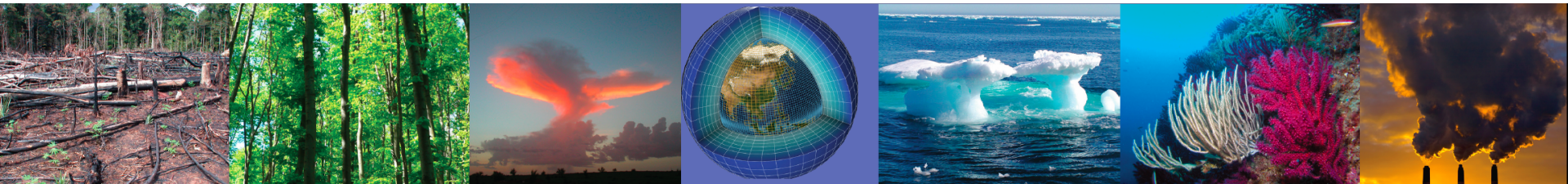
ESGF F2F
December 6th, 2016
Washington, USA

European Network for Earth System Modelling Data Infrastructure

Sylvie Joussaume
CNRS, IPSL
Coordinator IS-ENES2, Infrastructure for ENES



ENES

European Network for Earth System modelling

http://enes.org/

http://is.enes.org/

A network of European groups in climate/Earth system modelling
Launched in 2001 (MOU)

Ca 50 groups from academic, public and industrial world

Main focus :
discuss strategy to accelerate progress in climate/ Earth system modelling and understanding

Several EU projects

ENSEMBLES, COMBINE, EUCLIPSE, EMBRACE, SPECS, PRIMAVERA, CRESCENDO *PRISM, METAFOR,* IS-ENES (*1*& 2), ESIWACE

IS-ENES
Infrastructure for ENES

FP7 European projects
IS-ENES 2009-2013
IS-ENES2 2013-2017

Infrastructure
Models & their environment
Model data (ESGF-ESDOC)
Interface with HPC ecosystem

Users :
Climate modelling community (Global & regional)
Impact studies

IS-ENES : Infrastructure for ENES

FP7 project « Integrating Activities »

http://is.enes.org/

1rst phase: March 2009- Feb 2013 (7.6 M€), 18 partners
2nd phase: Apr 2013- March 2017 (8 M€), 23 partners

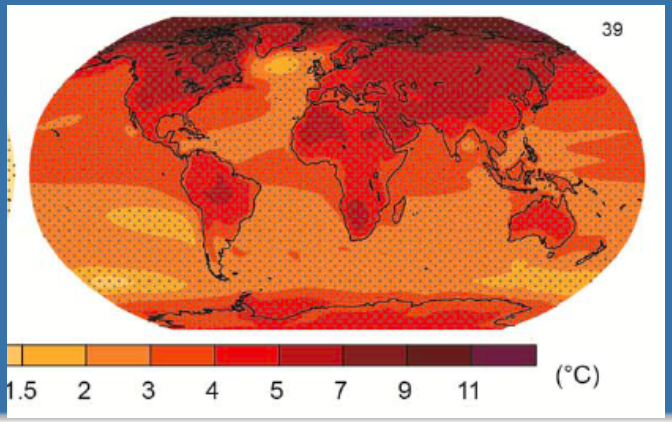
Better understand and predict climate variability & changes

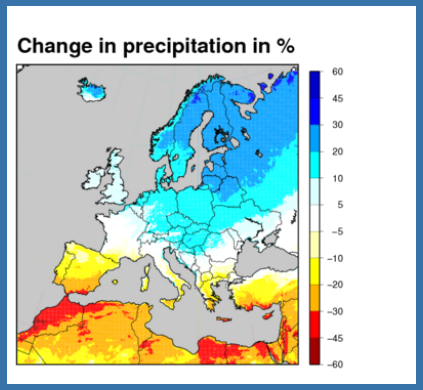
Foster:

- The integration of the European ESM community
- The development of ESMs and their environment
- High-end simulations
- The application of ESM simulations for climate change impacts

Support to international coordinated experiments for IPCC

CMIP5
7 european models





CORDEX

Euro-cordex
Africa Cordex

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

http://is.enes.org/

IS-ENES : Infrastructure for ENES

FP7 project « Integrating Activities »

