
NASA
HEASARC
The HEASARC Office
Alan Smale, Director
NASA Goddard Space Flight Center
Code 660.1

The HEASARC

The High Energy Astrophysics Science Archive Research Center

- Founded in 1990, the HEASARC is NASA's primary archive for High Energy Astrophysics and Cosmic Microwave Background data, supporting the broad science goals of the Physics of the Cosmos theme.
- Curates data from NASA, ESA, and JAXA space missions and associated suborbital experiments dealing with electromagnetic radiation from extremely energetic phenomena, ranging from black holes to the Big Bang.
- The HEASARC archive is now in excess of 90 TB, and is increasing in size by 10 TB/yr; in the coming 4 years will ingest data from 19 operating HEA and CMB missions/experiments, while serving data from 39 missions which are no longer operational.

Maintains mission datasets. Keeps them secure and accessible

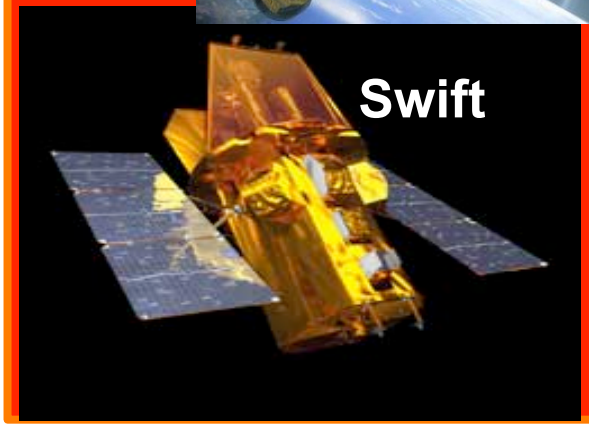
Much more than just an archive, the HEASARC enables world-class science

Operating Missions

Large Missions

Explorers

HEASARC



High Energy Astrophysics Data Holdings

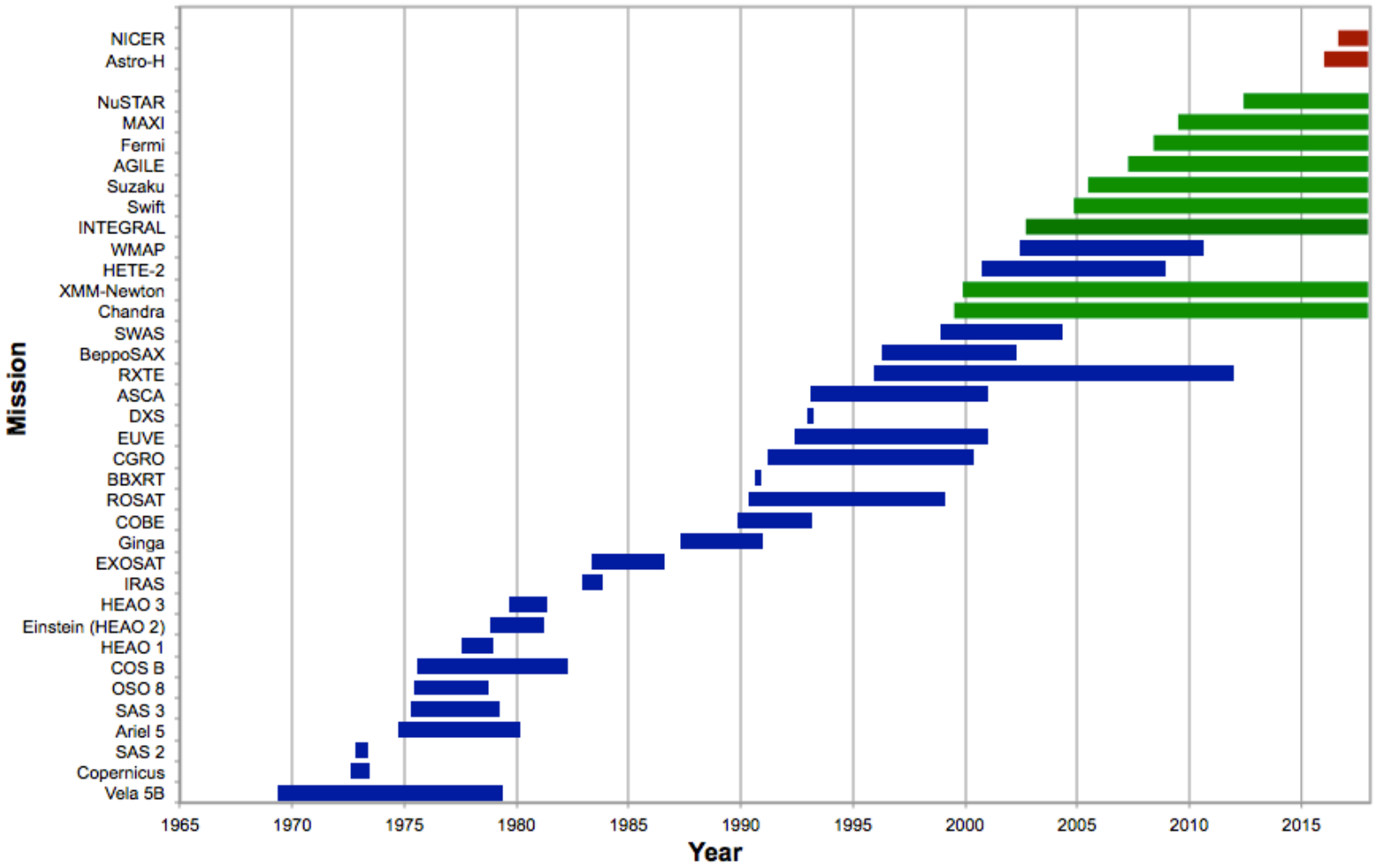
HEASARC

Currently-active, future, and prospective X-ray/gamma ray mission archives include:

Name	Archive Role	Energy range	Status/Launch	Characteristics
Chandra	Mirror	0.5–10 keV	Operational	High spectral resolution imaging, large effective area, and high resolution grating spectroscopy.
XMM-Newton	Mirror	0.1–15 keV Optical and UV	Operational	Very large collecting area with imaging capability and high resolution grating spectroscopy.
INTEGRAL	Mirror	15 keV –10 MeV	Operational	High spectral and spatial resolution. Simultaneous gamma-ray, X-ray, and Optical observations.
Swift	Main	0.3–150 keV Optical and UV	Operational	Autonomous spacecraft response to GRB. Accurate position estimates within minutes. Data immediately in the public archive.
Fermi	Main	20 MeV–200 GeV	Operational	Large field of view. Sensitivity 50 times more at 100 MeV than EGRET. GRB monitor.
NuSTAR	Main	5–80 keV	Operational	First hard X-ray observatory with good spatial resolution.
Hitomi	Main	0.3-600 keV	Archiving	High-resolution and wideband X-ray spectroscopy.
NICER	Main	0.2-12 keV	Early 2017	Precise pulsar timing
IXPE	Main	-	2020	X-ray Polarimetry - In Phase A Downselect
PRAXyS	Main	-	2020	X-ray Polarimetry - In Phase A Downselect

High Energy Astrophysics Data Holdings

HEASARC



HEASARC's Value to Scientists

HEASARC

The HEASARC provides a standard analysis system for many missions, improving the efficiency of scientists by reducing time spent learning new tools:

- Complete mission datasets in standard formats → **FITS, OGIP standards**
- Data and metadata that are logically arranged, easy to retrieve, ready to analyze → **Browse, Xamin, Hera, FTP**
- Associated calibrations in standard formats → **CalDB**
- Standard, mostly mission-independent analysis tools → **HEASOFT, including FTOOLS, XSPEC, XIMAGE etc**
- Readily available in-house expertise → **HEASARC Scientists, both HEA and CMB**

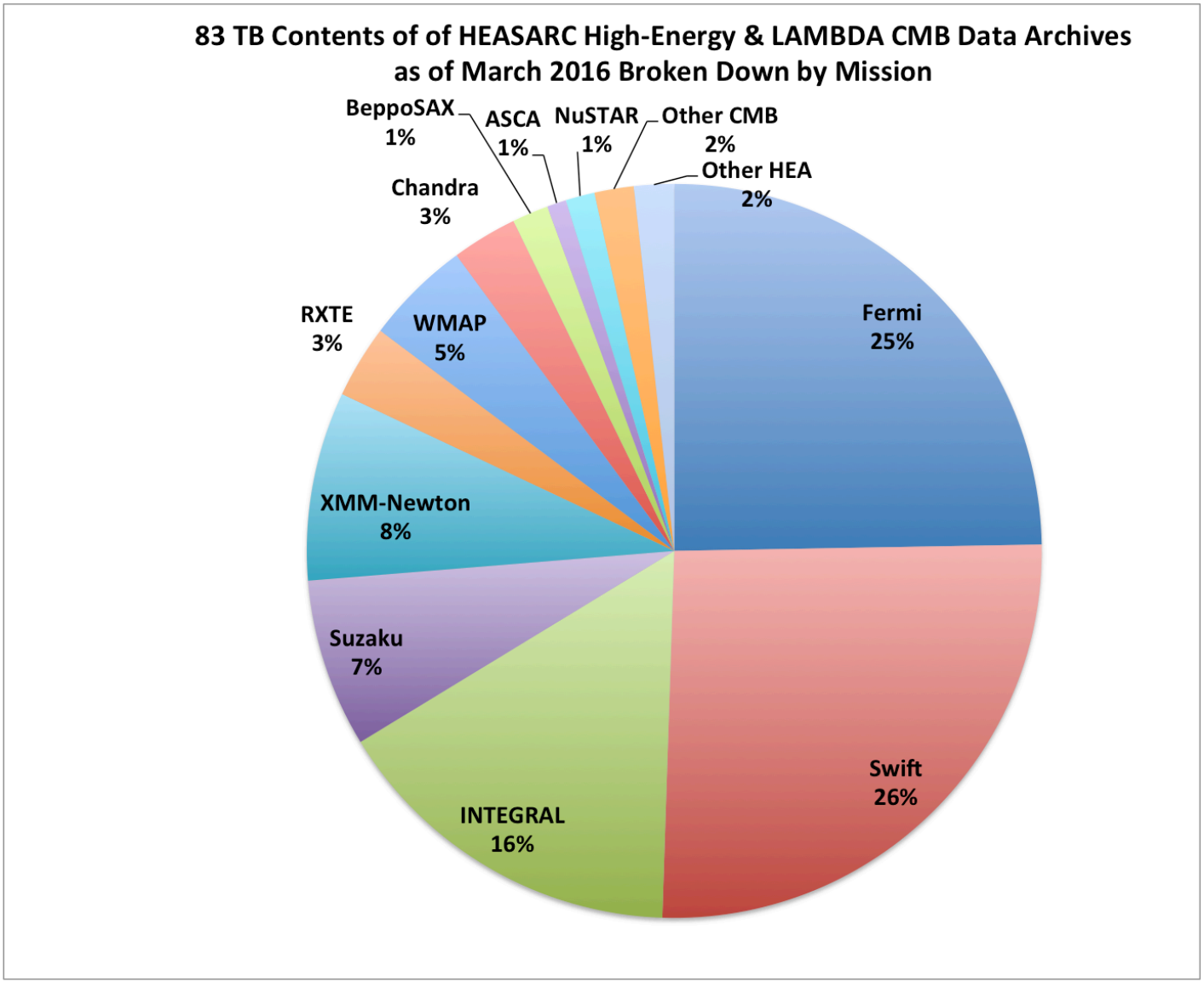
Many Missions – One Analysis System

HEASARC's Value to NASA

HEASARC

- The HEASARC archive enhances the return on NASA's investment in its missions and technology
  - Easy access to data from current missions
  - Access to data from older missions in modern formats
  - Better understanding of cosmic phenomena by comparing data from new and old missions. Possible to study variability of sources over nearly the entire history of high energy astrophysics
- A standard analysis system for multiple missions improves the efficiency of scientists by reducing time spent learning new tools – in NASA's interests too, since NASA funds scientists
- **A standard and easy-to-adopt framework for data archiving, pipeline processing, data analysis, calibration, and proposal entry and management reduces development costs and risks for new missions**

