

Draft Recommendation for
Space Data System Practices

|  |
| --- |
| Information CURATION Lifecycle framework |

PROPOSED Draft Recommended Practice

CCSDS 653.0-W-0.4

WHITE Book

* August 2015

AUTHORITY

|  |
| --- |
|  |
|  | Issue: | White Book, Issue 0.4-A |  |
|  | Date: | August 2015 |  |
|  | Location: | Not Applicable |  |
|  |

**(WHEN THIS RECOMMENDED PRACTICE IS FINALIZED, IT WILL CONTAIN THE FOLLOWING STATEMENT OF AUTHORITY:)**

This document has been approved for publication by the Management Council of the Consultative Committee for Space Data Systems (CCSDS) and represents the consensus technical agreement of the participating CCSDS Member Agencies. The procedure for review and authorization of CCSDS documents is detailed in *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4), and the record of Agency participation in the authorization of this document can be obtained from the CCSDS Secretariat at the address below.

This document is published and maintained by:

CCSDS Secretariat

Space Communications and Navigation Office, 7L70

Space Operations Mission Directorate

NASA Headquarters

Washington, DC 20546-0001, USA

FOREWORD

[Foreword text specific to this document goes here. The text below is boilerplate.]

Through the process of normal evolution, it is expected that expansion, deletion, or modification of this document may occur. This Recommended Practice is therefore subject to CCSDS document management and change control procedures, which are defined in *Organization and Processes for the Consultative Committee for Space Data Systems* (CCSDS A02.1-Y-4). Current versions of CCSDS documents are maintained at the CCSDS Web site:

http://www.ccsds.org/

Questions relating to the contents or status of this document should be addressed to the CCSDS Secretariat at the address indicated on page i.

At time of publication, the active Member and Observer Agencies of the CCSDS were:

Member Agencies

* Agenzia Spaziale Italiana (ASI)/Italy.
* Canadian Space Agency (CSA)/Canada.
* Centre National d’Etudes Spatiales (CNES)/France.
* China National Space Administration (CNSA)/People’s Republic of China.
* Deutsches Zentrum für Luft- und Raumfahrt (DLR)/Germany.
* European Space Agency (ESA)/Europe.
* Federal Space Agency (FSA)/Russian Federation.
* Instituto Nacional de Pesquisas Espaciais (INPE)/Brazil.
* Japan Aerospace Exploration Agency (JAXA)/Japan.
* National Aeronautics and Space Administration (NASA)/USA.
* UK Space Agency/United Kingdom.

Observer Agencies

* Austrian Space Agency (ASA)/Austria.
* Belgian Federal Science Policy Office (BFSPO)/Belgium.
* Central Research Institute of Machine Building (TsNIIMash)/Russian Federation.
* China Satellite Launch and Tracking Control General, Beijing Institute of Tracking and Telecommunications Technology (CLTC/BITTT)/China.
* Chinese Academy of Sciences (CAS)/China.
* Chinese Academy of Space Technology (CAST)/China.
* Commonwealth Scientific and Industrial Research Organization (CSIRO)/Australia.
* Danish National Space Center (DNSC)/Denmark.
* Departamento de Ciência e Tecnologia Aeroespacial (DCTA)/Brazil.
* European Organization for the Exploitation of Meteorological Satellites (EUMETSAT)/Europe.
* European Telecommunications Satellite Organization (EUTELSAT)/Europe.
* Geo-Informatics and Space Technology Development Agency (GISTDA)/Thailand.
* Hellenic National Space Committee (HNSC)/Greece.
* Indian Space Research Organization (ISRO)/India.
* Institute of Space Research (IKI)/Russian Federation.
* KFKI Research Institute for Particle & Nuclear Physics (KFKI)/Hungary.
* Korea Aerospace Research Institute (KARI)/Korea.
* Ministry of Communications (MOC)/Israel.
* National Institute of Information and Communications Technology (NICT)/Japan.
* National Oceanic and Atmospheric Administration (NOAA)/USA.
* National Space Agency of the Republic of Kazakhstan (NSARK)/Kazakhstan.
* National Space Organization (NSPO)/Chinese Taipei.
* Naval Center for Space Technology (NCST)/USA.
* Scientific and Technological Research Council of Turkey (TUBITAK)/Turkey.
* South African National Space Agency (SANSA)/Republic of South Africa.
* Space and Upper Atmosphere Research Commission (SUPARCO)/Pakistan.
* Swedish Space Corporation (SSC)/Sweden.
* Swiss Space Office (SSO)/Switzerland.
* United States Geological Survey (USGS)/USA.

