Documenting climate models
ES-DOC for CMIP6

Eric Guilyardi for the ES-DOC team

IS-ENES2 - January 2017

Community expectations and use cases

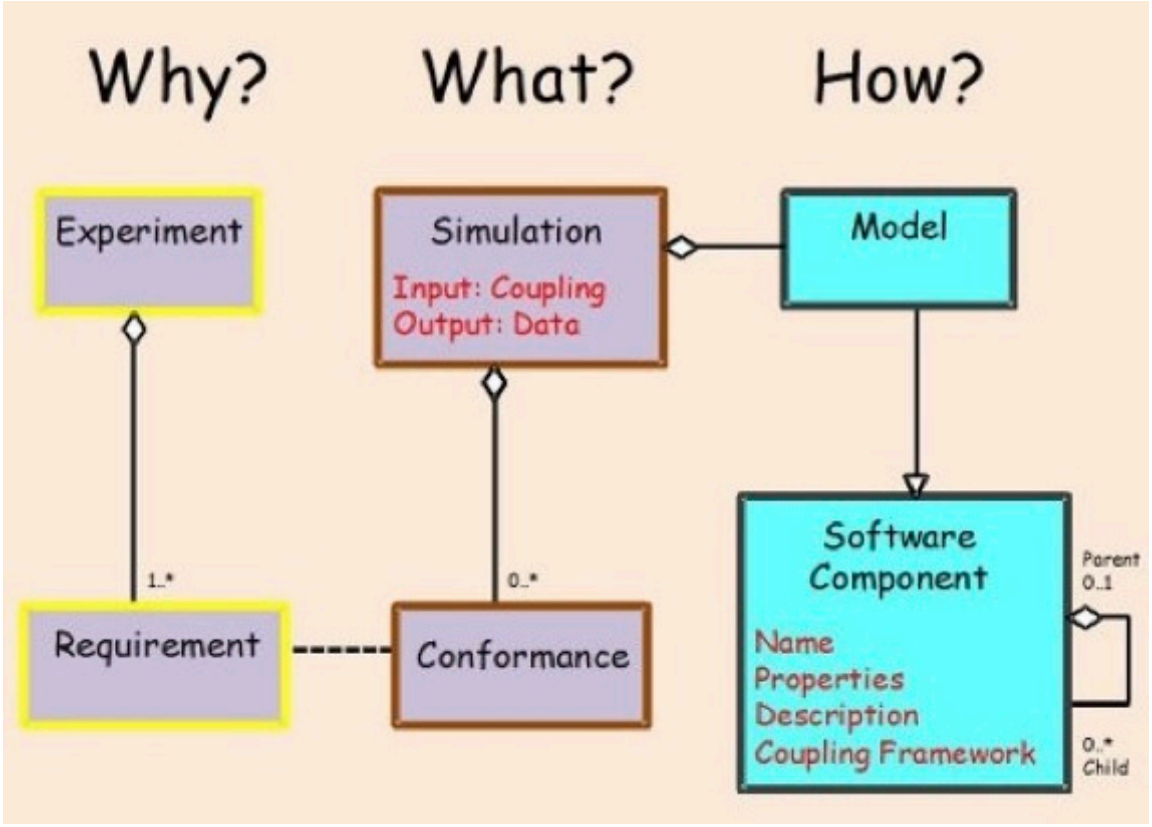
- Document CMIP models, simulations and experiments for wide array of end users
- WGCM specific requirements (e.g. tables)
- Evolution from CMIP5 – learn from own experience and data providers and users, improve, simplify, make robust, beta testing
- Use driven, not technology driven
- Document howtos and steps for all (workflows)
- WIP oversees integration of ES-DOC ecosystems in CMIP activities
- CMIP-driven but not restricted to CMIP

