



Experimental WIS 2.0 data exchange for data in WMO CF-NetCDF profiles

Kai-Thorsten Wirt (DWD)
Antje Schremmer (DWD)
Kevin O`Brian (NOAA)

WMO OMM

World Meteorological Organization

Organisation météorologique mondiale

Background

- Task Team on CF-NetCDF
 - Developed WMO extensions to CF-NetCDF conventions
 - Test and Evaluate CF-NetCDF for Ocean and Radar data
 - Asking for easy distribution and access to available data for users
- WMO Members
 - Whats the TTAAii for this?
 - Can we have that on the GTS?
 - We need to extend the Manual on the GTS to include a new T1
- WIS2 Group
 - Working on retirement of TTAAii headers and manual routing
 - Does it make sense to add new data in new data formats to something we want to retire
 - How can we make everybody happy :-)
- Idea
 - Have a gateway that is collecting the data and metadata and that is implementing WIS2 protocols



WMO OMM

Project objectives

Experiment with international real-time exchange of data

- using pub/sub protocols (and topic structure from TT-Protocols)
- distributed storage of the original data along with a central, mirrored repository
- discoverable using WIS Metadata
- assesses the advantages, disadvantages and usability of WMO Core Metadata Profile 1.3
- exposure of this data to commercial search engines.



WMO OMM

Project team

- Antje Schremmer (DWD)
- Axel Anderson (DWD)
- Daniel Michelson (ECCC)
- David Berry (NOC, UK)
- Hermann Asensio (DWD)
- Kai Wirt (DWD)
- Kevin O’Brien (U Washington, NOAA)
- Mark Curtis (BoM)

