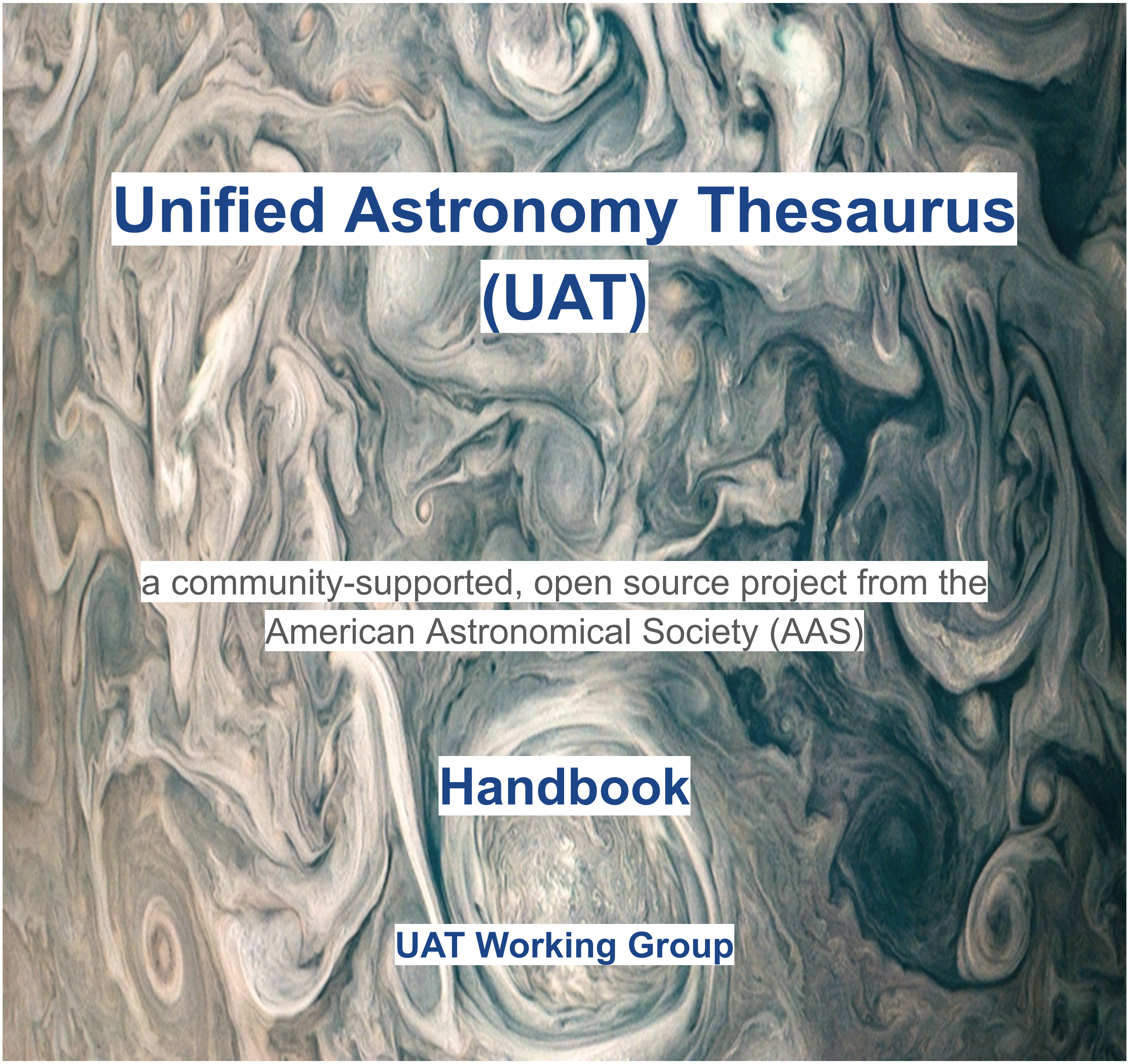

Unified Astronomy Thesaurus (UAT)

a community-supported, open source project from the American Astronomical Society (AAS)

Handbook

UAT Working Group

Version 1.0

August 2025

Table of Contents

Notes

- The UAT Handbook is a live document that is continuously revised and updated as needed.
- UAT website: astrothesaurus.org
- American Astronomical Society (AAS) Working Group on the Unified Astronomy Thesaurus: https://aas.org/comms/working-group-unified-astronomy-thesaurus-wguat

Executive Summary

The Unified Astronomy Thesaurus Handbook provides a discussion of the value of a controlled vocabulary of astronomical keywords, a historical background on the Unified Astronomy Thesaurus (UAT), a description of the UAT and how to use and implement it, and a discussion of the benefits the UAT offers to the astronomy and astrophysics research community.

The UAT is an open, interoperable community thesaurus that organizes astronomy concepts and their relationships into a single high-quality resource. The UAT is the most recent iteration of efforts to develop a keyword classification system to describe the astronomical concepts and entities found within literature, software, conferences, proposals, data sets, and other astronomical resources. The American Astronomical Society (AAS) has been owner and community steward of the UAT since 2012.

In the era of big data, well-chosen keywords are critical for turning a search query into meaningful results. High-quality keywords make literature searches faster and more effective; improve discoverability and help an author’s work stand out; assist readers in judging the suitability of a publication to their interests; and help science journalists and educators translate complex topics for a broader audience. The UAT broadens the impact of scientific work and facilitates interoperability with other systems, aiding in cross-disciplinary research.

The UAT offers many benefits to the astronomical community. For researchers in astronomy and astrophysics, use of the UAT in ADS aids the researcher in identifying relevant publications, as it includes term normalization, concept disambiguation, and concept hierarchy. For science policy and proposal management groups, it provides a tool to align internal terms with a community system. For publishers, the UAT provides an accurate and up-to-date vocabulary that reflects the view of subject matter experts. Use of the UAT ensures that the keywords and terms that describe publications improve the discovery of the content. High-quality metadata helps with indexing for searching and the growing number of automated discovery services that are based on AI.

For data archives, the UAT provides an unified vocabulary to aid in interoperability between archives. For bibliographers, the UAT provides a way to tag publications associated with a set of observations with standard astrophysical concepts, sources, and phenomena. For information science researchers, the UAT is a promising research tool for learning about the evolution and community dynamics of astronomy as an intellectual domain, as well as a platform for innovation in machine learning and natural language processing applications.

The UAT is a high-quality, freely available resource that is built upon authoritative sources and continuous community input, ensuring that its terminology is current and accurate. The primary mission of the UAT is to enhance semantic enrichment of the scholarly literature, improving how astronomical research is indexed, searched, and shared.

1 Introduction

The Unified Astronomy Thesaurus (UAT) is an open, interoperable, and community-supported thesaurus that reconciled divergent and isolated vocabularies from the fields of astronomy and astrophysics, such as the IAU Thesaurus, the Physics and Astronomy Classification Scheme, the Astronomy Subject Keywords, and others. The resulting Thesaurus is a single high-quality and freely available open thesaurus that formalizes astronomical concepts and their interrelationships. The UAT's primary mission is to support semantic enrichment of the literature, thereby enabling greater search and discovery across the astrophysics literature. In addition, the UAT is being used as a taxonomy with which to label other astronomical research products, such as software and datasets.

