



Marine Energy Data Pipelines

May 11, 2021

Carina Lansing
Maxwell Levin
Chitra Sivaraman



PNNL is operated by Battelle for the U.S. Departm



Introduction

Data Pipeline Team:

- Chitra Sivaraman – Project Lead
 - Carina Lansing – Chief Architect
 - Max Levin – Lead Software Developer
- Funded by EERE's Water Power Technologies Office
 - Part of a larger MHKiT team and project led by NREL and Sandia
 - Sponsored by the Portal and Repository for Information on Marine Renewable Energy (PRIMRE)
 - 1st Webinar: [Data Standards](#) – April 20th, 2021
 - 2nd Webinar: ME Data Ingestion Pipeline – May 11th, 2021

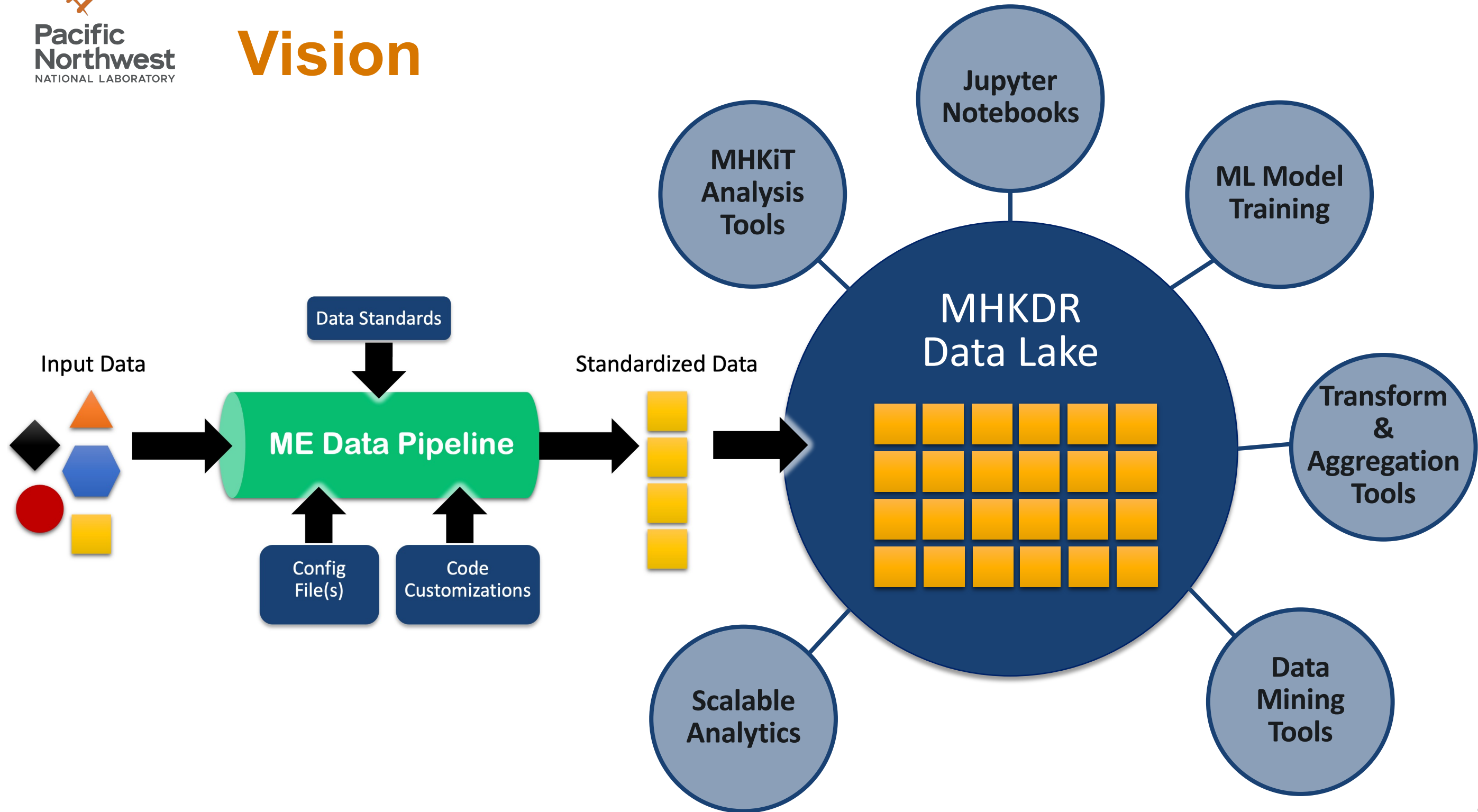




Outline

- Vision for data accessibility & interoperability
- Data standards recap
- Overview of tsdat data pipelines
- Live demonstration
- Roadmap
- Summary

