

DMAC/AI Briefing for IOOS FAC

Micah Wengren
Hassan Moustahfid
Mathew Biddle

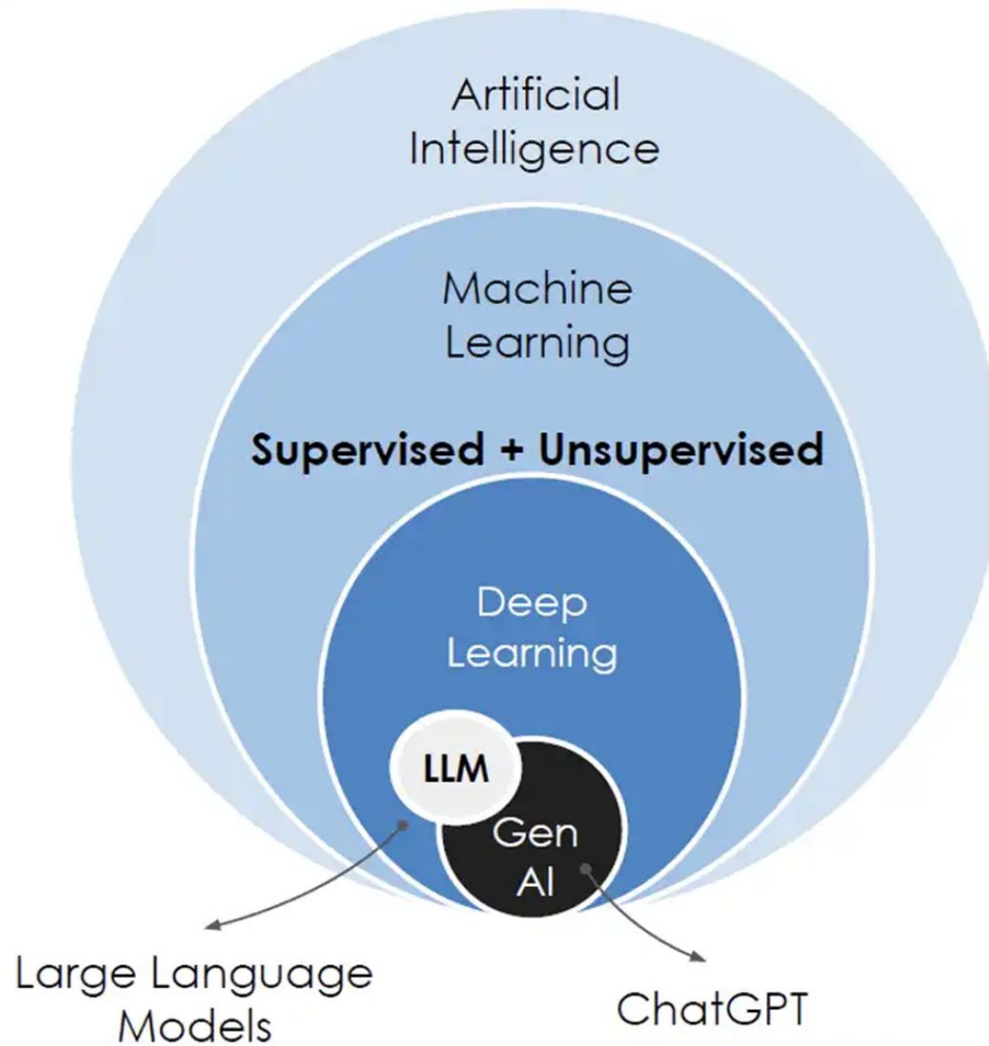
September 22, 2025



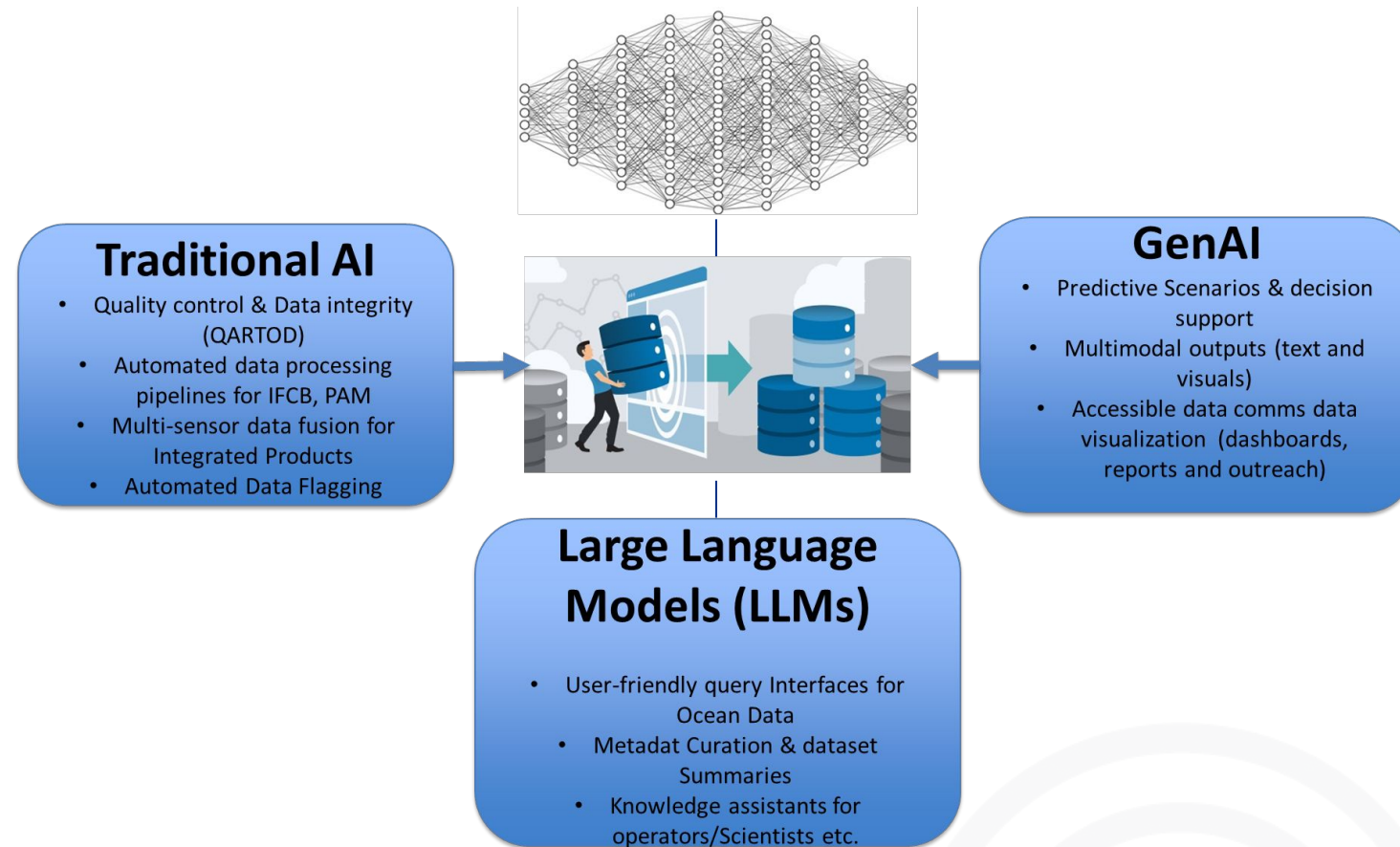
AI/ML: Overview and Context for DMAC Activities

Hassan Moustahfid

Demystifying AI: Key Differences Between Traditional AI, GenAI, and LLMs



AI IN IOOS DMAC



The relation between large language models (LLMs), generative AI, and other related learning schemes. Malandrucolo. – A Review. Pr. J. of Science. DOI: <https://doi.org/10.70389/PJS.100004>

Evolution of AI Technologies: Next Frontiers

PHYSICAL AI

SELF-DRIVING CARS
GENERAL ROBOTICS

AGENTIC AI

CODING ASSISTANT
CUSTOMER SERVICE
PATIENT CARE

GENERATIVE AI

DIGITAL MARKETING
CONTENT CREATION

PERCEPTION AI

SPEECH RECOGNITION
DEEP RECSYS
MEDICAL IMAGING

2012 ALEXNET



Nvidia CEO Jensen Huang outlining the evolution of AI technologies, from perception AI to generative AI, agentic AI, and the rise of physical AI, during his keynote at the Consumer Electronics Show (CES) 2025.

