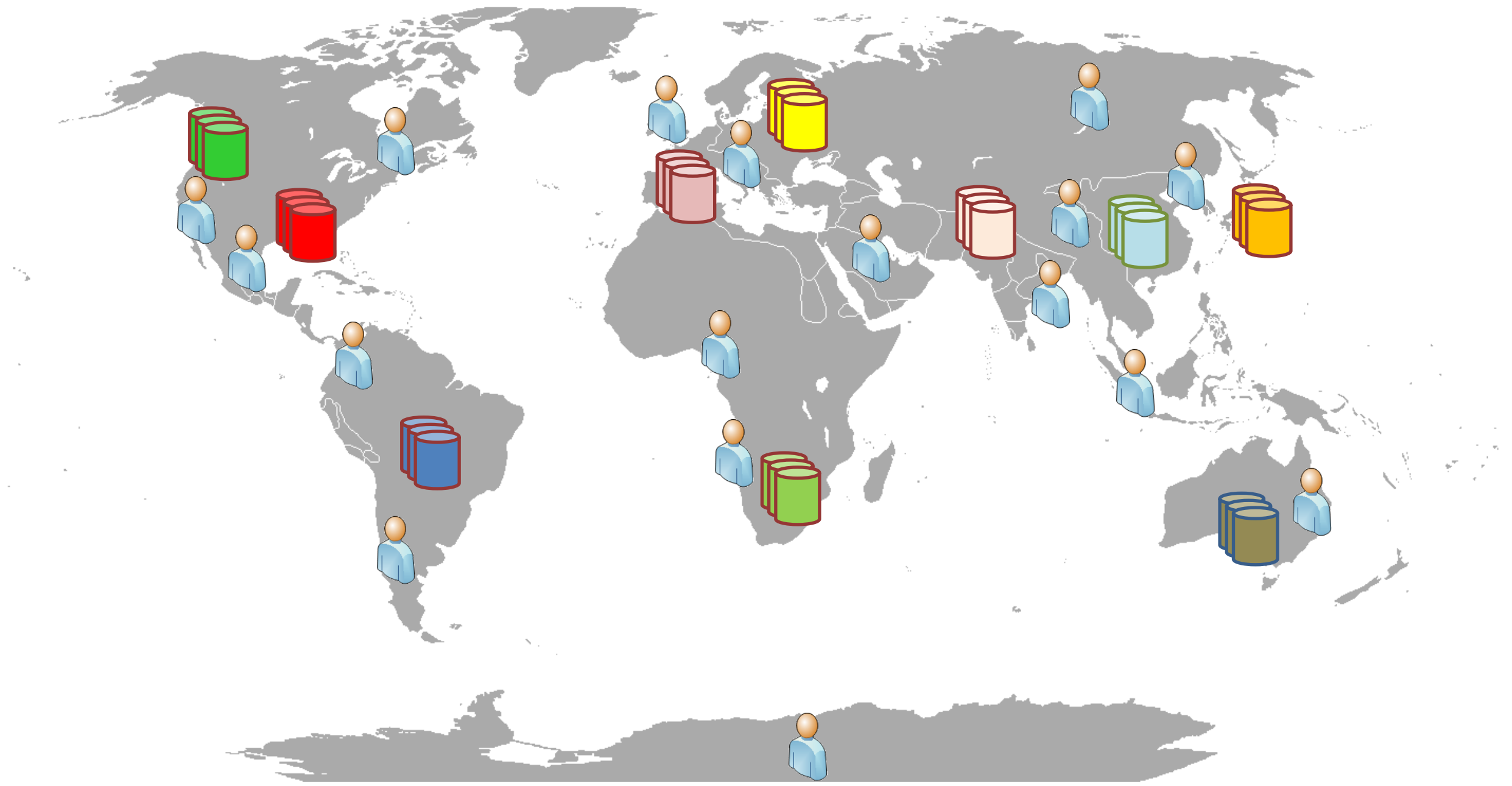


VO-DML and Mapping

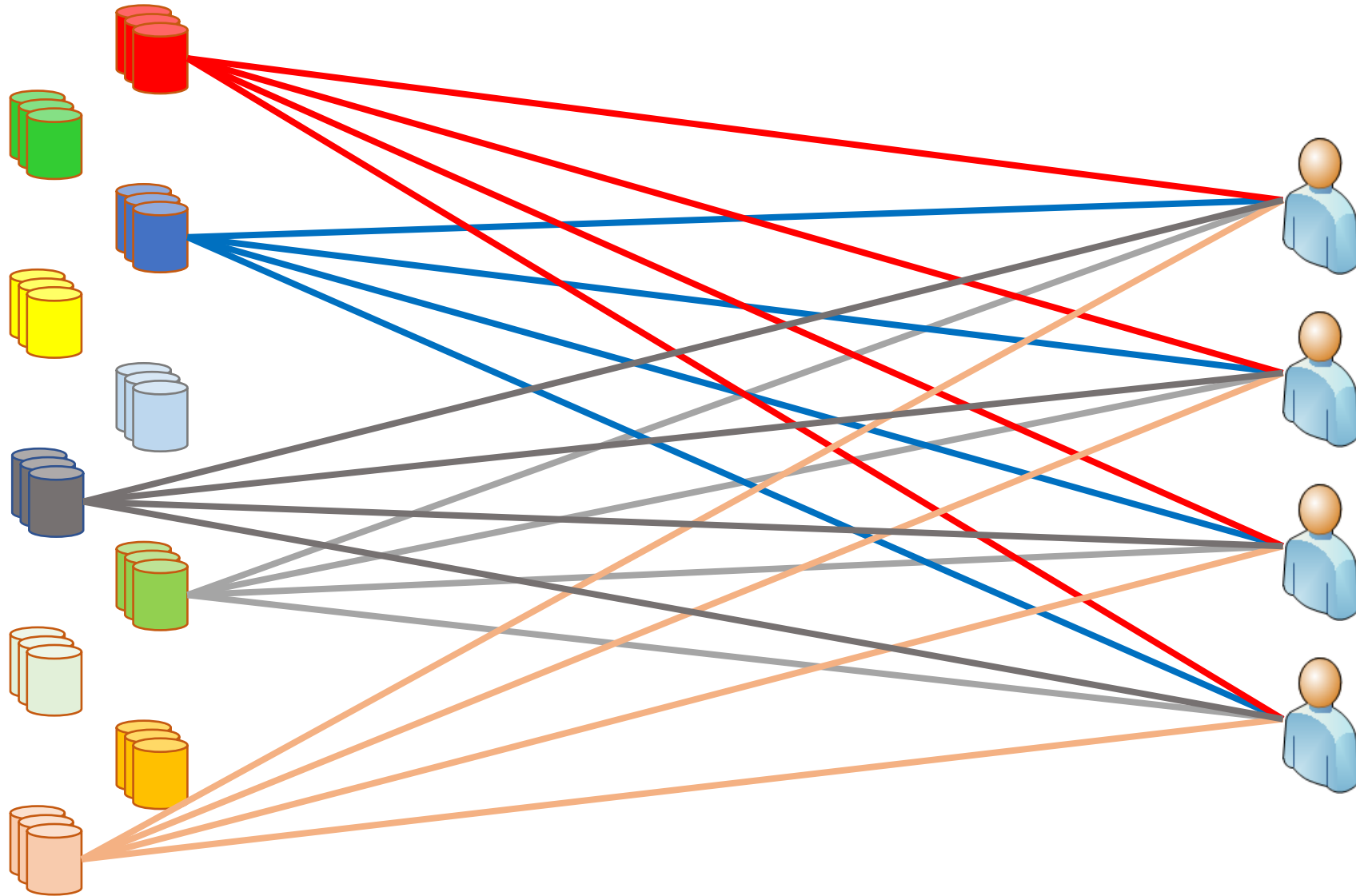
Data Integration

“... combining data residing in different sources and providing users with a unified view of these data.”

