A Review of Geospatial Semantic Information Modeling and Elicitation Approaches

Margarita Kokla [1],* and Eric Guilbert [2]

[1]   School of Rural and Surveying Engineering, National Technical University of Athens, 15780 Athens, Greece
[2]   Department of Geomatics Sciences, Université Laval, Québec, QC G1V 0A6, Canada;
      eric.guilbert@scg.ulaval.ca
*   Correspondence: mkokla@survey.ntua.gr

check for updates

Abstract: The present paper provides a review of two research topics that are central to geospatial semantics: information modeling and elicitation. The first topic deals with the development of ontologies at different levels of generality and formality, tailored to various needs and uses. The second topic involves a set of processes that aim to draw out latent knowledge from unstructured or semi-structured content: semantic-based extraction, enrichment, search, and analysis. These processes focus on eliciting a structured representation of information in various forms such as: semantic metadata, links to ontology concepts, a collection of topics, etc. The paper reviews the progress made over the last five years in these two very active areas of research. It discusses the problems and the challenges faced, highlights the types of semantic information formalized and extracted, as well as the methodologies and tools used, and identifies directions for future research.

1. Introduction

Geospatial semantics has been at the center of research over the last 20 years. The semantic problems occurring during the exchange, reuse, and integration of heterogeneous spatial data and the need to achieve semantic interoperability were among the first research challenges faced by the Geographic Information Science (GIScience) community [1]. In this endeavor, data about location and thematic attributes were not considered rich enough to fully represent the complexity of geographic entities and to support their integration and interpretation in different contexts. Hence, the discussion shifted from geospatial data to geospatial concepts and from data specifications to the nature and conceptualization of geographic entities and phenomena. Research on ontologies allowed for formal representations of geospatial concepts and enriched the discussion on the nature and conceptualization of geographic space.

Over these years, several reviews of geospatial semantics have highlighted different research directions based on the challenges of that specific period. In an early review of geospatial semantics, Kuhn [1] identified two aspects pertaining to this research field: Understanding Geographic Information Systems (GIS) contents and capturing this understanding in formal theories. He also highlighted that reasoning is even more important than formalization of meaning. He identified several challenges related to geospatial semantics and semantic interoperability: data discovery and evaluation, and service discovery, evaluation, and composition.

The interpretation of data across heterogeneous data sources and scientific domains became even more challenging in the context of the Semantic Web. The Semantic Web has further complicated the modeling, interpretation, reuse, and integration of knowledge, since "web knowledge has no primitives,

no core, no fundamental categories, no fixed structure, and is heavily context-dependent" [2] (p.4). Two main strands of research on geospatial semantics are prominent in this context: semantic modeling and semantic-based search, integration, and interoperability [3].

From a cognitive and linguistic perspective, notions, such as experiential realism, geographic information atoms, semantic reference systems, semantic datums, similarity measurement, and conceptual spaces, prevail in geospatial semantics research [4]. Under this perspective, the problem of semantics of geographic information is formulated on the basis of constraints on the use and interpretation of geographic terms [5].

Although, space has emerged as a fundamental integrative framework to support interdisciplinary research, spatial notions evoke different interpretations, across different domains and common-sense conceptualizations. Kuhn [6] proposed a set of 10 high-level concepts essential for understanding spatial information in a transdisciplinary context: location, neighborhood, field, object, network, event, granularity, accuracy, meaning, and value. These concepts may constitute the core for supporting a broader use of spatial information in science and society; however, formalizing and mapping their varying uses across disciplines is still a research challenge.

In a more recent review of geospatial semantics, Hu delineates six major, highly-interconnected research areas for which the meaning of geographic information constitutes a common core: semantic interoperability, digital gazetteers, geographic information retrieval, geospatial Semantic Web, place semantics, and cognitive geospatial concepts [7].

The present paper delves into two research challenges that are central to geospatial semantics: information modeling and elicitation. The first challenge deals with the development of ontologies at different levels to support various knowledge-based processes. At an upper level, research focuses on a richer semantic characterization of space, time, spatial and temporal concepts, and relations. These are central notions across upper-level ontologies. They act as unifying framework across domains and are fundamental for ontology development, reuse, and integration. At the domain level, interdisciplinary geospatial concepts (In the context of the paper, both for the semantic information formalization and elicitation processes, the term "concept" is used to denote a conceptual grouping of entities or instances based on a set of distinguishing properties, roles, or functions. The terms "instances" or "entities" are used to denote the "things" represented by a concept (e.g., the concept "city" compared to the instance Paris). present many challenges for semantic formalization. Ontological research has also moved beyond the strict formalization of geospatial concepts under specific domains, to the shared interpretation of their meaning across different contexts [3] and the development of lightweight and micro-ontologies tailored to specific needs [2,8,9].

The second challenge refers to the elicitation of semantic information from semi-structured and unstructured resources: places, regions, events, trajectories, topics, geospatial concepts and relations. Knowledge elicitation traditionally refers to a set of techniques that attempt to deduce the knowledge of a domain expert in order to obtain a concrete representation of this knowledge [10]. In this paper; however, the term is used more broadly to embrace a set of processes that aim to draw out latent knowledge from unstructured or semi-structured content: semantic-based extraction, enrichment, search, and analysis. These processes focus on eliciting a structured representation of information in various forms, such as: semantic metadata, links to ontology concepts, a collection of topics, a semantic network, etc.

The paper reviews the progress made over the last five years in these two very active areas of research. It discusses the various problems targeted and the challenges faced, highlights the types of semantic information formalized and extracted, as well as the methodologies and tools used, and identifies directions for future research.

2. Semantic Information Formalization

The formalization of semantics is a prevalent research issue for various processes, such as semantic mapping and integration, semantic search, and knowledge discovery. Although geographic space

may be viewed as a common ground to interconnect different conceptualizations, its inherent semantic complexity and vagueness, as well as the different domains and perspectives involved, cause semantic heterogeneities that need to be identified and resolved. Ontological approaches have been recognized as critical to model geospatial semantics, resolve semantic heterogeneities, integrate different semantic descriptions, and ground conceptualizations [11].

A co-citation and cluster analysis of subjects, journals, and authors based on 1533 articles published between 2001 and 2016, has shown that ontologies are prominent in GIScience research across several disciplines (computer science, engineering, geography, geosciences, etc.) [12]. Ontologies have been employed for solving various research problems such as GEographic Object-Based Image Analysis (GEOBIA) [13–15], information extraction and retrieval [16,17], information integration [18,19], linked data [20–22] and geoprocessing workflows [23], geospatial data provenance on the web [24], interpretation of natural language descriptions [25], automatic feature recognition from point clouds [26], and sketch map interpretation [27].

Traditionally, ontologies are classified according to two criteria: their level of formality and their level of generality. According to formality, ontologies range from informal to semiformal and formal, although the distinction is not meant to be a rigid one [28]. According to generality, four types of ontologies are distinguished: top-level, domain, task, and application ontologies [29]. Top-level, upper-level, or foundational ontologies define general, fundamental concepts, such as entity, property, relation, process, action, space, and time. These are domain-independent concepts and; therefore, are used as a framework for unfolding and associating more specialized domain knowledge. Domain ontologies define the concepts pertinent to a domain, such as oceanography, topography, meteorology, natural disasters, etc., whereas task ontologies define the concepts relating to a task, such as analysis, planning, monitoring, and forecasting. Both domain and task ontologies may specialize the concepts defined by a top-level ontology. Application ontologies on the other hand, define concepts relative to a particular domain and task, and may specialize both corresponding ontologies. For example, an application ontology on urban planning formalizes both domain knowledge on the urban environment and task knowledge on planning. An application ontology on earthquake emergency response specializes both domain knowledge on natural disasters and task knowledge on monitoring, warning, response, and recovery. Figure 1 shows examples of ontologies developed at different levels of generality.

2.1. Upper-Level Ontologies

Geographic ontologies may be considered as domain ontologies. However, top-level ontologies are also relevant to geospatial knowledge. They define central notions for the geospatial domain such as space, time, spatial regions, boundaries, and processes and investigate ontological issues regarding the hypostasis and dimensionality of geographic entities and their dependence on spatial regions and boundaries [30–32]. Prominent upper-level ontologies that have influenced geospatial ontology development and research are: Basic Formal Ontology (BFO) [33,34], Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) [35], General Formal Ontology (GFO) [36], Generalized Upper Model (GUM) [37], and Suggested Upper Merged Ontology (SUMO) [38]. WordNet [39], although not designed as an ontology but as a lexical database and semantic network, has also been used in the geospatial domain for ontology-based information extraction and retrieval, and semantic similarity computation.

Top-level ontologies are core components of formal ontology, the discipline that integrates aspects of philosophy, formal logic, and artificial intelligence [40]. Formal ontology is defined as the theory of a priori distinctions among the entities of the world (physical objects, events, processes, quantities, etc.) and among the meta-categories used to model the world (concepts, properties, qualities, states, etc.) [41]. An important ontological distinction upon which several top-level ontologies are built is that between continuants and occurrents. Continuants (also called endurants) are objects that are wholly present through time, whereas occurrents (also called perdurants) are objects that are temporally

restricted and have temporal parts, such as processes or events. Space, spatial regions, and spatial relations, as well as time and temporal phenomena are fundamental notions in upper-level ontologies, as are mereology, the theory of parthood relations [42], and topology, the theory of spatial continuity and compactness [43].

In the last five years, the discipline of formal ontology has focused on further formalization of these notions. Baumann et al. [44] introduced GFO-Space, a first-order formalization of an ontology of space for General Formal Ontology (GFO). The principles underlying the ontology originate from Brentano's ideas on space and continuum and focus on phenomenal space (i.e., space determined by material objects and relations between them), as appearing to the mind of a subject. GFO-Space accounts for mereotopological relations, boundaries and dimensionality of entities. The axiomatization of their theory is based on four primitives: The category "space region" (i.e., the basic entities of phenomenal space), and the relations: "spatial part of", "spatial boundary of", and "coincidence".
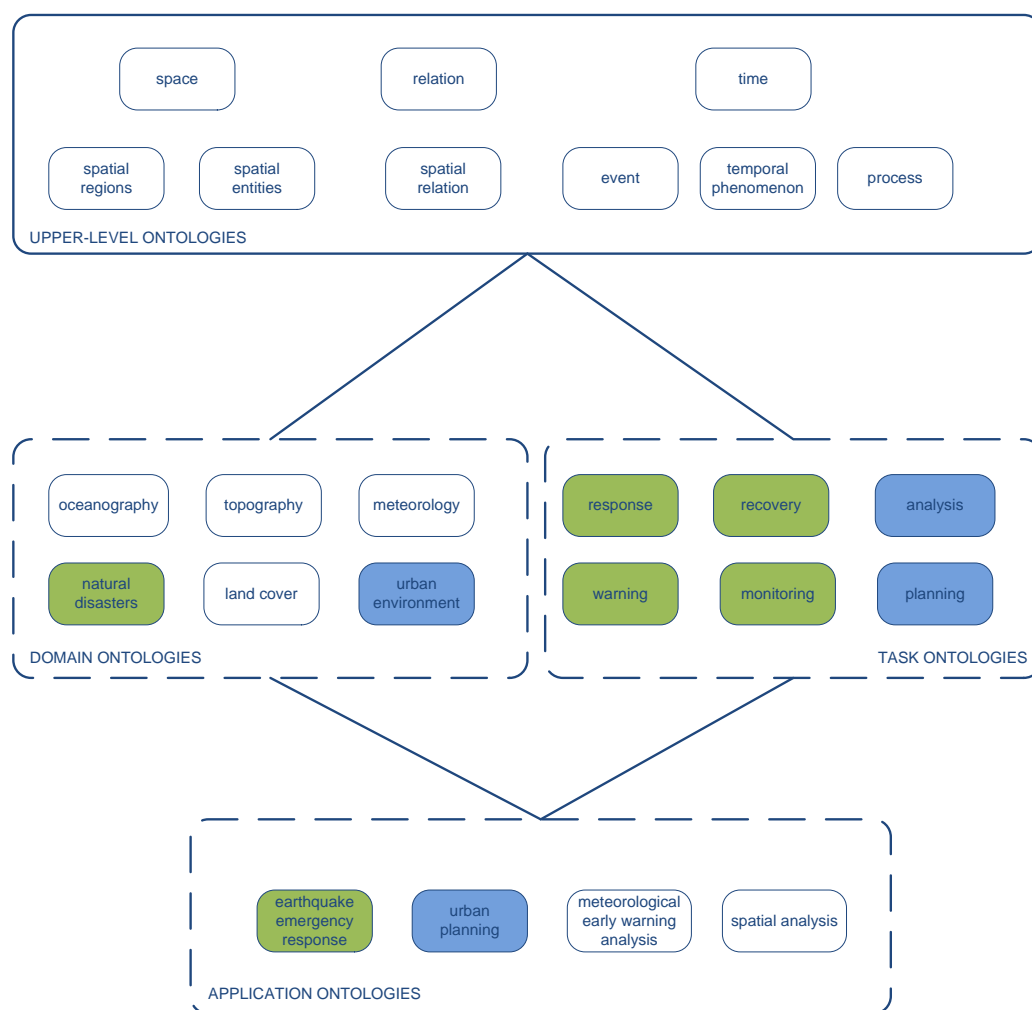

Figure 1. Indicative ontologies at different levels of generality.

Mereogeometry is the theory of space and spatial relations which formalizes geometric notions within a region-based mereological framework. On the other hand, mereotopological relations, such as contact, parthood, and overlap, are also fundamental for the qualitative representation of spatial information. Schmidtke [45] outlines a mereogeometric framework for the formalization of geometric relations of incidence, congruence, and parallelism over extended regions. Hahmann [46] extends the axiomatization of CODI (Containment and Dimension), a first-order logic ontology of multidimensional mereotopology, with the mereological operations intersection and difference that apply to pairs of

regions regardless of their dimensions. CODI provides a first-order formalization of the notions of spatial containment and relative spatial dimension and defines a set of six intuitive mereotopological relations: Containment and its refinement parthood as well as contact and its refinements partial overlap, incidence, and superficial contact.

Time and temporal phenomena are also prominent notions within upper-level ontologies. Key issues relating to how time and temporal phenomena are treated relate to: (a) Dimensionality (i.e., the treatment of instants and intervals), (b) frame-dependence (i.e., definition of time with respect to a reference frame), and (c) indexicality (i.e., reference to and distinction of the past, the present, and the future) [47].

In order to reconcile the data-modeling and process-modeling requirements of GIScience, Galton [48] laid the foundation for a formal theory of events and processes integrating two distinct ways of viewing time, referred as historical and experiential time. Historical time, presupposed by data-modeling functions, such as storage, retrieval, manipulation, and presentation, is considered as static and "frozen" and places emphasis on completed events. Experiential time on the other hand, presupposed by process-modeling functions, such as explanation, prediction, and simulation, is considered as active and "fluid" and places emphasis on ongoing processes. Temporal scale or granularity is considered important to delineate the relationship between processes and events. The theory distinguishes different types of processes, specified by activity conditions, and different types of events, specified by occurrence conditions. It also presents a formalization of operations for deriving events from processes, processes from events, and events from events.