Vision



Data Standards Recap

A standardized dataset has:

- Metadata (descriptive, provenance, quality control)
- Standard units
- Standard names for variables
- Standard format & file names

Standardized datasets can be extended by

Data Models which have:

- Sets of specific variables that should be included
- Sets of specific calculations that should be included
- Sets of specific metadata that should be included

Data Standards [document](#)

Data Standards Webinar [recording](#) and [slides](#)

Baseline Data Standards

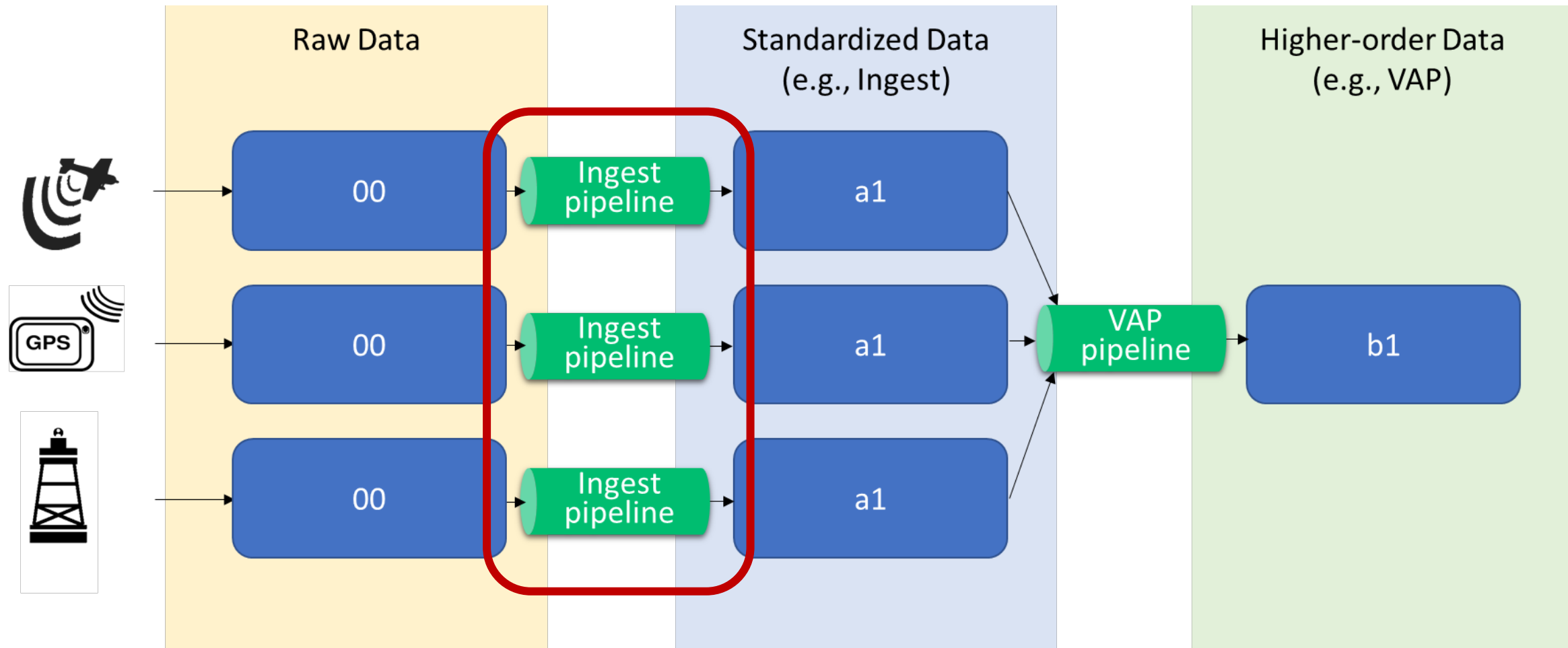
- Standard data format(s)
- Standard units
- Standard names
- Standard variables
- Data quality standards
- File naming conventions