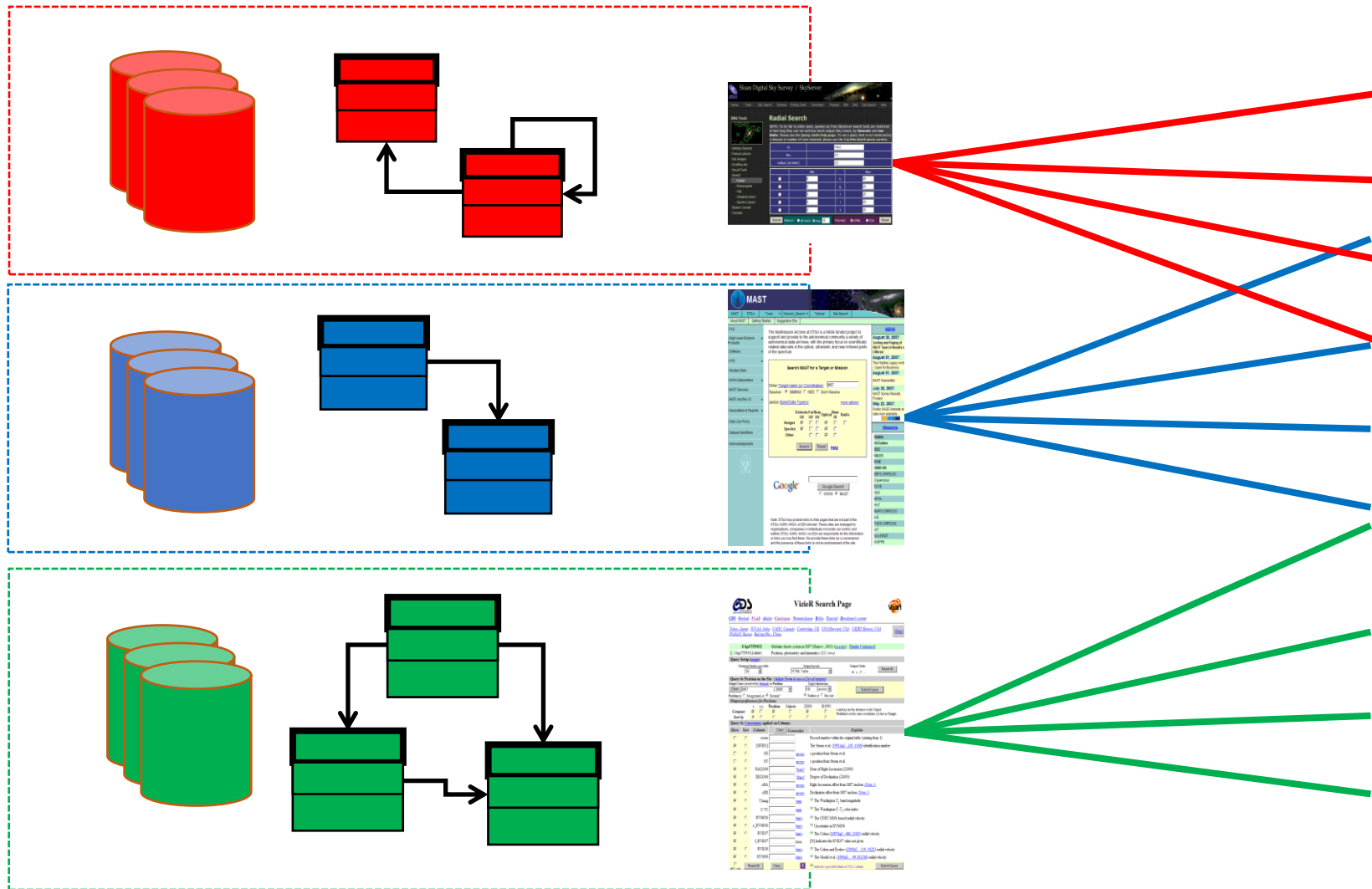
(wikipedia)



