

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

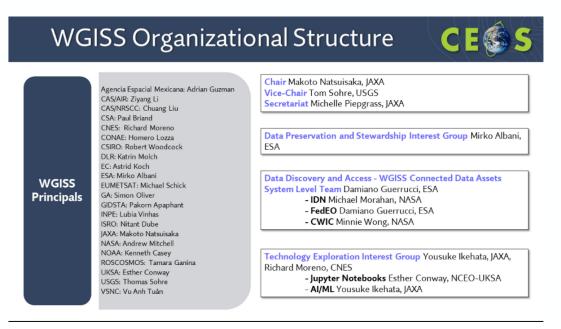
I. Maggio (Rhea for ESA) OLOS September 25, 2023



# WGISS Background and Scope



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



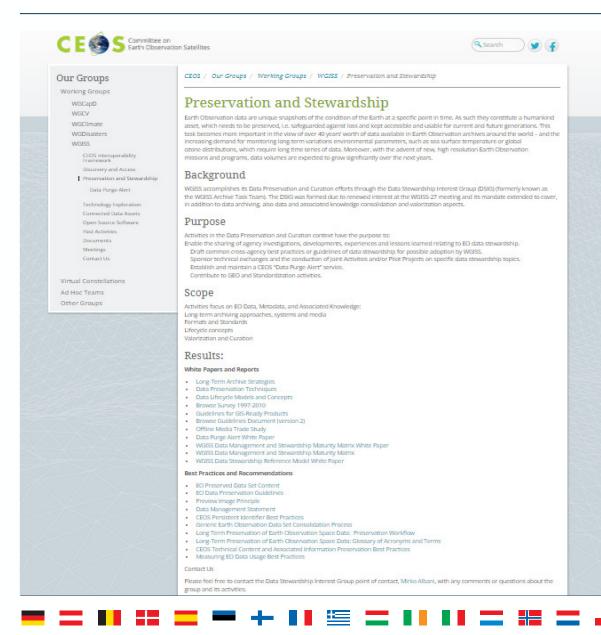
- ✓ Promotes collaboration in the development of systems and services that manage and supply Earth Observation data;
- ✓ Creates and demonstrates prototypes supporting CEOS and Group on Earth Observation (GEO) requirements;
- ✓ Addresses the internal management of EO data, the creation of information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

# **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.

# Data Management and Stewardship Maturity Matrix



Three questions come to light...

### What is the Maturity Matrix/Model?



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**

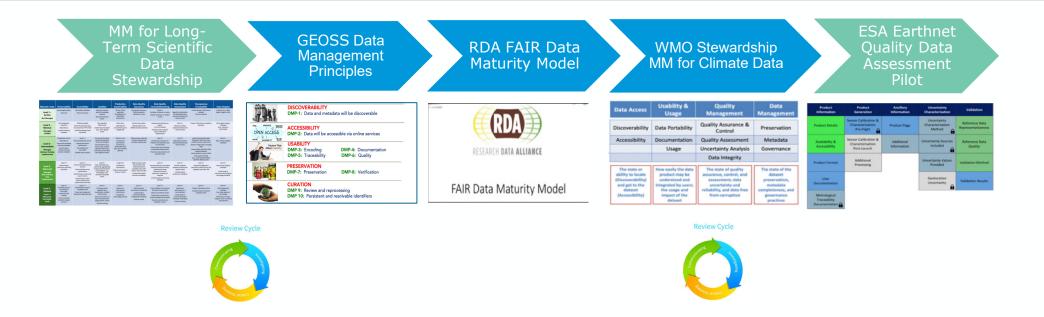


- ✓ Provides data quality, usability information to users, stakeholders, and decision makers;
- ✓ A reference model for stewardship planning and resource allocation;
- ✓ Creates a roadmap for scientific data stewardship improvement;
- ✓ Provides detailed guideline and recommendations for preservation;
- ✓ 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**





