



**STC: LEARNING  
THE EARTH  
WITH ARTIFICIAL  
INTELLIGENCE  
AND PHYSICS**



**LEAP**

**VIRTUAL SITE VISIT, NSF SCIENCE AND TECHNOLOGY CENTERS  
TUESDAY, SEPTEMBER 22-WEDNESDAY, SEPTEMBER 23**

# House Rules for Today's Site Visit

## Everyone

- Find appropriate Zoom links in agenda
  - **Website:** <https://leap.research.columbia.edu/>  
Username: NSF  
Password: LEAP
  - Different rooms throughout day
  - **Back-Up Plan:** WebEx link emailed
- Email [mb3952@columbia.edu](mailto:mb3952@columbia.edu) for logistics/ IT support

## LEAP Non-Presenters

- Camera off, mic off
- Protect bandwidth
- Only presenters answer questions





# Welcome + Director's Overview



Center Director and  
Diversity Co-Director  
**Pierre Gentine**

- Associate Professor  
Earth + Environmental Engineering  
Earth + Environmental Sciences, Columbia University
  - PhD, Civil + Environmental Engineering, MIT

### **Expertise**

- Hydrologic cycle
- Land-atmosphere interactions
- Machine learning for geoscience

### **Other Synergistic Experience**

- Co-Founder + Co-Lead, Data Science Working Group, US CLIVAR
- Co-Lead, NOAA Drought Task Force
- Member, Working Group on Seasonal to Interannual Prediction, World Climate Research Program

# My Presentation Will Explain:

1. Significant societal problems requiring LEAP's solutions
2. Center vision + integration of all activities
3. Leadership + management strategies promoting an effective culture
4. Role of team personnel + partner organizations

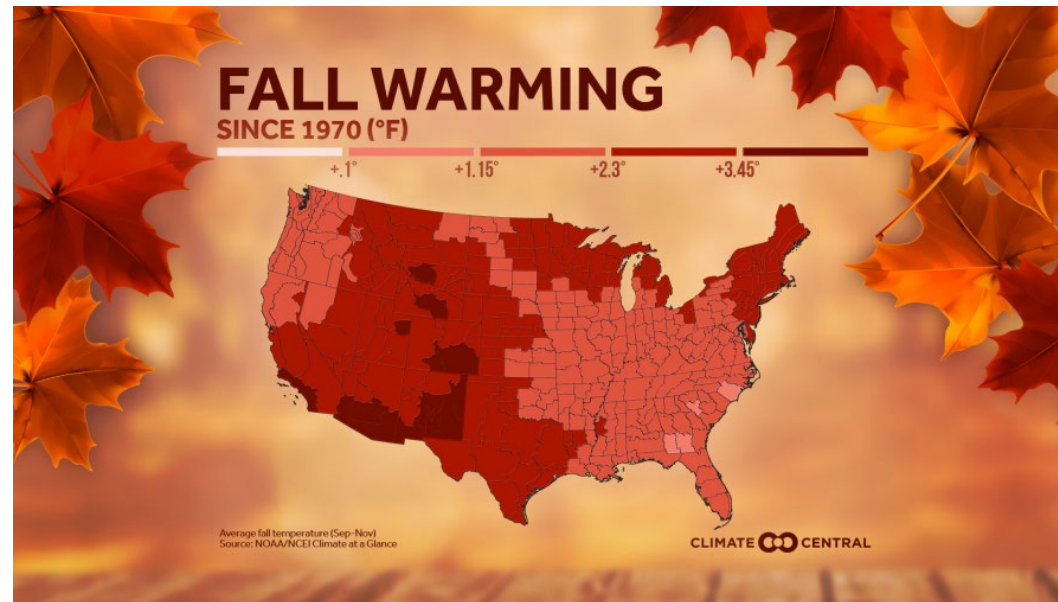
# Climate Change is *our* Grand Challenge



## Atmosphere

From 1880 - 2012, average  
**global temperatures**  
increased by **0.85°C**

- 2°C is maximum excess carbon budget



# Climate Change is *our* Grand Challenge



## Atmosphere

From 1880 - 2012, average global temperatures increased by **0.85°C**

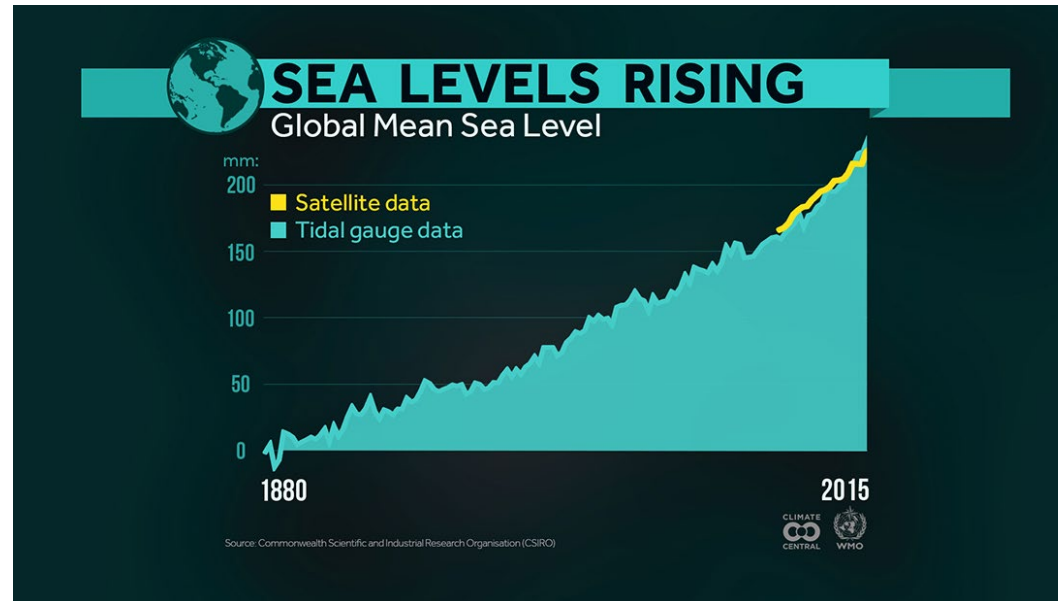
- 2°C is maximum excess carbon budget



## Ocean

Since 1850, **global sea levels rose by 19cm**

- 15cm more by 2031;  
30cm more by 2050
- Sea Walls: \$280 - \$500B projected annual cost



# Climate Change is *our* Grand Challenge



## Atmosphere

From 1880 - 2012, average global temperatures increased by **0.85°C**

- 2°C is maximum excess carbon budget



## Ocean

Since 1850, global sea levels rose by **19cm**

- 15cm more by 2031; 30cm more by 2050
- Sea Walls: \$280 - \$500B projected annual cost



## Cryosphere

Arctic ice sheets are melting

- **1991 - 2001:** 30 Gt/ year
- **2002 - 2011:** 147 Gt/ year



# Climate Change is *our* Grand Challenge



## Atmosphere

From 1880 - 2012, average global temperatures increased by **0.85°C**

- 2°C is maximum excess carbon budget



## Ocean

Since 1850, global sea levels rose by **19cm**

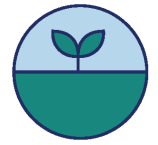
- 15cm more by 2031; 30cm more by 2050
- Sea Walls: \$280 - \$500B projected annual cost



## Cryosphere

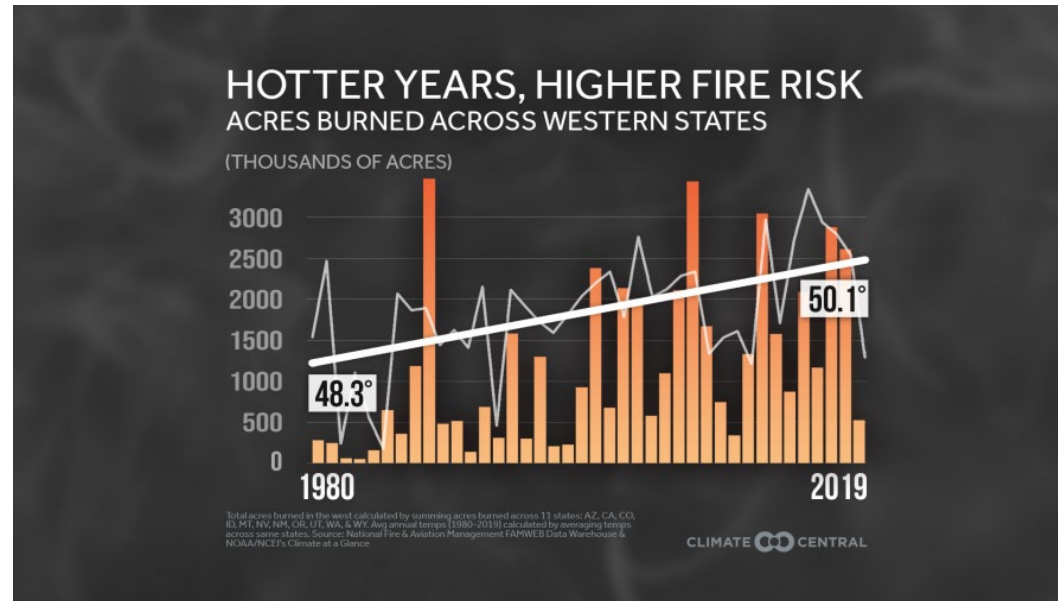
Arctic ice sheets are melting

- 1991 - 2001: 30 Gt/ year
- 2002 - 2011: 147 Gt/ year

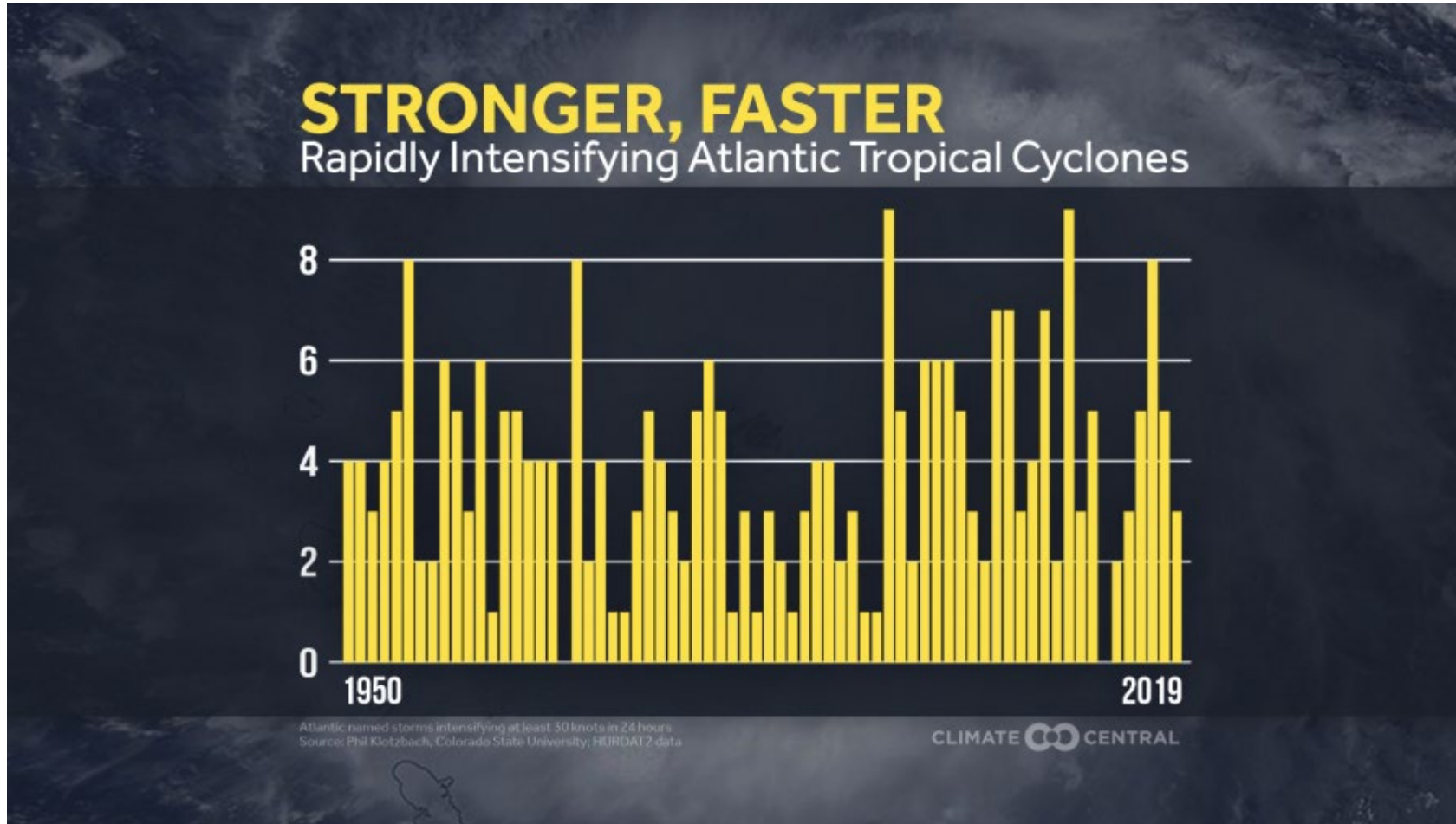


## Land

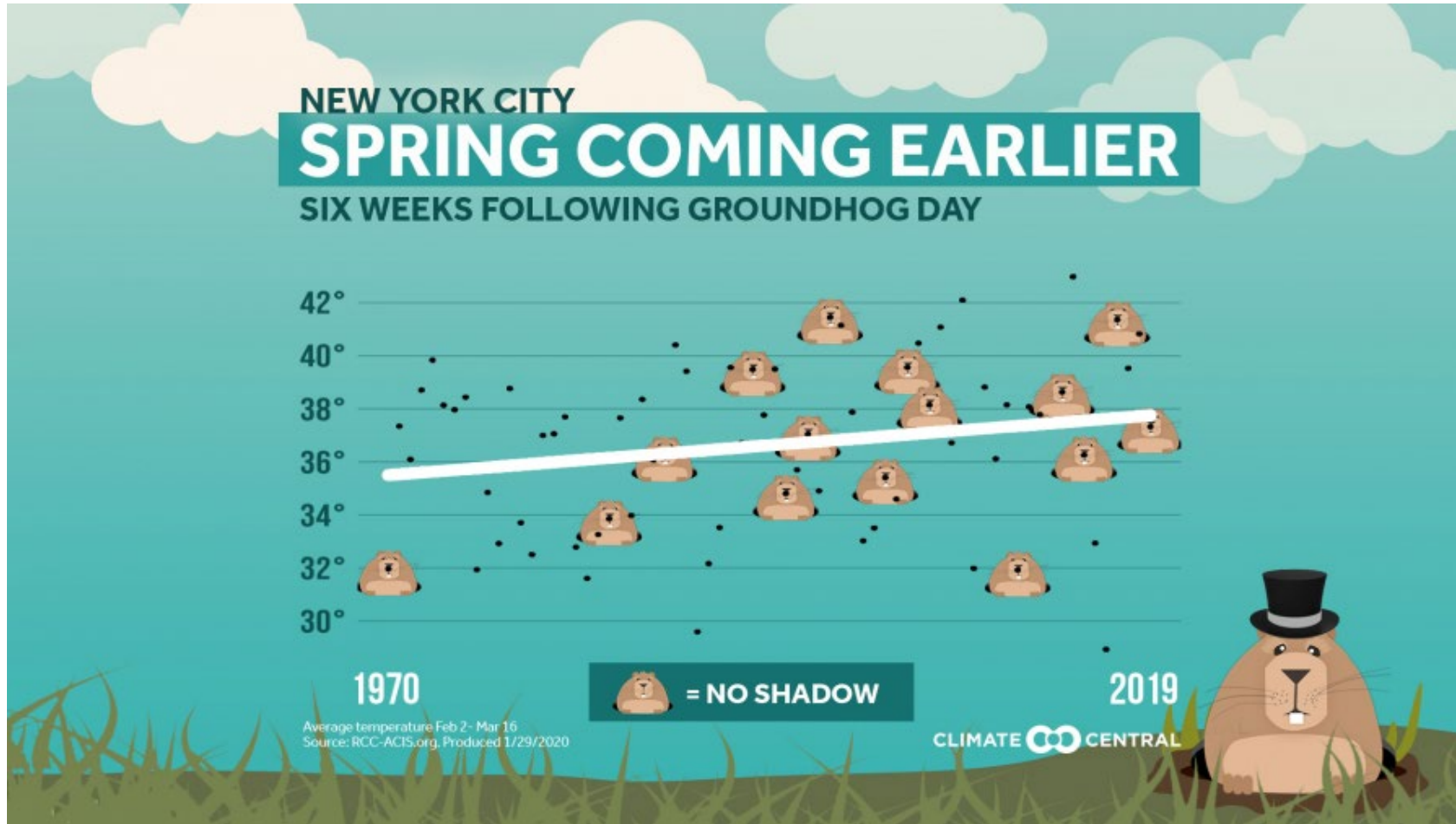
Wildfires + extreme events have increased in frequency



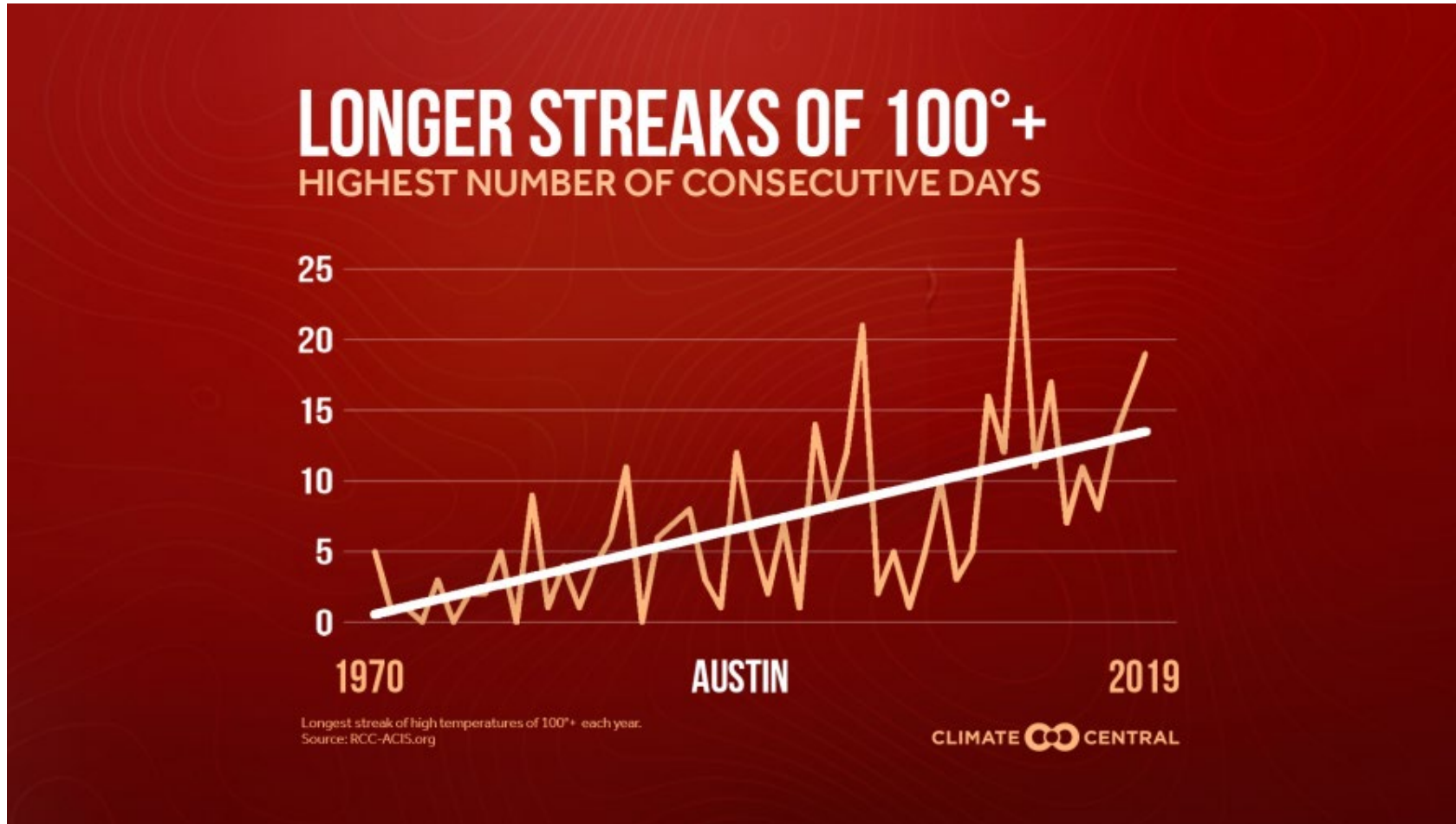
# American Lives are Directly and *Already* Impacted by Climate Change



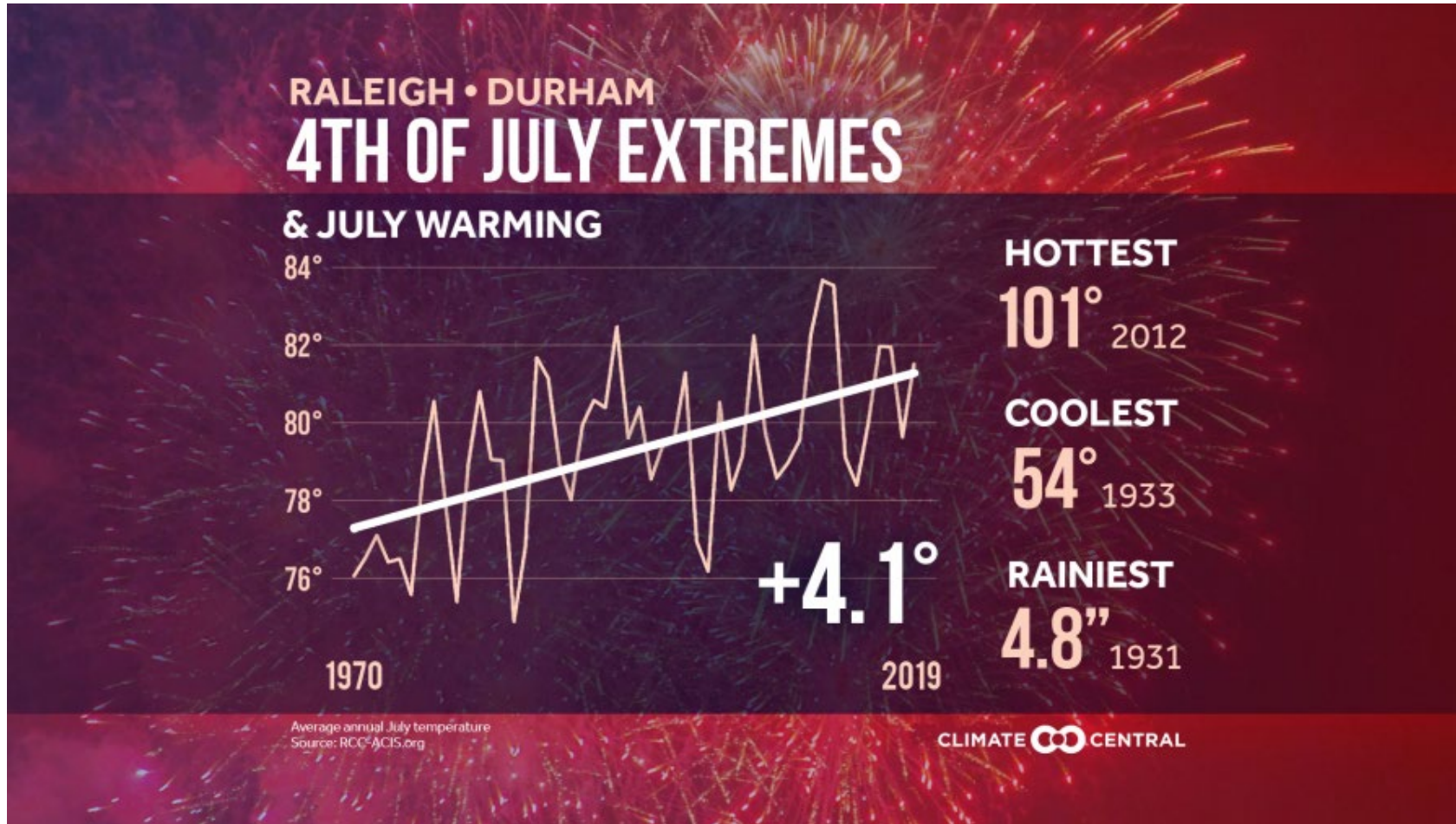
# American Lives are Directly and *Already* Impacted by Climate Change



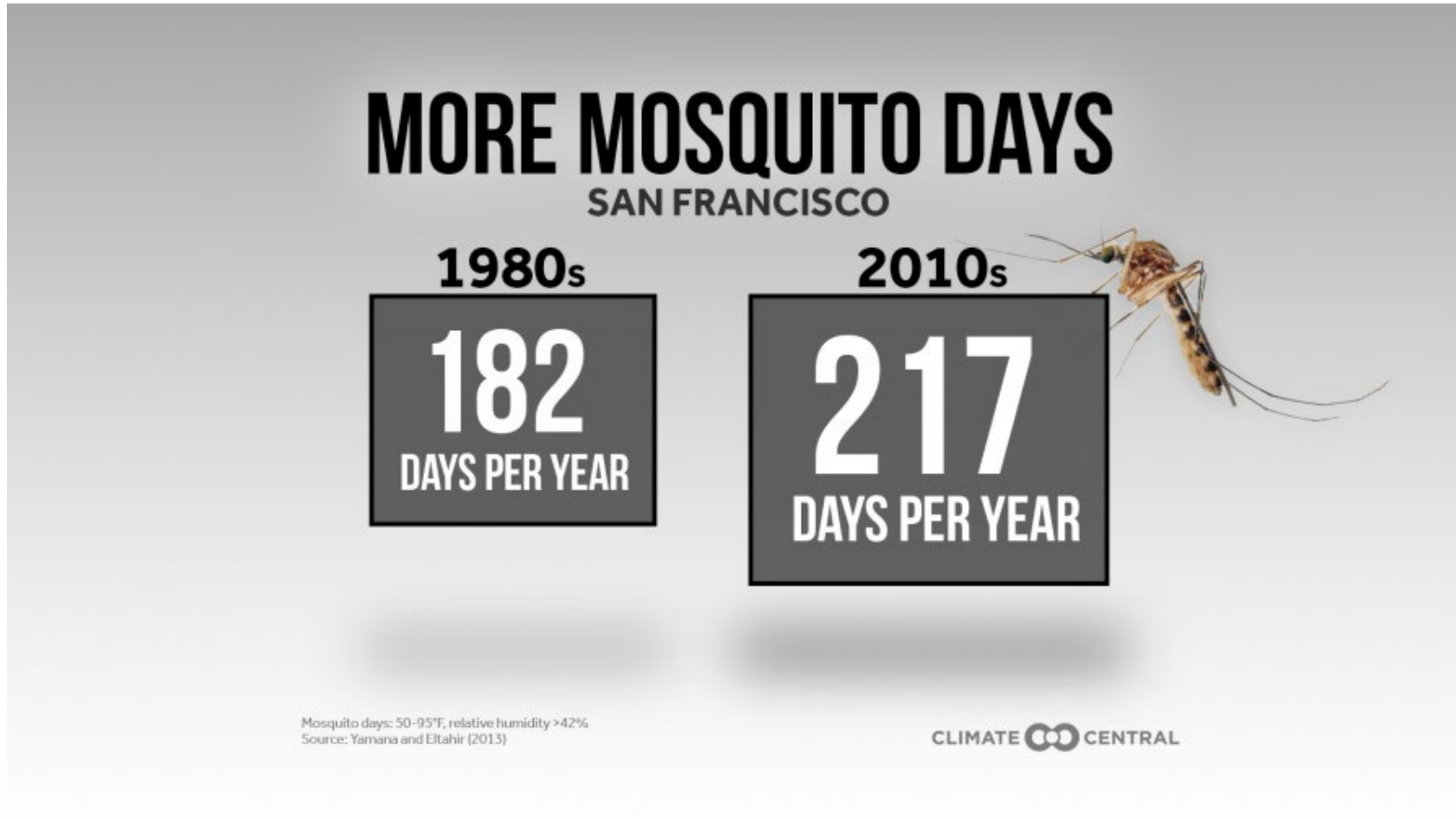
# American Lives are Directly and *Already* Impacted by Climate Change



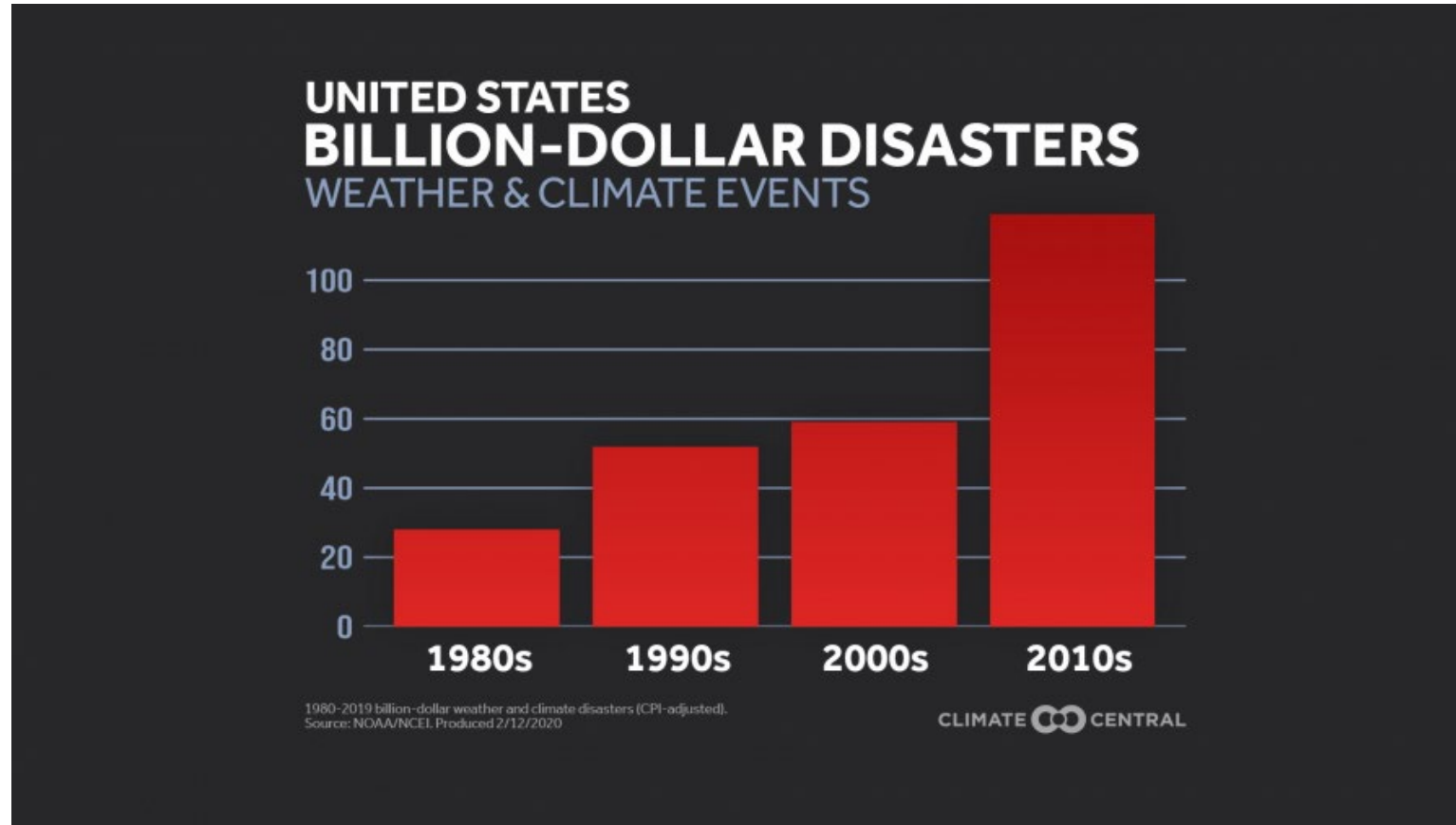
# American Lives are Directly and *Already* Impacted by Climate Change



# American Lives are Directly and *Already* Impacted by Climate Change



# Number of \$1B+ Disasters are Irrevocably Increasing



Reduction of future human + financial costs requires **adaptation**

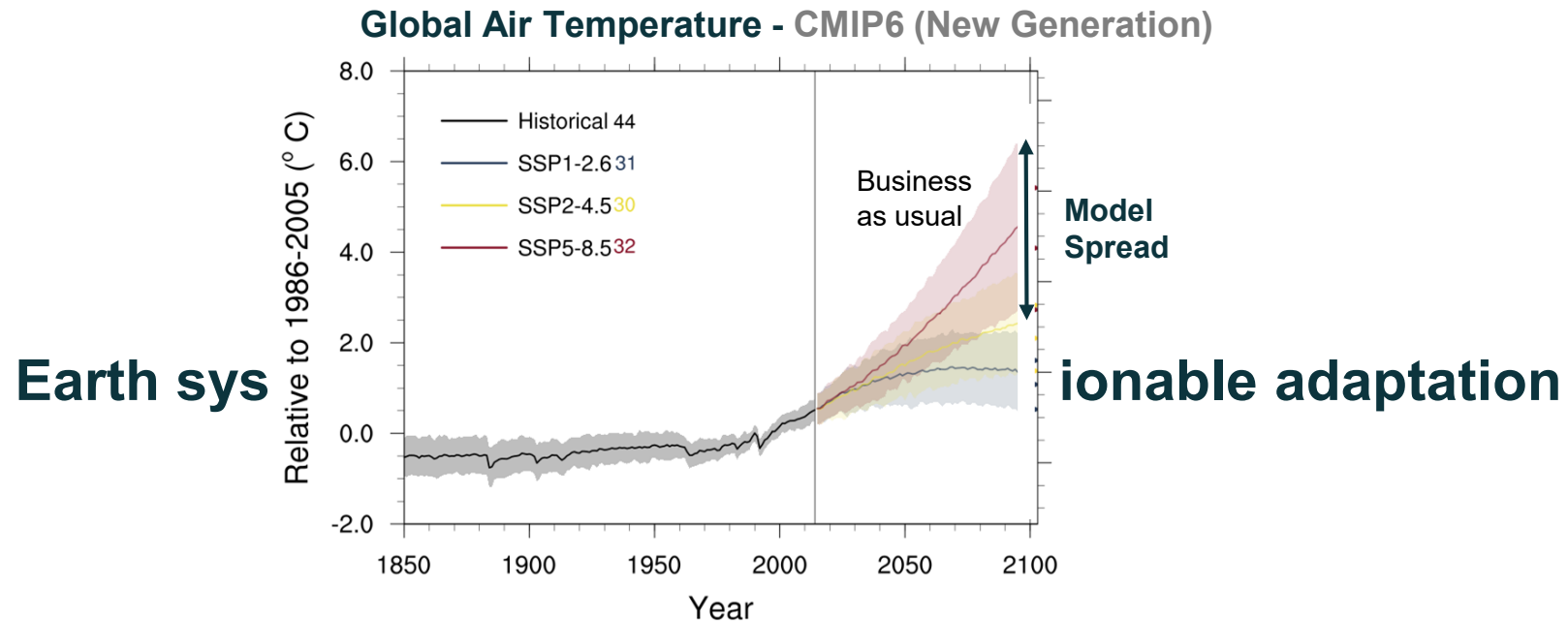


To Promote + Protect:  
National Security  
Economic Prosperity  
Human Health  
Environmental Sustainability

**We must predict + adapt to climate  
change now**

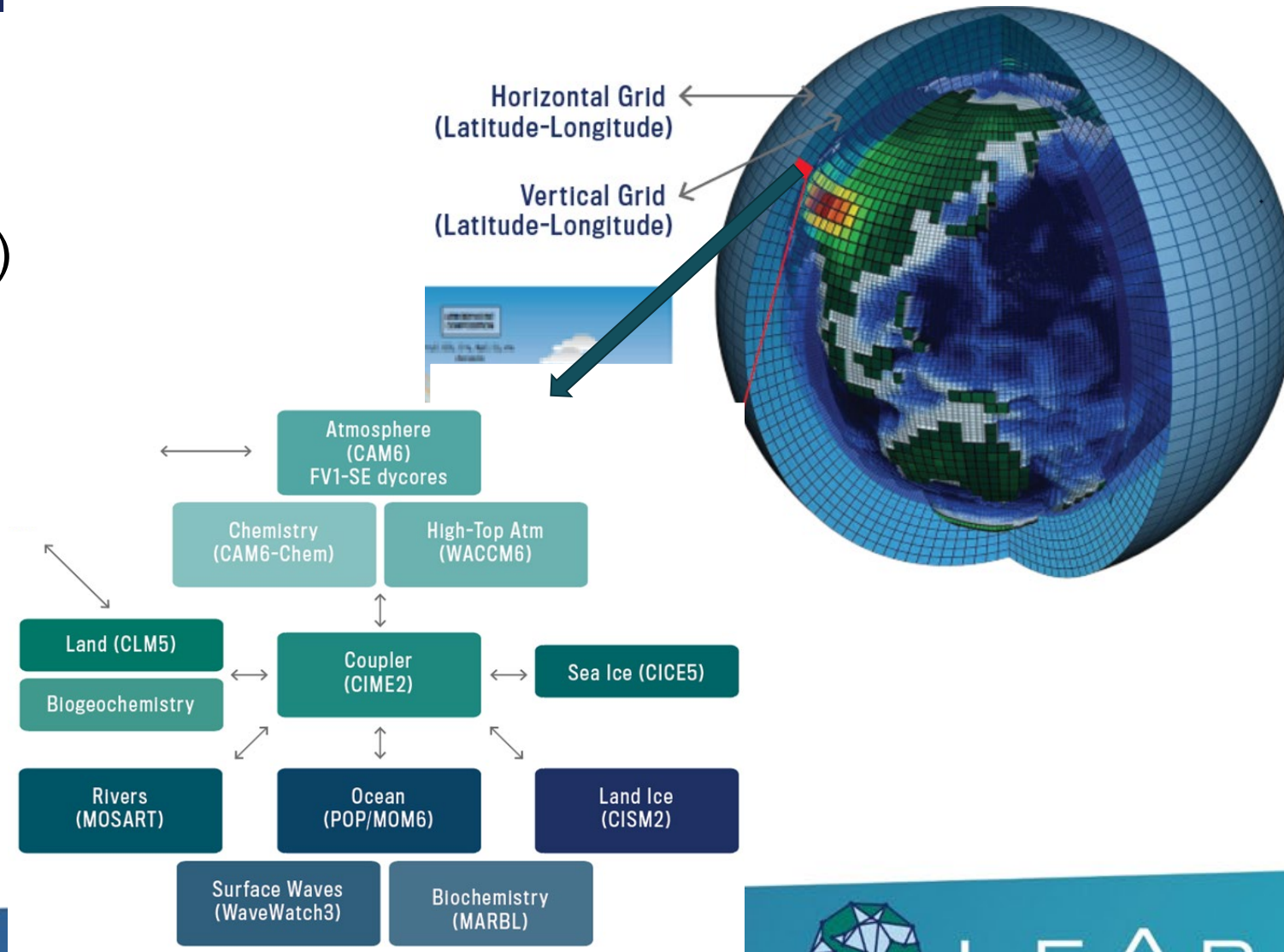
# Adaptation Requires Reliable Climate Projections (Forecasts)

- Projections made with **Earth System Models (ESMs)**
- Current models are *too* uncertain



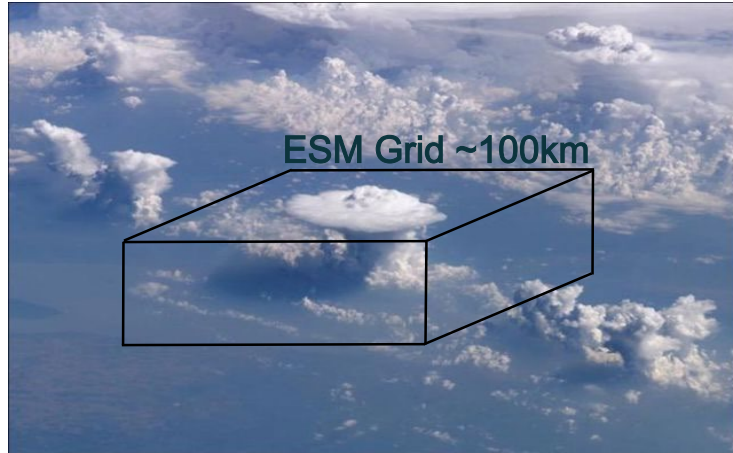
# An Earth System Model is a Complex Coupled System

- Simplified representation in Earth System Models (ESMs)
- 1000's of equations
- Parameters empirically fine-tuned
- Grid resolution ~100km



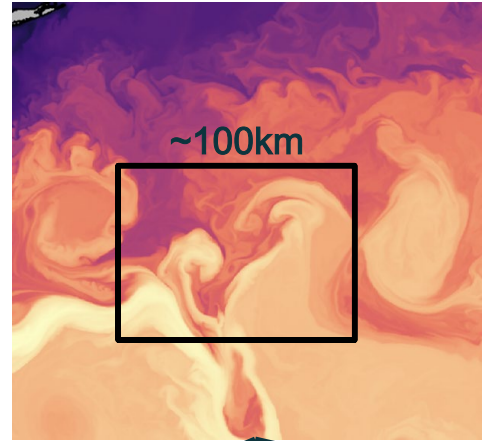
# Unresolved Small + Complex Processes Require "Parameterizations" Causing Projection Uncertainties

Clouds

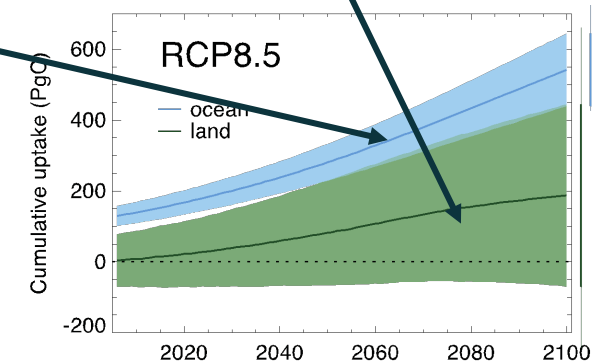
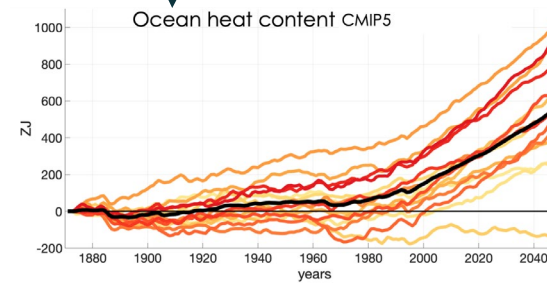
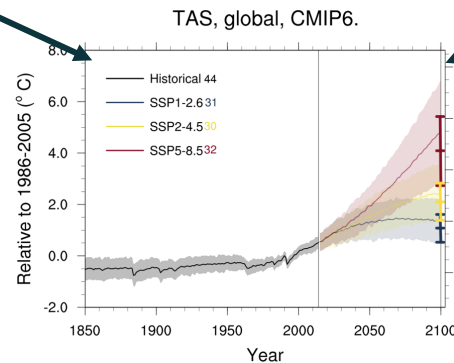
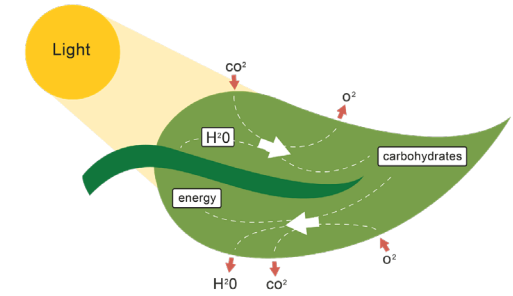


Physical + Biological Processes

Ocean Eddies



Photosynthesis



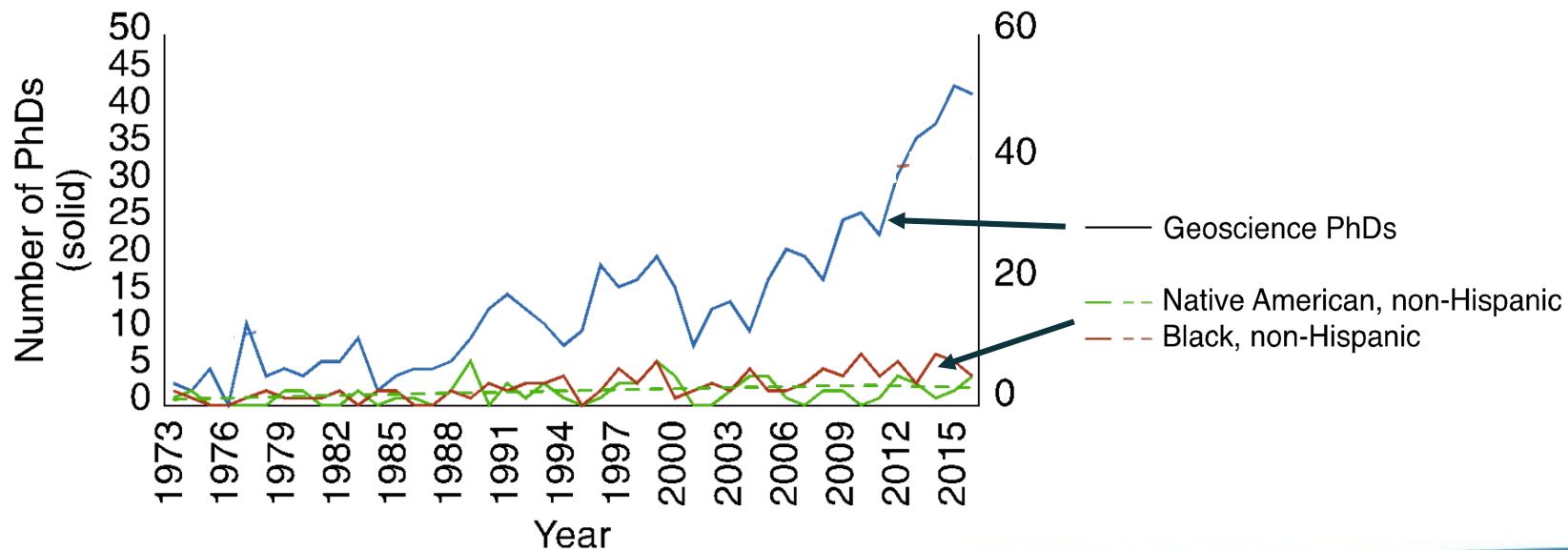
- Model errors dominate (>50%) uncertainties <40 years
- Massive amount of **data** to constrain those processes



# Other Biases: Lack of Diversity Handicaps **Reach + Utility** of Climate Modeling

- Climate science has made progress in gender equity, but is still **not racially/ethnically diverse**
  - Columbia's ADVANCE grant established best practices
- Yet, climate change *disproportionately* impacts underrepresented minorities

ADVANCE  
AT THE EARTH INSTITUTE AT COLUMBIA UNIVERSITY



## The New York Times

### Earth Science Has a Whiteness Problem

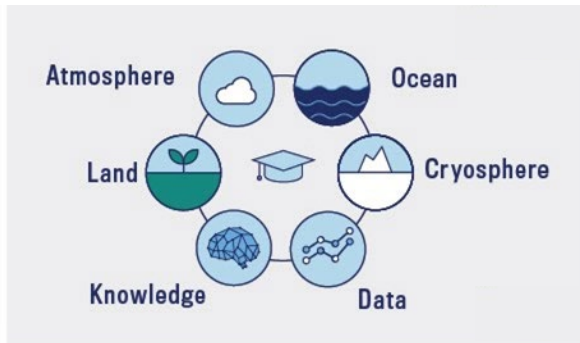
Barely 10 percent of doctoral degrees in the geosciences go to recipients of color. The lack of diversity limits the quality of research, many scientists say.



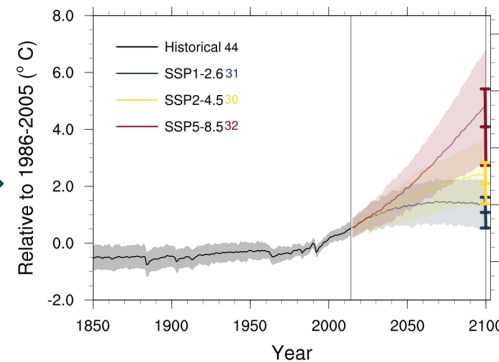
# Other Biases: Climate Communications are One-Way

- Data is not formatted/ available for accessibility + actionable knowledge
- Researchers have limited understanding of public + private stakeholders' concerns
- Decision-makers need dialogue + actionable information to adapt

## Climate Data Science Research



## Imprecise Projections



## Stakeholders

Sustainability-oriented startups

Press media

Sustainability-oriented investment companies

Independent journalists

Traditional investment companies

Community groups

Technology-oriented companies

Non-governmental organizations

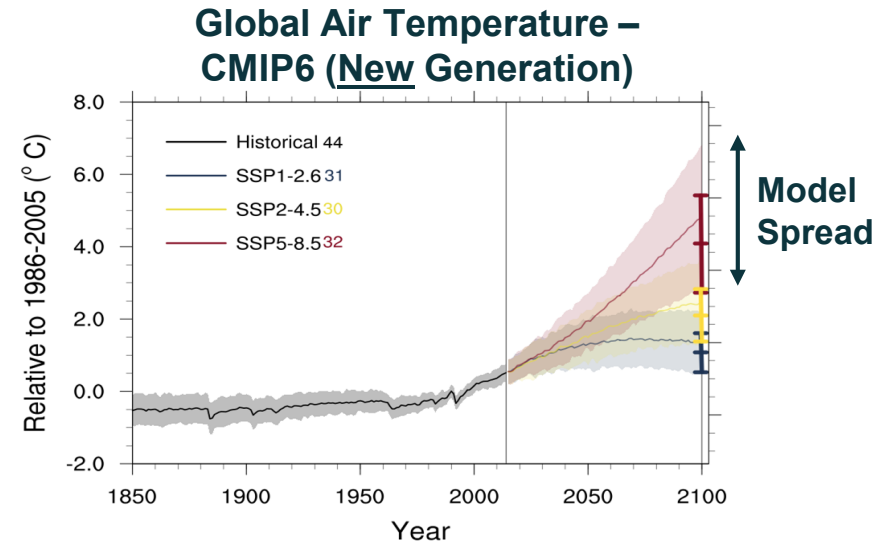
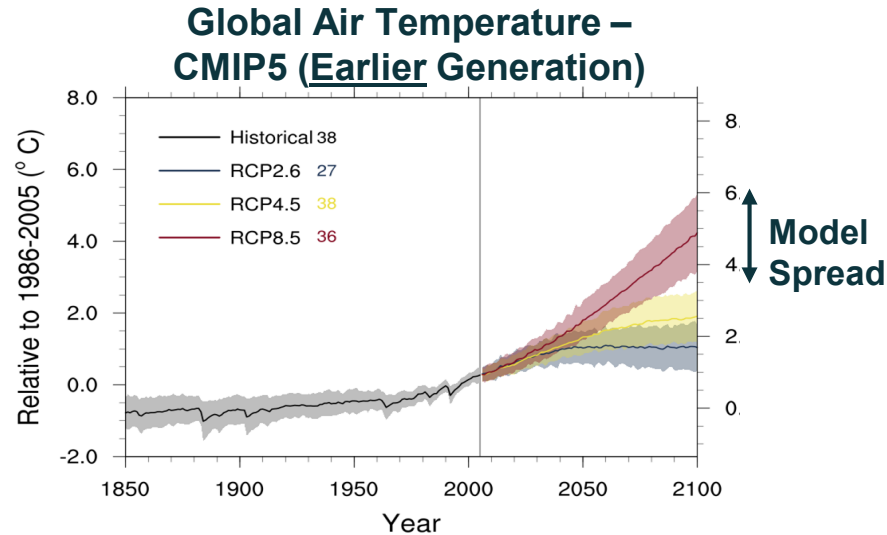
Traditional companies

Individual and diverse people



# Time is Running Out to Adapt

- ESM parameterization progress is *too slow*
  - On par with climate change



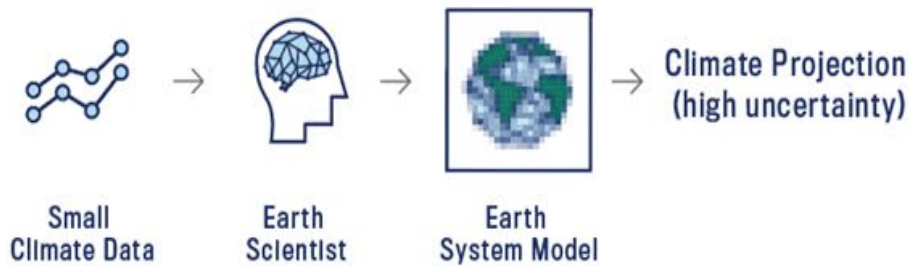
- Diversity + inclusion + communication progress is equally slow

**Reliable climate models need an inclusive STEM workforce + broad stakeholder engagement for national progress**



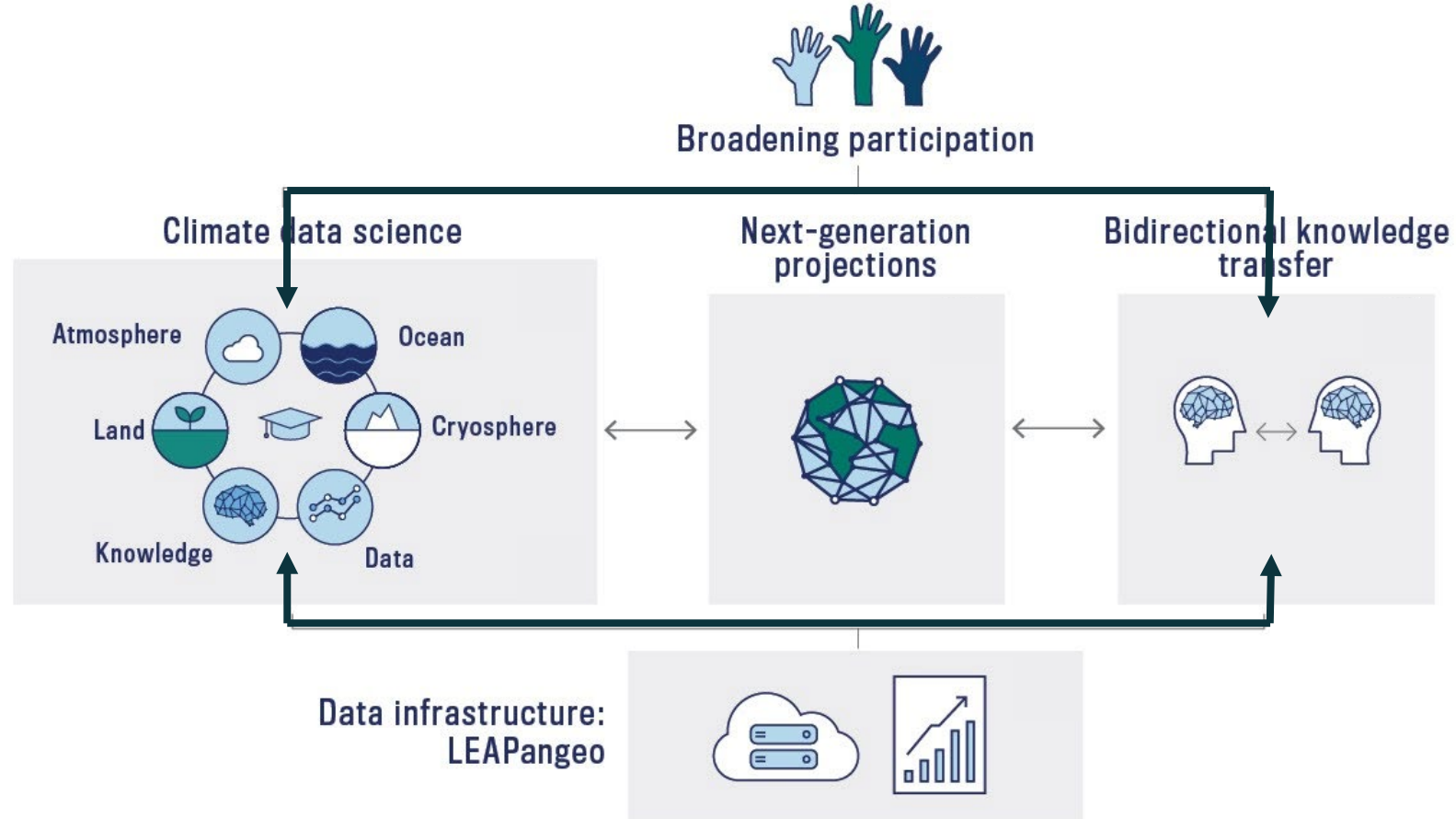
# LEAP Tackles these Challenges **Now**

## Learning the Earth with Artificial Intelligence + Physics



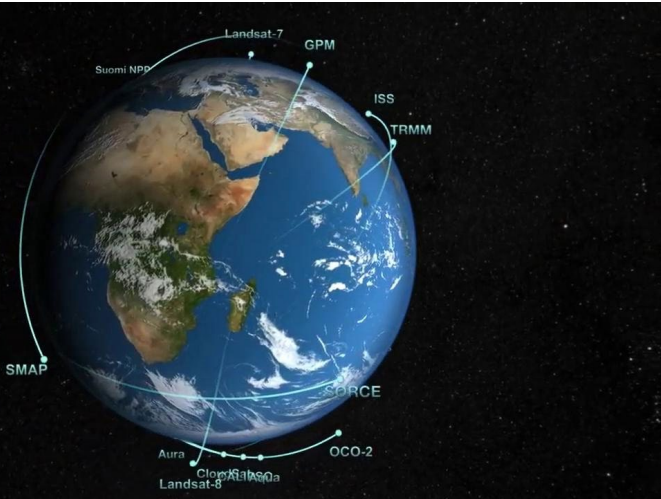
EARTH SYSTEM MODELING TODAY

# Vision: LEAP Forward in the Reliability, Utility + Reach of Climate Projections through Synergistic Innovations in Data Science + Climate Science



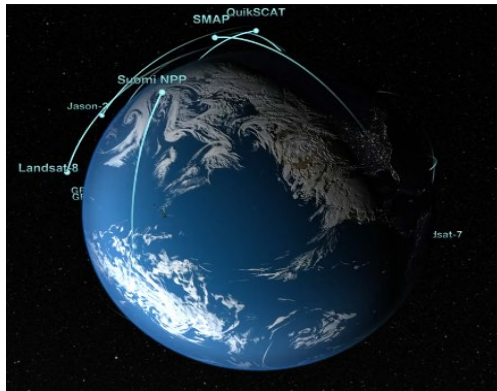
# Today is a Unique Moment of Opportunity

## Massive Data from Earth Observations

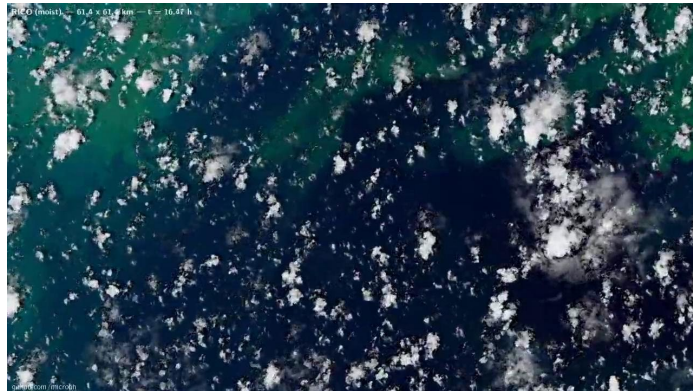


# Today is a Unique Moment of Opportunity

Massive Data  
from Earth  
Observations

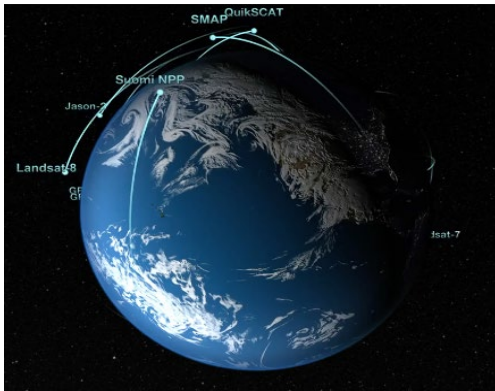


High-Resolution  
Process-Resolving  
Simulations

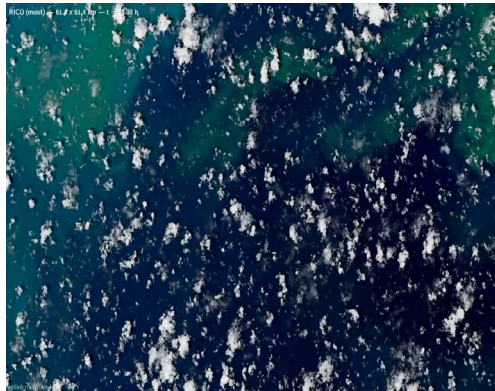


# Today is a Unique Moment of Opportunity

Massive Data  
from Earth  
Observations



High-Resolution  
Process-Resolving  
Simulations

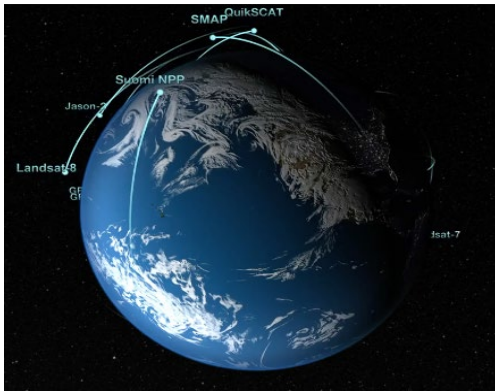


Progress  
in Machine  
Learning

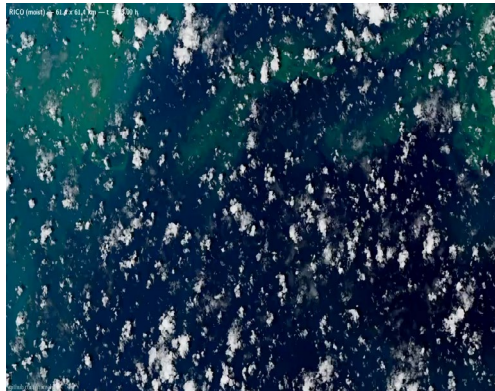


# Today is a Unique Moment of Opportunity

Massive Data  
from Earth  
Observations



High-Resolution  
Process-Resolving  
Simulations



Progress  
in Machine  
Learning

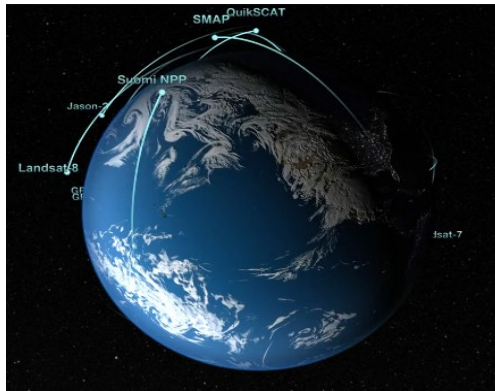


Public  
Debate on  
Climate  
Change +  
Diversity

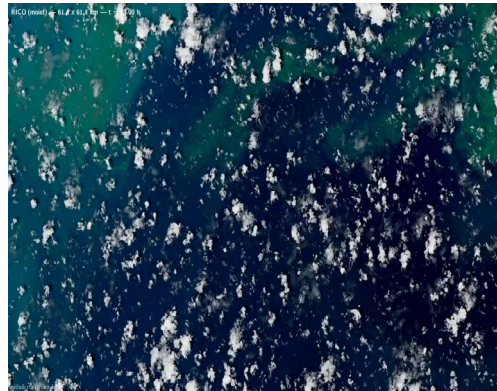


# Today is a Unique Moment of Opportunity

Massive Data  
from Earth  
Observations



High-Resolution  
Process-Resolving  
Simulations



Progress  
in Machine  
Learning



Public  
Debate on  
Climate  
Change +  
Diversity



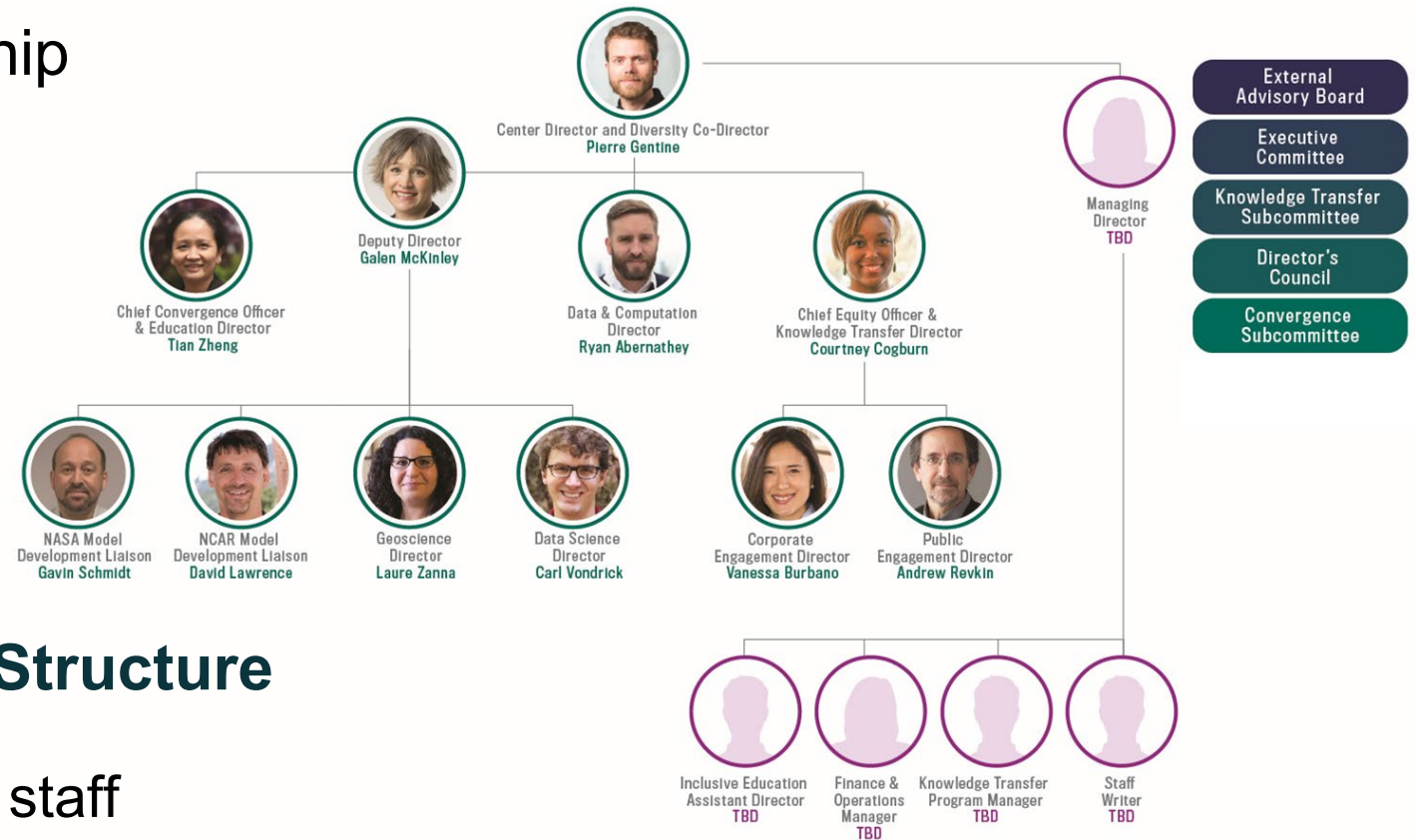
Private  
Sector is  
New Ally for  
Change



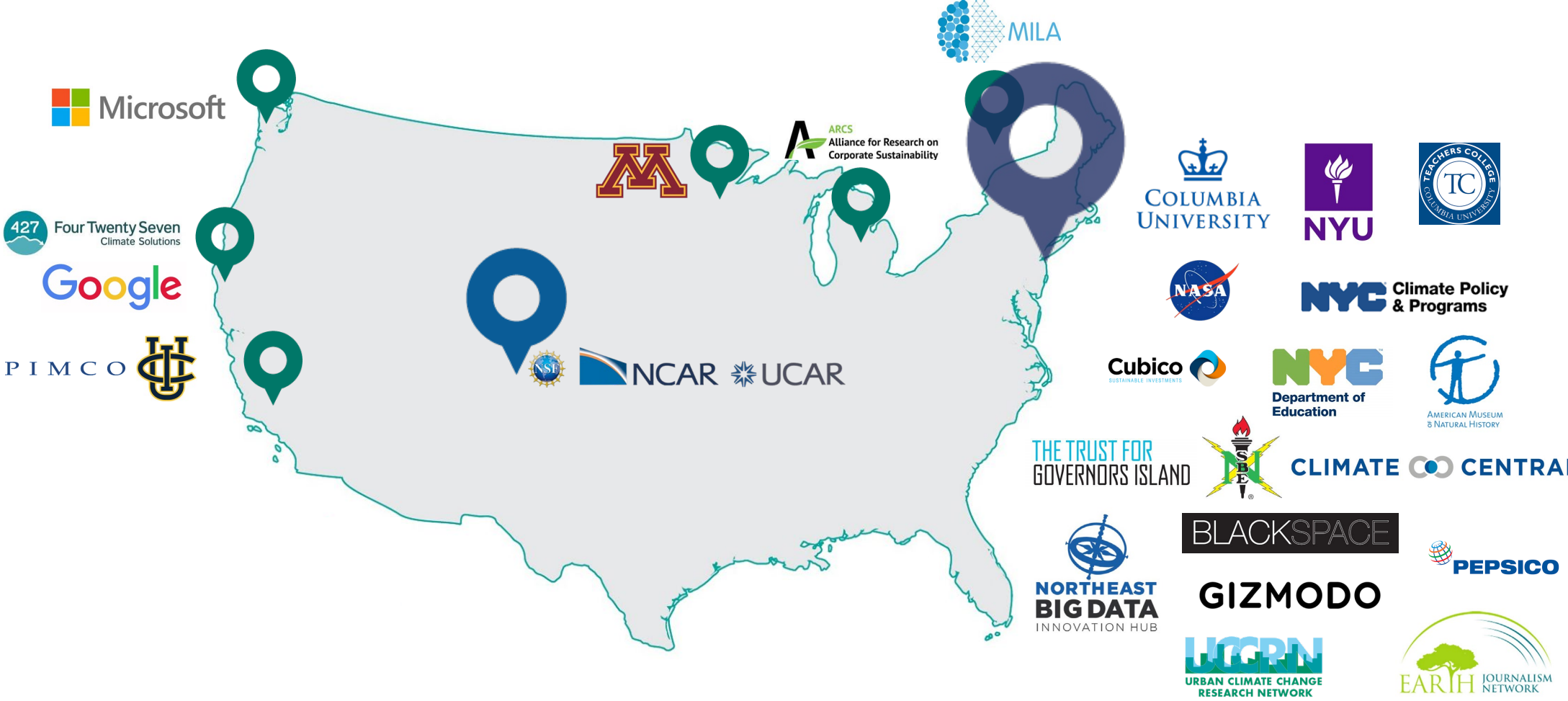
LEAP converges + elevates these opportunities to achieve its vision.

