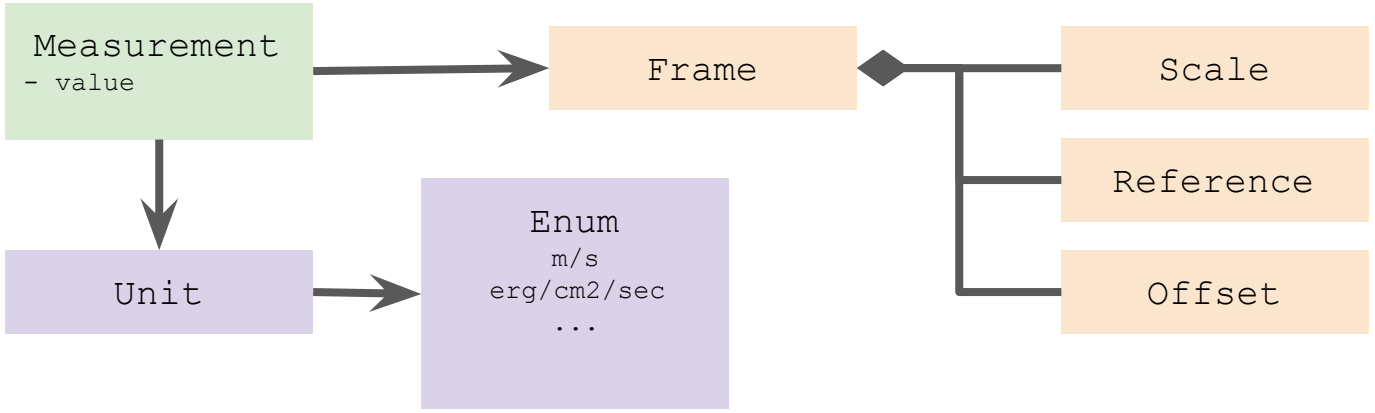
Model in the VO (Part1)

L. Michel

What Are Data Model For

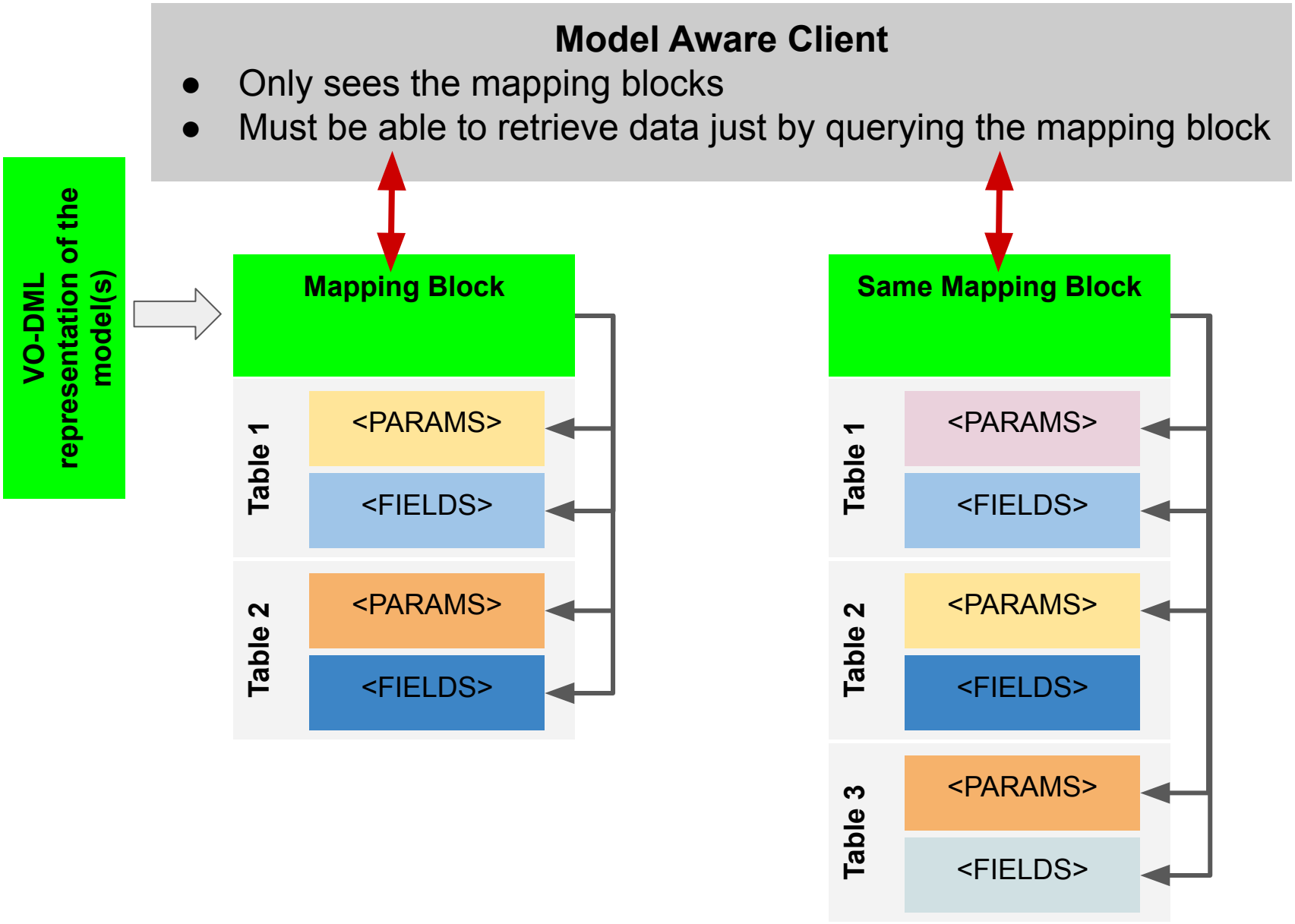
- **Formal** description of the quantities used by the experts in a domain
- What does the human knowledge (common sense) say:
  - A **`measurement`** is a value with a given **`unit`** that is valid in a given **`frame`**
- The Model gives a formal representation of that knowledge
  - The model defines the quantity classes, the names, the vocabulary and ….
    …. The relationships between those elements



What Are VO Data Models Used For

- **Documentation**
  - Developer guideline
    - Developers work with the DM standard on the table
    - Client, server, validator
  - DAL protocol design
    - Designing protocol where data responses are retrospectively compliant with a model

- **Interoperability** 
  - **Different data mapped on the same model can be combined or compared to each other**
    - **Data discovery (Obscore)**
    - **Stacked plots**
    - **Cross-match**


Model Aware Client
Only sees the mapping blocks
Must be able to retrieve data just by querying the mapping block
VO-DML representation of the model(s)
Mapping Block
Table 1
<PARAMS>
<FIELDS>
Table 2
<PARAMS>
<FIELDS>
Same Mapping Block
Table 1
<PARAMS>
<FIELDS>
Table 2
<PARAMS>
<FIELDS>
Table 3
<PARAMS>
<FIELDS>

The DM workshop

- **Data annotation with UTypes**
  - Data elements refer to model leaves
  - Data response kinked to models by a key mechanisim
- **Data annotation with VO-DML mapping**
  - Data response comes with a whole description of the model they refer to.
  - The client has enough material to build model instance from the data.
- **Embedding models in client code**
  - Client code is enable to interpret (properly display e.g.) data just by analysing a model

works well

a bit stuck for now

rather a dream

Chicken-Egg DM Deadlock

Let's consider we have working proposals for both models and annotation

Client developer

I'll support model mapping **when data providers** will annotate their data

Data provider

I'll annotate my data when **client developers** will be able to process annotated data

All Have Good Arguments

Client developer

What could I do with the models that I can't do now?

Dataset will never be all annotated; I'll ever have to process *raw* dataset

What happens when a model is updated?

Data provider

Annotating data is a big job for my service, what is the gain?

What happens when a model is updated?

All Have Good Arguments

What could I do with the models that I can't do now?

- Get a clean representation of the coordinate systems
- Get a clean representation of the errors
- Support cross-columns parameters
  - Columns grouping
  - Complex errors (pos + pm + parallax)
- Gather data from multi-table VOTable (sources + detection)
- Exchange model instance with SAMP

DM Workshop

DM Working group committed by the TCG to run a workshop to set a consensus that can get the  **ack on the road**

FEBRUARY 2021: use case workbench setup

- Use-case: raw data
- Proposal: Annotated VOTable + documentation
- Issues + Wiki

MARCH-APRIL 2021: open contributions

- Exercice
- Comment
- Proposals

APRIL 2021: Virtual workshop

MAY 2021: Report at Interop

Taken Positions

Un-entangled models (Markus Demleitner)
- A mapping block on the top of the VOTable
- Contains sparse model components not gathered in a global model

Works fine for the simple case

Product models (Mark Cresitello Dittmar)
- A mapping block on the top of the VOTable
- Maps the data on a product model (e.g. Time Series)
- The product model is made of model component classes

