UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

FERNANDO CICCONETO

GeoReservoir: An Ontology for Deep-marine Depositional System Description

Thesis presented in partial fulfillment of the requirements for the degree of Master of Computer Science

Advisor: Prof. Dr. Mara Abel

Porto Alegre
March 2021

CIP — CATALOGING-IN-PUBLICATION

Cicconeto, Fernando

GeoReservoir: An Ontology for Deep-marine Depositional System Description / Fernando Cicconeto. – Porto Alegre: PPGC da UFRGS, 2021.

94 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2021. Advisor: Mara Abel.

1. Ontology. 2. Artificial intelligence. 3. Deep-marine depositional system. 4. Turbidite. 5. Petroleum reservoir. 6. Geology. I. Abel, Mara. II. Título.

ACKNOWLEDGEMENTS

I thank my colleagues from BDI (Intelligent Databases) group, especially Alcides Gonçalves Lopes Junior, Cauã Roca Antunes, Fabrício Henrique Rodrigues, Givanildo Santana do Nascimento, Joel Luis Carbonera, Luan Fonseca Garcia, and Nicolau Oyhenard dos Santos, for giving me essential technical support and for participating in fruitful discussions throughout this work's development.

I would like to express my gratitude to the geologists who heavily collaborated in this work: Lucas Valadares Vieira (CENPES – Petrobras), Renata dos Santos Alvarenga (Geosciences Institute and BDI – UFRGS), Tiago Agne de Oliveira (CENPES – Petrobras), and Yuanwei Qu (University of Oslo). Without their contributions, it would be impossible for me, a computer scientist, to understand the geological concepts which this thesis required for its development.

I thank my advisor, Prof. Mara Abel, for allowing me to take the master's course that resulted in this thesis and for all the assistance and guidance that I needed to develop this work.

I acknowledge the Brazilian funding agencies CNPq and CAPES for financing this work in part.

ABSTRACT

Ontology is the logical theory that accounts for a domain's vocabulary intended meaning, which allows us to develop computational artifacts that explicitly and formally define the conceptualization associated with a particular set of vocabulary inside a domain. This master thesis presents the result of the ontological analysis of the terminology adopted in sedimentological studies and proposes a domain ontology to support the description of deep-marine depositional system geological occurrences. The domain in focus is of special interest primarily because of its economic importance: deep-marine deposits constitute the main type of petroleum reservoir explored nowadays in the world. Further on, the task in focus demands new approaches in conceptual modeling methodologies because the available terminology describes an interpretation of the studied occurrences instead of a rational description of the visually recognized objects. The ontological substrate helps in choose and define the precise vocabulary requested for the descriptive capturing of data. Therefore, the project aimed to collect and disambiguate the Geology terminology to describe reservoir occurrences in deep-marine depositional systems and build a domain ontology for reservoir description: the GeoReservoir. A team of professional reservoir geologists supported the knowledge acquisition process. We developed the ontology in iterative steps from an initial prototype to a complete taxonomy of entities, relations, and qualities, increasingly refined by the experts. The process implied the specialization of the concepts of two previously existent upper-level ontologies to build GeoReservoir: the Basic Formal Ontology (BFO) and the GeoCore ontology. This preexisting framework provided us the necessary structure and organization to focus our analysis in the domain, reducing the development effort. The resulting ontology includes a taxonomy of sedimentary units, lithologies, geological structures, and contact types. This ontology is complemented by a spatial relations ontology that formalizes the mereotopological relations of the described geological objects. The geological occurrences of Karoo turbidites in South Africa offer the data of a case study for testing and initial validation of the proposed ontology. The ontology is under continuous expansion through the inclusion of terminology for other depositional environments. A public accessible repository stores an implemented version of the model and the application case study.

Keywords: Ontology. Artificial intelligence. Deep-marine depositional system. Turbidite. Petroleum reservoir. Geology.

GeoReservoir: Uma Ontologia para a Descrição de Sistemas Deposicionais Marinhos Profundos

RESUMO

Ontologia é a teoria lógica que contempla o significado pretendido do vocabulário de um domínio, permitindo o desenvolvimento de artefatos computacionais que definem a conceitualização associada ao vocabulário de maneira explícita e formal. Esta dissertação de mestrado apresenta o resultado da análise ontológica da terminologia adotada em estudos de sistemas deposicionais marinhos profundos e propõe uma ontologia de domínio para suportar a descrição destas ocorrências geológicas. O domínio em foco é de especial interesse por conta de sua importância econômica: depósitos marinhos profundos constituem o principal tipo de reservatório de petróleo nos dias atuais. Além disto, a tarefa em foco demanda novas abordagens em modelagem conceitual porque a terminologia disponível descreve uma interpretação das ocorrências estudadas ao invés de uma descrição racional dos objetos visualmente reconhecidos. O substrato ontológico auxilia na escolha e definição precisa do vocabulário necessário para a captura descritiva dos dados. Sendo assim, o projeto teve como objetivo a captura e desambiguação de termos geológicos utilizados para descrever ocorrências de reservatórios e o desenvolvimento de uma ontologia de domínio: a GeoReservoir. Este projeto contou com o suporte de um time de geólogos de reservatório profissionais. A ontologia foi desenvolvida em passos iterativos, desde um protótipo inicial até uma taxonomia completa de entidades, relações e qualidades, refinados de forma incremental pelos especialistas. O processo implicou na especialização dos conceitos de duas ontologias previamente existentes: a Basic Formal Ontology (BFO) e a ontologia GeoCore. Este arcabouço preexistente proveu a estrutura e organização necessários para manter o foco da análise no domínio de interesse, reduzindo o esforço de desenvolvimento. A ontologia resultante inclui uma taxonomia de unidades sedimentares, litologias, estruturas geológicas e tipos de contatos. Esta ontologia é complementada por uma ontologia de relações espaciais que formaliza as relações mereotopológicas dos objetos geológicos descritos. Este trabalho inclui um estudo de caso para teste e validação inicial da ontologia proposta. A ontologia está sob expansão contínua através da inclusão das terminologias de outros ambientes deposicionais.

Palavras-chave: Ontologia. Inteligência artificial. Sistema deposicional marinho profundo. Turbidito. Reservatório de petróleo. Geologia.

LIST OF ABBREVIATIONS AND ACRONYMS

BFO	Basic Formal Ontology
DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
ER	Entity-Relationship
IAO	Information Artifact Ontology
MTD	Mass Transport Deposit
OBO	Open Biological and Biomedical Ontology
OWL	Ontology Web Language
RCC	Region Connection Calculus
RDF	Resource Description Framework
UFO	Unified Foundational Ontology
UML	Unified Modeling Language
XML	eXtensible Markup Language

LIST OF FIGURES

Figure 2.1 An illustration of some observation scales in Geology. This work focuses on descriptions made in the outcrop and seismic section scale.......................18

Figure 2.2 An overview of deep-marine depositional systems. Labels indicate some of the most significant sedimentary bodies that compose these systems.................18

Figure 2.3 A schematic representation of a channel with adjacent levees in a cross-section view. Yellow represents sand-rich sediments, green represents mud-rich sediments, brown represents mud-clast-rich sediments, and gray represents pre-existing sediments. 19

Figure 2.4 A hierarchical classification scheme for channels. This classification can also be analogously applied to lobes. 21

Figure 2.5 A drawing over satellite images representing the Karoo deposits outcrop (A) near Laingsburg, South Africa (B). The deposits' location is approximate.......22

Figure 2.6 The correlation panel featuring the representation of the channel-levee systems described in this section. The detailed panel is available in the source. The red rectangle outlines the part of the channel-levee system that we use in this work's case study..23

Figure 3.1 Diagram depicting some of the most relevant concepts from the ontology of Qu (2020). White boxes represent BFO and GeoCore concepts, while gray boxes represent the concepts defined by the ontology. Arrows represent "subsumed by" (i.e., taxonomical) relations..29

Figure 3.2 Representation of the relational database schema proposed by Cullis et al. (2019). Boxes represent tables, connecting lines represent relationships, and colored areas represent the three conceptual categories of data described by the authors: geological units, spatial relationships, and metadata......................32

Figure 3.3 Representation of the knowledge graph proposed by Le Bouteiller et al. (2019), where dots correspond to nodes, arrows correspond to directed edges, and lines correspond to undirected edges. 35

Figure 3.4 Hierarchical classification for channellized deposits according to Sprague et al. (2005), Moraes, Blaskovski and Paraizo (2006), and McHargue et al. (2011).38

Figure 3.5 An illustration of the Sedimentary Facies concept as defined by Lorenzatti et al. (2010) with its relations with the datatypes, the Fossil concept, and the Sedimentary Structure concept. 40

Figure 4.1 The Ullmann's triangle, which draws the relations between a symbol in a language, a concept in the mind of one or more users of this language, and a thing in reality. 45

Figure 4.2 Illustration of a modeling example of Rock and its related concepts. Boxes represent concepts, arrows with closed endings represent "is a" relations, and arrows with open endings represent other relations. 46

Figure 4.3 Diagram representing the ontology classification as addressed by this work. Rounded boxes bounded by solid lines represent ontology types, while rounded boxes bounded by dashed lines represent example ontologies for each ontology type. Arrows represent relations between the types or ontologies. It is interesting to notice that the GeoReservoir ontology directly specializes both the BFO and the GeoCore ontologies as it uses entities of these two ontologies in its formal definitions. 49

Figure 4.4 Diagram representing the BFO continuants. Boxes represent entities and arrows represent “is a” relations....50
Figure 4.5 Diagram representing the GeoCore continuants. Gray boxes represent GeoCore entities, while white boxes represent BFO entities. Arrows represent “is a” relations....53
Figure 4.6 Visual representation of the RCC-8 relations. Circles or ellipses represent spatial entities....56
Figure 4.7 Examples of distinct frames of reference for directional spatial relations. The relation illustrated in (A) uses the cardinal reference system as its frame of reference, while the relation illustrated in (B) assumes an intrinsic frame of reference....56

Figure 5.1 Diagram representing the relations between the ontologies depicted in this work. Gray rounded boxes represent the ontologies developed in this work, while white rounded boxes represent the ontologies reused by this work. Arrows represent relations between the ontologies....59
Figure 5.2 The GeoReservoir ontology continuous development process. Each iteration (i.e., from scope definition to evaluation) creates an evolving prototype that can be approved or not. If not approved, a new iteration is started....59
Figure 5.3 GeoReservoir ontology substantial entities’ taxonomical overview. White boxes represent BFO and GeoCore entities, while gray boxes represent GeoReservoir entities. All arrows represent “subsumed by” relations....64
Figure 5.4 GeoReservoir ontology moment entities’ taxonomical overview. White boxes represent BFO and GeoCore entities, while gray boxes represent GeoReservoir entities. All arrows represent “subsumed by” relations....67
Figure 5.5 Geological Spatial Relations ontology overview. White boxes represent BFO relations, while gray boxes represent our ontology’s relations. Arrows with closed endings represent Subrelation axioms, and arrows with open endings represent other axioms....71

Figure 6.1 The subset of Karoo basin geological data used in this work’s case study....77
Figure 6.2 The spatial distribution model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols, while arrows represent relations....78
Figure 6.3 The sedimentary facies model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols, while arrows represent relations....79
Figure 6.4 The sedimentary rocks model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols or entities when they don’t have chevron symbols. Arrows with closed ending represent Subsumed By relations, while arrows with open endings represent other relations....80

Figure 7.1 Figure that illustrates the two meanings of the “channel” term defined in this work. Boxes represent entities, while arrows represent relations....82

LIST OF TABLES

Table 3.1 Literature comparative analysis..27

CONTENTS

1 INTRODUCTION **12**
1.1 Objectives **13**
1.2 Thesis structure **15**
2 APPLICATION DOMAIN **17**
2.1 Deep-marine depositional systems **17**
2.1.1 Sedimentary body types 19
2.1.2 Hierarchical classifications 20
2.2 An instance of deep-marine depositional system **21**
2.3 Discussion about the domain terminology characteristics **23**
2.4 Final considerations **25**
3 LITERATURE REVIEW **26**
3.1 An ontology for deep-marine depositional system interpretation **27**
3.2 A database for quantitative analysis of deep-marine instances **31**
3.3 A knowledge model for description and intepretation of mass transport deposits (MTD) **34**
3.4 Conceptual models for deep-marine recurring features description **36**
3.5 An ontology for the description of visual sedimentological properties **39**
3.6 Models for general Geology and Geosciences description **40**
3.7 Final considerations **43**
4 ONTOLOGICAL FOUNDATIONS AND BACKGROUND **44**
4.1 What is an ontology? **44**
4.2 Structural ontological relations and metaproperties **45**
4.3 Ontology classification by scope **48**
4.4 The BFO ontology **49**
4.5 The GeoCore ontology **52**
4.6 Qualitative spatial representation **55**
4.7 Final considerations **57**
5 GEORESERVOIR ONTOLOGY DEVELOPMENT **58**
5.1 Methods and tools **58**
5.2 Formal notations **60**
5.3 GeoReservoir ontology presentation **62**
5.3.1 Substantial entities 63
5.3.2 Moment entities 67
5.3.3 Auxiliary relations 70
5.4 Geological Spatial Relations ontology presentation **70**
5.5 Final considerations **74**
6 ONTOLOGY APPLICATION **76**
6.1 Material instances and their mereotopology **77**
6.2 Sedimentary facies and lithological features **78**
6.3 Final considerations **79**
7 DISCUSSION **81**
7.1 GeoReservoir: a tool for language analysis and disambiguation in deep-marine settings **81**
7.2 Separation between geological description and process interpretation **83**
7.3 GeoReservoir's reuse potential in other depositional environments **84**
7.4 Comparison with the GeoSciML and SWEET approaches **85**
8 CONCLUSION **86**
8.1 Future work **86**

APPENDIX A – RESUMO EXPANDIDO EM PORTUGUÊS 88
REFERENCES..90

1 INTRODUCTION

Deep-marine sandstone deposits are among the most relevant types of hydrocarbon reservoirs around the world (BRUHN et al., 2003; SULLIVAN et al., 2004; MARTINSEN; LIEN; JACKSON, 2005). Much of the knowledge on these deposits is gathered from observation and interpretation of seismic data, calibrated by smaller-scale and higher-resolution well core and analogous outcrop data. This method is used due to the inaccessibility and large spatial scale of the sedimentary bodies that compose the deposits (ARNOTT, 2010).

As the petroleum industry's demand accelerated over the past decades, large volumes of geological data had been created for different usages and purposes (QU, 2020). While geologists investigated the data, they began to identify recurring features in distinct datasets and developed conceptual models, which were intended to help evaluate the geometry, architecture, and stacking pattern of most deep-marine deposits around the world, further enhancing the characterization and prediction of petroleum reservoirs (SPRAGUE et al., 2005; MORAES; BLASKOVSKI; PARAIZO, 2006; MCHARGUE et al., 2011; CULLIS et al., 2019; LE BOUTEILLER et al., 2019). However, variations in terminology and criteria applied to classify sedimentary bodies in distinct models became a hindrance in studying deep-marine deposits (CULLIS et al., 2018).

Comparing geological data from studies that commit to distinct models and even sharing knowledge between geologists from different schools or research interests is a common challenge in petroleum exploration. The geological language is not the result of logical evolution but a natural language that rose from many subjective concepts. Such a language is often ambiguous and misleading for those who do not master it (ABEL; PERRIN; CARBONERA, 2015; GARCIA et al., 2020). For example, in deep-marine depositional systems, a channel can be either a conduit for sediment currents or a depositional unit consisting of the consolidated sediments that fill a conduit.

This problem becomes especially evident when one intends to develop specialized software applications to capture and interpret reservoir data. Such an application needs to integrate heterogeneous data from different sources, while computers require a clear and unambiguous vocabulary and uniform data format to process and interpret large volumes of geological data using data analysis techniques. Geologists and scholars can depict the domain language semantics with some effort analyzing the context and combining with it previous knowledge, but computers cannot do that. To make this geological language processable, it is necessary to build a formal vocabulary that is not open to more than one interpretation (GARCIA et al., 2020). Such an explicit terminology would allow a more fluid development of specialized systems and AI models for applications such as reservoir visualization, pattern determination, analogous search, and many others.

This kind of application scenario in geological information systems is discussed by Abel, Perrin and Carbonera (2015), who claim that the interpretation heterogeneity in Geosciences can be amended using ontologies to analyze the conceptual models behind systems, geological reports, and languages. Furthermore, an empirical study conducted by Verdonck et al. (2019) demonstrated that conceptual modeling aided by ontology foundations makes novice modelers produce higher quality models than traditional methods and tools such as ER, UML, and others, especially when dealing with challenging and ambiguous aspects of the modeled domain. The use of ontologies leads modelers to reason more carefully about the domain and its classification in terms of concepts and their underlying relationships, making the models more suitable and truthful to a set of phenomena (VERDONCK et al., 2019).

In Computer Science, an ontology is a computational artifact built to formally organize and represent the structure of a domain of interest. An ontology takes the benefit of having philosophical foundations and making entities and relations explicit to describe their intended meaning and represent the worldview of a given community of stakeholders (STUDER; BENJAMINS; FENSEL, 1998; GUARINO; OBERLE; STAAB, 2009).

An ontology is composed of all the entities and relations of interest in a given domain. In the ontological sense, an entity represents a concept, which abstracts the necessary conditions that an instance must hold to match it in terms of properties and relations (GUIZZARDI, 2005). For example, a Rock concept would abstract properties and relations like being hard and constituted by minerals or biological matter. As so, an ontology is a suitable tool to describe the concepts behind terminologies in challenging domains such as Geology.

1.1 Objectives

This thesis's main objectives are to describe the GeoReservoir domain ontology development and to guide its usage. GeoReservoir is an ontology supporting the description of deep-marine deposits' geometry and lithology, which we put available for reuse[1].

---

[1] <https://github.com/BDI-UFRGS/GeoReservoirOntology>

This research covers the ontology development process and the domain ontology's main structural definitions. This approach's final goal is to allow the petroleum exploration team to populate a database of reservoir occurrences able to be compared and analyzed under a uniform reference view, without possible contamination of geological school of thought or interpretation tendencies. Although GeoReservoir's terminology embodies the domain semantics in a human view, it is intended to facilitate computer processing on geological occurrence data, which requires some level of simplification and uniformization over the collected data. We demonstrate our approach by defining the set of concepts on turbidite siliciclastic deposits, even though the semantic structure is intended to support the definition of all types of marine and lacustrine reservoir systems.

This ontology provides formal and clear definitions for describing the main geometric elements and lithologies detected in reservoir studies in outcrop and seismic study scales. We based our approach on the material philosophical perspective provided by the Basic Formal Ontology (BFO), which offers the basis for the representation of concrete entities (ARP; SMITH; SPEAR, 2015), and the GeoCore ontology, that derives the BFO perspective providing the definition of material entities that compound the main framework of geological substrates (GARCIA et al., 2020). These geological definitions provided by GeoCore include geological bodies, the earth materials that constitute them, their internal structure, the boundaries that separate one body from its adjacent bodies, and the types of contact that can be found between the bodies. We follow the same philosophical view and specialize these entities to offer a formal terminology defining the material endurants of reservoir geological occurrences. This aspect of our approach also allows us to envisage the GeoReservoir ontology integration with other ontologies that will be developed in the future extending GeoCore as well.

At this stage, the GeoReservoir ontology supports the description of any deep-marine channel and lobe system occurrence in all of its variations. To develop this ontology, we extensively reviewed the literature of deep-marine deposits and previous systematization projects of sedimentary deposit analysis. We also have counted on the support of experts with broad expertise in outcrop description and seismic interpretation for reservoir characterization in petroleum companies. In the future, the GeoReservoir ontology is intended to incorporate terminology from all depositional environments.

We orient our work through a philosophical view of considering Geology as a descriptive science, refusing, as much as possible, to adopt the hypothesis-driven approach that geologists use to analyze the geological occurrences. For example, Shanmugam

(2002) defines a turbidite as a sediment deposited by a turbidity current. Even though experts can simulate a turbidity current in research laboratories, it is not possible to see sediments and sedimentary rocks being formed by turbidity currents in nature because these processes take thousands or millions of years to happen. Geologists infer these processes based on their observations, adopting the most probable causes as the generating processes. Being so, we aim to capture the information as it is objectively described by the geologists independently of the sedimentary process interpretation, making it the most independent of tasks as possible. As there is greater agreement about the description than the interpretation, this descriptive focus favors integration and leverages data reuse potential.

As an initial research step to support future work such as a database of reservoir occurrence descriptions, this work has an exploratory focus. In other words, we do not validate any hypothesis here. The GeoReservoir ontology's actual benefits will take place in the future as applications will be developed using this ontology as their conceptual basis. Nevertheless, we probe the terminology's adequacy and truthfulness to reality by conducting a case study in which we describe a geological instance of deep-marine deposits using the GeoReservoir ontology. This instance, initially described by Hodgson et al. (2011), comprises two submarine channel-levee systems in the Karoo Basin that cropped out near Laingsburg, South Africa. It is used as an analogous in many deep-marine depositional system studies, configuring a suitable case for ontology application.

Besides building a domain ontology for deep-marine depositional systems, this work's inspiration is to show how ontological foundations can help us build high-quality conceptual models for a complex domain like Geology, which features an elevated level of terminological ambiguity and many language pitfalls. As conceptual modelers, we aspire to encourage the community to use ontologies in other complex modeling cases that arise in a wide variety of domains.

1.2 Thesis structure

This thesis is organized as follows:

- In Chapter 2, we describe the knowledge domain covered by this work in detail;
- In Chapter 3, we review other approaches to describe and integrate deep-marine depositional system knowledge and geological data in general;

- In Chapter 4, we study the necessary ontological foundations for this work's comprehension, as well as we summarize the BFO and GeoCore ontologies;
- In Chapter 5, we present the methods and tools that we used to develop the GeoReservoir ontology, as well as the ontology itself;
- In Chapter 6, we present the case study in which we applied the GeoReservoir to describe a part of a deep-marine depositional system;
- In Chapter 7, we discuss and validate our approach's benefits and limitations;
- In Chapter 8, we conclude this thesis and show some future work perspectives.

2 APPLICATION DOMAIN

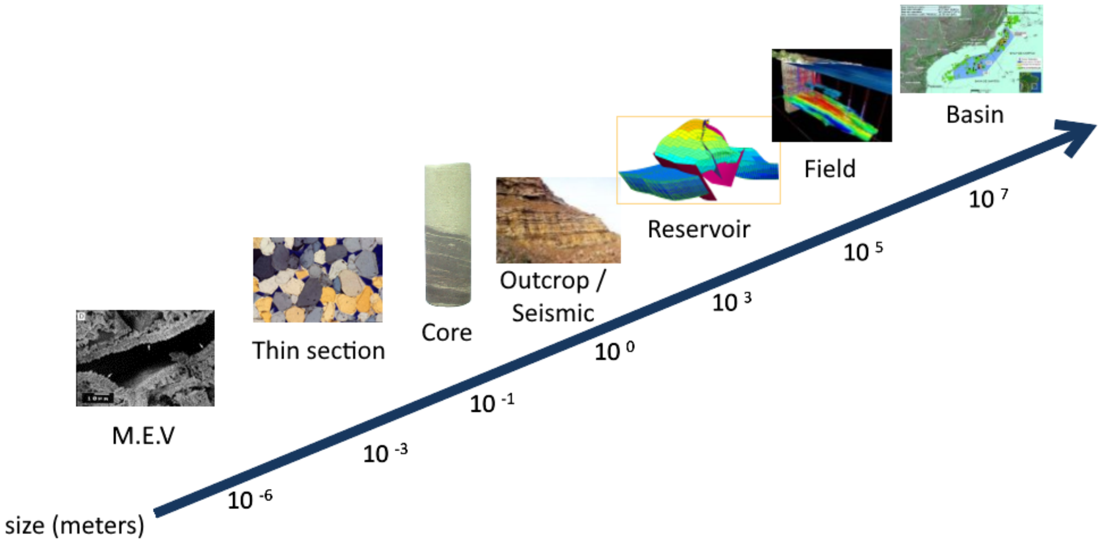
In this chapter, we describe the knowledge domain that this thesis covers. We present the main concepts that geologists use to describe deep-marine depositional systems. These concepts are scattered in several Geology branches, such as Sedimentology, Stratigraphy, and Petrology (NICHOLS, 2009). The knowledge in these branches is applied to depositional systems in general. However, here we describe the most common geological terms used to characterize deep-marine instances, keeping a descriptive focus instead of reviewing the sedimentary process interpretations available in the literature. Besides, this terminology covers the deep-marine deposit description at outcrop and seismic section scale, which constitutes this work's focus but can be extended in the future (Figure 2.1).

2.1 Deep-marine depositional systems

Deep-marine depositional systems (Figure 2.2) are groups of sedimentary bodies typically formed in continental slopes (ARNOTT, 2010). As mentioned before, the study of these deposits relies on several branches of Geology, such as Sedimentology, Stratigraphy, and Petrology. Sedimentology is the discipline concerned with studying the conditions in which the sediments and sedimentary rocks form, Stratigraphy is concerned with investigating the succession of sedimentary rock layers, and Petrology is concerned with examining the overall characteristics of rocks (NICHOLS, 2009). Another significant branch is Geophysics, as seismic imaging is essential due to the deposits' poor accessibility (ARNOTT, 2010).

