

Converting CAA to SPASE Metadata

Phil Richards and Chris Perry, RAL
SPASE Team Meeting
Cosener's House
Abingdon, UK
July 9-11, 2007



Background

- Integrate UK Cluster Data Centre into VO developments in the UK and Europe
- Investigate VO technologies – ESA SpaceGrid
- SpaceGrid – ESA project to assess relevance of Grid technology to space activities including Solar System Physics



SpaceGrid - Interoperability Issues

- Standardisation of the metadata between different data domains is fundamental
- Common data format for query results required (VOTable)
- Data Mining presents technical problems – e.g. data volume from a query on many data products



UK Cluster Data Centre

- Provides public, proprietary, and ancillary Cluster data and summary plots via web interface – event monitoring
- Provides Summary and Prime Parameter data and JSOC Orbit and Event data
- Supporting data standardisation initiatives:
 - Cluster Science Data System (CSDS)
 - SPASE
 - CAA



ESA Cluster Active Archive (CAA)

- Contains Cluster high resolution data from all instruments
- Suitable for detailed science investigations – best achievable quality
- Developed while all necessary expertise is available – preserve information in metadata
- Data product descriptions – CAA metadata
- Publicly accessible (high res)



CAA Metadata

- Evolution of metadata from:
ISTP, CSDS (CEF), CDPP and SPASE
- Metadata Dictionary (Harvey et al, 2006)
 - Semantic description (compatible with SPASE)
 - Parameter syntax and semantics – information for reading data records and plotting
 - Curation information on data products



Cluster Numerical Data Products

- UKCDC provides Summary and Prime Parameter data products for all Cluster field, particle and wave instruments
 - simple data products (scalars/vectors)
 - currently being ingested into CAA
 - CAA metadata descriptions exist
 - complete descriptions of dataset available
 - BUT SPASE Data Model has different structure and terminology



Converting CAA to SPASE Metadata

- CAA Datasets - SPASE Numerical Data Products
- Product level descriptions
 - no granules (data available via services)
 - ‘coarse-grained’ registry
- CAA Metadata richer than original CSDS metadata (CDF files)
- Different approaches – CAA concepts (hierarchy) and SPASE resources



CAA Concepts and SPASE Resources

CAA

Mission

Observatory

Experiment

Instrument

Dataset

Parameter

File

SPASE

(Observatory Group)

Observatory

-

Instrument

Numerical Data

Physical Parameter

Granule



CAA Parameters - I

- Keywords provide rich description of both semantics and syntax
- Described in terms of concepts:
 - Entity (e.g. Particles, Magnetic Field)
 - Property (e.g. Mass Density, Vector)
 - Fluctuations (e.g. Fourier Spectrum)
- Keywords for parameter description and identification



CAA Parameters - II

- Parameter types – Data or Support Data
- Semantics for vectors, tensors and arrays
 - representation (e.g. Cartesian)
 - sparseness of tensors
 - axis labels for plotting
- Coordinate reference frame
- SI units conversion
- Syntax: array sizes, value type and significant digits



SPASE Parameters - I

- Current Data Model focussed on semantics
- Simple concepts:
 - Measured or Support parameter
 - Measured parameter:
 - * Field
 - * Particle
 - * Photon
 - * Mixed



SPASE Parameters - II

- Parameter description and identification
- Semantics for vectors, tensors and arrays
 - Structure element allows description of each component
- Coordinate reference frame
- SI unit conversion
- Syntax: array sizes



SPASE Example – FGM SP Product I

Spase:

Version: 1.2.0

NumericalData:

*ResourceID: spase://caa/spaseregistry/ndr/CL_SP_FGM

ResourceHeader:

*ResourceName: CL_SP_FGM

ReleaseDate: 2007-02-08T18:54:47.0Z

*Description:

This dataset contains summary 1 minute averaged measurements of the magnetic field vector from the FGM experiment on the Cluster C3 spacecraft. These data have been converted into Cluster Exchange Format from the original Cluster Science Data System Common Data Format (CDF) Prime Parameter files that were made available through the Cluster Science Data System. The metadata has been updated from the CSDS/CDF standard to the CAA to aid compatibility with tools developed for the Cluster Active Archive. Version 01 of this dataset is the initial translation prepared for the launch of the CAA during the second half of 2005.

