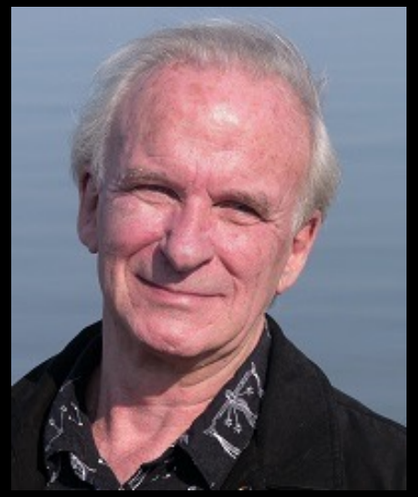


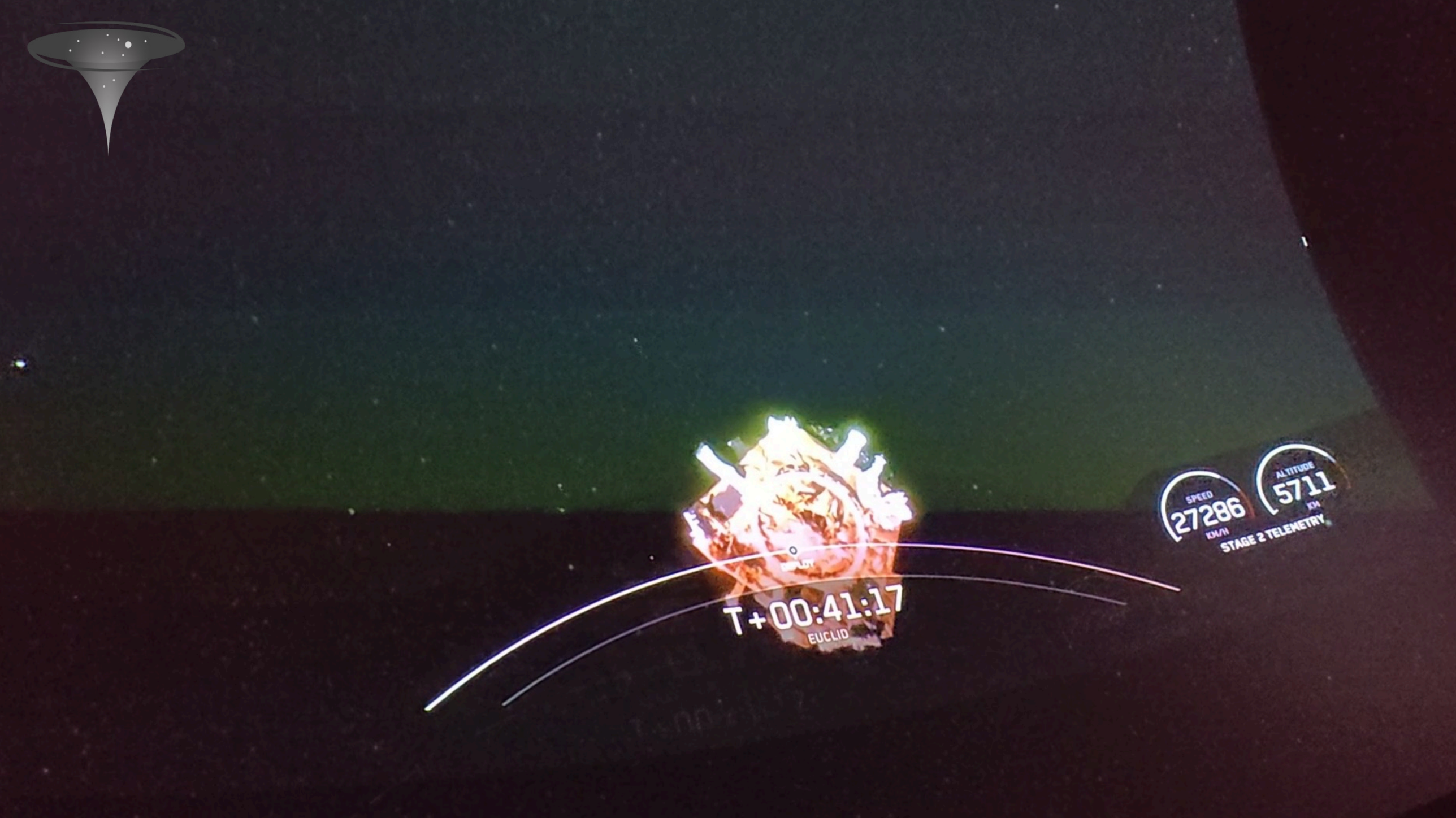
The Euclid NL- Science Data Center

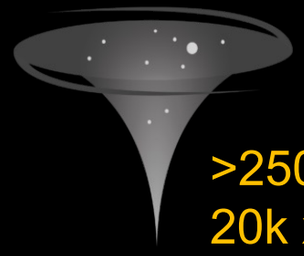
Edwin A. Valentijn
Rijks Universiteit Groningen - Kapteyn Astronomisch Instituut

NL – Euclid SGS Project manager

Big Data processing in Astronomy and Earth Observations
23 jan 2024 – SURF/NSO- Utrecht

T+00:00:08


SPEED
27286
KM/H
ALTITUDE
5711
KM
STAGE 2 TELEMETRY
T+00:41:17
EUCLID



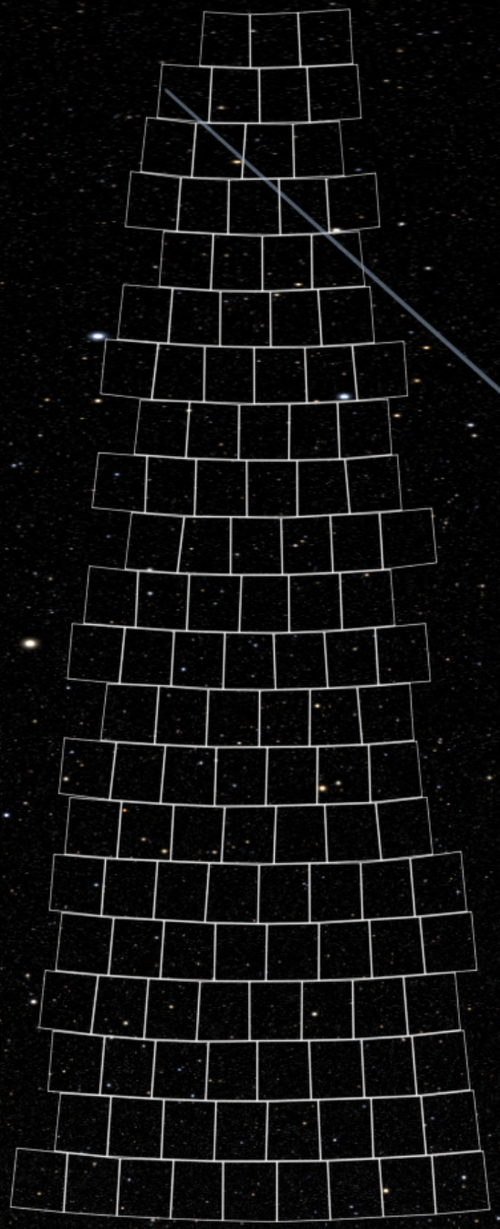
>250 days Hubble Legacy Field
20k x 20k
265.000 galaxies

