



Simple Spectral Data Model

Version 0.5

IVOA Working Draft

2004-05-21

This version:

<http://www.ivoa.net/Documents/WD/DAL/DMSSP-20031210.html>

Latest version:

<http://www.ivoa.net/Documents/WD/DAL/DMSSP.html>

Previous versions:

None.

Note: ivoa.net links are not yet active; the above are placeholders.

Editors:

Jonathan McDowell, Doug Tody

Authors:

Jonathan McDowell, Doug Tody, Tamas Budavari, Frank Valdes, Pavlos Protopapas, Arnold Rots, Markus Dolensky and the IVOA Data Access Layer Working Group

Abstract

We present a data model describing the structure of spectrophotometric datasets with spectral and temporal coordinates and associated metadata.

Status of this document

This is a Working Draft. The first release of this document was 2003 XXX XX.

This is an IVOA Working Draft for review by IVOA members and other interested parties. It is a draft document and may be updated, replaced or obsoleted by other documents at any time. It is inappropriate to use IVOA Working Drafts as reference materials or to cite them as other than "work in progress". A list of current IVOA recommendataions and other technical documents can be found at <http://www.ivoa.net/Documents/>.

Contents

1	Introduction and Motivation	3
2	Requirements	3
3	Spectral data model summary	3
4	Spectral data model Measurement objects	7
4.1	Spectrum and Time Series	7
4.2	Spectral coordinate	7
4.3	Flux (Spectral Intensity) Object	8
4.4	BackgroundModel Object	9
4.5	Time coordinate	9
4.6	Position coordinate	10
4.7	Accuracy Fields	10
5	Associated Metadata Fields	11
5.1	Coverage Fields	11
5.2	Frame fields	12
5.3	Derived Data Fields	13
5.4	Curation model	13
5.5	Data Identification model	13
6	Global metadata	14
6.1	SED attributes	14
6.2	Target model	14
6.3	Packaging model	15
7	Relationship to general VO data models	15
8	Serializations	16
8.1	FITS serialization	16
8.2	Alternate FITS serialization	19
8.3	VOTable Serialization	19
8.4	Direct XML serialization	21

1 Introduction and Motivation

Spectra are stored in many different ways within the astronomical community. In this document we present a proposed abstraction for spectra and serializations in VOTABLE and FITS, for use as a standard method of spectral data interchange.

We distinguish in several places between the implementation proposed in this document, referred to as Version 1, and capabilities proposed for possible later implementation.

2 Requirements

We need to represent a single 1-dimensional spectrum in sufficient detail to understand the differences between two spectra of the same object and between two spectra of different objects.

We need to represent time series photometry, with many photometry points of the same object at different times.

Finally, we need to represent spectral energy distributions (SED) which consist of multiple spectra and photometry points, usually for a single object.

3 Spectral data model summary

Our model for an SED is a set of spectra and/or time series, some of which may have only one or few data points (photometry) and each of which may have different contextual metadata (aperture, position, etc.). Specifically, a spectrum will have arrays of the following values:

- Spectral coordinate (e.g. wavelength), central and bin min and max
- Time coordinate, convertible to MJD UTC
- Flux value, with upper and lower statistical (uncorrelated) errors
- Quality mask
- Spectral resolution

and will have the following associated metadata:

- Instrument and filter ID
- Exposure time in seconds
- Position of aperture center, given as J2000 ICRS degrees.
- Aperture in arcseconds
- Systematic (correlated) error

- Bibcode

A general SED may be considered as consisting of **segments** for which the associated contextual metadata is constant.

In the following sections we elaborate these concepts in detail, including some complications that we explicitly do not attempt to handle in this version.

