

# Article A Semantic Approach for Quality Assurance and Assessment of Volunteered Geographic Information

Gloria Bordogna 回



**Abstract:** The paper analyses the characteristics of Volunteer Geographic Information (VGI) and the need to assure and assess its quality for a possible use and re-use. Ontologies and soft ontologies are presented as means to support quality assurance and assessment of VGI by highlighting their limitations. A proposal of a possibilistic approach using fuzzy ontology is finally illustrated that allows to model both imprecision and vagueness of domain knowledge and epistemic uncertainty affecting observations. A case study example is illustrated.

**Keywords:** volunteered geographic information; semantic quality assurance and assessment; ontologies; soft ontologies; epistemic uncertainty; vagueness/fuzziness

# 1. Introduction

Volunteered Geographic Information and neo-geography are nowadays common synonymous terms used to indicate the practice of volunteers engaged in creating and sharing information on the Web about observations relative to target objects or events in various forms and with different nature, such as textual notes, pictures, sound recording, videos, measurements of properties, by associating a geographic reference with them, i.e., a geographic footprints on the Earth surface, indicating where the object or event was observed [1,2].

Generally, such practice was first adopted within "Citizen Science" (CS) collaborative projects, carrying out some of the tasks by the involvement of a large number of volunteer citizens [3,4]. Thanks to the Internet, and to the diffusion of smart applications on mobile devices, a potential increase of the number of volunteers' from worldwide with timely contributions is nowadays possible. This makes CS a viable solution also for monitoring processes, thus constituting a new challenge not only for science but even for collaborative city management. Nevertheless, many users express criticism towards usability of VGI due to its questionable quality.

In an attempt to improve the quality of VGI, CS projects introduced either ex-ante, expost or hybrid methods to normalize the creation and selection of VGI based on controlled vocabularies, ontologies and database integrity constraints checks [5–7].

By applying such methods, while VGI is made compliant with the project's specification for which was created, it can hardly be re-used for other projects having different quality requirements: this is mainly due to the lack of metadata documenting both the confidence of the volunteer in his/her understanding of the problem and the contextual conditions of the observations, which may cause a degradation of VGI [6].

In facts, when creating VGI for a CS project, volunteers may be required to classify some objects of interest according to a taxonomy or ontology. This may be uneasy when volunteers do not have adequate knowledge of the problem, such as within the Space4Agri project, in which they were asked to classify the crops growing in agronomic parcels and their actual phenological stage by using a classic agronomic ontology [8]. In some cases, they could not assign a single crop and/or a single phenological stage to an observation; this uncertainty of volunteers was due to the imprecise and vague descriptions of shape



Citation: Bordogna, G. A Semantic Approach for Quality Assurance and Assessment of Volunteered Geographic Information. *Information* 2021, *12*, 492. https://doi.org/ 10.3390/info12120492

Academic Editor: Antonio Comi

Received: 11 October 2021 Accepted: 24 November 2021 Published: 25 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). characteristics of both crops and stages in the classic ontology, which observations did not perfectly match. In some cases, the inability was also due to the not optimal observation conditions, such as far point of view, which prevented them from distinguishing well the characteristics. This is epistemic uncertainty affecting the observation.

In this paper, the use of fuzzy ontologies is analyzed to allow implementing both ex-ante flexible quality assurance mechanisms and ex-post flexible fitness for use quality assessment. Flexibility is meant as the ability to model imprecision, vagueness and uncertainty. To this end, some limitations of current ontologies and soft ontologies are identified, and specifically the need to distinguish, represent and reason in presence of both epistemic uncertainty and imprecision/vagueness of domain knowledge.

In particular, possibilistic fuzzy databases [9,10] are suggested as a viable means to model epistemic uncertainty affecting data of VGI observations, so that the uncertainty expands the possible data values of VGI observations. Fuzzy ontologies are proposed to represent imprecise/vague domain knowledge, and thus as a suitable means to describe observations qualitatively [11]. Finally, fuzzy Description Logic (DL) is an appropriate framework to perform reasoning on a fuzzy ontology [12]. A case study example will be described to exemplify how these frameworks will be combined to define the proposal.

The next section introduces the problem of VGI quality assurance and assessment, and the notions of ontologies and fuzzy ontologies. Section 3 discusses the limitations of fuzzy ontologies and proposes a possibilistic method, drawn from fuzzy databases, that exploits fuzzy ontologies and fuzzy DL; the discussion section illustrates the application of the proposed method in a case study. Conclusions summarize the main contribution and limitation of the work.

#### 2. Materials and Methods

Many studies have been published to assess the quality of VGI based on either cross comparison with authoritative geospatial data and remote sensing data [13], or by manual moderation [14].

The quality assessment based on comparison to an authoritative dataset of the same area (that is considered as ground-truth) uses different measurements to represent the quality of VGI, such as the length of the common features and the difference between the attribute values [6,15].

VGI quality assessment has been mainly investigated by considering as quality indicators logical consistency, positional accuracy, temporal accuracy, thematic accuracy, and completeness [16], formally standardized by ISO 19113 for spatial data quality [17]. A comprehensive historic updated survey can be found in [18]. Some works expressed the overall VGI quality by combining and integrating several measures such as [18–21]. For example, the work in Ref. [18] proposes a comprehensive quality assessment framework for VGI linear features that integrates novel quality metrics with other commonly used geometric and topologic indicators.

Most works found that VGI quality is very heterogeneous, mostly affected by spatial incompleteness or sparseness, thus the importance of evaluating completeness [16,22], with greatest timeliness than official geospatial data in some localities [14]. For example, quality assessment of VGI within the OpenStreetMap (OSM) project, has been carried out in distinct localities and the results outline that urban areas can reach both higher positional accuracy and greater completeness than the authoritative datasets with respect to rural places [23]. Nevertheless, in general, it is difficult to draw a conclusion about the degree of adherence to a specific set of data quality standards or to assess usability of VGI in the case in which ground truth data are not available.

This is the reason why other approaches proposed to explore VGI quality assessment by evaluating intrinsic VGI characteristics, such as the number of edits, the number of the volunteers and volunteers' trustfulness [24] or a combination of them [16,25]. For example, ref. [26] found that the positional accuracy of OSM VGI increases with the number of volunteers contributing in the same area. Other approaches assessed VGI quality mainly by evaluating extrinsic VGI characteristics, such as the volunteers motivations, experience and recognition [27].

Finally, some approaches proposed to assess VGI quality by exploiting the consistency of VGI with its geographic context [28], by relying on the Tobler's first law of geography stating that "all things are related, but nearby things are more related than distant things" [29]. Very few works evaluated the semantic quality of VGI: for a rich survey on the main methods of VGI quality assessment see [6,15].

In [30] a "consumer centric" multi criteria decision making approach to VGI quality assessment was proposed based on the specification of flexible queries to a database storing VGI quality indicators, representing both intrinsic and extrinsic quality characteristics of VGI, in order to select VGI in decreasing order of its quality suitability, i.e., relevance to the flexible query. This approach enables the VGI quality assessment conceived as evaluating VGI "fitness for use", and thus is adequate to enable the re-use of VGI.