Data
Model

Data
Model

Data
Model

Types of Pipelines

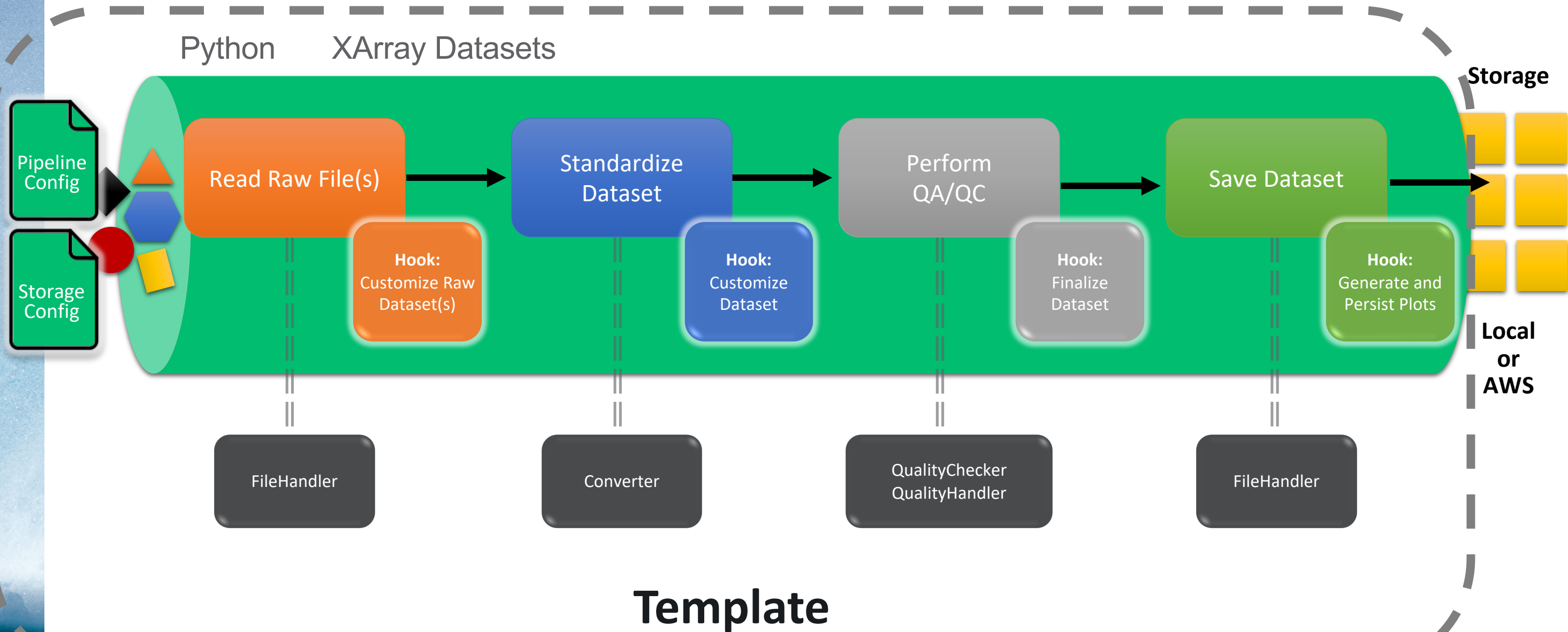


No changes to data produced by instruments or models

Data have been converted to standard format, and quality checks and corrections have been applied

Data have been combined and/or algorithms have been applied to translate the instrumental measurements to the geophysical quantities needed for scientific analysis

