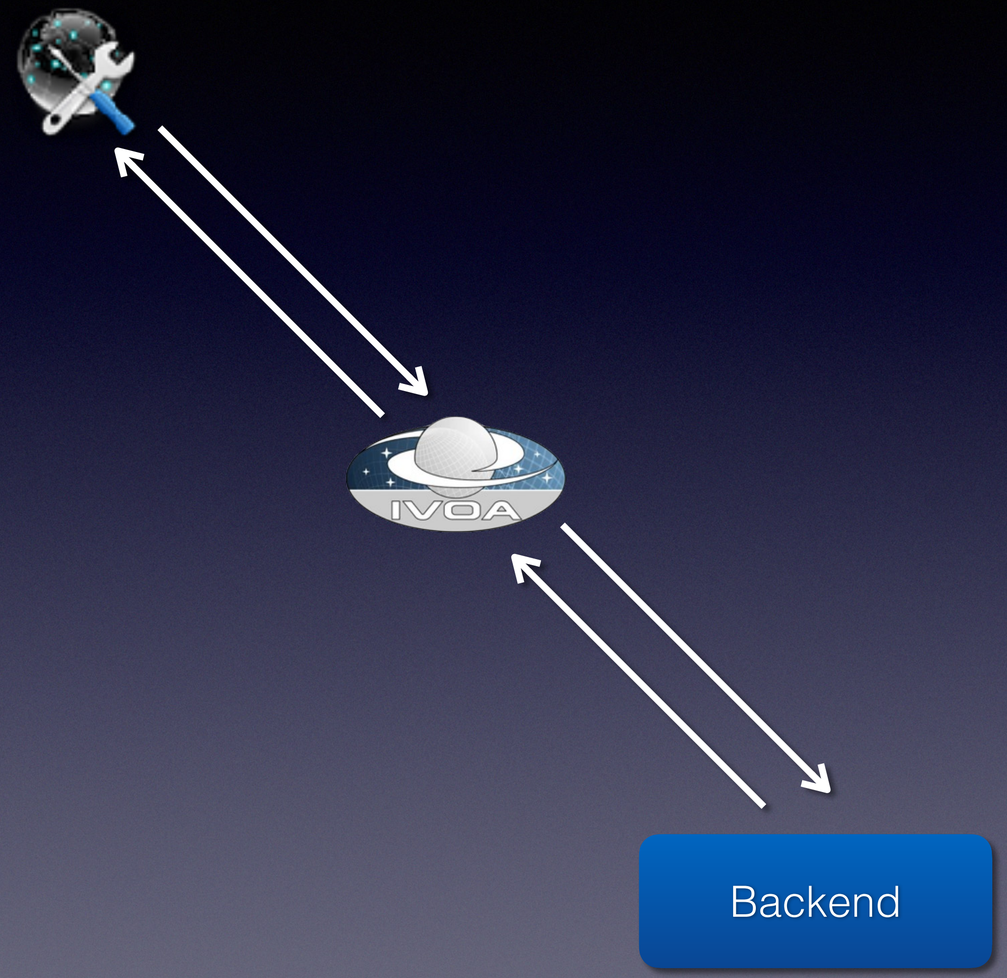
Specta Data Model

The SPHEREx Problem



Trey Roby, IRSA, IPAC, Caltech
roby@ipac.caltech.edu

The problem

- Firefly will be dealing many spectra at one time
- Use cases:
  - Upload table of object -> Return an array of spectra (2 per object)
  - Upload a table of spectrum
- UI shows
  - 2 spectra per Object
  - Cutout Images related to select points
- Download: the Spectra results

Issues

- Single spectrum file is too limited
- Need format that contains many spectra

Many Spectra

- Table: Each row contains 1 or more spectra
- Uses columns with array data

Multi Spectra Rough View

ra	dec	wavelength_1	flux_1	wavelength_2	flux_2	flux_3
1.233	3.422	[*array data*]	[*array data*]	[*array data*]	[*array data*]	[*array data*]
1.511	4.533	[*array data*]	[*array data*]	[*array data*]	[*array data*]	[*array data*]