# 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).

## **CEOS WGISS DMSMM**



	DISCOVERABILITY	ACCESSIBILITY				PRESERVA	PRESERVATION CURATI			CURATION	PRATION	
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMP5 Data Traceabi			MMP8 Data Quality Control	MMP9 Data Preservation	MMP10 Data Verification	MMP11 Data Processing/Reprocessing	MMP12 Persistent & Resolvable
Level-0 Not Manag	1) No catalogue available     2) No advertising available		Data Not Structured     Non-standard or     proprietary data format, or,     poorly-documented     standard file format.	Partial and incomplete mission documentation	Limited produc information avail: (not online)	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed	MMP10 Data Verification  No Data/Associated Information integrity, authenticity and readability	No control and monitoring check     No quality indicator in metadata     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.	Identifier  1) No persistent and resolvable identifiers available
				1) Already existent		4) Relevant information on Product Details Assessment not made available	check		Basic archiving for original		Calibration Algorithm - Calibration algorithm somewhat documented.     Calibration algorithm too simple to be judged "fit for purpose" in terms of the mission's stated performance.	
Level-1 Partially M:	1) Advertising available     2) Catalogue search     available at product level	1)Basic online services available for data and metadata access	Basic schema for automated data use     Data in documented standard file format. Non-standard naming conventions used.	mission documentation available and preserved for the long term 2) No link between mission documentation and data records	1) Product inform available (not onl	Basic archiving for original data records preservation     Assessment of SW preservation     Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	Basic data quality control and monitoring check     Johiminal set of quality     control procedures     documented and available	data records preservation 2) Assessment of SW preservation 3) Product Details Assessment: Any required information missing	Data Records/Associated Information integrity basic check	2] Geometric Processing - Geometric processing documented. Missing all to 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.	Persistent identifier assignment only for particular Data Records Collections     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	Simple Access Architecture through metadata     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.	Documentation produced, published and well described 2) Link between mission documentation and data records created and managed	Dataset tested presence of corre provenance meta 2) Well described product informati 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	Quality indicator post- processing available     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.	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 Mana	1) Catalogue accessible vis 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 metigeneration for provenance documentation 2) Complete and u data provenance available online	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.	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) Data quality control fully compliant with an nternational 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 relevan 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 "fif for purpose" in terms of the mission's stated performance.  2) Geometric Processing - Geometric processing well-documented. State-of-the-art methodology used, easily "fif 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 "fif 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
						required and recommended information available						

