LSA Data Center:
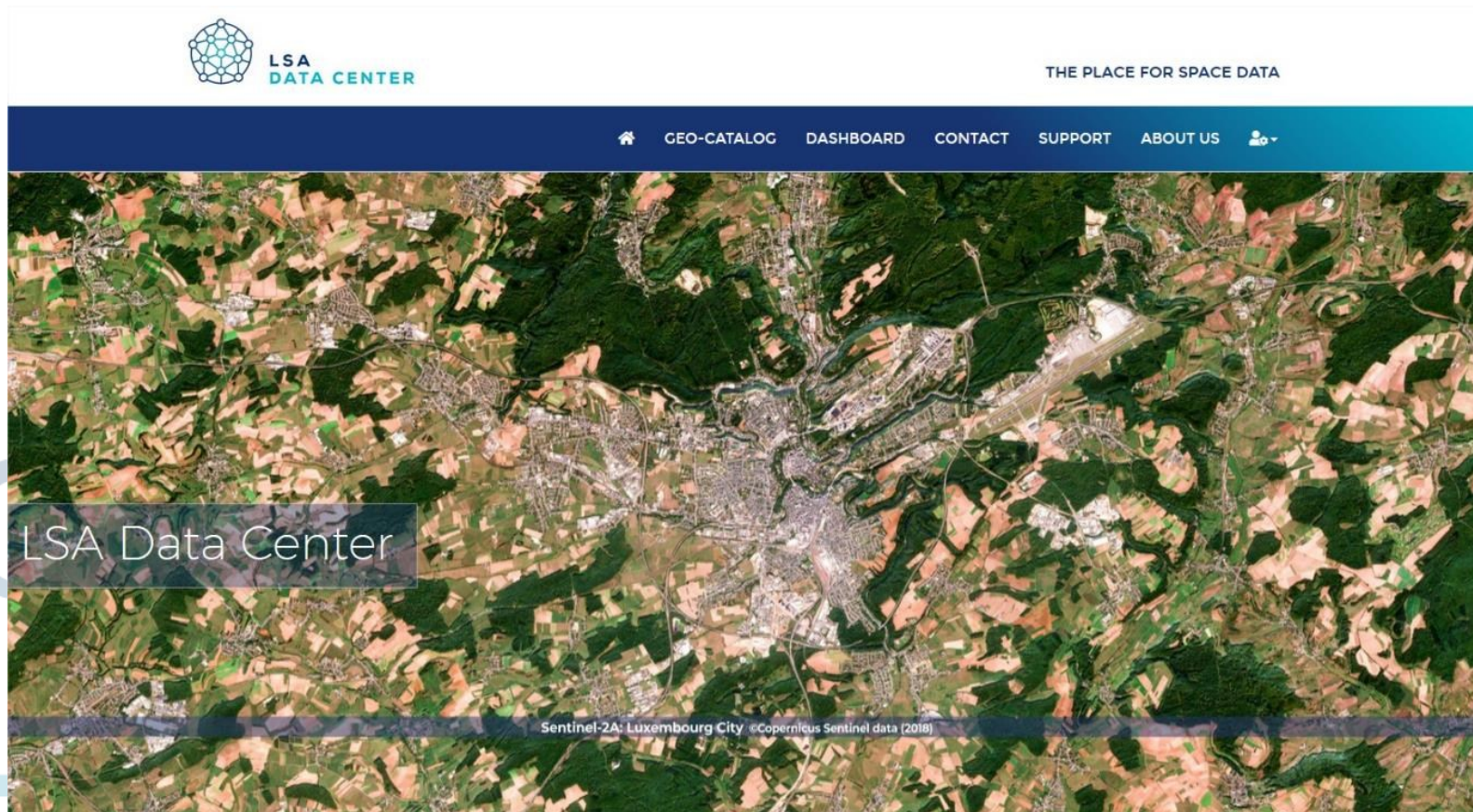# The Luxembourg Entry point for Sentinel Data

Giovanni Corato

Luxembourg Space Agency Data Center

LSA Data Center

➢ The LSA Data Center is a cornerstone of the Luxembourg Space Agency

➢ The LSA Data Center is the Luxembourg entry point to data products of the Copernicus Sentinel constellation.

➢ The LSA Data Centers' mission is to accelerate the development of the downstream sector of value-added applications and services using these data by facilitating the access to the Sentinel data.

➢ The LSA Data Center allows user to search, discover and download the data.

