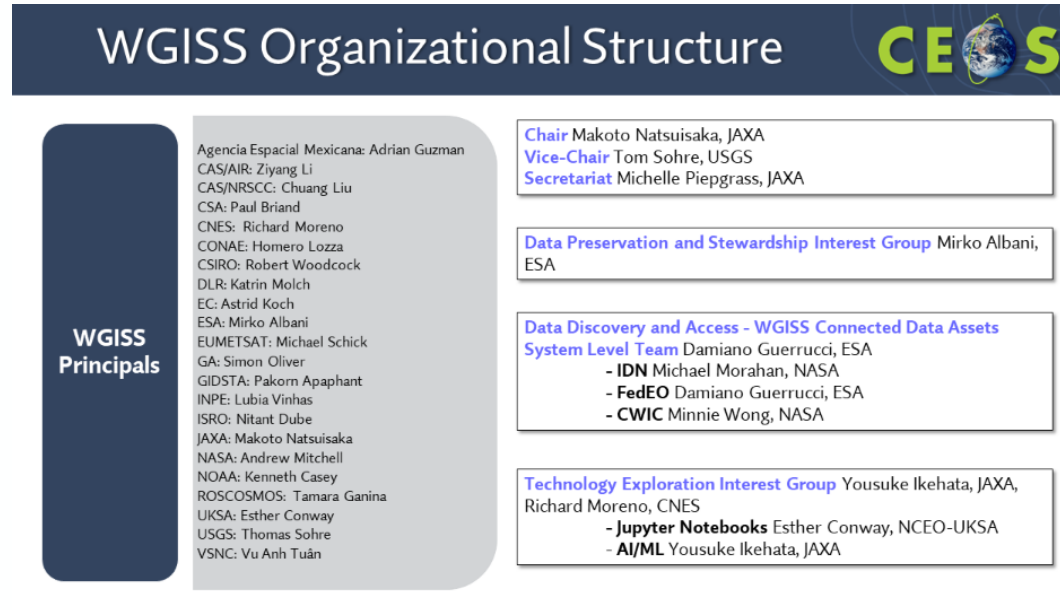# CEOS Data Management and Stewardship Maturity Matrix in support to Earth Observation data preservation and curation

I. Maggio (Rhea for ESA) PV 2023 May 2–4, 2023 - CERN

→ THE EUROPEAN SPACE AGENCY

WGISS Background and Scope

WGISS (The Working Group on Information Systems and Services) is a subsidiary body supporting CEOS.

- ✓ Promotes collaboration in the development of systems and services that manage and supply Earth Observation data;
- ✓ Creates and demonstrates prototypes supporting CEOS and Group on Earth Observation (GEO) requirements;
- ✓ Addresses the internal management of EO data, the creation of information systems and the delivery of interoperable services.

The activities and expertise of WGISS span the full range of the information life cycle from the requirements and metadata definition for the initial ingestion of satellite data into archives through to the incorporation of derived information into end-user applications.

https://ceos.org/ourwork/workinggroups/wgiss/preservation/

→ THE EUROPEAN SPACE AGENCY

DSIG Background and Scope

- Enable the sharing of agency investigations, developments, experiences and lessons learned relating to EO data stewardship.
- Draft common cross-agency best practices or guidelines of data stewardship for possible adoption by WGISS.
- Sponsor technical exchanges and the conduction of Joint Activities and/or Pilot Projects on specific data stewardship topics.
- Establish and maintain a CEOS "Data Purge Alert" service.
- Contribute to GEO and Standardization activities.
- Activities focus on EO Data, Metadata, and Associated Information.
- Long-term archiving approaches, systems and media.
- Data Formats and Standards.
- Preservation Lifecycle concepts.
- Data Valorization and Curation.

Three questions come to light…

What is the Maturity Matrix/Model?

Who could use it?

Why it should be used?

Maturity Matrix/Model - WHAT

All activities needed to preserve and improve the information content, quality, accessibility, and usability of data and metadata.

Maturity models/matrices are used to measure "levels of maturity" addressing the needs of specific domains. Examples:

- Capability Maturity Model Integration (CMMI)
- Levels of Maturity of Digital Repositories (e.g. ISO 16363)
- Climate Data Record Maturity Matrix (CDRMM)
- ESA TECHNOLOGY READINESS LEVELS (TRLs)
- ESA Scientific Readiness Levels (SRL)

Maturity Matrix for Long-Term Scientific Data Stewardship (2015, Ge Peng and Jeffrey L. Privette) covers the full scientific data lifecycle

