In [Levy1996], a language formalism is developed as a combination of Datalog with the DL *ALCNR*. The work is an extension of the work done in [Donini1998] by allowing more DL atoms in the rules, yielding in an undecidable formalism. A different approach is [Grosof2003], which works on the intersection between Horn clauses and the *SHOIN* DL rather than on the union of the two formalisms. The approach imposes restrictions on the use of DL, yielding in a decidable formalism whose rules cannot be altered. Other Datalog-based approaches such as [Rosati2005], defines framework for the integration of DL ontologies and $Datalog^{\neg\vee}$ following [Eiter1997].

Among other interesting approaches, the work presented in [Rosati2005] proposes a rewriting technique that preserves the semantics of the rules to be added to the DL. Although the technique imposes some syntactic restrictions to maintain regularity (as *admissible rules*) over a role hierarchy R_h, however, simplicity remains problematic where the technique does not guarantee that non-simple roles are not used in some concept inclusion axioms and role assertion axioms, as the simplicity constraint restricts [Horrocks2006].

Thus, there is a limited availability of systematic guidelines, to date, for translating FOL theories into SROIQ knowledge bases. In this chapter, *driven by the need to facilitate the practical application of FORT*, particularly in the semantic web, we aim to address this issue. To do so, we develop a translation procedure that takes an FOL theory as input and generates a corresponding *SROIQ* knowledge base as output. We illustrate the application of this procedure by translating FORT, as demonstrated in the subsequent sections.

7.3 The proposed translation procedure

In this section, we propose a procedure for translating FOL theories into decidable SROIQ TBoxes (Section 7.3.1), followed by showing an example of its use (Section 7.3.2).

7.3.1 Steps of the procedure

Figure 7.1 depicts the proposed procedure of steps which takes as input a set of axioms corresponding to a theory T formalized in FOL. Knowing that arbitrary FOL theories can contain predicates of any arity [1], we narrow down the inclusion to only unary and binary predciates by assuming first and second arity FOL theories only. Each step comprises ; an input set S_i resembling the theory's serialization in a specific logical formalism (e.g. S_1 is the set of axioms in Clausal Form) ; the operation(s) to be performed ; and an output set S_{i+1} which forms the input of the next step. The final output of the translation corresponds to a *SROIQ* serialization as a lite decidable fragment of the initial inputted FOL theory.

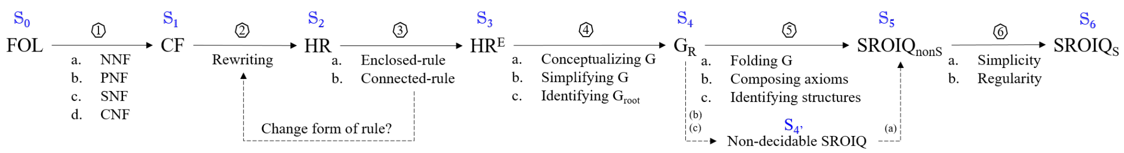


FIGURE 7.1 – *The six steps of the translation procedure and their corresponding operations, starting from a set (S_0) in first-order logic, and resulting in structured set (S_6) of SROIQ axioms.*

In the following we present, for each step, the operation(s) to be performed on each axiom A_j in the input S_i, where $0 \leqslant i \leqslant 5$ and $1 \leqslant j \leqslant n$ for a set of n axioms. It is possible to initiate the procedure at any step e.g. 1, 2, 3, 4, or 5, depending on the serialization of the initial theory inputted e.g. FOL, CF, HR, HR^E, or G_R respectively.

1. An m-arity theory is one that allows for predicates holding m-variables. For example, a three-arity theory allows for including predicates such as R(x,y,z)

A. Transforming to Clausal Form (CF)

- Input : S_0 set of n formulas in first-order logic. For each axiom, we apply the four operations :

A.1 Negation Normal Form (NNF) The first operation is to simplify the formula and move negations inwards so that the formula contains connectives of the form $\forall$, $\exists$, $\wedge$, $\vee$, and $\neg$ only. This is done using logical equivalences such as the implication identity ($A \rightarrow B \equiv \neg A \vee B$), De Morgan's Laws ($\neg(A \vee B) \equiv \neg A \wedge \neg B$, $\neg(A \wedge B) \equiv \neg A \vee \neg B$), those for negating quantifiers ($\neg(\exists x A(x)) \equiv \forall x \neg A(x)$, $\neg(\forall x A(x)) \equiv \exists x \neg A(x)$). This operation results in set S_{NNF} equivalent to S_0 ($S_{NNF} \equiv S_0$).

A.2 Prenex Normal Form (PNF) The second operation is to demarcate the variables of the formula and move all the quantifiers to the left. Unification takes place by renaming different variables that have same notation i.e. each quantifier should use a unique variable name. For example, the formula $\forall x(A(x) \rightarrow B(x)) \wedge \exists x(\neg C(x) \vee B(x))$ becomes $\forall x \exists y(A(x) \rightarrow B(x)) \wedge (\neg C(y) \vee B(y))$. This operation results in set S_{PNF} equivalent to S_{NNF} ($S_{PNF} \equiv S_{NNF}$).

A.3 Skolem Normal Form (SNF) The third operation is to remove all existential quantifiers. This is by replacing an existentially quantified variable e.g. x by a *Skolem constant* if x is not preceded by universally quantified variables, and by a *Skolem function* in terms of the universally quantified variables that precede x otherwise. For example, the formula $\exists x A(x)$ becomes $A(c)$ upon replacing x by the constant c as $\{x \leftarrow c\}$, while the formula $\forall y B(y) \exists x R(x,y)$ becomes $\forall y B(y) R(f(y),y)$ upon replacing x by the function f as $\{x \leftarrow f(y)\}$.
The resulting set of axioms S_{SNF} is not equivalent to S_{PNF} ($S_{PNF} \neq S_{NNF}$). However, it is worth noting that the skolem form is not precisely logically equivalent to the original statement S_{PNF}. Indeed, for the skolem (S_{SNF}) to be interpretative in terms of the original prenex form (S_{PNF}), the interpretation shall be extended by giving definitions of the added skolem constants/functions. This is however possible for any interpretation in which S_{PNF} was true. Fortunately, as the skolem form is more restrictive than the original formulation, S_{PNF} is subsumed by its skolem form S_{SNF} ($S_{SNF} \leq S_{PNF}$)

A.4 Clausal Normal Form (CNF) The fourth operation is to distribute disjunctions over conjunctions so that the formula becomes a set of conjunctions of disjunctions. For instance, the formula $A(x) \vee (B(x) \wedge C(x))$ is rewritten as $(A(x) \vee B(x)) \wedge (A(x) \vee C(x))$. We say the formula is a set of *clauses* linked by conjunctions. The *clausal form* basically removes these conjunctions and represents the original formula as the set of resulting clauses where each clause is a disjunction(s) of negated and non-negated atoms e.g. $(A(x) \vee B(x)) \wedge (A(x) \vee C(x))$ rewrites as $\{(A(x) \vee B(x)), (A(x) \vee C(x))\}$. The resulting set of clauses S_{CNF} of all axioms is equivalent to the inputted set S_{SNF} ($S_{CNF} \equiv S_{SNF}$).

All four operations formulate a sequence of transformations performed on an axiom a. This sequence of operations guarantees that the final form of a, as a set of clauses, is the simplest and best form for the rewriting of a as a group of rules as required by step 7.3.1. Although, a skolem form is not equivalent to its prenex original form, however, there are satisfiability preserving transformations from first order logic (the original formula S_0) to the clausal norm form (S_{CNF}). In other words, if a first order formula is satisfiable, then its clausal normal form is satisfiable. Conversely, if the clausal normal form of a of formula is unsatisfiable, then the formula is unsatisfiable. Since the resulting set S_{CNF} (S_1) is a restrictive set of original S_0, then S_1 is subsumed by S_0, denoted $S_1 \leq S_0$.

B. Imposing Horn form Rules (*HR*)

• Input : S_1 (= S_C) set of m clauses in clausal form. For each clause C_i in S_1, let n and m be the numbers of non-negated and negated atoms respectively. Let A, B, and C be atoms of C_i resembling unary/binary predicates.

☐ if $n \leqslant 1$, then C_i is a horn clause.

- if $n = 1$, then C_i which is a positive horn clause of the form $\neg A \vee \neg B \vee C$ is rewritten into $[\neg(A \wedge B) \vee C] \equiv [A \wedge B \rightarrow C]$. This rewriting is not the only option but the favorable rewriting in comparison to the other m possible rewriting options e.g. $[\neg(A \wedge \neg C) \vee \neg B] \equiv [A \wedge \neg C \rightarrow \neg B]$. So, in total $m+n$ options with a preference of the n^{th} option.
- if $n = 0$, then C_i which is a negative horn clause of the form $\neg A \vee \neg B \vee \neg C$ can be re-written into one of the following m favorable options ; $[\neg(A \wedge B) \vee \neg C] \equiv [A \wedge B \rightarrow \neg C]$, $[\neg(A \wedge C) \vee \neg B] \equiv [A \wedge C \rightarrow \neg B]$, or $[\neg(B \wedge C) \vee \neg A] \equiv [B \wedge C \rightarrow \neg A]$, in comparison to $[\neg(A \wedge B \wedge C)] \equiv [A \wedge B \wedge C \rightarrow \varnothing]$. So, in total $m+n$ options with a preference of one of the m options.

☐ if $n > 1$, then C_i is not a horn clause, and has the form $\neg A \vee B \vee C$. However, it is still possible to impose the rewriting within a horn implication form by qualifying any of the n atoms for the rule's head e.g. $[\neg(A \wedge \neg B) \vee C] \equiv [A \wedge \neg B \rightarrow C]$ as possible favorable options. It is also possible to qualify one of the m atoms e.g. $[\neg(\neg B \wedge \neg C) \vee \neg A] \equiv [\neg B \wedge \neg C \rightarrow \neg A]$ as a less favorable option of having negation in the rule's head. Whereas the least favorable option is to qualify all the n atoms to the rule's head e.g. $A \rightarrow B \vee C$.

This step is critical in establishing an implication form of C_i. Such a form is preliminary for the translation process in which the intended resulting rules shall have the form of *inclusion axioms*. The preference of one option over another is to avoid having non-negated atoms (e.g. case $n = 1$), an empty-operator (e.g. case $n = 0$), or disjunction operators (e.g. case $n > 1$) in the rule's head. The resulting set of horn rules S_{HR} (S_2) is equivalent to the inputted set (S_1) i.e. ($S_2 \equiv S_1$).

C. Qualifying Expressible rules (HR^E)

• Input : S_2 (= $S_H R$) set of m horn rules in their implication form. For each horn rule R_i of the form $a_1 \wedge a_2 \wedge .. \wedge a_n \rightarrow h$ in S_2, check whether it satisfies the two syntactic restrictions :

C.1 Enclosed-rule constraint This constraint restricts the variables present in the head of the rule h to be present in at least one of the body's atoms a_i for $1 \leqslant i \leqslant n$ i.e. enclosed-variables. For example, a rule of the form $R(x,y) \wedge S(y,a) \rightarrow R(x,b)$ is in-expressible since b does not appear in the rule's body, and thus cannot be qualified.

C.2 Connected-rule constraint This constraint assures that for each pair of variables x and y in r_i, there exists a sequence $z_1, z_2, ..., z_n$ such that $z_1 = x$, $z_n = y$, and for $1 \leqslant j \leqslant n$ there is a binary predicate R in r_i such that $R(z_i, z_{i+1})$ or $R^-(z_{i+1}, z_i)$, i.e. connected-variables. For instance, a rule of the form $R(x,y) \wedge S(a,b) \rightarrow R(x,b)$ is in-expressible since the variables a/b are not connected to x/y.

Since a variable that is not enclosed or not connected results in an inexpressible rule, then having a such mediating step in procedure, which qualifies rules according to the requirements of upcoming steps, is essential. In cases where the R_i could not be qualified, this means that the form of R_i does not comply syntactic constraints. However, it is possible to reexamine the form of R_i qualified from the previous step, and re-consider another form of the rule that satisfies the two constraints, as depicted in figure 7.1. For example a rule of the form $R(x,y) \wedge S(y,a) \rightarrow R(x,b)$ can be rewritten in the form $\neg R(x,b) \wedge S(y,a) \rightarrow R(x,y)$ according to step 7.3.1. The resulting set of expressible horn rules S_{HR^E} (S_3) is a restrictive set of the inputted set (S_2) i.e. the S_3 is subsumed by S_2, denoted $S_3 \leq S_2$.

D. Constructing the rule graph (G_R)

• Input : S_3 (= S_{HR^E}) set of *expressible* horn rules in their implication form. For each horn rule R_i apply the following steps to construct the graph G_{R_i} :

D.1 Conceptualizing G The first operation is to conceptualize the rule R_i as a directed labeled graph G_{R_i}. For the remaining part of the paper, we refer to the unary and binary predicates of R_i as concepts and roles respectively. A rule graph G_{R_i}, following [Gasse2008], is defined as $G = \langle V,E,L,H \rangle$ where ;

- V is a finite set of the variables of R_i resembling the vertices of the graph ;
- E is a finite set of the roles in the body of R_i resembling the edges of the graph, such that a role $S(x,y)$ is added to E in the form of an edge S_{xy} ;
- L is a finite set of label sets corresponding to each variable in R_i of the form $L = \{L_{x1} = \{C_1,C_2\}, L_{x2} = \{C_2,C_3\}, ..., L_{xn} = \{C_1,C_3\}\}$ where $L_{x1}, L_{x2}, ...,$ and L_{xn} are the label sets of the variables $x1, x2, ...,$ and xn respectively, C_1,C_2,C_3 are concepts in R_i which the variables satisfy, and L resembles the labels of all the vertices of the graph ;
- H is the head of R_i written in the form of an assertion ; either a concept assertion of the form $x : C$ if the rule is concept-headed i.e. $C(x)$, or a role assertion of the form $x,y : R$ if the rule is role-headed i.e. $R(x,y)$.

For example, a rule of the form $R(x,y) \wedge A(x) \wedge B(y) \wedge S(x,z) \wedge S(y,z) \rightarrow T(x,y)$ is conceptualized into $G = \langle V,E,L,H \rangle$ defined in terms of ; $V = \{x,y,z\}$, $E = \{R_{xy}, S_{xz}, S_{yz}, T_{xy}\}$, $L = \{L_x = \{A\}, L_y = \{B\}, L_z = \varnothing\}$, and $H = x,y : T$.

D.2 Simplifying G The second operation is to simplify the roles and concepts in E and L of the rule graph. Simplification is made by removing edges and labels that can be removed without altering the satisfiability of the rule. Such edges/labels correspond to roles/concepts that subsume other roles/concepts in E/L i.e. are entailed (implied) by other edges/labels within E/L. Thus, these edges/labels do not further constraint G, rather than being only implied within G (redundant). To simplify $G = \langle V,E,L,H \rangle$, first we proceed with simplifying the roles, followed by the concepts according to the rules below. Let $V = \{x,y,z\}$.

D.2.1 Simplifying roles : For each role in E, check whether the following rules apply, in order, until no rule is applicable anymore. Let $\mathscr{R}$ be the role hierarchy consisting of *role inclusion axioms* (*RIAs*) of the roles in E.

1. *if* (a) $\mathscr{R}$ contains a *RIA* of the form $R^1 \circ R^2 \circ ... \circ R^n \sqsubseteq S$, and (b) E contains the roles R^1_{x1x2}, R^2_{x2x3}, ..., R^n_{xnxn+1}, and S_{x1xn+1}, or their inverses, *then* remove S_{x1xn+1} from E.
2. *if* (a) $\mathscr{R}$ contains a *RIA* of the form $R \circ R \sqsubseteq R$ i.e. $Tra(R)$, and (b) E contains the roles R_{xy}, R_{yz}, and R_{xz}, or their inverses, *then* remove R_{xz} from E.
3. *if* (a) $\mathscr{R}$ contains a *RIA* of the form $S \sqsubseteq R^-$, and (b) E contains the roles R_{xy} and S_{yx}, *then* remove S_{yx} or R_{xy} from E.

4. *if* (a) $\mathscr{R}$ contains a *RIA* of the form $S \sqsubseteq R$, and (b) E contains the roles R_{xy}, and S_{xy}, and (c) $R \neq S$, *then* remove R_{xy} from E.

D.2.2 Simplifying concepts : For each variable labels' set $L_i = C_j for 1 \leqslant j \leqslant n$ in L, check whether the following rules apply, until no rule is applicable anymore. Let $\mathscr{T}$ be the concepts hierarchy consisting of *general concept inclusion axioms* (GCIs) of the concepts in L.

1. *if* $\bigcap_{C_j \in L_i} C_j \equiv \top$ in $\mathscr{T}$, *then* empty L_i i.e. $L_i = \varnothing$.
2. *if* $C_j \sqsubseteq C_{j'}$ in $\mathscr{T}$, *then* remove $C_{j'}$ i.e. $L_i = L_i - \{C_{j'}\}$.

D.3 Identifying the root G_{root} The third operation is to identify the root of G which is a (set of) variable(s) depending on the form of the rule's head H. If the rule is concept-headed i.e. $H = x : C$, then the root is a single variable expressed $G_{root} = \{x\}$. If the rule is role-headed i.e. $H = x, y : R$, then the root is a path of variables starting by x in G to y in G and encompassing all the vertices in between x and y, expressed $G_{root} = \{x, .., y\}$.

The intuition behind constructing a conceptualization of a rule as a graph of vertices and edges, is to visualize the links between the variables of the rule, figure-out the overall shape of the rule (e.g. a tree or a cycle), and make the necessary operations (simplification and identification of root) to convert it into the *SROIQ* serialization. The importance of roles' simplification relies in maximizing the possibility that G will satisfy the later semantic restrictions required by the next step. Whereas that of concepts' simplification relies in minimizing the label's set of each vertex of G by removing subsumers and maintaining (a list) of the *most specific concepts* that imply the subsumers concepts. And last, identifying the root of G is the key to the next step's conversion. The resulting set of rule graphs S_G (S_4) is equivalent to the inputted set S_3 i.e. $S_4 \equiv S_3$.

E. Converting into *SROIQ* axioms

• Input : S_4 (= S_G) set of rules graphs in their simplified form. For each rule graph G perform the subsequent operations to convert it to a *SROIQ* inclusion axiom A, and prepare next step's inputs I and $\mathscr{R}_{NS}$ resembling the proposition builders and the set of non-simple roles in S_4, respectively.

E.1 Folding G The first operation is to qualify tree rule graphs only, i.e. G must not contain cycles when considered as an undirected graph, and to shrink G to have V equivalent to its root G_{root}. It is referred to as folding since the vertices of G that are not in the G_{root} set, are folded back into a neighbor vertex [Gasse2008]. The folding uses the *rolling-up technique* [Tessaris2001] which allows tree-like structures to be expressed as concept expressions. Thus, for each vertex z of G that is a leaf node and does not appear in G_{root}, we fold z into a neighbor vertex y by using the edge between z and y i.e. R_{zy}. This is done by rolling-up z into y as follows;

1. eliminating R_{zy} from E; by rolling the edge into a concept expression and adding it to y's set of labels L_y :

   — if $L_z \neq \varnothing$, then $L_y = L_y \cup \exists R. \bigcap_{C_i \in L_z} C_i$

   — if $L_z = \varnothing$, then $L_y = L_y \cup \exists R.\top$

2. eliminating z from V; $V = V - \{z\}$

E.2 Composing axioms The second operation is to compose the axiom(s) A from G as *general concept inclusion axioms* or *role inclusion axioms* according to the rule graph G.

- ☐ if G is concept-headed i.e. it has the form $G = \langle\{x\}, \varnothing, L_x, x : C_{root}\rangle$, then G is converted into a single axiom A of the form $\bigcap_{C_i \in L_x} C_i \sqsubseteq C_{root}$ i.e. a *general concept inclusion axioms*.
- ☐ if G is role-headed i.e. it has the form $G = \langle\{x_1, ..., x_n\}, \{R_{x1,x2}, R_{x2,x3}, ..., R_{xn-1,xn}\}, \{L_{x1}, L_{x2}, ..., L_{xn}\}, x_1, x_n : R_{root}\rangle$, then G must not contain concepts but only roles to be converted to role insertion axiom(s). Thus we apply the following ;
  1. for every vertex x_i in V, rewrite its label $L_x i$ as a role expression using a fresh role, and fresh concept. Let
     - C' be a fresh concept that is the intersection of all x's labels as $C' = \bigcap_{C_i \in L_x} C_i$, and
     - RC' be an auxiliary property associated to C' as $C' \equiv \exists RC'.SELF$.

     Thus, each instance of C' will have the role RC' with itself, and the existence of such a loop implies that the individual upon which RC' loops over is an instance of C'. This results in emptying V, and G becoming a set of edges only in which there is a single path between x_1 and x_n.
  2. convert the list of roles in E into a *role inclusion axiom* $\Omega \sqsubseteq R_{root}$, where Ω is the concatenation of the roles in E in the form of $\Omega = S_{x1,x2} \circ ... \circ S_{xn-1,xn}$.

E.3 Identifying structures The third operation is to verify the syntax of *role inclusion axioms* to be compliant with one of the forms restricted by the simplicity and regularity constraint of *SROIQ*, without achieving decidability (yet), and to form the inputs of the next final step. Considering *RIAs* having the form $(a_j) : \Omega \sqsubseteq R_{root}$, we apply the following ;

- — restrict Ω to satisfy one of the following forms ; (i) $R \circ R$; or (ii) R^- ; or (iii) $S_1 \circ S_2 \circ .. \circ S_n$; or (iv) $R \circ S_1 \circ .. \circ S_n$; or (v) $S_1 \circ .. \circ S_n \circ R$; or (vi) S and S is simple.
- — if Ω is a role composition, then ;
  1. for every S_i in Ω, that is not an inverse role and different from R_{root} ; we define a "proposition builder" $\mathbb{I}_j$ to represent that the axiom a_j having index j is included in the TBox, and a_j holds some ordering relations between S_i and R_{root} as follows ; $\mathbb{I}_j \rightarrow (S_i \prec R_{root}) \wedge (S_i^- \prec R_{root}) \wedge ..$ for all $1 \leqslant i \leqslant n$, following [Botti Benevides2019]. Each proposition builder $\mathbb{I}_j$ is added to the set of proposition builders I.
  2. add R_{root} to the list of non-simple roles $\mathscr{R}_{NS}$ as follows ; $\mathscr{R}_{NS} = \mathscr{R}_{NS} \cup \{R_{root}\}$.

In this step, the folding of G ensures that all leaf nodes of G are rolled up so that $V = G_{root}$ and G consists of sole edges which link the variables of G_{root}. Then, composing A is straightforwardly done using rolling-up techniques of concepts into roles. After that, $\mathscr{R}_{NS}$ and I are built, and the syntax of each axiom A_i is guaranteed to fall within the regularity constraint of *SROIQ*. The resulting set of axioms $S_{SROIQ(nonS)}$ (S_5) is a restrictive sub-set of the inputted set S_4 i.e. S_5 is subsumed by S_4, denoted $S_5 \leq S_4$.

F. Establishing decidability - Generalization

• Inputs : S_5 ($= S_{SROIQ(nonS)}$) set of *SROIQ* axioms in their non-structured form, I set of proposition builders, and $\mathscr{R}_{NS}$ set of non-simple roles.

In this final step, the goal is to structure the TBox by extracting a decidable fragment of the inputted set i.e. from a non-structured TBox to a structured one. This is done by applying two rules imposed by the two syntactic constraints of *SROIQ* ; simplicity and regularity [Krötzsch2012b].

These constraints target the theory as a whole rather than each single axiom per separately, to guarantee that the reasoning algorithms are correct and do terminate [Krötzsch2012b], and that the satisfiability problem is decidable [Horrocks2006]. Thus, in contrast to the preceding steps in which we have treated the inputted sets S_i as per element, in the following, we deal with all the axioms of S_5 at once, for each rule.

F.1 Simplicity rule This operation applies the simplicity rule by tracking the occurrences of non-simple roles in S_5 and dropping axioms that violate decidability. For each role R_i in the set of non-simple role $\mathscr{R}_{NS}$, for each A_i in S_5, if A_i is of the form ; (i) $\exists R_i.SELF$; or (ii) $< | = | > nR_i.C$; or (iii) $Irr(R)$; or (vi) $Asy(R)$; or (v) $Dis(R,S)$, then drop A_i, and $S_{5'} = S_5 - \{A_i\}$. It is possible to argue another option of dealing with the simplicity rule, by dropping the *role inclusion axiom* $\Omega \sqsubseteq R_i$ which makes R_i non-simple in the first place, and keep *role assertion axioms*.

F.2 Regularity rule This operation applies the regularity rule by computing the incompatibilities of different regular orders in I and tracking their corresponding axioms to be dropped. We follow the approach proposed in [Botti Benevides2019] for tracking incompatibilities caused by irregularities of contracting partial orders over a role hierarchy. The approach builds a meta-theory in propositional logic and defines the problem of finding subsets of S_5 (i.e. the TBox) that satisfy regularity, as a decidable SAT problem.

