astrophysics data system

Content & Curation

The ADS Curation Team

ADS Users Group Meeting, 20-21 Nov. 2025

ADS/SciX Content & Curation Team

Carolyn Grant	Daniel Chivvis	Edwin Henneken	Jeffrey Pomerantz	Jenny Koch	Matt Templeton	Tom Allen	Harry Blom
Data Ingestion Lead	Digital Technologies Development Librarian	Content, Curation & Collaborations Lead	Digital Technologies Development Librarian	Digital Technologies Development Librarian	Data Integration Manager	Liaison with Dev Team	Senior Publishing Liaison (till 9/2025)

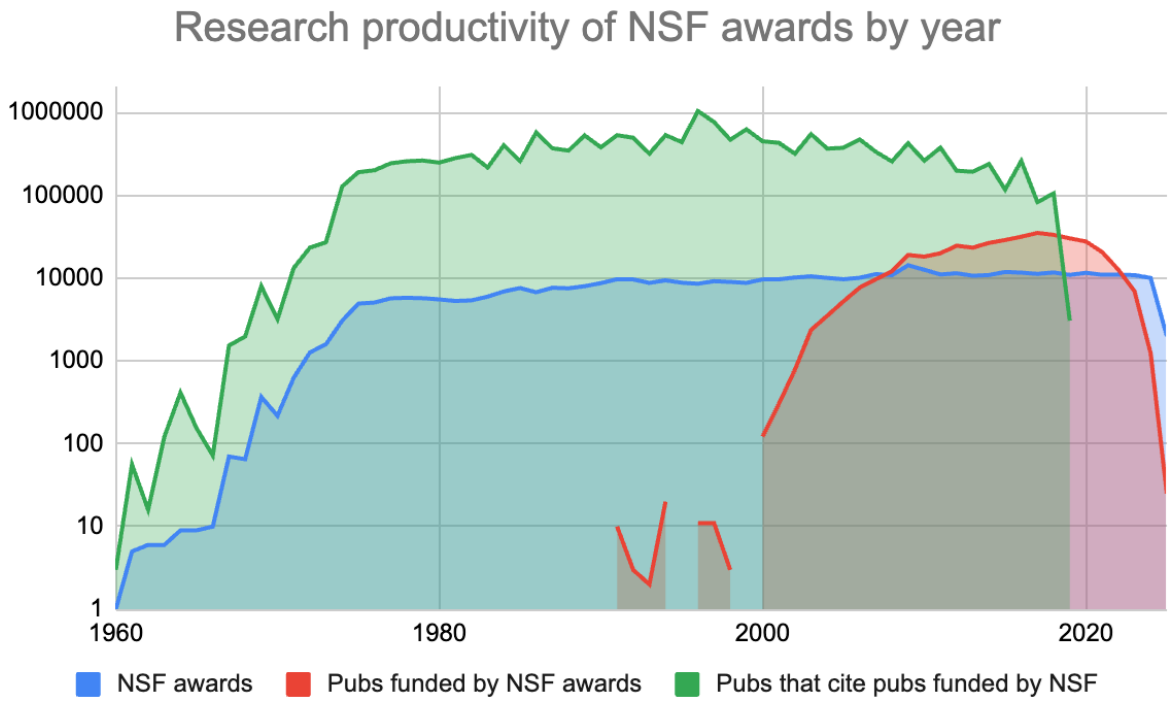
TOPICS

- **Some highlights from the last year**
  - Ingest - Carolyn Grant
  - Awards & proposals - Jeffrey Pomerantz
  - Data, Software & Grant Mentions - Edwin Henneken
  - Journal Information Inventory - Daniel Chivvis
  - Completeness - Matt Templeton
  - Bibliographies - Jenny Koch
  - User Support & Feedback - Edwin Henneken
- **Scenario 1: Plans if ADS & SciX funded**
- **Scenario 2: Plans if only ADS funded**

