THE EVOLUTION OF CHANDRA DATA USAGE OVER TIME AS SEEN THROUGH PAPERS:

trends, insights and lessons learned

Raffaele D'Abrusco, for the ArcOps team
Center for Astrophysics

The Chandra bibliography

The Chandra Data Archive (CDA) curates the Chandra Bibliography, a comprehensive list of scientific and technical publications related to Chandra. Publications are sourced from the Astrophysical Data System (ADS) and classified by expert bibliographers. Each paper is labeled based on various criteria, such as the significance and type of Chandra data usage and the nature of the publication (scientific or technical). Additionally, every paper is linked to the specific Chandra observations and/or Chandra Source Catalog (CSC) data products utilized.

Goals of the Chandra bibliography

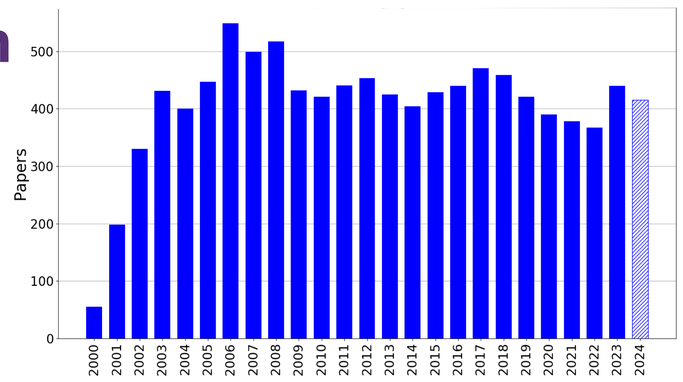
- Quantitatively assess the scientific impact of the mission
- Serve as a research tool for the Chandra community
- Investigate changes in Chandra data usage over time
- Highlight the importance of archival data for scientific research

An Incredibly Impactful Mission

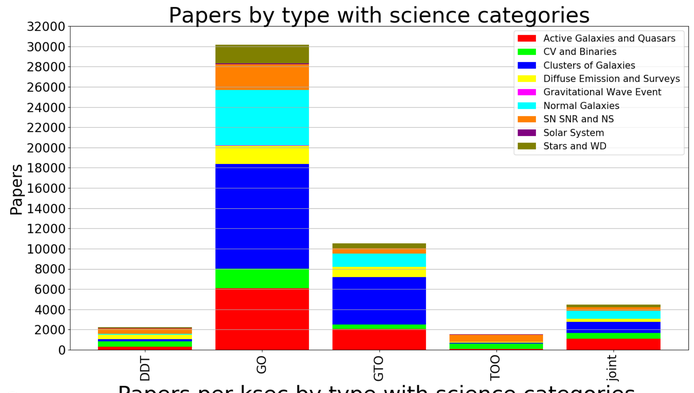
Recently, we classified the 10,000th Chandra Science Paper (CSP), representing publications that would not have been possible without Chandra. This extensive and unique collection of publications provides valuable insights into the scientific impact and productivity of the mission, as well as how these have evolved over time. Specifically, the granular linking of CSPs to the distinct Chandra datasets used allows for an in-depth investigation of bibliographic trends across all metadata.

Contact: arcops@cfa.harvard.edu

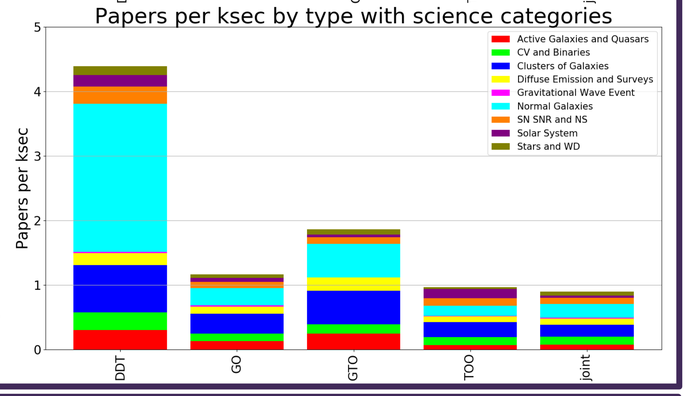
Public Chandra bibliometrics stats/plots:
https://cxc.cfa.harvard.edu/cda/bibstats/bibstats.html



- Chandra's scientific output, measured by the number of Chandra Science Papers (CSPs), has remained stable over the past 15 years, averaging ~ 400 CSPs per year.