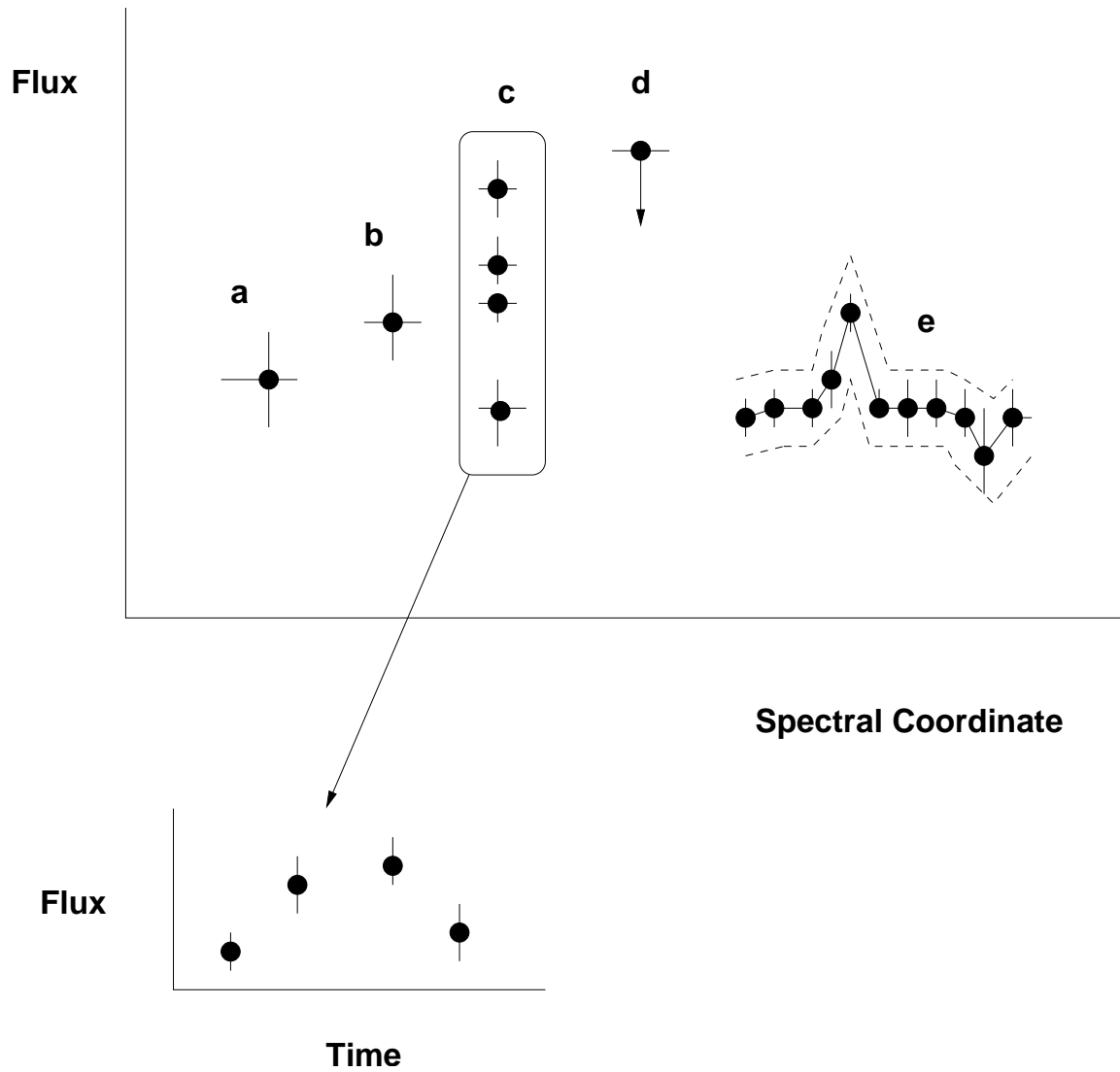


Figure 1: A moderately complicated SED with five segments a through e. segments **a** and **b** are photometry points with associated errors (vertical bars) and bandpass widths (horizontal bars). segment **c** is a time series with four measurements at different times with the same instrumental configuration. segment **d** is an upper limit measurement. segment **e** is a simple spectrum, with point-to-point statistical errors indicated by vertical bars and an overall (correlated) systematic error indicated by the dashed lines.