Understanding the models and theories subsumed by upper-level ontologies is critical for making the ontological commitments explicit, specifying relations to other upper-level ontologies and extending upper-level ontologies to create new domain-specific ones. Hanzal et al. [49] present a comparative survey of how foundational and Semantic Web ontologies (Event Ontology (http://purl.org/NET/c4dm/event.owl#), Simple Event Model Ontology [50], Linking Open Descriptions of Events (http://linkedevents.org/ontology/), etc.) formalize the notion of event. Based on modeling considerations followed by these ontologies, they provide an empirical classification of events in four categories: actions, happenings, planned "social" events, and structural components of temporal entities. Categorization relies on the distinction between object and relationship as formalized by the modeling language PURO [51]. Muñoz and Grüninger [52] apply the process of ontology verification on core temporal concepts of the Suggested Upper Merged Ontology (SUMO) to rule out unintended models and characterize missing intended ones. The verification process also supports ontology mappings to other time ontologies, PSL-Core ontology [53] and Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) upper ontology [35].

2.2. Domain, Task, and Application Ontologies

Besides upper-level ontologies, a variety of other ontologies have been developed in the geospatial domain at different levels of formality and generality (Figure 1). Domain ontologies have been developed at different levels of detail, covering whole domains, such as earth and environmental sciences [54], oceanography [20], land cover and land use [55], etc., as well as specific domain concepts, such as city, locality, and forest.

Research in the last five years addresses or revisits issues related to the formalization of interdisciplinary or vague geospatial concepts, grounding their meaning on upper-level ontologies. Furthermore, the formalization of geospatial concepts with a cognitive and linguistic basis such as places, landforms, or landscapes aims to bridge the gap between the qualitative human perception and the necessity to have rigorous, unambiguous definitions in view of their implementation.

Alvarez and Bennett [56] propose a framework for the formal definition of the highly vague and interdisciplinary concept "forest". The framework encapsulates different aspects of forest definitions pertaining to their classification (i.e., determining whether an object is an instance of a class), individuation (i.e., establishing how many distinct individual objects of a given type exist),

and demarcation (i.e., determining the boundary of an object), such as location, morphological, metrical, topological, and mereological restrictions, qualitative characteristics, scale, etc. The framework, which is implemented in a prototype prolog-based GIS, is based on supervaluation semantics to express the variety of possible meanings of the forest concept, from different perspectives. The fundamental idea of supervaluation semantics is that a vague language can be interpreted in many different precise ways, each of which can be modelled in terms of precise truth conditions for each predicate of the language, which is referred to as a precisification.

Ballatore [57] takes a preliminary step towards the formalization of an ontology of place by surveying the representation of this vague, polysemous, and culturally-dependent notion across a range of lightweight and formal ontologies. He also discusses issues (cultural and linguistic dependence of place, temporal dimension, social role, scale, and themes) and outlines conceptual tools for a multi-faceted formal ontology of place.

Calafiore et al. [58] focus on cities, which are viewed as systems of various types of urban artefacts interacting with human activities by playing multiple roles, probably at the same time. They perform an ontological analysis of urban artefacts and their social uses. In this context, a distinction is made between the intended and actual uses of urban artefacts and places are treated as social concepts. The ontological analysis is grounded on DOLCE foundational ontology due to its cognitive orientation. The general taxonomy in DOLCE is enriched with the notions of urban artefact, urban artefact types and roles, social practices, institutional places and social places and uses of artefacts are modeled in terms of roles theory.

In the context of historical geographic information systems, Garbacz et al. [59] identify diachronic criteria of identity for localities in order to model the type of changes they may undergo over time. In this context, localities are considered as endurants, which possess qualities relating to their name, location, type, and mereology, and participate in events or processes (considered as perdurants). Event-based criteria of identity are used to define qualitative transformations of localities with respect to one of these qualities.

Gharebaghi and Mostafavi [60] propose an ontology of the urban environment for modeling the interaction between humans and their social and physical environments and supporting the accessibility of people with disabilities. The ontology is based on the nature-development perspective. In this approach, concepts belong either to natural environment (such as forest and tree) or to developed environment (such as sidewalk and building) and have physical and social properties which are related to each other.

Stephen and Hahmann [61] developed an ontological framework for the formalization of surface and subsurface spatio-temporal processes that describe hydrologic flow (Figure 2). The framework extends static hydrogeological concepts from the Hydro Foundational Ontology (HyFO) [61] and also builds on the distinction between endurants and perdurants and the participation relation $PC(x, y, t)$ between them, as defined by DOLCE upper-level ontology. The participation relation $PC(x, y, t)$ expresses that an endurant x participates in a perdurant y at time t. Flow processes are distinguished based on how and what kinds of endurants can participate and are formalized using semantic roles such as theme participant, source participant, goal participant, and locative participant. The taxonomy and related concepts are axiomatized in first-order logic.

Several domain ontologies were also developed for landforms. Landforms are fuzzy objects whose perception and classification mainly depends on the context and on human cognition. Hence, it has been acknowledged since long that a universal ontology of landforms does not exist. Since landforms are mainly characterized by their shape, several ontologies for specific fields or languages were proposed. These ontologies structure landforms into general concepts that are further specialized with a more precise shape. Yan et al. [62] present an ontology of undersea features following the nomenclature from the International Hydrographic Organization. They first define broad categories such as depression and eminence that are then refined into features such as trench and reef based on some shape descriptors such as "elongated" or "deep". The ontology is applied to feature classification from bathymetric

data. Grenoble et al. [63] propose an ontology including both land and sea terms in the Kalaallisut language, categorizing features according to their shape and their function. More generally, Sinha et al. [64] present a broader landform reference ontology where landforms are also organized according to their shape and according to their dependence on the planetary body (Figure 3). The ontology is applied to mapping linguistic categories to landform categories.
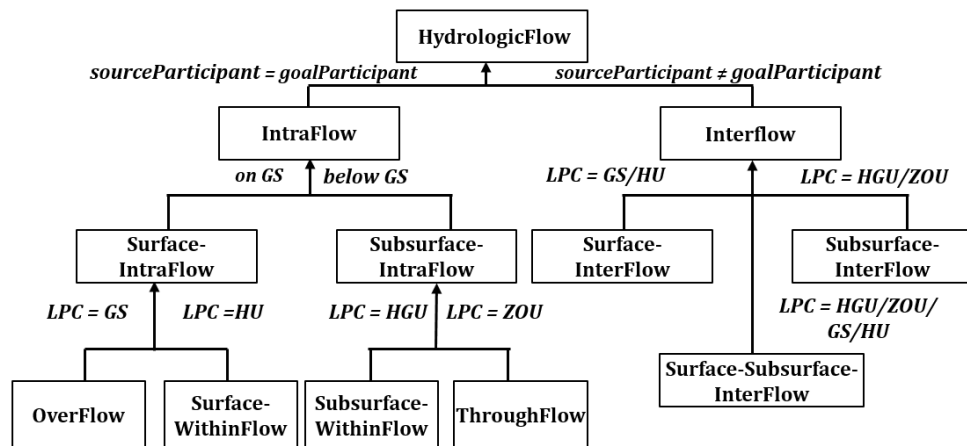

Figure 2. The taxonomy of hydrological flow concepts [61] by Stephen and Hahmann licensed under CC-BY 3.0. Hydrological flow is distinguished at the higher-level based on whether water moves within a single locative participant (intraflow) or between two distinct locative participants (interflow).
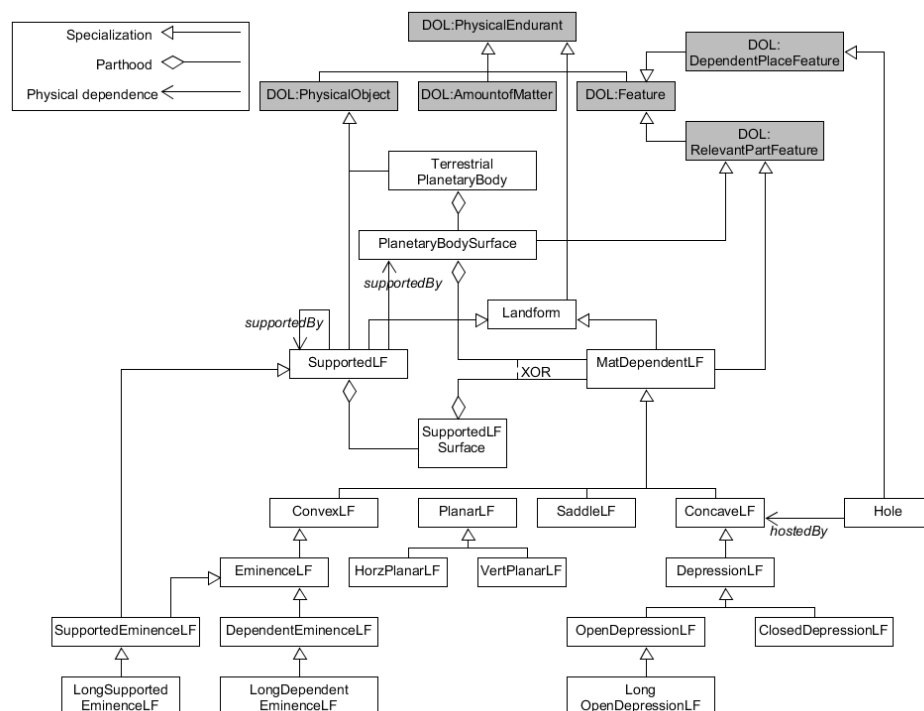

Figure 3. Landform categorization related to their shape, from the landform (LF) reference ontology [64] (by Sinha et al. licensed under CC-BY 3.0). DOL indicates categories from the DOLCE ontology.

One limitation of these approaches is that the shape of a landform is always defined within a context and terms such as "large" or "narrow" always depend on this context and the use of a generic landform ontology for classifying landforms from data is still a challenge. Hence, Guilbert et al. [65] propose to use salient elements of the terrain forming the backbone (or skeleton) of the landforms assuming that these elements can be translated into geometrical shapes (point or line) that

can be identified from a digital terrain model. The landform is then delineated by a region built around its skeleton. The model proposes to handle fuzziness by defining a core region surrounded by a larger region where its boundary is located (Figure 4). Guilbert and Moulin [66] describe a framework for applying this approach to landform description, illustrating it on specific landforms. However, the framework remains mainly a conceptual approach that needs to be implemented for generic purpose.
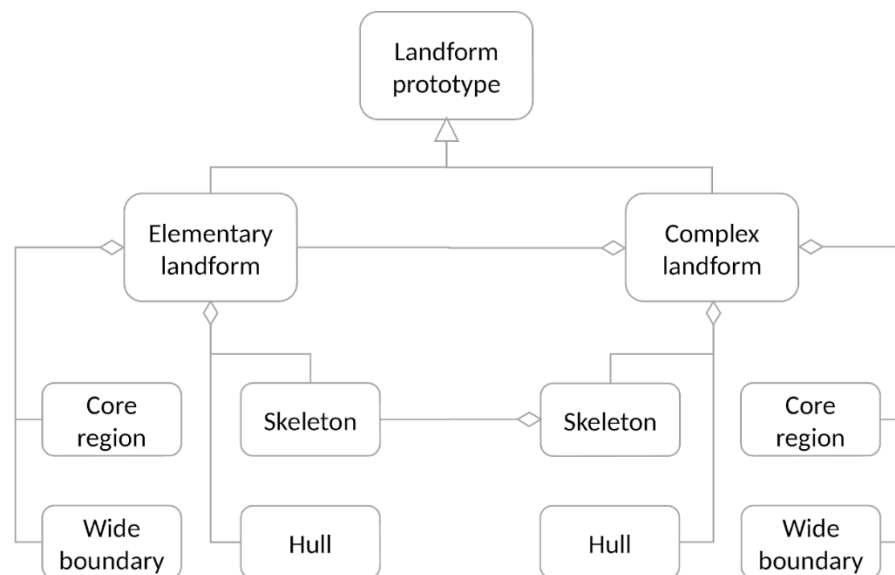
Figure 4. Landform description based on their saliences.

Domain ontologies also brought a lot of interest in remote sensing and Geographic Object-Based Image Analysis (GEOBIA). Most processing in remote sensing requires the classification of data (usually raster images) into significant segments based on geometrical, topological, and semantic attributes. While most methods are data-driven, the abundance of data in different formats and the need for generic approaches require more knowledge-driven approaches that would allow for the integration of expert knowledge in the process. For that purpose, domain ontologies are relevant since ontologies can store knowledge and help sharing this knowledge [67]. Ontologies can include both human concepts, based on cognition, and their translation into numerical concepts required for processing. By this way, a same concept can be mapped to different numerical definitions associated to different types of data. As mentioned by Arvor et al. (2019) [67], the use of ontologies does not improve classification results but provides a more reliable representation of expert knowledge. A geographical object is defined by its characteristics in a domain ontology. Each characteristic is then related to some characteristic values measured on the image (Figure 5). In the context of foreshore identification, Argyridis and Argialas [15] designed an ontology to formalize the implicit spectral, geometric, and spatial relationships described in the interpretation criteria, and employ them during identification. Rajbhandari et al. [14] also emphasize that ontologies allow for a better modularization of the methods: Common knowledge can be reused together with features more specific to an application. They also facilitate the automation of the classification since knowledge is transferred minimizing human intervention. They show in a case study that ontologies can be combined with machine learning approaches where the ontology stores generic knowledge and machine learning supplements the classification with specific rules. The application of some rules to the ontology often depends on the definition of some threshold values. These thresholds can be learned and applied for each specific case. The result of the classification can then be assessed by semantic similarity measures.
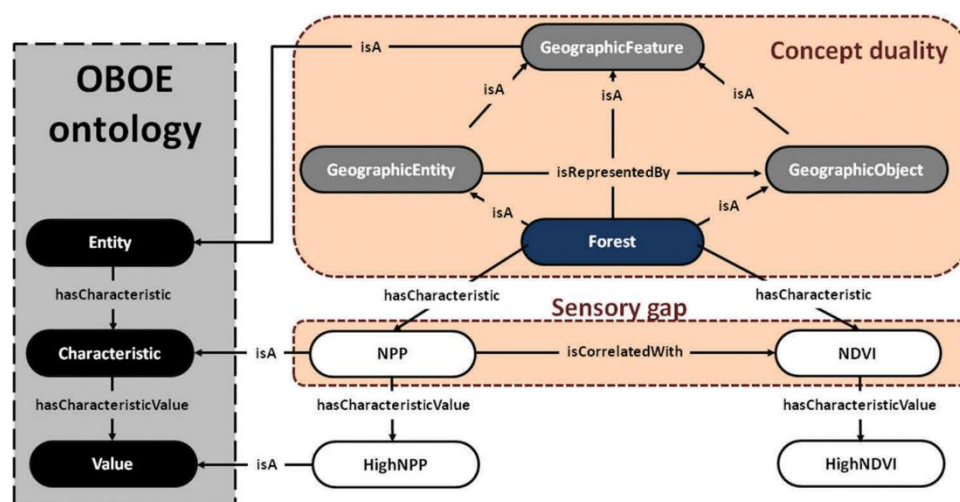
Figure 5. Example of the image representation of a concept for GEOBIA. The sensory gap is the gap between the real world and its computational description including indicators computed from the image (from [67], by Arvor et al. licensed under CC-BY-NC-ND).

