Ontological foundations for petroleum application modeling

Ricardo Werlang[1,2], Mara Abel[2], Michel Perrin[3], Joel Luis Carbonera[2], Sandro Rama Fiorini[2]

[1]ENDEEPER – UFRGS Scientific and Technological Park - CEI - Institute of Informatics – Porto Alegre – RS – Brazil

[2]Institute of Informatics – Universidade Federal do Rio Grande do Sul (UFRGS) Porto Alegre – RS – Brazil

[3] Geosiris SA – 78112, Fourqueux, France

ricardo.werlang@endeeper.com, {rwerlang, marabel, jlcarbonera, srfiorini}@inf.ufrgs.br, michel.perrin@geosiris.com

Abstract. *Before designing a software application, a conceptual model is defined in order to simplify the geological reality and to help understanding it. Such a model does not represent the reality itself but the geoscientist's conceptualization. For a geologist, the reality that lies behind the words "horizon" or "fault" corresponds to material objects (a layer of small thickness, an interruption of a strata set) observed in the field or on well bore samples. Conversely, for a geophysicist, the same words refer to specific arrangements of pixels on a seismic image. These distinct views hinder of merging or interoperating conceptualizations operated by various types of geoscientists, who intend to refer to the same reality. In order to face this interoperability problem, many efforts have been made to create file format standards to transfer of information from one application to another. For keeping the original intended meanings, ontologies were used for expliciting the semantic of the models and for integrating the data and files generated in the various stages of the exploration chain. In this article, we apply methodologies based on foundational ontologies for analyzing the communication standard formats (LAS, WITSML, PRODML, and RESQML) most used in the construction of reservoir models. We show how the notions of identity, rigidity, essentiality and unity applied to ontological concepts lead the modeler to more precisely define the geological objects in the model. By making explicit the identity properties of the modeled objects, the modeler who applies data standards, can overcome the ambiguities of the geological terminology. Our analysis has helped us in producing an adequate conceptual model of the objects entering into the geological models (e.g., well, wellbore, horizon, fault, lithological unit, porosity properties). The nature and properties of the objects are made explicit in each format and this guides the mapping of these objects from one application to another.*

1. Introduction

A large amount of heterogeneous data is generated every day from multiple sources related with petroleum exploration activities, notably seismic, well, drilling, transportation, and marketing data. The petroleum industry depends on the efficient use of these data to the construction of computer models in order to reduce decision-making uncertainty and risk. Earth models, that are three- (3D) or four-dimension (4D) representations of data and interpretation concerning subsurface resources, are key tools for identifying and characterizing potential hydrocarbon reservoirs.

Earth models are developed by geoscientists who are responsible for evolving a hydrocarbon prospect through various stages of modeling. Their final goal is the building of a reservoir model, which will be used for simulating oil accumulation in the underground. Currently, this final model is connected to the original raw data by a long chain of successive interpretations realized by different professionals (e.g. geophysicists, geologists, reservoir engineers, etc.) who may have different conceptualizations about the modeled objects. These different professionals use heterogeneous data management environments that rest on various data representations and encoding conventions for dealing with the same information in different parts of the workflow.

An integration approach to deal with these heterogeneous sources is still a significant problem in the petroleum industry. Although the original geologic objects in reality are the same, the process of representing and building an information system fails in guaranteeing interoperability among the systems along the petroleum exploration process. *Interoperability* is the capability of providing to the user a uniform access to data or information that were created and managed by different information systems (Wache, Vogele et al., 2001), i.e., the ability of distinct information system applications and services to communicate, exchange and share data, information and knowledge in safe and efficient way. The first step in an information integration process is to determine if the sources contain semantically related information, that is, information related to the same or to similar entities of reality. Ontologies are being applied as an approach to overcome the problem of semantic heterogeneity since they allow one to clarify the meaning of the represented concepts according to the intention of the geologists.

In this work, we claim that interoperability among earth models built and manipulated by different professionals and systems can be achieved by making apparent the meaning of the geological objects represented in the models. We show that domain ontologies developed with support of the theoretical background of foundational ontologies show to be an adequate tool to clarify the semantics of geology concepts. We exemplify this capability by analyzing the communication standard formats most used in the modeling chain (LAS, WITSML, PRODML, and RESQML), searching for entities semantically related with the geological concepts described in ontologies for Geosciences, in order to make explicit the nature and properties of the geological objects found in each format. We aim to identify which entities in the model are able to be mapped from one application to another. We will restrict our analysis to the exploration steps, although we could later extend the study to cover production too.

We will first describe hereafter in Section 2 the use of ontologies in the conceptual modeling process and show how the use of properties from foundational ontologies can lead the modeler to more precisely define the geological objects in the model. In Section 3, we will introduce several domain ontologies that were developed in the domain of Geology and petroleum exploration in the last few years, with a special regard, in

Section 3.1, to the Basic Geology ontology (Mastella, 2010). We will then describe with some detail in Section 4 the analyzed communication standard formats, presenting the results of the analysis performed with LAS (Section 4.1), WITSML (Section 4.2) and RESQML (Section 4.3). In Section 5 we will present the conclusions of our work.

2. Conceptual modeling and foundational ontology

The modeling process in reservoir analysis is carried out by the geologist, who tries to capture and represent parts of the geological reality of the subsurface. In order to cope with this task, the geologist creates mental models of his comprehension of the subsurface geology. A mental model abstracts the relevant aspects of the objects, omitting those considered irrelevant for the task in hands. The *conceptualization* represented in an earth model is an abstract entity that only exists in the geologist's mind. Consequently, what is finally represented in the model is the geologist's idea about reality and not the reality itself. In order for a model to be understood by a community of professionals along with the modeling chain, the set of concepts in the geologist's mind and the corresponding external representation need to share the same respective meanings accepted by the community. These shared concepts and the agreed representations are what we call an *ontology*.