Maturity Matrix/Model - WHO

Data providers

➤ to evaluate and improve the quality and usability of their products

Modelers, decision-makers, and scientists

➤ to improve their products
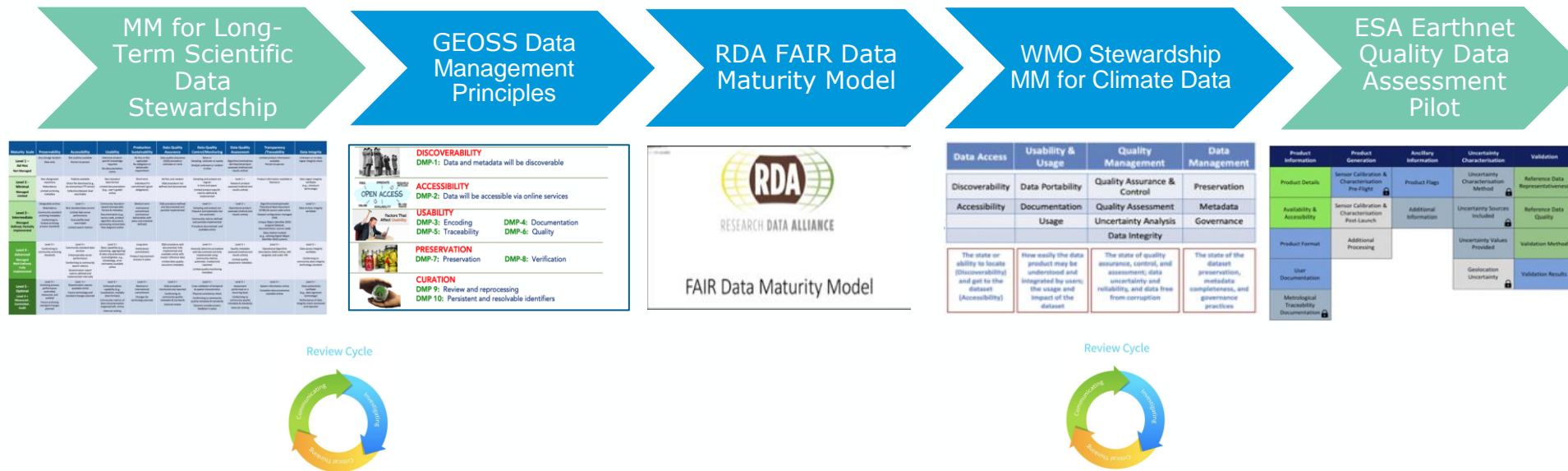
➤ to make investment and use decision

Data managers/stewards of data centers and repositories

➤ to validate their compliance or lack of stewardship practice or standards

➤ to assess the current state

➤ to create a roadmap forward to improve or enhance its stewardship maturity of practices applied to all its holdings

Maturity Matrix/Model - WHY

✓Provides data quality, usability information to users, stakeholders, and decision makers;

✓A reference model for stewardship planning and resource allocation;

✓Creates a roadmap for scientific data stewardship improvement;

✓Provides detailed guideline and recommendations for preservation;

✓Evaluates if the preservation follows best practices;

✓Gives a technical evaluation of the level of preservation and helps with self assessment of preservation;

✓Gives no numbers or average but a status;

✓Helps to break the problem down, and understand the costs associated with each;

✓Funding agencies can define goal levels;

✓Flexible and adaptable after a tailoring.

..… This is enough, isn't it???

CEOS WGISS DMSMM: Generation process

MM for Long-Term Scientific Data Stewardship → GEOSS Data Management Principles → RDA FAIR Data Maturity Model → WMO Stewardship MM for Climate Data → ESA Earthnet Quality Data Assessment Pilot

DMSMM defines all activities needed to preserve and improve the information content, quality, accessibility, and usability of data and metadata.

Data stewardship "encompasses all activities that preserve and improve the information content, accessibility, and usability of data and metadata" (National Research Council 2007).

Data management includes all activities for "planning, execution and oversight of policies, practices and projects that acquire, control, protect, deliver and enhance the value of data and information assets." (Mosely et al. 2009).

→ THE EUROPEAN SPACE AGENCY

CEOS WGISS DMSMM

Highlighted callout — PRESERVATION

MMP9 — Data Preservation	MMP10 — Data Verification
1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check
1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

