EDITO

European Digital Twin Ocean

# NCZarr Viewer

## Exploring and Subsetting Zarr & NetCDF Data

2 Sept 2025

Samuel Fooks — VLIZ


Funded by the European Union

This work is funded by the European Union under grant agreement no. 101227771. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

# NCZarr Viewer

- 📊 **Load and explore** NetCDF and Zarr datasets
- 🔍 **Browse variables** and dimensions through a simple interface
- ✂️ **Subset data** by time, space, and other dimensions visually
- 📈 **Visualize results** with interactive plots
- 🌍 **Access cloud data** directly from S3 buckets
- 🚀 **Work with large datasets** efficiently
- 🐋 **Containerized** for easy deployment and sharing
- In EDITO Datalab!

# 🏗️ Architecture Overview

👤 User Interface → 🚀 Dash App → 🗄️ Data Engine

↓ ↓ ↓

🖥️ Web Browser 🐍 Python Core 📊 Xarray

↓ ↓ ↓

🎨 Interactive UI 🔧 Data Manager 📁 NetCDF/Zarr

# 🛠️ Technology Stack

- **Frontend**: Dash + Bootstrap Components
- **Data Processing**: Xarray + NumPy
- **File Formats**: NetCDF4, Zarr
- **Visualization**: Plotly, Matplotlib, Cartopy
- **Cloud Access**: S3FS, FSSpec (cloud storage access)
- **Marine Data**: Copernicus Marine Toolbox integration

# 🚀 Quick Start

## On EDITO

**OR**

```
# Use Docker
docker run -p 8050:8050 samfooks/nczarr-viewer:latest
```

**OR**

```
# Local development (if you have Python)
git clone [https://github.com/EDITO-Infra/nczarr-viewer)](https://github.com/EDITO-Infra/nczarr-viewer)
cd nczarr-viewer
pip install -r requirements.txt
python run.py
```

**Access at:** http://localhost:8050

💡 **Tip**: Think of this as "R Shiny for NetCDF data" - but already built f[illegible]

# Supported Data Sources

- **ARCO data on EDITO**: ARCO datasets from the EDITO STAC
- **Personal Cloud Storage**: Minio storage on EDITO
- **Local Files**: NetCDF, Zarr

# 🔍 Core Features

## Data Exploration

- **Variable Browser**: See all variables, dimensions, and metadata
- **Dimension Handling**: Time, depth, latitude, longitude
- **Data Subsetting**: Interactive selection of regions and time periods

## Visualization

- **Interactive Maps**: Cartopy-based geographic plots
- **Time Series**: Plotly charts for temporal data
- **Statistical Analysis**: Basic stats, and summaries

# Marine Data Examples

## EDITO Integration

- **Biodiversity**: Species distribution data
- **Chemistry**: Water quality parameters
- **Geology**: Seafloor characteristics
- **STAC Access**: Browse collections and datasets

## Copernicus Marine

- **Direct Access**: CMEMS credentials integration (you will need an account)
- **Multiple Formats**: NetCDF, Zarr (and others in future)
- **Real-time Data**: Latest ocean observations

# 🚀 Performance Features

- **Chunked Processing**: Handle datasets larger than memory
- **Lazy Loading**: Only load data when needed
- **Cloud Optimization**: Efficient S3 data access

# 🔧 Configuration & Deployment

## Setup

To access CMEMS datasets you may need an account using Copernicus Marine Toolbox

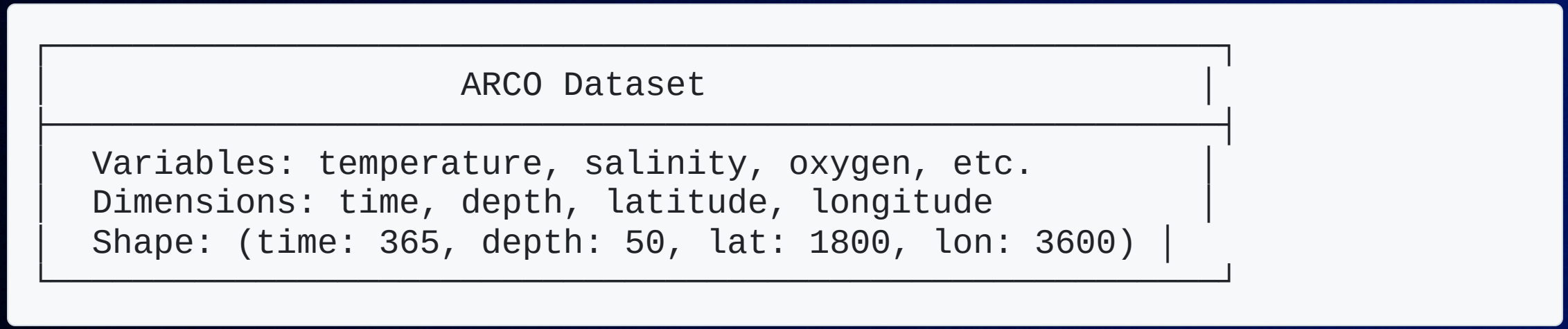
```
# CMEMS credentials
CMEMS_USERNAME=your_username
CMEMS_PASSWORD=your_password
```