I is the set of proposition builders of the form $\mathbb{I}_j \rightarrow (S_i \prec R) \wedge (S_i^- \prec R)$ capturing the statement ; the inclusion of axiom a_j in the TBox requires both orders $S_i \prec R$ and $S_i^- \prec R$ to hold in the role hierarchy $\mathscr{R}$. Using I, we track incompatibilities between each pair of proposition builders if there exists two different contradicting regular orders on roles, which lead to an irregular role hierarchy. An incompatibility between two proposition builders $\mathbb{I}_m$ and $\mathbb{I}_n$ corresponds to an incompatibility between the axioms a_m and a_n. We register this by build a meta axiom $\mathbf{m} \in \mathscr{M}$, of the form $\mathbf{m} : \mathbb{I}_m \rightarrow \neg \mathbb{I}_n$ stating that the inclusion of the axioms of one proposition builder infers the requirement to drop the axioms of the other i.e. either one of a_m or a_n shall be dropped. The axioms a_m and a_n are added to the set of unstructured axioms $\mathscr{U}$.

For example, the two proposition builders $\mathbb{I}_5 \rightarrow (S_1 \prec S_2) \wedge (S_1^- \prec S_2)$ and $\mathbb{I}_7 \rightarrow (S_2 \prec S_1) \wedge (S_2^- \prec S_1)$ signify that axioms a_5 and a_7 are incompatible. This is represented in the meta-theory as a meta axiom $\mathbf{m} : \mathbb{I}_5 \rightarrow \neg \mathbb{I}_7$ indicating that for the TBox to be decidable, one of the two axioms shall be dropped i.e. TBox $- \{a_5\}$ or TBox $- \{a_7\}$, and $\mathscr{U} = \mathscr{U} \cup \{a_5, a_7\}$.

After computing all the incompatibilities due the regularity rule, we have the following :

- $S_{5'} = \{a_i\}$; the modified unstructured set of axioms inputted (after applying the simplicity rule to the initial S_5 set).
- $\mathscr{M} = \{m_k\}$; the set of meta axioms from the meta level propositional theory.
- $\mathscr{U} = \{a_i\}$; the set of unstructured axioms from the incompatible proposition builders specified in the the meta axioms set.

As a consequence of what preceded, removing now the unstructured axioms in $\mathscr{U}$ from $S_{5'}$, yields in $S_{5'} - \mathscr{U}$ containing a set of structured (safe) axioms that comply with the syntactic restrictions, that is a subset of S_5 i.e. $S_{5'} \leq S_5$. The goal is however to find the maximal structured set of axioms by finding the structured subsets of $\mathscr{U}$ e.g. if a_1 and a_2 are incompatible, and a_3 and a_4 are incompatible, this does not mean that (a_1,a_3) or (a_1,a_4) or (a_2,a_3) or (a_2,a_4) are incompatible too. To find out axioms that are related, the meta axioms in $\mathscr{U}$ can be linked depending on the axioms they tackle. For instance, if **m1** which captures $\mathbb{I}_5$ and $\mathbb{I}_7$, is the only meta axiom that captures these two proposition builders while no other meta axiom does, then **m1** is to be extracted from $\mathscr{M}$ into a sub-theory $\mathbb{M}_1$, and axioms a_5 and a_7 are to be removed from $\mathscr{U}$ into a tuple $\mathbb{U}_1 = \langle\{a_5\},\{a_7\}\rangle$ indicating an ordered subset from $\mathscr{U}$ where one set of axioms is to be chosen to be included in the structured TBox (e.g. in our case either $\{a_5\}$ or $\{a_7\}$. In the case when another meta axiom, e.g. **m6**, captures either $\mathbb{I}_5$ or $\mathbb{I}_7$ or both, then both **m1** and **m6** are extracted from $\mathscr{M}$ int $\mathbb{M}_1$, and all axioms participating in the proposition builders of both meta axioms are extracted from $\mathscr{U}$ to $\mathbb{U}_1$.

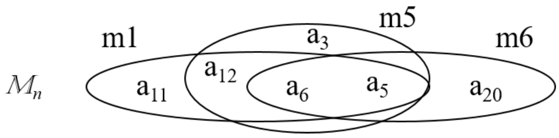


FIGURE 7.2 – *An example of a meta level propositional sub theory $\mathbb{M}_n$ consisting of three meta axioms $m-1$, m_5, and m_6. m_1 captures the axioms a_{11}, a_{12}, a_6 and a_5, having a_6 and a_5 common with m_6, and a_{12}, a_6 and a_5 common with m_5.*

To generalize what preceded, we apply the following :

1. divide the meta axioms in $\mathscr{M}$ into subsets of dependent meta axioms $\mathbb{M}_n$. Each set $\mathbb{M}_n$ contains the meta axioms m_k's that capture overlapping axioms a_i i.e. an $\mathbb{M}_n$ consists of m_k's overlapping over the roles that are tackled in their axioms. Figure 7.2 shows an example of a sub theory $\mathbb{M}_n$ of 3 dependent meta axioms (m_1, m_5, and m_6) overlapping over 6 axioms (a_{11}, a_{12}, a_6, a_3, a_5, and a_{20}).
2. for each set $\mathbb{M}_n$, specify subsets of axioms within a tuple $\mathbb{U}_n$ upon which one should be excluded.

The result is a number of tuples $\mathbb{U}$ equivalent to that of the meta theory subsets $\mathbb{M}$, where each tuple presents choices of sets of axioms that, together, violate decidability, and one set exactly shall be chosen to be excluded from the TBox.
Thus, $S_6 = S_{5'} - \bigcup_n Choice[\mathbb{U}_n]$, and S_6 is subsumed by $S_{5'}$ ($S_6 \leq S_{5'}$).

In Sections 7.4, we apply the procedure on FORT for its translation into SROIQ, in which we show how to compute each of the constructs of the meta level theory precisely. [1]

7.3.2 A walk-through example

Consider the formula below. It is a horn rule satisfying the syntactic restrictions for expressible rules. Thus, we start the translation procedure from step 7.3.1.

$$\forall(x,y,z,w)S(x,y) \wedge A(x) \wedge B(y) \wedge R(y,w) \wedge D(w) \wedge F(z) \wedge T(y,z) \rightarrow W(x,z) \tag{R}$$

$$G = \langle\{x,y,z,w\},\{S_{xy},R_{yw},T_{yz}\},\{L_x = \{A\},L_y = \{B\},L_z = \{F\},L_w = \{D\}\},x,z : W\rangle \tag{G}$$

1. For more information on the use of meta-level theories in propositional logic and the definition of satisfiability as a SAT problem, please refer to the paper [Botti Benevides2019], from which we adopt the sixth step of our procedure.

According to figure 7.3a, w must be folded i.e. roll-up w into y, where $L_w = \{D\}$. Thus $L_y = B \cup \exists R^-.D$. Hence, G becomes :

$$G = \langle \{x,y,z\}, \{S_{xy}, T_{yz}\}, \{L_x = \{A\}, L_y = \{B \cup \exists R^-.D\}, L_z = \{F\}\}, x, z : W \rangle \quad \text{(G)}$$

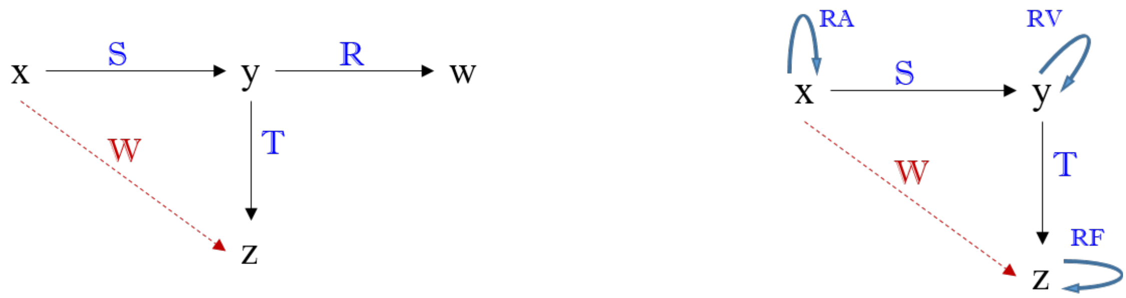
Since the graph is role-headed, then we shall empty L by rolling-up concepts into roles as following

$A \equiv R_A.SELF, where\ R_A\ is\ an\ auxiliary\ role.$
$F \equiv R_F.SELF, where\ R_F\ is\ an\ auxiliary\ role.$
$Let V \equiv B \cup \exists R^-.D \,\&\, V \equiv \exists R_V.SELF, where\ R_V\ is\ an\ auxiliary\ role.$

$$G = \langle \{x,y,z\}, \{S_{xy}, T_{yz}, RA_{xx}, RF_{zz}, RV_{yy}\}, \varnothing, x, z : W \rangle \quad \text{(G)}$$

Now that the rule graph is fully folded 7.3b, it is possible to convert it into an axiom A.

$$RA \circ S \circ RV \circ T \circ RF \sqsubseteq W \quad \text{(A)}$$



(a) *Before folding.*

(b) *After folding.*

FIGURE 7.3 – *A graph rule (a) before and (b) after the folding operation.*

7.4 Methodology step 3 : a SROIQ formalization [$FORT - lite_{SROIQ}$]

In this section we illustrate the translation of FORT through employing the proposed translation procedure. The original FOL theory, illustrated in Chapter 5, consists of 47 formulas, in addition to 24 mereo(topo)logical axioms and 26 location axioms, thus 97 axioms in total, composing the initial set $FORT_{S0}$. For steps 1 to 5, the transformations are long and straightforward, so we present only the translation of a single axiom ; the specific existential dependence axiom (*SED*) in 7.4.1. After that, we provide the output of fifth step ; $FORT_{S5}$ the non-structured set of *SROIQ* axioms in 7.4.2, and demonstrate the application of the sixth step on the whole theory resulting in $FORT_{S6}$ as the structured set of axioms in 7.4.3. The full SROIQ formalization of the theory is also available online on the GitHub repository ; FORT, in the *FORT-lite-SROIQ* section.

7.4.1 Steps (1-5) applied on a single axiom

$$\forall(x,y) SED(x,y) \rightarrow \forall t (E(x,t) \rightarrow E(y,t)) \wedge \neg(x = y) \wedge \exists t E(x,t) \quad \text{(D)}$$

In the following, we express a binary predicate $R(x,y)$ as R_{xy}, and the predicate $x = y$ as Eq_{xy}.

Transforming to Clausal Form :

Upon skolemization SNF, we substitute n and a by the skolem functions f as $\{n \leftarrow f(x,y,t)\}$. For simplicity we will write $f(x,y,m)$ as f.

$$\forall(x,y)[\neg SED_{xy} \vee (\forall t(\neg E_{xt} \vee E_{yt}) \wedge \neg Eq_{xy} \wedge \exists t E_{xt})] \quad \text{(NNF)}$$
$$\forall x,y,t \exists n \forall b[\neg SED_{xy} \vee ((\neg E_{xt} \vee E_{yt}) \wedge \neg Eq_{xy} \wedge E_{xn})] \quad \text{(PNF)}$$
$$\neg SED_{xy} \vee ((\neg E_{xt} \vee E_{yt}) \wedge \neg Eq_{xy} \wedge E_{xf})] \wedge \quad \text{(SNF)}$$
$$(\neg SED_{xy} \vee \neg E_{xt} \vee E_{yt}) \wedge (\neg SED_{xy} \vee \neg Eq_{xy}) \wedge (\neg SED_{xy} \vee E_{xf}) \quad \text{(CNF)}$$

Rewriting as Horn rules :

$$SED_{xy} \wedge E_{xt} \rightarrow E_{yt} \quad \text{(R1)}$$
$$SED_{xy} \rightarrow \neg Eq_{xy} \quad \text{(R2)}$$
$$SED_{xy} \rightarrow E_{xf} \quad \text{(R3)}$$

Qualifying Expressible Horn rules :

R3 does not qualify since the variables are not enclosed.

$$SED_{xy} \wedge E_{xt} \rightarrow E_{yt} \quad \text{(R1)}$$
$$SED_{xy} \rightarrow \neg Eq_{xy} \quad \text{(R2)}$$

Constructing the rule graphs :

$$G = \langle \{x,y,m\}, \{SED_{xy}, E_{yt}\}, \varnothing, y,t : E \rangle \quad \text{(G1)}$$
$$G = \langle \{x,y\}, \{SED_{xy}\}, \varnothing, x,y : \neg Eq \rangle \quad \text{(G2)}$$

Converting into axioms :

$$SED^{-} \circ E \sqsubseteq E \quad \text{(A1)}$$
$$SED \sqsubseteq nEq, wherenEq = \neg Eq \quad \text{(A2)}$$

7.4.2 The non-structured set of SROIQ axioms

We present below the output of step 5 ; $FORT_{S5}$ as the set of 120 non-structured axioms, and the structures R_{NS} and I resembling the set of non-simple roles and the set of proposition builders, respectively.

$$SED^{-} \circ E \sqsubseteq E \quad \text{(a1)}$$
$$SED \sqsubseteq \neg equal \quad \text{(a2)}$$
$$SED \circ negE \sqsubseteq \varnothing \quad \text{(a3)}$$
$$SED \circ SED \sqsubseteq SED \quad \text{(a4)}$$

$$componentOf \sqsubseteq partOf \quad (a5)$$
$$Tra(componentOf) \quad (a6)$$
$$Irr(componentOf) \quad (a7)$$
$$Asy(componentOf) \quad (a8)$$
$$componentOf \sqsubseteq properPartOf \quad (a9)$$
$$componentOf \circ \neg PartOf^{-} \sqsubseteq \varnothing \quad (a10)$$
$$overlaps \circ componentOf \sqsubseteq \varnothing \quad (a11)$$
$$elementOf \sqsubseteq partOf \quad (a12)$$
$$elementOf \sqsubseteq SED \quad (a13)$$
$$Tra(elementOf) \quad (a14)$$
$$Irr(elementOf) \quad (a15)$$
$$Asy(elementOf) \quad (a16)$$
$$elementOf \sqsubseteq properPartOf \quad (a17)$$
$$elementOf \circ \neg PartOf^{-} \sqsubseteq \varnothing \quad (a18)$$
$$overlaps \circ elementOf \sqsubseteq \varnothing \quad (a19)$$
$$Tra(partOf) \quad (a20)$$
$$Ref(partOf) \quad (a21)$$
$$equal \sqsubseteq partOf \quad (a22)$$
$$equal \sqsubseteq partOf^{-} \quad (a23)$$
$$Tra(equal) \quad (a24)$$
$$Ref(equal) \quad (a25)$$
$$Sym(equal) \quad (a26)$$
$$properPartOf \sqsubseteq partOf \quad (a27)$$
$$properPartOf \sqsubseteq \neg partOf^{-} \quad ||Dis(properPartOf, partOf^{-}) \quad (a28)$$
$$Tra(properPartOf) \quad (a29)$$
$$Irr(properPartOf) \quad (a30)$$
$$Asy(properPartOf) \quad (a31)$$
$$partOf^{-} \circ partOf \sqsubseteq overlaps \quad (a32)$$
$$Ref(overlaps) \quad (a33)$$
$$Sym(overlaps) \quad (a34)$$
$$partOf \circ partOf^{-} \sqsubseteq underlaps \quad (a35)$$
$$Ref(underlaps) \quad (a36)$$
$$Sym(underlaps) \quad (a37)$$
$$overcross \sqsubseteq overlaps \quad (a38)$$
$$overcross \sqsubseteq \neg partOf \quad ||Dis(overcross, partOf) \quad (a39)$$
$$Ref(overcross) \quad (a40)$$
$$Sym(overcross) \quad (a41)$$
$$undercross \sqsubseteq underlaps \quad (a42)$$
$$undercross \sqsubseteq \neg partOf^{-} \quad ||Dis(undercross, partOf^{-}) \quad (a43)$$
$$properOverlap \sqsubseteq overcross \quad (a44)$$
$$properOverlap \sqsubseteq overcross^{-} \quad (a45)$$

$negPartOf \equiv \neg partOf$ (a46)

$negOverlaps \equiv \neg overlaps$ (a47)

$properUnderlap \sqsubseteq undercross$ (a48)

$properUnderlap \sqsubseteq undercross^{-}$ (a49)

$properExtension \sqsubseteq \neg partOf \quad ||Dis(properExtension, partOf)$ (a50)

$properExtension \sqsubseteq partOf^{-}$ (a51)

$C_{overlaps} \equiv \exists overlaps.\top$ (a52)

$C_{overlaps} \sqsubseteq \exists R_{overlaps}.SELF$ (a53)

$R_{overlaps} \circ \neg overlaps \circ underlaps \sqsubseteq overlaps$ (a54)

$C_{partOf} \equiv \exists partOf.\top$ (a55)

$C_{partOf} \sqsubseteq \exists R_{partOf}.SELF$ (a56)

$R_{partOf} \circ \neg partOf \circ overlaps \sqsubseteq partOf$ (a57)

$Ref(connected)$ (a58)

$Sym(connected)$ (a59)

$externallyConnected \sqsubseteq connected$ (a60)

$externallyConnected \sqsubseteq \neg overlaps \quad ||Dis(externallyConnected, overlaps)$ (a61)

$tangentialPartOf \sqsubseteq partOf$ (a62)

$internalPartOf \sqsubseteq partOf$ (a63)

$internalPartOf \sqsubseteq \neg tangentialPartOf \quad ||Dis(internalPartOf, tangentialPartOf)$ (a64)

$Ref(EL)$ (a65)

$Tra(EL)$ (a66)

$partOf \sqsubseteq EL$ (a67)

$partOf \circ R_{EL} \sqsubseteq EL$ (a68)

$C_{EL} \equiv \exists EL.\top$ (a69)

$C_{EL} \sqsubseteq \exists R_{EL}.SELF$ (a70)

$EL \circ partOf \sqsubseteq EL$ (a71)

$EL \circ L \sqsubseteq WL$ (a72)

$L^{-} \circ EL \circ L \sqsubseteq partOf$ (a73)

$L^{-} \circ L \sqsubseteq equal$ (a74)

$partOf^{-} \circ L \sqsubseteq PL$ (a75)

$\neg partOf \circ PL \sqsubseteq \varnothing$ (a76)

$PL \circ L^{-} \sqsubseteq \varnothing$ (a77)

$tangentialPartOf^{-} \circ L \sqsubseteq TPL$ (a78)

$\neg tangentialPartOf \circ TPL \sqsubseteq \varnothing$ (a79)

$TPL \circ L^{-} \sqsubseteq \varnothing$ (a80)

$internalPartOf^{-} \circ L \sqsubseteq IPL$ (a81)

$\neg internalPartOf \circ IPL \sqsubseteq \varnothing$ (a82)

$IPL \circ L^{-} \sqsubseteq \varnothing$ (a83)

$L \circ partOf \sqsubseteq WL$ (a84)

$WL \circ \neg partOf^{-} \sqsubseteq \varnothing$ (a85)

$WL^{-} \circ \neg L \sqsubseteq \varnothing$ (a86)

$$L \circ tangentialPartOf \sqsubseteq TWL \tag{a87}$$
$$TWL \circ \neg tangentialPartOf^{-} \sqsubseteq \varnothing \tag{a88}$$
$$TWL^{-} \circ \neg L \sqsubseteq \varnothing \tag{a89}$$
$$L \circ internalPartOf \sqsubseteq IWL \tag{a90}$$
$$IWL \circ \neg internalPartOf^{-} \sqsubseteq \varnothing \tag{a91}$$
$$IWL^{-} \circ \neg L \sqsubseteq \varnothing \tag{a92}$$
$$L \sqsubseteq PL \tag{a93}$$
$$L \sqsubseteq WL \tag{a94}$$
$$L \circ partOf^{-} \sqsubseteq PL \tag{a95}$$
$$L \circ tangentialPartOf^{-} \sqsubseteq TPL \tag{a96}$$
$$L \circ internalPartOf^{-} \sqsubseteq IPL \tag{a97}$$
$$PL \circ partOf^{-} \sqsubseteq PL \tag{a98}$$
$$TPL \circ partOf^{-} \sqsubseteq TPL \tag{a99}$$
$$IPL \circ partOf^{-} \sqsubseteq IPL \tag{a100}$$
$$partOf \circ WL \sqsubseteq WL \tag{a101}$$
$$internalPartOf \circ IWL \sqsubseteq IWL \tag{a102}$$
$$partOf \circ PL \sqsubseteq PL \tag{a103}$$
$$IPL \sqsubseteq PL \tag{a104}$$
$$TPL \sqsubseteq PL \tag{a105}$$
$$IWL \sqsubseteq WL \tag{a106}$$
$$TWL \sqsubseteq WL \tag{a107}$$
$$Irr(memberOf) \tag{a108}$$
$$Asy(memberOf) \tag{a109}$$
$$memberOf \sqsubseteq properPartOf \tag{a110}$$
$$negMemberOf \equiv \neg memberOf \tag{a111}$$
$$properOverlap^{-} \circ memberOf \sqsubseteq \neg memberOf \tag{a112}$$
$$properPartOf^{-} \circ memberOf \sqsubseteq \neg memberOf \tag{a113}$$
$$properPartOf \circ memberOf \sqsubseteq \neg memberOf \tag{a114}$$
$$overlaps \circ memberOf \circ R_{memberOf} \sqsubseteq overlaps \tag{a115}$$
$$C_{memberOf} \equiv \exists memberOf^{-}.\top \tag{a116}$$
$$C_{memberOf} \sqsubseteq \exists R_{memberOf}.SELF \tag{a117}$$
$$Irr(constitutes) \tag{a118}$$
$$Asy(constitutes) \tag{a119}$$
$$Tra(constitutes) \tag{a120}$$
$$partOf \circ R_{constitutes} \sqsubseteq elementOf \tag{a121}$$
$$C_{constitutes} \equiv \exists constitutes.\top \tag{a122}$$
$$C_{constitutes} \sqsubseteq \exists R_{constitutes}.SELF \tag{a123}$$
$$negEqual \equiv \neg equal \tag{a124}$$

Based on the preceding 124 axioms, we specify the set of non-simple roles R_{NS}, and the set of proposition builders I (all proposition builder are stated afterwards) as follows. In R_{NS}, the denotation $rolename^n$ refers to a rolename that is added to R_{NS} because of axiom number n in S_5 i.e. axiom n lead to the non-simplicity of the role *rolename*.

$$R_{NS} = \{E^1, negEqual^2, SED^4, partOf^5, componentOf^6, elementOf^{14}, ppartOf^{9,17}, equal^{24}, overlaps^{32}, underlaps^{35}, EL^{66}, WL^{72}, PL^{75,95}, TPL^{78}, IPL^{81}, TWL^{87}, IWL^{90}, negMemberOf^{108}, constitutes^{114}\}.$$

$$I = \{\mathbb{I}_2, \mathbb{I}_5, \mathbb{I}_9, \mathbb{I}_{12}, \mathbb{I}_{13}, \mathbb{I}_{17}, \mathbb{I}_{22}, \mathbb{I}_{23}, \mathbb{I}_{27}, \mathbb{I}_{32}, \mathbb{I}_{35}, \mathbb{I}_{38}, \mathbb{I}_{42}, \mathbb{I}_{44}, \mathbb{I}_{45}, \mathbb{I}_{48}, \mathbb{I}_{49}, \mathbb{I}_{51}, \mathbb{I}_{54}, \mathbb{I}_{57}, \mathbb{I}_{60}, \mathbb{I}_{62}, \mathbb{I}_{63}, \mathbb{I}_{67}, \mathbb{I}_{68}, \mathbb{I}_{72}, \mathbb{I}_{73}, \mathbb{I}_{74}, \mathbb{I}_{75}, \mathbb{I}_{78}, \mathbb{I}_{81}, \mathbb{I}_{84}, \mathbb{I}_{87}, \mathbb{I}_{90}, \mathbb{I}_{93}, \mathbb{I}_{94}, \mathbb{I}_{95}, \mathbb{I}_{96}, \mathbb{I}_{97}, \mathbb{I}_{98}, \mathbb{I}_{99}, \mathbb{I}_{100}, \mathbb{I}_{101}, \mathbb{I}_{102}, \mathbb{I}_{103}, \mathbb{I}_{104}, \mathbb{I}_{105}, \mathbb{I}_{106}, \mathbb{I}_{107}, \mathbb{I}_{110}, \mathbb{I}_{112}, \mathbb{I}_{113}, \mathbb{I}_{114}, \mathbb{I}_{115}, \mathbb{I}_{121}\}.$$

$\mathbb{I}_2 \rightarrow (SED \prec negEqual) \wedge (SED^- \prec negEqual)$

$\mathbb{I}_5 \rightarrow (componentOf \prec partOf) \wedge (componentOf^- \prec partOf)$

$\mathbb{I}_9 \rightarrow (componentOf \prec ppartOf) \wedge (componentOf^- \prec ppartOf)$

$\mathbb{I}_{12} \rightarrow (elementOf \prec partOf) \wedge (elementOf^- \prec partOf)$

$\mathbb{I}_{13} \rightarrow (elementOf \prec SED) \wedge (elementOf^- \prec SED)$

$\mathbb{I}_{17} \rightarrow (elementOf \prec ppartOf) \wedge (elementOf^- \prec ppartOf)$

$\mathbb{I}_{22} \rightarrow (equal \prec partOf) \wedge (equal^- \prec partOf)$

$\mathbb{I}_{23} \rightarrow (equal \prec partOf^-) \wedge (equal^- \prec partOf^-)$

