

From the VLT to the ELT: Challenges and Strategies for Science Data Management

Michael Sterzik
European Southern Observatory
Data Management and Operations, Head



ELT construction at full speed



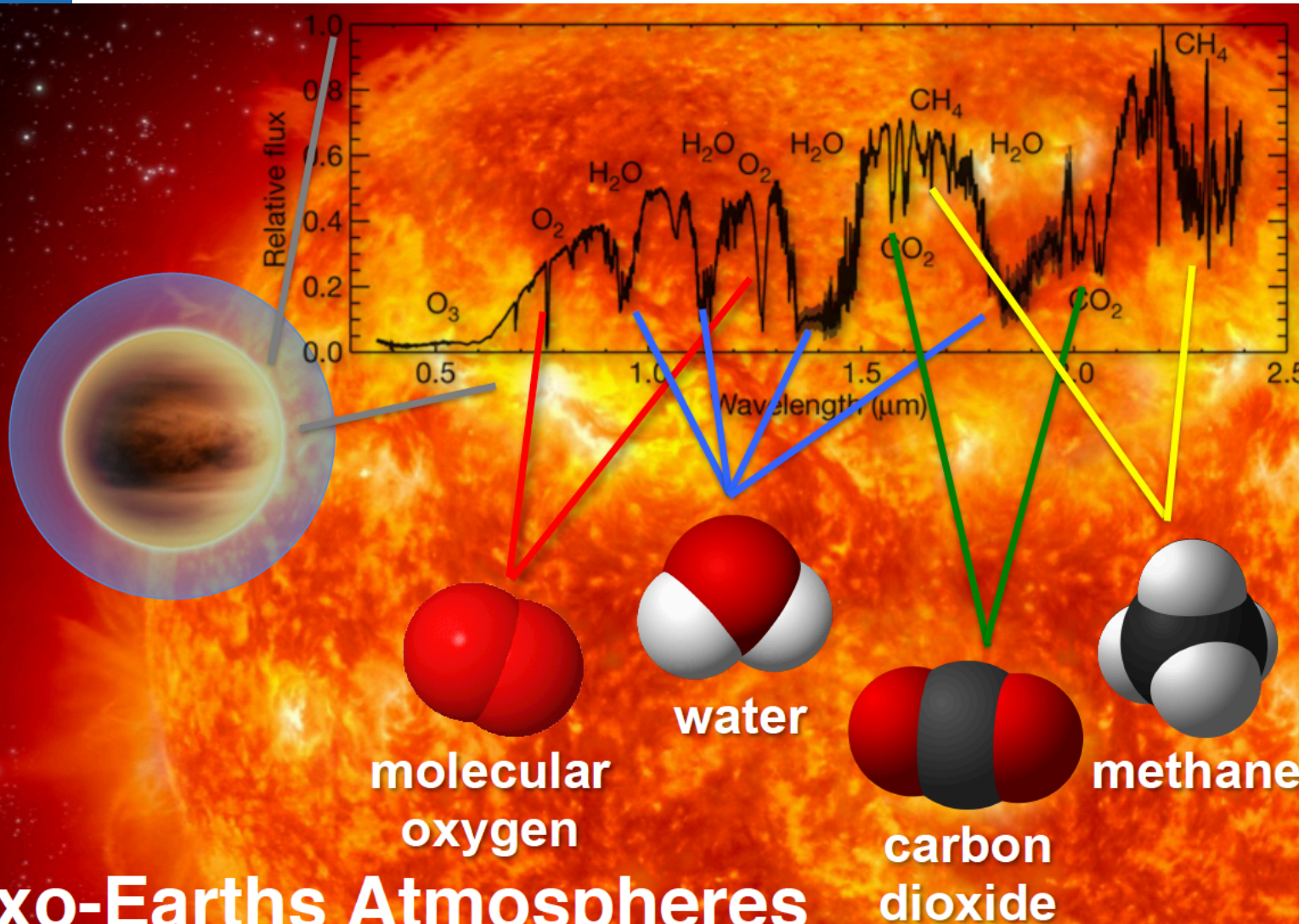


ELT First Light: 2024





Characterize exoplanets



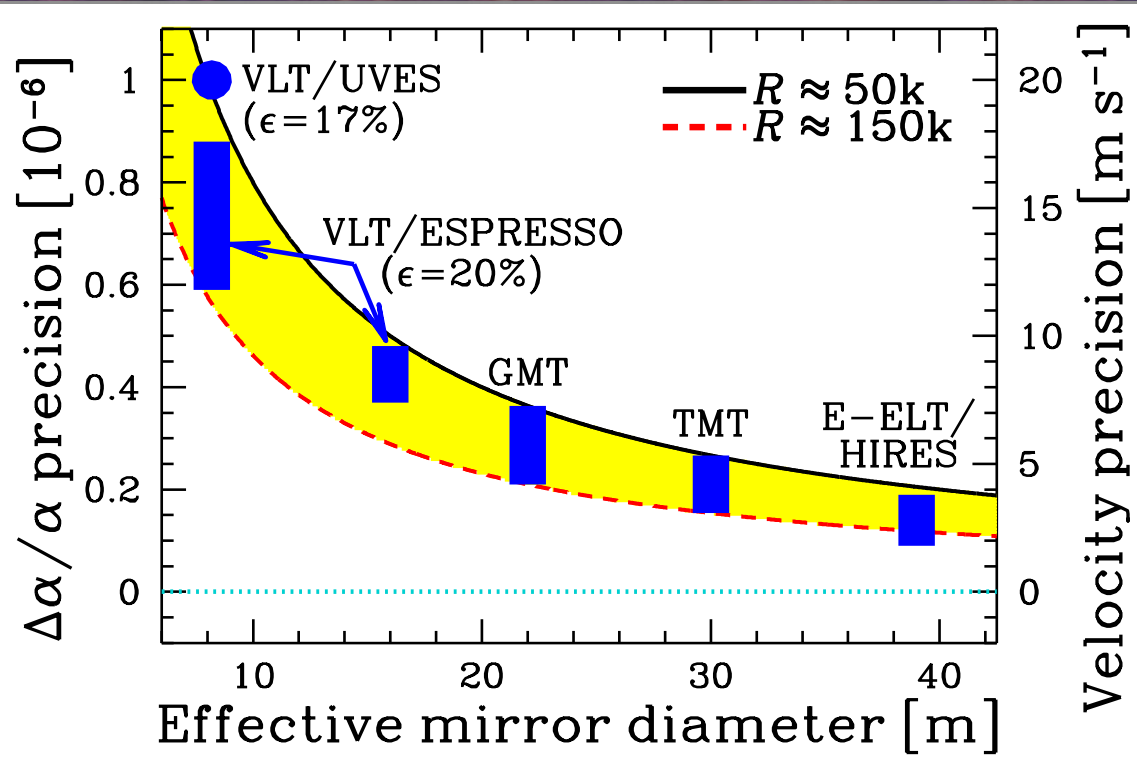
Exo-Earths Atmospheres
