

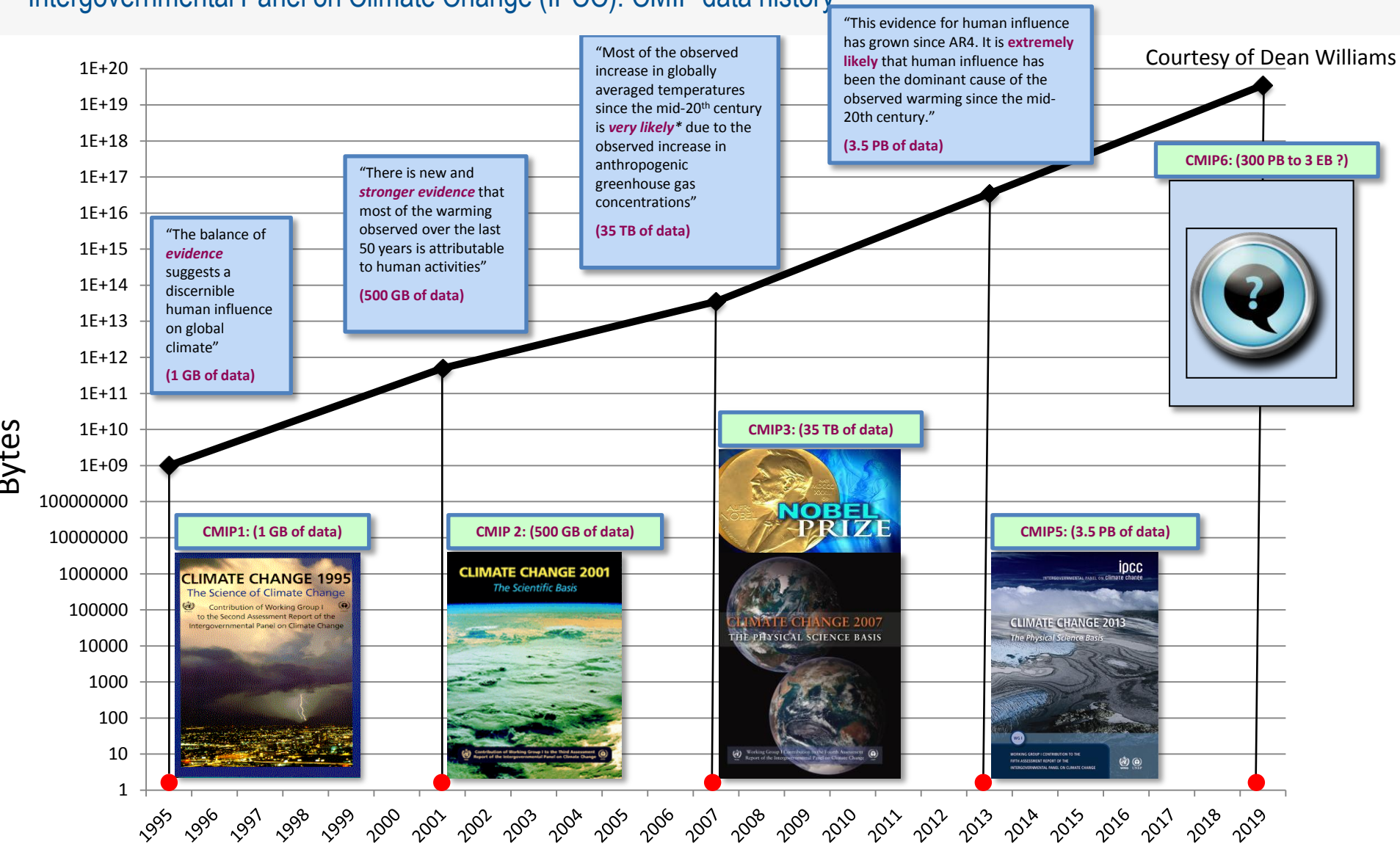
Digital Object Management for ENES: Challenges and Opportunities

