PROJECT EXECUTIVE SUMMARY

Title: NeuroX-Fusion: Unified Foundation Model of Brain for Transformative Neuroscience

PI and Co-PI(s): Shinjae Yoo (BNL, PI), Jiook Cha (SNU, Co-PI), Chun Kee Chung (SNU, Co-I), Jay-Yoon Lee (SNU, Co-I), Taesup Moon (SNU, Co-I), Franco Pestilli (UT Austin, Co-I)

Applying Institution/Organization: Brookhaven National Laboratory, The University of Texas at Austin, Seoul National University

Resource Name(s) and Number of Node Hours Requested: Aurora. 1,269,000 node-hours

Amount of Storage Requested: 600 TB

Executive Summary : Modern neuroscience is limited not by data scarcity but by the absence of a framework for integrating hundreds of terabytes scale, multimodal, multi-scale brain data into a single, computable object that captures the brain's dynamic architecture. *NeuroX-Fusion* will close this gap by training a 130 billion-parameter foundation model that unifies multimodal MRI and electrophysiology into one latent representation, enabling cross-modal decoding, mechanistic in-silico perturbations, and rapid hypothesis testing that were previously infeasible.
Our scientific program is a staged campaign to construct this unified model. We will first develop two Large Brain Models (LBMs): **NeuroX-MRI**, built on our novel 4D Swin Transformer architecture to decode mesoscale brain dynamics from multimodal MRI data, and **NeuroX-Ephys**, a channel-equivariant model to decipher millisecond-scale neural syntax from heterogeneous electrophysiology recordings (ECoG, EEG, sEEG, and MEG). These two models will then be aligned through the shared semantic space of a Large Language Model (LLM), enabling the translation of latent brain states into explicit linguistic concepts. Our scientific program pursues three breakthrough objectives:

Decode the Unspoken Mind: Generate continuous, high-fidelity read-outs of internal affective and cognitive states and health from multimodal MRI and electrophysiology, enabling transformative applications, including early biomarkers for psychiatric disorders, novel brain therapeutics, and brain computer interfaces.

Bridge Brain Activity Across Scales and Modalities: Create the first unified model that fuses the high-resolution spatial data from MRI with the millisecond-scale temporal precision of electrophysiology. By aligning the disparate data streams, we will create a "digital twin" of brain function to probe the links between large-scale network dynamics and high-speed neural codes.

Advance Exascale AI for Science: Develop open-source novel AI architectures and data pipelines optimized for leadership-class systems. These innovations will create a transferable blueprint for data-centric AI in other complex DOE-relevant domains, such as climate science and materials discovery. To ensure broad adoption and reproducibility, all models and data pipelines will be integrated and disseminated through our established open-science platform, brainlife.io, democratizing access to these powerful tools for the global research community.

Converging these models requires training on over 400 TB scale of datasets, a process that necessitates the extreme I/O bandwidth and parallel file system capabilities unique to leadership-class systems like Aurora. Through our previous DOE ALCC award, we demonstrated excellent strong-scaling efficiency up to 512 nodes on Aurora, showing our computational readiness. We request 1,269,000 node-hours on Aurora over three years.

By creating a quantitative, computable model of the mind, NeuroX-Fusion will establish a new paradigm for neuroscience. An INCITE allocation is indispensable to converge this innovative integration of brain and AI, and ultimately, to transform the unspoken interior of the human mind into the domain of measurable, tractable science.

PROJECT NARRATIVE

1 SIGNIFICANCE OF RESEARCH

1.1. Grand Scientific Challenge: From Brain Mapping to Brain Decoding. Modern neuroscience is data-rich but knowledge-poor. We have over 400 terabytes of brain data from fMRI and EEG, but lack the ability to read its language. The fundamental barrier is the brain's staggering multi-scale complexity; a single thought unfolds across milliseconds (electrical spikes captured by Ephys), seconds (blood flow changes in fMRI), and is constrained by static anatomical wiring (dMRI). To truly understand the brain, we must unify these disparate data streams–yet no such scientific framework exists. The consequences are a global crisis in mental and neurological health, with conditions still diagnosed by subjective observation and an economic burden projected to reach $16 trillion by 2030 [1]. This crisis demands a new class of scientific instrument: a unified, computational model of the brain. This proposal details the construction of that instrument, a foundation model designed to finally bridge the brain's multiple scales and translate its complex signals into measurable, tractable science.

1.2. Foundation Models for Predictive Models for Complex Systems. The fundamental gap between mapping and understanding the brain persists because the brain's complex, multi-scale dynamics render traditional methods ineffective [2, 3]. To address this, we propose to leverage foundation models, an AI paradigm that has become the essential tool for tackling such complexity. As demonstrated in domains from climate science to biology [4]-[5] and neuroscience [6, 7], large-scale models trained on massive heterogeneous data develop powerful emergent capabilities for prediction, decoding, and generation. This makes them uniquely suited to the challenges of modern neuroscience.

However, today's AI models for the brain remain fragmented. They are limited to single modalities and operate at capacities far below the regimes where advanced reasoning appears in language and vision [8]. Crucially, no existing model achieves the cross-modal, multi-scale integration needed to connect brain circuits, dynamics, and cognition altogether. Overcoming these limitations requires exascale computation. Our proposed NeuroX-Fusion model directly targets this gap, providing the first unified, 130 billion-parameter AI model designed to bridge disparate brain data with language representations at the exascale.

1.3. The Proposed Scientific Breakthrough: An Integrated, Multi-Scale Science Plan. Our project is built on a staged, three-pillar science plan to create a unified model of the mind.

Pillar I: NeuroX-MRI — Capturing the Dynamic Functional Connectome

The Scientific Gap: Systems neuroscience is dominated by the concept of the functional connectome, yet our analytical tools are fundamentally mismatched to its nature. We rely on methods that produce static, correlation-based "snapshots" of brain networks, which implicitly assume that connectivity is constant over the entire measurement period [3]. This approach fails to capture the brain's most critical property: its dynamic, moment-to-moment reconfiguration of neural circuits. This dynamic activity underpins the order and disorder of cognition, yet its neglect leaves us blind to the network instabilities that precede major neuropsychiatric disorders. Consequently, our understanding is limited to group-level biomarkers, falling short of the actionable, predictive insights that modern medicine demands.

Our Breakthrough Approach: We confront this limitation with NeuroX-MRI, a foundation model built on our novel Swin 4D fMRI Transformer architecture [9]. This approach is the first to learn directly from full-resolution 4D fMRI data (3D space+time), resolving a long-standing computational bottleneck that forced prior research to discard dynamic information [3]. In this proposal, by integrating fMRI with structural and diffusion MRI in a self-supervised way, NeuroX-MRI will learn the intrinsic "rules" of the brain's dynamic functional connectome—how network states evolve over time, constrained by their physical wiring. At a scale of tens of billions of parameters, it will possess an unprecedented capacity to capture population-level variability while detecting the subtle, transient connectivity changes that define cognition.

Scientific Payoff: NeuroX-MRI yields the first predictive AI model of the dynamic functional connectome, shifting neuroscience from static correlation to dynamic causation. It serves as a foundational, pre-trained "brain simulator" for the global neuroscience community, changing how research is done. Any researcher can leverage NeuroX-MRI with zero-shot or minimal fine-tuning for their downstream tasks, enabling two transformative shifts:

- **From Conceptual Theories to Testable Models:** Researchers will be able to go beyond just debating conceptual theories and directly test them. A hypothesis about how the brain switches between internal and external attention, for instance, can be translated into a task-prompt for the model, which can generate precise, falsifiable predictions of the resulting brain network reconfiguration. This will transform qualitative ideas into quantifiable science.
- **From Group Averages to Individualized Prediction:** The model will enable true precision neuroscience. By fine-tuning the model to specific clinical targets (e.g., Alzheimer's disease), a researcher will be able to identify subtle, subject-specific dynamic signatures that predict treatment response or relapse risk. This moves beyond the static, group-level biomarkers that dominate the field today and opens the door to personalized diagnostics and interventions.

In essence, NeuroX-MRI will provide a powerful, generalizable engine to dramatically accelerate hypothesis testing, biomarker discovery, and our fundamental understanding of brain function.

Pillar II: NeuroX-Ephys — Deciphering the Brain's High-Speed Neural Code

The Scientific Gap: While fMRI can capture meso-scale brain dynamics, it lacks the temporal precision to resolve the micro-scale (millisecond) neural codes that form the "language" of thought. Electrophysiology (Ephys) methods can capture this high-speed activity, but progress has been hindered by extreme data heterogeneity. Because electrode layouts differ for each individual, it has been difficult to integrate data across subjects to build a generalizable, population-level model needed for broad scientific and clinical impact.

Our Breakthrough Approach: We will develop NeuroX-Ephys, a novel foundation model designed to overcome this heterogeneity [10]. Its core innovation is a layout-agnostic backbone architecture that treats electrode recordings as a "set" rather than a fixed grid, making its representations equivariant to electrode orderings, able to generalize to arbitrary number and placement of electrodes. By incorporating flexible spatial encodings and attention mechanisms, the model can ingest and harmonize diverse Ephys data—from 20-channel EEG caps to hundreds of intracranial electrodes—enabling true generalization across individuals for the first time. We will train NeuroX-Ephys on an unprecedented scale of multimodal data, aligning Ephys recordings with ~50,000 hours of concurrent video capturing the participants' overt actions and behaviors to ground its representations in meaningful real-world events.