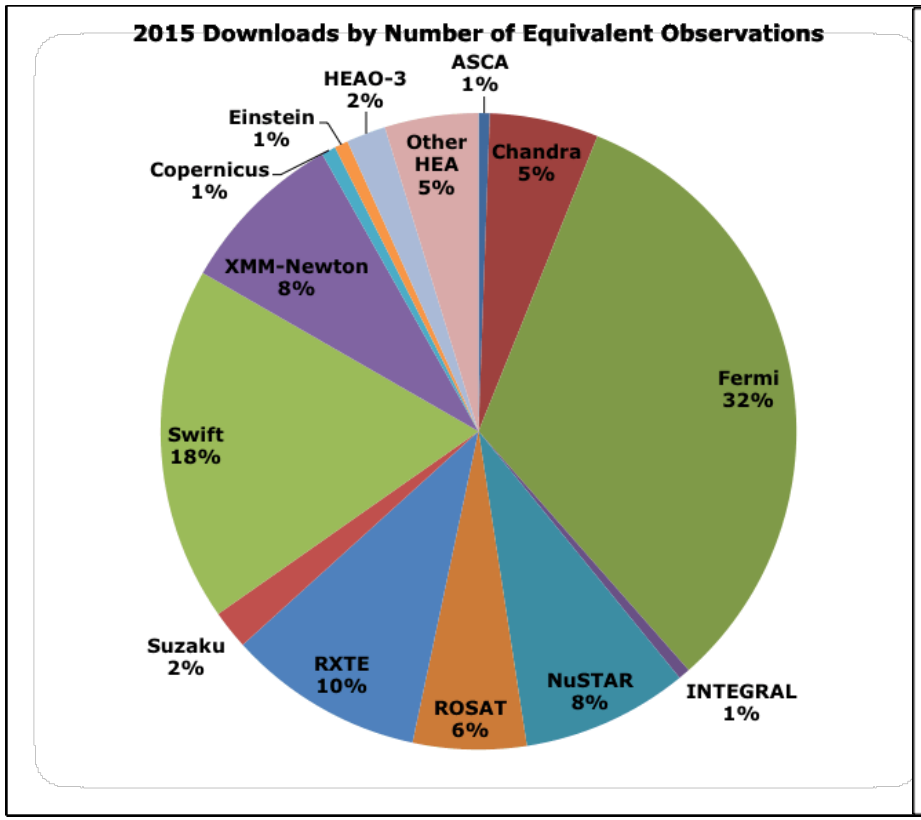
HEASARC Contents

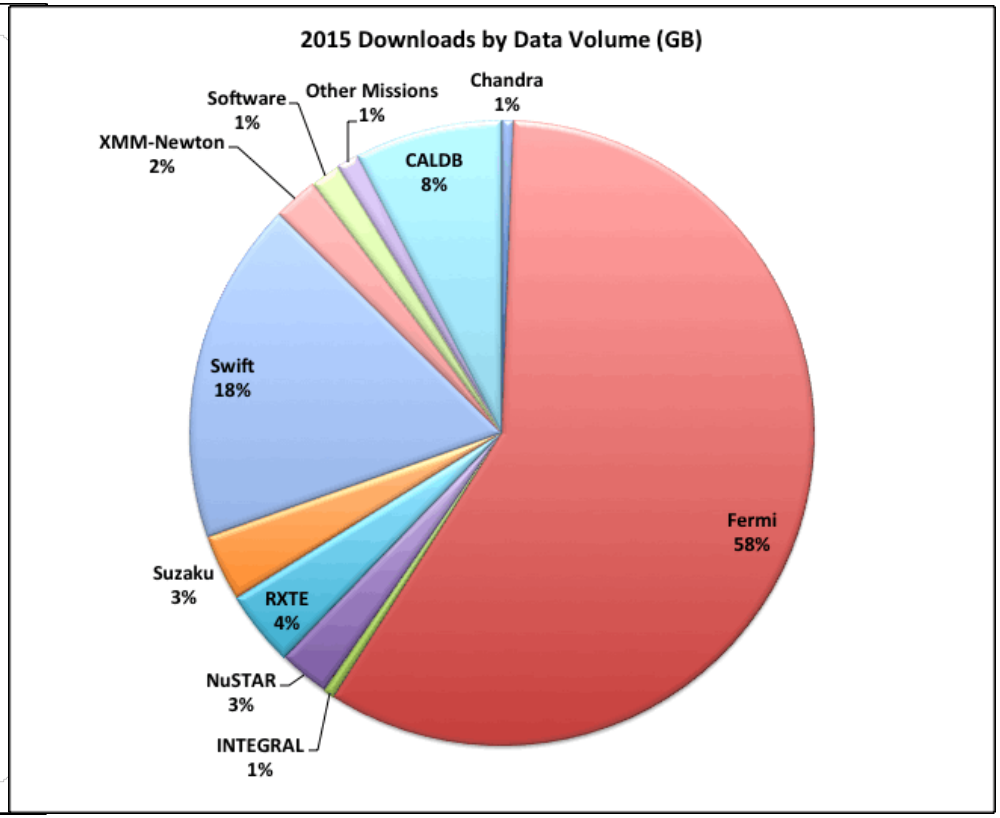


Data Downloads

HEASARC

Fermi (2008-), Swift (2004-) , RXTE (1995-2012), and XMM-Newton (1999-) are the most popular data sets by download volume, but data from older missions such as ROSAT (1990-1999) are still popular.