Nevertheless, the prerequisite for the applicability of this approach is the possibility to compute quality indicators of the single dimensions of VGI, which can denote both intrinsic and extrinsic characteristics. For this reason, in [5] we analyzed the possible causes of VGI quality degradation by identifying some relevant dimensions of VGI that may impact on quality assurance.

One first aspect that influences VGI quality degradation is due to the means by which VGI is created. Ex-ante approaches are generally provided with the objective of supporting volunteers with easy to access resources for a correct creation of information, and automatic mechanism for controlling data entry. These resources include templates with automatic error checking capabilities to enforce better data creation practices; controlled vocabularies; geographic gazetteers, and ontologies in the specific scientific domain of the project. A review of the policies adopted by citizen science projects for regulating, constraining, filtering and correcting the data entered by volunteers, so as to reduce, as much as possible, the uncertainty of their interpretation due to errors, ambiguities, incompleteness and inaccuracy can be found in [5,6,31].

Ex-ante semantic support for creating VGI has been addressed in two distinct ways: either by relying on domain ontologies [32,33] or by adopting the Semantic Web and Linked Data framework [34].

The use of gazetteers and ontologies is particularly diffused in CS projects that ask volunteers to identify objects of interest and to categorize them on the basis of the provided domain ontology [31]. These projects may tolerate inaccuracies about the objects geolocation while demanding certainty of the classifications of the observed objects. For example, in the *inaturalist* project (https://www.inaturalist.org) (accessed on 9 September 2021) the tolerance of the geo-localization of the observed object can be defined by specifying an uncertainty radius around a point identified by geographic coordinates. On the other side, in order to achieve a high quality in the classification of the observed objects, an hybrid approach for quality improvement is adopted. Ontologies enriched with visual information in the form of pictures and drawings of the species are provided to help volunteers create correct classifications of observations, for which they have to upload a photo that is then submitted to the possible revision by the community.

In addition, CS projects in which the positional accuracy is important, such as Open-StreetMap (OSM), started exploring the possibility of creating a semantic schema of the concepts by organizing them into ontologies [15]. For example, LinkedGeoData [34] enriches OSM data with other semantic datasets such as DBPedia or GeoNames and publishes them in a semantic form, and OSMonto [35], although designed for a specific application, organizes OSM concepts within an ontology with a rich number of relationships.

In order to improve VGI quality, once VGI has been already created and stored in the project database, ex-post approaches have been defined which clean errors, normalize data and discard VGI not satisfying integrity constraints, by the aid of external knowledge provided as ground truth data, authoritative information from administrative and commercial datasets, such as land cover, land use, Digital Elevation Models, or gazetteers and ontologies in the specific scientific domain [36,37]. Ex-post approaches are generally used to achieve semantic interoperability of VGI from different sources [38]. In facts, semantics is fundamental to understanding the meaning of data. The Semantic Web has developed methods and techniques to support semantic interoperability among applications [39,40].

Nevertheless, VGI creation by using classic ontologies cannot completely avoid VGI quality degradation. In facts, one import aspect of VGI creation is data contextualization, that means explaining the context in which data have been created [6]. Data contextualization implies explaining the conditions in which an observation has been taken, for instance, far point of view, presence of fog, low sun light, which may affect the quality of data with epistemic uncertainty. Another aspect of contextualization is explaining the meaning of the project and of the important concepts. Within the Space4Agri project for the improvement of the agronomic sector in northern Italy, we experienced the difficulties of volunteers to identify the crops and their actual crop growth stage when observing agronomic parcels by using a classic ontology [8]: they found difficulties, since crops may have different appearance from place to place within the same parcel, and this appearance may not correspond to a unique growth stage as described in the classic ontology of reference. This motivated our proposal of defining a fuzzy ontology in order to represent the fuzziness of crops growth stages to support the ex-ante creation of VGI [32].

On the other side, VGI consumers need to select VGI based on their specific quality needs. By taking ideas from [36], in which an ontology driven VGI filtering was proposed for disaster management, in the following sections we analyze the appropriateness of using current ontologies and soft ontologies to support both VGI quality assurance and assessment.

## **Ontologies and Soft Ontologies**

When data are distributed or shared among various data producers and consumers, as it happens for geo-referenced data created by the crowd and shared on the Web, it is necessary to have either a shared semantics through a common vocabulary to specify entities, concepts, and relationships (a domain ontology), or mapping rules between different local vocabularies (ontology alignment). To this aim, the role of ontologies is growing in significance with the growth of the social Web and the diversity of user generated contents.

Ontology is a formal and explicit specification of a shared conceptualization [41]. It constitutes a common context model and plays a key role in inferring semantics of user generated contents. "Conceptualization" refers to an abstract model of some part of the world that identifies the relevant concepts and relations. "Explicit" means that the type of concepts, the relationships between concepts, and the constraints on their usage, are explicitly defined. "Formal" refers to the fact that ontology should be defined in a machine-readable form. Finally, "shared" means that the ontology should reflect the understanding of a community. Specifically, in the Semantic Web context, ontologies provide a knowledge representation shared between agents, in order to understand each other in communication. Ontologies subsumes taxonomies, since, besides the hierarchical organization of concepts through inheritance relationship ("IsA"), they represent other kinds of relationships, such as "part of". There are several formal languages that can be used for expressing ontologies. The Web Ontology Language OWL and its extensions (OWL-2 a more expressive version of OWL) is the W3C recommendation for coding ontologies for the Web. The basis for OWL is Description Logics (DL) for representing structured knowledge and decidable reasoning based on a well-defined theoretic semantics [42]. OWL allows for intensional definition of classes (concepts in the formalism of DL), and properties, that is, binary relationships between classes (named roles in DL). Classes and properties constitute the Terminological Box (TBox) of a knowledge base of concepts. OWL also allows for the definition of instances (individuals), the extensional definition of classes and properties, which constitute the Assertional Box (ABox), i.e., the set of facts. Finally, OWL allows for role-centric definition of modeling features (the RBox). Specifically, the

recommendation defines three levels of extensions of OWL characterized by different trade-off of expressiveness and computational efficiency [43]. OWL-DLs can be viewed as machine-processable representation of DL, suitable for interoperability and scalability of systems, that promotes reuse of data and reasoning over the Web. The power of OWL-DL ontologies resides in the sound inference, i.e., reasoning, that is applied on the basis of the axioms in the knowledge base. The two basic inference on concepts expressions in DL are subsumption, i.e.,  $C \sqsubseteq E$ , the problem of checking whether the concept E is more general (includes) the concept C, and satisfiability, i.e., the problem of checking whether a concept is not the empty concept. For example, a deductive reasoning task uses the knowledge about a domain in the TBox in order to produce a classification of instances described in terms of observed properties in an ABox. Notice that OWL DL allows for representing the world in terms of binary membership of instances to crisp concepts (sets) and binary relationships among instances. In particular, a given statement can be either true or false. As a consequence, OWL is suitable to modelling domains in which concepts or relationships have a precise definition, and thus are not imprecise and vague by their very nature [11]. Nevertheless, there are concepts and relationships of the real world that are intrinsically imprecise and fuzzy, due to their gradual nature [44,45]. For instance, in the sentence "Most climbing trails are narrow", the concepts "most" and "narrow" involve some fuzziness for which a crisp definition is impossible or does not make sense. What is the width of a trail that makes it "narrow"? This is a matter of degrees depending on a subjective or collective interpretation and, certainly, there is not a crisp transition between a "narrow" and a "wide" trail that may be agreed upon by all. By the term "most" it is meant that there are exceptions, for instance, there are climbing trails which are "wide". Furthermore there may be cases in which one needs to define a fuzzy concept hierarchy, i.e., a fuzzy taxonomy, in which a class is a specialization to a degree of several super classes such as "In Italy churches besides being (1) places of worship are often (0.8) historical buildings". Finally, it may be also necessary to define fuzzy relationships between concepts such as "bell towers are very close to churches".

