Astronomical Data Analysis Software and Systems XVIII **O9.4**
ASP Conference Series, Vol. ???, 2009
D.A. Bohlender, D. Durand & P. Dowler

PREPRINT: to appear in Proceedings of ADASS XVIII, Bohlender et al., ASP, 2009

The SKUA project and the Semantic Web

Norman Gray[1], Tony Linde

Department of Physics and Astronomy, University of Leicester, UK

Kona Andrews

Royal Observatory Edinburgh, UK

Abstract. The Semantic Web promises much for software developers, but because its claimed benefits are rather abstract, there is little obvious incentive to master its unfamiliar technology. In contrast, many 'Social Web' applications seem rather trivial, and not obviously useful for astronomy.

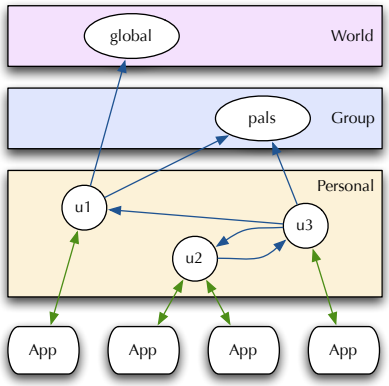
The SKUA project (Semantic Knowledge Underpinning Astronomy) is implementing a service which will realise the benefits of both these web technologies. This RESTful web service gives application authors ready access to simple persistence, simple (social) sharing, and lightweight semantics, at a low software-engineering cost. The SKUA service allows applications to persist assertions (such as bookmarks and ratings), and share them between users. On top of this, it provides lightweight, astronomy-specific, semantics to enhance the usefulness and retrieval of the users' data. We will demonstrate the service, and astronomical applications using it.

1. Introduction

For all its current fashionability, we can identify at least two reasons why the Semantic Web excites little interest among astronomical software developers. Firstly, there is so far no well-known 'killer app' for the semantic web, and the use-cases sometimes brandished in support of the Semantic Web's promise – involving machines booking hospital appointments, or comparing prices (Berners-Lee et al., 2001, and see `http://www.w3.org/2001/sw/`) – are not obviously relevant to astronomical applications development. Secondly, even when a potential application is dimly discernable – and it would surely be useful somehow for a machine to 'know' that a black hole is a type of compact object – there are multiple barriers of novel terminology and technology to be overcome before an idea can be turned into a useful software product.

In the SKUA project (`http://myskua.org`) we are developing an infrastructure which addresses both of these concerns. The SKUA infrastructure provides a mechanism for persisting and sharing a flexible range of application state, including annotations (of which we give examples below), in a way which lets applications transparently take advantage of lightweight semantic knowledge in the SKUA system. By combining the social aspects of the annotation

[1]Also Department of Physics and Astronomy, University of Glasgow, UK




```
@base <http://blah/my-sac>.
@prefix s:
  <http://myskua.org/claimtypes/1.0/>.
@prefix dc:
  <http://purl.org/dc/elements/1.1/>.
<#b1>
  a s:bookmark;
  s:ref [
    a s:webpage;
    s:url <http://www.w3.org/2001/sw/>;
    dc:title "SemWeb @ W3C";
    s:extended "The W3C nexus"
  ];
  s:tag "semanticweb", "rdf";
  s:time "2008-01-21T18:11:58Z".
```

Figure 1. SKUA's sharing architecture: on the left we show the relationships, both peer-to-peer and hierarchical, between annotatation stores, with double-headed arrows indicating read-write relationships with applications, and the single-headed arrows indicating the federation of queries between services; and on the right we illustrate a potential annotation type, in this case a URL bookmark, using the Turtle notation for RDF (Beckett & Berners-Lee 2008).

sharing and the lightweight semantics, the SKUA infrastructure can be regarded as a simple 'Web 3.0' application, to the extent that that term represents the anticipated melding of Web 2.0 applications with Semantic Web technologies.

2. The SKUA infrastructure

The SKUA infrastructure consists of a network of assertion services, which are either on serparate servers or are logically distinct entities on a single server, and to which applications write per-user state information – 'assertions' – such as annotations ('this paper is good'), or preferences ('I'm interested in pulsars'). These annotations can then be retrieved, using a SPARQL query (Prudhommeaux & Seaborne 2008), by the same application, by another instance of the same application, or by a cooperating application.

The infrastructure also allows these assertions to be shared between users or shared to a common annotation server, in such a way that an application's SPARQL query against its 'local' service is forwarded to the services it delegates to (see Fig. **??**). Thus if, in that figure, user 'u1' shares the assertion that 'paper X is good', then when an application belonging to user 'u3' looks for good papers, it picks up the corresponding assertion by 'u1'. This query federation will be permitted only if the user whose service this is explicitly allows it (which is important in the case where the assertion is something like 'the author of paper Y is clearly mad').

