Interopera

Digitale Interoperabilität in kollaborativen Wertschöpfungsnetzwerken der Industrie 4.0

SPECIFICATION

Submodel AI-Dataset

Version 1.0

23 06 2023

Submodel Template of the
Asset Adminstration Shell

Imprint

Publisher

Steinbeis Innovation gGmbH
Adornostr. 8
70599 Stuttgart
Germany

Source for Specification Document

Plattform Industrie 4.0
Bertolt-Brecht-Platz 3
10117 Berlin
Germany

Authors

Mike Reichardt
Daniel Buch
Philip Stricker
Dachuan Shi
Georg Güntner
Sebastian Baron
Christoph Legat
Jürgen Lenz
Nico Braunisch

Version history

2023-06-23	1.0	Submission to InterOpera Consortium

Content

List of Figures

List of Tables

Forward

We would like to thank all of the working group members for their help and support to develop the AI Dataset Submodel. The discussions were always helpful and we got lots of good ideas out of it.

A special thanks goes to the Interopera Consortium, that provided us with the opportunity do develop the AI Dataset Submodel.

General

1.1 About this document

This document is a part of a specification series. Each part specifies the contents of a Submodel template for the Asset Administration Shell (AAS). The AAS is described in [1-3] and [6]. First exemplary Submodel contents were described in [4], while the actual format of this document was derived by the "Administration Shell in Practice" [5]. The format aims to be very concise, giving only minimal necessary information for applying a Submodel template, while leaving deeper descriptions and specification of concepts, structures and mapping to the respective documents [1-6].

The target group of the specification are developers and editors of datasets for ML applications, which are describing a dataset by means of the Asset Administration Shell (AAS) and therefore need to create a Submodel instance with a hierarchy of SubmodelElements. This document especially details on the question, which SubmodelElements with which semantic identification shall be used for this purpose.

1.2 Scope of the Submodel

This Submodel template aims at interoperable provision of information describing a dataset and meta information about it in regard to the asset of the respective Asset Administration Shell. Central element is the provision of properties [7], ideally interoperable by the means of dictionaries such as ECLASS and IEC CDD (Common Data Dictionary). The purpose of this document is to make selected specifications of Submodels in such manner that information about assets can be exchanged in a meaningful way between partners in a value creation network. It aims to provide detailed information about a dataset used in AI context. This includes additional service information (the person responsible for the dataset).

The intended use-case is the provision of a standardized property structure for datasets used to train AI, which enables a standardized way to describe datasets. It also helps to recycle existing datasets, by providing meta information, for example environmental conditions. With the help of two additional Submodels (AI ModelNameplate and AI Deployment Submodels), it allows to monitor the complete AI lifecycle.

This concept can serve as a basis for standardizing the respective Submodel. The conception is based on existing norms, studies of common practices at enterprises, directives and standards so that a far-reaching acceptance can be achieved.

1.3 Relevant standards and sources of concepts for the Submodel template

According to [3], interoperable properties might be defined by standards, consortium specifications or manufacturer specifications. So called property dictionaries are used identify information elements (see Terms and Definitions of [6]). Such property dictionaries include:

- ECLASS, see: https://www.eclasscontent.com/
- IEC CDD, see: https://cdd.iec.ch/cdd/iec61987/iec61987.nsf and https://cdd.iec.ch/cdd/iec62683/cdddev.nsf

In this document, properties are aimed to be described by ECLASS.

Information set for Submodel AI-Dataset

While defining Submodels the following three aspects must be considered as suggested in [5]:

Use and economic relevance

The Submodel AI Dataset is designed for every AI application that uses datasets. Its use case attends to document datasets and allows an easier reuse of existing datasets, because of standardized access to additional meta information. Meta Information are for example the environment conditions, in which the data was created.

It assists with communication within a development process, because of the assignment of a person responsible for each dataset. Therefore ML experts can contact the dataset expert (responsible person) directly for additional information.

