International
Virtual
Observatory
Alliance

# MANGO: A Component and Association Based Model for representing data for astronomical sources

# Version 1.0

## IVOA Working Draft 2020-07-24

Working group
DM
This version
http://www.ivoa.net/documents/MANGO/20200724
Latest version
http://www.ivoa.net/documents/MANGO
Previous versions
This is the first public release
Author(s)
François Bonnarel, Gilles Landais, Laurent Michel, Jesus Salgado, Mireille Louys, Marco Molinaro
Editor(s)
Laurent Michel

## Abstract

The MANGO model proposes a flexible way to expose data related to astronomical source objects in an interoperable way. It takes into account the huge diversity of source data in terms of feature description, format and usage. The MANGO model attaches identifiers on astronomical sources and associates to each a flexible set of parameters (e.g. observed physical quantities) and other information like e.g. spectra, time series or preview images. Parameters usually appear in the columns of a source catalogue. Additional data products are bound to the source to contribute to the science analysis and enhance data understanding. Mango object parameters are built upon classes or extended classes of the IVOA Measure and Coordinates data model. Associated data can be simple URLs, VO service endpoints or VO data model instances. The roles of both parameters and associated data are qualified by semantic tags


## 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/.

## Contents

## Acknowledgments

We would like to thank all those who took the time to present their own use cases (INAF, CDS, CFA, ESAC) on which the model has been built. We would also like to thank all the people having tested MANGO on their own data.

## Model Name

This model was initially named with a very explicit but hard to remember acronym, `CAB-MSD` standing for Component and Association Based Model for Source Data. We decided later to rename it `MANGO` with reference to the inside out MANGO picture used to introduce the model in Groningen. As the tradition requires that such unexpected names are acronyms, let's assume that `MANGO` stands for Metadata ANnotation for Generic Objects (in astronomy).

## Conformance-related definitions

The words "MUST", "SHALL", "SHOULD", "MAY", "RECOMMENDED", and "OPTIONAL" (in upper or lower case) used in this document are to be interpreted as described in IETF standard RFC2119 (**?**).

The *Virtual Observatory (VO)* is a general term for a collection of federated resources that can be used to conduct astronomical research, education, and outreach. The International Virtual Observatory Alliance (IVOA) is a global collaboration of separately funded projects to develop standards and infrastructure that enable VO applications.

## 1 Introduction

Modeling data collected to study astronomical source objects has been a long term concern for the DM working group and more generally for the IVOA. In the past years, there were some proposals to design a global model for sources (Salgado and Lemson et al., 2016) as well as for catalogs (Osuna et al., 2006). Other proposals, more model-agnostic, were focused on the data annotation in VOTables (Demleitner and Ochsenbein et al., 2016) (Derriere, 2016). In this case the goal was no longer to design a source model but to provide a complete description of individual quantities (positions, velocity, fluxes, magnitudes…). None of these proposals have come to completion.

The source DM issue resurfaced at the spring 2018 Interop in Victoria during an hands-on session focusing on the tools available to work with VO data models and especially with VO-DML. The goal of this session was to

PDF fallback.
Sorry - your ImageMagick (convert) does not support SVG import. If on Linux, installing librsvg2-bin should remedy this. Otherwise, please commit your SVG and ask the ivoatex creators to do the the conversion.

*Figure 1:* Architecture diagram for this document

annotate data from different origins in order to make them interoperable with each other. One of the main concerns outside de tools necessary to workout this notation was the lack of models for source objects. This is a big paradox in the VO world: source data which represent the basic building blocks of astronomers' work, is not modelled. This paradox can be explained by the fact that the observation of source objects is multi-faceted. In a general way, the way features for source data are described and organised depends on the targeted science case. Principal investigators and archive designers set up the data profile and structure it according to this goal which varies from one project to another. Therefore this diversity cannot be served by a single static data model describing a source item for all possible cases. Having a global source model would lead to a very complex solution not usable in practice.

This standard proposes to overcome this paradox and presents a template model gathering independent components from VO existing models together with VO data products and files embedded on demand in a container. The template supports fine grain association by composing classes as well as coarse grain relations to data products and files distributed within projects archive. Mango is not designed to describe what a source is but to help clients to discover and to understand the quantities available for a particular source instance.

## 1.1 Role within the VO Architecture

Fig. 1 shows the role this document plays within the IVOA architecture (Arviset and Gaudet et al., 2010).

# 2 Representing observed astronomical objects : Use Cases and Requirements

## 2.1 Use Cases

The following uses-cases have been collected since 2019 from representatives of various astronomical missions, archive designers and tools developers. The contribution was totally open. This gave a good picture of the needs but we do not pretend that everything will be supported by this first version.

All the use-cases summarized below are detailed in appendix.

### 2.1.1 Gaia

Gaia mission is producing the largest and more precise 3D map of our galaxy.

A non-exhaustive list of properties required for Gaia use cases would be composed of:

- identifier
- sky reference position
- proper motion
- parallax and distance
- source extension
- radial velocity
- redshift
- photometry
- date of detection
- correlation
- multiple detection

### 2.1.2 Euclid

Euclid telescope has been designed to unveil some of the questions about the dark Universe, including dark matter and dark energy, what would include, e.g. quite accurate measurements of the expansion of the Universe.

- identifier
- sky position

- correlation with other catalogues
- photometry (ground + satellite )
- morphology class
- redshift
- photometric redshift

### 2.1.3 Exoplanets

Annotation of (exo-)planetary records in catalogues requires some specific metadata or model.

This was initially made clear when trying to allow discoverability of exoplanets time series (Molinaro and Benatti et al. (2020)) where (exercising the ObsCore model with specific use cases) information like stellar host characteristics together with stellar system details turned out to be required for a proper description of the catalogue records.

The use cases identified, besides some refinement needed in the ObsCore time axis annotation, the following metadata:

- the degree of confidence in the detection: exoplanets candidates w.r.t. confirmed ones, plus last update of the record content ;
- the method used in the discovery (since it affects the available stellar system description parameters);
- a set of stellar host characteristics (besides sky coordinates): activity, mass, type, metallicity, age, some systemic values, like the global RV (radial velocity) of the system, and so on;
- (exo-)planet parameters, like mass, orbital period, orbit's eccentricity, RV semi-amplitude, time at periastron (for RV detections) or central transit time (for transit method), longitude of periastron, and so on.

The scenario gained some further complexity, but also better stated the idea of modelling datasets and catalogue records in the (exo-)planetary systems subdomain, incorporating requirements from exoplanets atmosphere simulations and (first efforts) observations.

Specific metadata additions were not specified exactly, but a draft model to describe stellar systems was developed, specifically trying to solve the issue in adding metadata elements to describing orbiting celestial objects (Molinaro and Alei et al. (2019)).

The model identified the main concepts and classes as:

- Celestial Object, typed/subclassed as: Star, Binary Star, Planet, Satellite, Brown Dwarf, Trojan, and so on;

- Orbit, as the class keeping the information needed to describe the orbit of a couple (or more) Celestial Objects.

Specific (exo-)planets metadata included: atmosphere with molecular composition, bulk details (mass, radius, ...).

A catalogue service that could make use of model metadata is the Exo-MerCat (Alei and Claudi et al. (2020)) catalogue (or the other available exoplanets' catalogues), already available in a VO-aware solution[1], but using only simple UCD description of the exposed information.

**TODO:**
mention the involved projects : examples ? GAPS, TESS?

**TODO:**
For GAPS I can add specifics later, if implementation can (re)start

### 2.1.4 Morphologically Complex Structures

The ViaLactea Knowledge Base (VLKB, see Molinaro and Butora et al. (2016)) is a set of data resources and services built up to study the star formation regions and processes in the Milky Way. Besides 2-D images and 3-D radial velocity cubes, the VLKB exposes a bunch of source catalogues. A model that supports description of such catalogues will need a way to describe sources with:

- non-point-like positions;
- extended complex area, possibly as multiple detached areas;
- aggregation of sub-parts (that can be heterogeneous).

### 2.1.5 Chandra Archive

The Chandra Source Catalog(CSC) is the definitive catalog of serendipitous X-ray sources identified in publicly released imaging observations obtained by NASA's ChandraX-ray Observatory (CXO).

The catalog itself consists of approximately 1,700 columns covering properites at the individual observation and stacked analysis levels. Table 3 summarizes some of the basic catalog properties derived from standard CSCView queries.

[1] ivo://ia2.inaf.it/catalogues/exomercat

| |
|---|
| **Per Source:** |
| `Source name` |
| `Source position and position errors` |
| `Significance of the source (signal to noise)` |
| `Likelihood of the source (True, False, or Marginal detection)` |
| `Source extent flag` |
| `Variability flag` |
| `Spectral variability flag` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |
| `Long term (inter-obs) variability probability for each band` |
| `Spectral (hardness ratios) variability for each color` |
| **Per Detection (at the stack level):** |
| `Detection ID` |
| `Detection position and position errors` |
| `Flux significance of the detection (S/N)` |
| `Detection likelihood (True, False, or marginal detection?)` |
| `Source extent code (codification of source extent in different bands)` |
| `Variability flag` |
| `Spectral variability flag` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |
| `Long term (inter-obs) variability probability for each band` |
| `Spectral (hardness ratios) variability probability for each color` |
| **Per Detection (at the observation level):** |
| Note that source detection is done at the stack level, but properties are estimated for the detections at each observation using the detection region from the stack level. |
| `Detection ID` |
| `Detection position and position errors` |
| `Flux significance of the detection (S/N)` |
| `Detection likelihood` |
| `Source extent code (codification of source extent in different bands)` |
| `Variability code (applies to intra-obs only)` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |

*Table 1:* Example Chandra Source Catalog Properties

### 2.1.6 Vizier catalog archive

VizieR provides science ready catalogues coming from space agencies or articles and covering number of different science cases. Published data encompass a very large set of measures (position, photometry, redshift, source type, etc.) depending on their origin. They can result from observations, simulations, models or catalog compilations. Individual Vizier tables can contain data all related to one source (e.g. time series of positions or magnitudes) or to a set of sources (one row per source) or a mix of both.

The Mango model must be able to provide a standard representation of most of the metadata contained in Vizier query responses, whether native or computed by the CDS, simple quantities or associated complex data. Mango is not meant to replace the current management of the meta-data, it is a way to make those meta-data understandable for a wide panel of VO-compliant clients.

### 2.1.7 Client on (MT behalf)

Right now, the meta-data provided within the VOTable allow clients such Aladin or Topcat to run most of the functionalities expected by the user, either for data analysis of plotting. This information is often guess from UCDs, UTypes or columns name. It can also be given by the user. Clients have no expectations of working with full model instances but in some cases models can help to know how quantities in an input table relate to each other.

