Reasoning about Access Control

Norman Gray
VOTech/AstroGrid
University of Leicester, UK
(and University of Glasgow, UK)

IVOA Interop, Moscow, 2006 September 18–20

overview

Access control is a very natural ‘Semantic Web’ problem. It involves:

- reasoning, solving logic problems
- flexible querying
- interoperability (ie, scavenging & repurposing information)

Ingest information from everywhere; combine it; reason about it; query it. Metcalf’s law for RDF. Heavily standards-based.

rdf quick intro

- All the world is *triples*, consisting of *resources* named by URIs (`ivo:...` or `urn:example#Norman`)
- . . . which have *properties* whose *values* are resources or literals.
- RDF/RDFS/OWL describe these using `rdf:type`, `rdfs:subClassOf`, `owl:symmetricProperty`, and so on.

There is an analogy with XML Schemas, *but it is a loose one* – they're not addressing the same problem. Same for O-O.

rdf/owl/semweb – wins and losses

RDF/OWL/reasoning now largely stable (though The Semantic Web will forever be Vision). Now engineering rather than CS.

Using the architectural principles which let HTML take over the internet. Very open and flexible; has existing powerful query language. Did I mention standards?

RDB to XML to RDF – spectrum of strengths. XML is more natural than RDF where the information density is high, and the information regular or highly constrained; RDF/SW is natural for incomplete or ragged data.

access control

This talk is about what happens *after* you've authenticated.

Access control maps *very* naturally to an ontology-style question.

About answering the question 'is this user provably a member of the group which is allowed access to the resource?'

Two demos here: delegation/federation of access, and extracting information from X.509 certificates.

use cases

See `http://wiki.eurovotech.org/twiki/bin/view/VOTech/AccessControlUseCases`

- database subset queries
- chain of group membership
- local/remote delegation
- proxy and attribute assertion certificates
- quotas

non-ontology approaches

- ACLs in filesystems: confusing.
- PERMIS: well-known, but not naturally capable of federation/delegation (closed-world).
- Shibboleth: concerned with attribute transmission (which *is* part of the challenge).
- XACML: procedural, not declarative; no delegation.
- Rule-based (policy) systems: plenty; opaque; developing. PeerTrust/ProTune (includes negotiation – hard).

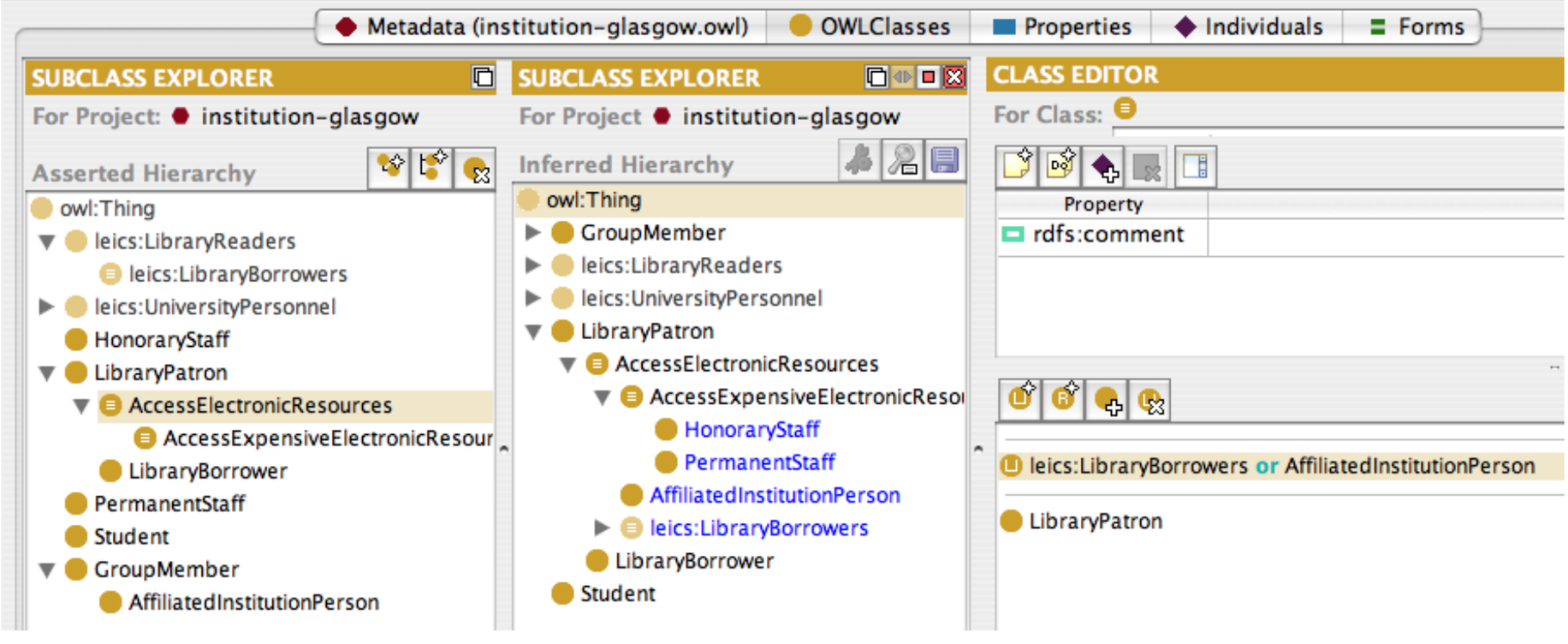
delegation: glasgow and leicester libraries

Glasgow lets honorary staff and permanent (Glasgow) staff access its expensive electronic resources, and additionally gives access to non-expensive resources to all members of the Leicester LibraryBorrowers class.