Detecting signatures of life



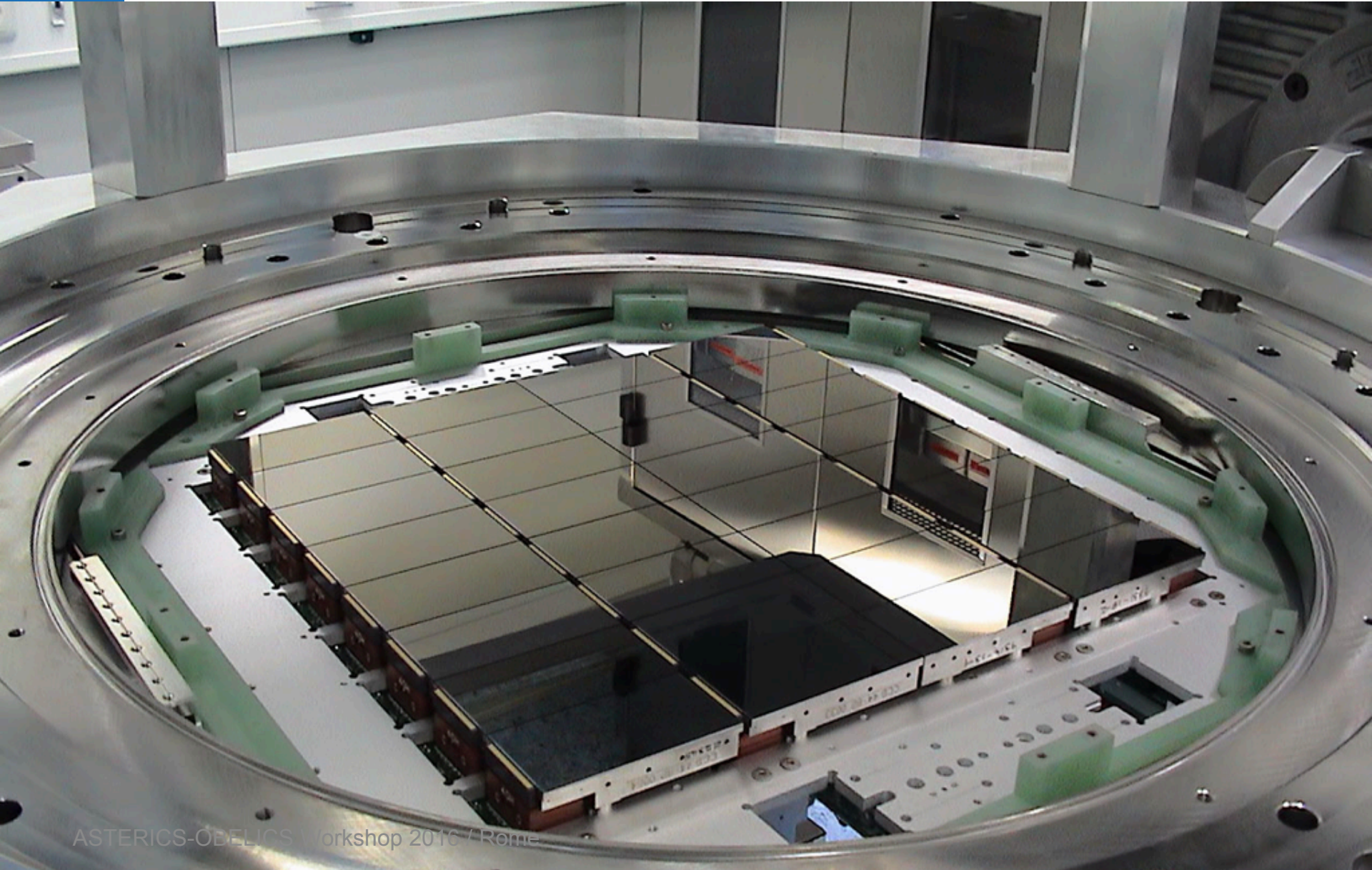
Fundamental constants: Mapping the dark universe

The E-ELT with a High-Resolution spectrograph can provide the required accuracy for reliable measurements of the fine-structure constant $\alpha \equiv 2\pi e^2/hc$ and proton–electron mass ratio $\mu \equiv m_p/m_e$ at the 0.3 part per million level.

