Ontology-driven representation of knowledge for geological maps

(Article begins on next page)

12 January 2025

ONTOLOGY-DRIVEN REPRESENTATION OF KNOWLEDGE FOR GEOLOGICAL MAPS

Alizia Mantovani[a,b], Fabrizio Piana[b*1], Vincenzo Lombardo[c,b]

[a] Dipartimento di Scienze della Terra, Università di Torino, Torino, Italy

[b] CNR-IGG, Istituto di Geoscienze e Georisorse, Torino, Italy

[c] Dipartimento di Informatica, Università di Torino, Torino, Italy

*Corresponding Author. E-mail address: *fabrizio.piana@cnr.it*

CNR IGG, Via Valperga Caluso 35 10125 Torino Italy

Highlights

1. Ontological approach is needed to reduce the ambiguity of geological concepts

2. Concepts and terms deriving from geosciences standard vocabularies have been encoded in the Ontology (OntoGeonous).

3. Ontology-Driven geodatabase implementation reduces the ambiguity of geological maps

---

[1] Contribution of Authors:

Alizia Mantovani: First Author, translation of GeoScienceML Scheme in machine-readable format (.owl by Protégé software), editing of text and figures, paper structure and ontology-driven geodatabase design
Vincenzo Lombardo: leader of the ontological analysis on geological knowledge, co-coordinator of the research, paper structure and ontology-driven geodatabase design
Fabrizio Piana: Corresponding Author, co-coordinator of the research, ontological analysis of geological mapping process, paper structure and ontology-driven geodatabase design

ONTOLOGY-DRIVEN REPRESENTATION OF KNOWLEDGE FOR GEOLOGICAL MAPS

Alizia Mantovani[a,b], Vincenzo Lombardo[c,b], Fabrizio Piana[b*]

[a] Dipartimento di Scienze della Terra, Università di Torino, Torino, Italy

[b] Dipartimento di Informatica, Università di Torino, Torino, Italy

[c] CNR-IGG, Istituto di Geoscienze e Georisorse, Torino, Italy

*Corresponding Author. E-mail address: *fabrizio.piana@cnr.it*

ABSTRACT

This paper presents an ontology-driven representation of knowledge for geological maps. The ontological formal language allows for a machine-readable encoding of the Earth scientist's interpretation through semantic categories and properties and is credited to support knowledge sharing and interoperability.

We introduce an ontology-driven method for the interpretation and the encoding of the map data that employs shared vocabularies and resources encoded through ontologies in order to prevent the use of ambiguous terms. The approach relies on a computational ontology of the geological knowledge (OntoGeonous), which formalizes a number of geological knowledge sources (including GeoScienceML), to guide the interpretation process. The design of the database underlying the map (OntoGeoBase) constrains the process of data entry to refer to the terminology conveyed by the taxonomic-axiomatic nature of the ontology.

This reduces the amount of implicit knowledge favouring a conceptual alignment of the ancillary documentation with the map, leading to a better comprehension of map and allowing the traceability of the interpretation.

1. INTRODUCTION

In the last two decades, the major methodological innovation in the production of geological maps has essentially concerned the usage of tablets and small size PC's, equipped with a GIS software, directly on the field (see, e.g., McCaffrey et al., 2005; Pavlis et al., 2010; Whitmeyer et al., 2010, DeDonatis et al., 2016). This technological advance favoured the acquisition of information in a digital format since the beginning of the mapping process in the field, in a suitable way for data reworking and sharing. However, the remarkable advantages brought about by these IT innovations have mainly addressed the digital recording and representation of spatial data. Some important features, such as the interoperability and the unambiguity of data, has not been addressed by these new technologies. This paper focuses mostly on the final step of the digital mapping process, i.e. the representation of geological knowledge in maps and in particular the disambiguation of concepts through the semantic formalization. Semantic formalization of concepts is conceived to address the peculiarities of the domain, allows abstracting objects from the real world to the information world, while the implementation of the resulting database will be driven by the logical constraints provided by a computational ontology.

The semantics-informed design of the database through a computational ontology, although largely debated in the literature (Uschold 2015), fits the problems of geological knowledge representation in maps, since it: (i) focuses on the formal definition of classes (or categories, where the instances are members of classes; (ii) provides strong constraints (axioms) to

convey data meaning, together with categories, for consistency and reasoning; (iii) comes with reasoning algorithms, to infer new information. In fact, the ontological approach is valuable when the encoding of human thinking is a crucial issue with respect to data encoding (Uschold 2015). We claim that the importance of the terminological issue as well as the engagement of the Earth scientists in the design process is paramount in geological mapping, especially in the light of terminological effort promoted by the international committees that goes together with the ontological development (more on this below).

It is known that geological mapping largely consists in a process of inferencing (Brodaric, 2004; Balestro & Piana, 2007; Loudon, 2009; 2011) because, in geology, rocks often document events that are inferred to have occurred, and not directly observed[2]. Consequently, from the observation of rocks and landforms on the field to the production of a geological map, many decisions are taken, and many of them are influenced by pre existing models of the geological evolution of the map area. Moreover, in the geological maps much of the knowledge is implicit (tacit knowledge, such as fundamental principles, intended meanings and assumptions). This working method may cause a loss of reproducibility of the data, because of the difficulty to separate data from interpretations, in contrast with the paradigm of the self-correcting nature of science. Therefore, we believe that the ontological approach, with its formal and explicit representation format, can effectively guide the representation of geological knowledge on maps. An ontology-consistent description of the mapped features requires to make explicit much information (namely, classes, properties, and axioms) and leads to a retractable path of interpretation. The ontology axioms can unambiguously encode the relationships of the geologic features with some Geologic Event, which is the key to the reconstruction of the

---

[2] "Geologists make sense of their observations by integrating them with through reasoning and interpretation" (Loudon, 2009)

geological history of the map area. All such explicit information is expressed in a machine-readable language, which allows for the automatic inference (reasoning), and the consistency checking.

In this paper, we employ ontologies for the formal representation of geologic knowledge and the consequent conceptual design of the database schema to address explicitly the interpretation of the mapped features. As a proof of concept, we also describe how to translate the conceptual schema into a logical database schema through a well-known GIS software. Further, the use of computational ontologies improves the employment of shared vocabularies, which in turn support interoperability and data sharing. Nowadays, effective data sharing, through the reference to a common framework (e.g., GeoScienceML[3]) is still rarely supported by geology data infrastructures. Several initiatives aimed at providing a knowledge infrastructure for geosciences addressed a number of issues in the literature. Certain approaches have provided wide scope analyses, which have sketched the scenarios of the infrastructure, such as business models, the assessment of the needs, the formal specification of the requirements (Buller, 2005; Brodaric and Gahegan, 2006; Raskin, 2006) Reitsma et al., 2009; Loudon, 2011). Other approaches have provided some concrete implementations, from the early database schemata (Laxton and Becken, 1996) to the tools for the collection of field data (Dey and Ghosh, 2008) and the definition of standard vocabularies for the harmonization of terminology (Raymond et al., 2016). Many of these initiatives have promoted the usage of ontologies as the major tool for the maintenance of the knowledge assets within the geoscience community or the survey organizations (Howard et al., 2009) or for addressing the issue of data heterogeneity (Abel et al., 2015) and

_____

[3] http://geosciml.org/ , visited on 12 September 2019

have developed ontologies for limited domains, such as, e.g., field activities (Brodaric, 2004; Hwang, 2012; Boyd, 2016), geochronological periods (Ma et al., 2011), lithological materials (Richard, 2006; Sinha et al., 2006; CGI SimpleLithology[4]), algorithmic interpretation of sedimentary facies for the individuation of geologic processes (Carbonera et al., 2015). Recently, we have been developing a logical framework for joining the efforts of the rigorousness of the ontological approach with the standardization of vocabularies, and we have applied the knowledge base to inform the terminology of a geological mapping process (Piana et al., 2017a,b; Lombardo et al., 2018). The method used here for the description and sharing of the geological knowledge in a map is to leverage on ontology OntoGeonous[5,6] (Lombardo et al., 2018), which axiomatizes the vocabularies, UML schemata and natural language definitions provided by GeoScienceML and other knowledge sources. We have designed a geodatabase, named OntoGeoBase, in which the process of data entry is terminologically constrained through the ontological terms.

In line with other approaches that pursue the alignment of representation (such as, e.g., Cox and Richard, 2015), we address the general representation of geological knowledge by encoding the general statements reported in the international standard documentation (such GeoScienceML, INSPIRE, SWEET) and by encoding the specific statements related to the geological map in a consistent way with respect to such general statements. Since the realisation of a geological map is a synthesys process that usually requires many decisions and choices among different interpretative solutions, we believe that an ontological approach grounded on a robust semantic knowledge could allow for a reduction of

---

[4] http://resource.geosciml.org/vocabulary/cgi/201211/simplelithology.rdf
[5] https://www.di.unito.it/wikigeo/index.php?title=Pagina_principale
[6] Note for the reader: here we report names of classes, items and properties from the ontology. All these names are always in the singular form and they are typed without the space between the words, in Camel case, as typical of computer languages: items and classes start with a capital letter (e.g., GeologicUnit), properties start with a lowercase letter (e.g., hasGeologicUnit).

ambiguities and/or implicit knowledge. We claim that the adherence to the international standards puts our approach in a wider perspective in terms of reusability and interoperability.

The paper is structured as follows: in Section 2 we report about the current state of the art, i.e. how digital geological mapping has become a common practice in the last decades, and how it shaped, in some countries, the operative framework of the national mapping projects. Section 3 presents the background technologies for our approach, namely the existing standard vocabularies and their encoding into the ontology OntoGeonous. Section 4, the core of the paper, reports on how the ontology can account for the interpretation of the geological knowledge to be reported in the map. As a proof of concept, Section 5 describes how we can shape the OntoGeoBase schema from the ontology axioms and properties. Section 6 discusses the novelties and the impact of our approach on the geological mapping process. Finally, Section 7 reports out conclusions.

2. STATE OF ART IN GEOLOGICAL MAPPING AND VOCABULARIES

In this section, we address the state of the art in the current methods of digital geological mapping, its standardization, and the role of shared vocabularies. With the expression "Digital geological mapping", we intend all the processes that lead to creation of a map, from the beginning of the working process (and directly in the field) to the sharing of the geological map and its contained knowledge. In this paper, however, we focus on the representation of the data contained in a geological map compliant with the standard vocabularies. As a proof of concept, in Section 5 we introduce some operative tools for the geological mapping.

