

living planet symposium | BONN 23–27 May 2022

TAKING THE PULSE
OF OUR PLANET FROM SPACE



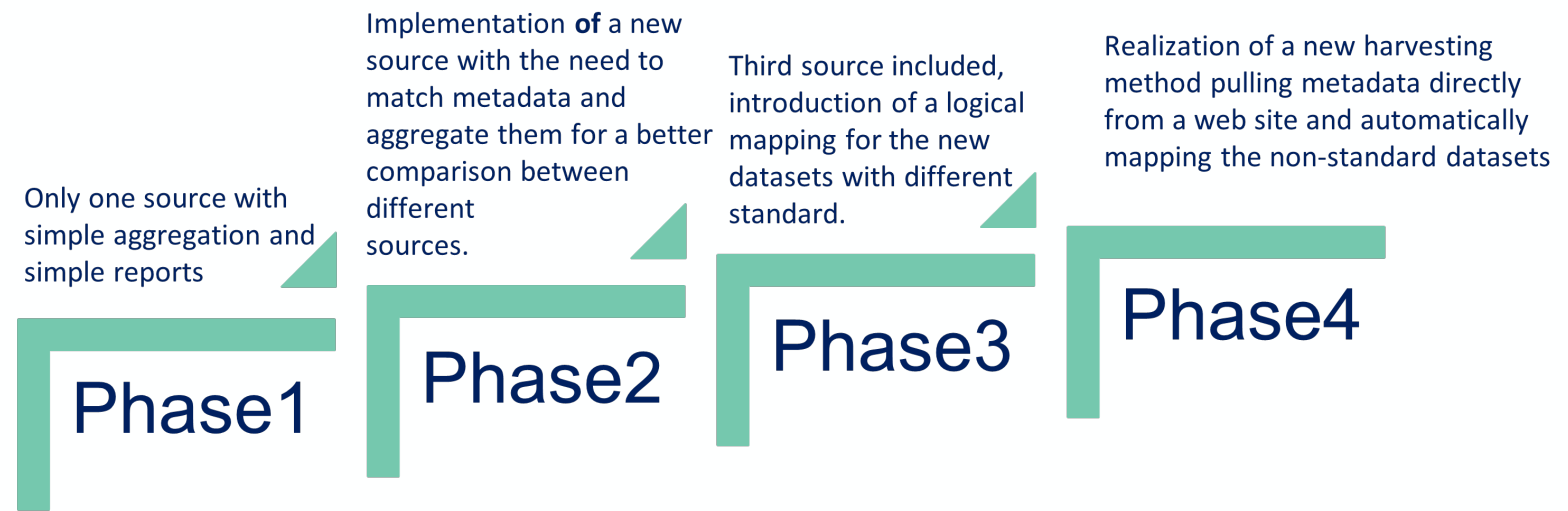
Earth Observation Data Information Service – EO DIS

Alessandra Paciucci
alessandra.paciucci@serco.com

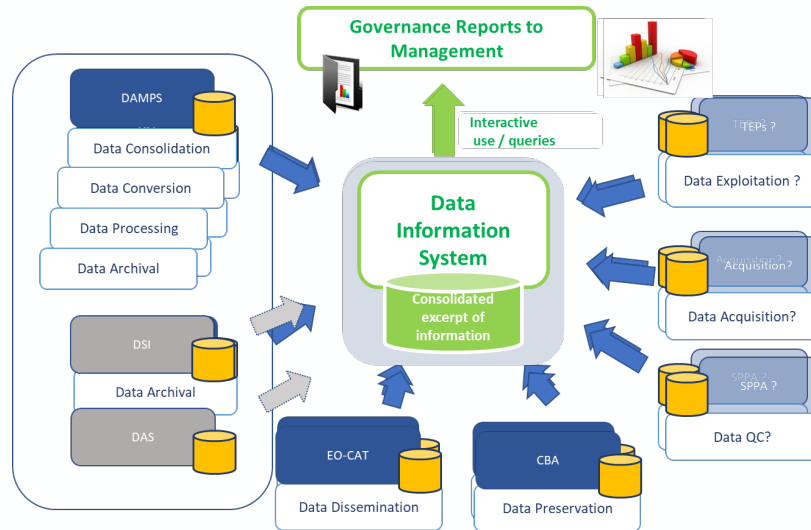
Authors: S. Papaleo, A. Paciucci, P. Boezi, R. Campbell, D. Castrovillari, G. Colamussi,
M. Douzal, S. Garofalo
24/05/2022

