Surface Network Ontology Design Patterns for Linked Topographic Data

Editor(s):
Solicited review(s):
Open review(s):

Gaurav Sinha[a,*], Dave Kolas[b], David Mark,[c], Boleslo E. Romero[d], E. Lynn Usery[e], Gary Berg-Cross[f], and Anand Padmanabhan[g]

[a] *Department of Geography, Ohio University, Athens, OH, USA.*
[b] *Raytheon BBN Technologies, Columbia, MD 21046, USA.*
[c] *Department of Geography & NCGIA, University at Buffalo, Buffalo, NY, USA.*
[d] *Department of Geography, University of California, Santa Barbara, CA, USA.*
[e] *U.S. Geological Survey, 1400 Independence Road, Rolla, MO, USA.*
[f] *Spatial Ontology Community of Practice (SOCOP) & Knowledge Strategies, Potomac, MD, USA.*
[g] *CyberInfrastructure and Geospatial Information Laboratory (CIGI) & National Center for Supercomputing Applications (NCSA), University of Illinois at Urbana-Champaign, Urbana, IL, USA.*

Abstract. The vision of Linked Topographic Data is critical for the Semantic Web, since topographic data are fundamental to a wide range of geoscientific analyses and mapping of geographic phenomena. Terrain datasets are probably the most important for Linked Topographic Data. Terrain is computationally represented using a continuous field data model (2-D surfaces), but the Semantic Web needs discrete objects for assigning URIs. This prevents terrain surface datasets from being shared on the Semantic Web—which is ironical given the availability of incredibly high resolution terrain data. Surface networks, which are a topologically connected set of shape-critical points, lines, and areal districts, can be used to share information about surfaces on the Semantic Web. In this paper, we reinterpret surface network theory for Linked Topographic Data, and present two OWL ontology design patterns. *Surface Network* is a template pattern intended for *any* type of surface network. It formalizes only the topological connections between surface network elements, since formalized of the metric properties requires commitment to a domain-specific spatial ontology. *Geospatial SNODP* extends SNODP for metric geographic space through alignment with the GeoSPARQL geospatial ontology. It can be used to annotate *any* geospatial surface network, but is expected to primarily serve as a core terrain ontology for Linked Topographic Data.

Keywords: Linked Topographic Data, Surface Network, Topography, Ontology Design Pattern, GeoSPARQL

1. Introduction

1.1 Why ontology design patterns for the Semantic Web?

The Semantic Web is a long-term vision of providing common frameworks for sharing, integrating, and reusing diverse data. Linked Data is the current practical expression of the Semantic Web for publishing data in accordance with a set of best practices. In words of the originator of both concepts, Tim Berners-Lee, Linked Data is “Semantic Web done right” [3]. Linked Data ideally breaks barriers between data and liberates them from hitherto inaccessible data silos. Linked Data is still mostly about instance level

*Corresponding author. Email: sinhag@ohio.edu

data to keep entry barriers low, but as argued by [27], sensible reuse of such data requires conceptual level integration between data sources and awareness of the societal context of the origins and reuse of data. In this regard, ontologies are essential pillars for the success of Linked Data. They define the semantics of data and clarify valid contexts of reuse. Misunderstandings and incorrect use of data are ultimately attributable to vague or insufficient descriptions and unintended interpretations by human designers [19]. For the heterogeneous Semantic Web, to be of any practical value, standardized approaches to ontology design need to be replaced or complemented with contextual and/or domain-oriented ontologies [24, 34]. Large, comprehensive ontologies are difficult to design and reuse due to complexity and lack of reason-ability [58]. An emerging consensus seems to be that Semantic Web ontologies will be most useful if they are minimalist in scope, and serve only to *constrain*, not completely specify, interpretations [28, 29]. Semantic Web ontologies will be necessarily heterogeneous, so multiple ontologies reflecting diverse ontological commitments will also be needed to capture socially situated, context-dependent interpretive frameworks [1, 6].

A significant challenge generally has been the lack of people with both domain expertise and ontology engineering skills. However, for the Semantic Web community, it seems to have been a blessing in disguise, since it created a synergy between domain experts and engineers resulting in an ontology design process wherein domain experts are actively engaged in the design process and not merely providing feedback to the engineers [24]. Ontology Design Patterns (ODPs) are a recent design paradigm to address this and several other issues discussed above. They are pragmatic alternatives to foundational ontologies, such as DOLCE [35], BFO [18, 55] or SUMO [39], which are too abstract to be understood by non-logicians interested only in Linked Data. The idea of a design pattern was first introduced in architecture as an archetypal solution to standard design problems. The idea also caught fire later in software engineering [10, 12] and DBMS design [20]. Solutions to commonly encountered software component design problems were published as abstract software patterns, which tremendously increased efficiency and robustness of the software design process. Similarly, ontology design patterns have become very popular in ontology engineering as a modeling solution for recurrent ontology design problems [13, 58]. ODPs are easy to understand and act as ready-to-use building blocks in larger scale ontology engineering efforts [14]. The primary benefits of ODPs are that they are very specific in scope serving as well-defined "units" of knowledge, and are easily comprehensible and reusable, factors critical for a fast-growing, but conceptually robust, and culturally diverse Semantic Web.

1.2 Surface network ontology patterns for Linked Topographic Data

Inspired by several of the ideas discussed above, since 2008, a series of *VoCamp* (Vocabulary Camp) workshops [1] have engaged academic scholars and practitioners from several types of institutions and industries to work in a few small groups to create lightweight vocabularies and/or ontologies for the Semantic Web. Between 2011 and 2013, seven *GeoVoCamps* were organized for creating *geospatial* vocabularies and ontology patterns, in response to the growing recognition of the Geospatial Semantic Web [4, 9, 30], which is a special interpretation of the Semantic Web centered on the spatiotemporal aspects of Linked Data and other Semantic Web resources [16, 26, 65]. GeoVoCamp participants include both ontology engineers and domain experts, who are cognizant of the benefits of philosophical, foundational ontologies, but also appreciate the bottom-up approach of capturing partial domain semantics. The goal is to focus on a tractable concept and create preliminary versions of a geospatial ontology pattern. The work is revised and finalized later by the collaborators. Most outcomes from GeoVoCamps are available online in some form for public review on the workshop wiki or website, and some are eventually reported in a formal publication (such as this paper and [23].

In this paper, we report on two related ontology design patterns: Surface Network ODP (SNODP) and Geospatial Surface Network ODP (Geospatial SNODP) that we developed at the *GeoVocampSOCop2012* workshop [2]. There are multiple motivations for choosing the niche and technical topic of surface networks, and as discussed later, the patterns have great relevance to the Semantic Web—but the primary motivation is to introduce and promote the concept of *Linked Topographic Data*. In a nutshell, it is our vision of making topographic datasets stored in various object, network, and field data models easily accessible and interoperable through Semantic Web technologies. The most fun-

[1] See http://www.vocamp.org for an overview of all VoCamps.
[2] URL: http://www.vocamp.org/wiki/GeoVoCampSOCoP2012.

damental component of topographic databases—and of Linked Topographic Data—is terrain, i.e., the shape of the earth's surface, which serves as the physical backdrop for all human activities, and is a critical factor in most geoscientific studies. From a data modeling perspective, terrain can be conceived as a spatially continuous (scalar) *field* of elevations measured with respect to a datum such as mean sea level. (A *scalar* geospatial field is represented mathematically as a surface that covers some part of the Earth's surface). Since scalar fields can be sampled in reality only at discrete locations, if a spatially continuous representation is desired, they must be mathematically approximated by *surfaces*, which are single-valued mathematical functions [$z = f(x, y)$] of position in 2-D space. The surface function is the basis of spatial interpolation at unsampled locations, and the form of the function depends on the spatial characteristics of the field it approximates. However, the field data model poses a general problem for the Semantic Web because fields store measurements of a spatial property without reference to any identifiable object. This makes it difficult to share field based datasets on the Semantic Web which requires stable *objects* that can be assigned independent identities (URIs) [25]. This is preventing an enormous amount of geoscientific data, stored as fields, to be shared on the Semantic Web.

This fundamental problem is the reason for us selecting SNODP and Geospatial SNODP as the founding patterns for Linked Topographic Data. Surface networks abstract the global spatial shape of *any* surface in terms of a topological network of shape-critical points (peaks, passes, and pits), lines (ridges, course, slope, and contour lines), and areas (hills, dales and territories). Surface network elements make it possible to share, at least, the basic shape semantics of their source surfaces (which are mathematical models of fields) on the Semantic Web. Surface networks substantially advance the vision of Linked Topographic Data because of the well-established importance of terrain shape (or geomorphometric) measurements in terrain analysis [66], and also in studying people's naïve, culturally and linguistically rooted cognition and communication about topography [33, 56]. Another major contribution of this work is the realization that surface network ontology can serve as the core ontology for both geoscientific and culturally based conceptualizations of the terrain.

Linked Topographic Data is not just a vision for the Semantic Web, but also aligns well with the public service goals of national mapping organizations, such as the United States Geological Survey (USGS) and the UK Ordnance Survey, which are developing topographic map ontologies to support semantic search for mapped topographic features [15, 37, 61]. Although these efforts are not coordinated across organizations, all are similarly challenged by the need to integrate field and object topographic data models. For example, if the current version of The National Map[3], an online mapping and topographic data exploration service from the USGS is queried for a named landform, the answer is effectively a point location associated with the named landform, not the spatial extent and other physical properties or descriptors of that landform. That problem can only be addressed by creating and storing *object*-based representations of landforms in topographic databases. However, extracting and characterizing landform objects from topographic data [53] and formalization of the semantics of these and many other types of topographic features [62] are complex goals, which demand collaborations between several research groups if Linked Topographic Data is to be successful.

