ODI-Open Data Interface for SAAPS, SEDAT, SPENVIS

D. Heynderickx

DH Consultancy, Leuven, Belgium

P. Wintoft, L. Eliasson

Swedish Institute of Space Physics, Lund, Sweden

H. Evans

ESA/ESTEC, Noordwijk, The Netherlands

Background

- Different data systems at ESA/ESTEC use different data storage and access systems
- ESA Contract ESTEC/RFQ 3-12487/08/NL/AT (1 year duration) to establish a common data backend for:
  - Spacecraft Anomaly Analysis and Prediction System (SAAPS): http://www.lund.irf.se/saaps
  - Space Environment Database and Analysis Tool (SEDAT): http://ww.ukssdc.ac.uk/sedat/
  - Space Environment Information System (SPENVIS) http://www.spenvis.oma.be/

System overview: data storage

- Central database in MySQL stores:
  - List of all datasets
  - Metadata for datasets and for all quantities in datasets
  - Data tables ingested from data files
  - List of ingested data files
- Data files are downloaded with wget, then parsed by php scripts; automated updates with cron jobs
- Metadata follow ISTP guidelines spdf.gsfc.nasa.gov/sp_use_of_cdf.html complemented with PRBEM craterre.onecert.fr/prbem/home.html

Data access

- (My)SQL is supported by most programming languages and tools: php, C++, Java, Perl, IDL, MATLAB, MS Excel, …
- SAAPS is built in Java, straightforward connection to SQL database
- SPENVIS currently uses IDL routines to parse data files; is being replaced by php script using mysqli
- SEDAT is a modular system built with Java (GUI), Perl, IDL (data access and processing), C++ (for data system); all data access will be done with MySQL queries

Data table creation

- Dataset description is done in CDF type skeleton files; CDF is NASA GSFC Common Data Format (http://cdf.gsfc.nasa.gov)
- Skeleton files are also used for non-cdf files so that a common parser routine can be used for all datasets
- Php routines parse skeleton files and create data tables and foreign keys
- Skeleton table completely describes dataset contents
- All datasets in ODI can be exported in CDF (and ASCII) files

```
#GLOBALattributes

! Attribute          Entry       Data
! Name               Number      Type        Value
! ---------          ------      ----        -----

  "Acknowledgement"     1:    CDF_CHAR     { "For terms and conditions " -
                                             "for the use and dissemination " -
                                             "of this dataset, please " -
                                             "contact the PI directly." } .

  "ADID_ref"            1:    CDF_CHAR     { "NSSD0241" } .

  "Data_type"           1:    CDF_CHAR     { "K0>Key Parameter" } .

  "Data_version"        1:    CDF_CHAR     { "1" } .

  "Descriptor"          1:    CDF_CHAR     { "SEM>Space Environment " -
                                             "Monitor" } .

  "Discipline"          1:    CDF_CHAR     { "Space " -
                                             "Physics>Magnetospheric " -
                                             "Science" } .

  "Generated_by"        1:    CDF_CHAR     { " " } .

  "Generation_date"     1:    CDF_CHAR     { " " } .

  "HTTP_LINK"           1:    CDF_CHAR     { "http://goes.ngdc.noaa.gov/data/avg/" } .

  "Instrument_type"     1:    CDF_CHAR     { "Particles (space)" }
                        2:    CDF_CHAR     { "Magnetic Fields (space)" } .

  "LINK_TEXT"           1:    CDF_CHAR     { "GOES05 SEM 5 minute average " -
                                             "data available at "} .
```

```
! Variable          Data      Number                    Record   Dimension
! Name              Type     Elements  Dims  Sizes  Variance  Variances
! --------          ----     --------  ----  -----  --------  ---------

  "Position"      CDF_FLOAT      1       1     3        T         T

  ! Attribute       Data
  ! Name            Type       Value
  ! --------        ----       -----

    "AVG_TYPE"    CDF_CHAR     { "standard" }
    "CATDESC"     CDF_CHAR     { "Position of the GOES05 satellite in " -
                                 "geographic coordinates" }
    "DEPEND_0"    CDF_CHAR     { "Epoch" }
    "DICT_KEY"    CDF_CHAR     { "position>geographic" }
    "DISPLAY_TYPE"
                  CDF_CHAR     { "time_series" }
    "FIELDNAM"    CDF_CHAR     { "Satellite position (GEO)" }
    "FILLVAL"     CDF_FLOAT    { -1.0e+31 }
    "FORMAT"      CDF_CHAR     { "F12.3" }
    "LABL_PTR_1"
                  CDF_CHAR     { "Position_LABL_1" }
    "SCALETYP"    CDF_CHAR     { "linear" }
    "SI_conversion"
                  CDF_CHAR     { "1.0e3>m" }
    "UNITS"       CDF_CHAR     { "km" }
    "VALIDMIN"    CDF_FLOAT    { -50000.0 }
    "VALIDMAX"    CDF_FLOAT    {  50000.0 }
    "VAR_NOTES"   CDF_CHAR     { "Origin = Earth's center of mass. X = " -
                                 "Intersection of Greenwich meridian and" -
                                 " geographic equator. Z = Geographic " -
                                 "North Pole. Y = completes a " -
                                 "right-handed Cartesian triad" }
    "VAR_TYPE"    CDF_CHAR     { "data" } .
```