Further a generic example use case is introduced. The AI Dataset Submodel provides information of the dataset location. In addition to the dataset location, other additional information is stored in the AAS Submodel. On the one hand, static parameters should be stored, such as the mean value or the median. On the other hand, information on the data format (e.g. PNG for image files) should be stored. Furthermore, information on labels, if available, shall be stored. Apart from data based directly on the training data, influencing factors in the application should also be covered. In particular, environmental conditions that can affect sensors should be mentioned here. Parameters would, for example, be the air humidity or the air temperature. Alongside the direct and indirect influences, service information must be recorded when a dataset is first created. These deal with the creator of the data and assign a direct unique ID to the data. Additional the date of creation of the dataset is registered and a responsible person for the dataset is defined.

Possible functions and interactions

The Submodel “AI Dataset” provides information from a dataset. Data scientists, AI experts can use the Submodel to obtain information about the dataset. Data scientists can select data for AI training based on metrics provided by the Submodel. Experts may be interested in the Submodel in order to investigate cases of damage and to draw conclusions about the cause. In addition, the Submodel provides information about a contact person and thus enables an easy exchange between user and creator, assisting the work of Data scientists.

The SMC "Labeled" contains information about the labelled data. The two children (SMC “Classification” and SMC “Regression”) of SMC “Labeled” are designed in such a way that a distinction is made between classification and regression data. This enables the characteristic mapping of the special features of both areas without reducing the freedom of the user. For example, both SMCs allow the upload of a single annotation file or multiple individual annotation files with additional meta-information.

The mandatory SMC "SizeInformation" contains information about the complete data set size. Furthermore, additional information about e.g. the training data size can be given. The SMC "MetaData" enables the user to store metadata. Depending on the type of data (image, audio,...), domain-specific meta information is stored in each case by a different child SMC. Information that cannot be assigned to a specific domain should be stored in the SMC "AdditionalInformation".

In order to be able to compare data sets with each other, there is the SMC "Metrics", which makes it possible to save various metrics of the dataset. For an overview of when the dataset has been used, there is the SMC "History". For each model trained with this dataset, a

separate SMC must be created in SMC History containing the access period and the reference to the SMC "ModelNameplate".

In the creation of datasets, there are influencing factors that cannot be measured directly, but which affect the informative value of a dataset. The SMC "BoundaryConditions" has been created to take these influencing factors into account. One identified influencing factor is all data collectors (e.g. sensors). These are collected in the SMC "DataCollectors", whereby each data collector receives its own SMC and existing AAS of the data collectors are referenced. A second identified influencing factor is environmental conditions, which are stored in the SMC "EnvironmentConditions", if they are not directly measured by a sensor but represent necessary background information for the dataset.

Property specification

See section 3 Submodel and Collections.

Submodel and Collections

3.1 Properties of the Submodel “AI-Dataset”

The figure below shows the UML-diagram defining the relevant properties which need to be set. Table 1: Properties of Submodel “Dataset" describes the details of the Submodel structure.

Figure 1: UML-Diagram for Submodel "AI-Dataset"

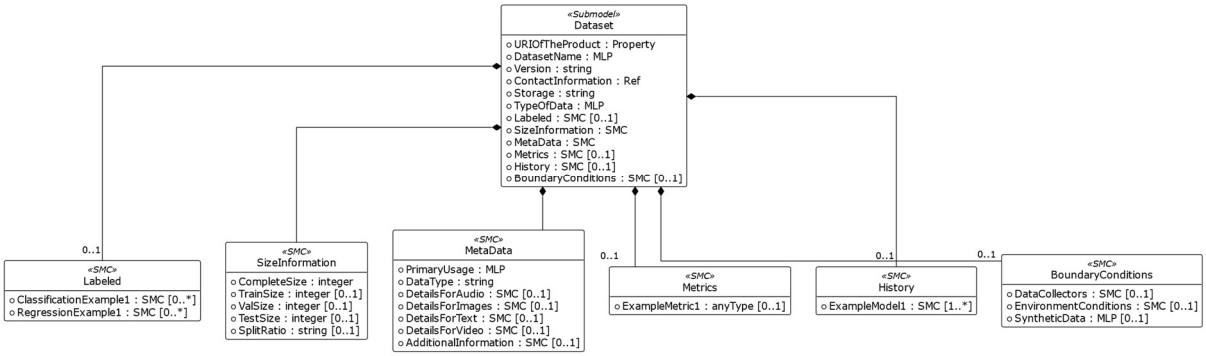

Table 1: Properties of Submodel “Dataset"

