LEAP

Name	Department	Institution	Center Role	Relevant Expertise
Pierre Gentine	Earth & Env. Eng.	Columbia	**Lead PI**, Center Director, Diversity Co-Director	Hydrologic, carbon cycles, land-atmosphere interaction, management
Galen McKinley	Earth & Env. Sci.	Columbia	**Co-PI**, Deputy Director	Carbon cycle, ocean biogeochemistry
Ryan Abernathey	Earth & Env. Sci.	Columbia	**Co-PI**, Data & Computation Director	Ocean circulation, climate dynamics, high-performance computing
Courtney Cogburn	Social Work	Columbia	**Co-PI**, Chief Equity Officer, Knowledge Transfer Director	Environmental justice, data science, psychology, virtual reality education
Tian Zheng	Statistics	Columbia	**Co-PI**, Chief Convergence Officer, Education Director	ML, spatiotemporal modeling, pattern identification, education, management
Vanessa Burbano	Management	Columbia	Corporate Engagement Director	Corporate sustainability, strategy
David Lawrence	Climate Dynamics	NCAR	Model Development Liaison	Arctic feedbacks, permafrost
Andrew Revkin	Earth Institute	Columbia	Public Engagement Director	Communication, partnerships
Gavin Schmidt	Goddard Institute	NASA	Model Development Liaison	Climate modeling, management
Carl Vondrick	Computer Sci.	Columbia	Data Science Director	Computer vision, machine learning
Laure Zanna	Mathematics	NYU	Geoscience Director	Oceanography, prediction, climate

Non-Lead Personnel (Total: 42)

Data Science (12): Bareinboim, Bengio, de la Peña, Fish, Gelman, Kaiser, Kim, Kpotufe, Kumar, Lipson, Rush, C.Zheng.

Geoscience (11): Bell, Camargo, Elsaesser, Giometto, Goddard, Kingslake, Lall, Pincus, Pritchard, Raymo, van Lier-Walqui.

Federal Lab (6): Gagne, Gettelman, Hamman, Lamarque, Long, Morrison.

Ecology (5): DeFries, Gomes, Menge, Uriarte, Weng.

Social Science (4): Blikstein, Lang, Pizmony-Levy, Purdie-Greenaway.

Executives (4): Joppa, Newman, Sukthankar, Zarrilli.

A. PROBLEM DESCRIPTION AND RATIONALE FOR CENTER APPROACH.

Climate change poses unprecedented risk for human sustainability [1,2]. This grand challenge threatens civil infrastructure, agriculture, public health, economic security, and international peace, with our world's most underserved communities disproportionately facing the greatest burden [3,4]. Global adaptation, such as building sea walls, is projected to cost between $280B and $500B (USD) per year [5]; early adaptation will maximize efficiency, resiliency, and cost-effectiveness [6]. However, most public and private climate adaptation decisions are stalled or misguided [7] because:

1. Climate-forcing factors (population and economic growth, technology development, land use) cannot be precisely predicted, leading to *scenario uncertainties*. These uncertainties dominate climate estimates after ~40 years in the future (Figure 1) [8].
2. Climate models, which project future climates, are *too imprecise* [9]. For example, by 2055 and for moderate emission pathways, the likely range of global temperature increase lies between 0.9° and 2.0°C; global sea level will rise between 0.19m and 0.33m [10]. Regional projections are even less precise [8]. Such model uncertainties exist across all timescales, but are dominant for the near-term (10-40 year), directly impacting mitigation strategies. Thus, climate projection uncertainties are deleteriously obstructing today's decisions regarding adaptation, investment, security, and resiliency [11].

Climate projection uncertainties stem from climate *model deficiencies* in representing physical and biological processes [8] because: 1) these processes occur at scales *too small* to be explicitly resolved on models' coarse (~100km) spatial grid; and 2) *basic scientific knowledge remains incomplete*. Overcoming these science and technology barriers is *the* grand challenge for climate scientists and governments alike.

But now, the proliferation of observational technologies continuously monitoring the Earth's system components – the **atmosphere**, **ocean**, **land**, and **cryosphere** [12] – provides an avalanche of new data that offer dramatic potential for placing observational constraints on these many physical and biological processes. Further, high-resolution models can now *simulate* key small-scale and fast processes (e.g., clouds, ocean eddies) with great accuracy, though only for brief periods of time and over small geographic areas due to their staggering high computational cost [13,14]. These timely revolutions combine to reveal a distinctive opportunity for harvesting data to improve *climate model projections*. Yet, only a small fraction of available climate data is currently used (Figure 2) because they are extremely challenging to fully exploit.

In comparably data-rich fields (e.g., health, security), modern data science has propelled efficient exploitation of big data. Similarly, **data science is uniquely poised to reshape climate modeling, dramatically increasing its societal utility**.

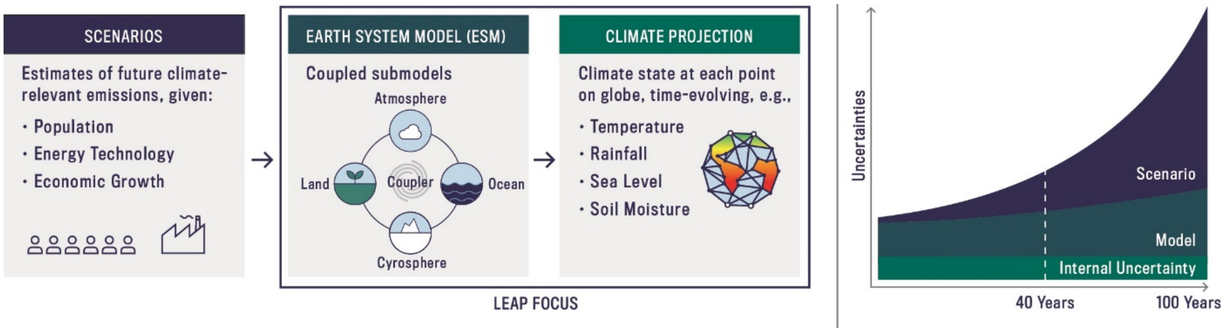


Figure 1: Projecting future climate requires the development of scenarios for different plausible socio-economic and technological developments (left). The resulting emissions and land use scenarios are fed into climate models (i.e. Earth system models) to make projections (middle). These projections are uncertain (right) because of internal (stochastic) variability, model errors, and scenario uncertainty. Model errors dominate total uncertainty for up to ~40 years and remain important on longer timescales [8].

Three substantial challenges stand in the way of reaching this ambitious goal: 1) machine learning must transcend traditional methods that efficiently interpolate but do not extrapolate to unseen conditions, and that do not respect physical constraints; 2) the need for efficient data use and access will require reshaping modern climate data infrastructure; and 3) more accurate projection knowledge must be transferred to public and private stakeholders to enable informed adaptation strategies.

Such ambitious goals are only achievable through a new **STC: Center for Learning the Earth with Artificial Intelligence and Physics (LEAP)**, tightly integrating geoscience with data science to revolutionize global and regional climate projections. By uniting these two disciplines, LEAP will dramatically improve climate model projections in ways impossible to achieve independently by firmly **embedding physical and biological knowledge into machine learning to transform Earth system modeling**. This integration will spawn a new era of machine learning (ML) algorithms using physical knowledge to robustly extrapolate to future conditions (Figure 2). By integrating these algorithms into the Community Earth System Model, LEAP pledges a high-impact legacy for the Earth system modeling community. Simultaneously, LEAP will transfer knowledge gleaned from its novel projections to a heterogeneous cohort of public and private sector audiences to promote effective climate adaptation, and equitably train the next generation of diverse learners in the emerging new discipline of **climate data science**.

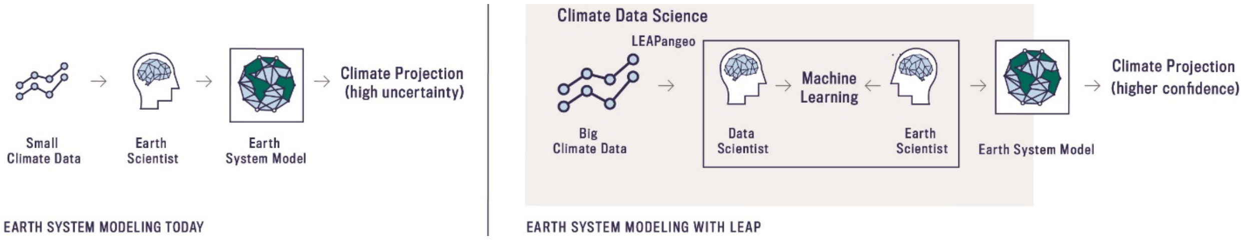


Figure 2: Current Earth System Modeling compared to LEAP's approach. By using ML to utilize much more climate data, LEAP will dramatically improve near-term (10-40 years) climate projections.

A2. Challenges and Opportunities – Better Leveraging Climate Data. Confronting these scientific challenges, LEAP will empower climate data science research by applying **five convergence strategies**:

1. **CS1:** Launch *LEAPangeo*, an open cloud computing platform engaging the broader community at the intersection of geoscientific and ML research (see Shared Experimental Facilities);
2. **CS2:** Develop next-generation ML algorithms, employing physical knowledge and causal mechanisms to robustly extrapolate and generalize, and improving near-term climate projections;
3. **CS3:** Harness new ML methodologies to improve the NSF-funded Community Earth System Model by leveraging the wealth of existing data to advance model process representation, developing new data products when needed, and applying novel model evaluation metrics;

4. **CS4:** Exploit these transdisciplinary studies to equitably and inclusively train the next generation of climate data science scholars [15]; and,
5. **CS5:** Establishing a bidirectional communication pathway between academia, industry, and the public, forging mutual stakeholdership in the science, utility, and humanity of climate projection.

A3. Integrative Partnerships. LEAP unites **48 researchers, five executive administrators**, **six academic institutions**, **two federal research labs**, and **27 institutional partners**, leveraging an expertise impossible to achieve independently; henceforth, all personnel and partner names are highlighted in **bold**.

The Center will be headquartered at **Columbia University in the City of New York**, a global leader in climate and data sciences. Housing ~1,000 climate faculty and researchers, primarily within its Lamont-Doherty Earth Observatory and Earth Institute, **Columbia** has contributed to more Intergovernmental Panel on Climate Change reports than any other American university. In September 2019, **Columbia**'s President launched a Climate Task Force, including five LEAP team members (**Gentine**, **Bell**, **Cogburn**, **DeFries**, and **Raymo**), to plan a new **Columbia** School of Climate. In tandem, **Columbia**'s Data Science Institute (DSI) spans 350 faculty advancing the state-of-the-art in data science and transforming disciplines for societal benefit. Most of LEAP's **Columbia** faculty are DSI affiliates. **T.Zheng** is DSI's Associate Director of Education, and **Cogburn** is Co-Director of DSI's Computational Social Science Working Group.

LEAP critically partners with the NSF's **National Center for Atmospheric Research** (NCAR), to transform the Community Earth System Model (CESM), a community global model providing state-of-the-art simulations of the Earth's past, present, and future climate. Second, NASA's **Goddard Institute for Space Studies** (GISS) modeling center, located four blocks from **Columbia**, will collaborate to integrate LEAP's algorithms into their ModelE. Though prohibited from receiving STC funding, both federal labs are fully committed partner organizations represented on LEAP's Executive Committee and Subcommittees.

Columbia, **NCAR**, and **GISS** are complemented by expertise from five academic partners: **NYU** provides trailblazing expertise in ML for ocean modeling; **UC Irvine** blends ML with computational atmospheric physics within a Hispanic-serving Institution; **Minnesota** brings physics-guided ML; **Montreal** brings world-class expertise in ML and its applications to climate change; **Teachers College** ensures integration with the **New York City Department of Education** and innovative environmental pedagogies. Finally, LEAP boasts an additional 27 external partnerships, most notably technology companies (**Microsoft** and **Google**), the **National Society of Black Engineers**, the **Trust for Governors Island**, and the **New York City Office of the Mayor**.

LEAP pledges a high-impact and sustained contribution to the United States' leadership in both climate science and artificial intelligence technology and innovation [16]. Through LEAPangeo, all LEAP datasets and research products will be made fully open-source and readily available for universal adoption. New algorithms will be developed in collaboration with **NCAR**'s CESM, fundamentally transforming how national climate centers utilize and integrate data (Figure 2). *This* is LEAP's legacy.

A4. Bolstering National Priorities. LEAP advances **six** NSF Big Ideas: It 1) *Harness[es] the Data Revolution* by developing new algorithms that include physical knowledge and causal mechanisms to better extrapolate, while leveraging the wealth of climate data; it develops 2) *Midscale Research Infrastructure* via cloud computing and accessibility. In recognizing that projection accuracy requires insight from traditionally separate disciplines, LEAP 3) *Grow[s] Convergence Research* by coordinating and sustaining climate data science as a new discipline, while preparing a diverse workforce for the 4) *Future of [Climate] Work* by fostering lifelong learning and technology literacy. LEAP's research will allow for 5) *Navigating the New Arctic* through improved projections of retreating ice sheets and better quantifying high-latitude land and ocean processes. Finally, LEAP 6) *INCLUDES* diverse trainees by spearheading formal and informal programming, mentorship, and recruitment to equitably support all learners in meeting their training objectives.

A5. Competitor Analysis and Synergies. LEAP's transformative approach to Earth system modeling has not been pursued before. Caltech's CLIMA effort develops systematic approaches to model parameter optimization using ML albeit relying upon traditional empirical parameterizations. Instead, LEAP unleashes ML to: 1) address model structural deficiencies with new parameterizations; and 2) develop new data products and model skill metrics. A new effort by VULCAN, a private company, focuses upon daily and weekly weather and seasonal to interannual climate forecasts. It uses ML for parameterizations of atmospheric processes, but does not tackle all Earth system components or multi-decadal climate change (10-40 years). **Gentine** is co-funded on USMILE, a European Research Council Synergy project using ML to develop superior observational products for the atmosphere (clouds and convection) and land (gridded carbon and water fluxes, soil moisture, and biomass), and to enhance data processing of climate model output via ML. LEAP will synergistically use these new land and atmosphere observational products for model evaluation (Section B6). **Kumar** completed an NSF Expeditions in Computing grant developing ML for climate and ecosystem data (climate model output, remote sensing, and *in situ*), refining local climate change impact. This project focused upon analyzing ESM outputs and data, not LEAP's focus of exploring ML-enhanced ESMs climate projections. Nevertheless, LEAP will leverage this Expeditions' ideas and momentum.

