

The GSFC IUE Final Archive Catalogue and IUE Label "Modification"

Karen Levay

September 19, 1990

1 Introduction

Plans for generating the IUE Final Archive are well underway at all three of the IUE Observatories; SERC, VILSPA, and GSFC. The IUE Final Archive will include over twelve years of IUE data, the IUE Final Archive Catalogue, and other documents. The IUE Final Archive Catalogue at GSFC will be an integral part of IUE research and the Final Archive Processing System. The catalogue will be built cooperatively with the VILSPA observing station staff and serve the following three functions:

- A data store for IUE observational data, processing and archival information and journal references. In addition to serving the community as a means of finding interesting data for research, the catalogue will supply correct data to augment the IUE image label.
- A means of tracking the status of the final archive data.
- An index of optical disk locations for data, scanned scripts, and on-line documentation.

These functions were identified from a survey made of existing IUE data repositories and potential catalogue users. Some requirements for the catalogue were defined by the Final Archive Definition Committee (FAD) and the IUE Data Coordination Group (DCG). The data elements identified were placed into a conceptual design which formulates an Entity-Relationship model. The Entity-Relationship model is discussed further in Section 3 of this report.

2 Survey

A survey of existing data repositories was made in the summer and fall of 1989. The results of the survey have been documented in the report "Current Database Environment" and distributed to GSFC IUE Observatory staff members. Major repositories identified were the IUE Automated Information Management System (IUEAIMS) databases, the image processing logs (IPL), the observing scripts, the image labels, the journal database, the proposer databases and the calibration database. The IUEAIMS databases contain scientific, spacecraft, processing and data management information for each image. The IUE Merged Logs are created from the IUEAIMS databases. The image processing logs are records of the processing status of each image. Since 1988, a machine readable version of the IPL is available. Prior to that time, only a handwritten IPL is available for each image. The journal database contains bibliographic information of papers written using or referring to IUE data. The proposer databases contain information concerning all proposals received by the IUE Project (e.g. titles, investigators, sponsoring institutions). The calibration database contains notes used by the IUE calibration group about all images acquired for the calibration effort.

The IUE Observatory staff members and IUE users were interviewed to identify additional data elements and ways that different groups in the observatory could use the catalogue. Suggestions were solicited from Guest Observers through the IUE Newsletter and the IUE Users Committee.

3 Conceptual Design

The second design phase was to create a conceptual model of the catalogue. A conceptual model involves the examination of the data elements, their relationships and grouping related data elements together to form entities. Entities were normalized, that is, designed so that each piece of data is stored once in a single location. Data is duplicated only for the purpose of relating the entities to each other. Some entities were formed so that data could be stored more efficiently, that is if a piece of information is relevant to a small number of images a separate entity should be formed.

The basic unit of IUE data is the spectral image. Images consist of one or more exposures. In order to normalize the data, data elements that are unique to an image were placed in an entity named IMAGE. Each row of

the IMAGE entity corresponds to a unique image and is identified by the camera and image number.

Data elements unique only within an exposure were placed in an entity named EXPOSURE. Each row of the EXPOSURE entity corresponds with a unique exposure and is identified by the camera, image number and the exposure start time. An exposure results in an identifiable unique spectrum on the image. Images acquired by parts (segmented images) result in one unique spectrum on the image and are defined as one row in the entity. The relationship between the IMAGE and EXPOSURE tables is a one to many relationship.

The other type of IUE data is the FES image. Data defining the FES image is stored in the FES table and is identified by the camera and image number. Many FES images are taken in association with a spectral image. If such a relationship exists, the FES entity is related to the IMAGE entity in a one to one relationship.

Target information stored in the IMAGE, EXPOSURE and FES entities will be that defined by the Guest Observer. Several potential users of the Final Archive Catalogue noted that many targets may have more than one valid classification (e.g. object classification and spectral type). The entity USER'S DESIGNATION will contain alternate data as contributed by users. It will also allow users of the data to comment on an image. This entity relates to EXPOSURE and IMAGE through the image number.

Several entities were defined to hold data for certain categories of exposures. Elements relating to only trailed exposures will be stored in the TRAIL entity, those relating to off-set reference points will be stored in the OFF-SET REFERENCE entity, and notes about images used for calibration will be stored in the CALIBRATION entity. Some images will have operation notes associated with them. The Telescope Operator (TO) comments from the image label will be stored for each image in the TO COMMENTS entity. Each row of these entities will be identified by the camera and image number.