$\mathbb{I}_{27} \rightarrow (ppartOf \prec partOf) \wedge (ppartOf^- \prec partOf)$

$\mathbb{I}_{32} \rightarrow (partOf \prec overlaps) \wedge (partOf^- \prec overlaps)$

$\mathbb{I}_{35} \rightarrow (partOf \prec underlaps) \wedge (partOf^- \prec underlaps)$

$\mathbb{I}_{38} \rightarrow (overcross \prec overlaps) \wedge (overcross^- \prec overlaps)$

$\mathbb{I}_{42} \rightarrow (undercross \prec underlaps) \wedge (undercross^- \prec underlaps)$

$\mathbb{I}_{44} \rightarrow (poverlaps \prec overcross) \wedge (poverlaps^- \prec overcross)$

$\mathbb{I}_{45} \rightarrow (poverlaps \prec overcross^-) \wedge (poverlaps^- \prec overcross^-)$

$\mathbb{I}_{48} \rightarrow (punderlaps \prec undercross) \wedge (punderlaps^- \prec undercross)$

$\mathbb{I}_{49} \rightarrow (punderlaps \prec undercross^-) \wedge (punderlaps^- \prec undercross^-)$

$\mathbb{I}_{51} \rightarrow (pExtension \prec partOf^-) \wedge (pExtension^- \prec partOf^-)$

$\mathbb{I}_{54} \rightarrow (R_{overlaps} \prec overlaps) \wedge (R^-_{overlaps} \prec overlaps) \wedge (negOverlaps \prec overlaps) \wedge (negOverlaps^- \prec overlaps) \wedge (underlaps \prec overlaps) \wedge (underlaps^- \prec overlaps)$

$\mathbb{I}_{57} \rightarrow (R_{partOf} \prec partOf) \wedge (R^-_{partOf} \prec partOf) \wedge (negPartOf \prec partOf) \wedge (negPartOf^- \prec partOf) \wedge (overlaps \prec partOf) \wedge (overlaps^- \prec partOf)$

$\mathbb{I}_{60} \rightarrow (externallyConnected \prec connected) \wedge (externallyConnected^- \prec connected)$

$\mathbb{I}_{62} \rightarrow (tangentialPartOf \prec partOf) \wedge (tangentialPartOf^- \prec partOf)$

$\mathbb{I}_{63} \rightarrow (internalPartOf \prec partOf) \wedge (internalPartOf^{-} \prec partOf)$

$\mathbb{I}_{67} \rightarrow (partOf \prec EL) \wedge (partOf^{-} \prec EL)$

$\mathbb{I}_{68} \rightarrow (R_{EL} \prec EL) \wedge (R_{EL}^{-} \prec EL) \wedge (partOf \prec EL) \wedge (partOf^{-} \prec EL)$

$\mathbb{I}_{72} \rightarrow (EL \prec WL) \wedge (EL^{-} \prec WL) \wedge (L \prec WL) \wedge (L^{-} \prec WL)$

$\mathbb{I}_{73} \rightarrow (L \prec partOf) \wedge (L^{-} \prec partOf) \wedge (EL \prec partOf) \wedge (EL^{-} \prec partOf)$

$\mathbb{I}_{74} \rightarrow (L \prec equal) \wedge (L^{-} \prec equal)$

$\mathbb{I}_{75} \rightarrow (partOf \prec PL) \wedge (partOf^{-} \prec PL) \wedge (L \prec PL) \wedge (L^{-} \prec PL)$

$\mathbb{I}_{78} \rightarrow (TP \prec TPL) \wedge (TP^{-} \prec TPL) \wedge (L \prec TPL) \wedge (L^{-} \prec TPL)$

$\mathbb{I}_{81} \rightarrow (IP \prec IPL) \wedge (IP^{-} \prec IPL) \wedge (L \prec IPL) \wedge (L^{-} \prec IPL)$

$\mathbb{I}_{84} \rightarrow (L \prec WL) \wedge (L^{-} \prec WL) \wedge (partOf \prec WL) \wedge (partOf^{-} \prec WL)$

$\mathbb{I}_{87} \rightarrow (L \prec TWL) \wedge (L^{-} \prec TWL) \wedge (TP \prec TWL) \wedge (TP^{-} \prec TWL)$

$\mathbb{I}_{90} \rightarrow (L \prec IWL) \wedge (L^{-} \prec IWL) \wedge (IP \prec IWL) \wedge (IP^{-} \prec IWL)$

$\mathbb{I}_{93} \rightarrow (L \prec PL) \wedge (L^{-} \prec PL)$

$\mathbb{I}_{94} \rightarrow (L \prec WL) \wedge (L^{-} \prec WL)$

$\mathbb{I}_{95} \rightarrow (L \prec PL) \wedge (L^{-} \prec PL) \wedge (partOf \prec PL) \wedge (partOf^{-} \prec PL)$

$\mathbb{I}_{96} \rightarrow (L \prec TPL) \wedge (L^{-} \prec TPL) \wedge (TP \prec TPL) \wedge (TP^{-} \prec TPL)$

$\mathbb{I}_{97} \rightarrow (L \prec IPL) \wedge (L^{-} \prec IPL) \wedge (IP \prec IPL) \wedge (IP^{-} \prec IPL)$

$\mathbb{I}_{98} \rightarrow (partOf \prec PL) \wedge (partOf^{-} \prec PL)$

$\mathbb{I}_{99} \rightarrow (partOf \prec TPL) \wedge (partOf^{-} \prec TPL)$

$\mathbb{I}_{100} \rightarrow (partOf \prec IPL) \wedge (partOf^{-} \prec IPL)$

$\mathbb{I}_{101} \rightarrow (partOf \prec WL) \wedge (partOf^{-} \prec WL)$

$\mathbb{I}_{102} \rightarrow (partOf \prec IWL) \wedge (partOf^{-} \prec IWL)$

$\mathbb{I}_{103} \rightarrow (partOf \prec PL) \wedge (partOf^{-} \prec PL)$

$\mathbb{I}_{104} \rightarrow (IPL \prec PL) \wedge (IPL^{-} \prec PL)$

$\mathbb{I}_{105} \rightarrow (TPL \prec PL) \wedge (TPL^{-} \prec PL)$

$\mathbb{I}_{106} \rightarrow (IWL \prec WL) \wedge (IWL^{-} \prec WL)$

$\mathbb{I}_{107} \rightarrow (TWL \prec WL) \wedge (TWL^{-} \prec WL)$

$\mathbb{I}_{110} \rightarrow (memberOf \prec ppartOf) \wedge (memberOf^{-} \prec ppartOf)$

$\mathbb{I}_{112} \rightarrow (properOverlap \prec negMemberOf) \wedge (properOverlap^{-} \prec negMemberOf)$
$\wedge (memberOf \prec negMemberOf) \wedge (memberOf^{-} \prec negMemberOf)$

$\mathbb{I}_{113} \rightarrow (ppartOf \prec negMemberOf) \wedge (ppartOf^{-} \prec negMemberOf) \wedge (memberOf \prec negMemberOf)$
$\wedge (memberOf^{-} \prec negMemberOf)$

$\mathbb{I}_{114} \rightarrow (ppartOf \prec negMemberOf) \wedge (ppartOf^{-} \prec negMemberOf) \wedge (memberOf \prec negMemberOf)$
$\wedge (memberOf^{-} \prec negMemberOf)$

$\mathbb{I}_{115} \rightarrow (memberOf \prec overlaps) \wedge (memberOf^{-} \prec overlaps) \wedge (R_{memberOf} \prec overlaps) \wedge (R_{memberOf}^{-}$
$\prec overlaps)$

$\mathbb{I}_{121} \rightarrow (partOf \prec elementOf) \wedge (partOf^{-} \prec elementOf) \wedge (R_{constitutes} \prec elementOf) \wedge (R_{constitutes}^{-}$
$\prec elementOf)$

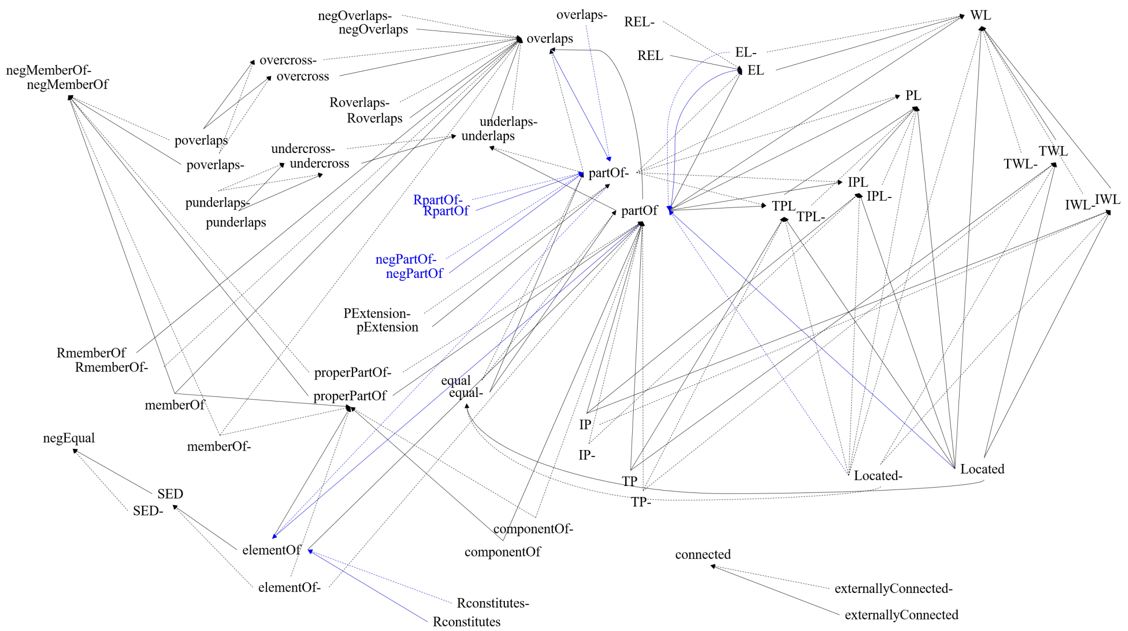


FIGURE 7.4 – *The role hierarchy of FORT presenting the regular orders in black, and the irregular ones in blue.*

7.4.3 Generalization and establishing decidability

We show below how step 6 is applied on $FORT_{S5}$ by applying both simplicity and regularity rules, and we provide the final output of the procedure ; the structured SROIQ set $FORT_{S6}$.

Applying the simplicity rule :

Upon checking the non-simple roles in R_{NS} and the corresponding axioms in S_5 in which they are found, the simplicity rule obliges the suppression of some axioms :
$S_{5'} = S_5 - \{a_7, a_8, a_{15}, a_{16}, a_{28}, a_{30}, a_{31}, a_{39}, a_{43}, a_{50}, a_{61}, a_{118}, a_{119}\}$

Applying the regularity rule :

Based on $S_{5'}$, we construct the (irregular) role hierarchy by translating each *RIA* into a regular order between its roles and add it into the role hierarchy. The resulting hierarchy is shown in figure 7.4. It shows three incompatibilities due to the orders shown *blue* which violate the totality of the regularity of the role hierarchy. We identify the following incompatibilities in function of the axioms that induced them as follows :

$m_1 : \mathbb{I}_{57} \rightarrow \neg\mathbb{I}_{32}$

$m_2 : \mathbb{I}_{73} \rightarrow \neg\mathbb{I}_{67} \wedge \neg\mathbb{I}_{68}$

$m_3 : \mathbb{I}_{121} \rightarrow \neg\mathbb{I}_{12}$

Thus, besides $S_{5'}$ we have the following structures :

$\mathscr{M} = \{m_1, m_2, m_3\}$

$\mathscr{U} = \{a_{57}, a_{32}, a_{73}, a_{67}, a_{68}, a_{121}, a_{12}\}$

Now, we are able to compute the structured subsets of $\mathscr{U}$ by computing the sub-theories :

$\mathbb{M}_1 = \{m_1\}, where \mathbb{U}_1 = \langle\{a_{57}\}, \{a_{32}\}\rangle$
$\mathbb{M}_2 = \{m_2\}, where \mathbb{U}_2 = \langle\{a_{73}\}, \{a_{67}, a_{68}\}\rangle$
$\mathbb{M}_3 = \{m_3\}, where \mathbb{U}_3 = \langle\{a_{121}\}, \{a_{12}\}\rangle$

As a last step, from each tuple $\mathbb{U}_n$, we make the choice of suppressing one set yielding in S_6 as a structured subset of $S_{5'}$, such that $S_6 = S_{5'} - \{a_{57}, a_{73}, a_{121}\}$.
The final structured subset S_6 consists of 108 axioms where the inputted set S_5 consisted of 124 axioms. Thus 16 axioms (a_7, a_8, a_{15}, a_{16}, a_{28}, a_{30}, a_{31}, a_{39}, a_{43}, a_{50}, a_{61}, a_{118}, a_{119}, a_{57}, a_{73}, a_{121}) were suppressed in total upon applying the simplicity and regularity rules.

Other DL formalizations of foundational ontologies

The OWL versions of DOLCE and BFO are available online as DOLCE-Lite [1] (an ontology of both concepts and relations making heavy use of object propoerties), and BFO2.0 [2] (as an ontology that is a taxonomy of categories whereas extensions with relations are made via the RO), to assist in the ontology development process, in which some domain ontologies that use them are listed on their sites. While DOLCE-Lite is encoded in the *SHI* DL, BFO is simpler encoded in *ALC*. However, neither one uses all OWL-DL capabilities of $SHOIN(D)$, nor all OWL 2 DL features. As for UFO, a *SROIQ* DL formalization, s the formal underpinning of OWL 2 DL.

7.5 Conclusion

In this Chapter, we have extracted a secondary decidable specification (FORT lightweight ontology) of the original expressive one (FORT reference ontology), by translating the original FOL formalization into a decidable fragment of $SROIQ - DL$ formalization [methodology-step-3].

For this translation, no tools or even principled works exist for providing guidelines for the translation of FOL theories into SROIQ knowledge bases. To put a step forward towards addressing this subject and to eventually translate FORT, we have proposed a procedure for translating FOL theories (with one or two arity predicates) into decidable *SROIQ* fragments. The procedure computes, at each step, a particular logical formalism of the initial theory (e.g. Clausal Normal Form, Horn Rules Form, Graph Rules Form, structured and non-structured *SROIQ* fragments). This is by performing operations such as rewriting formulas, syntactic/semantic checks, graph transformations, and rule-rolling techniques. The approach exploits the literature by re-using extant convenient techniques, such as constructing graph rules from [Gasse2008] and structuring a TBox from [Botti Benevides2019], and assembles them together with other operations to form a series of consecutive steps.

After that, we applied the procedure on FORT to translate it into a *SROIQ* fragment yielding in a *TBox* of 108 structured axioms. Thus, we have established in this Chapter [micro-objective-3] by obtaining a secondary decidable formalization of the our proposed language of relations.

1. http://www.loa.istc.cnr.it/ontologies/DOLCE-Lite.owl
2. https://github.com/BFO-ontology/BFO2

For the semantic web, developing ontological models starting from logical formalisms is widely growing nowadays. In cases where these logical formalisms are FOL-based, our procedure facilitates the task of extracting a decidable the SROIQ-based logical formalism, upon which the OWL2DL web ontology language is based. Thus, the procedure aids the shift from a the initial logical framework to to implement an eventual OWL2 ontology for practice.

Moreover, although the research in the field of knowledge representation and reasoning continues to add new constructs and relax other restrictions on DL families to maximize expressivity and still maintain decidability, there are still some forms of knowledge that cannot be captured by current DLs. Examples include the basic antisymmetry in mereological theories, as well as the restrictions on non-simple roles. These are indeed proof of the inadequacy of DLs in the specification of meta-ontologies as FOs and FORT. In practice, extant semantic web technologies have been developed to enhance expressivity abilities, such as the Semantic Web Rule Language (SWRL) and the Shapes Constraints Language (SHACL).

In the following Chapter 8, we focus on the FORT lightweight ontology, in which we move forward towards our last contribution addressing the last three steps of the proposed methodology.

8

A lightweight ontology of foundational relations ; application proposals

This Chapter is a design for the proposal's continuation, rather than a complete contribution.

Sommaire

8.1 Introduction

In the previous Chapter 7, we extracted a secondary decidable specification of FORT by translating the original FOL-formalization into a SROIQ knowledge base. Thus, we have obtained a decidable lite fragment of the proposed language of relations.

In this Chapter, we aim to cover the contribution of providing FORT as a lightweight ontology in practice [micro-objective-4], while showing its possible employment methods [micro-objective-5] and demonstrating its convenience through the cultural heritage application [micro-objective-6]. This could be achieved by implementing an OWL2 ontology of FORT [methodology-step-4], studying its possible importation and linkage scenarios [methodology-step-5], and demonstrating some scenario using a real world application [methodology-step-6].

We first provide in Section 8.2.2 the FORT OWL2 ontology. Then, we discuss in Section 8.3 the possible employment methods of FORT based on the application's setting. After that, based on limited time and inputs from the application domain, we discuss requirements for a complete application and conclude in Section 8.4.

8.2 Methodology step 4 : an OWL ontology [$FORT - lite_{OWL}$]

In this section, we initially discuss the significance of the Semantic Web (SW) and its associated technologies as a domain to contextualize our theory in. Indeed, the current FORT ontology allows for further exploration of alternative approaches, such as within a conceptual modeling framework employing UML for ontology-driven conceptual modeling tasks. Nevertheless, we opted to proceed with SW technologies, as elucidated below. Then, we introduce the OWL2DL implementation of FORT, serving as an ontology represented in a Semantic Web language to facilitate its application in semantic web-oriented applications.

8.2.1 Context : the relevance of the Semantic Web

The SW domain offers multiple advantages in terms how information is structured and shared, emphasizing the meaningful interpretation of data across diverse domains. Several technologies within the SW domain support knowledge representation in general, and ontology handling in particular, such as the Resource Description Framework (RDF) as a graph model for data, the RDF Schema (RDFS) and the Web Ontology Language (OWL) for defining vocabularies and ontologies, the SPARQL Protocol And RDF Query Language (SPARQL) as a query language dedicated to RDF, in addition to other supported services for inferences and graph validation.

Additionally, the adoption of the principles of the Semantic Web fosters integration with Linked Open Data (LOD) initiatives. LOD emphasizes the publication and interlinking of structured data on the web, creating a vast network of interconnected resources.

Furthermore, when it comes to interdisciplinary fields, as discussed in Chapter 1 there is a crucial need to contextualize data, link them to their different aspects (historical context, art context, material context, etc.) and to each other. Such contextualized and interlinked information allows for a holistic comprehension of an interdisciplinary domain facilitating further interdisciplinary analyses.

Taking the cultural heritage field as an interdisciplinary example, in [Vavliakis2012], the authors present a study which focus on the challenges imposed by the multidimensionality of information in this domain (which has been also presented in Chapter 1), and discuss the significance of the SW and LOD initiatives in addressing this multidimensionality by contextualizing and interlinking data.

For that, the authors provide an example that highlights the advantages of incorporating CH data into the SW and leveraging LOD access. The example pictures the case of archaeologist performing an excavation in an ancient Macedonian tomb in Northern Greece. While having their crew already unearthed the dome of the tomb and working towards revealing the rest of the structure, using their link to the web (and to the world), their application queries a Linked open data repository and finds information of similar excavation sites. Based on these resources, the application gives inferences that given the materials found in another similar tomb, the ongoing operations, and the reported humidity of the soil, there is a considerable risk of collapse. Then, a strong recommendation of halting the operation is given. Moreover, while a restorer was working on a damages golden crown found in that tomb, further knowledge about its sources was needed, whether is of Macedonian or Persian style. Using the object's report, which contains spectroscopic and μRaman analysis data identifying the object's materials, the application inferences that is most probably of Persian style, for that the orpiment mixture found in the object, was used to be extracted from the mines of Persia and at the creation time of the crown it was common among Persian blacksmiths.

This example not only underscores the importance of adopting LOD principles but also emphasizes the value of contextualizing data in a precise and interconnected manner to gain valuable insights. Moreover, it also serves as a compelling demonstration of the significance of studies focusing on the materiality of entities, such as analyzing the materials of tombs and soil to infer similarities between tombs, or examining the composition of crowns to determine the sources of their materials. And this underscores the importance of our proposed FORT approach which enables a detailed and expressive representation of tangible entities, like archaeological sites and artifacts, through relations. Thereby emphasizing the significance of the FORT framework and the importance of integrating it into the SW for future benefits.

8.2.2 The implementation of FORT

For implementing FORT, we use the Web Ontology Language (OWL)a knowledge representation language capable of defining ontologies with high expressiveness by employing constructs from Description Logic (DL). The OWL2 recommendation offers two levels of expressiveness, depending on the formal semantics employed : OWL2 DL and OWL2 Full. With the former having direct semantics, decidable properties, and underpinned by the SROIQ DL (in which FORT had been formalized in the previous Chapter 7), we opt to proceed with OWL2 DL.

The implementation is straightforward in which we assert the 108 structured axioms of the SROIQ-Tbox obtained upon translating FORT into its lightweight version yielding in an OWL2 DL ontology. For that, we used the Protégé software, a free ontology editor incorporating several reasoners. Figure 8.1 shows a snapshot from the Protégé software of the ObjectProperty hierarchy of the FORT ontology.

Some rules are then added, as shown in Listing 8.1, using Semantic Web Rule Language (SWRL), a SW technology allowing for writing semantic rules, applied to semantic Web and OWL ontologies [Horrocks2004b]. Although SWRL is a W3C "member submission" and not a "recommendation", meaning it has not yet been assessed for standardization process, it is particularly interesting since it addresses directly OWL ontologies. Other semantic rules *exchange* technologies include the Rule Interchange Format (RIF), a W3C specification for writing and exchanging rules in the SW but not for implementing rules for applications. Additionally, these include the SPARQL Inference Notation (SPIN a W3C member "submission") and the SHACL Advanced Features (SHACL-AF which is based on the SHACL W3C recommendation) semantic rules, which focus on RDF and not OWL (like SWRL) which make them applicable on RDF graphs of data (which contain OWL constructs) for validation of data at the data level rather than assertion of rules at the classes/properties level.

The OWL file is available online at the FORT GitHub repository accessible at : FORT-OWL ontology [1].

```
PartOf(?x, ?y), PartOf(?y, ?x) ->  SameAs (?x, ?y)
LocatedAt(?x, ?y) -> LocatedAt(?y, ?y)
```

Listing 8.1 – Some SWRL rules in FORT.

8.3 Methodology step 5 and 6 : applicability and convenience of FORT

In this section, our objective is to illustrate the practical implementation of FORT as an expressive language of relations. Specifically, we aim to demonstrate two potential methods of using FORT : as a relation language imported into user-based domain ontologies, and as a valuable meta-ontology within a system of ontologies such as interdisciplinary applications like cultural heritage.

To achieve this, we present two employment methods for FORT, each tailored to the application's context and intended goals. The first method involves a direct employment of FORT for use by a single ontology, whereas the second method adopts an indirect use of FORT, catering to interdisciplinary scenarios. These methods are depicted in Figure 8.2, and we provide detailed explanations for each in subsequent sections.

It is important to highlight that none of these systems are implemented as part of this thesis work. We only propose designs of the employment methods as exploration for later implementations and testing.

8.3.1 Direct employment method

The goal of the direct employment method is to use FORT's relations within a user-based domain ontology as an expressive language to semantically enhance domain relations. This is by importing FORT as an ontology (i.e. at the global ontology level) into the domain ontology of a user (i.e. at the domain ontology level), followed by an ontology alignment task as depicted in Figure 8.2, under the "direct method" section.