B. DESCRIPTION OF THE RESEARCH OBJECTIVES OF THE CENTER.

B1. Earth System Modeling and the Parameterization Deadlock. Today's climate models, called Earth System Models (ESMs), simulate physical and biogeochemical processes in the atmosphere, ocean, land, and cryosphere to replicate historical and project future climates. ESMs solve thousands of equations, involving thousands of parameters. Climate projection simulations are typically run on a coarse spatial grid (~100 km horizontal resolution, Figure 3) due to the limited computational capacity of present-day high-performance computing systems. Biological and physical processes occurring at scales smaller than the grid resolution are empirically represented using "**parameterizations**": mathematical representations of processes' impact on the coarse scale (e.g., heating due to subgrid clouds), based upon equations with adjustable free parameters (e.g., a diffusion equation with a diffusion coefficient parameter).

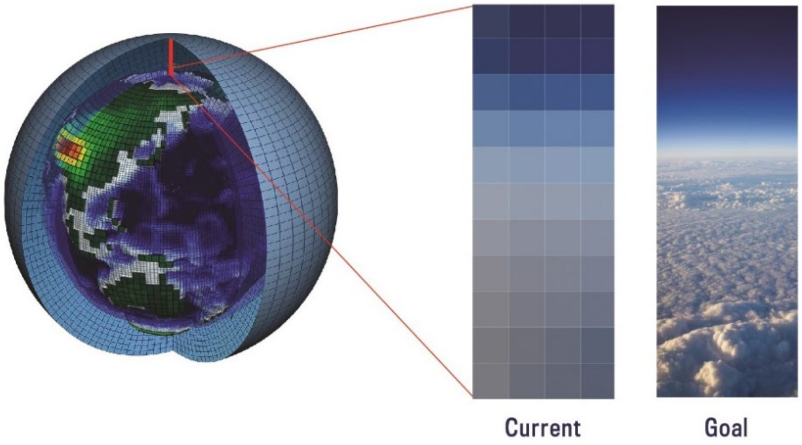


Figure 3: Today's ESMs (left) represent key climate processes such as clouds only coarsely (~100km resolution). LEAP will use ML to better represent climate effects of these and other small-scale processes (right). LEAP's goal is to achieve climate modeling with a quality equivalent or superior to high-resolution simulations.

Internal climate variability notwithstanding (Figure 1), near-term (10-40 year) climate projection uncertainties are primarily due to two parameterization errors [17]:

1. **Structural Errors** stemming from poor or missing representation of small-scale subgrid processes within the coarse ESM resolution (e.g., clouds, convection, turbulence) [20] or from incomplete scientific knowledge of these processes (e.g., land and ocean biology, ice flow).
2. **Model Parameter Errors**: ESM parameters are calibrated to achieve simulations that reasonably replicate observations, but this calibration step, called "**model tuning**," is insufficiently automated and leverages only a small subset of available data.

ESM developers compare model results against historical observations to evaluate and revise model structure and parameters (Figure 2). ESM subcomponents (atmosphere, ocean, etc.) are first run independently and evaluated against observations. Then adjustments to structures and parameters are made to improve comparisons with data, iterating until the model is considered adequate. Subcomponents are then coupled together, run again, and re-tuned to address new discrepancies. This cumbersome and labor-intensive process requires deep geoscientific and modeling expertise, but the resulting outputs *still* exhibit high biases (e.g., clouds [13,18,19], convection [20,21], and ocean eddies [22,23]) due to flawed model structures that cannot be corrected by parameter tuning alone [24]. This long-standing "**parameterization deadlock**" has impeded ESM climate projection for too long [18].

B2. Critical Proofs of Concept.

B2.1. ML for Earth System Modeling.

LEAP's team has pioneered using ML to: 1) improve ESM model structure; 2) estimate model parameters; and 3) develop new data products constraining key processes. Atmospheric convection is a long-standing roadblock for accurate climate projection [26,27], though many convection biases are alleviated in higher-resolution Cloud Resolving Model (CRM) simulations [14,25,27]. To reduce structural errors in coarse ESMs, **Gentine** and **Pritchard** demonstrated that deep neural networks could learn complex convection processes using CRM data, coarse-grained to the ESM's resolution (Figure 4). This novel, data-driven parameterization substantially improved upon traditional convection parameterizations, but still experienced difficulty generalizing to much warmer climates [25]. Similarly, **Zanna** pioneered applying deep neural networks to ocean eddy momentum transport, learning coarse-grained Reynolds stresses (i.e., shear) as a function of the coarse-scale state, using eddy-resolving ocean simulations [22]. This new ML-driven ocean eddy transport parameterization was successfully tested in an idealized coarse-resolution ocean model, yet the neural network led to unphysical instabilities when extrapolating to unseen conditions.

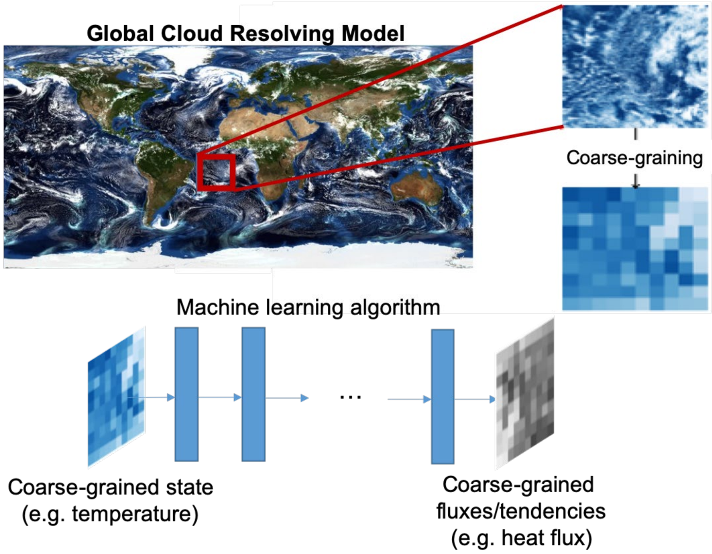


Figure 4: Example of strategy to use high-resolution global cloud-resolving model (CRM) simulations, similar to the successful approach of **Gentine** *and* **Pritchard** *[21,25]. High-resolution CRM data are coarse-grained to the scale of the ESM (~100km). The impact of convection on the resolved coarse-scale variables is learned by an ML algorithm and applied into the ESM.*

In parallel, **Elsaesser** and colleagues demonstrated that neural networks could systematically estimate ESM parameters by efficiently emulating perturbed ESM simulations [28]; the retrieved parameters greatly improved the **GISS** ModelE ESM's skill. Further, **McKinley** and **T.Zheng** have been developing new ML-based extrapolations of sparse ocean data into global gridded products, with a particular focus on quantifying uncertainty so as to make these products more useful as ESM metrics [29]. Together, these proofs of concept demonstrate that a ML able to extrapolate is well-positioned to transform data usage and better represent complex ESMs' physical and biological processes (Figure 2), thus finally breaking the parameterization deadlock. **Now is the time to expand and improve upon these efforts for full development and ESM implementation**.

B2.2. ML Benefiting from Domain Knowledge. LEAP's geoscience problems will spark new ML algorithms that robustly extrapolates to novel scenarios and conditions, pivotally advancing ML's applications and impact. The LEAP team has been leading the development of ML algorithms for learning predictive models from perceptual data. **Vondrick** showed that generative models trained upon large amounts of unlabeled video can anticipate future scene dynamics [30]. Explicit spatial constraints improve performance, allowing *extrapolation* of videos. In preliminary work, **C.Zheng** introduced ML models that generalize over chaotic dynamics, using multi-scale representations focusing upon large-scale prediction, whose dynamics are smooth and less sensitive to perturbations. **Lipson** pioneered the development of algorithms that retrieve equations behind laws of physics using only data [31]. **Google** collaborators developed data-driven ML algorithms for improving numerical solutions of partial differential equations [32]. Altogether, these results suggest that constraints on problem structures, when *informed by domain knowledge,* lead to better ML models and more robust generalization. Further advances will require deeper intertwining of physical and biological knowledge with ML so as to best take advantage of the large volumes of data. To realize this, an NSF Science and Technology Center is required that spans expertise across both geoscience and data science.

B3. Center Research Strategy. LEAP's strategy is dependent upon trade-offs between data availability and knowledge. When data are largely available and spans most conditions (i.e., the whole distribution), they can be harvested by established, mostly-interpolating, ML algorithms. When data are limited, physical

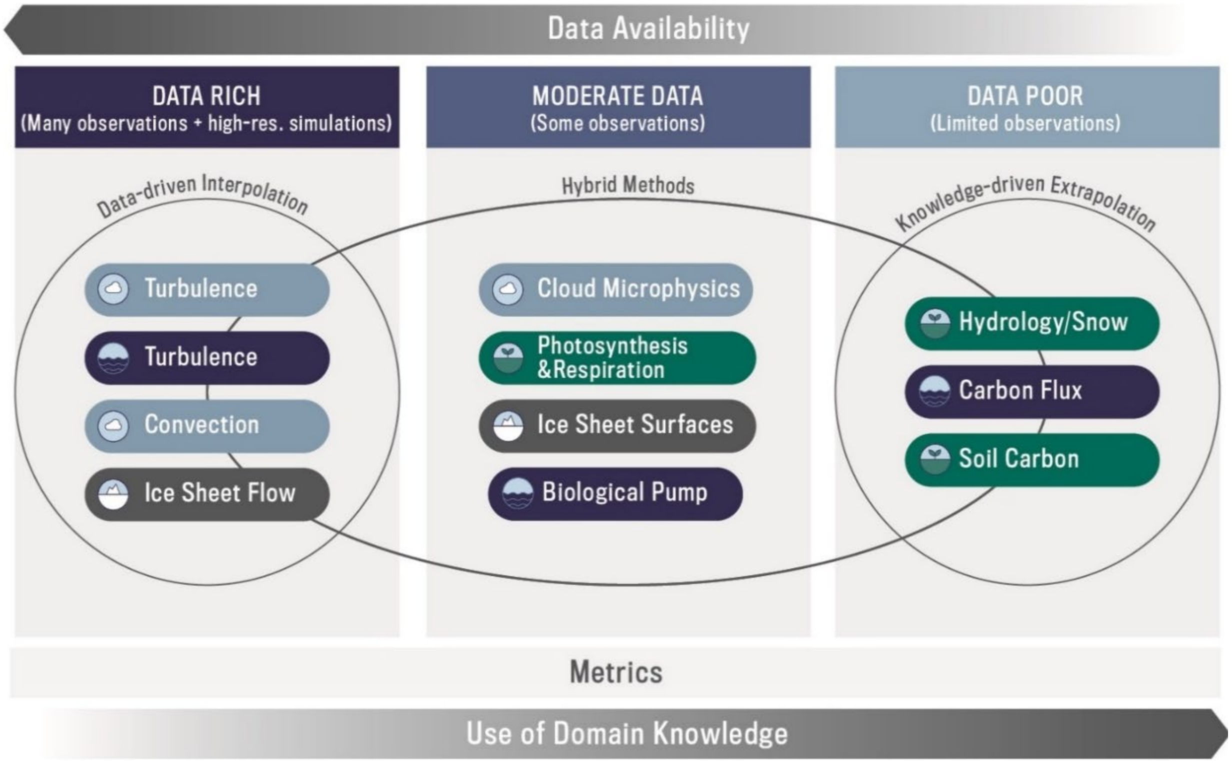


Figure 5: Schematic of LEAP's intersecting geoscience and data science research. ML strategies are determined by data richness of each Earth system problem. When data are poorer, more domain knowledge is used. Throughout, models are evaluated against data, "Metrics."

and biological knowledge can efficiently guide the learning process. LEAP will therefore develop and evaluate a spectrum of ML strategies within a "**Knowledge-Data Continuum**" (Figure 5) to reduce structural and parameter errors in ESM parameterizations, while developing novel data products and metrics for model evaluation. LEAP will use a range of approaches leveraging high-resolution simulations and observational data. LEAP's CESM research projects lie along this Knowledge-Data Continuum, eschewing traditional disciplinary divides that commonly handicap convergent research [33].

1. **Data Rich**: Here, short-term high-resolution simulations (e.g., Cloud Resolving Models) that alleviate many biases present in coarse-scale ESMs [27] can be generated over a wide range of conditions; the ML problem then becomes one of *interpolation*, wherein established methods excel. These can be further combined with observational data to provide strong constraints on processes. LEAP will learn ML parameterizations optimally based upon: 1) coarse-grained, high-resolution, process-resolving simulations (Figure 4); and 2) observational data (*in situ* or remote sensing). Observations will correct high-resolution simulations' remaining biases caused by partially-resolved processes. LEAP will develop ESM modules with qualities *equivalent or superior to high-resolution* process-resolving simulations (Figure 3), optimally leveraging both high-resolution and observational data.
2. **Moderate Data**: Here, governing equations are unknown (e.g., ecosystem respiration), and high-resolution simulations cannot be leveraged to gain understanding. However, observational data are available, though they do not span the full range of present and future environmental conditions. Thus, established ML methods will be challenged to extrapolate. In such cases, LEAP will use **causal inference** to drive the ML algorithms or **equation discovery** to learn data-driven equations that can be implemented to improve the ESM parameterization structure. This will be complemented by systematic **parameter estimation** using a Bayesian approach.
3. **Data Poor, Knowledge-Driven Extrapolation:** With minimal data but sufficient knowledge, LEAP will either: 1) use **process understanding, relating observations, and ML to extrapolate** from sparse measurements to develop novel gridded data products, and then use these products to refine ESM structure and parameters; or 2) use **knowledge-guided ML** (Section B4.3). Both techniques will be complemented by **parameter estimation**.