PREFACE

This document is a draft CCSDS Recommended Practice. Its ‘White Book’ status indicates that its contents are not stable, and several iterations resulting in substantial technical changes are likely to occur before it is considered to be sufficiently mature to be released for review by the CCSDS Agencies.

Implementers are cautioned **not** to fabricate any final equipment in accordance with this document’s technical content.

DOCUMENT CONTROL

|  |  |  |  |
| --- | --- | --- | --- |
| **Document** | **Title and Issue** | **Date** | **Status** |
| CCSDS 000.0-W-0.1 | Information Curation Process, Proposed Draft Recommended Practice, Issue 0.1 | April 2014 | Original proposed draft |
| CCSDS 000.0-W-0.1a | Information Curation Process, Proposed Draft Recommended Practice, Issue 0.1a | June 2014 | Added Scope and Purpose Text, Import Abbreviations and Terminology |
| CCSDS 000.0-W-0.2 | Information Curation Process, Proposed Draft Recommended Practice, Issue 0.2 | September 2014 | Reworked Abbreviations from other standards. Still need to incorporate Purpose and Scope and Terminology from other documents. |
| CCSDS 653.0-W-0.3 | Information Lifecycle Framework, Proposed Draft Recommended Practice, Issue 0.3 | June 2015 | Renamed document, Entire document reworked to include only material from the agreed project description document. |
| CCSDS 653.0-W-0.4 | Information Lifecycle Framework, Proposed Draft Recommended Practice, Issue 0.4 | June 2015 | Expanded descriptions of Lifecycle stages. Lifecycle activities section added. |
| CCSDS 653.0-W-0.4-A | Information Curation Lifecycle Framework, Proposed Draft Recommended Practice, Issue 0.4-A | August 2015 | Current draft. Re-focused onto curation aspects.  |
|  |  |  |  |
|  |  |  |  |

CONTENTS

Section Page

[DOCUMENT CONTROL v](#_Toc427527687)

[CONTENTS vi](#_Toc427527688)

[1 Introduction 1-1](#_Toc427527689)

[1.1 purpose and scope 1-1](#_Toc427527690)

[1.2 applicability 1-3](#_Toc427527691)

[1.3 rationale 1-3](#_Toc427527692)

[1.4 conformance 1-4](#_Toc427527693)

[1.5 document structure 1-4](#_Toc427527694)

[1.6 definitions 1-4](#_Toc427527695)

[1.6.1 terminology 1-4](#_Toc427527696)

[1.7 NOMENCLATURE 1-5](#_Toc427527697)

[1.7.1 NORMATIVE TEXT 1-5](#_Toc427527698)

[1.7.2 INFORMATIVE TEXT 1-5](#_Toc427527699)

[1.8 References 1-6](#_Toc427527700)

[2 Overview and COntext 2-1](#_Toc427527701)

[3 THE LIFECYCLE FRAMEWORK: THE MAIN STAGES and Information to be gathered 3-1](#_Toc427527702)

[3.1 Information Lifecycle stages 3-1](#_Toc427527703)

[3.2 Information topics 3-2](#_Toc427527704)

[3.3 Staged curation information 3-2](#_Toc427527705)

[3.3.1 Formulation Stage 3-1](#_Toc427527706)

[3.3.2 Implementation Stage 3-1](#_Toc427527707)

[3.3.3 Operation Stage 3-2](#_Toc427527708)

[3.3.4 Exploitation Stage 3-2](#_Toc427527709)

# Introduction

[Insert introductory subsections such as PURPOSE, SCOPE, APPLICABILITY, RATIONALE, etc. See CCSDS A20.0-Y-4, *CCSDS Publications Manual* (Yellow Book, Issue 4, April 2014) for the contents of section 1.]