1rst phase: March 2009- Feb 2013 (7.6 M€), 18 partners
2nd phase: Apr 2013- March 2017 (8 M€), 23 partners

Better understand and predict climate variability & changes

Foster:

- The integration of the European ESM community
- The development of ESMs and their environment
- High-end simulations
- The application of ESM simulations for climate change impacts

Institut Pierre Simon Laplace
Max-Planck-Institut für Meteorologie
CERFACS
DKRZ
National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL
The University of Manchester
MANCHESTER 1824
DMI Vejr, klima og hav
UNIVERSITAS BERGENSIS
Science & Technology Facilities Council
cmcc Centro euro-Mediterraneo sui Cambiamenti Climatici
Met Office
Royal Netherlands Meteorological Institute
Ministry of Transport, Public Works and Water Management
METEO FRANCE Toujours un temps d'avance
IC3
UC UNIVERSIDAD DE CANTABRIA
SMHI
LINKÖPINGS UNIVERSITET
BSC
WAGENINGEN UR For quality of life
INHGA
DLR
Meteorologisk institutt met.no

Enhance sharing of climate model data CMIP5/CORDEX

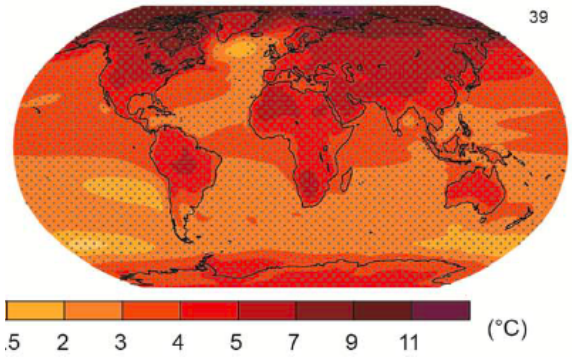
Support CMIP cycles

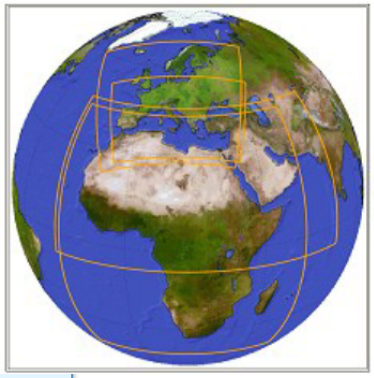
CMIP5

Preparation of CMIP6

NCC	NorESM1-M NorESM1-ME
MPI-M	MPI-ESM-LR MPI-ESM-MR MPI-ESM-P
MOHC (with INPE)	HadCM3 Hadcm3Q HadGEM2-A HadGEM2-CC HadGEM2-ES
EC-EARTH	EC-EARTH
IPSL	IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR
CNRM-CERFACS	CNRM-CM5 CNRS-CM5-2
CMCC	CMCC-CESM CMCC-CM CMCC-CMS

ESGF Earth System Grid Federation





Support sharing of CORDEX data through ESGF

http://is-enes-data.github.io/CORDEX_status.htm

Status at 23/08/2016

AFR 44 (8 RCMs), EUR 11 (8 RCMs), EUR 44 (8 RCMs),
Fewer for ARC 44 (5) and other domains
Mostly European models

DKRZ support

ISI-MIP

Inter_Sectoral Impact MIP

Enhance sharing of data

Contribute to development of ESGF & ESGF governance

Implementation Plan V1.0
May 2016
Dean Williams et al.

Table 1. *The current list of ESGF working teams and designated working team leads.*

