INTERNATIONAL STANDARD

ISO 19165-2

First edition
2020-07

Geographic information — Preservation of digital data and metadata —

Part 2: Content specifications for Earth observation data and derived digital products

Information géographique — Archivage des données numériques et des métadonnées —

Partie 2: Spécifications de contenu pour les données d'observation de la Terre et les produits numériques dérivés

Reference number
ISO 19165-2:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

A list of all parts in the ISO 19165 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Many agencies across the globe are generating important datasets by collecting measurements from instruments in-situ and on board aircraft and spacecraft, globally and constantly. The data resulting from such measurements and digital products derived from them are valuable resources that need to be preserved for the benefit of future generations. These observations are the primary record of the Earth's environment and are therefore the key to understanding how conditions in the future will compare to conditions today. Earth observational data, derived products and models are used to answer key questions such as "How is the global Earth system changing?", "What are the sources of change in the Earth systems and what are their magnitudes and trends?", "How will the Earth system change in the future?", and "How can Earth system science improve mitigation of and adaptation to global change?".

In the near-term, as long as the missions' data are being used actively for research and applications, it continues to be important to provide easy access to the data and services commensurate with current information technology. For the longer term, when the focus of the research community shifts towards new missions and observations, it is essential to preserve the previous mission data and associated information. This will enable a new user in the future to understand how the data were used for deriving information, knowledge and policy recommendations and to "repeat the experiment" to ascertain the validity and possible limitations of conclusions reached in the past as well as to provide confidence in long-term trends that depended on data from multiple missions.

Organizations that collect, process and utilize Earth observation data today have a responsibility to ensure that the data and associated content continue to be preserved by gathering this information and preserving it themselves, or by handing it off to other organizations. In order to ensure preservation of all the content necessary for understanding and reusing the data and derived digital products, a standard is needed that specifies this content. While there are existing standards that address archival and preservation in general, there are no existing international standards or specifications to address what content should be preserved.

Specifications for preservation of information content complement existing archive standards. Space agencies that are members of the International Consultative Committee for Space Data Systems (CCSDS) have long recognized the importance of developing information standards for use in long-term preservation of space-related data collections. Volunteers developed the Open Archival Information System Reference Model (OAIS-RM). Subsequent activities continue to expand through a range of related interests that reach towards more practical guidance for developing agency standards. An example of this is a recommended standard on packaging of data and metadata (XFDU), to facilitate information transfer and archiving[11]. The most recent update to the OAIS-RM is ISO 14721. The OAIS-RM provides a conceptual framework for archiving digital information. The CCSDS has also developed ISO 16363, which specifies requirements for certification of trustworthy digital repositories, based on the OAIS-RM, and ISO 16919, which describes how to audit archives for conformance with the requirements.

ISO 19115-1 provides a metadata model for describing geographic information and services, and ISO 19115-2 augments ISO 19115-1 with additional structure to describe the acquisition and processing of geographic imagery and gridded data. It provides the structure needed to represent properties of the instruments acquiring data, e.g. instrument geometry and production processes. The structure provided by ISO 19115-2 is useful for representing the preservation content intended to be specified with this document (ISO 19165-2).

ISO 19165-1 considers geographic information preservation in general and this document (ISO 19165-2) is its extension for Earth observation data and its derived products.

ISO 19165-1:2018, 7.3.1 indicates that specific content items needed to preserve the full provenance and context of data and associated data depend on the needs of the designated communities and types of datasets (e.g., maps, remotely sensed data from satellites and airborne instruments, physical samples). It also states that follow-up parts to ISO 19165-1 may be developed to provide details of content items appropriate to specific disciplines.

This document, as Part 2 of the ISO 19165 series, provides more detailed specifications for Earth observation data and derived digital products resulting from spaceborne and airborne remote sensing, as well as in situ observations.

This document benefits from the work performed by the Data Preservation and Stewardship Committee of the U.S. Earth Science Information Partners (ESIP)[5], NASA[6], ESA and CEOS WGISS[7]. The documents from these groups are integrated along with the ISO international standards mentioned above to provide specific content items to be preserved from Earth observing missions for the benefit of users. It is expected that if the content items specified by this document are preserved, users will have sufficient information to be able to understand, reuse, and, ideally, regenerate data products without the assistance of the original teams that were responsible for their initial generation.

Geographic information — Preservation of digital data and metadata —

Part 2:
Content specifications for Earth observation data and derived digital products

1 Scope

This document aims to extend the long-term preservation of digital geospatial data to provide details about content describing the provenance and context specific to data from missions that observe the Earth using spaceborne, airborne or in situ instruments.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 19115-1, *Geographic information — Metadata — Part 1: Fundamentals*

ISO 19115-2, *Geographic information — Metadata — Part 2: Extensions for acquisition and processing*

ISO 19115-3, *Geographic information — Metadata — Part 3: XML schema implementation for fundamental concepts*

ISO 19130-1, *Geographic information — Imagery sensor models for geopositioning — Part 1: Fundamentals*

ISO/TS 19130-2, *Geographic information — Imagery sensor models for geopositioning — Part 2: SAR, InSAR, lidar and sonar*

ISO 19157-1, *Geographic information — Data quality — Part 1: General requirements*

ISO 19157-2, *Geographic information — Data quality — Part 2: XML schema implementation*

ISO/TS 19159-1, *Geographic information — Calibration and validation of remote sensing imagery sensors and data — Part 1: Optical sensors*

ISO/TS 19159-2, *Geographic information — Calibration and validation of remote sensing imagery sensors and data — Part 2: Lidar*

ISO/TS 19159-3, *Geographic information — Calibration and validation of remote sensing imagery sensors and data — Part 3: SAR/InSAR*

ISO 19165-1, *Geographic information — Preservation of digital data and metadata — Part 1: Fundamentals*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 19165-1 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
analysis ready data
<earth observation> data that have been processed to a minimum set of requirements and organized into a form that allows immediate analysis with a minimum of additional user effort and interoperability both through time and with other datasets

Note 1 to entry: The definition from Committee on Earth Observation Satellites (CEOS) for CEOS Analysis Ready Data for Land (CARD4L) has been generalized here to include data other than satellite data by omitting the word "satellite" from the definition.