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

SQL query:

```
SELECT *
FROM `dataset`
LIMIT 0 , 30
```

[Edit] [Explain SQL] [Create PHP Code]

Query results operations

Print view Print view (with full texts) Export

Show : 30 row(s) starting from record # 0

in horizontal mode and repeat headers after 100 cells

Sort by key: None Go

←T→	id	name	epoch_start	epoch_end	creation_date
	3	dataset_ns41_bdd2r	63143625657974	63331718218041	2009-09-02 11:54:25
	6	dataset_index_dst	61756992000000	63421570800000	2009-09-03 12:13:37
	11	dataset_index_kpap_3h	60967987200000	63419468400000	2009-09-06 18:49:53
	13	dataset_index_kpap_1d	60967987200000	63418896000000	2009-09-06 19:17:02
	17	dataset_xmm_rm	63114205202095	63413470931973	2009-09-08 15:20:55
	21	dataset_goes_g05_5	62672140800000	62711452500000	2009-09-09 10:50:19
	23	dataset_index_omni2	61946294400000	63413622000000	2009-09-09 11:09:34

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

Showing rows 5820 - 5849 (5,916 total, Query took 0.0750 sec)

SQL query:

```
SELECT *
FROM `dataset_file`
LIMIT 5820 , 30
```

[Edit] [Explain SQL] [Create PHF

Query results operations

Print view Print view (with full texts) Export

<< < Show : 30 row(s) starting from record # 5850 > >>

in horizontal mode and repeat headers after 100 cells

Sort by key: None Go

←T→	id	filename	dataset_id	filedate
	6755	G0558701.TXT	21	2005-07-08 09:29:23
	6756	G0558702.TXT	21	2005-07-08 09:29:26
	6757	G0558703.TXT	21	2005-07-08 09:29:29
	6758	omni2_h0_mrg1hr_19630101_v01.cdf	23	2009-03-11 03:11:00
	6759	omni2_h0_mrg1hr_19630701_v01.cdf	23	2009-03-11 03:11:00
	6760	omni2_h0_mrg1hr_19640101_v01.cdf	23	2009-03-11 03:11:00
	6761	omni2_h0_mrg1hr_19640701_v01.cdf	23	2009-03-11 03:11:00
	6762	omni2_h0_mrg1hr_19650101_v01.cdf	23	2009-03-11 03:11:00

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

	Field	Type	Collation	Attributes	Null	Default
	cdf_epoch	double			No	
	epoch	datetime			Yes	NULL
	millisec	smallint(6)			Yes	NULL
	position_1	double			Yes	NULL
	position_2	double			Yes	NULL
	position_3	double			Yes	NULL
	position_quality	int(11)			Yes	NULL
	b_calc	double			Yes	NULL
	b_eq	double			Yes	NULL
	i	float			Yes	NULL
	l	float			Yes	NULL
	l_star	float			Yes	NULL
	mlt	float			Yes	NULL
	alpha	float			Yes	NULL
	alpha_eq	float			Yes	NULL
	fpdo_1	float			Yes	NULL
	fpdo_2	float			Yes	NULL
	fpdo_3	float			Yes	NULL
	fpdo_4	float			Yes	NULL
	fpdo_5	float			Yes	NULL
	fpdo_6	float			Yes	NULL
	fpdo_7	float			Yes	NULL
	fpdo_8	float			Yes	NULL
	fpdo_9	float			Yes	NULL
	fpdo_10	float			Yes	NULL
	fpdo_11	float			Yes	NULL
	fpdo_quality_1	int(11)			Yes	NULL
	fpdo_quality_2	int(11)			Yes	NULL
	fpdo_quality_3	int(11)			Yes	NULL

