Different ways of exposing data through Vizier’s votable output

Contributors:

Engineers: G.Landais
Documentalists: E.Perret, M.Brouty, T.Pouvreau, P.Vannier
Astronomers: P.Ocvirk
2 levels of description to describe a table

- Description applied to individual column:
  → type, unit, description & name, UCD

- Global description:
  - associate columns together according to a common purpose
    e.g.: error on a position
  - Fit the whole table description into a model
    adopt VO-standards for photometry, ObsCore

=> Difficult for VizieR to fit all tables into a pre-defined Data model

however, we can identify some common topics – but how to provide?
Basic adopted conventions

Take advantage of the columns nomenclature

- VizieR adopts a nomenclature (in particular for columns) for column description
  http://cdsarc.u-strasbg.fr/vizier/catstd/catstd-3.3.htx

  ex: RAJ2000, DEJ2000, Epoch
      Bmag, Vmag..

- Adopted also by AAS as an author recommendation for MRT table.

Machine Readable Tables

It is in the best interest of both the author and the reader for lengthy tables to appear in a machine readable table format. Machine readable tables (MRTs) consist of structured ASCII (non-binary) data with a meta-data header. Those MRTs published in the AAS Journals utilize very similar standards and styles as CDS’s VizieR tables. Indeed, VizieR harvests AAS Journal MRTs and makes these data discoverable and searchable via Virtual Observatory protocols, which is another benefit to using this data format.

Link columns together with one-letter-underscore-prefix: e_, l_, ...

- Positions: RAJ2000, e_RAJ2000 (error)
- Photometry: Bmag, e_Bmag (error), l_Bmag (limit)
## Basic adopted conventions

### Naming Convention

Describe columns association using the nomenclature convention

A parameter has frequently associated values, and we have adopted the rule of association with the one-letter-underscore prefix: if a column is obviously associated to another one — typically mean errors or uncertainty flags — we use one of the underscore prefixes listed in prefix.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Explanation</th>
<th>Default Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_\text{label})</td>
<td>aperture used for parameter label</td>
<td>(\geq 0)</td>
</tr>
<tr>
<td>(b_\text{label})</td>
<td>for an upper bound (maximal value) on parameter label</td>
<td></td>
</tr>
<tr>
<td>(l_\text{label})</td>
<td>for a lower bound (minimal value) on parameter label</td>
<td></td>
</tr>
<tr>
<td>(d_\text{label})</td>
<td>for a difference ((\Delta)) on parameter label (e.g. (O-C))</td>
<td></td>
</tr>
<tr>
<td>(u_\text{label})</td>
<td>number of degrees of freedom or number of digits on parameter label</td>
<td>(&gt; 0)</td>
</tr>
<tr>
<td>(e_\text{label})</td>
<td>mean error (upper limit) on parameter label</td>
<td>(\geq 0)</td>
</tr>
<tr>
<td>(\tau_\text{label})</td>
<td>mean error ((\sigma)) on parameter label</td>
<td>(&gt; 0)</td>
</tr>
<tr>
<td>(f_\text{label})</td>
<td>flag on parameter label</td>
<td></td>
</tr>
<tr>
<td>(l_\text{label})</td>
<td>Likelihood on parameter label</td>
<td></td>
</tr>
<tr>
<td>(m_\text{label})</td>
<td>multiplicity index on parameter label to resolve ambiguities</td>
<td>([\leq1])</td>
</tr>
<tr>
<td>(n_\text{label})</td>
<td>note (remark) on parameter label</td>
<td></td>
</tr>
<tr>
<td>(q_\text{label})</td>
<td>number of observations on parameter label</td>
<td>(&gt; 0)</td>
</tr>
<tr>
<td>(r_\text{label})</td>
<td>quality on parameter label</td>
<td></td>
</tr>
<tr>
<td>(s_\text{label})</td>
<td>reference (source) for parameter label</td>
<td></td>
</tr>
<tr>
<td>(u_\text{label})</td>
<td>dispersion ((\sigma)) on parameter label (the (\sigma) of a mean of (N) values is asymptotically equal to the dispersion divided by (\sqrt{N}))</td>
<td>(&gt; 0)</td>
</tr>
<tr>
<td>(v_\text{label})</td>
<td>uncertainty flag on parameter label</td>
<td>([\leq1])</td>
</tr>
<tr>
<td>(w_\text{label})</td>
<td>weight of parameter label</td>
<td>(&gt; 0)</td>
</tr>
<tr>
<td>(x_\text{label})</td>
<td>unit in which parameter label is expressed</td>
<td></td>
</tr>
</tbody>
</table>