Note 2 to entry: Adapted from Reference [10].

3.2
ancillary data
<earth observation> data which are not obtained from the sensor itself (usually provided in the science telemetry) and have the primary purpose to serve the processing of instrument data

Note 1 to entry: Ancillary data refers to data that exist purely to serve the data processing. *Auxiliary data* (3.3), while helping the process, are also datasets in their own right.

Note 2 to entry: Ancillary data may be 'engineering', 'core housekeeping' or 'subsystem' data obtained from other parts of the measurement platform. It may include parameters such as position and velocity, attitude and its range of change, time, temperatures, pressures, internally produced magnetic fields, and other environmental measurements.

Note 3 to entry: Some missions may treat ancillary data as datasets in their own right, thus blurring the distinction being made here between ancillary and *auxiliary data* (3.3).

Note 4 to entry: The definition in the original source is tailored to spaceborne Earth observations. It has been slightly modified here to be more general. The concept is that ancillary data are those not collected by the sensor itself and that their primary purpose to serve processing applies to Earth observation data from airborne and in situ instruments as well.

Note 5 to entry: Adapted from Reference [8].

3.3
auxiliary data
<earth observation> data which enhance processing and utilization of the Earth observing instrument data

Note 1 to entry: The auxiliary data are not captured by the same data collection process as the instrument data. Auxiliary data include data collected by any other platform or process, preferably in georeferenced digital format. Auxiliary data help in data processing, but are also datasets in their own right.

Note 2 to entry: Adapted from Reference [8].

3.4
dataset
identifiable collection of data

[SOURCE: ISO 19115-1:2014, 4.3, modified — Note 1 to entry has been deleted.]

3.5
dataset series
collection of datasets sharing common characteristics

[SOURCE: ISO 19115-1:2014, 4.4]

3.6
granule
smallest aggregation of data which is independently managed

Note 1 to entry: Granules may be managed (i.e. described, inventoried, retrievable) as logical granules and/or physical granules.

Note 2 to entry: Granule is often equivalent to *dataset* (3.4).

3.7
mission
<earth observation> activity that uses spaceborne, airborne or in situ instruments

Note 1 to entry: Some organizations reserve the term "mission" for satellite observation activities and refer to airborne and in situ observation activities as "investigations" and "field campaigns", respectively.

3.8
product level
<earth observation> number indicating the degree of processing that has been performed on the observed data

Note 1 to entry: Product levels 0 through 4 indicate the degree of processing performed on the raw data to convert them into more useful parameters and formats. The Committee on Earth Observation Satellites (CEOS) defines product levels as follows:

— Raw Data: Data in their original packets, as received from a satellite.

— Level 0: Reconstructed unprocessed instrument data at full space time resolution with all available supplemental information to be used in subsequent processing (e.g. ephemeris, health and safety) appended.

— Level 1: Unpacked, reformatted level 0 data, with all supplemental information to be used in subsequent processing appended. Optional radiometric and geometric correction applied to produce parameters in physical units. Data generally presented as full time/space resolution. A wide variety of sub level products are possible.

— Level 2: Retrieved environmental variables (e.g. ocean wave height, soil moisture, ice concentration) at the same resolution and location as the level 1 source data.

— Level 3: Data or retrieved environmental variables which have been spatially and/or temporally re-sampled (i.e. derived from level 1 or 2 products). Such re-sampling may include averaging and compositing.

— Level 4: Model output or results from analyses of lower level data (i.e. variables that are not directly measured by the instruments but are derived from these measurements).

Note 2 to entry: The product levels defined here are derived from satellite remote sensing heritage. For the case of airborne and in situ observations, these do not necessarily apply, but may be used as appropriate for indicating the degree of processing performed on the observed data.

Note 3 to entry: Adapted from Reference [9].

3.9
stage
<earth observation> well-defined part of the lifecycle of a mission

4 Symbols and abbreviated terms

Doc	Document
CEOS	Committee on Earth Observation Satellites
ESA	European Space Agency

ESIP	Earth Science Information Partners
GNSS	Global Navigation Satellite System
ICD	Interface Control Document
L0	Product Level 0
L1	Product Level 1
L2+	Product Level 2 or higher, including Analysis Ready Data
NASA	National Aeronautics and Space Administration (United States)
QA4EO	Quality Assurance framework for Earth Observations
SI	International System (of units)
SW	Software
WGISS	Working Group on Information Systems and Services

5 Conformance

In order to conform to this document, the abstract test suite in Annex A shall be used.