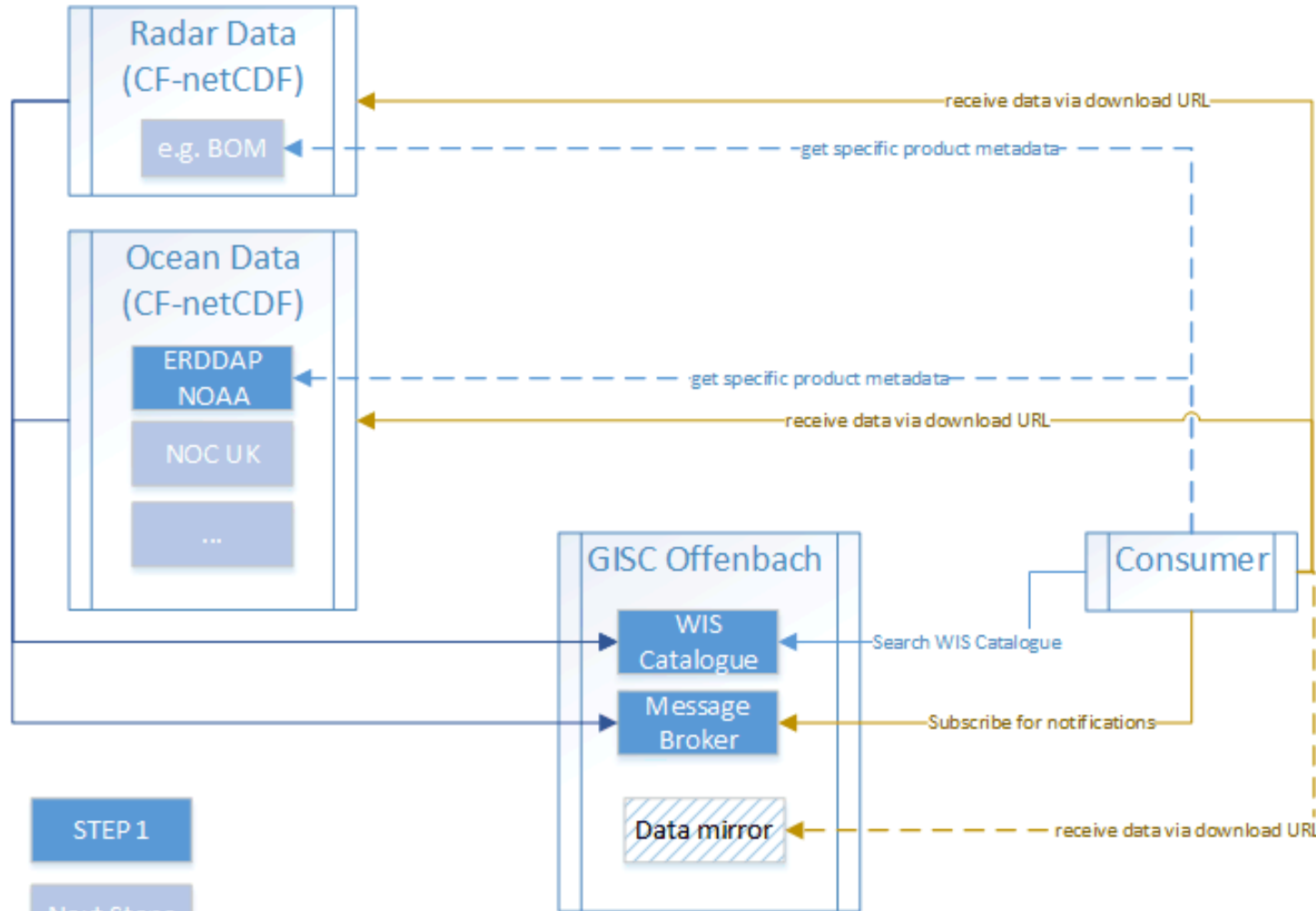
WMO Secretariat

- Dean Lockett
- Xiaoxia Chen



WMO OMM

Project plan (Overview)



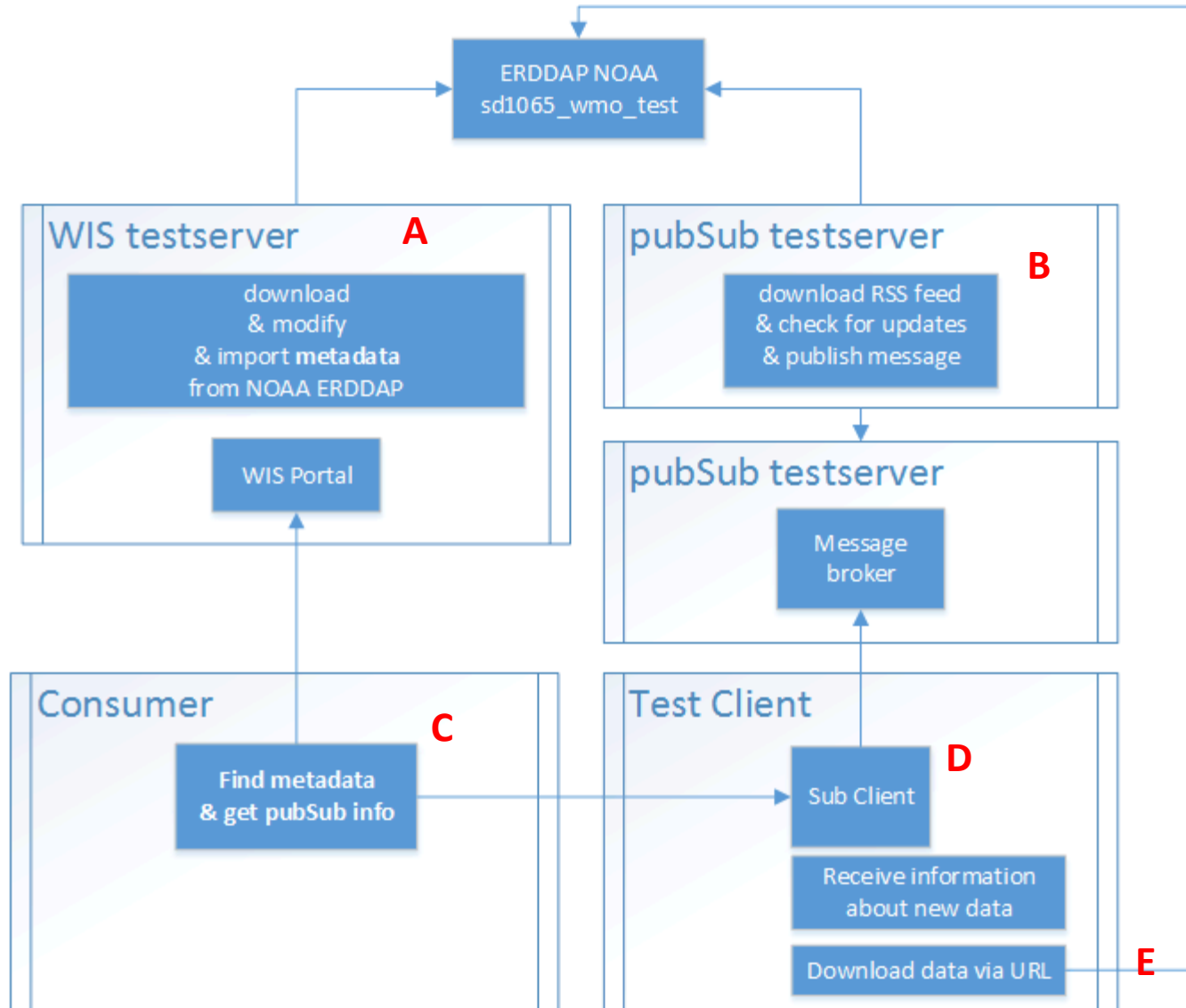
STEP1
share ocean / marine data
using the appropriate CF
profile this autumn / fall