A noteworthy point about our work is that the scope of the patterns we designed is actually wider than the cause of Linked Topographic Data that originally inspired it. Two-dimensional surfaces are easily associated with the surface topography because we live on it, but there are many other surfaces that can also be conceptualized. Thus, we consciously designed the patterns such that we could stay true to both the spirit of ODP design and the theory of surface networks, which is a general theory for *all* (not just terrain) surfaces. Our fundamental pattern (SNODP) can be used for capturing shape semantics of any type of surface, while our second pattern (Geospatial SNODP) is applicable to any *geospatial* surface, including surfaces representing the physical terrain. The topic of surface networks was also quite suitable for a GeoVocamp workshop, because its basic principles are relatively compact, well-understood, and unambiguous. Given the limited time at the workshop and diversity in backgrounds of participants at GeoVoCamps, it was important to select a topic that could be easily understood, demanded minimalistic shifts in participant's perspectives, and would yield ontology patterns that can benefit a large community of users. Surface network theory met these criteria very well.

[3] URL: http://www.nationalmap.gov

2. Review of surface network theory and applications

2.1 Surface network theory

The basic theory of the surface network was proposed more than 150 years ago by mathematicians and physicists [7, 36, 49]. Considering a two-dimensional continuous smooth closed surface floating in space with surface values relative to an internal reference point, Reech [49] described ideas surrounding three types of local extrema, or critical (singular) points, existing on the surface: maxima (i.e. *peaks*), minima (i.e. *pits*), and mixed extrema (*saddle points*) which are maxima across one axis and minima across another. The mixed extrema fall into two types, which Warntz [63] later called *passes* (lowest point between peaks) and *pales* (highest point between pits). Cayley [7] independently presented a theory of surfaces, focusing on *contour lines* and *slope lines*, where contour lines run horizontally and slope lines are orthogonal to contour lines running directly up and down slopes. He noted different shapes of the indicatrix and contour lines just surrounding each critical point. Peaks and pits have elliptical shapes and the saddles have hyperbolic shapes approaching the point where the contour has a knot and crosses itself. Two basic types of knotted contour lines exist: *outloops*, which are similar to a figure-eight, and *inloops*, which have one loop entirely contained within the other. Cayley [7] also identified special slope lines, *ridge lines* and *course lines*, which are particular slope lines that connect critical points. From each normal saddle point (pass), the slope lines of steepest ascent are called ridge lines, and ascend to peaks (rarely, to the same peak); the two slope lines of steepest descent from a saddle point (pale) are called course lines and descend to pits (rarely, to the same pit). The definitions and relations among these critical points and lines form the foundation of surface networks.

Maxwell [36] further developed surface network theory by describing and naming two types of regions (called *districts*) on a surface bounded by ridge-lines and/or course-lines. A district around a peak and bounded by course-lines form a *hill*, and similarly a region around a pit and bounded by ridge-lines was called a *dale*. All of the slope lines in a hill ascend to a single peak, and all of the slope-lines in a dale descent to a single pit, respectively. Using both special lines and districts along with the previous point relations, Maxwell proposed a set of numerical relations among the surface features. In the 1960s, William Warntz rediscovered the work of Cayley and Maxwell and synthesized much of the previous surface theory [63, 64]. Warntz [63] also extended Maxwell's [36] concepts of districts by introducing *territories* which are subsets of districts and are bounded, in general, by two ridge lines and two course lines. Finally, Warntz also tabulated the vergency of forces flowing down the surface to lower values: peaks, ridges, and hills have divergence; pits, courses, and dales have convergence; and passes, pales, and territories have mixed vergency.

In an independent line of work, apparently not linked to Cayley and Maxwell, Morse [38] established the mathematical theory of critical points on differentiable surfaces in Euclidean space and generalized the results to multiple dimensions. Another mathematician, Reeb [48] conceived the splitting and merging of contours at critical points as forming what is a now called Reeb Graph (also equivalent to the Contour Tree) with critical points as nodes, ridge lines and course lines as edges, and territories as faces. Peaks and pits, with infinitely small contours, are terminal nodes which are connected by edges, for ridge and course lines, to nodes representing the *inloop* and *outloop* contours of passes and pales. The pass and pale nodes represent points of topological change with hills or dales being split or merged in the graph. Later, Pfaltz [42] also developed a type of graph to represent relationships of surface network features and increase efficiency of storage and selection. Peaks and pits are connected via saddles with edges representing ridge and course lines. His mathematical interpretation and proofs strengthened the mathematical basis of the theory and led to its adoption in computer science fields such as medical imaging, and computer vision. Pfaltz [42] also proposed a graph-theoretic method called *homomorphic contraction* for simplifying surface networks in order to increase efficiency of storage, selection, and retrieval. This process identified and simplified sets of related features without creating or destroying any topological circuits. General patterns are maintained while redundant details could be stored as secondary linked data, accessed only if necessary. It is notable, however, that not all such simplifications resulted in general patterns that might be intuitively identified by people.

An important point to note about surface network discussions in the literature is that over the years different authors have presented the same concepts in slightly different ways and using different terminologies. For a comprehensive review of surface network theory, equivalent terminologies, and challenges, we recommend [45, 46].

2.2 Practical applications

Surface networks abstract surface-specific information about surface structure of single-valued fields in a highly condensed form using identifiable features such as critical points and special slope lines. Whether represented in a spatially embedded network or as a graph, this limited set of data offers efficiency in modes of storage, query, selection, visualization, and transfer. With surface network features available, most of the remaining surface data need not be stored or readily available for every stage of processing and analysis. This has benefits for visualizing surfaces (see a discussion in [46]). Some analyses may be performed with only the surface network features, such as identifying regions of high variability that may benefit from additional sampling or identifying courses or topographic watershed boundaries [31]. Also, a surface network is an objective and efficient aid for selecting values from the field data. For instance, using only the reduced set of network features, a manual or automatic selection may be performed to identify all of the relevant field data for a particular analysis. The "Very Important Point" algorithm for constructing Triangular Irregular Network (TIN) data models uses surface network theory to find a sparse set of points that maximize information about the surface configuration. Spatial or graphical relationships among surface network features can aid in such a selection, possibly finding all of the downslope values comprising a hill bounded by courses and associated with a particular peak or the inverse case for a pit. More complex features might also be identified. For example, generally defined patterns of surface network features may help to find volcanoes, craters, or other features with characteristic configurations. It may be possible to match surface network features to intuitive, identifiable features named by people in either general or specific terms with extended potential for location and navigation tasks, as well. A relatively unexplored application area of surface networks is comparison of surfaces [50, 51] and even tracking the morphology of a surface as it evolves in the real world or in digital animations and artificial simulations.

3. Representing and interpreting surface networks for the Semantic Web

3.1 Importance of surface network theory for the Semantic Web

Surface network theory is amenable for formalization as an ontology design pattern because of its compactness, and applicability to any kind of surface modeling. Apart from the above mentioned scientific applications, surface networks can benefit the Semantic Web specifically in many ways:

i. As already mentioned, many significant volumes of scientific data, especially in the geosciences, are stored as fields, but Semantic Web ontologies are object-based [25]. The basis of Linked Data is URIs assigned to every discrete entity on the Web, but field datasets are mere arrays of stored values, with no explicitly encoded objects that could be assigned URIs. Fields are also stored using special format data models (e.g., raster, TIN, netCDF) and using complex data reduction methodologies—conversion to RDF results in intractably large data files. Clearly, fields present tremendous interoperability challenges for Linked Data and the Semantic Web. In general, only objects extracted from fields should become part of Linked Data (e.g., landcover objects from remotely sensed electromagnetic fields, topographic features from terrain fields), but feature extraction methods are domain specific and highly context dependent. Surface network theory can help by allowing, at least, the shape semantics of any surface to be easily expressed and shared through sets of discrete objects.

ii. As noted in [41] notes, advances in computer visualization facilitate human comprehension of shapes, but computational reasoning with irregular shapes is a significant challenge. Because of its well-defined structure, a surface network provides a way to logically reason about geometric, topological, and mereo-topological relationships between distant features observed on a surface. It also forms a hierarchy that can be exploited for scale sensitive queries and abstraction of data.

iii. A surface network also serves as an objective data reduction knowledge pattern for sharing large surface datasets. Surface networks can also be further generalized. The net reduction could be as much as 90% according to one study of vector representations of surfaces [21]. This increases the efficiency of both human comprehension, and computational storage and processing of surfaces, especially for high resolution surfaces represented with ten to hundreds of millions pixels or data points, creating tremendous data processing and transmission problems.

iv. One of the critiques of the Semantic Web is that its ontologies often are engineering artifacts and lack cognitive validity [47]. Surface network theory scores highly on that account, since it has strong cognitive validity, despite being a mathematical construct. Its basic elements are easily conceptualized and visualized due to its reliance on shape patterns derived from topographic feature types (e.g., peak, pit, pass, ridge line, course line, and slope lines) that are simple and generic enough to make intuitive sense to most people, irrespective of their linguistic or cultural background.

v. The surface network patterns proposed here are not just ontology engineering artifacts, but encapsulations of well-established scientific concepts. The original theory is already known to have several scientific applications. Availability and popularity of ontology patterns for surface networks will hopefully spur more innovative approaches to surface data modeling and sharing, and advance the cause of e-Science on the Semantic Web [5].

vi. From the perspective of ontology and the Semantic Web, various domains, such as physics, mathematics, computer vision, geomorphology, hydrology, or human and physical geography, could use a common analytic vocabulary of surface network features and their relationships. With surface network ontology patterns available for Linked Data in the Semantic Web, general patterns and processes may be revealed within or across research domains.