2 Why the UAT

The UAT provides a common vocabulary for search and discovery and shared descriptions of scientific literature, datasets, software, and other research products.

Since the beginnings of modern astronomy publishing, no single keyword system has consistently served the astronomical community across all platforms. Leading astronomy journals used the Astronomical Subject Keywords from the 1970s into the 2010s. However, those terms were last updated in 2013 and do not cover the latest topics in the field. Furthermore, those "keywords" do not include definitions or relationships between concepts. Therefore, the American Astronomical Society (AAS) journals and the Publications of the Astronomical Society of the Pacific (PASP) elected to adopt the UAT as their keyword system of choice in 2019 and 2020, respectively.

For similar reasons, the Astrophysics Data System (ADS), a digital library used by nearly all astronomers daily, has adopted the UAT as its preferred keyword system for the scientific literature, datasets, and software that it indexes. Section 6 discusses the integration of the UAT with ADS to make scientific literature and other research products more findable.

While it is currently possible to search ADS by keywords provided by publishers, there is no single vocabulary that has been consistently used throughout the indexed literature in ADS. For this reason, ADS users will be able to browse results using left-side facets in the query results screen or conduct an initial search using UAT terms. The ADS Team's goal is to promote the use of UAT concepts as a standard way to describe and discover records in its astronomy collection.

3 Thesauri 101

3.1. Definition of Controlled Vocabulary

Controlled vocabularies and thesauri existed long before the advent of computers and were originally designed as independent tools to standardize terminology within specific domains. But in today's networked world, it is now common for vocabularies to interact and integrate and to evolve into interconnected information resources.[1] From the technological standpoint, controlled vocabulary can be seen as a structured and standardized set of terms and phrases used to categorize, index, and retrieve information consistently. It serves as a tool to describe resources by using a predefined list of terms that represent concepts such as ideas, physical attributes, people, places, events, and subject matters. This organized arrangement of terms facilitates effective browsing, searching, and tagging of content, ensuring consistency and precision in information retrieval (Harping 2010; Leise et al. 2008).

In the context of metadata, a controlled vocabulary includes both preferred and variant terms, allowing flexibility in representation while maintaining a defined scope that aligns with specific domains. It eliminates ambiguity by providing clear definitions and relationships between terms, often structured hierarchically or associatively, as seen in thesauri. This hierarchical arrangement not only supports indexing but also enables collaborative development by allowing domain experts to refine terms and definitions (National Information Standards Organization, 2010).

The function of controlled vocabularies extends beyond indexing to encompass subject analysis or determining the "aboutness" of resources. By representing both the inherent and intellectual attributes of a resource—such as titles, creators, language, and content—controlled vocabularies facilitate subject access points in retrieval systems. These access points serve to locate and collocate related resources effectively, ensuring comprehensive discoverability across systems (Chan & Salaba 2015).

Controlled vocabularies also impose control over term usage and modification. Only authorized catalogers or editors may update terms, ensuring stability and reliability within the vocabulary's scope. This controlled approach is particularly beneficial for descriptive cataloging, classification, and indexing, where consistency in terminology is crucial for maintaining metadata quality (Hedden 2010).

[1] Allemang, Hendler, & Gandon, 2020, p. 310

3.2. Benefits of Controlled Vocabulary

The benefits of controlled vocabularies are extensive and essential in ensuring effective information management.

- They provide consistency in terminology, which is critical in reducing variability and ambiguity when describing concepts. This consistency ensures that information indexed under a term remains accessible and understandable, regardless of who created or accessed it.[2]
- Controlled vocabularies improve discoverability by enabling accurate and comprehensive search results. In systems such as digital libraries and institutional repositories, standardized subject headings and metadata ensure that users can retrieve all relevant resources without being hindered by inconsistent terminology.[3]
- They enhance metadata quality by ensuring accurate descriptions of resources. This structured approach prevents errors and inconsistencies, which can significantly hinder access and retrieval. Quality metadata facilitates interoperability across systems, allowing different organizations to share and integrate data seamlessly.[4]
- Controlled vocabularies serve as foundational tools for advanced data systems, enabling the integration of complex datasets while maintaining clarity and usability.

3.3. Definition of Thesauri

A thesaurus is a type of controlled vocabulary that includes hierarchical, associative, and equivalence relationships between terms. It enhances discoverability by providing richer contextual information and facilitating precise indexing.[5] It organizes terms into broader and narrower categories, as well as identifying related terms and synonyms. This structure allows users to explore terms in context, making it particularly valuable for complex indexing and retrieval tasks. For example, a thesaurus might include “active galactic nuclei” as a broader term, “radio loud quasars” as a narrower term, and “radio sources” as a related term. Additionally, synonyms like “quasi-stellar object” or “QSO” may direct users to the preferred term, “quasars.”

By integrating these relationships, thesauri facilitate nuanced subject analysis and ensure that users can navigate through information in an organized and comprehensive manner. They are widely used in specialized fields such as research databases and digital libraries where precision and depth are paramount.

[2] Harpring, 2010, p. 12
[3] Chipangila et al., 2024, p. 140
[4] Leise et al., 2008, p. 121
[5] Hedden, 2010, p. 281

3.4. Relationship and Differences Between Controlled Vocabulary and Thesauri

Controlled vocabularies and thesauri are closely related tools designed to standardize terminology and improve information retrieval, but they differ in scope and application. Both ensure consistency and accuracy in describing and categorizing information, serving as essential resources in metadata creation and indexing. However, thesauri are a more specialized form of controlled vocabulary, offering additional features such as hierarchical, associative, and equivalence relationships.

Controlled vocabularies provide a straightforward list of approved terms used for indexing and retrieval. They focus on standardizing language to eliminate ambiguity and ensure consistent tagging of resources. These vocabularies are foundational tools in systems requiring uniform terminology, such as library catalogs and institutional repositories (Hedden 2010, p. 281).

In contrast, thesauri build on this foundation by including relationships between terms. For example, while a controlled vocabulary might list “automobiles” as a term, a thesaurus would further categorize it into narrower terms like “sports cars” and broader terms like “vehicles.” It also identifies related terms and includes synonyms, guiding users in exploring the subject comprehensively (Leise et al. 2008, p. 122). This distinction makes thesauri particularly suited for complex and detailed subject analysis, while controlled vocabularies are ideal for simpler, more uniform applications. Both tools, however, are indispensable for ensuring clarity, consistency, and efficiency in managing and retrieving information (Hedden 2010, p. 281; Chipangila et al. 2024, p. 140)

3.5. Benefits of Thesauri

- **Hierarchical Structure**: Provides a framework for understanding relationships between broader, narrower, and related terms, improving navigation and retrieval.[6]
- **Inclusion of Synonyms**: Accounts for synonyms and variant terms, enabling users to find information even when different terminologies are used.[7]
- **Improved Search Precision**: Facilitates context-based retrieval by offering term scope notes and associations.[8]
- **Support for Complex Queries**: Enables users to construct detailed and nuanced queries by leveraging term relationships.[9]

[6] Hedden, 2010, p. 281
[7] Bagov et al., 2022, p. 2
[8] Harpring, 2010, p. 24
[9] Leise et al., 2008, p. 122

4 UAT Basics

4.1. History & Origins of UAT

The UAT is the most recent iteration of efforts to develop a keyword classification system that can be used to describe the astronomical concepts and entities found within literature, software, conferences, proposals, data sets, and other astronomical resources. The International Astronomical Union (IAU) Thesaurus, compiled by Robyn Shobbrook of the Australian Astronomical Observatory and Robert Shobbrook of the University of Sydney in 1992, is often considered the first such attempt (Shobbrook & Shobbrook 1992). The IAU Thesaurus was last updated in 1995 and contained 2,551 terms organized into a loose hierarchy. In 2008, an update to the IAU Thesaurus was created by Rick Hessman and labeled the International Virtual Observatory Alliance, or IVOA, Thesaurus. Hessman's work added about 300 terms, bringing his proposed IVOA Thesaurus to 2,889 terms.

