

## The Theia Land Data Centre

Nicolas Baghdadi (1), Marc Leroy (2), Pierre Maurel (1), Selma Cherchali (2), Magali Stoll (3), Jean-François Faure (4), Jean-Christophe Desconnets (4), Olivier Hagolle (5), Jérôme Gasperi (2), Philippe Pacholczyk (2)

(1) *Irstea, Maison de la Télédétection, Montpellier*

(2) *Cnes, Toulouse*

(3) *IGN-Espace, Toulouse*

(4) *IRD, Maison de la Télédétection, Montpellier*

(5) *Cesbio, Toulouse*

**Highlights:** This paper describes the Theia Land Data Centre, its service and data infrastructure, network of scientific expertise centres, product portfolio and main activities. The paper emphasizes in its conclusion the connection between Theia and the Copernicus programme at European level, and the forthcoming Remote Sensing Data Infrastructure network at international level.

**Key words:** *remote sensing, data infrastructure*

### Background and Objectives

Ten French public institutions involved in Earth Observation, environmental studies and scientific research (Cea, Cerema, Cnes, Cirad, Cnrs, Ign, Inra, Ird, Irstea and Météo-France), have launched in 2012 the Theia Land Data Centre ([www.theia-land.fr](http://www.theia-land.fr)), pooling their expertise and resources to make satellite data available to the environmental research community and actors in charge of public policies. The Theia primary mission is (i) to build a national service and data infrastructure able to produce value-added space data over land and provide services fitted to users' needs, (ii) to support the sharing of experience and scientific knowledge on methodologies relevant to process and use space data for land thematic issues.

The overall objective of the Theia Land Data Centre is to enhance scientific knowledge and support the development of management capacities in domains related with anthropogenic action and climate impact on ecosystems and territories, the observation, quantification and modelling of water and carbon cycles, the monitoring and modelling of changes in societies and their activities (water management, urban planning, agriculture, forest management, ...), the understanding of biodiversity, its dynamics and related preservation strategies. To this end, Theia is working to produce data, products, methods and services linked to space observations of continental areas, from local (ecosystem and territory) to global scale, and make them available to the user community.

Theia's construction is based historically on the merge of activities linked to the Equipex Geosud project on the one hand, and to Cnes research and development on the other hand.

Equipex Geosud project was selected as part of the call for proposals "Equipment of Excellence" in the national Programme d'Investissements d'Avenir (large national bond issued in 2011). The project gathers a large number of institutional partners in research, high education and public management of the environment and territories. Equipex Geosud's mission is to develop a national satellite imagery infrastructure to serve the research on environment and territories and its applications in the management of public policies. The objective is to contribute to the full development of the potential of satellite imagery. Geosud is described in depth in a companion paper [1].

Cnes on its side has developed for more than a decade activities aiming at producing added value products from satellite imagery and distributing them to the scientific user community. Activities such as Kalideos, with long time series of Spot data on selected sites, Postel [2], a precursor of Theia, specialized in the production of low resolution time series of biophysical variables at global scale, Hydroweb, time series of large lakes and rivers level derived from altimetry data, Spot World Heritage, aiming at distributing freely to the international community a large fraction of the global archive of Spot data at orthorectified level, fall in this general category.

## Structure

### General

Theia is based on a distributed services and data infrastructure, and a network of thematic and regional expertise centres in various regions in France and in overseas French territories. Figure 1 shows Theia's general organisation and governance.

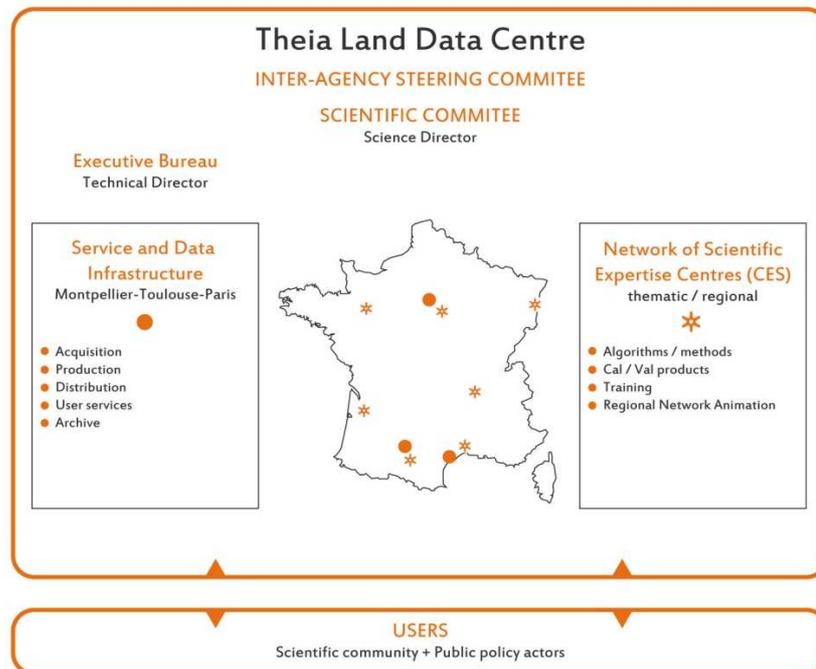


Figure 1 – Theia's structure and governance scheme

### Service and Data Infrastructure

Theia's infrastructure is based on a concept of database federation system (Figure 2). Currently, Theia's infrastructure relies on two data centres- the Cnes data centre located in Toulouse, and the Geosud data centre located in Montpellier, with an extension in Paris connected to the Geoportail data centre. The data centres are interconnected across a very high speed (10 Gbits/s) telecommunication network Renater, a High school and research national network. Each data centre is independent and responsible of pre-processing, storage, archiving and distribution of its own data to the users.

