

SERAD CNES Service for Data Referencing and Archiving

Danièle Boucon, Richard Moreno, Martine Larroque, Dominique Heulet, Paul Kopp, Michel Duplaa



Introduction

■SERAD

- Service for Data Referencing and Archiving
- ■will allow CNES to build a centralized repository structure that
 - collects, stores and disseminates information and metadata on all Data that are under CNES responsibility
- ■The objective for CNES is to better handle and to improve the access to his data patrimony.
- ■=> to identify all data which are relevant and to verify whether these data are properly archived
 - if not, then, to proceed to the archiving of these data.



Table of content

- **■SERAD** mission
- **■CNES** background
- **■**Functional definition
- **■Tools**
- **■**Definition of a minimal metadata model
- **■**Open points / To do



SERAD Mission

- ■To constitute and maintain an open and centralized metadata repository of all data that are under CNES responsibility
- ■To archive data when necessary
- ■To survey the data production centres in order to guaranty the long term preservation of these data even if (critical case) one of these centres has to be closed.
- ■This system will be built upon existing generic tools that will be customized
 - BDMS which is a kind of clearing house not limited to geographic information, but able to handle data from any thematic
 - * SIPAD-NG which is a generic tool allowing to give a full access (with search criteria) to (nearly) any kind of archived data

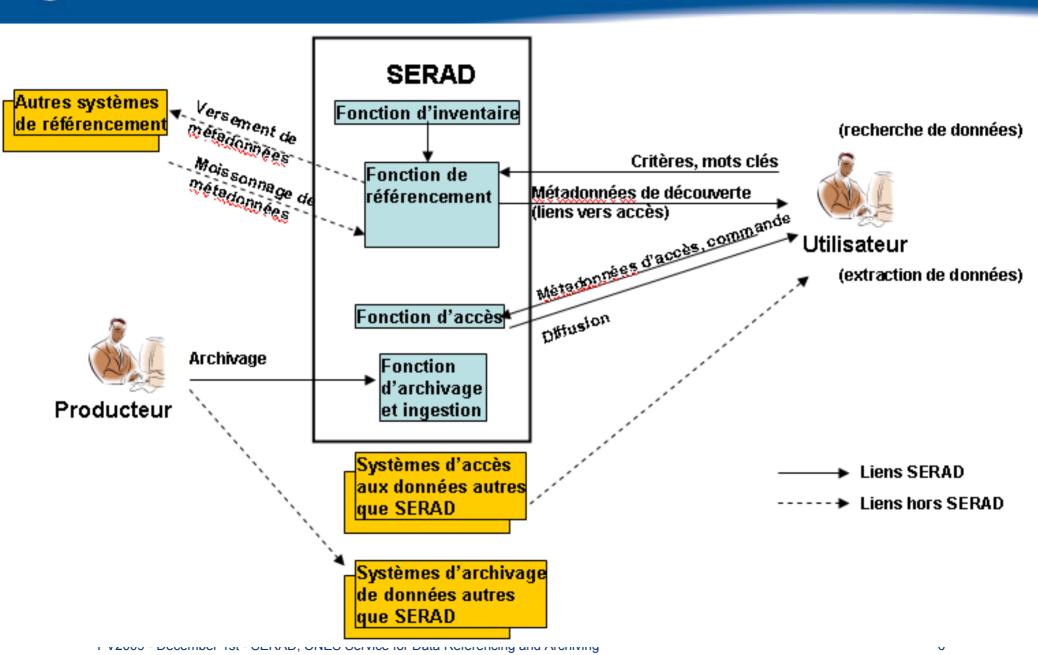




- ■For 40 years, in CNES, a big number of space missions have been producing a huge amount of data (hundreds of To).
- ■This data constitutes a valuable heritage that must be preserved because many of them are :
 - unique when related to an event that will never happen again or in a very long time
 - Halley comet period is 76 years!
 - integrated in long cycles of observations
 - including cycles climate change observation
 - mandatory to prepare future missions
 - GAIA take benefits from HIPPARCOS experience



Functional definition





Functional definition

■ Referencing function

- allows the user to search data of interest through the discovery metadata, using criteria (interests) and keywords.
- returns to the user the metadata for its selection.
 - These metadata contain descriptive information of potentially interesting data,
 - and the link to the data access function.
- will also broadcast, harvest other systems of referencing, as the IDN for example.

Access function

- allows users to order data.
- Its control is done using metadata access (on temporal criteria for example).
- Other services can be provided at that time (processing format for example).
- The access function allows the controlled release of data to the user.

■ Archiving and ingestion functions

• allows archiving data from a producer and ingestion of metadata access in the SIPAD-NG catalogue.

■ Inventory function

• provides an exhaustive inventory of the candidate data for referencing.