6 Mission stages

6.1 General

This document covers missions that observe the Earth using spaceborne, airborne or in situ instruments. The preservation of observations (i.e. measurements from such instruments) and digital data products derived from the observations is important for the benefit of future users. In addition, it is also important to preserve metadata and other content items such as ancillary data, planning documents and associated knowledge in order for future users to be able to understand and reuse the data and possibly reproduce the results from the missions. Such content items are created and gathered at different stages of the mission (defined below). Annex B provides a mapping between the stages and satellite mission phases defined by NASA and ESA. These stages are conceptually applicable to any type of mission. Along with the definitions for each stage, a few examples of the types of activities appropriate to the stage are given to clarify the definition. However, the specific contents, number of items to be preserved, and the complexity of the preservation activities depend on the specific type of mission. While the focus of this document is the preservation content, i.e. "what" needs to be preserved, enumerating the content associated with the mission stages helps address "when" the content items need to be preserved, as well.

6.2 Mission concept stage

The mission concept stage is the period when ideas for a mission are developed and proposed to funding entities. At this stage, the mission is defined to a level sufficient to show the scientific/applications value and technical feasibility. During this stage, science and applications requirements are identified. Additional activities may include identification of plans and tools to be used in preliminary system level studies. Feasibility verification documents, mission technology and programmatic estimates for the future mission stages may also be generated.

6.3 Mission definition stage

The mission definition stage is the period when mission scientific/applications requirements are defined in detail and technical solutions are selected for the system concept. During this stage, types of scientific measurements (e.g. spectral analysis, temperature measurement) may be identified and defined.

6.4 Mission implementation stage

The mission implementation stage is the period when the detailed design, implementation, and testing of the mission system and its components are realized. These may include:

— sensors/instruments;

— algorithms and their interfaces;

— methods of measurement;

— any other context necessary to perform measurements.

6.5 Mission operations stage

The mission operations stage is the period when

— data are captured;

— algorithms are revised and improved;

— input are analysed;

— calibration and validation of sensor/instrument as well as activities concerned with qualification of processed data are performed; and

— higher level derived digital products are generated.

6.6 Post mission stage

The post mission stage is the period after mission operations are completed, and includes the post-operations and preservation. The post mission stage may start with the satellite end of life (e.g. for an Earth observation mission with the event of satellite disposal or failure), the completion of the last planned aircraft flight in a series that constitutes a mission, or the last planned activity in a series of in situ measurement activities. The post mission stage focuses on:

— consolidation and appraisal of datasets (data and information);

— reprocessing of datasets to align to the latest version;

— ground segment and media disposal (depending on specific mission);

— migration of data and associated information to a long-term preservation environment.

During the post mission stage, a limited set of functions (e.g. data discovery and access) may be provided by the mission operations team until data migration to a long-term preservation environment. This stage can extend beyond the point where the preservation package has been prepared and archived and involve updates to algorithms and reprocessing. This stage also focuses on:

— historical data reuse and exploitation;

— preservation of data and related information against aging and technological changes; and

— data curation and enrichment.

7 Preservation content

7.1 General

The content to be preserved is discussed below in one subclause for each of the mission stages described above. Each subclause provides the rationale for preserving the listed content items and a list of items recommended to be preserved. Each of the items listed in the tables may correspond to one or more documents (Doc), software (SW) objects or data records. It is also possible for some missions that some of the documents specified in the tables can be combined into a single document. The list shall be tailored to each mission and the specifics of which items will be preserved to satisfy the content requirements indicated in this clause shall be documented as preservation metadata. Also, it is to be noted that many of the content items in a given stage may need to be updated during subsequent stages. It is assumed that such updates are made and the resulting items are preserved in a version-controlled manner. In Tables 1 to 5, shown in the following subclauses, the column headings are:

— ID: a short identifier for the content item;

— Need for: indicates for which type of mission the specified item is needed (satellite: SAT, aircraft: AIR, field campaign: FLD, all types: ALL);

— Type: indicates whether the item is a document, data record, software, etc.;

— Identification: a phrase identifying the content item (longer than ID);

— Description: a brief description of the content item;

— Examples of types of quality information: indicates information to be contained in the content item for the item to be considered to be of high quality.

The standards in the ISO 19115 series, ISO 19130 series, ISO 19157 series, 19159 series and 19165-1 shall be considered as appropriate in representing the content items listed in the Tables 1 to 5. Annex C (informative) illustrates how the items can be organized in an ISO 19115-1 DS_Series that holds references to all preservation information for a mission.

7.2 Mission concept stage

7.2.1 Rationale

Information produced during this stage provides a snapshot of the framework in which the mission was born. Mission and sensor requirements, assessment studies, technology readiness review and cost analysis are performed during this stage. Documents generated during this stage show the objectives and plans for the mission. Preserving this information would allow future users to have reference material for evaluation and definition of new missions. Traceability to this information is also useful for comparing initial expectations with mission results and for understanding changes that may have occurred between this stage and the following stages.

7.2.2 Content

The content required to be preserved by the end of the mission concept stage is identified and described in Table 1.