GEDE workshop,
Brussels, 2018-09-26

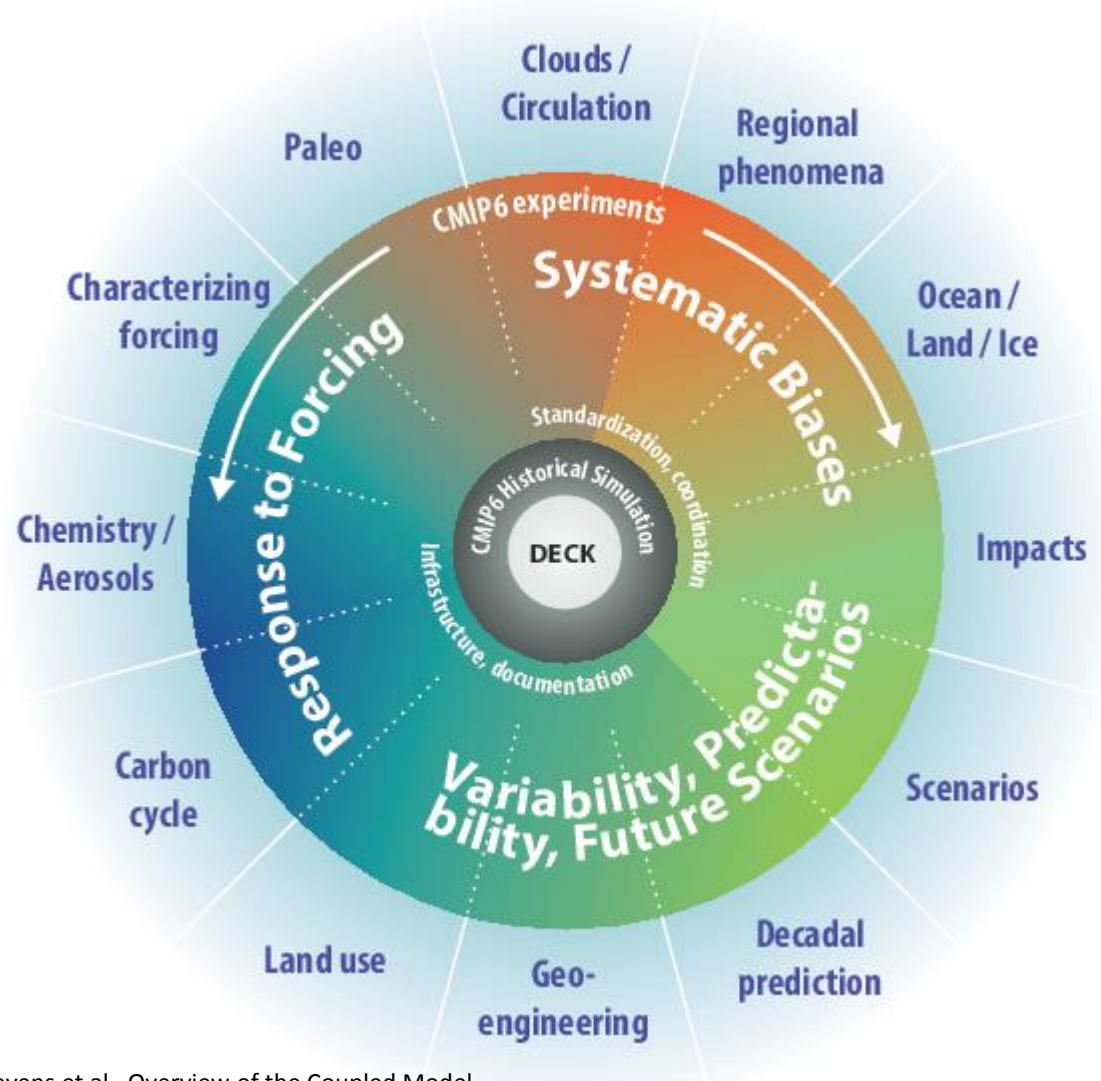
Tobias Weigel
Deutsches Klimarechenzentrum (DKRZ)