# 



	DISCOVERABILITY	BILITY ACCESSIBILITY USABILITY					PRESERVA	TION	CURATION			
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMPS 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	No catalogue available     No advertising     available	Data and metadata are not accessible online	Data Not     Structured     Non-standard or     proprietary data     format, or, poorly- documented     standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	Reference Data Representativeness - No validation     Reference Data Quality - No validation     Walidation Method - No validation     Validation Results - No validation     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.	No control and monitoring check     No quality indicator in metadata     No procedures documentation	I) Uncontrolled storage location.     Only data are stored     Data Records archiving not managed     Relevant information on Product Details     Assessment not made	No Data/Associated Information integrity, authenticity and readability check	No reprocessing activities planned     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
Partially	Advertising available     Catalogue search     available at product level	Basic online services available for data and metadata access	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)	Reference Data Representativeness: measurements assessed to be mostly representative of the satellite measurements?  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.	I) 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	Basic data quality control and monitoring check     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	I) Minor updates and bugs corrections of data records implemented 2) Data Records appealoging and/or reformatting 3) Pre-flight calibration & characterisation misses more important apects 4) Post-baunch calibration & characterisation misses some important apects 4) Post-baunch calibration & characterisation misses some important apects of instrument behaviour and/or is not entirely of a level of quality to be judged fif for purpose. 5) Additional processing steps documented. Some emportant additional processing steps may not be fit for stated purpose.	2) Basic landing
Level-2 Managed	1) Detailed catalogue search available at product level 2) Product netadata 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	I) 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 stellite 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 & classification.  2) Uncertainty Sources: All important sources of uncertainty included.  3) Uncertainty Values: Total uncertainty produced per pixel is provided, with basic breakdown of key components no errocovariance.	Quality indicator post-processing available 2. Quality control procedures documented and available online	Preservation repository certified internally 2 Community-standard for archiving metadata 3) Product Details. Assessment: All required information available, any recommended information missing.	Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing	1) Reprocessing for calibration and for algorithm improvement 2) Preflight calibration & characterisation covers all reasonable aspects 3) Post-shaunch calibration & characterisation covers all reasonable aspects of instrument behaviour to a quality that is -fif for purpose in terms of the mission's stated performance and uses appropriate community instructure/methods (CEOS/RMs). 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	via international or community agreed standards protocol 4) Data policy available in metadata 5) Periodic updates of metadata in the	1) International standard for Data and metadata access system 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 metadata accessible through a free and open access protocol	1) Accepted and Available semantic encoding standards for complete interoperability 2) Data and metadas uses FAR-compliant vocabularies 3) Analysis Ready Data standard	1) Standards based metadata for documentation 2) Link between mission documentation and data records published	Automatic metadata generation for provenance documentation 2) Complete and updated data provenance available online	1) Reference Data Representativenes: Reference measurements independently assessed to be fully representative of the satellite reasurements, covering the satellite's full range of measurements and under some satellite reasurement of uncertaintain carried out on a regular basis determined by product performance. 2) Reference Data Quality full uncertainty and error-correlation information, assessed following the GDM and funceable to \$1 3) Validation Methods assess satellite reasurements and reference data w.z.1 their error-covariance and validate those uncertainties. 4) Validation Results show excellent agreement between satellite and inference 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 byte provided with error-covariance information for all appropriate components.	1) Data quality control fully compliant with an international stand 2) Quality indicator pre and post processing available in the metadata 3) Quality metadata assessed	1) Preservation repository officially certified 2) Periodic technology certereshment 3) Identify and manage that participates of relevant mission SW, ensuring that preserved data can be recreated. 4) Continuity of service availability and the control of the con	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 review 4) Metadata includes information about the Icence 5) Pre-Fight: As Level-2, additionally calibration and characterisation includes the measurements needed to assess uncertainties at component level and characterisation includes the measurements needed to assess uncertainties at component level and their impact on the final product. 6) Post-baunch 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 way that it can be away 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	USABIUTY		PRESERVATION		CURATION					
	MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMPS 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 Identi
Level-0 Not Managed	No catalogue available     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)	Reference Data Representativeness - No validation 2) Reference Data Quality - No validation 3) Validation Method - No validation 4) Validation Results - No validation	Uncertainty Method: Uncertainty characterisation not performed, or method not documented.     Uncertainty Sources: Uncertainty characterisation not performed, or sources analysed not documented.     Uncertainty Values: No uncertainty information provided.	No control and monitoring check     No quality indicator in metadata     No procedures documentation	I) Uncontrolled storage location.     Only data are stored     Data Records archiving not managed     Relevant information on Product Details     Assessment not made	No Data/Associated information integrity, authenticity and readability check	No reprocessing activities planned     Pre-flight calibration & characterisation not documented or information not available.     Post-launch calibration & characterisation not documented or not available.     Processing: Additional processing steps not documented.	No persistent a resolvable identifiers avails
Level-1 Partially Managed	Advertising available     Catalogue search     available at product level	Basic online services available for data and metadata access	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 of 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 Value: Single uncertainty value provided for subsets of data	Basic data quality control and monitoring check     Minimal set of quality control procedures documented and available	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.	Persistent identifier assignment only for particular D. Records     Collections     Basic landing pages manager
Level-2 Managed	Detailed catalogue search available at product twel 2) Product twel 2) Product twest an 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	Use of non-proprietary international standards encodings for syntactic interoperability.     Periodically reportanting of archived data.     Other in well-documented standard file format, community naming convention standard first 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.	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.	Persistent identifier assignment to a disseminated D Records     Collections and metadata     2) Automatic landing page generation and extensive management of landing pages.
Level-3 Fully Managed	I) International standard for Product metadata 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 in the catalogue accession expensive period of the peri	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	Accepted and Available semantic encoding standards for complete interoperability     Data and metadata uses FAIR-compliant vocabularies	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.	Data quality control fully compliant with an international standard 2) Qualify indicator pre and post processing available in the metadata     Qualify 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 werification 2) Data authenticity 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 create for all accessibil data records an inetadata 2) Metadata includes the identifier for the data as offered in such way that it can harvested and indexed.

COMPONENTS	L0	L1	L2	L3
Metadata for Discovery	*	*	*	*
Online Access	*	*	*	*
Data Encoding	*	*	*	$\star$
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



	PRESERVATION							
	MMP9 Data Preservation	MMP10 Data Verification						
LO	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	Data Records/Associated Information content integrity check and verification     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						

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



	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	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	Data Records/Associated Information content integrity check and verification     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						

PRESERVA	ATION			
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.



ATION MMP10 **Data Verification** No Data/Associated Information integrity, authenticity and readability Data Records/Associated Information integrity basic check 1) Data Records/Associated Information content integrity check and verification 2) Media readability and accessibility testing 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 EUROPEAN SPACE AGENCY



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 record 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

L2

L3

PRESERVA	ATION			
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			
Preservation repository certified internally     Community-standard for archiving metadata     Product Details Assessment: All required information available, any recommended information missing	Data Records/Associated Information content integrity check and verification     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			