In the last years, there has been a research shift towards the development of task and application ontologies to define, share, and reuse more specific knowledge. These ontologies are especially relevant for defining complicated geospatial tasks and operations or simulating geospatial workflows necessary to solve complex geospatial problems. They usually require the combination of other ontologies. For example, Zhuang et al. [68] developed an ontology-based approach using various ontologies (task ontology, process ontology, GIS operation ontology, interface ontology, data type ontology, GIS data ontology, and GIService-type ontology) to support a task-oriented knowledge base for modeling meteorological early-warning (MEW) analysis (Figure 6). In the context of web-service approaches, Hofer et al. [69] developed a knowledge base to support spatial analysis workflow development. Yue et al. [23] present a linked data approach for discovering geospatial resources in the Web of Data to build geoprocessing workflows. The approach leverages existing ontologies and vocabularies to semantically describe different types of geospatial resources, such as sensors, observations, raster data, and geospatial services.

Rospocher [70] outlines the methodology for creating the Personalized Environmental Service Configuration and Delivery Orchestration (PESCaDO) Ontology for supporting personalized environmental decision support. The methodology involves requirements specification, reuse of existing models and ontologies, such as GeoSPARQL (http://www.ogc.org/standards/geosparql) and PROVenance Interchange Ontology (PROV-O) (https://www.w3.org/TR/prov-o/), terminology extraction, formalization, revision, and documentation. The upper-level of the PESCaDO ontology includes three interrelated components: (a) The problem component describing the user request, activity, and profile; (b) the data component describing meteorological, pollen, and air quality data; and (c) the conclusion component for encoding warnings, recommendation, and suggestions provided by the system.

Application ontologies have also been developed for managing natural disaster information [71,72]. In this context, ontologies are designed not only as a knowledge base but rather as a tool for assisting decision making in emergency response. Zhong et al. [72] implemented an application ontology where further information is inferred from semantic and spatial data. Linyao et al. [71] propose a tool for retrieving and managing disaster-related data. The objective is to automatically link some sources by measuring spatial and temporal similarities. Both approaches provide a system where data are collected or disseminated through web services.
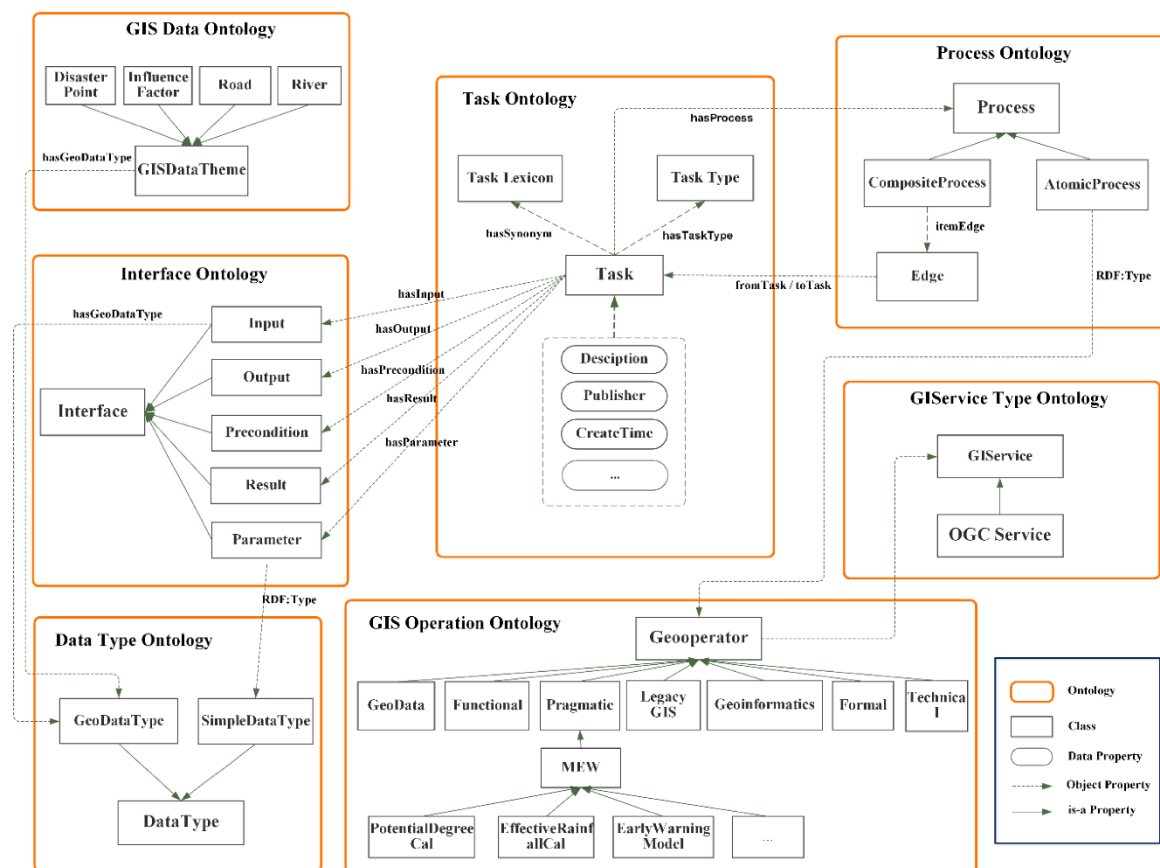
Figure 6. The relationships between different ontologies in a knowledge base for meteorological early-warning (MEW) analysis from [68], licensed under CC BY 4.0 by Zhuang et al.

Ontologies at different levels of generality are also built to formalize and support reasoning over cartographic knowledge. Earlier works mostly focused on the formalization of the cartographic generalization process [73] classifying cartographic features, spatial relations, and relational constraints to be observed by the generalization process. Gould and Mackaness [74] also include generalization algorithms in their ontology in order to have a knowledge base that could be used in an automated system. Yan et al. [75] propose an ontology for the generalization of isobaths on charts. In their ontology, they not only store cartographic elements such as soundings and isobaths, but also include features formed by groups of isobaths and soundings representing submarine features as perceived by the readers. They also integrate generalization constraints and operators that can apply to these larger features. The ontology was implemented in a triplestore and used to select and perform operations in a multi-agent system. However, these works were mostly considering the development of a knowledge base to support map generalization. More recently, Varanka and Usery [76] consider the map itself as a knowledge base and propose an ontology that would not only include data and design concepts but also semantic and logical knowledge that are also embedded in the map. They also present an architecture based on a triplestore where map elements can be retrieved with SPARQL and GeoSPARQL queries. Huang and Harrie [77] propose a broader model including visualization knowledge and new concepts describing scale and portrayal. They propose a system architecture with a web server retrieving geospatial and portrayal data from a knowledge base and producing a map with the required style for a client application. Hahmann and Usery [78] present a first-order logic formalization of contour semantics to support qualitative and quantitative reasoning about contours. The formalization comprises four fundamental concepts: Contour regions, contour lines, contour values, and contour sets, as well as their subclasses and relations between them. The ontology is developed in first-order

logic to support the detailed ontological analysis of contour semantics in general but is also translated into an OWL ontology for storing and querying information of particular contour maps.

From this review, we can see that, while geospatial ontologies were firstly designed as a tool to formalize expert knowledge, the last few years saw a growing focus towards the integration of domain and task ontologies as computational ontologies into larger systems for data retrieval and analysis. Latest works on geovisualization and disaster management make use of ontologies to develop interoperable systems and assist users with tailor-made solutions.

2.3. Ontology Design Patterns and Lightweight Ontologies

Ontologies presented above are referred as heavyweight ontologies. They are completely developed ontologies describing a whole domain or a domain concept. Recent developments are also concerned with the design of more modular ontologies applicable to different problems. This issue is addressed by two kinds of ontologies: lightweight ontologies and ontology design patterns (ODP).

The distinction between heavyweight and lightweight ontologies is based on the expressivity provided by ontologies: lightweight ontologies have restricted expressivity as they provide the simplest formalization of the simplest model of a domain, adequate for the task at hand; they typically consist of a hierarchy of concepts and a set of relations holding between those concepts [79]. Lightweight ontologies are also proposed as an alternative to foundational and domain ontologies in order to capture the semantics of a domain. Although they lack the expressivity of heavyweight ontologies, they are especially useful for supporting connectivity and interoperability across communities and platforms in the context of the Semantic Web and linked data [80] and may be applied in key applications such as document classification, semantic search, and data integration [81].

A lightweight ontology can be used to extend or integrate other ontologies. Hasan et al. [82] defined a lightweight ontology for earthquake engineering that was integrated to WordNet. Hong and Kuo [18] propose the development of lightweight ontologies to semantically integrate concepts from two domains (topography and land use). Natural language definitions of concepts are transformed into structured representations which are then compared and associated on the basis of four types of semantic relationships (exact, subset, superset, overlap, and null). A bridge ontology is used to formally represent the semantic relationships between the concepts of the lightweight ontologies.

Kuai et al. [83] propose a lightweight ontology for mapping concepts from topographic maps in English and Chinese. They define a hierarchy of concepts from natural language definitions. Concepts from both languages are then related based on their similarity. Kordjamshidi and Moens [84] propose a lightweight spatial ontology for locating objects in space based on a spatial annotation scheme. The ontology consists of spatial concepts (trajectory, landmark) and spatial relations (region, direction, and distance). Its aim is to define mappings between cognitive-linguistic spatial concepts in natural language and qualitative spatial representation models. A global supervised machine-learning model for ontology population is used to implement the mappings between natural language and the lightweight spatial ontology.

More recently, Couclelis [2] proposed a conceptual model for the development of micro-ontologies for specific user needs and purposes encompassing three interrelated views of information: (1) Measurements and formal operations on these, (2) semantics for defining the meaning of information, and (3) context for the interpretation and use of information. The model encompasses a process called "semantic contraction" which generates a sequence of representational layers varying in semantic richness to support the varying perspectives and interpretations of information and the resulting representational and informational requirements of different users.

The specification of ontology design patterns is pursued as an alternative to ontology design and more profound ontology integration approaches when semantic heterogeneities and conflicting requirements do not allow for the development of a common, integrated ontology. Ontology design patterns or content ontology design patterns are small, modular ontologies that may be used as building blocks in the ontology design process [85]. They facilitate ontology reuse, evaluation, alignment,

and mapping across domains. For example, Calafiore et al. [86] defined two ODP modeling the built environment and social behavior. Both ODP are then integrated together to mine social behavior patterns in urban areas from crowdsourced data.

The GeoLink project [19] has developed a framework for integrating seven data repositories mainly in the domain of ocean science using semantic technologies such as linked data and ontology design patterns (ODP). The GeoLink oceanography ontology [20] is a collection of ODPs used to model notions such as person, organization, dataset, cruise, feature, geofeature, place, etc. A pattern may also be aligned to another pattern or to an external ontology.

Because they are modular, ODP have also been used in topography and landform description. Indeed, the definition of a landform varies with the application domain and the perception of the user or the expert. Hence, instead of building a complete landform ontology, several authors propose to apply ODP for the description of the topography in different contexts. Sinha et al. [87] defined an ODP for surface water features that can be used to describe different water bodies with their shape and flow. Guilbert and Moulin [66] propose a landform ODP based on landform saliences together with a conceptual framework. They apply their approach on the modeling of landforms based on existing models. Such ODP allow for more flexibility for the description of water features and landforms. However, their usability has not been demonstrated yet by designing and integrating new ontologies based on these ODP.

Finally, Janowicz et al. [88] propose the Sensor, Observation, Sample, and Actuator (SOSA) lightweight ontology that provides specifications for modeling interactions with sensors and samplers. It is now a W3C recommendation and OGC standard. SOSA is based on an ODP that designs the core structure from which different perspectives are derived. The ODP provides a pattern for designing lightweight ontologies that can be integrated into larger ontologies.

2.4. Bottom-Up Ontological Approaches

Geographic ontologies are traditionally defined top-down by authoritative organizations or groups of experts using semantic modeling approaches. However, the amount of information available today has shifted the focus to bottom-up approaches for enriching existing geographic ontologies. Different approaches can be found including semantic information extraction and text mining techniques.

Bennett and Cialone [89] describe a methodology, called corpus-guided sense cluster analysis, for ontology development which combines two different modes of investigation: (a) Logic-based, formal semantic analysis of concepts and relations of expert knowledge; and (b) corpus-based statistical analysis of the actual use of terminology in natural language texts highlighting the range and frequency of senses associated with a lexical term. Although the methodology is general, it focuses on the construction of spatial ontologies: spatial entities like surfaces and cavities and spatial relations like "surrounding", "enclosing", and "containing" are used for the explication of the proposed approach. Instead of adhering to a single precise definition, the methodology is based on the notion of sense cluster where the referent of a conceptual term is modelled by a probability distribution over a set of precise definitions.

Hu and Janowicz [21] proposed a workflow to mine bottom-up geographic knowledge from the Linked Open Data (LOD) (https://lod-cloud.net/) cloud in order to enrich existing geographic ontologies developed top-down by groups of experts with diverse perspectives provided by general users. Geographic knowledge is expressed through instances and property-value pairs attributed to relevant target categories.

Zhu et al. [90] propose a bottom-up, data-driven approach for identifying similarities and differences in the semantics of geographic feature types to supplement existing top-down alignment methods. Three major gazetteers are used to demonstrate the proposed approach: DBpedia places [91], GeoNames [92], and Getty Thesaurus of Geographic Names [93]. The semantics of gazetteers are examined not on the basis of feature type terms or definitions but on the basis of instances belonging

to them. Three kinds of spatial statistical features are extracted from instances: spatial point patterns, spatial autocorrelations, and spatial interactions with other geographic features.

Kokla et al. [94] used a top-down and a bottom-up approach for enriching and populating a geospatial ontology in order to enable semantic information extraction. The top-down approach is applied in order to incorporate knowledge from existing ontologies. The bottom-up approach is applied in order to enrich and populate the geospatial ontology with semantic information (concepts, relations, and instances) extracted from domain-specific web content.

2.5. Modeling vs. Encoding

While ontologies are supposed to be designed for formalizing knowledge in view of using it in some implementation, Kuhn [95] pointed, in the Semantic Web context, that ontology tasks should be separated into: modeling and encoding. Through this review, one can see that ontologies were indeed developed with a focus on either of these tasks (Table 1). On one hand, ontologies were built to formalize expert knowledge in a specific area to provide unambiguous concepts and rules. For example, landform ontologies such as [63,64] propose concepts to structure landforms in lattices while [76] provides a formalization of existing rules for contour maps in first order logic. In such approaches, ontologies are used to constrain the interpretation of the different concepts.

