



DEL-03
AIREO Specification

Date: March 19th, 2021

From: Irish Centre for High-End Computing (ICHEC)

To: EUROPEAN SPACE AGENCY (ESA),
The European Space Research Institute (ESRIN),
Largo Galileo Galilei 1,
00044 Frascati (RM), Italy.
Att.: Patrick Griffiths

Activity: AIREO Training Datasets (AO/3-16408/20/I/NB)

Category: ESA Express Procurement - [EXPRO]

Reference: Proposal No. 19.381.16

Document: AIREO DEL-03



This spreadsheet contains the metadata elements for the AIREO Specification. Each dataset will contain at a minimum the required metadata for 'CoreElements' and the metadata for one other profile type. Further details on the metadata and how to implement it are contained in the AIREO Specification and AIREO Best Practice Guidelines Documents.

AIREO Version	0
Version Release Date	21-Jun-2021
AIREO Specification Document	URL TBD
AIREO Best Practice Guidelines Document	URL TBD

Metadata Defined for all AIREO Profiles								
	Domain	Required	Recommended	Optional	Reference	Other reference mentioned in the description	Description	Comments
aireo_version	Item	X			URL of Specification		0' in initial releases	
title	Collection	X			STAC basics		A human readable title describing the Item.	
description	Collection	X			STAC basics	CommonMark 0.29	Detailed multi-line description to fully explain the Item. CommonMark 0.29 syntax may be used for rich text representation.	
created	Collection	X			STAC date and time	RFC 339, section 5.6	Creation date and time of the data subassets Could be NULL	
updated	Collection		X		STAC date and time	RFC 339, section 5.6	date and time data was updated last update of any element	
start_datetime	Collection		X		STAC date range	RFC 339, section 5.6	STAC date range describes best practises to set it up according to RFC 339, section 5.6	
end_datetime	Collection		X		STAC date range	RFC 339, section 5.6	STAC date range describes best practises to set it up according to RFC 339, section 5.6	
license	Collection	X			STAC licensing	SPDX license list	Information about license of the data. Need to be used SPDX License identifier. In case it is not listed, "various" will be used if multiple licenses apply or "proprietary" for all other cases.	
license_url_list	Collection			X	STAC licensing		The license URL(s) for the item SHOULD be specified if the license field is set to proprietary or various. If not possible, it is RECOMMENDED to supplement the STAC Item with the license text in a separate file, and link to this file. This item REQUIRED for any license marked "various" or "proprietary".	
providers_name	Collection			X	STAC Provider		Name of the organization or individual capturing, producing, processing, hosting or publishing this data. Providers should be listed in chronological order with the most recent provider being the last element of the list.	
providers_description	Collection			X	STAC Provider	CommonMark 0.29	Description to add further information of the provider such as add processing details, hosting details.	
providers_role	Collection			X	STAC Provider		Roles of the providers. Possible values : "licensor", "producer", "processor" or "host"	
providers_url	Collection			X	STAC Provider		Homepage on which further information and contact information is published	
collection	Collection			X	STAC item fields	bounding box	Collection field provides a straightforward way to search for any item belonging to a specific collection. STAC collections provide a breakdown of best practises to follow being one of the most important to refer to provide a link back to the collection definition (STAC relation types)	
id	Collection	X			STAC item fields		Unique identifier. Highly recommended to use just the ID in case provider identifies following an identification scheme, UUID is recommended.	
links	Collection		X		STAC item fields		List of link objects to resources and related URLs. A link with the rel set to self is strongly recommended.	
type	Collection	X			STAC item fields		Value must be "Catalog" or "Collection". Required for STAC	
stac_extensions	Collection	X			STAC item fields		List of extensions the item implements	
Identifier metadata								
stac_version	Collection	X			STAC item fields		Recommended as best practices	
doi	Collection			X	STAC scientific extension		Provide if defined.	
provenance	Collection		X		STAC item fields		If the TDS is created using information from various sources, all of them should be mentioned here in a human readable form. Reference should also be made to the links in the providers section where the original dataset providers were mentioned.	
publications	Collection			X	STAC scientific extension		If the data is accompanied by a publication it needs to be mentioned here. Data identifier. Highly recommended to use just the ID in case provider identifies following an identification scheme	
Datasheet metadata								
purpose	Collection		X				Describes the goal of the TDS. If possible, include the list specific tasks/problems the dataset was created for	

