



University
of Glasgow

Vocabularies in the VO

Alasdair J G Gray

Norman Gray

Iadh Ounis





What is a Vocabulary?

Dictionary

vo•cab•u•lar•y |və(u)'kabjʊləri|

noun (pl. **-lar•ies**)

the body of words used in a particular language.

- a part of such a body of words used on a particular occasion or in a particular sphere : *the vocabulary of law* | *the term became part of business vocabulary*.
- the body of words known to an individual person : *he had a wide vocabulary*.
- a list of difficult or unfamiliar words with an explanation of their meanings, accompanying a piece of specialist or foreign-language text.
- a range of artistic or stylistic forms, techniques, or movements : *dance companies have their own vocabularies of movement*.

ORIGIN mid 16th cent. (denoting a list of words with definitions or translations): from medieval Latin **vocabularius**, from Latin **vocabulum** (see VOCABLE).

Thesaurus

vocabulary

noun

1 *technical vocabulary* LANGUAGE, lexicon, lexis, words; diction, terminology, phraseology, nomenclature, terms, expressions, parlance, idiom, jargon, vernacular, argot, cant; informal vocab, lingo, -speak, -ese.

2 *she is improving her vocabulary* WORD POWER, lexicon, command of language; informal vocab.



A set of terms with:

- Label
- Synonyms
- Definition
- Relationships to other terms:
 - Broader term
 - Narrower term
 - Related term

Example:

- Spiral galaxy
- Spiral nebula
- A galaxy having a spiral structure
- Relationships:
 - BT: Galaxy
 - NT: Barred spiral galaxy
 - RT: Spiral arm



Why do we need vocabularies?

- Limit and define terminology
- Improve search
 - Retrieve papers from ADS using keywords
 - All papers about “*Binary eclipsing stars*”
 - Process all VOEvents about “supernovae”
 - Should also process “1a supernovae”
 - Locating relevant resources in the registry/VizieR
 - Removes ambiguity
 - Improves precision



Problems:

- Plural/singular
- Case
- Abbreviations
- Different tags

Thanks to Sébastien Derriere
for this data.

75 Star	5 Pulsar	2 Interstellar medium
52 Galaxy	4 supernova	2 QSO
37 Stars	3 Clusters of Galaxies	2 QSOs
36 Galaxies	3 Infrared:stars	2 SNR
16 AGN	3 Quasars: general	2 Variable Star
12 Cluster of Galaxies	3 Supernova	2 White Dwarf
12 Nebulae	3 White dwarfs	2 clusters of galaxies
11 Planets	3 galaxies	2 stars
10 GRB	2 Comets	1 Asteroids
10 Globular Clusters	2 Cool stars	1 BL Lac
8 Star Cluster	2 Extragalactic Source	1 Be/X-ray binary stars
7 Nebula	2 Extragalactic objects	1 Binary stars
6 Variable stars	2 Infrared: stars	...
5 Hot stars		



Requirements:

- Provide term identifiers
 - Unambiguous tagging
- Capture semantic relationships
 - Poly-hierarchy structure
- Machine processable
 - Allows inter-operability
 - “Machine intelligence”
- Avoids problems of:
 - Spelling
 - Case
 - Plurality problems
 - Tags
- Automated reasoning:
 - Interested in all “Supernova”
 - Items tagged as “1a Supernova” also returned



- W3C standard for sharing vocabularies
- Based on RDF
 - Semantic model for describing resources
 - Provides URI for each term
 - Captures properties of terms
 - Encodes relationships between terms
 - Enables automated reasoning
 - Standard serialisations
 - “Looser” semantics than OWL
- Adopted by IVOA Vocabularies standard



Example SKOS Vocabulary Term

Example

“Spiral galaxy”

“Spiral nebula”

“A galaxy having a spiral structure”

Relationships:

BT: “Galaxy”

NT: “Barred spiral galaxy”

RT: “Spiral arm”

```
#spiralGalaxy a concept;  
prefLabel "Spiral galaxy"@en;  
altLabel "Spiral nebula"@en;  
definition "A galaxy having a  
spiral structure"@en;  
broader #galaxy;  
narrower #barredSpiralGalaxy;  
related #spiralArm .
```