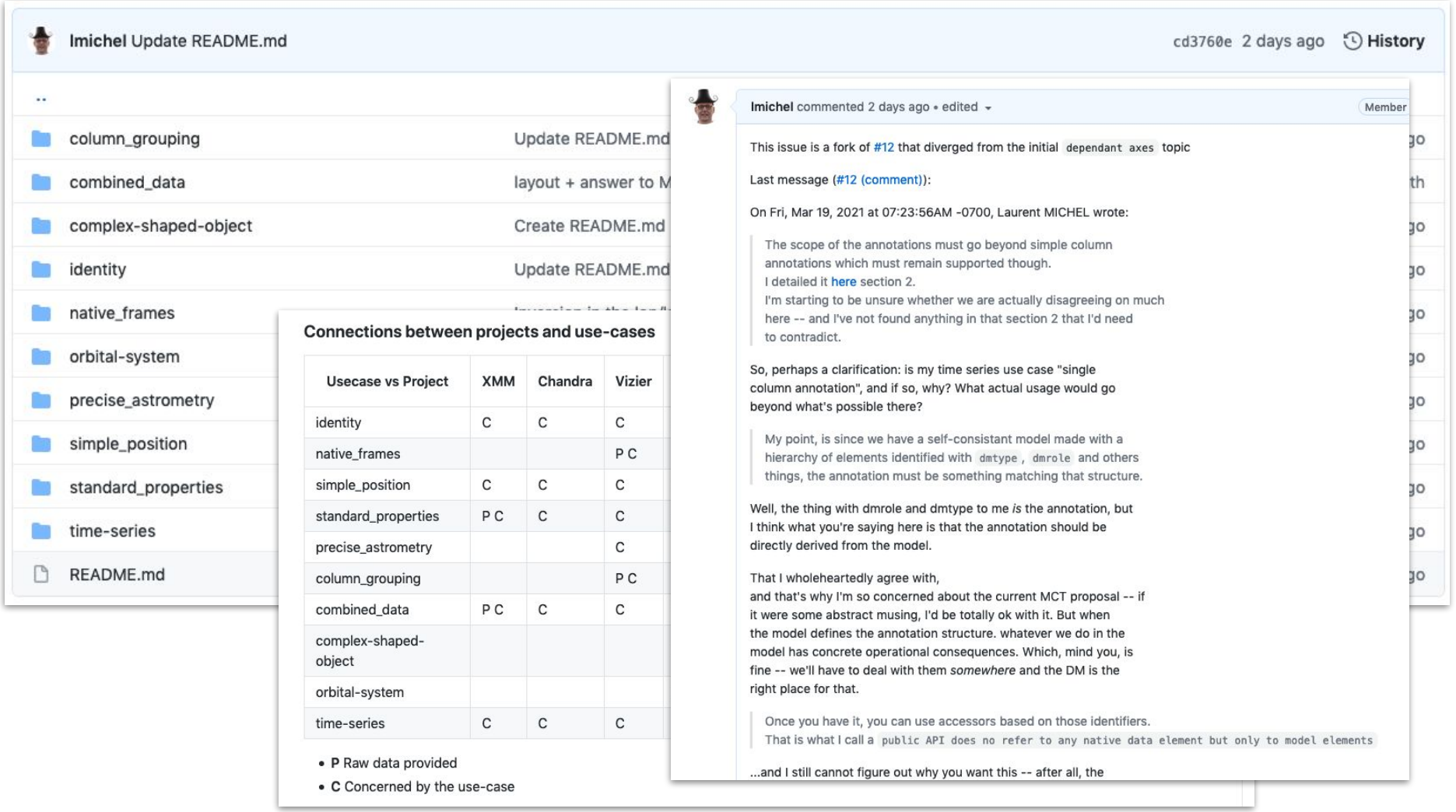
Common interface for any products of a given type

In between (Laurent Michel et al.)
- A mapping block on the top of the VOTable
- Contains sparse model components arranged in a container (MANGO)
- Support both parameters and associated data

Suited for archival data even with complex features

Please Contribute

https://github.com/ivoa/dm-usecases



lmichel Update README.md — cd3760e 2 days ago — History

column_grouping	Update README.md
combined_data	layout + answer to M
complex-shaped-object	Create README.md
identity	Update README.md
native_frames	
orbital-system	
precise_astrometry	
simple_position	
standard_properties	
time-series	
README.md	

Connections between projects and use-cases

Usecase vs Project	XMM	Chandra	Vizier
identity	C	C	C
native_frames			P C
simple_position	C	C	C
standard_properties	P C	C	C
precise_astrometry			C
column_grouping			P C
combined_data	P C	C	C
complex-shaped-object			
orbital-system			
time-series	C	C	C

- **P** Raw data provided
- **C** Concerned by the use-case

lmichel commented 2 days ago • edited — Member

This issue is a fork of #12 that diverged from the initial `dependant axes` topic

Last message (#12 (comment)):

On Fri, Mar 19, 2021 at 07:23:56AM -0700, Laurent MICHEL wrote:

> The scope of the annotations must go beyond simple column annotations which must remain supported though.
> I detailed it here section 2.
> I'm starting to be unsure whether we are actually disagreeing on much here -- and I've not found anything in that section 2 that I'd need to contradict.

So, perhaps a clarification: is my time series use case "single column annotation", and if so, why? What actual usage would go beyond what's possible there?

> My point, is since we have a self-consistant model made with a hierarchy of elements identified with `dmtype`, `dmrole` and others things, the annotation must be something matching that structure.

Well, the thing with dmrole and dmtype to me *is* the annotation, but I think what you're saying here is that the annotation should be directly derived from the model.

That I wholeheartedly agree with,
and that's why I'm so concerned about the current MCT proposal -- if it were some abstract musing, I'd be totally ok with it. But when the model defines the annotation structure. whatever we do in the model has concrete operational consequences. Which, mind you, is fine -- we'll have to deal with them *somewhere* and the DM is the right place for that.

> Once you have it, you can use accessors based on those identifiers.
> That is what I call a `public API does no refer to any native data element but only to model elements`

...and I still cannot figure out why you want this -- after all, the

Model in the VO (Part2)

Model for ANnotating Generic Objects

L. Michel F. bonnarel M. Molinaro M Louys J. Salgado G.Landais

Motivation for a Source Model

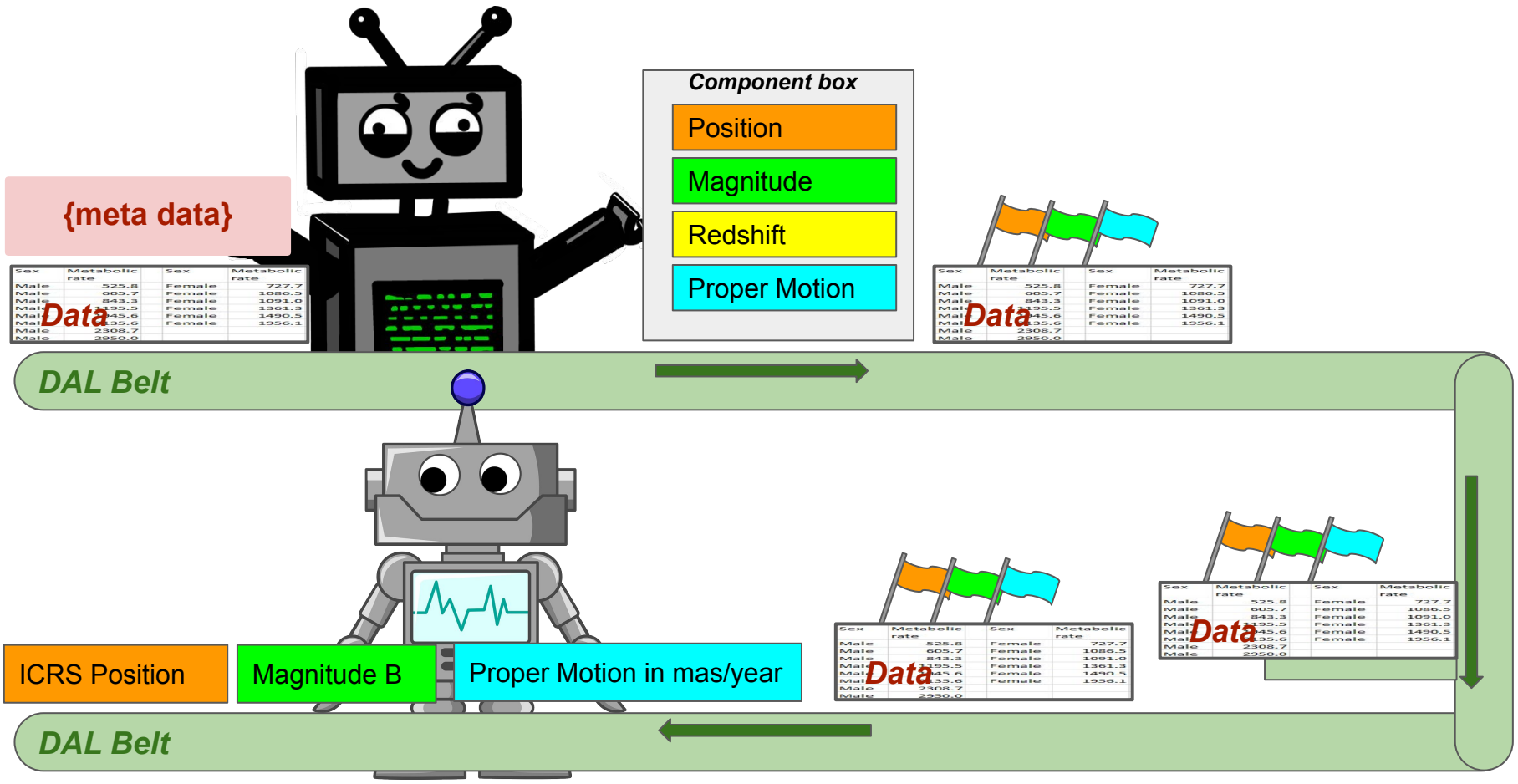
J/A+A/532/A103/IC4665 Photometry and proper motions in IC4665 (Lodieu+, 2011) 2011A&A
Post annotation
Coordinates, photometry, proper motions, and physical parameters for IC4665 stars *(tables A1, B1 and C1 of paper) (1533 rows)*

