Accelerating AI Applications in Environmental Sciences

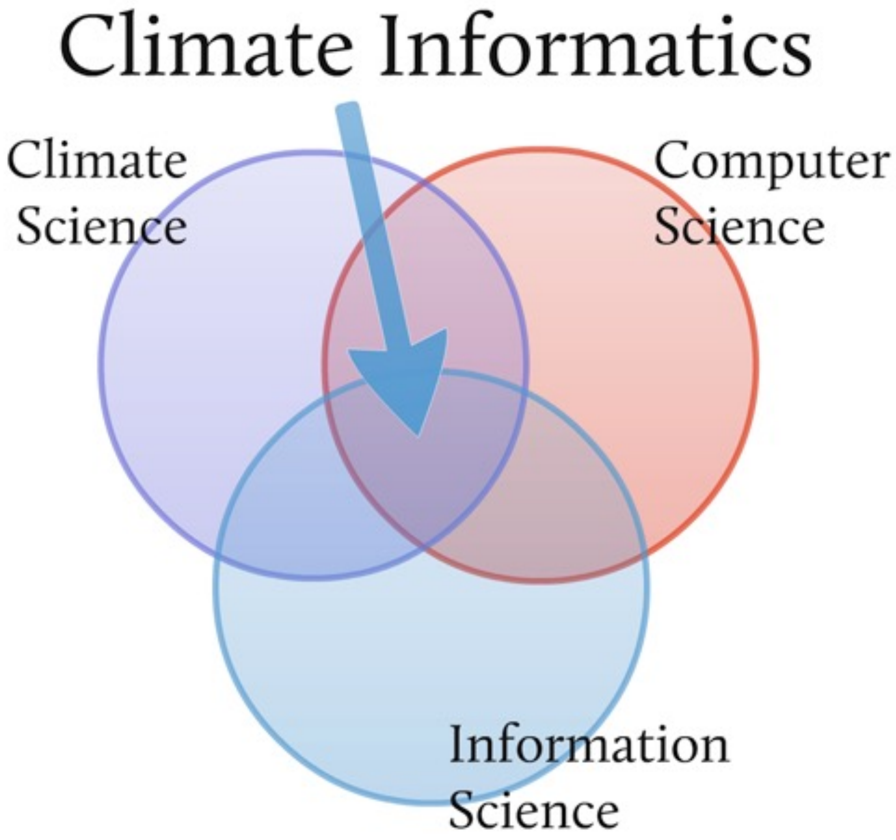
Yuhan "Douglas" Rao, PhD (douglas_rao@ncsu.edu)
Cooperative Institute for Satellite Earth System Studies/NOAA NCEI

