



Science & Technology
Facilities Council

Metadata Model

Experience of Preserving Diverse Information as well as Data

Charlotte Pascoe and Spiros Ventouras

QA4EO, October 18th 2011, Rutherford Appleton Laboratory



**British Atmospheric
Data Centre**

NATIONAL CENTRE FOR ATMOSPHERIC SCIENCE
NATURAL ENVIRONMENT RESEARCH COUNCIL



Overview

- About CEDA
- Computer Simulations
 - Metafor Common Information Model
- Observations of the real world
 - MOLES
 - Metadata Objects for Linking Environmental Sciences
- UK Greenhouse Gas Platform



CEDA data Centres

The Centre for Environmental Data Archival (CEDA) is based at the STFC Rutherford Appleton Laboratory and hosts a range of activities associated with environmental data archives.



The British Atmospheric Data Centre (BADC) , part of the National Centre for Atmospheric Science([NCAS](#)).



The NERC Earth Observation Data Centre, part of the National Center for Earth Observation (NCEO)



The UK [Defra](#) funded main site for the WMO/UNEP Intergovernmental Panel for Climate Change ([IPCC](#)) data distribution centre



CEDA: Major Projects



The UK Climate Impacts Programme



The European Union Framework Seven funded Metadata For Climate project

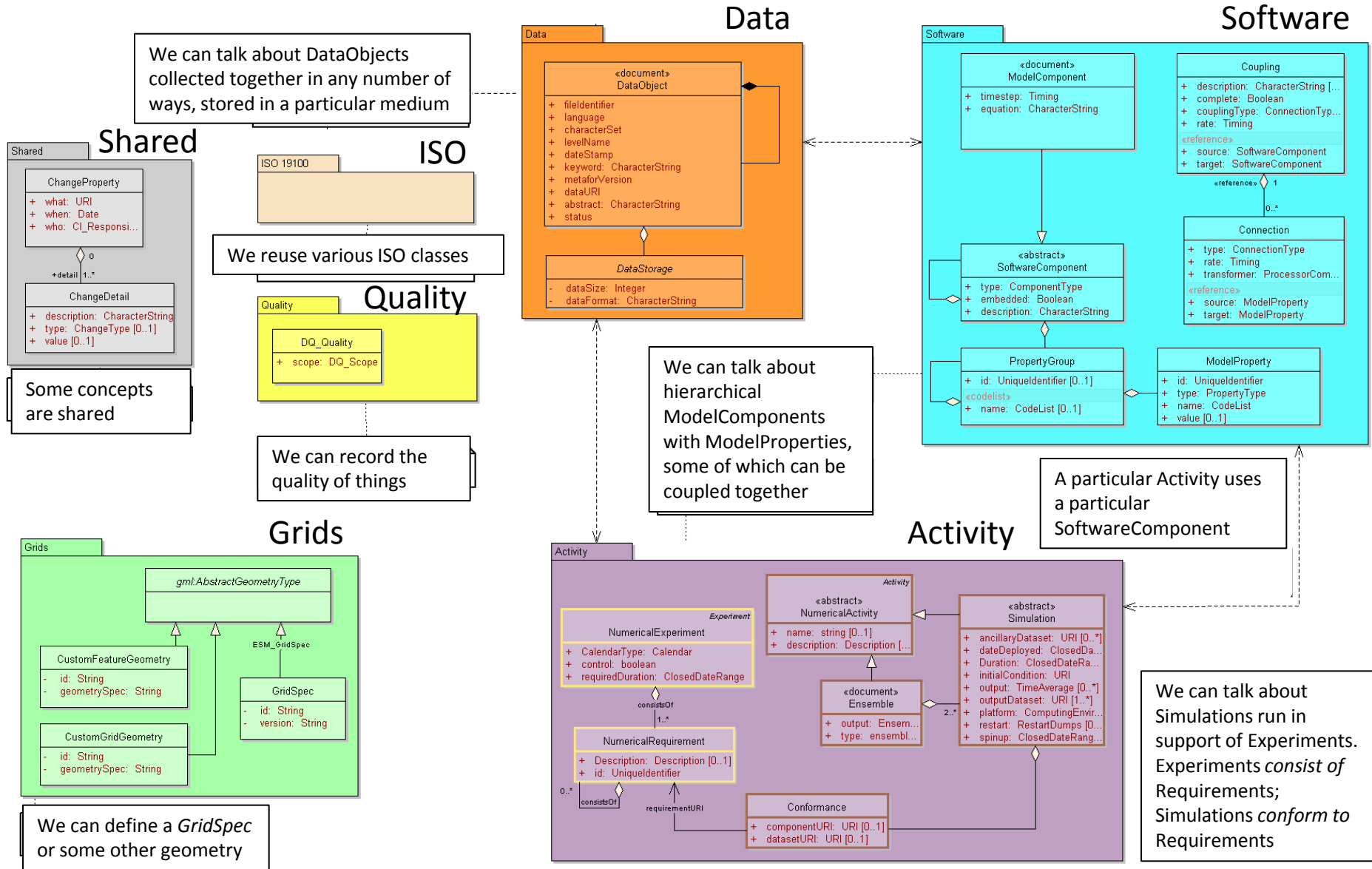


The Overlay Journal Infrastructure for Meteorological Sciences. A [JISC](#) funded project to create an open access journal for meteorological data exploiting overlay journal mechanics.



The European Facility for Airborne Research ([EUFAR](#)) aims at coordinating the operations of the European fleet of instrumented aircraft in the field of environmental research in the atmospheric, marine, terrestrial and Earth sciences.

Common Information Model



We can talk about DataObjects collected together in any number of ways, stored in a particular medium

We reuse various ISO classes

Some concepts are shared

We can record the quality of things

We can talk about hierarchical ModelComponents with ModelProperties, some of which can be coupled together

A particular Activity uses a particular SoftwareComponent

We can define a GridSpec or some other geometry

We can talk about Simulations run in support of Experiments. Experiments consist of Requirements; Simulations conform to Requirements

CMIP5

<http://cmip-pcmdi.llnl.gov/cmip5/index.html>

CMIP5 is meant to provide a framework for coordinated climate change experiments for the next five years and thus includes simulations for assessment in the AR5 as well as others that extend beyond the AR5.

CMIP5 promotes a standard set of model simulations in order to:

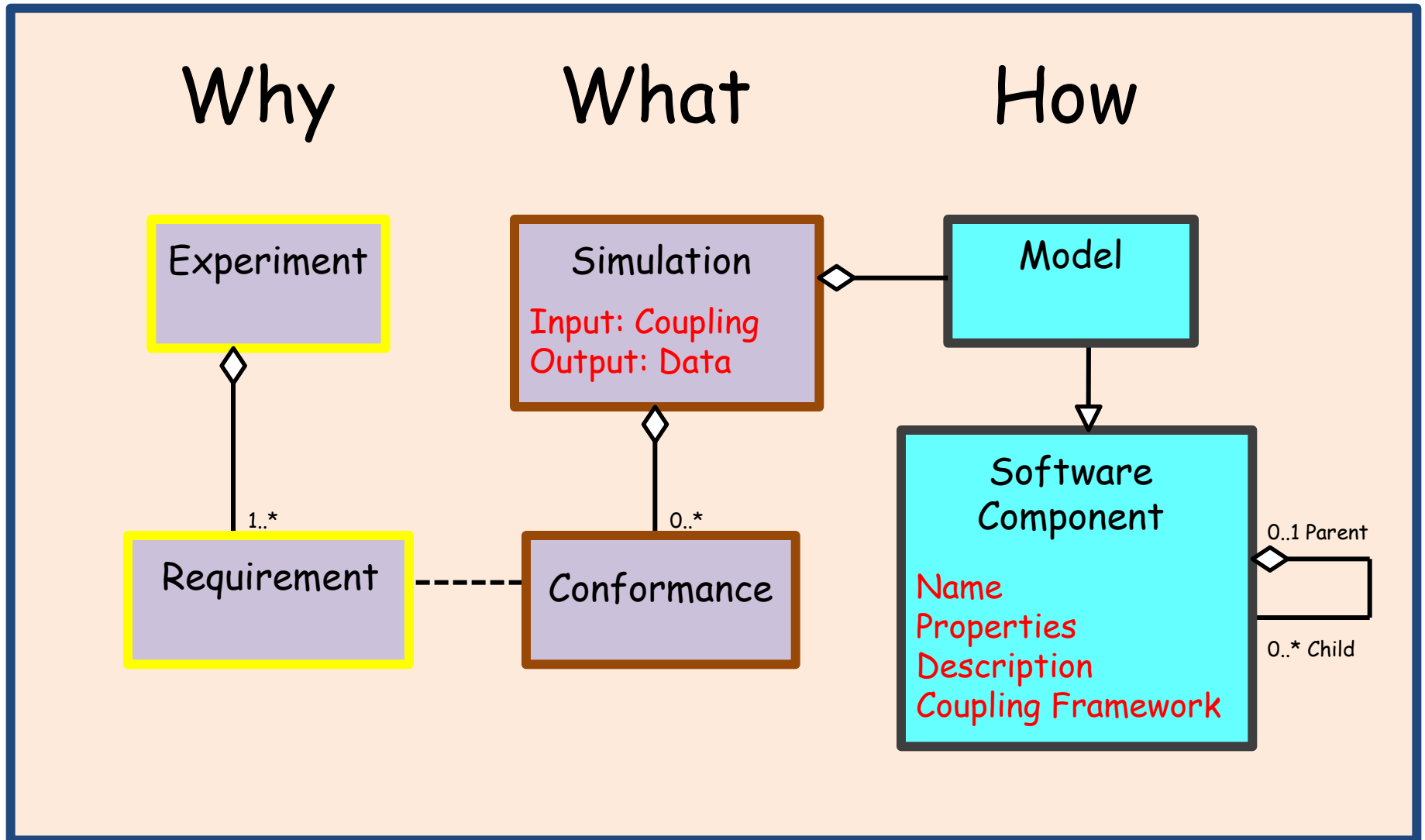
- evaluate how realistic the models are in simulating the recent past,
- provide projections of future climate change on two time scales,
 - near term (out to about 2035) and
 - long term (out to 2100 and beyond), and
- understand some of the factors responsible for differences in model projections,
 - some key feedbacks such as those involving clouds and the carbon cycle

