The history and future of vocabularies in the VO

Norman Gray, University of Glasgow

Astrosemantics workshop, Astroinformatics2012

Microsoft Research, Redmond, USA

2012 September 12

There are multiple thesauri developed and deployed within astronomy. This is no longer arcane; it's ready for deployment

vocabularies?
thesauri?
ontologies?

Mathilda is reading a paper online. She drags the paper into VOExplorer and asks for 'more like this'. VOExplorer calls out to a service which finds the AVM and Simbad equivalences of the paper's journal keywords, and uses the former to query a suitable service to find some pretty pictures (APOD), and the latter to query Simbad, presenting the two lists to Mathilda. There aren't many pretty pictures, so Mathilda asks to expand the search, and VOExplorer asks for pretty pictures corresponding to a more general term, found either directly in the AVM vocabulary, or finding a more general Simbad term and finding the AVM equivalent of that. The Simbad query, on the other hand, has produced far too many hits, so VOExplorer looks down the tree of Simbad terms which are 'narrower', and asks 'you were looking for compact objects: do you mean black holes, quasars, or...?' Once she has established a suitable keyword or keywords, she can make queries to find relevant registry entries and VOEvents. She finds some heterodyne observations, but she's an X-ray person, so is a bit vague, and curious, about just what that is – but oooh, there's a link to DBpedia and hence Wikipedia, so she goes there on the off-chance the article is decent.

What does this mean?

```
| pos.eq.ra | pos.eq.dec | phot.mag;em.opt.R |
+-----------+------------+-------------------+
|     12.34 |      56.78 |             16.12 |
```
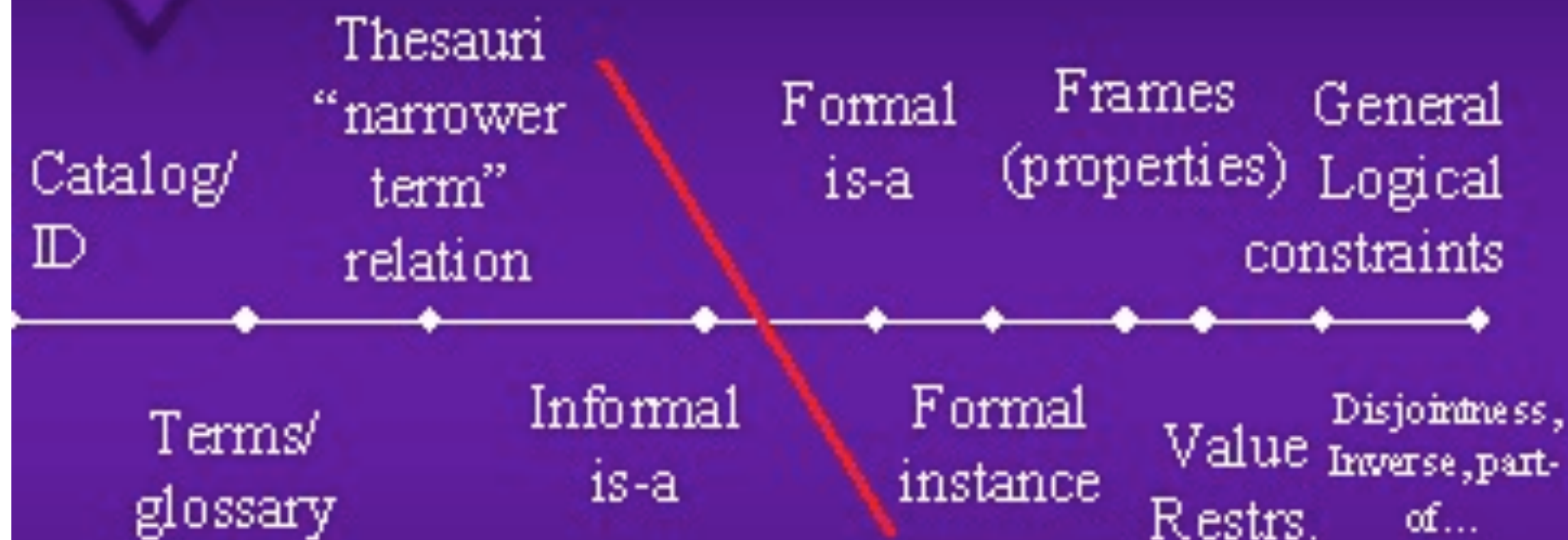
Observer: Dr Yseult Goatstrangler
OBS_DATE: 2010 March 19