STEP2
Other ocean data

NEXT STEPS
Radar data (Spring 2022)
data is more complicated in
both terms of instrumental
metadata and data policy



Project plan (Overview STEP1)



Parts

- A → get Metadata
- B → publish message

Client

- C → search for product
- D → subscribe to topic
- E → receive data



Project plan (Milestones/ Deliverables)

Part	Task	Status
A	Get Metadata from NOAA ERDDAP	✓
	Modify Metadata (WMCP is based on older ISO version)	✓
	Add pubSub information under transferOptions	✓
	Import Metadata in WIS Catalogue	✓
B	Receive RSS feed	✓
	Check for updates	✓
	Download data for integrity value and publish message (using data URL from ERDDAP)	✓
C	Metadata searchable in WIS Portal	✓
	Metadata searchable in commercial search engines	✓
	Metadata view shows additional transferOptions	✓



Project plan (Milestones/ Deliverables)

Part	Task	Status
D	Subscribe client script	✓
	Client received messages after sub to topic	✓
E	Client downloaded data from ERDDAP via link included in message	✓
	Downloaded data is available via testserver apache with restricted access	✓
F	CF-netcdf data quality check	open
G	Other ocean CF-netCDF data	open
H	Radar data CF profile	open



WMO OMM

WIS 2 Principles in the project

	WIS 2.0 Principle	
1	WIS 2.0: adopts Web technologies and leverages industry best practices and open standards.	✓
2	WIS 2.0: uses URLs to identify resources	✓
3	WIS 2.0: prioritizes use of public telecommunications networks	✓
4	WIS 2.0: requires provision of Web service(s) to access or interact with digital resources (e.g. data, information, products) published using WIS.	✓
5	WIS 2.0: encourages NCs and DCPCs to provide 'data reduction' services via WIS that process 'big data' to create results or products that are small enough to be conveniently downloaded and used by those with minimal technical infrastructure.	X
6	WIS 2.0: will add open standard messaging protocols that use the publish-subscribe message pattern to the list of data exchange mechanisms approved for use within WIS and GTS.	✓



WMO OMM

WIS 2 Principles in the project

	WIS 2.0 Principle	
7	WIS 2.0: will require all services that provide real-time distribution of messages (containing data or notifications about data availability) to cache/store the messages for a minimum of 24-hours, and allow users to request cached messages for download.	✓
8	WIS 2.0: will adopt direct data-exchange between provider and consumer.	✓
9	WIS 2.0: will phase out use of routing tables and bulletin headers.	✓
10	WIS 2.0: will provide a catalogue containing metadata that describes both data and the service(s) provided to access that data.	✓
11	WIS 2.0: encourages data providers to publish metadata describing their data and Web services in a way that can be indexed by commercial search engines.	✓



WMO OMM

Project data/metadata standards

Which data standards are used in the project and how:

Standard	Used for...
RSS Feed	check for updates
CF-netCDF	meteorological data format
ISO 19115-2:2009(E)	ERDDAP metadata
ISO 19115:2003/Cor. 1:2006 (WMCP version 1.3 based on...)	WIS Catalogue import, search function WIS Portal, WIS OAI
GTStoWIS2	Topic value for pubSub messaging
WMO message format (https://github.com/wmo-im/GTStoWIS2/blob/main/message_format/message-schema.json)	messaging