Tsdat Ingest Pipeline

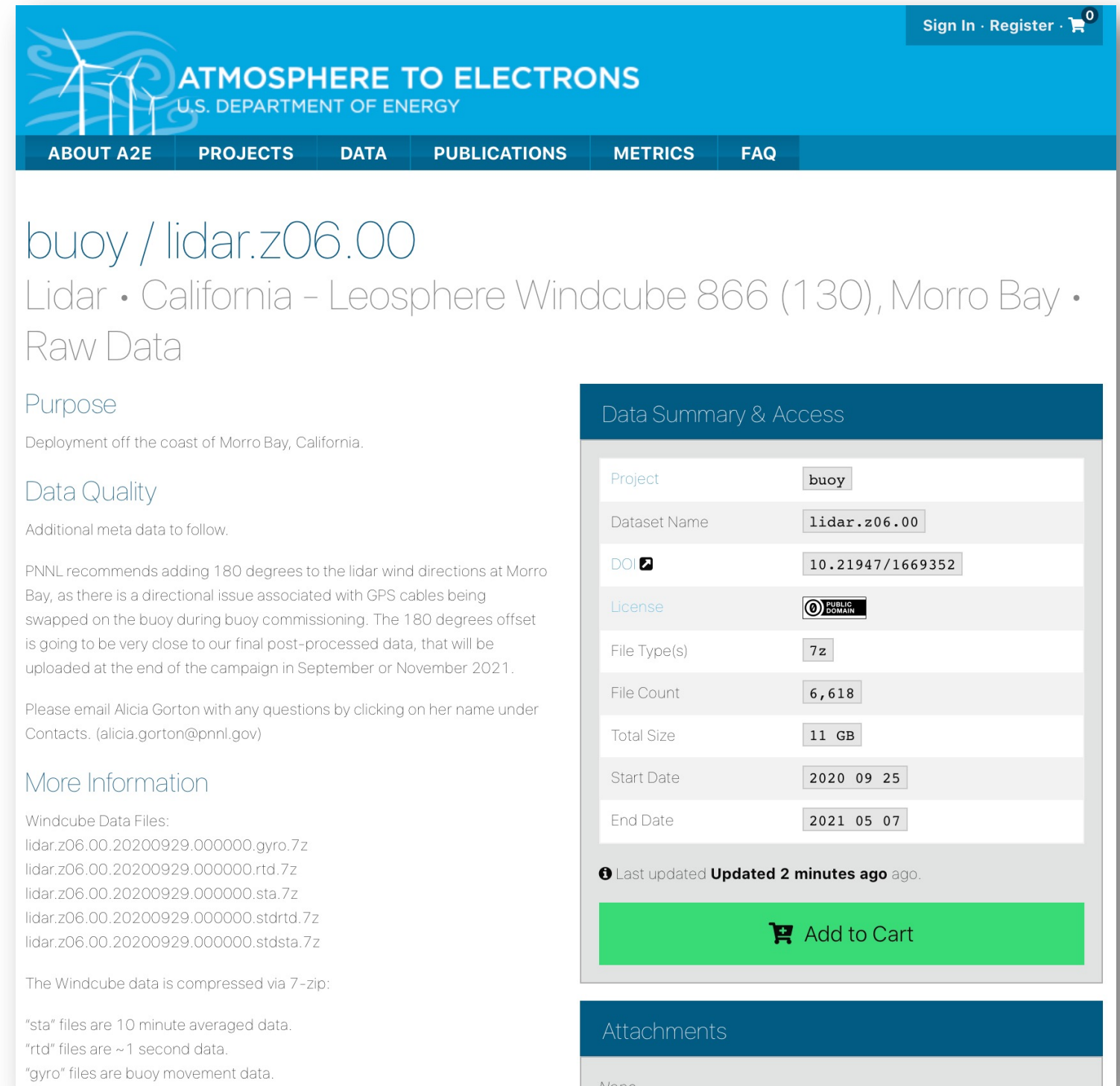



A high-angle, close-up shot of a massive, curling blue wave. The water is a vibrant, deep blue, and the crest is breaking into white foam. The wave's face is curved, creating a tunnel-like effect. The text "Live Demo" is centered over the wave in a bold, orange, sans-serif font.

Live Demo

Demo Overview

- Brief example of a tsdat ingest
 - Getting started, customizable components, running the pipeline
- Data from the A2E program
 - Buoy located in Morro Bay, CA
 - Lidar measuring wind vectors
 - “sta” 10-min averaged data
- Goals
 - embed metadata for ease-of-use
 - perform quality checks / controls
 - visualize processed data



Sign In · Register · 

ATMOSPHERE TO ELECTRONS
U.S. DEPARTMENT OF ENERGY

ABOUT A2E PROJECTS DATA PUBLICATIONS METRICS FAQ

buoy / lidar.z06.00

Lidar · California - Leosphere Windcube 866 (130), Morro Bay · Raw Data

Purpose
Deployment off the coast of Morro Bay, California.

Data Quality
Additional meta data to follow.

PNNL recommends adding 180 degrees to the lidar wind directions at Morro Bay, as there is a directional issue associated with GPS cables being swapped on the buoy during buoy commissioning. The 180 degrees offset is going to be very close to our final post-processed data, that will be uploaded at the end of the campaign in September or November 2021.



Please email Alicia Gorton with any questions by clicking on her name under Contacts. (alicia.gorton@pnnl.gov)


More Information
Windcube Data Files:
lidar.z06.00.20200929.000000.gyro.7z
lidar.z06.00.20200929.000000.rtd.7z
lidar.z06.00.20200929.000000.sta.7z
lidar.z06.00.20200929.000000.std.rtd.7z
lidar.z06.00.20200929.000000.stdsta.7z

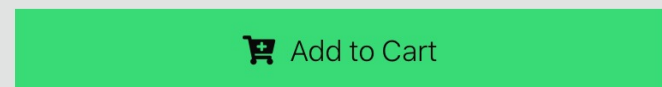
The Windcube data is compressed via 7-zip:

“sta” files are 10 minute averaged data.
“rtd” files are ~1 second data.
“gyro” files are buoy movement data.