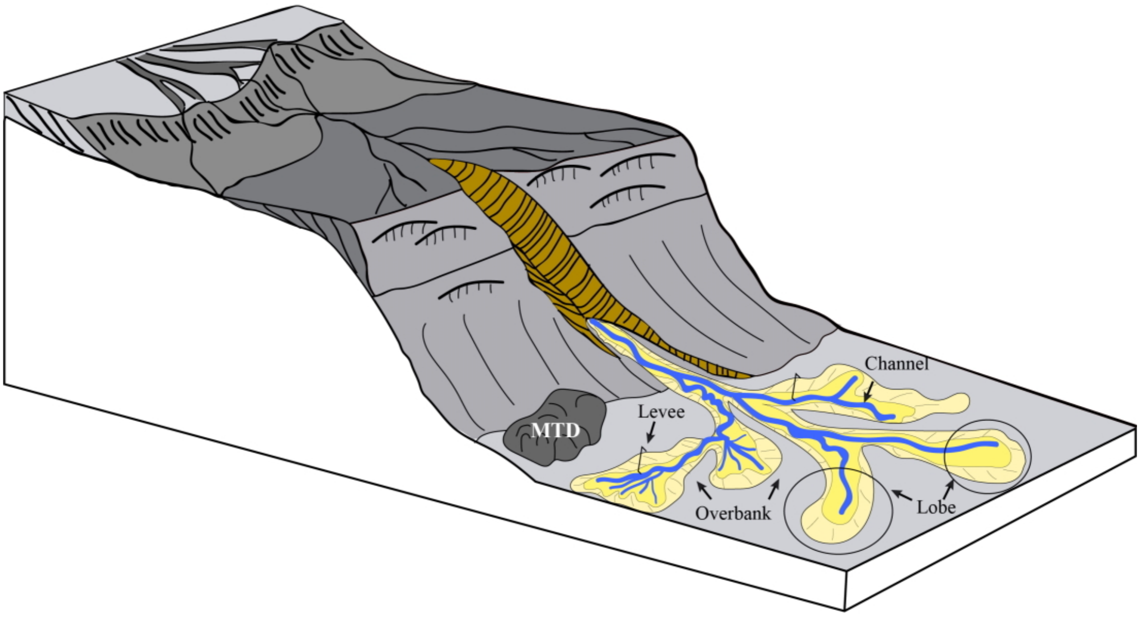
Geologists recognize and individualize sedimentary bodies according to their geometries and facies (MUTTI; NORMARK, 1987; PICKERING et al., 1995). Geometries are the three-dimensional shapes of sedimentary bodies whose identification requires the delimitation of bounding surfaces or unconformities, where facies change abruptly (PICKERING et al., 1995). Facies are combinations of features that bestow an aspect of the sedimentary body (WALKER; JAMES, 1992; DALRYMPLE, 2010). As these features usually have genetic significance, we focus on the facies' descriptive part. In this context, the features include sedimentary rock type, sedimentary structures such as planar or cross-stratification, bed thickness, and others (NICHOLS, 2009).

Figure 2.1 – An illustration of some observation scales in Geology. This work focuses on descriptions made in the outcrop and seismic section scale.



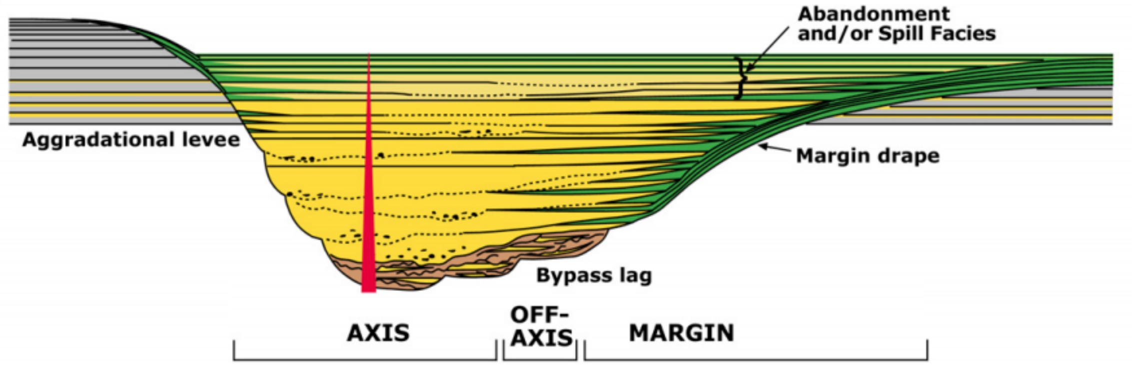
Source: modified from Della Favera (2001) and Jarna et al. (2015)

Figure 2.2 – An overview of deep-marine depositional systems. Labels indicate some of the most significant sedimentary bodies that compose these systems.



Source: modified from Posamentier and Walker (2006)

Figure 2.3 – A schematic representation of a channel with adjacent levees in a cross-section view. Yellow represents sand-rich sediments, green represents mud-rich sediments, brown represents mud-clast-rich sediments, and gray represents pre-existing sediments.



Source: modified from McHargue et al. (2011)

2.1.1 Sedimentary body types

Sedimentary bodies are classified based on the geometries and typical sedimentary facies found in deep-marine depositional systems. Some of the most common sedimentary body types are illustrated and labeled in Figure 2.2 and defined in the following.

A *channel* (Figure 2.3) is an elongated sedimentary body consisting of a concave-up basal surface and the sediments or sedimentary rocks that fill it (CULLIS et al., 2019). Individual channels feature dimensions that typically vary, with some uncertainty, from widths between 100 and 300 meters and thicknesses around 10 meters. These channels stack to form channel systems[1], which constitute one of the most common types of hydrocarbon reservoirs found in deep-marine sedimentary environments (MCHARGUE et al., 2011).

As depicted in Figure 2.3, channel fills are usually organized in three regions: axis, off-axis, and margin. In the literature, the aggregation of the sedimentary facies usually found in each region is called a “facies association”[2]. The axis association is found at the thickest part of the channel, has thicker beds than the other associations, and has the highest sand concentration. The margin association, on the other hand, has thinner beds and the lowest concentration of sand. The off-axis association has intermediate characteristics in these terms (MCHARGUE et al., 2011).

A *lobe* is a lobate-shaped sedimentary body having a lobate shape from above

---

[1] See Section 2.1.2. “Channel system” shall not be confounded with “depositional system”, which is the entire group of sedimentary bodies.

[2] “A group of sedimentary facies that are used to define a particular sedimentary environment.” Available at <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/facies-association-0> . Accessed at 29 dec. 2020.

and a lens-shaped cross-section, found at some laterals, terminus, or isolated (POSAMENTIER; WALKER, 2006). Lobes are much wider and less thick than channels. Such as occur with channels, the sediments that constitute lobes are organized in three facies associations: axis, off-axis, and fringe (cf. channel axis, off-axis, and margin).

A *levee* is a wedge or wing-shaped sedimentary body laterally adjacent to a channel. Generally, levees are found in the most distal parts of the channel systems, i.e., in the region closer to the lobes, which is considered the end of the depositional systems. In the proximal parts, channels are more erosional instead of leveed because they typically form inside canyon walls (see Figure 2.2). As the channels go down on the slopes, the levees become more prominent in a gradual way (ARNOTT, 2010; MCHARGUE et al., 2011).

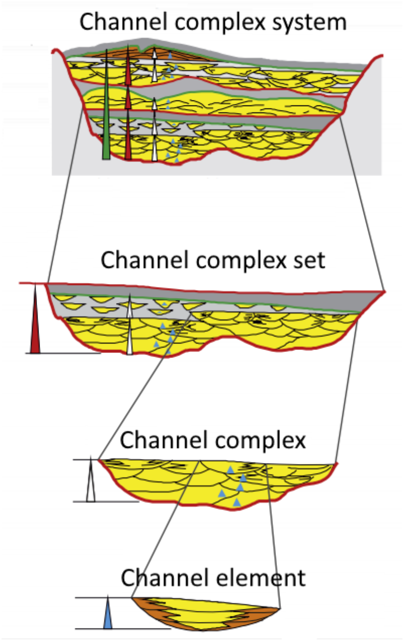
An *overbank deposit* is a sedimentary body with geometries such as lobate and waveform and found at the overbank area, i.e., a laterally-adjacent area around a channel. This type of sedimentary body includes sediment waves, crevasse splays, among other particular bodies. Crevasse splay is a lobate-shaped body that, differently from lobes, is found at the lateral of channel-levee complexes, starting from some point of rupture that passes through a levee. On its hand, a sediment wave is more isolated from its associated channel, as if the sediment flow that generated the channel fill is supposed to pass over the levee (POSAMENTIER; WALKER, 2006).

A *mass transport deposit (MTD)* is a sedimentary body with an irregular top surface, a chaotic internal structure, and constituted by sediments or sedimentary rocks that were rapidly remobilized and redeposited on the seafloor (POSAMENTIER; MARTINSEN, 2011). As these bodies can show a wide variety of geometries and lithological features, this term is more related to process interpretation than descriptive features. Nevertheless, MTD bodies are very relevant in deep-marine studies to provide information on hydrocarbon production potential (LE BOUTEILLER et al., 2019).

2.1.2 Hierarchical classifications

Channel and lobe bodies can be nested in increasingly larger bodies. Geologists assign hierarchical classifications to indicate the bodies' spatial order (SPRAGUE et al., 2005; CULLIS et al., 2018). In the geomorphological sense, these bodies are found stacked and nested in a fractal-like way, e.g., individual channels stack to form a complex and bigger body that also features a channel geometry (see Figure 2.4). There exists a wide variety of hierarchical classifications in the literature, and distinct classifications

Figure 2.4 – A hierarchical classification scheme for channels. This classification can also be analogously applied to lobes.



Source: Modified from Sprague et al. (2005) and Cullis et al. (2018)

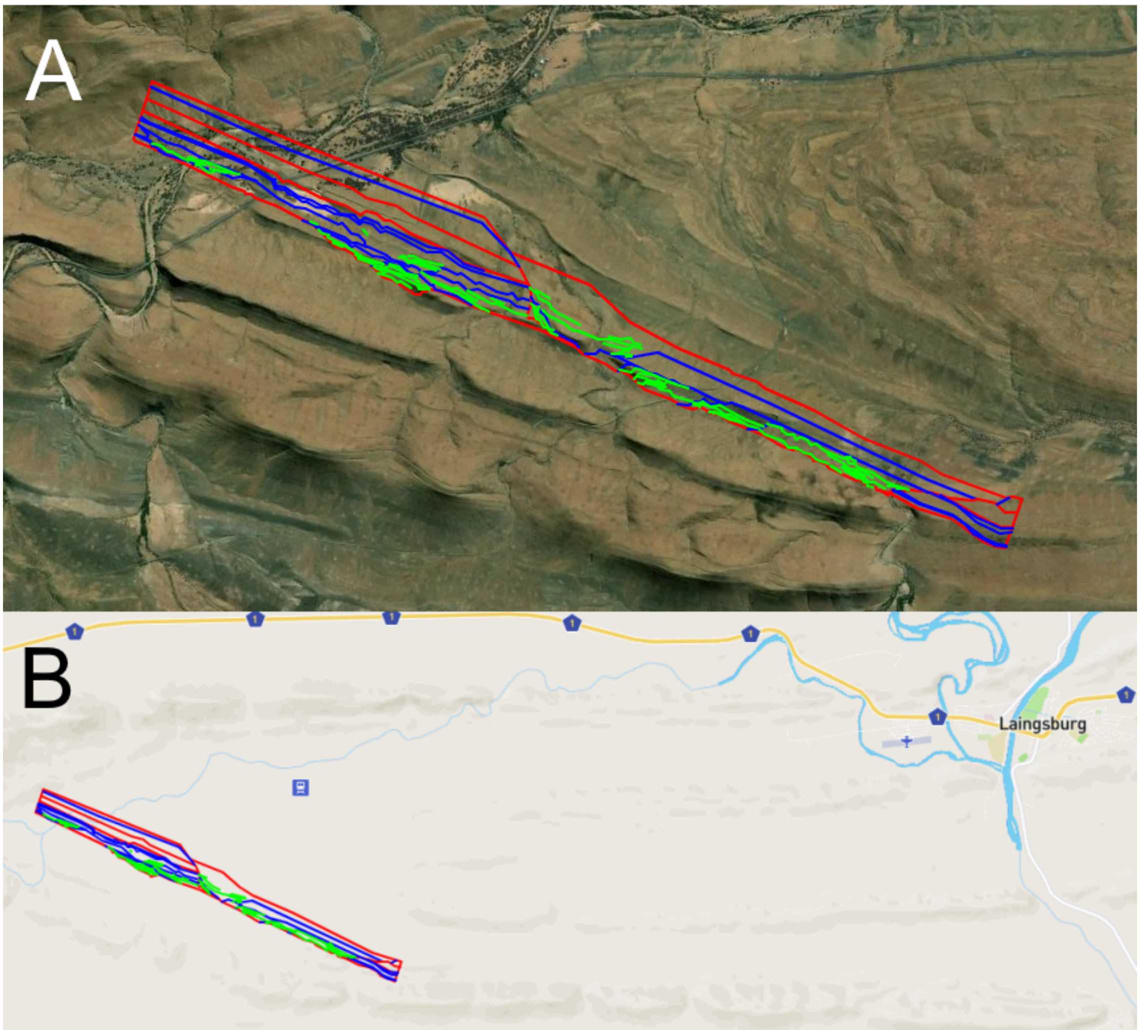
are not necessarily interchangeable because of variations in the criteria used to classify sedimentary bodies (CULLIS et al., 2018).

The classification depicted in Figure 2.4 is defined by Sprague et al. (2005) and amended by McHargue et al. (2011). It includes the following terms: channel/lobe element, channel/lobe complex, channel/lobe complex set, and channel/lobe complex system. In such a classification, a system is composed of at least two complex sets, which are composed of at least two complexes, which are composed of at least two elements. There exist spatially smaller hierarchical scales such as channel/lobe storey, bed, and lamina (SPRAGUE et al., 2005; MCHARGUE et al., 2011). However, we do not address hierarchical scales smaller than elements in this work because it is intended to support the description of sedimentary bodies at outcrop and seismic section scales, which cover from elements to complex systems. Nonetheless, one can extend this work in the future for description in more detailed scales.

2.2 An instance of deep-marine depositional system

The work of Hodgson et al. (2011) describes an outcrop of a deep-marine depositional system found in an area near the city of Laingsburg, South Africa (Figure 2.5).

Figure 2.5 – A drawing over satellite images representing the Karoo deposits outcrop (A) near Laingsburg, South Africa (B). The deposits' location is approximate.



Source: edited by the author using supplementary data from Hodgson et al. (2011) and Google Maps images <https://www.google.com/maps>.

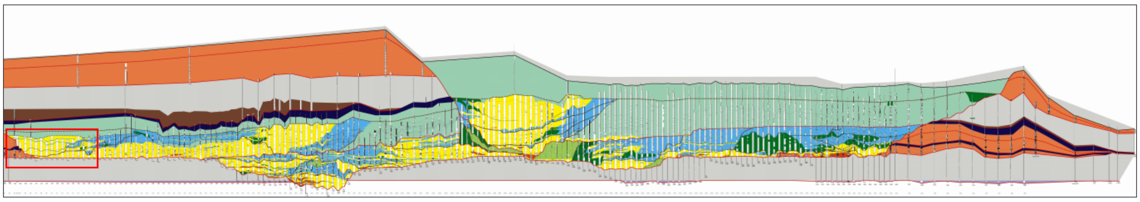
This outcrop features two seismic-scale channel-levee systems (named Units C and D) contained in the Karoo Supergroup, whose rocks cover almost two-thirds of the southern African land surface[3]. In the study of Hodgson et al. (2011), the authors present a correlation panel (Figure 2.6) built using data collected from more than two hundred measured sections and detailing the actual bodies' geomorphological and lithological features.

The Karoo Basin evolved from Permian to Triassic as a result of the subsidence of a retroarc region. In the Laingsburg depocentre, a 1.4 km thick Permian progradational basin floor to upper slope succession is well exposed and is being the object of several stratigraphic studies. The main lithostratigraphic units are: (1) close channel-

---

[3] Available at <https://en.wikipedia.org/wiki/Karoo_Supergroup>. Accessed at 10 jan. 2021.

Figure 2.6 – The correlation panel featuring the representation of the channel-levee systems described in this section. The detailed panel is available in the source. The red rectangle outlines the part of the channel-levee system that we use in this work's case study.



Source: Hodgson et al. (2011)

ized facies with cross-cutting erosion surfaces overlain by mudstone clast conglomerates, amalgamated fine-grained sandstones, mass-flow deposits (chaotic and folded strata), and thin-bedded ripple laminated beds that locally fine and thin onto erosion surfaces; and (2) non-channelized facies comprising mudstones, tabular, planar laminated siltstones, and climbing ripple laminated very fine and fine-grained sandstones. The basal datum for the sedimentary logs is a thin but extensive fine-grained sandstone unit overlain by mudstone. The whole site shows occurrences of a complex of channels, with the presence of an external levee and a lobe in the basal part (HODGSON et al., 2011).

Part of this data is used in this work in a case study to evaluate the GeoReservoir ontology adequacy and truthfulness (see Figure 2.6). We describe this subset with a higher detail level in Chapter 6.

2.3 Discussion about the domain terminology characteristics

It is a complex task to establish a standard reference to the knowledge in deep-marine depositional systems due to the variety of terminologies available in the literature, representing different views about the domain entities (ABEL; PERRIN; CARBONERA, 2015; CULLIS et al., 2018). As so, it is essential to distinguish and classify the terminology in two categories: descriptive and interpretative terms. Descriptive terms represent abstractions of visual and spatial aspects that geologists perceive in the sedimentary bodies, referring to the entities that exist when observed. Interpretative terms represent abstractions of processes that generate and modify these bodies, referring to entities that occurred or existed in the past (CARBONERA, 2012). An example of a descriptive term is "thin-bedded sandstone", which is based entirely on what the specialists observe in the geological field. On the other hand, "turbidity current" is an example of an interpretative term because geologists assume that some portion of rock was generated by a turbidity current by observing that portion's descriptive features, e.g., an instance of the Bouma sequence (BOUMA, 1962).

Such interpretations are hypothesis-driven and can evolve over time. This is why it is crucial to address the correct definition and storage of descriptive data when designing a geological information system. If some stored data is based on an interpretation, it is impossible to re-evaluate this interpretation when it becomes obsolete at some moment or when some geologist disagrees with it. On the other hand, storing the descriptive data in which an interpretation is based allows multiple evaluations. Such an approach leverages the data integration potential because there exists a higher level of agreement in descriptive knowledge than in interpretative knowledge, making it much more feasible to integrate descriptive data between different systems and geological reports. For these reasons, the present work focuses on separating and capturing the descriptive knowledge contained in the literature.

In deep-marine deposit studies, some terms have unclear connotations concerning this descriptive-versus-interpretative distinction. Terms such as "turbidite", "debrite", and "MTD" present both descriptive and interpretative meaning. When a geologist refers to some turbidite, he emphasizes the hypothetical fact that some portion of rock was generated by a turbidity current. Even though this hypothesis is endorsed by a set of descriptive features found in that portion of rock, the problem we address here is that these terms emphasize interpreted sedimentary processes instead of descriptive sedimentary characteristics.

Another relevant trait found in the domain terminology is the ambiguity presented by some terms. For example, the term "channel" sometimes refers to the conduits where sedimentary flows such as turbidity currents pass by and leave sediments that become channel fills (MUTTI; NORMARK, 1987). Other times, this same term refers to the channel fills and the surface where the sedimentation took place (i.e., their boundaries). Even though the two concepts are related, they refer to distinct entities in ontological terms. A computer application cannot deal with this ambiguity since the properties of the two entities differ.

Besides ambiguity, there exists a variety of distinct terms that refer to the same abstraction. For example, "channel fill" (SPRAGUE et al., 2005), "individual channel" (MORAES; BLASKOVSKI; PARAIZO, 2006), and "channel element" (MCHARGUE et al., 2011) refer to the same individual channelized sedimentary body that can stack to form more complex and bigger channelized bodies. This characteristic also represents a challenge for computer application development since software or ontology engineers can be misled to address the different terms as referring to distinct entities.

The terminological issues faced here were addressed by many works involving computation and data analysis in Geology, e.g., Carbonera (2012), Abel, Perrin and Carbonera (2015), and Garcia et al. (2020). The novelty in the present work is that we contextualize these problems in deep-marine deposit studies. To our knowledge, there does not exist a work that successfully addressed these problems in this domain at this thesis's publication date.

2.4 Final considerations

The application domain that is conceptually modeled in this work has many language pitfalls, such as the cases of ambiguity, the variety of classifications, the lack of a standard reference to the concepts, and the presence of interpretative meaning in the terminology. These particularities are an obstacle for data integration and make modeling these concepts a challenging task. This work is intended to ameliorate these issues by building the GeoReservoir ontology, which defines these concepts in a standard, unambiguous and formal way, supporting the description of deep-marine depositional systems instances and their further processing by computer applications.

3 LITERATURE REVIEW

Extensive research has been done on conceptual models for deep-marine depositional systems, Sedimentology, and Geology in general in the last two decades, aiming to achieve a uniform model for comparing reservoir data with the support of computer applications. In this chapter, we systematically review these conceptual models. We have selected the approaches that propose conceptual models for deep-marine depositional systems, especially turbidite deposits. We built our selection with studies in Stratigraphy, Sedimentology, and Informatics that focus on the systematic description of the deposits for further analysis by parametric search or artificial intelligence methods. These works present distinct levels of formalization of the published conceptual models, but they all contribute to raise and understand the main features that support reservoir geological interpretation.

We compared these approaches according to some comparison criteria defined according to the context and problem described in this thesis. The comparison criteria are listed in the following:

- *Focused on deep-marine*: indicates whether the model focuses on deep-marine channel and levee description. In the negative case, we consider that the model describes Geology in a broad sense.
- *Formal*: indicates whether the model is readable for computers. Formal models can be integrated into software applications and artificial intelligence models (STUDER; BENJAMINS; FENSEL, 1998).
- *Has public repository*: has at least one publicly available implementation persisted in some standard format such as OWL, UML, RDF, and others. This aspect is related with to intent to produce processable and reusable models.
- *Descriptive*: it is independent of geological process interpretation. To be considered descriptive, the model's geological material entity definitions (e.g., rocks, rock units, geological structures) should be comprehensible without needing to consult the definitions of the geological process that are supposed to generate or modify them. This distinction between descriptive and interpretative knowledge was discussed in Section 2.3.
- *Analyzes metaproperties*: the model analyzes at least one of the ontological metaproperties mentioned in Chapter 4. Ontological metaproperty analysis is essential to

Table 3.1 – Literature comparative analysis.

Source	**Focused on deep-marine**	**Has public repository**	**Formal**	**Descriptive**	**Analyzes meta-properties**	**Aligned with a foundational ontology**
SWEET (2005)	No	Yes	Yes	No	No	No
Sprague et al. (2005)	Yes	No	No	No	No	No
Moraes et al. (2006)	Yes	No	No	No	No	No
GeoSciML (2006)	No	Yes	Yes	Yes	No	No
Lorenzatti et al. (2010)	No	No	Yes	Yes	Yes	Yes
McHargue et al. (2011)	Yes	No	No	No	No	No
Cullis et al. (2019)	Yes	No	Yes	No	No	No
Le Bouteiller et al. (2019)	Yes	No	Yes	Yes	No	No
Qu (2020)	Yes	No	Yes	No	Yes	Yes

Source: the author

classify the entities according to what they are in reality and not to the geologists' point of view (GARCIA et al., 2020).

- *Aligned with foundational ontology*: the model is aligned with some foundational ontology such as BFO, UFO (GUIZZARDI, 2005), and others. Foundational ontology alignment eases integration with other ontologies and draws the benefits from the conceptual and philosophical basis to the domain ontology (GUARINO; OBERLE; STAAB, 2009; ARP; SMITH; SPEAR, 2015).

We list this comparison in Table 3.1 and detail each work in the following sections, starting from the most relevant models for this thesis's objectives.