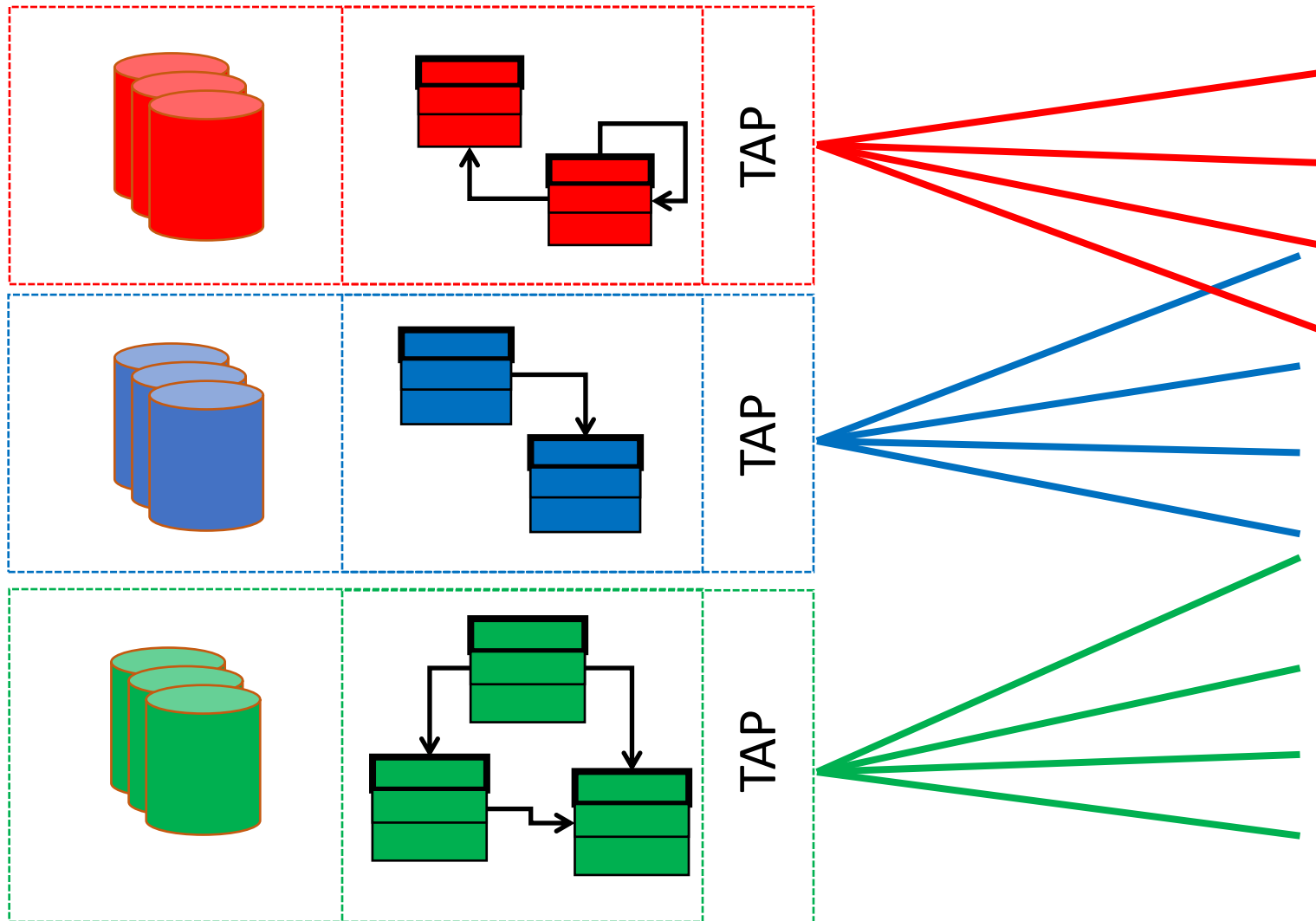
heterogeneity + scaling problem



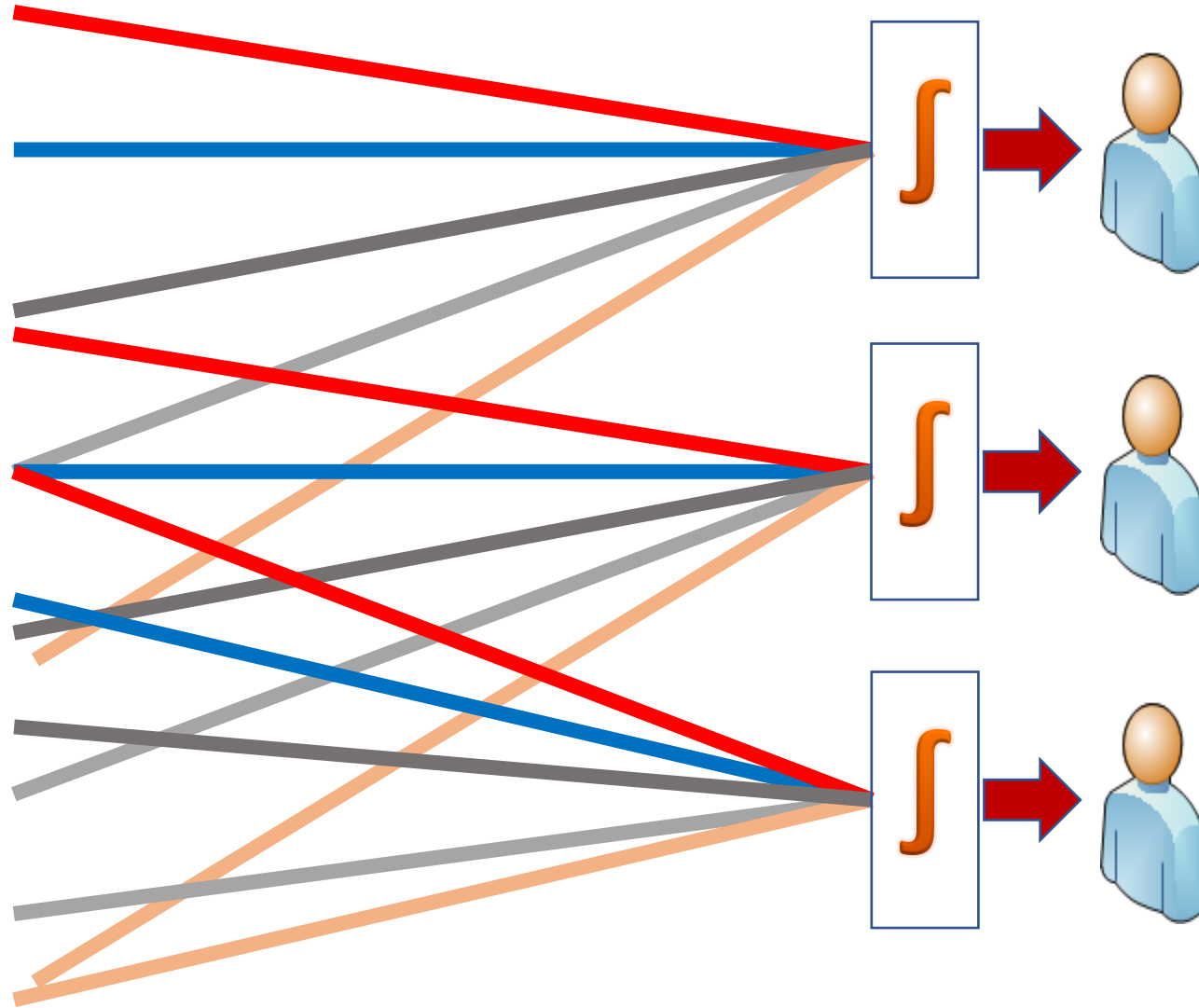
semantic: heterogeneous schemas
syntactic: custom access services