Usual mathematical functions may be specified in the label, with parentheses or a dot; for instance, the logarithm of the effective temperature could be labelled \(\log(\text{Tw})\) or \(\log\text{.Tw}\).

### Example Table

- `<TABLE ID="J_ApJ_788_125_table2" name="JApJ/788/125/table2">`
  - `<DESCRIPTION>`Photometry</DESCRIPTION>`
  - `<FIELD name="l_Bmag" ucd="meta.code.error" datatype="char" arraysize="1">`/FIELD>
  - `<FIELD name="Bmag" ucd="phot.mag;em.opt.B" datatype="float" width="5" precision="2" unit="mag">`/FIELD>
  - `<FIELD name="l_Vmag" ucd="meta.code.error" datatype="char" arraysize="1">`/FIELD>
  - `<FIELD name="Vmag" ucd="phot.mag;em.opt.V" datatype="float" width="5" precision="2" unit="mag">`/FIELD>
  - `<FIELD name="l_Jmag" ucd="meta.code.error" datatype="char" arraysize="1">`/FIELD>
  - `<FIELD name="Jmag" ucd="phot.mag;em.IR.J" datatype="float" width="5" precision="2" unit="mag">`/FIELD>
  - `<FIELD name="f_Jmag" ucd="meta.code" datatype="char" arraysize="2">`/FIELD>`
Positions description

Meta-data related to positions in VizieR

- meta-data applied to the whole table
e.g. : system (ICRS, Galactic, FK4..), equinox, epoch
- meta-data associated to positions
e.g. : epoch, proper motion, parallax

Several systems possible in the same table in VizieR!

→ define a main position (UCD pos.eq.ra;meta.main)
   including pos., err., prop mot., plx.

Main positions enables to gather all positions columns
(In vizieR this is possible only for main position!)
VOTable output

- Put meta.main only on columns position (RA,DEC)
- Use `<COOSYS>` in the output

```
<RESOURCE ID="yCat_1022003603" name="J/other/NewA/36.70">
  <DESCRIPTION>
    Astrometry of 3 vdBH open clusters (Orellana+, 2015)
  </DESCRIPTION>
  <COOSYS ID="J2000" system="eq_FK5" equinox="J2000">
  </COOSYS>
  <TABLE ID="J/other/NewA_36.70_table2" name="J/other/NewA/36.70/table2">
    <DESCRIPTION>
      Centre coordinates, mean proper motion, number of members N and diameters of the clusters
    </DESCRIPTION>
    +<FIELD name="RA[J2000]" ucd="pos.eq.ra;meta.main" ref="J2000" datatype="double" width="9" precision="5" unit="deg"/>
    +<FIELD name="e_RA[J2000]" ucd="stat.error;pos.eq.ra" datatype="double" width="8" precision="5" unit="deg"/>
    +<FIELD name="DE[J2000]" ucd="pos.eq.dec;meta.main" ref="J2000" datatype="double" width="9" precision="5" unit="deg"/>
    +<FIELD name="e_DE[J2000]" ucd="stat.error;pos.eq.dec" datatype="double" width="8" precision="5" unit="deg"/>
    +<FIELD name="pmRA" ucd="pos.pm;pos.eq.ra" ref="J2000" datatype="float" width="5" precision="2" unit="mas/yr"/>
    +<FIELD name="e_pmRA" ucd="stat.error;pos.pm;pos.eq.ra" datatype="float" width="5" precision="2" unit="mas/yr"/>
    +<FIELD name="pmDE" ucd="pos.pm;pos.eq.dec" ref="J2000" datatype="float" width="5" precision="2" unit="mas/yr"/>
    +<FIELD name="e_pmDE" ucd="stat.error;pos.pm;pos.eq.dec" datatype="float" width="5" precision="2" unit="mas/yr"/>
  </TABLE>
</RESOURCE>
```
Photometry output