DMAC Program Overview

Micah Wengren, DMAC System Architect

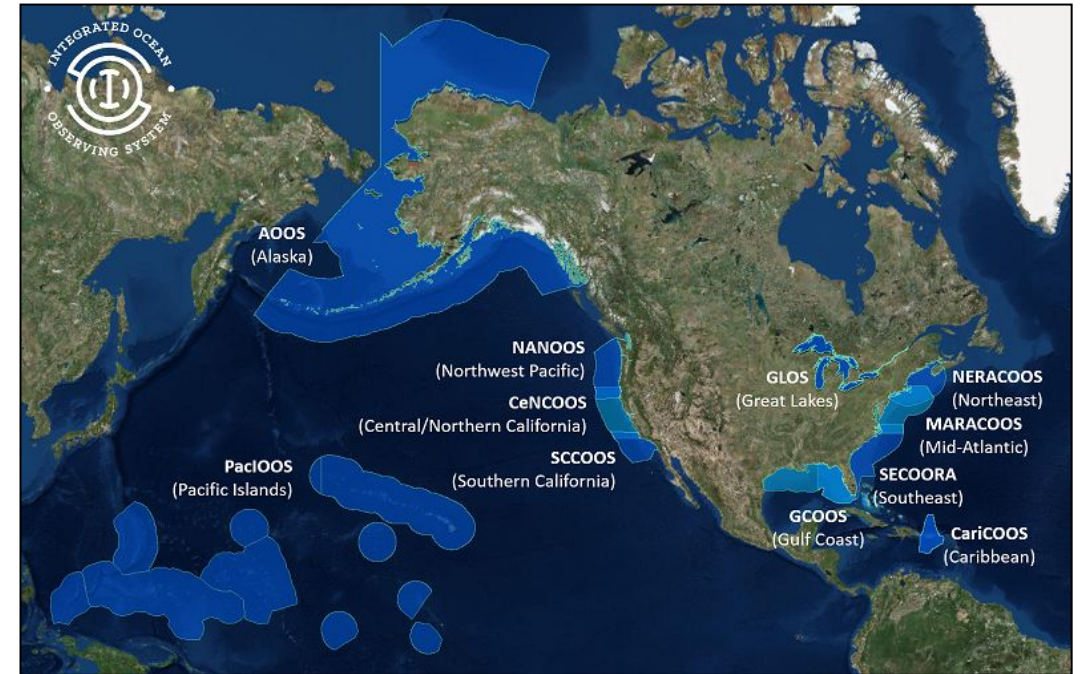
What is DMAC?

DMAC = Data Management and Cyberinfrastructure

DMAC is a community of practice: people, process, and technology that enable management and dissemination of ocean data following IOOS' guidelines.

DMAC spans IOOS' 11 RAs, public and private sector partners co-developing data standards, software & technology, and data management practices.

IOOS RAs implement DMAC data standards and serve as NOAA-certified, interoperable regional ocean data aggregators.



Technical Components & Resources of DMAC Community of Practice

Open Source Code: <https://github.com/ioos>

DMAC Data Standards: <https://ioos.github.io>

IOOS.us National Data Portal: <https://ioos.us>

Collaboration/Communication Platforms:

Mailing List: https://groups.google.com/g/ioos_tech

Slack Workspace: <https://ioos.slack.com>

IOOS DMAC Community of Practice



DMAC Annual Meetings

Annual

Project updates/briefings, presentations, small breakout discussions

<https://ioos.noaa.gov/project/dmac/>

IOOS Code Sprints

Bi-Annual

Hands-on coding on self-organized projects

<https://ioos.github.io/ioos-code-sprint/>

Mobilization Workshops

Annual

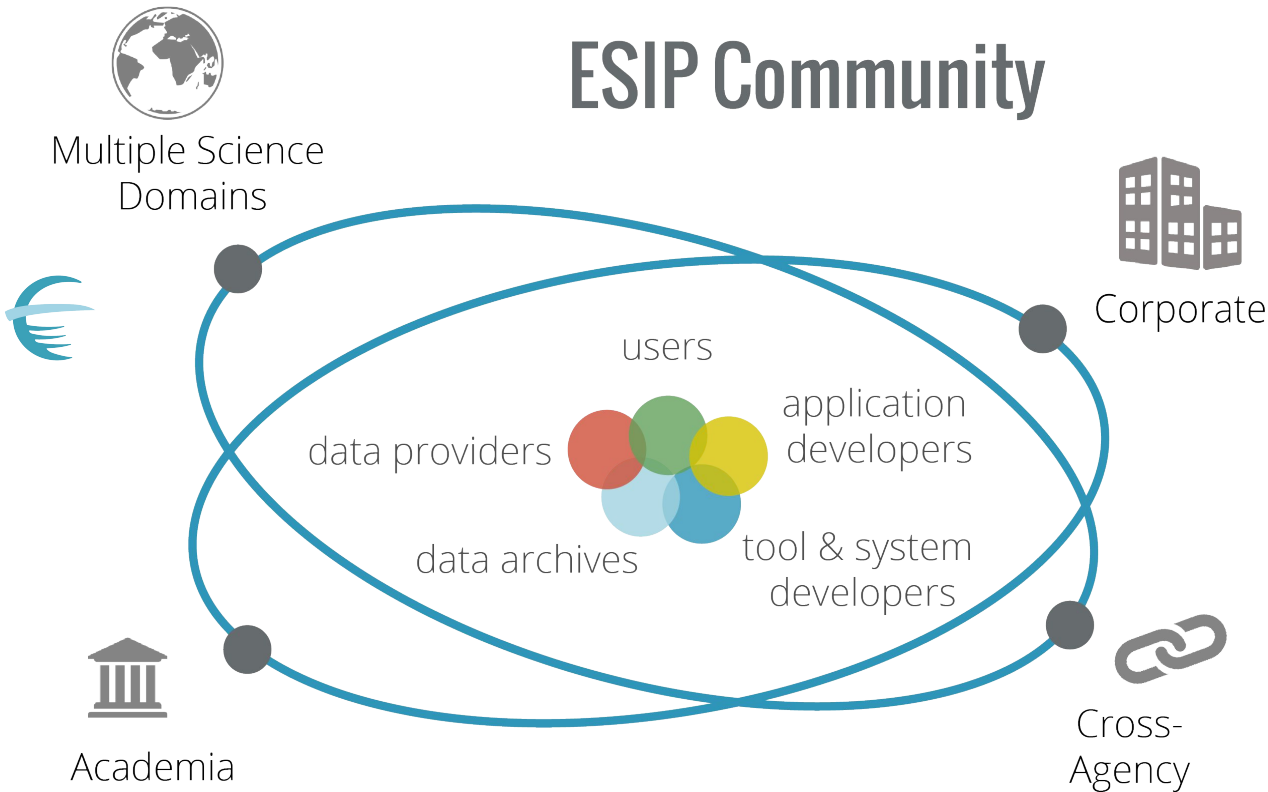
Hands-on data standardization workshop

https://ioos.github.io/bio_mobilization_workshop/

Supporting New Initiatives: Community-driven AI-Ready Data

ESIP is a brain trust of Earth science data & computing professionals

ESIP Community



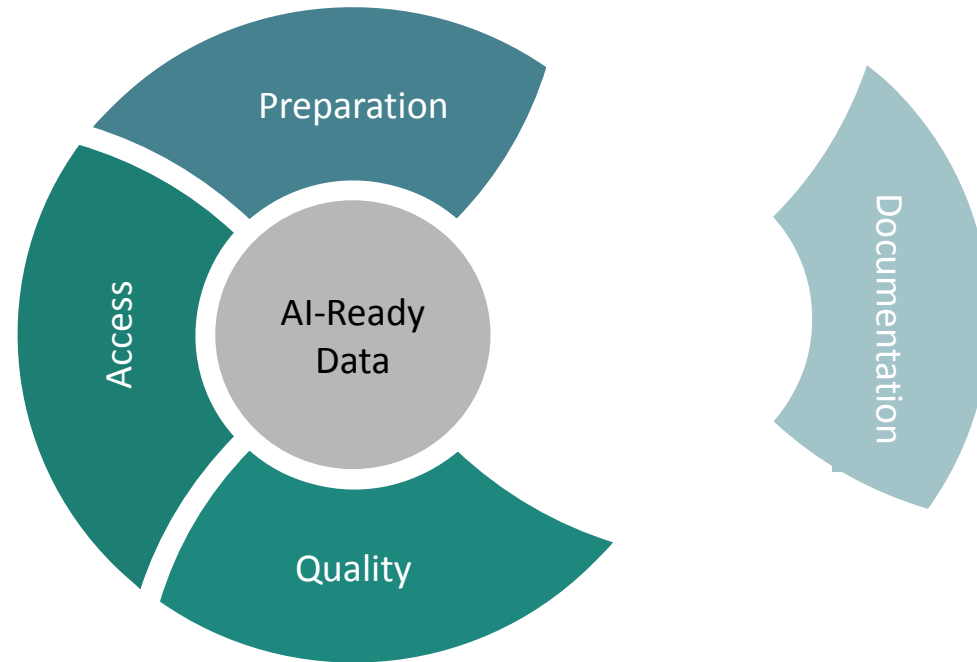
<https://www.esipfed.org/collaboration-areas/>