GSFC Primary Investigators are being solicited for previously unrecorded comments or insights about data acquired under their observing programs. Comments received will be stored in the GUEST OBSERVER COMMENTS entity and be related to the IMAGE entity by camera and image number.

An entity containing a homogeneous set of information for targets taken by the IUE will be called the HOMOGENEOUS TARGET entity. The HOMOGENEOUS TARGET entity will be related to the exposure table. The homogeneous data is being supplied by CDS. The requirements for the ho-

mogeneous data were set by Dr. Michael Barylak and VILSPA staff.

Another major component of the catalogue will be journal references. The references themselves will be stored in the JOURNAL entity. Associated co-authors, objects and images will be stored in separate entities and will be related to the JOURNAL entity by an identifying number. The ALIAS entity will contain the homogeneous object names and associated aliases for those names allowing the journal references to be related to the HOMOGENEOUS TARGET entity.

Another major component of the catalogue is the storage of processing and data product information. Entities storing processing information (PROCESSING), images requested for on-demand processing (REDO AND REDO REQUESTOR), and associated data products (DATA PRODUCT) will be related to each other and the IMAGE entity by camera and image number. The entity FINAL ARCHIVE TRACKING will contain flags indicating the status of the image (e.g. have the core items been verified, has the image been visually inspected).

The final major component of the catalogue are the indices to data locations. The entity LOCATION will store locations to image data and scanned scripts on the optical disk and other storage media. This entity relates to other entities in the catalogue by camera and image number. Locations of various documentation stored on-line will be kept in to be defined entity.

The resulting Entity-Relationship Diagram is contained in Appendix One and the associated data dictionary of entities of general interest to the community is contained in Appendix Two of this report.

4 Physical Design

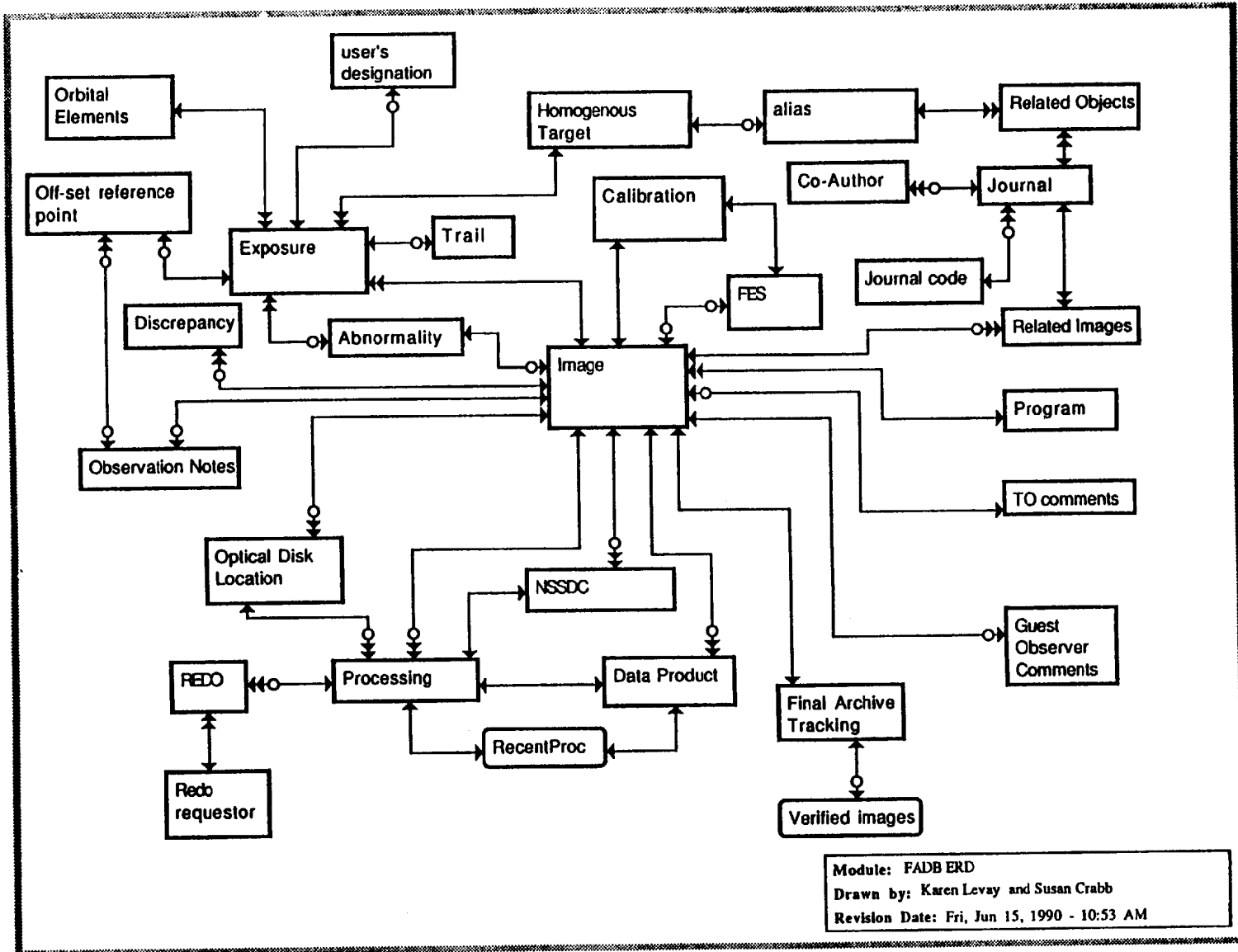
GSFC has completed the initial physical design of the GSFC Final Archive Catalogue using the Ingres relational database management system. The entities defined during the conceptual design phase have been mapped directly into Ingres tables to create a prototype catalogue. The prototype was populated from existing machine readable data repositories identified during survey of existing data repositories. Currently, efficient query access and report generation procedures are under development.

