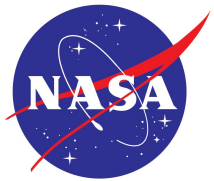
State of the NASA Science Explorer (SciX)

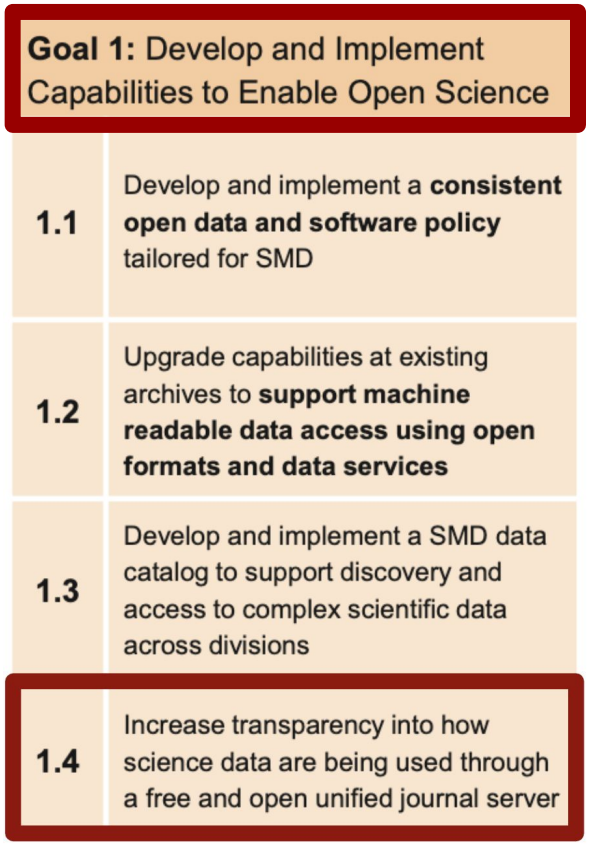
Alberto Accomazzi and the ADS Team

ADS Users Group Meeting, 16 Nov. 2023



SciX: ADS for all of NASA Science

NASA's Science Mission Directorate in 2019 calls for the creation of interdisciplinary literature portal spanning across SMD