- ESRIN, the European Space Agency's for Earth Observation has commissioned a centralized service to collect and store metadata relevant to Earth Observation data as generated in several systems and services running in the ESA Ground Segment.
- The service was designed, developed and operated inside the DSI (Data Service Initiative) contract led by Serco.
- The Data information Service (DIS) purpose is :
 - ✓ Centralise the information on the data produced by the various ESA services standardizing the handling of metadata
 - ✓ Provide a centralised end-to-end information on the data life-cycle managed by ESA Earth Observation
 - ✓ Monitor the data information across different ESA teams.
 - ✓ Give visibility of the Operations and Processes performed on the data
 - ✓ Provide Business Intelligence capabilities for data governance and data administration

- How the service has evolved through its various phases
 - The need of DSI Data management (Pilot)
 - Service approach requirements with only 1 source of metadata (Phase 1)
 - Additional source + matching metadata requirements (Phase 2)
 - Additional source + requirements for datasets with different standard (Phase 3)
 - Additional source + harvesting and automatization of mapping not standard datasets



- Service provides standard interfaces to allow data to be collected from additional external sources.
- Metadata extraction from several source systems and services and collection for every dataset and data product into a centralized system.
- All changes to the data to be managed and traceable, similar to any other items under standard configuration control.



Generic ICD

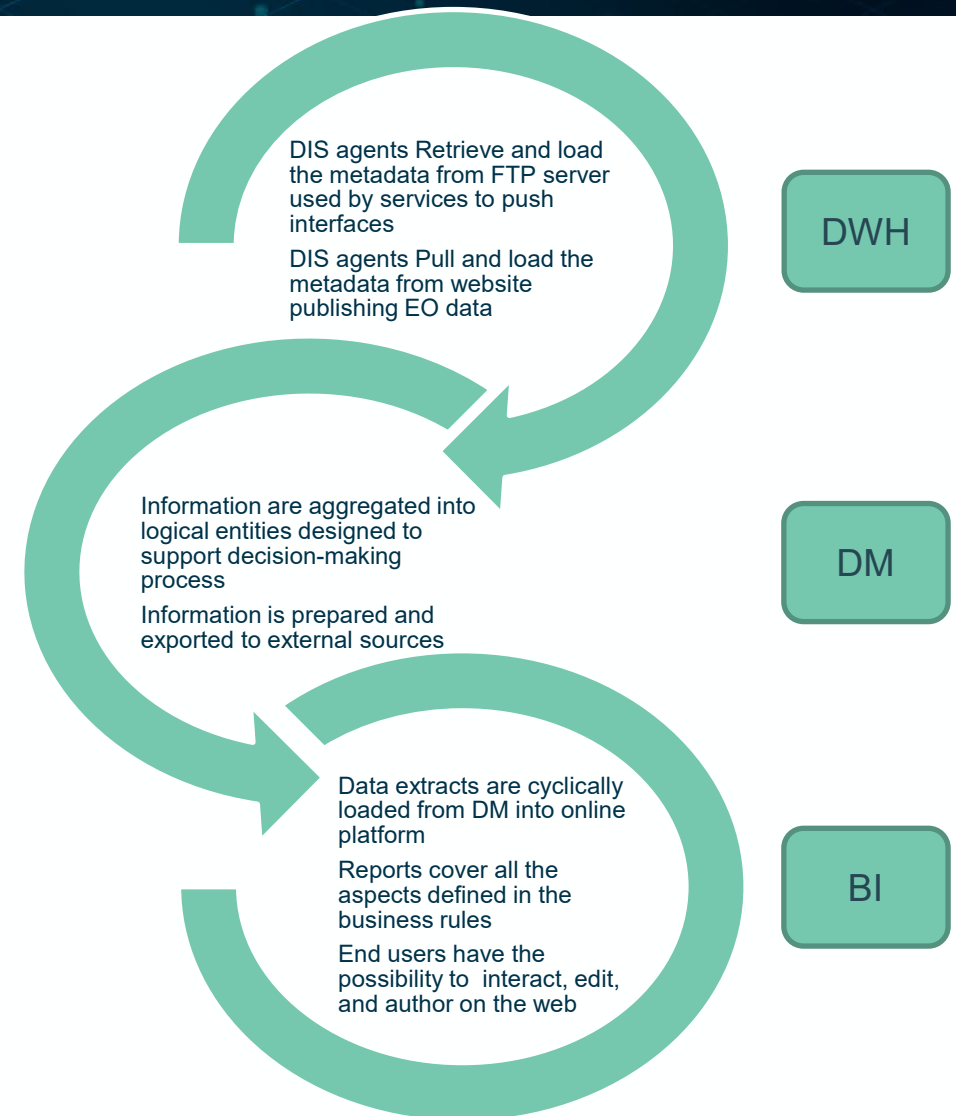
- ✓ Which data shall be provided by sources and how are packaged
- ✓ XSD schema definition for the xml files
- ✓ Error handling
- ✓ Communication rules