In the context of Computer Science, the term ontology usually designates an artifact that refers to some shared conceptualization and to its external representation in a computer processing language (Gruber, 1992). However, Nicholas Guarino (Guarino, 1998) reinforces that an ontology is a logical theory accounting for the *intended meaning* of a formal vocabulary utilized for representation. Indeed, ontology is the theory that helps us in keeping the correspondence between the geological reality and the models produced from this reality.

Domain ontologies refer to the concepts and meanings that a group of people needs to share when they communicate for solving problems in some restricted domain. Several domain ontologies were developed in the domain of Geology and petroleum exploration in the last few years, with the aim of defining standards for scientific communication, improving data integration and system interoperability. These efforts will be introduced in the Section 3 of this paper.

Foundational ontologies are meta-models that orient the way in which some conceptualization should be identified and modeled in a formal representation to build an artifact representing the domain ontology. The object of foundational ontologies is the study of *universals,* which are the intended meaning or the abstraction of the main properties that characterize a set of *individuals* recognized in the real world. Universals are equivalent to classes or objects in modeling languages, while individuals are their instances. Ontological analysis also provides an important distinction among the entities in terms of their behavior with refers to time: *endurants* and *perdurants*. Endurants correspond to individuals wholly present whenever they are present, i.e., *they are in time* (e.g., rock, reservoir and petroleum). Perdurants correspond to individuals composed of temporal parts, they *happen in time* in the sense that they extend in time accumulating temporal parts (e.g., deposition, oil migration or earthquake). Endurants have a relation of participation in perdurant occurrences, and these are commonly responsible for the creation of those. Also, foundational ontologies deal with some ontological properties drawn from Philosophy (e.g., identity, dependence, unity), i.e., with formal aspects of universal and individuals irrespective of their particular nature. The properties applied in this work are *essence, identity,* and *unity* (Guarino and Welty, 2009).

A property of an entity is *essential* to that entity if it must hold for it. It is a definitional property, that explains what makes some individual be an instance of a particular universal. For example, it is essential for a mineral being solid and having a crystalline structure at a normal temperature. Some essential properties are *rigid*, meaning that they are essential to all their possible instances. Some minerals, opal for instance, may not show a crystalline structure. However, if it ceases of being a solid, it will stop being a mineral. Being solid is a rigid property for a mineral that helps in identifying and defining it. Some properties are *anti-rigid*. For example, a *reservoir* can stop being considered a reservoir because it was fully depleted or became non-economic, but the concrete entity that was before a reservoir (the rock unit) is still there. Rigid properties are important because they identify the objects that are present in all geological models and can be mapped from one model to another. Also, they are the only properties that should be used for defining taxonomies (subsumption relationship) for structuring the model.

Identity refers to the issue of being able to recognize individual entities in the world as being the same (or different), and *unity* refers to the issue of being able to recognize all the parts that form an individual entity. Both ontological properties are crucial for geological interpretation and are used by geologists to interpret stratigraphic and structural correlations. Identity involves, for instance, the rigid properties of a rock unit that need to be considered for deciding, whether a body of rock corresponds or not to some geological unit although these two entities were possibly described in two different places (a wellbore and an outcrop, for example). Unity refers to the problem of describing the parts and boundaries of an object, so that we can decide what is part of the object, what is not, and under what conditions the object is a whole. For example, *water* is a concept whose instances were not wholes, except if they are limited by an instance of some other object, such as a cup, a bottle or a lake. In the same sense, rocks have no unity, since a rock can only be individualized by a core, a sample or a geological unit.

3. Standards and ontologies in geological modeling

Intense efforts were developed, during the last years, by various organizations (geological surveys, geoscience consortia, oil companies) for issuing codifications and formalizations of geological knowledge. According to the specific domains or activities that these organizations address, these can be classified in various categories.

Geological surveys are national or regional institutions, which are notably in charge of issuing geological maps. Their main goal is exchanging the information contained in field or laboratory observations and linking it with the objects that they intend to represent on a geological map. Between the geological surveys, a few models stand out. The NADM model - North American Geologic Map Data Model [1] (Richard, 2006) - and the derived GeoSciML[2] model are designed as ontologies for developing interoperable

[1] http://ngmdb.usgs.gov/www-nadm/
[2] http://www.geosciml.org/

geologic map-centered databases. The GeoSciML formalization is based in addition on the normative Geography Markup Language (GML) for the representation of geographic features and geometry. The GEON (Geosciences Network) [3] project is interested for its part in the problem of integrating geologic maps, whose source files contain geologic age or rock type information in the tables with different schemas and vocabularies.

Among the oil companies projects, stand out the IPP (Integrated Information Platform, (Omdal, 2006; Sandsmark and Mehta, 2004) project, comprising one of the largest ontologies ever developed for an industrial field for formalizing the terminology used in petroleum production. The project addresses many domains, such as subsea production equipment, seismic, drilling and logging, reservoir evaluation, but does not include earth sciences. Parts of the ontology are based on the ISO 15926 standard, for oil and gas *production* life-cycle data, which considerably differs from the oil and gas *exploration* life-cycle data, but they also include concepts issued from other terminologies.

For specific geoscience domains, many knowledge models were defined as well as specialized domain ontologies and upper level ontologies. Some of these models are to be found in (Sinha, 2006). In addition to those, stand out the ontologies proposed in: (Abel, 2001), for petrographic description of reservoir rocks; (Cox and Richard, 2005) and (Perrin et al., 2011) for geological time; and (Lorenzatti et al., 2009), for modeling of sedimentary structures and textual features of rocks. Some of them are proposed to be upper-level ontologies, which define more abstract objects in some domains, intended to support the organization and knowledge interchanging is some area of knowledge. That is the case of: the ontologies proposed by the project SWEET (Semantic Web for Earth and Environmental Terminology), which includes several thousand terms, spanning a broad extent of concepts from Earth system sciences and related concepts (Raskin and Pan, 2005); and the ontology proposed in (Mastella, 2010), the Basic Geology ontology, that describes and interconnects geological entities considered in reservoir modeling.

In Section 3.1 we will describe in more details the Basic Geology ontology, which is an ontology built around the concept *GeologicalObject* defining the main concepts that we have used in our analysis.

