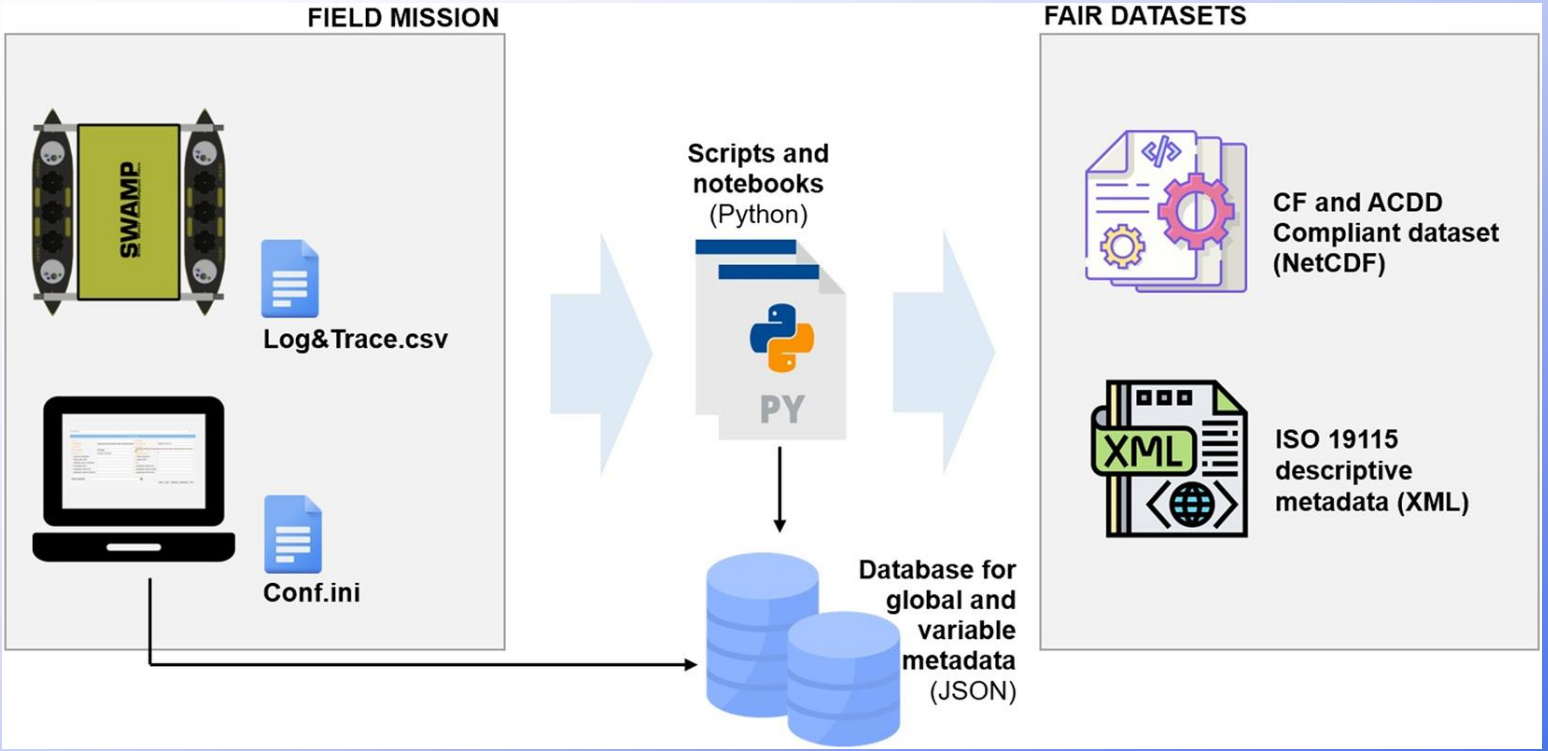


Massimo Caccia
Consiglio Nazionale delle Ricerche
Istituto di Ingegneria del Mare

FAIR data in marine robotics

Marine robotics and sustainability

Keep the Blue in the Green deal

Economic	Environmental	Social
Industry	**Marine science**	**Citizenship**
Robots from special tools to **standard devices**	**Robots** from instrument carriers to **services**	**Robots** from machines to **companions**
Standard procedures for system characterisation: • system dynamics (robotics) • manoeuvring performance (ITTC). Experiment replicability and quantitative performance indexes (GEMs)	New procedures for repeatable, accurate and adaptive in time and space sample and data collection. Robots provide detailed information on when, where and how data are collected.	Engagement of citizens in the research process
	AI-based methodologies for citizens-designed robot control	

