

ENES Data Task Force Meeting

Hamburg, 29th April, 2015

(Draft Minutes, BNL, 4th April, 2015)



DTF Members Attending: Sebastien Denvil, Sylvie Jousame (by phone for an hour), Martin Juckes, Stephan Kindermann, Bryan Lawrence, Frank Toussaint

(It is not clear what other DTF members were invited to take part by phone and were unavailable. Future meetings will be more formal.)

In attendance: Reinhard Budich (some of the time), David Hassell, Ag Stephens, Martina Stockhause, Nicolas Carenton, Michael Lautenschlager, Alan Iwi, Mark Greenslade, Guillaume Levavasseur, Tobias Weigel, Philip Kershaw.

This meeting took place within a two-day workshop held at DKRZ with attendance of core data management teams from DKRZ, IPSL and CEDA (BADG). The aim of that meeting was to look at synergies between the institutional activities in support of CMIP6, ESGF, ENES and more generally. In that context there was an opportunity to discuss DTF matters in advance of a future formal meeting.

General Discussion (without Sylvie and Reinhard)

(The main context was to consider how national funding from one country could leverage funding from other countries - how could a European data infrastructure be strengthened and exploit an ESFR?!)

1. Agreed that there should be a distinction (also in redmine) between the ENES task force and the IS-ENES2 activities (though there is considerable overlap, not least because IS-ENES2 will have much of the effort required to deliver DTF activities).
2. On the relationship between national funding and European funding: Things developed with national funding consist of “national capability” and “development projects”. Where the development projects become embedded in European infrastructure, this is seen as a “win” for national funding. European funding could be used for more experimental things rather than, as now, to deliver integral activities. One would need a roadmap to join national and international activities.
3. How can we exploit national services internationally?
4. To make progress one would need to know what the existing infrastructure embedded nationally looks like (going beyond ESGF, of which we have good visibility).
 - a. But no-one knows since our internal infrastructure is mostly invisible!
 - b. Could we agree a [RM-ODP](#)-like approach? Break things up into
 - i. Enterprise Viewpoint - This is why we do things, institutional context, the workflows we want to support.
 - ii. Information Viewpoint - What are the “information artifacts” that exist to deliver our activity?
 - iii. Computational Viewpoint - This is the software artifacts we use
 - iv. Engineering Viewpoint - The hardware we have deployed.
 - v. Deployment Viewpoint - includes a description of the computational services and APIs and how these things are joined together in support of the required workflows.

- c. Do so with two components: Status, and near term plans (whatever near term means)?
- d. Could we then synchronise our roadmap “reporting” (as a first step, this is not the same as synchronising the content of the roadmaps :-).
 - i. this could be a pre-requisite to the “communications papers discussed below”
 - ii. would broaden and harmonise to take on the wider European perspective ... relate to the foresight.
- e. **Action: Three centres to prepare this information in presentation format, for a small workshop, winter 2015 ... this could become a delayed version of MS114 (or at least the output could be). We would aim at the workshop to consolidate the language used for the roadmaps so we all understand what each other is trying to achieve.**
 - i. **SubAction: Get a time into the diaries, decide on a venue, and a pre-meeting documentation deadline.**
- f. **Action: We need to deliver the IS-ENES2 deliverable D5.1 asap, this would cover much of the general requirement, and could be updated further.**

5. What is the definition of “an ENES ESFRI”?

- a. It needs to be broader than CMIP6 (or indeed the CMIP process)?
- b. Clearly goes beyond ESGF as currently incarnated.
 - i. Need to allow, as a thought experiment, of joining things up at a European level in a much tighter way than would be done at a global scale, e.g what if we joined the EUDAT consortium, or had even closer integration using [irods](#) or [ipfs](#), etc.
 - ii. Would involve two levels of dependency:
 - 1. long term funding (bound by ESFRI)
 - 2. low level functionality (e.g. CF support)
- c. How would it interact with the HPC activity?
 - i. Climate model data production (configure and run) (connected to seamless workflow)
 - ii. It would not be all of “HPC-land” for climate (e.g. CoE: ESIWACE) and scientific development and process evaluation per se would vbe out of scope)

6. We need a “communications paper”/“policy brief” (or both) ... outlining our joint European activities thus far and roadmap (how is this different from the proposed foresight activity?)

- a. Could we aim for GMD rather than BAMS for the detailed view?
- b. Could we embed a detailed data foresight in that upcoming review?
- c. We haven’t expressed anything of the scale of what we do in venues visible to our scientific community?
 - i. There are modelling experts that interact with ESGF, those that don’t interact with ESGF (PI types), and downstream users (including policy) ... we probably need different messages for each.
 - ii. **Action: Bryan/Michael/Sebastien ... to consider how to take this forward in the data task force ...**

Cooperation with related activities (EUDAT, ...)