ES-DOC and CMIP5 metadata

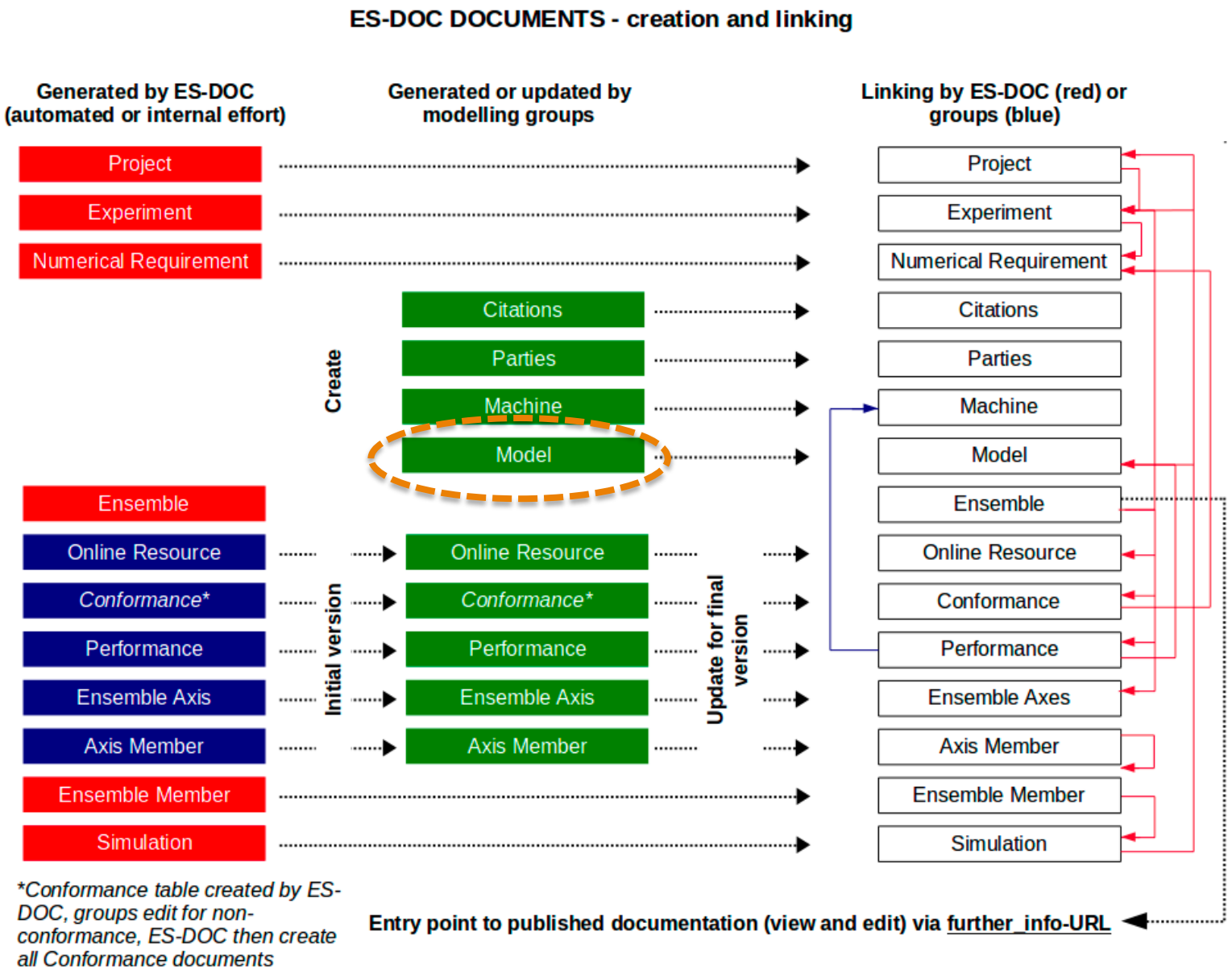
- What we had at the end of CMIP5:
  - The CIM: strategic community effort, several projects leveraging from it
  - A set of controlled vocabularies: specializations for particular user communities
  - CMIP5 documentation: 42 models and 600 simulations described in CIM database – a unique ressource – contributed to AR5 (model table)…
  - CIM tools: viewer and comparator (web service)
  - Governance: ES-DOC. EU-US close collaboration, management and development process
  - Community « buy in »