In most cases this is for visualisation, e.g.:

- what is the sky position for this row (what columns contain latitude and longitude, and what sky system are they in)
- what +/-ERR error bars should I plot for these points (what column is a simple error for column A)
- what error ellipses should I plot for these sky positions (what columns provide ra_error, dec_error, ra_dec_corr, or how can I derive those from columns that do exist)
- where do I get the grid information for a column containing a vector of samples so I can label the X axis of a spectrogram (what column or parameter contains an axis vector matching the sample vectors)
- does this table contain sky positions, or HEALPix tiles, or both? What's the best way to represent it on the sky?
- What is the meaning of such URL found out in a table?s

But there are some other places too:

- how do I propagate this sky position to a future epoch (what columns contain pmra, pmdec, and maybe all the associated errors and correlation coefficients)
- what is the error ellipse/oid to use for a sky/Cartesian crossmatch (what columns provide the relevant errors and, if available, correlations)

This usage shows that MANGO must be designed in a way that individual measurements or quantities can be easily be identified as such and manipulated independently of the whole instance.

### 2.1.8 Xmatch tool

The basic cross-match of two astronomical tables consists in associating pairs of sources – one from each table – fulfilling a given angular distance based criterion. In relational algebra terms, it is a theta-join on a distance predicate.

More generally, a cross-match is the association of sources from different tables given their proximity in an astrometric (but also possibly photometric, statistical, ...) parameter space (Pineau and Derriere et al., 2017) .

If proper motions (plus parallax and radial velocities) are available, the cross-match tool may propagate the positions of each table to a common epoch. It may also take into account positional uncertainties to reject the statistically unlikely associations.

In the latter case (cross-match between two tables taking into account positional errors), the tool needs to be able to retrieve the errors associated to the each position in each table.

UCDs may help in identifying the errors associated to a positional columns as shown in table, but this is not sufficient to table with more complex cases based on multi-parameter cases.

## 2.2 Requirements

### 2.2.1 Parameters and Associated Data

From the use-cases' description, two categories of features must provided or foreseen by the projects:

- The source *parameters* astronomers will investigate for their science. They are measures provided as numerical values or classification tags exposed as numbers or simple strings. Usually on measure corresponds to one individual column or one group of columns .

- The *Associated data* are generally science ready data-products either from the same project, or shared by other projects within the IVOA interoperability framework. They bring a complement of information to interpret the source's parameters under study and compare visually (or computationally) the sky neighbourhood of detected sources, their variation through time or spectral behaviour.

  Referencing such datasets by a URI or by a service endpoint already works for existing VO products. ( It has been promoted and recommended in Obscore / Spectrum Dm, . etc for access URIs, and DataLink for service endpoints REFS). Associated data can also present a complex structure designed to bring a very advanced context of interpretation (CTA model assumptions, energy model for Xray sources, etc. REFS) which need to be described in the attachement.

### 2.2.2 R01: Supported Quantities

- MANGO must provide unique source identifiers.
- MANGO must provide modelling classes for both parameters and associated data.
- The number of parameters attached to a MANGO instance must be free.
- The number of associated data attached to a MANGO instance must be free.

### 2.2.3 R02: Parameters

The concept of `Parameter` matches the concept of measure of the Meas model. MANGO may support Parameter classes that are not Meas classes though.

- MANGO must support explicit classes imported from an IVOA datamodel for the most used parameters.
- MANGO must provide a generic way to support parameters that do not enter the above category.
- MANGO instances must support multiple instances of the same parameter class.
- The presence of any parameter in MANGO instances must be optional.
- MANGO must provide a way to identify the role of each parameter.
- MANGO must provide a way to describe the meaning of flags or qualifiers.
- The role of each parameter should be machine-readable.
- It must be possible to group parameters in a free way. This allows to tag quantities with timestamps or a flags.

### 2.2.4 R03: Associated Data

The notion of associated data relates to any sort of complex data. This can be a pointer to a service or a data set, a data table or other data structure.

- MANGO must support references to external datasets.
- MANGO must support references to external services.

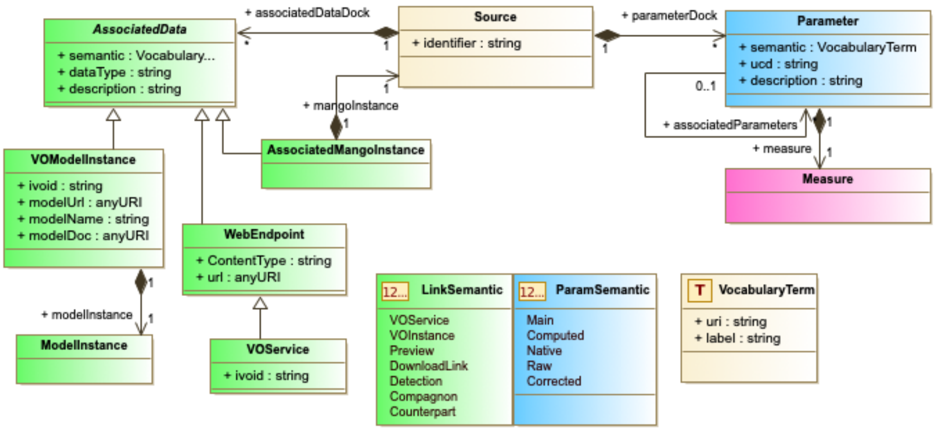


*Figure 2:* MANGO overview

- MANGO must support references to other MANGO instances.
- MANGO must support references to instances of models serialised in VO-DML.
- MANGO instances must support multiple instances of the same associated data class.
- The presence of any associated data in MANGO instances must be optional.
- MANGO must provide a way to identify the role of each associated data in order to explain which purpose is served when associating this data to the source object.
- The role of each associated data should be machine-readable.

## 3 Model Overview

Sky objects are represented by instances of the class `MangoObject` which has only one mandatory attribute, the source `identifier`. It is recommended the identifiers to be unique within a source collection e.g. a catalog.

The `MangoObject` is a dock for all source parameters and for all data associated to that source. The pattern of either parameters or associated data attached to a source is not specified by the model. It depends of the data set on which the model is applied.

The `MangoObject` has one connector for the parameters, the `parameters` relation, and another for the associatedData, the `associatedData` relations. Both connectors have an open-ended cardinality.

Each source parameter is hooked to the `MangoObject` by a wrapping class (`Parameter` ) that contains anything necessary to identify both its nature and its role.

Parameters can be linked together to represent logical parameter sets. These logical sets have no semantic. Their interpretation is in charge of the clients.

Each associated data or data pointer is hooked to the `MangoObject` by a wrapping class (`AssociatedData`) that contains anything necessary to identify its nature and its role.

## 3.1 Parameters

`Parameter` connectors are used to bind measures with the source. A parameters is composed by semantic tags in addition to the measure itself, instance of `Measure`.

`Measure` is an abstract class imported from `meas` model. Concrete classes referenced by a `Parameter` instances are either `meas` built-in classes or `Measure` sub-classes being part of Mango.

### 3.1.1 Parameters Identification

As the parameter set attached to a particular instance is not defined by Mango, the model must provide an accurate parameter description to allow the client to figure out what it can do with each of them. Mango provides 5 description levels for each parameter:

- **Measure class (vodml type)**: measures are modelled by specific `Measure` sub-classes. Knowing that class tells the clients how to interpret the corresponding measure but does not help much to get its role e.g. a position can be either a source position or a pointing direction. Furthermore, unusual measures e.g. magnetic field, are represented by `GenericMeasure` . In that case, the vodml type does not help at all.
- **UCD**: A valid UCD must be attached to each parameter. Mango provides a UCD space for each `Measure` sub-class . UCDs used for specific measures must be compliant with table 2. For generic measures, the UCD choice is in charge of the data provider. In any cases, the consistance between UCDs and measure is the responsibility of the data provider.
- **Reduction status (model enumeration)**: A reduction level of the parameter (e.g. a parameter can be calibrated or not or it can be a computed qualifier) may be attached to the parameter.

- **Semantic**: A reference to a valid vocabulary word may be attached to each parameter. The choice of that vocabulary is totally free as long as it is published.
- **Description**: A free text description may be attached to each parameter.

**TODO:**
TBC phys.luminosity vs phot.flux

| Parameter | Original model | UCDs 1+ first word |
|---|---|---|
| `GenericMeasure` | Measure | Appropriate UCD |
| `Position` | Measure | pos.* |
| `Velocity` | Measure | phys.veloc |
| `Proper motion` | Measure | pos.pm |
| `Time` | Measure | time.epoch |
| `Polarization` | Measure | phys.polarization |
| `LonLatSkyPosition` | MANGO | pos.eq |
| `Redshift` | MANGO | src.redshift |
| `Photometry` | MANGO | phot.* |
| `HardnessRatio` | MANGO | phot.flux;arith.ratio |
| `Shape` | MANGO | phys.area |
| `Flag` | MANGO | meta.code |
| `Orbit` | MANGO | src.orbital |
| `GenericStringMeasure` | MANGO | Appropriate UCD |
| `GenericVocabMeasure` | MANGO | Appropriate UCD |

*Table 2:* UCDs space to be used for the supported measures

### 3.1.2 Measure Extension

All `Measure` classes are built upon the MCT pattern (see Fig 3).

- `Measure` instances are made with an `Error` and a `Coordinate` that contains the measure value(s).
- The `Coordinate` includes a `CoordSys` instance describing the coordinate system relevant for that measure.
- The coordinate system has two components : the space ( `CoordSpace` class) that describes the axis and the frame (`CoordFrame` class).

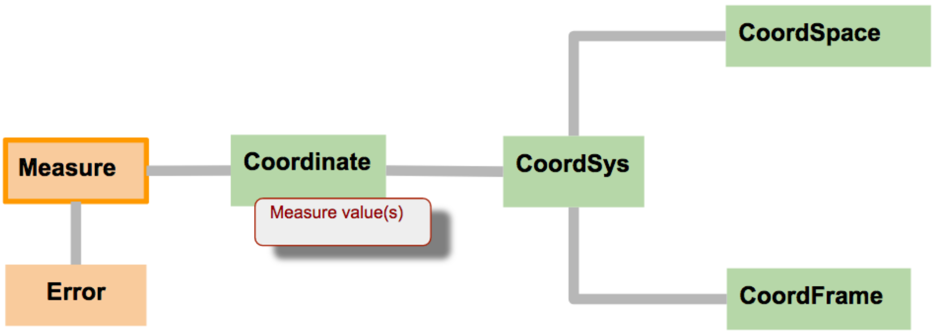


*Figure 3:* Measure/Coordinate pattern (simplified view)

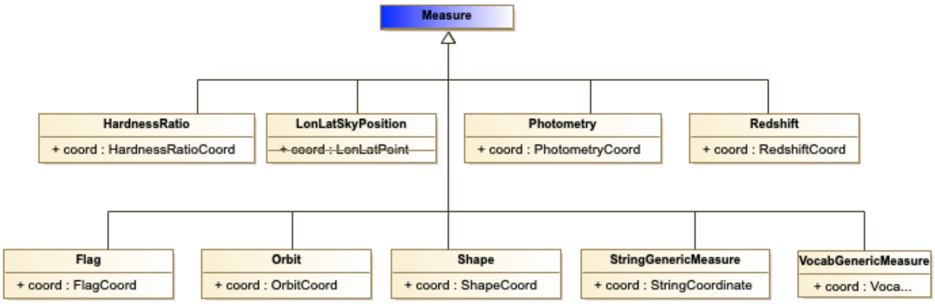


*Figure 4:* Mango extensions of Measure.

