

Space Science Data Archive: Case Study

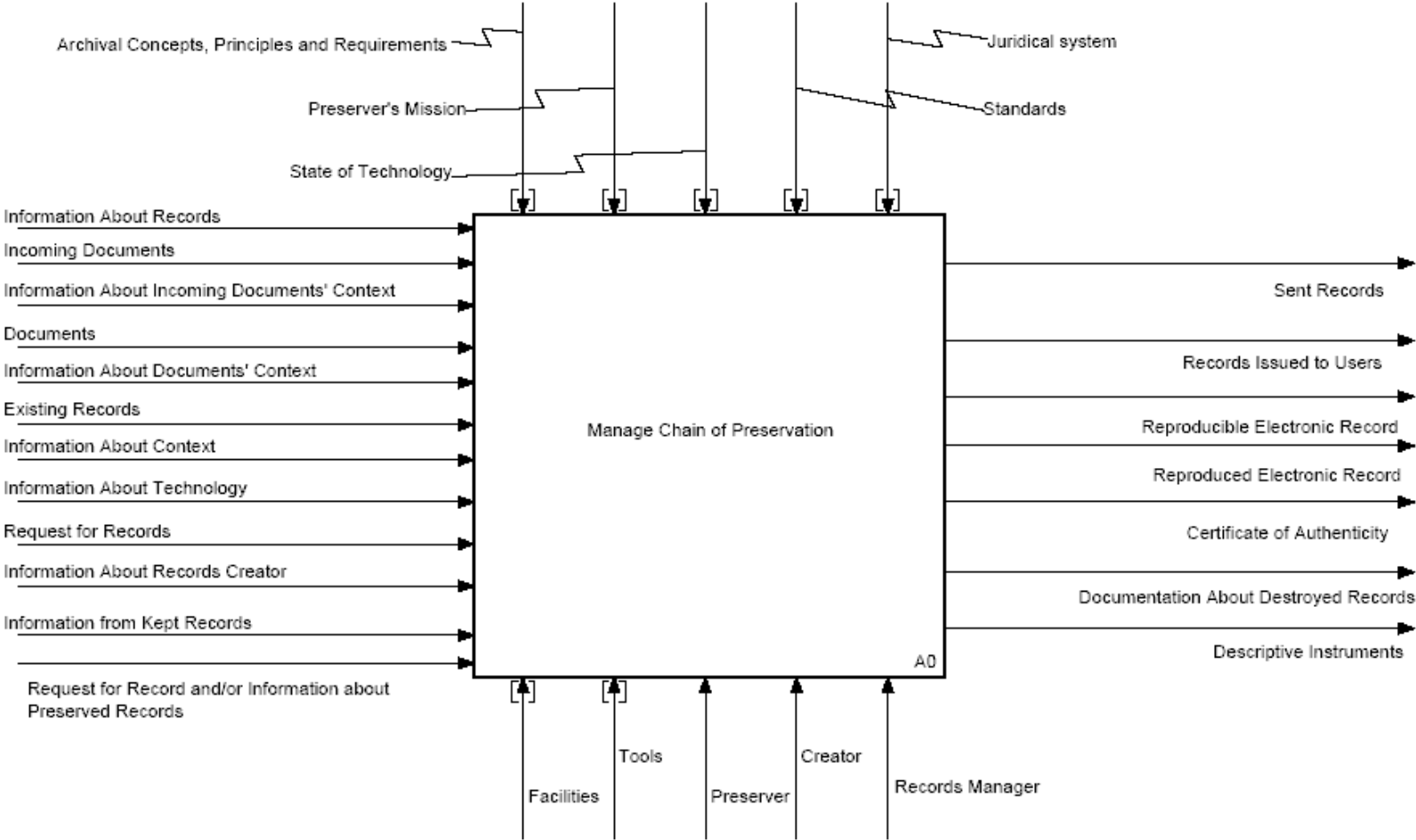
William Underwood
Georgia Tech Research Institute
Atlanta, Georgia USA

