

# Common Archive Observation Model, CAOM within the ESAC Science Data Centre

M. Arévalo<sup>1</sup>, J. Espinosa<sup>1</sup>, on behalf of ESDC team Patrick Dowler<sup>2</sup>, Séverin Gaudet<sup>2</sup>, on behalf of CADC team Brian McLean<sup>3</sup>, David Rodriguez<sup>3</sup>, on behalf of STScI team

<sup>1</sup> RHEA Group for ESA, <sup>2</sup> Canadian Astronomy Data Centre, <sup>3</sup> Space Telescope Science Institute

**ESCAPE** Data Provider Forum

23/11/2021

ESA UNCLASSIFIED – For ESA Official Use Only



## **ESAC Science Data centre**



#### ESA Digital Library of the Universe.

- Astro-, Planetary, Helio-, and Fundamental Physics
- Science data and added value tools
- All phases from development to Legacy



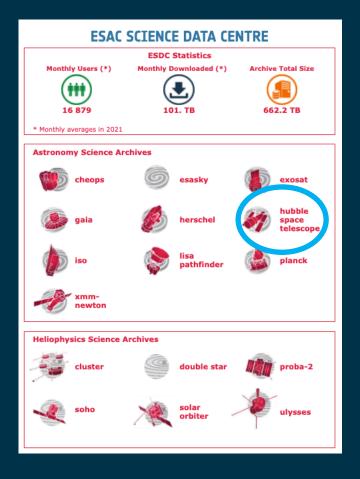
Enable maximum **scientific exploitation** of data sets



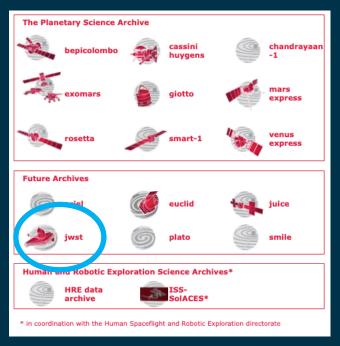
Enable efficient long-term preservation of data, software and knowledge, using modern technology



Enable cost-effective archive production by *integration in, and across, projects* 



Primary source of data for ESA's Space Science Missions



https://www.cosmos.esa.int/web/esdc

## **European HST and JWST Archives**



- Provide official ESA repository of public mission data
- Support the community (especially European) in archival data access and use

All HST (and in the future, JWST) data is made available via

- the Mikulski Archive for Space Telescopes (MAST at STScI)
- the Canadian Astronomy Data Centre (CADC)
- and ESA/ESAC at ESDC.