es-doc
Earth System Documentation

Capturing the concepts
The Common Information Model - CIM



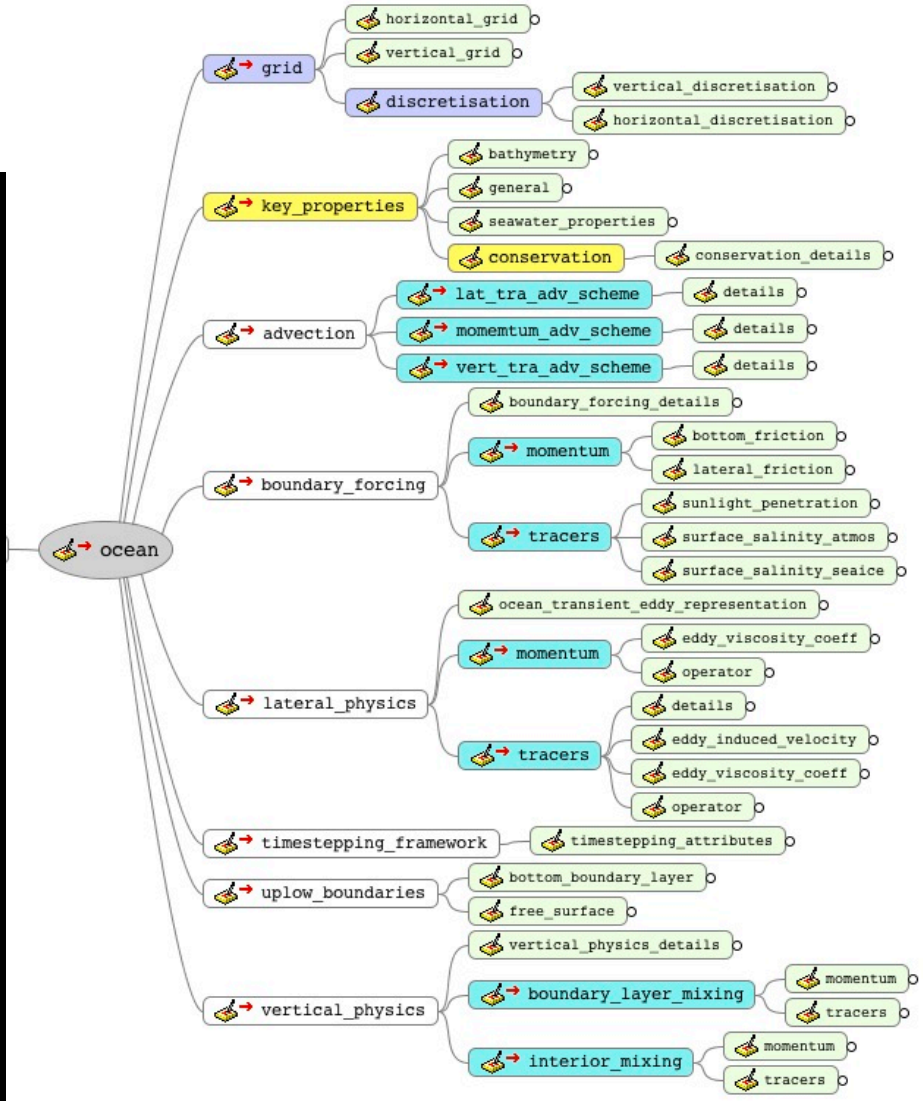
CMIP6 activity: who does what ?



CMIP6 model documentation

Update of CMIP5 vocabularies contents
Specialisation of CIM 2.0
New formalism (python) and tools
Separation of concerns, GitHub

```
# ----------------------------------------------------------------------
# PROCESS: DESCRIPTION
#
# Scientific context of the process
# ----------------------------------------------------------------------
DESCRIPTION = 'Properties of ocean advection'

# ----------------------------------------------------------------------
# PROCESS: DETAILS
#
# Sets of details for the process
# ----------------------------------------------------------------------
DETAILS = OrderedDict()

# ----------------------------------------------------------------------
# PROCESS: SUB PROCESSES
#
# Sets of discrete portions of the process
# ----------------------------------------------------------------------
SUB_PROCESSES = OrderedDict()

SUB_PROCESSES['momemtum_adv_scheme'] = {
    'description': 'Properties of lateral momemtum advection scheme in ocean',
    'details': ['details']
}

SUB_PROCESSES['lat_tra_adv_scheme'] = {
    'description': 'Properties of lateral tracer advection scheme in ocean',
    'details': ['details']
}

SUB_PROCESSES['vert_tra_adv_scheme'] = {
    'description': 'Properties of vertical momemtum advection scheme in ocean',
    'details': ['details']
}


# ----------------------------------------------------------------------
# PROCESS: SUB PROCESSES: DETAILS
#
# Sets of details for the sub processes
# ----------------------------------------------------------------------
SUB_PROCESS_DETAILS = OrderedDict()

SUB_PROCESS_DETAILS['momemtum_adv_scheme:details'] = {
    'description': 'Properties of lateral momemtum advection scheme in ocean',
    'properties': [
        ('type', 'ENUM:adv_mom_scheme_types', '1.1',
```