idShort:	Dataset		
Class:	Submodel		
semanticId:	[https://admin-shell.io/id/InterOpera/AIDataset]		
Parent:			
Explanation:			
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] URIOfTheProduct	[IRDI]0173-1#02-ABH173#001 URI of the product	[string]	1
[MLP] DatasetName	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/DatasetName Name of the dataset	[] @en	1
[Property] Version	[IRDI]0173-1#02-AAS354#002 Version	[string]	1
[Ref] ContactInformation	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/ContactInformation Reference to the Contact Information IDTA Submodel to describe the responsible person for the Submodel	[]	1
[Property] Storage	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/Storage	[string]	1

	Path to the dataset (e.g. local path, serverpath,...)		
[MLP] TypeOfData	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/TypeOfData Type of data in this dataset (e.g. audiofiles, images, videos, ...)	[] @en	1
[SMC] Labeled	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/Labeled Additional Information for labeled datasets	[] 2 elements	0..1
[SMC] SizeInformation	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation Collection about the number of dataset elements in the dataset and subsets	[] 5 elements	1
[SMC] MetaData	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData Collection of meta data information about the dataset	[] 7 elements	1
[SMC] Metrics	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/Metrics Collection of different metrics	[] 1 elements	0..1
[SMC] History	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/History Collection of models already trained with this dataset	[] 1 elements	0..1
[SMC] BoundaryConditions	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions Boundary conditions in which the dataset was created	[] 3 elements	0..1

3.2 Properties of the SMC “Labeled”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table and figure describe the details of the SMC structure.

Figure 2: UML-Diagram for SubmodelCollection "Labeled"

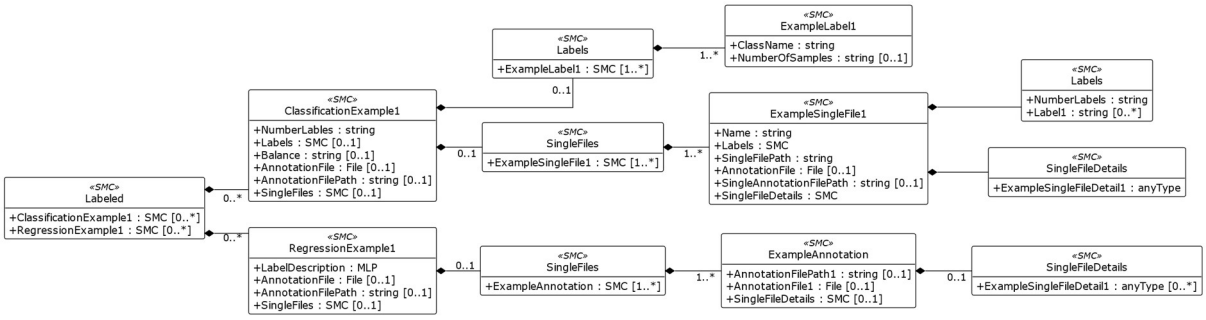

Table 2: Properties of SMC "Labeled"

idShort:	Labeled		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/Labeled		
Parent:	AIDataset		
Explanation:	Additional Information for labeled datasets@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] ClassificationExample1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification Information about labeled data for classification.	[] 6 elements	0..*
[SMC] RegressionExample1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression Information about labeled data for regression	[] 4 elements	0..*

3.3 Properties of the SMC “ClassificationExample1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 3: Properties of SMC " ClassificationExample1

idShort:	ClassificationExample1		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification		
Parent:	LabeId		