Slide courtesy of : Rob Redmon/NOAA NCAI

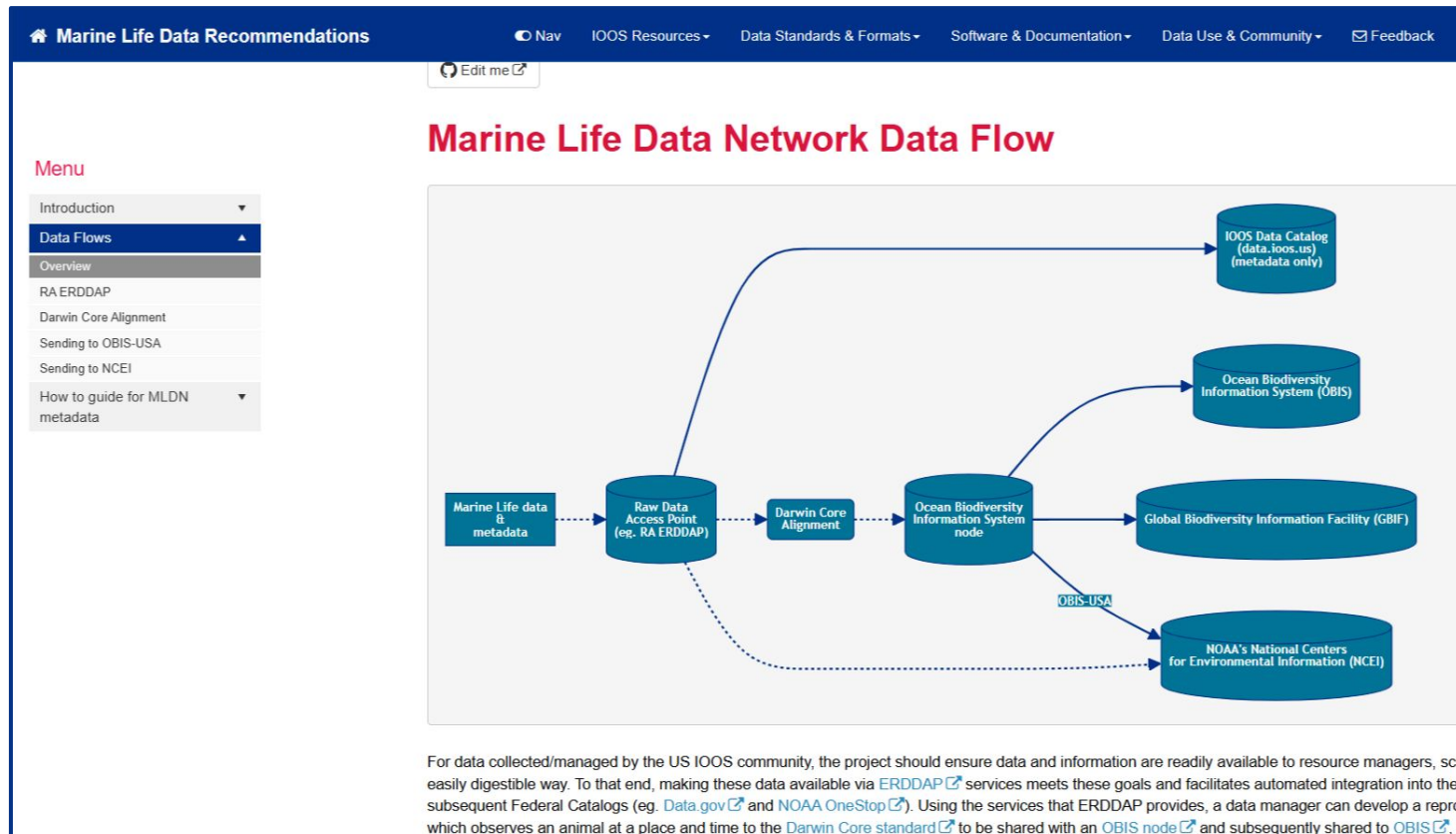


ESIP AI-readiness Checklist				
Dataset Name			Version 1.0, last updated 7/13/2023	
Dataset Version				
Dataset Location/Link				
Assessor Name				
Assessor Email				
		Responses	Definitions Usability Score	
Is this raw data or a derived/processed data product?		(please select option) ▼		
Is this observational data, simulation/model output, or synthetic data?		(please select option) ▼		
Is the data single-source or aggregated from several sources?		(please select option) ▼		
Data Quality				
Will the dataset be updated?		(please select option) ▼		
If the data will be updated, how often will it be updated?		(please select option) ▼		
If the data will be updated, choose the update frequency that best describes the dataset.		(please select option) ▼		
Will there be different stages of the update (e.g., updated with preliminary data first and replaced by a later update of the full record)?		(please select option) ▼		
If yes, what is the delay between different stages?		(short answer)		
Should the new version of the dataset supersede the current version?		(please select option) ▼		
Please provide an explanation for "Other" response.		(short answer)		
Is there any documentation about the completeness of the dataset?		(please select option) ▼	Completeness definition	
If there is documentation regarding the completeness of the dataset, please provide a link to report/document.		(link to report/document)		
How complete is the dataset compared to the expected spatial coverage?		(please select option) ▼		

1. Documentation



DMAC Data Standards, Data Flows, Data Mgmt Guidance



DMAC Data Standards & Technical Guidelines Portal:

<https://ioos.github.io/>

- technical documentation
- data formats/specifications
- software documentation
- data-type specific data management guidance

[IOOS Metadata Profile](#): Buoy & shore station-focused, GTS ingest requirements

[IOOS Glider Specification](#): Glider DAC format/guidance

[IOOS ATN Specification](#): ATN format/guidance

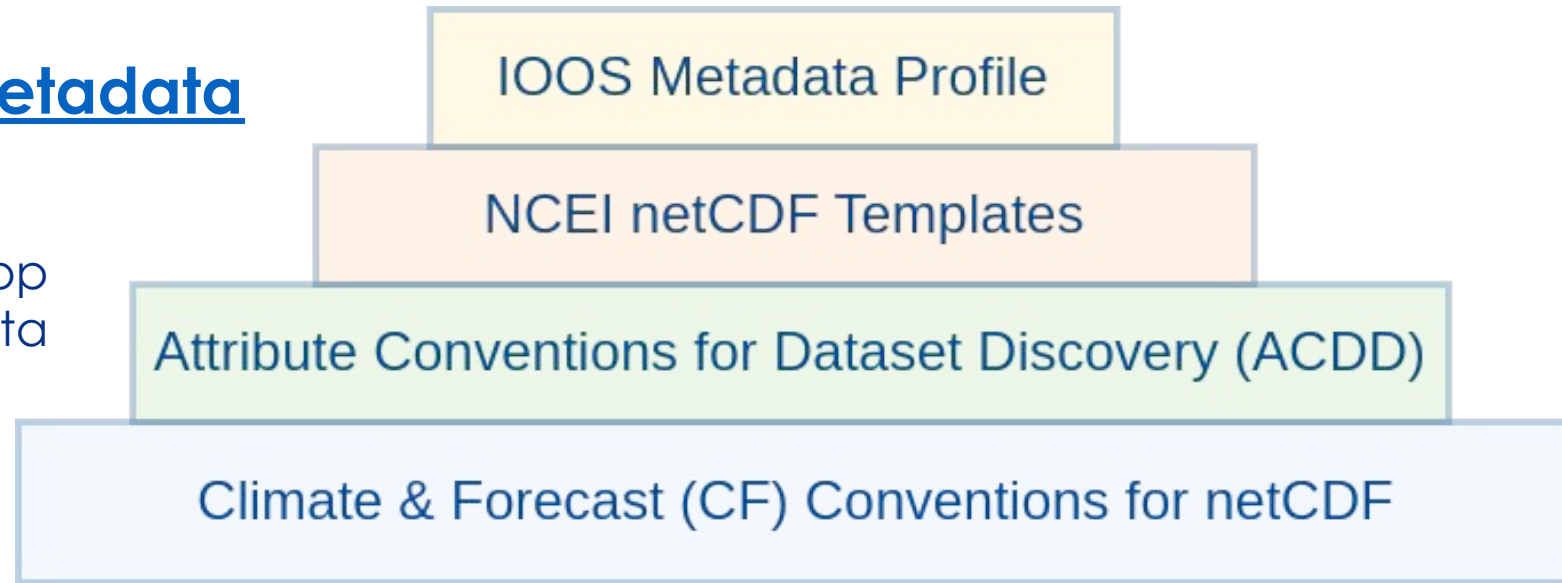
[IOOS Marine Life Data Network](#): Overarching guidance/tools for Marine Life data mgmt