Multi Spectra Table

```
<TABLE utype="ipac:MultiSpectrum">
```

Recognized by `utype`

VO Table columns

```
<FIELD ID="source_id" datatype="long" name="source_id" ucd="meta.id"/>
<FIELD ID="external_survey" datatype="unsignedByte" name="external_survey"/>
<FIELD ID="external_source_id" datatype="long" name="external_source_id" ucd="meta.id"/>
<FIELD ID="spherex_class" datatype="int" name="spherex_class"/>
<FIELD ID="ra" datatype="double" name="ra" precision="7" ucd="pos.eq.ra;meta.main"
unit="deg" width="11"/>
<FIELD ID="dec" datatype="double" name="dec" precision="7" ucd="pos.eq.dec;meta.main"
unit="deg" width="11"/>
<FIELD ID="pm_ra" datatype="float" name="pm_ra" unit="mas/yr"/>
<FIELD ID="pm_dec" datatype="float" name="pm_dec" unit="mas/yr"/>
<FIELD ID="parallax" datatype="float" name="parallax" unit="mas"/>
<FIELD ID="x_image" arraysize="*" datatype="float" name="x_image" unit="pix"/>
<FIELD ID="y_image" arraysize="*" datatype="float" name="y_image" unit="pix"/>
<FIELD ID="offset_ra" arraysize="*" datatype="float" name="offset_ra" unit="mas"/>
<FIELD ID="offset_dec" arraysize="*" datatype="float" name="offset_dec" unit="mas"/>
```

Groups

Each Spectrum Group `utype`

Each Axis `utype`

Each Axis `utype`

```
<GROUP ID="photometry" utype="ipac:Spectrum.ArrayData">
    <GROUP utype="spec:Spectrum.Data.SpectralAxis">
        <FIELDref ref="lambda"/>
        <GROUP utype="spec:Spectrum.Data.SpectralAxis.Accuracy">
            <FIELDref ref="lambda_width"/>
        </GROUP>
    </GROUP>
    <GROUP utype="spec:Spectrum.Data.FluxAxis">
        <FIELDref ref="flux"/>
        <GROUP utype="spec:Spectrum.Data.FluxAxis.Accuracy">
            <FIELDref ref="flux_err"/>
        </GROUP>
    </GROUP>
    <!-- additional related columns -->
    <FIELDref ref="x_image" />
    <FIELDref ref="y_image" />
    <FIELDref ref="offset_ra" />
    <FIELDref ref="offset_dec" />
    <FIELDref ref="mjd" />
    <FIELDref ref="flux_bkg" />
</GROUP>
```

Reusing Wavelength

- Some spectrum have the same wavelengths
- Our example: sphereX: interpolated spectrum
- Huge waste of space to repeat in every column

Using <PARAM/> for repeated column data

```
<PARAM ID="lambda_interp"
       name="lambda_interp"
       datatype="float"
       arraysize="102"
       ucd="em.wavenumber"
       unit="um"
       utype="spec:Spectrum.Data.SpectralAxis.Value"
       value="0.758951 0.7770666 … 14.665331 4.700469 … 4.8436894 4.8801713"
/>
```

- Array of 100+ entries
- shared by every interpolated spectrum

Interpolated spectrum group

```
<GROUP ID="interp"  utype="ipac:Spectrum.ArrayData">
    <GROUP utype="spec:Spectrum.Data.SpectralAxis">
        <PARAMref ref="lambda_interp"/>
    </GROUP>
    <GROUP utype="spec:Spectrum.Data.FluxAxis">
        <FIELDref ref="flux_interp"/>
        <GROUP utype="spec:Spectrum.Data.FluxAxis.Accuracy">
            <FIELDref ref="flux_interp_err"/>
        </GROUP>
    </GROUP>
    <FIELDref ref="x_image" />
    <FIELDref ref="y_image" />
    <FIELDref ref="offset_ra" />
    <FIELDref ref="offset_dec" />
    <FIELDref ref="mjd" />
    <FIELDref ref="flux_bkg" />
    <FIELDref ref="flags" />
 </GROUP>
```

Use PARAMref

Use FIELDref

Demo

very, very bleeding edge