Scientific Payoff: NeuroX-Ephys will serve as the first universal, pre-trained foundation model for high-speed neural dynamics, acting as a powerful engine for neurotechnology, clinical neurology, and fundamental neuroscience.

- **Revolutionizing Brain-Computer Interfaces (BCIs):** It will provide a universal "neural decoder" that works out-of-the-box, eliminating the need for painstaking, user-specific calibration. A neuroprosthetic arm, for example, could interpret a new user's motor-intent signals with minimal fine-tuning, dramatically expanding the accessibility of intuitive neurotechnology.
- **Enabling Proactive Clinical Intervention:** The model facilitates a shift from reactive to predictive neurology. A prime use-case is seizure forecasting, where learning subtle pre-ictal signatures from large ECoG/EEG datasets can provide a reliable warning minutes in advance—a life-changing capability not achievable today.
- **Unlocking the "Grammar" of Thought:** For basic science, the model provides an unprecedented tool to investigate the temporal "grammar" of the brain's code. By analyzing its internal representations, researchers can pinpoint the neural patterns underlying specific cognitive events and even generate semantic descriptions of neural activity.

Pillar III: NeuroX-Fusion — Unifying Modalities to Create a "Digital Twin" Brain

The Scientific Gap: A truly comprehensive and causal model of the brain—one that seamlessly bridges anatomy, multi-scale dynamics, and behavior—remains the holy grail of neuroscience. Such a model is currently impossible to build, primarily because it would require vast datasets where all modalities (MRI, Ephys) are recorded simultaneously, which are exceptionally scarce.

Our Breakthrough Approach: In this capstone pillar, we will achieve this unification not by waiting for impossible data, but through a landmark innovation. We will create NeuroX-Fusion, the first model of its kind, by fusing our specialized NeuroX-MRI and NeuroX-Ephys models. The key is a novel strategy that uses a Large Language Model (LLM) as a shared semantic "Rosetta Stone."

- **Language as a Semantic Bridge:** We circumvent the need for simultaneously-recorded data by aligning each brain modality independently to a common linguistic embedding space. For instance, the fMRI signal for "watching a movie" and the Ephys signal for the same activity will both be mapped to the LLM's representation of the concept 'watching a movie'. This allows us to link disparate data streams through their shared meaning, a completely new paradigm for multimodal fusion.
- **Triangulation with Vision:** This fusion will be further enriched by incorporating a pre-existing vision foundation model. This allows us to triangulate between modalities: if the MRI and Ephys models map a given experience (e.g., watching a suspenseful movie) to similar semantic embeddings, and the vision model confirms those embeddings from the video content, all components are synchronized on that brain state.

The architecture of NeuroX-Fusion will consist of our two transformer backbones with a higher-level integration module interfacing with the frozen LLM and vision models. This system will be trained via a staged process of adapter tuning. We will validate the final fused model by testing its ability to make novel, emergent cross-modal predictions. For instance, using a textual prompt like "smelling garlic while recalling a childhood memory", the model will be tasked to generate the predicted fMRI and Ephys signals for that experience. Success in such demanding tasks will provide proof that NeuroX-Fusion synthesizes disparate brain modalities.

Scientific Payoff: This creates the first end-to-end, computable model of the mind—a *"digital* *twin" of brain function* [11]. NeuroX-Fusion will map the entire causal pathway from brain structure, through multi-scale dynamics, to cognition. This provides a revolutionary platform for in silico experiments currently confined to the realm of science fiction. Researchers can, for the first time, simulate the effects of a structural lesion or targeted brain stimulation, predict the downstream consequences on high-speed neural codes, and mechanistically test hypotheses about brain function and dysfunction—tasks that are ethically prohibitive and technically impossible in human subjects today. NeuroX-Fusion represents a fundamental leap towards a quantitative, predictive science of the human mind.

1.4. Why Now? — A Convergence of Data, Algorithms, and DOE Exascale Resources. This ambitious project is only now feasible because of an unprecedented and timely confluence of three critical factors:

- **Data Explosion.** Multimodal neuroimaging consortia [12-14] now exceed 10 PB and more than half a million subjects, while large-language corpora surpass 5 trillion tokens. The *AI for Science* report flags precisely this data deluge and calls for foundation models that can learn across diverse modalities at exascale [11].
- **Algorithmic Readiness.** Our team is computationally ready. Grounded in a proven track record of scaling AI on DOE systems, we have already de-risked the core of this project. Under a prior ALCC award, we developed and scaled SwiFT, our 8.8B-parameter prototype, demonstrating its superior performance and scalability on leadership-class systems [15]. This prior work established both the core architecture and the neural scaling laws [16] for our domain, making the next logical step—scaling to >100B parameters—a well-defined engineering challenge rather than a speculative research endeavor.
- **Infrastructure Alignment.** The architectural capabilities of Aurora—its >2 EF peak performance, high-bandwidth memory, and extreme-throughput DAOS file system—are precisely the hardware prerequisites for training foundation models of this scale and complexity [11]. These systems were built for this class of problem.

1.5. Competitive Landscape, Broader Impact, and Readiness. This project is unparalleled in its ambition, scale, and commitment to open science. While other leading academic and industry labs pursue AI for neuroscience, their efforts are limited to a single modality, operate at a smaller scale, or rely on proprietary, closed models. NeuroX-Fusion is the only project targeting the triple-fusion of brain (MRI+Ephys) and language at this scale. Our commitment to open-sourcing the resulting models, disseminated through our neuroscience data platform (brainlife.io [17]), will provide an instrument for the entire global research community.

The **broader impact** will be transformative:

- **For Neuroscience and Medicine:** It will deliver a new class of predictive, objective biomarkers for diagnosing and tracking psychiatric and neurological illness, enabling a true precision medicine approach for brain health.
- **For AI and Technology:** It will open new avenues for AI systems that can better understand human intent, leading to symbiotic brain-computer interfaces.
- **For the DOE Mission:** It will deliver a transferable blueprint and reusable AI "building blocks" for data-centric science in other complex domains central to the DOE mission, from plasma physics and climate modeling to materials science and cosmology [11].

Proven Readiness: Our team is uniquely positioned for success and enters at a high state of technical readiness (TRL-4). We combine expertise in computational neuroscience with demonstrated leadership in scaling AI on DOE systems. This is complemented by our experience building the brainlife.io open-science platform, which serves thousands of researchers globally [17]. Our previous DOE ALCC award was crucial for de-risking this campaign: we successfully scaled our 4D fMRI model to 8.8B parameters and achieved 30.3% strong-scaling efficiency on 512 Aurora nodes, validating our codebase and readiness [15].

1.6. Indispensability of the INCITE Allocation. This research is impossible without a leadership-class system like Aurora. The scale and complexity of NeuroX-Fusion place it fundamentally beyond the capabilities of any other national resource. Three factors make an INCITE allocation indispensable:

- **Computational Scale:** Training the unified ~130B-parameter NeuroX-Fusion model requires millions of GPU-hours, an order of magnitude beyond the scope of other allocation programs. This is not just a matter of time, but of capability; only a leadership-class system can complete these runs within a scientifically meaningful timeframe.
- **Architectural Need:** Our model's design, particularly its Mixture-of-Experts (MoE) layers and hundreds of terabytes scale multimodal data ingestion, is a direct match for Aurora's unique architecture. The training process generates massive, complex communication patterns and extreme I/O demands that require the specific combination of Aurora's compute power, its high-bandwidth memory, and the >1.6 TB/s I/O of its DAOS parallel filesystem to run efficiently.
- **Timeliness and Urgency:** We stand at a unique convergence of data availability, algorithmic readiness, and hardware capability. Delaying this work to rely on smaller, slower clusters would stretch the training timeline from months to years, forfeiting a decisive first-mover advantage and missing a critical scientific window to establish a new paradigm for neuroscience.

In summary, NeuroX-Fusion directly answers DOE's AI grand challenges by creating a foundational instrument for predictive neuroscience. Its success, however, hinges entirely on the specific scale and architectural capabilities unique to the INCITE leadership-class systems. An INCITE award is therefore not merely beneficial—it is absolutely indispensable.

2 RESEARCH OBJECTIVES AND MILESTONES

Overview of objectives and milestones: In three years, we will deliver NeuroX-Fusion, a 130 B-parameter foundation model that fuses multimodal MRI and electrophysiology into a single language-aligned representation of the human brain. Two parallel streams—**NeuroX-MRI** and **NeuroX-Ephys**—will progressively scale from 10B to 50B, and ultimately to a unified 130B -parameter model, **NeuroX-Fusion**. The final model will be the first open model to map the full chain from brain structure through neural dynamics to behaviour, enabling population-scale, predictive neuroscience. All models and codebases will be released to the community at the end of each major milestone. Our three objectives are: 1) Multimodal MRI Foundation Model aligned with LLM, 2) Multimodal EPhys Model aligned with LLM via video, and 3) Unified Multimodal Neuroimaging Foundation Model (**Table 1, Figure 1**).

