



Implementation of the IVOA Spectral Data Model at IPAC

Vandana Desai

Science Lead
Caltech/NASA-IPAC Infrared
Science Archive

&

Tim Brooke
Gregory Dubois-Felsmann
Steve Groom
Justin Howell
Emmanuel Joliet
Joe Mazzarella
Serge Monkwewitz
Xiuqin Wu



NASA/IPAC Infrared Science Archive



Abstract

IPAC is home to several data centers that serve spectroscopic data: The InfraRed Science Archive (IRSA), the NASA Extragalactic Database (NED) and the Keck Observatory Archive (KOA). Together, these archives curate spectra from many observatories and projects, including Spitzer, SOFIA, Herschel, and COSMOS. In the future, Euclid and SPHEREx will add over a billion new spectra of asteroids, Galactic ices, stars, galaxies, and quasars. Gaia, Roman, and ALMA are also creating spectroscopic data sets with enormous archival science potential. In addition, NED curates photometric data and combines them to into SEDs spanning the electromagnetic spectrum. We present our experiences in beginning to adapt metadata for existing data sets into the IVOA Spectral Data Model. We also strongly advocate for the prioritization of work on the Spectral Data Model in light of upcoming missions.



Spectroscopy and SEDs at IPAC

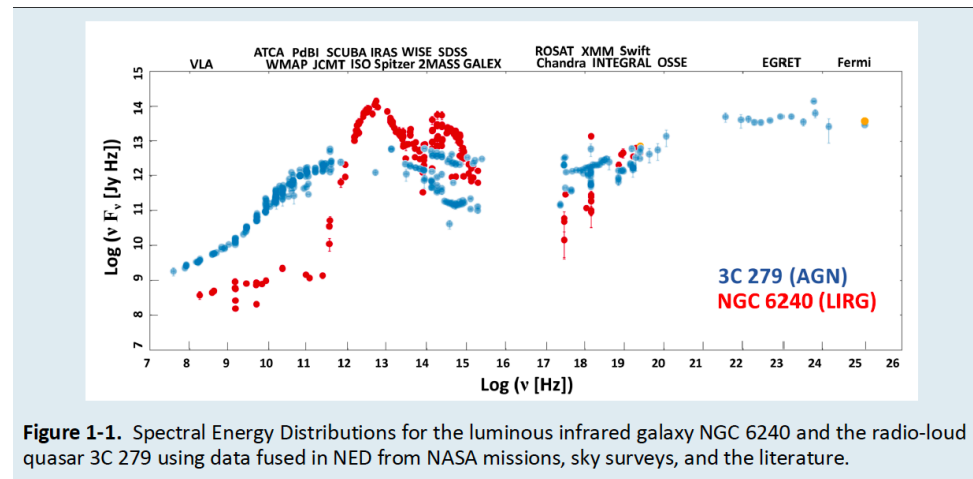


Figure 1-1. Spectral Energy Distributions for the luminous infrared galaxy NGC 6240 and the radio-loud quasar 3C 279 using data fused in NED from NASA missions, sky surveys, and the literature.

....from many different providers

Spectral data are about to get big

IRSA's Era of Spectroscopy

SPHEREx all-sky survey:

- 1 billion galaxy spectra
- 100 million stellar spectra
- 100 thousand ice absorption spectra
- 1 mission quasar spectra
- 10 thousand asteroid spectra

Euclid Wide and Deep surveys:

- 30 million galaxy spectra

...in addition to the smaller volume but unique spectroscopic capabilities of Spitzer, Herschel, SOFIA, IRTF, GUSTO