Probabilistic ontologies have been defined in order to be able to represent weighted relationships in an ontology: for example, "*A is a small subclass of B*", "*A and B are largely overlapped with each other*". BayesOWL [46] extends OWL based on Bayesian networks. It provides a set of rules and procedures for automatic translation of an OWL ontology into a Bayesian network, that is a directed acyclic graph. Algorithms are defined that incorporate probabilistic information about both classes and inter-class relationships into the conditional probability tables of the Bayesian network. One main problem of the probabilistic approach is that it heavily relies on the availability of probabilistic information for both translation of ontology to Bayesian network and ontology mapping, which is hardly available from domain experts.

An alternative to probabilistic ontologies are fuzzy ontologies, defined with several purposes, depending on the kind of imperfection they need to represent and manage in the application [47]. Samani and Shamsfard [48] distinguish the definitions of fuzzy ontologies into three main categories: those including just fuzzy relationships [49–52] or fuzzy attributes [53]; those including both fuzzy attributes and fuzzy relationships [54]; and those including fuzzy attributes, fuzzy relationships and fuzzy axioms defined by fuzzy IF-THEN rules admitting a degree of truth [12,55].

Much research has fuzzified the existing Description Logics (DL) and have defined fuzzy DL reasoners (see [12] for a review). The most up-to-date and complete fuzzyDL ontology reasoner has been proposed in [11,12,56] which supports an expressive language defined as a fuzzy extension of OWL 2. A useful plugin of the formal ontology software Protégé, named Fuzzy OWL, has been developed allowing to define fuzzy data types, fuzzy modifiers, fuzzy concepts and fuzzy axioms having degrees of truth in [0, 1] denoting a lower bound for the degrees of satisfaction of the axiom. This plugin is associated to a fuzzy reasoner, named FuzzyDL, which allows to compute degrees of truth of fuzzy propositions by choosing one among three fuzzy logics: Łukasiewicz logic, Gödel logic, and Zadeh logic.

In this proposal, the degree of truth of an axiom does not represent epistemic uncertainty modeling defects and limitations of the observations and observation means (for instance far point of view, variability of observed instances) but it expresses a degree of preferences or importance of an axiom.

In order to model epistemic uncertainty, soft ontologies have been defined within the possibilistic framework [57,58], that deals with certainty and possibility degrees of truth. Possibilistic ontologies aim at modeling the epistemic uncertainty characterizing experts' subjective knowledge and evaluation of the certainty of this knowledge. To this end, several possibilistic DL reasoners have been defined [57,58] which allows representation and reasoning on uncertain statements such as "*It is possible that this trail is a climbing trail*". To this end each concept, relationships and axiom is associated to a real value  $\alpha$  in (0, 1] representing its certainty level. Nevertheless, no soft ontology has been proposed for reasoning in presence of both fuzziness and epistemic uncertainty in the same framework.

#### 3. Results

When creating VGI, a possible source of imperfection occurs when the volunteer is not completely sure about his/her observation, which is the case of epistemic uncertainty. This may happen because he/she does not have adequate knowledge of the problem or because of deficiencies in the means of observation. This may also happen when the domain knowledge is encoded into a precise ontology. For example, consider the description of a trail *"as a path with a width greater than 1 mt"*. For an observer, it might be impossible to measure the width precisely. In such situations it might be questionable to state the precise data axioms about the trail width, while one could state an imprecise fact, for example that the trail width is small. Another possibility would be to associate a degree b in [0, 1] to a precise statement *"the path width's is 1 mt"* expressing the confidence, belief, on its truth. Finally, there are more complex situations that may involve both imprecise, fuzzy and incomplete knowledge on one side, when encoding domain knowledge, and epistemic uncertainty and imprecision of volunteer on the other side, when creating VGI instances.

# A Pragmatic Approach to Model Both Epistemic Uncertainty and Imprecision/Vagueness of VGI in a Fuzzy Ontology

In the context of possibilistic fuzzy databases, it is important to distinguish imprecision/vagueness from epistemic uncertainty [59]. As stated previously, there is often a misunderstanding on the modeling of uncertainty and imprecision or fuzziness [56]. Under uncertainty theory statements are either completely true or completely false but we do not know if they are true of false, so we define a probability or possibility distribution over the worlds. For example, the statement "x is a trail" is a crisp one: x can be either a trail or not, it cannot be partially a trail, the degree that we can associate to this statement is relative to our knowledge on the truth about x being a trail, which may depend on some deficiency of the observation. Conversely, imprecision/fuzziness of a proposition are modelled by fuzzy concepts represented by fuzzy sets defined on the domain of possible worlds, so that the more values of the domain with not-null membership degree to the fuzzy set, the more the concept is imprecise/fuzzy [45]. We can model observations affected by some deficiency by alternative statements: for example, by observing a flower from a far point of view one could either specify the uncertainty on the truth of a precise predicate such as: a) "I am only 0.2 certain that x has 5 Petals", or both an uncertain and fuzzy predicate such as: b) "I am 0.7 certain that x has 4 or 5 Petals", or equivalently a certain and fuzzy predicate such as: c) "I am sure that x has less than 10 Petals". In all these statements we can notice that the certainty degrees (0.2, 0.7, and 1 in statements (a), (b) and (c), respectively) are increasing and are directly proportional to the amount of imprecision of the property values (5 petals is precise, 4–5 petals is imprecise, less than 10 petals is highly imprecise). One can guess that the specificity [60] of the possibility distributions defining the property values in the three cases decreases with the increase of the certainty, where the specificity of a possibility distribution measures the degree to which the distribution allows one and only one element as its manifestation [60].

In fuzzy databases the occurrence of both uncertainty and imprecise/fuzzy values V has been modeled by combining the uncertainty u with the membership function  $\mu_V$  ( $\mu_V$ : X $\rightarrow$ [0, 1]) of the fuzzy value, so as to derive a modified membership function:

 $\mu_{V'}(x)$ = i(u,  $\mu_{V}(x)$ ), where i is an implication function the effect of which is to expand the imprecision/fuzziness of the fuzzy value V based on the uncertainty degree u by increasing its unspecificity [9,10]. Within a possibilistic framework the modification is defined as follows: i(u,  $\mu_{V}(x)$ ) = max (u,  $\mu_{V}(x)$ ) [61]. This way, in presence of an uncertainty u, all values of the domain of x are possible values of V at least to a degree u.