Small country big CGS

www.lsa-datacenter.lu

56.53 PB
volume of data

78.21 M
n. of products

➢ **Full collections of Sentinel-1 and Sentinel-2 products without time or spatial restrictions.**

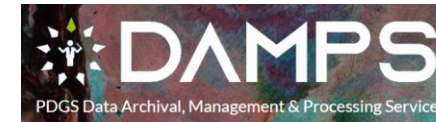➢ **All products are available online with a backup copy on tape in a second datacenter**

As today one of the few repositories having complete collections of Sentinel 1 and 2 data

LSA-DC archive as 1st Mar 2025

EO Data Vault

LSA DATA CENTER

56.53 PB
volume of data

78.21 M
n. of products

Sentinel-1 and Sentinel-2 CGS

DAMPS
PDGS Data Archival, Management & Processing Services

11.21 PB
volume of data

103.4 M
n. of products

Master Archive for Earth Explorer Data and reprocessing center

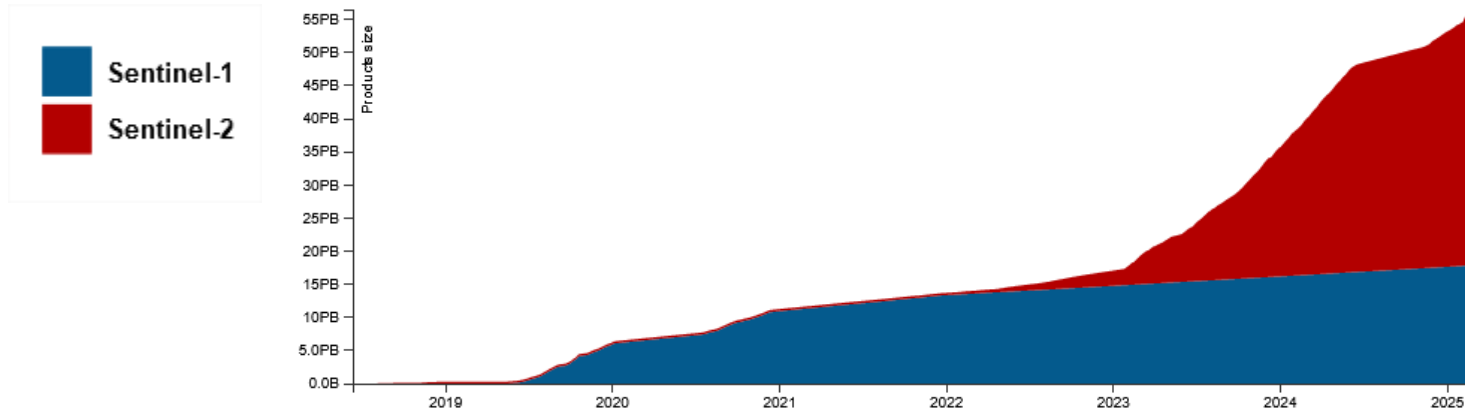**HIGHWAY EO DATA SPACE**

Mirror site and production center

Key points

1. **Migration to the cloud platform Terra Adwäis**

2. **Phase 2 Sentinel 2 Reprocessing**

3. **Since October 2023 LSA DC integrated in the CDSE offering DAD for some Sentinel 1 and 2 collections (~17 PB).**

