International

Virtual

Observatory

Alliance

IVOA Spectral Energy Distribution (SED) Data Model

Version 1.0
IVOA Working Draft, 2012 October 15

Abstract

We define a data model for Spectral Energy Distributions (SED) by extending the existing IVOA Spectral data model. A top level SED dataset is defined which describes the overall SED object. This may include a uniform SED computed from the individual SED segments (observations contributing to the SED). An optional aggregate SED is also defined which includes the individual SED segments, to support advanced use cases such as SED computation and editing.

Status of This Document

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 Recommendations and other technical documents* can be found at http://www.ivoa.net/Documents/.

Acknowledgements

Experience with the VAO SED project and manipulation of spectral energy distributions from NED (IPAC) contributed greatly to this standard by providing a real world use case and prototype for SED publication, creating, editing, and analysis. The Spanish VO (SVO) has contributed additional use cases to help define the data model described herein.

Contents

1   Introduction

The IVOA has identified creation and analysis of spectral energy distributions as a key science capability desired in the VO. In this document we present a proposed abstraction for spectral energy distribution data and serializations in VOTABLE and FITS for use as a standard method of SED data interchange.

1.1   Change Log

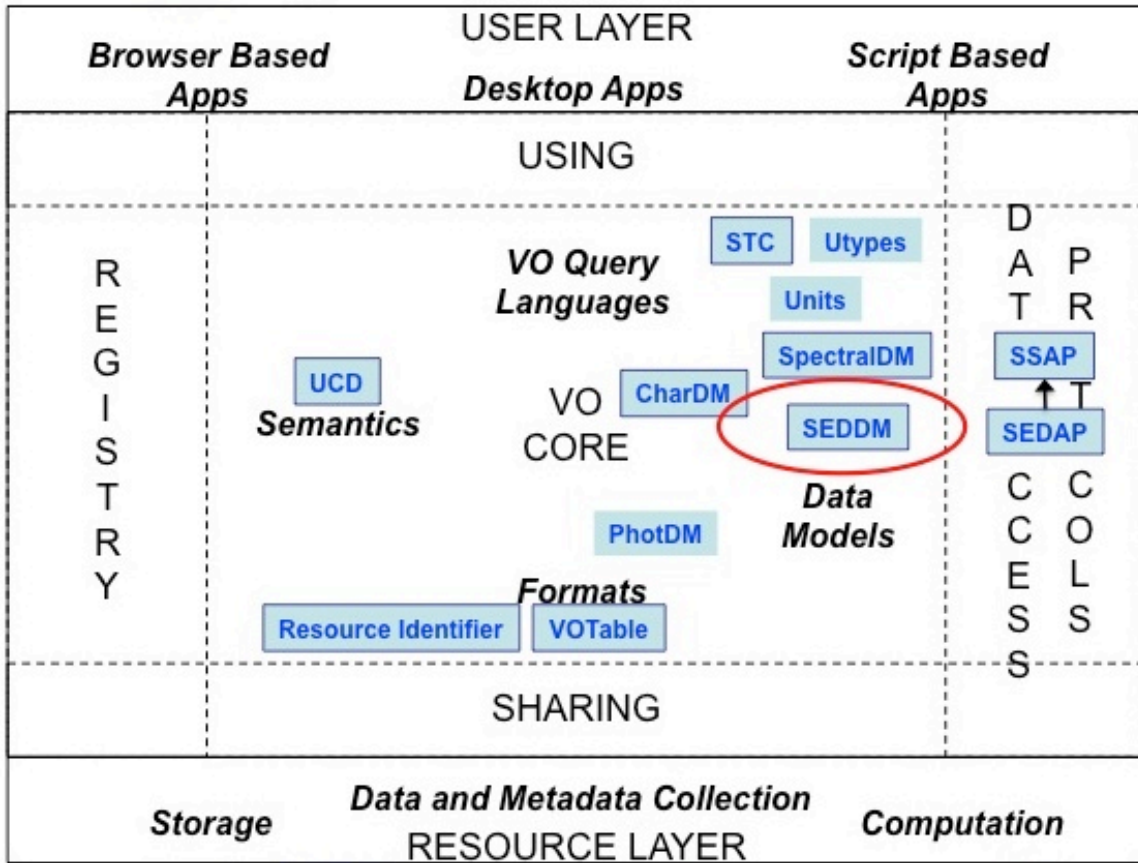|	
2010 Nov 4	Initial draft.
2011 May 15	Expanded draft addressing uniform and aggregate SEDs and serializations.
2012 May 15	Updated to support the draft SpectralDM 2.0 and Utype usage.
2012 Oct 15	Further updates based upon most recent Spectral DM 2.0.