Data/Metadata Standards

- Task Team on CF-NetCDF
 - Develop WMO extensions to CF-NetCDF conventions
 - Codify extensions for data profiles in WMO Data Manual
 - Increase standardization over conventional CF to support operational needs of WMO
 - Necessary to support exchange of data from multiple sources of similar data
 - CF-NetCDF profiles to act as templates for reporting data from different sources



WMO OMM

Data/Metadata Standards

Initially developed general regulations to ensure overall interoperability and consistency between different WMO-CF profiles

- Define standardization of general metadata and semantics



WMO OMM

WMO CF-Extensions

10 March 2021, version 0.1

DEFINITIONS

CF Conventions: The Climate and Forecast Conventions for netCDF (CF Conventions; e.g. Eaton et al., 2020) define a minimum set of metadata required to ensure that conforming netCDF files meet a basic level of self-description and interoperability. The required (minimal) set of metadata ensures that all variables in a dataset have "an associated description of what it represents, including physical units if appropriate, and that each value can be located in space (relative to earth-based coordinates) and time". Additional metadata are defined by the CF Conventions but are only recommended where they may not be needed or appropriate for all datasets.

WMO-CF Extensions: The WMO-CF extensions build on the CF Conventions to provide the framework for standardizing semantics and metadata, further reducing the effort involved in specifying data products and increasing interoperability. The WMO-CF:

1. Define additional metadata requirements or recommendations that are not defined by the CF Conventions;
2. Specify the set of optional CF Conventions metadata that this extension requires, making those optional metadata mandatory.

WMO-CF Profiles: The WMO-CF profiles implement the WMO-CF extensions for different data types by, *inter alia*: defining the standardized metadata and semantics; specifying the names of dimension and coordinate variables; and specifying the ordering of dimensions. The WMO-CF Profiles reduce the degrees of freedom available when creating netCDF files, increasing the standardization of data from different publishers for the same type of data.

FM System of Numbering WMO-CF Extensions and Profiles

Each WMO-CF profile or extension bears a number (X), preceded by the letters FM. This number is followed by the year (Y) and month (m) of operational implementation in the form FM X-Y-M.

Furthermore, an indicator term is used to designate the WMO-CF profile colloquially and is therefore called a "code name".

Notes on nomenclature:

To follow

FM SYSTEM OF WMO-CF PROFILES

Data/Metadata Standards

- Developed two specific CF-NetCDF Data Profiles
 - WMO-CF Radial
 - used for reporting of weather radar and lidar data in the native instrument-centric polar coordinates
 - WMO-CF Marine Trajectory
 - used for the reporting of meteorological and/or oceanographic observations along one or more trajectories, including both at or near the ocean surface and at depth.



WMO OMM

WMO CF-Extensions
10 March 2021, version 0.1

FM 302-XX WMO-CF MARINE TRAJECTORY

1. Scope

- This profile is intended for observations along one or more trajectories, including both at or near the ocean surface and at depth.
- A ragged array representation (e.g. see CF v1.8 conventions) shall be used.
- Only data for a single trajectory shall be used.
- Groups are not supported.

2. Global scope / root group

- Global attributes
 - The regulations in Table 1 shall be used.
- Station / platform identifier
 - For platforms with a unique identifier, the identifier shall be used.
 - For platforms with a common identifier, the identifier shall be used.
- Dimensions
 - Files containing dimensions:
 - obs, the name of the observation
 - trajec, the name of the trajectory
- Coordinate Variables
 - Table 2 lists the variables.
 - For platforms with a unique identifier, the observation identifier shall be used.
 - For platforms with a common identifier, the observation identifier shall be used.
- Trajectory identifier

WMO CF-Extensions
10 March 2021, version 0.1

FM 301-XX WMO-CF RADIAL

1. Scope

- This profile is for the representation of weather radar and lidar data in the native instrument-centric polar coordinates. Such data is the primary output of the radar/lidar signal processor known as "Level 2" data. This is the lowest level of output commonly available from operational instruments and is well suited to data exchange.
- The structure of this profile conforms to the WMO Information and Data Models for Radial Radar and Lidar Data. Effort has also been made to maximize compatibility with the CFRadial 2 format from which this profile has been derived.