## purpose and scope

The purpose of this Recommended Practice is to define a framework for gathering information throughout the information lifecycle, from the proposal to the long-tern re-use of the resulting information, focusing on the activities needed at each stage which will help to ensure that the data can be optimally exploited over the long term. It can form the basis on which Data Management Plans can be constructed. It should be applicable to both existing as well as data to be created in future. It should be of use to funders, researchers, archive managers and end-users by helping to reduce the effort and increase the efficacy of preservation and exploitation of data.

The Recommendation does not cover all aspects of the lifecycle and aspects of the activities it does specify do not have to be carried out strictly sequentially, and indeed some may be revisited and improved at several of the stages.

It will describe stages of the information curation lifecycle and within each stage, this recommendation identifies the information which should be collected, created or improved in order to be able to preserve and utilize information objects for the long-term. Within the framework, standards, best practices and software tools that could be used are identified.

This framework takes the view that curation and preservation are not separate activities to be considered at the end of an information production project, but as a set of actions that must be conducted throughout the information lifecycle.

Other aspects, such as costing, risk management, metadata management, data formats, policies and workflow, value-adding and service architectures, which are clearly of interest in preservation and curation, are not addressed except at a high-level, to provide context. It is expected that full treatment of these issues will require additional, more focused, standards.

While this process originates in the space community, it is being designed in a generic way and should be applicable to any science domain and to the wider library and archival communities.

This Recommended Practice accomplishes the following:

* identifies the main stages in the information curation lifecycle;
* defines the information to be collected at each of these stages,;
* forms a general methodological framework, which should be applicable and usable in any information stewardship, curation or preservation context (this general framework should provide sufficient flexibility to be applied to individual user’s situations);
* forms a basis for the identification and/or development of additional standards and implementation guides including those that address particular concerns in more detail;
* forms a basis for identification and/or development of a set of software tools that will assist the development, operation and checking of the different stages of the lifecycle.

This Recommended Standard fits into the context defined by:

* The *Reference Model for an Open Archival Information System (OAIS)* Recommended Practice (see reference [1]).
* The *Producer-Archive Interface Methodology Abstract Standard* (PAIMAS) Recommended Practice (see reference [2]).
* The *Producer-Archive Interface Specification* Recommended Standard (see reference [3])
* The *Auditing and Certification of Trustworthy Digital Repositories* Recommended Practice (see reference [4]).
* The *Requirements for Bodies Providing Auditing and Certification of Candidate Trustworthy Digital Repositories* Recommended Practice (see reference [5]).

The OAIS Best Practice Standard (see reference [1]) is one of the most widely recognized and applied archival standards available today. An OAIS is an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community. It meets a set of such responsibilities as defined in this document. The model provides a framework for the understanding and increased awareness of archival concepts needed for long-term digital information preservation and access, and for describing and comparing architectures and operations of existing and future archives. It also guides the identification and production of OAIS related standards.

The PAIMAS Best Practice Standard (see reference [2]) defines a methodology for transferring data from an Information Producer to an Archives based on the four following phases: Preliminary, Formal Definition, Transfer, Validation. Required activities during each phase are identified.

The PAIS Recommended Standard (see reference [3]) provides the abstract syntax and an XML implementation of descriptions of data to be sent to an archive. These descriptions are negotiated agreements between the data Producer and the Archive and facilitate production of agreed data by the Producer and validation of received data by the Archive. The Recommended Standard includes an abstract syntax and one possible concrete implementation for the packages.

The Auditing and Certification Recommended Practice (see reference [4]) provides metrics for use in assessing the trustworthiness of digital repositories and the Guidelines for Bodies Recommended Practice (see reference [5]) provides procedures to be followed when conducting audits of digital repositories using those metrics.

## applicability

The considerations/processes defined in this document applies to any activities producing data which is (or may be) re-used and preserved for significant periods. It is applicable to those organizations and individuals who create information that may need Long-Term Preservation and to organizations making information available for the Long Term.

## rationale

Data that is collected or created needs to have additional information associated with it if it is to be independently understandable, usable and trusted as being authentic. That additional information changes over time, as hardware, software, the general environment and users’ tacit knowledge changes. OAIS uses the terms Representation Information and Preservation Description Information for this associated information, and must be accumulated over the lifecycle of the data. For example Provenance Information will accumulate over time, recording the things which have happened to the data.

Funders are increasingly asking that Data Management Plans accompany any request for project funding, however these tend not to evolve with the project and are difficult to monitor.