A detection of varying fundamental couplings would be revolutionary: it would automatically prove that the Einstein Equivalence Principle is violated (i.e. gravity is not purely geometry), and that there is a fifth force.

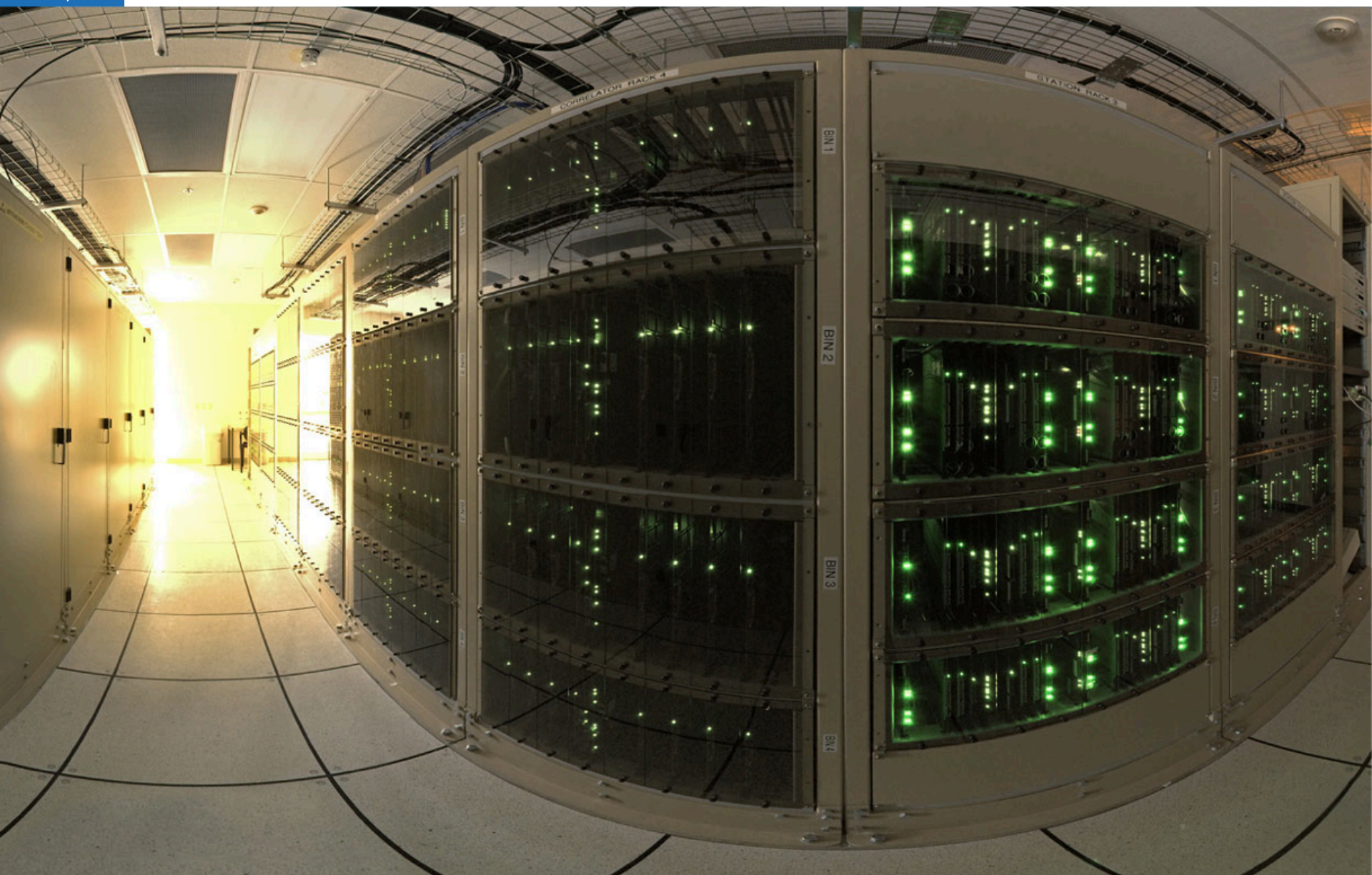


(Not a real) challenge for DM





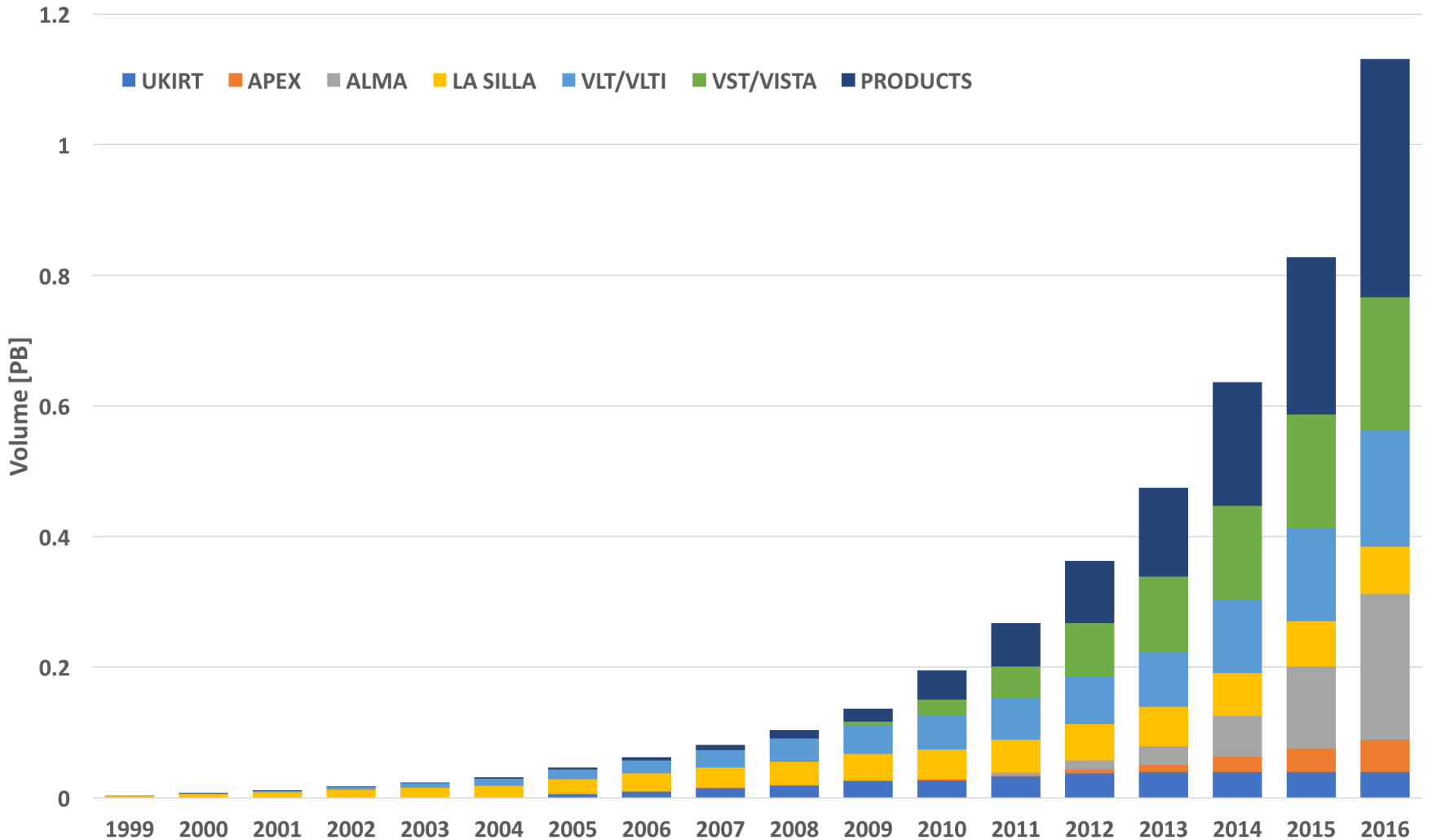