2. Overview

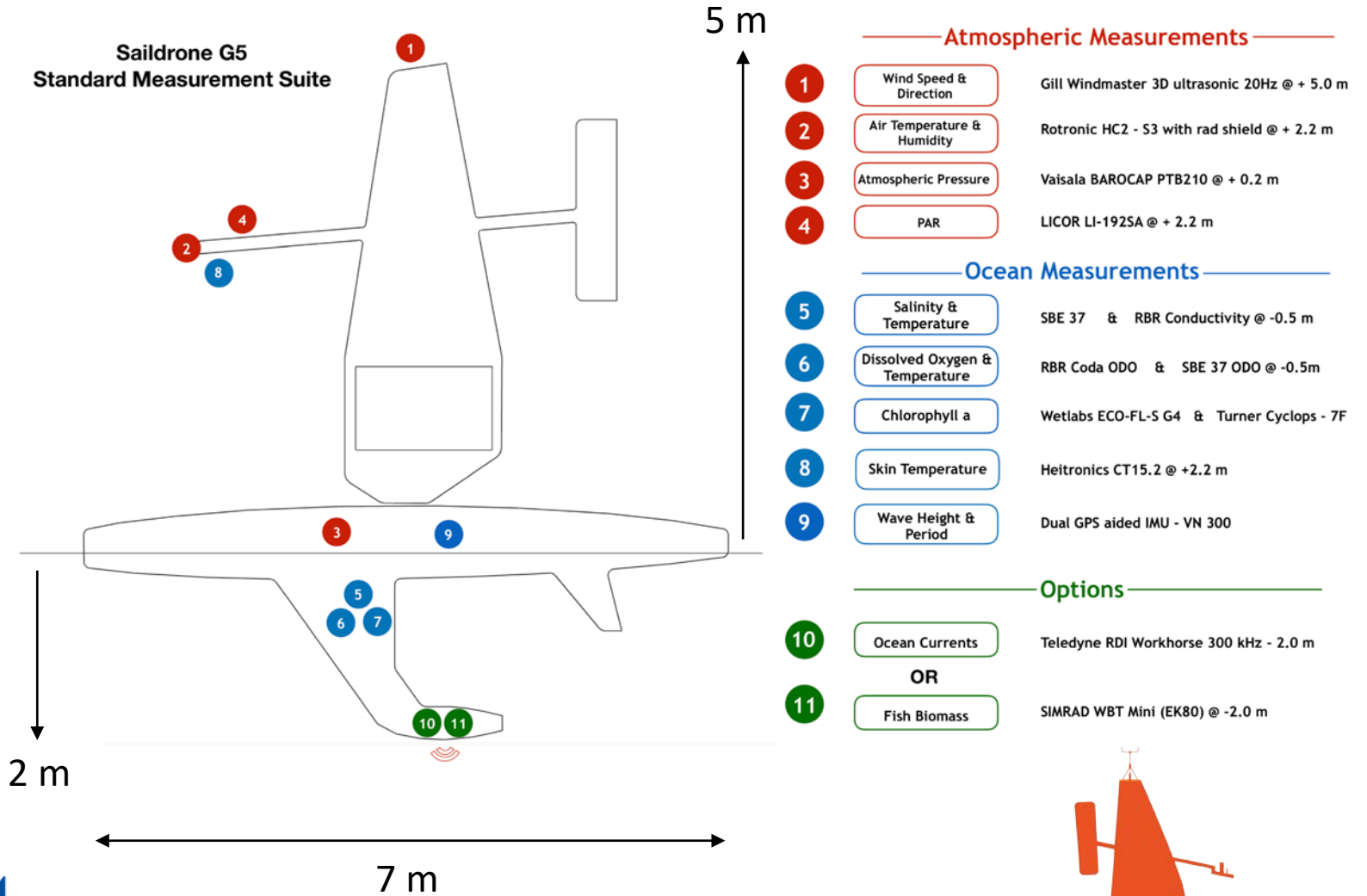
- Level 2 radar/lidar data may be conceptualized as a simple hierarchy of data objects where each object contains a collection of objects from the level below. These objects are:
 - Volume – The top-level object for the profile. A Volume is a collection of logically associated sweeps. Typically, these sweeps will represent a continuous or near-continuous series of observations acquired by the instrument during a single cycle of the scan schedule.
 - Sweep – Represents a subset of the data in the volume over which certain fundamental conditions remain constant. A common example is for a sweep to contain the data observed during a single 360-degree scan at a fixed elevation angle.
 - Ray – Represents a collection of data along a single direction of pointing from the instrument.
 - Range Bin – Represents a collection of data within a ray that are related to the same short window of range along the beam propagation path.
 - Dataset – A measured or calculated quantity that is associated with a range bin. Each Dataset will typically represent one of the measured radar moments such as reflectivity or Doppler velocity, but may also be used to store derived information such as quality control metrics.
- Within a Sweep all Range Bins contain the same collection of Datasets, and all Rays contain the same collection of Range Bins. This allows the lower three levels of the hierarchy to be collapsed into a collection of 2D variables. Each variable stores a single Dataset, with dimensions for Ray and Range Bin.
- To facilitate the hierarchical nature of the data to be represented, NetCDF groups are used. The global scope is used to store the Volume object, a group is used for each Sweep object, and a variable within each Sweep group is used for each Dataset. Coordinate variables and ancillary variables within the Sweep groups provide metadata related to the Ray and Range Bin objects.