Table 1 — Mission concept stage preservation content

ID	Need for	Type	Identification	Description	Examples of types of quality information
MC_1.0	ALL	Doc	Preservation metadata	List of items preserved at this stage to satisfy content requirements specified below. (Note that this item will be updated at each of the subsequent stages.)	Identification of items by title and persistent ID. Access method (e.g., link). Indication of access restrictions if any as well as whether and when the restrictions will be removed.
MC_1.1	ALL	Doc	Scientific/ applications scenario, data producer and user communities	Defines scientific/applications scenario and expected goals. Lists Principal Investigator, designated user communities and third-party actors.	Required uncertainty bounds for services and applications, lifetime, data availability, data accuracy, data latency, revisit frequency or data acquisition timeline, geographical coverage, spatial resolution. Names of key science team leads and product team members, roles, performing organization, contact information, sponsoring agencies or organizations are included (within the constraints of applicable privacy regulations and policies). As responsibility changes hands in subsequent stages, the names of individuals and periods during which they were responsible for various aspects of the product should be documented (within the constraints of applicable privacy regulations and policies).
MC_1.2	SAT	Doc	Mission requirements document	Defines scientific/applications mission and sensor requirements, processing methods and qualification methods. Includes instrument specifications (e.g., frequencies, bandwidths, polarizations, antenna size, and scan modes), data products to be produced, and operations concepts.	Calibration plan and quality assessment plan for the mission. Uncertainty requirements for instrument product (e.g. radiometric/geometric uncertainty bounds, coverage, revisit times, etc.). Justification for the design decisions (e.g. band selection).
MC_1.3	ALL	Doc	Mission operation plan	Defines the plan for how the mission will be conducted.	Initial operations concept.
MC_1.4	ALL	Doc	Mission cost and schedule	Defines planned cost and schedule for the mission.	Initial estimated costs and schedule timelines

7.3 Mission definition stage

7.3.1 Rationale

In this stage, detailed definition documents are produced for the entire mission and data, including sensor/instrument requirements, characteristics, calibration methods, etc. Preserving this information is fundamental to forming a baseline and to understanding changes that may have occurred during the mission operations stage.

7.3.2 Content

The content required to be preserved by the end of the mission definition stage is identified and described in Table 2 below.

Table 2 — Mission definition stage preservation content

ID	Need for	Type	Identification	Description	Examples of types of quality information
MD_1.0	ALL	Doc	Preservation metadata	List of items preserved at this stage to satisfy content requirements specified below. Updated version of MC_1.0	Identification of items by title and persistent ID. Access method (e.g., link). Indication of access restrictions if any as well as whether and when the restrictions will be removed.
MD_1.1	ALL	Doc	Mission requirements specifications	Defines mission requirements, functional and resource allocation between mission measurement platform and data capture systems (e.g., ground systems for satellite operations), and operational scenario (e.g. flight plans for airplanes, un-crewed aerial vehicles, and/or drones). This may be an updated version of documentation specified in MC_1.2.	Description of the information at a system level (e.g. revisit times and mission products uncertainty) and at a subsystem level (e.g., for instruments, instrument stray light, channel crosstalk, spatial sampling, field of view, observation mode, spectral channels).
MD_1.2	Primarily SAT	Doc	Space or aircraft to ground segment ICDs	Defines the main systems/segments, ICDs, system latency budget estimation and data flow, including transmission of data from spacecraft to ground, transfer of recorded data on board aircraft, handling of in situ measurement data, etc.	Error control (e.g. Cyclic Redundancy Check), data latency, data rate, quality flags, packets lost/damaged, timeliness etc. for different scenarios (e.g. Near-Real-Time, calibration mode, ground stations availability and relative position).
MD_1.3A	ALL	Doc	Sensor/instrument/platform requirements	Defines the sensor/instruments/platform requirements for design (e.g. spectral bands, bandwidths, scan modes, polarizations, performance, antenna size; for photogrammetry instruments: camera/sensors, GNSS receiver, inertial measurement unit, navigation system; requirements for platforms - satellite, aircraft, tower, buoy)	Sensor uncertainty budget based on previous knowledge. Specification of uncertainty associated with optical properties (e.g. noise, linearity, calibration accuracy, signal synchronization, electrostatic protection, and temperature and pressure ranges).

Table 2 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MD_1.3B	ALL	Doc/data record	Sensor/instrument processing characteristics	Characteristics for processing of acquired data, data processing model	Assessment of performance/acceptability including uncertainty, linearity, sun-glint, stray light. Documented model descriptions, version control.
MD_1.4A	ALL	Doc/data record	Sensor/instrument qualification process	Qualification process for sensor, captured data, processed data.	Documented procedure for validation. Validation of model and software. Validation by comparison with other models or reference datasets including simulated products and ground measurements. Data collected for supporting validation.
MD_1.4B	SAT	Doc/data record	Pre-launch/pre-operational calibration and characterization plan	Calibration requirements – documentation of pre-launch/pre-operational calibration methods and data from such calibration. Pre-launch calibration/ pre-operational calibration includes: optical tests, thermal test, external calibration test, field of view determination.	Identification of reference standards, pre-flight calibration methods, re-calibration intervals. Uncertainty goals, such as expected variances or probability distributions of error.
MD_1.4C	SAT	Doc / Data Record	Ground/ocean calibration reference and scientific base	Calibration requirements - including description of ground/ocean reference sites, accuracy, and stability of the site conditions. Data from such calibration sites.	Traceability to International System of Units (SI) via international reference standards: procedures, calibration certificates, traceability statement, and uncertainty analysis.