Working Team	Working Team Leads	Team Goals
1. CoG User Interface Working Team	Cecelia DeLuca (NOAA) and Luca Cinquini (NOAA)	Improve ESGF search and data cart management and interface
2. Metadata and Search Working Team	Luca Cinquini (NASA)	Implement ESGF search engine based on Solr5 and discoverable search metadata
3. Publication Working Team	Sasha Ames (DOE) and Rachana Ananthakrishnan	Enable capability for publishing CMIP and other project data sets to ESGF
4. Node Manager Working Team	Sasha Ames (DOE) and Prashanth Dwarakanath (IS-ENES)	Manage ESGF nodes and node communications
4a. Tracking/Feedback Notification Working Team	Sasha Ames (DOE) and Prashanth Dwarakanath (IS-ENES)	Implement user and node notification of changed data in the ESGF ecosystem
5. Identity Entitlement Access Management Working Team	Philip Kershaw (IS-ENES) and Rachana Ananthakrishnan (DOE)	Implement ESGF X.509 certificate-based authentication and improved interface
6. Compute Working Team	Charles Doutriaux (DOE) and Daniel Duffy (NASA)	Develop data analytics capability within ESGF
7. Quality Control Working Team	Martina Stockhause (IS-ENES) and Katharina Berger (IS-ENES)	Integrate external information into the ESGF portal
8. Installation Working Team	Nicolas Carenton and Prashanth Dwarakanath (IS-ENES)	Install the components of the ESGF software stack
9. Dashboard Working Team	Paola Nassisi (CMCC) and Sandro Fiore (IS-ENES)	Monitor the Earth System Grid Federation in terms of system metrics and data usage statistics
10. International Climate Network Working Group	Eli Dart (DOE/ESnet) and Mary Hester (DOE/ESnet)	Increase data transfer rates between the ESGF climate data centers
11. Data Transfer Working Team	Lukasz Lacinski (DOE) and Rachana Ananthakrishnan (DOE)	Enhance ESGF data transfer and web-based download
12. Software Security Working Team	George Rumney (NASA) and Dan Duffy (NASA)	Implement security measures to identify vulnerabilities in the ESGF software and provide continuous improvement to the ESGF software development life cycle.
13. Support Working Team	Torsten Rathmann (IS-ENES) and Matthew Harris (DOE)	Develop frequently asked questions regarding ESGF and housed data
14. Documentation Working Team	Matthew Harris (DOE) and Sam Fries (DOE)	Document the use of the ESGF software stack
15. Replication and Versioning Working Team	Stephan Kindermann (IS-ENES) and Tobias Weigel (IS-ENES)	Create replication tool for moving data from one ESGF center to another; in addition, preserve versioning history of the ESGF published data sets
16. Provenance Capture Working	Bibi Raju (DOE)	Enable ESGF provenance capture for

Infrastructure strategies to accelerate progress in Earth System modelling/observations and understanding ?

IS-ENES 1 &2 : common research infrastructure in Europe on climate modelling
Models and environment tools / HPC / data - WCRP experiments

Instrumental: accelerate R&D, allow joint developments & service on data
IS-ENES & national support
Limitation: IS-ENES on project funding

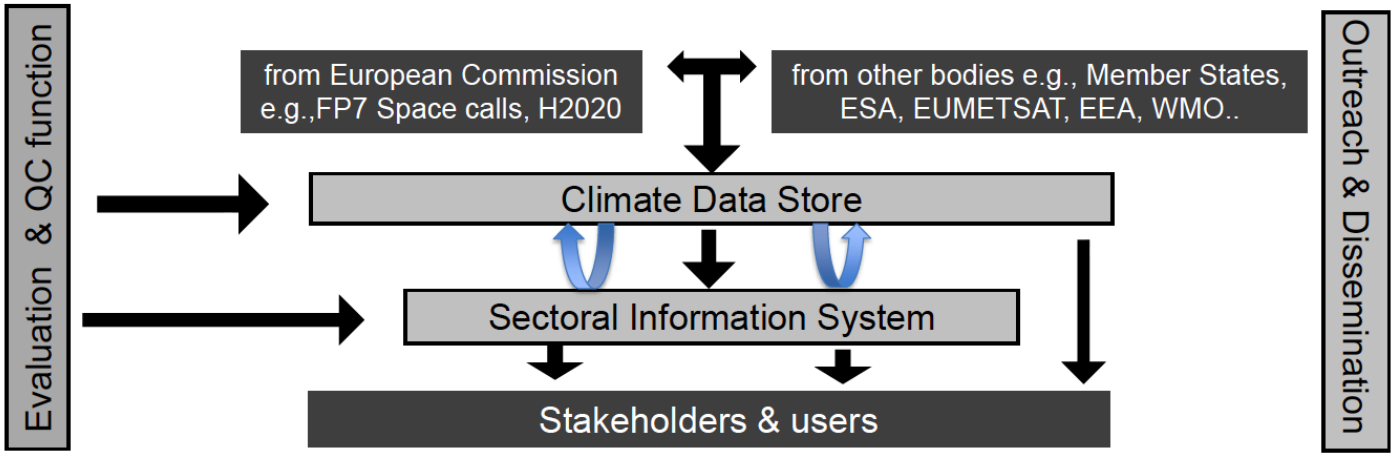
Infrastructure dimension also recognised at WGCM level with WIP
Probably still need be better accounted for in the overall design of CMIP