Explanation:	Information about labeled data for classification.@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] NumberLables	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/NumberLabels Number of the different labels in the dataset	[string]	1
[SMC] Labels	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/Labels every label as a property and the number of samples	[] 1 elements	0..1
[Property] Balance	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/Balance Balance between the classes	[string]	0..1
[File] AnnotationFile	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/AnnotationFile Annotation file of the dataset (e.g. CSV,JSON,...)	[]	0..1
[Property] AnnotationFilePath	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/AnnotationFilePath Path to the annotation file location	[string]	0..1
[SMC] SingleFiles	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles Collection containing every dataset element individuallly	[] 1 elements	0..1

3.4 Properties of the SMC “Labels”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 4: Properties of SMC "Labels"

idShort:	Labels		
Class:	SubmodelElementCollection		

semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabelId/Classification/Labels		
Parent:	Classification		
Explanation:	every label as a property and the number of samples@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] ExampleLabel1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabelId/Classification/Labels/Label Additional information about one label	[] 2 elements	1..*

3.5 Properties of the SMC “ExampleLabel1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 5: Properties of SMC " ExampleLabel1"

idShort:	ExampleLabel1		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabelId/Classification/Labels/Label		
Parent:	Labels		
Explanation:	Additional information about one label@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] ClassName	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabelId/Classification/Labels/Label/ClassName Name of the class	[string]	1
[Property] NumberOfSamples	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabelId/Classification/Labels/Label/NumberOfSamples Number of occurrences of the label in the dataset	[string]	0..1

3.6 Properties of the SMC “SingleFiles”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 6: Properties of SMC " SingleFiles"

idShort:	SingleFiles		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles		
Parent:	Classification		
Explanation:	Collection containing every dataset element individuallly@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] ExampleSingleFile1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/ Collection for features and the annotation file for a single dataset element.	[] 6 elements	1..*

3.7 Properties of the SMC “ExampleSingleFile1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 7: Properties of SMC " ExampleSingleFile1"

idShort:	ExampleSingleFile1		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile		
Parent:	SingleFiles		
Explanation:	Collection for features and the annotation file for a single dataset element.@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] Name	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/Name Name of the file	[string]	1
[SMC] Labels	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/Labels Collection containing all the labels within the dataset element	[] 2 elements	1

[Property] SingleFilePath	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/SingleFilePath Path to the dataset element	[string]	1
[File] AnnotationFile	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/AnnotationFile Annotation file with labeling information about the dataset element	[]	0..1
[Property] SingleAnnotationFilePath	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/SingleAnnotationFilePath Path to the annotation file location	[string]	0..1
[SMC] SingleFileDetails	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/SingleFileDetails Collection to add additional information about the dataset element (e.g. speaker information for audio data)	[] 1 elements	

3.8 Properties of the SMC “Labels”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 8: Properties of SMC " Labels"

idShort:	Labels		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/Labels		
Parent:	ExampleSingleFile1		
Explanation:	Collection containing all the labels within the dataset element@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	

[Property] NumberLabels	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/Labels/NumberLabels Number of labels in this dataset element	[string]	1
[Property] Label1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/Labels/Label ExampleLabel, this property should be existing for every label with the label name as value	[string]	0..*

3.9 Properties of the SMC “SingleFileDetails”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 9: Properties of SMC " SingleFileDetails"

idShort:	SingleFileDetails		
Class:	SubmodelElementCollection		
semanticId:	[IRI] https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/SingleFileDetails		
Parent:	SingleFile		
Explanation:	Collection to add additional information about the dataset element (e.g. speaker information for audio data)@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] ExampleSingleFileDetail1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Classification/SingleFiles/SingleFile/SingleFileDetails/SingleFileDetail additional information about the dataset element	[anyType]	1

3.10 Properties of the SMC “RegressionExample1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 10: Properties of SMC " RegressionExample1 "

idShort:	RegressionExample1		
Class:	SubmodelElementCollection		

semanticId:	[IRI] https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles		
Parent:	ExampleAnnotation		
Explanation:	Collection containing every dataset element individually@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] ExampleAnnotation	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation Collection for a single annotation of a dataset element	[] 3 elements	1..*