epoch	millisec	ap	ap_quality	cp	cp_quality	c9	c9_quality	ssn	ssn_quality	f107	f107_quality	dataset_file_id
1991-03-02 00:00:00	0	9	0	0.5	0	2	0	93	0	207.5	0	3114
1991-03-03 00:00:00	0	6	0	0.3	0	1	0	71	0	206.4	0	3114
1991-03-04 00:00:00	0	9	0	0.5	0	2	0	55	0	218.9	2	3114
1991-03-05 00:00:00	0	22	0	1.1	0	5	0	74	0	208.1	0	3114
1991-03-06 00:00:00	0	24	0	1.2	0	6	0	88	0	206.7	0	3114
1991-03-07 00:00:00	0	24	0	1.2	0	6	0	131	0	214.5	0	3114
1991-03-08 00:00:00	0	17	0	0.9	0	4	0	146	0	209.1	0	3114
1991-03-09 00:00:00	0	25	0	1.2	0	6	0	156	0	215.7	0	3114
1991-03-10 00:00:00	0	21	0	1.1	0	5	0	159	0	222.8	0	3114
1991-03-11 00:00:00	0	4	0	0.2	0	1	0	167	0	221.9	0	3114
1991-03-12 00:00:00	0	17	0	0.9	0	4	0	163	0	228.7	0	3114
1991-03-13 00:00:00	0	27	0	1.2	0	6	0	152	0	239.1	0	3114
1991-03-14 00:00:00	0	7	0	0.4	0	2	0	161	0	241.6	0	3114
1991-03-15 00:00:00	0	6	0	0.3	0	1	0	182	0	242.2	0	3114
1991-03-16 00:00:00	0	8	0	0.4	0	2	0	202	0	258.5	0	3114
1991-03-17 00:00:00	0	13	0	0.8	0	4	0	167	0	245.4	0	3114
1991-03-18 00:00:00	0	9	0	0.5	0	2	0	168	0	274.8	0	3114
1991-03-19 00:00:00	0	12	0	0.7	0	3	0	155	0	264.9	0	3114
1991-03-20 00:00:00	0	12	0	0.7	0	3	0	173	0	254.2	0	3114
1991-03-21 00:00:00	0	26	0	1.2	0	6	0	167	0	253.1	0	3114
1991-03-22 00:00:00	0	20	0	1	0	5	0	179	0	257.7	0	3114
1991-03-23 00:00:00	0	11	0	0.6	0	3	0	179	0	233.4	0	3114
1991-03-24 00:00:00	0	161	0	2	0	9	0	154	0	260.5	0	3114
1991-03-25 00:00:00	0	130	0	1.9	0	8	0	146	0	235.2	0	3114
1991-03-26 00:00:00	0	114	0	1.9	0	8	0	153	0	229.4	0	3114
1991-03-27 00:00:00	0	31	0	1.3	0	6	0	137	0	203	0	3114
1991-03-28 00:00:00	0	20	0	1	0	5	0	129	0	197.7	0	3114
1991-03-29 00:00:00	0	4	0	0.2	0	1	0	140	0	192.9	0	3114
1991-03-30 00:00:00	0	26	0	1.2	0	6	0	141	0	201.3	0	3114
1991-03-31 00:00:00	0	9	0	0.5	0	2	0	115	0	194.7	0	3114

