Thematic centers :

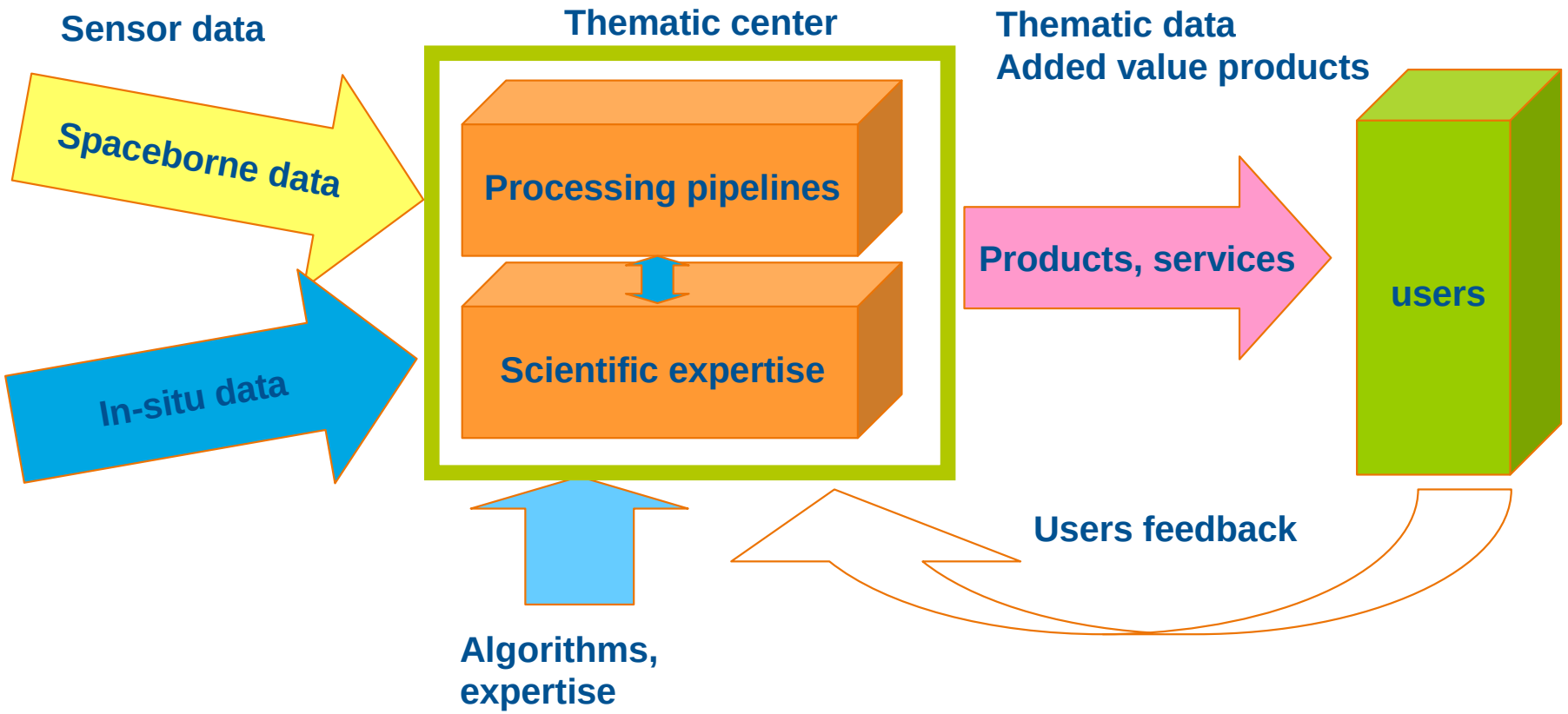
Functional model, data and services dissemination. Status and questions.

CNES

Workshop ESRIN – 11/10/2012

D. Roumiguières, T. Levoir

Generic functional model for the thematic centers



CNES is partner of 4 active centers :

- Ether, Icare, CDPP, MEDOC

Several “structures” working at the same functional level :

- SALP, IDS, CATDS, GIS-COOC, Centre Données astronomiques de Strasbourg (CDS)

1 center under construction (PTSC : land)

1 planned center (Formes et Mouvements de la terre = solid earth science)

Ether thematic center : overview
http://www.pole-ether.fr

- National thematic center
  - » Aiming to unite, coordinate and conduct activities dedicated to **Atmospheric Chemistry** (pollution and the interaction between chemistry and the climate)
  - » Serving the French, European and international scientific community
- Partnership CNES/CNRS-INSU
  - » Started in 1995 ; Operational in 2000
- Missions
  - » Makes available to the scientific community "Atmospheric Chemistry" products (spatial and non-spatial) and value-added services developed as part of Ether.
  - » **Develops applications, ensures the production, archiving** and provision of spatial and in situ data.
  - » Provides **expertise** on algorithms and data
  - » Provides access and support in the interpretation of data from the Atmospheric Chemistry
  - » Provides resources (software and tools) to process the data
  - » Gathers and matches the needs of users
  - » **Organizes networks of expertise** on atmospheric chemistry