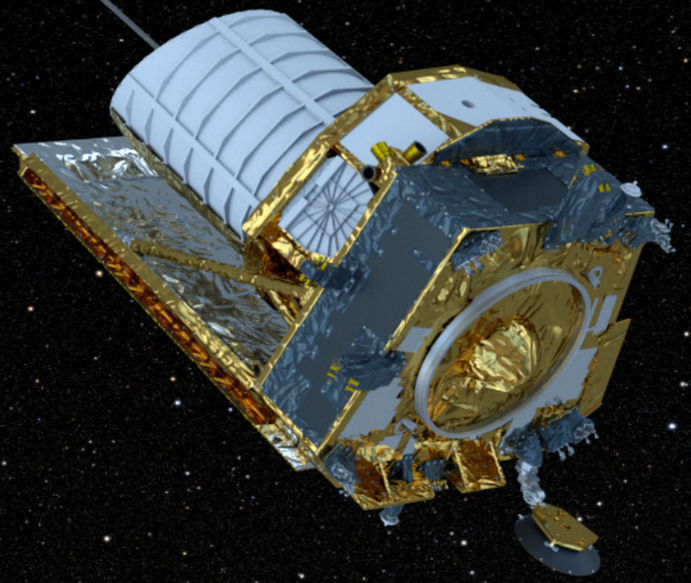
30’

4200 sec nominal 45’ x 45’ 0.57 Sq deg
VIS: 4x 565 sec + 2x100s 0.1”/pix 24k x 24 k 609 Mpix
NISP: 565 sec spectra + Y, J, H 112 sec
300.0000 galaxies
50.000 distances

30’

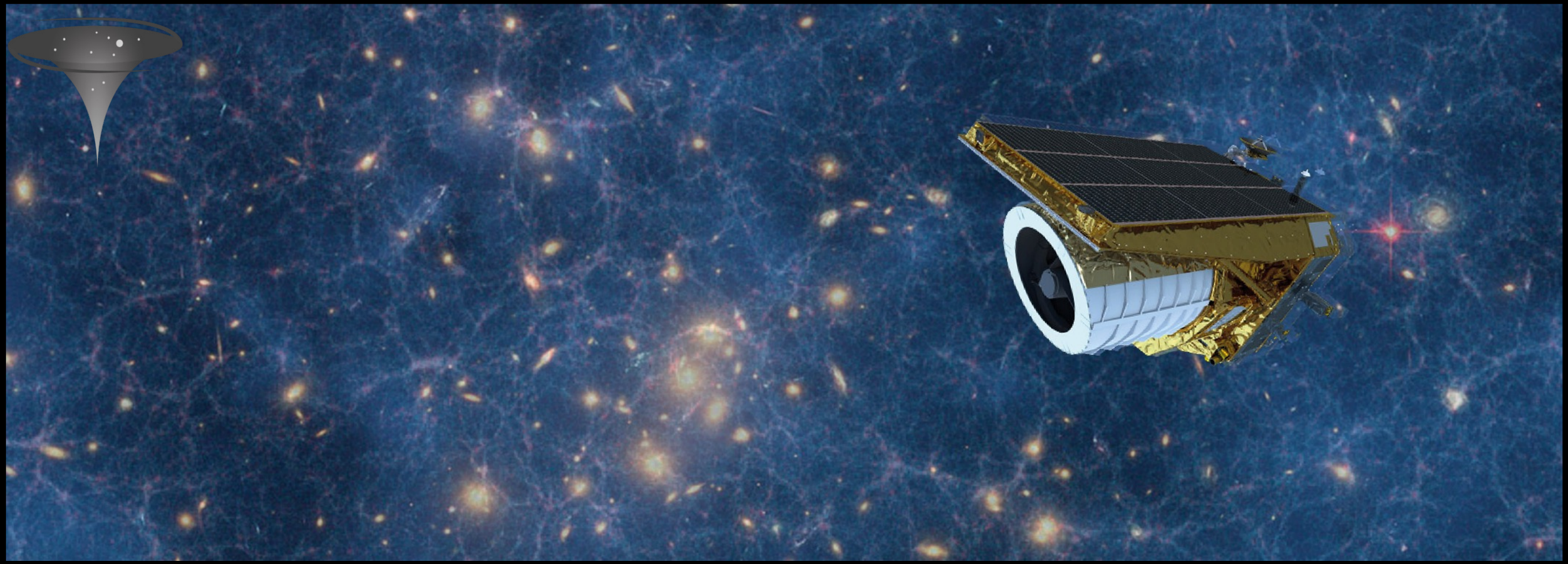
Credit: Hubble Legacy Field Image: NASA, ESA, and G. Illingworth and D. Magee (University of California, Santa Cruz); Moon Image: NASA, Goddard Space Flight Center and Arizona State





- 15.000 Sq deg Wide
- 50 Sq deg - 3 deep fields
- calibrations

  6 year



Cosmology :

Evolution $z < 2$ last 10 Billion year - 3 probes:

- Weak lensing tomography ~10-12 slices
- BAOs
- Large scale structure

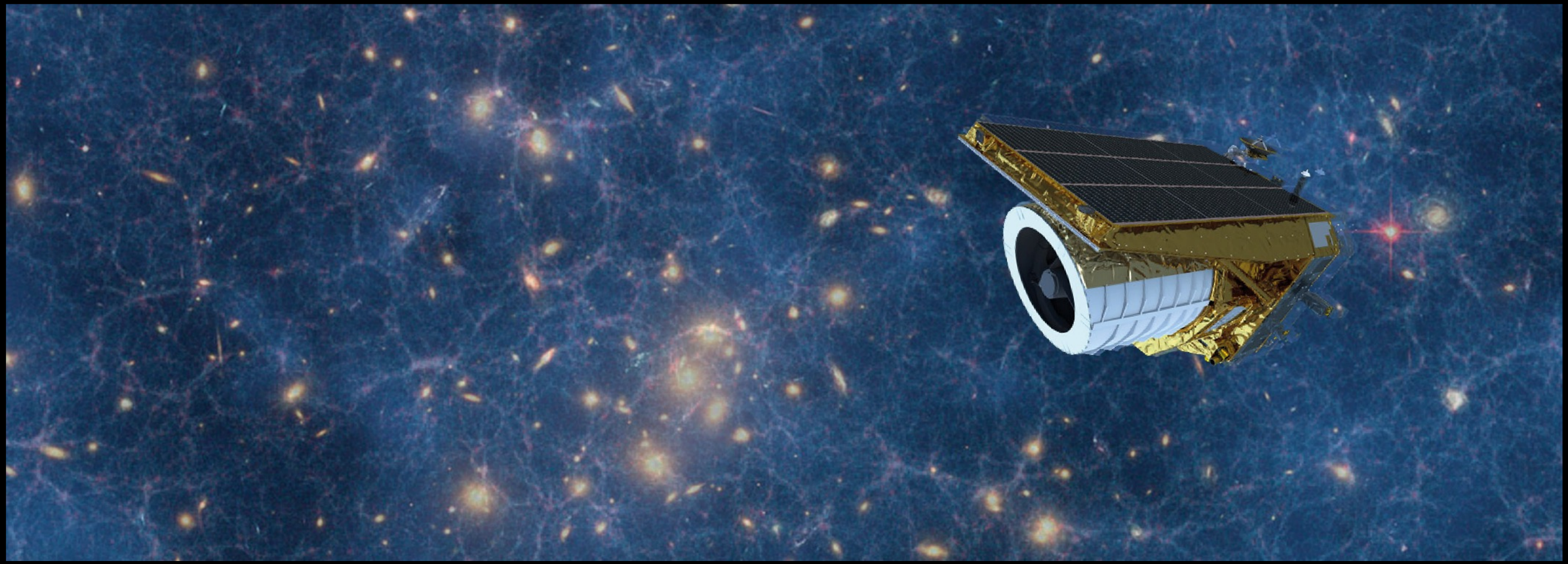
very high precision :

milli-arcsec milli-mag

PSF Phase diverstity cal.

Legacy Science:

Local Universe -LSB
Strong lensing
Theory
Galaxy, AGN evolution
Milky Way
Primeval Universe
Solar system

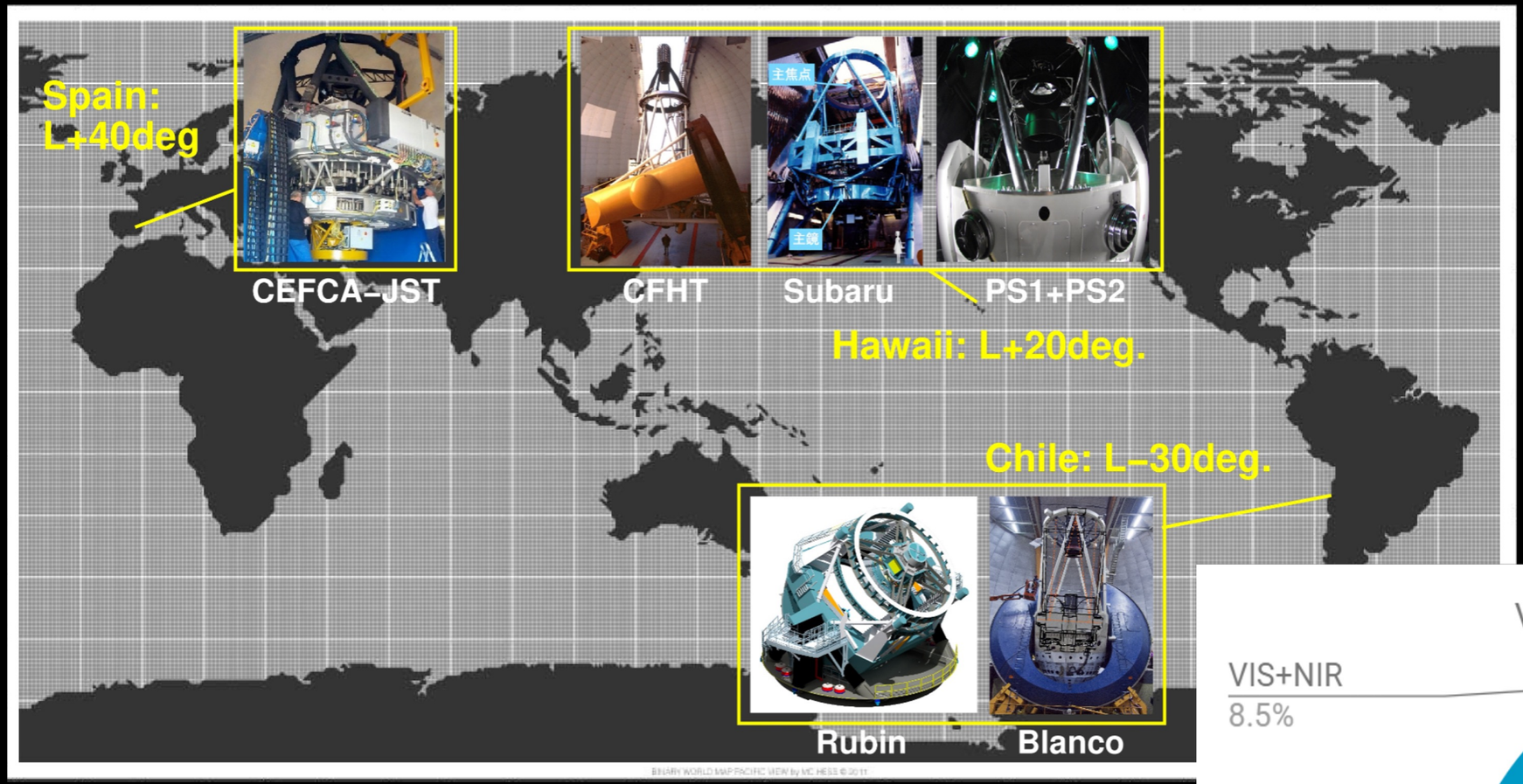


“ Euclid is Big Data from Space ”

2 billion galaxies – distortions and distances

200.000 exposures (1.8 Million images) – 2.5 billion pixels /4.400 sec > 3.700 Tbyte