(Not a real) challenge for DM



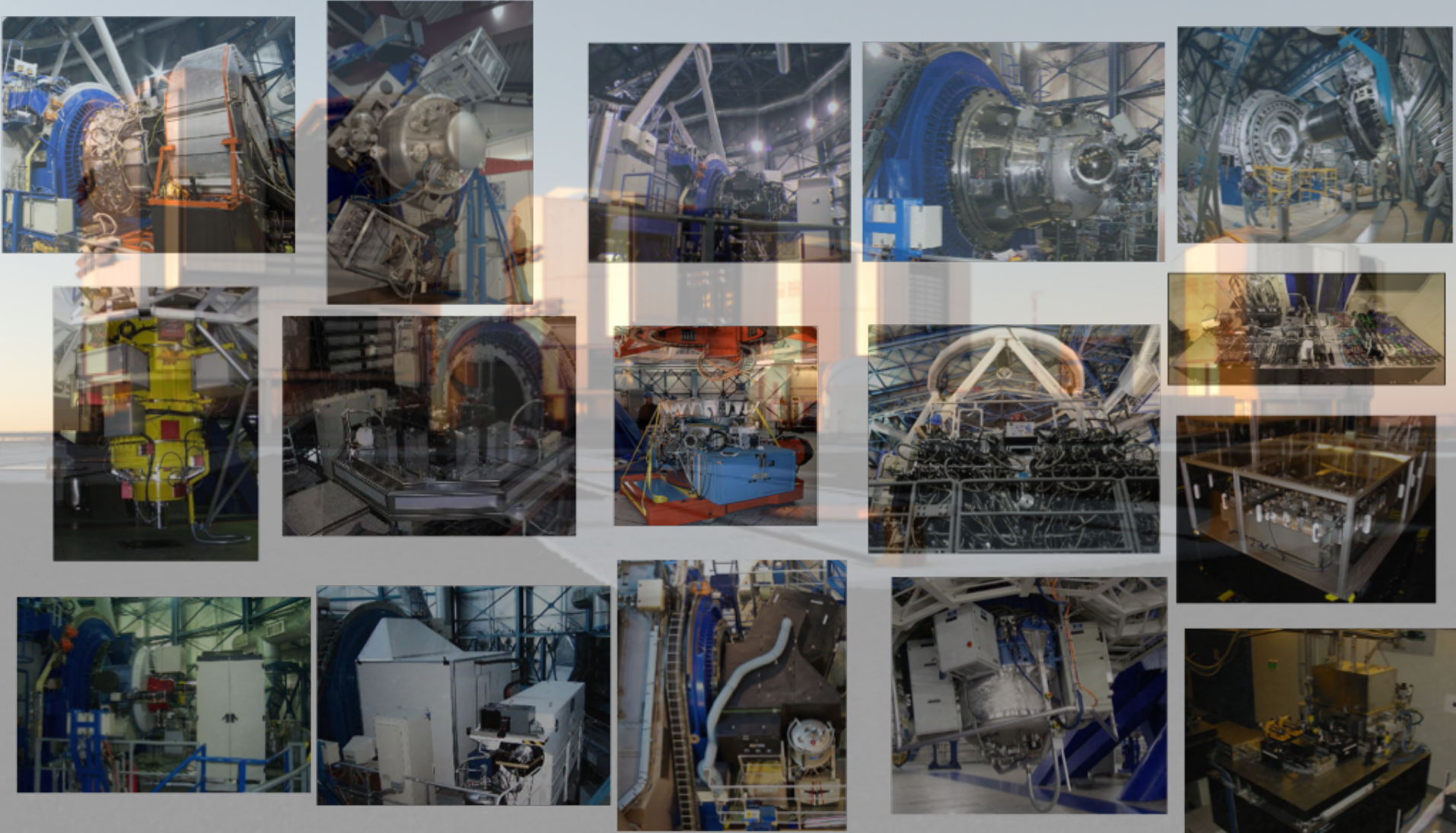


Data Volume in the ESO Archive

ESO Archive Assets

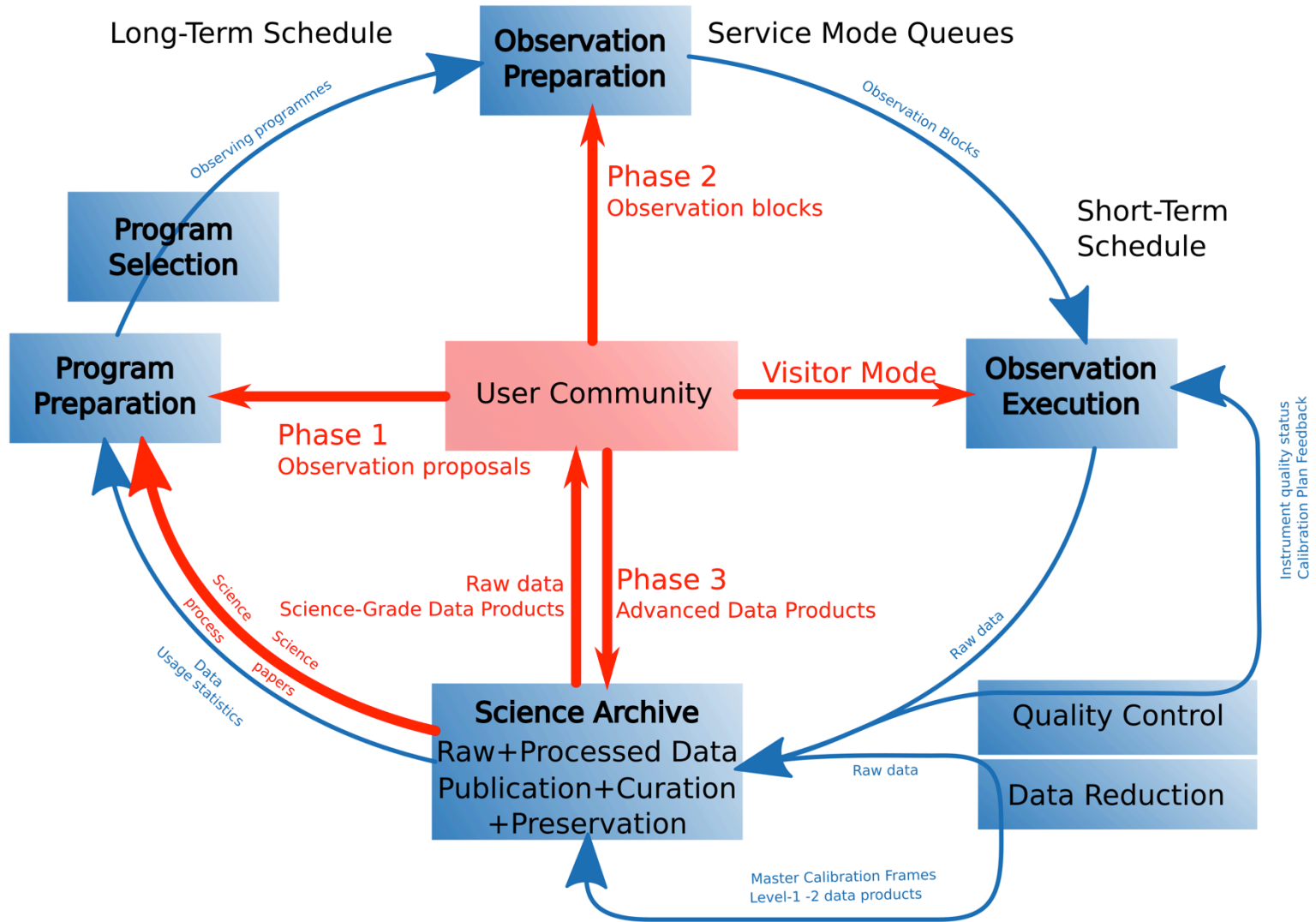


Challenges





E/VLT end-to-end operations model

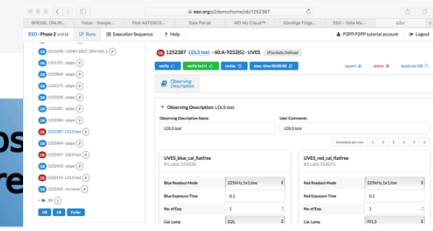




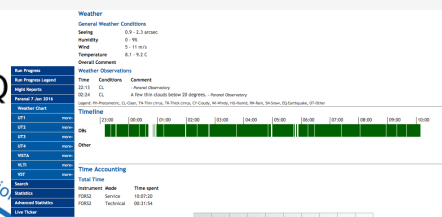
E/VLT end-to-end operations model



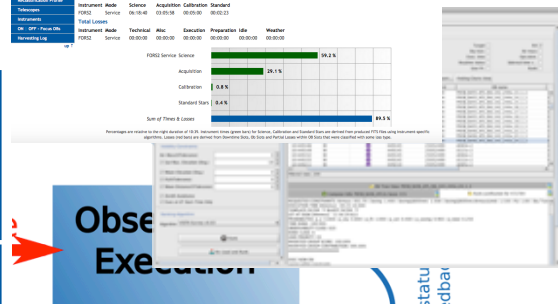