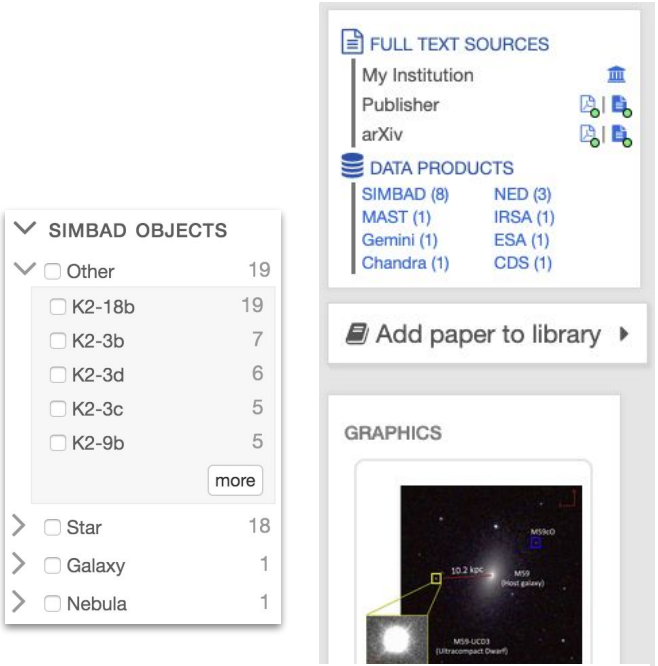


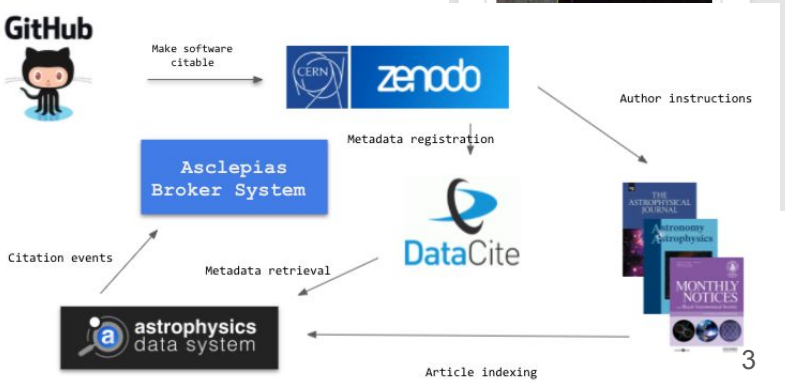
https://smd-cms.nasa.gov/wp-content/uploads/2023/05/SDMWGStrategy_Final.pdf

SciX: ADS for all of NASA Science

NASA's Science Mission Directorate in 2019 calls for the creation of interdisciplinary literature portal spanning across SMD

ADS has been selected from its support of open science goals: facilitating discovery and dissemination of OA publications, data, and software by aggregating and linking them





SciX: ADS for all of NASA Science

NASA's Science Mission Directorate in 2019 calls for the creation of interdisciplinary literature portal spanning across SMD

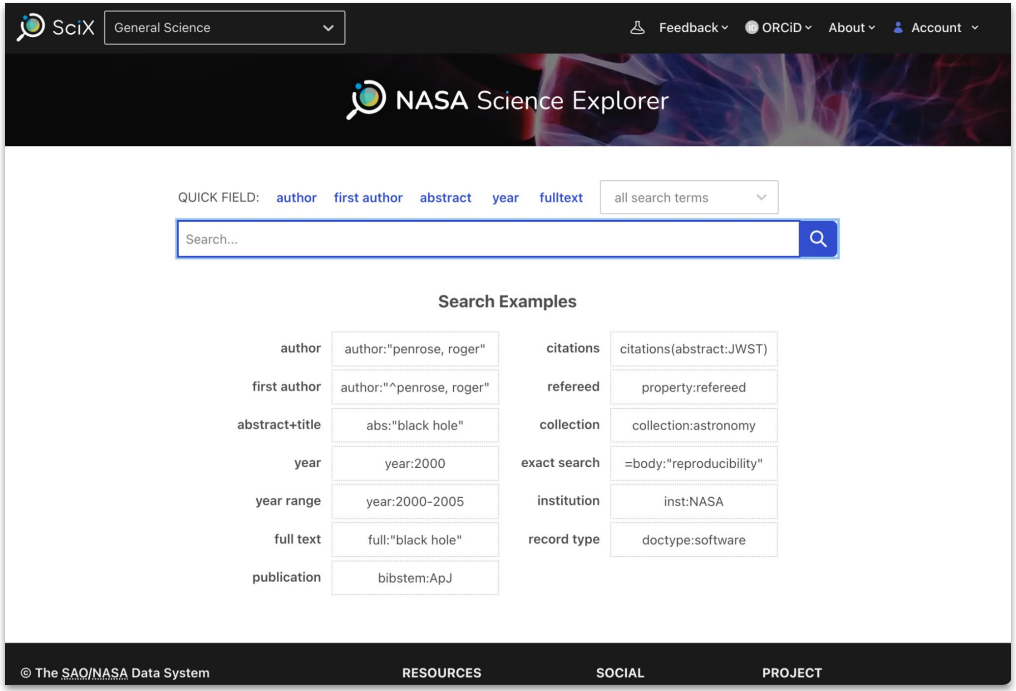
ADS has been selected from its support of open science goals: facilitating discovery and dissemination of OA publications, data, and software by aggregating and linking them

ADS will develop and operate a new digital library portal covering all SMD disciplines

The NASA Science Explorer (SciX)

NASA SciX will be a literature-based, open digital information system covering and unifying the fields of Astrophysics, Planetary Science, Heliophysics, and Earth Science. It will also cover NASA funded research in Biological and Physical Sciences.