On the other hand, domain ontologies were built to provide implementation standards [88] and facilitate data processing [67], or even, for bottom-up approaches, were directly built from data. They provide design patterns that allow for better transfer of knowledge from concepts to data dependent models and for data operability with the use of standard patterns. Hence, one major novelty in these last few years was the application of ontologies in other areas such as remote sensing and image processing where traditionally, semantics was given limited consideration. Ontologies are used to associate some concept with their data representation. For example, vegetation is associated with a vegetation index and a topographic eminence is associated to an elevation difference.

Table 1. Ontologies classified according to their main focus towards modeling or encoding.

Modeling	Encoding
Hydrology: Stephen and Hahmann 2017 [61]	Hydrology: Kuai et al. 2016 [83]
Forest: Álvarez and Bennett 2017 [56]	Oceanography: Krisnadhi et al. 2015 [20]
Landforms: Grenoble et al. 2019 [63], Sinha et al. 2018 [64], Guilbert et al. 2016 [65], Guilbert and Moulin 2017 [66]	Natural disasters: Linyao et al. 2015 [71], Zhong et al. 2017 [72]
Urban environment and place: Ballatore 2016 [57], Calafiore et al. 2017 [58], Garbacz et al. 2018 [59], Calafiore et al. 2018 [86]	Social and Physical environment: Gharebaghi and Mostafavi 2016 [60]
Cartography: Hahmann and Usery 2015 [78]	Cartography: Yan et al. 2015 [62], Touya et al. 2014 [73], Gould and Mackaness 2016 [74], Yan et al. 2017 [75], Varanka and Usery 2018 [76]
	GEOBIA: Rajbhandari et al. 2017 [14], Arvor et al. 2019 [67]
	Sensors: Janowicz et al. 2019 [88]
	Semantic integration: Hong and Kuo 2015 [18]
	Ontology enrichment: Hu and Janowicz [21]

Within this five-year period, the first few years mainly saw works related to modeling without much concern on the implementation while these last two years saw an interest in providing workable solutions. Hence a "qualitative-quantitative divide" [96] still exists but some works such as [71,77] show promising implementations. Recent research on GEOBIA shows that ontologies are no longer

the monopoly of a specialized community but are also of interest for researchers looking for solutions to cope with the inherent vagueness to common problems such as vegetation mapping or landform classification. However, as mentioned in [56], vagueness can be of two kinds: conceptual vagueness and threshold (or sorite) vagueness. Ontologies provide a way to handle the first one while, in [14] the second is handled with machine learning approaches.

While machine learning can achieve good prediction in setting threshold values for different attributes, their computation is fully based on data. It; thus, requires training for each type of data (according to data resolution, source …) and does not provide any elicitation of this vagueness at a conceptual level. Indeed, defining threshold values for distinguishing what should be big or small for example depends on the context in which the analysis is done. Domain ontologies on the modeling side can be designed to provide this contextual information. Hence, we believe that a better integration between "modeling" and "encoding" ontologies would help to define thresholds according to properties obtained from the concepts, leading to some semantic thresholds rather than data-driven thresholds.

3. Semantic Information Elicitation

In the last years, research has focused on eliciting meaning from the considerable amount of semi-structured content, such as html pages and metadata records and unstructured content, such as scientific reports, news articles, travel blogs, and historical archives. These sources provide a wealth of information on geospatial concepts, places, events, activities, etc. The term elicitation in the context of this paper is used to encompass a set of related processes that focus on making this knowledge explicit and discoverable: semantic information extraction, enrichment, and search.

Information extraction supports the automatic processing of unstructured or semi-structured resources and the identification of relevant entities, concepts, and relations. In addition to the extraction of these semantic elements, topic modeling techniques unveil latent abstract topics and semantic associations from large text collections.

Ontology-based information extraction (OBIE) [97,98] is an emerging subfield of information extraction, in which the information extraction process is guided by an ontology as a means to formally describe domain knowledge and assist the extraction of pre-defined concepts, properties, and instances. OBIE includes: (a) Approaches that use an ontology as a guide to acquire knowledge from unstructured and semi-structured text; and (b) approaches that seek to build the ontology by processing texts, also known as ontology learning and population approaches [98].

Eliciting places, events, concepts, relations, topics, etc. from texts in natural language and connecting these to their meaning through ontologies and knowledge bases supports semantic enrichment (or annotation or tagging), i.e., linking the content to other relevant resources based on an understanding of what each term is about. It is also significant for semantic search based not only on keywords used to formulate a query, but also on the meaning of terms behind the query.

Furthermore, it is used for computing semantic similarity and applying semantic analysis, classification, and spatialization techniques to explore numerous cognitive and semantic aspects, such as the relations among places, the historical evolution of cities, the progression of physical phenomena and social events, and people's perception of landscapes and regions.

Data-driven geospatial semantics refers to a bottom-up approach for eliciting geospatial knowledge from natural language texts that contain either explicit or implicit locations of places [99]. It is distinct from top-down or expert-driven approaches that extract geospatial knowledge from experts in a specific field. In terms of methodology and scientific tools used, data-driven approaches are also contrasted with conventional human-participants studies that elicit cognitive aspects, such as the meaning of geographic categories [100–102] or the representation of cognitive regions [103]. However, the increased availability of data and their analysis using statistical and data mining techniques does not necessarily ensure meaningful knowledge discovery, a task that presents many challenges and requires the support of synthetic and semantic tools [104].

3.1. Geospatial Semantic Information Extraction and Enrichment

Semantic information extraction aims at eliciting salient, specific types of information from unstructured or semi-structured data sources [105]. Entities, concepts, and/or semantic relations that are implicit in a given source are made explicit to support semantic annotation, content-based exploration, semantic search, and data-driven geographic analysis. Figure 7 shows spatial entities (in blue), spatial concepts (in green), and relations (underlined) that may be extracted from a small passage.
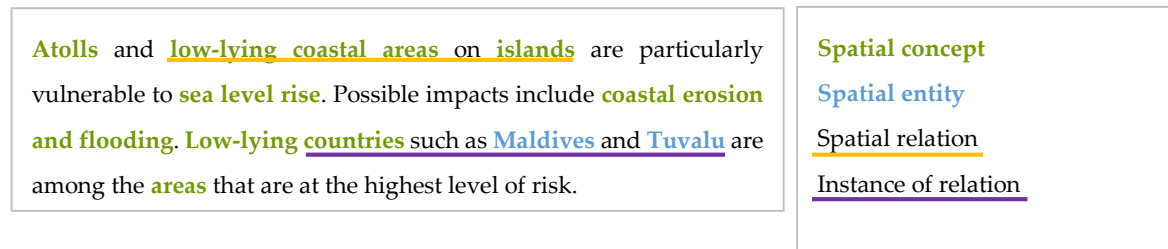

Atolls and <u>low-lying coastal areas</u> on islands are particularly vulnerable to **sea level rise**. Possible impacts include **coastal erosion and flooding**. **Low-lying** <u>countries</u> such as <u>Maldives</u> and <u>Tuvalu</u> are among the **areas** that are at the highest level of risk.	**Spatial concept** **Spatial entity** <u>Spatial relation</u> <u>Instance of relation</u>

Figure 7. An example of spatial entities, concepts, and relations extracted from the sentences on the left.

Named entities in this context refer to specific real-world objects or instances, such as persons, locations, and organizations, denoted by proper nouns. For example, in Figure 2, Maldives and Tuvalu are named spatial entities. Named entity recognition (NER) [105] is used to identify named entity mentions in text and classify them in pre-defined categories. Gazetteer lists that provide words or phrases representing individual instances of a specific category (e.g., location, time, and organization) are widely used in NER tasks. Named entity disambiguation (NED) [105] is used to determine the identity of a named entity mention and link it to a unique entity in a knowledge base. NED is crucial when the same entity mention may refer to different instances, e.g., the entity mention "Washington" may refer to different real-world entities such as George Washington, the State of Washington, or Washington D.C.

Concept extraction techniques are used to identify words or phrases that denote semantic classes, such as the concepts: "atoll", "coastal area", and "island" in Figure 2. Concept extraction involves related processes such as term extraction, keyword/key phrase extraction, and topic modeling [105]. These processes are applied on a text collection, which usually describes one or several domains. Term extraction provides insight into the core concepts of the domain(s) that the text collection is about. Keyword / key phrase extraction focuses on extracting domain concepts that describe a given text. Topic modeling aims at analyzing co-occurrences of related keywords that describe higher-level semantic topics of the text collection.

Concept extraction usually relies on pre-processing steps such as tokenization to split text into words, phrases, symbols, or other meaningful elements called tokens, sentence splitting to divide the texts into sentences, part-of-speech (POS) tagging [106] to mark up each phrase as corresponding to a particular part of speech, i.e., noun phrases, verb phrases, adjective phrases, adverb phrases, etc., and lemmatization to identify the base or dictionary form of a word (lemma). Then, some approaches called window-based, consider a window around terms and examine contextual information within this window to judge whether the term may indeed represent a specific concept. Other approaches define a set of rules in the form of regular expressions and lexico-syntactic patterns for the identification of predefined domain concepts.

Topic modeling techniques [107] such as latent Dirichlet allocation and latent semantic analysis (or else latent semantic indexing) are text mining [108] techniques employed to identify terms that frequently co-occur in given collections of documents. These term co-occurrences form clusters which represent latent abstract topics, e.g., the words cyclone, hurricane, typhoon, blizzard, and thunderstorm may co-occur under a meteorological disasters' topic. These techniques are used to identify the relevant context of a given corpus and aid the classification, search, and retrieval processes.

Relation extraction techniques are used to identify semantic relations between concepts and entities (i.e., hypernymy, synonymy, relatedness, etc.). For example, in Figure 7, Maldives and Tuvalu may be defined as instances of the concept "low-lying countries". Various techniques are used for the extraction of predefined relations based on text processing, mapping to external resources such as ontologies, thesauri, and computational lexicons or on a combination of both. In the first case, key methods involve: (a) the identification of the head/modifier terms of a phrase (e.g. the head of the phrase "low-lying countries" is "countries", while low-lying is a modifier that specializes the head and creates a hyponym, i.e., a subclass); and (b) formulation of lexico-syntactic patterns. For example, the Hearst pattern) [109] "Low-lying countries such as Maldives and Tuvalu.." signifies that Maldives and Tuvalu are instances of "low-lying countries". In the second case, ontologies or other external sources such as WordNet and DBpedia provide predefined relations between concepts (e.g., hypernyms/hyponyms, synonyms, and related concepts). However, identifying the correct sense of a term, called word sense disambiguation [110] is a very challenging issue greatly influencing the outcome of relation extraction. The extraction of a broader range of relations between entities is even more demanding requiring the combination of the previous mentioned techniques with techniques such as distant supervision, extraction of syntactic dependencies between entities, and calculation of the frequencies of n-grams occurring in the text window surrounding both entities [105]. For an extensive survey of information extraction techniques in a Semantic Web context refer to [105]. For comprehensive reviews of OBIE systems and the contribution of ontologies in information extraction refer to [97,98]. Figure 8 shows an overview of semantic information processes and extracted elements.
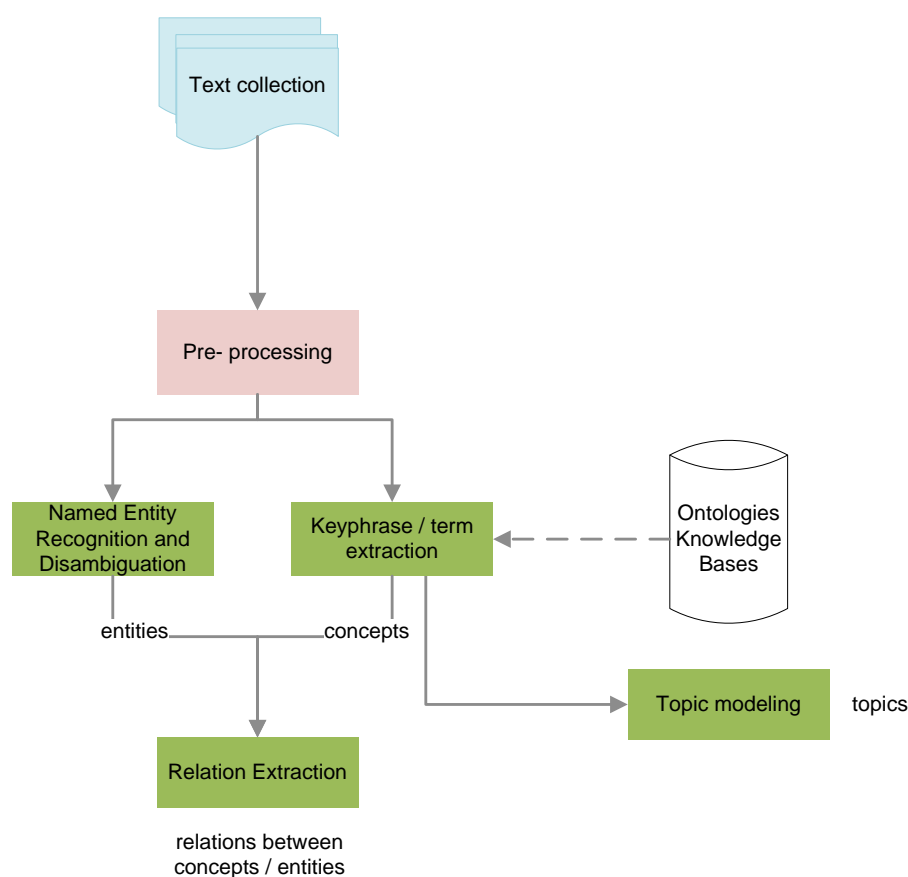

Figure 8. Overview of semantic information processes and extracted elements.

In the geospatial domain, place names (or toponyms) are the principal type of knowledge extracted from natural language texts using a process called geoparsing [111]. Geoparsing involves both the identification of place names in text and their disambiguation in order to be linked to unambiguous

spatial references using geographic gazetteers. For an extensive survey of geoparsing challenges and techniques refer to [112,113]. Besides place names, some applications, such as the analysis of historical archives, also require the extraction of temporal expressions and person names that refer to historical personalities. Regarding relations, emphasis has been placed on identifying spatial relations between objects for interpreting indirect geographic references and for associating an entity or event to a reference location. In natural language texts, spatial relations are usually specified qualitatively using spatial expressions such as at, close to, to the left of, north of, etc. These expressions convey various spatial relations, whose meaning is often vague and context dependent hindering their automatic extraction and disambiguation [114,115].

Wallgrün et al. [116] built a corpus of natural language expressions extracted from web documents for analyzing and modeling Spatial Relational Expressions (SREs). The corpus consists of geo-referenced triplets which contain a located object, a spatial relation and a reference object. Mirrezaei et al [117] developed TRIPLEX-ST, a system for the extraction of spatio-temporal information about entities and their properties from textual resources. TRIPLEX-ST is based on a distantly supervised approach, which uses linguistic annotations (e.g., dependency relations, named entities, and lexical constraints) together with information available in existing KBs (e.g., DBpedia and YAGO knowledge bases), to extract triples associated with temporal and spatial contexts and infer templates that capture spatiotemporal information from unseen sentences. Each template includes dependency relations, POS tags, named entities, WordNet synsets, subject and object types, and syntactic/semantic restrictions of verbs in dependency paths.