Marine robotics: change of perspective

Robots from tools/instruments to services

Industry

Marine science

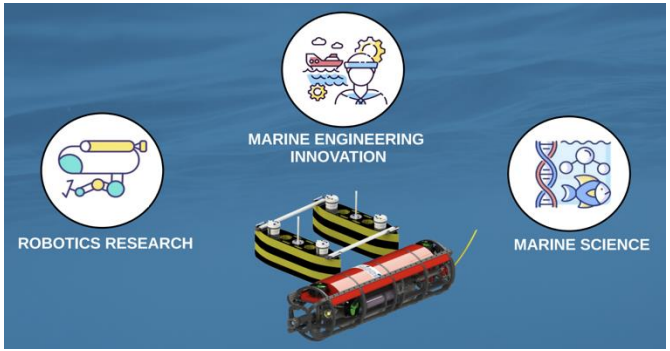
Citizenship

Beneficiaries
End-users

Co-developers

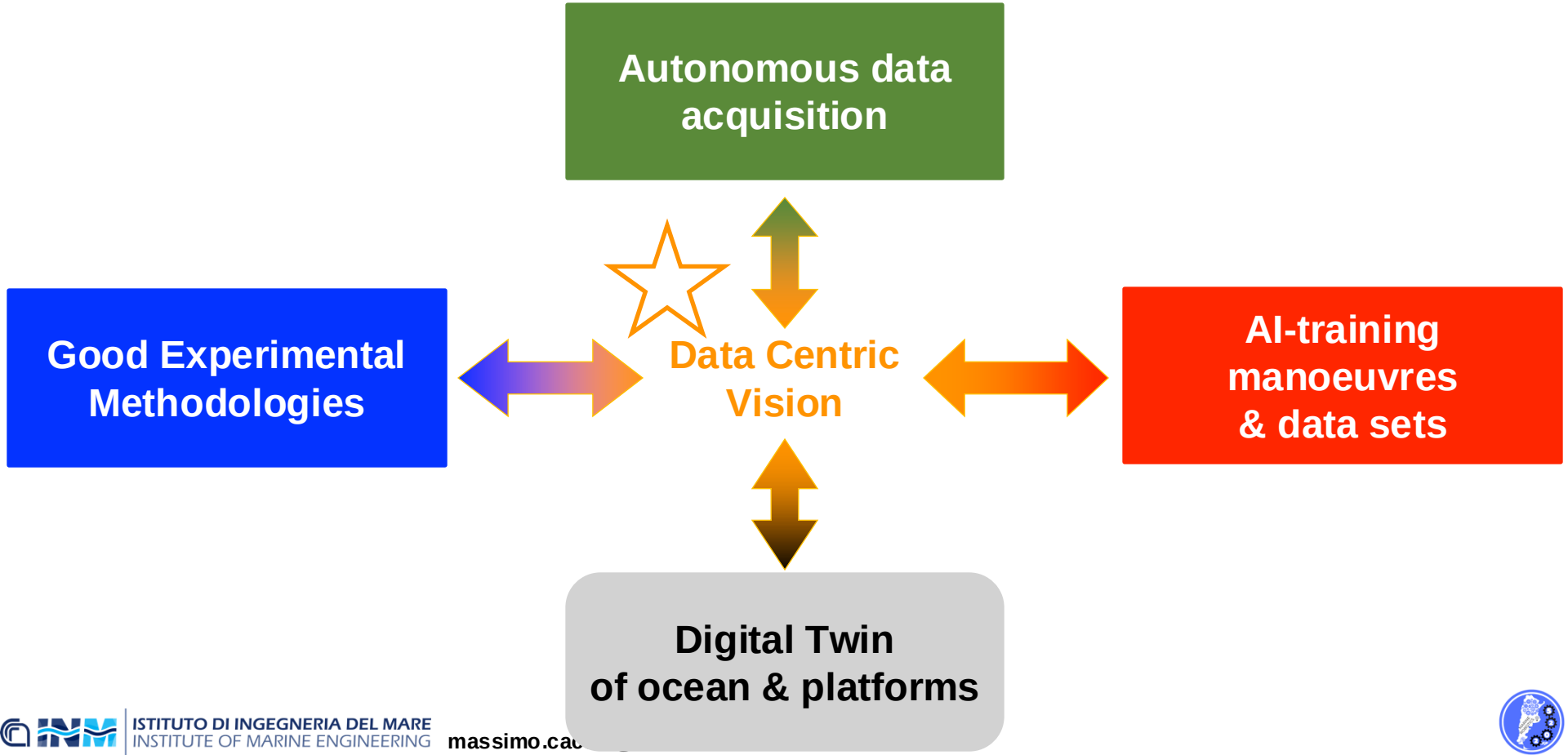
Multidisciplinarity

Interdisciplinarity

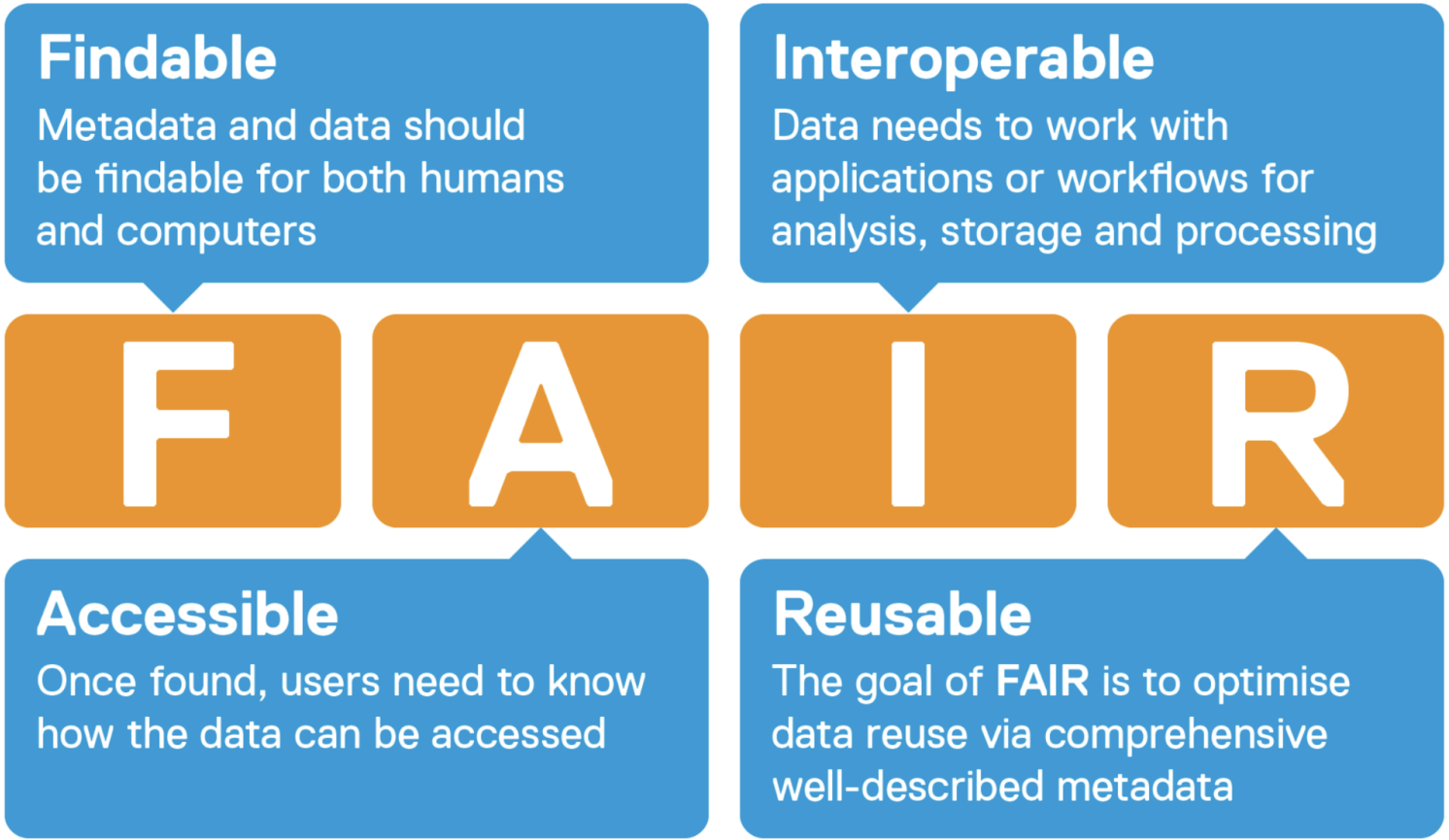


ISTITUTO DI INGEGNERIA DEL MARE
INSTITUTE OF MARINE ENGINEERING
massimo.caccia@cnr.it