In this case, the alignment task concerns building correspondence between the relations of FORT, denoted R_{FORT}, and those of the domain ontology, denoted R_{domain}, [Euzenat2007] :

— an equivalence correspondence denoted $\langle R_{domain}, \equiv, R_{FORT} \rangle$ relating two roles R_{domain} and R_{FORT}, if they are semantically equivalent ($R_{domain} \equiv R_{FORT}$)

— a subsumed-by correspondence denoted $\langle R_{domain}, \leq, R_{FORT} \rangle$ relating two roles R_{domain} and R_{FORT}, if R_{domain} is a sub property of R_{FORT} ($R_{domain} \sqsubseteq R_{FORT}$)

— a subsumed-by correspondence denoted $\langle \Omega_{domain}, \leq, R_{FORT} \rangle$ relating a role composition of n in the domain $\Omega_{domain} = R_1 \circ R_2 \circ .. \circ R_n$ to R_{FORT}, if the composition of roles Ω_{domain} yields in a role that is sub property of R_{FORT} ($\Omega_{domain} \sqsubseteq R_{FORT}$)

1. https://github.com/DanashFatima/FORT/tree/5c6390830c2e8cff52df2a6d41725dfd3d5dcf02/FORT-OWL-ontology

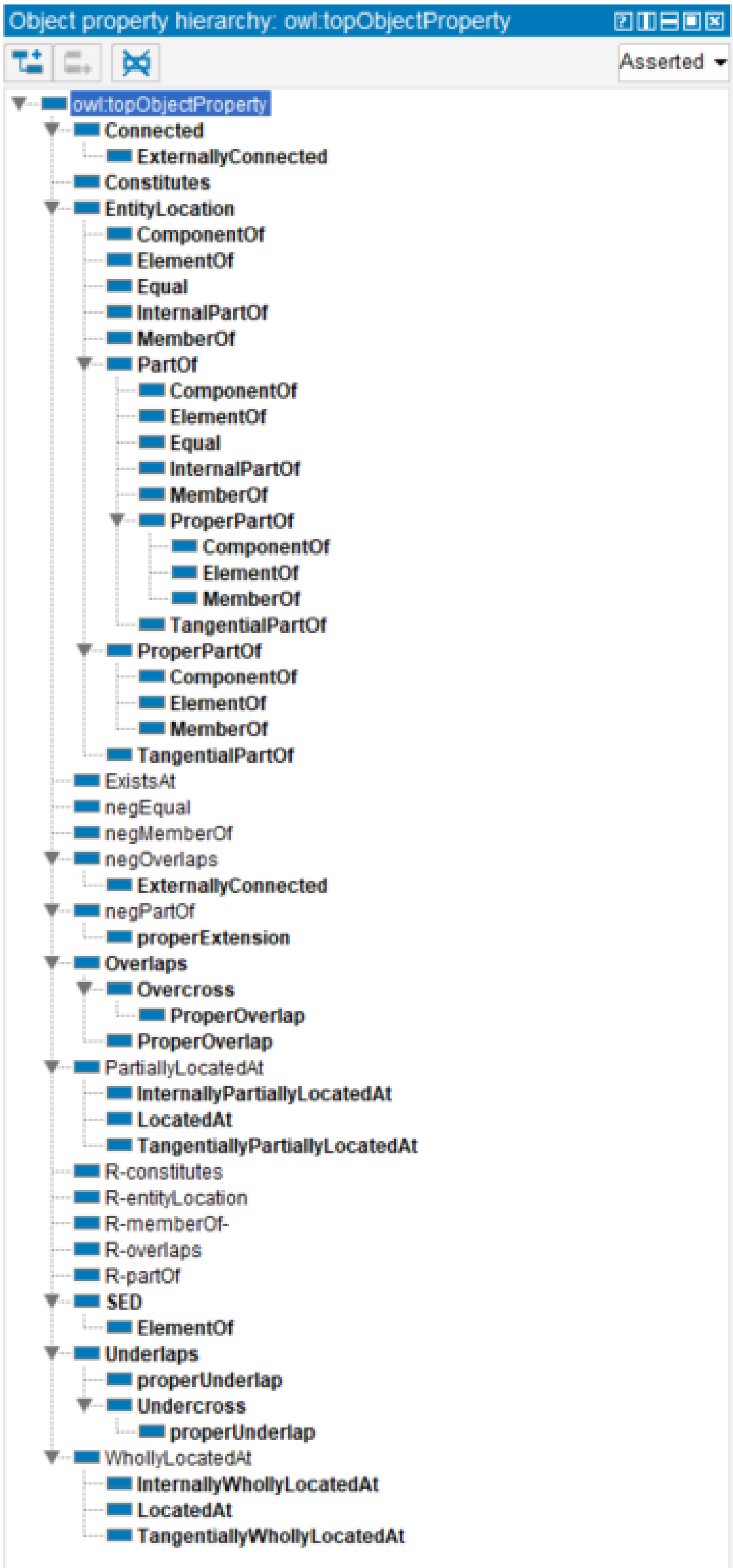


FIGURE 8.1 – *The ObjectProperty hierarchy of the FORT ontology, a snapshot from the Protégé software.*

Upon incorporating FORT, a network of correspondences between the relations of both ontologies is created, which makes room for new inferences. Consequently, this expansion in terms

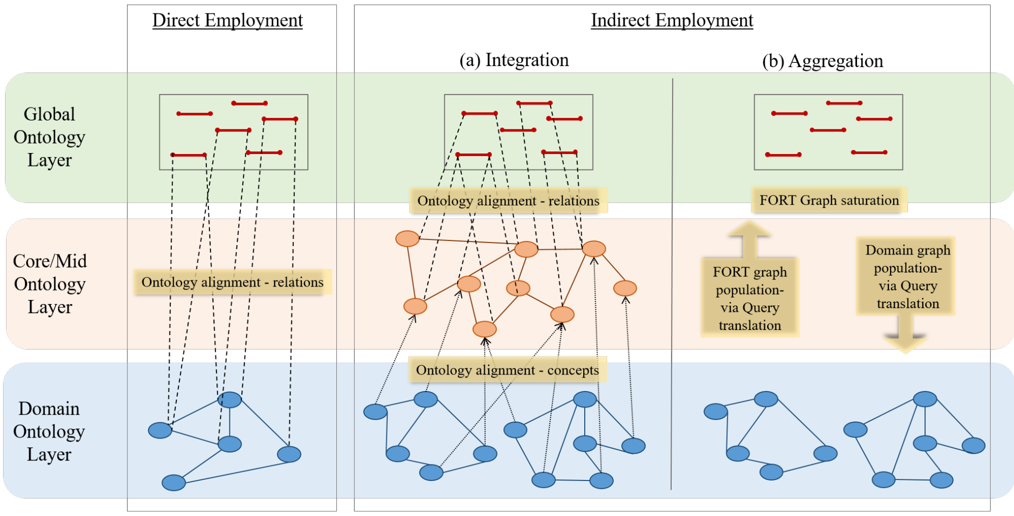


FIGURE 8.2 – *The possible employment methods of FORT, direct and indirect, based on the application's setting.*

in relations provides a wider range of navigation options within the ontology, considering that FORT's relations can be leveraged in the querying process.

8.3.2 Indirect employment method

The goal of the indirect employment method is to exploit the semantics of FORT's relations within an interdisciplinary application in which multiple domain ontologies exist at the domain ontology level, such as the cultural heritage field. Depending on the objectives of the application, data integration or aggregation, we define two indirect employment approaches. In both of which, FORT serves as a relation ontology at the global level, while domain ontologies at the domain level.

Integration-based approach

For an integration-based approach, the primary objective is to merge multiple ontologies at the domain level, thus creating a unified and coherent representation of data. And the use of meta-ontologies for interdisciplinary purposes, as discussed in the Preliminary Remarks section, serves this purpose of *integrating* data and models through ontology-based data integration paradigms.

Furthermore, as discussed in Chapter 1, achieving a shared understanding based on a consensus between the different ontologies is essential for an *interdisciplinary integration*. This necessitates having a global ontology (within an ontology-based data integration paradigm), to facilitate the integration of domain-specific ontologies comprehensively. And based on the analysis in [Ekaputra2017], the GaV approach is particularly relevant as it involves such a global abstraction level through which ontologies are integrated.

Now, we consider FORT as the ontology of relations at a global level for enhancing relation representations. However, alone, it cannot achieve a comprehensive integration for that a shared consensus hierarchy of categories is further needed, besides and hierarchy of relations (provided by FORT). The integration of two ontologies involves identifying some least common subsumer, denoted as C_{global}, that can generalize the fine-grained classes $C_{domain1}$ and $C_{domain2}$ from the domain ontologies before seeking common relations.

Thus, there arises a necessity for an ontology that offers this consensus hierarchy of categories. To address this need, and depending on each application, the addition of the core/mid-level ontology layer [1] is vital. Core ontologies are particularly interesting for multidisciplinary domains aiming for interdisciplinary integration since they define the minimal concepts required to understand the finer concepts across these domains [Falquet2011]. This results in an intended consensus of a hierarchy of concepts while maintaining a certain abstraction level generic enough to avoid domain-dependency. Mid-level ontologies serve a similar purpose, bridging the gap between domain and global levels.

Thus, as depicted in Figure 8.2 under the "Integration" section of the indirect employment method, we propose an approach that involves two steps : first, the construction or selection and mapping of a core/mid-level ontology to the domain ontologies, and second, the exploitation of FORT at the global level with links to the core/mid-level through the alignment of relations from both ontologies, following the direct method explained in Section 8.3.1.

In this approach, the navigation across user-based domain ontologies, which is an integral task during the integration process, is facilitated by the links established between each domain ontology and the corresponding core/mid-level ontology. Meanwhile, the exploitation of rich semantic relation, occurs at the core-mid-level, which incorporates the relations provided by FORT. Thus, domain ontologies can use FORT's relations in an indirect method via the integrating core/mid-level ontology.

Aggregation-based approach

For an aggregation-based approach, the primary objective is to combine data from multiple sources without necessarily achieving a consensus between their respective ontologies. It involves aggregating data from different ontologies into a single system, allowing access to the combined information by multiple domains. The ontologies remain independent and do not undergo modification or alignment.

However, the system in which the data is aggregated is capable of exploiting the semantics of both combined to further infer new data, based on some semantic rules over data graphs. Of course, the application still has its interdisciplinary characteristic where it is trying to make use of the data published by each ontology, to benefit each other ontology. just not the goal of integrating these domain ontologies at the domain level neither via a consensus hierarchy, but merging their results instead.

In this approach we propose the use of FORT, as a relations and rules language, within this aggregation system seeking to infer new data using FORT's semantics. This means that the interdisciplinary application is seeking links between their entities regarding an enhancement of their structural and spatial characteristics.

In the domain of cultural heritage, consider a scenario where there are two domain ontologies : one for archaeological sites and another for objects (findings).

The archaeological site ontology includes details about site descriptions, locations, findings,

1. Refer to Figure 1.7 for the placement of core/mid-level ontologies in the abstraction levels layers.

dates, population, schematic figures, parietal paintings found in the sites, on-site and off-site analysis of these figures, and the results of these analyses.

The findings ontology, on the other hand, describes various artifacts that are found in places and acquire a patrimonial value such as tools, crowns, small statues, painting materials, pots, etc., with describing their structural and spatial characteristics, and multiple types of analysis performed on them

An ontology aggregation system designed aims to combine these two domain ontologies into a unified system without explicitly linking the ontologies. The system aims to create connections and inferences between data elements from both ontologies to facilitate knowledge discovery across both. For example, the system aims to inferring identical locations, the findings at a same or overlapping locations, connections between painting matters found at the same locations, linking artifacts having similar analysis results, etc.

To achieve this, the system could exploits the domain-independent and expressive relations and rules offered in FORT within a SW framework using SW technologies e.g. SPARQL for queries,as well as SPIN/SHACL-AF for semantic rules (validation and inference services) at the data graph level. This required that data of domain ontologies (or relational databases) exist in the for of RDF graphs. [1] As depicted in Figure 8.2 under the "Aggregation" section of the indirect employment method, we propose the following approach :

1. Query Translation : Data from the domain ontologies is translated into FORT's relations, allowing for data aggregation at the global level based on query rewriting
2. Graph Saturation : FORT's rules, along with additional application-specific rules, are used for saturating the graph. The system populates the graph with inferred data, expanding the knowledge base and enabling connections between different entities.
3. Query Translation (Reverse) : Another query translation task ensures that the domain ontologies are updated with the findings from the global level, incorporating data from other ontologies as well.

By employing this approach, the ontology aggregation system effectively combines data from archaeological sites and objects, enabling cross-domain queries, inferences, and knowledge discovery based on query reformulation and saturation-based techniques which could be further investigated in [Bischof2014, Buron2019, Buron2020]. In such approaches, using FORT's relations and rules language plays an important role in achieving a domain-independent and rich structural and spatial based representations and links between entities.

8.4 Conclusion

In this Chapter, we have supported the practice of the FORT lightweight ontology in the Semantic Web by providing an OWL2 DL ontology, thus achieving [micro-objective-4].

Additionally, we have designed the possible employment methods of FORT based on the application's setup and objectives, resulting in direct employment for a single-ontology use, and indirect

1. *RDF stores* or *triplestore* is a type of database specially designed to store information in the form of triplets (subject, property, object), which the form of RDF data, in order to trate them as RDF data, and to retrieve this information through query mechanisms using SPARQL. Examples of these systems which based on SW technologies and used to store knowledge in the form of RDF graphs include GraphDB, a graph database and triplestore.

employment for a multiple ontology use in an interdisciplinary application based on integration or aggregation purposes. However, we have not implemented any of the proposed architectures due to time and input limitations. Thus, we have only designed ground proposals for future completion of [micro-objective-5] and [micro-objective-6], while keeping space for possible refinements.

In order to go further with FORT in an interdisciplinary application such as cultural heritage, it is essential to have rigid structures of data already present, so that the high-level formalization of FORT can capture precisely these structures. This concerns each of the included domain at the domain level and calls for building domain models to represent properly their view point of their entities. It is highly important to shed the light on the fact that only experts within these domain are able to precisely contextualize their data and build structures of these data, upon which a model can be built. It is only with the domain models at hand, along with the data aligning to these models, that the empirical tasks specified in steps [methodology-step-5] and [methodology-step-6] can be fulfilled.

In the following Chapter 9, we conclude this manuscript by recalling the contributions of this work, discussing some limitations, and listing some future research and development directions.

SECTION C :
***
CONCLUSION AND DICUSSION

This last Section concludes the thesis's main contributions and discusses possible future directions.

9

Conclusion

In this Chapter, we conclude this manuscript by recalling the contributions of our work and listing further research and development future directions.

Sommaire

"*Perfection is achieved not when there is nothing more to add, but when there is nothing left to take away.*"
Antoine de Saint-Exupéry.

In the field of patrimonial sciences, including cultural heritage studies, the multidisciplinary nature of the domain poses a fundamental challenge for integrating collaborative knowledge. However, it is through this multidisciplinary lens that insights about cultural heritage entities can be gained. These entities, acting as boundary objects, exhibit an additional cross-disciplinary nature i.e. they bear multiple interpretations while maintaining a recognizable shared identity. While achieving an interdisciplinary dialogue across disciplines is vital for cultural studies, a holistic comprehension and representation of cultural heritage entities is required. This can be only fulfilled by an integrated and interdisciplinary approach, driven by a *shared goal* across various disciplines.

Having the shared goal of studying and constructing the tangible discourse of a cultural heritage entity for transgenerational transmission, the *materiality of the entity* emerges as a significant focal point. It serves as a unifying factor that recognizes the entity's unique identity among the diverse interests of multiple disciplines. Thus, it provides a solid foundation for interdisciplinary integration.

To effectively understand and represent the materiality of tangible cultural heritage entities, *it becomes imperative to employ an ontological model that encompasses semantically rich and well-formalized structural and spatial relations*. Furthermore, for successful interdisciplinary integration, *a meta-ontology approach* focusing on these structural and spatial relations is essential to address the challenges posed by the heterogeneity of multiple disciplines and foster interoperability across various models.

Existing ontological models in the cultural heritage field often have shown either insufficient treatment of composition relations (particularly structural and spatial), data-centric approaches rather than entity-centric approaches, or poorly formalized relations. This highlights the necessity for *a robust and well-formalized set of foundational ontological relations that can accurately capture the composition of a cultural heritage entity* and enable precise representation of its materiality.

Shifting the investigation towards foundational ontological relations, the focus has primarily revolved around taxonomies and theories within the context of structural and spatial "part-whole" relations. While some individual approaches have demonstrated value in the applied ontology field, there need for *a unified theory that collectively integrates a set of foundational ontological relations* was recognized.

Therefore, this thesis goes beyond the scope of cultural heritage and generalizes the research problem to encompass **the representation and modeling of the composition of any tangible entity, using ontological structural and spatial relations through a meta-ontology**. This fundamental objective, along with the corresponding needs identified in the literature, gave rise to six micro-objectives, a selection of which has been addressed in our contributions, as summarized below.

9.1 Summary of contributions

We have proposed in this manuscript **FORT**, a **F**oundational **O**ntological **R**elations **T**heory, within an Applied Ontological approach. For the modelisation aspect of the approach, we have proposed FORT as a meta-ontology specifying a meta-conceptualization using a meta-modeling language. As for the employment aspect, FORT can be employed directly as a relation language, or indirectly for interdisciplinary applications within a GaV approach.

In order to formalize FORT and design its employment within a GaV paradigm for interdisciplinary applications, we constructed and adhered to a a six-steps ontology engineering methodology. Each step is introduced to accomplish a specific micro-objective and ultimately overcome the challenges at hand. This resulted in the following contributions, all of which revolving around FORT, as the key deliverable in our thesis.

⋙ First, **we have proposed an expressive and well-founded language of exclusive relations and rule constraints** [micro-objective-1]. To do so, we have specified and formalized FORT as a reference ontology of relations in the expressive First-order Logic (FOL) [methodology-step-1]. This resulted in *a FOL-formalization of the FORT reference ontology - theoretical level*. The formalization has been carried out with an extensive re-use and re-formalization of theories from the applied ontological literature.

FORT encompasses a minimal set of structural and spatial relations that are indispensable for the representation-of and reasoning-over the composition of a tangible entity. These relations cover our specified intended representations to models links between : an entity as a whole and its different inseparable parts (i.e. parthood and dependence) ; an entity as collective whole and the entities that it groups under certain semantics (i.e. membership) ; an entity and its constituents (i.e. constitution) ; an entity and the spatial region that locates it (i.e. location) ; and the spatial constraints among entities (i.e. entity-location). Hence, we are concerned with parthood (and connection), location, membership, and constitution relations, with the formal properties that characterize them.

FORT has been presented at two levels : a group of microtheories of relations which exploit the literature (e.g. mereology, mereotopology, and location) at a microlevel, a unified macrotheory which interlinks the multiple microtheories at a macrolevel. The forte contributions of FORT have been identified at the macrolevel as twofold.
First, FORT addresses the "requirements" specified for a meta-ontology of relations as being modular, meta, and of exclusive relations. The first requirement (**modular**) has been covered with FORT encompassing a number of relation microtheories, each being intralinked, as well as interlinked among each other. The second requirement (**meta**) was achieved with FORT being a meta-ontology specifying a meta-conceptualization (i.e. top-level abstractions) using a meta-modeling language (i.e. generic modeling primitives). And the third requirement (**exclusive relations**) was implemented during the formalization which does not account for ontological categories, yet ensures semantically well-formalized relation via axiomatic constraints on their domain and range.
Second, FORT offers a "minimal" set of foundational relations that is yet inclusive in its capability in covering the different possible representations of the composition of an entity : its internal structure, spatial conditions, and interrelations with other entities.

⋙ Second, **we have demonstrated the novelty and consistency of our proposed relations language** [micro-objective-2]. For that, we have analyzed the FORT reference ontology in view of extant ontology relation theories and validated its consistency at an empirical level according to [methodology-step-2]. This resulted in *a CL-serialization of the FORT reference ontology - an* *empirical level.*

For the analysis of FORT, we have specified the reasons behind proposing FORT and positioned it in view of the literature. In the former, we elaborated on three arguments behind constructing FORT based on our observations in the extant literature of meta-ontologies. These have argued the absence of a modular theory of exclusive relations, the need for an inclusive incorporation of the intended minimal set of relations, and the difficulties that arise with employing foundational ontologies for the particular use of foundational relations. In the latter, we elucidated a relation-based comparison between each module in FORT and its corresponding aligned relation in selected foundational ontologies.

For the validation of FORT's consistency, we formalized FORT using Common Logic (CL) as a CL-ontology, based on a CLIF serialization. We used the Hets online tool for checking the consistency of the CLIF files, validating thus the existence of models for FORT. Additionally, we translates FORT into other ontology serializations and ran automatic theorem proofs.

⋙ Third, **we have established a decidable lite formalization of our proposed language of relations** [micro-objective-3]. To do so, we have extracted a decidable fragment of the original FOL-formalization as a lightweight version, based on a proposed systematic and generic procedure [methodology-step-3]. This yielded in *a SROIQ-formalization of the FORT lightweight ontology - a theoretical level.*

For retrieving the secondary SROIQ formalization, we built a generic procedure for translating FOL-theories into decidable SROIQ-Tboxes which is the most expressive, yet decidable Description Logic (DL). The procedure computes, at each step, a particular logical formalism of the initial theory by performing operations such as rewriting formulas, syntactic/semantic checks, graph transformations, and rule-rolling techniques. The approach exploits the literature by re-using convenient techniques, such as constructing graph rules and structuring a TBox.

Motivated by the task of translating FORT, we applied the procedure on the FORT reference ontology (the FOL formalization) to obtain the FORT lightweight ontology as the SROIQ-Tbox.

⋙ Fourth, **we have supported the practice of the FORT lightweight ontology** [micro-objective-4]. This is by implmenting the SROIQ-Tbox obtained from the translation into the OWL2 DL Semantic Web language [methodology-step-4] resulting in an *a OWL2-ontology of the FORT lightweight ontology - an empirical level.*

For proving the applicability and convenience of FORT, we have designed two employment methods, direct for single-ontology uses, and indirect for interdisciplinary application based on either integration or aggregation purposes. These methods were presented only as proposals (due to time and input limitations) in which further investigation and implementation is needed.

In conclusion, it is acknowledged that our response to the cultural heritage application, which served as the foundation for this thesis work, may not be entirely comprehensive. Nevertheless, we have taken incremental steps towards addressing their requirement for an entity-centric approach that emphasizes the materiality of the entity. Through this research, we have furnished the necessary tools (the FORT ontology) encompassing ontology modeling and application design proposals, laying the groundwork to for further advancements.

9.2 Discussion : limitations and future directions

In the final section, we discuss some remarks regarding our approach and highlight certain areas that could benefit from significant further exploration. These propositions serve as future directions for both research and development, encompassing various aspects of our work.

— Regarding the FORT reference ontology at a theoretical level, further extensions of the theory could be considered (Propositions 9.2.1 and 9.2.2).

— Regarding the FORT lightweight ontology at an empirical level, an ontological tool to support users could be modeled (Proposition 9.2.3).

— Regarding the practical application of the FORT lightweight ontology for cultural heritage within the Semantic Web, an overall architecture involving both CIDOC CRM and CHARM could be employed (Proposition 9.2.4).

— Regarding the procedure which has been designed for translating FORT (from the reference to the lightweight ontology), a theoretical/empirical validation could be developed for generalization (Proposition 9.2.5).

9.2.1 From an atemporal to a temporal framework

One notable simplification made in Chapter 5 is the treatment of time, which is not explicitly addressed in FORT. This decision has been made for that the ontological choices regarding time should be approached after establishing framework for the theory's fundamental elements : foundational ontological relations. Additionally, temporal considerations are complex and need to be taken in a separate step. Thus, a first perspective to our work is to **extend FORT to a temporal framework**.

Indeed, expanding FORT with time would extend its capabilities from representing the composition of a tangible entity to representing the spatiotemporal evolution of this composition with time. And thus, to draw insights on the lifeline of tangible structures. Additionally, we would be able to study the behavior of each of FORT's relations upon events that alter tangible structures. Thus, drawing insights on the evolution of relations upon changes with time.

To achieve this, several aspects related to the incorporation of time need to be considered. We outline these aspects below, accompanied by references to relevant literature that can serve as starting points for further investigations.

Firstly, there is the question of incorporating time within the current FOL formalization of FORT. This can be addressed by either (1) introducing a time variable using ternary relations [1], (2) adopting an interval-based temporal logic such as the one proposed by Allen [Allen1983], or (3) expressing temporal constraints using a subset of first-order temporal logic, such as the Description Logics DLR extended with temporal operators like "since" and "until" [Artale2002, Artale2007]. These approaches would allow for the representation of entities (e.g. x and y, and the relation holding between them (e.g. R) at different time points (e.g. $t1$, $t2$, etc.), providing a snapshot of their state at each scenario.

Secondly, it is relevant to consider the events that can alter the structural and spatial representation of a tangible entity, and how to model these events. Examples of events can be drawn from

1. Note that in FORT's FOL formalization, we restricted the use of predicates to unary and binary ones only i.e. arity equal one and two. The choice goes back to facilitating that eventual shift from First-order Logic to the SROIQ Description Logic. Hence, upon allowing for ternary predicates within FOL theories, further complications arise.

the LODE ontology [Shaw2009], with a deep understanding of event modeling [Casati1997]. This enables the representation of the continuous lifeline of an entity, demonstrating the changes that have influenced its structural and spatial aspects.

Lastly, attention should be given to the behavior of entities linked by relation R at time $t1$ in response to a change event E. Some valuable insights can be derived from the work of Artale et al. [Artale2008]. This analysis allows for reasoning about the evolution of relations over time and predicting the nature of relation $R2$ at time $t2$ following the occurrence of change event E.

As a result of this *modelisation task, the extension of FORT with temporal constraints would be a significant contribution within the applied ontology field*. The application of this $FORT^{+time}$ would make further contributions to fields where its is employed. The ability to represent the spatiotemporal evolution of tangible entities would have implications in fields such as cultural heritage preservation, architectural design, urban planning, and many others. By capturing the dynamic aspects of tangible structures, this extension would enhance our understanding and enable inferences on the reasons and consequences of changes.

Furthermore, this could be closely examined in conjunction with causality theories, see [Sloman2005]. By considering the causal relationships between events and the resulting changes in the structure and spatial aspects of entities, we can gain deeper insights into the underlying mechanisms driving these changes. This integration of temporal representation and causality theories has the potential to enhance our understanding of how events shape the composition of tangible entities over time.