3.2 Importance of surface network theory for Linked Topographic Data

The Linked Topographic Data initiative will succeed if there is ubiquitous adoption, applicability and interoperability of different types of topographic data models. It is critical to start with methods that bridge the conceptual and technological divide between two historically disparate approaches to terrain data modeling: (a) 2-D *field-based data models* suited for scientific computations and the de facto standard for storing and sharing terrain data, and (b) the *object-based data models* that are needed to support intuitive, natural language driven terrain information retrieval [34]. Surface networks are clearly quite important to Linked Topographic Data. For example, surface network theory was used in [53] for linking elevation field to object representations of topographic eminences. The method can be similarly useful for linking field and object versions of watersheds, valleys and many other landforms. A recent pilot study by the USGS extracted and quantified characteristic attributes of Meteor Crater from a digitally scanned topographic map and an orthophoto, published the data on the Semantic Web, and also implemented a query interface [60]; however, for semi-automated topographic object extraction, elevation and derivative fields are preferable to digitally scanned topographic maps and satellite imagery. Terrain surface network elements can act as topographic markers [32] that can be used to extract landforms, for automated location and annotation in visual displays of topographic data, and for geospatial registration and alignment of many types of topographic datasets.

Many types of semantic queries about topographic relationships can be answered quickly once surface networks are created for the earth's surface. As shown in [52], watershed and mountain hierarchies can be derived from terrain surface networks, which when combined with the topological information, and toponym databases, allows determination of river networks, ridge networks, mountain-valley connections, constituting mountains of mountain ranges, etc. Such queries can be resolved by processing only the much more compact surface network, without reasoning with complex topographic feature geometries. When a surface network is examined in combination with its source surface, even more detailed information emerges about surfaces. For example, the spatial pattern of ridge and course lines may suggest certain erosion regimes of consequence to geomorphologic and hydrologic analyses. Advantages also accrue in computational data modeling of terrain. Algorithms for converting topographic surfaces from a raster to a TIN representation model rely on surface network theory to find Very Important Points (VIPs) (i.e., critical points), preservation of which as nodes in the TIN ensures that the basic topographic structure is maintained. This increases the visual fidelity of terrain based visualizations, and also maintains the spatial, topological, and hydrologic properties of terrain features.

3.3 *Understanding surface network semantics relevant for designing ontology patterns*

While Semantic Web ontologies were always intended to be reusable, designing *easily* reusable ontologies has proven more challenging for a number of reasons: size and complexity of the ontology, incompatible assumptions about the domain, required comprehension of concepts unrelated to the task at hand, and demanding inferential requirements. FOAF[4] and SKOS[5] are two popular ontologies that show the potential of small, portable, or "sustainable" ontologies. Ontology modularization [17] or knowledge patterns [8] can alleviate the problem through small ontologies that represent only a slice of a domain. As mentioned earlier, a similar approach is that of ontology design patterns (ODPs), which are designed as domain independent solutions to a general class of problems, and readily expressible in any logical language [13]. While ODPs can be of several types [14], only Content Ontology Design Patterns (CPs) are relevant here since they encode domain concepts using non-logical, domain-specific vocabulary—the other types of patterns pertain to abstracting ontology engineering processes. CPs are characterized as "computational, small, autonomous, hierarchical, cognitively relevant, linguistically relevant, and best practices", and often reused through specializations, extensions, or compositions [43].

The two Surface Network ODPs are examples of content patterns, and an example of a growing list of geospatially motivated patterns that act as small resources of geospatial domain knowledge on the Semantic Web. It is important to note that because this is the first known ontology of a surface network and created specifically for the Semantic Web, it encapsulates only the core theoretical concepts needed for data sharing. A much more expressive ontology of surface networks would be needed for creating a comprehensive ontology of all possible entities and relationships. *That* ontology would also not be small or modular enough, and too complex for only sharing data on the Semantic Web. As such, all choices were made to create a streamlined, maximally reusable ontology pattern, and eschew the trappings of a comprehensive reference ontology for surface networks.

The key to designing surface network patterns is to understand that unlike physical road, river or blood vessel networks, the surface network is a network of *abstract* entities. The surface from which it originates is similarly a mathematical entity that *represents* both material surfaces extending in physical space (e.g., earth' crustal surface, earth's gravitational (Geoidal) surface, exterior of an animal body, and microscopic surfaces as studied in physics and chemistry) and abstract surfaces extending in conceptual space (e.g., surfaces of: population density, crime potential, simulated terrain, and grayscale images). In case of a physical surface (e.g., the earth's surface), some of the pattern's classes may *correspond* to observable features of the real world surface (e.g., peaks, passes, ridge tops, valleys), but they remain mathematical entities that should not be confused with the real world physical entities that they idealize.

In a purely topological surface network, the spatial configuration of network objects may be avoided (i.e., no coordinates for points or lines are recorded). The exclusively graph theoretic representation still allows resolution of several queries related to the topology of the surface, by facilitating selective access and retrieval of information about surface shape [42]. However, the inability to embed the surface network in metric space means the lack of support for spatial analysis, visualization in metric space, and optimization of queries for surfaces. A purely topological network may still support storing the surface height associated with the critical points, which allows additional inferences to be made (e.g., calculating peak prominence and pit depths with respect to saddle points). Surface heights also allow the realization of a weighted surface network, where weights are calculated as the difference in surface values between the critical points [70].

Realization of metric surface networks [67] requires a surface network to remain spatially coregistered with its source surface, and spatial coordinates of *at least* the critical points need to be stored in the surface network. These points act as 'entry points' to the surface for information retrieval algorithms seeking to use surface networks in combination with surfaces [42]. The coordinates for the ridge and course lines may also be recorded to a desired degree of precision to yield their approximate geometric signatures. However, automated extraction of critical lines from digital surfaces is not trivial [44, 46]. When the geometric details of lines are not available, critical point coordinates can be used to compute approximate lengths of ridge/course lines, by calculating Euclidean (straight-line) distances between peaks/pits and saddle points and even stored

[4] FOAF Vocabulary Specification 0.98: http://xmlns.com/foaf/spec/

[5] SKOS Simple Knowledge Organization System Reference: http://www.w3.org/TR/skos-reference/

properties of edges. These calculated lengths can also be used as edge weights.

Accordingly, two surface network ODPs, one topological, and the other extended with geometric capabilities, are presented below. The topological, domain-independent *Surface Network ODP* (*SNODP*) captures only the semantics needed to create a topologically consistent surface network, with the additional capability to store surface heights for critical points. SNODP can be extended with metric properties, but only through a separate extension pattern that must first commit to an ontology of the space in which the surface network is supposed to be embedded. The *Geospatial SNODP* is such an extension of SNODP specifically for the geospatial domain. Both ODPs were developed in schematic form and later formalized in the OWL 2 language using the Protégé-OWL ontology editing software.

4. Surface Network ODP

The *Surface Network* pattern discussed here is based on [7, 36, 42, 63]. Figure 1 is a schematic representation of the classes that capture the essential semantics of the topological surface network and jointly form the first pattern called *Surface Network*. This pattern's class names are derived from Warntz's [63] paper, since his terms are quite similar to Maxwell's and his modifications are more appropriate for describing general surface semantics on the Semantic Web (cf. use of pit and peak, instead of Maxwell's terms summit and immit).[6] Class relationship names are based on descriptive phrases in the literature, but not necessarily attributable to a particular author. The OWL formalization of the SNODP ontology is hosted online at a resolvable URI[7]. In the interest of space, we will focus our discussion below on special conceptual issues that are not already evident from the schematic and the OWL ontology.

4.1 Surface

As mentioned earlier, surface datasets generally should not be serialized as RDF triples due to conceptual and technical reasons. Therefore, this class captures only the semantics of the fundamental relationship that exists between a surface and its surface networks, and nothing else. A surface is ontologically prior to the surface network because the former is a pre-requisite for the extraction of the latter. The surface network's component critical points, lines, and districts completely inherit their topological and spatial configuration from the source surface. The nature of this dependency is captured by the *embeds* and its inverse *isEmbeddedIn* properties, which are together meant to imply (albeit implicitly) that the location of all parts are fixed in space and derivable only from the specific locations in the surface where they are embedded. These properties are transitive under parthood, implying that every part of the embedded surface network must also be embedded in the surface. This is specified axiomatically using a property chain [22], a built-in functionality in OWL 2 to allow axiomatic definition of new properties by a chain of object properties—in this case *embeds* and *hasPart_directly*, where the latter property is defined in the W3C recommended best practice *SimplePartWhole* OWL pattern that is imported by SNODP to model straightforward cases of part-whole mereological relations.[8] Property chain reasoning for the inverse properties *isEmbeddedIn* and *partOf_directly* holds automatically in OWL. A limitation of using property chaining is that reasoners must be OWL 2 compliant to infer these axiomatically specified properties.

[6] The OWL file includes comments that mention alternative terms used in the literature to refer to exactly the same entities as represented by the classes of the SNODP.

[7] The *SurfaceNetwork* OWL file is available @: http://purl.org/geovocamp/ontology/SurfaceNetwork.