Marine robotics: a data-centric vision



Data management: FAIR data



massimo.caccia@cnr.it

Metadata

- **Global metadata**
  - domain agnostic

The name of the
controlled vocabulary
for the variable standard names

Name	Description	ACDD	Auto
Title	A brief title for the dataset	title	no
Abstract	A short summary for dataset, the content and potential linkages etc.	summary	no
keywords	A comma separated list of key words and phrases	keywords	no
Conventions	A comma-separated list of the conventions followed by the dataset.	conventions	no
keywords vocabulary	Guideline for the words/phrases in your "keywords" attribute, if any	keywords_vocabulary	no
PI name	Name of the PI	creator_name	no
PI email	Email to the PI	creator_email	no
PI institution	Affiliation of the PI	institution	no
Dataset start time	ISO8601 reference for the dataset	time_coverage_start	yes
Dataset end time	ISO8601 reference for the dataset	time_coverage_end	yes
Dataset northernmost latitude	Geographical northernmost position of the dataset	geospatial_lat_max	yes
Dataset southernmost latitude	Geographical southernmost position of the dataset	geospatial_lat_min	yes
Dataset latitude units	Further refinement of the box	geospatial_lat_units	yes
Dataset easternmost longitude	Geographical easternmost position of the dataset.	geospatial_lon_min	yes
Dataset westernmost longitude	Geographical westernmost position of the dataset.	geospatial_lon_max	yes
Dataset longitude units	Further refinement of the box	geospatial_lon_units	yes
date created	The date on which the data was created.	date_created	yes
Platform	Name of the platform(s) that supported the sensor data.	platform	no
Project long name	The scientific project that produced the data.	project	no
License	Describe the restrictions to data access and distribution	license	no
Dataset version	Version identifier of the data file or product.	product_version	no
Data center (URL)	URL to the data center hosting the data	publisher_url	no
Creator (URL)	URL to creator or to information	creator_url	no
standard_name_vocabulary	The name of the controlled vocabulary for the variable standard names	standard_name_vocabulary	no
time_coverage_duration	Describes the duration of the dataset. Use ISO 8601:2019 format	time_coverage_duration	yes
time_coverage_resolution	Describes the time period between each value. Use ISO 8601:2019 format	time_coverage_resolution	yes
id	An identifier for the dataset, it can be the DOI as well	id	no
processing_level	A textual description of the processing level of the data	processing_level	no

Variable metadata: domain dependent

- the nomenclature used for metadata and for the description of variables and their attributes has to comply with a **controlled vocabulary**
- a shared/**controlled vocabulary**, for a given discipline, ideally contains the standard names of all the variables that can be stored in a **FAIR** dataset

massimo.caccia@cnr.it

Log file example

given by the robot NGC system

given by the adopted Vocabulary

given by the user

log_name	…	robot_lat	robot_lon	…
standard_name	…	latitude	longitude	…
…	…	…	…	…
value at time t	…	42.43555800	18.76497800	…
…	…	…	…	…

Marine robotics vocabulary sources

- does a Vocabulary exist whose standard names include all the ones needed by marine robotics?
- how do we choose additional names, called long names?

Science

CF MetaData Conventions Discussion Governance

CF Metadata Conventions

cfconventions.org/Data/cf-standard-names/current/build/cf-standard-name-table.html

CF Standard Name Table

Version 85, 21 May 2024

Ship architecture

INTERNATIONAL TOWING TANK CONFERENCE

ITTC Symbols and Terminology List

Version 2021

June 2021

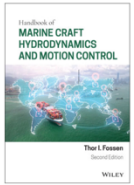
Supersedes all previous versions

SNAME (1950). Nomenclature for Treating the Motion of a Submerged Body Through a Fluid. *The Society of Naval Architects and Marine Engineers, Technical and Research Bulletin No. 1-5, April 1950, pp. 1-15.*

Cybernetics

Handbook of Marine Craft Hydrodynamics and Motion Control

The "Handbook of Marine Craft Hydrodynamics and Motion Control" offers a comprehensive exploration of cutting-edge developments in the fields of marine craft hydrodynamics and guidance, navigation, and control (GNC) systems. The text establishes how mathematical models and modern control theory can be implemented to simulate and verify control systems.



