

Designing a Framework to Support the Semantic Interoperability of Geospatial Information

Mark Gahegan, Boyan Brodaric and Masahiro Takatsuka

Department of Geography
302 Walker Building
The Pennsylvania State University
University Park, PA 16802, USA
mark@geog.psu.edu
<http://www.geog.psu.edu/~mark/>

Where is the wisdom we have lost in knowledge?

Where is the knowledge we have lost in information?

--T.S. Eliot, *The Rock*, 1934.

Data and information mean different things to different people. Within a large information community, those involved in information production are not necessarily the end users. Instead, teams and individuals work to produce useful information products, others are responsible for disseminating this information, and yet others use it operationally. Together they form a network of people and machines that fill the roles of producers, suppliers and consumers in an ‘information economy’. This is especially true when considering geospatial data, since it originates from a variety of sources (e.g. remotely-sensed imagery, field surveys, existing maps, aerial photography) and is applied in a variety of different settings (e.g. mapping, hydrology, landcover classification, disaster management). Each of these diverse applications involve different groups of analysts who bring unique perspectives on: the data, the systems and models in which the data resides and the underlying technology by which the data is uncovered, assembled, transformed and analyzed.

From a semantic perspective we can say that the *meaning* of geospatial information is constructed, shaped and changed by the interaction of people and systems. But this interaction is not captured, and information becomes easily separated from the context by which it is given meaning. This defining context has organizational, system and human dimensions. With the advent of interoperable GIS, this problem is compounded because it is now not difficult to remove a dataset (or indeed a method or a whole model) far from the context that initially gave it meaning. Within an organization, information is often constructed incrementally, with successive communities or individuals adding to or modifying meaning through a series of different processes. Those not directly involved in the creation or uncovering of information may not share in this evolving schema fully, neither is its essence conveyed in the existing data models of the systems in which the information resides, nor can it be associated with datasets or methods in any formal sense and made interoperable. One might even go so far as to say that a large part of the meaning exists only in the minds of key individuals, with no formal expression whatsoever in software or data. As a consequence, we are faced with the difficult task of interpreting meaning through a *veil* of implicit and missing knowledge. Such a situation makes it difficult to apply data correctly, repeat analysis and communicate results—three cornerstones of the scientific method. Needless to say, this is a great impediment to coordinated decision making.

This paper describes a novel research program to establish a *semantic framework* for managing the meaning associated with geospatial information. The framework describes the various transformations that the geospatial information undergoes, thereby facilitating better understanding, providing some measure of repeatability of analysis and improving communication in the hope of promoting best practice in bringing geospatial information to bear. The development of such a framework is described, along with mechanisms to encode associated rules and knowledge structures that might oversee the transformation and interpretation processes. Specifically, three major components are discussed (Figure 1):

1. A *framework* within which to manage the entire process of semantic representation and semantic transformation.

2. Extensions of existing systems and processing activities so that their semantic properties are represented and can be made available externally. These extensions are termed *cognitive schemata*.
3. A suite of operations (methods) that allow semantic information to be successfully applied and exchanged, and that can adapt when the data changes following transformations or interoperation.

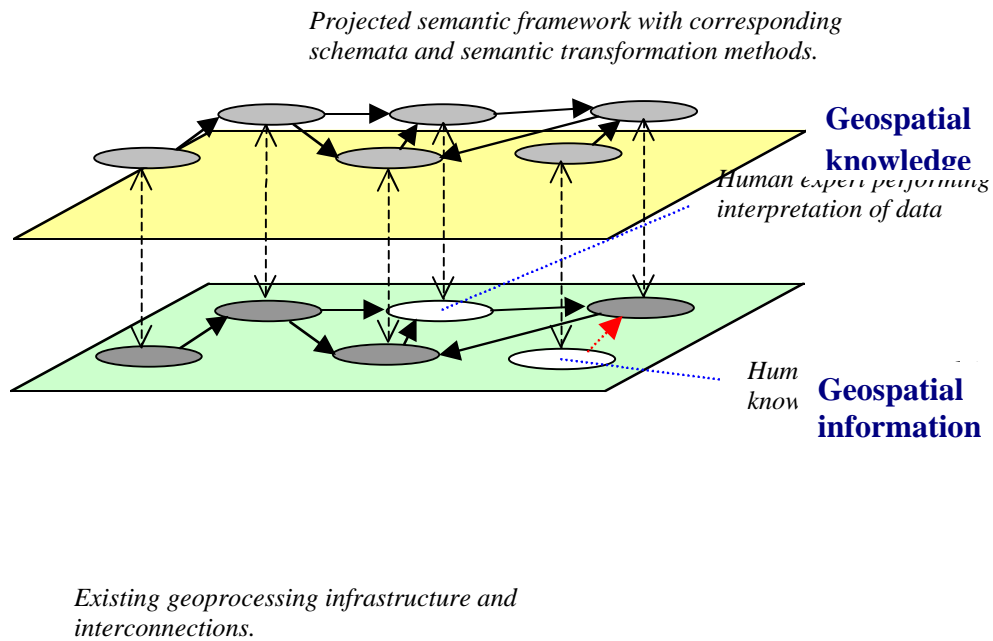


Figure 1. A semantic framework imposed on existing geospatial infrastructure.