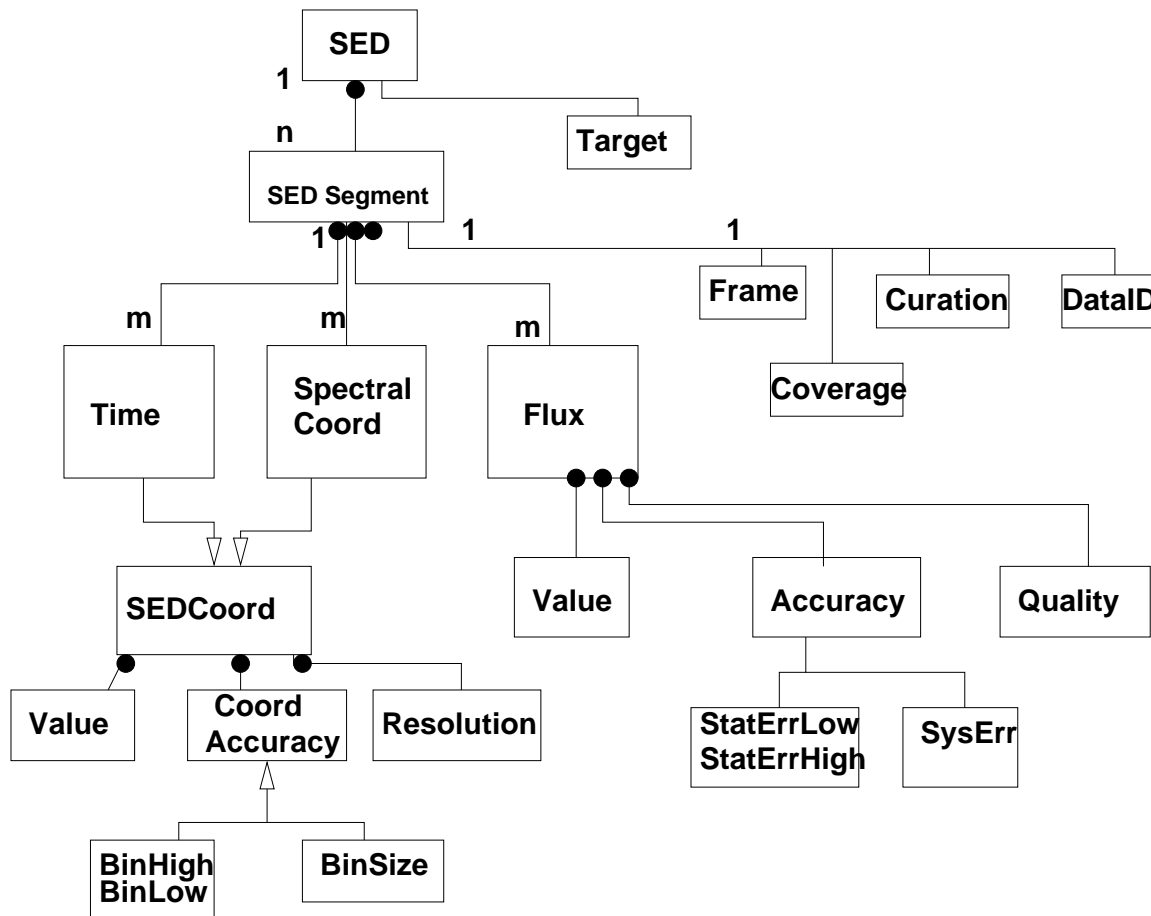


Figure 2: UML class diagram for the SED data model. The Frame, Coverage, Curation, DataID classes are not shown in detail.

4 Spectral data model Measurement objects

4.1 Spectrum and Time Series

A simple Spectrum is an SED object for which the spectral coordinate varies but the time coordinate is fixed, and for which the associated metadata are constant (i.e. a single segment with a fixed time.)

A simple Time Series is an SED object for which the spectral coordinate is constant but the time coordinate varies, and for which the associated metadata are constant (i.e. one set of metadata for the whole time series; a single segment at a fixed spectral coordinate).

Doug: your presentation suggested that TimeSeries did not have to have constant sky position. Was this intentional? I'm not sure if we want to get into moving objects at this stage...

4.2 Spectral coordinate

Astronomers use a number of different spectral coordinates to label the electromagnetic spectrum. The cases enumerated by Greisen et al. (2003) are listed below with their UCDS.

Exactly one of these fields should be present. We distinguish between the VO data model field name (which might be used for VOTABLE UTYPE), the FITS WCS name, and the UCD1+ names. For UCD1+, I propose UCDS in italics for those not currently covered by the document.

Note 1: For version 1, we adopt only the first three entries, Wavelength, Frequency, Energy. The others are considered reserved names to be considered for future implementations.

Note 2: For the velocity cases, I propose an *em.veloc* tree, rather than a *src.veloc* tree, because the velocity here is really a labelling of a spectral coordinate, and the link to the physical radial velocity of the different emission sources contributing to the spectrum is rather indirect.

Spectral coordinate options			
Field	FITS WCS	UCD1+	Meaning
SpectralCoord	WAVE	em.wl	Wavelength
	FREQ	em.freq	Frequency of photon
	ENER	em.energy	Photon energy
	-	instr.pixel;em.wl	Instrumental spectral bin
	WAVN	em.wavenumber	Wavenumber
	AWAV	em.wl.air	Wavelength
	WAVE-LOG	em.wl.log	Log Wavelength
	FREQ-LOG	em.freq.log	Log Frequency of photon
	ENER-LOG	em.energy.log	Log Photon energy
	VELO	em.veloc	Apparent radial velocity
	VRAD	em.veloc.radio	Radio velocity
	VOPT	em.veloc.optical	Optical velocity
	BETA	em.veloc.beta	Velocity (c=1)

4.3 Flux (Spectral Intensity) Object

There are two alternate representations of the flux: Flux (the background-subtracted flux) and TotalFlux (the source+background). In the current version of SSA, only one Flux or TotalFlux entry should be present for each segment.

For either case, there are many slightly different physical quantities covered by the general concept of Flux; we distinguish them by their UCD.

Flux coordinate options		
Field	UCD1+	Meaning
Flux		
TotalFlux		
	phot.fluxDens;em.wl	Flux density per unit wave.
	phot.fluxDens;em.freq	Flux density per unit freq.
	phot.fluxDens;em.energy	Flux density per energy interval
	phot.fluxDens,em.wl.log	Flux density per log wave interval
	phot.fluxDens.sb, em.wl	Surface brightness per unit wavelength
	phot.count	Counts in spectral channel
	<i>phot.rate</i>	Count rate in spectral channel
	arith.ratio,phot.fluxDens	Flux ratio of two spectra
	phys.luminosity;em.wl	Luminosity per unit wave
	phys.luminosity;em.freq	Luminosity per unit freq
	phys.luminosity;em.energy	Luminosity per unit energy
	phys.luminosity;em.energy.log	Luminosity per log frequency
	<i>phys.energy-density</i>	Radiation energy density per unit volume, per unit wave etc.
	<i>phys.intensity</i>	Flux per unit solid angle (at source)
	instr.antenna-temp	Antenna temperature
	<i>phot.brightness-temp</i>	Brightness temperature
	phot.mag	Magnitude in defined band
	phot.mag	AB (spectrophotometric) magnitude
	phot.flux.beam	Flux per resolution element (e.g. Jy/beam)
	phot.mag.sb	Surface brightness in magnitudes

4.4 BackgroundModel Object

The BackgroundModel array is required to have the same UCD and units as the Flux array. It represents a model for the expected flux values if the Target had zero flux.

Often, the BackgroundModel will be generated by taking a flux measurement at another location and rescaling it for any difference in exposure.

4.5 Time coordinate

The time coordinate is given by an elapsed time in some physical units (e.g. seconds or days) relative to a reference time, (see the Space-Time Coordinates document for details of time coordinate complexity). It refers to the midpoint of the sample interval.

