# A Consideration of Geospatial Feature Formation in Linked Open Vocabularies

Dalia VARANKA
U.S. Geological Survey, Email: dvaranka@usgs.gov

## Abstract

A geospatial feature is a central concept representing a range of material entities and abstractions based on those entities for geographic information systems (GIS). The design of geospatial features, called just 'features' in this work, is a key concern in the development of a geospatial vocabulary for linked open data (LOD) for the National Hydrography Dataset (NHD) of the U.S. Geological Survey (USGS). Background and analysis of the vocabulary are discussed to provide context for attempts to integrate feature concepts as described in the geographic information literature with upper ontology design principles. The presentation discusses implications of the vocabulary design.

The motivation for the NHD vocabulary project, called the GIS NHD because it consists of a representation of NHD as is modeled in GIS, is to articulate and support LOD through appropriate internal ontology relations, and externally to more widely established LOD vocabularies. GIS NHD has two predominant modules: NHD Feature and an instance module of the structure of an ontology-driven gazetteer. GIS NHD is not a general-purpose surface water domain ontology, but is based on the conversion of GIS data attributes to triples and reflects the schema from the NHD data model. A structure of the surface water domain can be abstracted from the NHD GIS data model with further processing. The gazetteer links to the Feature Module by geospatial feature type, and supports searches and data handling using GeoSPARQL. External linkages are possible to vocabularies for provenance, observations and measurements, and mapping, for example. A class called NHD Event supports linkages to hydrology data based on co-location.

The primary application for the GIS NHD ontology is to structure NHD data for users who are working with semantic technology. The vocabulary excludes instances, which are available by downloading and converting GIS data using a custom program. Feature geometry objects required by GeoSPARQL are compiled by the conversion program. NHD LOD users can optionally define and retrieve their own geospatial features from basic primitives by forming GeoSPARQL triple graph patterns. For example, features can be formed using other properties, such as everything with a certain name, although the complete data graph converted from a GIS table would be required, not just individual geometry objects.

In adapting the GIS NHD feature model to ontology, an upper ontology structure was used to maintain ontologically consistent classes and properties. An inconsistency of feature concepts as defined in the GIS literature in contrast to the upper ontology suggests that feature concepts may need to be clarified in light of the semantic data model. For example, a GIS feature does not differentiate between material objects whose existence is concrete,

though their identifying terms may vary, such as ‘hill,’ and realizable entities, which can include socially assigned roles, and whose identify may be more abstract, such as ‘ski slope.’ Features as used in GIS may benefit from a wider representation range of specific qualities; the literature on ‘real world’ ontology may benefit from a broader consideration of geographical space in the formalization of feature types.