Many data lifecycles have been proposed. However they do not focus on the activities needed at each stage which will help to ensure that the data can be optimally exploited over the long term.

There are a small number of generally applicable stages in the information lifecycle where, typically, the responsibility is handed on from one individual or team to another. Each of those teams has specific knowledge about the information which subsequent teams may not possess. Therefore there is a need to specify the information to be captured at each of those stages. Improvements may be needed to, for example, the Representation Information, which was recorded in an earlier stage; this may arise when the better understood or reformatted or re-processed in later stages.

Therefore there is a need for a curation lifecycle framework to provide guidance as to what additional information should be captured or improved through the various stages of the lifecycle.

This document should enable:

* the Producer (including for example scientists who create the data) to capture and record the relevant information in a timely manner;
* the Archive to be assured that it will receive adequate information to enable it to perform preservation activities and support exploitation of the information
* the user to re-use unfamiliar information more easily
* the funder/sponsor to be assured that the resources that they contribute to the creation of the information will have suitable pay-back
* etc

Digital Content requires active management throughout its entire period of use/data lifecycle.

## conformance

TBD text

## document structure

TBD text

## definitions

TBD text

### terminology

We expect to need definitions for:

Note: In common usage there is confusion/overlap between many of these terms. To avoid confusion and to allow precision, we will need specific definitions that will apply within the recommendation. Hopefully these definitions will resonate with others and will gain wider usage. However we expect that existing domains will continue with their own usage, but we expect that they will be easily able to map their usages to the usage within the recommendation.

SEE TABLE 3-1

We need definitions compatible with diagram in following section:

**Curation:** (wider concept that includes Preservation and Consolidation)

**Metadata:** text TBD

**Data Record:** text TBD

**Preserved Data Set Content (PDSC):** text TBD

**Curation Project :** an activity planned and designed to achieve a particular aim ranging from the creation of new data to the preservation of existing data with a particular preservation aim.

**Data Management Plan** (Wikipedia): A data management plan or DMP is a formal document that outlines how you will handle your data both during your research, and after the project is completed. The goal of a data management plan is to consider the many aspects of data management, metadata generation, data preservation, and analysis before the project begins. This ensures that data are well-managed in the present, and prepared for preservation in the future.

**Active Data Management Plan :** An evolving record of the metadata needed to be confident of the authenticity and re-usability of the information which has been created.

## NOMENCLATURE

### NORMATIVE TEXT

The following conventions apply for the normative specifications in this Recommended Standard:

1. the words ‘shall’ and ‘must’ imply a binding and verifiable specification;
2. the word ‘should’ implies an optional, but desirable, specification;
3. the word ‘may’ implies an optional specification;
4. the words ‘is’, ‘are’, and ‘will’ imply statements of fact.

NOTE – These conventions do not imply constraints on diction in text that is clearly informative in nature.

### INFORMATIVE TEXT

In the normative sections of this document (sections **Error! Reference source not found.**-**Error! Reference source not found.**), informative text is set off from the normative specifications either in notes or under one of the following subsection headings:

* Overview;
* Background;
* Rationale;
* Discussion.

## References

The following publications contain provisions which, through reference in this text, constitute provisions of this document. At the time of publication, the editions indicated were valid. All publications are subject to revision, and users of this document are encouraged to investigate the possibility of applying the most recent editions of the publications indicated below. The CCSDS Secretariat maintains a register of currently valid CCSDS publications.

[Only references required as part of the specification are listed in the References subsection. See CCSDS A20.0-Y-3, *CCSDS Publications Manual* (Yellow Book, Issue 3, December 2011) for additional information on this subsection.]

[Editor note: References below are from the project description document. I don’t think any of those except for the CCSDS Recommendations will actually be referenced. I dimmed the items that I don’t expect to be in the final reference list. It is also possible to add an information only reference list.]

