International Cartographic Association
Association Cartographique Internationale

A lightweight ontology for landmarks to assist rescue in mountainous areas

Ana-Maria Olteanu-Raimond [a,*], Mattia Bunel [a], Catherine Domingués [a], Cécile Duchêne [a], Laurence Jolivet [a], Sébastien Mustière [a], Marie-Dominique Van Damme [a], Olivier Favre [b], Véronique Gendner [a]

[a] *LASTIG, Univ Gustave Eiffel, IGN-ENSG, 94165-F Saint-Mande, France, ana-maria.raimond@ign.fr, mattia.bunel@ign.fr, catherine.domingues@ign.fr,cecile.duchene@ensg.eu, laurence.jolivet@ign.fr, sebastien.mustiere@ensg.eu, marie-dominique.van-damme@ensg.eu, veronique.gendner@e-tissage.net*
[b] *High Mountain Gendarmerie Platoon of Grenoble, 38420 Le Versoud, olivier.favre@gendarmerie.interieur.gouv.fr*

* Corresponding author

Abstract: When people are injured or lost in mountains during outdoor activities and when web-based locations are not available, they locate themselves by describing their environment, routes and activities. The description of their location is done using landmarks and spatial locations (e.g., “I am located in front of Punay Lake”, “I am near a protected area”). Landmarks used can be named (e.g., “Punay Lake”) or unnamed if the landmark has no name or if the victim does not know it (e.g., "area lake"). Landmarks are represented in geographic databases by name (if possible), type and geometry. To reduce the heterogeneity of landmark types present in oral language and geographic databases representing landmarks, and thus improve locating victims, our goal is to define a controlled vocabulary for landmarks. In this research, we present a lightweight ontology (i.e. ontology having generally less complexity and does not express formal constraints) of landmarks, named Landmark Ontology (OOR), describing landmark types. It is an application ontology, i.e. it is designed to support mountain rescue operations. The ontology construction is adapted from the SAMOD methodology for engineering ontology development and involves researchers and experts from mountain rescue teams. The construction of OOR is composed of four main phases: knowledge acquisition, conceptual formalisation, implementation, and testing. The implementation phase is carried out by an iterative and collaborative approach and using four formalised sources of knowledge (a landmark and a landform ontologies, and two other domain vocabularies), an un-formalised taxonomy of outdoor activities, and five authoritative and volunteered geographic information sources representing geographic data. The landmarks ontology contains 543 classes associated with 1739 labels: 1086 *prefLabel* (preferential label) in French and English, 321 *altLabel* (alternative label) in French, and 332 *altLabel* in English. The depth of the ontology varies from four for land cover, hydrological and land subdivision landmark types), to six for landform types, and eight for building types. Although the use of ontology is broader, in this paper we illustrate and test its use through three applications in the context of mountain rescue operations: semantic mapping, data instantiation and data matching.

Keywords: lightweight ontology, landmarks, semantic, mountain rescue operation, geographic database.

1 Introduction

Rescue in mountainous areas is a public right in many countries. France is one of these countries and the High Mountain Gendarmerie Platoon (PGHM in French) is the public service operating rescue. The alert phase is crucial, consisting of a dialogue between the person calling for help (i.e. victim(s) or third party) to determine both the appropriate means to be deployed for the rescue operation and the location to send the helicopter if the victims are injured, or to guide them to find the route, if they are lost. Nowadays, thanks to the development of the location tracking technologies, PGHM is able to locate the victims by using a location-based service application. Nevertheless, there are cases where this GPS based location is not possible (e.g., lack of a phone with GPS location or 3G signal, low battery power for 3G signal, the person calling is a third-party worried about a missing person in the mountain). In these specific cases, victims or third-parties describe their location to the rescue team by using clues composed of landmarks and locational relationships. In the clue “I am in front of the Robert hut”, the landmark is “Robert hut” and the location relationship is "in front of". This description is known in the literature as ‘indirect georeferencing’ (Hill and Zheng, 1999). One of the needs to improve victim location is to identify and organise the types of landmarks that may be used by victims in mountain to locate themselves (e.g., lake, summit, hut, waterfall), on the one hand, and the types of location relationships (e.g. "opposite", "north of"), on the other hand. This paper focuses only on types of landmarks.

Our goal is to define a controlled vocabulary to reduce the heterogeneities of the various types of landmarks present in oral language and geographic databases representing landmarks. The aim of the vocabulary is to ensure interoperability of tools and methods for localizing victims in mountain areas and to provide the widest possible entry point for searching by means of types of landmarks. For example, the rescue operator[1] can search by using a geo-visualisation tool which integrates the ontology a specific type of landmarks (e.g., waterfall) to visualise all data representing waterfalls from an area to ask a new question or to better understand the geographic context of the victims (Viry, 2022). The spatialisation of clues (i.e., define a corresponding area where the victim may be located knowing the clue) requires formalised and non-ambiguous types of landmarks (Bunel et al., 2019). The controlled vocabulary is also used to extract landmarks from text (Gaio and Moncla, 2019) to enrich landmark data sources or to compute semantic data quality indicators for heterogeneous landmarks data sources including crowdsourced data (Van Damme and Olteanu-Raimond, 2022) as well as for computing semantic similarities between types of landmarks needed for data matching processes (Van Damme and OlteanuRaimond, 2022).