Custom ICD

- ✓ Additional product information provided by source not included in the generic ICD
- ✓ Customization of EO-DIS procedures (schedule, naming requirements etc.)
- ✓ Specific communication rules
- ✓ Any other variation related to the external system characteristics

- The service ensures that data assets can be traced in any system that is federated to EO DIS – therefore providing a unified view of multiple archives and copies of the asset

- Data warehouse (DWH) – a staging area used to store information from different sources identifying data unequivocally across missions & services
- Data mart (DM) - containing the aggregated information modelled to be ready for the presentation layer
- Business Intelligence platform (BI) – based on Tableau online cloud. It is possible to access data on the web using customized reports tailored to the end-user needs.



➤ DIS platform is designed to be used in various ways:

- ✓ Reports : User driven reports are available in online platform allowing end-users to intuitively visualize, explore and discover aggregated data
- ✓ Queries : Customized queries upon specific needs
- ✓ Exports: Aggregated data can be exported for any further usage in several formats (CSV, XML, etc...).

- The front-end relies on Tableau online platform which provides a Software-as-a-Service model improving productivity in terms of efficiency, scalability and accessibility. The platform offers an intuitive user interface and a self-service approach.
- The main features of EO DIS BI solution are:
 - **Decision Making** - customized reports allow end-users to go deeply through satellite metadata and make data-driven decisions.
 - **Data Tracking**: it is possible with some reports to keep track of the information among all the external sources describing relationships between groups of data and across different environments.
 - **Data Mining**: ad-hoc report can be used to uncover trends and to classify data through categories

EO DIS User Front-end – General dashboards examples



Data Information System

Ingestion Status

DAMPS	Datasets : 1,258 Products : 67,201,809 Refresh Date : 04/05/2022
CBA	Datasets : 549 Products : 32,227,597 Refresh Date : 29/04/2022
DISS	Datasets : 101 Products : 16,464,740 Refresh Date : 09/03/2022