Level-2 fuzzy set have been defined in order to allow the explicit representation of both fuzziness/imprecision and epistemic uncertainty represented by belief of truth [59]. Given a fuzzy set A defined on a universal set D with  $\mu_A: D \rightarrow [0, 1]$ , given  $\wp(D)$  the set of all fuzzy sets defined on D, a level-2 fuzzy set  $\tilde{A}: \wp(D) \rightarrow [0, 1]$  is defined by:

$$\bar{A} = \{(A; \mu_{\bar{A}}(A)) \mid \forall A \in \wp(D) : \mu_{\bar{A}}(A) > 0\} \text{ and } A = \{(d; \mu_A(d)) \mid \forall d \in D : \mu_A(d) > 0\}$$
(1)

The membership degree  $\mu_{\bar{A}}$  (A) of the level-2 fuzzy set is called 'outer-layer' membership grade, and is interpreted as degree of belief about the truth of the proposition, whereas the membership degrees  $\mu_A(d)$  of the elements  $d \in D$  are called 'inner-layer' membership degrees, and are used for modelling imprecision/fuzziness of concepts. Thus, by means of level-2 fuzzy sets, we can represent explicitly the belief on the truth of an imprecise/fuzzy predicate.

An uncertain and fuzzy proposition denoted by (Is.A(x); b) can be defined by a level-2 fuzzy set with D being a concrete domain, and A being a fuzzy set on D:

$$(Is.A(x); b) = \{(Is, b = \mu_{Is}(A), ) \mid A = \{(d, \mu_A(d)) \mid d \in D\}\}$$
(2)

 $\mu_{Is}(A) = b$  defines the belief degree that A is the actual (imprecise/fuzzy) value of a property p of instance x. This fuzzy and uncertain proposition can be transformed into a certain and less specific proposition (Is.A<sub>b</sub>(x); 1) as follows:

$$(Is.A_b(x); 1) = \{(Is, 1) \mid A_b = \{(d, \mu_{Ab}(d)) = max(1 - b, \mu_A(d)) \; \forall d \in D\}\}$$
(3)

which can be rewritten in a simpler form as a normal fuzzy set:

$$A_{b} = \{(d, \mu_{Ab}(d)) \mid \mu_{Ab}(d) = \max(1 - b, \mu_{A}(d)) \forall d \in D\}$$

It can be seen that the degree of belief b is used to generate a possibility distribution  $A_b$  with greater unspecificity than A:  $\mu_{Ab}(d) \ge \mu_A(d) \ \forall d \in D$ , i.e.,  $A \subseteq A_b$ .

In case of full belief,  $\mu_{Is}(A) = 1$ , nothing is changed, and the level-2 fuzzy set notation can be replaced by normal fuzzy set notation; in the case of null belief,  $\mu_{Is}(A) = 0$  any value d is fully possible, and we have maximum unspecificity.

As pointed out in [9,10], although these proposals can be useful in many applications, unfortunately they are inappropriate when reasoning in possibilistic fuzzy databases. In fact, it is reasonable to think that a small uncertainty on a value means that one cannot exclude as possible some near values, but still one can exclude as possible the far values, i.e., very dissimilar values, from V.

To model this desired behavior, a pragmatic approach is to constrain the modification of a possibility distribution by relaxing it within an envelope so as to limit the values of the domain that are possible, i.e., with possibility degree above zero.

Following this idea, we propose to constrain Formula (3) within an envelope of the support of  $\mu_A$  defined depending on the value of  $b = \mu_{Is}(A)$  as follows (see an example of modification in Figure 1):

Is.<sub>Ab</sub>(x); 1) = {(Is, 1) | 
$$A_b = {(d, \mu_{Ab}(d) = max(1 - b, \mu_A(d)) \forall d \in env(A_b, \alpha) and \mu_{Ab}(d) = 0 otherwise }$$
 (4)

where  $env(A_b, \alpha) = [d_1, d_2] \subseteq D$  with  $\alpha \ge 0$  and

d1= Argmind 
$$\in$$
 D ( $\mu$ A(d) > 0) - (1 - b)\* $\alpha$ , d2 = Argmaxd  $\in$  D( $\mu$ A(d) > 0) + (1 - b)\* $\alpha$  (5)



**Figure 1.** Uncertain and fuzzy proposition (Is.width.medium; 0.4) in orange and its semantically equivalent certain and fuzzy proposition (Is.width.medium0.4; 1) in yellow, obtained by applying the simple transformation defined in Formula (4) with  $\alpha = 0$ .

 $\alpha \ge 0$  models the extent of maximum tolerable unspecificity. In the cases in which b = 1, full belief, nothing is changed and (4) reduces to (3). In the case of null belief, b = 0, all values in the envelope become fully possible:

$$\mu_{A}(d) = 1 \quad \forall d \in [\operatorname{Argmin}_{d \in D} (\mu_{A}(d) > 0) - \alpha, \operatorname{Argmax}_{d \in D} (\mu_{A}(d) > 0) + \alpha]$$
(6)

Notice that, when  $\alpha$  tends towards infinite Formula (4) reduces to Formula (3). Conversely, when  $\alpha = 0$ ,  $[d_1, d_2] \equiv [\operatorname{Argmin}_{d \in D} (\mu_A(d) > 0))$ ,  $\operatorname{Argmax}_{d \in D} (\mu_A(d) > 0)]$ , thus only values in the support of A are rewarded with an increase of possibility.

Formula (4) can be applied also in the case in which A is defined on a discrete ordinal scale. In this case the relaxation is applied to an envelope defined by a selected subset  $D_S$  of D ( $D_S \subseteq D$ ).  $D_S$  can be defined by considering a similarity between the domain values.

As discussed above, given that we can transform any uncertain and fuzzy predicate into a certain and fuzzy predicate with greater unspecificity, we can then infer new knowledge by applying fuzzy reasoning based on the fuzzy ontology framework and fuzzy DL without epistemic uncertainty [12]. This means that we can explicit all implicit knowledge by computing the closure of the set K of fuzzy predicates (which represent both the ground facts and knowledge) that in the worst case has a complexity equal to  $O(|K|^2)$ , with |\*| indicating the set cardinality [11]. Then, any fuzzy query can be answered as in a fuzzy database by returning VGI items ranked in decreasing order of their possibility to satisfy the query [9,11,59].

# 4. Discussion

The proposed approach can thus reuse the framework of fuzzy ontologies and fuzzy DL [12] by modeling epistemic uncertainty on fuzzy facts. In our proposed approach, domain knowledge is represented by a belief base, that is a fuzzy knowledge base or fuzzy ontology in which all fuzzy concepts and fuzzy relationships between concepts relevant in the domain are defined, while a volunteer stating an observation "x is A", where A is a concept defined in the fuzzy ontology, is an agent with an epistemic state, who has to answer the question "Am I sure that "x is A" is true on the basis of the belief base in which A is defined?.