5 Core Items

The image label is the major source of machine readable information about each image and was used to create the IUEAIMS database or old database. Some designated data was subsequently checked for accuracy. While generally accurate, the label may be incomplete or contain inaccurate data. Some useful information is not contained in the label but only in handwritten form on the observing scripts. The FAD committee recommended that some means of completing or correcting the image label without replacing existing data be implemented as part of the IUE Final Archive. The DCG committee defined a set of data elements or Core Items that are either required input for image processing or extremely valuable for scientific analysis (listed in Appendix 3). The Core Items will be extracted from the raw image label and the observing scripts and used to update the already populated IUE Final Archive Catalogue. The result of this process will provide an accurate source of processing parameters necessary for the Final Archive Processing effort and replaces the need to perform actual label modifications. These Core Items are to be certified as accurate and included as FITS keywords in the label of processed images in the IUE Final Archive, thus fulfilling the recommendation of the FAD committee. The final archive processing system will get the necessary processing parameters from the IUE Final Archive Catalogue using the Ingres interface software (SQL) and FORTRAN.

6 Verification of the IUE Final Archive Catalogue

The Core Items will be the initial focus of the IUE Final Archive Catalogue verification. GSFC and VILSPA will follow similar procedures for verifying the accuracy of the Core Items. The data will be re-extracted from the image label, compared with the data in the existing record of the data base and with the data on the observing scripts. The initial comparisons will be done automatically. Discrepancies will be resolved interactively.



7 Appendix 1 - Entity Relationship Diagram

8 Appendix 2 - Data Elements

The data elements listed below are a subset judged to be of the most interest to the IUE community. Other entities exist for internal use.

IMAGE ENTITY

The IMAGE entity holds data unique to a specific image.

- CAMERA
Camera number where 1=LWP, 2=LWR, 3=SWP, 4=SWR
- IMAGE
Image number
- NUMBER OF EXPOSURES
Number of unique spectra or exposures associated with the image
- THDA AT TIME OF READ
THDA (temperature of camera head amplifier) at time of read.
- OBSERVING STATION
Observing station:GSFC, VILSPA. For collaborative images the reading station will be used.
- UVC VOLTAGE
Voltage of the UVC of the camera
- OBSERVERS NAME
Name of the actual observer (not necessarily the primary investigator)
- PROGRAM ID
Program identification code
- RIGHT ASCENSION
Right ascension of the target as specified by the Guest Observer. Stored as radian degrees but available in other formats interactively or in printed output.

- **DECLINATION**

Declination of the target as specified by the Guest Observer. Stored as radian degrees but available in other formats interactively or in printed output.

- **TRACKING MODE**

- **READ GAIN**

Gain used for reading the image: High, Low.