Table 2 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MD_1.5A	SAT and AIR; FLD if available	Doc	Processing algorithms and data format specification	Defines: mathematical models and algorithms for mission data processing including algorithm theoretical basis; high-level data flow diagrams; assumptions about algorithm performance and limitations; auxiliary and ancillary data usage; data and product format requirements and standards; metadata to facilitate discovery, search, access, understanding and usage associated with each of the data products.	Documented descriptions of mathematical models and algorithms for mission data processing, including: assessment of performance/acceptability; peer reviewed papers; simulation for validation results; validation by comparison with test datasets; product validation criteria; auxiliary and ancillary data usage; data and product format requirements and standards including: metadata specifications (including quality information/ parameters), naming conventions, and version controls.
MD_1.5B	ALL	Doc	Data product specifications	Provides a detailed description of data products and their characteristics. It is recommended that ISO 19131 be followed for data product specifications. Descriptions of data products' structure, format, range of values and special fill or error values. Detailed description of output data products, sufficient to determine if products meet their specified requirements.	Description of uncertainty/quality indicators and method to provide uncertainty to different users. Includes content, format, latency, accuracy and quality.
MD_1.6	ALL	Doc	Data management plan (DMP)	Preliminary DMP describing how the data and derived products will be managed during the mission	Descriptions of responsibilities of individuals or organizations for managing data including capture, processing, archiving and distribution; data flows; identification of interfaces and interface control documents to be developed; data quality plan; long-term preservation plan.
MD_1.7	ALL	Doc	Mission cost and schedule	Defines planned cost and schedule for the mission.	Updated estimated costs and schedule timelines (See MC_1.4)

7.4 Mission implementation stage

7.4.1 Rationale

Preserving all the information produced during the mission implementation stage is necessary for understanding procedural impacts relative to instrument, algorithm and product implementation. Data acquired during the calibration and validation campaigns of instruments under construction (e.g. in a laboratory or dedicated field campaigns) are of critical importance as a reference for the future use of the data. Documentation of measurements made before deploying instruments will help establish a baseline and help users understand changes that may have occurred over time while in operation.

7.4.2 Content

The content required to be preserved by the end of the mission implementation stage is identified and described in Table 3 below.

Table 3 — Mission implementation stage preservation content

ID	Need for	Type	Identification	Description	Examples of types of quality information
MI_1.0	ALL	Doc	Preservation metadata	List of items preserved at this stage to satisfy content requirements specified below. Updated version of MD_1.0	Identification of items by title and persistent ID. Access method (e.g., link). Indication of access restrictions if any as well as whether and when the restrictions will be removed.
MI_1.1	ALL	Doc	Mission Design	Defines mission requirements specification and describes design as implemented.	Clear identification of technical procedure. Record of decisions made during implementation.
MI_1.2A	SAT and AIR	Doc	Detailed space or aircraft to ground segment operations concept and implementation	Defines the detailed operational implementation and any contingency procedure/plan needed.	Recording procedure for assuring the data integrity and quality. Storing of diagnostic information received.
MI_1.2B	ALL	Doc	Updated DMP	DMP specified in MD_1.6 updated with more details and "To Be Determined (TBD)" items from the preliminary DMP filled in.	Descriptions of responsibilities of individuals or organizations for managing data including capture, processing, archiving and distribution; data flows; identification of interfaces and interface control documents to be developed; data quality plan; long-term preservation plan.
MI_1.2C	SAT	Doc	On board processing	Documentation of on board processing, if any.	Algorithm description and software documentation as well as software validation information.
MI_1.3	SAT (AIR and FLD if applica-ble)	Doc	Sensor/instrument design and implementation	Defines the sensor/instrument platform design, implementation and performance.	Testing results including uncertainty. Uncertainty budget with supporting evidence (from on ground characterization).

Table 3 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MI_1.4	SAT, AIR	Doc / data records	Calibration and validation	Independent calibration and validation campaign method, data validation activities with simulated data. Numeric (digital data) files of instrument/sensor characteristics including pre-flight or pre-operational performance measurements (e.g., spectral response, instrument geometric calibration (geo-location offsets), and noise characteristics).	Pre-launch/Pre-deployment calibration results, uncertainty budget with supporting evidence, traceability to SI validation results.
MI_1.5A	SAT	Doc	Ground processor design, algorithm implementation and supporting information for data processing.	Defines the design and implementation of the ground data processors and the algorithm. Includes methods for geo-location, radiometric calibration, and computing geophysical parameters. Includes supporting information for data processing (e.g. ancillary, auxiliary data description and usage), and sampling or mapping algorithms used in creation of the product. Includes description of numerical implementation, including any issues with computationally intensive operations, including impact of workarounds on quality. Includes data processing software test plans and sources of test data.	Algorithm description and software documentation and validation for all software used on ground and on board. Design includes content, format, latency, accuracy and quality. Clear naming conventions and version control. Sufficient metadata to facilitate discovery, search, access, understanding and usage associated with each of the data products. (Some of the information called for here can be covered in the updated DMP – see MI_1.2B, in which case, a reference to the DMP would suffice).
MI_1.5B	ALL	Doc	Technical notes Scientific papers	Algorithm description and software validation for all software used on ground and on board. Metadata and naming conventions.	Properly controlled versions and their descriptions. Complete metadata to facilitate discovery, search, access, understanding and usage associated with each of the data products.
MI_1.6A	ALL	Doc	Data format specifications	Contains information that will allow the user to read and use the data. Includes data format standard(s) used.	Data format naming conventions, performance of compression algorithm, quality indicator specification.
MI_1.6B	ALL	Doc / Data Records	Supporting information for processing	Defines and identifies ancillary and auxiliary data and how they are to be used in processing. Documentation about and data records for all ancillary data or other datasets used in generation or calibration of the data or derived product at all levels.	Appropriate quality indicator for ancillary/auxiliary data to be used in the mission operations stage with the relevant metadata.

Table 3 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MI_1.7	ALL	Doc / Data Records	Qualification Process	Detailed qualification methods and data	Assessment of performance/ acceptability.

