Current status of the Characterisation data model

Characterisation metadata

- Should answer the question:

  Where, when, what, how precise and *reliable* are the data according to the usage I will make of them?

  - FoV, Bandpass, Resolution, Quantum Efficiency, etc…

- It is a summary of metadata to be used for data retrieval as in DAL but also for data analysis: resampling, source detections, multi-wavelength analysis, etc…

Organising metadata

- Lists the **properties** of an observation : **coverage, resolution , sampling precision, sensitivity, psf, transmission curve,** etc…
- Defines characterisation axes as ***space, time, wavelength***, '**observable**' which is the measured quantity like flux, photons, counts
- Organise them according to **increasing level of details**

Table representation

- All properties are described on each axis
- The deepest layers are not necessarily provided by data producers
- The first two levels are mandatory

Characterisation table for a 2D image

Axes →

Properties ↓

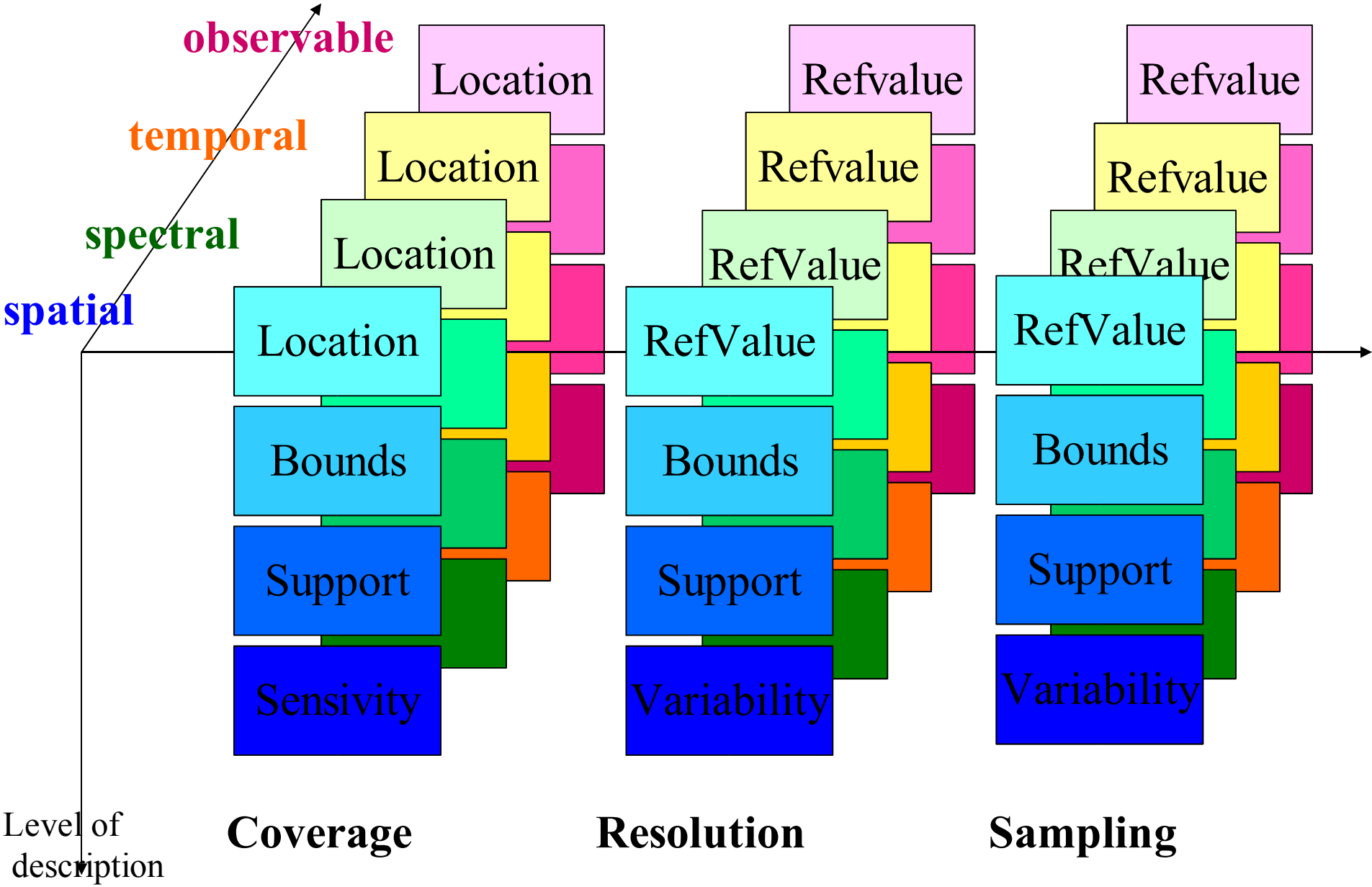
General	Spatial	Temporal	Spectral	Flux/Observable
Location	Central position	Mid- Time	Central wavelength	Average flux
Bounds	Min RA,DEC, Max RA,DEC	Start/stop time	Min Max Wavelength	Saturation, Limiting flux
Support	FOV as accurate array of poly	Array of time intervals	Array of WI intervals	-
Sensitivity	Quantum efficiency (x,y)		Transmission curve(λ)	Function property e.g. Linearity
Filling factor	Effective/ Total area	Dead time		
Resolution	PSF (x,y) or its FWHM	?	FWHM of the band	Statistical error
Sampling	Pixel scale (x,y)	?	FWHM of the band	Quantization