[NCEI Archiving Cookbook](#): Guide for RAs for NCEI archiving process

Enabling FAIR Data Access: IOOS Metadata Profile

<https://ioos.github.io/ioos-metadata>

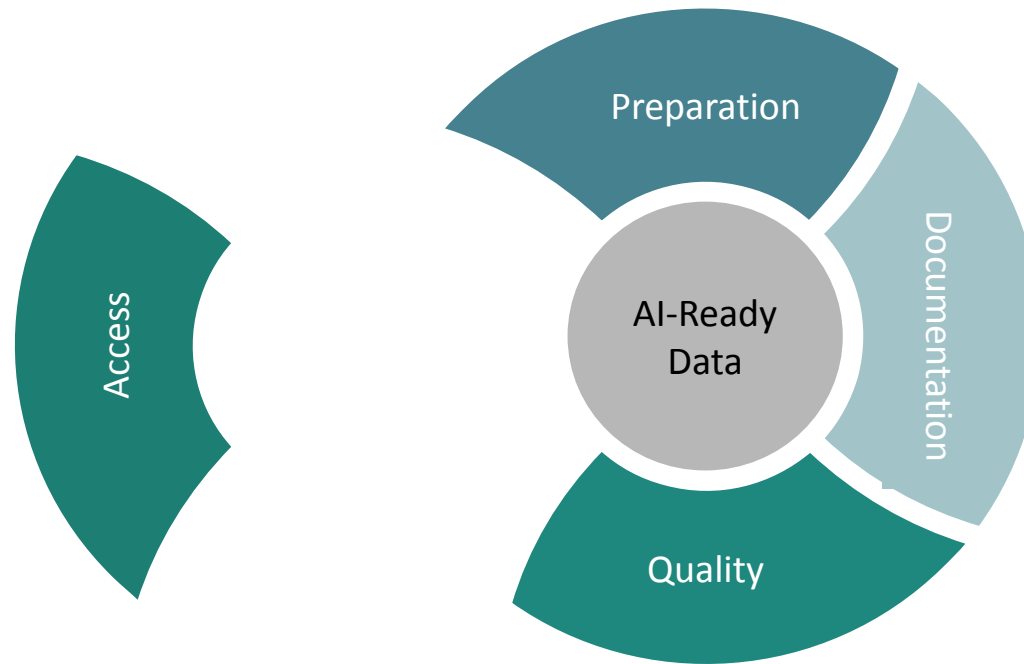
The IOOS Metadata Profile is built on top of and extends existing community data standards with IOOS' requirements



Key Features:

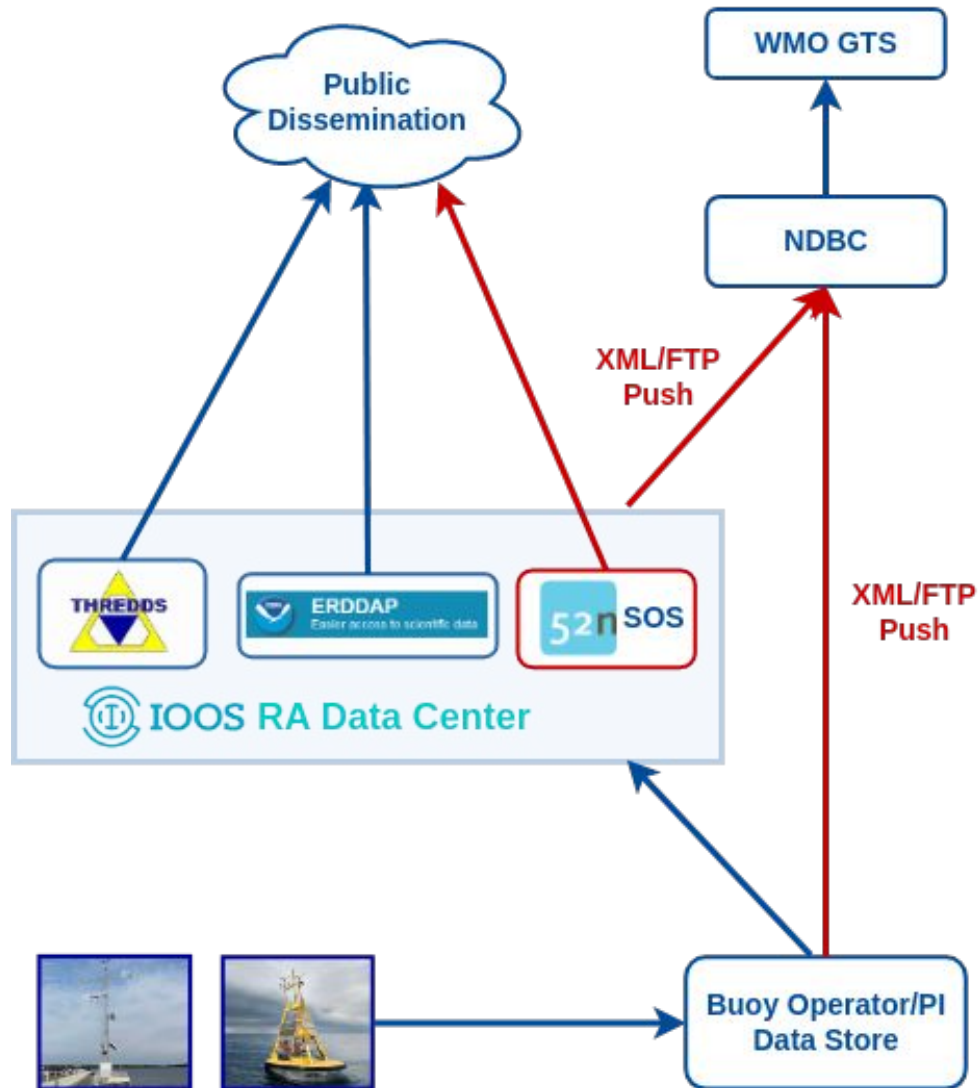
- Specifies attribute guidelines and restricts some CF flexibility to suit IOOS community use:
 - single observing platform (in [CF discrete sampling geometries](#)) per dataset
 - QARTOD flagging: requires all data variables to specify an ancillary variable with 'aggregate_quality_flag' Standard Name -> contributed to [CF Standard Names v72](#)
 - consistent, required global attribution for display in IOOS' national products
- Enables FAIR data access - more on this later
- Interoperability with other CF-based ocean obs data standards