start AladinLite plot the output query using TAP/SQL

Full	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"	Mm	Name	Zmag mag	Ymag mag	Jmag mag	Hmag mag	Kmag mag	pmRA mas/yr	pmDE mas/yr	logL [Lsun]	Mass Msun	GCS	Simbad	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"
1	17 42 05.9300	+05 24 13.900	C	174205.93+052413.9	15.113	14.535	13.849	13.200	12.831	-1.80	-19.56	-1.29	0.448	GCS	Simbad	17 42 05.93	+05 24 13.9
2	17 42 06.0200	+05 14 17.900	C	174206.02+051417.9	15.720	15.304	14.737	14.173	13.850	-56.48	24.63	-1.67	0.241	GCS	Simbad	17 42 06.02	+05 14 17.9
3	17 42 09.5800	+05 21 12.600	C	174209.58+052112.6	15.693	15.200	14.587	13.944	13.615	-36.93	-34.57	-1.61	0.267	GCS	Simbad	17 42 09.58	+05 21 12.6
4	17 42 16.9500	+05 26 51.300	C	174216.95+052651.3	15.670	15.269	14.676	14.082	13.761	-3.44	3.13	-1.65	0.251	GCS	Simbad	17 42 16.95	+05 26 51.3
5	17 42 17.7800	+05 56 26.200	C	174217.78+055626.2	16.616	16.024	15.344	14.767	14.377	-25.95	-12.29	-1.93	0.160	GCS	Simbad	17 42 17.78	+05 56 26.2
6	17 42 18.0000	+05 49 25.500	C	174218.00+054925.5	14.923	14.523	13.970	13.353	13.095	-8.36	4.92	-1.34	0.412	GCS	Simbad	17 42 18.00	+05 49 25.5
7	17 42 18.1900	+05 53 53.300	C	174218.19+055353.3	17.047	16.495	15.845	15.246	14.900	17.51	-8.99	-2.14	0.115	GCS	Simbad	17 42 18.19	+05 53 53.3
8	17 42 20.2900	+05 55 56.500	C	174220.29+055556.5	14.734	14.267	13.654	13.059	12.740	8.68	2.01	-1.21	0.508	GCS	Simbad	17 42 20.29	+05 55 56.5
9	17 42 20.7900	+05 46 35.600	C	174220.79+054635.6	15.508	15.061	14.449	13.906	13.586	-27.37	-2.79	-1.55	0.293	GCS	Simbad	17 42 20.79	+05 46 35.6
10	17 42 21.0800	+05 43 13.900	C	174221.08+054313.9	17.377	16.697	15.988	15.443	15.087	6.06	33.92	-2.20	0.104	GCS	Simbad	17 42 21.08	+05 43 13.9
11	17 42 23.5500	+05 38 23.500	C	174223.55+053823.5	15.344	14.881	14.291	13.705	13.403	1.01	-31.80	-1.48	0.328	GCS	Simbad	17 42 23.55	+05 38 23.5
12	17 42 24.8900	+05 06 06.100	C	174224.89+050606.1	16.389	15.827	15.163	14.587	14.272	0.98	-0.62	-1.85	0.181	GCS	Simbad	17 42 24.89	+05 06 06.1
13	17 42 25.4100	+06 21 05.300	C	174225.41+062105.3	14.731	14.381	13.831	13.177	12.885	-4.17	4.60	-1.29	0.453	GCS	Simbad	17 42 25.41	+06 21 05.3
14	17 42 25.6900	+05 29 47.200	C	174225.69+052947.2	16.944	16.476	15.847	15.208	14.889	10.33	10.36	-2.14	0.114	GCS	Simbad	17 42 25.69	+05 29 47.2
15	17 42 26.6000	+06 22 19.800	C	174226.60+062219.8	14.758	14.378	13.873	13.210	12.946	0.05	-6.42	-1.30	0.440	GCS	Simbad	17 42 26.60	+06 22 19.8
16	17 42 26.9300	+06 20 14.600	C	174226.93+062014.6	16.782	16.238	15.635	15.019	14.640	-4.73	11.58	-2.05	0.132	GCS	Simbad	17 42 26.93	+06 20 14.6
17	17 42 28.0300	+05 26 40.700	C	174228.03+052640.7	14.683	14.311	13.740	13.129	12.833	4.01	21.98	-1.25	0.481	GCS	Simbad	17 42 28.03	+05 26 40.7
18	17 42 28.9300	+05 54 53.800	C	174228.93+055453.8	17.691	17.048	16.427	15.888	15.495	28.31	79.04	-2.39	0.078	GCS	Simbad	17 42 28.93	+05 54 53.8
19	17 42 28.9400	+06 20 28.000	C	174228.94+062028.0	15.234	14.810	14.260	13.592	13.286	2.95	7.72	-1.47	0.336	GCS	Simbad	17 42 28.94	+06 20 28.0
20	17 42 31.9100	+06 18 49.500	C	174231.91+061849.5	14.560	14.139	13.606	12.988	12.679	-41.07	0.18	-1.19	0.524	GCS	Simbad	17 42 31.91	+06 18 49.5

How could a client process or even plot Position/Mags/PM/Mass entries of this catalog **without taking into consideration it comes from Vizier**?

2 Stakeholders: Data Provider/Consumer:



The model is so discreet in this diagram that one may wonder if it exists

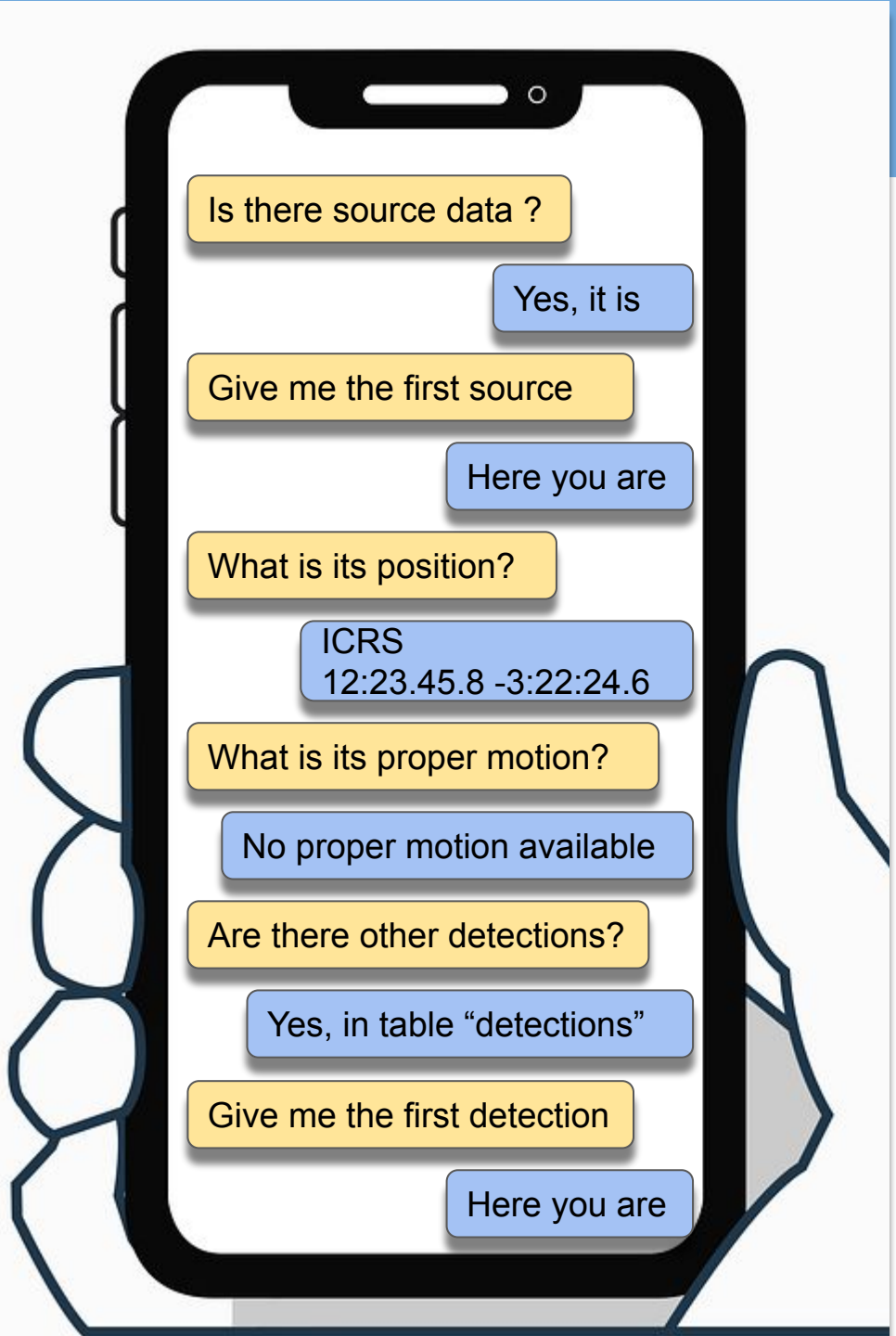
Client Perspecti

SourceDM — -bash — 80×24

```
MacBook-Pro-de-Laurent-MICHEL:SourceDM laurentmichel$
```

```
parser = Parser("My VoTable")
sources = parser.getInstanceSet("CABMSD");

while( sources.hasNext())
    source = sources.next()
    print(source.get("position"))
    print(source.get("mag.G"))
```

```
MacBook-Pro-de-Laurent-MICHEL:SourceDM laurentmichel$
```



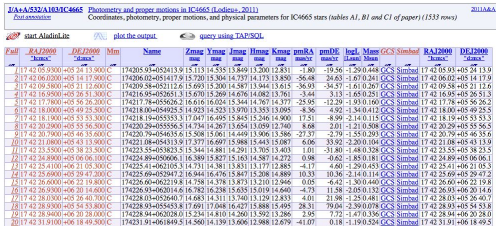
MANGO Guideline

We have to consider:

1. **The annotation content**
   a. Data modeling
   b. Serialization
2. **The data annotation process**
   a. Data provider point of view
   b. No hope to use the model as long as no data provider implements it
3. **The annotated data processing**
   a. Client developer point of view

All of these are parts of the MANGO project

The Model: Object Types and Params



2/A+A/532/A103/IC4665 Photometry and proper motions in IC4665 (Lodieu+, 2011) 2011A&A

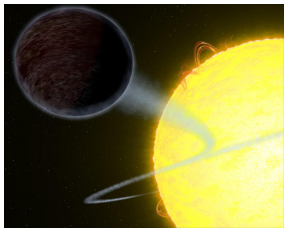
Post annotation

Coordinates, photometry, proper motions, and physical parameters for IC4665 stars (*tables A1, B1 and C1 of paper*) (1533 rows)