While the fuzzy ontology uses normal fuzzy sets to represent fuzzy concepts [45], fuzzy and uncertain assertion axioms, i.e., uncertain and fuzzy facts, are represented by level-2 fuzzy sets [59]: the outer-layer membership degrees are the degrees of belief on

the truth of the fuzzy axioms (necessity), while the inner-layer membership degrees are the possibility degrees by which the values of a concrete domain may be the values of the property, which are defined in the fuzzy ontology. Level-2 fuzzy sets are thus stored in a database to represent the fuzzy values of observed properties of VGI items with their associated belief degree.

In order to apply approximate reasoning, such as in a fuzzy ontology framework based on fuzzy DL [12], first we need to transform uncertain observations, into certain and fuzzy observations with greater unspecificity by applying the transformations method described in the previous section through application of Formula (4).

This approach can serve an ex-post quality assurance approach modeling "fitness-foruse" [30]. Filtering is activated by a user who specifies the maximum tolerable levels of uncertainty on the VGI items of interest.

In the following subsection an illustrative case study is discussed to exemplify the proposal.

#### Case Study Example

Let us consider a collaborative project that collects VGI items on the characteristics of paths in a region for sharing useful and rich information to hikers, climbers, cyclists, and tourists in general. Since it can be questionable for a volunteer to directly classify the path as either a climbing trail, a hiker's trail or a cycle trail, the idea is to collect observations of properties of the paths guided by a fuzzy ontology, such as the simple one illustrated in Figure 2. This ontology defines three types of trails by means of fuzzy concepts inclusion axioms: climber trails, hiker trails and cycle trails, and provides the definitions of the fuzzy concrete predicates, aka linguistic values of properties, through their membership functions as reported in Table 1.



**Figure 2.** The fuzzy ontology defines the concepts climbing trail, hikers trail, and cycle trail by means of fuzzy predicates on the properties' usage, slope, surface, width, and tortuosity which can assume fuzzy/imprecise linguistic values, whose meaning is specified by fuzzy sets. At the bottom of the figure the image shows a path and an uncertain and fuzzy observation.

Table 1. TBox of the fuzzy ontology and membership functions of the fuzzy concrete predicates depicted in Figure 2.

- 1. Is.climbingTrail  $\sqsubseteq$  Is.Trail  $\sqcap$  ( $\exists$  usage.(climbing  $\lor$  strolling  $\lor$  walking)  $\sqcap$  ( $\exists$  slope.*high*)  $\sqcap$  ( $\exists$  surface.(rocky  $\lor$  gravelly)  $\sqcap$  ( $\exists$  width.*small*)  $\sqcap$  ( $\exists$  tortuosity.*high*)
- 2. Is.hikersTrail  $\sqsubseteq$  Is.Trail  $\sqcap$  ( $\exists$  usage.(strolling  $\lor$  walking)  $\sqcap$  ( $\exists$  slope.*variable*)  $\sqcap$  ( $\exists$  surface.(stony  $\lor$  muddy)  $\sqcap$  ( $\exists$  width.*medium*)  $\sqcap$  ( $\exists$  tortuosity.*average*)
- 3. Is.cycleTrail  $\Box$  Is.Trail  $\sqcap$  ( $\exists$  usage.(biking $\lor$  riding  $\lor$  walking)  $\sqcap$  ( $\exists$  slope.*low*)  $\sqcap$  ( $\exists$  surface.(pebbly  $\lor$  gravelly  $\lor$  paved)  $\sqcap$  ( $\exists$  width.*large*)  $\sqcap$  ( $\exists$  tortuosity.*low*)

#### Membership functions definitions

$$width.small(x) = \begin{cases} e^{-\frac{(x-5)^2}{3}} & for \ x > 5\\ 1 & for \ x \le 5 \end{cases} \quad width.medium(x) = e^{-\frac{(x-10)^2}{3}} for \ \forall x \in R^+ \ width.large(x) = \begin{cases} e^{-\frac{(x-15)^2}{3}} & for \ x < 15\\ 1 & for \ x \ge 15 \end{cases}$$
$$slope.low(x) = \begin{cases} \frac{10-x}{10} & for \ x < 10\\ 0 & for \ x \ge 10 \end{cases} \quad slope.variable(x) = 1 \ \forall x \in [0,100] \qquad slope.high(x) = \begin{cases} 0 & for \ x \le 5\\ \frac{x-5}{5} & for \ 5 < x < 10\\ 1 & for \ x \ge 10 \end{cases}$$
$$tortuosity.low(x) = \begin{cases} e^{-\frac{x^2}{100}} for \ x > 0 \ tortuosity.average(x) = e^{-\frac{(x-05)^2}{005}} \ \forall x \in R^+ \ tortuosity.high(x) = \begin{cases} e^{-\frac{(x-1)^2}{3}} for \ x < 1 \end{cases}$$

In the literature [62], the manual elicitation of the membership functions from domain experts can be performed by applying one of the six common methods such as polling (do you agree that trail 01 is "muddy"? (Yes/No)); direct and reverse rating (How much trail 01 is muddy?/ which trail is muddy to degree 0.6?); interval estimation (give an interval in which you think the wideness of trail 01 lies); pairwise comparison, and direct membership function exemplification, which can use computer graphics to give exemplar membership functions to be modified by the experts. To achieve agreed definitions, averaging or aggregation of the responses from several experts' assessments is performed. Finally, semi-automatic approaches can exploit clustering [63] of a set of data to identify groups whose property values are characterized by both an ideal value and a variability, which are submitted to the experts' revision and validation.

Each VGI item created by a volunteer is georeferenced, with metadata of the author, and represented by a row in a relation such as the one in Table 2, in which the properties that are reported are those used in the fuzzy ontology to define the three distinct types of trails, plus a picture and a textual description.

A VGI database collects all observations of paths: the values of the paths' properties observed by volunteers are associated with a belief degree  $b \in (0, 1]$  expressing their confidence on the fact that the reported value of the property is true. For example, the item 01 in Table 2, is believed to degree 0.7 that one can *ride* or to degree 0.5 that one can *walk* on it, its slope is *variable* with belief 0.9, its surface is muddy with full belief 1, its width is *medium* and its tortuosity is low both with belief 0.4.

The representation of each uncertain and fuzzy fact, i.e., a level-2 fuzzy set, is interpreted by the reasoner so as to modify the membership function of the concrete fuzzy predicate, defined in the fuzzy ontology, based on the degrees of belief b by applying Formula (4). This step allows to generate certain fuzzy facts, aka an ABox of fuzzy assertion actions, so that given the fuzzy ontology TBox, it is possible to adopt a fuzzy *SHOIN(D)* reasoning framework [56] to associate each VGI items to each type of trail with a satisfiability degree as reported in Table 3.

Let us follow a simple numeric example. In the case of the uncertain properties, <Is.width. *medium*(01); 0.4> with the membership function defined in Table 1, by applying the transformation in Formula (4) we obtain the yellow membership function in Figure 1.

