

AstroTag

MUG meeting, STScI

December 2014

Data Tagging

- Storing associations between data sets and tags (words/phrases)
 - IPPSSOOT $\Leftrightarrow \{w_1, w_2, \dots, w_n\}$
- Tags have meaning
 - Identify types of objects within a data set
 - E.g. Dwarf irregular galaxies
 - Identify data related to a particular area of study
 - E.g. Galaxy interactions, galaxy formation

Tagging from a Thesaurus

- Tagging with concepts
 - Multiple labels in multiple languages can be used to express the same concept
 - Concept labels can change without changing concept/data set relations (& v.v.)
- Relationships between concepts add additional meaning
 - Broader term (BT), narrower term (NT), related term (RT), others

[[English](#) | [French](#) | [German](#) | [Italian](#) | [Spanish](#)]

[[GIANT STARS](#) | [ETOILE GEANTE](#) | [RIESENSTERNE](#) | [STELLE GIGANTI](#) | [ESTRELLAS GIGANTES](#)]

GIANT STARS

UF

[cs stars](#)

BT

[POSTMAINSEQUENCE STARS](#)

NT

[A GIANT STARS](#)

[B GIANT STARS](#)

[CH STARS](#)

[CN STARS](#)

[EXTREME CARBON STARS](#)

[F GIANT STARS](#)

[G GIANT STARS](#)

[LATE TYPE GIANT STARS](#)

[Lb VARIABLE STARS](#)

[M GIANT STARS](#)

[OH IR STARS](#)

[PECULIAR GIANT STARS](#)

[R STARS](#)

[S STARS](#)

[WEAK G BAND STARS](#)

RT

[CARBON BURNING](#)

[DREDGE UP](#)

[LITHIUM STARS](#)

[LUMINOSITY CLASSIFICATION](#)

[STELLAR EVOLUTION](#)

[ZIRCONIUM STARS](#)

<http://www.mso.anu.edu.au/library/thesaurus/english/GIANTSTARS.html>

Searching: **PsycINFO** | [Choose Databases](#)

Demonstration Customer

[Basic Search](#) | [Advanced Search](#) | [Search History](#)

Browsing: PsycINFO -- Thesaurus

 Term Begins With Term Contains Relevancy RankedPage: [◀ Previous](#) | [Next ▶](#)Select term, then add to search using: Explode Major Concept

(Click term to display details.)

<input type="checkbox"/>	Abandonment	<input type="checkbox"/>
<input type="checkbox"/>	Abdomen	<input type="checkbox"/>
<input type="checkbox"/>	Abdominal Wall	<input type="checkbox"/>
<input type="checkbox"/>	Abducens Nerve	<input type="checkbox"/>
<input type="checkbox"/>	Ability	<input checked="" type="checkbox"/>
<input type="checkbox"/>	Ability Grouping	<input type="checkbox"/>
<input type="checkbox"/>	Ability Level	<input type="checkbox"/>

[Ability Tests](#) Use [Aptitude Measures](#)

Generating Associations

- Concept labels <-> Papers <-> Datasets
 - Automatic matching of labels with papers/proposals
 - Manual association of papers with data sets
- Catalogs
 - Type, morphology, etc.
- User input (APT, etc.)

Thesaurus-enabled Features

- Hierarchical browsing
- Search (with browsing)
- Result Filtering
- Breadcrumbs
- Tag clouds
- * Customized data delivery
- * Topic ranking for data sets through citation mining.



Shop by Department

Search

Astronomy

astronomy textbooks

Books Advanced Search New Releases Best Sellers The New York Times® Best Sellers Children's Books Textbooks Textbook Rentals Sell Us Your B

1-12 of 2,006 results for Books : Science & Math : Astronomy & Space Science : Astronomy : "astronomy textbooks"