european space astronomy centre

## **European HST & JWST Archives evolution**



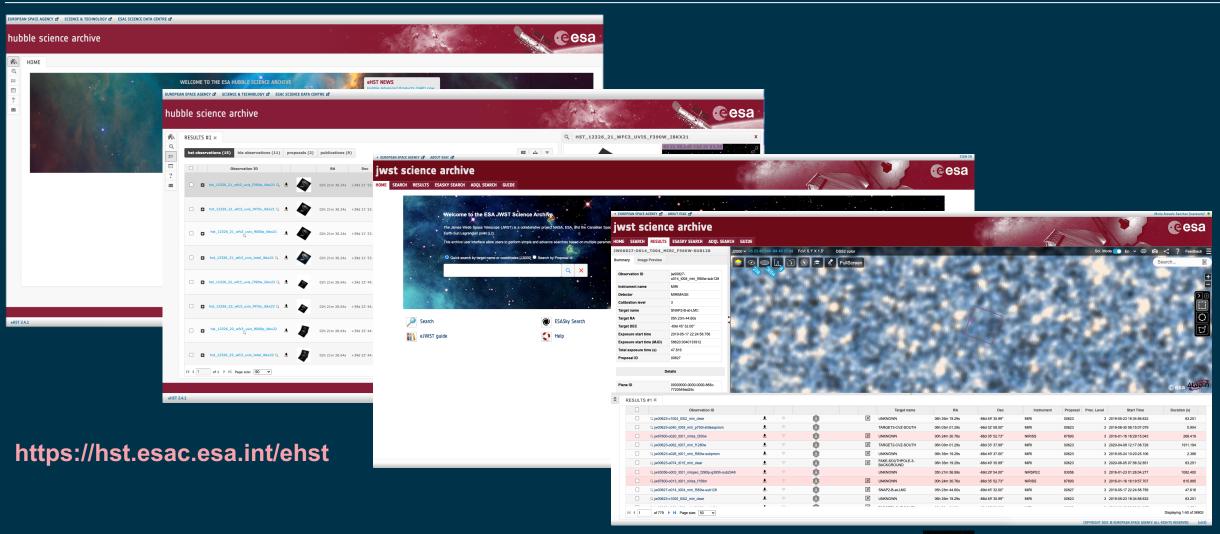
#### **Brief history of eHST**



- 1984-2010: The Space Telescope European Coordinating Facility (ST-ECF)
- 2011-2012: European HST Archive moved from ESO/ST-ECF to ESAC
- 2013 2015: eHST re-engineering project
- 2015: eHST 1.0 released
  - Including continuous processing of science data with the latest calibration software, local reprocessing pipelines at CADC and ESAC locally
- 2018: eHST 2.0 MAST is source of HST reprocessed metadata
  - CADC, MAST, ESDC common software tools
  - Same information all over three partners
  - CAOM designed by CADC is the main driver for metadata unification.
- 2021: Same approach implemented within the European JWST archive.
  - European JWST archive first public version will be available in December 2021

# **European HST and JWST Archives Uls**



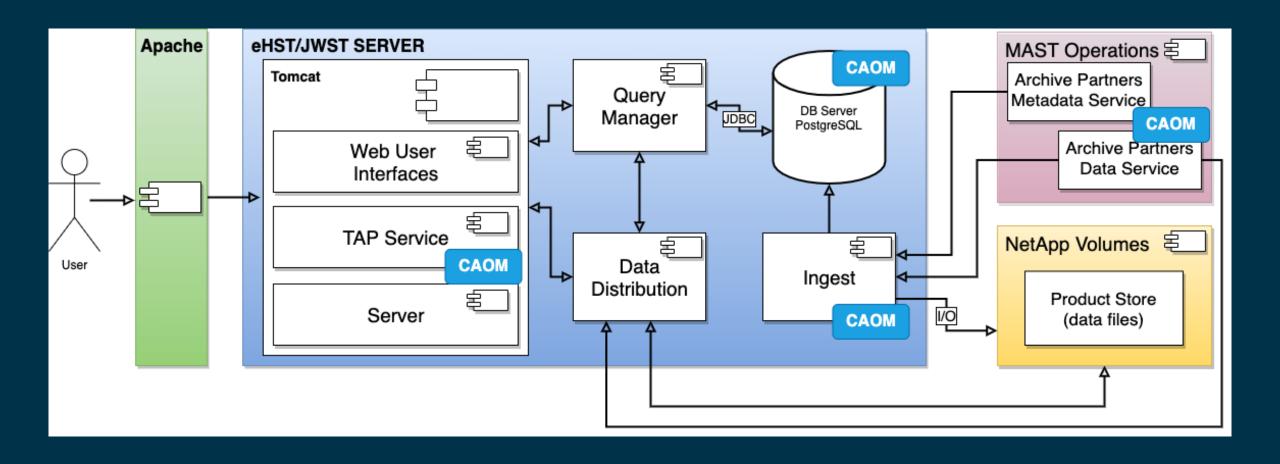


**URL** will be coming soon



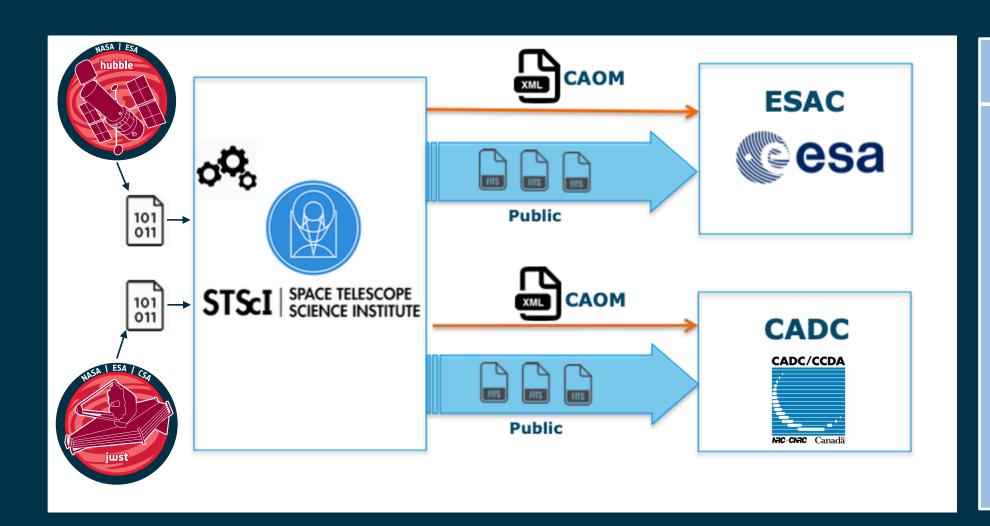
# **CAOM** inside European HST and JWST Archives