2. Data Access

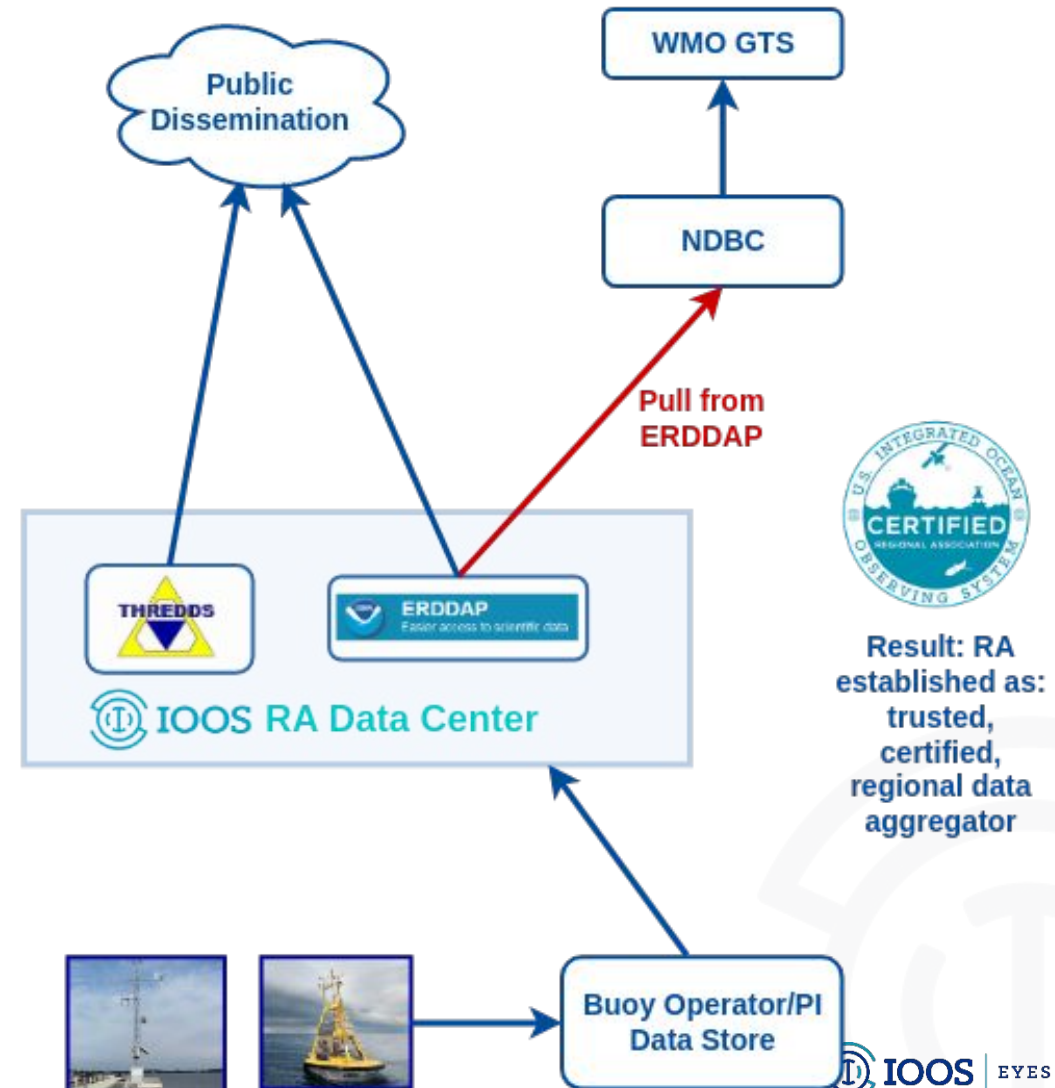


IOOS RAs use ERDDAP for multi-purpose data dissemination

Before



After



Year 1 of CO-DMAC National DMAC CI Project - 2025

Award: 500K FY2023-2026-funded BIL/IIJA funding to MARACOOS & TetraTech to implement a Cloud-Optimized DMAC (CO-DMAC) CI

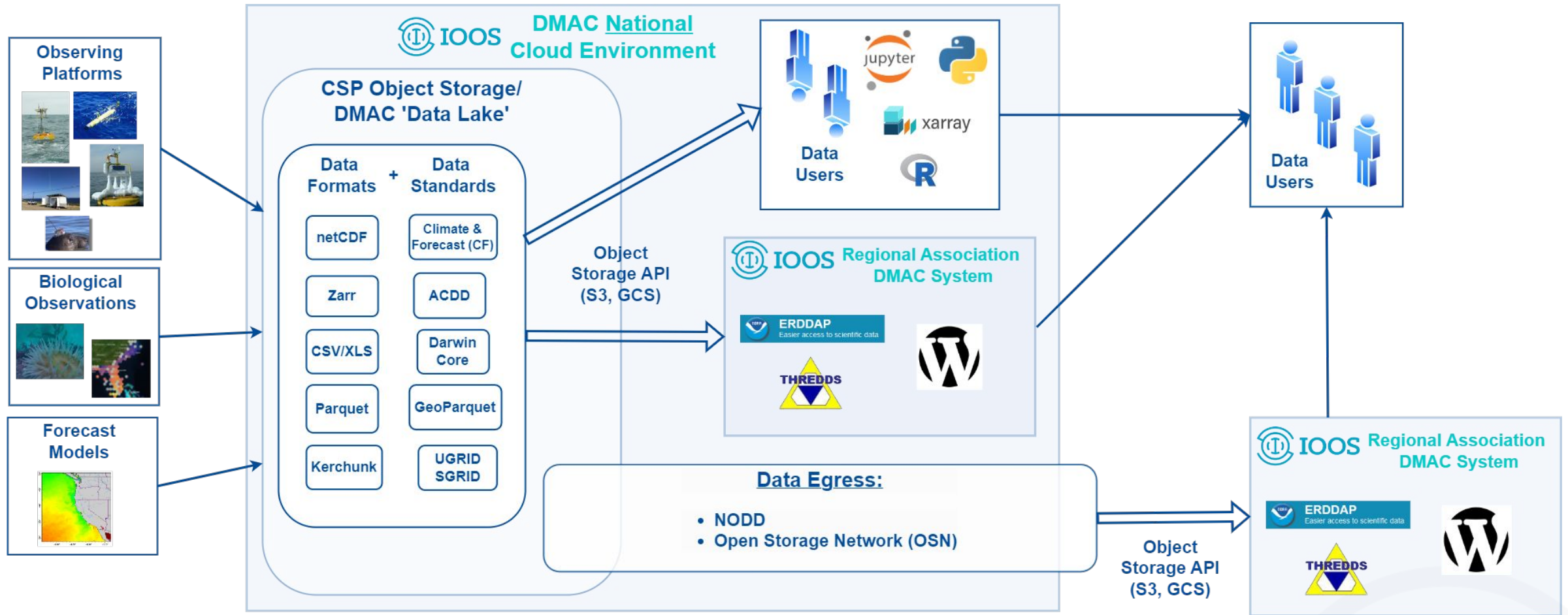
Project Goals:

- Determine feasibility of a cloud-based National DMAC cyberinfrastructure platform to allow RAs an alternative to operating their own DMAC systems
- Provide baseline DMAC capability & allow RAs a baseline capability to develop regionally-focused tools & products upon a shared CI
- Better serve National-scale users of IOOS' data by providing a means of access that crosses RA boundaries
- Support IRA implementation/connection with OTT Offshore Operations project
- Assess potential for Enterprise-wide DMAC cost savings

Stakeholders:

- MARACOOS (IIJA recipient), TetraTech & IOOS DMAC staff
- RA DMAC staff/data managers: provide requirements, co-develop prototypes, test, and validate with regional data

Conceptual DMAC National Cloud-based Architecture



Hypothetical National DMAC cloud-based platform including consolidated/shared shared object storage to host IOOS' data. Considerations:

- Data egress challenges, varying cloud 'user profiles', and individual RA appetite for adoption of cloud for their own data/product development - must support hybrid workflows

