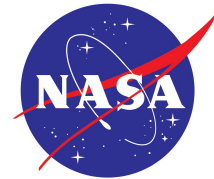


Data Curation, User Support and Outreach

*Edwin Henneken, Carolyn Grant, Matthew Templeton, Donna Thompson,
Jenny Koch and the ADS Team*

ADS Users Group Meeting, 9-10 Nov. 2022



Data Curation, User Support and Outreach

- Content & Curation Roadmap (EH)
- ADS Curation Model: Journal Selection (EH)
- Coverage reporting (EH)
- Content Summary 2022 (CG)
- Journal Database update (MT)
- Twitter Campaigns (JK)
- User Support (DT)

Content & Curation Roadmap

Content & Curation Roadmap - Topic areas

<p style="text-align: center;">Content</p> <ul style="list-style-type: none">● Mitigation of missing content● Discovery, retrieval & ingest of new content● Curation & (meta)data quality● Content analysis	<p style="text-align: center;">Tooling</p> <ul style="list-style-type: none">● Journals database● Remove Classic dependencies● Curation tools● Reporting
<p style="text-align: center;">Documentation</p> <ul style="list-style-type: none">● Write & discuss RFCs● Help pages maintenance● Document workflows● Authoring blogs & papers	<p style="text-align: center;">Outreach & Support</p> <ul style="list-style-type: none">● Collaborations● Social media● User/Collaborator/Supplier support● Exhibiting

ADS Curation Model: Journal Selection

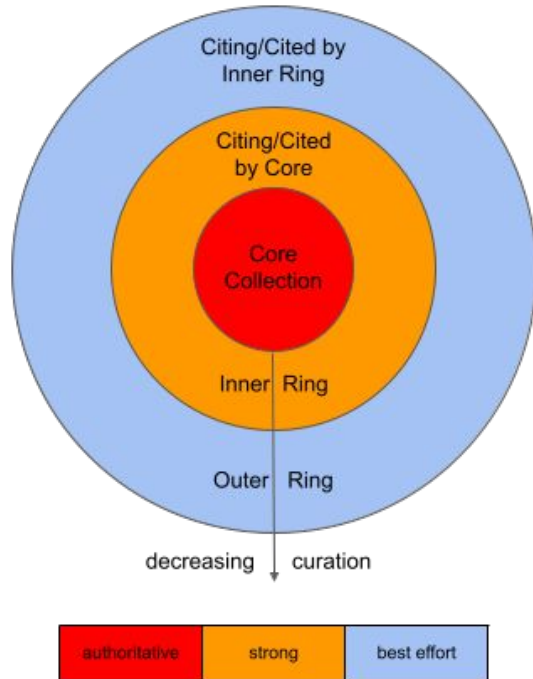


Figure 5. ADS's tiered curation model. The core collection represents disciplines where its curation is strongest and its coverage is authoritative. The surrounding tiers are connected to the core via the citation citation network.

Core: main journals for discipline

- Complete, authoritative coverage of the literature
- High-level data products and software indexed
- Links to datasets and archives

Inner Ring: closely related disciplines

- Most refereed journals covered
- Some conf proceedings, some gray literature
- Citation and fulltext coverage incomplete

Outer Ring: connected to inner content

- Incomplete coverage of the literature because outside of disciplinary scope
- Includes content from multidisciplinary journals (e.g. Nature) or repositories (arXiv)
- No curation applied other than basic error corrections

ADS Curation Model: Journal Selection

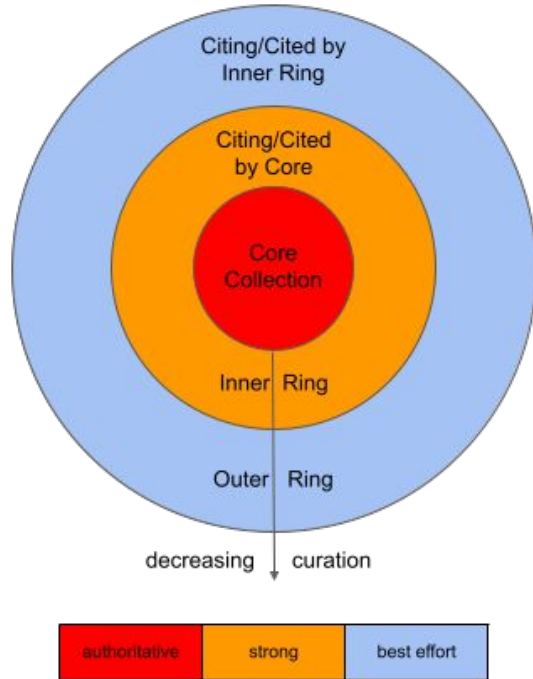
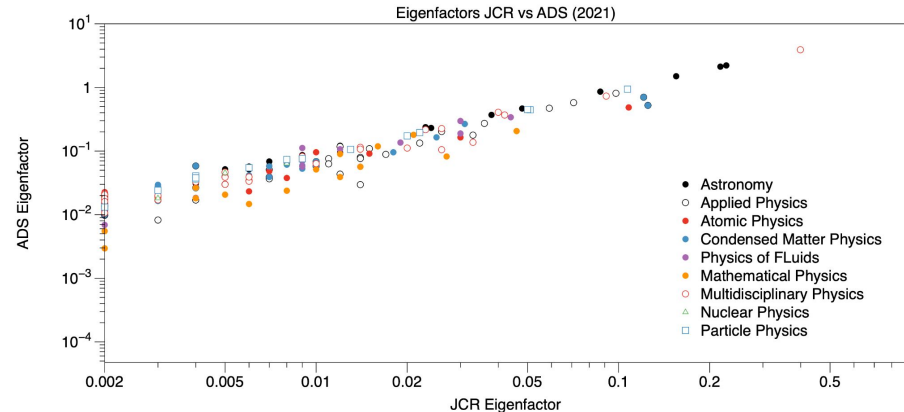