This is an optical measurement of a star (because `em.opt.R` is an optical filter, and only celestial bodies have (RA, Dec))

Dr Goatstrangler is an optical astronomer (because she's observed a star using an optical filter)

She was not at a radio observatory on 2010-03-19 (because she was at an optical observatory, and that's different from a radio observatory)

So you could answer 'who was at an optical observatory in March?'

If X has an RA and Dec, it's a celestial object

If X has a em.opt.<anything>, it's visible in optical

If X looks at stars in the optical, X is an astronomer

Optical astronomy != radio astronomy

Select ?y where { ?y a OpticalAstronomer. }

from Deborah L. McGuinness, "Ontologies come of age" (2003)

POS_EQ_RA

pos.eq.ra

file://localhost/checkouts/volute.googlecode.com/trunk/projects/vocabularies/vocabularies-1.19/Vocabular...

java   +pin

Vocabularies in the Virtual Observatory
Version 1.19

IVOA Recommendation, 2009 October 7

This version

http://www.ivoa.net/Documents/REC/Semantics/Vocabularies-20091007.html

Latest version

http://www.ivoa.net/Documents/latest/Vocabularies.html

Previous versions

http://www.ivoa.net/Documents/PR/Semantics/Vocabularies-20090825.html
http://www.ivoa.net/Documents/PR/Semantics/Vocabularies-20081104.html
http://www.ivoa.net/Documents/PR/Semantics/Vocabularies-20080912.html
http://www.ivoa.net/Documents/PR/Semantics/Vocabularies-20080729.html
http://www.ivoa.net/Documents/WD/Semantics/vocabularies-20080320.html

Working Group

Semantics

Editors

Alasdair J G Gray, University of Manchester, UK
Norman Gray, University of Leicester / University of Glasgow, UK
Frederic V Hessman, University of Göttingen, Germany
Andrea Preite Martinez, INAF, Italy

```
<#AbsoluteMagnitude>
  a skos:Concept;
  skos:broader <#Magnitudes>;
  skos:inScheme <http://www.ivoa.net/rdf/Vocabularies/IAUT93>;
  skos:prefLabel "ABSOLUTE HELLIGKEIT"@de,
    "ABSOLUTE MAGNITUDE"@en, "MAGNITUD ABSOLUTA"@es,
    "MAGNITUDE ABSOLUE"@fr, "MAGNITUDINE ASSOLUTA"@it;
  skos:related <#Distance>, <#DistanceModulus>,
    <#HertzsprungRussellDiagram>, <#Luminosity>,
    <#WilsonBappuEffect> .
```
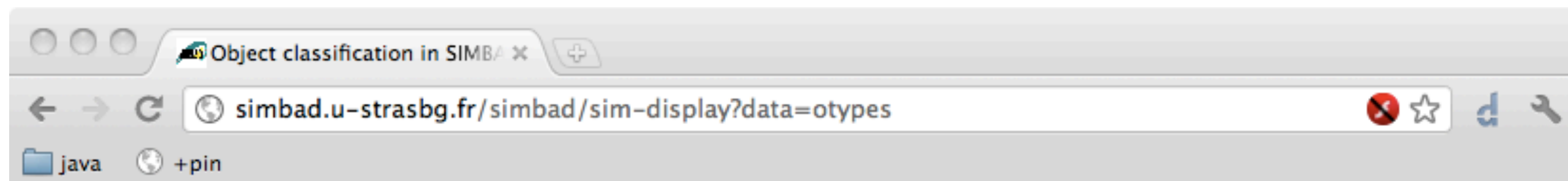
Gravitation	grv	Gravitational Source
· LensingEv	Lev	(Micro)Lensing Event
· Candidate_LensSystem	LS?	Possible gravitational lens System
· Candidate_Lens	Le?	Possible gravitational lens
· Possible_lensImage	LI?	Possible gravitationally lensed image
· GravLens	gLe	Gravitational Lens
· GravLensSystem	gLS	Gravitational Lens System (lens+images)