1.2 IVOA Architecture Context

Figure 1 - The SED data model in the IVOA architecture.

The SED data model is derived from the IVOA Spectral data model and shares the same dependencies as the Spectral DM.

2 Spectral Energy Distributions

A spectral energy distribution (SED) represents the energy output of an astronomical source over a wide spectral range. It differs from a traditional astronomical spectrum in that it is a composite data set, which may be derived from (or in the case of a model SED, compared to) multiple instruments, telescopes and observational epochs.

Individual observations used to create a SED may be single photometry points or 1-D spectra with flux values at many spectral coordinates. Each of these observations has its own metadata describing the observational conditions, etc. We refer to these individual observations as segments.

We distinguish two main varieties of SED:

- An aggregate SED, consisting of a logical collection of segments together with each segment's metadata; no attempt is made at this stage to reconcile the data in the different segments.

- A uniform (or rebinned) SED, consisting of a single merged segment. To go from an aggregate to a uniform SED requires transforming each segment to provide uniform units for the spectral axis and for flux. To go from an aggregate SED to a binned SED further requires rebinning the data on a fixed wavelength or frequency grid and handling overlaps (for example, you may have a V photometry point overlapping an optical spectrum, or multiple V photometry points at different times).

Some SED science is done with uniform SEDs, which are the simplest to handle; some science requires you to work with the aggregate SED directly. This is the usual methodological dichotomy in astronomy between correcting your data to compare directly with a model, or folding your model through a simulation of the observation process and comparing the result with the data. Both kinds of SED may therefore have model (theoretical) counterparts, but we normally think of a model SED as being of the uniform kind.

One can also have a SED described by an analytic function; we consider this to be outside our scope, and assume that such representations will be instantiated as a finite, discrete array of flux and spectral coordinate values.

2.1 Scope

In modern instruments, photometry and spectra are often obtained by extraction from two- or higher-dimensional detectors, including photometry points from CCD images or time series observations, and spectra from long slit images, Echelle datasets and spectral data cubes. The process of reducing such data to simple sets of flux-versus-spectral-coordinate arrays is out of our scope; we are only attempting at this stage to describe the extracted data, usually in calibrated form.

3 SED Data Model

3.1 Primary SED Object or Dataset

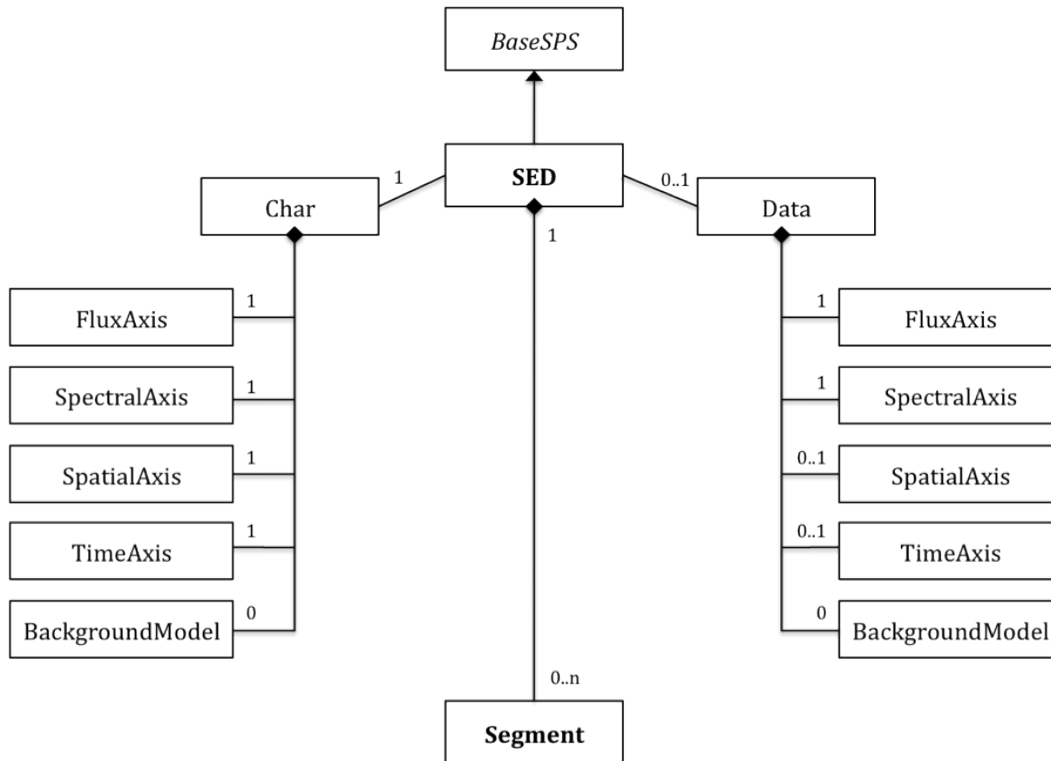SED extends the *BaseSPS* class of the Spectral data model (SpectralDM) [*Ref XX*] as shown in Figure 2 below.

