



Breakout Session on Data and Metadata

Stephan Kindermann (DKRZ)

Sebastien Denvil (IPSL)

Sandro Fiore (CMCC)

and the ENES Data Task Force



Topics

1. Coordination and planning on outreach and user requirements, user engagement

- initial outreach activities and workshop/coding sprints planning
- Per service we have to identify target audiences, communities and stakeholders (by competence, by knowledge, with a minimum ability to use the e-infra)
- Help define CDI climate data infrastructure and sustainability
- Documentation is a key element for user engagement
- KPIs (something on documentation, usage,)



From each tools (ESMvaltool, ESGF, ESDOC, Compute)

- Expose what we think we are doing, and will deliver (short term). Today's roadmap.
- User feedback on those
- Revised Requirements identified wrt feedback and available publicly, Requirements gap analysis. In two phases.
- OGC, GEOSS and EOSC and space agencies needs to talk together before (not only for compute) they are trying to do the same thing (not only)
- Copernicus & others requirements database
- What are the coordination requirements wrt user requirements?
- OGC, EOSS and EOSC and space agencies needs to talk together before (not only for compute) they are trying to do the same thing



Topics

2. Overall context/landscape for the “data” and “metadata” component of IS-ENES

(how it (i) relates to, (ii) can benefit from, (iii) can support, (iv) positions itself with, etc. other initiatives)

- EOSC, Copernicus, ESGF, RDA, EUDAT-CDI, FAIR, ENVRI+, Citation, Other initiatives?
- Repository, archive and curation functions and associated services have to be clear.
- Already exercised within ESGF (PID and Handle service and data citation)
- Which certification (repository, FAIR others...) should we invest in obtaining?
- When we are operating in external how do we work on representation of ENES → ENES Data Task Force? Or WP leaders.



Topics

2. Overall context/landscape for the “data” and “metadata” component of IS-ENES
(how it (i) relates to, (ii) can benefit from, (iii) can support, (iv) positions itself with, etc. other initiatives)

- Very good if we know what we expect from European initiatives like EOSC, EUDAT CDI that will feed ENES CDI.
- We need to analyze the service portfolio roadmap of EOSC and identify what we need from them.
- The only way to influence those is to engage with project.
- Being able to acknowledge or cite software, or infrastructure in addition to data would be great. DOIed paper describing ENES CDI could be useful for this.



Topics

3. Compute service roadmap and strategy

- Evaluate and discuss M12, adopt/prototype/test M24 and adapt.
 - Invite workshop on those topics before M12
 - Xarray + zarr + dask, Object store, post POSIX activities in ESIWACE 1 & 2
- ESGF Compute Challenge (CMCC, DKRZ, CERFACS, IPSL)
 - Averaging, interpolation, subsetting
 - Resource limitation and access management to be looked at closely.
 - Which backend to use? We need a mapping between the API and the tool.
 - We will learn a lot from that challenge.
 - How do we advertise compute capabilities.



Topics

3. Compute service roadmap and strategy

- Interaction with other initiatives outside IS-ENES3 (EOSC, Copernicus)
- Interaction and visibility within RDA, OGC
 - Idea to have an interest group for CMIP analysis
(infra+software+provenance+take the best of services that are data citation, errata, ...)
 - RDA have a topic on datacubes
 - Climate and weather and air quality interest group in RDA
 - OGC testbed 15 challenges that reference ESGF.
- VA/TNA support activity, software stack



Topics

4. VA and TNA

- TNA procedure and ideas for external service review board as well as selection panel members, evaluation criteria
 - Ready by M12
 - Schedule a meeting in the next 3 months.
 - What users expect from TNA? Find out within proposal.
 - What their feedback will be. Capture that.
- Access statistics and KPI collection: initial brainstorming on organization
 - Review what we have now. Summarize, collect and decide on devmt.



Topics

5. Support activities

- short summary of status in ESGF, C4I, ES-DOC, vocabs (incl. climate indices), eval areas, impact community support together with WP3/NA2 and WP5/NA4
 - Support infrastructure installation process.
 - ESGF
 - Compute service
- short discussion how to harmonize ENES-CDI support activities and next steps
 - Very important. Time were missing to cover. Plan discussions and meeting.