7.5 Mission operations stage

7.5.1 Rationale

The mission operations stage provides the data that will be analysed by the scientific (and other user) community and will be the core of the mission preservation objective. These data are the concrete heritage that the mission will leave to future generations. Software related to this mission stage needs to be preserved in order to understand, use, process, and exploit data in the future. Documents also need to be preserved to support users' comprehension of the data and enable future verification and validation of mission results.

7.5.2 Content

The content required to be preserved by the end of the mission operations stage is identified and described in Table 4 below.

Table 4 — Mission operations stage preservation content

ID	Need for	Type	Identification	Description	Examples of types of quality information
MO_1.0	ALL	Doc	Preservation metadata	List of items preserved at this stage to satisfy content requirements specified below. Updated version of MI_1.0	Identification of items by title and persistent ID. Access method (e.g., link). Indication of access restrictions if any as well as whether and when the restrictions will be removed.
MO_1.1	SAT	Doc	Mission data access and service requirements document and user handbook	Defines the data archival and processing/ reprocessing strategy, the data accessible to users and the service requirements & performance during the operations stage.	Clear identification of technical procedures.
MO_1.2	SAT	Doc	Sensor ground segment operations plan	Describes the actual implementation of the end-to-end mission operations.	—
MO_1.3	ALL	Doc/ data records	Mission operations acquisition plans and reports	Describes acquisition plans and reports for data from the mission sensor(s). Includes mission logs accounting for data gaps, manoeuvres, etc.	Availability of data, data quality, model evolution, calibration parameter evolution, geo-location performance, data anomalies.
MO_1.4	ALL	Data records	Raw/product level 0	Raw or level 0 data from the sensor or instrument data packets.	Completeness of data, timeline, certification of L0 processing.

Table 4 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MO_1.5A	ALL	Data Records	Product level 1	Processed science data L1 products. The final version of a derived product should be the version archived. If results reported in peer reviewed publications were based on earlier versions of the product, those versions or at least representative subsets of those versions should also be archived. At a minimum, the algorithm and software that generated earlier versions should be archived. Metadata at the dataset (granule) level should indicate which version of software was used for producing a given dataset (granule).	Associated uncertainties. Processing algorithm recorded and validated. Reference to calibration, traceability. For geometrically located area, geometric alignment and resampling information.
MO_1.5B	ALL	Data records	Product level 2 +	Processed science data L2 products and higher, including analysis ready data. The final version of a derived product should be the version archived. If results reported in peer reviewed publications were based on earlier versions of the product, those versions or at least representative subsets of those versions should also be archived. At a minimum, the algorithm and software that generated earlier versions should be archived. Metadata at the dataset (granule) level should indicate which version of software was used for producing a given dataset (granule). Metadata at the dataset (granule) level, or at dataset series (collection) level as appropriate, should include information about calibration parameters, precision orbit & attitude data; climatological norms, geophysical masks; first-guess fields from numerical weather or climate models; spectrum and transmittance information; detailed site descriptions and field methodology should be documented for in-situ sampling.	Associated uncertainties. Processing algorithm recorded and validated. Reference to calibration, atmospheric corrections, traceability. Reference to validation where relevant.
MO_1.6	ALL	Data records	Browse images	Browse digital catalogue and images, when generated.	Images to facilitate preview.

Table 4 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MO_1.7	SAT, AIR (FLD if needed)	Data records	Ancillary data	Attitude, ephemeris, navigation parameters, observation counters, orbital state vectors, times, sun position, temperatures sensor/CCD/amplifier noise, Earth relative position, azimuth, satellite or aircraft manoeuvres, instrument parameters (e.g. optical response), aircraft position, locations of in situ instruments.	Quality flags and performance parameters, e.g., orbit accuracy, temperature stability.
MO_1.8	SAT (AIR/FLD if available)	Data records	Auxiliary data	Band/multispectral/band-by-band parameters for algorithms, non-linearity correction factors, error/failure/gap correction factors, calibration curve/factors or look-up tables, scaling correction factors, atmospheric correction factors, geometry correction factors, drift factor, albedo parameters, instrument modes, incident angle, absolute calibration constants, solar radiance, moon temperature brightness, local seasonal variances, weather forecast/actual, wind, altimetry/geoid model DEM, any significant external events impacting observations.	Associated uncertainties and evidence where appropriate. Otherwise performance flags and parameters, e.g., drift. Sensitivity coefficients for L1 and L2 data to their parameters. Date range for auxiliary file version. (Some missions may include some of the parameters as a part of the L1 or L2+ data files.)
MO_1.9	ALL	Doc/data records	Calibration and validation data	Description of validation process, including identification of validation datasets, measurement protocols, data collection, analysis and accuracy reporting. Cal/Val (CALibration/VALidation) data acquired during mission operations (optical/radiometric stability, instrument availability, internal calibration, optic pointing pattern, etc.). Detailed history of validation activities and validation datasets along with metadata from previous validation exercises. Published data validation papers showing how well the data compare to the best-known correlative measurements. Accuracy of products, as measured by validation testing, and compared to accuracy requirements.	In-flight calibration reports, uncertainly with evidence, version report, instrument anomalies parameters evolution (degradation model, pixel response linearity, etc.). Instrument validation: Signal-to-Noise Ratio validation, absolute and relative radiometric vicarious calibration, modulation transfer function, geolocation, L2 products, etc. Validation reports, satellite uncertainties.