3. Global scope / root group

1 of 26

What is a saildrone?

- Vehicles deployed on this mission developed in cooperation between NOAA's Pacific Marine Environmental Laboratory (PMEL) and Saildrone, Inc.
- Harnesses wind for propulsion like traditional sailboats, except with a vertical rigid wing
- Uses solar panels, battery packs for powering onboard sensors, telecommunications & navigation
- Capable of measuring variables needed to estimate surface fluxes of heat, momentum and CO₂, along with other variables, like salinity and upper ocean currents (ADCP)



WMO OMM



Data discovery

Is the project implementing a data discovery function? How?

- Users search for netCDF data via WIS Catalogue,
- read pubSub information from metadata record, and then
- can use a subscribe client script to receive notifications about new data for this product.
- The actual data download is done via URLs sent in the messages direct from data originator



WMO OMM

Data exchange

Is the project implementing a data exchange function? How?

- Via a client script with download, an exchange of the actual data between originator and customer (or other centers) can take place



WMO OMM

WIS 2 Principles in the project

- How the principle is implemented in the project

1	Used technologies and standards: Apache, Tomcat, myCore OAI, python (with lib pika), ERDDAP, rabbitMQ
2	Uses URLs for: Get metadata, search WIS Catalogue, download data
3	Only using Internet
4	Web Services used: WIS Portal, WIS OAI Service, ERDDAP, RSS Feed
6	Use publish and subscribe messaging
7	Real-time distribution via pubSub messaging, using 24h Cache ERDDAP and downloaded data (client)
8	Direct data exchange for download between provider and consumer



WIS 2 Principles in the project

- How the principle is implemented in the project

9	phase out use of routing tables and bulletin headers: pilot project data exchange works without routing tables and bulletin headers
10	Metadata containing data description and additional pubSub information transferOptions
11	can be indexed by commercial search engines: Should be possible but needs metadata imported to live system



WMO OMM

Input to WIS2

Contributions, suggestions, technical specifications, standards used in the project that should be adopted as standard or recommended practises in WIS2

- pubSub messages
- CF-netCDF as meteorological data format
- Web services for the provision of product data (like ERDDAP)



WMO OMM

Input to WIS2 (Lessons Learned)

WMCP version 1.3 vs. WMCP WIS2

- pubSub messaging: single, separate attributes for
 - message broker URL
 - Exchange
 - Topic(so that scripts could use metadata to get pubSub values)



WMO OMM

Input to WIS2 (Lessons Learned)

GTS2WIS JSON Specification

- Integrity Value requires download of data
- How is this value set by data provided using a Web Service?

Authentication/ Authorization concepts

- How to register for pubSub messaging
- Who should be allowed to publish messages
- Access control to restricted data
- User Federation



WMO OMM

WEATHER CLIMATE WATER

TEMPS CLIMAT EAU



Thank you
Merci

WMO OMM

World Meteorological Organization

Organisation météorologique mondiale