To establish convergence and forge the new climate data science discipline, **each of the following 23 Knowledge-Data Continuum research projects pair geoscientists with data scientists**.

B4. Data Science Research. ML leverages data to develop statistical models. Key to LEAP's success will be: 1) efficiently unleashing *LEAPangeo* onto large datasets (see Shared Experimental Facilities); and 2) generalizing to situations outside of training sets while respecting physical constraints. To tackle this, **LEAP "hybridizes" ML to include Earth system knowledge**. Throughout, LEAP will investigate ML methods that use varying degrees of physical and biological knowledge to constrain algorithms and robustly capture system dynamics, thereby ensuring physically consistent projections. Physical constraints additionally reduce hypothesis space sizes, thus improving learning efficiency. Depicted in Figure 5, LEAP will employ ML strategies ranging from the data-driven to the knowledge-driven, with many hybrid approaches lying in between these two polarities.

B4.1. Data Rich: Data-Driven Interpolation.

In cases where many data are available, from high-resolution simulations or remote sensing observations, spanning most of the range of conditions (e.g., temperatures), LEAP will use *established* ML methods targeting *performant interpolation* and that do not require Earth system knowledge (e.g., deep or convolutional neural networks, random forests) [34], and will trigger improvements using:

Loss Functions (**Abernathey**, **Vondrick**, **Kpotufe**). The loss functions used for training the ML model must be robust to changes in the mean and extreme values (e.g., sea levels or extreme temperatures). Regression approaches typically estimate ML-model parameters by minimizing distance between the prediction and target (e.g., mean square error). However, metrics can be biased towards the data distribution's mean, underestimating the tails (extremes). Instead, LEAP will investigate other loss functions; for instance, "quantile regression" can return general quantiles of a data distribution or conditional distribution. Furthermore, learnable metrics offer a solution to efficiently capture climate variable distributions; here LEAP will investigate loss functions learned using conditional Generative Adversarial Networks (GAN) [35].

Physical Constraints (**C.Zheng, Gentine**, **Kumar**). *Soft* constraints are physical constraints that need not be exactly satisfied, but are based upon prior physical knowledge (e.g., turbulence is thought to follow a diffusion process). Here, LEAP will investigate ML methods that encode constraint violations as a penalty and add them as regularization terms during optimization. This can be done for both intermediate and final outputs of the ML model. LEAP will also develop ML methods that enforce laws that must be *strictly* satisfied (e.g., mass conservation) through constraints on the ML model's architecture and outputs. To enforce strictly-satisfied hard constraints, constrained optimization programs will be directly embedded as operations inside neural networks during both learning and inference [36].

Spatiotemporal Process Constraints (**Bengio**, **McKinley**, **Vondrick**, **C.Zheng**). Physical processes are governed by multi-scale spatiotemporal dependencies. For instance, atmospheric convection in one model grid cell exhibits memory [37] and depends upon the neighboring cells [38]. LEAP will investigate ML representations trained end-to-end that respect spatiotemporal dependencies. For example, since deep networks are optimized with gradient descent, spatiotemporal constraints can be enforced and improve the ML model performance, provided they are sub-differentiable. For *temporal processes*, LEAP will investigate multiscale representations (e.g., using long short-term memory (LSTM)) that capture long-term temporal relationships and integrate both fast and slower dynamics. For *spatial processes*, LEAP will investigate spatial dependencies (e.g., with convolutional neural networks (CNN), such as U-Net [39]). Spatiotemporal dependencies can also be combined (e.g., using convolutional LSTM [40]).

B4.2. Moderate Data Richness: Hybrid Methods.

When some data are available, but additional knowledge (e.g., a causal graph) is required, LEAP will develop hybrid methods that are both data- and knowledge-driven.

Causal Discovery and Inference (**Bareinboim**, **Kingslake**, **Vondrick**). Since generalization failure is often due to learning the wrong generating process, LEAP will pioneer using causal reasoning in ESMs, making causal mechanisms explicit. Typical ML algorithms are based upon correlations between predicted and input variables, yet this relationship may not be causal. To improve generalization, ML models dealing with physical systems must correctly attribute predictions to a causal variable, and not to spurious confounding factors between correlated variables. Causality can improve algorithmic efficiency, requiring fewer data, and can be combined with meta-learning from small datasets [43] to improve the model's prediction. *Transportability* is a formal mathematical framework identifying the conditions where causal experimental information can be accurately transferred to different environments wherein only observational data are available [44,45]. LEAP will integrate transportability within ML methods to build generalized causal and counterfactual (i.e., avoiding spurious correlations) predictions from numerical or observational experiments, so that predictive models are robust to changing conditions and causal and counterfactual generalizations. This approach will identify additional data needs or improved theory/ model equations.

Equation Discovery (**Lipson, Zanna**). For relatively rich datasets that cover only a limited range of conditions and not the whole distribution (e.g., historical temperatures are lower than future ones), LEAP will use alternative ML approaches to improve extrapolation and interpretability. Using *equation discovery*, LEAP will leverage data to discover functional relationships for physical and biological processes. Unlike traditional linear and nonlinear regression methods that fit parameters to an equation of a given form, symbolic regression simultaneously searches both the parameters and the form of equations [31]. Using this approach, LEAP will either determine closed-form equations for the coarse-scale effects of high-resolution processes (e.g., turbulence or atmospheric convection) or improve upon existing equations for poorly-understood processes (e.g., ecosystem photosynthesis). LEAP will pioneer the application of equation discovery from Earth system data, and embed the resulting equations into **NCAR**'s CESM. When these algorithms are either equivalent to established ML (Section B4.1) or numerically more efficient, they will be chosen as they will likely extrapolate better. Equation discovery is an excellent alternative when established ML methods (e.g., deep learning) do not provide adequate insight; this will enhance process understanding.

Transfer Learning (**Kpotufe**, **Lipson**, **Pritchard**). When a problem has limited data but shares its intrinsic structure with a data-rich problem, LEAP will extend transfer learning to physical and biological ESM parameterizations. Transfer learning has proven successful for object recognition [41] and, recently, for El Niño prediction [42]. Multiple transfer levels are possible: In *near transfer*, the target task is closely related to the source task, such as predicting atmospheric convection from observations based upon transferring knowledge from high-resolution models. In *far transfer*, the target task is only structurally related to source tasks, such as transferring fluid dynamics problems governed by the Navier-Stokes equations from the ocean to the atmosphere. In such cases, only the underlying structure is transferred, not the problem-specific ML-model parameters. To identify what tasks to transfer between, and which transfer approach to use, LEAP will use domain knowledge on structural similarity between tasks, coupled with established notions of similarity between data spaces developed in transfer learning and related fields.

B4.3. Data Poor: Knowledge-driven Extrapolation.

When data are limited, they can be exploited more efficiently using domain knowledge about physical or biological processes, and LEAP will develop methods that rely more heavily upon such knowledge.

Physics-Guided Machine Learning (**Kumar**, **Gentine**). When data are limited, it is possible to incorporate physical knowledge encoded in physics-based/ mechanistic models (i.e., a standard parameterization), as pioneered by **Kumar** [46] and used by **Gentine** [47] for streamflow. Mechanistic models are built based upon mechanistic knowledge, and therefore contain rich information about the expected process' dynamics. Though these mechanistic models may be inaccurate due to incomplete knowledge of physics or biology, they can still be used to guide the learning process when data are limited. This can be done as a regularization of the ML algorithm or by first training the ML model on the mechanistic model and then refining it using only a small amount of observed data. This approach to correct structural errors holds potential to better generalize within out-of-sample scenarios, given its effective use of the knowledge inherent in the mechanistic model combined with observations for refinement [46,47].

B4.4. Cross-Cutting Methods.

Finally, LEAP will develop techniques that span the data-richness spectrum.

Non-Dimensionalization (**Giometto**, **Fish**). Dimensional analysis has a rich history in fluid mechanics to identify similarity in seemingly different experiments [48]. It provides a framework for dimensionality reduction based upon physical dimensions. LEAP will systematically use dimensional analysis, reducing a system's dimensionality by combining variables according to their physical dimensions (e.g., length, time, and velocity) to transform problems into non-dimensional ones, thus reducing system dimensionality. This then helps match seemingly different distributions, limiting the need for out-of-sample generalization by transforming the problem back into an interpolation one.

Uncertainty Quantification and Characterization (**Gelman**, **McKinley**, **T.Zheng**). ML algorithms under a Bayesian framework may use posterior draws to create uncertainty estimates. Rather than giving a single numerical estimate or prediction, LEAP will produce probabilistic estimates with uncertainties. Two challenges that can arise are: 1) statistical model misspecification; and 2) possible understatement of uncertainty from approximate inference algorithms. LEAP will deal with (1) by structured prediction scoring [49] via cross-validation that adjusts for biases and uncertainty due to model misspecification and randomness in probabilistic algorithms. LEAP will deal with (2) by using ideas of importance sampling [50,51] and expectation propagation [52] to improve upon the uncertainty estimates obtained by variational inference. LEAP will quantify how uncertainty varies across situations (e.g., sea surface temperature) using functional regression methods [50]. Such a regression framework allows quantifying the influence of physical factors on uncertainties. Spatial dependence and other structural constraints will also be integrated.

Bayesian Parameter Estimation (**de la Peña**, **Elsaesser**, **Gelman**, **Lall**, **Rush, van Lier-Walqui**). Today, ESMs' components are sequentially tuned based upon expert knowledge and first-order comparisons to data (e.g., adjusting cloud fields [24,53]). This is a major roadblock to accurate climate projections. Given the increase in model complexity, traditional tuning has reached its limit, motivating automatic and systematic approaches [54]. The prohibitive cost of running ESMs or their individual components allows only very few parameter-perturbed simulation runs, rendering direct parameter estimation impossible. The high dimensionality in both space and time and noisy observations is a further challenge.

LEAP will build upon recent work from **Elsaesser**'s use of neural networks [28] as a surrogate model to emulate sensitivity of the coupled **GISS'** ModelE output to perturbations in uncertain ESM parameters. This surrogate was used with a Markov Chain Monte Carlo (MCMC) sampler and relevant satellite datasets to estimate parameter posterior probabilities given observational uncertainties in a Bayesian framework. LEAP will apply and improve upon this approach in the coupled **NCAR** CESM and NASA **GISS** ModelE, accounting for spatiotemporal dependence (Section B4.1) and correlations between variables in a hierarchical Bayesian framework. These methods will improve model fidelity by reducing parameter uncertainty, but do not address structural uncertainties, which are addressed by LEAP's other strategies (e.g., equation discovery or established ML). By applying Bayesian parameter estimation to CESM components in its initial years, LEAP will develop initial benchmarks on model performance with the existing structure. The Center's structural improvements, combined with subsequent parameter estimation, can be subsequently compared to these benchmarks. Systematic Bayesian approaches for parameter estimation allow for frequently updating distributions as information becomes available.

B5. Earth System Component Research. LEAP's data-driven climate research combines and implements consistent strategies and methodologies across different fields, creating natural synergy and accelerating progress. The examples that follow illustrate LEAP's convergent strategy (Figure 5), though the Center will support additional riskier and novel projects throughout its lifespan, accomplished through educational programs (Section C) and a research seed funding program (Section B9).

The geoscientific projects outlined below are dedicated to improving CESM components possessing sufficient data, which are expected to *yield substantial improvement* in the overall CESM projections. Results will first be evaluated offline in individual CESM components, then in the coupled CESM facilitated by LEAP Fellows and postdocs' dedicated visits at **NCAR**. To understand and mitigate potential numerical instability of ML-based parameterizations [55], LEAP will systematically develop offline error growth diagnostics, following recent work by **Pritchard** and collaborators.

B5.1. Data Rich Fields.

In LEAP's data rich fields, equations are mostly known at fine resolutions (e.g., turbulence is controlled by Navier-Stokes equations) but not at the coarse scale due to the truncation of small, subgrid scales. These projects will leverage both high-resolution simulations and observations. First, high-resolution simulations resolving the specific process will be run across a diverse range of conditions – this wide range of generated conditions ensures that **ML algorithms will mostly interpolate** and not extrapolate. Simulations will then be coarse-grained to the ESM's resolution (Figure 4). An ML algorithm will be trained to reproduce coarse-scale fluxes (e.g., Reynolds' stress, heat and moisture fluxes) or the variable of interest (e.g., cloud cover) with coarse-scale state variables as inputs (e.g., wind, temperature, total specific humidity, ice-sheet thickness). Throughout, algorithms will include soft (prior physical knowledge), hard (e.g., mass and energy conservations), and spatiotemporal constraints, as well as non-dimensionalization. This research category will focus upon established **ML interpolation methods** (e.g., deep or convolutional neural networks, Section B4.1). When data are still too limited or the ML model too opaque, LEAP will additionally combine those ML methods with **causal inference** or use **equation discovery methods** [22] (Section B4.2) to translate data-driven models into physical insights and to better extrapolate. Since resulting algorithms are trained upon high-resolution simulations that still exhibit biases against observations, they will then be refined and adjusted with sparser observational data using **transfer learning**. This section builds upon previous proofs of concept by **Gentine**, **Pritchard**, and **Zanna**.

Atmospheric Turbulence (**Gentine**, **Giometto**, **Kpotufe, Lipson, Vondrick**, **Zanna**). Due to the vast range of scales spanning atmospheric flows, underlying atmospheric equations cannot be entirely resolved. Therefore, an ESM parameterization of sub-grid turbulence is required, which impacts low clouds [56], critical to climate response [13]. Direct Numerical Simulations (DNS) and Large-Eddy simulations (LES), which resolve different scales of turbulence, will be used to inform the development of sub-grid-scale (SGS) parameterization of atmospheric turbulence using ML. These simulations will cover a wide range of atmospheric conditions, including very stable to unstable boundary layers, different shear, and buoyancy Reynolds numbers. **Gentine** and **Giometto** have already run many multi-million CPU-hour DNS and LES simulations covering various conditions used for training, which will be complemented by new ones covering more conditions. Then, LES subgrid turbulence will be refined by coarse-graining DNS simulations and learning the fluxes as a function of state with established ML (Section B4.1). This new LES will in turn define an ML representation of turbulence in the CESM, impacting multiple turbulence scales.