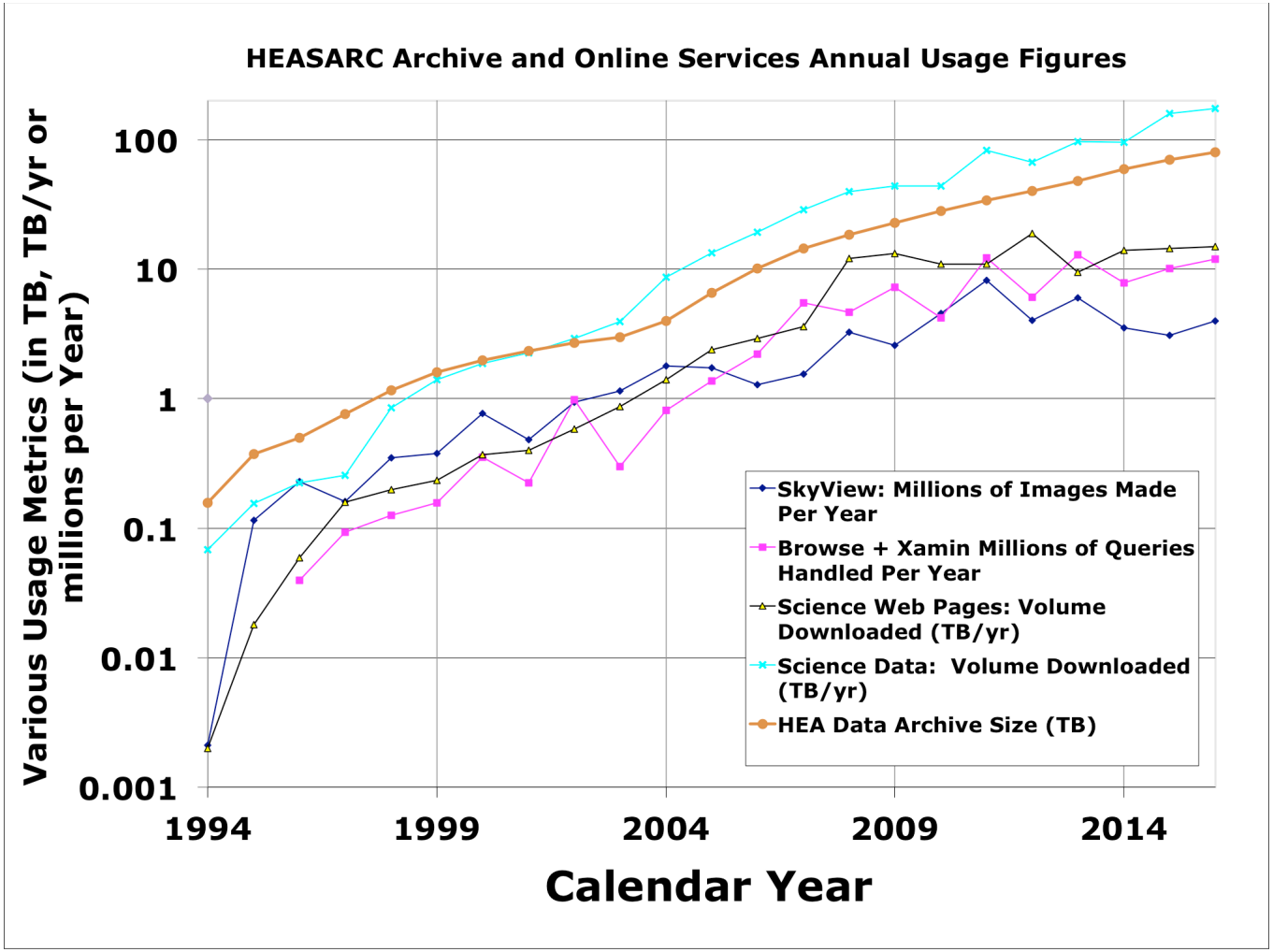




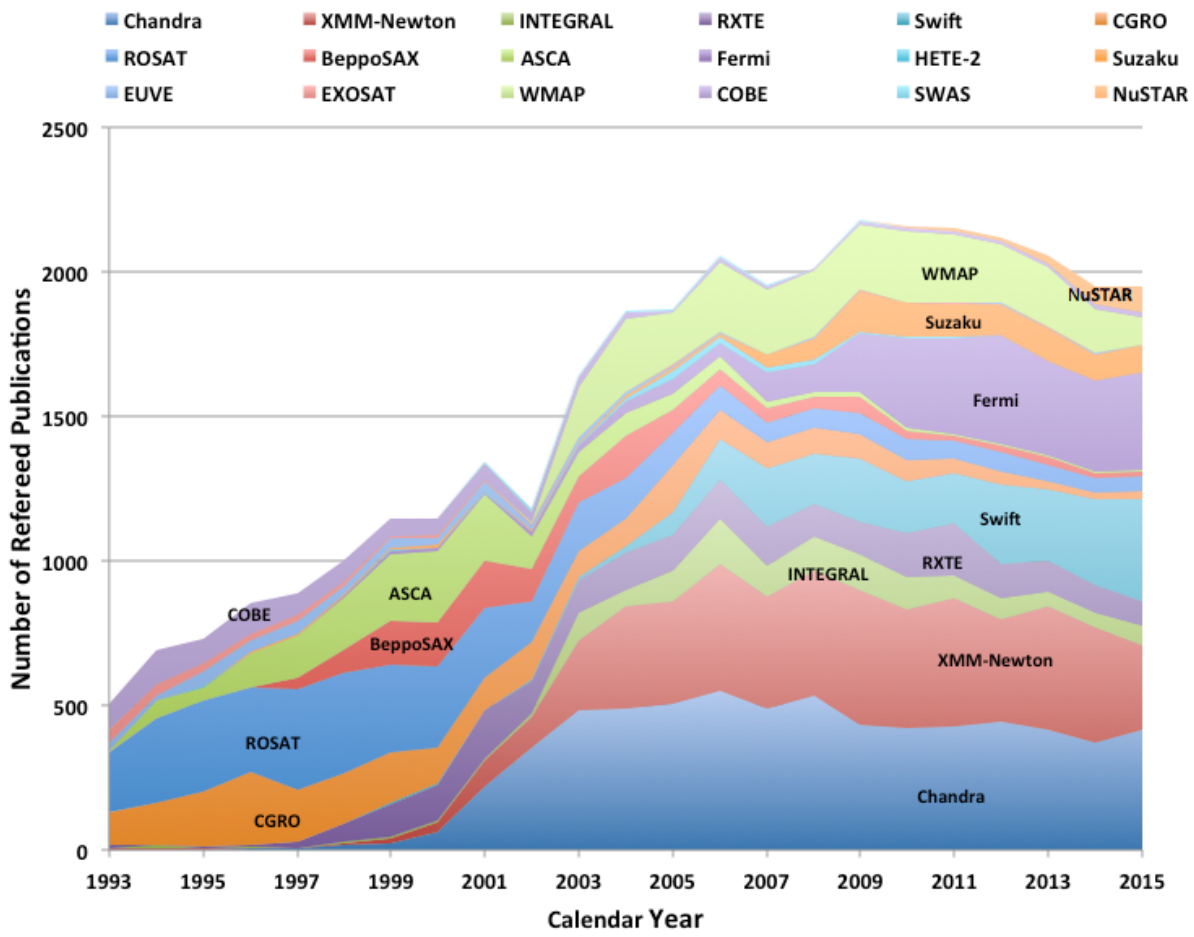
Data Content and Growth

HEASARC

The HEASARC archive has grown by an order of magnitude in the last decade in data volume and science information served per year.