NOAA AI-Ready Data Ecosystem Priorities

1

*Uplifting
high-priority
data sets*



3

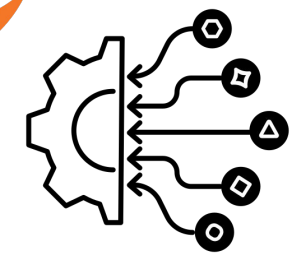
*Developing
tools and
services*

**NAIRR Pilot
Industry Proving Ground
World Data System**

**Climate Data Records
Tropical Cyclones
Offshore wind
Underwater Acoustics
Ocean Profiles**
...

2

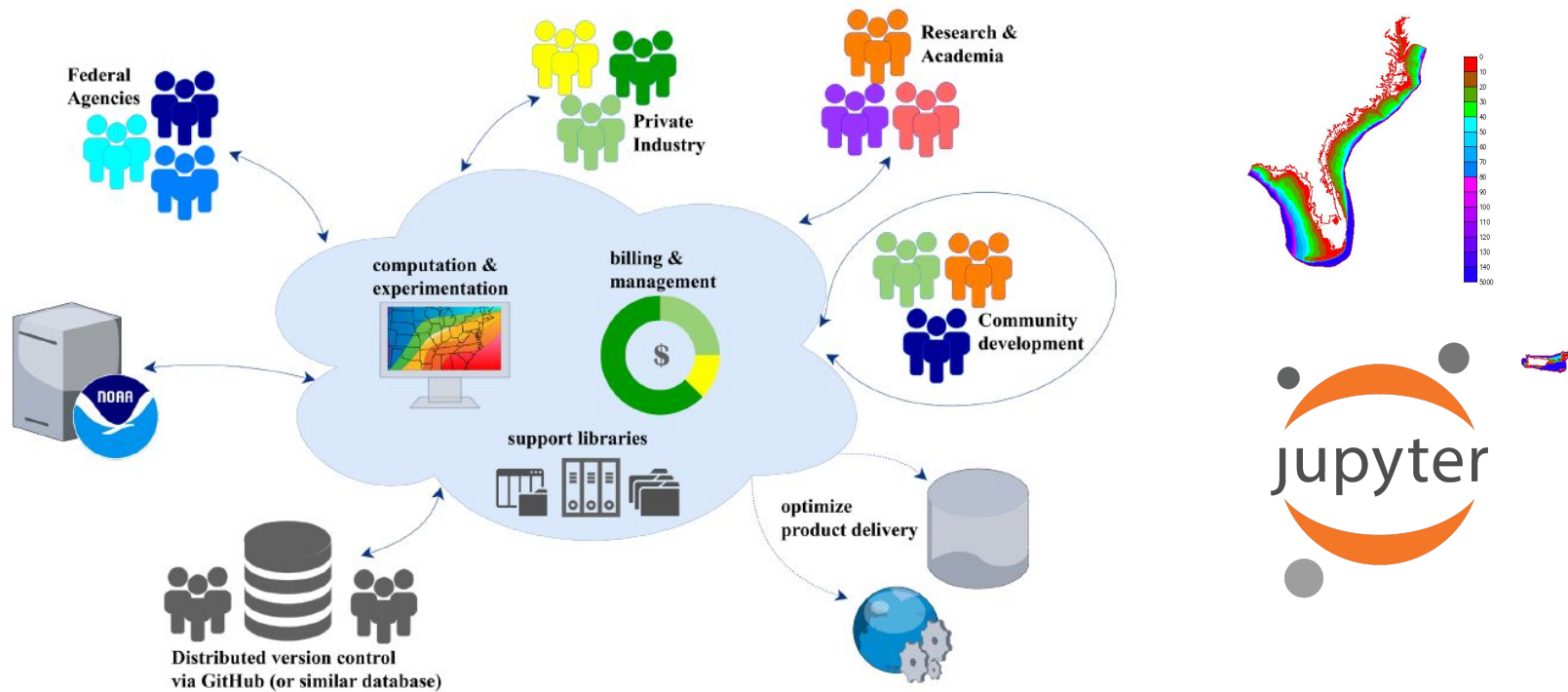
*Adopting
community
standards*



4

*Expanding
users' capacity*

Coastal Modeling Cloud Sandbox (BIL/IIJA)



Project Goals:

- Collaborative environment for developing new models
- Enable post-processing workflows adjacent to model outputs in cloud (JupyterHub)

NOS Programs:

IOOS, OCS, CO-OPS

NOAA Partner: **OCIO**

Developers:

TetraTech/Lynker, AWS

IOOS Cloud
Sandbox (RPS
hosted)



NOS Coastal
Coupling
Sandbox
(NOAA hosted)

DMAC CoP: AI Activities, Initiatives, Plans

Mathew Biddle, DMAC Community Manager

IOOS DMAC Meets ESIP AI Data Readiness



Data Preparation

- Have outliers been identified? Yes



Data Quality

- Have measures been taken to ensure completeness? YES - QARTOD



Data Documentation

- YES. IOOS Metadata Profile



Data Access

- YES. ERDDAP/THREDDS



Recent DMAC Community AI-Related Activities:

- DMAC Annual Meeting 2023, 2024 - several breakout discussion topics on AI/DMAC
- IOOS Code Sprint 2024 - [AI Ready Data Checker](#)

ESIP Data Readiness Cluster (2023): Checklist to Examine AI-readiness for Open Environmental Datasets v.1.0. ESIP.
Online resource. <https://doi.org/10.6084/m9.figshare.19983722.v1>

DMAC Community Developed Tooling: erddap2mcp

Usage Examples

RD Okay, use this dataset ID usf-sam-20210512T0000 and do a test. Give me detailed output so I can show my colleagues how cool this is

get_dataset_info

Automated Ocean Data Discovery & Analysis

Here's what our ERDDAP MCP server just discovered about this Gulf of Mexico research mission in seconds:

Mission Overview

This is **USF glider "sam"** - an autonomous underwater vehicle deployed by the **University of South Florida** as part of the **GOMIX** (Gulf of Mexico mixing) research program. The glider operated for **12 days** (May 12-24, 2021) collecting high-resolution oceanographic data.

Mission Coverage

- **Geographic Span:** ~120 nautical miles across the Gulf of Mexico
- **Latitude Range:** 27.69°N to 27.81°N (covering ~8 miles north-south)
- **Longitude Range:** 84.81°W to 83.26°W (covering ~95 miles east-west)
- **Location:** West Florida continental shelf - a critical oceanographic region

Scientific Instrumentation

The glider carried a **Sea-Bird CTD** (Conductivity-Temperature-Depth) sensor package collecting:

Physical Oceanography:

- **Sea water conductivity** → Salinity calculations
- **Temperature profiles** → Thermal structure
- **Density measurements** → Water mass identification
- **Depth profiles** → Vertical water column structure



Bob Currier
Data Scientist/Product Developer
robertdcurrier@tamu.edu



Why ERDDAP2MCP?

- ERDDAP is powerful but complex
- Scientists want to focus on research, not data access
- AI assistants need structured access to live data

File ID	File Name	File Type	File Size	File Date	File Description	File Status	File Location	File Access	File Permissions	File Owner	File Group	File Mode	File Creation	File Modification	File Deletion
...



Bob Currier
Data Scientist/Product Developer
robertdcurrier@tamu.edu



DMAC Tech Webinar recording from Bob Currier, Data Scientist/Product Developer @ GCOOS:
<https://ioos.noaa.gov/data/dmac-webinar-erddap2mcp-making-erddap-accessible-to-ai-assistants/>

GitHub repository: <https://github.com/robertdcurrier/erddap2mcp>

Ocean Technology Transition (OTT)

LOW COST SENSORS/ACCESSIBLE OCEAN TECHNOLOGY

Low cost sensors enable more observations over a wider area to fill data gaps and provide a lower bar to entry for new markets and developers.

APPLICATIONS OF MACHINE LEARNING AND AI

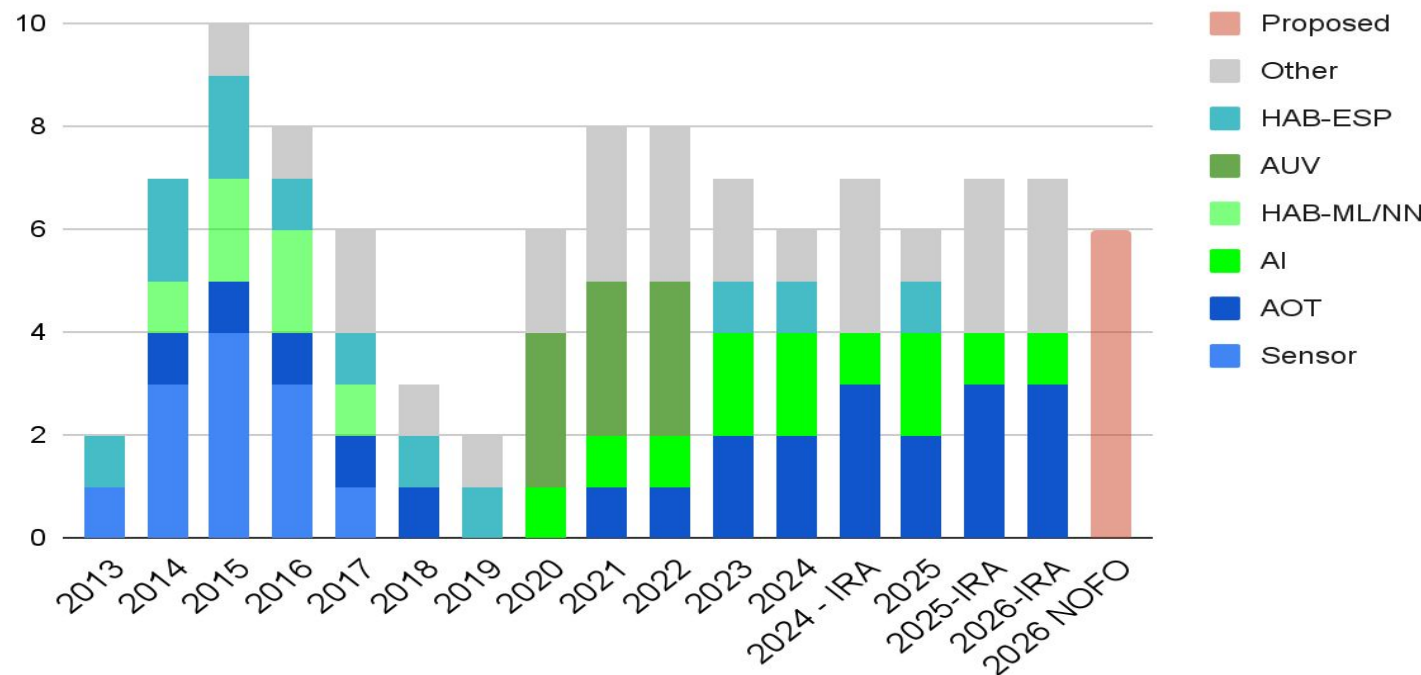
Advances in machine learning and artificial intelligence are enabling new ways to automate, interpret, and apply ocean observation data.