[8] Simple part-whole relations in OWL Ontologies. Accessed August 30, 2013 @ http://www.w3.org/2001/sw/BestPractices/OEP/SimplePartWhole/simple-part-whole-relations-v1.5.html

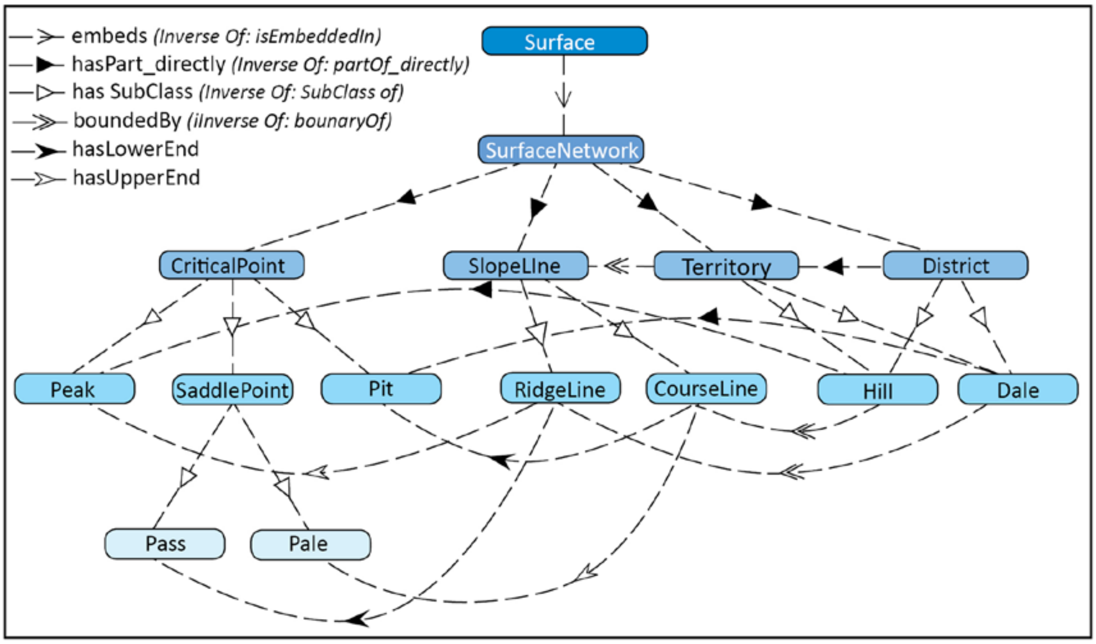

Fig. 1. Schematic representation of (only) the relationships between the classes of the Surface Network ODP. The inverse relationships are not shown in the diagram for maintaining clarity, but are mentioned in the diagram key. For other properties specific to a particular class only, consult the full ontology and the main text.

The *surfaceData* object property supported by the class is not a semantic property *per se*, but is needed for the technical purpose of storing a link to an external resource hosting the surface dataset. It takes If access to the external surface dataset does not need to be provided, the class can remain uninstantiated to eliminate the overhead of providing access to surface data. Note that because we support this flexibility, a design consequence is that the *isEmbeddedIn* property cannot be restricted with a cardinality of 1 to specify that the surface network and its parts are all embedded in *only one* surface. Restricting the cardinality would also entail the mandatory instantiation of the *Surface* class because otherwise OWL rules would infer an unknown class to honor the cardinality. The result would be a vacuous assertion for *every instance* in the ontology that it is embedded in some unknown object (actually supposed to be the surface). Not enforcing the cardinality is an elegant safeguard against such redundant assertions and gives users the flexibility to decide whether to provide access to surface datasets.

4.2 *SurfaceNetwork*

In addition to being the only class of the pattern related directly to the *Surface* class, the *SurfaceNetwork* class also serves as the collection class for all other classes needed to model the semantics of a topological surface network. As shown in Figure 1, the *hasPart_directly* and its inverse *partOf_*directly properties from the *SimplePartWhole* pattern are used to specify the partonomic relationship between the *SurfaceNetwork* class and *CriticalPoint*, *SlopeLine*, *District*, and *Territory* classes. These represent the four types of entities essential to the specification of the semantics of a surface network.

The most common use case of SNODP will be for sharing data for *only one* surface network. In such cases, instantiation of the *SurfaceNetwork* class is not necessary since it only leads to redundant assertions that *every* instantiated individual is part of a surface network. The real purpose of this class is to help partition the collection of different surface network parts, when multiple surface networks are generated and managed as part of the same resource (e.g., SPARQL graph, RDF document). However, an unintended consequence of not instantiating the *SurfaceNetwork* class is the disruption of the property chain needed to infer that the components or parts of the surface network are embedded in the surface. That renders the *Surface* instance as a redundant stand-alone object in the ontology with no relationships to any other object in the ontology.

4.2.1 *CriticalPoint*

CriticalPoint has three subclasses: *Peak*, *Pit*, and *SaddlePoint*, which model, respectively, local maxima, local minima, and saddle (i.e., non-extrema) stationary points (*i.e.*, points where the first derivative (gradient) in all directions on a surface is zero making the tangential plane at that point is parallel to the base plane of the surface). *SaddlePoint* has two subclasses: *Pass* and *Pale*. Since authors have used the term *pass* in two senses: i) to refer to *any* saddle point, and ii) the special case of saddle points that are the lowest point between two peaks, but are not bars (pales), in this pattern, we use *SaddlePoint* to cover all saddle points, and created *Pass* and *Bars* subclasses to distinguish between saddle points that are the lowest point connecting two peaks, from those which are the highest point connecting two pits, respectively. All *CriticalPoint* classes also support a *surfaceValue* datatype property (of type *double*) to record surface heights as measured along an axis oriented orthogonal to the base plane domain of the surface.

Critical points are the most basic elements of a surface network and in many cases, they are the only elements extracted when the full topological surface network is not needed or difficult to extract. Merely having knowledge of critical points or even a subset of them (e.g., only peaks or pits) is relevant in a lot of applications. For example, topographic maps generally need to shown only peaks and some passes of strategic importance; finding only the pits for a terrain surface approximated through a digital elevation model is an important step in hydrological analysis; and in [53], only peaks and saddles were needed to guide algorithms for extracting the spatial extents of topographic eminences.

4.2.2 SlopeLine

Critical points are connected via slope lines, whose semantics are modeled via the *SlopeLine* class. Almost every slope line trends on the surface between a peak and a pit—except ridge and course slope lines which have saddle points at one of the extremities and a peak or pit, respectively, at the other. *RidgeLine* and *CourseLine* are subclasses to recognize special slope lines which are the steepest slope lines between a saddle point and a peak or pit, respectively, and deemed critical to surface shape description. Other slope lines do not signify anything unique about surface shape, and they are normally not explicitly stored in databases. The exact shape of ridge and course lines is not of interest in SNODOP, since it only aims to capture the topological connection of the lines with critical points. The lack of metric space support in the topological SNODP also means that ridge and course lines cannot be ‘overlaid’ spatially on the source surface. However, the property chaining axiom for the *isEmbeddedIn* property of *SurfaceNetwork* instances ensures that all ridge and course lines are, at least, inferred to be embedded in the surface.

Topological connections between critical points and lines are modeled through two slope line properties: *hasUpperEnd* and *hasLowerEnd* which distinguish the upper and lower ends of ridge and course lines, where up and down directions correspond to the direction of convexities (maxima) and concavities (minima), respectively. The range of the two properties must be some *CriticalPoint*. Additionally, OWL exact cardinality restrictions are used to specify that a ridge line has exactly one peak and exactly one saddle point, while a course line has exactly one saddle point and exactly one pit, at the upper and lower ends, respectively. Ridge lines never reach pits, and course lines never reach peaks. The rest of the slope lines have a peak and pit at their upper and lower end, respectively, but this is left unspecified because no separate subclass is included in the pattern for such slope lines.

4.2.3 District and Territory

The *District* class captures the semantics of the areas that the surface is partitioned into by the network of ridge and course lines that form the boundaries of the districts. The class subsumes *Hill* and *Dale* classes. The *boundedBy* property with a minimum cardinality of 1 is used for specifying that every *Hill* has one or more *CourseLine* instances defining its boundary, and similarly every *Dale* has at least one, and often more, *RidgeLine* instances as its boundary. *boundaryOf* is declared in SNODP as the property inverse of *boundedBy*. Since a hill or dale is defined as the union of all slope lines converging at the same peak or pit, respectively, OWL cardinality constraints enforce that every *Hill* instance has exactly one *Peak*, and every *Dale* instance has exactly one *Pit* as a direct part. Every instance of *District* must also have at least one instance of *Territory* (typically many) as a direct part. Every *Territory* instance is always *boundedBy* by at least one instance of *SlopeLine*, and is also a direct part of exactly one *Hill* and also one *Dale* instance. A territory in surface network theory represents a contiguous shared area between a hill and a dale.

5. Extending SNODP for the geospatial domain

SNODP eschews spatial semantics to avoid commitment to a particular ontology of space and limiting the applicability of the pattern, since different contexts need different formalizations of space (and time) semantics. SNODP semantics will suffice when only the locations of critical points and lines are not of interest—but such use cases will be less common than those requiring explicit spatial embedding of the surface network. Hence, SNODP should be interpreted as a template ontology pattern that defines the fundamental topology semantics of a surface network, but it must be extended with spatial semantics such that the *locations* of critical points and (optionally of critical lines) can also be shared to explicitly embed the surface network in the space occupied by its source surface. We discuss here only one extension of SNODP for the entire geospatial domain, since the inspiration and majority of use cases still arise from geospatial analysis and visualization needs. Special attention is paid to topographic surface networks since the vision of Linked Topographic Data is the primary motivation for this project.