19th International CODATA Conference
Berlin, Germany, 7-10 November 2004

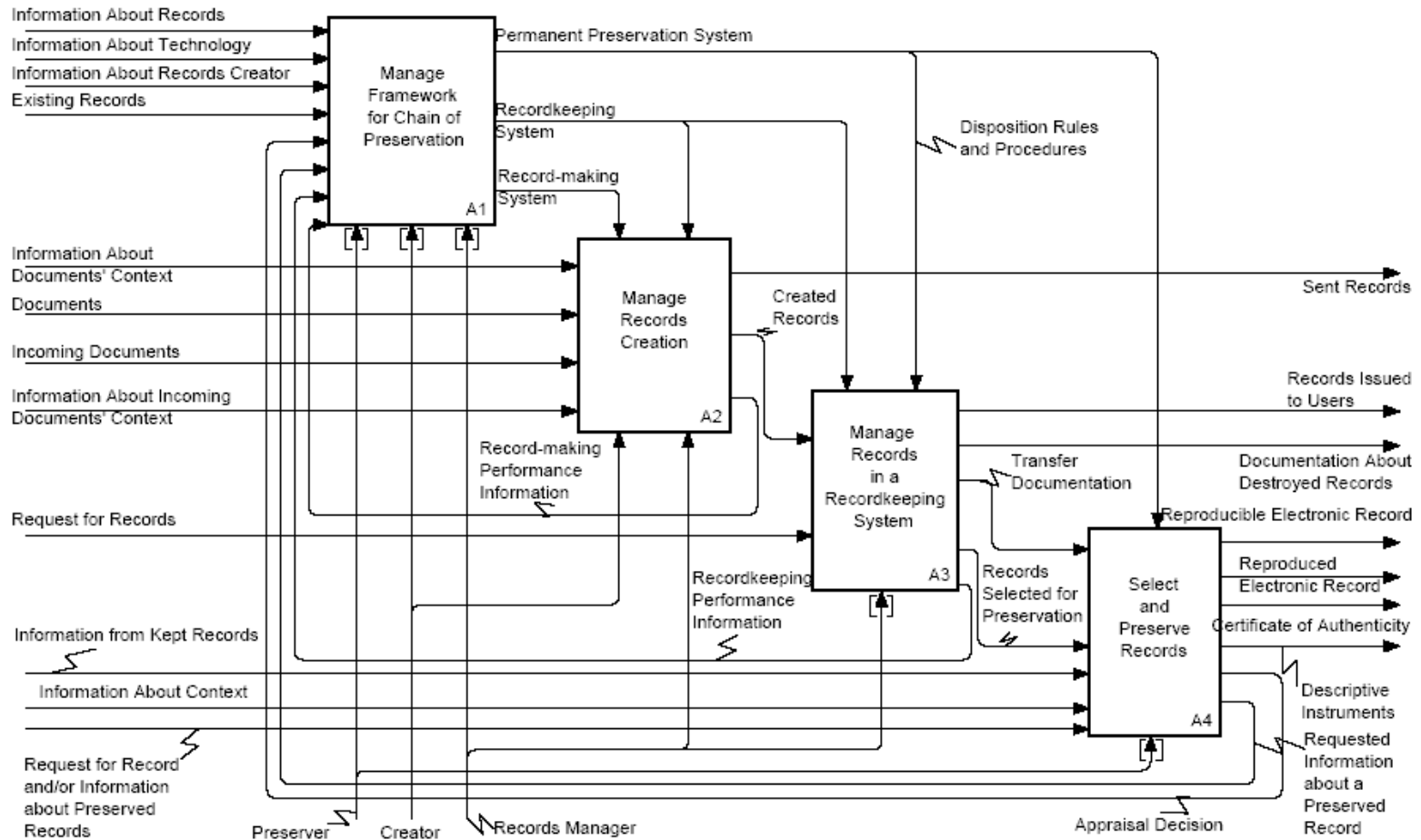
Overview

- Background: The InterPARES Chain of Preservation Model
- Validating the Model via Walkthrough
- The Planetary Data System (PDS)
- Reliability, Accuracy and Authenticity in the PDS
- Summary