Figure 5. ADS's tiered curation model. The core collection represents disciplines where its curation is strongest and its coverage is authoritative. The surrounding tiers are connected to the core via the citation citation network.

For any given discipline, in a given year, the Core Collection needs to consist of the top N most influential journals, such that this collection represents at least 90% of all citations represented in the set of journals considered. In other words: Core Collection is determined by a ranking and a cut-off threshold.

In progress: verify journal impact ranking from ADS data

First results indicate that ranking is the same as generated from WoS data from their Journal Citations Report (JCR)



Coverage reporting & metrics

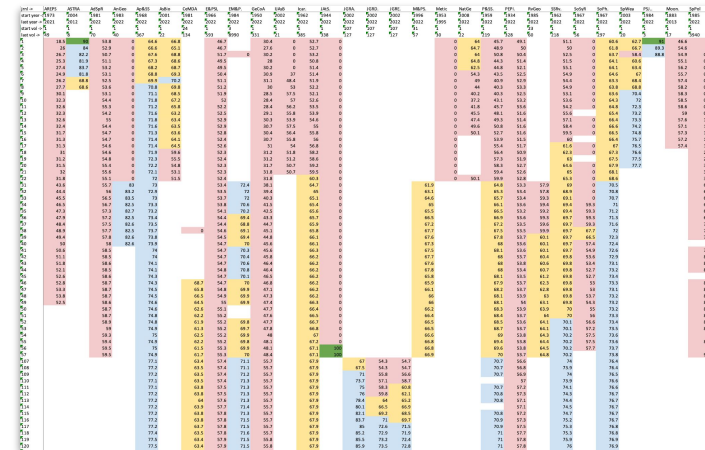
Coverage levels

Journal →

Volume ↓



Full text



Matched references

The ADS curation team has created a workflow that reports on coverage levels

Snapshots of coverage levels can be used to report progress

2022 Content update

- 9.5% growth in abstracts and citations
- Weekly updates
 - 2,100 abstracts/week (AST)
 - 10,000 abstracts/week (PHY)
 - 250,000 citations/week
- Planetary additions
 - PDS, LPI Bulletin and precursors, Oxford Research Encyclopedia
 - More monograph content, HOLLIS records, Princeton University Press, Elsevier book series
 - Updates to existing records (references and fulltext)
- Continued maintenance to affiliations (AST ~96%; PHY ~90%)

Journals Database

Back-office Journal metadata storage and management, including: journal identifiers (bibstems); ISSNs and Crossref collection IDs; publisher information; title histories; and summary data for reference sources and fulltext availability.