5.1 Extending SNODP with GeoSPARQL

SNODP covers most of the required semantics of surface networks. The extension for the geospatial domain only needs to additionally link the various surface network elements to their locations in geographic space, which already provides the framework for locating the source geospatial surface. For compatibility with the Semantic Web, a desirable end result would be serialization of locations in RDF triples. The Open Geospatial Consortium (OGC) recently proposed a standard called GeoSPARQL, which extends SPARQL, a W3C recommended RDF query language, with geospatial information representation and retrieval capabilities on the Semantic Web [40]. GeoSPARQL is based on existing OGC standards and also addresses several limitations with previously proposed geospatial vocabularies (see [2] for an extensive review). GeoSPARQL supports a small ontology for representing geospatial data semantics. This ontology can be attached to any ontology that needs to describe the location of entities. GeoSPARQL is not an ontology of the geospatial domain, but only a vocabulary for asserting and querying the location of geospatial entities using Euclidean geometry primitives. There are several properties of spatial entities that GeoSPARQL cannot handle. Nonetheless, for the modest technical goal of extending SNODP with geospatial capabilities, it is currently the best publicly available solution. Still, we do not discuss this alignment with GeoSPARQL prescriptively, but only to highlight some important conceptual issues related to Geospatial SNODP design that must be considered regardless of which specific spatial ontology is used.

5.2 Geospatial Surface Network ODP

Figure 2 schematizes Geospatial SNODP as derived by alignment with GeoSPARQL ontology. The OWL formalization of Geospatial SNODP is available online at a resolvable URI.[9] The primary class in GeoSPARQL is *geo:SpatialObject* which represents all spatial entities, and subsumes two subclasses: *geo:Feature* and *geo:Geometry*. These are declared to be *Disjoint* to clearly distinguish that geospatial entities are different from the geometric abstractions needed to represent them as objects in spatial databases. A *geo:Feature* is an abstraction of any entity which can have a real world location. No further specialization of this class is entailed since that is left to domain ontologies. The *hasGeometry* object property, which has *geo:Feature* as its domain and *geo:Geometry* as its range, links all *geo:Feature*s to their geometric representations. The *geo:Geometry* class subsumes all the geometry classes typically needed for representing the spatial extension of geospatial entities. A *geo:Feature* can have multiple *geo:Geometry*s to support different reasoning contexts, and one them (usually the most detailed) may also be declared as the *geo:defaultGeometry* for typical use cases.

[9] The *GeospatialSurfaceNetwork* OWL file is available @: http://purl.org/geovocamp/ontology/GeospatialSurfaceNetwork.

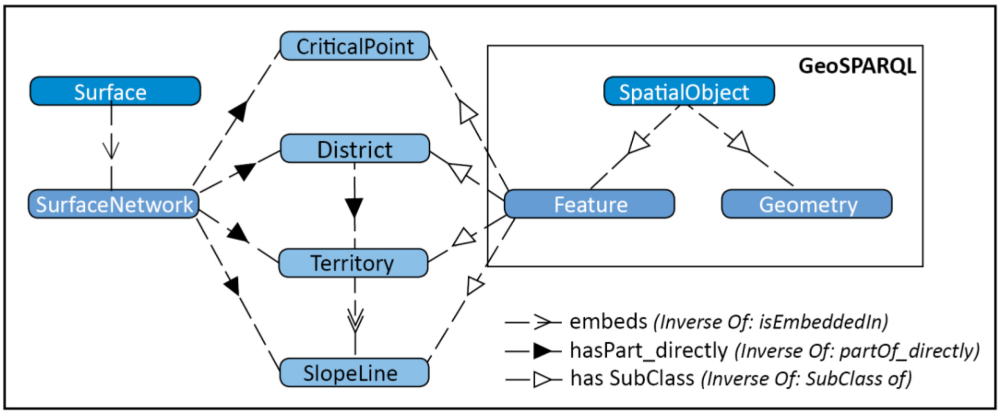

Fig. 2. Schematic representation of the top level classes of the Geospatial Surface Network ODP.

For RDF compatibility, the spatial location of *geo:Geometry* must be converted from traditional spatial database storage formats and serialized as a geometric literal, which can be based either on the Well-Known Text (WKT) [10] or the Geographic Markup Language (GML) [11] vector geometry representation standards. Unlike many other standards, GeoSPARQL supports multiple coordinate reference systems (CRSs) as defined by EPSG system. [12] The *geo:Geometry* class further depends on the *hasSerialization* data property, which has two subproperties: *geo:asWKT* and *geo:asGML*, to link to the appropriate WKT or GML geometric literal representation, respectively. Values for these properties use the *geo:wktLiteral* and *geo:gmlLiteral* data types respectively.

Any geospatial entity declared in other ontologies can be subsumed by *geo:Feature* to inherit geospatial properties and its location can be declared using the *geo:Geometry* subclasses. The geospatial 'feature' in most OGC standards, and the SDTS [13], a US Federal government standard, generally refers to an abstraction of a geospatial entity, which is typically assumed to have a physical extent in the real world. Geospatial surfaces present an interesting case in this regard. On one hand they are conceptual entities with no direct correspondence supposedly to physical entities; on the other, they acquire their significance only by being represented in geographic space, and in relation to other geographic entities. How should they be interpreted in GeoSPARQL? One might argue for surface network elements to be declared as subclasses of *geo:Geometry* since they are, after all, idealized points, lines, and areas. However, there are several reasons not to do so. First, that would limit the representation of a surface network element to only a *particular* geometric representation, precluding its representation at different levels of detail with multiple geometries, *all* linked via its *hasGeometry* property to the same *geo:Feature* (and to each other). Second, because multiple surface networks may be realized for a surface, if the elements were only instances of *geo:Geometry*, tracking identical, shared elements *across* multiple surface networks is precluded. Finally, and most importantly, such a choice would also be conceptually flawed since surface network elements are not *mere* geometric shapes; they also have specific spatial and topological properties, and there are many domain rules that apply to them.

It is evident that for both technical and conceptual reasons, surface network elements cannot be subclasses of *geo:Geometry*, and must be subsumed by *geo:Feature*. Accordingly, the Geospatial SNODP classes: *CriticalPoint*, *SlopeLine*, *District*, and *Territory* are declared subclasses of *geo:Feature*. Their subclasses automatically inherit *geo:Feature* properties as well. All other properties from SNODP are retained. The *geo:Geometry* classes can be used for instantiating the location and shape of all surface network elements in Geospatial SNODP—an option that is not available in SNODP. Note that the *SurfaceNetwork* class does not need to be aligned with GeoSPARQL ontology since only instances of its contained classes can have spatial representation.

Geospatial SNODP also supports *elevation* as a specialized sub-property of *surfaceValue* to support the most common use case of terrain surface net-

[10] WKT was originally specified in Simple Feature Access (Part 1: Common Architecture), an OGC® and ISO standard (19125).
[11] http://www.opengeospatial.org/standards/gml
[12] European Petroleum Survey Group (EPSG): http://www.epsg-registry.org
[13] http://mcmcweb.er.usgs.gov/sdts/

works. The property can simply be ignored for non-terrain geospatial surfaces. This design choice is an efficient compromise because it precludes the creation of a separate pattern for terrain on account of just one sub-property, especially when Geospatial SNODP is predominantly expected to be used for terrain data sharing. Terrain features are formed under the influence of gravity, and, therefore, terrain elevations must be *orthometric*, i.e., measured with respect to a level surface (e.g., Geoid or mean sea level) perpendicular to the gravity vector [36]. This is an indirect method of ensuring the correct vertical orientation of the surface in geographic space. Any orientation change of the surface effectively changes the shape of the surface and results in a different set of surface network elements. If the terrain surface is not constrained to be oriented in the gravitational "up" direction (by using orthometric elevations), the mathematically extracted shape elements will be just that—they will not be guaranteed to match in type or overlap in geographic space with observed terrain shapes. As one extreme case, if a surface were to be inverted, peaks get exchanged with pits, passes with pales, and course lines with ridge lines. Despite such constraints on the *elevation* sub-property, specifying the semantics of elevation is not within the scope of Geospatial SNODP—it must be addressed through another ontology pattern. We annotated the ontology file to describe the intent and terrain specific use of the *elevation* sub-property.

5.3 Geospatial SNODP as a "core" terrain ontology

Geospatial SNODP can be used for any type of geospatial surface network, physical or abstract, but we envision the primary use of this pattern to be for terrain surface data sharing. We contend here that geospatially embedded surface network elements are desirable as the primary constituents of a fundamental or core ontology that can underlie all other terrain ontologies because the shape elements can be clearly defined, and have clear correspondence with nearly universal topographic features. We outline the nature of these correspondences below.