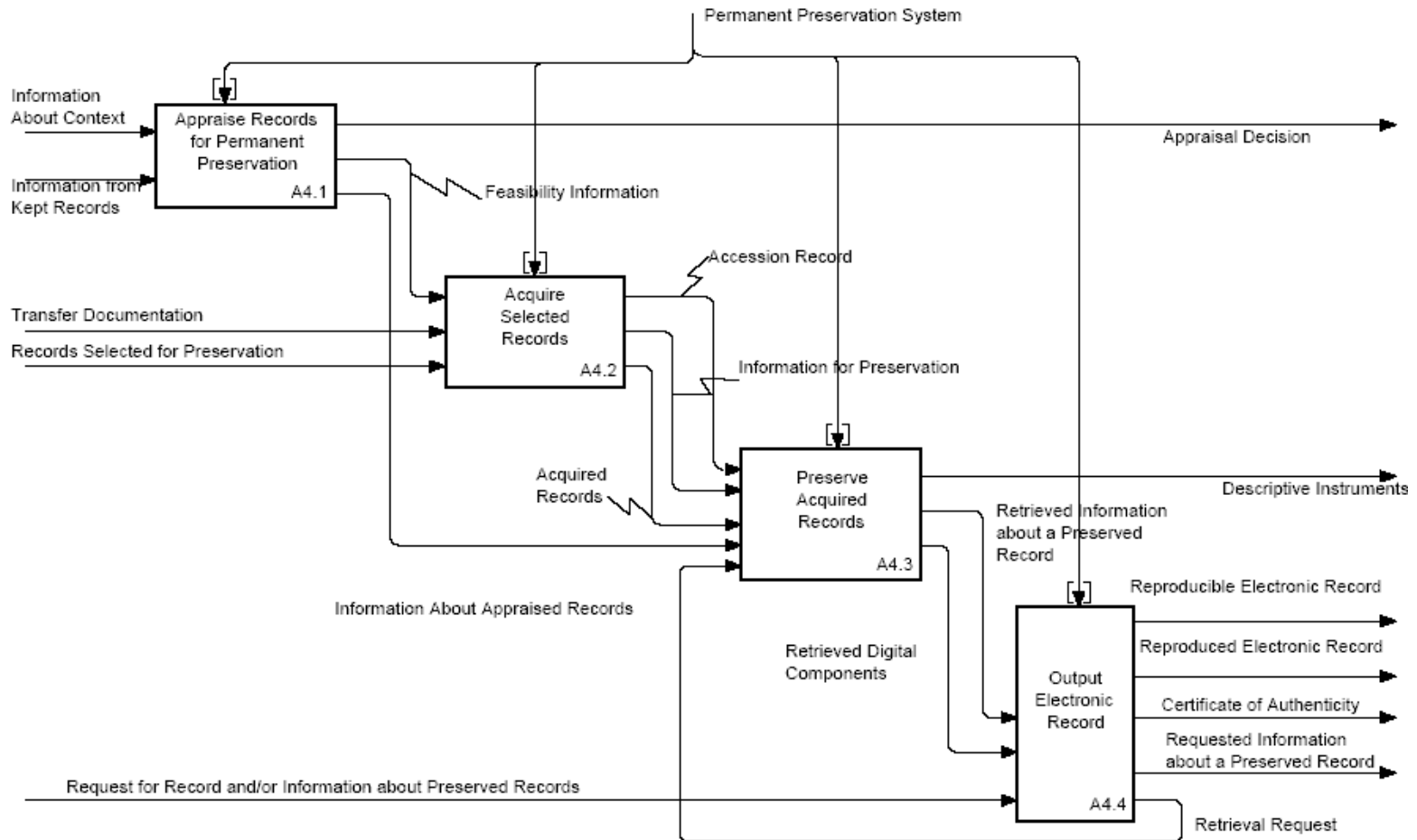
Chain of Preservation Model



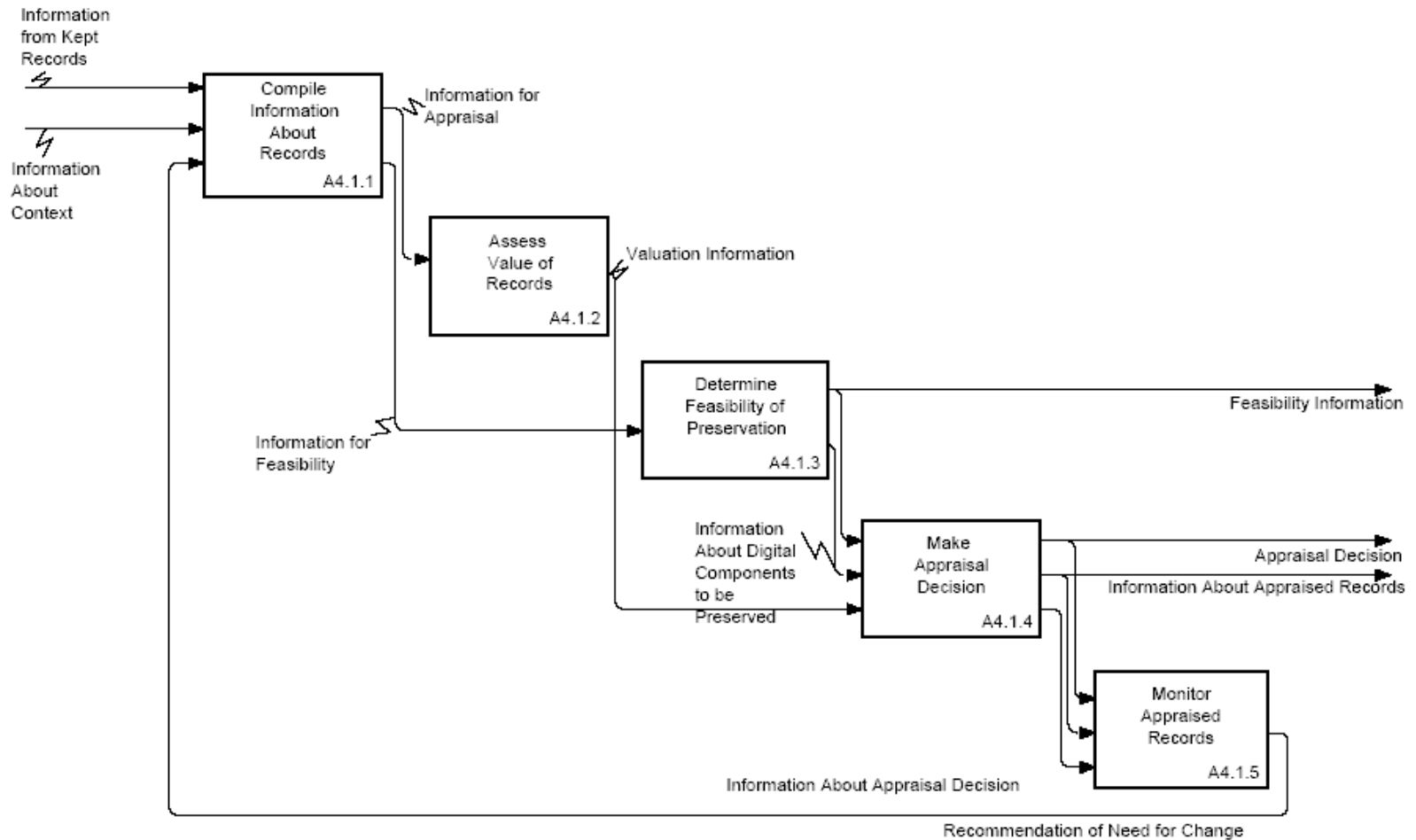
The Decomposition



