Standards Change Request 3-1073
===============================

Objective: Modify use of the ARCHIVE_STATUS keyword
     Date: 2007-01-17 (EN)
     Revised: 2007-01-19 (Simpson)
     Revised: 2007-02-27 (Sword)[b1]
     Revised: 2007-02-27 (Simpson)
     Revised: 2007-03-21 (Simpson, after Tech Session telecon)

Working Group
=============

     Steve Hughes (EN, lead)

     Data Tracking Working Group

Background
==========

ARCHIVE_STATUS was added to the DATA_SET_INFORMATION object (in the DATA_SET
object) and to the DATA_SET_RELEASE object several years ago. From the
beginning, there was confusion over both the function and use of the new
keyword. Some felt that it was primarily for the internal use of PDS -- to
track the progress of data sets through the archiving process. Others argued
that it was better suited to warn users about available, but only partially
archived, data. Still others suggested that it could do both, but that its
use needed to be under closer supervision.

The situation has been complicated by the fact that there are nine standard
values, few of which can be applied unambiguously. In addition, a modifier to
the standard values was authorized, meaning that the total number of possible
combinations is at least 18. Use of the modifier has been inconsistent; some
have insisted on using the modifier without any of the standard value roots.

Problem Statement
=================

Some data providers, charged with initializing the value of ARCHIVE_STATUS in
DATASET.CAT files, have been puzzled at their responsibilities (or lack of
same) as data sets moved through the review and ingestion steps and the value
of ARCHIVE_STATUS presumably needed revision. What is the process by which
values of ARCHIVE_STATUS are changed? Eventually, the PDS Management Council
directed that ARCHIVE_STATUS be removed from the DATA_SET_INFORMATION object
and that its future use be clarified.

Proposed Solution
=================

This Standards Change Request removes ARCHIVE_STATUS from the list of required
keywords in the DATA_SET_INFORMATION object, adds it to the list for the
DATA_SET_HOUSEKEEPING object, reduces the number of standard values, defines a
new keyword (DELIVERY_STATUS) to provide the previous function served by the
modifier, and assigns responsibility for maintaining current values of both
keywords.

By consensus at the 2007-03-21 Standards Telecon, this SCR was updated to
reflect final comments, then referred to a separate working group charged with
defining a PDS "data tracking" process. The DTWG will determine what the
future use of ARCHIVE_STATUS should be. Anne Raugh will submit a separate SCR

which will remove ARCHIVE_STATUS from DATA_SET_INFORMATION and note that MC's
"clarification" request falls conveniently within the charge to the DTWG.

Current Urgency
===============

In practical terms there is no urgency. Proposals to rescind the requirement
for ARCHIVE_STATUS were floated almost as soon as it was implemented more than
three years ago. If there is urgency, it is in responding to the Management
Council directive in a timely manner.

Recommendations
===============
The Working Group recommends the following actions:

  1. Amend the Planetary Science Data Dictionary (PSDD) to redefine the
      ARCHIVE_STATUS keyword and its standard values:

      -- replace IN QUEUE, PRE PEER REVIEW, IN PEER REVIEW, IN LIEN
         RESOLUTION and LOCALLY ARCHIVED with two standard values: "IN
         REVIEW" and "N/A".
      -- update descriptions for several standard values by substituting
         "curating node" for "CN" and "Central Node" and deleting references
         to NSSDC
      -- remove ACCUMULATING as a qualifier for standard values

  2. Amend the PSDD by adding the DELIVERY_STATUS keyword, which will
      qualify the state of a data set in terms of its current and future
      delivery schedule.

  3. Amend the Standards Reference and PSDD to remove ARCHIVE_STATUS as a
      required element of the DATA_SET_INFORMATION catalog object.

  4. Amend the PSDD to add the ARCHIVE_STATUS and DELIVERY_STATUS keywords
      as required elements of the DATA_SET_HOUSEKEEPING object.

  5. Amend the PSDD to add the DELIVERY_STATUS keyword as a required element
      of the DATA_SET_RELEASE object.

  6. Add the values of ARCHIVE_STATUS and DELIVERY_STATUS to display pages
      for individual data sets in the following:
      (a) the PDS subscription service:

      (b) The PDS Data Set View (we propose adding columns between Data Set
       and Instrument Host):

   Note: Although ARCHIVE_STATUS and DELIVERY_STATUS can be obtained from a
         DATA_SET_RELEASE object, their current values will be initialized and
         maintained by the curating node in the DATA_SET_HOUSEKEEPING
         object. Changes in DATA_SET_HOUSEKEEPING values do not necessarily
         trigger issuance of new DATA_SET_RELEASE objects.