# LEAP's Vision is Achieved Through Convergence

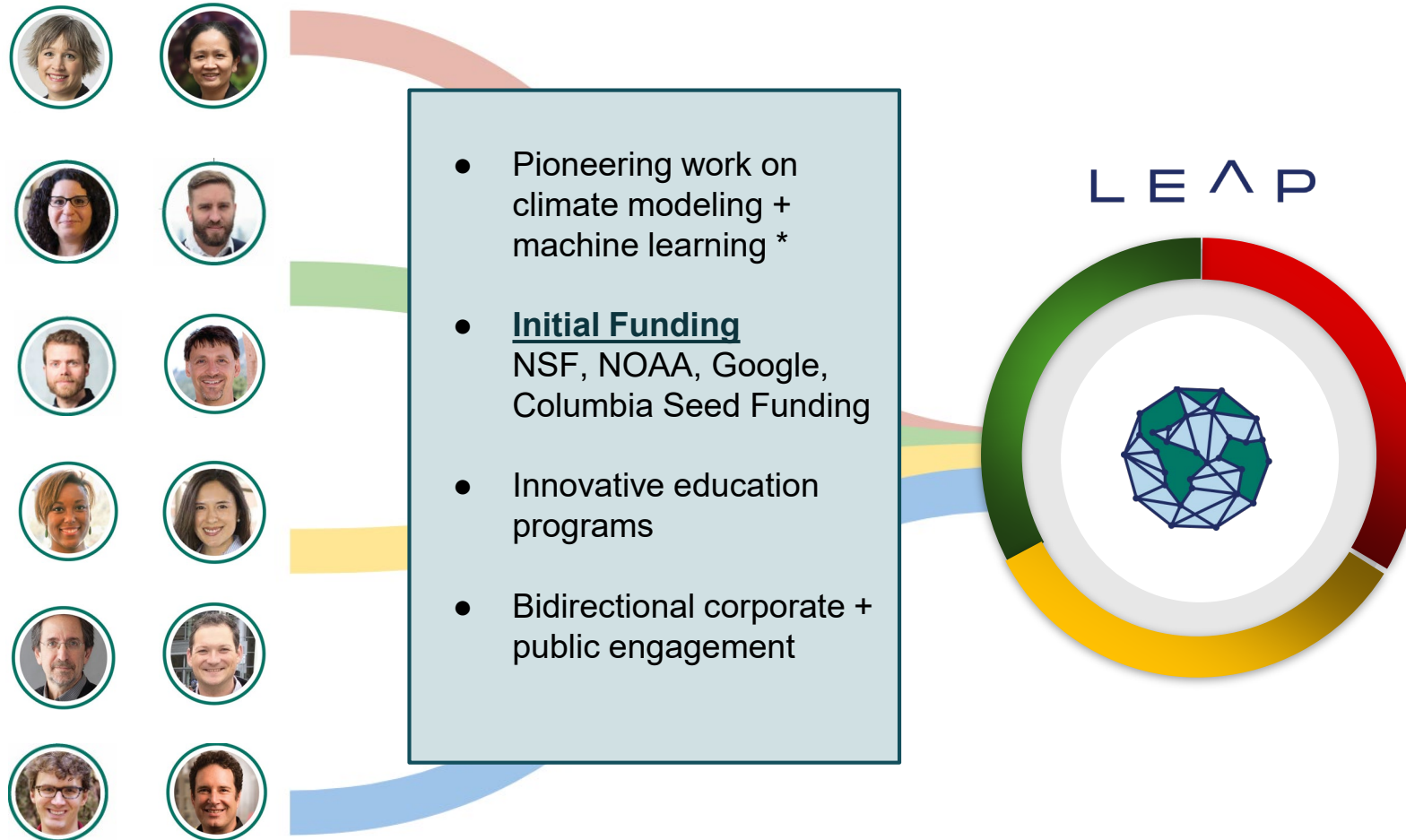
- Diverse + cohesive leadership
- **Highly-convergent team**
  - 28 climate scientists
  - 14 data scientists
  - 6 social scientists
  - 5 administrative leaders
- 5 university partners
- 20+ non university partners
- **Transparent Governance Structure**
  - 5 governance committees
  - 5 dedicated administrative staff



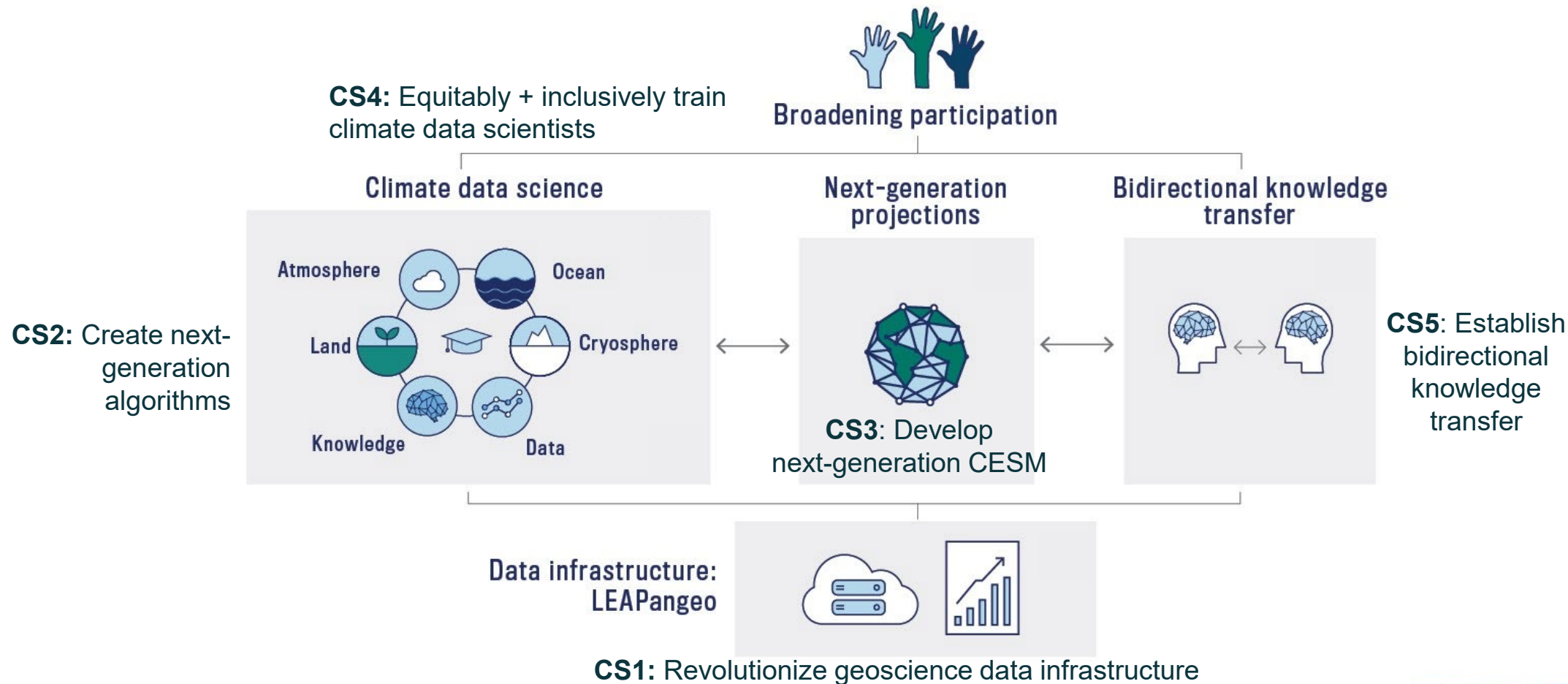
# LEAP Spans the Continent, Hubbed in NYC



# LEAP Builds upon Innovative + Convergent Collaborations



# Five Convergence Strategies (CS) Drive LEAP's Vision



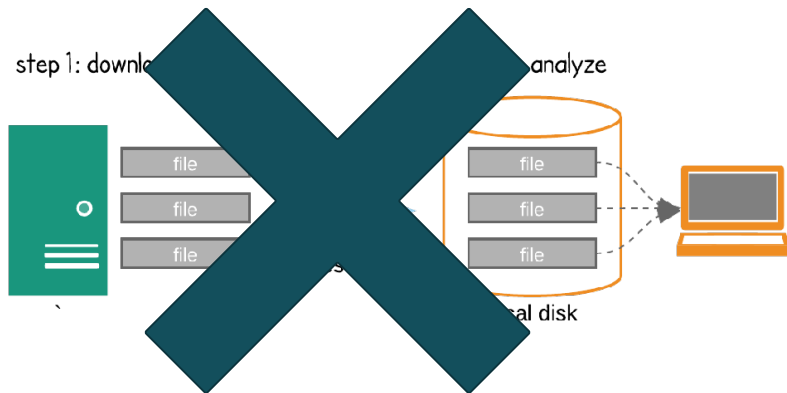
# CS1: Revolutionize Geoscience Data Infrastructure

- **Challenge:** Massive + heterogeneous datasets
- **Solution:** LEAPangeo: open-cloud computing platform

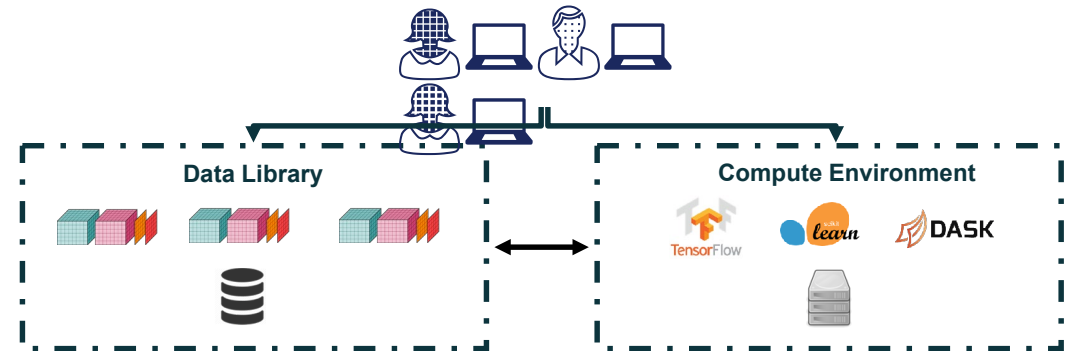


Data & Computation  
Director  
Ryan Abernathey

## Download Model

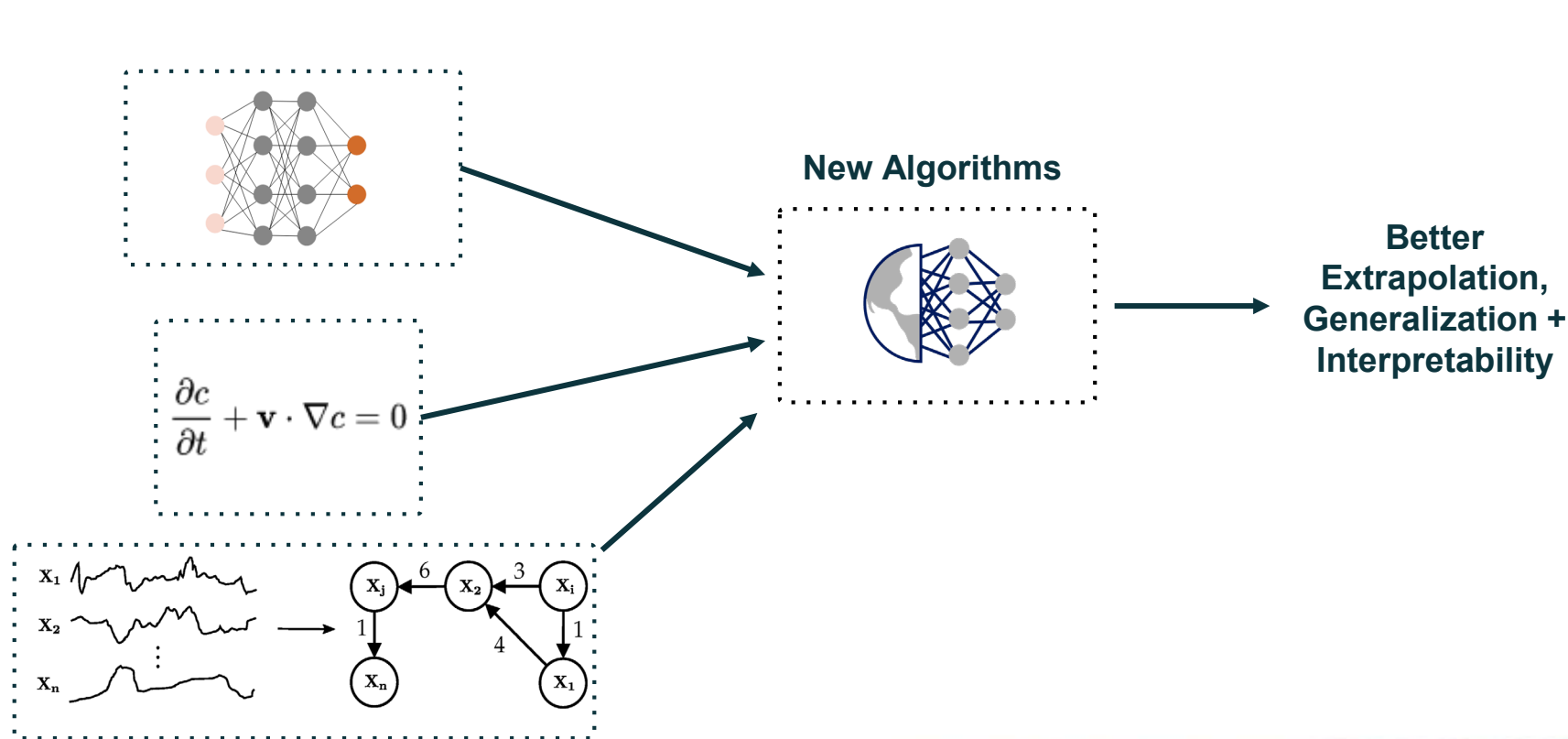


## Data-Proximate Computing



# CS2: Create Next-Generation Algorithms

- **Challenge:** ML algorithms need to respect physics + better extrapolate + be interpretable
- **Solution:** Include **physical + causal knowledge** to robustly extrapolate + understand



Deputy Director  
Galen McKinley



Chief Convergence Officer  
& Education Director  
Tian Zheng



Data Science  
Director  
Carl Vondrick

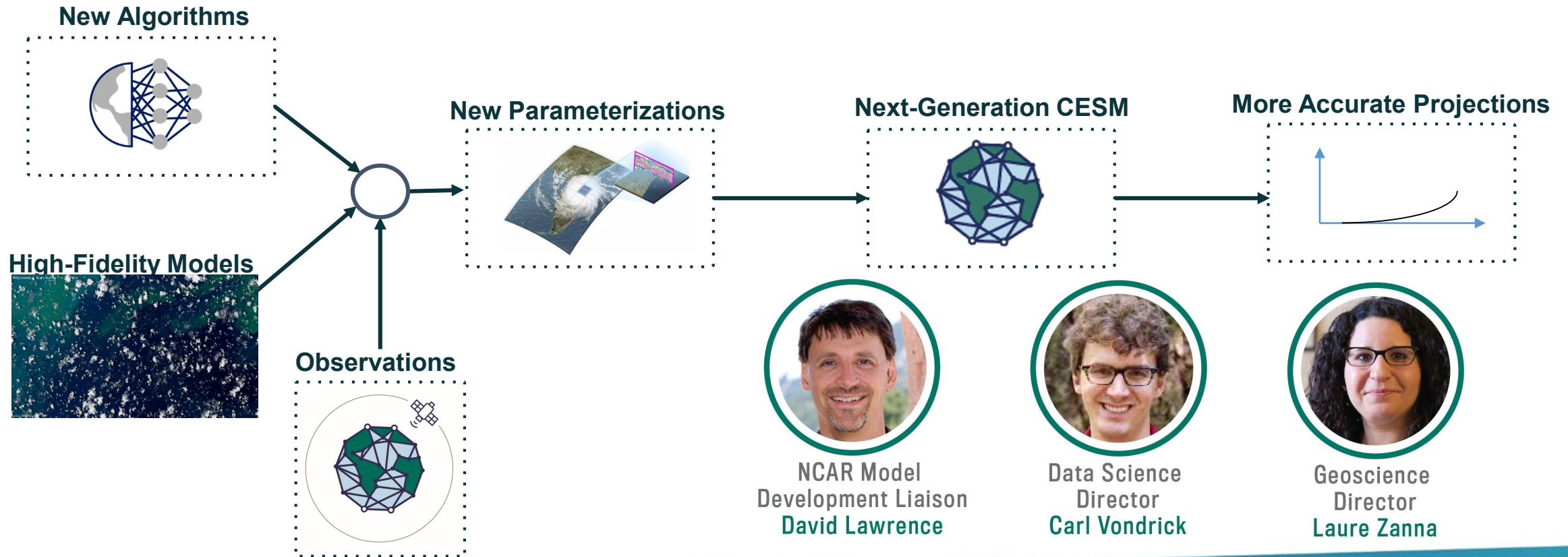


Geoscience  
Director  
Laure Zanna



# CS3: Develop Next-Generation CESM

- **Challenge:** Create high-reliability climate projections
- **Solution:** Harness new ML and data to transform the open NSF-funded Community Earth System Model (CESM)



# CS4: Equitably + Inclusively Train Climate Data Scientists

- **Challenge:** Distinct fields + vocabularies limit collaborations + progress



# CS4: Equitably + Inclusively Train Climate Data Scientists

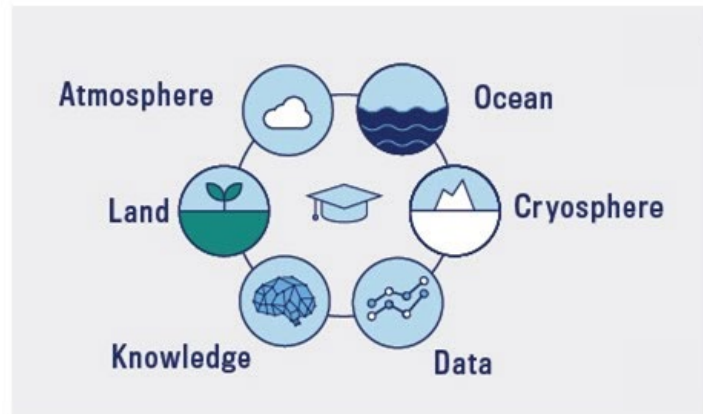
- **Challenge:** Distinct fields + vocabularies limit collaborations + progress
- **Solution:** Forge the new **climate data science** discipline via convergence between education + research
  - Expand on major ADVANCE lessons

ADVANCE  
AT THE EARTH INSTITUTE AT COLUMBIA UNIVERSITY



Broadening participation

Climate data science



Chief Convergence Officer  
& Education Director  
**Tian Zheng**

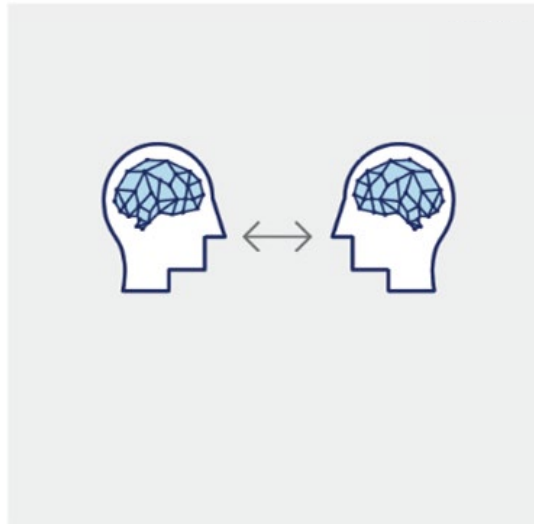


Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



# CS5: Establish Bidirectional Knowledge Transfer

- **Challenge:** Better climate modeling will not increase climate adaption alone
- **Solution:** Bidirectional communication across academia, modeling centers + diverse stakeholders
  - Advance climate modeling **utility, reach + relevance**



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



Corporate  
Engagement Director  
**Vanessa Burbano**



Public  
Engagement Director  
**Andrew Revkin**



# An STC is Necessary to Achieve LEAP's Vision

## **Innovative + Transformative**

- Creating reliable climate adaptation requires:
  1. Convergent research + education
    - Defining new discipline
  1. Equitable workforce + stakeholder engagement
  1. Bidirectional knowledge transfer

## **Complex + Large-Scale**

- LEAP subcomponents are interdependent + interconnected
  - Requires synergies impossible for smaller grants

## **Long-Term**

- LEAP needs national visibility + cohesive impact



# LEAP Addresses National Priorities



EXECUTIVE OFFICE OF THE PRESIDENT  
WASHINGTON, D.C.



August 14, 2020

- “Beyond healthcare, **use-inspired AI research** ... can help the Nation address future crises.
- “Agencies should prioritize, coordinate + collaborate to... **[improve] the theoretical understanding and practical utilization of [climate] predictability**... + [explore] advanced modeling capabilities using non-traditional approaches such as AI.”



Harnessing the Data Revolution

Midscale Research Infrastructure

Future of Work

Navigating the New Arctic

Growing Convergence Research

INCLUDES



# LEAP's 5 - 10 Year Legacy will be Significant

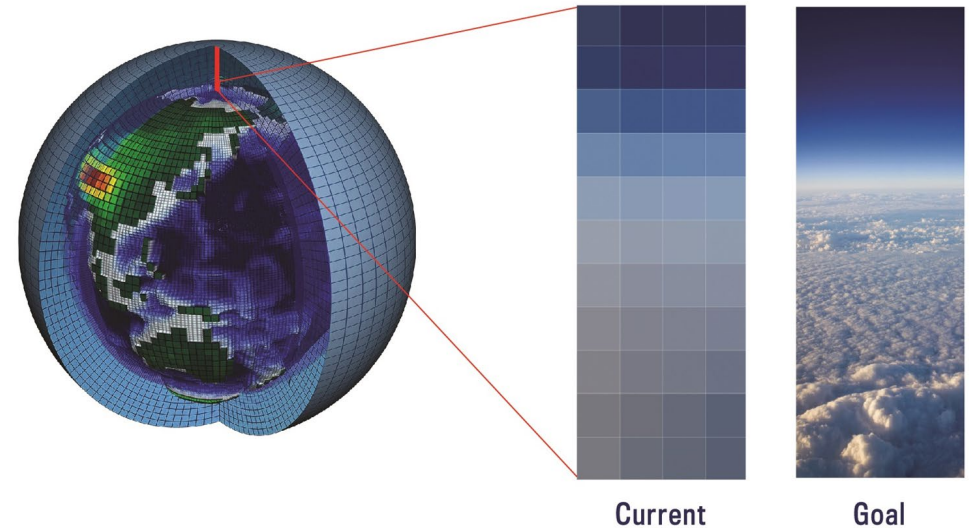
- Develop **next-generation climate models** with novel parameterizations
- Train new generation of **climate data scientists** to bridge academia + society
- Forge a significantly more **diverse cohorts of scientists**
- Establish **community + dialogue with heterogeneous stakeholders** around actionable climate model metrics

# Vision: LEAP Forward in the Reliability, Utility + Reach of Climate Projections through Synergistic Innovations in Data Science + Climate Science

For climate adaptations, we pledge a LEAP in:

- **Reliability**, by developing next-generation climate projections with data at its core.
- **Utility**, by creating a platform for two-way interactions with the broader community.
- **Reach**, by diversifying pipeline + reach, while training more diverse climate data scientists.

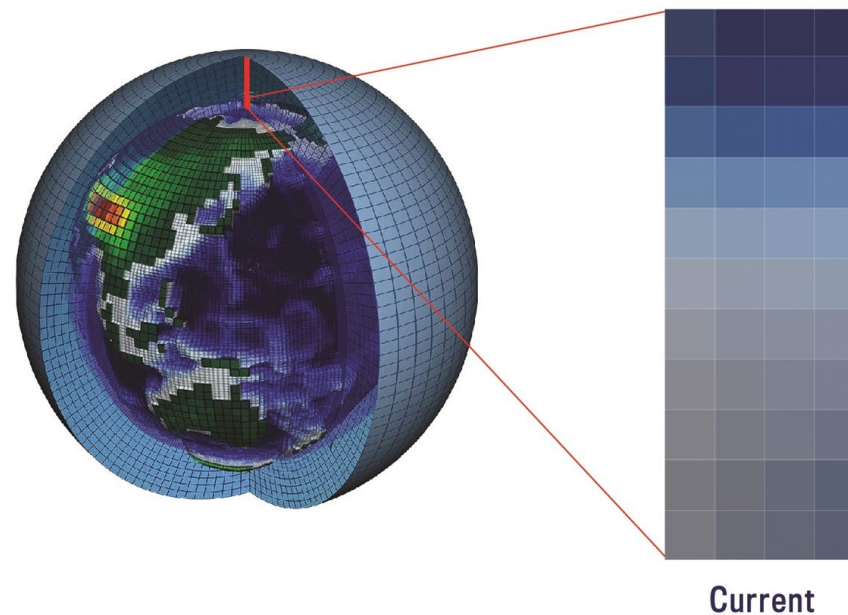
Achieved through convergence between  
data science + climate science




Questions?

# LEAP Can Break the Climate Reliability Deadlock

1. Develop next-generation **parametrizations** learnt from data + deeper understanding
2. Define better **metrics** to systematically evaluate + refine ESMs
3. LEAP forward in the **reliability, utility + reach** of climate projections





# Research Overview



Deputy Director  
**Galen McKinley**

## Professional Experience and Education

- Professor of Earth + Environmental Sciences, Columbia
  - Previously Professor of Atmospheric + Oceanic Sciences, Wisconsin
- PhD, Climate Physics + Chemistry, MIT

## Expertise

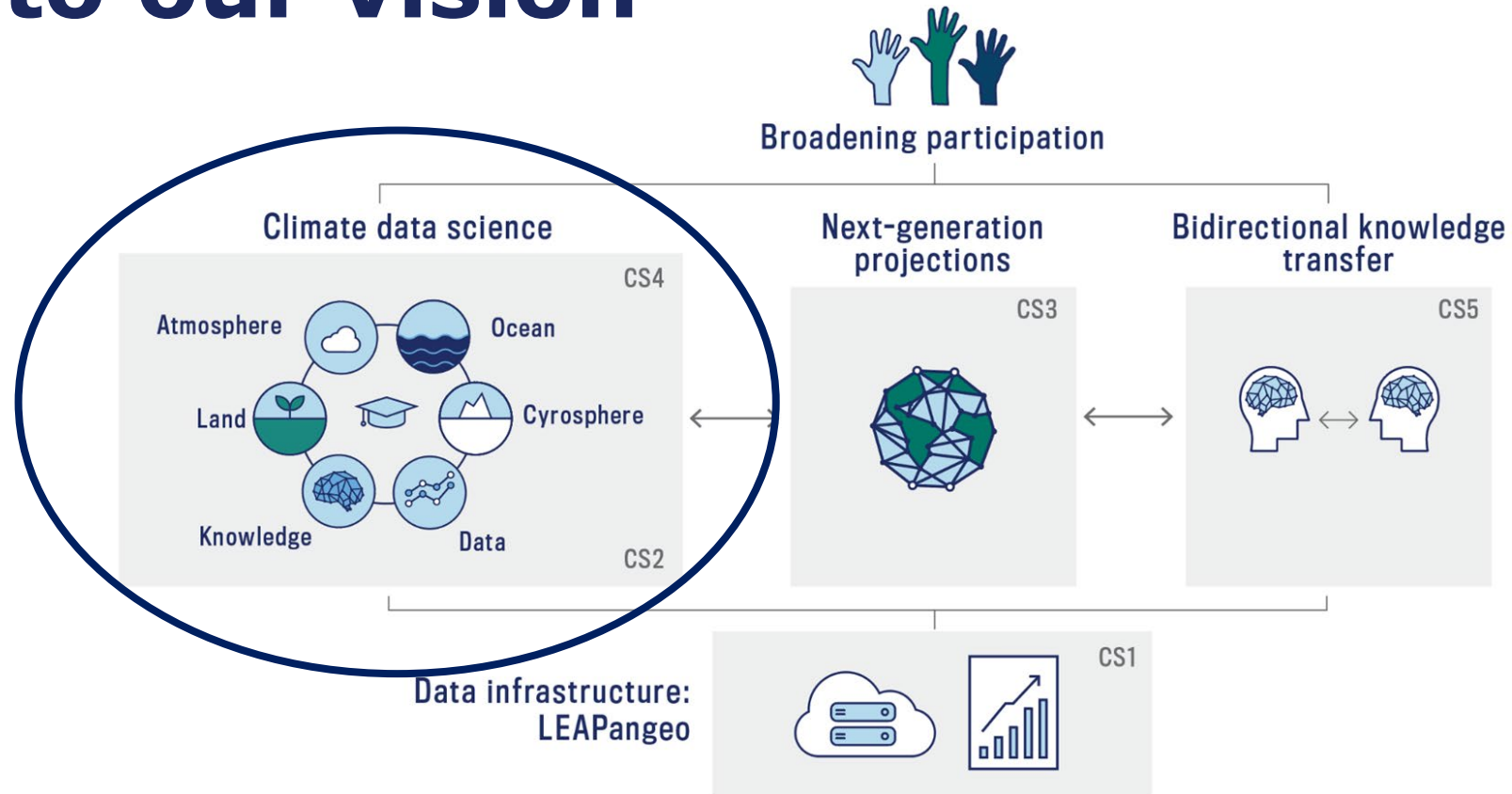
- Ocean carbon cycle, feedbacks with climate
- Ocean + climate modeling
- Data science

## Synergistic Activities

- The Oceanography Society, Councilor, 2020-present
- JASON Advisory Group, 2015-present
- Diversity Committee chair, EES department, 2019-present
- Co-chair for undergrad education, Wisconsin AOS, 2010-15

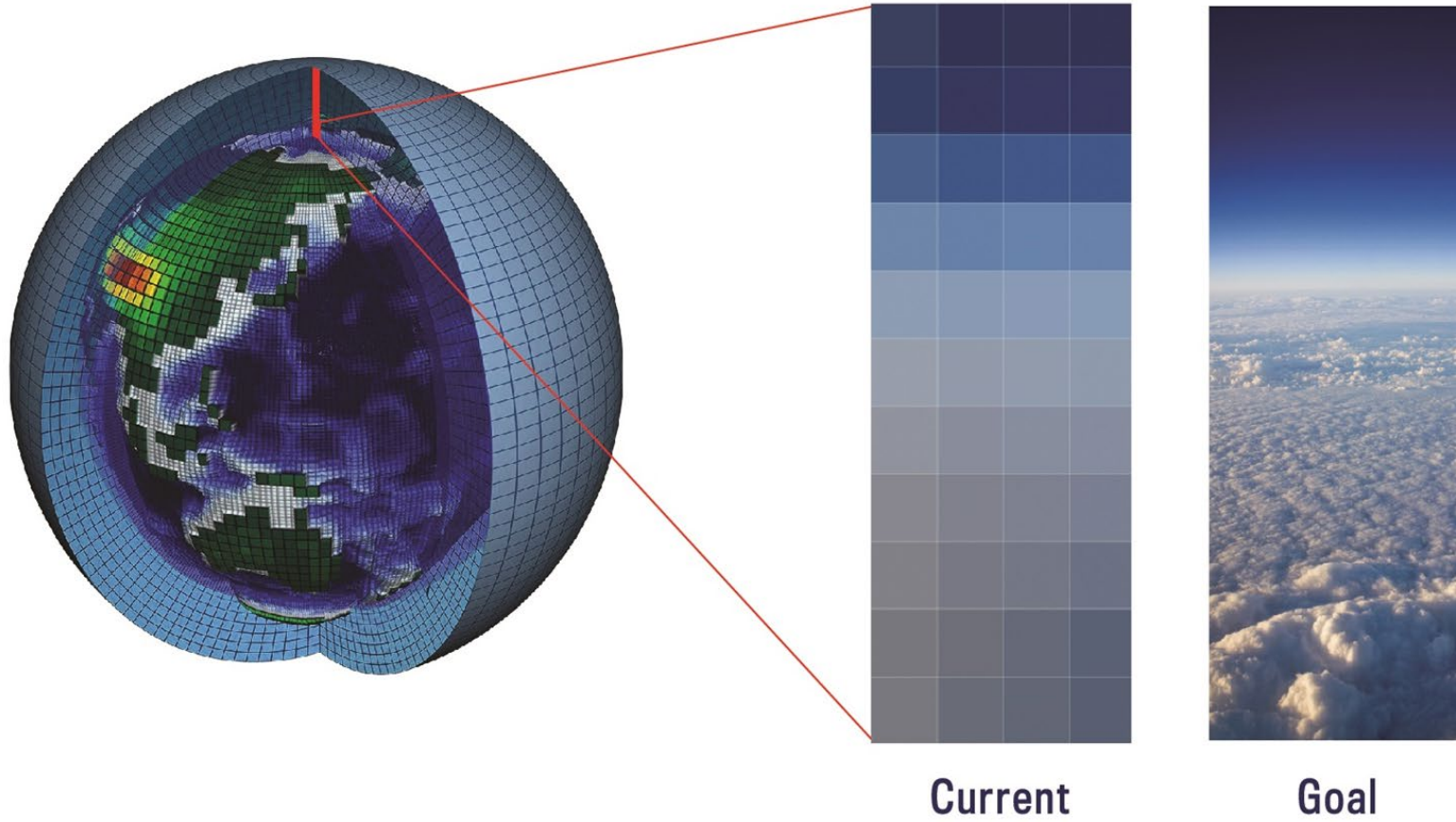


# Structured research in climate data science is key to our vision

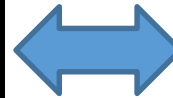
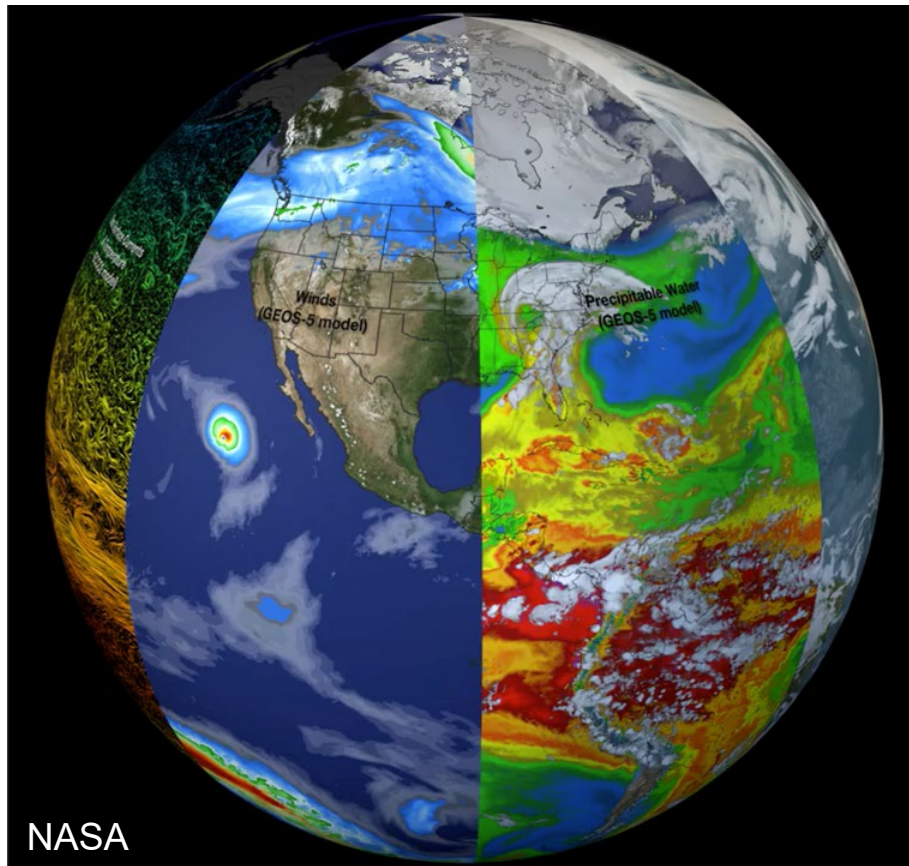


*LEAP forward in the **reliability**, utility, and reach of climate projections through **synergistic innovations** in data science and climate science.*

# LEAP research objective #1: **Next-generation parametrizations built on data**



# LEAP research objective #2: Metrics: enhanced datasets and improved approaches to systematically evaluate ESMs against data



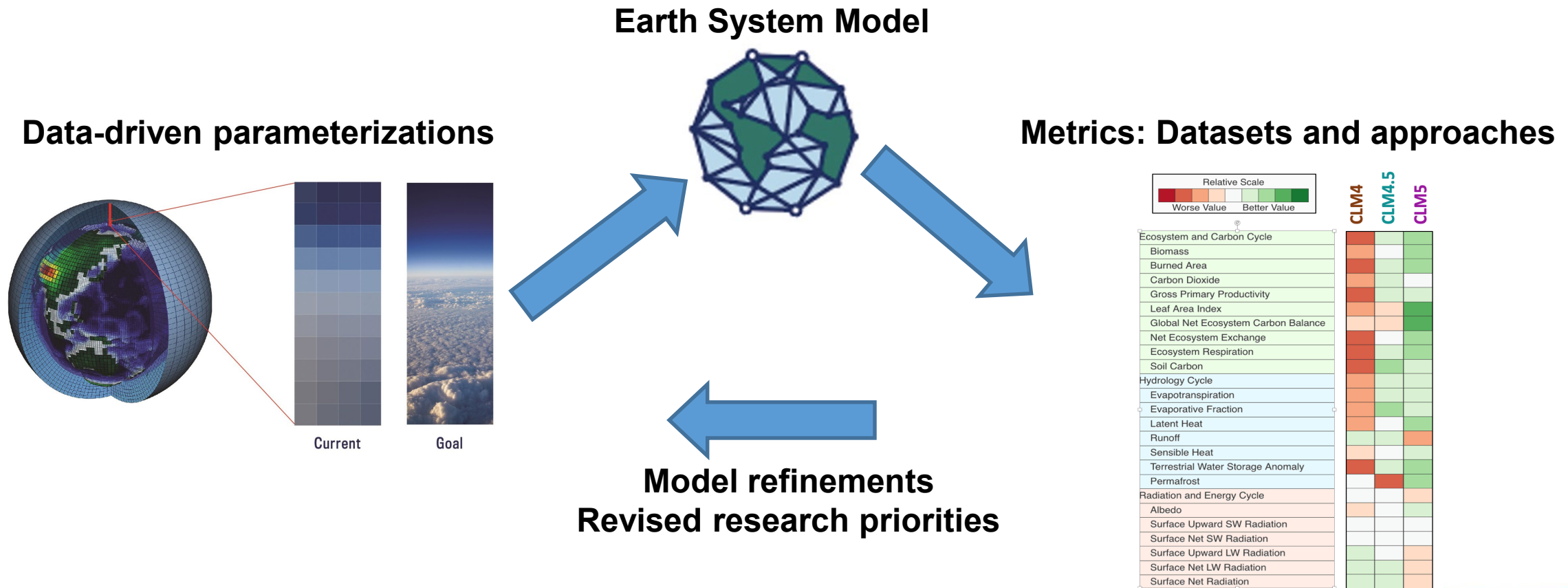
**Model skill assessment  
CESM's land module**

Relative Scale  
Worse Value Better Value

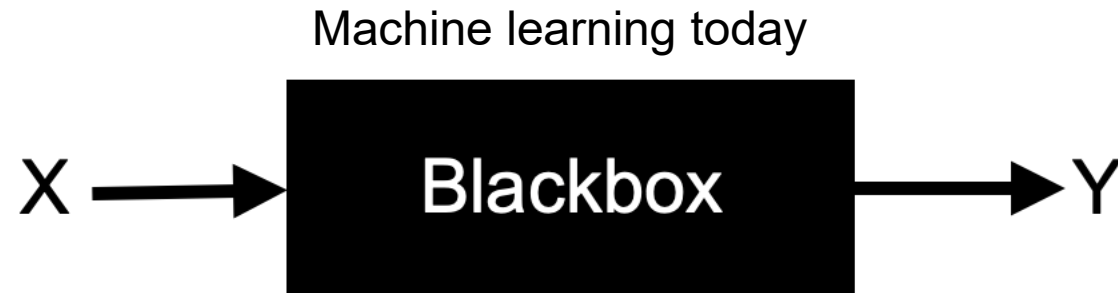
	CLM4*	CLM4.5	CLM5
Ecosystem and Carbon Cycle			
Biomass			
Burned Area			
Carbon Dioxide			
Gross Primary Productivity			
Leaf Area Index			
Global Net Ecosystem Carbon Balance			
Net Ecosystem Exchange			
Ecosystem Respiration			
Soil Carbon			
Hydrology Cycle			
Evapotranspiration			
Evaporative Fraction			
Latent Heat			
Runoff			
Sensible Heat			
Terrestrial Water Storage Anomaly			
Permafrost			
Radiation and Energy Cycle			
Albedo			
Surface Upward SW Radiation			
Surface Net SW Radiation			
Surface Upward LW Radiation			
Surface Net LW Radiation			
Surface Net Radiation			

\*CLM = Community Land Model

# Data-driven parameterizations and metrics catalyze ESM development



# ML must be evolved to respect physics and be scientifically interpretable



# ML must be evolved to respect physics and be scientifically interpretable

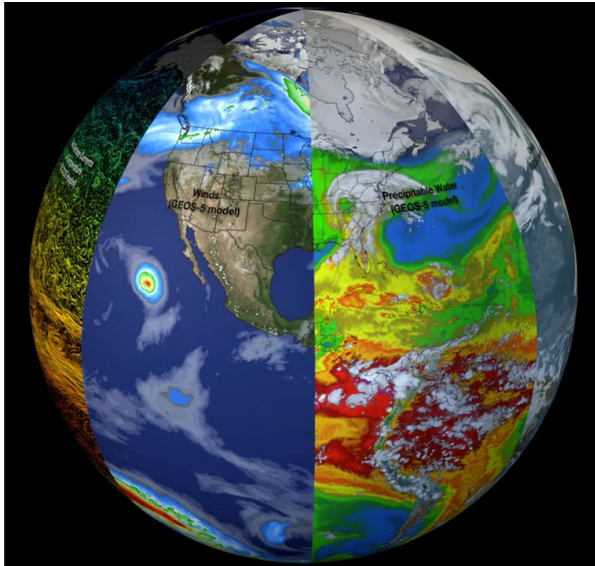


# ML must be evolved to respect physics and be scientifically interpretable



- **Why** does X lead to Y?
- How will **future change in X**, outside current distribution, impact Y?
- What are the **limits of ML** for a particular problem?

# Complex climate data provides a rich challenge for Machine Learning



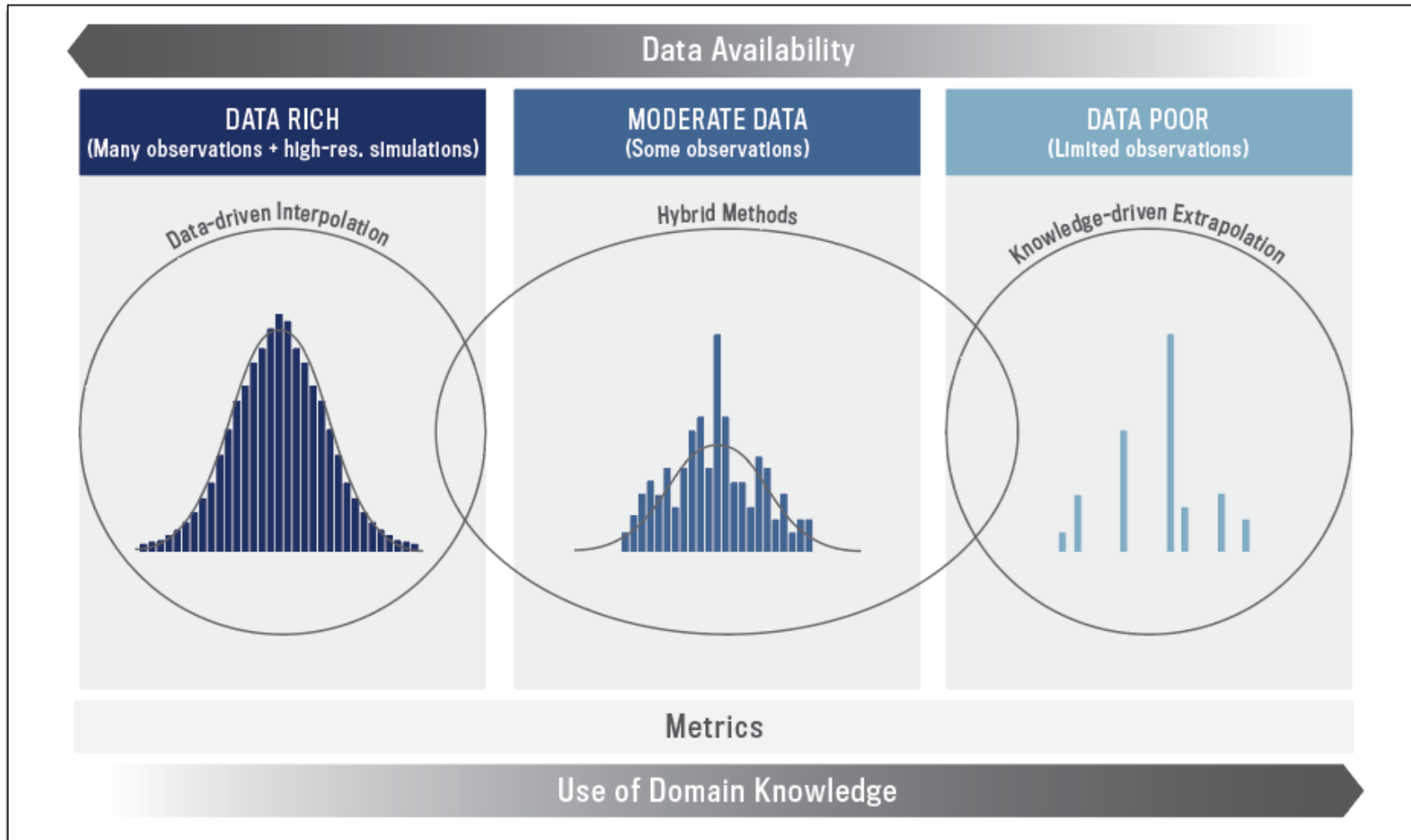
Structured, linked, dynamic datasets



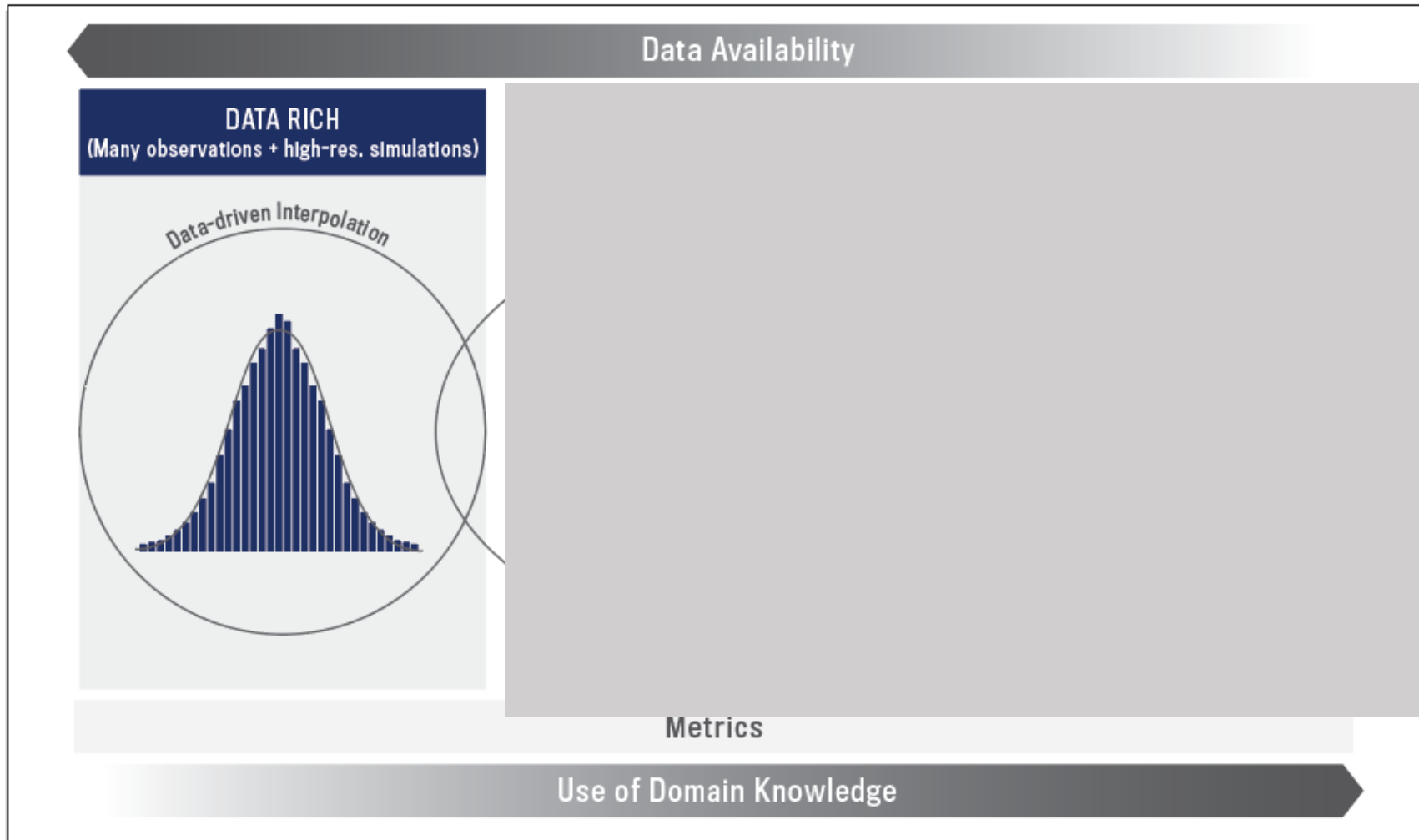
Progress in  
Machine  
Learning



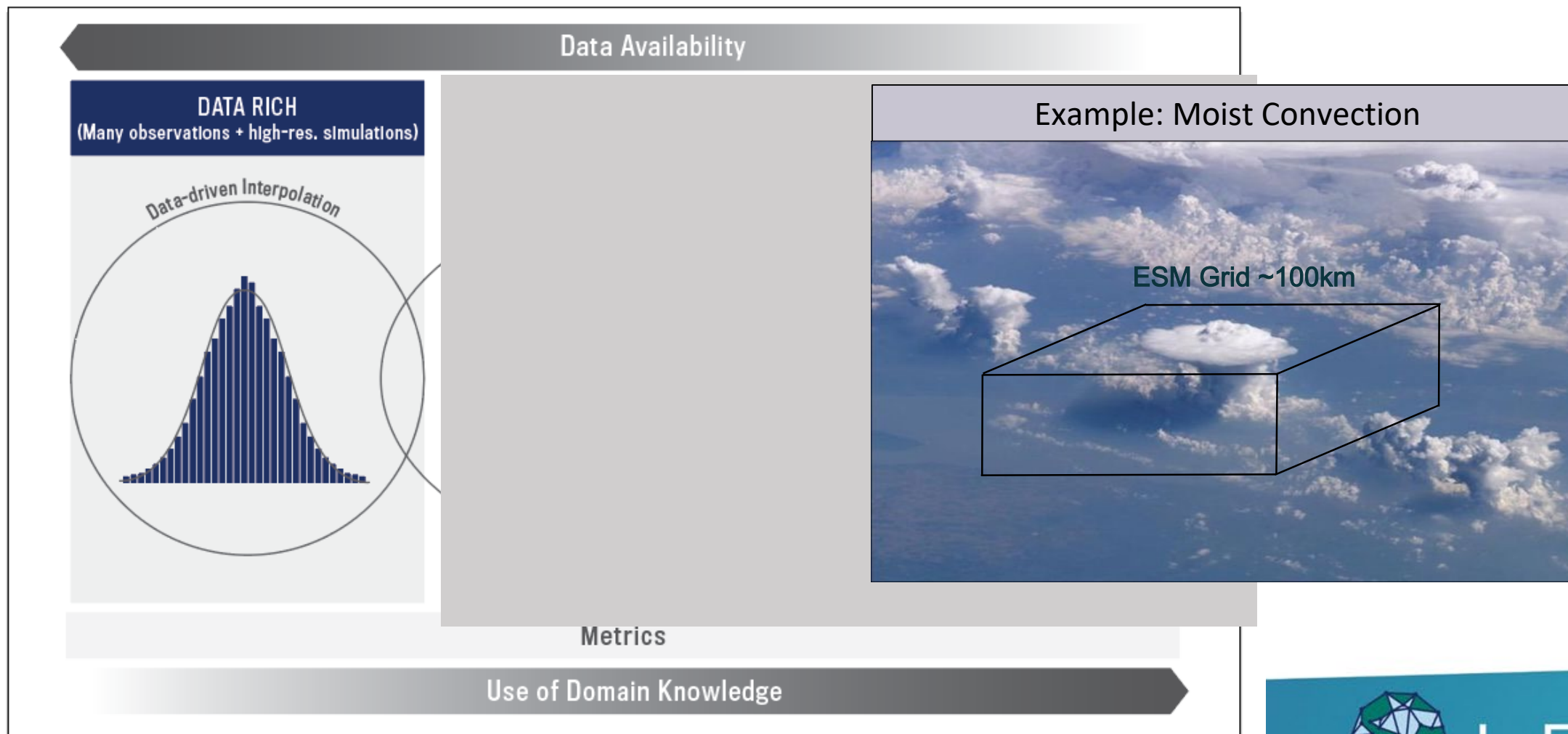
# LEAP research is organized on a spectrum of data availability and process knowledge



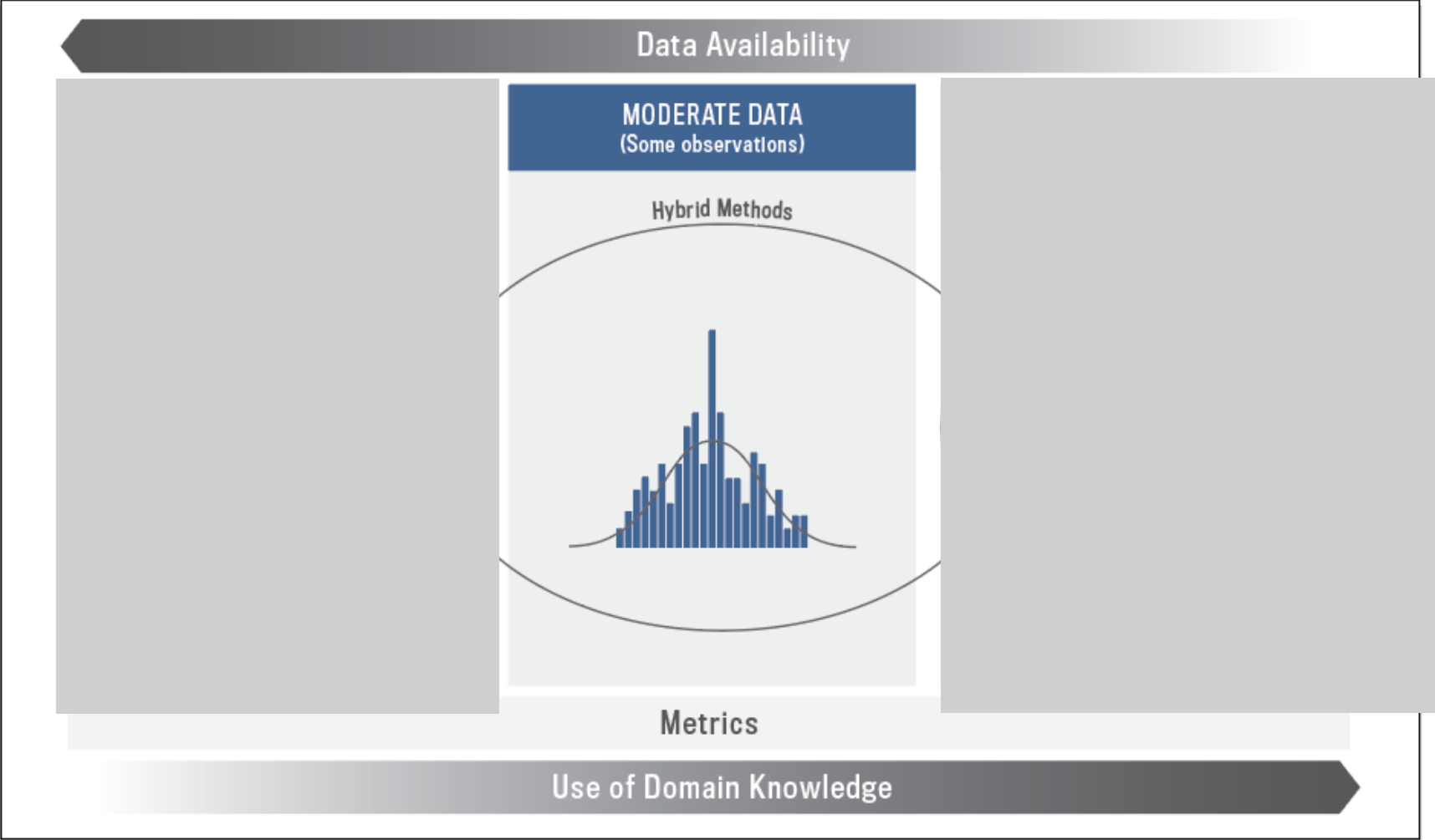
# For Data Rich, real or simulated data exist, but compute load disallows direct ESM application



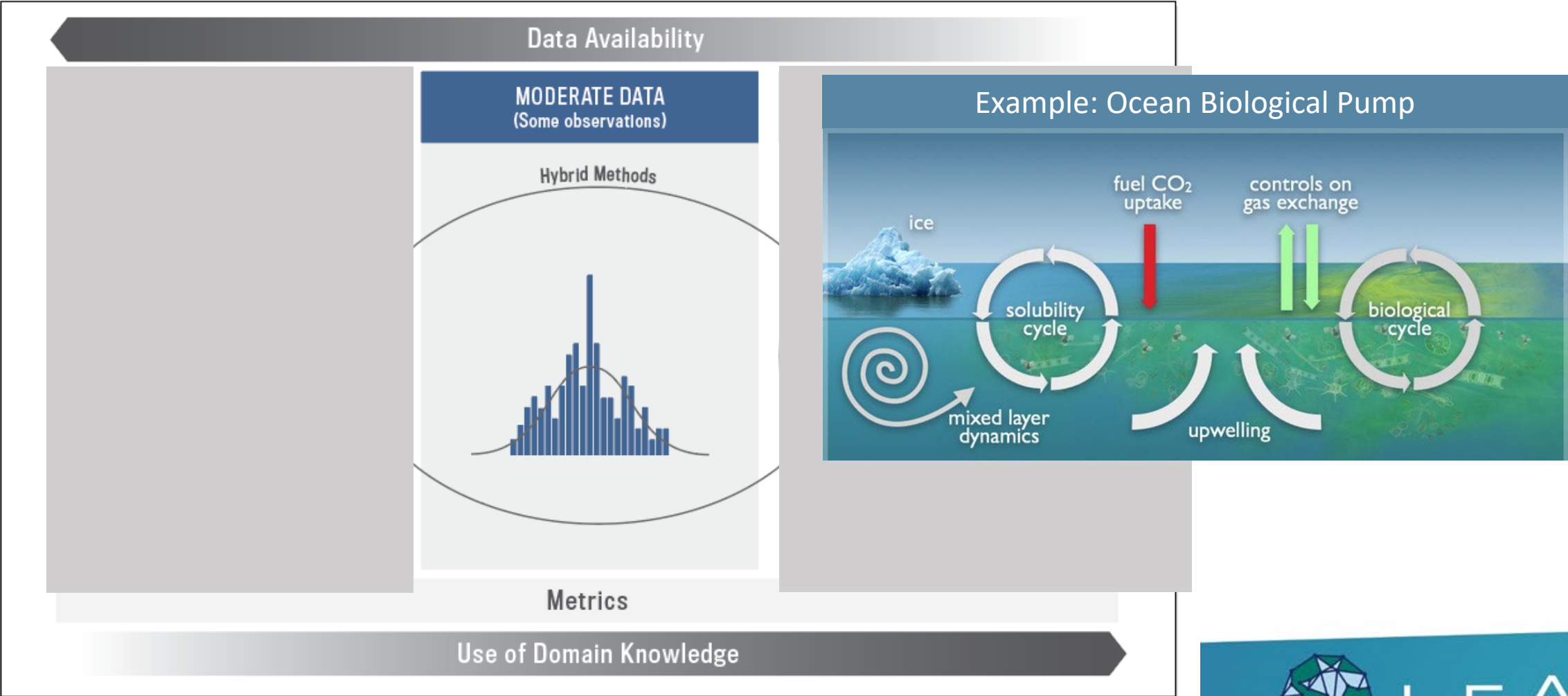
# For Data Rich, real or simulated data exist, but compute load disallows direct ESM application



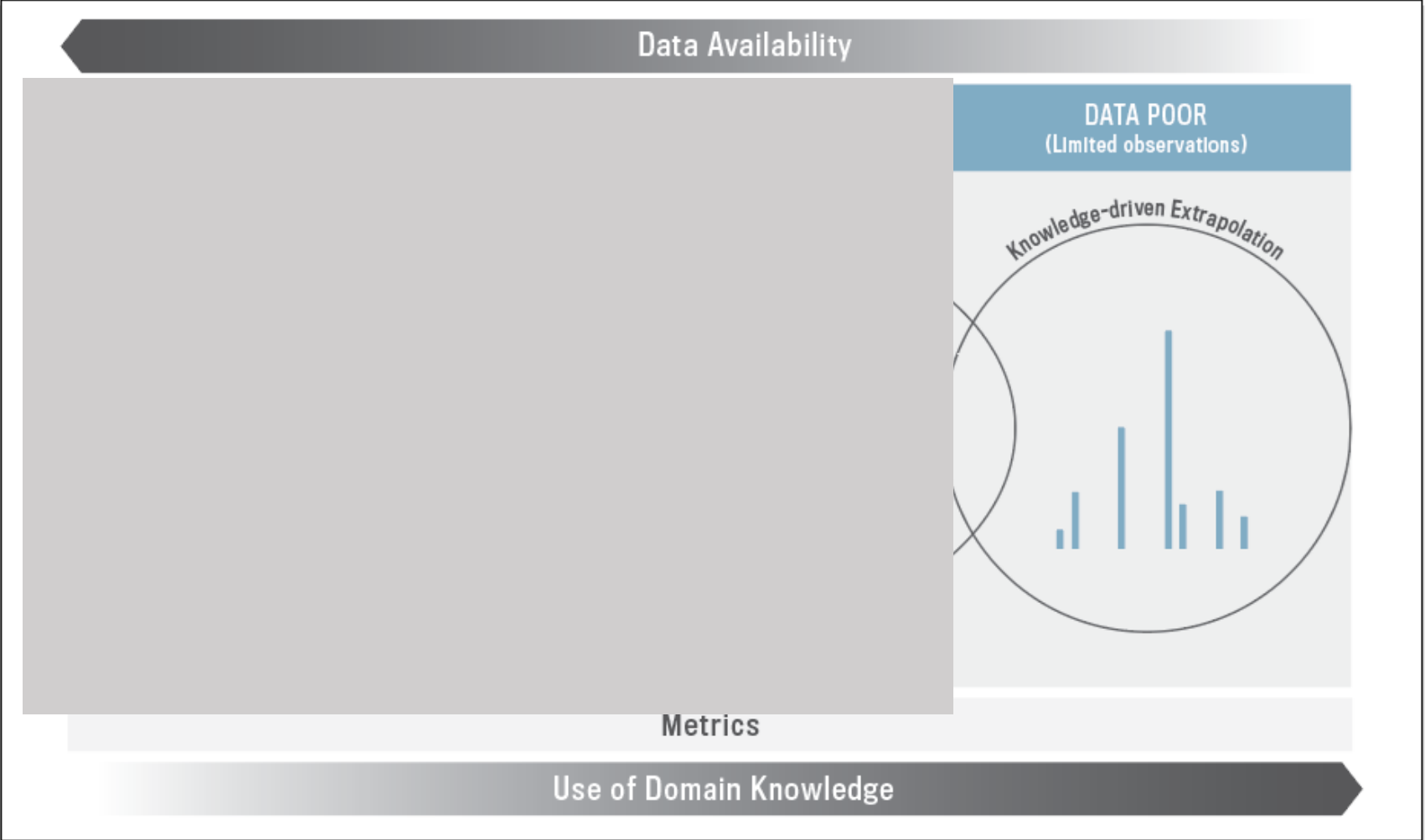
# For Moderate Data, data science can reveal causal links and increase generalizability



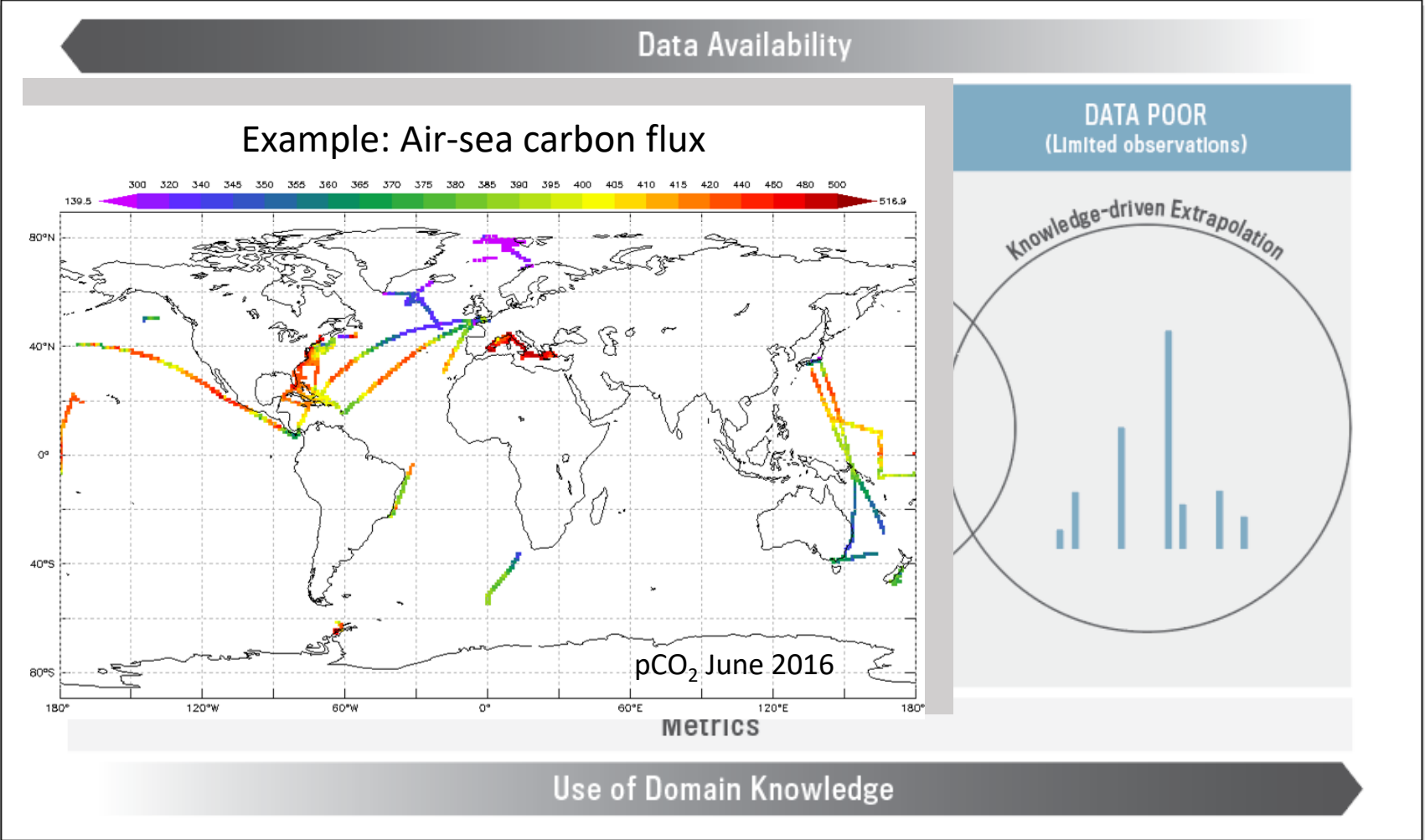
# For Moderate Data, data science can reveal causal links and increase generalizability



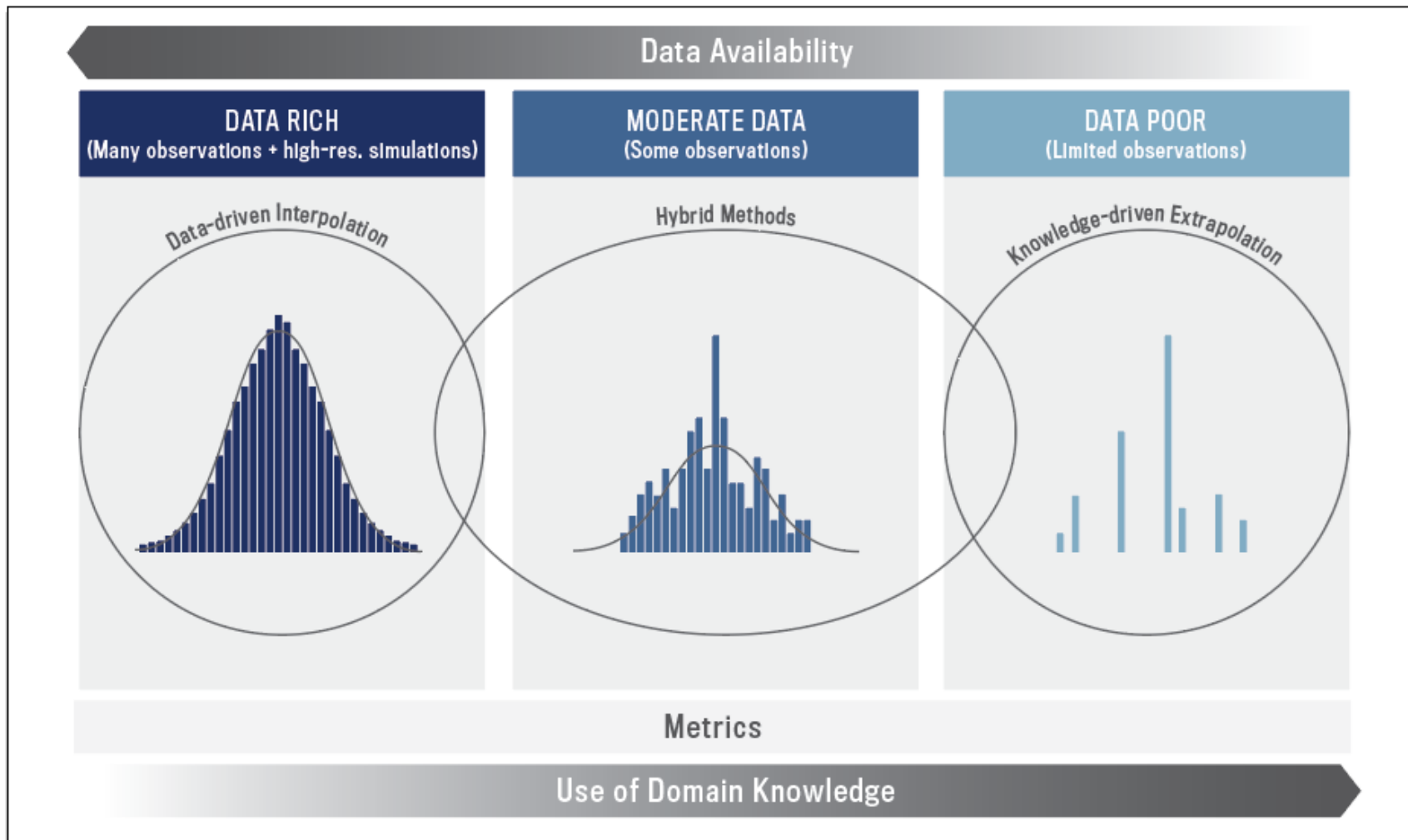
# For Data Poor, process knowledge will support parameterizations and uncertainty assessment



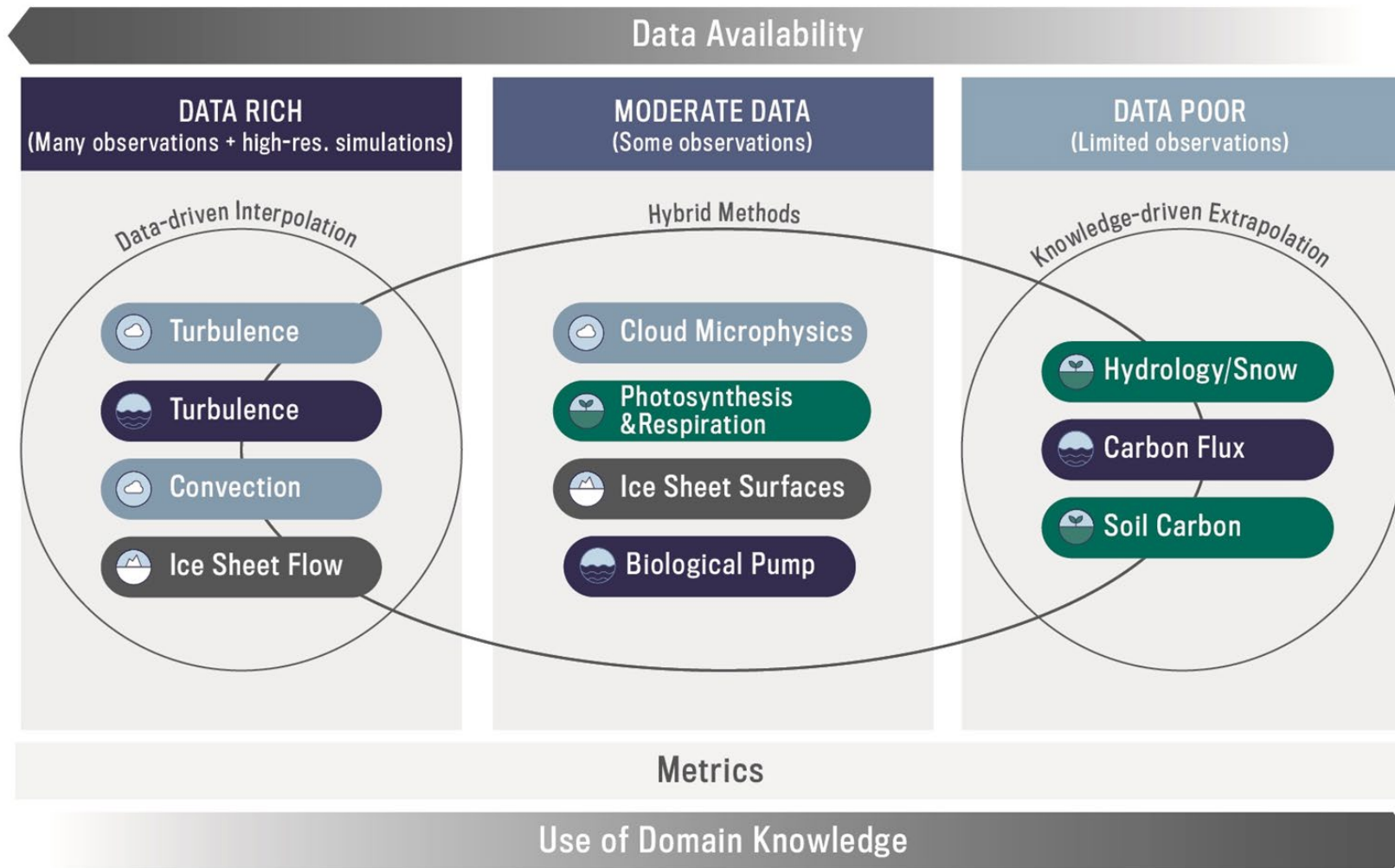
# For Data Poor, process knowledge will support parameterizations and uncertainty assessment



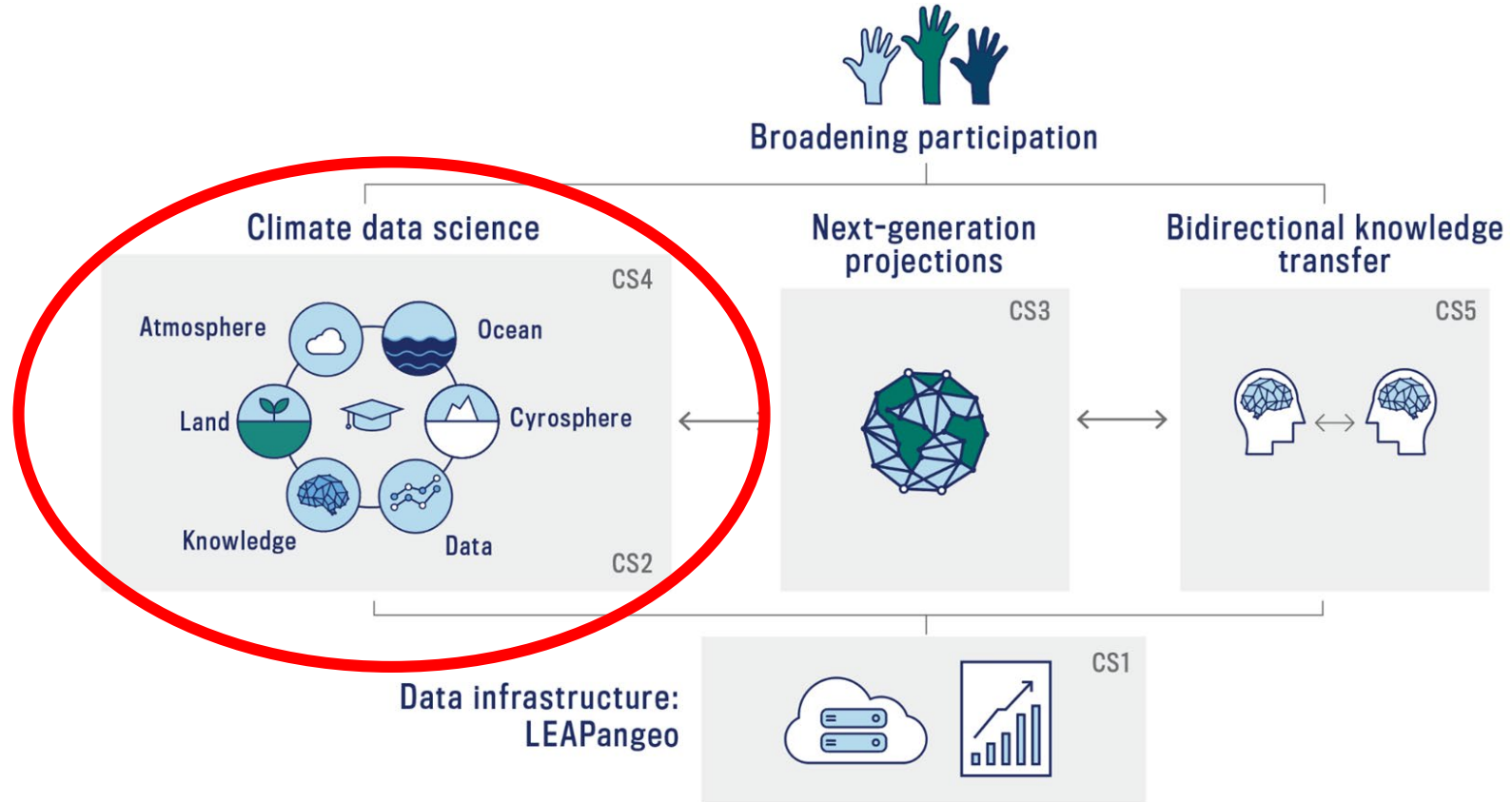
# LEAP research is organized on a spectrum of data availability and process knowledge



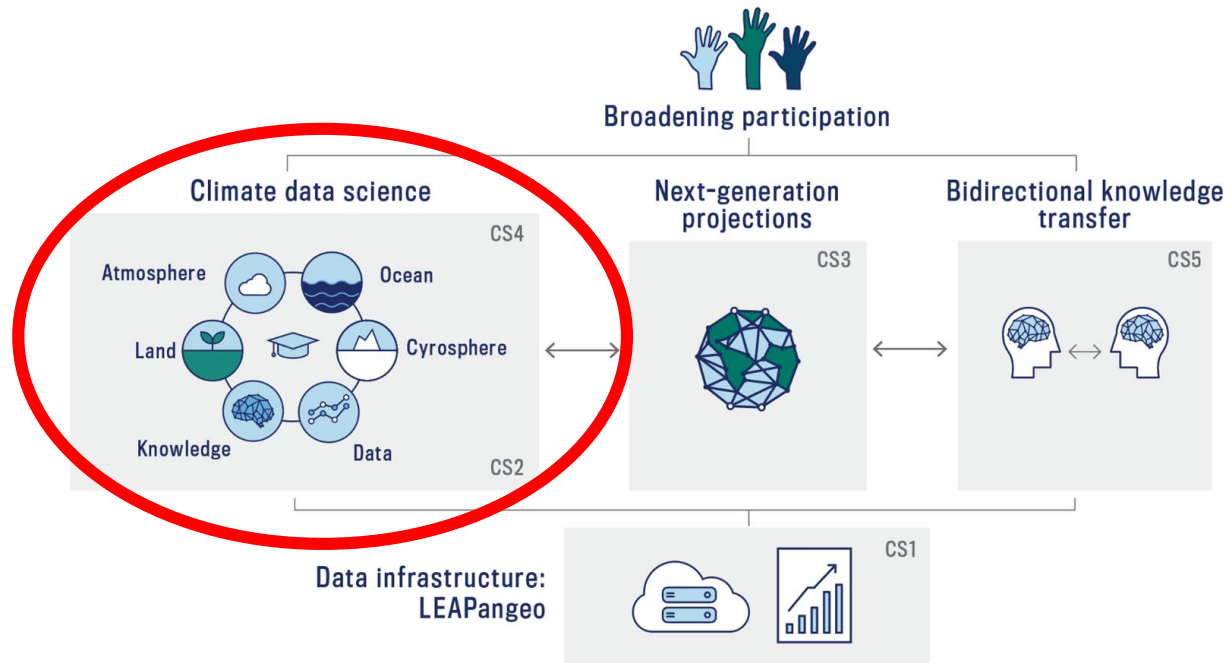
# LEAP research is organized on a spectrum of data availability and process knowledge