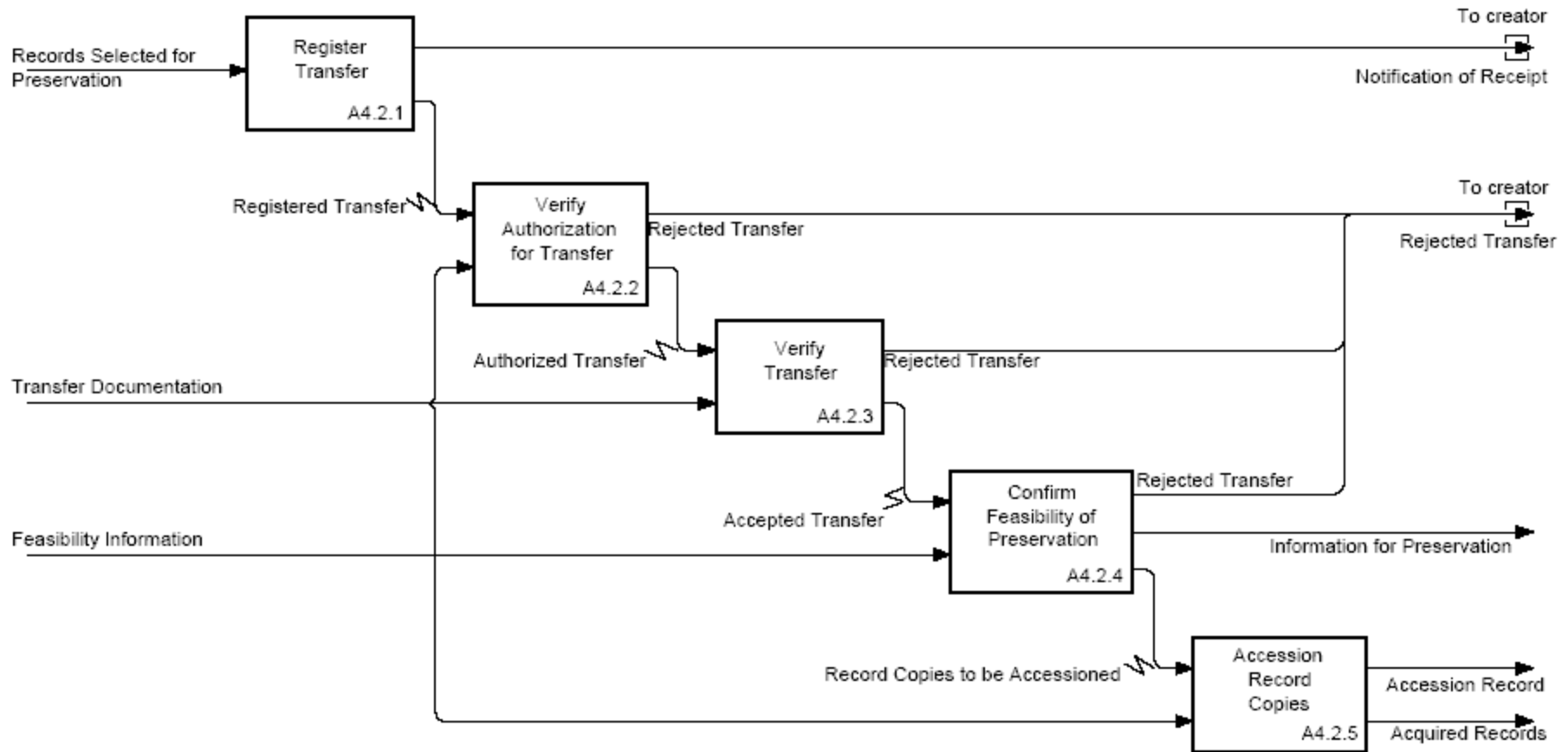
Select and Preserve Records



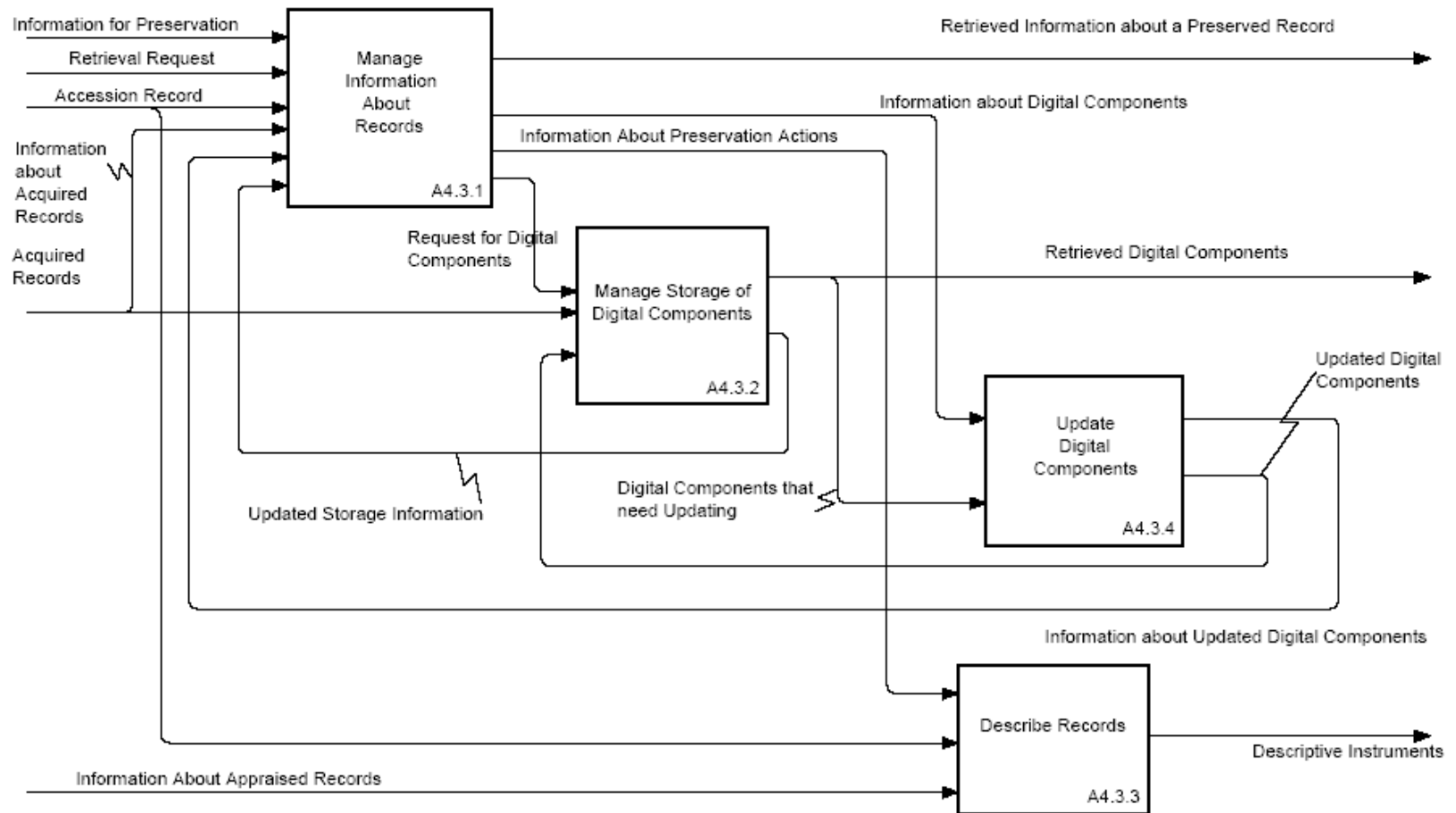
Appraise Records



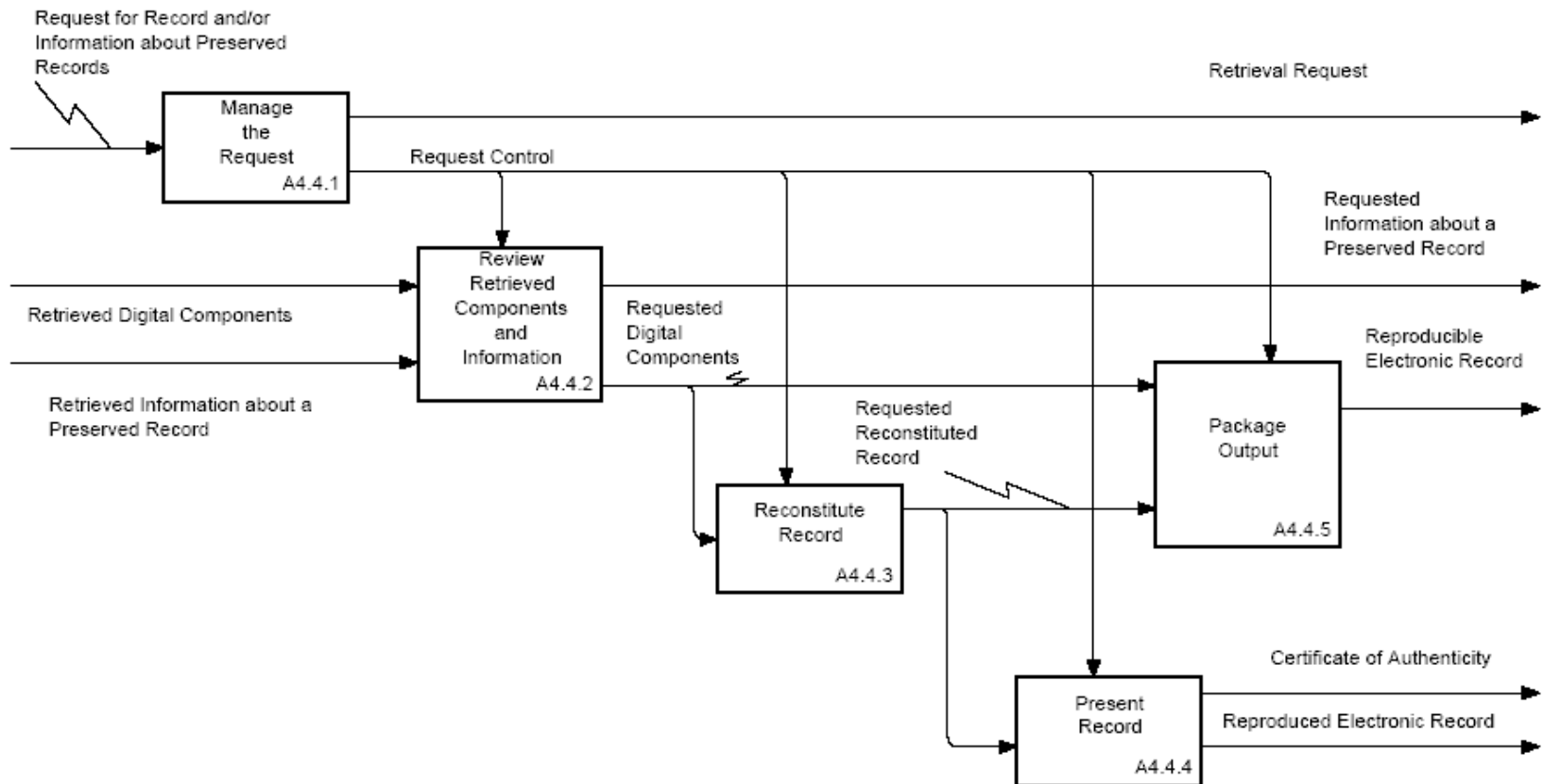
Acquire Selected Records



Preserve Acquired Records



Output Record



The Walkthrough

- Objectives
 - To demonstrate that the Model of Selection and Preservation works for an actual case of selection and preservation
 - To test and refine, and eventually validate, the model

