



## **Semantic Challenges facing AI Applications in Earth Sciences**

Stephen M. Richard (1), David A. Poole (2), Clinton P. Smyth (1), Christopher Ahern (1), and Victoria Wang (1)

(1) Minerva Intelligence, Semantics, Vancouver, BC, CA (srichard@minervaintelligence.com), (2) Department of Computer Science, Univ. of British Columbia, Vancouver, CA

Use of technical terminology for AI applications requires a level of precision and consistency in the use of terms that is different from spoken language. Progress in natural language processing has been steady, but the most straightforward approach to machine actionable knowledge representation remains the use of carefully constructed taxonomies that follow Aristotelian logic by defining concepts through binding to a more general concept (genus) and using a consistent set of differentiating properties (differentia) to distinguish concepts with the same genus. Property value vocabularies are constructed as taxonomies as well, which tend to be simpler and less problematic to define and agree upon. The asserted properties and their hierarchical relationships can then be used to consistently infer hierarchical relationships between concepts they are used to define. The resulting taxonomy is easier to maintain and accounts for multiple parent relationships. This approach has been successfully applied in applications matching geologic occurrence descriptions with models for mineral deposits to generate exploration targets, and models for landslide genesis to generate hazard maps. In recent years, significant progress has been made in the development of vocabularies for geoscience information interchange by the IUGS CGI Geoscience Terminology Working Group, and by the EU INSPIRE program. Implementation of these vocabularies as formal Aristotelian taxonomies is providing a foundation to improve software applications for more accurate model-instance matching and to extend their scope to other fields such as flood hazards and environmental remediation.