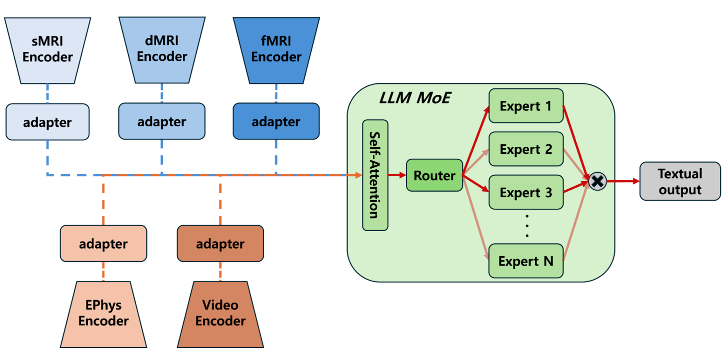

Figure 1. Schematic overview of NeuroX-Fusion. The model integrates five modalities through dedicated encoders: sMRI, dMRI, and fMRI, along with electrophysiology (EPhys) and video data. Each modality is processed through its respective encoder and adapter module before being fed into a Mixture of Experts (MoE) language model. The MoE system uses a router to dynamically select appropriate experts for processing multimodal inputs, generating textual responses. Dashed lines indicate the multimodal input pathways into the unified LLM framework.

Table 1. Summary of objectives and milestones

Year	Stream	Mile-stone	Model Size (MoE)	Data Modality	Key Evaluation Task	Major Risk → Mitigation	Node-hours
1	MRI	M1.1	10 B (MoE 8×)	resting-state (rs) fMRI, sMRI, dMRI	Decoding cognitive states decode (F1 ≥ 0.85) and Emotion states ($R^2 \geq 0.3$)	MoE training instability → Train a non-MoE model	288k
1	EPhys	M2.1	10 B (MoE 8×)	Scalp-EEG	5% perf. increase in various EEG tasks	Scaling law breakdown → Harder pretext task	122k
2	MRI +Text	M1.2	50 B (MoE 8×)	rsfMRI, active-fMRI sMRI, dMRI, Text	Multimodal cognitive tasks (HCP/HBN accuracy, modality ID, QA)	Modality encoders' MoE convergence failure → Removal of MoE	155k
1,2	EPhys	M2.2	50 B (MoE 8×)	Scalp-EEG, iEEG , MEG	Joint training outperforms single modality training	Insufficient MEG data → Removal/addition of MEG data	527k
2	EPhys +Text	M2.3	50 B (MoE 8×)	Scalp-EEG, iEEG, MEG, video paired with iEEG, Text	Cross-modal performance gains	Insufficient Paired Data → Attain more data from external source, utilize video captioning	38k
3	MRI +EPhys +Text	M3.1, M3.2	130 B (MoE > 8)	all MRI and EPhys modalities, Text	Cross-modal complex reasoning	Convergence failure → curriculum fine-tuning	139k

Objective 1: Build Multimodal MRI Foundation Model (NeuroX-MRI) aligned with LLM

Goal. Our initial step involves developing a 91B multimodal MRI foundation model, aligned with a Large Language Model (LLM). This model will capture both structural and functional brain representations by leveraging the general knowledge and neuroscientific priors present in established LLMs. We will begin by scaling up our existing SwiFT architecture, incorporating a compact Mixture-of-Experts (MoE) [18] with eight experts, and integrating structural and functional MRI data (**Figure 2**). Subsequently, we will expand the architecture to 91B using advanced MoE techniques [19, 20], adding task-fMRI data to encompass a wider range of functional dynamics. The MRI model will be aligned with the LLM through modality adapter modules, which will ground the brain representations in linguistic concepts (**Table 2**).

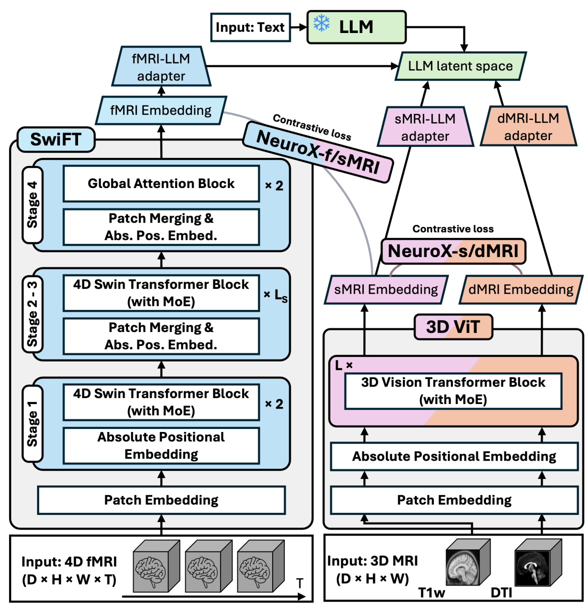

Figure 2. Our NeuroX-MRI model architecture. This framework unifies brain function (fMRI via 4D SwiFT) and structure (sMRI/dMRI via 3D ViT with MoE). The core mechanism involves two stages: **first**, contrastive learning aligns the MRI modalities. **Second**, adapter modules project these unified brain representations into a LLM's latent space, linking brain states to linguistic concepts.

Rationale. The main challenge in neuroscience is creating a model that jointly integrates the brain's anatomy (structural MRI), axonal architecture (diffusion MRI), and spatio-temporal dynamics (fMRI). Current pretrained models in neuroscience focus on a static modality or heavily reduce the features of the data, failing to produce a comprehensive picture of macro/micro brain structure and function [6, 21-24]. No model has yet achieved an end-to-end unification at scale [25].

Readiness. Our team developed SwiFT, the first 4-D Transformer designed to maintain near-linear computational complexity relative to sequence length [26]. Under a prior ALCC

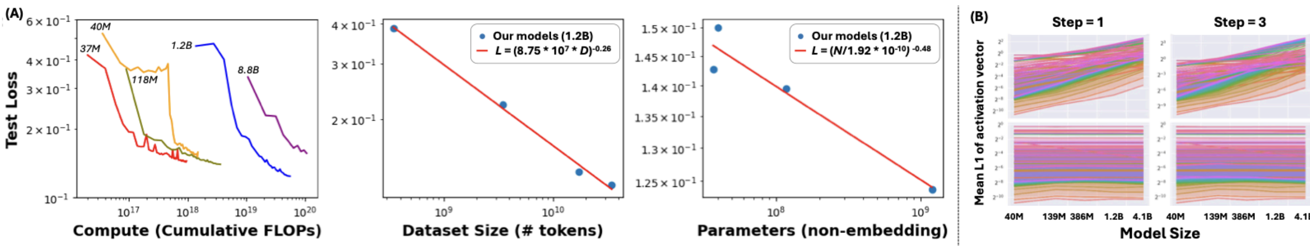

Figure 3. De-risking Large-Scale Training: Validated Scaling Laws and Training Stability. This figure validates our computational readiness, showing (A) predictable performance scaling and (B) robust training stability.
(A) Neural Scaling Laws (Frontier): Our fMRI Swin Transformer model's (SwiFT) performance predictably improves with more compute, data, and parameters [15]. This validates the scientific rationale for scaling up.
(B) Training Stability with μTransfer (Aurora): The key challenge of large model training is instability. The top row shows the activation divergence that occurs with standard training. In contrast, the bottom row shows that our implementation of μTransfer [27] stabilizes training across various model scales, **a critical prerequisite for the success of this project**.

project, we scaled SwiFT to 8.8B parameters, achieving 93% GPU utilization on Frontier. Our pre-training on fMRI data from 45k subjects demonstrated neural scaling laws [16] (**Figure 3A**). To enhance training stability across various model scales, we implemented μTransfer [27], a powerful hyperparameter tuning technique, reducing the tuning cost by 90%. Our implementation of μTransfer stabilized activation norms across model scales up to 4.1B parameters (**Figure 3B,** bottom row), preventing the divergence seen in standard training (**Figure 3B,** top row). Also, we successfully trained a 4.1B model on Aurora using 256 nodes, completing each epoch in 13 minutes with 20 TB of data, a notable performance enabled by DeepSpeed-ZeRO-2 [28], bfloat16 [29], and Flash-Attention 2 [30].

Table 2. Milestones of Objective 1

Milestone	Description	Node-hours	Key Evaluation Task
M1.1 (Y1)	57B NeuroX-f/sMRI & 34B NeuroX-s/dMRI with 8 experts MoE	288k	HBN valence regression $R^2 \geq 0.3$
M1.2 (Y2)	Build a 91B multimodal NeuroX-MRI and align with LLM	155k	Multimodal cognitive tasks (HCP/HBN accuracy, modality ID, QA)

Milestone 1.1 (Year 1): 57B-parameter NeuroX-f/sMRI and 34B-parameter NeuroX- s/dMRI with 8 experts MoE

Our initial milestone is to develop 57 billion-parameter foundation models that perform joint, end-to-end learning across the three core MRI modalities: macro-anatomy (structural MRI), micro-structural connectivity (diffusion MRI), and spatio-temporal dynamics (functional MRI).

Our strategy leverages structural MRI as a natural anchor for unification, as it is commonly acquired alongside both fMRI and dMRI. Using these paired datasets, we will develop two foundation models to learn the fundamental relationships across brain modalities: NeuroX-f/sMRI, to capture the coupling between brain structure and function, and NeuroX-s/dMRI, to bridge macro-anatomy with micro-structural connectivity. We will train both models using a contrastive learning objective (InfoNCE loss [31]). This method for multimodal integration will force the models to learn a shared representation that captures meaningful, subject-specific brain features. Architecturally, these models will use our SwiFT architecture for the fMRI encoder, and Vision Transformer [9] for sMRI/dMRI encoders. To achieve massive scale efficiently, we will integrate Mixture-of-Experts (MoE) layers into the transformer backbones. For downstream evaluation, embeddings from the frozen, pre-trained encoders will be fused via a lightweight linear layer. As a risk mitigation plan for MoE training instability, a dense feed-forward network will be kept ready as a fallback option.

