



A FAIR Framework for Management of Maritime Datasets to Assist in the Development of Common Maritime Data Bank

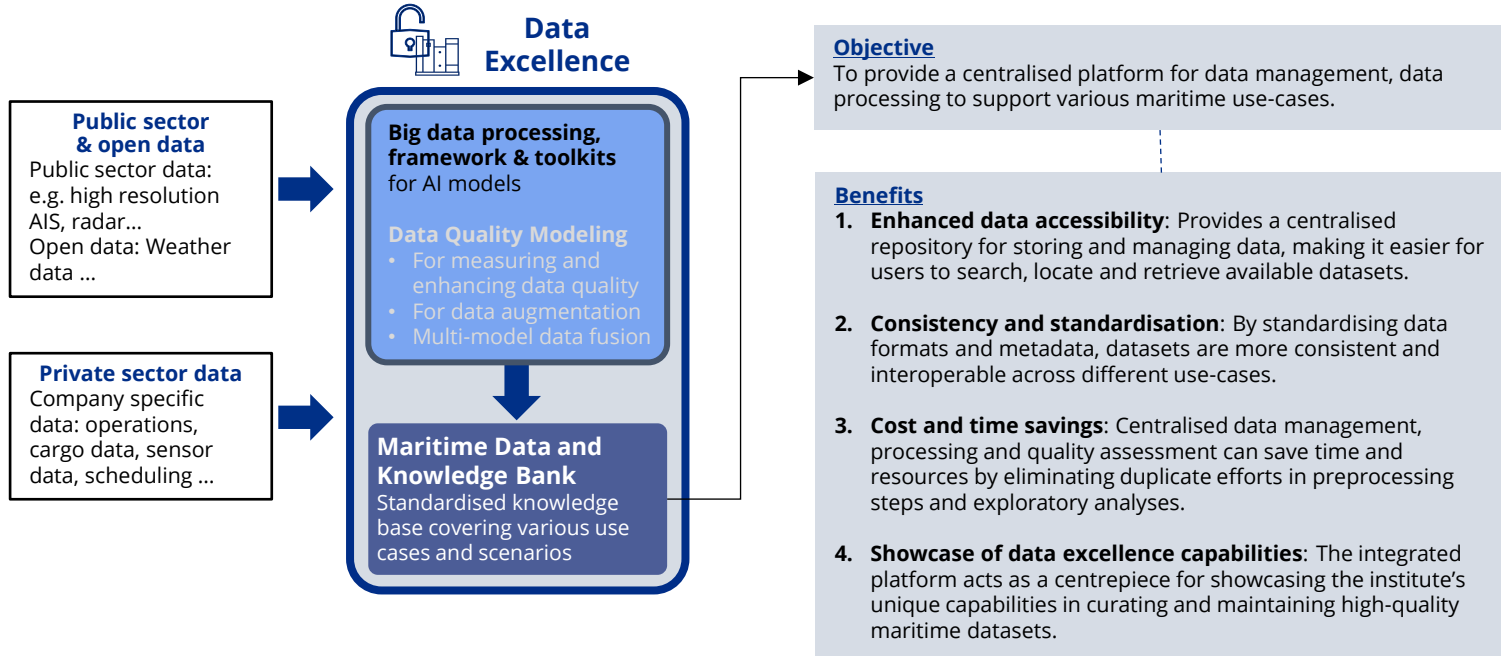
Ping Chong CHUA, Kelvin LEE, Ke WANG, Haiyan XU, Terrence TAN, Xiang LIU, Putu HANGGA, Xiuju FU, Zheng QIN