Show results for

- < Any Category
- < Books
- < Science & Math
- < Astronomy & Space Science

Astronomy

Refine by

Amazon Prime

Prime Eligible

Delivery Day

Get It Today

Get It by Tomorrow

New Releases

Last 30 days

Last 90 days

Coming Soon

Author

Eric Chaisson

Book Format

Hardcover

Paperback

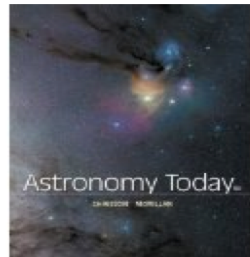
Kindle Edition



Save up to 50% on Popular Science Books

Until December 14, enjoy savings on popular science and astronomy titles

Book Format: Hardcover | Paperback | Kindle Edition



Astronomy Today (8th Edition) Sep 19, 2013

by Eric Chaisson and Steve McMillan

Hardcover

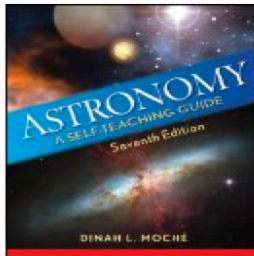
\$15.98 - \$68.57 to rent Prime

\$136.29 to buy Prime

Usually ships in 3 to 4 weeks

More Buying Choices

\$91.18 used & new (54 offers)



Astronomy: A Self-Teaching Guide, Seventh Edition (Wiley Self-Teaching

by Dinah L. Moche

Paperback

\$13.20 \$19.95 Prime

Get it by **Tuesday, Nov 25**

More Buying Choices

\$3.99 used & new (76 offers)

Kindle Edition

\$9.49

Auto-delivered wirelessly

See newer edition of this book

Choosing a Tag Set

- Shape of tree:
 - Top-level terms
 - Depth
 - Poly-hierarchy
- Are the right concepts with the right labels available?
 - Astroparticle physics vs. Particle astrophysics
- What level of the concept tree do we tag at?

The UAT






















- Unified Astronomy Thesaurus
(astrothesaurus.org)
- Community authored/edited thesaurus, was maintained by CfA, but is now maintained by AAS
- Combines IVOAT, PACS, journal keywords
- Creative Commons Licensed

UAT Continued

- Pros
 - 15 top-level terms (easy to browse)
 - Community buy-in
 - Web standards (SKOS/RDF)
- Caveats
 - Not necessarily designed with tagging and search in mind
 - Sometimes consensus means slow to change
 - Combining RDF and relational data can be tricky

Thesaurus Evaluation

	UAT	IVOAT
# concepts/labels	1909 / 3017	2890 / 3531
Max label length	8 “unidentified sources of radiation outside the solar system”	6 “low mass x-ray binary star”
Found papers/progs	99.1% / 87.3%	~100% / 99.2%
Not found papers/progs	458 / 1200	4 / 74
Not found concepts/labels	433 (22.7%) / 969 (32.1%)	548 (20%) / 745 (21.1%)

-  **Cosmology (2346)**
 -  **Narrower Terms**
 - +  **Astronomical radiation sources (0)**
 - +  **Cosmological phase transitions (0)**
 - +  **Curved space (0)**
 - +  **Bianchi cosmology (0)**
 - +  **Anthropic principle (0)**
 - +  **Plasma astrophysics (1)**
 -  **Cosmogony (10)**
 -  **Narrower Terms**
 -  **Related Terms**
 -  **Solar system (854)**
 - +  **Hierarchical cosmology (3)**
 - +  **Radiative transfer (1019)**
 - +  **Large-scale structure of the universe (15)**
 - +  **Non-Gaussianity (13)**
 -  **Redshift (9424)**
 -  **Narrower Terms**
 -  **Related Terms**
 -  **Redshift surveys (353)**
 -  **Cosmic background radiation (63)**

Next Steps

- Expanding vocabulary for search
 - Focusing on missing labels, not structure
 - How best to share data, collaborate/merge with UAT
- Tagging and search model

Open Questions

- Accuracy
 - Human / machine error will create bad associations. How harmful is a misidentification of a type X object?
- Completeness
 - Not everything will be tagged. How useful is a partial list of type X objects?
- Provenance
 - How best to provide the source of tags for a data set?
 - E.g. The list of papers containing the tags / and data set associations.

Thanks!

- sweissman@stsci.edu