



Future of the Unified Astronomy Thesaurus

Jennifer Lynn Bartlett,
Felix Grezes, Daniel Chivvis,
and the ADS Team

jennifer.bartlett@cfa.harvard.edu

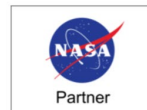
felix.grezes@cfa.harvard.edu

daniel.chivvis@cfa.harvard.edu

ADS Users Group Meeting, 05-06 Dec. 2024



CENTER FOR
ASTROPHYSICS
HARVARD & SMITHSONIAN



This work is openly licensed via [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/).



UNIFIED ASTRONOMY THESAURUS

a community supported, open source project from the American Astronomical Society

- **Simple Knowledge Organization System (SKOS) ontology for astronomy**
 - Owned and freely licensed by the American Astronomical Society (AAS)
 - Open, community-contributed resource
 - Primarily astronomy concepts, some planetary science & heliophysics content
- **Keyword system for**
 - AAS journals
 - *Publications of the Astronomical Society of the Pacific* (PASP)
 - *Icarus*
 - International Virtual Observatory Alliance (IVOA)
 - Space Telescope Science Institute (STScI)



<https://astrothesaurus.org/>

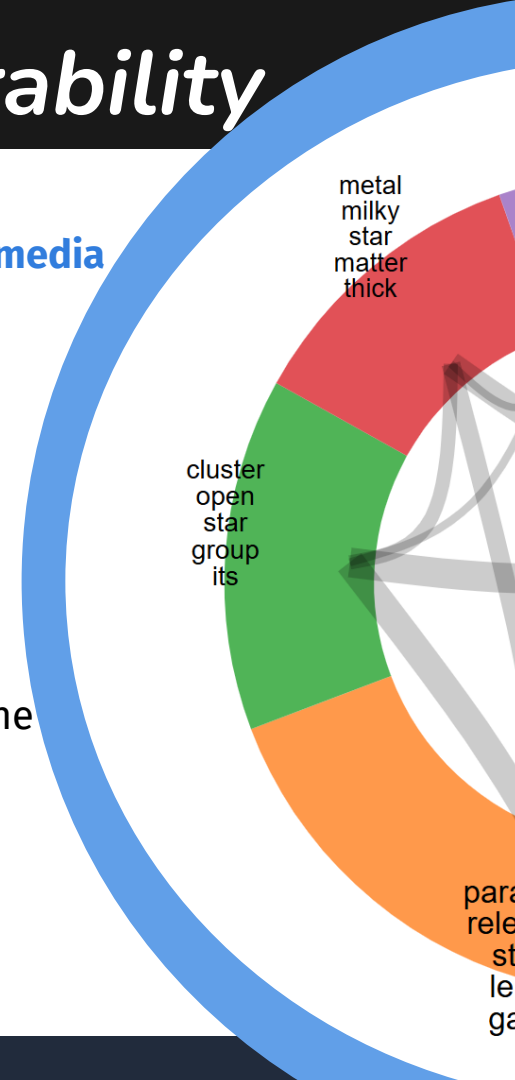


<https://github.com/astrothesaurus/UAT>



UAT for Enhanced Discoverability

- **Potential to be single keyword system for space sciences regardless of media**
 - Consistent, controlled vocabulary for searching
 - Describes literature, datasets, and software
 - Relates concepts to one another in hierarchy for exploration
- **Potential to connect to other ontologies for cross-disciplinary search**
- **ADS has promoted UAT use since 2013**
 - Goal: assign UAT keywords to all astronomy articles, including Older pre-UAT articles, articles in journals that have not adopted the system, and articles whose authors are parsimonious
 - Supplement, not supplant, existing keywords
 - Keyword AI Labeler At Scix (KAILAS) is developmental AI algorithm