Figure 2 - The SED data model

Our fundamental SED data model is as follows:

- SED extends the *BaseSPS* class to represent a SED object or dataset.

- SED defines additional SED-specific metadata (described further below) and otherwise inherits all metadata from the core Spectral data model.

- The data element of a SED may consist of the uniform SED (shown as *Data* in Figure 2), an aggregation of segments (the aggregate SED), or a combination of the two.

When we refer to the "overall" SED object or dataset we mean the object labeled "SED" in Figure 2, including any associated elements such as Char, Data or Segment. The terms "uniform SED" or "aggregate SED" refer to the same object but restrict the data element referenced to either the uniform or aggregate SED.

SED and Spectrum share the same core model and a SED may be treated as a special case of a Spectrum by Spectrum-aware software provided the extensions defined by SED are ignored (generic spectral analysis software will not

understand the aggregate SED hence the data for the uniform SED must be included to manipulate a SED as a spectrum).

3.2 Aggregate SED

The aggregate SED is a sequence of segments, one for each observation contributing to the overall spectral energy distribution.
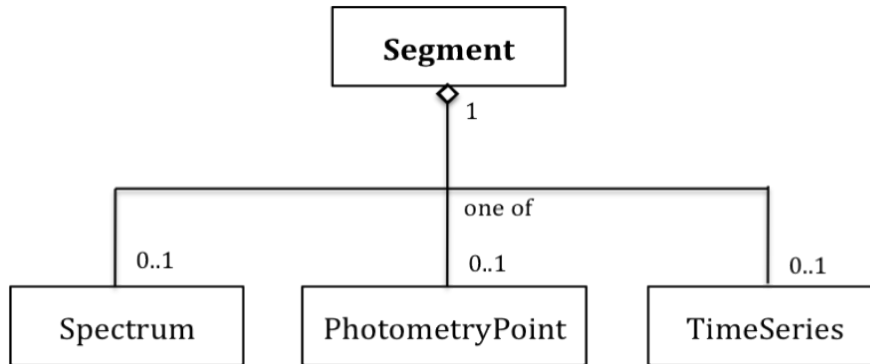

Figure 3 - Structure of a Segment in the Aggregate SED

For an aggregate SED:

- Each segment contains a photometry point, a spectrum, or a time series, or a URI pointing to such an instance stored externally.

- *Segment* is a generic container for the contained object instance. The contained object (photometry point, spectrum, or time series) may be extracted as a dataset or may be referenced externally, and may be manipulated by generic (non-SED) software without knowledge of how the object instance is used within the SED.

- A *Spectrum* object is described by the Spectrum extension to the IVOA Spectral data model with the spectral coordinate as the independent variable.

- A *PhotometryPoint* object is described by the IVOA Spectral data model extension for photometry point data. The IVOA Photometry model, which is integrated into the Spectral data model, defines the photometric band observed.

- A *TimeSeries* object is described by a time series extension to the IVOA Spectral data model with the time coordinate as the independent variable. *TimeSeries* is a placeholder for a future Time Series data model, which is still to be defined. A time series as used for a SED segment may measure flux in multiple photometric bands. A light curve is a special case of a time series measuring flux in a single band.

Spectrum, *TimeSeries*, *PhotometryPoint*, and *SED* are all VO dataset types derived from and extending the core Spectral data model.

If a SED dataset contains both a uniform SED and the aggregate SED data from which the uniform SED was derived, it is possible to associate individual segments of the aggregate SED with data points within the uniform SED. While in some cases there may be a simple (e.g. one to one) mapping between the two, in general a single point within the uniform SED may be a complicated function of points from multiple segments in the aggregate – for example time averaging of photometry, or combining overlapping spectral segments and/or photometry points.