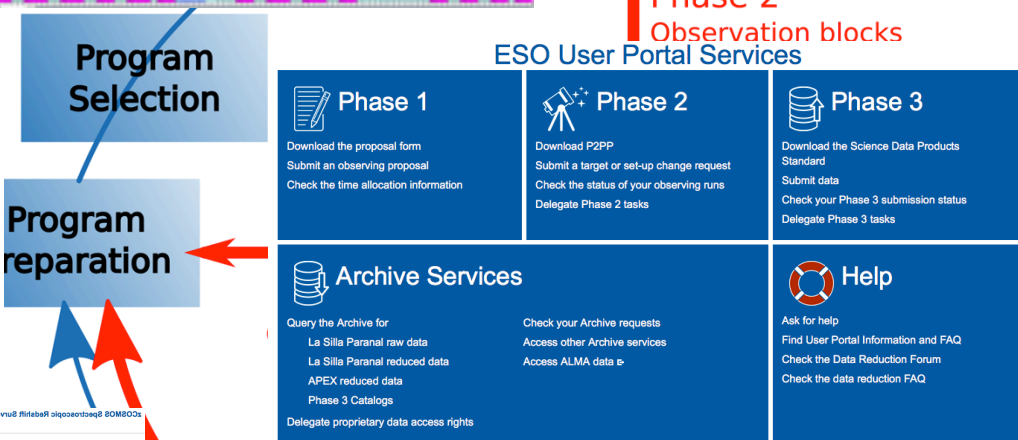
Obs Pre



Q



Phase 2 Observation blocks



FORS2 trending system: overview of scores

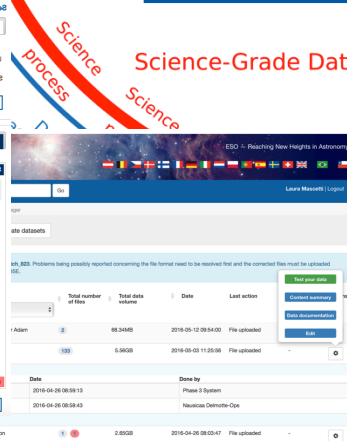
Last update: 2016-06-08T12:34:27 (LT)

General views: FORS2 scores

FORS2 instrument scores

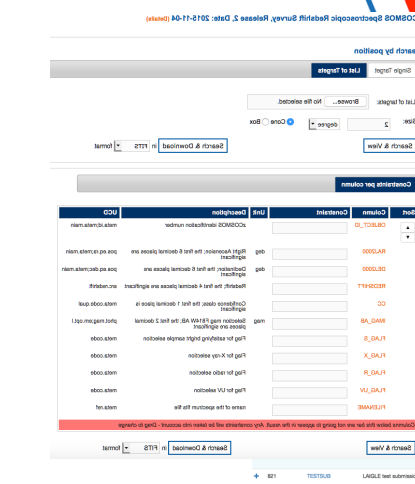
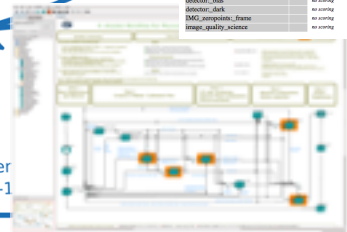
group	score	comment
detector_monitoring	1000	OK
LSS_wavelength_calibration	1000	OK
MOS_wavelength_calibration	1000	OK
polarization_optics	1000	OK
spec_efficiency	1000	OK
spec_resolution	1000	OK
detector_monitoring	1000	OK
spec_efficiency	1000	OK
detector_monitoring	1000	OK
spec_efficiency	1000	OK
detector_monitoring	1000	OK
spec_efficiency	1000	OK