Full maturity matrix

	DISCOVERABILITY — MMP1 Metadata for Discovery	ACCESSIBILITY — MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceability	MMP8 Data Quality Control	PRESERVATION — MMP9 Data Preservation	MMP10 Data Verification	CURATION — MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable Identifier
Level-0 Not Managed	1) No catalogue available 2) No advertising available	1) Data and metadata are not accessible online	1) Data Not Structured 2) Non-standard or proprietary data format, or, poorly-documented standard file format.	1) Partial and incomplete mission documentation	1) Limited product information avail... (not online)	1) No control and monitoring check 2) No quality indicator in metadata 3) No procedures documentation	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	1) No Data/Associated Information integrity, authenticity and readability check	1) Calibration Algorithm - Calibration algorithm not documented. 2) Geometric Processing - Geometric processing algorithm not documented. 3) Retrieval algorithm not documented. 4) Mission Specific Processing – Additional processing steps not documented.	1) No persistent and resolvable identifiers available
Level-1 Partially Managed	1) Advertising available 2) Catalogue search available at product level	1) Basic online services available for data and metadata access	1) Basic schema for automated data use 2) Data in documented standard file format. Non-standard naming conventions used.	1) Already existent mission documentation available and preserved for the long term 2) No link between mission documentation and data records	1) Product information available (not onl...)	1) Basic data quality control and monitoring check 2) Minimal set of quality control procedures documented and available	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	1) Data Records/Associated Information integrity basic check	1) Calibration Algorithm - Calibration algorithm somewhat documented. Calibration algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance. 2) Geometric Processing - Geometric processing documented. Missing all or part of the calibration parameters. Calibration algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance. Confidence in the calibration quality is minimal. 3) Retrieval algorithm somewhat documented. Retrieval algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance. 4) Mission Specific Processing – Additional processing steps documented. Additional processing steps not considered fit for stated purpose.	1) Persistent identifier assignment only for particular Data Records Collections 2) Basic landing pages management
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product metadata oriented towards an international standard 3) Data Collection and Associated Information searchable. 4) International standard for Collection metadata	1) Simple Access Architecture through metadata 2) Data access system oriented towards an international standard	1) Use of non-proprietary international standards encodings for syntactic interoperability. 2) Periodically repackaging/reformatting of archived data. 3) Data in well-documented standard file format, community naming convention standards.	1) Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	1) Dataset tested presence of corre... provenance meta... 2) Well described product informati... available online	1) Quality indicator post-processing available 2) Quality control procedures documented and available online	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Calibration Algorithm - Calibration algorithm documented. Calibration used "fit for purpose" in terms of the mission's stated performance all expected use cases. 2) Geometric Processing - Geometric processing documented. All input calibration parameters exist. Methodology used is considered "fit for purpose" in terms of the mission's stated performance for all expected use cases. Quality flags indicate good geometric accuracy with less than 5% exceptional. 3) Retrieval algorithm documented. Retrieval algorithm "fit for purpose" in terms of the mission's stated performance all expected use cases and validated performance against similar algorithms or with empirical evidence. 4) Mission Specific Processing – Additional processing steps documented. All additional processes steps considered fit for stated purpose.	1) Persistent identifier assignment to all disseminated Data Records Collections and metadata 2) Automatic landing page generation and extensive management of landing pages
Level-3 Fully Managed	1) Catalogue accessible via international or community agreed standards protocol 2) Data policy available in metadata 3) Periodic updates of metadata in the catalogue 4) Quality indicator metadata available and discoverable 5) Search results relevancy. 6) Seamless transition from discovery to access	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 4) Reporting system 5) Hosted processing 6) Quick adoption to new technologies and standards evolution 7) Data and metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadata uses FAIR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	1) Automatic meta... generation for provenance documentation 2) Complete and u... data provenance available online	1) Data quality control fully compliant with an international standard 2) Quality indicator pre and post processing available in the metadata 3) Quality metadata assessed	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting	1) Calibration Algorithm - Calibration algorithm well-documented. State-of-the-art calibration algorithm applied and considered "fit for purpose" in terms of the mission's stated performance. 2) Geometric Processing - Geometric processing well-documented. State-of-the-art methodology used, easily "fit for purpose" in terms of the mission's stated performance. Quality flags indicate excellent geometric accuracy. 3) Retrieval algorithm documented. State-of-the-art retrieval "fit for purpose" in terms of the mission's stated performance, full uncertainty budget derived and validated. 4) Mission Specific Processing – Additional processing steps documented. All additional processes steps considered fit for stated purpose.	1) Persistent identifier created for all accessible data records and metadata 2) Metadata includes the identifier for the data 3) Metadata is offered in such a way that it can be harvested and indexed