NASA SciX will combine a scalable, discipline-agnostic core with a set of discipline specific knowledge centers which will curate and enrich its content using the deep subject matter expertise which has been crucial to the success of the ADS.



SciX: a Digital Library for NASA Science

1. All discipline-specific research content is aggregated, connected, and indexed for each of the SMD divisions;
2. Relevant taxonomies are used to capture the knowledge and semantics of the subject disciplines;
3. Curation and machine learning-based text mining and enrichment are combined in a platform that is designed to scale without sacrificing accuracy and flexibility;
4. Digital collections are enriched with links to other research objects such as datasets, software, notebooks, and funding information;
5. Discipline-specific capabilities and analytic services are exposed to the relevant research communities;
6. Discoverability and access to NASA-funded research artifacts and derived data products are available to all from a public search portal;
7. New and existing initiatives are developed and supported in collaboration with NASA and other research organizations.

What the Expansion Entails

- Content Selection: complete coverage of PS, HP, add ES and NASA BPS
  - We estimate to grow by a factor of 2.5 in records indexed
  - We will ingest records for data products
- Partnerships and Collection Development: develop collaborations & outreach
  - Interface with NASA divisions, new publishers, and archives
  - Develop an outreach strategy in new disciplines
- Data Ingestion and Curation: improve efficiency
  - Improve ingest pipelines, text mining activities, ML-based metadata enrichment
  - Select collections of data products to index/link
  - Provide releases of Open Access content, training data, and models
- System Development and support: improve discovery
  - Improve Semantic Search capabilities (synonyms, taxonomies)
  - Improve author name disambiguation and ORCID integration
  - Develop UI enhancements tailored to specific disciplines

Our Solemn Pledge: we won’t break ADS!

- The ADS website is not disappearing
- It will still be available from the same URL you have bookmarked
- Your account and settings will transfer over (really, be shared)
- Your private/public libraries will transfer over (again, shared)
- The existing functionality will remain available
- You won’t need to change/update links to ADS/SciX (even links to ADS classic will continue to resolve)
- You will still be able to use the classic form
- We’ll take our time to make sure astronomers are comfortable with the SciX platform

HOWEVER:

- Future development will focus on the SciX platform
- New features will be made available on SciX
- Discipline-specific customizations will allow us to provide compatibility with current ADS