Scientific Driver: International Climate Model Intercomparison Projects

Intergovernmental Panel on Climate Change (IPCC): CMIP data history



CMIP6 experiment design

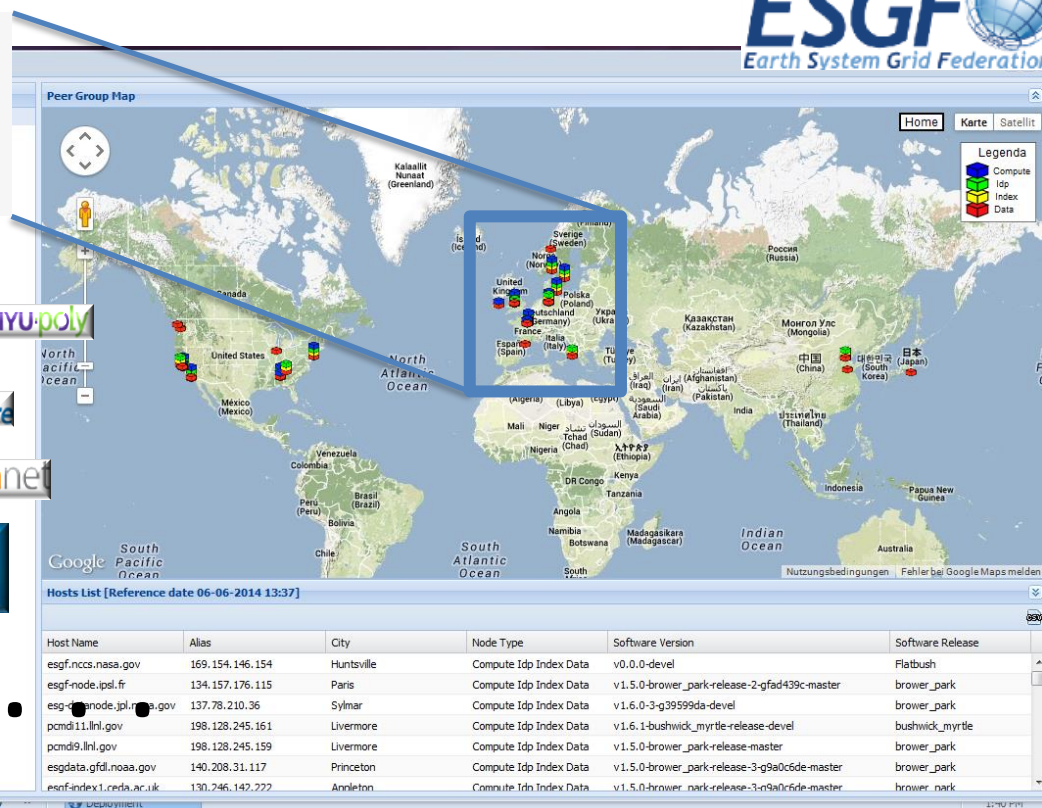


Eyring, Bony, Meehl, Senior, Stevens et al., Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geosci. Model Dev., EGU, 2016. doi:10.5194/gmd-9-1937-2016

The Earth System Grid Federation (ESGF)

ESGF is a coordinated multiagency, international collaboration of institutions that continually develop, deploy, and maintain software needed to facilitate and empower the study of climate

IS-ENES: European ESGF federation part



Courtesy of Dean Williams

Challenges and opportunities

1. Automated digital object management
2. Workflow support and provenance aggregation
3. Support for work at higher levels of abstraction
4. Services to new user communities
5. Sustainable funding and business models

