STScI | SPACE TELESCOPE SCIENCE INSTITUTE

EXPANDING THE FRONTIERS OF SPACE ASTRONOMY

Implementing the UAT in the Proposal Process at Space Telescope Science Institute (STScI): Downstream and Upstream Impact

Sarah Weissman, Senior Software Engineer
Jenny Novacescu, Chief Librarian

December 08, 2020 ● AAS-UAT

Who We Are

Space Telescope Science Institute (STScI)

We are the Space Telescope Science Institute in Baltimore, Maryland, operated by the Association of Universities for Research in Astronomy (AURA). We help humanity explore the universe with advanced space telescopes and ever-growing data archives.

Barbara A. Mikulski Archive for Space Telescopes (MAST)

MAST is a NASA funded project. MAST provides a variety of astronomical data archives focused on scientific data sets in the optical, ultraviolet, and near-infrared parts of the spectrum. MAST is located at STScI.

Background

STScI is responsible for:

- Science operations for the Hubble Space Telescope (HST), from proposal through data dissemination, 1990-present
- Science and flight operations of the James Webb Space Telescope (JWST), launch scheduled 2021
- Science operations for the Nancy Grace Roman Space Telescope (formerly WFIRST), launch scheduled mid-2020s
- Housing and making data discoverable from 20+ missions, past and present, including TESS, Kepler/K2, SDSS DR1, GALEX, IUE, EUVE, etc.

UAT life cycle from proposal to publication

1. Proposer selects UAT-aligned keywords that best describe broad scientific category and goals of the proposal.
2. Proposer selects UAT-aligned target description keywords.**
3. Proposal assigned to expert reviewers using scientific category keywords to match areas of expertise.
4. Target keywords ascribed to observational data in MAST via data pipelines and made searchable.
5. Published literature mapped to observational data through bibliography/DOI.
6. (In future) UAT keywords for publications compared to proposal and target keywords.
    - Increased understanding of full impact of programs and observations.
    - Improved searching in archives and literature; improved reporting.

