ADS 5-Year Plan

Alberto Accomazzi and the ADS Team
@aaccomazzi

ADS Users Group Meeting, 19-20 Nov. 2020

NASA Archives Programmatic Review

- Every 5 years NASA Astrophysics conducts a programmatic review of all its archives (HEASARC, MAST, IRSA, NED, NEA, NAVO, ADS)
- Goals: refine strategy, prioritize tasks, give direction to archives
- Evaluation: based on support for science activities, identification and curation of data and software, promoting use of NASA Astrophysics data, taking full advantage of state-of-the-art data management techniques
- Panel met on 10-12 March 2020, released report on June 15 2020: https://science.nasa.gov/science-pink/s3fs-public/atoms/files/NASA-AAPR2020_Public_Report-Final.pdf
- Based on ADS proposal and successful outcome of review process, NASA setting up ADS Cooperative Agreement for next 5-year period starting on 1 March 2021

Work Area 1: Data Ingest and Curation

- Backoffice rewrite: update the ADS back-office ingest and curation system to improve frequency of updates, data quality, and operational efficiencies
- Metadata Enrichment: expand the use of text mining techniques to identify, index and expose entities and topics present in the indexed literature
- Data Indexing: implement indexing of high-level data products published by NASA and other major archives, and track their citations

Work Area 2: Use and Access

- User Interface Refactor: refactor the internals of its User Interface to use current technology and frameworks, improving its overall usability, accessibility and maintainability
- Lead API Interoperability Efforts: work with the community to develop a full-featured client library and integrate it in the popular open-source package astropy
- System Reliability and Capacity: improve our system architecture to better protect it against computationally expensive queries and to allow the execution of long running queries (via e.g., asynchronous API calls)

Work Area 3: User Support

- Advanced ADS features: utilize social networks, blogs, newsletters, help pages and other alternatives (e.g., online support forums) to explain and promote lesser known functionalities
- Outreach: continue to engage the community through participation at conferences and other community events
- Terms of Service: define and publish Terms of Service for its content and operations, including coverage, completeness, currency and availability

Work Area 4: System Enhancements

- Improve Discovery: use state-of-the-art NLP techniques to improve our search, recommendation and notification systems
- Usability and Access: improve access of its system and make sure that all of its functionality is usable to its fullest extent by all people
- Redundancy: strive to maintain redundancy in its operations and document its architecture and components to ensure their long-term maintenance and support

AAPR2020 Outcome and Feedback

ADS rated “excellent” by panel, with a strong endorsement of our plan:

> *The software and overall architecture work ADS has done and plans to do is sound.*
>
> *The discussion of new technologies to implement in planning for the future (i.e., concept-based indexing, natural language processing, and cloud services) shows that ADS is properly exploring future trends and tasks that will be required to execute its core mission going forward.*
>
> *Moving to an agile workspace where no single person is key to the success of any ADS component is admirable and should be the standard for all archives, particularly in light of difficulties filling some positions and with turnover in a hot job market.*

AAPR2020 Recommendations (1/2)

Come up with clear plan/policy for indexing complex objects (notebooks, data, documentations, presentations)

On roadmap for 2021

Find a way to hire qualified personnel (said in the context of UI/UX)

Partly addressed earlier

Publish clear policy about tracking user behavior and privacy

On roadmap for this year (also, GDPR audit and updates)

Replace FTP access with more modern upload methods

On roadmap before end of 2020

AAPR2020 Recommendations (2/2)

Provide better way to link (not index) references outside of core disciplines

Will be considered in the context of a rewrite of the backoffice system

Leverage markup of objects, datasets, UAT terms from publishers

Enhancing fulltext extraction on roadmap for 2021

Advertise ADS outside of astronomy ("ADS Ambassadors"?)

Part of ADS Expansion effort addressed in Future Plans presentations

Teach advanced ADS functionality by example

Planned as part of UI redesign, waiting for new hire

Consider additional layer in architecture besides UI and back-end

API/microservice layer already providing this functionality