(The DTF is expected to work with, and represent ENES with other bodies.)

7. (There was not time to discuss the RDA Paris meeting on for research data for climate change.)

8. We had a presentation from Hannes Thiemann covering EUDAT and the b2service suite. DKRZ are currently involved in EUDAT representing the ENES community. CEDA have a very small engagement planned under the auspices of their parent body (STFC) who are also EUDAT partners. Those present remarked:
 - a. It is not yet clear where EUDAT is positioned in the European landscape, it needs formal relationships with projects like ENES (and IS-ENES2).
 - b. B2SAFE: It was thought that
 - i. It is not operating for, or targeting, the bigger institutions and nations, but the ones which do not have the same institutional infrastructure support we have; for the long-tail (although it is not obvious what part of the ENES long tail). Again, there needs to be a clear relationship (which might include an overlap) between EUDAT and any ENES infrastructure.
 - ii. It is not obvious that EUDAT has positioned itself well in terms of the scale of it's offerings. It needs to be positioned between the likes of Dropbox/Box (and European industrial equivalents) and the large discipline academic activities. The boundary between these will move with time, and EUDAT and ENES will have to move with it.
 - c. B2STAGE: We already do a lot of migration between remote sites and PRACE and other HPC sites. Could we expand this?
 - i. **Action: CEDA to look into this in the context of JASMIN**
 - d. B2SHARE: Clearly aiming at the long tail data. DKRZ will bring up a b2share for ENES users. Will come up with robust implementation plan in autumn 2015.

Futures of IS-ENES and ENES programme

(with Sylvie and Reinhard)

9. Most probably a break between IS-ENES2 and a possible IS-ENES3.
 - a. "IS-ENES3" is on the work programme for 2018 (which is not a guarantee) of anything.
 - b. Will need to fill up with national funding, hence the urgency to begin inter-national more integrated collaboration.
10. Need the foresight case to provide the science case ready for "before autumn" 2016 ...
 - a. Update the existing foresight, but update with another layer of actionable activities around data and HPC. Actionable plans will depend on the architecture and roadmap ideas we have coming out of the planned workshop above.
 - i. Really need to push the metrics of use outside the funded consortium.
 1. compare/compete numbers with EUDAT...
 - ii. Competition of resources: climeri/cmip6/future of is-enes3/communications plans. Currently, "everyone does everything".
 - iii. Expecting ESIWACE decision early June ... (if successful two months to do negotiations to start around October).
11. Copernicus
 - a. Important to recognise that ENES is engaged both downstream and upstream.
 - b. Expecting a climate data store call in Jun/Jul ...
 - i. What does "brokering with ESGF mean?" ... how does ESGF provide a service level (and how do European ESGF providers sustain that in the context of European funding).

- ii. What experiments will be involved? Will there be emphasis on, or priority around, data that has already been comprehensively evaluated (e.g. CMIP5/CORDEX)?
- iii. Dedicated experiments for Copernicus? An “ECMIP”?
- iv. The ENES consortium should coordinate a bid to deliver access to the global climate data projections; but how do we organise the governance for a few ENES sites to represent the community ... with o(1.5)M overall to one contractor to work with a small consortium of subcontractors. Risk management issues with funding tied to delivery and significant dependence on ESGF.
- v. Traceability/provenance interface to the sectorial interface systems? Why would we use the ESGF interface for this? Local data -> QC -> new interface ... What is the winnable argument that says “the ENES” ESGF interface is the right interface?
- vi. Clearly ENES is the right community to provide scientific guidance
- vii. ***Data Task Force Resolution: A compact set of partners is a necessary condition for delivering this proposed activity. It would not be good for Europe if we salami sliced this funding across groups. We need to make sure the upstream modelling groups recognise their benefit would not be direct funding *from this opportunity* for running their ESGF interfaces per se.***
 - What about the climate impact portal in this context? Several calls anticipated ... global/regional ... there will be a number of calls. These should be small numbers of groups, but not necessarily the same groups.
 - **Action: Sylvie to request Francesca to produce with a timeline of calls; and ask the ENES board to deal with a political decision about how to deal with avoiding political salami slicing. Sebastien to add some paragraphs to give context for each of the known calls.**
 - **Action: the board meeting in June ... following a brain-storming telco. DTF task force telco in late May to follow this discussion.**