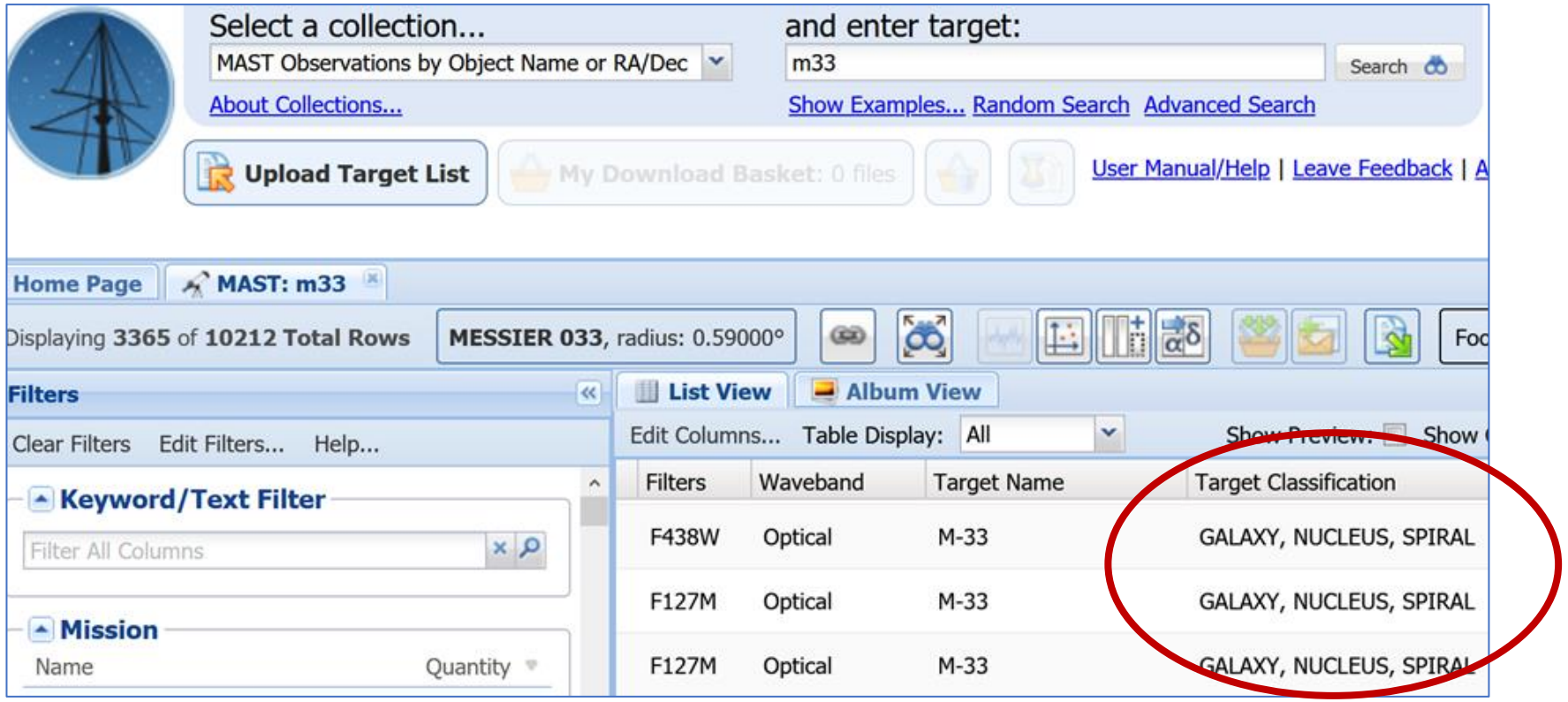
Astronomers Proposal Tool

** For JWST. HST target keyword selection happens after proposal acceptance in Phase II and keywords are not (yet) UAT-aligned.

UAT Integration Process at Space Telescope

UAT integration started with target description keywords (for JWST). See Frey & Accomazzi 2018ApJS..236...24F

- New mission + long-standing goal of finding a new vocab for targets and data tagging presented opportunity to integrate UAT.
- Target keywords most closely associated with observational data in MAST archive



Mapping Target Keywords

- UAT (2016) – 1047 “paths”, 9 levels, polyhierarchy
- Target vocab – 8 top level concepts, 238 concepts, 2ish levels, simple hierarchy
  - external/non-external
  - descriptors for all categories, e.g. ring, bar, jet, lobe
- Process
  - Manual mapping from old keywords to new
  - Stats on usage of each keyword (proposal count, last time used)
    - E.g. Star -> FK Comae, used 4 times, last time in 7/99
  - Review with 3 astronomers and 1 taxonomist
    - Include/Don’t include/Probably shouldn’t include
  - Score and rank

Updated Target Vocabulary

Updated vocabulary (https://jwst-docs.stsci.edu/jppom/targets/fixed-targets/target-descriptions)

- Currently just used for JWST

- 8 categories, 2 levels, ~400 concepts
- no external/non-external, no descriptors for all categories
- Some terms do not map to UAT (calibration, unidentified)
- Because of software limitations only human-readable (preferred) labels are stored
- A stars, A dwarfs, A giants, A subgiants, A supergiants vs. A0-A3 V-IV, A4-A9 V-IV, A0-A3 III-I
- Expansion of terms, esp. for ISM and Galaxy categories

Downstream Impact

After JWST target keyword implementation, it became apparent initial [HST Phase I] science keywords selected by proposers should also be UAT aligned. Impacted both HST and JWST, shared proposal keyword set.

- Assessed usage of existing 159 proposal keywords, grouped among 7 main proposal panels such as Solar System, Exoplanets, Galaxies, Stellar Physics
- Aligned each term to UAT equivalent; suggested addition term(s) to UAT if none existed.
- Evaluated new keyword set for narrower/broader terms to add from UAT, expanded from 159 to ~198, added one additional panel (IGM/CGM)

Benefits of UAT Keyword Alignment for Space Telescope

UAT alignment has benefited and will continue to benefit STScI/MAST in numerous ways:

- Systematic proposal keyword review and refresh
- Reporting: Easier to “roll up” keywords when requests for information about proposal metrics comes from Director’s Office, Science Policy Group, AURA corporate, or NASA/ESA.
- Links between Data, UAT, Publications provide a rich model for data/collections discovery
- Proposal keyword/reviewer matching: Currently dependent on text match of reviewers’ publication abstracts. In future, Phase I proposal keywords can be matched to reviewers’ literature keywords for further verification (dependent on ADS development). See Strolger et al. 2017AJ....153..181S

Benefits to UAT

UAT itself benefited from STScI keyword alignment. ~20% of Phase I keywords had no equivalent in in the UAT.

Alignment efforts at institution level helps expand UAT for worldwide community.

Examples of additions to UAT:

- Biosignatures
- disk formation keywords
- cooling flows
- quenched galaxies
- exoplanet terms (numerous)

Summary of UAT Mapping for Phase I Proposal Science Keywords

COUNT (OF 159)	DESCRIPTION/CATEGORY	APPROX. PERCENT
51	Exact Match	32.1%
74	Synonymous Match	46.5%
32	Additions to UAT	20.1%
2	Not Counted/Unique to Institution	1.3%