MANGO parameters are all based on this pattern. Native `Measure` classes are used whenever possible. Others Parameters are built by extending this class as shown in figure 3).

The scope of the MCT `Measure` concept has been extended to quantities that are not physical measurements such as e.g. flags, object types or orbits. This lexical deviation allows us to place all quantities describing a source in a single container since all measurements in the Mango sense are derived from the MCT `Measure` class. The same approach has been used for the coordinate systems since e.g the way to serialise a complex shape (e.g. as a MOC) is considered as a coordinate system in sense of `Coords`.

These extended classes are part of MANGO.

## 3.2 Associated Data

`AssociatedData` connectors are used to bind any sort of complex data with the source. One connector can only refer to one dataset. Associated data can be either URIs (VO services or not), other mango instances or reference

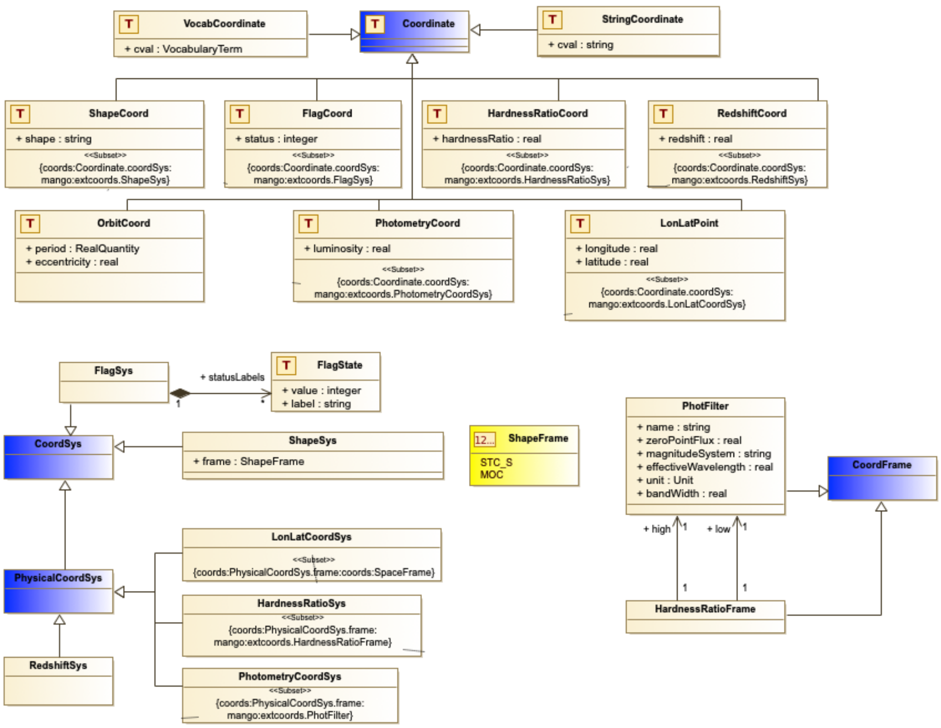


*Figure 5:* Mango extensions of Coordinates.

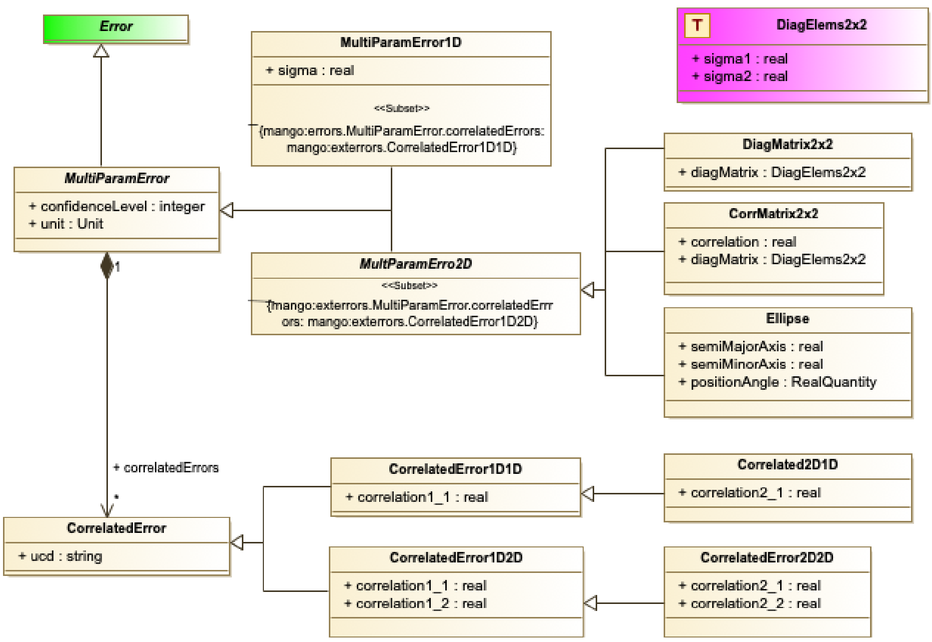


*Figure 6:* Mango correlated errors.

to instances of other VO models (e.g. Obscore, Provenance…).

### 3.2.1 AssociatedData Identification

As the associated data set attached to a particular instance is not defined by Mango, the model must provide an accurate parameter description to allow the client to figure out what it can do with each of them. Mango provides 3 description levels for each category of associated data:

- **Data class (vodml type)**: Knowing the class representing associated data tells the clients how to interpret it but does help to get its role.
- **Semantic**: A reference to a valid vocabulary word may be attached to each associated data set. The choice of that vocabulary is totally free as long as it is published.
- **Description**: A free text description may be attached to each associated data .

### 3.2.2 Associated Model Instances

The way to attach VO model instances to MANGO sources is very specific in a sense of that Mango gives the model reference by tells nothing about the way import the instance. The model of associated instances must be available as a vodml file. The `ModelInstance` is an empty class without sub classes.

# 4 Model: mango

Data model based oon components and data association for source data

## 4.1 Article

Reference article for the MANGO source

### 4.1.1 Article.editor

**vodml-id: Article.editor**
**type: ivoa:string**
**multiplicity: 1**
Article editor name

### 4.1.2 Article.article

**vodml-id: Article.article**
**type: ivoa:string**
**multiplicity: 1**
Bibcode or DOI of the reference article

## 4.2 AssociatedData (Abstract)

Abstract reference to a particular dataset associated to the Source. This class is used to specify the type of the dataset as well as its role.

### 4.2.1 AssociatedData.semantic

**vodml-id: AssociatedData.semantic**
**type: mango:VocabularyTerm**
**multiplicity: 1**
Reference to a semantic concept giving the nature of the associated data. As long as the vocabulary is not set, the possible values of this attribute are given by the LinkSemantic enumeration.

### 4.2.2 AssociatedData.description

**vodml-id: AssociatedData.description**
**type: ivoa:string**
**multiplicity: 1**
Free text description of the associated data

## 4.3 AssociatedMangoInstance

Reference to another MANGO instance that is part of the associated data.

### 4.3.1 AssociatedMangoInstance.mangoInstance

**vodml-id: AssociatedMangoInstance.mangoInstance**
**type: mango:Source**
**multiplicity: 1**
Composition link pointing on one MANGO instance associated with the source.

## 4.4 BitField

State of a property for which each value is represented by a bit so that several states can be included in the same numerical value. The values defined in the associated StatusValues must correspond to bit patterns. This constraint is not managed by the model.

## 4.5 Color

Photometric property describing a color of the MANGO source. color is not an intrinsic quantity of the MANGO source but a value relative to two filters or energy bands.

### 4.5.1 Color.colorDef

**vodml-id: Color.colorDef**
**type: mango:ColorDef**
**multiplicity: 1**
Color definition. Can be either a difference of magnitudes or an hardness ratio.

## 4.6 ColorDef

Definition of the colour to which the Color property applies. It contains the way the colour is calculated as well as the definition of the filters used.

### 4.6.1 ColorDef.definition

**vodml-id: ColorDef.definition**
**type: mango:ColorDefinition**
**multiplicity: 1**
TODO : Missing description : please, update your UML model asap.

### 4.6.2 ColorDef.high

**vodml-id: ColorDef.high**
**type: :PhotFilter**
**multiplicity: 1**
Reference to the `photdm:PhotFilter` corresponding the highest energy for that color.

### 4.6.3 ColorDef.low

**vodml-id: ColorDef.low**
**type: :PhotFilter**
**multiplicity: 1**
Reference to the `photdm:PhotFilter` corresponding the lowest energy for that color.

## 4.7 DataOrigin

Class representing the origin of the data following the DCP note (TBD)

### 4.7.1 DataOrigin.citation

**vodml-id: DataOrigin.citation**
**type: ivoa:string**
**multiplicity: 1**
Dataset idnetifier that can be used for citation (e.g. DOI)

### 4.7.2 DataOrigin.reference_url

**vodml-id: DataOrigin.reference_url**
**type: ivoa:string**
**multiplicity: 1**
Dataset landing page

### 4.7.3 DataOrigin.resource_version

**vodml-id: DataOrigin.resource_version**
**type: ivoa:string**
**multiplicity: 1**
Dataset version

#### 4.7.4 DataOrigin.creator

**vodml-id: DataOrigin.creator**
**type: ivoa:string**
**multiplicity: 1**
Person(s) mainly involved in the creation of the resource, generally the author.

#### 4.7.5 DataOrigin.cites

**vodml-id: DataOrigin.cites**
**type: ivoa:string**
**multiplicity: 1**
Identifier (IVOID, DOI or Bibcode) of a second resource using relation of type `cites` (https://www.ivoa.net/rdf/voresource/relationship_type/)

#### 4.7.6 DataOrigin.is_derived_from

**vodml-id: DataOrigin.is_derivedi_from**
**type: ivoa:string**
**multiplicity: 1**
Identifier (IVOID, DOI or Bibcode) of a second resource using relation of type `is_derived_from` (https://www.ivoa.net/rdf/voresource/relationship_type/)

#### 4.7.7 DataOrigin.original_date

**vodml-id: DataOrigin.original_date**
**type: ivoa:datetime**
**multiplicity: 1**
Date of the original resource from which the MANGO source instance is derived

#### 4.7.8 DataOrigin.query

**vodml-id: DataOrigin.query**
**type: mango:QueryOrigin**
**multiplicity: 0..1**
TODO : Missing description : please, update your UML model asap.

