



Fig. 1



Fig. 2

1. VOResource 1.1

(cf. Fig. 1)

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

(cf. Fig. 2)

An overhaul of the central schema of the Registry after almost 10 years

- Main goal: DOIs, ORCID's
- Secondary goal: Consistency with DataCite metadata
- And then some touches based on usage experience and community practices
- Overarching requirement: Strictly no breaking stuff

Many minor things, sorted by estimated contentiousness.

2. Technics

Currently an internal Working Draft.

I'd like to fast-track this to REC within this year. Please protest as early as you can.

Formatted WD: <http://docs.g-vo.org/VOResource.pdf>

Preferred source:

<http://volute.g-vo.org/svn/trunk/projects/registry/VOResource>

(including version history, diffs, and all) – please commit your fixes!

3. DataCite as Additional Parent?

DataCite core is a metadata schema for DOI registration.

It largely covers the same ground as VOResource, but differs in terminology and such. Since it's hard enough to convince people to come up with good metadata for one scheme, I'd really like to avoid trying to make them do two disjunct sets of metadata – and I hope we'll see a lot of DOIs on services and datasets.

VOResource so far pretended to be a serialisation of RM.

Proposal: Make it „RM+DataCite“

Alternative: Change RM to include DataCite.

4. DOIs, ORCID's

Only one new element: altIdentifier. Contrived example (identifiers are invented):

```
<ri:Resource created="2008-05-07T11:33:00"...>
  <title>GAVO Data Center Registry</title>
  <identifier>ivo://org.gavo.dc/registry</identifier>
  <altIdentifier>doi:10.5072/TWFya3Vz</altIdentifier>
  <altIdentifier>x-bibcode:2008gavo.conf...17D</altIdentifier>
  <curator>
    <publisher>GAVO Heidelberg Data Center</publisher>
    <creator>
      <name>Anton Y. Mous</name>
      <altIdentifier>orcid:0000-0000-0000-000X</orcid>
    </creator>
  </curator>
</ri:Resource>
```

5. altIdentifier questions

- Note that altIdentifier currently can sit in two places. Should it be possible in more places (capabilities perhaps? There's also the question of unifying party-like elements).
- What do people think of x-bibcode:?
- Instead of relying on URI schemes, should we have altIdentifier/@type?

6. Vocabularies iso Enumerations

VOResource 1.0 has vocabularies for:

- `date/@role` („What happened to the resource?“)
- `relationshipType` („Is this service-for or served-by?“)
- `content/type` („Is this a catalog or an image collection?“)
- `content/level` („Is this for kindergarten or postgrads?“)

In all cases,

1. I've lifted the restrictions in either schema or text,
2. created vocabularies
3. keeping VOResource 1.0 compatibility but moving towards DataCite.

7. Vocabulary Maintenance

I will not discuss vocabularies with more than 4 persons at a time (but any 3 are welcome at any time).

Valuable Discussion: *How do we maintain the vocabularies?*

SAMP MType maintenance seems to work fine. I'd like to copy that.

[i.e., Have a wiki page; since we have RDF vocabularies we'd only keep new terms there until they're moved to RDF]

There's also the elephant in the room: Shouldn't we just completely move to RDF and actually have URIs wherever there terms now? And if so, can we do this non-disruptively, e.g., by defining special rules on prepending URI fragments to each term? I won't do it, but if someone did the grunt work, I'd probably go along.

8. testQueryString

A new optional child of interface intended to help validators.

Currently: Registry extensions define structured `testQuery` types.

Problem: Services without an extension type don't have test queries (e.g., datalink).

Alternative: standard MAXREC=1 for discovery protocols, other might perhaps not need it?

9. IdentifierURI

IdentifierURI currently does not allow fragment/query parts.

```
ivo://ivoa.net/std/RegTAP#1.0
```

hence doesn't match. That's bad because, for instance, our new-style standard ids will look like that. Also, IdentifierURI is a bad name for two reasons: IVOA identifiers always allowed fragments and query parts, and there's non non-URI version any more anyway.

I'd like to deprecate IdentifierURI and instead have a type `IVOAIdentifier` (`xs:anyURI` with `ivo` scheme).

If necessary, a new `RegistryReference` could keep current IdentifierURI RE.

Didn't dare so far. Opinions?

10. Unified party type?

Content models for party-like entities in VOResource 1.0:

	contact	publisher	creator	contributor
<code>role_name</code>	<code>co/name</code>	<code>cp</code>	<code>cc/name</code>	<code>cb</code>
<code>role_ivoId</code>	<code>co/name/@ivo-id</code>	<code>cp/@ivo-id</code>	<code>cc/name/@ivo-id</code>	<code>cb/@ivo-id</code>
<code>address</code>	<code>co/address</code>	N/A	N/A	N/A
<code>email</code>	<code>co/email</code>	N/A	N/A	N/A
<code>telephone</code>	<code>co/telephone</code>	N/A	N/A	N/A
<code>logo</code>	<code>co/logo</code>	N/A	<code>cc/logo</code>	N/A

At least contact and creator could be unified with a single type, with publisher and creator following in VOResource 2.0. RegTAP already has all of these in one table.