Around the time of the initial development of the IAU Thesaurus, editors of the major astronomy and astrophysics journals worked together to create an organized set of keywords for classifying and indexing journal articles. This system is called the Astronomical Subject Keywords (ASK) and was adopted for use in 1992 (Accomazzi et al. 2014). Although the limitations of this classification scheme have become apparent in the intervening years, this set of terms is still currently in use by some of the journals in the field. The Astronomical Subject Keywords were last updated in 2013 and contain about 370 concepts organized by general subject area.

The Physics and Astronomy Classification Scheme (PACS) is a classification scheme containing roughly 400 astronomical terms, developed by the American Institute of Physics (AIP) in the 1970s to annotate the physics literature using a hierarchical set of codes. PACS was proposed in 1975, updated every two years, and served as the main organizational scheme for the major physics journals until it was decommissioned in 2011 in favor of developing a new system that would support modern indexing and searching technology.

In 2010, major publishers in physics and astronomy, together with other stakeholders from the IVOA, the NASA Astrophysics Data System (ADS), and the Harvard-Smithsonian Center for Astrophysics Library, began to explore new approaches to the various classification schemes in use. Specifically, the Institute of Physics (IOP) and AIP decided to collaborate, funding a project to merge the astronomy portion of their vocabularies with the enhanced IVOA Thesaurus to create a more robust thesaurus for astronomy. The resulting work was named the Unified Astronomy Thesaurus and was donated to the AAS, which holds ownership and copyright. To further develop and promote the UAT, the AAS established a Steering Committee, which reported to the AAS Publications Committee and included members from the AAS, ADS, the library and information science communities, and other astrophysics institutions. The AAS adopted the UAT in June 2019 for all of its journals as a replacement for the out-of-date ASK.

In 2024, an AAS Working Group on the Unified Astronomy Thesaurus (WGUAT) was established. This working group includes members from stakeholder representatives and general members. It is governed by a chair and vice-chair and reports to the AAS Board of Trustees.

4.2. UAT Scope Statement

Early on, it was decided that the UAT should contain all the terms necessary to completely describe the primary topics found in the mainstream refereed literature of astronomy and astrophysics. As the discipline expands over time, so, too, will the Thesaurus; for instance, additional terms describing exoplanets and gravitational waves will surely need to be incorporated into the Thesaurus within the next few years. In addition to deciding what should be included in the UAT, the boundaries of the Thesaurus have been intentionally delineated to define what must be excluded. Without a clearly articulated definition of scope, the task of maintaining a vocabulary can quickly get out of control as additional terms are incorporated. Concepts that are only tangentially connected to astronomy, or methods and techniques that are not specific to astronomy, are better described in external vocabularies, which can be referenced by, rather than incorporated into, the UAT. Similarly, existing taxonomies that detail named astronomical objects, instruments, and facilities are best kept outside of the UAT and managed by the existing groups that are currently curating them.

The curatorial decisions limiting which concepts are represented in the UAT and which are kept out do not necessarily mean that descriptiveness needs to be compromised for the sake of expedience: other thesauri can provide the concepts needed to properly describe a multidisciplinary work that could not be fully described using an astronomy-focused thesaurus. One of the features of the Semantic Web (Berners-Lee et al. 2001) is the ability to interlink distributed concept schemes, and the Simple Knowledge Organization System (SKOS; section 7.2) provides ways to formally map, reuse, or even delegate entire branches of a thesaurus to external knowledge systems.

Reusing a concept allows a system to “borrow” it from an external source via a simple referencing mechanism. A publisher could use this mechanism to say that the top-level UAT concepts represent the astronomy branch of a higher-level physics vocabulary, which could also include high-energy physics concepts to fully describe an interdisciplinary cosmology paper. Finally, the SKOS standard allows one to extend a thesaurus by delegating parts of it to outside knowledge systems rather than incorporating their contents. One useful application of this capability for the UAT could be the delegation of object types to the SIMBAD ontology (Derriere et al. 2010). This delegation works both ways and allows a publisher such as IOP to simply adopt the UAT itself as the astronomy branch of its physics thesaurus. It should be noted that the choice of relying on external knowledge systems involves making certain compromises, the most notable of which is a delegation of authority, and therefore control, over part of the domain being covered.

4.3. UAT Structure and Hierarchy

Although there are a range of options for organizing a controlled vocabulary—all the way from a flat list of unconnected concepts to a structured arrangement of precise relationships—the format of a thesaurus was determined to best support the goals of search and findability. According to the ANSI/NISO standard on controlled vocabularies, a thesaurus is a “controlled vocabulary arranged in a known order and structured so that the various relationships among terms are displayed clearly and identified by standardized relationship indicators” (ANSI/NISO Z39.19, 2010). This allows us to define some general relationships between concepts using standard SKOS properties (e.g., “broader,” “narrower,” “related,” “preferred label,” and “alternate label”) while avoiding overly complex structures.

Linking concepts together in child/parent relationships allows a user to follow a path from broader terms to narrower terms and to discover more specific terms to assist in finding the concepts that best describe their resource. This structure also helps to give additional context to each term, which can help clarify its meaning. More importantly, this hierarchy can inform searching and indexing applications that child terms are subsets of their parent terms and can be used to suggest related, alternative results to the user.

Relationships between concepts in the UAT are defined by ISO 25964-1 and ISO 25964-2. They include the following:

1) Broader – typically either a class/instance or whole/part relationship

2) Narrower – reciprocal with Broader, typically an instance/class or part/whole relationship

3) Use For – alternative labels, non-preferred terms, synonyms, aliases, acronyms

4) Related – connections across the hierarchy, but the nature of the relationship is not defined. Examples: “Solar flares” and “Stellar flares” are related because they are essentially the same thing but in different areas of study. “Horizontal branch” and “Horizontal branch stars” are also related, but in a very different way than the previous example. In this case, one describes an aspect of the theoretical model of stellar evolution, and the other describes stars/objects that would be found on that branch.

Using these relationships, the UAT is organized into a hierarchical structure under 11 main categories. It’s important to note that concept depth in this structure does not imply anything about the value of that concept. It’s simply an organizational structure, more typically related to how expansive current literature on the topic is or how “new” the field is. As an example, most exoplanet astronomy terms are no more than four levels deep. Fields of research with longer history, such as galactic and extragalactic astronomy, can be five or more levels deep. Terms that happen to fall at level 5 in the galactic section and terms at level 4 in the exoplanet or solar system sections are not more or less important than one another. Similarly, two concepts at the same level in different sections of the UAT does not imply that these two concepts are equally granular.