Ether : products and services overview

Models, Assimilation
(LATMOS, LA, CERFACS, CNRM, LMD)

In-situ data
(LA, SA, LSCE, LACy, LOA, LaMP…)

Balloons data
(LPC2E,LPMAA …)

Spaceborne data
(LATMOS, LPCE, CNRM…)

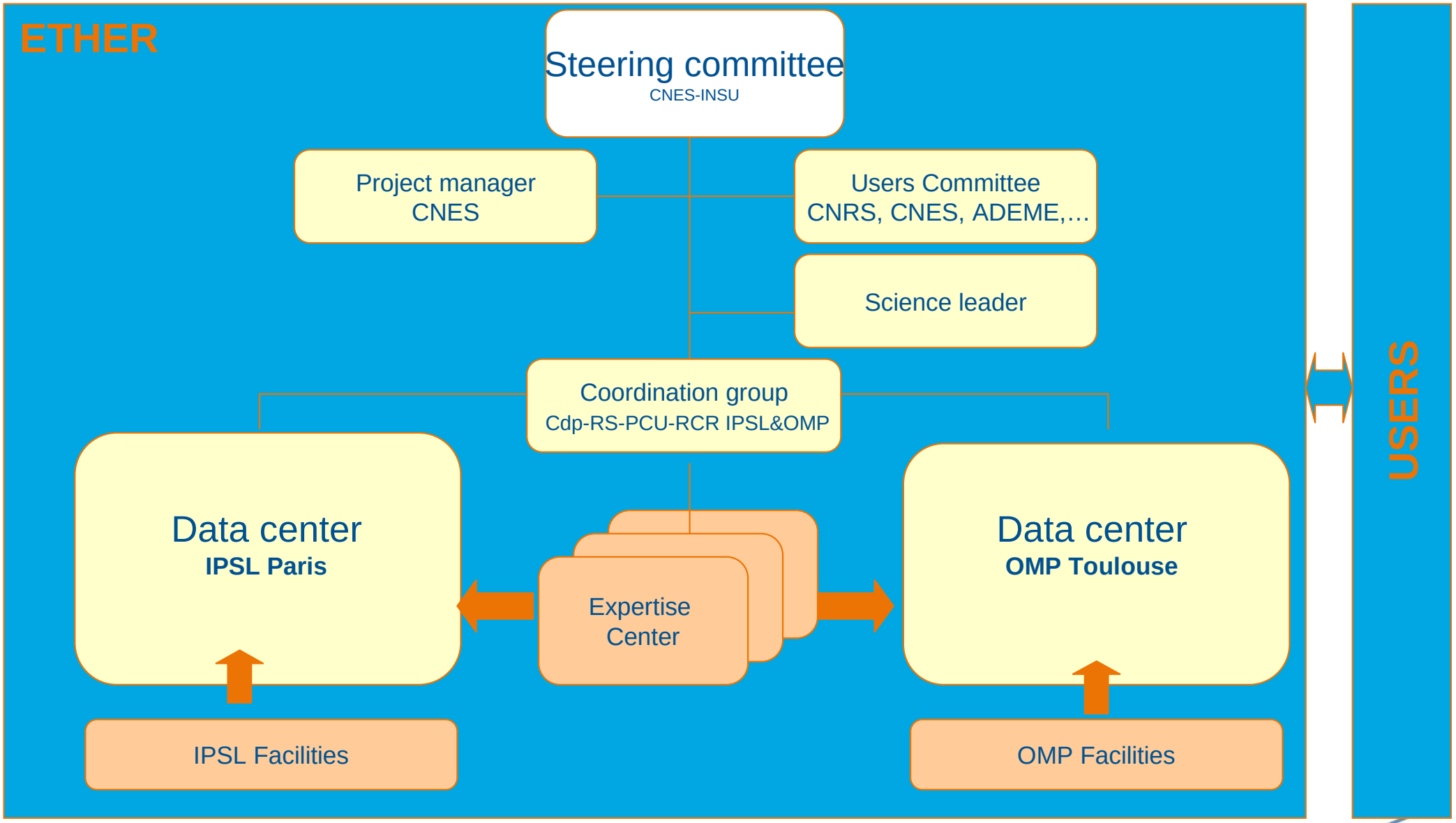
Data base
(LATMOS, LA, LMD, LISA…)

Educational tools

European or international

e.g. : ECCAD & IAGOS used by MACC (Monitoring Atmospheric Composition and Climate) current pre-operational atmospheric service of the European GMES programme

Ether : architecture / management



Thematic center ICARE : overview

Maximizing the use of satellite data to contribute to:

- Improve the scientific knowledge on **clouds, aerosols, their interaction with radiation, water and energy cycles**;
- Improvement of **weather and climate prediction**.
- Providing the national and international scientific community with a set of data, products and services in the thematic area « clouds, aerosols, water, radiation » to allow them an optimal use of space and in situ observations for climate modeling and study.