→ THE EUROPEAN SPACE AGENCY

Data Management & Stewardship Maturity Matrix - Scope

	DISCOVERABILITY	ACCESSIBILITY	USABILITY						PRESERVATION		CURATION	
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceability	MMP6 Data Validation	MMP7 Data Uncertainty	MMP8 Data Quality Control	MMP9 Data Preservation	MMP10 Data Verification	MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable Identifier
Level-0 Not Managed	1) No catalogue available 2) No advertising available	Data and metadata are not accessible online	1) Data Not Structured 2) Non-standard or proprietary data format, or, poorly-documented standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	1) Reference Data Representativeness - No validation 2) Reference Data Quality - No validation 3) Validation Method - No validation 4) Validation Results - No validation	1) Uncertainty Method: Uncertainty characterisation not performed, or method not documented. 2) Uncertainty Sources: Uncertainty characterisation not performed, or sources analysed not documented. 3) Uncertainty Values: No uncertainty information provided.	1) No control and monitoring check 2) No quality indicator in metadata 3) No procedures documentation	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available.	No Data/Associated Information integrity, authenticity and readability check	1) No reprocessing activities planned 2) Pre-flight calibration & characterisation not documented or information not available. 3) Post-launch calibration & characterisation not documented or not available. 4) Processing: Additional processing steps not documented.	No persistent and resolvable identifiers available
Level-1 Partially Managed	1) Advertising available 2) Catalogue search available at product level	Basic online services available for data and metadata access	1) Basic schema for automated data use 2) Data in documented standard file format. Non-standard naming conventions used.	1) Already existent mission documentation available and preserved for the long term 2) No link between documentation and data records	Product information available (not online)	1) Reference Data Representativeness: measurements assessed to be mostly representative of the satellite measurements 2) Reference Data Quality: single uncertainty for the entire dataset. 3) Validation Method: simple uncertainty estimated 4) Validation Results: Validation results show good agreement between satellite and reference measurements within uncertainties in most cases.	1) Uncertainty Method: Limited use of GUM approach, and/or, an expanded comparison to measurements by other sensors. 2) Uncertainty Sources: Most important sources of uncertainty included. 3) Uncertainty Values: Single uncertainty value provided for subsets of data	1) Basic data quality control and monitoring check 2) Minimal set of quality control procedures documented and available	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	1) Minor updates and bugs corrections of data records implemented 2) Data Records repackaging and/or reformatting 3) Pre-flight calibration & characterisation misses some important aspects 4) Post-launch calibration & characterisation misses some important aspects of instrument behaviour and/or is not entirely of a level of quality to be judged fit for purpose. 5) Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.	1) Persistent identifier assignment only for particular Data Records Collections 2) Basic landing pages management
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product metadata oriented towards an international standard 3) Data Collection and Associated Information searchable. 4) International standard for Collection metadata	1) Simple Access Architecture through metadata 2) Data access system oriented towards an international standard	1) Use of non-proprietary international standards encodings for syntactic interoperability. 2) Periodically repackaging/ reformatting of archived data. 3) Data in well-documented standard file format, community naming convention standards.	1) Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	Dataset tested for presence of correct provenance metadata. Well described product information available online	1) Reference Data Representativeness: measurements assessed to be well representative of the satellite measurements 2) Reference Data Quality: full uncertainty information 3) Validation Methods assess satellite measurements 4) Validation Results show excellent agreement between satellite and reference measurements, within uncertainties.	1) Uncertainty Method: GUM approach to estimate measurement uncertainty with full breakdown of components and separated as Type A or B classification. 2) Uncertainty Sources: All important sources of uncertainty included. 3) Uncertainty Values: Total uncertainty per pixel is provided, with basic breakdown of key components no error-covariance.	1) Quality indicator post-processing available 2) Quality control procedures documented and available online	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Reprocessing for calibration and/or algorithm improvement 2) Pre-flight calibration & characterisation covers all reasonable aspects 3) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance and uses appropriate community infrastructure/methods (CEOS/FRMs). 4) Additional processing steps documented.	1) Persistent identifier assignment to all disseminated Data Records Collections and metadata 2) Automatic landing page generation and extensive management of landing pages
Level-3 Fully Managed	1) International standard for Product metadata 2) International standard for Collection metadata 3) Catalogue accessible via international or community agreed standards protocol 4) Data policy available in metadata 5) Periodic updates of metadata in the catalogue 6) Quality indicator metadata available and discoverable 7) Search results relevancy. 8) Seamless transition	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 4) Reporting system 5) Hosted processing 6) Quick adoption to new technologies and standards evolution 7) Data and metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadata uses FAIR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	1) Automatic metadata generation for provenance documentation 2) Complete and updated data provenance available online	1) Reference Data Representativeness: Reference measurements independently assessed to be fully representative of the satellite measurements, covering the satellite's full range of measurements and with full assessment of uncertainties and carried out on a regular basis determined by product performance. 2) Reference Data Quality: full uncertainty and error-correlation information, assessed following the GUM and traceable to SI 3) Validation Methods assess satellite measurements and reference data w.r.t. their error- covariance and validates those uncertainties. 4) Validation Results show excellent agreement between satellite and reference measurements, within	1) Uncertainty Method: GUM approach to estimate measurement uncertainty, including a treatment of error-covariance. 2) Uncertainty Sources: All reasonable sources of uncertainty included. 3) Uncertainty Values: Uncertainties per pixel provided with error-covariance information for all appropriate components.	1) Data quality control fully compliant with an international standard 2) Quality indicator pre and post processing available in the metadata 3) Quality metadata assessed	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process including monitoring and reporting	1) Reprocessing for time-series creation 2) Roadmap for technology evolution 3) Plurality of accurate and relevant attributes are provided to allow reuse 4) Metadata includes information about the licence 5) Pre-Flight: As Level-2, additionally calibration and characterisation includes the measurements needed to assess uncertainties at component level and their impact on the final product. 6) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance. 7) All additional processing steps fully documented and state-of-the-art.	1) Persistent identifier created for all accessible data records and metadata 2) Metadata includes the identifier for the data 3) Metadata is offered in such a way that it can be harvested and indexed

- A way to measure the status of the Agency Data Stewardship processes in place

- A way to plan goals of Data Stewardship processes and projects

CEOS WGISS DMSMM pillars

5 Areas

- Discoverability
- Accessibility
- Usability
- Preservation
- Curation

4 Level of Maturity

- L0 Not Managed
- L1 Partially Managed
- L2 Managed
- L3 Fully Managed

12 Components

- Metadata for Discovery
- Online Access
- Data encoding
- Data Documentation
- Data Traceability
- Data Validation
- Data Metrology (e.g. Uncertainty)
- Data Quality Control
- Product Details
- Data Preservation
- Data Verification
- Data Processing/Reprocessing
- Persistent & Resolvable Identifier