1. *Reference Model for an Open Archival Information System (OAIS)*. Recommendation for Space Data System Practices, CCSDS 650.0-M-2. Blue Book. Issue 1. Washington, D.C.: CCSDS, June 2012. [Equivalent to ISO 14721:2012.] Available from: <http://public.ccsds.org/publications/archive/650x0m2.pdf>
2. *Producer-Archive Interface Methodology Abstract Standard*. Recommendation for Space Data System Practices, CCSDS 651.0-M-1. Magenta Book. Issue 1. Washington, D.C.: CCSDS, May 2004. [Equivalent to ISO 20652:2006.] .] Available from: <http://public.ccsds.org/publications/archive/651x0m1.pdf>
3. *Producer-Archive Ingest Specifications*. Recommendation for Space Data System Standards, CCSDS 651.1-B-1. Blue Book. Issue 1. Washington, D.C.: CCSDS, February 2014. [Equivalent to ISO/DIS 20104:2015] .] Available from: <http://public.ccsds.org/publications/archive/651x1b1.pdf>
4. *Audit and Certification of Trustworthy Digital Repositories*. Recommendation for Space Data System Practices, CCSDS 652.0-M-1. Magenta Book. Issue 1. Washington, D.C.: CCSDS, September 2011. [Equivalent to ISO 16363:2012.] Available from: <http://public.ccsds.org/publications/archive/652x0m1.pdf>
5. *Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories*. Recommendation for Space Data System Practices, CCSDS 652.1-M-2. Magenta Book. Issue 2. Washington, D.C.: CCSDS, March 2014. [Equivalent to ISO 16919:2014.] Available from: <http://public.ccsds.org/publications/archive/652x1m2.pdf>

# Overview and COntext

[Non-normative overview text appears in section 2. See CCSDS A20.0-Y-4, *CCSDS Publications Manual* (Yellow Book, Issue 4, April 2014) for the contents of section 2.]

This standard deals with the curation aspects of the entire information lifecycle.

There are many models, some general and some domain specific, which discuss all or parts of the information lifecycle; there are many separate terms used in this area such as stewardship. This standard does not further elucidate the relationship between these terms other than to indicate that curation includes preservation and adding-value.

# THE LIFECYCLE FRAMEWORK: THE MAIN STAGES and Information to be gathered

[Normative specifications appear in sections 3 through *n*. See CCSDS A20.0-Y-4, *CCSDS Publications Manual* (Yellow Book, Issue 4, April 2014).

In the following we use the term “curation project” to mean quite generally an activity “planned and designed to achieve a particular aim ranging from the creation of new data to the preservation of existing data with a particular preservation aim.

Section 3.1 outlines the stages of the lifecycle, section 3.2 identifies the major pieces of information related to curation which needs to be collected and section 3.3 shows the way in which that information evolves through the lifecycle.

## Information Lifecycle stages



Figure 3‑1 Stages of the Information Curation Lifecycle

Figure 3‑1 above identifies the stages of the information lifecycle:

1) The Formulation stage is where those responsible for proposing the project to acquire or generate, manage or support exploitation of data or information gain approval and resources to carry out that activity.

2) The Implementation stage is where those responsible prepare to carry out the activity. The project develops (or updates) systems (hardware, software, data and processes to meet the needs of the proposed activity.

3) The Operation stage is where those responsible carry out the activity to acquire or generate, manage, and support exploitation of data or information from this activity

4) The Exploitation stage is where those responsible maintain the data or information useable and adding value if possible and proposed as part of activity. It also handles de-accessioning of the data if appropriate.

## Information topics

Table 3‑1 indicates the topics under which the curation information is grouped.

Table 3‑1 Curation information to be collected

|  |  |
| --- | --- |
| **Topic** | **Brief description** |
| Data | The data created or collected which may be re-used and preserved |
| Representation Information | See OAIS Representation Information |
| Rights | See OAIS Access Rights |
| Value | The ways in which value may be added to the data |
| Proposal | The proposal which led to resources being made available for the data creation/ collecting |
| Provenance | See OAIS Provenance |
| Authenticity | See OAIS |
| Discoverability | To be added |
| Risk | See ISO definition |
| Publications | To be added |
| Handover | The process of preparing for and completing the ingest of the data to an archive |
| Quality checking | The process of checking the quality of the data and associated information. |
| Packaging | See OAIS |
| SIP | See OAIS |