Errata.pdf

ISBN: 978-1-119-57505-4

Thor I. Fossen

John Wiley & Sons Ltd., 2nd Edition, April 2021

Lecture Material

massimo.caccia@cnr.it

Earth Sciences: Climate Forecast convention

- atmospheric chemistry, atmosphere dynamics, carbon cycle, cloud, hydrology, ocean dynamics, radiation, sea ice, surface
- the CF metadata conventions are designed to promote the processing and sharing of files created with the **NetCDF API**

Network Common Data Form (NetCDF)

NetCDF (Network Common Data Form) is a set of software libraries and machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data. It is also a community standard for sharing scientific data. The Unidata Program Center supports and maintains netCDF programming interfaces for C, C++, Java, and Fortran. Programming interfaces are also available for Python, IDL, MATLAB, R, Ruby, and Perl.

Data in netCDF format is:

- **Self-Describing.** A netCDF file includes information about the data it contains.
- **Portable.** A netCDF file can be accessed by computers with different ways of storing integers, characters, and floating-point numbers.
- **Scalable.** Small subsets of large datasets in various formats may be accessed efficiently through netCDF interfaces, even from remote servers.
- **Appendable.** Data may be appended to a properly structured netCDF file without copying the dataset or redefining its structure.
- **Sharable.** One writer and multiple readers may simultaneously access the same netCDF file.
- **Archivable.** Access to all earlier forms of netCDF data will be supported by current and future versions of the software.

Earth Sciences: Climate Forecast convention

Guidelines for Construction of CF Standard Names

Version 1, 3 December 2008

- Component
  - the **direction of the spatial component of a vector** is indicated by one of the words **upward**, **downward**, **northward**, **southward**, **eastward**, **westward**, **x**, **y**. The last two indicate directions along the horizontal grid being used when they are not true longitude and latitude (if there is a rotated pole, for instance).

Units	Generic Name
m2	area
kg m-3	density
J	energy
s-1	frequency
kg	mass
kg m-2 s-1	mass_flux
s	period
W	power
Pa	pressure
1	probability
m s-1	speed
K	temperature
m s-1	velocity
m3	volume

Earth Sciences: Climate Forecast convention

Guidelines for Construction of CF Standard Names

Version 1, 3 December 2008

- Generic Names
  - the following names are used with consistent meanings and units as elements in other standard names, although they are themselves too general to be chosen as standard names. They are recorded here for reference only. ***These are not standard names***.