tasks	Collection	X			label/tasks		list of tasks possible with entire dataset: {'regression', 'classification', ... }	machine readable.
funding_info	Collection		X	X			Mention who funded the creation of the dataset	
collection_mechanism	Collection		X				Describes the procedures to collect the datasets. Free-text	
data_preprocessing	Collection	X					Describes how the preprocessing/cleaning was done on the data set.	
field_schema	Collection	X					Describes all the profiles (features and reference data) that make up the dataset and their corresponding data types {'input_a':['GeoreferencedEOImage'], 'target_a':['ReferenceData', 'Numeric']}	
example_definition	Collection	X					Define what is an individual training example (instance) in the TDS	
dataset_split	Collection		X				Describes the recommended Train/Validation/Test Split and the rationale behind them. Freetext	
data_completeness	Collection		X				Describe if the dataset has all the data needed in it, or it relies/points to external resources. Used to generate human-readable documentation	
data_sharing	Collection		X				Describes how the dataset will be distributed	
Data Elements for all AIREO profiles								
		Required	Recommended	Optional	Reference		Description	Comments
Quality Indicators for entire dataset								
		Required	Recommended	Optional	Reference		Description	Comments
AIREO Compliance level								
compliance_level		X					1 - required metadata present; 2 - required and recommended metadata present; 3 - required, recommended and optional metadata present	
FAIR indicators:								
findable				X			Example: a fully findable dataset will have the value [A,B,C,D] in the findable metadata element A - (Meta)data are assigned a globally unique and persistent identifier; B - Data are described with rich metadata; C - Metadata clearly and explicitly include the identifier of the data they describe; D - (Meta)data are registered or indexed in a searchable resource	
accessible				X			A - (Meta)data are retrievable by their identifier using a standardised communications protocol (open, free, and universally implementable or allows for authentication and authorisation procedure, where necessary); B - Metadata are accessible, even when the data are no longer available	
interoperable				X			A - (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation; B - (Meta)data use vocabularies that follow FAIR principles; C - (Meta)data include qualified references to other (meta)data	
reusable				X			A - Full metadata, B - provenance, C - licence information, D - using community standards, eg file formats	
Data provenance								
processing_history				X			Global check on all processing metadata in all profiles (Set of processes applied to source data to achieve the current version of a dataset, including tools and software used)	
feature_engineering_recipes			X				Global check on feature recipe metadata in all profiles (Set of processes involved in creation, extraction and selection of features, including tools and software used)	
data_source			X				Global check on data sources in all profiles	