288k node-hours are assigned to ***Milestone 1.1***, including (1) Architecture Exploration (70.4k), (2) Pre-training and Scaling (204.8k), and (3) Validation on downstream tasks (12.8k) .

Milestone 1.2 (Year 2): Build a 91B multimodal NeuroX-MRI and align with LLM

In this milestone, we will achieve **LLM-centered multimodal integration using cross-attention adapter modules** to enable the learning of complex interactions between MRI modalities (fMRI, sMRI, and dMRI) and text [32-34]. Building on the pre-trained MRI encoders from ***Milestone 1.1***, we will train adapter modules that project modality-specific features into the LLM's latent space for text alignment. The LLM will identify modality types through text instructions rather than parameter tuning, enabling efficient training with manageable GPU requirements.

After developing the LLM-aligned NeuroX-MRI and testing its scaling properties, we will implement MoE specifically designed for multimodal models. Building on recent studies showing MoE's potential in cross-attention adapter-based multimodal LLMs [20], we will scale up our model using advanced MoE techniques including decoupled shared/routed experts [35], bias-based load balancing [36], and high expert-dropout ratios [37] to enhance multimodal capacity, training stability, and convergence. As a risk mitigation plan, we will exclude MoE from modality-specific encoders if training becomes unstable.

Since no benchmark datasets exist for multimodal MRI models, we will evaluate LLM-aligned NeuroX-MRI in two ways (**Table 3**): 1) Quantitative evaluation comparing ***Milestone 1.1*** models with NeuroX-MRI's textual outputs using a question-answering framework, and 2) Qualitative evaluation against open-source multimodal models using sMRI for multimodal question-answering and reasoning tasks. We will assess modality identification and demographic profiling using text datasets created from questionnaires in our datasets.

155k node-hours are assigned to ***Milestone 1.2***, including (1) Pre-training and Scaling of LLM-aligned NeuroX-MRI (153.6k), (2) Validation on neuroscience & medical QA tasks (0.6k).<u>Model of Brain for Transformative Neuroscience</u>

Table 3. Validation Framework for Milestone 1.2 (NeuroX-MRI)

Objective	Method	Key Metrics	Comparison
Quantitative Performance	Question-answering framework on established tasks (HCP sex & age prediction; HBN valence regression)	Accuracy ≥94%, MAE ≤2.2, R^2 ≥0.3	***Milestone 1.1*** models
Multimodal Capabilities	Modality identification, multimodal QA, demographic profiling, MRI-based reasoning	Text quality, prediction accuracy, explanation detail	RadFM, PMC-CLIP, Med-Flamingo

Objective 2: Develop and Scale a Multimodal EPhys Model (NeuroX-EPhys)

Goal. We will build NeuroX-EPhys, a foundation model that unifies multimodal human electrophysiology data and then aligns these signals to an LLM. Year 1 trains a 10B MoE-8 baseline using 344k node-hours and achieving unprecedented results on various downstream tasks. Year 2 expands this to 50B parameters with a 8-expert MoE using 305k node-hours, adding 50 TB iEEG and modality-aware architectures. 38k node-hours supports brain-LLM alignment through both direct neural-text methods and video-mediated approaches (**Table 4**).

Rationale. Despite steady progress from linear regression to foundation models [38, 39], today's electrophysiology frameworks still suffer four limitations: (i) sub-optimal spatio-temporal modeling, (ii) insufficient scaling, (iii) lack of genuine multimodal fusion, and (iv) absence of naturalistic behavior alignment beyond task-specific fine-tuning. Resolving these issues will enable meaningful clinical translation and breakthroughs in real-time neural decoding.

Readiness. We are well-prepared to undertake this challenge. Our team has already addressed the critical limitation of existing EEG spatio-temporal modeling by developing **DIVER-0, a fully channel-permutation-equivariant model**. DIVER-0 [10] **(Figure 4**) outperformed the current SOTA [7] on the emotion recognition benchmark [40] by +4.1 pp with only 10% of the pre-training data. This successful de-risking demonstrates our readiness to tackle the remaining limitations of scale, fusion, and behavioral alignment in the milestones outlined below.

Table 4. Milestones of Objective 2

Milestone	Description	Node-hours	Key Evaluation Task
M2.1 (Y1)	EEG data 10 B NeuroX-EPhys v1 (8-expert MoE)	122k	5% performance increase in various EEG tasks
M2.2 (Y1-Y2)	Ephys data (EEG,iEEG,MEG) scale → 50 B (8-expert MoE)	527k	Cross-modal performance gains
M2.3 (Y2)	Brain–LLM alignment (contrastive)	38k	Cross-site Generalizability Multitask learning

Milestone 2.1 (Year 1): Scale DIVER Following Established Scaling Laws in EEG

The field of EPhys foundation models is currently hampered by significant under-scaling, despite consistent evidence that scaling improves performance [39, 41, 42]. To date, the largest models have utilized only 1 TB of data [7], while compute has been severely limited to efforts such as hundreds of node-hours [41, 43].

We will directly address this gap by systematically scaling DIVER [10], our novel channel-invariant transformer architecture (**Figure 4**), in a phased campaign. First, we will perform architecture optimization using computationally efficient, small-scale models to identify optimal configurations, leveraging Mixture-of-Experts (MoE) for capacity and μTransfer to inform larger runs. Following this, we will undertake a progressive scaling of the model to 1B, 4B, and finally 10B parameters. This multi-step approach ensures that performance gains are consistent and allows us to identify the optimal model scale for various downstream applications.

Success for this milestone is defined by a median performance gain of over 5% compared to current state-of-the-art models [7] across a suite of standardized EEG tasks. To ensure fair and reproducible evaluation, all comparisons will be conducted using the torcheeg library [44], and we will contribute our new benchmarks back to the community.

122k node-hours are assigned to ***Milestone 2.1***, including (1) Architecture Exploration (10.9k), (2) Pre-training and Scaling (107.9k), and (3) Validation (2.5K).

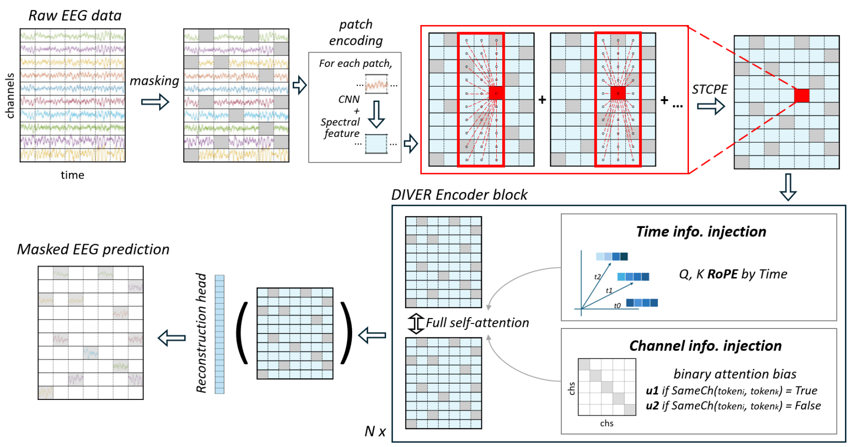

Figure 4: Overview of architecture and pretraining of our NeuroX-Ephys Model. Novel architecture of our NeuroX-Ephys mode [13]. It first converts EEG signals into patches, incorporating a novel Sliding Temporal Conditional Positional Embedding (STCPE) to provide positional context. These patched embeddings are then processed by transformer encoder blocks that use Rotary Position Embeddings (RoPE) [55] and binary attention biases to capture fine-grained spatio-temporal dynamics [50, 56].

Milestone 2.2 (Year 1~2): Achieve True Multimodal EPhys Integration

Current EPhys AI models lack true multimodal integration, treating EEG and iEEG identically without modality awareness despite evidence that joint training enhances performance [41], [45]. We will create **the first truly multimodal EPhys foundation model** by making our model architecturally modality-aware, including modality embedding (EEG, iEEG, MEG) and if needed, sub-embedding (grid, strip, depth). We will search for optimal multimodal strategies of 50M-parameter model with MoE, equipped with μTransfer. Scaling our Ephys transformer to 50B parameters using MoE 8x architecture [46, 47], this approach enables the model to learn from complementary strengths of each Ephys modality. Success will be measured by demonstrating that models trained on mixed modalities (iEEG 1 TB + EEG 1 TB) achieve statistically significant superior performance on both EEG and iEEG downstream tasks compared to single-modality models trained on equivalent total data (EEG 2 TB for EEG tasks, iEEG 2 TB for iEEG tasks). Effectiveness will be shown by cross-modal performance gains compared to single modality settings.

527k node-hours are assigned to ***Milestone 2.2*** (1) Architecture Exploration (3.8k) and (2) Pre-training and Scaling (217.5k) in year 1; (3) Pre-train Massive 50B Model (289.3k) and (4) Validation (15.8k) in year 2.

Milestone 2.3 (Year 2): Video-Mediated Brain-LLM Alignment and Instruction Tuning

To achieve true cross-task generalization, a capability lacking in current Ephys models, we will ground neural signals in the rich, semantic context of language. We will pursue this via two complementary strategies designed to create robust neural-to-text mappings.

Our primary route will be direct neural-text alignment, where we fine-tune NeuroX-Ephys using enhanced instruction-tuning techniques to map neural activity directly to corresponding textual descriptions [42, 48]. But this approach may be limited by the sparsity of time-locked annotations in clinical data.