Due to a huge improvement in the techniques of data acquisition, geologic data (observations, graphic representations and measurements) have been progressively transferred from paper to electronic devices supporting GIS software. Many geologists have experienced these techniques in different types of applications, such as the fieldwork for the digital geological mapping (Pavlis et al., 2010) and the geo-engineering works with a 3D representation (Pavlis et al., 2017; Thum and De Paoli, 2015; Ambrosi and Scapozza, 2015). The advantages of the digital geological mapping techniques are (i) a faster and more accurate data collecting, (ii) easier management of large amount of data (either before or during the field work), (iii) more efficient sharing of data (De Donatis et al., 2016). However, this condition is effective only in the case of a standard usage of the technologies and language (not so frequent at present, Mookerjee et al., 2015), and if the "collection" of data is strictly tailored to the project task (in geology, data collection is often discriminated and interpreted since the very beginning of their acquisition process).

Several National Geological Surveys have manifested the necessity of a standard approach and of controlled vocabularies for the geological mapping process, usually addressing the database information entry (Richard, 2003; Tudor and Gheucă, 2009; Tiainen et al., 2008; van Gasselt and Nass, 2010; NMBGMR, 2018[7]).

For example, during the 80's, the Italian Geological Survey launched a National Mapping Program (named CARG project[8]), to produce spatially continuous and coherent geological maps for the whole Italian territory at 1:50,000 scale, compliant with the European standards of those times, and provided with a common database. This allowed for a significant conceptual improvement of the geological mapping process. Nevertheless, the CARG geodatabase structure, lacking of controlled vocabularies or conceptual maps for the representation of the relations over the geological entities, displays a number of ambiguities in the representation of the geologic features.

More recent national geological mapping projects are from the UK and Austria. The DiGMapGB[9] by BGS is a synthesis geological map of UK, equipped with a database for the geologic units. However, the terms employed for the database implementation are only compliant with the British vocabularies, with "traditional" and well rooted terms[10] that are compliant with the Digital National Framework (DNF), "a model for the integration of geographic information of all kinds" (Holland, 2001). DNF only applies at a national level (UK), although it largely overlaps with EU INSPIRE in terms of consistency and interoperability (Laxton and Duffy, 2011). The Geological Survey of Austria recently shared

---

[7] NMBGMR Draft Geologic Data Model - v. 1.0.4 Overview; Created by: Adam S. Read, Geoff Rawling, Daniel J.Koning, Gretchen Hoffman, Sean D. Connell, J. Michael Timmons, David McCraw, Glen Jones, Mark Mansell, & Shannon Williams; Revised: 27-Jul-2018; https://geoinfo.nmt.edu/statemap/datamodel/#top
[8] http://www.isprambiente.gov.it/en/cartography/geological-and-geothematic-maps?set_language=en
[9] https://www.bgs.ac.uk/products/digitalmaps/DiGMapGB.html
[10] Principles of the BGS Rock Classification Scheme https://www.bgs.ac.uk/bgsrcs/home.html

online a Geological Map at 1:1.000.000 scale[11] by implementing a WebGIS Service, where the properties of several geological object classes can be queried. The geodatabase leans on a Thesaurus consisting of native vocabularies, partially consistent with some authoritative sources such as NADM - North American Data Model (2003), GeoScienceML and INSPIRE, as well as some reference papers (Neuendorf et al., 2005; Hintersberger et al., 2017).

The idea of a global standard for geological mapping was born in USA, with the goal of producing a unified data model (NADM) for geologic cartography, exploitable to align geological maps of all the US territory[12]. Its intention was also to provide guidelines for the design and filling of a database[13] (Johnson et al, 1999). NADM (2003) has provided a vocabulary with the definitions of the terms as well as UML schemata that represent the relations over the concepts. This work paved the road to the IUGS Commission for the Management and Application of Geoscience Information (CGI), which developed UML schemata for the geological knowledge, representing the relations between the objects (Geological Features) and concepts commonly used in the geological maps.

The development of digital mapping techniques of the last two decades and the consequent increased capacities of data sharing (Pavlis et al., 2010; De Donatis et al., 2016) has pushed forward the requirements for sharing and alignment of geoscientific information. In this context, the initiative carried out by the CGI Commission, named "GeoScienceML"[14] gains a notable relevance.

GeoScienceML is a data model standard for the general organization of the geological knowledge (Sen and Duffy, 2005), consisting of a number of vocabularies (provided by CGI, Simons et al., 2006) throughout many different categories. The vocabularies represent the

---

[11] http://geolba.maps.arcgis.com/apps/webappviewer/index.html?id=0e19d373a13d4eb19da3544ce15f35ec
[12] https://ncgmp.usgs.gov/
[13] https://ngmdb.usgs.gov/ngmdb/ngmdb_home.html
[14] https://www.seegrid.csiro.au/wiki/CGIModel/GeoSciML

geological concepts as classes, with their hierarchy, through broader-narrower relationship or through UML diagrams.

These standards have been reformulated by the INSPIRE directive[15], which aims to create a European Union spatial data infrastructure. INSPIRE embeds GeoScienceML as its data model standard for the exchange of geological information over the countries of the European Union (Data Specification on Geology, version 3.0[16]) by adopting some of the fifteen packages of GeoScienceML (e.g., GeologicUnit, GeologicStructure, Geomorphology, GeologicAge, Borehole, EarthMaterial..., see GeoScienceML Cookbook for INSPIRE[17]). Controlled vocabularies provide a standard terminology to be shared between countries (and languages).

Notwithstanding the initiatives above, the mapped features of geological maps (geologic units, geologic structures and geomorphological features) are, in the current geomapping practices, mostly described with a high level of detail, referring to complex hierarchy relations, but rarely in a way that is compliant with standard vocabularies. To describe the "4D" (3D + Time) geological complexity, geoscientists often use non-standardized terms (cf., e.g., Tectonostratigraphic Unit), thus leading to some level of ambiguity.

3. ONTOGEONOUS AND THE ONTOLOGIES FOR THE GEOSCIENCES

Computational ontologies has proved to be effective in sense disambiguation (see, e.g., Navigli and Velardi, 2005). Ontology OntoGeonous concerns the definition of the geological features, which are encoded into formally axiomatized classes, including properties

---

[15] https://inspire.ec.europa.eu/
[16] D2.8.II.4 INSPIRE Data Specification on Geology - Technical Guidelines v. 3.0. (10.12.2013); (http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_GE_v3.0.pdf).
[17] http://geosciml.org/doc/geosciml/3.2/documentation/cookbook/INSPIRE_GeoSciML_Cookbook%20_1.0.pdf

between concepts (see also Lombardo et al., 2018). Here we review ontology OntoGeonous and the major properties employed in the representation of geological knowledge in maps. Ontologies usually provide the semantic backbone for the knowledge graph of entities and concepts (Reitsma et al, 2009). There are a number of ontologies in the field of geosciences:

- a top-level ontology of the entities in the universe (SWEET ontology, Semantic Web for Earth and Environmental Terminology, Raskin and Pan, 2005; Raskin, 2006[18]) was created by NASA to organize a huge amount of data concerning the planetary realm, and can provide the upper level knowledge that contextualizes the geologic features;

- some ontologies have been developed in some specific knowledge domains of geological sciences, e.g., ontology of fracture (Zhong et al., 2009), ontology for geologic time scale (Ma, 2011), Structural Geo-Ontology (Babaie et al.,2006), Simple Lithology[19] and others; differently from SWEET, the latter provide a detailed knowledge encoding for some limited domain of interest.

OntoGeonous, the ontology we developed in the context of the geological mapping process, is positioned at a "middle" layer, linking the top-level SWEET ontology with the specific domain ontologies, through the encoding of the GeoScienceML schemata and vocabularies[20]

.

---

[18] sweetontology.net

[19] http://resource.geosciml.org/vocabulary/cgi/201211/simplelithology.rdf

[20] As far as we know, this is the first case of middle layer, defined as comprehensive of geoscience concepts that are not related to a specific application and connected to both top and domain layers. By "middle layer ontologies" we intend ontologies that are limited to a discipline, such as geosciences, kept distinct from top-layer concepts (such as, e.g., process, chemical substance, or geometrical object), domain ontologies (e.g., petrographic ontology) and application ontologies (e.g., GEON ontology), respectively.
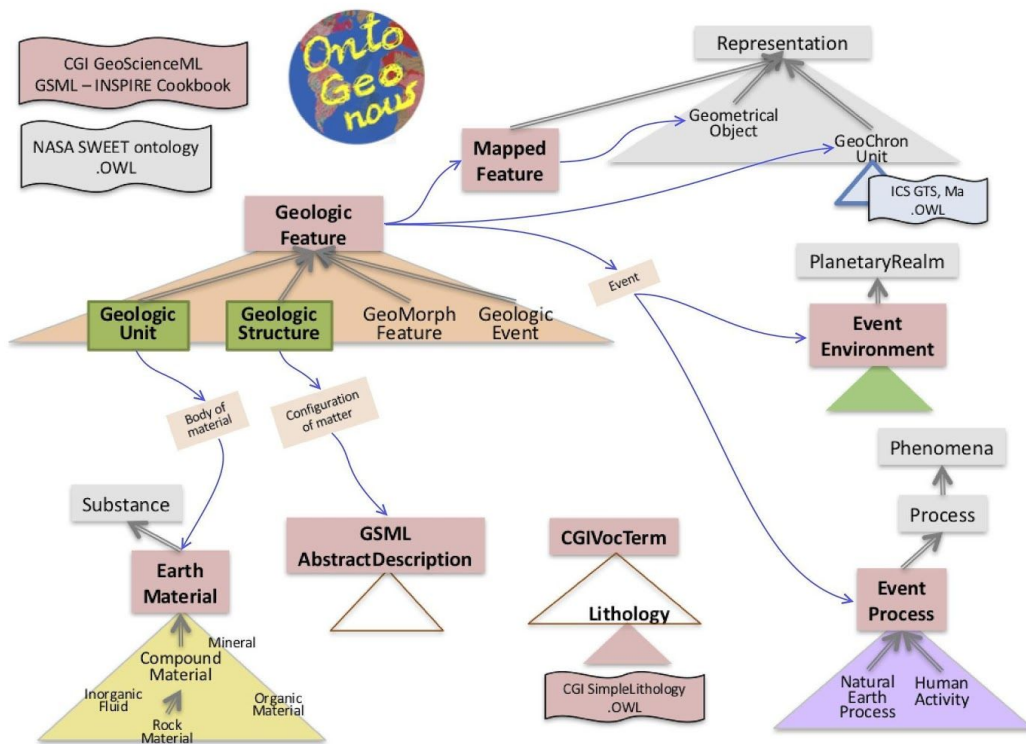
Fig.1 Overview of OntoGeonous (modified after Lombardo et al., 2017). The triangles (labelled with rectangles) represent the main classes included in the ontology; grey bold arrows link the various subclasses to the main classes; the relations (properties) between classes are represented by blue thin arrows.