5 Benefits of a Standardized Vocabulary in the Research Community

5.1. Benefits to Researchers in Astronomy and Astrophysics

The concepts in the UAT create a standardized and consistent set of keywords that help to codify the language used to describe astronomy. Drawing descriptive metadata from a controlled vocabulary helps to assure that researchers are all using the same words and concepts to describe similar astronomical phenomena. Even though full text searching is becoming more common, if two researchers use vastly different words to describe similar events in their latest papers, even the most sophisticated AI program might not be able to make connections between their two documents.

Use of the UAT in ADS will aid the researcher in identifying relevant publications as it will include term normalization (allowing for multiple expressions of the same concept to be captured in a single entry), concept disambiguation (allowing for different meanings of the same words to be specified depending on their context), and concept hierarchy (providing a structured way to navigate research topics).

As more publishers adopt the UAT and the UAT is integrated into ADS, the WGUAT envisions a future where a researcher can click on a hierarchical concept and then expand or drill down to additional concepts to aid in discovery and research using the literature, a data archive, or proposal software system.

Having an expansive, polyhierarchical vocabulary at one’s fingertips enables researchers to better describe their own research findings; they are not tied to a finite, antiquated, or imprecise list of keywords that may or may adequately capture new phenomena or a deeper understanding of existing concepts. As noted in Section 2, drawing descriptive metadata from a controlled vocabulary helps to assure that researchers are all using the same words and concepts to describe similar astronomical phenomena, thereby aiding each other's mutual discovery of new research.

5.2. Benefits to Science Policy and Proposal Groups

A community-shared and -controlled vocabulary benefits science policy groups and those responsible for managing the proposal process at their institution in a number of ways. By aligning your internal terms with a community system, individuals naturally adopt a systematic process to review, update, and continually automate the maintenance of terms. This helps avoid a stagnation in internal vocabularies. By having the entire community's input and vocabulary terms available in a hierarchical structure, it is also easier to be more or less expansive, depending on your needs.

With a more precise and hierarchical vocabulary, it is easier to match proposers to those most capable of reviewing the content from a scientific perspective. This will improve even more once the UAT is fully integrated in ADS and reviewers' literature is pre-mapped to UAT terms.

The hierarchical nature of the UAT has simplified the reporting process. When asked to report on how many proposals support a particular field such as galaxy formation, it is now possible to group terms from within related UAT branches or group parent and child terms.

5.3. Benefits for Publishers

The American Astronomical Society (AAS) has been owner and community steward of the open-source UAT since 2012 and established the UAT Steering Committee (now WGUAT). The AAS research journals include *The Astronomical Journal, The Astrophysical Journal, The Astrophysical Journal Letters, The Astrophysical Journal Supplements, The Planetary Science Journal,* and *Research Notes of the AAS*. The Astronomical Society of the Pacific's *Publications of the ASP* has also adopted the UAT. AAS authors are now asked to assign UAT concepts to their submitted manuscripts.

It is a challenge for publishers to ensure that their vocabularies grow to reflect new research areas. The community ownership of the UAT helps in many ways:

- Community ownership ensures that the coverage of the UAT meets the community's needs, providing publishers with an accurate and up-to-date vocabulary that reflects the view of the subject matter experts. Also of value is the process by which the community agrees to the correct labels to describe terms. There has been some debate on this already in several areas of the UAT, in particular planetary and exoplanetary terms. Publishers would find it slow and costly to get this level of subject expertise for proprietary vocabularies.

- With wide adoption the UAT offers a single vocabulary to provide a more cohesive user experience. This should extend well beyond publishers, to grant funding, data gathering, data publication, preprints, searching, and alerting. Publishers can be confident that their metadata aligns with community applications and that users are familiar with the terms through other systems.
- A strong motivation for publishers to adopt the UAT is to ensure that the keywords and terms that describe publications improve the discovery of the content. High-quality metadata helps with indexing for searching and the growing number of automated discovery services that are based on AI. The UAT makes a high-quality, detailed vocabulary accessible to all discovery services due to its interoperability and open source code.

5.4. Benefits to Data Archives

Adoption of the UAT is being pursued by several data archives and interoperability standards, including the following:

- Catalogs in the CDS's VizieR archive are assigned searchable UAT keywords. For example, search for *uat:cataclysmic* here and see the UAT field on the description page for each catalog.
- The IVOA's interoperability standards have adopted a mirror of the UAT as the vocabulary for the *subject* field in Virtual Observatory (VO) resources. A thing can be considered a resource if it is "*a VO element that can be described in terms of who curates or maintains it and which can be given a name and a unique identifier ... it can be an abstract idea, such as sky coverage or an instrumental setup, or it can be fairly concrete, like an organisation or a data collection*" (Plante et al. 2025).
- STScI's JWST proposal keywords, including target classification and science keywords, are mostly aligned to the UAT, and the former propagate into the MAST archive. These keywords are also used to help assign reviewers to JWST proposals.
- Different sections of STScI are investigating the possibility of further adopting the UAT in various contexts, including the MAST archive and the Hubble Spectroscopic Legacy Archive data collection.

Tagging astronomical datasets with science areas or target/object type classifications is often a requirement for observatories and their data archives, with the potential to enable user queries like "*show me all observations where the intended target was a cataclysmic variable*." The use of such metadata in search applications has at times been stymied by databases that use multiple contradictory vocabularies, with some observations tagged with "cataclysmic variable," others with "cataclysmic binaries," and others with "CataclyV," all meaning the same thing. This is a particular problem for interoperability between archives, as well as within archives hosting observations from multiple observatories, spanning multiple teams and decades during which vocabularies and metadata standards shifted.

The UAT offers a resolution to this conundrum: convert or map all legacy keywords into stable URIs from a unified vocabulary and encourage providers of new incoming data to switch to providing stable URIs from a unified vocabulary. For example, this URI — https://astrothesaurus.org/uat/203 — will always correspond to the concept of cataclysmic variables, even if someday we swap out our currently preferred label "Cataclysmic variable stars" for one of our alternative labels like “Cataclysmic binaries.” Furthermore, search applications could take advantage of the hierarchical structure of the UAT to empower users with options for query expansion to related keywords, such as expanding a search for “Cataclysmic variable stars” to also include data tagged with “Novae.”

Adoption of the UAT by journals and ADS (Sections 5.3 and 6) also offers exciting opportunities for cross-travel between data and the literature, using UAT concept keywords as the bridge. The scientific significance of archival observations often evolves over the life of the data archive and extends well beyond that anticipated by the proposers of the observations. Keeping track of this evolution by tagging observations with terms reflecting their later use in the linked literature makes it possible to track the evolving significance and scientific legacy of data holdings and find relevant observations based on common vocabularies.

5.5. Benefits to Bibliographers

Curated mission bibliographies are powerful tools to assess the scientific impact of a mission and allow astronomers to back-search for observations that have contributed to one or more publications. The granular linking of observations and publications that lies at the core of the bibliographic effort is asymmetric, however: the wealth of metadata and properties associated with each observation is not matched by a similarly comprehensive characterization of the scientific concepts discussed in the linked publications. In other words, it is relatively easy to retrieve publications linked to a selected set of observations, but the opposite is not true, leaving only connections to strictly bibliographic information (bibcode, journal, authors info, authors' institutions, year of publication, etc.). The UAT provides a great opportunity to tag publications associated with a set of observations with astrophysical concepts, sources, and phenomena during the data linking procedure, either by ingesting the UAT keywords applied to the manuscript during publication or by manually classifying the research paper if no UAT keywords have already been applied.