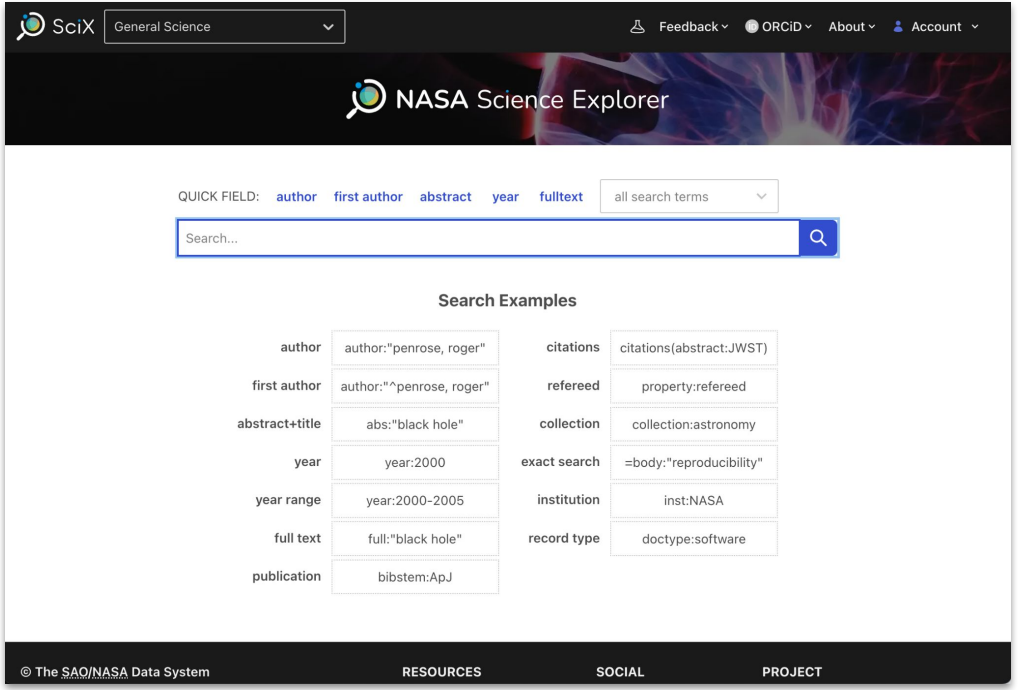
Budget

	Period of Performance	SciX budget	SciX FTEs	ADS budget		New Hires
Year 1(*)	6/1/22-2/28/23	$1,600K	8.5	$3,312K		4 new FTEs
Year 2	3/1/23-2/28/24	$2,500K	10.5	$4,487K		2 new FTEs
Year 3	3/1/24-2/28/25	$3,500K	14.0	$4,400K	(+)	3 new FTEs
Year 4 (#)	3/1/25-2/28/26	$4,000K	16.0	$4,500K	(+)	2 new FTEs

(*) Pro-rated ADS budget for the given PoP (annual budget is $4,431K)
(+) Revised budget from NASA as of April 2023 cuts Y3/4 budgets by 4% and keeps them flat in the future (no inflation)
(#) NASA requested a new proposal to fund Year 4 and beyond, the provided figures are their base guideline

Where are we now?

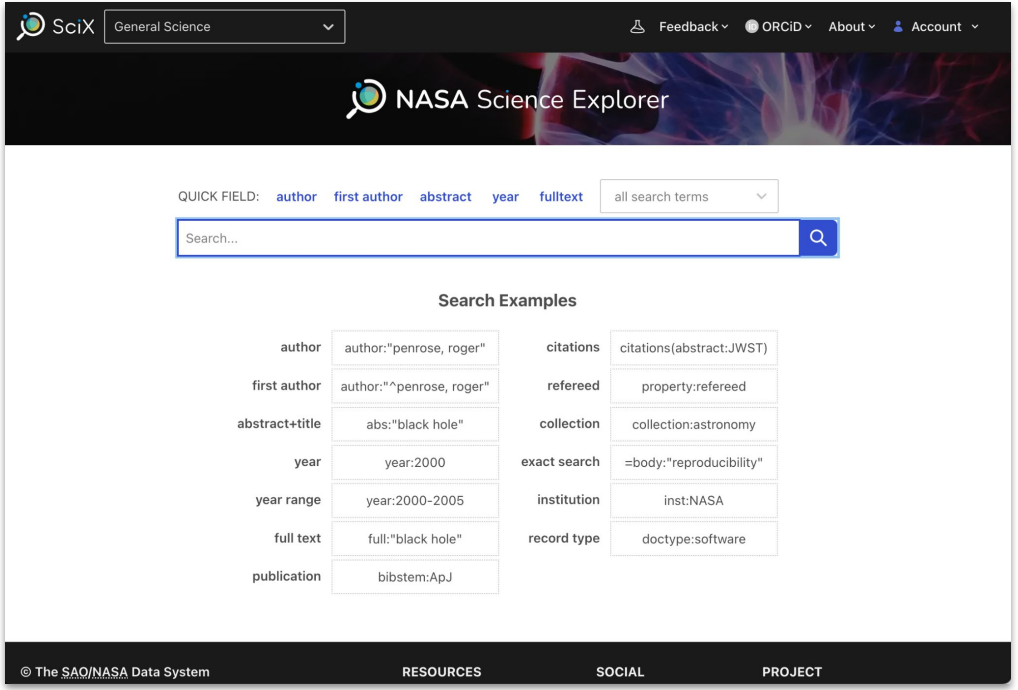
- We have an ambitious plan to grow ADS into a multidisciplinary system, and have received full support from NASA and SAO
- We are using the current ADS infrastructure, with an enhanced UI to provide content to the new system
- We are recruiting expertise and making connections with the relevant communities



https://SciXplorer.org

What have we accomplished in the last year?