IHPC, A*STAR
30th October 2024

Agenda

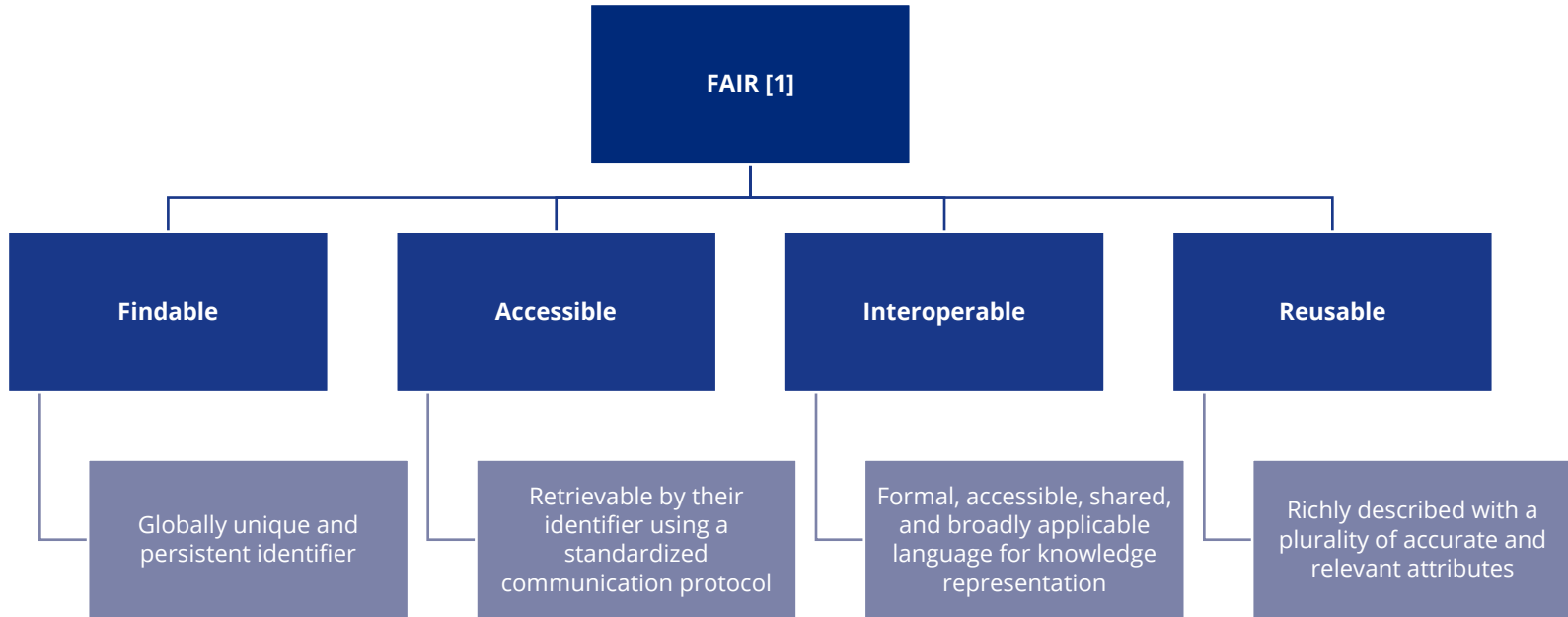
- Overview
- Methodology
- Preliminary Results
 - Predictive maintenance
- Future Work