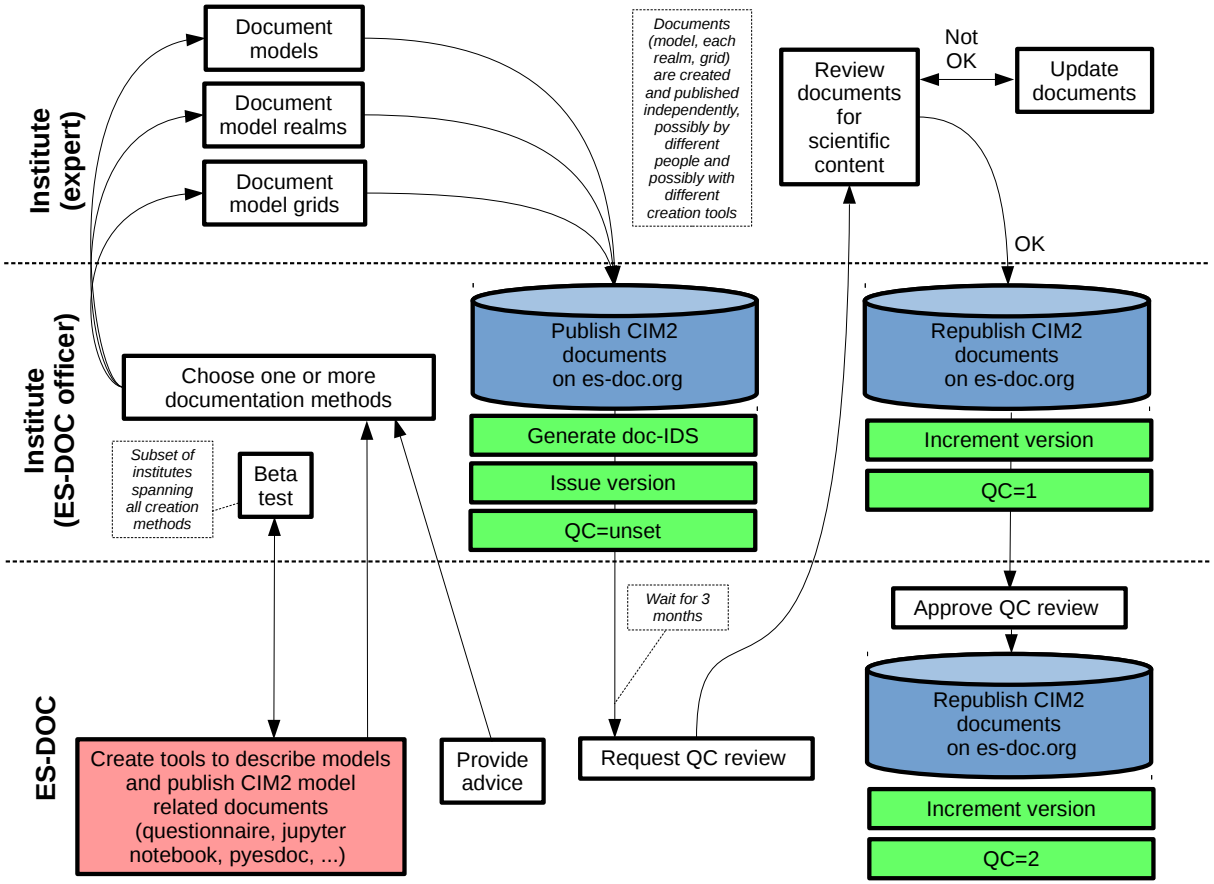


CMIP6 model documentation

Different tools to create documents (py-esdoc, questionnaire, iPython notebooks, spreadsheets)