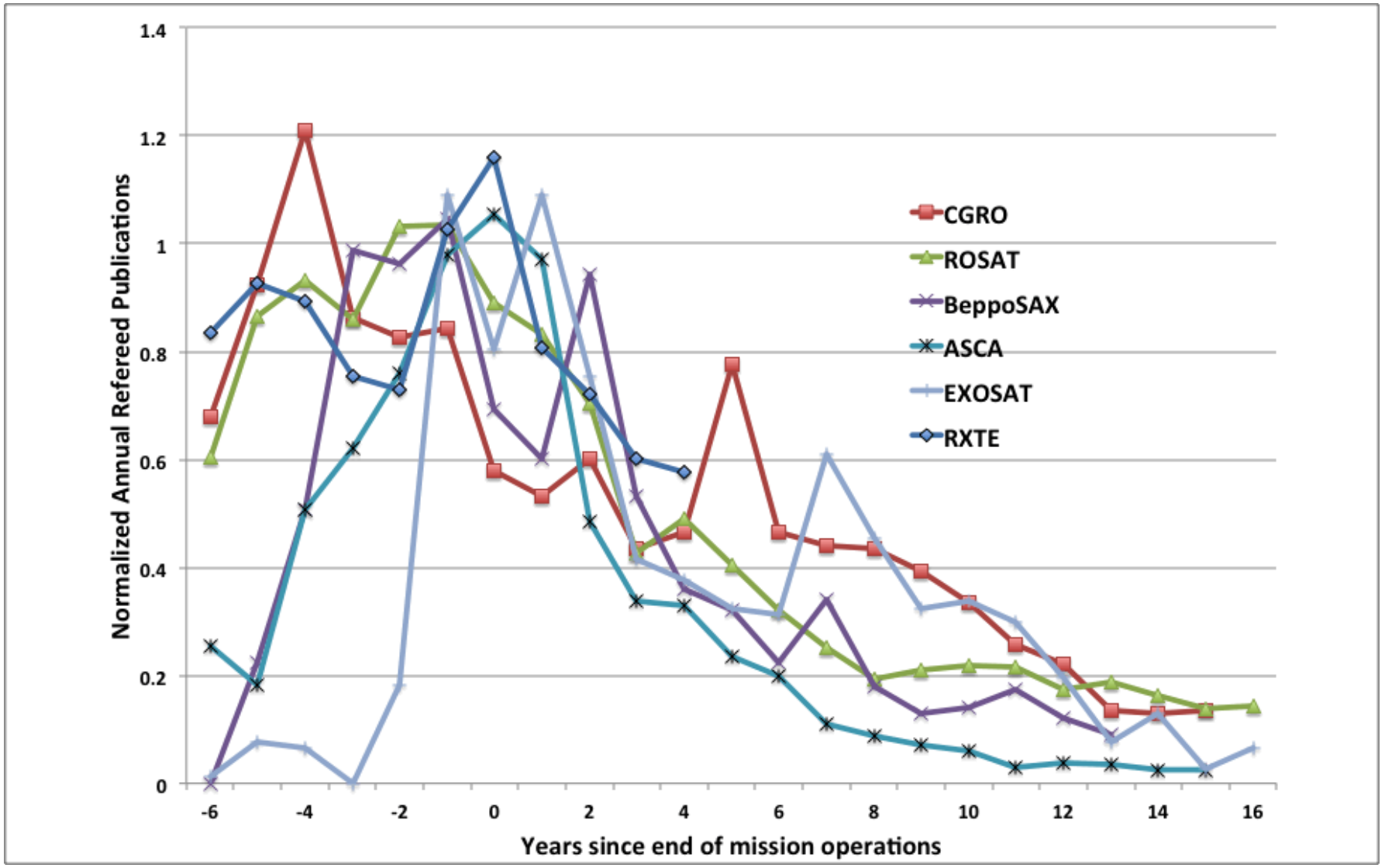
Refereed Publications using HEASARC Data



Papers written using HEASARC data comprise ~10% of the total astronomical literature and include some of the most highly cited papers in the field.

