

Article

# Implicit, Formal, and Powerful Semantics in Geoinformation

Gloria Bordogna , Cristiano Fugazza \* , Paolo Tagliolato Acquaviva d’Aragona  and Paola Carrara 

Institute for Electromagnetic Sensing of the Environment, National Research Council of Italy, via Bassini 15, I20133 Milan, Italy; bordogna.g@irea.cnr.it (G.B.); tagliolato.p@irea.cnr.it (P.T.A.d.); carrara.p@irea.cnr.it (P.C.)

\* Correspondence: fugazza.c@irea.cnr.it

**Abstract:** Distinct, alternative forms of geosemantics, whose classification is often ill-defined, emerge in the management of geospatial information. This paper proposes a workflow to identify patterns in the different practices and methods dealing with geoinformation. From a meta-review of the state of the art in geosemantics, this paper first pinpoints “keywords” representing key concepts, challenges, methods, and technologies. Then, we illustrate several case studies, following the categorization into implicit, formal, and powerful (i.e., soft) semantics depending on the kind of their input. Finally, we associate the case studies with the previously identified keywords and compute their similarities in order to ascertain if distinguishing methodologies, techniques, and challenges can be related to the three distinct forms of semantics. The outcomes of the analysis sheds some light on the diverse methods and technologies that are more suited to model and deal with specific forms of geosemantics.

**Keywords:** geosemantics; implicit semantics; formal semantics; powerful semantics



**Citation:** Bordogna, G.; Fugazza, C.; Tagliolato Acquaviva d’Aragona, P.; Carrara, P. Implicit, Formal, and Powerful Semantics in Geoinformation. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 330. <https://doi.org/10.3390/ijgi10050330>

Academic Editor: Wolfgang Kainz

Received: 12 March 2021

Accepted: 1 May 2021

Published: 13 May 2021

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



**Copyright:** © 2021 by the authors. 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/>).

## 1. Introduction

Semantics is cornerstone in state-of-the-art data management, notwithstanding the specific domain; without semantics, we would helplessly drown in a deluge of unintelligible Big Data. Let aside the enormous literature on this topic in the field of Linguistics and, even before that, in Philosophy, representing and managing semantics is frequently regarded to as the solution to heterogeneity in data retrieval and exploitation in Computer Science (CS) [1–3]. This paper relates to a specific domain in the landscape of semantics-aware CS, i.e., geospatial information provided in the form of both data and metadata. This is a particularly challenging domain as the non-textual nature of most geospatial data means that the indexing practices of generalist search engines are ineffective; hence the need for semantics representation and management.

Both Sheth et al. [4] and Uschold [5] provide a coarse-grained categorization of semantics; the latter includes the following four categories: (i) implicit semantics, (ii) informally expressed semantics, (iii) formally expressed semantics for human consumption, and (iv) formally expressed semantics for machine processing. In practice, the first three levels fall in the first category defined in [4], which proposes the following classification:

**Implicit semantics** is the semantics not explicitly represented, i.e., not directly usable by machines to derive new knowledge.

**Formal semantics** when semantics is represented in some sort of formalism, in order to be machine readable and processable, e.g., in the form of ontologies.

**Powerful (soft) semantics** when semantics is represented in forms that enable overcoming crisp set-based formalisms, allowing representing degrees of memberships and certainty, e.g., by using fuzzy approaches and contextual time-varying semantics.

In our opinion, the second classification not only includes the first one but, at the same time, empowers the fourth level of the first by offering two distinct representation classifications, opening towards methods that better mimic the human soft and flexible

approaches to reasoning and decision making. This is the main motivation for adopting this second classification method, exporting its concepts in the geospatial domain and reflecting them in the forthcoming Sections of this paper. Albeit there is apparently a broad spectrum of technologies that fall under the umbrella of each of these categories (in fact, Almeida et al. [6] elaborate on the notion of *semantic continuum*), we will discuss their common traits.

As regards high level categorization of the forms of semantics, Gärdenfors [7] distinguishes between “symbolic”, “associationist”, and “conceptual”, providing the latter with a spatial characterization. His *cognitive spaces* feature interesting analogies with notions that are typical of the geospatial domain (e.g., spatial intersection). Still, non-symbolic approaches are mostly contained in the category of implicit semantics according to the classification by Sheth.

The ultimate purpose of this work is to provide the reader with awareness of directions on the main issues, challenges, and possible solutions to address the different categories of semantics defined by Sheth in the domain of geoinformation. In a nutshell, this paper outlines which technologies are more appropriate to consider when tackling a given research problem. The importance of this topic for the geospatial community is attested by the increasing relevance of semantics as the “glue” between heterogeneous thematic domains and also across their individual workflows. On the one hand, inter-disciplinary interoperability requires mapping of the individual terminologies used for annotating data. On the other, effective discovery and provision of geospatial data requires fine-grained characterization of resources (i.e., semantic metadata) not only for data, but also for services, APIs, instruments, data providers, etc.

The hypothesis behind our work is that Sheth’s three forms of semantics are also reflected in the geosemantics context. The objective of our paper is then to identify technologies, methodologies, challenges, and solutions that are distinctive for the implicit, the formal, and the powerful geosemantics in order to orient the reader in problem solving. To achieve this, by analyzing recent reviews and editorial papers on geosemantics, we first mine which are the main technologies, methodologies, research challenges, and solutions presented by the authors, regarding them as keywords (Section 2.3).

Successively, we perform a two-step analysis by first discussing selected case studies involving the management of implicit, formal, and powerful geosemantics. The choice of the case studies has been performed by taking into account both their belonging to one of the semantic categories of Sheth (depending on the characteristics of their inputs) and the variety and representativeness of application domains as outlined in [8]. Specifically, the varied and most representative applications to which geomatics can be put include urban planning, disaster management, assessment of biodiversity, and land administration. We then associate the keywords with the case studies and assess whether Sheth’s categories are characterized by distinguishing keywords, i.e., specific methods, technologies and solutions, thus allowing for a more distinctive clustering of the keywords with respect to what emerged from the metareview in Section 2.3.

A contribution of this paper is also the methodological workflow we followed in order to characterize the forms of semantics in geoinformation with their preferred/elective approaches.

## 2. Materials and Methods

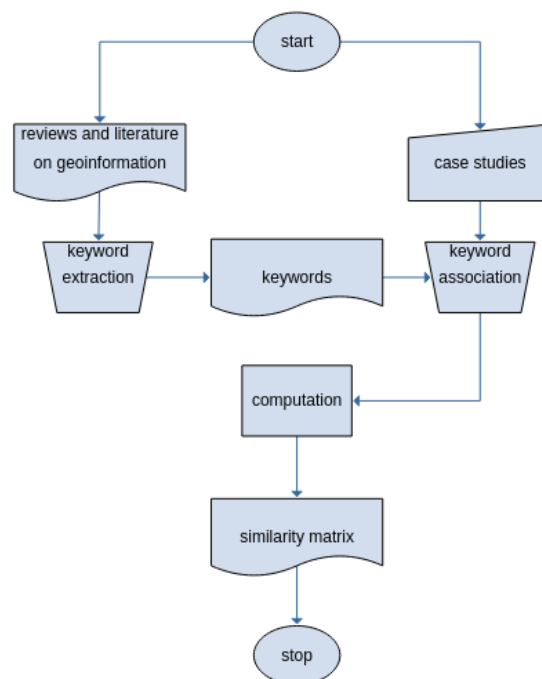
This Section is organized as follows: Section 2.1 details our aim and the workflow we followed to confirm our hypothesis. Section 2.2 explains the categorization of semantics in the main reference work [4] inspiring this paper; then, Section 2.3 presents a meta-analysis of the literature on geosemantics as discussed in recent surveys and review papers. Sections 2.4–2.6 present the case studies we selected according to the criteria expressed above.

### 2.1. Workflow

In this work, we aim at investigating whether the three forms of semantics by Sheth et al. [4] can be related to distinguishing methodologies, techniques, and knowl-

edge sources among those found in the literature on geospatial information. This is by no means a foregone conclusion and these distinguishing methodologies may not be the same as in other contexts. In fact, the geospatial domain sometimes diverges from current trends because of its specificities (e.g., proposing service-oriented architectures as opposed to resource-oriented ones).

To this aim, we define the workflow whose main phases are depicted in Figure 1: Top-left, a meta-review of recent surveys of papers illustrating applications of geospatial information management is performed (Section 2.3). The meta-review allows for identifying topics, research challenges, and solutions; these are considered to be keywords and represented on the right side of Figure 2. On the top-right hand side, assuming as starting point of our analysis the aforementioned three forms of semantics (whose definitions are clarified in Section 2.2), we select and analyze several case studies, categorizing them according to these three forms of semantics on the basis of the characteristics of their inputs (Sections 2.4–2.6).



**Figure 1.** Depiction of the workflow followed.

Finally, in order to substantiate the hypothesis behind this work—that Sheth’s categories are also reflected in the geosemantics context—the results yielded by the two previous independent phases are cross-referenced in order to compute a similarity matrix on the basis of the keywords associated with the case studies. Specifically, this is achieved by verifying that the intra-similarities (similarity degrees between pairs of case studies belonging to the same form of geosemantics) are greater than the inter-similarity degrees between pairs of study cases classified as different forms of geosemantics. The greater the intra-similarity with respect to the inter-similarity, the more distinctive the methods and technologies characterizing the three forms of geosemantics.

## 2.2. Three Shades of Semantics

Looking at geosemantics through the lenses proposed by Sheth allows for categorizing in a minimal set of classes the broad (and ever-growing) landscape of topics (comprising both methods and technologies) that populate this domain. Otherwise, the implications of information source heterogeneity (as far as genre and nature are concerned), information multidimensionality, and domain knowledge dependency easily yield a multiplicity of

classes that configures a semantic continuum *à la* Almeida [6]. Since data source heterogeneity, cross-domain interaction, data/process imperfection, and big data volumes are common traits in the geospatial domain, distinguishing between *implicit*, *formal*, and *powerful* semantics allows us to divide the presented case studies in three categories with a clear solution of continuity.

*Implicit* semantics refers to the kind that is implicit in data and that is not represented explicitly in any machine-processable syntax. It is typically related to concepts and relationships between them that are not represented in a formal way but are embedded in multimedia documents, i.e., their “meaning is conveyed based on a shared understanding derived from human consensus” [5]. These can be natural language documents, multi-spectral images, time series of measurements, video frame sequences, audio recordings, undocumented tabular data, etc. The main objective of extracting implicit semantics is to cope with the inherent ambiguity characterizing it. In fact, terms in texts, visual aspects in images, etc., can mean different things depending on both context and knowledge of people [9]. It should be noted that *implicit* does not mean missing a knowledge-based underpinning but that the latter does not (or cannot) be given a formal representation, such as in the assessment by a domain expert.

In more general terms, we can state that semantics that are represented in some well-formed syntax (governed by syntax rules) is referred to as formal semantics. In 2001, Berners-Lee et al. [10] stated that “The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation”. As such, the Semantic Web (SW) is the most apparent embodiment of semantics in the field of Internet-mediated contents and applications. Here, the inflection we give to this term is that of *formal* semantics, specifically those provided by decidable fragments of First-Order Logic (FOL) [11]. In fact, “formal” is the category name that Sheth et al. give to this kind of semantics [4], analogous to “formal semantics for machine-processing” in Uschold’s categorization [5]. Explicit representations of formal semantics include knowledge graphs, ontologies, and the like.

Finally, Sheth et al. introduce the concept of *powerful* semantics, intended as formal semantics which is empowered with the ability to represent not only precise and well-defined concepts and relationships, but also imprecise and uncertain concepts and gradual relationships, whose meaning can be subjective, vague, and variable depending on several contextual conditions [12]. The ability of formal frameworks to represent and manage powerful semantics is indeed aimed at performing approximate and qualitative reasoning in order to discover implicit concepts and relationships, possibly uncertain and imprecise too, although accurate enough to be useful to solve some needed task.

### 2.3. A Meta-Analysis Perspective

Timothy Tambassi, in his Preface to the book “The Philosophy of GIS” [13], pointed out that the literature on GIS is heterogeneous and scattered, primarily because of the multiple branches of knowledge that use, manage, and create geographic information. This is also true for geosemantics, whose literature configures a conceptual ‘forest’ of issues, topics, technologies, methodologies, challenges, and solutions where it is easy to lose orientation. To frame approaches in the field of geosemantics, we have taken into account some stimulating overview papers on this subject which appeared in the last decade and tried to examine and categorize the topics, research challenges, and solutions described. It is a meta-analysis exercise that considered the papers described in the following.

Kokla et al. [14] offers a comprehensive review of the contributions that represent a progress in geospatial semantics since 2015; it focuses around two main topics, i.e., information modeling (ontologies and their development) and (latent) knowledge elicitation (from unstructured or semi-structured content, based in particular on textual contents). This paper reviews more than 150 works; among them are papers that present categorizations of methods and approaches to geosemantics, such as [15–19]. Other cited contributions report on the efforts for describing the methods at hand: [20–32]. Furthermore, in this

review the reader can find many works that exemplify the former within a great number of applications; among these [33–55].

Hu [56] provides an overview and a review of important contributions dealing with six major research areas in geospatial semantics, i.e., “semantic interoperability and ontologies” [16,24,38,57–70], “digital gazetteers” [71–86], “Geographic Information Retrieval” [32,87–106], “geospatial Semantic Web and Linked Data” [43,107–114], “place semantics” [47,115–121], “cognitive geographic concepts and qualitative reasoning” [70,119–124].

Janowicz et al. [125] is a rich overview of the geosemantics landscape focusing on some selected topics that the authors deem of particular interest; the contributions reviewed are organized according to these. With respect to the question on what kinds of Geospatial Classes should be distinguished, they cite [16–18,63,65,66,68,126–129]; instead, the question on how to reference Geospatial Phenomena is supported by [113,123,130–132]. Discovering events and accounting for geographic change are faced and fostered in [133–136], Handling places and moving object trajectories is dealt with in [70,77,90,133,137–143]. The following papers are cited with reference to comparison, alignment, and translation of Geospatial Classes [15,69,90,144–149]. Finally, the issues raised by processing, publishing, and retrieving geodata are tackled by in [150–156].

The approach changes in [157]: Rather than reviewing papers dealing with projects and issues related to geosemantics, it reviews ideas rooted in cognitive science and linguistics for sketching their application to semantics of geographic information. It discusses notions from 1990 to 2010 and shows why and how these ideas have been productive for dealing with semantics.

We also considered a couple of papers that are not strictly reviews but, in our opinion, are worth being included as they offer a landscape of trends and contributions in geosemantics. Janowicz et al. [112], an editorial paper on the Semantic Web, outlines the research field of geospatial semantics, highlights major research directions and trends, and takes a glance at future challenges. Another editorial paper [158], considers VGI (Voluntary Geographic Information) and claims that geospatial Linked Data and Knowledge Graphs, when used for implementing intelligent data search, can result in precise data-sharing services.

The less recent work we considered is [159], where the author observes that the main approaches to overcome semantic heterogeneity rely on ontologies that, having a priori definitions, are decontextualized. On the contrary, he affirms that semantics reconciliation needs to take into account context-based meanings. Since “meaning and context are dynamically emergent from activity and interaction, determined in the moment and in the doing”. He further highlights the limitations of representational approaches. In fact, the latter assume that context is stable, delimited information that can be known and encoded in just another information layer or another ontology in an information system. These are the reasons why this work encourages non-representational modelling formalisms to cope with semantic interoperability in sharing and integrating geographic information.