The reference time should be specified by a date in ISO-8601 format. The default value of the reference time is -4713-11-24T11:59:27.81, corresponding to the origin of Julian Day Number on the TT (Terrestrial Time) timescale. Using TT is preferable to UTC because it does not contain leap seconds, so the elapsed time in days is just

equal to the JD.

The time unit is specified by a string, and the only valid values are 's' (seconds) and 'd' (days).

4.6 Position coordinate

In general we may consider position coordinates as part of the measurement, but this is not included in the current document. The position of each segment is given in the Coverage.Location field.

4.7 Accuracy Fields

We include accuracy models for both the coordinates (spectral and temporal) and the fluxes.

We express the bandpass for each spectral bin as a low and high value for the spectral coordinate, or as a width. The same is done for photometry points, which amounts to approximating a filter by a rectangular bandpass. Time bins are also given as low and high values or as a width.

We also use a very simple error model for the fluxes: we include plus and minus flux errors, and a quality flag. The errors are understood as 1 sigma gaussian errors which are uncorrelated for different points in the spectrum. If the data provider has only upper limit information, it should be represented by setting the flux value and the lower error value equal to the limit. In general applications may choose to render measurements as upper limits if the flux value is less than some multiple (e.g. 3) of the lower error. We also allow a systematic error value, assumed constant across a given spectrum and fully correlated (so that, e.g. it does not enter into estimating spectral slopes). We also include a trivial resolution model: a single number nominally representing a FWHM spectral or time resolution expressed in the same units as the spectral or time coordinate. The spatial resolution may be useful to know if it exceeds the aperture size.

The Quality model represents quality by an integer, with the following meanings: 0 is good data, 1 is data which is bad for an unspecified reason, and other positive integers may be used for data which is bad or dubious for specific reasons. Scalar string-valued metadata fields Quality.2, Quality.3... may be used to define those reasons on a per-segment basis.

Accuracy fields		
Field	UCD1+	Meaning
SpectralCoord.BinLow	em.*;stat.min	Wavelength etc
SpectralCoord.BinHigh	em.*;stat.max	Wavelength etc
SpectralCoord.BinSize	em.*;?	Wavelength bin size
SpectralCoord.Resolution	instr.spectr.resolution	Spectral resolution FWHM
SpectralCoord.Error	em.*; stat.error	Spectral coord measurement error
Time.BinLow	em.*;stat.min	Time bin start
Time.BinHigh	em.*;stat.max	Time bin stop
Time.BinSize	em.*;?	Time bin size
Time.Resolution	time.resolution	Temporal resolution FWHM
Time.Error	time; stat.error	Time coord measurement error
Spatial.Resolution	?	Spatial resolution of data
Spatial.Error	pos.eq;stat.error	Astrometric error
Flux.ErrorLow	phot.flux;em....;stat.error	Negative error
Flux.ErrorHigh	phot.flux;em....;stat.error	Positive error
Flux.SysError	phot.flux;stat.error. <i>sys</i>	Systematic error
Flux.Quality	meta.code.qual;phot.flux,em....	Quality mask
Flux.Quality.n		String value, for n = 0,1,2...; meaning of quality value

5 Associated Metadata Fields

5.1 Coverage Fields

The coverage fields will have a constant value for a given spectrum.

We give a value giving the effective exposure time (useful for selecting among multiple spectra from the same instrument). The aperture is important to determine what part of an extended object is contributing to the spectrum; we allow a simple aperture description consisting of a single number representing the aperture diameter in decimal degrees; it is anticipated that a full region description based on the CoordArea object will be supported in later versions.

Coverage fields		
Field	UCD1+	Meaning
Coverage.Location.Sky	pos.eq	RA and Dec
Coverage.Location.Time	time. <i>obs</i>	Midpoint of exposure
Coverage.Location.Spectral	instr.bandpass	Band, consistent with RSM
Coverage.Extent.Sky	instr.fov	Aperture angular diameter
Coverage.Extent.Time	time.expo	Exposure time
Coverage.Extent.Spectral	instr.bandwidth	Width of spectrum in A or other spec. coord.
Coverage.Region.Sky		Aperture region
Coverage.Region.Time	time.expo.start,time.expo.end	Start and stop time
Coverage.Region.Spectral	em.*;stat.min,stat.max	Start/Stop in spectral coordinate
Coverage.Fill.Sky	stat. <i>fill</i> ;pos.eq	Sampling Filling factor
Coverage.Fill.Time	time;stat. <i>fill</i> ;time	Sampling Filling factor
Coverage.Fill.Spectral	stat. <i>fill</i> ;em.*	Sampling Filling factor

5.2 Frame fields

The Frame object is a simplified instance of the STC CoordSystem object.

For all the numeric fields, units must be supplied. For the spectral coordinate, position and time we need to add further metadata to define the frames used. Here we try and match existing FITS conventions (TIMESYS is a convention used in the X-ray community).

Other frame information needed for velocity spectral coordinates include the observation-fixed spectral frame, the observatory location, the source redshift, and the velocity zero point (in Greisen et al, SSYSOBS, OBSGEO, VELOSYS, REST-FRQ/RESTWAV).