Timeline

- 2002: several French organizations (CNES, CNRS/INSU, Lille-1 University, the Nord-Pas-de-Calais region) start the cooperation to design Icare.
- November 2003: these 4 organizations sign the Constitutive agreement of the Icare Thematic center, with the support of the European commission.
- November 2005: the Icare Data center becomes a UMS (formal partnership), by agreement between CNRS, CNES and Lille-1 University).
- February 2009: the 4 organizations sign an extended version of the Constitutive agreement of the Icare Thematic center.

ICARE : Products and services
http://www.icare.univ-lille1.fr/

Operational :

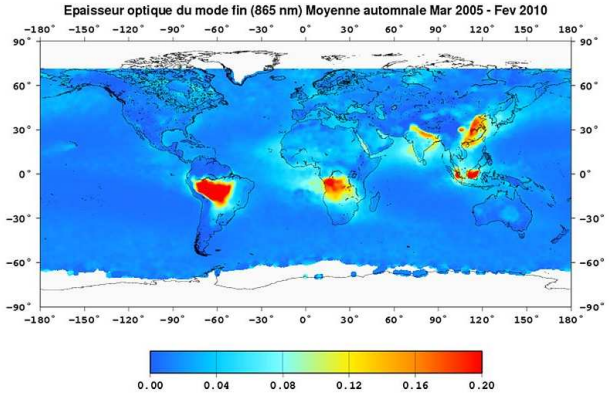
- PARASOL:
  - ✦ processing and distribution of level 2, 3 and 4 products (aerosols, clouds, radiation, merged Modis-Parasol products).
- CALIPSO:
  - ✦ Icare center acts as a mirror of the NASA ASDC center. Distribution of Calipso level 1 and 2 products: lidar profiles, IR and VIS radiances, cloud and aerosol products. Processing and distribution of browses. New products.
- Non-CNES other space missions:
  - ✦ MSG/Seviri (aerosol products), MODIS level 1, Cloudsat (cloud products)…

Ongoing

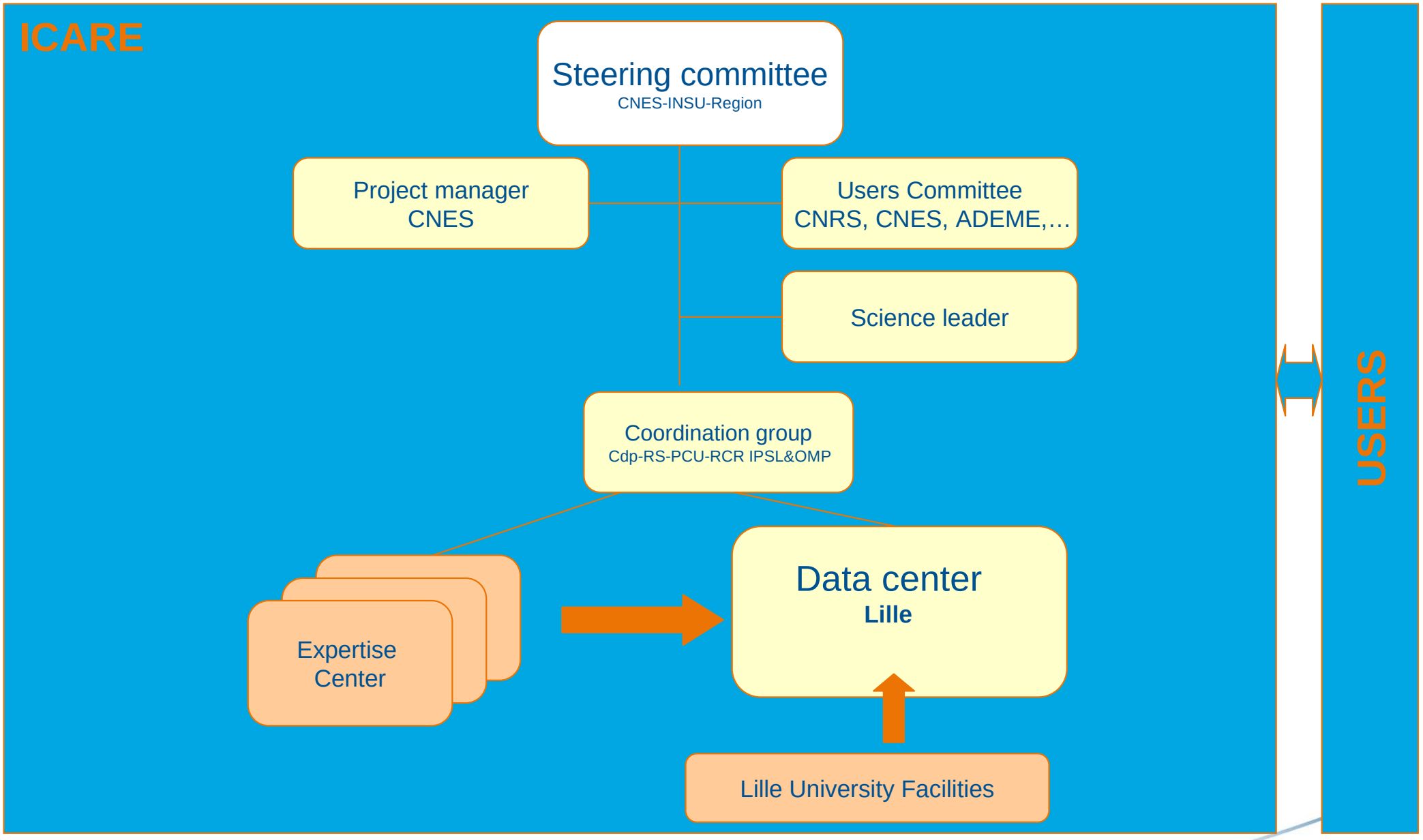
- MEGHA-TROPIQUES:
  - ✦ processing and distribution of level 2, 3 and 4 products (precipitation, water vapor, radiative budget, convective clouds…).
- Merged CALIPSO-CLOUDSAT products on ice-cloud microphysics.
- Systematic processing of the EPSAT-SG rainfall algorithm (MSG rain algorithm) over Africa during AMMA.
- Implementation of the STRAT lidar algorithm.
- Intercomparison of different aerosol satellite products (Parasol, Modis, Caliop, MSG/Seviri, Envisat/Meris) and comparison with Aeronet (ground remote sensing) products.
- Impact of cirrus clouds on the determination of NO2 with Aura/OMI.
- Intercalibration of the Calipso/IIR radiometer with infrared radiances of Modis and MSG/Seviri.
- A lot of processing and visualization tools.