# Creating a community in *climate data science* is essential

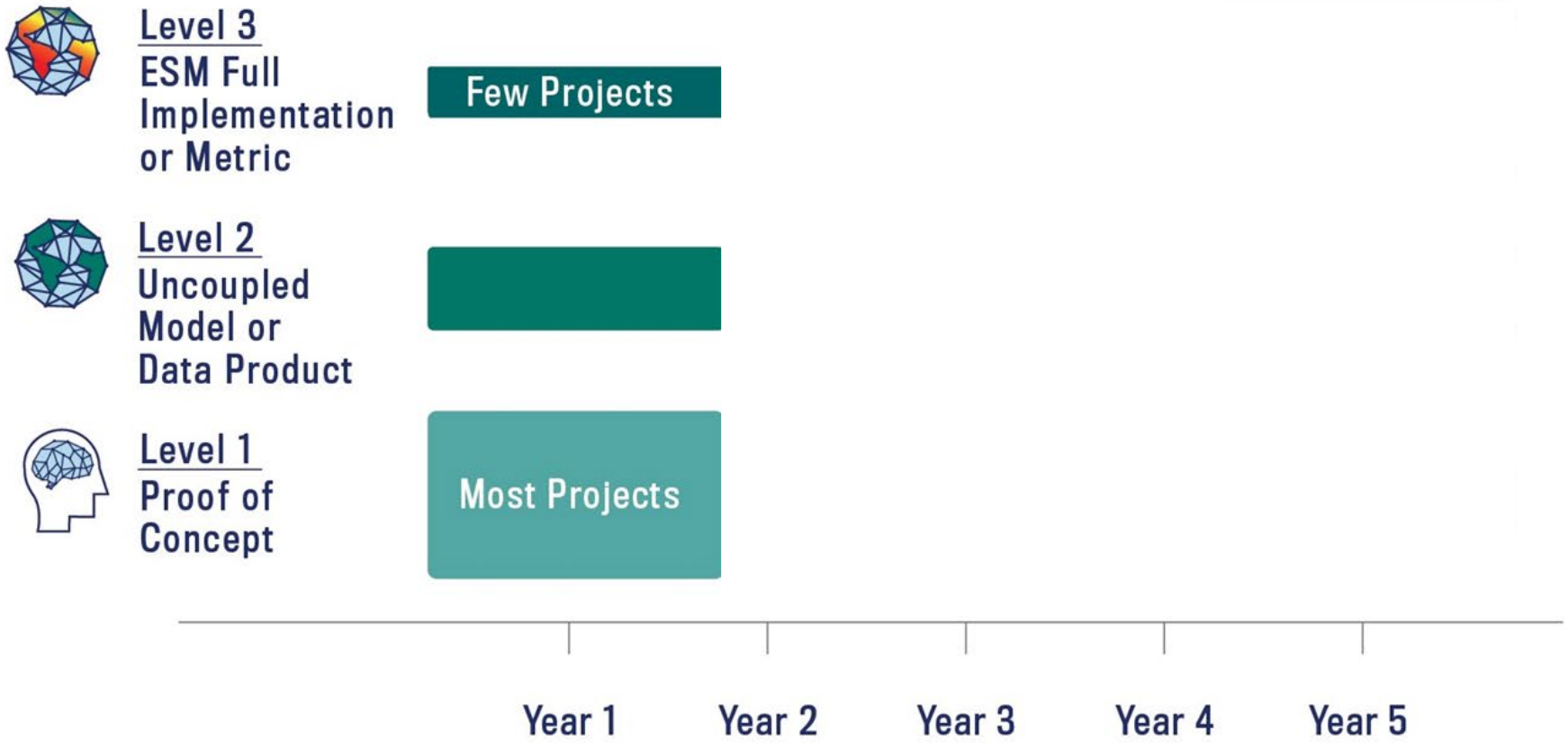


# Creating a community in *climate data science* is essential

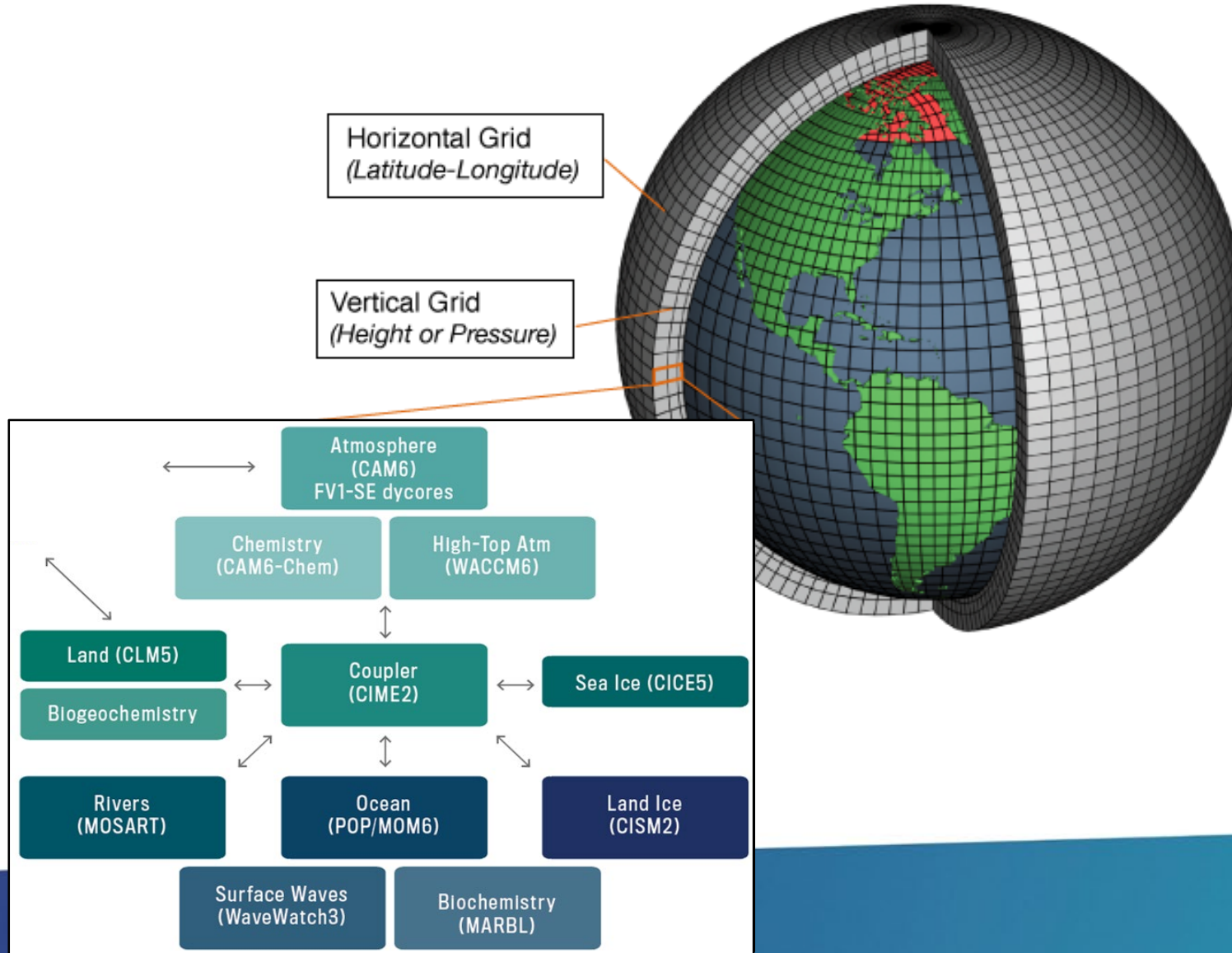


- Education programs from K-12 to postdoctoral
- Research Working Groups
- Annual Meetings
- Comprehensive commitment to diversity, equity and inclusion (DEI)

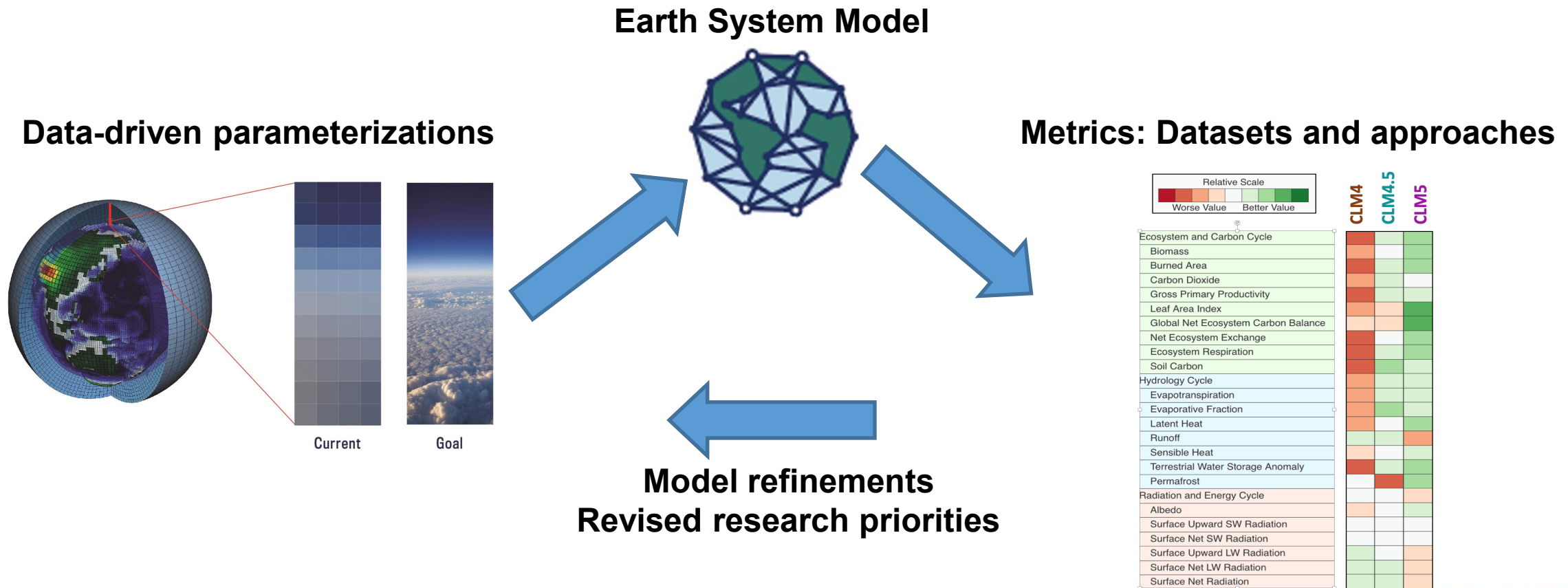
# Research will be structured to progress from proof of concept to ESM implementation



# Potential research bottleneck: Coupling



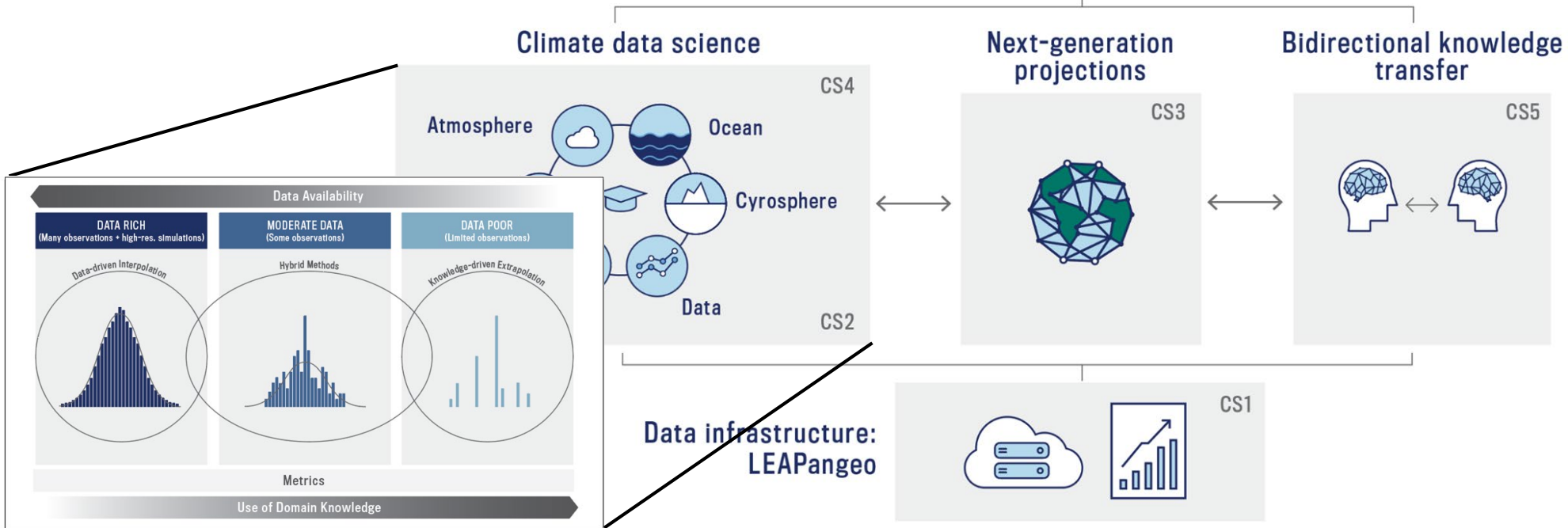
# Data-driven parameterizations and metrics catalyze ESM development



# Structured research in climate data science is key to our vision



Broadening participation



*LEAP forward in the **reliability**, utility, and reach of climate projections through **synergistic innovations** in data science and climate science.*



# Questions and Feedback?

# Data-Rich Research



Geoscience  
Director  
**Laure Zanna**

- Professor of Mathematics + Atmosphere/ Ocean Science, NYU
- PhD, Earth + Planetary Sciences, Harvard

### **Expertise**

- Large-scale climate dynamics
- Ocean Turbulence
- Uncertainty quantification

### **Synergistic activities**

- Lead Principal Investigator, NSF-NOAA Climate Process Team on Ocean Eddy Parameterizations, 2019-
- Member, Ocean Model Development Panel, International CLIVAR, 2019-
- Steering Committee, FAFMIP (Flux-anomaly-forced model intercomparison project, part of CMIP6), 2019-





Data Science  
Director  
**Carl Vondrick**

- Assistant Professor of Computer Science, Columbia
- PhD, Computer Science, MIT

### **Expertise**

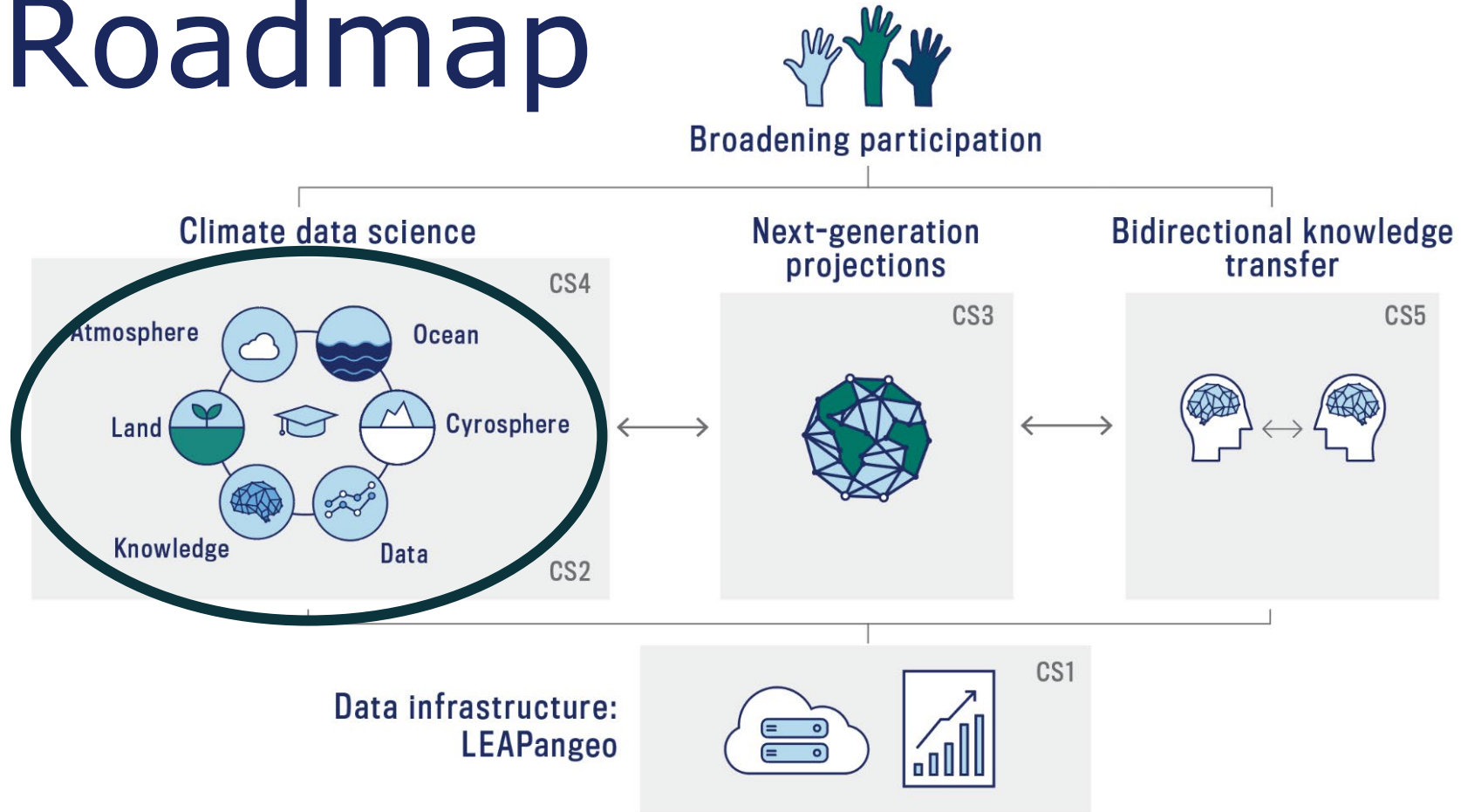
- Machine learning
- Computer vision
- Interpretable models

### **Synergistic activities**

- Research Scientist, Google, 2017-2018
- Co-host of AI4All at Columbia for computer vision
- Best Paper Finalist, IEEE Conference on Computer Vision and Pattern Recognition, 2019

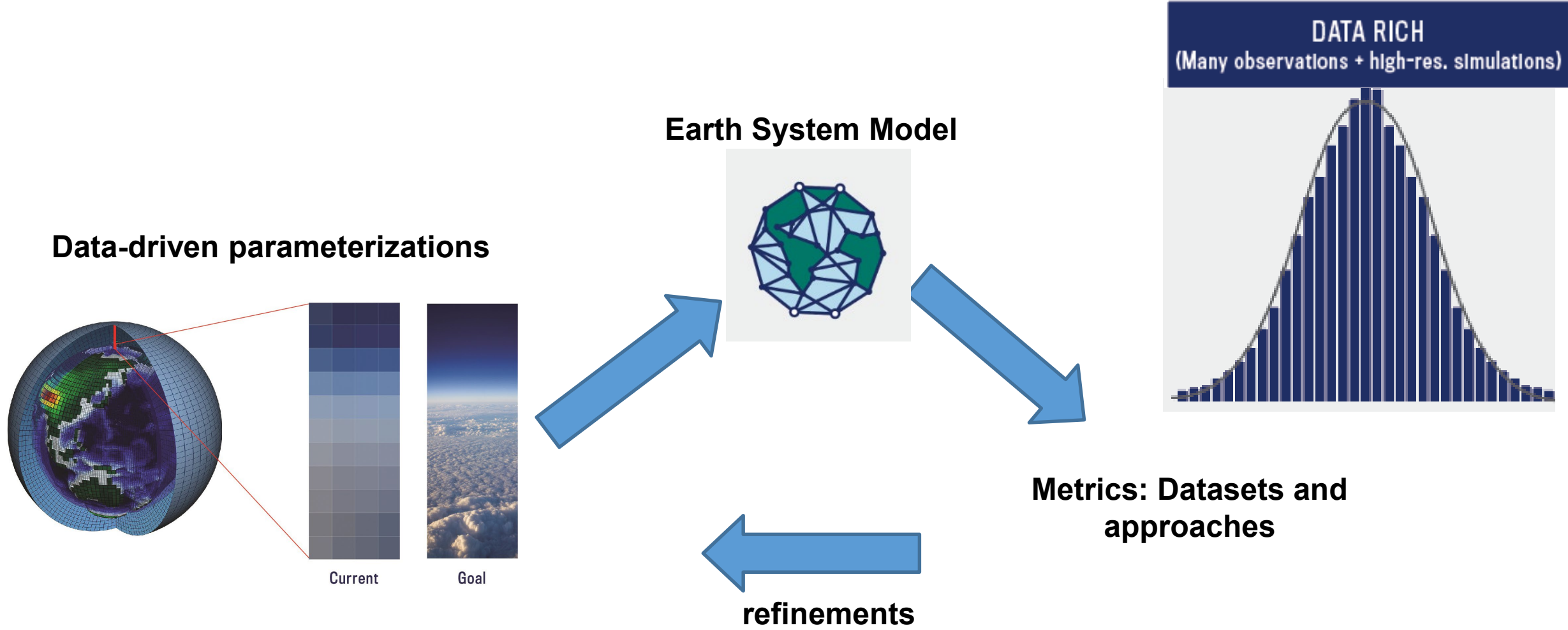


# Our Roadmap

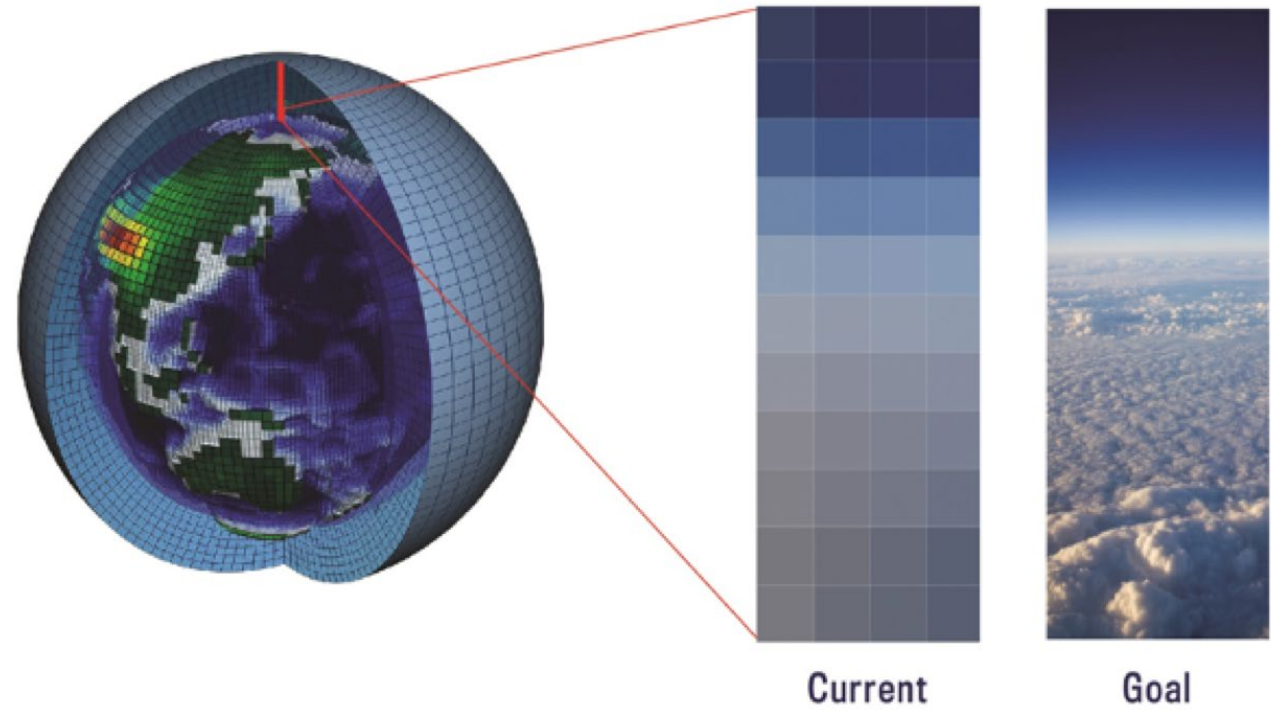
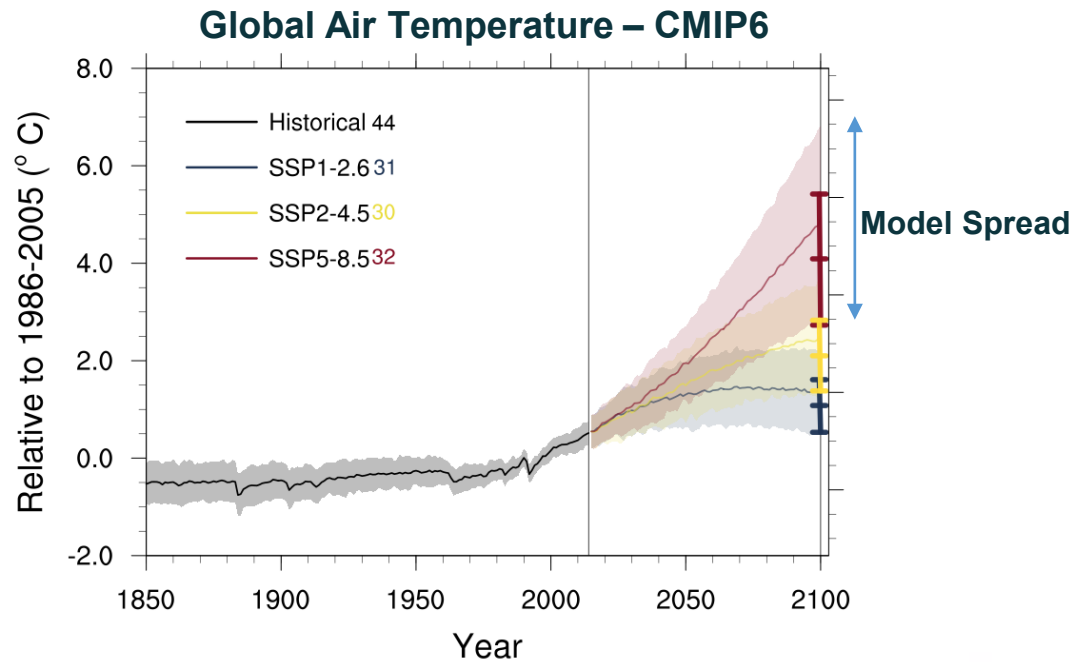


*LEAP forward in the **reliability**, utility, and reach of climate projections through **synergistic innovations in data science and climate science**.*

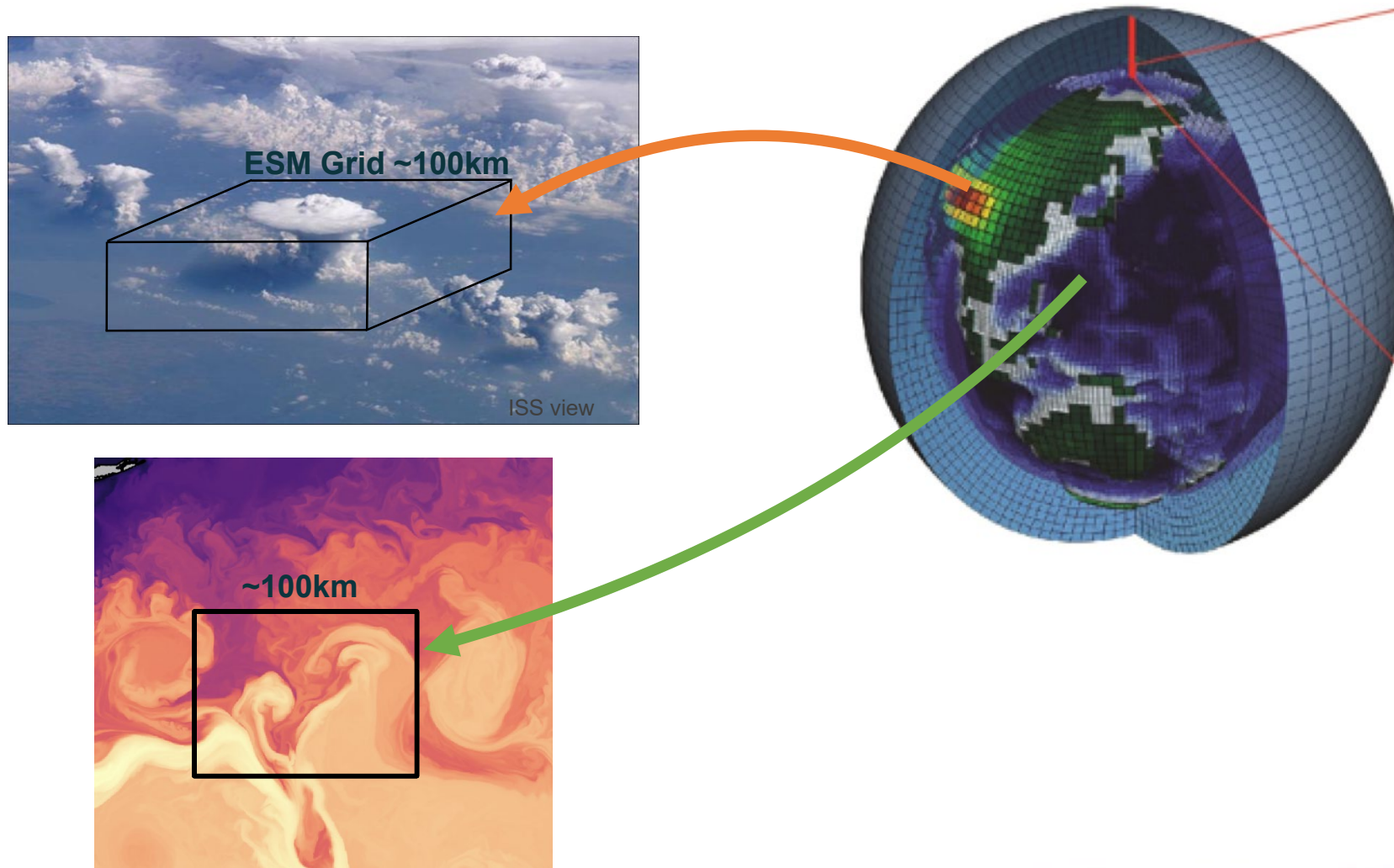
# Next generation of parameterizations & metrics for climate simulations



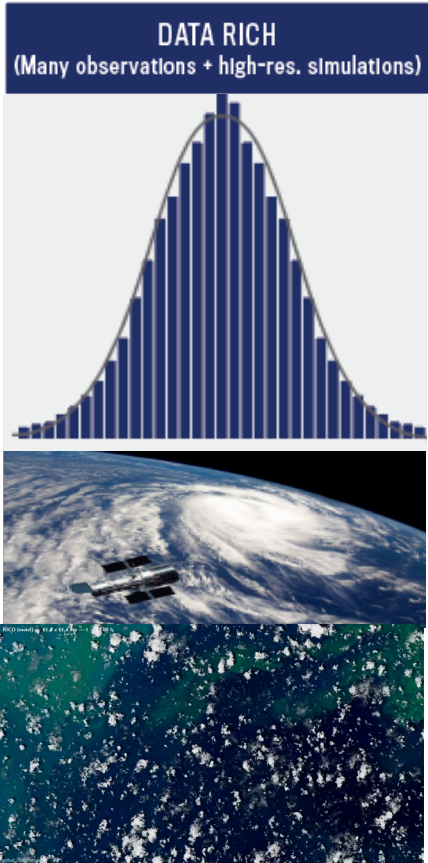
# Uncertainty in projections due to limited computing resources



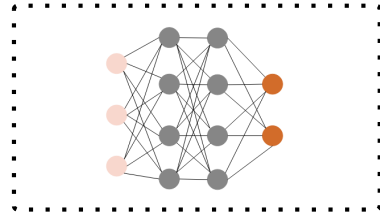
# Parameterizations are rough approximations of missing physics



# LEAP will integrate data science with domain-knowledge to create a new generation of ML parameterizations



Optimized ML  
algorithms



+

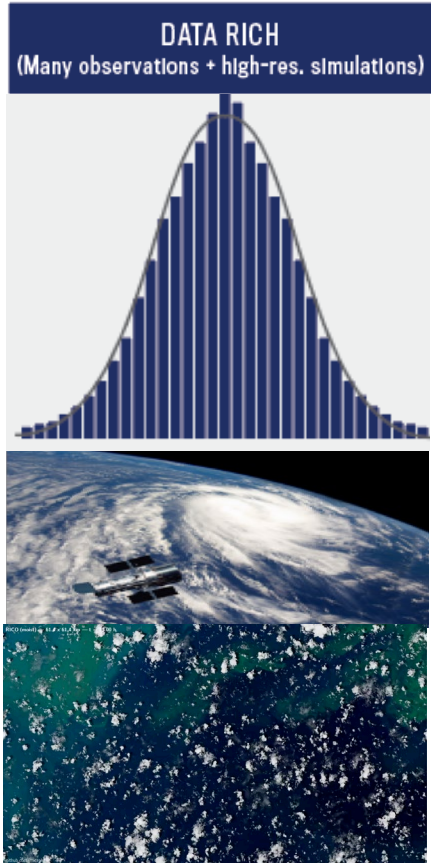
=

$$\frac{\partial c}{\partial t} + \mathbf{v} \cdot \nabla c = 0$$

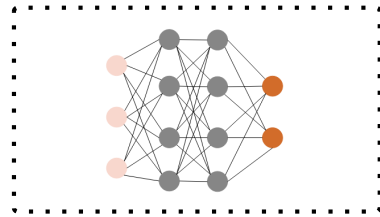
Physics  
constraints



# LEAP will integrate data science with domain-knowledge to create a new generation of ML parameterizations



Optimized ML algorithms



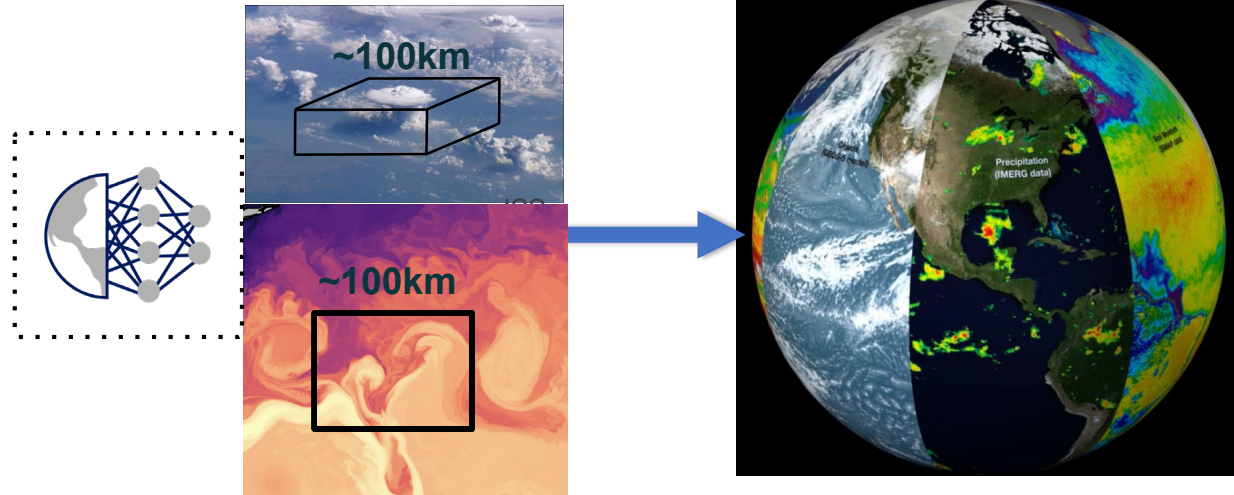
+

=

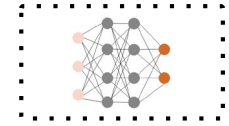
$$\frac{\partial c}{\partial t} + \mathbf{v} \cdot \nabla c = 0$$

Physics constraints

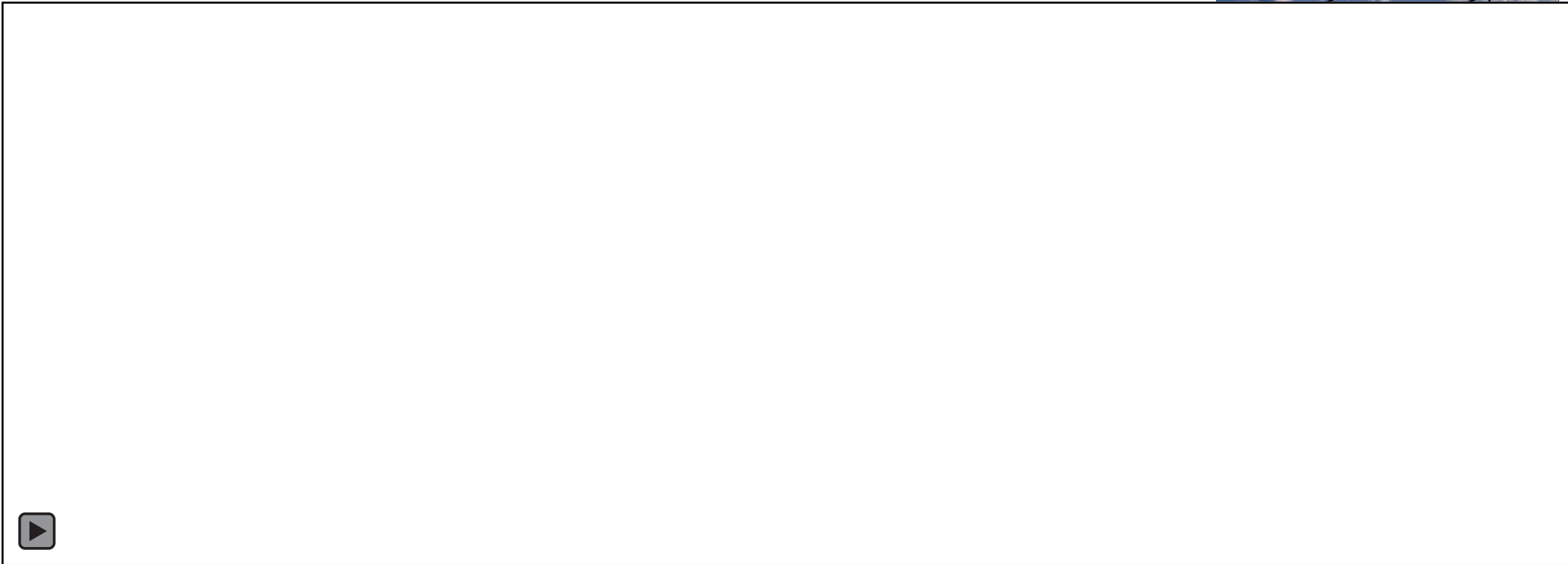
New generation of physics-guided parameterizations



# Proof-of-Concept #1: Convection parametrization

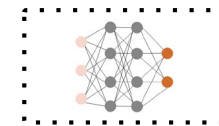


**Atmospheric** convection & clouds set regional + global temperatures, precipitations

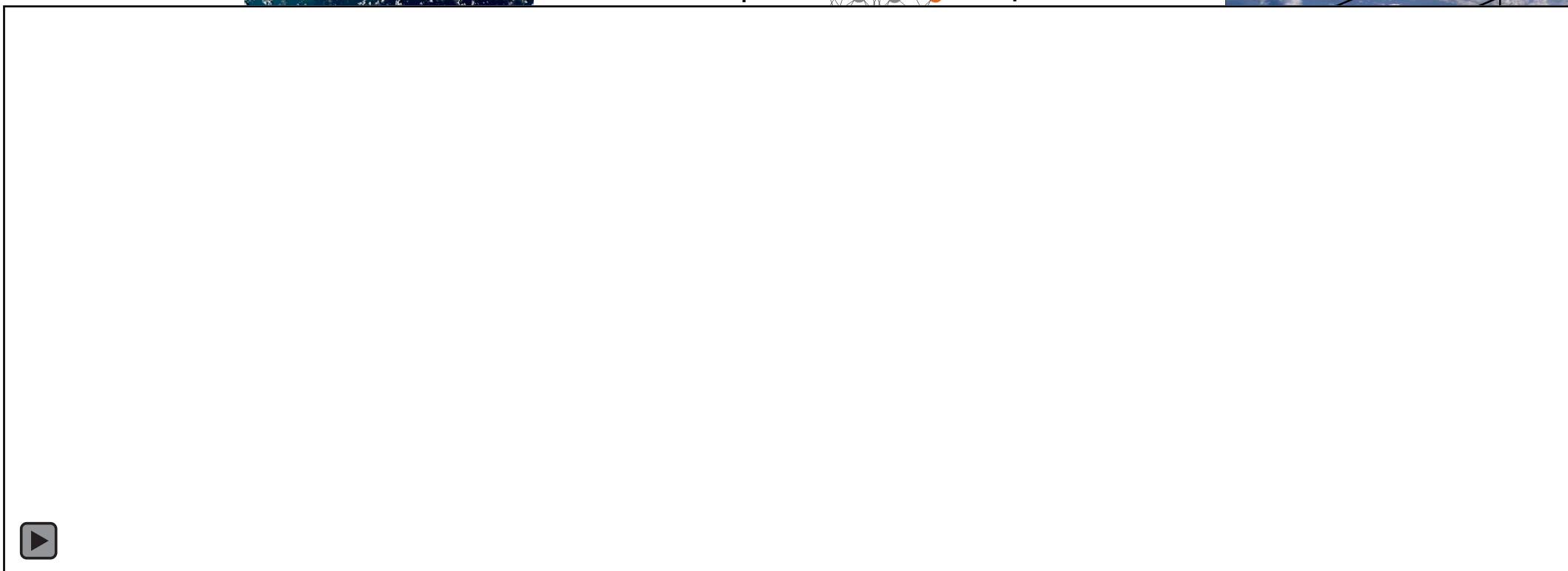
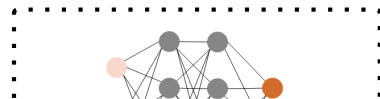


LEAP

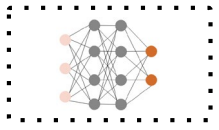
# Proof-of-Concept #1: Convection parametrization



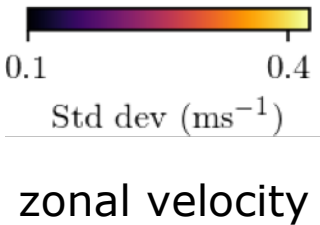
**Atmospheric** convection & clouds set regional + global temperatures, precipitations



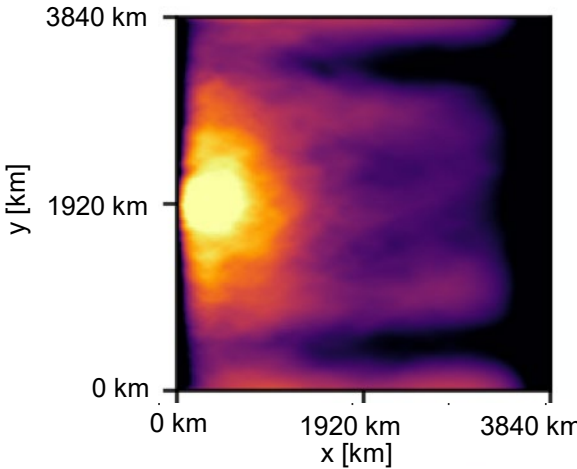
# Proof-of-Concept #2: Ocean turbulence parameterization



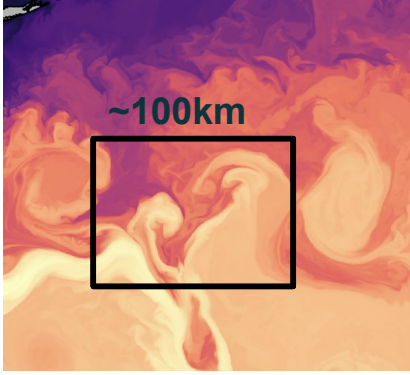
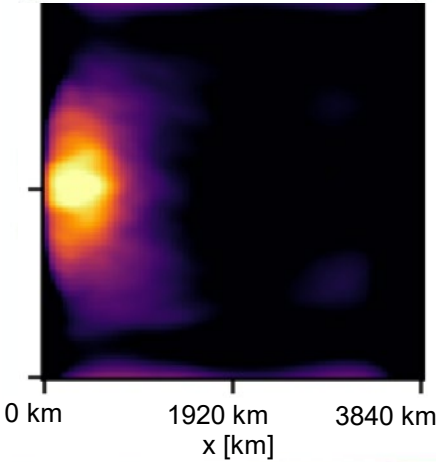
**Ocean eddies:** turbulent features 10-100 km, crucial heat and carbon uptake, sea level



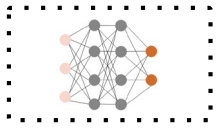
high-res



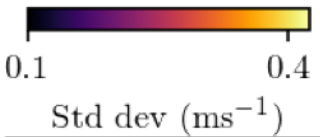
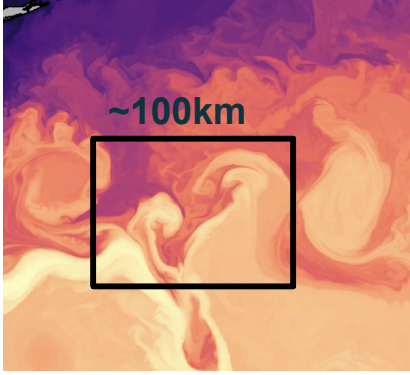
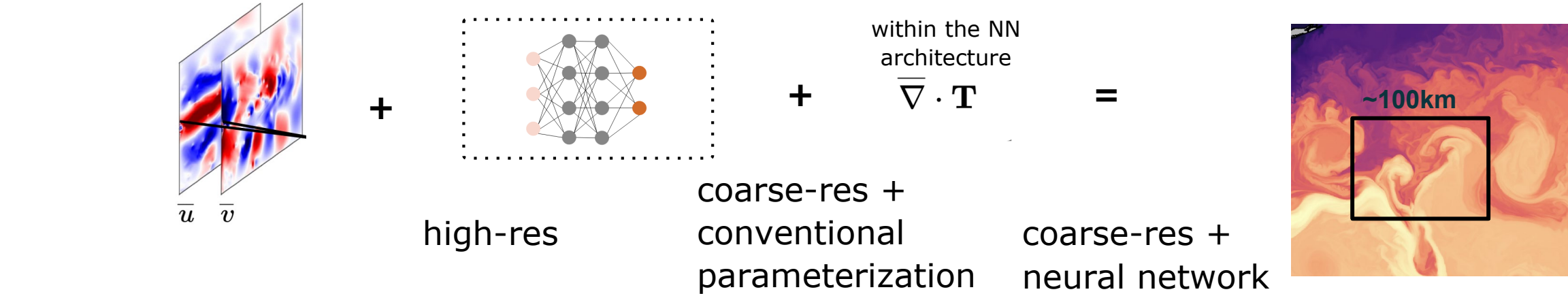
coarse-res +  
conventional  
parameterization



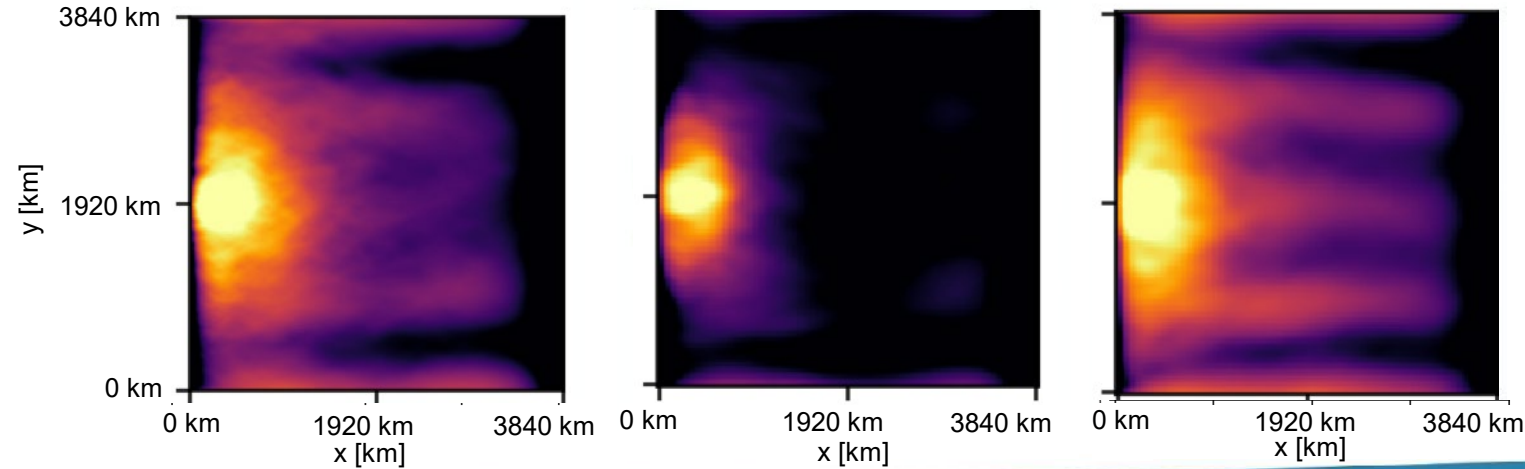
# Proof-of-Concept #2: Ocean turbulence parameterization



**Ocean eddies:** turbulent features 10-100 km, crucial heat and carbon uptake, sea level

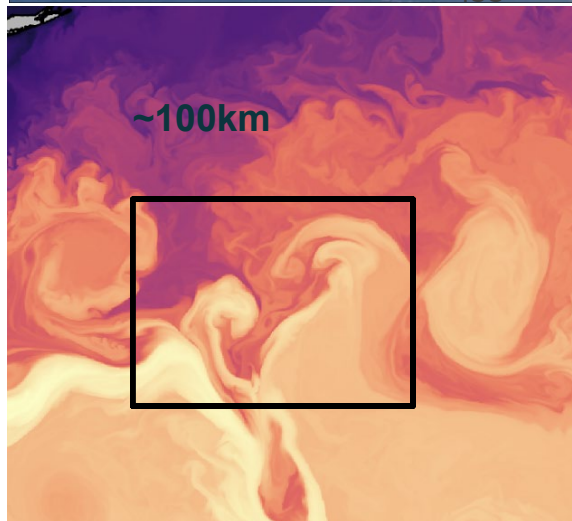
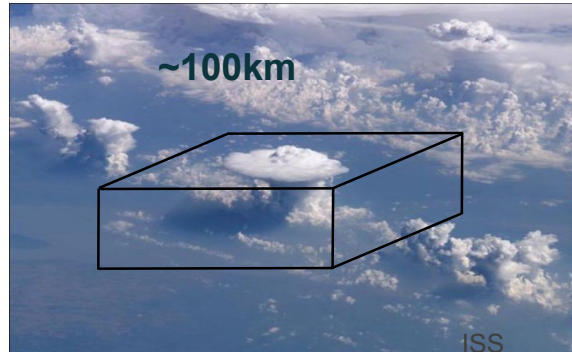


zonal velocity



# LEAPing forward: Promise & Challenges

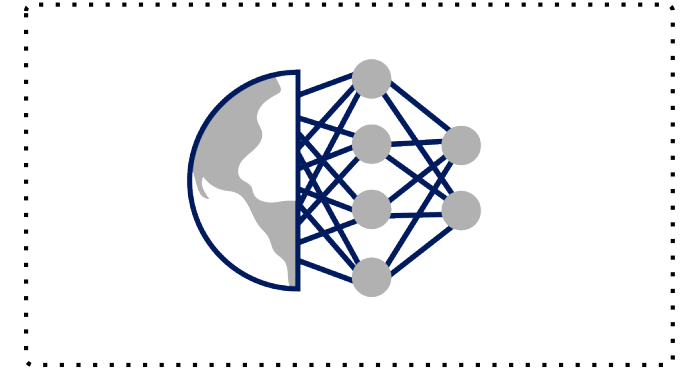
## Exciting proof-of-concepts



## Nexti-generation CESM



## New algorithms



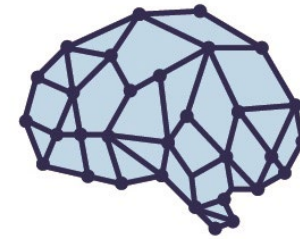
### Pushing ML forward:

- Physics-guided/ generalization
- Uncertainty quantification
- Efficiency
- Metrics

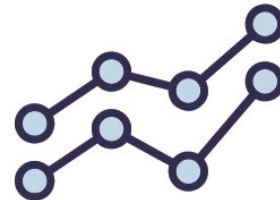


# LEAP tackles four challenges in deep learning for climate science

1. Physical Reliability
2. Training Efficiency
3. Quantification of Uncertainty
4. Metrics



Knowledge

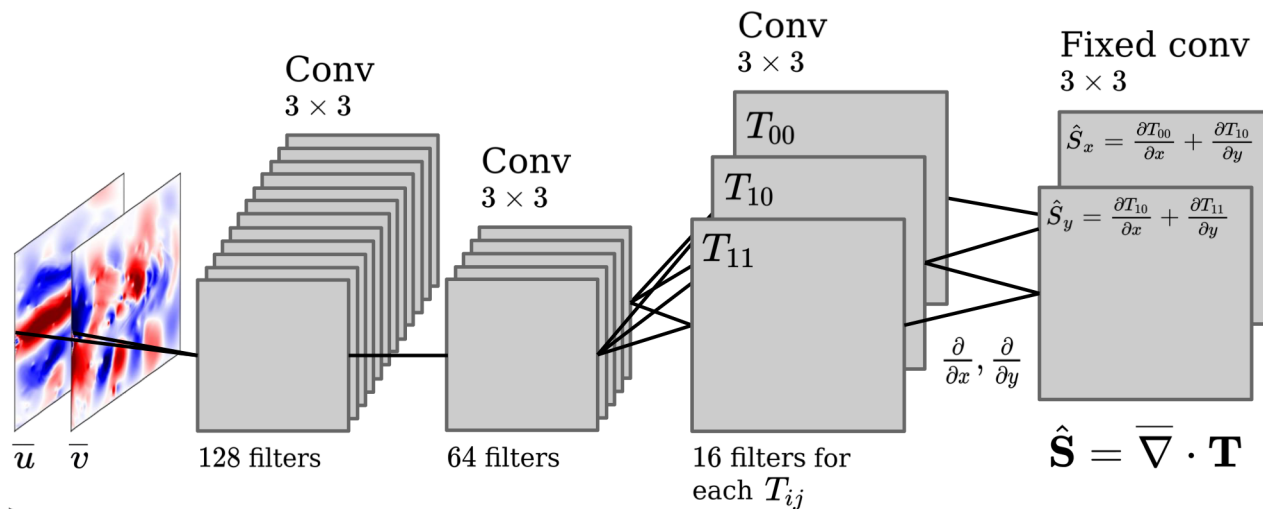


Data



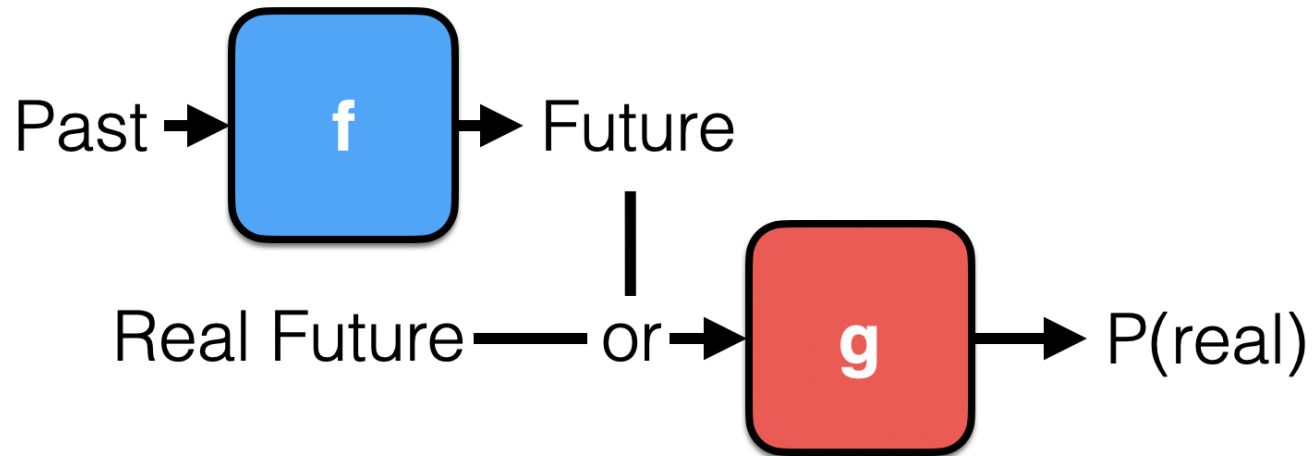
LEAP

# Physical constraints improve reliability and efficiency of machine learning models



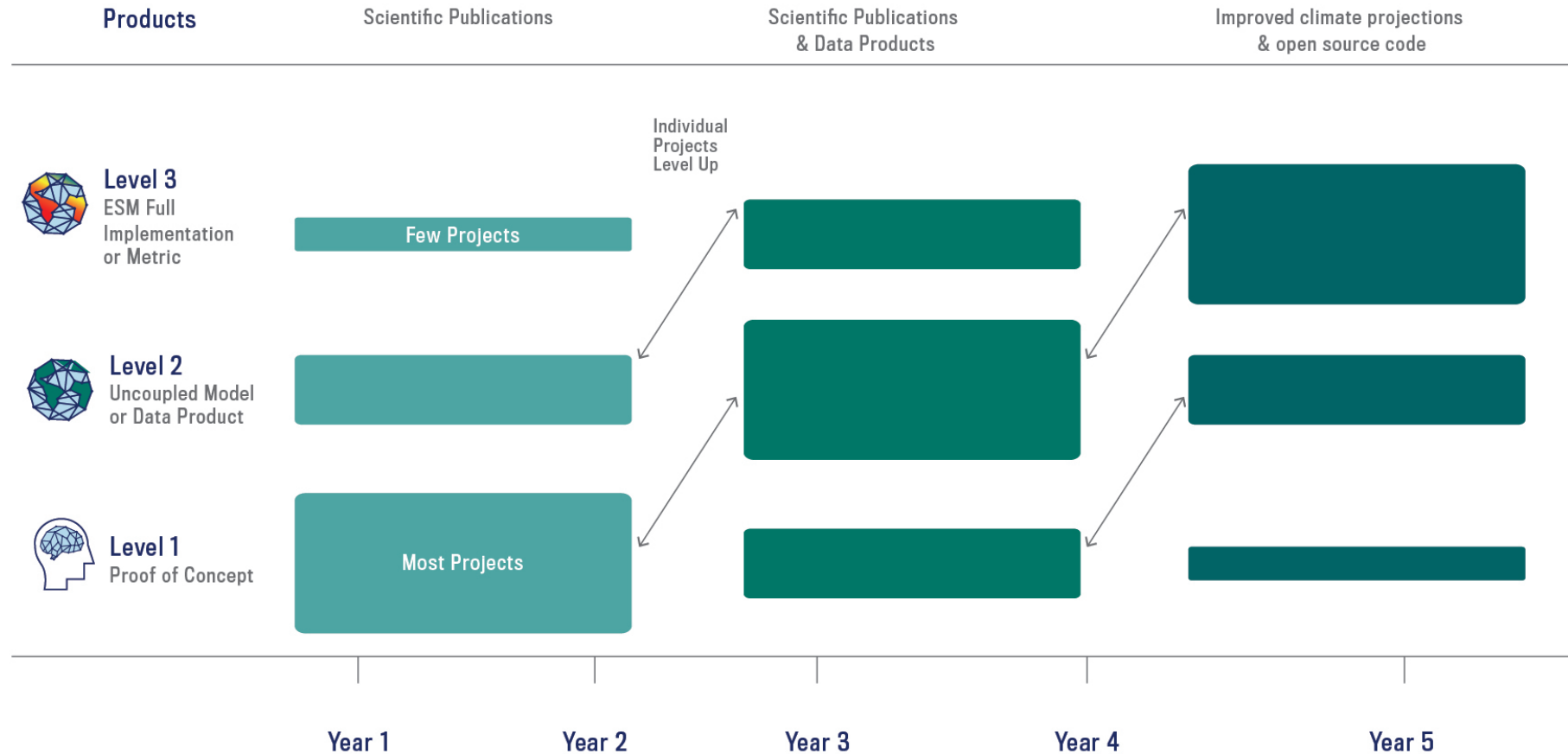
1. Soft constraints
2. Hard constraints by construction
3. Embedded optimization problem inside network

# LEAP incorporates physics and data into metrics and loss functions



1. Higher order moments
2. Extreme values
3. Uncertainty estimates from simulation
4. Adversarial learning

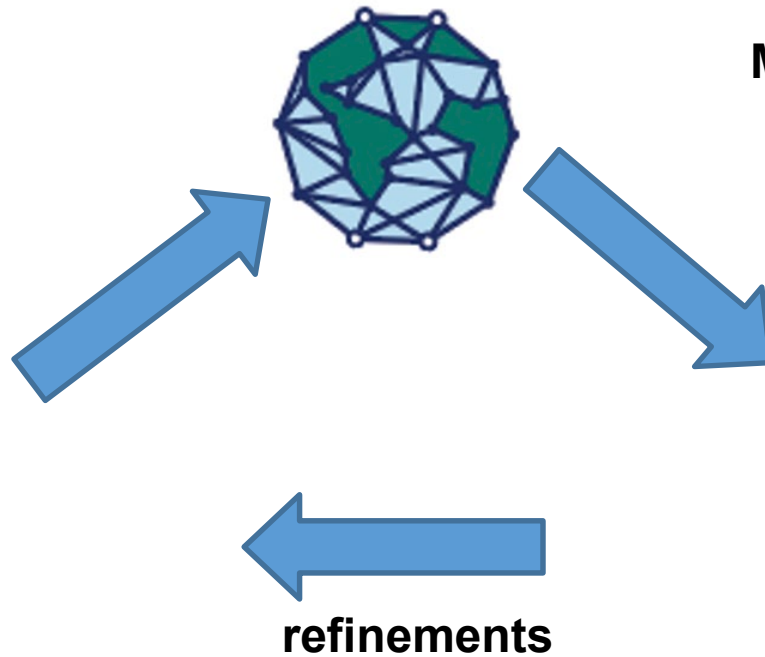
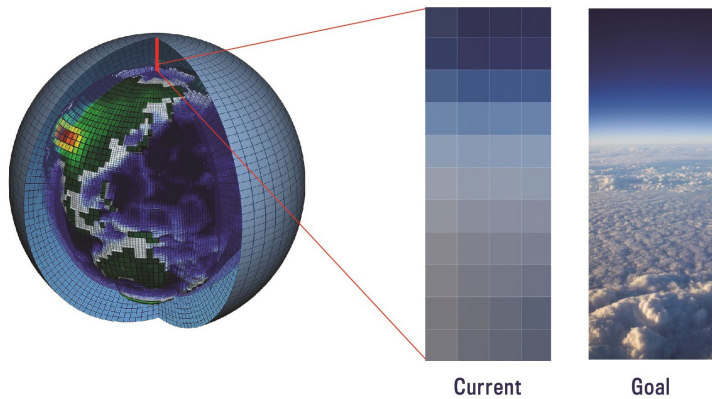
# LEAP will progressively structure research to achieve impact on ESMs



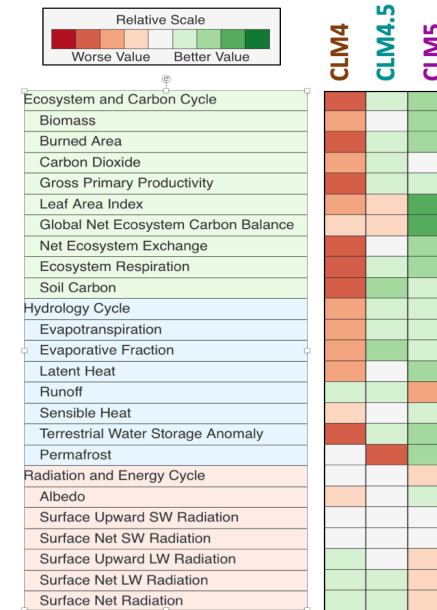
# LEAP uses rich datasets, physical constraints, vetted metrics to improve parameterizations + climate simulations

## Earth System Model

### Data-driven parameterizations



### Metrics: Datasets and approaches



# Questions?

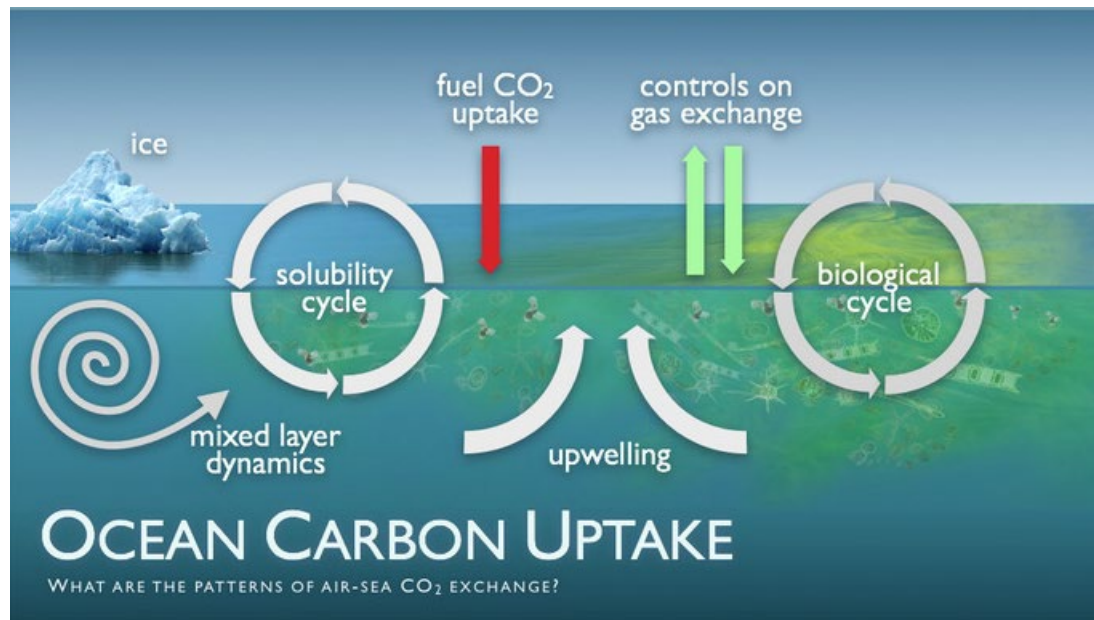
Dr. Laure Zanna, Professor at NYU  
Dr. Carl Vondrick, Asst. Professor at Columbia  
Research Co-Directors



# Data-Moderate Research

Dr. Laure Zanna, Professor at NYU  
Dr. Carl Vondrick, Asst. Professor at Columbia  
Research Co-Directors

# Equations are often unknown



- Only approximate solutions
- Incorrect representation in climate models
- Moderate observational data available

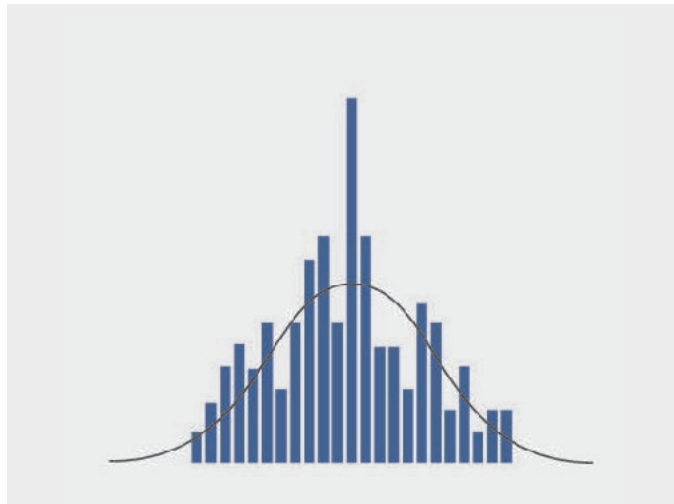
# What's missing in machine learning?



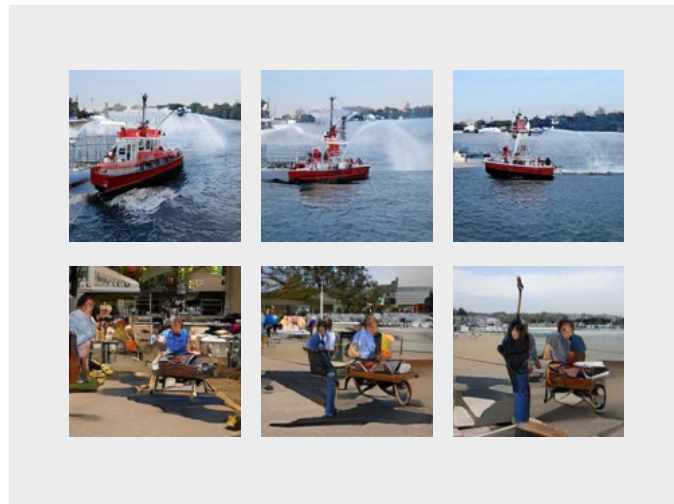
1. Generalization
2. Causal reasoning
3. Interpretability



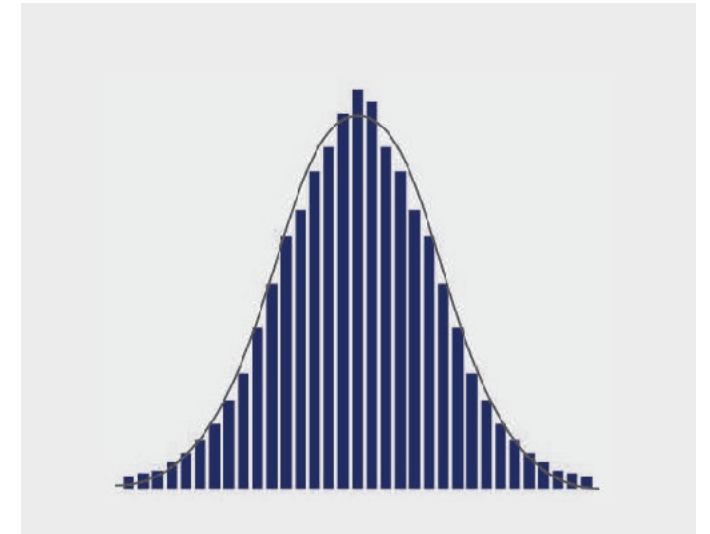
# Neural synthesis drives ML generalization



Observational  
Distribution



Neural Synthesis



Interventional  
Distribution



# What's missing in machine learning?



1. Generalization

2. Causal reasoning

3. Interpretability

# Causal inference enables richer analysis

<b>Level</b>	<b>Typical Question</b>	<b>Examples</b>
Associational	How are X and Y related?	What do clouds tell us about rain?



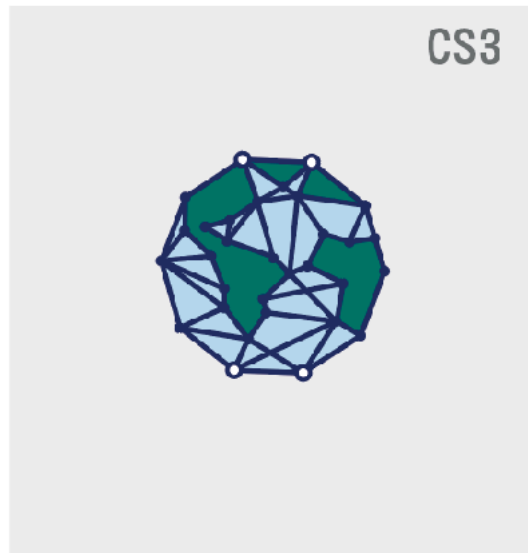
# LEAP reasons about causal structure

Level	Typical Question	Examples
Associational	How are X and Y related?	What do clouds tell us about rain?
Interventional	What if X?	If more clouds form, will there be more rain?
Counterfactual	Why X?	Did the clouds create rain?

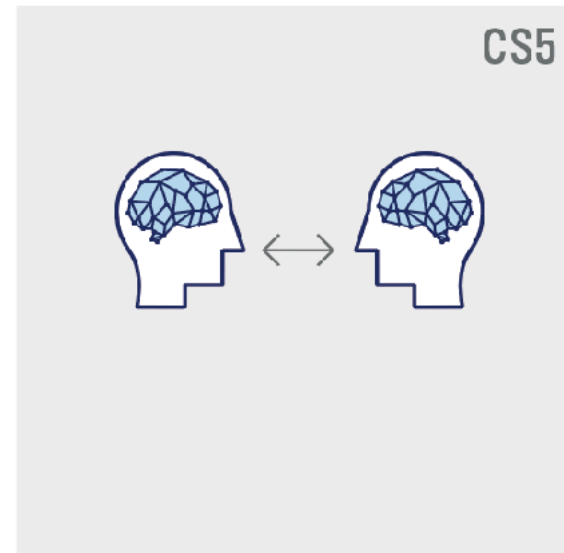


# Causality impacts both research and knowledge transfer

Next-generation  
projections



Bidirectional knowledge  
transfer



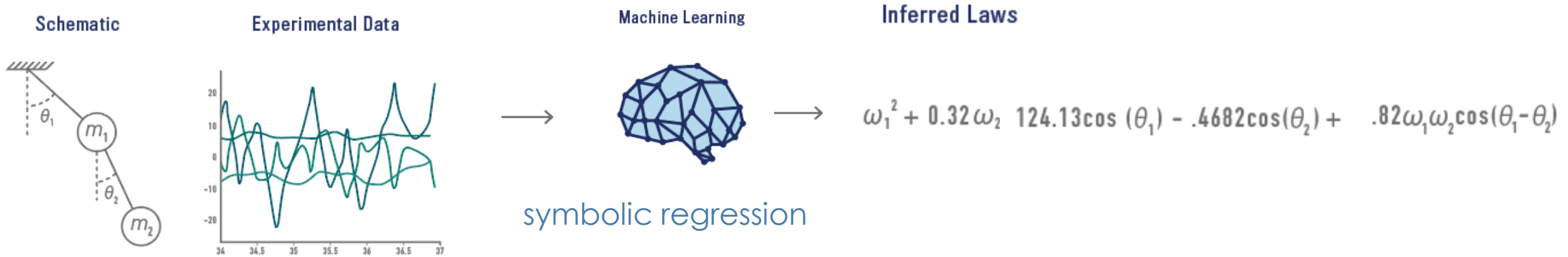
# What's missing in machine learning?