By analyzing the above overviews, we have extracted a list of terms that the authors pinpointed as topics of interest, research challenges, or solutions, which we regard as keywords. The correspondence between the keywords and the respective originating reviews can be found in supplementary material. The keywords are listed on the right side of the diagram in Figure 2. The list is wide enough to suggest how large is the playground offered by geosemantics.

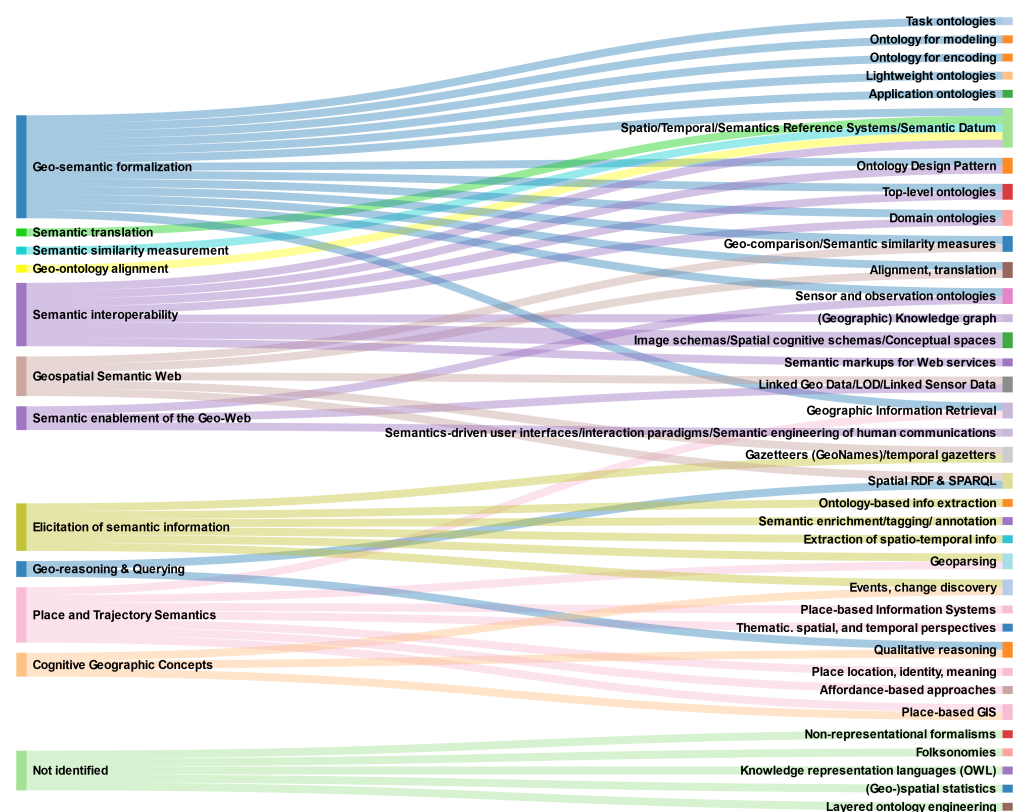
Still, this list may be biased, being based on authors’ views and reviews of a rapidly evolving literature, and some terms can have overlapping meanings. For instance, more recent reviews, such as Kokla et al. [14], produced an increase in this term list, due to the emergence of mobile and social applications, IoT, AI, etc. in the last five years. These research fields introduced novel concepts, such as lightweight ontologies. This increase is also due to the paradigm shift, dating back in 2012 [20], from the general-purpose Web to communities and their specific perspectives, pushed in turn by the movement of Critical GIS [160]. With reference to the notion of Digital Earth, in [161] the authors



solicited “a network of theories that fosters interoperability without giving up on semantic heterogeneity”. As such, it is possible that more recent works may further populate the list in Figure 2.

In the papers we examined, the authors suggested a grouping of these keywords according to some categories, listed on the left side of the diagram. Some keywords can be related to multiple categories as they can be good suggestions in diverse application scenarios. As an example, term “gazetteers” has been presented in some works as dealing with either “geospatial Semantic Web” or “elicitation of semantic information”; “domain ontologies” have been used in works coping with “geo-semantic formalization” and “semantic interoperability”. On the other hand, there are categories that can be tackled with multiple strategies; for example, geosemantics issues falling under category “cognitive geographic concepts” have been dealt with in projects on either “events-change discovery”, “place-based GIS”, or “qualitative reasoning”.

Figure 2 makes it apparent that the categories on the left are not associated with distinguishing topics and solutions on the right, i.e., the reviews did not succeed in letting patterns emerge in the geosemantics “forest”, thus making order in the diverse practices.



**Figure 2.** Diagram connecting keywords in geosemantics (right) and their categorization (left), as found in the reviews taken into consideration.

#### 2.4. Implicit Geosemantics

In [14], the extraction of implicit geosemantics is named “elicitation of semantic information”. Under this interpretation, the term is used in a broader sense to encompass processes aimed to make latent knowledge explicit from unstructured or semi-structured contents. These processes focus on eliciting a structured representation of information in various forms, such as semantic metadata, links to ontology concepts, collections of topics, geotagged maps and images, etc. Sources of implicit geosemantics are multimedia documents, in the form of unstructured and semi-structured textual documents, pictures taken from cameras, images from remote sensing, audio and video files. In most cases,

metadata are available but are generally insufficient to representing and understanding the contents.

Typically, unstructured texts, posts in social networks, and news streams may refer to geographic names into their contents to describe events, points of interest (POIs), and places. The discipline that extracts geographic contents from unstructured and semistructured texts in order to index them and enable the evaluation of both content and spatial queries is Geographic Information Retrieval (GIR) [87]. Images are another potential source of geosemantic information. Photos may depict geographic places without explicitly mentioning their name or geolocation. With regard to video files, we can consider TV news reporting events relative to specific geographic areas. Finally, remote sensing images may contain representations of the status of the environment with respect to the occurrence of geo-temporal phenomena and events going on in a given area. The segmentation of images in order to extract geographic footprints of places and events can be performed by applying spatio-temporal analysis. The latter is primarily based on (i) domain experts' knowledge; (ii) statistical and machine learning approaches, or (iii) hybrid approaches combining the previous two [162].

Some important challenges of implicit geosemantics extraction within multimedia documents are related to three main objectives:

- (i) reconciling the place and space conceptualizations of geosemantics: while the "platial" (based on place) perspective is usually defined within texts by textual place names, linguistic descriptions, and the semantic relationships between places, the spatial perspective typical of georeferenced maps explicitly represents the geometries by their coordinates, distances, topology, and directions, but mostly lack descriptions of their meanings. This reconciliation from platial to spatial and vice-versa requires modeling uncertainty of the recognition process;
- (ii) increasing human perception of the semantics of geoinformation by considering users' spatial, temporal, and content needs and preferences. This amounts to identifying and summarizing geographic contents on the basis of distinct spatial, temporal, and content granularities;
- (iii) enhancing interoperability of the geoinformation semantics representation in order to be able to re-use it within different contexts and applications. This is achieved by adopting standards and domain/task/application ontologies.

Basically, artificial intelligence approaches comprising different methodologies (such as soft computing, clustering, genetic algorithms, geostatistic analysis, neural networks, support vector machines, and the like) are applied to extract implicit semantics from multimedia documents. Knowledge bases are used to support the analysis: These may take the form of gazetteers, DBpedia (<https://wiki.dbpedia.org/> accessed on 1 April 2021), generic and domain thesauri such as WordNet (<https://wordnet.princeton.edu/> accessed on 1 April 2021), geo ontologies, and thematic geospatial information. In the following, we present some case studies focused on to the above challenges which consider different genres of geographic contents (basically, objects, events and moving objects' trajectories) within distinct categories of multimedia documents (textual documents and social media posts).

A synoptic view of the four case studies dealing with an implicit form of semantics is reported in Table 1: Besides the identifier of the case study, its acronym, and a brief description, the table reports the type of input, the method it applies, the type of generated output, and its potential use. It can be noticed that the type of input is either unstructured textual documents or social media documents, a kind of data that typically contain the implicit form of semantics. It can be also noticed that the outputs contain more explicit geosemantics, constituted by geofootprints of documents, spatio-temporal clusters of events, trajectories, and georeferenced placenames.

As for the application domains that are covered by the case studies, that in Section 2.4.1 is related to retrieval of georeferenced information, providing urban planners with effective means for mining knowledge of territorial resources. The case study in

Section 2.4.2 performs trajectory mining to support mobility planning for tourists. The case study in Section 2.4.4 shows that disaster management can be fostered by timely event detection and, finally, the case study in Section 2.4.4 is about geo-gazetteer creation from VGI, in support of land administration.

#### 2.4.1. From “Place” to “Space”: Representing Uncertainty of Geoinformation within Texts to Support Geographic Information Retrieval

In [163], a GIR system was proposed that allows for extracting implicit geosemantics within contents of textual documents through the identification of *fuzzy geographic footprints*, i.e., the distinct locations on Earth referred to by documents.

The GIR model applies soft computing methods; specifically, the evaluation of multiple bipolar criteria [164,165] aggregated based on a p-norm operator [166] to extract the fuzzy footprints of documents representing their geographic focus. In a nutshell, some criteria have a positive influence on the selection of geographic names within the text as footprints of a document (for example, when the initial characters of the term is a capital letter, when the term occurrence is close to positive anchor terms such as “street”, “city”, “nation”, etc.). Others have a negative influence (for example when the term is preceded by negative anchor terms such as “Sir”, “Mr”, “Mrs”, etc.).

The prototypical system, has the classic structure of an Information Retrieval System (IRS) [163], consisting of two main components: the *Indexing Module* and the *Retrieval Module*. The Indexing Module has two main sub-modules: the *Full-Text Indexing* and the *GeoIndexing* sub-modules. The former performs full text indexing of the documents to represent their significant contents, and generates the textual inverted index to enable content based searches. Instead, the GeoIndexing sub-module identifies the fuzzy footprints of documents by the support of a knowledge base that comprises both a geo-ontology and a rule-base that encodes the heuristic knowledge required to cope with geo/non-geo ambiguities during geoparsing, and with geo/geo ambiguities during geocoding. An example of geo/non-geo ambiguity is the case of a place name having also a non geographic meaning such as “Nice” (France), “Crema, Brindisi” (Italy), and “Of” (Turkey). Instead, geo/geo ambiguities are due to distinct locations on Earth having the same place name, such as Rome, Paris, London, etc.). The disambiguation rules take into account both the geographic context, based on the shared assumption that “close places are more closely related than far places”, and the textual context, based on the consideration that distinct geographic names appearing close in text are also closely related in geographic space. This way, place names within documents are associated with a fuzzy footprint in the geographic space, thus reconciling the two conceptualizations of geosemantics and enabling both content and spatial searchers.



**Table 1.** Dimensions of case studies (Implicit geosemantics).

Subsection	Name	Description	Input Data	Method Classification	Output Data	Potential Use
Section 2.4.1	GeoFinder	GIR extracting and modeling uncertainty of geofootprints in textual documents	textual documents from etherogeneous sources (tested on a collection of articles about Energy power stations and on CLEF2008 collection)	fuzzy computational intelligence	uncertain geofootprints of textual documents, i.e., fuzzy sets of geographic points with uncertainty degrees in [0,1] associated with the index terms of textual documents	Performing spatial queries (expressing metrical and topological conditions) in combination with content-based queries on textual collections
Section 2.4.2	Events spatio-temporal footprint	footprint detection of events' popularity (from Twitter)	semi-structured text (Twitter) messages with explicit geotags	unsupervised learning	clusters of punctual geo-temporal footprints of an event or topic identified by a set of keywords in a given time-lapse	geo temporal analysis of events reported in social networks
Section 2.4.3	Tour miner	mining popular tourists' tours (from Twitter)	semi-structured text (Twitter) messages with explicit geotags	knowledge-based semi-supervised learning	popular tours identified by a hierarchy of clusters containing sets of "close" paths (a path being an ordered lists of geographic entity names)	geo temporal analysis of tourists' mobility based on social network messages
Section 2.4.4	Eliciting Geographic Gazetteer	Extracting place names and their footprints from social networks	images, captions and metadata from Flickr	unsupervised learning	geographic gazetteer of place names	updating or creating geographic gazetteer

#### 2.4.2. Detecting Periodic/Episodic Events from Social Networks with Desired Spatio-Temporal Granularity

The paper by [167] proposes an approach to discover events of interest from social media by modeling the distinct spatio-temporal granularity. The main characteristic of this study is flexibility in detecting events characterized by either an hypothetical periodic or episodic timestamp, thus allowing confirming a priori knowledge of their possible geotemporal regularities. Given a set of sources of spatio-temporal information, such as Twitter, the methodology first performs a focused crawling of the selected social media contents to collect candidate messages related to an event of interest; successively, the collected messages are analyzed by means of an original, density-based spatio-temporal clustering algorithm. The latter is defined by extending the DBSCAN algorithm to group messages densely located in the spatio-temporal domain. Its output is a set of spatio-temporal clusters with arbitrary shapes: these identify the areas on Earth where an event matching the keywords (i.e., the parameters used to filter the messages) occurred within a given time span, possibly with a given periodicity.

The exploration is interactive and multi-granular, allowing analysts to customize not only the topics of interest, i.e., the category, but also the time period and the spatial density so as to fit different spatio-temporal scales. One can specify (i) a set of keywords of interest to filter the messages about an event or a topic (e.g., traffic jam, hurricane, landslide, football match), (ii) the desired granularity of the time period of analysis (such as each day, month, year) and (iii) the desired spatial granularity needed to form a cluster, defined by spatio-temporal density of messages. Each cluster generated by the algorithm can be identified by the list of the most representative keywords that were found in the messages of the cluster, thus representing the cluster's semantics. The use of thesauri [168] helps identify the more general terms expressing the meaning of the specific terms found in individual messages of the cluster. As far as the representation of the geographic footprint of each cluster is concerned, a convex hull can be computed from the geographic coordinates of the messages in each cluster to obtain a polygon representation of the geo-footprint.

#### 2.4.3. Discovering and Summarizing Moving Object Trajectories from Twitter

The work described in [169] proposes an approach to identify, track, and analyze popular tours of tourists visiting a Region Of Interest (ROI) based on the Tweets they publish.

The solution is constituted by two main suites of tools: the FollowMe suite for tourist identification and tracking and the TripsAnalysis suite for popular tour mining.

The FollowMe suite allows users to submit spatial queries to the Twitter API to find *hang* tweets, i.e., tweets posted in the area of the monitored airports. For each user identified by means of hang tweets, the FollowMe suite queries (through the Twitter API) his/her timeline, i.e., the history of tweets posted by the user, to get tweets tracked.

Given a ROI, trips that occur in the ROI are reconstructed and extracted by querying hang tweets and tracked tweets previously stored in the local data base. Reconstructed trips are represented by a list of geographic coordinates, ordered according to message creation time and are exported through the web service interface.

The Trip Analysis Suite performs the activities of knowledge discovery on trips collected by the FollowMe Suite. A knowledge-based trajectory clustering method allows analyzing trips based on customizable semantics. The analyst can specify both the desired granularity and semantics of the analysis by providing a vector layer of geographic slots (geo-slots) of interest. These are drawn from external interoperable sources that the algorithm exploits to conflate the trips' points to ease their grouping. For example, it is possible to conflate and then analyse trips with respect to the visited municipalities, regions, countries, city's neighborhoods, ZIP codes, etc. This way, the algorithm first geo-partitions the trips represented based on the ordered sequence of geographic coordinates into a conflated trip representation consisting of an ordered sequence of geo-slot identifiers, i.e., a string. This way, different geo-slots partitions provide different interpretations, scales, and semantics of the analysis.