To overcome this limitation, we will pioneer an innovative video-bridge alignment strategy. This approach addresses a fundamental challenge: many complex cognitive states resist simple textual labels, but are naturally captured by video. Our strategy first learns to map iEEG signals to their concurrent, time-locked video streams using a CLIP-style contrastive learning framework [31]. This crucial step grounds the neural data in observable, real-world events. Subsequently, we connect this neurally-aligned model to a pre-existing multimodal LLM. By leveraging the LLM's understanding of video-language relationships, NeuroX-Ephys inherits the ability to infer linguistic descriptions from neural patterns, using video as a powerful semantic intermediary.

Both strategies are enabled by our unique 50 TB clinically annotated iEEG dataset, which, with 8.4 TB of concurrent video, provides orders-of-magnitude richer supervision than any public alternative [49, 50]. The final model's success will be validated by its generalizability across diverse multi-site and multi-cultural benchmarks, with performance evaluated in collaboration with neurologists in the US and Korea.

38k node-hours are assigned to ***Milestone 2.3***: (1) Architecture Exploration (3.8k), (2) Direct neural-text alignment (10.2k), (3) Indirect video bridge approach (20.5k), (4) Validation (2.5k).

Objective 3. Unified Multimodal Neuroimaging Foundation Model

Goal & Rationale. The project culminates in the construction of NeuroX-Fusion, a unified foundation model integrating our NeuroX-MRI and NeuroX-Ephys models. This approach is designed to circumvent the central bottleneck–the scarcity of paired MRI-Ephys data for end-to-end training. Our innovation is to use an LLM as a semantic "Rosetta Stone" to bridge the two models. Having already aligned both modalities to a common LLM latent space (Objectives 1 & 2), we can unify them through this shared representation. Consequently, the limited open-source paired data is not used for training but is instead reserved for rigorously evaluating the model's emergent cross-modal reasoning (**Table 5**).

Table 5. Milestones of Objective 3

Milestone		Node-hours	Key Evaluation Task
M3.1 (Y3)	Train a unified 130B-parameter model integrating MRI, EPhys, video, and language	134k	Evaluate cross-modal reasoning on a held-out, paired dataset
M3.2 (Y3)	Validate the unified model on high-impact tasks requiring comprehensive understanding across all modalities	5k	Demonstrate the unified model's performance in complex reasoning and test-time-training

Milestone 3.1 (Year 3): Implement a unification architecture via LLM latent space to fuse the ultimate MRI and EPhys LBMs and create the final 130B model

Main challenges include training a multimodal 130B MoE model with unpaired data. We will extend adapter techniques from *Objective 1 & 2*, projecting modalities into a shared semantic space using language as an anchor, then integrate LLM-aligned NeuroX-MRI and NeuroX-EPhys models using the LLM as a bridge. To ensure training stability, we will adopt a three-stage process [20, 51]: first, training only the adapter modules; second, training additional modality-specific experts within the LLM's MoE system; and finally, fine-tuning the full model using Quantized Low Rank Approximation [52, 55].

133k node-hours assigned to ***Milestone 3.1***: (1) Pre-training (133.1k) and (2) Validation (0.1k).

Risk Mitigation and Fallback Plan. We will run milestone-gated checks during unification. If the unified model shows insufficient gains, instability, or weak cross-modal reasoning, the final-year allocation pivots to independently scaling NeuroX-MRI and NeuroX-Ephys and conducting a joint meta-analysis. To further enhance risk mitigation, we will implement an incremental cross-modal integration strategy, progressively validating compatibility at partial integration stages (e.g., shared latent spaces). This adaptive approach, combined with milestone-driven evaluations, guarantees delivery of two leadership-class brain foundation models, thus de-risking the INCITE investment even if full unification proves infeasible.

Milestone 3.2 (Year 3): Validate the unified model on high-impact tasks requiring comprehensive synergy across MRI, EPhys, Video, and Language

We will test NeuroX-Fusion's cross-modal, zero-shot inference on held-out paired MRI-EPhys validation data. The model will be tasked with predicting features of one modality (e.g., EPhys dynamics) from another (e.g., structural MRI) to prove successful fusion. We will also assess its ability to generate plausible mechanistic hypotheses in response to in-silico experimental prompts, with outcomes evaluated by domain experts for scientific validity.

The pioneering nature of NeuroX-Fusion—unifying diverse MRI and electrophysiology within an LLM framework—necessitates a new benchmark for evaluating its reasoning capabilities. We will therefore create a purpose-built dataset using text-based prompts to assess performance on key tasks, including: 1) identifying data modalities, 2) detecting anomalies by integrating brain, behavioral, and demographic data, 3) profiling individual clinical risk, 4) decoding actions and emotions from brain signals, and 5) contextualizing brain data with scientific literature.

5k node-hours are assigned to ***Milestone 3.2*** (1) Evaluate NeuroX-Fusion on each task (1.5k), and (2) Explore potential of NeuroX-Fusion on Test-Time-Training (3.5k).

3 TECHNICAL ASSESSMENT

3.1 USE OF RESOURCES REQUESTED. Building foundation models at this scale is a grand computational challenge fundamentally tied to DOE's Leadership Computing Facilities. This section justifies our 1.27M node-hour request on Aurora by providing a scientific rationale for our target model scale, leveraging lessons from our prior ALCC award to de-risk the campaign, and detailing the computational runs for each milestone.

Table 6. Resource Usage Plan

Objective	Mile-stone	Yr	Run description (a, b, and c denote the run types)	Node-hours calculation[1]				# of data tokens[3]	Node-hours
				Nodes	Hours	Epochs	Runs[2]		
1. NeuroX -MRI	M1.1	1	(a) Architecture Exploration (70.4k)	16	[1, 1.5]	40	[25, 20] × 2	4.7B	288k
			(b) Pre-training & scaling (204.8k)	512	1	40	5 x 2	11.9B	
			(c) Evaluation (12.8k)	4	0.4	40	100 x 2	200M	
	M1.2	2	(a) Pre-train LLM-aligned model (153.6k)	512	1.5	20	10	11.9B	155k
			(b) Evaluation (0.6k)	4	0.5	1	300	85.9M	
2. NeuroX -EPhys	M2.1	1	(a) Search configurations (10.9k)	16	0.34	15	100	1.3B	122k
			(b) Pretrain 1B, 4B, 10B model (107.9k)	[128, 256, 512]	[1, 3.1, 5.8]	[20, 15, 10]	[5, 3, 3]	13.1B	
			(c) Evaluation (2.5k)	1	0.5	50	50	1.1M	
	M2.2	1	(a) Search multimodality strategy (3.8k)	16	0.8	15	20	2.7B	527k
			(b) Pre-training & scaling (217.5k)	[128, 256, 512]	[2.1, 6.5, 11.3]	[20, 15, 10]	[5, 3, 2]	27.1B	
		2	(b) Pretrain 50B model (289.3k)	512	56.5	10	1	27.1B	
			(c) Evaluation (15.8k)	1	3.5	30	5 x 30	159M	
	M2.3	2	(a) Search Alignment Strategy (3.8k)	128	0.3	10	10	13M	38k
			(b) Align 50B LBM-LLM (10.2k)	512	1	10	2	13M	
			(c) Align LBM-Video / Video-LLM (20.5k)	512	1	10	2 x 2	14.8M	
			(d) Evaluation (2.5k)	1	1	50	5 x 10	60M	
3. NeuroX	M3.1	3	(a) Pretrain modality encoders (51.2k)	256	5	20	2	39B	134k

-Fusion			(b) Pretrain LLM's modality experts (56.3k)	256	5.5	20	2	39B	
			(c) Fine-tune LLM with QLoRA (25.6k)	256	2.5	20	2	39B	
			(d) Evaluation (0.1k)	1	0.5	1	200	145.9M	
	M3.2	3	(a) Complex reasoning tasks (1.5k)	1	1	1	300 × 5	145.9M	5k
			(b) Test-Time-Training tasks (3.5k)	1	1	1	700 × 5	14.6M	
Total									**1.269M**

1. Node-hour Calculation: Estimates are derived from FMRI transformer (Aurora) and Ephys transformer (Polaris) benchmarks. Numbers in brackets [x, y, z] denote different model-size experiments. Calculations account for a) a 2:1 Polaris-to-Aurora conversion factor, b) a 1.7x training time increase for MoE heads, and c) a 1.6× performance gain from multi-node parallelization.
2. Runs: This column refers to (repetitions) × (number of models, methods, or tasks, as applicable).
3. Data Tokens: Token counts are modality-specific. For MRI, one token equals a 16^3 voxel block; for Ephys, one channel-second. Video tokens are 16^2 patches. Ephys-text tokens are estimated by multiplying text pairs by 10 (word-to-pair ratio).

Our readiness based on ALCC project experience Our readiness for this INCITE campaign is grounded in concrete outcomes and lessons learned from our "NeuroX" ALCC award, which significantly mitigates key risks for this project:

1. **Validation of Scaling Laws:** Our prior work confirmed that our foundational MRI model, SwiFT V2, adheres to neural scaling laws. This provides a predictable, empirical basis for our projections and justifies the scientific value of scaling to larger parameter counts.
2. **Identification of HPC Bottlenecks:** Through scaling tests on Aurora and Frontier, we learned that at the multi-hundred node scale, I/O and data loading become primary bottlenecks, not just raw computation. This critical insight directly informs our strategic reliance on Aurora's DAOS file system and tools like Copper for this proposal.
3. **Architectural and Code Readiness:** We have successfully developed, benchmarked, and scaled our core architectures (SwiFT V2, DIVER-0) on LCFs. Our Aurora-optimized codebase, which utilizes DeepSpeed and oneCCL, has already demonstrated stable training and high performance, enabling production-level runs immediately in Year 1.