IVOA homogenizes syntax: e.g. TAP

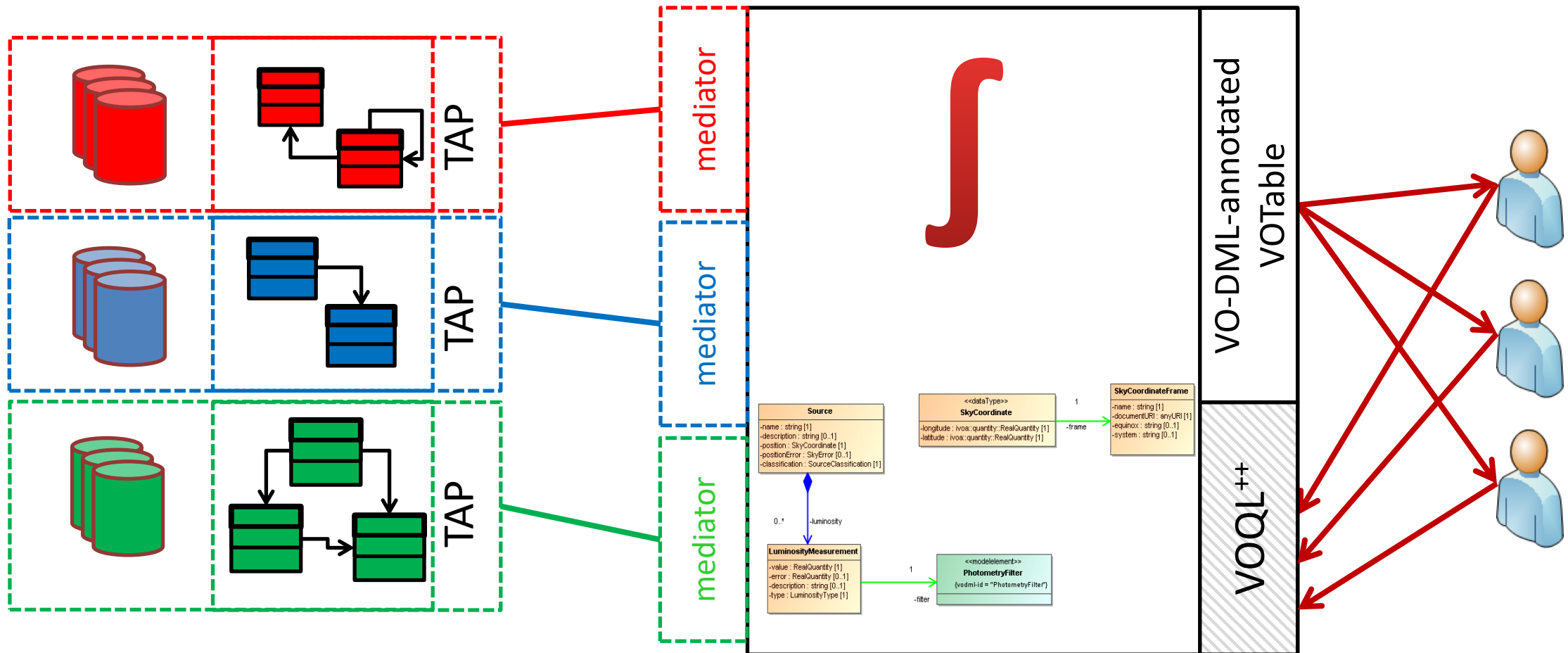


semantic heterogeneity requires individual data integration



Sketch of Integration Solution:

common (global) schema +
TAP + mediation + VOQL⁺⁺
 (see lots of CS literature)



Global Schema(s)

==

common data model(s)

- The *unified view* of the data sources
- Defined using VO-DML
 - VO-Data Modelling Language
 - very restricted subset of UML Class Diagrams
 - supports model dependency/reuse
 - simplified, XML serialization language: machine readability!
 - Faithful representations possible
 - XSD, Java, Python, JSON?, YAML?
 - VO-URP: RDB, TAP_SCHEMA
 - Accepted as IVOA standard

International **V**irtual **O**bservatory **A**lliance

IVOA Documents



VO-DML: a consistent modeling language for IVOA data models
Version 1.0

IVOA Recommendation 10 September 2018

Interest/Working Group:

<http://www.ivoa.net/twiki/bin/view/IVOA/IvoaDataModel>

Author(s):

Gerard Lemson, Omar Laurino, Laurent Bourges, Mark Cresitello-Dittmar, Markus Demleitner, Tom Donaldson, Patrick Dowler, Matthew Graham, Norman Gray, Laurent Michel, Jesus Salgado

Editor(s):

Gerard Lemson, Omar Laurino

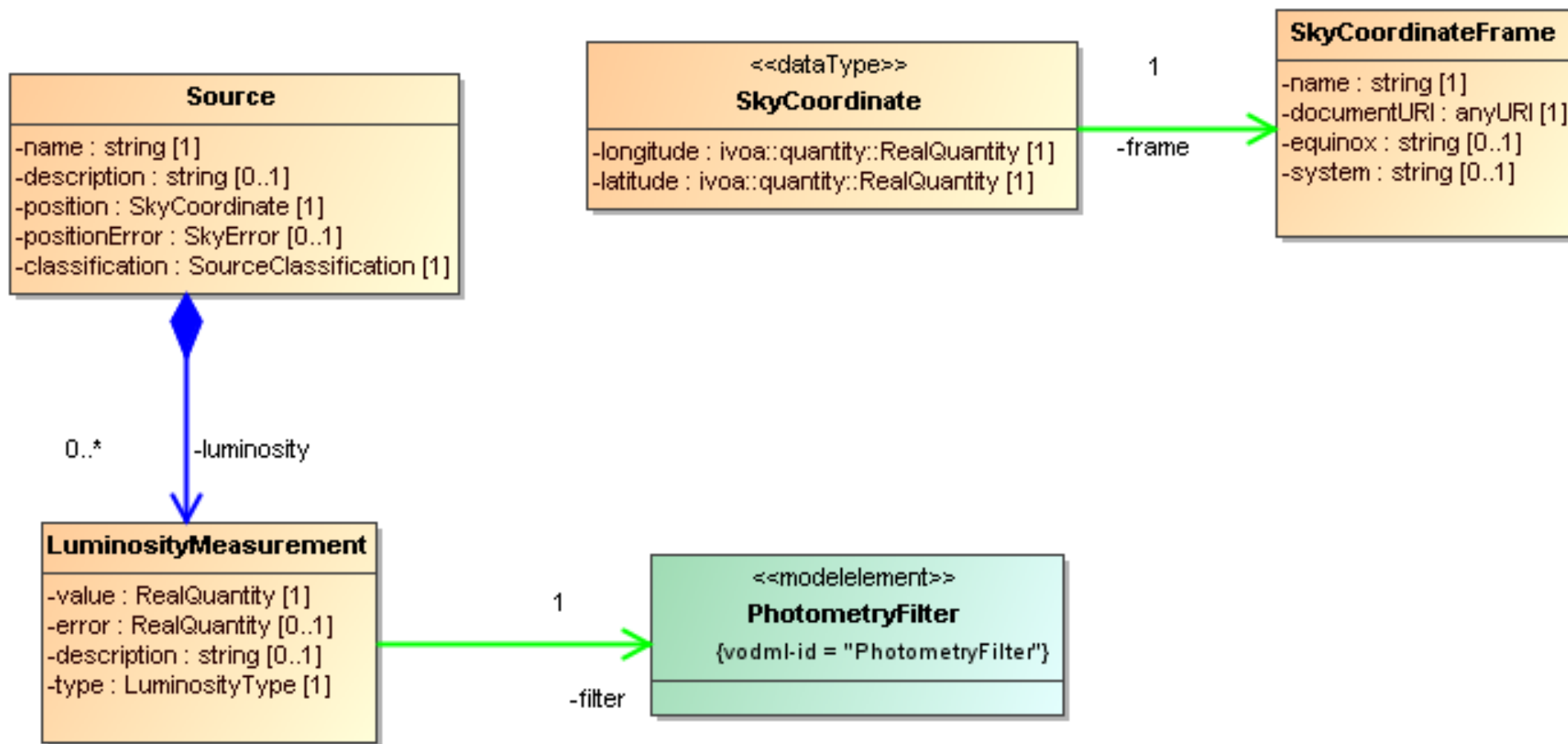
DOI:

10.5479/ADS/bib/2018ivoa.spec.0910L

(VO-DML) mapping

- Expresses how instances of a data model (expressed as VO-DML) are represented in a tabular representation
 - VOTable
 - TAP schema

Example: Simple source data model



objid	ra	dec	u	g	r	i	z	run	rerun	camcol	field	specobjid	class	redshift	plate	mjd	fiberid
1237680191504712292	319.42017295	-2.91605515	19.453272	17.512213	16.453272	16.453272	16.453272	5	45	4933710530549838848	GALAXY	0.091911	4382	55742	62		
1237680191504842797	319.6726666	-2.89320328	18.006258	18.408295	17.026258	17.026258	17.026258	5	47	4933700634945188864	STAR	-9.116632E-5	4382	55742	26		
1237660241388240997	51.95792979	0.44178806	17.90674	16.767498	16.247498	16.247498	16.247498	5	146	2329593891403098112	STAR	-2.435169E-4	2069	53376	389		
1237660241388371981	52.15864799	0.5100779	18.619341	17.314531	16.767498	16.767498	16.767498	5	148	2329602687496120320	STAR	-1.707261E-5	2069	53376	421		
1237660241925505040	52.72667419	0.88746662	17.930399	16.900446	16.418163	16.214106	16.1061	3438	301	6	152	2329600488472864768	STAR	-1.622409E-4	2069	53376	413
1237660241925505156	52.85661769	0.97756273	18.178764	16.997499	16.512629	16.314194	16.207306	3438	301	6	152	2329599663839143936	STAR	-1.234436E-4	2069	53376	410
1237662305111507089	202.55299093	39.86892911	17.820675	16.164869	15.296254	14.812856	14.419583	3919	301	1	16	5299625250001449984	GALAXY	0.048569	4707	55653	52
123766323879787997	52.05059022	0.14966321	19.351822	18.277271	18.06134	17.999191	17.999123	4136	301	4	165	2329595265792632832	STAR	-5.184785E-4	2069	53376	394
1237651271358108122	158.78373508	63.9613952	19.283352	17.41073	16.419657	16.042131	15.731997	1350	301	1	295	550602195343534080	GALAXY	0.11802	489	51930	135
1237651271358801125	158.82992158	63.94061555	19.297565	17.409573	16.431635	16.044048	15.707916	1350	301	1	295	550601370709813248	GALAXY	0.117888	489	51930	132

SDSS

```

Source
- name : string [1]
- description : string [0..1]
- position : SkyCoordinate [1]
- positionError : SkyError [0..1]
- classification : SourceClassification [1]

```

```

<<dataType>>
SkyCoordinate
- longitude : ivoa:quantity:RealQuantity [1]
- latitude : ivoa:quantity:RealQuantity [1]

```

```

CoordinateFrame
- name : string [1]
- documentURI : any
- equinox : string [0..1]
- system : string [0..1]

```

Identifying that a table contains Sources

[I/239/hip_main](#) The Hipparcos and Tycho Catalogues (ESA 1997)
[1 annotation\(s\)](#) - [post](#) The Hipparcos Main Catalogue (118218 rows)

Full	RAJ2000	DEJ2000	HIP	RAhms	DEdms	Vmag	RA(ICRS)	DE(ICRS)	Plx	pmRA	pmDE	e Plx	BTmag	e	VTmag	e	B-V	Hpmag	e
	"h:m:s"	"d:m:s"				mag	deg	deg	mas	mas/yr	mas/yr	mas	mag	mag	mag	mag	mag	mag	mag
1	00 00 00.216	+01 05 20.43	1	00 00 00.22	+01 05 20.4	9.10	0.00091185	1.08901332	3.54	-5.20	-1.88	1.39	9.643	0.020	9.130	0.019	0.482	9.2043	0.0020
2	00 00 01.024	-19 29 55.82	2	00 00 00.91	-19 29 55.8	9.27	0.00379737	-19.49883745	21.90	181.21	-0.93	3.10	10.519	0.033	9.378	0.021	0.999	9.4017	0.0017
3	00 00 01.206	+38 51 33.40	3	00 00 01.20	+38 51 33.4	6.61	0.00500795	38.85928608	2.81	5.24	-2.91	0.63	6.576	0.004	6.621	0.005	-0.019	6.6081	0.0007
4	00 00 02.071	-51 53 36.76	4	00 00 02.01	-51 53 36.8	8.06	0.00838170	-51.89354612	7.75	62.85	0.16	0.97	8.471	0.007	8.092	0.007	0.370	8.1498	0.0011
5	00 00 02.394	-40 35 28.33	5	00 00 02.39	-40 35 28.4	8.55	0.00996534	-40.59122440	2.87	2.53	9.07	1.11	9.693	0.014	8.656	0.010	0.902	8.7077	0.0018
6	00 00 04.486	+03 56 47.25	6														1.336	12.4488	0.0085
7	00 00 05.283	+20 02 10.01	7										5.542	0.039	9.679	0.030	0.740	9.6795	0.0021
8	00 00 06.562	+25 53 11.26	8										4.433	0.055	9.151	0.029	1.102	8.5522	0.1671
9	00 00 08.477	+36 35 09.45	9										9.962	0.025	8.711	0.015	1.067	8.7534	0.0018
10	00 00 08.740	-50 52 01.11	10										1.140	0.011	8.630	0.010	0.489	8.6994	0.0020
11	00 00 08.961	+46 56 23.99	11	00 00 08.95	+46 56 24.0	7.34	0.03729695	46.94000154	4.29	11.09	-2.02	0.84	7.446	0.005	7.364	0.005	0.081	7.3777	0.0010
12	00 00 09.816	-35 57 36.81	12	00 00 09.82	-35 57 36.8	8.43	0.04091756	-35.96022482	4.06	-5.99	-0.10	1.16	10.369	0.023	8.588	0.010	1.484	8.5598	0.0012
13	00 00 10.008	-22 35 40.94	13	00 00 10.00	-22 35 40.9	8.80	0.04167970	-22.59468060	3.49	8.45	-10.07	1.48	10.216	0.026	8.887	0.014	1.128	8.9707	0.0017