The conflated trips can be easily clustered using a complete-link hierarchical trajectory clustering algorithm using a string-similarity matching. Matching is applied to the concatenated identifiers of the geo-slots in the conflated trips' representation. Finally, popular tours can be identified by selecting a partition of the clusters' hierarchy by specifying either a threshold on the minimum desired inter-similarity of conflated trips within a popular tour, or a minimum number of trips that a popular tour must contain.

#### 2.4.4. Creation of Geographic Gazetteers by Volunteered Geographic Information Analysis

Constructing geographic gazetteers is very costly in terms of human effort and, once created, they need to be constantly updated. The work [81] proposes to exploit data science for the extraction of semantic information on toponyms, places, and POIs from big geoinformation created by volunteers on the Web, specifically from geotagged Flickr pictures. The aim is to enrich and update current gazetteers by automatically creating digital gazetteers of georeferenced place names such as "city center", "shopping district", and POIs associated with keywords and geofootprints. The ultimate purpose is to support diverse applications, such as geographic information retrieval (GIR), digital library services, and systems using spatio-temporal knowledge. The geographic footprints are extracted from the GPS locations of Flickr pictures while place descriptions are distilled from their tags. Close GPS locations associated with similar textual descriptions created by distinct volunteers are assumed as identifying the same place. These locations are generally not perfectly matching but usually have a cluster structure in space. This suggested the authors to use a distance-decaying function to measure the membership of candidate point locations assigned to a place so as to present an intuitive user reputation model for trust evaluation.

#### 2.5. Formal Geosemantics

The reason the use case in Section 2.5.1 is exemplar to the transition from implicit to formal geosemantics is twofold. On the one hand, it upholds ontologies as the formalization means, offering less constrained expressiveness to the modeling of geospatial entities; on the other, it tackles a research issue, that of next generation maps, that has roots in cartography and, as such, is typically bound to the interpretation of implicit information mediated by the domain expertise of end users. Most applications of semantics to geospatial information use "lowercase" semantics, such as that of SKOS vocabularies [170] which are not harnessing full expressiveness of ontology languages; others mistake RDF encoding for semantics. Instead, it is important to keep in mind that far more expressive modeling criteria (ontology languages) and inference tools (reasoners) exist. Section 2.5.1 provides both a conceptual model for geo-entities and an exemplar implementation.

*Discovery*, in the sense of "retrieval of geospatial information", is largely dependent on metadata. In turn, semantic characterization of metadata is regarded to as the primary means to achieve interoperability [171] in a domain that is otherwise fraught with heterogeneities [14,56]. Unleashing this potential typically amounts to relating metadata items to entities in the Web of Data (the Linked Open Data Cloud: <https://lod-cloud.net/> accessed on 1 April 2021), such as terms from SKOS vocabularies, people and organizations in FOAF representations [172], etc. Whereas this step may not be strictly necessary for semantics-aware discovery [173], leveraging on these categories of data structures can easily yield semantics-aware resource descriptions. The advantages of this practice are manifold. On the one hand, these data structures may greatly improve user experience in metadata production. On the other hand, traditional metadata can be enriched in order to enable smarter discovery criteria. This is the focus of Section 2.5.2.

Let aside the aforementioned virtuous data structures, there is a large corpus of web-accessible data structures that does not take advantage of ontologies expressed in OWL/OWL2, such as those mentioned above, or schema languages compatible with them (e.g., RDF Schema). As an example, consider the Microdata that is typically embedded in web pages or the XML/JSON data structures that are often used in the enactment of APIs. Section 2.5.3 proposes creation of "semantic twins" of JSON data structures to allow for

transparently accessing heterogeneous data sources. It should be noted that although we already considered the JSON format in the previous Section, in this context the semantics underlying the JSON data (its implicit schema) is made explicit by the mapping to RDF, assuming an interpretation. Some of the (augmented) information contained in the RDF data structures could be fed back to the original JSON ones so as to realize a JSON-LD [174] representation of resources.

Finally, Section 2.5.4 describes a model for *semantic mediation* with the aim of improving geospatial discovery, e.g., by exploiting the smarter metadata originating from creation methodologies akin to those presented in Sections 2.5.2 and 2.5.3. In fact, it is apparent that discovery constitutes a “crucial first step” in the enactment of Spatial Data Infrastructures (SDIs) and nevertheless is “mostly neglected and approached following old paradigms” [112]. Beside harnessing the richer information entailed by semantic characterization of metadata, another key objective of this practice is to implement geospatial data management as a machine-processable API, thus fostering FAIR access to geospatial resources [175]. The rationale for this is that it makes little sense to strive for semantic characterization of metadata and not accomplish the last mile toward their full exploitation by automated agents. The synoptic view of the case studies analysed in relation to formal geosemantics is reported in Table 2.

#### 2.5.1. Holistic Map Representation with Geographic Scenarios

The work in [176] illustrates *Geographic Scenarios* [177], a notion developed on the basis of General System Theory [178] integrating spatial, process, and relational information related to geographical elements and georeferenced events. In contrast with reductionist approaches (such as those dividing geo-entities into themes), Geographic Scenarios propose a holistic view that should be better suited to represent hierarchical connections among geo-entities. Moreover, by favoring space over time, state-of-the-art GIS may fall short of portraying dynamic relationships and causalities.

Basing the conceptual framework of Geographic Scenarios on an ontology allows for expressing multi-hierarchy categorizations and fuzzy boundaries, portraying diverse and complex entities at different scales and dimensions. Geo-characterization is the process by means of which scenarios as well as their individual components are assigned properties and relationships not only on the basis of traditional notions, such as regionalization and classification, but also according to ecology and human-orientation (that are often regarded to as mere thematic dimensions). Events are made first-class citizens in the ontological modeling of geographic scenarios, thus allowing attribution of dynamic relationships between geo-entities.

From a technical viewpoint, the realization that is presented combines relational data with ontology classes and properties by applying SWRL rules [179]; the resulting information is stored in a graph database for querying. Whereas the proposed example does not fully demonstrate the augmented capabilities of geographic scenarios, modularity of the possible semantic underpinning (the ontology) and the scalable solution for storage (a graph database) suggest more extensive implementations.

**Table 2.** Dimensions of case studies (Formal geosemantics).

Subsection	Name	Description	Input Data	Method Classification	Output Data	Potential Use
Section 2.5.1	Holistic map representation with Geographic Scenarios	Ontology-based ingestion of geo-entities into graph databases.	relational databases	rule-based system	graph databases	express diacronic relations and causalities with respect to traditional GIS
Section 2.5.2	EDI/Liftboy	semantic augmentation of geospatial metadata	structured text (metadata documents, XML Schema-based)	information retrieval in a graph database	metadata enriched (aka annotated) with links to RDF entities (often defined by authoritative sources) in semantic-web (RDF) resources (organized in graphs)	semantic discovery (e.g., multilingualism, semantic expansion); disambiguation and preservation of information meaning (in the future and with respect to different audiences)
Section 2.5.3	Semantic twins	augmentation of geospatial metadata based on heterogeneous sources	structured and semi-structured data, JSON- and HTML-based)	information retrieval in a graph database	metadata enriched (i.e., annotated) with links to json entities (exploiting semantic twins that grant consistency of metadata items)	same as previous
Section 2.5.4	Semantic mediation for FAIR access to resources	machine-actionable search of geospatial resources [once geospatial service interfaces (e.g., standard CSW) has been extended to semantic machine actionable API]	REST service interface (API) definitions enriched with semantics	information retrieval based on a task-ontology	machine-actionable semantic augmentation of REST API definitions expressed in Hydra	enablement of semantic agent



### 2.5.2. Ex-Ante and Ex-Post Semantic Characterization of Metadata

In the last decade, our work group has been tasked with the development of the SDI for a national flagship project on marine research. The key approaches were (i) creation of a decentralized network of nodes providing data [180] and (ii) the extensive use of semantics-aware technologies in metadata management [181]. The latter entailed development of a metadata editor that could easily adapt to the ever-changing landscape of metadata formats and profiles [182].

Since no tool in the state-of-the-art allowed for this degree of flexibility, we decided to develop EDI, a brand new metadata editor [183]. Beside allowing for an extremely user-friendly interface for metadata provision, the tool allows for both compliance with any XML or text-based metadata format as well as pluggability of heterogeneous RDF-based resources (made available as SPARQL [184] endpoints) as the reference data sources for providing auto-completion functionalities. This feature allows for the integration of a broad range of third-party data structures (e.g., code lists, controlled vocabularies, gazetteers, and registries) in the Web of Data.

Field values can also be generated on demand, can duplicate the content of another field, and even use generic XPath functions in order to mix-and-match values taken from the output XML document. Finally, this output document can be fed into an arbitrary chain of XSLT transformations (e.g., to generate a text-based output, such as JSON). All these functionalities are governed by a *template*, expressed in XML, that regulates production of the output document, defines the external data sources to be accessed via SPARQL, etc. Please refer to [185] for a comprehensive description of the template language.

Addressing semantic augmentation of metadata at editing-time (i.e., ex-ante) leaves an enormous amount of resource descriptions not featuring this important characteristic. As a consequence of this, important capabilities enabled by semantically enriched metadata (e.g., multilingualism, query expansion) could not be implemented by geoportals in discovery workflows. Then, we started working on offline, ex-post semantic lift of metadata records and realized it was possible to employ templates the other way around to search traditional XML metadata for correspondences in RDF data sources. The resulting application, named Liftboy, is described in [186] and made available on GitHub (<https://github.com/IREA-CNR-MI/liftboy-python> accessed on 1 April 2021) in its newer, improved implementation.

As a final note, we want to stress the importance of semantic characterization of metadata. Typically, this is seen as a solution to semantic heterogeneity and an opportunity for applying query expansion in information retrieval (in [186] the authors provide examples for both of these). In our opinion, semantic metadata can serve a higher purpose, that of “normalizing” resource description by conflation into a kind of pointer instead of repeatedly duplicating metadata property values (such as keywords, names, e-mail addresses of people, etc.) that frequently lead to inconsistencies, a practice we named *metadata delegation* [187]. It would be easier if all references to a keyword provided by a well-known controlled vocabulary were tagged with a unique identifier for that term (the URI of a skos:Concept [170]), if all references to a researcher pointed to her FOAF record [172], creating a web of decentralized metadata.

### 2.5.3. Exploiting Non-Rdf Data Structures for Semantic Metadata Creation

This case study builds on a software named SPARQL-Generate (<https://ci.mines-stetienne.fr/sparql-generate/> accessed on 1 April 2021) [188] that extends the syntax of SPARQL 1.1 [189] with constructs that allow for extracting data from heterogeneous data structures and generating RDF descriptions. The application to the geospatial domain we describe is production of metadata for *samples* (also called *specimens*) in the International Geo Sample Number (IGSN) format [190]. The target data structures are the entities made available by the European Long Term Ecological Research Network (eLTER) in its Sites and Data Registry (DEIMS-SDR) [191,192] (specifically, the entities representing *activities*, *sites*, and *sensors*).

We wanted to build on EDI, the metadata editor presented in the previous Section, but the originating sources are in JSON format and thus could not be directly integrated in the autocompletion functionalities provided by the former. We then decided to create RDF descriptions as signpost for the aforementioned entities and relate samples to them by plugging-in these RDF “semantic twins” in a custom EDI template. Then, the metadata maintainer can access the HTML5 interface generated by the EDI client and select the entities in the originating data structures via the many widgets made available by the software, drawing information from external data structures.

#### 2.5.4. Semantic Mediation for FAIR Access to Resources

This case study considers the articulation of geospatial discovery as a web API in order to make catalogs accessible by automated agents. One may argue that the Catalogue Service for the Web (CSW) by OGC [193] serves this purpose and, of course, when the automated agent knows where the endpoint is and which protocol to use, resource harvesting and search are straightforward. Still, when the agent only knows the homepage of the data provider and no information on the protocol applying, these operations may get difficult to achieve.

The problem (and the link to the subject of this paper, i.e., semantics) is that the Web, as experienced by human agents, is unlike web APIs in that there is a *semantic gap* to be bridged [194] before machines can fully participate. Overcoming this gap requires internalizing the key principles of REST (REpresentational State Transfer) as expressed by Roy Fielding in his Ph.D. dissertation [195]; specifically:

1. identification of resources
2. self-descriptive messages
3. hypermedia as the engine of application state

Please refer to Chapter 5 of the dissertation for an explanation of these. The attentive reader may already have spotted how the breadth of this research topic can be extended so as to encompass FAIR (Findable, Accessible, Interoperable, and Reusable) practices [175].

Since their inception, the FAIR principles have been deeply rooted in the notion of machine-actionability. Among the technologies for a machine-actionable Web, it is generally acknowledged that, despite the apparent differences, there is a broad overlapping between REST principles and FAIR practices (FORCE11 Guiding Principles for Findable, Accessible, Interoperable and Re-usable Data Publishing: <https://www.force11.org/fairprinciples> 1 April 2021). In fact [196], the machine-actionable behaviors of REST match the requirements of (at least) the first three letters in “FAIR”, as both recur to specification of semantics for their enactment and both rely on resolvable identifiers.

In order to achieve machine-actionability for geospatial services, the European Plate Observing System research infrastructure [197,198] exploits Hydra [199], an RDF vocabulary that is capable of expressing the mechanics of APIs in a way that is both intelligible to automated agents and also semantically rich. Please refer to the Hydra Core Vocabulary (<https://www.hydra-cg.com/spec/latest/core> 1 April 2021) for a more thorough descriptions of the features of this formalism.

The potential of this characterization of APIs is apparent. As an example, search for processing services matching a given set of parameters, such as the Normalized Difference Vegetation Index (NDVI) for a specific bounding box can greatly take advantage of semantics-aware service description [200]. Moreover, automated workflow composition on the basis of more precisely defined inputs and outputs can be easier than with other technologies [201].

#### 2.6. Powerful Geosemantics

There are concepts and relationships in the real world that are intrinsically imprecise and fuzzy, due to their gradual nature. This characteristic is particularly evident in the geographic context, in which natural entities and spatio-temporal phenomena are characterized by blurred and time-varying contours. For instance, it is impossible to encode in a

classic ontology based on OWL vague concepts like “most streets in Naples center are very narrow”, which involve some fuzziness for which a crisp definition does not make sense. What is the size of a street that makes it “narrow”? This is a matter of degrees depending on a subjective interpretation and, certainly, there is not a crisp transition between a street being *large* and *narrow* that may be agreed upon by all observers. The term *most* means that there are exceptions, i.e., a few streets are large, but its hard to quantify a crisp percentage. Furthermore there may be cases in which one needs to define a fuzzy concept hierarchy, a fuzzy taxonomy, in which a class is a specialization to a degree of several super classes such as “In Italy churches, beside being (1) places of worship, are often (0.8) historical buildings”. Furthermore, it may be necessary to define fuzzy relationships between concepts such as in “bell towers are very close to churches”.

Another possible source of imperfection occurs when an ontology is used for quality assurance to tag observations such as in Citizen Science (CS) projects. Such projects are at present a common practice to collect geospatial data in many domains such as natural sciences by involving volunteers to create georeferenced observations of objects of interest. A volunteer may be not completely sure about his/her observation, which is the case of epistemic uncertainty. This may happen because (s)he 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 precise.

Finally, there are more complex situations that may involve both ill-defined knowledge and epistemic uncertainty [202].