This experience informs our scientifically-driven target of a **~130B-parameter** unified model, a scale we hypothesize is necessary to integrate our specialized ~50-90B LBMs and enable emergent, cross-modal reasoning. Our detailed computational plan classifies the required runs into three types: (a) small to medium-scale architecture exploration, (b) large-scale pre-training informed by μTransfer, and (c) fine-tuning for validation.

Our resource usage is strategically distributed, with a peak usage in **Year 2 (697k node-hours)** for scaling the LBMs to 50B+ parameters and integrating new data modalities. This plan is fundamentally tied to **Aurora's unique capabilities**. The immense scale of our models and data (≈450 TB) requires the aggregated memory and computational power unique to LCFs. Aurora's high-performance Slingshot-11 interconnect and DAOS file system are essential for heavy Mixture-of-Experts (MoE) communication and the extreme neuroimaging I/O—a requirement our prior benchmarks confirmed is critical for success.

Table 7. NeuroX Data Summary by Modality

Modality	Total Subjects	Required Storage (TB)
Structural MRI (T1/T2)	83,101	1.7
Diffusion MRI	83,101	128.9
Functional MRI - resting state	56,317	120.8
Functional MRI - active state (e.g., tasks, movies)	21,421	50.1
Electrophysiology	40,048	139.6
Video (Ecog-linked)	100	8.4
Total		**449.5**

Data requirements. Our campaign requires a total storage size of 600 TB: scratch peaks at 600 TB in Y3, including 150 TB of model checkpoints. Our campaign will stage ≈450 TB of data - about 302 TB for ***Milestone 1*** (NeuroX-MRI), 148 TB for ***Milestone 2*** (NeuroX-EPhys), and 150 TB of model checkpoints for each milestone (50 TB x 3 = 150 TB). Scratch usage is projected to peak twice—first in Year 1 during multimodal pre-training and again in Year 3 when merging into NeuroX-Fusion—with a short-lived maximum up to ≈600 TB (**Table 7**).

3.2 COMPUTATIONAL APPROACH. Our computational framework leverages Aurora-optimized exascale AI software. It integrates Python and PyTorch, parallelized via Microsoft's DeepSpeed, MONAI for healthcare imaging, and Intel XPU-optimized PyTorch for large-scale training. Our parallel programming model relies on Aurora's MPICH and the oneAPI Collective Communications Library (oneCCL), with PBS scripts configured to bind MPI ranks to CPU cores for optimal performance. Our workflow comprises four main stages:

(1) **Data Preprocessing and Harmonization:** Data will be transferred to the Lustre Flare via Globus, and then moved to DAOS storage system on Aurora. MRI data has been preprocessed using established Human Connectome Project protocols. Electrophysiological data has been preprocessed using validated protocols [53, 54]. Behavioral annotations for participants' video will undergo automated or manual quality inspections. To ensure data standardization and provenance tracking across across our 450 TB dataset, we have adapted and used the data ingestion and management microservices originally developed for brainlife.io [17], which are compliant with community standards.

    Merging multimodal datasets from many sites introduces scanner- and site-specific batch effects that can bias the model. To ensure model generalizability, we will deploy a data harmonization pipeline. For MRI, we will use a Combat harmonization method [58] to remove site-specific variance while preserving biological variability.

(2) **HPC Optimization and Model Training:** We have already implemented DeepSpeed's ZeRO and lower-precision arithmetic (BF16, FP16), alongside *μ*Transfer for optimized hyperparameter tuning. Our implementation of ***μ*Transfer** significantly reduces the compute budget required to find optimal settings for large models (**Figure 5**). Performance tracking utilizes Aurora's *xpu-smi* for GPU utilization and *unitrace* for detailed profiling.

(3) **Model Evaluation and Interpretation:** Post-training validation of models will target milestone-specific downstream tasks. Model interpretability will leverage PCA/UMAP [55] for embedding visualization and Explainable AI (XAI) techniques (Integrated Gradients [56] and SmoothGrad [57] from the Captum library [58]).

(4) **Deployment and Dissemination:** All models and key findings will be version-controlled and openly shared via GitHub and Hugging Face, consistent with our ALCC-established practices. The final models will be also packaged as 'Apps' in the brainlife.io platform [17], allowing domain researchers to easily apply our models to their own data through a web-based interface, lowering the barrier to entry for exascale AI in neuroscience.

Workflow Automation and Experiment Management. To manage our complex, multi-stage campaign, we use a workflow solution built on bash PBS scripts for systematic experiment automation, combined with an AI developer platform (e.g., Neptune) for comprehensive logging, visualization, and real-time monitoring.

I/O and Data Management Strategy. For terabyte-scale data, we use NIfTI (MRI), .cnt/.edt (EPhys), and .mp4 (video) with JSON/CSV metadata. From 512-node SwiFT V2 tests we pinpointed Lustre read bottlenecks and, in consultation with ALCF, we plan to use DAOS. Structured data access employs HDF5, while Copper accelerates Python library loading. All model checkpoints are stored on DAOS to ensure throughput efficiency and resilience.

3.3 PARALLEL PERFORMANCE. Our readiness for an INCITE-scale campaign is demonstrated by extensive strong scaling benchmarks of our foundational 4.1B-parameter SwiFT V2 model. As shown in **Figure 5**, we evaluated our code on representative 4D-fMRI data using an optimized stack (DeepSpeed ZeRO-2, BF16) on both ALCF's Aurora and OLCF's Frontier. The results validate our code's exceptional performance and scalability, particularly on our primary target system, Aurora. Key findings include:

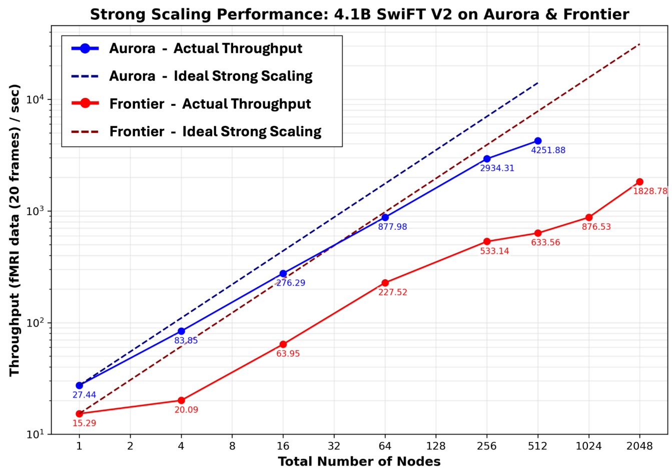

Figure 5. Strong Scaling performance of SwiFT V2 on Frontier and Aurora. We tested a scaling efficiency our 4.1B fMRI transformer model (SwiFT V2) [15]. Using 1 node as a baseline, our model showed 30.8% of strong scaling efficiency on Aurora.

- **Superior Throughput on Aurora:** On 512 nodes, we achieved a throughput of 4,252 samples/sec on Aurora, a 6.7x throughput improvement over an equivalent run on Frontier.
- **Architectural Optimization:** This demonstrates a 4.47x superior performance per GPU rank, confirming our code's high level of optimization for the Aurora architecture.

These benchmarks, which incorporate our full production I/O workflow, show robust, near-ideal strong scaling up to 512 nodes (6,144 GPU-ranks). This performance provides a confident baseline for our node-hour calculations and confirms our ability to efficiently leverage larger node counts to meet the demanding timelines of this research campaign.

3.4 DEVELOPMENTAL WORK. This project is de-risked by a prior ALCC project: our backbone models—fMRI [9,15] and ephys transformers [10]—are developed and validated. We have built and benchmarked a robust, Aurora-optimized PyTorch-DeepSpeed codebase, successfully scaling a 4.1B-parameter model and demonstrating strong-scaling efficiency on up to 512 Aurora nodes. This proven readiness allows our INCITE effort to focus on the following targeted, novel developments. Our proposed campaign will execute three core developmental tasks:

- **Mixture-of-Experts (MoE) Integration and Scaling**: In Year 1, our primary development will be implementing and optimizing MoE layers within the fMRI transformer (SwiFT V2) and ephys transformer (DIVER-0) backbones. We will begin with an 8-expert Swin-Transformer MoE and scale to a high-complexity architecture for the unified model in Year 3.
- **Advanced Multimodal Alignment Modules:** We will develop a suite of alignment frameworks. This includes contrastive learning modules to fuse MRI modalities (Year 1), cross-attention adapters for deep, LLM-centric integration (Year 2), and a novel CLIP-style framework to align EPhys signals with large-scale behavioral video data (Year 2).
- **The Unification Architecture:** The capstone developmental effort in Year 3 will be to design and implement the unification framework that fuses the NeuroX-MRI and NeuroX-EPhys models. This will involve creating and validating a bridging mechanism based on the shared LLM latent space or a hierarchical MoE router.

This developmental work is tightly integrated into the exploration and pre-training phases of our milestones. Each component will be validated by its ability to sustain stable, scalable training on Aurora and by the modes' performance on their corresponding scientific evaluation tasks.

REFERENCES

[1] J. Campion *et al.*, “Public mental health: required actions to address implementation failure in the context of COVID-19,” *Lancet Psychiatry*, vol. 9, no. 2, pp. 169–182, Feb. 2022, doi: 10.1016/S2215-0366(21)00199-1.