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

id	dataset_id	attribute	data_type_id	entry_number	value
581	21	Acknowledgement	14	1	NOAA
582	21	ADID_ref	14	1	NSSD0241
583	21	Data_type	14	1	K0>Key Parameter
584	21	Data_version	14	1	1
585	21	Descriptor	14	1	SEM>Space Environment Monitor
586	21	Discipline	14	1	Space Physics>Magnetospheric Science
587	21	Generated_by	14	1	
588	21	Generation_date	14	1	
589	21	HTTP_LINK	14	1	http://goes.ngdc.noaa.gov /data/avg/
590	21	Instrument_type	14	1	Particles (space)
591	21	Instrument_type	14	2	Magnetic Fields (space)
592	21	LINK_TEXT	14	1	GOES05 SEM 5 minute average data available at
593	21	LINK_TITLE	14	1	NGDC
594	21	Logical_file_id	14	1	
595	21	Logical_source	14	1	G0E505_SEM_G5
596	21	Logical_source_description	14	1	5 minute resolution data from SEM instrument onboa...
597	21	Mission_group	14	1	GOES
598	21	MODS	14	1	
599	21	Parents	14	1	
600	21	PI_affiliation	14	1	NOAA
601	21	PI_name	14	1	Harold Leinbach
602	21	Project	14	1	ISTP>International Solar-Terrestrial Physics
603	21	Rules_of_use	14	1	For terms and conditions for the use and dissemina...

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

SQL query:

```
SELECT *
FROM `data_type`
LIMIT 0 , 30
```

[Edit] [Explain SQL] [Create PH

Query results operations

Print view Print view (with full texts) Export

Show : 30 row(s) starting from record # 0

in horizontal mode and repeat headers after 100 cells

Sort by key: None Go

←T→	id	cdf_type	mysql_type	sedat_type	idl_type
	1	CDF_BYTE	TINYINT	I	BYTE
	2	CDF_INT1	TINYINT	I	INT
	3	CDF_UINT1	TINYINT UNSIGNED	I	UINT
	4	CDF_INT2	SMALLINT	I	INT
	5	CDF_UINT2	SMALLINT UNSIGNED	I	UINT
	6	CDF_INT4	INT	I	LONG
	7	CDF_UINT4	INT UNSIGNED	I	ULONG
	8	CDF_REAL4	FLOAT	F	FLOAT
	9	CDF_FLOAT	FLOAT	F	FLOAT
	10	CDF_REAL8	DOUBLE	F	DOUBLE
	11	CDF_DOUBLE	DOUBLE	F	DOUBLE
	12	CDF_EPOCH	DATETIME_millisec_SMALLINT	F	DOUBLE
	13	CDF_EPOCH16	DATETIME_pico_BIGINT	F	DOUBLE
	14	CDF_CHAR	TEXT	C	STRING
	15	CDF_UCHAR	TEXT	C	STRING

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

id	dataset_id	data_type_id	name	number_elements	dims	sizes	record_var
139	11	12	Epoch	1	0		T
140	11	9	Kp	1	0		T
141	11	4	Kp_Quality	1	0		T
142	11	4	ap	1	0		T
143	11	4	ap_Quality	1	0		T
155	13	12	Epoch	1	0		T
156	13	4	Ap	1	0		T
157	13	4	Ap_Quality	1	0		T
158	13	9	Cp	1	0		T
159	13	4	Cp_Quality	1	0		T
160	13	4	C9	1	0		T
161	13	4	C9_Quality	1	0		T
162	13	4	SSN	1	0		T
163	13	4	SSN_Quality	1	0		T
164	13	9	F107	1	0		T
165	13	4	F107_Quality	1	0		T
262	17	12	Epoch	1	0		T
263	17	11	Position	1	1	3	T
264	17	14	Position_LABL_1	4	1	3	F
265	17	6	Position_Quality	1	0		T
266	17	11	B_Calc	1	0		T
267	17	11	B_Eq	1	0		T
268	17	9	I	1	0		T
269	17	9	L	1	0		T
270	17	9	L_star	1	0		T
271	17	9	MLT	1	0		T
272	17	9	Alpha	1	0		T
273	17	9	Alpha_Eq	1	0		T
274	17	9	FPDO	1	1	11	T
275	17	9	FPDO_Energy	1	2	11,2	F