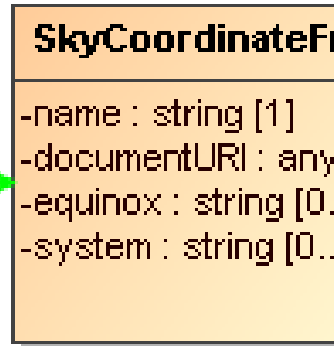
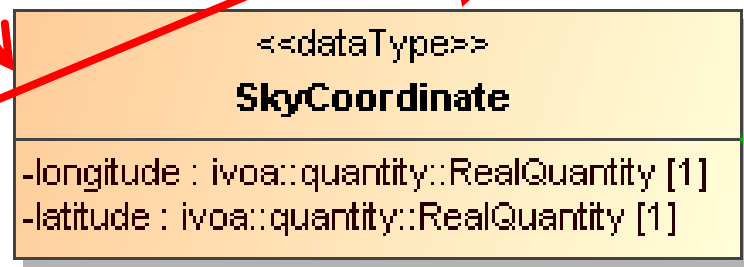
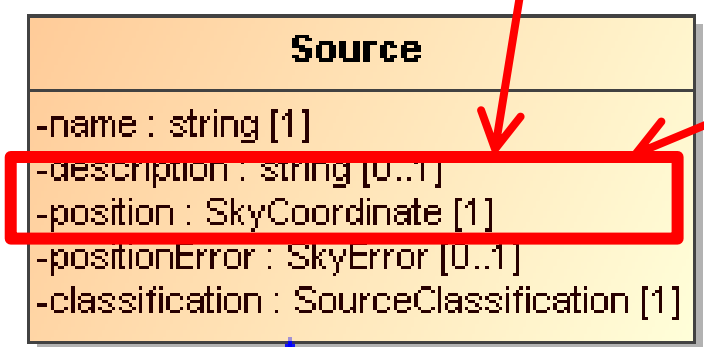
Hipparcos@VizieR

objid	ra	dec	u
1237680191504712292	319.42017295	-2.91605515	19.45
1237680191504842797	319.6726666	-2.89320328	18.00
1237660241388240997	51.95792979	0.44178806	17.90
1237660241388371981	52.15864799	0.5100779	18.61

I/239/hip_main

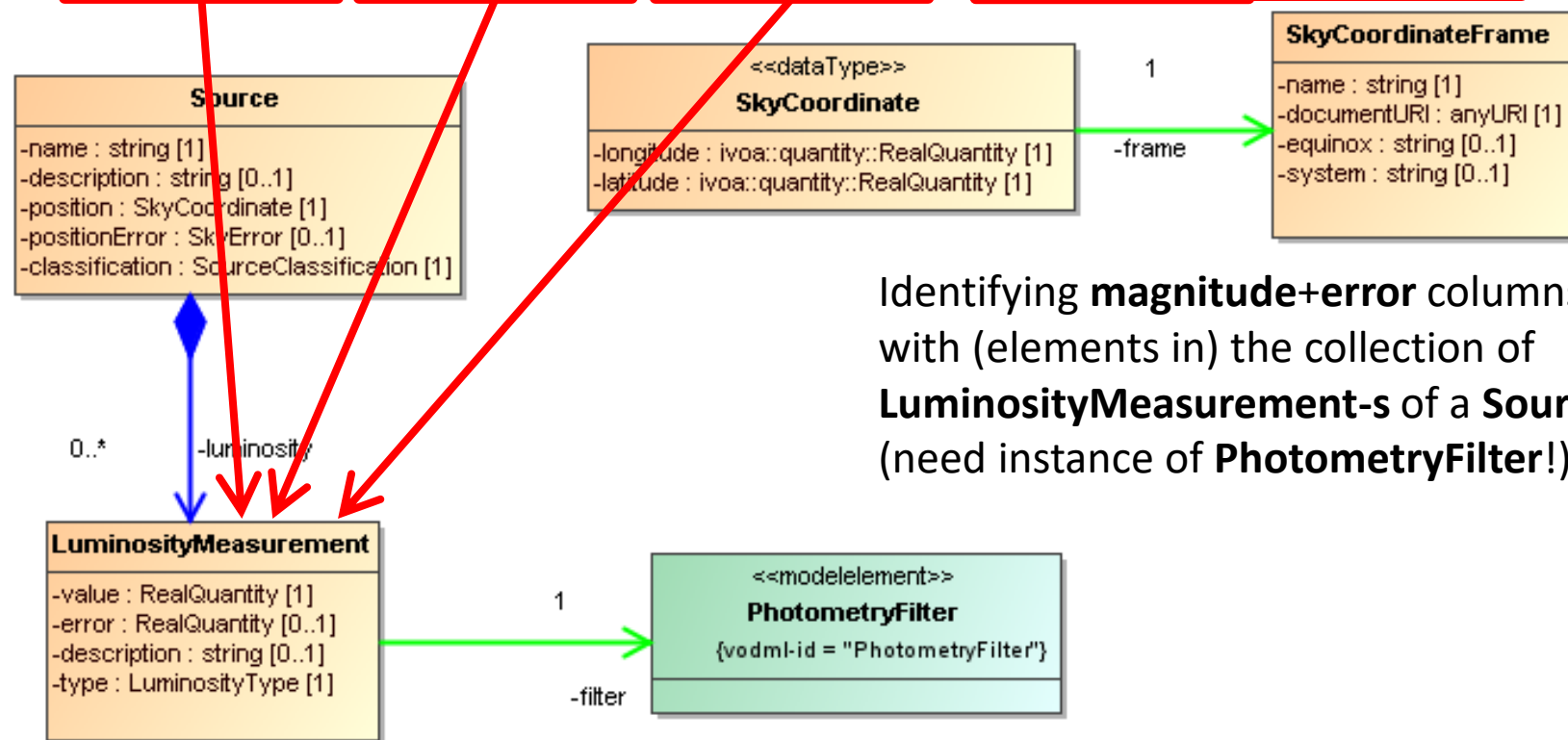
Annotation(s) - post

Full	RAJ2000	DEJ2000	HI
	"h:m:s"	"d:m:s"	
1	00 00 00.216	+01 05 20.43	
2	00 00 01.024	-19 29 55.82	
3	00 00 01.206	+38 51 33.40	
4	00 00 02.071	-51 53 36.76	
5	00 00 02.394	-40 35 28.33	



Identifying two columns as the **position** of a **Source**, a **SkyCoordinate**

ec	u	g	r	i	z	ru
505515	19.453272	17.512213	16.457823	15.911284	15.466995	808
820328	18.006258	18.408295	17.027491	16.594183	15.47355	808
78806	17.90674	16.767498	16.243202	16.038309	15.933592	343
00779	18.619341	17.314531	16.763399	16.53455	16.394312	343
46662	17.030300	16.000446	16.418163	16.214106	16.1061	343



