

CICS Science Conference
November 29, 30 & December 1, 2016
College Park, MD, USA

Abstract: Linking netCDF Data with the Semantic Web - Enhancing Data Discovery across Domains

Jim Biard, Jonathan Yu, Mark Hedley, Simon J D Cox, Adam Leadbetter, Nicholas John Car, Kelsey A Druken, Stefano Nativi, and Ethan Davis

Geophysical data communities are publishing large quantities of data across a wide variety of scientific domains which are overlapping more and more. Whilst netCDF is a common format for many of these communities, it is only one of a large number of data storage and transfer formats. One of the major challenges ahead is finding ways to leverage these diverse data sets to advance our understanding of complex problems.

We describe a methodology for incorporating Resource Description Framework (RDF) triples into netCDF files called netCDF-LD (netCDF Linked Data). NetCDF-LD explicitly connects the contents of netCDF files - both data and metadata, with external web-based resources, including vocabularies, standards definitions, and data collections, and through them, a whole host of related information. This approach also preserves and enhances the self-describing essence of the netCDF format and its metadata, whilst addressing the challenge of integrating various conventions into files.

We present a case study illustrating how reasoning over RDF graphs can empower researchers to discover datasets across domain boundaries.