Fig. 1

Fig. 2

Fig. 3

1. Towards RegTAP 1.1

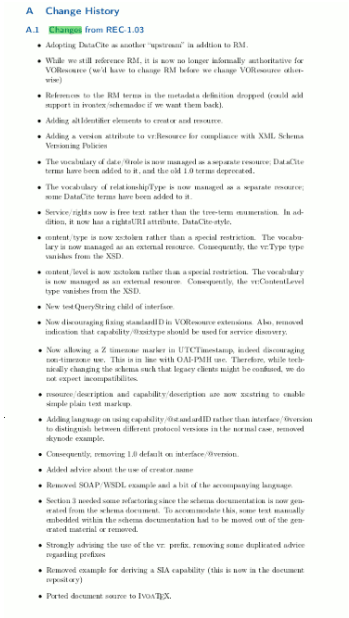
(cf. Fig. 1)

Markus Demleitner
msdemlei@ari.uni-heidelberg.de

(cf. Fig. 2)

- What this is about
- an alt_identifier table
- a standard_key table
- new interface.mirror_url
- per-capability testQueryString in res_details
- rights, rightsURI in res_details

(cf. Fig. 3)


A Change History

A.1 Changes from REC-1.03

- Adopting DataCite as another "upstream" in addition to RM.
- While we still reference RM, it is now no longer informally authoritative for VOResource (we'd have to change RM before we change VOResource otherwise)
- References to the RM terms in the metadata definition dropped (could add support in rconfex/schemadoc if we want them back).
- Adding altIdentifier elements to creator and resource.
- Adding a version attribute to vr:Resource for compliance with XML Schema Versioning Policies
- The vocabulary of date/@role is now managed as a separate resource; DataCite terms have been added to it, and the old 1.0 terms deprecated.
- The vocabulary of relationship@Type is now managed as a separate resource; some DataCite terms have been added to it.
- Service/rights now is free text rather than the tree-term enumeration. In addition, it now has a rightsURI attribute, DataCite-style.
- content/type is now xs:token rather than a special restriction. The vocabulary is now managed as an external resource. Consequently, the vr:Type type vanishes from the XSD.
- content/level is now xs:token rather than a special restriction. The vocabulary is now managed as an external resource. Consequently, the vr:ContentLevel type vanishes from the XSD.
- New testQueryString child of interface.
- Now discouraging fixing standardID in VOResource extensions. Also, removed indication that capability/@xsitype should be used for service discovery.
- Now allowing a Z timezone marker in UTCTimestamp, indeed discouraging non-timezone use. This is in line with OAI-PMH use. Therefore, while technically changing the schema such that legacy clients might be confused, we do not expect incompatibilites.
- resource/description and capability/description are now xs:string to enable simple plain text markup.
- Adding language on using capability/@standardID rather than interface/@version to distinguish between different protocol versions in the normal case, removed dynamic example.
- Consequently, removing 1.0 default on interface/@version.
- Added advice about the use of creator.name
- Removed SOAP/WSDL example and a bit of the accompanying language.
- Section 3 needed some refactoring since the schema documentation is now generated from the schema document. To accommodate this, some text manually embedded within the schema documentation had to be moved out of the generated material or removed.
- Strongly advising the use of the vr: prefix, removing some duplicated advice regarding prefixes
- Removed example for deriving a SIA capability (this is now in the document repository)
- Ported document source to IVOATEX.

Fig. 4

What this is about

VOResource 1.1 is around the corner with $\frac{3}{2}$ pages of changelog.

RegTAP, as the common "user interface" to VOResource, has to catch up.

Fortunately, RegTAP doesn't notice most of the changes, as they mainly concern clarifications, vocabulary changes, type refinements, etc., which don't show in the database.

Background: To keep the schema manageable, we've traditionally done a manual mapping of VOResource to a relational model. In some extreme cases, we've even restricted what's expressible in RegTAP vs. what VOResource lets people do. This tactic has so far served us well.

So, this is about finding workable compromises.

Also, this is about providing implementation experience for VOResource RFC.

You're not supposed to read this:

(cf. Fig. 4)

3. New alt_identifier Table

ivoid	alt_identifier
ivo://ex/res1	orcid:0000-0000-0000-000X
ivo://ex/res1	orcid:0000-0000-0000-001V
ivo://ex/res1	doi:10.5072/ex/res1
ivo://other/q	doi:10.5072/hurgl

But: No way to work out what *role* the referenced entity plays wrt the resource.