4. **Refactoring data catalog and data access in progress**

Data repatriation

Archived products

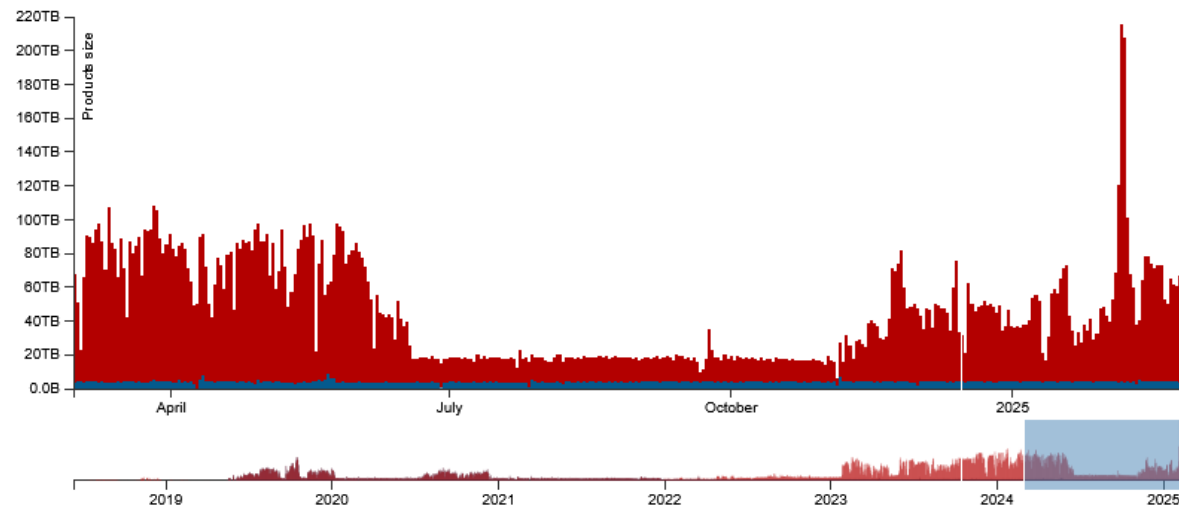## 56.5 PB   78.2 M

volume of data   n. of products

Downloaded products