3.1. Basic Geology ontology

The Geological Object ontology summarizes a diversified amount of endurant geological objects that can be simple or complex (examples among many other are: a *stratified sedimentary unit,* a *reef,* a *diapir,* a *fault network*, etc.). Complex geological objects can be made of a various number of atomic geological objects. The Basic Geology Ontology was proposed with the important role of providing high level, general geological objects that helps in linking further knowledge and data representations. We will describe here a modified version of the Basic Ontology adherent to ontological principles of modeling. The original definition of the ontology can be found in the site of project E-WOK HUB (Environmental Web Ontology

[3] http://www.geongrid.org/

Knowledge Hub).[4] The RDFS/OWL version of the ontologies related to this project can be downloaded from this same site.[5]

There are two kinds of elementary geological objects (Mastella, 2010):

- 2D objects, corresponding to **Geological Boundaries**, such as the erosion surface *E*, the fault *F* and the upper and lower boundaries *bu* and *bl* on Figure 01;
- 3D objects, which are **Geological Units**, such as the sedimentary unit *U* limited by the boundaries *bu* and *bl* on Figure 01. Geological Unit is a volume of continuous geological matter limited by one or several Geological Boundaries.

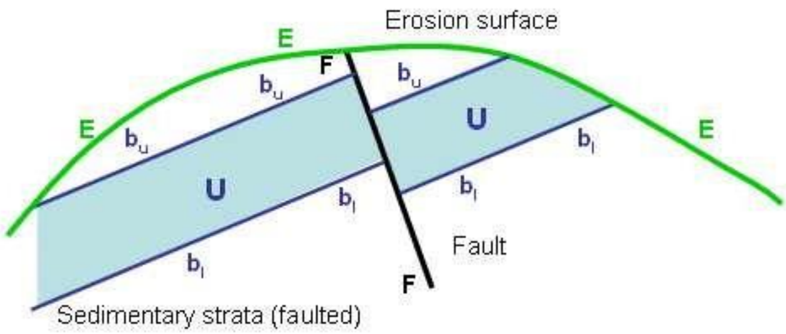

Figure 01: Geological objects: Erosion surface *E*, fault *F* and Sedimentary strata unit *U*, constituted of volume and boundaries *bu* and *bl* (Mastella, 2010).

Geological boundaries, units and *structures* are rigid objects that preserve the unity property. This means that their instances in reality cannot stop being instances of these objects or they would disappear. Along with the *substance* that specializes *rock,* these are the most important concepts for data. We will detail the ontological properties and relationships of these objects in the way they are usually conceived by geologists.

Rocks are the substance that makes possible geological units to exist in space. The identity of a unit is given by its spatial properties (e.g., location, size, format). A geological unit provides unity to an instance of rock, but not identity. The identity of a rock is given by its internal properties, such as, composition, texture, fabric. A sub-class of a substance needs to preserve the same identity properties. A taxonomy of the concept rock would include carbonate and siliciclastic rocks, but rock cannot subsume the concept *sill* (a horizontal intrusion of rock) as a sub-class of rock, since the identity of sill is not provided by a lithology, but by some lithological unit (considering properties like size and position).

The relationship between a geological unit and a rock is called constitution. When we say that a geological unit has some property like *granulometry*, we are meaning that the substance that constitutes a lithological unit has this property, but not the unit itself. When we assert that a lake is salty, we indeed refer to the property of the water that is inside the lake. Taking care of preserving the independence and the specific properties of two concepts collocated in space such as *rock* and *geological unit* will allow us to integrate them within other applications in a much easier way.

---

[4] http://www.inria.fr/sophia/edelweiss/projects/ewok/

[5] http://www-sop.inria.fr/edelweiss/projects/ewok/ontologyview/ontologies.html/

In contrast with geological units, *geological boundaries* have no internal constitution. However they are still rigid concepts that are an inseparable part of a geological unit (according to the classification of (Guizzardi, 2005)). Geological structures are also rigid objects and inseparable parts of geological units, whose identity is defined by geometric internal and external properties.

These few objects - boundaries, structures, units and substance - are omnipresent in earth models and are specialized, extended and derived to represent all the different aspects that some model aims to emphasize. It is important to stress that only these objects and the objects that they subsume (their subclasses) have instances, which means that any instance manipulated by some software application can be mapped to these few objects and further integrated. That statement is the basis for our approach.

Other concepts that will be further analyzed in this work have an anti-rigid identity, which means that they can preserve their existence when they lose their identity. Such concepts are existentially dependent of another concept to exist. A reservoir, a source-rock and a trap are examples of roles played by rock units and geological structures, without configuring a specialization of rock unit. Any instance of reservoir or a source rock is necessarily an instance of a *GeologicalUnit*, and an instance of a trap is necessarily an instance of *GeologicalStructure*.

Besides the endurant universals described above, the Basic Geology Ontology describe the perdurant concept *GeologicalEvent* that instantiates individuals that are composed by temporal parts and have no physical representations in reality. A geological event may consist of a single geological process (e.g. the deposition of a sedimentary unit) or be composed of multiple geological processes (e.g., a graben filled with syntectonic sediments deposited while the faults limiting the graben are in movement). The various specializations of *GeologicalProcess*, corresponding to creation, destruction or transformation of geological matter, are detailed in the Geological Process sub-ontology.

The Basic Geology Ontology is complemented by universals that have no identity and are existentially dependent of other objects. They describe the *properties* that characterize the concepts and are usually of special interest in reservoir characterization. Attribute domains are associated to conceptual spaces and integrated through rigid objects. For a question of space, we will not discuss here how to deal with interoperability in the case of attributes.