Table 4 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MO_1.10	ALL	Doc/ data records	Quality parameters	Quality assessment of instrument, raw data and products. Flowed-through effects of sensor noise, calibration errors, spatial and spectral errors, and/ or un-modelled or neglected geophysical phenomena on the quality of product	Assessment of performance/ acceptability.
MO_1.11	ALL	Doc/ Data Records	Metadata	Metadata digital inventory. Information about data to facilitate discovery, search, access and use associated with each of the data products.	
MO_1.12	SAT	SW Code	Level 0 consolidation	Software used for generating level 0 data from raw data.	
MO_1.13	SAT, AIR	Doc/ SW Code	Data processing software	Instrument processing algorithms, context and source codes, testing context. Description of lineage, i.e., input data and attributes covering all input data used by the algorithm - primary sensor data, ancillary data, forward models and look-up tables. Lineage information at granule level, i.e., all inputs (including ancillary or other data granules, calibration files, look-up tables, ground control, climatology etc.) used to generate the product. All information needed to verify what output data was created by a run, including data volume and file sizes. Documentation of expected exceptions, and how they are identified, trapped, and handled. Source for values of constants and look-up tables used in the algorithm, or explanation of how they were calculated. Documentation of processing history and production version history (which versions were used when, why different versions came about and what the improvements and changes were from version to version). Descriptions of datasets used for software verification and validation.	Algorithm description. Algorithms and software verification/validation, version control.

Table 4 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
				Test reports or summary of test results in sufficient detail to show that products meet requirements. Software release notes, including references to versions of operating systems, compilers, commercial or other software libraries used in the code.	
MO_1.13 (Continued)				Description of potential future enhancements to the algorithm, the limitations they will mitigate. For all products held in the archive, the versions of source code used to produce the products. Where different versions of ancillary, input data, or calibration were used, the history of those changes should be available as part of the processing history.	
MO_1.14	SAT	SW Code	Quality control software	Software used for quality assessment and developing quality indicators and/or quality flags.	Algorithms and software verification/validation, version control.
MO_1.15A	SAT	SW Code	Science data tools	Product access (reader) and analysis tools. Source code to facilitate use of the calibration data, ancillary data and the data products at all levels. Source code useful for creating programmes to read and display the calibration data, ancillary data and product data and metadata values. References to applicable dependency tools and libraries, and version numbers. Release notes, sample inputs and corresponding output results.	
MO_1.15B	SAT	SW Code	Visualization tools	Processing and visualizing tools.	Software validation and version control. Algorithms and software verification/ validation.
MO_1.16	SAT	SW Code	Value-added software	Software used for developing value-added products from the basic products (see MO_1.5A and MO_1.5B).	Software validation and version control. Algorithms and software verification/validation.

Table 4 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
MO_1.17	SAT	Doc	Product qualification and quality assurance monitoring reports	Defines the product qualification process outputs. Includes test reports, and results of reviews and appraisals.	Assessment of performance/accept ability based on relevant quality parameters such as uncertainty levels, flags etc.
MO_1.18	ALL	Doc	Sensor/instrument evolution and history records	Describes any instrument event that might affect data quality (e.g. upgrading, downgrading, look-up tables). It includes also known errors and limits of sensors/instruments. This is essential, especially for airborne and in-situ observations where sensors are updated or replaced during a mission. All calibration and quality assessment results/processes/reports should be explicitly connected to each instance of sensors/instruments.	Instrument timeline. Documented supporting evidence for decisions.
MO_1.19	ALL	Doc	Referred publications and papers	Referred publications, articles and technical notes clearly referencing the used datasets. References to published articles about the use of the data. User feedback about products.	Persistent identifiers.
MO_1.20	ALL	Doc	Tandem and/or combined campaigns, comparisons	Data and reports.	Uncertainty budgets with supporting evidence. Comparisons report following QA4EO Guideline 7[12]
MO_1.21	ALL	Doc/ data records	Cross- campaign, cross-comparisons and cross-calibration activities documentation and data	Describes the cross-campaign scenario and operational context. Also describes any cross-calibration activities.	Evidence of participation in appropriate comparisons. Comparison report following QA4EO Guidelines 4 and 7[12].
MO_1.22	ALL	Doc	Data access policy	Describes the data access policy for mission in the operational stage.	

7.6 Post mission stage

7.6.1 Rationale

After the end of a mission, datasets acquired during the operational stage shall be consolidated and aligned to the latest available version of the processors and/or improved version. All the evolution activities carried out in the previous stages and the changes to the data and associated information are properly assessed and consolidated during this stage for end-to-end consistency/coherency/provenance based on the documentation produced and preserved in the previous stages. During this stage the user communities will still need to analyse and process data. Enhanced algorithms and processor improvements could be implemented to improve data exploitation and processing performance.

7.6.2 Content

The content required to be preserved by the end of the post mission stage is identified and described in Table 5.

