Unified Astronomy Thesaurus

C. Erdmann, K. Frey & Team
IAU 2015: Community Integration - Joining Forces
August 6, 2015
History and Background

- **IAU Thesaurus:**
  - Last updated 1995, missing more current terms
  - No community update mechanism

- **Astronomical Subject Keywords:**
  - Established mechanism to keep vocabulary updated

- **Physics and Astronomy Classification Scheme (PACS):**
  - Retired by AIP in 2010
  - Astronomy terms added from literature

- IoP and AIP contracted with Access Innovations to update their projects separately, but asked AI to merge their keyword lists into a unified thesaurus.
The Gardener & Community

- Gardener
- Community (feedback)
- Reviewers/Editors
- Stewardship and Outreach Groups
- ADS (platform)

Governance is evolving, see http://astrothesaurus.org/
AAS Custodianship
2014ASPC..485..461 / arXiv:1403.6656
Vocabulary Overlaps

Unified Astronomy Thesaurus: 1915
Physics & Astronomy Classification: ~3500
Astronomical Subject Keywords: ~380
IAU Thesaurus: ~2900
SPIE Keywords: 94
UAT Information

- Current statistics:
  - 1915 concepts (plus 5 deprecated terms)
  - 15 top level concepts
  - 12 level depth, most concepts are in the 4th level

- Highly structured, polyhierarchy
- Tentative update cycle: every six months
Platform: VocBench

- Open source
- SKOS standards
- Access via browser
- Custom relationships
- User roles
- Workflow process
- Provenance (term history)
- User community
- Backed by FAO (UN) & the University of Rome

[Image of the VocBench platform]

http://astrothesaurus.org/vocbench/
**Review and Release Process**

**Suggestion Workflow** (constantly occurring process for every new suggestion)

1. **Propose** (anyone)
2. **Assign** (gardener)
3. **Evaluate** (editors)
4. **Review** (gardener)
5. **Validate** (gardener)

   - **discrepancies**
     - yes: resolve discrepancies (gardener & editors)
     - no: proceed to next step

**Release Workflow** (occurs when preparing for a new version release)

1. **Publish Changes**
2. **Release New Version**
3. **Push Updates**
4. **Archive Previous Version**
Use Cases

- UAT concept indexing of data, software, articles, conferences, & proposals
  - Zenodo, GitHub, ADS, CodeMeta, Chandra proposals
- Connections to other resources: SIMBAD, WikiData, MeSH, DBpedia
- Relevant article discovery for early career / post-grads
- Concept definitions (e.g. expand UAT with from Paris Observatory & Wordnik)
- Parallel database for telescope taxonomies (instruments, locations, etc)
- Bibliometrics beyond the word cloud…
  - concept trend lines (e.g. Culturomic’s ArXiv n-grams or Google n-grams)
  - content analysis and comparison
UAT APIs & Tools

- OpenRDF Sesame & Workbench: http://astrothesaurus.org/openrdf-workbench/
- Autocomplete widget: http://astrothesaurus.org/tools/autocomplete-widget/
- Visual Editor: http://www.altbibl.io/astrothesaurus/sort/
- Unified Astronomy Thesaurus API: http://uat.altbibl.io/
Workshop on Shared Curation of the Astronomical Literature

In this 1.5 day workshop (BY INVITATION ONLY), from June 12 – 13, we will explore current curation practices for telescope bibliographies and discuss how the community can work together towards improved workflows. We will look at new digital curation techniques and tools, particularly the Unified Astronomy Thesaurus (UAT), and discuss next steps ahead of the IAU General Assembly.
Next Steps

- Organize the UAT
- Identify collaborators
- Refine tools & techniques
- Test & scrub
- AAS/UAT releasefest

Collaborators: Jill Lagerstrom, Sarah Weissman, Barbara Kern, Katie Frey, Chris Erdmann, Jane Holmquist, Dianne Dietrich, Kayleigh Bohemier, Sharon Hunt, Josh Peek
Thank you!