To reach this goal, the very well-known OWL formalism is used to build a *lightweight ontology* (LO), define landmark types, and organise or use them in mountain areas. An ontology is defined as "a formal representation of knowledge that forms a shared conceptualisation of a given domain" (Hogan, 2020). A *lightweight* ontology is an ontology having generally less complexity and does not express formal constraints. It can be considered an ontology that consists of a backbone structure (i.e., a hierarchy of concepts and a set of relations between concepts) and is used to capture the semantic of a specific application (Davies, 2010). Thus, our proposed lightweight ontology is an application ontology (which is defined as an ontology focusing on a specific domain, i.e mountain rescue). It describes and organises landmark types to assists rescue in mountainous area. Note that only “rdfs:subClassOf” relationships are defined.

The paper is organised as follows. Section 2 presents the literature review describing both formalisation of landmarks and methods used to construct ontology. Section 3 describes the methodology we propose to build the landmark ontology for mountain rescue. Before concluding, section 4 describes the ontology and illustrates some use cases.

2 Related work of landmark representation

Landmarks can be defined from different points of view. From a cognitive point of view, landmarks are remarkable real-world elements that people use to understand space and better orient themselves (Lynch, 1960). From a space description perspective, landmarks are considered as references to describe the space (Zhou et al., 2017), whereas, in the field of information retrieval, landmarks represent placenames considered as anchors to geocode proper names (Moncla et al., 2014).

Based on these definitions, we define a landmark as a landscape reference feature, named or unnamed, natural or built, which can be seen, known or used to practice an outdoor activity and used by victims in mountainous areas for locating themselves. In this work, we focus on landmarks which can be represented in a geographic database by features.

Different taxonomies or ontologies for landmarks are already proposed. GeOnto is an ontology representing landmark types (Mustière et al., 2011), where concepts were first semi-automatically extracted by natural language processing of specifications of databases and textual descriptions of routes, and completed by terms coming from RAMEAU thesaurus[2]. The concepts has been organised manually. Despite the fact that GeOnto is formalized by using the OWL and RDFS (Resource Description Framework Schema) formalisms, it was not published on the Web. Another French ontology defining elements of the territory and its infrastructure was recently published[3] (Abadie et al., 2019). This ontology was not retained in our work because it describes only elements from a topographic point of view described by BDTOPO[4] (an authoritative database produced by the French National Mapping Agency).

Land Form Reference (LFR) Ontology was defined for automated mapping and delineation of landmarks and their formalized categories from Digital Elevation Models (DEM)(Sinha et al., 2018). The landforms are organized according to their shape and their relationships on the Earth. The ontology is not published on the web, only the conceptual schema is available. Note that the types of landforms are not exhaustive in LFR ontology. Some concepts from this ontology were used in our work to classify landforms in mountainous area. Gazetters indexe localized place-names and may be considered as a relevant formalization for landmarks. Each place-name is characterized at least by three main characteristics: toponyms[5] (i.e., name of the place-name), the feature type (i.e., type of the feature), and geographic coordinates (Hill, 2006). However, gazetteers have relied on both spatial and semantic indexing, as for example Geonames[6], whose ontology has a depth of 2 for the feature codes structure. 50k Gazetter[7] is based on a simple ontology composed of 13 concepts of which only 10 define geographical types of features. DBPedia gazetteer (Lehmann et al., 2015) is based on an

[1] A dialog is established by the rescue operator with the person calling for help. Relevant terms are extracted from the spoken interaction by the operator.

[2] https://rameau.bnf.fr/

[3] https://data.ign.fr

[4] https://geoservices.ign.fr/bdtopo

[5] Several alternative names, endonyms and exonyms, may be proposed.

[6] https://www.geonames.org/ontology/documentation.html

[7] https://data.ordnancesurvey.co.uk/ontology/50kGazetteer/

31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa. This contribution underwent double-blind peer review based on the full paper. https://doi.org/10.5194/ica-adv-4-15-2023

ontology[7], defined by a bottom-up approach, i.e., it was generated by semantically annotating the data extracted from Wikipedia. The last three ontologies are published on the Web, but do not meet our needs well. 50k Gazetteer and Geonames only define some concepts and there is no sophisticated hierarchy in either of them. Concerning DBPedia, the main inconvenience is linked to the omission (only types of place-names are defined) and commission (i.e., most of concepts are out of our scope), as well as the flat hierarchy of concepts representing natural features.

We have described above some ontologies considered as relevant for our needs. We have noticed that most of them are not published on the Web, which is an inconvenience for their re-use. Moreover, the organisation of concepts, when there is one, is less specialised and not adapted to mountain rescue needs, and is based on a flat hierarchy. The ontology closest to our needs, including concepts (few concepts are out of scope), is GeOnto. Thus, GeOnto was used as a first input to our work.

Research papers propose an overview of different methodologies to define an ontology (Gomez-Perez, 1999; Sure et al., 2009; ElHassouni and Qadi, 2022). The last concludes that no methodology receives consensus and suggests using ontology design patterns. This methodology is not adaptable to our purpose since the landmark types are difficult to be split into modules. The agile methodology for ontology development (AMOD) (Abdelghany et al., 2019) is often criticized for being time consuming; but its simplified version (SAMOD) (Peroni, 2016) which is an iterative methodology, bringing together designers and end users, seems to meet our needs, and is easy to adapt.