Raw data | Phase 3 Science-Grade Data Products | Advanced Data Products



master level-1

Reduction





E/VLT Development Roadmap

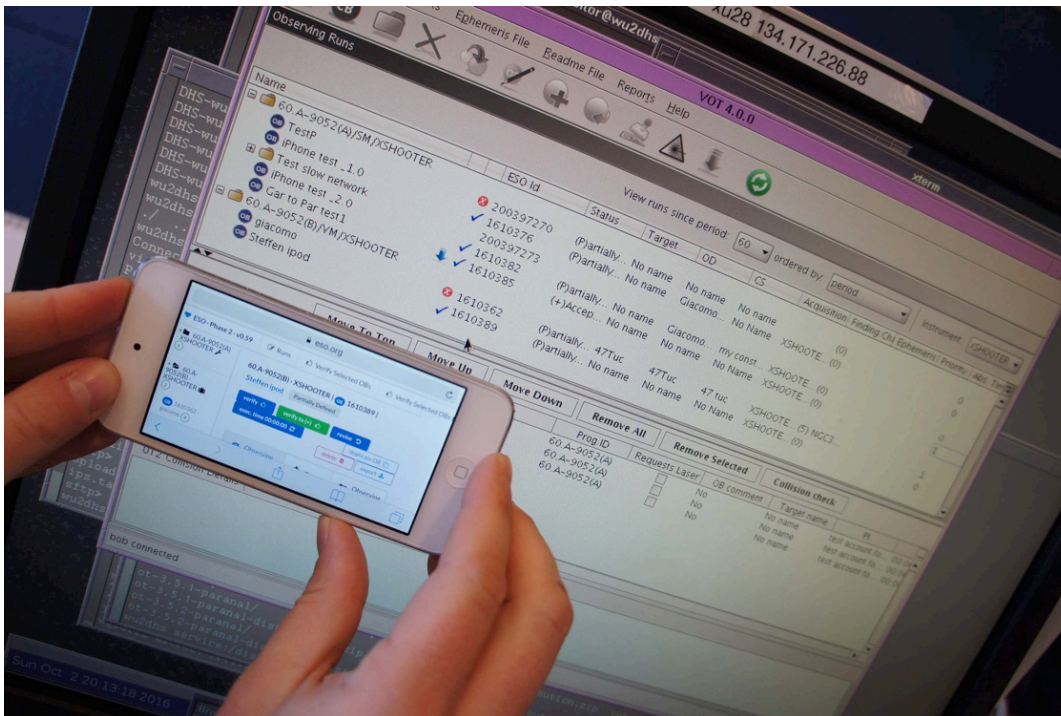
		2015	2016	2017	2018	2019	2020	2021	2022	2023	2024
	ELT VLT			◆ Instruments' PDR		◆ Instruments' FDR					First light ◆
		◆ Sphere	◆ Gravity	◆ Espresso	◆ Cires+	◆ Eris	◆ Moons	◆ 4Most			
			◆ AOF	◆ Matisse							
PHASE 1	P1web TAtwo	Pre-Pre-Pre-A	A A A B	B B B B	B B B B	B B B B	B B B B				
PHASE 2	P2fO GuideCam ETC	Pre-Pre-A A	A A A B	B B B B	B B B B	B B B B	B B B B				
Observations	OT4 + STS Remote Control AstroClimate Interface ELT			Pre-Pre-Pre-Pre-A A	A A B B	B B					
QC and Data Processing	QC PL libraries PL infrastructure			Pre-Pre-Pre-Pre-Pre-A A A A	A A A A	A A A A	B B B B	B B B B			
Archive	Archive Services NGAS+	Pre-Pre-Pre-Pre-Pre-A A A	B B B B	B B B B							
General Operations	DHS+ DOME+		Pre-Pre-Pre-Pre-A A A A	B B B B							



Towards a future-proof system

NEW Phase 2 Preparation Solution:

- dynamic WEB application, based on Angular 2
- zero-install, real-time, bi-directional DB handling
 - The API empowers the community to built their own tools for flexible observation specification and modification



Production of Science Data

■ In-house generation of Data Products (IDPs)

- enabled through standardized acquisition and quality control
 - near-real time quality control process ensures certified master calibrations
- un-attended processing w/ certified pipelines, process is QC'ld
- goal: science grade data for all popular instrument modes
 - UVES, XSHOOTER, HARPS, FLAMES/GIRAFFE, MUSE, HAWK-I, FEROS
 - imminent: VIMOS (IMG), PIONIER, KMOS, FORS2

■ External Data Products (EDPs)

- provided by public surveys and large programs (deliverables)
- programs selected by their high legacy value
- most use dedicated (non-ESO) user-pipes (eg CASU)
- goal: advanced products (wide, deep, merged catalogs)
- perspective: users *at large* contribute EDPs
 - quality assurance: published datasets only?
 - acknowledgement: DOIs?