## Staged curation information

The next subsections will provide brief descriptions of the information to be collected at each stage as well as the high-level actors and activities.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Topic** |  | **Formulation** | **Implementation** | **Operation** | **Exploitation** |
| Data | Volume that would require preservation | Rough idea | Increasingly detailed | Becoming complete | Complete |
| Types of data (raw, processed etc) which should be preserved | Rough idea | Increasingly detailed | Becoming complete | Complete |
| Choice of data format | Rough idea | Increasingly detailed | Becoming complete | Complete |
| Semantics of the data elements | Rough idea | Increasingly detailed | Becoming complete | Almost complete |
| Inventory of data produced/expected | Rough idea | Increasingly detailed | Becoming complete | Complete |
| Representation Information | Outline of background concepts needed to understand the project | Rough idea | Increasingly detailed | Becoming complete | Up to date and accumulating |
| Applicable standards | Rough idea | Increasingly detailed | Becoming complete | Up to date and accumulating |
| Software dependencies | Rough idea | Increasingly detailed | Becoming complete | Up to date and accumulating |
| Data dictionaries and other semantics | Rough idea | Increasingly detailed | Becoming complete | Up to date and accumulating |
| Format definitions and formal descriptions | Rough idea | Increasingly detailed | Becoming complete | Up to date and accumulating |
|  |  |  |  |  |
| Rights | What are the restrictions on access in the long term | Fairly firm | Very detailed | Complete | Complete |
| Licences involved | Fairly firm | Very detailed | Complete | Complete |
| Owners of the data – who can authorise hand-over | Fairly firm | Very detailed | Complete | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| (Adding) value | Related data which may in the future be combined with this data |  | Increasingly detailed | Increasingly detailed | Growing |
| Other software which may be used on the data |  | Increasingly detailed | Increasingly detailed | Growing |
| Interfaces that are applicable to the data e.g. images, tables. |  | Increasingly detailed | Increasingly detailed | Almost complete |
| Potential value of the data and likely business case for sustainability | Rough idea | Rough idea | Developing | Developing |
|  |  |  |  |  |
|  |  |  |  |  |
| Proposal | Record of origins of the project e.g. in a CRIS system | Fairly firm | Complete | Completed | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Provenance | Documentation about the hardware and software used to create the data, including a history of the changes in these over time  |  | Increasingly detailed | Becoming complete | Up to date and accumulating |
| Processing workflow |  | Increasingly detailed | Becoming complete | Complete |
| Processing inputs |  | Increasingly detailed | Becoming complete | Complete |
| Processing parameters |  | Increasingly detailed | Becoming complete | Complete |
| When each stage was performed |  | Increasingly detailed | Becoming complete | Complete |
| Record of any special hardware needed |  | Increasingly detailed | Becoming complete | Complete |
| Authenticity | Who was responsible for each stage of processing |  | Complete | Complete | Complete |
| Fixity (e.g. CRC or digest) of data which may be preserved |  | Complete  | Complete | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Discoverability | What may be used in future to identify the data | Fairly firm | Increasingly detailed | Increasingly detailed | Evolving |
| Methods for exploration/quick-look at the data | Fairly firm | Increasingly detailed | Increasingly detailed | Evolving |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Risk | Potential risks to preservation and exploitation of the data |  |  | Rough idea | Evolving |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Publications | Publications related to the data | Rough idea | Increasingly detailed |  | Evolving |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Handover | Pointers to the components to be transferred to the archive | Rough idea | Increasingly detailed | Increasingly detailed | Complete |
| Identification of archives which are likely to be able to host the data | Rough idea | Increasingly detailed | Increasingly detailed | Complete |
| Potential preservation aims of the archive | Rough idea | Increasingly detailed | Increasingly detailed | Complete |
| Inputs to definition of the Designated Community | Rough idea | Increasingly detailed | Increasingly detailed | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
| Quality checking | Quality checks which may be performed on the data by non-experts |  | Increasingly detailed | Increasingly detailed | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| Packaging | Details of the way components are packaged together for delivery to a repository |  |  | Increasingly detailed | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| SIP | Definition of SIPs |  |  | Developing | Complete |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

### Formulation Stage

The participants in the formulation stage will certainly include Sponsors and Proposers and may also include Data Managers and Archives.

Sponsors typically have a well-defined charter that identifies their area of interest, within which, for example, programs are identified for specific research topics. Sponsors may bring in domain experts from existing projects and interest groups to develop strategic plans and objectives for the programs as well as long term schedules to fulfill the plans and objectives.