Highlights: Ingest

- Astronomy content growth steady at ~12% for last several years
  - > 4% increase in journal articles
  - Software, data, proposals, grey literature, etc.
- Earth Science content growth continuing at ~100K per week
  - Grey literature (40K AMS confs; 65K Goldschmidt confs)
- Significant metadata improvement due to new publisher agreements
  - Wiley, Taylor & Francis, MDPI, Frontiers
  - 15 additional smaller publishers
  - IEEE (bigger impact on AST once categorized)
- Matched and added green access links to ~1.2 million PubMed Central (PMC) identifiers
  - esources: (PMC_HTML or PMC_PDF)
- Finalizing ingest by highly cited DOIs (~ 500K)
- Parsing Crossref content with python - only a few perl loading scripts left in production

- Added records for all NSF awards, approx. 470k (bibstem:nsf)
- In process of adding records for all NASA awards and submitted proposals, approx. 7k and growing (do
- Next steps: Add awards fro other funding agencies: DOD, DOE, EPA, USDA, etc
- Records link to publications funded by awards
  - Shows long-term research productivity o government funding
- Part of a larger effort to ingest grey literature to increase SciX completenes



Full Text Sources
Publisher
Data Products
Related Materials
Abstract
Citations
References
Credits 20
Mentions 1
Co-Reads
Similar Papers
Volume Content
Graphics
Metrics
Export Citation

Dormant black holes and neutron stars in stellar binaries

El-Badry, Kareem J *show details*

Proposal

Black holes (BHs) are perhaps the most bizarre prediction of Einstein's theory of general relativity: regions of spacetime where gravity is so strong that nothing, not even light, can escape. Models predict that more than 100 million BHs exist in the Milky Way, yet only about 20 have been discovered to date. This proposal seeks to increase that number by using telescopes on the ground, data from the Gaia spacecraft, and novel data analysis methods to detect BHs via their gravitation effects on stars that orbit them. A research group at the California Institute of Technology will carry out the investigation using a variety of observational and theoretical tools developed by that group. Along the way, the researchers also expect to discover an unprecedented sample of neutron stars -- close cousins of BHs with lower masses -- orbited by stars like the Sun. The ultimate goal is a census of our Galaxy's BH and neutron star population, which will ultimately improve our understanding of the physics of stellar evolution, supernovae, and compact object formation. A majority of the research will be carried out by a PhD student and will form the central pillar of that student's thesis. The group will also develop open-source material for undergraduate labs centered around orbital dynamics. The PI will continue to give public talks and work with the Future Ignited program at Caltech to support students who are members of under-represented minorities in STEM (for example, Black, African-American, Native American and Pacific Islanders).
The team will use a range of ground- and space-based facilities to discover and characterize the first statistical sample of dormant black holes (BHs) and neutron stars (NSs) in binaries with widely separated stellar companions. Candidates will be selected from the recent 3rd data release of the Gaia mission, which provides orbital solutions for an unprecedented 350,000 astrometric and spectroscopic binaries. The dataset represents a factor of 50 increase in sample size over all previous work and thus provides novel opportunities to search for rare objects. The group will use ground-based telescopes to follow-up the best candidates spectroscopically. The main goals of the program are to (1) obtain spectroscopy and multi-epoch radial velocities for all ~100 binaries in the Gaia sample with high probability of having BH or NS components; (2) measure the mass distribution of BHs and NSs from 1 to 15 solar masses, detecting or ruling out the mass gap; (3) constrain the frequency and magnitude of natal kicks for BHs and NSs from binary eccentricities and separations; and (4) detect or rule out the presence of unseen tertiaries within binaries containing BHs. This project advances the goals of the NSF Windows on the Universe Big Idea.
This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

Publication	NSF Award Number 2307232. Directorate for Mathematical and Physical Sciences, Division Of Astronomical Sciences. 2023.
Publication Date	August 2023
Bibcode	2023nsf....2307232E
Collection	astronomy
Comment(s)	Award type: Standard Grant; Award amount: $475,345; Project start date: 2023-08-15; Project end date: 2026-07-31; Directorate: Directorate for Mathematical and Physical Sciences; Division: Division Of Astronomical Sciences; Program: WoU-Windows on the Universe: T, STELLAR ASTRONOMY & ASTROPHYSC