To cope with the above issues, powerful semantics approaches are needed which “extend” classic ontologies with the ability to represent and manage uncertainty and imprecision: To this end, the literature proposes soft ontologies [12]. In particular, there are three main groups defined on the basis of the probabilistic, the fuzzy, and the possibilistic or evidential frameworks. They have been adopted for extending propositional logic with probability, possibility, belief, or truth of a statement.

Fuzzy ontologies have been defined to model ill-defined knowledge with several purposes, depending on the kind of imperfection they need to represent and manage in the application [202]. Although a standard representation of a fuzzy ontology is still to come, a lot of researches have fuzzified the existing Description Logics (DL) and have defined fuzzy DL reasoners. The most up-to-date and complete fuzzyDL ontology reasoner has been proposed in [203].

To model epistemic uncertainty, fuzzy ontologies have been defined within a possibilistic framework that deals with certainty and possibility degrees of truth thus modeling the epistemic uncertainty characterizing experts’ subjective knowledge and the evaluation of the certainty of this knowledge. To this end, several possibilistic DL reasoners have been defined [204], which allow for representing and reasoning on uncertain statements such as “It is possible that this town is an Historic Area”. To this end, each concept, relation, and axiom is associated with a real value  $u$  in  $(0, 1]$  representing its certainty level.

Nevertheless, fuzzy ontologies do not allow to model the time varying nature of concepts and their context-dependent meaning. Specifically, most geographic concepts are represented by prototypes that vary with time: The prototypical modern city to an Italian person has changed during centuries, and it different for Chinese people. Fuzzy set theory cannot completely model how humans use concepts, in particular the fact that their meaning is influenced by context and states that vary with human knowledge in time. To this end, the framework known as state-context-property (SCOP) based on quantum mechanics [205] has been defined to map elements taken from operational foundations of quantum mechanics (like states, measurements, and observables) onto concepts and contexts.

In the following Subsections, we recap three case studies exploiting powerful semantics. Their synoptic view is reported in Table 3. They have been selected as representative of distinct application domains such as the creation of biodiversity observations (Section 2.6.1), remote sensing to aid disaster management (Section 2.6.2), and dynamic urban planning (Section 2.6.3). The first two of them exploit a fuzzy ontology encoding epistemic uncertainty of volunteers

when creating georeferenced observations (i.e., VGI) and the vague and incomplete knowledge of experts when interpreting a phenomenon from remote sensing evidence, respectively. By representing epistemic uncertainty and vagueness of knowledge, it is possible to model the distinct quality of the results of a decision process.

The last case study illustrates the application of the SCOP framework to model retrieval of maps within a GIR with increasing precision, achieved by exploiting the varying states of knowledge of user needs.

#### 2.6.1. A Fuzzy Ontology to Support Volunteered Geographic Information Creation and Search

Within the *Space4agri* [206] project, agronomists surveyed agronomic fields by tagging the observed crops and their phenological growth stages based on an agronomic ontology [207]. In this process, texts or pictures were added to report a difficulty or doubt of the agronomists when selecting a phenological growth stage from the ontology. This is due to different reasons:

- doubt in interpreting the meaning of the descriptions in the ontology;
- difficulty to distinguish the characteristic aspects of a phenological stage in the observed crop sample, because of a deficiency of the observation means (e.g., a far point of view);
- hesitancy to select a unique growth stage for several observed crop samples close in space within the same parcel, because of variability of their characteristics.

This suggested the need for extending the classic ontology-based reasoning by representing the epistemic uncertainty of the agronomists in creating VGI items (i.e., when selecting tags from the ontology [207]). Specifically, volunteers can create georeferenced annotations of crops they are observing in situ with the support of a fuzzy ontology. They are bound to select linguistic predicates, possibly fuzzy, to tag the observed crops and with each selected predicate they can associate a degree  $d$  in  $[0,1]$  representing the overall deficiency of their observation. This way, they can represent epistemic uncertainty due to both limitations of the means of observation (e.g., a far point of view, low resolution of the means of observation) and difficulty of precisely quantifying some properties of the observed crops. The linguistic predicates such as “crop has large leaf”, “crop has long stamen”, “crop has many branches” describe possibly fuzzy properties of the distinct kinds of crops: For example, a rice crop during its germination can appear with “elongated and thin branches” and “very small seeds”. The semantics of these linguistic predicates can be defined by level-1 fuzzy sets (whose membership degrees are numeric in the range  $[0,1]$ ). The fuzzy ontology can then explicitly represent linguistic concepts in both symbolic form (encoded by the linguistic terms “large”, “long”, “many”) and quantitative form. The latter is expressed by the membership functions defined on the numeric domains of the properties: For example, “large” is defined with a membership function on numeric values in cm. In the fuzzy ontology, compatibility between linguistic predicates is represented by Level-2 fuzzy relations, i.e., fuzzy sets on multidimensional basic domains whose membership degrees are not numbers but linguistic values. Fuzzy relations between linguistic predicates are used to perform approximate reasoning in the fuzzy ontology to automatically classify the crops, possibly into distinct types with different membership degrees. The defect degrees are interpreted as minimum thresholds, i.e., uncertainty levels, on the compatibility degrees between the linguistic predicates so that the final membership to a type of crop is modified by epistemic uncertainty. When formulating queries to the database of georeferenced crop observations, for example, requesting to map “rice crop fields”, the stored observations can be mapped Onto different shades of color depending on their membership degrees to type “rice crop”, thus accounting for both fuzziness and observation uncertainty.

**Table 3.** Dimensions of case studies (Powerful geosemantics).

Subsection	Name	Description	Input Data	Method Classification	Output Data	Potential Use
Section 2.6.1	Fuzzy ontology supporting VGI	VGI quality assurance and assessment by modeling both imprecision/vagueness of domain knowledge and uncertainty of volunteer's perceptions	VGI items created by selecting linguistic predicates from a fuzzy ontology and by associating uncertainty of observations	fuzzy computational intelligence	VGI quality assessment based on qualitative reasoning (level-based uncertainty reasoning)	quality assurance and assessment; ontology enrichment
Section 2.6.2	Environmental status indicator mapping	fuzzy classification of standing water from remote sensing images	remote sensing images and in situ observations plus incomplete (fuzzy) ontology (contributing factors and their (soft) constraints to derive partial evidence of watered water)	fuzzy computational intelligence+ machine learning	identification of watered areas and the fuzzy ontology enrichment	monitoring water bodies; ontology enrichment
Section 2.6.3	Modeling user interaction in GIR	modeling user intention and concepts' status	State-Context-Property representation of concepts and user queries	context-sensitive measurement of conceptual distance	identification of collapsed states representing answers to the user intention	modeling evolving and context dependent geographic concepts



### 2.6.2. Fuzzy Ontology to Support Remote Sensing Image Interpretation

In remote sensing, Geographic Object-Based Image Analysis (GEOBIA) groups techniques aiming at segmenting and classifying objects and phenomena (represented by groups of pixels sharing common properties) in satellite images based on image analysis procedures that rely on a priori expert knowledge [35]. In recent years, application of ontologies encoding experts' knowledge is emerging [14]. Ontologies are used to associate some perceived concepts with their data representation [35]. A widely applied approach to detect the geographic footprint of environmental phenomena is to compute spectral indexes (SI) maps. SI values integrate reflectance measurements at different wavelengths into a synthetic feature that can highlight some perceived aspects of the phenomenon in each pixel. SI maps are then segmented to identify target phenomena, such as vegetation presence and vigor (biomass presence, Leaf Area Index, Chlorophyll content, etc.), bare soil condition, and soil properties composition, burned areas, water presence, and so on. The segmentation consists of thresholding the pixel SI values by different thresholds specified in the ontology to define the different environmental phenomena.

Nevertheless, using the same ontology to segment a given phenomenon such as "green areas" in a new image may cause inaccuracies with many omissions and commission errors, since the value of the threshold must be tuned depending on several factors, such as the context and observation conditions. In fact, accurate calibration is needed to set a proper threshold for each study area. Thus uncertainty and imprecision must be represented since the kind of knowledge is perceptual by very nature [35]. These are the reasons why powerful semantics approaches are appealing. In fact, these techniques allow for explicit representation of perceptual characteristics of phenomena in images by means of fuzzy ontologies. Thus, they can cope with the limitations of both traditional GEOBIA solutions using ontologies and machine learning techniques requiring huge amounts of training data often unavailable.

In [162], an approach based on powerful semantics was proposed to map standing water areas from optical multispectral remote sensing images. Ill-defined knowledge of experts on the perceptual characteristics of standing water within optical images is represented by defining fuzzy sets on spectral indexes identified as features. The membership functions of these fuzzy sets relax the crisp segmentation thresholds defined in the vast literature on standing water mapping so as to tolerate imprecision and uncertainty. A fuzzy ontology is thus defined describing standing water in terms of fuzzy sets on spectral indexes. For each spatial unit with given values of spectral indexes, partial evidence degrees of standing water are computed by evaluating the membership degrees to the fuzzy sets in the fuzzy ontology. Finally, the partial evidence degrees in each spatial unit are combined by applying a fuzzy aggregation operator, learnt by a shallow machine learning algorithm trained on a small reference data set. Beside not requiring big training data, the approach offers the advantage of explicating the criteria used to map standing water, allowing discovering how many spectral indexes, which of them, and to which extent they contributed to map standing water in each spatial unit. The fuzzy ontology with new fuzzy relationships between fuzzy concepts.

### 2.6.3. State-Context-Property Framework to Model Human Interaction within a Geographic Information Retrieval System

According to [208], human-computer interaction is based on the exchange of words (or graphical tokens on maps) which are interpreted in the context of the conversation. The words used may originally have a broad meaning; through conversation the context becomes more precise and the concepts obtain more specific meanings. The authors present a proof of concept that shows the selection of several predetermined map types (e.g., street map, political map, map for hiking, ski routes) in a GIR by formalizing their approach in SCOP [205]. Specifically, SCOP is applied to predict an answer to the question: "Which map is appropriate for a given context?" where the context is declined as the intended purpose of the user.

A concept and a context serve as input parameters to the inference model that calculates the collapsed state and returns it. In this collapsed state, probability values for prototypes of the concept can be calculated. A use case is illustrated, in which a user states to a GIR query interface that she needs a map, without stating the kind of map. So far the concept “map” is in ground state, where all maps have some non-zero probability to be relevant. The user then states the intended usage that is to go on a bicycling trip. Now the state of the concept “map” collapses into a bicycling map. The user interaction may continue to indicate the region where the trip is planned, and this new information further restricts the map to an area. The application of SCOP is still at its early stage; it needs further developments and investigations to be practically applied, but its potential is great as far as prototypical modeling of contexts and states is concerned.

### 3. Results and Discussion

To organize the material, we started from the notion of semantics as a function that maps the world of syntax onto the world of meaning, in analogy with the studies on denotational semantics [209]. Once put on these lenses, we analyzed the presented case studies considering the original information they dealt with (syntactic objects with a certain amount of semantics), the meaning that is extracted and formalized (the new semantic objects), and the techniques that are applied to map the former onto the latter (the incremental semantic mapping function). This analysis of the case studies is presented in Table 4, where each row resumes one of them.

The first two columns identify the case study by indicating the corresponding subsection and a short name. In the columns that follow, one can find information about the mapping of the input information onto the new semantic objects: Specifically, column 3 contains the description of the input information pertaining the case study; column 4 provides the incremental semantic function that is used to map the original information with partial semantics onto the output with augmented meaning; finally, column 5 indicates the final information, i.e., the semantic domain of the case study. Column 6 indicates the delta between the input and output information; finally, column 7 enumerates the keywords, among those on the right side of Figure 2, that can be related to the case study: The more relevant keywords are in bold font and are assigned a weight  $w = 2$  in the analysis that follows.

Here, complexity degree is intended as the level at which semantics is made explicit in either the input or the output data structures considered by the specific case study. Specifically, the complexity degree is an integer in the range 1–7, following the principle of indiscernibility of Miller [210]. The general criterion for attributing this value is that complexity lower than 4 accounts for objects presenting scarce or no machine understandable information about their meaning; values between 4 and 7 indicate that meaning is more and more machine understandable and processable. For instance, the most simple case is that of unstructured text (complexity = 1), such as in case study 3.1 where input is constituted by free text keywords. The degree increases when more information is added such as in case study 3.2 and 3.3 (complexity = 2) where input is enriched both by the presence of structure (JSON documents) and by geographic coordinates. When the previous information is further augmented, complexity increases (complexity = 3) such as in the output data of case study 3.1 where uncertainty degrees are added. The next step in explication of semantics may involve schema information or categorization of data (complexity = 4). Then, when relationships among the entities (topological, order, metric, broader/narrower) are taken into account, complexity increases to 5. Complexity is 6 when vague and uncertain concepts and relationships are represented. Finally, when information can be generated by approximate reasoning or has fully reached semantic interoperability, complexity is 7.

Table 4. Dimensions of case studies.

Subsection	Name	Complexity Degree of Input Data Explicit Semantics	Incremental Semantic Function (Methods, Techniques)	Complexity Degree of Output Data Explicit Semantics	Added Semantic Value	Keywords in Geosemantics (cfr.Sankey Diagram in Figure2)
Section 2.4.1	GeoFinder	1	heuristic rules and explicit geographic information in gazetteer	3	2	<b>Geographic Information Retrieval</b> ; semantic enrichment/tagging/annotation; <b>gazetteer</b> ; <b>geoparsing</b> ; <b>geonames</b> ; <b>place-based information systems</b> ; geo-comparison; extraction of spatio-temporal information
Section 2.4.2	Events spatio-temporal footprint	2	clustering	4	2	semantic enrichment/tagging/annotation; <b>geo-comparison</b> ; <b>extraction of spatio-temporal information</b> ; <b>events</b> , <b>change discovery</b> ; thematic/spatial and temporal perspective
Section 2.4.3	Tour miner	2	knowledge-based clustering	5	3	<b>geo-comparison</b> ; <b>folksonomies</b> ; semantic enrichment/tagging/annotation; <b>extraction of spatio-temporal information</b> ; <b>events</b> , <b>change discovery</b> ; thematic/spatial and temporal perspective; <b>(Geographic) Knowledge graph</b> ; geospatial statistics
Section 2.4.4	Eliciting Geographic Gazetteer	2	statistic analysis and clustering	4	2	<b>Geographic Information Retrieval</b> ; semantic enrichment/tagging/annotation; <b>gazetteer</b> ; <b>place-based information systems</b> ; <b>folksonomies</b> ; geo-comparison; extraction of spatio-temporal information; Place location, identity, meaning; geospatial statistics; top-level ontologies
Section 2.5.1	Holistic map representation with Geographic Scenarios	4	SWRL rules combining relational data and ontologies	5	1	<b>Ontology for modeling</b> ; Domain ontologies; <b>Ontology-based info extraction</b> ; Semantic enrichment/tagging/ annotation; Thematic, spatial, and temporal perspectives; <b>Knowledge representation languages (OWL)</b>
Section 2.5.2	EDI/Liftboy	5	entity annotation. Specific metadata profile specifying where and how to find the Semantic Web (RDF) resources	7	2	Domain ontologies; <b>Linked Geo Data/LOD/Linked Sensor Data</b> ; Gazetteers (GeoNames)/temporal gazetteers; Spatial RDF and SPARQL; <b>Ontology-based info extraction</b> ; <b>Semantic enrichment/tagging/ annotation</b> ; Knowledge representation languages (OWL); <b>Semantics-driven user interfaces/interaction paradigms/Semantic engineering of human communications</b>
Section 2.5.3	Semantic twins	4	same as example EDI/Liftboy, but to grant consistency of metadata items the semantic twins of JSON entities is exploited	7	3	Domain ontologies; <b>Linked Geo Data/LOD/Linked Sensor Data</b> ; Gazetteers (GeoNames)/temporal gazetteers; <b>Spatial RDF and SPARQL</b> ; <b>Ontology-based info extraction</b> ; <b>Semantic enrichment/tagging/ annotation</b> ; Knowledge representation languages (OWL)
Section 2.5.4	Semantic mediation for FAIR access to resources	5	information retrieval based on an application ontology	7	2	Ontology for encoding; Application ontologies; Geographic Information Retrieval; <b>Linked Geo Data/LOD/Linked Sensor Data</b> ; <b>Semantic enrichment/tagging/ annotation</b> ; Knowledge representation languages (OWL); Semantics-driven user interfaces/interaction paradigms/Semantic engineering of human communications; <b>Semantic markups for Web services</b>
Section 2.6.1	Fuzzy ontology supporting VGI	6	fuzzy rule based inference engine	7	1	<b>task-ontology</b> ; <b>ontology for modeling</b> ; <b>ontology for encoding</b> ; sensor and observation ontology; <b>qualitative reasoning</b> ; <b>application ontologies</b> ; ontology design pattern; ontology-based information extraction; semantics-driven user interfaces
Section 2.6.2	Environmental status indicator mapping	5	incomplete (fuzzy) ontology + machine learning exploiting in situ classified data	7	2	task ontology; <b>ontology for modeling</b> ; <b>ontology for encoding</b> ; <b>qualitative reasoning</b> ; <b>application ontologies</b> ; ontology design pattern; ontology-based information extraction
Section 2.6.3	Modeling user interaction in GIR	5	inference in SCOP framework	7	2	<b>conceptual space</b> ; <b>semantic engineering of human communications</b> ; <b>qualitative reasoning</b> ; Non-representational formalisms

The four case studies presented in Section 2.4 share the same type of input geoinformation, which is essentially not explicit, being dispersed within unstructured and loosely structured texts. In the output geoinformation of these case studies, semantics is made explicit but not always in a standard, interoperable format; because of this, it may be difficult or even impossible to reuse the results in different contexts.