Metadata for the GeoreferencedEOImage Profile							
	Domain	Required	Recommended	Optional	Reference	Description	OGC Metadata Name
EO Metadata					https://docs.opengeospatial.org/is/10-157r4/10-157r4.htm#22		
eop:EarthObservationMetaData							
identifier	Collection	X			eop:EarthObservationMetaData	Identifier for item	identifier
doi	Collection			X	eop:EarthObservationMetaData	Digital Object Identifier	doi
parent_identifier	Collection	X			eop:EarthObservationMetaData	Collection Identifier	parentIdentifier
product_type	Collection	X			eop:EarthObservationMetaData	Describes the product type in case mixed types are present	productType
acquisition_date	Collection	X			eop:EarthObservationMetaData	Acquisition date time	downloadedTo/ DownlinkInformation/ acquisitionDate
product_quality_report_url	Collection			X	eop:EarthObservationMetaData	URL reference to an external quality report	productQualityReportURL
histogram_band_id	Collection			X	eop:EarthObservationMetaData	Identifier of the spectral band used to compute histogram values	histograms/Histogram/ bandId
histogram_min	Collection			X	eop:EarthObservationMetaData	Histogram minimum value	histograms/Histogram/ min
histogram_max	Collection			X	eop:EarthObservationMetaData	Histogram maximum value	histograms/Histogram/ max
histogram_mean	Collection			X	eop:EarthObservationMetaData	Histogram mean value	histograms/Histogram/ mean
histogram_std	Collection			X	eop:EarthObservationMetaData	Histogram standard deviation value	histograms/Histogram/ stdDeviation
composed_of	Collection			X	eop:EarthObservationMetaData	Link to an EO product that is part of this EO product (e.g. a phr.DataStrip is composed of one or more phr.Scene)	composedOf
subset_of	Collection			X	eop:EarthObservationMetaData	Link to the "father" EO product (e.g. a phr.Scene is a subset of a phr.DataStrip)	subsetOf
linked_with	Collection			X	eop:EarthObservationMetaData	Specify a link to another EO product (e.g. ERS1 and ERS2 interferometric pair)	linkedWith
processing_date	Collection			X	eop:EarthObservationMetaData	Processing date time	processing/ProcessingInformation/ processingDate
composite_type	Collection		X		eop:EarthObservationMetaData	Composite product expressed as timeperiod	processing/ProcessingInformation/ compositeType
processing_info_method			X		eop:EarthObservationMetaData	Method used to compute datalayer. (e.g. Kalman filtering, ROSE)	processing/ProcessingInformation/ method
processing_info_version	Collection		X		eop:EarthObservationMetaData	Method version	processing/ProcessingInformation/ method Version
processing_info_name	Collection			X	eop:EarthObservationMetaData	Processor software name	processing/ProcessingInformation/ processorName
processing_info_processing_version	Collection			X	eop:EarthObservationMetaData	Processor software version	processing/ProcessingInformation/ processingVersion
processing_info_native_product_format	Collection			X	eop:EarthObservationMetaData	Native product format	processing/ProcessingInformation/ nativeProductFormat
processing_info_processing_level	Collection	X			eop:EarthObservationMetaData	Processing level applied to the product	processing/ProcessingInformation/ processingLevel
processing_info_auxiliary_dataset_filename	Collection			X	eop:EarthObservationMetaData	Name of auxiliary dataset used.	processing/ProcessingInformation/ auxiliaryDataSetFileName
EarthObservationEquipment : ALL METADATA CONTAINING INFORMATION RELATIVE TO THE MECHANISM USED DURING THE OBSERVATION							
platform_short_name	Collection		X		eop:EarthObservationEquipment	Platform short name	platform/Platform/shortName
platform_serial_identifier	Collection		X		eop:EarthObservationEquipment	Platform serial identifier	platform/ Platform/ serialIdentifier
platform_orbit_type	Collection			X	eop:EarthObservationEquipment	High level characterisation of main mission types (GEO, LEO)	platform/ Platform/ orbitType

instrument_short_name	Collection		X		eop:EarthObservationEquipment	Sensor name	instrument/ Instrument/ shortName
instrument_description	Collection			X	eop:EarthObservationEquipment	Instrument description	instrument/ Instrument/ description
instrument_instrument_type	Collection			X	eop:EarthObservationEquipment	Instrument type	instrument/ Instrument/ instrumentType
sensor_type	Collection		X		eop:EarthObservationEquipment	Sensor type based (OPTICAL, RADAR, ALTIMETRIC, ATMOSPHERIC,...)	sensor/ Sensor/ sensorType
sensor_operational_mode	Collection		X		eop:EarthObservationEquipment	Sensor mode	sensor/ Sensor/ operationalMode
sensor_resolution	Collection		X		eop:EarthObservationEquipment	Sensor resolution	sensor/ Sensor/ resolution
sensor_discrete_wavelengths	Collection			X	eop:EarthObservationEquipment	List of Wavelengths observed in the product	sensor/ Sensor/ wavelengthInformation/ WavelengthInformation/ discreteWavelengths
sensor_start_wavelength	Collection			X	eop:EarthObservationEquipment	Start of the observed Wavelength range	sensor/ Sensor/ wavelengthInformation/ WavelengthInformation/ startWavelength
sensor_end_wavelength	Collection			X	eop:EarthObservationEquipment	End of the observed Wavelength range	sensor/ Sensor/ wavelengthInformation/ WavelengthInformation/ endWavelength
sensor_spectral_range	Collection			X	eop:EarthObservationEquipment	Observed spectral range (IR, NIR, UV, V,...)	sensor/ Sensor/ wavelengthInformation/ WavelengthInformation/ spectralRange
acquisition_parameters_orbit_number	Collection			X	eop:EarthObservationEquipment	Acquisition orbit number	acquisitionParameters/ Acquisition/ orbitNumber
acquisition_parameters_ascending_node_longitude	Collection			X	eop:EarthObservationEquipment	Longitude at ascending node orbit	acquisitionParameters/ Acquisition/ ascendingNodeLongitude
acquisition_parameters_incidence_angle	Collection			X	eop:EarthObservationEquipment	Incidence angle	acquisitionParameters/ Acquisition/ incidenceAngle
acquisition_parameters_illumination_zenith_angle	Collection			X	eop:EarthObservationEquipment	Mean illumination/solar zenith angle given in degrees.	acquisitionParameters/ Acquisition/ illuminationZenithAngle
EarthObservationResult: All metadata containing information about observed results							