Peaks are universally known topographic features, while many morphologically salient passes are recognized, at least, by some groups, as part of paths from one mountain to another. Bodies of standing water fill basins or low lying areas, but if those areas were imagined to be dry, the lowest point within each basin would correspond to a pit in the surface network. If outlets for the standing water bodies exist, they may be recognized as "natural" gateways from one basin to another—these are idealized as pales in surface networks. The shorelines of those water bodies would be contour lines that self-cross (or self-touch) at those pales. Watercourses (i.e., stream or river beds) are places where streams and rivers flow, and these should coincide with course lines. Drainage basins are perceived more for their function than morphology—they are units of land that drain water 'together' into watercourses that also lie within the extent of these basins. Drainage basins are equivalent to surface network dales, and are separated by drainage divides, which coincide often with ridge lines. Course and ridge lines, therefore, correspond to *bona fide* physical boundaries on the earth's surface [57]. People also recognize mountain sides and valley walls as shared areas—these are territories in a surface network. There is probably very little intuitive value to (surface network) hills, which represent land parcels partitioned by watercourses. Neither do those parcels correspond to natural, functional spatial units for explaining a spatial process, nor do they have a characteristic morphologic shape. In terrain with high relief, they may correspond approximately to areas occupied by mountains and hills, but in most other areas hill units may just contain, not exactly correspond to landforms. It is also unlikely that people perceive the topological network that can be formed between the elements, but, individually, surface network elements clearly have a correspondence to commonly understood terrain features. Surface network elements therefore can function as the common vocabulary, and more formally, as a 'core' terrain ontology underlying both cultural (object-based) and geoscientific (field-based) topographic ontologies.

Geospatial SNODP is well-positioned to serve as the 'core' terrain ontology since it covers most of the basic concepts needed for terrain ontology. However, it is not a comprehensive ontology of terrain surface networks, since there are some terrain features and relationships that are unaccounted for in the traditional theory. We list below several ways in which Geospatial SNODP cannot capture what holds for real terrain, because of limitations of the original theory it formalizes.

– In surface network theory, saddle points are locations where exactly two ridge and two course lines meet. For real terrain, sometimes more than two ridge and course lines can meet at "monkey saddles" and neither is it necessary that ridges and valleys always meet at saddle points.

– The real terrain often has flat areas which may act as local peaks, pits or saddle points, but they are hard to resolve during surface network construction because of how algorithms generally encode how critical points are detected.
– The mathematically defined ridge lines must trend along paths of steepest ascent between passes and peaks. However, in reality, many localized spurs and ridge networks that do not make it up to the peak or a pass can be observed— these features cannot be represented as ridge lines in a surface network. Similarly, if there is a long near-planar valley wall, gullies can develop that do not extend up to the ridge and therefore do not have saddles at their tops, so they cannot be course lines even if they are ravines with large streams. Other cases of stream channels not connecting to passes also may result.
– Conversely, many ridges and course lines in a surface network may have no noticeable terrain expression if the terrain is near-planar, so some surface network elements are practically redundant if the ontology must include only morphologic features of interest. This was shown in [53] as causing significant problems for computational analysis of DEMs for detecting the boundaries of isolated topographic eminences standing above planar lands. However, the ridge and course lines will still play an important role in hydrological analysis of the terrain.
– In the mathematical theory of surface networks, peaks/pits, passes/pales, and ridge line/course line systems are supposed to be complementary and equivalent duals of each other. However, when the actual terrain surface of the earth is considered, there are several differences between the semantics of terrain features and those of idealized surface network elements. On earth there are fewer pits, whereas highlands are much more dissected and there are many more peaks than pits on the earth's surface. Watercourses serve an important geophysical function in the landscape since water flows downhill under the influence of gravity, eroding and transporting sedimentary material. Watercourses are, therefore, mostly monotonically connected terrain features because their morphology is reinforced over time with sustained hydrological (or glacial) flow. In comparison, ridges are merely divergent morphologic features which lose water and where minimum erosion occurs. Because nothing flows through ridges and they are often discontinuous features, which makes it difficult to detect them as topologically continuous features extending from peak to pass.

The implication of all these caveats is that for a complete representation of topographic semantics, Geospatial SNODP must be further extended with additional surface network semantics, and complemented with other topography specific ontology patterns. Some authors have suggested adding new elements and modifying or adapting the definitions of the classical elements to address these limitations [52, 70]. It may be worthwhile to examine the suggestions, and if necessary, extend Geospatial SNODP with another pattern that is specialized exclusively for terrain surface networks.

5.4 Geospatial SNODP as a "core" ontology for Linked Topographic Data

Topographic information can be available in object, network, and field based geospatial datasets which must be all made to interoperate transparently to realize Linked Topographic Data. Ontologies that define semantics of surfaces or fields are yet to be designed, but they will be scientific in scope and relatively free of linguistic and cultural impacts. In contrast, ontologies of terrain objects will be quite diverse, reflecting people's varying ontological commitments rooted in cultural and linguistic differences. Several ontologies will be necessary to cater to the specialized semantics of different topographic sub-domains (e.g., eminences, surface water, land cover), and to account for varying ontological commitments. This heterogeneity is nothing new for the Semantic Web, but fundamental, commonly shared concepts should be condensed as "core" ontologies that can be imported by all other topographic ontologies.

We recognize Geospatial SNODP as one such core ontology that will be critical to the Linked Topography Data initiative. On one hand, its purpose is to represent information about the geometric and topological structure of surfaces and make it accessible on the Semantic Web. On the other, Geospatial SNODP formalizes semantics of those entities that can be treated as the building blocks of higher level terrain objects (landforms) that have social importance. Thus, in a Linked Topographic Data context, Geospatial SNODP is positioned to play the pivotal role of a mid-tier ontology functioning to "vertically align" semantically sparse field ontologies at the lowest tier with semantically rich landform object ontologies at the uppermost tier. Tiered ontology design has been suggested for spatiotemporal database design in [11], and particularly in the context of hierarchically integrating terrain field and object database models [54]. We recommend this approach, for Linked Topographic Data, in general, because queries about topographic objects or locations often will need to be run across multiple linked topographic databases. Surface networks provide only a partial account of a surface; many use cases will use surface network elements to limit the search domain on the original surface. Similarly, even if topographic landforms or land cover areas are available as objects with pre-calculated properties, users may still need access to the original elevation field and imagery for the covered area to calculate new properties or for visualization purposes. Without ontologies to explain how such databases and objects relate to each other, there cannot be 'Linked' Topographic Data.

6. Discussion

A universal ontology of space would (ideally) eliminate the need for domain-specific spatial ontologies, but the ontology of space (and time) has been argued philosophically for a long time, and it seems unlikely that a universal ontology will emerge anytime soon. Currently, there is not even an established ontology for the geospatial domain—only standards for geometric representation and reasoning are available. Recently *Descartes-Core* was proposed at one of the GeoVocamps as a community-wide collection of geo-ontology patterns and vocabularies, best-practice guides, examples and case studies, software and services.[14] We are hopeful that the effort will yield ontology patterns that can benefit the SNODP patterns. There are three patterns that we believe will greatly benefit surface network patterns. The biggest benefit will, however, be realized if SNODP can be aligned with a fundamental ontology pattern for 2D fields or surfaces. That would eliminate the need for the ad-hoc *Surface* class and empower users with semantic reasoning with both surfaces and surface networks. Second, the semantics of topological and mereo-topological connections are not explicitly specified in SNODP, and so it stands to undoubtedly benefit if an OWL ontology pattern formalizing mereo-topological semantics can be incorporated in the future. Final Geospatial SNODP will benefit from its alignment with a general purpose Semantic Web ontology pattern that captures the general semantics of *any* network embedded in metric space.

Geospatial SNODP's reliance on GeoSPARQL is a practical choice for enabling the pattern with modest geospatial capabilities. But, even if a comprehensive and foundation ontology of space were to replace GeoSPARQL in the distant future and the patterns were extended to cover all special cases ignored currently, still only semantics pertaining to the *shape* of the surface can be encoded using SNODP and derivative patterns. There also will always remain other surface semantics, which can be extracted only by contextualizing surface values with background domain knowledge. For example, two topographic peaks at elevations of 1000 meters and 6000 meters will be associated with starkly different environmental conditions, but the semantics of terrain elevation as an environmental factor cannot ever be captured within Geospatial SNODP. In fact, even specifying basic elevation semantics (i.e., how it is defined or measured) is not within its scope. This is why surface network ontology must be complemented with domain-specific surface ontologies for sharing surface information.

We also debated the value in aligning SNODP or Geospatial SNODP to a foundational ontology such as DOLCE [35], BFO [18, 55] or SUMO [39]. While ontologies cannot truly fix meaning, they are supposed to restrict the interpretations of the classes and relations of a particular ontology. For example, in SUMO, a surface network element would be declared explicitly as an *Abstract* object. In BFO critical points, lines and areas could be declared as zero, one, or two dimensional *spatial region* entities, respectively, but only if "*spatial region*" covers not just the physical space but also abstract mathematical spaces. In DOLCE, probably surface network elements should be defined as *Abstract* entities, although we wonder if surface network elements corresponding to physical entities, as for terrain features, would be better described as *Mental Objects*, which are *Non-Physical Endurant*, but not *Abstract* entities. The problem seems to be though that none of the choices can completely capture some essential meta-level semantics that can help in general pattern interpretation. In our opinion, the benefits of aligning SNODP to a foundational ontology would be minimal and may even restrict the intended interpretation. Moreover, alignment with these ontologies implies commitment to a much larger body of assertions than is required to complete the relevant tasks. Since the patterns are well documented and complemented by

[14] http://vocamp.org/wiki/GeoVoCampSB2013.

this paper, interested parties can easily align SNODP and its extensions with one or more of the foundational ontologies, if it suits their objectives.