3000 visitors/month

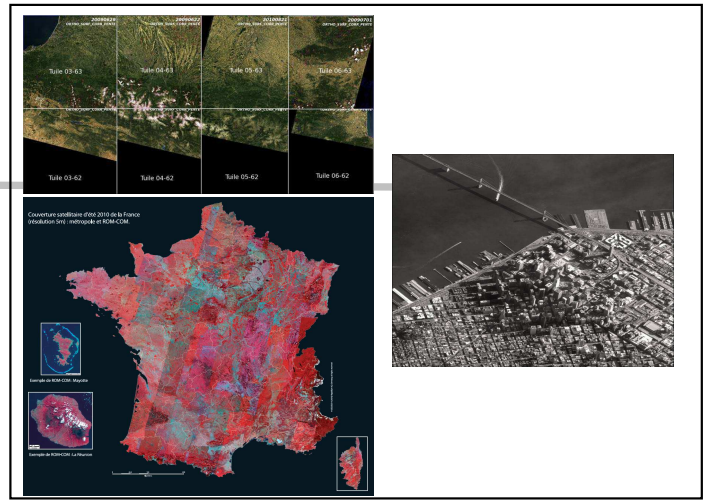
- 300 registered users for FTP data transfer
- 6 To/month of data FTP-transferred to users
- Remote processing unit designed for users.



ICARE : architecture / management



Land Thematic Center (PTSC) : overview



Addresses a wide range of scales and themes

- Both the high to the low resolutions
- scientific users and social issues of resource management

A strong national consensus :

- PTSC draft Convention prepared by six partners: CNES, IRSTEA, INSU, IRD, CIRAD, IGN

Building a Data Center (STSC) bi-located

- Toulouse : CNES-based investments, with the support of CESBIO
- Montpellier : based on the developments of infrastructure GEOSUD

Covers not only the scientific community but also public policy actors

Sentinel-2 data processing should be a major focus of development from 2015

- The priority is to cover France (including French overseas territories) with Sentinel-2 products daily (level 2) and monthly (Level 3) with production facility Muscat
- Possible future :
  - ✦ Sentinel-2 production may be extended to Europe in a cooperative PTSC / GMES
  - ✦ Sentinel-2 production could be extended to Africa in a cooperative PTSC / GMES-Africa