#### 4.7.9 DataOrigin.rights

**vodml-id: DataOrigin.rights**
**type: mango:License**

**multiplicity: 0..1**
TODO : Missing description : please, update your UML model asap.

### 4.7.10 DataOrigin.article

**vodml-id: DataOrigin.article**
**type: mango:Article**
**multiplicity: 0..1**
TODO : Missing description : please, update your UML model asap.

## 4.8 EpochPosition

Class grouping all parameters needed to define an object position at a given epoch. The space coordinate system is common to all attributes to insure the consistance between all of the instance parameters.

### 4.8.1 EpochPosition.longitude

**vodml-id: EpochPosition.longitude**
**type: ivoa:RealQuantity**
**multiplicity: 1**
The longitude of the Point, as a Quantity with angular units (see `coords:LonLatPoint.lon`.

### 4.8.2 EpochPosition.latitude

**vodml-id: EpochPosition.latitude**
**type: ivoa:RealQuantity**
**multiplicity: 1**
The latitude of the Point, as a Quantity with angular units (see `coords:LonLatPoint.lat`.

### 4.8.3 EpochPosition.parallax

**vodml-id: EpochPosition.parallax**
**type: ivoa:RealQuantity**
**multiplicity: 1**
The measured parallax value in the coordinate system of the `EpochPosition` instance.

### 4.8.4 EpochPosition.radialVelocity

**vodml-id: EpochPosition.radialVelocity**
**type: ivoa:RealQuantity**
**multiplicity: 1**
The measured Velocity value along of the radius axis (see texttmeas:Velocity.coord).

### 4.8.5 EpochPosition.pmLongitude

**vodml-id: EpochPosition.pmLongitude**
**type: ivoa:RealQuantity**
**multiplicity: 1**
Velocity along the longitude axis in angular distance per unit time (see `meas:ProperMotion.coord`). The current version of the model only allows a representation in the Polar coordinate space.

### 4.8.6 EpochPosition.pmLatitude

**vodml-id: EpochPosition.pmLatitude**
**type: ivoa:RealQuantity**
**multiplicity: 1**
Velocity along the latitude axis in angular distance per unit time (see `meas:ProperMotion.coord`). The current version of the model only allows a representation in the Polar coordinate space.

### 4.8.7 EpochPosition.epoch

**vodml-id: EpochPosition.epoch**
**type: coords:Epoch**
**multiplicity: 1**
Class grouping all parameters needed to define an object position at a given epoch. The space coordinate system is common to all attributes to insure the consistance between all of the instance parameters.

### 4.8.8 EpochPosition.pmCosDeltApplied

**vodml-id: EpochPosition.pmCosDeltApplied**
**type: ivoa:boolean**
**multiplicity: 1**
It is common, though not universal, practice to quote longitudinal proper motion pre-multiplied by cos(latitude) so that the magnitude of the quantity is not affected by its longitudinal position. We do not constrain the value to one form or the other in this model. Instead, this flag enables providers to convey whether or not the factor has been applied (see `meas:ProperMotion.cosLat_applied`)

### 4.8.9 EpochPosition.error

**vodml-id: EpochPosition.error**
**type: mango:EpochPositionError**
**multiplicity: 0..1**
Class grouping all parameters needed to define an object position at a given

epoch. The space coordinate system is common to all attributes to insure the consistance between all of the instance parameters.

#### 4.8.10 EpochPosition.coordSys

**vodml-id: EpochPosition.coordSys**
**type: coords:SpaceSys**
**multiplicity: 0..1**
Class grouping all parameters needed to define an object position at a given epoch. The space coordinate system is common to all attributes to insure the consistance between all of the instance parameters.

### 4.9 EpochPositionError

Class for the error attached to a EpochPosition. The component in this class represent the errors of individual parameters as well as errors due to the parameter correlations.

#### 4.9.1 EpochPositionError.parallaxError

**vodml-id: EpochPositionError.parallaxError**
**type: mango:ErrorTypes.PropertyError1D**
**multiplicity: 0..1**
Parallax error. This error is meant to be symetrical.

#### 4.9.2 EpochPositionError.radialVelocityError

**vodml-id: EpochPositionError.radialVelocityError**
**type: mango:ErrorTypes.PropertyError1D**
**multiplicity: 0..1**
Error in the radial velocity. This error is meant to be symetrical.

#### 4.9.3 EpochPositionError.pmPosError

**vodml-id: EpochPositionError.pmPosError**
**type: mango:ErrorTypes.Correlated22Error**
**multiplicity: 0..1**
Correlated error between the proper motion and the position. This error is represented by a 2D matrix. $error\_propermotion = f(position)$

#### 4.9.4 EpochPositionError.radialVelocityPosError

**vodml-id: EpochPositionError.radialVelocityPosError**
**type: mango:ErrorTypes.Correlated12Error**
**multiplicity: 0..1**

Correlated error between the radial velocity and the position. The radial velocity error might depend on the sky position. $error\_radialVelocity = f(position)$

### 4.9.5 EpochPositionError.parallaxPosError

**vodml-id: EpochPositionError.parallaxPosError**
**type: mango:ErrorTypes.Correlated12Error**
**multiplicity: 0..1**
Correlated error between the parallax and the position. The parallax error might depend on the sky position. $error\_parallax = f(position)$

### 4.9.6 EpochPositionError.parallaxPmError

**vodml-id: EpochPositionError.parallaxPmError**
**type: mango:ErrorTypes.Correlated12Error**
**multiplicity: 0..1**
Correlated error between the parallax and the proper motion. The parallax error might depend on the proper motion vector. $error\_parallax = f(properMotion)$

## 4.10 Label

Free text label seen as a MANGO source property.

### 4.10.1 Label.text

**vodml-id: Label.text**
**type: ivoa:string**
**multiplicity: 1**
Text of label seen as a MANGO source property.

## 4.11 License

Place holder for the license covering the MANGO instance

### 4.11.1 License.rights_uri

**vodml-id: License.rights_uri**
**type: ivoa:string**
**multiplicity: 1**
Licence URI following the regostry practice. This should come from SPDX `https://spdx.org/licenses`. Creatives Commons (`https://creativecommons.org`) are also accepted.

#### 4.11.2 License.rights

**vodml-id: License.rights**
**type: ivoa:string**
**multiplicity: 1**
License or Copyright text

### 4.12 PhotometricProperty

Observed brightness of the MANGO source. The distinction between fluxes of a magnitude is made by the unit. This property should refer to a photometric calibration as defined by the `PhotDM` model.

#### 4.12.1 PhotometricProperty.value

**vodml-id: PhotometricProperty.value**
**type: ivoa:RealQuantity**
**multiplicity: 1**
Value of the photometric property associated with a photometric calibration as defined by the `PhotDM` model.

#### 4.12.2 PhotometricProperty.photCal

**vodml-id: PhotometricProperty.photCal**
**type: :PhotCal**
**multiplicity: 1**
Photometric calibration that applies to the photometric property. It must be an instance of `photdm:PhotCal`.

### 4.13 PhysicalProperty

Place holder for any quantity that can be hold by measure as defined in the `Astronomical Measurements Model`.

#### 4.13.1 PhysicalProperty.calibrationLevel

**vodml-id: PhysicalProperty.calibrationLevel**
**type: mango:CalibrationLevel**
**multiplicity: 1**
TODO : Missing description : please, update your UML model asap.

#### 4.13.2 PhysicalProperty.measure

**vodml-id: PhysicalProperty.measure**
**type: meas:Measure**
**multiplicity: 1**
Instance of `Astronomical Measurements Model` that holds the Property value(s).

### 4.14 Property

Reference to a particular measure of the Source. This class is used to specify the type of the measure as well as its role.
**constraint**
**detail: Property.One association at the time**

#### 4.14.1 Property.semantic

**vodml-id: Property.semantic**
**type: mango:VocabularyTerm**
**multiplicity: 1**
Reference to a semantic concept giving the nature of the parameter As long as the vocabulary is not set, the possible values of this attribute are given by the ParamSemantic enumeration.

#### 4.14.2 Property.description

**vodml-id: Property.description**
**type: ivoa:string**
**multiplicity: 1**
Free text description of the measure

#### 4.14.3 Property.measure

**vodml-id: Property.measure**
**type: meas:Measure**
**multiplicity: 1**
Composition link pointing to the meas:Measure instance

#### 4.14.4 Property.associatedProperties

**vodml-id: Property.associatedProperties**
**type: mango:Property**
**multiplicity: 1..***
<Enter note text here>

## 4.15 QueryOrigin

TODO : Missing description : please, update your UML model asap.

### 4.15.1 QueryOrigin.ivoid

**vodml-id: QueryOrigin.ivoid**
**type: ivoa:string**
**multiplicity: 1**
IVOID of the underlying data collection

### 4.15.2 QueryOrigin.publisher

**vodml-id: QueryOrigin.publisher**
**type: ivoa:string**
**multiplicity: 1**
Data center that produced the MANGOI source instance

### 4.15.3 QueryOrigin.server_software

**vodml-id: QueryOrigin.server_software**
**type: ivoa:string**
**multiplicity: 1**
Version of the software the produced the MANO source instance. It is encouraged to follow `https://ivoa.net/documents/Notes/softid/index.html`.

### 4.15.4 QueryOrigin.service_protocol

**vodml-id: QueryOrigin.service_protocol**
**type: ivoa:string**
**multiplicity: 1**
IVOID of the protocol through which the data was retrieved

### 4.15.5 QueryOrigin.request

**vodml-id: QueryOrigin.request**
**type: ivoa:string**
**multiplicity: 1**
Full request URL including a query string. For the simple protocols,put the url-encoded form of the query parameters. For TAP queries, use the /sync UWS URL. The format is free for others request types.

#### 4.15.6 QueryOrigin.query

**vodml-id: QueryOrigin.query**
**type: ivoa:string**
**multiplicity: 1**
Input query in a formal langage such as ADQL.

#### 4.15.7 QueryOrigin.request_date

**vodml-id: QueryOrigin.request_date**
**type: ivoa:datetime**
**multiplicity: 1**
Query execution date

#### 4.15.8 QueryOrigin.contact

**vodml-id: QueryOrigin.contact**
**type: ivoa:string**
**multiplicity: 1**
Email or URL to contact the publisher

### 4.16 SatusValues

TODO : Missing description : please, update your UML model asap.

#### 4.16.1 SatusValues.values

**vodml-id: SatusValues.values**
**type: mango:StatusValue**
**multiplicity: 1..***
TODO : Missing description : please, update your UML model asap.

### 4.17 Shape

Description of the spatial extension of the MANGO source (for e.g. dust clouds).

#### 4.17.1 Shape.shape

**vodml-id: Shape.shape**
**type: ivoa:string**
**multiplicity: 1**
String serialization of the spatial extension of the MANGO source.

### 4.17.2 Shape.serialization

**vodml-id: Shape.serialization**
**type: mango:ShapeSerialization**
**multiplicity: 1**
Serialization mode of the spatial extension of the MANGO source.