CEOS WGISS DMSMM: Implementation

	DISCOVERABILITY	ACCESSIBILITY	USABILITY						PRESERVATION		CURATION	
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceability	MMP6 Data Validation	MMP7 Data Uncertainty	MMP8 Data Quality Control	MMP9 Data Preservation	MMP10 Data Verification	MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable Identifier
Level-0 Not Managed	1) No catalogue available 2) No advertising available	Data and metadata are not accessible online	1) Data Not Structured 2) Non-standard or proprietary data format, or, poorly-documented standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	1) Reference Data Representativeness - No validation 2) Reference Data Quality - No validation 3) Validation Method - No validation 4) Validation Results - No validation	1) Uncertainty Method: Uncertainty characterisation not performed, or method not documented. 2) Uncertainty Sources: Uncertainty characterisation not performed, or sources analysed not documented. 3) Uncertainty Values: No uncertainty information provided.	1) No control and monitoring check 2) No quality indicator in metadata 3) No procedures documentation	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available.	No Data/Associated Information integrity, authenticity and readability check	1) No reprocessing activities planned 2) Pre-flight calibration & characterisation not documented or information not available. 3) Post-launch calibration & characterisation not documented or not available. 4) Processing: Additional processing steps not documented.	No persistent and resolvable identifiers available
Level-1 Partially Managed	1) Advertising available 2) Catalogue search available at product level	Basic online services available for data and metadata access	1) Basic schema for automated data use 2) Data in documented standard file format. Non-standard naming conventions used.	1) Already existent mission documentation available and preserved for the long term 2) No link between mission documentation and data records	Product information available (not online)	1) Reference Data Representativeness: measurements assessed to be mostly representative of the satellite measurements 2) Reference Data Quality: single uncertainty for the entire dataset. 3) Validation Method: simple uncertainty estimated 4) Validation Results: Validation results show good agreement between satellite and reference measurements within uncertainties in most cases.	1) Uncertainty Method: Limited use of GUM approach, and/or, an expanded comparison to measurements by other sensors. 2) Uncertainty Sources: Most important sources of uncertainty included. 3) Uncertainty Values: Single uncertainty value provided for subsets of data	1) Basic data quality control and monitoring check 2) Minimal set of quality control procedures documented and available	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	1) Minor updates and bugs corrections of data records implemented 2) Data Records repackaging and/or reformatting 3) Pre-flight calibration & characterisation misses some important aspects 4) Post-launch calibration & characterisation misses some important aspects of instrument behaviour and/or is not entirely of a level of quality to be judged fit for purpose. 5) Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.	1) Persistent identifier assignment only for particular Data Records Collections 2) Basic landing pages management
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product metadata oriented towards an international standard 3) Data Collection and Associated Information searchable. 4) International standard for Collection metadata	1) Simple Access Architecture through metadata 2) Data access system oriented towards an international standard	1) Use of non-proprietary international standards encodings for syntactic interoperability. 2) Periodically repackaging/reformatting of archived data. 3) Data in well-documented standard file format, community naming convention standards.	1) Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	Dataset tested for presence of correct provenance metadata. Well described product information available online	1) Reference Data Representativeness: measurements assessed to be well representative of the satellite measurements 2) Reference Data Quality: full uncertainty information 3) Validation Methods assess satellite measurements 4) Validation Results show excellent agreement between satellite and reference measurements, within uncertainties.	1) Uncertainty Method: GUM approach to estimate measurement uncertainty with full breakdown of components and separated as Type A or B classification. 2) Uncertainty Sources: All important sources of uncertainty included. 3) Uncertainty Values: Total uncertainty per pixel is provided, with basic breakdown of key components no error-covariance.	1) Quality indicator post-processing available 2) Quality control procedures documented and available online	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Reprocessing for calibration and/or algorithm improvement 2) Pre-flight calibration & characterisation covers all reasonable aspects 3) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance and uses appropriate community infrastructure/methods (CEOS/FRMs). 4) Additional processing steps documented.	1) Persistent identifier assignment to all disseminated Data Records Collections and metadata 2) Automatic landing page generation and extensive management of landing pages
Level-3 Fully Managed	1) International standard for Product metadata 2) International standard for Collection metadata 3) Catalogue accessible via international or community agreed standards protocol 4) Data policy available in metadata 5) Periodic updates of metadata in the catalogue 6) Quality indicator metadata available and discoverable 7) Search results relevancy. 8) Seamless transition	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 4) Reporting system 5) Hosted processing 6) Quick adoption to new technologies and standards evolution 7) Data and metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadata uses FAIR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	1) Automatic metadata generation for provenance documentation 2) Complete and updated data provenance available online	1) Reference Data Representativeness: Reference measurements independently assessed to be fully representative of the satellite measurements, covering the satellite's full range of measurements and with full assessment of uncertainties and carried out on a regular basis determined by product performance. 2) Reference Data Quality: full uncertainty and error-correlation information, assessed following the GUM and traceable to SI 3) Validation Methods assess satellite measurements and reference data w.r.t. their error- covariance and validates those uncertainties. 4) Validation Results show excellent agreement between satellite and reference measurements, within	1) Uncertainty Method: GUM approach to estimate measurement uncertainty, including a treatment of error-covariance. 2) Uncertainty Sources: All reasonable sources of uncertainty included. 3) Uncertainty Values: Uncertainties per pixel provided with error-covariance information for all appropriate components.	1) Data quality control fully compliant with an international standard 2) Quality indicator pre and post processing available in the metadata 3) Quality metadata assessed	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting	1) Reprocessing for time-series creation 2) Roadmap for technology evolution 3) Plurality of accurate and relevant attributes are provided to allow reuse 4) Metadata includes information about the licence 5) Pre-Flight: As Level-2, additionally calibration and characterisation includes the measurements needed to assess uncertainties at component level and their impact on the final product. 6) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance. 7) All additional processing steps fully documented and state-of-the-art.	1) Persistent identifier created for all accessible data records and metadata 2) Metadata includes the identifier for the data 3) Metadata is offered in such a way that it can be harvested and indexed