Dittrich et al. [118] presented a classification schema for disambiguating spatial from non-spatial uses of an extensive list of English prepositions such as above, below, next to, in (the) front of, north of, etc. The classification schema involves rules for identifying and excluding abstract entities (abstract locatum or abstract relatum) involved in locative expressions based on knowledge sources such as WordNet and DBpedia and non-spatial uses of prepositions such as a temporal relation, the material of an object, the agent of an action, or the topic of communication based on dependency trees provided by a natural language processing (NLP) parser. Radke et al. [119] modify a machine-learning approach for generic spatial role labelling [120] to automatically extract prepositions that are used specifically in a geospatial sense and distinguish geospatial uses of prepositions from other spatial and non-spatial uses. A geospatial relation is defined as one in which the preposition has a spatial sense and the reference object to which the preposition applies is a named place or a geographic feature type. Place names are identified using Geonames gazetteer, whereas geographic feature types are identified using a dictionary of geographic feature types. The approach is evaluated on the basis of a corpus of 1876 instances of preposition usage manually annotated as geospatial, spatial (but not geospatial), and non-spatial.

Derungs and Purves [115] extract spatial expressions of the nearness relation using an analysis of n-grams—contiguous sequences of tokens retrieved from a large corpus of the form A near B where A and B are both place names, referring to different spatial granularities. The analysis of n-grams allows estimations of the probabilities of phrases expressing the nearness relation and the exploration of what is considered to be near at different scales.

The combined extraction of spatial and temporal information from documents supports the discovery of complex information such as events or trajectories from historical documents or travel literature. Abraham et al. [121] developed an approach for extracting spatial, temporal, and attributive information for analyzing the movement of rival troops during the anti-colonial resistance war of 1904 in Namibia from historical documents. Named entity recognition is used for the extraction of place names, data expressions, and person names based on self-tailored gazetteers for tagging historical place names and temporal expressions that are not included in traditional gazetteers. Pattern-based rules are formulated to extract spatio-temporal relationships.

Wang and Stewart [16] extracted spatiotemporal and semantic information on natural hazards from web news reports. The semantic information extraction process was based on a hazard ontology manually constructed from authoritative sources on hazards, which was integrated with spatial,

temporal, and semantic gazetteers that capture these three aspects of hazards. The extraction process was implemented with GATE 8.0 (General Architecture for Text Engineering) [122] and involved several steps: linguistic processing, named entity recognition to identify spatial and temporal information, manual annotation to build the semantic gazetteer, mapping the terms from the semantic gazetteer to the corresponding ontology classes, rule-based association of events to spatial and temporal information, and geocoding the results for their subsequent visualization.

In order to explore the emotional structure of place classes such as city, forest, and road, Ballatore and Adams [123] developed a vocabulary of place nouns of natural and built places based on DBpedia [91], GeoNames [92], and WordNet [39]. A natural language processing (NLP) technique was implemented to extract place emotions from a corpus of travel blog posts based on the emotion vocabulary WordNet-Affect [124]. These are used for constructing emotion vectors based on term co-occurrences within a context window around place terms.

Egorova et al. [125] developed linguistic rules based on morpho-syntactic patterns for the automated extraction and classification of three types of fictive motion (actual motion of the observer, description of a vista somewhere along the way or encyclopedic knowledge) based on an annotated corpus of alpine narratives.

Deriving spatial concepts and entities from texts also relates to the research on geospatial semantic analysis and spatialization of text corpora. Derungs and Purves [126] analyzed a large corpus describing Swiss alpine landscapes in order to explore the description of landscapes. The approach extracts toponyms using the Swissnames (https://shop.swisstopo.admin.ch/en/products/landscape/names3D) gazetteer which are linked to nouns referring to natural features using linguistic processing and manual annotation based on predefined rules. In another study of people's perceptions of landscapes and their prominent characteristics, Wartmann et al. [127] collected landscape descriptions of five landscape types across three different sources (participant free lists obtained through interviews, hiking blogs, and Flickr tags). The sources were manually annotated based on a coding scheme including toponyms, biophysical, cultural, and perceptual landscape elements, activities, sense of place, and people. The comparison of sources based on coded terms using cosine similarity showed that descriptions from the same source were significantly more similar irrespectively of the landscape type.

Cooper et al. [128] proposed Geographical Text Analysis (GTA) for the spatialization and analysis of digital texts. The information extraction process identifies place names, as well as thematic tags representing topics (e.g., education, warfare, and farming using the UCREL Semantic Analysis System (USAS)). Bruggmann and Fabrikant [129,130] extracted place names (toponyms) and temporal information from an online dictionary about Swiss history in order to visualize temporal relations between Swiss toponyms using spatialization techniques. Place names are extracted using the Swissnames gazetteer and temporal information (such as dates and periods of time) is extracted using HeidelTime temporal tagger. Salvini and Fabrikant [131] used semantic information analysis and spatialization techniques of Wikipedia content for the construction and the empirical investigation of a multi-relational world city network. User-generated tags, which categorize the content of the articles and their organization using hypernymic and hyponymic relationships, are used to compute similarity relationships between shared articles and group them into thematically relevant clusters using the topic modeling technique.

Semantic enrichment (also known as semantic tagging or semantic annotation) is a related process aiming at adding semantic metadata to help machines make sense of the content and reveal latent relations and other semantic information. It is used for information organization and retrieval, semantic search and knowledge discovery, and ontology development and population. Semantic enrichment has been used to add semantic metadata to different types of content, including: unstructured documents [132], maps [133], images [134,135]), metadata [136], and videos [137]. Semantic metadata describe the meaning of content in terms of abstract concepts and entities, such as people, things, and places.

Lemmens et al. [22] discuss ways to semantically enrich unstructured VGI content with concepts from both informal structures (folksonomies) and formal structures (ontologies) in order to support information retrieval. Tardy et al. [135] present a method for the semantic enrichment of photo tags with place characteristics such as feature type, function, use, shape, material, appearance, etc., paying special emphasis on places in urban areas with a small number of photos. The method uses a combination of geometric and linguistic techniques to classify the tags and then select those that would most likely semantically enrich feature descriptions. In order to improve the classification of events in tweets, Romero and Becker [138] use semantic enrichment to identify entities and relevant vocabulary from tweets and related web pages and associate these features with concepts extracted from the LoD cloud. A pruning technique is then applied in order to discard too generic or too specific semantic features and select those with the most discriminative power for event classification.

Besides unstructured and semi-structured content such as texts, tweets, and photo tags, other types of spatiotemporal data, such as trajectories and movement data can also be semantically enriched. In this context, trajectories and movement data are not considered as mere sequences of geographic coordinates but as patterns with meaning that needs to be made explicit in order to support a more profound interpretation of the movement performed [139]. The Baquara2 framework is developed to semantically enrich and analyze movement data based on a multilevel hierarchical ontological model [140]. The semantic enrichment process annotates movement segments with references to concepts and instances from ontologies and linked open data collections (e.g., DBpedia, LinkedGeoData (http://linkedgeodata.org)), which describe the place, event, goal, or environmental conditions of the movement segment.

3.2. Semantic Search and Knowledge Discovery

Semantic-based search is a prevalent research topic in the context of the Semantic Web aiming to overcome the limitations of keyword-based search. In traditional keyword-based approaches, resources that are semantically related to a user's query but described differently from the query keyword are considered irrelevant and are excluded from the search results. Semantic-based search seeks to identify relevant sources by not only matching the keywords used but also based on the meaning behind a user's query. This could potentially result in identifying relevant sources that do not necessarily contain the keyword used in the query. For example, a semantic search for resources related to "geological phenomena in the Mediterranean" would ideally return resources including "earthquakes in Ionian Islands", "volcanic eruption of Mount Vesuvius" and "soil erosion in Spain" although these resources may not explicitly mention the terms "geological phenomena" or "Mediterranean".

The semantic description of both source data and users' queries is critical for achieving this and associating a search request with the most relevant source. Semantic information extraction is used to educe this semantic description from unstructured or semi-structured sources, whereas semantic annotation is used for query expansion, retrieval, and ranking. Query expansion consists in selecting and adding relevant terms to the user's query for minimizing query-document mismatch and improving retrieval performance [141]. An ontology's conceptual hierarchy enables the generalization / specialization of semantic search based on concepts and entities and the relations between them, especially hypernymy/hyponymy and synonymy relations.

Geographic information retrieval (GIR) [142,143] is another relevant field which deals with the automated interpretation of place names and spatial relationships in queries and in documents and with indexing, relevance ranking, and retrieval of the relevant content. GIR faces significant challenges such as detecting geographical references and associated spatial natural language qualifiers, disambiguating place names or other geographic information, and ranking resources with respect to spatial, temporal, and thematic relevance. Although GIR is a distinct research field beyond the scope of this paper, some aspects of semantic search are also relevant to GIR, such as the semantic description of source data and users' queries, query expansion, and semantic relatedness.

Semantic search in the geospatial domain is based on different approaches: top-down ontological approaches, bottom-up data mining approaches, or a combination of both. In the first case, ontologies are used to interpret users' queries and enrich them with other meaningful terms using ontology concepts and their between relations [144,145]. Concept-based information retrieval represents both documents and queries using semantic concepts instead of keywords and performs retrieval in concept space [146]. Synonymous terms, as well as more general, more specific or semantically related concepts are used for query expansion. In the second case, topic modeling techniques such as latent Dirichlet allocation and latent semantic analysis are used to analyze large corpora and discover abstract topics that co-occur in these collections of documents.

De Andrade et al. [147] propose a framework for improving geographic information retrieval in spatial data infrastructures. The framework includes a tagging process for enriching metadata with information about the space, theme, and time related to each feature type offered by a service. Thematic tagging relates the themes of feature types to ontology concepts. To facilitate semantic retrieval of remote sensing services, Nys et al. [148] created a self-learning knowledge graph that structures the concepts used in related queries. Natural language processing techniques (part-of-speech tagging and lemmatization) are used to reduce the complexity of both queries and service descriptions. Queries are then expanded based on the UNESCO thesaurus and a subgraph of the nearest neighbors of selected concepts is extracted including broader, narrower, and related concepts. The spatial aspect of queries is also used for expansion using the location on a map and administrative subdivisions provided by GeoNames.

Zaila and Montesi [149] developed GeoNW, a geographic ontology combining GeoNames, WordNet, and Wikipedia to support toponym extraction and disambiguation. The ontology includes physical or administrative places associated with synonyms derived from GeoNames and WordNet, and nationality adjectives from Wikipedia. This information is subsequently used to identify related terms of a geographic entity in a document. This approach is based on the assumption that the more related a geographic entity is to other geographic terms in the document, the more likely it is that the geographic term is associated to the document.

In order to enable semantic search and knowledge discovery for ArcGIS Online, Hu et al. [133] designed a specific ontology based on the ArcGIS Online schema and used two semantic annotation systems, DBpedia Spotlight [150] and OpenCalais [151], to extract entities and classes from map titles and descriptions. Semantically similar terms were also taken into account for query expansion based on the UMBC Semantic Similarity Service [152]. Jiang et al. [153] proposed an approach to discover semantic relationships between domain-specific vocabularies in order to obtain the semantic context of a user's query. The approach integrates the results from four different methods: user search history analysis, clickstream analysis, metadata analysis, and ontology concept similarity and is implemented in oceanographic data discovery.

Li et al. [154] developed a semantic search tool integrated in GeoNetwork (https://geonetwork-opensource.org/) to support search and retrieval of polar datasets. The semantic search tool is based on latent semantic indexing (LSI) for full-text indexing, searching, and ranking relevant metadata records. Semantic associations between dataset metadata terminologies are identified and stored in a semantic matrix maintained within the catalogue. The tool also adopts a revised cosine similarity measure for ranking relevant results.

Frankenplace [155] is an interactive thematic map search engine that supports exploratory search tasks. An indexing approach using a discrete global grid and term boosting as well as topic modeling are used to support information exploration at different scales (zooming in and out) both geographically and thematically.

Huang [17] combines ontologies with latent semantic analysis (LSA) in order to categorize geographic features from text documents. LSA is used to capture the semantic context of terms, while ontologies are used to represent domain knowledge and support the extraction of terms from text, as well as the identification of query terms that represent the predefined categories.

Cross-linguistic information discovery is also an issue attracting attention especially in the context of geoportals. The use of different languages both for the documentation of data sources and the formulation of queries by different users further complicates the process of semantic search and knowledge discovery. Multilingual thesauri and ontologies are used to enrich metadata with synonyms and term translations, as well as query terms for improving metadata discovery [156]. Laurini [157] uses geometric characteristics and toponyms derived from gazetteers to formulate inference rules for matching multilingual ontological concepts. Since toponyms and concept types across different languages are not always strictly equivalent, the matching is based on homology relations which are reflexive and symmetric, instead of equivalence relations which are reflexive, symmetric, and transitive.

The notion of semantic similarity is central to information search and knowledge discovery. Similar entities and concepts are used for linking data among different sources, answering user queries, and disambiguating word senses and place names. Adams [158] proposed an observation-to-generalization model that distinguishes between observed attributes of the environment at a specific location and generalized attributes about places that are inferred from these observations. Within this model, a place is defined by a six-tuple: a set of toponyms, a type, a set of spatial footprints, a set of associated observations, a set of generalizations, and a set of relations to other places. The model and a suite of operations based on the invariance of generalized place attributes are used to address the problem of similar-place search.

Kim et al. [159] proposed an approach for extracting and disambiguating place names from unstructured place descriptions. The approach uses natural language parsing to extract place names and spatial relations between them from triplets consisting of an object to be located, a reference object and their spatial relation [160,161]. Graph matching based on string, linguistic, and spatial similarities between places is further used to find corresponding spatial objects.

Besides semantic similarity, the notion of semantic relatedness has also begun to attract the attention of GIScience. Semantic relatedness deals with concepts that are not necessarily similar, but are somehow related through various types of relations, such as the part-of relation between forests and trees and the caused-by relation between tsunami and earthquake.

Semantic similarity/relatedness have been studied in the context of spatial data infrastructures to identify similar geographic data sources on the basis of their metadata records [162]. The similarity/relatedness measures applied are usually based on external sources like thesauri and computational lexicons such as WordNet and also take into account the semi-structured descriptive information of metadata records such as title, keywords, spatial and temporal coverage.

Unstructured data may also be used to infer knowledge about places and other spatial entities and semantically enrich more structured representations. Hu et al. [163] developed a computational framework to detect semantic relatedness between cities mentioned in the same news articles. The framework is based on the labeled latent Dirichlet allocation (LLDA) model to extract likelihood scores for different topics (e.g., politics, culture, business, and sports) from news articles. The semantic relatedness between two cities under a topic is quantified as the number of news articles containing the co-occurrences of the two cities and also discussing the topic. The extracted semantic relatedness is then aggregated according to the topic (e.g., culture) or to the time period using the number of news articles published in that time period.