75.7 PB   115.8 M

volume of data   n. of products

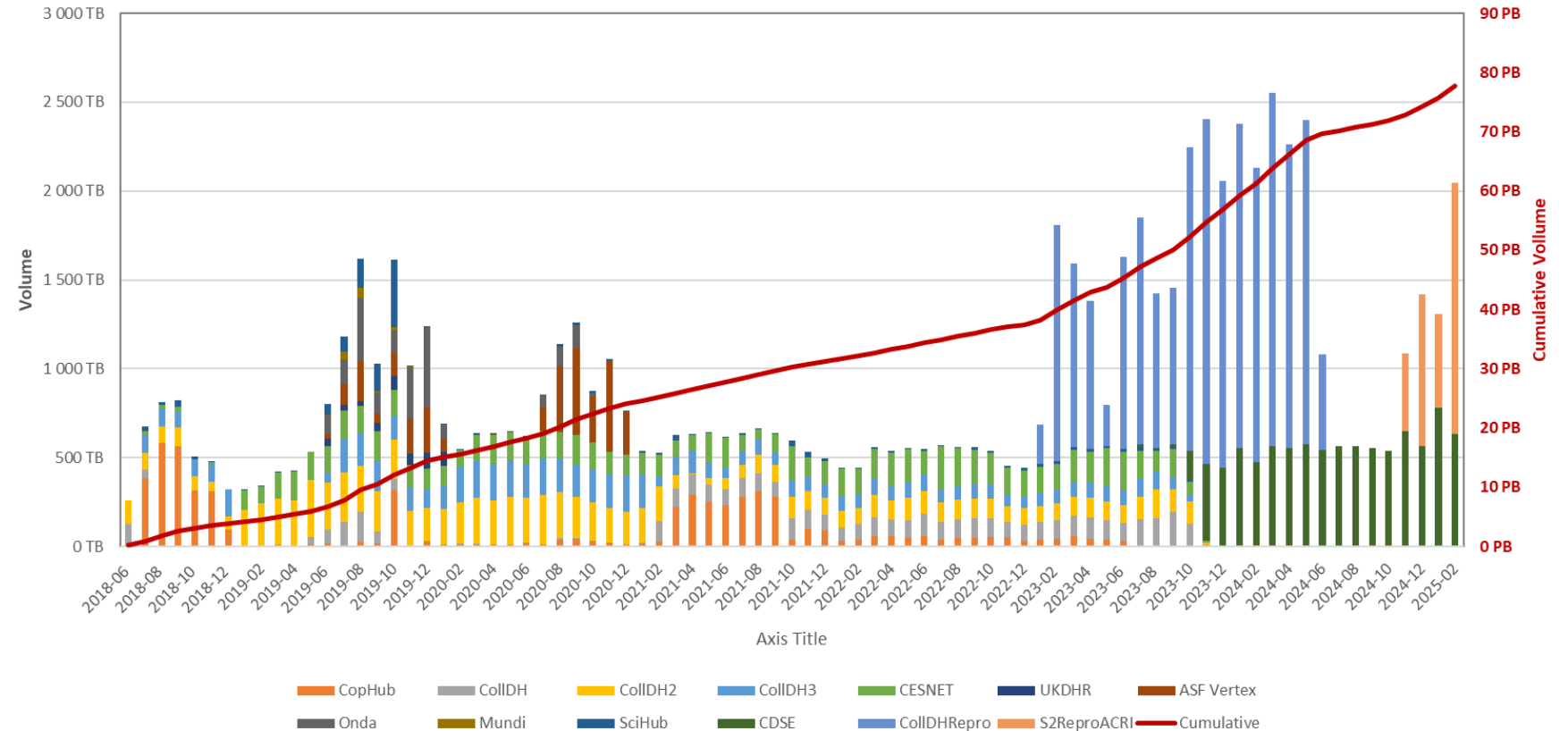**Dramatic increase of download rate due to Sentinel 2 reprocessing**
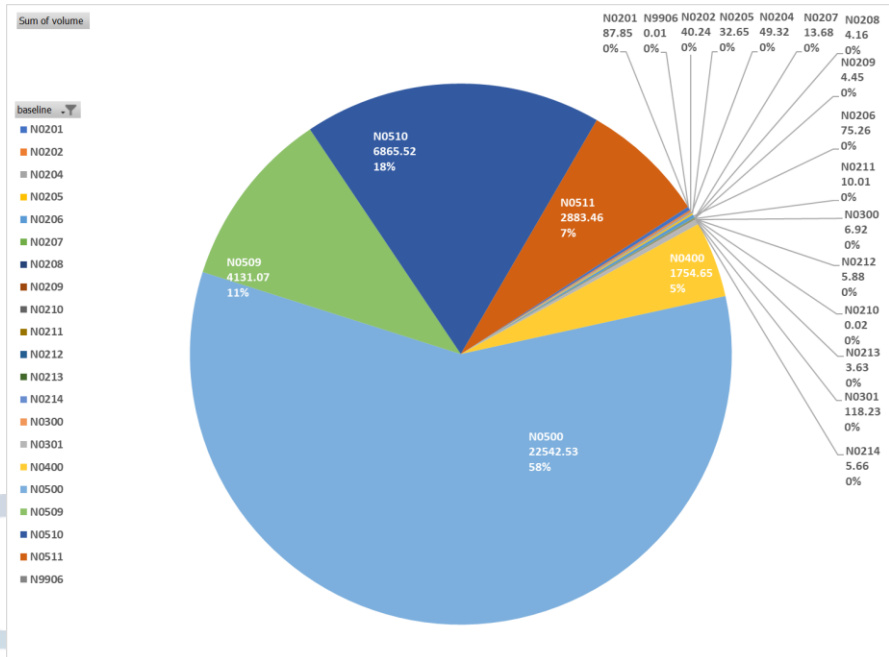
Data repatriation