### 4.17.3 Shape.coordSys

**vodml-id: Shape.coordSys**
**type: coords:SpaceSys**
**multiplicity: 0..1**
Coordinate system the applies to the spatial extension of the MANGO source. This component is imported from the `coords` model (see `coords:SpaceSys`).

## 4.18 Source

Root class of the model. MANGO instances are meant of be Source instances. A source is something with an identifier and two docks: one for the parameters and one for the associated data.

### 4.18.1 Source.identifier

**vodml-id: Source.identifier**
**type: ivoa:string**
**multiplicity: 1**
Unique identifier for a Source. The uniqueness of that identifier is not managed by the model. The format is free.

### 4.18.2 Source.associatedDataDock

**vodml-id: Source.associatedDataDock**
**type: mango:AssociatedData**
**multiplicity: 0..***
Composition link pointing on all data associated with the source.

### 4.18.3 Source.propertyDock

**vodml-id: Source.propertyDock**
**type: mango:Property**
**multiplicity: 0..***
Composition link pointing on all parameters attached to the source.

### 4.18.4 Source.dataOrigin

**vodml-id: Source.dataOrigin**
**type: mango:DataOrigin**
**multiplicity: 0..1**
Description of the origin of the MANGO object. The parameters contained in this component come from the work of the DCP-IG.

## 4.19 Status

Property representing a status defined by a integer number that can only take on a defined number of values, each with its own description. Boolean status can be represented by `StatusValues` with 2 values e.g. 0 for False and 1 for True.

### 4.19.1 Status.status

**vodml-id: Status.status**
**type: ivoa:integer**
**multiplicity: 1**
Actual value of the status.

### 4.19.2 Status.allowedValues

**vodml-id: Status.allowedValues**
**type: mango:SatusValues**
**multiplicity: 0..1**
List of the allowed values for the status. Each value has its own textual description.

## 4.20 StatusValue

List of the allowed values for the status. Each value has its own textual description.

### 4.20.1 StatusValue.value

**vodml-id: StatusValue.value**
**type: ivoa:integer**
**multiplicity: 1**
TODO : Missing description : please, update your UML model asap.

#### 4.20.2 StatusValue.description

**vodml-id: StatusValue.description**
**type: ivoa:string**
**multiplicity: 1**
TODO : Missing description : please, update your UML model asap.

### 4.21 VocabularyTerm

Term of a standardised vocabulary that applies to property.

#### 4.21.1 VocabularyTerm.uri

**vodml-id: VocabularyTerm.uri**
**type: ivoa:string**
**multiplicity: 1**
URI the vocabulary term.

#### 4.21.2 VocabularyTerm.label

**vodml-id: VocabularyTerm.label**
**type: ivoa:string**
**multiplicity: 1**
Label attached to the vocabulary term. This is necessary because the URI may not contain an explicit label. This was the case for the IUA vocabulary until the Registry WG introduced rewriting rules that fix the issue.

### 4.22 WebEndpoint

Class for associated data referenced by an URL

#### 4.22.1 WebEndpoint.ContentType

**vodml-id: WebEndpoint.ContentType**
**type: ivoa:string**
**multiplicity: 1**
Mime type of the URL

#### 4.22.2 WebEndpoint.url

**vodml-id: WebEndpoint.url**
**type: ivoa:anyURI**
**multiplicity: 1**
Web endpoint

## 4.23 ShapeFrame

Possible options to encode a shape in a string.
Enumeration Literals

**STC_S** : **vodml-id:** ShapeFrame.STC_S
**description:** STCs serialisation

**MOC** : **vodml-id:** ShapeFrame.MOC
**description:** MOC serialisation

## 4.24 ShapeSerialization

Enumeration of the supported serialization modes for the shapes
Enumeration Literals

**MOC** : **vodml-id:** ShapeSerialization.MOC
**description:** Label indicating that the shape has been serialized as a S-MOC

**STCS** : **vodml-id:** ShapeSerialization.STCS
**description:** Label indicating that the shape has been serialized as a STCS string

**POLYGON** : **vodml-id:** ShapeSerialization.POLYGON
**description:** Label indicating that the shape has been serialized as a polygon (cf xtypes)

## 4.25 CalibrationLevel

Science ready data with the instrument signature removed (ObsCore).
Enumeration Literals

**Raw** : **vodml-id:** CalibrationLevel.Raw
**description:** Raw instrumental data, in a proprietary or internal data provider defined format, that needs instrument specific tools to be handled (ObsCore).

**Instrumental** : **vodml-id:** CalibrationLevel.Instrumental
**description:** Instrumental data in a standard format which could be manipulated with standard astronomical packages (ObsCore).

**Calibrated** : **vodml-id:** CalibrationLevel.Calibrated
**description:** Science ready data with the instrument signature removed (ObsCore).

## 4.26 ColorDefinition

Enumeration of the different color types supported by the model.
Enumeration Literals

**ColorIndex** : **vodml-id:** ColorDefinition.ColorIndex
**description:** Difference of mangnitudes: typically $M\_B - M\_v$

**HardnessRatio** : **vodml-id:** ColorDefinition.HardnessRatio
**description:** Normalized ratio of fluxes: $(F\_EB2 - F\_EB1)/(F\_EB2 + F\_EB1)$

# 5 Package: ErrorTypes

TODO : Missing description : please, update your UML model asap.

## 5.1 Correlated11Error

Error of a 1D property (A) correlated with a 1D parameter (B): $error$
$_{A} = correlation$
$_{A}1B1 * value$
$_{B}$

### 5.1.1 Correlated11Error.a1b1

**vodml-id: ErrorTypes.Correlated11Error.a1b1**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the first axis of `B` to the error on the first axis of `A`.

## 5.2 Correlated12Error

Error of a 1D property (A) correlated with a 2D parameter (B): $error$
$_{A} = correlation$
$_{A}1B1 * value$
$_{B}\_1 + correlation$
$_{A}1B2 * value$
$_{B}\_2$

### 5.2.1 Correlated12Error.a1b1

**vodml-id: ErrorTypes.Correlated12Error.a1b1**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the first axis of `B` to the error on the first axis of `A`.

### 5.2.2 Correlated12Error.a1b2

**vodml-id: ErrorTypes.Correlated12Error.a1b2**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the second axis of `B` to the error on the first axis of `A`.

## 5.3 Correlated21Error

Error of a 2D property (A) correlated with a 1D parameter (B): $error$
$_{A}\_1 = correlation$
$_{A}1B1 * value$
$_{B}\ error$
$_{A}\_2 = correlation$
$_{A}12B1 * value$
$_{B}$

### 5.3.1 Correlated21Error.a2b1

**vodml-id: ErrorTypes.Correlated21Error.a2b1**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the second axis of `B` to the error on the first axis of `A`.

### 5.3.2 Correlated21Error.a1b1

**vodml-id: ErrorTypes.Correlated21Error.a1b1**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the first axis of `B` to the error on the first axis of `A`.

## 5.4 Correlated22Error

Error of a 2D property (A) correlated with a 2D parameter (B): $error$
$_{A}\_1 = correlation$
$_{A}1B1 * value$
$_{B}\_1 + correlation$
$_{A}1B2 * value$
$_{B}\_2\ error$
$_{A}\_2 = correlation$
$_{A}2B1 * value$
$_{B}\_1 + correlation$
$_{A}2B2 * value$
$_{B}\_2$

### 5.4.1 Correlated22Error.a2b1

**vodml-id: ErrorTypes.Correlated22Error.a2b1**
**type: ivoa:real**
**multiplicity: 1**

Correlation coefficient giving the contribution of the first axis of `B` to the error on the second axis of `A`.

### 5.4.2 Correlated22Error.a2b2

**vodml-id: ErrorTypes.Correlated22Error.a2b2**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the second axis of `B` to the error on the second axis of `A`.

### 5.4.3 Correlated22Error.a1b1

**vodml-id: ErrorTypes.Correlated22Error.a1b1**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the first axis of `B` to the error on the first axis of `A`.

### 5.4.4 Correlated22Error.a1b2

**vodml-id: ErrorTypes.Correlated22Error.a1b2**
**type: ivoa:real**
**multiplicity: 1**
Correlation coefficient giving the contribution of the second axis of `B` to the error on the first axis of `A`.

## 5.5 CorrelatedError (Abstract)

Abstract root class for the correlated errors. In the subclasses nomenclature the error relates to a property named `A` in correlation with a property named B).

### 5.5.1 CorrelatedError.ucd_B

**vodml-id: ErrorTypes.CorrelatedError.ucd_B**
**type: ivoa:string**
**multiplicity: 0..1**
Optional parameter giving the UCD of the correlated property. This attribute may facilitate the interpretation of complex error patterns.

## 5.6 DiagElems2x2

Datatype containing the 2 diagonal elements of a 2x2 matrix. Attributes are named $\sigma$ because this datatype is mostly used in the context of complex errors.

### 5.6.1 DiagElems2x2.sigma1

**vodml-id: ErrorTypes.DiagElems2x2.sigma1**
**type: ivoa:real**
**multiplicity: 1**
Variance on the first axis.

### 5.6.2 DiagElems2x2.sigma2

**vodml-id: ErrorTypes.DiagElems2x2.sigma2**
**type: ivoa:real**
**multiplicity: 1**
Variance on the second axis

## 5.7 Ellipse

Elliptical errror for 2D parameters such as the sky positions.

### 5.7.1 Ellipse.semiMajorAxis

**vodml-id: ErrorTypes.Ellipse.semiMajorAxis**
**type: ivoa:real**
**multiplicity: 1**
Half of the ellipse major axis.

### 5.7.2 Ellipse.semiMinorAxis

**vodml-id: ErrorTypes.Ellipse.semiMinorAxis**
**type: ivoa:real**
**multiplicity: 1**
Half of the ellipse minor axis.

### 5.7.3 Ellipse.angle

**vodml-id: ErrorTypes.Ellipse.angle**
**type: ivoa:RealQuantity**
**multiplicity: 1**
Angle between the North Polar Cape (NCP) and the major axis. This angle must be positive taking into account that angles are positive from North to the East. The angle has its own unit.

## 5.8 ErrorMatrix

Diagonal 2D matrix. Non diagonal elements are null.

### 5.8.1 ErrorMatrix.sigma1

**vodml-id: ErrorTypes.ErrorMatrix.sigma1**
**type: ivoa:real**
**multiplicity: 1**
First diagonal element ($(x\_11)$)

### 5.8.2 ErrorMatrix.sigma2

**vodml-id: ErrorTypes.ErrorMatrix.sigma2**
**type: ivoa:real**
**multiplicity: 1**
Second diagonal element ($(x\_22)$)

### 5.8.3 ErrorMatrix.covariance

**vodml-id: ErrorTypes.ErrorMatrix.covariance**
**type: ivoa:real**
**multiplicity: 1**
TODO : Missing description : please, update your UML model asap.

## 5.9 PropertyError (Abstract)

Root (abstract) class of the errors that can be attached to a MA?GO property. The class inherit from `meas:uncertainty` in order to be usable in the context of properties based on `Measures` classes.