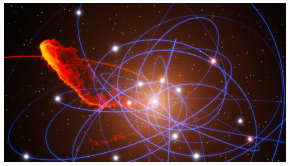
start AladinLite plot the output query using TAP/SQL

Full	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"	Mm	Name	Zmag mag	Ymag mag	Jmag mag	Hmag mag	Kmag mag	pmRA mas/yr	pmDE mas/yr	logL [Lsun]	Mass Msun	GCS Simbad	RAJ2000 "h:m:s"	DEJ2000 "d:m:s"
J	17 42 05.930	+05 24 13.900	C	174205.93+052413.9	15.713	14.535	13.849	13.200	12.831	-1.80	-19.56	-1.29	0.448	GCS Simbad	17 42 05.93	+05 24 13.9
J	17 42 06.020	+05 14 17.900	C	174206.02+051417.9	15.720	15.304	14.737	14.173	13.850	-56.68	26.63	-1.67	0.241	GCS Simbad	17 42 06.02	+05 14 17.9
J	17 42 09.580	+05 21 12.600	C	174209.58+052112.6	15.693	15.200	14.587	13.944	13.615	-36.93	-36.57	-1.61	0.267	GCS Simbad	17 42 09.58	+05 21 12.6
J	17 42 16.950	+05 26 51.300	C	174216.95+052651.3	15.670	15.269	14.676	14.082	13.761	-3.44	3.13	-1.65	0.251	GCS Simbad	17 42 16.95	+05 26 51.3
J	17 42 17.780	+05 56 26.300	C	174217.78+055626.2	16.616	16.024	15.344	14.767	14.377	-25.95	-12.29	-1.93	0.160	GCS Simbad	17 42 17.78	+05 56 26.2
J	17 42 18.000	+05 49 25.500	C	174218.00+054925.5	14.923	14.523	13.970	13.355	13.095	-8.36	4.92	-1.34	0.412	GCS Simbad	17 42 18.00	+05 49 25.5
J	17 42 18.130	+05 53 53.300	C	174218.19+055353.3	17.047	16.495	15.845	15.246	14.900	17.51	-8.99	-2.14	0.115	GCS Simbad	17 42 18.19	+05 53 53.3
J	17 42 20.290	+05 55 56.500	C	174220.29+055556.5	14.734	14.267	13.654	13.059	12.740	8.68	2.01	-1.21	0.508	GCS Simbad	17 42 20.29	+05 55 56.5
J	17 42 20.790	+05 46 35.600	C	174220.79+054635.6	15.508	15.063	14.449	13.906	13.586	-27.37	-2.79	-1.55	0.293	GCS Simbad	17 42 20.79	+05 46 35.6
J	17 42 21.080	+05 43 13.900	C	174221.08+054313.9	17.377	16.697	15.988	15.443	15.087	6.06	33.92	-2.20	0.104	GCS Simbad	17 42 21.08	+05 43 13.9
J	17 42 23.550	+05 38 23.500	C	174223.55+053823.5	15.344	14.881	14.291	13.705	13.403	1.01	-31.80	-1.48	0.328	GCS Simbad	17 42 23.55	+05 38 23.5
J	17 42 24.890	+05 06 06.100	C	174224.89+050606.1	16.389	15.827	15.163	14.587	14.272	0.98	-0.62	-1.85	0.181	GCS Simbad	17 42 24.89	+05 06 06.1
J	17 42 25.410	+06 21 05.300	C	174225.41+062105.3	14.731	14.381	13.831	13.177	12.885	-4.17	4.60	-1.29	0.453	GCS Simbad	17 42 25.41	+06 21 05.3
J	17 42 25.690	+05 29 47.300	C	174225.69+052947.2	16.944	16.476	15.847	15.208	14.889	10.53	10.36	-2.14	0.114	GCS Simbad	17 42 25.69	+05 29 47.2
J	17 42 26.600	+06 22 19.800	C	174226.60+062219.8	14.758	14.378	13.873	13.210	12.946	0.05	-6.42	-1.30	0.440	GCS Simbad	17 42 26.60	+06 22 19.8
J	17 42 26.930	+06 20 14.600	C	174226.93+062014.6	16.782	16.238	15.635	15.019	14.640	-4.73	11.58	-2.05	0.132	GCS Simbad	17 42 26.93	+06 20 14.6
J	17 42 28.030	+05 26 40.700	C	174228.03+052640.7	14.683	14.311	13.740	13.129	12.833	4.01	21.98	-1.25	0.481	GCS Simbad	17 42 28.03	+05 26 40.7
J	17 42 28.930	+05 54 53.800	C	174228.93+055453.8	17.691	17.048	16.427	15.888	15.495	28.31	79.04	-2.39	0.078	GCS Simbad	17 42 28.93	+05 54 53.8
J	17 42 28.940	+06 20 28.000	C	174228.94+062028.0	15.234	14.810	14.260	13.592	13.286	2.95	7.72	-1.67	0.336	GCS Simbad	17 42 28.94	+06 20 28.0
J	17 42 31.910	+06 18 49.500	C	174231.91+061849.5	14.560	14.139	13.606	12.988	12.679	-41.07	0.18	-1.19	0.524	GCS Simbad	17 42 31.91	+06 18 49.5

Standard Parameters

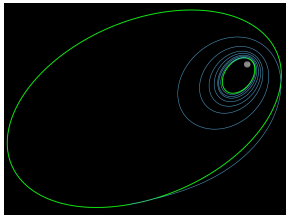


Exoplanets



Orbiting stars

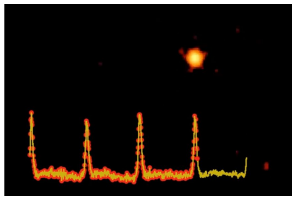
Complex shaped objects



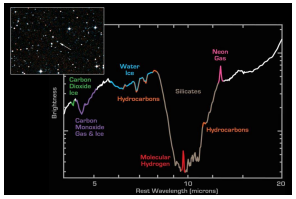
Complex errors

The Model: Associated Data

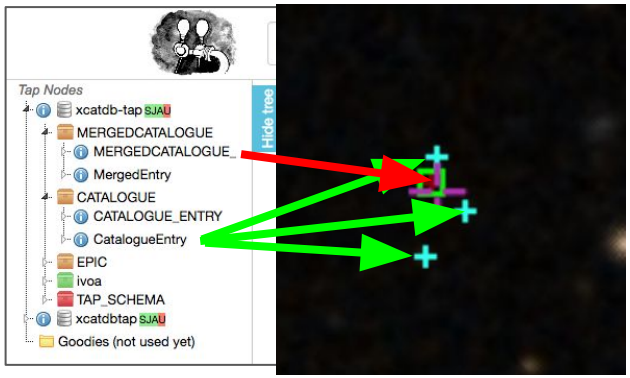
Provenance



Time Series



Spectrum

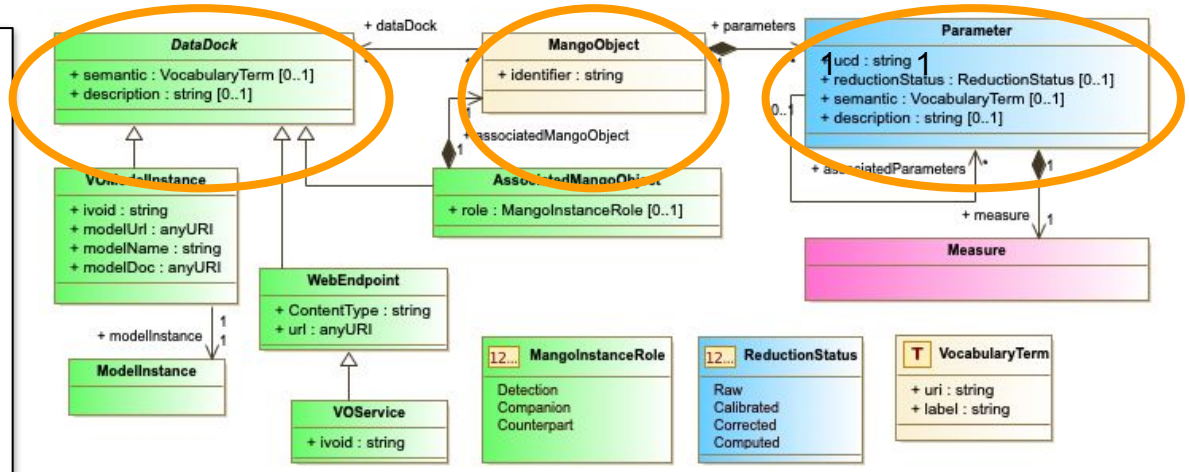


Multiple detections
Correlations

Mango Skeleton

3 components

- **One source identifier**
- **2 Docks**
  - The content of the docks are not defined by the model
  - The model lists possible objects that can be attached to a dock