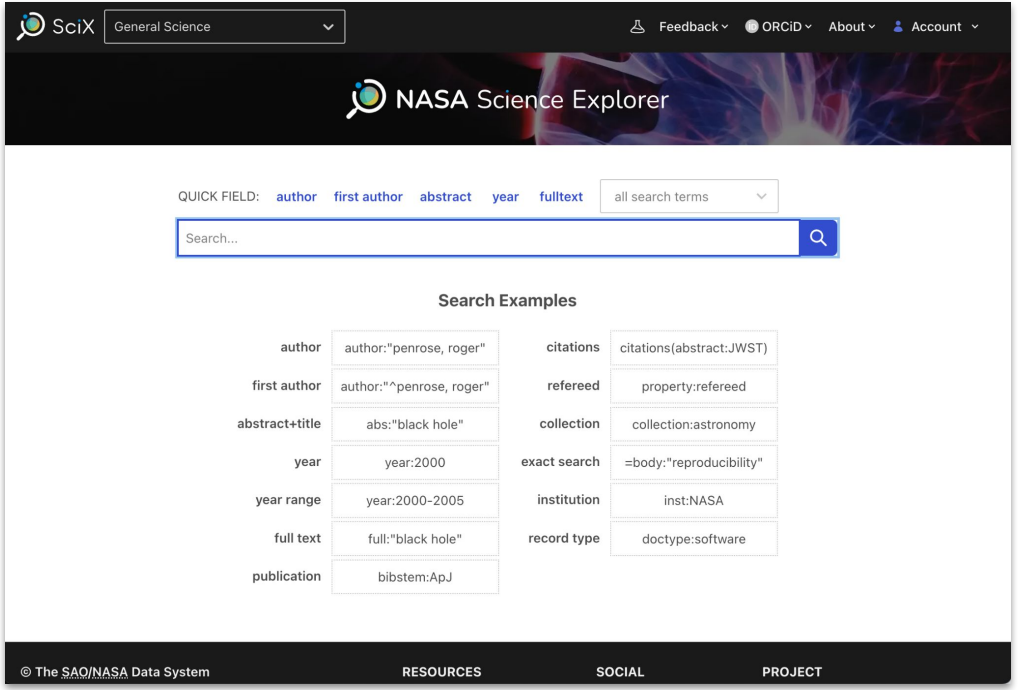
- We have restructured our team
- We suffered a resignation, hired six new employees and are on our way to add two more in the next 6 months
- We have engaged consultants for publisher relations, community outreach efforts and API development
- We have made decisions about what to ingest to support expansion
- We have gained further visibility within NASA and are expanding our collaborations