Table 1: 2D-Image Characterization

Characterisation table for a 1D spectrum

General	Spatial	Temporal	Spectral	Flux/Observable
Location	Central position	Mid Time	Central wavelength	Average flux
Bounds	Slit Min Max RA,DEC	Start/stop time	Min Max Wavelength	Saturation, Limiting flux
Support	Slit as accurate array of poly	Array of time intervals	Array of WI intervals	Lowest and highest value
Sensitivity	Response(x,y) n the slit		Quantum eff (lambda)	Function property e.g. Linearity
Filling factor	Effective/ Total area	Dead time		
Resolution	Bound "Size"		LSF or its FWHM	Statistical error
Sampling	Bound "size"		Pixel scale in lambda	Quantization

Table 2: 1D-Spectrum Characterization


observable
temporal
spectral
spatial
Location
Location
Location
Location
Bounds
Support
Sensivity
Refvalue
Refvalue
RefValue
RefValue
Bounds
Support
Variability
Refvalue
Refvalue
RefValue
RefValue
Bounds
Support
Variability
Level of description
Coverage
Resolution
Sampling

Multiple levels needed

- Under **Resolution** concept for spatial axis, we can hook the PSF information
- Using a fine to coarse approach:
  - Lev 4: **Variability** variation of resolution with x,y, λ,T,…
    - e.g. a profile for any position on the characterisation axes, that is one curve for each position, -> Complex maps for diff λ, T etc…
    - A value (FWHM) for each position -> simple map
    - An averaged profile for all positions
  - Lev 3: **Support** intervals of possible values for resolution/ PSF estimation