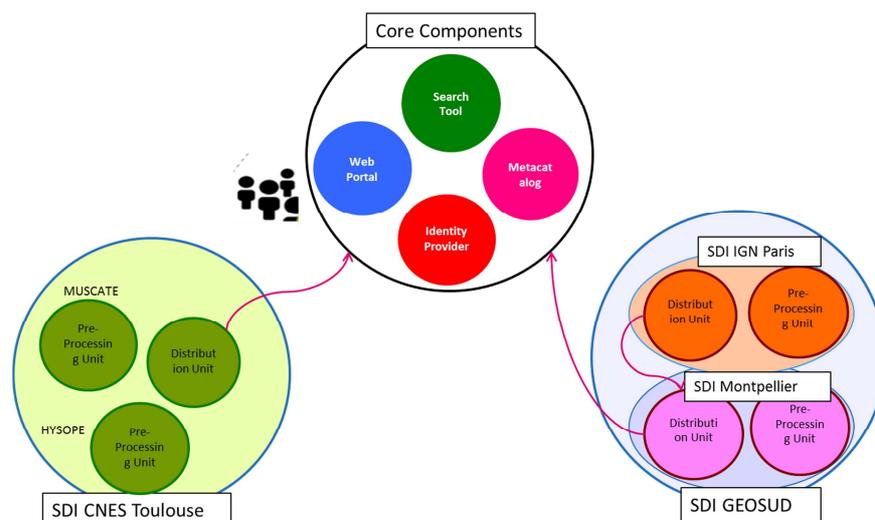


Figure 2 – Theia's Service & Data Infrastructure

Data are organized by project – usually by sensor. Depending on the project, the level of processing can go up to level 1 (e.g., orthorectification), 2 (e.g., atmospheric corrections), 3 (time composites) or more (e.g., land cover maps). Once pre-processed, data are archived and prepared for the distribution layer. At Theia's level,

interoperability between partners is achieved by the use of a standardized metadata model and a common vocabulary. The communication between data centres is ensured by sharing mutualized components used by all members of the federation : the Web portal, the Identity Provider, the Metacatalog and the Search Tool. Figure 2 shows the links between mutualized components ('core components') and the various data centres.

The association of new data centres to Theia's data infrastructure is possible. It is subject to the publication of data and metadata using REST web interface to ensure their inter-connection to shared services. Specifications of web services, metadata model and common vocabulary provide the technical framework for binding a new data centre to the Theia infrastructure, which can be easily integrated in a larger federation.

### Network of Scientific Expertise Centres

The Scientific Expertise Centres (SECs) are laboratories or groups of national laboratories developing innovative processes to use space data for land surfaces issues. They potentially have a dual purpose "thematic" and / or "regional".

The "thematic" SECs are focused on added-value products, possibly with services associated with these products. These are mono or multi-team, spread over one or more regions. Their objectives are to take part in the validation of the products provided by the Service & Data Infrastructure, develop processes to use the data and demonstrate new applications; contribute to network and federate the scientific players at regional, national and even international levels, around thematic fields: farming, forestry, urban, coastal, mountain, global approaches...

"Thematic" SECs have been identified at this stage around the following added-value products : Surface Reflectance, Albedo, Land cover, Vegetation biophysic variables, Evapotranspiration, Irrigated surfaces, Digital soil mapping, Soil humidity, Forest biomass and changes in forest cover, Water levels of lakes and rivers, Continental waters colours, Snow-covered surface, Urbanisation / Artificialisation, Risks associated with infectious diseases, High frequency change detection.

	TYPE	PRODUCT	AREA	PERIOD	ACCESS	AVAILABILITY
IMAGES VERY HIGH SPATIAL RESOLUTION		Pléiades	Small areas in France and other sites	2012 - 2014	National public actors	Available
		Spot 6	France and other sites	2014 - present	National public actors	Available
		Spot 6/7	Direct reception over France	2015	National public actors	July 2015
IMAGES HIGH SPATIAL RESOLUTION		Spot 4 (Take 5)	45 sites worldwide	Feb - June 2013	All users	Available
		Spot 5 (Take 5)	100 sites worldwide	April - August 2015	All users	July 2015
		Rapid Eye, Spot 1-5	France	1995 - 2013	National public actors	Available
		Spot World Heritage	More than 100 000 images worldwide	1986 - 2008	All users (non-commercial use)	1st semester 2015
		Kalideos	Spot time series for 4 sites in France	2000 - 2012	Users group	Available
		Landsat	Time series for France	2005 - 2011 - April 2013 - present	All users	Available
		Sentinel-2	Western Europe and other areas worldwide	From Sentinel launch	All users	Early 2016
OTHER PRODUCTS		Venµs	50 sites in the world	From Venµs launch	All users	Dec. 2016
		AVHRR	Global Leaf Area Index and fAPAR	1978 - 2013	All users	Dec. 2015
		Hydroweb	Water level, lakes and rivers : Jason and others	2000 - present	All users	Available
		Lidar	Some sites worldwide	2003 - 2009	All users	Available
		Radar	Some sites in France and abroad	2013 - 2014	All users	Available