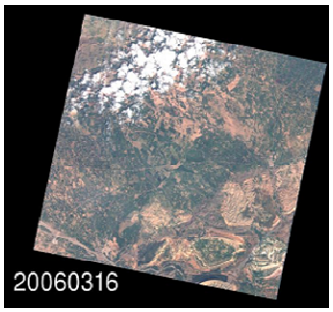
PTSC : Products and services overview

Products : mainly levels 2 and 3

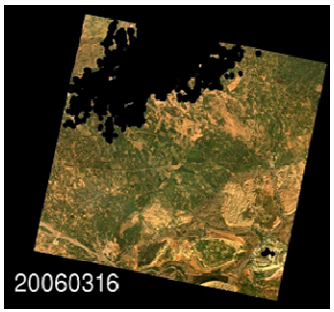
- Time series of surface reflectance at high or very high resolution, in France (including including French overseas territories) and other regions of interest
- Time series of biogeophysical products (e. g. Leaf Area Index, water depths, biomass, soil moisture ...) on a global scale

Services : priority is given to products access for users

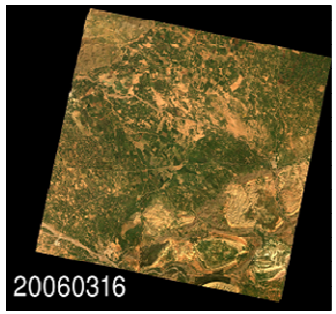
- Viewing products
- Product Downloads
- User Management
- Dissemination of tools and methods



Niveau 1C:
Réflectance Top of Atmosphere orthorectifiée

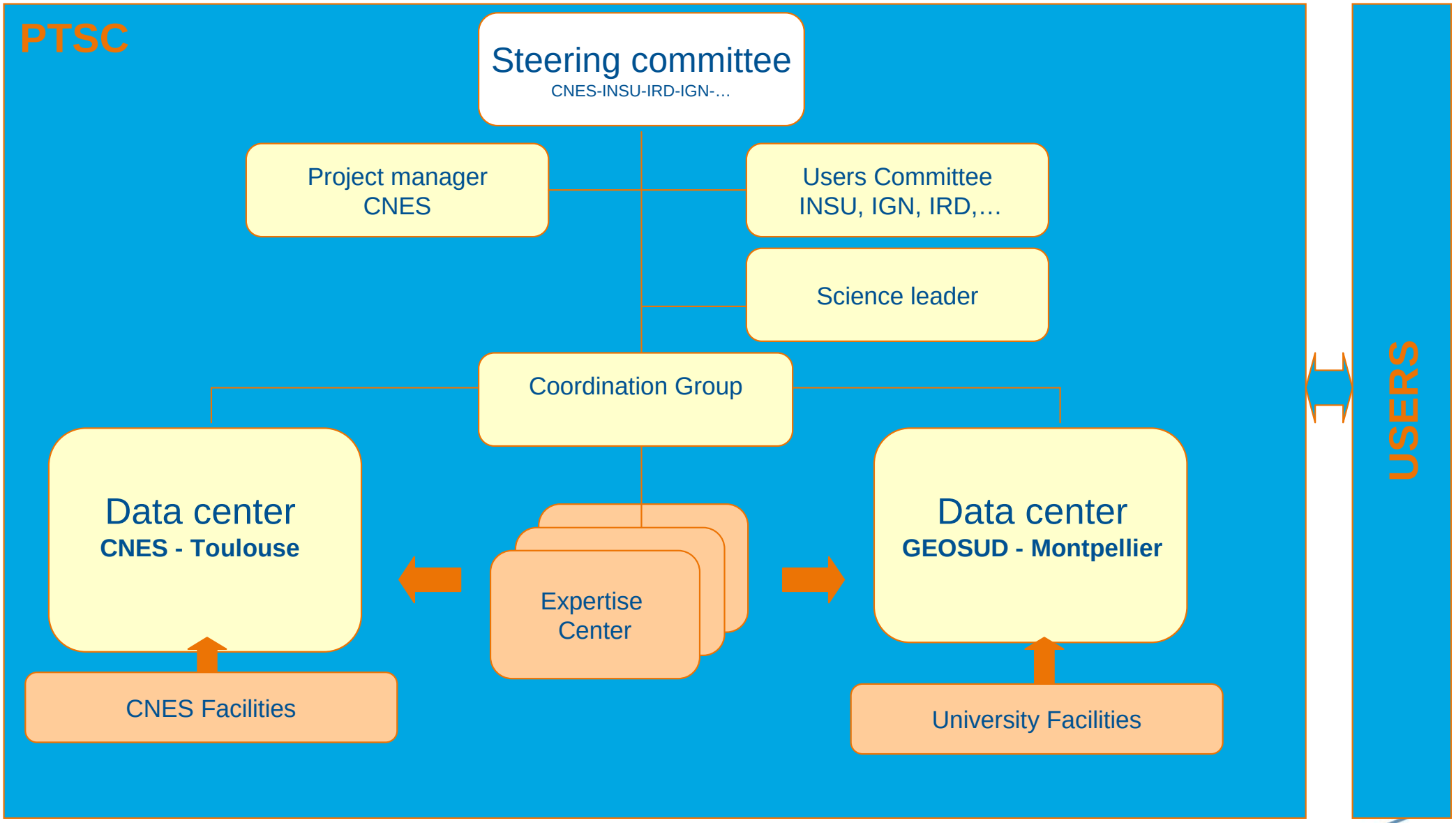


Niveau 2A:
Réflectance de surface journalière après correction atmosphérique et masquage des nuages