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



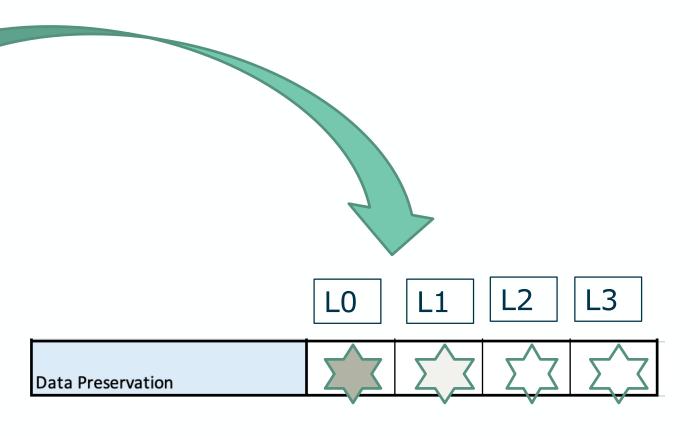
	PRESERV <i>A</i>	ATION			
	MMP9 Data Preservation	MMP10 Data Verification			
L0	1) Uncontrolled storage location. 2) Only data are stored 3) Data Records archiving not managed 4) Relevant information on Product Details Assessment not made available	No Data/Associated Information integrity, authenticity and readability check			
L1	Basic archiving for original data records preservation     Assessment of SW preservation     Product Details Assessment: Any required information missing	Data Records/Associated Inform 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			

PRESERV <i>I</i>	TION
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	Data Records/Associated     Information content integrity check     and verification     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 a task is not already implemented it remains in black and at the end the cell, partially covered, becomes light green.



	PRESERVA	ATION
	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 integrity, auther eadability check
L1	Basic archiving for original data records preservation     Assessment of SW preservation     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





	7	DISCOVERABILITY	ACCESSIBILITY				USABILITY			PRESERVATION		CURATION	
		MMP1 Metadata for Discovery	MMP2 Online Access	MMP3 Data Encoding	MMP4 Data Documentation	MMPS 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
		No catalogue available     No advertising     available	Data and metadata are not accessible online	Data Not     Structured     Non-standard or     proprietary data     format, or, poorly- documented     standard file format.	Partial and incomplete mission documentation	Limited product information available (not online)	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	Uncontrolled storage location.     Only data are stored     Data Records archiving not managed     Relevant information on Product Details     Assessment not made	No Data/Associated Information integrity, authenticity and readability check	No reprocessing activities planned     Pre-flight calibration & characterisation not documented or information not available.     Post-launch calibration & characterisation not documented or not available.     Processing: Additional processing steps not documented.	No persistent and resolvable identifiers available
Par	rtially	Advertising available     Catalogue search     available at product level	Basic online services available for data and metadata access	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)	Reference Data Representativeness: measurements assessed to be mostly representative of the safellite measurements?  Reference Data Quality: single uncertainty for the entire dataset.  3) Validation Method: simple uncertainty estimated  4) Validation Results: Validation results show good agreement between safellite 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	Basic data quality control and monitoring check     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	Information	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 fif for purpose. 5) Additional processing steps documented. Some important additional processing steps may not be fit for stated purpose.	Persistent identifier assignment only for particular Data Records     Collections     Basic landing pages management
	evel-2 inaged	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	I) Use of non- proprietary international standards encodings for syntactic interoperability.     2) Periodically respectaging/ 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.	Quality indicator post-processing available 2) Quality control procedures documented and available online	Preservation repository certified internally 21 Community-standard for archiving metadata 3) Product Details.     Assessment: All required information available, any recommended information missing.	Information content integrity check and verification	1) Reprocessing for calibration and/or algorithm improvement 2) Pre-fight 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/RMs). 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
F	evel-3 Fully inaged	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.	1) International standard for Data and metadata access system 2) Data policy available in the metadata. 3) Visualisation services 3) Visualisation services 6) Quick adoption to new technologies and standards evolution to new technologies and standards evolution to the visualisation of visualisation o	1) Accepted and Available semantic encoding standards for complete interoperability 2] Data and metadata uses FAIR-compliant vocabularies 3] Avalysis Ready Data standard	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, overling the satellite measurements, overling 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 GIMM and traceoble to 51  3) Validation Methods assess satellite measurements and reference data wx.t. their error-covariance and validates those uncertainties.  4) Validation Results show excellent agreement between statellite and reference measurements, within	1) Uncertainty Method: GUM approach to estimate measurement uncertainty, including a treatment of error-covariance. Juncertainty 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 recommendation 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	LO	L1	L2	L3
Metadata for Discovery	*	*	*	$\star$
Online Access	*	*	*	*
Data Encoding	*	*	*	$\star$
Data Documentation	*	*	*	$\star$
Data Traceability	*	*	*	$\star$
Data Validation	*	*	*	*
Data Uncertainty	*	*	$\star$	$\star$
Data Quality Control	*	*	*	*
Data Preservation	*	*	*	*
Data Verification	*	*	*	*
Data Processing/Reprocessing	*	*	*	*
Persistent & Resolvable Identifier	*	*	*	*