ESGF publication and versioning

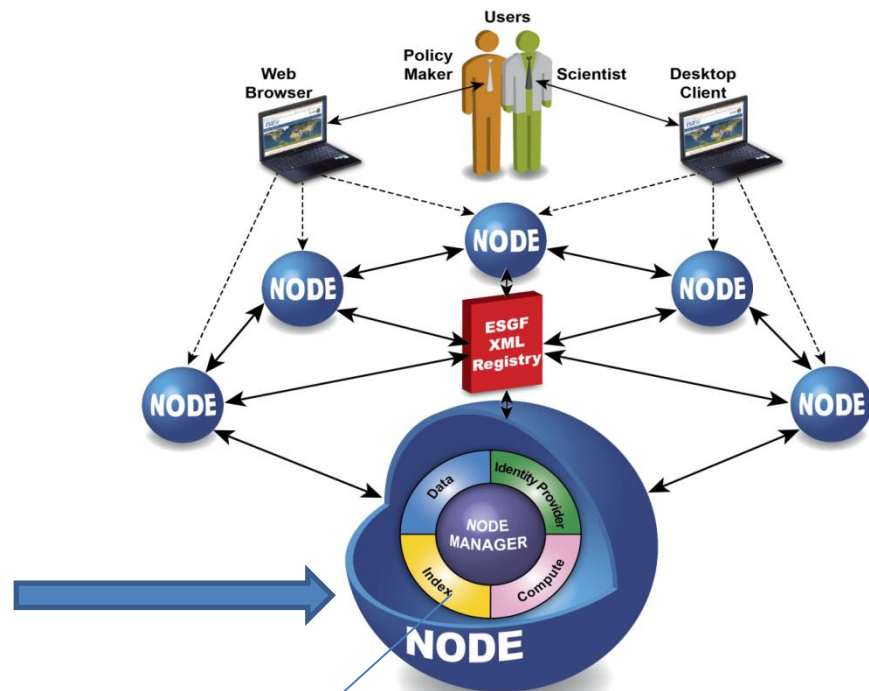
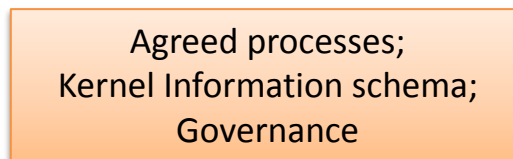
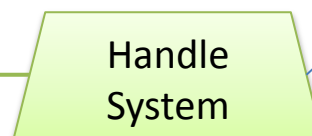
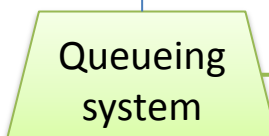
raw data (model data, obs data)