Candidates	..?	Candidate objects
· Possible_SClG	SC?	Possible Supercluster of Galaxies
· Possible_ClG	Cl?	Possible Cluster of Galaxies
· Possible_GrG	Gr?	Possible Group of Galaxies
· Candidate_**	**?	Physical Binary Candidate
· · Candidate_EB*	EB?	Eclipsing Binary Candidate
· · Candidate_CV*	CV?	Cataclysmic Binary Candidate
· · Candidate_XB*	XB?	X-ray binary Candidate
· · · Candidate_LMXB	LX?	Low-Mass X-ray binary Candidate
· · · Candidate_HMXB	HX?	High-Mass X-ray binary Candidate
· Candidate_Pec*	Pec?	Possible Peculiar Star

ontology of astronomical object types

- `http://purl.org/astronomy/ont/object-types`

- ...by Laurent Cambrésy, Sébastien Derriere, Paolo Padovani, Andrea Preite-Martinez, Alexandre Richard

- 2008 May (last updated 2012 April)

- See `http://www.ivoa.net/Documents/Notes/AstrObjectOntology/` and `http://www.ivoa.net/Documents/Notes/AstrObjectOntologyUseCases/`

BBC

News | Sport | Weather | Travel | Future | TV | Radio | More... | Search

SPORT

12 September 2012 Last updated at 17:01 GMT

Home | Football | Formula 1 | Cricket | Rugby U | Rugby L | Tennis | Golf | More Sports

SNOOKER An investigation is launched into Steve Davis's recent 4-0 win over Thepchaiya Un-Nooh due to concerns over betting patterns

Cameron apologises over Hillsborough

The prime minister says he is "profoundly sorry" for what he called the double injustice of the Hillsborough disaster.