The evaluation of semantic similarity/ relatedness results is also a critical factor, which is measured by the degree to which a computational measure approximates human-generated judgements. Ballatore et al. [164] have developed the Geo Relatedness and Similarity Dataset (GeReSiD), an open dataset for the evaluation of computational measures of geo-semantic relatedness and similarity. Geo-semantic relatedness focuses on relations involving at least one term with a spatial dimension. The dataset includes 97 geographic terms forming 50 term pairs and was developed based on human judgments.

4. Conclusions and Future Challenges

The paper provides a review of semantic-based approaches to information modeling and elicitation spanning the last five years of research. In this context, information modeling refers to the development of ontologies at different levels of generality and formality, tailored to various needs and uses. Elicitation on the other hand includes a set of processes educing semantic information from semi-structured and unstructured resources. Modeling and elicitation may be used synergistically to enhance both: extracted information is used to improve and enrich an ontology, which in turn is used to refine the information extraction results. In the last five years, information modeling and elicitation focused on a variety of spatial concepts: Interdisciplinary and vague geospatial concepts such as places and forests, dynamic concepts such as phenomena, events, and trajectories, and concepts with a cognitive and linguistic basis such as social roles associated with cities, emotions evoked by places, and characteristics of landscapes. Such concepts may be defined in a more traditional way in top-down approaches fitted for a formalization of expert knowledge but also in bottom-up approaches built more easily from user perceptions and sense of their spatial environment, as expressed in natural language texts.

While remarkable progress has been made over these last five years, much is still to be done to solve problems caused by different conceptualizations and interpretations of geospatial information and facilitate knowledge exchange also across different domains and languages, as outlined below.

Although ontologies have provided a sound basis for the semantic formalization of domain knowledge, many interdisciplinary geospatial concepts are formalized in different ways by different domains. Multidomain knowledge integration is critical to solve complex scientific problems that require geospatial resources pertaining to multiple domains (e.g., geography, environmental science, Earth sciences, hydrology, forestry, spatial planning, and natural resource management). The arrival of the Internet of Things has pushed for the development of new ontologies but also calls for a better integration of these ontologies. Existing works have shown the direction with the development of ontology design patterns, lightweight or micro-ontologies as an alternative to upper level and domain ontologies. These alternative approaches, varying in semantic expressivity, focus on representing different perspectives and interpretations of geospatial information.

The development of methods capable of expressing a wide range of cross-domain knowledge and deeper semantic information elicitation processes for geospatial and other relevant concepts would greatly enhance the ability of scientists and general users to search for and combine information. Existing geospatial information extraction approaches focus mostly on locations, concepts, and topics, and to a lesser extent on predefined properties and relations between concepts. The extraction of properties and interrelations between entities, as well as the joint extraction and modeling of various elements and their interrelations is a more complicated and challenging research topic which is also critical for further populating both domain-specific ontologies and general-knowledge KBs such as DBpedia [91], YAGO [165], and BabelNet [166].

The majority of information extraction approaches rely on language-specific tools that are developed primarily for the English language followed by other common languages such as French, Spanish, and Chinese. However, there is a growing need for extracting semantic information from sources written in other languages, as well as for integrating cross-lingual information. Besides the development of extraction tools tailored to other languages, this requires the use of multilingual KBs, such as DBpedia or BabelNet, and the development of more robust tools for multi- or cross-lingual information extraction, search, and retrieval. Moreover, since information extraction and search approaches are typically complex, requiring different inputs, resulting in different outputs based on different methods, their evaluation is critical for assessing not only their value, but also their reproducibility for different resources and languages and portability to different domains.

Information elicitation can help bridge the gap between the user and the machine, also allowing the user to express qualitative high-level concepts. The volume and variety of available resources, most of them in semi-structured and unstructured format, raises additional challenges as to the interpretations

and valid uses of this information. The inherent ambiguity, polysemy, and context-dependence of natural language further complicates the semantic-based extraction, enrichment, retrieval, and analysis of these resources, especially in a multilingual and multisource context. The need to overcome these barriers in knowledge discovery is a future research challenge highly relevant to data linking in the context of the Semantic Web and the advancement of geo-portals. In this context, besides more powerful techniques for multilingual multi-context semantic-based extraction, enrichment, and search, users would greatly benefit from novel approaches to interactive knowledge exploration and visualization. These approaches could provide users contextual information for a deeper comprehension of query entities, concepts, and semantic relations between them and thus enable them to formulate more relevant and effective queries.

Author Contributions: Conceptualization for the overall paper, writing Abstract, Introduction and Conclusions Margarita Kokla and Eric Guilbert. Writing Sections 2.1–2.4 and 3 including Figures 1, 7 and 8 Margarita Kokla. Writing Sections 2.2 and 2.4 including Figure 4, Eric Guilbert. Review & Editing Margarita Kokla and Eric Guilbert. All authors have read and agreed to the published version of the manuscript.

References

1. Kuhn, W. Geospatial Semantics: Why, of What, and How? In Proceedings of the Journal on Data Semantics III; Spaccapietra, S., Zimányi, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2005; pp. 1–24.

2. Couclelis, H. Unpacking the "I" in GIS: Information, Ontology, and the Geographic World. In *The Philosophy of GIS*; Springer Geography; Tambassi, T., Ed.; Springer: Cham, Switzerland, 2019; pp. 3–24, ISBN 978-3-030-16829-2.

3. 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]

4. Kuhn, W. Cognitive and Linguistic Ideas in Geographic Information Semantics. In *Cognitive and Linguistic Aspects of Geographic Space: New Perspectives on Geographic Information Research*; Lecture Notes in Geoinformation and Cartography; Raubal, M., Mark, D.M., Frank, A.U., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 159–174, ISBN 978-3-642-34359-9.

5. Kuhn, W. Semantic Engineering. In *Research Trends in Geographic Information Science*; Lecture Notes in Geoinformation and Cartography; Navratil, G., Ed.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 63–76, ISBN 978-3-540-88244-2.

6. Kuhn, W. Core concepts of spatial information for transdisciplinary research. *Int. J. Geogr. Inf. Sci.* **2012**, *26*, 2267–2276. [CrossRef]

7. Hu, Y. Geospatial Semantics. In *Comprehensive Geographic Information Systems*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 80–94, ISBN 978-0-12-804793-4.

8. Adams, B.; Janowicz, K. Constructing geo-ontologies by reification of observation data. In Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems—GIS'11, Chicago, IL, USA, 1–4 November 2011; p. 309.

9. Janowicz, K.; Hitzler, P. The Digital Earth as Knowledge Engine. *Semant. Web Interoper. Usability Appl.* **2012**, *3*, 213–221. [CrossRef]

10. Shadbolt, N.; Smart, P.R. Knowledge Elicitation: Methods, Tools and Techniques. In *Evaluation of Human Work*; Wilson, J.R., Sharples, S., Eds.; CRC Press: Boca Raton, FL, USA, 2015; pp. 163–200, ISBN 978-0-415-26757-1.

11. Kavouras, M.; Kokla, M. *Theories of Geographic Concepts: Ontological Approaches to Semantic Integration*; CRC Press: Boca Raton, FL, USA, 2008; ISBN 978-0-429-12928-5.

12. Liu, Y.; Li, L.; Shen, H.; Yang, H.; Luo, F. A Co-Citation and Cluster Analysis of Scientometrics of Geographic Information Ontology. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 120. [CrossRef]

13. Huang, H.; Chen, J.; Li, Z.; Gong, F.; Chen, N. Ontology-Guided Image Interpretation for GEOBIA of High Spatial Resolution Remote Sense Imagery: A Coastal Area Case Study. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 105. [CrossRef]

14. 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]

15. Argyridis, A.; Argialas, D. Fuzzy ontology-based foreshore identification from digital terrain models and very high-resolution airborne imagery through GEOBIA multi-scale analysis. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 2153–2169. [CrossRef]

16. Wang, W.; Stewart, K. Spatiotemporal and semantic information extraction from Web news reports about natural hazards. *Comput. Environ. Urban Syst.* **2015**, *50*, 30–40. [CrossRef]

17. Huang, Y. Conceptually categorizing geographic features from text based on latent semantic analysis and ontologies. *Ann. GIS* **2016**, *22*, 113–127. [CrossRef]

18. Hong, J.-H.; Kuo, C.-L. 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]

19. Krisnadhi, A. Ontology Design Patterns for Data Integration: The GeoLink Experience. In *Ontology Engineering with Ontology Design Patterns: Foundations and Applications*; IOS Press: Amsterdam, The Netherlands, 2016; pp. 267–278.

20. 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 *The Semantic Web—ISWC 2015*; Arenas, M., Corcho, O., Simperl, E., Strohmaier, M., d'Aquin, M., Srinivas, K., Groth, P., Dumontier, M., Heflin, J., Thirunarayan, K., et al., Eds.; Springer: Cham, Switzerland, 2015; Volume 9367, pp. 301–309, ISBN 978-3-319-25009-0.

21. Hu, Y.; Janowicz, K. Enriching top-down geo-ontologies using bottom-up knowledge mined from Linked Data. In *Advancing Geographic Information Science: The Past and Next Twenty Years*; Onsrud, H., Kuhn, W., Eds.; GSDI Association Press: Needham, MA, USA, 2016; pp. 183–198, ISBN 978-0-9852444-4-6.

22. Lemmens, R.; Falquet, G.; De Sabbata, S.; Jiang, B.; Bucher, B. Querying VGI by semantic enrichment. In *European Handbook of Crowdsourced Geographic Information*; Capineri, C., Haklay, M., Huang, H., Antoniou, V., Kettunen, J., Ostermann, F., Purves, R., Eds.; Ubiquity Press: London, UK, 2016; ISBN 978-1-909188-79-2.

23. Yue, P.; Guo, X.; Zhang, M.; Jiang, L.; Zhai, X. Linked Data and SDI: The case on Web geoprocessing workflows. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 245–257. [CrossRef]

24. Jiang, L.; Yue, P.; Kuhn, W.; Zhang, C.; Yu, C.; Guo, X. Advancing interoperability of geospatial data provenance on the web: Gap analysis and strategies. *Comput. Geosci.* **2018**, *117*, 21–31. [CrossRef]

25. Stock, K.; Yousaf, J. Context-aware automated interpretation of elaborate natural language descriptions of location through learning from empirical data. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 1087–1116. [CrossRef]

26. Xing, X.-F.; Mostafavi, M.-A.; Chavoshi, S. A Knowledge Base for Automatic Feature Recognition from Point Clouds in an Urban Scene. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 28. [CrossRef]

27. Wang, J.; Worboys, M. Ontologies and representation spaces for sketch map interpretation. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 1697–1721. [CrossRef]

28. Gruber, T. Every Ontology Is a Treaty. Interview at http://www.sigsemis.org from Miltiadis Lytras on 2004-11-22. Available online: http://lists.w3.org/Archives/Public/www-annotation/2004JulDec/0032.html (accessed on 24 October 2019).

29. Guarino, N. Formal ontology in information systems: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy. In *Frontiers in Artificial Intelligence and Applications*; Guarino, N., Ed.; IOS Press: Amsterdam, The Netherlands, 1998; pp. 3–15, ISBN 978-90-5199-399-8.

30. Casati, R.; Varzi, A.C. Spatial Entities. In *Spatial and Temporal Reasoning*; Stock, O., Ed.; Springer: Dordrecht, The Netherlands, 1997; ISBN 978-0-7923-4716-3.

31. Casati, R.; Smith, B.; Varzi, A.C. Ontological tools for geographic representation. In *Formal Ontology in Information Systems: Proceedings of the First International Conference (FOIS'98), June 6-8, Trento, Italy*; Frontiers in artificial intelligence and applications; Guarino, N., Ed.; IOS Press: Amsterdam, The Netherlands, 1998; pp. 77–85, ISBN 978-90-5199-399-8.

32. Muller, P. Space-time as a primitive for space and motion. In *Formal Ontology in Information Systems: PROCEEDINGS of the First International Conference (FOIS'98), June 6-8, Trento, Italy*; Frontiers in artificial intelligence and applications; Guarino, N., Ed.; IOS Press: Amsterdam, The Netherlands, 1998; pp. 63–76, ISBN 978-90-5199-399-8.

33. Grenon, P. The formal Ontology of Spatio-Temporal Reality and its Formalization. In *Foundations and Applications of Spatio-Temporal Reasoning (FATSR)*; AAAI Spring Symposium Technical Reports Series; AAAI Press: Menlo Park, CA, USA, 2003; pp. 27–34.

34. Grenon, P.; Smith, B. SNAP and SPAN: Towards Dynamic Spatial Ontology. *Spat. Cogn. Comput.* **2004**, *4*, 69–104. [CrossRef]

35. Masolo, C.; Borgo, S.; Gangemi, A.; Guarino, N.; Oltramari, A. *WonderWeb Deliverable D18 Ontology Library (final)*; Technical report No. 2001-33052; WonderWeb Ontology Infrastructure for the Semantic Web: Manchester, UK, 2003.

36. Herre, H.; Heller, B.; Burek, P.; Hoehndorf, R.; Loebe, F.; Michalek, H. *General Formal Ontology (GFO): A Foundational Ontology Integrating Objects and Processes. Part I: Basic Principles*; Research Group Ontologies in Medicine (Onto-Med): Leipzig, Germany, 2006.

37. Bateman, J.A.; Hois, J.; Ross, R.; Tenbrink, T. A linguistic ontology of space for natural language processing. *Artif. Intell.* **2010**, *174*, 1027–1071. [CrossRef]

38. Niles, I.; Pease, A. Towards a standard upper ontology. In Proceedings of the International Conference on Formal Ontology in Information Systems—FOIS'01, Ogunquit, ME, USA, 17–19 October 2001; Volume 2001, pp. 2–9.

39. Fellbaum, C. *WordNet: An Electronic Lexical Database*; MIT Press: Cambridge, UK, 1998; ISBN 978-0-262-27255-1.

40. Herre, H. Formal Ontology-A New Discipline Between Philosophy, Formal Logic, and Artificial Intelligence. In *Kybernetik, Logik, Semiotik Philosophische Sichtweisen*; Trafo Wissenschaftsverlag: Berlin, Germany, 2016; pp. 375–389, ISBN 978-3-86464-095-7.

41. Guarino, N. Formal ontology, conceptual analysis and knowledge representation. *Int. J. Hum.-Comput. Stud.* **1995**, *43*, 625–640. [CrossRef]

42. Varzi, A. Mereology. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Metaphysics Research Lab: Stanford, CA, USA, 2019.

43. Casati, R.; Varzi, A. *Parts and Places*; The MIT Press: Cambridge, MA, USA, 2003.

44. Baumann, R.; Loebe, F.; Herre, H. Towards an Ontology of Space for GFO. In Proceedings of the Formal Ontology in Information Systems; IOS Press: Amsterdam, The Netherlands, 2016; Volume 283, pp. 53–66.

45. Schmidtke, H.R. Granular Mereogeometry. In Proceedings of the Formal Ontology in Information Systems—Proceedings of the 9th International Conference (FOIS 2016), Annecy, France, 6–9 July 2016.