HABs and HYPOXIA

Improved sensors address water quality issues including Harmful Algal Blooms (HAB) and Hypoxia.

PHYSICAL OCEAN OBSERVATIONS

Observing physical oceanographic variables such as sea surface currents, temperature, salinity, external and internal tides, and surface waves support a vast array of stakeholders and missions.

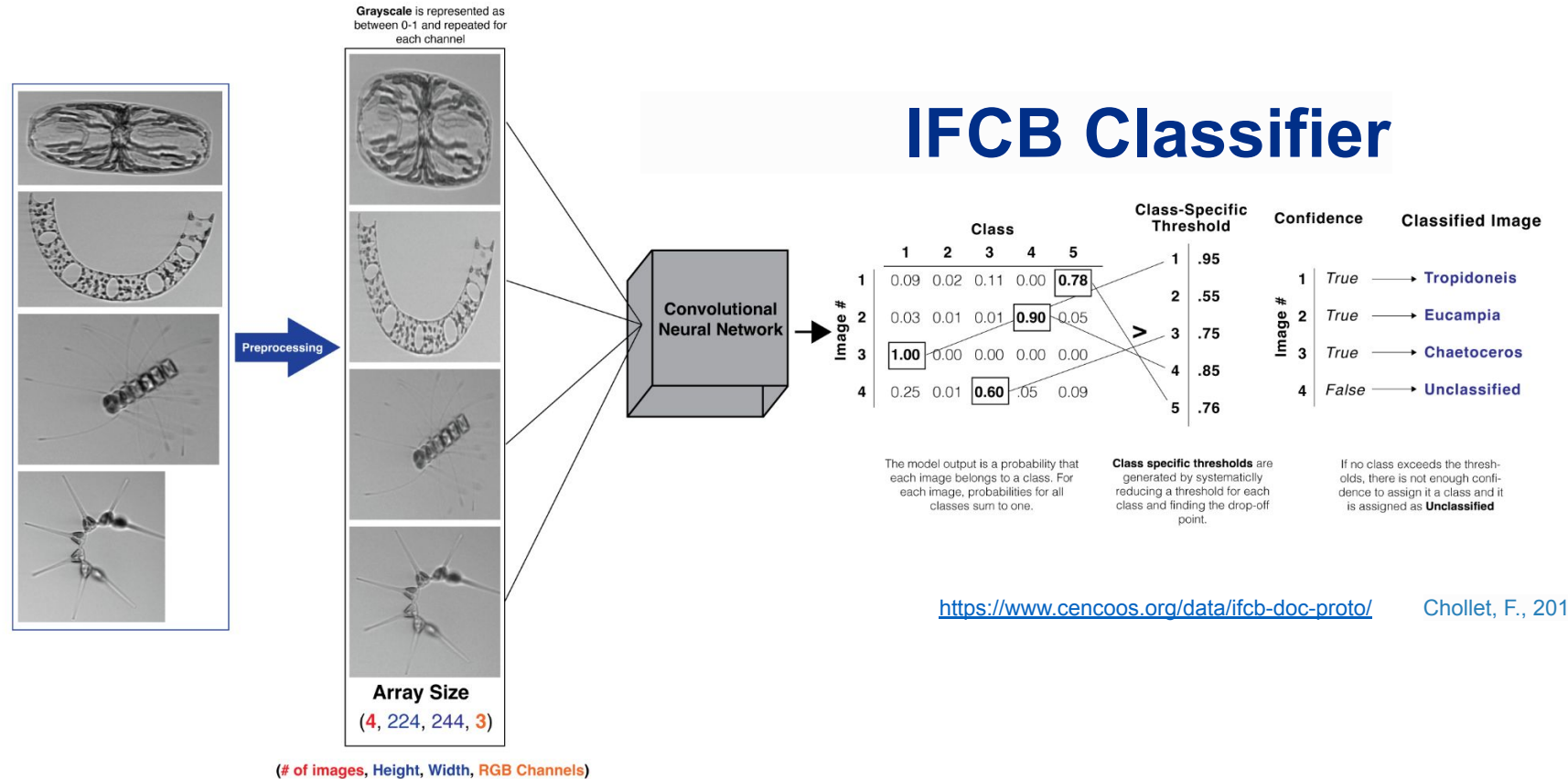


Classifying Images Collected by Imaging FlowCytobots (IFCB)



CA Imaging FlowCytoBot (IFCB) Network
 6 IFCBs - OPC Funding (PI: C. Anderson)
 1 IFCB - OCSD Funding (Robertson)
 1 IFCB - ECOHAB Funding (PI: A. Allen)
 1 IFCB - Simons Funding (PI: A. Barton)
 3 IFCB - NOAA Funding (PI: R. Kudela)

California IFCB Network (12 IFCBs)



<https://www.cencoos.org/data/ifcb-doc-PROTO/> Chollet, F., 2017



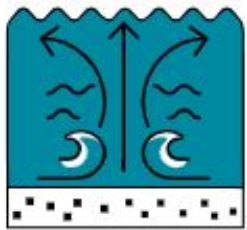
AI Methods to Detect Rip Currents

- New AI methods detect rip currents in real time using surf zone video imagery
- Developed through NOAA IOOS-funded [WebCOOS](#), in partnership with SECOORA, CO-OPS, and UC Santa Cruz
- Can be used to improve NOAA's rip current forecast model and inform NWS surf zone forecasts



de Silva, A., I. Mori, G. Dusek, J. Davis, and A. Pang, (2021). [Automated Rip Current Detection with Region based Convolutional Neural Networks](#). *Coastal Engineering*.

WebCOOS partners with communities to use webcams to:



Identify Rip Currents



Study Beach Erosion



Monitor Beach Usage



Flood Monitoring



de Silva, A., M. Zhao, D. Stewart, F. Kahn, G. Dusek, J. Davis and A. Pang. (2023). [RipViz: Finding Rip Currents by Learning Pathline Behavior](#). *IEEE Transactions on Visualization and Computer Graphics*.

NOAA/NOS AI Activities & DMAC Connections

Hassan Moustahfid



2025

Key High Priority Themes from the Workshop:

- Workforce Readiness & Education: Many sessions emphasized the need to build internal expertise and training, particularly around AI/ML and cloud tools. There's demand for better education for non-experts around topics like omics data and metadata standards.
- Governance Challenges: Speakers and participants noted governance, funding, and policy alignment (especially regarding AI) as ongoing hurdles.
- Tool Awareness: Attendees showed strong interest in learning about tool features and applications (e.g., NOAA's Collection Metadata Editing Tool (CoMET), Great Lakes Atlas of Multi-Omics Research (GLAMR), Environmental DNA To Ocean Biodiversity Information System (eDNA2OBIS))
- Cross-cutting Issues: Sessions frequently bridged IT and data management, highlighting the importance of collaborative governance structures.

Key Topics Across Sessions

- AI and Machine Learning
- Cloud Data Infrastructure
- Metadata Management
- Cross-Line Office Collaboration and Communication
- Data Value, Access, and Usability

SUMMARY RESULTS FROM NOAA ENTERPRISE DATA MANAGEMENT WORKSHOP

Integrate AI/ML and GenAI

Handbook

Comprehensive Guidance on Trustworthy AI, Managing AI-synthesized data, Defining FAIR and AI-ready concepts for GenAI, Interactions between third-party AI agents and NOAA's data services.