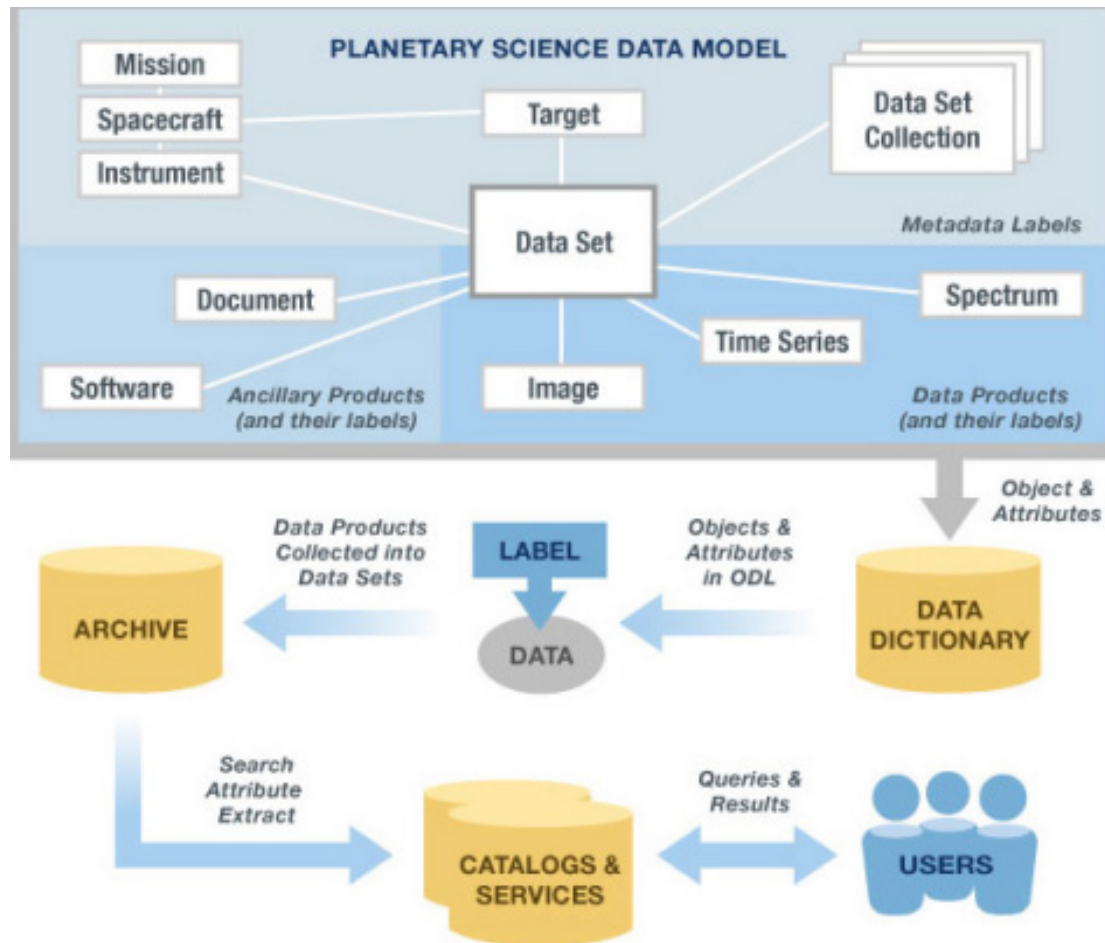
Walkthrough Team

- Presenter
- Reviewers
- Case Study Expert
- Secretary

Walkthrough Method

- Review the activity definitions and the input, output, and control definitions.
- Identify the data elements of the labels on the input and out arrows.
- Determine the transformations of inputs to outputs.
- Determine the values of the data elements that are related to the specific body of records.
- Record the results and any problems or issues that arise and any suggested solutions.

Planetary Data System (PDS): A Scientific Data Archive



Space Science Disciplines and Data Formats

Astrophysics	Astronomers	FITS
Planetary Physics	Solar System	PDS
Space Physics	Solar Wind, Charged Particles	CDF

PDS Peer Review

All data incorporated into the PDS archives must undergo a peer review. The purpose of the review is to determine that:

- The data is accurate, complete and reliable
- The data are suitable for archiving
- The PDS standards have been followed

Terms and Conditions of Transfer

- Data Preparation Workbook
- Project Data Management Plan
- Archive and Transfer Plan
- Software Interface Specification

PDS Data Types

- Array
- Cube
- Qube
- Spectrum
- Spreadsheet
- Table
- Image
- Text
- Time Series