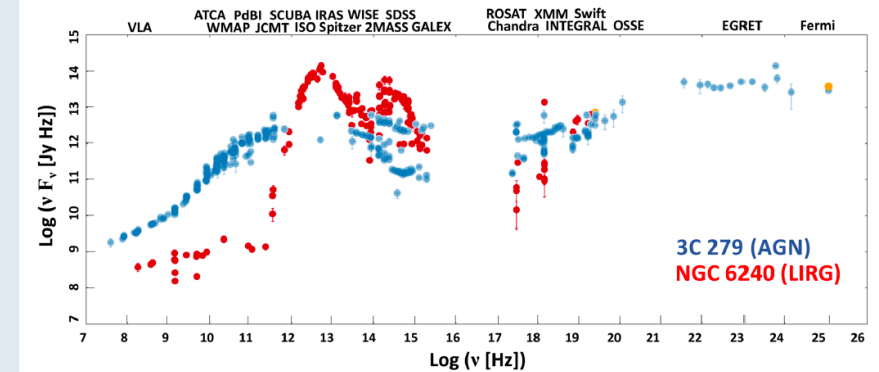


Figure 1-1. Spectral Energy Distributions for the luminous infrared galaxy NGC 6240 and the radio-loud quasar 3C 279 using data fused in NED from NASA missions, sky surveys, and the literature.

- **Webb**
- **Roman**
- **Gaia**



Spectral DM

- **Nothing currently in progress.**
- **Most recent approved version: v1.1 (2011)**
 - *“This data model may be used to represent spectra, time series data, segments of SED (Spectral Energy Distributions) and other spectral or temporal associations.”*
- **Most recent proposed recommendation: v2.0 (2016)**
 - *“Work on this Data model has concluded. . . . it is expected that the Spectral model will be revisited, possibly expanding the scope to cover other products such as Eschelle spectra and TimeSeries.”*

Title	Most stable	In progress	Version history																			
Spectral DM - IVOA Spectral Data Model	1.1		2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	1.1	1.1	1.1	1.03	1.02	1.01	1.01	1.01	1.01	1.00



What will users do with these spectra?

- Search
 - IPAC web tools
 - Other archive search tools
 - Python
- Visualize
 - Firefly by IPAC
- Analyze
 - Mission-provided software
 - Open-source software
- Download
 - Can include thousands at a time!

1. Identify a spectral data product

Firefly

Images Catalogs TAP Searches Charts Upload Background Monitor

Upload file Upload from URL

Choose File SPITZER_S0_25343744_0001_3_E7173899_tune.votable

Format: VO_TABLE Size: 66.5234375K KB Parts: 1

Index	Type	Description
<input checked="" type="checkbox"/>	0 Table	SPITZER_S0_25343744_0001_3_E7173899_t

Table Meta

desc Spitzer IRS SL PBCD spectrum
name SPITZER_S0_25343744_0001_3_E7173899_tune.votable
utype spec:Spectrum