COMPONENTS	L0	L1	L2	L3
Metadata for Discovery	★	★	★	★
Online Access	★	★	★	★
Data Encoding	★	★	★	★
Data Documentation	★	★	★	★
Data Traceability	★	★	★	★
Data Validation	★	★	★	★
Data Uncertainty	★	★	★	★
Data Quality Control	★	★	★	★
Data Preservation	★	★	★	★
Data Verification	★	★	★	★
Data Processing/Reprocessing	★	★	★	★
Persistent & Resolvable Identifier	★	★	★	★

Conclusions

- The goal of the DMSMM is to provide a holistic, consistent, quantifiable, and scalable measure of data stewardship maturity for end users and stakeholders including data providers and decision-makers.

- The DMSMM should be tailored for each organization and eventually specific dataset to properly take into account individual Data Stewardship and Data Management needs.

- The DMSMM helps data stewards and curators to get a consistent and quantifiable measure of an organisation's data holdings maturity.

- The ratings of DMSMM will help to validate compliance with applicable regulations on stewarding digital environmental geospatial data. The results can be used to identify potential areas for improvement and to create a roadmap for enhancing maturity of selected datasets in the identified areas by following community-accepted best practices.

- The evaluation of the DSMMM of a product can be used to build a stewardship cost model for planning purposes – based on the difference between the current maturity levels of key components and relevant stewardship requirements – prior to beginning the archive and data governance process.

- The DMSMM can be utilised by data providers or scientific stewards seeking to evaluate and improve the quality and usability of their products. The results can be also used by scientists to better understand the upstream data and data quality management practices applied to their input datasets.

Step by step in DMSMM

→ THE EUROPEAN SPACE AGENCY

Perform a verification for each task putting in green the tasks already implemented.

PRESERVATION (left table)

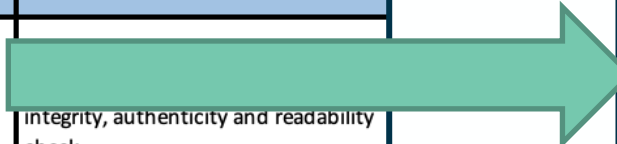|	MMP9 Data Preservation	MMP10 Data Verification
L0	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	integrity, authenticity and readability check
L1	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
L2	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
L3	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

PRESERVATION (right table)

	MMP9 Data Preservation	MMP10 Data Verification
	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check
	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

Perform this kind of verification for each task in the maturity level under analysis.

PRESERVATION	
MMP9 **Data Preservation**	**MMP10** **Data Verification**
1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	integrity, authenticity and readability check
1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

L0
L1
L2
L3

PRESERVATION	
MMP9 **Data Preservation**	**MMP10** **Data Verification**
1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check
1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

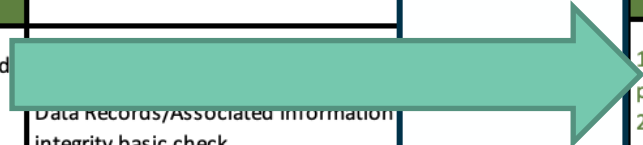When all tasks in the analysed maturity level are fully covered, the whole cell becomes green.

The analysis continues going on to the next level of maturity.

Left Table:

PRESERVATION	
MMP9 Data Preservation	**MMP10** Data Verification
L0: 1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check
L1: 1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
L2: 1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
L3: 1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

Right Table:

PRESERVATION	
MMP9 Data Preservation	**MMP10** Data Verification
L0: 1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check
L1: 1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
L2: 1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
L3: 1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

When a task is not already implemented it remains in black and at the end the cell, partially covered, becomes light green.