- **EXPOSURE GAIN**

Camera gain: Maximum, Medium, Minimum.

- **RELEASE DATE**

Date image becomes available for public use.

EXPOSURE ENTITY

The EXPOSURE entity contains a row for each exposure made within an image, so a common single exposure would have one row, a standard double exposure would have two rows, while a multiple exposure would have a row for each unique spectrum in that row. The VILSPA station has not yet completed the definition of their database and may not include a row for each multiple exposure.

- **CAMERA**

Camera number where 1=LWP, 2=LWR, 3=SWP, 4=SWR

- **IMAGE**

Image number

- **SPECNUM**

Number of the order in which the spectrum was acquired for the image.

- **DATETIME**

Date and exposure start time for the exposure.

- **DISPERSION**

Dispersion used: High, Low.

- **APERTURE**
Aperture used: Large Small.
- **APERTURE STATUS**
Status of the large aperture: Open, Closed.
- **EXPOSURE MODE**
Exposure mode and observing technique: Point or Extended Source, Trailed Spectrum, Multiple Exposure, Segmented Exposure (image acquired in parts).
- **LENGTH OF EXPOSURE**
Length of the exposure.
- **THDA START**
THDA (temperature of camera head amplifier) at starting time of the exposure.
- **FES MODE**
Working mode of the FES: Fast/Slow track - Overlap/Underlap.
- **FES COUNT**
Fine Error Sensor counts of the target at the reference point.
- **FOCUS**
Step value at starting time of the observation.
- **CATALOG**
Catalog of the target as specified by the Guest Observer.
- **OBJECT NAME**
Object name of the target as specified by the Guest Observer.
- **MAGNITUDE**
Magnitude of the target as specified by the Guest Observer for GSFC images.
- **LUMINOSITY**
Luminosity of the target as specified by the Guest Observer for GSFC images.

- **SPECTYPE**
Spectral type of the target as specified by the Guest Observer for GSFC images.
- **EBV**
Color excess of the target as specified by the Guest Observer for GSFC images.
- **OBJ CLASS**
Object classification as specified by the Guest Observer.
- **GSX**
FES x-position of the guide star.
- **GSY**
FES y-position of the guide star.
- **GSCOUNTS**
FES counts of the guide star.
- **GMODE**
FES mode: Fast/Slow track - Overlap/Underlap, No guiding.
- **OBSERVING STATION**
Observing station:GSFC, VILSPA. For collaborative images the reading station will be used.
- **DATABKG**
Maximum DN value measured in the spectrum. (Value determined by processing system and will be consistent between both stations).
- **DATAcnt**
Measured value of background in DN's. (Value determined by processing system and will be consistent between both stations).

FES ENTITY

The FES entity holds information unique to FES images.

- **CAMERA**
Camera number: 8=FES1, 9=FES2.
- **IMAGE**
Image number.
- **SIZE**
FES Size.
- **NUMBER**
- **BITRATE**
Bitrate.
- **SAMPLES**
Samples.
- **DATETIME**
Date and exposure start time for the exposure.
- **RIGHT ASCENSION**
Right ascension of the target as specified by the Guest Observer. Stored as radian degrees but available in other formats interactively or in printed output.
- **DECLINATION**
Declination of the target as specified by the Guest Observer. Stored as radian degrees but available in other formats interactively or in printed output.
- **CATALOG**
Catalog of the target as specified by the Guest Observer.
- **OBJECT NAME**
Object name of the target as specified by the Guest Observer.
- **MAGNITUDE**
Magnitude of the target as specified by the Guest Observer for GSFC images.

- **LUMINOSITY**

Luminosity of the target as specified by the Guest Observer for GSFC images.

- **SPECTYPE**

Spectral type of the target as specified by the Guest Observer for GSFC images.

- **EBV**

Color excess of the target as specified by the Guest Observer for GSFC images.

- **OBJ CLASS**

Object classification as specified by the Guest Observer.

- **OBSERVERS NAME**

Name of the actual observer (not necessarily the primary investigator)

- **PROGRAM ID**

Program identification code

- **FOCUS**

Step value at starting time of the observation.

- **COMMENTS**

Comments.

- **ASSOCIATED CAM**

Camera number of associated spectral image.

- **ASSOCIATED IMAGENO**

Image number of associated spectral image.