Finally, we comment about the current status of the technical feasibility of realizing surface network instances from surfaces. Computational surfaces are digital approximations of the ideal, smooth, continuous surface, which is often considered a pre-requisite for deriving surface networks. However, in reality, digital surfaces such as DEMs or TINs of topographic surfaces are discrete approximations that will not guarantee the extraction of a consistent surface network for which all theoretical properties hold. The methods of surface networks extraction employed so far include simple manual [68], triangulation [59], and complex surface fitting techniques [69]. Yet, for all practical purposes, fully automated extraction of a complete and consistent surface network is still extremely difficult to implement successfully. Thus, in the near future, users should be ready to accept only subsets of surface network elements (e.g., only critical points) being extracted from surface datasets. Our view is that even such partial surface network data is a step forward and may suffice for many types of surface related queries. It is quite common to instantiate ontologies only partially on the Semantic Web, and it is with this flexibility in mind that we proceeded to design the two patterns.

7. Conclusion

The primary motivation for designing SNODP and Geospatial SNODP is Linked Topographic Data, for which a critical problem remains the interoperability of object and field representations of terrain data. The patterns, if used and adopted by the Semantic Web community, will unlock a wealth of information in surface datasets, currently outside the realm of Semantic Web technologies. Overall, we believe that our patterns meet the general expectations of an ODP:

i. *Expressive*: SNODP clearly captures surface network element relationships and also specifies their embedding in the surface. The spatial properties of GeoSPARQL are quite adequate to cover the needs of the Geospatial SNODP pattern. Both patterns are quite expressive since OWL is the primary Semantic Web language and its use opens up the patterns to a wide range of inferences.

ii. *Simple:* We deliberately eschewed many advanced surface network issues in designing SNODP and Geospatial SNODP since ontology patterns are about data sharing, not comprehensive formalization of all domain semantics. The patterns support a minimum number of classes that are needed to semantically annotate typical surface network datasets included in the patterns. We decided to create two separate patterns to isolate the core topological foundations from the extensions that would additionally support metric space capabilities. These are patterns for sharing scientific data, but anybody familiar with even the basics of surface networks can reuse these patterns easily. The names of the classes are in accordance with the literature and property names are chosen to clearly communicate their function. We foresee very little unintended interpretations of this pattern.

iii. *Reusable:* The two patterns are quite generalized to be reusable easily. SNODP is applicable to any kind of surface, whereas Geospatial SNODP is specialized only for geospatial surfaces, not just for terrain data. The use of GeoSPARQL is also not prescriptive in any way since it can be easily replaced with a geospatial ontology of choice. For other domains, available spatial ontologies must be substituted for GeoSPARQL to realize extensions of SNODP that support representation of metric spatial properties in accordance with domain principles.

iv. *Scalable:* Surface networks are data reduction patterns since they are much more compact than any other form of expressing the surface. Surface networks can themselves be further generalized, but multiple surface networks can be stored in one combined dataset and still be semantically annotated by the patterns. The separation of the *Feature* and *Geometry* concepts in GeoSPARQL also ensures that multiple geometric realizations of a surface network can be linked and share common critical points.

v. *Well-documented:* The OWL ontology files contain comments to explain our intentions, but this paper should really serve to clarify the contexts of use that inspired us to design the patterns, and how we intend them to be used.

As stressed in this paper, surface network ontologies bring to the fore a fundamental problem that cuts across all topographic sub-domains—i.e., the technical problem of integrating field and object based conceptualizations. However, for Linked Topographic Data to matter and be functional, we clearly need *several* other topographic and landscape reasoning ontologies for representing both geoscientific and culturally based concepts relating to topographic eminences, hydrological features, maritime features, vegetation, soil, lithology, settlements, and other subdomains. There exist data related to such domains already, but there is no standardized way to describe the semantics of those topographic sub-domains. Through this work, we hope to foster more discussion and encourage others to think about this problem that prevents the Semantic Web from harnessing information in surface datasets of so many different types. Based on our experience, we recommend GeoVocamps as an appropriate venue and vehicle for exploring ontology patterns for Linked Topographic Data.

Acknowledgements

The authors acknowledge the USGS for providing its Reston office facilities for the GeoVoCampSOCoP12 workshop. They also thank GeoVoCamp participants and organizers, including members of SOCoP, for organizing this workshop and providing initial feedback to this work. Gaurav Sinha acknowledges funds provided by Ohio University for attending the workshop. Boleslo Romero acknowledges funding and support of the USGS Center of Excellence for Geospatial Information Science (CEGIS) and the CESU program. Gary Berg-Cross acknowledges funding from the National Science Foundation (NSF) under Grant Number 0955816 (INTEROP - Spatial Ontology Community of Practice). Anand Padmanabhan acknowledges NSF funding under Grant Number OCI-1047916.

References

[1] Barsalou L, 2003. Situated Simulation in the Human Conceptual System. *Language and Cognitive Processes*, *5(6)*, 513–562.

[2] Battle R, and Kolas D, 2012. Enabling the Geospatial Semantic Web with Parliament and GeoSPARQL. *Semantic Web, 3(4)*, 355-370.

[3] Berners-Lee T, 2008. *Linked Open Data*. Accessed August 30, 2013 @ http://www.w3.org/2008/Talks/0617-lod-tbl/#(1)

[4] Bishr Y, 2006. Geospatial Semantic Web. In: Rana S, and Sharma J (eds.), *Frontiers of Geographic Information Technology*, 139-154, Springer-Verlag, Berlin Heidelberg.

[5] Brodaric B, and Gahegan MF, 2010. Ontology Use for Semantic e-Science. *Semantic Web*, *1(1-2)*, 149-153.

[6] Brodaric B, and Gahegan MF, 2007. Experiments to Examine the Situated Nature of Geoscientific Concepts. *Spatial Cognition & Computation*, *7(1)*, 61-95.

[7] Cayley A, 1859. On Contour Lines and Slope Lines. *London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *18*, 264-268.

[8] Clark P, Thompson J, and Porter B, 2000. Knowledge Patterns. In: Cohn AG, Giunchiglia F, and Selman B, (edts.), *KR2000: Principles of Knowledge Representation and Reasoning*, 591 – 600, Morgan Kaufmann, San Francisco, USA.

[9] Egenhofer MJ, 2002. Toward the Semantic Geospatial Web. In: *Proceedings of the 10th ACM international symposium on Advances in Geographic Information Systems, GIS '02*, ACM, 2002, 1-4.

[10] Fowler M, 1997. *Analysis Patterns*: *Reusable Object Patterns*, Addison Wesley, Boston, MA, USA.

[11] Frank A, 2001. Ontology for Spatio-temporal Databases. In: Sellis T et al. (eds.): *Spatio-Temporal Databases*: *The CHOROCHRONOS Approach, LNCS 2520*, 9-77, Springer-Verlag, Berlin, Heidelberg.

[12] Gamma E, Helm R, Johnson R, and Vlissides J, 1995. *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley, Boston MA, USA.

[13] Gangemi A, 2005. Ontology Design Patterns for Semantic Web Content. In: Gil Y, Motta E, Benjamins VR, and Musen MA (eds.), *The Semantic Web – ISWC 2005*, *Lecture Notes in Computer Science* *Volume 3729*, 262-276, Springer-Verlag, Berlin Heidelberg,.

[14] Gangemi A, and Presutti V, 2009. Ontology Design Patterns. In: Staab S and Studer R (eds.), *Handbook on Ontologies, International Handbooks on Information Systems*, *2nd Edition*, 221-243, Springer, Berlin.

[15] Goodwin J, 2005. What have Ontologies ever done for us – Potential Applications at a National Mapping Agency. In: *Proceedings of OWL: Experiences and Directions (OWLED) 2005*, November 11-12, 2005, Galway, Ireland.

[16] Goodwin J, Dolbear C, and Hart G, 2008. Geographical Linked Data: The Administrative Geography of Great Britain on the Semantic Web. *Transactions in GIS*, *12*, 19-30.

[17] Grau BC, Horrocks I, Kazakov Y, and Sattler U, 2008. Modular Reuse of Ontologies: Theory and Practice. *Journal of Artificial Intelligence Research*, *31(1)*, 273–318.

[18] Grenon P, 2003. BFO in a Nutshell: A Bi-Categorial Axiomatization of BFO and Comparison with DOLCE. *Technical report, Universitat Leipzig, Faculty of Medicine, Institute for Formal Ontology and Medical Information Science (IFOMIS)*, June 6, 2003. Accessed August 30, 2013 @ http://www.ifomis.org/Research/IFOMISReports/IFOMIS%20Report%2006_2003.pdf

[19] Guarino N, 1998. Formal Ontology and Information Systems. In: Guarino N (ed.), International Conference on Formal Ontology in Information Systems (FOIS1998), 3-15, IOS Press, Trento, Italy.

[20] Hay DC, 1996. *Data Model Patterns: Conventions of Thought*, Dorset House Publishing.

[21] Helman JL, and Hesselink L, 1991. Visualizing Vector Field Topology in Fluid Flows. *IEEE Computer Graphics and Applications*, *11(3)*, 36-46.

[22] Hitzler P, Krötzsch M, Parsia B, Patel-Schneider PF, and Rudolph S, 2012. *OWL 2 Web Ontology Language Primer (Second Edition)*. Accessed August 30, 2013 @ http://www.w3.org/TR/owl2-primer

[23] Hu Y, Janowicz K, Carral D, Scheider S, Kuhn W, Berg-Cross G, Hitzler P, Dean M, and Kolas D, 2013. A Geo-Ontology Design Pattern for Semantic Trajectories. In: *Proceedings of*

COSIT 2013, Conference on Spatial Information Theory, Scarborough, UK, September 2-6, 2013.

