# Progress Toward an Al-Ready Agency

### NOAA Center for Artificial Intelligence

NIC AND ATMOSPHE

NOAA

97MENT OF COMM

MINISTRATION

October 5, 2022

Rob Redmon (NCAI Lead, noaa.gov/ai) And the NCAI Team

# Agenda

### **NOAA's AI Strategy**

- Call to Action
- NOAA's many Use Cases
- NOAA's Approach
  - National AI Initiative, Strategies and Plans

### **NOAA's Center for Artificial Intelligence**

- Community of Practice
- Developing an "Al-ready" data standard
- Pilot Projects (just a subset)
- Training the Workforce
- 4th NOAA AI Workshop

### noaa.gov/ai



A place for publicly connecting to NOAA's 700+ member Community of Practice around AI for Earth system science to develop synergies and partnerships NCAI Mailing List: <u>tinyurl.com/y2ehvhfg</u>

## **AI Call to Action**

#### Needs:

#### Integrated AI Lifecycle

- Agencies trained on building Trustworthy & Equitable AI applications
  - For Digital Twin Earth systems fueling economic & equitable climate policies
  - Automating and augmenting stewardship of our national critical ecosystem resources and environmental observations
- AI/Data science-ready open environmental datasets
  - To power the Digital Blue Economy, Climate Ready Nation, Social Environmental Justice, and federated National AI Research Resources

### **Recommendations:**

- Develop AI training standards centered around interactive training tools as "Learning Journeys" demonstrating Lessons Learned
  - This ensures AI applications are Trustworthy, maximizes technology transfer and directly benefits Research to Applications and Sustainment.
- Develop AI-ready data standards for open environmental data
  - This is the future of data stewardship, building upon Findable, Accessible, Interoperable and Reusable (FAIR principles) and Analysis Ready partnerships



ESIP: How much time do you spend to find, access and pre-process AI data?



### U.S. 2021 Billion-Dollar Weather and Climate Disasters



NOAA 2022 AI Data Call - Application Domains



NOAA





### Example AI @ NOAA

Marine Life Speciation using Video Image Analytics for the Marine Environment (VIAME)



VIAME helps automate the detection and identification of fish species captured by video

Video adapted from: https://videos.fisheries.noaa.gov/detail/videos/science-technology/video/6255809190001/video-image-analytics-for-the-marine-environment



### Example AI @ NOAA

### Automated Rip Current Detection with Region based Convolutional Neural Networks



#### Flow-Based Rip Current Detection and Visualization (IEEE) doi:10.1109/ACCESS.2022.3140340 Gregory Dusek (NOS) and UC Santa Cruz

Debra Hernandez, Southeast Coastal Ocean Observing Regional Association (SECOORA) Executive Director:

"Whether it's identifying a right whale or a rip current or shoreline erosion, we need faster analysis for more effective alerts to inform decision-makers."

secoora.org/noaa-launches-a-new-life-saving-rip-current-model

Video: https://arxiv.org/pdf/2102.02902.pdf

# NOAA's Approach



<u>National AI Initiative Act of 2020</u>: "The Administrator of NOAA [...] shall establish, a Center for Artificial Intelligence"

#### Several Executive Orders, including:

- "Maintaining American Leadership in Artificial Intelligence"
- "Tackling the Climate Crisis at Home and Abroad"
- "Protecting Public Health and the Environment and Restoring Science To Tackle the Climate Crisis"

#### **Related NOAA Strategic Plan Goals & Objectives**

#### Foster an Information-Based Blue Economy:

NOAA will introduce innovation to data collection through various in-situ methods for species detection and explore AI/ML and data visualization technologies... Ensure accessibility and enable an enterprise climate information framework to meet the needs of NOAA's users:

Analysis-ready datasets available (or percentage of existing satellite/other observational data made Al/ML ready on the cloud for climate, weather, oceans, etc. products and services)



NUAA Artificial Intelligence Strategy

sciencecouncil.noaa.gov

#### 2022 Data Call included 261 projects (188 in 2020)





NOAA Center for Artificial Intelligence noaa.gov/ai

# **Community of Practice 705 Members**

Members in NOAA's AI Community of Practice from USG, Academia and Industry are looking to NCAI to facilitate conversations around infusing AI into Climate, Wx, Ecosystems and Environmental Justice.

