State of the ADS

Alberto Accomazzi and the ADS Team

ADS Users Group Meeting, 15-16 Nov. 2021

Overview

- Project Funding
- ADS Staffing
- Programmatic Goals for FY21/FY22/FY23
- Review of 2020 ADSUG Recommendations

Project Funding

Updates from last year:

- NASA Cooperative Agreement started on March 1, 2021
  - Includes funding for a full FTE position for Project Scientist
  - Includes funding for ML engineer and UI/UX specialist

- A solicited NASA supplemental proposal for expansion of ADS into Planetary Science and Heliophysics was funded in May 2021
  - Adds 3 FTEs to the ADS staff
  - Includes a request for a study to explore a further expansion into Earth Sciences and Biological and Physical Sciences
  - 5-year expansion proposal due to NASA in early 2022, with potential funding to start in FY23

ADS Staffing

With pandemic changing the rules of the game, ADS management decided to make all new positions eligible for 100% remote work for at least a one year. This extended the pool of candidates significantly and allowed multiple positions to be filled or extended:

- Four new team members work remotely (UI/UX, ML, curator, data support)
- Two new team members are local (back-office development)
- Michael Kurtz (PS Emeritus) stayed on to help with expansion proposals

ADS Staff and Roles

- **Alberto Accomazzi**, PI and Program Manager
- **Michael J. Kurtz**, Project Scientist Emeritus
- **Pavlos Protopapas**, Interim Project Scientist for Data Science (part time)
- **Edwin A. Henneken**, Content, Curation and Collaborations Lead
- **Sergi Blanco-Cuaresma**, Software Development and Cloud Operations Lead
- **Carolyn S. Grant**, Senior Curator
- **Roman Chyla**, System Architect and Senior Developer
- **Donna M. Thompson**, Data Curation Librarian
- **Steve McDonald**, Pipeline Development and Operations (part time, retiring)
- **Tim Hostetler**, User Interface and Front-end Developer
- **Golnaz Shapurian**, Senior Developer, Applications & Services
- **Matthew Templeton**, Data Ingest and Curation
- **Kelly Lockhart**, Back-End Development and Community Outreach
- **Nemanja Martinovic**, System Operations, Cloud Computing and DevOps
- **Jennifer Chen**, User Experience and Front-End Development
- **Felix Grezes**, Machine Learning and Natural Language Processing
- **Jenny Koch**, Digital Technologies Development Librarian
- **Tom Allen**, Astrophysics Data Support (part time)
- **Taylor Jacovich**, Back-office Development and Data Analysis
- **Thomas Fehring**, Back-office Development and Data Analysis

Programmatic Goals FY2021

✓ Metadata Enrichment: build a language model to support identification of entities and concepts found in the literature (UAT)

✓ Usability and Access: progressive enhancements have improved website accessibility and usability (especially from mobile devices)

✓ User Interface refactor: build prototype system to test new javascript framework, evaluating its overall usability, accessibility and maintainability

● Legacy System Updates: rewrite of ADS Article service has been planned, pipeline legacy code updated and ported to modern computing platforms

✓ API Interoperability Efforts: engage the community to document API, develop a full-featured client library, and integrate it in astropy

✓ Outreach: we continued to engage the community through participation at DPS, AGU, AAS, LPSC meetings

✓ Redundancy: increasing redundancy in operations and documenting system architecture to ensure long-term maintenance and support

✓ Security: replacing FTP and HTTP access with secure protocols, close access to servers in LAN

Programmatic Goals FY2022/23

- Update curation workflows, ingest system, and data model
- Implement indexing of high-level data products published by NASA and other major archives
- Utilize social networks, blogs, newsletters, help pages and other alternatives (e.g., online support forums) to explain and promote lesser known functionalities
- Define and update Terms of Service including coverage, completeness, currency and availability
- Improve system architecture to better protect it against computationally expensive queries
- Incorporate Machine Learning efforts in metadata enrichment and curation workflows

From Planetary Science (PS) and Heliophysics (HP) Expansion Augmentation:
- Ingest 100% of available refereed literature in PS and HP, up to 50% of available gray literature
- Index content from ESSOAr preprints; index/link data and software records cited in PS and HP lit.
- Improve citation processing for PS and HP content
- Complete census of literature for Earth Sciences and Biological and Physical Sciences
- Complete evaluation of current infrastructure, documentation, and processes
- Deliver Updated Proposal for 5-year Expansion Plan

ADSUG 2020 Recommendations (1/7)

ADS should continue to prioritize system reliability
- *Consider proactive measures, like IT emergency response drills, that help assess staff readiness for situations affecting ADS services.*
- *In functional areas of the system where expertise is limited to few staff members, consider training options to fill gaps.*

  Ongoing - Addressed in IT Infrastructure Overview

Retire legacy components
- *Continuing to rely upon the classic codebase and extend it presents an operational risk and a source for additional technical debt*
- *We recommend minimizing the time spent running a hybrid system (i.e. old and new systems running in parallel) during the transition.*

