Cube/Dataset in HE context

Mark Cresitello-Dittmar

Nov. 11, 2023

Preamble

- The Cube data model is the core model for N-Dimensional data. This covers the range from simple images to sparse data cubes.

- The model has been exercised in the context of Spectra, TimeSeries, and Chandra Event lists.

- Recently, there has been an uptick in interest by the High Energy community to have an IVOA model representing their data products, specifically the Event List.

High Energy Data Club

- A small group met at the May 2023 interop in Bologna to discuss the possibility of starting an informal group of High Energy projects to discuss the specific needs of that domain and how well the IVOA standards support these needs. This group includes representatives from X-Ray, Gamma Ray, and Neutrino projects.

- Presentations and meetings:

  - **2023-05-11**: Presentation at the interop, "DM for High Energy Astrophysics", M. Servillat

  - **2023-06-28**: French-VO Workshop in Paris, "IVOA Standards for High Energy Astrophysics"

  - **2023-09-14**: Overview presentation of the Dataset/Cube data models and their applicability to High Energy datasets

  - **2023-10-18**: Discussion and presentations of various HE event data and dependencies on response functions (IRF,ARF,RMF).

- https://wiki.ivoa.net/twiki/bin/view/IVOA/HEGroup

Cube model compatibility

- In general, the cube model seems compatible with the core properties of HE Event lists

  - Time: event detection time

  - Position: local instrument -> celestial

  - Energy: detected 'channel', physical energy

- However, there are varying degrees of dependency on ancillary data, particularly the Response data

  - No dependency, the response is folded in during processing and physical energy included as a property in the event list

  - Heavy dependency, with multiple IRFs associated with events as a function of time (GTI)

Challenge

- It is unclear how to represent this dependency at the IVOA level.

  - The algorithm for applying the IRF data to the detected data is complex and instrument specific. So it is similar, but not quite like Transforms which have well-defined and generally implementable algorithms.

    - How would we describe the algorithm?

  - How to associate the event record to the ancillary data at the proper level.

    - They are not the same shape, and an IRF can be associated with >1 event.

Other interesting questions

- Common issue of different definitions of certain terms

  - Dataset vs DataProduct - no common definition among the group

    - Dataset: Metadata associated with one or more DataProducts

    - DataProduct: The structural description of a particular kind of data. This remains constant regardless of how it was obtained (e.g. Observational, Simulated)

  - "Would not have looked at the Cube model for Event List data.. it isn't binned."