Science Legacy

HEASARC



Data Archiving and Analysis: Software Support

HEASARC

- In addition to data processing and archiving support for active and legacy missions, the HEASARC provides comprehensive software and calibration support:
  - HEASARC programmers coordinate 1-2 major public releases per year of the HEASoft data analysis software package (6.17 – Aug 2015, 6.18 – Feb 2016).
  - HEASoft contains ~3M lines of code in 750 individual analysis tasks for high energy missions supported by the HEASARC, as well as general analysis of astronomical data from other missions.
  - Includes XSPEC, XRONOS, XIMAGE.
  - HEASoft is open source – users have access to the source code.
  - In 2015 approximately 5,000 registered individual or institutional users installed the HEASoft package on their computers.
  - HEASARC's CALDB – calibration database – is updated regularly (a total of 109 updates over the past four years, of the Chandra, Fermi, NuSTAR, Suzaku etc CALDB areas.

LAMBDA

HEASARC

- Legacy Archive for Microwave Background Data and Analysis: a hub for the CMB community. Established in 2002.
- Since 2008, a thematic archive for CMB research within the HEASARC.
- LAMBDA maintains CMB subject matter experts and leadership, and a distinct website. Its outside advisory panel provides advice, acts as liaison to CMB research groups and users.
- Permanent archive for NASA mission data from WMAP (1, 3, 5, 7, and 9-year), COBE, IRAS, SWAS. Links to IRSA Planck data.
- Permanent archive for suborbital CMB mission data. Suborbital missions have high-impact science and are sites of significant NASA technology investment for next-generation space missions.

Name	Years	Multipoles	Freq. (GHz)	Platform	Data
COBE	1989-1993	2–40	32–90	Satellite	
MSAM	1992-1997	69–362	150–650	Balloon	
QMASK	1993-1996	40–335	26–46	Suborbital	
MAXIMA	1995–1999	50–700	150–420	Balloon	
QMAP	1996	30–850	30–140	Balloon	
BOOMERanG	1997–2003	25–1025	90–420	Balloon	
Archeops	1999–2002	10–700	143–545	Balloon	
DASI	2001–2003	200–900	26–36	Ground	
ARCADE	2001–2006	N/A	3–90	Balloon	
ACBAR	2001–2008	60–2700	150–274	Ground	
WMAP	2001–2010	2–1200	23–94	Satellite	
VSA	2002–2004	130–1800	26–36	Ground	
CAPMAP	2002–2008	500–1500	40, 90	Ground	
CBI	2002–2008	300–3000	26–36	Ground	
QUaD	2005–2010	200–2000	100, 150	Ground	
BICEP	2006–2008	21–335	100–220	Ground	
SPT	2007–2011	650–9500	95–220	Ground	
QUIET	2008–2010	60–3500	40, 90	Ground	
ACT	2008–2011	500–10000	148–277	Ground	
Planck	2009–date	2–2500	30–857	Satellite	
BICEP2	2009–2012	21–335	150	Ground	
ACTpol	2011–date	250–8395	148	Ground	
SPTPol	2011-date	650–9500	150	Ground	
POLARBEAR	2012–date	50–2000	150	Ground	
Keck	2015–	21–300	150	Ground	

Products on LAMBDA before the 2011 Senior Review
Updated since SR11
New Products after SR11

LAMBDA: Recent Releases and Future Data

HEASARC

Since April 2015:

- **ACTPol:** CMB polarization likelihoods
- **BICEP/Keck:** October 2015 release
- **QUIET:** CMB polarization maps in galactic fields
- **SPTPol:** 2015 BB bandpowers
- Archival work: Supported an outside investigator by recalling DIRBE warm mission data from the tape archive. Added previously unavailable FIRAS raw data.
- 3D galactic dust model (Green et al. 2015); AKARI All-sky FIR Maps; and IRIS All-sky IR Maps.
- **SPIDER:** analysis + data release pending. LAMBDA has exclusive distribution of survey region in footprint tool.
- **QUIET:** LAMBDA has exclusive CMB field coverage maps in footprint tool.

Upcoming, with LAMBDA identified as archive:

- **Suborbital (Balloon experiments):**
  - EBEx, PIPER, SPIDER.
- **Suborbital (Ground-based):**
  - B-mode experiments at large angular scales: ABS, BICEP/Keck, and CLASS.
  - Polarimeters sensitive to smaller angular scales: ACTPol, PolarBear/Simons, SPTpol/SPT 3G.
- **Future Explorer-class Missions:**
  - JAXA's LiteBIRD
  - PIXIE
  - DARE
- **Future Probe-class Missions**: will seek archiving involvement.

LAMBDA holdings will be critical in the design and implementation of future space polarimetry/spectral missions.

HEASARC Response to the 2015 Senior Review

HEASARC

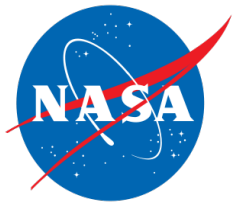
Challenges across all archives

- *"The infrastructure and the technological approaches that are being used will certainly be obsolete at the end of the next 4-5 year review cycle."*
  - The HEASARC stays current with new technologies, agrees that it is critical to assure our infrastructure is refreshed regularly. We were an early adopter of Virtual Machine Environment; major refresh in 2014; regular hard disk storage upgrades; budget contains provisions for future upgrades to serves, data storage, other infrastructure necessary to support a leading-edge astrophysics archive.
- *"Network bandwidths available to the data centers will soon be two generations behind the current standard for research internet."*
  - The HEASARC web servers are currently connected to the public zone web using 1 Gb/sec Network Interface Controllers (2 of them). Backend network connection to our disk storage is 10 Gb/sec. As of this time last year, Goddard CNE planned to replace existing data center switch with several 10 Gb switches. Currently, network management is transitioning to MSFC; CNE funding for infrastructure upgrades has been correspondingly reduced.
- *"Data centers need to raise concerns about sustainability where they exist, regardless of budgetary constraint."*
  - Fear not, we won't be shy…