Ocean Turbulence (**Abernathey**, **Kpotufe**, **Lipson**, **Vondrick**, **Zanna**). For ocean turbulence, LEAP will employ a strategy similar to atmospheric turbulence, and the two groups will closely collaborate, naturally following **Zanna**'s pioneering work [22]. Using coarse-grained, high-resolution ocean simulations, LEAP will obtain the true sub-grid fluxes at the scale of the ESM ocean model, targeting both subgrid horizontal and vertical tracer and momentum transport, key processes in setting heat and carbon uptakes [57]. Simulated datasets to feed development of ML subgrid ocean transport parameterizations will be generated using idealized Modular Ocean Model 6 (MOM6) configurations, the future ocean model for **NCAR** and NOAA's Geophysical Fluid Dynamics Laboratory, and targeted LES turbulent-resolving simulations (described above). MOM6 simulations will be either meso- or submeso-scale resolving, using different forcing, dissipation, topography, etc. LEAP will run simulations across a range of Reynolds, Richardson, and Rossby numbers to span the phase space. The idealized datasets will be complemented by drawing from existing and planned mesoscale-resolution global simulations, including ocean-only and coupled runs. The highest-resolution dataset will come from a global 1/48-degree ocean-only simulation. Coupled runs will be drawn from existing CESM simulations with 1/10-degree POP ocean components [98]. Mesoscale-resolving satellite products (e.g., sea-surface temperature and height) will further refine ML SGS parameterizations using transfer learning. Before being tested in the CESM configuration, the ocean turbulence model will be initially implemented into the idealized MOM6-based configurations.

Moist Convection (**Elsaesser**, **Gentine**, **Pritchard**, **van Lier-Walqui**). ESMs exhibit large biases in their intensity, modes of variability, and diurnal cycles of convection [58,59], mostly alleviated in Cloud Resolving Models (CRMs) (at ~1km resolution) that explicitly represent deep moist convection [27]. LEAP will use global CRM simulations to develop a refined ML-based parameterization of moist convection (deep clouds), expanding from work by **Gentine** and **Pritchard**. LEAP will use a super-parameterized approach, wherein 2D (1D horizontal and vertical) CRMs are embedded within a coarse-scale climate model [18,60,61]. These simulations will efficiently span diverse conditions (e.g., sea surface temperatures). The super-parameterization will provide a baseline to derive the initial ML algorithm. This ML algorithm will be refined with transfer learning using: 1) fully 3D (much more computationally expensive) CRM simulations [62] that can only be run for a few days but have full spatial dependency; and 2) observational data (e.g., precipitation). By incorporating LEAP's Bayesian approach to microphysics (Section B5.2), the convection parameterization will avoid the limits of the incomplete knowledge of cloud microphysical processes.

Ice-Sheet Flow (**Bell**, **Kingslake**, **Lipson**, **Raymo**, **C.Zheng**). Learning from the turbulence and convection work, in the CESM's Community Ice Sheet Model (CISM), LEAP will replace parameterizations of key processes (e.g., basal sliding, fracturing, subglacial hydrology) with ML algorithms trained upon output from high-resolution physics-based models [63,64]. Basal sliding is controlled by skin drag and flow around bumps in the bed, called "form drag" [82]. Basal topography varies over smaller scales than ISMs can resolve [65], so they lump together sub-resolution form drag with skin drag to parameterize and invert for basal shear stress, τ, from observations [66]. However, τ may become inaccurate as ice accelerates, since parameterizations are simplified [67] and form drag varies spatially [68]. Inspired by **Gentine** et al.'s approach to parameterizing convection [25,64], and using established ML (Section B4.1) and, later, equation discovery (Section B4.2) methods, LEAP will pioneer the first ML model for τ by training a neural network on output from high-resolution, 3D physics-based flow simulations. This approach will provide estimates of τ as ice flow accelerates, improving sea-level predictions, and will be extended to grounding line flux and supraglacial and subglacial hydrology.

B5.2. Moderate Data Availability.

In moderate data fields, a reference model or set of equations at the fine scale is unknown, so high-resolution models cannot be run to refine process understanding and span more conditions. Observations do exist, but only for a subset of conditions and across short timeframes. Here, LEAP will use causal inference, equation discovery, transfer learning, and parameter estimation (Section B4.4) to refine ESM parameterizations. This approach builds upon work by **Zanna, Morrison, van Lier-Walqui**, and **Gentine**.

Microphysics (**Gettelman**, **Lipson**, **Morrison**, **Pincus**, **van Lier-Walqui**). Cloud microphysics (microscale cloud and precipitation processes) is a key source of uncertainty for global temperature projections and its recent ESM spread [69], partly because the complete governing equations are unknown. However, observational sources (e.g., weather radars, *in situ* particle measurements, and laboratory experiments) provide rich information for ML-based inference. For microphysics, LEAP will utilize a hierarchy of microphysical parameterization schemes: At the beginning of the hierarchy, a high-accuracy but computationally expensive Lagrangian scheme will provide a reference model. At the end of the hierarchy, a computationally efficient bulk scheme will quantify structural and parametric uncertainty across a range of conditions [70,71]. Recent work in microphysics employed power-series representations for microphysical process rates [70,71]. This will be extended to *equation discovery*. The detailed Lagrangian scheme will additionally support bulk microphysics schemes by providing a reference, free from particle size-distribution approximations and with small numerical errors.

Biological Pump (**Gomes**, **Long**, **McKinley**, **van Lier-Walqui**). Biology is critical to ocean carbon uptake and storage [87], yet biological carbon transport to the deep ocean is difficult to predict because it is the product of complex ecosystem functioning with unknown governing equations. Approximate equations are used in state-of-the-art biological pump models [88], requiring numerous tracers to resolve the many phytoplankton and detritus sizes and interactions of sinking particles. LEAP will run these models over limited spatial domains with variable environmental conditions to develop an ML training dataset for an ML-based parameterization of biological pump function. Further refinement will be made using biological pump data from the 2018-2020 NASA EXPORTS field program [89] *using parameter estimation*.

Photosynthesis and Respiration (**DeFries**, **Gentine**, **Menge**, **Uriarte**, **Weng, T.Zheng**). One third of annual CO_2 emissions are absorbed by the terrestrial biosphere [72]. ESM projections of terrestrial carbon uptake through 2100 diverge greatly: From a strongly growing sink (9 PgC yr^{-1} storage) to a strong additional carbon source (6 PgC yr^{-1} release) [73]. Leaf photosynthesis represents the largest flux of carbon (photosynthesis) between the land and atmosphere [74]. Detailed observations at a few field sites provide abundant observations of these fast processes. Land model formulations for plant carbon and water flux resistance, and their dependence on environmental conditions are empirical [75–77], yet can be estimated using ML informed by field data, as shown by **Gentine** [78]. To improve generalization, *causal inference* will be used to better attribute causing variables, limiting spurious correlations. LEAP will apply these new surface resistance formulations in the CESM Community Land Model (CLM). To improve generalization to elevated CO_2, LEAP will use: 1) *equation discovery* to define canopy resistances' analytical structure; or 2) *transfer learning*, using Free-Air Carbon dioxide Enrichment experiments (FACE) data [79]. Though less understood, ecosystem respiration is also critical to net terrestrial carbon uptake. To better model respiration, LEAP will evaluate two strategies: 1) parameter estimation with uncertainty quantification for the temperature sensitivity of respiration, using *in situ* soil data [80,81] and ecosystem respiration data from eddy-covariance sites [82]); combined with 2) an equation discovery approach for the reference respiration rate as a function of environmental drivers (e.g., soil moisture, temperature).

Ice-Sheet Surfaces (**Bell**, **Kingslake**, **Lipson, Raymo**). LEAP will exploit new polar remote sensing datasets to test and implement CISM sub-grid parameterizations. For example, brittle fracture controls ice-shelf collapse [83] and iceberg calving, yet its ISM representations are highly-simplified [84,85]. First, LEAP will map surface fractures using a deep convolutional neural network with new high-resolution (<1m) Digital Globe imagery. Then, ice shelf stresses, computed from observations, will be compared to predictions of fracture parameterizations based upon traditional physics-only approaches (e.g., linear elastic fracture mechanics) or a ML-model derived from this new dataset [86]. This will be the first time any fracture parameterizations have been tested against comprehensive, ice sheet-wide fracture data. Parameterizations that perform well will be implemented in the CISM, ultimately leading to improved sea-level projections. Other imagery-ML applications will include modeling surface meltwater, quantifying melt volumes, and interpreting englacial structures from ice-penetrating radar (**NASA**'s Operation IceBridge, and NSF and **Columbia**'s ROSETTA/ Ice project). With **Minnesota**'s NSF-funded Polar Geospatial Center, LEAP will lead international efforts to unlock the potential of these critical polar datasets with ML.

B5.3. Limited Data Fields.

When globally- or regionally-gridded products are unavailable to parameterize or evaluate the ESM, LEAP will: 1) employ a physics-guided ML, guided by a mechanistic model; 2) develop new gridded products using extrapolation from *in situ* sparse data; these products are required to refine model parameterizations because comprehensive global observations are lacking. These datasets will complement or replace existing datasets in the ESM evaluation set (Section B6). Compared to existing products, these will include uncertainty quantification, allowing robust ESM parameterization uncertainty estimates. This builds upon **T.Zheng** and **McKinley**'s recent work.

Hydrology and Snow (**Gentine**, **Hamman**, **Kumar**, **Lawrence**, **Weng**). ML with memory captures flooding and drought events better than the ESMs' current hydrological models [47,90]. LEAP will replace the CLM's existing streamflow routing schemes with a physics-guided ML model (Section B4.3), pioneered by **Kumar** and used by **Gentine**, as a dynamic bias correction (e.g., using Long-Short-Term Memory) trained upon observed streamflow data. Snow cover and snow water equivalent exhibit biases [91] due to imperfect process representation. However, SNOTEL *in situ* observations and remote sensing data provide observational estimates that will correct the CLM, similarly to streamflow, using ML as a dynamic (state and time dependent) post-processing of the snow module.

Air-Sea Carbon Fluxes (**McKinley**, **T.Zheng**). Due to the rise of atmospheric CO_2, oceans have accumulated one third of all fossil CO_2 emissions [92]. Increasing ESM projection confidence for future ocean carbon uptake requires high-quality data products. The critical data are *in situ* pCO_2, but these cover only 2% of the ocean [93]. Various off-the-shelf ML methods, including neural networks, have been used to extrapolate into full-coverage data products using satellite-observed temperature, chlorophyll, and other variables [94]. However, the air-sea carbon flux estimates from these pCO_2 products diverge. Along with insufficient uncertainty quantification, this limits the data products' utility for ESM evaluation [95]. **McKinley** and **T.Zheng** are developing next-generation ML techniques using neural networks, random forests, and extreme gradient boosting, and also approaches for uncertainty quantification [29]. LEAP will build upon their initial efforts to develop superior data products with spatially-resolved uncertainty.

Soil Carbon (**Menge**, **Weng**, **T.Zheng**). Soil stores three times more carbon than vegetation [96], and this carbon release would dramatically enhance warming. LEAP will extrapolate site-level carbon data to global scales, learning from ocean carbon fluxes' ML approaches. For carbon content and decomposition, latent variables (vegetation type, soil texture/ depth) can reduce dimensionality, further improving extrapolation. This data product with uncertainty quantification (Section B4.4) will then support CLM parameter optimization (e.g., residence time, temperature sensitivity).

B6. ESM Integration (**Elsaesser**, **Fish**, **Gagne**, **Gettleman**, **Lamarque**, **Lawrence**, **Long**). LEAP's ultimate objective is the systematic development and integration of novel ML algorithms into the coupled

CESM, refined with physics and observations. Developing new algorithms that generalize by combining domain knowledge and constraints with ML (Section B4.3), thereafter integrated within CESM components, will substantially reduce structural and parameter errors and improve multi-decadal climate projections. One significant advantage of integrating ML algorithms with the CESM is that it leverages mostly unused hardware (GPUs) for acceleration. For the same cost, more simulations can be run, thus better characterizing internal climate variability uncertainties (through the spread across ensemble members). Interactions between different ESM components (i.e., "coupling"), which only arise in the full ESM context, are a research area ripe for ML intervention. Both before and after coupling of individual components, the CESM will be rigorously evaluated against skill metrics. Implementation into the CESM will be supported by NCAR and visiting LEAP Fellows and postdocs**.**

B6.1. Coupling (**Elsaesser**, **Giometto**, **Kaiser**, **Kim**). Coupling ESM components is a labor-intensive step in model development, wherein individual ESM components, first evaluated and calibrated individually, are manually re-tuned post-coupling. The coupled model must be recalibrated against observations because of inherent component biases and error growth. Two main error sources impact ESM coupling: 1) structural errors in model subcomponents; and 2) parameter uncertainty in these subcomponents. By substantially reducing structural errors in CESM subcomponents (e.g., atmosphere, ocean), the coupling process will be simplified due to *reduced biases* and by having *fewer key parameters* to tune. A systematic use of ML and causal techniques (which limit spurious correlations) (Section B4.2) to estimate covarying parameter probability distributions will further reduce coupled ESM biases compared to observations. For LEAP-developed parameterizations and traditional parameterizations, subcomponents will be developed and tuned individually before linking and re-tuning in the coupled model, thus leveraging the uncoupled components' lower computational costs. The coupling stage will use the uncoupled tuning as *a priori* estimates of subcomponent uncertainty, further refined by coupled experiments informed by analysis of causally-correlated subcomponents. For coupled Bayesian parameter estimation, instead of considering the full variation of parameters, LEAP will introduce discrete or continuous valued latent variables to cluster variables (e.g., cloud type, plant functional types, soil types) and reduce dimensionality. Considering a multi-level Bayesian modeling structure [97], each ML model will be introduced and tested for each major conditioning step, then integrated across the modeling chain. Uncertainty propagation across building blocks on each model chain level will be accomplished in the Bayesian conditioning framework.