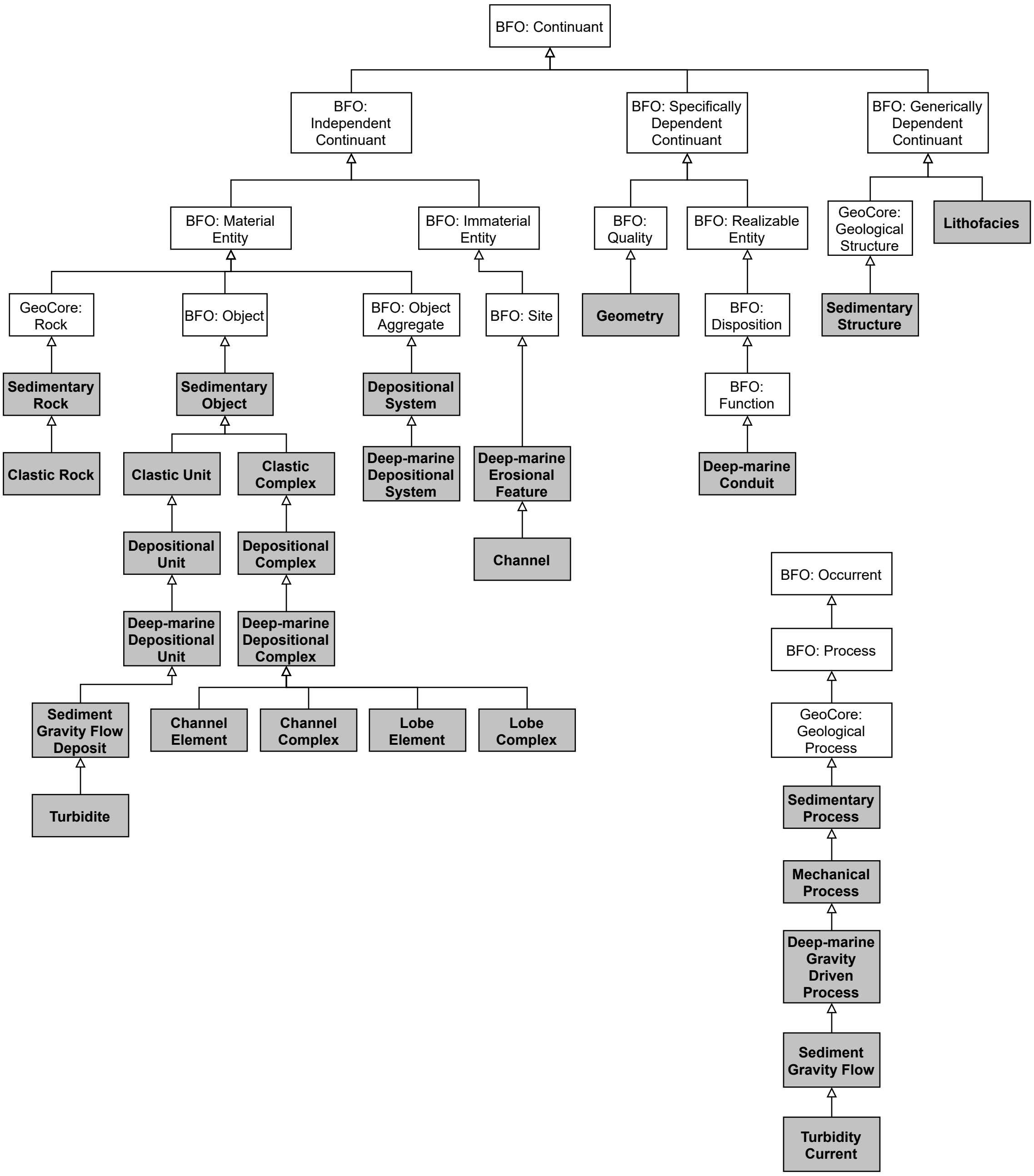
3.1 An ontology for deep-marine depositional system interpretation

The master's thesis in Geology presented by Qu (2020), which was developed with our collaboration, proposes an ontology that formalizes the deep-marine depositional system concepts required for computer-assisted geological process interpretation. This ontology features a systematic definition of the geological descriptions and interpreted sedimentary processes in deep-marine settings, resting on the ontological foundations of the BFO, GeoCore, and IAO ontologies, as well as taking advantage of the ontological metaproperty analysis. Qu's work intends to identify what are the required descriptions to support sedimentary process interpretation and formally defines the entities needed for the geological description.

Qu divides his ontology (Figure 3.1) into two main sections: general clastic depositional system and deep-marine clastic depositional system. The former is intended to be applicable to any kind of depositional system, such as deep-marine, aeolian, shallow-marine, fluvial, and deltaic, while the latter is focused on deep-marine. For all concepts' complete definition, one should refer to Qu (2020). The GeoCore and BFO concepts mentioned in the following are reviewed in detail in Chapter 4. In our work's perspective, some of the most relevant concepts in the general part of Qu's ontology are:

- *Sedimentary Rock*: a GeoCore Rock constituted by a collection of grains and generated by some Sedimentary Process;
- *Clastic Rock*: a Sedimentary Rock constituted by a collection of grains originated from a pre-existing Rock and generated by some Mechanical Process;
- *Sedimentary Object*: a GeoCore Geological Object constituted by some Sedimentary Rock and generated by some Sedimentary Process;
- *Clastic Unit*: a Sedimentary Object constituted by some Clastic Rock;
- *Clastic Complex*: a Sedimentary Object composed of more than one Clastic Units;
- *Depositional Unit*: a Clastic Unit that is a member of a Depositional System;
- *Depositional Complex*: a Clastic Complex composed of more than one Depositional Units;
- *Depositional System*: a BFO Object Aggregate whose members are genetically-related Depositional Units;
- *Sedimentary Structure*: a GeoCore Geological Structure dependent on some Clastic Unit that is the pattern of the internal arrangement of that unit and is generated by some Sedimentary Process;
- *Lithofacies*: a BFO Generically Dependent Continuant dependent on some Clastic Unit and concretized as multiple BFO Qualities of that unit and the GeoCore Rock that constitutes it, such as Grain Size, Sorting, and Angularity;

Figure 3.1 – Diagram depicting some of the most relevant concepts from the ontology of Qu (2020). White boxes represent BFO and GeoCore concepts, while gray boxes represent the concepts defined by the ontology. Arrows represent “subsumed by” (i.e., taxonomical) relations.



Source: modified from Qu (2020)

- *Geometry*: a BFO Quality of some GeoCore Geological Object that describes its external three-dimensional form;
- *Sedimentary Process*: a GeoCore Geological Process that generates some Sedimentary Rock, involving the deposition of a collection of grains from a pre-existing rock, the cementation of mineral or organic particles, and the subsequent lithification;
- *Mechanical Process*: a Sedimentary Process which erodes, transports, deposits, and compacts a collection of grains from some pre-existing Rock, and having erosional and depositional processes as temporal parts.

The deep-marine clastic depositional system, in its turn, features the following concepts among the most relevant ones:

- *Deep-marine Depositional Unit*: a Depositional Unit that is a member of a Deep-marine Depositional System;
- *Deep-marine Depositional Complex*: a Depositional Complex composed of more than one Deep-marine Depositional Units;
- *Deep-marine Depositional System*: a Depositional System located in the deep-marine environment, which is over 200 m deep from the continental shelf edge down to the basin floor;
- *Deep-marine Erosional Feature*: a BFO Site confined by some GeoCore Rock, having the function of a Deep-marine Conduit and generated by some Deep-marine Gravity Driven Process;
- *Sediment Gravity Flow Deposit*: a Deep-marine Depositional Unit generated by some Sediment Gravity Flow;
- *Turbidite*: a Sediment Gravity Flow Deposit generated by some Turbidity Current;
- *Channel*: a Deep-marine Erosional Feature located on the relatively moderate-gradient continental slope or basin floor and characterized by a concave-up shape in transverse profile;
- *Deep-marine Conduit*: a BFO Function of being a conduit for some Deep-marine Gravity Driven Process and transporting a collection of grains;
- *Channel Element*: a Deep-marine Depositional Complex with steady aggradation, located in some Channel, and having a Geometry with “channel” nominal value;
- *Channel Complex*: a Deep-marine Depositional Complex located in a Channel and composed of more than one Channel Element;

- *Lobe Element*: a Deep-marine Depositional Complex with steady aggradation, generated by some Deep-marine Gravity Driven Process;
- *Lobe Complex*: a Deep-marine Depositional Complex composed of more than one Lobe Element;
- *Deep-marine Gravity Driven Process*: a Mechanical Process that happens at under 200 m of depth below the sea surface, driven by gravity, and having some water, Clastic Rock, and a collection of grains as participants;
- *Sediment Gravity Flow*: a Deep-marine Gravity Driven Process that generates some Sediment Gravity Flow Deposit;
- *Turbidity Current*: a non-cohesive, turbulent, and Newtonian Sediment Gravity Flow that generates some Turbidite.

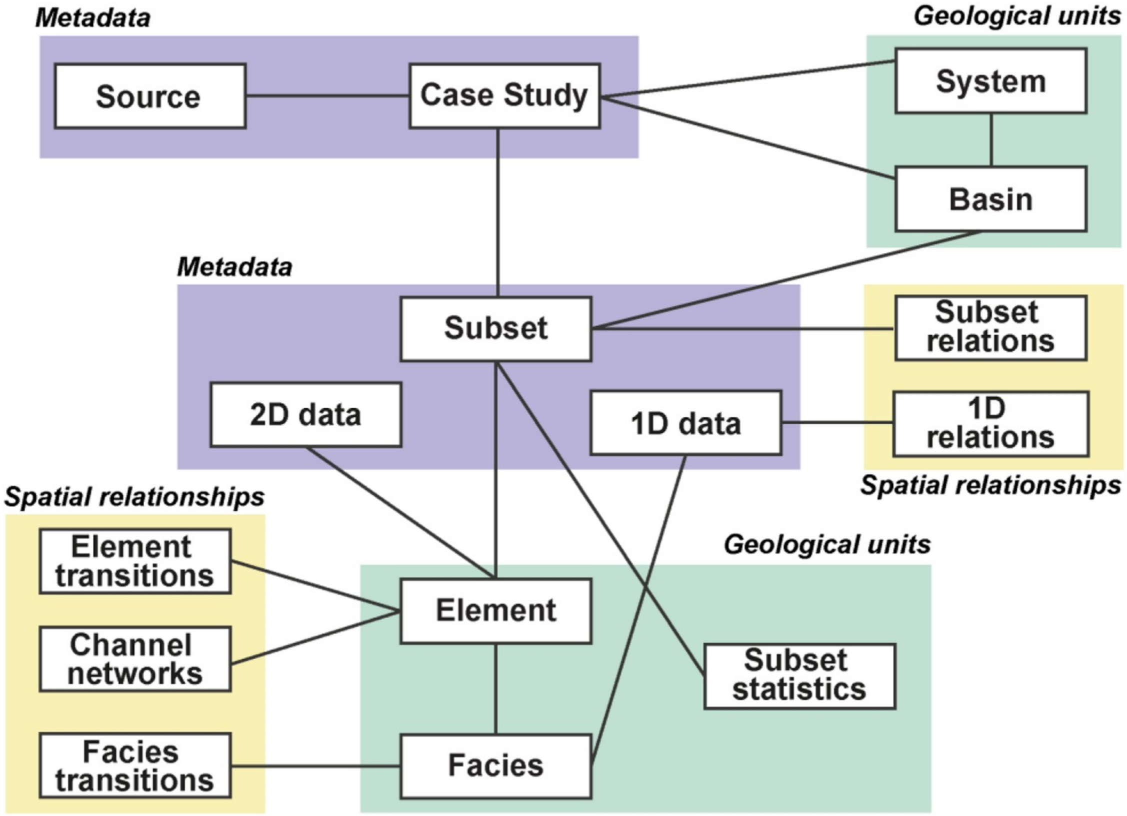
Qu's work offers a systematic view of the literature about deep-marine depositional systems, formalized and analyzed under the same ontological foundations that we use to build the GeoReservoir ontology. Nevertheless, it is not descriptive under this thesis's sense because many definitions of entities observed in the geological field have both descriptive and interpretative sense, e.g., Turbidite. Being so, in our work, we analyze these descriptive features further and define them independently of the relations with the sedimentary processes and genetic connotation as much as possible.

3.2 A database for quantitative analysis of deep-marine instances

The work of Cullis et al. (2019) describes a relational database for deep-marine occurrence quantitative analysis. As such, its relational schema (Figure 3.2) has an underlying conceptual model, which is also described in the paper. The work's main objective is to demonstrate how the integration of deep-marine deposit data can allow the application of meaningful queries to discover new knowledge about these depositional systems. It is an extension of another approach initially proposed by Baas, McCaffrey and Knipe (2005), which also addresses integration for deep-marine deposit data. The works of Cullis et al. (2019) and Baas, McCaffrey and Knipe (2005) have the same ultimate goal that our thesis presents: capturing descriptions of geological occurrences in a systematic way in order to support computer applications to query, extract patterns, and detect analogous occurrences.

The authors divide the data schema (Figure 3.2) into three conceptual categories:

Figure 3.2 – Representation of the relational database schema proposed by Cullis et al. (2019). Boxes represent tables, connecting lines represent relationships, and colored areas represent the three conceptual categories of data described by the authors: geological units, spatial relationships, and metadata.



Source: Cullis et al. (2019)

data on geological units, data on spatial relationships between geological units, and associated metadata. The first category comprises systems, basins, elements, and facies. A system is defined as a unit that extends itself from the slope-break to the most distal point of deposition. An element is a unit with a distinct architectural or geomorphological expression, reflecting a particular suite of processes occurring in a specific deep-marine sub-environment. Elements are characterized by features like geometry, sinuosity, paleoflow, gradient, age, and net-to-gross ratio, and are further classified by general types such as channel, levee, lateral splay, terminal deposit, and MTD. Facies are the smallest units in the database, each one contained in a single element and characterized by lithological and textural features. Although not conceptually defined in the work, a basin is characterized by features like tectonic setting, mechanisms of formation, geological evolution, and physiography. When data cannot be associated with an individual element, it is stored in the "Subset statistics" table.

The category of spatial relationships comprises element transitions, facies transitions, channel networks, subset relation, and 1D relations. An element transition is a spatial relationship between elements, while a facies transition is a spatial relationship between facies. Both types of transitions relate units whose boundaries or gradational changes are in contact. A channel network is a particular kind of transition between channel elements bounded by points of avulsion or branching. Subset relations and 1D relations are transitions between study areas, which are included in the last conceptual category.

This last category, namely metadata, includes sources, case studies, subsets, 2D data, and 1D data. A source is a record that describes the published or unpublished work that originated the associated data. A case study refers to a system or a portion thereof, which has been the subject of study by one or more groups of authors. A subset is a part of a case study distinguished on the type of information it provides and can be linked to specific metadata when it is sourced by a 2D (e.g., cross-section) or 1D (e.g., well log) dataset. As stated above, subsets can be spatially related through the "subset relation" and "1D relation" entities in the spatial relationships category.

Although formal, some of the database's conceptual definitions are not descriptive since they are vague or linked to process definitions that are not addressed. For example, the authors define an "element" as "a geological unit with a distinct architectural or geomorphological expression, which reflects a particular suite of processes occurring in a specific deep-marine sub-environment" (CULLIS et al., 2019). The model is not intended to be an ontology, so no ontological analysis or foundational ontology alignment is presented. Consequently, the ambiguity and vagueness of the original vocabulary affect the use case descriptions and the queries. In our analysis, this model has further conceptual issues, such as a "channel network" being presented as a type of "transition", which are clearly distinct entities. The presentation lacks cardinality definitions and an exact list of attributes that compose each table, which cannot be accessed because the data model is not publicly available. In summary, further formal definition and disambiguation are required for our work's purposes.

3.3 A knowledge model for description and intepretation of mass transport deposits (MTD)

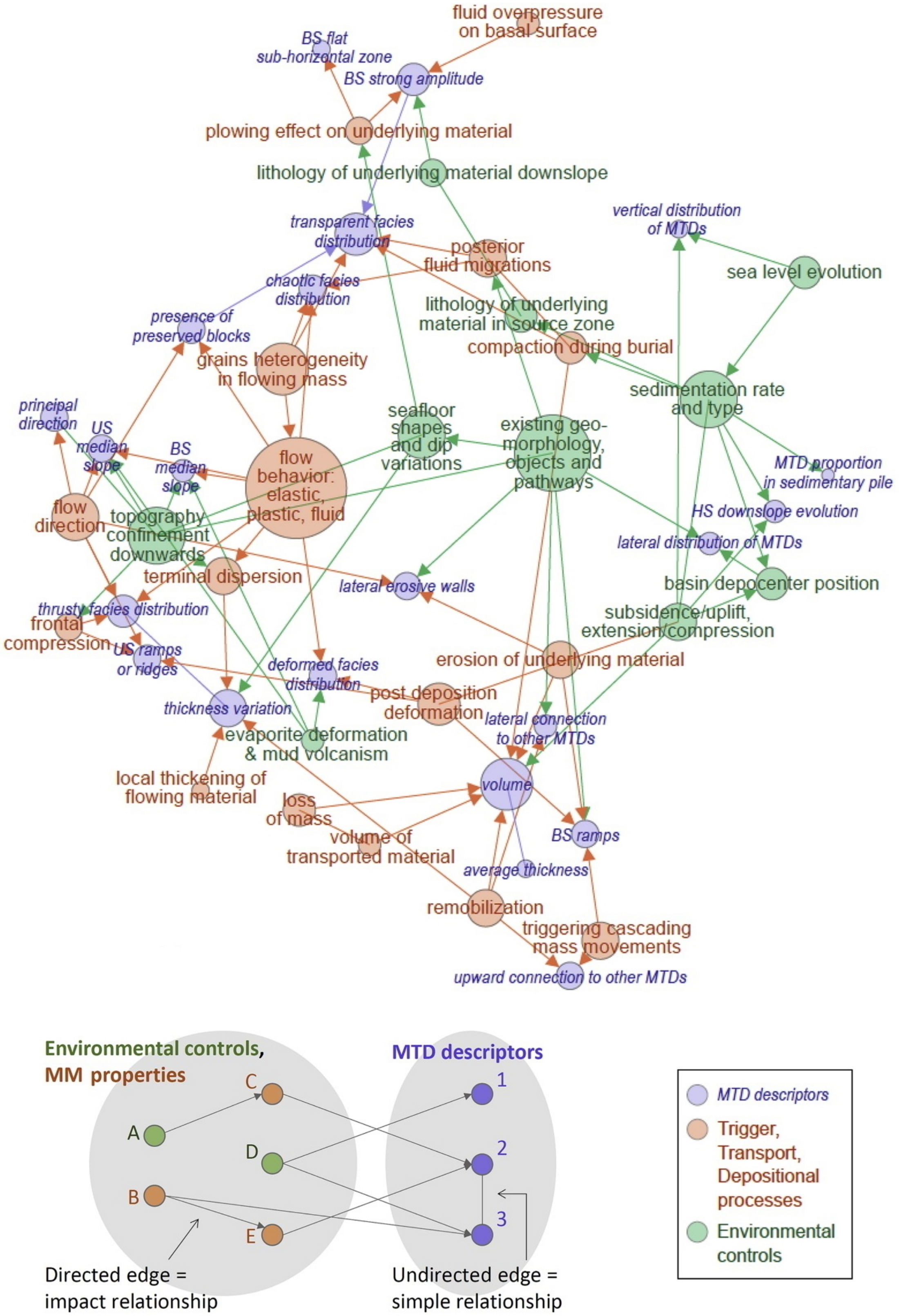
In the approach described by Le Bouteiller et al. (2019), the authors build a knowledge base, structured as a graph, to support the description of MTD bodies and the interpretation of mass transport processes. In this graph (Figure 3.3), nodes represent MTD descriptors, depositional process features, and environmental controls. Directed edges link pairs of causally related nodes, e.g., a process to a descriptor in which the process has some impact or influence. Undirected edges link pairs of related nodes in which the nodes mutually influence themselves.

Descriptors are dimensions that compose MTD properties. The knowledge base features seven MTD properties: morphology, basal surface, upper surface, position, headscarp, internal facies distributions, and global environment. For example, "volume" is a descriptor of the "morphology" property. Identically, the depositional process features compose depositional process properties. The depositional properties are: trigger phase, transport phase, and deposition phase. Environmental controls are considered larger-scale processes associated with the regional or global environment, e.g., "sea-level evolution".

Along with the knowledge graph, the authors propose an interpretation methodology relying on the graph. In this methodology, they first characterize the MTD units in terms of the properties described in the graph. Then, based on the edges, they look for possible causes or explanations for each descriptor's value. Lastly, for each possible cause, they evaluate their certainty based on the number of pointing to the cause. Therefore, the authors specify that the results are not necessarily interpretations, possibly being hypotheses depending on the case.

The work of Le Bouteiller et al. (2019) shows several limitations concerning the

Figure 3.3 – Representation of the knowledge graph proposed by Le Bouteiller et al. (2019), where dots correspond to nodes, arrows correspond to directed edges, and lines correspond to undirected edges.



Source: Le Bouteiller et al. (2019)

maintenance of the knowledge model and further integration with other geological models and applications. The natural lack of structure of the graph formalism and no ontological foundation in the conceptual model building guarantee a high grade of freedom in the model design and evolution. However, the drawback is the difficulty in keeping consistency and avoiding semantic redundancy on the model evolution, both of which are qualities aimed by this thesis's ontological approach. Besides that, the descriptive capture of MTD features is not the main interest of Le Bouteiller and colleagues' work. Therefore, their work does not detail the MTD features' meaning and focuses on MTD units, while our work is concerned with deep-marine channels, levees, lobes, and mounds in general. Nevertheless, the study domain in the work of Le Bouteiller et al. (2019) intersects with the set of depositional unit geometrical and lithological properties described in this thesis.

3.4 Conceptual models for deep-marine recurring features description

The non-formal conceptual models proposed by Sprague et al. (2005), Moraes, Blaskovski and Paraizo (2006), and McHargue et al. (2011) are relevant sources of information about deep-marine deposit recurring features like geometries, architectures, hierarchical classifications, stacking patterns, and lithological characters. There exists a wide variety of these models in the literature, which was extensively reviewed by Cullis et al. (2018). Therefore, we selected these three works for having the most practical application for the GeoReservoir ontology goals.

The three models establish a hierarchical framework for channelized deposits (Figure 3.4), such as the hierarchical classification presented in Chapter 2. This hierarchical framework consists of the orders: (1) channel fill (SPRAGUE et al., 2005), individual channel (MORAES; BLASKOVSKI; PARAIZO, 2006), or channel element (MCHARGUE et al., 2011); (2) channel complex (SPRAGUE et al., 2005; MCHARGUE et al., 2011) or composite channel (MORAES; BLASKOVSKI; PARAIZO, 2006); (3) channel complex set (SPRAGUE et al., 2005; MCHARGUE et al., 2011) or channel complex (MORAES; BLASKOVSKI; PARAIZO, 2006); and (4) channel complex system (SPRAGUE et al., 2005). According to Sprague et al. (2005), this hierarchy can be analogously applied to lobes, consisting of lobe, lobe complex, lobe complex set, and lobe complex system. Channel elements (order 1) are the volumes of sediment deposited in single channel filling and abandonment cycles. Channel complexes (order 2) are composed of genetically related channel elements having a similar architectural style. The subsequent orders (orders 3 and 4) are composed of genetically related units of lower order. Levees are further classified into two categories: inner levee, which is often adjacent to channel elements (order 1), and outer levee, which is often adjacent to channel complex sets (order 3).

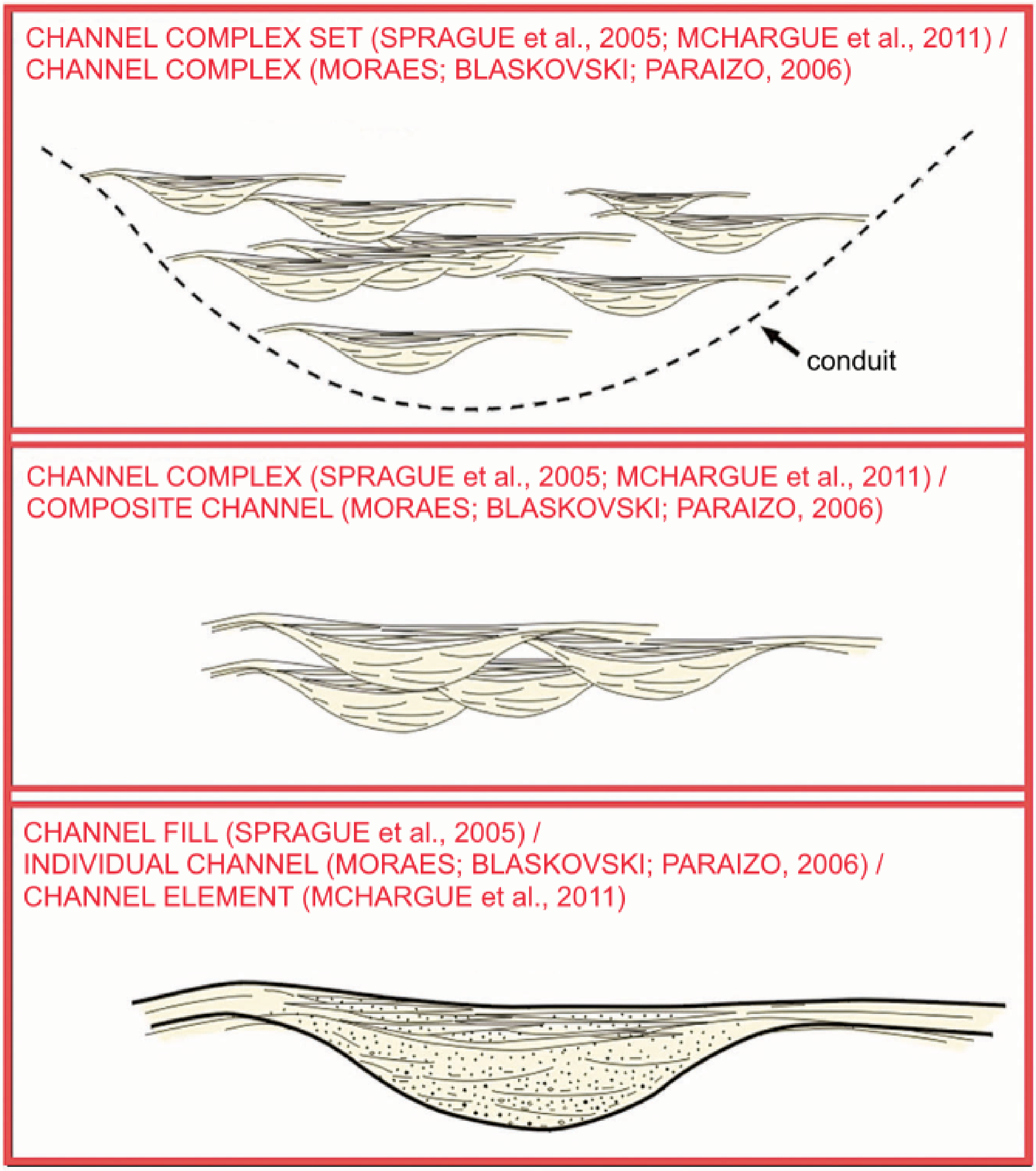
The models also establish architectural patterns according to the deposits' relative position to the continental slope: the most proximal channel complexes are confined while the most distal are distributary, i.e., as the channels get father from the slope break, they gradually lose their lateral walls or levees while they become wider and thinner until they reach the terminal lobes. Sprague et al. (2005) classify the most confined channels as "bypass" and "leveed". "Bypass" refers to the erosional nature of some channels, in conformance with Arnott (2010), whose lateral confinement mostly consists of erosional excavation at the surface. "Leveed" refers to the presence of wedge-shaped sedimentary units at the channel laterals, as the depositional channels described by Arnott (2010). Proximal channels can be erosional, leveed, or both.