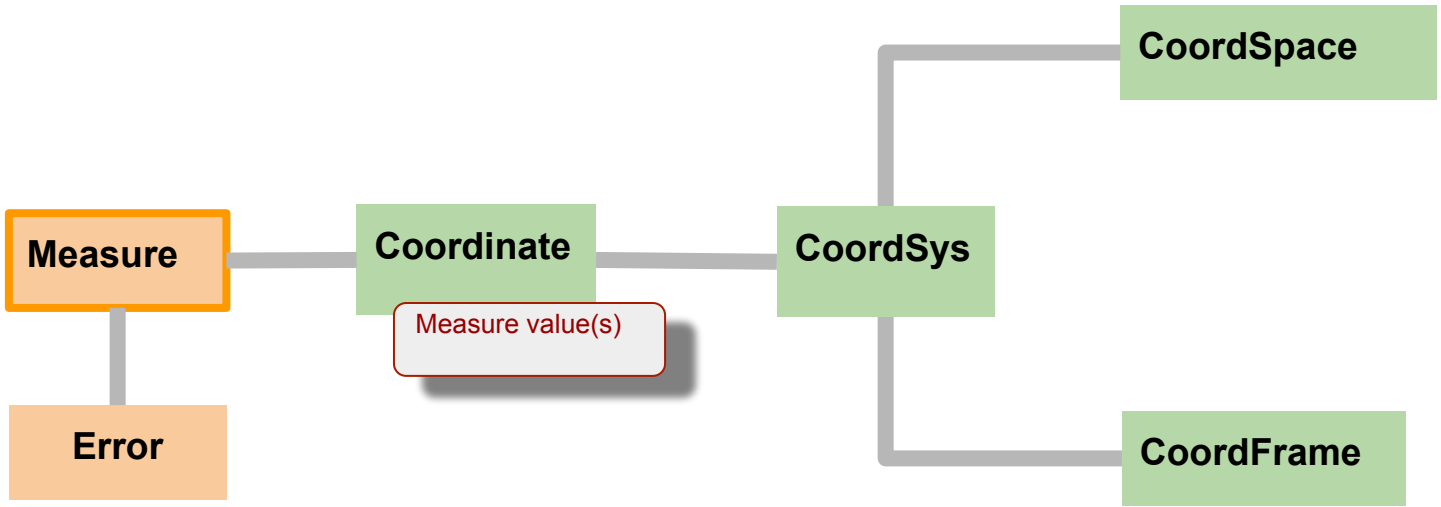
Docks are open ended data containers

- The model describes quantities that can be dropped off on a dock
- It does no say which ones have to be there or not
- The content of the docks varies from a dataset to another
- We can have several instances of the same quantity on a dock
  - Multiple positions
  - Multiple counterparts

STC Measure sub classes

- **Time**
- **Position**
- **Velocity**
- **Proper motion**
- **Polarization**
- **Luminosity**
- **Hardness ratio**
- **Status**
- **Spherical Position**
- **shape**

STC (simplified) Pattern

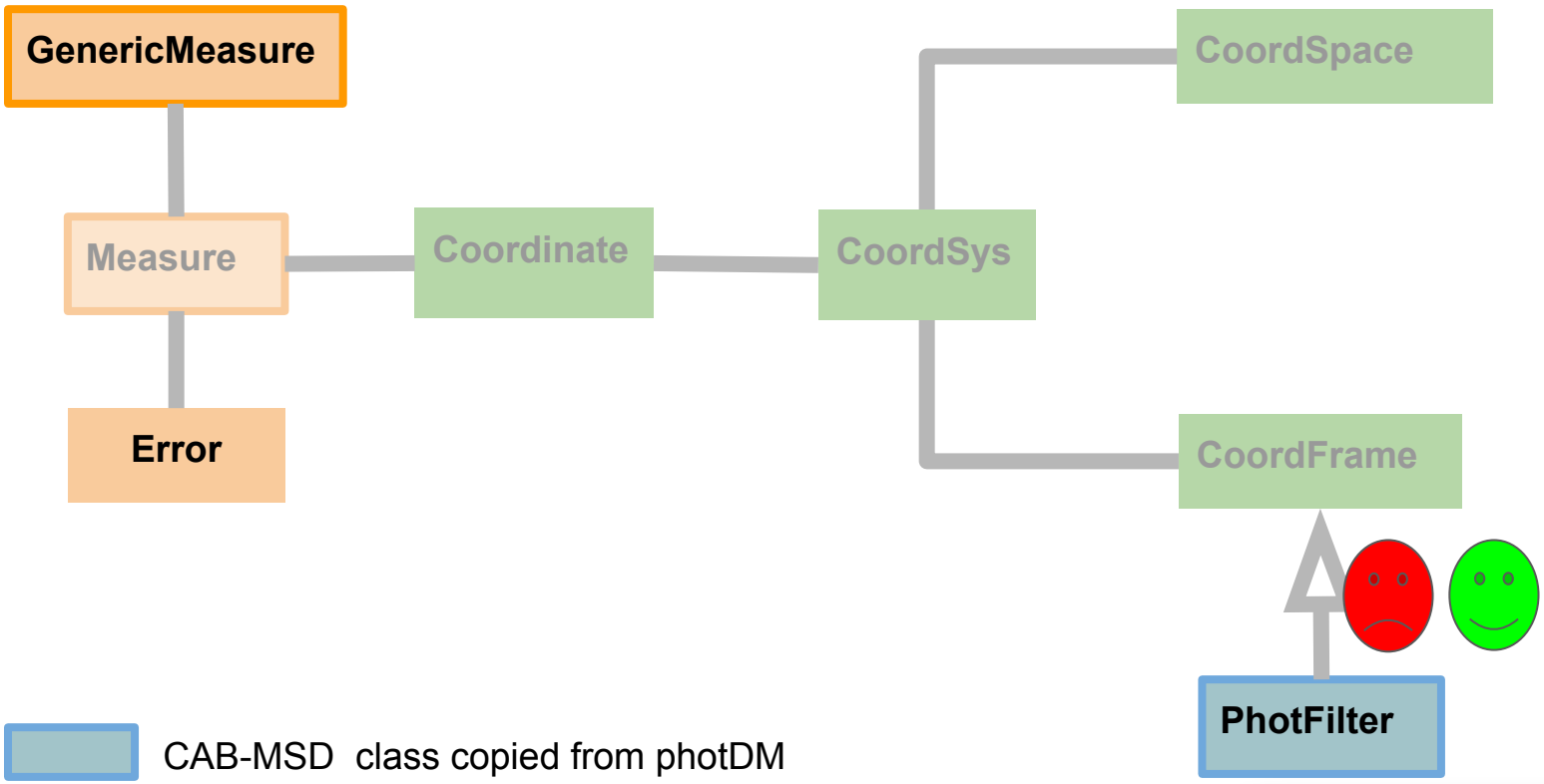


NOTE: Measure natively defined in STC are used as much as possible

Grey links are logical links
The detail of the Meas/Coord is hidden.

General relation of the Meas/Coords objects.

STC Extension: Luminosity



CAB-MSD class copied from photDM

STC class

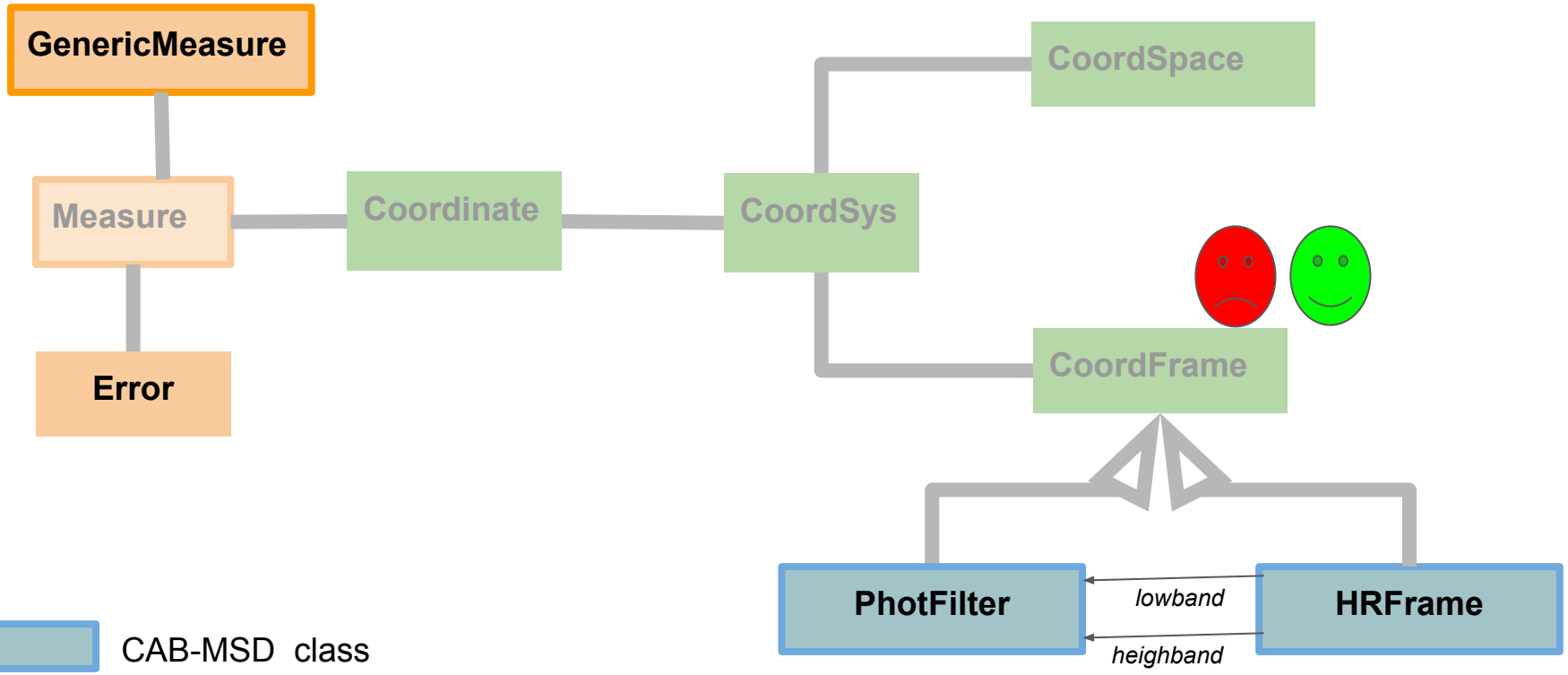
General relation of the Meas/Coords objects.

CABMSD Extension of the Meas/Coords objects.

Grey links are logical links
The detail of the Meas/Coord is hidden.