9.2.2 Transitivity in multi-type relations

Within the FORT framework, we have successfully addressed the transitivity of each individual relation, enabling reasoning on the composition of uni-typed relations, specifically $R1 \circ R1 \rightarrow R1$. A possible perspective regarding this topic, is to **explicitly tackle the issue of multi-typed relation composition**. For example, given $R1(x,y) \wedge R2(y,z)$, it could be the case where the relation holding between x and z could be further determined, if any.

Gaining insights into relations through relation composition offers several key benefits. Firstly, it allows us to deduce new data that may not be explicitly stated within the knowledge base of relationships. Secondly, it facilitates the discovery of patterns or regularities among relationships, enabling us to observe their application to real-world data and utilize these insights for further enhancement. These advantages have far-reaching implications across various domains, including knowledge representation, ontology engineering, data analysis, and artificial intelligence.

To study the composition of relations, a commonly employed method would be constructing a composition table for the different relations, analogous to the approach taken with RCC8 relations [1] as shown in Table 9.1. It could be possible to study the set constraints, say $S1$, projected by $R1$ on y as its range, versus the set of constraints $S2$, projected by $R2$ on y as its domain. It is possible that these constraints contradict, meaning the range of $R1$ is basically disjoint from that of $R2$, and thus no possible composition relation exists, and so on.

Similar to Proposition 9.2.1, the *outcome of this modeling task would add further value to FORT within the applied ontology field*. It will expand its capabilities by introducing additional axioms that enable implicit reasoning about the relations existing between entities. This advance-

1. https://en.wikipedia.org/wiki/Region_connection_calculus

R1\R2	SED(y,z)	Component-of(y,z)	Element-of(y,z)	Located-at(y,z)	Entity-located(y,z)	Member-of(y,z)	Constitutes(y,z)
SED(y,z)	T						
Component-of(y,z)		T					
Element-of(y,z)			T				
Located-at(y,z)				$\neg T$			
Entity-located(y,z)					T		
Member-of(y,z)						$\neg T$	
Constitutes(y,z)							T

TABLEAU 9.1 – *The composition table of FORT's relations.*

ment is particularly significant when explicit knowledge is available regarding the relations shared by another common entity between them.

9.2.3 A (semi-)automatic decision procedure for *directly* importing FORT

In order to employ FORT as a robust and expressive language for representing relations, we have discussed several employment scenarios in Section 8.3 of Chapter 8. And based on the direct employment scenario, in which an ontology aims at importing FORT as a expressive language of relations, one perspective of our work is to **design and develop a tool to support the direct employment method for FORT within domain ontologies**.

Indeed, it would also be convenient for FORT to be easily imported into any domain ontology to support the representation and reasoning of relations. this could be the case of OWL2DL formalized ontologies whose modelers/user may not possess expertise in foundational ontological relations. To facilitate the integration of FORT into user-based models, a decision procedure can be developed to assist in selecting a relevant FORT relation between the categories of the user's ontology. This approach is similar to the work presented in [Morales-González2015]. The objective would be to simplify the process of choosing the appropriate relation by abstracting the complexity of logic formalization and presenting it as a series of steps in natural language.

From a modeling perspective, the decision procedure can leverage the semantics of relations, specifically the constraints on the domain and range of relationships, to determine the most suitable relation based on these characteristics. This can be achieved through the use of a tree diagram, as exemplified in Figure 9.1. At each node, a binary question is formulated to address specific attributes of the relations. Based on the output, a YES or NO answer, the procedure progresses to the next node until a relationship is selected, if applicable. Given that users are generally more familiar with the semantics of their own categories than those of FORT's relations, the questions can be designed to pertain to the user's categories as the intended domain and range of the relation to be selected, considering the relation's axioms.

From an implementation standpoint, the tool can be realized as a user-friendly interface application using a programming language such as C# with the Jena ontology API. The interface would initially ask the user to load the OWL ontology file. Subsequently, the user would select the categories between which a relationship is to be employed. The tool would then present the corresponding questions based on the designed decision procedure, which could be accompanied by examples to facilitate comprehension. Based on the user's answers, a relationship from FORT's would be eventually suggested. Finally, the list of relationships to be imported, along with their associated domain and range categories in the user's ontology, would be saved in a new ontology file.

This proposition encompasses both *a modelisation task and an implementation task, which*

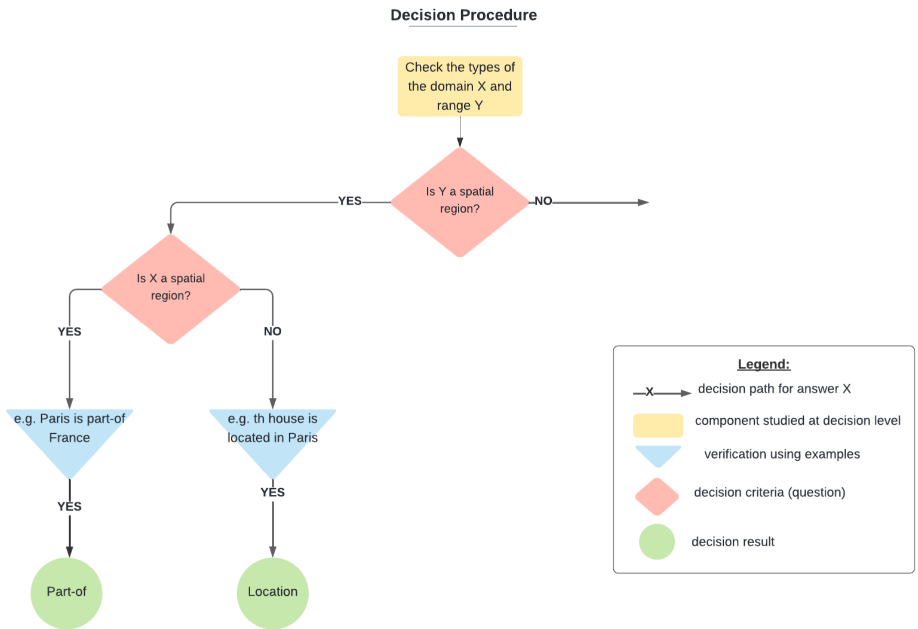


FIGURE 9.1 – *A tree diagram example for designing the decision procedure for selecting FORT relations.*

together contribute to an employment method of FORT into domain ontologies, which in turn enhances representation and reasoning tasks in OWL2 ontologies.

9.2.4 An ontology architecture for a cultural heritage application using FORT

A crucial perspective for this thesis is **to design and implement an ontology architecture for cultural heritage, employing FORT and CIDOC CRM, within a GaV interdisciplinary integration approach**. This objective can be achieved through the use of FORT's indirect employment method, as discussed in Section 8.3 of Chapter 8.

Indeed, FORT provides an expressive means of representing the structural and spatial characteristics of tangible cultural heritage. Let's refer to these as "compositional characteristics" of cultural heritage entities, which can be represented using FORT's "compositional relations". However, it is important to consider other aspects of tangible cultural heritage entities that hold equal significance, such as its symbol, value, meaning, color, creator, population, dating, etc. Let's refer to these as "descriptive characteristics", which are beyond the scope of FORT's coverage. These descriptive characteristics can be represented using "description relations" and others. Thus, (1) there is need to articulate FORT with an additional ontology model tat could support such characteristics for a cultural heritage application.

Furthermore, to achieve interdisciplinary integration, it is necessary to establish agreement between different models regarding the integration of their respective **categories**. This entails the adoption of a shared hierarchy of categories within this interdisciplinary approach, forming the basis for semantic agreement among the various disciplines involved. Thus, (2) an additional ontology is required which can serve as a core ontology in the cultural heritage field, recalling that a core ontology defines multiple domains (e.g. field within cultural heritage) and consist of the minimal generic concepts required to understand domain concepts across these domains.

Thus, based on the aforementioned considerations (1) and (2), there is a need for a core/mid-level ontology in the cultural heritage domain that can garner consensus across different fields beside FORT. The CIDOC CRM serves as an exemplary case in the cultural heritage community, widely accepted and used. It offers a comprehensive range of representations, including the required descriptive representations mentioned earlier, as well as representations of spatiotemporal entities and evolutionary characteristics, as discussed in Chapter 2.

Combining FORT and CIDOC CRM within a GaV approach for cultural heritage results in a synergistic integration of FORT's semantically rich compositional representations with the comprehensive and widely accepted cultural heritage ontology, as depicted in Figure 9.2.

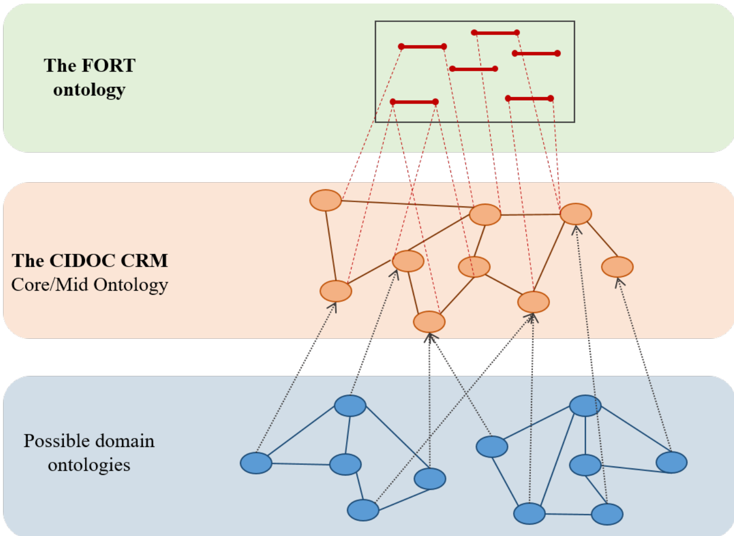


FIGURE 9.2 – *An ontology architecture for cultural heritage based on the GaV paradigm for interdisciplinary integration. This architecture uses FORT (via the indirect employment method) at the global level, and the CIDOC CRM at an additional core/mid-level.*

To implement this architecture, two main tasks need to be undertaken, both centered around ontology mapping. The first task involves mapping between FORT and CIDOC CRM, which would primarily focus on object properties. The second task, dependent on the number and domains of ontologies at the domain level, pertains to mapping these domain-specific ontologies to CIDOC CRM.

In cases where users employ relational databases rather than ontologies, the process may involve transforming the databases into RDF graphs compliant with CIDOC CRM. Alternatively, existing domain ontologies in the cultural heritage field, such as the CHARM model discussed in Chapter 2, can be used too. For instance, CHARM provides comprehensive representations for archaeological entities at a generic level and proves relevant for the archaeological field.

This perspective necessitates a deep understanding of both cultural heritage and ontology modeling. If implemented, it would make *a significant contribution to the cultural heritage field by providing a comprehensive ontology architecture*. This architecture incorporates the widely recognized and extensively used CIDOC CRM, while enhancing it with semantically rich relations offered by FORT. The resulting *entity-centric approach* for cultural heritage, which we advocated for in Chapter 2, would encompass both compositional and descriptive representations, as well as time constraints, offering thus a comprehensive framework.

9.2.5 An empirical validation of the translation procedure

We have developed in Chapter 7 a design of a very basic procedure described in natural language for translating FOL theories into SROIQ Tboxes. However, it is acknowledged that further research and development are necessary to refine its study at lower granularity. To address this, a possible future work is **to advance the procedure by formalizing it as an algorithm and conducting empirical validation to investigate its properties**.

The proposal raises an intriguing idea, as it offers practical implications for interpreting domain knowledge within logical theories using a relatively expressive yet efficient logic framework. Nevertheless, there remains a need to establish a clearer understanding of the properties that can be expected from such a procedure including termination, soundness, and maximality.

These properties could be studied at a theoretical perspective first, dealing with the procedure P as a deductive system deriving a set of axioms $A = \{a_i\} \forall 0 \leq i \leq n, n \in \mathbb{N}$ from a formula A_0. Taking into account the notions of truth "$\models$" and provability "$\vdash$", soundness and termination could be studied.

For the former, P can be assessed for soundness by examining if, $A_0 \vdash_P A_6 \implies A_0 \models A_6$. This means checking that if A_6 is derived from A_0 by P, then A_6 is true in all models of A_0.

For the latter, P can be assessed if complete by studying if $A_0 \models A_6 \implies A_0 \vdash_P A_6$. This means checking that if A_6 is true in all models of A_0, then A_6 is derived from A_0 by P.

As for maximality, it could be tested by considering the final step as a SAT problem using a SAT solver to find all the possible subsets of the structured axioms and proving that the output of P yields in the maximal structured set, following the work in [Botti Benevides2019].

In practical terms, additional validation can be achieved by implementing the individual steps of the procedure as an algorithm per se. This implementation would enable testing using case studies such foundational ontologies and other meta-ontologies that have already been formalized in FOL, as well as supplementary OWL2DL ontologies (i.e. the reference of inputs and outputs of these case studies are available).

The result of this *computational task would provide a valuable contribution by serving as a practical tool for bridging two knowledge representation languages : FOL and SRROIQ*. This direction of bridging, specifically from FOL to SROIQ, is particularly important as it facilitates the integration of logical theories into the semantic web framework. This contribution would be highly valued within the semantic web communities, as it addresses the theoretical foundations of semantic web technologies, such as the OWL2DL web ontology language.

These propositions provide valuable avenues for future exploration and development, encompassing both theoretical and practical aspects of our approach.

Bibliographie

[2012019] *An Overview of the Common Core Ontologies*. 2019.

[Abiteboul1998] Serge Abiteboul et Oliver M Duschka. *Complexity of answering queries using materialized views*. In Proceedings of the seventeenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, pages 254–263, 1998.

[Akkerman2011] Sanne F Akkerman et Arthur Bakker. *Boundary crossing and boundary objects*. Review of educational research, vol. 81, no. 2, pages 132–169, 2011.

[Alexiev2015] Vladimir Alexiev, Joan Cobb, Gregg Garcia et Patricia Harpring. Getty vocabularies linked open data : Semantic representation. 03 2015.

[Allen1983] James F Allen. *Maintaining knowledge about temporal intervals*. Communications of the ACM, vol. 26, no. 11, pages 832–843, 1983.

[Allen1984] James F Allen. *Towards a general theory of action and time*. Artificial intelligence, vol. 23, no. 2, pages 123–154, 1984.

[Aloia2011] Nicola Aloia, Cesare Concordia et Carlo Meghini. *Europeana v1. 0*. In Digital Libraries and Archives : 7th Italian Research Conference, IRCDL 2011, Pisa, Italy, January 20-21, 2011. Revised Papers 7, pages 127–129. Springer, 2011.

[Arano2005] Silvia Arano. *Thesauruses and ontologies*. Hipertext. net, vol. 3, page 24, 2005.

[Armstrong2018] David Malet Armstrong. Universals : An opinionated introduction. Routledge, 2018.

[Arp2015] Robert Arp, Barry Smith et Andrew D Spear. Building ontologies with basic formal ontology. Mit Press, 2015.

[Artale1996a] Alessandro Artale, Enrico Franconi et Nicola Guarino. *Open problems for part-whole relations*. In Description Logics, pages 70–73, 1996.

[Artale1996b] Alessandro Artale, Enrico Franconi, Nicola Guarino et Luca Pazzi. *Part-whole relations in object-centered systems : An overview*. Data & Knowledge Engineering, vol. 20, no. 3, pages 347–383, 1996.

[Artale2002] Alessandro Artale, Enrico Franconi, Frank Wolter et Michael Zakharyaschev. *A temporal description logic for reasoning over conceptual schemas and queries*. In European Workshop on Logics in Artificial Intelligence, pages 98–110. Springer, 2002.

[Artale2007] Alessandro Artale, Christine Parent et Stefano Spaccapietra. *Evolving objects in temporal information systems*. Annals of Mathematics and Artificial Intelligence, vol. 50, no. 1-2, pages 5–38, 2007.

[Artale2008] Alessandro Artale, Nicola Guarino et C Maria Keet. *Formalising temporal constraints on part-whole relations.* In KR, pages 673–683, 2008.

[ARTstor2006] J Paul Getty Trust ARTstor. *CDWA Lite : Specification for an XML Schema for Contributing Records via the OAI Harvesting Protocol*. Retrieved January, vol. 2, page 2019, 2006.

[Astesiano2002] Egidio Astesiano, Michel Bidoit, Hélene Kirchner, Bernd Krieg-Brückner, Peter D Mosses, Donald Sannella et Andrzej Tarlecki. *CASL : the common algebraic specification language*. Theoretical Computer Science, vol. 286, no. 2, pages 153–196, 2002.

[Baader2003] Franz Baader, Diego Calvanese, Deborah McGuinness, Peter Patel-Schneider, Daniele Nardi*et al.* The description logic handbook : Theory, implementation and applications. Cambridge university press, 2003.

[Baker2000] Lynne Rudder Baker*et al.* Persons and bodies : A constitution view. Cambridge University Press, 2000.

[Barbier2003] Franck Barbier, Brian Henderson-Sellers, Annig Le Parc-Lacayrelle et J-M Bruel. *Formalization of the whole-part relationship in the unified modeling language*. IEEE Transactions on Software Engineering, vol. 29, no. 5, pages 459–470, 2003.

[Bateman1995] John A Bateman, Renate Henschel et Fabio Rinaldi. *The generalized upper model 2.0*. In Proceedings of the ECAI94 Workshop : Comparison of Implemented Ontologies, 1995.

[Bateman2010] John A Bateman, Joana Hois, Robert Ross et Thora Tenbrink. *A linguistic ontology of space for natural language processing*. Artificial Intelligence, vol. 174, no. 14, pages 1027–1071, 2010.

[Bechhofer2004] Sean Bechhofer, Frank Van Harmelen, Jim Hendler, Ian Horrocks, Deborah L McGuinness, Peter F Patel-Schneider, Lynn Andrea Stein*et al.* *OWL web ontology language reference*. W3C recommendation, vol. 10, no. 2, pages 1–53, 2004.

[Berardi2005] Daniela Berardi, Diego Calvanese et Giuseppe De Giacomo. *Reasoning on UML class diagrams*. Artificial intelligence, vol. 168, no. 1-2, pages 70–118, 2005.

[Bergamaschi2001] Sonia Bergamaschi, Silvana Castano, Maurizio Vincini et Domenico Beneventano. *Semantic integration of heterogeneous information sources*. Data & Knowledge Engineering, vol. 36, no. 3, pages 215–249, 2001.

[Berghammer2010] Rudolf Berghammer et Georg Struth. *On automated program construction and verification*. In Mathematics of Program Construction : 10th International Conference, MPC 2010, Québec City, Canada, June 21-23, 2010. Proceedings 10, pages 22–41. Springer, 2010.

[Berners-Lee2001] Tim Berners-Lee, James Hendler et Ora Lassila. *The semantic web*. Scientific american, vol. 284, no. 5, pages 34–43, 2001.

[Bidoit2004] Michel Bidoit et Peter D Mosses. *CASL User Manual, volume 2900 of LNCS*, 2004.

[Binding2010a] Ceri Binding. *Implementing archaeological time periods using CIDOC CRM and SKOS*. In The Semantic Web : Research and Applications : 7th Extended Semantic Web Conference, ESWC 2010, Heraklion, Crete, Greece, May 30–June 3, 2010, Proceedings, Part I 7, pages 273–287. Springer, 2010.

[Binding2010b] Ceri Binding, Keith May, Renato Souza, Douglas Tudhope et Andreas Vlachidis. *Semantic technologies for archaeology resources : results from the STAR project*. In Th Annual Conference on Computer Applications and Quantitative Methods in Archaeology (CAA2010), Granada : Archaeopress, pages 555–561. CAA, 2010.

[Bischof2014] Stefan Bischof, Markus Krötzsch, Axel Polleres et Sebastian Rudolph. *Schema-agnostic query rewriting in SPARQL 1.1*. In The Semantic Web–ISWC 2014 : 13th International Semantic Web Conference, Riva del Garda, Italy, October 19-23, 2014. Proceedings, Part I 13, pages 584–600. Springer, 2014.

[Bittner2001] Thomas Bittner et Barry Smith. *A taxonomy of granular partitions*. In International Conference on Spatial Information Theory, pages 28–43. Springer, 2001.

[Bittner2003] Thomas Bittner et Barry Smith. *A theory of granular partitions*. Foundations of geographic information science, vol. 7, pages 124–125, 2003.

[Bittner2004a] Thomas Bittner. *Axioms for parthood and containment relations in bio-ontologies*. In KR-MED, pages 4–11, 2004.

[Bittner2004b] Thomas Bittner, Maureen Donnelly et Barry Smith. *Endurants and perdurants in directly depicting ontologies*. AI Communications, vol. 17, no. 4, pages 247–258, 2004.

[Bittner2004c] Thomas Bittner, Maureen Donnelly et Barry Smith. *On the Foundational Relations of Ontology*. In Formal Ontology in Information Systems : Proceedings of the Third International Conference (FOIS-2004), page 37. IOS Press, 2004.

[Bittner2005] Thomas Bittner et Maureen Donnelly. *Computational ontologies of parthood, componenthood, and containment*. In International Joint Conference on Artificial Intelligence, volume 19, page 382. LAWRENCE ERLBAUM ASSOCIATES LTD, 2005.

[Bloom1996] Paul Bloom. *Intention, history, and artifact concepts*. Cognition, vol. 60, no. 1, pages 1–29, 1996.

[BOOK2002] BLUE BOOK. *Reference Model for an Open Archival Information System (OAIS)*. 2002.

[Bordet2021] Pierre Bordet, Florian Kergourlay, Ariane Pinto, Nils Blanc et Pauline Martinetto. *Applying multivariate analysis to X-ray diffraction computed tomography : the study of medieval applied brocades*. Journal of Analytical Atomic Spectrometry, vol. 36, no. 8, pages 1724–1734, 2021.

[Borgo2008] Stefano Borgo et Leonardo Lesmo. *The attractiveness of foundational ontologies in industry*. Frontiers in Artificial Intelligence and Applications, vol. 174, page 1, 2008.

[Borgo2022a] Stefano Borgo, Roberta Ferrario, Aldo Gangemi, Nicola Guarino, Claudio Masolo, Daniele Porello, Emilio M Sanfilippo et Laure

Vieu. *DOLCE : A descriptive ontology for linguistic and cognitive engineering*. Applied ontology, vol. 17, no. 1, pages 45–69, 2022.

[Borgo2022b] Stefano Borgo, Antony Galton et Oliver Kutz. *Foundational ontologies in action*. Applied ontology, no. Preprint, pages 1–16, 2022.

[Borst1997] Pim Borst, Hans Akkermans et Jan Top. *Engineering ontologies*. International journal of human-computer studies, vol. 46, no. 2-3, pages 365–406, 1997.

[Borst1999] Willem Nico Borst. *Construction of engineering ontologies for knowledge sharing and reuse*. 1999.

[Botti Benevides2019] Alessander Botti Benevides, Jean-Rémi Bourguet, Giancarlo Guizzardi, Rafael Penaloza et João Paulo A Almeida. *Representing a reference foundational ontology of events in SROIQ*. Applied Ontology, vol. 14, no. 3, pages 293–334, 2019.

[Bouchenaki2003] Mounir Bouchenaki. *The interdependency of the tangible and intangible cultural heritage*. 2003.

[Bower1994] James M Bower, Murtha Baca, Vocabulary Coordination Group*et al.* *Union list of artist names*. 1994.

[Bowker2000] Geoffrey C Bowker et Susan Leigh Star. Sorting things out : Classification and its consequences. MIT press, 2000.

[Brachman1978] Ronald J Brachman. *A Structural Paradigm for Representing Knowledge*. Rapport technique, BOLT BERANEK AND NEWMAN INC CAMBRIDGE MASS, 1978.

[Brachman1983] Ronald J. Brachman. *What IS-A is and isn't : An analysis of taxonomic links in semantic networks*. Computer, vol. 16, no. 10, pages 30–36, 1983.

[Brachman1989] Ronald J Brachman et James G Schmolze. *An overview of the KL-ONE knowledge representation system*. Readings in artificial intelligence and databases, pages 207–230, 1989.

[Brickhouse2016] David Brickhouse. *Tin ImitatingTextile : A Study of the Applied Brocade Technique Using Technical Reconstructions*. Thèse de Doctorat, University of Delaware, 2016.

[Brickley2014] Dan Brickley et Ramanathan Guha. *RDF Schema 1.1*. W3C recommendation, W3C, February 2014. https ://www.w3.org/TR/2014/REC-rdf-schema-20140225/.

[Brogan2005] Martha L Brogan. A survey of digital library aggregation services, 2003. Digital Library Federation, 2005.

[Bruseker2017] George Bruseker, Nicola Carboni et Anaïs Guillem. *Cultural heritage data management : the role of formal ontology and CIDOC CRM*. Heritage and archaeology in the digital age : acquisition, curation, and dissemination of spatial cultural heritage data, pages 93–131, 2017.

[Buccell2005] Agustina Buccell, Alejandra Cechich et Nieves R Brisaboa. *Ontology-based data integration methods : A framework for comparison*. Revista Colombiana de Computacion, vol. 6, no. 1, pages 1–24, 2005.

[Bunge1977] Mario Bunge. Treatise on basic philosophy : Ontology i : the furniture of the world, volume 3. Springer Science & Business Media, 1977.