Niveau 3A:
Synthèse sur 15 jours des réflectances de surface

PTSC : architecture / management



Analyze / questions

Thematic structure

- Science is structured according to disciplines / themes
- Thematic centers are structured the same way
- when declined to low levels data this **thematic affiliation may appear as a limitation** :
  - » Interoperability between centers from different themes (catalog and data) is not easy
  - » data sharing between disciplines is difficult
  - » interdisciplinary products are not well hosted

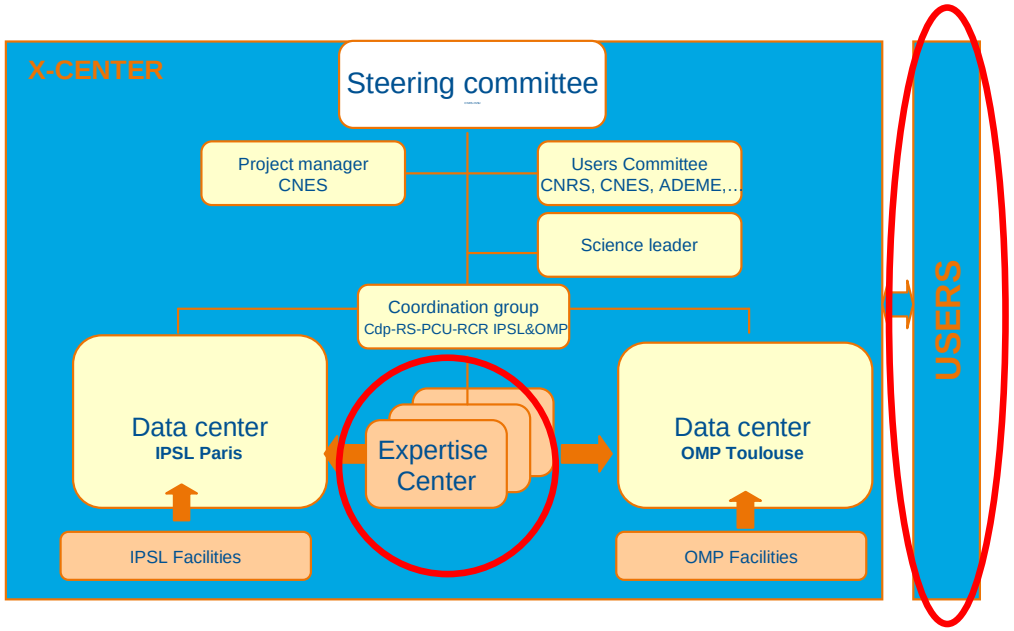
Some other constraints may be relevant when gathering data in a same data center :

- Data acquisition mode
- Data preservation constraints
- Data policy
- Help desk experts proximity
- …

The theme appears as a data user point of view, not a data producer point of view

- SMOS : Soil moisture et Ocean salinity
- Doris : auxiliary data for many missions and data for geodesy
- SWOT : oceanography and hydrology

Analyze / questions



The thematic centers model addresses two different communities for one theme :

- internal community (experts)

  &

- external community (users)

➔ issue : To be **IN** or **OUT** ?

At national level, this split of the scientific communities is not relevant :

- the center is a resource providing services **to a single community** :
  - Data acquisition, Data archive capacity, Data production hosting capacity, Data diffusion, … etc

Do we need a different model of the data center, a different architecture ?

- The range of provided services must be extended to
  - Data long term preservation,
  - process development facilities,
  - scientific communication,
  - outreach
  - etc …

Analyze / Questions

Ability to take operational commitments?

- To grow up, the downstream sector (including scientific) must have confidence in the **sustainability** of the upstream providers
- For the time being, long term data preservation and operational production commitments are not taken into account in the framework of thematic centers
- The structure of thematic centers must be improved to enable this kind of service commitments.

The architecture of thematic centers is based on data dissemination.