### 5.9.1 PropertyError.confidenceLevel

**vodml-id: ErrorTypes.PropertyError.confidenceLevel**
**type: ivoa:integer**
**multiplicity: 1**
Confidence level of the error, expressed in $\sigma$.

### 5.9.2 PropertyError.unit

**vodml-id: ErrorTypes.PropertyError.unit**
**type: ivoa:Unit**
**multiplicity: 1**
Error unit. It must be compliant with the `VOUnit` standard. The error unit must be consistant with the unit of the property the error is attached with. This is not checked at the model level.

### 5.10 PropertyError1D

Symetrical error for 1D parameter.

#### 5.10.1 PropertyError1D.sigma

**vodml-id: ErrorTypes.PropertyError1D.sigma**
**type: ivoa:real**
**multiplicity: 1**
Error amplitude.

## 6 TAP and MANGO

This not normative section gives possible tips to save and discover MANGO instances in TAP services. We suppose that the TAP service hosts catalogs which sources are MANGO instances. These catalogs are named *MANGO Catalogs*.

### 6.1 Storing MANGO Catalogs in TAP

For now this section only concerns the parameter. The associated data will be taken into account later.

- One master table for the catalogs with various meta-data out of the MANGO scope plus a unique identifier (primary key)
- One master sources table for the source instances with the catalog identifier and a primary key safer than the MANGO identifier.
- One table for each supported parameter with a foreign key for the join with the master source table

Although the model of the measures is hierarchical, it should be possible to flatten them in one single table considering that the model structure can be retrieved with the TAP_SCHEMA annotations (TBC)

This schema requires the server to explore all the parameter tables to retrieve whole MANGO instances. This process can be speed up by using the *MANGOCore* table.

### 6.2 *MANGOCore* Table

The discovery of *MANGO Catalogs* can be helped by a *MANGOCore* table located in the *schema* schema. As MANGO is not dedicated to any specific domain, we cannot define a set of core parameters, but parameters can be flagged as *Core Parameter*. This selection is left at the discretion of the curator. The *MANGOCore* table has set of columns per parameter class plus one for the catalog ID. It has one row per stored catalog. Each parameter has at least 2 columns: one with the UCD and one with the *Core* flag. TBC

# A Gaia

Gaia mission is producing the largest and more precise 3D map of our galaxy.

Gaia core solution is able to solve the astrometric solution of more than 1 billion sources by complex models and algorithms (Lindegren and Lammers et al., 2012). Using a minimisation problem approach, different detections identified on different scans can be associated to the same astronomical source. Some of the properties would be direct measurements on single scans (e.g. positions or magnitudes). Also other properties like radial velocity (measured in redshift units) are also obtained at integration time of the scans.

Once detections on different scans are associated to a single astronomical source, these direct measurements can be combined to generate derived measurements. Trajectories on the sky for galactic objects are seen as spirals including the combination of the proper motion of the object (derived by the main vector of movement of the source) and the parallax (derived by the radius of the apparent spiral produced by the different angles of the Gaia observations at different periods).

From these properties, others could be also derived like, e.g. the distance, although the exact transformation from parallax to distance requires the use of accurate calculations and calibrations so, in general, only direct astrometric properties, like parallax, are usually provided into the catalogues.

Finally, other properties will be also obtained by the cross-identification of detections as a single astronomical source. For example, time series could be combined as the result of measurements of magnitudes from different scans detections.

Although it is not its main scientific target and apart from stellar objects, other extra-galactic sources could be also studied with Gaia. For example, QSOs are observed by Gaia as point-like sources with zero proper motion. For these kind of sources, typical extra-galactic properties like, redshift, could be also provided.

A non-exhaustive list of properties required for Gaia use cases would be composed of:

- identifier
- sky reference position
- proper motion

- parallax and distance
- source extension
- radial velocity
- redshift
- photometry
- date of detection
- correlation
- multiple detection

## B Euclid

Euclid telescope has been designed to unveil some of the questions about the dark Universe, including dark matter and dark energy, what would include, e.g. quite accurate measurements of the expansion of the Universe.

Euclid will mainly observe extragalactic objects providing, e.g. information of the shapes of galaxies, gravitational lensing, baryon acoustic oscillations and distances to galaxies using spectroscopic data.

For this mission, and apart from the common metadata provided for extra galactic sources into astronomical catalogues, a good support for object taxonomy and shapes of objects will be required. As known due to general relativity effects, shapes far galaxies could be deformed due to gravitational lensing effects, producing convergence (visual displacements on the position) and rear (deformation of the shape) effects. All these metadata should be ready for annotations and, also, correlated to theoretical or real metadata in other datasets.

Finally, crossmatch information with other catalogues will be of crucial interest as data from other satellites and, more importantly, from ground based observatories will be combined with Euclid data to produce consistent scientific datasets.

A non-exhaustive list of properties required for Euclid use cases would be composed of:

- identifier
- sky position
- correlation with other catalogues
- photometry (ground + satellite )

- morphology class
- redshift
- photometric redshift

## C Chandra Archive

The Chandra Source Catalog(CSC) is the definitive catalog of serendipitous X-ray sources identified in publicly released imaging observations obtained by NASA's ChandraX-ray Observatory (CXO). The CSC is developed and published by the ChandraX-ray Center (CXC) and is supported by NASA contract NAS 8-3060 to the Smithsonian Astrophysical Observatory for operation of the CXC. CSC Release 2.0 (Oct. 2019) includes properties for approximately 316,000 X-ray sources on the sky extracted from about 375,000 detections.

The catalog itself consists of approximately 1,700 columns covering properites at the individual observation and stacked analysis levels. Table 3 summarizes some of the basic catalog properties derived from standard CSCView queries.

The following are some example usage threads which can be facilitated by this model. We provide a high-level summary here described in generic terms. For each case, we are generating detailed threads describing how each would be implemented using the Chandra Source Catalog.

Note: Most of these cases have been simplified to a level appropriate for this version of the model. More complete and accurate science threads could be defined using more complex types and associations which are out of the the current scope.

1. Searching for spectrally variable or flaring point sources:
    - Identifying spectrally variable point sources
        (a) From a catalog, identify a set of point sources with multiple spectrally-resolved observations
            - Sources are classified as point sources or have a source extent that is consistent with being a point source
            - Observations include detection properties in at least two user-specified wavebands
            - There are multiple observations associated with the source
            - If the catalog differentiates between detections and sources, each included detection must be uniquely associated with the source

| |
|---|
| **Per Source:** |
| `Source name` |
| `Source position and position errors` |
| `Significance of the source (signal to noise)` |
| `Likelihood of the source (True, False, or Marginal detection)` |
| `Source extent flag` |
| `Variability flag` |
| `Spectral variability flag` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |
| `Long term (inter-obs) variability probability for each band` |
| `Spectral (hardness ratios) variability for each color` |
| **Per Detection (at the stack level):** |
| `Detection ID` |
| `Detection position and position errors` |
| `Flux significance of the detection (S/N)` |
| `Detection likelihood (True, False, or marginal detection?)` |
| `Source extent code (codification of source extent in different bands)` |
| `Variability flag` |
| `Spectral variability flag` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |
| `Long term (inter-obs) variability probability for each band` |
| `Spectral (hardness ratios) variability probability for each color` |
| **Per Detection (at the observation level):** |
| Note that source detection is done at the stack level, but properties are estimated for the detections at each observation using the detection region from the stack level. |
| `Detection ID` |
| `Detection position and position errors` |
| `Flux significance of the detection (S/N)` |
| `Detection likelihood` |
| `Source extent code (codification of source extent in different bands)` |
| `Variability code (applies to intra-obs only)` |
| `Fluxes and flux errors in ACIS bands b, h, m, s, u` |
| `Flux and flux error in HRC band w` |
| `Hardness ratios and errors for hm, hs, ms colors` |
| `Short term (intra-obs) variability probability for each band` |

*Table 3:* Example Chandra Source Catalog Properties

(b) From the set identified in (a), identify the subset of sources with significant spectral variability between observations
   - If the catalog includes hardness ratios in the user-specified wavebands, then identify sources with hardness ratios that vary by more than a user-specified confidence limit between observations
   - If the catalog does not include hardness ratios in the user-specified wavebands, then extract individual observation fluxes and associated confidence limits from the catalog and construct hardness ratios and associated confidence limits for each of the observations and then compare with the user-specified confidence limit
   - If hardness ratios and/or fluxes are computed using mul-

tiple methods, allow the user to specify which properties to use
    - If measurement probability density functions are provided rather than confidence intervals, allow those to be used instead

- Identifying flaring point sources
  (a) From a catalog, identify a set of point sources
    - Sources are classified as point sources or have a source extent that is consistent with being a point source
    - If the catalog differentiates between detections and sources, each included detection must be uniquely associated with the source

  (b) From the set identified in (a), identify the subset of sources with intra-observation variability greater than a user-specified amount in any waveband
  (c) From the set identified in (b), extract the intra-observation lightcurves and associated confidence intervals
    - Apply a user-specified matched filter to the lightcurve and confidence information to search for flares

2. Find sources with changing properties:
   Look for sources with changes of spectral slope and column density between observations so as function of time; this can easily be done across X-ray catalogs provided that the same spectral model (absorbed power-law) is used in the different catalogs. The changes in spectral slope and column density are measured in sigma using the errors as well on each quantity to evaluate the statistical significance of the changes.

   Properties needed:

   - time of observation
   - spectral parameters and their errors
   - measure of the quality of the spectral fit

   Procedure:
   For each source..

   (a) retrieve all available spectral properties as function of time
   (b) compare the properties
   (c) select sources with extreme changes (3 sigma difference with respect to the average)

3. Building a lgN-lgS:
   Source number counts are really important for comparison with population synthesis models. Moreover, number counts tell us the total resolved fraction of the CXB for comparison between different catalogs.
   Properties needed:

   - Fluxes computed in a homegeneous way by all the different catalogs and reported in the same band.
     (e.g. flux in the same 90% PSF fraction for each telescope and computed with the same spectral model)

   Procedure:
   For each source..

   (a) retrieve the fluxes and compute the lgN-lgS given the catalog sensitivity
   (b) compare between telescopes, only possible if fluxes are computed with the same assumptions for sources in different X-ray telescopes.

4. Finding Tidal Disruption Events in the CSC:
   Tidal Disruption Events (TDEs) happen when a star falls into the tidal radius of a super-massive black hole and gets disrupted by the ensuing forces exerted. Non-thermal X-ray emission is produced from a relativistic jet associated with the accretion, but thermal X-ray emission is also generated in the inner part of the accretion disk formed from the stellar debris. The transition between non-thermal and thermal emission results in an outburst that is spectrally soft during the luminosity peak, and stays soft in timescales of several years as their X-ray luminosity decreases. In order to spot these objects, we need to look at how the hardness ratios change over time.
   Properties needed:

   - A measurement of the time-dependent fluxes of the source in different bands (for example hard, medium and soft energy bands), for all epochs in which the source is observed. Source model should either provide access to band-specific fluxes and flux errors for each epoch the source was observed, or to measurements of the hardness ratios (colors) for the different epochs, with associated errors.
   - If only the per-epoch flux measurements are available, the hardness ratio variability can be estimated probabilistically as a likelihood ratio test between the null hypothesis of constant hardness ratios vs variable hardness ratios. In this case, the hardness ratio probability density function (or confidence interval) should first

be estimated from the confidence intervals of the fluxes, and then the confidence intervals derived for the hardness ratios are confronted with two hypotheses: one in which all hardness ratios measured for a source are consistent with a single true value, and one in which the true value is allowed to vary.

