is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

Institut Pierre Simon Laplace

ISENES Semantic Landscape

Mark A. Greenslade, Sebastien Denvil

Institut Pierre Simon Laplace (IPSL)

es-doc
Earth System Documentation

Agence Nationale de la Recherche
ANR - CONVERGENCE
MODÈLES NUMÉRIQUES

Context

Context - ISENES

Infrastructure for the European Network for Earth System Modelling

Experiments → Models → Simulations → Datasets

Full data life cycle - very long term archives

Strong links with global partners

EU Funded – 3rd funding cycle imminent

Context - Requirements

Experiments –> Models → Simulations → Datasets

What are the datasets for a certain ensemble axis ?
What set of experiments are common to projects A & B ?
What is the estimated output (in Pb) experiment X ?
What are the boundary forcing requirements for experiment Y ?
What is a model's ocean advection schema ?
How is the model parametrizing sea-ice radiation ?
Upon which computational platform was the simulation ran ?

Activites

Activities

Data Request

Dataset Publication

Citation Service

Dataset Errata

Documentation

Activities - Data Request

Metadata describing relationships between model inter-comparison projects, experiments & requested variables

Useful for deriving data output size (in Terabytes)

Essential for normalizing output variables across models

Activities - Citation Service

Centralized citation management application

Links datasets to citations via persistent identifiers

Activities - Dataset Publication

Earth System Grid Federation (ESG-F)

Globally distributed P2P dataset publication system

Hosts model & observational datasets

Extensive community & tooling eco-system

Node types = Index | Identity | Compute

Activities - Dataset Errata

Documentation of dataset errata

Typically published post-simulation

Multi project support

Integrates with PID handle service using B2Handle

Activities - Documentation

Detailed documentation of projects, experiments, models, ensembles, & platforms

Ontology + Specializations + Vocabularies

Extensive web infrastructure & tooling eco-system

Well run project with good access to senior scientists

Implementations

Implementations

Data Request = XML & Python

Citation Service = JSON, PostgreSQL, PID, REST, HTML, Python

Dataset Publication = JSON, SOLR, PostgreSQL, REST, HTML, Python

Dataset Errata = JSON, PostgreSQL, REST, HTML, Python, Javascript

ES-DOC = JSON, Python, Javascript, PostgreSQL, REST, meta-programming

Takeaways

Takeaways

Fluid narratives versus static ontologies

High Bus Factor

Institutional Fragmentation

Algorithm design != semantic web

Software engineering as a discipline in its own right

Simplicity is undervalued

Are you using or are you planning to use the existing EUDAT semantic services?

B2 services are being used within the context of dataset errata

Which potential needs for additional semantic services or functionalities of existing service would you need?

Data life cycle & provenance could be of interest

OWL & SKOS engineering might be a useful step towards unification