## Metadata and data transfer





**CAOM** 



Metadata and data synchronisation between eHST/JWST partners:

- Common Data model
- Software libraries.

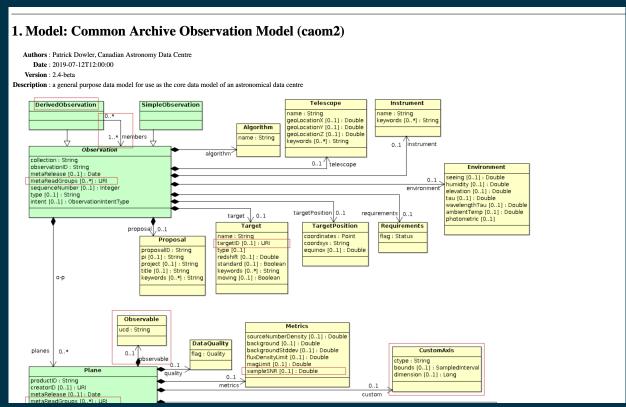
Regular meetings to

- Detect synchronization issues
- Inform about new datasets and reprocessings
- Share status

## **CAOM Data Model**



- CAOM: Common Archive Observation Model
- Standard model for metadata between partners.
- CAOM Architecture is based in three main components and their relationships:
  - Conditions of an observation.
  - Associated datasets.
  - Data files containing observational data.
- Support for IVOA Standards:
  - ObsCore DM (Observations + Planes).
  - DataLink (Planes + Artifacts).
  - SODA (Artifacts).
  - SIA.
- Common software for metadata and data sync:
  - Open source (https://github.com/opencadc).
  - A three-way collaboration between STScI, ESAC and CADC.

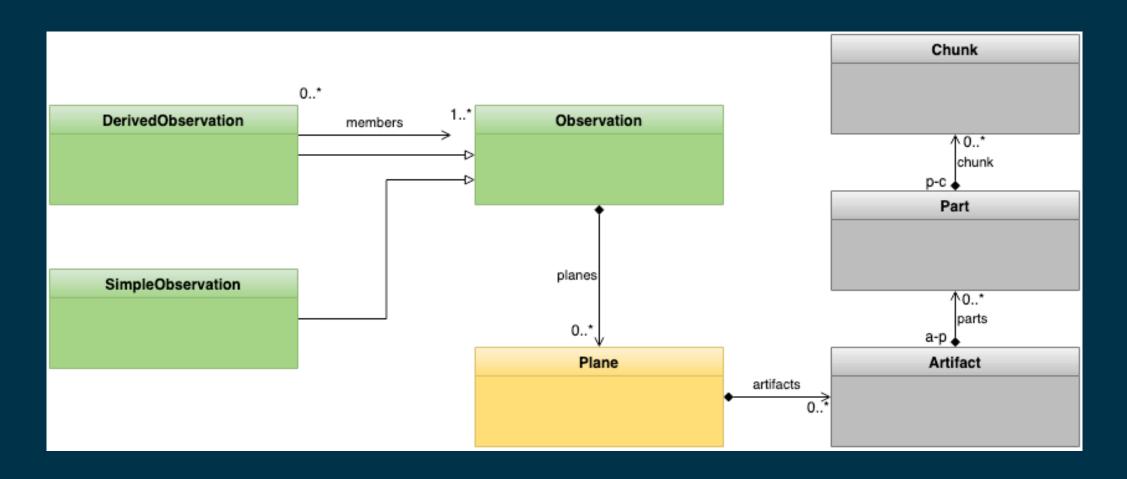


#### Sources:

http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/doc/caom/ https://wiki.ivoa.net/internal/IVOA/InterOpMay2019DM/caom-developments-2019.pdf https://www.opencadc.org/caom2/

## **CAOM Data Model**





#### Sources:

http://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/doc/caom/ https://www.opencadc.org/caom2/

## Synchronization processes



#### Metadata

- Standard mode: incremental harvesting observations.
- Skip: reprocessing failed observations
- Validation: check observations between local and service.
  - Checksums.
  - Last modified dates.
- (FULL): download all observations.