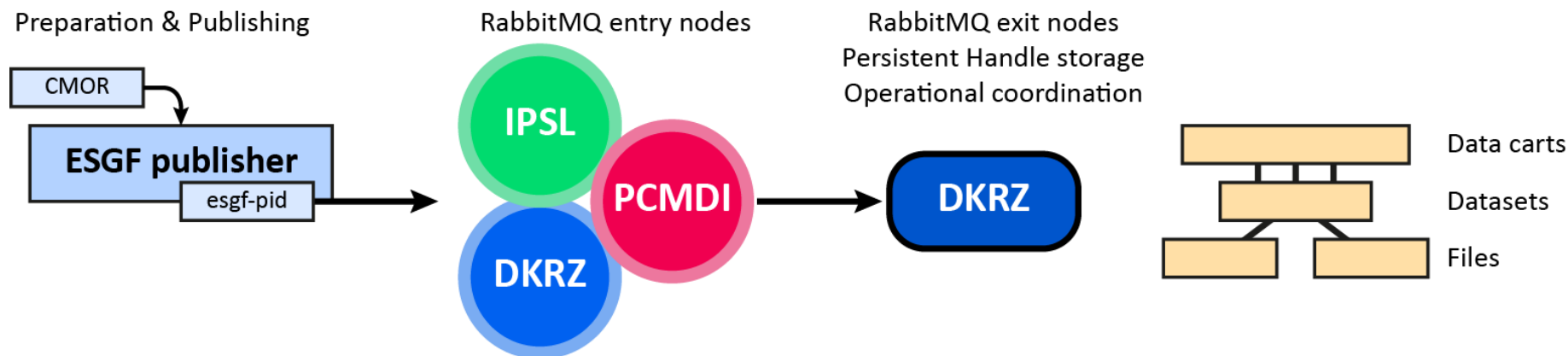
„ESGF publishable“ files



Iterate on new versions



ESGF PID services

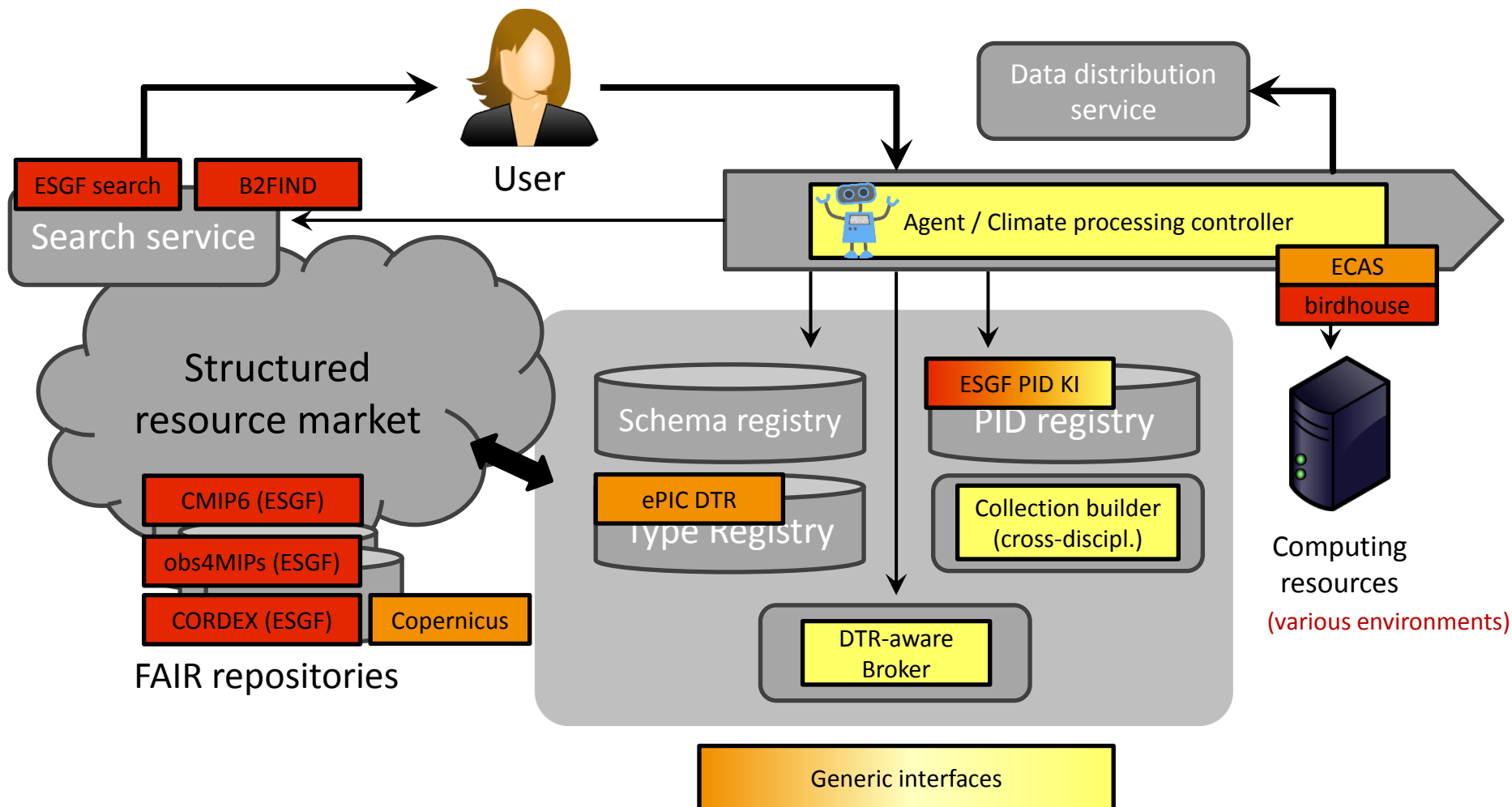


- Scalability, reliability, governance
- Future option: Replication support
 - package – replicate – verify
- Will require clear interfaces such as the DOIP

Automated DO management & Workflow support

- Example: Replication support
- Example: HPC workflow support
 - Models should be able to record who they are and what they did
- Example: Workflow brokering, matching, data transformations
 - We discussed this in the frame of T-TAP in the past