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

←T→	id	name	variable_metadata_id	data_type_id	value
	112	CATDESC	9	14	Default time
	113	DICT_KEY	9	14	time>Epoch
	114	FIELDNAM	9	14	Epoch
	115	FILLVAL	9	11	-1.0e+31
	116	SCALETYP	9	14	linear
	117	SI_conversion	9	14	1.0e-3>s
	118	UNITS	9	14	ms
	119	VALIDMIN	9	12	10-Dec-2000 00:00:57.974
	120	VALIDMAX	9	12	01-Jan-2020 00:00:00.000
	121	VAR_TYPE	9	14	support_data
	122	VAR_NOTES	9	14	Epoch, UT
	123	AVG_TYPE	10	14	standard
	124	CATDESC	10	14	Position of the satellite in geographic coordinate...
	125	DEPEND_0	10	14	Epoch
	126	DICT_KEY	10	14	position>geographic
	127	DISPLAY_TYPE	10	14	time_series
	128	FIELDNAM	10	14	Satellite position (GEO)
	129	FILLVAL	10	9	-1.0e+31
	130	FORMAT	10	14	F8.3
	131	LABL_PTR_1	10	14	Position_LABL_1
	132	SCALETYP	10	14	linear
	133	SI_conversion	10	14	1.0e3>m
	134	UNITS	10	14	km
	135	VALIDMIN	10	9	-30000.0
	136	VALIDMAX	10	9	30000.0
	137	VAR_TYPE	10	14	data
	138	VAR_NOTES	10	14	Origin = Earths center of mass. X = Intersection o...
	139	CATDESC	11	14	Position labels
	140	DICT_KEY	11	14	label>position
	141	FIELDNAM	11	14	Position_LABL_1

phpMyAdmin

Database

ODI (15)

ODI (15)

- dataset
- dataset_file
- dataset_goes_g05_5
- dataset_index_dst
- dataset_index_kpap_1d
- dataset_index_kpap_3h
- dataset_index_omni2
- dataset_metadata
- dataset_ns41_bdd2r
- dataset_xmm_rm
- data_type
- variable_attribute
- variable_depend
- variable_metadata
- variable_nrv

id	variable_metadata_id	nrv_index	nrv_pos	nrv_value
1	11	1	1	Xgeo
2	11	2	2	Ygeo
3	11	3	3	Zgeo
4	22	1	1,1	1.300
5	22	2	1,2	5.300
6	23	1	1	1.3 - 5.3 MeV
7	25	1	1	1.0
8	27	1	1	1.18
9	27	2	2	1.66
10	27	3	3	2.29
11	27	4	4	3.52
12	27	5	5	4.75
13	28	1	1	1.180 MeV
14	28	2	2	1.660 MeV
15	28	3	3	2.290 MeV
16	28	4	4	3.520 MeV
17	28	5	5	4.750 MeV
18	30	1	1	1.0
19	30	2	2	1.0
20	30	3	3	1.0
21	30	4	4	1.0
22	30	5	5	1.0
23	32	1	1	0.1
24	32	2	2	0.19
25	32	3	3	0.35
26	32	4	4	0.65
27	32	5	5	1.22
28	32	6	6	2.29
29	32	7	7	4.28
30	32	8	8	8.0

UP

SPENVIS Project: TEST

Data base interface

Access

Output

Help

Data bases on the space environment

Available data sets

	Dataset	Start Time	End Time
○	High resolution energetic particle data from the BDD2R instrument onboard the NS41 satellite	2000-12-10 00:00:57.974	2006-11-25 23:56:58.041
○	WDC-Kyoto Dst Index	1957-01-01 00:00:00.000	2009-09-30 23:00:00.000
○	NGDC Kp and ap indices	1932-01-01 00:00:00.000	2009-09-06 15:00:00.000
○	NGDC Kp and ap indices	1932-01-01 00:00:00.000	2009-08-31 00:00:00.000
○	High resolution particle flux data from ERMD instrument onboard XMM satellite	2000-01-04 11:40:02.095	2009-06-29 05:02:11.973
◉	5 minute resolution data from SEM instrument onboard GOES05 satellite (uncorrected G file series	1986-01-01 00:00:00.000	1987-03-31 23:55:00.000
○	OMNI Combined, Definitive, 1AU Hourly IMF, Plasma, Energetic Proton Fluxes, and Solar and Magnetic Indices	1963-01-01 00:00:00.000	2009-06-30 23:00:00.000

Channels >

Tool developed by

BIRA-IASB DCH