



Data integration with the Climate Science Modelling Language

Andrew Woolf (1), Bryan Lawrence (2), Roy Lowry (3), Kerstin Kleese van Dam (1), Ray Cramer (3), Marta Gutierrez (2), Siva Kondapalli (3), Susan Latham (2), Dominic Lowe (2), Kevin O'Neill (1), Ag Stephens (2)

(1) CCLRC e-Science Centre (2) British Atmospheric Data Centre (3) British Oceanographic Data Centre

The Climate Science Modelling Language (CSML) has been developed by the NERC DataGrid (NDG) project as a standards-based data model and XML markup for describing and constructing climate science datasets. It uses conceptual models from emerging standards in GIS to define a number of feature types, and adopts schemas of the Geography Markup Language (GML) where possible for encoding.

A prototype deployment of CSML is being trialled across the curated archives of the British Atmospheric and Oceanographic Data Centres. These data include a wide range of data types - both observational and model - and heterogeneous file-based storage systems.

CSML provides a semantic abstraction layer for data files, and is exposed through higher level data delivery services. In NDG these will include file instantiation services (for formats of choice) and the web services of the Open Geospatial Consortium (OGC).