In turtle notation



- **Folksonomies**
 - Keyword tags, freely chosen – e.g. VizieR subjects
- **Vocabulary**
 - Controlled list of words with definitions
- **Taxonomy**
 - Relationships: Broader/Narrower/Related
- **Thesaurus**
 - Synonyms, antonyms, see also
- **Ontology**
 - Formal specification of a shared conceptualisation – OWL

**“Vocabulary” used in IVOA to cover
vocabularies, taxonomies, and thesauri.**



- **Journal Keywords**
 - Developed for tagging papers
 - 311 terms
 - Actively used
- **Astronomy Visualization Metadata (AVM)**
 - Tagging images
 - 217 terms
 - Actively used
- **IAU Thesaurus**
 - Developed for libraries in 1993
 - 2,551 terms
 - Never really used
- **Unified Content Descriptor (UCD)**
 - Tagging resource data
 - 473 terms
 - Actively used

Published as SKOS vocabularies by the IVOA



Which vocabulary should I use?

- Closest match to your needs
- Relate vocabulary terms using mappings
 - Part of the SKOS standard
 - One mapping file per pair of vocabularies
 - Chains of mappings
 - A → B → C
 - A ⇒ C

Inter-vocabulary mappings

- **Broad match:** more general term
- **Narrow match:** more specific term
- **Related match:** associated term
- **Exact match:** equivalent term
- **Close match:** similar but not equivalent term



Mapping Editor

SKOS Mapping Editor

File

Vocabulary for Astronomy & Astrophysics Journal keywords (Version WD-1.11)

- ▶ Vocabulary for Astronomy & Astrophysics Journal keywords (Version WD-1.11)
 - ▶ Astrometry and celestial mechanics
 - ▶ Astronomical data bases
 - ▶ Astronomical instrumentation, methods and techniques
 - ▶ Cosmology
 - ▶ Galaxies
 - ▶ Interstellar medium, nebulae
 - ▶ Physical data and processes
 - ▶ Solar system
 - ▶ Comets: general
 - ▶ Earth
 - ▶ Interplanetary medium
 - ▶ Kuiper Belt
 - ▶ Meteors, meteoroids
 - ▶ Minor planets, asteroids
 - ▶ Moon
 - ▶ Oort Cloud

<http://www.ivoa.net/rdf/Vocabularies/AVM>

- ▶ 1. Planet
- ▶ 2. Interplanetary Body
 - ▶ 2.1. Dwarf planet
 - ▶ 2.2. Comet
 - ▶ 2.2.1. Nucleus
 - ▶ 2.2.2. Coma
 - ▶ 2.2.3. Tail

2.2. Comet

Alternative Labels

- Comet
- Interplanetary Body: Comet

Broader Terms

2. Interplanetary Body

Narrower Terms

2.2.1. Nucleus
2.2.2. Coma
2.2.3. Tail

Comets: general

Broader Terms

Solar system

Selected	Left Vocabulary Term
<input checked="" type="checkbox"/>	TheGalaxy
<input type="checkbox"/>	Galaxies

Match Type	Right Vocabulary Term
broadMatch	avm-5
exactMatch	avm-5

Save All

Save Selected

Delete Selected

Delete All



- Use vocabulary concepts for
 - Tagging (using URI)
 - Resources in the registry
 - VOEvent packets
 - Searching by vocabulary concept
 - User keyword search converted to vocabulary URI
- Provides semantic advantages
 - Reasoning about terms
 - Relationships
 - Mappings
- Requires mechanism to convert string to concept



- Search and browse vocabularies
 - Configure
 - Vocabularies
 - Mappings
- Based on Information Retrieval techniques
 - Matching mechanisms
 - Ranking results

The screenshot shows a web browser window titled 'Vocabulary Explorer'. In the search bar, the word 'planets' is entered. Below the search bar, a message says 'Results 1 - 10 of 74 for "planets"'.

The main content area displays the following results:

- ▶ inner planet
- ▶ OUTER PLANETS
- ▼ 1. Planet
 - Alternative Labels**
1
 - Planet
 - Vocabulary**
 - The list of AVM keywords in SKOS format (Version 1.13)
 - Equivalent Concepts**
 - [Planets and satellites: general](#)

To the right, a 'Configuration' sidebar is visible, listing several checked options related to vocabularies and mappings, such as 'IVOAT Thesaurus - Version Tue Jul 29 11:39:52 2008' and 'The list of AVM keywords in SKOS format (Version 1.13)'.