===
PDS Standards Reference Changes
===

(1) Section B.1.6 - Example (DATA_SET)

(a) Remove "ARCHIVE_STATUS = ARCHIVED" from the example

            Change:
                    OBJECT = DATA_SET
                      DATA_SET_ID = "MGN-V-RDRS-5-GVDR-V1.0"

                      OBJECT = DATA_SET_INFORMATION
                        DATA_SET_NAME = "MGN V RDRS DERIVED GLOBAL VECTOR DATA
RECORD V1.0"

                        DATA_SET_COLLECTION_MEMBER_FLG = "N"
                        DATA_OBJECT_TYPE = TABLE
                        ARCHIVE_STATUS = ARCHIVED
                        START_TIME = 1990-08-01T00:00:00
                        STOP_TIME   = 1993-12-31T23:59:59
                        DATA_SET_RELEASE_DATE   = 1994-07-01
                        PRODUCER_FULL_NAME = "MICHAEL J. MAURER"
                        DETAILED_CATALOG_FLAG = "N"

            To:
                    OBJECT = DATA_SET
                      DATA_SET_ID = "MGN-V-RDRS-5-GVDR-V1.0"

                      OBJECT = DATA_SET_INFORMATION
                        DATA_SET_NAME = "MGN V RDRS DERIVED GLOBAL VECTOR DATA
RECORD V1.0"

                        DATA_SET_COLLECTION_MEMBER_FLG = "N"
                        DATA_OBJECT_TYPE = TABLE
                        START_TIME   = 1990-08-01T00:00:00
                        STOP_TIME = 1993-12-31T23:59:59
                        DATA_SET_RELEASE_DATE = 1994-07-01
                        PRODUCER_FULL_NAME   = "MICHAEL J. MAURER"
                        DETAILED_CATALOG_FLAG = "N"

(2) Section B.7.1 - Required Keywords (DATA_SET_INFORMATION)

    (a) Remove "ARCHIVE_STATUS" from the list of Required Keywords:

        Change:
        B.7.1 Required Keywords

        1. ABSTRACT_DESC
        2. ARCHIVE_STATUS
        3. CITATION_DESC
        4. CONFIDENCE_LEVEL_NOTE
        5. DATA_OBJECT_TYPE
        6. DATA_SET_COLLECTION_MEMBER_FLG
        7. DATA_SET_DESC
        8. DATA_SET_NAME
        9. DATA_SET_RELEASE_DATE
        10. DATA_SET_TERSE_DESC
        11. DETAILED_CATALOG_FLAG
        12. PRODUCER_FULL_NAME
        13. START_TIME
        14. STOP_TIME

        To:
        B.7.1 Required Keywords

        1. ABSTRACT_DESC
        2. CITATION_DESC
        3. CONFIDENCE_LEVEL_NOTE
        4. DATA_OBJECT_TYPE
        5. DATA_SET_COLLECTION_MEMBER_FLG

```
            6. DATA_SET_DESC
            7. DATA_SET_NAME
            8. DATA_SET_RELEASE_DATE
            9. DATA_SET_TERSE_DESC
           10. DETAILED_CATALOG_FLAG
           11. PRODUCER_FULL_NAME
           12. START_TIME
           13. STOP_TIME
```

(3) Appendix G - SAVED Data

    - No change: The following sentence is not affected by the proposed
             revisions:

     In rare cases data will be encountered in the PDS archive which
     are classified as having ARCHIVE_STATUS = SAVED.

```
=======================================================================
PDS Data Dictionary Changes
=======================================================================
```

  (1) Amend the DATA_SET_INFORMATION object definition by removing the
     ARCHIVE_STATUS keyword.

```
      OBJECT = SPECIFIC_OBJECT_DEFINITION
        NAME = DATA_SET_INFORMATION
        STATUS_TYPE = PENDING
        DESCRIPTION = "
           The DATASETINFO  (Data Set Information) catalog object is
           completed for the data set specified in the DATASET catalog
           object, in order to provide information specific to the
           given data set, such as the boundaries of the data set."
        REQUIRED_ELEMENT_SET = {
           ABSTRACT_DESC,
           CITATION_DESC,
           CONFIDENCE_LEVEL_NOTE,
           DATA_OBJECT_TYPE,
           DATA_SET_COLLECTION_MEMBER_FLG,
           DATA_SET_DESC,
           DATA_SET_NAME,
           DATA_SET_RELEASE_DATE,
           DATA_SET_TERSE_DESC,
           DETAILED_CATALOG_FLAG,
           PRODUCER_FULL_NAME,
           START_TIME,
           STOP_TIME}
        OBJECT_TYPE = SPECIFIC
      END_OBJECT = SPECIFIC_OBJECT_DEFINITION
      END
```