Regarding facies, the models organize and describe the most common facies found in deep-marine depositional systems. For example, the most confined channels usually contain thin-bedded and coarse-grained sedimentary rocks, which constitute lower quality reservoirs, while the less confined channels usually contain sandy rocks with a high net-to-gross ratio (SPRAGUE et al., 2005).

These models are used by McHargue et al. (2011) as the conceptual bases for the assertion of "rules", which are observations and hypotheses of events that generate and modify deep-marine depositional systems. Based on these rules, the authors construct reproducible forward models that simulate the reservoirs' evolution, enhancing the geologists' understanding of deep-marine deposit geometry and architecture.

This category of conceptual models is crucial for the geological studies and knowledge dissemination about deep-marine depositional systems and Sedimentology in general. However, for our work's purposes, these models lack formality and have many of the conceptual issues discussed in Section 2.3. They are not descriptive in this work's sense; e.g., a channel element is defined as a "volume of sediment deposited in a single cycle of channel filling and abandonment" (SPRAGUE et al., 2005). Such a definition requires further knowledge of what a "cycle of channel filling and abandonment" is. As a geological process, this kind of cycle is interpreted and not observed. As discussed throughout our thesis, descriptions that merge observation and interpretation create problems like contaminating the data with subjective hypotheses, hiding the descriptive data

Figure 3.4 – Hierarchical classification for channellized deposits according to Sprague et al. (2005), Moraes, Blaskovski and Paraizo (2006), and McHargue et al. (2011).



Source: modified from Moraes, Blaskovski and Paraizo (2006) that allowed the interpretation, and limiting the integration of data described under different hypotheses.

The lack of formality also demands an amount of effort from both experts and non-specialists to understand the conceptual basis behind the geological definitions and hampers the integration of data from different sources in computer applications. As such, these models require further ontological analysis for this work's objectives.

It is also essential to notice that predictive definitions such as "the most confined channels usually contain thin-bedded and coarse-grained sedimentary rocks, which constitute lower quality reservoirs" (SPRAGUE et al., 2005) comprise epistemological definitions instead of ontological ones. In our thesis, we focus on building ontological definitions, which will serve as the basis for future predictive models and interpretations.

3.5 An ontology for the description of visual sedimentological properties

The work of Lorenzatti et al. (2010) describes an ontology for the representation of visual knowledge in the Sedimentary Stratigraphy domain. Their work assumes that geologists rely on their visual perceptions to describe geological occurrences and that visual knowledge representation can support the domain documentation, communication, and construction of knowledge-based systems. The ontology defines a set of primitives that combine visual and textual descriptions and formalizes them using the UFO ontology's conceptual basis (GUIZZARDI, 2005).

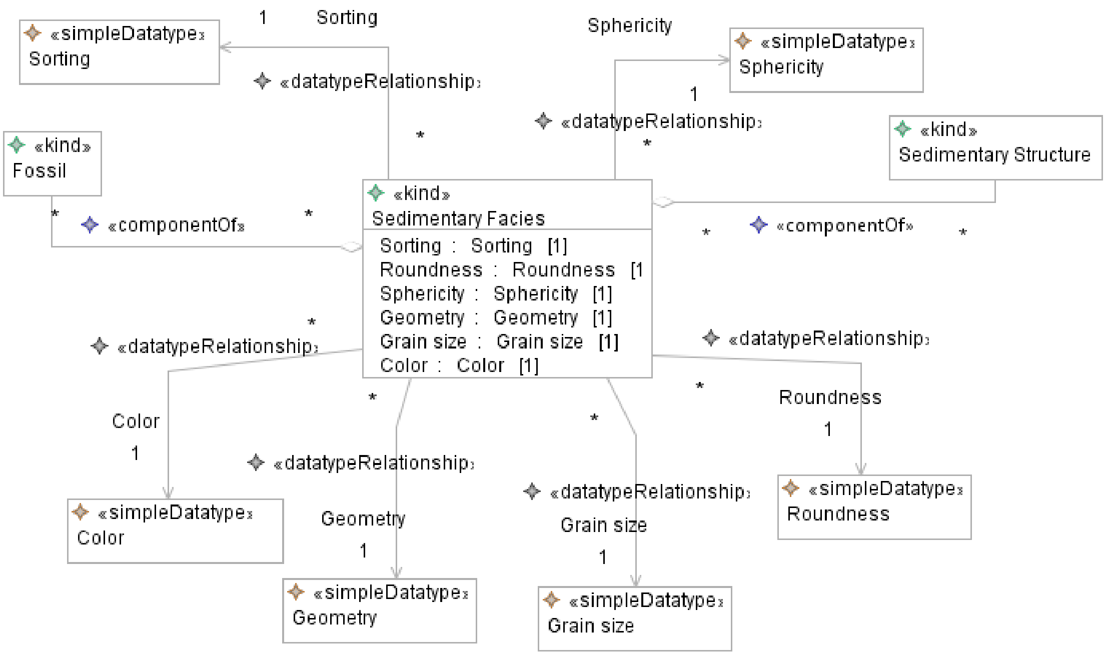
In the ontology proposed by the authors (Figure 3.5), a *Sedimentary Facies* is a rigid sortal with own identity[1] that groups together a set of a sedimentary rock's visual aspects, which are strongly connected with the depositional conditions in which the rock was created (LORENZATTI et al., 2010). A Sedimentary Facies is characterized by a set of qualities, namely Sorting, Roundness, Sphericity, Geometry, Grain Size, and Color. Although not necessary, a Sedimentary Facies can also be composed of some *Sedimentary Structure* (i.e., the visual aspect of some internal spatial arrangement of the rock grains) and some *Fossil*.

As such, the ontology proposed by Lorenzatti et al. (2010) works with the notion that a sedimentary facies is the sum of the visual aspects of *one* rock body. Although this is a descriptive view, in the GeoReservoir ontology, we work with a different notion: a sedimentary facies is a pattern concretized in one or many sedimentary geological objects

---

[1] We review the meanings of rigidity and identity in Chapter 4.

Figure 3.5 – An illustration of the Sedimentary Facies concept as defined by Lorenzatti et al. (2010) with its relations with the datatypes, the Fossil concept, and the Sedimentary Structure concept.



Source: Lorenzatti et al. (2010)

constituted by some sedimentary rock thereof (see Chapter 5). Nonetheless, the sedimentary facies ontology of Lorenzatti et al. (2010) features foundational ontology alignment and ontological metaproperty analysis, which ensures the evaluation of the ontological nature of the entities being modeled, e.g., separating the visual aspects perceived by the geologists from the actual rock bodies. This ontology misses core ontology alignment, which we address in this thesis, enabling these concepts' further integration with other geological ontologies.

3.6 Models for general Geology and Geosciences description

SWEET (RASKIN; PAN, 2005) is an ontology network that aims to represent the full domain of Earth Sciences. The SWEET ontologies include an extensive collection of terms, which are divided into *facets*. Each *facet* features a set of ontologies concerning a specific sub-field of Earth Sciences. The facets are described as follows (RASKIN; PAN, 2005):

- *EarthRealm*: concepts about the spheres of the Earth, e.g., atmosphere, lithosphere, and hydrosphere;

- *NonLivingSubstances*: non-living building blocks of nature such as particles, electromagnetic radiation, and chemical compounds;
- *LivingSubstances*: plants and animal species, i.e., the biosphere;
- *PhysicalProcesses*: processes that affect living and non-living substances, such as diffusion and evaporation;
- *PhysicalProperties*: properties that apply to living and non-living substances, such as temperature, pressure, and height;
- *Units*: measurement units such as kilometer;
- *Time*: temporal extents, such as duration, season, and century, and temporal relations, such as after and before;
- *Space*: spatial extents, such as country and equator, and spatial relations, such as above and north of;
- *Numerics*: numerical extents, such as interval and point, and numerical relations, such as greater than;
- *PhysicalPhenomena*: phenomena associated with concepts from other facets (time, space, living and non-living substances) such as hurricane, earthquake, and volcanism;
- *HumanActivities*: activities that humans engage in, such as commerce and fishery, regarding their impact in nature;
- *Data*: dataset concepts such as representation, storage, and modeling.

The entire ontology network is available in a public Web repository[2]. For our work's interest, the most relevant *facet* is *NonLivingSubstances*, which includes rocks, their chemical properties, and associated processes. The SWEET ontologies are integrated between themselves, e.g., rock is related to the solid state concept, which is used by other ontologies. However, further integration with other domains using foundational ontologies is not addressed in this approach. The ontologies lack definitions that could explain and clarify each concept's meaning, as it defines only the formal relations between the concepts. They also do not feature geological unit concepts such as "sedimentary unit", which play a central role in the GeoReservoir ontology's definitions.

The GeoSciML (SIMONS et al., 2006; RICHARD et al., 2007) is a data model and data exchange format intended to support the building and integration of geological maps, containing geological descriptions attached to geospatial referencing. These

---

[2] <https://github.com/ESIPFed/sweet>

geological descriptions include instances of units and rocks with their compositions and structures such as folds and foliations. The data format is available as an XML schema and documented as a UML model[3]. The model evolved since its first release during the year 2005 and influenced important integration standards such as the RESQML of Energistics[4] (KING et al., 2012). The latest version at this thesis's publication date is GeoSciML 4.1, published in 2017.

GeoSciML is divided into a set of packages, which are defined as follows:

- *GeoSciMLBasic*: represents fundamental geological map features and the relations between them;
- *GeologicTime*: describes geological periods, boundaries, and the relationships between them;
- *GeoSciMLExtension*: provides further descriptions of basic classes by adding more properties and relations;
- *Borehole*: describes boreholes and related artifacts;
- *LaboratoryAnalysis-Specimen*: describes processes and results related to the analysis of geological samples using instruments;
- *GeoSciMLLite*: a separated package that provides a simplified view of the other packages, organized as a set of flat tables.

Some of the most relevant classes for this thesis are specified as follows. A *GeologicFeature* is a conceptual feature that represents some particular Earth phenomenon or observation thereof. It is formally related to a *MappedFeature*, i.e., an instance that carries a geometry or shape and holds a geospatial reference. A *GeologicUnit* is a *GeologicFeature* representing a body of some *EarthMaterial*, which represents a naturally occurring substance in the Earth, independent of quantity and location. A *GeologicStructure* is a *GeologicFeature* that describes a configuration of matter in the Earth based on a describable heterogeneity or pattern, such as a contact, a foliation, and a fault. *Description* classes represent further details about many of these entities and are formally related to the classes they help describe.

Compared to SWEET in this thesis's terms, GeoSciML features a good clarification on the concepts that it defines. As occur with SWEET, GeoSciML also lacks ontological metaproperty analysis and foundational ontology alignment, which would as-

---

[3] <http://geosciml.org/doc/geosciml/4.1/documentation/html>

[4] <https://www.energistics.org/resqml-standards/>

sist the integration with other ontologies and domains. As a core model for Geology, GeoSciML's issues are discussed in more detail by Garcia et al. (2020), which present the GeoCore ontology and discuss the conceptual similarities and differences between GeoCore and GeoSciML. As stated in Chapter 1 and further discussed in Chapter 4, GeoCore is a core ontology for the Geology domain that we adopt as GeoReservoir's conceptual basis by virtue of the previously made ontological analysis, disambiguation, and foundational ontology alignment.

3.7 Final considerations

The previous works reviewed in this chapter give us useful insights into our model's development. However, they do not cover all of this project's requirements. We consider that a strict focus on descriptive view over the domain, with the least contamination of process interpretation as possible, will allow us to create a reusable database of geological occurrences that can support several applications. Moreover, we acknowledge that a strong ontological foundation and the alignment with previously developed foundational and core ontologies builds a solid basis for our model's definition and integration with further effort for Geology ontology development in the industry and academy.

Therefore, we need to develop a new ontology focused on describing deep-marine reservoir geometry and lithology that reuses the GeoCore and the BFO ontologies' conceptual framework. Nonetheless, much of the knowledge described by the GeoReservoir ontology is based on the aggregation and further ontological analysis of the concepts reviewed in this chapter.

4 ONTOLOGICAL FOUNDATIONS AND BACKGROUND

In order to build a model providing a standard, unambiguous, and integrable language for the domain described in Chapter 2, we use the framework of ontology theories, which we review in this chapter. Moreover, we define the ontological terms that we use throughout this thesis and review the BFO and GeoCore ontologies, which compose the reusable conceptual basis that supports the GeoReservoir ontology.

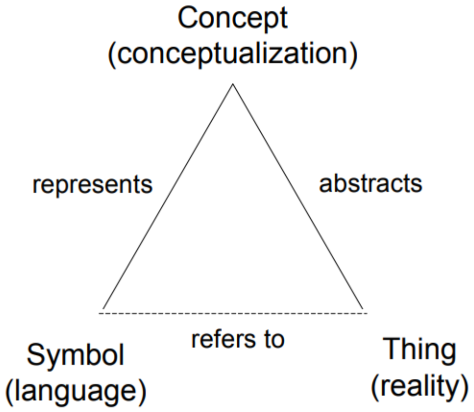
4.1 What is an ontology?

The term “ontology” has at least two meanings depending on the context in which it appears. Here, we adopt the same distinction of Guarino, Oberle and Staab (2009), which differentiates between the meanings of “ontology” in Philosophy and in Computer Science. In Philosophy, Ontology (with uppercase “O”) is the branch that studies the nature and structure of reality, such as the attributes of entities that belong to them because of their essence. Unlike epistemological sciences concerned with describing reality as being perceived under a particular perspective, Ontology focuses on describing things as they are.

In Computer Science, an ontology (with lowercase “o”) is a logical theory that accounts for a domain’s vocabulary intended meaning. When we apply this theory to build a conceptual model, we refer to the resulting ontology as a computational artifact that describes a domain’s concepts and their relations (GUARINO; OBERLE; STAAB, 2009). Under this view, Studer, Benjamins and Fensel (1998) defined ontology as a “formal and explicit specification of a shared conceptualization”. *Formal* means that an ontology is computer-readable, allowing ontology-driven data processing and automated reasoning. *Explicit* means that the definition contained in an ontology should be exact and precise. *Shared* indicates that an ontology must capture the consensual knowledge accepted by a group and not just an individual. *Conceptualization* means that an ontology captures an abstract and intensional model of some phenomenon in the world.

As so, ontologies in Computer Science assume a more pragmatic view than the Ontology in Philosophy: things in reality are perceived under the lenses of concepts. As stated by Guizzardi (2005), concepts are abstract entities that only exist in the mind of a user or a community of users of a language. Each *symbol* in a language does not represent a *thing* in reality but actually represents a *concept* that abstracts something in

Figure 4.1 – The Ullmann's triangle, which draws the relations between a symbol in a language, a concept in the mind of one or more users of this language, and a thing in reality.



Source: Ullmann (1972) and Guizzardi (2005)

reality. These relations between symbol, concept, and thing are depicted by the Ullmann's triangle (Figure 4.1). A concept does not abstract a single individual in the world but rather abstracts every individual that holds the necessary conditions to be considered a sample of it. An example of a concept in this sense would be Rock, which abstracts all the individuals in the world that hold properties like being hard, being constituted by minerals or biological matter, and so forth.

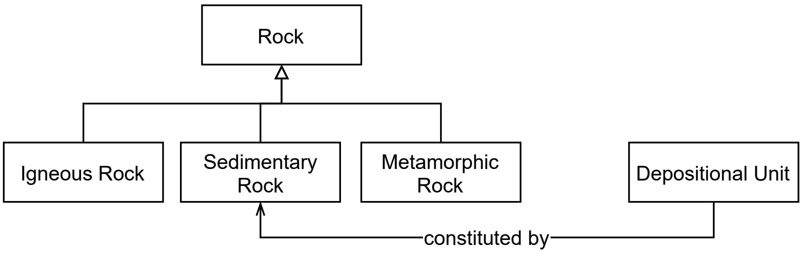
It is crucial to notice that all models have an *ontological commitment* independently of being ontologies or not. A model's ontological commitment is the mapping of its symbols or terms to the concepts that they represent. The ontologies' role is to make the models' ontological commitments explicit, a practice that contributes to the development of more unambiguous and reusable terminologies (GUIZZARDI, 2005; GUARINO; GUIZZARDI; MYLOPOULOS, 2019).

4.2 Structural ontological relations and metaproperties

In this thesis, we consider as an *entity* any unary predicate applicable to individuals that abstracts the necessary conditions for an individual to be considered a sample of it. This "entity" definition represents the "concept" definition presented in Section 4.1 in a logic formalism. Having this formal representation, we can apply it to an individual x by asserting, for example, that "x is a Rock".

A *relation* is a binary predicate that connects two instances, two entities, or an instance to an entity (ARP; SMITH; SPEAR, 2015). Relations are directional so that if a relation r connects x to y, it does not necessarily connect y to x. An example is

Figure 4.2 – Illustration of a modeling example of Rock and its related concepts. Boxes represent concepts, arrows with closed endings represent "is a" relations, and arrows with open endings represent other relations.



Source: the author

the constitution relation (GARCIA et al., 2019), which connects a material entity like Depositional Unit to the matter that constitutes all these objects like Sedimentary Rock (see Figure 4.2).

One particular and significant structural relation is *subsumption*. It is a taxonomical relation where, given two entities, p and q, if p subsumes q, all instances of q are also instances of p (GUARINO; WELTY, 2002b). Let us suppose we have the concept of Rock and three basic types of rock that model the concepts: Igneous Rock, Sedimentary Rock, and Metamorphic Rock. All instances of these three basic types are instances of Rock. Hence, Rock subsumes Igneous Rock, Sedimentary Rock, and Metamorphic Rock. The inverse relation of Subsumes is "Is A", i.e., if Rock subsumes Sedimentary Rock, then Sedimentary Rock is a Rock (see Figure 4.2).

To model subsumption relations consistently, it is useful to analyze the concepts' *metaproperties* and how they reflect in the entities' representation in the model. A metaproperty is a predicate that abstracts some aspect of the ontological nature of a concept. Here, we use the following metaproperties for ontological analysis (GUARINO; WELTY, 2002a; ABEL; PERRIN; CARBONERA, 2015):

- *Rigidity*: a rigid concept is essential for all its instances so that every individual that instantiates that concept must instantiate it while it exists. For example, Rock is a rigid concept because no individual can cease being a rock without ceasing to exist. On the other hand, a not rigid concept is not necessarily essential for all its instances; e.g., Hard is essential for rocks for not for reefs. Some concepts are not essential for all their instances; these are anti-rigid. Reservoir Rock is anti-rigid because every reservoir rock can cease being a reservoir rock once its petroleum is extracted (yet it continues being a rock).
- *Carrying identity*: identity refers to the criteria that we use to tell whether multiple

individuals are the same one. When a concept carries identity, it means that all its instances hold identity criteria, whatever it is. This metaproperty differentiates sortal concepts like Rock from categories or mixins like Porous Rock.

- *Providing identity*: some concepts that carry identity also provide their own identity. For example, Rock provides its own identity while Reservoir Rock carries it. In other words, the criteria that we use to distinguish reservoir rocks are the same that we use to distinguish rocks. Every individual must instantiate exactly one concept that provides identity.
- *Carrying unity*: unity refers to the criteria that we use to recognize an entity as a whole composed of parts and delimited by boundaries. When a concept carries unity, all its instances hold unity criteria. On the other hand, when a concept carries anti-unity, it means that all its instances do not hold unity criteria. This metaproperty helps us distinguish bounded objects like Depositional Unit from the matter or substance that constitutes them, such as Sedimentary Rock.
- *Existential dependence*: denotes that all instances of a concept depend on other entities' existence to exist. For example, the porosity of a rock depends on the existence of that rock to exist. In this case, Porosity is existentially dependent while Rock is the bearer of Porosity.

The practical interests of ontological metaproperty analysis are many. For example, let us suppose that a modeler depicted that Sedimentary Rock subsumes Depositional Unit, i.e., all depositional units are rocks. If we analyze these concepts' metaproperties, we should assign "carrying anti-unity" to Sedimentary Rock and "carrying unity" to Depositional Unit. In other words, all sedimentary rocks carry no unity criteria because of their ontological nature as substances, while all depositional units hold unity criteria for being bounded objects that have parts such as layers and depositional surfaces. Such a model is considered inconsistent by this analysis because it creates a logical contradiction: all sedimentary rocks hold no unity, all depositional units hold unity, and all depositional units are sedimentary rocks, which hold no unity. Hence, Sedimentary Rock cannot subsume Depositional Unit, and the relation between them must be another, such as constitution (GUARINO; WELTY, 2004; GARCIA et al., 2019).

As presented in this section and also discussed by Abel, Perrin and Carbonera (2015), assigning these metaproperties to concepts helps us clarify each term's intended meaning. Besides, capturing the concepts that are rigid and provide their own identity assists modelers to develop integrable models since the corresponding entities are more prone to arise in different systems and contexts. For instance, Rock is a concept that is represented in almost all geological studies, while Reservoir Rock most likely appears in petroleum exploration and production studies.

4.3 Ontology classification by scope

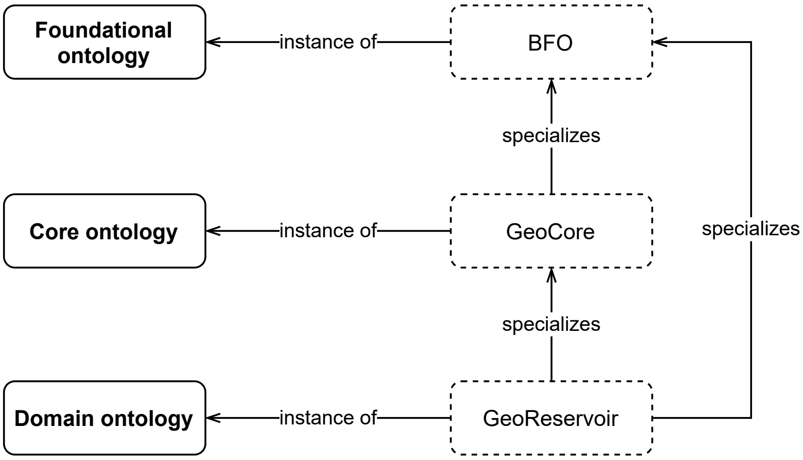
In this work, we adopt a classification of ontologies according to their generalization level (GUARINO, 1998; OBERLE, 2006; GUARINO; OBERLE; STAAB, 2009). A *foundational ontology* is an ontology that aims to provide a broad view of the world. Such an ontology is independent of domain because it describes very general entities, which can include ontological metaproperty connotations such as in the UFO ontology (GUIZZARDI, 2005), or taxonomies and properties of general categories like in DOLCE and BFO ontologies (GANGEMI et al., 2002; ARP; SMITH; SPEAR, 2015).

A *core ontology* is an ontology that describes a domain's most general entities. A core ontology allows integrating the manifold sub-fields of a domain, such as Stratigraphy, Sedimentology, and Petrography in Geology (OBERLE, 2006; GARCIA et al., 2020). A core ontology can specialize a foundational ontology to describe its domain, allowing further integration with other core ontologies and taking advantage of the foundational ontology's philosophical and conceptual basis.

A *domain ontology* is an ontology that covers a restricted domain of interest, such as the deep-marine reservoir geometry covered by this work. A domain ontology can specialize both a foundational and a core ontology, in the same fashion that a core ontology can specialize a foundational ontology. In this work, we develop the GeoReservoir ontology, which specializes the GeoCore (a core ontology) and the BFO (a foundational ontology). This organization is illustrated in Figure 4.3.

As stated by Oberle (2006), the borderline between the types of ontologies is not clearly defined because there is no exhaustive enumeration of all the knowledge domains and their sub-fields. Nevertheless, this distinction is intuitively meaningful and useful for building ontologies. In this thesis's research context, it is essential to have such a distinction for scope delimitation purposes, as we intend to build an integrated ontology network covering many fields of interest for petroleum exploration and production.

Figure 4.3 – Diagram representing the ontology classification as addressed by this work. Rounded boxes bounded by solid lines represent ontology types, while rounded boxes bounded by dashed lines represent example ontologies for each ontology type. Arrows represent relations between the types or ontologies. It is interesting to notice that the GeoReservoir ontology directly specializes both the BFO and the GeoCore ontologies as it uses entities of these two ontologies in its formal definitions.