TRAIL ENTITY

The trailed entity will contain information about those images that were trailed.

- **CAMERA**

- **IMAGE**

- **TRL RATE**
Trail rate.
- **TRAILS**
Number of trails.

ABNORMAL ENTITY

The abnormality codes are a warning to the user of IUE archived data to examine the data carefully before use. To help the user with this examination, the IUE Project will attempt to flag images with certain known abnormalities, but can not guarantee that all images with abnormalities will be flagged. The following abnormalities are felt to be significant enough for flagging.

- **CAMERA**
- **IMAGE NUMBER**
- **POSSIBLE BAD QUALITY AND/OR MISSING MINOR FRAMES.**
This flag could be set by the processing system. No attempt will be made to distinguish between bad or missing minor frames in the spectral image versus the background.
- **MICROPHONICS.**
Flag set if microphonics are present. There will be no attempt to distinguish between microphonics in the spectral image versus the background.
- **NON-STANDARD IMAGE.**
This flag would be set for images taken in a non-standard manner. Examples of non-standard conditions are a camera grid turned off and medium-gain exposures. Other examples can be identified later. These images would be flagged at the time the image and data base are verified.
- **LWP BAD SCANS.**
This flag may be required for processing as the bad scan starts may be associated with large reseaux shifts. The effectiveness of the cross correlation algorithm could be improved for these images if the internal parameters were adjusted for this condition prior to processing.

- **LWR 4-MINUTE HEATER WARM-UPS.**
Image was acquired after 4-minute heater warm-up for ping avoidance.
- **READS AT OTHER THAN 20 KB.**
Image was read at unusual rate.
- **LWR 4.5 UVC VOLTAGE.**
Image was acquired at this UVC VOLTAGE
- **HISTORY REPLAY.**
Users need to be warned of history replays so that they can avoid the incomplete and possibly corrupt image label.
- **PARTIAL READS.**
The image was acquired with a partial read.
- **OTHER.** This flag would be set for other distorted or incomplete images e.g. images with bad reconstruction. This flag would be set at the time the image is visually inspected.

GO COMMENTS ENTITY

The GO COMMENTS entity contains any comments by Guest Observers sent to the IUE Project concerning the images acquired for their programs.

- **CAMERA**
- **IMAGE NUMBER**
- **COMMENTS**
Any additional comments a Guest Observer may have.

OFF SET REFERENCES ENTITY

Any off set references will be stored in the off-set reference entity.

- **CAMERA**
- **IMAGE NUMBER**
- **SPECNUM**
Spectrum number of the exposure within the image.

- OFR-X
Off set X reference.
- OFR-Y
Off set Y reference.

HOMOGENEOUS ENTITY

The homogeneous entity has not been fully defined but will include co-ordinates, catalog and object name, magnitude calculated from FES counts and other scientific information about the object as supplied from CDS.

PROGRAM ENTITY

The PROGRAM entity contains information concerning each observing program.

- PROGID
Program identification number.
- PI
Primary Investigator.
- TITLE
Title of the observing program.
- EPISODE
Episode for which the program was approved.
- SPONSOR
Sponsoring institution of the program.

JOURNAL ENTITY

The JOURNAL table contains references of articles using or referring to IUE data.

- ARTICLE NUMBER
Uniquely assigned identifying number.
- JOURNAL
Journal in which article appeared.

- **VOLUME**
Journal volume.
- **PAGE**
Page number.
- **YEAR**
Year article appeared.
- **AUTHOR NAME**
Primary authors name.
- **TITLE**
Article title.

CO-AUTHOR ENTITY

This table contains co-authors associated with each journal article.

- **ARTICLE NUMBER**
Uniquely assigned identifying number.
- **COAUTHOR NAME**
Name of co-author.

RELATED OBJECTS ENTITY

Objects specifically mentioned in journal references stored in the JOURNAL entity will be stored in this entity.

- **ARTICLE NUMBER**
Uniquely assigned identifying number.
- **OBJECT NAME**
Name of target in article.

RELATED IMAGES ENTITY

IUE images specifically mentioned in journal references stored in the JOURNAL entity will be stored in this entity.

- **ARTICLE NUMBER**
Uniquely assigned identifying number.
- **CAMERA**
Camera number of image referred to in the article.
- **IMAGE**
Image number of image referred to in the article.