						Reference to File or OGC Web Service In case the products are offered from FTP or HTTP URLs, the xlink:href attribute is used to encode the full URL to the product and the ows:RequestMessage element is left blank. In case the products are offered through WMS or WCS using HTTP GET with KeyValuePair encoding, the xlink:href attribute is used to encode the full URL including the KVP and the ows:RequestMessage element is left blank. In case the products are offered through a service supporting HTTP POST or SOAP the xlink:href attribute is used to encode the service endpoint (online resource and the ows:RequestMessage shall contain the XML encoded message (including the SOAP Header in case of SOAP messaging).	product/ ProductInformation/ filename
product_info_filename	Collection	X			eop:EarthObservationResults		
product_info_version	Collection			X	eop:EarthObservationResults	Product version	product/ ProductInformation/ version
product_info_size	Collection			X	eop:EarthObservationResults	Product size (bytes) allowing the user to realise how long a download is likely to take	product/ ProductInformation/ size
product_info_timeliness	Collection			X	eop:EarthObservationResults	timeliness of the product ("Near real time", "real time")	product/ ProductInformation/ timeliness
mask_info_type	Collection	X			eop:EarthObservationResults	Mask type. Possible values are : SNOW, CLOUD and QUALITY	mask/ MaskInformation/ type
mask_info_sub_type	Collection			X	eop:EarthObservationResults	Further information on Mask type	mask/ MaskInformation/ subType
mask_info_format	Collection	X			eop:EarthObservationResults	Mask format (RASTER, VECTOR)	mask/ MaskInformation/ format
						Reference to File or OGC Web Service In case the masks are offered from FTP or HTTP URLs, the xlink:href attribute is used to encode the full URL to the product and the ows:RequestMessage element is left blank. In case the masks are offered through WMS or WCS using HTTP GET with KeyValuePair encoding, the xlink:href attribute is used to encode the full URL including the KVP and the ows:RequestMessage element is left blank. Preferably the masks are encoded using GML 3.2.1 following the model that is specified in Annex F. In case the masks are offered through a service supporting HTTP POST or SOAP the xlink:href attribute is used to encode the service endpoint (online resource and the ows:RequestMessage shall contain the XML encoded message (including the SOAP Header in case of SOAP messaging).	mask/ MaskInformation/ filename
mask_info_filename	Collection	X			eop:EarthObservationResults		
mask_info_multi_extent_of	Collection	X			eop:EarthObservationResults	Contains inline encoded mask polygon geometries using the gml:MultiSurface/gml:surfaceMembers/gml:Polygon constructs.	mask/MaskInformation/ multiExtentOf