Procedure:
The following procedure assumes that the X-ray catalog measures fluxes and hardness ratios in at least two bands, a hard band, and a soft band. With a broad band flux we refer to the band that encompasses the majority of valid photon detections, i.e., those falling in an energy window with good nominal quantum efficiencies and effective areas. For Chandra, this is within 0.5 and 7.0 keV.

(a) From an X-ray catalog, identify all extra-galactic ($|b|>5$) sources that have measured long-term spectral variability, i.e., sources should be detected in more than one observation, and in these observations the hardness ratios between at least two bands should be different at a significant level. In CSC2 for example, this means var_inter_hard_flag should be set. Alternatively, if no variability flag is available, the spectral variability probability should be larger than 0.5.
  - Sources need to be detected at a significance larger than 3 (S/N>3) in at least 2 observations
  - Sources need to be compact, either slightly extended or point-like. Extremely extended sources are those where the extent of the emission is more than about 20% of the PSF size should be excluded.
  - Sources should have $|b| > 5$ arcseconds (extragalactic)

(b) TDEs have a transient nature with an initial increase in luminosity of between 1 and 3 orders of magnitude, and a slower decrease in luminosity that can last several years. Therefore, for this search, sources that have a flux variability probability in the broad band larger that 50%, and that have inter-observation flux differences of at least a factor of 3 (this is a conservative limit) should be selected.

(c) Ideally, observations should cover several years. If this is not the case, it suffices to determine if the source has had an increase in luminosity and is currently in a luminous, soft state. These sources can be identified as sources that are soft in at least one observation (e.g for Chandra, o.hard_hs $< 0.5$), and that either were not detected in previous existing observations, or that were previously detected with a harder state and a broad band flux that is at least 3 times dimmer than in the soft state.

(d) For the selected sources, take the set of the per-observation b-band fluxes and hardness ratios (for all available hardness ratios). Generate plots that allow to visualize simultaneously the time evolution of flux and hardness ratio. As opposed to Active Galactic Nuclei, for which the hardness ratio hardens as they become less luminous, for TDEs, the emission tends to remain soft.

(e) Identify TDEs and separate them from AGN flares. TDEs will likely appear as transient soft sources that remain soft regardless of epoch and regardless of luminosity. They are also typically softer than AGN flares. AGN flares will show an anti-correlation between the broad band flux and the hard-to-soft hardness ratio. As the flare becomes dimmer, it also becomes harder. In order to discriminate between TDEs and AGN flares, do the following: 1. Select the observations at peak luminosity and after the peak. 2. Compute the standard deviation of the hardness ratios for the observations 3. Calculate the average of the individual hardness ratio uncertainties 4. If the standard deviation is less than the average, flag the sources as a TDE candidate. This condition guarantees that the variation of the hardness ratios over time is slow.

(f) As long as the latter condition is true for at least one of the hard-to-soft hardness ratios, flag the source as a TDE candidate. Otherwise, flag it as a possible AGN flare.

(g) If the catalog does not contain a measure of hardness ratio variability probability, but it contains either a probability density function for the hardness ratio values, or their confidence intervals, then a measure of hardness ratio variability can be obtained from a likelihood analysis. One can assume that the hardness ratio is variable if the likelihood of all individual measurements being drawn from a single true flux (assuming Gaussian errors) is smaller that the likelihood of each individual measurement being produced by a different true flux.

5. Quick, rough identification of AGN, galaxies, and stars:
Properties needed:

- CSC columns required: At stack level: flux_<model>_aper_i
- XMM columns required: At stack level: EP_FLUX

Procedure:
Applicable to all CSC2/4XMM sources with optical counter parts.

(a) With CSC2/4XMM and other optical catalogs (e.g., SDSS, Legacy, PS),

   (b) Define energy bands, X-ray emission model for X-ray (and optical)
   (c) Cross-check CSC2 and 4XMM data in different energy bands/-models
   (d) If necessary, convert 4XMM (or CSC2) to the common energy-band/model
   (e) Cross-check different optical catalogs in different energy bands
   (f) Calc fx/fo for individual CSC2/4XMM sources with optical counter parts.
   (g) Also provide the range of pre-defined fx/fo for known samples.

6. Lx: useful for many follow-up research
   Properties needed:

   - CSC columns required: At stack level: flux_<model>_aper_i
   - From optical catalogs: spec-z or photo-z

   Procedure:
   Applicable to all CSC2/4XMM sources with known spec-z or photo-z

   (a) From the cross-matches, identify spec-z or photo-z
   (b) Cross-check photo-z for those with both
   (c) If possible, preset models which is most appropriate for a source type (e.g., given by fx/fo)
   (d) Calc Lx in multiple energy bands/models
   (e) If necessary, convert 4XMM (or CSC2) to the common energy band/model

7. Lx: Spectral decomposition of X-ray sources
   Analog to bulge-disk decomposition by fitting the optical radial profile.

   Usage:

   - separation XRB and hot gas emissions from the entire galaxy
   - separation AGN and hot gas emissions from the central region

   Procedure:
   Applicable to X-ray sources identified as (A) galaxy and (B) AGN by cross-matches or by fx/fo

   (a) Mapping CSC2/4XMM spectral info (e.g., hardness ratio) into a combination of softer hot gas (~1 keV APEC) and harder point source (7 keV BREM or photon index 1.7 power-law)

8. Using CSC 2.0 data to create Color-Color-Intensity plots(CCI) to try and identify the nature of extragalactic XRBs.

   **TODO:**
   Details TBD

# D Vizier

VizieR provides science ready catalogues coming from space agencies or articles and covering number of different science cases. Published data encompass a very large set of measures (position, photometry, redshift, source type, etc.) depending on their origin. They can result from observations, simulations, models or catalog compilations. Individual Vizier tables can contain data all related to one source (e.g. time series of positions or magnitudes) or to a set of sources (one row per source) or a mix of both. Data sets are ingested in Vizier on author request. Before to be put online they are processed and documented by documentalists so as to ensure a certain level of interoperability. This work relies on the analyse of both data content and scientific paper.

- Missing meta-data, e.g. space frames or filters, are added when available in the paper.
- Columns are renamed following the Vizier nomenclature in order make them compliant with the DBMS and to facilitate the grouping of all values related to one particular quantities (e.g. quality flag for a radial velocity)
- UCDs are checked or set for all columns.
- README files are generated. A README is a text file with a specific layout making it machine readable.
- Some values, not part of the original data but assigned by the CDS, are added to the tables (e.g. identifier, ICRS positions)
- Ancillary data pointing on associated data (e.g. spectra) or on linked services (e.g. visualisation), can also be added to enrich the table content.

All Vizier meta data are gathered in a specific resource in a way to facilitate the localisation of data of interest. The main specificity of the data is their heterogeneity

- Huge variety of data provenance and processing
- Meta data heterogeneity
- Table content heterogeneity
- Huge variety of possible measures
- Huge variety of patterns of measure groups

- Use of different coordinate systems
- Large variety of associated data

The Mango model must be able to provide a standard representation of most of the metadata contained in Vizier query responses, whether native or computed by the CDS, simple quantities or associated complex data. Mango is not meant to replace the current management of the meta-data, it is a way to make those meta-data understandable for a wide panel of VO-compliant clients.

Vizier gathers and delivers a curated version of published catalogs from various missions and experiments. It also distributes results of scientific papers, based on the computation , comparison and classification of sources extracted from archived data after science analysis. Vizier handles a very large set of measures in position, photometry, redshift, source type, etc. It adds value to it by recomputing additional quantities in various reference frames or equivalent spectral bands, units conversions , etc . It binds the resulting object description to other data sets representing the object, or its counterparts, or neighbourhood on sky (image), its spectral behaviour (spectrum, spectral energy distribution) or evolution through time (light curve, radial velocity curve, time series, etc.). Currently the binding and structure of the quantities is done by column grouping.

- pre-existing data
- grouping columns
- lots of available metadata
- column name formatting
- one column different frames

# E Morphologically Complex Structures

The ViaLactea Knowledge Base (VLKB, see Molinaro and Butora et al. (2016)) is a set of data resources and services built up to study the star formation regions and processes in the Milky Way. Besides 2-D images and 3-D radial velocity cubes, the VLKB exposes a bunch of source catalogues. These catalogues can be categorised as:

- compact sources catalogues: where each source can be simply described by position and (usually) ellipsoidal extension plus some photometric flux(es);

- a band merged catalogue: where each source record combines the single band source of the previous type into a multi-center/multi-positional record;

- diffuse objects of two types, bubbles and filaments, where the positional description cannot be simplified by a point in the sky and some geometric error/shape value.

The first type of catalogue presents only a small issue, related to the ellipsoidal description of the extension of the single records/sources. For these catalogues the ellipsoidal (or circular) counterpart of the position is not to be considered as a positional uncertainty, but actually some level of confidence on the actual extension of the source at the observed frequency band, thus, besides a position on the celestial sphere, a minor axis, a major axis and a positional angle are needed to describe the source.

The band merged catalogue requires an extra step, that is the aggregagtion of multiple single band sources (potentially degenerate depending on the observational waveband) to be able to provide a sort of SED of the aggregated source.

A different type of challenge is presented by filaments and bubbles where (not approximating bubbles to simple circles) the source morphology requires or a complex geometrical shape, like single or multiple polygons, or a description based on tessellation (where the order of the tessels should fit the resolution/uncertainty of the source borders). Moreover, while bubbles can be fully described by a centroid and a single complex (multi-)polygon or set of tessels, filaments are themselves complex objects such that their description includes so-called branches, spines and nodes (areas, broken-lines, points in spherical geometry).

Thus, a model that supports description of such catalogues will need a way to describe sources with:

- non-point-like positions;

- extended complex area, possibly as multiple detached areas;

- aggregation of sub-parts (that can be heterogeneous).

The VLKB tried[2] to describe, both by (custom) contour polygons and MOCs, the bubbles and filaments having the main use case of cross-matching or distance measuring diffuse objects w.r.t. compact sources. However, work progressend slowly and anyway the proper model annotations were never investigated properly.

[2] if you want to have a look at the plain content, there's a TAP service at http://vlkb.dev.neanias.eu:8080/vlkb/tap (http://saada.unistra.fr/taphandle?url=http%3A//vlkb.dev.neanias.eu%3A8080/vlkb/tap/) that can be accessed, even if not complete with descriptions nor registered

MANGO could be useful to serve nicely annotated and usable records for sources as the ones described above.