[2] L. Cocchi, L. L. Gollo, A. Zalesky, and M. Breakspear, “Criticality in the brain: A synthesis of neurobiology, models and cognition,” *Prog Neurobiol*, vol. 158, pp. 132–152, Nov. 2017, doi: 10.1016/j.pneurobio.2017.07.002.

[3] S. Marek *et al.*, “Reproducible brain-wide association studies require thousands of individuals,” *Nature*, vol. 603, no. 7902, pp. 654–660, Mar. 2022, doi: 10.1038/s41586-022-04492-9.

[4] X. Wang *et al.*, “ORBIT-2: Scaling exascale vision foundation models for weather and climate downscaling,” *arXiv [cs.LG]*, May 07, 2025. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2505.04802

[5] H. Cui *et al.*, “scGPT: toward building a foundation model for single-cell multi-omics using generative AI,” *Nat Methods*, vol. 21, no. 8, pp. 1470–1480, Aug. 2024, doi: 10.1038/s41592-024-02201-0.

[6] S. Wang *et al.*, “Triad: Vision foundation model for 3D magnetic resonance imaging,” *arXiv [cs.CV]*, Feb. 19, 2025. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2502.14064

[7] J. Wang *et al.*, “CBraMod: A Criss-cross brain foundation model for EEG decoding,” *arXiv [eess.SP]*, Dec. 10, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2412.07236

[8] X. Zhou *et al.*, “Brain foundation models: A survey on advancements in neural signal processing and brain discovery,” *arXiv [cs.LG]*, 2025. doi: 10.48550/ARXIV.2503.00580.

[9] P. Y. Kim *et al.*, “SwiFT: Swin 4D fMRI Transformer,” presented at the Advances in Neural Information Processing Systems, in 36. 2023. doi: 10.48550/ARXIV.2307.05916.

[10] D. D. Han, A. L. Lee, S. Lee, J. Cha, “DIVER-0: A fully channel equivariant EEG foundation model,” in *In Proceedings of the ICLR GenBio Workshop 2025 (Spotlight)*, 2025.

[11] R. Stevens, V. Taylor, J. Nichols, A. Maccabe, K. Yelick, and D. Brown, “AI for science: Report on the department of energy (DOE) town halls on artificial intelligence (AI) for science,” Argonne National Laboratory (ANL), Feb. 2020. doi: 10.2172/1604756.

[12] C. Sudlow *et al.*, “UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age,” *PLoS Med*, vol. 12, no. 3, p. e1001779, Mar. 2015, doi: 10.1371/journal.pmed.1001779.

[13] J. S. Elam *et al.*, “The Human Connectome Project: A retrospective,” *Neuroimage*, vol. 244, p. 118543, Dec. 2021, doi: 10.1016/j.neuroimage.2021.118543.

[14] B. J. Casey *et al.*, “The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites,” *Dev Cogn Neurosci*, vol. 32, pp. 43–54, Aug. 2018, doi: 10.1016/j.dcn.2018.03.001.

[15] Choi, J., Wang, H., Kwon, J., Yoo, S., & Cha, J, “SwiFT V2: Towards large-scale foundation model for functional MRI. Proceedings of the 2025 Conference on Computational and Cognitive Neuroscience,” presented at the Cognitive Computational Neurosicence, 2025.

[16] J. Kaplan *et al.*, “Scaling laws for neural language models,” *arXiv [cs.LG]*, Jan. 22, 2020. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/2001.08361

[17] S. Hayashi *et al.*, “Brainlife.Io: A decentralized and open-source cloud platform to support neuroscience research,” *Nat. Methods*, vol. 21, no. 5, pp. 809–813, May 2024, doi: 10.1038/s41592-024-02237-2.

[18] C. Hwang *et al.*, “Tutel: Adaptive mixture-of-experts at scale,” *arXiv [cs.DC]*, Jun. 07, 2022. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/2206.03382

[19] D. Dai *et al.*, “DeepSeekMoE: Towards ultimate expert specialization in Mixture-of-experts

language models," *arXiv [cs.CL]*, Jan. 11, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2401.06066
[20] Y. Li *et al.*, "Uni-MoE: Scaling unified multimodal LLMs with Mixture of experts," *arXiv [cs.AI]*, May 18, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2405.11273
[21] J. O. Caro *et al.*, "BrainLM: A foundation model for brain activity recordings," *bioRxiv*, Sep. 13, 2023. doi: 10.1101/2023.09.12.557460.
[22] Z. Dong *et al.*, "Brain-JEPA: Brain dynamics foundation model with Gradient Positioning and Spatiotemporal Masking," *arXiv [q-bio.NC]*, Sep. 28, 2024. [Online]. Available: http://arxiv.org/abs/2409.19407
[23] Y. Sun, L. Wang, G. Li, W. Lin, and L. Wang, "A foundation model for enhancing magnetic resonance images and downstream segmentation, registration and diagnostic tasks," *Nat. Biomed. Eng.*, vol. 9, no. 4, pp. 521–538, Apr. 2025, doi: 10.1038/s41551-024-01283-7.
[24] L. Wu, J. Zhuang, and H. Chen, "VoCo: A simple-yet-effective volume contrastive learning framework for 3D medical image analysis," *arXiv [eess.IV]*, Feb. 27, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2402.17300
[25] G. Qu, Z. Zhou, V. D. Calhoun, A. Zhang, and Y.-P. Wang, "Integrated brain connectivity analysis with fMRI, DTI, and sMRI powered by interpretable graph neural networks," *arXiv [q-bio.NC]*, Aug. 26, 2024. doi: 10.1016/j.media.2025.103570.
[26] Z. Liu *et al.*, "Swin transformer: Hierarchical vision transformer using shifted windows," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2021, pp. 9992–10002. doi: 10.1109/iccv48922.2021.00986.
[27] G. Yang *et al.*, "Tuning Large Neural Networks via Zero-Shot Hyperparameter Transfer," *Advances in Neural Information Processing Systems*, vol. 34, pp. 17084–17097, Dec. 2021, Accessed: Jun. 13, 2025. [Online]. Available: https://proceedings.neurips.cc/paper/2021/hash/8df7c2e3c3c3be098ef7b382bd2c37ba-Abstract.html
[28] J. Rasley, S. Rajbhandari, O. Ruwase, and Y. He, "DeepSpeed," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, New York, NY, USA: ACM, Aug. 2020. doi: 10.1145/3394486.3406703.
[29] D. Kalamkar *et al.*, "A study of BFLOAT16 for Deep Learning training," *arXiv [cs.LG]*, 2019. doi: 10.48550/ARXIV.1905.12322.
[30] T. Dao, "FlashAttention-2: Faster attention with better parallelism and work partitioning," *arXiv [cs.LG]*, 2023. doi: 10.48550/ARXIV.2307.08691.
[31] A. Radford *et al.*, "Learning transferable visual models from natural language supervision," *ICML*, vol. 139, pp. 8748–8763, Feb. 2021, [Online]. Available: https://proceedings.mlr.press/v139/radford21a.html
[32] F. Chen *et al.*, "X-LLM: Bootstrapping advanced large language models by treating multi-modalities as foreign languages," *arXiv [cs.CL]*, 2023. doi: 10.48550/ARXIV.2305.04160.
[33] J. Li, D. Li, S. Savarese, and S. Hoi, "BLIP-2: Bootstrapping language-image pre-training with frozen image encoders and large language models," *arXiv [cs.CV]*, 2023. doi: 10.48550/ARXIV.2301.12597.
[34] A. Panagopoulou *et al.*, "X-InstructBLIP: A Framework for aligning X-Modal instruction-aware representations to LLMs and Emergent Cross-modal Reasoning," *arXiv [cs.CV]*, 2023. doi: 10.48550/ARXIV.2311.18799.
[35] B. Lin *et al.*, "MoE-LLaVA: Mixture of experts for Large Vision-Language Models," *arXiv [cs.CV]*, Jan. 29, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2401.15947
[36] S. Rajbhandari *et al.*, "DeepSpeed-MoE: Advancing Mixture-of-Experts inference and training to power next-generation AI scale," *arXiv [cs.LG]*, Jan. 14, 2022. Accessed: Jun.