Data Summary & Access

Project	buoy
Dataset Name	lidar.z06.00
DOI 	10.21947/1669352
License	
File Type(s)	7z
File Count	6,618
Total Size	11 GB
Start Date	2020 09 25
End Date	2021 05 07

 Last updated **Updated 2 minutes ago**.

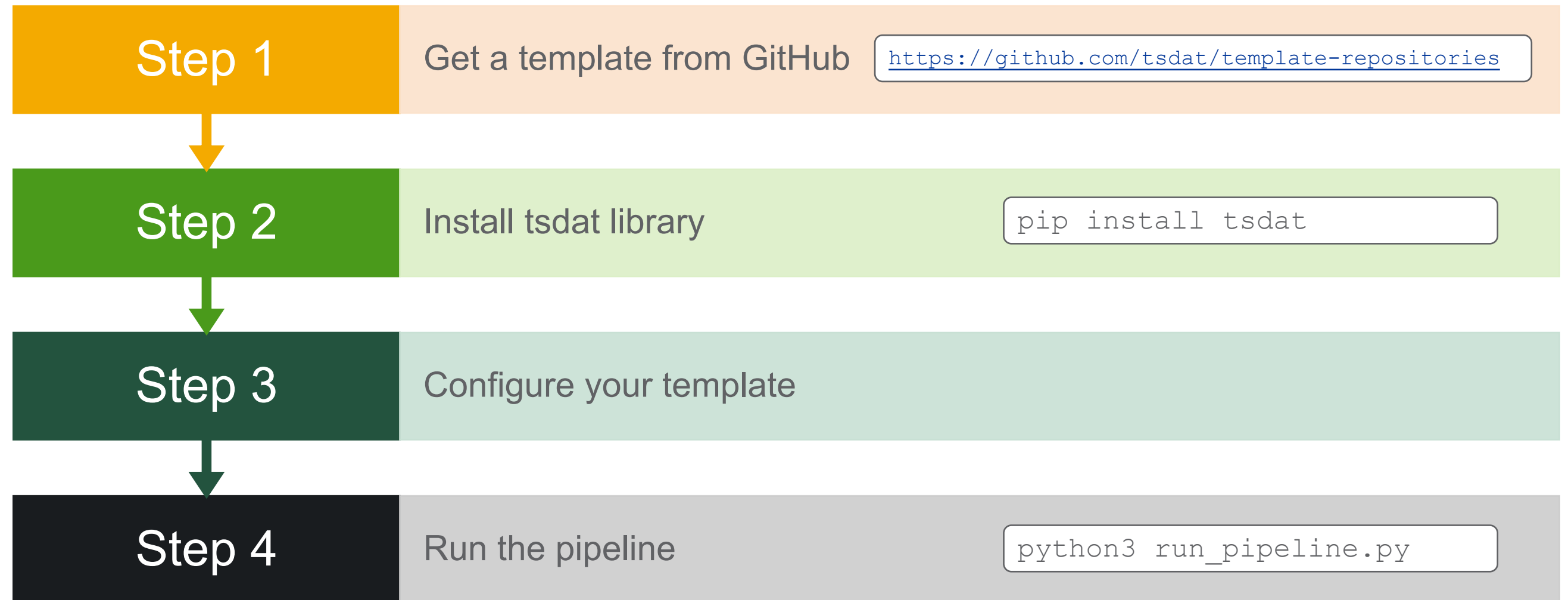


Attachments

None

Tsdat - Getting Started

Prerequisite: Python 3





Questions & Answers

Roadmap

FY21	FY22	FY23	FY24
Ingest pipeline framework and templates	VAP pipeline framework and templates	Tools for curating, versioning, and reprocessing erroneous data	Tools and APIs for big data analytics
	MHKDR Integration		
Draft data standards	Select data model template development		
Community outreach, documentation, webinars, and tutorials			
	Data lake design and implementation		
Review data standards and prioritize data models		Data discovery dashboard	
	Data model standards development		
	Implement select data pipelines		

Summary

- Tsdats code & templates: <https://github.com/tsdat>
- Tsdats documentation: <https://tsdat.readthedocs.io/>
- We are recruiting early adopters, so please reach out to us if we can assist!
tsdat@pnnl.gov



**Pacific
Northwest**
NATIONAL LABORATORY

Thank you