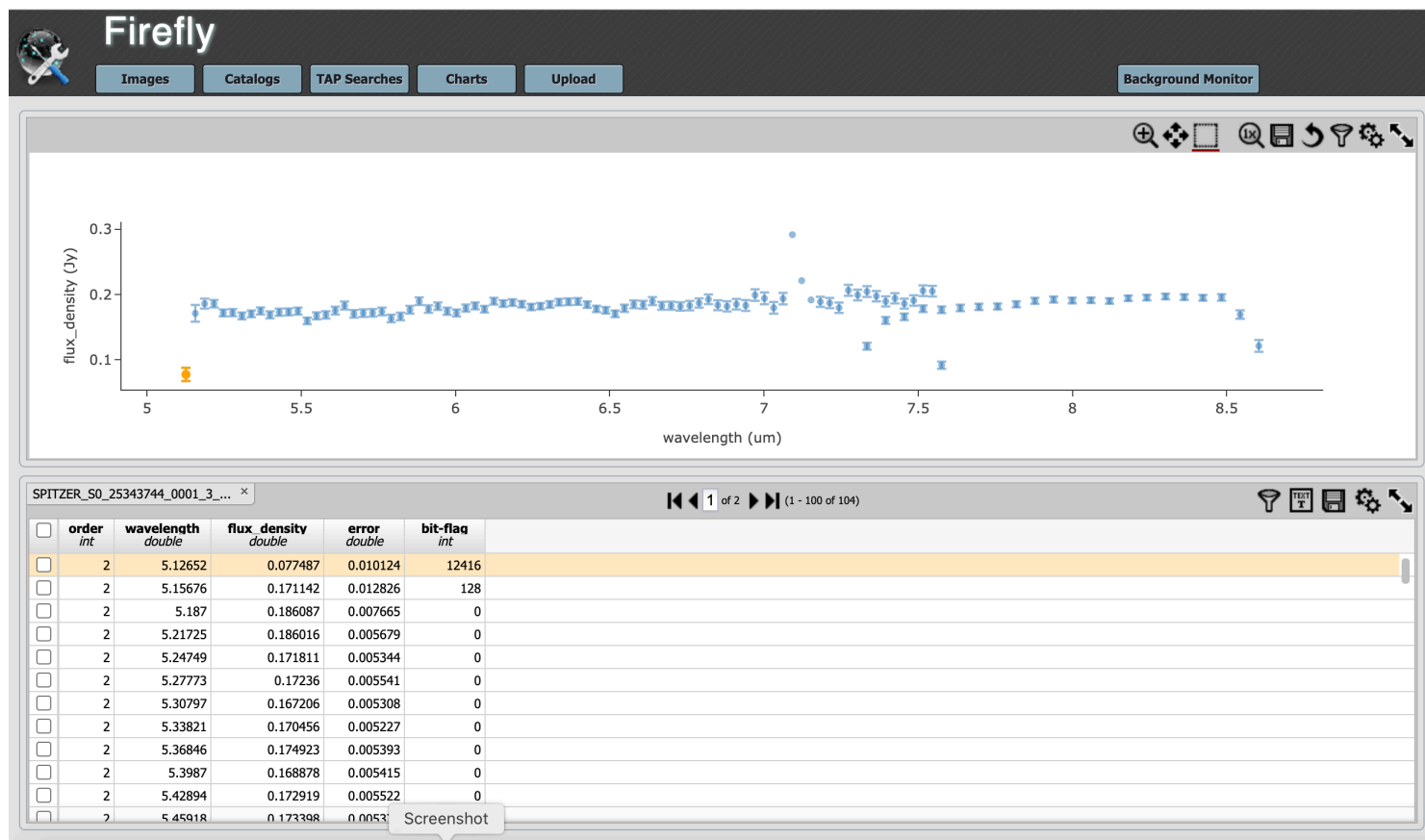
Table Params

name	type	unit	desc
order	int		
wavelength	double	um	
flux_density	double	Jy	
error	double	Jy	
bit-flag	int		