List of IOOS Regional Associations (RA) AI Activities

IOOS RAs	AI/ML Activity	Description
SECOORA	WebCOOS- Webcams for Coastal Observations and Operational Support	3 teams/application areas using beach webcams or cameras of opportunity for AI/ML feature detection- https://secoora.org/?s=WebCOOS and https://ioos.noaa.gov/project/ott-webcoos/
GCOOS	HABscopeV2- An AI tool for use by citizen scientists to facilitate early warning of HABs	HABscopeV1 drives the NOAA respiratory forecast which is visualized by GCOOS at https://habforecast.gcoos.org
CenCOOS/SCCOOS	Classifying images collected by IFCB	The UCSC IFCB Classifier is a trained classification model used for classifying images collected by an Imaging FlowCytobot (IFCB)- https://www.cencoos.org/data/ifcb-doc-protol/
CenCOOS	FathomNet	Catalog of images and AI models for classifying marine life. Generalization of code for use by all users.
GLOS	Identification and enumeration of freshwater phytoplankton, specifically cyanobacteria	CNN trained to detect Microcystis from holographic images and work is underway to add more species to the classification.
NANOOS	Salmon IRA	Web-based annotation and CNN classifiers from WHOI
NERACOOS	High-Frequency Ocean Surface Current Radar - Advanced Identification of First-Order Line Settings Using Image Processing Methods	The ImageFOL methodology processes the HFR spectra as images to define the bounds of the first-order/Bragg region. A number of empirical ML adjustments are made within a feedback loop that works to prepare the image for segmentation analysis as well as analyze the results to group segments to define the first-order line (FOL) region.
SCCOOS	Kelp canopy classification (satellite imagery)	Random forest and CNN techniques
AOOS	Unpiloted Aerial System (UAS) Salmon Counts	CNN and other ML methods. Data generated from this project will assist the Yup'ik village of Quinhagak in determining total salmon counts to prevent overfishing and ensure conservation, especially regarding the king and chum salmon depletion in the Yukon and Kuskokwim drainages.
MARACOOS	Oceanographic High-Frequency Radar Data Preservation in Wind Turbine Interference Mitigation	Based on ML techniques. The focus to use empirical machine learning methods, aided by the physics of wind turbine interference (WTI), to estimate WTI signals to a high degree of accuracy such that they can be separated and removed from the oceanographic signal echoes of high-frequency radar ocean surface current and wave sensors;

IOOS work on Accelerating the use of AI and ML in Ocean Obs

IOOS, in collaboration with NOAA line offices and partners organize teams every year to participate in NOAA NVIDIA and Open ACC AI Hackathons since 2020. **Mostly in kind work (no funding)** .

2020 - CoralML team (IOOS & NMFS)

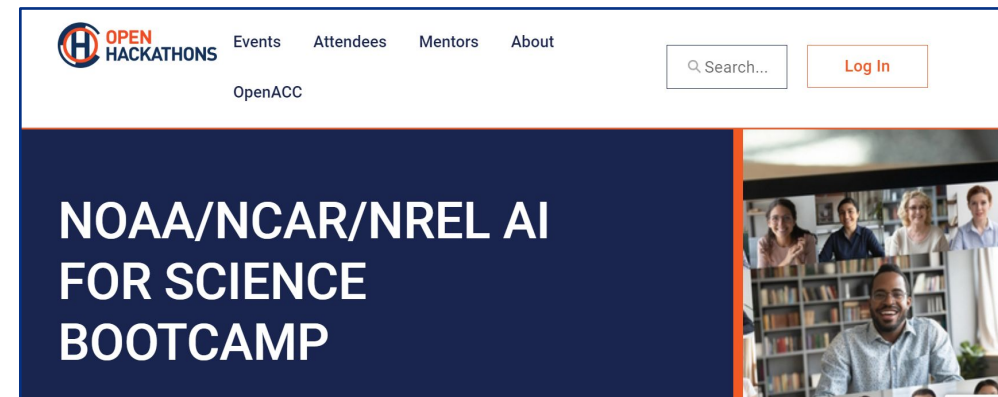
2021 - Plankton Automation team (IOOS & NMFS)

2022 - Metagenomics team (IOOS & OAR)

2023 - Water Level QC AI team (IOOS & NOS)

2024 - SeagrassAI team (IOOS & NESDIS)

2025 - Ocean Innovators AI- data Fusion and AI Emulation For Forecasting (IOOS & NOS)



For additional information:

(1) [NOAA CIO](#), (2) [Open Hackathons / NOAA/NCAR/NREL AI for Science Bootcamp](#)

Partnerships: Unlocking AI's Potential in IOOS



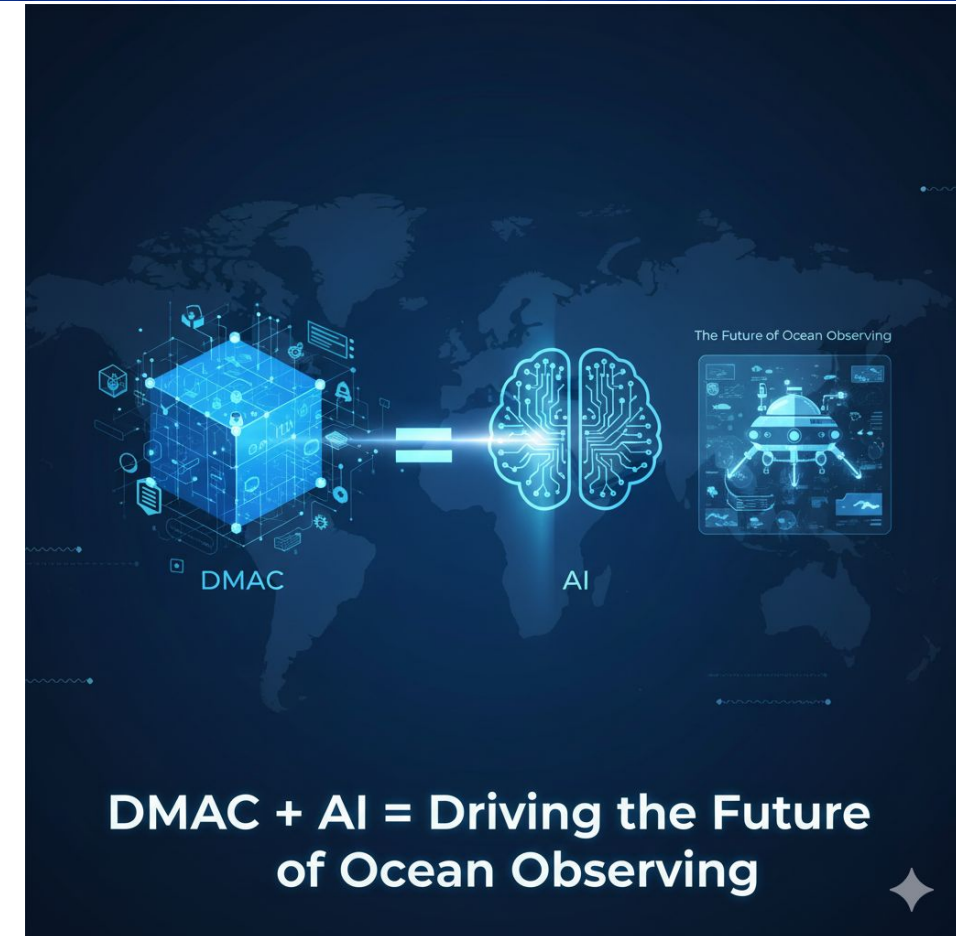
IOOS AI Leadership and Representation

- Active member of GenAI WG
- Active Member of NOAA AI Executive Committee (NAIEC)
- Chair NOS AI WORKING GROUP
- Co-Chair of the International of the Exploitation of The seas (ICES) Working Group on Machine learning for marine Science (WGMLEARN)

- Chairing Sessions at the ICES ASC 2026- “Rethinking Marine AI: Limitations, Failures and Future Directions”
Session at the CIAC 2025” AI: Unlocking New Frontiers in Marine Research

Conclusion

- **With limited and mostly in-kind resources, IOOS is already leveraging AI/ML to advance DMAC Cyberinfrastructure and Products development.**
- **IOOS is contributing to setting national standards and practices and, in some areas, leading AI implementation within NOAA for ocean observing systems.**
- **IOOS is building enduring AI capacity within NOAA, ensuring sustainable innovation for ocean data management and applications.**

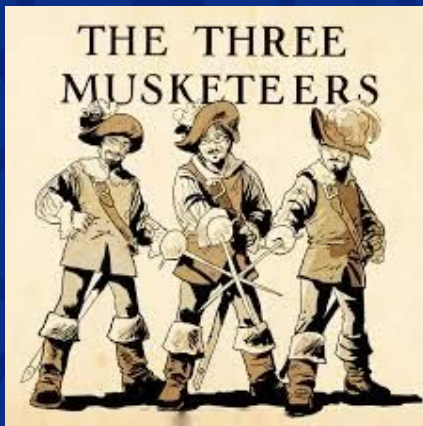


Challenges → Opportunities

- **Integration with Infrastructure** → AI models must fit diverse, legacy systems and real-time data feeds.
- **Stakeholder Coordination** → IOOS Office, RAs etc. Private partners must align for accessibility & usability.
- **Trustworthy AI Standards** → Need for transparency, fairness, and accountability in model predictions and decisions.
- **Resources & Workforce** → need funding to accelerate AI and need for skilled staff.
- **Computational Infrastructure** → Sustaining access to cloud/HPC while managing cost and security.



Thank You



Micah Wengren - micah.wengren@noaa.gov

Hassan Moustahfid - hassan.moustahfid@noaa.gov

Mathew Biddle - mathew.biddle@noaa.gov