Operational Information

Operations are nominal

Missions

ALOS ADEOS UK-DMC WindSat Deimos

Landsat **SMOS**

Cryosat

Envisat

ERS Swarm COSMO-SkyMed ESACampaign OceanSat QuikSCAT TERRA QuikSCAT SeaWiFS

Reports Last Extracts Publication : 5/7/2022 5:04:59 PM

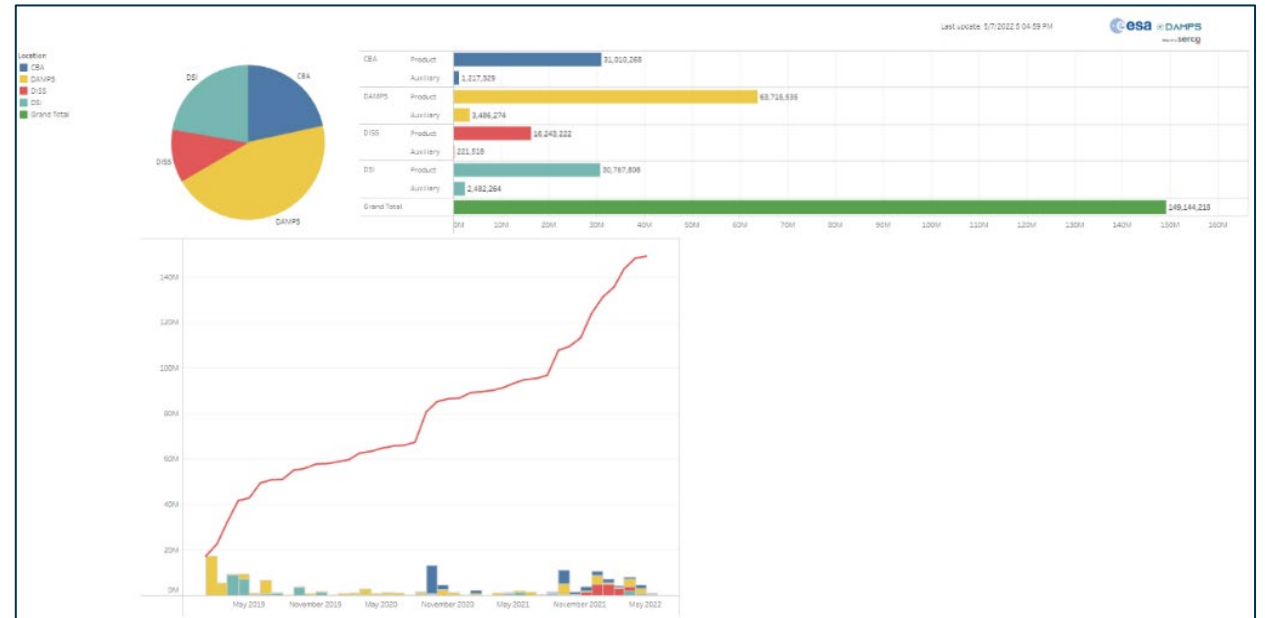
Management Reports

Operations Reports

Administration Reports

For any queries, contact Us : eodis@damps.info - For any issues, contact TellUs

EO DIS Welcome page

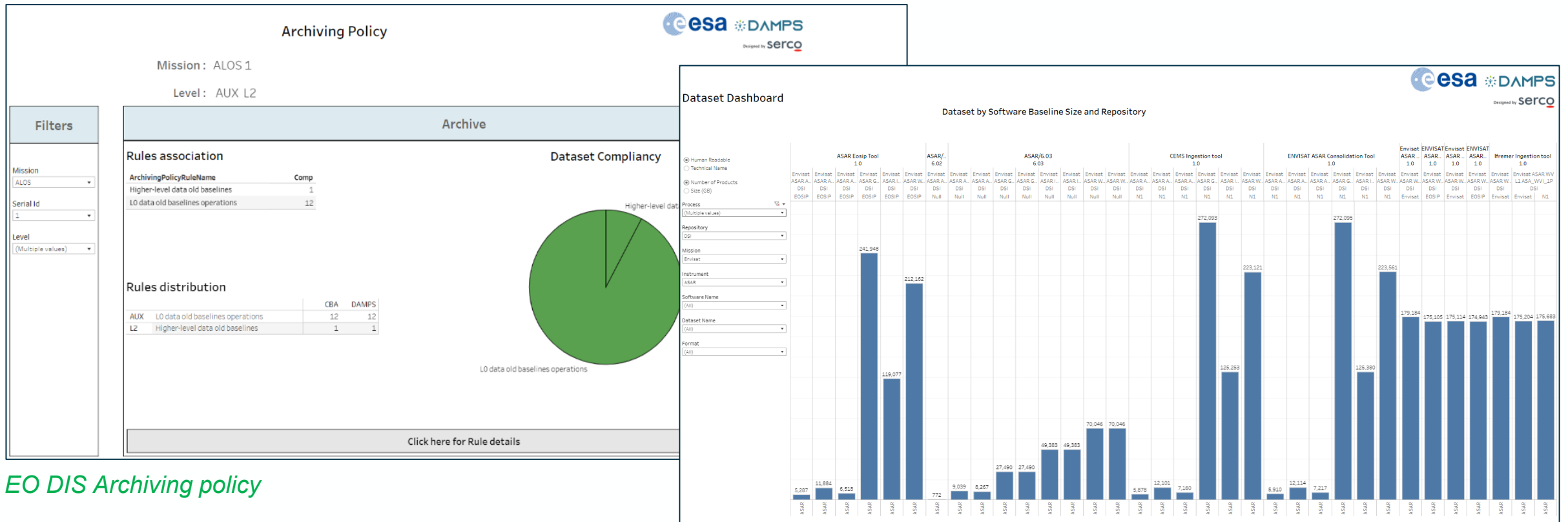


EO DIS Administration report example

EO DIS User Front-end – Decision Making examples



- The stakeholders include mostly Management, which DIS assists with decision-making, and Operations, which are supported by a detailed and accurate account of the location, status and cross-relationship of the data.