3.3 Namespace

The data model namespace "sed" is to be associated with objects defined by the SED data model. Usage should conform to application of the relevant IVOA conventions as defined by the Spectral data model.

Dataset instances within a segment should conform to the specification for the specific class of data contained therein. For example, a Spectrum instance stored in a segment would use the namespace "spec" internally for any Spectrum-defined metadata.

3.4 Axis Requirements

The requirements for the Char and Data elements of the SED dataset are as follows:

Flavor	Char Requirement	Data Requirement
SpectralAxis	Exactly one must be defined.	Exactly one must be defined.
FluxAxis	Exactly one must be defined.	Exactly one must be defined.
TimeAxis	Exactly one must be defined.	Exactly one may be defined.
SpatialAxis	Exactly one must be defined.	Exactly one may be defined.
BackgroundModel	None may be defined.	None may be defined.

The requirements for the Data element defined here are for the uniform SED and apply only if a data for the uniform SED is provided; otherwise the entire Data element is omitted. Spectral and flux data axes are required as for Spectrum. Time and spatial data axes are optional, but are permitted to allow specification of the time of observation and observed spatial position for each data point in the uniform SED; if the data point is composite a null value should be given.

3.5 UTypes

We define the UType "SED" for use, as needed, to tag the top-level container node in serializations. A SED instance may use any UType from the base Spectral data model. SED further extends the Spectral model to define additional SED-specific dataset and data element metadata, as defined in the following sections.

3.5.1 Dataset Metadata

The SED object or dataset should include the following global UTypes in addition to those defined Spectral DM (mandatory SED dataset fields are shown in bold text).

Dataset.dataModel	"SED-1.0"
Dataset.type	"SED"
Dataset.length	Number of data points in the uniform SED

SED.uniformSed	Data for the uniform SED is included ("true" or "false")
SED.aggregateSed	Data for the aggregate SED is included ("true" or "false")
SED.nSegments	Number of segments in the aggregate SED
SED.aggregateSedURI	URI to aggregate SED if applicable (FITS keyword: SEDORIG)

If the SED has a uniform SED data element then *SED.uniformSed* will be "true" and *Dataset.length* will be nonzero. If data for the aggregate SED is included in the dataset then *SED.aggregateSed* will be "true" and *SED.nSegments* will specify the number of segments. If data for the aggregate SED is available externally then *SED.aggregateSedURI* should point to an external SED dataset instance containing the aggregate SED and *SED.nSegments* should again be specified. Regardless of whether an aggregate SED is available the number of observations or other data sources contributing to the SED must be specified in *SED.nSegments* (for something like a theoretical model there would be a single source but no aggregate SED).

All attributes of the Spectral data model refer to the overall SED dataset, e.g., the spatial, spectral, and time coverage of the overall SED should be characterized, and the dataset identification, curation, and access groups should refer to the overall SED dataset.

3.5.2 Data-Element Metadata

Additional metadata beyond that defined by the generic SpectralDM is required to describe the data in the uniform or binned SED since this data is derived from segment information. Metadata may be SED-specific (*e.g., SegmentType*), may be segment-specific metadata carried over from the SpectralDM, or may be custom metadata defined by the SED creator. Any attribute of the SpectralDM may be included as a data column in the uniform SED. Any metadata included in data columns refers to the individual SED data point and not to the overall uniform SED.

SED defines the following additional UTypes to describe the uniform SED. Mandatory elements are shown in bold.

Data.segmentType	One of "PhotometryPoint", "TimeSeries", "Spectrum", or "Composite".
Data.segments	The segment number, a comma-delimited list of segment numbers, or NULL if a more complex combination of segments are used to compute the given photometric value. The first segment of the aggregate SED is number one.
Data.comments	Additional information about the computed data value or segment.

SegmentType is the same as *Spectrum.Type* for a reference to a single segment, but should have the value "composite" if multiple segments were used to compute the data point of the uniform SED. Where possible *segments* should specify the segment or segments used to compute the given data point. If no valid value can be specified for a given metadata attribute, for example in the case of a composite segment type, then a null value should be given.

3.5.3 Custom Metadata

The data provider may define additional custom metadata to describe their SED data products. These custom metadata attributes may differ for different data providers or data collections.