#### **Artifacts**

- Discover: incrementally harvesting artifacts.
- Download: download the associated files.
- Validation: check artifacts between local and service (File checksums, length, content)
- (DIFF): discover missing artifacts (periodic validation).

#### Status Record

- Metadata synchronization statistics.
- Artifact synchronization statistics.
- Last synchronization dates for Observations and Artifacts from different sources.

#### PP Synchronization

- Proposals linked to observations.
- Publications linked to proposals.
- This way, we have publications linked to observations.
- Additional Service offered by STScI for eHST

## **CAOM** highlights



#### Standardization:

- Mature model with 10+ years of operational experience.
- Widely used in CADC (31 different collections) and in MAST.
- Improvements are being implemented (e.g. new kinds of data).
- All partners use the same services/tools to harvest metadata and to synchronize files.
- This ensures common content for all partners to distribute.
- Model is public: <a href="https://www.opencadc.org/caom2/">https://www.opencadc.org/caom2/</a>

### Interoperability:

- ESAC took benefit from this when we harvested HLA metadata from CADC.
- Potentially, metadata from any caom2-ready-mission can be accessed through the same interface.
- Same processes across missions.

#### Confidence:

- Same metadata and files in all mirrors.
- Increase the trust users have in the contents and therefore, in the archives.

# Thank you!

# Questions?

www.esa.int



## Resources



- CAOM datamodel
  - https://www.opencadc.org/caom2/
- OpenCADC
  - https://github.com/orgs/opencadc/repositories
- HST Metadata & data synchronisation
  - https://ui.adsabs.harvard.edu/abs/2019ASPC..523..425D/abstract
- ESAC Science Data Centre
  - https://www.cosmos.esa.int/web/esdc
  - https://www.cosmos.esa.int/web/hst
- HST Science Archives User Interfaces
  - https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html
  - https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/search/?collection=HST&noexec=true
  - http://hst.esac.esa.int/ehst/#home

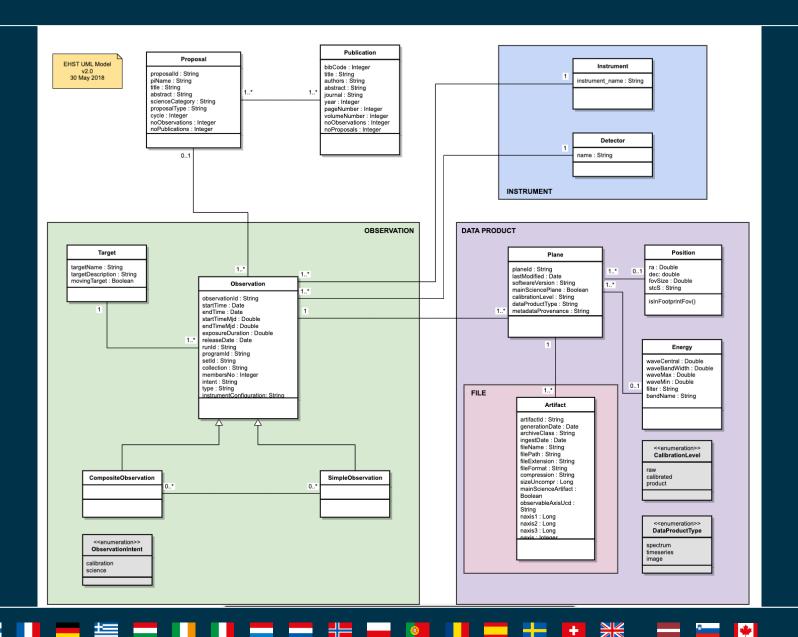
# **Backup Slides**

www.esa.int



## **CAOM** extended with Publications in eHST





# Synchronization: CAOM in eHST and eJWST