[24] Janowicz K, 2010. The Role of Space and Time for Knowledge Organization on the Semantic Web. *Semantic Web, 1(1)*, 25-32.

[25] Janowicz K, and Hitzler P, 2012. The Digital Earth as a Knowledge Engine (Editorial), *Semantic Web, 3(3)*, 213-221.

[26] Janowicz K, Scheider SA, Pehle TB, and Hart G, 2012. Geospatial Semantics and Linked Spatiotemporal Data – Past, Present, and Future (Editorial), *Semantic Web, 3(4)*, 321-332.

[27] Jhingran A, 2008. Web 2.0, Enterprise 2.0 and Information Management. Or Different Approaches in making 1 + 1 = 11. *Linked Data Planet Conference*, New York, USA, June 17-18.

[28] Kuhn W, 2010. Modeling vs. Encoding for the Semantic Web. *Semantic Web, 1(1)*, 11-15.

[29] Kuhn W, 2009. Semantic Engineering. In: Navratil G (ed.), *Research Trends in Geographic Information Science. Springer-Verlag, Lecture Notes in Geoinformation and Cartography*, 63-74.

[30] Kuhn W, 2005. Geospatial Semantics: Why, of What, and How? *Journal on Data Semantics, 3*, 1-24.

[31] Mark DM, 1978. Topological Properties of Geographic Surfaces: Applications in Computer Cartography. *Harvard Papers on Geographic Information Systems*, Laboratory for Computer Graphics and Spatial Analysis, Harvard University, Cambridge, MA, 5, Mark 1-Mark 11.

[32] Mark DM, and Sinha G, 2006. Ontology of Landforms: Delimitation and Classification of Topographic Eminences. In: Raubal M, Miller H, Frank A, and Goodchild M (eds.), *Geographic Information Science, 4th International Conference, GIScience 2006, Extended Abstracts*, Münster, Germany, September 20 - 23, 2006.

[33] Mark DM, and Smith B, 2004. A Science of Topography: From Qualitative Ontology to Digital Representations. In: Bishop MP and Shroder JF (eds.), *Geographic Information Science and Mountain Geomorphology*, 75-100, Chichester, Springer-Praxis, UK.

[34] Marshall CC, and Shipman FM, 2003. Which Semantic Web? In: *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia HYPERTEXT '03*, August 26-30, 2003, Nottingham, UK.

[35] Masolo C, Borgo S, Gangemi A, Guarino N, and Oltramari A, 2003. *WonderWeb Deliverable D18 Ontology Library (final)*. Accessed August 30, 2013 @ http://wonderweb.semanticweb.org/deliverables/documents/D18.pdf.

[36] Maxwell JC, 1870. On Hills and Dales. *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, 40*, 421-427.

[37] Mizen H, Dolbear C, and Hart G, 2005. Ontology Ontogeny: Understanding How an Ontology is Created and Developed. In: Rodríguez MA, Cruz IF, Egenhofer MJ, and Levashk S (eds.), *Proceedings of the First International Conference on GeoSpatial Semantics (GeoS) 2005*, November 29-30, 2005, Mexico City, Mexico, 15-29.

[38] Morse M, 1925. Relations between the Critical Points of a Real Function on n Independent Variables. *Transactions of the American Mathematical Society, 27*, 345-396.

[39] Niles I, Pease A, 2001. Towards a Standard Upper Ontology. In: *Proceedings of the International Conference on Formal Ontology in Information Systems (FOIS)*, 2-9, Ogunquit, Maine, USA.

[40] Open Geospatial Consortium (OGC®), 2012. *OGC GeoSPARQL - A Geographic Query Language for RDF Data. OGC® Document No. 11-052r4, Version 1.0.* URL: http://www.opengis.net/doc/IS/geosparql/1.0

[41] Pfaltz J, 2004. Foreword. In: Rana S (Ed.), *Topological Data Structures: An Introduction to Geographical Information Science*, John Wiley & Sons, Ltd., Chichester, England.

[42] Pfaltz J, 1976. Surface Networks. *Geographical Analysis, 8*, 77-93.

[43] Presutti V, and Gangemi A, 2010. Content Ontology Design Patterns as Practical Building Blocks for Web Ontologies. In: Spaccapietra S et al. (eds.), *Proceedings of ER2008, 27th International Conference on Conceptual Modeling*, Springer-Verlag Berlin, Heidelberg, 128-141.

[44] Rana SS, 2004a. Surface Networks: New Techniques for their Automated Extraction, Generalisation and Application. *Unpublished PhD Dissertation*, Department of Geomatic Engineering, University College London, University of London, UK.

[45] Rana, S. (Ed.), 2004b. *Topological Data Structures: An Introduction to Geographical Information Science*, John Wiley & Sons, Ltd., Chichester, England.

[46] Rana SS, and Morley J, 2002. Surface Networks. *UCL Working Paper Series, Paper 43*, Center for Advanced Spatial Analysis (CASA), University College of London, University of London, UK. Accessed August 30, 2013 @ http//www.casa.ucl.ac.uk/working_papers/paper43.pdf.

[47] Raubal M, and Adams B, 2010. The Semantic Web Needs More Cognition. *The Semantic Web, 1(1)*, 69-74.

[48] Reeb G, 1946. Sur les Points Singuliers d'une Forme de Pfaff Completement Intergrableou d'une Fonction Numerique. *Computes Rendus Acad. Sciences Paris, 222*, 847-849.

[49] Reech M, 1858. Proprieté générale des surfaces fermées. *Ecole. Polytech J., 37*, 169-178.

[50] Sadahiro Y, 2001. Analysis of Surface Changes using Primitive Events, *International Journal of Geographical Information Science, 15(6)*, 523-538.

[51] Sadahiro Y, and Masui M, 2004. Analysis of Qualitative Similarity between Surfaces, *Geographical Analysis, 36 (3)*, 217-233.

[52] Schneider B, 2005. Extraction of Hierarchical Surface Networks from Bilinear Surface Patches. *Geographic Analysis, 37(2)*, 244-263.

[53] Sinha G, 2008. *Delineation, Characterization, and Classification of Topographic Eminences*. Unpublished PhD Dissertation, Department of Geography, University at Buffalo, Buffalo, NY, USA.

[54] Sinha G, and Mark DM, 2010. Cognition-Based Extraction and Modelling of Topographic Eminences. *Cartographica: The International Journal for Geographic Information and Geovisualization, 45(2)*, 105-112.

[55] Smith B, 1998. The Basic Tools of Formal Ontology. In: *Formal Ontology in Information Systems Frontiers in Artificial Intelligence and Applications*, IOS Press, Washington, DC, 19–28.

[56] Smith B, and Mark DM, 2003. Do Mountains Exist: Towards an Ontology of Landforms. *Environment and Planning B: Planning and Design, 30*, 411-427.

[57] Smith B, and Varzi A, 2000. Fiat and Bona Fide Boundaries. *Philosophy and Phenomenological Research, 60(2)*, 401-420.

[58] Svatek V, 2004. Design Patterns for Semantic Web Ontologies: Motivation and Discussion. In: *Proceedings of the 7th Conference on Business Information Systems, BIS04*, Poznan, Poland, April 21-23, 2004.

[59] Takahashi S, Ikeda T, Shinagawa Y, Kunii TL, and Ueda M, 1995. Algorithms for Extracting Correct Critical Points and Constructing Topological Graphs from Discrete Geographical Elevation Data. *Computer Graphics Forum, 14(3)*, 181-192.

[60] Usery EL, and Varanka DE, 2012. Design and Development of Linked Data for The National Map. *The Semantic Web, 3(4)*, 371-384.

[61] Varanka DE, and Usery EL, 2010. Ontological Issues for The National Map. *Cartographica: The International Journal for Geographic Information and Visualization*, *45(2)*, 103-104.

[62] Varanka DE, 2011. Ontology Design Patterns for Complex Topographic Feature Types. *Cartography and Geographic Information Science, 38(2)*, 126-136.

[63] Warntz W, 1966. The Topology of a Socio-Economic Terrain and Spatial Flows. *Papers, Regional Science Association, 17*, 47-61.

[64] Warntz W, and Woldenberg M, 1967. Concepts and Applications -- Spatial order. *Harvard Papers in Theoretical Geography No. 1*, *Office of Naval Research Technical Report, Project 389-147*, Harvard University, Cambridge, MA, USA.

[65] Wiegand N, Kolas D, and Cross GB, 2010. Guest Editorial: Intersecting Semantic Web and Geospatial Technologies. *Transactions in GIS, 14(2)*, 93-95.

[66] Wilson JP and Gallant JC, 2000. *Terrain Analysis: Principles and Applications*. John Wiley & Sons, Inc., New York, NY, USA.

[67] Wolf GW, 1990. Metric Surface Networks. In: *Proceedings of the 4th International Symposium on Spatial Data Handling*, Zurich, Switzerland, 844-856.

[68] Wolf GW, 1984. A Mathematical Model of Cartographic Generalization. *Geo-Processing*, *2 (3)*, 271-286.

[69] Wood JD, 1998. Modelling the Continuity of Surface Form Using Digital Elevation Models. In: *Proceedings of 8th International Symposium on Spatial Data Handling*, Vancouver, Canada, 725-736.

[70] Wood JD, 2000. Construction of Weighted Surface Networks for the Representation and Analysis of Surface Topology. In: *Proceedings, 5th International Conference on GeoComputation, 2000*, University of Greenwich, UK.