On the other hand: Currently evident from alt_identifier type.

Alternative: Add role column?

4. New standard_key Table

ivoid	key_name	key_description
ivo://ivoa.net/std/regtap	table-1.0	The data model for...
ivo://gavo/std/example	vodml-prefix	foo
ivo://gavo/std/example	vodml-dmuri	http://.../vodml/foo-1
ivo://gavo/std/example	vodml-prefix	bar
ivo://gavo/std/example	vodml-dmuri	http://.../vodml/bar-1

StandardKeys come from StandardsRegExt and are intended for term enumerations coming with standards. These days, they're mostly used to tell various versions of a standard from each other (e.g., query-1.0 vs. query-1.1).

DM has a use case for discovery on StandardKeys. This table is a straight mapping, but the DM use case is a bit of a shortcut to their real problem.

Alternative: (a) Keep it unmapped, tell DM to define their own table. (b) Perhaps abusing capability?

5. Mirror URLs

In `rr.interface`:

primary key	...	mirror_url
ivo://id1, 1, 1	...	spiegel.de/svc#mirror.us/svc#spaijel.in/svc-m
ivo://id1, 1, 2	...	mirror.br/extra#extra.fr/extra
ivo://id2, 1, 1	...	NULL

Yes, array simulation sucks. But it's safe here, since there's no way # can be part of an access URL.

Alternatives: (a) Extra rr.mirrors table (yikes!); (b) support for arrays for var-length strings in VOTable.

6. interface/testQueryString

Again in `rr.interface`:

primary key	...	test_query_string
ivo://id1, 1, 1	...	REQUEST=doQuery&POS=23,1&SIZE=0.01
ivo://id1, 1, 2	...	MAXREC=1
ivo://id2, 1, 1	...	NULL

This is not yet implemented, as until a few weeks ago testQueryString in VOResource was 0...n, and then this wouldn't nearly be good enough. But there's no case for more than one testQueryString, I'd say.

Alternatives: None. Really.

7. rights, rightsURI

ivoid	detail_xpath	detail_value
ivo://.../hsoy/	/rights/@rightsURI	http://cc.org/publicdomain/zero/
ivo://.../hsoy/	/rights	Licensed under CC-0

RegTAP 1.0 has rr.resource.rights as hash-separated list.

VOResource 1.1 has free-text rights and rights/@rightsURI, so hash separation doesn't work any more.

Minor (because existing rights isn't terribly useful and thus hasn't been used to my knowledge) incompatible change: I've taken out resource.rights. rights and rightsURI now in res_details.

Alternatives: keep resource.rights but deprecate it and fill with NULL.

8. Mapping Vocabulary Terms

VOResource 1.1 has synonyms in order to keep old records valid but move to uniform term syntax.

date_role	
representative	Collected
creation	Created
update	Updated
relationship_type	
mirror-of	IsIdenticalTo
service-for	IsServiceFor
served-by	IsServedBy
derived-from	IsDerivedFrom

Proposal: Make RegTAP map to the *new* terms.

Advantage: Clients do not need to worry about vocabularies at all, and they can use new (DataCite) terms from the start exclusively.

Disadvantages: Incompatible change (but: there are probably very few clients actually using relationships and none using date roles at the moment). Also, mapping won't be good enough any more as soon as we do a bit more advanced semantics (hyponymy, "subclassing"). Then client smarts would be necessary anyway.

Alternatives: (a) Map to old terms? (b) Don't map at all?

9. Enabling Sensible Author Search

Right now:

```
select creator_seq
from rr.resource
natural join rr.res_role
where base_role='creator'
and role_name like 'Van der Waerden%'
```

So, you're supposed to take an (ordered) author string from `rr.resource`, and search within `rr.res_role`, where supposedly there's an expectable author form.

In reality, there's still horribly mess in creator.name, so it's questionable whether the fielded authors are helpful in the first place.

Recommend full-text index on `creator_seq` and matching with `ivo_hasword(creator_seq, 'Van der Waerden')`?

10. Kind Requests in Parting

- Ponder over changes and alternatives discussed here
- Use new VOResource features in your registry records
- Fix up your author names to be Last, F.I. (and perhaps list 2nd and further authors)

Thanks!