The first case study in Section 2.5 portrays a model exploiting semantics at its full potential, via ontologies. The second applies to semi-structured geoinformation in the form of metadata, possibly compliant with OGC standards. The third case study involves structured (JSON) and semi-structured (HTML) information that lacks the relations between the entities involved (e.g., between descriptions of sensors and the corresponding points of contact) and, in general, can not be easily reused in a Web of Data context. Finally, the fourth case study applies to unstructured information intended to the human agent (i.e., the specification of computer interaction protocols). For each of these, the output is information that can be shared and reused in an interoperable way by enabling querying and retrieval in a Linked Data perspective. The first two case studies in Section 2.6 involve explicit and rich geoinformation in the form of soft ontologies, while the last case study uses the SCOP formalism. All these case studies enable qualitative and approximate reasoning to deduce novel geoinformation automatically.

A preliminary observation that can be made is that complexity of the inputs is lower for the case studies in Section 2.4, medium for those in Section 2.5, and maximum for the case studies in Section 2.6. The same for the outputs. More insights come from cross-referencing of the case studies and the keywords listed on the right of Figure 2, yielding the representation in Figure 3. This last figure illustrates the weighted associations between case studies and keywords: Case studies within the same Section (i.e., associated with the same form of geosemantics) are characterized by shades of the same color (yellow for implicit, blue for formal, and grey for powerful geosemantics). On the x axis, the length of the bar represents the different importance of the method/technique in the case study while the pair hue-color uniquely identifies both the case study and its belonging semantic category. It can be visually noticed that the case studies classified in the same form of geosemantics are mostly associated with distinctive keywords. For example, the case studies in Section 2.6 (powerful geosemantics) are associated with “Non-representational formalisms”, “Task ontologies”, and “Qualitative reasoning”. Nevertheless, some keywords (e.g., “Semantic enrichment/tagging/annotation”) are associated with case studies classified in “adjacent” forms of semantics.

To confirm the conjecture suggested by Figure 3, i.e., that the three geosemantics forms are good categorizations for the keywords, we also computed the similarity measure known as Jaccard coefficient between any pair of case studies on the basis of the aforementioned weighted keywords, as shown in Figure 4. The figure clearly shows that the intra-similarities (regarding pairs of case studies belonging to the same form of geosemantics, grouped within the colored rectangles) are greater than the inter-similarity degrees between pairs of case studies classified as different forms of geosemantics (i.e., appearing outside the colored rectangles).

It can be noticed that all case studies have greatest intra-similarity with another case study of the same geosemantics form. Only case studies in the yellow group share some inter-similarity with those of the blue group, which is anyway an order of magnitude lower than the intra-similarity. Specifically, as far as the case studies dealing with the implicit form of geosemantics are concerned, their overall intra-similarity, computed as percentage of shared keywords among all the case studies of the same category, reaches 54.3%, while their overall inter-similarity with any other case study of the others two categories is only 1.7%; as far as the case studies dealing with the explicit form are concerned, they have an overall intra-similarity of 58% and an overall inter-similarity of 2.6%; finally the case studies dealing with the powerful form have overall intra-similarity of 37% and an overall inter-similarity of only 0.9%. These findings confirm our hypothesis that the three forms

of semantics are characterized by distinguishing techniques, methods, and knowledge sources in the geospatial domain.

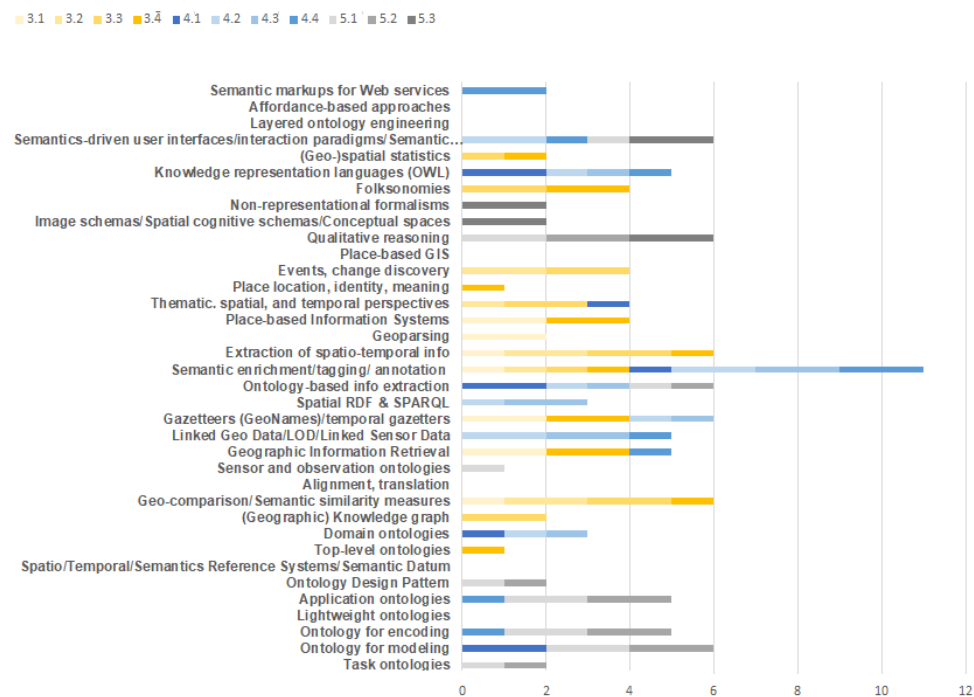


Figure 3. Case studies and the keywords representing their main activities and technologies.

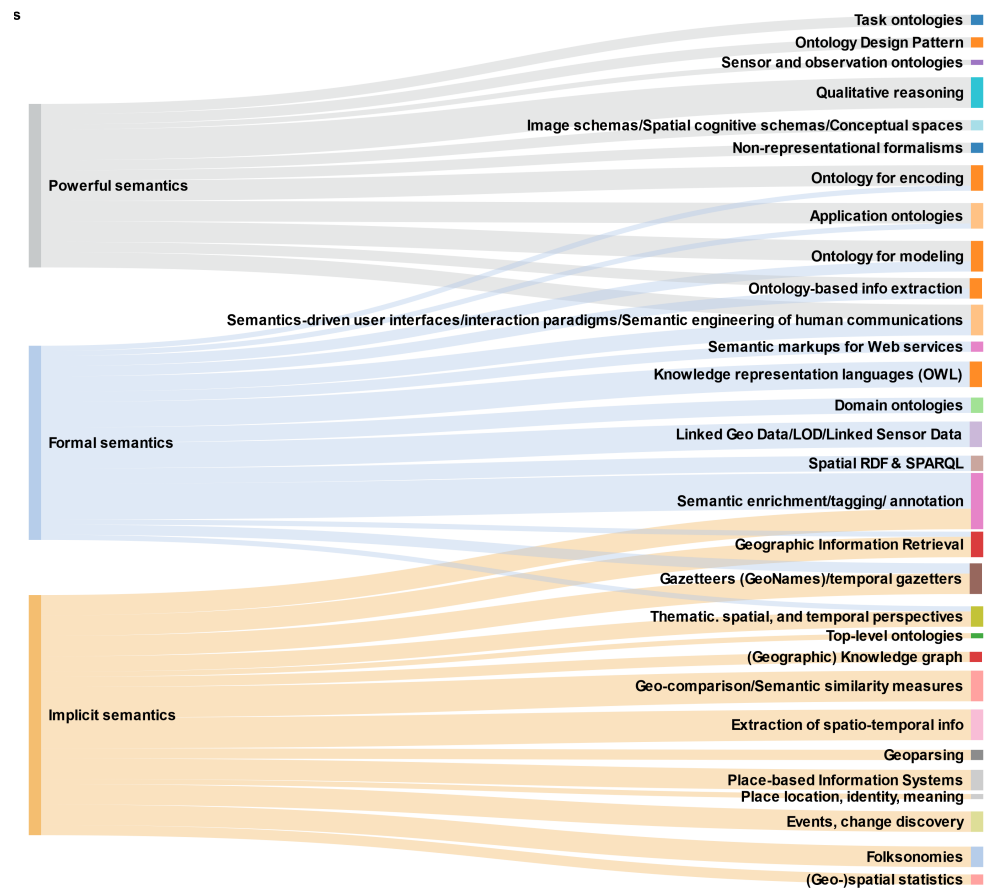
Similarity matrix between case studies (based on shared keywords)	3.1				4.1				5.1		
	3.1	3.2	3.3	3.4	4.1	4.2	4.3	4.4	5.1	5.2	5.3
3.1	1.00	0.19	0.14	0.13	0.03	0.03	0.03	0.02	0.00	0.00	0.00
3.2		1.00	0.57	0.14	0.03	0.03	0.03	0.03	0.00	0.00	0.00
3.3			1.00	0.27	0.03	0.03	0.03	0.03	0.00	0.00	0.00
3.4				1.00	0.05	0.04	0.03	0.03	0.00	0.00	0.00
4.1					1.00	0.25	0.24	0.09	0.00	0.00	0.00
4.2						1.00	0.75	0.24	0.00	0.00	0.00
4.3							1.00	0.25	0.00	0.00	0.00
4.4								1.00	0.15	0.10	0.00
5.1									1.00	0.11	0.23
5.2										1.00	0.12
5.3											1.00

Figure 4. Jaccard similarity between study cases represented as fuzzy sets of keywords

Besides revealing the distinguishing features of the geosemantics forms, we also found in this analysis that case studies related to implicit and formal semantics have many activities in common, identified by the shared keywords “Thematic spatial and temporal perspectives”, “semantic enrichment/tagging/annotation”, “Gazetteers (GeoNames)/temporal gazetteers”, and “Geographic Information Retrieval”. Formal and powerful semantics share “Semantics-driven user interfaces/interaction paradigms/...”, “ontology based information extraction”, “Application ontologies”, “Ontology for encoding”, and “Ontology for modeling”. This means that there is not a clear-cut partition between the forms of semantics. This shows that a “semantic continuum” is present, gently blending the groups, moving from implicit to powerful semantics. Conversely, the approaches related to powerful and implicit semantics share no keywords. These findings reveal that the ordering of categories introduced by Sheth [4] also seems to emerge from our analysis even in the context of geographic information.



Figure 5 provides an even more synoptic view on the relations between the keywords and the three forms of semantics, complementing Figure 2 with the findings described in this section. In fact, the figure clearly visualizes that, once the keywords are grouped according to the forms of semantics that are associated with the case studies presented in this paper, they are much more clustered. This means that patterns emerge in the geosemantics “forest”, thus making order among the diverse practices.



**Figure 5.** Comparison between the grouping of keywords in Figure 2 (on the right-hand side) and the grouping induced by the three forms of geosemantics (via the case studies) makes it apparent their greater distinguishing power.

Of course, this analysis can be enriched both by extending the meta-review to encompass more methodologies, techniques, and knowledge bases and by analyzing other case studies in the literature. Nevertheless we think that this contribution has the merit of setting a methodological workflow to characterize the forms of semantics in geoinformation and their preferred/elective approaches.

#### 4. Conclusions

This paper applied the categories of semantics defined by Sheth to the domain of geoinformation in order to orient the reader in problem solving. We first analyzed recent reviews and editorial papers on geosemantics, mining which are the main technologies, methodologies, research challenges, and solutions presented by the authors. Then, we discussed selected case studies for the implicit, formal, and powerful geosemantics, respectively. The two-step analysis culminates with cross-referencing these two sources in order to confirm that the three forms of geosemantics are characterized by distinguishing techniques, methods, and knowledge sources.

The subsistence of this conjecture is attested by the Jaccard distances computed between members of the same/different categories of semantics (see Figure 4). This can also be visually assessed by looking at Figures 3 and 5. In the latter, it is also apparent that there are fringe keywords associated with “adjacent” categories (i.e., categories with similar semantics explicitation degrees). This paper contributes to structuring the approaches to semantics in geoinformation, partitioning the semantic continuum suggested in [6] in discrete, distinguishing techniques and methods.

Further insight may come from categorizing in the three forms of semantics the papers considered in the meta-review (Section 2.3) according to the associated keywords. Future work will also investigate scaling-up of the workflow by applying content representation methods used in information retrieval. In fact, these can automatically identify the keywords from the text of the reviewed literature.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/article/10.3390/ijgi10050330/s1>.

**Author Contributions:** All authors equally contributed to the writing and revising of this paper. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

CS	Citizen Science
CSW	Catalogue Service for the Web
eLTER	European Long-Term Ecological Research
FAIR	Findable, Accessible, Interoperable, and Reusable
FOAF	Friend Of A Friend
FOL	First-Order Logic
GEOBIA	GEographic Object-Based Image Analysis
GIR	Geographic Information Retrieval
IGSN	International Geo-Sampling Number
IRS	Information Retrieval System
NDVI	Normalized Difference Vegetation Index
NER	Named Entity Recognition
PID	Persistent IDentifier
POI	Point Of Interest
REST	REpresentational State Transfer
ROI	Region Of Interest
SEM	Simple Event Model
SI	Spectral Indexes
WKT	Well-Known Text

## References

1. Guha, R.; McCool, R.; Miller, E. Semantic search. In Proceedings of the WWW '03, 12th international conference on World Wide Web, Budapest, Hungary, 20–24 May 2003; ACM Press: New York, NY, USA, 2003; pp. 700–709.
2. Duval, E.; Hodgins, W. Metadata principles and practicalities. *D-Lib Mag.* **2002**, *8*, 1–10. [CrossRef]
3. Noguera-Iso, J.; Muro-Medrano, P.; Zarazaga-Soria, F. *Geographic Information Metadata for Spatial Data Infrastructures: Resources, Interoperability and Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2005.
4. Sheth, A.; Ramakrishnan, C.; Thomas, C. Semantics for the Semantic Web: The Implicit, the Formal and the Powerful. *Int. J. Semant. Web Inf. Syst.* **2005**, *1*, 1–18. [CrossRef]
5. Uschold, M. Where are the semantics in the semantic web. *AI Mag.* **2003**, *24*, 25.
6. Almeida, M.; Rocha Souza, R.; Fonseca, F. Semantics in the Semantic Web: A Critical Evaluation. *Knowl. Organ. J.* **2011**, *38*, 187–203. [CrossRef]
7. Gärdenfors, P. How to make the Semantic Web more semantic. In *Formal Ontology in Information Systems*; Varzi, A., Vieu, L., Eds.; IOS Press: Amsterdam, The Netherlands, 2004; pp. 19–36.