Each group to experiment and choose

Start from CMIP5 model documentation translated into CIM2



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

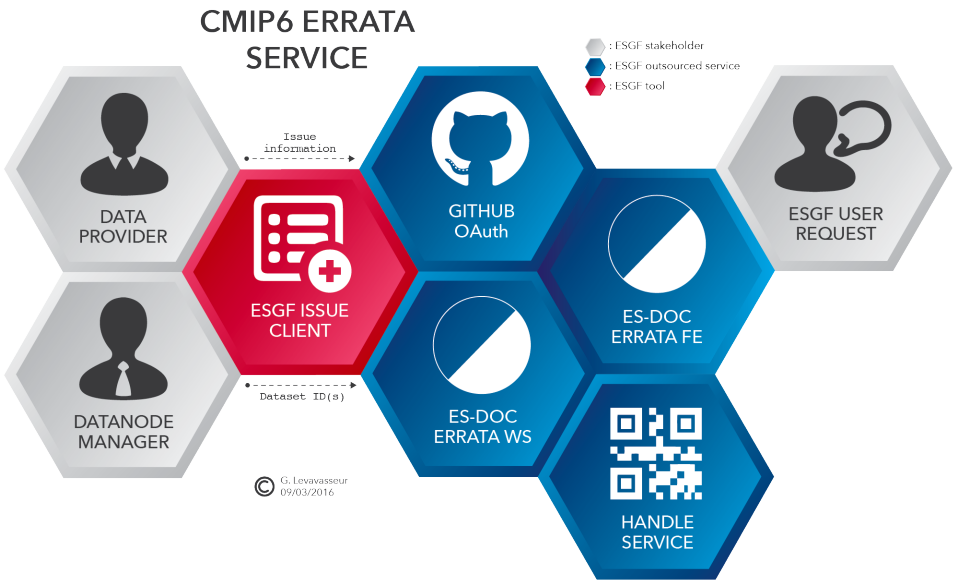
es-doc
Earth System Documentation

ES-DOC Errata Service

es-doc
Earth System Documentation

Keeping track of changes in CMIP6 data files

- Ensures archive quality by providing issue status tracking.
- Relies on the upcoming PID feature for retracing past and future versions of datasets and/or files.
- Divided into two major pieces: Remote server and associated web-services and local client.
- Supports duplicate issue related meta-data storage in local files and in remote server database.



- Full community release: March 2017

WCRP World Climate Research Programme
is-enes INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

ES-DOC for CMIP6 status

- **The collecting process will be different than for CMIP5:**
  - Large fraction of documents (Exp., Simul., Ens.,...) automated (following ESGF publishing)
  - The others (model, conformance to protocol, forcings, responsible party,...) produced by groups when ready – joined together via the "futher_info_URL" attribute
  - Multiple tools to create these documents (python library or notebooks, questionnaire,…).
- **Ready for community review (Jan 2017):**
  - Documentation workflow for CMIP6 and type of information to be collected
  - WIP white paper describing overvall processs (modelling groups to review)
    - Ready this week, also IS-ENES2 D11.4
  - Ocean and sea-ice realms (including short table for papers)
- **Currently in internal review:**
  - Documents creation tools, both automated and UIs (py-esdoc, questionnaire, cdf2cim,…)
  - Top level, atmosphere, atmosphere chemistry, ocean bcg, land surface, realms
- **Working on:**
  - Connecting the pieces together (coding sprint next week) – March 2017
  - Update science contents of other realms (with the community) – Feb 2017
  - Beta testing

ES-DOC for CMIP6 status

- Project to document CMIP6 well underway
- Building on CMIP5 experience (both good and bad !)
- Clear set of uses cases
- Community review formalised (internal, WIP/WGCM, wider)
- Designed so that process much easier for groups:
  - Large fraction is automated and/or provided by ES-DOC
  - Starting model description from CMIP5 version
  - Beta testing of 5 months (Oct 2016 – March 2017, UKMO, GFDL, IPSL)
    - Additional beta testing by CCCMA + 2 other groups (Asia)
  - Documentation for all steps (+ overview as WIP white paper)
- Operational phase 1 release: April 2017
- Connects to errata, data citation,… services
- Looking ahead (post CMIP6) to include other « realms »:
  - Regional models, downscaling
  - Evaluation & metrics, obs4MIP

Looking ahead: strengths and challenges

- Strong expertise fostered by PRISM/METAFOR/IS-ENES support and legacy – common understanding of issues
- International coordination, overseen by WIP
- Number of operational tools – maturing ecosystem, CIM2
- CMIP6 documentation workflow well defined and organised around modelling groups perspective

- Long term sustainability
  - How to continue funding this joint effort
  - Hiring of people comfortable with the full range of skills
  - Coordination and project management
  - Standards, codes, scientific realms governance (separation of concerns)
  - Leveraging from CLIMERI, Copernicus,… ?