[Burkhardt1991] Hans Burkhardt et Carlos A. Dufour. *Part/Whole I : History*. In Hans Burkhardt et Barry Smith, editeurs, Handbook of Metaphysics and Ontology, pages 2–663. Philosophia Verlag, 1991.

[Buron2019] Maxime Buron, François Goasdoué, Ioana Manolescu et Marie-Laure Mugnier. *Reformulation-based query answering for RDF graphs with RDFS ontologies*. In The Semantic Web : 16th International Conference, ESWC 2019, Portorož, Slovenia, June 2–6, 2019, Proceedings 16, pages 19–35. Springer, 2019.

[Buron2020] Maxime Buron, François Goasdoué, Ioana Manolescu, Tayeb Merabti et Marie-Laure Mugnier. *Revisiting RDF storage layouts for efficient query answering*. In SSWS 2020-13th International Workshop on Scalable Semantic Web Knowledge Base Systems. inria saclay, 2020.

[Caccamo2022] Marta Caccamo, Daniel Pittino et Fredrik Tell. *Boundary objects, knowledge integration, and innovation management : A systematic review of the literature*. Technovation, page 102645, 2022.

[Calì2001] Andrea Calì, Diego Calvanese, Giuseppe De Giacomo et Maurizio Lenzerini. *Accessing data integration systems through conceptual schemas*. In Conceptual Modeling—ER 2001 : 20th International Conference on Conceptual Modeling Yokohama, Japan, November 27–30, 2001 Proceedings 20, pages 270–284. Springer, 2001.

[Calvanese2000] Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini et Moshe Y Vardi. *Answering regular path queries using views*. In Proceedings of 16th International Conference on Data Engineering (Cat. No. 00CB37073), pages 389–398. IEEE, 2000.

[Calvanese2001] Diego Calvanese, Giuseppe De Giacomo et Maurizio Lenzerini. *A Framework for Ontology Integration*. In The emerging semantic web, pages 201–214. Citeseer, 2001.

[Calvanese2003] Diego Calvanese et Giuseppe De Giacomo. *Expressive description logics*. In The description logic handbook : theory, implementation, and applications, pages 178–218. 2003.

[Candela2007] Leonardo Candela, Donatella Castelli, Nicola Ferro, Yannis Ioannidis, Georgia Koutrika, Carlo Meghini, Pasquale Pagano, Seamus Ross, Dagobert Soergel, Maristella Agosti*et al. The DELOS digital library reference model. Foundations for digital libraries*. 2007.

[Casati1997] Roberto Casati et Achille C Varzi. *50 Years of Events : An Annotated Bibliography, 1947 to 1997*. 1997.

[Casati1998] Roberto Casati, Barry Smith et Achille Varzi. *Ontological Tools for Geographic Representation*. In N. Guarino, editeur, Formal Ontology in Information Systems, pages 77–85. Amsterdam : IOS Press, 1998.

[Casati1999] Roberto Casati et Achille C Varzi. Parts and places : The structures of spatial representation. MIT press, 1999.

[Ceusters2015] Werner Ceusters et Barry Smith. *Aboutness : Towards foundations for the information artifact ontology*. 2015.

[Chalmin2019] Emilie Chalmin et Jillian Huntley. *Characterizing Rock Art Pigments*. In The Oxford Handbook of the Archaeology and Anthropology of Rock Art. Oxford University Press, 01 2019.

[Charles2013] Valentine Charles, Antoine Isaac, Vassilis Tzouvaras et Steffen Hennicke. *Mapping cross-domain metadata to the Europeana data model (EDM)*. In Research and Advanced Technology for Digital Libraries : International Conference on Theory and Practice of Digital Libraries, TPDL 2013, Valletta, Malta, September 22-26, 2013. Proceedings 3, pages 484–485. Springer, 2013.

[Chen1976] Peter Pin-Shan Chen. *The entity-relationship model—toward a unified view of data*. ACM transactions on database systems (TODS), vol. 1, no. 1, pages 9–36, 1976.

[Chisholm1975] Roderick M Chisholm. *Mereological Essentialism : Some Further Considerations*. The Review of Metaphysics, pages 477–484, 1975.

[Chisholm1996] Roderick M Chisholm. A realistic theory of categories : An essay on ontology. Cambridge University Press, 1996.

[Coburn2010] Erin Coburn, Richard Light, Gordon McKenna, Regine Stein et Axel Vitzthum. *LIDO-lightweight information describing objects version 1.0*. ICOM International Committee of Museums, 2010.

[Cohen2021] S. Marc Cohen et C. D. C. Reeve. *Aristotle's Metaphysics*. In Edward N. Zalta, editeur, The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Winter 2021 edition, 2021.

[Cruse1979] D Alan Cruse. *On the transitivity of the part-whole relation1*. Journal of linguistics, vol. 15, no. 1, pages 29–38, 1979.

[Cruse1986] David Alan Cruse. Lexical semantics. Cambridge university press, 1986.

[Cruz2005] Isabel F Cruz, Huiyong Xiao*et al. The role of ontologies in data integration*. Engineering intelligent systems for electrical engineering and communications, vol. 13, no. 4, page 245, 2005.

[Cui1993] Zhan Cui, Anthony G Cohn et David A Randell. *Qualitative and topological relationships in spatial databases*. In Advances in Spatial Databases : Third International Symposium, SSD'93 Singapore, June 23–25, 1993 Proceedings 3, pages 296–315. Springer, 1993.

[Cummings2017] Joel Cummings et Deborah Stacey. *The mid level data collection ontology (dco)*. Proceedings of KEOD, 2017.

[Cyganiak2014] Richard Cyganiak, David Wood et Markus Lanthaler. *RDF 1.1 Concepts and Abstract Syntax*. W3C recommendation, W3C, February 2014. https ://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/.

[Dancause2018] Renée Dancause, Janet Wagner et Jan Vuori. Caring for textiles and costumes. Canadian Heritage= Patrimoine canadien, 2018.

[David2002] Bruno David. Landscapes, rock-art and the dreaming : an archaeology of preunderstanding. London, 2002.

[David2018] Bruno David et Ian J McNiven. The oxford handbook of the archaeology and anthropology of rock art. Oxford University Press, 2018.

[De Bruijn2006] Jos De Bruijn, Marc Ehrig, Cristina Feier, Francisco Martín-Recuerda, François Scharffe et Moritz Weiten. *Ontology mediation, merging and aligning*. Semantic web technologies, pages 95–113, 2006.

[De Giacomo2018] Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Antonella Poggi et Riccardo Rosati. *Using ontologies for semantic data integration*. A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years, pages 187–202, 2018.

[de Keizer2000] Nicolette F de Keizer, Ameen Abu-Hanna et JHM Zwetsloot-Schonk. *Understanding terminological systems I : terminology and typology*. Methods of information in medicine, vol. 39, no. 01, pages 16–21, 2000.

[de Kergommeaux2021] Aurélie Chassin de Kergommeaux, Hélène Salomon, Julien Monney, Claire Chanteraud, Pradeau Jean-Victor, Éric Goemaere, Thierry Leduc et Emilie Chalmin. *Référencement géologique des ressources en matières colorantes entre l'Ardèche et le Gardon. Un outil pour appréhender les paysages vécus au cours du Paléolithique supérieur*. Ardèche archéologie, no. 38, pages 14–23, 2021.

[De Kleer1984] Johan De Kleer. *How circuits work*. Artificial intelligence, vol. 24, no. 1-3, pages 205–280, 1984.

[Decker1999] Stefan Decker, Michael Erdmann, Dieter Fensel et Rudi Studer. Ontobroker : Ontology based access to distributed and semi-structured information. Springer, 1999.

[Defrasne2019] Claudia Defrasne, Emilie Chalmin, Ludovic Bellot-Gurlet, Eric Thirault et Guy André. *From archeological layers to schematic rock art ? Integrated study of the Neolithic pigments and pigmented rocks at the Rocher du Château (Western Alps, Savoie, France)*. Archaeological and Anthropological Sciences, vol. 11, pages 6065–6091, 2019.

[Defrasne2023] Claudia Defrasne. *The Cognitive Processus Behind Neolithic Schematic Rock Art. Archaeological Implications and Research Hypothesis*. Cambridge Archaeological Journal, vol. 33, no. 1, pages 157–179, 2023.

[Degen2001] Wolfgang Degen, Barbara Heller, Heinrich Herre et Barry Smith. *GOL : toward an axiomatized upper-level ontology*. In Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001, pages 34–46, 2001.

[Dipert1993] Randall R. Dipert. Artifacts, art works, and agency. Philadelphia : Temple University Press, 1993.

[Doan2012] AnHai Doan, Alon Halevy et Zachary Ives. Principles of data integration. Elsevier, 2012.

[Doerr2000] Martin Doerr. *Mapping of the Dublin Core metadata element set to the CIDOC CRM*. ICS/FORTH, Techincal Report, vol. 274, 2000.

[Doerr2002] M Doerr et A Sarris. *The Digital Heritage of Archaeology*. In Proc. CAA, 2002.

[Doerr2003a] Martin Doerr. *The CIDOC conceptual reference module : an ontological approach to semantic interoperability of metadata*. AI magazine, vol. 24, no. 3, pages 75–75, 2003.

[Doerr2003b] Martin Doerr, Jane Hunter et Carl Lagoze. *Towards a core ontology for information integration*. 2003.

[Doerr2009] Martin Doerr. Ontologies for cultural heritage, pages 463–486. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[Doerr2010] Martin Doerr, Stefan Gradmann, Steffen Hennicke, Antoine Isaac, Carlo Meghini et Herbert Van de Sompel. *The europeana data model (edm)*. In World Library and Information Congress : 76th IFLA general conference and assembly, volume 10, page 15, 2010.

[Doerr2013] Martin Doerr, Gerald Hiebel et Øyvind Eide. *CRMgeo : Linking the CIDOC CRM to GeoSPARQL through a spatiotemporal refinement*. Rapport technique, TECHNICAL REPORT : ICS-FORTH/TR-435, 2013.

[Doerr2018] Martin Doerr, Achille Felicetti, Sorin Hermon, Gerald Hiebel, Athina Kritsotaki, Anja Masur, Keith May, Paola Ronzino, Wolfgang Schmidle, Maria Theodoridou*et al.* *Definition of the CRMarchaeo*. An Extension of CIDOC CRM to Support the Archaeological Excavation Process. Version-1.5. 0 http ://www. cidoc-crm. org/crmarchaeo/ModelVersion/version-1.5. 0, 2018.

[Donini1998] Francesco M Donini, Maurizio Lenzerini, Daniele Nardi et Andrea Schaerf. *AL-log : Integrating datalog and description logics*. Journal of Intelligent Information Systems, vol. 10, no. 3, pages 227–252, 1998.

[Donnelly2006] Maureen Donnelly, Thomas Bittner et Cornelius Rosse. *A formal theory for spatial representation and reasoning in biomedical ontologies*. Artificial Intelligence in Medicine, vol. 36, no. 1, pages 1–27, 2006.

[Dubinin2014] Victor Dubinin, Valeriy Vyatkin, Chen-Wei Yang et Cheng Pang. *Automatic generation of automation applications based on ontology transformations*. In Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA), pages 1–4. IEEE, 2014.

[Dudley2012] Sandra H Dudley. Museum objects : Experiencing the properties of things. Routledge, 2012.

[Edmund1970] Husserl Edmund. *Logical investigations*, 1970.

[Eiter1997] Thomas Eiter, Georg Gottlob et Heikki Mannila. *Disjunctive datalog*. ACM Transactions on Database Systems (TODS), vol. 22, no. 3, pages 364–418, 1997.

[Ekaputra2016] Fajar J Ekaputra, Marta Sabou, Estefanía Serral et Stefan Biffl. *Knowledge change management and analysis during the engineering of cyber physical production systems : A use case of hydro power plants*. In Proceedings of the 12th International Conference on Semantic Systems, pages 105–112, 2016.

[Ekaputra2017] Fajar Ekaputra, Marta Sabou, Estefanía Serral Asensio, Elmar Kiesling et Stefan Biffl. *Ontology-based data integration in multi-disciplinary engineering environments : A review*. Open Journal of Information Systems, vol. 4, no. 1, pages 1–26, 2017.

[Emeruem2022] Chiadika Emeruem, C Maria Keet, Zubeida C Dawood et Steve Wang. *BFO Classifier : Aligning domain ontologies to BFO*. 2022.

[Euzenat2007] Jérôme Euzenat, Pavel Shvaiko*et al.* Ontology matching, volume 18. Springer, 2007.

[Falquet2011] Gilles Falquet, Claudine Métral, Jacques Teller, Christopher Tweed, Catherine Roussey, Francois Pinet, Myoung Ah Kang et Oscar Corcho. *An introduction to ontologies and ontology engineering*. Ontologies in Urban development projects, pages 9–38, 2011.

[Fernández-López1997] Mariano Fernández-López, Asunción Gómez-Pérez et Natalia Juristo. *Methontology : from ontological art towards ontological engineering*. 1997.

[Fernández-López1999] Mariano Fernández-López. *Overview of methodologies for building ontologies*. 1999.

[Fernández-López2008] Mariano Fernández-López, Asunción Gómez-Pérez et Mari Carmen Suárez-Figueroa. *Selecting and customizing a mereology ontology for its reuse in a pharmaceutical product ontology*. 2008.

[Fernández-López2002] Mariano Fernández-López, Asuncion Gomez-Perez, Jerome Euzenat, Aldo Gangemi, Y Kalfoglou, Domenico Pisanelli, M Schorlemmer, G Steve, Ljiljana Stojanovic, Gerd Stumme et York Sure-Vetter. *Ontoweb : A survey on methodologies for developing, maintaining, integrating, evaluating and reengineering ontologies*, 01 2002.

[Fine1983] K Fine et B Smith. *Husserl (Part One) : The Pure Theory*. mimeograph, Manchester, 1983.

[Fine1999] Kit Fine. *Things and their parts*. Midwest Studies in Philosophy, vol. 23, pages 61–74, 1999.

[Fonseca2019] Claudenir M Fonseca, Daniele Porello, Giancarlo Guizzardi, João Paulo A Almeida et Nicola Guarino. *Relations in ontology-driven conceptual modeling*. In Conceptual Modeling : 38th International Conference, ER 2019, Salvador, Brazil, November 4–7, 2019, Proceedings 38, pages 28–42. Springer, 2019.

[Force2010] UML Revision Task Force. *OMG unified modeling language : Superstructure*. Object Management Group (OMG), 2010.

[Gagnon2007] Michel Gagnon. *Ontology-based integration of data sources*. In 2007 10th International Conference on Information Fusion, pages 1–8. IEEE, 2007.

[Galton2010] Antony Galton. *How is a collection related to its members*. Interdisciplinary ontology, vol. 3, pages 9–17, 2010.

[Gangemi1998] A Gangemi, D Pisanelli et G Steve. *Ontology alignment : experiences with medical terminologies*. 1998.

[Gangemi2001] Aldo Gangemi, Nicola Guarino, Claudio Masolo et Alessandro Oltramari. *Understanding top-level ontological distinctions*. In In Proceedings of IJCAI-01 Workshop on Ontologies and Information Sharing (OIS)@ IJCAI, 2001.

[Gangemi2002] Aldo Gangemi, Nicola Guarino, Claudio Masolo, Alessandro Oltramari et Luc Schneider. *Sweetening ontologies with DOLCE*. In Knowledge Engineering and Knowledge Management : Ontologies and the Semantic Web : 13th International Conference, EKAW 2002 Sigüenza, Spain, October 1–4, 2002 Proceedings 13, pages 166–181. Springer, 2002.

[Garbacz2007] Pawel Garbacz. *A first order theory of functional parthood*. Journal of philosophical logic, vol. 36, pages 309–337, 2007.

[Gasse2008] Francis Gasse, Ulrike Sattler et Volker Haarslev. *Rewriting Rules into SROIQ Axioms*. Description Logics, vol. 353, 2008.

[Gell1998] Alfred Gell. Art and agency : an anthropological theory. Clarendon Press, 1998.

[Gergatsoulis2010] Manolis Gergatsoulis, Lina Bountouri, Panorea Gaitanou et Christos Papatheodorou. *Mapping cultural metadata schemas to CIDOC conceptual reference model*. In Artificial Intelligence : Theories, Models and Applications : 6th Hellenic Conference on AI, SETN 2010, Athens, Greece, May 4-7, 2010. Proceedings 6, pages 321–326. Springer, 2010.

[Gerstl1995] Peter Gerstl et Simone Pribbenow. *Midwinters, end games, and body parts : a classification of part-whole relations*. International journal of human-computer studies, vol. 43, no. 5-6, pages 865–889, 1995.

[Gerstl1996] Peter Gerstl et Simone Pribbenow. *A conceptual theory of part-whole relations and its applications*. Data & Knowledge Engineering, vol. 20, no. 3, pages 305–322, 1996.

[Gilchrist2003] Alan Gilchrist. *Thesauri, taxonomies and ontologies–an etymological note*. Journal of documentation, 2003.

[Gilmore2018] Cody Gilmore. *Location and Mereology*. In Edward N. Zalta, editeur, The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall 2018 edition, 2018.

[Goh1997] Cheng Hian Goh. *Representing and reasoning about semantic conflicts in heterogeneous information systems*. Thèse de Doctorat, Massachusetts Institute of Technology, 1997.

[Gómez-Pérez2006] Asunción Gómez-Pérez, Mariano Fernández-López et Oscar Corcho. Ontological engineering : with examples from the areas of knowledge management, e-commerce and the semantic web. Springer Science & Business Media, 2006.

[Gómez-Pérez2019] Asunción Gómez-Pérez. *Knowledge sharing and reuse*. In The handbook of applied expert systems, pages 10–1. CRC Press, 2019.

[Gonzalez-Perez2012] Cesar Gonzalez-Perez. *A conceptual modelling language for the humanities and social sciences*. In 2012 Sixth International Conference on Research Challenges in Information Science (RCIS), pages 1–6. IEEE, 2012.

[Gonzalez-Perez2018] Cesar Gonzalez-Perez. *Information modelling for archaeology and anthropology*. Software Engineering Principles for Cultural Heritage, 2018.

[Graham2012] Brian Graham et Peter Howard. The ashgate research companion to heritage and identity. Ashgate Publishing, Ltd., 2012.

[Grant1994] Alice Grant et Jeff Cowton. Spectrum : The uk museum documentation standard. 1994.

[Grant1995] Alice Grant, Joséphine Nieuwenhuis et Toni Petersen. International guidelines for museum object information : the cidoc information categories. International Committee for Documentation of the International Council of . . . , 1995.

[Grimm2009] Stephan Grimm. *Knowledge representation and ontologies*. In Scientific data mining and knowledge discovery : principles and foundations, pages 111–137. Springer, 2009.

[Grosof2003] Benjamin N Grosof, Ian Horrocks, Raphael Volz et Stefan Decker. *Description logic programs : Combining logic programs with description logic*. In Proceedings of the 12th international conference on World Wide Web, pages 48–57, 2003.

[Group2009] W3C Owl Working Group*et al.* *OWL 2 web ontology language document overview*. http ://www. w3. org/TR/owl2-overview/, 2009.

[Gruber1993] Thomas R Gruber. *A translation approach to portable ontology specifications*. Knowledge acquisition, vol. 5, no. 2, pages 199–220, 1993.

[Gruber1995] Thomas R Gruber. *Toward principles for the design of ontologies used for knowledge sharing ?* International journal of human-computer studies, vol. 43, no. 5-6, pages 907–928, 1995.

[Gruninger1995] Michael Gruninger. *Methodology for the design and evaluation of ontologies*. In Proc. IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, 1995.

[Gruninger1996] Michael Gruninger et Mark S Fox. *The logic of enterprise modelling*. In Modelling and Methodologies for Enterprise Integration : Proceedings of the IFIP TC5 Working Conference on Models and Methodologies for Enterprise Integration, Queensland, Australia, November 1995, pages 140–157. Springer, 1996.

[Gruninger2002] Michael Gruninger. *Ontology : applications and design*. Commun. ACM, vol. 45, no. 2, 2002.

[Grüninger2009] Michael Grüninger. *COLORE : Common logic ontology repository*, 2009.

[Grüninger2012] Michael Grüninger, Torsten Hahmann, Ali Hashemi, Darren Ong et Atalay Ozgovde. *Modular first-order ontologies via repositories*. Applied Ontology, vol. 7, no. 2, pages 169–209, 2012.

[Grüninger2014] Michael Grüninger, Torsten Hahmann, Megan Katsumi et Carmen Chui. *A Sideways Look at Upper Ontologies.* In FOIS, pages 9–22, 2014.

[Grüninger2022] Michael Grüninger, Yi Ru et Jona Thai. *TUpper : A top level ontology within standards*. Applied Ontology, no. Preprint, pages 1–23, 2022.

[Guarino1995a] Nicola Guarino. *Formal ontology, conceptual analysis and knowledge representation*. International journal of human-computer studies, vol. 43, no. 5-6, pages 625–640, 1995.

[Guarino1995b] Nicola Guarino et Pierdaniele Giaretta. *Ontologies and knowledge bases*. Towards very large knowledge bases, pages 1–2, 1995.

[Guarino1997a] Nicola Guarino. *Semantic matching : Formal ontological distinctions for information organization, extraction, and integration*. In Information Extraction A Multidisciplinary Approach to an Emerging Information Technology : International Summer School, SCIE-97 Frascati, Italy, July 14–18, 1997, pages 139–170. Springer, 1997.

[Guarino1997b] Nicola Guarino. *Understanding, building and using ontologies*. International journal of human-computer studies, vol. 46, no. 2-3, pages 293–310, 1997.

[Guarino1998a] Nicola Guarino. Formal ontology in information systems : Proceedings of the first international conference (fois'98), june 6-8, trento, italy, volume 46. IOS press, 1998.

[Guarino1998b] Nicola Guarino. *Some ontological principles for designing upper level lexical resources*. arXiv preprint cmp-lg/9809002, 1998.

[Guarino2000a] Nicola Guarino et Christopher Welty. *A formal ontology of properties*. In International Conference on Knowledge Engineering and Knowledge Management, pages 97–112. Springer, 2000.

[Guarino2000b] Nicola Guarino, Christopher Welty*et al.* *Identity, unity, and individuality : Towards a formal toolkit for ontological analysis*. In ECAI, volume 2000, pages 219–223, 2000.

[Guarino2008] Nicola Guarino. *Ontologies and Ontological Analysis*. FOIS, 2008.

[Guarino2009a] Nicola Guarino. *The ontological level : Revisiting 30 years of knowledge representation*. Conceptual Modeling : Foundations and Applications : Essays in Honor of John Mylopoulos, pages 52–67, 2009.

[Guarino2009b] Nicola Guarino, Daniel Oberle et Steffen Staab. *What is an ontology ?* Handbook on ontologies, pages 1–17, 2009.

[Guarino2012] Nicola Guarino. *Introduction to applied ontology and ontological analysis*. 2012.

[Guarino2017] Nicola Guarino. *BFO and DOLCE : So far, so close....* Cosmos+ Taxis, vol. 4, no. 4, 2017.

[Guizzardi2005] Giancarlo Guizzardi. *Ontological foundations for structural conceptual models*. 2005.

[Guizzardi2006a] Giancarlo Guizzardi. *The role of foundational ontologies for conceptual modeling and domain ontology representation*. In 2006 7th International Baltic conference on databases and information systems, pages 17–25. IEEE, 2006.

[Guizzardi2006b] Giancarlo Guizzardi, Claudio Masolo et Stefano Borgo. *In defense of a trope-based ontology for conceptual modeling : An example with the foundations of attributes, weak entities and datatypes*. In International Conference on Conceptual Modeling, pages 112–125. Springer, 2006.

[Guizzardi2007] Giancarlo Guizzardi. *On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta)Models*. In Databases and Information Systems IV : Selected Papers from the Seventh International Baltic Conference, DB&IS'2006, volume 155, page 18. IOS Press, 2007.

[Guizzardi2008] Giancarlo Guizzardi et Terry Halpin. *Ontological foundations for conceptual modelling*. Applied Ontology, vol. 3, no. 1-2, pages 1–12, 2008.

[Guizzardi2009] Giancarlo Guizzardi. *The problem of transitivity of part-whole relations in conceptual modeling revisited*. In International Conference on Advanced Information Systems Engineering, pages 94–109. Springer, 2009.

[Guizzardi2011] Giancarlo Guizzardi. *Ontological foundations for conceptual part-whole relations : The case of collectives and their parts*. In International Conference on Advanced Information Systems Engineering, pages 138–153. Springer, 2011.

[Guizzardi2015] Giancarlo Guizzardi, Gerd Wagner, João Paulo Andrade Almeida et Renata SS Guizzardi. *Towards ontological foundations for conceptual modeling : The unified foundational ontology (UFO) story*. Applied ontology, vol. 10, no. 3-4, pages 259–271, 2015.

[Guizzardi2022] Giancarlo Guizzardi, Alessander Botti Benevides, Claudenir M Fonseca, Daniele Porello, João Paulo A Almeida et Tiago Prince Sales. *UFO : Unified foundational ontology*. Applied ontology, no. Preprint, pages 1–44, 2022.

[Hakimpour2001] Farshad Hakimpour et Andreas Geppert. *Resolving semantic heterogeneity in schema integration*. In Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001, pages 297–308, 2001.