STC Extension: Hardness Ratio

$$HR = \frac{F_{heighband} - F_{lowband}}{F_{heighband} + F_{lowband}}$$



CAB-MSD class

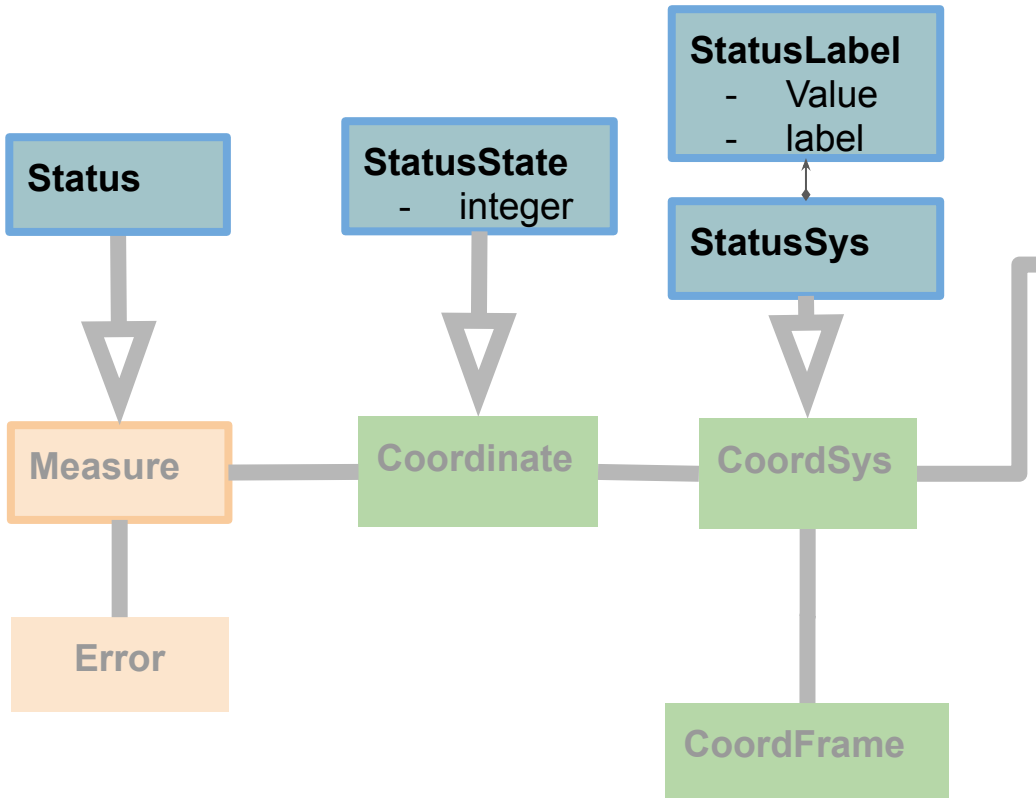
STC class

General relation of the Meas/Coords objects.

CABMSD Extension of the Meas/Coords objects.

Grey links are logical links
The detail of the Meas/Coord is hidden.

STC extension: Flag




```
XMM detection status example
      0: good
      1: Sources parameters may
         be affected
      2: possibly spurious
      3: located in an area
         where spurious
         detection may occur
```

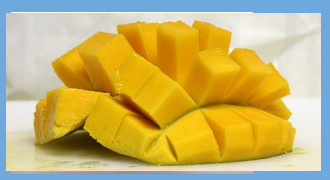
CAB-MSD class

STC class

General relation of the Meas/Coords objects.

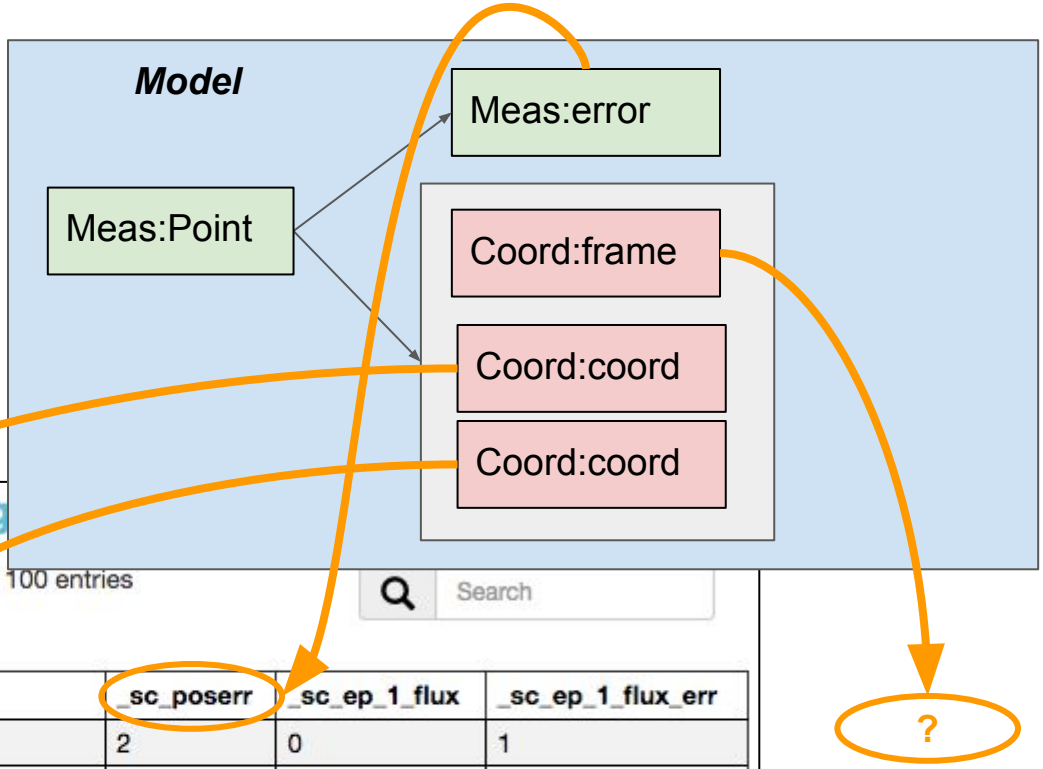
CABMSD Extension of the Meas/Coords objects.

Grey links are logical links
The detail of the Meas/Coord is hidden.



Data Annotation: Bridge connecting model with Data

Bridge between data and model
Must be able to set missing data



xcatdb-tap>MERGEDCATALOGUE>Merg

Show 5 entries

Showing 1 to 5 of 100 entries

Search

_iauname	_sc_ra	_sc_dec	_sc_poserr	_sc_ep_1_flux	_sc_ep_1_flux_err
3XMM J061628.5+223000	94.118827	22.500138	2	0	1
3XMM J061640.9+222438	94.170480	22.410767	5	6	6
3XMM J061704.2+090924	94.267916	9.1566678	1	0	1
3XMM J061706.7+224815	94.278027	22.804341	4	0	2
3XMM J061713.8+222914	94.307609	22.487336	1	0	6

Data annotation : “Don’t be evil” (Larry Page)

- **Shy Annotations**
  - **#1:** Able to be ignored
    - Do not break working things
    - The parser implementation shouldn’t alter the existing code
    - The annotation implementation shouldn’t alter the original data
  - **#2:** To provide what is still missing in VOTables
    - A clear indication of the nature of the VOTable content
  - **#3:** Parser helper: Can be used at different levels
    - Just get the type of the VOTable content
    - Just the meta data
    - Just get column mapping
    - Get everything through the model

Mapping Guidelines: `ModelInstInVot`

- **We need a convenient way to exercise Mango on Real data**

- **Client requirements**
  - Retrieving data with generic code (no dependency with any particular service)
  - Getting a data presentation that facilitates the comparison with different datasets
  - Being able to restore data hierarchies faithful to the model
  - Being able to gather data spread out within the VOTable

- **Provider requirements**
  - Facilitate(*) the annotation of heterogeneous and frozen datasets

- **In between <GROUP> and a pure ORM (Object Relational Mapping)**
  - Compactness
  - Human readability
  - Better than GROUPs to map hierarchical data
  - Do not pretend support a round trip validation (model -> votable -> model)
    - This allows major simplifications

(*) As much as possible

Mapping Block Structure

- **One block located in the top of the VOTable**
- **One block maps data for one model**

```
<MODEL_INSTANCE>
    <MODEL>
     URI + name of the instanciated model
    </MODEL>
    <GLOBALS>
     Model instances with a global scope
        -   Datatypes
        -   Coord systems
    </GLOBALS>
    <TABLE_MAPPING tableref="Table1">
     Mapping of the data contained in the table labeled Table1
    </TABLE_MAPPING>
    <TABLE_MAPPING tableref="Table2">
     Mapping of the data contained in the table labeled Table2
    </TABLE_MAPPING>
</MODEL_INSTANCE>
```

- **Any complex data hierarchy can be exchanged pair key values, tuples and lists**
  - See JSON based Web applications

Key/value pair

```
<ATTRIBUTE
    dmrole="mango:stcextend.PhotFilter.effectiveWavlength"
    dmtype="ivoa:real" value="7740.87" />
```

Tuple

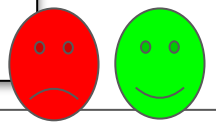
```
<INSTANCE dmrole="mango:Parameter.measure"
    dmtype="mango:Parameter">
    <ATTRIBUTE dmrole="mango:Parameter.semantic"
        dmtype="ivoa:string" value="#position" />
    <ATTRIBUTE dmrole="mango:Parameter.ucd"
        dmtype="ivoa:string" value="pos.eq;meta.main" />
    <ATTRIBUTE dmrole="mango:Parameter.description"
        dmtype="ivoa:string" value="this is the position" />
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="meas:Measure.error"..
</INSTANCE>
```

List

```
<COLLECTION size="-1"
    dmrole="mango:MangoObject.parameters">
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="mango:Parameter.measure"..
    <INSTANCE dmrole="mango:Parameter.measure"..
</COLLECTION>
```

Other Features

Shortcuts: Model components that are parts of a standard can be folded in compact XML elements



```
<INSTANCE dmrole="coords:Point.axis1" dmtype="ivoa:RealQuantity">
  <ATTRIBUTE dmrole="ivoa:RealQuantity.value" dmtype="ivoa:real" ref="RA_ICRS"/>
  <ATTRIBUTE dmrole="ivoa:Quantity.unit" dmtype="ivoa:Unit" value="deg"/>
</INSTANCE>
```

```
<SC_REALQUANTITY dmrole="coords:Point.axis1"
    ref="RA_ICRS" unit="deg" />
```

Row filtering: Only processing data with a certain field value

```
<COLLECTION dmrole="dmrole">
    <TABLE_ROW_TEMPLATE>
        <FILTER ref="ref" value="value"/>
        <INSTANCE dmref="dmref" dmrole="dmrole" />
    </TABLE_ROW_TEMPLATE>
</COLLECTION>
```

Foreign keys: Joining data from different tables

```
<TABLE_MAPPING tableref="fgdgfddf">
    <COLLECTION dmrole="dmrole">
        <JOIN tableref="tableref" primary="primary" foreign="foreign">
            <INSTANCE dmref="dmref" dmrole="dmrole" />
        </JOIN>
    </COLLECTION>
</TABLE_MAPPING>
```

Row grouping: Grouping data of the same source spread over multiple rows

```
<GROUPBY ref="ref">
    <INSTANCE dmref="dmref" dmrole="dmrole">
        <ATTRIBUTE ref="ref" dmrole="dmrole" dmtype="dmtype"
            value="dmvalue" />
    </INSTANCE>
</GROUPBY>
```

AstroPy Wrapper

The MANGO validation requires to show up a good level of compliance with AstroPy.