Costs of data handling ~ costs instruments

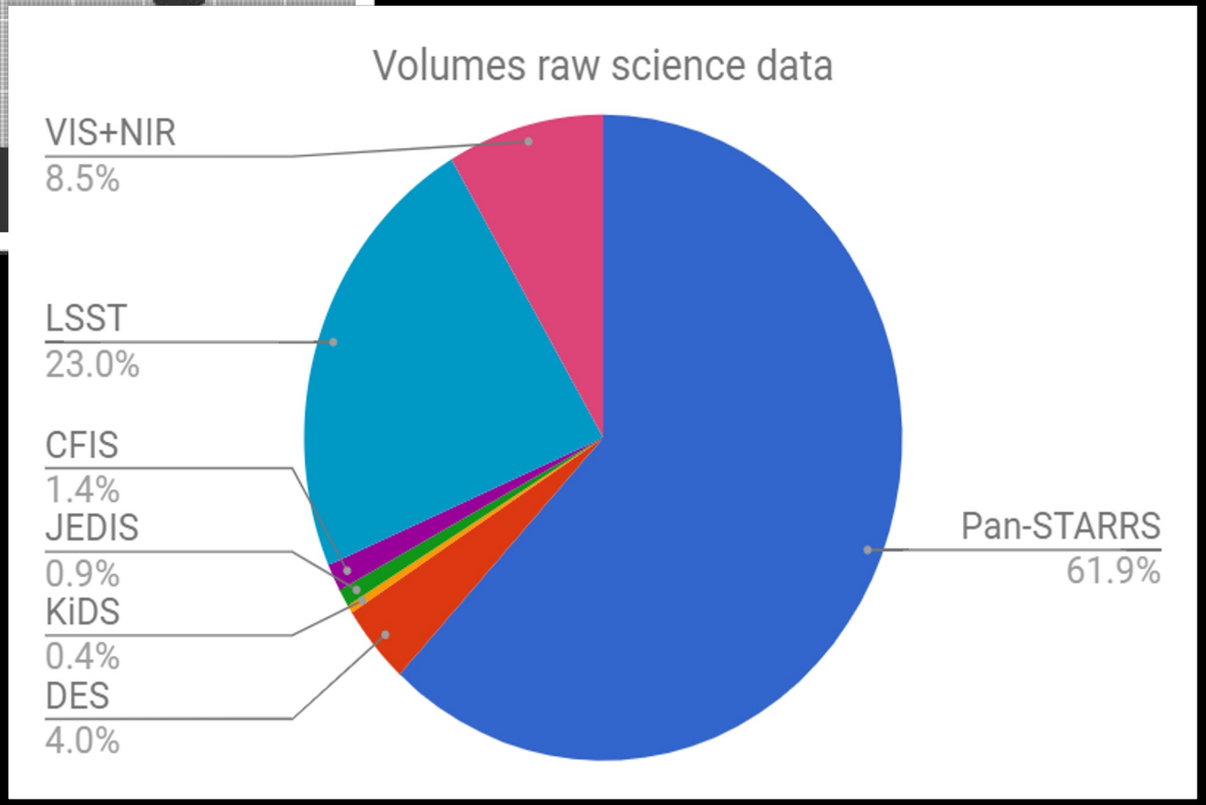


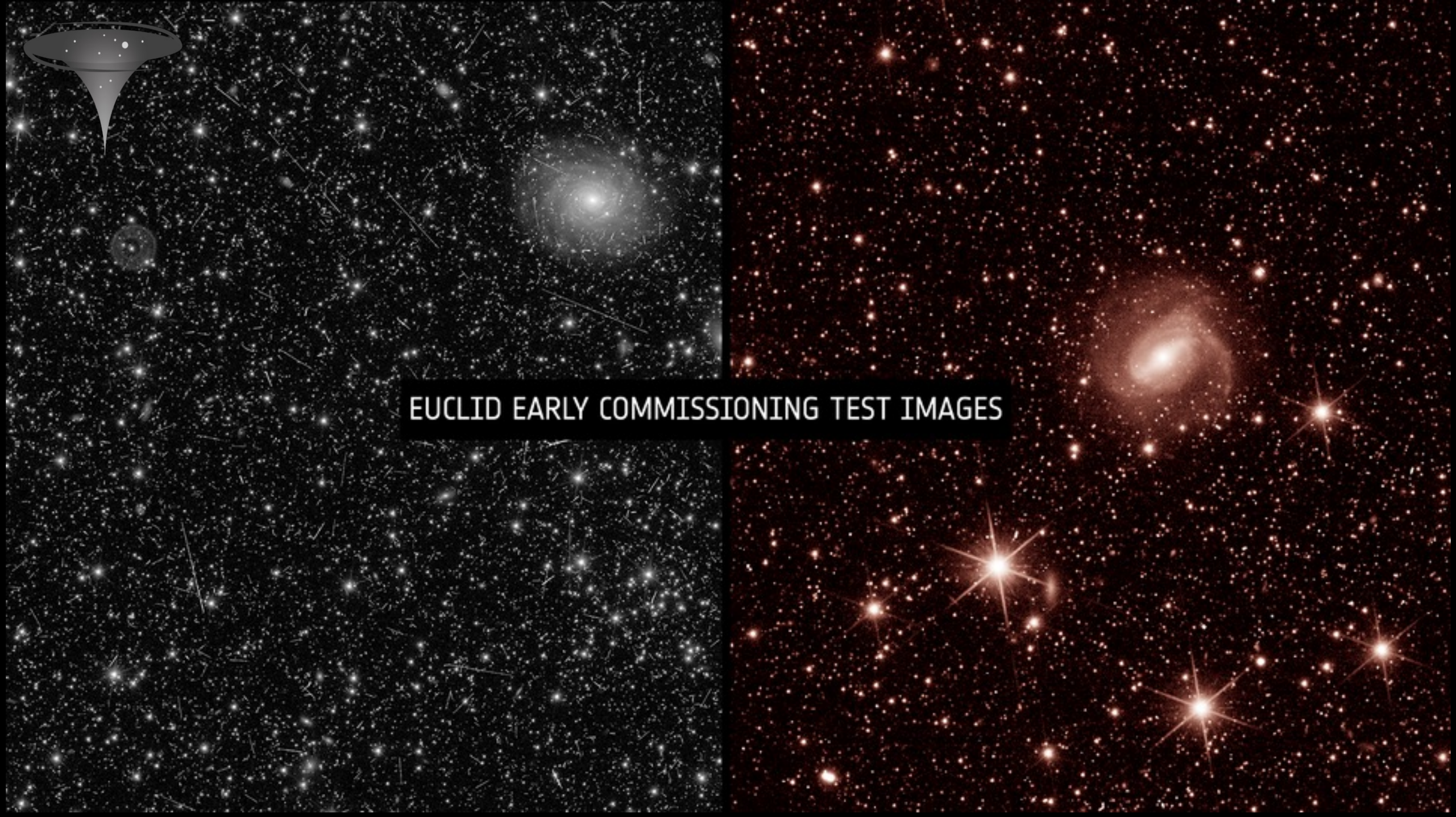
Euclid images > 3.7 Petabyte

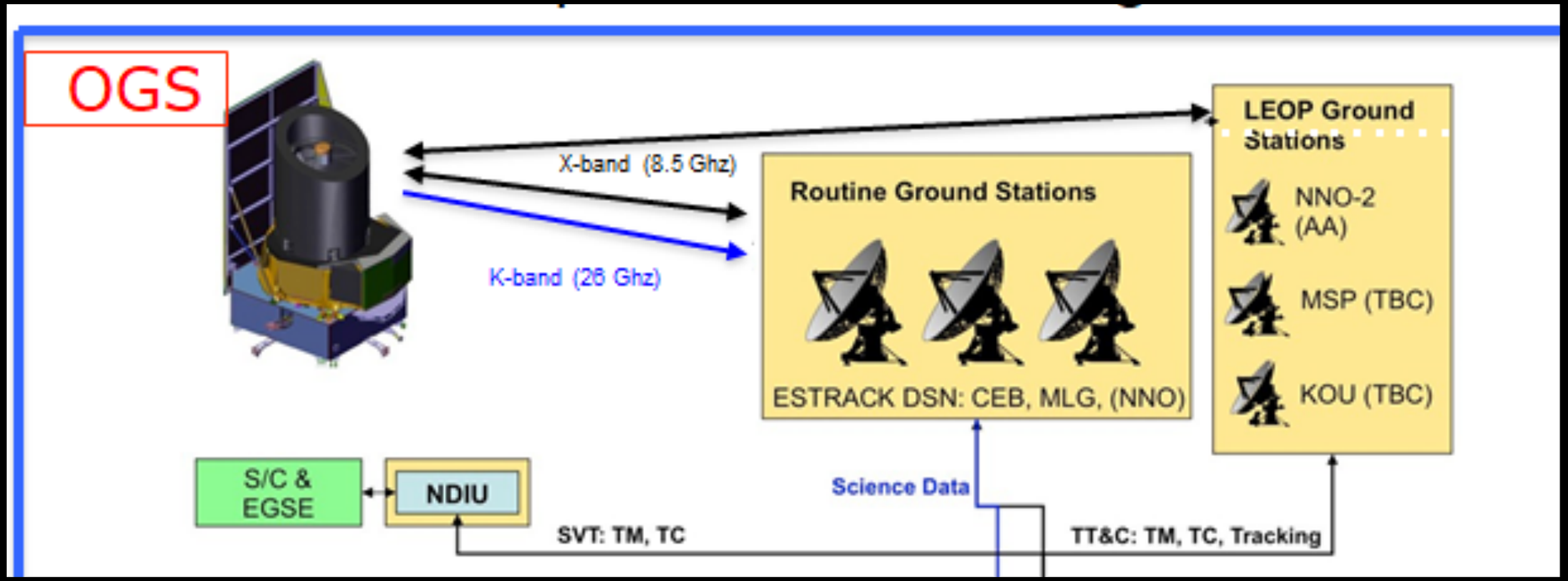
Photometric redshifts- ground based data