Overview



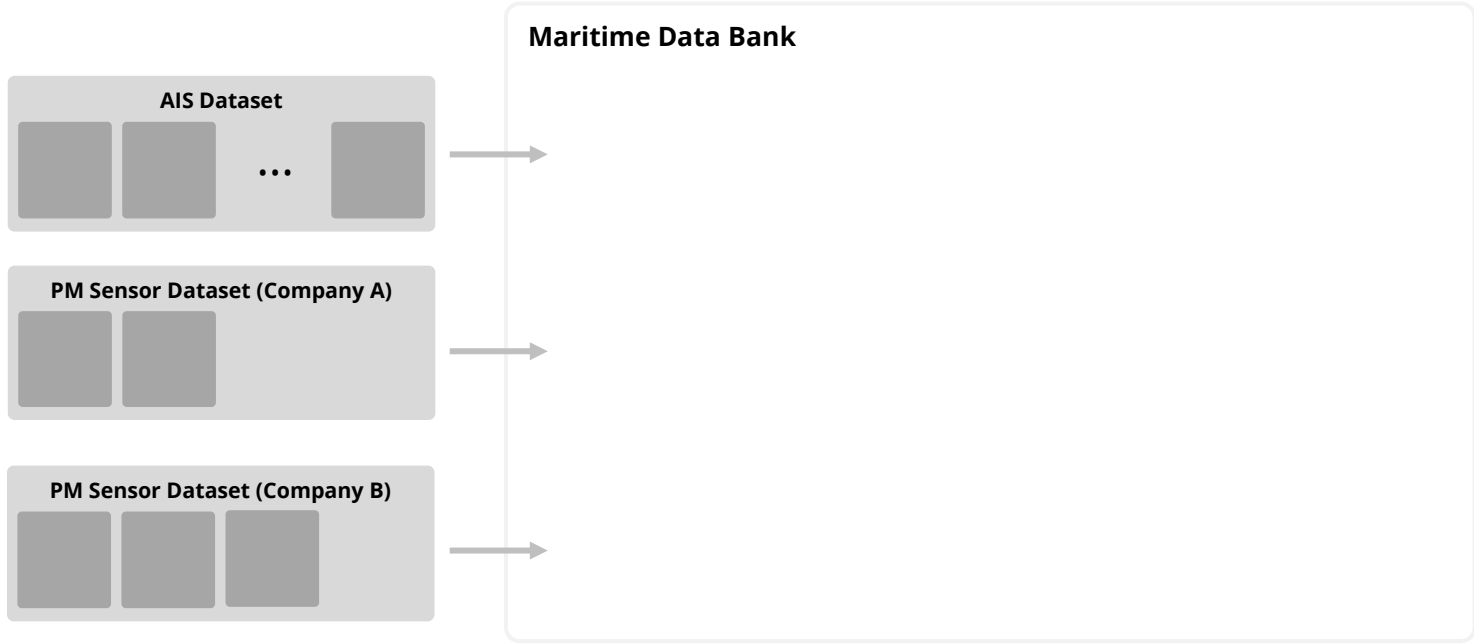
Methodology

- FAIR Principle Towards Development of Maritime Data Bank

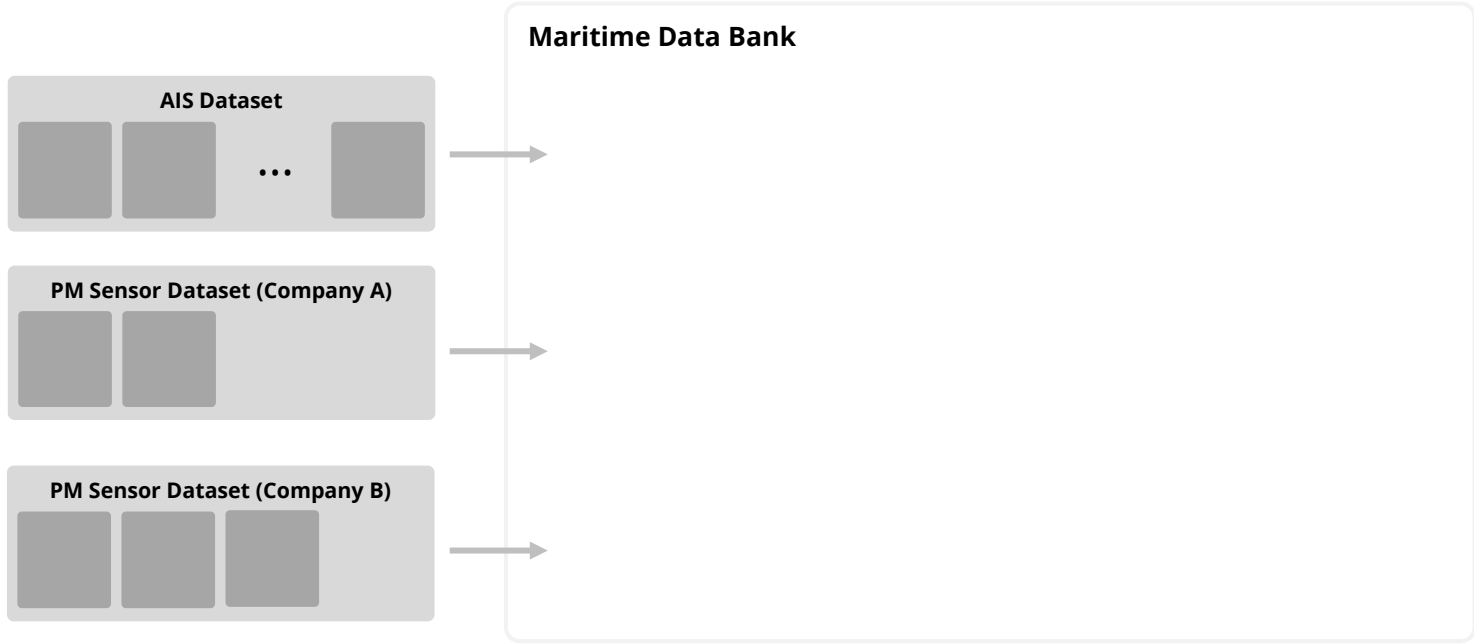


[1] Musen, M. A., O'Connor, M. J., Schultes, E., Martínez-Romero, M., Hardi, J., and Graybeal, J., 2022. Modeling community standards for metadata as templates makes data FAIR. *Sci Data*, **9**: 696. DOI: 10.1038/s41597-022-01815-3