Frame fields			
Field	FITS WCS	UCD1+	Meaning
Frame.Sky.Type	RA,DEC	pos.frame	e.g. Equatorial
Frame.Sky.Equinox	EQUINOX	time.equinox;pos.frame	e.g. 2000.0
Frame.Sky.System	RADECSYS	?	e.g. ICRS or FK5
Frame.Time.Type	TIMESYS	time.scale	Timescale, UTC TT etc
Frame.Time.Zero	MJDREF	time;arith.zp	Zero point of timescale in MJD
Frame.SpectralCoord.System	SPECSYS	?	e.g. BARYCENT;REST use Greisen et al values

5.3 Derived Data Fields

The Derived Data object has useful, and optional, summary information about the segment. For now, we include the option of adding signal-to-noise and variability indicators.

Frame fields		
Field	UCD1+	Meaning
Derived.SNR	stat.snr	Signal-to-noise for segment
Derived.VarAmpl	stat.ampl	Variability amplitude as fraction of mean (0 to 1)

5.4 Curation model

The Curation field is an object derived from the VO Resource Metadata document.

Curation fields		
Field	UCD1+	Meaning
Publisher		Publisher
PubID		URI for VO Publisher
Logo		URL for creator logo
Contributor		Contributor (may be many)

5.5 Data Identification model

The Data Identification model gives the dataset ID and its membership of larger collections.

Packaging fields		
Field	UCD1+	Meaning
Title		Dataset Title
Creator		VO Creator ID
Collection		Collection name(s)
DatasetID	meta.id,meta.dataset	Dataset ID
Date	time;meta.dataset	Data processing/creation date
Version		Version of dataset
Instrument	meta.id;instr	Instrument ID

The dataset is associated with one or more Collections (instrument name, survey name. etc.) indicating some degree of compatibility with other datasets sharing the same Collection properties.

6 Global metadata

6.1 SED attributes

The overall SED object will contain values indicating the number of SED segments and curation information about their assembly into a single SED, as well as their overall spectral range (the union of the segment bandpass coverages).

We introduce the concept of an dataset creation type, which can have one of the following three values:

- Archival, indicating that it is one of a collection of datasets (in this case SEDs) generated in a systematic, homogeneous way and stored statically (or at least versioned). It will be possible to regenerate this dataset at a later date.
- Dynamic, indicating that the dataset was created ‘on-the-fly’, possibly by assembling the latest available data and applying a possibly changeable processing algorithm. A future attempt to generate this dataset in the same way may give different results.
- Custom, indicating that the dataset was created by manual analysis and processing, whether by data center staff or as part of a scientist’s research project. Traceability of the processing may be incomplete.

SED fields		
Field	UCD1+	Meaning
Creator	meta.id	Person or organization creating the SED
CreatorID	meta.id	URL for documentation
Date	time;meta.dataset	Data processing/creation date
NSegments	arith.factor	Number of segments
Bandpass.Min	em.wl;stat.min	Total spectral coord range, wavelength, meters
Bandpass.Max	em.wl;stat.max	Total spectral coord range
CreationType		dataset creation type

6.2 Target model

In spectral data it is particularly important to be able to specify the target of the observation, which may be an astronomical source or some other target (calibration, diffuse background, etc.). By explicitly including a target model in the SED object we can not only facilitate searches on particular types of target but also support archives of model spectra for which the Coverage fields may not be relevant.

A single SED object is assumed to correspond to data for a single ‘target’, i.e. the same target for each segment.

Target fields		
Field	UCD1+	Meaning
Target.Name	meta.id;src	Target name
Target.Class	src.class	Target or object class
Target.spectralClass	src.spType?	Object spectral class
Target.redshift	src.redshift	Target redshift
Target.pos	pos.eq;src	Target RA and Dec
Target.VarAmpl	src.var.amplitude	Target variability amplitude, typical

At the moment there is no international standard list of valid values for Target class and spectral class. Nevertheless an initial deployment of the VO would gain some benefit from using archive-specific classes, and provide a framework for converging on a standard list.

6.3 Packaging model

The simple Packaging model for SSA describes the format of the associated dataset. Allowed values for the format (detailed serialization to be specified in a separate document:)

- VOTABLE
- XML (native XML for web services and XML tools)
- FITS (standard BINTABLE for SED, to be defined)
- text (simple text table with columns of data and no markup)
- text/html
- graphics; a JPG, GIF etc. representation of the data
- metadata; only the XML metadata.

7 Relationship to general VO data models

The Spectrum model involves objects addressed by the proposed VO Observation and Quantity data models. Although these models have not yet been fully worked out, we may note that a single Spectrum maps to the Observation model, which will include the Curation and Coverage objects. The Flux and the spectral coordinate entries together with their associated errors and quality will be special cases of the Quantity model, as will the simpler individual parameters. The field structure presented here is consistent with current drafts of the models.

8 Serializations

8.1 FITS serialization

We define a reference serialization of this data model as a FITS binary table. The format is similar in spirit to the X-ray PHA type II dataset. It represents each spectrum or photometry point as a single row of the table. Variable-length arrays are used to contain the array quantities.

Here we give the mapping of data model fields to FITS columns and keywords. For each column, the standard keywords TTYPE_n, TUNIT_n, TFORM_n should be provided. In addition, we define a new keyword TUCD_n which should contain the UCD1+ string for each column. Order of keywords and columns is not significant, except that it is strongly recommended that RA and Dec be in adjacent columns or keywords.