Interoperability with observations – Obs4MIP, ana4MIP in ESGF
ESGF as WCRP strategy for data

Models & HPC dimension
Sharing of SW; Prepare for future architectures; Access computing facilities
Started with IS-ENES, now extended through the
ESIWACE Center of Excellence on weather and climate (Coord J. Biercamp)

Copernicus
Climate Change Service
C3S
at ECMWF

Funded by
the European Union



Climate Data Store
Reanalyses
Satellite data (ESA CCI)
Seasonal Forecasts
Projections (Global/Regional)

Sectoral Information System
Energy, water, agriculture, health ….

Copernicus Climate Change Service C3S at ECMWF

Funded by the European Union

enes EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

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

Climate Services

Build operational services on ESGF and IS-ENES

Climate Information Portal for Copernicus
CLIP-C project
Coordination M. Juckes

Precursor project
2013-2016

C3S Climate Data Store

Global projections : broker on ESGF
CP4CDS: ESGF node for C3S (2016-2019)
coord B. Lawrence (CEDA, DKRZ & IPSL)

Contribute to further developments :

- ESGF data node maintenance (IPSL)
- CP4CDS compute node (DKRZ)
- Integration of Information and tools (CEDA)

Regional projections:
Submitted proposal to also broker on ESGF
Coord S. Denvil (IPSL, DKRZ, CEDA, SMHI, UCAN, ENEA, LIU)

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

What are the key things that are difficult to do today and are impeding scientific progress or productivity and the sharing of data?

How to perform multi-model analyses on ESGF ? need for local multi-model archives

How to ease model evaluation ? Sharing of common standard analyses ; Open access APIs with full provenance for data and methods;

Ease access to other communities
e.g. climate impact research

Climate4impact platform:

- Explore climate data and perform analysis
- Need for in-depth documentation and guidance
- Use cases from impact researchers
- Perform calculations / Data processing

http://climate4impact.eu/

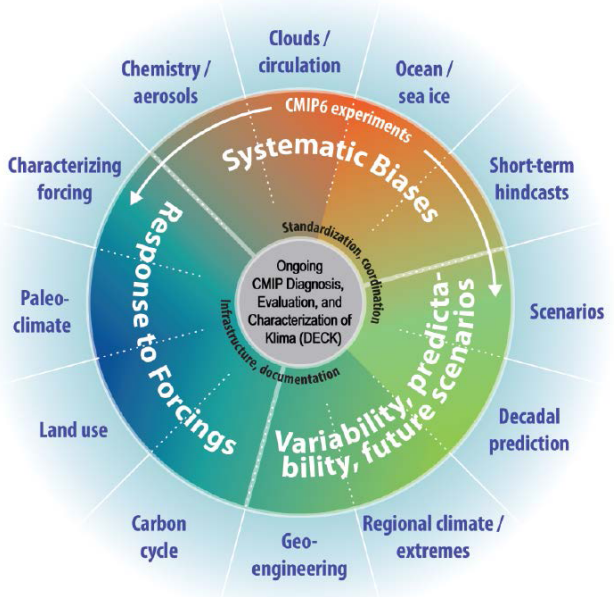
is-enes
INFRASTRUCTURE FOR THE EUROPEAN NETWORK FOR EARTH SYSTEM MODELLING

What is your timeline for data production and distribution from climate model and observations, high-performance computer, network, and storage facilities needs and investments?

European models in CMIP6

Country	name of model (CMIP5)
Consortium	EC-EARTH
France	IPSLCM5
France	CNRM-Cerfacs
Germany	MPI-ESM
Italy	C-ESM
UK	HadGEM2
Norway	NorESM