Search Cancel

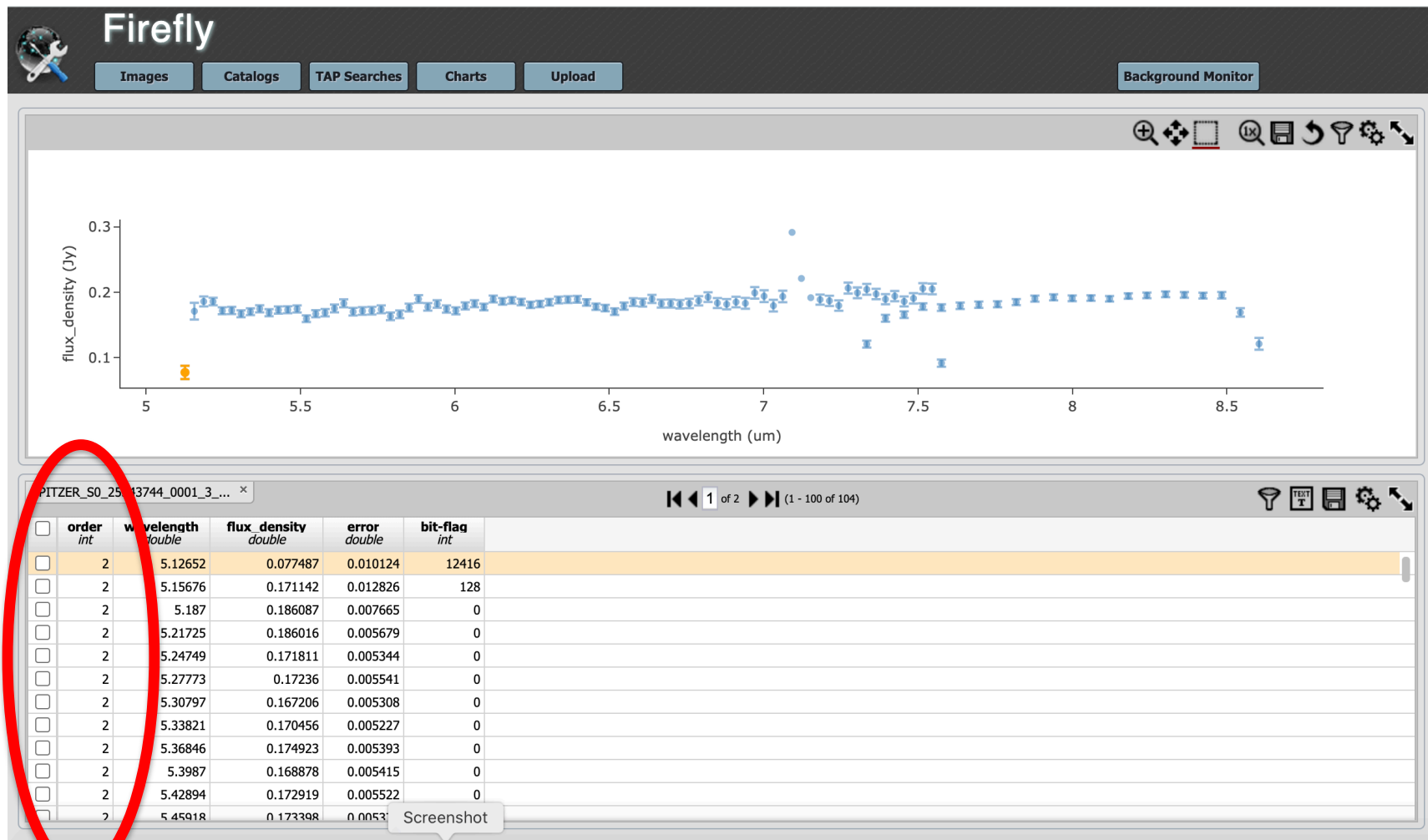
✓ Table utype = spec:Spectrum

2. Plot flux vs wavelength with errors



- ✓ Param utypes:
 - ✓ Spectrum.Data.SpectralAxis
 - ✓ Spectrum.Char.FluxAxis
 - ✓ Spectrum.Data.FluxAxis.Accuracy.StatError