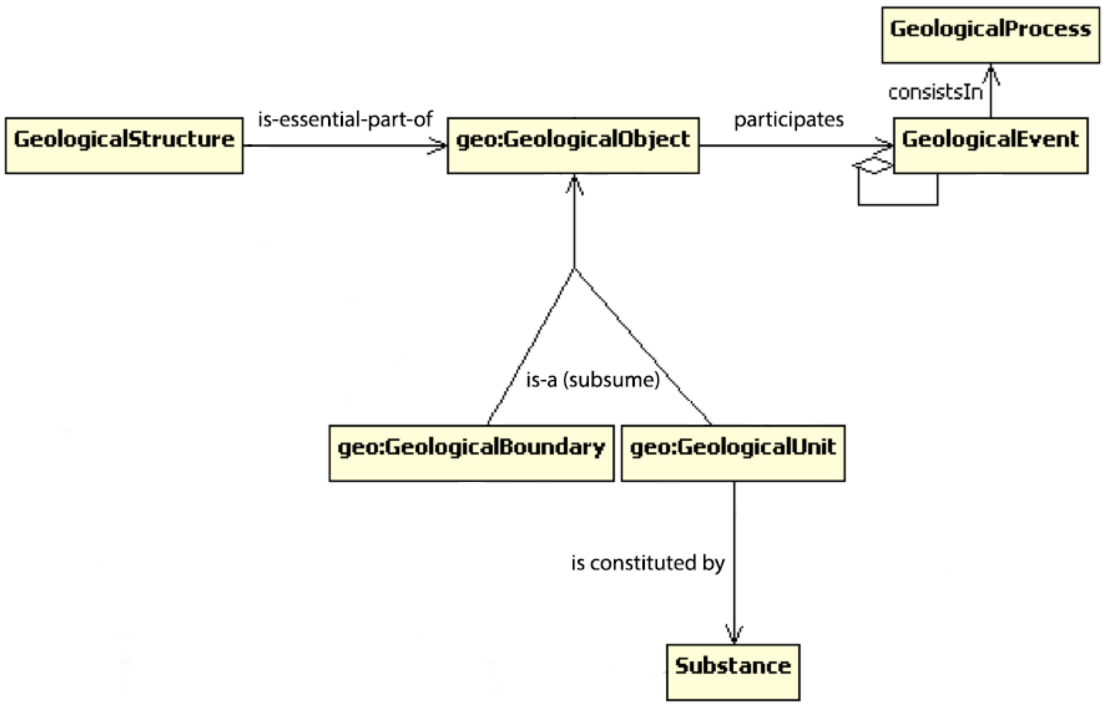

Figure 02: The top level Basic Geology ontology modified from (Mastella, 2010) to describe the ontological meaning of relationships.

4. Communication standard formats

A huge amount of data formats and standards are being used for data exchange in reservoir characterization models. Usually, these data are stored in different file formats and represented in different formats.

The Geological File Ontology defines the file formats used for storing information related to geological data. Basically, all data are stored in textual documents or classified as one-dimensional, two-dimensional, or three-dimensional documents, according to the way in which the data are stored (one, two or three dimensional arrays). Among those formats, the LAS and the WITSML standards store geophysical log data in one-dimensional documents. The RESQML standard, proposed by the Energistics Consortium[6] stores its data in three-dimensional documents, divided by three model types: '*Reservoir*', '*Structural*' and '*Stratigraphic*'. Figure 03 represents a part of the Geological File Ontology with the concepts related with the RESQML standard.

PRODML (PRODuction-ML)[7] is an industry standard which supports data exchange representing the flow of fluids from the point they enter into the wellbore to the point of custody transfer, together with production operation workflows, in a vendor-neutral, open format. Since this model doesn't refer to geological objects, we will not analyze it in this paper.

[6] http://www.energistics.org/

[7] http://www.energistics.org/production/prodml-standards/

In Section 4.1, we will present the analysis accomplished about the LAS standard, describing the standard itself and its data set sections. In Section 4.2, we present the analysis realized with the WITSML standard, emphasizing the data objects with possible mapping with the summarized ontologies. In the Section 4.3, we present the analysis of RESQML standard.

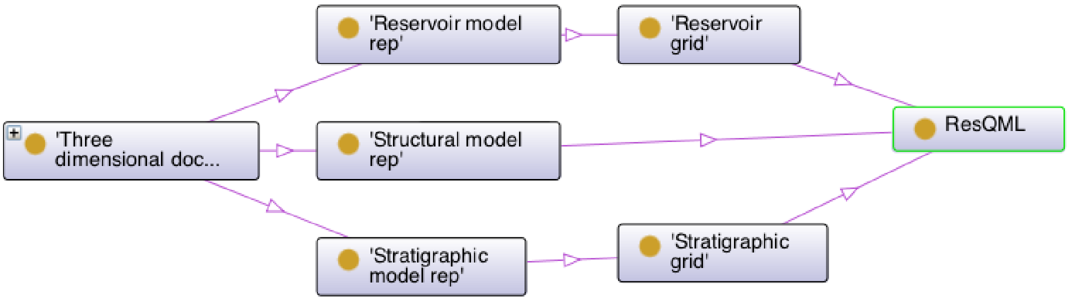

Figure 03: The concepts related with the RESQML standard in Geological File Ontology.

4.1. LAS

The *Log ASCII Standard* (LAS) format began with the aim for a simple format for exchanging well log data. The worldwide acceptance of LAS proved the need for such a format. The Canadian Well Logging Society[8] introduced the LAS standard, in 1989, which is now in the third version. LAS 3.0 clarifies several of the poorly defined specifications of LAS 2.0 and provides expanded data storage capabilities, but has seen limited implementation.

LAS 3.0 files are divided into logical sections. Each section begins with a title line, which is marked with a tilde ("~") at the beginning of the line. Sections contain lines where data is described and/or stored. There are several types of sections and several types of lines within sections. The LAS 3.0 standard defines which combinations of sections must exist in LAS files, and in which order. For example, the **~Version** and **~Well** section must exist in that order in any file of the version 3.0 LAS. As in LAS version 2.0, only one well is described within a single file.

As we saw before in the Geological File Ontology, the LAS standard stores 'Well Log' data in one-dimensional ASCII documents, meaning that the data are stored in one-dimensional arrays. However, when representing information through arrays, LAS version 3.0 allows that these documents have data stored as one, two or three-dimensional arrays, something that was not possible until this version. The data are usually indexed to depth or time, but may be presented as discrete measurements if required. Also, data are grouped by type into related sections, as they relate to the well where the data was acquired. Types include depth and time indexing Logging, Core, Inclinometry, Drilling, Formation tops, test data, user defined types, etc.