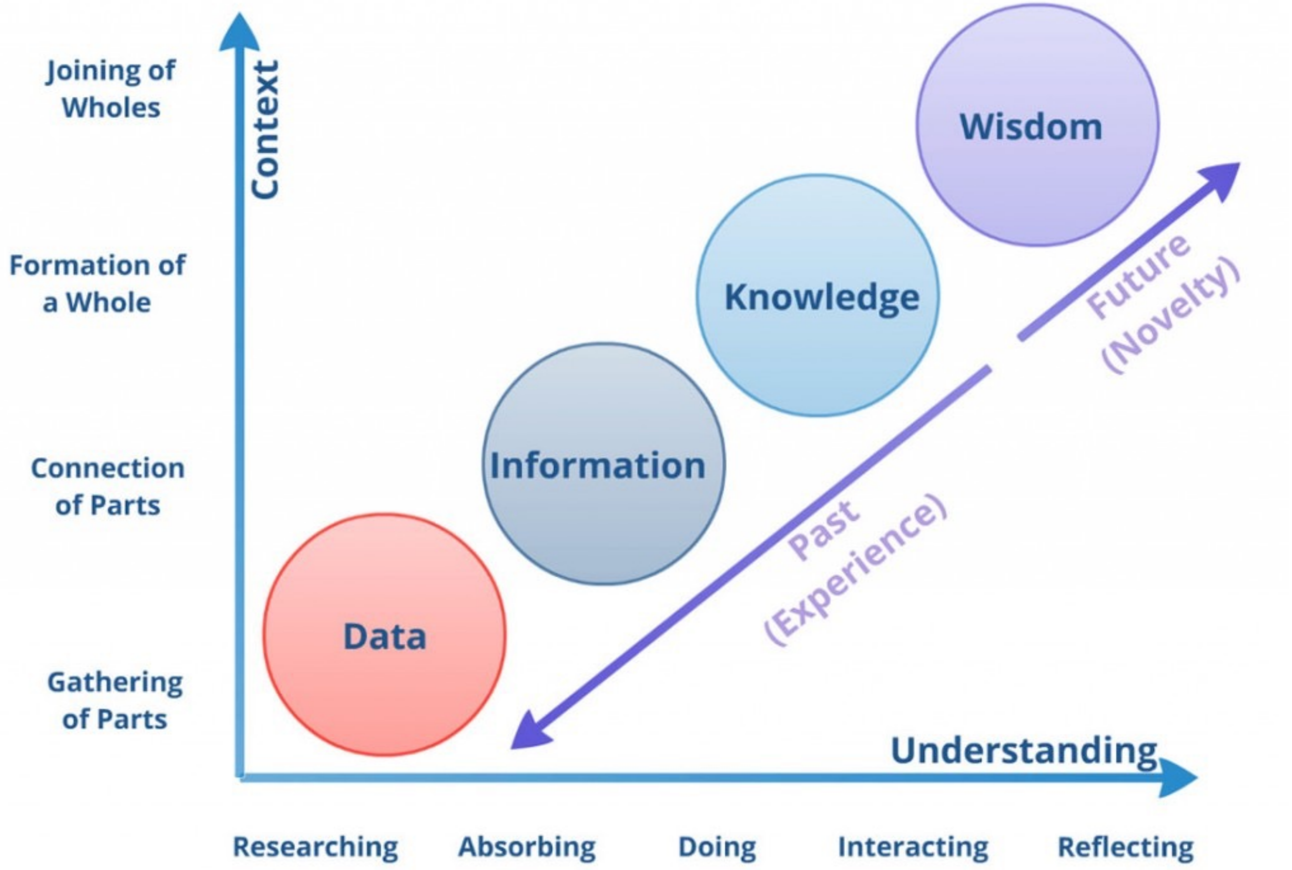
Rob Redmon (NOAA Center for AI), Eric Kihn (NOAA NCEI)

NOAA CENTER FOR ARTIFICIAL INTELLIGENCE

Climate Informatics

Climate Informatics





Source: S. Easterbrook, https://www.easterbrook.ca/steve/2012/09/what-is-climate-informatics/

Climate Informatics 2023

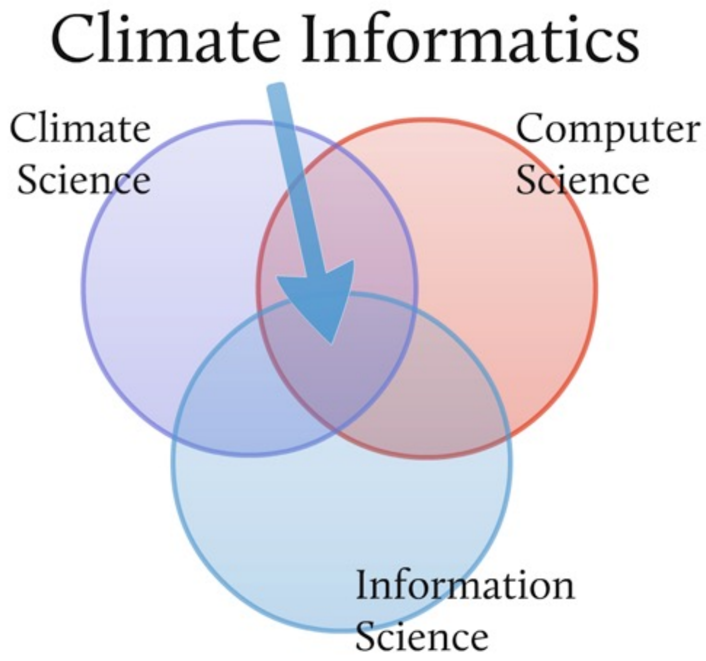
Data-driven prediction & forecast

Sub-seasonal to seasonal (S2S)

Extreme events

Model downscaling

Calibration & emulation



Generative model

xAI

Unsupervised model

Knowledge discovery

Causal inference

Uncertainty quantification

Trustworthiness & ethics of Socio-technical systems

Reproducibility & Replicability

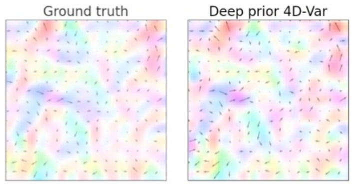
Data & Skills

Climate Informatics Reproducibility Challenge

ENVIRONMENTAL DATA SCIENCE

Ocean | Modelling | Special Issue | Python

Variational data assimilation with deep prior (CIRC23)