Custom metadata must follow the rules defined for extending VO data models defined using the UType metadata extension mechanism as summarized in the Spectral data model document.   Custom data provider-defined metadata attributes must not share the same namespace as the standard SpectralDM and SED data models. In particular, any UTypes for custom data provider-defined attributes must not use the "sdm", "sed", "spec" or "phot" namespace prefixes. Services and client software should if possible preserve any custom metadata and pass it through to applications.

Custom metadata may be inherently distinct, *e.g.*, specific to a given SED building application or data provider, or may be more general and hence a candidate for eventual inclusion in the standard model. Reserving the data model namespace to only metadata defined by the standard model is necessary to avoid name collisions and to allow changes as new attributes are added to the standard model, as well as make it clear to a data consumer where a given bit of metadata originates and is defined.

3.6 UCDs and Units

In addition to summarizing SED UTypes, Appendix A provides the appropriate UCD string to use for various model fields. All other UCDs and units should be as defined by the SpectralDM or the specification for the type of data object stored in a segment.

3.7 Required Fields

[*TBD*]

4 SED Serialization

We define two serializations of the SED data model: VOTable and FITS binary table.

4.1 VOTable Serialization

A SED serialized as a VOTable is a dataset of type "SED", containing at least one RESOURCE element that in turn contains an initial TABLE element describing the overall SED dataset. TABLEDATA is included only if data for the uniform SED is provided.

Additional TABLE elements may be provided for segment data. Each TABLE element corresponds to a single segment and contains an instance of a PhotometryPoint, Spectrum, or TimeSeries.

- A SED dataset that includes data only for the uniform SED will contain a single TABLE element with *Dataset.type* set to "SED" and *SED.uniformSed* set to "true". A reference to external aggregate SED data may optionally be provided.

- A SED dataset including data only for the aggregate SED will contain an initial TABLE element of type "SED" defining metadata for the overall SED dataset but omitting TABLEDATA, followed by *nSegments* TABLE elements, one for each segment.

- A SED dataset containing both the uniform SED as well as the corresponding aggregate SED will contain an initial TABLE element of type "SED" including data for the uniform SED, followed by *nSegments* TABLE elements, one for each segment.

An example follows, showing only the major VOTable elements [*legacy xmlns elements are included as a placeholder until we define a new mechanism to define the data model namespace*].

```
<?xml version="1.0" encoding="UTF-8"?>
<VOTABLE version="1.2"
   xmlns:xsi=http://www.w3.org/2001/XMLSchema-instance
   xsi:noNamespaceSchemaLocation="xmlns:http://www.ivoa.net/xml/VOTable/VOTable-1.1.xsd"
   xmlns:sed=http://www.ivoa.net/xml/SEDDM/v1.0
   xmlns:sdm=http://www.ivoa.net/xml/SpectralDM/v2.0
   xmlns="http://www.ivoa.net/xml/VOTable/v1.1">
<RESOURCE utype="sed:SED">
   <TABLE utype="sed:SED"></TABLE>
     <PARAM name="DatasetType" utype="sed:Dataset.type " value="SED" ... </PARAM>
   <TABLE utype="sdm:PhotometryPoint"></TABLE>
     <PARAM name="DatasetType" utype="sdm:Dataset.type" value="PhotometryPoint" ...
   <TABLE utype="sdm:Spectrum"></TABLE>
     <PARAM name="DatasetType" utype="sdm:Dataset.type" value="Spectrum" ...
 </RESOURCE>
 </VOTABLE>
```

At the top level, both RESOURCE and the initial TABLE element specify the overall VO Dataset (object) type, in this case SED; for other types of datasets we would have Spectrum, TimeSeries, and so forth. Individual segments specify their object type separately as shown, and can be referenced or extracted as a standalone dataset unchanged. In this example all allowable dataset or object types are defined by the SED and SpectralDM (sdm) data models.

4.2 FITS Serialization

The overall SED, including data for the uniform SED if provided, is serialized according to the Spectral data model as a FITS binary table. Data for the aggregate SED can either be appended as additional FITS extensions, or a link can be provided pointing to an external SED dataset containing data for the aggregate SED. [*details and/or an example should be added, but this looks pretty straightforward given a mapping to FITS keywords for the few additional UTypes defined by the SED DM*]

Appendix A: "Appendix Title"

Insert appendix here

References

[*TBA*]

[1] R. Hanisch, *Resource Metadata for the Virtual Observatory* , http://www.ivoa.net/Documents/latest/RM.html
[2] R. Hanisch, M. Dolensky, M. Leoni, *Document Standards Management: Guidelines and Procedure* , http://www.ivoa.net/Documents/latest/DocStdProc.html