Germany:
AWI-CM
EMAC



Short name of MIP	AWI-CM	CMCC	CNRM	EC-Earth	EMAC	IPSL	MPI-ESM	Nor ESM	UK ESM	Had GEM3
AerChem MIP	0	0	1	1	1	1	0	1	1	0
C4MIP	0	1	1	2	0	1	1	1	1	0
CFMIP	0	0	1	1	0	1	1	1	0	1
DAMIP	0	0	1	0	0	1	0	1	0	1
DCPP	0	1	1	1	0	1	1	1	0	1
FAFMIP	0	0	1	0	0	1	1	0	0	1
GeoMIP	0	0	1	1	0	1	1	2	1	0
GMMIP	0	1	1	0	0	1	1	0	0	1
HighRes MIP	1	1	1	1	0	2	1	2	0	1
ISMIP6	0	0	1	1	0	1	1	0	1	0
LS3MIP	0	1	1	1	0	1	1	2	1	0
LUMIP	0	1	1	1	0	1	1	1	1	0
OMIP	1	1	1	0	0	1	1	1	1	0
PMIP	1	0	1	1	0	1	1	1	1	0
RFMIP	0	0	1	0	0	1	1	1	0	1
Scenario MIP	0†	1	1	1	0	1	1	1	1	0
VolMIP	0	0	0	1	0	1	1	1	1	0
CORDEX	1	0	1	1	0	1	0	0	1	0
DynVar	0	0	0	1	0	1	1	0	0	1
SIMIP	1	1	1	1	0	1	1	1	1	0
VIACS AB	0	1	0	1	0	0	0	1	1	1

Source Veronika Eyring, WGCM 2016

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

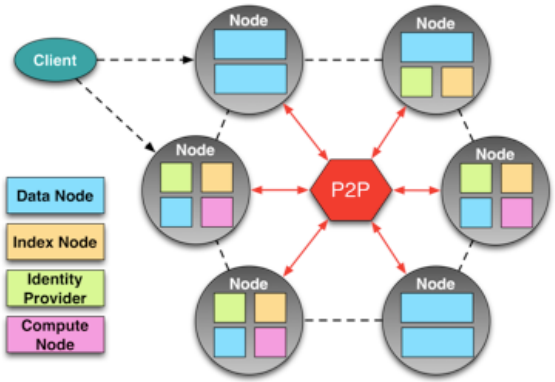
What is your timeline for data production and estimated size of distributed archive ?

Estimated numbers to be confirmed (ENES HPC and Data TF and CDNOT)

Center	Core hours (Mh)					Data (PB)	Data ESGF	Data ESGF
	Total	2016	2017	2018	2019	Total	Amount	start ?
CMCC (IT)	**100**					**1**	**1**	2017
MetOffice NCAS UK)	**500**					**12**	***1 to 2 provisioned at CEDA***	
CERFACS	**11**					**0,7**	**2**	Q2 2017
CNRM (FR)	**150**					**2**		
IPSL (FR)	**290**					**14**	**2**	mid-2017
BSC (ES)	**58**					**3**	**3**	follow production
MPI-M/DKRZ DE	**84**					**6**	**2**	S1 2017
Bergen / Met Norway (NO)	**110**					**1**	***1 2 provisioned***	
Total	**1303**					**39,7**	**12 to 14**	

Germany also includes simulations done at AWI, DLR and DWD

What are your common developments, sharing of expertise, and accelerated developments?



IS-ENES services & development

CMIP5 European Index nodes
CEDA / DKRZ / IPSL/ CMCC

Euro-cordex index nodes
CEDA/DKRZ/IPSL/DMI/LIU

Security SW	***CEDA***
QC SW	***DKRZ/CEDA/SMHI***
Monitoring	***CMCC***
Replication SW	***IPSL***
Replication sites	***CEDA/DKRZ***
Access to metadata	***IPSL/DKRZ/CEDA***

Installation SW	***IPSL/LIU/DKRZ/ CEDA***
Help/FAQ	***DKRZ/BADC****/PCMDI*

Joint Research Activity

Data citation	***DKRZ***
Controlled vocabularies	***NCAS/BADC****/NOAA*

C4I interface

Ease access to model data
Processing tools
KNMI/CERFACS

What are the administrative/sponsor requirements that arise from each project (basically, metrics collection and reporting)?