OntoGeonous (Lombardo et al., 2018) is a merged set of computational ontologies that has been realized through the OWL encoding of the definitions reported in authoritative resources, here listed (see Fig.1 for an overview and the dedicated Wiki pages[21] for the axiom encoding process):

- GeoScienceML (Version 4.1, 2015[22]) and INSPIRE (Infrastructure for Spatial Information in Europe)[23] condensed into the GeoScienceML + INSPIRE cookbook (addressing the major vocabularies);

- top-level ontology NASA SWEET[24] for environmental and Earth system science terms;

- various vocabularies of specific subdomains of geologic knowledge[25], e.g., the lithology domain vocabulary "SimpleLithology"[26];

- the ICS Geological Time Scale Ontology (Ma, 2011) as a subtaxonomy of the Geochronologic Unit class of SWEET Representation (actually the hierarchical path Representation - NumericalEntity - Interval - Duration - GeochronologicUnit).

As devised by the UML schemata in GeoScienceML, the core of the geologic knowledge in OntoGeonous is a taxonomy rooted by the class GeologicFeature, which encompasses the geologic core knowledge, related to 1) MappedFeature, i.e. the spatial extent of the geologic feature on the map, 2) GeoChronologicUnit, rooted by the ICS GTS taxonomy, 3) CGIVocabularyTerm (an OntoGeonous taxonomy for CGI vocabularies), which provide specific concepts for several subdomains, such as those of Earth materials, and other abstract descriptions in GeoScienceML.

GeologicFeature is subdivided into four sub-taxonomies, namely GeologicUnit (the bodies of some material), GeologicStructure (configurations or patterns in which the geologic units are arranged), GeomorphologicFeature (the landforms), GeologicEvent (relevant events in geologic history). These classes have in turn subclasses. The relations between the classes are represented by the properties, distinguished in two types, namely ObjectProperties (OP, connecting two instances) and DataProperties (DP, when the range is not a conceptual class, but a scalar value of some type, such as, a boolean, a string, a number). Tab.1 shows some examples of properties (namely, the ones necessary for the classification of an instance in classes Shear Displacement Structure and Contact).

---

[24] https://sweet.jpl.nasa.gov/
[25] http://resource.geosciml.org/vocabulary/cgi/201211/
[26] http://resource.geosciml.org/vocabulary/cgi/201211/simplelithology.rdf

N.	Name	Type	Description	Accepted Values	Defining Class	DB Label
1	has Geologic Unit	OP	It describes the relation of the structure with a unit, e.g., "Contact bounds GeologicUnit"; comprises both its sub-properties "bounds" and "is within"	items from Geologic Unit class	Contact SDS	GEOL_ UNITn
2	has Geometrical Object	OP	It describes the shape of the object	items from GeometricalObject class (2DStructure,..)	Contact SDS	GEOM _OBJ
3	has Movement	DP	It indicates if movement between two rock parts occurred parallel to the structure	Boolean (True or False)	Contact SDS	MOVEM
4	has Deformation Style	OP	It describes the mechanical effects of tectonically induced shape change of a rock body	items from CGIVocabulary Deforma-tionStyle (Brittle, Ductile or Mixed)	SDS	DEF_ STYLE
5	has Hiatus In Deposition	DP	It indicates if is present a hiatus in deposition	Boolean (True or False)	Contact	HIATUS
6	has Movement Type	DP	It refers to values indicating the kind of movement	Extraction, Normal, Oblique-Slip, Reverse, Scissor, Strike-Slip	SDS	MOVEM _Ty
7	has Dip	DP	It is the value of the planar orientation of the object	integer (0-90)	Contact SDS	DIP

Tab.1 *List of the properties needed for the classification of the instances of GeologicStructure class and subclasses. Each property has a type (Object property: OP; Data property: DP), a textual description of its meaning, the accepted values and the Geologic Structure type which the property refers to (defining classes) and the label used in the columns of OntoGeoBase. In the column "Defining Class", SDS is the abbreviation for "ShearDisplacementStructure"*

Every term and concept of the ontologies and UML schemata becomes a class in one of the OntoGeonous taxonomies and it is provided with a definition expressed through an axiom written in a machine-readable language. The axiom, reporting the necessary and sufficient conditions for an item to be included in a given class, is composed by a number of properties organized with the operators AND/OR. The classes instantiate the individuals to describe specific cases, the individuals must satisfy the axioms of the class definition in order to be classified properly through an automated reasoning process. Fig.2 shows the relations (expressed by the properties - arrows) between the individuals (instances) of the ontology. Some are real objects from a geological map ("Canova Fault" and "Areniti di Tonengo"; Geological Map of Piemonte, Piana et al., 2017a), others are terms from standard

vocabularies (e.g., Lithology). Through the ontology, the relations between the items and the terms that describe them are explicit, formalized and unambiguous.
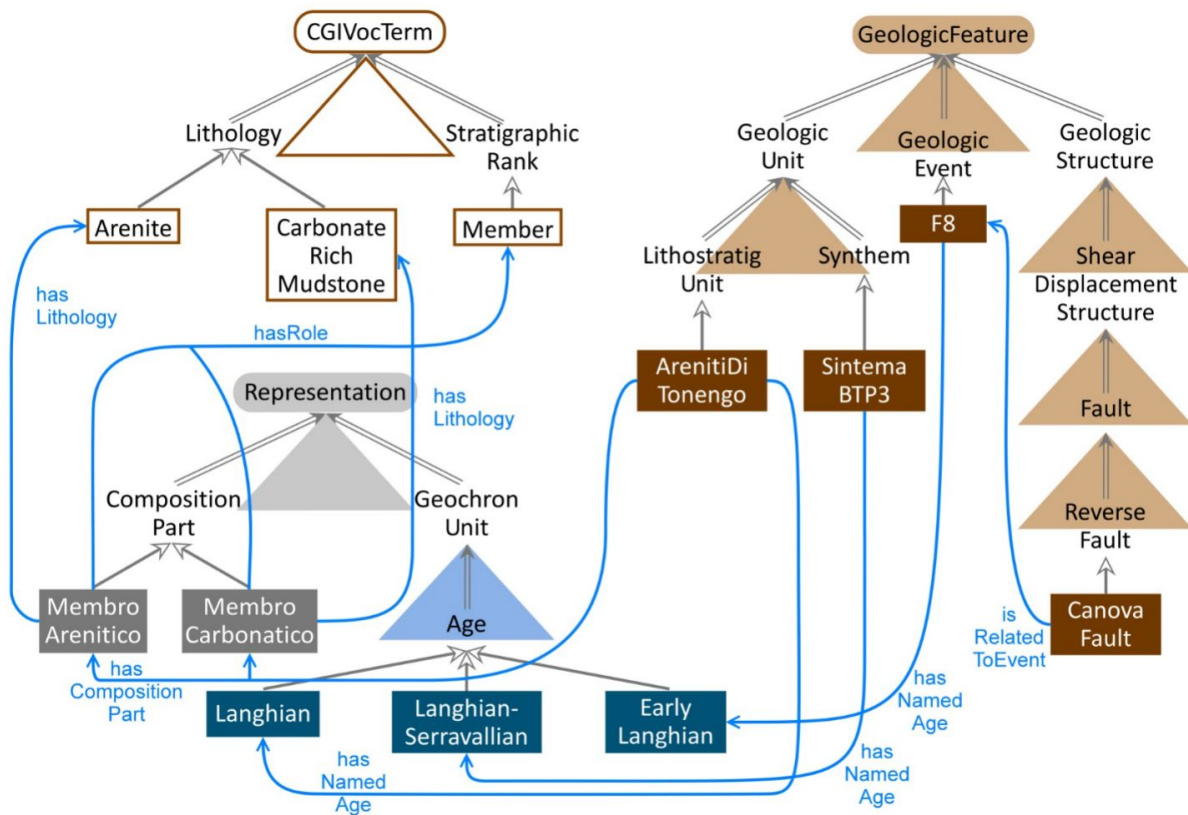

Fig.2 Relations between some individual instances of OntoGeonous (from Geological Map of Piemonte, Piana et al., 2017a). The instance CanovaFault (GeologicStructure class) and Areniti di Tonengo (GeologicUnit class) belong to hierarchical classes (represented by triangles) and have relations (blue arrows) with other instances (represented by sharp rectangles). The colour of the graphics indicates which ontology a given class belongs to: CGI in white, GeologicFeature (GeoScienceML) in brown, Representation (SWEET) in grey and its subclass GeochronologicUnit in light blue.

The use of the formal language has the advantage of making explicit concepts (in the form of classes) and relations over classes and instances; moreover, the machine-readable encoding could allow for setting up automated services that account for a number of tasks.

In addition, building an ontology based on international standards (CGI vocabularies and GeoScienceML schema) make the ontological resource interoperable with other domains. For example, the application of the SKOS framework[27] allows for conceptual terms to be associated with lexical labels; In particular, there are tags for identifying the language to which some label belongs. For example, "en" stands for English, "ja" for Japanese, and "it" for Italian.

At the current stage of development, despite the existence of standard controlled vocabularies licensed by international committees and ontologies to connect them to, a general method that applies the encoded knowledge is still lacking. This paper proposes an approach for a knowledge-driven geological mapping, with an implemented prototype as a proof of concept, where the OntoGeonous knowledge base drives the design of a geodatabase (OntoGeoBase).

4. ONTOGEONOUS FOR THE GEOLOGICAL MAP REPRESENTATION

The geological map is a synthesis and interpretation of a wide range of data sources (Harrison, 1963), implemented in the frame of interpretative and historically based geological reasoning (Frodeman, 1995). Different vocabularies and concepts are to be considered, depending on which kind of geological features we want to highlight in the map. One of the relevant difficulties in understanding geological maps relies on the fact that they are actually four-dimensional data systems, where the fourth dimension of time is crucial to understand the relations between the represented geological features.

A basic geological map is usually intended as a lithostratigraphic map, representing two groups of items: geologic units and geologic structures, closely dependent on each other. A

---

geological map also represents geomorphological features (landforms) and many other specific punctual features (such as springs, fossil-rich sites, …). All these instances were originated by some geologic event in the Earth's history, which should be encoded in a chronological list of all the geologic events occurred in a given geologic time span. These four essential group of items are the four major classes of OntoGeonous, borrowed from GeoScienceML: GeologicStructure, GeologicUnit and GeomorphologicFeature correspond to actual features in the map, GeologicEvent is a conceptual feature not directly represented in the map, but elicited by such features and reported as an attribute of the corresponding specific instances.