1. Generalization
2. Causal reasoning
3. Interpretability

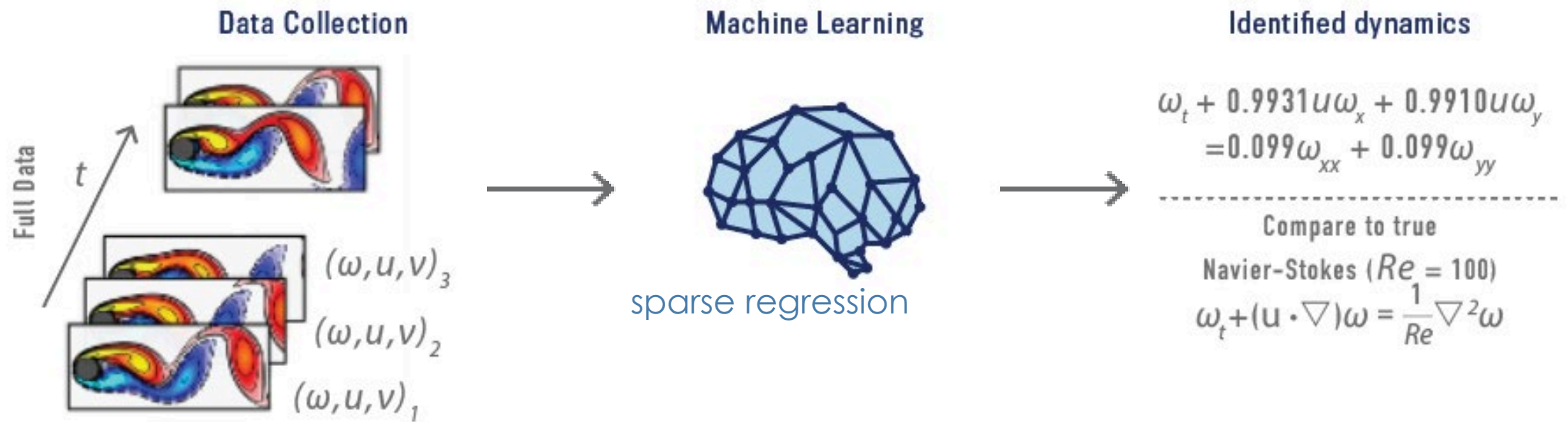
# Learning governing equations from data

- Conservations, symmetries & invariants



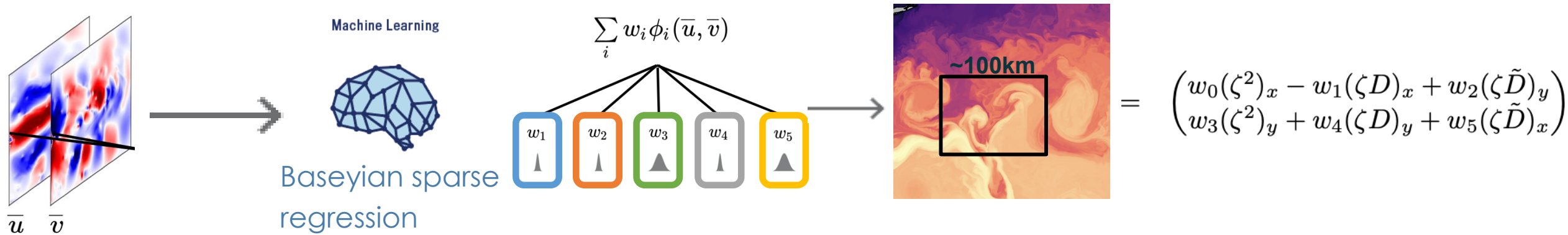
# Learning governing equations from data

- Conservations, symmetries & invariants
- Learning complex time-evolution of known systems

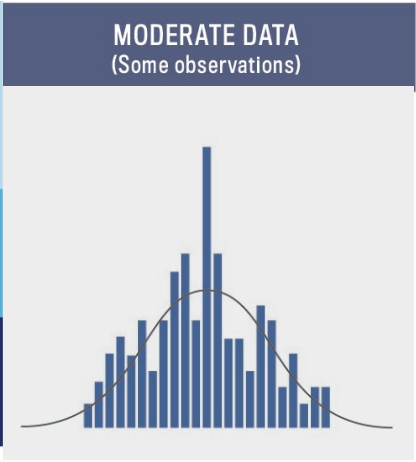
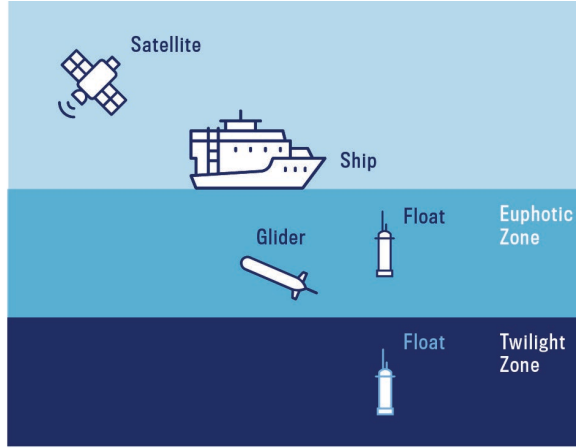


# Learning governing equations from data

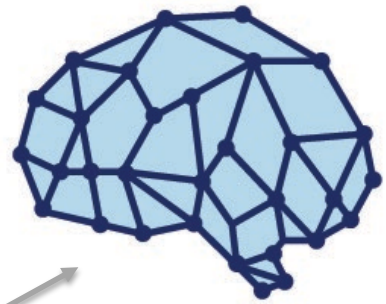
- Conservations, symmetries & invariants
- Learning complex time-evolution of known systems
- Unknown processes



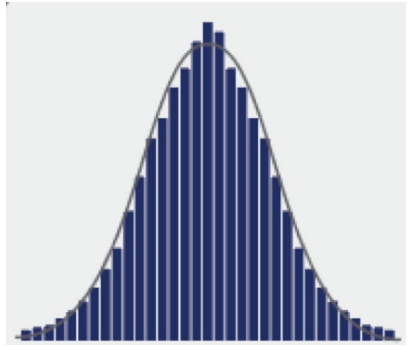
# LEAP will discover equations governing the biological pump



Machine Learning

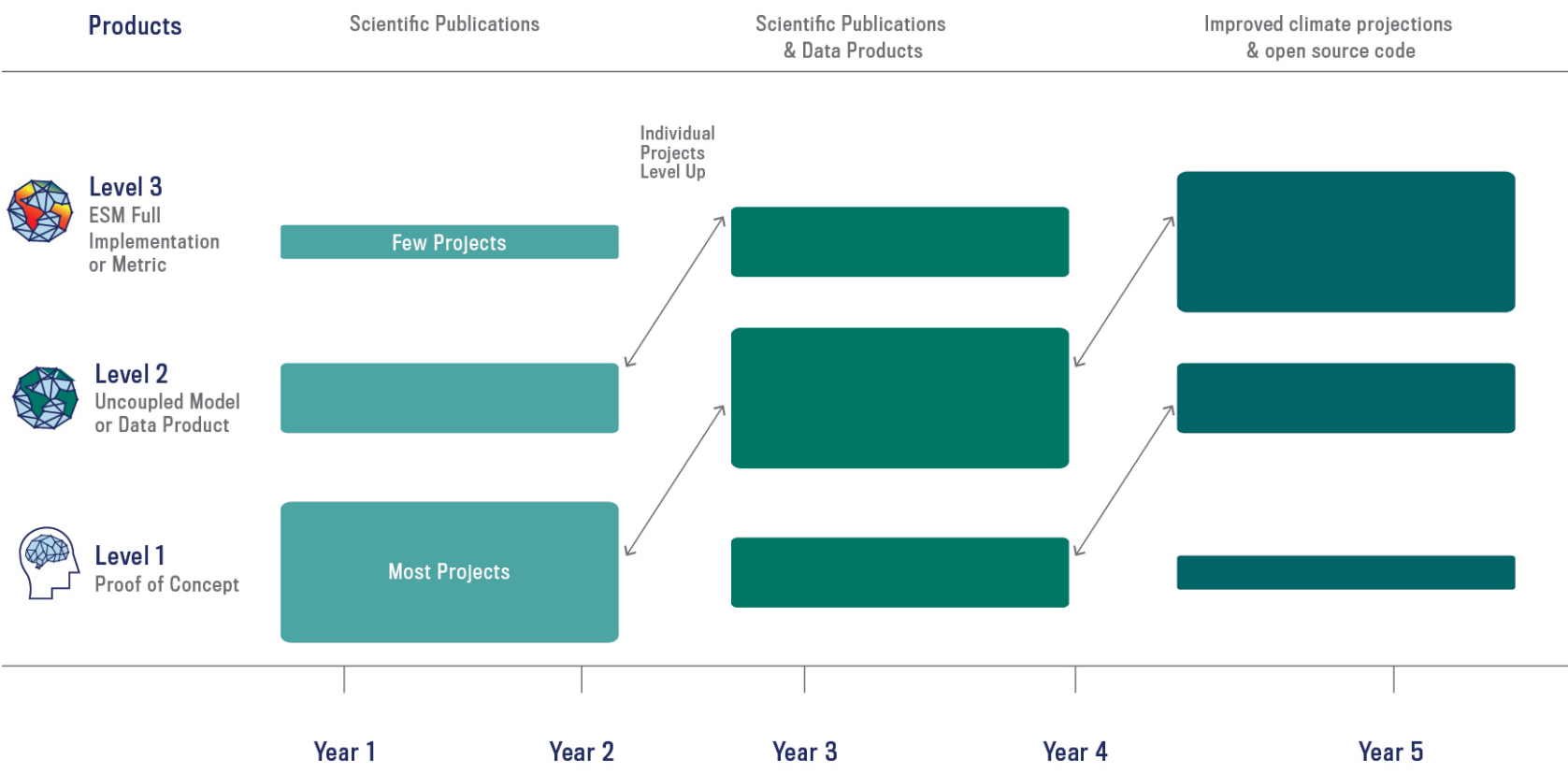
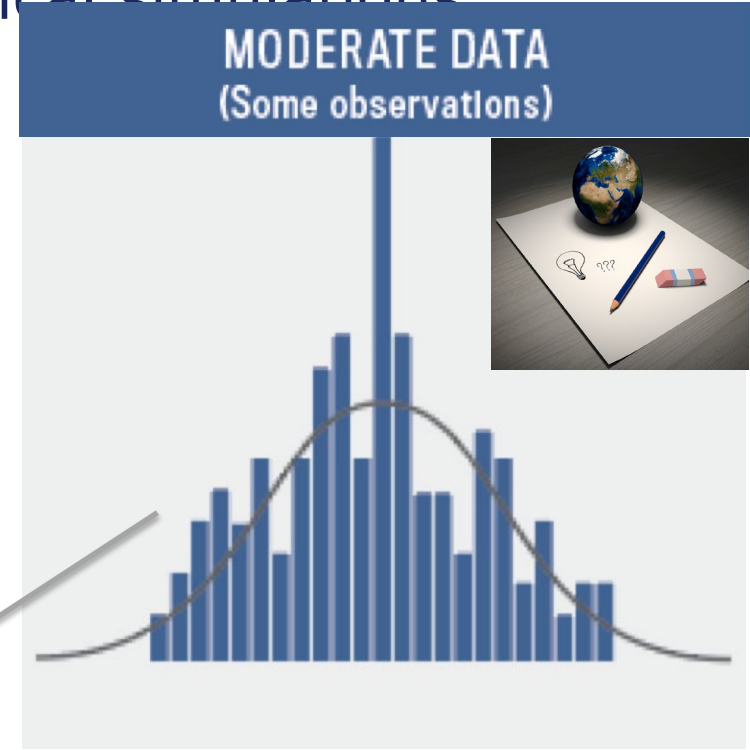


Equation for biological pump



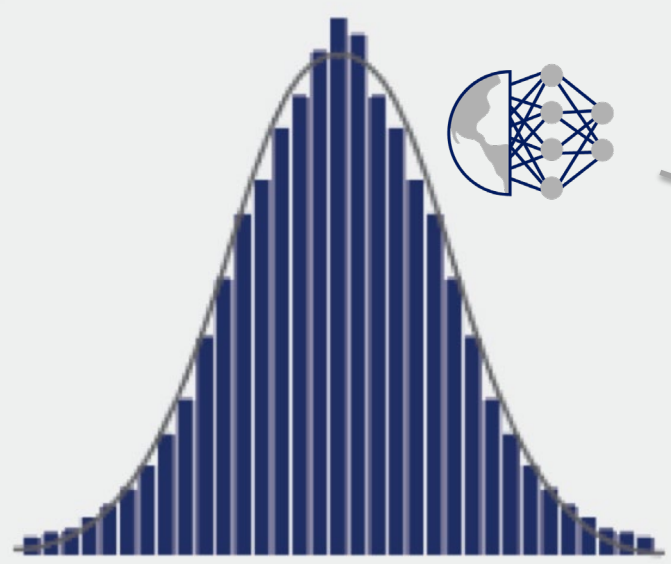
Transfer learning

# LEAP will advance simultaneously machine learning algorithms, our understanding of Earth system processes & their representation in numerical simulations

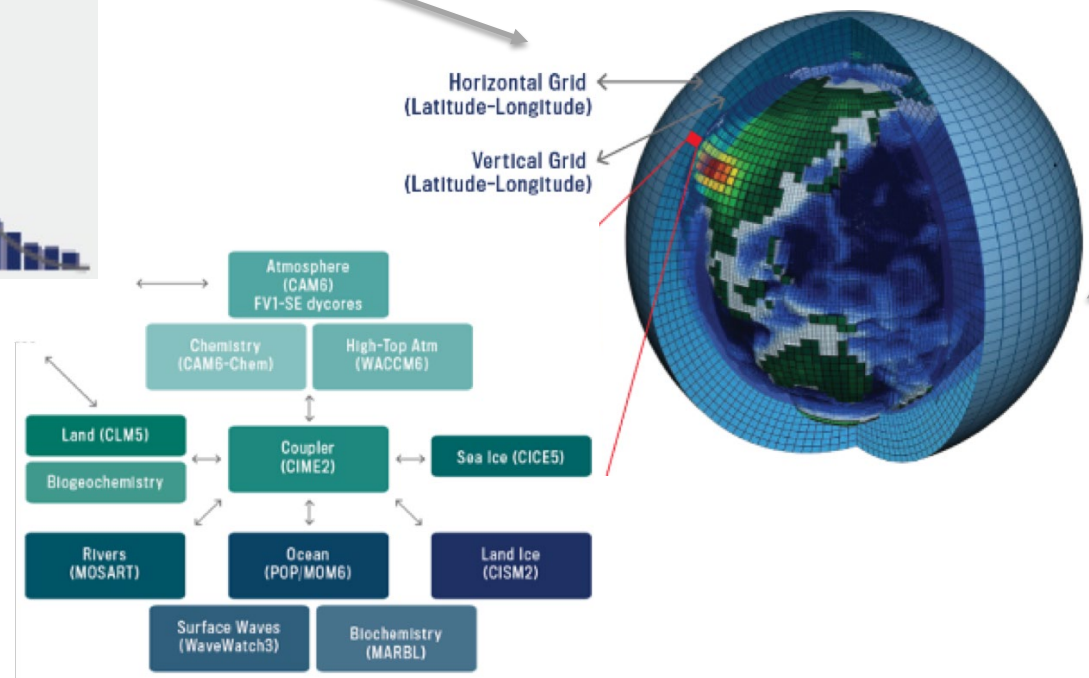
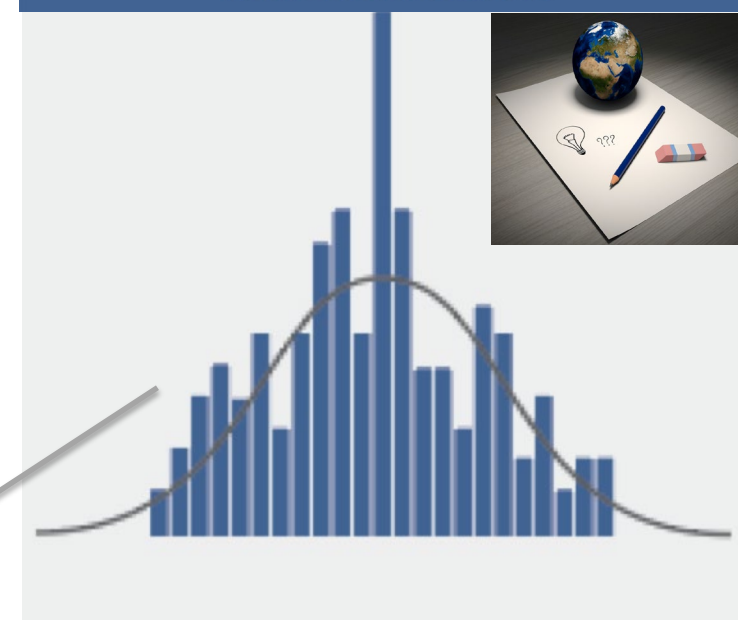


# LEAP will advance simultaneously machine learning algorithms, our understanding of Earth system processes & their representation in numerical simulations

**DATA RICH**  
(Many observations + high-res. simulations)



**MODERATE DATA**  
(Some observations)



# Questions?

Dr. Laure Zanna, Professor at NYU  
Dr. Carl Vondrick, Asst. Professor at Columbia  
Research Co-Directors



# Data Poor Research



Deputy Director  
**Galen McKinley**

## Professional Experience + Education

- Professor of Earth + Environmental Sciences, Columbia
  - Previously Professor of Atmospheric + Oceanic Science, Wisconsin
- PhD, Climate Physics + Chemistry, MIT

## Expertise

- Ocean carbon cycle, feedbacks with climate
- Ocean and climate modeling
- Data science

## Synergistic Activities

- The Oceanography Society, Councilor, 2020-present
- JASON Advisory Group, 2015-present
- Diversity Committee chair, EES department, 2019-present
- Co-chair for undergrad education, Wisconsin AOS, 2010-15





**Chief Convergence Officer  
& Education Director**  
**Tian Zheng**

- Professor + Chair of Statistics, Columbia
- PhD, Statistics, Columbia

### **Expertise**

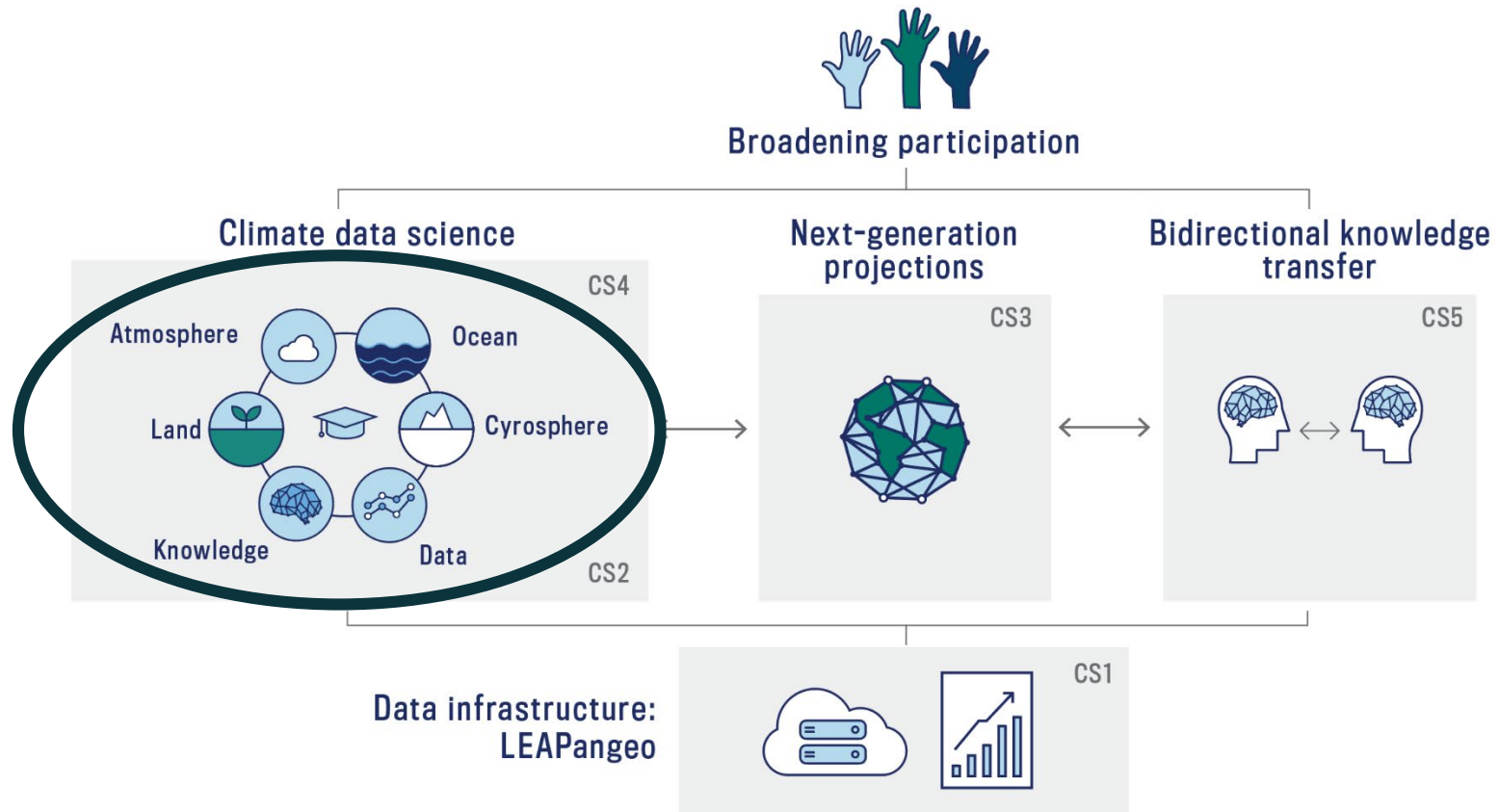
- Spatiotemporal modeling
- Statistical machine learning
- Data science education
- Management

### **Synergistic Experience**

- Associate Director for Education, Data Science Institute (DSI), 2017-2020
- Creator, DSI Scholars Program, 2018-2020
- Leadership Team Member, Collaboratory Program
- Faculty Advisor, Columbia Statistics Club 2016-2020

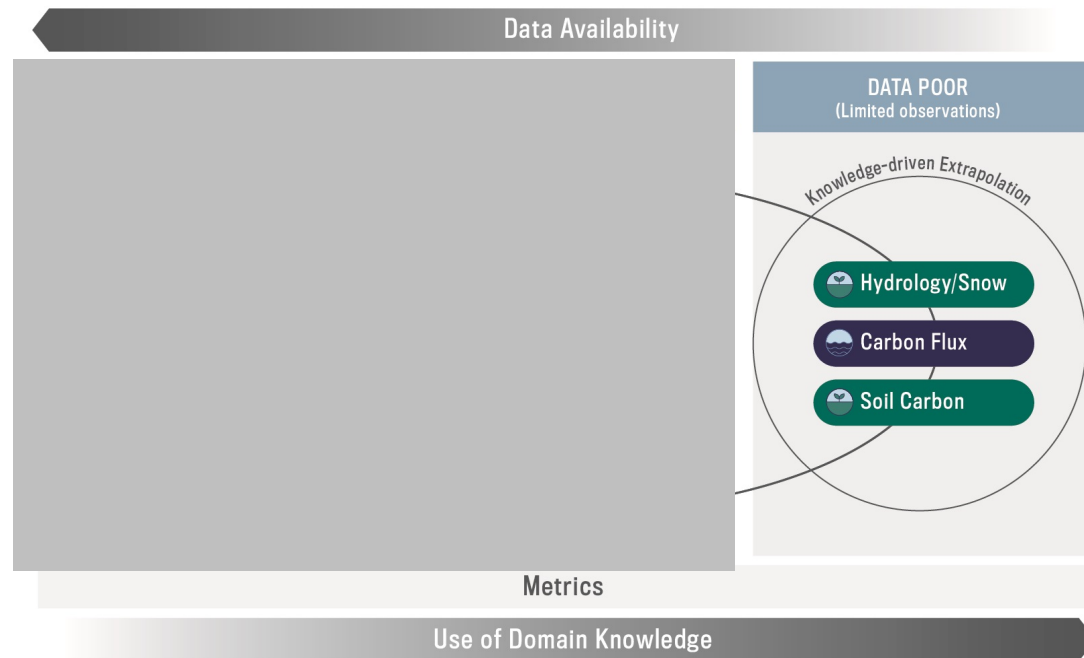


# Implementing our vision



*LEAP forward in the **reliability**, utility, and reach of climate projections through **synergistic innovations in data science and climate science**.*

# Understanding complex systems is often complicated by a lack of data: “Data Poor”



## Physics Guided ML



Hydrology and Snow

## ML-based data products w/ uncertainty



Air-sea CO<sub>2</sub> Flux



Soil Carbon



# Understanding complex systems is often complicated by a lack of data: "Data Poor"

Inland lake – PGML Example



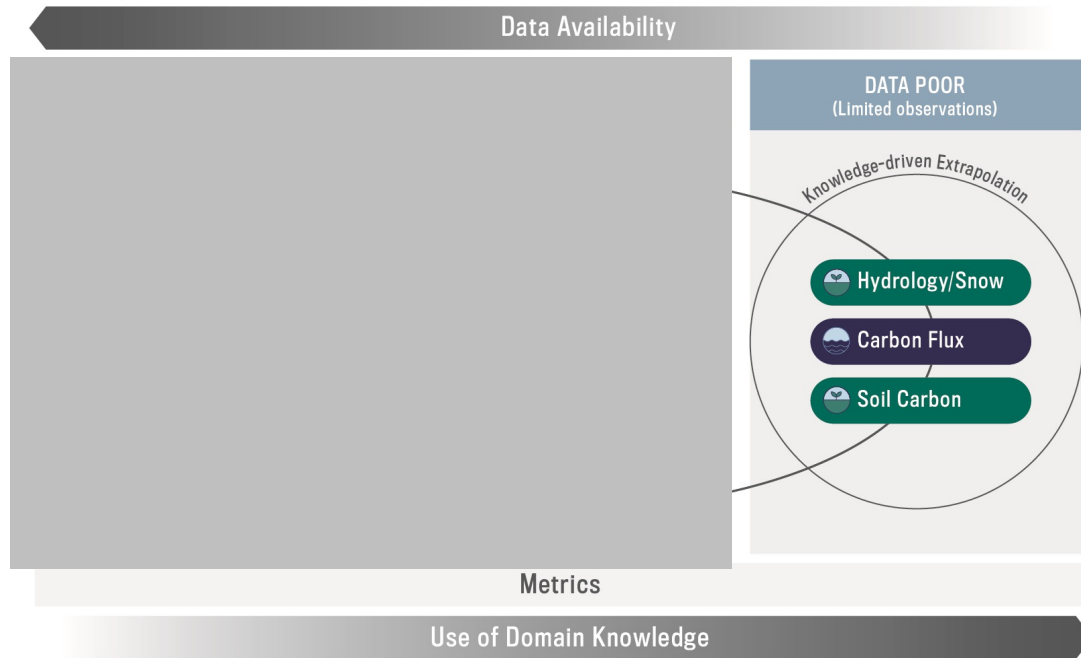
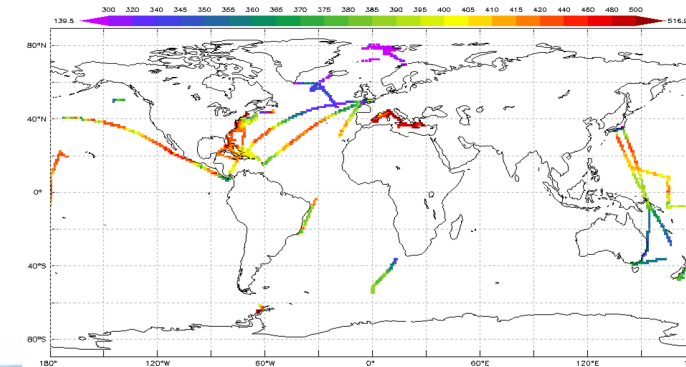
## Physics Guided ML

 Hydrology and Snow

## ML-based data products w/ uncertainty

 Air-sea CO<sub>2</sub> Flux

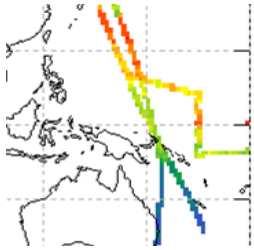
 Soil Carbon



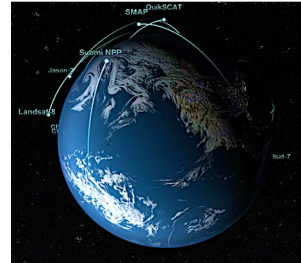
# Challenges



# Opportunities



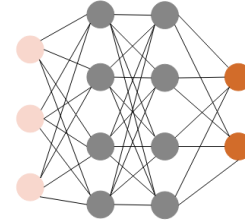
Sparse Data



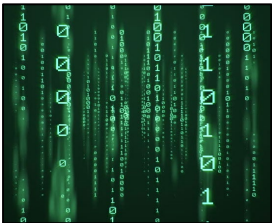
Advanced observing systems and new infrastructures to link datasets



Complex Processes



Advances in machine learning  
Deep Learning, Physics-Guided ML,  
Modeling of Functional Objects



Computational Intensity

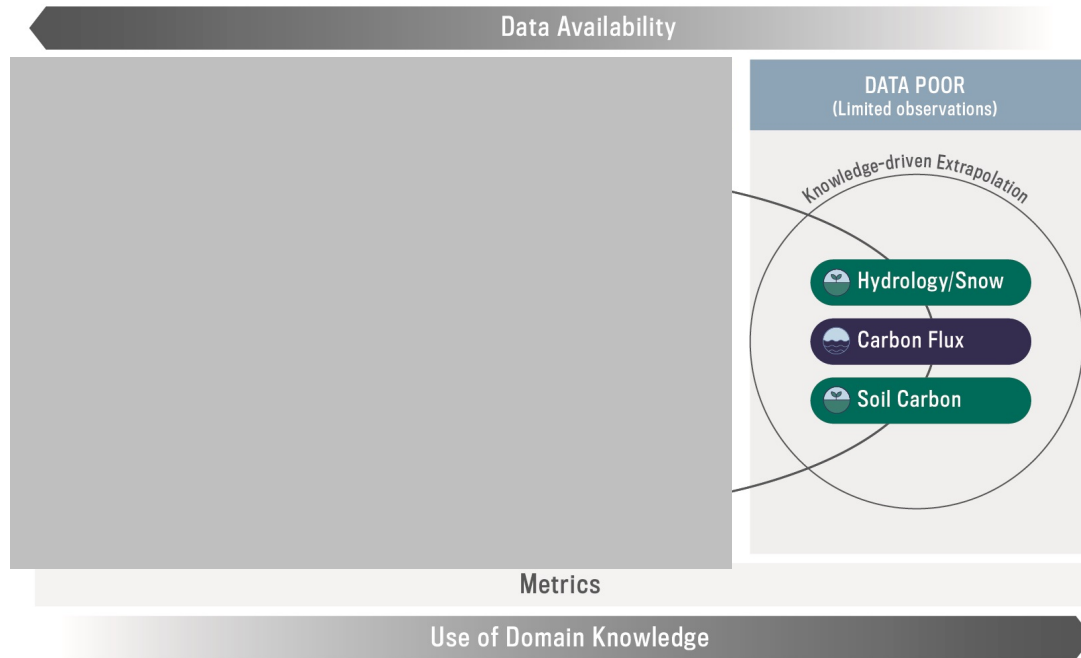


Advances in optimization  
and cloud computing



# Understanding complex systems is often complicated by a lack of data: “Data Poor”

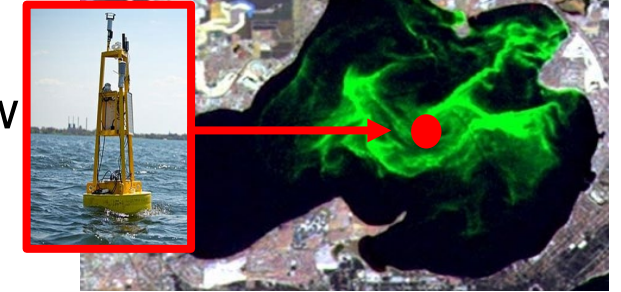
Inland lake – PGML Example



## Physics Guided ML



Hydrology and Snow



## ML-based data products w/ uncertainty



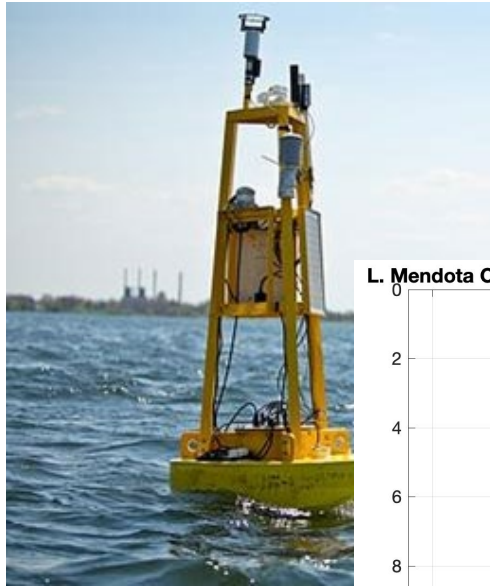
Air-sea CO<sub>2</sub> Flux



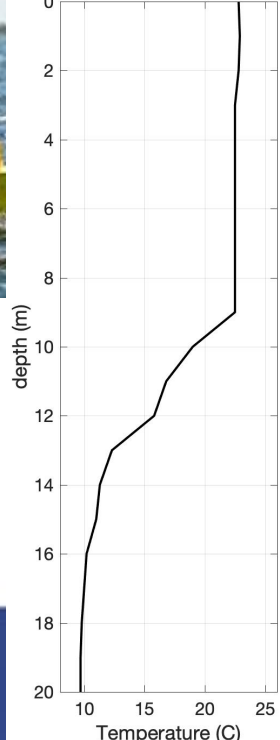
Soil Carbon

# Status Quo: Process model to predict lake 1D temperatures

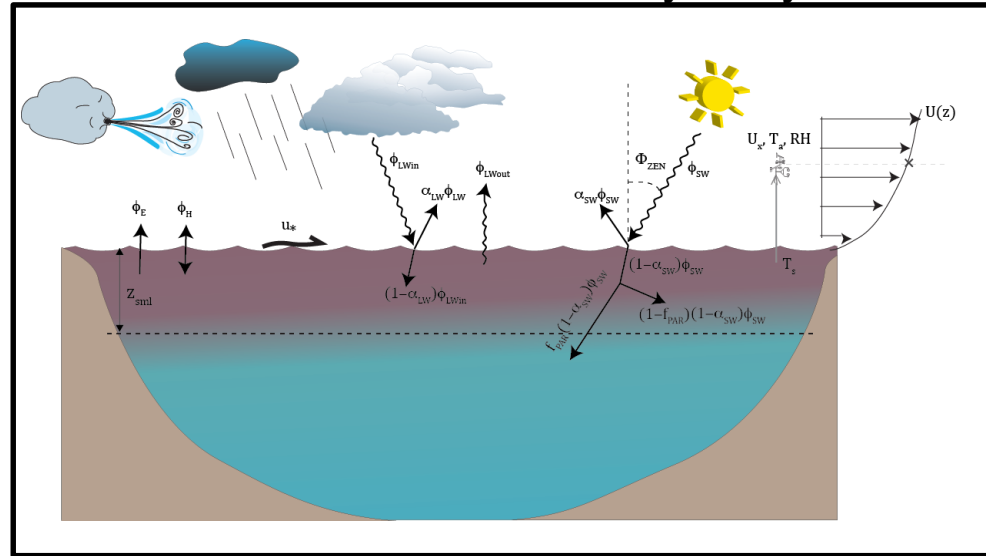
Lake Mendota, WI



L. Mendota Center, 9 Sept 2014

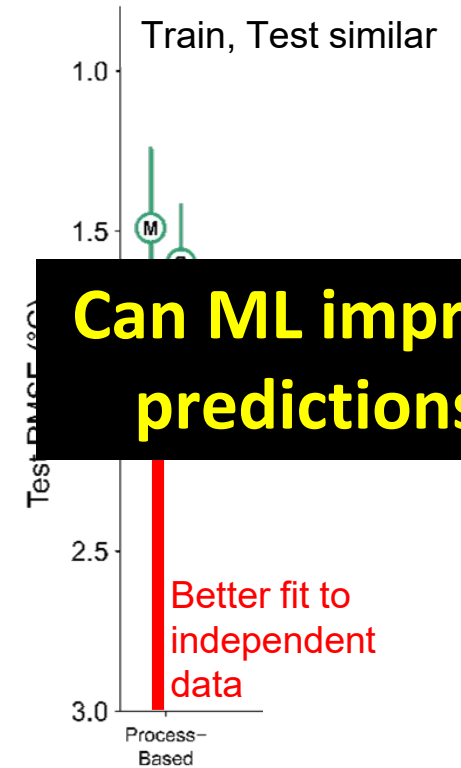


Process Model for 1D lake hydrodynamics



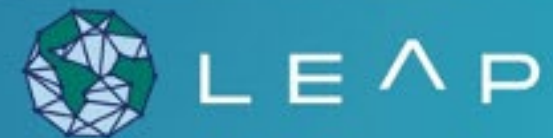
## Results

Train, Test similar



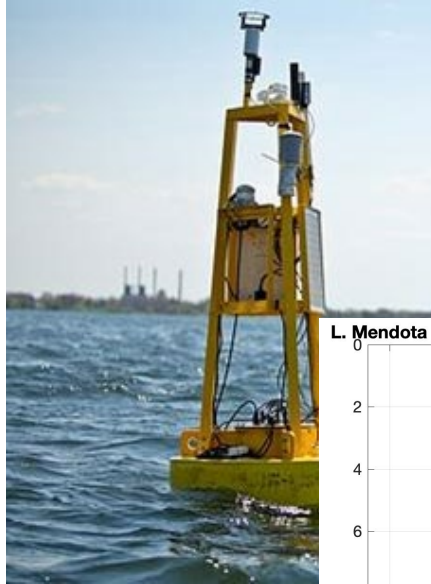
**Can ML improve predictions?**

Read, [Kumar et al. 2019, Water Resour. Res.](#)

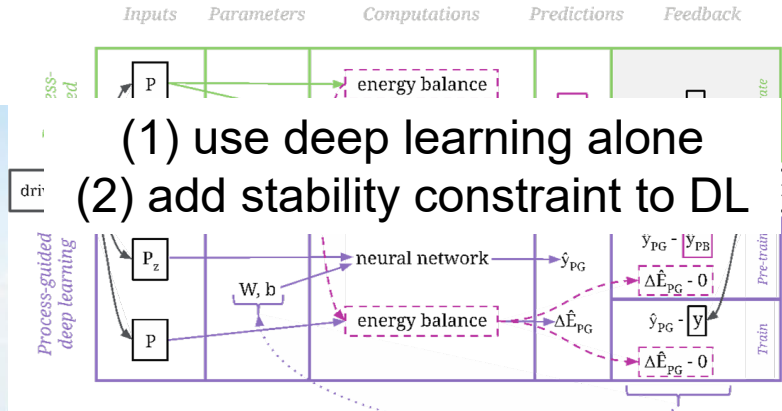
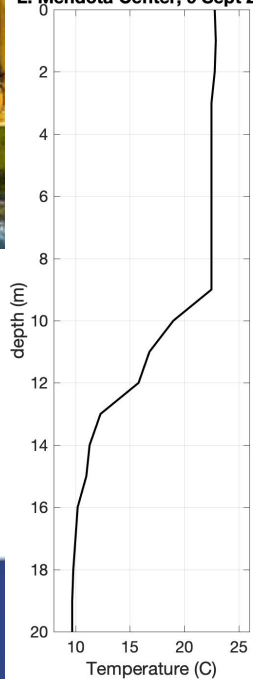


# Physics-Guided ML blends process knowledge and deep learning for more robust predictions

Lake Mendota, WI

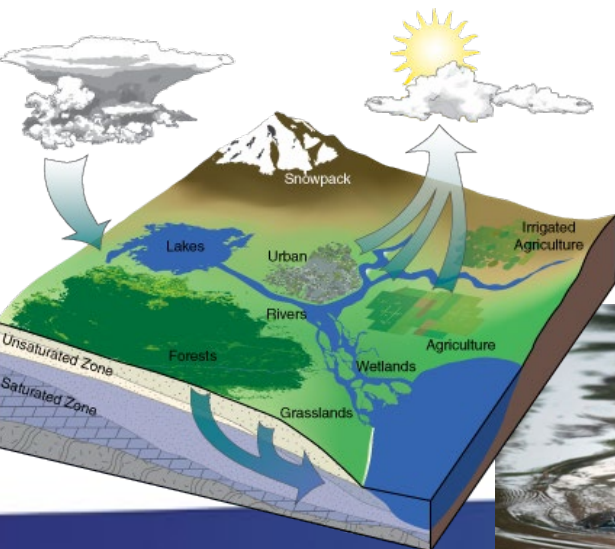
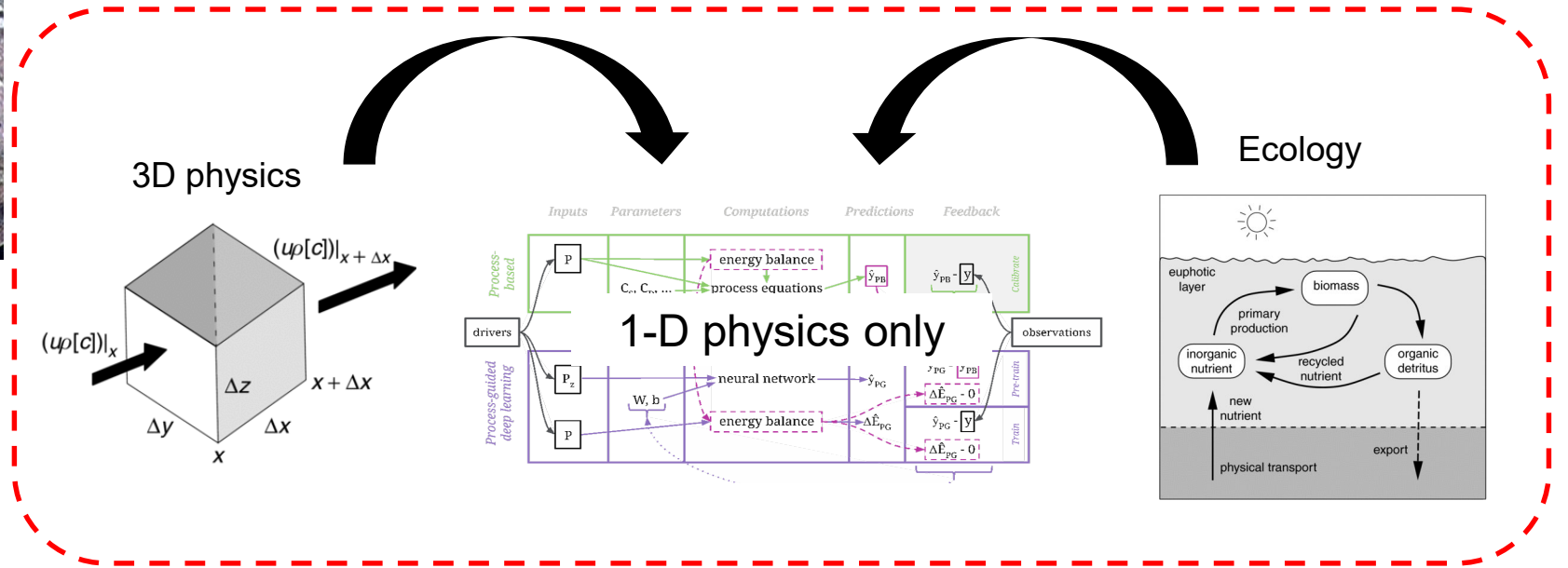
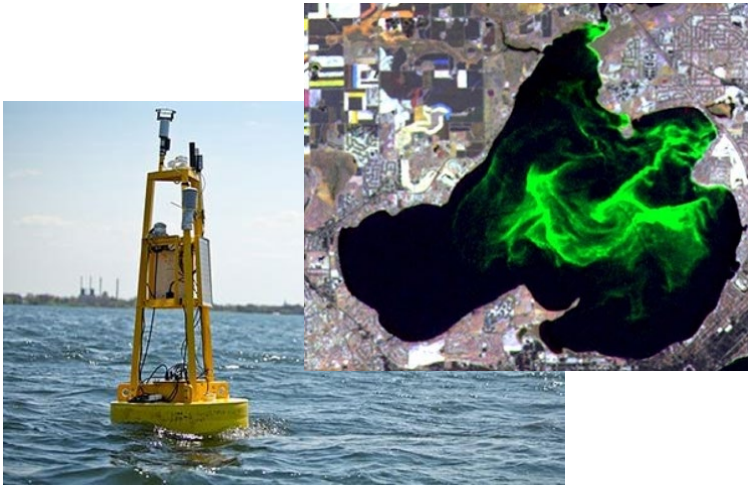


L. Mendota Center, 9 Sept 2014



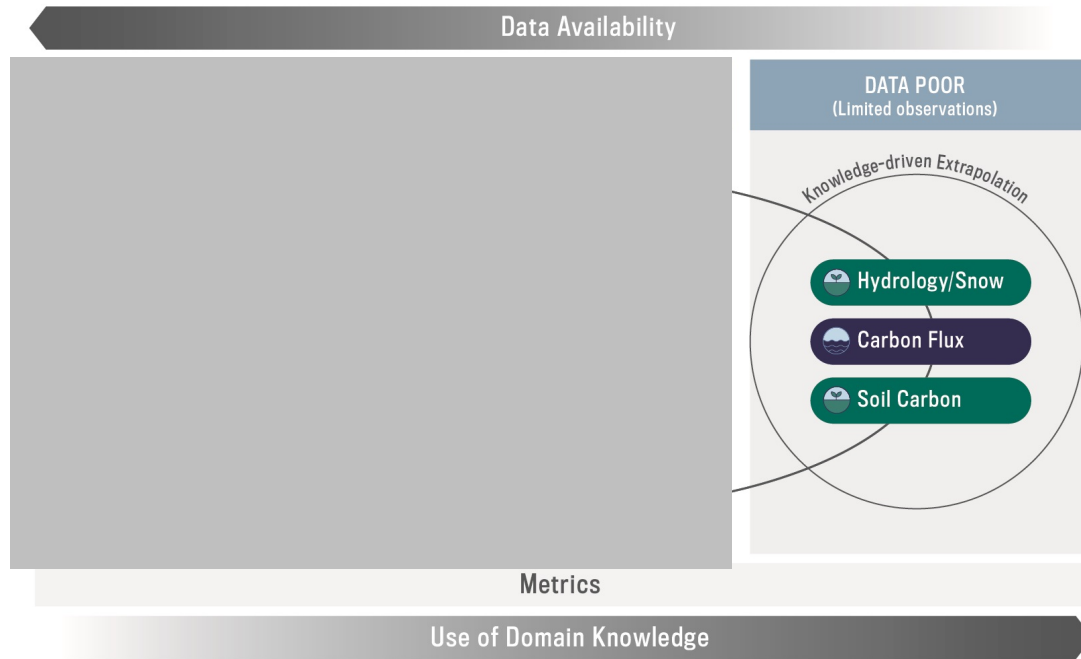
Read, [Kumar et al. 2019](#), *Water Resour. Res.*

# Next step: extend Physics-Guided ML to multiple interacting processes and spatial scales



Global hydrology: Yang, [Gentine](#) et al. 2019, *ERL*

# Understanding complex systems is often complicated by a lack of data: “Data Poor”



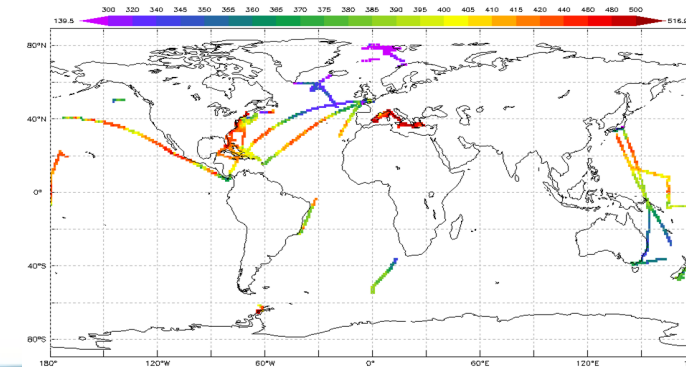
## Physics Guided ML

 Hydrology and Snow

## ML-based data products w/ uncertainty

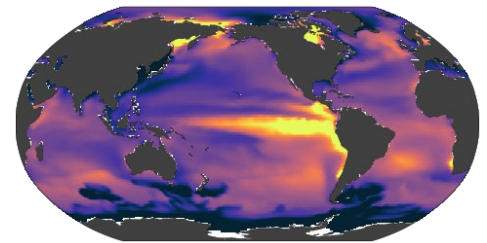
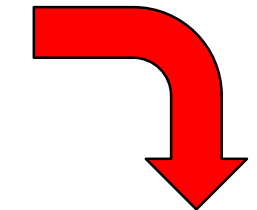
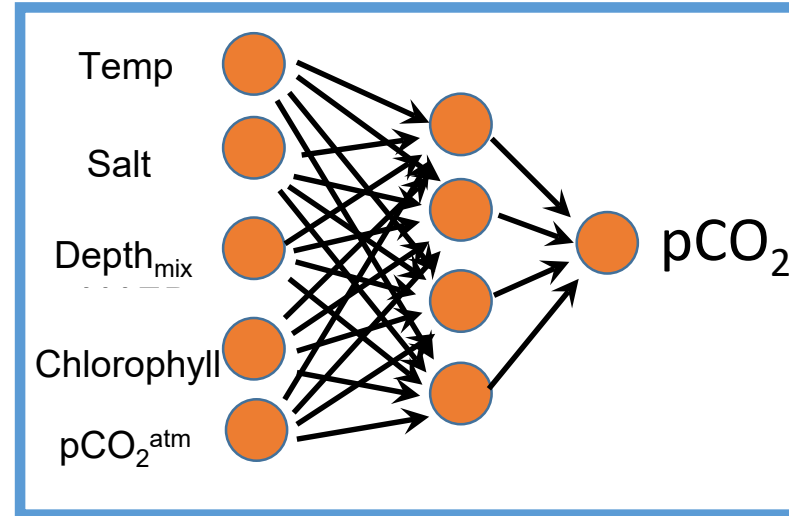
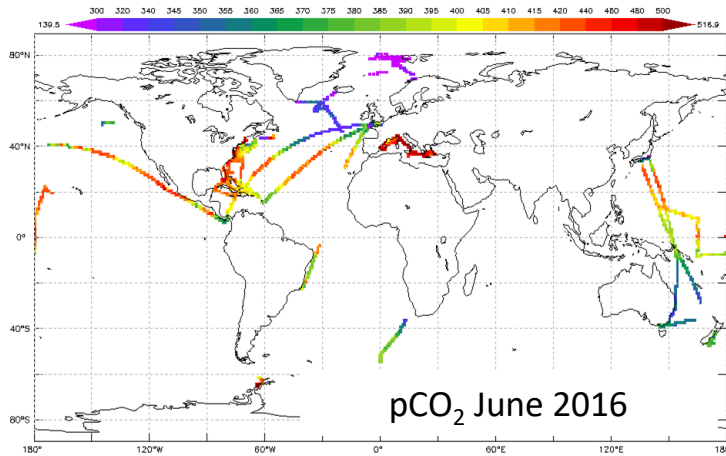
 Air-sea CO<sub>2</sub> Flux

 Soil Carbon



# Status Quo: Apply existing ML to extrapolate

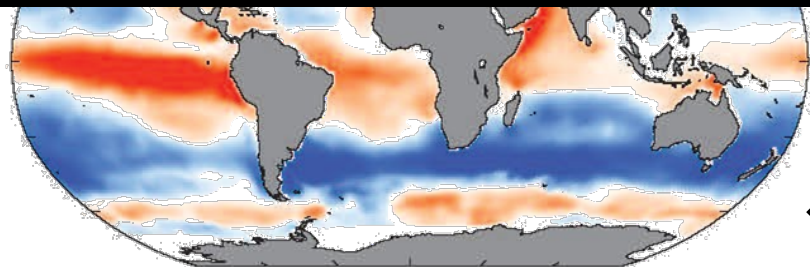
1. Train an off-the-shelf Neural Network



2. Predict pCO<sub>2</sub> everywhere

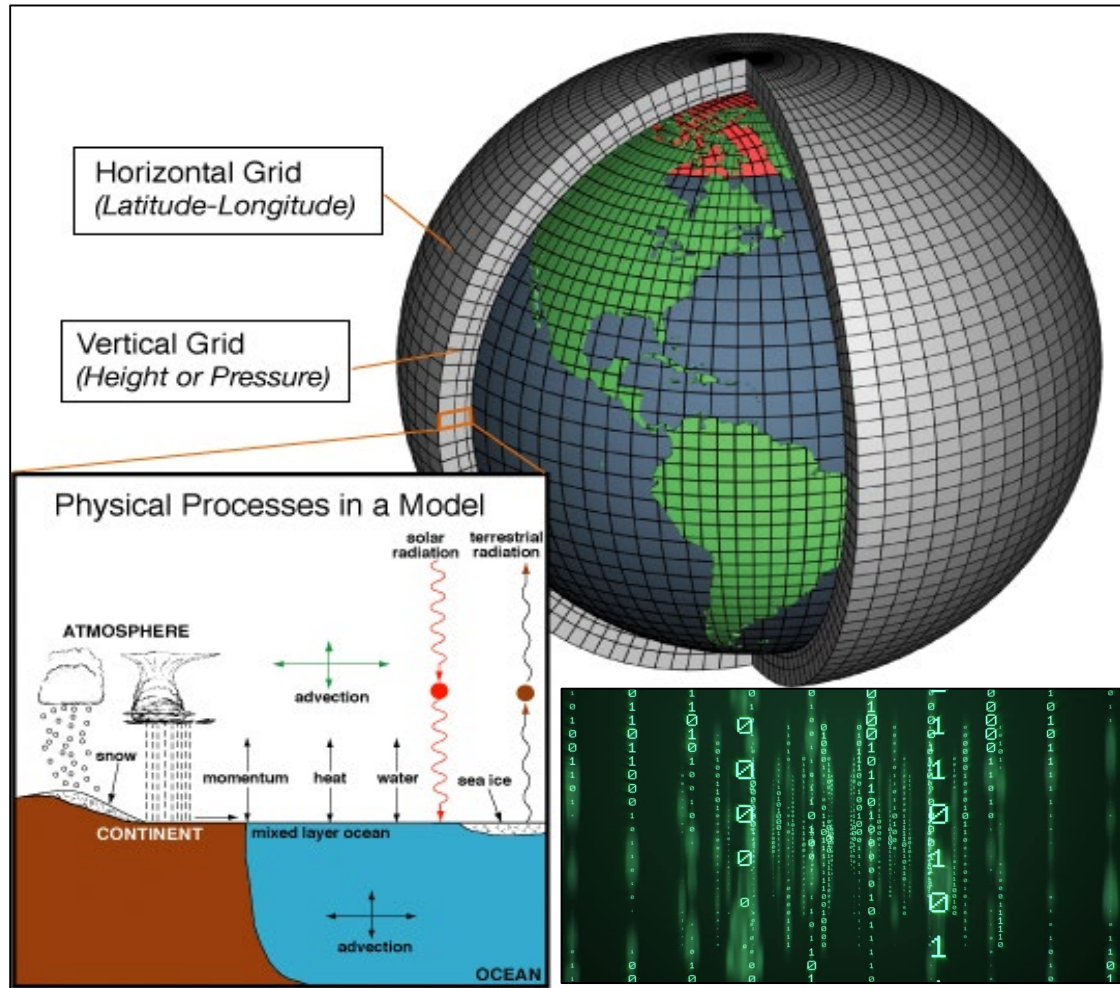
3. Calculate CO<sub>2</sub> flux

**How are the uncertainties structured?**

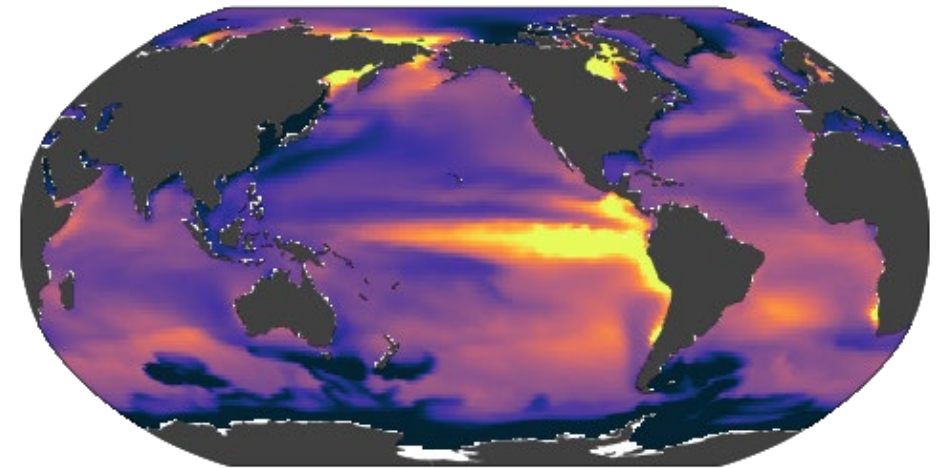


Landschutzer et al. (2014), GBC

# Earth System Models provide ample simulated data with which ML uncertainties can be evaluated

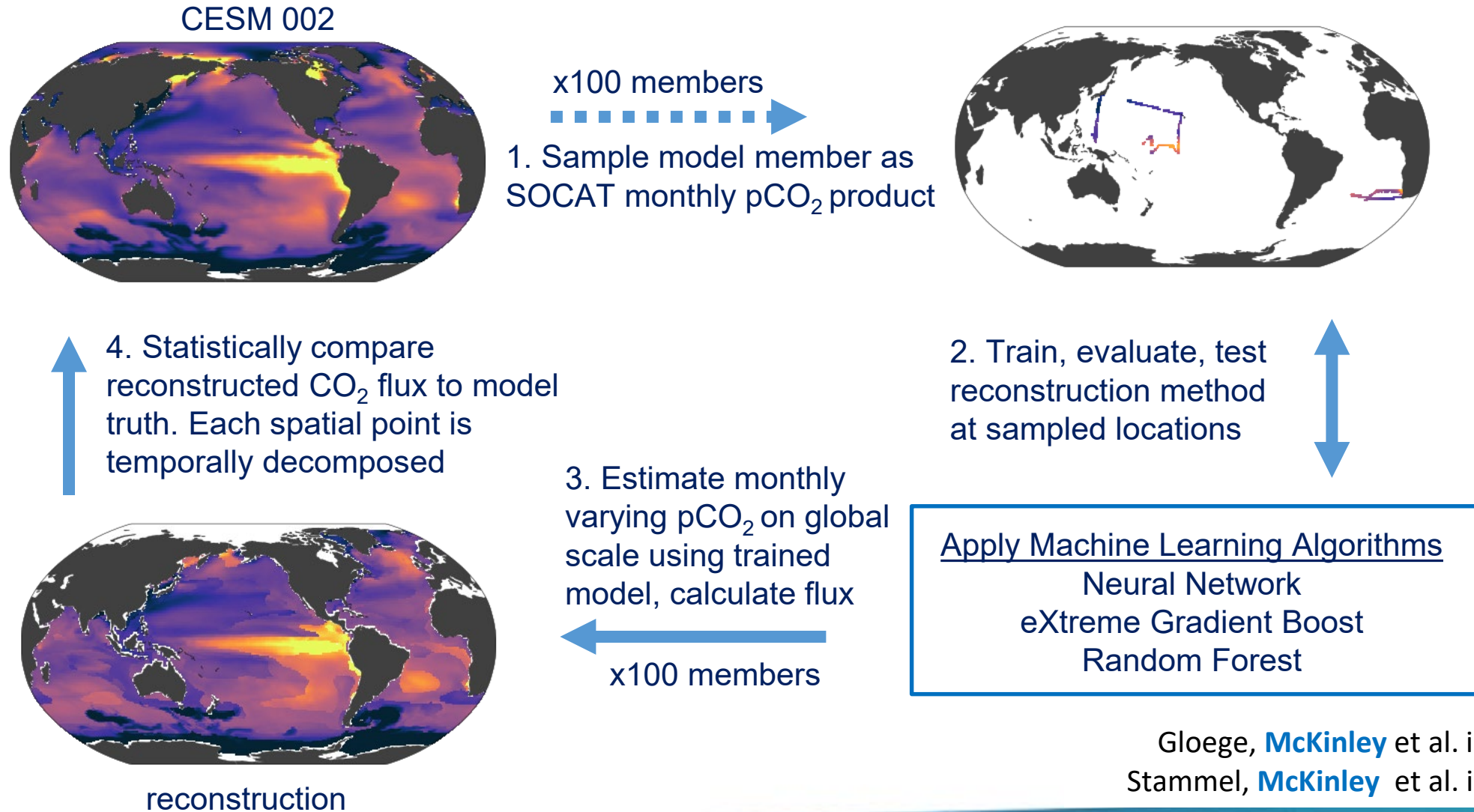


*Ocean pCO<sub>2</sub> model output*



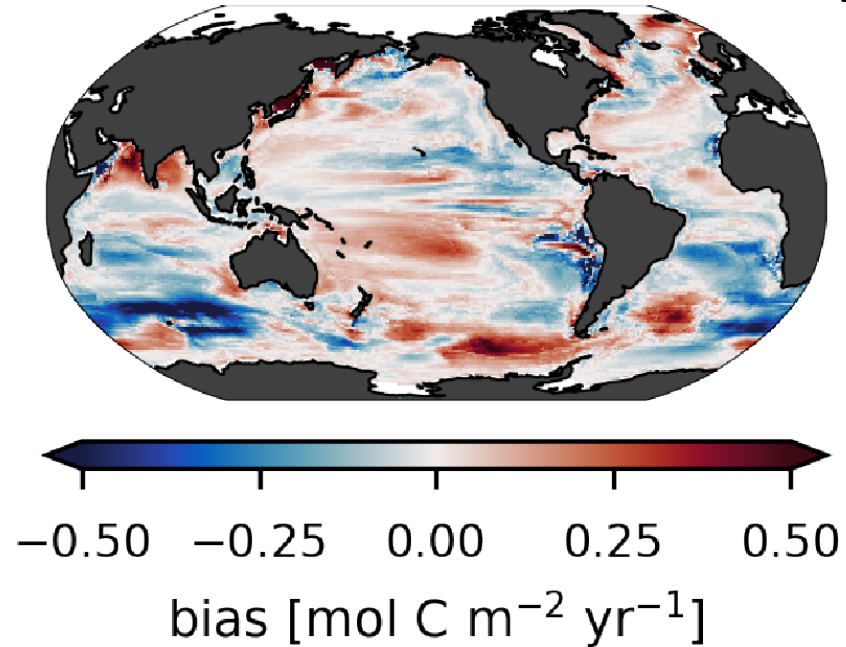
- 4 Large Ensembles (CESM, CanESM, GFDL, MPI)
- 25 members x 4 = 100 members
- 1985-2016

# Using ESMs to test skill of ML extrapolation



# Neural Network bias is small globally, but large where data are most sparse

Global mean bias =  $-0.01 \text{ molC m}^{-2} \text{ y}$

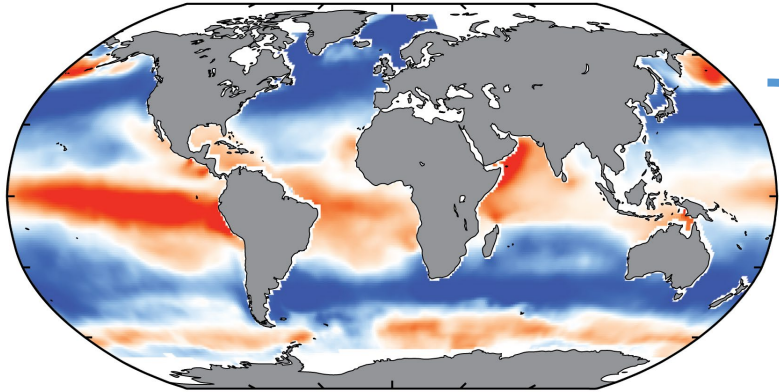


Mean over 100-members and 1985-2016

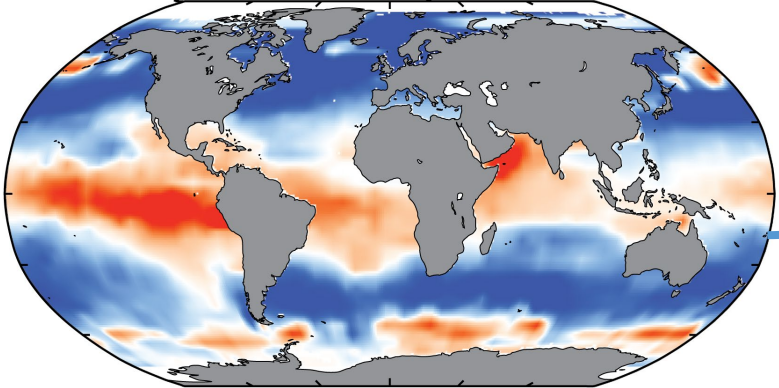
Gloege, [McKinley](#) et al. in review

# LEAP will develop approaches to quantify and reduce extrapolation uncertainties

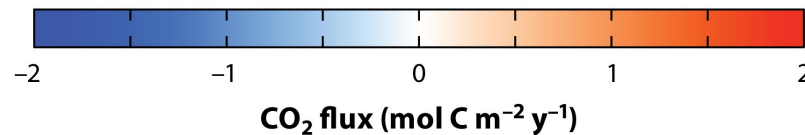
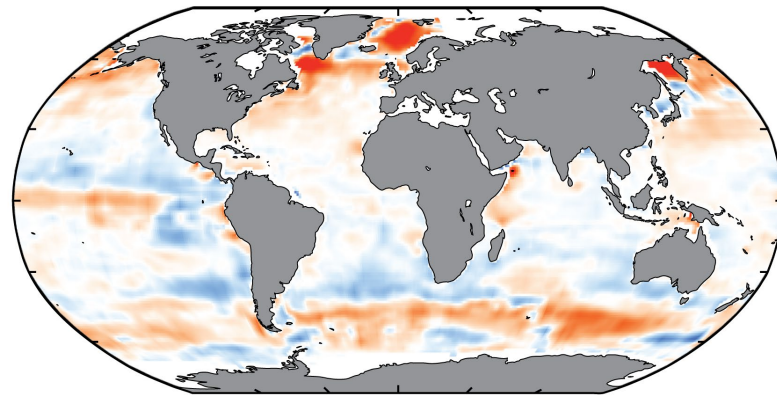
ML-based data product



possible ESM prediction



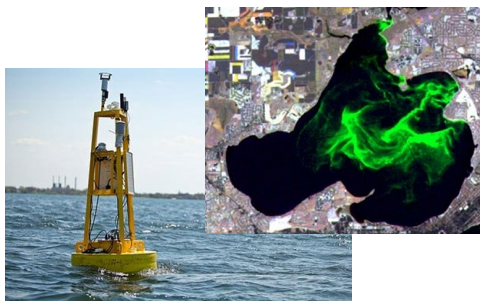
Data product - ESM



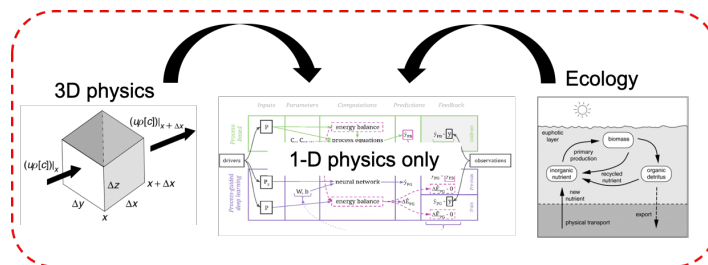
Is this difference meaningful given the uncertainty of the ML-based extrapolation?

# LEAP will push the limits of Machine Learning

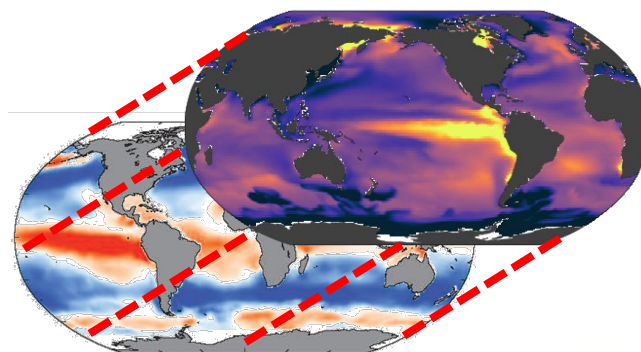
More Types  
of “Data”



More Physics as  
Constraints



Uncertainty  
Quantification



Innovative workflows based on  
ML building blocks

Scalable algorithms for  
applications of real-world scale

Functional model evaluation for  
identifying new ML frontiers





# Questions or Feedback?

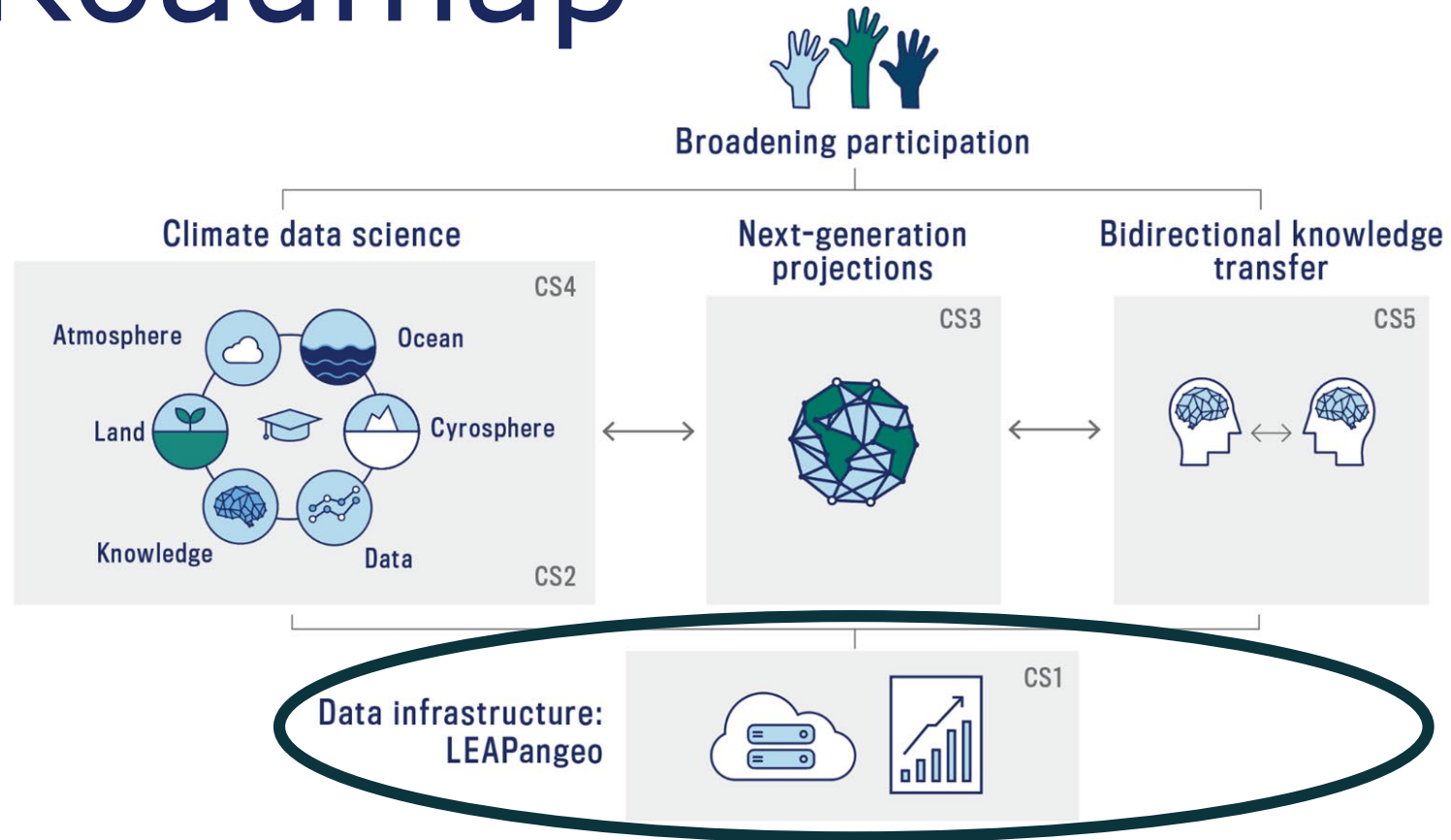
**Galen McKinley**, Deputy Director  
**Tian Zheng**, Chief Convergence Officer + Education Director



# *LEAP*angeo: Shared Experimental Facilities

Dr. Ryan Abernathey  
Director of Data and Computing  
Associate Professor of Earth and  
Environmental Sciences

# Our Roadmap





Data & Computation  
Director  
**Ryan Abernathy**

- Associate Professor of Earth + Environmental Sciences, Columbia
- PhD, Climate Physics + Chemistry, MIT

### **Expertise**

- Ocean transport and mixing
- Ocean modeling
- Data analysis, visualization, cloud computing

### **Notable Experience**

- NASA SWOT Science Team
- Lead PI, Pangeo (NSF EarthCube Project)



# My Presentation Will Explain:

- Why LEAPangeo—a shared data and computing platform—is needed to achieve LEAP's goals
- Architecture, development, and management of the platform
- Integration of this platform with research, education, and knowledge transfer



# Why is climate data hard?

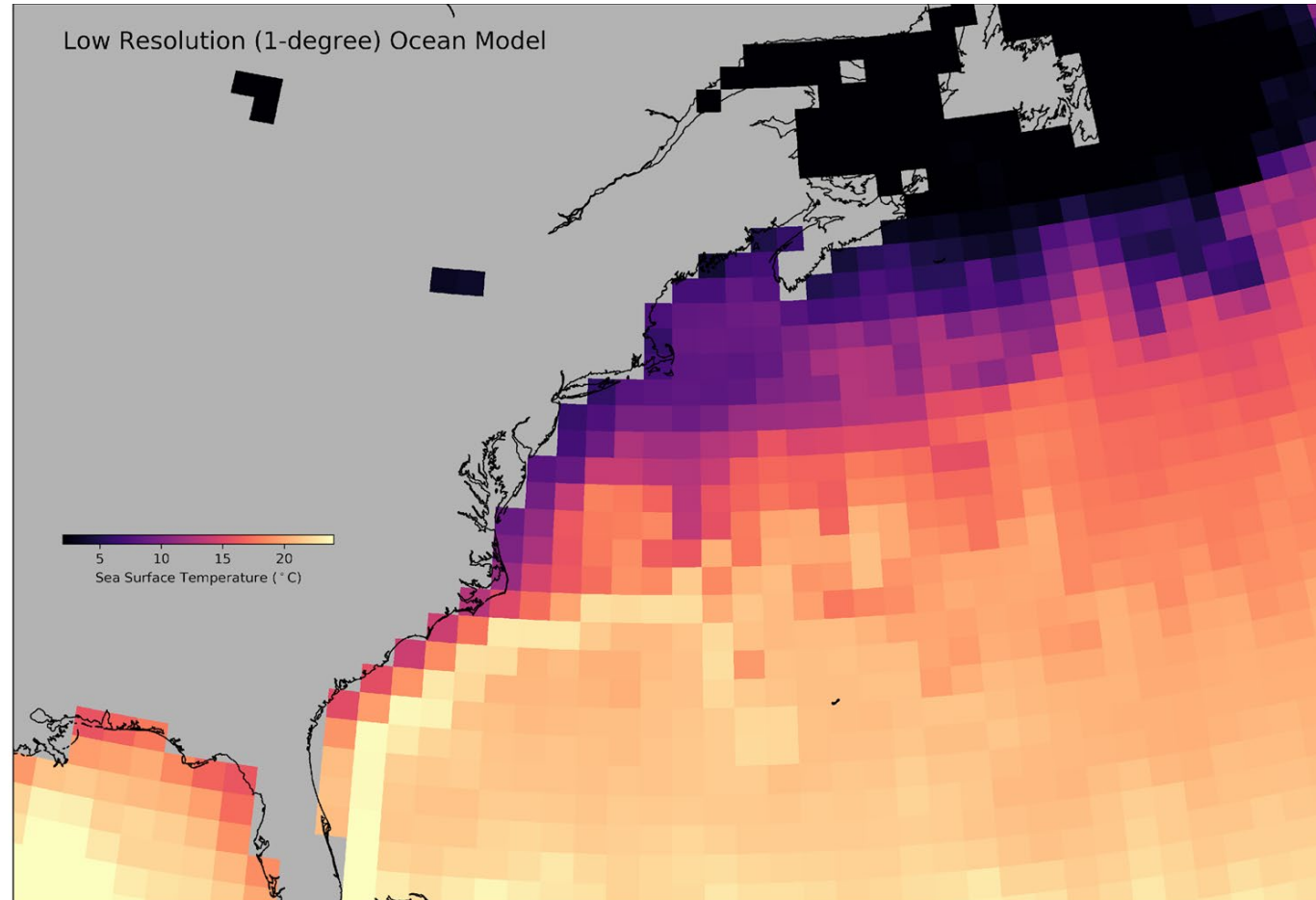
**Size**

**Complexity**

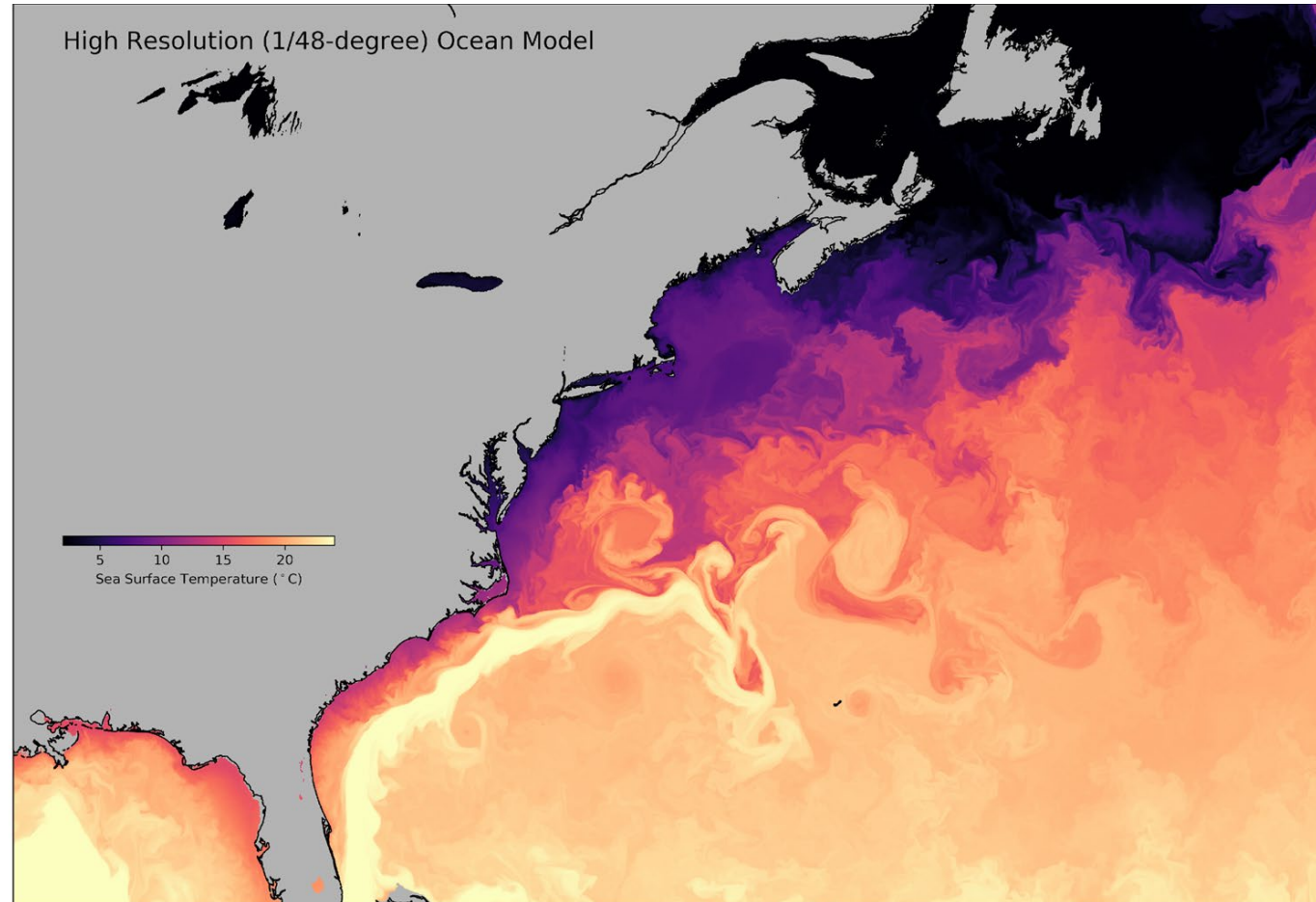
**Geometry**



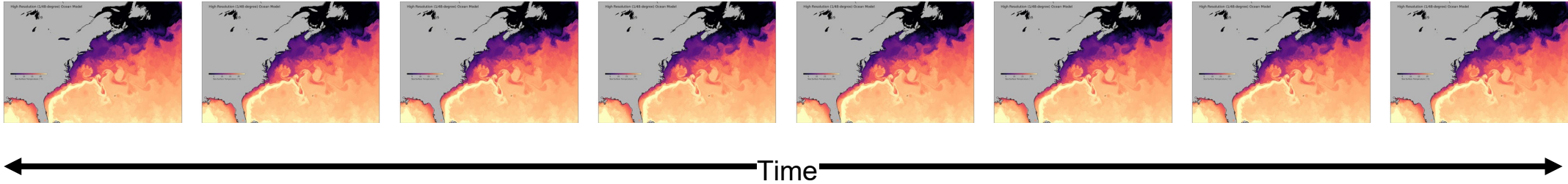
# Why is climate data hard?