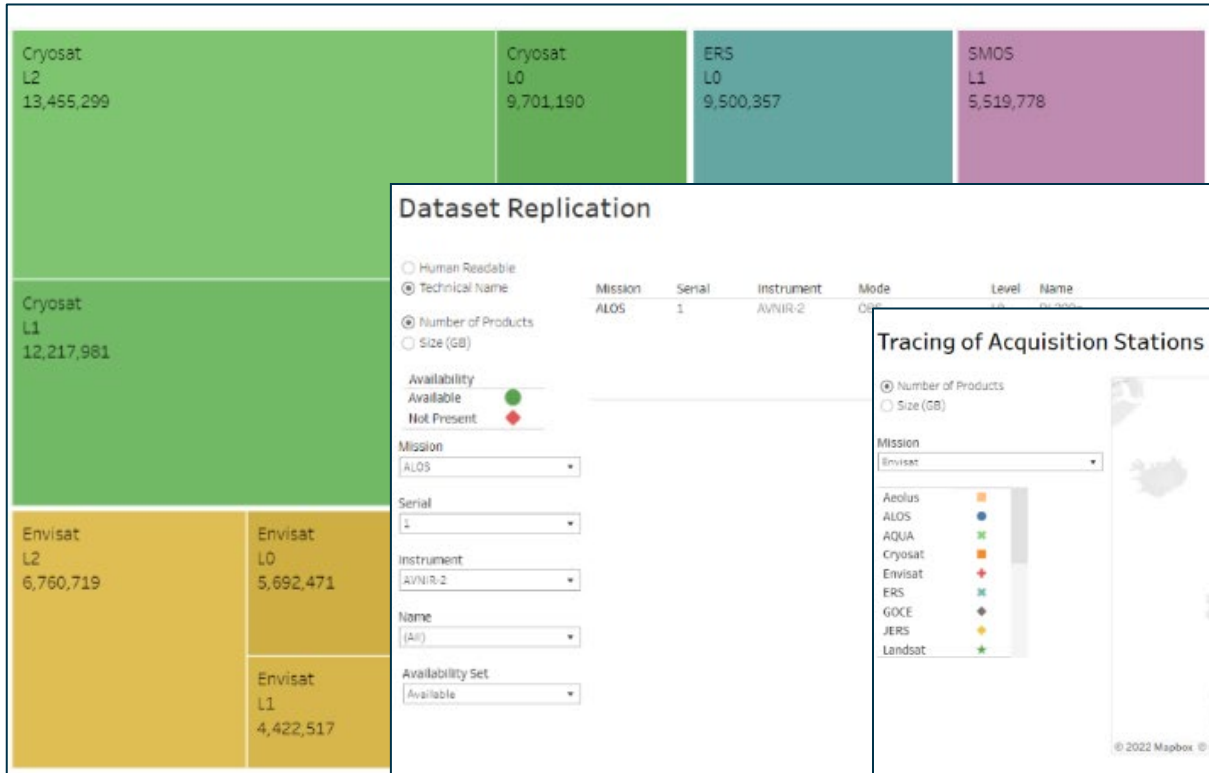


EO DIS Archiving policy

EO DIS Dataset Overview



EO DIS User Front-end – Data Tracking examples



EO DIS Data Holding

Dataset Replication

Human Readable
 Technical Name
 Number of Products
 Size (GB)

Availability: ● Available / ● Not Present

Mission: ALOS

Serial: 1

Instrument: AVNIR-2

Name: (All)

Availability Set: Available

Mission	Serial	Instrument	Mode	Level	Name	Version	Repository
ALOS	1	AVNIR-2	Orb	L0	AVNIR-2	1.0.0	DIS

EO DIS Dataset Replication

Tracing of Acquisition Stations

Number of Products
 Size (GB)

Mission: Envisat

- Aeolus
- ALOS
- AQUA
- Cryosat
- Envisat
- ERS
- GOCE
- JERS
- Landsat

Acquisition Station	Country	Station	Operator	Volume
ESR	Italy	Pracati	ESRIN	517,808
MTI	Italy	Matera	Matera	27,323
NSG	Germany	Neustrelitz	DLR	2
PDE	Italy	Pracati	ESRIN	353,645
PDHS-E	Italy	Pracati	ESRIN	497,596
PDHS-E,PDHS-K	Italy	Pracati	ESRIN	1,620
PDHSE	Italy	Pracati	ESRIN	54

Product Acquisition and Cumulative Total

EO DIS Tracking of Acquisition Station



EO DIS User Front-end – Data Mining examples



Added Value Timeline

Number of Products
 Size (GB)

Process
 Archival
 Collection
 Consolidation
 Processing
 Repackaging

Process
 (All)

Mission
 Envisat

Serial
 (All)