B6.2. Standard Metrics for Model Improvements. LEAP will employ a variety of metrics to continually assess the new ML-based CESM's skill and potential for improvement. By definition, data exists only from the *historical* climate, though it is critical to constrain the *future* climate. LEAP will follow established diagnostics for model-data comparison, while also developing new data products and ML-based metrics. Though contemporary ESMs use many metrics, current validation metrics do not adequately constrain the future [69], partly because compensating errors across ESM subcomponents lead to incorrect model parameter calibration. To address this, LEAP will improve model diagnostics by developing new metrics that take advantage of ML and of the available wealth of data (e.g., daily cloud cover). LEAP will run high-resolution simulations under future conditions (e.g., higher sea surface temperatures) to test algorithmic abilities to extrapolate to new conditions.

Standard Metrics. In Years 1-3, LEAP will employ standard metrics similar to those currently used at **NCAR** and **GISS**. These data are mostly based upon remote sensing or compilations of historical weather products, called "reanalysis." Integrated evaluation tools already exist within the ESM ValTool [98], and will focus upon climatologies of global outputs (e.g., top of atmosphere radiative fluxes, precipitation, and water vapor), climate variability with the **NCAR** Climate Variability Diagnostic Package (CVDP) [99] (e.g., El Niño Southern Oscillations), and teleconnections to ENSO, Madden Julian Oscillation, Quasi Biennial Oscillation, and Pacific Decadal Oscillation. The ESM ValTool also allows measuring common ESM biases (e.g., coupled tropical climate variability, monsoons, Southern Ocean processes, or continental dry biases). For the CLM, LEAP will use the recent ILAMB [100] model evaluation toolbox that compares model performance across a suite of metrics to *in situ* data, satellite products, or carbon, energy, and water fluxes gridded products based upon an ML upscaling from *in situ* data, as used in CLM5.

B6.3. Enriched Metrics for Model Improvements (**Abernathey**, **Camargo**, **Lamarque**, **Vondrick**). In Years 3-5, LEAP will leverage new ocean and soil carbon flux products, including anticipated USMILE products (e.g., carbon and water fluxes, biomass and carbon stocks, and soil moisture). For oceans, regional patterns of air-sea carbon fluxes and temporal variability with robust uncertainty bounds will be ready for application (Section B5.3). LEAP will refine ESM evaluation metrics: Current metrics primarily focus upon the mean or variance (e.g., seasonal cycle, ENSO), and do not fully quantify the distribution of the process; therefore, *existing observations go underutilized*. Better characterizing process distributions and temporal and spatial spectra can transform model performance metrics, as illustrated by **Abernathey** [101]. However, even spectral analyses metrics cannot represent the full subtlety of variability; for instance, small-scale clustering or aggregation [102] is largely not "seen" by power spectra.

LEAP will develop novel ML-based metrics to address model evaluation. Several important benchmarks use ML to analyze outputs and decipher ESM information (e.g., regions of the globe most indicative of climate change [103]). LEAP will systematically learn the right metric for comparing observations and ESM outputs. This challenge can be framed in the context of problems in computer vision, wherein it is difficult to define distance functions between two images or dimensionally heterogeneous variables (e.g., temperature, humidity). Recently, collaborators at **Columbia** showed that deep learning is a powerful tool for comparing complex cosmological images [102]. LEAP will improve upon this approach using generative models aimed at learning the distribution between source and target and how to distinguish them. This latter discriminator can optimally define the metric and accurately compare models to observations, better reproducing the distribution instead of prescribing *ad hoc* metrics targeting the mean (e.g., mean square errors). For fast processes (e.g., clouds), this strategy will better leverage existing data, especially detailed spatial and temporal information from remote sensing data.

B7. Research Roles and Support. Across its five years, LEAP's projects will be supported by approximately 40 graduate fellows. Postdoctoral research scholars will hold either three year (Years 1-3) or two year (Years 4-5) appointments, so that the five-year award will support approximately 16 scholars.

For *Data Rich* fields, data scientists **Bengio**, **Kumar**, **Kpotufe, Lipson**, **C.Zheng**, and **Vondrick** will work with 0.5 graduate students per year. Geoscientists **Abernathey**, **Bell**, **Elsaesser**, **Gentine**, **Giometto, Kingslake**, **Pritchard, Raymo, van Lier-Walqui,** and **Zanna** will work with three graduate students and one postdoc per year. For fields with *Moderate Data* availability, LEAP will develop equation discovery and transfer learning, and apply these to ESM problems. Data scientists **Bareinboim, de la Peña**, **Kpotufe**, **Gelman**, **Lall**, **Lipson**, **Rush**, **Vondrick** and **T.Zheng** will work with two graduate students per year. Geoscientists **Bell**, **DeFries**, **Elsaesser**, **Gettleman**, **Gomes**, **Kingslake**, **Long**, **McKinley, Menge**, **Morrison**, **Pritchard, Pincus**, **Raymo**, **Uriate**, **van Lier-Walqui** (25% time), and **Weng** (25% time) will work with one graduate student and two postdocs per year. For fields with *Limited Data*, LEAP will focus upon knowledge-driven extrapolation. Data scientists **Kumar** and **T.Zheng** will work with one graduate student per year. Geoscientists **Gentine, Hamman**, **Lawrence, McKinley**, **Menge**, and **Weng** will work with 1.5 graduate student for Years 1-3, and 1.5 postdocs per year. Uncertainty quantification, led by **de la Peña**, **Lall**, **Rush**, and **van Lier-Walqui** will work with one graduate student per year. Parameter estimation will be led by **de la Peña**, **Elsaesser, Gelman, Lall**, **Rush,** and **van Lier-Walqui** (25% time). Enriched metrics will be developed by **Abernathey**, **Camargo**, **Lamarque,** and **Vondrick** with one postdoc per year. Coupling will become a greater focus in Years 3-5, with **Elsaesser, Fish**, **Gagne, Giometto**, **Gettleman, Kaiser, Kim, Lamarque**, **Lawrence,** and **Long** assisted by two postdocs per year and one graduate student in Years 4-5.

B8. Research Timeline. In Years 1-2, LEAP's primary approach will be to develop and implement new ESM parameterizations in stand-alone, uncoupled models. In Years 3-4, LEAP will separately couple new parameterizations into the CESM. In Year 5, LEAP will couple all components *together* into the CESM. Data products will be developed primarily in Years 1-3. Evaluation against metrics will occur continuously so as to assess progress from an initial baseline of CESM *without* LEAP to CESM *with the continuous deployment of* LEAP's ML-enhanced parameterizations.

B9. Research Seed Funding. Each September, the Convergence Subcommittee (Section F6) will release a Request for Applications for research seed funding to support collaborations between one faculty sponsor and one graduate student; though the faculty sponsor must be Senior Personnel on the project, the student can be from any of LEAP's participating academic institutions. Transdisciplinary collaborations will be strongly encouraged. Based upon LEAP's Broadening Participation (Section D) commitments, applications will be fully anonymized. Teams will propose 6- to 12-month collaborative projects, culminating in a final research product published open-source to the Center's website. The Convergence Subcommittee will evaluate applications for intellectual merit and broader impact, and applicants' self-reported attendance of DEI workshops (Section D2).

C. DESCRIPTION OF THE EDUCATION AND HUMAN RESOURCE DEVELOPMENT OBJECTIVES OF THE CENTER.

(**T.Zheng**, **Blikstein**, **Cogburn**, **DeFries**, **Kim**, **Lang**, **Pizmony-Levy**, **Purdie-Greenaway**, **Uriarte**).

LEAP's educational and human resource development programs (Figure 6) establish convergence to achieve Center objectives by: 1) collaboratively defining and advancing climate data science; 2) integrating research, education, broadening participation, and knowledge transfer initiatives; and 3) partnering with academic, federal, corporate, NGO, and local government stakeholders towards shared objectives. LEAP's trainees will unify all Center initiatives, guided by and formatively defining **three educational objectives**:

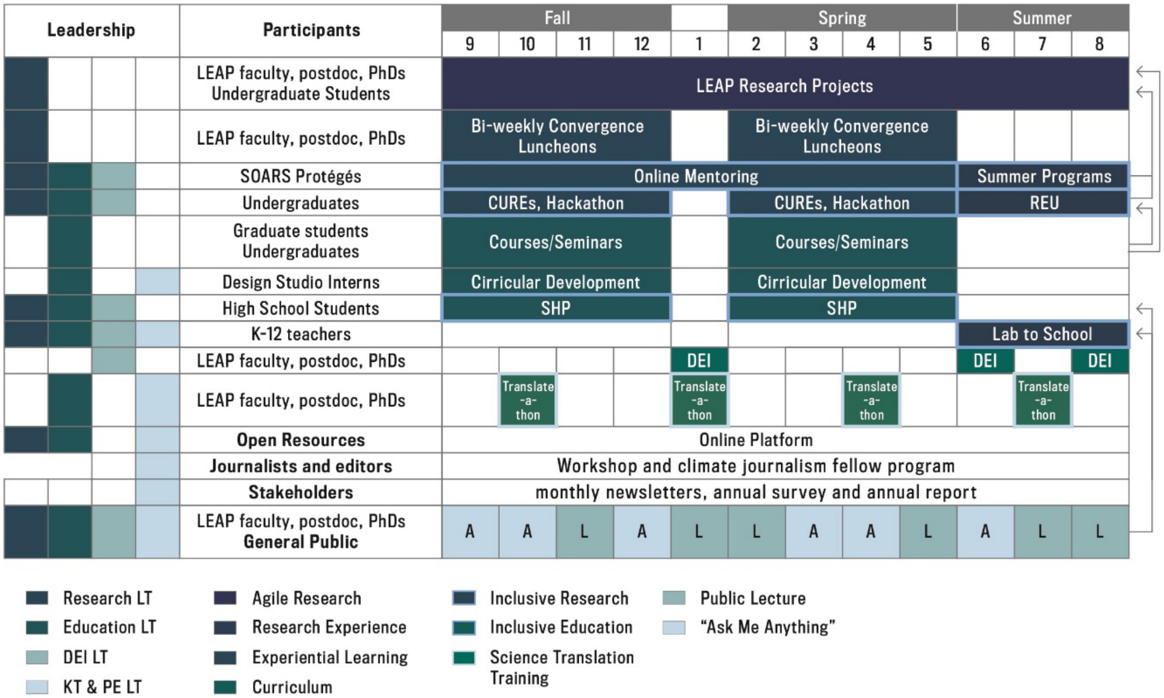
- **EO1:** Develop research-intensive undergraduate and graduate curricula, providing multiple transdisciplinary entryways for students from diverse backgrounds and with varied career goals;
- **EO2:** Offer immersive research experiences to student, teacher, parent, and workforce trainees;
- **EO3:** Broaden research impact through communication, translation, human-centered design, curriculum development, and evaluation and assessment.

EO1: New Entry Points into Climate Data Science.

C1. Certificate Program. LEAP will launch a **Certificate of Professional Achievement in Climate Data Science** immediately available to all partner academic institutions, plus all New York City doctoral students through the existing NYC Inter-University Doctoral Consortium; all LEAP Fellows (Section C2) will complete this Certificate, consisting of four graduate courses that pave pathways from literacy and core competency to advanced research skills: 1) one of three existing cross-disciplinary electives: **McKinley**'s *Humans & the Carbon Cycle*, **T.Zheng**'s *Statistical Machine Learning*, or **Gentine**'s *Machine Learning for Environmental Engineers and Scientists*; 2) a new transdisciplinary course, *Climate Projection Challenges*; 3) an introductory research seminar (Section C4); and 4) an advanced research seminar (Section C4). In Year 1, LEAP will acquire approval from the Provost, formally launching the Certificate in Year 2. For its first two years, the Certificate's courses will be taught only at **Columbia**; subsequently, and sustainably after STC funding expires, LEAP will translate courses into fully-online options, facilitated by the award-winning **Columbia** Video Network portal, thus allowing global engagement. The Certificate will be marketed by **Columbia** Business School's executive education program to recruit corporate trainees; full tuition remission will be provided to Bridge post-baccalaureate students (Section D5), graduate students affiliated with the **National Society of Black Engineers**, and each semester's Executives (Section E3) and Storytellers (Section E7) in Residence. LEAP budgets financial aid for up to five underrepresented minority graduate students per year to take *Challenges* at no cost; this opportunity will be marketed across each of LEAP's academic institutions through local chapters of the **National Society of Black Engineers**.

C2. Doctoral Fellowship. Each year, between six and eleven graduate students will be placed across LEAP's research projects as "**LEAP Fellows**" selected by the Convergence Subcommittee (Section F6). Fellows will receive full tuition, stipend, and benefits support for 12 months, with appointments renewable depending upon satisfactory progress and advisor endorsement. Students completing the Certificate may apply for Fellowships after having completed *Challenges*. Complementing research, LEAP Fellows will attend formal and informal training and professional development forums guided by the student-run Climate Justice Leadership Board (Section F8); each such forum and opportunity is defined below.

C3. New Course. Modeled after **T.Zheng**'s existing project-based learning course, *Climate Prediction* "*Challenges*" will offer Course-based Undergraduate Research Experiences (CUREs) featuring evidence-supported, scalable, and accessible pedagogy [104]. Enabled by peer-learning [105], common task frameworks [106], and discussion-based learning [107], *Challenges* will advance student competencies in critical thinking, research ethics, communication, and team science. Through sequential team projects, *Challenges* will guide students through ML workflows to address Earth system problems. Because real-world research is not naturally classroom-ready, each project will be explicitly mapped and scoped against the course's educational goals in the Design Studio (Section C8), along with assessment rubrics. This course will be accessible for students with heterogeneous workforce development goals by: 1) serving as a gateway for research students to enter LEAP's two seminars; and 2) offering advanced research experiences to students from the social sciences and humanities disciplines, and the corporate, non-profit, and policy fields, whose current climate data science training opportunities are comparatively data science-light.