Table 5 — Post mission stage preservation content

ID	Need for	Type	Identification	Description	Examples of types of quality information
PM_1.0	ALL	Doc	Preservation metadata	List of items preserved at this stage to satisfy content requirements specified below. Updated version of MO_1.0	Identification of items by title and persistent ID. Access method (e.g., link). Indication of access restrictions if any as well as whether and when the restrictions will be removed.
PM_1.1A	SAT	Doc	Data consolidation & reprocessing strategy, implementation plans and consolidated/ reprocessed data. Processing.	Description of final processing and/or calibration changes including provenance and context.	Algorithms and software validation, version control. Clear description of motivation for reprocessing and improvements gained.
PM_1.1B	SAT	Doc/data records	Data consolidation & reprocessing strategy, implementation plans and consolidated/ reprocessed data. Ancillary, auxiliary data.	Updated ancillary, auxiliary data and their description. Name and location of the ancillary/auxiliary data archive facility if ancillary/ auxiliary data will not be stored with the products.	Associated uncertainties and evidence, version control.
PM_1.1C	SAT	Doc/data records	Data consolidation & reprocessing strategy, implementation plans and consolidated/ reprocessed data. QA	Quality information updated as part of reprocessing.	Assessment of performance/acceptability.
PM_1.2	ALL	Data records (Reprocessed dataset)	Data consolidation & reprocessing strategy, implementation plans, and consolidated/reprocessed data. L0, L1, L2+	Reprocessed data & products. The final version of a derived product should be the version archived. If results reported in peer reviewed publications were based on earlier versions of the product, those versions or at least representative subsets of those versions should also be archived. At a minimum, the algorithm and software that generated such earlier versions should be archived.	Associated uncertainties and evidence, version control.

Table 5 *(continued)*

ID	Need for	Type	Identification	Description	Examples of types of quality information
PM_1.3	ALL	Data	Data consolidation and reprocessing strategy, implementation plans and consolidated/reprocessed data. Metadata.	Metadata inventory.	
PM_1.4	ALL	Doc	Referred publications and papers.	Referred publications, articles and technical notes clearly referencing the used datasets.	Persistent Identifiers.
PM_1.5	SAT	Doc	Historical data access policy.	Describes the data access policy for the historical mission in the preservation stage.	
PM_1.6	SAT	Doc	Historical mission user handbook	Describes the consolidated end-to-end mission description, data formats, operational scenarios, and all information necessary for future data use. It also includes the appraisal of the mission datasets (i.e. their value).	Summary of quality information approach within mission/instrument.

Annex A
(normative)

Abstract test suite

A.1 Semantics

Conformance to this document consists of satisfying the requirements listed in Tables 1 through 5. Conformance can be demonstrated by using checklists naming the preserved documents and/or data records that map to the requirements in Tables 1 through 5. Correspondingly, the abstract test suite has five conformance classes:

a) Mission concept stage

b) Mission definition stage

c) Mission implementation stage

d) Mission operations stage

e) Post mission stage

A.2 Mission concept stage

a) Test purpose: Determine if all requirements for gathering preservation content from the mission concept stage are met.

b) Test method: Inspect documents and/or data records gathered at the end of the stage.

c) Reference: Table 1.

A.3 Mission definition stage

a) Test purpose: Determine if all requirements for gathering preservation content from the mission definition stage are met.

b) Test method: Inspect documents and/or data records gathered at the end of the stage.

c) Reference: Table 2.

A.4 Mission implementation stage

a) Test purpose: Determine if all requirements for gathering preservation content from the mission implementation stage are met.

b) Test method: Inspect documents and/or data records gathered at the end of the stage.

c) Reference: Table 3.

A.5 Mission operations stage

a) Test purpose: Determine if all requirements for gathering preservation content from the mission operations stage are met.

b) Test method: Inspect documents and/or data records gathered at the end of the stage.

c) Reference: Table 4.

A.6 Post mission stage

a) Test purpose: Determine if all requirements for gathering preservation content from the post mission stage are met.

b) Test method: Inspect documents and/or data records gathered at the end of the stage.

c) Reference: Table 5.

Annex B
(informative)

Stages and phases

This annex provides a mapping from the mission stages defined in Clause 6 to the phases defined by NASA and ESA in relation to their satellite missions.

Table B.1 — Stages and phases

Stage	Phase
Mission concept	A
Mission definition	B
Mission implementation	C and D
Mission operations	E
Post mission	F

Annex C
(informative)

XML representation for ISO 19165-2 (this document)

C.1 Introduction

ISO 19165-1 describes a complete model for mapping ISO 19115-1 DS_Dataset and DS_Series into the OAIS preservation model and for extending ISO 19115-1 MD_Metadata with preservation information in the GP_PreservationMetadata class (ISO 19165-1:2018, Figures 2 and 3). As indicated in the Introduction to this document, ISO 19165-1:2018, 7.3.1 indicates that specific content items needed to preserve the full provenance and context of the data and associated data depend on the needs of the designated communities and types of datasets (e.g., maps, remotely sensed data from satellites and airborne instruments, physical samples). It also states that follow-up parts to ISO 19165-1 may be developed to provide details of content items appropriate to specific disciplines. This document identifies the content that should be created and managed in order to ensure preservation of data and information from Earth observing missions (remote sensing with airborne and spaceborne instruments, and in situ measurements such as those from field campaigns). This annex describes how that information can be organized in an ISO 19115-1 DS_Series that holds references to all preservation information for a mission.

While the classes GP_Usage, GP_ValueAndPreservationStrategy, GP_Acquisition, GP_AssociatedResource, and GP_PackagingInformation defined in ISO 19165-1:2018, Figure 3 under the class GP_PreservationMetadata can be used as required in meeting the content preservation recommendations detailed in this document (ISO 19165-2), the primary focus of this document is on the content specific to Earth observation missions. This can be organized in an ISO 19115-1 DS_Series that holds references to all preservation information for a mission.

C.2 Structure of the mission metadata

The mission metadata is held as a group of DS_Initiative sections contained within a DS_Series object. Each of the DS_Initiative classes corresponds to a specific stage defined in this document and each stage includes an MD_Metadata record that contains the metadata for that stage.

The overall structure is shown in Figure C.1 and described below.