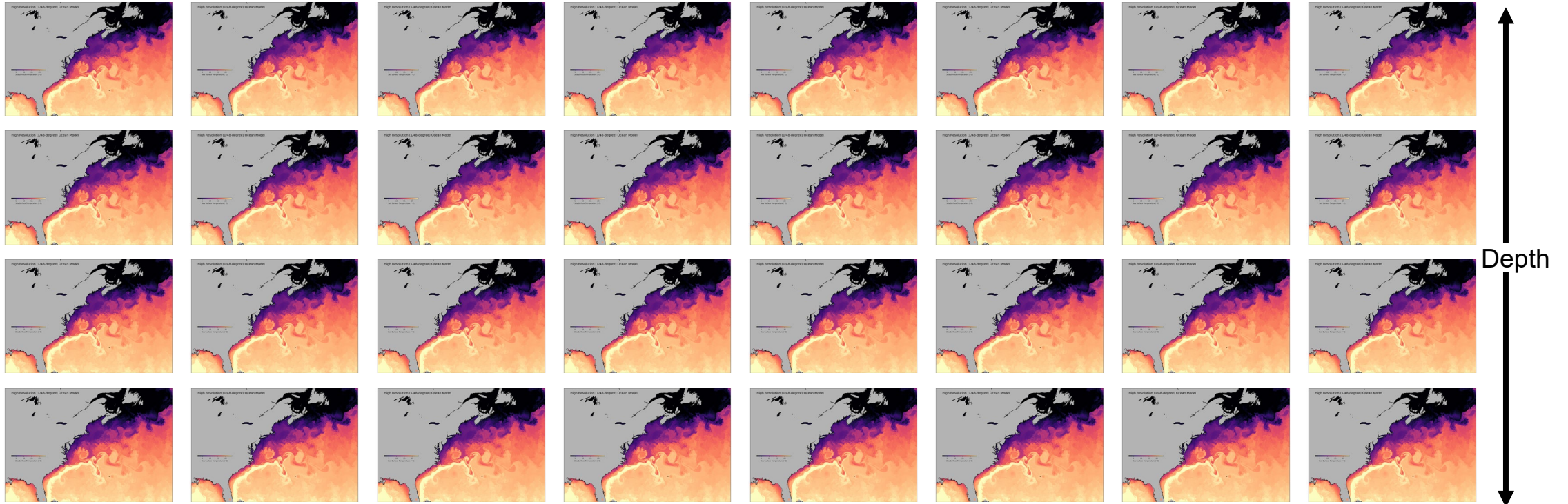
# Why is climate data hard?



# Why is climate data hard?



# Why is climate data hard?



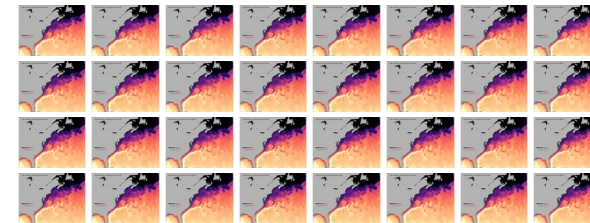
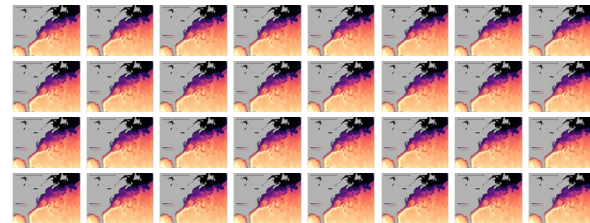
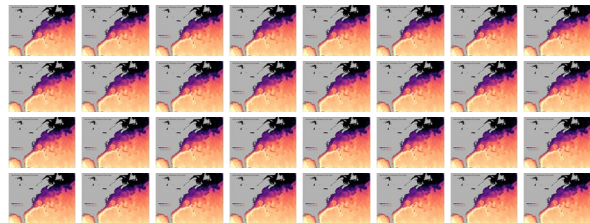
**Size**

# Why is climate data hard?

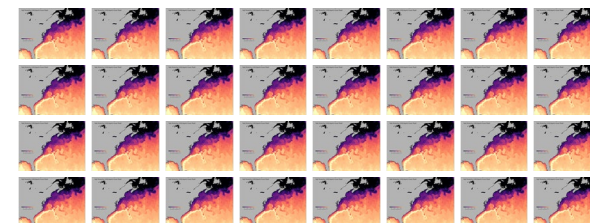
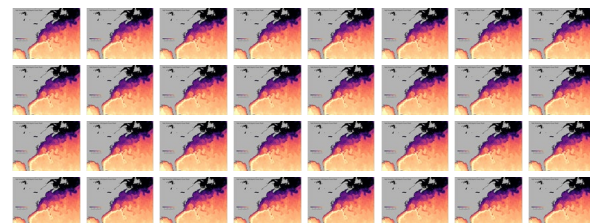
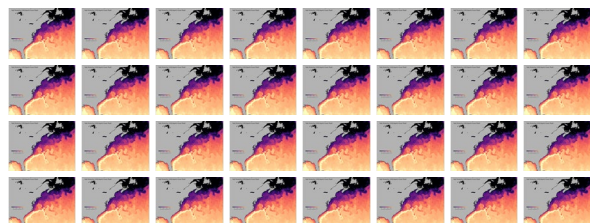
Scenario A

Scenario B

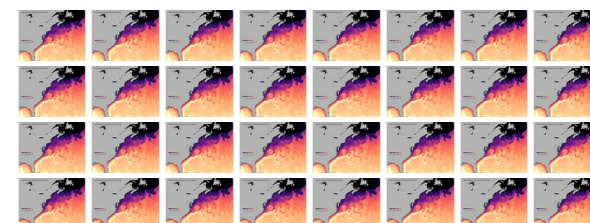
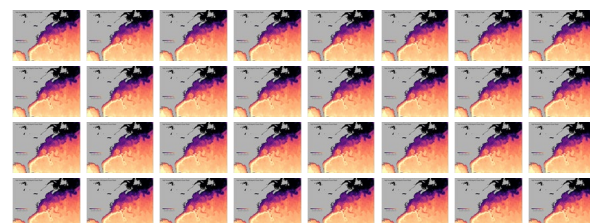
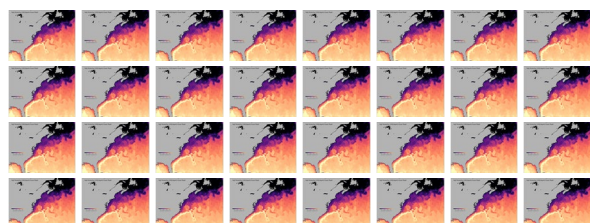
Scenario C



Variable A



Variable B



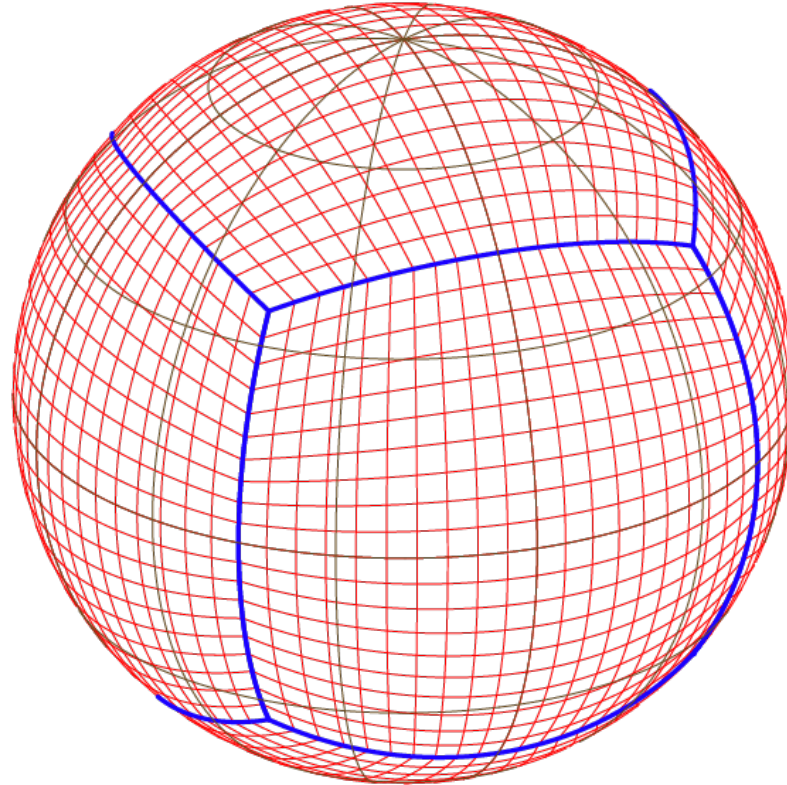
Variable C

# Complexity



LEAP

# Why is climate data hard?

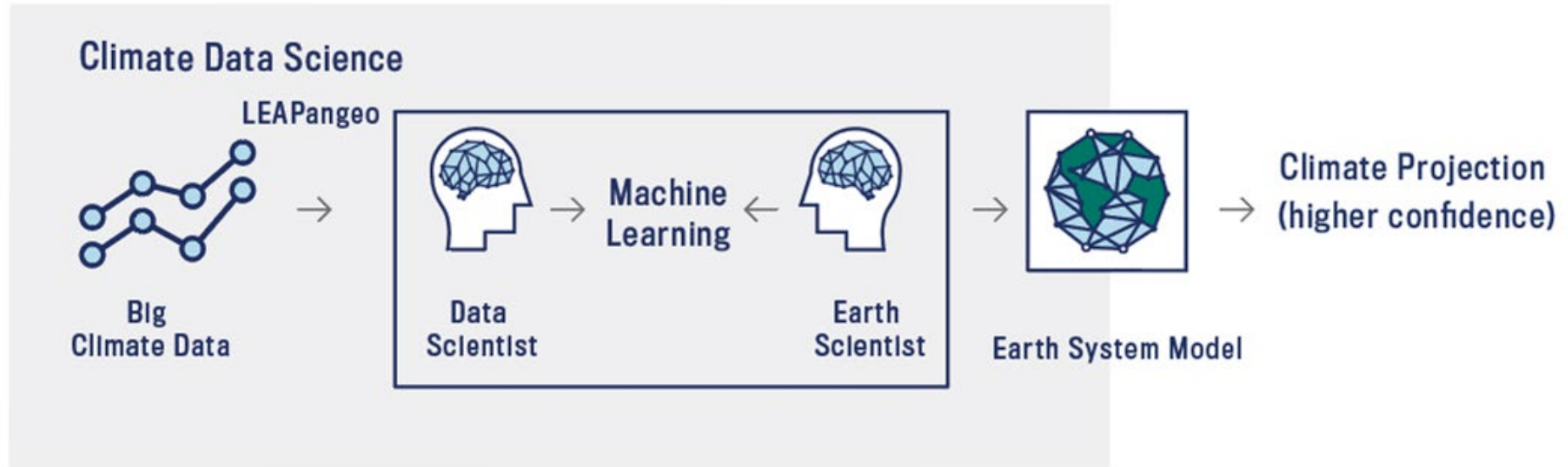


## Geometry



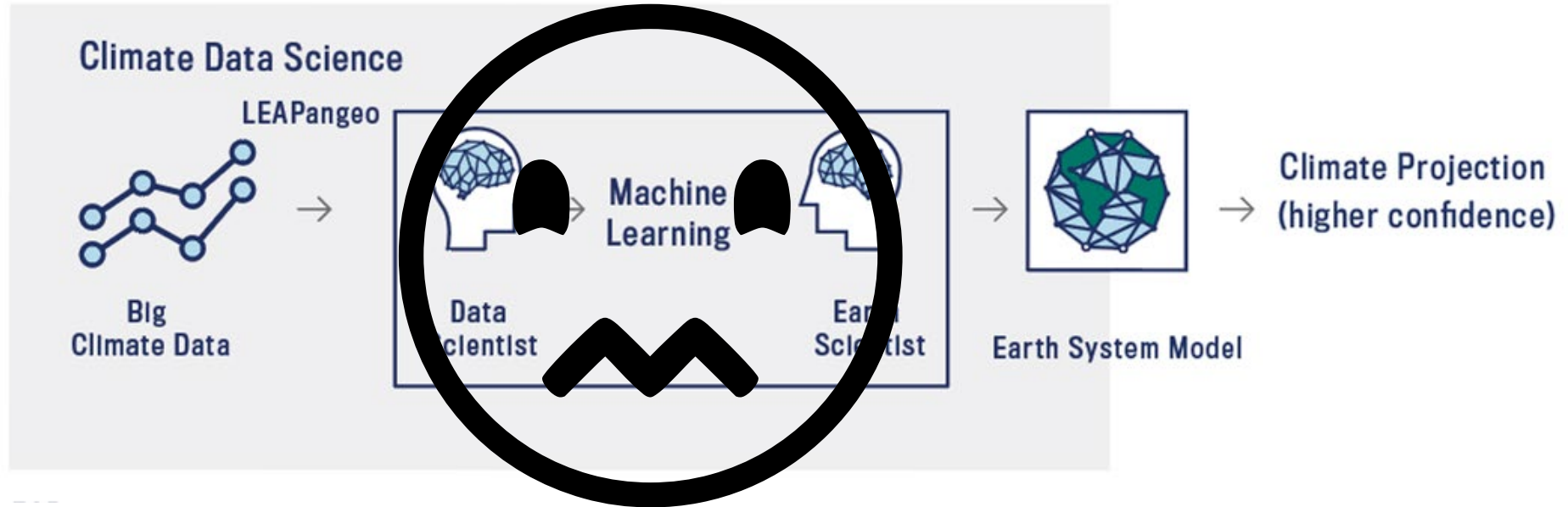
LEAP

# The Leap Data / Computing Challenge



- Very large, complex datasets
- Cutting-edge techniques / algorithms
- Diverse group of participants (different disciplines / expertise levels)

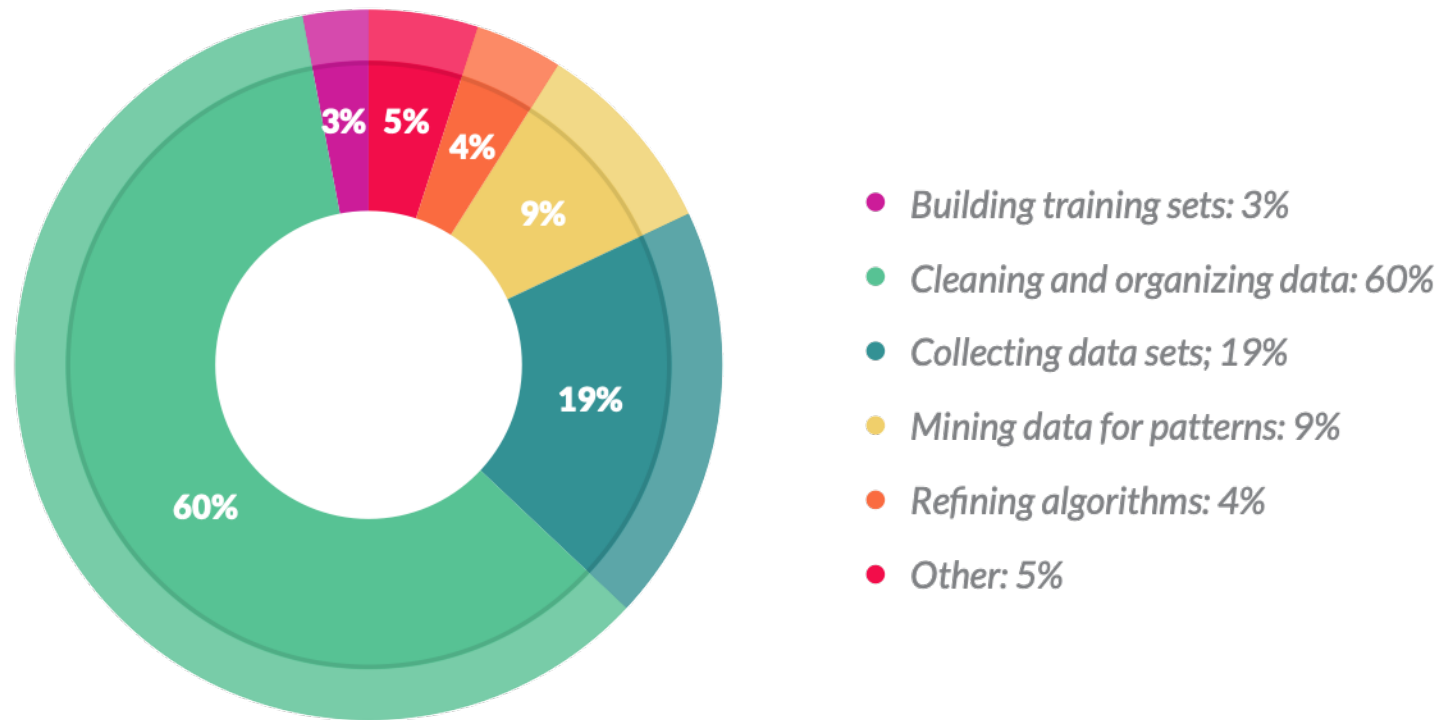
# The Leap Data / Computing Challenge



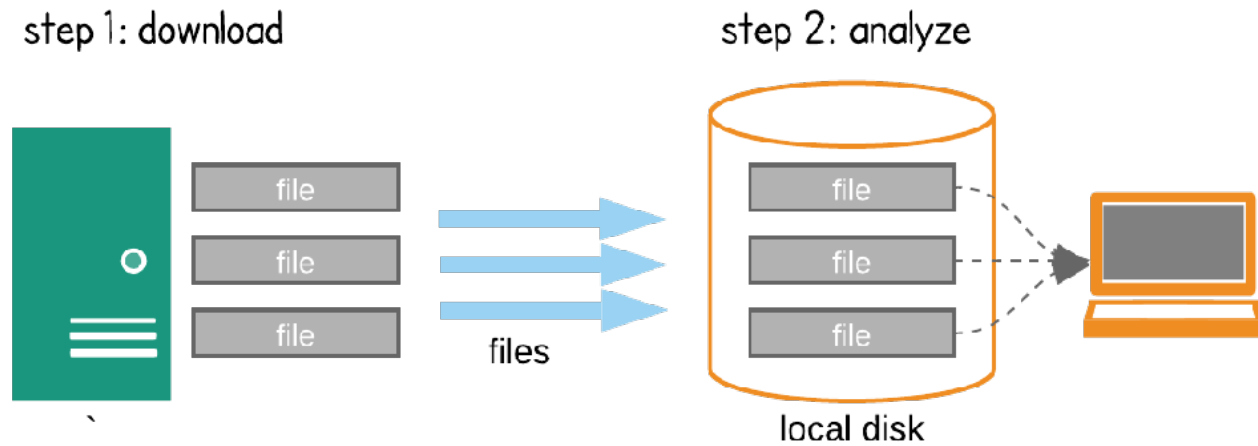
- Very large, complex datasets
- Cutting-edge techniques / algorithms
- Diverse group of participants (different disciplines / expertise levels)

# 80/20 Rule of Data Science

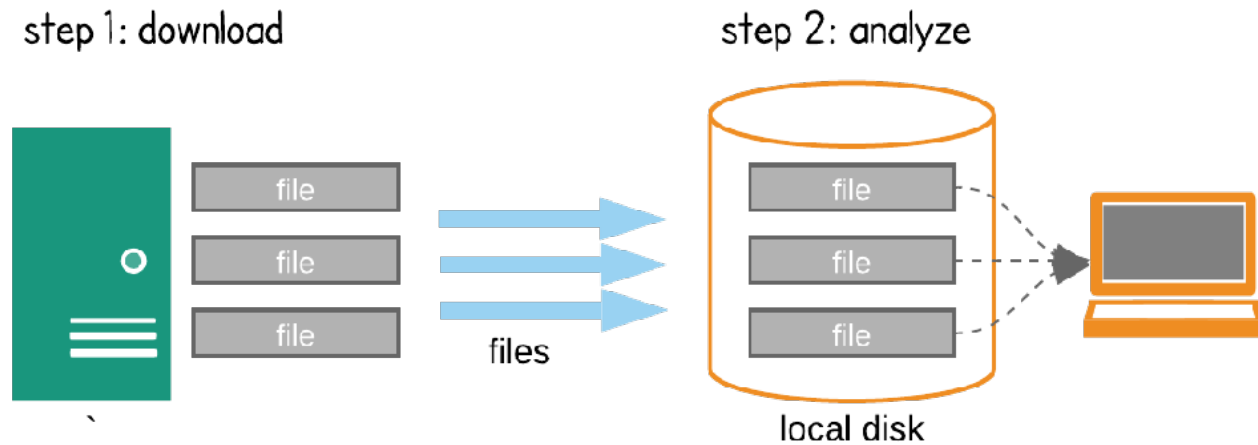
How do data scientists spend their time?



# Status Quo in Data Science: The “Download Model”

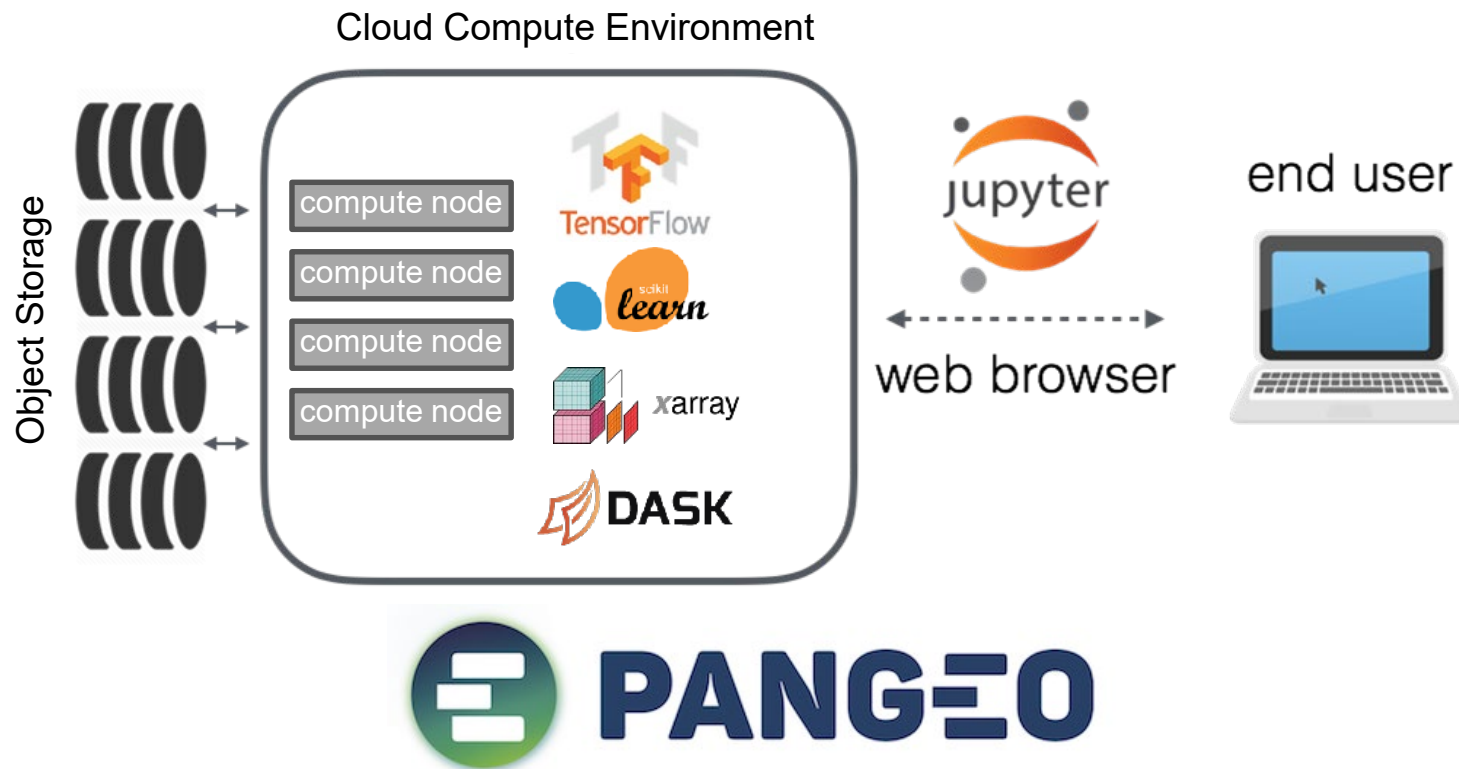


# Status Quo in Data Science: The “Download Model”



- ✗ Not scalable
- ✗ Not reproducible
- ✗ Barrier to collaboration
- ✗ Limits creativity
- ✗ Inefficient / duplicative
- ✗ Limits inclusion

# The Future of Data Science: Data-Proximate Computing



- ✓ On-demand scaling
- ✓ Inherently reproducible
- ✓ Highly collaborative
- ✓ Unleashes creativity
- ✓ Cost effective
- ✓ Accessible to all





## A Community Platform for Big-Data Geoscience

<http://pangeo.io/>

- Founded at Columbia in 2016
- Grass-roots collaboration between scientists, software developers around open-source tools for solving real problems
- Supported by NSF EarthCube, NASA ACCESS
- International partners, industry connections



Lamont-Doherty Earth Observatory  
COLUMBIA UNIVERSITY | EARTH INSTITUTE

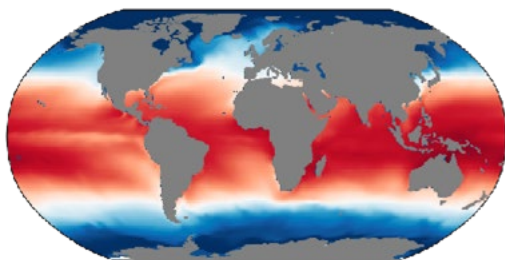
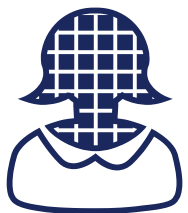


LEAP

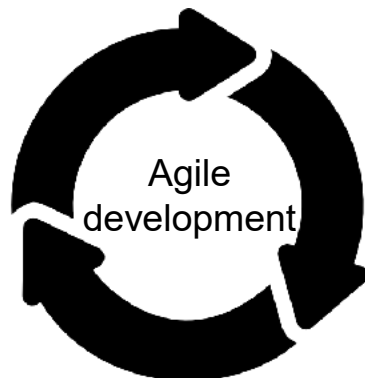


# PANGEO: Open Source, Community Development

## scientific users / use cases



- Define science questions
- Use software / infrastructure
- Identify bugs / bottlenecks
- Provide feedback to developers



## open-source software libraries



- Contribute widely to the open source scientific python ecosystem
- Maintain / extend existing libraries, start new ones reluctantly
- Solve integration challenges



## HPC and cloud infrastructure

- Deploy interactive analysis environments
- Curate analysis-ready datasets
- Platform agnostic



Google Cloud Platform



LEAP



# PANGEO : Real Scientists, Real Impacts

**Erin Dougherty** @edougherty\_

My #dayofscience: paper revisions 🙄 and using @pangeo\_data to analyze massive amounts of high-res climate data to understand floods in a current and future climate over the U.S. When not doing this, I ❤️ observing #wx directly via field work and watching storms. 🌩️



2:38 PM · Oct 15, 2019 · Twitter for iPhone

5 Retweets 51 Likes

pangeo Retweeted

**Patrick Gray** @clifgray

One of the highlights of #AGU19 for me was a workshop on @pangeo\_data. If you're interested in earth sci + geospatial analysis at scale I can't recommend their tutorial enough. Find it at [github.com/pangeo-data/pa...](https://github.com/pangeo-data/pangeo-tutorial) Clear and concise intro to #xarray, #dask, #geopandas, and #intake.



pangeo-data/pangeo-tutorial  
Interactive jupyter notebooks for pangeo tutorial events - pangeo-data/pangeo-tutorial  
[github.com](https://github.com)

2:03 PM · Dec 16, 2019 · Twitter Web App

15 Retweets 64 Likes

**Andrew Williams** @AndrewWilliams

It's taken a while, but I think that the whole xarray/Dask/cloud thing has finally clicked!

I've been working with CESM Large Ensemble data for a few months now - moved onto a Pangeo server and managed to speed up my workflow massively!

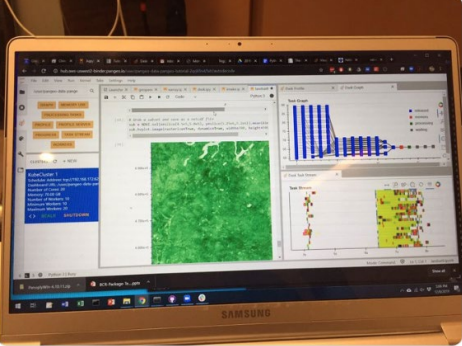
[@xarray\\_dev](#), [@dask\\_dev](#), [@pangeo\\_data](#)

10:23 AM · Jan 12, 2020 · Twitter Web App

6 Retweets 43 Likes

**Chelle Gentemann** @ChelleGentemann

Rocking 70GB Landsat data at the @pangeo\_data #AGU2019 tutorial. Such a powerful #OpenSource software stack.



10:45 PM · Dec 8, 2019 · Twitter for iPhone

17 Retweets 92 Likes

**Scott Collis** @Cyclogenesis\_au · Nov 11, 2019

Teaching a course on #@ONScience #OpenScience with @Shobenase at @MonashUni with 20 very diverse attendees @environmentca @argonne @armnewsteam



1 3 13

pangeo Retweeted

**Scott Collis** @Cyclogenesis\_au

Replying to @Cyclogenesis\_au @Shobenase and 4 others

Forgot to mention all done on @pangeo\_data ! Making future Pangeans!

9:15 PM · Nov 11, 2019 · Twitter Web App

1 Retweet 5 Likes

pangeo Retweeted

**Andrew Pauling** @andrewp109

#cmip6hack is just wrapping up, and has changed the way I will think about, and hopefully do, climate model analysis in the future. The @pangeo\_data infrastructure makes it all so easy.

5:29 PM · Oct 18, 2019 · Twitter Web App

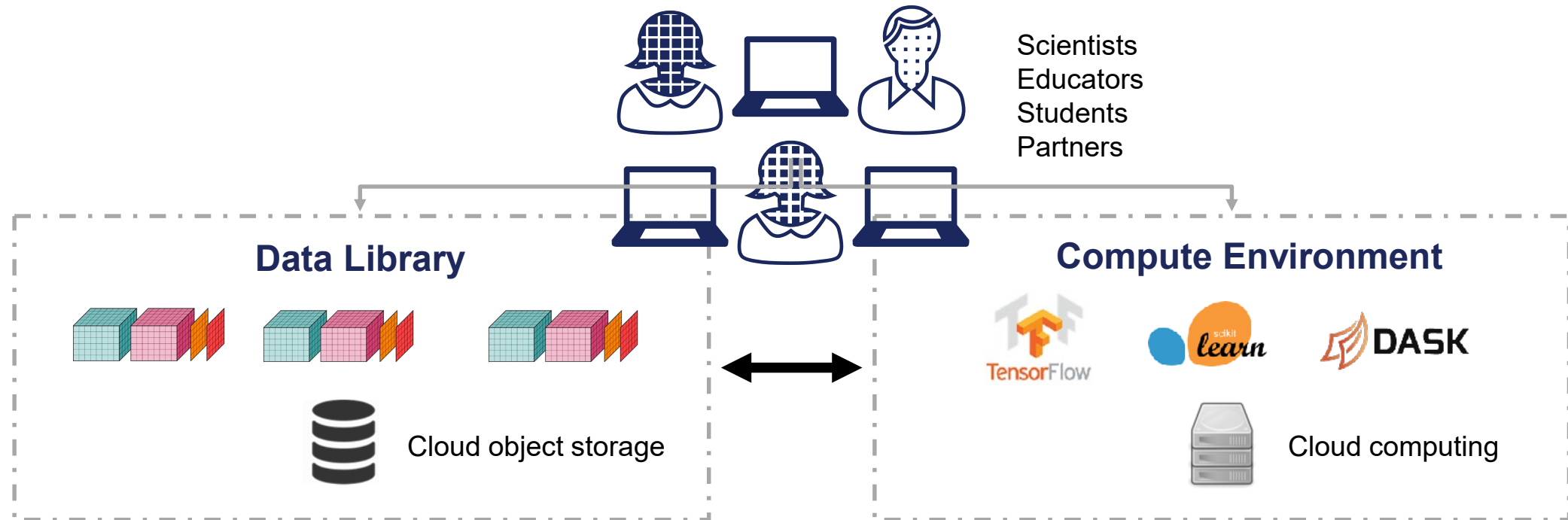
6 Retweets 24 Likes



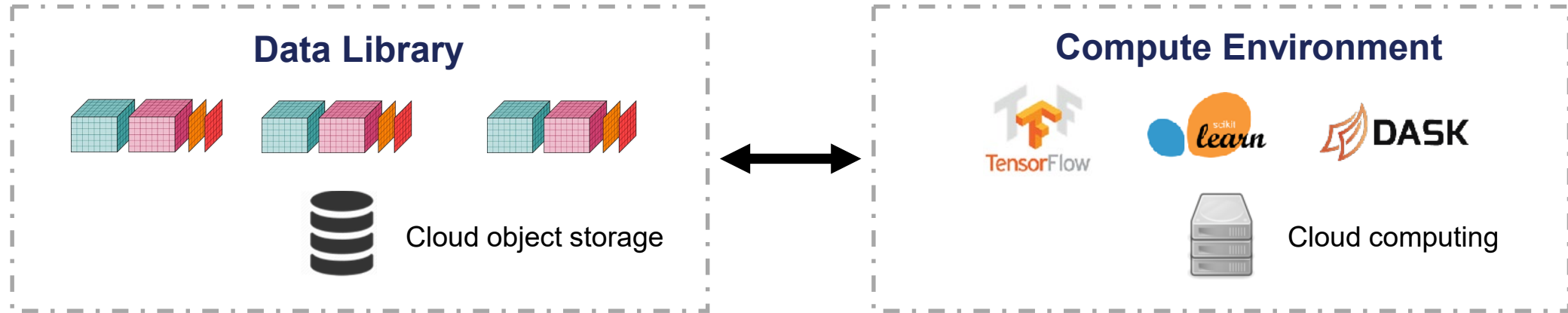


# LEAP + PANGEO = LEAPangeo

collaboration platform for the LEAP project



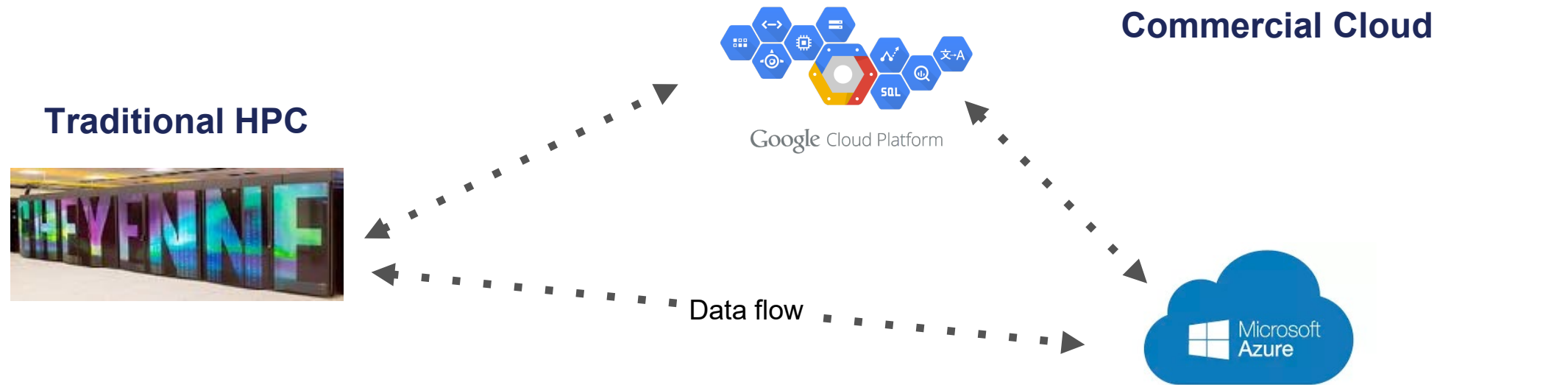
# LEAPangeo Architecture is Open and Modular



- Analysis ready data, optimized for ease of access and computing
- Used to train / validate ML models
- Hosted by cloud providers: fully public datasets accessible to the world
- Impact: lower the barriers to working with climate data

- Uniform software environments for all collaborators w/ latest ML tools
- Access to ML-optimized hardware (e.g. GPUs, TPUs)
- Impact: enhance collaboration and facilitate reproducibility

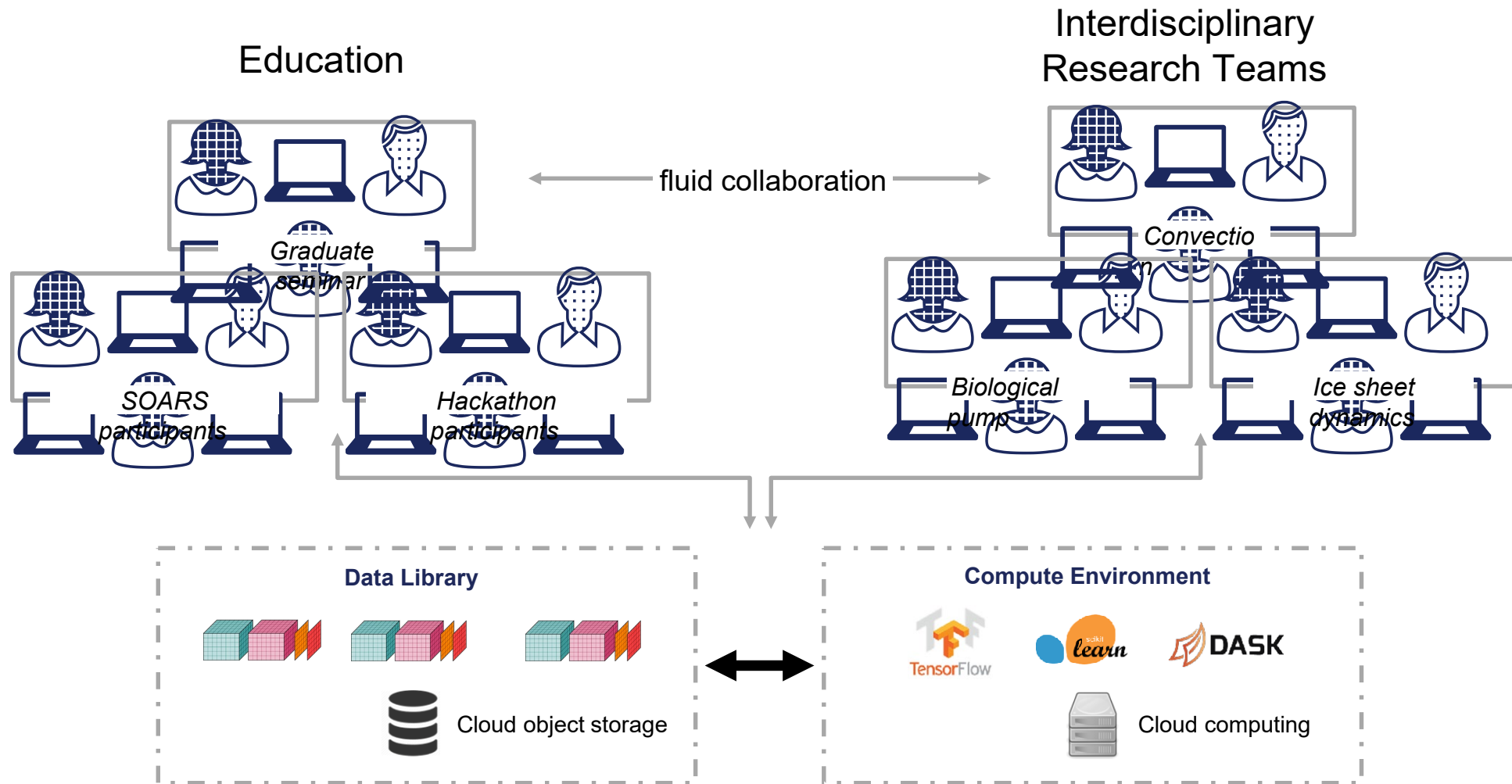
# Cloud + HPC: Complementary Tools



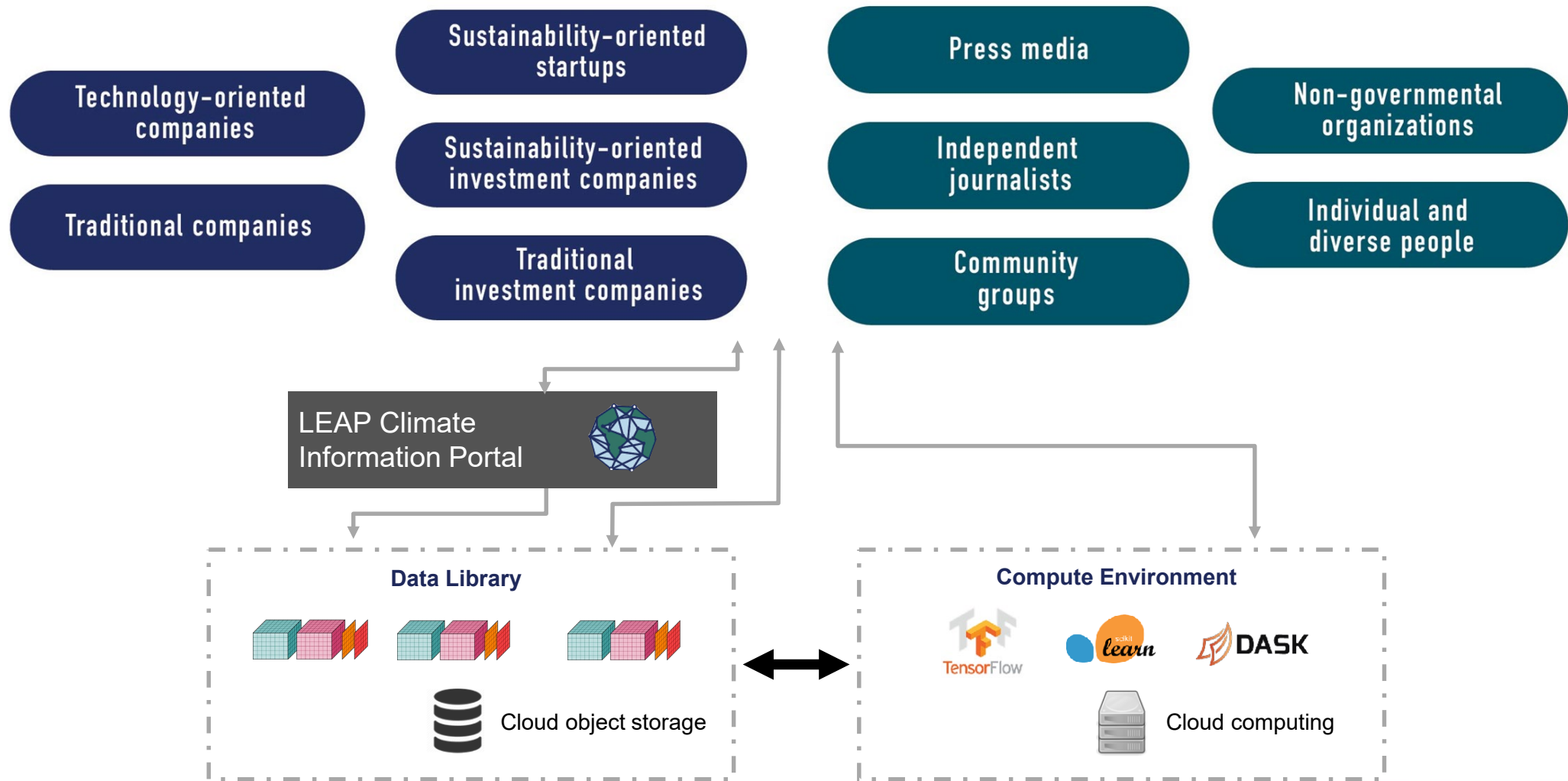
- Used to *run* large-scale simulations
- Generate training data (high-res simulations)
- Test parameterizations online in CESM
- Restricted access to data and compute

- Used for interactive data analysis
- Analyze data / train models
- Collaborate with industry ML teams
- Open access

# LEAPangeo Supports Education and Research



# LEAPangeo Enables Knowledge Transfer



# Management for Success

- Two FT engineers at Columbia:  
data library + compute environment
- Open development / Agile approach  
➔ □ low risk, immediate impact
- Contribution to upstream open-source libraries  
➔ □ sustainability and broader impacts across NSF research
- Partnership w/ Google + Microsoft:  
Cloud computing credits, joint development
- Bi-annual training bootcamps for all participants



Google Cloud Platform

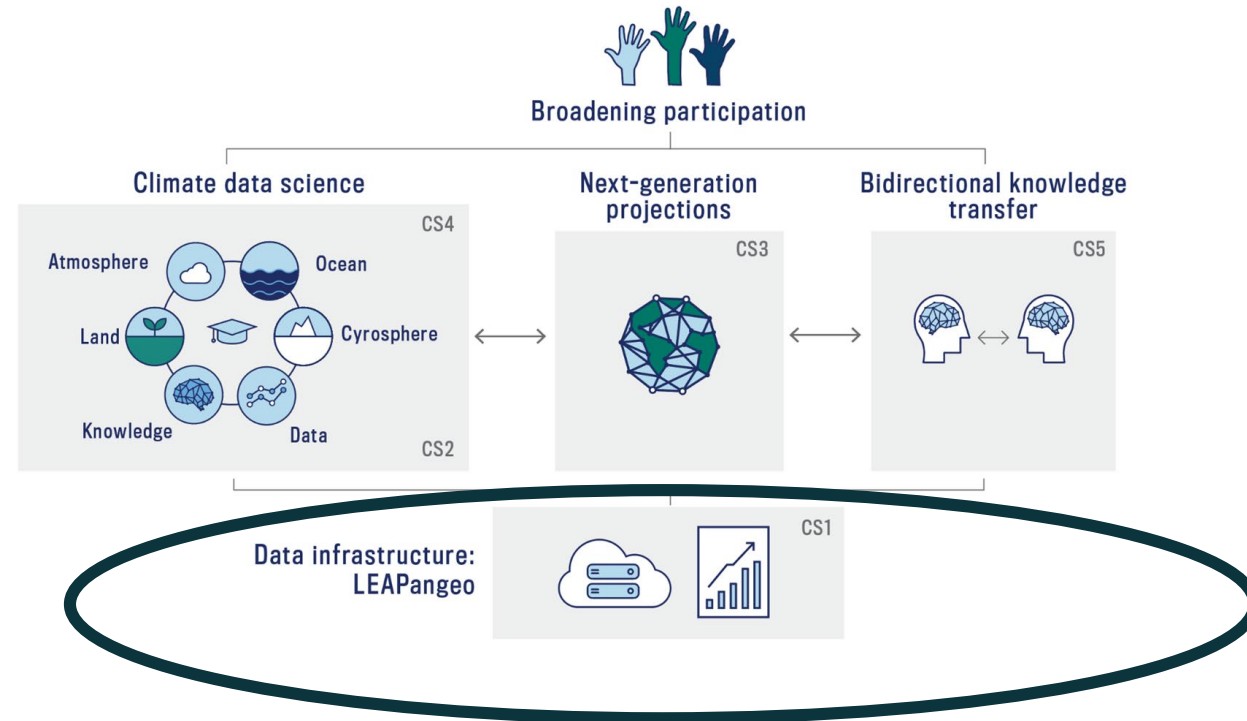


LEAP

# LEAPangeo is Sustainable

- Underlying software is 100% open source and modular
- Implementation is cloud agnostic: no lock-in
- Data can be moved between commercial cloud, traditional HPC, and emerging platforms (e.g. NSF OpenStorageNetwork)
- Wide range of government and industry partners

# LEAPangeo will transform science!



*Enable the ambitious LEAP research / education / knowledge-transfer goals.  
Transform infrastructure for collaboration in Data Science.*

# Earth System Model Coupling + Integration



NCAR Model  
Development Liaison  
**David Lawrence**

- Senior Scientist and Section Head for Terrestrial Sciences, Climate + Global Dynamics Lab, NCAR
- PhD, Atmospheric + Ocean Science, Colorado

### **Expertise**

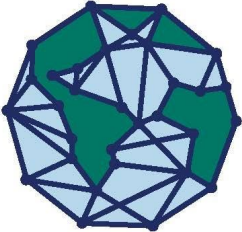
- Earth system modeling
- Land-atmosphere interactions
- Land use change
- Arctic climate feedbacks

### **Synergistic Activities**

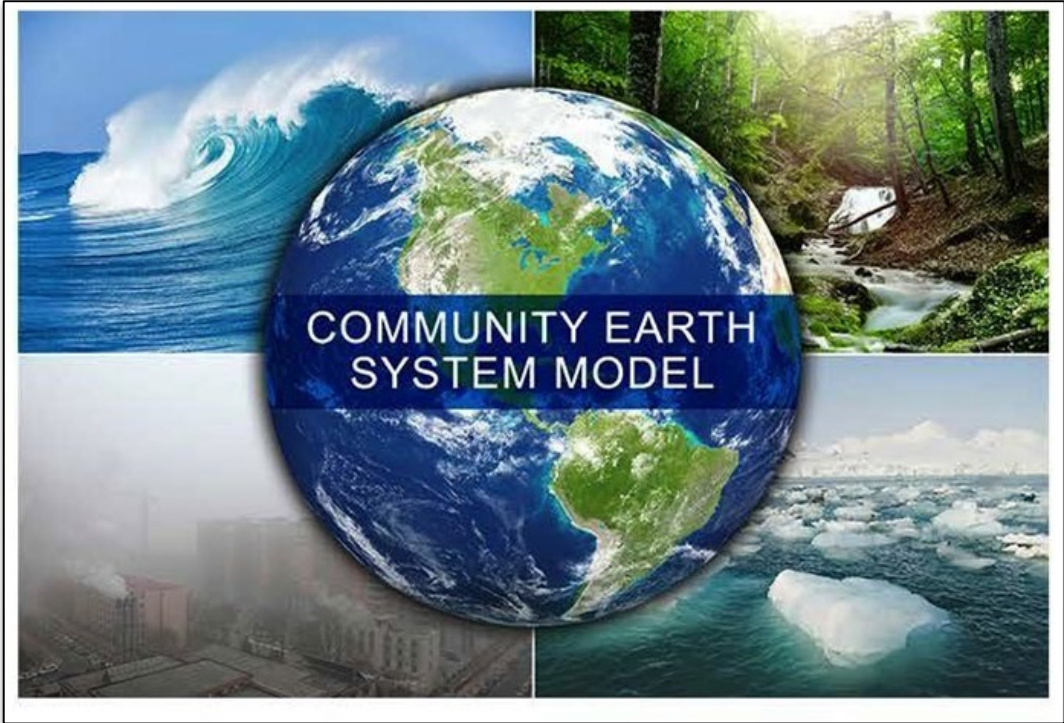
- Co-Lead, Community Terrestrial Systems Model
- Co-Chair, Land Use Model Intercomparison Project
- CESM Scientific Steering Committee (2008-2019)



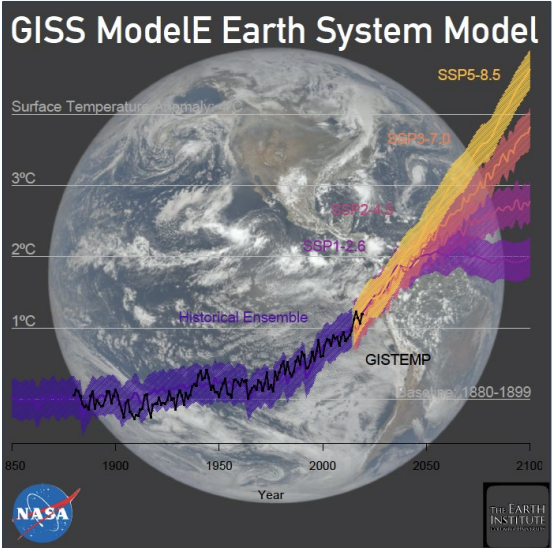
# My presentation will explain how



LEAP

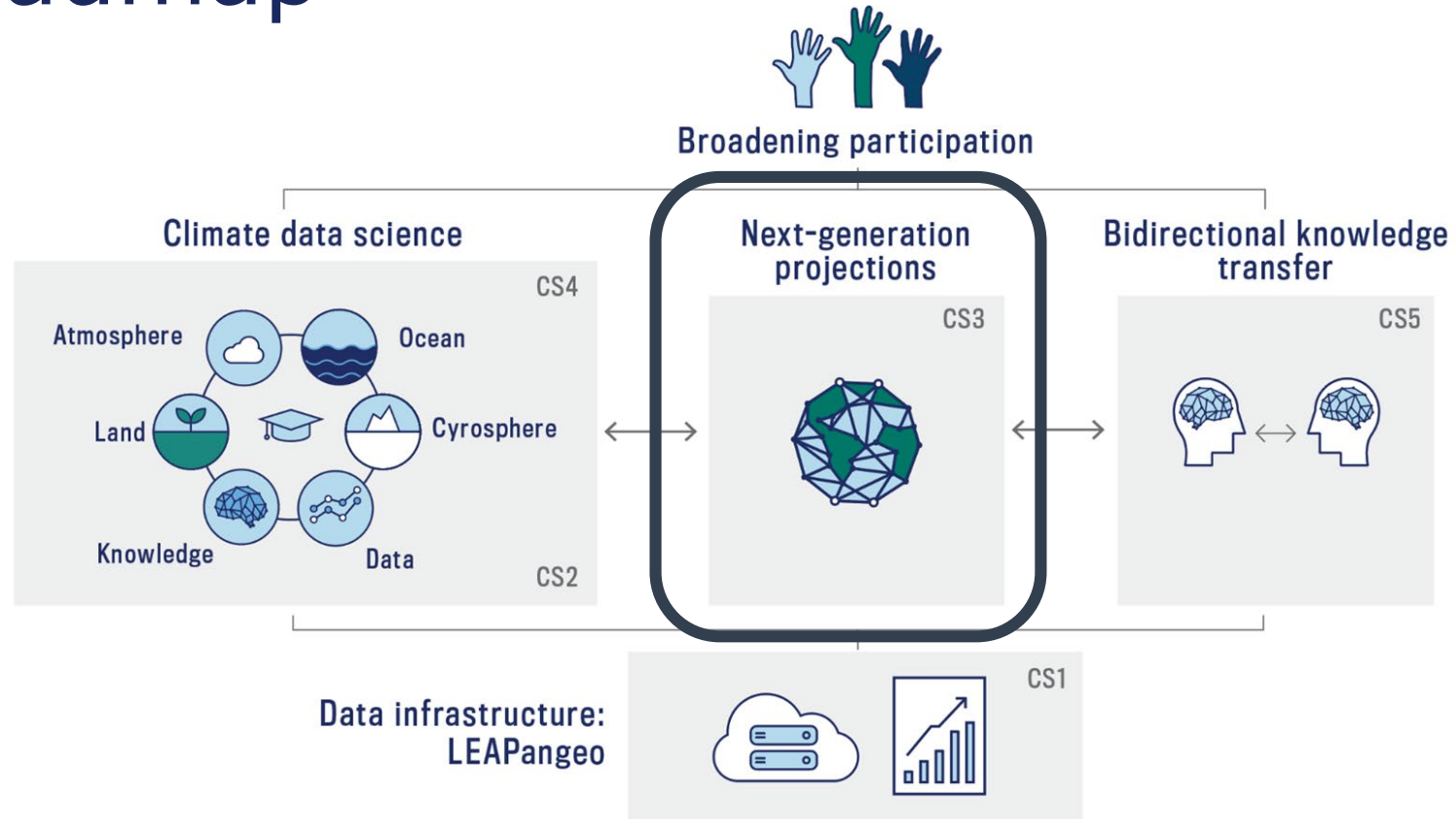


Gavin Schmidt



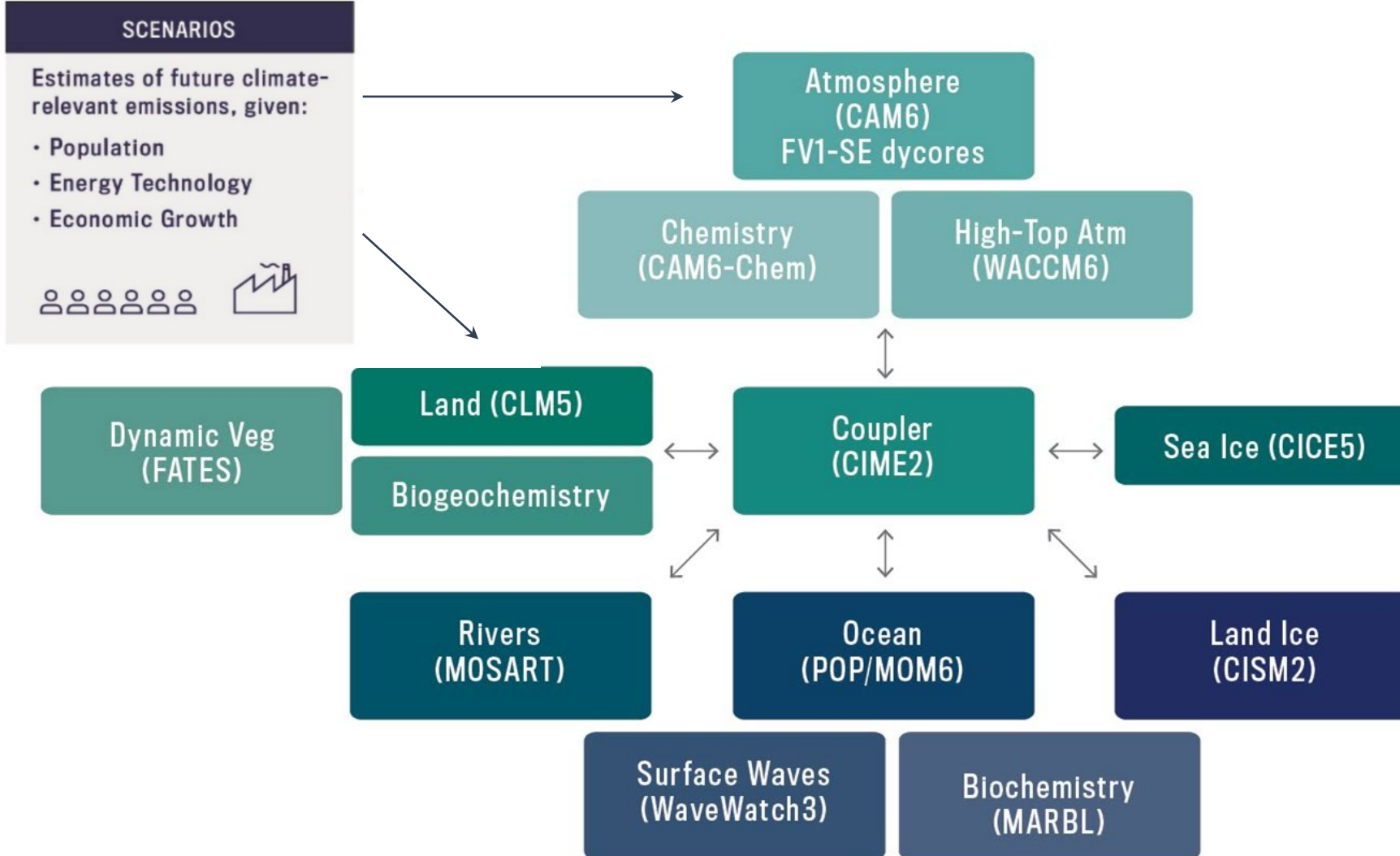
LEAP

# Our Roadmap



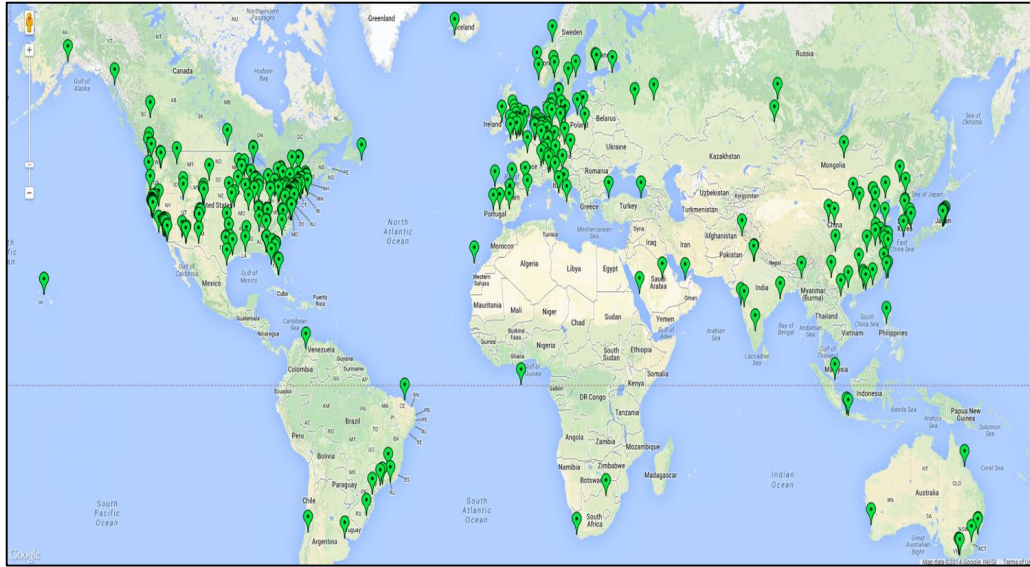
*LEAP forward in the **reliability**, **utility**, and **reach** of climate projections through synergistic innovations in data science and climate science.*

# Community Earth System Model (CESM2)

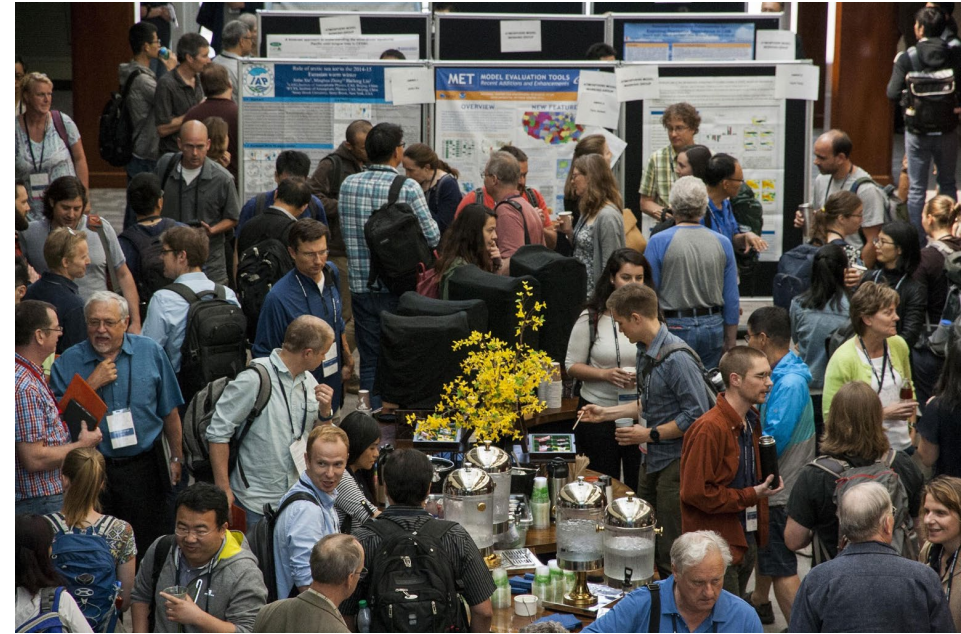


- Highly flexible system
- Multiple supported resolutions
- Data assimilation
- > 2 million lines of code
- > 1000 geophysical output variables
- **20+ year legacy of NSF investment**

# CESM supports a global research community



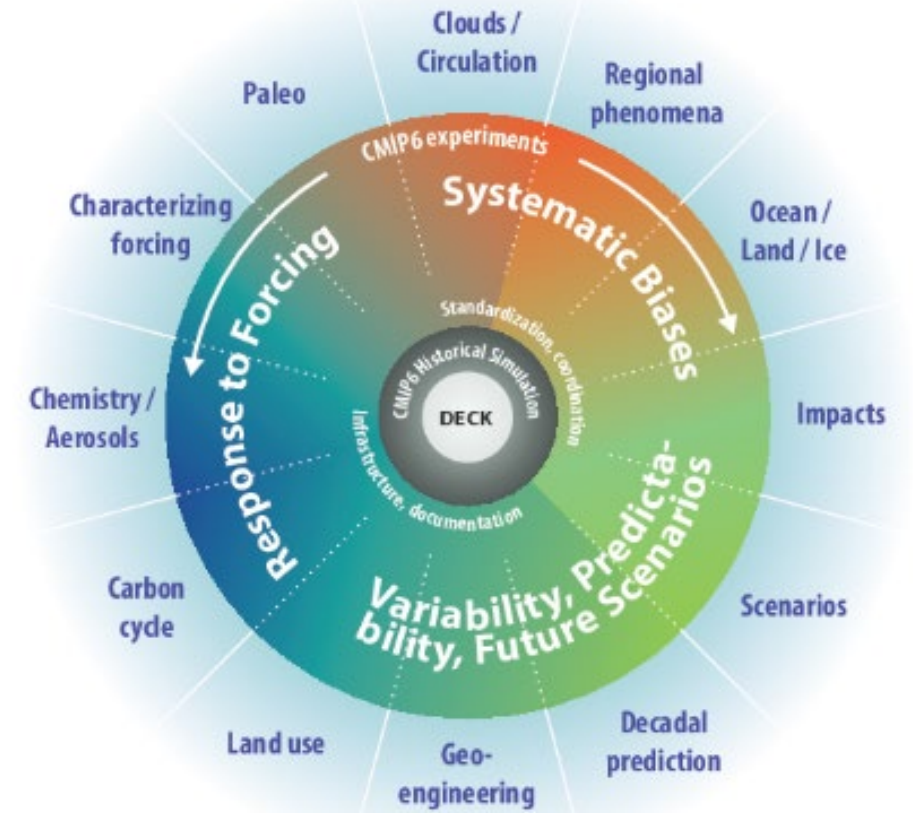
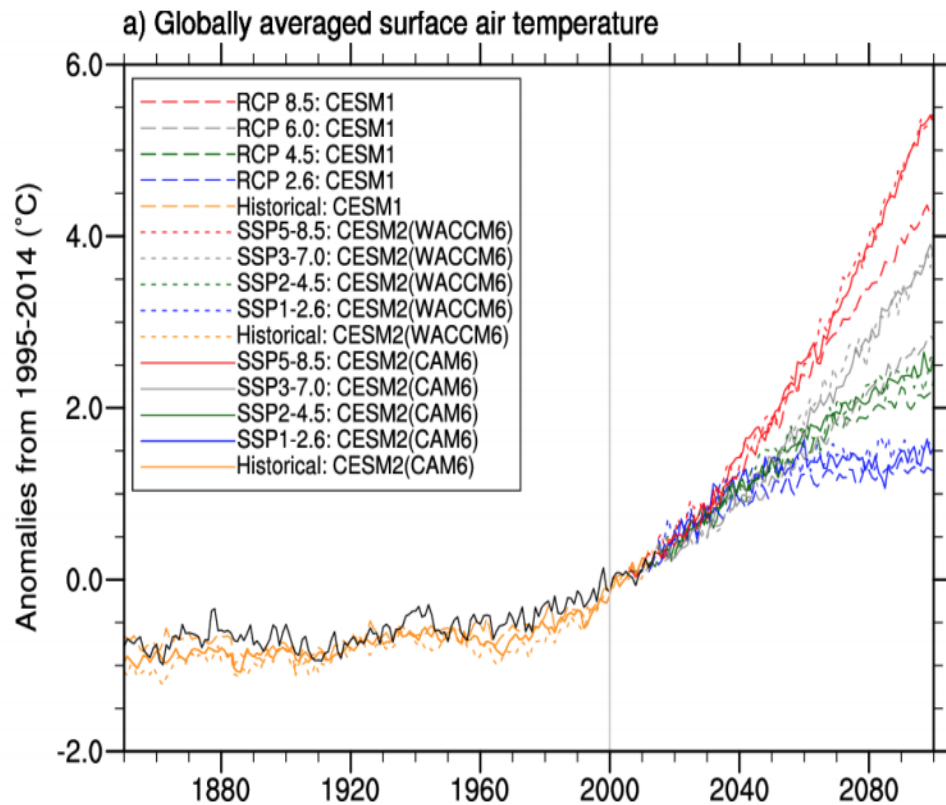
Download of released version since 2010



- Workshops and tutorials
- Easy-to-use open source model
- User support
- Publicly available model output

# CESM used for wide range of climate research

# Coupled Model Intercomparison Project (CMIP6)



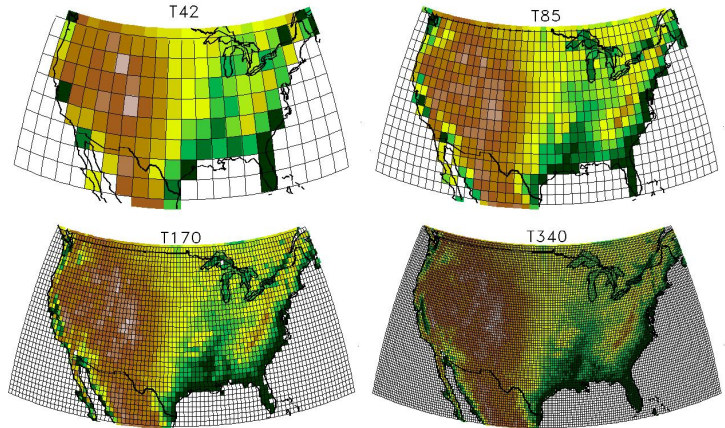
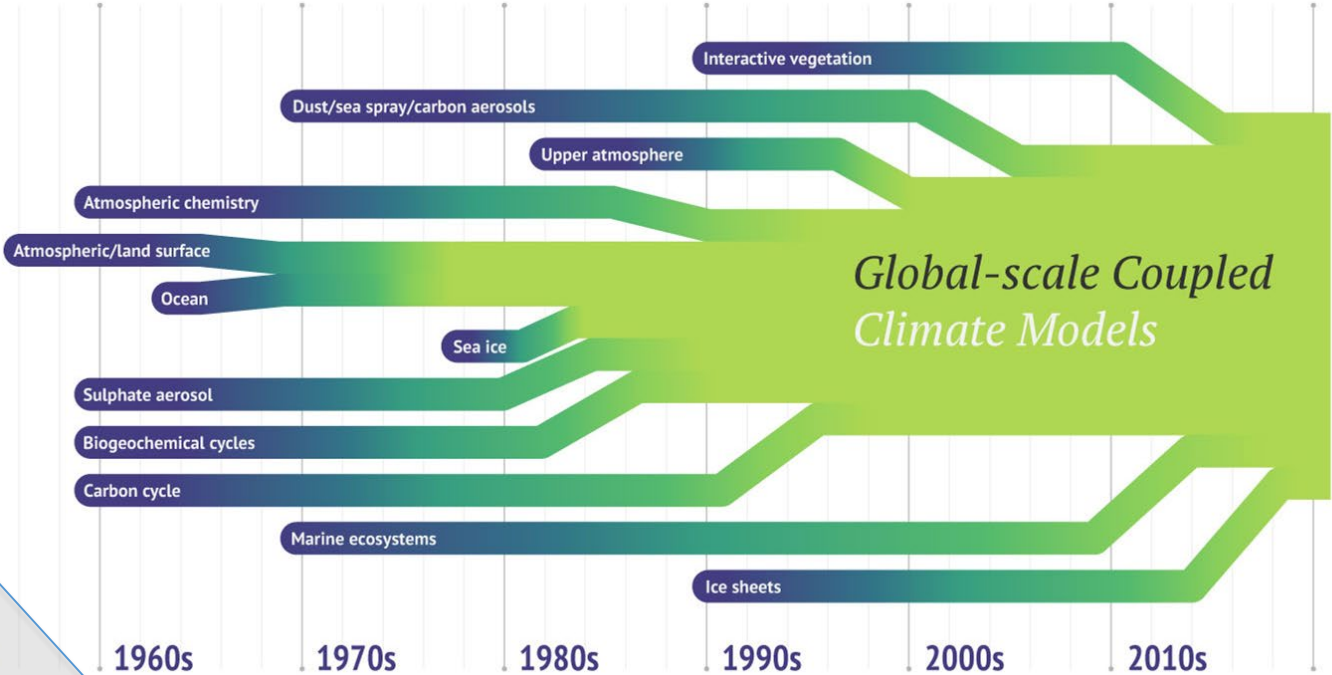
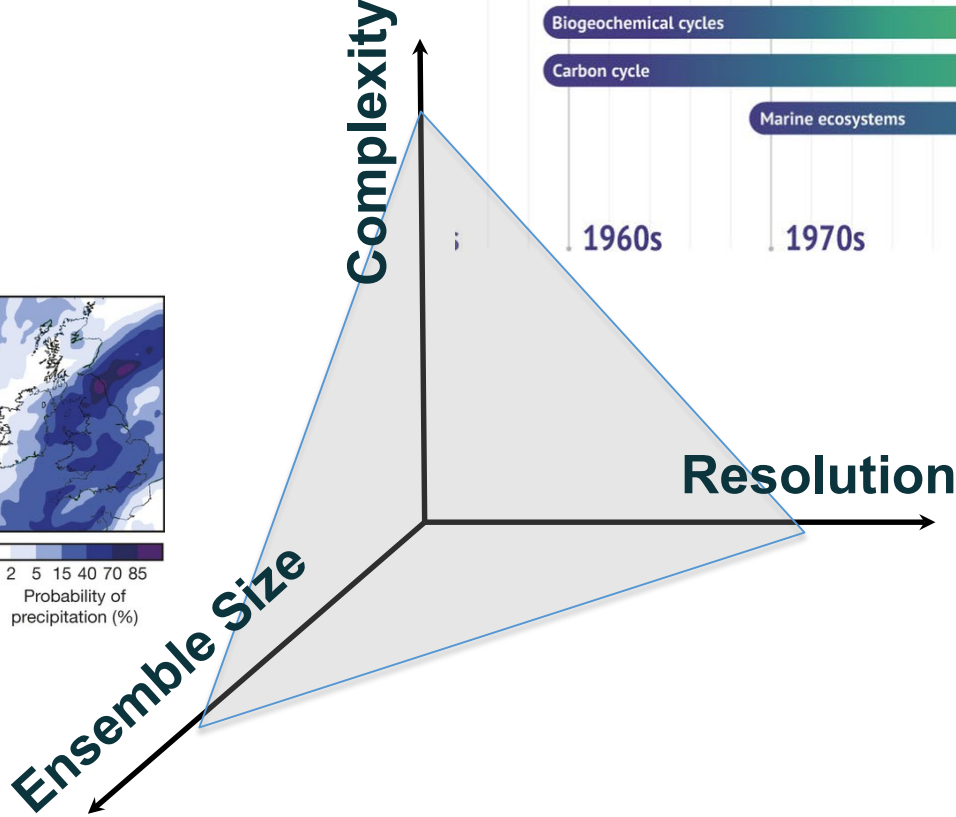
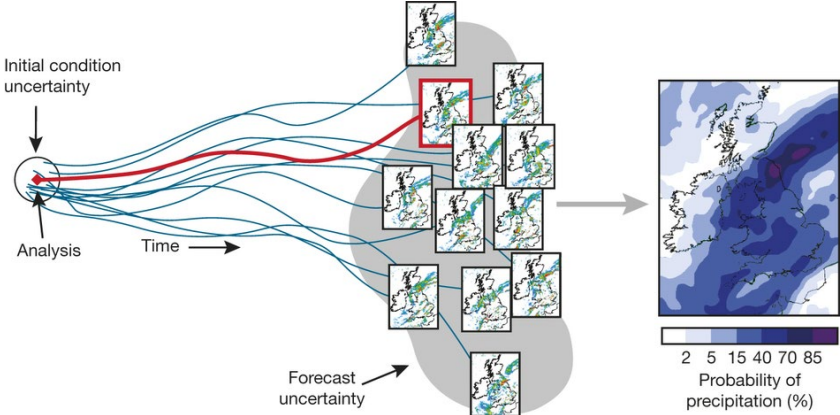
- > 1000 model experiments
- > 25,000 years of simulation
- ~2 PB of compressed data

# CESM as research tool

- Climate variability and change
- Earth System prediction
- Biogeochemical cycles
- Process understanding
- Impacts research
- ...