## Docker Deployment

```
docker build -t nczarr-viewer .
docker run -p 8050:8050 nczarr-viewer
```

# 🔍 Subsetting ARCO Data: Core Concepts

## Multidimensional Data Structure




## Subsetting Operations

- **Variable**: Pick specific parameters
- **Temporal**: Select specific dates
- **Spatial**: Choose latitude/longitude boundaries

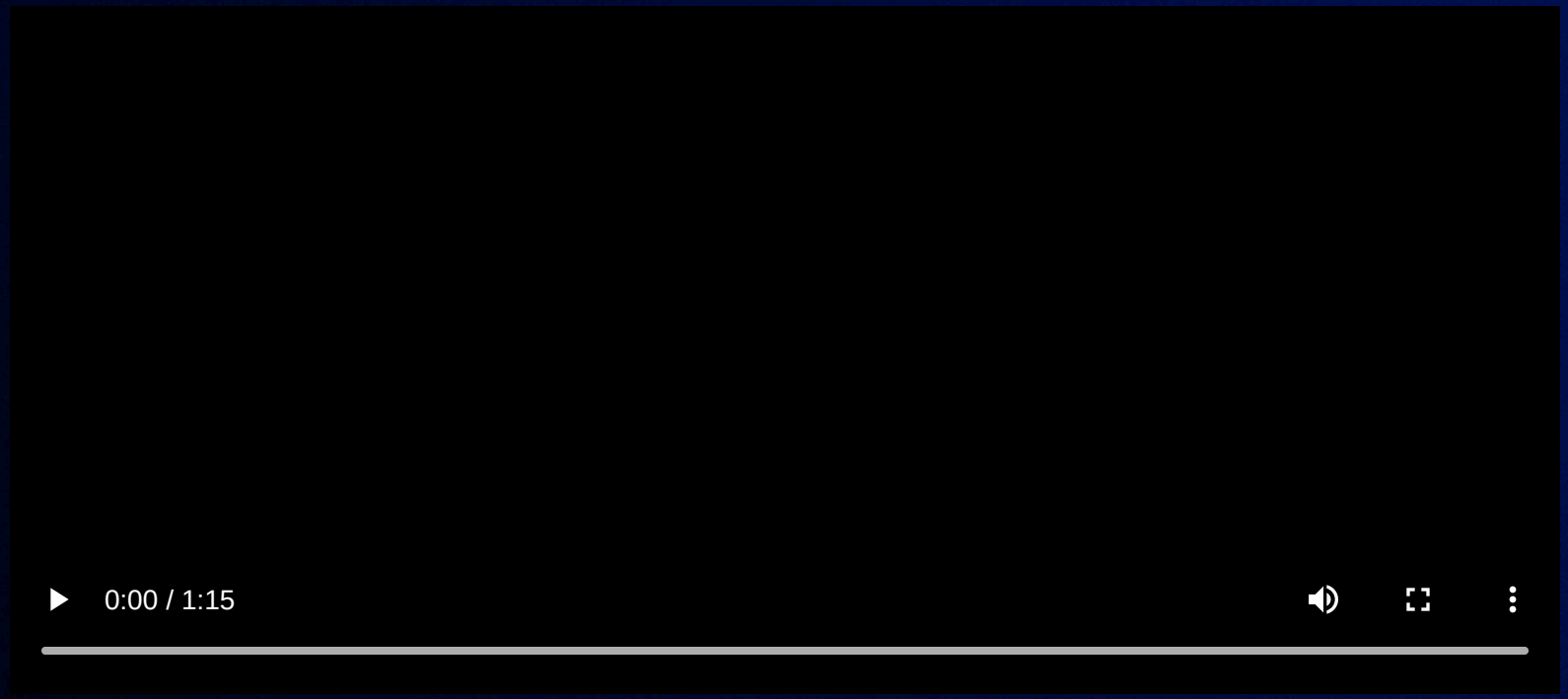
# 🎯 Subsetting in Practice

## Example: Extract Surface Temperature for North Sea

```
import xarray as xr

# Load ARCO dataset
ds = xr.open_zarr("s3://arco-data/ocean-temp.zarr")

# Variable: surface temperature
temp_surface = ds['temperature'].sel(depth=0)

# Temporal: August 30
august30_data = temp_surface.sel(
    time='2025-08-30'
)
# Spatial bounds: North Sea region
north_sea = august30_data.sel(
    latitude=slice(51.0, 61.0),    # 51°N to 61°N
    longitude=slice(-5.0, 15.0)    # 5°W to 15°E
```