Leadership				Participants	Fall					Spring				Summer		
					9	10	11	12	1	2	3	4	5	6	7	8
■				LEAP faculty, postdoc, PhDs Undergraduate Students	LEAP Research Projects											
■				LEAP faculty, postdoc, PhDs	Bi-weekly Convergence Luncheons					Bi-weekly Convergence Luncheons						
■	■	■		SOARS Protégés	Online Mentoring									Summer Programs		
■	■	■		Undergraduates	CUREs, Hackathon					CUREs, Hackathon				REU		
	■			Graduate students Undergraduates	Courses/Seminars					Courses/Seminars						
	■		■	Design Studio Interns	Cirricular Development					Cirricular Development						
■	■	■		High School Students	SHP					SHP						
■	■	■	■	K-12 teachers										Lab to School		
		■		LEAP faculty, postdoc, PhDs					DEI					DEI		DEI
	■		■	LEAP faculty, postdoc, PhDs		Translate-a-thon			Translate-a-thon			Translate-a-thon			Translate-a-thon	
■	■		■	**Open Resources**	Online Platform											
			■	**Journalists and editors**	Workshop and climate journalism fellow program											
			■	**Stakeholders**	monthly newsletters, annual survey and annual report											
■	■	■	■	LEAP faculty, postdoc, PhDs **General Public**	A	A	L	A	L	L	A	A	L	A	L	L

- Research LT
- Education LT
- DEI LT
- KT & PE LT
- Agile Research
- Research Experience
- Experiential Learning
- Curriculum
- Inclusive Research
- Inclusive Education
- Science Translation Training
- Public Lecture
- "Ask Me Anything"

Figure 6: LEAP's timeline of convergent education and human resources development programs.

C4. Two Convergent Seminars. LEAP will launch a doctoral research seminar series co-taught by **Gentine**, **McKinley**, **Abernathey**, **Vondrick**, and **T.Zheng** to: 1) expose geoscience students to cutting-edge data science methodologies; 2) expose data science students to open climate and ESM parameterization problems; and 3) collaboratively engage students from both disciplines in project-based learning, emphasizing reproducibility and ethics.

The Fall seminar will be a "journal club" intersecting ML and climate science. ML's rapid growth means that no standard textbooks or curriculum yet exists for such a course; rather, instructors and students will collaboratively determine the semester's reading list. The Spring seminar will immerse student teams in large, data-intensive research projects executing modern ML on large geoscientific datasets, made accessible in an ML-friendly format through LEAPangeo. Projects will be guided by LEAP's overall research strategy (Figure 5). Students will finish their semesters with a publication-grade project including a "reproducibility pack" of code and data, and an ethics audit for fairness, transparency, and accountability. As with all LEAP products, these projects will be posted open-source to the Center's website.

EO2: Research Experiences for Multiple Trainee Groups.

LEAP will embed four trainee categories within each of its research, broadening participation, and knowledge transfer initiatives in order to establish convergence and sustain engagement between:

1. **Postdocs:** PhD-level scholars will hold two- to three-year appointments at **Columbia** and **Irvine**, and frequently visit **GISS**. Each semester, two postdocs will visit **NCAR** to implement modules into the CESM. Communications technologies (Slack, Zoom, Asana) will sustain distance collaboration.
2. **Graduate Students:** PhD students at **Columbia**, **NYU**, and **Minnesota** will participate in Research Seminars and develop dissertation proposals jointly-supervised by LEAP faculty. Each summer, three students will visit **NCAR** to pursue implementation of CESM parameterizations.
3. **Undergraduate Students:** LEAP will offer summer research experiences to undergraduates via **Columbia**'s Summer@SEAS. Each December, undergraduate and faculty teams will propose REU projects to the Convergence Subcommittee; all undergraduates having completed introductory geoscience and/ or data science courses are eligible. Proposals will be required to outline the project's research

goals, its relevance to LEAP's research strategy, the scope of work, learning opportunities, mentoring plan, relevance to broadening participation commitments, and project management schedule.

4. **High School Students:** Beginning in Year 1, LEAP will integrate a new experiential climate data science course into the **Columbia** Science Honors Program (SHP), a high school STEM enrichment program on Saturdays. Course materials will be developed in the Design Studio. LEAP graduate and undergraduate students will lead this new SHP, thus providing trainees with teaching experiences.

C5. "Train the Trainer" Faculty Workshops. Each year, LEAP will engage interested faculty from all its academic institutions in a two-day workshop. Using LEAPangeo-based datasets, on Day 1 a mini-hackathon will assess LEAP's physics- and biology-guided ML algorithms; on Day 2, small teams and Design Studio (Section C8) interns will design a teaching module that includes climate projection questions, two or more challenging data sources, potential data science tools, and evaluation rubrics. These modules will introduce climate data science to students across both STEM and non-STEM courses. Workshop materials and products will be posted open-source on the LEAP website, and disseminated through **Teachers College**'s network of education colleges.

C6. "Lab-to-School" Summer Institute and Parent Training. LEAP will engage New York City high schools through a **Summer Institute on Climate Data Science**, spearheaded by **Teachers College**'s Center for Sustainable Futures, which facilitates research on educational practices and translation of research into action through dissemination to educators, parents, and the public. The Center forges LEAP's research-practice partnership with the **New York City Department of Education** to galvanize school support towards the City's climate goals [108]. Since 2009, New York City requires that all public schools appoint a Sustainability Coordinator to plan and implement energy and waste mitigation and communication initiatives, revealing a special opportunity for LEAP to achieve broad influence and impact: At no cost and open to all Sustainability Coordinators, the **Summer Institute** will be a five-day professional development workshop introducing teachers to climate data science for prototyping teaching units aligned with New York State science standards. All teaching units and related materials will be posted online and disseminated through a webinar series; these materials will each be translated into Spanish, broadening potential participation from historically underserved communities. **Pizmony-Levy** will develop case studies allowing educational scholars to reflect upon and learn from LEAP's educational experiments with novel audiences.

Open to all parents of K-12 students and free of charge, a subsequent **Parents' Conference** will introduce participants to data-driven climate research and **New York City Department of Education** climate policy and programs. The Conference will feature presentations by Summer Institute graduates and LEAP's faculty, postdoc, graduate student, policy, press, and corporate communities.

C7. Science Museum Partnership. LEAP will partner with the **American Museum of Natural History** (AMNH) to equitably support students in two ways: 1) BridgeUP, supporting data-driven natural science through scholarships, fellowships, and hackathons for high school girls and post-baccalaureate women; and 2) Urban Advantage (UA), a partnership between the **New York City Department of Education** and informal urban science institutions bringing investigation-based science into classrooms [109], annually reaching 93,000 eighth graders (90% self-reporting as black or Hispanic). LEAPangeo products will provide accessible and relevant resources for UA teachers and students to investigate climate data science. Data products and ML codes will support projects for advanced BridgeUP participants. Project outcomes will be incorporated into public-facing, interactive exhibition elements in **AMNH**'s Hall of Planet Earth.

EO3: Education and Human Resource Development Evaluation and Dissemination.

C8. Design Studio. Chief Convergence Officer **T.Zheng** and evaluator **Lang** will launch a Design Studio training graduate interns to develop case studies and formal curricula. Receiving course credit, interns will learn human-centered design, acquire experience with raw geoscience datasets, and develop metacognitive skills that advance their climate data science skills. In Years 2-5, the Design Studio will update existing curricula and develop new case studies based upon LEAP's Knowledge-Data Continuum, bolstering knowledge transfer. All curricula will be shared open-access, summarized, and highlighted in the monthly newsletter, annual report (Section E12), and biennial **Microsoft** AI4Earth conference (Section E4).

C9. Evaluation Plan. Evaluator **Lang** will systematically assess LEAP's non-research programs, building upon scholarship in data science and interdisciplinary education outcomes [110]. Specifically, project-based exploration and discovery evaluation will be guided by the Association of American Colleges and Universities' rubrics on quantitative reasoning, problem solving, and teamwork [111]. This will grow a library of assessment tools and rubrics, as shared resources for the full LEAP community and beyond.

Type	Evaluation Methods for Student Outcomes
Formative Assessment	Feedback during the project on problem setup, methodology choice and the design of validation studies; Observational study of student questions when approaching a new Climate Data Science topic before and after the CUREs experience; peer-reviews of team collaboration.
Short-term Summative	Evaluation of the final product with designed grading rubrics on creating scientific value, presentation, reproducibility, and ethics. An ethics audit will be required as part of the final report.
Long-term Summative	Longitudinal follow-up, observational study on students' willingness to further climate data science study, willingness to apply for internships, participation in the Design Studio, participation in LEAP research.

In addition to the above assessment, LEAP's Executive Committee (Section F5) will semiannually and summatively evaluate convergence outcomes both to determine resource allocation and report to outside partners and stakeholders. This summative work will be aided by ongoing formative project evaluations conducted by two LEAP subcommittees: The Convergence Subcommittee for research and education (Section F6) and the Knowledge Transfer Subcommittee for corporate, public, and local government engagement (Section F7). Subcommittee members will be trained in both implementation and progress evaluation, including objective definition and evaluation metric development led by **Lang**, who will be responsible for collecting formative assessments and disseminating metrics and objectives.

D. DESCRIPTION OF THE BROADENING PARTICIPATION OBJECTIVES OF THE CENTER.

(**Cogburn**, **Gentine**, **Bell**, **Burbano**, **Camargo**, **Lang**, **McKinley**, **Purdie-Greenaway**, **Revkin**, **T.Zheng**). Geoscience's lack of diversity and inclusion is one of its "*largest cultural problem[s]*" [15], garnering media coverage and handicapping its knowledge transfer capacity [112]. STEM's potential to positively impact our society's most vulnerable is inextricably linked to promoting diverse perspectives in every aspect of climate data science research, education, and bidirectional knowledge transfer. Thus, "broadening participation" is not *sufficiently* achieved through diversifying talent pools (the "pipeline"); instead, this requires deep cultural and structural solutions [113]. Therefore, Center Director **Gentine** will co-lead LEAP's diversity, equity, and inclusion (DEI) initiatives alongside Chief Equity Officer and Knowledge Transfer Director **Cogburn**. LEAP's broadening participation initiatives set three objectives:

- **DEIO1:** Motivate LEAP's vision, recruitment, research, education, and knowledge transfer efforts;
- **DEIO2:** Increase representation of URMs in the burgeoning climate data science discipline; and
- **DEIO3:** Rigorously assess, review, and revise all LEAP activities through an explicitly DEI lens.

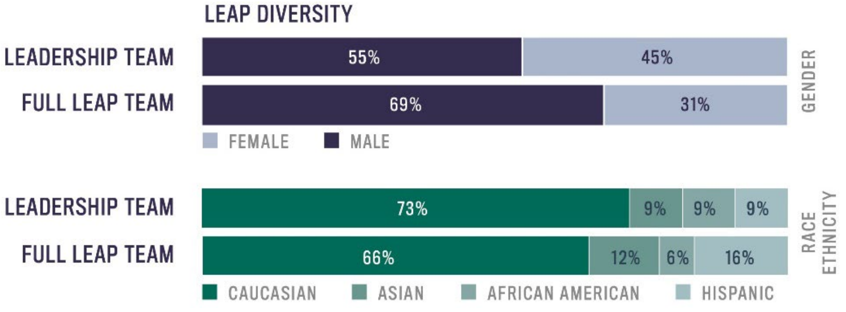


Figure 7: Diversity snapshot of LEAP's leadership and full teams.

LEAP's faculty team is designed to optimize diversity in gender, race, and ethnicity (Figure 7); the Center's DEI commitments, defined below, are evident in decisions already made to elevate members of the team into leadership positions. Ensuring a diverse faculty team enables and emboldens LEAP to reach and, preferably, *exceed* the following broadening participation objectives [114].

D1. A Transdisciplinary Ethic. LEAP will be an experimental testbed for team formation, combining inputs to achieve meaningful equity outcomes, and systematically and quantitatively assessing DEI's causal impact on research, knowledge transfer, and education. This requires training STEM researchers and educators in approaching question and methodology design, data analysis, and communications from an ethical perspective rooted in equity and parity across all racial, ethnic, gender, (dis)ability, and other positionalities.

D2. DEI Strategy, Execution, and Reflection. All LEAP researchers, trainees, and administrators will complete three annual DEI workshops; LEAP will strongly prioritize research seed funding, fellowships, and other support towards recipients who self-report having attended most or all workshops.

- **Equity Literacy & Visioning.** Paired with LEAP's Ethics Workshop held each August (see Ethics Plan), this Workshop recognizes and re-imagines LEAP's role not *only* in addressing pipeline issues, but in more meaningfully integrating diverse perspectives into climate data science. This workshop will establish and formatively guide LEAP's Center-wide DEI commitments, supporting participants through designing and enhancing equity literacy, developing an understanding of DEI's importance to science's quality and societal utility, and exploring the contributions of structural and social inequities across race, class, gender identity, sexual orientation, (dis)ability, language, and other positionalities.
- **Equity Planning.** Each January, a second workshop will develop short- and long-term strategic plans for integrating DEI principles across all Center policies, procedures, and research, education, and knowledge transfer initiatives. LEAP will develop and critically assess the Center's recruitment and data acquisition, use, and dissemination policies, and its historic and planned future partnerships. This workshop will translate the Equity Literacy & Visioning Workshop's ideas into strategic implementation plans and policy guidelines.
- **Performance Review.** Each June, a third workshop will review the Equity Assessment's findings (see below) and revise corresponding strategic plans. Subsequent workshops will address growth areas identified through the equity assessments (Section D9).

DEIO2: Increasing Minority Representation in Climate Data Science.

D3. Integrating DEI into Recruitment and Engagement. To equitably attract top talent, the Staff Writer and Inclusive Education Assistant Director will integrate LEAP's DEI convictions into trainee recruitment materials, and form a discussion guide provided to faculty involved in recruitment and partnership formation. LEAP will quantitatively evaluate strategy effectiveness across recruitment frames, assessing the effectiveness of job postings with and without an inclusive statement, and with and without exhaustive qualification descriptions [115]. Faculty and administrators supporting personnel hiring will refine their recruitment and hiring objectives through the Literacy & Visioning and Equity Planning workshops.