3 Approach to build the landmarks ontology

The *lightweight* ontology describing types of landmarks is named Ontology of Landmarks Objects (OOR, "Ontologie d'objets de repère" in French). Lightweight is not part of the name since our aim is to enrich OOR by adding new relationships such as "confusion", "is part of". Hereinafter, we will refer to this ontology as OOR. The methodology for defining OOR is adapted from SAMOD methodology: it is made iteratively; the domain experts are mountain rescue teams (PGHM) and ontology engineers are researchers, authors of the current paper.

The approach is composed of four main phases: 1) Knowledge acquisition, 2) Conceptual formalization, 3) Implementation, and 4) Testing after a first release of OOR. This last phase consists in applying the ontology for different applications and is described in Section 4.

3.1 Knowledge acquisition

The goal of this first phase is to acquire knowledge necessary to define the application ontology. It includes both the researchers and rescue teams. For our OOR, we identified four sources of knowledge for formalizing and organizing OOR and three additional sources for providing an ontology as exhaustive as possible. Source of knowledge are:

3.1.1 *GeOnto Ontology*

As described in Section 2, GeoOnto describes landmarks from a topographic point of view. It is formalized in OWL format and contains 890 classes; each class is described by labels in French and English (not systematically) and organised into different parent classes. The first level of classes contains "topographic artificial entity" and "topographic natural entity". The main reasons for choosing GeoOnto are it is already in OWL format and it contains many topographic landmarks compliant with authoritative databases which will be easily used to instantiate OOR.

3.1.2 *Land Form Reference Ontology* As mentioned in Section 2, LFRO provides a description of landforms being formalized by using UML classes (Sinha et al., 2018) and contains 22 classes. A landform is defined as being part of the surface of the Earth (*MatDependentLF class*) or physically supported by or on the surface (*SupportedLF class)*. The landforms are categorized based on theirs shapes: convex, concave, horizontal vertical planar, and saddle. Although LFRO does not contain types of landmarks (e.g., mountain, valley), we used this categorisation as basis to organise the natural landmarks since it fits the way geography science defines landforms.

3.1.3 *SANDRE dictionary*

The SANDRE dictionary[8] (i.e., National Administration Service for Water Data and Repositories) describes hydrography entities present on the French territory to produce a national repository for locating water-related data. The use of the dictionary to organise hydrography landmarks was recommended by PGHM for three main reasons that it is: INSPIRE compliant; recognized as a reference at national scale; and used by data producers and users, including PGHM. The hydrography elements are classified into different categories. Only five of them are selected for OOR: water course, wetland, hydrographic node, water body. We added a new category named *permanent snow and ice surface* defined by in the dictionary proposed by the French National Committee of Geographic Information (CNIG).

3.1.4 *Outdoor activities nomenclature*

The National Resource Center Sports in Nature [9] published a list of outdoor activities such as running, paragliding, rafting, etc. Most of these activities include outdoor infrastructure or signage. This source of knowledge is relevant for our needs since the outdoor structure for outdoor activities are well known both by practitioners and rescue teams. Most of outdoor activities are taken into account in OOR.

3.1.5 *Knowledge from domain experts*

A team from PGHM of Grenoble organised a brainstorming workshop with rescue staff to define a list of landmarks known in mountain environment. In total, 183

[7] http://mappings.dbpedia.org/server/ontology/classes/

[8] http://id.eaufrance.fr/ddd/ETH/2002-1

[9] https://www.sportsdenature.gouv.fr/

types of landmarks are identified about well-known (e.g., lake, summit), local (e.g. karst, moraine) and specialized (e.g., cairn, via ferrata) type of landmarks.

3.1.6 *Corpus of emergency calls*

A corpus of emergency calls is selected by PGHM contains 45 audio anonymised emergency calls registered during rescue operations. The calls are first transcribed by some authors of the current paper (Bunel et al., 2019). Then, types of landmarks used by the victims are manually identified and tagged. In total, this source of knowledge identified 52 types of landmarks.

3.1.7 *Authoritative and VGI databases* To populate OOR with geographic data, different potential data sources are analysed, one authoritative and three volunteered geographic information (VGI) sources. The authoritative database, named *Protected areas* (PA) is a local database produced by French public institutions managing natural landscapes for outdoors activities. In total 14 types of landmarks are identified.

OpenStreetMap (OSM) is one of the most famous volunteered geographic information (VGI) projects (Jokar Arsanjani et al., 2015). OSM provides geographic information at the global scale. Based on OSM wiki[10], a list of 216 types of landmarks was extracted. Camptocamp (C2C) is a VGI initiative dedicated to outdoor activities. Practitioners share data about landmarks (e.g., summit, lake, pass) and routes (e.g. GPS tracks). It contains 13 types of landmarks. The VGI website Refuges.info provides information about mountain shelters (e.g., name, position, number of places) as well as some types of landmarks (e.g., water points, summits). In total eight types of landmarks are identified. For detailed description of these above sources, please see (Van Damme and Olteanu-Raimond, 2022). Sources such as GeoNames and DBPedia were not considered a priority to be analysed since they contain only place names.