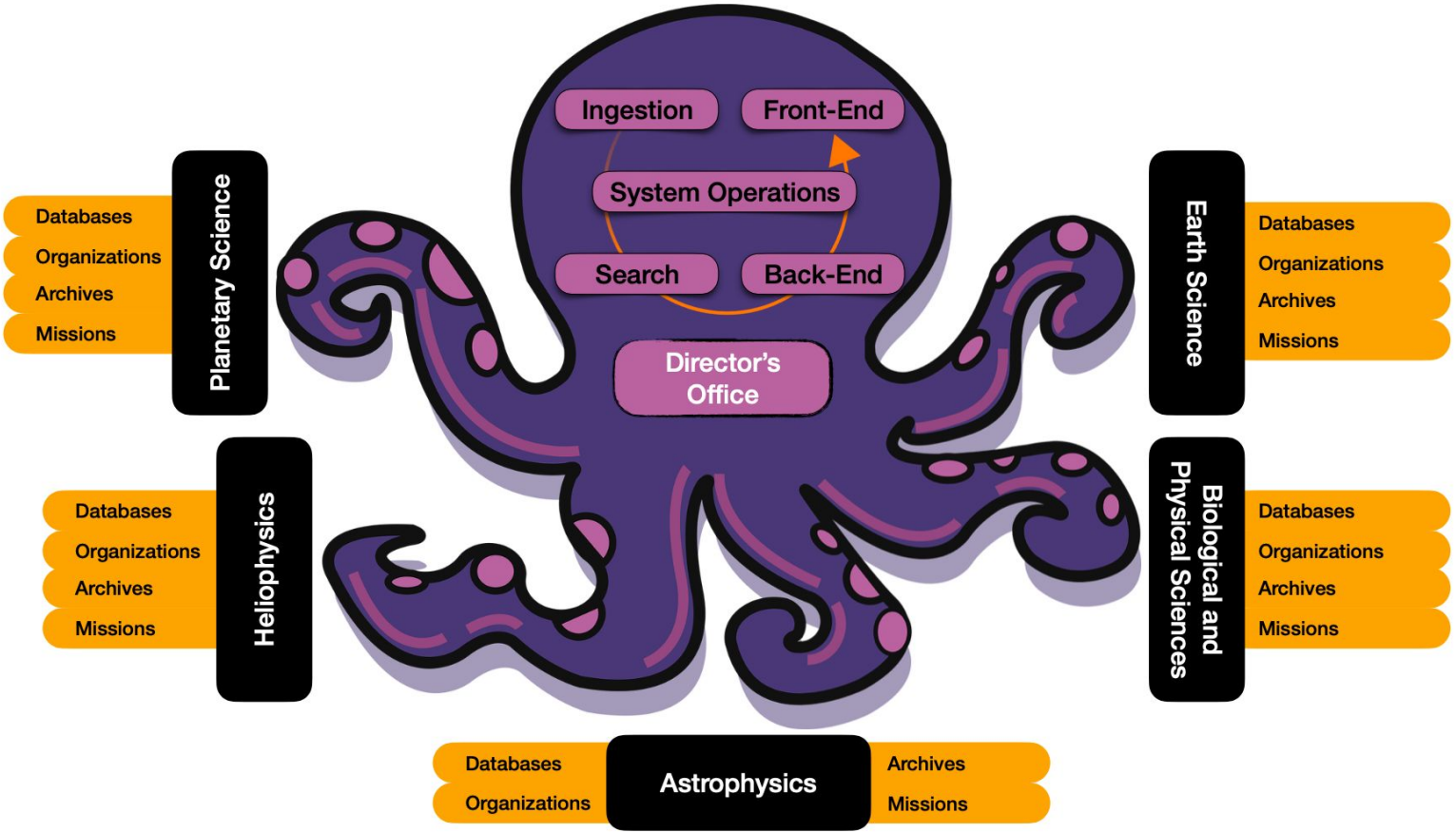
https://SciXplorer.org

What's coming next: Launch of SciX

- We have a new user interface to support the expansion, to be **beta** launched at the AGU meeting
- We are enriching metadata by text mining of data links, document classification, planetary names, UAT tagging
- A lot of features, workflows, decisions and fine-tuning still need to be finalized, these will be our highest priorities going forward



https://SciXplorer.org


Ingestion
Front-End
System Operations
Search
Back-End
Director's Office
Planetary Science
Databases
Organizations
Archives
Missions
Earth Science
Databases
Organizations
Archives
Missions
Heliophysics
Databases
Organizations
Archives
Missions
Biological and Physical Sciences
Databases
Organizations
Archives
Missions
Astrophysics
Databases
Organizations
Archives
Missions