Make Corrections

Full Text Sources
Publisher
Data Products
Related Materials

- Abstract
- Citations
- References
- Credits 20
- Mentions 1
- Co-Reads
- Similar Papers
- Volume Content
- Graphics
- Metrics
- Export Citation

Papers that credited
Dormant black holes and neutron stars in stellar binaries

View as search results

1 **Gaia's binary star renaissance**
El-Badry, Kareem; *show details*
2024/06 · New Astronomy Reviews · cited: 52

2 **A population of neutron star candidates in wide orbits from Gaia astrometry**
El-Badry, Kareem; Rix, Hans-Walter; Latham, David W.; Shahaf, Sahar; Mazeh, Tsevi; Bieryla, Allyson; Buchhave, Lars A.; Andrae, René; Yamaguchi, Natsuko; Isaacson, Howard; *and 3 more*
2024/07 · The Open Journal of Astrophysics · cited: 43

3 **A generative model for Gaia astrometric orbit catalogs: selection functions for binary stars, giant planets, and compact object companions**
El-Badry, Kareem; Lam, Casey; Holl, Berry; Halbwachs, Jean-Louis; Rix, Hans-Walter; Mazeh, Tsevi; Shahaf, Sahar; *show details*
2024/11 · The Open Journal of Astrophysics · cited: 20

4 **The black hole low-mass X-ray binary V404 Cygni is part of a wide triple**
Burdge, Kevin B.; El-Badry, Kareem; Kara, Erin; Canizares, Claude; Chakrabarty, Deepto; Frebel, Anna; Millholland, Sarah C.; Rappaport, Saul; Simcoe, Rob; Vanderburg, Andrew; *show details*
2024/11 · Nature · cited: 26

5 **On the formation of a 33 solar-mass black hole in a low-metallicity binary**
El-Badry, Kareem; *show details*
2024/05 · The Open Journal of Astrophysics · cited: 24

6 **A 1.9 solar-mass neutron star candidate in a 2-year orbit**
El-Badry, Kareem; Simon, Joshua D.; Reggiani, Henrique; Rix, Hans-Walter; Latham, David W.; Bieryla, Allyson; Buchhave, Lars A.; Shahaf, Sahar; Mazeh, Tsevi; Chakrabarti, Sukanya; *and 3 more*
2024/04 · The Open Journal of Astrophysics · cited: 32

7 **Triple Evolution Pathways to Black Hole Low-mass X-Ray Binaries: Insights from V404 Cygni**
Shariat, Cheyanne; Naoz, Smadar; El-Badry, Kareem; Rocha, Kyle Akira; Kalogera, Vicky; Stephan, Alexander P.; Burdge, Kevin B.; Angelo, Isabel; *show details*
2025/04 · The Astrophysical Journal · cited: 13

Full Text Sources
Elsevier
Preprint
Data Products
Related Materials
Abstract
Citations 52
References 230
Credits
Mentions 1
Co-Reads
Similar Papers
Volume Content
Graphics
Metrics
Export Citation

Gaia's binary star renaissance

El-Badry, Kareem show details

Refereed Article

Stellar multiplicity is among the oldest and richest problems in astrophysics. Binary stars are a cornerstone of stellar mass and radius measurements that underpin modern stellar evolutionary models. Binaries are the progenitors of many of the most interesting and exotic astrophysical phenomena, ranging from type Ia supernovae to gamma ray bursts, hypervelocity stars, and most detectable stellar black holes. They are also ubiquitous, accounting for about half of all stars in the Universe. In the era of gravitational waves, wide-field surveys, and open-source stellar models, binaries are coming back stronger than a nineties trend. Much of the progress in the last decade has been enabled by the Gaia mission, which provides high-precision astrometry for more than a billion stars in the Milky Way. The Gaia data probe a wider range of binary separations and mass ratios than most previous surveys, enabling both an improved binary population census and discovery of rare objects. I summarize recent results in the study of binary stars brought about by Gaia, focusing in particular on developments related to wide ($a \gtrsim 100$au) binaries, evidence of binarity from astrometric noise and proper motion anomaly, astrometric and radial velocity orbits from Gaia DR3, and binaries containing non-accreting compact objects. Limitations of the Gaia data, the importance of ground-based follow-up, and anticipated improvements with Gaia DR4 are also discussed.