  Ongoing - Addressed in Classic back-office rewrite

ADSUG 2020 Recommendations (2/7)

Document and improve API access
- *We encourage ADS to work with the open-source community to broaden the capabilities of clients that make use of the APIs that ADS has to offer*
- *Use services that generate API web documentation from code (e.g., swagger)*

  Done - addressed in API Interoperability efforts

Staffing
- *Time to hire a full-time Project Scientist*
- *Concern that delays in filling open positions may indicate that the process for recruitment is flawed.*

  Ongoing - PS position evolving, addressed in Expansion Proposal section;
  Recruitment more successful with remote work option

ADSUG 2020 Recommendations (3/7)

Educate User Community
- *Consider UX methods of highlighting new and sophisticated features in the UI itself, such as persistent tooltips on new elements or options that need to be dismissed [...]*
    - **Under consideration for new UI refresh**
- *Consider advertising new features and past promotions through society newsletters that link back to the ADS help pages, blog, and newsletter subscription page [...]*
    - **Ongoing as new features are implemented**
- *Consider ways to facilitate the community to help itself troubleshoot and share information.*
    - **Ongoing via existing social media platforms**
- *Consider implementing a program similar to the AAS Ambassadors/Agent program.*
    - **Under consideration (collaboration with librarians already an ongoing activity)**
- *ADS should continue to promote its indexing capabilities and functionality for end users through video tutorials, social media presence (like Astronomers facebook group), and continue to offer major workshops, sessions, etc. at major meetings*
    - **Ongoing - outreach activities described in User Support and Outreach**

ADSUG 2020 Recommendations (4/7)

Accessibility and Usability
- *ADS should ultimately meet the WCAG 2.0 standard. One possible approach is that ADS should set for itself accessibility milestones to be achieved in each coming year, to ensure steady ongoing progress toward a longer-term goal.*
- *We note the distinct usability and accessibility issues that arise on different types of devices (desktop, tablet, smartphone, screenreader, etc.)*
    - **Ongoing - addressed in the Website Updates presentation**

Enhancements and Improvements
- *Recommendations are difficult and should not be a priority, given the existence of myADS and other tools (e.g. arXivsorter, arXangel).*
- *Relevant and potentially more useful efforts that fall within this general area include integration of UAT, ability to search embedded publication tags for telescopes/facilities/software, etc.*
    - **Done - work on recommendations now low priority, work on NLP and document understanding now high priority**

ADSUG 2020 Recommendations (5/7)

Expansion

- *Expansion offers tremendous opportunities and risks for ADS - its current success depends on a deep understanding of the community it serves and attention to that community's changing needs*
- *Before doing work, understand deeply the cultures of the fields which they are expanding into, since each is different.*
- *Therefore we recommend moving field by field - in planetary and, to some extent heliophysics*
- *Realism is needed to understand what resources and time are needed to provide a useful service to each division's scientific community. This may mean providing services at different levels to different communities in the short and medium term.*
- *Any substantial increase in the number of staff will also need to account for an increase in management capacity.*

      Done - Expansion in Planetary and Heliophysics started,
      all other recommendations being incorporated in upcoming proposal

ADSUG 2020 Recommendations (6/7)

Wish List for our consideration:

- *A push to index and link to more theses, including those from outside the USA.*

  Ongoing - Addressed in the Data Curation, User Support and Outreach section

- *Linking to data and research artefacts deserves more focus. In particular, the ADSUG agrees that workflows and pipelines to index data sets, notebooks, software are essential in order to serve the community well in the next one-two years. Conversations with major entities such as CrossRef, DataCite and Zenodo will be essential to doing this well.*

  Ongoing - Conversations with STScI, AAS, AGU, DataCite

- *Consideration of indexing other material authored by astronomers, such as articles in The Conversation and Astrobites, and possibly AAS Nova as well.*

  Ongoing - AAS Nova indexed, contemplating linking vs. indexing other sources

- *Augmentation of frequency of reindexing and updates to indexed records, so that updates do not lag up to 1-2 weeks.*

  Done - Addressed in the Data Curation, User Support and Outreach section

ADSUG 2020 Recommendations (7/7)

Wish List for our consideration:

- *Offering example notebooks that demonstrate how to conduct custom or complex calculations on bespoke citation metrics.*

  Done-ish - Addressed in the API Interoperability Efforts presentation

- *Indexing of conference materials (e.g., talk slides, conference programs, e-posters, speaker videos)*

  On hold - waiting to see these artifacts properly published and preserved

- *Ongoing improvement to feedback form interface (e.g., copying information over from the referring page, as currently happens for the "Missing/Incorrect record form"), or even facilitating common requests (e.g. article/arXiv linking) with clicks directly from the search results.*

  Ongoing - current feedback workflow considered adequate by curators, with updates to address new use cases (e.g. third party ORCID claims)

- *Links to referee reports (where available) as an associated work or otherwise linked to indexed publication entry*

  On hold - low community interest and adoption so far