D4. Candidate Evaluation. Implicit and explicit biases emerge at the delicate job application evaluation stage: By overestimating academic qualifications of candidates from elite institutions, failing to critically determine characteristics and skills needed for the position, and believing that hiring criteria must be remediated to achieve diversity goals [116]. LEAP's candidate evaluation criteria will include a multi-phase hiring process, including: 1) *blinding applicant names and institutional affiliations* for all LEAP Fellowships, postdoctoral, or REU positions, Design Studio internships, and the research seed funding program; 2) work sample tasks (where applicable); and 3) use of standardized interviews. A randomized field experiment will randomly assign applications to one of three conditions: Control (open and unstructured review); Treatment 1 (single phase: blinded CV/ resumé); or Treatment 2 (multi-phase: blinded CV/ resumé, work sample tasks determined by primary hiring unit, and standardized interviews). Findings will be published in business and psychology journals and academic magazines.

D5. Bridge Program. LEAP will partner with **Columbia**'s Bridge-to-PhD program to increase the participation of students from underrepresented groups in STEM doctoral programs. LEAP will support two cohorts of three Bridge scholars each. These post-baccalaureate scholars will spend two years conducting research, taking *Certificate* courses to strengthen their doctoral applications, and receiving a salary and full tuition remission. **Cogburn** and **Purdie-Greenaway** will co-mentor Bridge scholars. Bridge staff will provide professional development and support in graduate school application writing.

D6. SOARS Integration. LEAP will closely collaborate with the **University Corporation for Atmospheric Research** (UCAR)'s *Significant Opportunities in Atmospheric Research and Science* (SOARS), one of the nation's longest-running STEM bridge programs, financially supporting and exposing underrepresented students to climate research and mentorship. Recognizing the importance for SOARS Protégés to spend their first summer conducting research at **NCAR**/ **UCAR**, LEAP will offer ten-week climate data science summer internships to SOARS Protégés returning for their second year. This internship will be complemented by weekly workshops in professional development and research computing. In Year 1, during LEAP

Fellows and postdocs' **NCAR** summer visitations, they will learn SOARS best practices and lead a workshop for perspective SOARS Protégés. In Years 2-5, LEAP Fellows and postdocs will mentor interested SOARS Protégés via monthly teleconferences; evaluator **Lang** will annually review the program. Further, SOARS alumni will be encouraged to attend the AI4Earth conference in New York City with travel and lodging scholarships, and be eligible for reduced tuition for any of the Certificate's four courses.

D7. Hackathons. Partnering with the **National Society of Black Engineers** and the **New York City Office of the Mayor**, in Years 1-3 LEAP will execute a Hackathon for climate data science, rotating across the Center's Knowledge-Data Continuum and hosted by the **Trust for Governors Island**, providing life-relevant learning [117] opportunities highlighting the immediate dangers of climate change (e.g., sea level rise). This hackathon will reach out to the broader "AI for Good" data science community, and will be facilitated by the LEAPangeo infrastructure. LEAP faculty and Fellows will teach pre-hackathon tutorials for prospective participants. Group products will be judged by the Knowledge Transfer Subcommittee and evaluated for their innovation, technical achievement, and potential societal impact. Winning teams will be invited to present at **Microsoft**'s AI4Earth conference.

D8. Outreach and Education Alignment. LEAP will leverage **Columbia**'s Community Benefits Agreement by giving priority consideration to student, postdoc, and administrative applicants living in Harlem, whenever possible and in close collaboration with **Columbia**'s Employment Information Center. LEAP's Knowledge Transfer Subcommittee will utilize this Center's partnerships to recruit a talent pool of diverse backgrounds, life histories, and experience levels.

DEIO3: Formatively Assessing DEI Impact.

D9. Years 1-4: Equity Assessment. **Cogburn** and **Purdie-Greenaway** will annually assess and develop a Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis of DEI strategy implementation, gathering formative data from in-person surveys, interviews, lab-based ethnography, and the comprehensive review of recruitment, hiring, and inclusive practices. They will build an **Inclusive Science Module**, including training spanning anti-racism/ -bias awareness, the history of race and gender in STEM, implicit bias mitigation, and growth mindset. Positioning LEAP as a testbed for equitable team training and development, this Module will be published open-source to the Center website.

E. DESCRIPTION OF THE KNOWLEDGE TRANSFER OBJECTIVES OF THE CENTER.

(**Cogburn**, **Burbano**, **Revkin**, **Abernathey**, **DeFries**, **Joppa**, **McKinley**, **Lang**, **Newman**, **Rush**, **Sukthankar**, **Uriarte**, **Zanna**, **Zarrilli**).

LEAP's knowledge transfer initiative grows out of and bolsters its broadening participation commitments. Pairing these two seminal STC components – and having **Cogburn** serve the dual role of Chief Equity Officer *and* Knowledge Transfer Director – is essential for LEAP's goal achievement and the sustained impact of all Center-wide initiatives. LEAP's knowledge transfer program sets three objectives:

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

KTO1: Bidirectional Knowledge Transfer.

E1. Communications Programming and Partnerships. The communities most vulnerable to climate change impacts have often the most limited capacity for reporting and absorbing climate news [118]. To provide public, corporate, and policy audiences access to ML-informed climate projections, and to train external audiences in navigating climate data science, LEAP will develop a *two-way* engagement strategy between academia, organizational stakeholders, and the individual citizenry. Administered by the Knowledge Transfer Program Manager and Staff Writer (Section F2), event programming and social media storytelling will foster mutual awareness and fluency. Initiating this bidirectional flow, in Year 1:

1. The **Council on Foreign Relations**, **Climate Central**, **Climate Matters**, and Public Engagement Director **Revkin** will co-host a workshop reviewing insights in covering climate change for American journalists whose outlets have shrinking capacity to cover climate news;

2. **Columbia**'s Earth Institute will launch a year-long Climate Journalism Fellowship program training reporters and editors in incorporating climate data science into compelling stories, visualizations, maps, and other media forms. Journalism Fellows will be closely mentored by **Revkin**.

E2. Applied Management Research. The private sector is greatly affected by climate change [119], and requires accurate climate projections distilled to be most pertinent to their needs. For instance, some companies consider investments in locations increasingly at risk for uncertain flooding and sea-level rise, and would tangibly benefit from more accurate projections of flooding occurrence and intensity. Likewise, many current supply chains could face catastrophic disruption due to extreme weather. Effective climate communications enable corporations to achieve business goals and ensure free flow of goods and services, but requires more knowledge of corporate stakeholders' climate adaptation expectations and needs: For investors [120], employees [121], consumers [122], regulators [123], and activists [124]. Thus, LEAP must integrate not just knowledge *transfer*, but also knowledge *generation* through research.

Corporate Engagement Director **Burbano** and **Bareinboim** will conduct experimental laboratory, survey, and within-company ethnographic research to establish causal inference (Section B4.2) in stakeholder response to climate information [121,125–129]; the Knowledge Transfer Subcommittee will develop research objectives and questions emerging from two overarching goals: 1) examine mechanisms under which communication initiatives influence perceptions and behaviors of key corporate stakeholders; and 2) inform the framing of climate projection content when engaging heterogeneous groups. Findings will inform causally-supported best practices suggesting whether, under what circumstances, and how firms *should* communicate climate information to stakeholders to best influence behavior.

An annual survey to 200 companies through the **Northeast Big Data Innovation Hub**'s listserv will map the state of climate information from private sector stakeholder perspectives, and identify the information type and format most useful for various sectors. Beginning in Year 2, this survey will transfer knowledge by reporting LEAP's web and digital outreach efforts, soliciting user preference for additional data or formats. The **Alliance for Research on Corporate Sustainability** and **Gizmodo** will disseminate opportunities and results through their multi-platform and social media channels.

E3. Executives in Residence Program. In each Fall and Spring semester, LEAP will support an Executives in Residence program, wherein executives from partner institutions will assume semester-long "visiting research scholar" appointments; during this time, Executives will teach one new graduate course, co-supervise students, participate in the *Public Conversations* series, and participate in the Knowledge Transfer Subcommittee. **Joppa** and **Newman** will be inaugural Residents in Spring and Fall 2022, respectively. Future Executives in Residence will be selected by the Knowledge Transfer Subcommittee.

E4. AI4Earth Conference. Microsoft's AI4Earth program provides seed funding, computational resources, and partnerships supporting the intersections of geophysics and data science; their annual AI4Earth conference hosts 300+ attendees. In Years 2 and 4, and as part of the Center's large partnership with **Microsoft**, LEAP will co-host this conference in New York City, thus opening the program to a new, larger, and more diverse audience. The conference will be co-planned by the Knowledge Transfer Subcommittee and all LEAP administrators; the **New York City Office of the Mayor** and the **National Society of Black Engineers** will co-host roundtable discussions with LEAP affiliates. LEAP will also provide travel grants and registration waivers to teachers, parents, and K-12 students participating in its educational programs.

E5. Convergence Luncheons. LEAP will annually host six Convergence Luncheons for its faculty, trainees, and partners, facilitating shared vocabulary development across traditional disciplines. Planned by the Convergence Subcommittee (Section F6), and supported by the Inclusive Education Assistant Director (Section F2), Luncheons will contain "lightning talks" by LEAP researchers, followed by discussion. The Program Manager will publish a meeting report for each Luncheon to the Center website.

E6. "Ask Me Anything." LEAP will initiate "Ask Me Anything" conversations on Reddit and Twitter using hashtags *#LEAPforclimate* or *#climateLEAP*. Similarly, LEAP will utilize Instagram's Question Stickers feature in its communications, allowing users to ask questions and receive LEAP's public answers. LEAP's Program Manager will lead researchers in content generation, interview organization, response transcription, and general venue moderation.

KTO2: Storytelling and Social Media Innovation.

E7. Storytellers in Residence. LEAP will launch a Storytellers in Residence program, each semester bringing a Harlem-based artist, writer, or journalist to embed and creatively translate LEAP's purpose and key findings to internal and external audiences. Residents will learn how climate data science tools are developed, thereafter supporting the Staff Writer in crafting storytelling content across owned and earned media. Residents will temporarily join the Knowledge Transfer Subcommittee as *ex officio* members, and take *Challenges* with full tuition remission. **Gizmodo** will designate a senior writer to serve as the inaugural Spring 2022 Storyteller; future Storytellers will be selected by the Knowledge Transfer Subcommittee.

E8. "Translate-a-Thons." The Staff Writer (Section F2) and Storytellers in Residence, in partnership with the **American Museum of Natural History** and **Gizmodo**, will organize quarterly, half-day intensive sessions on scientific translations, wherein LEAP faculty, postdocs, and Fellows will collaboratively develop visualization, videos, and interactive web tools. Public Engagement Director **Revkin** will maintain a list of existing knowledge needs, and prepare task descriptions, tutorials, and tools before each event, then review and refine tools and products afterwards.

E9. Public Lectures. With **Columbia**'s Center for Science and Society, LEAP's Climate Justice Leadership Board (Section F8) will organize a "*Public Conversations*" *in Climate Data Science* series connecting natural and social scientists, policy and business scholars, and the general public. Four events per year will forge convergent research and knowledge transfer, and will each host 200 attendees in-person and 500 live-streaming, each featuring a cross-disciplinary panel including an Executive or Storyteller in Residence. Events will be live-tweeted through LEAP's social media channels and recorded; the Knowledge Transfer Program Manager will post all video recordings LEAP's YouTube channel.

E10. Sharing Products. LEAP will create, organize, translate, and disseminate research findings to aid public servants, community organizations, and corporations in adapting and innovating in response to climate variability. Providing superior climate projections for 10-40 years is crucial for firms in real estate, finance, and reinsurance, and for those not yet versed in exploiting and interpreting climate data. For climate-dependent investors (**AllianceBernstein**), reinsurance companies (**SwissRe**, **AonRe**), and start-ups (**Four Twenty Seven**), access to superior climate data in an easily analyzable form creates value for core business propositions. Climate change is of increasing interest to large technology companies (**Google**, **Microsoft**, **BlackRock**). LEAP's city and community partners will directly benefit from improved climate projections when setting climate policy, programming, and fiscal budgets, particularly the new $20B investment launching the **New York City Office of the Mayor**'s and **Zarrilli**'s OneNYC 2050 Livable Climate initiative, the largest and most prominent urban climate response program in the United States.

E11. Monthly Newsletter. The Staff Writer will publish a monthly online newsletter targeting Center affiliates and external policy makers and private firms. Each newsletter will describe one key climate change indicator (e.g., NYC sea level rise, heat wave frequency), best estimates of ESM projections and uncertainties, and practical implications for various sectors (e.g., real estate investment, supply chain).

E12. Annual Report. LEAP will translate appropriate information from its NSF reports into a public and graphics-heavy annual report for the Center's website, including letters from Center Director and Diversity Co-Director **Gentine**, Deputy Director **McKinley**, and one of its Executives or Storytellers having taken residence over the year. It will highlight annual milestones and key scientific discoveries, reporting research and education seed grants activity; hackathon and translate-a-thon results; and quantified *At a Glance* accomplishments (e.g., code created, committee membership, new partnerships, and website user traffic).

F. DESCRIPTION OF THE MANAGEMENT PLAN FOR THE RESEARCH, EDUCATION, BROADENING PARTICIPATION, AND KNOWLEDGE TRANSFER ACTIVITIES OF THE CENTER.

LEAP will be administratively housed within **Columbia** Engineering's Office of the Dean, affording a greater quantity and quality of resources than a single department alone could provide; thus ensuring broad and equitable engagement. LEAP's management structure encourages shared knowledge development and frequent interactions between geoscientists and data scientists, trainees, and all partners. LEAP's programs and dedicated headquarters space will sustain collaborative interaction across disciplines (see Facilities Statement). Per Figure 8, LEAP financially prioritizes all four initiatives, and strongly support its integrated broader impacts; these budget ratios map onto space allocations in the present work.