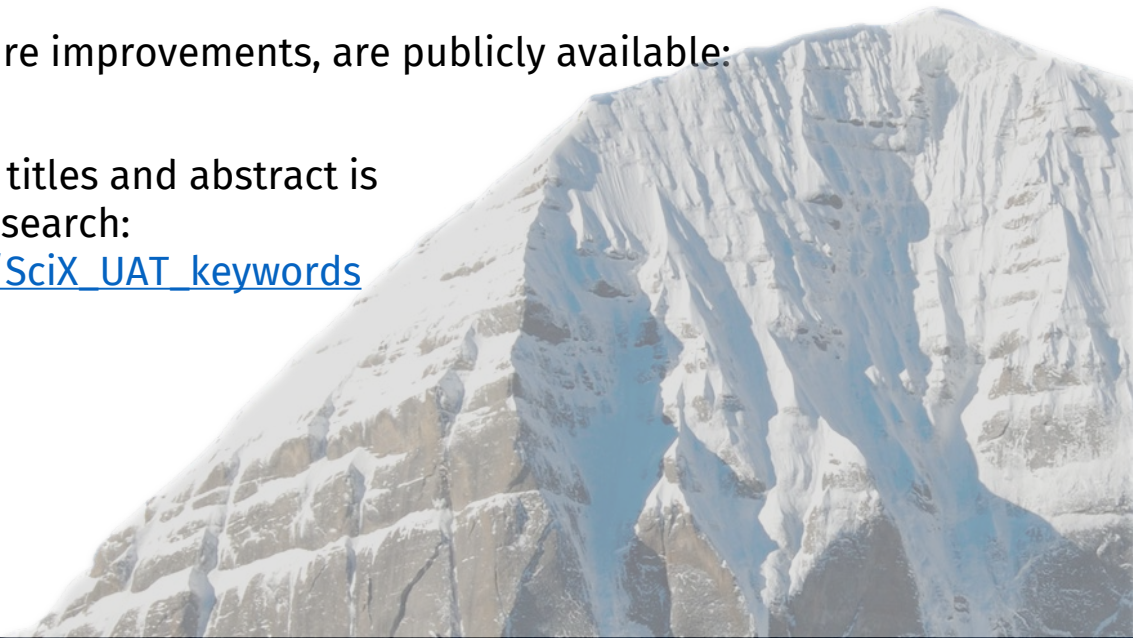


KAILAS: Keyword AI Labeler At Scix

We are building KAILAS to automatically assign UAT keywords to astronomy papers.

- KAILAS is a RoBERTa-based model trained on the text of ~18K astronomy papers in our database that have been tagged by authors/publishers with UAT keywords
- The trained model, as well as future improvements, are publicly available: huggingface.co/adsabs/KAILAS
- A version of the dataset with only titles and abstract is publicly available for academic research: huggingface.co/datasets/adsabs/SciX_UAT_keywords

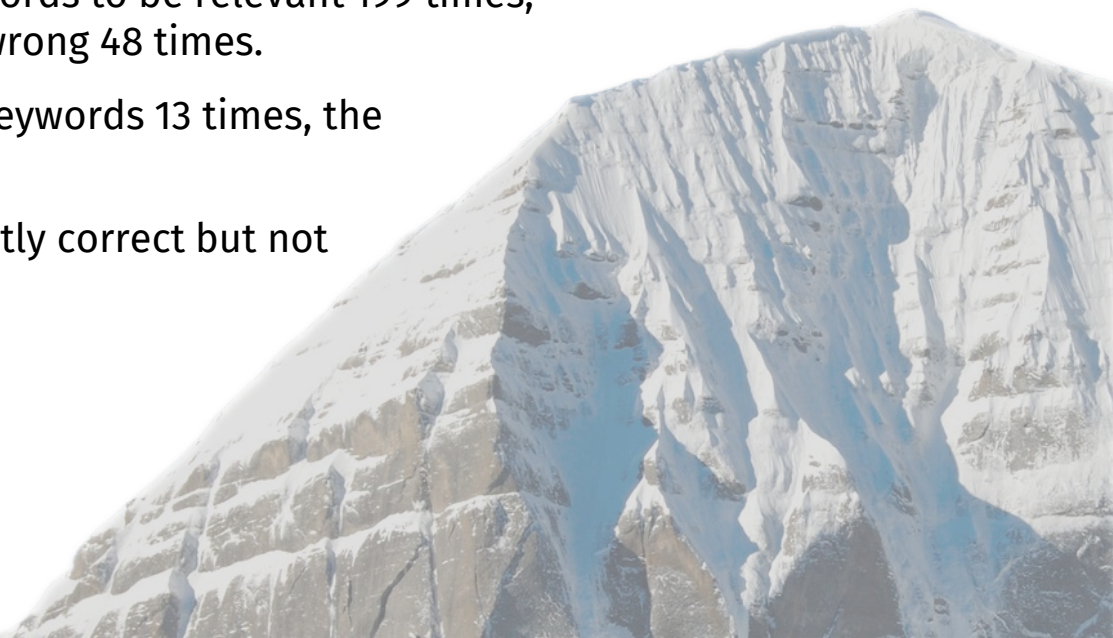
Mount Kailas from the Himalayas



KAILAS: Keyword AI Labeler At Scix

Early feedback from the Working Group

- total of 50 papers + 1 control reviewed by 4 experts
- The reviewers found KAILAS's keywords to be relevant 199 times, potentially relevant 34 times and wrong 48 times.
- The reviewers preferred KAILAS's keywords 13 times, the authors' 21 times with 20 ties.
- keywords assigned are varied, mostly correct but not better than human yet.