8. Lemmens, M. Geo-information, Technologies, Applications and the Environment. In *Geotechnologies and the Environment (GEOTECH)*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 5, p. 349.
9. Ait-Ameur, Y.; Gibson, J.; Mery, D. On Implicit and Explicit Semantics: Integration Issues in Proof-Based Development of Systems. In Proceedings of the 6th International on Symposium Leveraging Applications of Formal Methods, Verification and Validation—ISO/FA 2014, Corfu, Greece, 8–11 October 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 604–618.
10. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [[CrossRef](#)]
11. Baader, F.; Calvanese, D.; McGuinness, D.; Nardi, D.; Patel-Schneider, P. *The Description Logic Handbook: Theory, Implementation, and Applications*; Cambridge University Press: Cambridge, UK, 2007.
12. Cagliani, M.; Fusco, G. Formal Ontologies and Uncertainty. In geographical knowledge. *TeMA J. Land Use Mobility Environ.* **2014**, *187*–198. [[CrossRef](#)]
13. Tambassi, T. (Ed.) *The Philosophy of GIS*; Springer International Publishing: Cham, Switzerland, 2019. [[CrossRef](#)]
14. Kokla, M.; Guilbert, E. A Review of Geospatial Semantic Information Modeling and Elicitation Approaches. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 146. [[CrossRef](#)]
15. Kuhn, W. Geospatial semantics: Why, of what, and how? In *Journal on Data Semantics III*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1–24.
16. Janowicz, K. Observation-Driven Geo-Ontology Engineering. *Trans. GIS* **2012**, *16*, 351–374. [[CrossRef](#)]
17. Kuhn, W. Semantic engineering. In *Research Trends in Geographic Information Science*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 63–76.
18. Kuhn, W. Core concepts of spatial information for transdisciplinary research. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 2267–2276. [[CrossRef](#)]
19. Hu, Y. 1.07—Geospatial Semantics. In *Comprehensive Geographic Information Systems*; Huang, B., Ed.; Elsevier: Oxford, UK, 2018; pp. 80–94. [[CrossRef](#)]
20. Janowicz, K.; Hitzler, P. The Digital Earth as knowledge engine. *Semant. Web J.* **2012**, *3*, 213–221. [[CrossRef](#)]
21. Shadbolt, N.; Smart, P. *Knowledge Elicitation: Methods, Tools and Techniques*; CRC Press: Boca Raton, FL, USA, 2015.
22. Kavouras, M.; Kokla, M. *Theories of Geographic Concepts: Ontological Approaches to Semantic Integration*; CRC Press: Boca Raton, FL, USA, 2007.
23. Casati, R.; Smith, B.; Varzi, A.C. *Ontological Tools for Geographic Representation*; IOS Press: Amsterdam, The Netherlands, 1998.
24. Hu, Y.; Janowicz, K. Enriching top-down geo-ontologies using bottom-up knowledge mined from linked data. *Adv. Geogr. Inf. Sci. Past Next Twenty Years* **2016**, *183*–198.
25. Hong, J.H.; Kuo, C. A semi-automatic lightweight ontology bridging for the semantic integration of cross-domain geospatial information. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 2223–2247. [[CrossRef](#)]
26. Gandon, F.L. A survey of the first 20 years of research on semantic Web and linked data. *Ingénierie des Systèmes d'Inf.* **2018**, *23*, 11–38. [[CrossRef](#)]
27. Giunchiglia, F.; Zaihrayeu, I. Lightweight Ontologies. In *Encyclopedia of Database Systems*; Springer: New York, NY, USA, 2009.
28. Gangemi, A.; Presutti, V. Ontology design patterns. In *Handbook on Ontologies*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 221–243.
29. Martinez-Rodriguez, J.L.; Hogan, A.; Lopez-Arevalo, I. Information extraction meets the semantic web: A survey. *Semant. Web* **2020**, *11*, 1–81. [[CrossRef](#)]
30. Allahyari, M.; Pouriyeh, S.; Assefi, M.; Safaei, S.; Trippe, E.D.; Gutierrez, J.B.; Kochut, K. A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *arXiv* **2017**, arXiv:1707.02919.
31. Monteiro, B.R.; Davis, C.A., Jr.; Fonseca, F. A survey on the geographic scope of textual documents. *Comput. Geosci.* **2016**, *96*, 23–34. [[CrossRef](#)]
32. Purves, R.S.; Clough, P.; Jones, C.B.; Hall, M.H.; Murdock, V. Geographic information retrieval: Progress and challenges in spatial search of text. *Found. Trends Inf. Retr.* **2018**, *12*, 164–318. [[CrossRef](#)]
33. Rajbhandari, S.; Aryal, J.; Osborn, J.; Musk, R.; Lucieer, A. Benchmarking the applicability of ontology in geographic object-based image analysis. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 386. [[CrossRef](#)]
34. Guilbert, É.; Moulin, B. Towards a Common Framework for the Identification of Landforms on Terrain Models. *ISPRS Int. J. Geo Inf.* **2017**, *6*, 12. [[CrossRef](#)]
35. Arvor, D.; Belgiu, M.; Falomir, Z.; Mougnot, I.; Durieux, L. Ontologies to interpret remote sensing images: Why do we need them? *GISci. Remote Sens.* **2019**, *56*, 911–939. [[CrossRef](#)]
36. Ballatore, A. Prolegomena for an Ontology of Place. In *Advancing Geographic Information Science*; GSDI Association Press: Needham, MA, USA, 2016.
37. Garbacz, P.; Lawrynowicz, A.; Szady, B. Identity Criteria for Localities. In Proceedings of the FOIS, Cape Town, South Africa, 19–21 September 2018.
38. Krisnadhi, A.; Hu, Y.; Janowicz, K.; Hitzler, P.; Arko, R.; Carbotte, S.; Chandler, C.; Cheatham, M.; Fils, D.; Finin, T.; et al. The GeoLink modular oceanography ontology. In Proceedings of the International Semantic Web Conference, Bethlehem, PA, USA, 11–15 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 301–309.
39. Gould, N.; Mackaness, W. From taxonomies to ontologies: formalizing generalization knowledge for on-demand mapping. *Cartogr. Geogr. Inf. Sci.* **2016**, *43*, 208–222. [[CrossRef](#)]

40. Yan, J.; Guilbert, É.; Saux, E. An ontology-driven multi-agent system for nautical chart generalization. *Cartogr. Geogr. Inf. Sci.* **2017**, *44*, 201–215. [[CrossRef](#)]
41. Varanka, D.; Usery, E.L. The map as knowledge base. *Int. J. Cartogr.* **2018**, *4*, 201–223. [[CrossRef](#)]
42. Janowicz, K.; Haller, A.; Cox, S.J.; Le Phuoc, D.; Lefrançois, M. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *J. Web Semant.* **2019**, *56*, 1–10. [[CrossRef](#)]
43. Auer, S.; Lehmann, J.; Hellmann, S. LinkedGeoData: Adding a Spatial Dimension to the Web of Data. In Proceedings of the International Semantic Web Conference, Chantilly, VA, USA, 25–29 October 2009.
44. Vatant, B.; Wick, M. GeoNames Ontology. 2012. Available online: <http://www.geonames.org/ontology/> (accessed on 1 April 2021)
45. J. Paul Getty Trust [Los Angeles, CA]. Getty Thesaurus of Geographic Names. [Software, E-Resource] Retrieved from the Library of Congress. 1999. Available online: <https://lcn.loc.gov/99483604> (accessed on 1 April 2021).
46. Derungs, C.; Purves, R. Mining nearness relations from an n-grams Web corpus in geographical space. *Spat. Cogn. Comput.* **2016**, *16*, 301–322. [[CrossRef](#)]
47. Ballatore, A. Extracting Place Emotions from Travel Blogs. In Proceedings of the AGILE, Helsinki, Finland, 25–29 May 2015.
48. Strapparava, C.; Valitutti, A. WordNet Affect: An Affective Extension of WordNet. In Proceedings of the LREC, Lisbon, Portugal, 26–28 May 2004.
49. Derungs, C.; Purves, R. From text to landscape: locating, identifying and mapping the use of landscape features in a Swiss Alpine corpus. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1272–1293. [[CrossRef](#)]
50. Wartmann, F.; Acheson, E.; Purves, R. Describing and comparing landscapes using tags, texts, and free lists: an interdisciplinary approach. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 1572–1592. [[CrossRef](#)]
51. Ilarri, S.; Stojanovic, D.; Ray, C. Semantic management of moving objects: A vision towards smart mobility. *Expert Syst. Appl.* **2015**, *42*, 1418–1435. [[CrossRef](#)]
52. Fileto, R.; May, C.; Renso, C.; Pelekis, N.; Klein, D.; Theodoridis, Y. The Baquara2 knowledge-based framework for semantic enrichment and analysis of movement data. *Data Knowl. Eng.* **2015**, *98*, 104–122. [[CrossRef](#)]
53. Han, L.; Kashyap, A.L.; Finin, T.W.; Mayfield, J.; Weese, J. UMBC-EBIQUITY-CORE: Semantic Textual Similarity Systems. In Proceedings of the \*SEM@NAACL-HLT, Atlanta, GA, USA, 13–14 June 2013.
54. Jiang, Y.; Li, Y.; Yang, C.; Liu, K.; Armstrong, E.; Huang, T.; Moroni, D.; Finch, C. A comprehensive methodology for discovering semantic relationships among geospatial vocabularies using oceanographic data discovery as an example. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2310–2328. [[CrossRef](#)]
55. Li, W.; Bhatia, V.; Cao, K. Intelligent polar cyberinfrastructure: Enabling semantic search in geospatial metadata catalogue to support polar data discovery. *Earth Sci. Inform.* **2015**, *8*, 111–123. [[CrossRef](#)]
56. Hu, Y. Geospatial Semantics. *Compr. Geogr. Inf. Syst.* **2018**, 80–94. [[CrossRef](#)]
57. Hakimpour, F.; Timpf, S. Using Ontologies for resolution of Semantic Heterogeneity in GIS. In Proceedings of the 4th AGILE Conference on Geographic Information Science, Brno, Czech Republic, 19–21 April 2001.
58. Fallahi, G.; Frank, A.; Mesgari, M.; Rajabifard, A. An ontological structure for semantic interoperability of GIS and environmental modeling. *Int. J. Appl. Earth Obs. Geoinf.* **2008**, *10*, 342–357. [[CrossRef](#)]
59. Fonseca, F.; Câmara, G.; Monteiro, A. A Framework for Measuring the Interoperability of Geo-Ontologies. *Spat. Cogn. Comput.* **2006**, *6*, 309–331. [[CrossRef](#)]
60. Brodaric, B. The design of GSC FieldLog: ontology-based software for computer aided geological field mapping. *Comput. Geosci.* **2004**, *30*, 5–20. [[CrossRef](#)]
61. Schuurman, N. Social Dimensions of Object Definition in GIS. In *Re-Presenting GIS*; Fisher, P., Unwin, D., Eds.; John Wiley & Sons: Hoboken, NJ, USA, 2005.
62. Baglioni, M.; Masserotti, M.V.; Renso, C.; Spinsanti, L. Building Geospatial Ontologies from Geographical Databases. In Proceedings of the GeoS, Mexico City, Mexico, 29–30 November 2007.
63. Scheider, S. Grounding Geographic Information in Perceptual Operations. In *Frontiers in Artificial Intelligence and Applications*; IOS Press: Amsterdam, The Netherlands, 2012.
64. Worboys, M.; Stewart, K. From Objects to Events: GEM, the Geospatial Event Model. In Proceedings of the GIScience, Adelphi, MD, USA, 20–23 October 2004.
65. Raskin, R.G.; Pan, M.J. Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Comput. Geosci.* **2005**, *31*, 1119–1125. [[CrossRef](#)]
66. Couclelis, H. Ontologies of geographic information. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 1785–1809. [[CrossRef](#)]
67. Hu, Y.; Janowicz, K.; Carral, D.; Scheider, S.; Kuhn, W.; Berg-Cross, G.; Hitzler, P.; Dean, M.; Kolas, D. A Geo-ontology Design Pattern for Semantic Trajectories. In Proceedings of the COSIT, Scarborough, UK, 2–6 September 2013.
68. Carral, D.; Scheider, S.; Janowicz, K.; Vardeman, C.; Krisnadhi, A.A.; Hitzler, P. An Ontology Design Pattern for Cartographic Map Scaling. In Proceedings of the ESWC, Montpellier, France, 26–30 May 2013.
69. Cruz, I.; Sunna, W.; Makar, N.; Bathala, S. A visual tool for ontology alignment to enable geospatial interoperability. *J. Vis. Lang. Comput.* **2007**, *18*, 230–254. [[CrossRef](#)]
70. Goodchild, M. *Formalizing Place in Geographic Information Systems*; Springer: New York, NY, USA, 2011.
71. Goodchild, M.; Hill, L. Introduction to digital gazetteer research. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 1039–1044. [[CrossRef](#)]