9 Appendix 3 - Core Items

Table 1: Working DB Formats of Core Data Items (Image Processing)		
Keyword	Format	Possible Values
CAMERA	Int 1	1=LWP, 2=LWR, 3=SWP, 4=SWR, 8=FES1, 9=FES2
IMAGE	Int 5	11223, 01396, ...
DATE-OBS	Int 6	900212 (YYMMDD), ...
TIME-OBS	Int 6	121355 (HHMMSS), ...
MJD-OBS	Flt 10.4	(JD - 2400000.5)
DISPERSN	Char 1	H=High, L=Low, B=Both
APERTURE	Char 1	S=Small, L=Large, B=Both
EXP-MODE		
_Point/Extended	Char 1	P=Point, E=Extended
_Trailed	Int 1	0=no trail, 1=trail-x, 2=trail-y
_Multiple	Char 1	Y=At constant Y, A=Along aper. axis, N=No
_Segmented	Char 1	Y=Yes, N=No
ABNCODE		
_Minor frames	Char 1	Q=Bad quality, M=Missing, B=Both, N=None
_Microphonics	Char 1	Y=Yes, N=No
_Non-standard	Char 1	Y=Yes, N=No
_LWP bad scans	Char 1	Y=Yes, N=No
_LWR Heater warmups	Char 1	Y=Yes, N=No
_Read other 20KB	Char 1	Y=Yes, N=No
_LWR 4.5 UVC volt	Char 1	Y=Yes, N=No
_History replay	Char 1	Y=Yes, N=No
_Partial read	Char 1	Y=Yes, N=No
_Other	Char 1	Y=Yes, N=No
READMODE	Char 1	P=Partial, F=Full, ...
READGAIN	Char 1	H=High, L=Low
EXPOGAIN	Char 1	X=Maximum, D=Medium, N=Minimum
EXPTIME	Flt 8.2	281.22, ...
THDASTRT	Flt 7.4	10.5000, ...
THDAREAD	Flt 7.4	9.2300, ...
RA	Int 8	21355892 (HHMMSSss), ...
DEC	Int 8	+3432121 (\pm DDMMSSs), -7523654, ...
EPOCH	Flt 6.1	1950.0
STATION	Char 1	V=VILSPA, G=GSFC
SCHEME0	Char 5	CxxxxF, ...
UVC-VOLT	Int 5	4500, 4643, 5000, ...
ORBEOCH	Int 6	860108 (YYMMDD), ...
ORB-A	Flt 7.1	42161.5 (km), 42155.3, ...
ORBECEN	Flt 9.7	0.1890308, 0.1794419, ...
ORBINCLI	Flt 7.3	029.871 (degrees), 030.184, ...
ORBASCEN	Flt 7.3	147.475 (degrees), 141.374, ...
ORBPERIG	Flt 7.3	316.400 (degrees), 323.848, ...
ORBANOMA	Flt 7.3	008.366 (degrees), 000.067, ...

Table 2: Working DB Formats of Core Data Items (Science Analysis)		
Keyword	Format	Possible Values
LAPSTAT	Char 1	O=open, C=closed
TFLOOD	Char 1	Y=Yes, N=No
FES-MODE	Char 2	FU,FO,SU,SO,BO
FESCOUNT	Int 5	22321, ...
POSANGLE	Flt 5.1	261.9, ...
TARGET	Char 16	SATURN.RNG
TARG-RA	Int 7	0054327 (HHMMSSs), ...
TARG-DEC	Int 7	-000674 (\pm DDMMSS), ...
OBJECT	Char 28	ZZ SATURN RING
PGM-ID	Char 5	PHCAL, ...
IUE-CLAS	Int 2	05, 12, 99, ...
TRAIL-RT	Flt 10.5	0.10000, ...
TRAIL-NR	Int 3	001, 003, ...
Y-SHIFT	Flt 6.2	1.54, ...
DATABKG	Int 3	080, 040, ...
DATAcnt	Int 3	100, 098, ...
SCHEME	Char ?	???
FOCUS	Flt 7.2	-1.69, -3.61, ...
FPM	Flt 7.2	0.08, 0.34, ...
GSTAR-X	Int 4	012, -032, ..
GSTAR-Y	Int 4	-237, 700, ..
GSTAR-CN	Int 5	00200, 12432,...
GSTAR-MD	Char 2	FO,FU,SO,SU,NO
COMMENT	Char 60	
COMMENT	Char 60	