3. Connect Points by Spectral Order

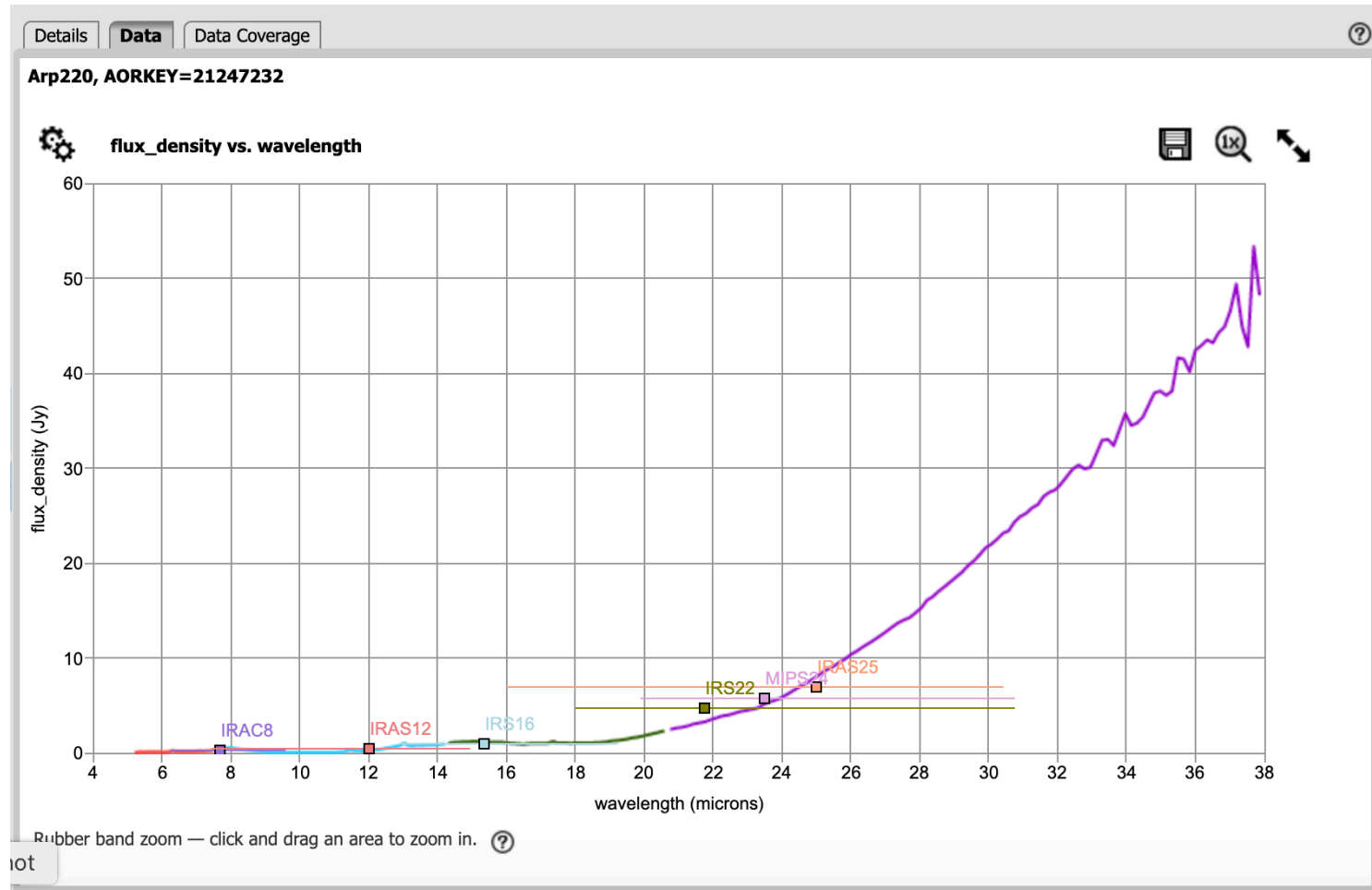


- Existing data model doesn't address order
- UCD="instr.order" does exist
- Natural to include order as a column in a single table

4. Plot upper & lower limits for SEDs

- `<FIELD ID="DataFluxDensityUpperError" datatype="double" name="Upper limit of uncertainty" ucd="stat.error;phot.flux.density;em;stat.max" utype="spec:Spectrum.Data.FluxAxis.Accuracy.StatErrLow" unit="Jy" >`
 - `<DESCRIPTION>The NED upper limit of uncertainty for photometric data.</DESCRIPTION>`
 - `</FIELD>`
 - `<FIELD ID="DataFluxDensityLowerError" datatype="double" name="Lower limit of uncertainty" ucd="stat.error;phot.flux.density;em;stat.min" utype="spec:Spectrum.Data.FluxAxis.Accuracy.StatErrLow" unit="Jy">`
 - `<DESCRIPTION>The NED lower limit of uncertainty for photometric data.</DESCRIPTION>`
 - `</FIELD>`
-
- **Standard utype and UCD needed**