72. Alani, H.; Jones, C.; Tudhope, D. Voronoi-based region approximation for geographical information retrieval with gazetteers. *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 287–306. [[CrossRef](#)]
73. Rice, M.T.; Aburizaiza, A.O.; Jacobson, R.; Shore, B.M.; Paez, F.I. Supporting Accessibility for Blind and Vision-impaired People With a Localized Gazetteer and Open Source Geotechnology. *Trans. GIS* **2012**, *16*, 177–190. [[CrossRef](#)]
74. Schlieder, C.; Vögele, T.; Visser, U. Qualitative Spatial Representation for Information Retrieval by Gazetteers. In Proceedings of the COSIT, Morro Bay, CA, USA, 19–23 September 2001.
75. Janowicz, K.; Keßler, C. The role of ontology in improving gazetteer interaction. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 1129–1157. [[CrossRef](#)]
76. Davies, C.; Holt, I.; Green, J.; Harding, J.; Diamond, L. User Needs and Implications for Modelling Vague Named Places. *Spat. Cogn. Comput.* **2009**, *9*, 174–194. [[CrossRef](#)]
77. Hollenstein, L.; Purves, R. Exploring place through user-generated content: Using Flickr tags to describe city cores. *J. Spat. Inf. Sci.* **2010**, *1*, 21–48.
78. Grothe, C.; Schaab, J. Automated Footprint Generation from Geotags with Kernel Density Estimation and Support Vector Machines. *Spat. Cogn. Comput.* **2009**, *9*, 195–211. [[CrossRef](#)]
79. Keßler, C.; Maué, P.; Heuer, J.T.; Bartoschek, T. Bottom-up gazetteers: Learning from the implicit semantics of geotags. In Proceedings of the International Conference on GeoSpatial Semantics, Mexico City, Mexico, 3–4 December 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 83–102.
80. Li, L.; Goodchild, M.F. Constructing places from spatial footprints. In Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information, Redondo Beach, CA, USA, 7–9 November 2012; pp. 15–21.
81. Gao, S.; Li, L.; Li, W.; Janowicz, K.; Zhang, Y. Constructing gazetteers from volunteered Big Geo-Data based on Hadoop. *Comput. Environ. Urban Syst.* **2017**, *61*, 172–186. [[CrossRef](#)]
82. Uryupina, O. Semi-supervised learning of geographical gazetteers from the internet. In Proceedings of the HLT-NAACL 2003, Edmonton, AB, Canada, 27 May–1 June 2003.
83. Zhu, R.; Hu, Y.; Janowicz, K.; McKenzie, G. Spatial signatures for geographic feature types: examining gazetteer ontologies using spatial statistics. *Trans. GIS* **2016**, *20*, 333–355. [[CrossRef](#)]
84. Samal, A.; Seth, S.; Cueto, K. A feature-based approach to conflation of geospatial sources. *Int. J. Geogr. Inf. Sci.* **2004**, *18*, 459–489. [[CrossRef](#)]
85. Sehgal, V.; Getoor, L.; Viechnicki, P. Entity resolution in geospatial data integration. In Proceedings of the GIS '06, Arlington, VA, USA, 10–11 November 2006.
86. Hastings, J. Automated conflation of digital gazetteer data. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 1109–1127. [[CrossRef](#)]
87. Jones, C.; Purves, R. Geographical information retrieval. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 219–228. [[CrossRef](#)]
88. Jones, R.; Zhang, W.; Rey, B.; Jhala, P.; Stipp, E. Geographic intention and modification in web search. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 229–246. [[CrossRef](#)]
89. Sanderson, M.; Kohler, J. Analyzing geographic queries. In Proceedings of the SIGIR Workshop on Geographic Information Retrieval, Sheffield, UK, 25–29 July 2004.
90. Janowicz, K.; Raubal, M.; Kuhn, W. The semantics of similarity in geographic information retrieval. *J. Spat. Inf. Sci.* **2011**, *2*, 29–57. [[CrossRef](#)]
91. Hu, Y.; Janowicz, K.; Prasad, S. Improving wikipedia-based place name disambiguation in short texts using structured data from DBpedia. In Proceedings of the GIR '14, Dallas, TX, USA, 4 November 2014.
92. Cucerzan, S. Large-Scale Named Entity Disambiguation Based on Wikipedia Data. In Proceedings of the EMNLP-CoNLL, Prague, Czech Republic, 28–30 June 2007.
93. Overell, S.; Rüger, S. Using co-occurrence models for placename disambiguation. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 265–287. [[CrossRef](#)]
94. Leidner, J.L. *Toponym Resolution in Text: Annotation, Evaluation and Applications of Spatial Grounding of Place Names*; Universal-Publishers: Irvine, CA, USA, 2008.
95. Ju, Y.; Adams, B.; Janowicz, K.; Hu, Y.; Yan, B.; McKenzie, G. Things and Strings: Improving Place Name Disambiguation from Short Texts by Combining Entity Co-Occurrence with Topic Modeling. In Proceedings of the EKAW, Bologna, Italy, 19–23 November 2016.
96. Gelernter, J.; Balaji, S. An algorithm for local geoparsing of microtext. *GeoInformatica* **2013**, *17*, 635–667. [[CrossRef](#)]
97. Vasardani, M.; Winter, S.; Richter, K.F. Locating place names from place descriptions. *Int. J. Geogr. Inf. Sci.* **2013**, *27*, 2509–2532. [[CrossRef](#)]
98. Hu, Y.; Janowicz, K.; Prasad, S.; Gao, S. Metadata Topic Harmonization and Semantic Search for Linked-Data-Driven Geoportals: A Case Study Using ArcGIS Online. *Trans. GIS* **2015**, *19*, 398–416. [[CrossRef](#)]
99. Li, W.; Goodchild, M.F.; Raskin, R. Towards geospatial semantic search: exploiting latent semantic relations in geospatial data. *Int. J. Digit. Earth* **2014**, *7*, 17–37. [[CrossRef](#)]
100. Amitay, E.; Har'El, N.; Sivan, R.; Soffer, A. Web-a-where: geotagging web content. In Proceedings of the SIGIR '04, Sheffield, UK, 25–29 July 2004.
101. Silva, M.; Martins, B.; Chaves, M.; Afonso, A.; Cardoso, N. Adding geographic scopes to web resources. *Comput. Environ. Urban Syst.* **2006**, *30*, 378–399. [[CrossRef](#)]



102. Wang, C.; Xie, X.; Wang, L.; Lu, Y.; Ma, W. Detecting geographic locations from web resources. In Proceedings of the GIR '05, Bremen, Germany, 4 November 2005.
103. Frontiera, P.; Larson, R.; Radke, J. A comparison of geometric approaches to assessing spatial similarity for GIR. *Int. J. Geogr. Inf. Sci.* **2008**, *22*, 337–360. [[CrossRef](#)]
104. Jones, C.; Purves, R.; Ruas, A.; Sanderson, M.; Sester, M.; Kreveld, M.V.; Weibel, R. Spatial information retrieval and geographical ontologies an overview of the SPIRIT project. In Proceedings of the SIGIR '02, Tampere, Finland, 11–15 August 2002.
105. Keßler, C.; Janowicz, K.; Bishr, M. An agenda for the next generation gazetteer: Geographic information contribution and retrieval. In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Seattle, WA, USA, 3–6 November 2009; pp. 91–100.
106. Gey, F.; Larson, R.; Sanderson, M.; Joho, H.; Clough, P.D. GeoCLEF: the CLEF 2005 Cross-Language Geographic Information Retrieval Track. In Proceedings of the CLEF, Vienna, Austria, 21–23 September 2005.
107. Egenhofer, M. Toward the semantic geospatial web. In Proceedings of the GIS '02, McLean, VA, USA, 8–9 November 2002.
108. Hart, G.; Dolbear, C. *Linked Data: A Geographic Perspective*; CRC Press: Boca Raton, FL, USA, 2013.
109. Kuhn, W.; Kauppinen, T.; Janowicz, K. Linked data—a paradigm shift for geographic information science. In Proceedings of the International Conference on Geographic Information Science, Vienna, Austria, 24–26 September 2014; Springer: Berlin/Heidelberg, Germany, 2014; pp. 173–186.
110. Goodwin, J.; Dolbear, C.; Hart, G. Geographical Linked Data: The Administrative Geography of Great Britain on the Semantic Web. *Trans. Gis* **2008**, *12*, 19–30. [[CrossRef](#)]
111. Patroumpas, K.; Alexakis, M.; Giannopoulos, G.; Athanasiou, S. TripleGeo: An ETL Tool for Transforming Geospatial Data into RDF Triples. In Proceedings of the Edbt/Icdt Workshops, Citeseer, Athens, Greece, 24–28 March 2014; pp. 275–278.
112. Janowicz, K.; Scheider, S.; Pehle, T.; Hart, G. Geospatial semantics and linked spatiotemporal data—Past, present, and future. *Semant. Web* **2012**, *3*, 321–332. [[CrossRef](#)]
113. Battle, R.; Kolas, D. GeoSPARQL: Enabling a Geospatial Semantic Web. *Semantic Web J.* **2011**, *3*, 355–370. [[CrossRef](#)]
114. Athanasis, N.; Kalabokidis, K.; Vaitis, M.; Soulakellis, N. Towards a semantics-based approach in the development of geographic portals. *Comput. Geosci.* **2009**, *35*, 301–308. [[CrossRef](#)]
115. Purves, R.; Edwardes, A.; Wood, J. Describing place through user generated content. *First Monday* **2011**, *16*. [[CrossRef](#)]
116. Manning, C.D.; Surdeanu, M.; Bauer, J.; Finkel, J.R.; Bethard, S.; McClosky, D. The Stanford CoreNLP natural language processing toolkit. In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, Baltimore, MD, USA, 23–24 June 2014; pp. 55–60.
117. Adams, B.; McKenzie, G.; Gahegan, M. Frankenplace: Interactive Thematic Mapping for Ad Hoc Exploratory Search. In Proceedings of the 24th International Conference on World Wide Web, WWW '15, Florence, Italy, 18–22 May 2015; International World Wide Web Conferences Steering Committee: Geneva, Switzerland, 2015; pp. 12–22. [[CrossRef](#)]
118. Kim, J.; Vasardani, M.; Winter, S. Similarity matching for integrating spatial information extracted from place descriptions. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 56–80. [[CrossRef](#)]
119. Ye, M.; Shou, D.; Lee, W.; Yin, P.; Janowicz, K. On the semantic annotation of places in location-based social networks. In Proceedings of the KDD, San Diego, CA, USA, 21–24 August 2011.
120. Rattenbury, T.; Naaman, M. Methods for extracting place semantics from Flickr tags. *ACM Trans. Web (TWEB)* **2009**, *3*, 1–30. [[CrossRef](#)]
121. Hu, Y.; Gao, S.; Janowicz, K.; Yu, B.; Li, W.; Prasad, S. Extracting and understanding urban areas of interest using geotagged photos. *Comput. Environ. Urban Syst.* **2015**, *54*, 240–254. [[CrossRef](#)]
122. Klippel, A.; Tappe, H.; Kulik, L.; Lee, P.U. Wayfinding choremes—A language for modeling conceptual route knowledge. *J. Vis. Lang. Comput.* **2005**, *16*, 311–329. [[CrossRef](#)]
123. Renz, J.; Nebel, B. Qualitative Spatial Reasoning Using Constraint Calculi. In *Handbook of Spatial Logics*; Springer: Berlin/Heidelberg, Germany, 2007.
124. Gao, S.; Janowicz, K.; McKenzie, G.; Li, L. Towards Platial Joins and Buffers in Place-Based GIS. In Proceedings of the COMP '13, London, UK, 12–13 March 2013.
125. Janowicz, K.; Scheider, S.; Adams, B. A Geo-semantics Flyby. In Proceedings of the Reasoning Web. Semantic Technologies for Intelligent Data Access, Reasoning Web 2013, Mannheim, Germany, 30 July–2 August 2013; Rudolph, S., Gottlob, G., Horrocks, I., van Harmelen, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8067, pp. 230–250. [[CrossRef](#)]
126. Bateman, J.; Farrar, S. Towards a generic foundation for spatial ontology. In *Proceedings of the Third International Conference on Formal Ontology in Information Systems*; IOS Press: Amsterdam, The Netherlands, 2004.
127. Bittner, T.; Donnelly, M.; Smith, B. A spatio-temporal ontology for geographic information integration. *Int. J. Geogr. Inf. Sci.* **2009**, *23*, 765–798. [[CrossRef](#)]
128. Brodaric, B.; Gahegan, M. Experiments to Examine the Situated Nature of Geoscientific Concepts. *Spat. Cogn. Comput.* **2007**, *7*, 61–95. [[CrossRef](#)]
129. Bennett, B.; Mallenby, D.; Third, A. An Ontology for Grounding Vague Geographic Terms. In Proceedings of the FOIS, Saarbrücken, Germany, 31 October–3 November 2008.
130. Chrisman, N. *Exploring Geographic Information Systems*; Wiley: Hoboken, NJ, USA, 2001.
131. Kuhn, W. Semantic reference systems. *Int. J. Geogr. Inf. Sci.* **2003**, *17*, 405–409. [[CrossRef](#)]

132. Mark, D.M.; Smith, B.; Egenhofer, M.; Hirtle, S. Ontological foundations for geographic information science. *Res. Challenges Geogr. Inf. Sci.* **2004**, 335–350.
133. Kauppinen, T.; Hyvönen, E. Modeling and Reasoning About Changes in Ontology Time Series. In *Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems*; Sharman, R., Kishore, R., Ramesh, R., Eds.; Springer: Boston, MA, USA, 2007; pp. 319–338. [[CrossRef](#)]
134. Christakos, G.; Bogaert, P.; Serre, M. *Temporal GIS: Advanced Functions for Field-Based Applications*; Springer: Berlin/Heidelberg, Germany, 2002.
135. Galton, A.; Mizoguchi, R. The water falls but the waterfall does not fall: New perspectives on objects, processes and events. *Appl. Ontol.* **2009**, 4, 71–107. [[CrossRef](#)]
136. Hage, W.V.; Malaisé, V.; Segers, R.; Hollink, L.; Schreiber, G. Design and use of the Simple Event Model (SEM). *J. Web Semant.* **2011**, 9, 128–136. [[CrossRef](#)]
137. Montello, D.R.; Goodchild, M.F.; Gottsegen, J.; Fohl, P. Where's downtown?: Behavioral methods for determining referents of vague spatial queries. *Spat. Cogn. Comput.* **2003**, 3, 185–204. [[CrossRef](#)]
138. Abdelmoty, A.I.; Smart, P.; Jones, C.B. Building Place Ontologies for the Semantic Web: Issues and Approaches. In Proceedings of the 4th ACM Workshop on Geographical Information Retrieval, GIR'07, Lisbon, Portugal, 9 November 2007; Association for Computing Machinery: New York, NY, USA, 2007; pp. 7–12. [[CrossRef](#)]
139. Lutz, M.; Klien, E. Ontology-based retrieval of geographic information. *Int. J. Geogr. Inf. Sci.* **2006**, 20, 233–260. [[CrossRef](#)]
140. Jones, C.; Alani, H.; Tudhope, D. Geographical Information Retrieval with Ontologies of Place. In Proceedings of the COSIT, Morro Bay, CA, USA, 19–23 September 2001.
141. Jordan, T.; Raubal, M.; Gartrell, B.; Egenhofer, M. An Affordance-Based Model of Place in GIS. 1999. Available online: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.4.8628&rep=rep1&type=pdf> (accessed on 1 April 2021)
142. Alazzawi, A.; Abdelmoty, A.; Jones, C. What can I do there? Towards the automatic discovery of place-related services and activities. *Int. J. Geogr. Inf. Sci.* **2012**, 26, 345–364. [[CrossRef](#)]
143. Ying, J.; Lee, W.; Weng, T.C.; Tseng, V.S. Semantic trajectory mining for location prediction. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, Chicago, IL, USA, 1–4 November 2011.
144. Harvey, F.; Kuhn, W.; Pundt, H.; Bishr, Y.; Riedemann, C. Semantic interoperability: A central issue for sharing geographic information. *Ann. Reg. Sci.* **1999**, 33, 213–232. [[CrossRef](#)]
145. Dou, D.; McDermott, D.; Qi, P. Ontology Translation on the Semantic Web. *J. Data Semant.* **2005**, 2, 35–57.
146. Raubal, M. Formalizing conceptual spaces. Formal ontology in information systems. In Proceedings of the Third International Conference (FOIS 2004), Torino, Italy, 4–6 November 2004; Volume 114, pp. 153–164.
147. Rodriguez, M.A.; Egenhofer, M. Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. *Int. J. Geogr. Inf. Sci.* **2004**, 18, 229–256. [[CrossRef](#)]
148. Schwering, A.; Raubal, M. Measuring Semantic Similarity Between Geospatial Conceptual Regions. In Proceedings of the GeoS, Mexico City, Mexico, 29–30 November 2005.
149. Li, B.; Fonseca, F. TDD: A comprehensive model for qualitative spatial similarity assessment. *Spat. Cogn. Comput.* **2006**, 6, 31–62. [[CrossRef](#)]
150. Martin, D.; Burstein, M.; McDermott, D.; McIlraith, S.; Paolucci, M.; Sycara, K.; McGuinness, D.L.; Sirin, E.; Srinivasan, N. Bringing semantics to web services with OWL-S. *World Wide Web* **2007**, 10, 243–277. [[CrossRef](#)]
151. Fensel, D.; Bussler, C. The Web Service Modeling Framework WSMF. *Electron. Commer. Res. Appl.* **2002**, 1, 113–137. [[CrossRef](#)]
152. Vaccari, L.; Shvaiko, P.; Marchese, M. A geo-service semantic integration in Spatial Data Infrastructures. *Int. J. Spat. Data Infrastruct. Res.* **2009**, 4, 24–51.
153. Lemmens, R.; Wytzisk, A.; de By, R.; Granell, C.; Gould, M.; Van Oosterom, P. Integrating semantic and syntactic descriptions to chain geographic services. *IEEE Internet Comput.* **2006**, 10, 42–52. [[CrossRef](#)]
154. Lutz, M. Ontology-based descriptions for semantic discovery and composition of geoprocessing services. *Geoinformatica* **2007**, 11, 1–36. [[CrossRef](#)]
155. Janowicz, K.; Schade, S.; Bröring, A.; Keßler, C.; Maué, P.; Stasch, C. Semantic Enablement for Spatial Data Infrastructures. *Trans. GIS* **2010**, 14, 111–129. [[CrossRef](#)]
156. Klien, E. A Rule-Based Strategy for the Semantic Annotation of Geodata. *Trans. GIS* **2007**, 11, 437–452. [[CrossRef](#)]
157. Kuhn, W. Cognitive and Linguistic Ideas in Geographic Information Semantics. In *Cognitive and Linguistic Aspects of Geographic Space*; Lecture Notes in Geoinformation and Cartography; Springer: Berlin/Heidelberg, Germany, 2013; pp. 159–174. [[CrossRef](#)]
158. Zhu, Y. Geospatial semantics, ontology and knowledge graphs for big Earth data. *Big Earth Data* **2019**, 3, 187–190. [[CrossRef](#)]
159. Di Donato, P. Geospatial Semantics: A Critical Review. In *Computational Science and Its Applications—ICCSA 2010*; Taniar, D., Gervasi, O., Murgante, B., Pardede, E., Apduhan, B.O., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6016. [[CrossRef](#)]
160. Schuurman, N. Formalization Matters: Critical GIS and Ontology Research. *Ann. Assoc. Am. Geogr.* **2006**, 96, 726–739. [[CrossRef](#)]
161. Wang, C.; Kantor, C.M.; Mitchell, J.T.; Bacastow, T.S. Digital Earth Education. In *Manual of Digital Earth*; Springer Nature: Berlin/Heidelberg, Germany, 2020; pp. 755–783.
162. Goffi, A.; Bordogna, G.; Stroppiana, D.; Boschetti, M.; Brivio, P.A. Knowledge and Data-Driven Mapping of Environmental Status Indicators from Remote Sensing and VGI. *Remote Sens.* **2020**, 12, 495. [[CrossRef](#)]

