Geophysical Research Abstracts
Vol. 14, EGU2012-10415, 2012
EGU General Assembly 2012
© Author(s) 2012

# Code lists for interoperability – Principles and best practices in INSPIRE

M. Lutz (1), C. Portele (2), S. Cox (3), and K. Murray (4)

(1) EC Joint Research Centre (JRC), Institute for Environment and Sustainability, Italy (michael.lutz@jrc.ec.europa.eu), (2) interactive instruments, Germany, (3) CSIRO Earth Science & Resource Engineering, Australia, (4) INSPIRE Data Specifications Drafting Team

Using free text for attribute values when exchanging geoscience data can lead to a number of problems, e.g. because different data providers and consumers use different languages, terminology or spellings. To overcome these issues, well-defined schemes of codes or concepts, known as code lists[1], are preferred to free text in defining the value domain of an attribute. The "code list" concept is well established in geospatial modelling standards (e.g. ISO 19103[2]), however, it has been used in many different ways. Here we present some considerations relating to code lists and related interoperability requirements in spatial data infrastructures (SDIs), in particular as discussed in the INSPIRE[3] data specifications working groups. These will form the basis for the specification of code list requirements in the INSPIRE Implementing Rules on interoperability of spatial data sets and services[4], which provide binding legal obligations for EU Member States for the interoperable provision of data related to the environment.

Requirements or recommendations for code lists should address the following aspects:

- *Governance*: When modeling an application domain, for each feature attribute whose value is a 'term', should we re-use an existing code list or specify a new code list for the SDI initiative? Use of existing code lists is likely to maximize cross-initiative interoperability.
- *Level of obligation*: For each use of a code list, what is the level of obligation? Is use of a specified code list(s) mandatory or just recommended? This is particularly important where the specifications carry a legal mandate (as in the case of INSPIRE).
- *Extensibility*: Must data providers use only the specified values or may they extend the code list? Are arbitrary extensions allowed or do additional values have to be specialisations of existing values?
- *Specifying values*: For each code list, the allowed values have to be specified, either directly in the specification, or by reference to an existing external vocabulary. In the former case, for each value, an external identifier, one or more labels (possibly in different languages), a definition and other metadata should be specified. In the latter case, the external vocabulary should be characterised, e.g. by specifying
  - the version to be used,
  - the format(s) in which the vocabulary is available,
  - possible constraints (e.g. if only as specific part of the external list is to be used),
  - rules for using values in the encoding of instance data, and
  - the maintenance rules applied to the external vocabulary.

This information is crucial for enabling implementation and interoperability in distributed systems (such as SDIs) and should be made available through a code list registry.

While thus the information on allowed code list values is usually managed outside the UML application schema, we recommend inclusion of «codeList»-stereotyped classes in the model for semantic clarity. Information on the obligation, extensibility and a reference to the specified values should be provided through tagged values.

[1] While traditionally 'code list' refers to a non-hierarchical list of codes, we use the term here to denote also more structured vocabularies or thesauri.

[2] ISO/TS 19103:2005. Geographic information — Conceptual schema language.

[3] Infrastructure for spatial information in Europe, http://inspire.jrc.ec.europa.eu/

[4] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:323:0011:0102:EN:PDF

**Acknowledgements:** The authors would like to thank the INSPIRE Thematic Working Groups, the Data Specifications Drafting Team and the JRC Contact Points for their contributions to the discussions on code lists in INSPIRE and to this abstract.