To avoid ambiguities while representing these GeologicFeature, we formalized them in OntoGeonous by encoding their standard definitions, i.e., creating their axioms. As an example, the definition of LithotectonicUnit ("Geologic unit defined on the basis of structural or deformation features, mutual relations, origin or historical evolution") is translated as follows:

ONTOGEONOUS AXIOM:

```
LithotectonicUnit EQUIVALENT TO
GeologicUnit
AND
  (hasInternalGeologicStructure some Foliation)
AND
  (isBoundedBy min 2 ShearDisplacementStructure)
AND
(isRelatedToEvent some GeologicEvent)
```


The axiom states that an item must satisfy four conditions at the same time to be classified as an instance of the class LithotectonicUnit (because of the "AND" operator between the conditions); they are:

1. *"GeologicUnit"*: the item must be a generic geologic unit, so it must satisfy all the conditions required for that class;

"AND"

2. *(hasInternalGeologicStructure some Foliation)*: a LithotectonicUnit must have a typical internal geometrical arrangement due to a tectonic deformation, expressed by a Foliation.

"AND"

3. *(isBoundedBy min 2 ShearDisplacementStructure)*: two properties must hold, respectively, between the geologic unit and two different instances of ShearDisplacementStructure class.

"AND"

4. *(isRelatedToEvent some GeologicEvent)*: this specifies that a lithotectonic unit must be related to an instance of the GeologicEvent class.

The upper part of Fig.5 reports a graphic expression of the ontological representation for the instance of a geologic structure named "CanovaFault". To be classified as a Reverse fault, it must satisfy all the conditions required by its superclasses, which are in turn Fault, ShearDisplacementStructure, and GeologicStructure. For the axioms related to the latter three classes, check the upper part of Fig.7: you can notice that the instances and properties required for the definition of CanovaFault are reported in Fig.5.

The axiom explicitly expresses the relations of some item and its defining features. For example, a LithotectonicUnit is related to its internal and bounding GeologicStructure and to its GeologicEvent; CanovaFault is related to some GeologicEvent and a number of other entities. In the context of the geological map interpretation, the most important relation to

be provided with an explicit reference is the one to a GeologicEvent, due to its relevant role in the traceability for the interpretation of the geologic history of the map area. As acknowledged in the literature, these formal categories and properties provide a classification for the data of a geologic map (Brodaric et al., 2004) and are conveniently stored in a relational database (Richard, 2003) that is usually integrated and accessed through GIS software. Now we show how the data of the geological map can be encoded through the ontology-driven representation of the geological knowledge, in order to support the alignment of different types of information for future interoperability.

5. PROOF OF CONCEPT: FROM ONTOGEONOUS TO ONTOGEOBASE

In this section, we apply the ontology classes, axioms, and properties to the design of a database logical schema, named OntoGeoBase, and we present its implementation into a well-known GIS application.

OntoGeoBase is a relational database, i.e. a number of tables (the so-called relations) consisting of rows and columns. Each row represents an entity (or instance, in ontological terms) and the columns represent the attributes of such entity; the whole table represents an entity type (or class). The methodology for deploying the OntoGeoBase schema from OntoGeonous is described in the following pages.

Attribute Tables: Most of the ontological classes that are relevant for the geological maps (GeologicUnit, GeologicStructure[28]) correspond to one or more shapefile in the GIS maps, each one having a dedicated attribute table format (Fig.3). The table columns are tailored for each specific class and/or subclass, depending on which are the necessary

---

[28] GeolomorphologicFeature class is not concerned in this paper, since it is not yet implemented in OntoGeonous and OntoGeoBase

definitory conditions of that class. The columns are grouped into three main categories: "Taxonomy columns", "Definitory Property columns" and "Other columns" (Fig.4).
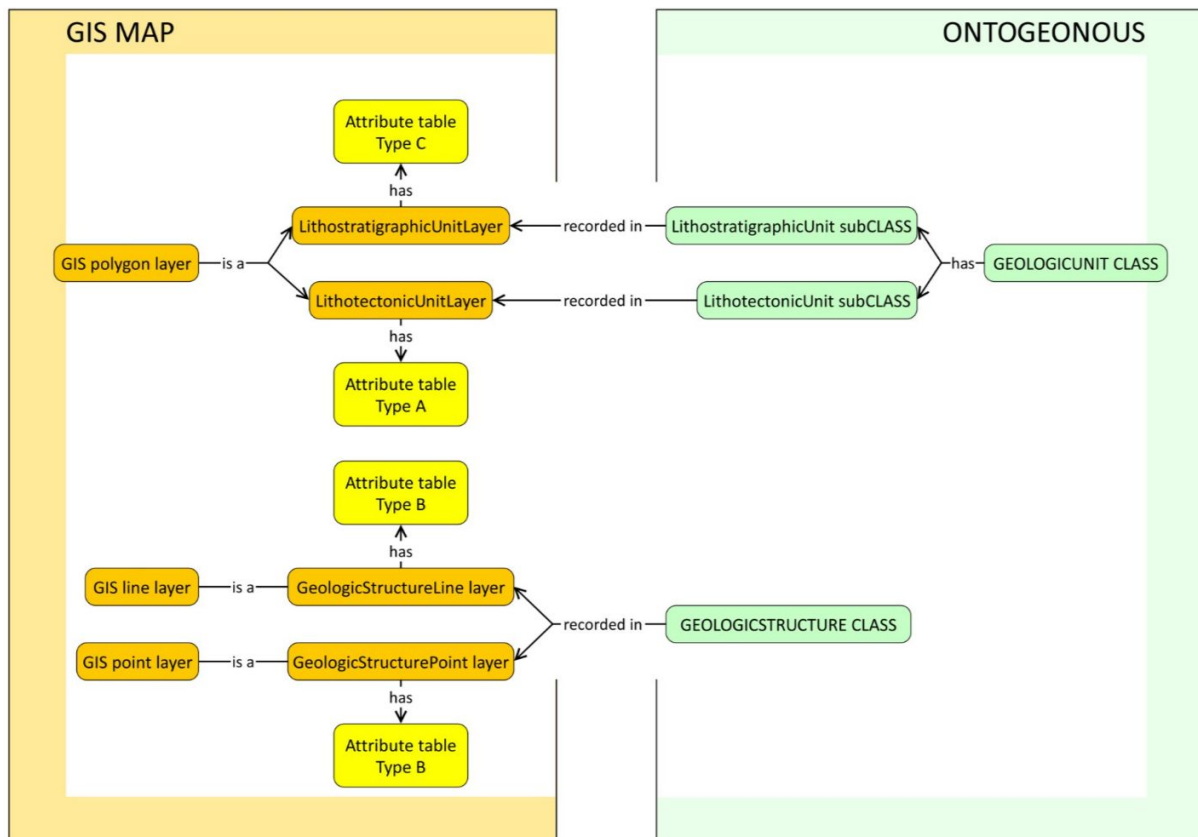
Fig.3 Relations between the conceptual classes of the ontology and the correspondent physical layers on the GIS map.

Taxonomy columns: in OntoGeonous, the main classes have a number of subclasses, e.g. the class GeologicStructure subsumes the classes Contact, ShearDisplacementStructure, Foliation, Lineation, Fold and Fracture[29]. All the instances belonging to different subclasses of a class will be inserted into the same table, in which a number of columns account for the relevant hierarchy levels (chain of subclasses, Fig.2 and Fig.5). This is why these are called taxonomy columns (Fig.4b): the goal is to yield a classification of the items at the most detailed hierarchy level. The property from OntoGeonous that relates these columns is the general property "isSubclassOf" (GS_SUBCL2 isSubclassOf GS_SUBCL1 in Fig.4b).

---

[29] https://www.di.unito.it/wikigeo/index.php?title=GeologicStructure#Subdivisions_of_GeologicStructure

Definitory Property columns (Required Values): definitory properties are those employed by the ontology to define axiomatized classes through the necessary and sufficient conditions that must hold for an instance to belong to that class. Each table of OntoGeoBase has thus a dedicated columns for each property required by OntoGeonous for the classification of the instances. As an example, here follows the transposition of the LithotectonicUnit axiom (check it in the Section 4) as the "Definitory Property columns" section. Since the axiom refers to the conditions for the classification of an instance in the LithotectonicUnit class, we translate those conditions as columns of the Lithotectonic table:

1. column INTERNAL_GS *(hasInternalGeologicStructure),* with values restricted to Foliation;

2. at least 2 columns labelled BOUNDED_BY *(isBoundedBy)* to be filled in the table and only with two different instances of ShearDisplacementStructure class, respectively;

3. column named EVENT *(isRelatedToEvent)* filled with a Geologic Event instance.

Other columns. To satisfy specific needs, it is possible to add some columns to the OntoGeoBase tables. For instance, to represent n-order of geologic units to which a smaller GeologicUnit belongs, it is possible to add n-columns, named "GEOL_UNITn" ($GEOL\_UNIT_{(n)}$ isPartOf $GEOL\_UNIT_{(n+1)}$). A column "Age" is also included for the chronological characterization of the instances and of the GeologicEvent that originated them. Other property classes of OntoGeoBase tables derive from the INSPIRE Directive (Data Specification on Geology):

- Description purpose, to specify the intended purpose/level of abstraction for a given instance. The possible values are instance, typicalNorm, and definingNorm;

- Observation method, to specify the approach used to acquire the attribute values for an instance (e.g. air photo interpretation, field observation, ...);

- Source of the data, to record the source (bibliographic reference) of data (this is not a specific property of OntoGeonous).

Finally, notice that GeologicEvent, one of the four topmost ontological classes, does not have a dedicated GIS layer because it never corresponds, directly, to any MappedFeature or GeometricalObject. However, GeologicEvent is a conceptual category that can have instances (belonging to an encoded list) that have induced remarkable geologic or geo-environmental changes in Earth's history. These instances must have some definitory properties (age, process, environment), which are reported in a dedicated table; some of them are also reported in some columns of the attribute tables of other ontological classes (i.e. in the GIS layers of the GeologicStructure and GeologicUnit class), if a given instance of those classes has been related, by inference or interpretation, to one of the encoded GeologicEvent. The OntoGeoBase schema implements attributes that descends not only from the ontological representation of OntoGeonous, but also from the top-level and domain ontologies merged in OntoGeonous. Fig.4 reports three examples of tables, graphically split into the taxonomic, definitory and other columns, respectively; a colour indicates whether the columns descend from OntoGeonous properties or from some other ontology.