BDTF Questions

HEASARC

- *"What are the processes for planning for future (5-10 yr) capabilities of the HEASARC? How and from whom do you gather input for this planning process and where does input typically come from? What new features have highest priority?"*
  - NASA Archives Programmatic/Senior Review process. Prioritized Archive Objectives; HQ PPBE process; associated discussions with Program Officers.
  - User Group input; ADEC meetings; ADASS meetings; other astronomy meetings (informal discussions with community).
  - Intra-Goddard discussions with Code 700; SEDVME group; special working groups (see next slide). Input from Office of the CIO.
  - Highest priorities: Support of missions in operation, development, formulation; software innovations; Python; Cloud; continuing vigilance on data security/ integrity; increasing interoperability via VO services.
- *"What feature(s) of your service would you like to stop performing? How do you gather input for making such decisions and where does such input typically come from? What is preventing you from stopping?"*
  - No features we wish to terminate.
- *What steps you are taking to make your data interoperable with allied data sets from other data sites in and out of NASA? How do you find allied data sets and what criteria make data sets candidates for enabling interoperability?*
  - See next slides.



HEASARC in the Cloud

HEASARC

- The Cloud represents a paradigm shift in how data are stored and accessed. We may soon be able to take advantage of the Cloud for HEASARC data archiving and data (and software) access.
  - Prices for full HEASARC implementation in the Cloud have become competitive with local installation, although security/data integrity concerns still exist, and the devil is in the details.
  - HEASARC is a part of the Goddard Private Cloud Working Group (consisting of members from Codes 400/500/600/700, driven from 600), and the HEASARC is one of the Use Cases under study to analyze governance requirements and cost models.
- The HEASARC envisages a gradual transition:
  - Transfer of deep archive copy from Iron Mountain to [Provider].
  - Support for key services using the Cloud at relatively low cost.
  - HEASARC backup using the Cloud: significant cost but turnkey access.
  - Full HEASARC services in the Cloud.
- We will proceed with the implementation of the deep archive copy, and take further steps as soon as practicable.

Inter-archive Cooperation: The ADEC

HEASARC

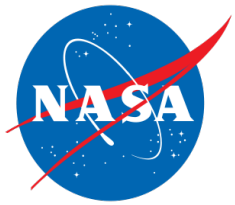
- NASA Astrophysics supports an integrated system of science data archives.
- HEASARC, MAST, IRSA/ NED etc are required to “compete” in the Senior (Programmatic) Review, but collaboration is by far our most usual mode.
- The archives voluntarily share information, seek out synergies and new opportunities, ensure interoperability, promote multi-wavelength and multi-mission data and software standards.
- The ADEC – Astronomy Data Centers Executive Committee. Informal organization of archive leads and selected designees. Includes members from Chandra, Fermi, NSSDC… Rotating chairship.
- The ADEC holds monthly telecons; healthy email exchange; face-to-face discussions at astronomy meetings, ADASS etc. Most recent dedicated meeting was an Archives Coordination Meeting hosted by Caltech in November 2014 to discuss our responses to the then-upcoming NASA Archives Senior Review…

Inter-archive Cooperation: NAVO

HEASARC

NASA Astronomical Virtual Observatories

- Sustaining the infrastructure of the VO, building on the framework of software and standards developed under the joint NASA/NSF Virtual Astronomical Observatory activity.
- A partnership between NASA's astrophysics domain archives: HEASARC, MAST, IRSA, NED. Established October 2014. Funded through archives' in-guide budgets, plus modest augmentations awarded after proposal review in 2013-2014.
- We encourage participation by other NASA and US astronomy institutions where feasible/appropriate.
- NAVO Mission: to facilitate the maximum science return for NASA astronomy data and resources using internationally agreed-upon standards. We:
  - Develop comprehensive and consistent access to NASA data through VO protocols.
  - Represent NASA and US interests in the development of astronomy data standards – continues the dialog with other VO groups and organizations.
  - Maintain the infrastructure needed to discover and exploit VO services in the US and worldwide.
- Project Scientist: Tom McGlynn (HEASARC); IVOA Liaison: Bruce Berriman (IPAC).