F1. Convergent Leadership. Center Director and Diversity Co-Director **Gentine** pioneers using ML for climate convection modeling and terrestrial biosphere modeling. He currently Co-Leads NOAA's Drought Task Force, is member of the World Climate Research Program (WCRP)'s Working Group on Seasonal to Interannual Prediction and the USCLIVAR Data Science Group, and **Columbia**'s representative to the Consortium of Universities for the Advancement of Hydrologic Science. As Center Director, **Gentine** will be responsible for all Center research, broader impacts, and operations. He will Chair the Center's Executive Committee, and will establish a Director's Council. His co-leading LEAP's broadening participation efforts ensures that equity will be achieved across all Center operations. If Center Director **Gentine** must step down, Deputy Director **McKinley** will serve as the new Center Director. If **McKinley** is then unable to continue leading, the Executive Committee will nominate two LEAP engineering faculty to assume the Directorship, and the final incumbent will be selected by the External Advisory Board at its annual meeting.

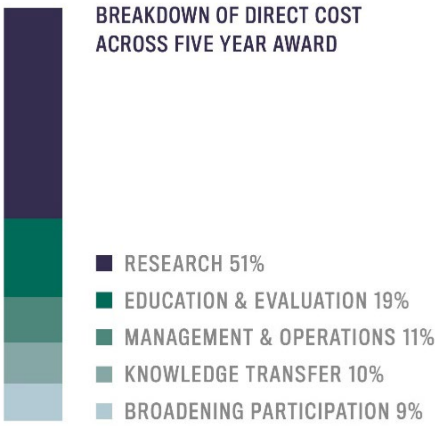


Figure 8: Breakdown of direct cost across five year award.

Deputy Director **McKinley** holds expertise in ocean carbon and biogeochemistry, facilitating knowledge transfer between geophysicists and data scientists. She is the current chair of an *Ocean Carbon & Biogeochemistry Program* Working Group on Ocean Carbon Gaps, and formerly served on the scientific steering committee for the Global Carbon Project. She has extensive University leadership experience, including chairing of the Diversity Committee for **Columbia**'s Department of Earth and Environmental Sciences. **McKinley** additionally has substantial experience communicating with policy-makers and the public. She will co-chair the Convergence Subcommittee alongside Chief Convergence Officer **T.Zheng**, and supervise Research Directors **Vondrick** and **Zanna**.

Data & Computation Director **Abernathey** is a physical oceanographer with expertise in mesoscale turbulence and deep-ocean transport. He will oversee the LEAPangeo cloud platform, and ensure integration across stakeholder computational resources, notably **Columbia**'s Data Science Institute and its Office of Information Technology. He will supervise two cloud engineers, and report jointly to Center Director **Gentine** and Deputy Director **McKinley**.

Chief Equity Officer and Knowledge Transfer Director **Cogburn** is a social worker, psychologist, and data scientist with expertise in the psychological and cultural foundations of racism, with interests in media-based racism and its effects on population health. She will oversee three workshops and the DEI consultant, the annual Equity Assessment, and personnel recruitment and evaluation protocols. She will supervise both Engagement Directors (**Burbano** and **Revkin**) and co-supervise the Knowledge Transfer Program Manager. She will chair the Knowledge Transfer Subcommittee.

Chief Convergence Officer and Education Director **T.Zheng** holds expertise in ML, spatiotemporal modeling, pattern identification, and educational innovation. As **Columbia**'s Chair of Statistics and its Data Science Institute's Associate Director for Education, she will develop innovative pedagogical and evaluation initiatives positioning LEAP's faculty, trainees, and K-12 stakeholder as agents to unify all Center efforts. She will supervise the evaluation of LEAP's education, broader impacts, and knowledge transfer activities.

Data Science Director **Vondrick** will supervise ML and causality development; Geoscience Director **Zanna** will oversee geoscientific research spanning all four Earth subcomponents. Both will report to Deputy Director **McKinley** and join the Convergence Subcommittee.

Two Engagement Directors will report to Chief Equity Officer and Knowledge Transfer Director **Cogburn**: Corporate Engagement Director **Burbano** studies corporate sustainability, will govern website and print communications targeting LEAP's industry partners, and will conduct research into corporate communication design. Public Engagement Director **Revkin** is Founding Director of **Columbia**'s Initiative on Communication and Sustainability, for 30 years previously a climate journalist and scientific visualization expert. He will strategize internal and external storytelling efforts, including coordinating the Storyteller in Residence program, developing public programs, and building knowledge transfer partnerships.

NCAR Model Development Liaison **Lawrence** will supervise integration within the CESM; **GISS** Model Development Liaison **Schmidt** will supervise integration of LEAP's parameter estimation strategy into ModelE. Both will co-supervise LEAP Fellows and join the Executive Committee.

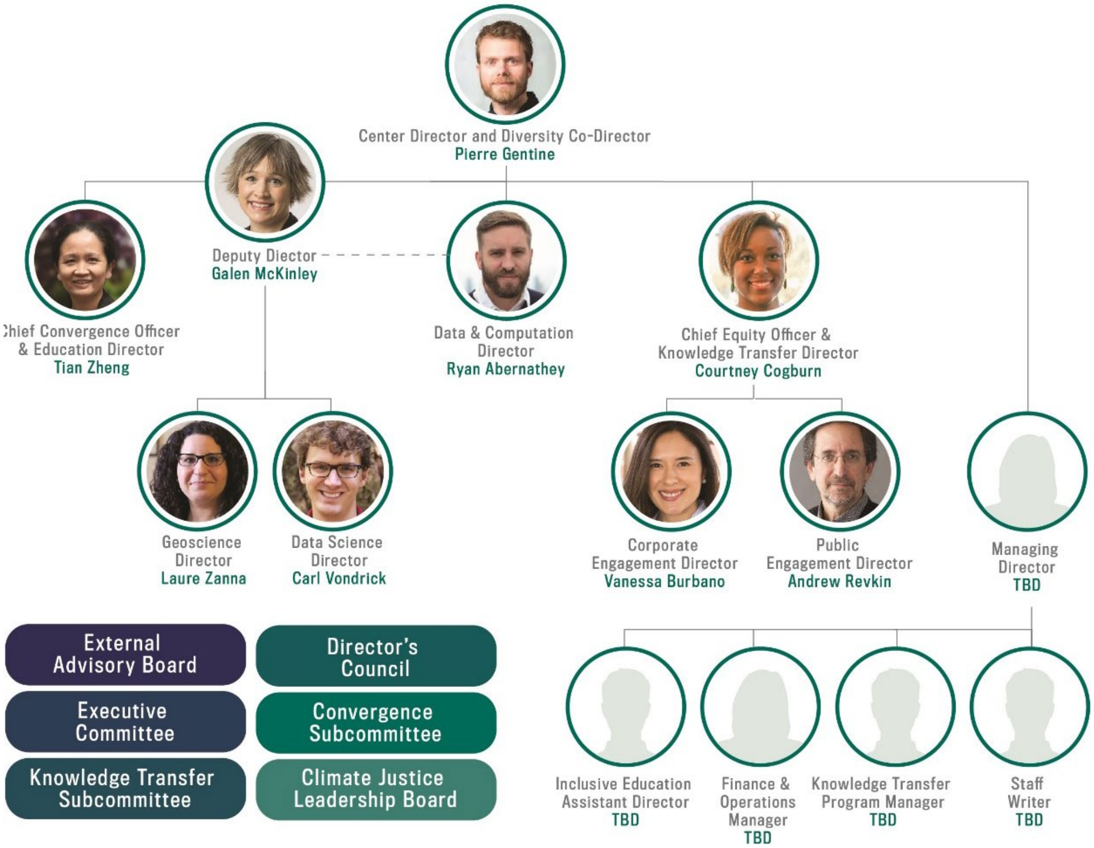


Figure 9: Organizational chart of LEAP faculty and governance committees.

F2. Administrative Management. A **Managing Director** will report to Center Director and Diversity Co-Director **Gentine**, serving as the chief guardian of Center resources and policies; the incumbent will assume responsibility for all Center operations, staff supervision, grants administration and compliance, NSF reporting, annual Site Visit planning, staffing all governance bodies, and undertaking special projects. The Managing Director will be expected to participate actively in professional research administrator societies.

Supporting the Managing Director will be four full-time administrators: An **Inclusive Education Assistant Director** will support LEAP's educational programs for faculty, graduate, undergraduate, K-12, teacher, and parent trainees, will facilitate rotation of LEAP Fellows to **NCAR**, and will support the Climate Justice Leadership Board. A **Knowledge Transfer Program Manager** will support programming linking academia with external stakeholders; the incumbent will manage the Center's website and social media platforms, author publications and opinion pieces, and steward and establish external partnerships with industry, state, and local governments, and NGOs. A **Finance & Operations Manager** will serve as LEAP's financial accounting, procurement, and HR specialist. The incumbent will be responsible for trainee hiring and payroll, and the two Residence programs. Finally, LEAP's **Staff Writer** will author content targeting corporate, non-profit, and public audiences. The incumbent will author a monthly newsletter and the annual report, publish articles across owned and earned media outlets, and supervise communications vendors.

Convergent Governance Committees.

Six groups will govern LEAP's research, education, broader impacts, knowledge transfer, and operational initiatives. To steward convergence, each group's membership spans disciplines and institutions.

F3. External Advisory Committee. A 6-person External Advisory Committee (EAC) of non-affiliate executives will meet annually, timed with the NSF's Site Visit, and will contain two representatives from diversity in science organizations, one director of an expired STC, one chief executive officer of a private firm, one senior geoscientist, and one senior data scientist. The EAC will receive an annual review, first reviewed by the Director’s Council. The EAC will advise on strategic partnership formation, the external competitive environment, and on complementary NSF grants to bolster its broader impacts (e.g., NRT).

F4. Director’s Council. Six to eight senior executives across LEAP’s core institutions – primarily Deans and Institute Directors – will meet each March or April to receive a preliminary annual report in advance of the EAC, and provide feedback for improvement, identify areas for providing administrative guidance, and support continuous and streamlined inter-institutional integration. This Council will include Center Director **Gentine** and **McKinley** as *ex officio* members, and will be staffed by the Managing Director.

F5. Executive Committee. Chaired by Center Director and Diversity Co-Director **Gentine**, an Executive Committee (EC) will have nine other voting members: **McKinley**, **Abernathey**, **Burbano**, **Cogburn**, **T.Zheng**, **Lawrence**, **Schmidt**, **Vondrick**, and **Zanna**. Staffed by the Managing Director and meeting monthly, the EC will be LEAP’s principal forum for strategic and operational decision-making. The EC will review Center progress, select LEAP Fellows, steward existing and forge new institutional partnerships, resolve internal conflicts, and receive operational updates from the Managing Director.

F6. Convergence Subcommittee. Co-chaired by Deputy Director **McKinley** and Chief Convergence Officer **T.Zheng**, and with Research Directors **Vondrick** and **Zanna**, Chief Equity Officer **Cogburn**, and evaluator **Lang**, a Convergence Subcommittee will meet monthly to review research progress, evaluate overall Center integration, select LEAP Fellows, award research seed funds, and plan the monthly *Convergence Luncheons*. Importantly, the Convergence Subcommittee will be responsible for: 1) advancing promising research projects; 2) launching new projects based upon formative evaluation data; and 3) determining whether existing projects should be expired and its resources distributed elsewhere. Projects will be expired depending upon NSF feedback, waning faculty or trainee interest, emergence of new projects that more directly align with LEAP’s Knowledge-Data Continuum, and/or improper productivity.

F7. Knowledge Transfer Subcommittee. Chaired by Chief Equity Officer and Knowledge Transfer Director **Cogburn**, including Engagement Directors **Burbano** and **Revkin** and Senior Personnel **Goddard**, **Sukthankar**, **Joppa**, **Newman**, and **Zarrilli**, and staffed by the Knowledge Transfer Program Manager, a Knowledge Transfer Subcommittee will: 1) govern the Storytellers in Residence program; 2) plan the DEI Workshops, Hackathon, Translate-a-thon, and AI4Earth symposium, particularly stewarding corporate and public engagement; and 3) track website usage. Subcommittee members will support the Convergence Subcommittee in selecting trainees and in anonymizing personal identifying information, and support Director **Gentine** and Deputy Director **McKinley** in selecting Executives and Storytellers in Residence.

F8. Climate Justice Leadership Board. LEAP will support over 40 Fellows across its initial award period. This graduate community will annually elect a President, Vice President, and Communications Director to form a Climate Justice Leadership Board (CJLB), with each position holding one-year terms. The CJLB will work with Chief Equity Officer **Cogburn** and evaluator **Lang** to accomplish two objectives: 1) annually partner with an NYC-based climate organization that targets vulnerable populations to engage in broadening participation, education, and knowledge transfer initiatives, and semiannually co-author opinion pieces on the nexus of climate prediction and social justice, for publication in earned and owned media, with editorial support by the Staff Writer; and 2) led by **Lang**, be trained and gain experience in evaluating LEAP’s comprehensive impact. Each August and January, evaluator **Lang** will lead a daylong pro-seminar training the CJLB in formative and summative evaluation methods and data collection. CJLB members will receive additional compensation for designing their own studies, collecting data through mixed methods, and developing an annual SWOT Analysis for presentation to the EAC. In addition to its formative evaluation work, the CJLB will coordinate *Conversations*, including speaker invitation, marketing, and evaluation.

F9. Quantum LEAP. Today’s climate projections are insufficient to support human needs and societal resilience in the face of mounting climate risk. Our team sees tremendous promise in developing a convergent approach to ML-guided parameterizations that dramatically improve climate projections for the CESM: This highly-ambitious endeavor is only possible with trailblazing engagement by our nation’s top climate research laboratories, the talent of new and diverse trainees, and bidirectional engagement with a broad community of public and private decision-makers who would benefit from more accurate future projections. Ameliorating critical shortcomings in current climate models is a worthy and necessary venture for our nation, and will pay dividends for many generations to come. *This is LEAP*.