Proposers develop proposals that describe potential projects The proposals are likely to provide estimates of the cost and a preliminary risk assessment. The proposer may also prepare a data management plan that identifies the information gathering system, data processing system, and the types of products that will be produced. The data management plan will include estimates of data volumes and a schedule of activities and deliveries.

The proposer will often have to work with an archive to understand the standards in place and the mechanisms for delivering information objects. The archive may be a co-signatory for the proposal or at least the Data Management Plan. The archive uses the data management plan to develop a support plan that identifies the cost and resources that will be required to support the Project. These resource requirements need to be integrated into the archive's long term budget and schedule.

Some projects may be able to use the Proposal and Data Management Plan as their system design. But for large and complex projects the selection is only the beginning of a substantial design process. We identify a System Design document to represent what may actually represent a sequence of requirements and design documents. The initial project design includes all elements of the acquisition, transfer, processing, distribution and analysis systems. For large projects there may be substantial negotiation and iteration between the sponsor and the project during the design process.

“Information produced during this stage provides a snapshot of the scientific and technical framework in which the project was born. Project … requirements, assessment studies, technology readiness reviews and cost analysis are performed during this stage. Preserving this information – both for approved and not approved projects – would allow future users to have reference material for new missions evaluation and definition. Traceability of this information is also useful to compare initial expectations to what was actually achieved by the mission and to understand which changes occurred between the pre-mission and the next stages.” (LTDP PDSC document).

### Implementation Stage

The implementation stage is performed by the project team. Once the project proposal is approved, the project implementation team develops the information acquisition systems or provides access to existing data sources. This could be as complicated as building a spacecraft or as simple as accessing an online database. The project implementation team may provide command and control systems to interact with the information collection systems. The project implementation team develops processing software and scripts. It provides data processing and distribution systems which include analysis and access software. It documents the information content of all these systems via interface specifications, information models and data dictionaries.

“This stage produces the entire project and data detailed definition documents. It includes Sensor/Instrument requirements, characteristics, calibration methods, etc. Preserving this information is fundamental to understand changes that may have occurred over time while in operation. It is also needed to understand procedural impacts relative to instrument, algorithm and product implementation. Data acquired during the calibration and validation campaigns of the instrument under construction (e.g. in a laboratory or dedicated campaigns) is of critical importance as a reference for the future use of the data.” (LTDP PDSC document).

### Operation Stage

The operation stage is performed by the project team possibly with support from the archive. It operates the acquisition systems to gather data. The project team receives and processes the data to support both command and control and data analysis. The project team maintains a record of all acquisition and processing activities that might be useful for understanding and interpreting the data. It performs data analysis and publishes results. The project team operates the implemented distribution mechanisms for transferring data to the project team and other users. ItThe project team prepares Submission Information Packages (SIPs) for delivery to the long-term archive and participates in peer reviews.

### Exploitation Stage

The exploitation stage is performed by the long-term archive possibly with support from the project. The archive receives and validates SIP deliveries from the project. It performs any conversion or consolidation necessary to integrate the data and metadata content of the SIPs into the archive collection. It transfers the data contents of the SIPs to the archive repository. It extracts metadata from the SIPs and loads it into the archive registry or database. It provides an access capability for users to be able to search, select and download both metadata and data from the archive collection. It provides for long term preservation by insuring the integrity, usability and accessibility of the data.

The differences between the operation and the exploitation stages are in the source of funding, the designated community and maturity of the products. The project and the archive may have completely separate funding sources. The project users may be limited to a small team whereas the archive may represent a large community. Special hardware and software capabilities may be provided to project teams but not to archive users. The archive may provide capabilities for exploiting the project data in concert with other data products.

1. [ANNEX TITLE]

[EITHER Normative or Informative]

[Annexes contain ancillary information. Normative annexes precede informative annexes. Informative references are placed in an informative annex. See CCSDS A20.0-Y-4, *CCSDS Publications Manual* (Yellow Book, Issue 4, April 2014) for discussion of the kinds of material contained in annexes.]

1. Security Considerations

(Informative)
	1. Introduction
	2. security concerns with respect to the CCSDS document
		1. Data privacy
		2. Data integrity
		3. Authentication of communicating entities
		4. Control of access to resources
		5. Availability of resources
		6. Auditing of resource usage
	3. Potential threats and attack scenarios
	4. Consequences of not applying security to the technology