  (2) Amend the definition of the ARCHIVE_STATUS data element. Remove
     IN QUEUE, PRE PEER REVIEW, IN PEER REVIEW, IN LIEN RESOLUTION, and
     LOCALLY_ARCHIVED from the list of standard values, and remove
     ACCUMULATING as a qualifier for standard values. Add IN REVIEW to
     the list of standard values. Edit the description of ARCHIVED to
     delete requirement that data be
     delivered to NSSDC. Edit description of SUPERSEDED to discourage
     use except for historical purposes. Edit description of IN LIEN
     RESOLUTION to include the sentence "Use with caution." Other editorial
     improvements as indicated.

```
OBJECT = ELEMENT_DEFINITION
  NAME = ARCHIVE_STATUS
  STATUS_TYPE = PENDING
  GENERAL_DATA_TYPE = CHARACTER
  UNIT_ID = NONE
  STANDARD_VALUE_TYPE = DYNAMIC
  MAXIMUM_LENGTH = 30
  DESCRIPTION = "
The archive_status element provides the status of a data set that
has been submitted for inclusion in the PDS archive.

The archive_status_note element is available to expand on the
archive_status value in more detail.

STANDARD VALUES

    IN DEVELOPMENT - The data set has been defined but no products have
    been received at the curating node.

IN REVIEW - The data set has been partially or wholly received at the
curating node and formal review and lien correction have either not been
initiated or are incomplete.  This data set should be used with caution.

ARCHIVED - Passed peer review with all liens resolved. Available
through the Data Set Catalog.

SUPERSEDED - Superseded by a new version of the data set. The
data set should not to be used except for historical reasons.

SAVED - Received by the PDS; no evaluation is planned and data will
not be formally archived.  This data set should be used with caution.

    WITHDRAWN - The data set has been removed from the archive collection.
    In some cases no data may have been delivered.  In others, the data
    were found to be unusable after the archiving process began (or was
    completed); there were no options for creating a replacement data
    set, so the data set could not be SUPERSEDED.

    Note: If a data set is being archived in segments, select the lower
    of the two levels, IN REVIEW and ARCHIVED.  IN REVIEW is assumed
    to be the lower level and ARCHIVED the higher.  For example, if any
    portion of a data set has not been declared ARCHIVED, assign the value
    of IN REVIEW to ARCHIVE_STATUS even though other portions may have
    completed the archive process.

    STANDARD_VALUE_SET = {
        "ARCHIVED",
        "IN DEVELOPMENT",
        "IN REVIEW",
        "SAVED",
        "SUPERSEDED",
        "WITHDRAWN"}
END_OBJECT = ELEMENT_DEFINITION
END
```

(3) Create a new DELIVERY_STATUS data element, which qualifies the state of
    a data set with regard to its current and future delivery schedule.

```
    OBJECT = ELEMENT_DEFINITION
      NAME = DELIVERY_STATUS
```

```
        STATUS_TYPE = PENDING
        GENERAL_DATA_TYPE = IDENTIFIER
        UNIT_ID = NONE
        STANDARD_VALUE_TYPE = STATIC
        MINIMUM_LENGTH = 1
        MAXIMUM_LENGTH = 25
        DESCRIPTION = "
   The DELIVERY_STATUS element describes the state of the
   intended delivery or deliveries of a data set that will be or
   has been submitted for inclusion in the PDS archive.  Changes in the
   value are usually initiated by the data provider.

   STANDARD VALUES

   SINGLE-PENDING    - the data set will be submitted as a single
                       delivery, or the delivery is in progress.

   SINGLE-DELIVERED - the data set has been submitted as a single delivery
                       and the delivery is complete.

   MULTIPLE-PENDING - the data set will be submitted as multiple
                       deliveries.

   MULTIPLE-ACCUMULATING - the data set is being submitted as multiple
                           deliveries; at least one delivery has been
                           started, and the final delivery has not been
                           completed.

   MULTIPLE-DELIVERED - the data set was submitted as multiple
                        deliveries.  All deliveries are complete."

        STANDARD_VALUE_SET = {
           "SINGLE-PENDING",
           "SINGLE-DELIVERED",
           "MULTIPLE-PENDING",
           "MULTIPLE-ACCUMULATING",
           "MULTIPLE-DELIVERED"}
      END_OBJECT = ELEMENT_DEFINITION
      END


(4) Amend the DATA_SET_HOUSEKEEPING object definition such that the
    ARCHIVE_STATUS and DELIVERY_STATUS keywords are mandatory
    elements.  Expand the description to mention other information
    beyond resources.

    OBJECT = GENERIC_OBJECT_DEFINITION
      NAME = DATA_SET_HOUSEKEEPING
      STATUS_TYPE = PENDING
      DESCRIPTION = "
         The DATA_SET_HOUSEKEEPING object contains information about
         the delivery, archiving, and curational status of a single
         data set and its associated resources (e.g. web pages, data set
         browsers, etc). The resources are used to select and download data
         products or other files associated with the data set."
      REQUIRED_ELEMENT_SET = {
         ARCHIVE_STATUS,
         DELIVERY_STATUS,
         CURATING_NODE_ID,
         DATA_SET_ID}
      REQUIRED_OBJECT_SET = {
         RESOURCE_INFORMATION}
      OBJECT_TYPE = GENERIC
      END_OBJECT = GENERIC_OBJECT_DEFINITION
```

```
        END

   (5) Amend the DATA_SET_RELEASE object such that the
       DELIVERY_STATUS keyword is a mandatory element.

       OBJECT = GENERIC_OBJECT_DEFINITION
         NAME = DATA_SET_RELEASE
         STATUS_TYPE = PENDING
         DESCRIPTION = "
           The DATA_SET_RELEASE object provides information on the
           release of a data set or portion of a data set being made
           available for online access."
         REQUIRED_ELEMENT_SET = {
           ARCHIVE_STATUS,
           DATA_PROVIDER_NAME,
           DATA_SET_ID,
           DELIVERY_STATUS,
           DESCRIPTION,
           DISTRIBUTION_TYPE,
           PRODUCT_TYPE,
           RELEASE_DATE,
           RELEASE_ID,
           RELEASE_MEDIUM,
           RELEASE_PARAMETER_TEXT}
         OBJECT_TYPE = GENERIC
         END_OBJECT = GENERIC_OBJECT_DESCRIPTION
```