- 705 (as of October 2022): more than half are NOAA
- 109 Organizations
  - o NASA, USGS, USCG, WYO, USDA, US Navy, NREL
  - University of Colorado and Colorado State, Hawaii, North Carolina State, Albany, Alaska, Montana, Massachusetts, Montana, UC San Diego, Texas, Exeter and many more
  - AECOM, AccuWeather, BAH, IBSScorp, Riverside, tomorrow.io, Raytheon, and many more

### Join the mailing List: tinyurl.com/y2ehvhfg

NCAI Community of Practice Affiliation



Timestamp - Year-Month

Al-Ready Data Why?, What? and How?



30

Time [seconds]

NOAA acoustic data used in deep learning to identify whale songs

fisheries.noaa.gov/science-blog/ok-google-find-humpbackwhales

### Goal: users spend less time data wrangling, more time on AI / ML

How can users find data that is easy to use in AI/ML? How can data providers assess and improve usability?

#### What's needed:

- Community-driven definition of AI-readiness requirements
- Assessment tools for data providers and stewards and users
- Representation of readiness level conditioned on use / application
- Feedback and iteration to improve the standard
- Ideally, a formal standard (or set of standards)

### **ESIP Data Readiness Cluster:** This Earth Science Information Partners (ESIP) collaboration is working to to develop an AI-ready data standard. Membership includes: US Federal agencies, academia, NGOs, private sector, and international (wiki.esipfed.org/Data\_Readiness)

Click here for <u>AI-</u> <u>ready "glossy"</u>, including project summary and the latest AI-ready checklist.

## • OK Google: Find the Humpback Whales

NOAA



### ESIP AI-Ready Data Collaboration (July 19-22)

ESIP Session: "Enabling AI Application for Climate: Developing A Collection of AI-ready Open Climate Data – Data-A-Thon"

### Goals

- Test AI-ready data assessment tool through community participation
- Assess AI-readiness for climate datasets from NOAA, NASA, USGS, and others, including NOAA climate data records (CDRs)

### Outcomes

- More than 40 participants evaluated AI-readiness for selected climate datasets
- Key feedback takeaways:
  - Training materials are needed to better onboard the community for Alreadiness assessment
  - Requirements for AI-readiness can be use case driven; need to balance specificity and general data management improvements
  - Position AI-readiness assessment as a process to guide future data/services development and data improvement







### NCAI Pilot Project: Create Tropical Cyclone Model Training Dataset

#### Challenge

Al-ready and accessible benchmark satellite datasets are needed to drive the future of tropical cyclone trajectory, intensity and coastal impact prediction (e.g. coastal flooding and other infrastructure damage).

#### **Description and Expected Outcomes**

Evaluate a new dataset's AI-readiness against NCAI draft standards and make necessary changes to brand it as AI-ready. The new dataset, Tropical Cyclone PRecipitation, Infrared, Microwave, and Environmental Dataset (<u>TC PRIMED</u>), collocates and subsets LEO/GEO satellite imagery with ancillary model information to create a 22-yr dataset of TC-centric scenes. This dataset will supersede NCEI's <u>HURSAT</u>.

#### **NCAI Benefits**

Al-ready standard maturation; Lesson Learned via interactive Python notebook ("Learning Journey"); increased collaboration with NOAA's NODD (previously BDP).

POC: Chris Slocum (NOAA/NESDIS/STAR), christopher.slocum@noaa.gov



A sampling of TC PRIMED products from Typhoon Maria (2018) at 10:13 UTC on 9 July 2018 in the western Pacific, where a) is GPROF, b) is GPM DPR precipitation rate, c) GPM DPR reflectivity, d) 36.6 GHz, e) 89 GHZ, and IR from Himawari-8.



### Initiative: Training the Workforce Powering Discovery and Innovation

### Training + Al-ready data $\rightarrow$ Trustworthy + Equitable Services

Need **NOAA-specific training material** using NOAA data and computing resources to **remove common barriers** to the "Research to Operations, Applications, and Services" pipeline.

To address needs, resource creation should be prioritized to **convert NOAA AI success stories** into interactive training material in a **sandbox computing environment** that allows the workforce to apply learning outcomes to support NOAA's mission via the AI Strategic Plan.

Factsheet: <a href="mailto:noaa.gov/ai/training">noaa.gov/ai/training</a>



NOAA training action priority lifecycle highlighted by workforce role and relationship to AI.

## **Training and Learning Journeys**

#### **NCAI Github status**

 Github SOP signed and in effect; NCAI Github is expected to be established in FY23 Q1.