46. Hahmann, T. *On Decomposition Operations in a Theory of Multidimensional Qualitative Space*; Frontiers in Artificial Intelligence and Applications; IOS Press: Amsterdam, The Netherlands, 2018; Volume 306, pp. 173–186.

47. Galton, A. The Treatment of Time in Upper Ontologies. In Proceedings of the 10th International Conference on Formal Ontology in Information Systems, Cape Town, South Africa, 17–21 September 2018; pp. 33–46.

48. Galton, A. Outline of a Formal Theory of Processes and Events, and Why GIScience Needs One. In Proceedings of the International Conference on Spatial Information Theory, Santa Fee, NM, USA, 12–16 October 2015; Fabrikant, S.I., Raubal, M., Bertolotto, M., Davies, C., Freundschuh, S., Bell, S., Eds.; Springer: Cham, Switzerland, 2015; pp. 3–22.

49. Hanzal, T.; Svátek, V.; Vacura, M. Event Categories on the Semantic Web and Their Relationship/Object Distinction. *Front. Artif. Intell. Appl.* **2016**, *283*, 183–196.

50. Van Hage, W.R.; Malaisé, V.; Segers, R.H.; Hollink, L.; Schreiber, G. *Design and Use of the Simple Event Model (SEM)*; Social Science Research Network: Rochester, NY, USA, 2011.

51. Svatek, V.; Homola, M.; Kluka, J.; Vacura, M. Metamodeling-Based Coherence Checking of OWL Vocabulary Background Models. In Proceedings of the 10th International Workshop on OWL: Experiences and Directions (OWLED 2013), Montpellier, France, 26–27 May 2013.

52. Muñoz, L.S.; Grüninger, M. *Mapping and Verification of the Time Ontology in SUMO*; Frontiers in Artificial Intelligence and Applications; IOS Press: Amsterdam, The Netherlands, 2016; Volume 283, pp. 109–122.

53. Chui, C.; Grüninger, M. Merging the DOLCE and PSL Upper Ontologies. In Proceedings of the International Conference on Knowledge Engineering and Ontology Development, Rome, Italy, 21–24 October 2014; pp. 16–26.

54. Raskin, R.; Pan, M. Semantic web for earth and environmental terminology (SWEET). In Proceedings of the Workshop on Semantic Web Technologies for Searching and Retrieving Scientific Data, Sanibel Island, FL, USA, 20 October 2003.

55.	Ahlqvist, O.; Varanka, D.; Fritz, S.; Janowicz, K. *Land Use and Land Cover Semantics: Principles, Best Practices, and Prospects*; CRC Press: Boca Raton, FL, USA, 2017; ISBN 978-1-138-74799-9.

56.	Álvarez, L.G.; Bennett, B. Classification, Individuation and Demarcation of Forests: Formalising the Multi-Faceted Semantics of Geographic Terms. In Proceedings of the 13th International Conference on Spatial Information Theory (COSIT 2017), L'Aquila, Italy, 4–8 September 2017; Clementini, E., Donnelly, M., Yuan, M., Kray, C., Fogliaroni, P., Ballatore, A., Eds.; Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik: Dagstuhl, Germany, 2017; Volume 86, pp. 8:1–8:15.

57.	Ballatore, A. Prolegomena for an Ontology of Place. In *Advancing Geographic Information Science*; GSDI Association Press: Santa Barbara, CA, USA, 2016.

58.	Calafiore, A.; Boella, G.; Borgo, S.; Guarino, N. *Urban Artefacts and Their Social Roles: Towards an Ontology of Social Practices*; Dagstuhl Publishing: Dagstuhl, Germany, 2017; p. 13.

59.	Garbacz, P.; Ławrynowicz, A.; Bogumił, S. *Identity Criteria for Localities*; Frontiers in Artificial Intelligence and Applications; IOS Press: Amsterdam, The Netherlands, 2018; Volume 306, pp. 47–54.

60.	Gharebaghi, A.; Mostafavi, M.A. A New Ontological Perspective for Integration of Social and Physical Environments: Disability and Rehabilitation Context. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III–2*, 137–142. [CrossRef]

61.	Stephen, S.; Hahmann, T. An Ontological Framework for Characterizing Hydrological Flow Processes. In Proceedings of the 13th International Conference on Spatial Information Theory (COSIT 2017), L'Aquila, Italy, 4–8 September 2017; p. 14.

62.	Yan, J.; Guilbert, E.; Saux, E. An Ontology of the Submarine Relief for Analysis and Representation on Nautical Charts. *Cartogr. J.* **2015**, *52*, 58–66. [CrossRef]

63.	Grenoble, L.A.; McMahan, H.; Petrussen, A.K. An Ontology of Landscape and Seascape in Greenland: The Linguistic Encoding of Land in Kalaallisut. *Int. J. Am. Linguist.* **2019**, *85*, 1–43. [CrossRef]

64.	Sinha, G.; Arundel, S.T.; Hahmann, T.; Usery, E.L.; Stewart, K.; Mark, D.M. *The Landform Reference Ontology (LFRO): A Foundation for Exploring Linguistic and Geospatial Conceptualization of Landforms*; Dagstuhl Publishing: Dagstuhl, Germany, 2018; p. 7.

65.	Guilbert, E.; Moulin, B.; Cortés Murcia, A. A Conceptual Model for the Representation of Landforms Using Ontology Design Patterns. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *III–2*, 15–22. [CrossRef]

66.	Guilbert, E.; Moulin, B. Towards a Common Framework for the Identification of Landforms on Terrain Models. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 12. [CrossRef]

67.	Arvor, D.; Belgiu, M.; Falomir, Z.; Mougenot, I.; Durieux, L. Ontologies to interpret remote sensing images: Why do we need them? *GISci. Remote Sens.* **2019**, *56*, 911–939. [CrossRef]

68.	Zhuang, C.; Xie, Z.; Ma, K.; Guo, M.; Wu, L. A Task-Oriented Knowledge Base for Geospatial Problem-Solving. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 423. [CrossRef]

69.	Hofer, B.; Mäs, S.; Brauner, J.; Bernard, L. Towards a knowledge base to support geoprocessing workflow development. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 694–716. [CrossRef]

70.	Rospocher, M. An ontology for personalized environmental decision support. In Proceedings of the 8th International Conference on Formal Ontology in Information Systems (FOIS 2014), Rio de Janeiro, Brazil, 22–25 September 2014.

71.	Linyao, Q.; Zhiqiang, D.; Qing, Z. A Task-Oriented Disaster Information Correlation Method. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *II-4/W2*, 169–176. [CrossRef]

72.	Zhong, S.; Fang, Z.; Zhu, M.; Huang, Q. A geo-ontology-based approach to decision-making in emergency management of meteorological disasters. *Nat. Hazards* **2017**, *89*, 531–554. [CrossRef]

73.	Touya, G.; Bucher, B.; Falquet, G.; Jaara, K.; Steiniger, S. Modelling Geographic Relationships in Automated Environments. In *Abstracting Geographic Information in a Data Rich World: Methodologies and Applications of Map Generalisation*; Lecture Notes in Geoinformation and Cartography; Burghardt, D., Duchêne, C., Mackaness, W., Eds.; Springer: Cham, Switzerland, 2014; pp. 53–82, ISBN 978-3-319-00203-3.

74.	Gould, N.; Mackaness, W. From taxonomies to ontologies: Formalizing generalization knowledge for on-demand mapping. *Cartogr. Geogr. Inf. Sci.* **2016**, *43*, 208–222. [CrossRef]

75.	Yan, J.; Guilbert, E.; Saux, E. An ontology-driven multi-agent system for nautical chart generalization. *Cartogr. Geogr. Inf. Sci.* **2017**, *44*, 201–215. [CrossRef]

76.	Varanka, D.E.; Usery, E.L. The map as knowledge base. *Int. J. Cartogr.* **2018**, *4*, 201–223. [CrossRef]

77. Huang, W.; Harrie, L. Towards knowledge-based geovisualisation using Semantic Web technologies: A knowledge representation approach coupling ontologies and rules. *Int. J. Digit. Earth* **2019**. [CrossRef]

78. Hahmann, T.; Usery, E.L. What is in a Contour Map? In Proceedings of the Spatial Information Theory, Santa Fe, NM, USA, 12–16 October 2015; Fabrikant, S.I., Raubal, M., Bertolotto, M., Davies, C., Freundschuh, S., Bell, S., Eds.; Springer: Cham, Switzerland, 2015; pp. 375–399.

79. Davies, J. Lightweight Ontologies. In *Theory and Applications of Ontology: Computer Applications*; Poli, R., Healy, M., Kameas, A., Eds.; Springer: Dordrecht, The Netherlands, 2010; pp. 197–229, ISBN 978-90-481-8846-8.

80. Gandon, F. A survey of the first 20 years of research on semantic Web and linked data. *Ingénierie Systèmes Inf.* **2018**, *23*, 11–38. [CrossRef]

81. Giunchiglia, F.; Zaihrayeu, I. Lightweight Ontologies. In *Encyclopedia of Database Systems*; Springer reference; Liu, L., Özsu, M.T., Eds.; Springer: New York, NY, USA, 2009; ISBN 978-0-387-35544-3.

82. Hasan, R.; Farazi, F.; Bursi, O.S.; Reza, M.S. A Faceted Lightweight Ontology for Earthquake Engineering Research Projects and Experiments. In *Experimental Research in Earthquake Engineering: EU-SERIES Concluding Workshop*; Geotechnical, Geological and Earthquake Engineering; Taucer, F., Apostolska, R., Eds.; Springer: Cham, Switzerland, 2015; pp. 11–19, ISBN 978-3-319-10136-1.

83. Kuai, X.; Li, L.; Luo, H.; Hang, S.; Zhang, Z.; Liu, Y. Geospatial Information Categories Mapping in a Cross-lingual Environment: A Case Study of "Surface Water" Categories in Chinese and American Topographic Maps. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 90. [CrossRef]

84. Kordjamshidi, P.; Moens, M.-F. Global Machine Learning for Spatial Ontology Population. *J. Web Semant.* **2015**, *30*, 3–21. [CrossRef]

85. Gangemi, A.; Presutti, V. Ontology Design Patterns. In *Handbook on Ontologies*; Staab, S., Studer, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 221–243, ISBN 978-3-540-70999-2.

86. Calafiore, A.; Boella, G.; van der Torre, L. From Georeferenced Data to Socio-Spatial Knowledge. Ontology Design Patterns to Discover Domain-Specific Knowledge from Crowdsourced Data. In Proceedings of the Knowledge Engineering and Knowledge Management, Nancy, France, 12–16 November 2018; Faron Zucker, C., Ghidini, C., Napoli, A., Toussaint, Y., Eds.; Springer Nature Switzerland AG: Cham, Switzerland, 2018; pp. 34–49.

87. Sinha, G.; Mark, D.; Kolas, D.; Varanka, D.; Romero, B.E.; Feng, C.-C.; Usery, E.L.; Liebermann, J.; Sorokine, A. An Ontology Design Pattern for Surface Water Features. In Proceedings of the 8th International Conference on Geographic Information Science, Vienna, Austria, 24–26 September 2014; Duckham, M., Pebesma, E., Stewart, K., Frank, A.U., Eds.; Springer: Cham, Switzerland, 2014; pp. 187–203.

88. Janowicz, K.; Haller, A.; Cox, S.J.D.; Le Phuoc, D.; Lefrançois, M. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *J. Web Semant.* **2019**, *56*, 1–10. [CrossRef]

89. Bennett, B.; Cialone, C. Corpus Guided Sense Cluster Analysis: A methodology for ontology development (with examples from the spatial domain). In Proceedings of the 8th International Conference on Formal Ontology in Information Systems, Rio de Janeiro, Brazil, 22–25 September 2014; Frontiers in Artificial Intelligence and Applications. IOS Press: Amsterdam, The Netherlands, 2014; Volume 267, pp. 213–226.

90. 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]

91. Auer, S.; Bizer, C.; Kobilarov, G.; Lehmann, J.; Cyganiak, R.; Ives, Z. DBpedia: A Nucleus for a Web of Open Data. In *The Semantic Web*; Aberer, K., Choi, K.-S., Noy, N., Allemang, D., Lee, K.-I., Nixon, L., Golbeck, J., Mika, P., Maynard, D., Mizoguchi, R., et al., Eds.; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4825, pp. 722–735, ISBN 978-3-540-76297-3.

92. GeoNames. Available online: http://www.geonames.org (accessed on 24 October 2019).

93. Getty Thesaurus of Geographic Names (Getty Research Institute). Available online: http://www.getty.edu/research/tools/vocabularies/tgn/index.html (accessed on 28 January 2020).

94. Kokla, M.; Papadias, V.; Tomai, E. Enrichment and Population of a Geospatial Ontology for Semantic Information Extraction. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *XLII–4*, 309–314. [CrossRef]

95. Kuhn, W. Modeling vs encoding for the Semantic Web. *Semant. Web* **2010**, *1*, 11–15. [CrossRef]

96. Mark, D.; Smith, B. A science of topography: From qualitative ontology to digital representations. In *Geographic Information Science and Mountain Geomorphology*; Bishop, M.P., Shroder, J.F., Eds.; Praxis publishing: Chichester, UK, 2004; pp. 75–100.

97. Karkaletsis, V.; Fragkou, P.; Petasis, G.; Iosif, E. Ontology Based Information Extraction from Text. In *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution: Bridging the Semantic Gap*; Lecture Notes in Computer Science; Paliouras, G., Spyropoulos, C.D., Tsatsaronis, G., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; pp. 89–109, ISBN 978-3-642-20795-2.

98. Wimalasuriya, D.C.; Dou, D. Ontology-based information extraction: An introduction and a survey of current approaches. *J. Inf. Sci.* **2010**, *36*, 306–323. [CrossRef]

99. Hu, Y. Geo-text data and data-driven geospatial semantics. *Geogr. Compass* **2018**, *12*, e12404. [CrossRef]

100. Mark, D.M.; Smith, B.; Tversky, B. Ontology and Geographic Objects: An Empirical Study of Cognitive Categorization. In Proceedings of the Spatial Information Theory. Cognitive and Computational Foundations of Geographic Information Science, Stade, Germany, 25–29 August 1999; Freksa, C., Mark, D.M., Eds.; Springer: Berlin/Heidelberg, Germany, 1999; pp. 283–298.

101. Smith, B.; Mark, D.M. Geographical categories: An ontological investigation. *Int. J. Geogr. Inf. Sci.* **2001**, *15*, 591–612. [CrossRef]

102. Giannakopoulou, L.; Kavouras, M.; Kokla, M.; Mark, D. From Compasses and Maps to Mountains and Territories: Experimental Results on Geographic Cognitive Categorization. In *Cognitive and Linguistic Aspects of Geographic Space: New Perspectives on Geographic Information Research*; Lecture Notes in Geoinformation and Cartography; Raubal, M., Mark, D.M., Frank, A.U., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 63–81, ISBN 978-3-642-34359-9.