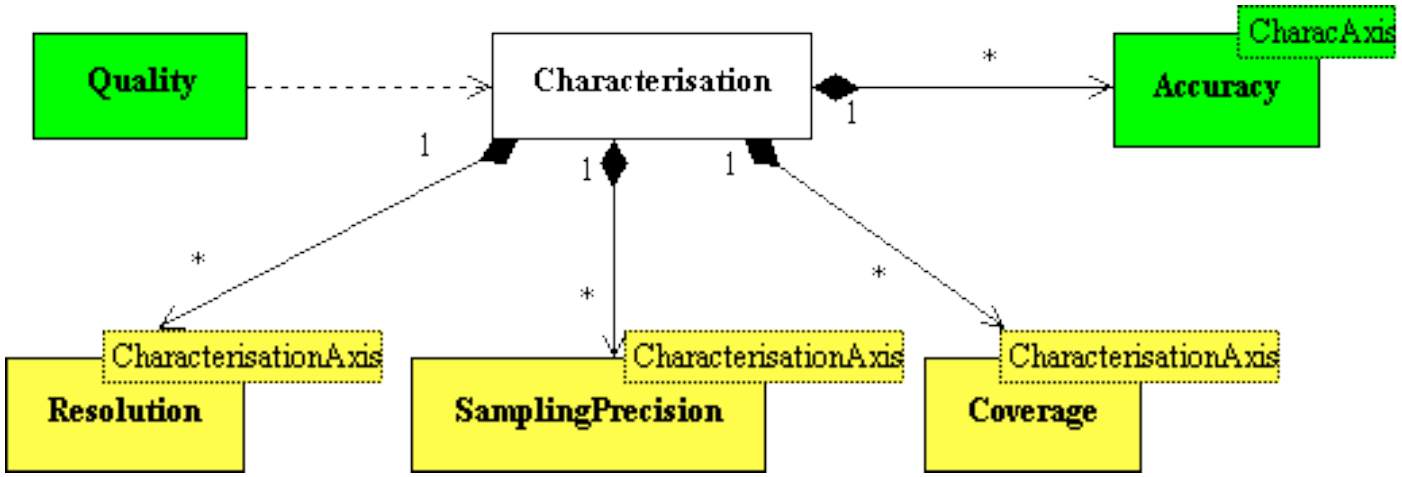
    Could be a list
  - Lev 2: **Bounds** min max possible values within the sampled area
  - Lev 1: **Reference Value** a reference value for the whole field of view (FWHM)
- Similar approach for Line Spread Function
- Use a **Map** object
- A general concept that is useful for many instrumental / ancillary metadata

UML encoding

- The model is a framework to define the various facets of the characterisation
- It suggests a structure for place holders
- On line at

  http://alinda.u-strasbg.fr/Model/Characterisation/DM

Scope of Characterisation (1)



- **Accuracy** is a concept close to **Coverage**
- **Quality** may be expressed as a set of flags defined from Characterisation elements, the instrumental chain and calibration pipeline → It does not depend only on characterisation

Scope of Characterisation (2)

- The **Coverage** , **Sampling** , and **Resolution** concepts apply to different types of observations
  - Images, 1D-spectrum, etc…
  - On-going exercise on IFU, datacubes, etc…
- Characterisation also allows for characterising a data collection using the first 2 levels + Coverage support
- Reuse the model concepts when modelling the different data types for a more homogeneous framework and easier navigation between the different data

Still to be done

- Characterisation does not fully handle complex dependencies of one concept upon several axes (**Maps,** not multi-variable functions)
- Ex: **Sensitivity**
  - One sensitivity variation along each axis assumes implicitely that these are independant which is not always the case
  - It is a multidimensional function of all the characterisation axes.
- Investigate how to describe functions in XML
  - MathML, etc…

Data model and Utypes

- The Utypes correspond to roles in the metadata structure
- They are identified by a path in the DM like:
  - Coverage/spatial/support
  - Or SPATIAL/RESOLUTION/VARIABILITY

- Homogenize with Utype list definition (Alberto Micol)

Links between Characterisation and other parts of DM

- The Characterisation DM relies on the coordinates and region description of *STC*
- Reuses *Quantity* objects for scalar values, maps,....
- Many possibilities currently investigated for XML serialisation (coll. Brian Thomas)
  - Needs some more iterations but in progress

Next steps

- Characterisation Draft document to be reviewed mid-July 05
- Finalize convergence of the U-type list and the DM classes and stabilize them
- Finalize XML serialisation (cf F. Bonnarel's talk)
- Provide reference implementation :
  - VOTable, lib of Java classes ??

Conclusion

- Collaborative effort within the subgroup
- Integration of previous modelling attempts and current discussions ideas
- Many inputs, discussions and feedback between the subgroup participants during the iterations