Methodology - Maritime Data Bank

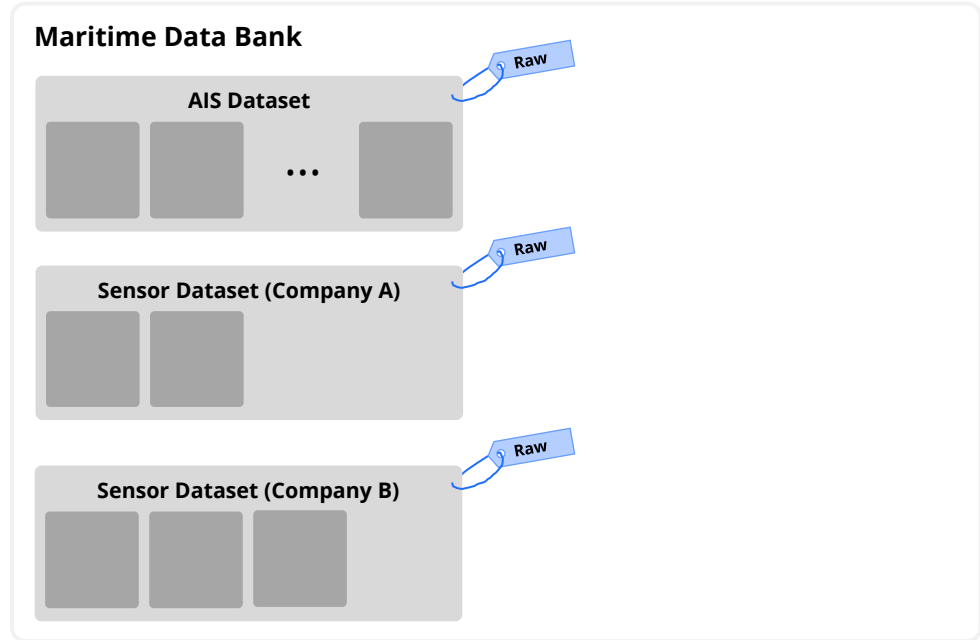


Methodology - Maritime Data Bank



Methodology - Maritime Data Bank

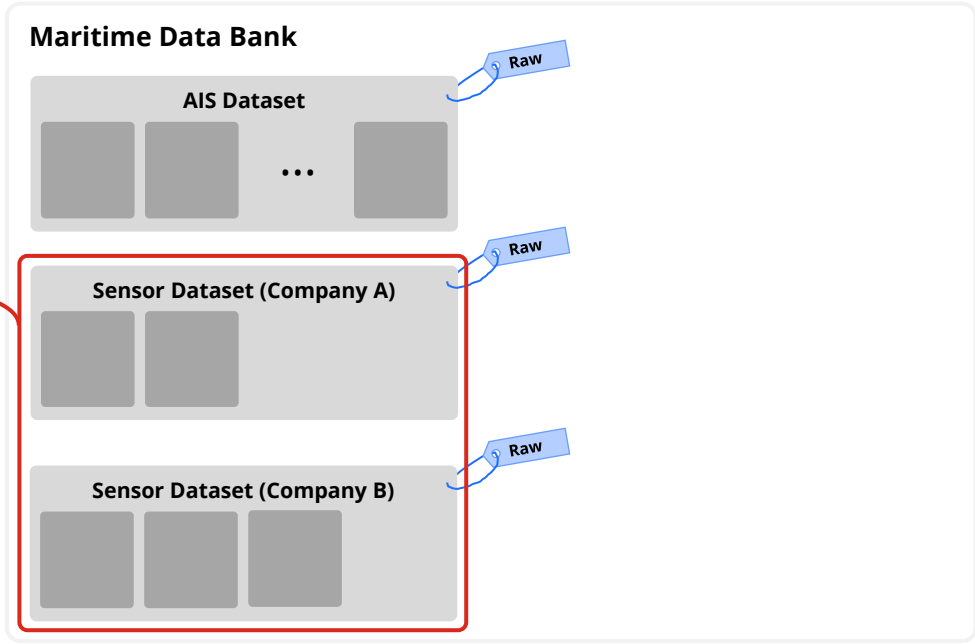
Datasets can be in different forms (tabular/shapes/video, single/multi-file).