(a) LITHOTECTONIC UNIT TABLE

	NAME
NAME	Murisengo Unit

	BOUNDED BY1	BOUNDED BY2	INTERNAL_GS	EVENT
DEFNITORY COLUMNS	Murisengo Fault	SanCandido Fault	Foliation	F11

	GEOL_UNIT1	GEOL_UNIT2	GEOL_UNIT3	AGE	SOURCE	OBS_METHOD	PURPOSE
OTHER COLUMNS	BTP Domain	Monferrato	BTP1 Synthem	Late Messinian	Dela Pierre et al., 2003	Surface data interpretation	null

(b) GEOLOGIC STRUCTURE TABLE

	NAME	GEOSTR_Ty	GS_SUBCL1	GS_SUBCL2
NAME + TAXONOMY COLUMNS	Canova Fault	SDS	Fault	Reverse Fault

	BOUND UNIT1	BOUND UNIT2	GEOM_OBJ	MOVEM	DEF_STYLE	HIATUS	MOVEM_Ty	DIP	EVENT
DEFNITORY COLUMNS	null	null	Plane	T	Brittle	null	Reverse	null	F8

	AGE	SOURCE	OBS_METHOD	PURPOSE
OTHER COLUMNS	Early Langhian	Dela Pierre et al., 2003	Surface data interpretation	null

(c) LITHOSTRATIGRAPHIC UNIT TABLE

	NAME
NAME	MarneDi Monte Piano

	LITHOLOGY	STRAT_ROLE
DEFNITORY COLUMNS	CarbonateRich Mudstone; Arenite	OnlyPart

	GEOL_UNIT1	GEOL_UNIT2	GEOL_UNIT3	AGE	METAMOR	SOURCE	OBS_METHOD	PURPOSE
OTHER COLUMNS	BTP	BorberaGrue, TorinoHill, Monferrato	BTP0 Synthem	Priabonian	non-Meta_Unit	Dela Pierre et al., 2003	Surface data interpretation	null

	DOMAIN	PALEOGEO	DESCRIPTON
NOT ENCODED CONCEPTS	Synthem BTP0	Syn Orogenic Basins	Marl limeston with intercalations of sand...nes ar...lcarenite

Fig.4 Attribute tables of OntoGeoBase classes. The columns of OntoGeoBase attribute tables for the LithotectonicUnit (4a), GeologicStructure (4b) and LithostratigraphicUnit (4c) are grouped into three categories: Name + Taxonomy Columns, Definitory Columns and Other Columns. Some columns are not derived from a property of OntoGeonous (blue highlighted), but are included anyway in OntoGeoBase as they are relevant for the geologic feature description. The acronym of some columns are reported as "DB Label" in Tab.1 or listed in Note[30]

Fig.5 (upper part) reports a graphic expression of the ontological knowledge for the instance CanovaFault: the GeoScienceML-derived classes are in brown rectangles; classes from the CGI vocabularies are in white; classes from the SWEET ontology are in grey; classes from the ICS ontology are in blue; Data Properties and/or Object properties are in green and blue arrows, respectively. The hierarchical representation of the instance CanovaFault, which belongs to the classes ReverseFault/Fault/ShearDisplacementStructure/GeologicStructure of GeoScienceML, leads to a direct entry in a row of OntoGeoBase (Fig.5, lower part). The item has attributes (reported in the columns of OntoGeoBase) corresponding to the ontology properties (as indicated by the blue, green and orange lines), i.e. every column of the Taxonomy columns and Definitory Property columns sections of OntoGeoBase is required by a formalized concept of OntoGeonous.

---

[30] Acronyms of the OntoGeoBase columns not included in Tab.1: INTERNAL_GS: internal GeologicStructure; GEOSTR_Ty: Geologic Structure Type; GS_SUBCL1-2: GeologicStructure_Subclass1-2; SOURCE: Source; OBS_METHOD: ObservationMethod sensu INSPIRE Data Specification on Geology; PURPOSE: Purpose sensu INSPIRE Data Specification on Geology; DOMAIN: Tectono-stratigraphic Domain; PALEOGEO: Paleogeographic environment or domain ; DESCRIPTION: Description
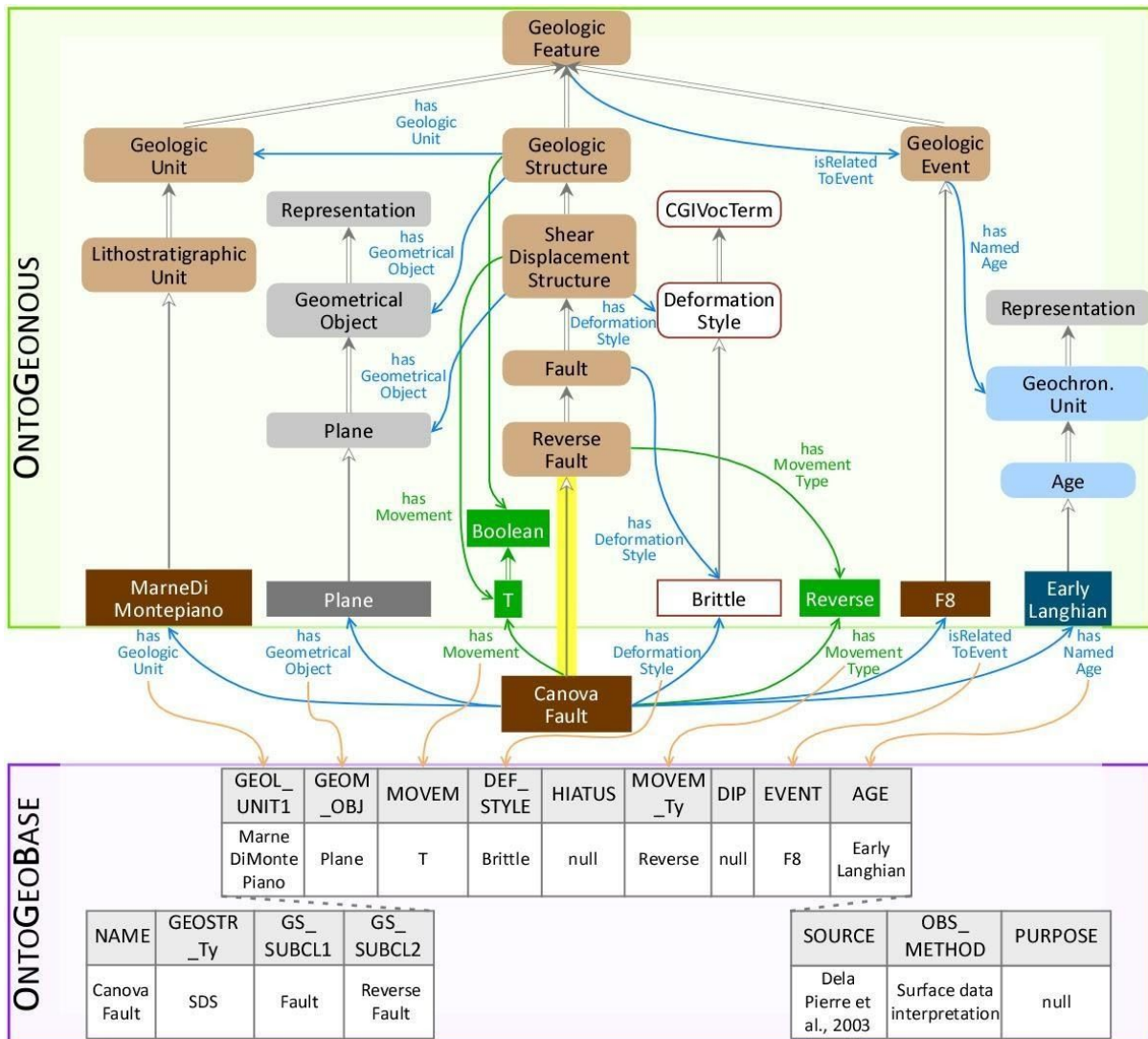
 Fig.5 Ontology-driven description of an item (Canova Fault) of the ReverseFault class and its representation in the database. In the upper part, the classes (light brown rounded squares) are linked through the object properties (represented by blue arrows) to other classes or to instances (sharp squares); the classes are linked by the data properties (green arrows) to encoded values (sharp green squares). The colour of the items represent the ontologies which they are part of: CGI vocabularies in white, GeoScienceML GeologicFeature in brown, Representation from SWEET ontology in grey, together with its subclass GeochronologicUnit in light blue. Every property that link the classes to the values, has a dedicated column of OntoGeoBase (lower part), linked by orange arrows to its defining property. For the acronyms of the columns, see Tab.1 and Note[30].

26

The consistency of a geodatabase is enforced through the design of an appropriate interface that constrains the form filling process: in this perspective, some data entry forms have been developed, in GIS software, for the geomapping relevant ontological classes (GeologicStructure, LithostratigraphicUnit and LithotectonicUnit), so that when entering a new item in the map, a window pops up. The window shows a Data Entry Form (Fig.6) that constrains the classification of a given item as an instance of a certain encoded class (and/or subclass) by the choice of proper Required Values (encoded in OntoGeonous) shown in the drop-down menu of each field. The Data Entry Form puts the classification process at work, incrementally, since the very first steps of the geological mapping process.

Fig.6 shows the structure of the interface for the GeologicStructure table, developed in Open Source GIS software (GIS 2.8.7-Wien). In Tab.1, all the properties for the description of the GeologicStructure instances from the Geological Map of Piemonte (Piana et al., 2017a), which is our case study to test the database implementation process, are listed. The fields of the Data Entry Form are grouped into five data sheets: NAME, GEOSTR_TY, GS_SUBCL1-2, OTHER_C. Among those, the GEOSTR_TY, GS_SUBCL1 and GS_SUBCL2 are dedicated to the classification of the items into their proper class and they are named as the corresponding OntoGeoBase taxonomy columns. These sheets are organized in two insets, one for the choice of the correct class (Taxonomy column, the same reported in the label of the sheet), and the other for the compilation of the related Required Values (Definitory property columns).

Each field corresponds to a column of OntoGeoBase and thus to a property or a class of OntoGeonous.

Fig.6 Data Entry Form for the GeologicStructure class (Geological Map of Piemonte case study). Two of the data sheets are represented: in GEOSTR_TY (left) are listed all the properties for the classification of the direct subclasses of GeologicStructure (Contact and ShearDisplacementStructure in the case study); GS_SUBCL1 (right) contains the properties required for the classification of the possible subclasses of the previous (Fault, DuctileShearStructure, UnconformableContact and ConformableContact in the case study). For the acronyms of the fields and labels, see Tab.1 and Note[30]

5.1. An example of item classification in OntoGeoBase (GeologicStructure class, Geological Map of Piemonte case study)

Now we apply the OntoGeoBase structure to an example from the Geological Map of Piemonte (Piana et al., 2017a); in particular, we show how some instances are classified as a

subclass of GeologicStructure (Contact or ShearDisplacementStructure) giving the properties required by the axioms (see Fig.7). Fig.7 represents three extracts from OntoGeoBase during the process of compilation, which is driven by the Data Entry Form (Fig.6). The two windows of Fig.6 are respectively the Data Entry Form for the step represented in Fig.7a (Fig.6 left) and Fig.7b (Fig.6 right).