Current state of mapping proposal

- OL, GL, TD, MC-D, LM etal
- Late PDF: http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-mapping/doc/VO-DML_mapping_WD.pdf
- XSD: <https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/xsd/ext/VODML-mapping.xsd>
- IN PROGRESS: Port to ivoatex <https://github.com/ivoa/mapping-vodml>

```
<VODML>
⊕ <MODEL>...
⊕ <GLOBALS>...
⊕ <TEMPLATES>...
</VODML>
```

```
<MODEL>
⊕ <NAME>
⊕ <URL>
⊕ <IDENTIFIER>
</MODEL>
```

```
<CONTAINER>
( <IDREF> | <REMOTEREERENCE> | <FOREIGNKEY> )
</CONTAINER>
```

```
<REFERENCE>
( <IDREF> | <REMOTEREERENCE> | <FOREIGNKEY> )...
</REFERENCE>
```

```
<GLOBALS>
⊕ <INSTANCE>...
</GLOBALS>
```

```
<INSTANCE>
⊕ <PRIMARYKEY>
⊕ <CONTAINER>
⊕ <ATTRIBUTE>...
⊕ <COMPOSITION>...
⊕ <REFERENCE>...
</INSTANCE>
```

```
<ATTRIBUTE>
↳ ( <COLUMN> | <CONSTANT> | <LITERAL> )...
↳ <INSTANCE>...
</ATTRIBUTE>
```

```
<TEMPLATES>
⊕ <INSTANCE>...
</TEMPLATES>
```

```
<COMPOSITION>
⊕ <INSTANCE>...
⊕ <EXTINSTANCES>...
</COMPOSITION>
```

```
<FOREIGNKEY>
⊕ <PKFIELD>...
⊕ <TARGETID>
</FOREIGNKEY>
```

```
<PRIMARYKEY>
⊕ <PKFIELD>...
</PRIMARYKEY>
```

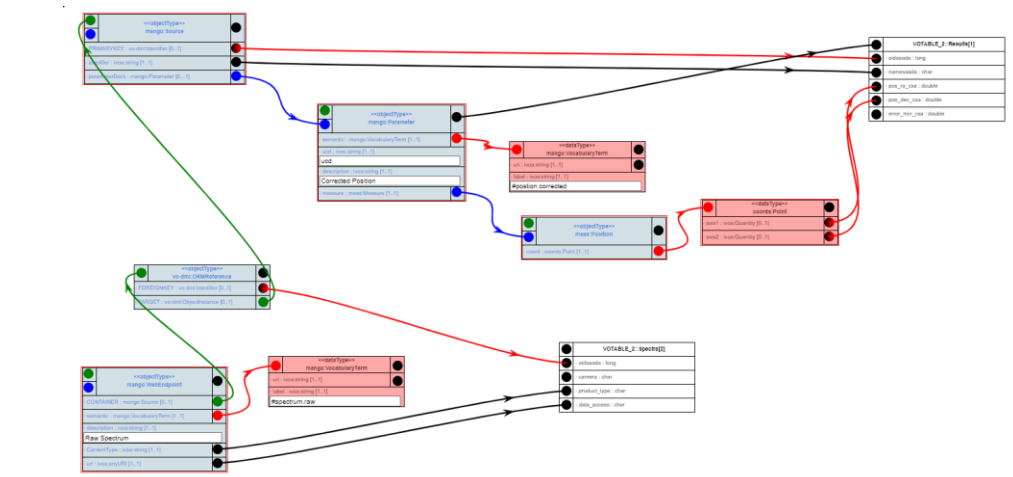
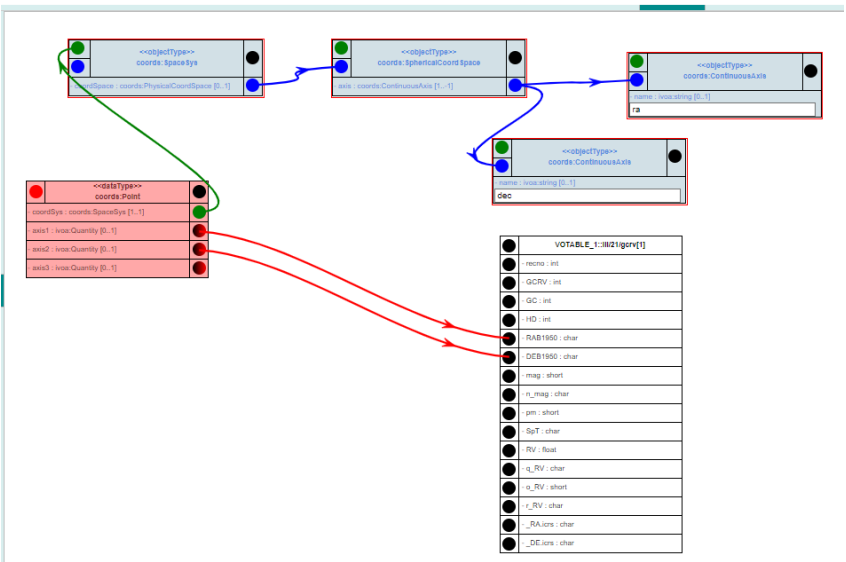
```
<PKFIELD>
( <COLUMN> | <CONSTANT> | <LITERAL> )
</PKFIELD>
```

```
<COLUMN>
<OPTIONMAPPING>...
</COLUMN>
```