IS-ENES2 : Key performance Indicators

Downloaded data volumes/files per month (pb: dashboard needed)

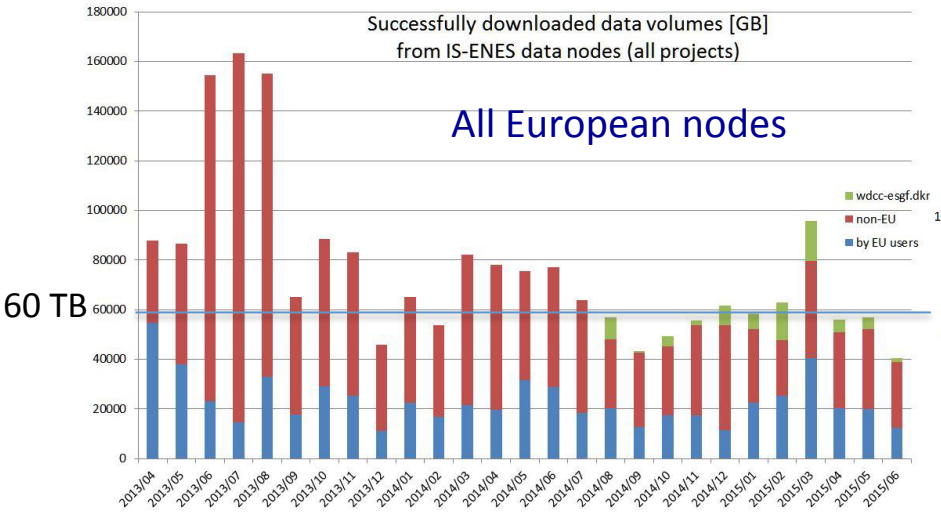
Statistics on active users of IS-ENES nodes

Number/percentage of reply threads to international ESGF

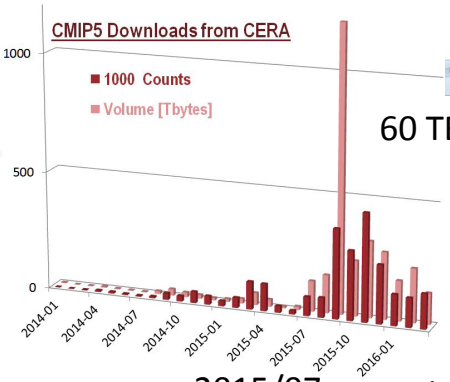
Statistics on use of climate4impact portal

Issue of visibility : hidden behind ESGF

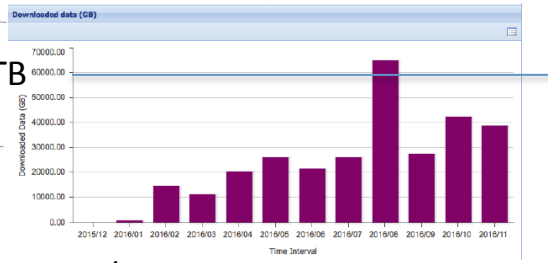
Downloaded data volumes – IS-ENES data nodes



CERA only during ESGF downtime



Only DKRZ, IPSL, DMI



What are your expected strategic roadmap and ESGF funding levels for the short-term (1 to 3 years), mid-term (3 to 5 years), and long-term (5 to 10 years)

IS-ENES2: project support until end of March 2017

ENES Data Infrastructure: Continue to work together

Apply for IS-ENES3 (but not before 2018)
Other joint projects: Copernicus climate change service project(s)

IS-ENES only part of support
Also national support (still to be quantified)

Governance: ENES Data Task Force (M. Lautenschlager Chair)

Need investigate possibilities for a long-term European research infrastructure
A multi-lateral agreement

What is the political landscape to be made aware of?