Methodology - Maritime Data Bank

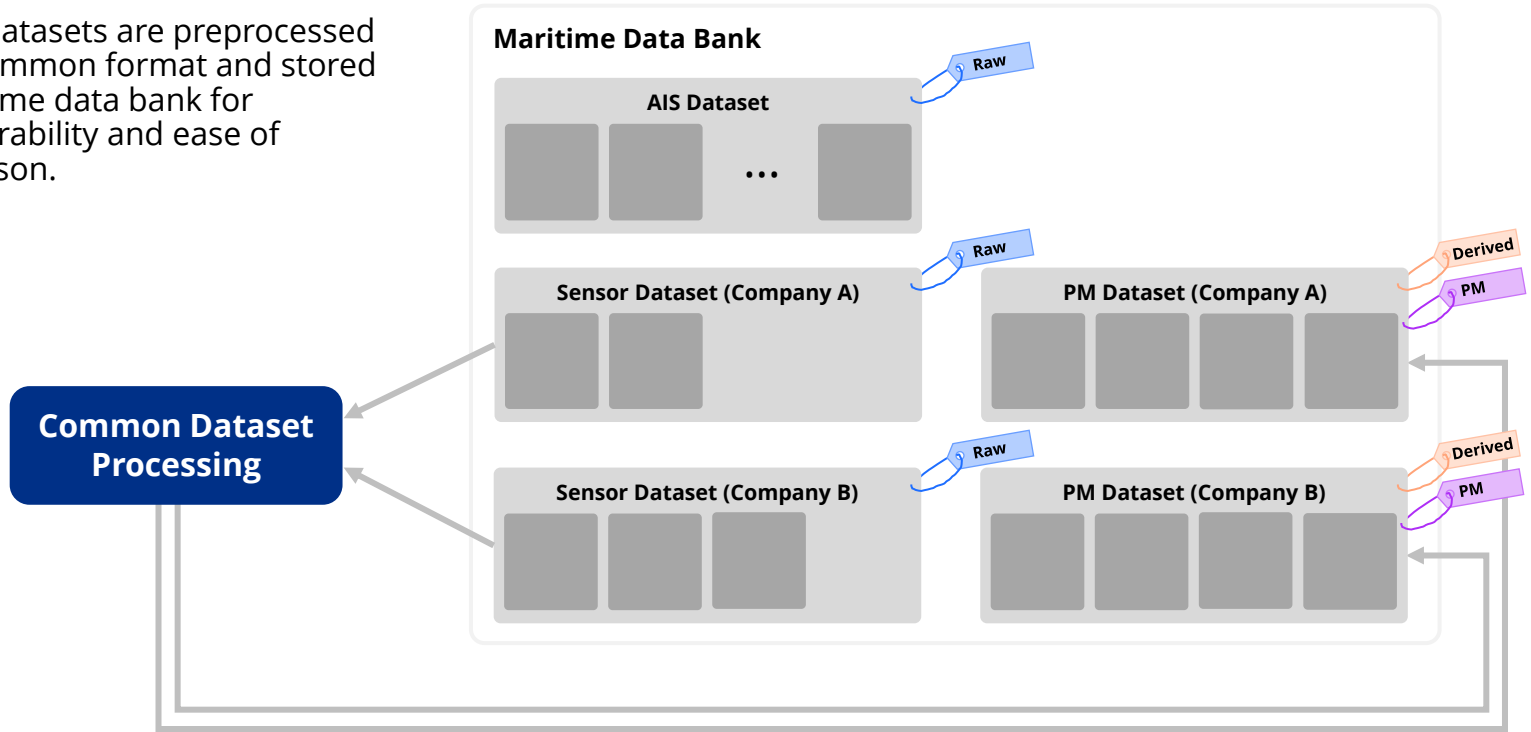


Similar datasets for predictive maintenance use-case



Methodology - Common Data Processing

Similar datasets are preprocessed into a common format and stored in the same data bank for interoperability and ease of comparison.