The column labeled Geostr_Ty (Fig.7a) specifies whether the item is a ShearDisplacementStructure or a Contact (both direct subclasses of GeologicStructure); hence in the Data Entry Form (Fig.6, left), the corresponding sheet provides all the properties and values required by the OntoGeonous axioms (red squares in Fig.7a, left) for this classification.

GS_SUBCL1 (Fig.7b) column can include all of the subclasses of Contact (ConformableContact and UnconformableContact) and ShearDisplacementStructure (Fault and DuctileShearStructure); Fig.6 (right) shows the corresponding data sheet). Finally, GS_SUBCL2 (Fig.7c) could contain all the subclasses of the classes from GS_SUBCL1, with the required properties and values that be effectively found by the operator using the dedicated Data Entry Form, as previously explained for the Geostr_Ty and GS_SUBCL1.

Supported by the Data Entry Form shown in Fig.6, an operator could fill in the OntoGeoBase tables, even ignoring the structure of OntoGeonous (which operates at a lower level), but exploiting all its constraints.

a

GEOLOGIC STRUCTURE
GeologicFeature
1 and (hasGeologicUnit some GeologicUnit)
2 and (hasGeometricalObject exactly 1 GeometricalObject)
3 and (hasMovement some boolean)
SHEAR DISPLACEMENT STRUCTURE
GeologicStructure
4 and (hasDeformationStyle some DeformationStyle)
2 and (hasGeometricalObject exactly 1 Plane)
3 and (hasMovement value true)
CONTACT
GeologicStructure
1 and (bounds min 2 GeologicUnit)
2 and (hasGeometricalObject exactly 1 Plane)
3 and (hasMovement value false)

GEOLOGIC STRUCTURE TABLE

NAME	GEOSTR_Ty	GS_SUBCL1	GS_SUBCL2	EVENT	AGE	DEF_STYLE
Canova Fault	Shear Displacement Structure			F8	Early Langhian	Brittle
D3	Contact			F8	Late Burdig.-Early Langhian	null

ONTOGEONOUS TAXONOMY

GEOLOGIC STRUCTURE (Geological Map of Piemonte)
- CONTACT
- SDS
  - DDS
  - FAULT
    - HIGH ANGLE
    - NORMAL
    - REVERSE
    - ...

REQUIRED VALUES

1) SintemaBTP2, SintemaBTP3	1) MarneDiMontePiano
2) Plane	2) Plane
3) F	3) T
4) null	4) Brittle

b

FAULT
ShearDisplacementStructure
and
4 ((hasDeformationStyle value brittle) or
  (hasDeformationStyle value mixed_brittle_ductile))
DUCTILE SHEAR STRUCTURE
ShearDisplacementStructure
4 and (hasDeformationStyle value ductile)

GEOLOGIC STRUCTURE TABLE

NAME	GEOSTR_Ty	GS_SUBCL1	GS_SUBCL2	EVENT	AGE	DEF_STYLE
Canova Fault	Shear Displacement Structure	Fault		F8	Early Langhian	Brittle
DSS1	Shear Displacement Structure	Ductile Shear Structure		F8	Late Burdig.-Early Langhian	null

ONTOGEONOUS TAXONOMY

GEOLOGIC STRUCTURE (Geological Map of Piemonte)
- CONTACT
- SDS
  - DDS
  - FAULT
    - HIGH ANGLE
    - NORMAL
    - REVERSE
    - ...

REQUIRED VALUES

4) Ductile	4) Brittle

c

REVERSE FAULT
Fault
6 and (hasMovementType value "Reverse")
HIGH ANGLE FAULT
Fault
7 and (hasDip exactly 1 integer[>=45])

GEOLOGIC STRUCTURE TABLE

NAME	GEOSTR_Ty	GS_SUBCL1	GS_SUBCL2	EVENT	AGE	DEF_STYLE
Canova Fault	Shear Displacement Structure	Fault	Reverse Fault	F8	Early Langhian	Brittle
HAF2	Shear Displacement Structure	Fault	High Angle Fault	F8	Late Burdig.-Early Langhian	null

ONTOGEONOUS TAXONOMY

GEOLOGIC STRUCTURE (Geological Map of Piemonte)
- CONTACT
- SDS
  - DDS
  - FAULT
    - HIGH ANGLE
    - NORMAL
    - REVERSE
    - ...

REQUIRED VALUES

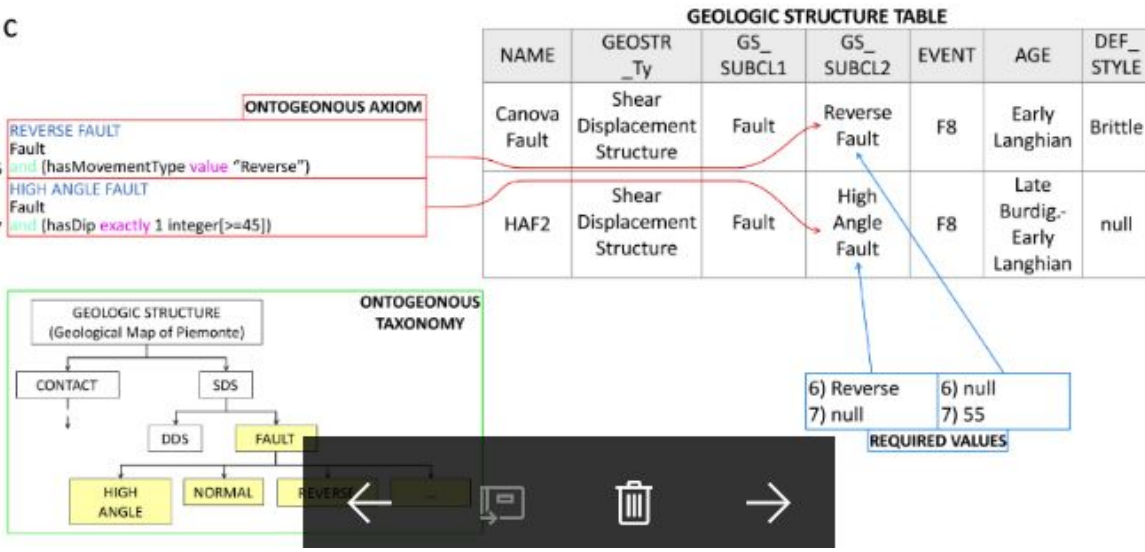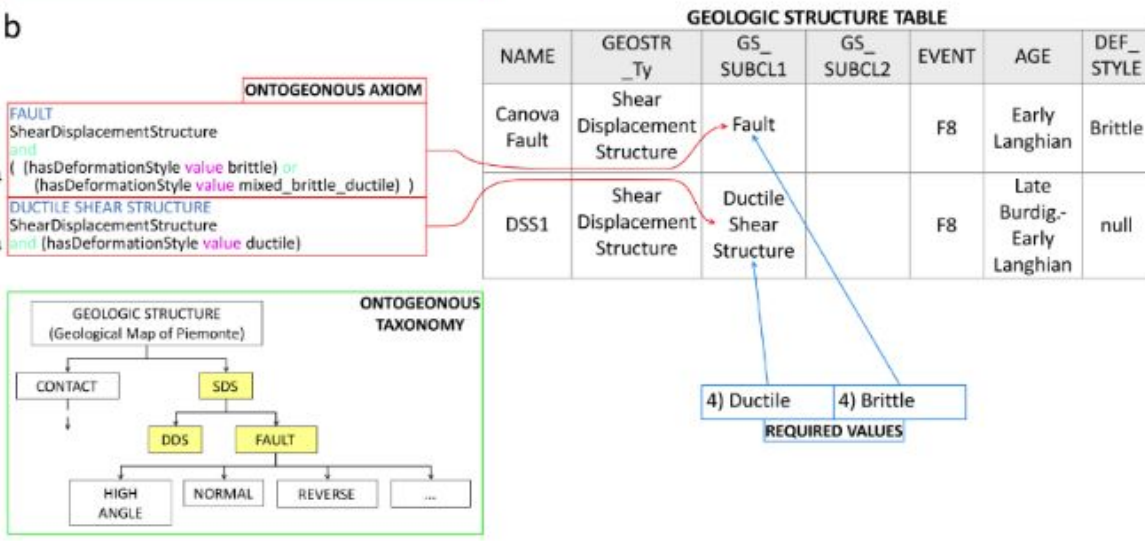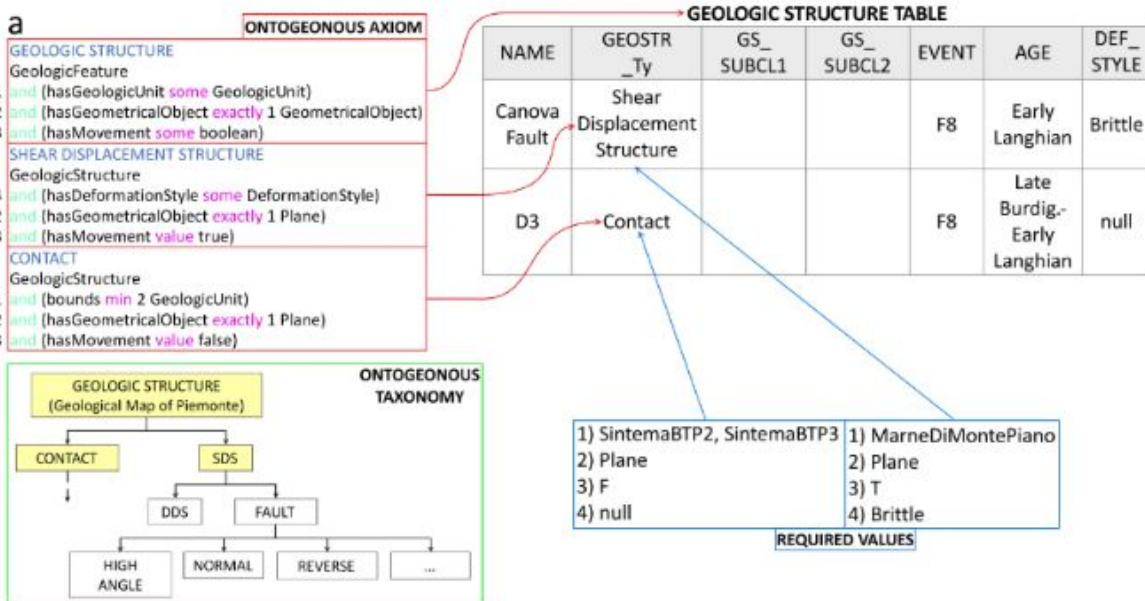|	
6) Reverse	6) null
7) null	7) 55

Fig.7 Extracts of GeologicStructure table. The tables (in gray) show the taxonomy columns of the OntoGeoBase GeologicStructureType (a), GeologicStructure_Subclass1 (b) and GeologicStructure_Subclass2 (c). The classes that the instance belongs to are highlighted in yellow inside the schema of the relevant taxonomy of OntoGeonous (green squares) , while their axiom are in the red square box labelled "OntoGeonous Axioms". The values for the classification of the instance are listed in the blue squares "Required Values", in which the numbers refer to the ID of "OntoGeonous Axioms" properties (see also Tab.1). For the acronyms of the columns see Tab.1 and Note[30]. SDS and DSS are the abbreviations for "ShearDisplacementStructure" and "DuctileShearStructure" respectively.