Management and Hiring

- After a long journey, we had six new staff members join us during the last six months!
- Kelly Lockhart has taken over tech lead role and is now busier than ever
- Tom Allen is moving to the development team, still focusing on IE/ML backoffice tasks
- We hope to have an Earth Scientist onboard soon and then look for a system architect
- New hires are all on-site, but complemented by remote contractors for specific tasks

astrophysics data system

Dr. Stephanie Jarmak

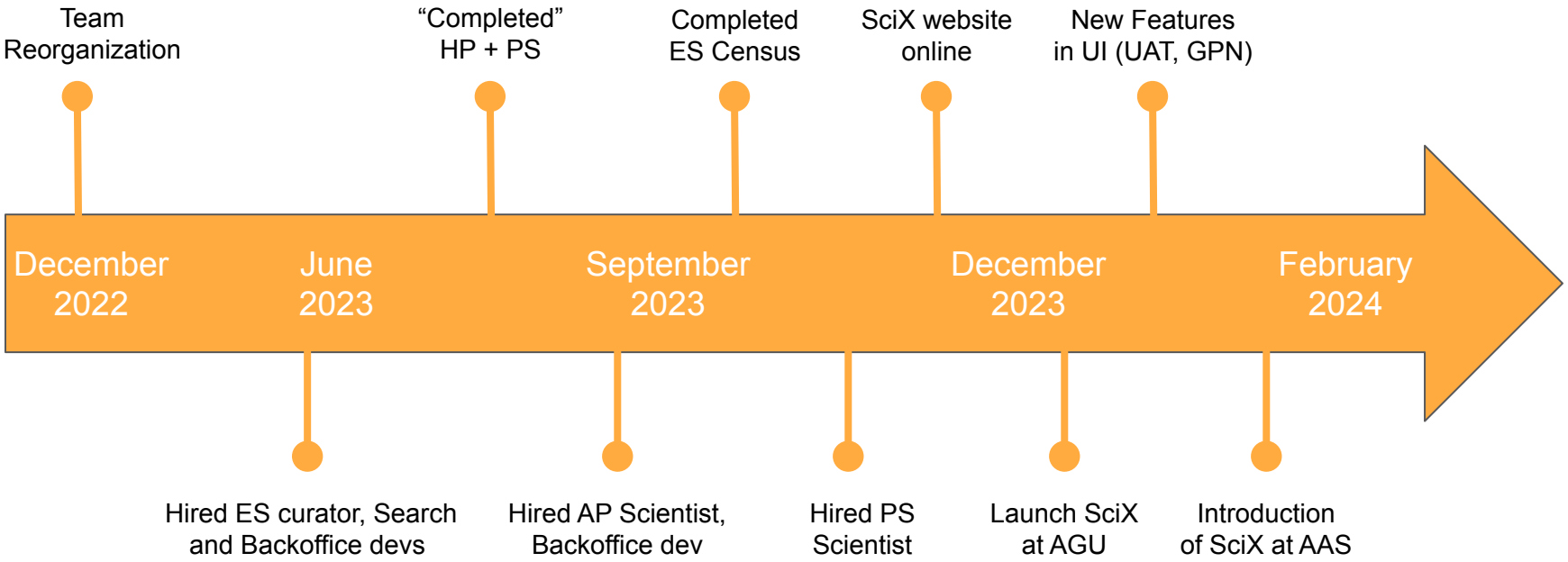
Stephanie Jarmak is the ADS project scientist for planetary science.

Dr. Jarmak studies asteroids, planetary rings, and grainflow processes on planetary bodies and contributes to the planetary community through her committee and social media group leadership roles.

Dr. Jarmak obtained a Ph.D. in Physics - Planetary Sciences from the University of Central Florida in 2020, a Master's in Physics from Texas A&M-Commerce in 2015, and a Bachelor of Science in Earth, Atmospheric and Planetary Sciences from MIT in 2013.

stephanie.jarmak [at] cfa.harvard.edu

Major Milestones in 2023



Overarching Major Milestones (**done** | in progress)

9/22 - 3/23: **Reorganization**, **recruit PS scientist, search & backoffice devs**.
Complete PS+HP refereed, **census of ES**, collaborate with SMD archives.