Mission
 Envisat

Mission Statistics per Product Type

Sensing Stop
 Fact Date

August 1, 2018 – November 19, 2019

Size (GB)

1,499.89

Products

76,245

Repository: DAS

Mission: Aeolus

Instrument: ALADIN

Product Type: (All)

Mission	Instrument	Product Type	Number of Records	Size (GB)
Aeolus	ALADIN	ALD_U_N_0_	7,260	203.3329
Aeolus	ALADIN	ALD_U_N_1A	7,717	236.5461
Aeolus	ALADIN	ALD_U_N_1B	7,715	322.6264
Aeolus	ALADIN	ALD_U_N_2A	6,972	49.6048
Aeolus	ALADIN	ALD_U_N_2B	7,678	95.3196
Aeolus	ALADIN	ALD_U_N_2C	5,516	76.9105
Aeolus	ALADIN	AUX_CAL_L2	4	0.1144
Aeolus	ALADIN	AUX_CHAR_	5	0.0006
Aeolus	ALADIN	AUX_CLM_L2	3	0.0008
Aeolus	ALADIN	AUX_CSR_1B	67	0.0274
Aeolus	ALADIN	AUX_DCC_1A	73	0.1452
Aeolus	ALADIN	AUX_DCC_1B	73	0.0336

EO DIS Added Value Timeline

EO DIS Product Type Overview

EO DIS Aeolus Mission Statistics per Product Type



- As the amount of data managed increases, the need to maintain knowledge of the data becomes more challenging. DIS is a helpful tool for retaining visibility and understanding of these EO data assets, as well as for comparing the contents of different archives to always ensure that datasets are fully identified and aligned, Detection of misalignments in archives trigger immediate resolution of such issues.