# CESM as research tool



# LEAP's impact on CESM development and research



- brings advanced data science into CESM development, consistent with NCAR Strategic Plan priorities
- innovates in ways that NCAR could not pursue on own
- provides avenue for more reliable Earth System projections and predictions
- educates next generation of model developers
- **will deliver a transformative advance for a significant NSF investment**



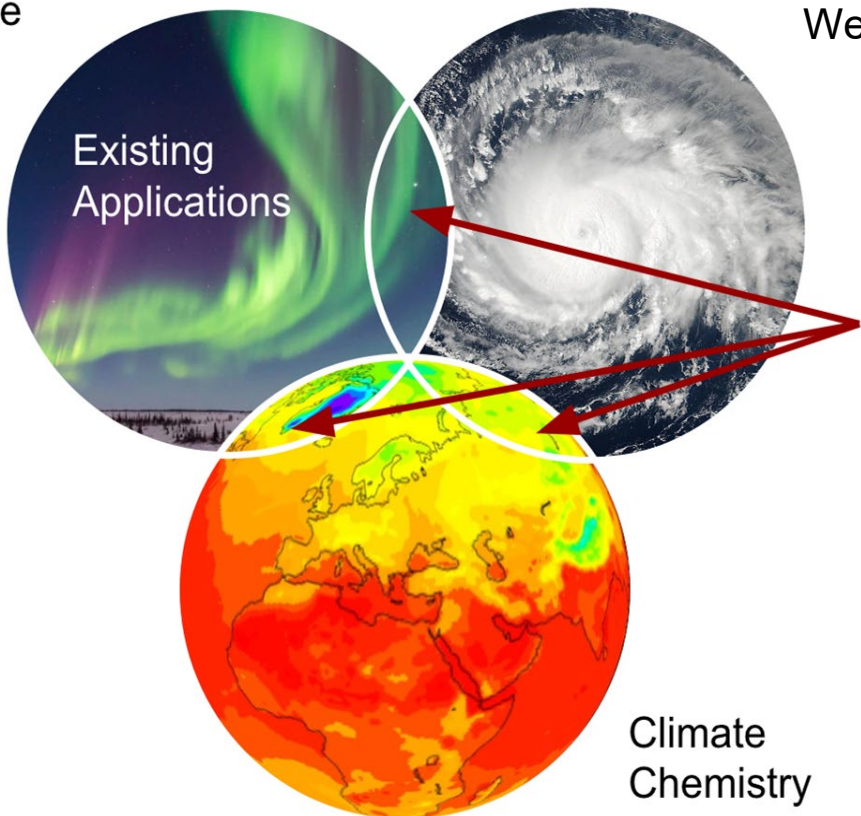
# LEAP's impact will extend beyond climate modeling

## NCAR transition to unified modeling systems

### SIMA

#### System for Integrated Modeling of the Atmosphere

Geospace



Existing Applications

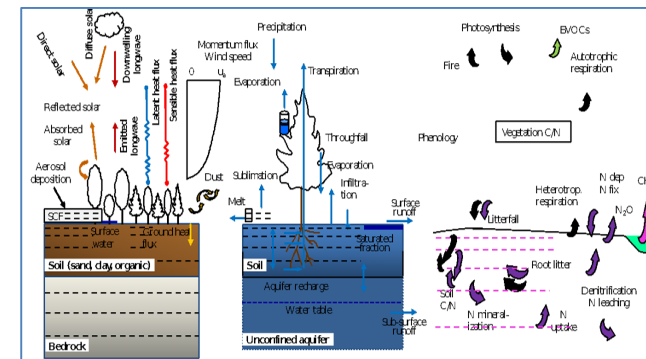
Weather

Climate Chemistry

### CTSM

#### Community Terrestrial Systems Model

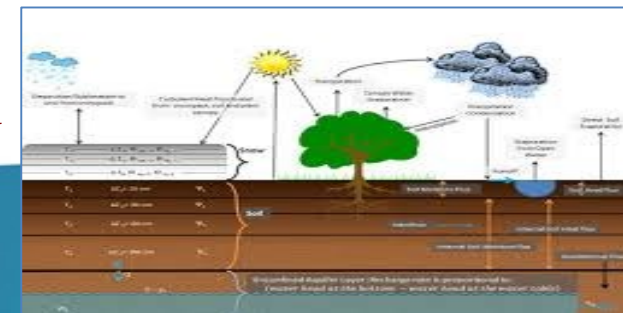
CLM climate



SUMMA concepts

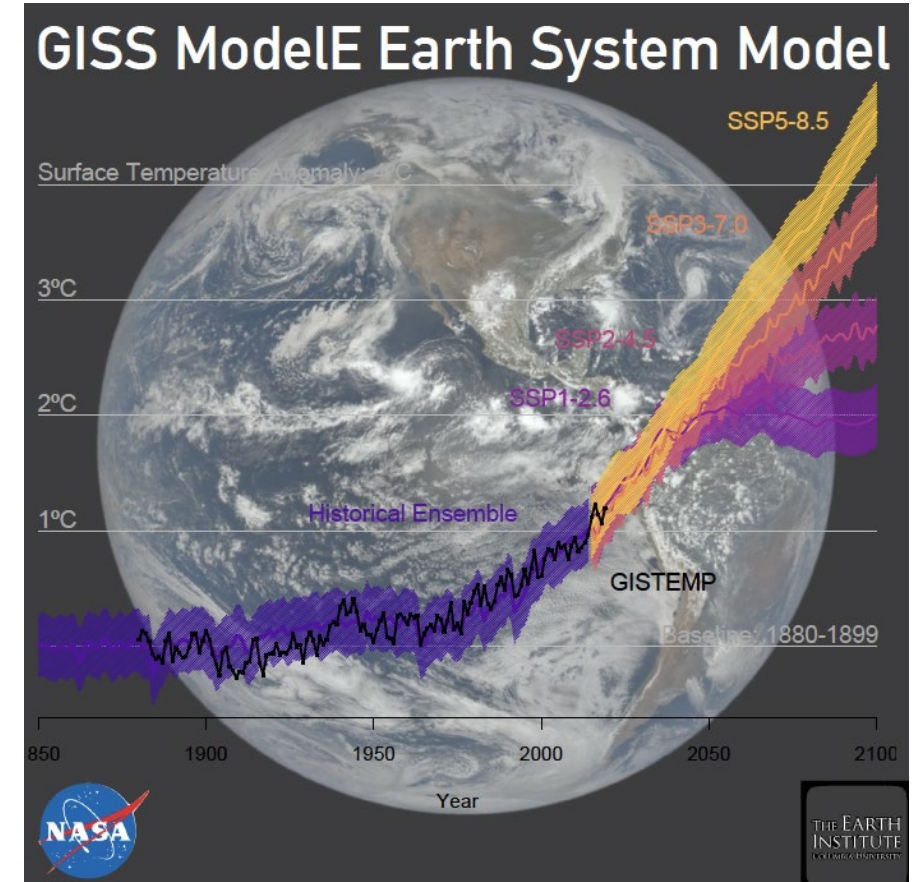
CTSM

Noah-MP weather, hydrology



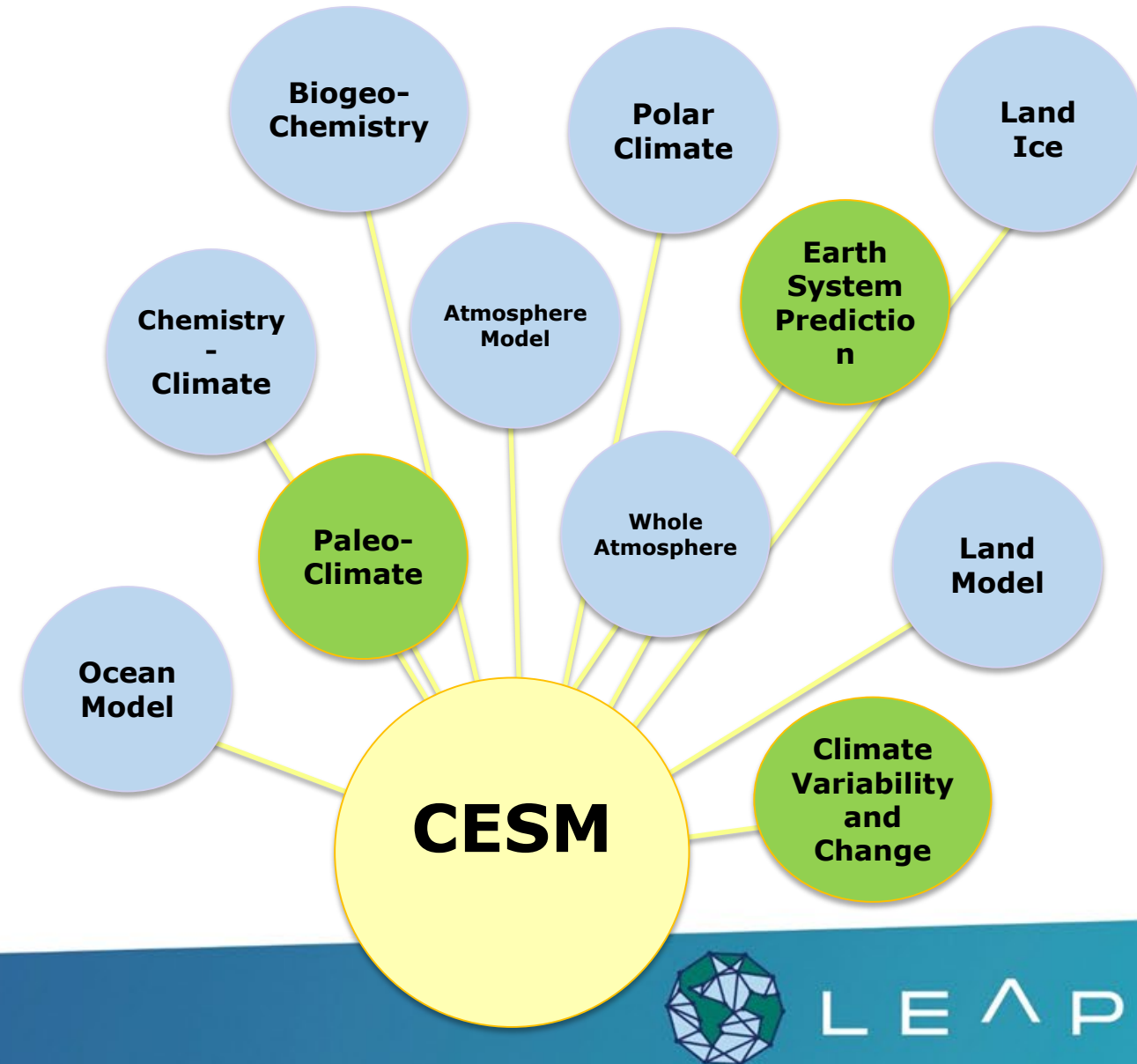
# LEAP – NASA GISS ModelE partnership

- Most promising LEAP advances will be tested in ModelE as well as CESM
- Assessment in multiple models yields insight about impact and robustness of ML parameterizations in context of structural uncertainty
- Take advantage of existing relationships between LEAP and GISS scientists

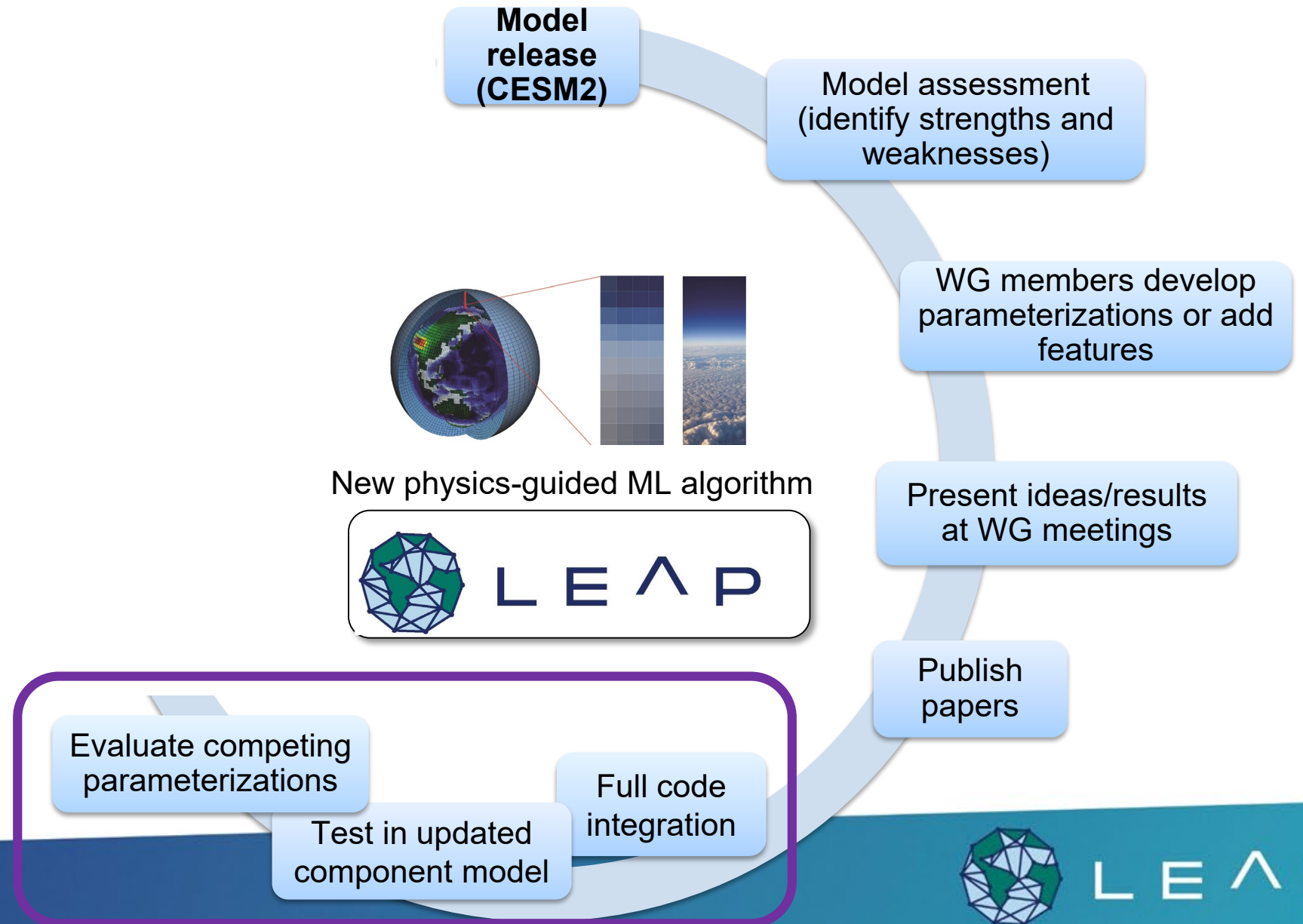


# LEAP integration into CESM development process

- LEAP scientists will participate in CESM Working Groups
- Short- and long-term collaborative visits to NCAR
- Dedicated software engineering supported by LEAP to integrate LEAP code into CESM
- Scientific support from NCAR staff scientists in LEAP team



# LEAP integration into CESM development process



# Compatibility with other new developments

## Developments for Community Land Model (CLM5)

1. Flexible leaf stoichiometry
2. Leaf N optimize for photosynthesis
3. Carbon costs for plant N uptake
4. Plant hydraulics w/ hydraulic redistribution
5. Spatially explicit soil depth
6. Dry surface layer, revised groundwater
7. Revised canopy interception
8. MOSART river model (hillslope  tributary  main channel)
9. Canopy snow storage and radiation
10. Fresh snow density (T, wind)
11. Simple firn model
12. Global crop model (8 crop types)
13. Transient irrigation and fertilization
14. Dynamic landunits (nat veg  crop, glacier  nat veg)
15. Urban heating and AC, heat stress indices
16. Deforestation, cropland fire
17. New PFT and CFT distributions
18. Carbon isotopes
19. *Coupled fire trace gas emissions*
20. *Ecosystem demography (FATES),*
21. *Ozone damage to plants*
22. *Shifting cultivation*



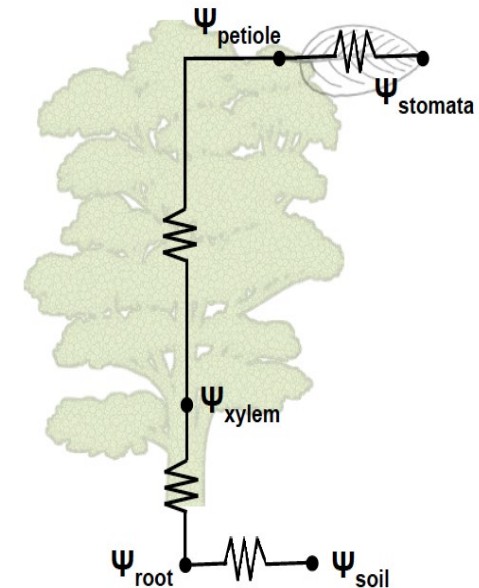
# Compatibility with other new developments

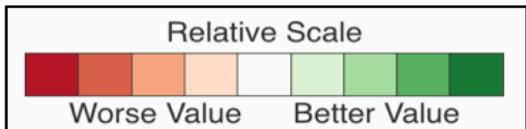
## Developments for Community Land Model (CLM5)

1. Flexible leaf stoichiometry
2. Leaf N optimize for photosynthesis
3. Carbon costs for plant N uptake
- 4. Plant hydraulics w/ hydraulic redistribution**
5. Spatially explicit soil depth
6. Dry surface layer, revised groundwater
7. Revised canopy interception
8. MOSART river model (hillslope  tributary  main channel)
9. Canopy snow storage and radiation
10. Fresh snow density (T, wind)
11. Simple firn model
12. Global crop model (8 crop types)
13. Transient irrigation and fertilization
14. Dynamic landunits (nat veg  crop, glacier  nat veg)
15. Urban heating and AC, heat stress indices
16. Deforestation, cropland fire
17. New PFT and CFT distributions
18. Carbon isotopes
19. Coupled fire trace gas emissions demography (FATES),  
*Ecosystem*
20. Ozone damage to plants
21. Shifting cultivation

## Implementing plant hydraulics

- Collaboration with Pierre Gentine's student
- 2+ year process
- Many interdependencies needed to be resolved





	CLM4	CLM4.5	CLM5
<b>Ecosystem and Carbon Cycle</b>			
Biomass	Red	Light Green	Green
Burned Area	Red	Light Green	Green
Carbon Dioxide	Orange	Light Green	White
Gross Primary Productivity	Red	Light Green	Light Green
Leaf Area Index	Orange	Light Orange	Green
Global Net Ecosystem Carbon Balance	Light Orange	Light Orange	Green
Net Ecosystem Exchange	Red	White	Light Green
Ecosystem Respiration	Red	Light Green	Light Green
Soil Carbon	Red	Light Green	Light Green
<b>Hydrology Cycle</b>			
Evapotranspiration	Orange	Light Green	Light Green
Evaporative Fraction	Orange	Light Green	Light Green
Latent Heat	Orange	White	Light Green
Runoff	Light Green	Light Green	Orange
Sensible Heat	Light Orange	White	Light Green
Terrestrial Water Storage Anomaly	Red	Light Green	Light Green
Permafrost	White	Red	Light Green
<b>Radiation and Energy Cycle</b>			
Albedo	Light Orange	White	Light Green
Surface Upward SW Radiation	White	White	White
Surface Net SW Radiation	White	White	White
Surface Upward LW Radiation	Light Green	White	Light Orange
Surface Net LW Radiation	Light Green	Light Green	Light Orange
Surface Net Radiation	Light Green	Light Green	Light Orange

# Model performance tracking

## Standardized metrics for evaluation

- Climate variability diagnostics package (CVDP)
- CESM component model diagnostics packages
- **International Land Model Benchmarking (ILAMB)**
- ESMValTool
- ...
- New LEAP machine-learning based metrics derived from big data



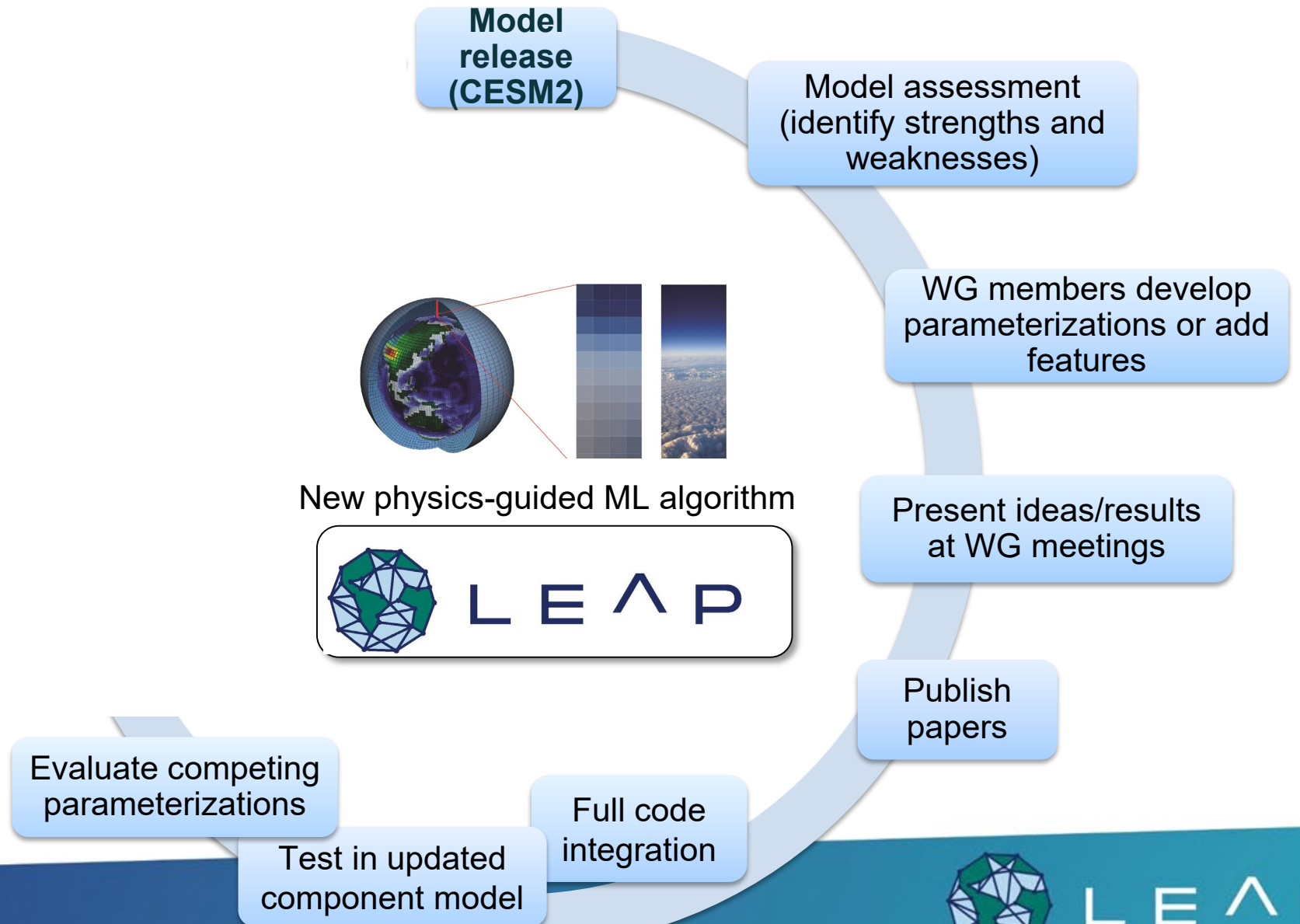
**RUBISCO**

Lawrence et al., 2019

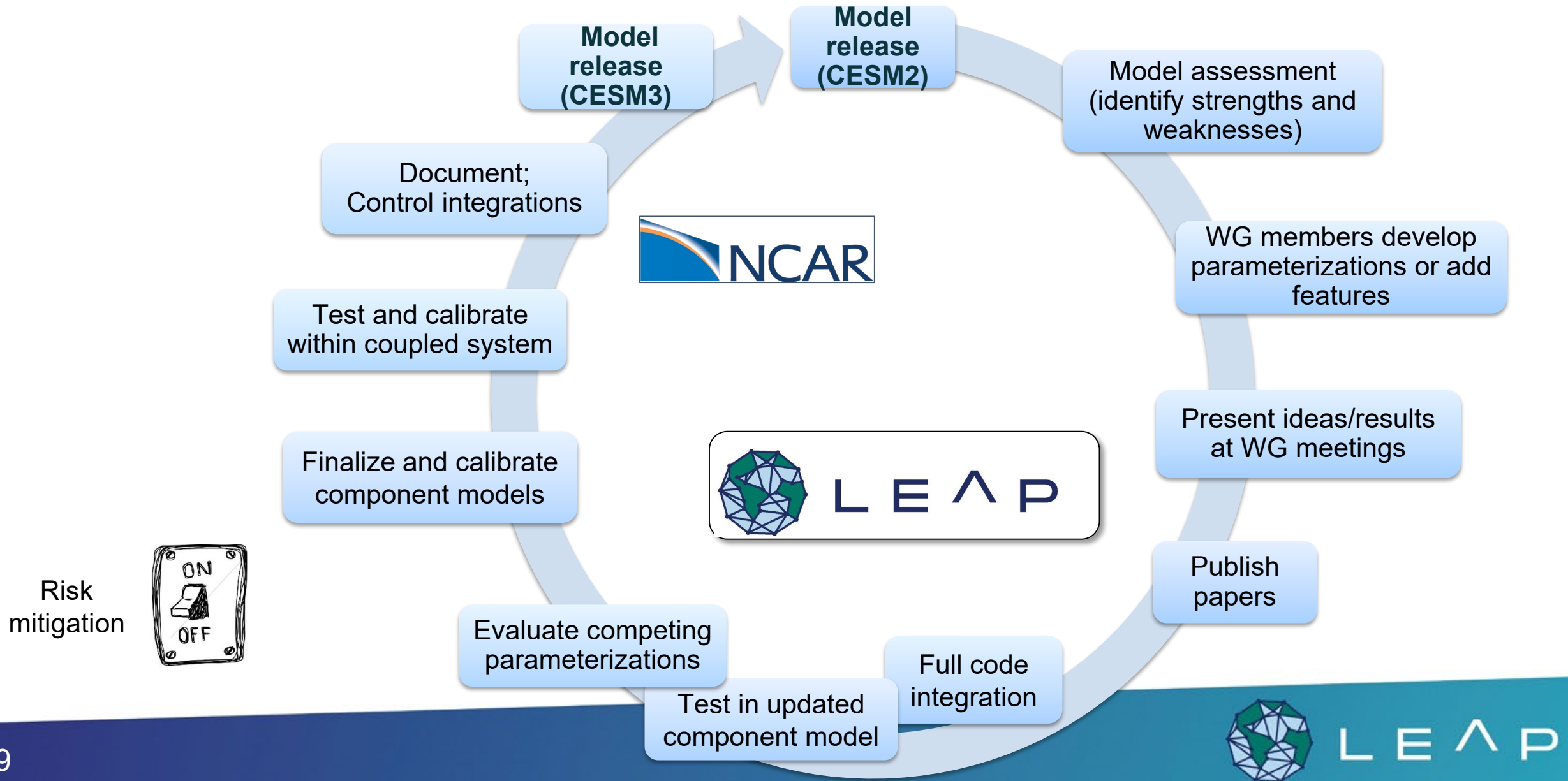


LEAP

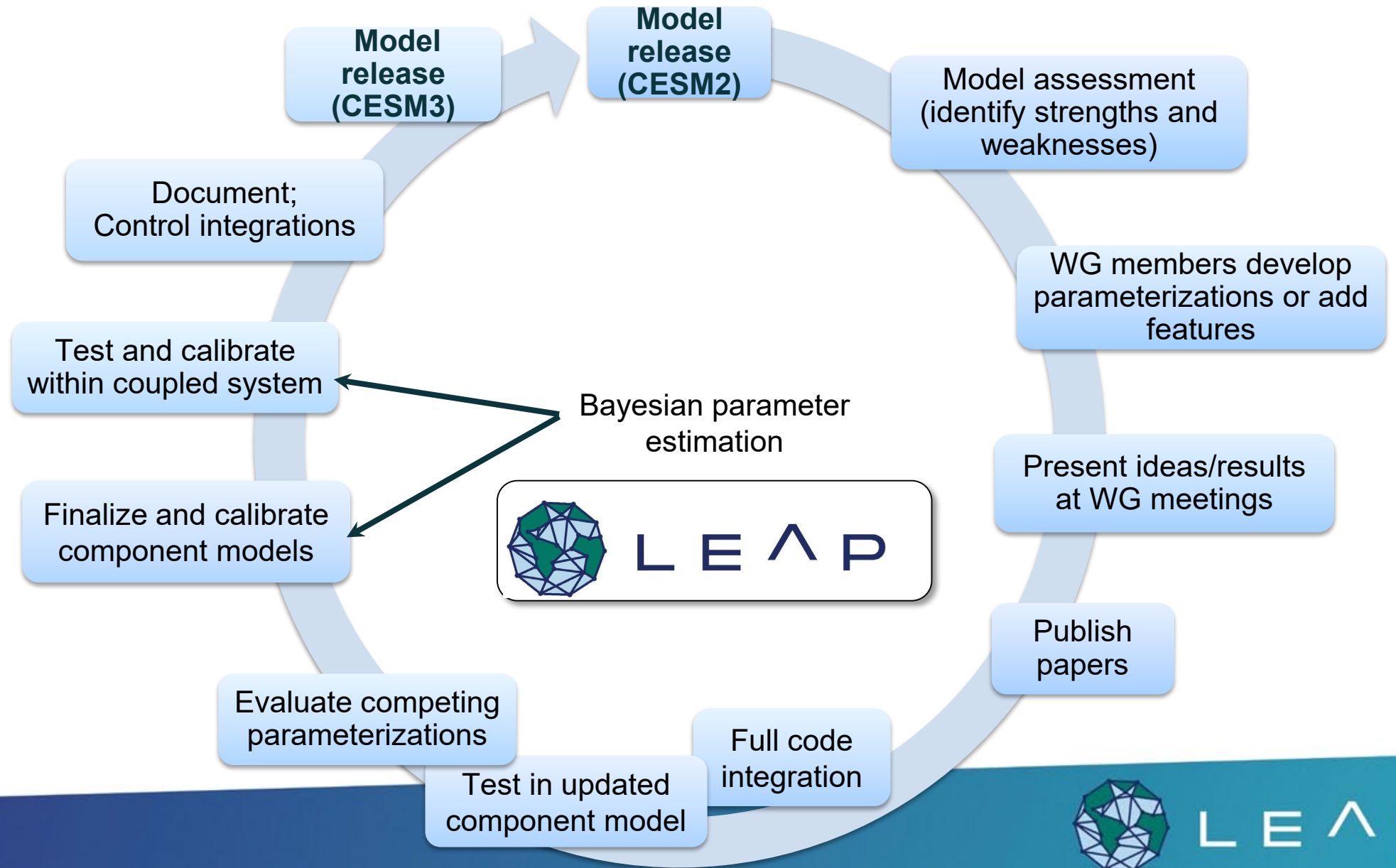
# LEAP integration into CESM development process

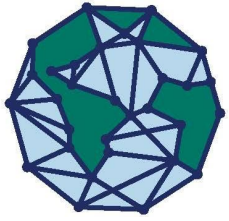


# LEAP integration into CESM development process



# LEAP integration into CESM development process

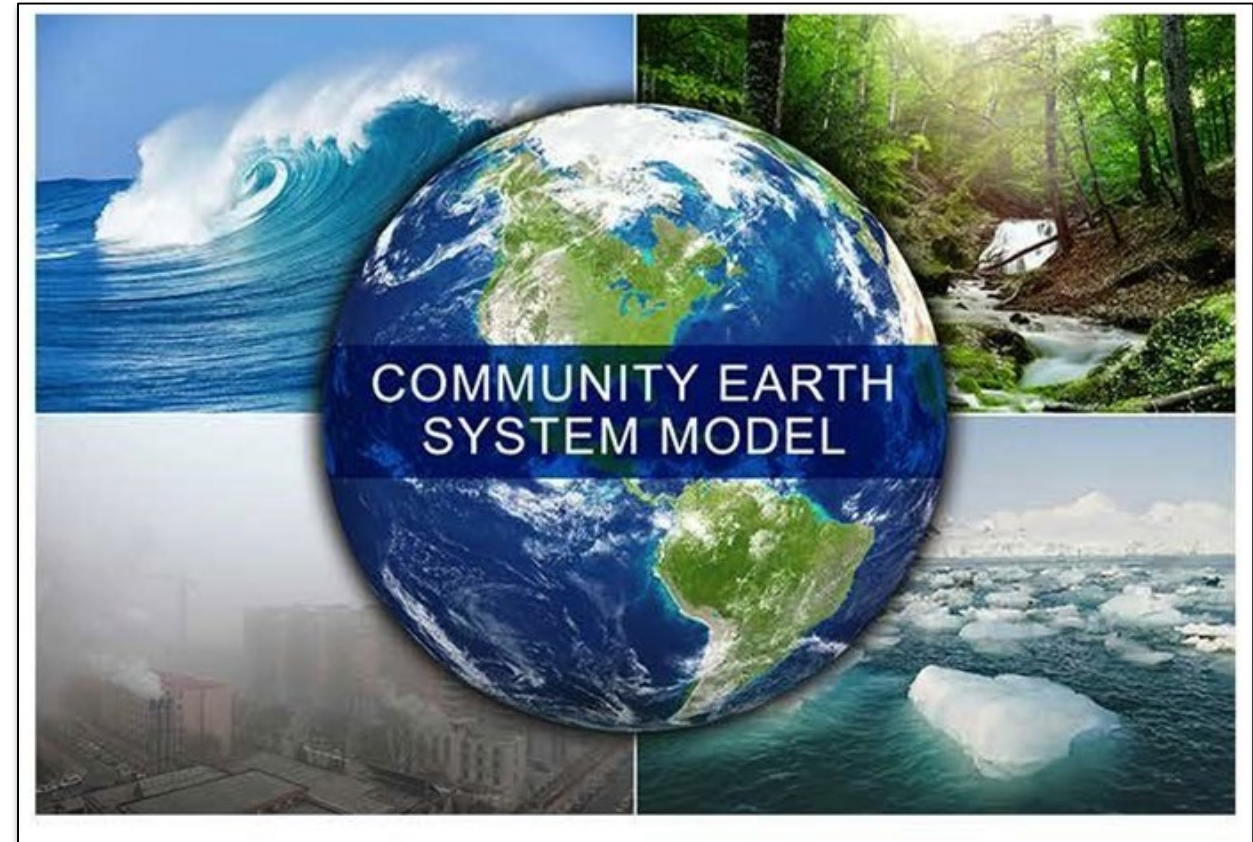




LEAP



- LEAP provides transformative advances in Earth System Modeling for the NSF-supported CESM
- Partnership with NCAR ensures that LEAP development is integrated into CESM for community use





Questions?  
Feedback?

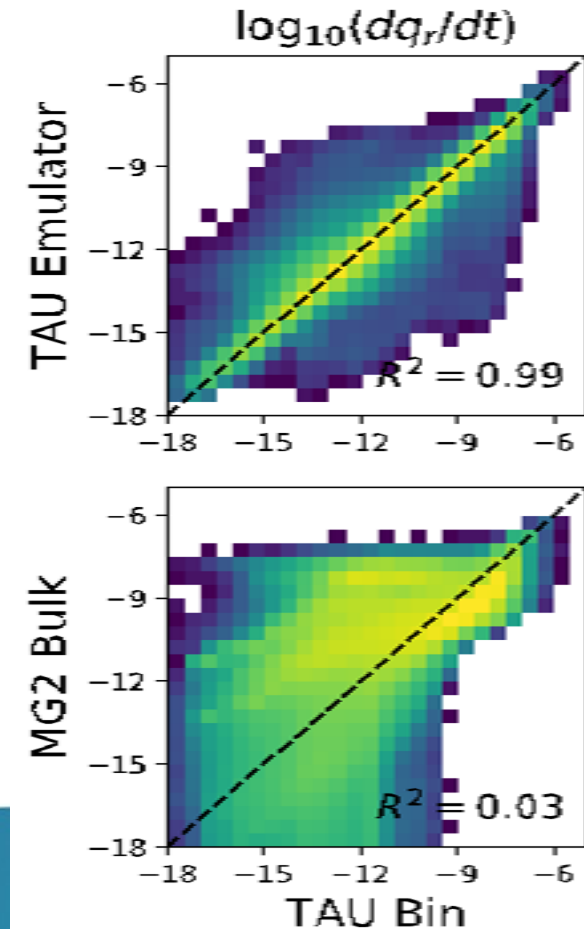
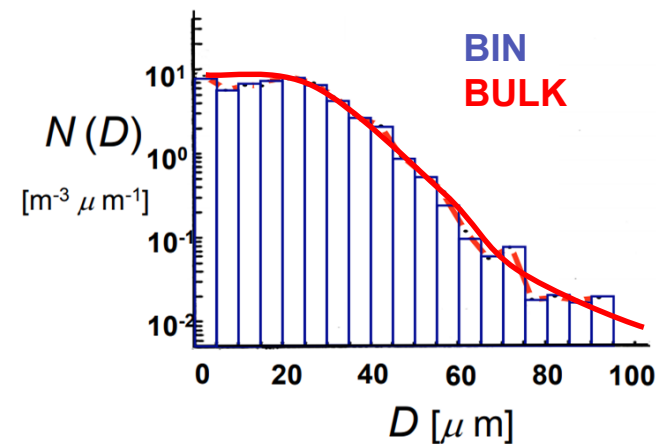


# ML parameterization pilot study in CESM

1. NCAR has a small Machine Learning group (led by David Gagne, also in LEAP) and some pilot projects, including with CESM
2. We have learned a few things about how to integrate ML code into CESM with pilot. This presents opportunities, but we have not solved all the challenges (LEAP will help!)
3. NCAR is Primed with basic infrastructure in CESM to do this and have an example (cloud microphysics, shown) which is interesting and being evaluated now
4. Pilot suggests ML work in CESM low risk, but high opportunity to advance further

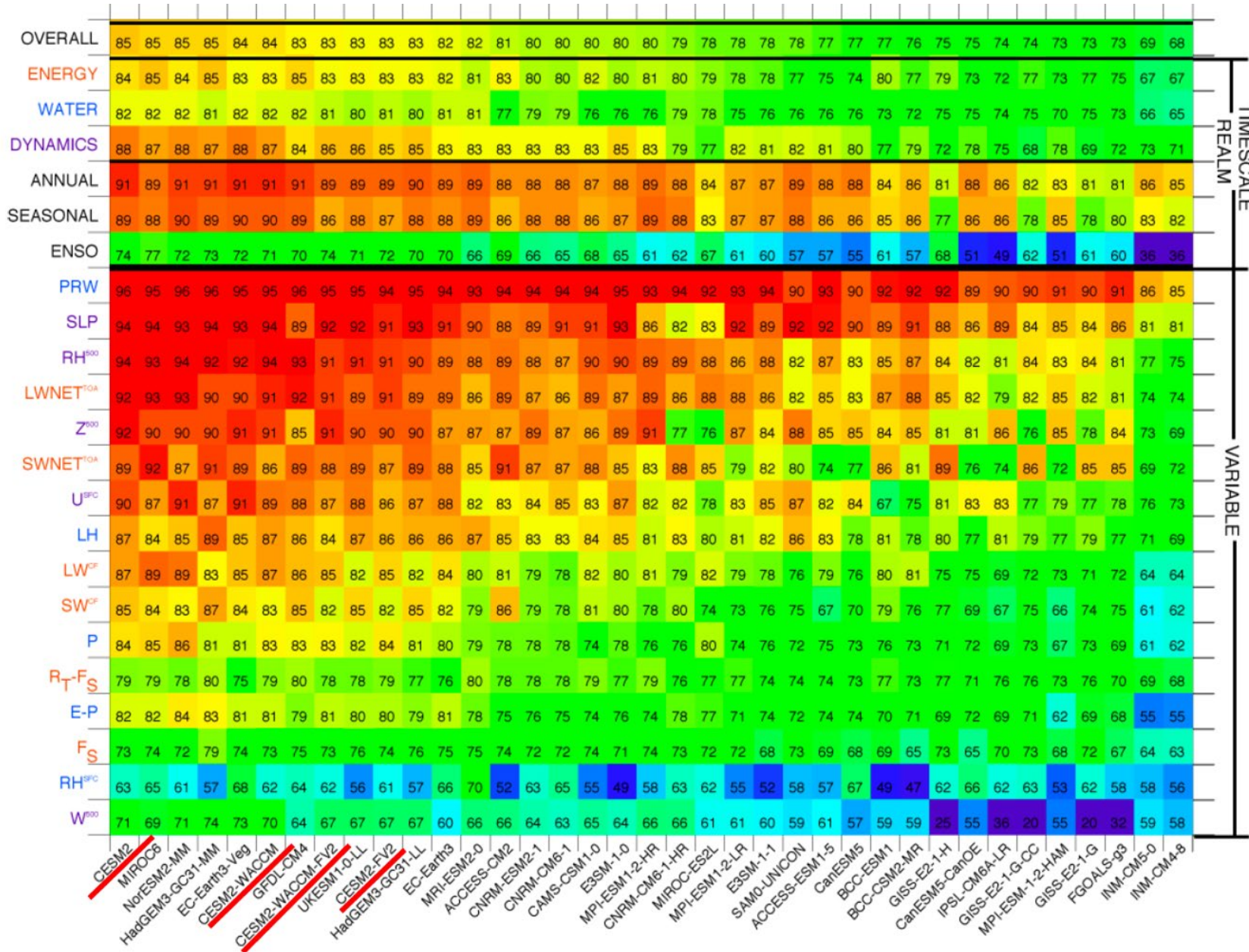
## Example

- Can we do the warm rain process better?
- Replace **BULK** precipitation processes with detailed treatment from **BIN** microphysics code.
- Run inside CAM, then emulate results
- Preliminary Results: ML code works, is pretty well optimized, reproduces detailed code



# Assessing quality of CESM simulations

Model Assessment Toolkit

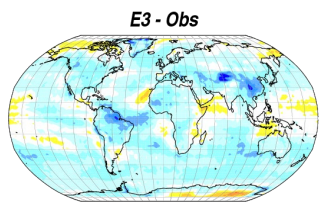
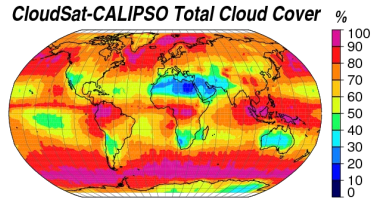


CESM is among best performing models, but

...

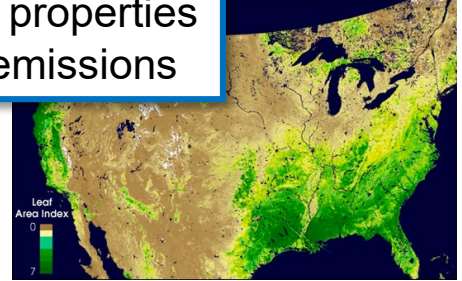
- Persistent biases and other limitations remain
- Stakeholders struggle to effectively understand and use model data

# GISS Modeling is unique in how it uses NASA data



Model inputs

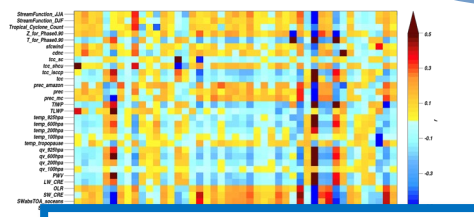
Land surface properties  
topography, emissions



Global and regional  
evaluation across multiple  
variables & teleconnections

Model  
evaluation

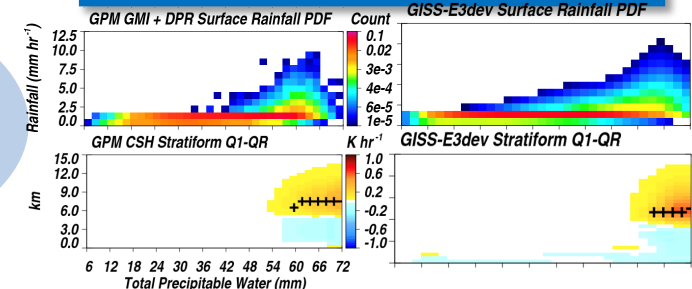
Parameterization  
development



Perturbed physics ensembles plus  
Machine Learning to match satellite  
metrics (with uncertainties) across  
40+ model parameters

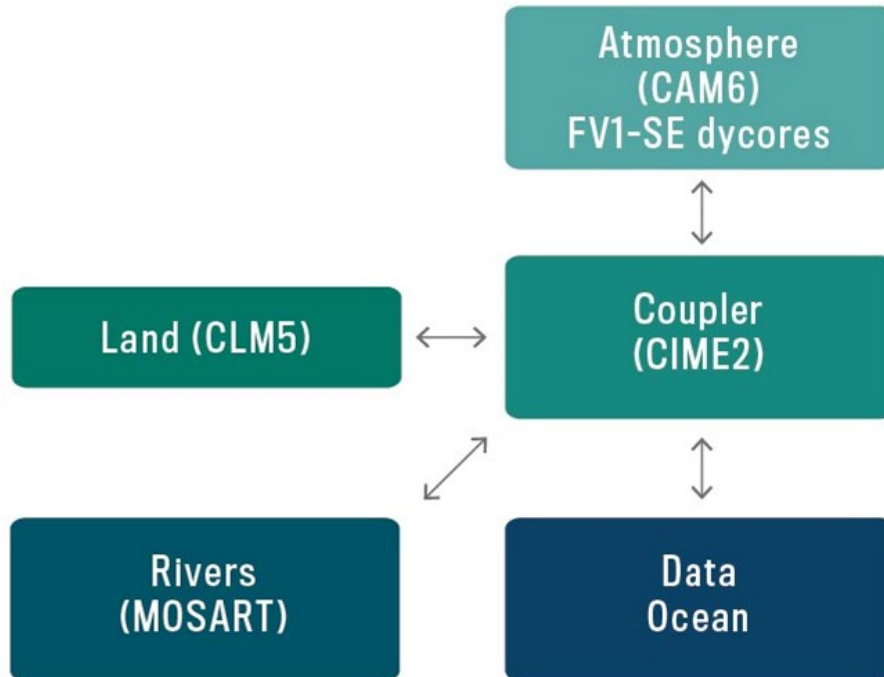
Global  
tuning

Process-related diagnostics  
from low earth orbit and flight  
campaigns

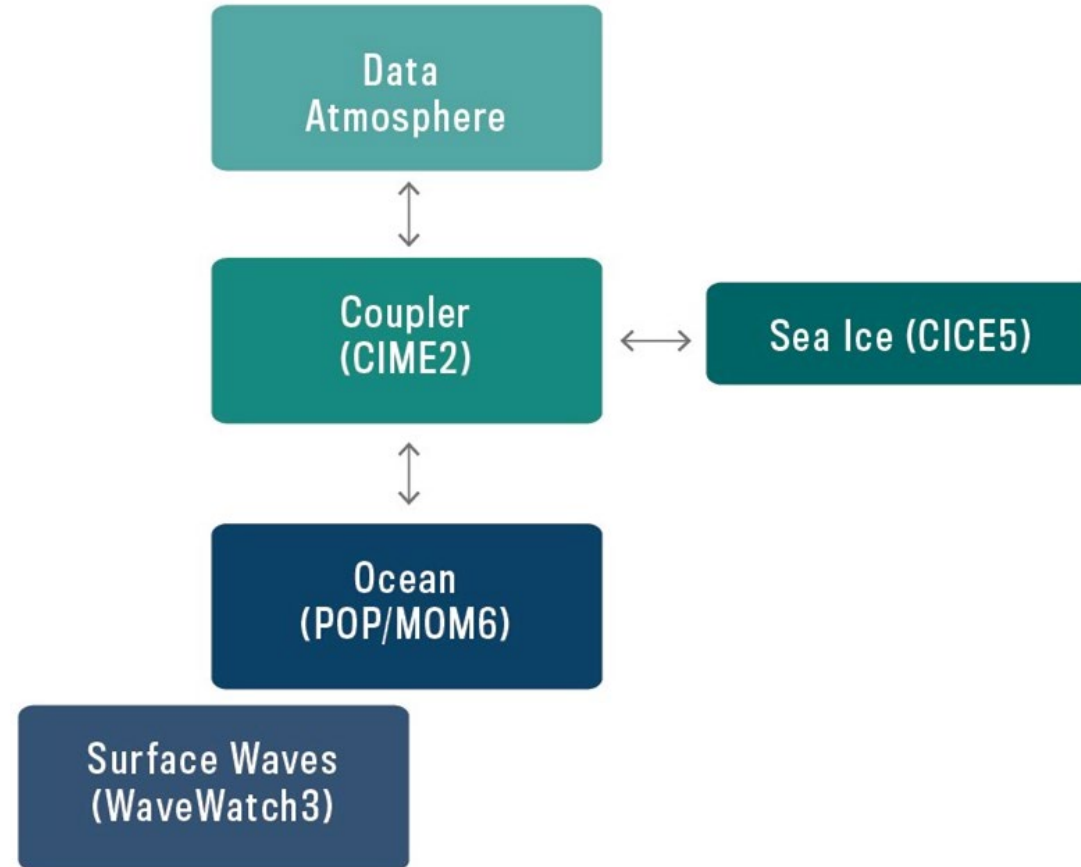


# Community Earth System Model (CESM2)

- Highly flexible system



# Community Earth System Model (CESM2)



- Highly flexible system





Questions?  
Feedback?





# Facilities + Physical Infrastructure



Deputy Director  
**Galen McKinley**

## Professional Experience and Education

- Professor of Earth + Environmental Sciences, Columbia
  - Previously Professor of Atmospheric + Oceanic Sciences, Wisconsin
- PhD, Climate Physics + Chemistry, MIT

## Expertise

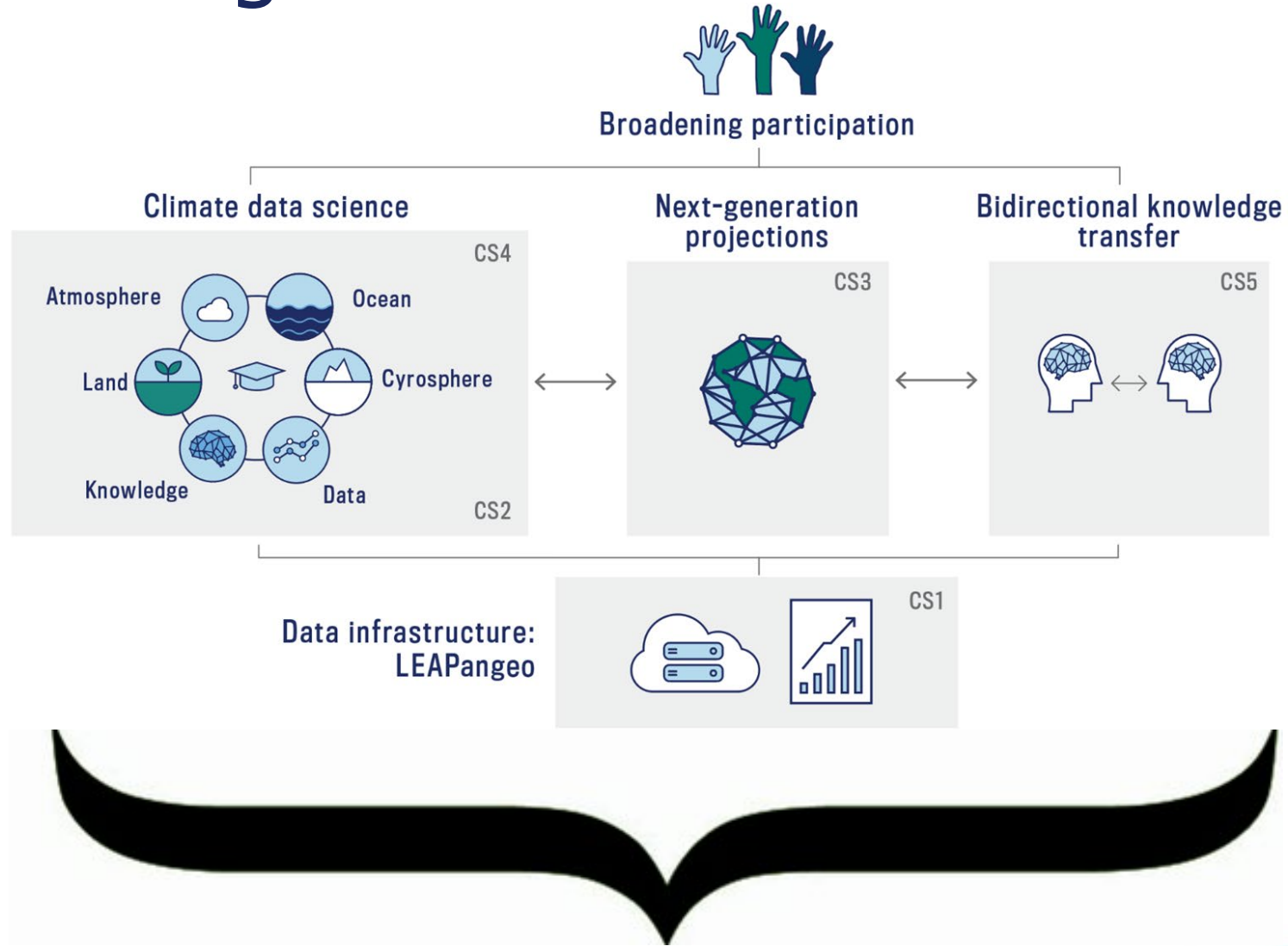
- Ocean carbon cycle, feedbacks with climate
- Ocean + climate modeling
- Data science

## Synergistic Activities

- The Oceanography Society, Councilor, 2020-present
- JASON Advisory Group, 2015-present
- Diversity Committee chair, EES department, 2019-present
- Co-chair for undergrad education, Wisconsin AOS, 2010-15

# Implementing our Vision

*LEAP forward in the reliability, utility, and reach of climate projections through synergistic innovations in data science and climate science.*



# My Presentation Will Explain:

- Institutional commitments necessary for achieving cross-institutional and sustained impact
- LEAP's dedicated headquarters spaces' role in broadening participation + transferring knowledge

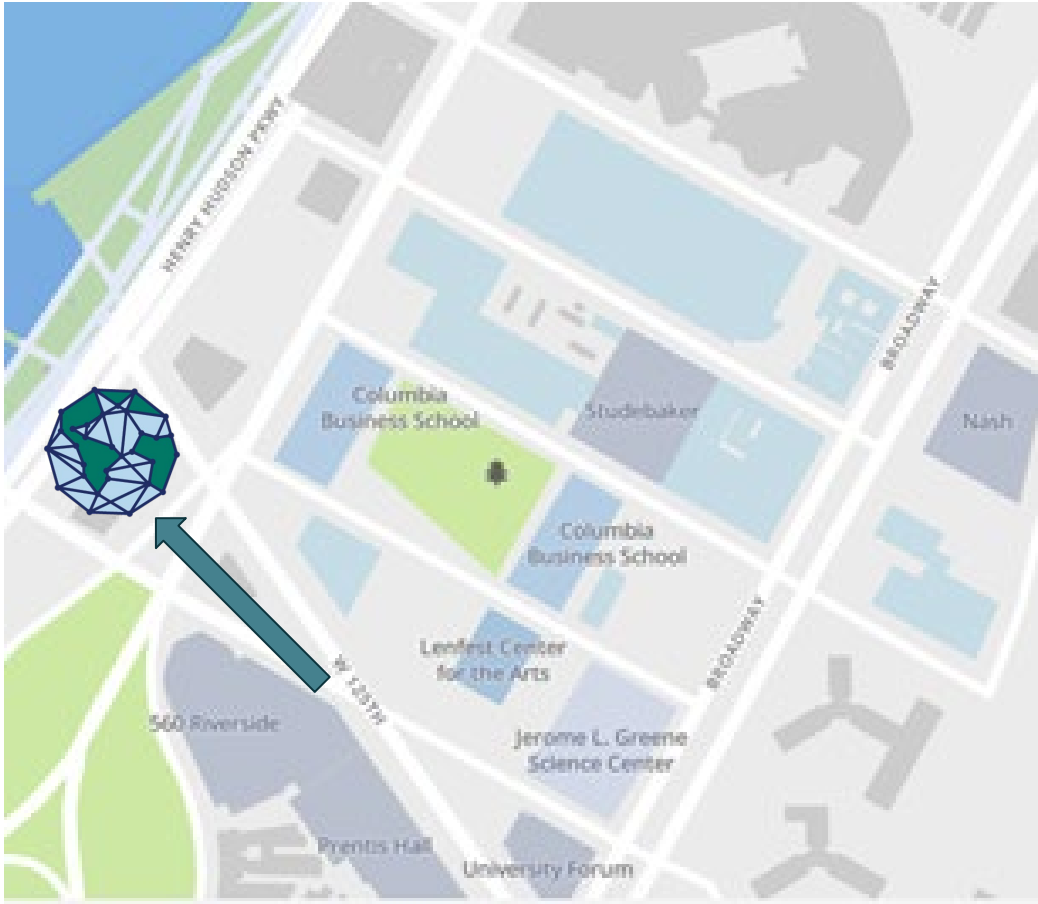
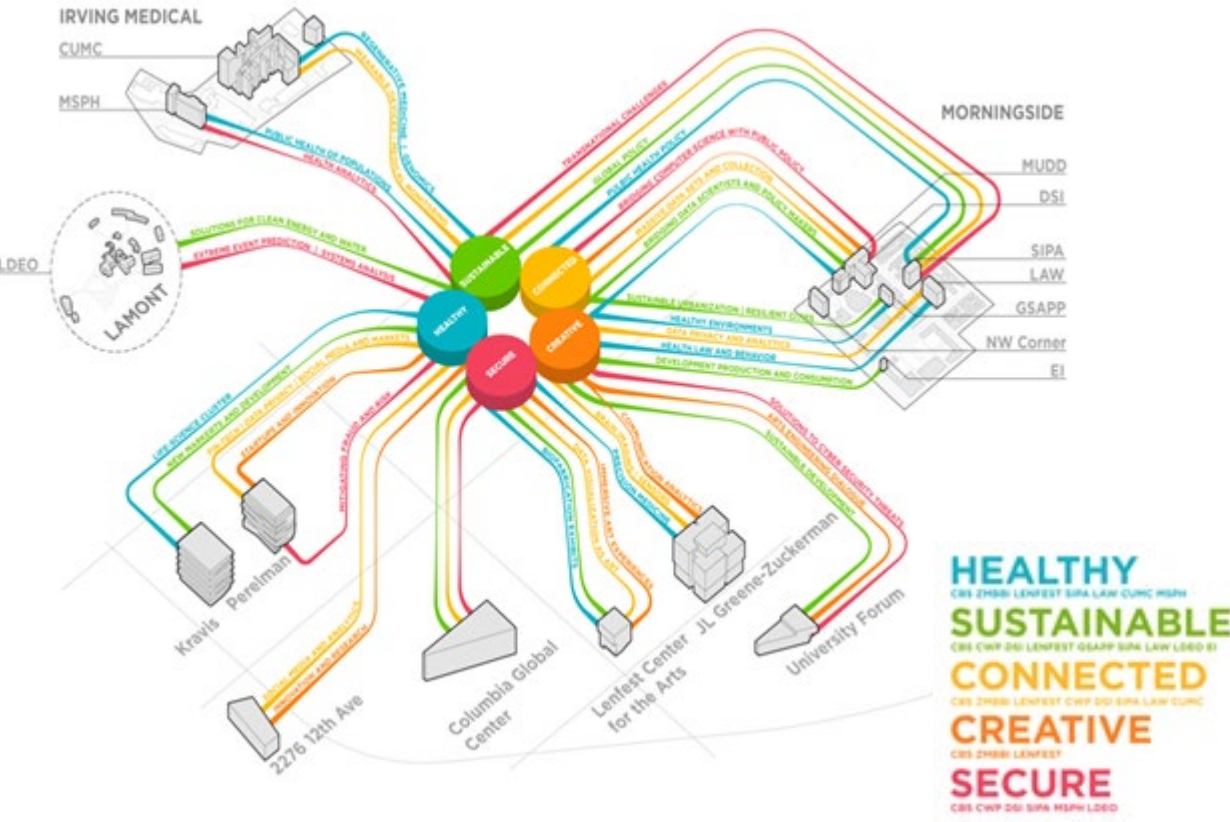


# LEAP's Headquarters are Large + New



- **Space**
  - 3 offices
  - 10 workstation cubicles
  - 16-person conference room
  - 70-person training space
- Opening December 2021

# LEAP is Strategically Placed in the New Manhattanville Campus



# LEAP will Leverage Future Columbia Climate School's Visibility

- **July 2020:** Trustees unanimously approve future creation of Columbia Climate School
  - First school ever devoted exclusively to climate
- Centralizing coordination + visibility of existing efforts across Columbia
  - Potentially including LEAP
- 5 LEAP members involved in Climate School pre-planning
  - Chief Equity Officer Cogburn on new Design Committee



# Lamont-Doherty Earth Observatory Commits Work + Meeting Spaces

## Campus Square Footage

- Buildings: 378k
- Offices: 82k
- Labs: 98k
- **Auditoriums Sizes for LEAP Meetings (by # seats):**  
200, 180, 100, 50, 25
- Free shuttle transportation between  
Morningside + Lamont campuses
  - 10x/ day
- All climate science postdocs + students  
have committed workstations
  - In addition to workstations at LEAP's  
Manhattanville headquarters



# LEAP Researchers take Hold Short-, Medium-, and Long-term Visits to Federal Labs

## NASA GISS

- 4 blocks south of Columbia
- 40 Columbia researchers semi-permanently visiting
  - Including 3 LEAP senior personnel
- **Since 2010:** 100 Columbia semi-permanent visits
- Shared offices distributed as projects come online
- Will permanently host 3 LEAP research scientists

## NCAR

- 100+ short (3-week) researcher visitations per year
- 5-10 long faculty visitations per year
- Dedicated desk space + supercomputer access to all visitors
- LEAP: 4-6 grad students + 4-6 postdocs visiting NCAR each summer
  - Travel + housing in NSF budget request



# Partner Institutions Commit Ample Computational Resources to LEAPangeo

## NCAR Computational + Information Systems Lab

- 145k-core “Cheyenne” supercomputer
- 5.34 petaflops
- Will support LEAP through University Large Allocations + CESM Working Group Allocations
- NCAR Wyoming Supercomputing Center installed in 2022

## NASA Center for Climate Simulation

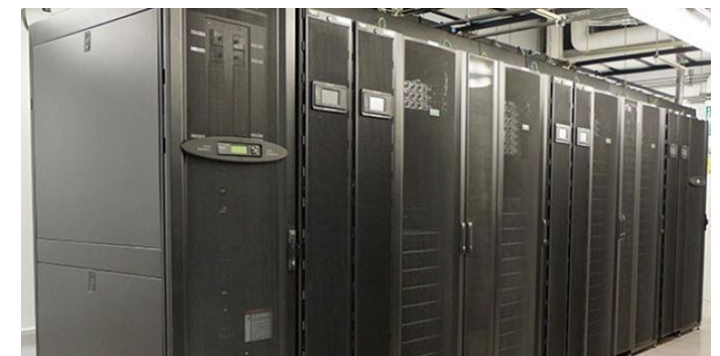
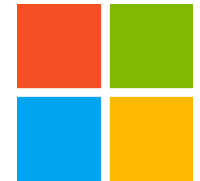
- 129k-core “Discover” supercomputer
- 6.8 petaflops
- GISS has access to 30% of capacity

## Industry Commitments

- Microsoft + Google donating sufficient cloud credits to support all LEAP activities

## Columbia High-Performance Computing

- 3.8k-core “Terremoto” cluster
- 137 nodes, 3,288 cores
- Governed by faculty committees
- Supported by 7 full-time engineers

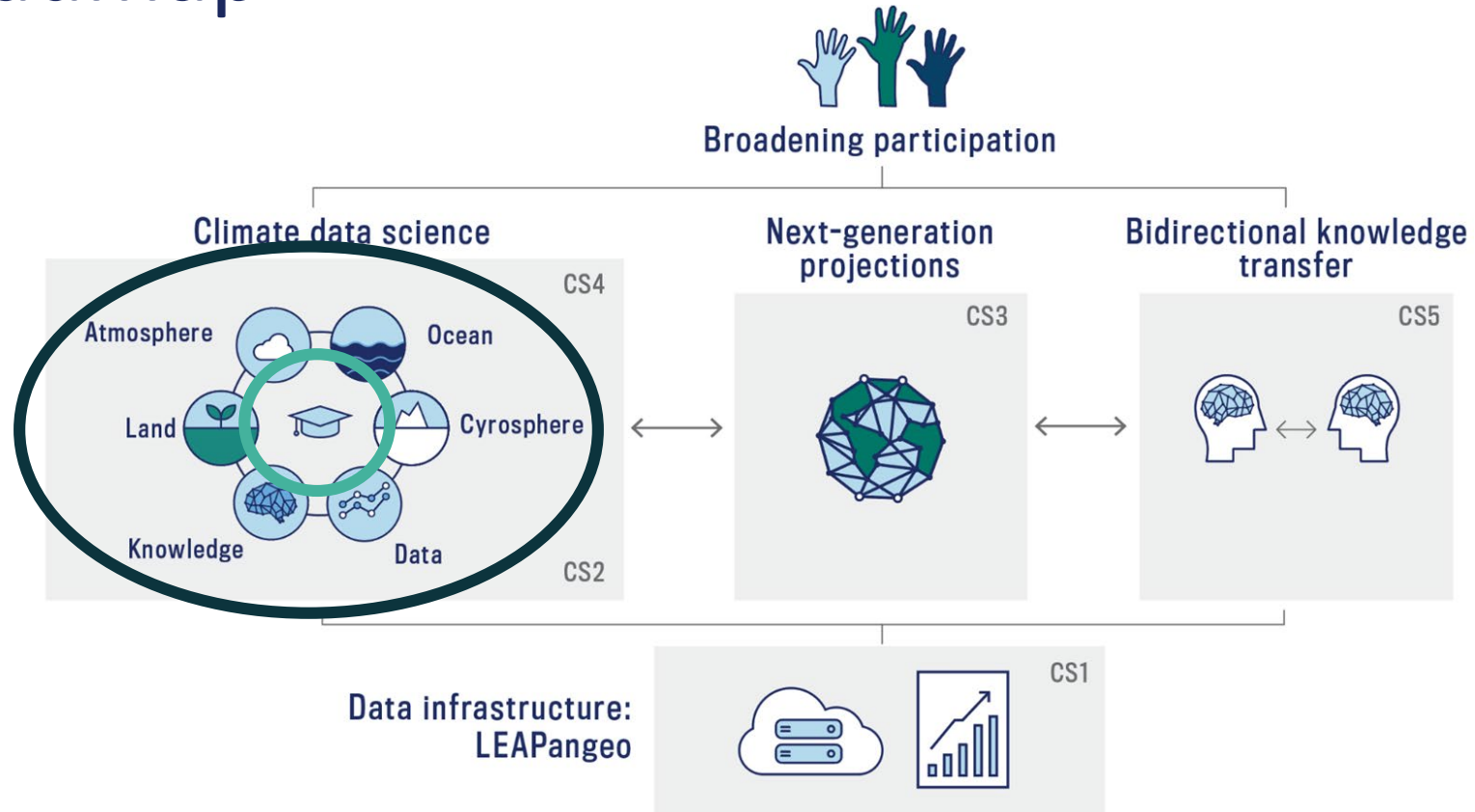


Questions?



# Integrating Research + Education

# Our Roadmap



*LEAP forward in the **accuracy, utility, and reach** of climate projections through **synergistic innovations** in data science and climate science.*



**Chief Convergence Officer  
& Education Director**  
**Tian Zheng**

- Professor + Chair of Statistics, Columbia
- PhD, Statistics, Columbia

### **Expertise**

- Spatiotemporal modeling
- Statistical machine learning
- Data science education
- Management

### **Synergistic Experiences**

- Associate Director for Education, Data Science Institute, 2017-2020
- Creator, DSI Scholars Program, 2018-2020
- Leadership Team Member, Collaboratory Program
- Faculty Advisor, Columbia Statistics Club 2016-2020

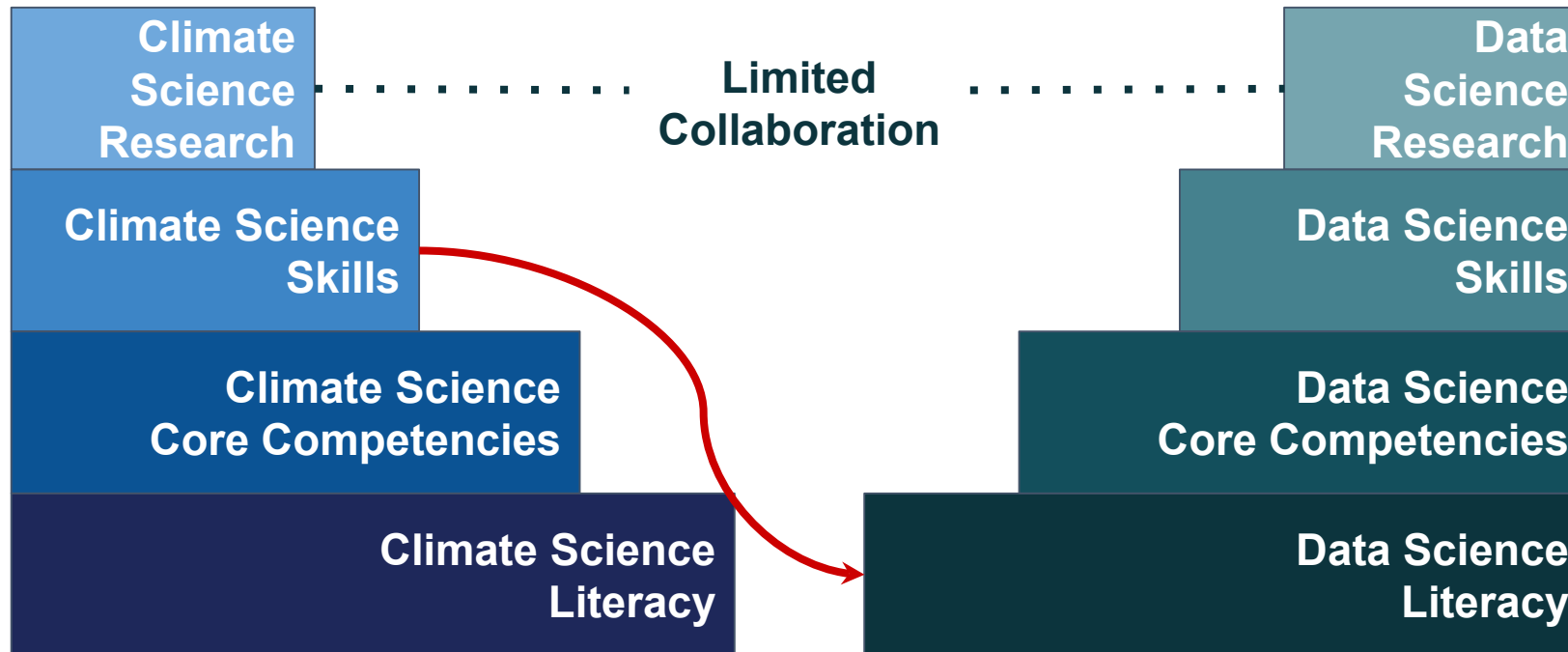


# My Presentation Will Explain:

- Integration of research + education
- Innovation of educational activities
- How education contributes to LEAP's unified mission
- How evaluation allows correction + development

# Climate Scientists + Data Scientists Face Collaboration Barriers

Disconnect between core disciplines exacerbates skills gaps + narrows pathways for convergent learning + research



# Research + Education Integration Drives Convergence



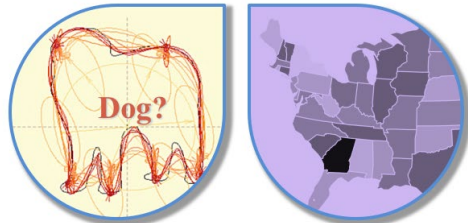
- Education *for* Research
  - **Relevant training** for convergent research
- Research *for* Education
  - **Meaningful convergent experiences** embedded in education

# Research + Education Integration Drives Convergence

Education

Research

Statistics gives you



<http://tzstatsads.github.io/>

Applied Data Science @ Columbia

## PROJECT TOXICDOCS

Columbia University and the City University of New York

Millions of Pages of Previously Classified Documents on Industrial Poisons (and Counting)

Free of Charge. Open to All.

Applied Data Science:  
A project-based learning course



# Research + Education Integration Drives Convergence

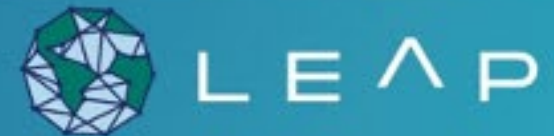
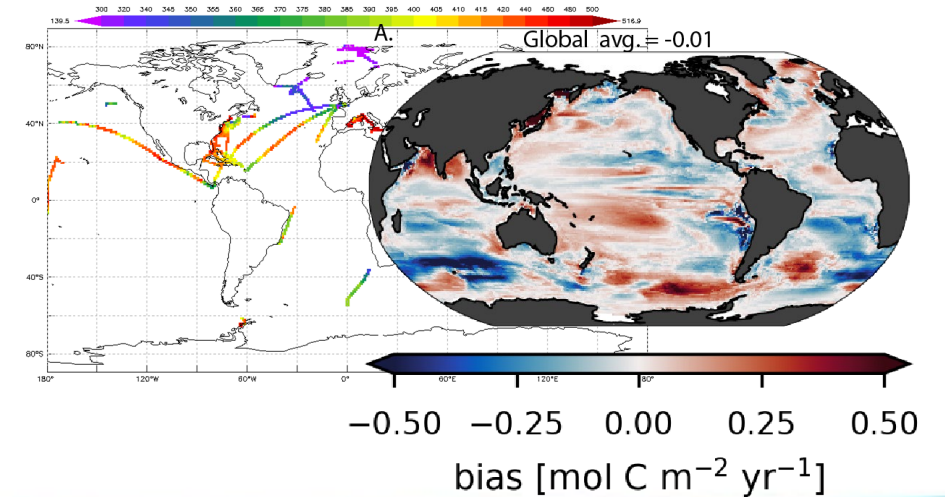
Education

Research



DATA ANALYTICS PROJECT  
DSI Scholars Program  
Enriched research experience in Data Science  
for the next generation of scholars

Data Science Capstone Project



# Next-Generation Climate Data Scientists Will Create Broad Impacts

EO1: Develop **research-intensive** undergraduate + graduate curricula, providing multiple transdisciplinary **entryways** for students from diverse backgrounds + with varied career goals.

EO2: Offer **immersive research** experiences to student, teacher, parent + workforce trainees.

EO3: **Broaden research** impact through communication, translation, human-centered design, curriculum development, evaluation + assessment.