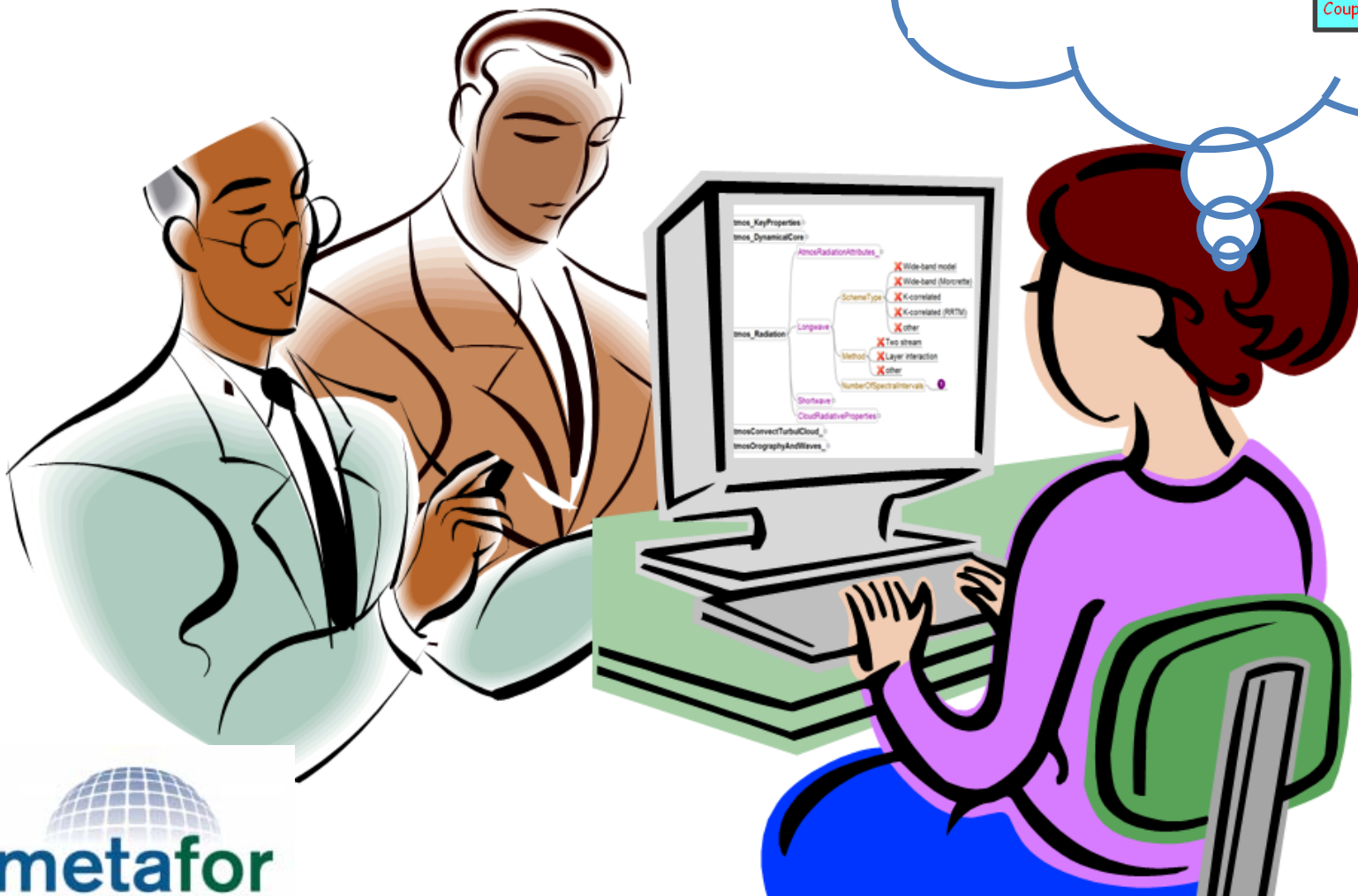
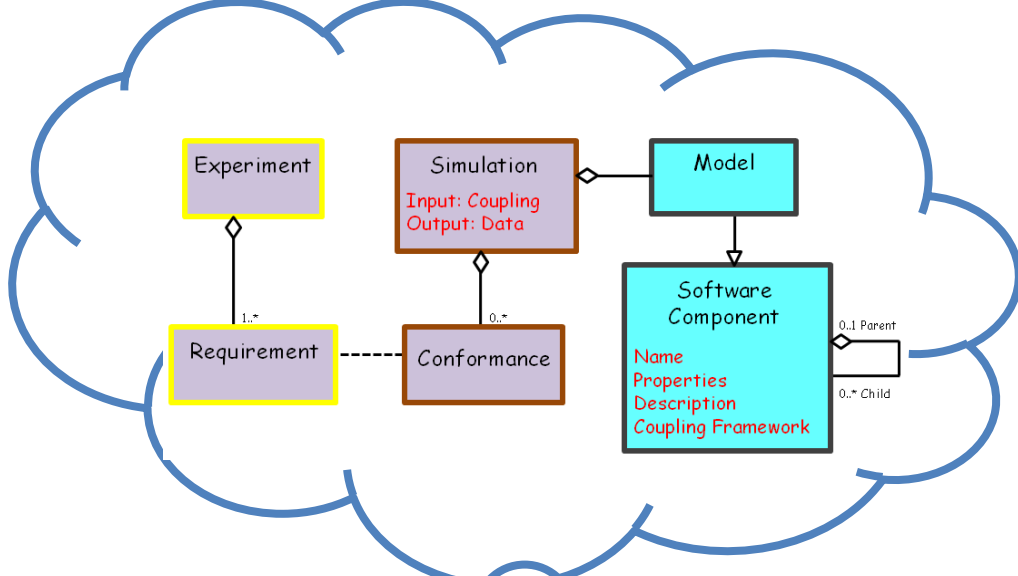
The CMIP5 (CMIP Phase 5) experiment design has following suites of experiments:

- I Decadal Hindcasts and Predictions simulations,
- II "long-term" simulations,
- III "atmosphere-only" (prescribed SST) simulations for especially computationally-demanding models.

CIM for CMIP5

<http://metaforclimate.eu/trac/browser/CIM/tags/version-1.5>





CMIP5 docs in the CIM Portal

TOOLS - CIM VIEWER

CIM XML File:

Simulation: rcp45

Overview Information

shortName	rcp45
longName	future simulation - representative concentration pathway 4.5
description	Representative concentration pathway - radiative forcing stabilises at ~4.5 W/m-2 after 2100
Project	CMIP5
Simulation Authors	M. Sanderson, J. Hughes, C. Jones

Supporting Experiment

Responsible Parties

Conformance information

Document Information

Model: HadGEM2-ES

Gridspecs

CMIP5 Interim Metadata



MODEL ID, VINTAGE	INSTITUTION	ATMOSPHERE TOP RESOLUTION REFERENCES	OCEAN RESOLUTION Z COORD., TOP BC REFERENCES	SEAICE	COUPLING	LAND
ACCESS1.0 ,2011	Centre for Australian Weather and Climate Research			rheology: EVP		
CanESM2 ,2010	Canadian Centre for Climate Modelling and Analysis	top = 0.5 hPa T63L35	256 X 192 depth ,other			
BCC_CSM1.1 ,2011	Beijing Climate Center, China Meteorological Administration	top = 2.917hPa T42 T42L26	1° with enhanced resolution in the meridional direction in the tropics (1/3° meridional resolution at the equator) Z-coordinate ,linear split-explicit	rheology: EVP		
CMCC-CESM ,2009	Centro Euro-Mediterraneo per I Cambiamenti Climatici		2° zonal resolution, meridional resolution varying from 0.5° at the equator to 2° cos / south of 20°S depth ,linear implicit Madec et al. (1998);	rheology: visco-plastic Fichefet (1997); Fichefet (1999); Timmermann et al (2005);		
CNRM-CM5 ,2010	Centre National de Recherches Meteorologiques - Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique.	top = 0 hPa none tl127r ARPEGE-Climat_V5;		rheology: EVP salas_melia_2002;		Masson et al. 2003; surfex_doc 2009;



MOLES – V3.4

Metadata Objects for Linking Environmental Sciences

Key components of MOLES v3.4 include:

- **Project** descriptions
- The **observation** event itself (ISO 19156), and
- The **processes** used to acquire or generate the observation (ISO 19156, 19115 and 19115-2)

ISO 19156

An **observation** is an act that results in the estimation of the value of a feature property, and involves the application of a specified **procedure**, such as a sensor, instrument, algorithm or process chain.

An **observation** is an act associated with a discrete time instant or period through which a number, term or other symbol is assigned to a phenomenon .

The **result** of an observation is an estimate of the value of a property of some feature, so the details of the observation are metadata concerning the value of the feature property.



UK GHG Platform

<http://www.ghgplatform.org.uk>

Agricultural
UK GHG PLATFORM

InveN₂Ory
& ResearCH₄
Synthesis

UK Agricultural
Emission of Green
House Gases

- N₂O
- CH₄



- Provenance / Pedigree
- Process Chain
- How are estimates of GHG emissions calculated?



Summary

Metadata Models should make use of existing standards

There is a balance between developing something so specific that it can't be used anywhere else and something that is so generic it isn't useful

If information models are to be fit for purpose then it very much helps if they are developed alongside an application environment.

Community Engagement and Iteration are Essential

Comprehensive metadata helps others to use your data