Earth Sciences: Climate Forecast convention

- an example: temperature

CF Standard Name Table

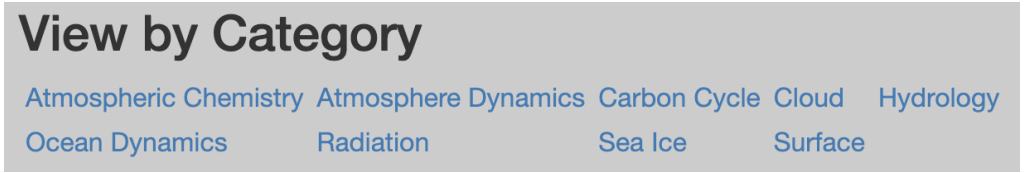
Version 85, 21 May 2024

sea_water_temperature	Sea water temperature is the in situ temperature of the sea water. To specify the depth at which the temperature applies use a vertical coordinate variable or scalar coordinate variable. There are standard names for sea_surface_temperature, sea_surface_skin_temperature, sea_surface_subskin_temperature and sea_surface_foundation_temperature which can be used to describe data located at the specified surfaces. For observed data, depending on the period during which the observation was made, the measured in situ temperature was recorded against standard "scales". These historical scales include the International Practical Temperature Scale of 1948 (IPTS-48; 1948-1967), the International Practical Temperature Scale of 1968 (IPTS-68, Barber, 1969; 1968-1989) and the International Temperature Scale of 1990 (ITS-90, Saunders 1990; 1990 onwards). Conversion of data between these scales follows t68 = t48 - (4.4 x 10e-6) * t48(100 - t - 48); t90 = 0.99976 * t68. Observations made prior to 1948 (IPTS-48) have not been documented and therefore a conversion cannot be certain. Differences between t90 and t68 can be up to 0.01 at temperatures of 40 C and above; differences of 0.002-0.007 occur across the standard range of ocean temperatures (-10 - 30 C). The International Equation of State of Seawater 1980 (EOS-80, UNESCO, 1981) and the Practical Salinity Scale (PSS-78) were both based on IPTS-68, while the Thermodynamic Equation of Seawater 2010 (TEOS-10) is based on ITS-90. References: Barber, 1969, doi: 10.1088/0026-1394/5/2/001; UNESCO, 1981; Saunders, 1990, WOCE Newsletter, 10, September 1990. It is strongly recommended that a variable with this standard name should have a units_metadata attribute, with one of the values "on-scale" or "difference", whichever is appropriate for the data, because it is essential to know whether the temperature is on-scale (meaning relative to the origin of the scale indicated by the units) or refers to temperature differences (implying that the origin of the temperature scale is irrevelant), in order to convert the units correctly (cf. https://cfconventions.org/cf-conventions/cf-conventions.html#temperature-units).	K

massimo.caccia@cnr.it

CF Standard Name table

- are there robotics-related standard names?
  - **there is no specific category**



  - "robot", "vessel", "ship", "boat" searches provide no results
  - while searching "**platform**" you obtain results
    - a "**platform**" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes,ships, buoy, instruments, ground stations, and masts.

CF Standard Name table: latitude and longitude

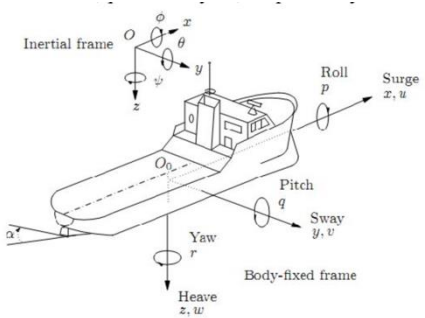
latitude		
Latitude is positive northward; its units of degree_north (or equivalent) indicate this explicitly. In a latitude-longitude system defined with respect to a rotated North Pole, the standard name of grid_latitude should be used instead of latitude. Grid latitude is positive in the grid-northward direction, but its units should be plain degree.	degree_north	latitude
longitude		
Longitude is positive eastward; its units of degree_east (or equivalent) indicate this explicitly. In a latitude-longitude system defined with respect to a rotated North Pole, the standard name of grid_longitude should be used instead of longitude. Grid longitude is positive in the grid-eastward direction, but its units should be plain degree.	degree_east	longitude

CF Standard Name table: x & y coordinates

projection_x_coordinate	
"x" indicates a vector component along the grid x-axis, when this is not true longitude, positive with increasing x. Projection coordinates are distances in the x- and y-directions on a plane onto which the surface of the Earth has been projected according to a map projection. The relationship between the projection coordinates and latitude and longitude is described by the grid_mapping.	m
projection_y_coordinate	
"y" indicates a vector component along the grid y-axis, when this is not true latitude, positive with increasing y. Projection coordinates are distances in the x- and y-directions on a plane onto which the surface of the Earth has been projected according to a map projection. The relationship between the projection coordinates and latitude and longitude is described by the grid_mapping.	m

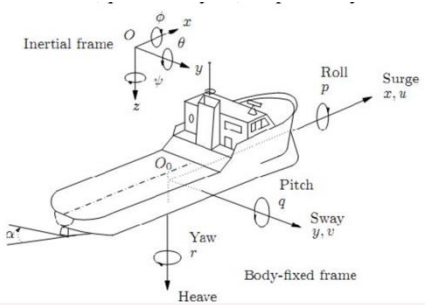
depth	
Depth is the vertical distance below the surface.	m