Figure 3 – Theia Product Portfolio

"Regional" SEC objectives are to unify and coordinate users (scientists and public stakeholders) at regional level, participate in community training efforts, particularly concerning added-value products developed by the "thematic" SECs. Regional SECs have been set up so far in Alsace, Aquitaine, Languedoc- Roussillon and Midi-Pyrénées.

## Product portfolio

Products and services provided by Theia are intended to be quality-controlled, to cover broad territories and long periods: annual satellite coverage of the national territory, high or very high resolution surface reflectance time series, biogeophysical variables (biomass, water levels, surface humidity, etc) time series and products at global scale, visualization and data processing tools, processing methods and algorithms, validation procedures, methodological guidelines and frameworks for thematic applications. The products portfolio for the period 2013 – 2016 is shown in Figure 3. It includes images mainly in the optical domain at very high, high and low spatial resolution. Specific features of the portfolio are outlined below.

### *Satellite data acquisition over the national territory and abroad*

A programme of satellite data acquisition over the national territory has started, mainly driven by the Equipex Geosud project. The program includes so far a yearly cover of the national territory at 5 m resolution with Spot, Rapid Eye and IRS data for the years 1997, 2005, 2009 to 2012, the same at 1.5 m resolution with Spot 6 in 2014. It will continue for the years to come through a specific partnership between 6 public institutions already members of Theia, and the acquisition of a direct reception antenna and a terminal receiving Spot 6/7 data at Maison de la Télédétection at Montpellier. The data acquisition also includes the very high resolution coverage of a number of cities and sensitive areas (rivers, coastal line, habitats...) with Pléiades data. A specificity of the project is that the data can be made available free of charge to the whole national public sector. User workshops are organised to exchange information and feedback on user needs. Images are also acquired over areas in southern countries in support to scientific partnerships abroad.

### *The Muscate production centre*

A processing centre called Muscate has been developed by Cnes to transform data in surface reflectances (level 2), using an innovative atmosphere correction scheme based on physical principles able to produce from time series of raw data smooth temporal profiles of reflectances for each pixel. The processing line can process data from Landsat, the Spot 1-5 series in the framework of the Spot World Heritage programme, and Sentinel-2 data. The target of the Theia Land Data Centre is to process Sentinel-2 at level 2 (atmospheric corrections) and level 3 (monthly nearly cloud free composites) in a systematic way over Western Europe and other interest areas. Muscate has already processed Landsat data over France, and from the Spot 4 / Take 5 experiment designed to simulate the repetitiveness and coverage of Sentinel-2 data. It is currently processing data from a new similar Take 5 experiment made this time with the Spot 5 satellite at the end of its commercial life.

### *Future activities*

The lines of the product portfolio of Figure 3 are not commented further for the sake of conciseness. In the future we intend to continue along the main directions outlined above, to extend the portfolio with new value-added products developed by the Science Expertise Centres, strengthen the services and data infrastructure, and develop an access to in-situ data useful for researchers.

## European and international policy

Theia is a national effort and one of its goals is to become an element of a European network of actors carrying similar activities and to be well articulated with the Copernicus European programme. Theia's portfolio is therefore conceived to be complementary to that of Copernicus. At the same time, we gear the product development activities so that the developed products have if possible a European future. This has been the case for global biophysical vegetation variables developed first in the Postel context and which are now part of the Global Land Copernicus service. We will see whether Theia's atmospheric correction products for Landsat and Sentinel-2, and lakes and rivers water level, can become part of the Copernicus services in the near or mid-term future. At the same time we attempt, through the reply to call for tenders issued by ESA or EU to insert Theia in a network of European data centres having similar goals. It is felt that such a goal is essential to reach a full interoperability between data infrastructures developed at country level.

Beyond the European level is the international one. At this stage, it is felt that the best answer that can be given to the need of relations between actors at international level aiming at developing Earth Observation applications at institutional level is to start the building of the Remote Sensing Data Infrastructure international network which is the subject of the present workshop.

## References

- [1] Maurel, P., et al. Geosud. This issue.
- [2] Leroy, M., P. Bicheron, R. Lacaze, F. Niño, and G. Bégni (2005). Postel, an initiative to develop biogeophysical geocoded products. *Journal of Computational Technologies*, Volume 10, Partie 2, pp. 3-11.