3.2 Conceptual formalisation

The goal of this second phase is to define the landmarks from a conceptual point of view to be implemented in the OOR. A type landmark is a "concept" defining a conceptual instance based on a set of properties, roles (or functions), or relationships. Each concept is represented in OWL by two types of classes. The first concerns *abstract* classes allowing organising and structuring the landmarks (e.g., "construction" class is an *abstract* class grouping classes such as building, routes). The second introduces classes representing the types of landmarks (e.g., summit, lake, hut), noted here as *typeLandmark* classes. Each class is described by a set of characteristics listed as follows.

a) URI: unified resource identifier.

b) A preferential label. The *prefLabel* proposed in RAMEAU thesaurus was used to define the preferential label of the landmarks. It is described by the property *skos:prefLabel* using Simple Knowledge Organization System (SKOS). The last offers a vocabulary for expressing the basic structure and content of concept schemes (Baker et al., 2013).

c) 0 to n alternative labels. The *altLabel* proposed by the same RAMEAU thesaurus was used to model the fact that a landmark can have different alternative labels such as synonyms, local or vernacular labels. The alternative labels are described by the property *skos:altLabel*.

d) A definition describing the signification of classes. We used the RDFS property *rdfs:comment.*

e) 0 to 1 provenance information. Allows specifying the provenance of the definition. The property *rdfs: idDefinedBy* was used.

f) 1..n subclass relationships. Defines the hierarchy between classes (*rdfs:subClassOf*). The hierarchy is not strict, as in some cases, classes can have several parents. Thus, the hierarchical organisation is a general principle and not a modelling constraint.

g) 0 to n links to other published ontologies. Identifies links with other classes from other ontologies. The property *owl:equivalentClass* is used.

Figure 1 illustrates an example of formalization for the concept summit.

```
<owl:Class rdf:about =
      "http://purl.org/choucas.ign.fr/oor#sommet">
      <rdfs:subClassOf rdf: resource =
       "http://purl.org/choucas.ign. fr / oor#convexe"/>
    <rdfs:comment xml:lang="fr"> Point le plus élevé d'une
          forme de relief</rdfs:comment>
    <rdfs:comment xml:lang="en">Highest point of a
       landform</rdfs:comment>
    <rdfs:isDefinedBy>https://www.cnrtl.fr/definition/sommet<
       /rdfs:isDefinedBy>
    <skos:prefLabel xml:lang="fr">sommet</skos:prefLabel>
    <skos:prefLabel xml:lang="en">summit</skos:prefLabel>
    <skos:altLabel xml:lang="fr">antécime</skos:altLabel>
    <skos:altLabel xml:lang="fr">bric</skos:altLabel>
    <skos:altLabel xml:lang="fr">cime</skos:altLabel>
    <skos:altLabel xml:lang="fr">montagne</skos:altLabel>
    <skos:altLabel xml:lang="en">mountain</skos:altLabel>
    <skos:altLabel xml:lang="fr">pointe</skos:altLabel>
    <skos:altLabel xml:lang="fr">sommité</skos:altLabel>
  </owl:Class>
```

Figure 1. Example of formalized concept "summit".

3.3 Implementation of OOR

The goal of this third phase is to build the ontology by using the knowledge acquired in phase 1 and applying the conceptualisation defined in phase 2. The ontology was built iteratively by researchers and experts. The followed steps are carried out (see Figure 2).

[10] https://wiki.openstreetmap.org/wiki/Main_Page

31st International Cartographic Conference (ICC 2023), 13–18 August 2023, Cape Town, South Africa. This contribution underwent double-blind peer review based on the full paper. https://doi.org/10.5194/ica-adv-4-15-2023

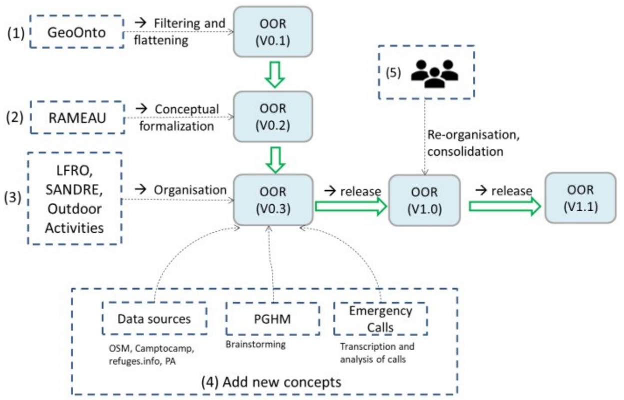

Figure 2. Implementation of OOR.

Step 1: Filtering and flattening GeOnto ontology. As we mentioned earlier, GeOnto is the base OOR. Since, the hierarchy and organisation of GeOnto is not adapted to our need, *abstract* classes are removed and only *typeLandmark* classes are retained. Then, some *typeLandmark* classes are removed if they are determined to be out of scope (e.g. country, city center, neighbourhood).

Step 2: Conceptual formalisation. The formalisms defined in phase 2 (section 3.2) is applied to all remained classes from GeOnto. Note that at this step, OOR has a flat hierarchy and few “rdfs:subClassOf” relationships are added.

Step 3: Organise *typeLandmark* classes by grouping them with respect to different criteria and adding *abstract* classes.