Leicester lets AcademicStaff or Students be LibraryBorrowers. `ng59@le.ac.uk` in DepartmentOfPhysicsAndAstronomyStaff, so in UniversityStaff and LibraryBorrower.

So `ng59@le.ac.uk` is allowed access to GU electronic resources, but initially not expensive ones, even though `norman@astro.gla.ac.uk`, being GU HonoraryStaff, is.

delegation: picture



Generic SPARQL endpoint (uses Jena and Tomcat); API is pure HTTP GET/POST/PUT/DELETE.

[demo]

quaestor demo

Interface at `http://192.168.169.216:8080/quaestor/`

Snapshot of knowledgebases at
`http://192.168.169.216:8080/quaestor/kb/`

Get knowledgebase:

```
% curl http://host:8080/quaestor/kb/delegation
```

querying using sparql

```
% curl http://localhost:8080/quaestor/kb/delegation \
    --header content-type:application/sparql-query \
    --data-binary @access.rq
...
%
```

Query access to all data

```
prefix gla: <http://ns.eurovotech.org/access-control
  /institution-glasgow.owl#>
select ?person
where {
  ?person a gla:AccessElectronicResources
}
```

access-control bis: x.509

An X.509 certificate is an identity assertion, but we can also subvert the set of attributes inside it to make it a source of *reliable* RDF triples, too.

That done, we can reason with the result.

Very SemWeb: it doesn't matter where your data comes from, as long as you can massage it into subject-predicate-object form.

Or use proxy certificates. Or, possibly better, use X.509 Attribute Assertion certificates (coming soon).

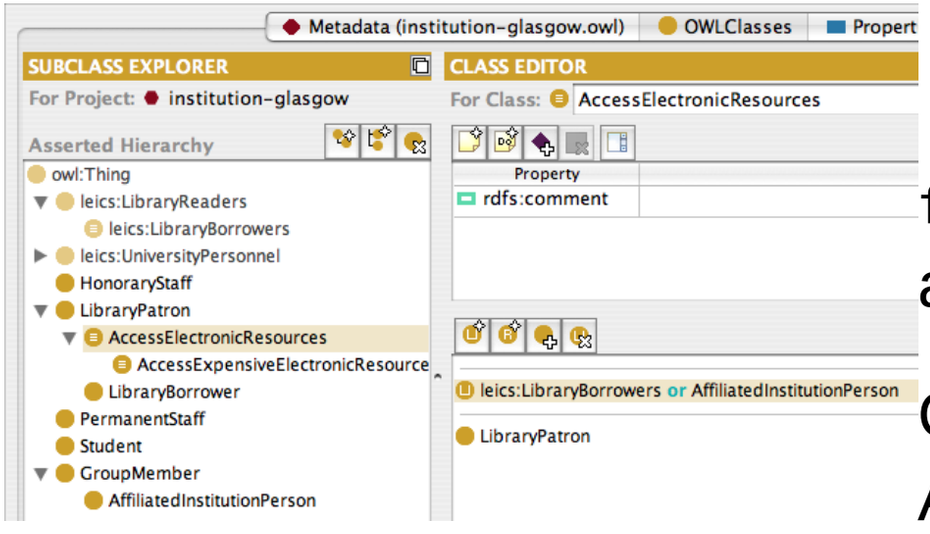
Or SAML assertions.

x.509 demo

[demo]

- We queried `foaf:name`, implied by the X.500 ontology – we didn't have to care that this wasn't originally a FOAF assertion. Cf, distinction between EEC and Proxy cert.
- Could do the reasoning – that is, express the policy – within the SPARQL query. . .
- . . . or we could do it within an ontology (eg, class membership is defined by presence of a particular attribute, or some more complicated logical predicate).

x.509 and access



`fred.bloggs@example.edu` has an eScience certificate.

Can he AccessElectronicResources?

[demo]

to do

What's next? Some suggestions:

- Ingest Proxy and Attribute Assertion certificates
- Ingest SAML
- LDAP-to-RDF?
- UI support for making assertions
- Toolkits and templates for expressing policies
- Keep an eye on ProTune

summary

I’ve presented an approach, rather than a tool.

Keep the making of assertions, the transmission of assertions, and the reasoning as decoupled as possible. Open-world.

Give resource owners flexibility.

RDF helps here, by being the ‘highest common factor’ of multiple systems.