Impact Assessment
=================

In addition to the Standards Reference and PSDD changes above, the following
tasks have been identified as being required to implement SCR 3-1073:

1. PDS Tool Suite - no impact

2. Ingestion tools - update the software that ingests DATA_SET_HOUSEKEEPING
                     objects to include the ARCHIVE_STATUS and
                     DELIVERY_STATUS keywords
                   - update the software that ingests DATA_SET_RELEASE
                     objects to include the DELIVERY_STATUS keyword.

3. "Release Manager" web page - update the software that ingests the
                     DATA_SET_RELEASE objects to include the new
                     DELIVERY_STATUS keyword and the amended
                     ARCHIVE_STATUS standard values.

4. "Resource Manager" web page - update the software that ingests the
                     DATA_SET_HOUSEKEEPING objects to include the new
                     DELIVERY_STATUS keyword and the amended
                     ARCHIVE_STATUS standard values.

5. "PDS Subscription Service" - add ARCHIVE_STATUS and DELIVERY_STATUS
                                to the display page

6. "PDS Data Set View" - add ARCHIVE_STATUS and DELIVERY_STATUS
                         to the display page

7. EN documentation - update the (internal) EN Ingestion manual to reflect
                     the change in the DATA_SET_HOUSEKEEPING and
                     DATA_SET_RELEASE objects.

       The recommendations above do not impact the PDS validation software
       as the modifications to the PSDD drive the creation of pdsdd.full

which, in turn, drives the validation software.

Adding the ARCHIVE_STATUS and DELIVERY_STATUS keywords to the
DATA_SET_HOUSEKEEPING object will affect the ingestion software,
but the change is minimal and will not require much effort.

Adding ARCHIVE_STATUS and DELIVERY_STATUS to the Subscription Service
and Data Set View pages will impact the software used to generate
those pages, but the change is minimal and will not require much
effort.

The changes to the DATA_SET_INFORMATION, DATA_SET_RELEASE and
DATA_SET_HOUSEKEEPING objects will impact the catalog ingestion
software, but the updates are minimal and will not require much
effort.

This proposal presents minimal disruption to the current operational
system, accommodates the design of the PDS, is tightly constrained,
and presents minimal impact on tool development.

8. Catalog update - Updating the value of ARCHIVE_STATUS and
   establishing initial values for DELIVERY_STATUS for existing data sets
   presents an unknown impact on resources. At the time of this writing,
   PDS holdings include data sets in the following categories:
   ```
   ARCHIVED:           746
   IN PEER REVIEW:      17
   IN QUEUE      :     169
   PRE PEER REVIEW:    168
   IN LIEN RESOLUTION:  12
   LOCALLY ARCHIVED:    98
   SUPERSEDED:          79
   SAVED                 6
   ```
   Those IN PEER REVIEW, IN QUEUE, PRE PEER REVIEW and IN LIEN RESOLUTION
   could almost be assumed to fit the IN REVIEW category and thus easily
   updated. Those in LOCALLY ARCHIVED should be looked at to determine if
   they belong in the ARCHIVED or IN REVIEW category. Those ARCHIVED or
   SUPERCEDED should not in general need to change.

Deleted: [b1]Revised:
2007-01-31(Sword)