Criterion 1: a first criterion is to roughly distinguish between the main types of landmarks. Thus, the first level of the hierarchy was set by defining five *abstract* classes (see figure 3): *hydrography* (i.e., hydrographic features and buildings whose apparent function is related to hydrography), *cover* (i.e., represents plant and mineral elements that cover the ground and are relevant for mountain rescue), *landform* (i.e., describes natural landforms), *construction* (i.e., all human constructions whose apparent function does not allow them to be assigned to the hydrography, landcover and landform classes), and *land_divisions* (i.e., non-visible limits of the territory but represented in geographic databases such as municipality, industrial zone. Although built-up areas are included in a landcover product, here we choose to consider them separately since from a semantic point of view there are very different.

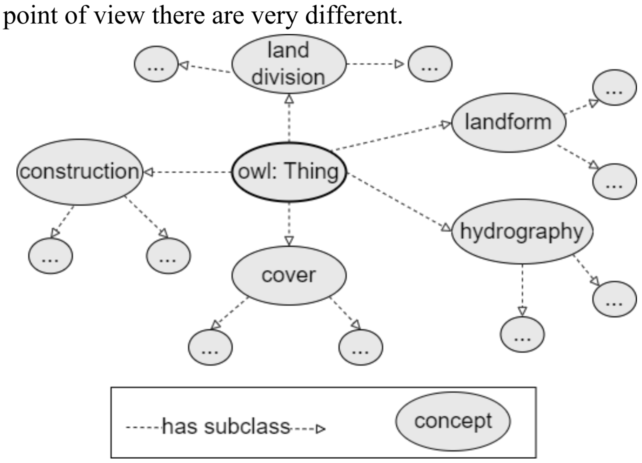

Figure 3. First level abstract classes of OOR.

Criterion 2: A second criterion is to organize the landmarks by considering their use and accessibility for built entities and shape for natural entities (except the natural hydrology landmarks such as natural lakes).

Criterion 3. To balance between generalisation and specialisation, a relevant criterion is specialisation. It allows choosing between specialising a concept into sub-concepts and associating several *altlabels* with the same concept. Thus, we limit the specialisation of a class in subclasses where terms are considered to be known by an informed public, and not necessarily only known to a specialist in the domain.

Finally, as mentioned in Section 3.1, for landforms, hydrography, and outdoor activities, LFRO ontology, SANDRE dictionary and outdoor activities nomenclature, are respectively used. Regarding LFRO ontology, much enrichment is made by adding supplementary levels. For example, based on the notion of slope and altitude, we defined new criteria to consider the fact that there are types of relief that are not perfectly horizontal or vertical.

Step 4: Add new classes from data sources described in Section 3.1. Here, the implementation process was guided by the need to harmonize different landmarks datasets having different semantics for the same entity in the real world and in some cases close but not entirely similar semantics (e.g., shelter, hut) or aggregated semantics (e.g., summit type encodes peaks, summits, rock mountain, mountain). As landmark types can be redundant from one source to another, each term was first labelled with a concept, and then only unique concepts were added in OOR.

Step 5: Consolidation and re-organisation of classes. In steps 1 to 4 only researchers worked together to build the ontology. In this final step, experts from the mountain rescue team worked collaboratively with researchers during face to face meetings to re-organise and consolidate the ontology. New abstract classes have been added to fit the mountain rescue expertise and improve the hierarchy levels. Some new *typeLandmark* classes are added and exhaustivity is checked. Moreover, the various applications of OOR (see section 4) have helped to consolidate and improve OOR by transforming an *altLabel* into a *typeLandmark* class or reducing ambiguity between concepts.

4 Results

This section describes the OOR ontology and gives some applications that are tested and assesses in the testing phase of the SAMOD methodology.

4.1 OOR overview

The ontology is defined as a Resource Description Framework (RDF) data model by using Protégé editor (Musen et al., 1993) and OWL format. The RDF format was completed with SKOS elements to define the *pref* and *alt* labels.

OOR contains 543 classes associated to 1739 labels: 1086 *prefLabel* in French and English, 321 *altLabel* in French and 332 *altLabel* in English.

The depth of the ontology varies from 4 (*cover*, *land_division*, and *hydrography*) to 8 (*construction*).

Particular attention was paid to the class of relief (depth equal to 6) because the activity of mountaineering uses relief and the class of buildings and facilities. According to the experience of the PGHM, the victims of the high mountain often use these classes to locate themselves easily (e.g., cliff, mountain, summit, peak, etc.) and isolated building (e.g., hut, habert), leisure infrastructure (e.g., picnic site, luge truck), etc. For landforms, the organisation of classes is mainly based on their shape.

Figure 4 illustrates the hierarchy for summit. A summit is a landform landmark and its class can be specified into two subclasses: pointed and rounded summits. Peak, is for example, a pointed summit.