VGI Items	Usage	Slope	Surface	Width	Tortuosity	Picture	Textual Description
01	riding 0.7 walking 0.5	variable 0.9	muddy 1	medium 0.4	low 0.4		Muddy public footpath.
02	walking 0.9	high 0.8	paved 0.3	small 0.8	average 0.2		Mountain path, quite steep, that runs along an overhang
03	biking 1, riding 0.8	low 0.8	gravelly 1	large 0.8	low 0.7	1	Scenic country route disclosed to the sun
04	climbing 0.8, walking 1	low 0.3	rocky 1, stony 1	medium 0.5	high 0.4		Spectacular path, partly dug directly into the limestone rocks
05	walking 1	variable 0.9	stony 0.5, pebbly 0.9	medium 0.5	average 0.7	A COL	World Heritage Site by Unesco, "Sentiero Azzurro" Cinque Terre, Italy

Table 2. Relation representing VGI items in a database of paths descriptions.

Table 3. VGI items of Table 1 as uncertain instances of the three concepts in the fuzzy ontology.

VGI Items	Climbing Trail	Hikers Trail	Cycle Trail
01	0	1	0
02	0.7	0.2	0
03	0	0	1
04	0.5	0.6	0
05	0.3	1	0.5

Indeed, <Is.width. *medium*(01); 0.4> represents an uncertain and vague assertion. The user states by the degree 0.4 the level of belief of the assertion, while the semantics of *medium* for the property width is defined in the fuzzy ontology as a concrete fuzzy predicate with the membership function  $\mu_{width \cdot medium}$  on the basic domain of positive real values. Within possibility theory [61], this assertion is equivalent to <Is.width. medium(01); 1> having full belief and relaxed membership function  $\mu_{width \cdot medium} = \max(\mu_{width \cdot medium}, 0.6)$ . By this relaxation, all values of the basic domain can be possible values of width at least to degree 0.6, due to the uncertainty (1–0.4) in the original assertion. Similarly to possibilistic fuzzy databases, our proposal limits the relaxation of the membership function to an envelope as defined by Formula (4), so as to include the support of  $\mu_{width \cdot medium}$ . This way some values of the basis domain can be excluded as possible values proportionally to the degree of belief, a greater belief determines a greater number of impossible values. By assuming this interpretation, the reasoner translates uncertain and vague assertions to certain and vague assertions so as to use fuzzy DL for further reasoning.

The same transformation is applied to all uncertain values of observed properties of item 01 so that their membership functions are modified according to the respective degrees of belief. These modified membership functions constitute the fuzzy ABox to perform the

reasoning task over the fuzzy TBox in Table 1. The result is a fuzzy classification of the VGI items into the three types of trails, where the membership degree is interpreted as the degree of belief that the VGI item belongs to the type of trail and can be stored in the database so as to ease VGI filtering based on "fitness for use".

Coming back to the example, let us consider the needs of a bike sharing company that wants to collect descriptions and pictures of trails suitable for cycling with a maximum tolerable uncertainty equal to 0.5.

A quality filtering is applied that selects all "cycle trails" in the database whose belief degree is equal-above 0.5, thus retrieving VGI items 03 and 05 in decreasing order of their associated belief degrees 1 and 0.5, respectively.

#### 5. Conclusions

This work analyzed several aspects of imperfection affecting VGI. More closely, it discussed how VGI quality assurance can be modeled by means of an ontology and a fuzzy ontology to constrain VGI creation, the latter employed to cope with imprecise and fuzzy knowledge. Finally, a practical semantic approach that allows to model both epistemic uncertainty affecting volunteers' observations and imprecision fuzziness of knowledge was proposed. This approach adopts the possibilistic framework of fuzzy databases to generate fuzzy assertion axioms from the uncertain and fuzzy observations by means of a relaxation of the membership functions of the fuzzy concrete predicates. The adoption of this method offers several advantages with respect to previous approaches:

- it models both precise and uncertain creation of VGI, so coping with the limitations of the observation means and context;
- it supports unexperienced volunteers who are unable to interpret the meaning of some concepts in the ontology by allowing them to select imprecise and fuzzy concepts;
- when analyzing the created VGI stored in a database, users we can filter VGI items based on the maximum levels of uncertainty they can tolerate for their application needs.

The proposed method is then suitable for implementing a quality-based "fitness for use" assurance and assessment mechanism.

The approach is practical since it reuses the frameworks of fuzzy DL and fuzzy databases. While reusing a formal framework is an advantage, since it guarantees consistency of the reasoning task, it also inherits its limitations which are mainly questionable efficiency when the fuzzy ontology is complex.

Moreover, the proposal relies on the availability of domain knowledge represented by means of a fuzzy ontology, with membership functions defined to represent the semantic of concrete fuzzy predicates. This may be seeing as the main obstacle to the applicability of the approach, since generally domain knowledge is hardly represented by fuzzy ontologies. Generally, methods to construct ontologies and fuzzy ontologies still require manual intervention, that can be a burden task for domain experts. However, with the progress of machine learning, the research on automatic or semi-automatic ontology development is flourishing [64]. Thus, methods to automatically help experts to elicit their knowledge and to represent it by means of fuzzy ontologies are being defined, for example, to learn fuzzy data types by using clustering techniques [63].

Funding: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

# References

- 1. Goodchild, M.F. Citizens as Voluntary Sensors: Spatial Data Infrastructure in the World of Web 2.0. *Int. J. Spat. Data Infrastruct. Res.* **2001**, *2*, 24–32.
- Haklay, M. Citizen Science and Volunteered Geographic Information—Overview and typology of participation. In Volunteered Geographic Information, Public Participation, and Crowdsourced Production of Geographic Knowledge; Sui, D.Z., Elwood, S., Goodchild, M.F., Eds.; Springer: Berlin/Heidelberg, Germany, 2012.