Source: the author

4.4 The BFO ontology

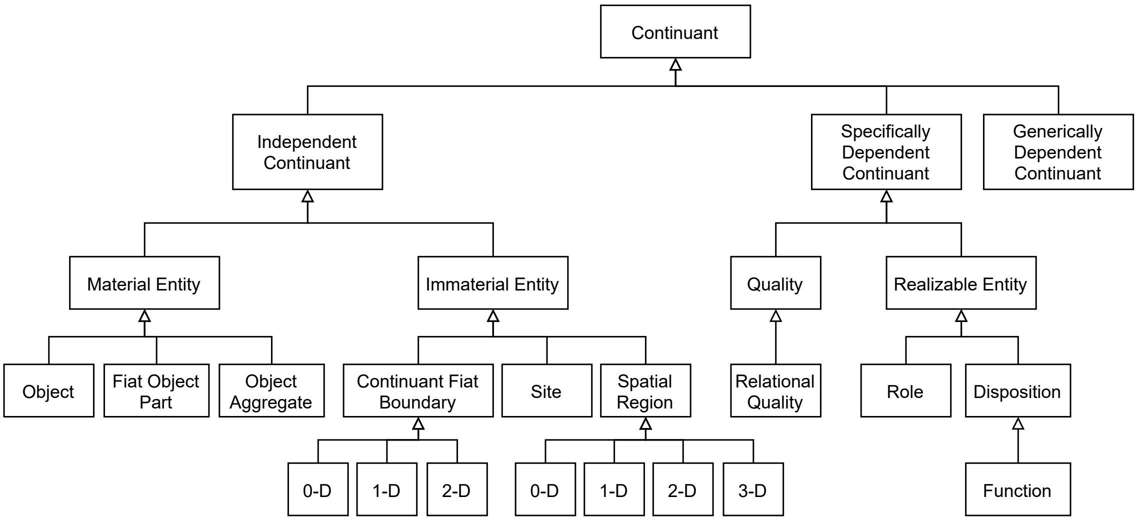
The BFO (Basic Formal Ontology) is a foundational ontology designed to support the description and integration of entities usually found in scientific domains (ARP; SMITH; SPEAR, 2015). It defines a small set of domain-neutral entities that helps ontologists address terminology selection, definition, and classification. The BFO rests on the principle of Realism, which is defined as “a philosophical position according to which reality and its constituents exist independently of our (linguistic, conceptual, theoretical, cultural) representations and can be known, for example, through perceptual experience and through application of the scientific method” (ARP; SMITH; SPEAR, 2015). In other words, it assumes that a sedimentary layer, for example, exists whether geologists assume that it composes a rock body or a stratigraphic succession (GARCIA et al., 2020). Furthermore, the BFO underlies a wide range of domain ontologies openly available on the Web (e.g., the OBO Foundry[1] ontologies).

The BFO ontology divides entities into two main categories: Continuant and Occurrent. A *continuant* is an entity that continues or persists through time, while an *occurrent* is an entity that occurs or happens in time. These two basic categories are also represented by the “endurant” and “perdurant” terms in the literature (GANGEMI et al., 2002; GUIZZARDI, 2005). Occurrents have temporal parts and continuant participants,

---

[1] <http://www.obofoundry.org>

Figure 4.4 – Diagram representing the BFO continuants. Boxes represent entities and arrows represent "is a" relations.



Source: modified from Arp, Smith and Spear (2015)

e.g., a geological process like a given folding occurs in a particular geological age and modifies the shape of a sedimentary layer, which is a participant in this process (ARP; SMITH; SPEAR, 2015; GARCIA et al., 2020).

Since this thesis accounts for the deep-marine reservoir geometry and lithology description, we define continuant entities in the GeoReservoir ontology. We do not define occurrents such as sedimentary processes, even though these continuants are anticipated to support further modeling of occurrents in which they participate. Hence, we review the Continuant subtypes defined by the BFO in this section (Figure 4.4). For a complete reference on all the entities that the BFO defines, one should refer to Arp, Smith and Spear (2015). We depict the BFO continuant entities' summarized definitions as follows:

- *Independent Continuant*: a Continuant that holds no existential dependence and can be the bearer of qualities and other existentially dependent continuants;
- *Material Entity*: an Independent Continuant that has some portion of matter as a part;
- *Object*: a Material Entity that is spatially extended in three dimensions and carries unity criteria, i.e., is a whole composed of unified parts;
- *Fiat Object Part*: a Material Entity that is a part of some Object but is not demarcated from the remainder of that object by any physical discontinuities;
- *Object Aggregate*: a Material Entity made up of a collection of separate Objects;
- *Immaterial Entity*: an Independent Continuant having no Material Entity as part;
- *Continuant Fiat Boundary*: an Immaterial Entity of zero, one, or two dimensions

that does not include a Spatial Region as a part and intuitively exists where a Material Entity meets its surroundings;

- *Site*: an Immaterial Entity in which Material Entities are or can be contained and is bounded by a Material Entity or is a three-dimensional immaterial part of another Site thereof;
- *Spatial Region*: an Immaterial Entity of zero, one, two, or three dimensions that is a part of space;
- *Specifically Dependent Continuant*: a Continuant holding existential dependence on one or more specific Independent Continuants;
- *Quality*: a Specifically Dependent Continuant that is exhibited in its bearers in the whole time;
- *Relational Quality*: a Quality that holds existential dependence on more than one bearer, denoting some relation between its bearers;
- *Realizable Entity*: a Specifically Dependent Continuant that is exhibited in its bearers during some process Occurrent only;
- *Role*: a Realizable Entity that is optional to its bearers, i.e., exists because of some external circumstance in which the bearers do not have to be necessarily involved;
- *Disposition*: a Realizable Entity that exists because of some physical setting of its bearers, i.e., if the disposition ceases to exist, then the bearer is physically changed;
- *Function*: a Disposition that exists because of the conception of its bearers, either natural or by design;
- *Generically Dependent Continuant*: a Continuant that holds existential dependence on one or more bearers and can migrate between bearers.

According to Garcia et al. (2020), using the BFO allows us to take advantage of these entities that are already defined and focus exclusively on the geological domain's main aspects. Besides, other BFO's benefits are that it is small, simple, and well-documented. A few examples using sedimentological terms are presented as follows[2]:

- Depositional Unit: it is an Object for being a whole delimited by depositional surfaces or unconformities and having layers or other units as parts;
- Sedimentary Rock: a Material Entity for being an amount of matter that is not unified;

---

[2] We present these examples in order to explain the advantages of using BFO in the sedimentological domain. These examples do not necessarily reflect the GeoReservoir ontology definitions.

- Depositional Surface: a Continuant Fiat Part of two dimensions for existing where Depositional Unit meets their surroundings;
- Channel (element): an Object such as a Depositional Unit because it consists of a channel-form depositional surface and the sediments that fill it;
- Channel (conduit): a Continuant Fiat Boundary for including no Spatial Region as a part and being an erosional surface in which depositional processes can occur;
- Porosity: a Quality that inheres in some Sedimentary Rock;
- Permeability: a Disposition that inheres in some Sedimentary Rock and is exhibited as it lets fluids pass through it;
- Cross-stratification: a Generically Dependent Continuant since it is a pattern that is repeatedly and continuously exhibited in many Depositional Units.

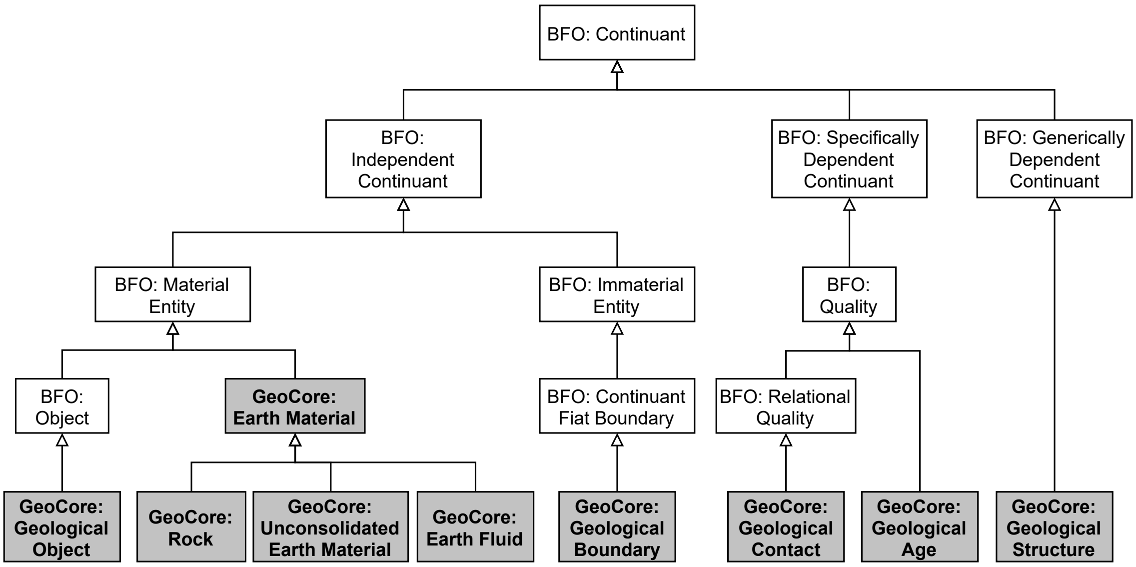
It is interesting to notice how the BFO also serves as a tool for terminology disambiguation, such as in the Channel example above. As discussed in Chapter 2, this term is often ambiguous in the literature because sometimes it refers to the conduits where sedimentary flows pass by, and other times it refers to the channel sedimentary fills and their boundaries. Classifying a term like this under the BFO taxonomy is a practice that helps us depict its multiple meanings.

4.5 The GeoCore ontology

The GeoCore is a core ontology for the Geology domain (GARCIA et al., 2020). It is designed to precisely define a limited set of general entities that permeate the whole Petroleum Geology domain. The authors offer a sound and general-use ontology that aids in integrating knowledge about distinct sub-domains of Geology, Reservoir Engineering, Geophysics, among others. The core ontology specializes BFO entities in terms of geological entities. It aims to cover the main entities, but it is not a complete partition, allowing its users to include entities that specialize BFO's entities directly, such as qualities. Hence, users can take advantage of both GeoCore and BFO's conceptual basis to build specialized domain ontologies, such as we do in this work.

Considering the BFO's distinction between continuants and occurrents and our strict focus on the continuants (see Section 4.4), we review the GeoCore continuant definitions in this section (Figure 4.5). For a complete reference on all definitions, one should refer to Garcia et al. (2020). Some definitions rest on metaproperties, which are reviewed

Figure 4.5 – Diagram representing the GeoCore continuants. Gray boxes represent GeoCore entities, while white boxes represent BFO entities. Arrows represent “is a” relations.



Source: modified from Garcia et al. (2020)

in Section 4.2. We summarize the GeoCore continuant definitions as follows:

- *Geological Object*: a natural BFO Object generated by some geological process and constituted by some Earth Material, providing its own identity, carrying unity, and being rigid;
- *Earth Material*: a BFO Material Entity of natural matter, either solid, fluid, or unconsolidated, generated by some geological process, providing its own identity and being rigid but carrying no unity;
- *Rock*: a solid and consolidated Earth Material made of polycrystalline, monocrystalline, or amorphous mineral matter or material of biological origin;
- *Unconsolidated Earth Material*: an Earth Material’ constituted by an aggregate of solid particles but not consolidated as a Rock itself;
- *Earth Fluid*: a fluid Earth Material such as water, oil, gas, and mixtures of those;
- *Geological Boundary*: a BFO Continuant Fiat Boundary corresponding to a physical discontinuity of any nature, located on a Geological Object’s external surface;
- *Geological Contact*: a BFO Relational Quality that relates two Geological Objects whose external boundaries are adjacent;
- *Geological Age*: a BFO Quality of some Geological Object related to a geological time interval during which some geological process generated that object;
- *Geological Structure*: a BFO Generically Dependent Continuant that describes the internal arrangement of some Geological Object, i.e., the configuration and mutual

relations between its different parts, such as a fold.

GeoCore also rests on the constitution relation defined by Garcia et al. (2019), i.e., a Geological Object is *constituted by* an Earth Material. This assumption determines that an object and the matter that constitutes it are different entities, even though they occupy the same place in space. For example, let us consider a channel element as of McHargue et al. (2011), which is a channel-form sedimentary body delimited by depositional surfaces or discontinuities, i.e., its boundaries. Under the GeoCore ontology, this sedimentary body is a Geological Object, while the portion of sedimentary rock contained within these boundaries is a Rock. In other words, this view assumes that the Rock constituting some Geological Object can be transported between different bodies or even cease constituting a body.

Analyzing the ontological features of Geological Objects and Rocks, we can clearly distinguish these two entities. A Rock has its identity defined by its composition and the geological process that generated it, has uniform properties across all of its spatial extension, and has no unifying relation among its parts. A Geological Object, by its hand, has a unifying relation among its parts (e.g., the portions of Rock that constitute it) and holds its own identity criteria.

This distinction between Rock and Geological Object is essential because the existentially dependent entities that characterize rocks and geological objects are different. Geologists characterize rocks by their intrinsic properties, such as porosity and permeability, which are inherent to the amount of matter generated in the original geological process and are preserved in their portions even when physically disconnected. On its hand, a Geological Object defines a *whole* and has its identity associated with its boundaries and expression in a three-dimensional space, being characterized by its geometric shape, dimensions, location, and other geometrical properties. When analyzing the architecture and geometry of depositional systems in outcrop and seismic scales, geologists are usually interested in properties and relations that hold between Geological Objects. Acknowledging this distinction is vital for the accurate modeling of the domain addressed by this thesis.

4.6 Qualitative spatial representation

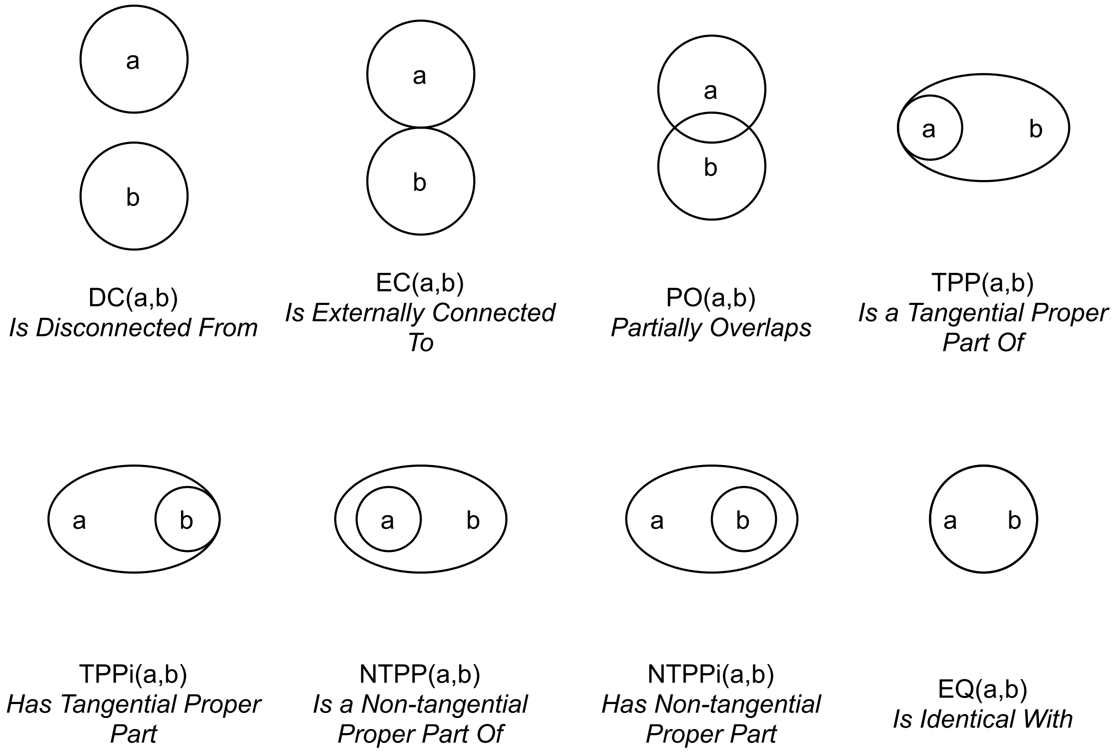
As the GeoReservoir ontology is intended to support the description of geometrical and lithological properties of deep-marine depositional systems, we include spatial relations in this work in order to represent the sedimentary bodies' spatial distribution. Specifically, we need to represent two spatial aspects: the *mereotopology*, i.e., the parthood relations by virtue of the bodies' spatial arrangement, and the *orientation* relations, which define in what directions the bodies are located relative to each other (COHN; RENZ, 2008). These two aspects are sufficient to describe the depositional systems' spatial distribution and were exhaustively studied, validated, and documented in the literature.

The mereotopological aspect asserts that two spatial entities are *connected* if the intersection of the spaces occupied by both is not empty. Considering this, a spatial entity x is a mereotopological part of another spatial entity y if and only if whatever entity connected to x is also connected to y. Based on these notions, the RCC-8 (Region Connection Calculus) is a set of qualitative spatial relations that consists of the following definitions (see Figure 4.6): Is Disconnected From (DC), Is Externally Connected To (EC), Partially Overlaps (PO), Is a Tangential Proper Part Of (TPP), Has Tangential Proper Part (TPPi), Is a Non-tangential Proper Part (NTPP), Has Non-tangential Proper Part (NTPPi), and Is Identical With (EQ) (COHN et al., 1997; COHN; RENZ, 2008).

The directional relations express where an object is located relative to another object and a frame of reference. As such, these relations can be ternary such that the relata are the two related entities and the frame of reference, or can be binary and presuppose a specific frame of reference. Moreover, a frame of reference can be extrinsic to the related entities such as the cardinal reference system (directions North, South, East, and West), or intrinsic to one of the related entities, e.g., when we say that a person is in the front of a vehicle (Figure 4.7). Whatever the case, it is necessary to explicitly assume a frame of reference to ensure that the directional relations' meanings are precise (COHN; RENZ, 2008).

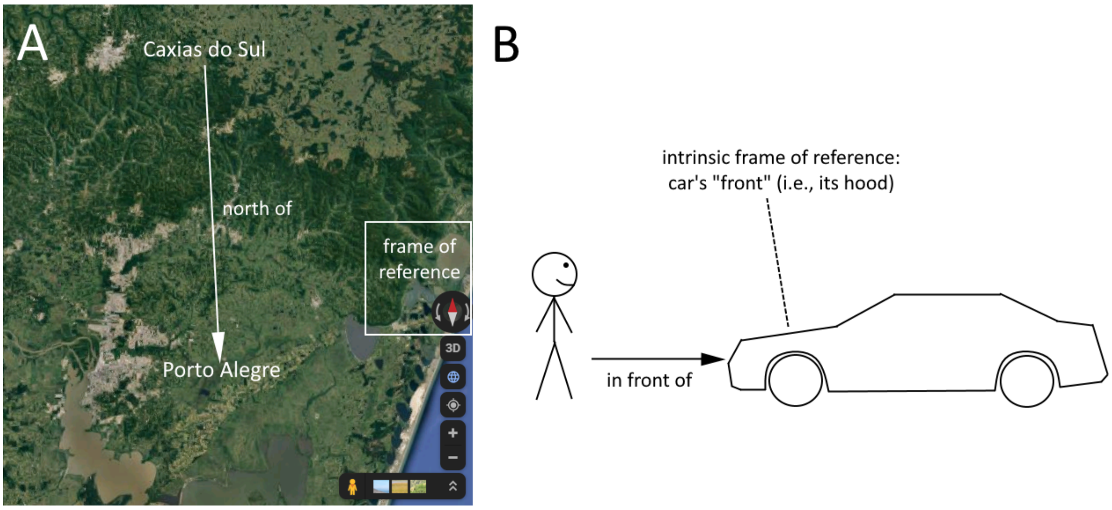
We assume that mereotopological and directional relations are sufficient for depositional systems to describe the sedimentary bodies' spatial distribution. Furthermore, spatial relations are correlated with some sedimentary features. For example, when a channel-form sedimentary unit is a proper part of another sedimentary unit having the same geometry, the former holds a lower hierarchical order than the latter unit. Know-

Figure 4.6 – Visual representation of the RCC-8 relations. Circles or ellipses represent spatial entities.



Source: modified from Cohn et al. (1997)

Figure 4.7 – Examples of distinct frames of reference for directional spatial relations. The relation illustrated in (A) uses the cardinal reference system as its frame of reference, while the relation illustrated in (B) assumes an intrinsic frame of reference.



Source: the author. Part (A) was drawn over a screenshot taken from the Google Maps application <https://www.google.com/maps>.

ing the orientation of a depositional system, it is possible to establish the relative ages of sedimentary units, e.g., the uppermost units are younger than the lower ones because they were the last to be deposited in that system. Being so, spatial relations compose a crucial aspect to describe geological occurrences' descriptive features, which is this thesis's focus, and further interpret these occurrences in the future (CICCONETO et al., 2020).

4.7 Final considerations

Relying on structural ontological foundations, metaproperties, and previously developed ontologies, we compose a solid and robust framework to analyze the sedimentological terms and build a sound and unambiguous terminology to support the description of deep-marine depositional system lithological and geometrical properties. The approach presented in this thesis depicts how we can take advantage of these foundations to build a knowledge model coping with the complexity and difficulties that arise in the effort to define the domain's most relevant entities.

5 GEORESERVOIR ONTOLOGY DEVELOPMENT

This chapter presents the GeoReservoir ontology, which supports the description of deep-marine depositional systems' geometrical and lithological properties, as presented in Chapter 2. The GeoReservoir ontology extends the BFO and GeoCore ontologies, which were reviewed and described in Chapter 4.

In this project, we understood that spatial relation definitions could serve several sub-domains in Geology and not only the domain presented in this thesis. Hence, we separated the spatial relations in another ontology distinct from GeoReservoir and developed this separate ontology as a complement to GeoCore (Figure 5.1). This spatial relations ontology, which embodies the relations presented in Section 4.6, was published and presented at the Ontobras 2020 conference, which occurred in Vitória, Brazil (CICCONETO et al., 2020).

We start this section by presenting the methods and tools employed to build the GeoReservoir and spatial relation ontologies. Next, we describe the notations that we use to present the ontologies formally. Finally, we present the GeoReservoir and spatial relation ontologies themselves.

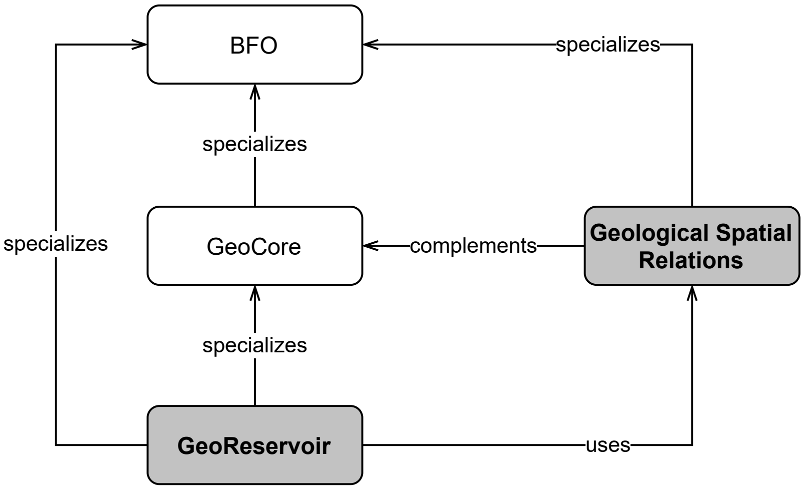
5.1 Methods and tools

To develop the GeoReservoir ontology, we did a collaborative effort involving ontology engineers and domain experts trained in ontology foundations. This collaboration included a series of meetings in which the stakeholders identified the key concepts and discussed their ontological nature based on the foundations described in Chapter 4.

We structured the ontology development (Figure 5.2) as an iterative process, in which each iteration creates an evolving prototype. This “evolving prototype” idea is based on the METHONTOLOGY process (FERNANDEZ-LOPEZ; GOMEZ-PEREZ; JURISTO, 1997), in which each ontology development cycle produces a prototype that is evaluated and further developed in subsequent cycles. Even after this thesis's publication, further iterations can be executed to evolve and release new versions of the ontologies whose development is described here.

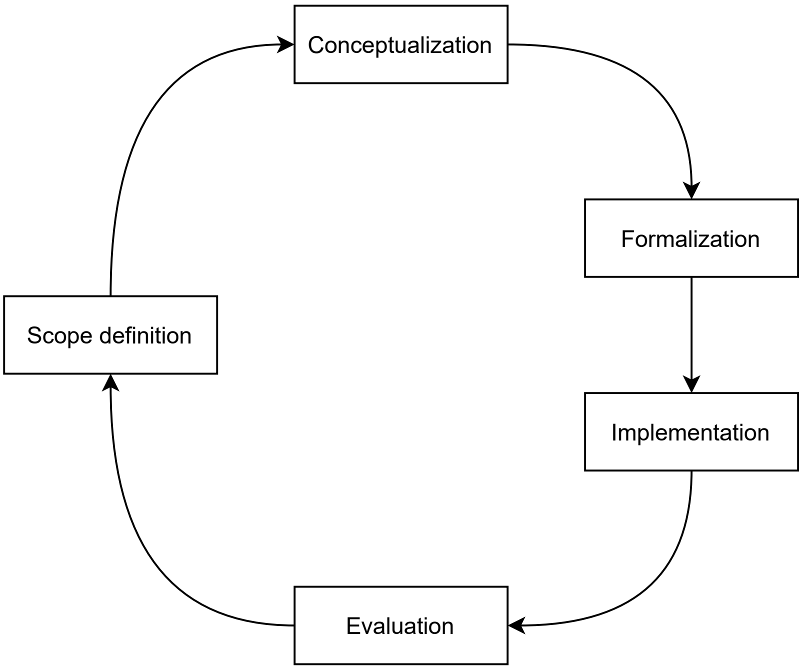
Each development iteration is divided into a series of steps, which we structured as follows (Figure 5.2):

Figure 5.1 – Diagram representing the relations between the ontologies depicted in this work. Gray rounded boxes represent the ontologies developed in this work, while white rounded boxes represent the ontologies reused by this work. Arrows represent relations between the ontologies.