We analyzed each data section set, including all parameter and definition data found for these data sections, searching for possible mappings with the Basic Geology ontology concepts. The concepts in LAS format map to three geological concepts: *GeologicalUnit*, *GeologicalBoundary* and *Substance*. The Basic Geology Ontology

[8] http://www.cwls.org/las_info.php

does not describe a concept to represent the different bodies of rocks created by human actions, such as cores and samples. We assume that a concept in a distinct domain ontology related to well development would subsumes the concepts *Core* and *Sample* utilized in our classification. We named it *SampleOfRock* to be mentioned in this paper and it represents all partial exposition of a rock, spatially delimited by human action, such as core, samples and thin-section. Also, we consider that the concept *Rock* is a specialization of *Substance*. The result of the analysis is presented in the Table 01.

Data Section	Description	Ontology Concept
Version	Identifies information belonging to the file itself.	Not a Geology concept.
Well	Contains data that uniquely identifies the Well bore data.	Not a Geology concept. Well development ontology not described here
Log	Includes information about log curves, like gamma ray. We can consider this information as a property of the rock around the borehole.	Not a Geology concept. Well development ontology not described here. The properties of this concept has relationship with Rock, which is specialization of Substance
Core	Includes information about the core, like the core source and type, primary formation cored, etc. Also, core properties like porosity, permeability, oil and water volume and saturation .	Needs to be mapped to two different concepts: the properties *core source*, *core type*, and *primary formation cored* are properties of Core; the properties porosity, permeability, oil and water volume and saturation are properties of Rock, which is an specialization of Substance.
Inclinometry	Information use for dip calculation, like Borehole Azimuth and Deviation.	It is related to a property attached to Geological Structure
Drilling	Contains data about the process of drilling, including the utilized rig and details about the process itself.	Not a Geology concept
Tops	Information about the formation tops. Formation tops are the depths in a well where formations are found in the subsurface.	*GeologicalBoundary* specialization
Test	Includes common test data annotations normally	It is a property of

	associated with petrophysical analysis presentations.	Rock, which is a specialization of Substance

Table 01: The result of the analysis realized with LAS version 3.0 data sections.

For the analysis of the LAS standard, we considered the description of each data set section and the parameter data and definition data described in the *LAS ASCII Standard document – File Structures*. Our next analysis shows the classification of the geological concepts found in the selected data set sections, according to the properties of identity, rigidity, essentiality and unity. The analyzed concepts were: *formation, core, oil* and *water.*

A lithological unit is a body of rock that is sufficiently distinctive and continuous for being mapped. The concept *LithologicalUnit* offers the principle of identity for its instances, since geologists are able to distinguish one instance from another by observing their visual characteristics. Also, this concept provides the principle of unity for their instances. This principle involves the observation of discontinuities in the visual characteristics of some body of rock, allowing the geologist to distinguish different lithological units. Furthermore, this concept is rigid, since its instances cannot fail to be so, unless it ceases to exist.

If the age and stratigraphic position of a lithological unit can be determined, this unit is referred as a formation and is given a name in order to support stratigraphic correlation. Even if the identity of the concept *Formation* is defined by rigid properties (lithological properties age), there are no instances of the concept *Formation* that are not instances of *LithologicalUnit*. Actually, a formation does not exist until some lithological unit is identified as such by geologists. This shows an existential dependence that is not acceptable for a rigid concept. Consequently, *Formation* is an anti-rigid concept and plays a "role" for some lithological unit. So, in terms of integration, a formation should be mapped to a lithological unit in other models.

A *Core* is a cylindrical sample extracted from a lithological unit that is being drilled. It is composed by one or more types of rocks. It is a specialization of *SampleOfROck* and it has a relationship of constitution with *Rock*. It is a rigid concept that preserves identity and unity.

Among the analyzed concepts, the concepts *Oil* and *Water* do not have unity, since they can only be individualized by other concepts (like a barrel of oil, a bottle of water, or a sample of rock). They are rigid concepts: they cannot stop being water and oil, without stopping to exist. But they are both uncountable objects, whose instances need to be individuated through a container, for instance, a reservoir. Water and oil are interesting for modeling in view of their internal properties. When present in some application, rigid and not unified objects preserve a relation of container or location with some rigid and whole entity (e.g., *oil is present within a reservoir)*. So, rigid and not unified objects are identified by their internal properties rather than by their spatial characteristics.

Less intuitively, the concept of *Rock* (specialization of *Substance*) in Geology is similar in ontological terms to *Oil* and *Water*. *Rock* is rigid concept with no unity property. It is individuated by some other entity, typically a *lithological unit*, a *core*, a *sample* or a *layer* in outcrop, with whom rocks have the ontological relation of *constitution*.

The result of this analysis is presented in the Table 02. The signs + indicates that the concept preserve that property, while – means the opposite.

Type	Supplies Identity (O)	Carries Identity (I)	Rigidity (R)	Relational Dependence (D)	Unity (U)
Oil	+	+	+	-	-
Water	+	+	+	-	-
Rock	+	+	+	-	-
Lithological Unit	+	+	+	-	+
Core	+	+	+	-	+
Log	-	-	-	+	-
Inclinometry	-	-	-	+	-
Tops	+	+	+	-	+
Test	-	-	-	+	-

Table 02: The result of the classification realized with LAS types.

4.2. WITSML

The Wellsite Information Transfer Standard Markup Language (WITSML) [9] is a standard used for sending well site information in an XML document format developed to promote the right-time, seamless flow of well data between operators and service companies, as well as regulatory agencies, in order to speed and enhance decision-making and reporting.