3.11 Properties of the SMC “ExampleAnnotation”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 11: Properties of SMC " ExampleAnnotation "

idShort:	ExampleAnnotation		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation		
Parent:	ExampleAnnotation		
Explanation:	Collection for a single annotation of a dataset element@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] AnnotationFilePath1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation/AnnotationFilePath Path to the annotation file location	[string]	0..1
[File] AnnotationFile1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation/AnnotationFile Annotation file with labeling information about the dataset element	[]	0..1
[SMC] SingleFileDetails	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation/SongleFileDetails	[] 1 elements	0..1

	Collection to add additional information about the dataset element		

3.12 Properties of the SMC “SingleFileDetails”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 12: Properties of SMC " SingleFileDetails"

idShort:	SingleFileDetails		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation/SingleFileDetails		
Parent:	ExampleAnnotation		
Explanation:	Collection to add additional information about the dataset element@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] ExampleSingleFileDetail1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/LabeId/Regression/SingleFiles/Annotation/SingleFileDetails/Detail This is a property for a specific metainformation for single regression files.	[anyType]	0..*

3.13 Properties of the SMC “SizeInformation”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 13: Properties of SMC " SizeInformation"

idShort:	SizeInformation		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation		
Parent:	AIDataset		
Explanation:	Collection about the number of dataset elements in the dataset and subsets@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	

[Property] CompleteSize	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation/CompleteSize Number of all dataset elements in the dataset	[integer]	1
[Property] TrainSize	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation/TrainSize Number of all dataset elements in the training subset	[integer]	0..1
[Property] ValSize	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformationValSize Number of all dataset elements in the validation subset	[integer]	0..1
[Property] TestSize	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation/TestSize Number of all dataset elements in the test subset	[integer]	0..1
[Property] SplitRatio	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/SizeInformation/SplitRatio Splitratio between the sets, notate as value1:value2:value3 if all three exist, else value1:value2	[string]	0..1

3.14 Properties of the SMC “MetaData”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table and figure describe the details of the SMC structure.

Figure 3: UML-Diagram for SubmodelCollection "MetaData"

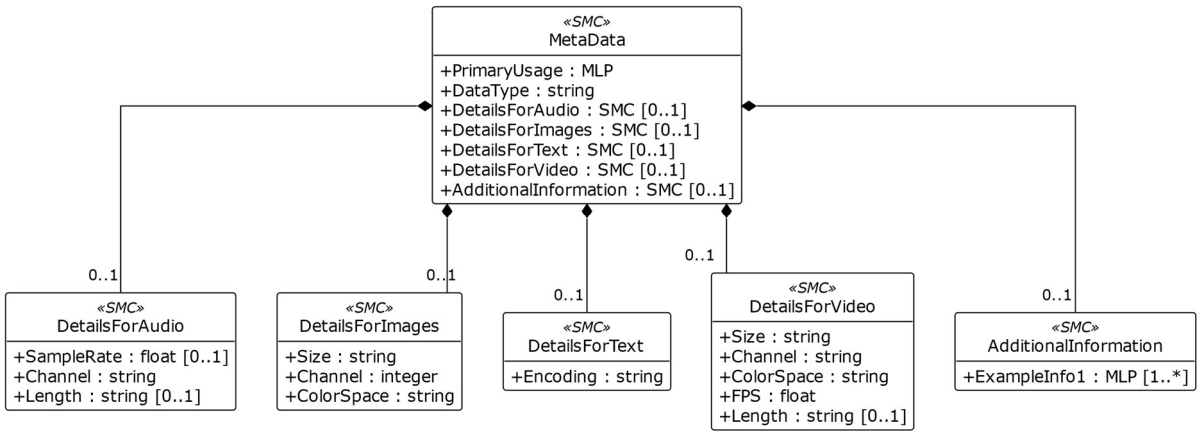

Table 14: Properties of SMC "MetaData"

idShort:	MetaData		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData		
Parent:	AIDataset		
Explanation:	Collection of meta data information about the dataset@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[MLP] PrimaryUsage	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/PrimaryUsage Purpose for creating the dataset (e.g. image recognition, keyword spotting, object detection, ...)	[] @en	1
[Property] DataType	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DataType Data type of the dataset elements (e.g. .WAV, .JPEG, ...)	[string]	1
[SMC] DetailsForAudio	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForAudio Specific collection for audio data	[] 3 elements	0..1
[SMC] DetailsForImages	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForImages Specific collection for image data	[] 3 elements	0..1
[SMC] DetailsForText	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForText Specific collection for text data	[] 1 elements	0..1
[SMC] DetailsForVideo	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo Specific collection for video data	[] 5 elements	0..1