6. SUMMARY OF THE METHOD AND DISCUSSION

Finally, we sum up the method we have implemented to represent the knowledge underlying a geological map through a controlled vocabulary licensed by ontology, claimed as a basic condition for the perspective of data sharing, and discuss pros and cons. The aim of the paper is to present how the geological knowledge can be represented without ambiguities and compliant with standard vocabularies through the use of an ontology. The properties in the axioms represent explicit relations over the geological feature contained in a geological map. The most important encoded concept is the one regarding the geologic event, as it is crucial to the reconstruction of the evolution and geologic history of the map area. However, we have to remark that our method does not yet cover all the geological concepts commonly used in a traditional geological mapping processes, as *geological map (and the associated map legend) is not simply the sum of the items stored in the relevant geodatabase*. Rules for the graphical representation of the geological features can be derived also from the general, not-encoded, geologic knowledge of the mapped region, as well as of its evolution through geological times. Although many efforts have been addressed to encode the most of terms and concepts, some knowledge

31

contents usually referred to draw a geological map (e.g "paleogeographic domain" or

"post-orogenic event", see Fig.4c "Not Encoded Concepts") although not encoded yet, are

still necessary and thus widely used. They will possibly be encoded in some ontological

format in the future and contribute directly to the method.

Finally, to apply this method to a real project of geological mapping we have proposed a

method to translate the ontology into database columns and related data entry form for the

compilation of the database itself. The whole method is summarized in Fig.8.
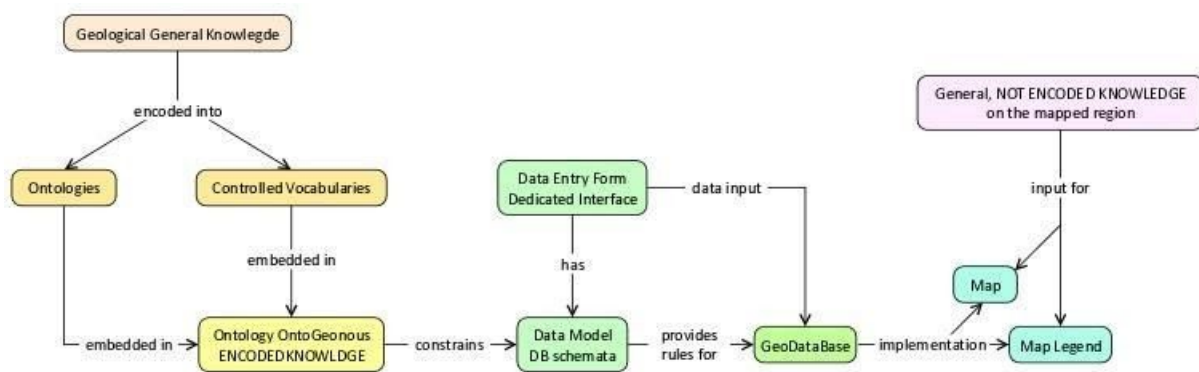
Fig.8 Conceptual map of the proposed methodology workflow.

The proposed approach, can lead to an improvement of the representation of geological

knowledge for three main groups of reasons:

a) *Encoding of geological theories applied at large*: the organization provided by

OntoGeonous allows to represent in a consistent way several types of knowledge

concerning on the geological mapping process. The classification of the geological

objects as instances belonging to some classes and the encoded representation of

theis properties (which explicit their mutual relations) favours the alignment

between the peculiar attributes given to the geologic features of the map and the

more general knowledge about the geological evolution of the mapped region. The

latter is often very complex and it is usually conceived by geologists at the

knowledge level of a "GeoScienceTheory" sensu Brodaric (2012), in which there are
concepts (prototypes) founded on theories, and real objects (individuals) depicted
on the map and founded on significant data. This means that the inference chains
necessary to derive a Map Legend from a model of the real world (concerning a
given mapped area) consist in theory-driven propositions and predictions, as well as
in data-driven inductions. In traditional geological maps, a graphic legend
(description of the instances and their parental relations) and a number of ancillary
maps and graphs are usually used to share the understandings of the geological
setting of the mapped area (with relevant uncertainty, Balestro and Piana, 2007) and
how it repeatedly changed through geological times. In the proposed approach, the
discretization of the geological evolution of a region into some encoded Geologic
Events, which have explicit properties (age, process, environment…) and into
encoded hierarchy relations between the mapped features, should allow for a formal
representation, in an ontology-driven database structure, of the geological evolution
of that region. In this frame, the geologic features can be clearly represented as they
are now (by their geometrical and physical properties) and also as the result of some
geologic events which have determined their relations, and from which they have
inherited some properties.

b) *Encoding of geological objects by ontological axioms and inferences into table*
attributes. OntoGeoBase is designed over the structure of the ontology: every
information that is planned to be recorded as a column in the tables is a
transposition of the axioms of the classes that include all the types of mapped
features. Hence, the reasoning processes of the ontology, following its rigorous

rules, can support the classification of the items, and can be re-traced back afterwards, to verify or update the inferences.

c) The OntoGeoBase structure can be integrated on every kind of GIS system, in order to fill in the geo database directly in the field, forcing the geologist to classify the instances in the proper logical and semantic framework from the very beginning. The ontology OntoGeonous relies on a number of properties that relate the geological features one another. The same links are expressed in the OntoGeoBase through a number of columns meant to be filled with defined values listed in drop-down menu of a Data Entry Form. These ensure the exclusive use of encoded terms, avoiding the use of non-standard terms or definitions. However, as mentioned above, not all the geological knowledge is already encoded, and many non-formalized terms and concepts are still used in the maps. For instance, in several geomapping projects the conceptual alignment between the main map and its ancillary maps is usually poor and the use of standard concepts in a non-fully conformable way is recurrent. An effort toward an encoding of knowledge could thus lead to a wider alignment between the implicit knowledge used to conceive a geological map and its actual representation, consisting of many individual mapped features, each one with its own formalized attributes.

7. CONCLUSION

In this paper, we propose a method for the representation of the geological knowledge underlying a map, supported by the formal structure of an ontology. Ontologies address human thinking in a formal language and a machine-readable representation, open to reasoning procedures and traceability of information.

The feasibility of the method is demonstrated through the implementation of an

ontology-driven design database; axioms, properties, and relations of the ontology become

columns in the database, thus every inserted item is described and classified according to

the rules of the ontology.

The use of the standard vocabularies enforced by the ontology provides unambiguous

definitions for the items, but also for the types of relations among them. This would

guarantee a rigorous semantic organization and provide rules for the reasoning processes,

which are essential for the item classification task.

The proposed approach aims to favour data sharing and interoperability, and a suitable use

of data for different purposes and applications. This methodology would make easier the

replicability of the logic processes of knowledge acquisition and will get the inference

processes, which are intrinsic to the geological mapping task, more explicit.

Computer Code Availability

To demonstrate the feasibility of the approach presented here, we have devised an

interface simulation (shown in Fig.6) as a proof of concept. However, the development of an

original running software has been not yet done.

The interface (Data Entry Form) has been created using a tool of the Open Source software

Quantum GIS 2.8.9 "Wien", downloaded from https://www.qgis.org/it/site/.

OntoGeobase was also developed using the above-cited software.

OntoGeonous was implemented using the Protégé resource, which is supported by grant GM10331601 from the National Institute of General Medical Sciences of the United States National Institutes of Health (Musen, 2015), downloaded from

https://protege.stanford.edu/

All the software resources mentioned above are available through the URL

https://di.unito.it/ontogeonouscode

In particular, at the URL, it is possible to find:

1) the OntoGeonous ontology (file ontogeonous.owl), which can be opened through the Protégé editor;

2) the database (OntoGeoBase) structure through a QGIS project folder (zip file GISproject_Database.zip) by clicking on the file CaGeo_DB.qgs it will be possible to access the structure of the Attribute Table of layers GeologicStructure, LithostratigraphicUnit and Lithotectonic Unit;

3) the Data Entry Form structure through a QGIS project folder (zip file GISproject_DataEntryForm.zip) by clicking on the file CaGeo_DEF.qgs.

REFERENCES

Abel, M., Perrin, M., Carbonera, J., 2015. Ontological analysis for information integration in geomodeling. Earth Science Informatics. 8. 10.1007/s12145-015-0211-9.

Ambrosi, C., Scapozza, C., 2015. Improvements in 3-D digital mapping for geomorphological and quaternary geological cartography. Geogr. Helv. 70, 121–133. https://doi.org/10.5194/gh-70-121-2015

Babaie, H.A., Oldow, J.S., Babaei, A., Lallemant, H.G.A., Watkinson, A.J., 2006. Designing a

modular architecture for the structural geology ontology. Geoinformatics Data to Knowl.

Geological Society of America Special Papers, v.397, p. 269-282. 10.1130/2006.2397(21)

Balestro, G., Piana, F., 2007. The representation of geological knowledge and uncertainty in

databases of GIS geological maps. Boll. Soc. Geol. It. (Ital. J. Geosci.) 126 (3), 487–495.

Boyd, T., 2016. Ontology of Geological Mapping. Thesis, Georgia State University.

https://scholarworks.gsu.edu/geosciences_theses/94

Brodaric, B., 2004. The design of GSC FieldLog: Ontology-based software for computer aided

geological field mapping. Comput. Geosci. 30, 5–20

https://doi.org/10.1016/j.cageo.2003.08.009

Brodaric, B., Gahegan, M., Harrap, R., 2004. The art and science of mapping: computing

geological categories from field data. Computers & Geosciences, 30 (7), 719–740.

Brodaric, B., Gahegan, M., 2006. Representing geoscientific knowledge in cyberinfrastructure:

Some challenges, approaches, and implementations. Geological Society of America

Special Papers, v.397, p. 1-20. 10.1130/2006.2397(01).

Brodaric, B., 2012. Characterizing and representing inference histories in geologic mapping.