# LEAP Creates a Climate Data Science Curriculum

Literacy

Core Competencies

Skills

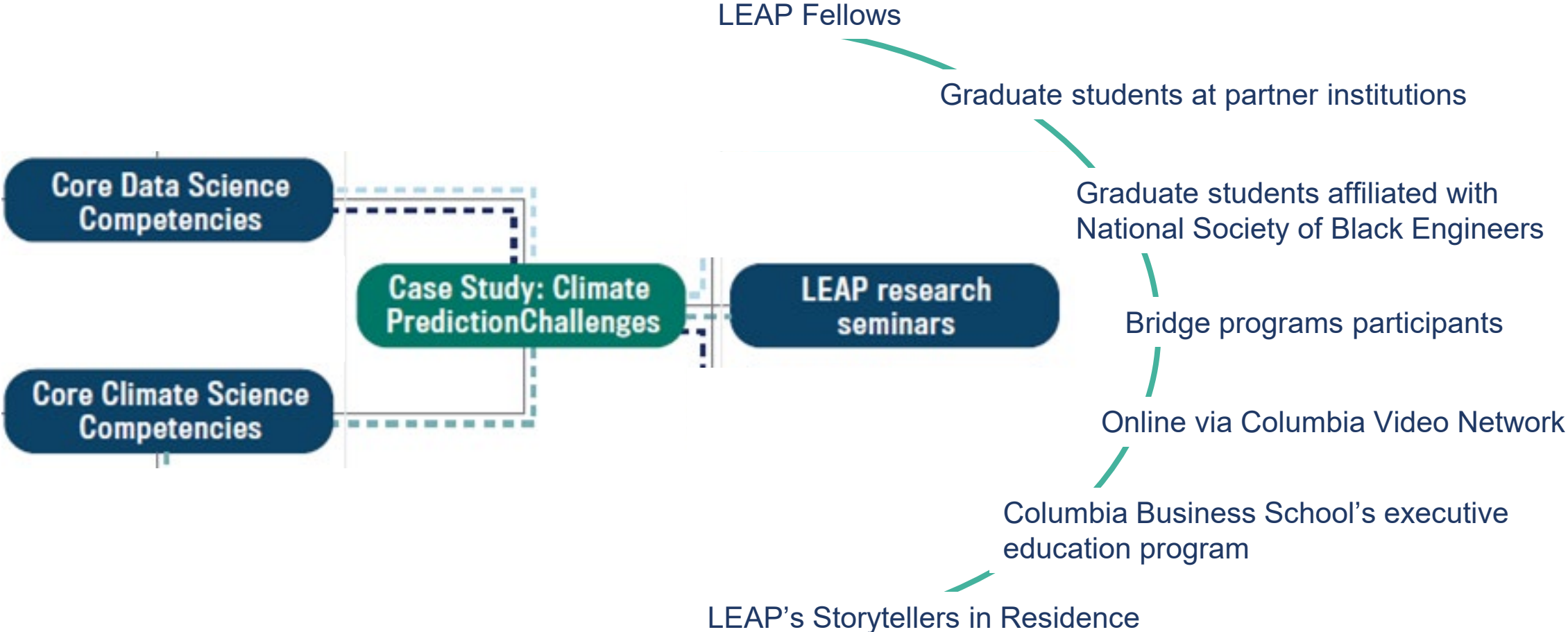
Research

Case Study:  
Climate + Big Data

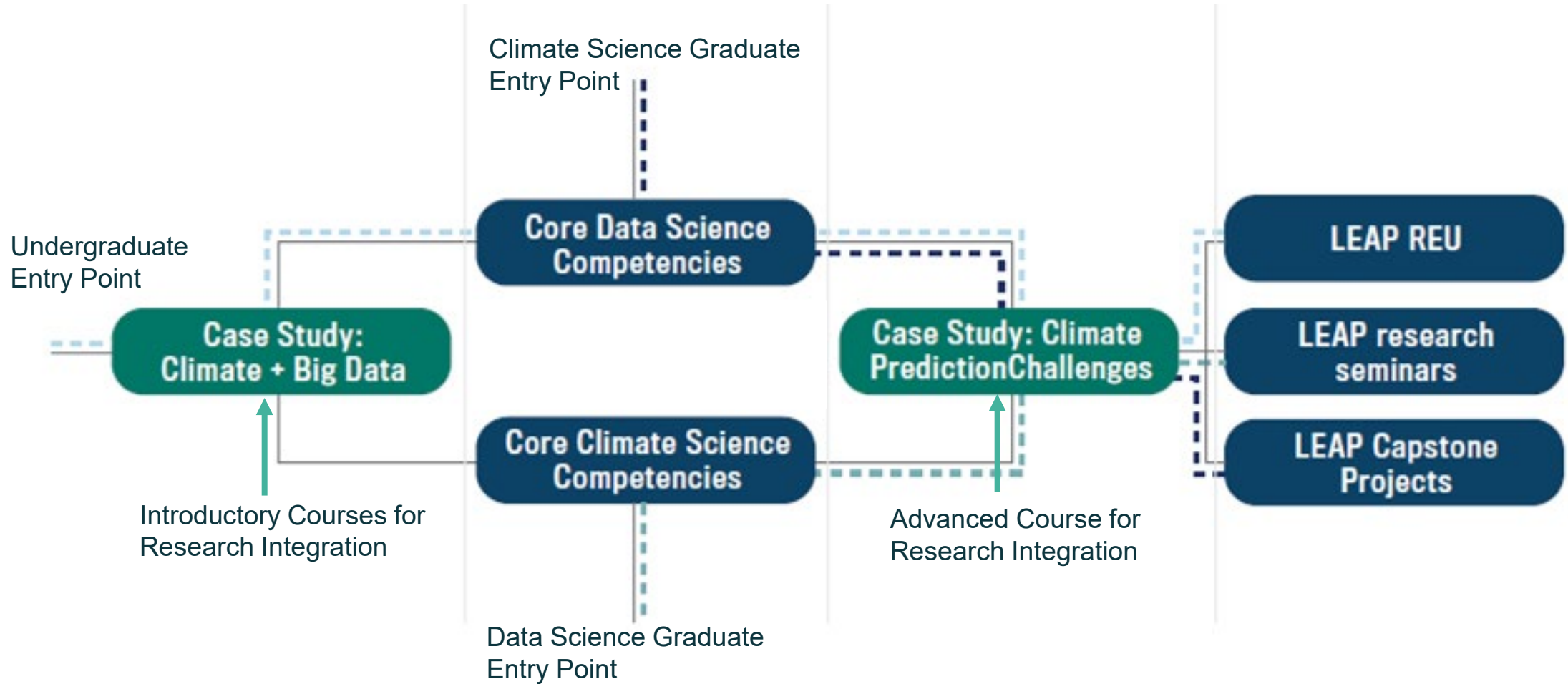


LEAP

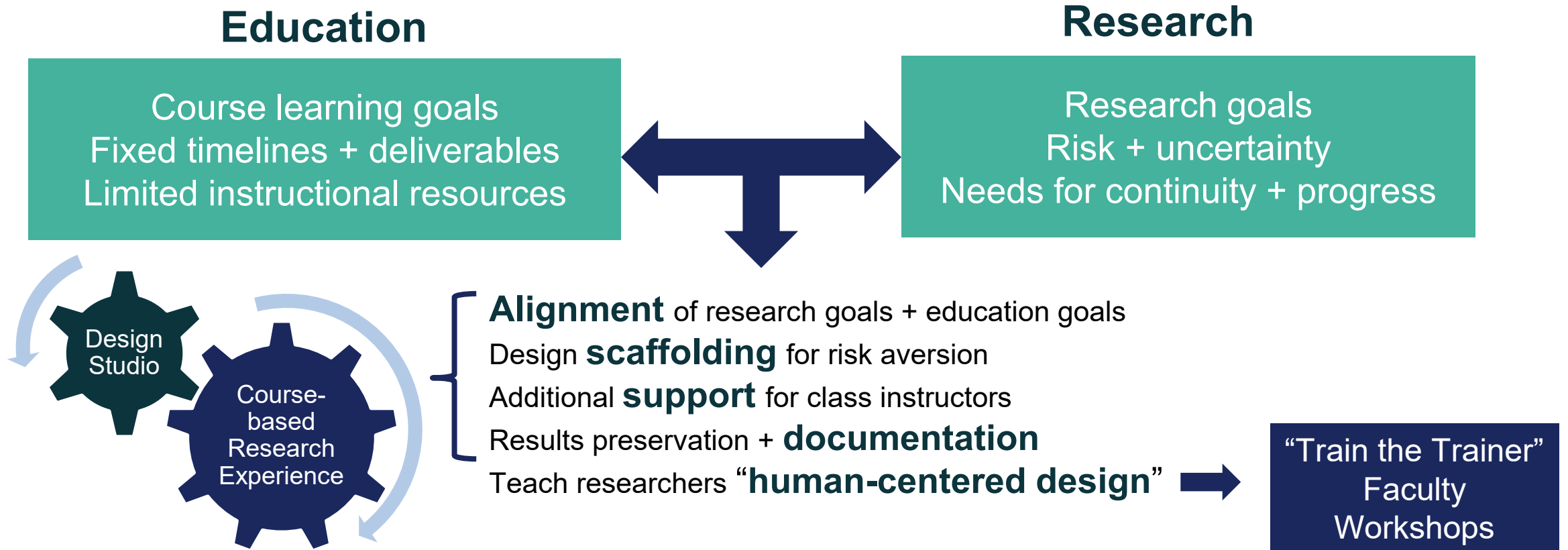
# A Graduate Certificate in Climate Data Science Reaches Diverse Trainees



# This Curriculum Offers Multiple Entry Points



# Research + Education is Integrated by Design



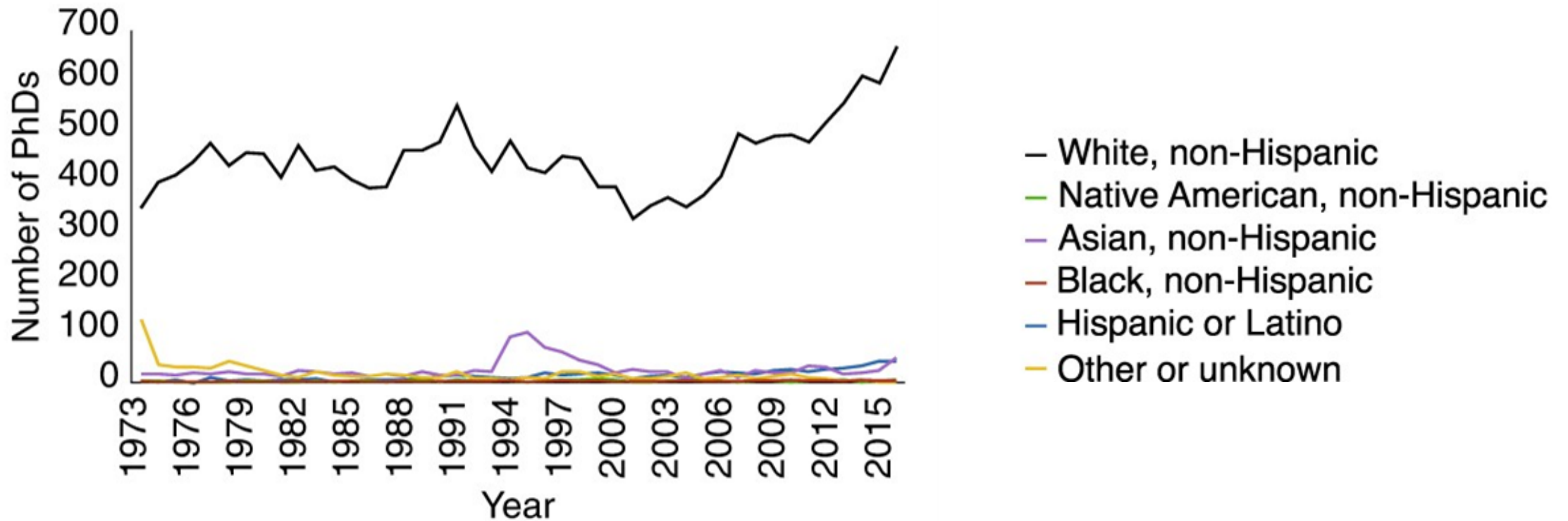
# Transdisciplinary Education Offers a Greater Quality + Quantity of Opportunities

- Inclusive curriculum
- Overcome barriers + embrace differences
- Collaboration + open science
- Fusing communities



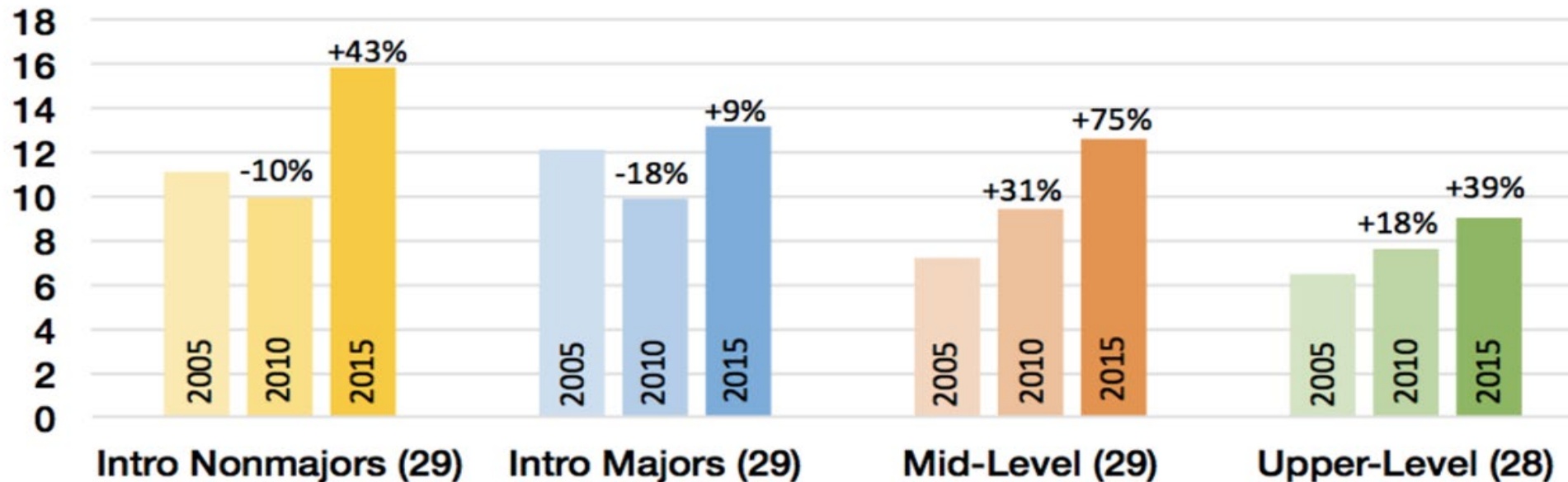
# Climate Science is Not Diverse or Equitable + Has Not Changed

**b** By race and ethnicity (subfields combined)



# Data Science Curriculum has Improved but Still Poses Barriers

## Underrepresented Minorities



Source: CRA Enrollment Survey



# Next-Generation Climate Data Scientists Will Create Broad Impacts

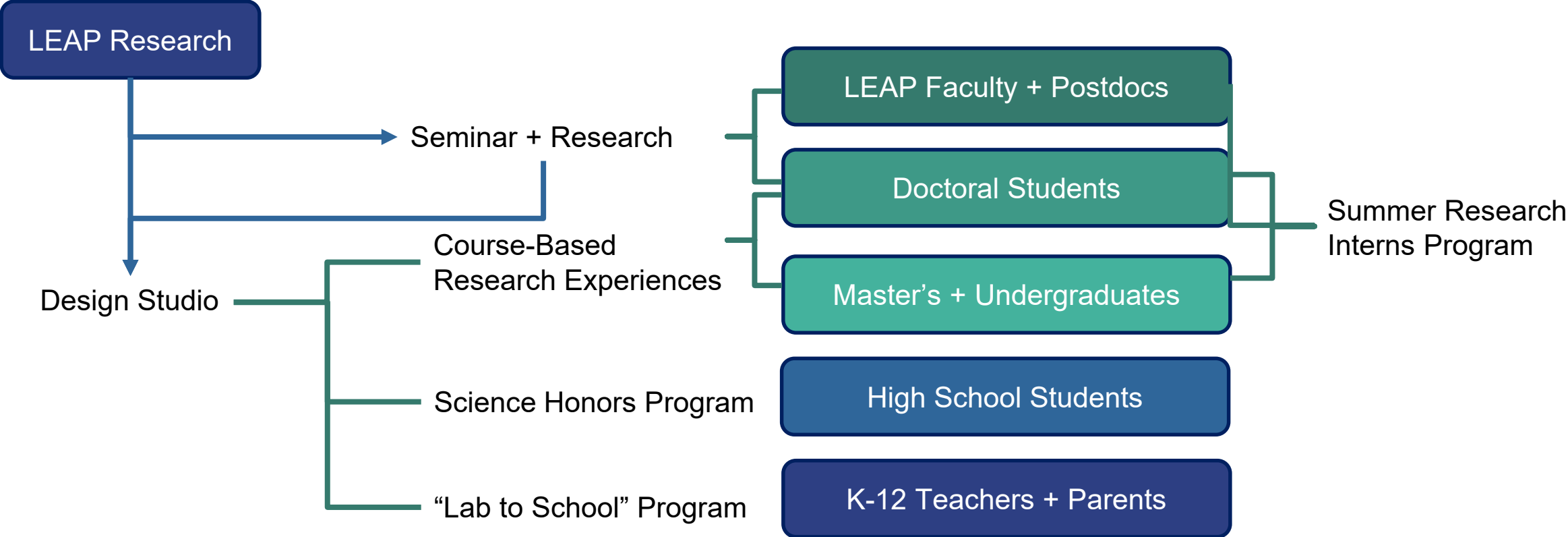
EO1: Develop **research-intensive** undergraduate + graduate curricula, providing multiple transdisciplinary **entryways** for students from diverse backgrounds + with varied career goals.

EO2: Offer **immersive research** experiences to student, teacher, parent + workforce trainees.

EO3: **Broaden research** impact through communication, translation, human-centered design, curriculum development, evaluation + assessment.



# Education Programs are Vertically Integrated



# LEAP Fellows Mentoring Plan

- 40 LEAP Fellowships across five year award
  - Fellowships renewable based on progress + strategy
- Faculty co-advising
- **Research Seminars:** research skills, communication, reproducibility + data ethics
- **Climate Justice Leadership Board:** leadership, program assessment + evaluation

# Postdoc Mentoring Plan

- 7-8 postdocs per year at Columbia + Irvine
- Faculty mentoring via bimonthly meetings
- Onboarding self-assessment
- **Office of Postdoctoral Affairs**
  - Orientation
  - Career exploration
  - Job preparation + interview skills
  - Publications + presentations



# Next-Generation Climate Data Scientists Will Create Broad Impacts

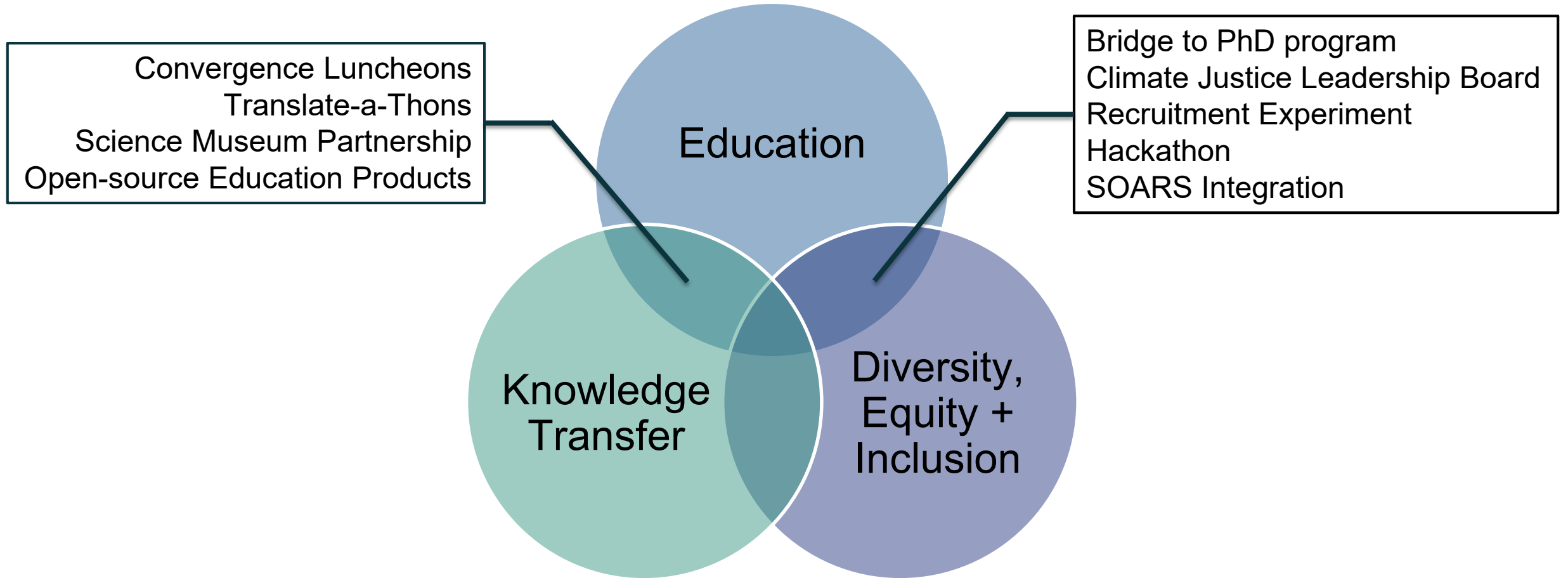
EO1: Develop **research-intensive** undergraduate + graduate curricula, providing multiple transdisciplinary **entryways** for students from diverse backgrounds + with varied career goals.

EO2: Offer **immersive research** experiences to student, teacher, parent + workforce trainees.

EO3: **Broaden research** impact through communication, translation, human-centered design, curriculum development, evaluation + assessment.



# LEAP Integrates Education, Broadening Participation + Knowledge Transfer



# LEAP Partners with Established Diversity Bridge Programs to Extend Reach

## Bridge to PhD Program



Columbia University's Bridge to the Ph.D.  
Program in STEM

## SOARS Integration



# LEAP Delivers Open-Source Educational Products

- Curricular materials + data sets
- Formative assessment tools + rubrics
- Designed course-based research experiences
- “Train the trainer” workshop materials
- “Lab to Schools” summer institute materials
  - Translated into Spanish
- **LEAP’s Open-Source Commitment:** Explained in Management + Administration presentation

# LEAP Disseminates Convergent Educational Best Practices

- Publishing in **journals** on climate science education + data science education
- Present at **professional conferences + meetings** for academic leaders
- “Lab to School” **webinar** series



# A LEAP Year in Education

Participants	Fall				1	Spring				Summer		
	9	10	11	12		2	3	4	5	6	7	8
LEAP Fellows and Postdocs	Mentored Research and Professional Development											
Undergraduate students										Research Interns		
Graduate students	Courses/Seminars					Courses/Seminars						
Undergraduate students	Courses/Seminars					Courses/Seminars						
Design Studio Interns	Curricular Development					Curricular Development						
Faculty Instructors and Researchers				Train the Trainers					Train the Trainers			
High School Students	SHP					SHP						
K-12 teachers										Lab to School		
Public	Science Museum Partnership											

# LEAP Leverages + Develops Existing Strengths



# LEAP Leverages + Develops Existing Strengths

COLLABORATORY

AT COLUMBIA UNIVERSITY  
Interdisciplinary Data Science Literacy

☑ Curriculum Development



CENTER *for* SUSTAINABLE FUTURES

TEACHERS COLLEGE COLUMBIA UNIVERSITY

☑ K-12 Education



# LEAP Leverages + Develops Existing Strengths



Online Collaboration


**Pangeo:** Open-Source Big Data Climate Science Platform



✓ Online Learning



# LEAP Leverages + Develops Existing Strengths

 COLUMBIA | ENGINEERING  
Student Research Program

REU

**SUMMER@SEAS**

**Columbia University  
Science Honors Program**

High School  
Students



Bridge Programs



Columbia University's Bric  
Program in STE



# Implementation: Year 1

## - Research + Course Development

### YEAR 1

Participants	Fall				1	Spring				Summer		
	9	10	11	12		2	3	4	5	6	7	8
LEAP Fellows and Postdocs	Mentored Research and Professional Development											
Undergraduate students											Research Interns	
Graduate students	Offer Existing Courses in Core Competencies											
Undergraduate students	Develop and Pilot New Course											
Design Studio Interns	Curricular Development				Curricular Development							
Faculty Instructors and Researchers												
High School Students												
K-12 teachers												
Public												

#### Open-Source Products

- Data sets
- Research products
- Course-based REU designs

Putting together **proposal** for Certificate for Professional Achievement in Climate Data Science



# Implementation: Year 2 - Scale Up + Expand

## YEAR 2

Participants	Fall				1	Spring				Summer		
	9	10	11	12		2	3	4	5	6	7	8
LEAP Fellows and Postdocs	Mentored Research and Professional Development											
Undergraduate students											Research Interns	
Graduate students	Courses/Seminars				Courses/Seminars							
Undergraduate students	Courses/Seminars				Courses/Seminars							
Design Studio Interns	Curricular Development				Curricular Development							
Faculty Instructors and Researchers				Train the Trainers					Train the Trainers			
High School Students	Science Museum Partnership											
K-12 teachers											Lab to School	
Public	Science Museum Partnership											

### Open-Source Products

- Data sets
- Research products
- Course-based REU designs
- Course evaluation materials
- Faculty + teacher training materials

Launch Certificate for Professional Achievement in Climate Data Science

LEAP and NYC

Bridge Programs

# Implementation: Stable State

## YEARS 3-5

Participants	Fall				1	Spring				Summer		
	9	10	11	12		2	3	4	5	6	7	8
LEAP Fellows and Postdocs	Mentored Research and Professional Development											
Undergraduate students										Research Interns		
Graduate students	Courses/Seminars					Courses/Seminars						
Undergraduate students												
Design Studio Interns	Curricular Development					Curricular Development						
Faculty Instructors and Researchers				Train the Trainers					Train the Trainers			
High School Students	SHP					SHP						
K-12 teachers										Lab to School		
Public	Science Museum Partnership											

### Open-Source Products

- Data sets
- Research products
- Course-based REU designs
- Course evaluation materials
- Faculty + teacher training materials

Launch Certificate for Professional Achievement in Climate Data Science

LEAP + NYC

Bridge to PHD

Online

SHP: Science Honors Program

# Education Evaluation: Student Outcomes

Assessment Needs	Evaluation Methods
Knowledge + skills in climate data science	Assessment of the quality of <ul style="list-style-type: none"><li>● Climate prediction challenges <b>project products</b>.</li><li>● Questions when <b>approaching a new climate data science topic</b>.</li><li>● <b>Decision-making process</b>.</li></ul>

# Education Evaluation: Program Outcomes

Assessment Needs	Evaluation Methods
Education Success	<ul style="list-style-type: none"><li>● Aggregated statistics on <b>student performance</b></li><li>● Aggregated statistics on <b>student participation</b></li><li>● Surveys + interviews of students' <b>learning experience</b></li></ul>



# Formative Reviews + Course Correction

## Student Outcomes

- **Frequency:** Once per semester
- **Goal:** Focus on outcomes contributing to convergence
- **Distribution:** Reviewed by Convergence Subcommittee

## Program Outcomes

- **Frequency:** Once per year
- **Goal:** Determine program resource allocation
- **Distribution:** Shared in annual newsletter



# Next-Generation Climate Data Scientists Will Create Broad Impacts

EO1: Develop **research-intensive** undergraduate + graduate curricula, providing multiple transdisciplinary **entryways** for students from diverse backgrounds + with varied career goals.

EO2: Offer **immersive research** experiences to student, teacher, parent + workforce trainees.

EO3: **Broaden research** impact through communication, translation, human-centered design, curriculum development, evaluation + assessment.

# LEAP Aspires to Create a New Discipline: Climate Data Science

6 Disciplinary Attributes *	LEAP's Translation
Defined Object of Research	Data-Driven Parameterization + Projection in Atmosphere, Ocean, Land, Cryosphere

\* Source: Inside Higher Ed, *6 Attributes of an Academic Discipline*  
<https://www.insidehighered.com/blogs/technology-and-learning/6-attributes-academic-discipline>



Questions + Feedback?



# Developing Human Resources + Broadening Participation



Center Director and  
Diversity Co-Director  
**Pierre Gentine**

- Associate Professor of Earth + Environmental Engineering, Associate Professor of Earth + Environmental Sciences, Columbia University
- PhD, Civil + Environmental Engineering, MIT

### **Expertise**

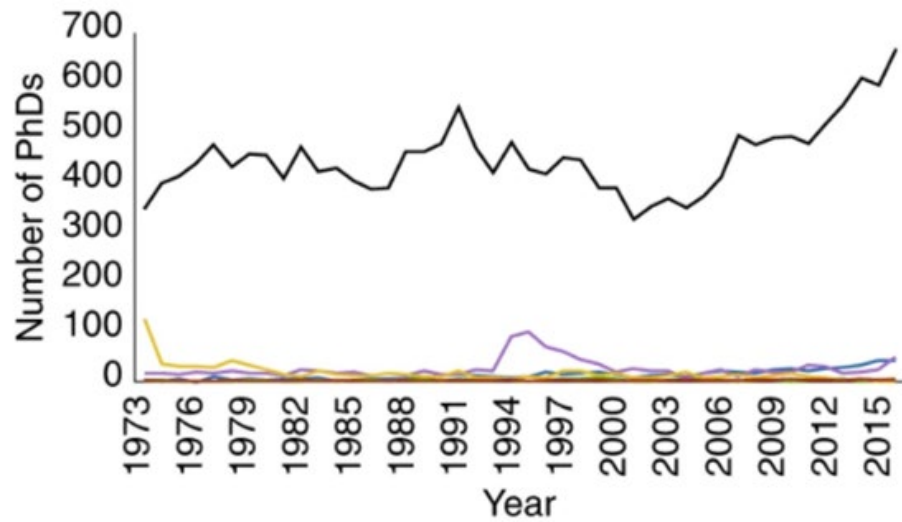
- Hydrologic cycle
- Land-atmosphere interactions
- Machine learning for geoscience

### **Other Synergistic Experience**

- Co-Founder + Co-Lead, Data Science Working Group, US CLIVAR
- Co-Lead, NOAA Drought Task Force
- Member, Working Group on Seasonal to Interannual Prediction, World Climate Research Program



# Geoscience's Homogeneity Impedes Progress



- White, non-Hispanic
- Native American, non-Hispanic
- Asian, non-Hispanic
- Black, non-Hispanic
- Hispanic or Latino
- Other or unknown

- Geosciences are not diverse, equitable, or inclusive
- Racial diversity has not increased in **40 years**
- Underrepresented minorities exit at higher rates
- Prior efforts have been unsuccessful
  - **Why** is there underrepresentation?



# Inequity in Geoscience is Highly-Visible

The New York Times

## *Earth Science Has a Whiteness Problem*

Barely 10 percent of doctoral degrees in the geosciences go to recipients of color. The lack of diversity limits the quality of research, many scientists say.



*“A lack of diversity + inclusion is the **single largest cultural problem** facing the geosciences today” - Dr. Kuheli Dutt, Lamont-Doherty Earth Observatory*

*“[The] lack of representation in turn **affects the quality + focus of earth science research, especially on climate change**”  
- Emma Goldberg, NYTimes*



LEAP



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**

- Associate Professor of Social Work, Columbia
  - Faculty, Population Research Research Center
  - Core Member, Data Science Institute
- PhD, Education + Psychology, Michigan

### **Expertise**

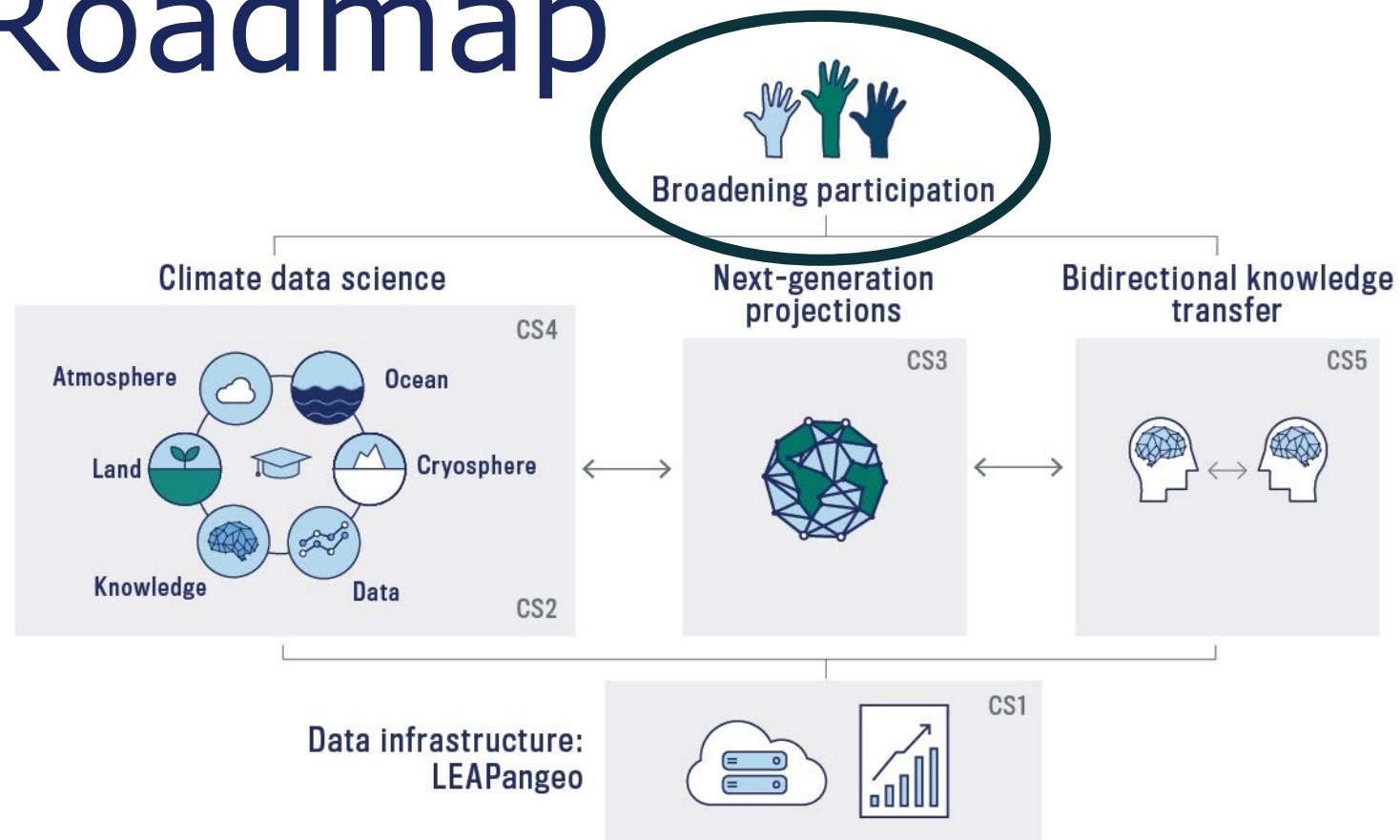
- Structural + cultural racism
- Race-based health inequities
- Virtual reality education, emerging tech + social justice

### **Other Synergistic Experience**

- Member, Columbia Climate School Faculty Task Force
- Member, Columbia Climate School Design Committee
- Faculty Advisor, SW Climate Action Initiative
- Co-Chair, Computational Social Science Working Group
- Co-Director, Justice, Equity + Tech Laboratory
- Lead Creator VR Experience, *1000 Cut Journey*



# Our Roadmap



*LEAP forward in the reliability, **utility**, and **reach** of climate projections through synergistic innovations in data science and climate science.*

# Broadening Participation Overview

- Explain why DEI is critical to LEAP's **intellectual merit + broader impact goals**
- Highlight LEAP's **vision** for Broadening Participation + Knowledge Transfer strategies
- How + why does LEAP promote “**bidirectional knowledge**” exchange with external stakeholders?



# Science \* Humanity \* Society

- Climate change will cause an **imbalance in human suffering**
- Focusing on most marginalized + vulnerable will be necessary to meet the hallmarks of **democracy, prosperity + sustainability**
- LEAP science cannot exist in a vacuum - broadening participation is **integral to climate data science**

***DEI + LEAP is critical.***



LEAP

# DEI is Critical for LEAP's **Broader Impacts**

- **Address needs** of most vulnerable communities
- **Engage** with multiple public + private stakeholders
- **Bidirectional Knowledge Transfer:** vulnerable communities + other stakeholders possess valuable knowledge + strategy
- Climate Data Science must meaningfully + substantively **reach beyond** the university



# Why Center the **Marginalized + Vulnerable**?

- **Black + Hispanic** communities in the US are **exposed to more air pollution** than they produce
- Climate change will cause the **most economic harm** in the nation's poorest counties - home to mostly **people of color**
- **Coastal (largely Black) communities** in the South are at greatest risk from sea level rise
- **Discriminatory housing policies** (e.g. “redlining”) are tied to Black communities having **more pavement, fewer trees + higher average temperatures**



# DEI is Critical for LEAP's **Research**

- Questions asked, problems + solutions identified, are **influenced by who we are**
- Richer questions, problems + solutions emerge from diverse teams
  - US needs teams equipped to address complex social ecologies

**LEAP's ability to impact society is improved  
if LEAP *looks* like society**



# We Must Interrogate *Why* the Geosciences are Inequitable

## Established Problems

1. Leaky pipeline
2. Insufficient mentorship
3. Implicit bias in recruitment

## Expanded Framing

1. Homogeneous networks
2. Segregated diversity efforts
3. Focus on bias, not **systems**

LEAP targets all three identified problems + employs an **expanded** approach.





# Leveraging DEI Best Practices

# LEAP Sets 3 Measurable DEI Objectives

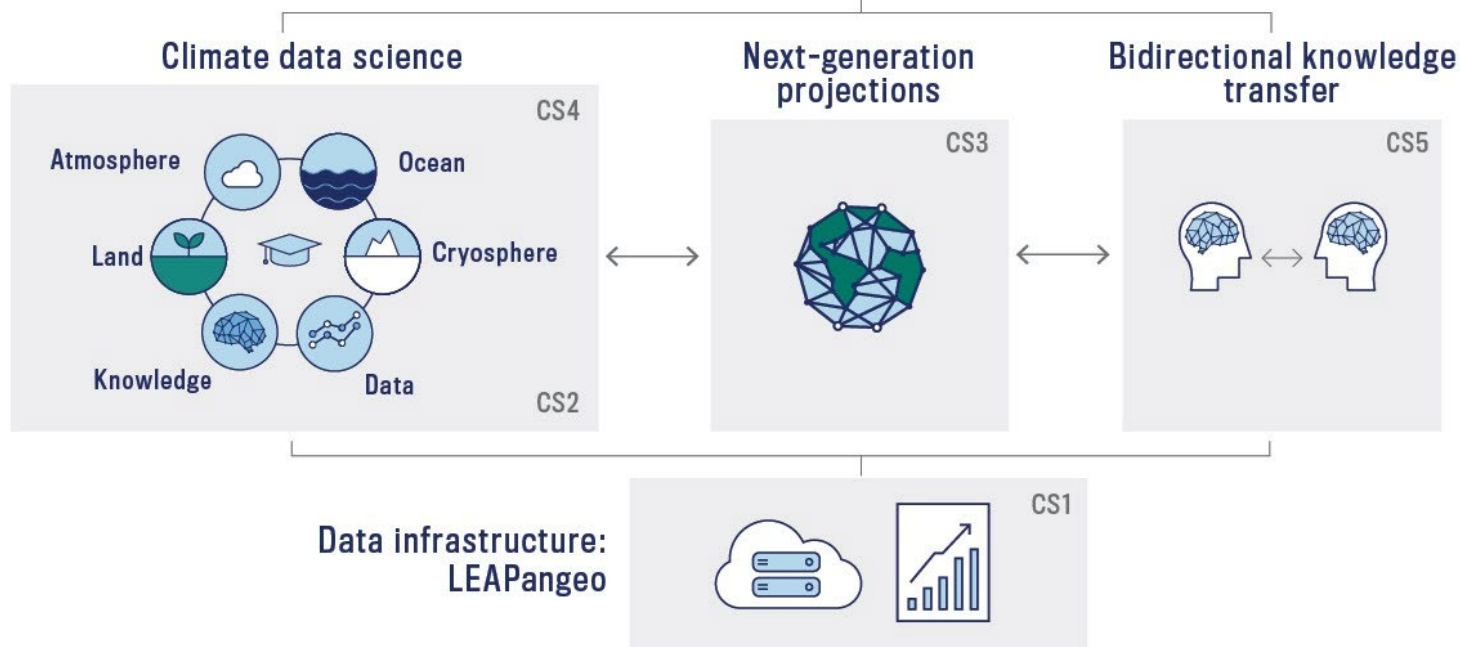
- **DEIO1: Integrate** LEAP's vision into recruitment, research, education + knowledge transfer efforts
- **DEIO2: Increase representation** of URM's in the burgeoning climate data science discipline
- **DEIO3: Assess**, review + revise all LEAP activities through an explicitly DEI lens



# LEAP's Approach: *DEI in Everything*



Broadening participation



- Strategically leveraging best practices + innovation
- DEI ecosystem focused on structural + cultural solutions
- Invest in increasing URMs in climate data science
- Assess, share + adapt



LEAP

# LEAP Leverages Best DEI Practices

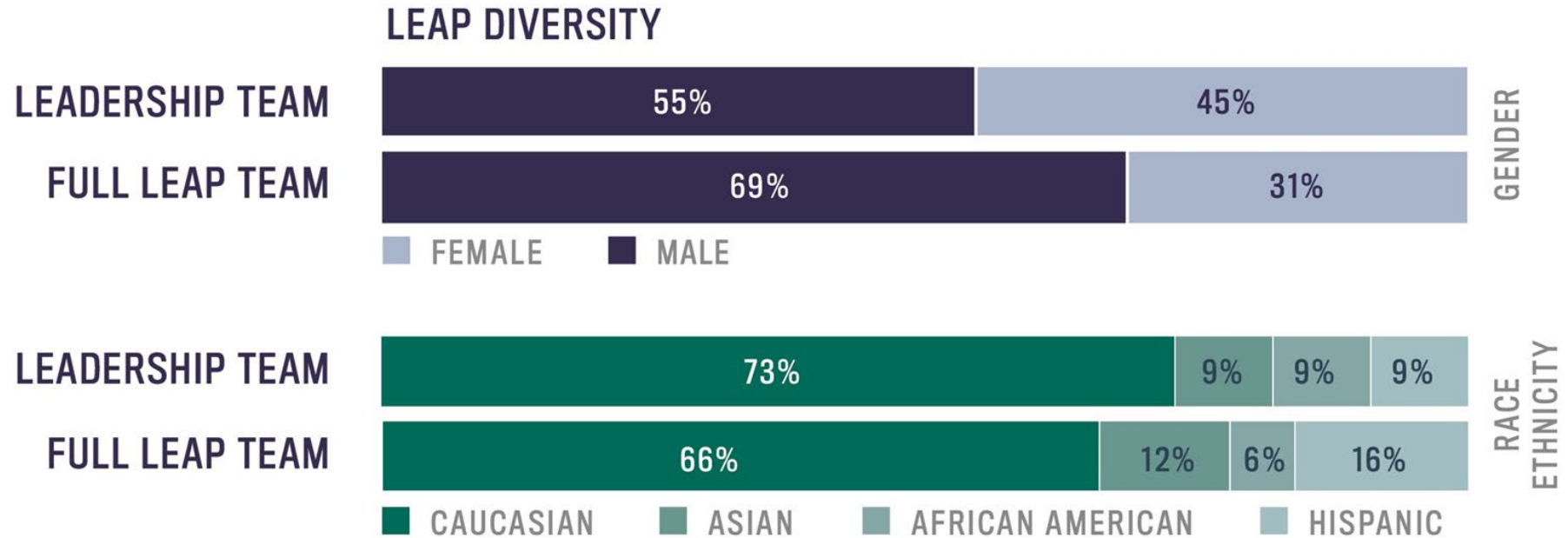
## Best Practices

- Diverse leadership
- Strengthen the pipeline
- Leverage key partnerships
- Integrate DEI into recruitment + engagement

## LEAP Response

- Leadership is 45% women, 27% POC
- HS to Post-Doc Educational programming
- Partnerships with racial + ethnic minority-led institution
- Partners with strong recruitment infrastructure
- Emphasizing broader impact goals in recruitment with:
  - Staff Writer
  - Assistant Director for Inclusive Education
  - Knowledge Transfer Program Manager

# LEAP's Team is Highly-Diverse, but We're Still Growing...



But we can do better.

We *must* do better.



LEAP



# Leveraging Prior DEI Successes

# LEAP Builds Upon Groundbreaking Work of its NSF ADVANCE Program



ADVANCE Lead PI  
LEAP Senior Personnel  
Robin Bell



DEI  
Assessment Specialist  
Valerie Purdie-Greenaway

- \$4.2M NSF investment in Columbia, 2004-2009
  - **Lead PI:** Robin Bell, LEAP
  - **Implicit Bias Trainer:** Valerie Purdie-Greenaway, LEAP
  - Began addressing cultural + systemic barriers for women in environmental science

## Synergistic Outcomes

- Women environmental science teaching faculty increased from 2% → 40%
  - 2% → 30% of environmental research faculty
- New Lamont-Doherty Office of Diversity
- Environmental Science Cluster Hire
- Implicit bias training for recruitment
- Maureen Raymo as first Lamont-Doherty Director → first woman, first climate scientist

## Best Practices for LEAP

- Gender ≠ race/ ethnicity
- Individual recruitment does not create lasting change
  - Cultural + systemic reflection is necessary
- *Science* of diversity dialogue positions teams for sustained equity
- Must use evidence-based DEI interventions, specifically with recruitment



LEAP

# “Bridge” Postbacs Strengthen a Pipeline of Graduate Researchers

- Columbia’s Bridge to PhD Postbac Program
  - URMs in STEM labs for two-year assistantships
  - Offers professional development
    - Interview skills + presentations coaching
  - Mentors graduate applications
- **Bridge works!** Since 2008:
  - 57 students trained
  - 68% PhD acceptance
- **LEAP will supports 3 Bridge Postbacs per year**

Race/ Ethnicity or Gender	Number of Bridge Alumni
Black	23
Latinx	20
Middle Eastern	5
Asian	4
White	4
Female	33
Male	24



# LEAP Will Mentor Diverse Pre-College Climate Scientists

- *Significant Opportunities in Atmospheric Research + Science*
  - UCAR's Bridge program for URM students
  - 90% of SOARS proteges enter graduate school
- **Year 1:** LEAP-led workshop for SOARS protégés
- **Years 2-5:** LEAP postdocs + students virtually mentor protégés
  - Mentorship program annually evaluated + refined





# Pushing DEI Innovation

# 3 Areas of Innovation

- **Breadth** and **Depth** of Broadening Participation
- **DEI Ecosystem:** *DEI in Everything*
- Assessing, sharing + adapting DEI strategies

# Pushing Innovation

## Breadth and Scope of Broadening Participation

- **Comprehensive Pipeline Programming**
  - High school students
  - Undergrad + grad students
  - Postdoc
  - Parents
  - Teachers
  - Journalists
  - Companies
- Journalism Workshops
- Design Studio
- Science Museum Outreach Programs
- Train-the-Trainer
- Translate-a-thon
- Lab to School
- Public Conversations
- Expanded partner network
- Storytellers in Residence
- Convergence Luncheons
- “Ask me Anything”



# Annual Hack-a-thon = Life-Relevant Learning

- Years 1-3 Hack-a-thon spanning Knowledge-Data Continuum
  - Heavily utilizing LEAPangeo
- Co-organized with NYC Office of the Mayor + BlackSpace
- Marketed by National Society of Black Engineers
- Hosted on Governors Island to showcase climate change risks in action

BLACKSPACE



**NYC** Climate Policy  
& Programs

THE TRUST FOR  
GOVERNORS ISLAND



LEAP

# BLACKSPACE



## SEEK PEOPLE AT THE MARGINS

Acknowledge the structures that create, maintain and uphold inequity. Learn and practice new ways of intentionally making space for marginalized voices, stories, and bodies.



## PLAN WITH, DESIGN WITH

Walk with people as they imagine and realize their own futures. Be connectors, conveners, and collaborators—not representatives.



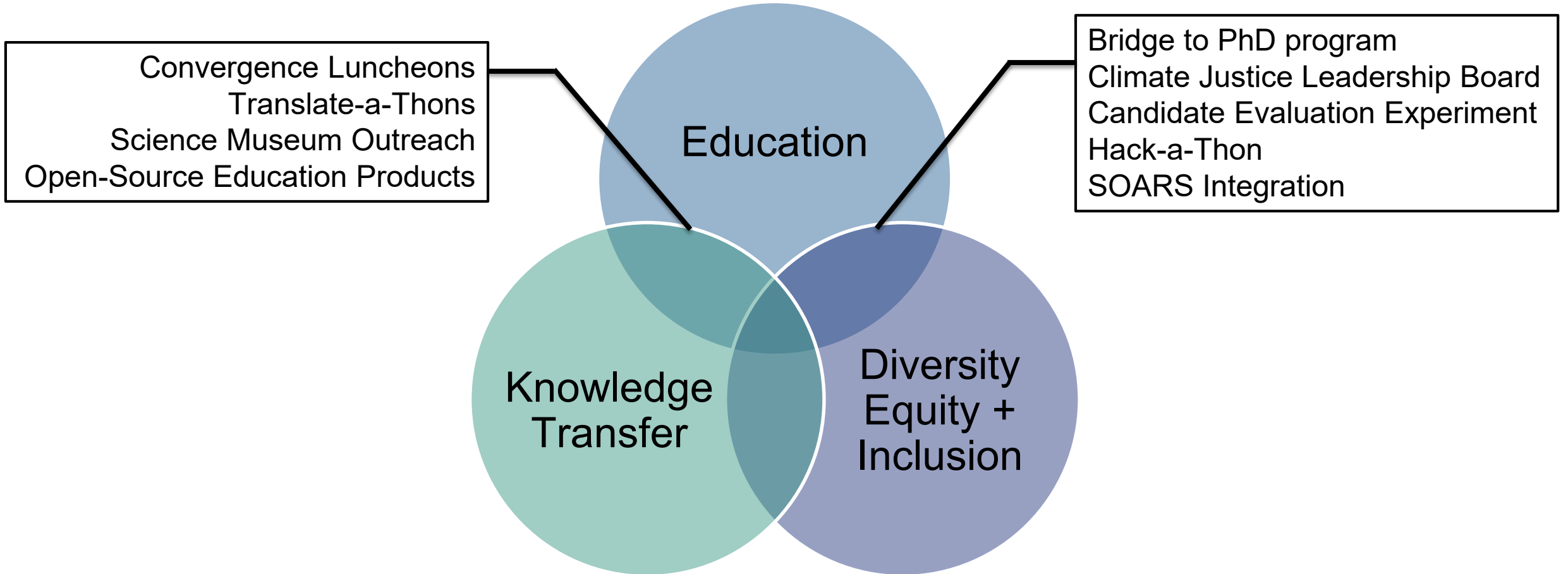
## MOVE AT THE SPEED OF TRUST

Grow trust and move together with fluidity at whatever speed is necessary.\*



LEAP

# Suffusion: LEAP Integrates Education, DEI + KT



# Pushing Innovation

## Suffusion

- Establish DEI grounded center culture - transdisciplinarity, transparency, comprehensive broadening participation, education + knowledge transfer initiatives
- Management Structure, subcommittees, Climate Justice Leadership Board
- **DEI Orientation:** Establishing a shared vision + strategy



# Pushing Innovation

## DEI Orientation

**Equity Literacy + Visioning**



**Equity Planning**



**Annual Performance Review**

- All LEAP researchers, trainees + administrators will complete DEI workshops
- Funding prioritized to applicants who completed workshops



# Pushing Innovation

## Assess, Share + Adapt

- Blind applicant names + institutional affiliations for:
  - Fellowships
  - Postdoc traineeships
  - REU programs
  - Internships
  - Seed funding
- Utilize work samples
- Standardized interviews
- Randomized field experiment



# Recruitment is *Itself* a Research Frontier

- Multi-phase hiring
- Blinding applicant name + institutional affiliations
- Work samples
- Standardized interviews



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



DEI  
Assessment Specialist  
**Valerie Purdie-Greenaway**



# Recruitment is *Itself* a Research Frontier

## Randomized Field Experiment

- **Pilot 1:** REU
- **Pilot 2:** Seed Funding
- **Research Design**
  - **Treatment 1:** Single-phase
    - Blinded CV/ resume
  - **Treatment 2:** Multi-phase
    - T1 + work sample, standardized interviews
- Present + Publish



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



DEI  
Assessment Specialist  
**Valerie Purdie-Greenaway**



# DEI Strategy: Assess, Share + Adapt

- Multi-methodological Strengths, Weakness, Opportunities + Threats (SWOT) Analysis
- Survey, interviews, lab-based ethnography, comprehensive review
- Build Inclusive Science Module
- Publish + adapt



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



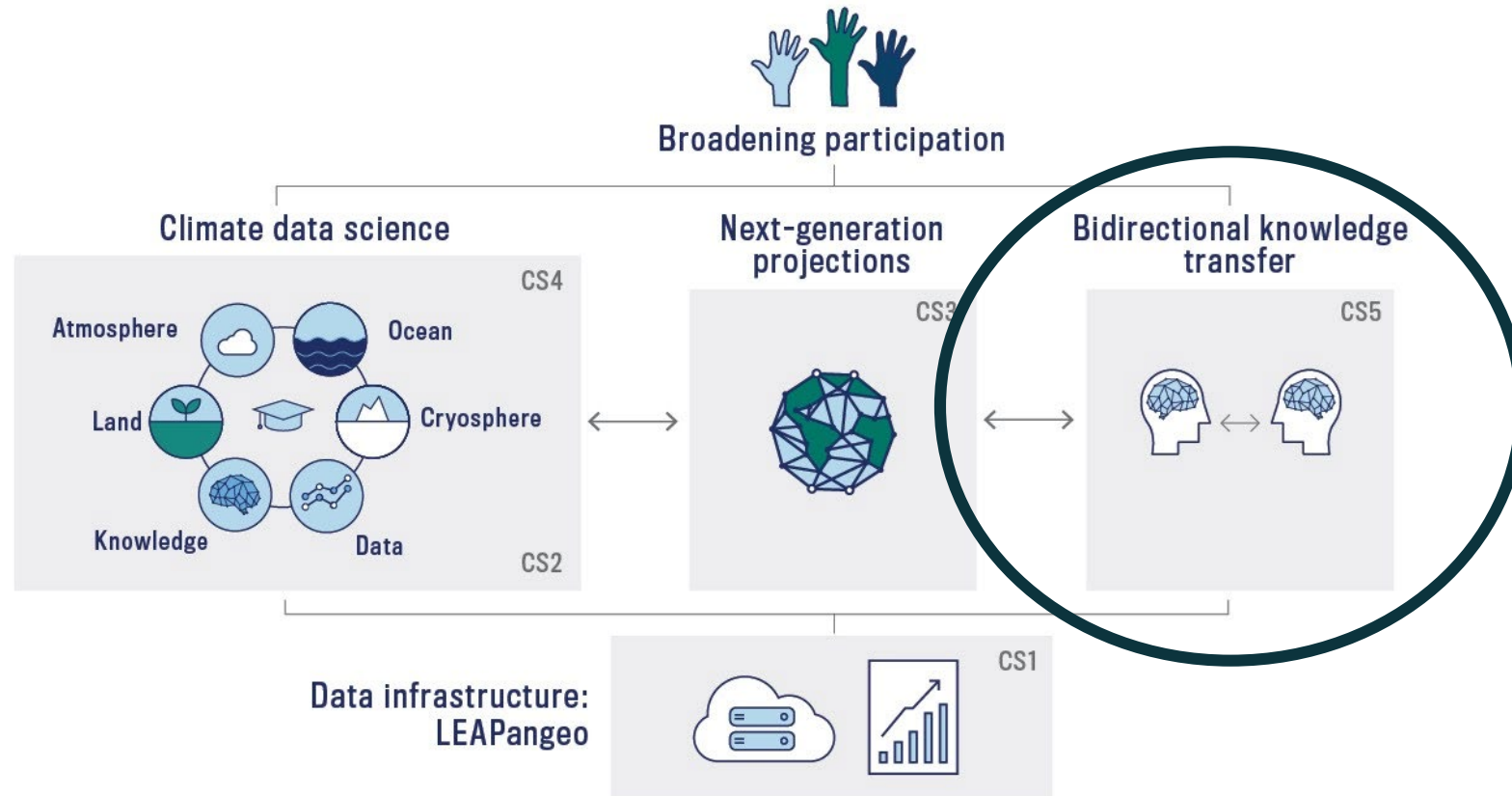
DEI  
Assessment Specialist  
**Valerie Purdie-Greenaway**





# Knowledge Transfer *is* DEI

# LEAP's Knowledge Transfer Grows Out of its *DEI in Everything* Commitment



# LEAP Sets Three Knowledge Transfer Objectives (KTO)

- **KTO1:** Establish a **bidirectional dialogue** between academia, industry, non-profits + the public
- **KTO2:** Innovate **storytelling, visualization + social media** for broad stakeholder intake
- **KTO3:** Make all LEAP research **open-source + broadly accessible**



# Knowledge Transfer Stems from LEAP's DEI Commitments

Research + education + knowledge transfer seeks societal betterment

+ Better societies are *more diverse, equitable + inclusive*

= **DEI in Everything**

- Education, Corporate + Public Engagement, DEI are all Knowledge Transfer
- LEAP must understand + engage heterogeneous interests in climate
- **Knowledge transfer must emerge from DEI commitments**
  - Cogburn as Chief Equity Officer + Knowledge Transfer Director



# Our Knowledge Transfer Team Spans Research + Practice

## Three Broad Audiences

- Private Industry
- Public + Media
- Communities of Color + Climate Action Communities



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



Corporate  
Engagement Director  
**Vanessa Burbano**



Public  
Engagement Director  
**Andrew Revkin**

## Our Experts

- Structural + cultural **racism** (Cogburn)
- Strategic implications of **CSR** (Burbano)
- 30+ years of climate **journalism** (Revkin)



# The Climate Justice Leadership Board Positions Students as Evaluators + Change Agents



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



Evaluator  
**Charles Lang**

- President, Vice President, Communications Director elected from + by LEAP grad students
- Partner with climate groups working with vulnerable populations + climate justice
- Get trained in evaluation + assessment
  - Receive supplemental funding for evaluation
- Coordinate *Public Conversations* events
- Meets monthly
- Supported by Cogburn + Lang



*“There are those who solve the problems of their disciplines and **those who solve the problems of the world.**”*

- Unknown

***This*** is LEAP.



Questions?  
Feedback?



# Corporate Engagement & Partnerships



Corporate  
Engagement Director  
**Vanessa Burbano**

- Assistant Professor of Management, Columbia
- PhD, Management, UCLA

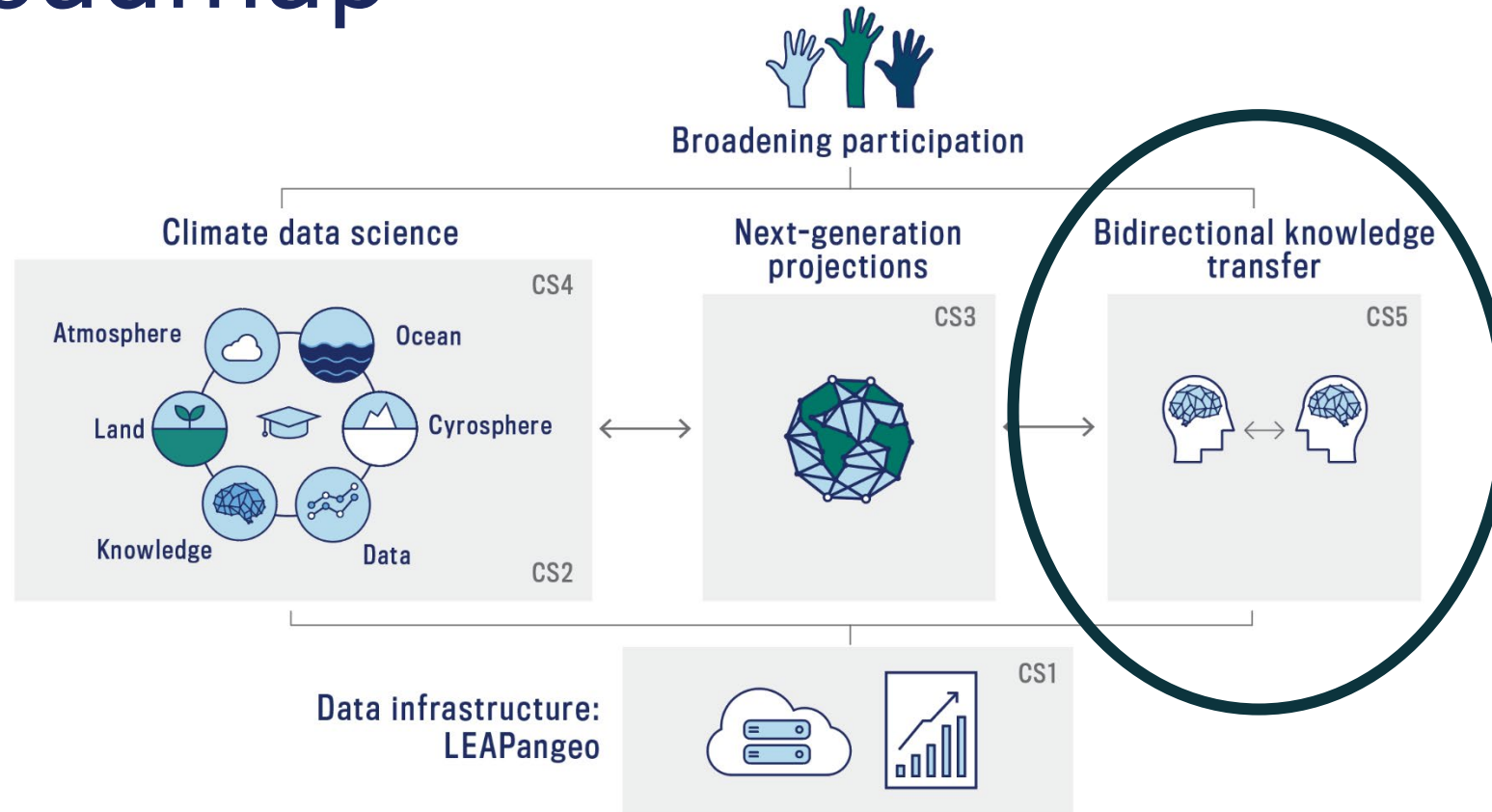
### Expertise

- Corporate social responsibility
- Sustainability strategy
- Employee motivation

### Other Synergistic Experience

- Leadership Team, Academy of Management's Strategy Division
- Leadership Team, Strategic Management Society's Stakeholder Strategy Group
- Board Member, Alliance for Research on Corporate Sustainability

# Our Roadmap



*LEAP forward in the reliability, **utility**, and **reach** of climate projections through synergistic innovations in data science and climate science.*

# My Presentation Will Explain:

- Why + how LEAP promotes bidirectional knowledge exchange with corporate stakeholders
- Role + integration of corporate + information-dissemination partners

# The Private Sector is Important for LEAP

- Corporations either exacerbate, or help to mitigate, climate change
- Only if knowledge outputs are relevant to + usable by corporate audiences will LEAP have an impact through the private sector



# Content + Format of Knowledge Outputs Will Vary to Ensure Relevance for a Range of Corporate Audiences

LEAP Climate  
Information Portal



**Granularity of Knowledge Output of  
Relevance**

High

Low

Raw datasets +  
tutorials

Aggregated  
data in exec-  
friendly format

Executive summaries  
of findings  
“translated” for a  
corporate audience

Research-informed  
tools for  
communication +  
design of initiatives



# LEAP Will Generate Research-Informed Tools for Communication + Design of Sustainability Initiatives

## Why Important?

- Companies either exacerbate, or help to mitigate climate change
- Challenge: conflicting goals
- Research on “doing well by doing good” inconclusive

## What is Missing?

- Need to understand **mechanisms** through which sustainability-oriented initiatives can influence **company stakeholders** \*
- Need to understand how **design** and **communication** of initiatives can influence stakeholders\*\*
- Need for **causal** studies\*\*\*

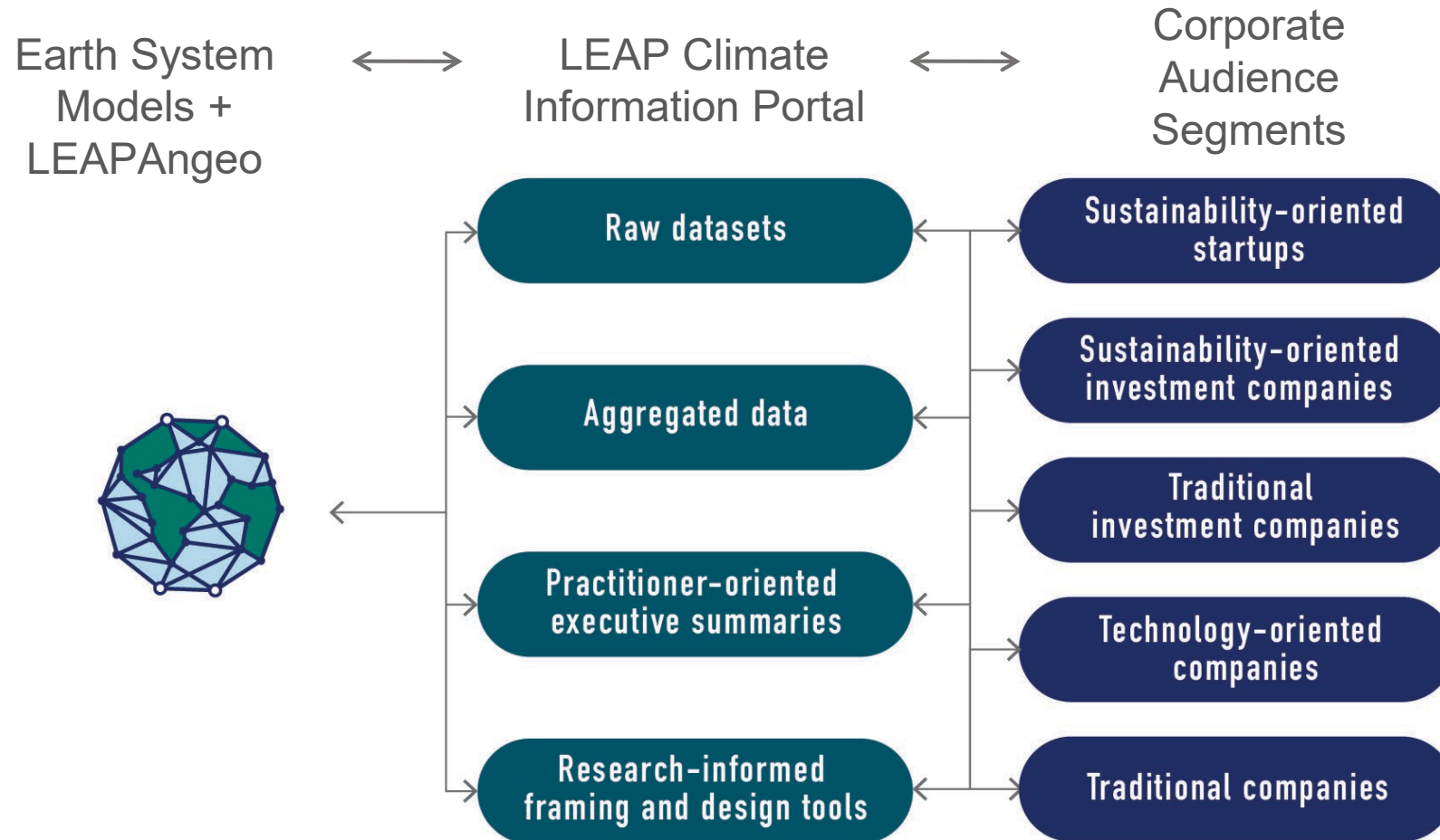


## Opportunity for LEAP

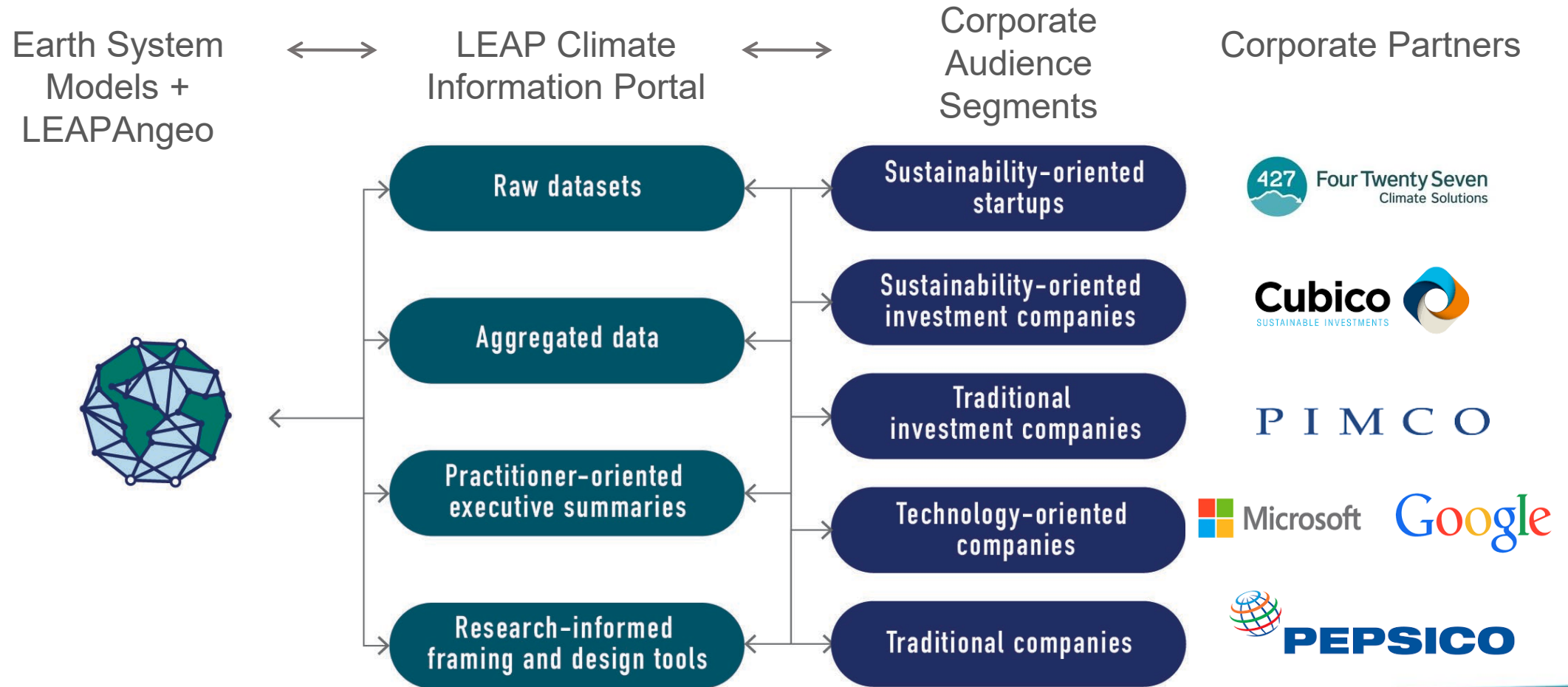
Create, aggregate + “translate” relevant research



# LEAP Adjusts Knowledge Outputs to Ensure Relevance to a Range of Corporate Audiences



# LEAP's Corporate Partners are Representative of this Range of Corporate Audience Segments



# LEAP's Process for Corporate Engagement

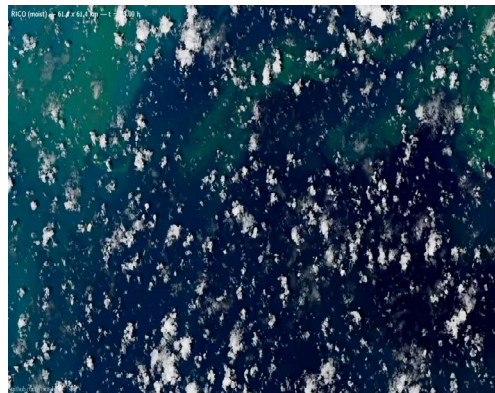
Year 1				Year 2				Year 3				Year 4				Year 5			
Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4

# Today is a Unique Moment of Opportunity for Corporate Engagement on Climate Change

Massive Data from Earth Observations



High-Resolution Process-Resolving Simulations



Progress in Machine Learning



Public Debate on Climate Change + Diversity



Private Sector New Ally for Change



# Increasing Climate Models' Utility + Reach Requires Corporate Stakeholder Engagement + Segmentation + Research

- LEAP's online platform must be a key “go-to” source for data, information + practitioner-oriented tools on climate issues
  - Annual evaluation based on survey responses
- This will facilitate private sector consideration of + adaptation to climate realities

Questions?



# Public Engagement + Partnerships



Public  
Engagement Director  
**Andrew Revkin**

- Founding Director, Initiative on Communication + Sustainability, Earth Institute, Columbia
- MS, Journalism, Columbia

### **Expertise**

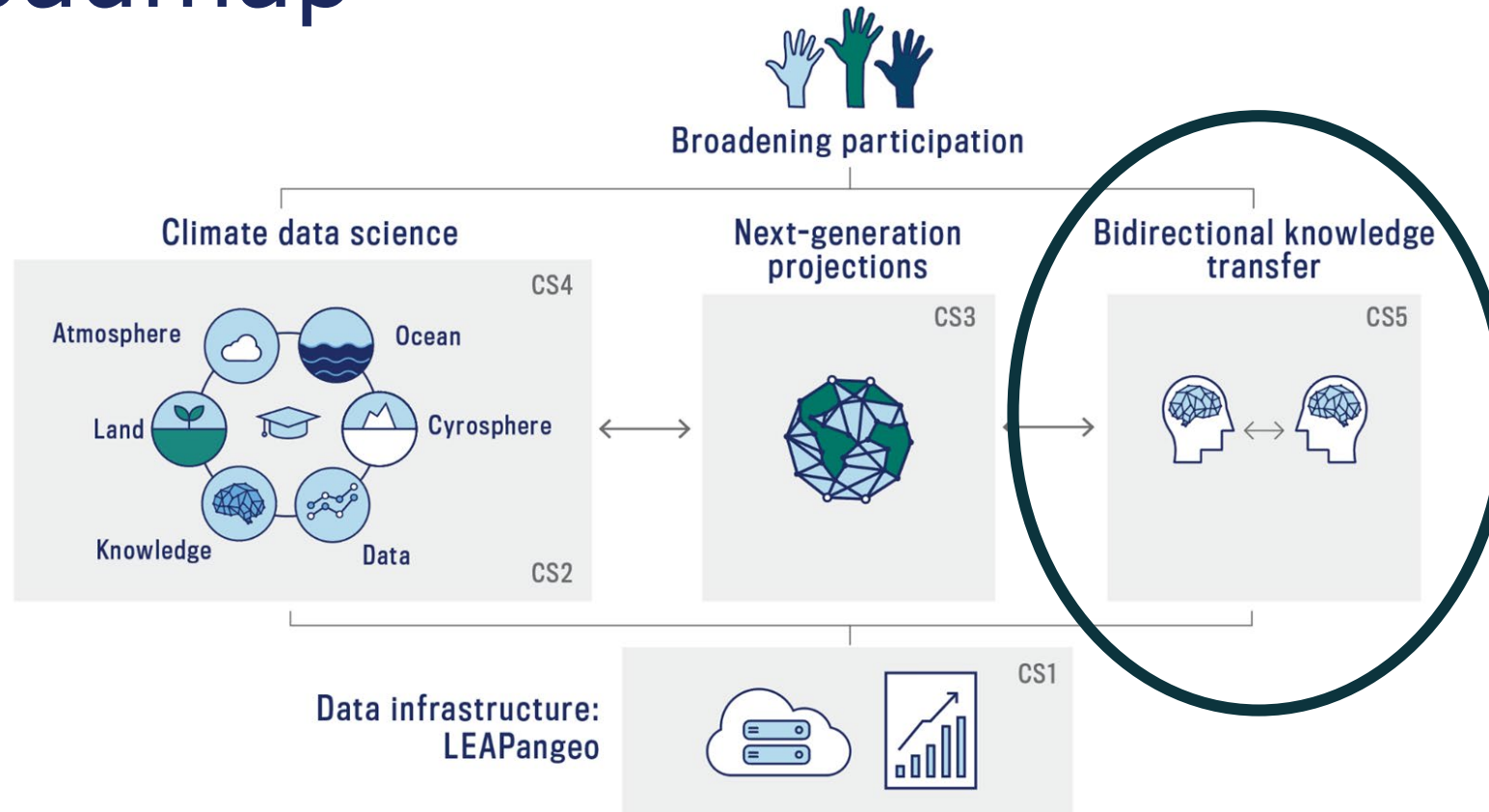
- 35 years of climate journalism
- Sustainable development
- Blogging + social media (teaching, training)

### **Other Synergistic Experience**

- Strategic Advisor, Science Journalism, National Geographic
- Reporter, Blogger, The New York Times, 1995-2016
- Member, Anthropocene Working Group, 2010-2016
- AGU Career Science Journalism Award, 2015
- NAS Communication Awards, 2003, 2011
- NAS Climate Communication Initiative Committee



# Our Roadmap



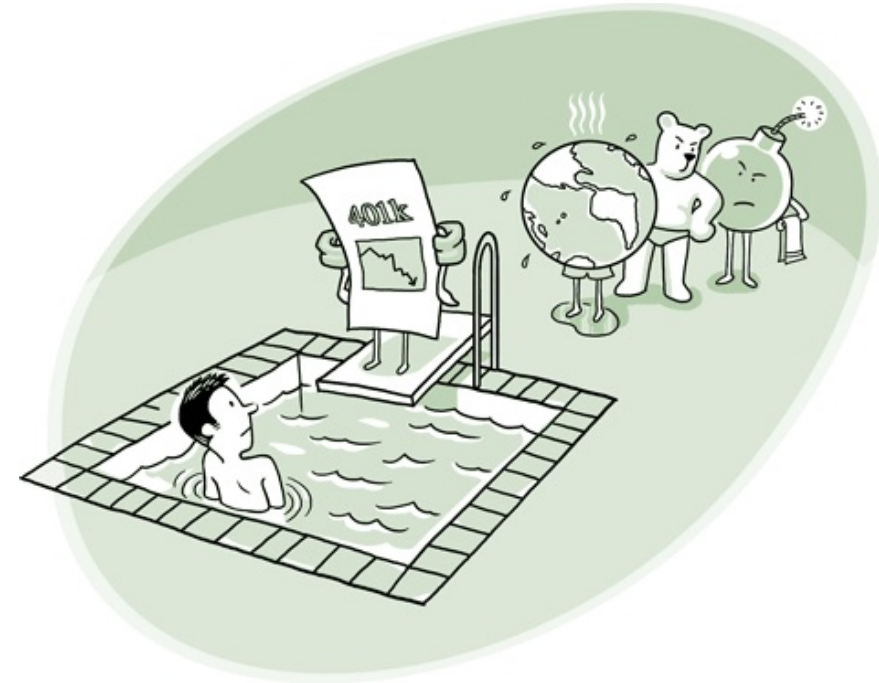
*LEAP forward in the reliability, **utility**, and **reach** of climate projections through synergistic innovations in data science and climate science.*

# My Presentation Will Explain:

- Why + how LEAP promotes bidirectional knowledge exchange with public + media stakeholders
- The role + integration of partners for public + media impact
- How this activity propels LEAP toward its Knowledge Transfer objectives

# 30 Years of Climate Communication Had Minimal Impact on Climate Response

- **We swim in a “Finite Pool of Worry”**
  - Climate warning competes with realtime challenges
  - People need access to actionable data + knowledge
- **People are Different**
  - Cultural values shape risk perceptions + responses
  - Climate science literacy does not correlate with climate change concern \*
- **Engagement Emerges via Convergence**
  - Diverse stakeholders require diverse dialogues
  - Communicators must understand the *science* of diversity
  - Climate science + data science must connect



# Public Stakeholders Need **Early Bidirectional** Data Dialogue

- Shapes **inquiry + output** for diverse stakeholders
- **Builds trust + understanding** in emerging fields amid polarizing debate
- Forges new pathways to **broadening participation in climate data science**
- **Empowers vulnerable communities** through access to actionable data



# Public Stakeholders Need **Early Bidirectional Data Dialogue**

- Shapes **inquiry + output** for diverse stakeholders
- **Builds trust + understanding** in emerging fields amid polarizing debate
- Forges new pathways to **broadening participation in climate data science**
- **Empowers vulnerable communities** through access to actionable data



Columbia Earth Institute Brainstorming Session:  
Maya Lin, Designer



Columbia Earth Institute Training Workshop:  
Bangladesh Meteorological Department

# Media Also Need **Early Bidirectional Data Dialogue**

- Newsroom capacity is plunging
  - “News deserts” spread where climate insights are most critical
- **Off-Deadline Engagement** Columbia Journalism Review. **Life in a News Desert**
  - Builds convergent relationships
  - Supports innovative translation capacities
  - Creates fluency for when news breaks
- Early efforts promising \*

Initiative on Communication and Sustainability  
EARTH INSTITUTE | COLUMBIA UNIVERSITY



CLIMATE  CENTRAL



# Three Knowledge Transfer Goals Guide LEAP's Public Engagement

- **KTO1:** Establish a **bidirectional dialogue** between academia, industry, non-profits + the public.
- **KTO2:** Innovate in **storytelling, visualization + social media** for broad stakeholder intake.
- **KTO3:** Make all LEAP research **open-source + broadly accessible**.

