

Action Plan following from the 5th GO-ESSP Community Workshop (Livermore, June 2006)

Notes from Jeremy Tandy and Bryan Lawrence

Dreadfully late: October 6, 2006.

A plenary session of the meeting was devoted to constructing an action plan for the community over the next one to three years. It was agreed that the thrust of the activities in this period should revolve around providing improved access to data for the proposed AR5 activity of the IPCC.

We agreed to re-establish the GO-ESSP working group around some specific activities, and for each we would attempt to:

- Establish relevance to AR5
- Set targets (Specific, Measurable, Achievable, ideally on one and three year timescales)
- Address what if any funding is available to achieve these goals, and suggest mechanisms for international collaboration.
- Identify who would do what.

In practice, no attempt to establish time-limited targets occurred in the meeting due to meeting fatigue, and it was suggested that the working groups should do that as a first item of business. In addition, it was agreed that funding opportunities are likely to be limited to national initiatives, but it would be helpful to establish some specific documentation of the international scope which would be referable in national applications.

Working Groups

1. GridSpec: Aim to describe complicated grids in ways that allow multiple components of models to interact.
 - Highly relevant to AR5, as at least one model (GFDL) will have non-traditional grids. Strong correlation with CF activities.
 - Specific actions:
 - i. Add gridspec code to the LAS (Steve H)
 - ii. Publish gridspec (in UML), along with implementation and tooling (Balaji)
 - iii. Develop convention for how to represent a staggered grid (Rich S)
 - iv. Explore unambiguous identification of vectors (Rich S)
2. Authorisation: Aim to come up with a common authorisation protocol.
 - Highly relevant to AR5, given the expectation that we would need to deploy a distributed archive.
 - Specific actions:
 - i. NCAR and NDG to test interaction (Luca and Bryan)
 - ii. CDP and ESG interoperating with the the Thredds Data Server at the authorisation level (Luca)
 - iii. Handling AAA for cross-project registration? What kind of model is appropriate? What do we need to graft onto community tools (Don)
3. Services and Outreach
 - Both underlying technology services, and outreach which has relevance beyond working group one (climate modeller) requirements of AR5.
 - Specific Actions

- i. Work with the Curator project & THREDDS catalogue to configure the live access server (Steve H)
 - ii. Explore forecast collections and ensembles via the (future) TDS (Kevin)
 - iii. Move the go-essp wiki to something better (exploit the CF wiki or NDG trac? ... discussion implies that we shouldn't use the CF wiki) (Bryan L)
 - iv. go-essp members need access to the ISO documents (as a community) (Bryan L)
 - v. Develop tooling to 'summarize' our content (for education and outreach) ... In particular, formalise the 8-bit image + scale idea as one of the 'magic' output formats (Annette)
 - vi. Server-side virtual data binding to operations ... dataset-service binding & publication ... and implications for vocabularies (Heinrich).
 - Activities that we thought important but which are unlikely to be actioned as no one took responsibility (or if they did, we failed to identify them):
 - vii. Create a tool registry and cookbook online, e.g. on the go-essp website) (No one volunteered to take this on).
 - viii. Build bridges between observations and modelling communities (not sure what we meant by this?)
 - ix. Dataset tagging via diverse methods like taxonomies, ontologies, folksonomies ...
4. Metadata
- In this context we will be concentrating on issues associated with documenting differences between simulations, and on finding data at disparate locations, both very relevant for AR5.
 - Specific Actions
 - i. NMM - Curator compatible metadata. Explore ability to harvest this metadata for AR5. (Balaji, Lois)
 - ii. Develop a semantic structure for CF and NMM (Benno, working with Luis and others)
 - iii. Develop a glossary of (go-essp community) terms (Leo, Katherine and others)
 - iv. Develop a datapedia using wikipedia entries (Russ)
 - v. Exploring how to project 'local' dialects into the community space (Roy, Luis, Alison)
 - vi. Create a reliable source of information for controlled vocabularies and mappings (Roy, Luis, Michael B)
 - vii. Harvest WMO ISO 19139 records into CDP and NDG (Michael, Bryan)
 - viii. Integrate GFDL metadata into the OAI based discovery services (including contacting NDG, Sergei)
 - ix. Integrate the WDC into the OAI based discovery services (Michael)
 - x. Explore coastal and ocean models for NMM and gridspec intergration (Katherine)
 - xi. Explore vertical domain descriptions (in Grib, Katherine, Bryan, Balaji)
 - xii. Explore convergence of NMM and NumSim (Bryan and Katherine)
 - xiii. Creation of scientific vocabulary for climatic models ... and apply this to some 'other' model (i.e. European UM model) (Sergei and Lois)
 - xiv. Develop a 'hydrology' ontology (Annette)
5. Data Modelling
- A key component of all autonomic systems to manipulate data is the underlying data model. This is also crucial for understanding interoperability (as may be required for AR5).
 - Specific Actions:
 - i. While the CF-API work is being begun at Unidata, there was discussion as to whether this could be continued as an open-source development. The specific action was to consider how this could be done (Russ and Bryan)
 - ii. Progress reference Implementations of CF and stimulate convergence of data models (Bob D).
 - iii. NetCDF4: complete, and publish examples (Russ)
 - iv. Explore whether the CSML parser can process the output from TDS (or more likely, what

would be needed to make this happen, Roland)

- v. Unstick progress on CF issues (John C)
 - vi. Evolve a high-level common datatype model motivated by OGC, and carry out some knowledge transfer based on UML modelling (John C, Bryan)
 - vii. Consider a workshop (Boulder, Jan 2007, possibly later now?) to expedite this UML modelling of the datamodels & progressing to the APIs and reference implementation (Bryan L)
 - i. Explore continuing interoperability with GIS systems (and enabling 'other' client tooling like MatLAB, i.e. via the netCDF Java api, python, matlab and IDL enabled clients ... including integration of clients with data catalogues for browse). Roy, Rich, Bryan, Kyle
6. Specific Preparation for AR5
- o Main thrust of this activity will be roadmapping how these activities can be contribute to AR5 (led by Bob D)
 - o Specific Actions
 - i. Overall plan (Bob D, Michael L, Bryan L)
 - ii. Support of CF governance (Bryan, Karl)
 - iii. Issues associated with a distributed archive (Michael L)
 - iv. Address issues associated with CMOR (Karl)
 - Possible conflict with similar tools from other Model Intercomparison Projects and eventual reference implementations of CF
 - Development of CMOR including convincing all AR5 working groups of the importance of metadata standardisation.
 - v. Develop a white paper on the importance of this activity for AR5 (Karl, Bryan, Micheal)