#### Learning journey progress

- Facilitate ongoing contributions from NOAA AI WG and communities of practices (rip current, MagNet, probabilistic forecasting & habitat modeling).
- Exploring training sandbox environments on the cloud for NOAA employees/affiliates

#### Partnership development

- Collaboration with AI2ES on the Trustworthy AI Summer School for Environmental Science.
  - more than 800 participants
  - space weather notebook for trust-a-thon
- Explore partnership with NCAR Pythia Project.

#### Classification Modeling for Hazardous Rip Current Prediction

- Andre van der Westhuysen<sup>1</sup>, Gregory Dusek<sup>2</sup>, Jung-Sun Im<sup>3</sup>, Mike Churma<sup>3</sup>, and John F. Kuhn<sup>4</sup>
- <sup>1</sup>IMSG at NOAA/NWS/NCEP Environmental Modeling Center, <sup>2</sup>NOAA/NOS/CO-OPS, <sup>3</sup>NOAA/NWS/Meteorological Development Lab,
   <sup>4</sup>NOAA/NWS/Analyze, Forecast and Support Office
- Version 0.1

#### Overview

This notebook explores binomial (two classes) classification modeling, using as application area the prediction of hazardoxs in currents in San Diego, CA. Rip currents are the leading source of drownings on U.S. beaches, with approximately 100 fatales per year. However, It is computationally expensive to predict using traditional physics-based models that resolve the detailed fluid opmarics that create the nip current structures. Instead. turning the prediction problem into a classification task using ML methods provides actionable guidance to NOAA forecasters at low computational cost.

#### Prerequisites

The student is assumed to have a basic understanding of statistics, in particular of linear regression. Fluency in Python and data manupulation with Pandas is assumed.

Learning Outcomes

From this notebook, the student will learn:

- · How a geophyscal modeling task can be restated as a classification problem, and how it can be solved using different ML models
- How to collect, clean, and transform the environmental data needed to solve the classification problem.
- · How to compare the quality of different classification models.
- How to set a probability threshold for identifying events, based on a trade-off between precision and recall.

Upon completion of this training, the student will be able to

- · Identify when a given geophysical problem can be cast as a classification problem.
- Locate, load and manipulate the necessary data for classification.
- Employ and evaluate different ML models to solve the classification problem.
- Quantify the performance of different ML classification models in terms of practical metrics such as the percentage of positive events
  detected (recall), and the accuracy rate of positive predictions (precision).

NCAI plans to expand training partnership via engagement with the Office of Education (e.g., NOAA scholarship) and LO Learning Officers (e.g., CLC programs) on workforce development.

### **Areas Resources could Accelerate Training R2X:** sandbox environments, development tools, Learning Journeys from NOAA's AI Inventory, etc.

Background on Hazardous Rip Current Prediction

# **Collaboration Building via Public Participation**



Fire Weather & Impacts

Community-driven use case ideation & development

Collecting feedback on NOAA data and Al-ready data need for fireWx research

Five use cases formed among 50+ active participants

Pathfinding for the emerging topic of digital twin for Earth system (DT4ES)

Interoperable Digital

Twins for Earth System

Community-driven discussion across the value chain of DT4ES

Active public-private-academic engagement & discussion

Exploring the current landscape and future directions of AI for ocean

Conservation

(Endorsed by UN Decade of Ocean)

Focused discussion on data and R2X challenges and needs

AI for Ocean Sciences &

Hands-on tutorials and training on emerging tools.

# **Workshop Highlights**



### More than 800 Registered

### **Invited speakers:**

Sarah Kapnick (NOAA) Stephen Belcher (Met Office) Nadine Alameh (OGC) Peter Dueben (ECMWF) Jacqueline Le Moigne (NASA) Evan Howell (NMFS) Mark Osler (NOS) Anthony Hoogs (Kitware) Margrit Betke (Boston Uni) Vinay Viswambharan (Esri) Everett Hinkley (US Forest Service) Ko Barrett (NOAA) Robyn Heffernan (NWS) David Danks (UC San Diego) Christine Curtis (Partnership on AI)



### **Connect With NCAI**

### noaa.gov/ai



NCAI Mailing List: tinyurl.com/y2ehvhfg



NOAA Center for Artificial Intelligence home

About NCAI

Events

Al/machine learning projects

NCAI Community of Practice

Training and resources