Type-Triggered Automated Processing (T-TAP): Status for climate data



red: operational / ready

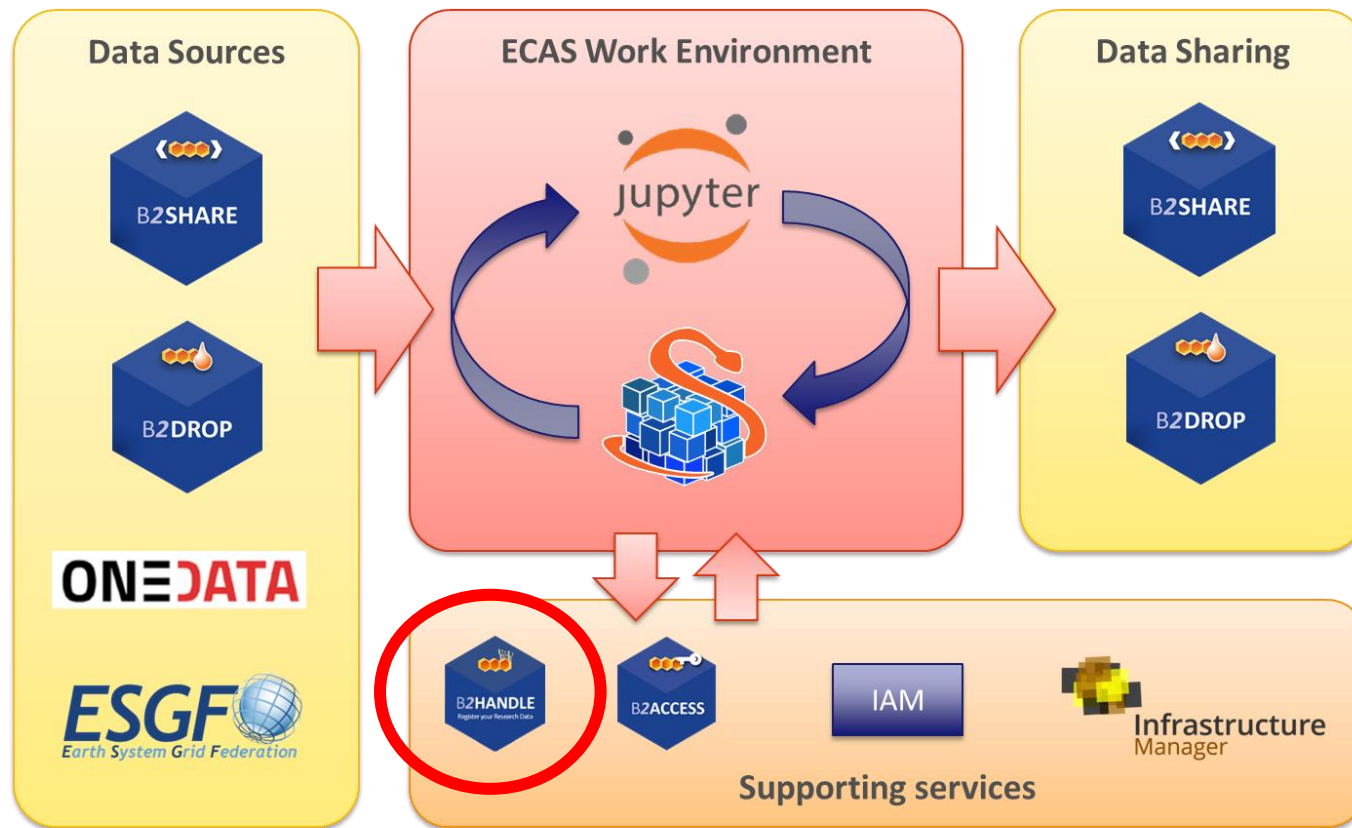
orange: under construction (e.g. via confirmed projects), but likely to become operational

yellow: more work to be done

Support for work at higher levels of abstraction

- DOs as primary citizen in ENES
- But: Abstraction not limited to DO concept
- Users should concentrate on analysis problems, not data wrangling
- Example: Data I/O layer for Jupyter environments
- Example: Machine Learning support VRE

Bridging one gap: Processing services (ECAS)



- Opportunity to put Kernel Information in place
- Envisaged development for mid 2019

Support for new user communities

- Knowledge of limitations and assumptions not obvious to non-ENES users
 - social sciences, public administration, policy making
- DO angle: Abstraction & Research Object approach

Thank you for your attention!

weigel@dkrz.de