Source: the author

Figure 5.2 – The GeoReservoir ontology continuous development process. Each iteration (i.e., from scope definition to evaluation) creates an evolving prototype that can be approved or not. If not approved, a new iteration is started.



Source: the author

1. *Scope definition*: in this step, we define the ontology's scope and purposes, which we described in Chapters 1 and 2 of this thesis.
2. *Conceptualization*: here, we identify, define, and organize the relevant concepts of the domain. In this step, we build a taxonomical structure, align the entities with BFO and GeoCore ontology, and establish relations between them, resulting in a semi-formal ontology.
3. *Formalization*: in this step, we formalize the ontology using Aristotelic definitions (i.e., "an A is a B that C"), assign metaproperties, and establish the necessary formal relations, i.e., the relations that the entities necessarily hold when they exist. The formal ontology is detailed in Subsections 5.3 and 5.4.
4. *Implementation*: here, we implement the ontology in OWL 2 language using Protégé, which is detailed below. This implementation is available in a Web repository[1].
5. *Evaluation*: in this final step, we do a user-focused evaluation[2], validating whether the domain specialists consider that the ontology is detailed enough to represent all the relevant reservoir features and support geological interpretations.

The Protégé tool, which we use in step 4, is a free and publicly available knowledge model editor (MUSEN, 2015). It persists models in OWL 2 and other available languages. Although Protégé and OWL 2 have some limitations in expressing the whole model semantics, they hold the advantages of interoperability and extensibility since BFO and GeoCore also have OWL 2 implementations, allowing us to import and integrate them with our ontology. Protégé also features an embedded automated reasoner, which we employ to validate the ontology's logical consistency.

5.2 Formal notations

We adopt the notation described in the following to formalize the GeoReservoir ontology entities. This notation aims to provide all formal information to the reader and clarify the entities' meaning explicitly and clearly. When a definition reuses some previously existent entity from another ontology such as BFO and GeoCore, we point out this reuse quoting in parentheses. Indeed, we keep the alignment with these previously

---

[1] <https://github.com/BDI-UFRGS/GeoReservoirOntology>

[2] This "user-focused evaluation" step is based on the ontology evaluation methods described by Sure, Staab and Studer (2009).

defined ontologies as much as possible and include only our application domain's definitions. The ontological foundations and structural relations needed to understand the formal definitions are detailed in Chapter 4.

Each definition is formally described by the following sequence of information:

- A brief textual definition;
- **Subsumed by**: indicates that the entity's instances are also instances of another entity;
- **Disjoint with**: indicates that the entity's instances cannot be instances of the listed entities simultaneously;
- **Metaproperties**: indicates the entity's metaproperties using a notation that we describe below;
- **Relation**: defines the necessary relations for the entity using a notation that we describe below;
- **Equivalent to**: for entities that are synonyms, we list the equivalent terms;
- **Notes**: some discussion or usage example.

For metaproperties, we use the following notation[3]:

- $R+$ / $R-$ / $R\sim$: rigid / not rigid / anti-rigid;
- $I+$ / $I-$: carries / does not carry identity;
- $O+$ / $O-$: provides / does not provide identity;
- $U+$ / $U-$: carries / does not carry unity;
- $ED+$ / $ED-$: existentially dependent / independent.

For formal relations, we use a notation similar to that used by Protégé. One example of a relation using this notation is "*Constituted By (GeoCore) some Sedimentary Rock (Definition 8)*", which indicates that every instance of the target entity must be constituted by at least one instance of Sedimentary Rock. The complete notation term list is defined as follows:

- *some*: existential quantifier (at least one instance);
- *only*: universal quantifier (all instances);
- *min n*: existential quantifier with minimum cardinality;
- *not*: logical negation;

---

[3] A very similar notation is used by Guarino and Welty (2004) and Abel, Perrin and Carbonera (2015).

- *and*: logical conjunction;
- *or* : logical disjunction (not exclusive);
- *decimal*: real number without constraint;
- *decimal[constraint]*: real number with some constraint;
- *1* : a literal number;
- *string*: any sequence of characters;
- *“abc”*: a literal string;
- *{“abc”, “def”, 1, 2}*: a disjoint set of literals.

For relation definitions, we use the following structure:

- A brief textual definition;
- **Subrelation of**: indicates that, when the relation holds for x and y, some other relation also holds for x and y;
- **Domain and image**: indicate the entities that can hold the relation;
- **Inverse of**: indicates that, when the relation holds for x and y, some other relations holds for y and x in the inverse direction.
- **Properties**: *transitive* indicates that, when the relation holds for x and y and also for y and z, it also holds for x and z; and *symmetric* indicates that, when the relation holds for x and y, it also holds for y and x in the inverse direction.

5.3 GeoReservoir ontology presentation

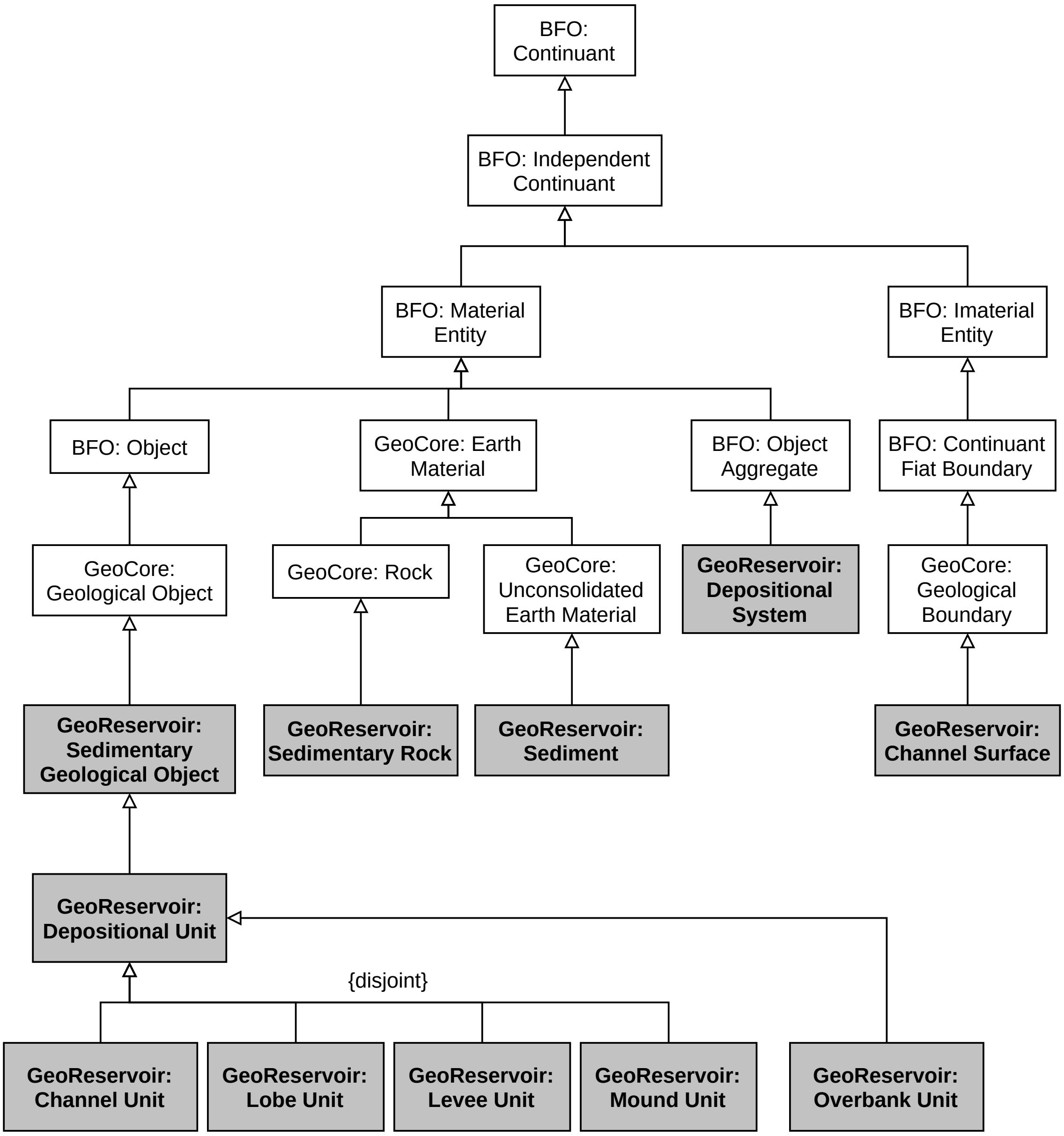
In the following, we present the GeoReservoir ontology entities and relations (Figures 5.3 and 5.4). Some definitions require to report to the GeoCore (GARCIA et al., 2020) or the BFO (ARP; SMITH; SPEAR, 2015) ontologies to get a complete understanding.

An important distinction we make in our model is the separation between substantial entities and moments. Substantials are existentially independent entities, such as a depositional unit or a rock (Figure 5.3). Moments are existentially dependent entities, such as the sinuosity of a channel, needing a substantial entity as its bearer to exist (Figure 5.4). This distinction affects the modeling because, to get a consistent usage of the model, we can only instantiate a moment having instantiated its bearer. Considering the substantial entities, another distinction is significant: identifying the substantial entities that supply ontological identity and those that carry identity (GUARINO; WELTY, 2002b). In our model, most of the substantial entities derive from Geological Object and Rock, which are defined in GeoCore and supply identity. These two GeoCore entities are present in most geological models and anchor different conceptual models in the Geology domain to the same ontological reference, ensuring the models' integrability.

5.3.1 Substantial entities

1. **Sedimentary Geological Object** A Geological Object constituted by some Sedimentary Rock or Sediment.
   Subsumed by: Geological Object (GeoCore)
   Metaproperties: R+ I+ O− U+ ED−
   Relation: Constituted By (GeoCore) only Sedimentary Rock (Definition 8)
2. **Depositional Unit** A Sedimentary Geological Object recognizable in a mapping scale of at least 1:1000 m.
   Subsumed by: Sedimentary Geological Object (Definition 1)
   Metaproperties: R+ I+ O− U+ ED−
3. **Channel Unit** An elongated Depositional Unit having some Channel Surface as its boundary and constituted by some Sediment or Sedimentary Rock that fills it (MCHARGUE et al., 2011).
   Subsumed by: Depositional Unit (Definition 2)
   Disjoint with: Lobe Unit (Definition 4), Levee Unit (Definition 5), Mound Unit (Definition 6)
   Metaproperties: R+ I+ O− U+ ED−
   Relation: Has 2D Boundary (BFO) some Channel Surface (Definition 11)
   Relation: Has Quality (BFO) some Channel Geometry (Definition 18)
   Notes: in stratigraphic terms, a channel unit is constituted by the sediments or sedimentary rocks that fill a channel. It is important to notice that, in more general terms, a channel is a passageway for fluids or sediments, i.e., an immaterial entity. Shall one elaborate about the erosion and deposition processes that generate these units (MCHARGUE et al., 2011), it is essential to have this disambiguation.
   Notes: in this work, we decided not to model the named hierarchical scales as entities (e.g., Channel Element, Channel Complex, Channel Complex Set) because

Figure 5.3 – GeoReservoir ontology substantial entities' taxonomical overview. White boxes represent BFO and GeoCore entities, while gray boxes represent GeoReservoir entities. All arrows represent "subsumed by" relations.



Source: the author

there exists a wide variety of hierarchical schemes in the literature, and they are not necessarily interoperable (CULLIS et al., 2018). Instead, we focus on the fact that units can be fractally nested: a unit can be a Proper Spatial Part Of (Definition 34) another unit, which can be a Proper Spatial Part Of another unit, and so forth. We consider that this approach solves the problem of implementing a solution for the repeatability of the geological objects in multiple scales. As so, a Channel Unit can be decomposed in other Channel Units many times as necessary, keeping the representation structure. In the future, we can study the integration of hierarchical schemes and try to infer scales from the spatial relations.

4. **Lobe Unit** A Depositional Unit having some Lobe Geometry.
   Subsumed by: Depositional Unit (Definition 2)
   Disjoint with: Channel Unit (Definition 3), Levee Unit (Definition 5), Mound Unit (Definition 6)
   Metaproperties: R$+$ I$+$ O$-$ U$+$ ED$-$
   Relation: Has Quality (BFO) some Lobe Geometry (Definition 19)
   Notes: these depositional units are usually found at deep-marine depositional system terminal regions, or at channel laterals as Overbank Units (Definition 7).
   Notes: for Lobe Unit hierarchical scales, we adopted the same approach as in Channel Unit (Definition 3).
5. **Levee Unit** A Depositional Unit having Wedge Geometry and forming a bank or ridge geomorphology associated with some Channel Unit.
   Subsumed by: Depositional Unit (Definition 2)
   Disjoint with: Channel Unit (Definition 3), Lobe Unit (Definition 4), Mound Unit (Definition 6)
   Metaproperties: R$\sim$ I$+$ O$-$ U$+$ ED$-$
   Relation: Has Quality (BFO) some Wedge Geometry (Definition 21)
   Notes: this term is not consensual as some authors refer to these units as the wedge-shape embankments around channels, which is the definition that we adopt here, while other authors refer to them as not necessarily being these embankments and possibly having wing shapes, among other definitions.
   Notes: this term refers specifically to the geomorphological association of some wedge-shaped unit with some channel unit, even if the channel does not exist at the moment of the observation by the geologist and the levee is probabilistically identified by tendencies (e.g., typical facies). In other words, the term “levee” onto-

logically denotes this relation even though it cannot be deterministically observed.

6. **Mound Unit** A Depositional Unit having Mound Geometry with an irregular top surface and internal chaotic facies.

   Subsumed by: Depositional Unit (Definition 2)

   Disjoint with: Channel Unit (Definition 3), Lobe Unit (Definition 4), Levee Unit (Definition 5)

   Metaproperties: R+ I+ O− U+ ED−

   Relation: Has Quality (BFO) some Mound Geometry (Definition 20)

7. **Overbank Unit** A Depositional Unit found at some overbank area, which is geomorphologically associated with some Channel Unit or Levee Unit.

   Subsumed by: Depositional Unit (Definition 2)

   Metaproperties: R∼ I+ O− U+ ED−

   Notes: this term refers to the units that usually result from the sedimentary current overflows and have geometries such as lobate or waveform. In the same way that occurs with Levee Unit (Definition 5), this term refers to a geomorphological association with some other unit that may not exist at the observation moment.

8. **Sedimentary Rock** A Rock constituted by some collection of sedimentary grains or particles (GARCIA et al., 2019).

   Subsumed by: Rock (GeoCore)

   Metaproperties: R+ I+ O− U− ED−

9. **Sediment** An Unconsolidated Earth Material constituted by some collection of sedimentary grains or particles (GARCIA et al., 2019).

   Subsumed by: Unconsolidated Earth Material (GeoCore)

   Metaproperties: R+ I+ O− U− ED−

10. **Depositional System** An Object Aggregate whose members are mereotopologically linked Depositional Units.

    Subsumed by: Object Aggregate (BFO)

    Metaproperties: R+ I+ O+ U+ ED−

    Relation: Has Member (BFO) min 2 Depositional Unit (Definition 2)

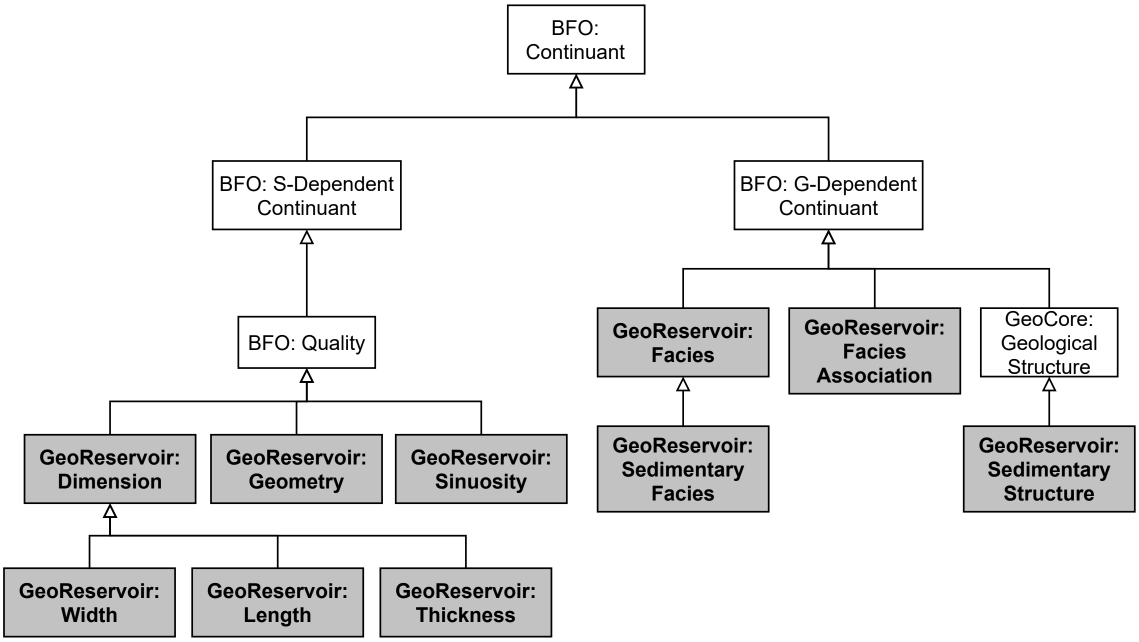
11. **Channel Surface** A Geological Boundary that forms the basis of a Channel Unit, having a concave-up shape unless truncated by overlying depositional units.

    Subsumed by: Geological Boundary (GeoCore)

    Metaproperties: R+ I+ O− U+ ED−

    Relation: 2D Boundary Of (BFO) some Channel Unit (Definition 3)

Figure 5.4 – GeoReservoir ontology moment entities' taxonomical overview. White boxes represent BFO and GeoCore entities, while gray boxes represent GeoReservoir entities. All arrows represent "subsumed by" relations.



Source: the author

5.3.2 Moment entities

All the entities below correspond to existentially dependent (ED+) entities, whose ontological identities are somewhat linked to their respective bearers. Hence, we do not present their ontological metaproperties individually.

12. **Dimension** A Quality that inheres in an Independent Continuant by virtue of one of its one-dimensional extents. It is a category that includes at least the three dimensions defined in this ontology: Width, Length, and Thickness. All dimensions are expressed by a real number indicating its value and a literal string indicating its measurement unit.

    Subsumed by: Quality (BFO)

    Relation: Quality Of (BFO) exactly 1 Independent Continuant (BFO) and Quality Of (BFO) only Independent Continuant (BFO)

    Relation: Has Quality Value (Definition 27) some decimal and Has Quality Value (Definition 27) only decimal

    Relation: Has Measurement Unit (Definition 28) some string and Has Measurement Unit (Definition 28) only string

13. **Width** A Dimension that inheres in an Independent Continuant by virtue of the

distance between its two lateral extremes.

Subsumed by: Dimension (Definition 12)

14. **Length** A Dimension that inheres in an Independent Continuant by virtue of the distance between its two extremes along its longest side.

    Subsumed by: Dimension (Definition 12)

15. **Thickness** A Dimension that inheres in an Independent Continuant by virtue of the distance between its two extremes perpendicularly to its width and length.

    Subsumed by: Dimension (Definition 12)

16. **Geometry** A Quality that inheres in a Depositional Unit by virtue of its external three-dimensional shape. It does not reflect the exact mathematical specifications of a geometric shape but abstracts and simplifies these specifications (ROVETTO, 2011).

    Subsumed by: Quality (BFO)

    Relation: Quality Of (BFO) exactly 1 Depositional Unit (Definition 2) and Quality Of (BFO) only Depositional Unit (Definition 2)

    Relation: Has Quality Value (Definition 27) exactly 1 Geometry Value (Definition 17) and Has Quality Value (Definition 27) only Geometry Value (Definition 17)

17. **Geometry Value** An enumeration of the possible values for Geometry instances.

    Equivalent to: {“channel”, “irregular”, “lens”, “lobe”, “mound”, “scour”, “sheet”, “sigmoid”, “waveform”, “wedge”, “wing”}

18. **Channel Geometry** A synonym for Geometry with “channel” value.

    Equivalent to: Geometry (Definition 16) and Has Quality Value (Definition 27) only {“channel”}

19. **Lobe Geometry** A synonym for Geometry with “lobe” value.

    Equivalent to: Geometry (Definition 16) and Has Quality Value (Definition 27) only {“lobe”}

20. **Mound Geometry** A synonym for Geometry with “mound” value.

    Equivalent to: Geometry (Definition 16) and Has Quality Value (Definition 27) only {“mound”}

21. **Wedge Geometry** A synonym for Geometry with “wedge” value.

    Equivalent to: Geometry (Definition 16) and Has Quality Value (Definition 27) only {“wedge”}

22. **Sinuosity** A Quality that inheres in a Channel Unit by virtue of how wavy it is

across its length. It is expressed by a real number given by its curved length divided by its length in a straight line.

Subsumed by: Quality (BFO)

Relation: Quality Of (BFO) exactly 1 Channel Unit (Definition 3) and Quality Of (BFO) exactly 1 Channel Unit (Definition 3)

Relation: Has Quality Value (Definition 27) some decimal[$\geq$ 1] and Has Quality Value (Definition 27) only decimal[$\geq$ 1]

23. **Facies** A visual pattern of properties of Geological Objects.

    Subsumed by: Generically Dependent Continuant (BFO)

    Relation: Generically Depends On (BFO) only Geological Object (GeoCore)

    Notes: this entity is defined in the literature as a combination of features that differentiates a geological object from its adjacent geological objects (WALKER; JAMES, 1992). However, a Facies does not describe a single Geological Object's features: it might repeat in several objects as a pattern. This particular aspect matches the definition of Generically Dependent Continuant in BFO. Another important aspect is that the features that compose facies vary depending on the research context, e.g., Sedimentary Facies (Definition 24), Petrofacies (ROS; GOLDBERG, 2007), Lithofacies, Ichnofacies, and Biofacies (NICHOLS, 2009).

24. **Sedimentary Facies** A Facies consisting of the sum of the sedimentological characteristics of Depositional Units (MIDDLETON, 1973; NICHOLS, 2009; CARBONERA; ABEL; SCHERER, 2015).

    Subsumed by: Facies (Definition 23)

    Relation: Generically Depends On (BFO) only Depositional Unit (Definition 2)

    Notes: as an example, if we would model Cross-bedded Sandstone (NICHOLS, 2009), it would be an entity subsumed by Sedimentary Facies having the following relations: (1) Generically Depends On (BFO) some (Constituted By (GeoCore) some Sandstone (a)); and (2) Has Part (BFO) some Cross-bedding Structure (b). (a): Sandstone would be an entity subsumed by Sedimentary Rock (Definition 8). (b): Cross-bedding Structure would be an entity subsumed by Sedimentary Structure (Definition 26).

25. **Facies Association** A Generically Dependent Continuant that consists of a collective of Facies.

    Subsumed by: Generically Dependent Continuant (BFO)

    Relation: Has Member (BFO) min 2 Facies (Definition 23)

Notes: it usually describes the Facies that are typically found in specific regions or sub-environments. For example, the Channel Axis Association (MCHARGUE et al., 2011) describes Sedimentary Facies typically found at the center part of channels (in a cross-section point of view).

26. **Sedimentary Structure** A Geological Structure consisting of a pattern of the internal arrangement of Depositional Units.
    Subsumed by: Geological Structure (GeoCore)
    Relation: Generically Depends On (BFO) only Depositional Unit (Definition 2)
    Notes: an example would be Cross-bedding Structure as noted in Definition 24.

5.3.3 Auxiliary relations

27. **Has Quality Value** A relation between a Quality and its value.
    Domain: Quality (BFO)
28. **Has Measurement Unit** A relation between a measurable Quality and a literal string that represents its measurement unit.
    Domain: Quality (BFO)
    Image: string

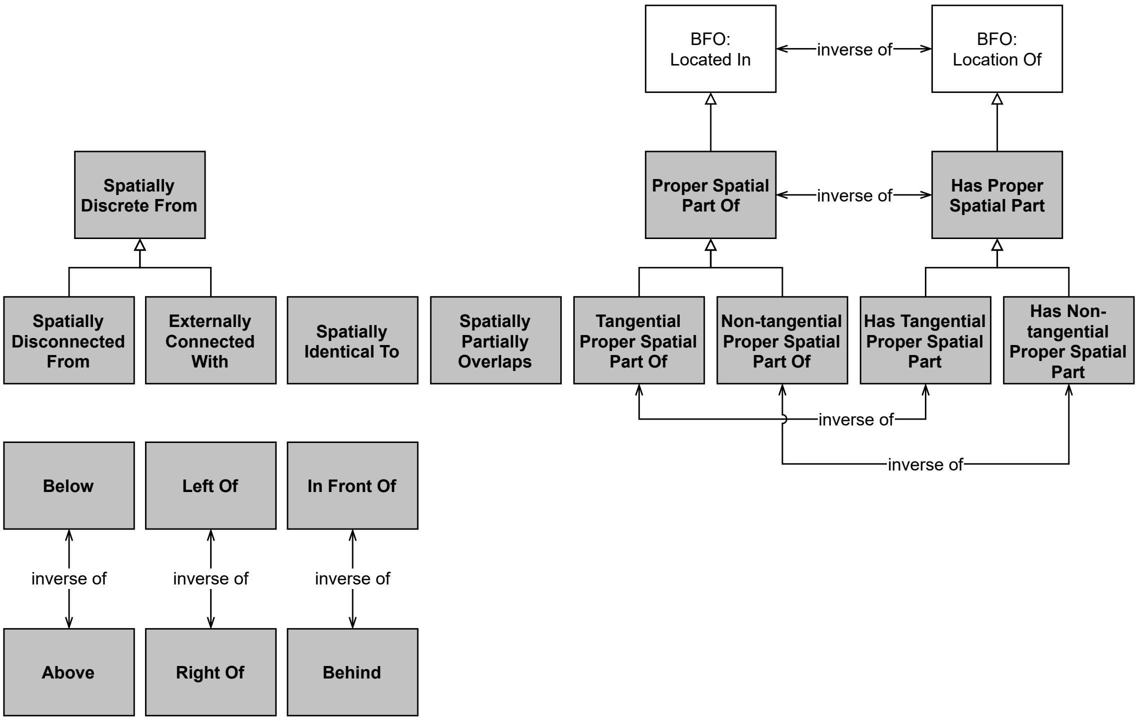
5.4 Geological Spatial Relations ontology presentation

In the following, we present the Geological Spatial Relations ontology (Figure 5.5), which complements the GeoCore ontology with the relations required to support the geological bodies' spatial distribution description. Some definitions require to report to the BFO ontology (ARP; SMITH; SPEAR, 2015) to get a complete understanding.