We adopt the convention that columns which are constant (same value for all rows) may if desired be omitted and the value given as a keyword instead. (e.g. the column TTYPE_n=’INSTRUME’ is replaced by a keyword INSTRUME = ’value’). This is a trivial overhead in the FITS reading interface.

We add a new keyword VOCLASS to describe the VO object represented by the FITS table.

FITS keyword	Data model field	Value
VOCLASS	Spectrum	Spectrum (required)
VOCREATE	SED.Creator	
VOCRID	SED.CreatorID	
DATE	SED.Date	
BANDPASS	SED.Bandpass	
OBJECT	Target.Name	
SRCCLASS	Target.Class	
SPECTYPE	Target.spectralClass	
REDSHIFT	Target.redshift	
DS_IDENT	DataID.DatasetID	
TITLE	DataID.Title	
VERSION	DataID.Version	
VOPUB	Curation.Publisher	
VOPUBID	Curation.PubID	
VOLOGO	Curation.Logo	
CONTRIB _n	Curation.Contributor	

The following fields are scalar columns. *They may be used as keywords if they are constant for the whole table.* The Position.Type field is not explicitly serialized: the names of the two columns used for Position are used to infer the type.

FITS TTYPE _n	Data model field	
FLUX_UCD	Flux.ucd	
EQUINOX	Frame.Sky.Equinox	2000.0 (required)
RADECSYS	Frame.Sky.System	either ICRS or FK5
TIMESYS	Frame.Time.Type	TT (required)
TIMEUNIT	(HEA convention)	
MJDREF	Frame.Time.Zero	(required)
SPECSYS	Frame.SpectralCoord.System	(see Greisen et al)
INSTRUME	DataID.Instrument	
COLLECT _n	DataID.Collection	
RA_NOM	Coverage.Location.Sky	
DEC_NOM	Coverage.Location.Sky	
RA_TARG	Target.pos	
DEC_TARG	Target.pos	
SPECBAND	Coverage.Location.Spectral	
APERTURE	Coverage.Region.Sky	
EXPOSURE	Coverage.Exposure	
SYS_ERR	Flux.SysError	
SKY_RES	Spatial.Resolution	
TSTART	Coverage.Region.Time	
TSTOP	Coverage.Region.Time	

The following fields are variable-length array columns. Rather than have a single spectral coordinate tag, we use the different tags WAVE, ENERGY, FREQ to be consistent with WCS paper 3.

FITS TTYPE _n	Data model field
WAVE	SpectralCoord
WAVE_LO	SpectralCoord.BinLow
WAVE_HI	SpectralCoord.BinHigh
ENERGY	SpectralCoord
ENERG_LO	SpectralCoord.BinLow
ENERG_HI	SpectralCoord.BinHigh
FREQ	SpectralCoord
FREQ_LO	SpectralCoord.BinLow
FREQ_HI	SpectralCoord.BinHigh
TIME	Time
TIME_LO	Time.BinLow
TIME_HI	Time.BinHigh
FLUX	Flux
ERR_LO	Flux.ErrorLow
ERR_HI	Flux.ErrorHigh
QUALITY	Flux.Quality
TIME_RES	Time.Resolution
SPEC_RES	SpectralCoord.Resolution

Note: The ENERG_LO, ENERG_HI columns are already used in the X-ray community.