■ Hillsborough report: Key findings   ■ Analysis: MPs united in shame

Headlines

TENNIS
Spaniard Ferrero to quit tennis

CYCLING
Cavendish wins stage to lead Tour 💬 7

SNOOKER
Davis v Un-Nooh match under investigation

FOOTBALL
Hodgson confident of England progress 💬 19

TAYSIDE AND CENTRAL SCOTLAND
Cheers as Murray returns to UK

MOTOGP
Spies confirms Ducati MotoGP move

FOOTBALL
Paladini interested in Birmingham

RUGBY LEAGUE

Video & Audio

Pietersen row very sad - Vaughan

Cameron apology over Hillsborough

Murray targets more Grand Slam wins 💬 119
TENNIS

From Perry to Murray: Ending the 76-year wait
TENNIS

Scores, Results & Fixtures

Latest Football

UPCOMING FIXTURES

SAT 15 SEP 2012 - PREMIER LEAGUE

linkeddata.org

As of September 2011

Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
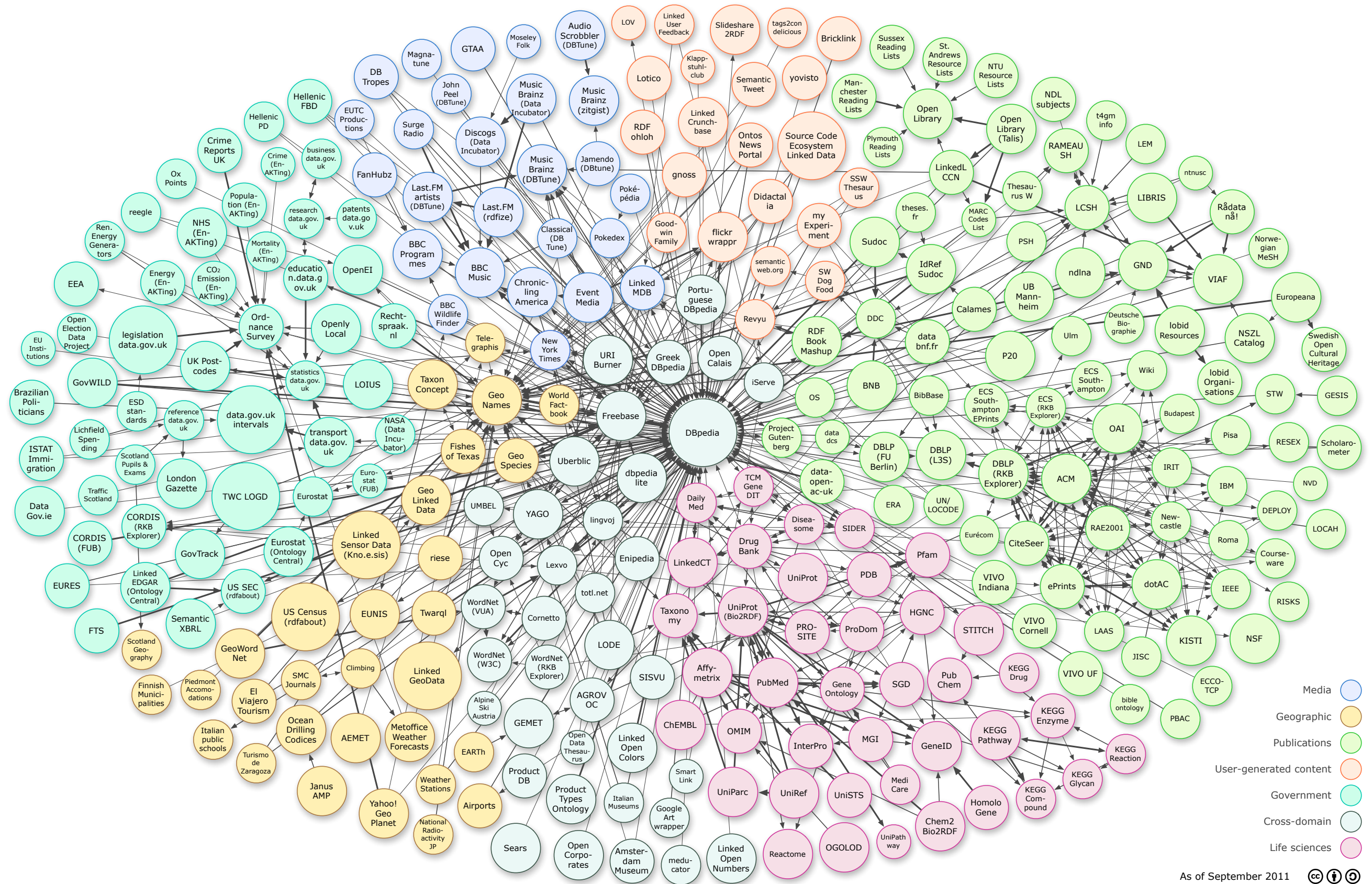
unified astronomy thesaurus

Norman Gray (Glasgow University, UK)

Alberto Accomazzi (ADS/Harvard-Smithsonian, US)

Graham McCann (Institute of Physics Publishing, UK)

Mark Cassar (American Institute of Physics, US)

Chris Biemesderfer (American Astronomical Society, US)

Chris Erdmann (Harvard Libraries, US)

Justin Soles (McGill University, Canada)

If you have an opinion, shout.

A consistent keyword system covering all publishers represented in ADS

Text-mining ADS's full-text, building on thesaurus concepts

Semantic query expansion

Intelligent faceted filtering

Recommendations and notifications

- Pass good metadata to ADS; helps discoverability

- Enhance end-user experience in a publisher-agnostic way

- Richer links to datasets

- Experiment with thesauri; think about ontology possibilities

- Create links to grants and funders (eg FundRef)

- identify & resolve incompatibilities between (astronomy parts of) existing thesauri (IAU93 and Rick Hessman's IVOAT are in their lineage)

- 'productise' the resulting thesaurus (need 'full' + 'core' division?)

- develop a stable maintenance process, mediating between publishers' and community's needs and expertise (technical and procedural problems)

- identify and prototype exemplar applications

- doesn't have to be Complete (Matthew's point)

- AIP and IOP will donate the astronomy parts of their thesauri to AAS.

- AAS owns the thesaurus, anticipating IAU blessing

- ...watched over by a panel of community stakeholders (archives, observatories, etc).