A WITSML document consists in one or more complete WITSML data-objects that correspond to a logical representation and organization of the data items associated with the major components and operations involved in well drilling —such as well, wellbore or log— represented as an XML document, which is essentially a text string. So each WITSML data-object is defined by an XML schema and is its own document. Each schema defines a set of data that can be transmitted within a single XML document and represents a cohesive subset (e.g. well, wellbore, rig, etc.) of an overall logical schema related to a single domain (well). Data object schemas contain attributes, elements, and included component sub-schemas. Figure 05 represents the WITSML data objects relationships.

We have chosen for analysis the components that have relation with a *GeologicalUnit* or a *GeologicalBoundary,* as well as, those that can be mapped with the Basic Geology ontology and its sub-ontologies. The WITSML standard enumerates 113 types with valid values for those elements. Analyzing these enumerated types, we identified 10 types whose values can be mapped to concepts from the Basic Geology ontology. A brief description of the identified type names is listed in the Table 03.

[9] http://www.energistics.org/drilling-completions-interventions/witsml-standards

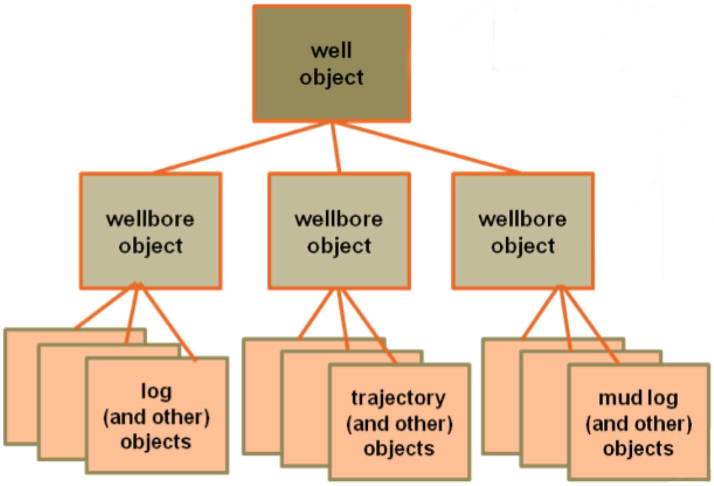

Figure 05: The WITSML data objects relationships.[10]

Type Name	Description	Ontology Concept
Chronostratigraphy Unit	Specifies the unit of chronostratigraphy.	Time ontology not described here
LithologyType	The geological name for the type of lithology from the enum table listing a subset of the OneGeology / CGI defined formation types.	Rock, which is an specialization of Substance
LithostratigraphyUnit	Specifies the lithostratigraphy unit.	GeologicalUnit
MatrixCementType	Lithology matrix/cement description.	Property of Rock, which is an specialization of Substance
MudClass	Defines the class of a drilling fluid.	Substance
MudSubClass	Mud Subtype at event occurrence.	Substance
QualifierType	Represent the type of qualifier in lithology.	Property of Rock, which is an specialization of Substance
StimFluidSubtype	The fluid sub type.	Substance
StimFluidType	The fluid used for some stage of the stimulation job.	Substance
WellFluid	The type of fluid being produced from or injected into a well facility.	Substance

Table 03: The selected type names and their descriptions.

We also analyzed all WITSML data objects (27 data objects) searching for possible mappings not found with the previously analysis. The data objects with elements that can be mapped are described in the Table 04.

[10] Energistics and the WITSML SIG, 2012: WITSML STORE Application Program Interface (API) Version 1.4.1

Data Object	Description	Ontology Concept
bhaRun	Used to capture information about one run of the drill string into and out of the hole.	- Substance - Hydrogeology
convCore	Used to capture information about a conventional core. This includes a description of the geology and lithology over the interval.	- GeologicalUnits - Core - Properties of Rock as a specialization of Substance
drillReport	Used to capture a daily drilling report focused on reporting from the operator to partners or to a governmental agency.	- Substance - Production concepts not described here
fluidsReport	Used to capture an analysis of the drilling mud.	- Substance - Production concepts not described here
formationMarker	Used to capture information about a geologic formation that was encountered in a wellbore.	- GeologicalUnits - GeologicalTime
mudLog	Used to capture information in a mud log. This includes lithologic information derived from cuttings encountered on a shaker screen at a site.	- GeologicalUnits - GeologicalTime - Properties of Rock as a specialization of Substance
opsReport	Used to capture a daily drilling report focused on reporting from the service company to the operator.	- Substance
sidewallCore	Used to capture information about a core from the side of a borehole.	- Properties of Rock as a specialization of Substance
stimJob	Used to capture a post job summary report about a stimulation (fracture) job.	- Properties of Rock as a specialization of Substance
well	Used to capture the general information about a well. This might sometimes be called a well header.	- Not a Geology concept. - Production ontology not described here

Table 04: The result of the analysis realized with WITSML data objects.

As we did with LAS types, we classified the geological concepts found in the selected WITSML data objects according to the presented properties from foundational ontologies. The analyzed concepts are: oil, water, gas, rock, cuttings samples, geology interval, formation, core, mud, well and wellbore. The concepts oil, water, rock, formation and core are present in the LAS format and were already classified in Section 4.1 (result presented in the Table 02). The concept *GeologyInterval* is a definition related to a time ontology. In brief, a *GeologyInterval* is a rigid concept with its own identity that preserves unity. The concepts that remain to be classified are thus: *gas*, *cutting sample*, *mud*, *well*, *wellbore* and *GeochronologicUnit*.

The type *gas* has no unity like the rock, oil and water types, since all these types can only be individualized by other concepts. We consider gas as a specialization of Substance. *Cuttings samples* are samples of rocks that were extracted from the wellbore during the drilling process. So, this type offers the identity and unity principles for its instances. Also, it is a rigid concept, since its instances cannot fail to be so, unless it ceases to exist. We classify it as a specialization of the proposed concept *SampleOfRock*. *Mud* is a term that is generally synonymous with drilling fluid and that encompasses most fluids used in hydrocarbon drilling operations, especially fluids that contain significant amounts of suspended solids (cuttings samples or samples of rocks), emulsified water or oil. Mud has not unity and includes all types of water-base, oil-base and synthetic-base drilling fluids, which means that its instances are the instances of several other rigid concepts which have their own identity. Mud corresponds to the ontological type *category* (rigid concepts that have not proper identity and group other instances using some property of interest). The types *Well* and *Wellbore* are rigid and offer unity principles, but they will not be further analyzed here because they are not Geology concepts. Finally, the concept *GeochronologicUnit* is a rigid concept having identity and unity, which corresponds to a set of *GeologicalUnit* that are grouped since they correspond to a given *GeologicalTimeInterval*. The results for the concepts that were not previously classified are presented in Table 05. The signs + indicates that the concept preserve that property, while – means the opposite. The sign ~ indicates that for some instances the property is preserved, but for others no.

