

HANDLING GEOLOGICAL TERMINOLOGY WITHIN AN ENVIRONMENTAL THESAURUS – THE EARTH THESAURUS

DI FRANCO SABINA ¹, PLINI PAOLO ¹, MAZZOCCHI FULVIO ¹

1 - CNR-IIA--EKOLab, Montelibretti (RM)

Presenter e-mail: plini@iia.cnr.it

Key terms: thesaurus; Earth Sciences; terminology; classification

Around 2.500 terms, coming from different sources, related to Earth sciences but environmentally relevant have been included in the EARTH thesaurus which is the result of a research project at CNR. Terms have been selected, checked, validated and then put into the thesaurus.

The organization of environmental knowledge should rely on a referral grid for the analysis of reality, for the classification of concepts and for the explanation of their "more essential" meanings.

Each concept has an intrinsic semantic complexity which emphasizes particular and contingent aspects. Solid conceptual foundations are essential for the management and the integration of conceptual and terminological heterogeneity of information systems.

EARTH is based on a faceted structure. Its classification scheme was built utilizing a deductive-inductive approach and is based on the identification and adoption of a system of categories. The Thesaurus vertical structure is founded on categories; it is organized in a framework of different levels and classification knots, and it comprises hierarchical relationships.

Terms have been arranged in a vertical structure acting as a tool that highlights the "primary meaning" of terms. The result is a semantic reference system, stable and partially independent from the context. Due to the rigidity of the vertical structure the planned model envisages the possibility of complementing the faceted structure with a system of themes crossing the vertical structure and forming a matrix system. Themes' function is to reassemble terms that in the vertical structure have been scattered under the more general referral concept.

The traditional semantic relationships (associative, equivalence and hierarchical) has been enriched through an extended set of subrelationships. The enrichment and the logical clarification of the relationships, reinforce their potential utilization for navigating on a conceptual basis.

The increase and the specification of the associative relationships will strengthen the capacity of the system to represent the interconnected feature of the conceptual area.

Earth sciences terminology has historically posed some critical points, well-known from the first attempt to apply the Linnaeus taxonomy of organisms to rocks, minerals and soils (E. Mendes da Costa). These criticalities are due to the effort to organize them in a strictly hierarchical scheme. Moreover the Earth sciences include a great number of different topics: from paleontology to hydrogeology, from mineralogy to engineering geology; each of these applied disciplines has its own specific terminology and poses different problems in terms of conceptual organization. In EARTH the geological terms, as all the others, belong to three macrocategories: Entities, Attributes and Dynamic aspects; they have been distributed in different hierarchical branches under specific facets and node labels. In addition a system of related terms establishes a direct association between terms that are not hierarchically-based linked. The result is a network of interrelated concepts, complex in its structure but easy to browse avoiding unnecessary noise and without losing the precision in the navigation.

I.e. the term "coal" has been included both in the following path ENTITIES - MATERIAL ENTITIES - NON LIVING ENTITIES - ARTIFICIAL ENTITIES - < MATERIALS AND PRODUCTS >, to indicate the coal as a material extracted, manufactured and used while the

path for "coal (rocks)" is ENTITIES - MATERIAL ENTITIES - NON LIVING ENTITIES - NATURAL ENTITIES - < EARTH CONSTITUENTS AND MATERIALS > to point out the meaning of the term "coal" as a sedimentary organic rock. The semantic duplicity of the term "coal" has been overcome applying two different parenthetical qualifiers to the terms in order to differentiate the artificial material and the rock, thus avoiding polyhierarchy.