NAVO and the HEASARC

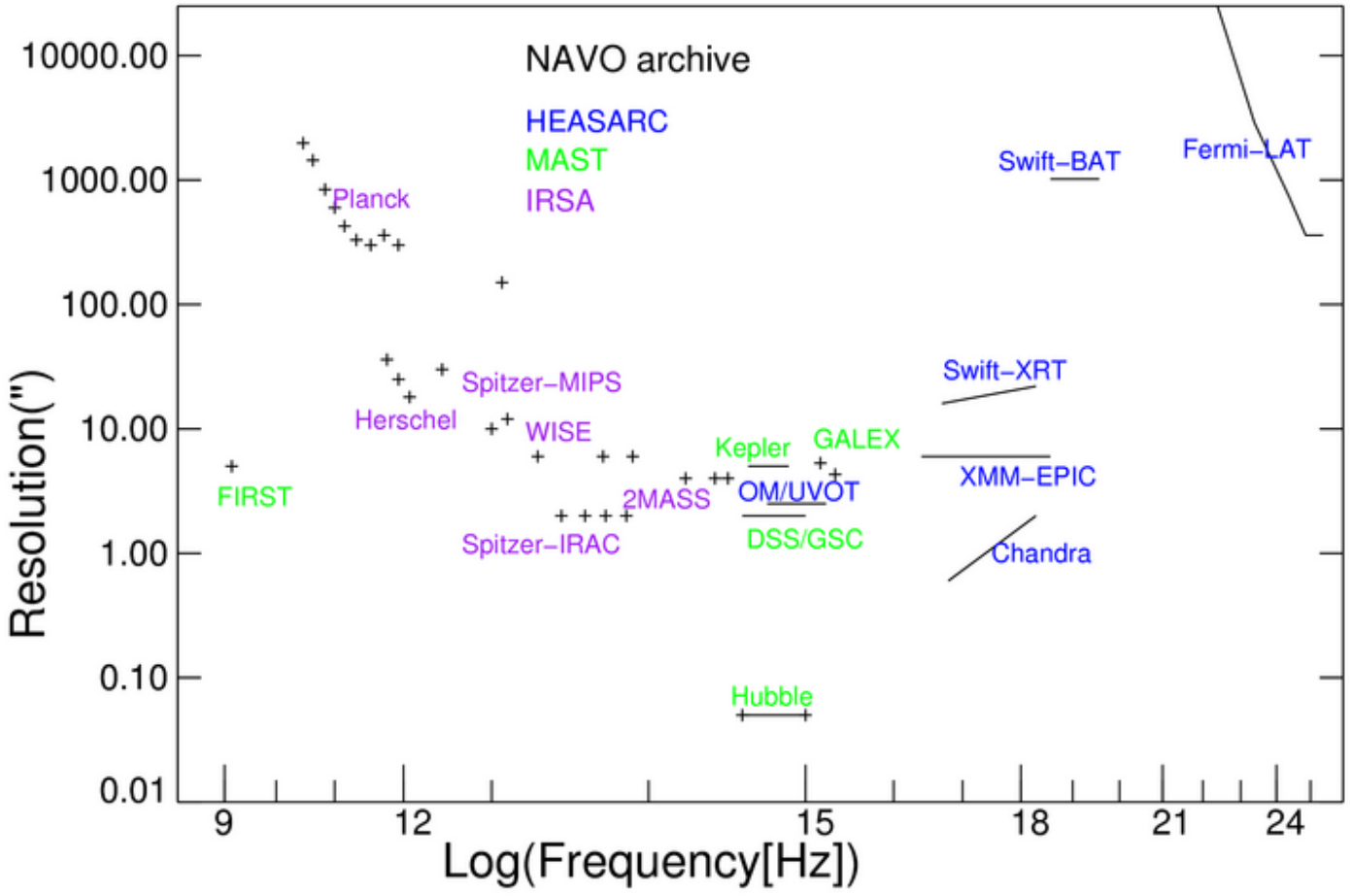
HEASARC

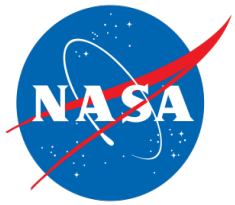
- "Under the hood" (i.e. invisible to most users), the VO infrastructure is an important means for discovering and retrieving NASA astrophysics data.
  - At the HEASARC, most user queries come through VO protocols; MAST's registry of VO resources provides a detailed index of astronomical resources throughout the world; IRSA offers millions of infrared images and billions of rows of catalog data; NED provides object parameters, images, and spectral energy distributions for millions of extragalactic sources, all through VO protocols.
  - HEASARC has VO-accessible table queries, images and spectra for most missions via VO TAP, SIA, and SSA protocols. Goals: fully searchable standardized metadata for all NASA astrophysics missions; cross-archive access in HEASARC/IRSA/MAST/NED tools (DataLink, ObsCore).
  - In addition to our work implementing VO protocols and our overall management role, the HEASARC also provides monitoring and validation tools which ensure data remain available and that services are working. A monitoring service checks hourly whether sites are up; validation service checks monthly for conformance to VO protocols; notification service enables providers to publish info about anticipated or current issues affecting the availability of services. All info available via public web sites and services intended to be used programmatically.

NAVO

HEASARC

Enables a consistent and comprehensive view of NASA astrophysics data.

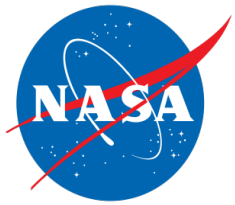




HEASARC Summary

HEASARC

- The HEASARC effectively serves NASA and the broader science community, and promotes first-class science, by:
  - Providing a consistent and coherent archive for dozens of HEA and CMB missions.
  - Maintaining a comprehensive multi-mission analysis environment that facilitates multi-mission science.
  - Developing and adopting community data, software and standards.
  - Permanently maintaining data and scientific expertise.
  - Enhancing the science productivity of numerous past, present, and future missions.
  - Collaborating closely with other NASA Archives to enhance interoperability and promote common approaches.



Supplementary materials

NAVO Work Areas

HEASARC

1. VO Infrastructure

1.1 Registry

1.2 Monitoring and Validation

1.3 VAO Legacy

2. Publishing Data

2.1 Existing Interfaces

2.2 Developing a Common NASA Astrophysics VO framework

2.3 Access to Large Tables

3. Community Engagement

3.1 IVOA Liaison

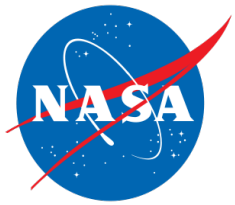
3.2 Registry updates

3.3 Community support

4. Management

4.1 General Infrastructure and Meetings

4.2 Project Scientist Effort



HEASARC Advisory Groups

HEASARC

HEASARC Advisory Group

- Greg Madejski (SLAC, Chair)
- Sarah Gallagher (U. Western Ontario)
- Tesla Jeltema (UC Santa Cruz)
- Paul Ray (NRL)
- Colleen Wilson-Hodge (MSFC)

- May increase group size, if additional expertise is needed.
- Several members request we ask for feedback on specific issues, rather than merely 'distributing information.'
- All but one of the above group read the HEASARC PR proposal (one was on sabbatical), and were in broad agreement with our goals and future objectives.

LAMBDA Advisory Group

- Brian Keating (UC San Diego)
- Ted Bunn (U. Richmond)
- David Chuss (Villanova U.)
- Joanna Dunkley (Oxford U.)
- Gary Hinshaw (U. British Columbia)

- Senior Review 2011 recommended that LAMBDA and HEASARC maintain separate advisory groups.
- We plan to extend this LAMBDA group in the coming years to include members of additional CMB research teams.