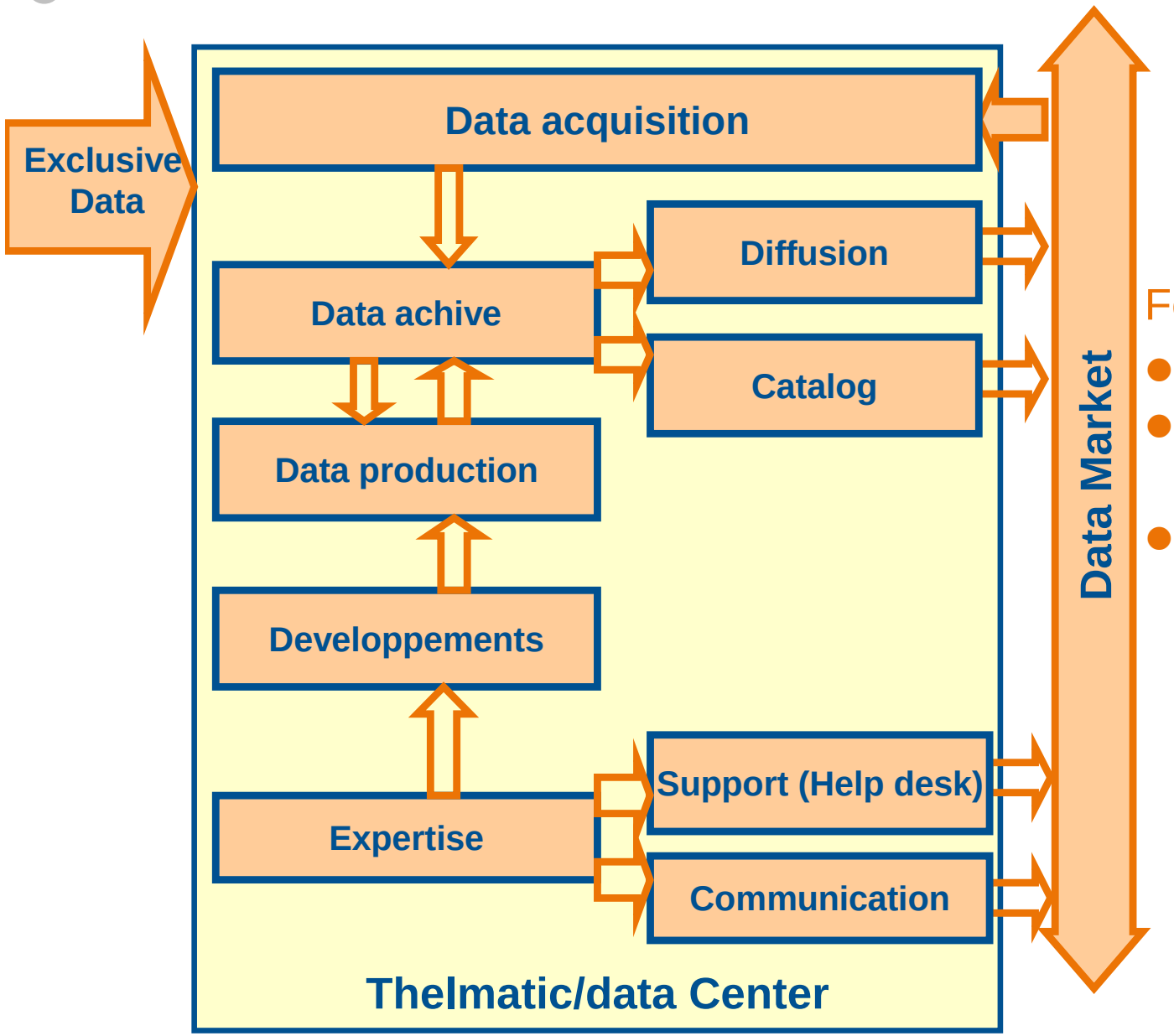
- But the data volumes created by future missions will increasingly constraint the exchange of such data
- Hosting processing code must be provided as a service to **avoid data transmission**.
  - ✦ (Some of the thematic centers already provide this service).

Thematic centers are designed to serve scientific community,

- but some others partners have a role in the downstream services
  - ✦ Actors of public policy
  - ✦ commercial compagnies

➔ Thematic centers must be prepared to involve non-scientific partners and to share data and services with them.