> 40 Petabyte

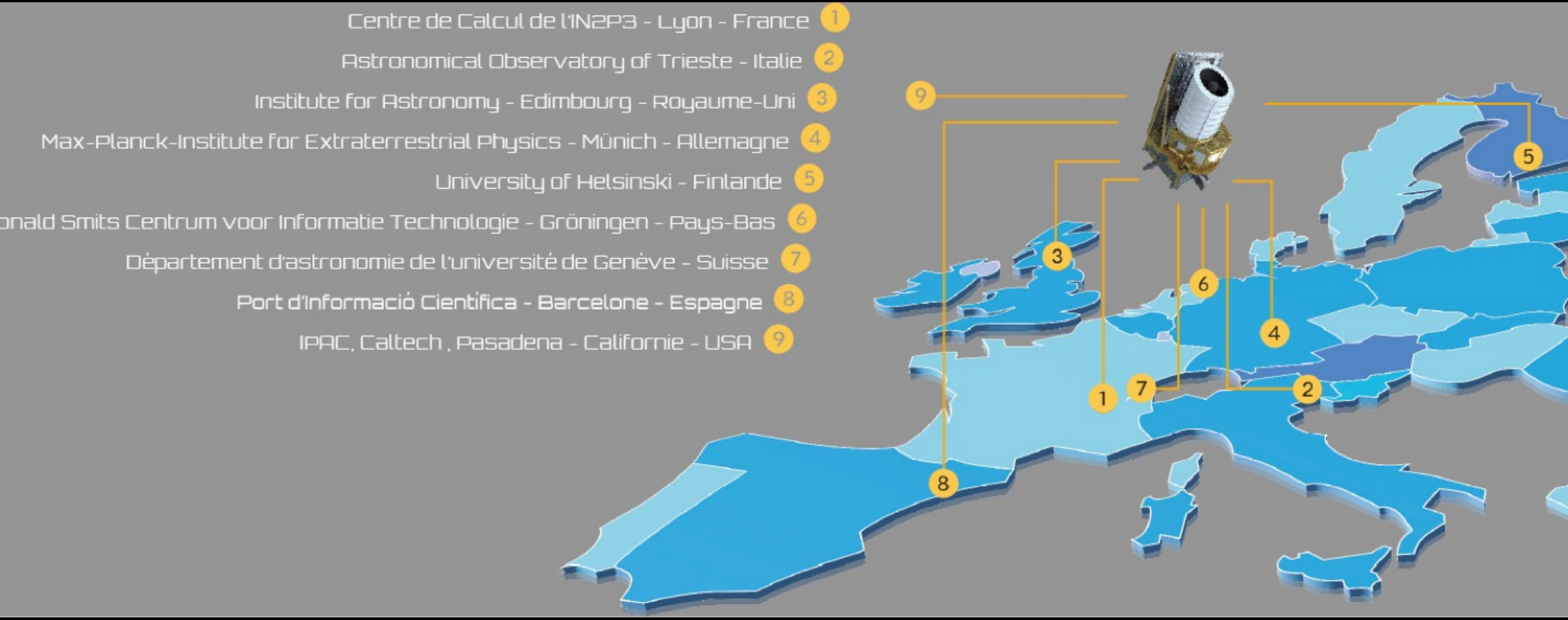
NL- OU EXT




EUCLID EARLY COMMISSIONING TEST IMAGES



K-Band (26 GHz) - L2 - high gain antenna's groundstations