The ModelInstanceInVot code includes an Astropy wrapper

- Produces Astropy objects from MANGO annotations
- Very few features for now

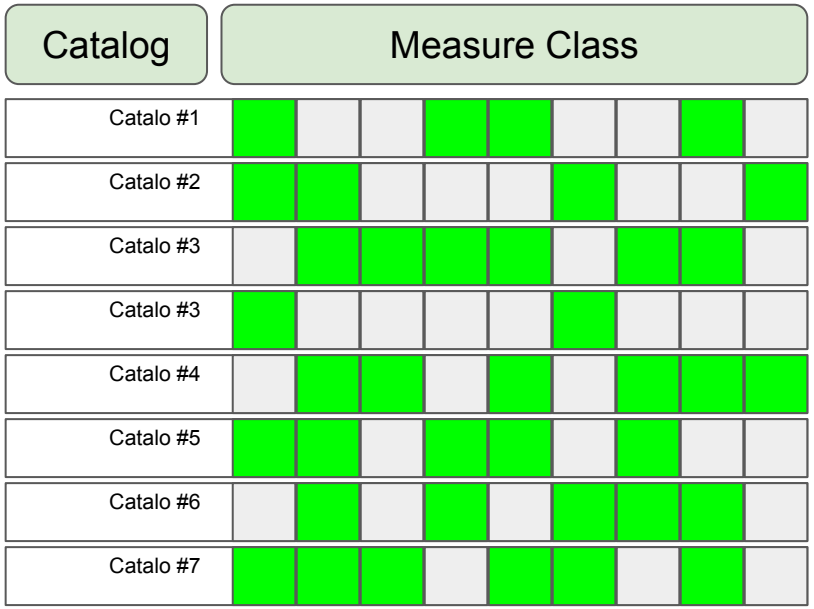
```
wrapper = AstropyWrapper(vodml_instance, mapper_name)

print(f"Astropy space frame: {wrapper.get_space_frame(inst)}")
print(f"Astropy time frame: {wrapper.get_time_frame(inst)}")

'''
# output

Astropy space frame: <ICRS Frame>
Astropy time frame: ('tcb', <EarthLocation (0., 0., 0.) m>, 'mid')
'''
```

Discovering Catalog data in TAP services with MANGO

Issue raise by Christophe Arviset (ESA) at last interop

- **Could be similar to Obscore (`MangoCore`)**
  - Rows: catalog identifiers
  - Columns: MANGO parameters
    - Ranges of simple booleans
- **Not easy to to do because parameter sets are open ended**
- **Must see how to refer to associated data.**

WE should have a look at whether there is a way to tag Mango parameters within the `TAP_SCHEMA`.

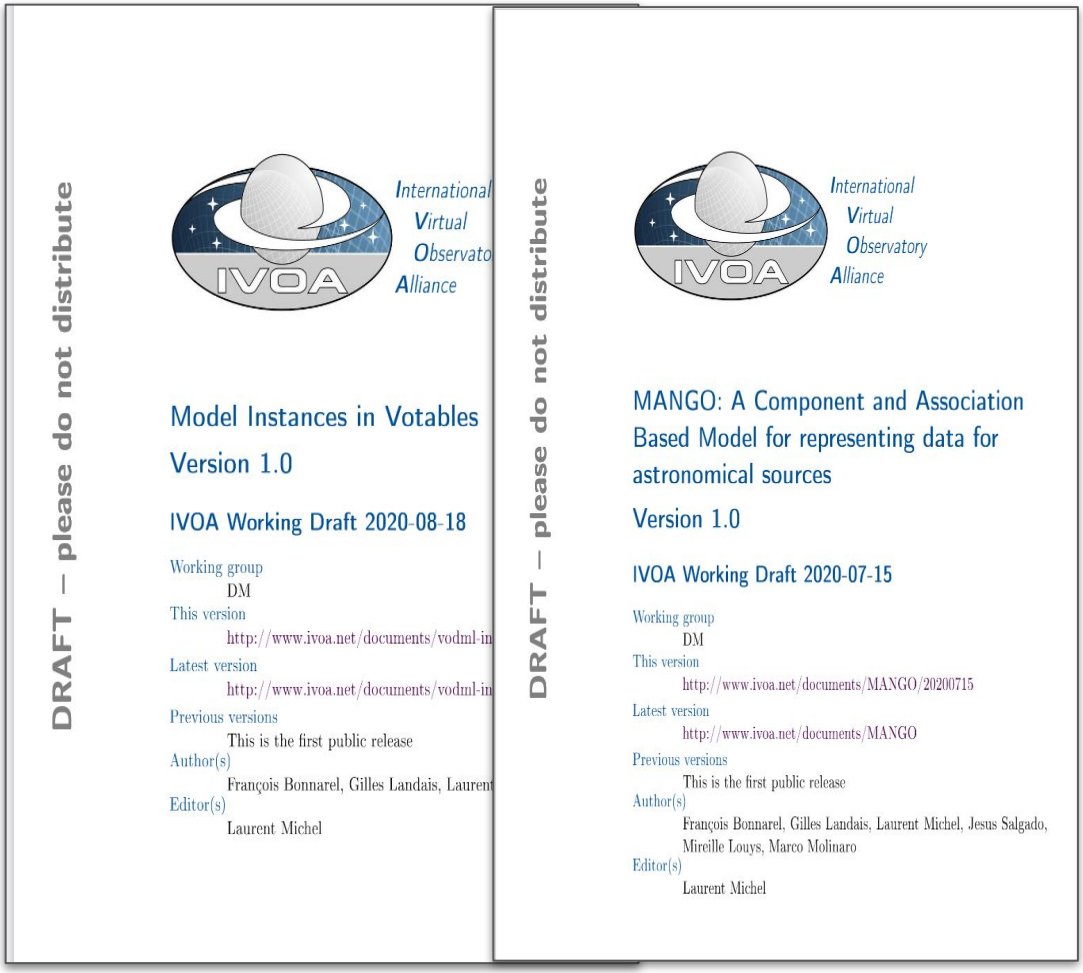


Possible `ivoa.MangoCore` table

Status and Prospects

https://github.com/ivoa-std/MANGO
https://github.com/ivoa-std/ModelInstanceInVot/

- **MANGO**
  - UML - Modelio + VO-DML
  - Document in progress
  - Available on GITHub
- **Mapping**
  - Schema
    - XSD 1.1 ready
    - Lots of unit tests
  - Document in progress
  - Test bench in permanent progress
- **The standard is developed along with concrete implementations**
  - Slow down or speed up the process
    - Question of point of view



FINI

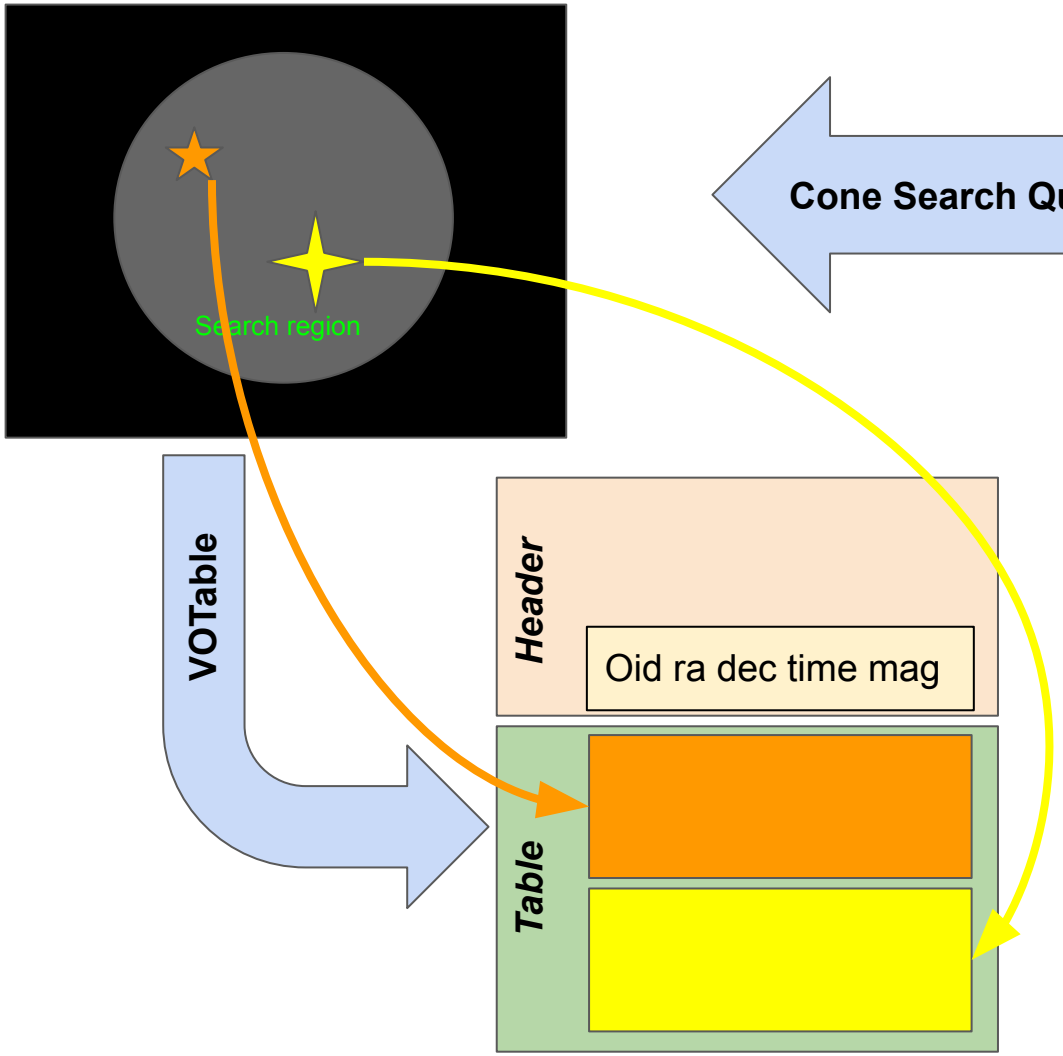
Mapping Process Overhead

- **Not Critical: Mission database**
  - A few number of different products
  - The source model mapping can be done once.