5.6. Potential for Integrating Data Archives with Publication Archives

Currently, public data can be searched through a combination of spatial, instrumental, time and target-related criteria that often do not carry hints about the information content and the astrophysical value of the data. Even in the case of data tagged with science categories and target types from their associated observing proposals, these tags only mirror the a priori assumptions about the scientific content of the observation. Unpredictable usage of the data, serendipitous discoveries and insights derived from the aggregation of heterogeneous observations cannot be reflected in any of the properties usually employed to identify the observations. The gap between the original predicted content of observations and their accumulated, post-analysis scientific value only increases with time, making the same set of "classical search criteria" more and more inadequate as archives grow.

One way to work around this fundamental issue is by leveraging the community-based semantic enrichment of the literature allowed by the UAT. This can be done by incorporating the papers' UAT annotations into the archives, linking the annotations with the relevant observation and by making the UAT terms searchable. In this way, the UAT will bridge the divide between the nominal properties of data and their actual scientific value.

The archives of many existing missions already link their data to scientific publications as a way to build a comprehensive bibliographic record of their impact. The UAT provides an opportunity to revolutionize how searches are performed and serve the astronomical community with an evolving concept- and idea-driven new framework for archival queries that extends into regions of the search parameter space inaccessible in any other way.

5.7. Benefits to Information Science Researchers

As a curated and openly available controlled vocabulary, the UAT is also a promising research tool for learning about the evolution and community dynamics of astronomy as an intellectual domain, as well as a platform for innovation in machine learning and natural language processing applications.

Broadly speaking, controlled vocabularies offer robust capabilities for subject classification and efficient database searching. Classification schemes like the UAT also inevitably reflect epistemic social structures that can be analyzed by social and information scientists as cultural information systems. Beyond subject classification and societal impact, controlled vocabularies have been used to measure field development in other domains, such as medicine and physics. Furthermore, in the areas of natural language processing (NLP) and machine learning, ongoing innovations in supervised and unsupervised machine learning, text analysis, and computational linguistics techniques have opened a new research frontier in which text data can be used on a large scale to address pertinent research questions.

UAT terms will soon be automatically assigned to previously published documents, utilizing modern methodologies through which it should be possible to accurately predict terms for untagged documents, and at the same time methodologically informing other disciplines attempting a similar task. Finally, the UAT can also be used to observe how the use of concepts changes over time through what can be referred to as "semantic drift" or "concept drift." Overall, these and other research efforts depend on widespread use of the UAT over time, ensuring that a sufficient number of documents are tagged with UAT terms, to fully exploit the UAT as a research tool for social science and information science.

6 ADS Integration

ADS is a NASA-funded digital library that has served astronomers since the 1990s. Globally, the average astronomer uses ADS daily to find scholarly literature, datasets, and software. Its astronomy collection includes 3 million records,[10] including scanned articles dating to the 19th century. At the request of NASA, ADS is expanding as the Science Explorer (SciX) to serve researchers in heliophysics, planetary science, and earth science as well. Consequently, the overall SciX holdings are over 26 million records.[11] While enabling astronomers to find relevant scientific research has always been the ADS mission, increasing content brings the challenge of potentially irrelevant results and the opportunity of greater cross-disciplinary discovery.

As mentioned earlier, a single consistent keyword for astronomy literature and research products would increase the findability of such works. Inconsistent or unavailable keywords reduce the results returned from a search. With the inclusion of earth science research products in SciX, the need to describe astronomy research well becomes critical. Increasing the findability of and access to research output should also increase the reuse of valuable scientific assets.

SciX is developing Artificial Intelligence (AI)-enhanced tools to assign UAT keywords to the research papers and other research products it indexes. The AI-generated keywords are visible now on the SciX abstract view of a paper. Because SciX retains a curatorial staff, corrections to individual records are possible with such feedback improving future algorithms. SciX is considering additional exploratory features using the UAT keywords for future development.

SciX assigns UAT URIs to papers but displays the preferred label as the keyword. Therefore, changes in preferred labels will propagate seamlessly.

For other disciplines, SciX plans to develop additional AI-tools to assign other keyword systems, such as the Global Change Master Directory (GCMD) keywords used in earth science,[12] to

[10] Holdings 2025 May 20
[11] Holdings 2025 May 20
[12] NASA Earth Data, GCMD Keyword Viewer, https://gcmd.earthdata.nasa.gov/KeywordViewer/scheme/all?gtm_scheme=all

papers in those collections. Eventually, SciX will link similar concepts in the different controlled vocabularies to facilitate cross-disciplinary discovery.

7 Linked Data, Interoperability, and FAIR

7.1. FAIR Vocabularies

In 2016, a group of scientists and organizations that support open practices and the reusability of data published an article (Wilkinson et al. 2016) introducing the FAIR Data Principles. FAIR is an acronym for Findable, Accessible, Interoperable, and Reusable that represents a set of guiding principles describing how to make data open and available, especially as more complex data are being created in larger volume and faster than ever before.

Since then, these principles have been adapted to apply to information resources outside datasets, including a set of guidelines that govern FAIR Vocabularies. For a vocabulary to be FAIR, it should be

- **Findable**: be registered (indexed, listed) in a vocabulary service
- **Accessible**: be available on the web, downloadable
- **Interoperable**: encoded in a standard representation, such as the Web Ontology Language (OWL) or SKOS and domain-specific extensions
- **Reusable**: licensed and maintained, ideally with an open license

7.2. Semantic Web / Linked Data / Interoperability

The UAT is built using Semantic Web and Linked Data standards. Publishing a vocabulary as Linked Data means it is on the web and findable using a unique identifier, or URI, and that when the URI is retrieved, a machine-readable representation of the concept is returned.

To be part of the Semantic Web, the UAT uses SKOS, which is a data model that provides a standard, lightweight, and intuitive language for developing and sharing vocabularies, taxonomies, and thesauri. SKOS includes a list of specific metadata labels, such as “skos:broader,” “skos:narrower,” and “skos:related” which create semantic links between different concepts within the UAT.

SKOS also provides mapping properties, which are used to define links between SKOS concepts found in different vocabularies or thesauri. These relationships (such as “skos:closeMatch,” “skos:exactMatch,” and “skos:broadMatch”) can be used to connect the UAT to concepts in other schema such as Wikidata, OpenAlex, or PhySH (the Physics Subject Headings, from the American Physical Society).

Concept mapping in this manner provides a way to establish semantic similarity between a term in the UAT and an external knowledge base. Using concept mapping allows a curator for a project such as WikiData (Vrandečić & Krötzsch 2014) to link concepts in Wikipedia to corresponding concepts in the UAT, stating, for instance, that the UAT concept for “G stars” is equivalent to the Wikipedia category “G-type star.” As more resources are linked to WikiData, such curation allows for cross-vocabulary mapping, which will increase interoperability across knowledge systems and disciplines. This is the approach behind the Linked Data effort (Heath & Bizer 2011), which attempts to publish and interlink structured data on the web. While this is, in general, a good thing, one should be aware of the fact that any such mapping is subject to interpretation and approximation, which means that the result of these efforts may present new challenges to those attempting to leverage them.