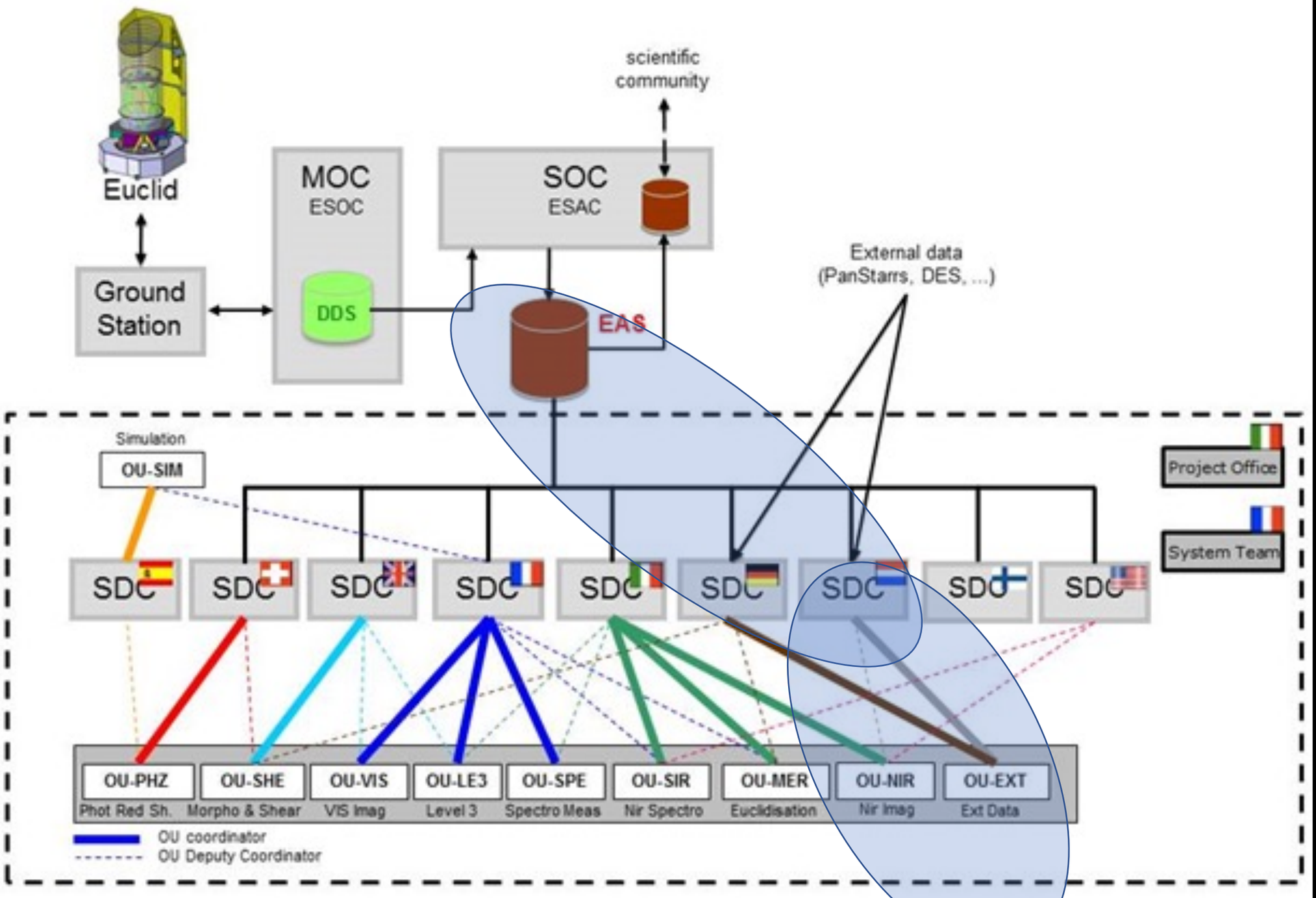


Network of 8 data centers in Europe – 1 in US

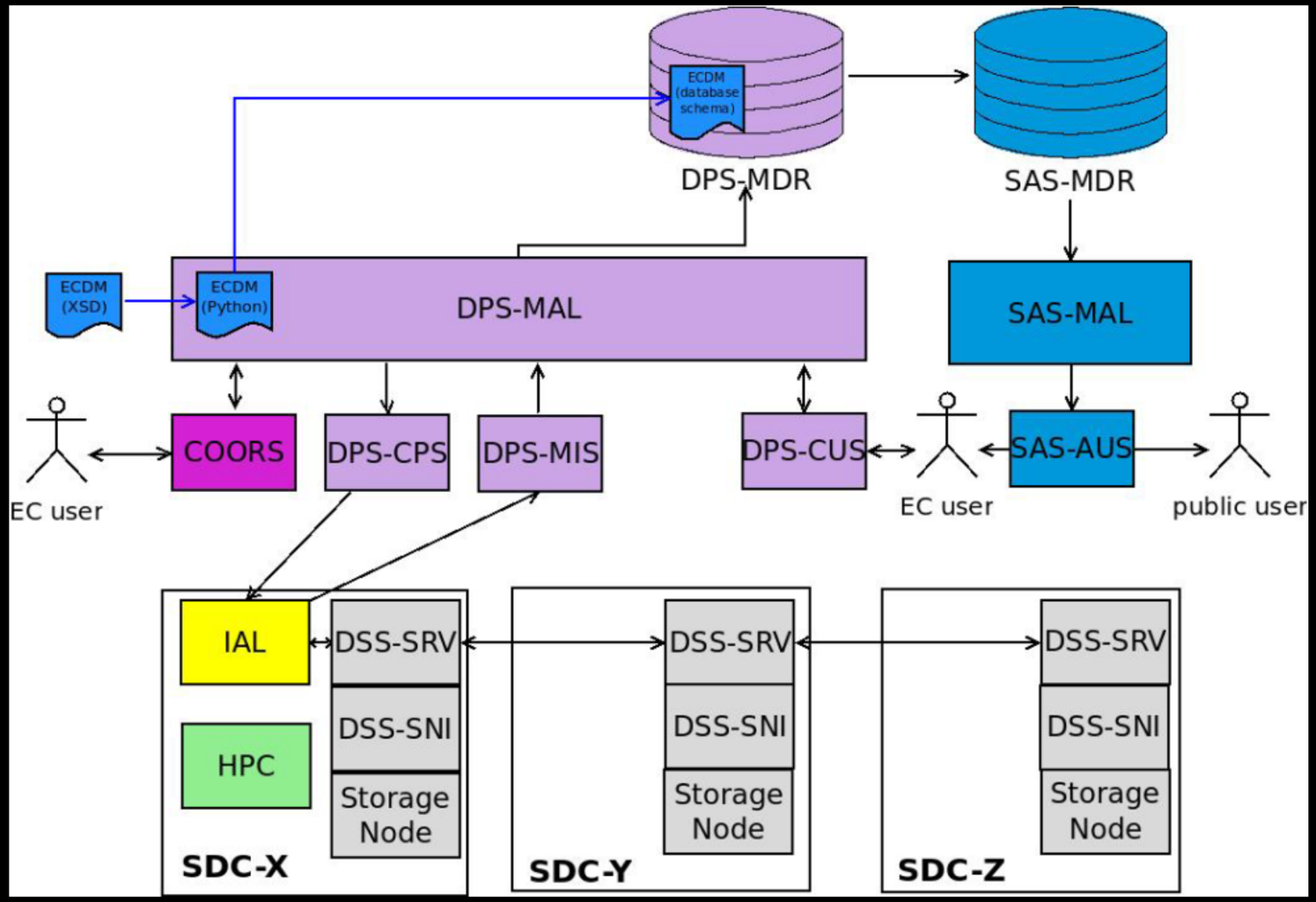
1 integrated system FAIR AstroWISE technology (2003, OmegaCAM, Lofar, +)