CF Standard Name table: surge rate



platform_surge_rate	"Surge rate" is the rate of displacement along an axis that is perpendicular to the local vertical axis and is coplanar with the nominal forward motion direction of the platform. Surge rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. The standard name platform_surge_rate should be chosen only if the sign convention of the data is unknown. For cases where the sign convention of the surge rate is known, a standard name of platform_surge_rate_fore or platform_surge_rate_aft should be chosen, as appropriate. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1
platform_surge_rate_aft	"Surge rate" is the rate of displacement along an axis that is perpendicular to the local vertical axis and is coplanar with the nominal forward motion direction of the platform. Surge rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. "Aft" indicates that positive values of surge rate represent the platform moving backward as viewed by an observer on top of the platform facing forward. The standard name platform_surge_rate_fore should be used for data having the opposite sign convention. The standard name platform_surge_rate should be chosen only if the sign convention of the data is unknown. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1
platform_surge_rate_fore	"Surge rate" is the rate of displacement along an axis that is perpendicular to the local vertical axis and is coplanar with the nominal forward motion direction of the platform. Surge rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. "Fore" indicates that positive values of surge rate represent the platform moving forward as viewed by an observer on top of the platform facing forward. The standard name platform_surge_rate_aft should be used for data having the opposite sign convention. The standard name platform_surge_rate should be chosen only if the sign convention of the data is unknown. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1

CF Standard Name table: sway rate



platform_sway_rate	"Sway rate" is the rate of displacement along an axis that is perpendicular to both the local vertical axis and the nominal forward motion direction of the platform. Sway rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. The standard name platform_sway_rate should be chosen only if the sign convention of the data is unknown. For cases where the sign convention of the sway rate is known, a standard name of platform_sway_rate_starboard or platform_sway_rate_port should be chosen, as appropriate. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1
platform_sway_rate_port	"Sway rate" is the rate of displacement along an axis that is perpendicular to both the local vertical axis and the nominal forward motion direction of the platform. Sway rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. "Port" indicates that positive values of sway rate represent the platform moving left as viewed by an observer on top of the platform facing forward. The standard name platform_sway_rate_starboard should be used for data having the opposite sign convention. The standard name platform_sway_rate should be chosen only if the sign convention of the data is unknown. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1
platform_sway_rate_starboard	"Sway rate" is the rate of displacement along an axis that is perpendicular to both the local vertical axis and the nominal forward motion direction of the platform. Sway rate might not include changes to the "at rest" position of the platform with respect to the axis of displacement, which may change over time. "Starboard" indicates that positive values of sway rate represent the platform moving right as viewed by an observer on top of the platform facing forward. The standard name platform_sway_rate_port should be used for data having the opposite sign convention. The standard name platform_sway_rate should be chosen only if the sign convention of the data is unknown. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1

CF Standard Name table: course and speed

platform_course	
Course is the clockwise angle with respect to North of the nominal forward motion direction of the platform (not necessarily the same as the direction in which it is pointing, called "platform_orientation"). A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	degree
platform_speed_wrt_ground	
Speed is the magnitude of velocity. The abbreviation "wrt" means with respect to. The platform speed with respect to ground is relative to the solid Earth beneath it, i.e. the sea floor for a ship. It is often called the "ground speed" of the platform. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1
platform_speed_wrt_sea_water	
Speed is the magnitude of velocity. The abbreviation "wrt" means with respect to. A "platform" is a structure or vehicle that serves as a base for mounting sensors. Platforms include, but are not limited to, satellites, aeroplanes, ships, buoys, instruments, ground stations, and masts.	m s-1

massimo.caccia@cnr.it

CF Standard Name table: course and speed

- Hydrodynamics force and moment involve the linear and angular velocities of the vehicle with respect to the water
  - platform_surge_rate_fore_wrt_sea_water
  - platform_sway_rate_starboard_wrt_sea_water
    - are not part of CF Standard Name Table
    - **marine robotics vocabulary** should include them as **long names**

massimo.caccia@cnr.it

Platform dynamics: force and moment

- **CF Standard Name Table** does not include platform dynamics data
- **ITTC Symbols and Terminology List** includes force and moment

K, M_x , F^1_1 , F_4	K, M(1), F1(1), F(4)	Moment around body axis x		Nm
M, M_y , F^1_2 , F_5	M, M(2), F1(2), F(5)	Moment around body axis y		Nm
N, M_z , FN^1_3 , F_6	N, M(3), F1(3), F(6)	Moment around body axis z		Nm
X, F_x , F^0_1 , F_1	X, FX, F0(1), F(1)	Force in direction of body axis x		Nm
Y, F_y , F^0_2 , F_2	Y, FY, F0(2), F(2)	Force in direction of body axis y		Nm
Z, F_z , F^0_3 , F_3	Z, FZ, F0(3), F(3)	Force in direction of body axis z		Nm