163. Bordogna, G.; Ghisalberti, G.; Psaila, G. Geographic information retrieval: Modeling uncertainty of user's context. *Fuzzy Sets Syst. FSS* **2012**, *196*. [[CrossRef](#)]
164. Zadrozny, S.; Kacprzyk, J. Bipolar Queries Using Various Interpretations of Logical Connectives. In Proceedings of the Foundations of Fuzzy Logic and Soft Computing, 12th International Fuzzy Systems Association World Congress, IFSA 2007, Cancun, Mexico, 18–21 June 2007; Melin, P., Castillo, O., Aguilar, L.T., Kacprzyk, J., Pedrycz, W., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4529, pp. 181–190. [[CrossRef](#)]
165. Dubois, D.; Prade, H. An introduction to bipolar representations of information and preference. *Int. J. Intell. Syst.* **2008**, *23*, 866–877. [[CrossRef](#)]
166. Dujmović, J.; Larsen, H. Generalized conjunction/disjunction. *Int. J. Approx. Reason.* **2007**, *46*, 423–446. [[CrossRef](#)]
167. Arcaini, P.; Bordogna, G.; Ienco, D.; Sterlacchini, S. User-driven geo-temporal density-based exploration of periodic and not periodic events reported in social networks. *Inf. Sci.* **2016**, *340–341*, 122–143. [[CrossRef](#)]
168. Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*; Language, Speech, and Communication; MIT Press: Cambridge, MA, USA, 1998.
169. Psaila, G.; Toccu, M.; Bordogna, G.; Frigerio, L.; Cuzzocrea, A. An Interoperable Open Data Framework for Discovering Popular Tours Based on Geo-Tagged Tweets. *Int. J. Intell. Inf. Database Syst.* **2017**, *10*, 1. [[CrossRef](#)]
170. Miles, A.; Bechhofer, S. SKOS Simple Knowledge Organization System Reference. In Proceedings of the W3C Recommendation, W3C, Maputo, Mozambique, 1–2 April 2009.
171. Perego, A.; Fugazza, C.; Vaccari, L.; Lutz, M.; Smits, P.; Kanellopoulos, I.; Schade, S. Harmonization and Interoperability of EU Environmental Information and Services. *IEEE Intell. Syst.* **2012**, *27*, 33–39. [[CrossRef](#)]
172. Brickley, D.; Miller, L. The Friend Of A Friend (FOAF) Vocabulary Specification. 2007. Available online: <http://xmlns.com/foaf/spec/> (accessed on 1 April 2021)
173. Santoro, M.; Mazzetti, P.; Nativi, S.; Fugazza, C.; Granell, C.; Díaz, L. Methodologies for augmented discovery of geospatial resources. In *Discovery of Geospatial Resources: Methodologies, Technologies, and Emergent Applications*; Díaz, L., Granell, C., Huerta, J., Eds.; IGI Global: Hershey, PA, USA, 2012; Chapter 9, pp. 172–203.
174. Sporny, M.; Longley, D.; Kellogg, G.; Lanthaler, M.; Champin, P.A.; Lindström, N. JSON-LD 1.1 A JSON-Based Serialization for Linked Data. Ph.D. Thesis, W3C Recommendation, W3C, Cambridge, MA, USA, 2020.
175. Wilkinson, M.D.; Dumontier, M.; Aalbersberg, I.J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.W.; da Silva Santos, L.B.; Bourne, P.E.; et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **2016**, *3*, 1–9. [[CrossRef](#)]
176. Huang, Y.; Yuan, M.; Sheng, Y.; Min, X.; Cao, Y. Using Geographic Ontologies and Geo-Characterization to Represent Geographic Scenarios. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 566. [[CrossRef](#)]
177. Lü, G.; Batty, M.; Strobl, J.; Lin, H.; Zhu, A.X.; Chen, M. Reflections and speculations on the progress in Geographic Information Systems (GIS): A geographic perspective. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 346–367. [[CrossRef](#)]
178. Bertalanffy, L.V. *General Systems Theory: Foundations, Development, Applications* / by Ludwig von Bertalanffy, rev. ed.; Braziller: New York, NY, USA, 1968.
179. Horrocks, I.; Patel-Schneider, P.F.; Boley, H.; Tabet, S.; Grosz, B.; Dean, M. SWRL: A Semantic Web Rule Language Combining OWL and RuleML. *W3C Memb. Submiss.* **2004**, *21*, 1–31.
180. Fugazza, C.; Menegon, S.; Pepe, M.; Oggioni, A.; Carrara, P. The RITMARE Starter Kit—Bottom-up Capacity Building for Geospatial Data Providers. In Proceedings of the 9th International Conference on Software Paradigm Trends (ICSOFT), INSTICC, Vienna, Austria, 29–31 August 2014; SciTePress: Setúbal, Portugal, 2014; Volume 1, pp. 169–176. [[CrossRef](#)]
181. Fugazza, C.; Basoni, A.; Menegon, S.; Oggioni, A.; Pavesi, F.; Pepe, M.; Sarretta, A.; Carrara, P. RITMARE: Semantics-aware Harmonisation of Data in Italian Marine Research. *Procedia Comput. Sci.* **2014**, *33*, 261–265. [[CrossRef](#)]
182. Tagliolato, P.; Fugazza, C.; Oggioni, A.; Carrara, P. Semantic Profiles for Easing SensorML Description: Review and Proposal. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 340. [[CrossRef](#)]
183. Pavesi, F.; Basoni, A.; Fugazza, C.; Menegon, S.; Oggioni, A.; Pepe, M.; Tagliolato, P.; Carrara, P. EDI—A Template-Driven Metadata Editor for Research Data. *JORS J. Open Res. Softw.* **2016**, *4*, e40. [[CrossRef](#)]
184. SPARQL Working Group. SPARQL 1.1 Query Language. W3C Recommendation 21 March 2013, World Wide Web Consortium. 2013. Available online: <http://www.w3.org/TR/sparql11-query/> (accessed on 1 April 2021).
185. Fugazza, C.; Pepe, M.; Oggioni, A.; Tagliolato, P.; Pavesi, F.; Carrara, P. Describing Geospatial Assets in the Web of Data: A Metadata Management Scenario. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 229. [[CrossRef](#)]
186. Fugazza, C.; Tagliolato, P.; Frigerio, L.; Carrara, P. Web-Scale Normalization of Geospatial Metadata Based on Semantics-Aware Data Sources. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 354. [[CrossRef](#)]
187. Fugazza, C.; Pepe, M.; Oggioni, A.; Tagliolato, P.; Carrara, P. Raising Semantics-Awareness in Geospatial Metadata Management. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 370. [[CrossRef](#)]
188. Lefrançois, M.; Zimmermann, A.; Bakerally, N. A SPARQL extension for generating RDF from heterogeneous formats. In Proceedings of the Extended Semantic Web Conference (ESWC'17), Portorož, Slovenia, 28 May–1 June 2017.
189. Prud'hommeaux, E.; Harris, S.; Seaborne, A. SPARQL 1.1 Query Language; Technical Report W3C. 2013. Available online: <https://www.w3.org/TR/sparql11-query/> (accessed on 1 April 2021).

190. Lehnert, K. IGSN: International Geo Sample Number. Unambiguous Citation of Physical Samples. 2015. Available online: <https://zenodo.org/record/31788> (accessed on 1 April 2021).
191. Wohner, C.; Peterseil, J.; Poursanidis, D.; Kliment, T.; Wilson, M.; Mirtl, M.; Chrysoulakis, N. DEIMS-SDR—A web portal to document research sites and their associated data. *Ecol. Inform.* **2019**, *51*, 15–24. [[CrossRef](#)]
192. Zilioli, M.; Oggioni, A.; Tagliolato, P.; Pugnetti, A.; Carrara, P. Feeding Essential Biodiversity Variables (EBVs): Actual and potential contributions from LTER-Italy. *Nat. Conserv.* **2019**, *34*, 477–503. [[CrossRef](#)]
193. Percivall, G. OGC's Open Standards for Geospatial Interoperability. In *Encyclopedia of GIS*; Springer International Publishing: Cham, Switzerland, 2017; pp. 1466–1473. [[CrossRef](#)]
194. Richardson, L.; Amundsen, M.; Ruby, S. *RESTful Web APIs*; O'Reilly Media, Inc.: Newton, MA, USA, 2013.
195. Fielding, R.T. REST: Architectural Styles and the Design of Network-based Software Architectures. Ph.D. Thesis, University of California, Irvine, CA, USA, 2000.
196. Wilkinson, M.; Verborgh, R.; Bonino da Silva Santos, L.O.; Clark, T.; Swertz, M.; Kelpin, F.; Gray, A.; Schultes, E.; van Mulligen, E.M.; Ciccacese, P.; et al. Interoperability and FAIRness through a novel combination of Web technologies. *PeerJ Comput. Sci.* **2016**. [[CrossRef](#)]
197. Trani, L.; Atkinson, M.; Bailo, D.; Paciello, R.; Filgueira, R. Establishing Core Concepts for Information-Powered Collaborations. *Future Gener. Comput. Syst.* **2018**, *89*, 421–437. [[CrossRef](#)]
198. Trani, L.; Paciello, R.; Sbarra, M.; Ulbricht, D. Representing Core Concepts for solid-Earth sciences with DCAT—The EPOS-DCAT Application Profile. In Proceedings of the EGU General Assembly Conference Abstracts, Vienna, Austria, 8–13 April 2018; p. 9797.
199. Lanthaler, M.; Gütl, C. Hydra: A Vocabulary for Hypermedia-Driven Web APIs. In Proceedings of the WWW2013 Workshop on Linked Data on the Web, Rio de Janeiro, Brazil, 14 May 2013; Volume 996, ISSN 1613-0073.
200. Lanucara, S.; Fugazza, C.; Tagliolato, P.; Oggioni, A. Information Systems for Precision Agriculture: Monitoring Computation of Prescription Maps. *ERCIM News* **2018**, *113*, 24–25.
201. Ventura, D.; Verborgh, R.; Catania, V.; Mannens, E. Autonomous Composition and Execution of REST APIs for Smart Sensors. In Proceedings of the Joint Proceedings of the 1st Joint International Workshop on Semantic Sensor Networks and Terra Cognita and the 4th International Workshop on Ordering and Reasoning, CEUR Workshop Proceedings, Bethlehem, PA, USA, 11–12 October 2015; Volume 1488, pp. 25–30.
202. Cross, V.; Chen, S. *Fuzzy Information Processing; Communications in Computer and Information Science; Chapter Fuzzy Ontologies: The State of the Art revised*; Springer: Berlin/Heidelberg, Germany, 2018.
203. Bobillo, F.; Straccia, U. The fuzzy ontology reasoner fuzzyDL. *Knowl.-Based Syst.* **2016**, *95*, 12–34. [[CrossRef](#)]
204. Safia, B.-B.; Aicha, M. Poss-OWL 2: Possibilistic Extension of OWL 2 for an uncertain geographic ontology. In Proceedings of the 18th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems—KES2014, Gdynia, Poland, 16–18 September 2014; Elsevier Procedia: Amsterdam, The Netherlands, 2014; pp. 407–416.
205. Aerts, D.; Gabora, L. A theory of concepts and their combinations II: A Hilbert space representation. *Kybernetes* **2005**, *34*, 192–221. [[CrossRef](#)]
206. Bordogna, G.; Kliment, T.; Frigerio, L.; Brivio, P.A.; Crema, A.; Stroppiana, D.; Boschetti, M.; Sterlacchini, S. A 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](#)]
207. Bordogna, G.; Frigerio, L.; Kliment, T.; Brivio, P.; Hossard, L.; Manfron, G.; Sterlacchini, S. Contextualized VGI Creation and Management to Cope with Uncertainty and Imprecision. *Int. J. Geo-Inf.* **2016**, *5*, 234–253. [[CrossRef](#)]
208. Hahn, J.; Frank, A.U. Select the appropriate map depending on context in a hilbert space model (scop). In *International Symposium on Quantum Interaction; Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8369. [[CrossRef](#)]
209. Slonneger, K.; Kurtz, B. *Formal Syntax and Semantics of Programming Languages: A Laboratory Based Approach*, 1st ed.; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, USA, 1995.
210. Miller, G. *The Magical Number Seven, Plus Or Minus Two: Some Limits on Our Capacity for Processing Information*; Bobbs-Merrill Reprint Series in the Social Sciences; P-241, College Division of Bobbs-Merrill Company: Indianapolis, IN, USA, 1975.