We summarize this with a full sample FITS extension header.

```

XTENSION= 'BINTABLE'           / binary table extension
BITPIX  =          8           / 8-bit bytes
NAXIS   =          2           / 2-dimensional binary table
NAXIS1  =          80          / width of table in bytes
NAXIS2  =         2048         / number of rows in table
PCOUNT  =          208         / size of special data area
GCOUNT  =           1         / one data group (required keyword)
TFIELDS =          14         / number of fields in each row
EXTNAME = 'SPECTRUM'          / name of this binary table extension
VOCLASS = 'Spectrum'         / VO Data Model
VOCREATE= 'MMT Archive'      / VO Creator
VOCRID  = 'ivoa://cfa.harvard.edu' / VO Publisher ID URI
DATE    = '2004-08-30T14:18:17' / Date and time of file creation
BANDPASS= ' '                / SED.Bandpass
RA_NOM  =          233.72789197 / [deg] Nominal RA
DEC_NOM =          23.49792615  / [deg] Nominal Dec
OBJECT  = 'ARP 220'          / Source name
SRCCLASS= 'Galaxy'           /
SPECTYPE= 'ULIRG'           /
RA_TARG =          233.73791700 / Observer's specified target RA
DEC_TARG =          23.50333300 / Observer's specified target Dec
REDSHIFT=          0.01812     / Emission redshift
DS_IDENT= 'cfa://whatever'    /
TITLE   = 'Observations of Merging Galaxies' /
VERSION = 2                  / Reprocessed 2004 Aug
VOPUB   = 'Cfa Archive'      / VO Publisher authority
VOLOGO  = 'http://cfa.harvard.edu/vo/cfalogo.jpg' / VO Creator logo
FLUX_UCD= 'phot.fluxDens;em.wl' /
EQUINOX = 2.0000000000000E+03 / default
RADECSYS= 'ICRS'           / default
TIMESYS  = 'TT'            / Time system
TIMEUNIT = 's'             / Time unit
SPECSYS  = 'TOPOCENT'      / Wavelengths are as observed
MJDREF   = 0.0             / MJD zero point for times
SPEC_RES =          5.0     / [Angstrom] Spectral resolution
SKY_RES  =          1.0     / [arcsec] Spatial Resolution
TELESCOP= 'MMT'           / Telescope
INSTRUME = 'BCS'          / Instrument
FILTER    = 'G220'        / Grating
TTYPE1    = 'INSTRUME'     / Instrument ID
TFORM1    = '8A'          / format of field
TTYPE2    = 'FILTER'      / Filter ID
TFORM2    = '8A'          / Format
TTYPE2    = 'RA'          / Position RA of aperture center
TFORM2    = '1D'          /
TUNIT2    = 'deg'         /
TTYPE3    = 'DEC'         / Position Dec of aperture center
TFORM3    = '1D'         /
TUNIT3    = 'deg'         /
TTYPE4    = 'APERTURE'    / Aperture diameter (physical or extraction)
TFORM4    = '1E'         /
TUNIT4    = 'arcsec'      /
TTYPE5    = 'TIME'       /
TFORM5    = '1D'         /
TUNIT5    = 'd'          / MJD days
TTYPE6    = 'EXPOSURE'    / Effective exposure time
TFORM6    = '1E'         /
TUNIT6    = 's'          /
TTYPE7    = 'SYS_ERR'     / Fractional systematic error
TFORM7    = '1E'         /
TUNIT7    = 's'          /
TTYPE8    = 'WAVE'        / Wavelength
TFORM8    = '1PE'        /
TUNIT8    = 'Angstrom'   /
TTYPE9    = 'WAVE_LO'     /
TFORM9    = '1PE'        /
TUNIT9    = 'Angstrom'   /
TTYPE10   = 'WAVE_HI'     /
TFORM10   = '1PE'        /
TUNIT10   = 'Angstrom'   /
TTYPE11   = 'FLUX'       /
TFORM11   = '1PE'        /
TUNIT11   = 'erg cm^-2 s^-1 Angstrom^-1'
TTYPE12   = 'ERR_LO'     /
TFORM12   = '1PE'        /
TUNIT12   = 'erg cm^-2 s^-1 Angstrom^-1'
TTYPE13   = 'ERR_HI'     /
TFORM13   = '1PE'        /
TUNIT13   = 'erg cm^-2 s^-1 Angstrom^-1'
TTYPE14   = 'QUALITY'     /
TFORM14   = '1PI'        /

```

The data would look like

```
MMT/BCS  G300  233.73791  23.50333  2.0 52984.301203 1500.0 0.15  3200.0 3195.0 3205.0 1.48E-12 2.0E-14 2.0E-14  0
                                                3210.0 3205.0 3215.0 1.52E-12 3.0E-14 3.0E-14  0
                                                3220.0 3215.0 3225.0 0.38E-12 0.38E-12 0.0  0
                                                3230.0 3225.0 3235.0 1.62E-12 3.0E-14 3.0E-14  0
MMT/BCS  G300  233.73792  23.50334  2.0 52102.103211 1480.0 0.15  3200.0 3195.0 3205.0 3.48E-12 2.0E-14 2.0E-14  0
                                                3210.0 3205.0 3215.0 2.52E-12 3.0E-14 3.0E-14  0
                                                3220.0 3215.0 3225.0 1.38E-12 0.38E-13 0.38E-13 1
                                                3230.0 3225.0 3235.0 1.62E-12 3.0E-14 3.0E-14  0
FLW0/4S  B    233.73791  23.50333  4.5 48776.001234  300.0 0.05  4400.0 4200.0 4600.0 1.82E-12 1.2E-14 3.1E-14  0
FLW0/4S  V    233.73791  23.50333  4.5 48776.012012  300.0 0.05  5400.0 5200.0 5600.0 3.82E-12 1.3E-14 1.5E-14  0
FLW0/4S  R    233.73791  23.50333  4.5 48776.019013  240.0 0.05  7000.0 6200.0 7500.0 5.82E-12 1.3E-13 2.1E-13  0
FLW0/4S  I    233.73791  23.50333  4.5 48776.024988  240.0 0.08  9000.0 8200.0 9900.0 8.12E-12 3.3E-13 3.4E-13  0
```

8.2 Alternate FITS serialization

Some implementors may find variable length arrays inconvenient. An alternate ‘relational normalization’ serialization, considered but currently rejected by the VO group, is recorded here for information. It would consist of two binary or ASCII tables, each with only scalar columns. The Spectrum table would contain the same scalar columns as above, but the variable array columns would be moved to a second table, called SpecData, and used as scalar columns arranged vertically. A new integer column ”OBS_ID” would be added to each table; the values in the Spectrum table would be unique and the corresponding values in the SpecData table would indicate which data points belonged to which rows in the Spectrum table.

Another approach would be to have one FITS HDU per spectrum or photometry point. However this was rejected as unworkable, as the overhead of 5760 bytes (2 FITS blocks) per photometry point would inflate the data for the photometry-only SED case by factors of around 50-100.

8.3 VOTable Serialization

The VOTable version of Spectrum will represent an SED by a series of tables, one for each individual spectrum. The data model fields described above as arrays map to VOTable FIELDS, while the remaining fields map to PARAM.

