

LAMBDA Advisory Group 2026 Report

Executive Summary.....	1
Purpose and value of LAMBDA.....	1
Updates since last report.....	2
Data capacity.....	2
LAMBDA mission statement.....	2
DOIs.....	3
Newly archived data.....	3
Globus and data modernization.....	3
The FORNAX Initiative.....	3
Resources/Graphics.....	4
LAMBDA Advisory Group Members.....	4

Executive Summary

With 1 PB of backed-up storage capacity recently obtained for use by LAMBDA, this now satisfies near future needs. This also satisfies our highest priority recommendation from the last report.

As such, **our new top priority for improving LAMBDA is to create a Globus endpoint.**

Current methods for downloading and accessing data do not scale well for synchronizing large modern CMB data sets between their archive on LAMBDA and the compute clusters on which they are used.

We provide a statement on the purpose and value of the LAMBDA archive to NASA and the wider CMB community. In addition to our recommendation about Globus, we also comment on unifying diverse astronomical data through the FORNAX Initiative being developed by other NASA archives. Finally, we comment on updates and operations at LAMBDA since our last report.

Purpose and value of LAMBDA

NASA's Legacy Archive for Microwave Background Data Analysis (LAMBDA) is a Cosmic Microwave Background (CMB) data center focused on preserving data from some legacy space missions, including foundational CMB data from the Cosmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP). Beyond preserving the data, LAMBDA serves as an interactive and user-accessible specialized repository for scientists studying the large-scale properties of the universe. To support the needs of research scientists, LAMBDA

also hosts several additional closely related data products, including but not limited to suborbital CMB data. LAMBDA distributes specialized open-source analysis software and data visualization utilities for the use of the scientific community.

LAMBDA operates as part of NASA's High Energy Astrophysics Science Archive Research Center (HEASARC) at the Goddard Space Flight Center, under the NASA Science Mission Directorate (SMD) Astrophysics Science Division. A User's Group with rotating membership provides guidance and suggestions to the LAMBDA staff from the scientific user's perspective.

Archives like LAMBDA are critical for enhancing and preserving the value of NASA science missions, both past and future. For past missions like COBE and WMAP, LAMBDA allows their irreplaceable data to continue to impact new research. For future missions, LAMBDA is a crucial piece of NASA's open science strategy, providing transparent data access to other researchers as well as the general public. Preservation of high quality raw data is also important for training of Artificial Intelligence (AI) and Machine Learning (ML) models.

Updates since last report

Data capacity

HEASARC has obtained an additional 2 PB of backed-up storage capacity, of which LAMBDA can use up to 1 PB. The LAMBDA Advisory Group applauds this achievement, as increasing capacity to meet imminent demand was the highest priority recommendation in our previous report. While available capacity now meets expected demand for the near term, investment into growth will remain a priority in the future, as CMB data volumes continue to grow as well.

LAMBDA mission statement

As per the recommendation from a previous meeting of the LAMBDA Advisory Group, the [published mission statement](#) has been amended to include new text (highlighted below) in its second line:

- Develops and maintains data access and analysis tools **to enhance data accessibility**.

While the former fourth line has been removed:

- ~~• Carries out data-intensive processing of vital importance to NASA's CMB community~~

We commend these changes. The scope of CMB analysis tools and data processing have grown well beyond what LAMBDA can support. The changes reflect a focus of resources towards servicing LAMBDA's primary missions of data archiving and access.

DOIs

It is now possible to associate Digital Object Identifiers (DOIs) with data on LAMBDA. This helps to persistently identify published data products, of utility to science teams, publications, and funding agencies.

Newly archived data

LAMBDA continues to ingest new data. Since the last report this includes updates from Planck (third party), ACT, CLASS, and SPT-3G.

Globus and data modernization

Consistent with the recommendations of the previous LAG report, there continues to be a strong need to modernize LAMBDA's data capacity and accessibility. As data volumes grow from major CMB experiments, improving accessibility is critical for the research community. Both Globus and FORNAX offer modernization pathways. The LAG believes implementing a Globus endpoint for LAMBDA continues to be a high priority. While FORNAX provides value for education, outreach, and quick exploration of data products if co-hosted with LAMBDA (see section on FORNAX), Globus is vastly more useful for core scientific research since it enables syncing data products to centers where production runs of data analysis will happen.

This Globus infrastructure is particularly critical for managing and distributing massive data collections, such as ACT and Simons Observatory “depth-1” (per-epoch or time-domain) datasets as well as large suites of simulations. Depth-1 data is inherently harder for external researchers to reproduce than simulations: it requires expensive maximum-likelihood map making. Therefore, hosting these time-domain datasets on a permanent archive represents high value to the community.

At the moment, the hyperlink and zip-file-based data hosting is a severe limitation, and so LAG once again highlights Globus data access as one of the key recommendations of this report.

The FORNAX Initiative

The [FORNAX initiative](#) is bringing together a diverse set of astronomical data sets and analysis tools to facilitate future discoveries. While the primary focus of FORNAX will be supporting analysis of data from current NASA missions (e.g., JWST, SPHEREx, EUCLID, Roman), this is an exciting opportunity to enable joint analyses and cross-correlation analyses with archival data sets, including some of those hosted on LAMBDA. The most compelling way to support these joint analysis efforts would be to “enhance data accessibility” by co-hosting the LAMBDA data sets with the FORNAX servers. This would substantially reduce the barrier for users who

would otherwise be required to make their own copies of data from LAMBDA on their own accounts, which is a slow process that uses significant data storage for each user.

We advocate for pursuing co-hosting of LAMBDA data with the FORNAX servers to facilitate joint analyses of CMB data with the exciting new satellite data sets.

Resources/Graphics

While LAMBDA's main purpose and value is hosting and preserving CMB data, we support continued maintenance of other sections of the website. In particular, the "Resources/Graphics" section provides a useful starting point and summary for junior researchers who are the next generation of users of the datasets hosted by LAMBDA. We estimate the current depth and breadth of the material to be sufficient for this purpose.

Many of these resources have been contributed by researchers and educators outside of LAMBDA. Like the tools section reviewed in the last report, this allows LAMDA to usefully serve as a community index, without needing to develop all of the material itself.

LAMBDA Advisory Group Members

Steven Benton - Chair	Princeton University
Charles Bennett	Johns Hopkins University
Sebastian Bocquet	Ludwig Maximilians Universität
Mathew Madhavacheril	University of Pennsylvania
Michael Niemack	Cornell University