Bi-level Framework for Describing Maritime Datasets

Maritime Data Bank



⋮



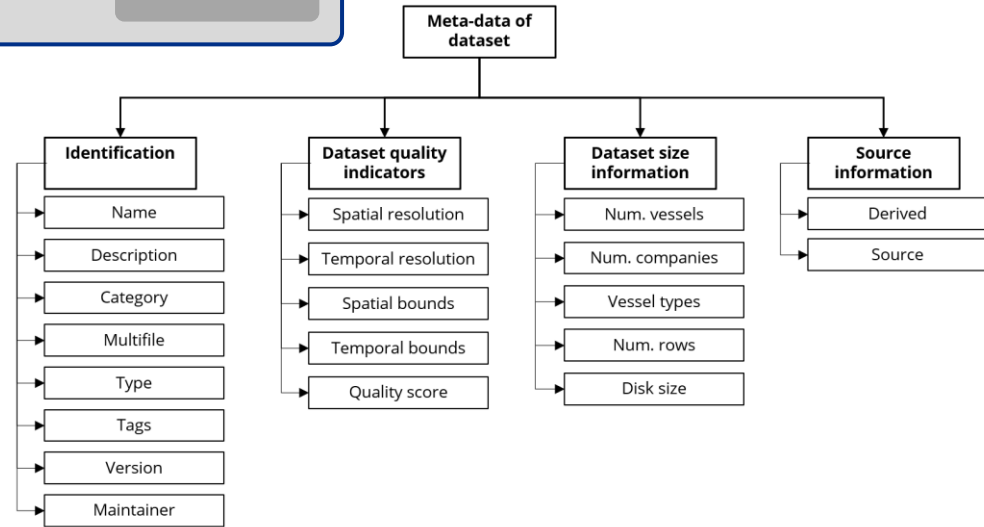
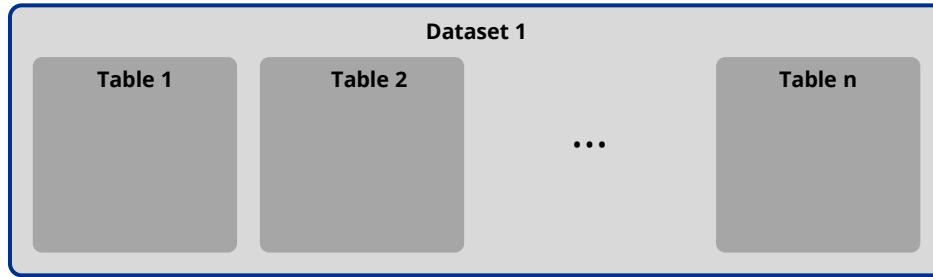
Top-level: Dataset-level Metadata

Unifies various datasets (single/multifile, tabular/shape/file etc.) under a single framework

Bottom-level: Table-level Metadata

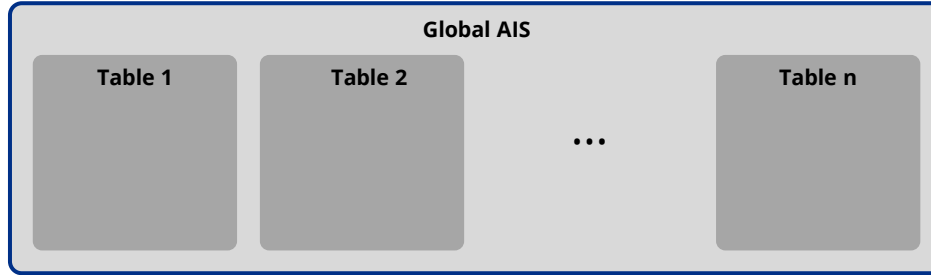
Standardizes the way each table is defined

Top-level: Dataset-level Metadata



Top-level example 1:

Generic Tabular Dataset with Multiple Files



Identification/

General information

Name: Global AIS

Description:

Category: **AIS**

Multifile: **Y**

Type: **Tabular**

Tags:

Version:

Maintainer:

Dataset quality indicators

Spatial resolution:

Temporal resolution:

Spatial bounds:

Temporal bounds:

Quality score:

Dataset size information

Num. vessels:

Num. companies:

Vessel types:

Num. rows:

Disk size:

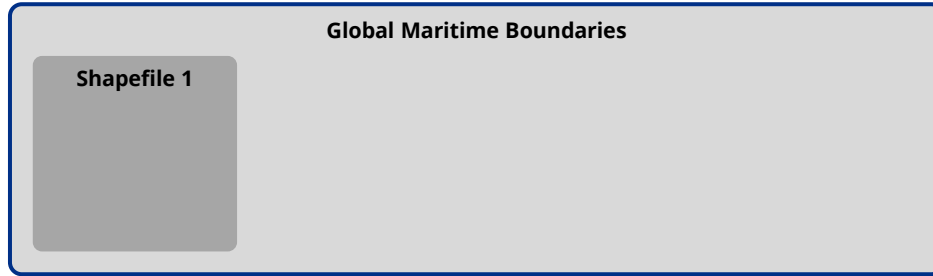
Source information

Derived:

Source:

Top-level example 2:

Generic Shapefile



Identification/

General information

Name: Global MB

Description:

Category: **Boundaries**

Multifile: **N**

Type: **Shapefile**

Tags:

Version:

Maintainer:

Dataset quality indicators

Spatial resolution:

Temporal resolution:

Spatial bounds:

Temporal bounds:

Quality score:

Dataset size information

Num. vessels:

Num. companies:

Vessel types:

Num. rows:

Disk size:

Source information

Derived:

Source:

Top-level example 3:

Generic Derived Shapefile



Identification/

General information

Name: **Asian MB**
 Description:
 Category: Boundaries
 Multifile: N
 Type: Shapefile
 Tags:
 Version:
 Maintainer:

Dataset quality indicators

Spatial resolution:
 Temporal resolution:
 Spatial bounds:
 Temporal bounds:
 Quality score:

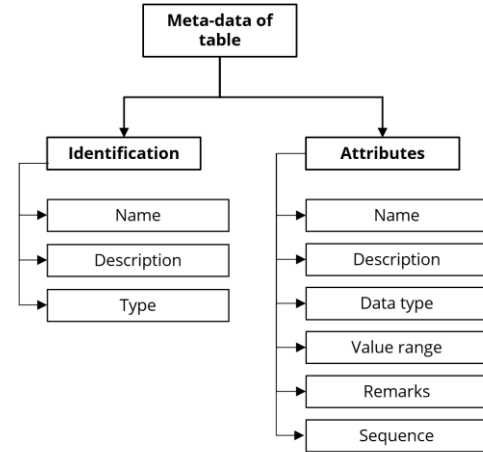
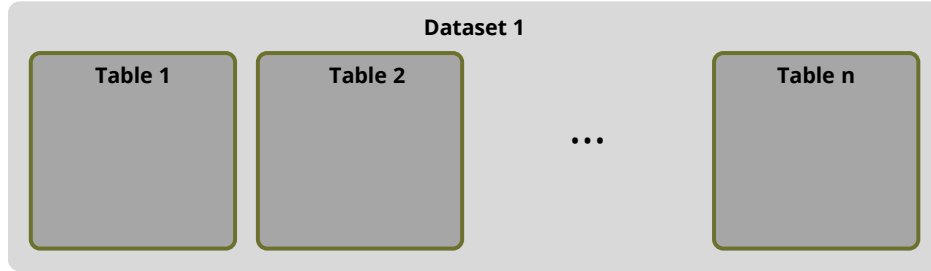
Dataset size information

Num. vessels:
 Num. companies:
 Vessel types:
 Num. rows:
 Disk size:

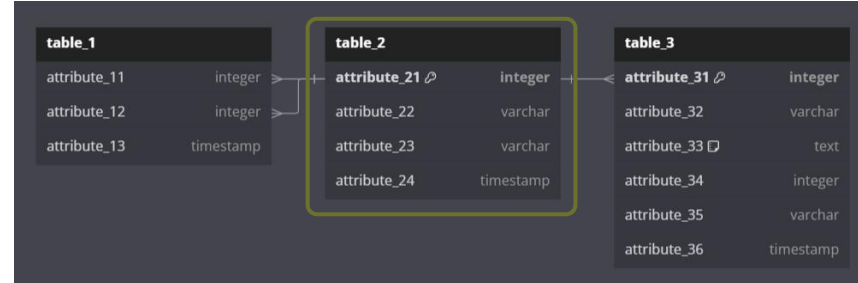
Source information

Derived: **Y**
 Source: **Global MB**

Bottom-level: Table-level Metadata



Bottom-level: Table-level Metadata



Identification

Name: table_2

Description: table_2

Type: Tabular

Attribute Metadata

Name	Description	Data type	Value range	Remarks	Seq
attribute_21	...	integer	-	PK AI	1
attribute_22	...	varchar	-		2
attribute_23	...	varchar	-	-	3
attribute_24	...	string	-	-	4
...

Preliminary Results: Sensor Data



- Dataset-level Metadata

id	name	description	category	tag	doi	version	updated	changelog
1	Tugboat_1_from_20220101	Sensor reading data of main engine of Tugboat_1 for January 2022	Time series	-	-	1.0	12/4/2023	
2	Tugboat_1_from_20220201	Sensor reading data of main engine of Tugboat_1 for February 2022	Time series	-	-	1.0	12/4/2023	
3	Tugboat_1_from_20220301	Sensor reading data of main engine of Tugboat_1 for March 2022	Time series	-	-	1.0	12/4/2023	
4	Tugboat_1_from_20220401	Sensor reading data of main engine of Tugboat_1 for April 2022	Time series	-	-	1.0	12/4/2023	
5	Tugboat_1_from_20220501	Sensor reading data of main engine of Tugboat_1 for May 2022	Time series	-	-	1.0	12/4/2023	
6	Tugboat_1_from_20220601	Sensor reading data of main engine of Tugboat_1 for June 2022	Time series	-	-	1.0	12/4/2023	
7	Tugboat_1_from_20220701	Sensor reading data of main engine of Tugboat_1 for July 2022	Time series	-	-	1.0	12/4/2023	