# KTO1: Annual Workshop Engages Journalists + Practitioners in Climate Data Science

- Full-day event for LEAP researchers, journalists, public communicators, city planners + public
- **Key Priorities**
  - Data literacy + LEAPangeo access
  - Integrating climate modeling into urban design + planning
  - *DEI in Everything*
  - Urban + regional resilience
  - Establishing co-production community
- Attendees recruited by UCCRN, Climate Central + Earth Journalism Network
- Administered by Knowledge Transfer Program Manager + Staff Writer



# KTO1: Journalism Fellowships Train Reporters in Climate Innovations



- **Yearlong + Online Fellowships**

- 3 per year, potentially more after
  - Learn risk quantification, data science + solution mapping
    - Educated via *DEI in Everything*
  - Engage communities facing high climate risk
  - Mentored by Earth Institute
  - Stories amplified through EI's *Resilience Media Project*
- Fellows recruited by Climate Central + Earth Journalism Network

Earth Institute Announces Resilience Media Project

BY EARTH INSTITUTE | MAY 11, 2020

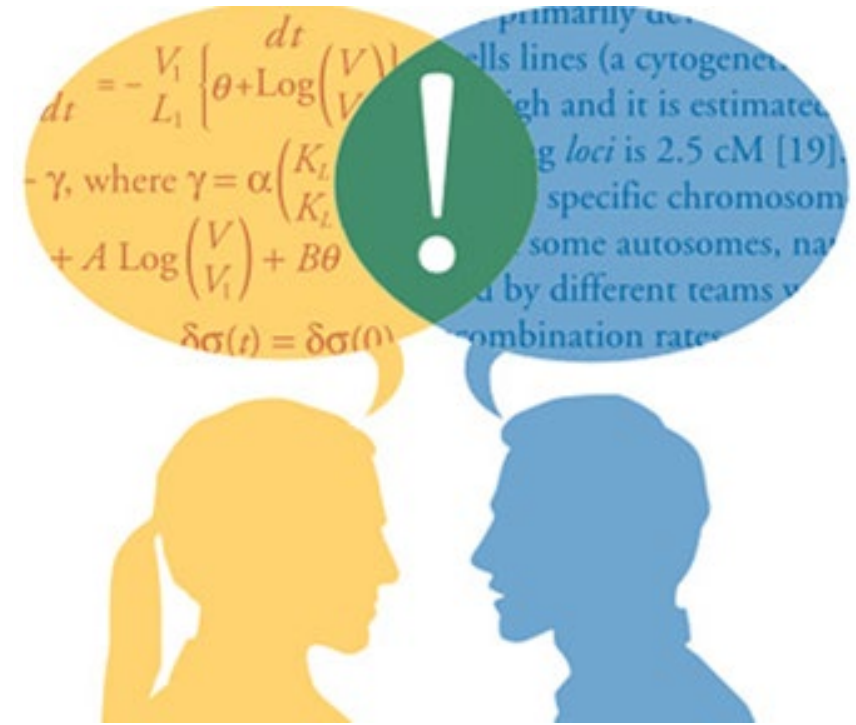


Dale Willman, who will lead the new Earth Institute Resilience Media Project, here runs a discussion of resilience journalism in Miami. Photo: [Kate Stein, WLRN radio](#)



# KTO1: Convergence Luncheons Facilitate Internal Collaboration

- 6 *Convergence Luncheons* per year for LEAP researchers, trainees + partners
- Forge shared vocabulary across climate science, data science, social science + outreach
- Lightning research talks spark discussion
- Knowledge Transfer Program Manager publishes meeting reports to website
- Planned by Assistant Director for Inclusive Education



# KTO1+2: Storytellers in Residence Spur Communication Innovations + Transparency

- Harlem-based artists, writers, or journalists embed for semester
  - Working space in LEAP headquarters
  - **Inaugural Storyteller:** Brian Kahn, Gizmodo
  - **Future Storytellers:** Selected by Knowledge Transfer Subcommittee
- **Activities + Responsibilities**
  - Translate research processes + products to external stakeholders
  - Acquire climate data science literacy
  - Support Staff Writer in developing content for broad dissemination
  - Join Knowledge Transfer Subcommittee *ex officio*



Managing Editor, Earthier  
Gizmodo  
Brian Kahn

**GIZMODO**

# KTO1+2: Storytellers in Residence Spur Communication Innovations + Transparency

- Harlem-based artists, writers, or journalists embed for semester
  - Working space in LEAP headquarters
  - **Inaugural Storyteller:** Brian Kahn, Gizmodo
  - **Future Storytellers:** Selected by Knowledge Transfer Subcommittee
- **Activities + Responsibilities**
  - Translate research processes + products to external stakeholders
  - Acquire climate data science literacy
  - Support Staff Writer in developing content for broad dissemination
  - Join Knowledge Transfer Subcommittee *ex officio*



Managing Editor, Earther  
Gizmodo  
**Brian Kahn**

**GIZMODO**



Prior Storyteller in Residence:  
Mary Heglar



# KTO1: *Public Conversations* Facilitate Discussion between Diverse Audiences + Climate Data Scientists

- Four *Public Conversations in Climate Data Science* per year
- Co-administered by Columbia's Center for Science + Society, Climate Justice Leadership Board
- **Invited Audiences**
  - Climate data scientists
  - Social scientists
  - Journalists
  - Teachers + parents from education programs
  - Policy + business scholars
  - General public
- **Each Event**
  - Features cross-disciplinary panel
  - 200 attendees in-person, 500 live-streaming
  - Events live-tweeted + recorded
    - Recordings posted to website + YouTube



# KTO2: Social Media Innovations Experiment with New Engagement Mechanisms



Jamison Ervin  
@jamisonervin

Andrew @Revkin - congrats for making top 50 of global influencers on #sustainability! Some tough competition, with @HelenClarkNZ, @PaulPolman, @ASteiner and @jrockstrom, among others!  
[onalytica.com/blog/posts/sus...](https://onalytica.com/blog/posts/sus...)



10:23 PM · Apr 12, 2018 · Twitter Web Client

- Public initiates dialogue + questions with LEAP faculty
- “Ask Me Anything” on Reddit, Twitter + Instagram
  - **Hashtag: #climateLEAP**
- **LEAP’s Dedicated Administrators**
  - Support content generation
  - Organize interviews
  - Transcribe responses
  - Moderate digital venues



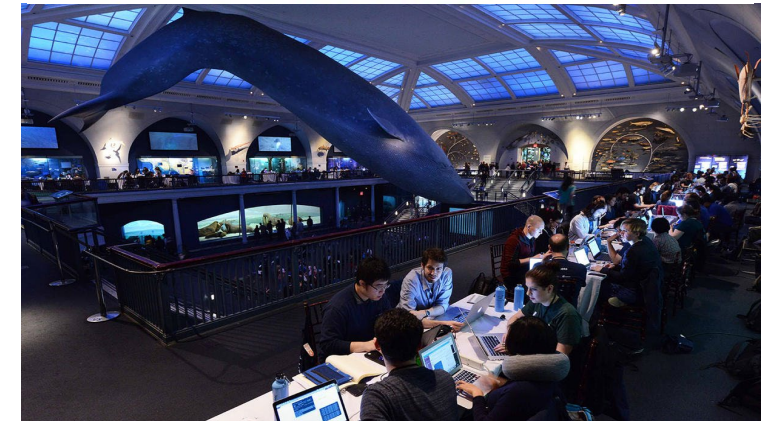
# KTO2: Translate-a-thons Train Students + Partners in Science Communications

- Quarterly, half-day sessions on scientific translations
- LEAP faculty + trainees + partners collaboratively develop:
  - Visualizations
  - Videos
  - Interactive web tools
- Co-hosted by Gizmodo, BlackSpace + AMNH
- Events guided by list of existing knowledge needs
- Supported by Assistant Director + Program Manager

**GIZMODO**

**BLACKSPACE**

 **AMERICAN MUSEUM  
OF NATURAL HISTORY**



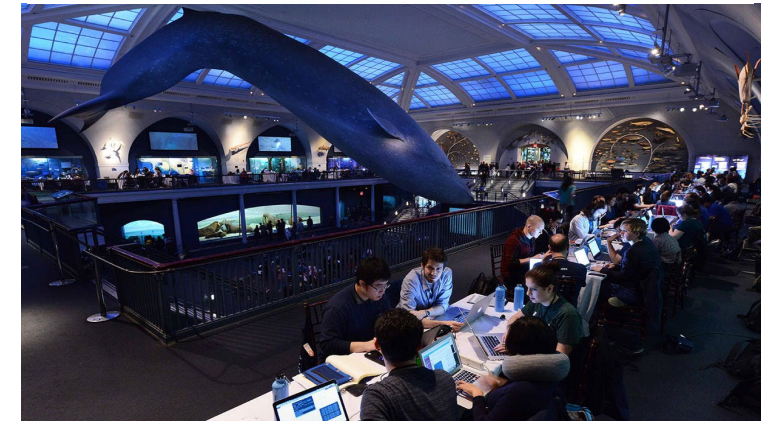
# KTO2: Translate-a-thons Train Students + Partners in Science Communications

- Quarterly, half-day sessions on scientific translations
- LEAP faculty + trainees + partners collaboratively develop:
  - Visualizations
  - Videos
  - Interactive web tools
- Co-hosted by Gizmodo, BlackSpace + AMNH
- Events guided by list of existing knowledge needs
- Supported by Assistant Director + Program Manager

**GIZMODO**

**BLACKSPACE**

 **AMERICAN MUSEUM  
OF NATURAL HISTORY**



@CarbonWrangler of  
@ColumbiaUEnergy

# KTO2: Leveraging Existing Programs + Relationships Accelerates Broadest Impact

- Earth Institute's ***Sustain What? Webcasts***
  - 85 episodes
  - 250k viewers
- Earth Institute blog, website + events amplify LEAP
- “Covering Climate Now”:  
Columbia Journalism Review program + blog
- Digital Storytelling Lab
- **NYC Office of the Mayor:** Leverage LEAP research for *OneNYC 2050*, NYC’s long-term strategic plan

**NYC** Climate Policy & Programs

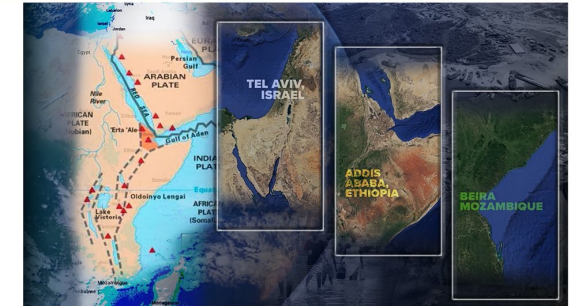
EARTH INSTITUTE  
**LIVE** 

**Sustain What: Urban Futures Along the Great Rift Valley**  
Weds., 22 July, 11 am US ET Info: [j.mp/sustainwhatlive](https://www.j.mp/sustainwhatlive)

Join host Andy @Revkin of the Earth Institute and a global team of Columbia design pioneers from GSAPP



ColumbiaUrbanDesign  
@ColumbiaUD Follows you



Join Columbia Architecture & Urban Design Studio students & faculty in a visual exploration of resilient thriving cities in progress

Info: [j.mp/sustainwhatlive](https://www.j.mp/sustainwhatlive)



Initiative on Communication and Sustainability  
EARTH INSTITUTE | COLUMBIA UNIVERSITY

**OneNYC 2050**  
BUILDING A STRONG AND FAIR CITY  
VOLUME 1 OF 9  
APRIL 2019

THE CITY OF NEW YORK  
MAYOR BILLY DE BLASIO  
DEAN FULFORD  
FIRST DEPUTY MAYOR  
DOMINIC WILLIAMS  
CHIEF POLICY ADVISOR  
DANIEL A. ZARRELLI  
DIRECTOR

**NYC**

**A LIVABLE CLIMATE**  
VOLUME 7 OF 9

New York City will lead a just transition to achieve carbon neutrality and adapt the city to withstand and emerge stronger from the impacts of climate change.

**NYC**

# My Career Tracks the Switch from Unidirectional to Bidirectional Communications

- LEAP's opportunity is illustrated in this shift

1985



1988



1995-2009



# My Career Tracks the Switch from Unidirectional to Bidirectional Communications

- LEAP's opportunity is illustrated in this shift

1985



1988



1995-2009



2007-2016



# My Career Tracks the Switch from Unidirectional to Bidirectional Communications

- LEAP's opportunity is illustrated in this shift

1985  
2020



1988



1995-2009



DOT EARTH

Nine Billion People, One Planet.

July 24, 2009, 1:47 PM

A Postcard From the Pleistocene

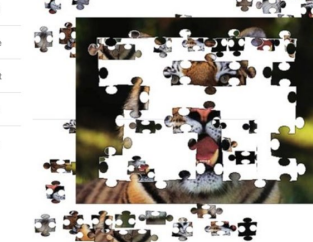
By ANDREW C. REVKIN



Certainties, Uncertainties and Choices with Global Warming

By ANDREW C. REVKIN

SEPTEMBER 26, 2014 11:00 AM

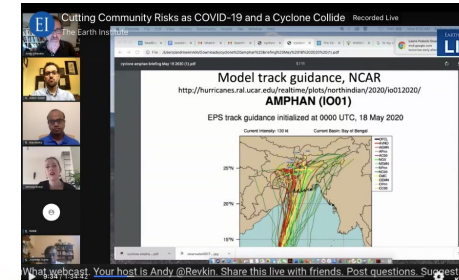


Mario Molina, a Nobel laureate in chemistry, uses this image representing a part completed jigsaw puzzle to convey the state of understanding of human-driven climate change. Mario Molina.

2007-2016

COLUMBIA UNIVERSITY  
Home About Climate Communications Network

Climate and Sustainability  
Communications Network



Questions?  
Feedback?



# Management + Administration

# LEAP PI + Co-PI team



Center Director and  
Diversity Co-Director  
**Pierre Gentine**



Deputy Director  
**Galen McKinley**



Chief Convergence Officer  
& Education Director  
**Tian Zheng**



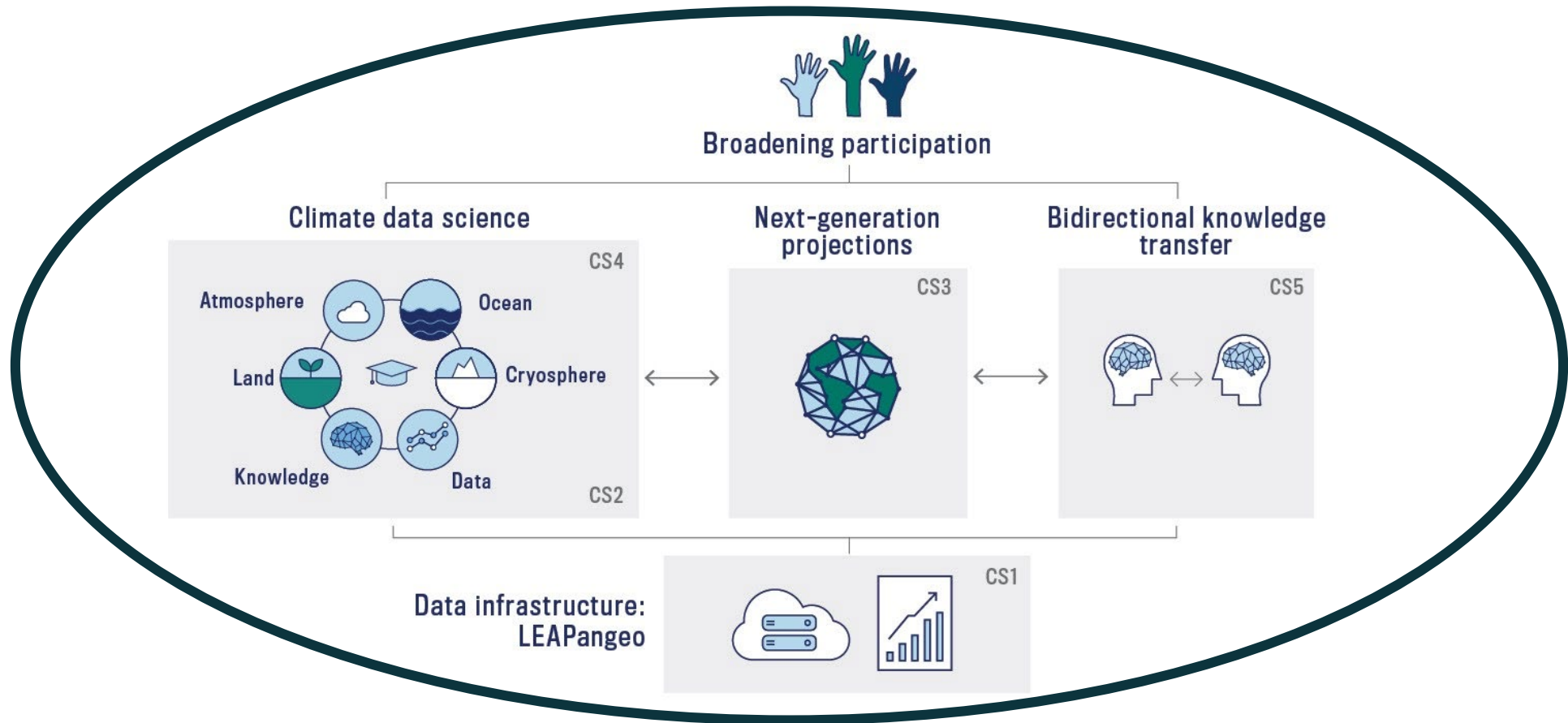
Data & Computation  
Director  
**Ryan Abernathey**



Chief Equity Officer &  
Knowledge Transfer Director  
**Courtney Cogburn**



# Implementing our vision



LEAP forward in the reliability, utility, and reach of climate projections through synergistic innovations in data science and climate science.

# LEAP spans the continent, with NYC as hub

The map shows the following logos and location pins:

- West Coast:** Microsoft, 427 Four Twenty Seven Climate Solutions, Google, PIMCO, UCI.
- Central:** NCAR, UCAR.
- Midwest:** M (University of Minnesota).
- East Coast:** MILA, ARCS Alliance for Research on Corporate Sustainability, COLUMBIA UNIVERSITY, NYU, TEACHERS COLLEGE COLUMBIA UNIVERSITY, NASA, NYC Climate Policy & Programs, Cubico Sustainable Investments, NYC Department of Education, AMERICAN MUSEUM OF NATURAL HISTORY, THE TRUST FOR GOVERNORS ISLAND, CLIMATE CENTRAL, BLACKSPACE, PEPSICO, GIZMODO, UCCRN URBAN CLIMATE CHANGE RESEARCH NETWORK, EARTH JOURNALISM NETWORK.

# Complementary university capabilities integrated through structured programs

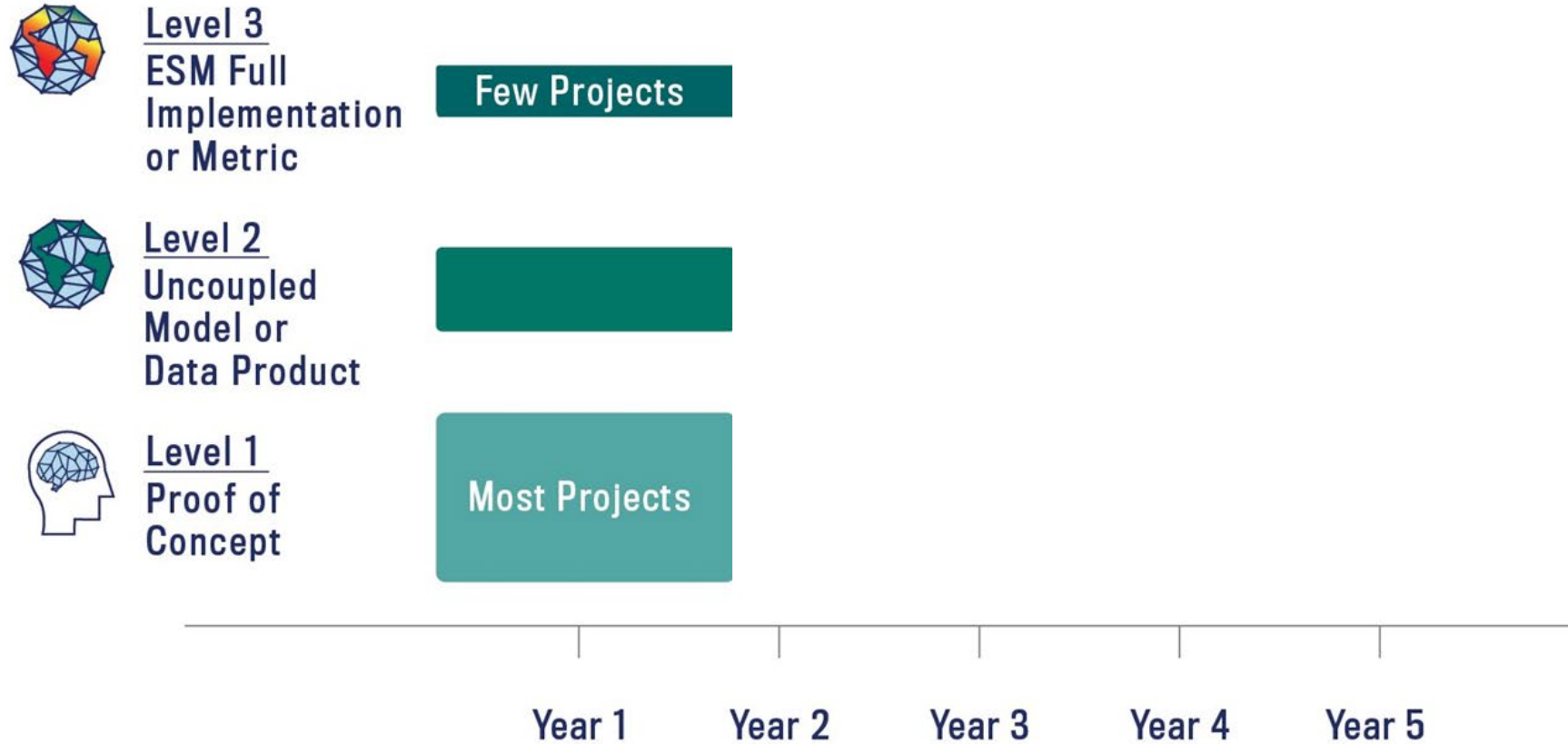
University	Expertise
Columbia	Climate science, data science, psychology, management, education, communications → build on transdisciplinary culture
NYU	Machine learning for ocean modeling
UC Irvine	Computational atmospheric physics
Minnesota	Physics-guided machine learning
Montreal (MILA)	Machine learning applications to climate change
Teachers College	Environmental education, education policy



# Values inform LEAP's management approach

- **Transparent**, shared governance to leverage broad expertise
- **Proactive** steps for diversity, equity + inclusion
- Research structured to ensure **impact** on climate projections
- **Open-source**
- **Bi-directional** knowledge transfer to continually engage stakeholders

# Research strategy will move projects from Proof of Concept to ESM Implementation



# LEAP is 100% Open-Source

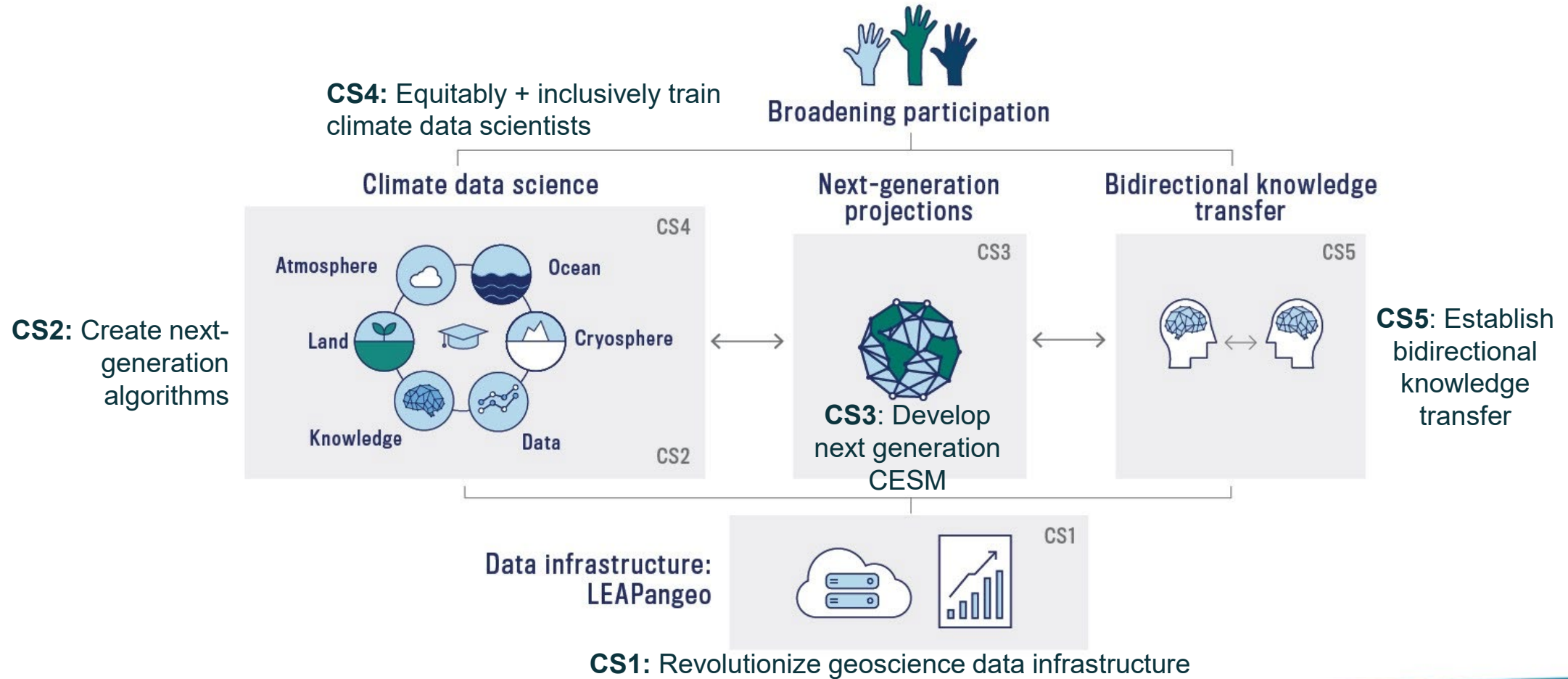
Goal	Vehicle
Shared Data	LEAPangeo, Website
Shared Code	GitHub
Shared Ownership, Transparent Operations	Operations Manual, Annual Report, Newsletter
Transparent Funding Decisions	Funding Review Panel, Ethics Handbook, Website

- **All innovations publicly disclosed**
  - Open licensing → inclusive, broadening participation, and knowledge transfer
  - Published as soon as possible
  - Underlying code published to website
- Consistent with NCAR's CESM + NSF's Pangeo open-source models

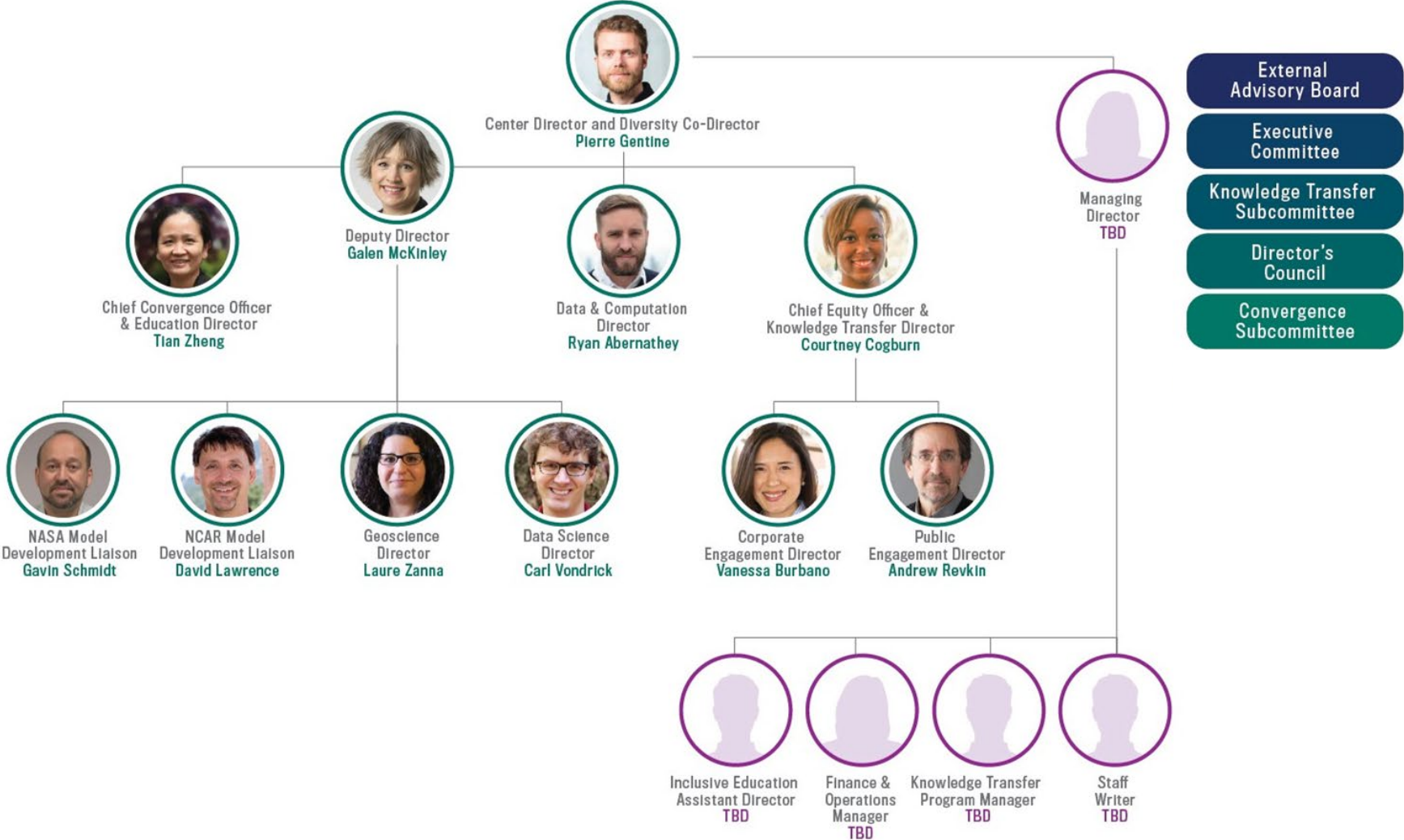
# LEAP is well positioned for virtual work

- Mainly *computational* research
- *LEAPangeo* provides cloud-based distributed data infrastructure
- Team highly-experienced in virtual meetings + distanced collaboration
- Will continue to adapt with emerging technology
- This Site Visit showcases team ability to work effectively at a distance

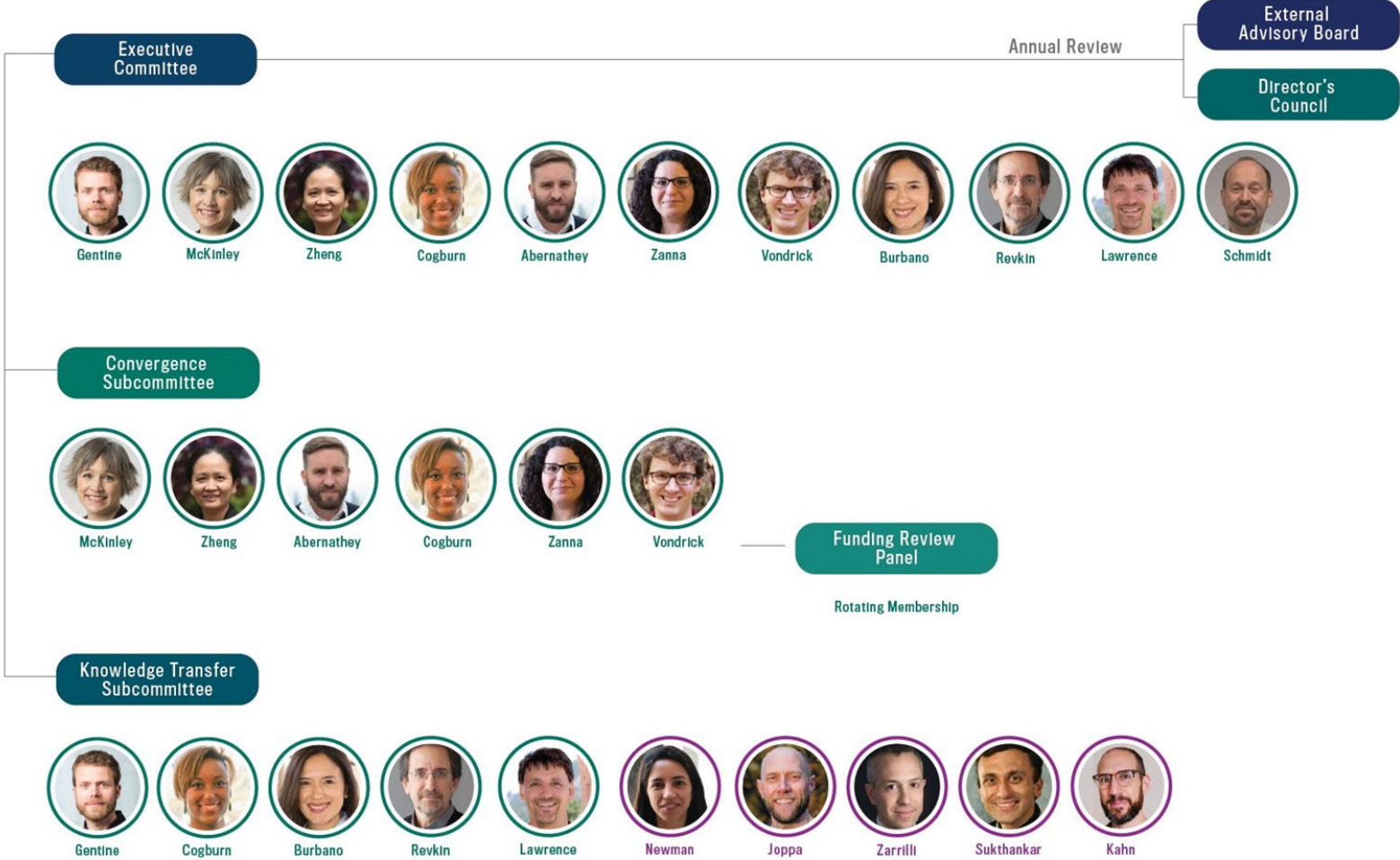
# Governance structure implements Vision through the 5 Convergence Strategies



# LEAP governance is shared + distributed



# LEAP Governance Committees



# Director's Responsibilities



Center Director and Diversity Co-Director  
**Pierre Gentine**

## **Center Director + Diversity Co-Director Gentine**

Chairs Executive Committee

Liaises with NSF + Partners

Co-Leads Broadening Participation Programs

Insures Integration of all Convergence Strategies (CS)



# Deputy Director's Responsibilities



Deputy Director -  
Galen McKinley



Center Director and Diversity Co-Director  
Pierre Gentine

**Deputy Director McKinley**  
Coordinates Research Activities  
Ensures Research + Education Integration  
Co-Chairs Convergence Subcommittee  
Supervises CS2 (algorithms) + CS3 (CESM)



# Chief Convergence Officer + Education Director Responsibilities



## Chief Convergence Officer + Education Director Zheng

Oversees All Educational Programs

Supervises Evaluation

Co-Chairs Convergence Subcommittee

Co-Supervises CS4 (training + broader impacts) with Cogburn

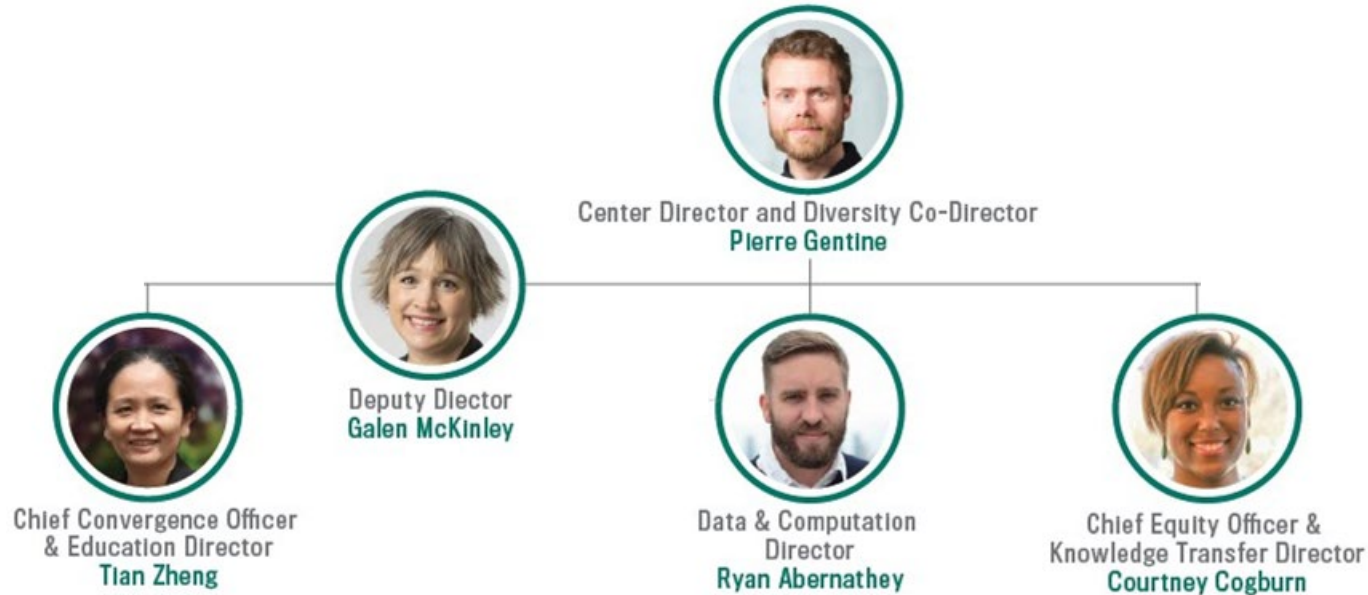
# Data + Computation Director's Responsibilities



## Data + Computation Director Abernathey

- Oversees LEAPangeo
- Supervises Two Staff Engineers
- Ensures Integration of Data Infrastructure + Research
- Co-supervises CS1 (LEAPangeo)

# Chief Equity Officer + Knowledge Transfer Director's Responsibilities



## Chief Equity Officer + Knowledge Transfer Director Cogburn

Chairs Knowledge Transfer Subcommittee

Oversees DEI Assessment of Center + Recruitment

Co-supervises CS4 (training + broader impact) with Zheng

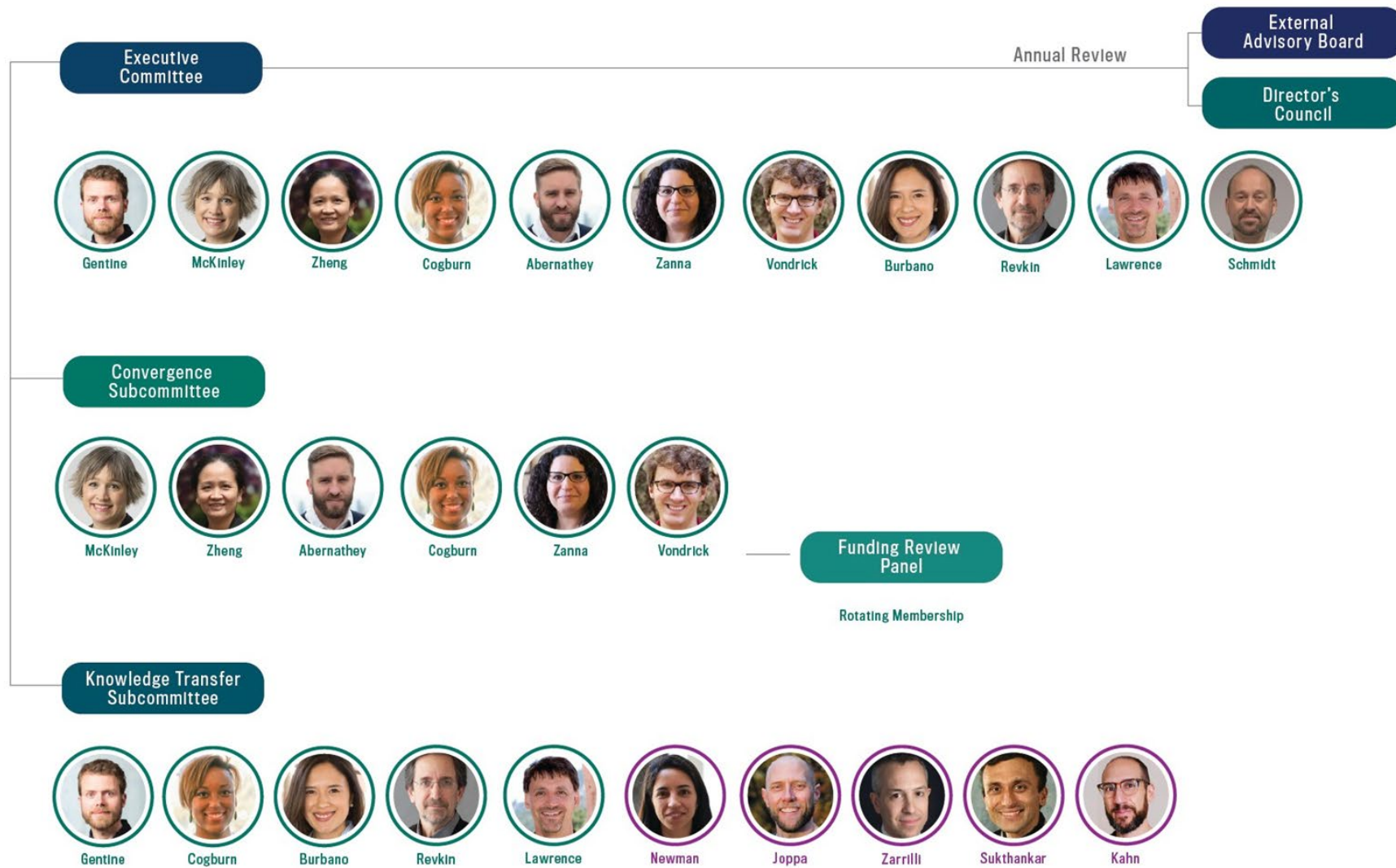
Co-supervises CS5 (knowledge transfer) with Burbano + Revkin



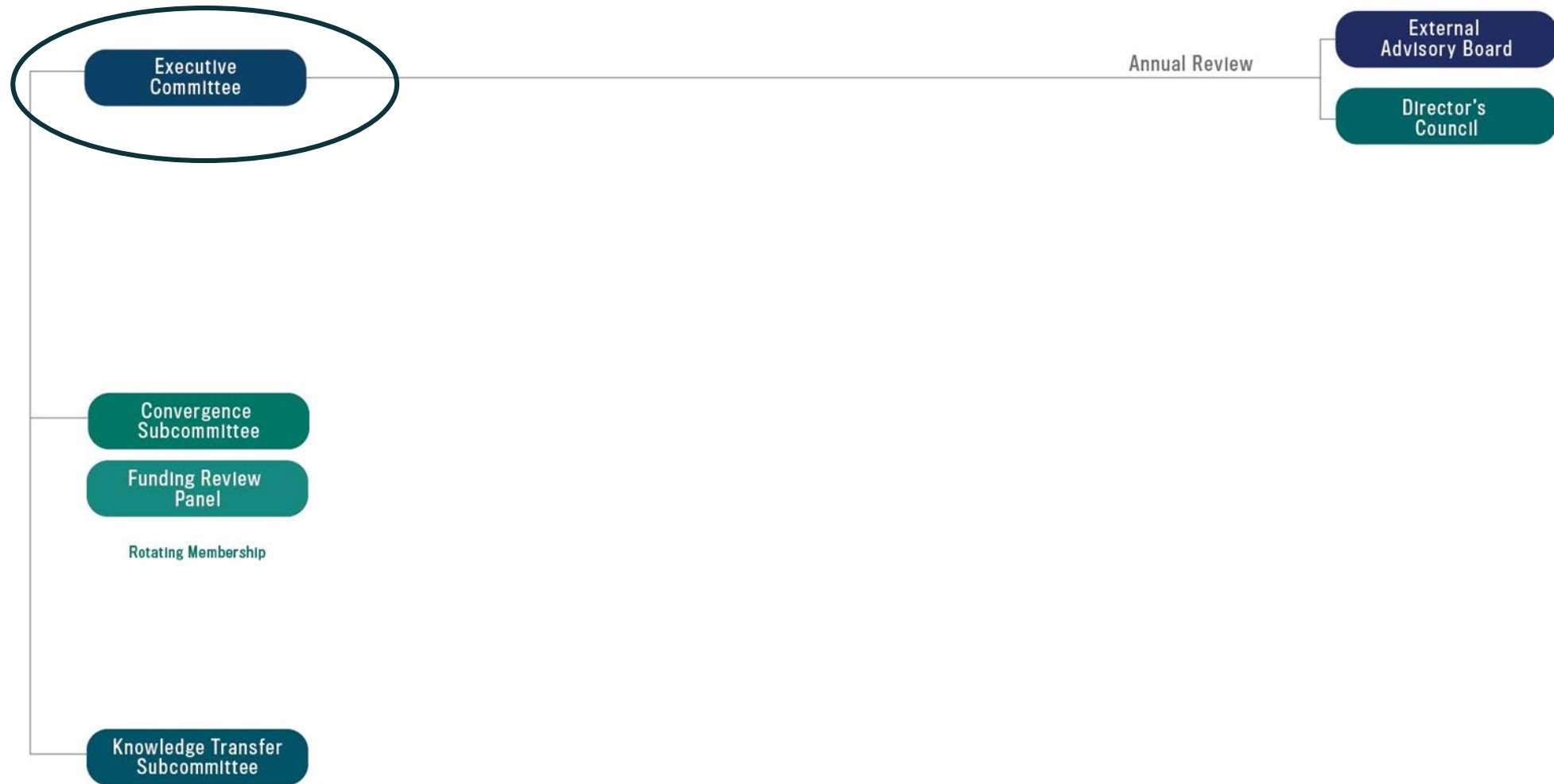
# If needed, directorship succession is clear

- Lead PI Gentine is **initial Center Director**
- If Gentine must withdraw, **Deputy Director McKinley** is Center Director
- If McKinley must withdraw, **Executive Committee votes** on six month-term Interim Center Director
- At six months, **full LEAP team votes** for permanent Center Director

# LEAP Governance Committees



# LEAP Governance Committees



# Executive Committee is the primary decision-making body

## Membership



Gentine



McKinley



Zheng



Cogburn



Abernathy



Zanna



Vondrick



Burbano



Revkin



Lawrence



Schmidt

## Responsibilities

- Strategic planning
- Distribution of funding
- Partnership stewardship
- Conflict resolution
- Meets monthly
- Equal voting
- Chaired by Gentine
- Staffed by Managing Director



# LEAP Governance Committees



# Convergence Subcommittee integrates Research + Education

## Membership



McKinley



Zheng



Abernathey



Cogburn



Zanna



Vondrick

## Responsibilities

- Coordinates review of research and seed funding proposals
- Reviews research progress and sets research priorities
- Integrates research + education
- Plans *Convergence Luncheons*
- Meets monthly
- Co-chaired by McKinley + Zheng
- Staffed by Managing Director

# Formed annually, the Funding Review Panel will score proposed research



- Membership drawn from full LEAP team
  - Changes annually

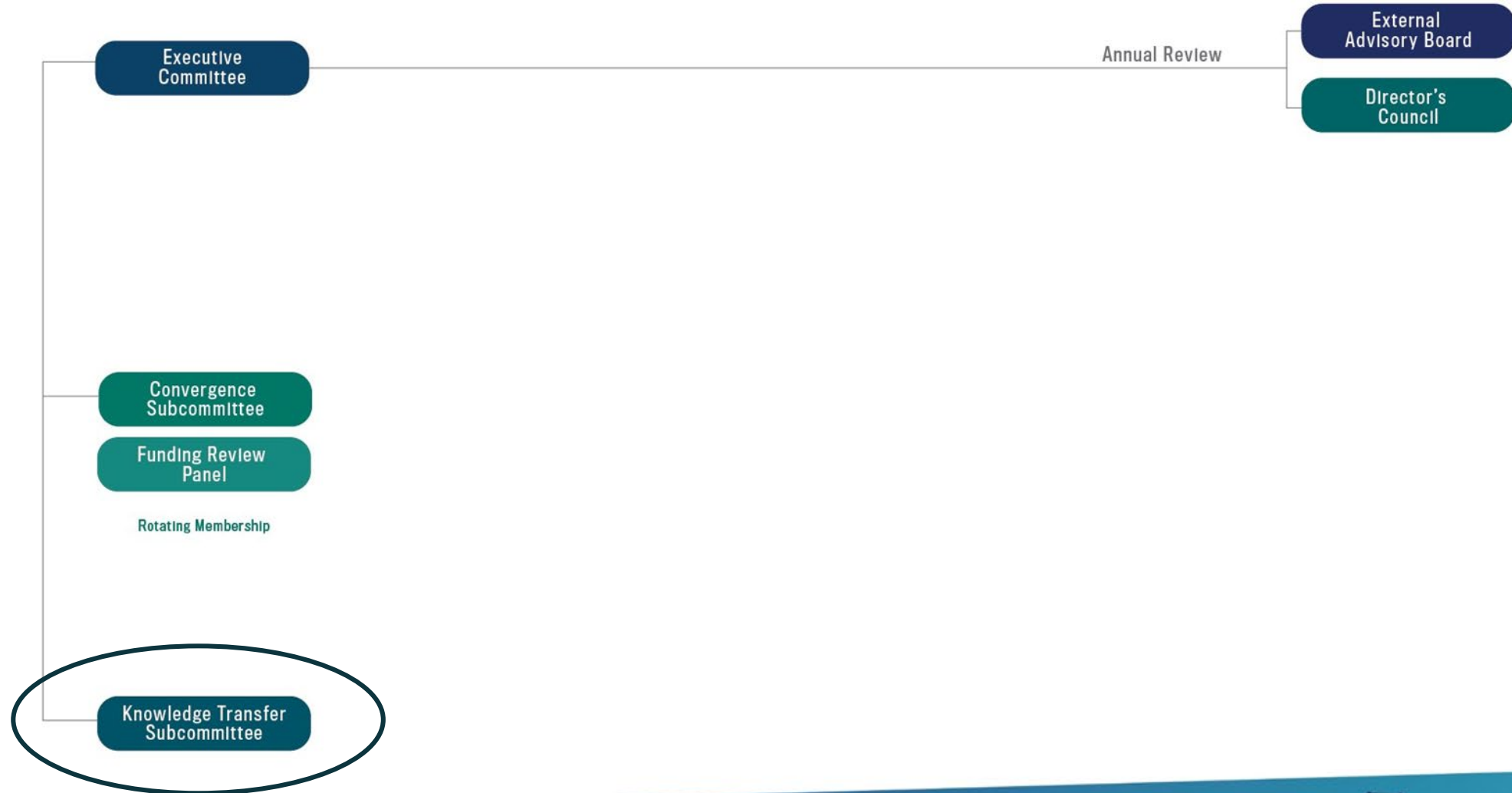
## Membership

- 6 Senior Personnel from all LEAP institutions
  - 1 student + 1 postdoc
- Co-chaired by member of Convergence Subcommittee

## Responsibility

- Rank proposed research + seed funding proposals
- Meets every January
- Staffed by Assistant Director

# LEAP Governance Committees



# Knowledge Transfer Subcommittee drives DEI initiatives and engages stakeholders

## Membership



Cogburn



Gentine



Burbano



Revkin



Lawrence



Newman



Joppa



Zarrilli



Sukthankar



Kahn

## Responsibilities

- Establishes bidirectional dialogue with stakeholders
- Governs *Storytellers* program
- Plans DEI day for annual meeting, Translate-a-thon, Hackathon
- Anonymization for hiring + proposal funding
- Meets monthly
- Chaired by Cogburn
- Includes **5 corporate + public partners**
- Staffed by Knowledge Transfer Program Manager



# Director's Council advises across academic institutions



## Membership

- 6-8 executives across LEAP's universities

## Responsibilities

- March meeting for preliminary Annual Report
- Advises on institutional integration + removing barriers
- Staffed by Gentine, McKinley + Managing Director



# External Advisory Board evaluates LEAP and provides strategic advice

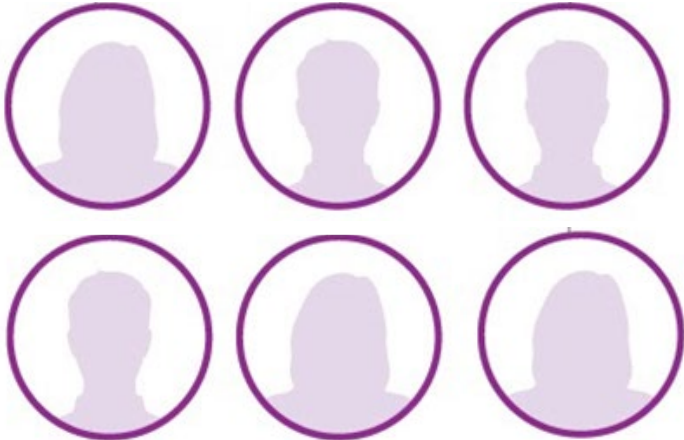


## Membership

- 6 senior researchers + leaders from non-affiliate institutions
- Selected by Executive Committee

## Responsibilities

- Reviews Annual Report in May
- Advises on new partnerships, center strategy + sustainability
- Staffed by Gentine, McKinley + Managing Director



# Continued professional leadership development will support effective center management

## LEAP Management Institute

- LEAP Executive Committee in Year 1, 10 trainees per year thereafter
- 15 hours of courses on leadership + inclusive management

## Executive Coaching

- Gentine + McKinley, 1 hour/ month
- Coach: Management Professor at Columbia Business School

## Dean's Mentorship Program

- Monthly group meeting with Columbia Center Directors
- Led by Columbia's Engineering Dean

## Provost Leadership Fellows

- Two-year program for mid-career Columbia faculty
- Insight into campus strategic initiatives, strategic planning + budget issues
- Zheng finishes in 2021/ Gentine guaranteed spot in 2021-23



# LEAP Staff will provide structured support



Managing  
Director  
TBD

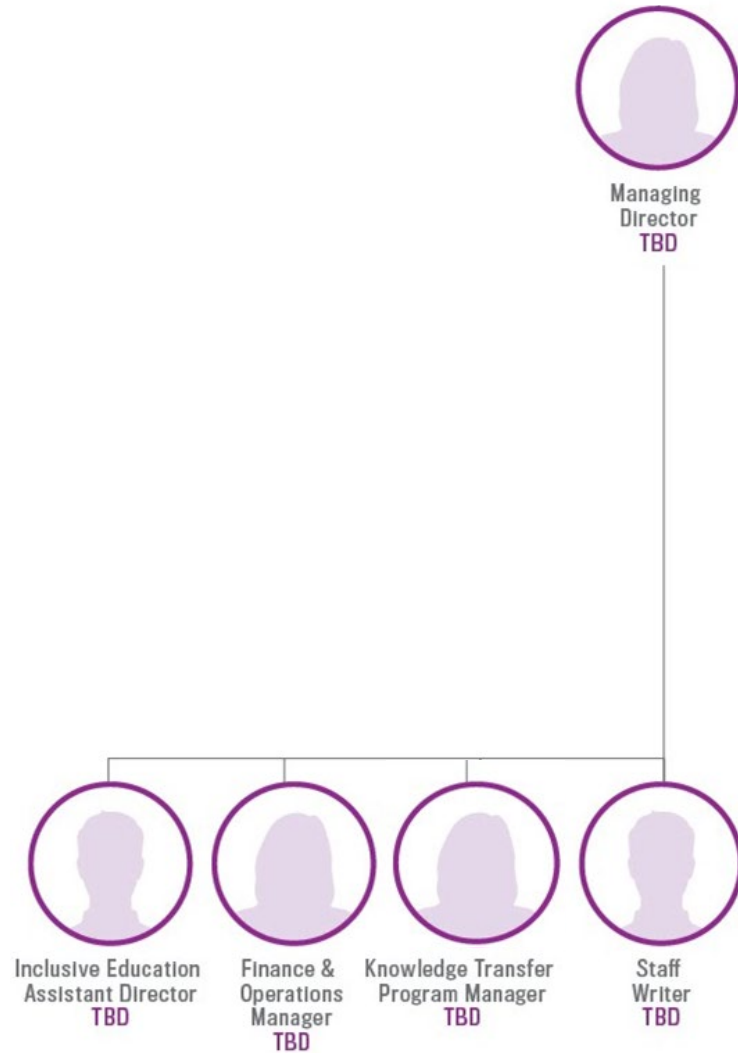
## Managing Director

- Reports to Gentine + McKinley
- Staffs Executive Committee, Advisory Boards
- Responsible for Financial Management, Reporting
- Must Hold Grad Degree in LEAP-related Field
- Employee of Columbia Engineering Dean's Office
- Demonstrated track record on DEI commitment



# LEAP Staff will provide structured support

Managing Director

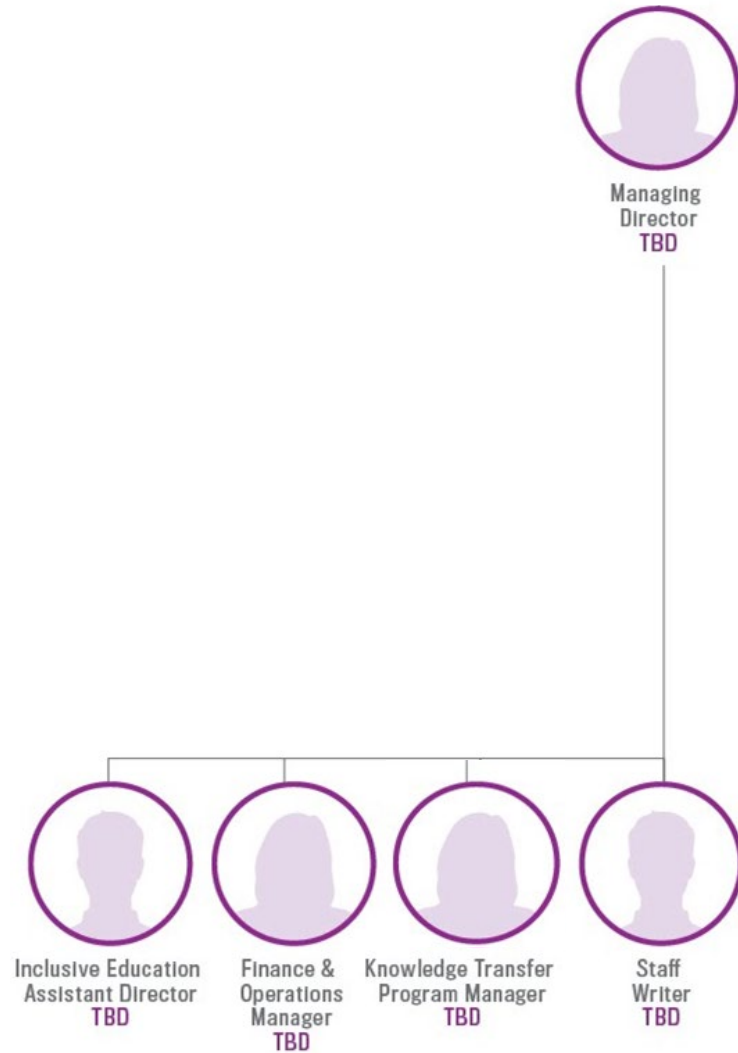


## Inclusive Education Assistant Director

- Implements education and broadening participation programming
- Facilitates student + postdoc visitation to NCAR

# LEAP Staff will provide structured support

Managing Director



## Inclusive Education Assistant Director

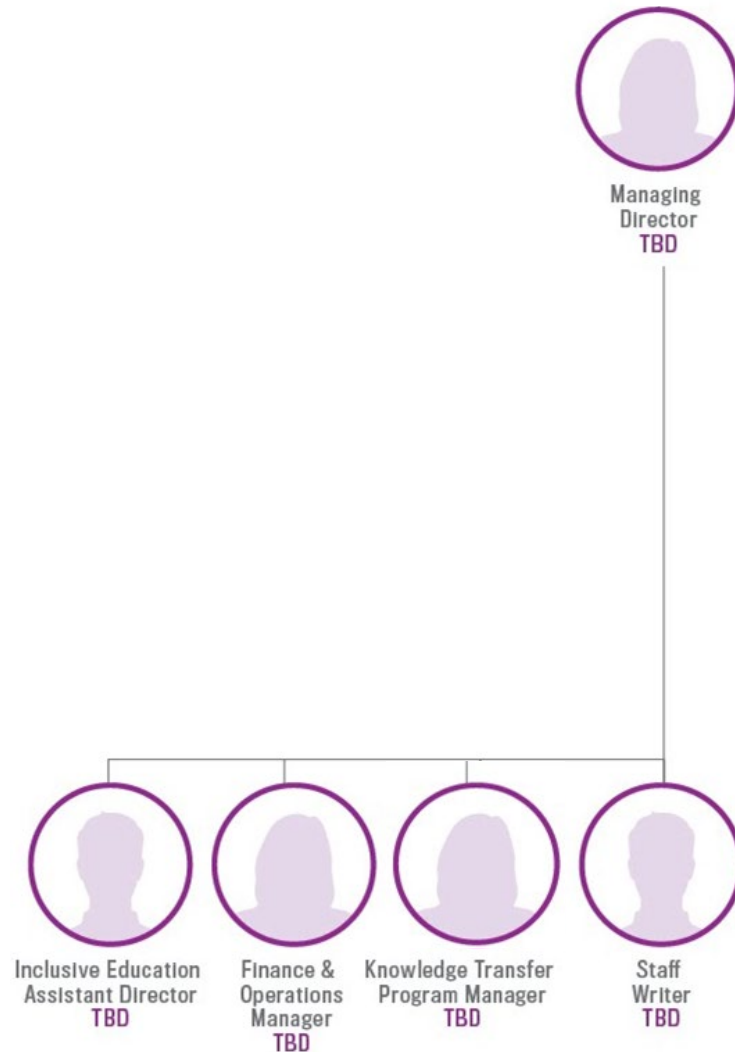
- Implements education and broadening participation programming
- Facilitates student + postdoc visitation to NCAR

## Finance + Operations Manager

- Supports comprehensive financial management, reporting
- Manages Center-wide human resources

# LEAP Staff will provide structured support

Managing Director



## Inclusive Education Assistant Director

- Implements education and broadening participation programming
- Facilitates student + postdoc visitation to NCAR

## Finance + Operations Manager

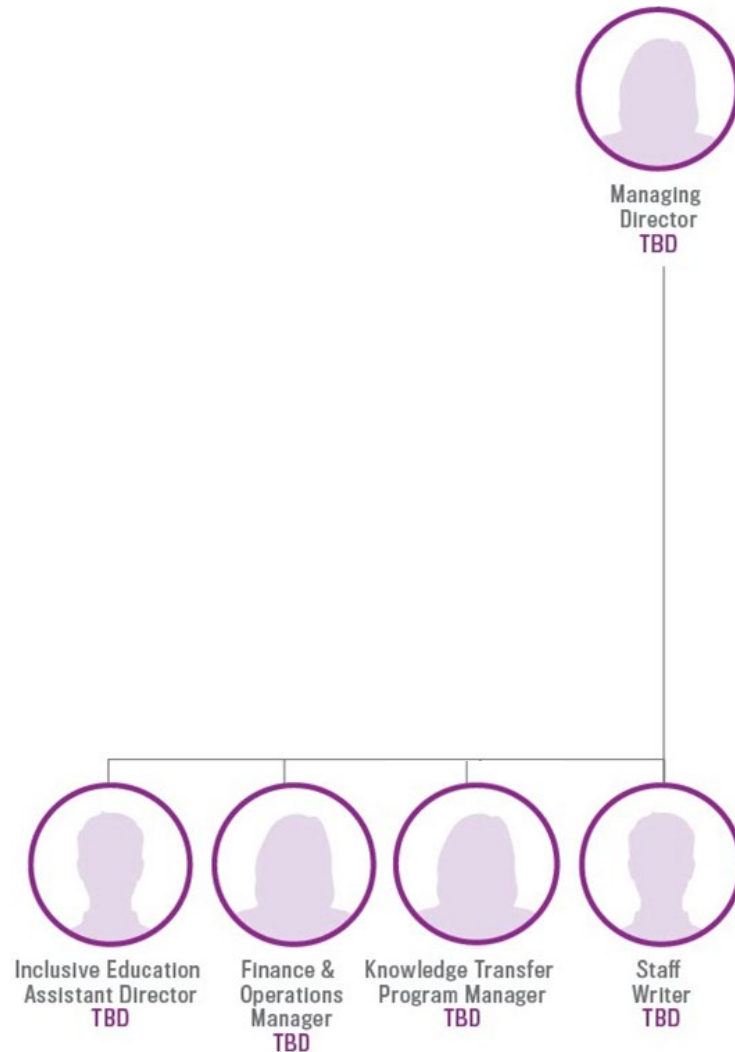
- Supports comprehensive financial management, reporting
- Manages Center-wide human resources

## Knowledge Transfer Program Manager

- Builds + maintains website
- Plans *Convergence Luncheons* + *Public Conversations*
- Staffs Knowledge Transfer Subcommittee

# LEAP Staff will provide structured support

Managing Director



## Inclusive Education Assistant Director

- Implements education and broadening participation programming
- Facilitates student + postdoc visitation to NCAR

## Finance + Operations Manager

- Supports comprehensive financial management, reporting
- Manages Center-wide human resources

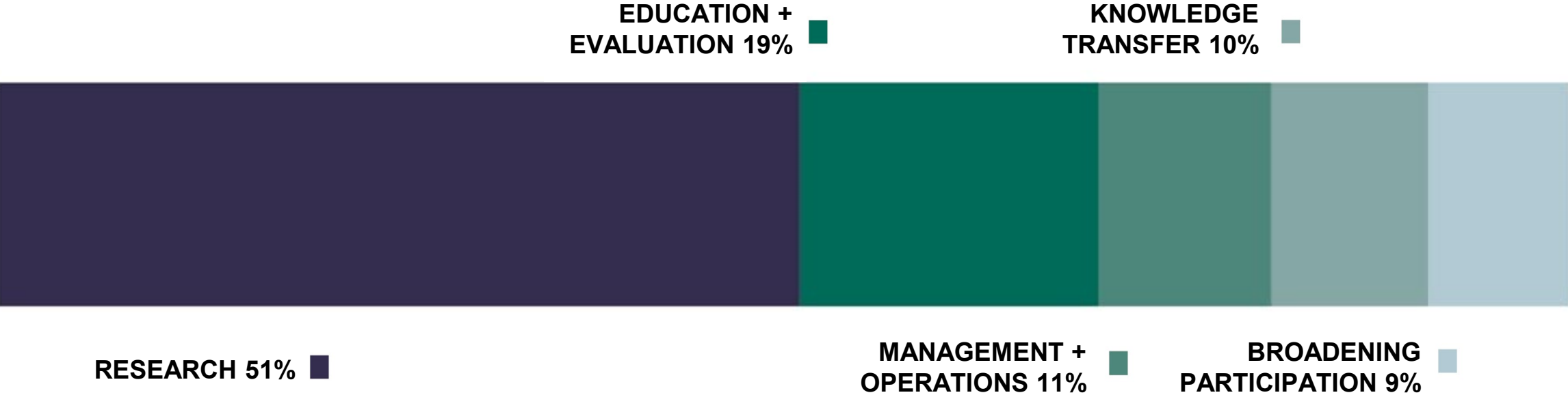
## Knowledge Transfer Program Manager

- Builds + maintains website
- Plans Convergence Luncheons + Public Conversations
- Staffs Knowledge Transfer Subcommittee

## Staff Writer

- Develops articles and OpEds for corporate + public audiences
- Plans *Storytellers in Residence* program
- Primary point of contact for media
- Staffs Knowledge Transfer Subcommittee

# Budget reflects LEAP priorities



# Columbia Institutional commitments are substantial

- **Headquarters Space**
- **Administrative Positions (2)**
  - Finance + Operations Manager
  - Staff Writer
- **NCAR-Based Software Engineer (1)**
- **Research Scientist (1)**
- **“Bridge to PhD” Postbaccalaureate Trainees (3)**
- **Storytellers in Residence (2)**
- **Management Institute (1 per year)**

# Annual meetings allow for reflection, redirection + building Center culture

## Kick-Off Meeting (Month 3)

- Team building
- Setting up goals and procedures (+ docs)

## Subsequent Annual Meetings (August)

- **Day 1:** Review of scientific progress
- **Day 2:** DEI focused training + assessment
- **Day 3:** Reflection + adjusting priorities



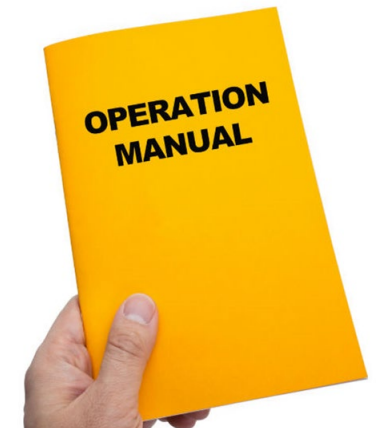
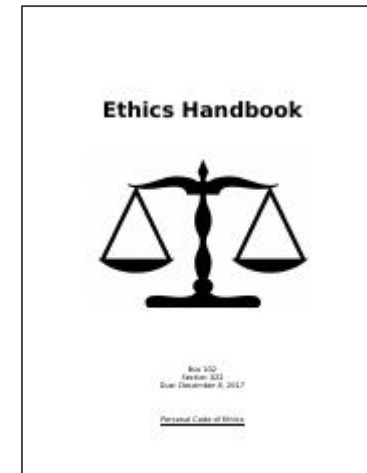
# Documents support efficiency + transparency

## Ethics Handbook

- Expectations will be clearly stated
- **Example Topics**
  - Collaborative roles + responsibilities
  - Authorship policies
  - Mentorship expectations

## Operations Manual

- Standard procedures will be described in one place
- **Example Topics**
  - Annual funding allocation proposals
  - Committee roles + responsibilities
  - Staff roles + responsibilities
  - Conflict resolution



**Operational documents will be updated as needed + reviewed formally at annual meeting**

# Internal + external communication will be multi-faceted

## Annual Report

- Scientific progress
- Education, proactive and integrative DEI initiatives + outcomes

## Monthly Newsletter

- Climate change indicators
- Team highlights
- Practical implications for stakeholders

## Website

- Research + education products
- Event engagement opportunities



Communications supported by Knowledge Transfer Program Manager + Staff Writer

# LEAP is prepared to address challenges as we enter uncharted territories

- **Research:** Applicability of machine learning in CESM  
Potential *bottleneck*: numerical instability + generalization  
*Solutions*: experience with current parameterizations  
+ eigenvalues + regularization + causality

# LEAP is prepared to address challenges as we enter uncharted territories

- **Research:** Applicability of machine learning in CESM  
Potential *bottlenecks*: numerical instability + generalization  
*Solutions*: experience with current parameterizations  
+ eigenvalues + regularization + causality
- **Management**  
Potential *bottleneck*: conflicts from internal funding process  
*Solution*: Seek best practices (CU seed funding)  
+ define priorities + transparent funding review + reduced COI

# LEAP is prepared to address challenges as we enter uncharted territories

- **Research:** Applicability of machine learning in CESM  
Potential *bottlenecks*: numerical instability + generalization  
*Solutions*: experience with current parameterizations  
+ eigenvalues + regularization + causality
- **Management**  
Potential *bottleneck*: conflicts from internal funding process  
*Solution*: Seek best practices (CU seed funding)  
+ define priorities + transparent funding review + reduce COI
- **Diversity**  
Potential *bottleneck*: limited diversity after a few (e.g. 2) years  
*Solution*: increase collaboration with successful pipeline programs  
+ expand reach in NYC's diverse communities  
+ partnership with NSBE (5 URM graduate/year)  
+ SWOT analysis to reallocate funds

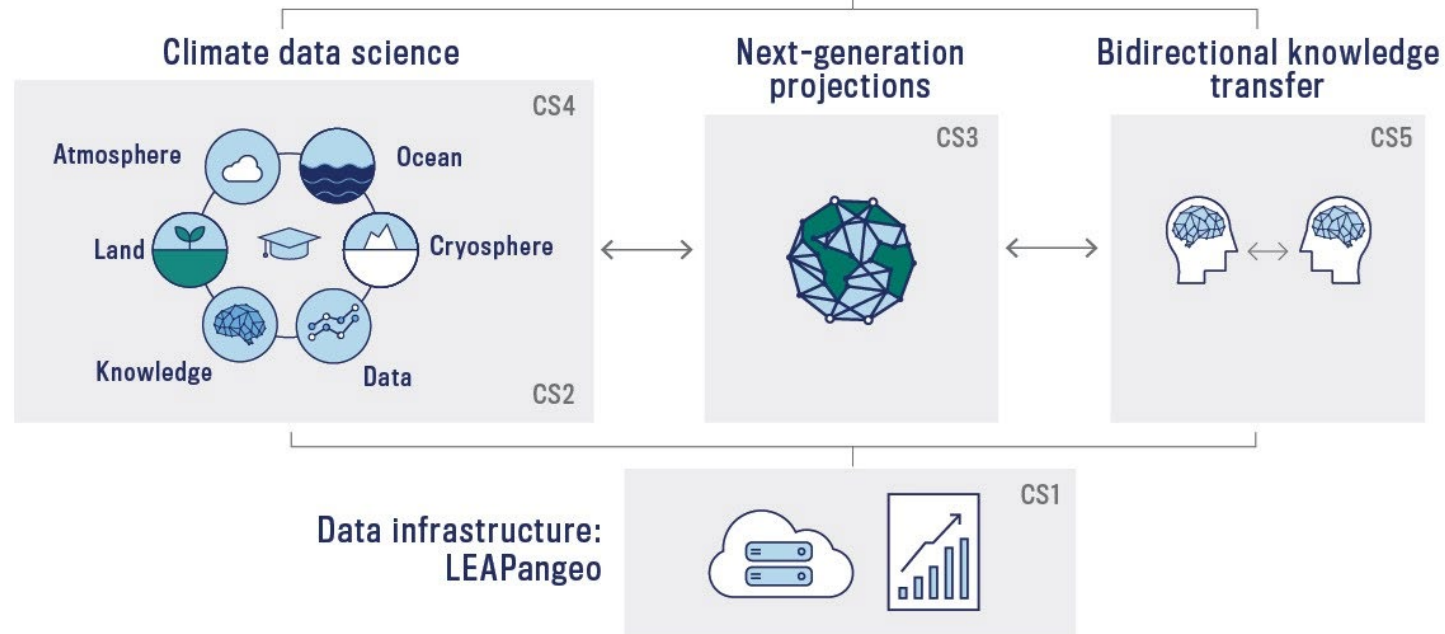
# We are ready for this



# LEAP



Broadening participation



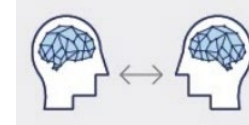
# LEAP

# We are ready for this



LEAP

- LEAP brings *transformative impacts*:



- The proposal process has demonstrated we are *a team*:
  - *Collaboratively* defined governance + priorities + funding
  - Made hard *decisions*
  - Created unique *dialogue* (climate + data science, broader impacts)

- We are ready and uniquely positioned to execute our vision:

**LEAP forward in the reliability, utility, and reach of climate projections through synergistic innovations in data science and climate science**



LEAP

Questions?