EUROPE

Strong emphasis on climate services

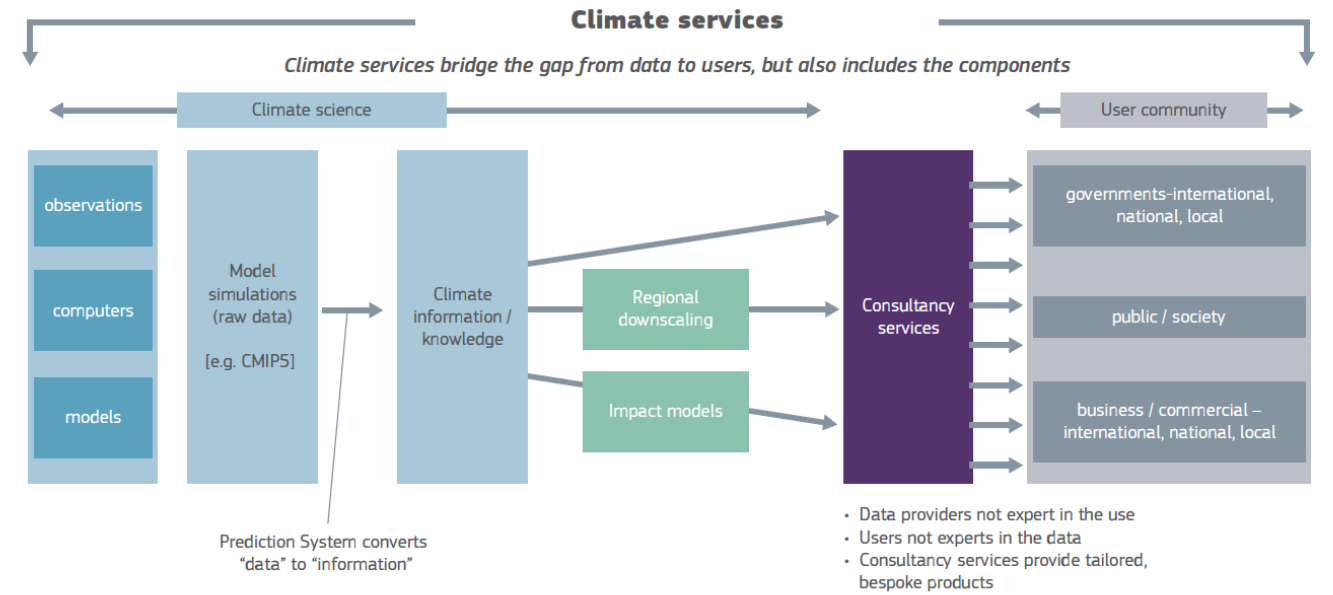
Climate modelling as part of the overall chain

H2020 EC framework programme / Copernicus Climate Change Service

Joint Programming on Climate (Member States)

A European Research and Innovation Roadmap for Climate Services (2015)

Figure 1. The Essence of Climate Services



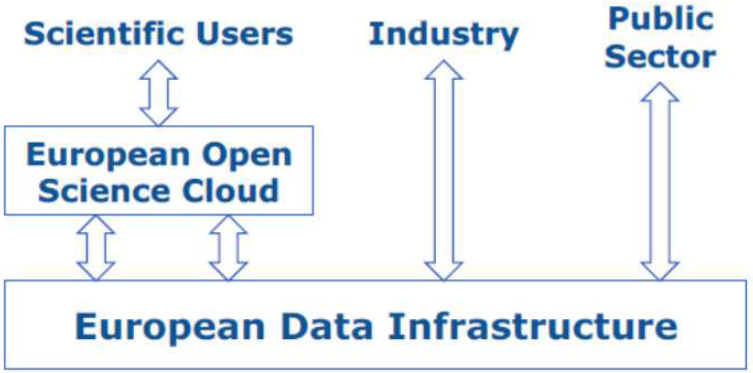
What is the political landscape to be made aware of?

European Science Cloud Initiative
Collaboration with EU e-infrastructures (eg PRACE and EUDAT)

European Cloud Initiative (highlights)

- *European Open Science Cloud (EOSC)*
  - **Integration and consolidation of e-infrastructures**
  - **Federation of existing research infrastructures and scientific clouds**
  - **Development of cloud-based services for Open Science**
  - **Connection of ESFRIs to the EOSC**
- *European Data Infrastructure*
  - **Development and deployment of large-scale European HPC, data and network infrastructure**
- *Widening access and building trust*
  - **SMEs, Government as a Service, Standards**

Static View on Year 2020 (over-simplification)



HPC + data centers + network
2 pre-exascale systems in 2020
2 exascale systems in 2022

EC Communication 2016
European Cloud Initiative:
Building a competitive data and knowledge economy in Europe

Conclusions

Need to better recognise the infrastructure dimension of CMIP WCRP experiments:

Long-term support of data infrastructure

But also better integration in the design of CMIP

CMIP : a strong infrastructure dimension for climate research

Europe: committed through ENES data infrastructure

Projects (IS-ENES, Copernicus) & National support

Still challenges to get a data infrastructure fulfilling science needs

Evaluation / Multi-model analyses

Key infrastructure for climate research but also other communities

Climate impact research / climate services

European Open Science Cloud: ESGF an asset

Integrate data and HPC / integrate with other communities