

Data Collections On The Cloud

On this page...

- [The MAST AWS Open Data Bucket](#)
 - [Available Data Collections](#)
 - [How To Use These Cloud-Hosted Data Collections](#)
- [For Further Reading...](#)

The MAST AWS Open Data Bucket

Many data collections at MAST are made available through the [Amazon Web Services Open Data Program](#). The AWS Open Data Program hosts hundreds of datasets on a [wide range of topics](#), astronomical datasets from MAST being one of them. The data are free to access within the AWS cloud and to download out of the cloud onto your computer, should you choose to do so.



An important note: the AWS Public Datasets are not a replacement of the MAST Archive. Data are, *and will always be*, available free of charge from MAST. By distributing data on AWS, MAST is exploring a new kind of archive service: one where the data are **highly available** through bulk, high-speed access in proximity to the vast computational resources of AWS. This enables us to create new services, like [Cloud Science Platforms](#), and gives users a chance to access "big data" far more quickly and reliably than can be done through conventional downloads.



There may be delays between when data products are available at MAST and copies are available on the AWS cloud. HST data might change hourly as reprocessing occurs. TESS data are delivered on a monthly cadence but the data take time to upload into the AWS open data bucket. Generally, updated or new data products are available on cloud within an hour of being available on-premise at MAST; however, some very large products like TESS image cubes might take a day or two to fully copy into the cloud.

Available Data Collections

Data Collection	Description	Update Frequency
GALEX at AWS	Available: All pipeline data products through Galex Release 6+7. Not Available: Data collected during the CAUSE phase of the mission when the spacecraft was operated directly by CalTech for a year at the end of the mission.	Never (mission complete)
Hubble at AWS	Available: All PUBLIC data for ACTIVE instruments (ACS, WFC3, COS, STIS, FGS). Not Available: Data that are still under exclusive access. Data from non-active instruments (FOC, FOS, GHRS, HSP, NICMOS, WFPC1, WFPC2).	Hourly
Kepler at AWS	Available: All light curves, target pixel files, and full frame images. Not Available: Data validation files, engineering data, collateral pixel files, co-trending basis vectors, target catalogs, and any output products from the planet search pipeline.	Never (mission complete)
K2 at AWS	Available: All light curves, target pixel files, and full frame images. Not Available: Data validation files, engineering data, collateral pixel files, co-trending basis vectors, and target catalogs.	Never (mission complete)
TESS at AWS	Available: All light curves, target pixel files, full frame images, co-trending basis vectors, collateral pixel files, full-frame-image-cubes (used by Astrocut/TESSCut to make image cutouts), and all the data validation files <i>except</i> the dvr.pdf and dvm.pdf files. The image cubes created by MAST from the full frame images used by Astrocut to create image cutouts are also available. Not Available: Most engineering data and all target catalogs.	Monthly

How To Use These Cloud-Hosted Data Collections

There are several ways to use these data hosted on AWS, including bringing your code to the data using our Science Platforms or retrieving AWS-hosted data through [astroquery.mast](#). For more information, consult our [Cloud Services Manual](#).

For Further Reading...

- [MAST Data Holdings](#)
- [Cloud Services Manual](#)
- [Enabling Cloud Access With astroquery.mast](#)