parameter_info_unit_of_measure	Collection	X			eop:EarthObservationResults	Unit of measure of observed	parameter/ParameterInformation/unitOfMeasure
parameter_info_phenomenon	Collection	X			eop:EarthObservationResults	A SWE 1.0 Phenomenon. Could be a single SWE Phenomenon (such as Sea Surface Height) or a SWE ConstrainedPhenomenon, such as a list of particular radiance bands, or a SWECompositePhenomeon which groups several discrete phenomena	parameter/ParameterInformation/phenomenon
coverage	Collection			X	eop:EarthObservationResults	Reference to coverage exploitation metadata (domainSet, RangeType, ...) as offered by a corresponding WCS using a HTTP GET encoded DescribeCoverage Request.	coverage
Quality Indicators for GeoreferencedEOImage Profile							
		Required	Recommended	Optional	Reference	Description	
qi_accuracy_outliers			X				
qi_accuracy_data_alignment			X			Assessment of co-alignment in altitude, latitude, longitude and time. In this case, alignment between multiple sources of EO data product if present.	
qi_completeness_missing_values			X			% of data as NAN	
qi_completeness_metadata			X			% of recommended metadata filled, % of optional metadata filled	
qi_eo_available			X			Pixel/granule level QI reported if available	
Data Elements for the GeoreferencedEOImage Profile							
	Domain	Required	Recommended	Optional	Reference	Description	
start_datetime	Item		X		STAC date range	RFC 339, section 5.6	
end_datetime	Item		X		STAC date range	RFC 339, section 5.6	
gsd	Item		X		STAC instrument	Ground sample distance at sensor	
bbox	Item	X			STAC item fields	REQUIRED if geometry is not null. Bounding Box of the asset represented by this Item, formatted according to RFC 7946, section 5.	
geometry	Collection	X			STAC item fields	GeoJSON Geometry object: Defines the full footprint of the asset represented by this item, formatted according to RFC 7946, section 3.1. The footprint should be the default GeoJSON geometry, though additional geometries can be included. Coordinates are specified in Longitude/Latitude or Longitude/Latitude/Elevation based on WGS 84.	
Controlled Vocabularies for the GeoreferencedEOImage Profile							
	Domain	Required	Recommended	Optional	Reference	Description	

Metadata for the GeoreferencedEOData Profile							
	Domain	Required	Recommended	Optional	Reference	Description	Comments
identifier	Collection	X			eop:EarthObservationMetaData	Identifier for item, as provided by provider	
parent_identifier	Collection	X			eop:EarthObservationMetaData	Collection Identifier, from provider	
product_type	Collection	X			eop:EarthObservationMetaData	Describes the product type	
composed_of	Collection			X	eop:EarthObservationMetaData	Link to an EO product that is part of this EO product	
subset_of	Collection			X	eop:EarthObservationMetaData	Link to the "father" EO product	
linked_with	Collection		X		eop:EarthObservationMetaData	Specify a link to another EO product	
processing_info_native_product_format	Collection			X	eop:EarthObservationMetaData	Native product format	
processing_info_processing_level	Collection		X		eop:EarthObservationMetaData	Processing level applied to the product	
processing_info_auxiliary_dataset_filename	Collection			X	eop:EarthObservationMetaData	Name of auxiliary dataset used.	
Quality Indicators for GeoreferencedEOImage Profile							
		Required	Recommended	Optional	Reference	Description	
qi_accuracy_outliers			X			% of data as outliers	
qi_accuracy_data_alignment			X			Assessment of co-alignment in altitude, latitude, longitude and time. In this case, alignment between multiple sources of EO data product if present.	
qi_accuracy_outliers			X			% of data as NAN	
qi_completeness_metadata			X			% of recommended metadata filled, % of optional metadata filled	
qi_eo_available			X			Pixel/granule level QI reported if available	
Controlled Vocabularies for the GeoreferencedEOData Profile							
	Domain	Required	Recommended	Optional	Reference	Description	Comments
start_datetime	Item		X		STAC date range	RFC 339, section 5.6	As per CoreElement, but specific to this item
end_datetime	Item		X		STAC date range	RFC 339, section 5.6	As per CoreElement, but specific to this item
gsd	Item		X		STAC instrument	Ground sample distance at sensor	
bbox	Item	X			STAC item fields	REQUIRED if geometry is not null. Bounding Box of the asset represented by this Item, formatted according to RFC 7946, section 5.	Formatted following RFC 7946, section 5. All GeoJSON coordinates shall use World Geodetic System 1984 (WGS84), with latitude/Longitude units in decimal. The equivalent CRS identified by OGC URN: urn:ogc:def:crs:OGC::CRS84
geometry	Collection	X			STAC item fields	GeoJSON Geometry object: Defines the full footprint of the asset represented by this item, formatted according to RFC 7946, section 3.1. The footprint should be the default GeoJSON geometry, though additional geometries can be included. Coordinates are specified in Longitude/Latitude or Longitude/Latitude/Elevation based on WGS 84.	