4. Overplot Photometry or SED



Requires Phot DM

Inconsistency in UCD1+ and SpectrumDM 1.1

E | phot.flux.density

| Flux density (per wl/freq/energy interval)

4. “E” means a photometric quantity, and can be followed by a word describing a part of the electromagnetic spectrum
5. “C” is a colour index, and can be followed by two successive word describing a part of the electromagnetic spectrum;

From SpectrumDM1.1:

Field	FITS	UCD1+	Meaning	Req	Default
Accuracy Fields - per data point (default to corresponding Spectrum.Char values)					
Data.FluxAxis.Accuracy.StatError	TTYPEn = 'ERR'	stat.error;phot.flux.density; em.*	symmetric error	OPT	(Char)
Data.FluxAxis.Accuracy.StatErrLow	TTYPEn='ERR.LO'	stat.error;phot.flux.density; em.*;stat.min	Lower error	OPT	(Char)
Data.FluxAxis.Accuracy.StatErrHigh	TTYPEn='ERR.HI'	stat.error;phot.flux.density; em.*;stat.max	Upper error	OPT	(Char)



4. Change units on both axes

- ? Model specifies OGIP unit convention, not VOUnit, established 2014
- If a user changes flux, we'd like the flux error to change automatically
- ```
<GROUP ID="Data.FluxAxis" name="Data.FluxAxis" utype="spec:Spectrum.Data.FluxAxis">
```

  - ```
<DESCRIPTION>Flux Axis Data</DESCRIPTION>
```
 - ```
<FIELDref ref="flux_density"/>
```
  - ```
<FIELDref ref="error"/>
```
 - ```
<PARAM ID="DataFluxUcd" datatype="char" name="DataFluxUcd"
 utype="spec:Spectrum.Data.FluxAxis.Ucd" value="phot.flux.density" arraysize="*">
```

    - ```
<DESCRIPTION>UCD for flux</DESCRIPTION> </PARAM>
```
 - ```
<PARAM ID="DataFluxUnit" datatype="char" name="DataFluxUnit"
 utype="spec:Spectrum.Data.FluxAxis.Unit" value="Jy" arraysize="*">
```

    - ```
<DESCRIPTION>Unit for flux</DESCRIPTION> </PARAM>
```
- ```
</GROUP>
```

## 7. Display upper limits for SEDs

- **utype and UCD needed:**
- What we are using now:
- ```
<FIELD ID="DataFluxDensityUpperError" datatype="double"
name="Upper limit of uncertainty" ucd="stat.error;phot.flux.density;
em;stat.max"
utype="spec:Spectrum.Data.FluxAxis.Accuracy.StatErrLow" unit="Jy" >
<DESCRIPTION>The NED upper limit of uncertainty for photometric
data.</DESCRIPTION>
</FIELD>
```
- ```
<FIELD ID="DataFluxDensityLowerError" datatype="double"
name="Lower limit of uncertainty" ucd="stat.error;phot.flux.density;
em;stat.min"
utype="spec:Spectrum.Data.FluxAxis.Accuracy.StatErrLow" unit="Jy">
<DESCRIPTION>The NED lower limit of uncertainty for photometric
data.</DESCRIPTION>
```

# Future “Big Spectra”

- How to represent billions of spectra?
- VO Table format not efficient for this use case.
- How does Spectrum DM apply?
- Is SSA sufficient?





# Summary

- Spectral DM is becoming really important.
- Need more guidance:
  - Spectral Order
  - Synthetic Photometry
- Should the following be updated?
  - OGIP units versus VOUnit
  - UCD for mjd
- How do we deal with future “Big Spectra”?
- What can we use now for SEDs if not SpectrumDM?