Future Initiatives

- Increase awareness of the UAT among research staff and in community:
  - Understand how to search, how to contribute; consider local integrations
  - Tag personal web pages; other research outputs not subject to vocab control
- Improve Archival Discovery
  - VO Registry keyword alignment. In progress (organized by Theresa Dower at MAST)
  - MAST collection discovery via graph databases, POC implementation in 2020. See: Weissman,"Building a Search Discovery System Powered By Graphs at MAST", ADASS XXX. https://schedule.adass2020.es/adass2020/talk/E8YXBY/
  - Additional crosswalks to other MAST metadata (High-Level Science Product)
- Align HST Phase II keywords to UAT to standardize across internal system
- Implement systematic maintenance of local vocabularies
  - Direct integration with UAT (URIs); possible with UAT API. Minimizes need for annual review and manual verification of terms currently in use
  - Establish annual review of proposal and target terms to ensure currency and inclusiveness for new science subfields
  - Develop process for deprecating terms from local vocabularies based on usage

Contacts

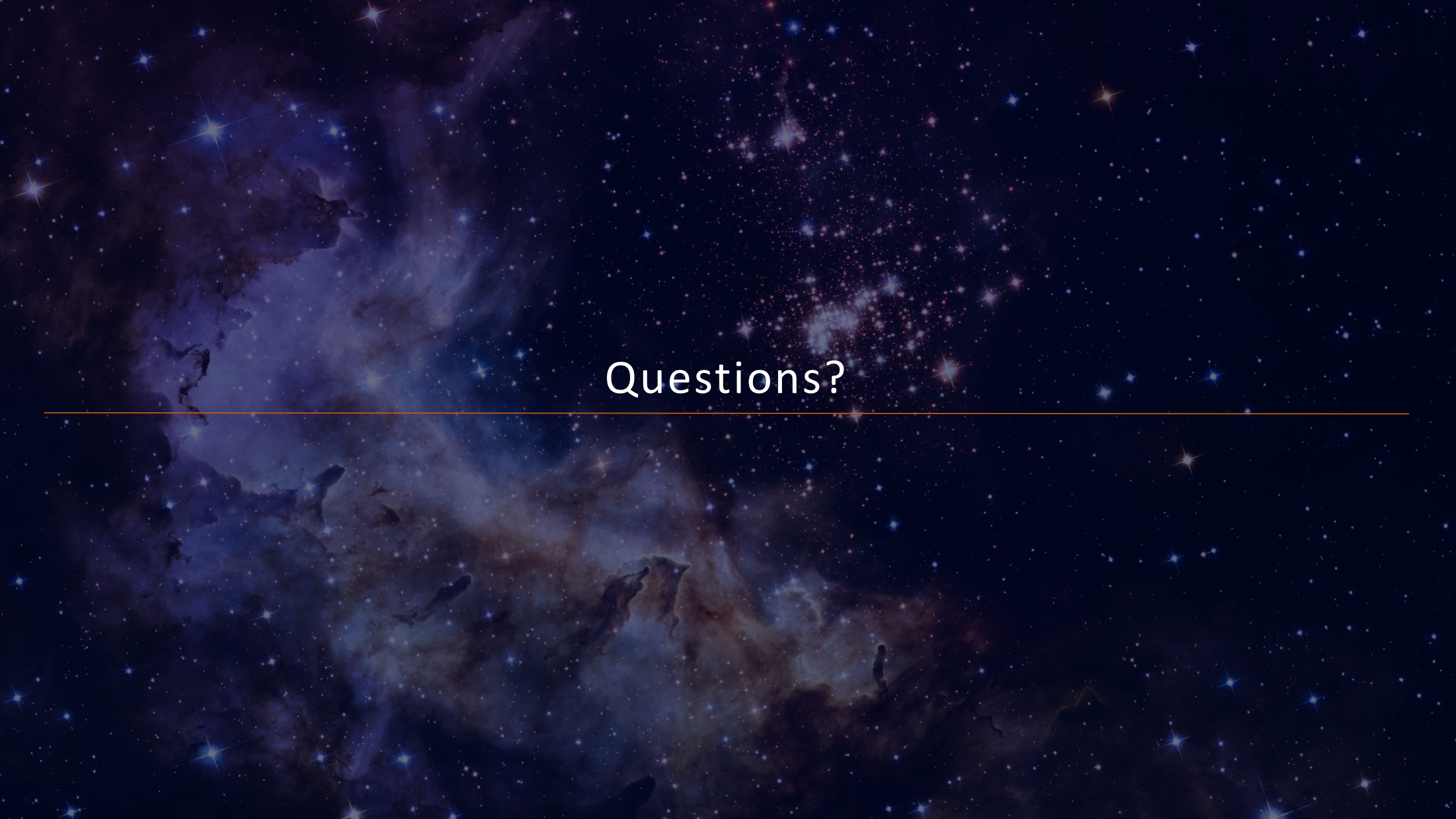
Sarah Weissman, Senior Software Engineer (MAST): sweissman@stsci.edu

Jenny Novacescu, Chief Librarian (STScI): jnovacescu@stsci.edu

General

STScI https://www.stsci.edu/ library@stsci.edu

MAST https://archive.stsci.edu/ archive@stsci.edu


Questions?


EXPANDING THE FRONTIERS OF SPACE ASTRONOMY