Current Status of the Unified Astronomy Thesaurus

Alberto Accomazzi
NASA Astrophysics Data System
Harvard-Smithsonian Center for Astrophysics

15 May 2013
IVOA Interop, Heidelberg

Plan

- Merge existing divergent thesauri into an open, interoperable thesaurus covering all areas of A&A
- Work out legal issues related to copyrights, re-use
- Develop a stable maintenance process that maximizes UAT adoption and use
- Allow for input from community to be incorporated in a timely and transparent way
- Provide well-curated, periodic releases of the UAT
- Seek IAU blessing of Thesaurus

Accomplished So Far

- The American Institute of Physics (AIP), the Institute of Physics (IOP), ADS (AA), IVOA (NG) and AAS (CB) have been discussing a Unified Thesaurus since 2011
- IOP developed a thesaurus covering its literature, incorporated much of IVOAT in the process
- AIP and IOP merged and donated the astronomy parts of their thesauri to AAS which in turns released it under CC-BY
- The CfA library in consultation with ADS is currently managing the curation and dissemination of the thesaurus
- Website and beta thesaurus available: http://astrothesaurus.org

Next Steps

- Expand stewardship group which now includes IVOA, AAS, ADS, CfA Library, with input from IOP and AIP. We expect further participation from RAS, CDS, Paris Observatory, ESO, others?
- We are seeking astronomers to act as editors for branches of the thesaurus. Volunteers?
- Input from community to be submitted through a web-based portal, then vetted by editors. Still looking for a platform to facilitate this workflow, any suggestions/volunteers?
- Revisions and updates regularly incorporated (by UAT curator) via a formal release mechanism. Suggestions for a schedule?

UAT Lineage

- The current draft version of the UAT is essentially the astronomy portion of the IOP thesaurus enriched with terms from IVOAT
- What has been kept are *astronomy* concepts that appear in the astronomy journals published by IOP, i.e. *ApJ* and *AJ*
- Additionally, restructuring was done to group concepts, "balance" the thesaurus and facilitate indexing/retrieval
- This means that some sections of the thesaurus were not retained, others were moved or heavily modified

What we have now

- UAT:
  - Total number of concepts: 1907
  - Top-level concepts: 15
  - Total number of paths: 3552
- IVOAT:
  - Total number of concepts: 2889
  - Top-level concepts: 274
  - Total number of paths: 5202

Examples: matching terms

- Chondrules in IVOAT:
  Meteor
    Meteorite
      Chondrite
        Chondrules

- Chondrules in UAT:
  Astronomical objects
    Solar system
      Meteoroids
        Meteorites
          Chondrites
            Chondrules

- Earth in IVOAT:
  Rotating body
    Planet
      Earth

- Earth in UAT:
  Astronomical objects
    Star systems
      Single star systems
        Solar system
          Solar system planets
            Inner planets
              Earth (planet)

Examples: unique terms

In UAT but not IVOAT:

String Theory

Astroparticle physics

Black hole physics

Trojan asteroids

Hot Jupiters

Radio jets

Planetary science

In IVOAT but not UAT:

Atom

Camera

CCD detector

Cosmic

Energy

Life

Television

All the Details

- Katie Frey (CfA library, UAT curator) spent some time reconciling UAT and IVOAT
- A hierarchical list of terms in UAT and IVOAT: https://docs.google.com/spreadsheet/ccc?key=0Aidbx2pXIZLGdFN0QTlKMlV6ZGdCRkdXSnpITnNEcVE&usp=sharing
- Color-coded comparison of UAT/IVOAT terms: https://docs.google.com/spreadsheet/ccc?key=0Aidbx2pXIZLGdG5ibEhGZ3AyUlZjTjUyS3dlMzhOenc&usp=sharing

Going forward

- Guiding principle: the thesaurus should contain astronomical concepts found in the literature
- Depending on how we define "literature" things may be added or even taken away. Some options:
  - The core journals: ApJ, AJ, A&A, MNRAS (~ 240K)
  - The extended core set: add PASP, PhRvD, PASJ, PhRvL, Nature, Science (~ 330K)
  - The refereed literature in ADS astronomy db (~ 1M)
  - The entire ADS astronomy db (~ 2M)

Useful links

- Browse the UAT:
  http://astrothesaurus.org/thesaurus/
  http://astrothesaurus.org/alphabetical-browse/

- Browse the IVOAT:
  http://www.astro.physik.uni-goettingen.de/~hessman/rdf/IVOAT/

- The Paris Observatory Dictionary:
  http://dictionary.obspm.fr/
  http://astroconcepts.obspm.fr/