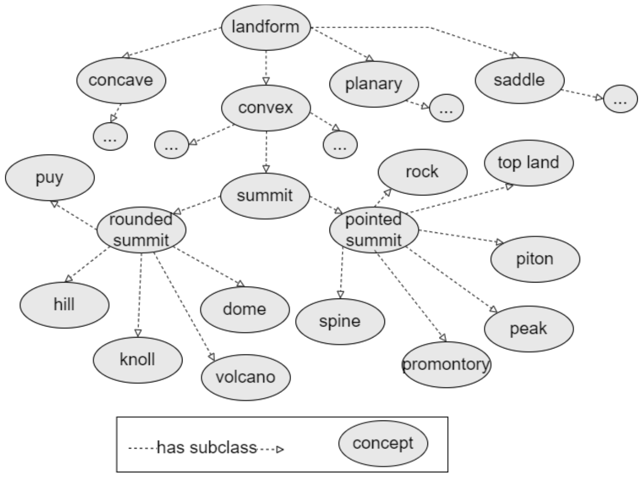

Figure 4 Illustration of landform organisation: summit concept.

Figure 5 illustrates the organisation of landmarks representing accommodation. Its sub-classes are organised according to accessibility, into isolated accommodations (e.g. shelter, hut refuge) and easy accessible (e.g. lodge, hotel) by using the main road network.

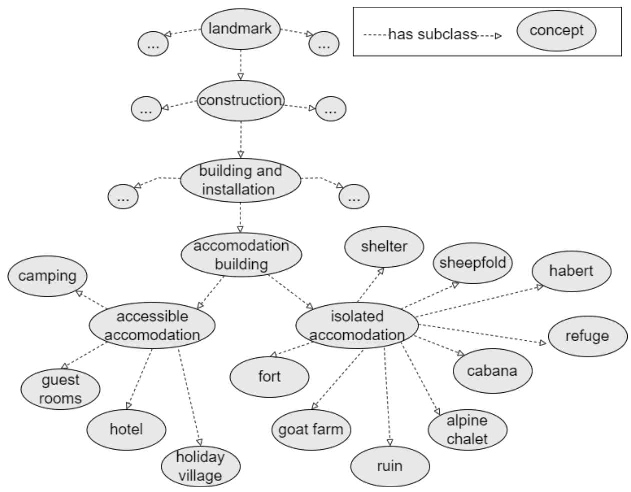

Figure 5. Illustration of accommodation building landform.

Finally, as mentioned earlier, some classes can inherit from two parent classes. Figure 6 shows such an example for class: reservoir. This one has two parents: freshwater catchment infrastructure and surface water.

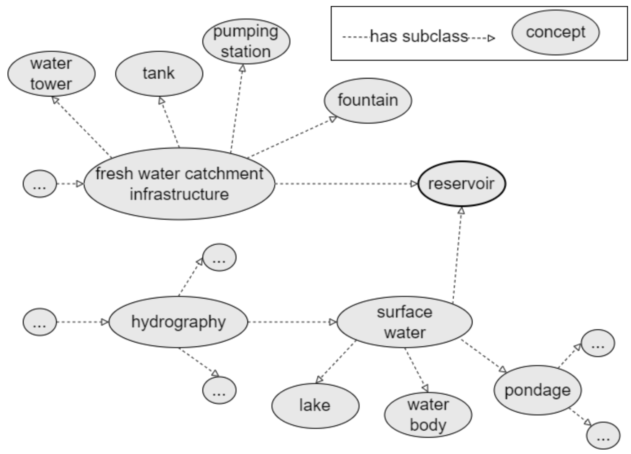

Figure 6. Illustration of a multiple heritage.

Indeed, a water body can have natural or anthropogenic origins. Nevertheless, victims can locate themselves in relation to a reservoir without knowing or mentioning its origin.

4.2 OOR application

This section describes several applications of OOR which correspond to the *Test phase* of the OOR construction. Most of the tests were carried out as part of the construction of a geographical database of landmarks for mountain rescue. The testing phase was iterative: each application allowed us to improve the ontology and define the version V1.1.

4.2.1 *Semantic mapping*

As described in Section 1.3.7, four data sources are identified as relevant for mountain rescue. The theme Point of Interest from BDTOPO database produced by IGN is also considered for building a landmarks geographic database. The semantic mapping is the alignment between landmark types of all data sources.

OOR is used as a pivot, as each feature type of each data source is assigned to a corresponding URI from OOR. The schema instantiation is done manually and collaboratively. In total 543 types are aligned with OOR (21 for C2C, 230 for OSM, 278 for BDTOPO among 295, seven among 28 for PA, seven among eight for refuges). Some landmark types are not aligned with the OOR classes, as they do not have a correspondent in OOR. The analysis of these types allows us to conclude that they are out of scope for our application and, consequently, they have not been added to OOR. This application has improved the organisation of classes in OOR.

For more information about this application, the reader can see (Van Damme and Olteanu-Raimond, 2022). The alignment files are available in Choucas ZENODO community (https://doi.org/10.5281/zenodo.6481338).

4.2.2 *Data instantiation*

This application concerns the instantiation of the ontology with landmark data coming from the five sources of data described above. Instances denote the features represented by a concept (e.g., the concept “lake” versus the instance “The White Lake” represented in BDTOPO by a point or a surface). OOR was used to build a geographic database where data are stored in a property graph database, a first version was described in Gendner et al., (2021). The instantiation consists of defining association links between each landmark in a data source and an OOR class by using "is-a-kind-of" relationships. Both *prefLabel* and *altLabel* are used to instantiate OOR. The final database account considers semantic heterogeneities between data sources by using OOR as a common vocabulary. The total number of instances is: 2289 for C2C, 1906 for PA, 659 for Refuges.info, 41454 for OSM and 17769 for BDTOPO. This application improved OOR in the following ways: *altLabel* are transformed into classes and become *prefLab*, the organisation of classes is improved, some typos are corrected, and the definitions are completed.