LSA DATA CENTER

Repatriation Sentinel 2
Collection 1

➢ Jan 2023-June2024 Phase1

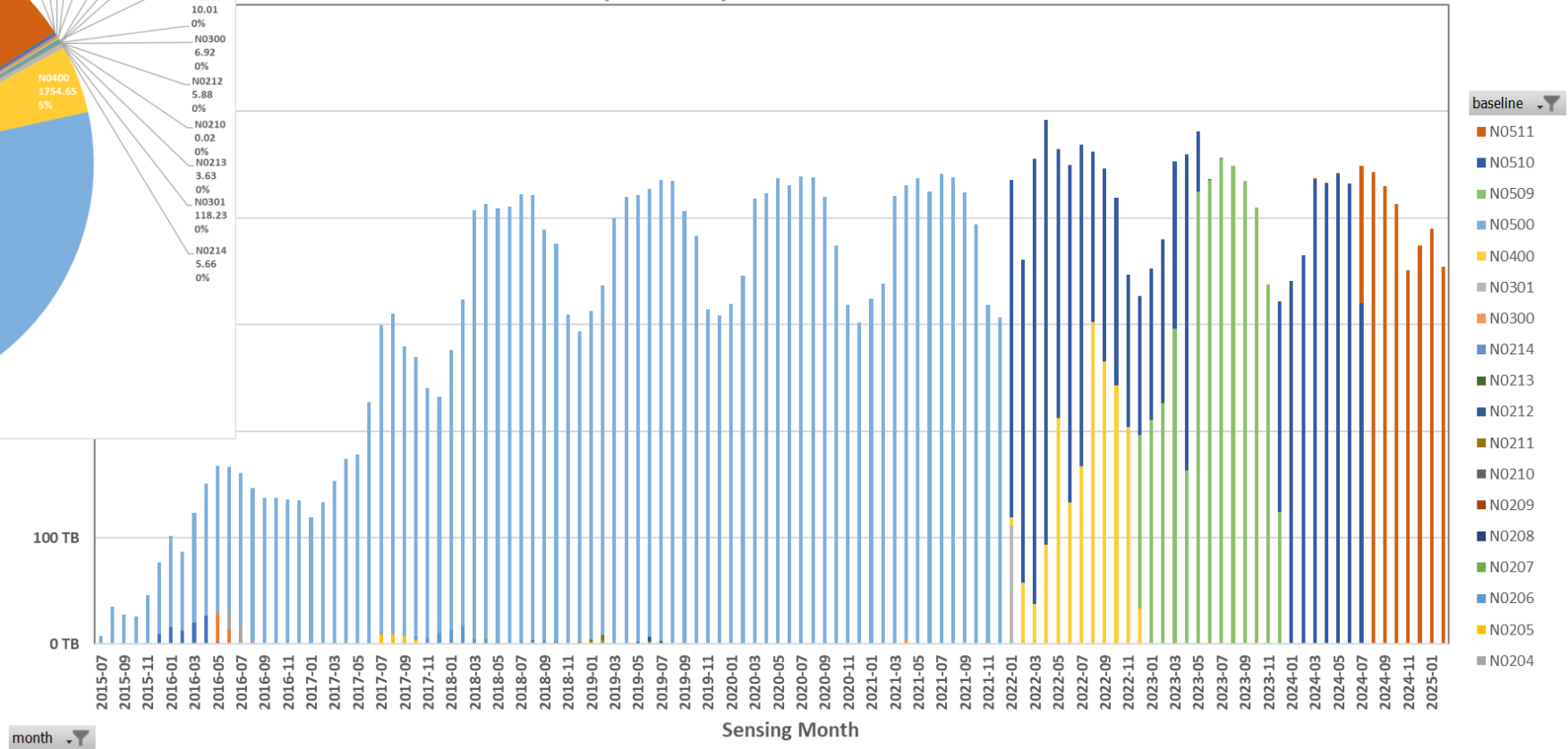➢ Since Nov 2024 Phase2

Monthly downloaded volumes

Legend: CopHub, CollDH, CollDH2, CollDH3, CESNET, UKDHR, ASF Vertex, Onda, Mundi, SciHub, CDSE, CollDHRepro, S2ReproACRI, Cumulative

Sentinel-2 baseline composition

Monthly volume per baseline

Sentinel-2 baseline composition

Not all the products with baseline > N0301 were replaced by the Phase1 Reprocessing: **343.9 TB with baseline < N0301**

Sentinel-2 reprocessing challenges

➢ Increase of download rate, but data flow kept internal to the cloud

➢ Backup and recycling of tapes:

  ✓ when data are erased from a tape the space is wasted if at least a segment remain in the tape. Recovery of wasted space using native tools is very long procedure

  ✓ Separated tape storage policies per product type, implemented in 2021 mitigated the impact

  ✓ The disk copy allowed to recover wasted space in a smart and quick way without increasing the service cost

Migration to cloud

Migration to the cloud platform Terra Adwäis

➤ Still running in TIER IV datacenters in Luxembourg.

➤ Optimization of egress/ingress:
  - ✓ Phase2 Reprocessed data repatriating internally to the cloud
  - ✓ Some of the clients of the LSA DC such as Sentinel 2 Global Mosaic run in the same cloud.

Same datacenter of MeluXina, the Luxembourg HPC, and the upcoming MeluXina-Q and AI Factory.

> Satellite products are data and data to be turned into valuable information need to be processed. Luxembourg invests also in processing and having the data close to HPC and AI Factory facilities is an opportunity.

Ongoing activities

Refactoring Data catalog:

✓ Extensive usage of table partition to improve database performances

✓ Refactoring of geoserver on going, portioning by product type allows to speed up the performances of the OpenSearch catalog but they are margin to further improve wms service performances.

✓ Limit cross collection search for OData I/F to maximize the benefit of partitioning.

✓ STAC catalog under development, here the challenge is to optimize the computing resources and guarantee a good level of performance.

Refactoring Data access

✓ Legacy FTP I/F to be dismissed, access full relaying on https.

Analysis of impact migration from SAFE to EOPF

From SAFE to EOPF

Challenges:

➢ Migration: data are on disk but still migrate 60 PB looks challenging, test with converter planned for March.

➢ Data in disk storage can be easily replaced but for data on tapes (i.e. backup) data replacement can problematic.

➢ Dramatic increase of items to be repatriated and preserved if data not packed (but EOPF data model foresees the possibility of pack the product in a zip container)

➢ Modification of catalog

➢ Modification of data access services

Conclusion

Duplicate 60 PB of data is necessary?

- LSA DC today contributes to secure the Copernicus data: one more copy may be a security;

- Having data close to processing facilities such as HPC, AI factory and Quantum computing is an opportunity in the era of AI;

Questions