SKUA is therefore an architecture for saving and sharing annotations of astronomical and other resources, which can be updated and queried via a RESTful API. It is targetted at application programmers rather than end-users, and is intended to help them make their applications better without having to make

significant speculative investments in semantic technology. The infrastructure is therefore a network-based, shareable semantic 'memory' for an application, which is agnostic about what information is stored.

3. Example applications

Part of the SKUA project is to develop applications which use the project's infrastructure, both as a way of validating the approach, and for their intrinsic usefulness. As well, we are cooperating with the developers of existing applications to support them in adding SKUA interfaces where appropriate.

In particular, we are developing *Spacebook* (Linde et al., 2009), as an adaptation of the myExperiment code-base (De Roure and Goble 2007, see also `http://myexperiment.org/`) which allows scientists to share digital objects of various kinds, supporting the development of communities. Spacebook builds on this by adding integration with AstroGrid's Taverna workflows, and lets users tagging resources using the SKUA infrastructure.

On top of that we are adapting, or have plans to adapt, VOExplorer (Tedds et al. 2008) and Paperscope (`http://paperscope.sourceforge.net/`). Basic support for SKUA exists in development versions of VOExplorer, replacing the local-only tagging and labelling support already present.

In the released versions of VOExplorer, there is a box beside the registry entry details where the user can specify a highlight colour, notes about the resource, an alternative title, and tags. In the SKUA-enabled version these facilities are also available, though stored in a SKUA server rather than in a local file, with the addition that, if the users attach keywords from one of the existing IVOA vocabularies (Gray et al., 2008; Gray et al., 2009) a subsequent search on the SKUA store is able to take advantage of the lightweight semantics associated with these keywords. For example, if a user annotates a resource with `aakeys:Ephemerides`, they can later make a SPARQL query for terms which have `AstrometryAndCelestialMechanics` as a broader term, and in doing so pick up resources tagged with `Astrometry`, `CelestialMechanics`, `Eclipses`, `Ephemerides`, `Occultations`, `ReferenceSystems` or `Time`.

We emphasise that we are not expecting users to write SPARQL queries for themselves, but instead expect applications to issue them on the user's behalf, based on simple query templates.

The SKUA software is available at `http://skua.googlecode.com`. The current version, at the time of writing, supports updating, persistence, querying and delegation; vocabulary-aware querying is available but undocumented; easier sharing and per-user security are in development.

Acknowledgments. The SKUA project is funded by the UK's Joint Information Systems Committee (JISC, `http://www.jisc.ac.uk`).

References

Beckett, D. & Berners-Lee, T. *Turtle – the Terse RDF Triple Language*, January 2008. W3C Team Submission: `http://www.w3.org/TeamSubmission/turtle/` (last retrieved January 2009)

Berners-Lee, T., Hendler, J. & Lassila, O. *The Semantic Web*, May 2001, Scientific American

De Roure, D. & Goble, C. 2007 *myExperiment – A Web 2.0 Virtual Research Environment*. In: International Workshop on Virtual Research Environments and Collaborative Work Environments, May 2007, Edinburgh, UK

Gray, A. J. G., Gray, N., Hessman, F. V. & Preite Martinez, A. (2008) *Vocabularies in the Virtual Observatory*, IVOA Proposed Recommendation `http://www.ivoa.net/Documents/latest/Vocabularies.html` (last retrieved January 2009)

Gray, A. J. G., Gray, N. & Ounis, I. 2009, in ASP Conf. Ser. XXX, ADASS XVIII, ed. D.A. Bohlender, D. Durand & P. Dowler (San Francisco: ASP), [O9.4]

Linde, T., Gray, N. & Andrews, K. 2009, in ASP Conf. Ser. XXX, ADASS XVIII, ed. D.A. Bohlender, D. Durand & P. Dowler (San Francisco: ASP), [E.20]

Prudhommeaux, E. & Seaborne, A. *SPARQL Query Language for RDF*, January 2008. W3C Recommendation: `http://www.w3.org/TR/rdf-sparql-query/` (last retrieved January 2009)

Tedds, J. A., Winstanley, N., Lawrence, A., Walton, N., Auden, E. & Dalla, S. 2008, in Argyle, R. W., Bunclark, P. S. & Lewis, J. R. 2008, eds. ASP Conf. Ser. 394, ADASS XVII (San Francisco: ASP), p159–162