All relations described by the ontology connect particulars that instantiate the type Independent Continuant (BFO). We made this modeling decision because Independent Continuant is a category that covers all the types of spatial entities that can be related, such as Geological Object, Site, and Geological Boundary. This modeling decision also increases the ontology's reuse and integration potential.

29. **Spatially Discrete From** A spatial relation between two Independent Continuants in which both do not share the same spatial region, either wholly or partially.
    Domain: Independent Continuant (BFO)

Figure 5.5 – Geological Spatial Relations ontology overview. White boxes represent BFO relations, while gray boxes represent our ontology's relations. Arrows with closed endings represent Subrelation axioms, and arrows with open endings represent other axioms.



Source: modified from Cicconeto et al. (2020)

Image: Independent Continuant (BFO)

Properties: symmetric

30. **Spatially Disconnected From** A spatial relation between two Independent Continuants that are Spatially Discrete From each other and whose external boundaries are not adjacent.

    Subrelation of: Spatially Discrete From (Definition 29)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Properties: symmetric

31. **Externally Connected With** A spatial relation between two Independent Continuants that are Spatially Discrete From each other and whose external boundaries are adjacent.

    Subrelation of: Spatially Discrete From (Definition 29)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Properties: symmetric

32. **Spatially Identical To** A spatial relation between two Independent Continuants in

which both occupy precisely the same spatial region.

Domain: Independent Continuant (BFO)

Image: Independent Continuant (BFO)

Properties: symmetric

33. **Spatially Partially Overlaps** A spatial relation between two Independent Continuants in which both share a part of the spatial regions they occupy.

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Properties: symmetric

34. **Proper Spatial Part Of** A spatial relation between two Independent Continuants, x and y, in which the spatial region that x occupies is entirely inside of the spatial region that y occupies.

    Subrelation of: Located In (BFO)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Has Proper Spatial Part (Definition 37)

    Properties: transitive

35. **Tangential Proper Spatial Part Of** A spatial relation between two Independent Continuants, x and y, in which x is a Proper Spatial Part Of y and whose external boundaries are adjacent.

    Subrelation of: Proper Spatial Part Of (Definition 34)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Has Tangential Proper Spatial Part (Definition 38)

36. **Non-tangential Proper Spatial Part Of** A spatial relation between two Independent Continuants, x and y, in which x is a Proper Spatial Part Of y and whose external boundaries are not adjacent.

    Subrelation of: Proper Spatial Part Of (Definition 34)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Has Non-tangential Proper Spatial Part (Definition 39)

37. **Has Proper Spatial Part** A spatial relation between two Independent Continuants, x and y, in which the spatial region that y occupies is entirely inside the spatial region that x occupies.

Subrelation of: Location Of (BFO)

Domain: Independent Continuant (BFO)

Image: Independent Continuant (BFO)

Inverse of: Proper Spatial Part Of (Definition 34)

Properties: transitive

38. **Has Tangential Proper Spatial Part** A spatial relation between two Independent Continuants, x and y, in which x has y as a Proper Spatial Part and whose external boundaries are adjacent.

    Subrelation of: Has Proper Spatial Part (Definition 37)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Tangential Proper Spatial Part Of (Definition 35)

39. **Has Non-tangential Proper Spatial Part** A spatial relation between two Independent Continuants, x and y, in which x has y as a Proper Spatial Part and whose external boundaries are not adjacent.

    Subrelation of: Has Proper Spatial Part (Definition 37)

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Non-tangential Proper Spatial Part Of (Definition 36)

40. **Below** A spatial relation between two Independent Continuants, x and y, in which x has a location lower than the location of y in the vertical axis corresponding to the same frame of reference.

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Above (Definition 41)

41. **Above** A spatial relation between two Independent Continuants, x and y, in which x has a location higher than the location of y in the vertical axis corresponding to the same frame of reference.

    Domain: Independent Continuant (BFO)

    Image: Independent Continuant (BFO)

    Inverse of: Below (Definition 40)

42. **Left Of** A spatial relation between two Independent Continuants, x and y, in which x has a location to the east of the location of y in the horizontal axis corresponding to the same frame of reference.

Domain: Independent Continuant (BFO)
Image: Independent Continuant (BFO)
Inverse of: Right Of (Definition 43)

43. **Right Of** A spatial relation between two Independent Continuants, x and y, in which x has a location to the west of the location of y in the horizontal axis corresponding to the same frame of reference.
Domain: Independent Continuant (BFO)
Image: Independent Continuant (BFO)
Inverse of: Left Of (Definition 42)
44. **In Front Of** A spatial relation between two Independent Continuants, x and y, in which x has a location that makes it nearer than y in the longitudinal axis corresponding to the same frame of reference.
Domain: Independent Continuant (BFO)
Image: Independent Continuant (BFO)
Inverse of: Behind (Definition 45)
45. **Behind** A spatial relation between two Independent Continuants, x and y, in which y has a location that makes it nearer than x in the longitudinal axis corresponding to the same frame of reference.
Domain: Independent Continuant (BFO)
Image: Independent Continuant (BFO)
Inverse of: In Front Of (Definition 44)

5.5 Final considerations

The GeoReservoir ontology presented in this chapter aims to provide formal and clear definitions to support the description of deep-marine depositional system lithological and geometrical properties. As so, it is proposed to ameliorate several terminological issues in the domain by presenting semantical disambiguation, supporting the uniform data collection, and separating geological description from interpretation. Furthermore, the ontology ensures interoperability with other models and applications anchored at the same ontological framework by extending the BFO and GeoCore ontologies. The GeoReservoir ontology model implemented in OWL 2 is available in a public repository at <https://github.com/BDI-UFRGS/GeoReservoirOntology>.

The next step of the present work, which is reported in the next chapter, is to validate the terminology's adequacy and truthfulness by applying the GeoReservoir ontology in a geological case study.

6 ONTOLOGY APPLICATION

This chapter presents the GeoReservoir ontology application in a case study to evaluate the adequacy, truthfulness, and limitations in describing the domain entities. Reservoir geologists evaluated the ontology's competence on reservoir description using the restricted terminology to describe a geological occurrence of a deep-marine deposit. The source of data was not the real outcrop but a previous detailed study described in a professional technical report. This report is originally presented by Hodgson et al. (2011) and summarized in Section 2.2. The intention was to verify if the terminology could capture all the detailed descriptions and sketches as presented in the technical document. For simplicity, we selected a section of the technical report and not the entire description.

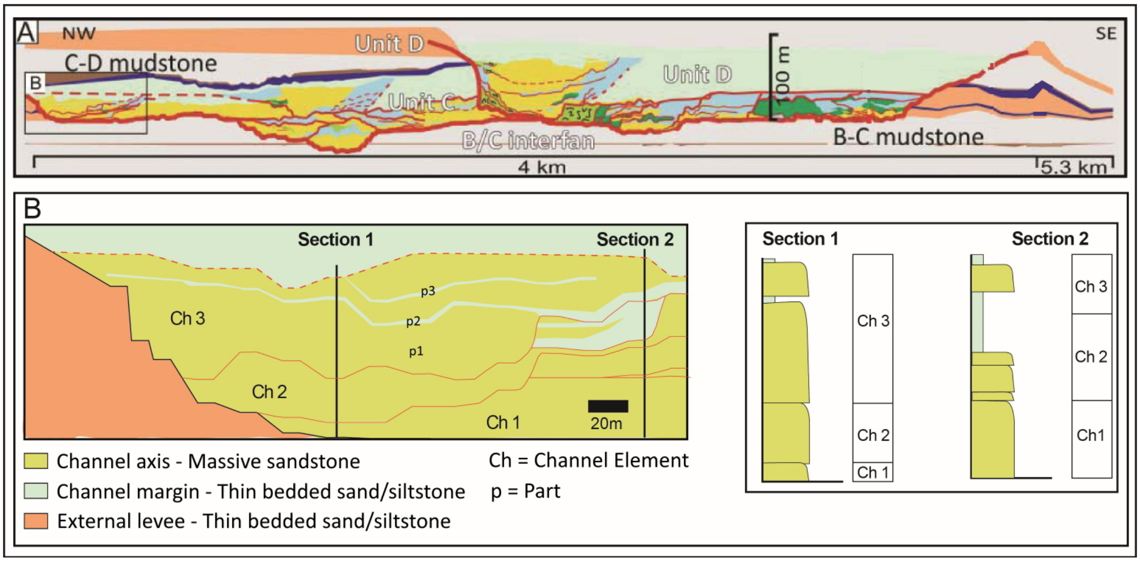
This section (Figure 6.1) features a few depositional bodies in different hierarchical scales, which are named "channel element" and "channel complex". Erosion surfaces are represented as solid red lines (or dashed red lines when they were inferred) and delimit the channel units. We chose to describe here, as a demonstration, the channel elements 1, 2, and 3 since they show the whole geometry in the outcrop section under analysis. We annotate them as Ch1, Ch2, and Ch3. The three other channels on the right side of the figure do not show their central axes in this particular, so we do not describe them here. Channel 3, even keeping its unity, also contains internal thin-bedded silt-sandstone layers preserved after some oscillation of the energy of deposition. We annotated these layers as p3 and p2, while the p1 is the main channel facies. The three parts compose together the Channel 3 unit. The three channel elements that we consider here compose the channel complex. All these bodies compose a larger-scale body, which is not entirely described here, in the "channel complex set" scale. There is also an "external levee" that is laterally adjacent to the channelized bodies. The facies are represented by distinct colors and named in the figure's legend.

Similar to the GeoReservoir ontology, this case study was implemented using Protégé and persisted in OWL 2 code. The implementation is available in the same Web repository that contains the GeoReservoir ontology[1]. In the following of this chapter, we describe the afore summarized subset using GeoReservoir terms. When a description refers to an entity defined in one of the ontologies described or reviewed in this thesis, this term is presented in an *italic* text followed by the referenced ontology or definition in parentheses.

---

[1] <https://github.com/BDI-UFRGS/GeoReservoirOntology>

Figure 6.1 – The subset of Karoo basin geological data used in this work's case study.



Source: modified from Hodgson et al. (2011) and Cicconeto et al. (2020)

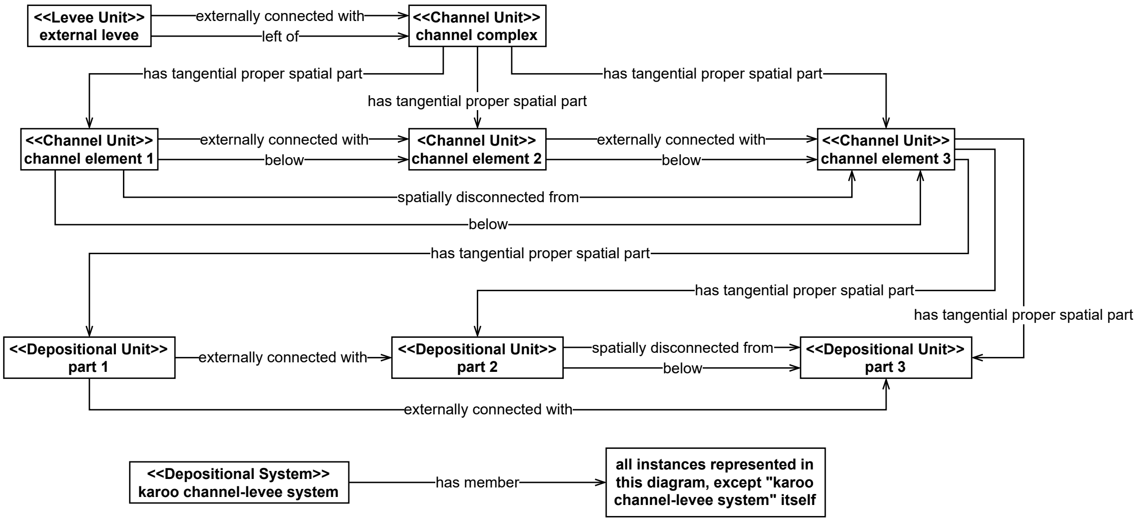
6.1 Material instances and their mereotopology

In terms of the GeoReservoir ontology, depositional bodies in all scales are instances of *Depositional Unit* (Definition 2). Channel elements and the complex are *Channel Unit* (Definition 3) instances, while the external levee is an instance of *Levee Unit* (Definition 5).

These units' spatial distribution model is illustrated in Figure 6.2. About the hierarchical relations, lower-order units are *Proper Spatial Parts Of* (Definition 34) higher-order units, e.g., the channel elements are *Tangential Proper Spatial Parts Of* (Definition 35) the channel complex. Depositional units in the same scale are always *Spatially Discrete From* (Definition 29) each other, being *Externally Connected With* (Definition 31) if they are adjacent or *Spatially Disconnected From* (Definition 30) otherwise. Furthermore, they are related by direction according to their relative distribution, e.g., *Left Of* (Definition 42) and *Below* (Definition 40). The external levee is *Externally Connected With* (Definition 31) the channel complex. Additionally, all these mereotopologically-related objects are *Members Of* (BFO) a Depositional System (Definition 10) instance.

As described in Section 4.6, it is necessary to have a frame of reference in order to represent directional spatial relations such as *Left Of* (Definition 42) and *Below* (Definition 40). In this work, we assume that depositional systems hold an intrinsic frame of reference fixed by the direction in which the units are deposited, which can be observed, for example, by the channel units' orientation. Although this frame of reference is rela-

Figure 6.2 – The spatial distribution model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols, while arrows represent relations.



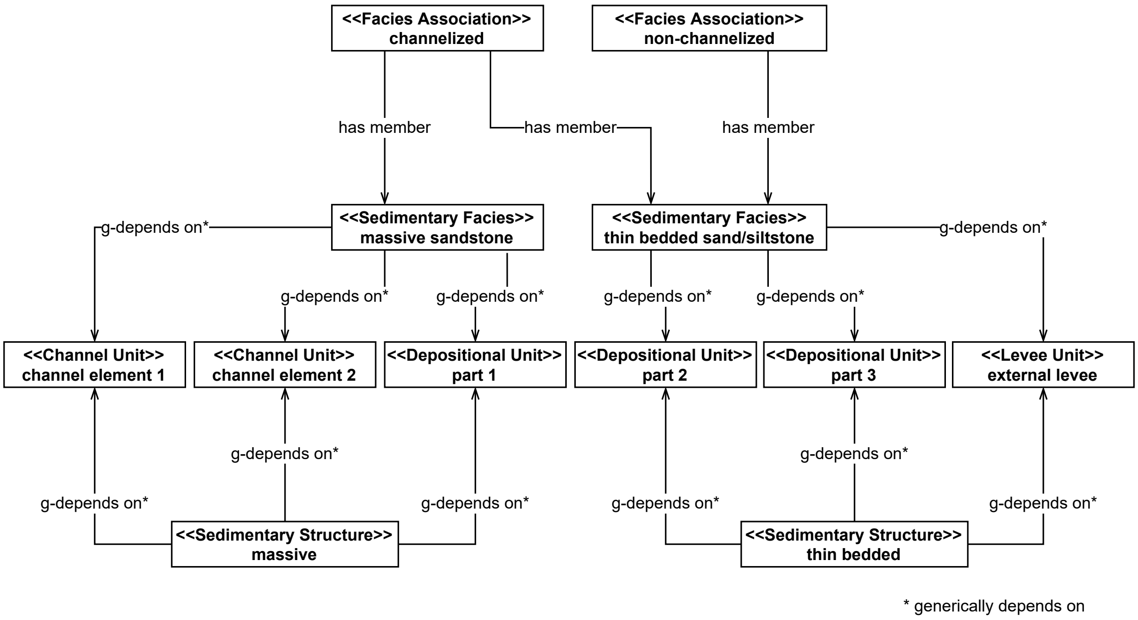
Source: the author

tive and can vary in different depositional systems, we assume that it does not affect the modeling of the geometry and architecture of the depositional system, still keeping the model's simplicity. Any future software application supported by the GeoReservoir ontology should include additional geographical reference information for the depositional system as a whole, which will allow calculating the real position of the geological bodies according to a geographical reference. This external frame of reference is essential for provenance interpretation, advanced geospatial reasoning, and data integration with other applications.

6.2 Sedimentary facies and lithological features

Regarding lithology, this Karoo subset features two distinct facies: Massive Sandstone and Thin Bedded Sand/Siltstone (Figure 6.3). Each of these two facies is an instance of *Sedimentary Facies* (Definition 24). We include the two facies associations defined by Hodgson et al. (2011), namely Channelized and Non-channelized, as *Facies Associations* (Definition 25). We relate the Channelized association with the two facies and the Non-channelized association with the Thin Bedded Sand/Siltstone facies, in both cases using the *Has Member* (BFO) relation. We also relate each facies to its respective *Depositional Units* (Definition 2) as colored in Figure 6.1: the Massive Sandstone facies *Generically Depends On* (BFO) the "Ch 1", "Ch 2", and "p1" units, while the Thin Bed-

Figure 6.3 – The sedimentary facies model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols, while arrows represent relations.



Source: the author

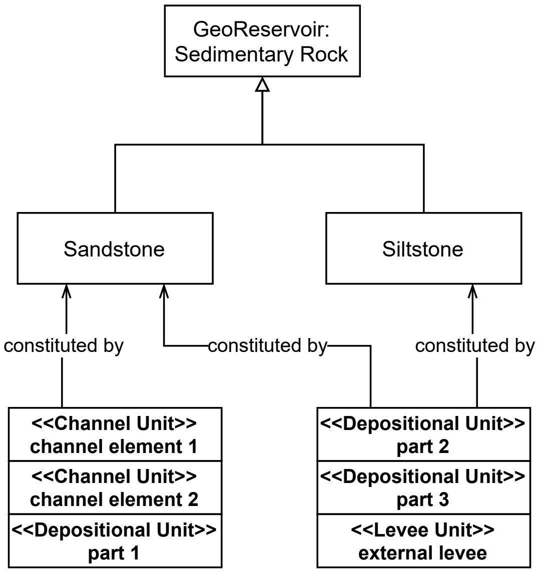
ded Sand/Siltstone facies *Generically Depends On* (BFO) the "p2", "p3", and "external levee" units.

The terms "thin bedded" and "massive" imply the presence of some sedimentary structures. Hence, we have two *Sedimentary Structure* (Definition 26) instances, which also *Generically Depend On* (BFO) the same units as their respective facies. Furthermore, the terms "sandstone" and "sand/siltstone" imply that the units in which the facies generically depend on are *Constituted By* (GeoCore) some subtypes of *Sedimentary Rock* (Definition 8): Sandstone in the former case, and both Sandstone and Siltstone in the latter case. In the future, we plan to create a complete taxonomy of rocks. For this work's purposes, we just include Sandstone and Siltstone in this case study scope (see Figure 6.4).

6.3 Final considerations

This case study demonstrates how we can use the GeoReservoir and its related ontologies to depict the descriptive features of a deep-marine channel-levee system. An ontology-controlled description restricts the geologists' freedom in using figurative language and imposes a scale and level of detail to the description. Even so, this approach

Figure 6.4 – The sedimentary rocks model of the Karoo channel-levee case study. Boxes represent instances with the entities that they instantiate between chevron symbols or entities when they don't have chevron symbols. Arrows with closed ending represent Subsumed By relations, while arrows with open endings represent other relations.



Source: the author

shows the significant advantage of creating a uniform record that a software application can search and compare. We consider that these gains outweigh the disadvantages since the approach allows to extract patterns, tendencies, and repetitions that may be not perceived in free textual reports. Moreover, the ontologies give us a sound and truthful terminology that we can use to precisely describe the geometrical and lithological aspects of deep-marine depositional systems.

A complete demonstration of the ontology capabilities would require representing a large set of channel, levee, and lobe deposits, showing how we could integrate and search over the data. This is the next step in the project in which this work is inserted, aiming to produce a useful library of deep-marine deposit data with the support of a description system. Nevertheless, we discuss the benefits and limitations of this thesis's approach in the next chapter.

7 DISCUSSION

In sedimentological studies, geologists describe the lithological features and spatial relations among the units based in a low-precision language complemented by figures and pictures of outcrops or seismic sections. The result is a non-uniform record with many language ambiguity and unclearness cases, offering low automatic processing capability. We present the GeoReservoir ontology as a solution to improve this scenario by defining a formal and precise domain terminology.

This chapter discusses the GeoReservoir's practical benefit in ameliorating the geological language's issues, considering the literature review, methodology, and case study presented in this work. Such as we did in Chapter 6, when we refer to an entity defined in one of the ontologies described or reviewed in this thesis, this term is presented in an *italic* text followed by the referenced ontology or definition in parentheses.

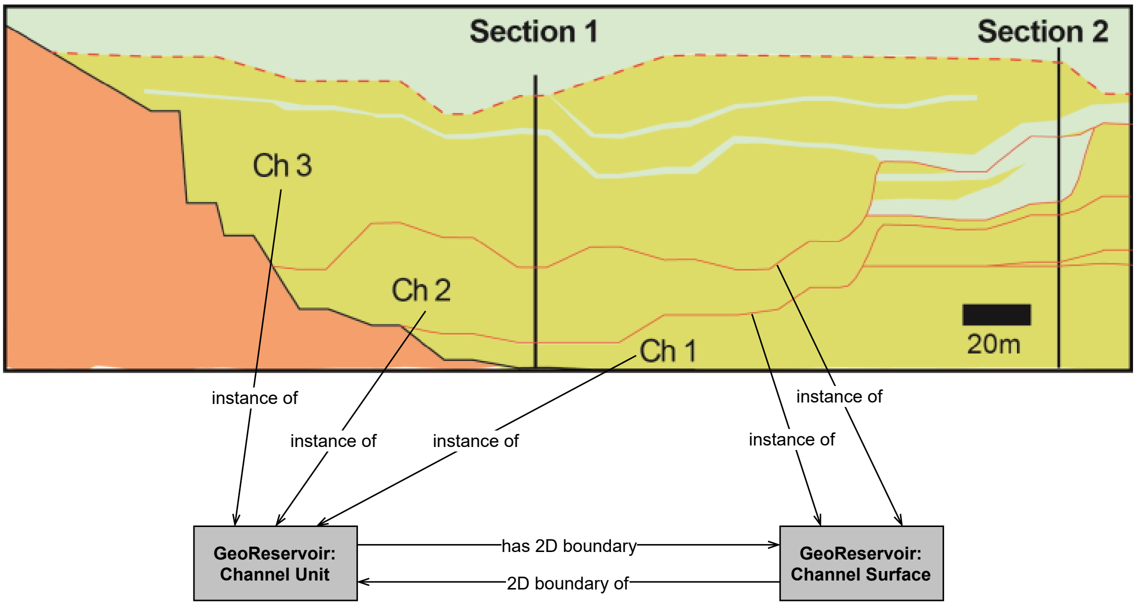
7.1 GeoReservoir: a tool for language analysis and disambiguation in deep-marine settings

The GeoReservoir ontology is not intended to define a new terminology but to disambiguate and formalize the existing geological terminology, as well as to enable its integration with other vocabularies formalized under the same ontological basis that we employed in this work. As so, the ontology clarifies the meaning variations of ambiguous terms, such as "channel".

As stated in Chapter 2, the term "channel" sometimes refers to the conduits where sedimentary flows pass by and leave sediments that become channel fills. This "channel fill" term leads us to the second meaning of the "channel" term, which is the concave-up shaped sedimentary unit constituted by these sediments. In this work, we define both meanings: the former as *Channel Surface* (Definition 11), and the latter as *Channel Unit* (Definition 3). Even though both terms are related (see Figure 7.1), when the term "channel" appears singly in a geological report, it is necessary to analyze the context in which it appears to decide what definition it fits.

For example, in the approach presented by Cullis et al. (2019), the term "channel" is associated with the "element" definition, which is the same kind of entity that the *Depositional Unit* (Definition 2) term refers to. As so, the "channel" term of Cullis et al. (2019) is equivalent to the *Channel Unit* (Definition 3) term. This meaning makes

Figure 7.1 – Figure that illustrates the two meanings of the “channel” term defined in this work. Boxes represent entities, while arrows represent relations.