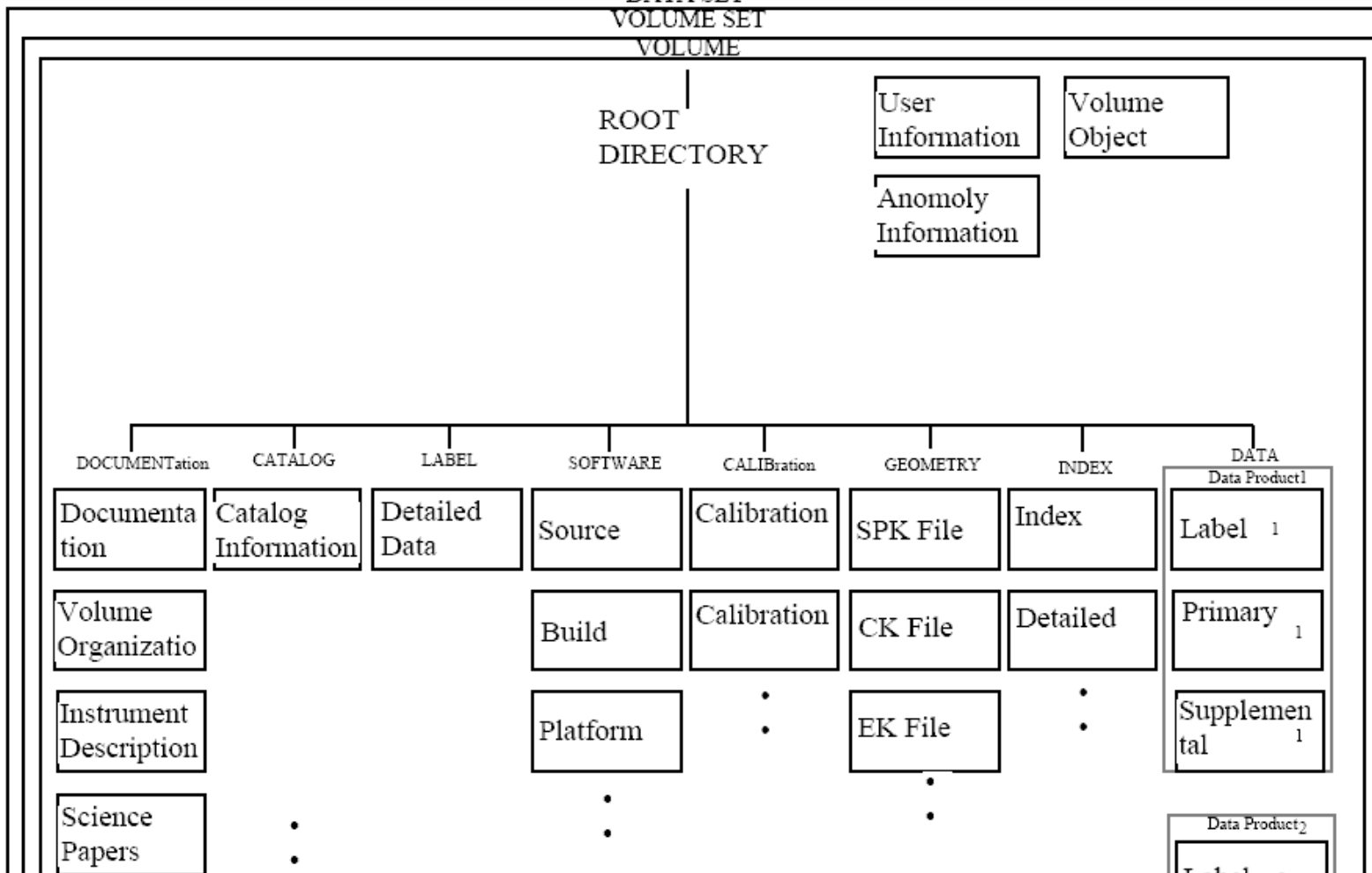
Sample ODL Definition of a File

```
/* File Format and Length */
    RECORD_TYPE          = FIXED_LENGTH
    RECORD_BYTES        = 800
    FILE_RECORDS        = 860
/* Pointer to First Record of Major Objects in File */
    ^IMAGE               = 40
    ^IMAGE_HISTOGRAM     = 840
    ^ANCILLARY_TABLE     = 842
/* Image Description */
    SPACECRAFT_NAME     = VOYAGER_2
    TARGET_NAME         = IO
    IMAGE_ID            = "0514J2-00"
    IMAGE_TIME          = 1979-07-08T05:19:11Z
    INSTRUMENT_NAME     = NARROW_ANGLE_CAMERA
    EXPOSURE_DURATION  = 1.9200<SECONDS>
    NOTE                = "Routine multispectral longitude
                        coverage, 1 of 7 frames"
/* Description of the Objects Contained in the File */
    OBJECT              = IMAGE
    LINES               = 800
    LINE_SAMPLES        = 800
    SAMPLE_TYPE         = UNSIGNED_INTEGER
    SAMPLE_BITS         = 8
    END_OBJECT          = IMAGE

    OBJECT              = IMAGE_HISTOGRAM
    ITEMS               = 25
    ITEM_TYPE           = INTEGER
    ITEM_BITS           = 32
    END_OBJECT          = IMAGE_HISTOGRAM

    OBJECT              = ANCILLARY_TABLE
    ^STRUCTURE          = "TABLE.FMT"
    END_OBJECT          = ANCILLARY_TABLE
END
```

DATA SET
VOLUME SET
VOLUME



The PDS is a Persistent Archive

- A persistent archive is an archive of digital objects that are persistent over time.
- The PDS achieves persistent data objects through encoding of data descriptions in the ODL and interpretation of these with routines in its Object Access Library which is written in C, thus achieving a high degree of software and hardware independence.

Summary

- Walkthroughs of the InterPARES Chain of Preservation model using case study data support refinement and validation of the model.
- Walkthroughs also provide concrete examples of the application of the model.
- Historically, the PDS is appears to be the first example of a Persistent Data Archive.

References

- Walkthrough of the InterPARES Model for Selecting and Preserving Electronic Records. (forthcoming) www.interpares.org
- Planetary Data System Data Preparation Workbook, Version 3.1, JPL D-7669, Part 1, Feb. 1995, Jet Propulsion Laboratory, Pasadena.
- Planetary Data System Standards Reference, Version 3.6, JPL D-7669, Part 2, Aug. 2003, Jet Propulsion Laboratory, Pasadena.