

3. Qualitative Data Support

Summary

At the moment OpenMI only supports exchange of quantitative data, which is not sufficient for certain environmental domains. Using OpenMI in these domains, or in projects that try to integrate such domains with others, is by consequence difficult. Support for other types of data, in particular qualitative data is needed.

Qualitative is in essence non numerical data (i.e. text, audio, video, categorical), in contrast to quantitative data (i.e. interval and ratio). Ordinal data (i.e. enums) can be considered to be somewhere in between. The previous examples of qualitative data (land use, crop types) can indeed be easily (to some extent) caught as enums (although useful implementation would have some problems; e.g. who defines the enum and how is it shared between models). Other examples (i.e. data as free form text) do not fit into this solution.

In the agro-environmental domain models are used (e.g. CAPRI, CLUE, EfiScen, Integrator) that for example forecast future land use based on scenarios of policy options. Or that generate rotations (sequences of crops farmers will plant) based on input parameters like soil conditions, technological development, weather conditions, etc. Also social and landscape indicators are used and processed that are very difficult to "quantify". For example a person expressing her/his opinion about a possible future landscape.

How to address in Version 1

The proposed solution (as implemented in Alterra's internal OpenMI Java version) is to add a super interface for Data, with IQuality and IQuantity as specialisations. A Quality consists of a number of categories, which can be ordered or not. The definition of the categories and the proper use of them is a responsibility of the models (modeller) not of the OpenMI.

How to address in Version 2

The meta-information about the data (this includes the definition of the categories for qualities, but also of units for quantities) could be stored in an ontology or some other type of glossary making it more flexible to work with and easier for models to use the same definitions. Also IQuality can be improved to support more types of qualitative information.