- 3. Hand, E. Citizen science: People power. Nature 2010, 466, 685–687. [CrossRef]
- 4. Bowser, A.; Cooper, C.; de Sherbinin, A.; Wiggins, A.; Brenton, P.; Chuang, T.R.; Faustman, E.; Haklay, M.; Meloche, M. Still in Need of Norms: The State of the Data in Citizen Science. *Citiz. Sci. Theory Pract.* **2020**, *5*, 1–18. [CrossRef]
- Bordogna, G.; Carrara, P.; Criscuolo, L.; Pepe, M.; Rampini, A. On predicting and improving the quality of Volunteer Geographic Information projects. *Int. J. Digit. Earth* 2016, 9, 134–155. [CrossRef]
- Balázs, B.; Mooney, P.; Nováková, E.; Bastin, L.; Arsanjani, J.J. Data Quality in Citizen Science. In *The Science of Citizen Science*; Vohland, K., Land-Zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., Wagenknecht, K., Eds.; Springer Nature: Berlin/Heidelberg, Germany, 2021; pp. 139–156. [CrossRef]
- 7. Trojan, J.; Schade, S.; Lemmens, R.; Frantál, B. Citizen science as a new approach in Geography and beyond: Review and reflections. *Morav. Geogr. Rep.* 2019, 27, 254–264. [CrossRef]
- Bordogna, G.; Kliment, K.; Frigerio, L.; Stroppiana, D.; Brivio, P.A.; Crema, A.; Boschetti, M.; Sterlacchini, S. Spatial Data Infrastructure integrating multisource heterogeneous geospatial data and time series: A study case in agriculture. *ISPRS Int. J. Geo-Inf.* 2016, 5, 73. [CrossRef]
- Bordogna, G.; Pasi, G. Modeling Linguistic Qualifiers of Uncertainty in a Fuzzy Database. Int. J. Intell. Sys. 2000, 15, 995–1014. [CrossRef]
- 10. Gonzalez, A.; Marın, N.; Pons, O.; Vila, M.A. Fuzzy Certainty on Fuzzy Values. Control Cybern. 2009, 38, 311–339.
- 11. Straccia, U. Foundations of Fuzzy Logic and Semantic Web Languages. In *CRC Studies in Informatics Series*; Zhang, G.Q., Ed.; Chapman & Hall CRC Press Taylor and Francis Group: Boca Raton, FL, USA, 2013.
- 12. Bobillo, F.; Straccia, U. The fuzzy ontology reasoner fuzzyDL. Knowl.-Based Syst. 2016, 95, 12–34. [CrossRef]
- 13. Barrington-Leigh, C.; Millard-Ball, A. The world's user-generated road map is more than 80% complete. *PLoS ONE* 2017, 12, e0180698. [CrossRef]
- 14. Antoniou, V.; Skopeliti, A. Measures and Indicators of VGI quality: An overview. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 2015, *II-3/W5*, 345–351. [CrossRef]
- 15. Vandecasteele, A.; Devillers, R. Improving volunteered geographic data quality using semantic similarity measurements. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *1*, 143–148. [CrossRef]
- 16. Yeboah, G.; Porto de Albuquerque, J.; Troilo, R.; Tregonning, G.; Perera, S.; Ahmed, S.A.; Yusuf, R. Analysis of OpenStreetMap Data Quality at Different Stages of a Participatory Mapping Process: Evidence from Slums in Africa and Asia. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 265. [CrossRef]
- 17. ISO 19113. Geographic Information—Quality Principles; ISO: Geneva, Switzerland, 2002.
- 18. Wu, H.; Lin, A.; Clarke, K.C.; Shi, W.; Cardenas-Tristan, A.; Tu, Z. A comprehensive quality assessment framework for linear features from Volunteered Geographic Information. *Int. J. Geogr. Inf. Sci.* **2021**, *35*, 1826–1847. [CrossRef]
- 19. Arsanjani, J.J.; Barron, C.; Nakillah, M.; Helbich, M. Assessing the Quality of OSM Contributors together with their Contributions. In Proceedings of the 16th AGILE International Conference of Geographic Information Science, Leuven, Belgium, 14–17 May 2013.
- 20. Barron, C.; Neis, P.; Zipf, A. A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Trans. GIS* **2014**, *18*, 877–895. [CrossRef]
- 21. Forghani, M.; Delavar, M.R. A Quality Study of the OSM Dataset for Tehran. ISPRS Int. J. Geo-Inf. 2014, 3, 750–763. [CrossRef]
- 22. Chehreghan, A.; Abbaspour, R.A. An evaluation of data completeness of VGI through geometric similarity assessment. *Int. J. Image Data Fusion* **2018**, *9*, 319–337. [CrossRef]
- Arsanjani, J.J.; Mooney, P.; Zipf, A.; Schauss, A. Quality Assessment of the Contributed Land Use Information from OSM Versus Authoritative Datasets. In *OpenStreetMap in GIScience*; Arsanjani, J.J., Zipf, A., Mooney, P., Helbich, M., Eds.; Lecture Notes in Geoinformation and Cartography; Springer: Berlin/Heidelberg, Germany, 2015.
- 24. Fogliaroni, P.; D'Antonio, F.; Clementini, E. Data trustworthiness and user reputation as indicators of VGI quality. *Geo-Spat. Inf. Sci.* 2018, *21*, 213–233. [CrossRef]
- 25. Mooney, P.; Corcoran, P. Characteristics of heavily edited objects in OSM. Future Internet 2012, 4, 285–305. [CrossRef]
- Haklay, M.; Basiouka, S.; Antoniou, V.; Ather, A. How Many Volunteers does it Take to Map an Area Well? The Validity of Linus' Law to VGI. *Cartogr. J.* 2010, 47, 315–322. [CrossRef]
- Rehrl, K.; Gröechenig, S.; Hochmair, H.; Leitinger, S.; Steinmann, R.; Wagner, A. A Conceptual Model for Analyzing Contribution Patterns in the Context of VGI. In *Progress in Location-Based Services*; Krisp, J.M., Ed.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 373–388.
- 28. Goodchild, M.F.; Li, L. Assuring the quality of volunteered geographic information. Spat. Stat. 2012, 1, 110–120. [CrossRef]
- 29. Tobler, W. A Computer Movie Simulating Urban Growth in the Detroit Region. Econ. Geogr. 1970, 46, 234–240. [CrossRef]
- 30. Bordogna, G.; Carrara, P.; Criscuolo, L.; Pepe, M.; Rampini, A. A linguistic decision making approach to assess the quality of volunteer geographic information for citizen science. *Inf. Sci.* 2014, 258, 312–327. [CrossRef]
- 31. Schade, S.; Tsinaraki, C. Survey report: Data management in Citizen Science projects. In *JRC Technical Report for European Commission;* Publication Office of the European Union: Luxembourg, 2016.
- 32. Bordogna, G.; Frigerio, L.; Kliment, T.; Brivio, P.A.; Hossard, L.; Manfron, G.; Sterlacchini, S. "Contextualized VGI" Creation and Management to Cope with Uncertainty and Imprecision. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 234. [CrossRef]