- **Critical: Archival Database (e.g. Vizier)**
  - Huge number of different products
    - Daily updated
  - Mapping a source model comes in addition to usual work
    - Must be done each time a new dataset is published
  - Must be a lightweight task
    - By minimising the amount of meta-data to be mapped
    - By using small **reusable components**

- **Very Critical: TAP services**
  - The possibility of automate the model mapping must be considered
    - This would be very useful for all VO stakeholders
  - There is no concrete proposal yet but (some ideas anyway)
    - Avoid to use show stopper features: **modularity**

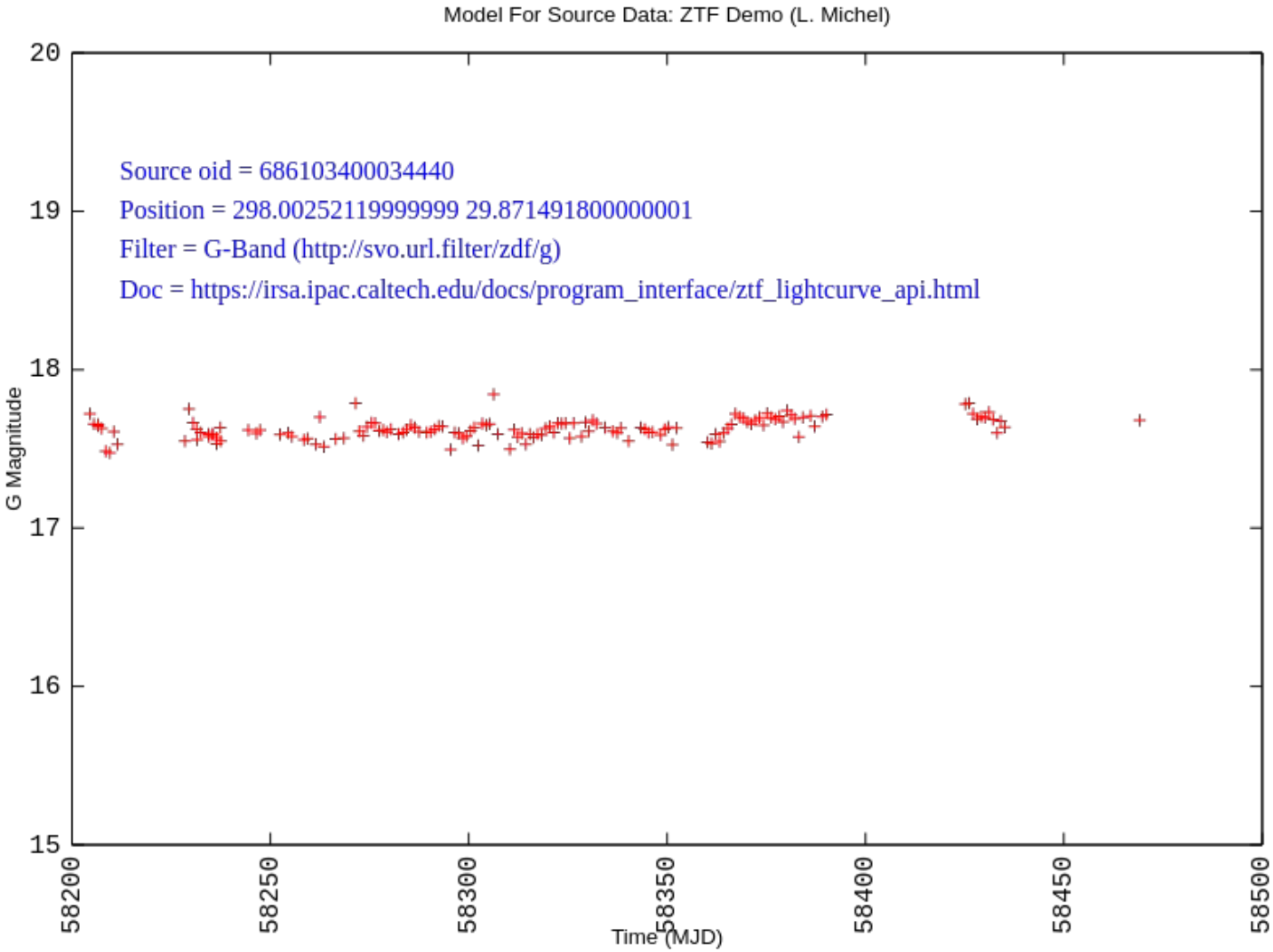
ZTF example



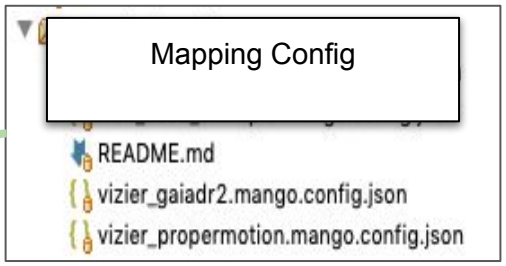
Challenge

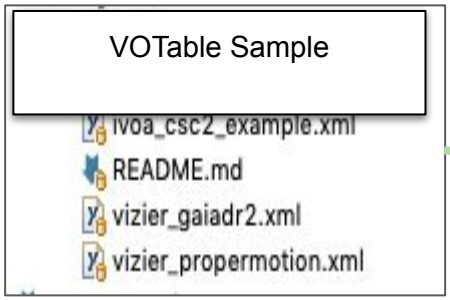
Mapping this VOTable on CABMSD to retrieve the Time Series of the **Orange** star

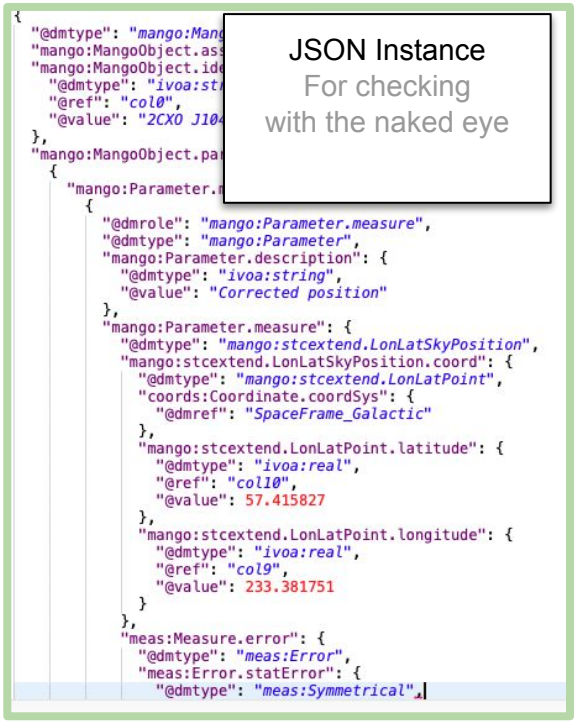
ZTF Example: GNUPlot Output



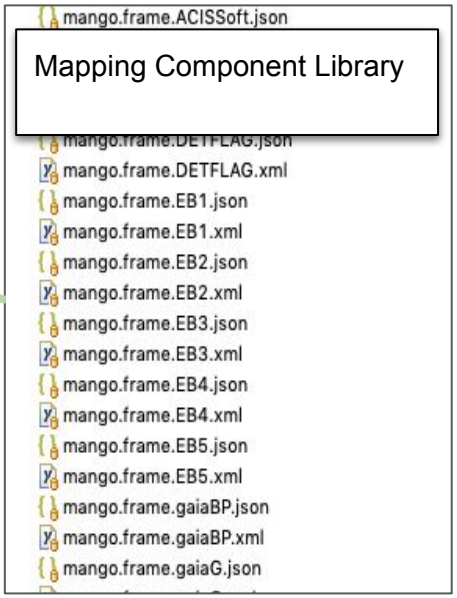
Test Bench

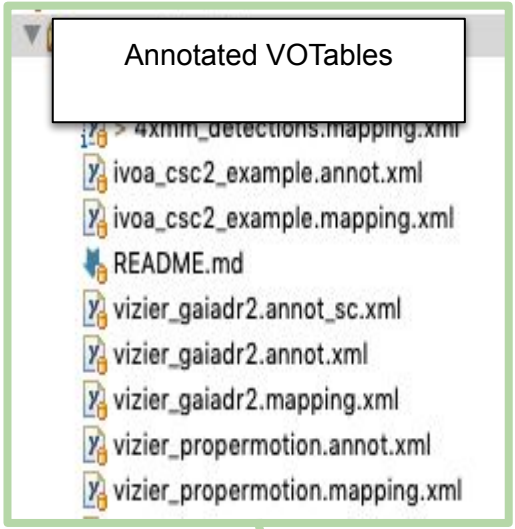






Annoter





Client