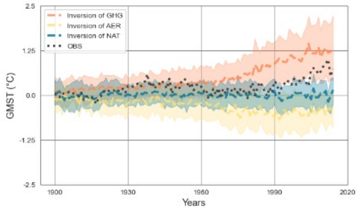


Pahari *et al.* (2023)

license MIT | launch binder | render passing | DOI 10.5281/zenodo.8339299

General | Modelling | Special Issue | Python

Deep learning and variational inversion for climate science (CIRC23)

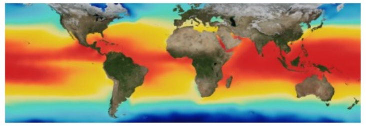


Domazetoski *et al.* (2023)

license MIT | launch binder | render passing | DOI 10.5281/zenodo.8330771

Ocean | Modelling | Special Issue | Python

Underlying physics of the ocean's temperature (CIRC23)

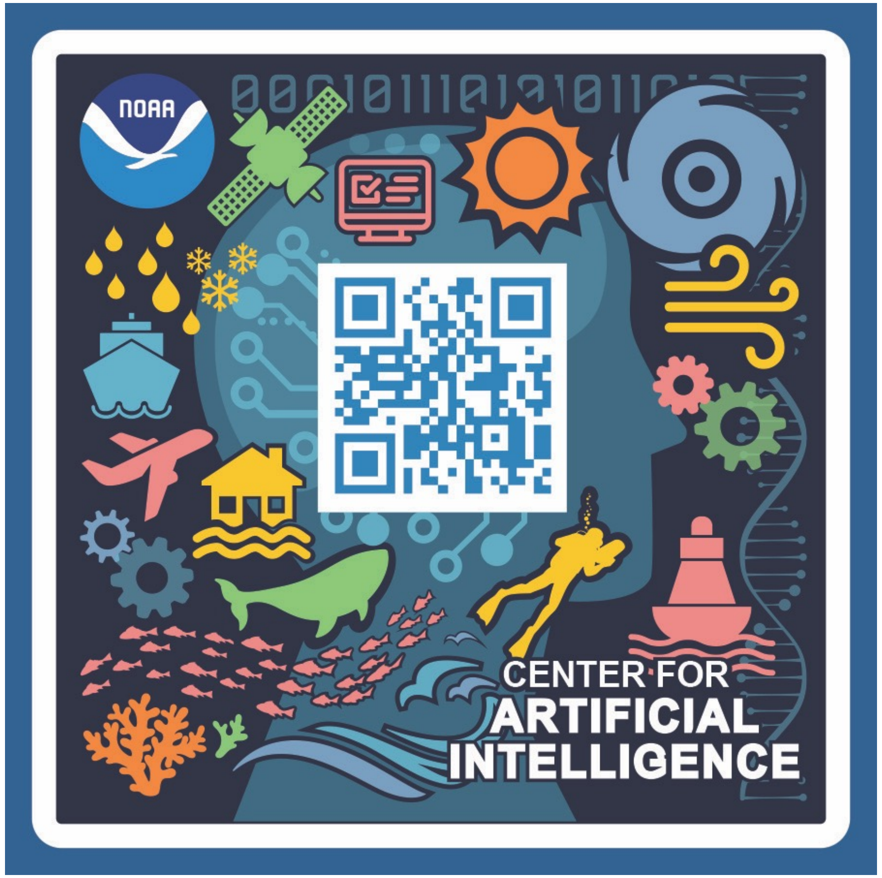


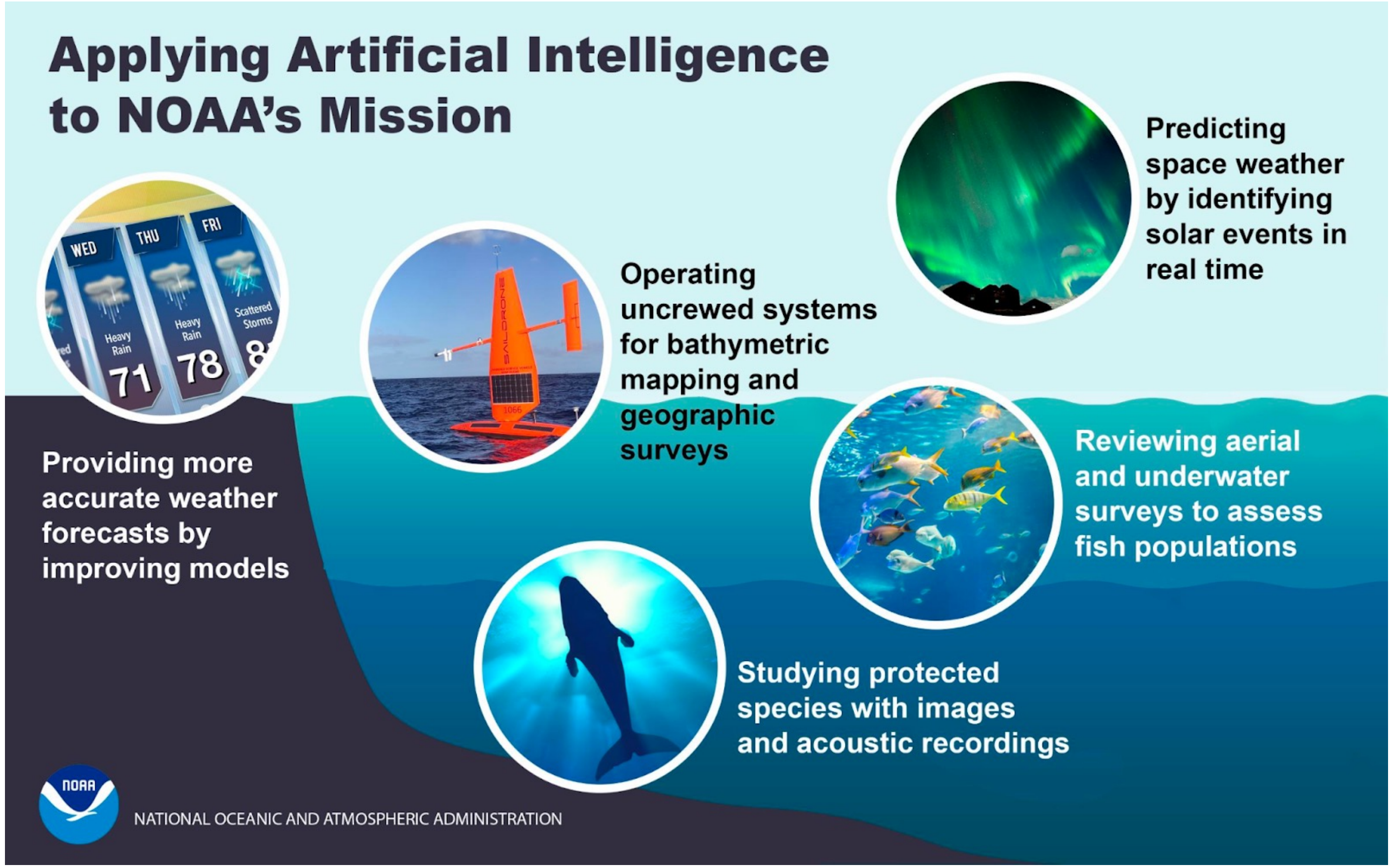
Malhotra *et al.* (2023)

license MIT | launch binder | render passing | DOI 10.5281/zenodo.8314669

From conventional scientific publication to computational notebooks with well documented reproducible workflows.