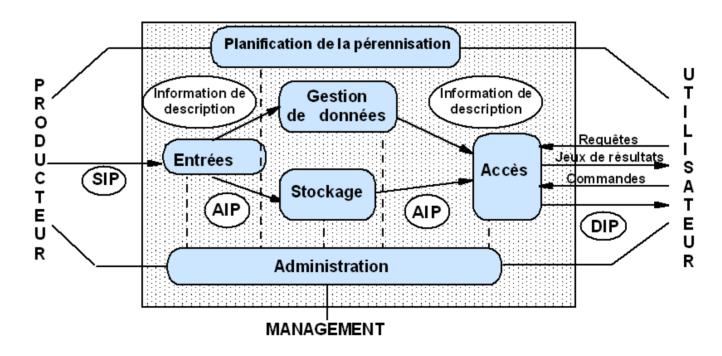
Tools



- ■The overall mechanism of referencing and archiving CNES data from CNES is based on ::
- ■TdB (Tableau de bord) which is a tool allowing to help SERAD management
- ■BDMS (Bureau des Métadonnées et des Services),
 - The BDMS is a referencing tool (clearing house) installed on the CNES Information System Direction machines.
 - It will provide an accurate description of the data sets and services that can be associated to them,
- ■SIPAD-NG (Système d'Information, de Préservation et d'Accès aux Données Nouvelle Génération)
 - generic tool for data management, which can operate in different contexts. It implements the full OAIS model.



■SIPAD-NG is a generic tool for data archive management. It implements the full OAIS model

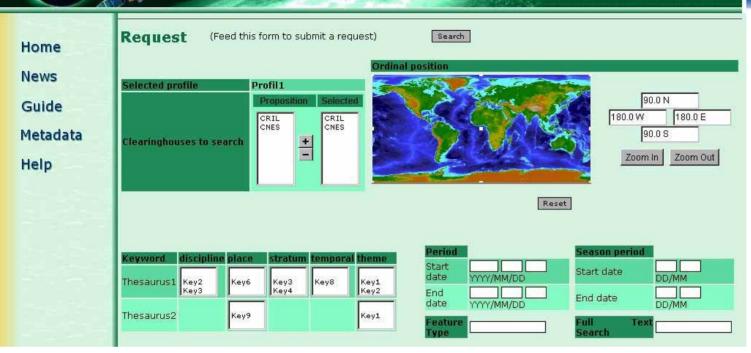


SIP = Submission Information Package / Paquet d'information à verser

AIP = Archival Information Package / Paquet d'informations archivé

DIP = Dissemination Information Package / Paquet d'informations diffusé





- ■BDMS is a tool which aims at allowing users to discover that a datasets of its center of interest exists.
 - This tools is based on the ISO19115 standards.
 - This standards is dedicated to Geographic Information.
 - It will then be profiled and enriched in order to be able to fit the needs of all thematic.
 - This metadata model enrichment will be based on the thematic classification and interview work

Definition of a minimal metadata model

■ This work includes the definition of:

Classification

• The objective is to define the centers of interest which are representative, discriminating and consistent to the terminology of the future users. This required a precise classification job of all the science thematics. This has been done thanks to the contribution of all CNES thematic responsible from the CNES Programmes Direction. This work is also done with the help of the C-S company

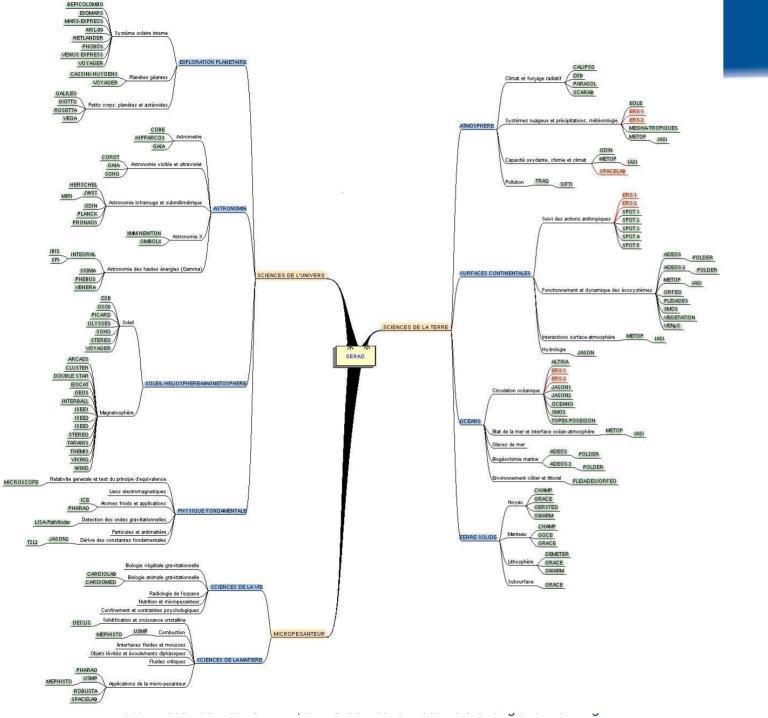
Thesaurus

- The goal is to have an ordered list of keywords recognized by communities of users in order:
 - To speak a common language understandable to those communities,
 - to be interoperable with other systems, precisely target datasets.
- This work is conducted in parallel with the classification of science thematics. It is a difficult
 exercise which will be expanded over time, data sets that will be introduced gradually in the
 SERAD, and from elements collected so far (thesaurus CNES, ...). This work is conducted with
 the support of CS-SI

Exhaustive list of missions

• The goal is to have an exhaustive list of CNES missions, and to be able to decide what to do with the datasets. These missions are also integrated in the classification tree

Classification



13



Open points / to do

- **■**Choice for the tools and their integration
- **■CEOS**: ho to be referenced by IDN through harvesting
- **■**Finalization of CNES data policy
- **■Planning**
 - Opening in early 2012
 - + In 2009/2010
 - Choice of the tools & system implantation
 - +2011
 - SMOS: Qualificatin