Type	Supplies Identity (O)	Carries Identity (I)	Rigidity (R)	Relational Dependence (D)	Unity (U)
Cutting sample	+	+	+	-	-
Gas	+	+	+	-	-
Mud	-	+	+	-	-
Well	-	+	+	-	+
Wellbore	-	+	+	-	+
GeochronologicUnit	+	+	+	-	+

Table 05: The result of the classification realized with LAS types.

4.3. RESQML

RESQML is an industry initiative to provide open, non-proprietary data exchange standards for reservoir characterization, earth and reservoir models proposed by Energistics's Special Interest Groups (SIG)[11]. RESQML is an XML-based data exchange standard that helps addressing the data-incompatibility and data-integrity challenges faced by professionals in petroleum industry when using the multiple software technologies required along the entire subsurface workflow, for analysis, interpretation, modeling, and simulation.

The release available for the public is RESQML V1.1, which is strictly focused on data exchange. However, recently, a new version is being developed, RESQML V2 (Endres et al., 2013; Deny et al., 2013). The key goal of this version is to provide a mechanism for transferring relationship information (between data-objects, such as faults, horizons and grids), while continuing to expand the fundamental data types within the standard, for example, unstructured simulation grids.

[11] www.energistics.org/reservoir/resqml-sig

We analyzed the data types referenced by elements and attributes in RESQML V1.1 and V2 data objects and component schemas that have restricted content, as we did with WITSML data types. Among the 6 data types analyzed below, we do not identify any possible mapping with Basic Geology ontology and its sub-ontologies. However, it is important to note that the RESQML V1.1 includes integration with just a few WITSML data objects (well, wellbore, trajectory, formation marker, and well log). The full integration of RESQML V2 and WITSML V1.4.1.1 is still under development by the time we wrote this paper. As presented previously in the WITSML section, some of the concepts that already have integration with RESQML V1.1 (formation maker and well) can be mapped with the presented ontology.

After that, we analyzed all RESQML V1.1 data objects (6 data objects) searching for possibly mappings. The data objects with elements that can be mapped are described in the Table 06.

Data Object	Description	Ontology Concept
areaOfInterest	A GML based feature representing the area of interest. This contains a mandatory bounding box and an optional polygon outline.	- *GeologicalLocation*
interfaceFeatureSet	A set of interface features and their geometry. These by themselves do not represent a coherent model. They build blocks[12] that can be used to create a model.	- *GeologicalBoundary* - *GeologicalObject* - Geological Unit
griddedVolumeSet	A container for 3D grids and related properties.	- *GeologicalObject* - Properties of Geological Unit

Table 06: The result of the analysis realized with RESQML data objects.

The RESQML V1.1 type *InterfaceFeatureSet* abstracts the several elements that instantiate to limits in the 3D Model, such as *fault*. The element fault can be mapped with the geological object named *Fault*, which is a *GeologicalBoundary*, a specialization of Geological Object in the Basic Geology Ontology. Also, the element *Fault* is characterized by having a *GeologicalStructure*, named *FaultStructure.* Another element of the type *InterfaceFeatureSet* is *horizon*. The element *horizon* can be mapped to *Horizon* with is a specialization of *GeologicalBoundary*. Here we will present our classification of these two concepts (*fault* and *horizon*), according to the properties of identity, rigidity, essentiality and unity. Also, we will present the classification of the concept *geological location*, which is a property of the concepts *Fault* and *Horizon*. However, for a better understanding of our classification, it is important to understand the meaning of these concepts:

- *Fault*: is a break within geological material across which there exists an observable displacement (corresponding to the offset of segments or points that were once continuous or adjacent). As an object, a fault can be approximated as

[12] A *geological block* is a volume of contiguous material points all belonging to the same geological formation and fully limited by a set of geological surfaces or by an external boundary of the model (Perrin and Rainaud, 2013).

a very thin tabular volume possibly made of brittle rock. Within earth models, faults are currently represented as mere surfaces or as volumes of zero thickness. Depending on the relative direction of displacement between the rocks and fault blocks, on either side of the fault, the fault movement is described as normal, reverse or strike-slip. Thus, we can assert that a fault separates different geological units that are not parts of the faults but only neighboring it. However, even when an observed geological discontinuity does not separate different geological units, it may still be a fault, since a fault can locally separate rock entities that belong to one thick unit.

- *Horizon*: according to the definition given in the glossary of the book "Shared Earth Modeling" (Perrin and Rainaud, 2013), *horizon* is "a term used for designating either a unit boundary or a remarkable bed of small thickness. It can also correspond to a sedimentary boundary" A horizon may be a surface where seismic waves are reflected and thus correspond to a seismic horizon. . In this paper, we adopt only the second meaning, and then a horizon has not substance.

In view of this, it is possible (and essential for geomodeling of geological mapping) to consider faults and horizons as objects, because things that can be observed can materialize both concepts. Then, we can say that they offer the principle of identity and unity for their instances. Also, both concepts are rigid, since their instances will remain what they are along the whole existence. A fault or a horizon may change its name or its age but there is practically no chance that it changes its nature. A fault will always remain a fault; a horizon will always remain a horizon.

Finally, the concept *GeoLocation* represents an essential property of an instance of fault and/or horizon. Although it is not an independent object since it inheres to a fault or horizon, this property is essential to all its possible instances. Thus, it is a rigid property. The result of this analysis is presented in the Table 07.