3/23 - 2/24: Hire **ES curator,** ES scientist, system architect
Ingest ES literature, preprints, **NASA STI**, launch SciX at AGU meeting
Updated search, author name lookup, metadata enrichment, **open datasets**.
Develop website, community engagement material and activities
Write NASA Proposal

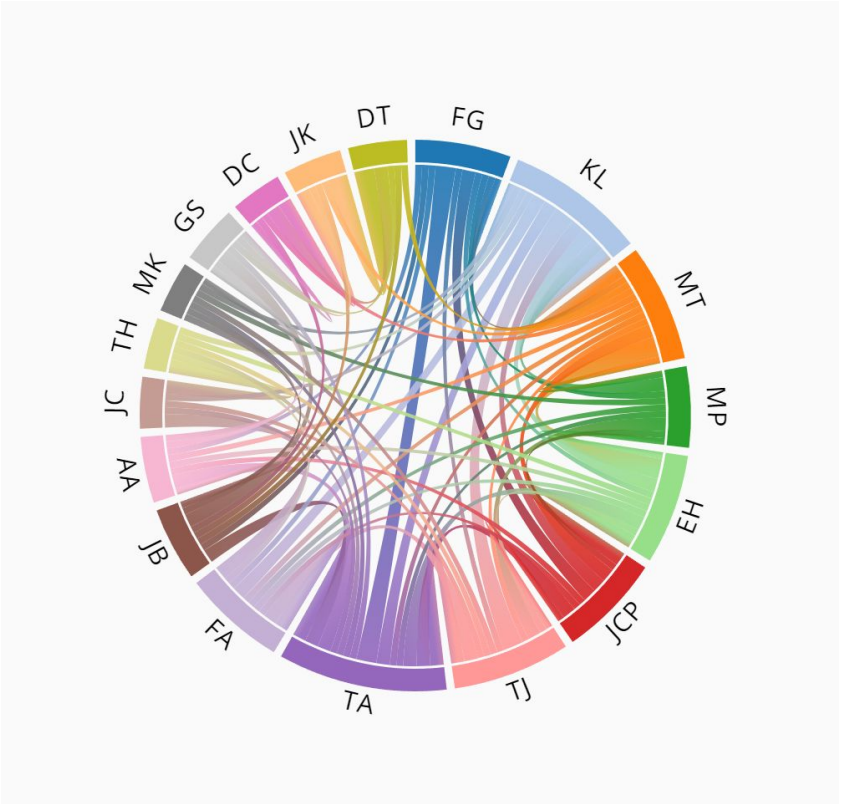
3/24 - 2/25: Hire ES/BPS librarian, curator, devops engineer
Census of BPS, ES gray literature, data indexing
Metadata enrichment for ES, search author profiles, PS object search

3/25 - 3/26: Hire R&D developer, back-end developer (interoperability)
Ingest BPS literature, cited ES literature, preprints
Public author profiles, authorship suggestions, filtering via taxonomies, language models

Team building and retreats

In 2023, we held three week-long in-person team retreats at CfA, the last of which was in September. Highlights:

- 22 People
- 18 Lightning Talks
- 5 Blue Sky Projects
- 6 Collaboration Time Projects
- Several One-on-one discussions
- Training on meetings that work
- Brainstorming with colleagues from NASA Science Discovery Engine



Backup Slides

Focus for September - December 2023

Complete the reorganization of our team:

- System architect
- ES Project Scientists
- Program Manager (PT)

Set up an advisory committee for SciX.

Ingest 80% of ES refereed literature

Establish collaborations with the ES archives

Make substantial progress on proposal writing

Implement an initial metadata enrichment pipeline incorporating AP + PS taxonomies:

- UAT
- Planetary Names

Achieve feature parity between Nectar and Bumblebee

Launch the new SciX interface at the fall 2023 AGU meeting with discipline-specific UI capabilities.

Scope of the Expansion

Astronomy: steady as we go

Planetary: some requests for better coverage of MPECs, gray literature

Helio: attempting to get better coverage for fulltext, citations

Earth: biggest effort so far - go after refereed literature (metadata, references, fulltext), then citations, gray literature, bibliographies

BPS: Not our focus at the moment, but on the radar for 2024