Publication	New Astronomy Reviews, Volume 98, id.101694
Publication Date	June 2024
DOI	10.1016/j.newar.2024.101694 10.48550/arXiv.2403.12146
arXiv	arXiv:2403.12146
Bibcode	2024NewAR..9801694E
Collection	astronomy
Keywords	Binaries: visual, Binaries: spectroscopic, Binaries: astrometric, Stars: black holes, White dwarfs, Astrophysics - Solar and Stellar Astrophysics, Astrophysics - Astrophysics of Galaxies, Astrophysics - High Energy Astrophysical Phenomena
UAT Keywords (generated) BETA	spectroscopic binary stars, stellar evolution
E-Print Comment(s)	Accepted to New Astronomy Reviews (Special Issue "Gaia, the first crop of discoveries"). 28 pages, 13 figures

Full Text Sources

Elsevier

Preprint

Data Products

Related Materials

Abstract

Citations 52

References 230

Credits

Mentions 1

Co-Reads

Similar Papers

Volume Content

Graphics

Metrics

Export Citation

Papers mentioned by

Gaia's binary star renaissance

View as search results

1

Dormant black holes and neutron stars in stellar binaries

El-Badry, Kareem J; *show details*

2023/08 · NSF Award · credited: 20

Highlights: Data, Software & Grant Mentions

A work, indexed in ADS/SciX, is acknowledged/referred to in another work, also indexed in ADS/SciX (outside of the bibliography)

We support the following mentions at the moment

Software

- From data supplied by the ASCL ("Described by", "Used in") - continuously updated
- From mining DOIs from Data Availability Statement sections - continuously updated

Datasets

- From mining DOIs from Data Availability Statement sections - continuously updated

Grants

- From Crossref and DataCite metadata mining - needs structural mining process

Currently we process Data Availability Statement sections for 375 journals.

The query **`has:mention`** returns those records that mention ADS/SciX records of the types above

The relationship `a--[mention]-->b` is automatically inverted to give `b--[credit]-->a`

(note that the queries **`has:mention`** and **`has:credit`** will return different numbers of records, about 320k versus 74k resp.)

Credits

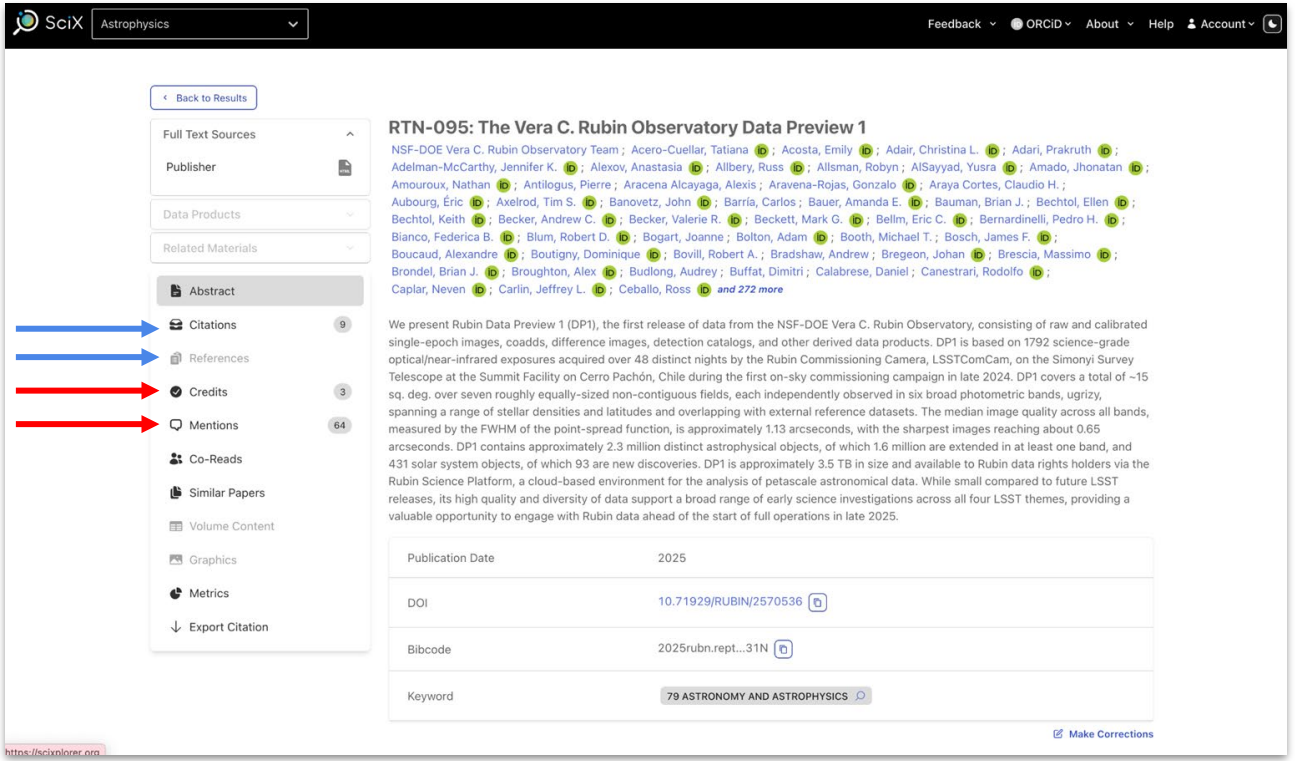
	Astronomy	Earth Science
software	3700 (9.5%)	12 (1.6%)
datasets	235 (0.97%)	946 (2.9%)
grants	6,426 (10.7%)	60,266 (15.0%)