AI in Environmental Sciences

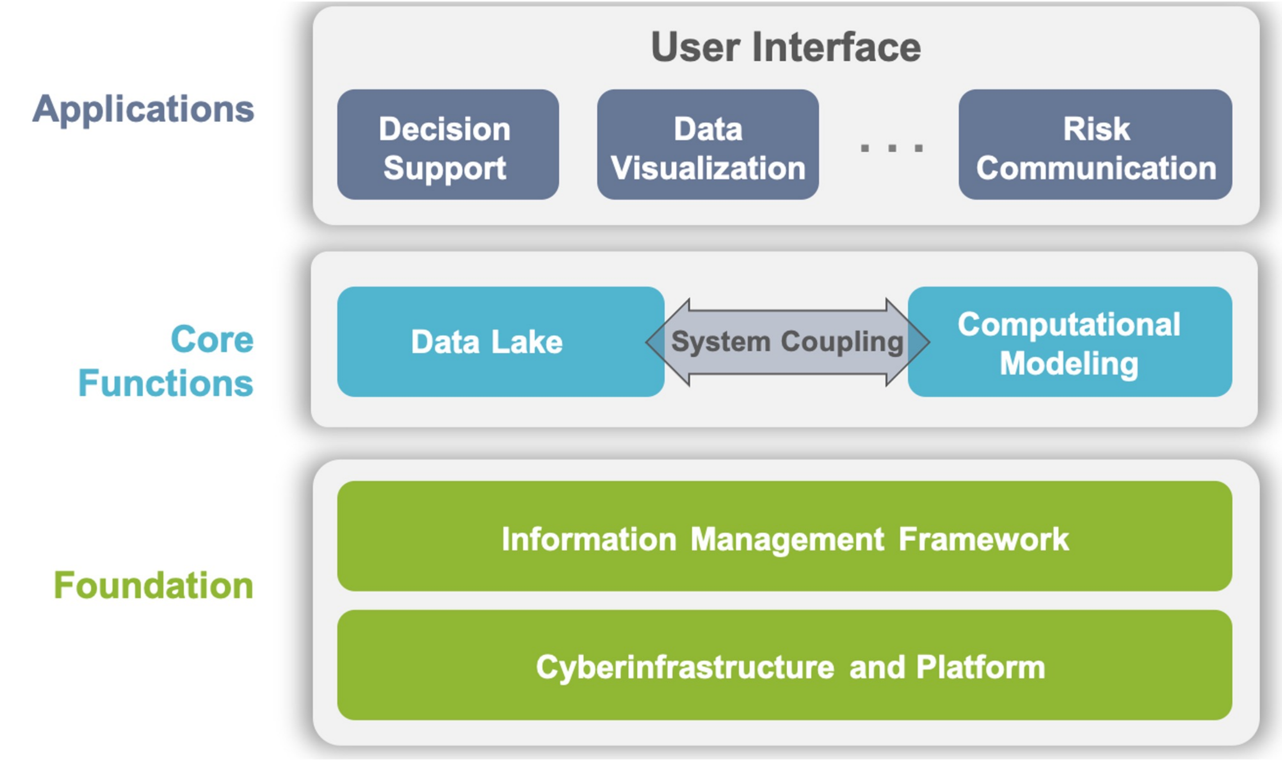




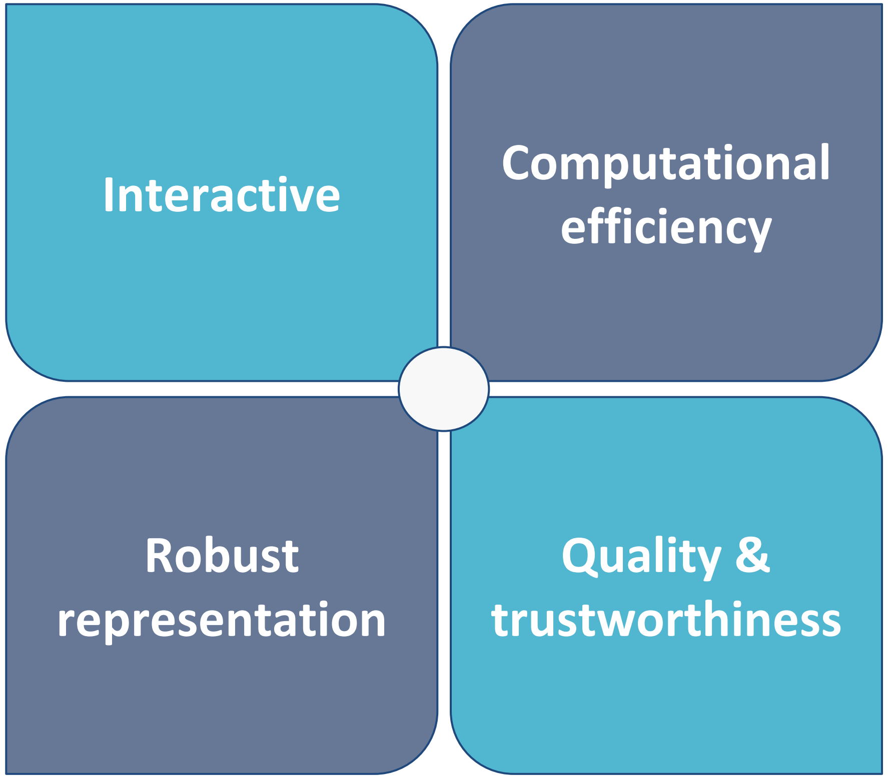
NOAA Center for AI strives to benefit NOAA's mission by proliferating the use of responsible AI through coordinated development, engagement, and partnership.

