Finding Data Resources in a Virtual Observatory using SKOS Vocabularies

Alasdair J G Gray¹, Norman Gray², and Iadh Ounis¹ ¹Department of Computing Science, University of Glasgow ²Department of Physics and Astronomy, University of Leicester http://explicator.dcs.gla.ac.uk



Overview There are a large number of data resources available to astronomers. We show how their search for "relevant" data sources can be improved by using controlled vocabularies.

1. Introduction

3

υ

៧

σ

Ø

A Virtual Observatory (VO) aims to allow seamless access to distributed data sources. Astronomers need help in *identifying* which resources contain *relevant* data for their needs.

Astronomers currently search using VOExplorer (Fig. 1):

- Resources are tagged by providers using arbitrary terms
- · Users search for relevant resources with keywords
- Results displayed according to *facets* generated from the ragged meta-data about the resources.

2. Problems

- **Terminology**: Disparate astronomical fields have developed specialist terms.
- Mismatch: Tags and search text are not given definitions, leading to ambiguity.
- Classification: Facets generated from tags, leading to "pot luck" groups.

Fig. 2: SKOS Encoded Vocabulary Concept

@base <http://...>
@prefix iaut: <http://.../iaut.ttl>
@prefix skos: <http://.../skos/core#>

skos:inScheme <>; skos:narrower <#Moon>, <#Titan>; skos:prefLabel "Planet satellite"@en, "Satellite planetaire"@fr; skos:related <#ArtificialSatellite>; skos:exactMatch iaut:PlanetSatellites .

4. Search Using Vocabularies (Fig. 3)

Tagging:

- Resource R₁ declares it is about "Natural Satellites"
- "Natural Satellites" corresponds to the vocabulary concept v:PlanetSatellite
- R₁ tagged with v:PlanetSatellite

Searching

- Looking for resources about the "Moon"
- "Moon" corresponds to the vocabulary concept v:Moon
 Since no resources tagged with v:Moon, the search is
- automatically expanded to broader concepts
- v:Moon has v:PlanetSatellite as a broader concept
- Resource R₁ returned as a search result

5. Conclusions

- Using vocabularies to aid search could lead:
- · Improved precision by exploiting definitions
- Increased recall by exploiting relationships
- Refinement of facets by using vocabulary groupings
- A prototype service for matching vocabulary concepts: http://explicator.dcs.gla.ac.uk/WebVocabularyExplorer

Fig. 1: VOExplorer Screenshot



University of Glasgow

Computing Science

Department of

3. Controlled Vocabularies

Vocabularies formalise the terminology used within a domain of discourse as a set of *concepts*. The concepts form a poly-hierarchy based on their semantic relationships.

SKOS encodes the vocabulary in a machine understandable way, providing the *loose semantic relationships* required for a vocabulary and a URI to uniquely identify a concept (Fig. 2).

Within a single application domain several vocabularies can exist for different purposes, e.g. there are four astronomical vocabularies with a SKOS encoding. Concepts in different vocabularies can be related using *mappings*, similar to the semantic relationships, e.g. skos:exactMatch.

Fig. 3: Semantic Search



EPSRC Engineering and Physical Sciences Research Council

This work is funded by UK EPSRC grant EP/E01142X/1.