Didn't dare so far. Opinions?

11. contributorType

While we're at it: DataCite has a `contributorType` with values like

- `DataManager`, `Distributor`
- `Editor`, `Sponsor`, `Funder`,
- `HostingInstitution`
- `ProjectLeader`, `ProjectManager`, `ProjectMember`,
- or even `RegistrationAuthority` and `RightsHolder`.

Do we want this? Perhaps even on the unified party type?

12. Language

Moet ons praat oor Afrikaanstalige records?

Ĉu ni parolu pri esperantlingva records?

Ar chir dinn labhairt faoi thaifid Gaeilge?

Sollten wir über deutschsprachige Records reden?

Should we talk about records with non-English content?

And should we worry about `vr:Organisation` vs. `vs:CatalogService`?

Came up during education work. DataCite lets people use `xml:lang`.

On the question of British vs. US spelling I'd say variations there are annoying, but overall rare enough that we can afford not to worry.

13. interfaces: multi-accessURL

VOResource 1.0 allows multiple accessURLs per interface. And multiple interfaces per capability, and multiple capabilities per resource. Last time I checked, nobody was using it.

But it was a headache in RegTAP design. Strike it?

Or actually use it for mirror management? In that case, we should probably mark a mirrors/secondary URLs.

Again, I will not actively pursue this topic, but people who want mirrors in the registry should speak up now.

14. interfaces: version

When VOResource was designed, multi-version support was expected to be:

```
capability(SIA)
  interface v-1.0
  interface v-2.0
  interface web
  common metadata
```

In VO practice, things rely much more on standardIDs:

```
capability(SIA-v1)
  interface
  metadata-v1
```

```
capability(SIA-v2-sync)
  interface
  metadata-v2
```

```
capability
  interface web
```

Hence, struck "1.0" default for `interface/@version` and changed some prose.

15. interfaces: std

VOResource 1.0 implies that clients should discover standard interfaces by looking for `capability[standardID='...']/interface[role='std']`.

In reality, some services neglected to have the `std`, and clients didn't bother to check it.

Now discouraging having non-standard interfaces in capabilities with standardIDs.

16. rights

In VOResource, there's `vr:Rights`, which is one of

public secure proprietary.

Proposed now:

```
<rights rightsURI="http://creativecommons.org/licenses/by/3.0">
  The images from the X survey are copyright 2016, the X project.
  They are published under the creative commons attribution 3.0 unported
  license. If you use this data, please cite doi:10.5072/7273288.
```

```
  Images are under embargo for one year after their addition to the
  repository .
```

```
</rights>
```

Anyone in for a collection of useful license URIs?

17. Light description markup

```
select count(*) from rr.resource where res_description like '%----%'
```

yields 78 records in the current Registry. People are putting ASCII markup into descriptions.

However: `description` is `xs:token`, so all whitespace is equivalent to just one blank.

Now making them `xs:string` and recommending to show them non-reflowed.

18. Timestamp Format

VOResource 1.0 has

2016-04-28T15:50.00

as timestamp format. OAI-PMH (and more and more of the VO) has

2016-04-28T15:50.00Z.

Now allowing and severely recommending the trailing Z.

19. Creator.name format

In the current Registry, you can see creator.names like these:

```
Kornilov, V.G., Volkov, I.M., Zakharov, A.I., Kozyreva, V.S., Kornil...
NRAO, generated by J.J. Condon, J.J. Broderick and G.A. Seielstad, D...
CompTel Instrument Team. Maps generated \n\n by Andrew Strong, Max-...
Helio
The Two Micron All Sky Survey is a joint project of the University o...
Data taken by ROE, AAO, and CalTech, Compression\nand distribution b...
All data is downloaded from the <a href=http://galax.stsci.edu>\nMAS...
Max Planck Institute for Exterrestrial Physics (Garching FRG)
PI: Ata Sarajedini
ASB branch STScI
Akylas et al.
```

Now encouraging single-name, *last-name first* formats.

20. securityMethod

securityMethod is essentially unused in the current VO.

Do we want to use this opportunity to fix it?

(VOResource 1.0 had expected some extension schema to deal with it)

(I'm waiting for input from GWS)

21. Teaser Picture

So... package managers let people view screenshots or similarly representative images.

Would we want this? Would people provide images?

I'd like this.

But I don't care enough...

22. You've made it!

Other proposals?
Interested in vocabulary work?
Simple questions?
Complicated questions?

Thanks.