[Halevy2001] Alon Y Halevy. *Answering queries using views : A survey*. The VLDB Journal, vol. 10, pages 270–294, 2001.

[Halpin2005] Terry Halpin. *ORM 2*. In On the Move to Meaningful Internet Systems 2005 : OTM 2005 Workshops : OTM Confederated Internationl Workshops and Posters, AWeSOMe, CAMS, GADA, MIOS+ INTEROP, ORM, PhDS, SeBGIS, SWWS, and WOSE 2005, Agia Napa, Cyprus, October 31-November 4, 2005. Proceedings, pages 676–687. Springer, 2005.

[Halpin2010] Terry Halpin. *Object-role modeling : Principles and benefits*. International Journal of Information System Modeling and Design (IJISMD), vol. 1, no. 1, pages 33–57, 2010.

[Harris2020] Keith Harris. *How individuals constitute group agents*. Canadian Journal of Philosophy, vol. 50, no. 3, pages 350–364, 2020.

[Harrison2013] Rodney Harrison. Heritage : critical approaches. Routledge, 2013.

[Haslhofer2010] Bernhard Haslhofer, Elaheh Momeni, Manuel Gay et Rainer Simon. *Augmenting Europeana content with linked data resources*. In Proceedings of the 6th International Conference on Semantic Systems, pages 1–3, 2010.

[Haslhofer2011] Bernhard Haslhofer et Antoine Isaac. *data. europeana. eu : The europeana linked open data pilot*. In International conference on dublin core and metadata applications, pages 94–104, 2011.

[Hastings2014] Janna Hastings, Gwen A Frishkoff, Barry Smith, Mark Jensen, Russell A Poldrack, Jane Lomax, Anita Bandrowski, Fahim Imam, Jessica A Turner et Maryann E Martone. *Interdisciplinary perspectives on the development, integration, and application of cognitive ontologies*. Frontiers in neuroinformatics, vol. 8, page 62, 2014.

[Hayes1979] Patrick J Hayes. *The naive physics manifesto*. Expert systems in the microelectronic age, 1979.

[Hayes1985] Patrick J Hayes. *Ontology for liquids*. Formal theories of the commonsense world, pages 71–107, 1985.

[Heflin2001] Jeff Heflin et James Hendler. *A portrait of the Semantic Web in action*. IEEE Intelligent Systems, vol. 16, no. 2, pages 54–59, 2001.

[Heller2004] Barbara Heller et Heinrich Herre. *Ontological categories in GOL*. Axiomathes, vol. 14, no. 1-3, pages 57–76, 2004.

[Henare2007] Amiria Henare, Martin Holbraad et Sari Wastell. Thinking through things : theorising artefacts ethnographically. Routledge, 2007.

[Hendler2001] James Hendler. *Agents and the semantic web*. IEEE Intelligent systems, vol. 16, no. 2, pages 30–37, 2001.

[Hermon2002] Sorin Hermon et Franco Niccolucci. *A fuzzy logic approach to typology in archaeological research*. The digital heritage of archaeology : CAA, pages 307–310, 2002.

[Herre2010] Heinrich Herre. *General Formal Ontology (GFO) : A foundational ontology for conceptual modelling*. In Theory and applications of ontology : computer applications, pages 297–345. Springer, 2010.

[Hiebel2017] Gerald Hiebel, Martin Doerr et Øyvind Eide. *CRMgeo : A spatio-temporal extension of CIDOC-CRM*. International Journal on Digital Libraries, vol. 18, pages 271–279, 2017.

[Hill2004] Linda Hill, Olha Buchel, Greg Janee et Zeng Lei. *Integration of knowledge organization systems into digital library architectures*. Data Analysis and Knowledge Discovery, vol. 20, no. 1, pages 4–8, 2004.

[Hitzler2020] Pascal Hitzler, Federico Bianchi, Monireh Ebrahimi et Md Kamruzzaman Sarker. *Neural-symbolic integration and the Semantic Web*. Semantic Web, vol. 11, no. 1, pages 3–11, 2020.

[Hodge2000a] Gail Hodge. Systems of knowledge organization for digital libraries : beyond traditional authority files, volume 91. Digital Library Federation, 2000.

[Hodge2000b] Gail Hodge. *Taxonomy of knowledge organization sources/systems*. CLIRPub91, 2000.

[Horrocks1999] Ian Horrocks, Ulrike Sattler et Stephan Tobies. *Practical reasoning for expressive description logics*. In International Conference on Logic for Programming Artificial Intelligence and Reasoning, pages 161–180. Springer, 1999.

[Horrocks2004a] Ian Horrocks et Peter F Patel-Schneider. *A proposal for an OWL rules language*. In Proceedings of the 13th international conference on World Wide Web, pages 723–731, 2004.

[Horrocks2004b] Ian Horrocks, Peter F Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean*et al.* *SWRL : A semantic web rule language combining OWL and RuleML*. W3C Member submission, vol. 21, no. 79, pages 1–31, 2004.

[Horrocks2005] Ian Horrocks, Oliver Kutz et Ulrike Sattler. *The Irresistible SRIQ*. In OWLED, 2005.

[Horrocks2006] Ian Horrocks, Oliver Kutz et Ulrike Sattler. *The Even More Irresistible SROIQ*. Kr, vol. 6, pages 57–67, 2006.

[Horrocks2008] Ian Horrocks. *Ontologies and the semantic web*. Communications of the ACM, vol. 51, no. 12, pages 58–67, 2008.

[Hull1997] Richard Hull. *Managing semantic heterogeneity in databases : a theoretical prospective*. In Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems, pages 51–61, 1997.

[Hunter2003] Jane Hunter. *Enhancing the semantic interoperability of multimedia through a core ontology*. IEEE Transactions on Circuits and Systems for Video Technology, vol. 13, no. 1, pages 49–58, 2003.

[Hyvönen2005] Eero Hyvönen, Eetu Mäkelä, Mirva Salminen, Arttu Valo, Kim Viljanen, Samppa Saarela, Miikka Junnila et Suvi Kettula. *Museum-Finland—Finnish museums on the semantic web*. Journal of Web Semantics, vol. 3, no. 2-3, pages 224–241, 2005.

[Hyvönen2006a] Eero Hyvönen. *FinnONTO—Building the Basis for a National Semantic Web Infrastructure in Finland*. In Proceedings of the 12th Finnish AI Conference STeP, volume 6, 2006.

[Hyvönen2006b] Eero Hyvönen, Tuukka Ruotsalo, Thomas Häggström, Mirva Salminen, Miikka Junnila, Mikko Virkkilä, Mikko Haaramo, Eetu Mäkelä, Tomi Kauppinen et Kim Viljanen. *Culturesampo–finnish culture on the semantic web : The vision and first results*. In Developments in Artificial Intelligence and the Semantic Web-Proceedings of the 12th Finnish AI Conference STeP, pages 26–27. Citeseer, 2006.

[Hyvönen2007] Eero Hyvönen, Kim Viljanen et Osma Suominen. *HealthFinland—Finnish health information on the semantic web*. In The Semantic Web : 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007+ ASWC 2007, Busan, Korea, November 11-15, 2007. Proceedings, pages 778–791. Springer, 2007.

[Hyvönen2008] Eero Hyvönen, Kim Viljanen, Jouni Tuominen et Katri Seppälä. *Building a national semantic web ontology and ontology service infrastructure–the FinnONTO approach*. In The Semantic Web : Research and Applications : 5th European Semantic Web Conference, ESWC 2008, Tenerife, Canary Islands, Spain, June 1-5, 2008 Proceedings 5, pages 95–109. Springer, 2008.

[Hyvönen2009a] Eero Hyvönen, Eetu Mäkelä, Tomi Kauppinen, Olli Alm, Jussi Kurki, Tuukka Ruotsalo, Katri Seppälä, Joeli Takala, Kimmo Puputti, Heini Kuittinen*et al.* *Culturesampo : A national publication system of cultural heritage on the semantic web 2.0*. In The Semantic Web : Research and Applications : 6th European Semantic Web Conference, ESWC 2009 Heraklion, Crete, Greece, May 31–June 4, 2009 Proceedings 6, pages 851–856. Springer, 2009.

[Hyvönen2009b] Eero Hyvönen, Eetu Mäkelä, Tomi Kauppinen, Olli Alm, Jussi Kurki, Tuukka Ruotsalo, Katri Seppälä, Joeli Takala, Kimmo Puputti, Heini Kuittinen*et al.* *CultureSampo—Finnish culture on the Semantic Web 2.0. Thematic perspectives for the end-user*. In Proceedings, museums and the web, pages 15–18, 2009.

[Hyvönen2010] Eero Hyvönen. *Developing and Using a National Cross-Domain Semantic Web Infrastructure*. Semantic Computing, pages 421–438, 2010.

[IEE1996] *IEEE Guide for Software Quality Assurance Planning*. IEEE Std 730.1-1995, pages 1–20, 1996.

[INSPIRE2013] INSPIRE. *D2.5 : Generic Conceptual Model, Version 3.4rc3*, 2013. June 2023.

[Iris1986] M.A. Iris, B.E. Litowitz et M.W. Evens. *The part-whole relation in the lexicon : an investigation of semantic primitives*. Unpublished manuscript, 1986.

[Isaac2013] Antoine Isaac*et al.* *Europeana data model primer*. 2013.

[ISO211272014] ISO21127. *Information and documentation — A reference ontology for the interchange of cultural heritage information*, 2014.

[ISO/IEC195012005] ISO/IEC19501. *Information technology — Open Distributed Processing — Unified Modeling Language (UML) Version 1.4.2*, 2005.

[ISO/IEC247072007] ISO/IEC24707. *Information Technology—Common Logic (CL) : A Framework for a Family of Logic-Based Languages*, 2007.

[Iyer1995] Hemalata Iyer. Classificatory structures : concepts, relations and representation, volume 2. Indeks Verlag, 1995.

[Jackson2012] Daniel Jackson. Software abstractions : logic, language, and analysis. MIT press, 2012.

[Johansson2004] Ingvar Johansson. *On the transitivity of the parthood relations*. 2004.

[Kakali2007] Constantia Kakali, Irene Lourdi, Thomais Stasinopoulou, Lina Bountouri, Christos Papatheodorou, Martin Doerr et Manolis Gergatsoulis. *Integrating Dublin Core metadata for cultural heritage collections using ontologies*. In International conference on Dublin core and metadata applications, pages 128–139, 2007.

[Katsumi2011] Megan Katsumi. A methodology for the development and verification of expressive ontologies. University of Toronto, 2011.

[Keet2006a] C Maria Keet. *Introduction to part-whole relations : mereology, conceptual modelling and mathematical aspects*. KRDB Research Centre Technical Report KRDB06-3 (Tutorial reader), Faculty of Computer Science, Free University of Bozen-Bolzano, Italy, vol. 2, 2006.

[Keet2006b] C Maria Keet. *Part-whole relations in object-role models*. In OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", pages 1118–1127. Springer, 2006.

[Keet2006c] C Maria Keet. *A taxonomy of types of granularity.* In GrC, pages 106–111, 2006.

[Keet2008] C Maria Keet et Alessandro Artale. *Representing and reasoning over a taxonomy of part–whole relations*. Applied Ontology, vol. 3, no. 1-2, pages 91–110, 2008.

[Keet2011] C Maria Keet*et al. The Use of Foundational Ontologies in Ontology Development : An Empirical Assessment.* In ESWC (1), pages 321–335, 2011.

[Keet2012] C Maria Keet, Francis C Fernández-Reyes et Annette Morales-González. *Representing mereotopological relations in OWL ontologies with ontoparts*. In Extended Semantic Web Conference, pages 240–254. Springer, 2012.

[Keet2017] C Maria Keet et Oliver Kutz. *Orchestrating a Network of Mereo (topo) logical Theories*. In Proceedings of the Knowledge Capture Conference, pages 1–8, 2017.

[Keet2018] C Maria Keet. An introduction to ontology engineering. College Publications, Cape Town, SA, 2018.

[Keet2022] C Maria Keet et Zubeida Khan. *Foundational ontologies : From theory to practice and back.* 2022.

[Khan2012] Zubeida Khan et C Maria Keet. *ONSET : Automated foundational ontology selection and explanation*. In Knowledge Engineering and Knowledge Management : 18th International Conference, EKAW 2012, Galway City, Ireland, October 8-12, 2012. Proceedings 18, pages 237–251. Springer, 2012.

[Khan2015] Zubeida C Khan et C Maria Keet. *Foundational ontology mediation in ROMULUS*. In Knowledge Discovery, Knowledge Engineering and Knowledge Management : 5th International Joint Conference, IC3K 2013, Vilamoura, Portugal, September 19-22, 2013. Revised Selected Papers 5, pages 132–152. Springer, 2015.

[Khan2016] Zubeida Casmod Khan et C Maria Keet. *Romulus : The repository of ontologies for multiple uses populated with mediated foundational ontologies*. Journal on Data Semantics, vol. 5, no. 1, pages 19–36, 2016.

[Klein1996] Julie Thompson Klein. Crossing boundaries : Knowledge, disciplinarities, and interdisciplinarities. University of Virginia Press, 1996.

[Klein2008] Julie T Klein. *Evaluation of interdisciplinary and transdisciplinary research : a literature review*. American journal of preventive medicine, vol. 35, no. 2, pages S116–S123, 2008.

[Klein2010] Julie Thompson Klein. *A taxonomy of interdisciplinarity*. The Oxford handbook of interdisciplinarity, vol. 15, no. 6, page 15, 2010.

[Knight1994] Kevin Knight et Steve K Luk. *Building a large-scale knowledge base for machine translation*. In AAAI, volume 94, pages 773–778, 1994.

[Knight1995] Kevin Knight, Ishwar Chander, Matthew Haines, Vasileios Hatzivassiloglou, Eduard Hovy, Masayo Iida, Steve K Luk, Richard Whitney et Kenji Yamada. *Filling knowledge gaps in a broad-coverage machine translation system*. arXiv preprint cmp-lg/9506009, 1995.

[Koch2004] Traugott Koch et Douglas Tudhope. *User-centred approaches to Networked Knowledge Organization Systems/Services (NKOS) : Background*, 2004.

[Kopytoff1986] Igor Kopytoff. *The cultural biography of things : commoditization as process*. The social life of things : Commodities in cultural perspective, vol. 68, pages 70–73, 1986.

[Krötzsch2012a] Markus Krötzsch, Sebastian Rudolph, Peter Patel-Schneider, Bijan Parsia et Pascal Hitzler. *OWL 2 Web Ontology Language Primer (Second Edition)*. Rapport technique, W3C, December 2012. https ://www.w3.org/TR/2012/REC-owl2-primer-20121211/.

[Krötzsch2012b] Markus Krötzsch, Frantisek Simancik et Ian Horrocks. *A description logic primer*. arXiv preprint arXiv :1201.4089, 2012.

[Kumar1998] Amruth N Kumar et Shambhu J Upadhyaya. *Component-ontological representation of function for reasoning about devices*. Artificial Intelligence in Engineering, vol. 12, no. 4, pages 399–415, 1998.

[Lach2014] Denise Lach. *Challenges of interdisciplinary research : Reconciling qualitative and quantitative methods for understanding human–landscape systems*. Environmental management, vol. 53, pages 88–93, 2014.

[Lagoze2002] Carl Lagoze et Jane Hunter. *The ABC ontology and model*. 2002.

[Lambrix2000] Patrick Lambrix. Part-whole reasoning in an object-centered framework. Springer Science & Business Media, 2000.

[Lei Zeng2004] Marcia Lei Zeng et Lois Mai Chan. *Trends and issues in establishing interoperability among knowledge organization systems*. Journal of the American Society for information science and technology, vol. 55, no. 5, pages 377–395, 2004.

[Leigh Star2010] Susan Leigh Star. *This is not a boundary object : Reflections on the origin of a concept*. Science, technology, & human values, vol. 35, no. 5, pages 601–617, 2010.

[Lelong2021] Florence Lelong, Emeline Pouyet, Sophie Champdavoine, Thomas Guiblain, Pauline Martinetto, Philippe Walter, Hélène Rousselière et Marine Cotte. *Des «brocarts appliqués» dans la sculpture savoyarde. Vers une caractérisation interdisciplinaire*. In CeROArt. Conservation, exposition, Restauration d'Objets d'Art, numéro HS. Association CeROArt asbl, 2021.

[Lenat1989] Douglas B Lenat et Ramanathan V Guha. Building large knowledge-based systems ; representation and inference in the cyc project. Addison-Wesley Longman Publishing Co., Inc., 1989.

[Lenzerini2002] Maurizio Lenzerini. *Data integration : A theoretical perspective*. In Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, pages 233–246, 2002.

[Leśniewski1991] Stanisław Leśniewski*et al.* *Stanisław Leśniewski's Collected Works*. 1991.

[Levy1996] Alon Y Levy et Marie-Christine Rousset. *CARIN : A representation language combining horn rules and description logics'*. In ECAI, pages 323–327. Citeseer, 1996.

[Lin2007] Hsiao-Kang Lin et Jenny A Harding. *A manufacturing system engineering ontology model on the semantic web for inter-enterprise collaboration*. Computers in Industry, vol. 58, no. 5, pages 428–437, 2007.

[Logan2007] William S Logan. Closing pandora's box : human rights conundrums in cultural heritage protection. Springer, 2007.

[Lowenthal1996] David Lowenthal. *Possessed by the Past the Heritage Crusade and the Spoils of History*. 1996.

[Lyons1977] John Lyons. Semantics : Volume 2, volume 2. Cambridge university press, 1977.

[Madsen2009] Bodil Nistrup Madsen et Hanne Erdman Thomsen. *Ontologies vs. classification systems*. In Proceedings of the NODALIDA 2009 workshop WordNets and other Lexical Semantic Resources—between Lexical Semantics, Lexicography, Terminology and Formal Ontologies, pages 27–32, 2009.

[Mahesh1996] Kavi Mahesh, Stephen Helmreich et Lori Wilson. Ontology development for machine translation : Ideology and methodology. New Mexico State University, Computing Research Laboratory Las Cruces, NM, 1996.

[Mandran2018] Nadine Mandran et Sophie Dupuy-Chessa. *Supporting experimental methods in information system research*. In 2018 12th International Conference on Research Challenges in Information Science (RCIS), pages 1–12. IEEE, 2018.

[Mandran2022] Nadine Mandran, Mathieu Vermeulen et Estelle Prior. *THEDRE's Framework : Empowering PhD Candidates to Efficiently Implement Design-Based Research*. Education and Information Technologies, vol. 27, no. 7, pages 9563–9586, 2022.

[Masolo2003] Claudio Masolo, Stefano Borgo, Aldo Gangemi, Nicola Guarino et Alessandro Oltramari. *WonderWeb Deliverable D18 Ontology Library (final)*. Rapport technique, IST Project 2001-33052 WonderWeb : Ontology Infrastructure for the Semantic Web, 2003.

[Masolo2020] Claudio Masolo, Laure Vieu, Roberta Ferrario, Stefano Borgo et Daniele Porello. *Pluralities, Collectives, and Composites*. In Formal Ontology in Information Systems : Proceedings of the 11th International Conference (FOIS 2020), volume 330, page 186. IOS Press, 10 2020.

[McCarthy1959] John McCarthy. *Programs with common sense*, 1959.

[McCune2007] William McCune. *Prover9 and Mace4*, 2007.

[Meskell2018] Lynn Meskell. A future in ruins : Unesco, world heritage, and the dream of peace. Oxford University Press, 2018.

[Miller1997] Daniel Miller. Material culture and mass consumerism. John Wiley & Sons, 1997.

[Miller2005] Daniel Miller*et al. Materiality : an introduction*. Materiality, vol. 1, pages 1–50, 2005.

[Mizoguchi2009] Riichiro Mizoguchi. *Yet another top-level ontology : Yato*. In Proceedings of the 2nd Interdisciplinary Ontology Meeting, volume 2, page 101. Citeseer, 2009.

[Mizoguchi2022] Riichiro Mizoguchi et Stefano Borgo. *YAMATO : Yet-another more advanced top-level ontology*. Applied ontology, vol. 17, no. 1, pages 211–232, 2022.

[Mäkelä2012] Eetu Mäkelä, Eero Hyvönen et Tuukka Ruotsalo. *How to deal with massively heterogeneous cultural heritage data - Lessons learned in CultureSampo*. Semantic Web, vol. 3, pages 85–109, 01 2012.

[Modoni2017] Gianfranco E Modoni, Michalis Doukas, Walter Terkaj, Marco Sacco et Dimitris Mourtzis. *Enhancing factory data integration through the development of an ontology : from the reference models reuse to the semantic conversion of the legacy models*. International Journal of Computer Integrated Manufacturing, vol. 30, no. 10, pages 1043–1059, 2017.

[Moraitou2019] Efthymia Moraitou, John Aliprantis, Yannis Christodoulou, Alexandros Teneketzis et George Caridakis. *Semantic bridging of cultural heritage disciplines and tasks*. Heritage, vol. 2, no. 1, pages 611–630, 2019.

[Moraitou2022] Efthymia Moraitou, Yannis Christodoulou et George Caridakis. *Semantic models and services for conservation and restoration of cultural heritage : A comprehensive survey*. Semantic Web, vol. 14, pages 1–31, 08 2022.

[Morales-González2015] Annette Morales-González, Francis C Fernández-Reyes et C Maria Keet. *OntoPartS : a tool to select part-whole relations in OWL ontologies*. In The Semantic Web : ESWC 2012 Satellite Events : ESWC 2012 Satellite Events, Heraklion, Crete, Greece, May 27-31, 2012. Revised Selected Papers 9, pages 452–457. Springer, 2015.

[Moser2009] Thomas Moser. *Semantic integration of engineering environments using an engineering knowledge base*. Thèse de Doctorat, 2009.

[Moser2016] Thomas Moser. The engineering knowledge base approach, pages 85–103. Springer International Publishing, Cham, 2016.

[Mossakowski2005] Till Mossakowski. *Heterogeneous specification and the heterogeneous tool set*. Thèse de Doctorat, Citeseer, 2005.

[Mossakowski2007] Till Mossakowski, Christian Maeder et Klaus Lüttich. *The heterogeneous tool set, Hets*. In Tools and Algorithms for the Construction and Analysis of Systems : 13th International Conference, TACAS 2007, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2007 Braga, Portugal, March 24-April 1, 2007. Proceedings 13, pages 519–522. Springer, 2007.

[Mossakowski2013] Till Mossakowski, Christian Maeder, Mihai Codescu et Dominik Lucke. *Hets user guide-version 0.99*. DKFI GmbH, Bremen, Germany, 2013.

[Mossakowski2014] Till Mossakowski, Mihai Codescu, Oliver Kutz, Christoph Lange et Michael Grüninger. *Proof Support for Common Logic*. In ARQNL@ IJCAR, pages 42–58, 2014.

[Mossakowski2015] Till Mossakowski, Mihai Codescu, Fabian Neuhaus et Oliver Kutz. *The distributed ontology, modeling and specification language–DOL*. The Road to Universal Logic : Festschrift for the 50th Birthday of Jean-Yves Béziau Volume II, pages 489–520, 2015.

[Motschnig-Pitrik1999] Renate Motschnig-Pitrik et Jens Kaasboll. *Part-whole relationship categories and their application in object-oriented analysis*. IEEE Transactions on Knowledge and Data Engineering, vol. 11, no. 5, pages 779–797, 1999.

[Munjeri2004] Dawson Munjeri. *Tangible and intangible heritage : From difference to convergence*. Museum international, vol. 56, no. 1-2, pages 12–20, 2004.

[Navarrete2013] Trilce Navarrete. *Digital cultural heritage*. In Handbook on the economics of cultural heritage, pages 251–271. Edward Elgar Publishing, 2013.

[Navigli2006a] Roberto Navigli et Paola Velardi. *Enriching a formal ontology with a thesaurus : an application in the cultural heritage domain*. In Proceedings of the 2nd workshop on ontology learning and population : bridging the gap between text and knowledge, pages 1–9, 2006.

[Navigli2006b] Roberto Navigli et Paola Velardi. *Ontology enrichment through automatic semantic annotation of on-line glossaries*. In Managing Knowledge in a World of Networks : 15th International Conference, EKAW 2006, Poděbrady, Czech Republic, October 2-6, 2006. Proceedings 15, pages 126–140. Springer, 2006.

[Niles2001] Ian Niles et Adam Pease. *Towards a standard upper ontology*. In Proceedings of the international conference on Formal Ontology in Information Systems-Volume 2001, pages 2–9, 2001.

[NISO2005] NISO. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*, 2005.

[Noor2019] Salma Noor, Lubna Shah, Mohammad Adil, Neelam Gohar, Gul E Saman, Sehrish Jamil et Fawad Qayum. *Modeling and representation of built cultural heritage data using semantic web technologies and building information model*. Computational and Mathematical Organization Theory, vol. 25, pages 247–270, 2019.

[Noy1997] Natalya Fridman Noy et Carole D Hafner. *The state of the art in ontology design : A survey and comparative review*. AI magazine, vol. 18, no. 3, pages 53–53, 1997.

[Noy2001] Natalya F Noy, Deborah L McGuinness*et al. Ontology development 101 : A guide to creating your first ontology*, 2001.

[Odell1998] James J Odell. Advanced object-oriented analysis and design using uml, volume 12. Cambridge University Press, 1998.