Platform dynamics: force and moment

- Marine Robotics add-ons to CF vocabulary
  - platform_surge_force_fore
  - platform_sway_force_starboard
  - platform_heave_force_down
  - platform_roll_moment
  - platform _pitch_moment
  - platform_yaw_moment

K, M_x, F^1_1, F_4	K, M(1), F1(1), F(4)	Moment around body axis x
M, M_y, F^1_2, F_5	M, M(2), F1(2), F(5)	Moment around body axis y
N, M_z, FN^1_3, F_6	N, M(3), F1(3), F(6)	Moment around body axis z
X, F_x, F^0_1, F_1	X, FX, F0(1), F(1)	Force in direction of body axis x
Y, F_y, F^0_2, F_2	Y, FY, F0(2), F(2)	Force in direction of body axis y
Z, F_z, F^0_3, F_3	Z, FZ, F0(3), F(3)	Force in direction of body axis z

Guidelines for Construction of CF Standard Names

Version 1, 3 December 2008

INTERNATIONAL TOWING TANK CONFERENCE

ITTC Symbols and Terminology List

Version 2021

June 2021

Supersedes all previous versions

ISTITUTO DI INGEGNERIA DEL MARE
INSTITUTE OF MARINE ENGINEERING
massimo.caccia@cnr.it

FAIR marine robots

www.nature.com/scientificdata/

www.nature.com/scientificdata

scientific data

OPEN

ARTICLE

A framework for FAIR robotic datasets

Corrado Motta[1,2,5], Simona Aracri[1,5✉], Roberta Ferretti[1,5], Marco Bibuli[1], Gabriele Bruzzone[1], Massimo Caccia[1], Angelo Odetti[1], Fausto Ferreira[3] & Francesca de Pascalis[4]

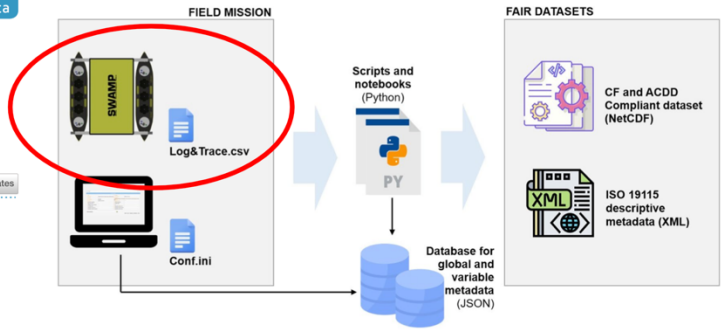


Fig. 1 Schema for the automation and application of the method.

1. *Log&Trace.csv*: the vehicle's telemetry data, in the form of a raw log table, provided by SWAMP
2. *Conf.ini*: the descriptive metadata, in the form of a configuration file, generated by the Human-Computer Interface (HCI) used to remotely control the vehicle

The telemetry file is a text file storing the values of each variable in a dedicated column. For each column, we set a 2-level name, followed by all the values. The first level name is the custom name given to a specific variable (what we used to do already), whilst the second level name is the *long_name*, which indicates the standard variable name it refers to. Since the *long_name* is also the unique ID of the variable database, such information can be used by the scripts to retrieve from the database all the attributes of the pointed variable. It is fundamental to use a 2-level name for the columns, as it is very common to have multiple measurements for the same type of data. For example, SWAMP contains more than one Global Navigation Satellite System (GNSS) on board. One

Log&Trace.csv file construction

given by the robot NGC system

given by the adopted Vocabulary

given by the user

log_name	…	robot_lat	robot_lon	…
long_name	…	latitude	longitude	…
…	…	…	…	…
value at time t	…	42.43555800	18.76497800	…
…	…	…	…	…

Log name construction: an example

- **vehicle position expressed in latitude and longitude**
  - estimated vehicle position
  - measured vehicle position

log_name	estimated_latitude	estimated_longitude	measured_latitude	measured_longitude
long_name	latitude	longitude	latitude	longitude

- **vehicle position can be measured by different sensors**
  - measured vehicle position by ZED-F9P GNSS module
    measured vehicle position by NEO-M9N

log_name	ZED_F9P_measured_latitude	ZED_F9P_measured_longitude	NEO_M9N_measured_latitude	NEO_M9N_measured_longitude
long_name	latitude	longitude	latitude	longitude

massimo.caccia@cnr.it