103. Montello, D.R.; Friedman, A.; Phillips, D.W. Vague cognitive regions in geography and geographic information science. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 1802–1820. [CrossRef]

104. Scheider, S.; Ostermann, F.O.; Adams, B. Why good data analysts need to be critical synthesists. Determining the role of semantics in data analysis. *Future Gener. Comput. Syst.* **2017**, *72*, 11–22. [CrossRef]

105. Martinez-Rodriguez, J.L.; Hogan, A.; Lopez-Arevalo, I. Information extraction meets the Semantic Web: A survey. *Semant. Web* **2018**, *11*, 1–81. [CrossRef]

106. Ghosh, S.; Gunning, D. *Natural Language Processing Fundamentals: Build Intelligent Applications that Can Interpret the Human Language to Deliver Impactful Results*; Packt Publishing: Birmingham, UK, 2019; ISBN 978-1-78995-404-3.

107. Blei, D.M.; Lafferty, J.D.; Lafferty, J.D. Topic Models. Available online: https://www.taylorfrancis.com/ (accessed on 31 December 2019).

108. Allahyari, M.; Pouriyeh, S.A.; Assefi, M.; Safaei, S.; Trippe, E.D.; Gutierrez, J.B.; Kochut, K.J. A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *arXiv* **2017**, arXiv:1707.02919.

109. Hearst, M.A. Automatic Acquisition of Hyponyms from Large Text Corpora. In Proceedings of the 14th Conference on Computational Linguistics, COLING'92, Nantes, France, 23–28 August 1992; Association for Computational Linguistics: Stroudsburg, PA, USA, 1992; Volume 2.

110. Navigli, R. Word sense disambiguation: A survey. *ACM Comput. Surv. CSUR* **2009**, *41*, 1–69. [CrossRef]

111. Richter, L.; Geiß, J.; Spitz, A.; Gertz, M. HeidelPlace: An Extensible Framework for Geoparsing. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Copenhagen, Denmark, 9–11 September 2017; pp. 85–90.

112. Gritta, M.; Pilehvar, M.T.; Limsopatham, N.; Collier, N. What's missing in geographical parsing? *Lang. Resour. Eval.* **2018**, *52*, 603–623. [CrossRef]

113. Monteiro, B.R.; Davis, C.A.; Fonseca, F. A survey on the geographic scope of textual documents. *Comput. Geosci.* **2016**, *96*, 23–34. [CrossRef]

114. Hahn, J.; Fogliaroni, P.; Frank, A.U.; Navratil, G. A Computational Model for Context and Spatial Concepts. In Proceedings of the Geospatial Data in a Changing World, Helsinki, Finland, 14–17 June 2016; Sarjakoski, T., Santos, M.Y., Sarjakoski, L.T., Eds.; Springer: Cham, Swizerland, 2016; pp. 3–19.

115. Derungs, C.; Purves, R.S. Mining nearness relations from an n-grams Web corpus in geographical space. *Spat. Cogn. Comput.* **2016**, *16*, 301–322. [CrossRef]

116. Wallgrün, J.O.; Klippel, A.; Baldwin, T. Building a corpus of spatial relational expressions extracted from web documents. In Proceedings of the 8th Workshop on Geographic Information Retrieval—GIR'14, Dallas, TX, USA, 4–7 November 2014; pp. 1–8.

117. Mirrezaei, S.I.; Martins, B.; Cruz, I.F. A distantly supervised method for extracting spatio-temporal information from text. In Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems—GIS'16, Burlingame, CA, USA, 31 October–3 November 2016; pp. 1–4.

118. Dittrich, A.; Vasardani, M.; Winter, S.; Baldwin, T.; Liu, F. A Classification Schema for Fast Disambiguation of Spatial Prepositions. In Proceedings of the 6th ACM SIGSPATIAL International Workshop on GeoStreaming—IWGS'15, Bellevue, WA, USA, 3–6 November 2015; pp. 78–86.

119. Radke, M.; Das, P.; Stock, K.; Jones, C.B. Detecting the Geospatialness of Prepositions from Natural Language Text. In Proceedings of the 14th International Conference on Spatial Information Theory (COSIT 2019), Regensburg, Germany, 9–13 September 2019; Timpf, S., Schlieder, C., Kattenbeck, M., Ludwig, B., Stewart, K., Eds.; Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik: Dagstuhl, Germany, 2019; Volume 142, pp. 11:1–11:8.

120. Kordjamshidi, P.; Van Otterlo, M.; Moens, M.F. Spatial role labeling. *ACM Trans. Speech Lang. Process. TSLP* **2011**, *8*, 1–36. [CrossRef]

121. Abraham, S.; Mäs, S.; Bernard, L. Extraction of spatio-temporal data about historical events from text documents. *Trans. GIS* **2018**, *22*, 677–696. [CrossRef]

122. General Architecture for Text Engineering (GATE). Available online: https://gate.ac.uk/ (accessed on 24 October 2019).

123. Ballatore, A.; Adams, B. Extracting Place Emotions from Travel Blogs. In Proceedings of the AGILE 2015: Geographic Information Science as an Enabler of Smarter Cities and Communities, Lecture Notes in Geoinformation and Cartography, Lisboa, Portugal, 9–12 June 2015; Bacao, F., Santos, M.Y., Painho, M., Eds.; Springer: Cham, Switzerland, 2015; pp. 1–5.

124. Strapparava, C.; Valitutti, A. Wordnet-affect: An affective extension of wordnet. In Proceedings of the 4th International Conference on Language Resources and Evaluation, Lisbon, Portugal, 26–28 May 2004; European Language Resources Association (ELRA): Paris, France, 2004; pp. 1083–1086.

125. Egorova, E.; Moncla, L.; Gaio, M.; Claramunt, C.; Purves, R.S. Fictive motion extraction and classification. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 2247–2271. [CrossRef]

126. Derungs, C.; Purves, R.S. 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]

127. Wartmann, F.M.; Acheson, E.; Purves, R.S. Describing and comparing landscapes using tags, texts, and free lists: An interdisciplinary approach. *Int. J. Geogr. Inf. Sci.* **2018**, *32*, 1572–1592. [CrossRef]

128. Cooper, D.C.; Gregory, I.N.; Hardie, A. Spatializing and Analyzing Digital Texts: Corpora, GIS, and Places. In *Deep Maps and Spatial Narratives*; Bodenhamer, D.J., Corrigan, J., Harris, T.M., Eds.; Indiana University Press: Bloomington, IN, USA, 2015; ISBN 978-0-253-01555-6.

129. Bruggmann, A.; Fabrikant, S.I. How to visualize the geography of Swiss history. In Proceedings of the 17th Agile International Conference on Geographic Information Science Connecting a Digital Europe through Location and Place, Castellón, Spain, 3–6 June 2014; Huerta, J.J., Schade, S., Granel, C., Eds.; Springer: Cham, Switzerland, 2014.

130. Bruggmann, A.; Fabrikant, S.I. Spatializing time in a history text corpus. In Proceedings of the GIScience 2014: Eighth International Conference on Geographic Information Science, Vienna, Austria, 24–26 September 2014.

131. Salvini, M.M.; Fabrikant, S.I. Spatialization of user-generated content to uncover the multirelational world city network. *Environ. Plan. B Plan. Des.* **2016**, *43*, 228–248. [CrossRef]

132. Pernelle, N. Semantic Enrichment of Data: Annotation and Data Linking. Habilitation Thesis, Université Paris Sud, Orsay, France, 2016.

133. Hu, Y.; Janowicz, K.; Prasad, S.; Gao, S. Enabling Semantic Search and Knowledge Discovery for ArcGIS Online: A Linked-Data-Driven Approach. In *AGILE 2015: Geographic Information Science as an Enabler of Smarter Cities and Communities*; Lecture Notes in Geoinformation and Cartography; Bacao, F., Santos, M.Y., Painho, M., Eds.; Springer: Cham, Switzerland, 2015; pp. 107–124, ISBN 978-3-319-16787-9.

134. Ennis, A.; Nugent, C.; Morrow, P.; Chen, L.; Ioannidis, G.; Stan, A.; Rachev, P. A Geospatial Semantic Enrichment and Query Service for Geotagged Photographs. *Sensors* **2015**, *15*, 17470–17482. [CrossRef]

135. Tardy, C.; Falquet, G.; Moccozet, L. Semantic enrichment of places with VGI sources: A knowledge based approach. In Proceedings of the 10th Workshop on Geographic Information Retrieval—GIR'16, Burlingame, CA, USA, 31 October 2016; pp. 1–2.

136. Manguinhas, H.; Freire, N.; Isaac, A.; Stiller, J.; Charles, V.; Soroa, A.; Simon, R.; Alexiev, V. Exploring Comparative Evaluation of Semantic Enrichment Tools for Cultural Heritage Metadata. In Proceedings of the Research and Advanced Technology for Digital Libraries, Hannover, Germany, 5–9 September 2016; Fuhr, N., Kovács, L., Risse, T., Nejdl, W., Eds.; Springer: Cham, Switzerland, 2016; pp. 266–278.

137. Nixon, L.; Bauer, M.; Bara, C. Connected Media Experiences: Web Based Interactive Video Using Linked Data. In Proceedings of the 22nd International Conference on World Wide Web, Rio de Janeiro, Brazil, 13–17 May 2013; pp. 309–312.

138. Romero, S.; Becker, K. Improving the classification of events in tweets using semantic enrichment. In Proceedings of the International Conference on Web Intelligence—WI'17, Leipzig, Germany, 23–26 August 2017; pp. 581–588.

139. Ilarri, S.; Stojanovic, D.; Ray, C. Semantic management of moving objects: A vision towards smart mobility. *Expert Syst. Appl.* **2015**, *42*, 1418–1435. [CrossRef]

140. 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]

141. Vechtomova, O. Query Expansion for Information Retrieval. In *Encyclopedia of Database Systems*; Liu, L., Özsu, M.T., Eds.; Springer: Boston, MA, USA, 2009; pp. 2254–2257, ISBN 978-0-387-39940-9.

142. Purves, R.; Jones, C. Geographic Information Retrieval. *SIGSPATIAL Spec.* **2011**, *3*, 2–4. [CrossRef]

143. 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]

144. Meštrović, A.; Calì, A. An Ontology-Based Approach to Information Retrieval. In Proceedings of the Semantic Keyword-Based Search on Structured Data Sources, Gdańsk, Poland, 11–12 September 2017; Calì, A., Gorgan, D., Ugarte, M., Eds.; Springer: Cham, Switzerland, 2017; pp. 150–156.

145. Munir, K.; Sheraz Anjum, M. The use of ontologies for effective knowledge modelling and information retrieval. *Appl. Comput. Inform.* **2018**, *14*, 116–126. [CrossRef]

146. Subarani, D. Concept Based Information Retrieval from Text Documents. *IOSR J. Comput. Eng.* **2012**, *2*, 38–48. [CrossRef]

147. De Andrade, F.G.; de Souza Baptista, C.; Davis, C.A. Improving geographic information retrieval in spatial data infrastructures. *GeoInformatica* **2014**, *18*, 793–818. [CrossRef]

148. Nys, G.-A.; Kasprzyk, J.-P.; Hallot, P.; Billen, R. A Semantic Retrieval System in Remote Sensing Web Platforms. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *XLII-2/W13*, 1593–1599. [CrossRef]

149. Zaila, Y.L.; Montesi, D. Geographic information extraction, disambiguation and ranking techniques. In Proceedings of the 9th Workshop on Geographic Information Retrieval—GIR'15, Paris, France, 26–27 November 2015; pp. 1–7.

150. Mendes, P.N.; Jakob, M.; García-Silva, A.; Bizer, C. DBpedia spotlight: Shedding light on the web of documents. In Proceedings of the 7th International Conference on Semantic Systems—I-Semantics'11, Graz, Austria, 7–9 September 2011; pp. 1–8.

151. Butuc, M.G. Semantically enriching content using OpenCalais. *EDITIA* **2009**, *9*, 77–88.

152. Han, L.; Kashyap, A.L.; Finin, T.; Mayfield, J.; Weese, J. UMBC_EBIQUITY-CORE: Semantic Textual Similarity Systems. In Second Joint Conference on Lexical and Computational Semantics (*SEM): Proceedings of the Main Conference and the Shared Task: Semantic Textual Similarity, Atlanta, GA, USA, 13–14 June 2013.

153. Jiang, Y.; Li, Y.; Yang, C.; Liu, K.; Armstrong, E.M.; Huang, T.; Moroni, D.F.; Finch, C.J. 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]

154. 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]

155. 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; pp. 12–22.

156. Vockner, B.; Mittlböck, M. Geo-Enrichment and Semantic Enhancement of Metadata Sets to Augment Discovery in Geoportals. *ISPRS Int. J. Geo-Inf.* **2014**, *3*, 345–367. [CrossRef]

157. Laurini, R. Geographic Ontologies, Gazetteers and Multilingualism. *Future Internet* **2015**, *7*, 1–23. [CrossRef]

158. Adams, B. Finding similar places using the observation-to-generalization place model. *J. Geogr. Syst.* **2015**, *17*, 137–156. [CrossRef]

159. 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]

160. Khan, A.; Vasardani, M.; Winter, S. Extracting Spatial Information from Place Descriptions. In Proceedings of the First ACM SIGSPATIAL International Workshop on Computational Models of Place—COMP'13, Orlando, FL, USA, 5–8 November 2013; pp. 62–69.

161. Liu, F.; Vasardani, M.; Baldwin, T. Automatic Identification of Locative Expressions from Social Media Text: A Comparative Analysis. In Proceedings of the 4th International Workshop on Location and the Web, Shanghai, China, 3–7 November 2014; pp. 9–16.

162. Chen, Z.; Yang, Y. Semantic relatedness algorithm for keyword sets of geographic metadata. *Cartogr. Geogr. Inf. Sci.* **2020**, *47*, 125–140. [CrossRef]

163. Hu, Y.; Ye, X.; Shaw, S.-L. Extracting and analyzing semantic relatedness between cities using news articles. *Int. J. Geogr. Inf. Sci.* **2017**, *31*, 2427–2451. [CrossRef]

164. Ballatore, A.; Bertolotto, M.; Wilson, D.C. An evaluative baseline for geo-semantic relatedness and similarity. *GeoInformatica* **2014**, *18*, 747–767. [CrossRef]

165. Rebele, T.; Suchanek, F.; Hoffart, J.; Biega, J.; Kuzey, E.; Weikum, G. YAGO: A Multilingual Knowledge Base from Wikipedia, Wordnet, and Geonames. In Proceedings of the 15th International Semantic Web Conference ISWC 2016, Kobe, Japan, 17–21 October 2016; Groth, P., Simperl, E., Gray, A., Sabou, M., Krötzsch, M., Lecue, F., Flöck, F., Gil, Y., Eds.; Springer: Cham, Switzerland, 2016; Volume 9982, pp. 177–185, ISBN 978-3-319-46546-3.

166. Navigli, R.; Ponzetto, S.P. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artif. Intell.* **2012**, *193*, 217–250. [CrossRef]