[of Europe2009] Council of Europe. Heritage and beyond. Council of Europe Publishing, 2009.

[Opdahl2001] Andreas L. Opdahl, Brian Henderson-Sellers et Franck Barbier. *Ontological analysis of whole–part relationships in OO-models*. Information and Software Technology, vol. 43, no. 6, pages 387–399, 2001.

[Parcero-Oubiña2013] César Parcero-Oubiña, Pastor Fábrega-Álvarez, Juan Manuel Vicent-García, Antonio Uriarte-González, Alfonso Fraguas-Bravo, Isabel del Bosque-González, Carlos Fernández-Freire et Esther Pérez-Asensio. *Conceptual basis for a cultural heritage data model for INSPIRE.* Rev. Int. Géomatique, vol. 23, no. 3-4, pages 445–467, 2013.

[Parsons2007] Josh Parsons. *Theories of location*. Oxford studies in metaphysics, vol. 3, pages 201–232, 2007.

[Partridge2020] Chris Partridge, Andrew Mitchell, Al Cook, Jan Sullivan et Matthew West. *A Survey of Top-Level Ontologies-to inform the ontological choices for a Foundation Data Model*. 2020.

[Patel2005] Manjula Patel, Traugott Koch, Martin Doerr, Chrisa Tsinaraki, Nektarios Gioldasis, Koraljka Golub et Doug Tudhope. *Semantic interoperability in digital library systems*, 2005.

[Pearce2017] Susan Pearce. Museums, objects, and collections : A cultural study. Smithsonian Institution, 2017.

[Peroni2013] Silvio Peroni, Francesca Tomasi et Fabio Vitali. *The aggregation of heterogeneous metadata in web-based cultural heritage collections : a case study*. International Journal of Web Engineering and Technology 8, vol. 8, no. 4, pages 412–432, 2013.

[Petersen1990] T. Petersen et J. Paul Getty Trust / Art History Information Program. Art & architecture thesaurus. Numeéro v. 1 in The Getty Art History Information Program. Oxford University Press, 1990.

[Phillips2008] JD Phillips et David Stanovskỳ. *Automated theorem proving in loop theory*. In Proceedings of the CICM Workshop on Empirically Successful Automated Reasoning in Mathematics, numéro 378, pages 42–54, 2008.

[Pinto2021] Ariane Pinto. *Par-dessus l'épaule du maître. Approche matérielle de l'adaptabilité artisanale médiévale : l'exemple des «brocarts appliqués» savoyards*. Bulletin du centre d'études médiévales d'Auxerrel BUCEMA, no. 25.2, 2021.

[Pinto2022] Ariane Pinto et Florian Bouquet. *Imiter des tissus de prestige par une technique prestigieuse : les «brocarts appliqués»*. e-Phaïstos. Revue d'histoire des techniques/Journal of the history of technology, no. X-2, 2022.

[Pisanelli2004] Domenico M Pisanelli. Ontologies in medicine, volume 102. IOS press, 2004.

[Plantinga1975] Alvin Plantinga. *On mereological essentialism.* The Review of Metaphysics, pages 468–476, 1975.

[Plassard1998] Marie-France Plassard*et al.* Functional requirements for bibliographic records (1998). KG Saur Verlag GmbH & Co. KG, 1998.

[Pohl2007] Christian Pohl et Gertrude Hirsch Hadorn. Principles for designing transdisciplinary research. oekom Munich, 2007.

[Pribbenow2002] Simone Pribbenow. *Meronymic relationships : From classical mereology to complex part-whole relations*. In The semantics of relationships, pages 35–50. Springer, 2002.

[Query2012] OGCOGCGA Query. *Language for RDF Data*. Open Geopatial Consortium, 2012.

[Quine1948] Willard V Quine. *On what there is*. The review of metaphysics, vol. 2, no. 1, pages 21–38, 1948.

[Quine1969] Willard Van Orman Quine. Ontological relativity and other essays. Numeéro 1. Columbia University Press, 1969.

[Rahm2001] Erhard Rahm et Philip A Bernstein. *A survey of approaches to automatic schema matching*. the VLDB Journal, vol. 10, pages 334–350, 2001.

[Randell1989] David A Randell et Anthony G Cohn. *Modelling Topological and Metrical Properties in Physical Processes.* KR, vol. 89, pages 357–368, 1989.

[Randell1992] David A Randell, Zhan Cui et Anthony G Cohn. *A spatial logic based on regions and connection.* KR, vol. 92, pages 165–176, 1992.

[Rector2006a] Alan Rector, Jeremy Rogers et Thomas Bittner. *Granularity, scale and collectivity : When size does and does not matter*. Journal of biomedical informatics, vol. 39, no. 3, pages 333–349, 2006.

[Rector2006b] Alan Rector, Jeremy Rogers et Thomas Bittner. *Granularity, scale and collectivity : When size does and does not matter*. Journal of biomedical informatics, vol. 39, no. 3, pages 333–349, 2006.

[Reed1995] Patricia Ann Reed. Cidoc relational data model : A guide. International Documentation Committee of the International Council of Museums, 1995.

[Riva2018] Pat Riva, Patrick Le Boeuf, Maja Žumer*et al. IFLA library reference model : A conceptual model for bibliographic information*. 2018.

[Robinson2001] Alan JA Robinson et Andrei Voronkov. Handbook of automated reasoning, volume 1. Elsevier, 2001.

[Rosati2005] Riccardo Rosati. *On the decidability and complexity of integrating ontologies and rules*. Journal of Web Semantics, vol. 3, no. 1, pages 61–73, 2005.

[Rosch1978] Eleanor Rosch et Barbara Bloom Lloyd. *Cognition and categorization*. 1978.

[Roy2001] Utpal Roy, Nilmani Pramanik, Rachuri Sudarsan, Ram D Sriram et Kevin W Lyons. *Function-to-form mapping : model, representation and applications in design synthesis*. Computer-Aided Design, vol. 33, no. 10, pages 699–719, 2001.

[Salmon1984] Wesley C Salmon. Scientific explanation and the causal structure of the world. Princeton University Press, 1984.

[Salomon2022] Hélène Salomon, Emilie Chalmin, Claire Chanteraud, Aurélie Chassin de Kergommeaux, Marianne Le Turnier, Giorgia Sardelli, Bernard Schmitt, Mathilde Waymel, Yvan Coquinot, Claudia Defrasne*et al. Constitution d'une" Pigmentothèque" : un outil pour comprendre l'approvisionnement en matériaux colorants durant la Préhistoire. Rapport intermédiaire 2021*. Thèse de Doctorat, DRAC-SRA Auvergne Rhône Alpes, 2022.

[Sattler2000] Ulrike Sattler. *Description logics for the representation of aggregated objects*. In ECAI, volume 9, pages 239–243, 2000.

[Schoeser2007] Mary Schoeser. Silk. Yale University Press, 2007.

[Schreiber2003] Guus Schreiber, Bob Wielinga et Wouter Jansweijer. *The KACTUS View on the 'O' Word*. 05 2003.

[Schulz2000] Stefan Schulz, Udo Hahn et Martin Romacker. *Modeling anatomical spatial relations with description logics*. In Proceedings of the AMIA Symposium, page 779. American Medical Informatics Association, 2000.

[Seyed2009] A Patrice Seyed. *BFO/DOLCE primitive relation comparison*. Nature Precedings, pages 1–1, 2009.

[Shadbolt2006] Nigel Shadbolt, Tim Berners-Lee et Wendy Hall. *The semantic web revisited*. IEEE intelligent systems, vol. 21, no. 3, pages 96–101, 2006.

[Shanks2004] Graeme Shanks, Elizabeth Tansley et Ron Weber. *Representing composites in conceptual modeling*. Communications of the ACM, vol. 47, no. 7, pages 77–80, 2004.

[Shaw2009] Ryan Shaw, Raphaël Troncy et Lynda Hardman. *Lode : Linking open descriptions of events*. In The Semantic Web : Fourth Asian Conference, ASWC 2009, Shanghai, China, December 6-9, 2009. Proceedings 4, pages 153–167. Springer, 2009.

[Shvaiko2005] Pavel Shvaiko et Jérôme Euzenat. *A survey of schema-based matching approaches*. In Journal on data semantics IV, pages 146–171. Springer, 2005.

[Simons1982] Peter Simons. *The Formalisation of Husserl ?s Theory of Wholes and Parts*. In Barry Smith, editeur, Parts and Moments. Studies in Logic and Formal Ontology, pages 111–159. Munich : Philosophia, 1982.

[Simons1987] Peter Simons. Parts : A study in ontology. Oxford, England : Oxford University Press, 1987.

[Sloman2005] Steven Sloman. Causal models : How people think about the world and its alternatives. Oxford University Press, 2005.

[Smith1997] Barry Smith. *On substances, accidents and universals : In defence of a constituent ontology*. Philosophical Papers, vol. 26, no. 1, pages 105–127, 1997.

[Smith1998] Barry Smith. *Basic concepts of formal ontology*. 1998.

[Smith2002] Barry Smith, Pierre Grenon, Holger Stenzhorn et Andrew Spear. *BFO Basic Formal Ontology*. Rapport technique, 2002.

[Smith2004a] Barry Smith. Ontology, chapitre 11, pages 153–166. John Wiley & Sons, Ltd, 2004.

[Smith2004b] Barry Smith et Cornelius Rosse. *The role of foundational relations in the alignment of biomedical ontologies*. 2004.

[Smith2005] Barry Smith, Werner Ceusters, Bert Klagges, Jacob Köhler, Anand Kumar, Jane Lomax, Chris Mungall, Fabian Neuhaus, Alan L Rector et Cornelius Rosse. *Relations in biomedical ontologies*. Genome biology, vol. 6, no. 5, pages 1–15, 2005.

[Smith2006] Laurajane Smith. Uses of heritage. Routledge, 2006.

[Smith2007] Barry Smith, Michael Ashburner, Cornelius Rosse, Jonathan Bard, William Bug, Werner Ceusters, Louis J Goldberg, Karen Eilbeck, Amelia Ireland, Christopher J Mungall*et al.* *The OBO Foundry : coordinated evolution of ontologies to support biomedical data integration*. Nature biotechnology, vol. 25, no. 11, pages 1251–1255, 2007.

[Smith2012] Barry Smith. *Ontology*. In The furniture of the world, pages 47–68. Brill, 2012.

[Smullyan1995] Raymond M Smullyan. First-order logic. Courier Corporation, 1995.

[Soergel1985] Dagobert Soergel. Organizing information : Principles of data base and retrieval systems. Elsevier, 1985.

[Sørensen2009] Estrid Sørensen. The materiality of learning : Technology and knowledge in educational practice. Cambridge University Press, 2009.

[Star1989] Susan Leigh Star et James R Griesemer. *Institutional ecology,translations' and boundary objects : Amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39*. Social studies of science, vol. 19, no. 3, pages 387–420, 1989.

[Studer1998] Rudi Studer, V Richard Benjamins et Dieter Fensel. *Knowledge engineering : Principles and methods*. Data & knowledge engineering, vol. 25, no. 1-2, pages 161–197, 1998.

[Sullivan2015] Ann Marie Sullivan. *Cultural heritage & new media : a future for the past*. J. Marshall Rev. Intell. Prop. L., vol. 15, page 604, 2015.

[Suominen2007] Osma Suominen, Kim Viljanen, Eero Hyvönen, Markus Holi et Petri Lindgren. *TerveSuomi.fi :n metatietomäärittely (TerveSuomi.fi metadata specification)*. Rapport technique, Semantic Computing Research Group, FinnONTO project, Jan 26 2007.

[Sure2003] York Sure, Hans Akkermans, Jeen Broekstra, John Davies, Ying Ding, Alistair Duke, Robert Engels, Dieter Fensel, Ian Horrocks, Victor Iosif*et al.* *On-To-knowledge : semantic web-enabled knowledge management*. In Web Intelligence, pages 277–300. Springer, 2003.

[Sure2004] York Sure, Steffen Staab et Rudi Studer. *On-to-knowledge methodology (OTKM)*. Handbook on ontologies, pages 117–132, 2004.

[Sutcliffe2017] Geoff Sutcliffe. *The TPTP problem library and associated infrastructure*. Journal of Automated Reasoning, vol. 59, no. 4, pages 483–502, 2017.

[Swartout1999] William Swartout et Austin Tate. *Guest editors' introduction : Ontologies*. IEEE Intelligent Systems, vol. 14, no. 01, pages 18–19, 1999.

[Tahko2020] Tuomas E. Tahko et E. Jonathan Lowe. *Ontological Dependence*. In Edward N. Zalta, editeur, The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, Fall 2020 edition, 2020.

[Temal2010] Lynda Temal, Arnaud Rosier, Olivier Dameron et Anita Burgun. *Mapping bfo and dolce*. In MEDINFO 2010, pages 1065–1069. IOS Press, 2010.

[Tessaris2001] Sergio Tessaris. Questions and answers : reasoning and querying in description logic. The University of Manchester (United Kingdom), 2001.

[Thomasson1999] Amie L Thomasson. Fiction and metaphysics. Cambridge University Press, 1999.

[Thomson1983] Judith Jarvis Thomson. *Parthood and identity across time*. the Journal of Philosophy, vol. 80, no. 4, pages 201–220, 1983.

[Thomson1998] Judith Jarvis Thomson. *The statue and the clay*. Noûs, vol. 32, no. 2, pages 149–173, 1998.

[Tudhope2004a] Douglas Tudhope et Ceri Binding. A case study of a faceted approach to knowledge organisation and retrieval in the cultural heritage sector. na, 2004.

[Tudhope2004b] Douglas Tudhope et Traugott Koch. *New Applications of Knowledge Organization Systems : introduction to a special issue*. 2004.

[Ullman1997] Jeffrey D Ullman. *Information integration using logical views*. In Database Theory—ICDT'97 : 6th International Conference Delphi, Greece, January 8–10, 1997 Proceedings 6, pages 19–40. Springer, 1997.

[Uschold1995] Michael Uschold et Martin King. Towards a methodology for building ontologies. Citeseer, 1995.

[Uschold1996] Mike Uschold et Michael Gruninger. *Ontologies : Principles, methods and applications*. The knowledge engineering review, vol. 11, no. 2, pages 93–136, 1996.

[Uschold1998] Mike Uschold, Martin King, Stuart Moralee et Yannis Zorgios. *The enterprise ontology*. The knowledge engineering review, vol. 13, no. 1, pages 31–89, 1998.

[van Assem2004] Mark van Assem, Maarten R. Menken, Guus Schreiber, Jan Wielemaker et Bob Wielinga. *A Method for Converting Thesauri to RDF/OWL*. In Sheila A. McIlraith, Dimitris Plexousakis et Frank van Harmelen, editeurs, The Semantic Web – ISWC 2004, pages 17–31, Berlin, Heidelberg, 2004. Springer Berlin Heidelberg.

[Van Duijn2012] Esther Van Duijn et Jessica Roeders. *Gold-brocaded velvets in paintings by Cornelis Engebrechtsz*. Journal of Historians of Netherlandish Art, vol. 4, no. 1, pages 1–28, 2012.

[Varzi1993] Achille C Varzi. On the boundary between mereology and topology. Istituto per la Ricerca Scientifica e Tecnologica, 1993.

[Varzi1996] Achille C Varzi. *Parts, wholes, and part-whole relations : The prospects of mereotopology*. Data & Knowledge Engineering, vol. 20, no. 3, pages 259–286, 1996.

[Varzi2003] Achille Varzi. *Mereology*. In Edward N. Zalta, editeur, The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, 2003.

[Varzi2006] Achille C Varzi. *A note on the transitivity of parthood*. Applied ontology, vol. 1, no. 2, pages 141–146, 2006.

[Varzi2007] Achille C Varzi. *Spatial reasoning and ontology : Parts, wholes, and locations*. In Handbook of spatial logics, pages 945–1038. Springer, 2007.

[Vavliakis2012] Konstantinos N Vavliakis, Georgios Th Karagiannis et Pericles A Mitkas. *Semantic Web in cultural heritage after 2020*. In Proceedings of the 11th International Semantic Web Conference (ISWC), Boston, MA, USA, pages 11–15, 2012.

[Velios2021] Athanasios Velios*et al. Definition of the CIDOC Conceptual Reference Model*, 2021.

[Verdonck2016] Michael Verdonck et Frederik Gailly. *Insights on the use and application of ontology and conceptual modeling languages in ontology-driven conceptual modeling*. In Conceptual Modeling : 35th International Conference, ER 2016, Gifu, Japan, November 14-17, 2016, Proceedings 35, pages 83–97. Springer, 2016.

[Verdonck2019] Michaël Verdonck, Frederik Gailly, Robert Pergl, Giancarlo Guizzardi, Beatriz Martins et Oscar Pastor. *Comparing traditional conceptual modeling with ontology-driven conceptual modeling : An empirical study*. Information Systems, vol. 81, pages 92–103, 2019.

[Vieu2007] Laure Vieu et Michel Aurnague. *Part-of relations, functionality and dependence*. The categorization of spatial entities in language and cognition, vol. 20, pages 307–336, 2007.

[Viljanen2008] Kim Viljanen, Jouni Tuominen, Teppo Känsälä et Eero Hyvönen. *Distributed semantic content creation and publication for cultural heritage legacy systems*. In Proceedings of the 2008 IEEE International Conference on Distibuted Human-Machine Systems, Athens, Greece, IEEE Press, 2008.

[Viñas2002] Salvador Muñoz Viñas. *Contemporary theory of conservation*. Studies in Conservation, vol. 47, no. sup1, pages 25–34, 2002.

[Vögele2003] Thomas Vögele, Sebastian Hübner et Gerhard Schuster. *BUSTER - An Information Broker for the Semantic Web*. KI, vol. 17, no. 3, page 31, 2003.

[Wache1999] Holger Wache, Th Scholz, Helge Stieghahn et Brigitta Konig-Ries. *An integration method for the specification of rule-oriented mediators*. In Proceedings 1999 International Symposium on Database Applications in Non-Traditional Environments (DANTE'99)(Cat. No. PR00496), pages 109–112. IEEE, 1999.

[Wache2001] Holger Wache, Thomas Voegele, Ubbo Visser, Heiner Stuckenschmidt, Gerhard Schuster, Holger Neumann et Sebastian Hübner. *Ontology-based integration of information-a survey of existing approaches.* In Ois@ ijcai, 2001.

[Warmer1998] Jos Warmer et Anneke Kleppe. The object constraint language : precise modeling with uml. Addison-Wesley Longman Publishing Co., Inc., 1998.

[Whitehead1925] Alfred North Whitehead. An enquiry concerning the principles of natural knowledge. University press, 1925.

[Wielinga2001] Bob J Wielinga, A Th Schreiber, Jan Wielemaker et JAC Sandberg. *From thesaurus to ontology*. In Proceedings of the 1st international conference on Knowledge capture, pages 194–201, 2001.

[Winston1987] Morton E Winston, Roger Chaffin et Douglas Herrmann. *A taxonomy of part-whole relations*. Cognitive science, vol. 11, no. 4, pages 417–444, 1987.

[Zemmouchi-Ghomari2012] Leila Zemmouchi-Ghomari et Abdessamed Réda Ghomari. *Ontology versus terminology, from the perspective of ontologists*. International journal of Web science, vol. 1, no. 4, pages 315–331, 2012.

[Zeng2008] Marcia Lei Zeng. *Knowledge organization systems (KOS)*. KO Knowledge Organization, vol. 35, no. 2-3, pages 160–182, 2008.

Table des figures

Liste des tableaux

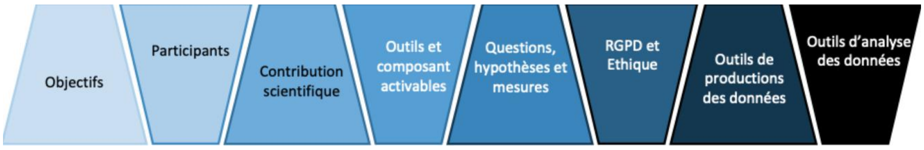
Listings

A

Annex 1 : Interviews with Patrimalp domain experts

In this Annex, we attach the documents used for carrying out the experimental procedure with Patrimalp's domain experts ; the experimental protocol and the questionnaire.

« Write an Experimental Protocol»



Why: Write the experimental protocol for producing data and specify the measures and methods for producing and analyzing data
When: The tools to be tested are produced, when the data production methods have been chosen. To be used to write the methodological parts of the thesis.

This document can be used in the thesis to explain how and why the experiments were implemented.

Objectives and goals of the experiment

Describe the scientific knowledge involved in the experiment

The terms used by the users to answer the questions as describing their work will serve in understanding their instances of work as; the concepts and the relations between them.

Indicate how the experiment will advance scientific knowledge

The produced information will serve the development of the intended ontology: conceptualizing the examples of entities that they describe to serve as "templates", and the properties between these entities and within the entities themselves, to serve as "relations".

Indicate the state of scientific knowledge (to be developed, improved, evaluated, etc.)

To be developed.

Questions or assumptions and measures
Indicate the questions and/or hypotheses that will have to be answered during this experiment

A list of questions is built as a separate document; "The Patrimalp Questionnaire".

List the measures that will be taken during the experiment: activities, expectations, needs, changes in practices, satisfaction, etc.

The description of the spatial and structural setup of the entity.

Participants in the experiment
Describe the profile of the participants in the experiment

Participants are domain experts in the Patrimalp Project e.g. material scientists, geologists, archaeologists, etc.

Méthode THEDRE, N.Mandran 2020, Version 2

Indicate the state of knowledge we have of the participants

Each participant is a specialist is his/her domain. They acquire different backgrounds, knowledge, and viewpoints of cultural heritage entities. However, they share common interests towards these entities.

Indicate why these participants are mobilized: observe in real situations, discuss with them, quantify their practices, make them share ideas with each other, etc.

Interviews will be held with each specialist individually to avoid the deviation of the discussion towards a pure social sciences debate, and to assure the accuracy and perspective-based information.

Indicate what will be asked of them during the experiment

Answer and elaborate on some questions.

Indicate the number of participants

5 participants.

Indicate where the experiment take place (in lab, in situ,...)

In Grenoble, at the LIG laboratory (IMAG), and in Chambery at the Edytem laboratory.

Indicate whether participants are consulted alone or in groups

Individual consultations.

Recruitment:

Indicate how participant's recruitment is done

They are members in the Patrimalp project.

Indicate the steps taken with the CNIL or an ethics committee to declare the experiment

None.

Temporality / measurement period

Timeliness of teaching and timing of monitoring or evaluation

Each interview lasts for around an hour.

Existing data production tools

Indicate whether questionnaires or measurement tools exist in the literature

None.

Data production tools to be produced
- Indicate if you need to create your own data production tools

Interviews.

- Using the flowchart (MATUI), indicate the given production methods you will use

No need for this Chart.

- List the experimental material to be built to carry out the experiment (e.g., presentation, questionnaire)

Questionnaire.

Data produced

- Indicate all the material and data produced during this experiment (diagram, audio, traces)

Voice records (with the consent of participants), written descriptions, and answers to checklist/options-list.

- Describe the format of the traces to be produced if a digital tool is used

For developing the list of examples, and the relations used to describe their composition.

Data analysis tools

- List the analytical tools and methods that will be used to analyze the data. Specify the provisional treatment plan

Thematic analysis

Technical equipment

- Indicate the technical equipment required for data capture (e. g. camera, recorder)

Recorder.

Participant:

Lab & Research field:

Hypothesis:

Objective: The objective of this interview is to understand example of cultural heritage entities and acquire the necessary information to represent their composition, from the viewpoint of domain experts.

Duration: 1 hour

1- What examples of cultural heritage entities do you consider within you research field? Choose one to describe, and elaborate on your interests and studies regarding this entity.

2- Describe the composition of this entity along with its spatial conditions. How does the study of this composition serve you interests?

3- While decomposing this entity to understand its structure; What examples of parts do you regard? How can you describe this part? Its relation to the whole entity? Is there existential dependence between the two entities?

4- What locative (spatial) relations do you consider while studying the conditions of the entity? Are there any correlations between entities found within the same region? Between entities

5- What changes (events) may an entity undergo e.g. structural or location changes? How do these changes affect the entity? Its identity? Its parts?

B

Annex 2 : Consistency Checks of FORT micro-theories

In this Annex, we attach the result of the consistency checks run over the micro-theories of FORT.

Results

return return to DGraph

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/FORT_all

Results for (Consistent)

ttps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/unified_entity_definition

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/memberOf_theorems

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/memberOf_root

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/existing

Results for (Consistent)

:tps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/entity_location_theorems

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/entity_location_root

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/elementOf_theorems

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/elementOf_definition

Results for (Consistent)

ıttps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/dependence_theorems

Results for (Consistent)

ttps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/dependence_definitions

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/constitution_C

Results for (Consistent)

ttps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/componentOf_theorems

Results for (Consistent)

ttps://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/componentOf_definition

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/colore/mereotopology/definitions/mereotopology_definitions

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/colore/mereotopology/cem_mt_mereotopology

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/colore/mereology/definitions/mereology_definitions

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/colore/mereology/cem_mereology

Results for (Consistent)

https://raw.githubusercontent.com/DanashFatima/FORT/main/FORT-CL-ontology/colore/location_varzi/location_L

Results for (Consistent)