*Acknowledgement:

Please acknowledge the FGM team and ESA Cluster Active Archive in any publication based upon use of this data



SPASE Example – FGM SP Product II

```
Contact:
  *PersonID:    spase://caa/spaseregistry/pr/ElizabethLucek
  *Role:        PrincipalInvestigator
InformationURL:
  URL:          http://caa.estec.esa.int/caa
AccessInformation:
  RepositoryID: spase://caa/spaseregistry/repr/caa
  Availability: Online
  AccessRights: Open
  AccessURL:
    URL:        http://caa.estec.esa.int/caa
  *Format:      CEF
*ProviderResourceName:    CL_SP_FGM
*ProviderProcessingLevel:  Calibrated
  ProviderVersion:        N/A
*InstrumentID:            spase://caa/spaseregistry/ir/FGM3
*MeasurementType:        MagneticField
TemporalDescription:
  TimeSpan:
    StartDate:            2000-10-01T00:00:00Z
    RelativeEndDate:      -P3M
  *Cadence:              PT60.000000S
```



SPASE Example – FGM SP Product III

```
ObservedRegion:      Earth.Magnetosphere
ObservedRegion:      Earth.Magnetosheath
ObservedRegion:      Earth.Magnetosphere.Magnetotail
ObservedRegion:      Earth.Magnetosphere.Main
ObservedRegion:      Earth.Magnetosphere.Polar
*Caveats:             *C3_CQ_FGM_CAVF
PhysicalParameter:
  *Name:              time_tags
  *ParameterKey:      time_tags__CL_SP_FGM
  *Description:       Interval centred time tag
  *Cadence:           PT60.000000S
  *Units:              s
  *UnitsConversion:   1.0
  Structure:
    *StructureType:   Scalar
    *Size:             1
    *Description:     A scalar quantity of type ISO_TIME
  *Support:           Temporal
```



SPASE Example – FGM SP Product IV

PhysicalParameter:

```
*Name: B_xyz_gse
*ParameterKey: B_xyz_gse__CL_SP_FGM
*Description: Summary (CSDS SP) Cluster C3, Magnetic Field Vector, 1 minute average in GSE
*Cadence: PT60.000000S
*Units: nT
*UnitsConversion: 1.0E-9>T
CoordinateSystem:
  *CoordinateRepresentation: Cartesian
  *CoordinateSystemName: GSE
Structure:
  *StructureType: Vector
  *Size: 3
  *Description: A vector quantity with 3 components of type FLOAT
Measured:
  Field:
    *FieldQualifier: Vector
    *FieldQuantity: Magnetic
```



SPASE Example – FGM SP Product V

PhysicalParameter:

```
*Name: B_nsigma_t
*ParameterKey: B_nsigma_t__CL_SP_FGM
*Description: Summary (CSDS SP) Cluster C3, Normalised magnetic variance: summed component
variances
*Cadence: PT60.000000S
*Units: unitless
*UnitsConversion: 1.0>(variance)
Structure:
  *StructureType: Scalar
  *Size: 1
  *Description: A scalar quantity of type FLOAT
Measured:
  Field:
    *FieldQualifier: Variance
    *FieldQuantity: Magnetic
```



Future Developments

- SPASE Data Descriptions for Cluster high resolution products (CAA)
- Build data provider registry for Cluster data products (UKCDC and CAA)
- Graphical Data Product Descriptions
- Event data – Catalogue descriptions(?)



Summary

- Common inter-lingua essential for interoperability
- SPASE
- CAA metadata provides comprehensive descriptions of CAA datasets
- CAA metadata and SPASE metadata differ in terminology not content – translation feasible
- Tools exist to create SPASE registry of CAA data products

