

A 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)

IVOA Interop, Urbana, 2012 May 21

JISC

	public.		
Research tools - projects to develop sustainable and open vocabularies for research and information management	Projects to build upon existing vocabularies that support academic disciplines, to develop exemplars demonstrating the support of the semantics of a discipline.	Total £300,000 Up to five projects Up to £70,000 each Jun 2012 – Jan 2013	Appendix B3
Research tools -	A project to synthesise lessons	Total £50,000	Appendix B4

So, a nice bounded project, to tie together existing and developing astronomy thesauri, and think about sustaining it for the future.

Then JISC got its funds for this call halved at the last moment

booooo!!!

hisssssss!!!

So, what _was_ plan a?

where we are

- IAU's 1993 thesaurus has ~3000 terms, but is out of date and unmaintained (too large to be useful?)
- Rick Hessman's IVOAT as an IAU-93 update
- IVOA REC: <http://www.ivoa.net/Documents/Latest/Vocabularies.html> (inc. 4 x 300-term thesauri)
- IoPP and AIP have advanced drafts of professionally-developed all-physics thesauri (donate astro subsets)
- VOTheory WG is developing thesauri for data access (Le Petit talk, Tuesday 09:00; & Wednesday 14:00)

norman gray

...useful with then technology

donation offer still open?

REC includes vocabs

IOP/AIP thesauri both have IAU-93 & journals in their lineage

- 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

use-cases for publishers

- 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
- '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

- Need centralised editorial and release management
- ...while encouraging contributions/edits/additions/deletions from the community
- ...and ideally not developing any new software.
- Reconciling the existing thesauri will prototype this process: indicates what long-term maintenance looks like

plan b

any ideas?

- We thought this was 12–15 staff months of effort; does that sound plausible?
- Anyone disagree with the status and goals?
- How much community involvement is necessary, or desirable, or manageable?

Norman Gray, <http://nxg.me.uk>

norman gray

we asked for 9 staff-months; more from IoP & ADS