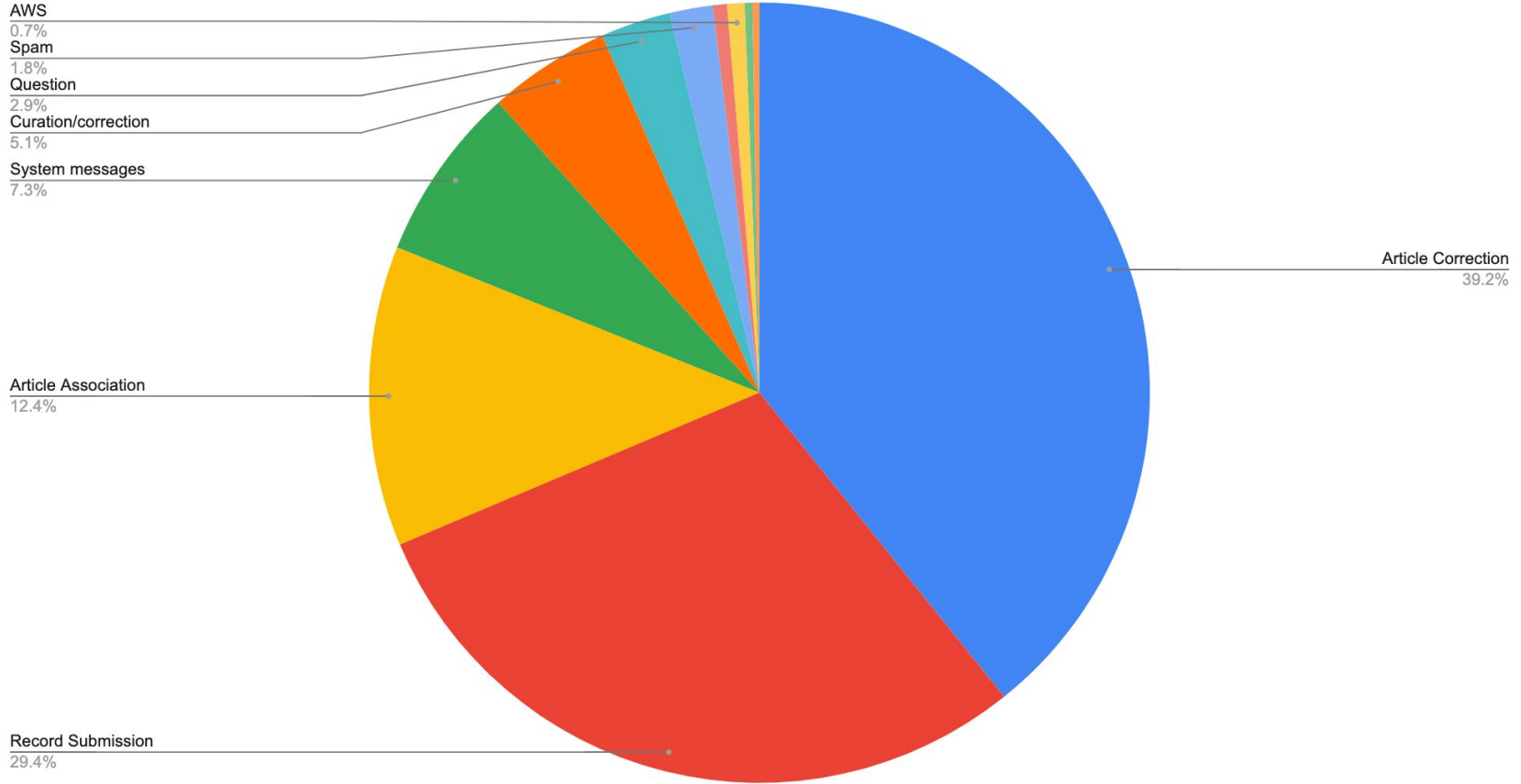
- Curation interface: bibstems (since January 2022)
- Additional backoffice curation: ISSNs, Crossref Collection IDs, Publisher info, collection info (topic-core-ring data)
- Public-facing API and documentation now online
- Shifting focus from software dev to curation
- Work in progress: Completeness data

Top HIVER Topics

January-September 2022

Article Correction	1473
Record Submission	1105
Article Association	467
System messages	273
Curation/correction	190
Question	109
Spam	67
Position application	22
AWS	27
myADS	12
Data issue	11

HIVER Activity January - September 2022



Twitter Outreach Campaigns

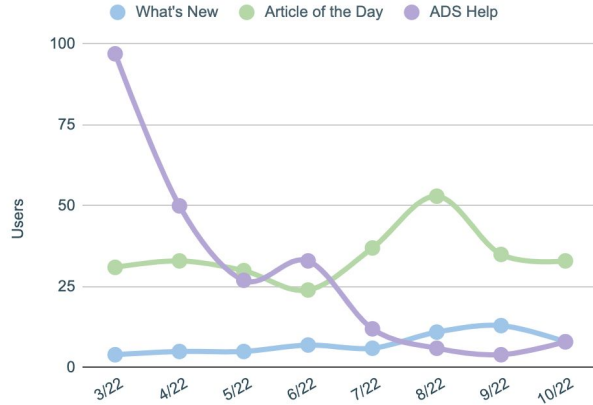


- [ADS Help](#) (#ADSHelp)
 - Ran series of ~21 tweets from March to June 2022
 - Tips on ADS services & search features (4-5 posted per week)
Examples include: ‘inst’ search, positional search, how to export custom reports and author affils, use of ORCiD, etc.
 - Newsletter & Help Page implementation (coming soon)
- Ask ADS Office Hours ([#AskADS](#))
 - Held monthly “Office Hours” on Twitter in [March](#), [April](#), & [May](#) 2022
 - Tested various times and days to gauge level of reach and engagement
 - Attendance was poor

Twitter Outreach Campaigns



Campaign Posts - Users



Followers



Account Interactions