Number of records of a particular doctype in a particular collection with credits.
Percentage of all records of that doctype in the collection

Highlights: Data, Software & Grant Mentions 2/2

Note: mentions and credits are only visible in the Science Explorer

Citations/references and credits/mentions are complementary



Highlights: Journal Information Inventory

- 8,235 journals with at least 1 ISSN
- Metadata enrichment in progress
  - Journal history
    - Title, Publisher, Volume (by year)
  - Language metadata + translations
  - Subject headings
  - Bibliometrics, rankings
- Structured in JSON for Journals DB import


```
{
  "journal": {
    "canonical_name": "Astronomy and Astrophysics",
    "journal_id": 1757,
    "classic_bibstem": "A&A",
    "previous_titles": [
      {
        "canonical_name": "Annales d'Astrophysique",
        "years": "1968-1975",
        "classic_bibstem": "AnAp",
        "issn": {
          "print": "0365-0499"
    },
      }
    ],
    "issn": {
      "print": "0004-6361",
      "electronic": "1432-0746"
    },
    "publisher_history": [
      {
        "publisher": "Gauthier-Villars",
        "years": "1968-1975"
      },
      {
        "publisher": "EDP Sciences",
        "years": "1975-present"
      }
    ],
    "languages": [
      "English",
      "French"
    ],
    "impact_factor": "5.0",
    "website": "https://www.aanda.org/"
  }
}
```

Highlights: Completeness Statistics

We cover **many** more journals now than before: *how complete is our metadata coverage across all journals in all disciplines*?

- **Weekly** Crossref harvest for comparison data – what records we *should* have
- Process new Crossref records and check previously unmatched Crossref records to determine whether we now have them in SciX
- Calculate journal/volume completeness statistics
- Results available via API ("journals" endpoint)

Core astronomy/physics: ~ 100%; rapidly improving for other collections


```
},
{
  "volume": "536",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "536L",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "537",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "537L",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "538",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "538L",
  "volume_completeness_fraction": 1.0
},
{
  "volume": "539",
  "volume_completeness_fraction": 1.0
},
{
```