- 33. Klien, E. A rule-based strategy for the semantic annotation of geodata. Trans. GIS 2007, 11, 437–452. [CrossRef]
- 34. Stadler, C.; Lehmann, J.; Höffner, K.; Auer, S. LinkedGeoData: A core for a web of spatial open data. *Semant. Web* 2012, *3*, 333–353. [CrossRef]
- Codescu, M.; Horsinka, G.; Kutz, O.; Mossakowski, T.; Rau, R. DO-ROAM: Activity-Oriented Search and Navigation with OSM. In *International Conference on GeoSpatial Sematics*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 88–107.
- 36. Koswatte, S.; Mcdougall, K.; Liu, X. Ontology driven VGI filtering to empower next generation SDIs for disaster management. *Proc. R@Locat.* 2014. Available online: http://ceur-ws.org/Vol-1142/RatLoc14.pdf (accessed on 9 September 2021).
- 37. Bakillah, M.; Liang, S.H.L.; Zipf, A.; Arsanjani, J.J. Semantic Interoperability of Sensor Data with Volunteered Geographic Information: A Unified Model. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 766–796. [CrossRef]
- Díaz, L.; Núñez-Redó, M.; González, D.; Gil, J.; Aragó, P.; Pultar, E.; Huerta, J. Interoperable Search Mechanisms for Web 2.0 Resources. Int. J. Spat. Data Infrastruct. Res. 2012, 7, 277–299.
- CCI Cross Community Interoperability-Annex B-CCI Thread Scope. 2015. Available online: http://www.opengeospatial.org/ pub/www/ows10/rfq/annexb-cci.html#cci-thread-scope (accessed on 8 September 2021).
- 40. Cho, W.C.; Richards, D. Ontology construction and concept reuse with formal concept analysis for improved web document retrieval. *Web Intell. Agent Syst. Int. J.* 2007, *5*, 109–126.
- 41. Gruber, T.R. Toward principles for the design of ontologies used for knowledge sharing. *Int. J. Hum. Comput. Stud.* **1995**, *43*, 907–928. [CrossRef]
- Hitzler, P.; Krötzsch, M.; Rudolph, S. Foundations of Semantic Web Technologies; Chapman & Hall/CRC: Boca Raton, FL, USA, 2009; pp. 11–18.
- 43. Hitzler, P.; Krötzsch, M.; Parsia, B.; Patel-Schneider, P.F.; Rudolph, S. OWL 2 web ontology language primer. *W3C Recomm.* 2009, 27, 123. Available online: http://www.w3.org/TR/owl2-primer/ (accessed on 9 September 2021).
- 44. Gottwald, S. Set theory for fuzzy sets of higher level. Fuzzy Sets Syst. 1979, 2, 125–151. [CrossRef]
- 45. Zadeh, L.A. Quantitative fuzzy semantics. Inf. Sci. 1971, 3, 177–200. [CrossRef]
- 46. Ding, Z.; Peng, Y.; Pan, R. BayesOWL: A Prototype System for Uncertainty in Semantic Web. In *Soft Computing in Ontologies and Semantic Web*; Ma, Z., Ed.; Studies in Fuzziness and Soft Computing; Springer: Berlin/Heidelberg, Germany, 2006; Volume 204.
- Cross, V.V. Fuzzy Ontologies: The State of the Art. In Proceedings of the 2014 IEEE Conference on Norbert Wiener in the 21st Century (21CW), Boston, MA, USA, 24 June 2014; pp. 1–8.
- 48. Samani, Z.R.; Shamsfard, M. The State of the Art in Developing Fuzzy Ontologies: A Survey. arXiv 2018, arXiv:1805.02290.
- Calegari, S.; Ciucci, D. Fuzzy Ontology, Fuzzy Description Logics and Fuzzy-OWL. In Proceedings of the International Workshop on Fuzzy Logic and Applications, Camogli, Italy, 7–10 July 2007.
- Lau, R.Y.; Song, D.; Li, Y.; Cheung, T.C.; Hao, J.X. Toward a fuzzy domain ontology extraction method for adaptive e-learning. *IEEE Trans. Knowl. Data Eng.* 2009, 21, 800–813. [CrossRef]
- 51. Sanchez, E.; Yamanoi, T.; D'Amico, E. Fuzzy Ontologies for the Semantic Web. In International Conference on Flexible Query Answering Systems (FQAS); Springer: Berlin/Heidelberg, Germany, 2006; pp. 691–699.
- 52. Tho, Q.T.; Hui, S.C.; Fong, A.C.M.; Cao, T.H. Automatic fuzzy ontology generation for semantic web. *IEEE Trans. Knowl. Data Eng.* **2006**, *18*, 842–856. [CrossRef]
- 53. Samani, Z.R.; Shamsfard, M. On the application of fuzzy ontology for qualitative spatial reasoning. *Int. J. Next Gener. Inf. Technol.* **2012**, *3*, 9–17.
- Dey, L.; Abulaish, M. Fuzzy ontologies for handling uncertainties and inconsistencies in domain knowledge description. In Proceedings of the 2008 IEEE International Conference on Fuzzy Systems (FUZZIEEE), Hong Kong, China, 1–6 June 2008; pp. 1366–1373.
- 55. Zhai, J.; Shen, L.; Liang, Y.; Jiang, J. Application of fuzzy ontology to information retrieval for electronic commerce. In Proceedings of the 2008 International Symposium on Electronic Commerce and Security, Guangzhou, China, 3–5 August 2008; pp. 221–225.
- Straccia, U. Towards a Fuzzy Description Logic for the Semantic Web. In *European Semantic Web Conference (ESWC)*; Gomez-Perez, A., Euzenat, J., Eds.; LNCS 3532; Springer: Berlin/Heidelberg, Germany, 2005; pp. 167–181.
- Bal-Bourai, S.; Mokhtari, A. Poss-OWL 2: Possibilistic Extension of OWL 2 for an uncertain geographic ontology. In Proceedings of 18th International Confonference on Knowledge-Based and Intelligent Information & Engineering Systems—(KES2014), Procedia Computer Science; Elsevier: Amsterdam, The Netherlands, 2014; Volume 35, pp. 407–416.
- Qi, G.; Ji, Q.; Pan, J.Z.; Du, J. PossDL—A Possibilistic DL Reasoner for Uncertainty Reasoning and Inconsistency Handling. In Lecture Notes in Computer Science (LNCS 6089), Proceedings of the 7th Extended Semantic Web Conference, ESWC 2010, Heraklion, Crete, Greece, 30 May–3 June 2010; Aroyo, L., Antoniou, G., Hyvönen, E., ten Teije, A., Stuckenschmidt, H., Cabral, L., Tudorache, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 416–420.
- 59. De Trè, G.; de Caluwe, R. Level-2 fuzzy sets and their usefulness in object-oriented database modelling. *Fuzzy Sets Syst.* 2003, 140, 29–49. [CrossRef]
- 60. Yager, R.R. On the specificity of a possibility distribution. Fuzzy Sets Syst. 1992, 50, 279–292. [CrossRef]
- Dubois, D.; Prade, H. Possibility Theory, Probability Theory and Multiple-valued Logics: A Clarification. *Ann. Math. Artif. Intell.* 2002, 32, 35–66. [CrossRef]

- 62. Bilgiç, T.; Burhan Türkşen, I. Measurement of membership functions: Theoretical and empirical works. In *The Handbooks of Fuzzy Set Series, Vol. 1: Fundamentals of Fuzzy Sets*; Dubois, D., Prade, H., Eds.; Chapter 3; Kluwer Academic Publishers: London, UK, 1999; pp. 195–202.
- Huitzil, I.; Straccia, U.; Díaz-Rodríguez, N.; Bobillo, F. Datil: Learning Fuzzy Ontology Datatypes. In Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations, Proceedings of the 17th International Conference, IPMU 2018, Cádiz, Spain, 11–15 June 2018; Medina, J., Ojeda-Aciego, M., Luis Verdegay, J., Pelta, D.A., Cabrera, I.P., Bouchon-Meunier, B., Yager, R.R., Eds.; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2018; Volume 854, pp. 100–112. [CrossRef]
- 64. Yun, W.; Zhang, X.; Li, Z.; Liu, H.; Han, M. Knowledge modeling: A survey of processes and techniques. *Int. J. Intell. Syst.* 2021, 36, 1686–1720. [CrossRef]