The VizieR photometry viewer is a global output of VizieR tables for which documentalists found filters used in observation (or similar filter)

Photometry meta-data

→ A reference table of filters (SVO) applied to magnitudes columns
Photometry output

- Use the Photometry Model (IVOA note) "Providing Photometric Data Measurements Description in VOTables" (S.Derriere)

```xml
<RESOURCE ID="VizieR_S610644484" name="VizieR(2019-05-08T15:28:04)"

<DESCRIPTION>
VizieR database maintained by CDS, see http://vizier.u-strasbg.fr
</DESCRIPTION>
<COOYSYS ID="J2000" system="eq_FK5" equinox="J2000"/>

<TABLE ID="VizieR_0" name="allVizieR">

<GROUP ID="gsed" name="_sed" ucd="phot" utype="spec:PhotometryPoint">

<DESCRIPTION>
The SED group is made of 4 columns: mean frequency, flux, flux error, and filter designation
</DESCRIPTION>
<FIELD ref="sed_flux" uctype="spec:PhotometryPoint"/>
<FIELD ref="sed_eflux" uctype="spec:PhotometryPointError"/>
<FIELD ref="sed_filter" uctype="photdm:PhotometryFilter.Identifier"/>

</GROUP>
</TABLE>
</RESOURCE>
```

Not a standard
Time output

- For a year, documentalist assigns Time description to time-column according to the IVOA spec.:
  - scale (TDB, TAI, GMT...)
  - frame (BARYCENTER, HELIOCENTER, ...)
  - systematic_error, offset, uncertainty
- Meta-data applied to the whole table on times columns
- Several time columns possible in the same table (with different meta-data)
  e.g. Gaia DR2:
  http://vizier.unistra.fr/viz-bin/VizieR-3?-source=I/345/transits

→ gather columns related to each time description possible in VizieR
Calibrated FoV transit photometry for CU5, consolidated and provided by CU7 for variable stars in Gaia names in green) (17712391 rows)

| Scale: TCB | Frame: BARYCENTER | Offset: 2455197.50 | Uncertainty: 44 |
| scale: TCB | frame: BARYCENTER | offset: 2455197.50 | uncertainty: 5 |
| scale: TCB | frame: BARYCENTER | offset: 2455197.50 | uncertainty: 5 |
Time output

Votable 1.4 in beta-release

- Use `<TIMESYS>`
- Gather columns related to the same time-column with GROUP
- Note: particular case for columns having a Coordinate+Time description

⚠️ needs 2 ref. ( `<COOSYS>`+ `<TIMESYS>` )
not possible yet with XSD schema : update asked..?
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TIMESYS</td>
<td>group</td>
</tr>
<tr>
<td>TIME</td>
<td>No DM available yet</td>
</tr>
</tbody>
</table>

**TIMESYS Group**

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>Calibration of FoV transit photometry for CUS, consolidated and provided by CUS for variable stars in Gaia DR2 (epoch_phot)</td>
</tr>
</tbody>
</table>

**Field Definitions**

- TIME
- E_TIME
- e_TIME
- TIME_BPM
- e_TIME_BPM
- TIME_FBP
- e_TIME_FBP
- TIME_BPmag
- e_TIME_BPmag
- TIME_RPmag
- e_TIME_RPmag
- TIME_Pmag
- e_TIME_Pmag
• Different ways to expose data

• No possible to fit every table in DM due to the VizieR heterogeneity

• VizieR is open for all VOTable serialisation if meta-data are findable!
  → not always available in literature!

• Each (new) meta-data (like Time, photometry, ObsCore for images/spectra) is a significant effort asked to CDS
  => it has a cost!
→ Needs a method to associate columns in VOTable
e.g.: error on columns, flags on columns...

Enriched grammar suggestion applied on attributes in FIELDS

Associate columns with attributes: is_error_of, is_related_to...

- associate <FIELDS> together
- semantic to describe the <FIELDS> relationships

(grammar inspired from DOI XML schema: is_referenced_to, is_part_of....)