8 Steps/Considerations to Implementing UAT

8.1. Technical Recommendation

This outline describes recommended approaches for implementing the UAT. It does not describe the code or functions necessary to integrate the UAT into a platform; rather it is meant to highlight what a developer should or could aim for when working on their software.

Initial Implementation

- Download the UAT.rdf file from the most recent release on the GitHub Repository
- Upload it as a local file for your implementation.
- Build an application suitable for your platform that uses your local RDF copy of the UAT as its source for keywords. (See “Recommended Levels of Implementation” below.)
- Make sure to include the concept identifier or URI along with the human readable labels.
    - For publishers, it is highly recommended that concepts be provided to ADS as full URIs. If URIs are not technically feasible, then concepts should be provided in the alternative format, which consists of the human readable label followed by a space and then the ID in parentheses.

      Examples:
        - Full URI (preferred): http://astrothesaurus.org/uat/1988
        - Alternative format: Active solar corona (1988)

Updating Your Implementation

- In a developer or test environment for your application, replace your local RDF copy of the UAT with the new version.
- Test your application to make sure everything works as expected.
- Update your application as needed.
- On, or shortly after, the official release date, implement the updated application on your production environment and replace your local RDF copy of the UAT with the new version.

Standalone Tools

Concept Selection Widget: https://astrothesaurus.org/concept-select/

Source Code: https://github.com/AASJournals/uat-chooser

Recommended Levels of Implementation

Flattened Concept List

- Your application includes all concepts from the UAT.
- Your application always keeps the URI or numeric identifier attached to a concept at all levels of usage.
    - The numeric identifier is the set of numbers found at the end of a concept URI.
        - URI: http://astrothesaurus.org/uat/1988
        - Identifier: 1988
    - The preferred format consists of the human readable label followed by a space and then the ID in parentheses.
        - Example preferred format: Active solar corona (1988)
    - Your application allows users to view information about a concept, including its parent and child concepts.
    - Your application allows users to select concepts to apply to content.

Basic Hierarchical Support

All of the above, plus

- Your application may allow parent/child concepts to be navigable and viewed.

Full Hierarchical Support

All of the above, plus

- Your application includes the ability to travel up and down the hierarchy, viewing content associated at these various levels.
- Your application allows you to roll up or group content by parent concepts.
  - See cautionary note under "Does the UAT support transitive inference?" FAQ question below.

Full SKOS/XML Implementation

All of the above, plus

- Your application allows for the usage of multiple SKOS-based vocabularies, which could be used to fill in gaps.
  - For example, a journal on astronomy education may wish to include pedagogy concepts that are not in scope for the UAT.

Specialized Implementations

These specialized implementation guidelines would all require additional maintenance above and beyond simply updating the UAT RDF file.

Partial Subset

- Only recommended for situations in which content is naturally limited, such as data systems, telescope proposal systems, etc.
  - For example, certain telescopes can only observe certain targets, or are funded for specific types of research.
- Publishers are strongly encouraged to utilize the full UAT in their implementation so as to not impose artificial limitations on authors.

Custom URIs

- A UAT mirror with custom URIs. (IVOA)
- Only recommended for edge cases where numeric IDs are not technically feasible.
- Each concept should point to the original UAT URI using the skos:exactMatch or owl:sameAs relationships.

Custom Organization

- Similar to custom URIs, but in addition to using new URIs, concepts are also organized in a custom fashion.
- Only recommended in edge cases where the existing UAT hierarchy is not useful.

9 FAQs

Is the UAT Open Source?

Yes! The UAT is available under the CC-BY-SA license, meaning that anyone is allowed to use, share, and adapt the UAT, provided that they give appropriate credit to the UAT and the American Astronomical Society, and apply the same license to any adaptations. The UAT is meant to be a tool for the astronomical community, and the more it is adopted, the more useful it is for everyone.

Is the UAT FAIR?

Yes! See Section 7.1 Fair Vocabularies for more information.

Does the UAT change?

Yes! The UAT is actively curated according to its Curation Process, a process that evolves with community needs. New versions are released approximately annually according to the UAT Release Cycle.

Can I use the UAT to categorize things other than astronomy research articles?

Yes! Organizations are using the UAT to categorize things such as observing proposals and datasets as well as research articles. In general, if it makes logical sense to apply a UAT term to something, go ahead! By using the UAT, you then make it possible to interrelate your things to all of the other articles, datasets, proposals, etc., that have also been UAT tagged.

Who is in charge of the UAT?

The UAT is overseen by the AAS Working Group on the Unified Astronomy Thesaurus (WGUAT). Membership of this group includes stakeholders representing publishers, libraries, data providers, and the astronomy community.

There isn't a concept in the UAT that exactly covers my topic. Now what?

In this case, there are a few approaches you can take:

1) Use two concepts that together cover the topic you are interested in. For example, a paper about the atmosphere of Venus could use the concepts "Planetary atmosphere" and "Venus" to fully describe their work.

2) Search for broader or narrower concepts.

3) Recommend a change to the UAT. The UAT thrives on community input! See the UAT Contributions web page for information about how to get involved. In most cases you might start by creating a GitHub issue to recommend a new concept or suggest other changes. Please include contextual information, such as background information, and suggest where new concepts could live within the hierarchy.

Some new concepts might be appropriate for your use case, but not within scope for inclusion into the UAT. In these situations, consider tagging your content with multiple vocabularies, as described below. If you envision a larger contribution, adding or reorganizing many related terms, you can still start with GitHub, or you might wish to engage more directly with one or more members of the WGUAT. A major responsibility of the Working Group is to understand the needs of UAT users and guide its development to meet them.

My journal or data archive contains both astronomy and geophysical (or some other topic) content, but the UAT is only focused on astronomy. How do I describe non-astronomy content?

As a content publisher or manager, consider SKOS vocabularies more generally as a solution for descriptive metadata. As mentioned in Section 7. Linked Data, the UAT follows common standards for building vocabularies and thesauri, and these standards are built with the idea that different vocabularies can—and should—be used in concert.

Is there a mapping between SIMBAD's object type vocabulary and the UAT?

A mirror of the UAT is the IVOA's subject vocabulary for VO Resources, while a mirror of SIMBAD's object type vocabulary is the IVOA's preliminary draft for an astronomical object type vocabulary. One can imagine, therefore, situations where you might want to convert between the two, while using the UAT in some contexts and SIMBAD's object type vocabulary in other contexts. Fortunately, our IVOA colleagues have mapped most SIMBAD concepts to their corresponding UAT concepts in the underlying RDF file available from the IVOA, using the *skos:exactMatch* property (see Section 7.2). It thus should be relatively easy for databases and downstream applications to convert from SIMBAD concepts to UAT concepts.

The reverse, converting from UAT concepts to SIMBAD concepts, may be more challenging at this time. The UAT includes many concepts, including more specific subtypes and components of astronomical object types, that are intentionally not within the scope of SIMBAD's vocabulary. Many narrow UAT concepts surely have a broader SIMBAD concept that could be mapped to with the *skos:broadMatch* property, but at this time we have not performed that mapping.