[SMC] AdditionalInformation	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/AdditionalInformation Collection of additional meta information of the dataset	[] 1 elements	0..1

3.15 Properties of the SMC “DetailsForAudio”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 15: Properties of SMC " DetailsForAudio"

idShort:	DetailsForAudio		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForAudio		
Parent:	MetaData		
Explanation:	Specific collection for audio data@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] SampleRate	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForAudio/SampleRate Sample rate of the dataset	[float]	0..1
[Property] Channel	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForAudio/Channel Number of the channels (e.g. 1 for mono or 2 for stereo)	[string]	1
[Property] Length	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForAudio/Length Duration of a dataset element	[string]	0..1

3.16 Properties of the SMC “DetailsForImages”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 16: Properties of SMC " DetailsForImages"

idShort:	DetailsForImages		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForImages		
Parent:	MetaData		
Explanation:	Specific collection for image data@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] Size	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForImages/Size Resolution of an image (notate as e.g. 128x128)	[string]	1
[Property] Channel	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForImages/Channel Number of color channels (e.g. 3 for RGB)	[integer]	1
[Property] ColorSpace	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForImages/ColorSpace Color space of the images (RGB, CYMK, Grayscale, ...)	[string]	1

3.17 Properties of the SMC “DetailsForText”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 17: Properties of SMC " DetailsForText"

idShort:	DetailsForText		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForText		
Parent:	MetaData		
Explanation:	Specific collection for text data@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] Encoding	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForText/Encoding	[string]	1

	Encoding information for text data.		

3.18 Properties of the SMC “DetailsForVideo”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 18: Properties of SMC " DetailsForVideo"

idShort:	DetailsForVideo		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo		
Parent:	MetaData		
Explanation:	Specific collection for video data@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] Size	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo/Size Resolution of the videos (e.g. 128x128)	[string]	1
[Property] Channel	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo/Channel Number of color channels (e.g. 3 for RGB)	[string]	1
[Property] ColorSpace	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo/ColorSpace Color space of the videos	[string]	1
[Property] FPS	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo/FPS Frames per second of the videos	[float]	1
[Property] Length	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/DetailsForVideo/Length Duration of a dataset element	[string]	0..1

3.19 Properties of the SMC “AdditionalInformation”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 19: Properties of SMC "AdditionalInformation"

idShort:	AdditionalInformation		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/AdditionalInformation		
Parent:	MetaData		
Explanation:	Collection of additional meta information of the dataset@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[MLP] ExampleInfo1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/AdditionalInformation/Info Meta data information of the dataset	[] @en	1..*

3.20 Properties of the SMC “Metrics”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 20: Properties of SMC " Metrics"

idShort:	Metrics		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/Metrics		
Parent:	AIDataset		
Explanation:	Collection of different metrics@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] ExampleMetric1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/MetaData/Metrics/Metric metric of the dataset	[anyType]	0..1

3.21 Properties of the SMC “History”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 21: Properties of SMC " History"

idShort:	History		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/History		
Parent:	AIDataset		
Explanation:	Collection of models already trained with this dataset@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] ExampleModel1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/History/Model Collection of a model trained with this dataset	[] 2 elements	1..*

3.22 Properties of the SMC “ExampleModel1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 22: Properties of SMC " ExampleModel1"

idShort:	ExampleModel1		
Class:	SubmodelElementCollection		
semanticId:	[IRI] https://admin-shell.io/id/InterOpera/AIDataset/History/Model		
Parent:	History		
Explanation:	Collection of a model trained with this dataset@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Ref] ModelNameplate	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/History/Model/ModeINameplate Reference to the ModelNameplate Submodel of the trained model	[]	1
[Range] TimeStamp	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/History/Model/TimeStamp Temporal range of the data used to train the model	[] ..	1

3.23 Properties of the SMC “BoundaryConditions”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table and figure describe the details of the SMC structure.