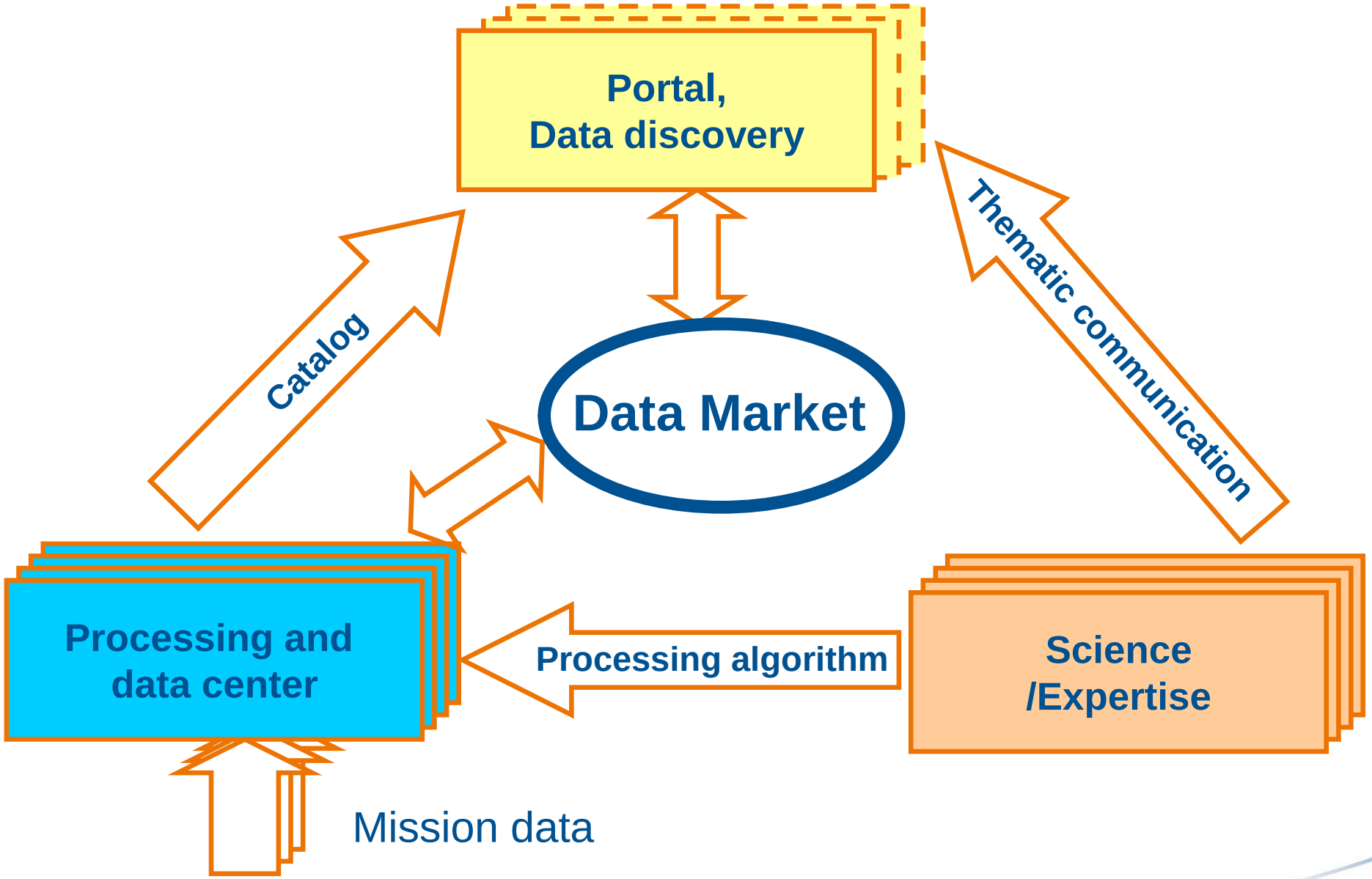
Functional diagram of a data center



For each functions : Questions :

- Is it in the data center?
- How is it implemented?
  - Infrastructure ?
- What is the governance model ?

Next step : Consolidation of main functions ?



Sharing hardware, virtual computing : a way ?

Consolidation of all functions on a "super computer center" (Virtua), with dynamic allocation of resources: with the following capacity

- Hosting scientific computations, data production
- Hosting developments Mechanism ("sand box“ non-intrusive on the central system)
- Hosting data (all disciplines)
- Hosting portal and data discovery functions

Benefits:

- Sharing of databases (avoid transfers)
- Sharing models and processing (avoids re-development)
- Management at the center level of heavy development technology (security, high availability of resources, interoperability, archiving, sustainability, standards compliance etc ...)
- High performance computing power available (peaks absorption capability e.g. for reprocessing)

Encouraged interdisciplinary collaboration:

- Data sharing without exchange
- Project communication hosted at the portal level (wide visibility)
- Sharing methods (algorithms and Designs)
- Interdisciplinary Data Discovery (centralized portal)
- Collaborative mechanisms at the center level : forum, wiki, blog, etc ...

Conclusion

In the forthcoming years, in order to adapt to the changing environment and the evolving needs of users, the French thematic centers must progress in 4 directions:

- Proposing a main thematic access to user communities while consolidating data and functions based on a logic of sharing functions and infrastructure.
- Moving towards a range of services not limited to the dissemination of data but extended to data preservation, hosting users processing pipelines, diffusion of thematic communication, …
- Strengthening the infrastructure and resources to address operational commitments
- Preparing cooperation with industry players and public policy actors in parallel of scientific users