## F Exoplanets

Annotation of (exo-)planetary records in catalogues requires some specific metadata or model.

This was initially made clear when trying to allow discoverability of exoplanets time series (Molinaro and Benatti et al. (2020)) where (exercising the ObsCore model with specific use cases) information like stellar host characteristics together with stellar system details turned out to be required for a proper description of the catalogue records.

The use cases identified, besides some refinement needed in the ObsCore time axis annotation, the following metadata:

- the degree of confidence in the detection: exoplanets candidates w.r.t. confirmed ones, plus last update of the record content ;
- the method used in the discovery (since it affects the available stellar system description parameters);
- a set of stellar host characteristics (besides sky coordinates): activity, mass, type, metallicity, age, some systemic values, like the global RV (radial velocity) of the system, and so on;
- (exo-)planet parameters, like mass, orbital period, orbit's eccentricity, RV semi-amplitude, time at periastron (for RV detections) or central transit time (for transit method), longitude of periastron, and so on.

The scenario gained some further complexity, but also better stated the idea of modelling datasets and catalogue records in the (exo-)planetary systems subdomain, incorporating requirements from exoplanets atmosphere simulations and (first efforts) observations.

Specific metadata additions were not specified exactly, but a draft model to describe stellar systems was developed, specifically trying to solve the issue in adding metadata elements to describing orbiting celestial objects (Molinaro and Alei et al. (2019)).

The model identified the main concepts and classes as:

- Celestial Object, typed/subclassed as: Star, Binary Star, Planet, Satellite, Brown Dwarf, Trojan, and so on;
- Orbit, as the class keeping the information needed to describe the orbit of a couple (or more) Celestial Objects.

Specific (exo-)planets metadata included: atmosphere with molecular composition, bulk details (mass, radius, . . . ).

A catalogue service that could make use of model metadata is the ExoMerCat (Alei and Claudi et al. (2020)) catalogue (or the other available exoplanets' catalogues), already available in a VO-aware solution[3], but using only simple UCD description of the exposed information.

**TODO:**
mention the involved projects : examples ? GAPS, TESS?

**TODO:**
For GAPS I can add specifics later, if implementation can (re)start

### F.0.1 Client on (MT behalf)

Right now, the meta-data provided within the VOTable allow clients such Aladin or Topcat to run most of the functionalities expected by the user, either for data analysis of plotting. This information is often guess from UCDs, UTypes or columns name. It can also be given by the user. Clients have no expectations of working with full model instances but in some cases models can help to know how quantities in an input table relate to each other. In most cases this is for visualisation, e.g.:

- what is the sky position for this row (what columns contain latitude and longitude, and what sky system are they in)
- what +/-ERR error bars should I plot for these points (what column is a simple error for column A)
- what error ellipses should I plot for these sky positions (what columns provide ra_error, dec_error, ra_dec_corr, or how can I derive those from columns that do exist)
- where do I get the grid information for a column containing a vector of samples so I can label the X axis of a spectrogram (what column or parameter contains an axis vector matching the sample vectors)
- does this table contain sky positions, or HEALPix tiles, or both? What's the best way to represent it on the sky?
- What is the meaning of such URL found out in a table?s

But there are some other places too:

- how do I propagate this sky position to a future epoch (what columns contain pmra, pmdec, and maybe all the associated errors and correlation coefficients)

[3] ivo://ia2.inaf.it/catalogues/exomercat

- what is the error ellipse/oid to use for a sky/Cartesian crossmatch (what columns provide the relevant errors and, if available, correlations)

This usage shows that MANGO must be designed in a way that individual measurements or quantities can be easily be identified as such and manipulated independently of the whole instance.

## G Xmatch tool

The basic cross-match of two astronomical tables consists in associating pairs of sources – one from each table – fulfilling a given angular distance based criterion. In relational algebra terms, it is a theta-join on a distance predicate.

More generally, a cross-match is the association of sources from different tables given their proximity in an astrometric (but also possibly photometric, statistical, ...) parameter space (Pineau and Derriere et al., 2017) .

If proper motions (plus parallax and radial velocities) are available, the cross-match tool may propagate the positions of each table to a common epoch. It may also take into account positional uncertainties to reject the statistically unlikely associations.

In the latter case (cross-match between two tables taking into account positional errors), the tool needs to be able to retrieve the errors associated to the each position in each table.

UCDs may help in identifying the errors associated to a positional columns as shown in table 4.

But this is not sufficient to table with more complex cases based on multi-parameter cases:

- Catalogues like AllWISE provides a co-sigma instead of the correlation factor of the covariance matrix. Co-sigma is the sign of the correlation factor time the square root of the covariance.

- Table fields UCDs may be too loose: for example `stat.error;pos.eq` is often used in place of `phys.angSize.smajAxis;pos.errorEllipse` or `phys.angSize.sminAxis;pos.errorEllipse`.

- The location of the column to be used for the Xmatch can be ambiguous. For instance, if several pairs of position are provided in a table, there is currently no way to associate unambiguously uncertainties with the (right) pair of coordinates.

- Possible ambiguity for circular errors. When the provided uncertainty is the parameter of a circular error, it may be:

| Error type Parameters | UCD | Description |
|---|---|---|
| **Circular error** | | |
| epos | `stat.error;pos.eq` | See "possible ambiguity for circular errors" |
| **Uncorrelated errors** | | |
| eRA | `stat.error;pos.eq.ra` | Error on 'RA cos(Dec)' |
| eDec | `stat.error;pos.eq.dec` | Error on 'Dec' |
| **Correlated errors** | | |
| eRa | `stat.error;pos.eq.ra` | Error on 'RA cos(Dec)' |
| eDec | `stat.error;pos.eq.dec` | Error on 'Dec' |
| corRADec | `stat.covariance; pos.eq.ra; pos.eq.dec` | Correlation factor |
| **Oriented Ellipse** | | |
| a | `phys.angSize.smajAxis; pos.errorEllipse` | Error ellipse semi-major axis |
| b | `phys.angSize.sminAxis; pos.errorEllipse` | Error ellipse semi-minor axis |
| theta | `pos.posAng; pos.errorEllipse` | Error ellipse position angle |

*Table 4:* Table of the different possible representations of positional errors that can be found in astronomical catalogues

  - either the 1 dimensional component on each axis of a symmetric 2-dimensional Gaussian distribution: $\sigma = \sigma_{\alpha\cos\delta} = \sigma_\delta$;
  - or the parameter of the radial error distribution (i.e. of the Rayleigh distribution): $\sigma = \sqrt{\sigma^2_{\alpha\cos\delta} + \sigma^2_\delta}$ with $\sigma_{\alpha\cos\delta} = \sigma_\delta$

- Possible ambiguity on the confidence level The provided error is usually the 1sigma error. It (theoretically) means that the "true" position has:
  - either 68% chances to be at a distance lower than the radial error from the position's mean value.
  - or 39% chances to be inside the error ellipse (or circle) around the position's mean value.

  But depending on the catalogue, the provided error parameters can correspond to different confidence levels.

# H Imported Models Instance

# I Changes from Previous Versions

No previous versions yet.

# References


Alei, E., Claudi, R., Bignamini, A. and Molinaro, M. (2020), 'Exo-MerCat: A merged exoplanet catalog with Virtual Observatory connection', *Astronomy and Computing* **31**, 100370, arXiv:2002.01834.
http://doi.org/10.1016/j.ascom.2020.100370

Arviset, C., Gaudet, S. and IVOA Technical Coordination Group (2010), 'IVOA Architecture Version 1.0', IVOA Note 23 November 2010.
http://doi.org/10.5479/ADS/bib/2010ivoa.rept.1123A

Demleitner, M., Ochsenbein, F., McDowell, J. and Rots, A. (2016), 'Referencing stc in votable', Note 2010-06-18.
http://ivoa.net/documents/Notes/VOTableSTC/

Derriere, S. (2016), 'Referencing stc in votable', Note 2011-05-12.
https://wiki.ivoa.net/internal/IVOA/PhotometryDataModel/NOTE-PPDMDesc-0.2-20110512.pdf

Lindegren, L., Lammers, U., Hobbs, D., O'Mullane, W., Bastian, U. and Hernández, J. (2012), 'The astrometric core solution for the Gaia mission. Overview of models, algorithms, and software implementation', *aap* **538**, A78, arXiv:1112.4139.
http://doi.org/10.1051/0004-6361/201117905

Molinaro, M., Alei, E., Benatti, S., Bignamini, A., Bonnarel, F., Damasso, M., Louys, M., Maris, M. and Nascimbeni, V. (2019), Starting Up a Data Model for Exoplanetary Data, *in* P. J. Teuben, M. W. Pound, B. A. Thomas and E. M. Warner, eds, 'Astronomical Data Analysis Software and Systems XXVII', Vol. 523 of *Astronomical Society of the Pacific Conference Series*, p. 597.
https://ui.adsabs.harvard.edu/abs/2019ASPC..523..597M

Molinaro, M., Benatti, S., Bignamini, A. and Claudi, R. (2020), Discover and Access GAPS Time Series Prototyping for Interoperability, *in* P. Ballester, J. Ibsen, M. Solar and K. Shortridge, eds, 'Astronomical Data Analysis Software and Systems XXVII', Vol. 522 of *Astronomical Society of the Pacific Conference Series*, p. 703.
https://ui.adsabs.harvard.edu/abs/2020ASPC..522..703M

Molinaro, M., Butora, R., Bandieramonte, M., Becciani, U., Brescia, M., Cavuoti, S., Costa, A., Di Giorgio, A. M., Elia, D., Hajnal, A., Gabor,

H., Kacsuk, P., Liu, S. J., Molinari, S., Riccio, G., Schisano, E., Sciacca, E., Smareglia, R. and Vitello, F. (2016), VIALACTEA knowledge base homogenizing access to Milky Way data, *in* G. Chiozzi and J. C. Guzman, eds, 'Software and Cyberinfrastructure for Astronomy IV', Vol. 9913 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 99130H.
http://doi.org/10.1117/12.2231674

Osuna, P. et al. (2006), 'Catalog data model', Catalog Data Model Work Package.
https://wiki.ivoa.net/twiki/bin/view/IVOA/IVAODMCatalogsWP

Pineau, F. X., Derriere, S., Motch, C., Carrera, F. J., Genova, F., Michel, L., Mingo, B., Mints, A., Nebot Gómez-Morán, A., Rosen, S. R. and Ruiz Camuñas, A. (2017), 'Probabilistic multi-catalogue positional cross-match', *aap* **597**, A89, arXiv:1609.00818.
http://doi.org/10.1051/0004-6361/201629219

Salgado, J., Lemson, G. and Demleitner, M. (2016), 'Source dm', Source DM 2016-05-10.
https://wiki.ivoa.net/twiki/bin/view/IVOA/SourceDataModel