

Semantics of Landscape: Providing Information About the Landscape via The Semantic Web

David M. Mark
NCGIA & Geography
University at Buffalo, NY 14261 USA
dmark@buffalo.edu

Semantics is all about meaning, and semantic interoperability is about the preservation of meaning as information is exchanged among software, databases, and people. Words in natural language play a critical role in the transmission of meaning from one person to another, in the sharing of concepts across individuals within speech communities. Recent developments in ontologies for knowledge representation also have words and their meanings playing a central role, and mainly formalize the definitions of and relations among words (terms) for a specific domain. The key role of words and language in semantics is recognized in the ubiquitous semiotic triangle, which relates signs (symbols, including words), concepts (mental ideas), and referents (in the real world). The Semantic Web promises to take the world-wide web to new levels of utility by building upon ontologies, inference procedures, and definitions, but if universally agreed-upon definitions are not available, 'Tower of Babel' problems will still occur.

The natural landscape is a very important domain for human activities, and is particularly important in geospatial intelligence. The landscape may be loosely defined as the larger elements of the human environment, those elements that can be perceived and understood from a distance of hundreds of meters or more, such as hills and valleys, lakes and rivers, forests and deserts. However, representation and exchange of information about the landscape in word-centered ontologies such as those alluded to above, for delivery through the Semantic Web, presents a number of significant research challenges that are rare or absent for other domains.

- **Semantics usually deals with objects but geospatial information is often fields or images.** While the semantic web is word-centered, the great majority of existing geospatial information about natural landscapes is in the form of fields (digital elevation models) or images (mainly from remote sensing). Serving of field-based or image-based geospatial information via the semantic web will require robust computational methods for identifying and delimiting landscape features present implicitly in those fields and images.
- **Classifications of landscape elements vary across cultures and languages.** For inorganic natural domains such as landforms and water bodies, natural variation usually forms a continuum, and categories for landscape elements thus have a somewhat arbitrary component, contingent on landscape variation, lifestyle of the people, and linguistic and cultural history.
- **Dependency of landscape objects on classifications.** For many domains, the entities dealt with are bona fide objects that exist as objects independently of their definitions. For example, individual higher organisms such as birds or butterflies are easily perceived as 'things' even if not identified as to type. In contrast, many landscape elements are not bona fide objects, but are parts of a continuous land

- surface, and thus different definitions of landscape element types may lead to a different subdivision of the land into objects.
- **The above issues interact!** The same landscape, represented by the same field-like data models, may produce quite different delimited objects and object classes according to the definitions of words in different languages. Feature-extraction methods probably will need to be language-specific. A word-based ontology of a type of natural landscape developed on the basis of one natural language will likely be inadequate for other languages.

The extent of these 'problems' is not yet known but several solid examples suggest that these issues will be a substantial impediments to the development of multilingual semantic interoperability for landscape data using the word-centered approaches now common for knowledge-representation ontologies and for the Semantic Web.