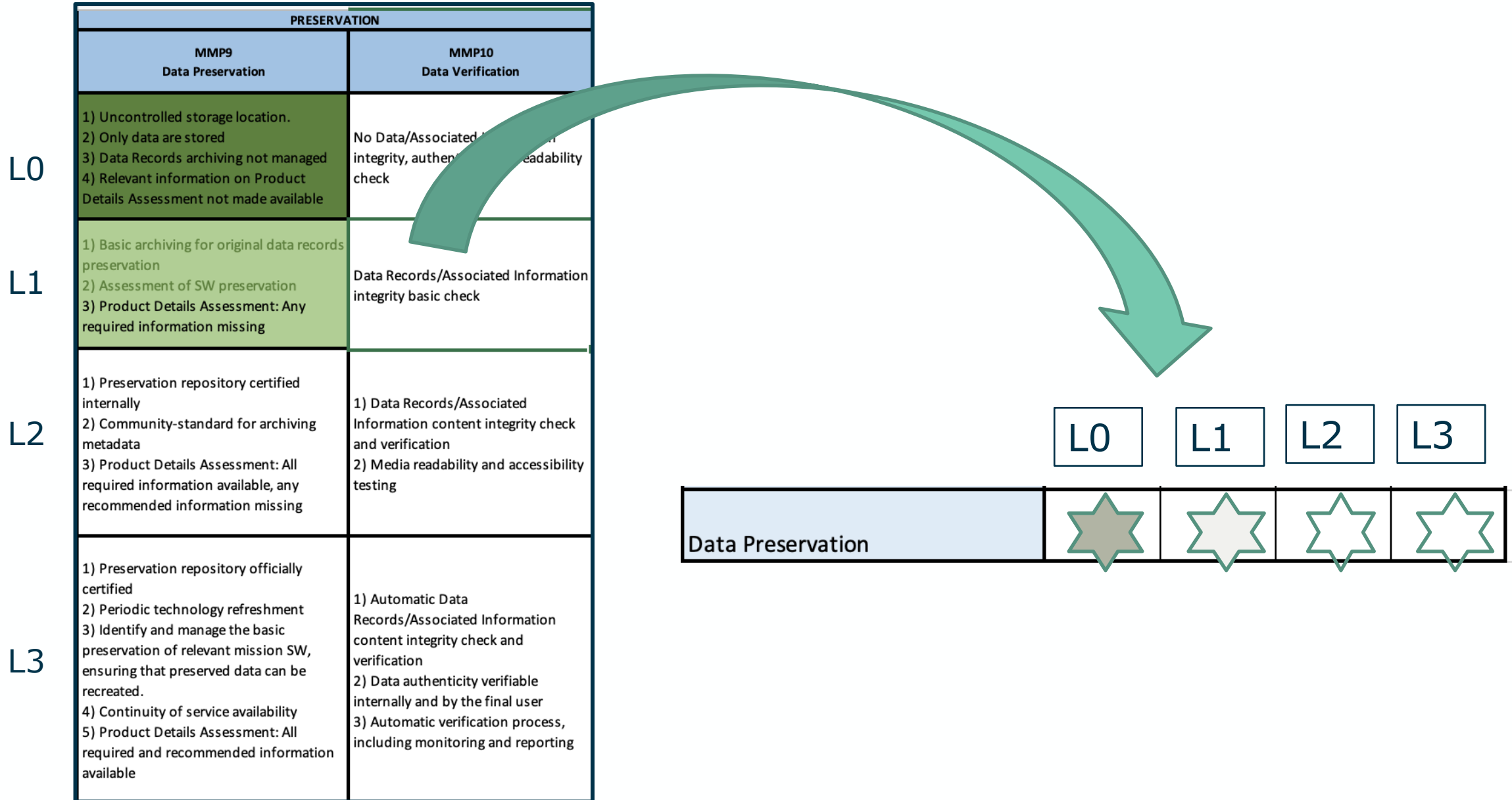| PRESERVATION	
MMP9 **Data Preservation**	**MMP10** **Data Verification**
L0 1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity, readability check
L1 1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check
L2 1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing
L3 1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting

L0 L1 L2 L3

Data Preservation

	DISCOVERABILITY	ACCESSIBILITY	USABILITY						PRESERVATION		CURATION	
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceability	MMP6 Data Validation	MMP7 Data Uncertainty	MMP8 Data Quality Control	MMP9 Data Preservation	MMP10 Data Verification	MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable Identifier
Level-0 Not Managed	1) No catalogue available 2) No advertising available	Data and metadata are not accessible online	1) Data Not Structured 2) Non-standard or proprietary data format, or, poorly-documented standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	1) Reference Data Representativeness - No validation 2) Reference Data Quality - No validation 3) Validation Method - No validation 4) Validation Results - No validation	1) Uncertainty Method: Uncertainty characterisation not performed, or method not documented. 2) Uncertainty Sources: Uncertainty characterisation not performed, or sources analysed not documented. 3) Uncertainty Values: No uncertainty information provided.	1) No control and monitoring check 2) No quality indicator in metadata 3) No procedures documentation	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated information integrity, authenticity and readability check	1) No reprocessing activities planned 2) Pre-flight calibration & characterisation not documented or information not available. 3) Post-launch calibration & characterisation not documented or not available. 4) Processing: Additional processing steps not documented.	No persistent and resolvable identifiers available
Level-1 Partially Managed	1) Advertising available 2) Catalogue search available at product level	Basic online services available for data and metadata access	1) Basic schema for automated data use 2) Data in documented standard file format. Non-standard naming conventions used.	1) Already existent mission documentation available and preserved for the long term 2) No link between mission documentation and data records	Product information available (not online)	1) Reference Data Representativeness: measurements assessed to be mostly representative of the satellite measurements 2) Reference Data Quality: single uncertainty for the entire dataset. 3) Validation Method: simple uncertainty estimated 4) Validation Results: Validation results show good agreement between satellite and reference measurements within uncertainties in most cases.	1) Uncertainty Method: Limited use of GUM approach, and/or, an expanded comparison to measurements by other sensors. 2) Uncertainty Sources: Most important sources of uncertainty included. 3) Uncertainty Values: Single uncertainty value provided for subsets of data	1) Basic data quality control and monitoring check 2) Minimal set of quality control procedures documented and available	1) Basic archiving for original data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	1) Minor updates and bugs corrections of data records implemented 2) Data Records repackaging and/or reformatting 3) Pre-flight calibration & characterisation misses some important aspects 4) Post-launch calibration & characterisation misses some important aspects of instrument behaviour and/or is not entirely of a level of quality to be judged fit for purpose. 5) Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.	1) Persistent identifier assignment only for particular Data Records 2) Basic landing pages management
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product metadata oriented towards an international standard 3) Data Collection and Associated Information searchable. 4) International standard for Collection metadata	1) Simple Access Architecture through metadata 2) Data access system oriented towards an international standard	1) Use of non-proprietary international standards encodings for syntactic interoperability. 2) Periodically repackaging/ reformatting of archived data. 3) Data in well-documented standard file format, community naming convention standards.	1) Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	Dataset tested for presence of correct provenance metadata. Well described product information available online	1) Reference Data Representativeness: measurements assessed to be well representative of the satellite measurements 2) Reference Data Quality: full uncertainty information 3) Validation Methods assess satellite measurements 4) Validation Results show excellent agreement between satellite and reference measurements, within uncertainties.	1) Uncertainty Method: GUM approach to estimate measurement uncertainty with full breakdown of components and separated as Type A or B classification. 2) Uncertainty Sources: All important sources of uncertainty included. 3) Uncertainty Values: Total uncertainty per pixel is provided, with basic breakdown of key components no error-covariance.	1) Quality indicator post-processing available 2) Quality control procedures documented and available online	1) Preservation repository certified internally 2) Community-standard for archiving metadata 3) Product Details Assessment: All required information available, any recommended information missing	1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Reprocessing for calibration and/or algorithm improvement 2) Pre-flight calibration & characterisation covers all reasonable aspects 3) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance and uses appropriate community infrastructure/methods (CEOS/FRMs). 4) Additional processing steps documented.	1) Persistent identifier assignment to all disseminated Data Records Collections and metadata 2) Automatic landing page generation and extensive management of landing pages
Level-3 Fully Managed	1) International standard for Product metadata 2) International standard for Collection metadata 3) Catalogue accessible via international or community agreed standards protocol 4) Data policy available in metadata 5) Periodic updates of metadata in the catalogue 6) Quality indicator metadata available and discoverable 7) Search results relevancy. 8) Seamless transition	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 4) Reporting system 5) Hosted processing 6) Quick adoption to new technologies and standards evolution 7) Data and metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadata uses FAIR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	1) Automatic metadata generation for provenance documentation 2) Complete and updated data provenance available online	1) Reference Data Representativeness: Reference measurements independently assessed to be fully representative of the satellite measurements, covering the satellite's full range of measurements and with full assessment of uncertainties and carried out on a regular basis determined by product performance. 2) Reference Data Quality: full uncertainty and error-correlation information, assessed following the GUM and traceable to SI 3) Validation Methods assess satellite measurements and reference data w.r.t. their error-covariance and validates those uncertainties. 4) Validation Results show excellent agreement between satellite and reference measurements, within	1) Uncertainty Method: GUM approach to estimate measurement uncertainty, including a treatment of error-covariance. 2) Uncertainty Sources: All reasonable sources of uncertainty included. 3) Uncertainty Values: Uncertainties per pixel provided with error-covariance information for all appropriate components.	1) Data quality control fully compliant with an international standard 2) Quality indicator pre and post processing 3) Continuity of service availability 4) Product Details Assessment: All required and recommended information available	1) Preservation repository officially certified 2) Periodic technology refreshment 3) Identify and manage the basic preservation of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability 5) Product Details Assessment: All required and recommended information available	1) Automatic Data Records/Associated Information content integrity check and verification 2) Data authenticity verifiable internally and by the final user 3) Automatic verification process, including monitoring and reporting	1) Reprocessing for time-series creation 2) Roadmap for technology evolution 3) Plurality of accurate and relevant attributes are provided to allow reuse 4) Metadata includes information about the licence 5) Pre-Flight: As Level-2, additionally calibration and characterisation includes the measurements needed to assess uncertainties at component level and their impact on the final product. 6) Post-launch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is "fit for purpose" in terms of the mission's stated performance. 7) All additional processing steps fully documented and state-of-the-art.	1) Persistent identifier created for all accessible data records and metadata 2) Metadata includes the identifier for the data 3) Metadata is offered in such a way that it can be harvested and indexed

COMPONENTS	L0	L1	L2	L3
Metadata for Discovery	★	★	★	☆
Online Access	★	★	★	★
Data Encoding	★	★	★	☆
Data Documentation	★	★	★	☆
Data Traceability	★	★	★	☆
Data Validation	★	★	★	★
Data Uncertainty	★	☆	☆	☆
Data Quality Control	★	★	★	★
Data Preservation	★	★	★	★
Data Verification	★	★	★	★
Data Processing/Reprocessing	★	★	★	☆
Persistent & Resolvable Identifier	★	★	★	★

→ THE EUROPEAN SPACE AGENCY