Can the UAT support transitive inference, with automatic query expansion from a concept to all of its descendants in the vocabulary's SKOS hierarchy? Put another way,

the contextual meaning of some UAT concepts seems to change depending on the path through the tree that you took to get to them; what's going on?

We plan to develop support for transitive inference in the future, but currently the UAT's concept hierarchy does *not reliably* support transitive inference. At the moment, caution is warranted if you are considering automatic query expansion to a concept's hierarchical descendants.

Imagine, for example, that a concept "Active galactic nuclei" has a descendant concept "Radio jets," either as a direct child concept or as a more distant descendant. Automatic query expansion to a search term's descendants would cause a search for "Active galactic nuclei" to yield all results tagged with "Radio jets" too. This may be dangerous, because radio jets appear in other astrophysical contexts (e.g., protostars) unrelated to AGN, polluting your AGN researcher's search results with unwanted protostar content.

We plan to revise the UAT to change this, which will take a lot of work! If you notice other examples, you can help us out by reporting them in a Github issue.

10 References

Resources Cited

Accomazzi, A., Gray, N., Erdmann, C., Biemesderfer, C., Frey, K., and Soles, J., 05/2014, The Unified Astronomy Thesaurus, *Astronomical Data Analysis Software and Systems XXIII*. Proceedings of a meeting held 29 September – 3 October 2013 at Waikoloa Beach Marriott, Hawaii, USA. Edited by N. Manset and P. Forshay ASP conference series, vol. 485, p. 461

Allemang, D., J. Hendler, and F. Gandon, 2020, *Semantic Web for the Working Ontologist: Effective Modeling for Linked Data, Rdfs, and Owl*, New York, Association for Computing Machinery. doi:10.1145/3382097

Bagov, I., C. Greiner, and N. Garbedian, 2022, Collaborative Metadata Definition Using Controlled Vocabularies, and Ontologies, *Research Ideas and Outcomes,* 8:e94931, doi:10.3897/rio.8.e94931

Chan, L. M., and A. Salaba, 2015, Cataloging and Classification: An Introduction, Rowman & Littlefield Publishers

Chipangila, B., et al., 2024, Controlled Vocabularies in Digital Libraries: Challenges and Solutions for Increased Discoverability of Digital Objects, *International Journal on Digital Libraries*, vol. 25, issue 10. doi:10.1007/s00799-023-00374-1

Harping, P., 2010, *Introduction to Controlled Vocabularies: Terminologies for Art, Architecture, and Other Cultural Works*, Getty Publications

Heath, T., and C. Bizer, 2011, *Linked Data: Evolving the Web into a Global Data Space*, Springer Nature, https://link.springer.com/book/10.1007/978-3-031-79432-2

Hedden, H., 2010, Taxonomies and Controlled Vocabularies Best Practices for Metadata, *Journal of Digital Asset Management*, vol. 6, 279-284. doi:10.1057/dam.2010.29

Leise, F., K. Martes, and N. Badgett, 2008, *Indexing for Editors and Authors: A Practical Guide to Understanding Indexes*, Information Today Inc.

National Information Standards Organization, 2010, ANSI/NISO Z39.19-2005 (R2010). Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. doi:10.3789/ansi.niso.z39.19-2005R2010

Plante, R., et al., 2025, ed. M. Demleitner, *VOResource: An XML Encoding Schema for Resource Metadata*, Version 1.2, IVOA, https://www.ivoa.net/documents/VOResource/

Shobbrook, R., & R. Shobbrook, 1992, The I.A.U. Thesaurus for Improved On-Line Access to Information, *Proceedings of the Astronomical Society of Australia*, vol. 10, issue 2, p. 134

Wilkinson, M.D., et al., 2016, The FAIR Guiding Principles for Scientific Data Management and Stewardship, *Sci Data* 3, 160018. doi:10.1038/sdata.2016.18

Vrandečić, D., & Krötzsch, M., 2014, Wikidata: a free collaborative knowledgebase, *Communications of the ACM,* vol. 57, issue 10, p. 78. doi:10.1145/2629489

UAT Publications

The Unified Astronomy Thesaurus (UAT) Publication public library in ADS lists publications relating to the UAT.

Accomazzi, A., Gray, N., Erdmann, C., Biemesderfer, C., Frey, K., and Soles, J., 05/2014, The Unified Astronomy Thesaurus, *Astronomical Data Analysis Software and Systems XXIII*. Proceedings of a meeting held 29 September – 3 October 2013 at Waikoloa Beach Marriott, Hawaii, USA. Edited by N. Manset and P. Forshay ASP conference series, vol. 485, 2014, p.461

Accomazzi, Alberto, 01/2024, Decades of Transformation: Evolution of the NASA Astrophysics Data System's Infrastructure, arXiv:2401.09685

Accomazzi, Alberto, Bouquin, Daina, D'Abrusco, Raffaele, Derriere, Sébastien, Frey, Katie, Kern, Barbara, McCann, Graham, Novacescu, Jenny, Spoon, Henrik, Stahlman, Gretchen, Steffen, Julie, Weissman, Sarah, and Williams, Peter K. G., 04/2022, Building the UAT as a Community, *Bulletin of the AAS*, Volume 54, Issue 2, article id.2022n2i019

Allen, M. G., Dowler, P., Evans, J. D., Cui, C., Jenness, T., Merin, B., Berriman, G. B., and Kavelaars, J.J., 01/2020, The International Virtual Observatory Alliance in 2019, *Astronomical Data Analysis Software and Systems XXIX*. ASP Conference Series, Vol. 527, proceedings of a conference held (6–10 October 2019) at the Martini Plaza, Groningen, the Netherlands. Edited by Roberto Pizzo, Erik R. Deul, Jan David Mol, Jelle de Plaa, and Harro Verkouter. San Francisco: Astronomical Society of the Pacific, 2020, p.795

Biemesderfer, C., 08/2012, Astronomy Thesaurus and IAU WG Libraries, Working Group on Libraries, IAU Division XII Commission 5. Special Session 15: Data-Intensive Astronomy, held 23-24 August 2012 in Beijing, China. Chaired by R. Hanisch and M. Bishop. Edited by L. Utley, p.69

Bouquin, Daina, Carver, Nico, Frey, Katie, Williams, Allie, Damkalis, Alfredos-Panagiotis, Papadeas, Pierros, and Tsiligiannis, Vasilis, 04/2022, LSTN: Library Space Technology Network, *Bulletin of the AAS*, Volume 54, Issue 2, article id.2022n2i014

Bouquin, Daina, Frey, Katie, Henneken, Edwin, McEachern, Maria, McGrath, Alex, Guarracino, Daniel, Koch, Jennifer, Damon, James, Brownell, Eric, Zrull, Lindsay Smith, and Daina Bouquin, 01/2018, Project PHaEDRA: Preserving Harvard's Early Data and Research in Astronomy, American Astronomical Society, AAS Meeting #231, id. 143.03

Bouquin, Daina, Frey, Katie, McEachern, Maria, Damon, James, Guarracino, Daniel, McGrath, Alex, Henneken, Edwin, and Zrull, Lindsay Smith, 07/2018, Project PHaEDRA: Preserving Harvard's Early Data and Research in Astronomy, Library and Information Services in Astronomy VIII: "Astronomy Librarianship in the Era of Big Data and Open Science", Strasbourg, France, Edited by Lesteven, S.; Kern, B.; D'Abrusco, R.; Dorch, B.; EPJ Web of Conferences, Volume 186, id.07003