Figure 4: UML-Diagram for SubmodelCollection "BoundaryConditions"

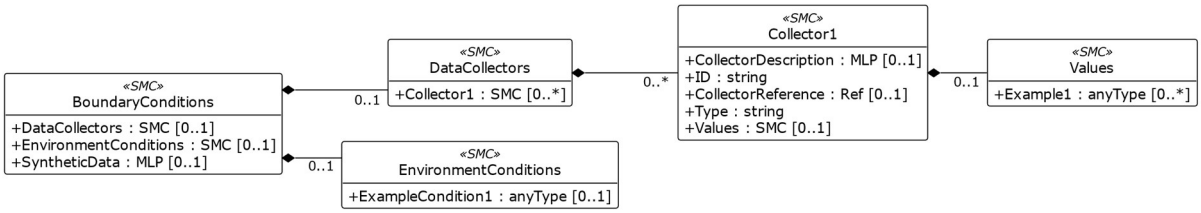

Table 23: Properties of SMC " BoundaryConditions "

idShort:	BoundaryConditions		
Class:	SubmodelElementCollection		
semanticId:	[IRI] https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions		
Parent:	AIDataset		
Explanation:	Boundary conditions in which the dataset was created@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] DataCollectors	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors Information about the data collector (e.g. a sensor)	[] 1 elements	0..1
[SMC] EnvironmentConditions	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/EnviromentConditions Collection about environmental conditions the dataset was created with	[] 1 elements	0..1
[MLP] SyntheticData	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/SyntheticData Information, if the dataset is synthetic, real or mixed	[]	0..1

3.24 Properties of the SMC “DataCollectors”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 24: Properties of SMC " DataCollectors"

idShort:	DataCollectors		

Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors		
Parent:	BoundaryConditions		
Explanation:	Information about the Datacollectors (e.g. a sensor)@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[SMC] Collector1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector Collection with information about a data collector	[] 5 elements	0..*

3.25 Properties of the SMC “Collector1”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 25: Properties of SMC " Collector1"

idShort:	Collector1		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector		
Parent:	DataCollectors		
Explanation:	Collection with information about a data collector@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[MLP] CollectorDescription	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/CollectorDescription Description of the collector	[] @en	0..1
[Property] ID	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/ID Identifier of the collector	[string]	1
[Ref] CollectorReference	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/CollectorReference	[]	0..1

	Reference to an existing Submodel of the collector		
[Property] Type	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/Type Type of collector (e.g. humidity sensor)	[string]	1
[SMC] Values	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/Values Collection of interesting values of the collector	[] 1 elements	0..1

3.26 Properties of the SMC “Values”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 26: Properties of SMC " Values"

idShort:	Values		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/Values		
Parent:	Collector		
Explanation:	Collection of interesting values of the collector@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] Example1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/DataCollectors/Collector/Values/CollectorValue Interesting value of the collector	[anyType]	0..*

3.27 Properties of the SMC “EnvironmentConditions”

Figure 1: UML-Diagram for Submodel "AI-Dataset" shows the UML-diagram defining the relevant properties which need to be set. The following table describes the details of the SMC structure.

Table 27: Properties of SMC " EnvironmentConditions "

idShort:	EnvironmentConditions		
Class:	SubmodelElementCollection		
semanticId:	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/EnviromentConditions		

Parent:	BoundaryConditions		
Explanation:	Collection about environmental conditions the dataset was created with@en		
[SME type]	semanticId = [idType]value	[valueType]	card.
idShort	**Description@en**	**example**	
[Property] ExampleCondition1	[IRI]https://admin-shell.io/id/InterOpera/AIDataset/BoundaryConditions/EnviromentConditions/Condition Condition about the environment.	[anyType]	0..1

Annex A: Explanations on used table formats

General

The used tables in this document try to outline information as concise as possible. They do not convey all information on Submodels and SubmodelElements. For this purpose, the definitive definitions are given by a separate file in form of an AASX file of the Submodel template and its elements.