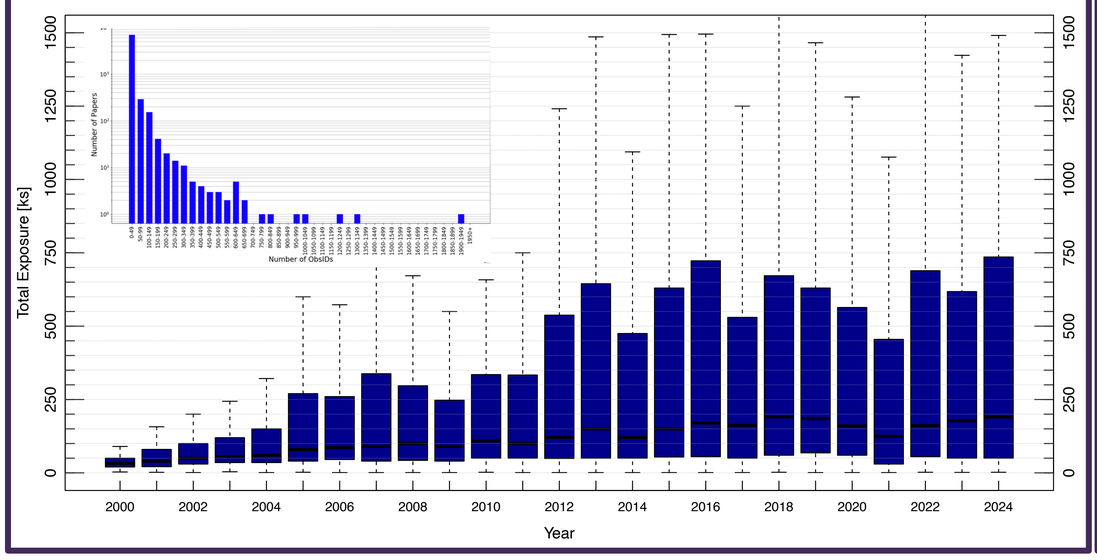
- While GO observations are used in the largest absolute number of CSPs, DDT proposals are the most productive per ks, outperforming all other types of proposals.



Data complexity

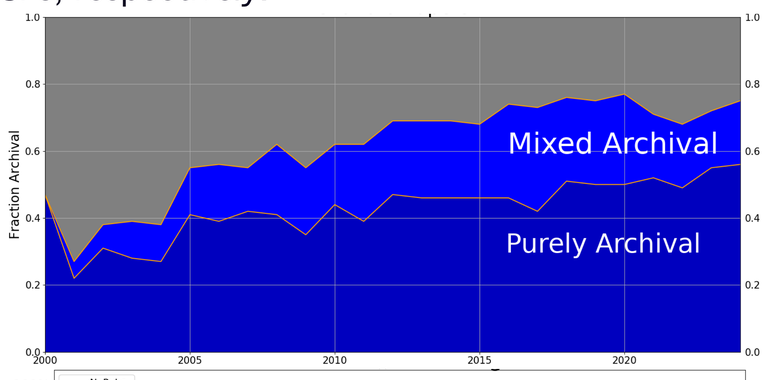
The distribution of the number of archival observations linked to CSPs (inset) is lopsided: ~90% of all CSPs use 50 or fewer observations, ~7% utilize between 50 and 250 observations, and only ~3% of papers are linked to a larger number of observations. However, the total exposure time in CSPs has steadily increased over time, driven by the growing availability of archival data and a focus on "data intensive" research topics like:

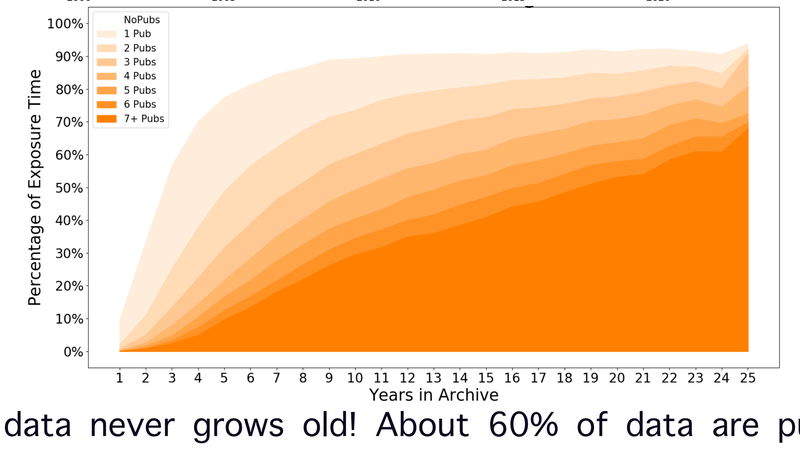
- Time-dependent behavior of the same source observed multiple times
- Comparative analysis of different sources of the same class
- Characterization of entire populations of astrophysical objects that can only be addressed via a statistical approach



Archival science

The well-curated, accessible Chandra archive has broadened the scientific community beyond PIs and their teams. Analyzing CSP authors reveals three categories: PI-led, mixed archival, and purely archival papers. Over time, fully archival and mixed archival CSPs have grown to ~50% and 75% of annual CSPs, respectively.





Chandra data never grows old! About 60% of data are published by 4 papers after 10 years since taken, and nearly half of all archival observations are published over 7 times after 20 years.