4th NOAA AI Workshop – Digital Twin Whitepaper

Core components



Core features



Rao et al., https://doi.org/10.48550/arXiv.2306.11175

5th NOAA AI Workshop

AI Benchmarking Frameworks

Order 1 requirements

- R1: Data available online without access restrictions
- R2: Clear problem statement for meaningful task in atmospheric science
- R3: Data input into high level open data science language provided
- R4: Evaluation metrics defined analytically and in code

Order 2 requirements

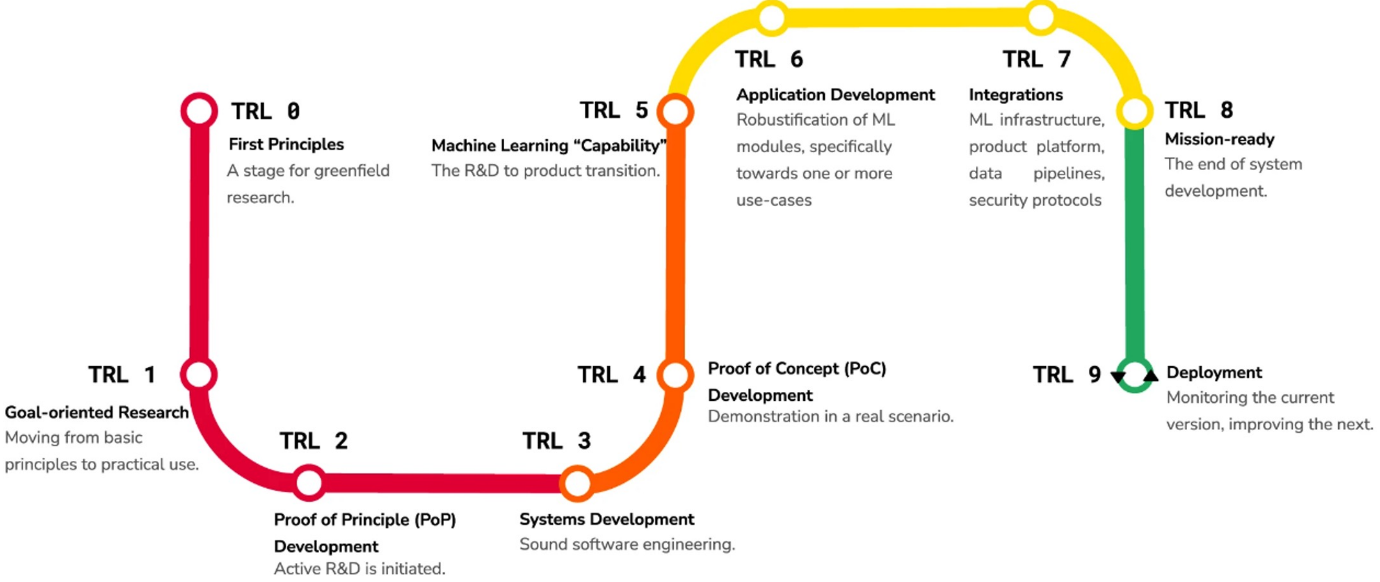
- R5: Simple example machine learning solution provided in code
- R6: Visualisation and diagnostics provided in code
- R7: Tests for physical consistency and explainability provided
- R8: Benchmarks for the computational performance provided

Dueben et al. (2022) https://doi.org/10.1175/AIES-D-21-0002.1

Facilitating R2X transitions for AI Research

Fig. 1: MLTRL spans research (red) through prototyping (orange), productization (yellow), and deployment (green).

From: Technology readiness levels for machine learning systems



Lavin et al. (2022): https://doi.org/10.1038/s41467-022-33128-9

Preliminary Take-aways

Identifying core use cases: community-driven core use cases for AI benchmarking framework development including data, tools, use case specific metrics (beyond RMSE).

Socio- and Cyber-infrastructure: efforts are needed to address the segmented infrastructures for AI R&D and social mechanisms to support community uptake of the benchmarking.

Addressing the need of AI-ready data: accessibility, quality, and documentation are critical for enabling benchmarking and R2X transitions.

Workforce and capacity development: critical needs to scale up education efforts for both current and future workforces with the emphasis on diverse AI workforce.

Evaluation & governance: critical needs to develop use case aware metrics to objectively measure and monitor AI benchmarking frameworks.

Ethics and risk management: need to adopting a common framework to evaluate, measure, and address risks associated with AI applications.