Formal NL contribution Euclid = datahandling:

- Euclid Archive System = central database – distributed processing + storage
- Ground based data
- NIR pipeline


Euclid
Ground Station
MOC
ESOC
DDS
SOC
ESAC
scientific community
EAS
External data
(PanStarrs, DES, ...)
Simulation
OU-SIM
SDC
SDC
SDC
SDC
SDC
SDC
SDC
SDC
SDC
Project Office
System Team
OU-PHZ
Phot Red Sh.
OU-SHE
Morpho & Shear
OU-VIS
VIS Imag
OU-LE3
Level 3
OU-SPE
Spectro Meas
OU-SIR
Nir Spectro
OU-MER
Euclidisation
OU-NIR
Nir Imag
OU-EXT
Ext Data
OU coordinator
OU Deputy Coordinator

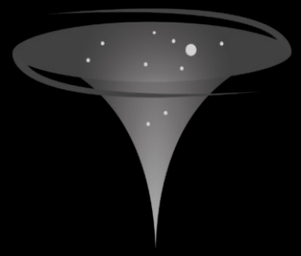
Euclid Archive System



- Central DB F A I **R** (**R**eproducible)
- Connect 9 datacenters
- AstroWISE 2003 – KiDS, LoWISE, MUSEWISE

- No dedicated hardware
- Many access layers

- Data Processing System - DPS
  - at the heart of Euclid data processing
  - Stores Metadata all files, including lineage (calibrations) and processing
  - Production system hosted by ESAC
  - Development system by SDC-NL
- Distributed Storage System - DSS
  - Stores the data files
  - Distributed across all national SDCs
- Science Archive System – SAS access to
  - external users -official data releases
  - internal users - internal data releases



DB view
EAS = DPS + DSS
db mirror at NL-SDC

DpdVisRawFrame

Order by	Data.VisSequence ○ descending ○ ascending
2nd Order by	<none> ○ descending ○ ascending
Maximum number of rows	100
Show only data within project	◉ yes ○ no
Show expanded attributes	◉ yes ○ no
When expanding attributes show empty joins	○ yes ◉ no
Expand desc list attributes	○ yes ◉ no
Show checkboxes	○ yes ◉ no
Export result as	HTML

Submit

▷ DpdVisRawFrame

Home | Contact | Help | user EVALENTI | project EUCLID | Tables | EAS page | Preferences