<http://explicator.dcs.gla.ac.uk/WebVocabularyExplorer/>



Vocabulary Explorer

http://explicator.dcs.gla.ac.uk/WebVocabularyExplorer/

Vocabulary Explorer

Enter an astronomical search term and select search

Configuration

▼ Vocabularies

- IVOAT Thesaurus - Version Tue Jul 29 11:39:52 2008
- The list of AVM keywords in SKOS format (Version 1.13)
- Vocabulary for the IVOA Unified Content Descriptors (Version 1.13, 2008-07-29)
- Vocabulary for Astronomy & Astrophysics Journal keywords (Version 1.13)
- IAUT93 Thesaurus - Version Tue Jul 29 11:39:56 2008

[Select All](#) [Clear All](#)

► Mappings

► Settings

Developed as part of the [Explicator](#) project



Vocabulary Explorer

http://explicator.dcs.gla.ac.uk/WebVocabularyExplorer/#display%3D1

planet

Results 1 - 10 of 74 for "planet"

- ▶ [inner planet](#)
- ▶ [OUTER PLANETS](#)
- ▶ [1. Planet](#)
- ▶ [INNER PLANETS](#)
- ▶ [giant planet](#)
- ▶ [hypothetical planet](#)
- ▶ [Jovian planet](#)
- ▶ [outer planet](#)
- ▶ [planet satellite](#)
- ▶ [transplutonian planet](#)

1 [2](#) [3](#) [4](#) [5](#) [Next](#)

Configuration

▼ [Vocabularies](#)

- IVOAT Thesaurus - Version Tue Jul 29 11:39:52 2008
- The list of AVM keywords in SKOS format (Version 1.13)
- Vocabulary for the IVOA Unified Content Descriptors (Version 1.13, 2008-07-29)
- Vocabulary for Astronomy & Astrophysics Journal keywords (Version 1.13)
- IAUT93 Thesaurus - Version Tue Jul 29 11:39:56 2008

[Select All](#) [Clear All](#)

▶ [Mappings](#)

▶ [Settings](#)

Developed as part of the [Explicator](#) project



Vocabulary Explorer

Results 1 - 10 of 74 for "planet"

▶ [inner planet](#)
▶ [OUTER PLANETS](#)
▼ 1. [Planet](#)

Alternative Labels

1

Planet

Vocabulary

- The list of AVM keywords in SKOS format (Version 1.13)

Equivalent Concepts

- [Planets and satellites: general](#)

Narrower Terms

- [1.1.1. Terrestrial](#)
- [1.1.2. Gas Giant](#)
- [1.2.1. Surface](#)
- [1.2.2. Atmosphere](#)

► [Vocabularies](#)
► [Mappings](#)
► [Settings](#)



The screenshot shows a Mac OS X-style window titled "Vocabulary Explorer". The URL in the address bar is <http://explicator.dcs.gla.ac.uk/WebVocabularyExplorer/#browse%3Dhttp%3A%2F%2Fexplicator.dcs.gla.ac.uk%2FWebVocabularyExplorer%2F#browse%3Dhttp%3A%2F%2Fexplicator.dcs.gla.ac.uk%2FWebVocabularyExplorer%2F>. The window contains the following content:

Vocabulary Explorer

Search bar: planet

Planets and satellites: general

Vocabulary

- Vocabulary for Astronomy & Astrophysics Journal keywords (Version 1.13)

Equivalent Concepts

- [1. Planet](#)

Broader Terms

- [Solar system](#)

Configuration

- ▶ [Vocabularies](#)
- ▶ [Mappings](#)
- ▶ [Settings](#)

Developed as part of the [Explicator project](#)



- **Vocabularies improve search**
 - Remove ambiguity
 - Increase precision and recall
 - Enable
 - Reasoning about relevance
 - Faceted browsing
- **Provided tools for working with vocabularies**
 - Reliable search from keyword string to vocabulary term
 - Exploration of vocabularies
 - Mapping terms across vocabularies

Future Work

- Provide a string to concept transformation service
- Improve multi-vocabulary ranking results



University
of Glasgow

PRACTICAL SEMANTIC ASTRONOMY 2009

2-5 MARCH 2009

GLASGOW, UK



<http://www.dcs.gla.ac.uk/workshops/semast09/>