Source: modified from Hodgson et al. (2011) and Cicconeto et al. (2020)

sense for their approach because the authors focus on developing a database for deep-marine deposit qualitative and quantitative analysis, in which the geologist is responsible for filtering the correct meaning of the data. However, if we imagine an application for geological process simulation that requires the *Channel Surface* (Definition 11) term, this definition is missing and can even be misunderstood as a Channel Unit. Being so, the GeoReservoir approach gains on reusability as the ontological analysis presented in this work addresses such terminological issues. The ontology vocabulary can be reused by data analysis, geological process simulation, and many other applications, further allowing intrinsic integrability between them or at least clarifying the vocabulary distinctions.

Another term that requires careful clarification is “facies”. As discussed in Chapters 3 and 5, a *Facies* (Definition 23) is an entity that can inhere in many *Geological Objects (GeoCore)*. In the approaches described by Lorenzatti et al. (2010) and Cullis et al. (2019), a facies is a unit contained in a single element, in conformance with most of the geological literature (WALKER; JAMES, 1992; DALRYMPLE, 2010). This distinction is relevant because it makes the approach described in this work naturally able to answer questions like “What depositional units present ‘thin bedded sand/silstone’ facies?”. The approaches presented by Lorenzatti et al. (2010) and Cullis et al. (2019), on the other hand, require further effort to quantitatively analyze the depositional unit’s characters in order to answer the same question.

7.2 Separation between geological description and process interpretation

One of this work's relevant contributions is to separate the terminology's genetic connotation from its descriptive meaning. As Geology studies the Earth's structure, composition, and the processes that generate and modify the Earth, geologists frequently define their vocabulary based on these genetic properties instead of what they observe on Earth.

For example, when defining the meaning of "depositional system", a geologist will postulate that it is the collection of the sedimentary bodies generated by a suite of related processes in a determined area. Although this definition is trivial for an experienced geologist, it does not clarify the descriptive features that make a collection of sedimentary bodies a depositional system, making it difficult for non-specialists and computer applications to understand how to distinguish these entities as these genetic features are highly interpretative.

This semantic separation between descriptive and interpretative terms is expected to ease computer application and artificial intelligence model development because it requests the geologists to bring the descriptive data that allow interpretation. Furthermore, it improves the possibilities of exploring Sedimentology and Stratigraphy as descriptive sciences, which strongly expands the reuse potential of collected data as time passes, since it is not contaminated by the evolution of geological theories.

Le Bouteiller et al. (2019) did a similar approach in separating and relating depositional features and sedimentary processes. However, the approach presented in this thesis is intended to cover all deep-marine deposits, while Le Bouteiller et al. (2019) is specialized in MTD features and processes. Our approach also features a higher level of formality as GeoReservoir and the other ontologies presented in this thesis define all relations' semantics.

These relations' semantics were successfully defined by Qu (2020). However, his work did not formally describe the descriptive properties of depositional features, relying primarily on process interpretation to compose the definitions. Such an approach holds a downside as it can hide the descriptive data that backed the interpretations.

An important limitation of this approach is that the explicit formalization of the descriptive features is not always completely possible. Let us consider the definition of *Sedimentary Rock* (Definition 8). According to this definition, a sedimentary rock is constituted by some collection of sedimentary grains or particles. However, this "sedimentary” term can be thought of as referring to the geological processes that carry these grains or particles to some deposition site. Actually, this term refers to the rock’s compositional and geometrical aspects that geologists identify in sedimentary rocks. However, there is no exact specification of these aspects in our knowledge, so that we cannot formalize them.

Another limitation is that we could not define some relevant terms found in the literature, such as MTD, turbidite, and debrite, since we focused on the descriptive terms’ definition. These terms have a high genetic connotation and need the process definitions to be properly formalized. We also did not define any hierarchical classification schema because of the many variations available in the literature. We considered that, for this initial research step, it is more important to define and validate the formal terminology for depositional units’ spatial distribution, as the hierarchical orders in any schema can be derived in the future from the spatial relations formalized in this work.

7.3 GeoReservoir’s reuse potential in other depositional environments

As presented in Chapter 3, in the approach presented by Qu (2020), the author separated his ontology into two main sections: general clastic depositional system and deep-marine clastic depositional system.

At the beginning of the work described in this thesis, we intended to divide the GeoReservoir ontology in the same sense. However, during the development, we noticed that the deep-marine environment does not have a specific terminology. A *Depositional Unit* (Definition 2), for instance, has the same essential properties and identity criteria regardless of depositional environment. In other words, the criteria that geologists use to differentiate a sedimentary unit in deep-marine, fluvial, or shallow-marine settings are the same, and the environment in which the unit is found does not add essential properties to it.

As stated by Smith (2006), an ontology should only define terms that are either expected that its intended users will need or required to fill some gap in order to complete the formal definitions. In our approach, we managed to build a single ontology aiming to describe deep-marine depositional systems but keeping its potential for reuse in other depositional environments.

7.4 Comparison with the GeoSciML and SWEET approaches

The models presented in the GeoSciML format (SIMONS et al., 2006; RICHARD et al., 2007) and the SWEET ontologies (RASKIN; PAN, 2005) provide useful standards for geological descriptions. In particular, the academy and industry widely adopted GeoSciML, which constitutes an important initial step for terminology definition in Geology. However, as discussed in Chapter 3 and by Garcia et al. (2020), these models have missed the ontological analysis as done in the state of the art conceptual modeling and applied in this thesis for terminological ambiguity reduction.

Furthermore, GeoReservoir defines a set of sedimentology-specific entities, diminishing the effort to align previous data with the ontologies and enriching the semantics associated with the aligned data. For example, it is possible to assign the *Levee Unit* (Definition 5) entity to some chunk of data, bringing along its formal definitions, such as being adjacent to some *Channel Unit* (Definition 3). In GeoSciML and SWEET, such a unit would be a GeologicUnit and a Rock Body respectively, configuring generic semantics.

8 CONCLUSION

As the geological language evolved over the past decades, a series of ambiguities and language pitfalls became a challenge in Geology, especially for those who intended to develop computer applications. Software applications need formal and explicit terminologies in order to process large volumes of data and discover useful knowledge. In deep-marine sedimentological studies, these challenges are potentialized by the fact that the data is very noisy because of the deposit's poor accessibility and wide scale.

In this thesis, we described the GeoReservoir ontology and guided its use. The GeoReservoir is an ontology intended to support the description of deep-marine depositional system geometrical and lithological features, providing a formal and precise language for deposits in outcrop and seismic section scales. Our approach separates descriptive properties from interpretative features, addressing the problem in which the geologists systematically rely on process interpretation to describe geological data.

We strongly relied on ontological foundations and methods to build the ontology, such as the metaproperties described by Guarino and Welty (2002a) and Guarino and Welty (2004). Moreover, we reused the conceptual basis built by the BFO and GeoCore ontologies, which provides a robust framework for the description of Geology in general. This approach allowed us to take advantage of previously analyzed and reviewed definitions, reducing the effort inherent to the development process of any conceptual model. Moreover, this approach will allow the integration of the GeoReservoir ontology with other models developed in the future using the same conceptual basis.

Although a formal specification cannot represent all the complexity of the nature in reality, the conceptual framework of ontologies helps us not to oversimplify this complexity and keep an interesting tradeoff. Furthermore, the work present in this thesis is an example of how ontological foundations can be used to model complex domains like Geology, which feature many cases of terminological ambiguity and unprecise definitions.

8.1 Future work

The present work is inserted in a larger project, which encompasses the development of a reservoir geometry and lithology database. In the future, the GeoReservoir ontology will support the application of artificial intelligence methods in order to assist petroleum exploration decisions. These methods include clustering similar deposits and depositional systems, discovering patterns and tendencies in reservoir occurrences, and others.

Other future work perspectives include developing a complete ontology of rocks, defining not only sedimentary rocks that are relevant for GeoReservoir but also other types of rocks, such as carbonate and salt, which are important for petroleum exploration and production. This perspective is likely to produce another research topic, as there exist many rock classifications in the literature, and further ontological analysis is required. Furthermore, the same kind of analysis can be applied to common facies classifications and sedimentary structures, producing standard and useful taxonomies and relations.

Computer-assisted geological process interpretation is also an interesting research topic to explore. To do so, an ontology of geological processes would be helpful, as well as the relations between the processes and the observational features defined in the GeoReservoir ontology. This topic is relevant because a complete series of ontologies supporting observation and interpretation would help geologists in studying the formation and evolution of petroleum reservoirs.

APPENDIX A – RESUMO EXPANDIDO EM PORTUGUÊS

Sistemas deposicionais marinhos profundos constituem os mais importantes tipos de reservatórios de hidrocarbonetos do mundo. Muito do conhecimento nestes depósitos sedimentares é extraído através da observação e interpretação de dados de sísmica, calibrados por dados de testemunhos e afloramentos, os quais possuem maior resolução e detalhamento. Estes métodos são utilizados por conta da inacessibilidade e grande escala dos corpos sedimentares que compõem estes depósitos.

Com o passar dos anos, geólogos identificaram características recorrentes nestes dados e desenvolveram modelos conceituais descrevendo os padrões mais comuns encontrados em depósitos marinhos profundos. Porém, variações de terminologia e interpretações distintas tornaram-se um problema no estudo dos depósitos e no desenvolvimento de aplicações de software com o objetivo de auxiliar na interpretação de grandes quantidades de dados geológicos. Tais aplicações demandam uma linguagem clara e não-ambígua.

Conforme já estudado em trabalhos anteriores que abordaram o desenvolvimento de aplicações e terminologias para a Geologia, a terminologia utilizada pelos geólogos apresenta muitos casos de ambiguidade e termos distintos que possuem o mesmo significado. Além disto, a Geologia é uma ciência que estuda a história da Terra e como a sua composição rochosa evoluiu ao longo desta história. Sendo assim, uma parte significativa da linguagem utilizada pelos geólogos é baseada em interpretações de como os processos geológicos ocorreram. Estas interpretações evoluem juntamente aos estudos científicos, tornando fundamental a correta definição e armazenamento dos dados descritivos que suportam estas interpretações.

Neste contexto, o objetivo da presente dissertação é descrever o desenvolvimento e a utilização da ontologia GeoReservoir, a qual foi projetada para abordar os problemas terminológicos deste domínio. A contribuição apresentada é relevante porque preenche lacunas apresentadas por trabalhos anteriores em modelagem conceitual e análise de dados para sistemas deposicionais marinhos profundos. Além disto, esta ontologia é o passo inicial de um projeto que pretende integrar de maneira uniforme e formal os dados de sistemas deposicionais ao redor do mundo, suportando um banco de dados e um sistema descritivo que permitirão a aplicação de técnicas de Inteligência Artificial para a descoberta de tendências e padrões em reservatórios de petróleo.

A metodologia utilizada prevê a reutilização das ontologias BFO (Basic Formal

Ontology) e GeoCore, além da aplicação de análise ontológica através da atribuição de metapropriedades sobre os termos definidos. A BFO é uma ontologia de fundamentação amplamente utilizada para a descrição de conhecimento em domínios científicos, enquanto a GeoCore é uma ontologia *core* que prevê a integração de todos os diversos sub-domínios da Geologia de Petróleo. Não obstante, a abordagem incorporou um conjunto de relações espaciais mereotopológicas e direcionais em uma ontologia distinta e independente, complementando o trabalho e alavancando o seu potencial de reúso. A terminologia foi desenvolvida em uma série de iterações que produziram protótipos evolutivos, os quais foram desenvolvidos e avaliados com o apoio de geólogos com experiência na indústria petrolífera.

Para avaliar a precisão, adequação e completude da terminologia, o trabalho inclui um estudo de caso utilizando dados de um sistema deposicional da Bacia do Karoo na África do Sul. Neste estudo, os dados foram descritos utilizando o vocabulário da ontologia GeoReservoir, demonstrando um exemplo de aplicação e comprovando a validade dos termos definidos.

O resultado do presente trabalho é uma ontologia completa, formal e clara para suportar a descrição dos aspectos observacionais de sistemas deposicionais marinho profundos. Como principais contribuições, a abordagem apresenta: (1) a desambiguação de diversos termos utilizados em Sedimentologia e Estratigrafia; (2) a separação entre termos descritivos e interpretativos da Geologia de Petróleo; (3) uma ontologia que pode ser reutilizada e expandida para outros ambientes deposicionais, como fluvial e marinho raso; (4) definição de uma semântica processável que pode ser associada a dados de relatórios geológicos, permitindo o raciocínio computadorizado sobre estes relatórios; e (5) a garantia de integrabilidade com outros modelos que reutilizem o mesmo arcabouço ontológico.

Como próximo passo do projeto no qual o presente trabalho está inserido, será realizada a integração de uma grande quantidade de relatórios geológicos com o suporte da ontologia aqui desenvolvida, gerando assim uma base de dados sobre a qual poderá ser realizada a descoberta de conhecimento através de métodos de mineração de dados e outras técnicas computacionais. Outras perspectivas de trabalhos futuros incluem: (1) o desenvolvimento de ontologias completas de rochas, fácies e estruturas sedimentares, os quais também requerem análise ontológica e desambiguação; e (2) o desenvolvimento de ontologias que definam termos de interpretação geológica tais como os processos sedimentares que hipoteticamente geram as entidades materiais descritas na ontologia GeoReservoir.

REFERENCES

ABEL, M.; PERRIN, M.; CARBONERA, J. L. Ontological analysis for information integration in geomodeling. **Earth Science Informatics**, v. 8, p. 21–36, mar. 2015.

ARNOTT, R. W. C. Deep-marine sediments and sedimentary systems. In: DALRYMPLE, R. W.; JAMES, N. P. (Ed.). **Facies models 4**. St. John's: Geological Association of Canada, 2010. p. 295–322.

ARP, R.; SMITH, B.; SPEAR, A. D. **Building Ontologies with Basic Formal Ontology**. Cambridge: The MIT Press, 2015.

BAAS, J. H.; MCCAFFREY, W. D.; KNIPE, R. J. The deep-water architecture knowledge base: Towards an objective comparison of deep-marine sedimentary systems. **Petroleum Geoscience**, v. 11, n. 4, p. 309–320, oct. 2005.

BOUMA, A. H. **Sedimentology of Some Flysch Deposits: A Graphic Approach to Facies Interpretation**. Amsterdam: Elsevier Publishing Company, 1962.

BRUHN, C. H. L. et al. Campos basin: Reservoir characterization and management – historical overview and future challenges. In: OFFSHORE TECHNOLOGY CONFERENCE, 2003, Houston, USA. [S.l.]: Offshore Technology Conference, 2003.

CARBONERA, J. L. **Raciocínio sobre conhecimento visual: Um estudo em estratigrafia sedimentar**. Dissertation (Master) — Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, 2012.

CARBONERA, J. L.; ABEL, M.; SCHERER, C. M. S. Visual interpretation of events in petroleum exploration: An approach supported by well-founded ontologies. **Expert Systems With Applications**, v. 42, n. 2, p. 2749–2763, apr. 2015.

CICCONETO, F. et al. A spatial relation ontology for deep-water depositional system description in geology. In: ONTOBRAS – ONTOLOGY RESEARCH IN BRAZIL, 13., 2020, Vitória, Brazil. [S.l.]: CEUR Workshop Proceedings, 2020. p. 35–47.

COHN, A. G. et al. Qualitative spatial representation and reasoning with the region connection calculus. **GeoInformatica**, v. 1, n. 3, p. 275–316, oct. 1997.

COHN, A. G.; RENZ, J. Qualitative spatial representation and reasoning. In: VAN HARMELEN, F.; LIFSCHITZ, V.; PORTER, B. (Ed.). **Handbook of Knowledge Representation**. [S.l.]: Elsevier, 2008. p. 551–596.

CULLIS, S. et al. Hierarchical classifications of the sedimentary architecture of deep-marine depositional systems. **Earth-Science Reviews**, v. 179, p. 38–71, apr. 2018.

CULLIS, S. et al. A database solution for the quantitative characterisation and comparison of deep-marine siliciclastic depositional systems. **Marine and Petroleum Geology**, v. 102, p. 321–339, apr. 2019.

DALRYMPLE, R. W. Interpreting sedimentary successions: Facies, facies analysis and facies models. In: DALRYMPLE, R. W.; JAMES, N. P. (Ed.). **Facies models 4**. St. John's: Geological Association of Canada, 2010. p. 3–18.

DELLA FAVERA, J. C. **Fundamentos de Estratigrafia Moderna**. Rio de Janeiro: Editora da Universidade do Estado do Rio de Janeiro, 2001.

FERNANDEZ-LOPEZ, M.; GOMEZ-PEREZ, A.; JURISTO, N. Methontology: From ontological art towards ontological engineering. In: AAAI SPRING SYMPOSIUM: ONTOLOGICAL ENGINEERING, 1997, Stanford. Menlo Park: AAAI Press, 1997. p. 33–40.

GANGEMI, A. et al. Sweetening ontologies with dolce. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT (EKAW), 13., 2002, Sigüenza, Spain. Berlin, Heidelberg: Springer, 2002. p. 166–181.

GARCIA, L. F. et al. The geocore ontology: A core ontology for general use in geology. **Computers & Geosciences**, v. 135, feb. 2020.

GARCIA, L. F. et al. What rocks are made of: Towards an ontological pattern for material constitution in the geological domain. In: ER: INTERNATIONAL CONFERENCE ON CONCEPTUAL MODELING, 38., 2019, Salvador. Cham: Springer, 2019. p. 275–286.

GUARINO, N. Formal ontology and information systems. In: FORMAL ONTOLOGY IN INFORMATION SYSTEMS (FOIS), 1., 1998, Trento. Amsterdam: IOS Press, 1998. p. 3–18.

GUARINO, N.; GUIZZARDI, G.; MYLOPOULOS, J. On the philosophical foundations of conceptual models. In: INTERNATIONAL CONFERENCE ON INFORMATION MODELLING AND KNOWLEDGE BASES (EJC), 29., 2019, Lappeenranta. Amsterdam: IOS Press, 2019. p. 1–15.

GUARINO, N.; OBERLE, D.; STAAB, S. What is an ontology? In: STAAB, S.; STUDER, R. (Ed.). **Handbook on Ontologies, Second Edition**. Berlin, Heidelberg: Springer-Verlag, 2009. p. 1–17.

GUARINO, N.; WELTY, C. A formal ontology of properties. In: INTERNATIONAL CONFERENCE ON KNOWLEDGE ENGINEERING AND KNOWLEDGE MANAGEMENT (EKAW), 12., 2000, Juan-les-Pins, France. Berlin, Heidelberg: Springer, 2002. p. 97–112.

GUARINO, N.; WELTY, C. Identity and subsumption. In: GREEN, R.; BEAN, C. A.; MYAENG, S. H. (Ed.). **The Semantics of Relationships**. Dordrecht: Springer, 2002. p. 111–126.

GUARINO, N.; WELTY, C. A. An overview of ontoclean. In: STAAB, S.; STUDER, R. (Ed.). **Handbook on Ontologies**. Berlin, Heidelberg: Springer, 2004. p. 151–171.

GUIZZARDI, G. **Ontological Foundations for Structural Conceptual Models**. Enschede: Centre for Telematics and Information Technology, University of Twente, 2005.

HODGSON, D. M. et al. Submarine slope degradation and aggradation and the stratigraphic evolution of channel–levee systems. **Journal of the Geological Society**, v. 168, n. 3, p. 625–628, may. 2011.

JARNA, A. et al. Dimensional geological mapping and modeling activities at the geological survey of norway. In: JOINT INTERNATIONAL GEOINFORMATION CONFERENCE, 2015, Kuala Lumpur, Malaysia. [S.l.], 2015.

KING, M. J. et al. Reservoir modeling: From rescue to resqml. **SPE Reservoir Evaluation & Engineering**, v. 15, n. 2, p. 127–138, apr. 2012.

LE BOUTEILLER, P. et al. A new conceptual methodology for interpretation of mass transport processes from seismic data. **Marine and Petroleum Geology**, v. 103, p. 438–455, may. 2019.

LORENZATTI, A. et al. Ontological primitives for visual knowledge. In: BRAZILIAN SYMPOSIUM ON ARTIFICIAL INTELLIGENCE (SBIA), 20., 2010, São Bernardo do Campo. Berlin, Heidelberg: Springer, 2010. p. 1–10.

MARTINSEN, O. J.; LIEN, T.; JACKSON, C. Cretaceous and palaeogene turbidite systems in the north sea and norwegian sea basins: Source, staging area and basin physiography controls on reservoir development. In: PETROLEUM GEOLOGY CONFERENCE, 6., 2003, London. Bath: The Geological Society Publishing House, 2005. p. 1147–1164.

MCHARGUE, T. et al. Architecture of turbidite channel systems on the continental slope: Patterns and predictions. **Marine and Petroleum Geology**, v. 28, n. 3, p. 728–743, mar. 2011.

MIDDLETON, G. V. Johannes walther's law of the correlation of facies. **Bulletin of the Geological Society of America**, v. 84, n. 3, p. 979–988, mar. 1973.

MORAES, M. A. S.; BLASKOVSKI, P. R.; PARAIZO, P. L. B. Arquitetura de reservatórios de águas profundas. **Boletim de Geociências da Petrobras**, v. 14, p. 7–25, 2006.

MUSEN, M. A. The protégé project: A look back and a look forward. **AI Matters**, v. 1, n. 4, p. 4–12, jun. 2015.

MUTTI, E.; NORMARK, W. R. Comparing examples of modern and ancient turbidite systems: Problems and concepts. In: LEGGETT, J. K.; ZUFFA, G. G. (Ed.). **Marine Clastic Sedimentology**. Dordrecht: Springer, 1987. p. 1–38.

NICHOLS, G. **Sedimentology and Stratigraphy, 2nd Edition**. Chichester: Wiley-Blackwell, 2009.

OBERLE, D. **Semantic Management of Middleware**. New York: Springer Science, 2006.

PICKERING, K. T. et al. Architectural element analysis of turbidite systems, and selected topical problems for sand-prone deep-water systems. In: PICKERING, K. T. et al. (Ed.). **Atlas of Deep Water Environments**. Dordrecht: Springer, 1995. p. 1–10.

POSAMENTIER, H. W.; MARTINSEN, O. J. The character and genesis of submarine mass-transport deposits: Insights from outcrop and 3d seismic data. In: SHIPP, R. C.; WEIMER, P.; POSAMENTIER, H. W. (Ed.). **Mass-Transport Deposits in Deepwater Settings**. Tulsa: SEPM Society for Sedimentary Geology, 2011. p. 7–38.

POSAMENTIER, H. W.; WALKER, R. G. Deep-water turbidites and submarine fans. In: POSAMENTIER, H. W.; WALKER, R. G. (Ed.). **Facies Models Revisited**. Tulsa: SEPM Society for Sedimentary Geology, 2006. p. 399–520.

QU, Y. **An ontology-based knowledge model for the deep-marine clastic depositional system**. Dissertation (Master) — Department of Geoscience, University of Oslo, Oslo, Norway, 2020.

RASKIN, R. G.; PAN, M. J. Knowledge representation in the semantic web for earth and environmental terminology (sweet). **Computers & Geosciences**, v. 31, n. 9, p. 1119–1125, nov. 2005.

RICHARD, S. M. et al. Geosciml – a gml application for geoscience information interchange. In: DIGITAL MAPPING TECHNIQUES '06, 2006, Columbus. Reston: U.S. Geological Survey, 2007. p. 47–59.

ROS, L. F. D.; GOLDBERG, K. Reservoir petrofacies: A tool for quality characterization and prediction. In: AAPG ANNUAL CONFERENCE AND EXHIBITION, 2007, Long Beach. [S.l.], 2007.

ROVETTO, R. The shape of shapes: An ontological exploration. In: INTERDISCIPLINARY WORKSHOP ON SHAPES, 1., 2011, Karlsruhe, Germany. [S.l.]: CEUR Workshop Proceedings, 2011.

SHANMUGAM, G. Ten turbidite myths. **Earth-Science Reviews**, v. 58, n. 3-4, p. 311–341, oct. 2002.

SIMONS, B. et al. Geosciml: Enabling the exchange of geological map data. **ASEG Extended Abstracts**, v. 2006, n. 1, p. 1–4, 2006.

SMITH, B. Against idiosyncrasy in ontology development. In: FORMAL ONTOLOGY IN INFORMATION SYSTEMS (FOIS), 4., 2006, Baltimore. Amsterdam: IOS Press, 2006. p. 15–26.

SPRAGUE, A. R. G. et al. Integrated slope channel depositional models: The key to successful prediction of reservoir presence and quality in offshore west africa. In: CIPM, CUARTO E-EXITEP, 2005, Veracruz. [S.l.], 2005.

STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: Principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1-2, p. 161–197, mar. 1998.

SULLIVAN, M. D. et al. An integrated approach to characterization and modeling of deep-water reservoirs, diana field, western gulf of mexico. In: GRAMMER, G. M.; HARRIS, P. M.; EBERLI, G. P. (Ed.). **Integration of Outcrop and Modern Analogs in Reservoir Modeling**. Tulsa: The American Association of Petroleum Geologists, 2004. p. 215–234.

SURE, Y.; STAAB, S.; STUDER, R. Ontology engineering methodology. In: STAAB, S.; STUDER, R. (Ed.). **Handbook on Ontologies, Second Edition**. Berlin, Heidelberg: Springer-Verlag, 2009. p. 135–152.

ULLMANN, S. **Semantics: An Introduction to the Science of Meaning**. Oxford: Basil Blackwell, 1972.

VERDONCK, M. et al. Comparing traditional conceptual modeling with ontology-driven conceptual modeling: An empirical study. **Information Systems**, v. 81, p. 92–103, mar. 2019.

WALKER, R. G.; JAMES, N. P. **Facies Models: Response to Sea Level Change**. St. John's: Geological Association of Canada, 1992.