KAILAS: Keyword AI Labeler At Scix

Input

TITLE: Ultrafast Variability in AGN Jets: Intermittency and Lighthouse Effect
ABSTRACT: Gamma-ray flares from active galactic nuclei (AGNs) show substantial variability on ultrafast timescales (i.e., shorter than the light-crossing time of the AGN's supermassive black hole). We propose that ultrafast variability is a by-product of the turbulent dissipation of the jet Poynting flux. Due to the intermittency of the turbulent cascade, the dissipation is concentrated in a set of reconnecting current sheets. Electrons energized by reconnection have a strong pitch-angle anisotropy, i.e., their velocity is [...]

Output

KAILAS ASSIGNED KEYWORDS:
blazars
active galactic nuclei
jets
relativistic jets
plasma astrophysics
high energy astrophysics
magnetic fields

Compare

AUTHOR ASSIGNED KEYWORDS:
gamma-rays
blazars
galaxy jets



UAT at Crossroads

Wolbach Library closing in March eliminated part-time UAT curator

- previously supported by CfA
- ADS absorbed some technical support responsibilities
- AAS reorganized management of project
- ADS & AAS funded temporary curation contract for knowledge transfer



Working Group on UAT

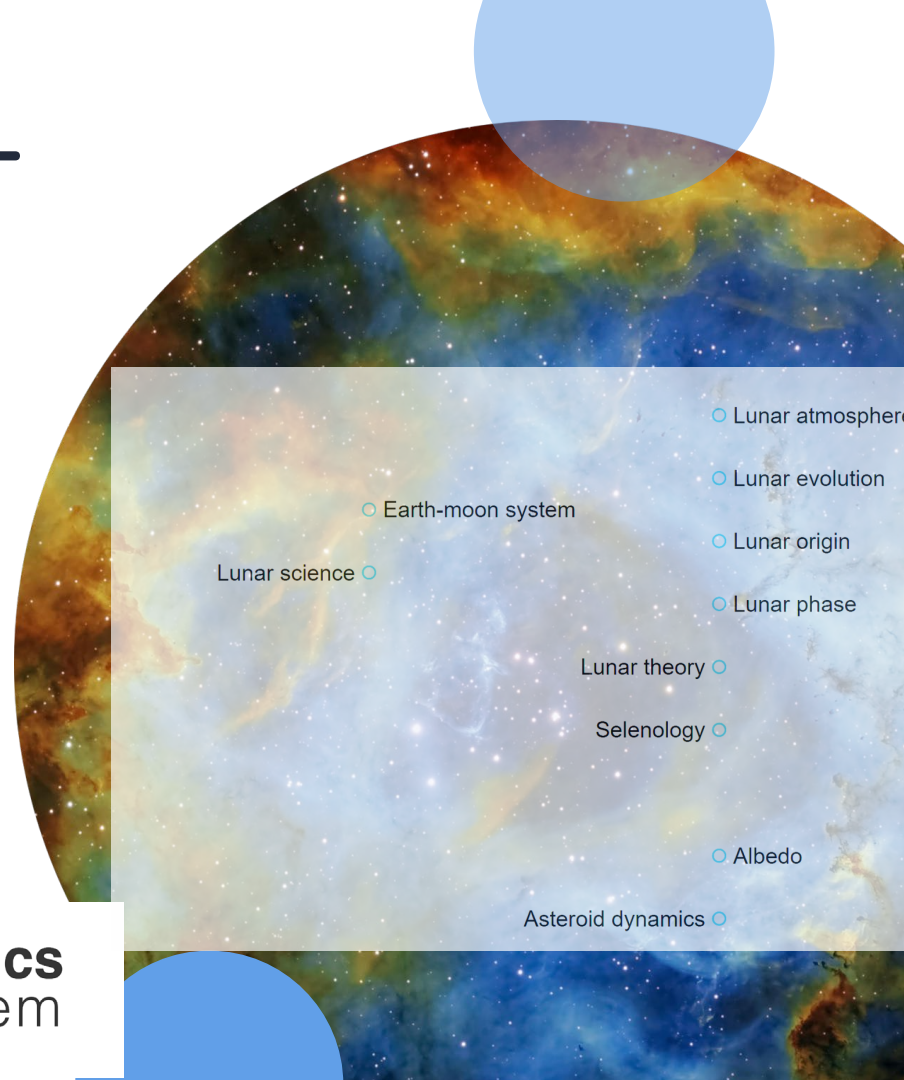
5 years to initiate strategic plan for thesaurus