Dataset Differences

Dataset	Version	ALOS	AUX	Dataset ID	Version	Status
DL210a	1.0.0			AUX_COI_DT	DL236b	1.0.0 ✓
				AUX_COI_PD	DL236a	1.0.0 ✓
				AUX_COI_SF	DL236c	1.0.0 ✓
				AUX_COI_SP	DL236d	1.0.0 ✓
				AUX_CTM_AX	DL236e	1.0.0 ✓
				AUX_PAD_AX	DL236f	1.0.0 ✓
				AUX_PDI_AX	DL236g	1.0.0 ✓
				AUX_PDI_AX	DL236h	1.0.0 ✓
				AUX_TWD_AX	DL236i	1.0.0 ✓
				AUX_TWD_AX	DL236j	1.0.0 ✓
Mission: ALOS				AUX_COI_DT	DL236b	1.0.0 ✓
				AUX_COI_PD	DL236a	1.0.0 ✓
				AUX_COI_SF	DL236c	1.0.0 ✓
				AUX_COI_SP	DL236d	1.0.0 ✓
				AUX_CTM_AX	DL236e	1.0.0 ✓
				AUX_PAD_AX	DL236f	1.0.0 ✓
				AUX_PDI_AX	DL236g	1.0.0 ✓
				AUX_PDI_AX	DL236h	1.0.0 ✓
				AUX_TWD_AX	DL236i	1.0.0 ✓
				AUX_TWD_AX	DL236j	1.0.0 ✓
Dataset Name: AVNIR-2				ALL_AV2_SF	DL003a	2.0.2 ✓
				AV2_OBS_SF	DL209a	1.0.0 ✓
				AV2_OBS_IC	DL244c	1.0.0 ✓
				FBS_DWL_SF	DL209a	1.0.0 ✓
				FBS_DWL_SF	DL346c	1.0.0 ✓
				PLR_DWL_SF	DL209c	1.0.0 ✓
				PLR_DWL_SF	DL346a	1.0.0 ✓
				PSR_WB1_1S	DL300a	1.0.0 ✓
				WB1_DWL_SF	DL346d	1.0.0 ✓
				PSM_DWL_SF	DL210a	1.0.0 ✓
Dataset Name: PALSAR				FBS_DWL_SF	DL209a	1.0.0 ✓
				FBS_DWL_SF	DL346c	1.0.0 ✓
				PLR_DWL_SF	DL209c	1.0.0 ✓
				PLR_DWL_SF	DL346a	1.0.0 ✓
				PSR_WB1_1S	DL300a	1.0.0 ✓
				WB1_DWL_SF	DL346d	1.0.0 ✓
				PSM_DWL_SF	DL210a	1.0.0 ✓
				PSM_DWL_SF	DL210a	1.0.0 ✓
				PSM_DWL_SF	DL210a	1.0.0 ✓
				PSM_DWL_SF	DL210a	1.0.0 ✓

Provenance

Dataset: (All)

- ALOS AVNIR-2
- ALOS PRISM
- CRYOSAT DORIS
- CRYOSAT SIRAL Calibration
- CRYOSAT SIRAL LO Consolidation
- CRYOSAT SIRAL Processing
- CRYOSAT STARTRACKER Calibration
- ENVISAT ASAR
- ENVISAT Meris
- ENVISAT MIPAS
- ENVISAT RA2/MWR**
- Landsat ETM
- Landsat MSS
- Landsat TM

The diagram illustrates the provenance of data for the ENVISAT RA2/MWR dataset. It shows a network of nodes representing datasets and processes. Key nodes include DL106a18, DL317a, DL183a, DL301a, and DL301b. Processes such as Consolidation, Processing, and Archival are shown as connecting lines between these datasets. The diagram is color-coded by mission: ALOS (purple), AVNIR-2 (green), PALSAR (orange), and PRISM (red).

EO DIS Dataset Differences

EO DIS Provenance

- The service provides a consolidated view of all data holdings throughout all federated systems
- Lending itself to providing a unique operational and management asset monitoring and control
- As the age of data increases and new data is generated from multiple sources the task of maintaining control of the data assets becomes more challenging
- The service provides a means to preserve the complete information of the product along its lifetime including all the references to the various preservation elements
- The service is fully scalable and can also be used for management and operational support as well as a pure asset management tool

Thanks