Type	Supplies Identity (O)	Carries Identity (I)	Rigidity (R)	Relational Dependence (D)	Unity (U)
Fault	+	+	+	-	+
Horizon	+	+	+	-	+
GeoLocation	-	-	+	+	+

Table 07: The result of the classification realized with RESQML types.

RESQML is built with several other objects and properties that cannot be mapped to the concepts of Basic Geology Ontology and therefore they were not described here. Also, for a question of space and comprehension, we present only a subset of the Geology concepts required to model a petroleum prospect. A deeper analysis will be detailed in further reports related to our project.

5. Conclusion

We presented in this paper a brief view of a methodology for ontological analysis of earth models, aiming in providing support for information integration and software interoperability. We founded our analysis in the philosophical background of ontological studies about the nature of the existence of being, updated with recent development in Computer Sciences.

Our claim is that the interoperability among earth models built and manipulated by different professionals and systems can be achieved by making apparent the intended meaning of the geological objects represented in the models. We describe the ontological meaning of Geology concepts and relationships previously modeled in the Basic Geology Ontology. The paper does not intend to cover the whole Geology domain in proper extension and detail. Thus, the few concepts that have been analyzed are those which are more frequently found in Earth models and which are of central importance for anchoring petroleum exploration models into raw data and into entities existing in the reality.

We showed that few ontological properties inhered to concepts - identity, rigidity, essentiality and unity – are enough to clarify the meaning of the modeled entities and to define the similarities and identities between models and data. These properties elucidate which are the concepts that are essential to the modeled reality and can be used for providing a unified view over this reality. In addition, the understanding of the modeling principles can lead to clean the future models of the entities that have no proper existence in reality (i.e., that have no instances). This allows modelers to develop naturally integrable models based on a common framework of the essential rigid concepts.

We also discussed several misconceptions about relationship meaning commonly found in Earth models. From these, the more significant are: the mix of the concepts *geological unit* (a delimited object in 3D space and time) and *rock* (the substance of what this object is made), the relation between geological unit and their spatial limits (boundaries, faults) and internal structure (geological structures, sedimentary structures), which are rigid concepts that are inseparable parts of lithological units and not properties of them.

Due to space, we have only analyzed in this paper some main concepts and properties attached to geological objects leaving apart some other important properties (permeability, granulometry, age, etc.) that we have studied. Moreover, significant work remains to be done for studying the ontological properties of all the objects that have to be considered when building earth models. The authors of this paper will go on working to achieve this goal.

6. References

Abel, M. (2001). The study of expertise in Sedimentary Petrography and its significance for knowledge engineering (Doctoral Thesis). *Computer Pos-Graduation Program*, UFRGS, Porto Alegre, 239.

Cox, S. J., & Richard, S. M. (2005). A formal model for the geologic time scale and global stratotype section and point, compatible with geospatial information transfer standards. *Geosphere*, *1*(3), 119-137.

Deny, L., Verney, P., Schey, J., King, M. J., Henri-Bally, R. C., & Rainaud, J. F. (2013, September). Delivering key WITSML Measured and Interpreted Well Information

Through RESQML V2 for Reservoir Characterization and Flow Simulation. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

Endres, D. M., King, M. J., Schey, J., Rainaud, J. F., Deny, L., Verney, P., & Morandini, F. (2013, June). How RESQML Version 2 Can Facilitate the Update of a Reservoir Model. In *EAGE Annual Conference & Exhibition incorporating SPE Europec*. Society of Petroleum Engineers.

Guarino, N. (Ed.). (1998). *Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy* (Vol. 46). IOS press.

Guarino, N., & Welty, C. A. (2009). An overview of OntoClean. In *Handbook on ontologies* (pp. 201-220). Springer Berlin Heidelberg.

Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. CTIT, Centre for Telematics and Information Technology.

Gruber, T. R. (1992). *Ontolingua: A mechanism to support portable ontologies*. Stanford University, Knowledge Systems Laboratory.

Lorenzatti, A., Abel, M., Nunes, B. R., & Scherer, C. M. (2009). Ontology for imagistic domains: Combining textual and pictorial primitives. In *Advances in Conceptual Modeling-Challenging Perspectives* (pp. 169-178). Springer Berlin Heidelberg.

Mastella, L. (2010). *Semantic exploitation of engineering models: application to petroleum reservoir models* (Doctoral dissertation, Thesis, March 2010, Ecole des Mines de Paris (ENSMP), Paris, France, p 247).

Omdal, S. (2006, May). The integrated information platform (iip) for resevoir and subsea production systems. Technical report, Norwegian Oil Industry Association (OLF) 2005, POSC IntOPS SIG Regional Meeting.

Perrin, M., Mastella, L. S., Morel, O., & Lorenzatti, A. (2011). Geological time formalization: an improved formal model for describing time successions and their correlation. *Earth Science Informatics*, *4*(2), 81-96.

Perrin, M., & Rainaud, J. F. (Eds.). (2013). *Shared Earth Modeling: Knowledge Driven Solutions for Building and Managing Subsurface 3D Geological Models*. TECHNIP OPHRYS EDITIONS.

Raskin, R. G., & Pan, M. J. (2005). Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers & Geosciences*, *31*(9), 1119-1125.

Richard S.M. (2006). Geoscience concept models in inha, A. K. ed (2006). *Geoinformatics: data to knowledge*. Geological Society of America, especial Paper n° 397, Boulder, USA, pp. 81-107

Sandsmark, N., & Mehta, S. (2004, October). Integrated information platform for reservoir and subsea production systems. In *Proceedings of the 13th Product Data Technology Europe Symposium (PDT 2004), October, Stockholm*.

Sinha, A. K. ed (2006). *Geoinformatics: data to knowledge*. Geological Society of America, especial Paper n° 397, Boulder, USA, 282 p.

Wache, H., Voegele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., & Hübner, S. (2001, August). Ontology-based integration of information-a survey of existing approaches. In *IJCAI-01 workshop: ontologies and information sharing* (Vol. 2001, pp. 108-117).