4.2.3 *Data matching*

Data matching consists of defining homologous features with respect to a reference data source. Four data matchings are made, each data source being matched with BDTOPO, considered as a reference.

Among different criteria used for data matching, the semantic criterion consists of computing semantic distances between type features by using Wu-Palmer distance (Olteanu Raimond et al., 2015). To assess whether OOR improves semantic measures, we compared the semantic distances obtained using both OOR and GeOnto ontologies for each couple data matching. This analysis shows that OOR better measures features that are semantically different (values around 1) or similar (values around 0) and the semantic difference is better highlighted. For example, the semantic distance between summit and valley is equal to respectively 0.6 and .33 for OOR and GeOnto, which means that OOR can say that from a semantic point of view, summit and valley are different. The correlations between semantic distances are computed for all four data matching and correlation coefficients (*p-value*=0.05) equal to respectively 0.85 and 0.8 for BDTOPO-C2C and BDTOPO-PA data matching, and 0.6 for BDTOPO-refuges. This shows strong correlations, but not equal to 1 meaning the interest of OOR for improving similarity measures.

For more information about this application, see (Van Damme and Olteanu-Raimond, 2022). The matching results files are available in Choucas ZENODO community (https://doi.org/10.5281/zenodo.6518363).

5 Conclusion and future works

This research proposes an application of lightweight ontology (OOR) to describe landmarks in mountain area and improve the rescue operations in mountains.

OOR is defined in iterative and collaborative ways by adapting the SAMOD methodology and using different sources of knowledge. Both researchers and experts in mountain areas participated to the construction of the ontology. The ontology is already applied in different steps of semantic data integration with semantic mapping and data matching techniques and was instantiated by using two authoritative (BDTOPO and PA) and three VGI (OSM, C2C, and refuges.info) data sources.

The main difference from a purely SAMOD methodology is that the testing phase was not done after a first version, but after each step. This is due to our internal organisation and the time constraints that required for a first version within a given timeframe. The iterative and collaborative construction included several working sessions that brought researchers together and was very fruitful in our case. The choices of classes, *prefLabel*, *altLabel*, definitions and hierarchical organisation of *prefLabel* were discussed in detail during the working sessions and all participants agreed on the decisions. This way of working was appreciated by the participants and ensured motivation during the long process of ontology construction. For efficiency and organisation reasons, only one researcher participated in the working sessions with the experts.

From an application perspective, the first goal of the ontology is for use by rescue team that listens to the victim's description and interprets it in real time. Thanks to the rescuer, the description is clarified and expanded, by using the information from the ontology, data and location information contained in the victim's call. Indeed, OOR can be used by a rescuer to search types of landmarks. For example, if a victim gives the clue "I'm 1500 m from a hut", the rescue operator can search the huts in the area and display the instances to better understand the environment of the victim and ask for new clues. Knowing that the hut is an *altLabel* of cabana, the query is converted to a concept (class *cabana*) and the answer is computed as the set of instances whose types are more specific or equivalent to the concept of the query.

Future work is to add new relationships such as composition, confusion, etc. Confusion relationships can improve the selection of instances during an emergency call. For example, victims mention that they see a power line, but it is a ski lift. A confusion relationship between *cable car* and *electric line* will allow the rescuer to ask the victim to validate the type or to select instances representing both the ski lift and the power lines.

To ensure interoperability and enhance the use of OOR, we plan to define links between OOR and DBPedia, and Geonames by using "isEquivalent" relationship.

Finally, we plan to help the mountain rescue team integrate the ontology into their system called Gend'Loc and help them use it. This will ensure the validation of our ontology from the experts' point of view. User feedback will be considered and a new version will be published.

6 Acknowledgment

This work was supported by the ANR under grant agreement no. ANR-16-CE23-0018 (CHOUCAS research Project: Heterogeneous data integration and spatial reasoning for localizing victims in mountain area).

The authors would like to acknowledge the reviewers to improving the quality of this paper.

7 References

Abadie, N., John McGibbney, L., Atemezing, G., 2019. Ontologie des éléments du territoire et de ses infrastructures de l'IGN. Revision: Version 1.1 - 201902-12.

Abdelghany, A., Darwish, N., Hefni, H., 2019. An Agile Methodology for Ontology Development. *International Journal of Intelligent Engineering & Systems* 12, 170–181. https://doi.org/10.22266/ijies2019.0430.17

Baker, T., Bechhofer, S., Isaac, A., Miles, A., Schreiber, G., Summers, E., 2013. Key choices in the design of Simple Knowledge Organization System (SKOS). *Journal of Web Semantics* 20, 35–49. https://doi.org/10.1016/j.websem.2013.05.001

Bunel, M., Duchêne, C., Olteanu-Raimond, A.-M., Villanova-Oliver, M., Bonhoure, G., Jouan, T., 2019. An approach for identifying and analysing reference features and spatial relations used in mountain emergency calls. *Proceedings of the ICA* 2, 1–8. https://doi.org/10.5194/ica-proc-2-12-2019