We use nested GROUP constructs to delimit data model objects within the main object, and PARAM and FIELD tags for attributes. Names of fields and parameters are left to the data provider. The utype and ucd attributes are used to denote data model and UCD tags. I have not included NAME fields for the PARAM; the name fields are free to be whatever the data provider wants, allowing compatibility with local archive nomenclature.

```
<VOTABLE>
<RESOURCE>
<TABLE utype="Spectrum"> # Defines local namespace for utype
<PARAM utype="date">2004-05-10</PARAM>
<PARAM utype="creationtype">Archival</PARAM>
<GROUP utype="SpectralCoord">
  <FIELD utype="Value" ucd="em.wavelength" unit="Angstrom"/>
  <FIELD utype="BinLow" ucd="stat.min,em.wavelength" unit="Angstrom"/>
  <FIELD utype="BinHigh" ucd="stat.max,em.wavelength" unit="Angstrom"/>
```

```

</GROUP>
<GROUP utype="Flux">
  <FIELD utype="Value" ucd="phot.flux,em.wavelength" unit="erg cm-2 s-1 Angstrom-1">
  <FIELD utype="ErrorLow" unit="erg cm-2 s-1 Angstrom-1">
  <FIELD utype="ErrorHigh" unit="erg cm-2 s-1 Angstrom-1">
  <FIELD utype="Quality">
# In this case Resolution is demoted from Field to Param since it is constant
  <PARAM utype="Resolution" unit="Angstrom">14.2</PARAM>
  <PARAM utype="SysError" unit="">0.05</PARAM>
</GROUP>
<GROUP utype="Coverage">
  <GROUP utype="Location">
    <GROUP utype="Sky">
      <PARAM utype="Value" ucd="pos.eq" unit="deg">132.4210, 12.1232</PARAM>
    </GROUP>
    <GROUP utype="Time">
      <PARAM utype="Value" ucd="time.obs">52148.3252</PARAM>
    </GROUP>
  </GROUP>
  <GROUP utype="Extent">
    <PARAM utype="Sky" ucd="pos.region.diameter" unit="arcsec">20</PARAM>
    <PARAM utype="Time" ucd="time.interval,phot.spectrum" unit="s">1500.0</PARAM>
    <PARAM utype="Spectral" ucd="instr.bandwidth" unit="Angstrom">3000.0</PARAM>
  </GROUP>
</GROUP>
<GROUP utype="Target">
  <PARAM utype="Name">Arp 220</PARAM>
  <PARAM utype="pos" unit="deg">233.737917 23.503330</PARAM>
</GROUP>
<GROUP utype="Frame">
  <GROUP utype="Sky">
    <PARAM utype="Equinox" ucd="time.equinox,pos.eq">2000.0</PARAM>
    <PARAM utype="System" ucd="frame.pos.system">ICRS</PARAM>
  </GROUP>
  <GROUP utype="Time">
    <PARAM utype="Type" ucd="frame.time.scale">UTC</PARAM>
    <PARAM utype="Zero" ucd="frame.time.zero">0.0</PARAM>
  </GROUP>
  <GROUP utype="SpectralCoord">
    <PARAM utype="Frame.SpectralCoord.System" ucd="frame.em.system">Barycentric</PARAM>
  </GROUP>
</GROUP>
</GROUP>
<GROUP utype="Derived">
  <PARAM utype="SNR">3.0</PARAM>
</GROUP>
<GROUP utype="Curation">

```

```

<PARAM utype="Publisher" ucd="human.publisher,meta.curation">SAO</PARAM>
<PARAM utype="PubID" ucd="meta.curation.pubid">ivoa://cfa.harvard.edu</PARAM>
<PARAM utype="Logo" ucd="meta.curation.logo">http://cfa-www.harvard.edu/nvo/cfalogo.jpg</PARAM>
</GROUP>
<GROUP utype="DataID">
  <PARAM utype="Title">"Arp 220 SED"</PARAM>
  <PARAM utype="Creator" ucd="meta.curation.creator">SAO/FLWO</PARAM>
  <PARAM utype="Date" ucd="time,soft.dataset,meta.curation">2003-12-31T14:00:02</PARAM>
  <PARAM utype="Version" ucd="soft.dataset.version,meta.curation">1</PARAM>
  <PARAM utype="Instrument" ucd="inst.id">BCS</PARAM>
  <PARAM utype="Collection" ucd="inst.filter.id">G300</PARAM>
</GROUP>
<TABDATA>
# Note slightly nonlinear wavelength solution
# Second row is upper limit
# Third row has quality mask set
<TR><TD>3200.0<TD>3195.0<TD>3205.0<TD>1.38E-12<TD>5.2E-14<TD>6.2E-14 <TD>0
<TR><TD>3210.5<TD>3205.0<TD>3216.0<TD>1.12E-12<TD>1.12E-12<TD>0<TD>0
<TR><TD>3222.0<TD>3216.0<TD>3228.0<TD>1.42E-12<TD>1.3E-14<TD>0.2E-14<TD>3
</TABDATA>
</TABLE>
</RESOURCE>
</VOTABLE>

```

8.4 Direct XML serialization

I'll leave this bit for Tamas.

References

Greisen, EW, Valdes F G, Calabretta M R and Allen S L 2003, in prep., www.aoc.nrao.edu/~egreisen/scs_1
Hanisch, R., (ed)., Resource Metadata for the VO 2003 Oct 16 Draft, Do Not Reference (hah!), www.ivoa.net/Documents/WD/ResMetadata/WD-RM-20031016.pdf