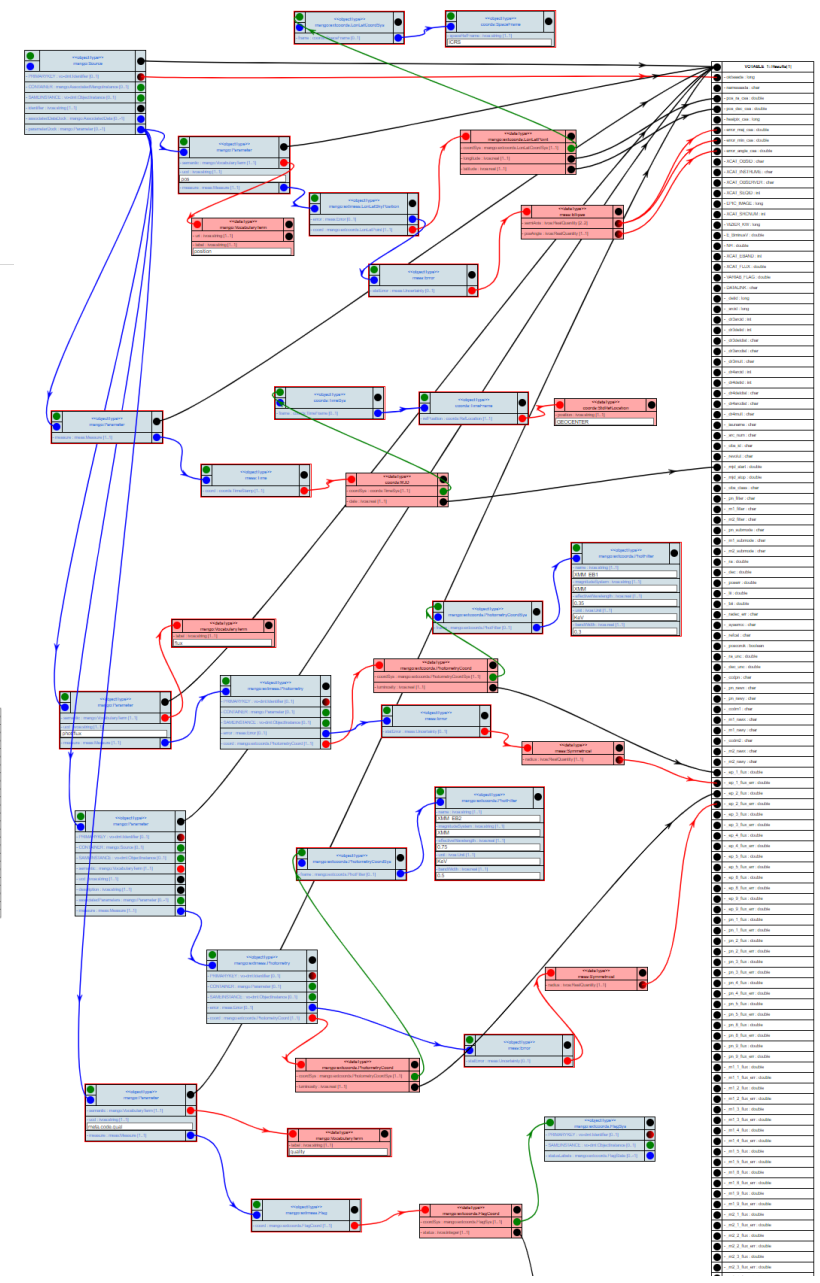
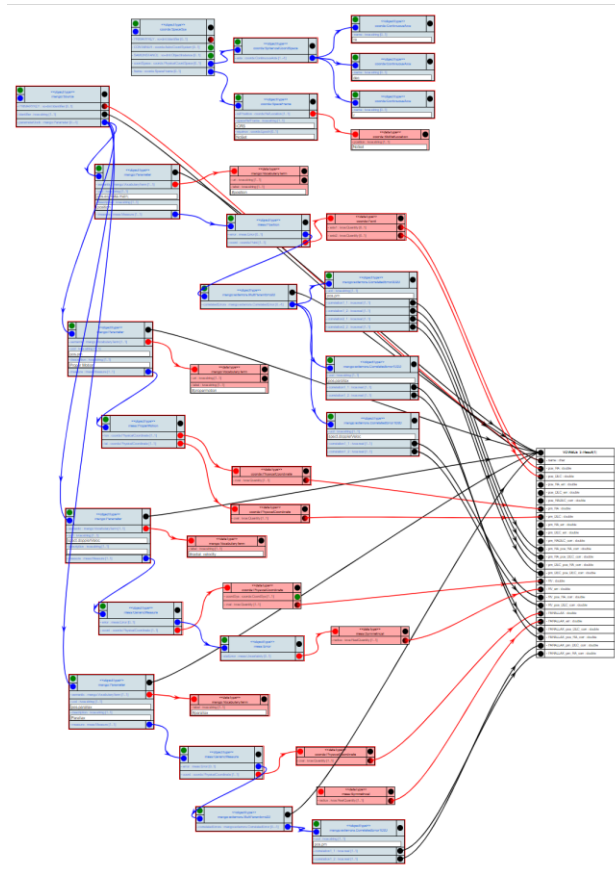
```
<CONSTANT>
<OPTIONMAPPING>...
</CONSTANT>
```

```
<LITERAL>
<OPTIONMAPPING>...
</LITERAL>
```

```
<OPTIONMAPPING>
⊕ <MAPPEDOPTION>
↳ <ENUMLITERAL>
↳ <SEMANTICCONCEPT>
</OPTIONMAPPING>
```

Use cases



How did we get here?

I: understanding UTYPE

- 2004: VOTable 1.1 introduces @utype: “Pointer into data model”
 - Ochsenbein, Bonnarel, Louys, Derriere, others discuss representing data models in VOTable
TODO add links
- ... - 2012: What is a UTYPE really?
 - Norman Gray:
<https://ivoa.net/documents/Notes/UTypes/utype-uri-20070302.html>
 - VO-URP for SimDM: utype is explicitly defined identifier of a (vo-urp) data model element
<https://ivoa.net/documents/SimDM/20120503/uml/intermediateModel.xsd>
- 2012-2013 : UTYPE Tiger Team (UTT) investigation
 - 1. There is no standard yet
<https://ivoa.net/documents/Notes/UTypesUsage/index.html>
 - 2. Question should be: how to express that a VOTable, a TAP_SCHEMA contains data model instances?
http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/doc/previously/VO-DML_and_UTILITY%20and%20VOTable-v0.2.pdf
- UTT: need something to point at: VO-DML
 - Need standardized representation of data model
 - Must have unique identifiers for all components
 - Almost identical to VO-URP representation of Simulation DM (2008-2012)
 - <http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/doc/previously/VO-DML-WD-v.0.x-20130416.pdf>
- UTT: need to define how to point: Mapping
 - Through annotations of EXISTING VOTable elements: GROUP, FIELDref, PARAM(ref)
 - UTYPE identifies the (VO-DML) data model element
 - <http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/doc/previously/UTYPES-WD-v.0.5-20130422.pdf>

II: VO-DML and Mapping

VODML:

- 2018: VO-DML becomes standard
 - <https://ivoa.net/documents/VODML/20180910>
 - used in various models since then →→→→→→→→→→
 - VODSL (PH)

Mapping VO-DML:

- 2013-2014: Mapping schema extensions:
 - None. Just use @utype and GROUP hierarchy
 - Madrid: don't use @utype, use @dmrole, @dmtype or so
 - Banff: use <VODML> instead
(<http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/doc/samples/votable/>)
- 2015 - 2016:
 - Write document
 - Realize need special elements to identify serialization-specific elements. E.g. for identifiers, containers, references, ORM
 - Tools: jovial, VODML-Mapper
- 2017: TOO COMPLEX !
- 2017- Present
 - Expansion of VOTable schema adding **explicit** components almost 1-1 with VO-DML:
<https://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-org/xsd/ext/VODML-mapping.xsd>
http://volute.g-vo.org/svn/trunk/projects/dm/vo-dml-mapping/doc/VO-DML_mapping_WD.pdf
- 2018: TOO BIG!
 - Alternative (LM) by now as big, maybe even bigger, and very similar in any case

VO DML	VO Data Model Language	
└_schema	VODML schema	v1.0
└_schematron	VODML schematron	v1.0
└_models		
└_ivoa	IVOA	v1.0
└_provenance	provenance	v1.0
└_coords	Astronomical Coordinates and Coordinate Systems	v1.0
└_meas	Astronomical Measurements Model	v1.0
└_ssldm	ssldm	v2.0
└_Trans	WCS transform model	v1.0

The end