empowered by

Query results for table DpdVisRawFrame

Shown: 100 rows out of 5200 entries, from only project EUCLID

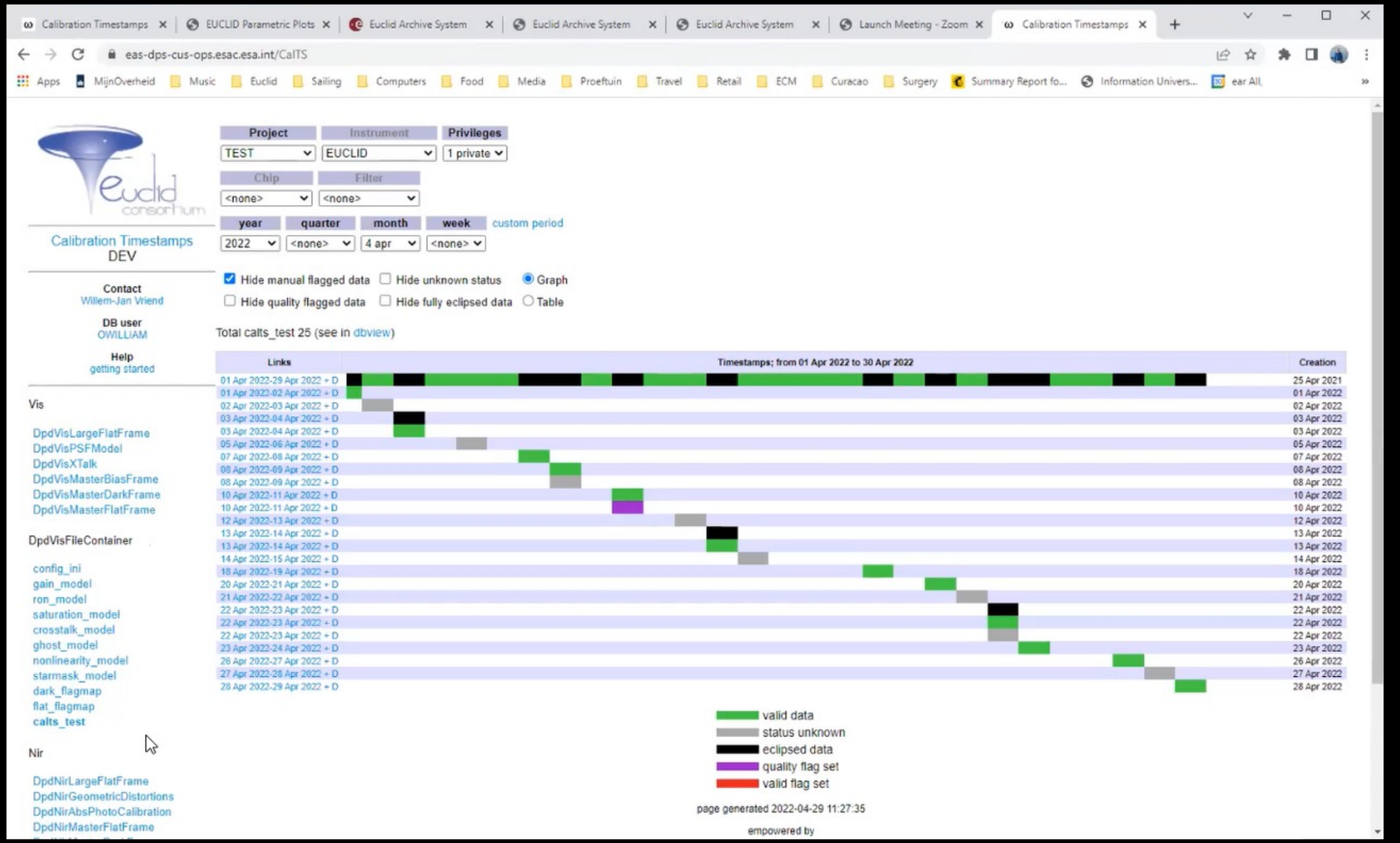
Send via SAMP | Help SAMP

ROWNUM	object_id	Lineage view	+PRIVILEGES	project_id	Data.AxisNumber	Data.CompressionAlgorithm	Data.DataLength	Data.DataSize	Data.ImgNumber	Data.InstrumentMode	Data.ObservationMode	Data.OnBoardS
1	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
2	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
3	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
4	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
5	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
6	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
7	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
8	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
9	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
10	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
11	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
12	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
13	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
14	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
15	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
16	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
17	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
18	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
19	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
20	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
21	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
22	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
23	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
24	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
25	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
26	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
27	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	CALIBRATION	3.1.2
28	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2
29	object view	lineage view	3	30	2	CCSDS121_REORDERING	4439008	16	144	Science	WIDE	3.1.2

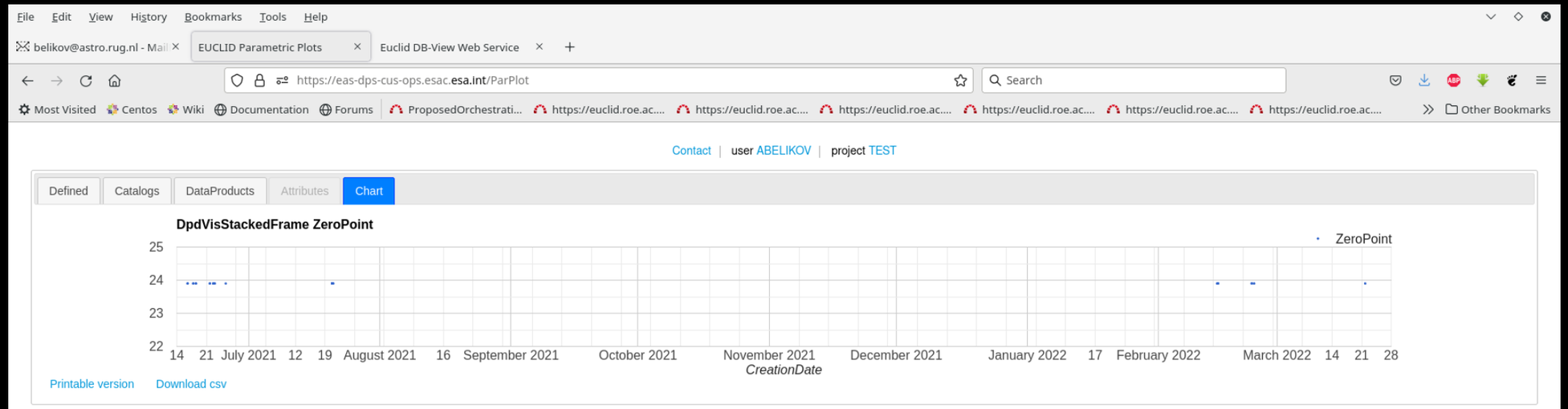
EAS:
~80 NL-EC
2.800 EC

later:
Public ESA - SAS

Calibration Files - timestamping



ParPlot e.g. trending



Jupyter Notebooks - Gitlab

Home Page - Select or create a | Untitled - Jupyter Notebook

https://apceuclidccjupyter.in2p3.fr/user/tnutma/notebooks/Untitled.ipynb?kernel_name=eas-dps-env_0.8

jupyterhub **Untitled** Last Checkpoint: 4 minutes ago (unsaved changes) | Logout | Control Panel

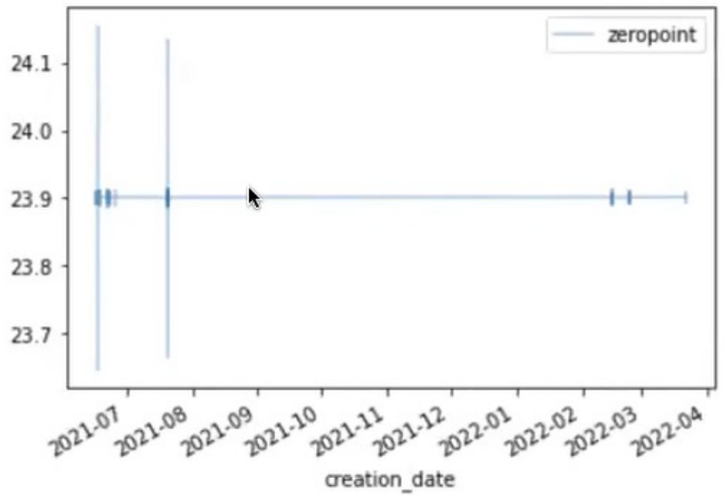
File Edit View Insert Cell Kernel Help | Trusted | Python 3 (EAS-DPS 0.8)

Run | Code

2021-07-20 17:29:48.271	2021-07-20 00:04:23.703	2021-07-20 02:09:53.747	23.9	0.01141	mag	VIS
2021-07-20 18:36:36.578	2021-07-20 03:40:47.094	2021-07-20 04:10:23.209	23.9	0.01317	mag	VIS

In [10]:
```
# Plot the zero point vs creation time.
df.plot(y='zeropoint', yerr='zeropointerror', alpha=0.4)
```

Out[10]: <AxesSubplot:xlabel='creation_date'>



In [ ]:

SDC-NL: Facilities at CIT

3 PByte storage

80 TByte SSD storage and two ORACLE database servers

8 servers Data Handling System - Euclid Archive System

640 cpu cores Habrok High Performance Cluster
+ **18,000 shared cores**

Contacts NL -Science Data Center

SDC-NL@list.rug.nl

Valentijn
lead

Saifollahi
Notebooks

Verdoes Kleijn
External data

de Jong
External Data

Eduardo Balbinot
External Data

Pablo Corcho Caballero
External Data

Williams
CIT lead

Vriend
User Interfaces

Boxhoorn
Databasing

Dröge
HPC

Tsyganov
Databasing

Belikov
Euclid Archive

Bouwens- NISP

Venemans- NISP

Organizing distributed communities – Open Science

Cathedral

Database

Data model

Access layers

DPS MAL

DPS x

IAL

DSS x

Containers -data Meta data

Bazar

Hardware

processors

storage

Jupiter Notebooks