International Journal of Geographical Information Science, 26 (2), 265-281

Buller, G., 2005. A Conceptual Approach to the Development of Digital Geological Field Data

Collection. https://www.researchgate.net/publication/242122464

Carbonera, J., Abel, M., Scherer, C., 2015. Visual interpretation of events in petroleum

exploration: An approach supported by well-founded ontologies. Expert Systems with

Applications. 42. 2749. 10.1016/j.eswa.2014.11.021.

Cox, S.J.D., Richard, S.M., 2015. A geologic timescale ontology and service. Earth Science

Informatics (2015) 8: 5. https://doi.org/10.1007/s12145-014-0170-6

De Donatis, M., Alberti, M., Cesarini, C., Menichetti, M., Susini, S., 2016. Open source GIS for

geological field mapping: research and teaching experience. Peer J Prepr. 4, e2258v3.

https://doi.org/10.7287/peerj.preprints.2258v3

Dey, S., Ghosh, P., 2008. GRDM-A digital field-mapping tool for management and analysis of

field geological data. Comput. Geosci. 34, 464–478.

https://doi.org/10.1016/j.cageo.2007.05.014

Frodeman, R., 1995. Geological reasoning: geology as an interpretive and historical science.

Geological Society of America Bulletin, 107 (8), 960–968.

Harrison, J.M., 1963. Nature and significance of geological maps, in Albritton, C.C., ed., The

Fabric of Geology: Stanford, California, Freeman, Cooper & Co., p. 225–232.

Hintersberger, E., Iglseder, C., Schuster, R., & Huet, B., 2017. The new database "Tectonic

Boundaries" at the Geological Survey of Austria. Jb. Geol. B.-A.,157, 1-4, 195-207, ISSN

0016-7800

Holland, D., 2001. The Digital National Framework - Evolving a Framework for Interoperability

across All Kinds of Information: Southampton, Ordnance Survey White Paper, 39 p.

Howard, A.S., Hatton, B., Reitsma, F., Lawrie, K.I.G., 2009. Developing a geoscience knowledge

framework for a national geological survey organisation. Comput. Geosci. 35, 820–835.

https://doi.org/10.1016/j.cageo.2008.06.004

Hwang, J., Nam, K.W., Ryu, K.H., 2012. Designing and implementing a geologic information

system using a spatiotemporal ontology model for a geologic map of Korea. Comput.

Geosci., 48, 173–186.

Johnson, B.R., Brodaric, B., Raines, G.L., Hastings, J.T., Wahl, R., 1999. Digital Geologic Map Data

Model Version 4.3. https://ngmdb.usgs.gov/www-nadm/about.html

Laxton, J.L., Becken, K., 1996. The design and implementation of a spatial database for the

production of geological maps. Comput. Geosci. 22, 723–733

https://doi.org/10.1016/0098-3004(95)00128-X

Laxton, J.L., Duffy, T.R., 2011. Developing and implementing international geoscience

standards—A domestic perspective. Geol. Soc. Am. 2482, 87–97.

https://doi.org/10.1130/2011.2482(09)

Lombardo, V., Piana, F., Mimmo, D., Mensa, E., Radicioni, D.P., 2017. Semantic Models for the

Geological Mapping Process. In: F. Esposito et al. (Eds.): AI*IA 2017, Advances in Artificial

Intelligence, LNAI 10640, Proceedings XVIth Int. Conf. Italian Ass. Artificial Intelligence,

Bari, Italy, November 14–17, 2017, 123, 295–306, Springer International Publishing.

Lombardo, V., Piana, F., Mimmo, D., 2018. Semantics–informed geological maps: Conceptual

modeling and knowledge encoding. Comput. Geosci. 116, 12–22.

https://doi.org/10.1016/j.cageo.2018.04.001

Loudon, T.V., 2009. Four interacting aspects of a geological survey knowledge system.

Computers & Geosciences 35, 700-705 doi:10.1016/j.cageo.2007.12.009

Loudon, T. V., 2011. A scenario for systems geology: suggestions concerning the emerging

geoscience knowledge system and the future geological map. Nottingham, UK, British

Geological Survey, 375pp. (RR/11/005).

Ma, X., 2011. Ontology Spectrum for Geological Data Interoperability. PhD Dissertation.

University of Twente.

Ma, X., Asch, K., Laxton, J.L., Richard, S.M., Asato, C.G., Carranza, E.J.M., Van Der Meer, F.D.,

Wu, C., Duclaux, G., Wakita, K., 2011. Data exchange facilitated. Nat. Geosci. 4, 814.

https://doi.org/10.1038/ngeo1335

McCaffrey, K.J.W., Jones, R.R., Holdsworth, R.E., Wilson, R.W., Clegg, P., Imber, J., Holliman, N.,

Trinks, I., 2005. Unlocking the spatial dimension: digital technologies and the future of

geoscience fieldwork. Journal of the Geological Society, 162, 927-938, 1 December

2005, https://doi.org/10.1144/0016-764905-017

Mookerjee, M., Vieira, D., Chan, M.A., Gil, Y., Pavlis, T.L., Spear, F.S., Tikoff, B., 2015. Field Data

Management : Integrating Cyberscience and Geoscience. Eos (Washington. DC). 96,

18–21. https://doi.org/10.1029/2015EO036703

Musen, M.A., 2015. The Protégé project: A look back and a look forward. AI Matters.

Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4),

June 2015. DOI: 10.1145/2557001.25757003.

NADM, 2003. NADM Conceptual Model 1.0 A Conceptual Model for Geologic Map Information:

U.S. Geological Survey Open-File Report 2004-1334, North American Geologic Map Data

Model (NADM) Steering Committee, Reston, VA, USA (2004)

Navigli, R., Velardi, p., 2005. Structural Semantic Interconnections: a Knowledge-Based

Approach to Word Sense Disambiguation. IEEE Transactions on Pattern Analysis and

Machine Intelligence (TPAMI), 27(7).

Neuendorf, K.K.E., Mehl, J.P. & Jackson, J.A., 2005. Glossary of Geology – 5th Edition, American

Geological Institute, 1-779, ISBN#: 922152-76-4

Pavlis, T.L., Langford, R., Hurtado, J., Serpa, L., 2010. Computer-based data acquisition and

visualization systems in field geology: Results from 12 years of experimentation and

future potential. Geosphere 6, 275–294. https://doi.org/10.1130/GES00503.1

Pavlis, T.L., Mason, K.A., 2017. The new world of 3D geologic mapping. GSA Today 27, 4–10.

https://doi.org/10.1130/GSATG313A.1

Piana, F., Fioraso, G., Irace, A., Mosca, P., D'Atri, A., Barale, L., Falletti, P., Monegato, G.,

Morelli, M., Tallone, S., Vigna, G.B., 2017 (a). Geology of Piemonte region (NW Italy,

Alps–Apennines interference zone). J. Maps 13, 395–405.

https://doi.org/10.1080/17445647.2017.1316218

Piana, F., Lombardo, V., Mimmo, D., Mulazzano, E., Barale, L., D'Atri, A., Irace, A., Morelli, M.,

Mosca, P., Tallone, S., 2017 (b). The geodatabase of the Piemonte geological map:

conceptual design for knowledge encoding. Rendiconti Online Società Geologica Italiana
42, 85–89

Raskin, R., Pan, M., 2005. Knowledge Representation in the Semantic Web for Earth and
Environmental Terminology (SWEET). Comput. Geosci. 31 (9), 1119–1125 sweet -
http://ceur-ws.org/vol-83/sia_7.pdf.

Raskin, R., 2006. Development of ontologies for earth system science. Geological Society of
America Special Papers, v.397, p.195-19910.1130/2006.2397(14).

Raymond, O.L., Boisvert, E., Sen, M., Duffy, T., Laxton, J.L., Cipolloni, C., Vuollo, J., Richard, S.,
Simons, B.A., Ritchie, A.B.H., 2016. GeoSciML version 4. 0 - a coming of age. 35th Int. Geol.
Congr. Cape Town, South Africa 737.

Reitsma, F., Laxton, J., Ballard, S., Kuhn, W., Abdelmoty, A., 2009. Semantics, ontologies and
eScience for the geosciences. Comput. Geosci. 35, 706–709.
https://doi.org/10.1016/j.cageo.2008.03.014

Richard, S.M., 2003. Geologic map database implementation in the ESRI Geodatabase
environment. Digital Mapping Techniques '03 — Workshop Proceedings U.S. Geological
Survey Open-File Report 03–471

Richard, S.M., 2006. Geoscience concept models. Geological Society of America Special Papers,
v.397, p. 81-107. 10.1130/2006.2397(07).

Sen, M., Duffy, T., 2005. GeoSciML: development of a generic Geoscience Markup Language.
Computers and Geosciences, 31 (9), 1095–1103.

Simons, B.A., Boisvert, E., Brodaric, B., Cox, S.J.D., Duffy, T.R., Johnson, B.R., Laxton, J.L.,
Richard, S.M., 2006. GeoSciML: enabling the exchange of geological Map Data. ASEG Ext.
Abstr. 1–4. https://doi.org/10.1071/ASEG2006ab162

Sinha, A.K, Zendel, A., Brodaric, B., Barnes, C., Najdi, J., 2006. Schema to ontology for igneous
rocks. Special Paper of the Geological Society of America. V.397. p.169-182.
10.1130/2006.2397(12).

Thum, L., De Paoli, R., 2015. 2D and 3D GIS-based geological and geomechanical survey during

 tunnel excavation. Eng. Geol. 192, 19–25. https://doi.org/10.1016/j.enggeo.2015.03.013

Tiainen, M., Rantala, O., Kahra, A. & Kuosmanen, E., 2008. Data management and map

 production in the geological mapping projects in Mozambique. Geological Survey of

 Finland, Special Paper 48, 23–33.

Tudor, G., Gheucă, I., 2009. GIS database model for geological maps. Romanian Journal of

 Mineralogy, 84, 86–88.

Uschold, M., 2015. Ontology and database schema: What's the difference? Applied Ontology

 10(3-4): 243-258

van Gasselt, S., Nass, A., 2010. Object-Relational Datamodel components for geologic mapping

 conduct. Special Joint Symposium of ISPRS Technical Commission IV and AutoCarto 2010,

 Nov 15-19, 2010, Orlando, Florida, USA.

Whitmeyer, S., Nicoletti, J., De Paor, D., 2010. The digital revolution in geologic mapping. Gsa

 Today. 20. 4-10. 10.1130/GSATG70A.1.

Zhong, J., Aydina, A., McGuinness, D.L., 2009. Ontology of fractures. J. Struct. Geol. 31,

 251–259. https://doi.org/10.1016/j.jsg.2009.01.008