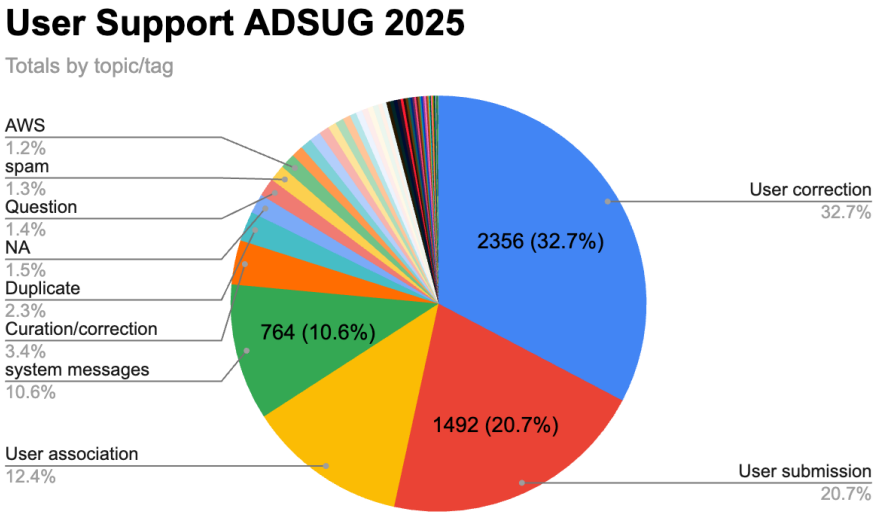
Highlights: Bibliographies

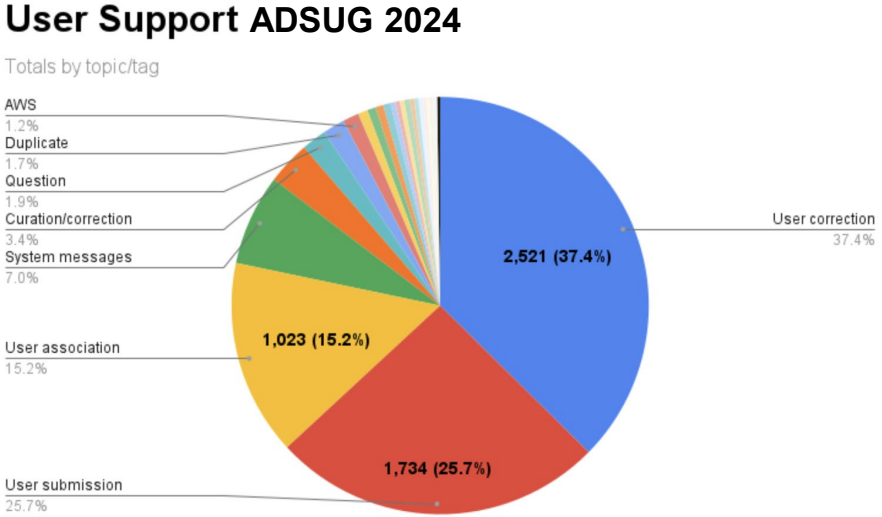
Coverage of bibliographies help ensure relevant content is represented in our collections, and in the process we build relationships with communities of researchers and scientists

Bibliography	Total citations	Matched	Curation	Status	Notes
NASA (GES DISC) Goddard Earth Sciences Data Information and Services Center	14,389	11,940	2,300	Complete	
NASA Ames Astrophysics & Astrochemistry Lab	788	765	23	Complete	
NASA Ames Space Sciences & Astrobiology	2,622		100	Complete	
NASA Astrobiology Division	4,778	4,720	58	Complete	
NASA Goddard (SED) Science & Exploration Directorate: 610 Earth Science	15,675	12,932	2,742	In progress	
NASA Proposals	6,766	1,387	5,379	Complete	Also added 1837 award profiles linked from USASpending
NASA PubSpace	37,410	31,783	3,780	Complete	
NASA SEDAC	5,927	3,397	2,530	Complete	
NASA Space Life Sciences Library	101,994	37,115	64,870	In progress	
NASA STI/NTRS (NASA Report Series)	106,592	2,048	104,545	Complete	+770 books curated
National Advisory Committee for Aeronautics (NACA)	13,787	13	13,774	In progress	
National Academies of Sciences publications (reports)	15,261		15,261	Complete	
Oak Ridge National Lab (ORNL) DAAC	3,900	3,170	700	Complete	
USGS Publications Warehouse	183,000	51,000	130,000	Complete	