- Table-level Metadata

id	name	description	dataset_id
1	Tugboat_1_20220101c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 1st January 2022	1
2	Tugboat_1_20220102c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 2nd January 2022	1
3	Tugboat_1_20220103c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 3rd January 2022	1
...
29	Tugboat_1_20220129c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 29th January 2022	1
30	Tugboat_1_20220130c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 30th January 2022	1
31	Tugboat_1_20220131c8391b97c6cb4c25bef4746a58e78b78-.csv	Sensor reading data of Tugboat_1 on 31st January 2022	1
32	Tugboat_1_20220201393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 1st February 2022	2
33	Tugboat_1_20220202393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 2nd February 2022	2
34	Tugboat_1_20220203393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 3rd February 2022	2
...
57	Tugboat_1_20220226393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 26th February 2022	2
58	Tugboat_1_20220227393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 27th February 2022	2
59	Tugboat_1_20220228393cad939d7e4f69bb2a0628695e6bb9-.csv	Sensor reading data of Tugboat_1 on 28th February 2022	2
60	Tugboat_1_2022030137b4b88031f645d28b8b0c0d3493e6b2-.csv	Sensor reading data of Tugboat_1 on 1st March 2022	3
61	Tugboat_1_2022030237b4b88031f645d28b8b0c0d3493e6b2-.csv	Sensor reading data of Tugboat_1 on 2nd March 2022	3
62	Tugboat_1_2022030337b4b88031f645d28b8b0c0d3493e6b2-.csv	Sensor reading data of Tugboat_1 on 3rd March 2022	3
...

Preliminary Results: Sensor Data (Cont'd)



- Attribute-level Metadata

id	name	description	table_id	data_type	value_range	remarks
1	timestamp	Current timestamp	1	DATETIME		
2	Latitude	Latitude location of tugboat	1	FLOAT	[-90,90]	
3	Longitude	Longitude location of tugboat	1	FLOAT	[0,180]	
...
29	SOG	Current speed in knots	1	FLOAT	[0,60]	
...
44	RPM	Current number of revolutions	1	INTEGER	[0.800]	

A Glimpse...

systemmax



System max / maxDataBank

Datasets

FILTER: Noon report Sensor data AIS data Weather Vessel information

DATASET	PURPOSE	FORMAT	MULTIFILE	SPATIAL COV.	TEMPORAL COV.	# RECORDS	QUALITY
AIS data,Vessel Features,Time-Related factors v0.1 Nanyang Technological University	AIS data	Tabular	N	Global	1Y	13,050,017	67
Vessel detection Datasets v0.1 Wuhan University of Technology	AIS data	Tabular	N	Global	1Y	9,991,376	54
Noon_report_from_company v0.1 Simon Fraser University	Noon report	Tabular	N	Global	1Y	4,631,377	-
AIS data v0.1 University of Exeter	AIS data	Tabular	N	Global	1Y	16,821,106	31
AIS data v0.3 Shanghai Maritime University	AIS data	Tabular	N	Regional	1Y	2,183,955	43
Tianjin Port (TJP) dataset v0.1 National University of Singapore	AIS data	Tabular	N	Local	1Y	17,629,377	45
Chengshan Cape (CSC) dataset v0.1 National University of Singapore	AIS data	Tabular	N	Local	1Y	9,939,871	55

A Glimpse...

system max

maxAssessor

maxProcessor

maxDataBank

maxModelStore

maxGPT

System max / maxDataBank

Datasets

FILTER: Noon report **Sensor data** AIS data Weather Vessel information

DATASET	PURPOSE	FORMAT	MULTIFILE	SPATIAL COV.	TEMPORAL COV.	# RECORDS	QUALITY
sensor_on_board_dataset v0.1 University of Shanghai for Science and Technology	Sensor data	Tabular	Y	Global	6M	12,820,029	-

Future Work

1. **Further develop data bank tool**
2. **Enrich the data bank to more applications**
 - Noon report
 - Weather data
3. **Provide guidance/template provided for text-based fields such as description in filling up the metadata.**

Contact Us



Kelvin LEE

Scientist

Email: kelvin_lee@ihpc.a-star.edu.sg



Ping Chong CHUA

Scientist

Email: chua_ping_chong@ihpc.a-star.edu.sg



Xiuju FU

Senior Principal Scientist

Email: fuxj@ihpc.a-star.edu.sg



THANK YOU