Chen, Tracy X., Schmitz, Marion, Mazzarella, Joseph M., Wu, Xiuqin, van Eyken, Julian C., Accomazzi, Alberto, Akeson, Rachel L., Allen, Mark, Beaton, Rachael, Berriman, G. Bruce, Boyle, Andrew W., Brouty, Marianne, Chan, Ben H. P., Christiansen, Jessie L., Ciardi, David R., Cook, David, D'Abrusco, Raffaele, Ebert, Rick, Frayer, Cren, Fulton, Benjamin J., Gelino, Christopher, Helou, George, Henderson, Calen B., Howell, Justin, Kim, Joyce, Landais, Gilles, Lo, Tak, Loup, Cécile, Madore, Barry, Monari, Giacomo, Muench, August, Oberto, Anaïs, Ocvirk, Pierre, Peek, Joshua E. G., Perret, Emmanuelle, Pevunova, Olga, Ramirez, Solange V., Rebull, Luisa, Shemmer, Ohad, Smale, Alan, Tam, Raymond, Terek, Scott, Van Orsow, Doug, Vannier, Patricia, and Wang, Shin-Ywan, 05/2022, Best Practices for Data Publication in the Astronomical Literature, *The Astrophysical Journal Supplement Series*, Volume 260, Issue 1, id.5, 13 pp.

Demleitner, M. and Frey, K., 07/2022, Adopting the UAT as an IVOA vocabulary Version 1.0, IVOA Endorsed Note 22 July 2022

Erdmann, Christopher and Frey, Katie, 08/2015, The Unified Astronomy Thesaurus, IAU General Assembly, Meeting #29, id.2254692

Frey, K., Erdmann, C., Accomazzi, A., Rubin, L., Biemesderfer, C., Gray, N., and Soles, J., 04/2015, Management of the Unified Astronomy Thesaurus, Open Science at the Frontiers of Librarianship. Proceedings of a conference held 17-20 June 2014, at Astronomical Observatory of Capodimonte, Naples, Italy 17-20. Edited by András Holl, Soizick Lesteven, Dianne Dietrich, and Antonella Gasperini. ASP Conference Series, Vol. 492. San Francisco: Astronomical Society of the Pacific, 2015, p.204

Frey, Katie and Accomazzi, Alberto, 05/2018, The Unified Astronomy Thesaurus: Semantic Metadata for Astronomy and Astrophysics, *The Astrophysical Journal Supplement Series*, Volume 236, Issue 1, article id. 24, 7 pp. (2018).

Frey, Katie and Bukovi, Kris, 04/2020, Unified Astronomy Thesaurus: Usage Stats and Release Cycle, *Research Notes of the AAS*, Volume 4, Issue 4, id.58

Frey, Katie, 04/2022, Implementation of The Unified Astronomy Thesaurus, *Bulletin of the AAS*, Volume 54, Issue 2, article id.2022n2i018

Frey, Katie, Erdmann, Chris, Accomazzi, Alberto, Rubin, Louise, Biemesderfer, Chris, and Gray, Norman, 06/2014, Classifying Infinity… and Beyond!, Posters from the 7th Library and Information Services in Astronomy (LISA VII), held 17–20 June 2014, in Naples, Italy, id. 11

Frey, Katie, Weissman, Sarah, Kern, Barbara, Lagerstrom, Jill, Peek, Josh, and Accomazzi, Alberto, 07/2018, Managing Contributions to the Unified Astronomy Thesaurus, Library and Information Services in Astronomy VIII: "Astronomy Librarianship in the era of Big Data and Open Science", Strasbourg, France, Edited by Lesteven, S.; Kern, B.; D'Abrusco, R.; Dorch, B.; EPJ Web of Conferences, Volume 186, id.09003

Graham, Matthew J., Djorgovski, S. G., Donalek, Ciro, Drake, Andrew J., Mahabal, Ashish A., Plante, Raymond L., Kantor, Jeffrey, and Good, John C., 09/2012, Connecting the Time Domain Community with the Virtual Astronomical Observatory, Observatory Operations: Strategies, Processes, and Systems IV. Proceedings of the SPIE, Volume 8448, article id. 84480P, 8 pp. (2012).

Gray, Norman, Erdmann, C., Accomazzi, A., Soles, J., McCann, G., Cassar, M., and Biemesderfer, C., 01/2013, Creation and Maintenance of a Unified Astronomy Thesaurus, American Astronomical Society, AAS Meeting #221, id.240.28

Heidorn, P. Bryan, Stahlman, Gretchen R., and Steffen, Julie, 05/2018, Astrolabe: Curating, Linking, and Computing Astronomy's Dark Data, *The Astrophysical Journal Supplement Series*, Volume 236, Issue 1, article id. 3, 12 pp. (2018).

Kurtz, Michael, Accomazzi, Alberto, and Henneken, Edwin, 05/2021, Enabling Synergy: Improving the Information Infrastructure for Planetary Science, Planetary Science and Astrobiology Decadal Survey 2023-2032 white paper e-id. 470; *Bulletin of the AAS*, Vol. 53, Issue 4, e-id. 470 (2021)

Lockhart, Kelly, Blanco-Cuaresma, Sergi, Accomazzi, Alberto, Kurtz, Michael, Henneken, Edwin, Shapurian, Golnaz, Grezes, Felix, Allen, Thomas, and NASA ADS team, 02/2024, Utilizing Artificial Intelligence Techniques for Information Retrieval and Data Enrichment in NASA ADS, American Astronomical Society, AAS Meeting #243, id. 111.07.

Molinaro, Marco, Allen, Mark, Bonnarel, François, Genova, Françoise, Demleitner, Markus, Graf, Kay, Morris, Dave, Solano, Enrique, and Schaaff, André, 11/2021, Supporting FAIR Principles in the Astrophysics Community: The European Experience, arXiv:2111.14468

Raugh, Anne, Hughes, John S., Guinness, Edward A., and Gordon, Mitchell K., 12/2022, First Steps in Developing an Ontology of Targeting for the Planetary Data System (PDS) Archives, AGU Fall Meeting 2022, held in Chicago, IL, 12-16 December 2022, id. IN22D-0330.

Stahlman, Gretchen, Bryan Heidorn, P., and Steffen, Julie, 07/2018, The Astrolabe Project: Identifying and Curating Astronomical 'Dark Data' through Development of Cyberinfrastructure Resources, Library and Information Services in Astronomy VIII: "Astronomy Librarianship in the era of Big Data and Open Science", Strasbourg, France, Edited by Lesteven, S.; Kern, B.; D'Abrusco, R.; Dorch, B.; EPJ Web of Conferences, Volume 186, id.03003

Teuben, Peter, Allen, Alice, and Berriman, G. Bruce, 07/2022, How to Better Describe Software for Discovery and Citation, Astronomical Data Analysis Software and Systems XXX. ASP Conference Series, Vol. 532, Proceedings of a virtual conference held 8–12 November 2020. Edited by Jose Enrique Ruiz, Francesco Pierfedereci, and Peter Teuben. San Francisco: Astronomical Society of the Pacific, 2022, p.627