- Additions and edits from community

- ...screened/triaged by community editors

- ...periodically tidied/refactored (by librarian) for formal release.

theory vocabularies

Franck Le Petit, Nicolas Moreau
David Languignon, Carlo-Maria Zwölf
(Paris Observatory, Meudon)

Norman Gray
(Glasgow)

SimDM is about retrieving the results of simulations

So they want to search for simulations by specifying the types of input parameter and of output results

...intelligently

As described in the IVOA standard, Simulation Data Model, registrations of theoretical services, require to provide several URIs corresponding to semantics keywords describing services and simulations. VO-Theory concepts are based on SKOS description as recommended by the IVOA Semantic Working Group.

Example of a VO-Theory URIs : http://purl.org/astronomy/vocab/Algorithms/GaussSeidel

This website is dedicated to the discovery of these URIs. Navigate through the broader, narrower, related terms to discover the most precise concept you wish.

To suggest new concepts or corrections, contact : VOTheory.semantics @ obspm.fr.

The development of the VO-Theory vocabularies rely on Poolparty, a thesaurus management system and a SKOS editor developed by punkt.netServices.

poolparty

Request

Choose a vocabulary

Search a concept

✓ Algorithms
Astronomical object
Input parameters
Physical processes
Physical quantities
Data object types
Astronomical dictionnary sample
Chemical species
Pdr Parameters
Parameters

All concepts

3+1 Formalism
8-Wave Scheme
Accelerated Lambda Iter
Adaptive Mesh Refineme
Advection Upstream Splitting Method
Algorithm
Alternating Direction Implicit
BiConjugate Gradient
BiConjugate Gradient Stabilized
Block Based AMR
Bulirsch-Stoer
Cell Based AMR
Cell Centred
Central Difference Scheme
Chebyshev Iteration

VAMDC molecules

Thesaurus:

```
@base <http://purl.org/astronomy/vocab/molecules/>

<XLYOFNOQVPJJNP-UHFFFAOYSA-N>
    a :Concept;
    :prefLabel "H2O"@zxx, "Water"@en;
    :altLabel "Dihydrogen monoxide"@en, "H2(16O)"@zxx,
        "Water vapor"@en;
    :narrower <XLYOFNOQVPJJNP-DQGQKLTASA-N>,
        <XLYOFNOQVPJJNP-XKYOGGAFSA-N>;
    :notation "H<sub>2</sub>O"^^<http://www.w3.org/1999/xhtml>,
        "InChI=1S/H2O/h1H2"^^<http://www.iupac.org/inchi/>;
    rdfs:seeAlso chebi:CHEBI_15377;
    m:isSKOSConceptFor ont:XLYOFNOQVPJJNP-UHFFFAOYSA-N .
```

norman gray

Ontology:

```
@base <http://purl.org/astronomy/ont/molecules/>

<XLYOFNOQVPJJNP-UHFFFAOYSA-N>
    m:hasSKOSConcept thes:XLYOFNOQVPJJNP-UHFFFAOYSA-N;
    a owl:Class;
    rdfs:comment "The class of molecules H2O";
    rdfs:seeAlso thes:XLYOFNOQVPJJNP-UHFFFAOYSA-N;
    rdfs:subClassOf chebi:CHEBI_23367;
    owl:equivalentClass chebi:CHEBI_15377 .

<XLYOFNOQVPJJNP-DYCDLGHISA-N>
    m:hasSKOSConcept thes:XLYOFNOQVPJJNP-DYCDLGHISA-N;
    a owl:Class;
    rdfs:comment "The class of molecules HD(16O)";
    rdfs:subClassOf <XLYOFNOQVPJJNP-UHFFFAOYSA-N> .
```

- a little structure goes a long way

- many of the thesauri here were preexisting ones repurposed into SKOS

- multiple thesauri are fine (no need for completeness or consensus)

- thesaurus + ontology sometimes necessary

- this is all about linking in the *machine-readable* web

norman gray

There are already multiple thesauri developed and
This is no longer arcane; it's ready for deployment
deployed within astronomy, with more to come.

So go forth and thesaurise!