Davies, J., 2010. Lightweight Ontologies, in: Poli, R., Healy, M., Kameas, A. (Eds.), *Theory and Applications of Ontology: Computer Applications*. Springer Netherlands, Dordrecht, pp. 197–229. https://doi.org/10.1007/978-90-481-8847-5_9

ElHassouni, J., Qadi, A.E., 2022. Ontology Engineering Methodologies: State of the Art, in: Lazaar, M., Duvallet, C., Touhafi, A., Al Achhab, M. (Eds.), *Proceedings of the 5th International Conference on Big Data and Internet of Things*, Lecture Notes in Networks and Systems. Springer International Publishing, Cham, pp. 59–72. https://doi.org/10.1007/978-3-031-07969-6_5

Gaio, M., Moncla, L., 2019. Geoparsing and geocoding places in a dynamic space context: The case of hiking descriptions, in: Aurnague, M., Stosic, D. (Eds.), Human Cognitive Processing. John Benjamins Publishing Company, Amsterdam, pp. 354–386. https://doi.org/10.1075/hcp.66.10gai

Gendner, V., Van Damme, M.-D., Olteanu-Raimond, A.M., 2021. Modelling and building of a graph database of multi-source landmarks to help emergency mountain rescuers. Abstr. Int. Cartogr. Assoc. 3, 1–2. https://doi.org/10.5194/ica-abs-3-90-2021

Gomez-Perez, A., 1999. Ontological Engineering: A state of the Art. British Computer Society, Londres, Inglaterra, pp. 33–43.

Hill, L., Zheng, Q., 1999. Indirect Geospatial Referencing through Place Names in the Digital Library: Alexandria Digital Library Experience with Developing and Implementing Gazetteers. Presented at the *Proceedings of the 62nd Annual Meeting of the American Society for Information Science*, Washington, DC, USA, pp. 57–69.

Hill, L.., 2006. Georeferencing: The Geographic Associations of Information. The MIT Press. https://doi.org/10.7551/mitpress/3260.001.0001

Hogan, A., 2020. The Web of Data. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-030-51580-5

Jokar Arsanjani, J., Zipf, A., Mooney, P., Helbich, M. (Eds.), 2015. OpenStreetMap in GIScience, Lecture Notes in Geoinformation and Cartography. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-14280-7

Lynch, K., 1960. The Image of the City. MIT Press.

Moncla, L., Renteria-Agualimpia, W., Nogueras-Iso, J., Gaio, M., 2014. Geocoding for texts with fine-grain toponyms: an experiment on a geoparsed hiking descriptions corpus, in: ACM (Ed.), ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL 2014). Dallas, Texas, United States. https://doi.org/10.1145/2666310.2666386

Musen, M., Tu, S., Eriksson, H., Gennari, J., Puerta, A., 1993. Protege-ii: An environment for reusable problemsolving methods and domain ontologies. Presented at the International Joint Conference on Artificial Intelligence (IJCAI), Savoie, France.

Mustière, S., Abadie, N., Aussenac-Gilles, N., Bessagnet, M.-N., Kamiel, M., Kergosien, E., Reynaud, C., Safar, B., Sallaberry, C., 2011. Analyses linguistiques et techniques d'alignement pour créer et enrichir une ontologie topographique. Revue internationale de géomatique 21, 155–179. https://doi.org/10.3166/rig.21.155-179

Olteanu Raimond, A.-M., Mustiere, S., Ruas, A., 2015. Knowledge formalization for vector data matching using Belief Theory. Journal of Spatial Information Science. https://doi.org/10.5311/JOSIS.2015.10.194

Olteanu-Raimond, A.-M., Favre, O., Gaio, M., Van Damme, Duchêne, C., Bunel, M., Viry, M., Halilali, M., Villanova-Oliver, M., Gouardères, E., Davoine, P.-A., 2022. Séminaire de clôture du projet ANR CHOUCAS - présentations.

Peroni, S., 2016. A Simplified Agile Methodology for Ontology Development. Presented at the OWL: Experiences and Directions – Reasoner Evaluation,Bologna, Italy.

Sinha, G., Arundel, S.T., Hahmann, T., Usery, E.L., Stewart, K., Mark, D.M., 2018. The Landform Reference Ontology (LFRO): A Foundation for Exploring Linguistic and Geospatial Conceptualization of Landforms (Short Paper). Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany. https://doi.org/10.4230/lipics.giscience.2018.59

Sure, Y., Staab, S., Studer, R., 2009. Ontology Engineering Methodology, in: Staab, S., Studer, R. (Eds.), Handbook on Ontologies. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 135–152. https://doi.org/10.1007/978-3-540-92673-3_6

Van Damme, M.-D., Olteanu-Raimond, A.-M., 2022. A method to produce metadata describing and assessing the quality of spatial landmark datasets in mountain area. AGILE GIScience Ser. 3, 1–11. https://doi.org/10.5194/agile-giss-3-17-2022

Viry, M., 2022. GASPAR —Geoviz Application for Searching and Rescuing People. mappemonde. https://doi.org/10.4000/mappemonde.8063

Zhou, Sha, Winter, S., Vasardani, M., Zhou, Shunping, 2017. Place descriptions by landmarks. Journal of Spatial Science 62, 47–67. https://doi.org/10.1080/14498596.2016.1196623