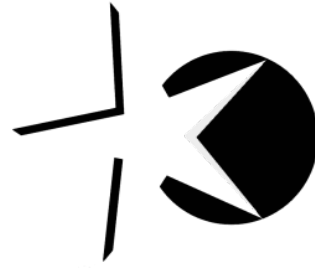
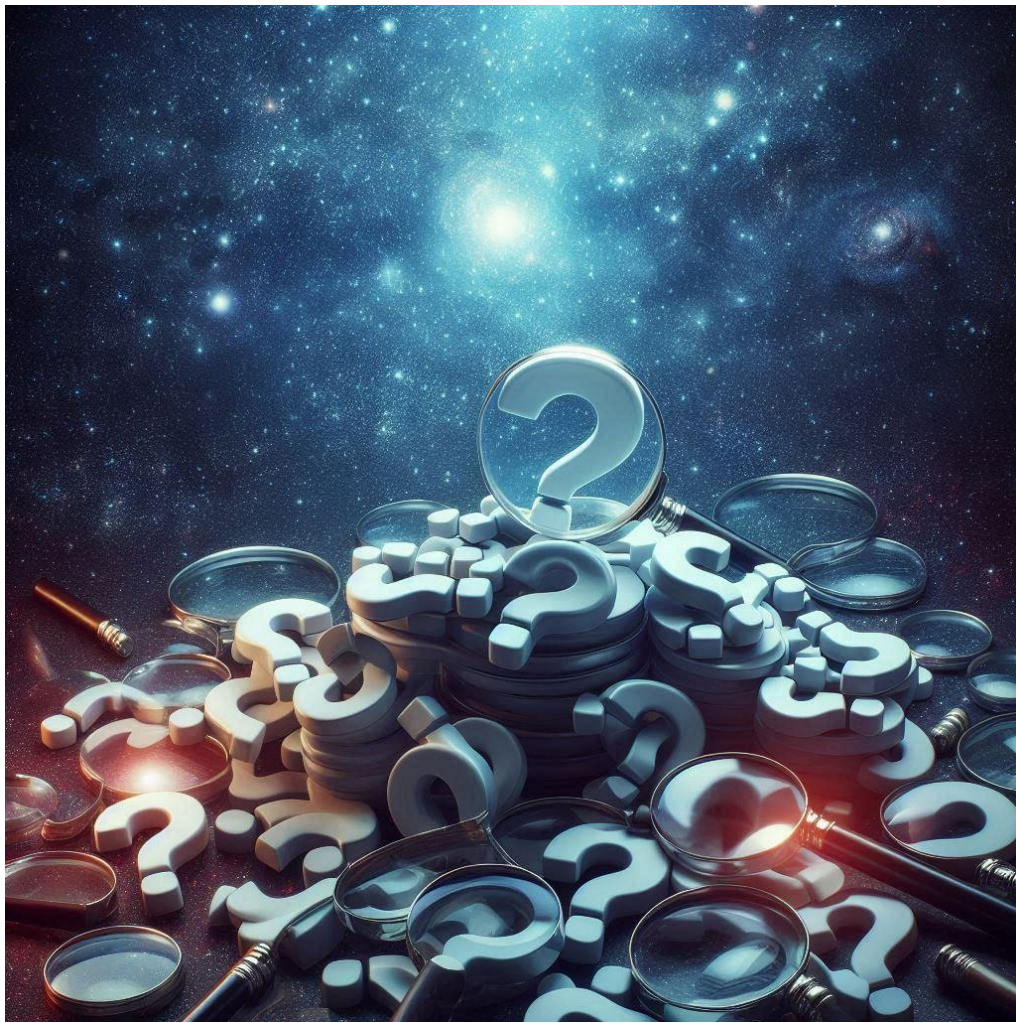
- Released v. 5.1 over summer
- Planning expanded coverage for planetary science & heliophysics
- Investigating alternative curation models
- Modernizing technology infrastructure
- Seeking additional funding sources

Daniel Chivvis, member
Jennifer Lynn Bartlett, vice chair



astrophysics
data system





Questions?