Tables on Submodels and SubmodelElements

For clarity and brevity, a set of rules is used for the tables for describing Submodels and SubmodelElements.

- The tables follow in principle the same conventions as in [5].
- The table heads abbreviate 'cardinality' with 'card'.
- The tables often place two informations in different rows of the same table cell. In this case, the first information is marked out by sharp brackets [] form the second information. A special case are the semanticIds, which are marked out by the format: (type)(local)[idType]value.
- The types of SubmodelElements are abbreviated: SME

SME type Submodel	Element type
Property	Property
MLP	MultiLanguageProperty
Range	Range
File	File
Blob	Blob
Ref	ReferenceElement
Rel	RelationshipElement
SMC	SubmodelElementCollection

- If an idShort ends with '{00}', this indicates a suffix of the respective length (here: 2) of decimal digits, in order to make the idShort unique. A different idShort might be choosen, as long as it is unique in the parent’s context.
- The Keys of semanticId in the main section feature only idType and value, such as: [IRI]https://admin-shell.io/vdi/2770/1/0/DocumentId/Id. The attributes "type" and "local" (typically "ConceptDescription" and "(local)" or "GlobalReference" and (no-local)") need to be set accordingly; see [6].
- If a table does not contain a column with "parent" heading, all represented attributes share the same parent. This parent is denoted in the head of the table.
- Multi-language strings are represented by the text value, followed by '@'-character and the ISO 639 language code: example@de.
- The [valueType] is only given for Properties.

Bibliography

[1] "Recommendations for implementing the strategic initiative INDUSTRIE 4.0", acatech, April 2013. [Online]. Available: https://www.acatech.de/Publikation/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group/

[2] "Implementation Strategy Industrie 4.0: Report on the results of the Industrie 4.0 Platform"; BITKOM e.V. / VDMA e.V., /ZVEI e.V., April 2015. [Online]. Available: https://www.bitkom.org/noindex/Publikationen/2016/Sonstiges/Implementation-Strategy-Industrie-40/2016-01-Implementation-Strategy-Industrie40.pdf

[3] "The Structure of the Administration Shell: TRILATERAL PERSPECTIVES from France, Italy and Germany", March 2018, [Online]. Available: https://www.plattform-i40.de/I40/Redaktion/EN/Downloads/Publikation/hm-2018-trilaterale-coop.html

[4] "Beispiele zur Verwaltungsschale der Industrie 4.0-Komponente – Basisteil (German)"; ZVEI e.V., Whitepaper, November 2016. [Online]. Available: https://www.zvei.org/presse-medien/publikationen/beispiele-zur-verwaltungsschale-der-industrie-40-komponente-basisteil/

[5] "Verwaltungsschale in der Praxis. Wie definiere ich Teilmodelle, beispielhafte Teilmodelle und Interaktion zwischen Verwaltungsschalen (in German)", Version 1.0, April 2019, Plattform Industrie 4.0 in Kooperation mit VDE GMA Fachausschuss 7.20, Federal Ministry for Economic Affairs and Energy (BMWi), Available: https://www.plattform-i40.de/PI40/Redaktion/DE/Downloads/Publikation/2019-verwaltungsschale-in-der-praxis.html

[6] "Details of the Asset Administration Shell; Part 1 - The exchange of information between partners in the value chain of Industrie 4.0 (Version 3.0RC01)", November 2020, [Online]. Available: https://www.plattform-i40.de/PI40/Redaktion/EN/Downloads/Publikation/Details-of-the-Asset-Administration-Shell-Part1.html

[7] "Semantic interoperability: challenges in the digital transformation age"; IEC, International Electronical Commission; 2019. [Online]. Available:https://basecamp.iec.ch/download/iec-white-paper-semantic-nteroperability-challenges-in-the-digital-transformation-age-en/