Science Data Process

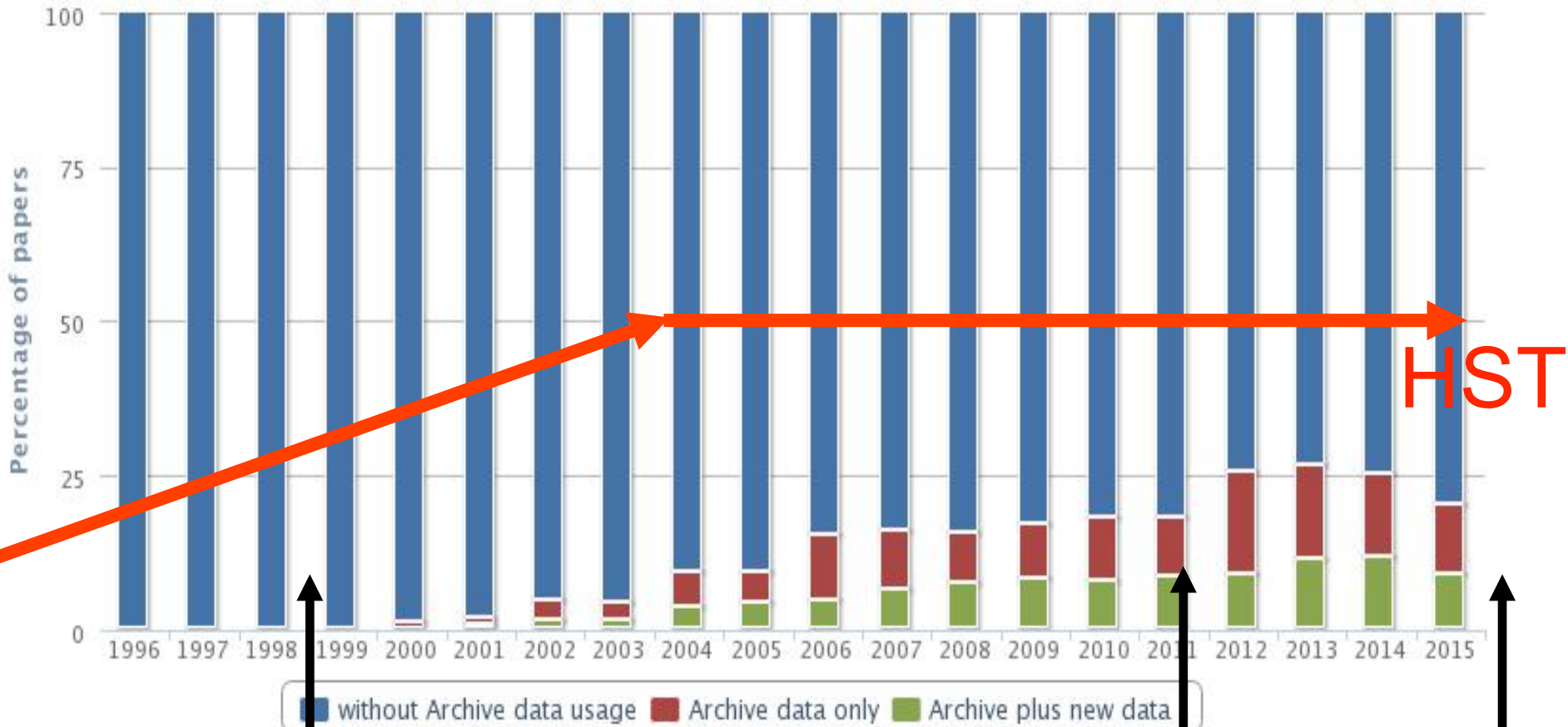
- ESO Phase 3 process enables
 - preparation, submission, validation and ingestion of science data products for storage in the ESO Science Archive Facility (SAF), and subsequent publication to the scientific community.
- ESO Science Data Product Standard is **required** for coherence of EDPs *and* IDPs in the SAF
 - defines format, meta-data, keywords, quality descriptors and processing provenance
 - generally derived from “VO” standards, when available
 - www.eso.org/sci/observing/phase3/p3sdpstd.pdf
- added-value through validated and curated content
- ESO SDPS innovations
 - multi-epoch photometry (surveys, timeseries, NGTS)
 - processing provenance
 - 3D/IFU cubes (KMOS, MUSE!)
 - sub-mm/radio maps (APEX/ATLASGAL)



Science Archive as a Resource

U. Grothkopf et al., <http://www.eso.org/sci/libraries/edocs/ESO/ESOstats.pdf>

Archive Data Usage
Source: telbib, 1996 - 2012



start of facility operations

start archive population with DP

archive services interoperability



... and costs?

(fraction of total operation costs)

- data archive operations
 - archive infrastructure TCO (1PB, 3 safe copies) 0.3-1%
 - content management (production, curation) ~10%

- “systemic” data generation
 - facility (VLT) time for calibrations ~ 4%

- favorable cost-benefit relation
 - close monitoring, metrics...
 - effective use of resources (FTE and \$)

Interaction welcome!

■ ASTERICS Project Contact

- Martino Romaniello (Back-end Operations, Head)
- Olivier Hainaut (End-to-end Operations Scientist)
- Thomas Bierwirth (Dataflow Project Manager)
- Stefano Zampieri (Dataflow Software Eng., Lead)
- Michael Sterzik (Coordination)

■ active exchange with CDS and ESA is ongoing

Acknowledgement

Support of accommodation:

- H2020-Astronomy ESFRI and Research Infrastructure Cluster (Grant Agreement number: 653477).