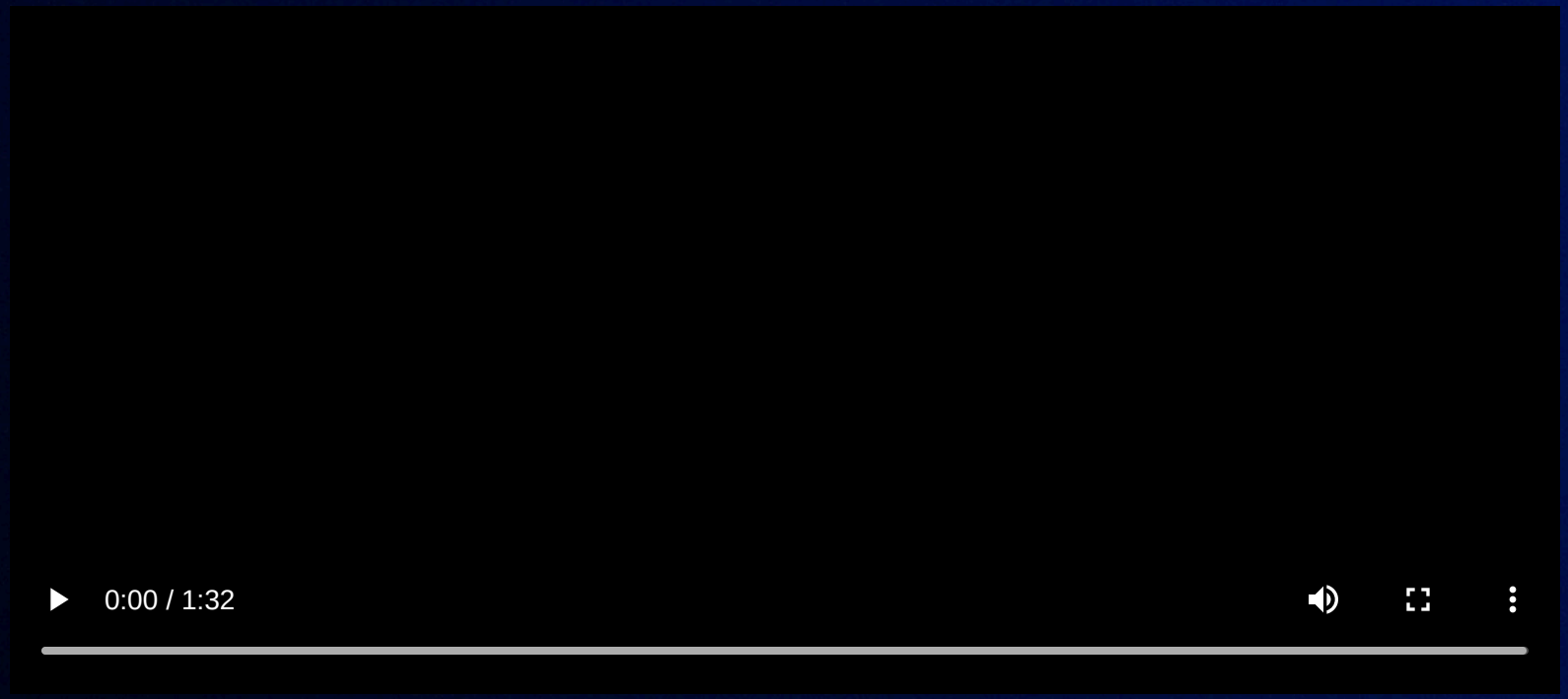
# Visual Representation

# Use the NCZarr Viewer locally or on EDITO




Video link

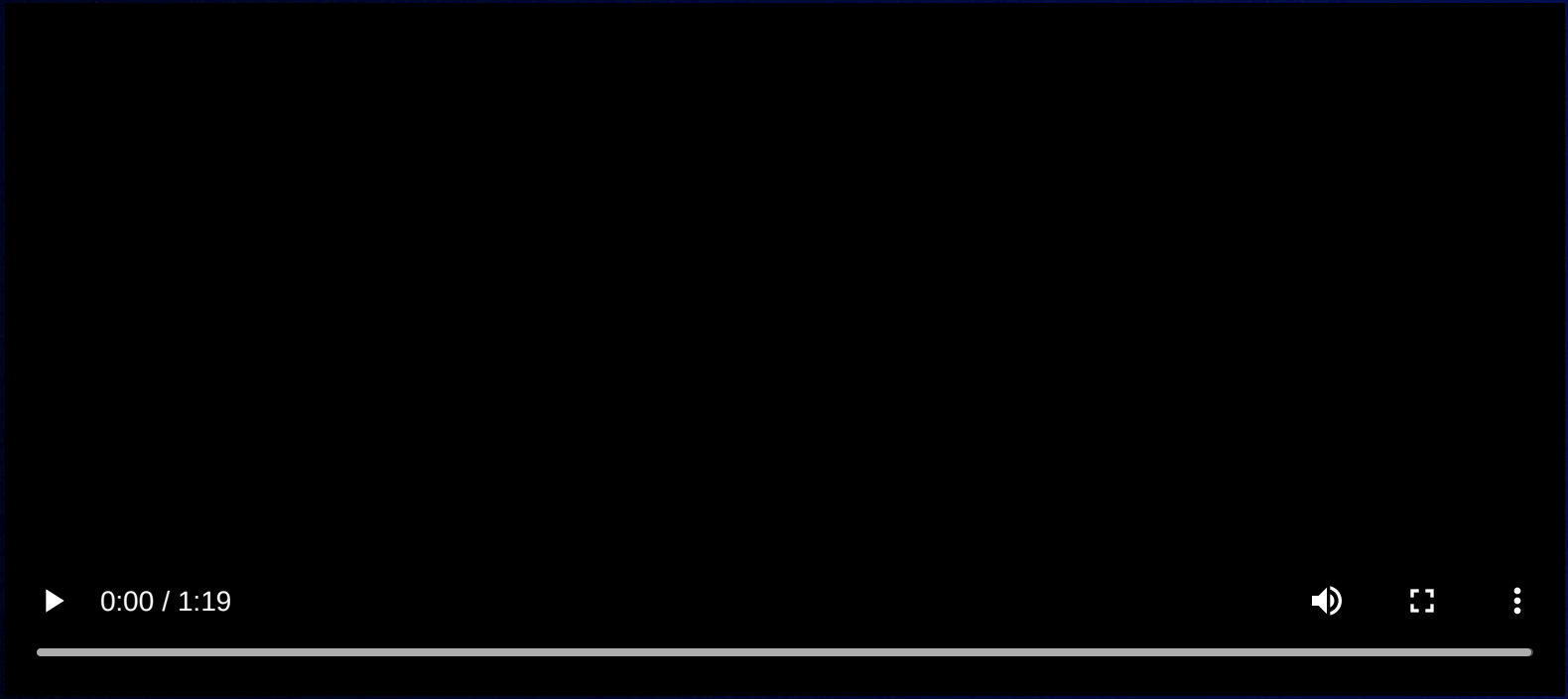
# Explore a NetCDF from your Minio bucket on EDITO




Video link

# Live Demo Time!

## Explore CMEMs dataset using zarr link from EDITO STAC




Video link

# 🔮 Future Developments

- **More Interactive Visualization**: More interactive global maps and plots
- **Advanced Analytics**: Statistical modeling tools/plugins
- **New ARCO Data types**: Parquet, Geoparquet
- **Collaboration**: Multi-user editing and sharing

# Thank You!

**Samuel Fooks** - samuel.fooks@vliz.be

**GitHub**: https://github.com/EDITO-Infra/nczarr-viewer

**Docker Hub**: samfooks/nczarr-viewer

**Questions?**