Highlights: User Support & Feedback

There were **233k** references submitted in the last year, compared to **44k** the year before





Ingest, Curation & augmentation

- Data products & software from ES, HP, PS, BPS
  - Indexing, citations & mentions
- Grey literature for the above disciplines (includes conferences, reports, thesis)
- Bibliographies for NASA divisions, projects and missions
- Ingest of NASA Technical Reports Server (NTRS) content
- Ingest of awards (from NASA and other agencies)
- Tagging publications with
  - Grant IDs
  - Geolocation information
  - License info
- Curation of citation coverage for all core journals in the disciplines covered by SciX
- User and community support
- UAT expansion
- Support for OA journals
- Improved coverage for "Outer Rim" of Astronomy (Astrobiology, Astrochemistry, Astrogeology, Laboratory Astrophysics, Celestial Navigation, …)

Scenario 1 Plans: Funding for SciX & ADS 2/2

Operations & tooling

- Removing dependency on legacy components
  - Reference extraction for arXiv using new infrastructure
  - Reference matching using new infrastructure
  - Classic indexing deprecated as "master"
  - The bibcode will no longer be the canonical identifier
- Curation of institution database (for normalization of affiliations)
- Mine information from full text documents (grants, license, mentions)
- Docmatching pipeline extended to finding errata, retractions and series
- Journals database as the canonical source for all journal related information

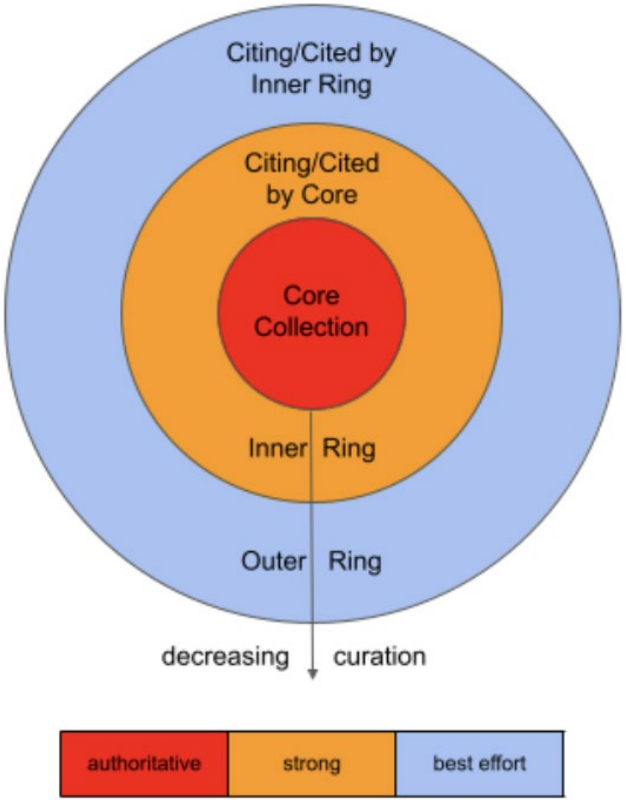
Scenario 2 Plans: Funding for just ADS

Overall impact

- The original goal of NASA SciX being an **authoritative, curated, open access metadata system** covering **all five SMD disciplines** is **no longer viable** with only ADS funding.
- **Earth Science and NASA BPS** will be **deprioritized**, moving outwards from the Core Collection in the ADS Curation Model.
- **All activities will be negatively impacted** due to **decreased resources**.

Impact on content & curation

- The **mandate changes** with only **ADS funding**, necessitating **internal discussions** to redefine operational **priorities and scope**.
- **Ingest and curation focus will revert** to the **pre-expansion time**.
- **The full scope of curation activities will be limited to Astrophysics only** under this funding model.
- **Discipline-agnostic efforts** (especially for **operations and tooling**) will continue at a much reduced pace. User and service support **will suffer** from **reduced staffing levels**.



The ADS that remains in this scenario **will not be the same** as the pre-expansion ADS.