13, 2025. [Online]. Available: http://arxiv.org/abs/2201.05596
[37] W. Fedus, B. Zoph, and N. Shazeer, “Switch transformers: Scaling to trillion parameter models with simple and efficient sparsity,” *arXiv [cs.LG]*, Jan. 11, 2021. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2101.03961
[38] D. Kostas, S. Aroca-Ouellette, and F. Rudzicz, “BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data,” *Front. Hum. Neurosci.*, vol. 15, p. 653659, Jun. 2021, doi: 10.3389/fnhum.2021.653659.
[39] W. Jiang, L. Zhao, and B.-L. Lu, “Large Brain Model for Learning Generic Representations with Tremendous EEG Data in BCI,” in *The Twelfth International Conference on Learning Representations*, Oct. 2023. Accessed: Jun. 13, 2025. [Online]. Available: https://openreview.net/forum?id=QzTpTRVtrP
[40] J. Chen, X. Wang, C. Huang, X. Hu, X. Shen, and D. Zhang, “A large Finer-grained Affective Computing EEG dataset,” *Sci. Data*, vol. 10, no. 1, p. 740, Oct. 2023, doi: 10.1038/s41597-023-02650-w.
[41] Z. Yuan, F. Shen, M. Li, Y. Yu, C. Tan, and Y. Yang, “BrainWave: A brain signal foundation model for clinical applications,” *arXiv [q-bio.NC]*, Feb. 15, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2402.10251
[42] G. Wang, W. Liu, Y. He, C. Xu, L. Ma, and H. Li, “EEGPT: Pretrained transformer for universal and reliable representation of EEG signals,” *Neural Inf Process Syst*, vol. 37, pp. 39249–39280, 2024, [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2024/hash/4540d267eeec4e5dbd9dae9448f0b739-Abstract-Conference.html
[43] D. Zhang, Z. Yuan, Y. Yang, J. Chen, J. Wang, and Y. Li, “Brant: Foundation model for intracranial neural signal,” *Neural Inf Process Syst*, 2023, [Online]. Available: http://yangy.org/works/brainnet/NeurIPS23_Brant.pdf
[44] Z. Zhang, S.-H. Zhong, and Y. Liu, “TorchEEGEMO: A deep learning toolbox towards EEG-based emotion recognition,” *Expert Syst. Appl.*, vol. 249, no. 123550, p. 123550, Mar. 2024, doi: 10.1016/j.eswa.2024.123550.
[45] E. Shi *et al.*, “FoME: A Foundation Model for EEG using adaptive temporal-lateral attention scaling,” *arXiv [cs.LG]*, Sep. 19, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2409.12454
[46] X. Liu *et al.*, “Moirai-MoE: Empowering time series foundation models with sparse mixture of experts,” *arXiv [cs.LG]*, Oct. 14, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2410.10469
[47] X. Shi *et al.*, “Time-MoE: Billion-scale time series foundation models with mixture of experts,” *arXiv [cs.LG]*, Sep. 24, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2409.16040
[48] W.-B. Jiang, Y. Wang, B.-L. Lu, and D. Li, “NeuroLM: A universal multi-task foundation model for bridging the gap between language and EEG signals,” *arXiv [eess.SP]*, Aug. 27, 2024. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2409.00101
[49] T. C. N’dir and R. T. Schirrmeister, “EEG-CLIP : Learning EEG representations from natural language descriptions,” *arXiv [cs.CL]*, Mar. 18, 2025. Accessed: Jun. 13, 2025. [Online]. Available: http://arxiv.org/abs/2503.16531
[50] A.-K. Kiessner, R. T. Schirrmeister, L. A. W. Gemein, J. Boedecker, and T. Ball, “An extended clinical EEG dataset with 15,300 automatically labelled recordings for pathology decoding,” *NeuroImage Clin.*, vol. 39, no. 103482, p. 103482, Jul. 2023, doi: 10.1016/j.nicl.2023.103482.
[51] “The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation,” Meta AI. Accessed: Jun. 13, 2025. [Online]. Available: https://ai.meta.com/blog/llama-4-multimodal-intelligence/
[52] T. Dettmers, A. Pagnoni, A. Holtzman, and L. Zettlemoyer, “QLoRA: Efficient Finetuning of

Quantized LLMs," *arXiv [cs.LG]*, May 23, 2023. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/2305.14314
[53] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004, doi: 10.1016/j.jneumeth.2003.10.009.
[54] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen, "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data," *Comput. Intell. Neurosci.*, vol. 2011, no. 1, p. 156869, Jan. 2011, doi: 10.1155/2011/156869.
[55] L. McInnes, J. Healy, and J. Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv [stat.ML]*, Feb. 09, 2018. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/1802.03426
[56] M. Sundararajan, A. Taly, and Q. Yan, "Axiomatic attribution for deep networks," *ICML*, vol. 70, pp. 3319–3328, Mar. 2017, [Online]. Available: https://proceedings.mlr.press/v70/sundararajan17a.html
[57] D. Smilkov, N. Thorat, B. Kim, F. Viégas, and M. Wattenberg, "SmoothGrad: removing noise by adding noise," *arXiv [cs.LG]*, Jun. 12, 2017. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/1706.03825
[58] N. Kokhlikyan *et al.*, "Captum: A unified and generic model interpretability library for PyTorch," *arXiv [cs.LG]*, Sep. 16, 2020. Accessed: Jun. 16, 2025. [Online]. Available: http://arxiv.org/abs/2009.07896

PERSONNEL JUSTIFICATION AND MANAGEMENT PLAN

PERSONNEL JUSTIFICATION - TOTAL 387 PM

This project brings together a world-class, interdisciplinary team with demonstrated expertise in large-scale AI, computational neuroscience, clinical neurophysiology, and high-performance computing on DOE Leadership Computing Facilities. All key personnel are in place and have a proven track record of collaboration.

Leadership Team – 30 PM

- **Dr. Shinjae Yoo (PI)** will provide overall leadership, ensuring the coordinated execution of all milestones and driving the delivery of the final scientific outputs. He will manage the project's technical direction and interactions with ALCF (3PM over three years).

- **Prof. Jiook Cha (SNU, Co-PI)** is co-director of the NeuroX-Fusion project, and a world leading expert in large-scale computational human neuroimaging. In the project, Prof. Cha will contribute to scientific planning and coordination of the project, and the development of all models (9PM).

- **Prof. Chun Kee Chung (SNU, Co-I)** is an internationally recognized neurosurgeon-scientist who established the country's largest intracranial-EEG repository. In this project, he will provide the curated iEEG dataset, oversee its quality-control pipeline, and serve as the lead domain scientist guiding the engineering team in modelling clinical electrophysiology (3 PM).

- **Prof. Jay-Yoon Lee (SNU, Co-I)** is a leading expert in constraint-aware, energy-based neural modeling and semi-supervised learning. In the project, Prof. Lee will develop knowledge-constrained loss functions and alignment algorithms that embed anatomical priors, reducing labeled-data demands and stabilizing cross-modal training (6 PM).

- **Prof. Taesup Moon (SNU, Co-I)** is an authority on large-scale transformer optimization and efficient multimodal pre-training on exascale systems. In the project, Prof. Moon will scale the SwiFT 4D Transformer to 130B parameters and implement mixed-precision kernels and adaptive sampling to double Aurora training throughput (6 PM).

- **Prof. Franco Pestilli (UT Austin, Co-I)** is a leading expert in diffusion imaging, tractography, and open-science neuroimaging platforms. In the project, he will lead the neuroimaging data curation and sharing infrastructure, leveraging his extensive work on the BRAIN Initiative-sponsored brainlife platform and BIDS standards to ensure seamless integration of diffusion MRI data with the multimodal modeling pipeline (3 PM).

Scientific personnel – 384 PM

Dr. Shinjae Yoo (PI)

- **David Kitae Park**, PhD (Research Staff Level II, BNL), will lead the development of the NeuroX-EPhys stream (Objective 2), overseeing the multimodal integration of EEG modalities. Dr. Park will also scale up the model using MoE architecture with μTransfer, and its subsequent alignment with the LLM (12PM).

Prof. Jiook Cha (SNU, Co-PI)

- **Jubin Choi** (PhD student) will lead HPC and workflow optimization. He will manage the project's programming environments, oversee data transfer and storage logistics including the use of DAOS, and investigate and implement the optimal parallel training configurations for all deep learning experiments on Aurora (24PM).

- **Heehwan Wang** (PhD student) will lead the MRI-LLM alignment work within Objective 1 (Milestone 1.2). He will implement and train the cross-attention adapter modules required to align the 91B NeuroX-MRI model with the LLM, and will lead the validation on multimodal question-answering and reasoning tasks (24PM).

- **Danny Dongyeop Han** (PhD student) will lead the development of the NeuroX-EPhys model (Objective 2). He will implement the multimodal MoE DIVER-EPhys architecture, integrate iEEG and scalp EEG, and lead the development of the novel video-mediated LLM alignment (24PM).

- **Ahhyun Lucy Lee** (PhD student) will take the lead on Objective 3 (NeuroX-Fusion). She will focus on implementing and validating the final unification architecture that fuses the MRI and EPhys LBMs via the LLM-bridged latent space (24PM).

- **Bo-Gyeom Kim** (PhD student) will serve as downstream-task lead for the MRI stream (Objective 1). She will curate task-fMRI and clinical benchmark datasets, design large-scale fine-tuning pipelines, and evaluate NeuroX-MRI representations on cognition and affect prediction tasks (24PM).

- **Sangyoon Bae** (PhD student) will conduct research on developing biologically plausible brain foundation models. She will leverage proven expertise in neuroscience, refine brain-inspired loss functions that leverage neural dynamics, and develop innovative pre-training strategies incorporating properties that respect underlying neural circuit principles (24PM).

- **Seungju Lee** (PhD student) will perform the clinical downstream task evaluation, adapting representations from NeuroX-MRI (Objective 1), NeuroX-EPhys (Objective 2), and NeuroX-Fusion (Objective 3). Her responsibilities include curating datasets for and building predictive models of mental disorders, aiming for early diagnosis and treatment response prediction (24PM).

- **Jungwoo Seo** (PhD student) will focus on the exploration of NeuroX as a generative model for MRI and multimodal brain data. He will implement diffusion-based generation pipelines, integrate cross-modal conditioning, and evaluate the model across imaging modalities (24PM).

- **Seokjin Moon** (PhD student) will conduct the downstream predictive modeling in NeuroX-MRI (Objective 1) using clinical and affective datasets. He is developing pipelines to benchmark NeuroX representations on cognitive and affective outcomes, validating their relevance and generalizability (24PM).

- **Sebin Lee** (MS student) will conduct research on diverse downstream tasks, including resting-state to task fMRI prediction with NeuroX-MRI (Objective 1). She will extend this task into variable prediction, multimodal and MoE integration (24PM).

- **Jun-Eun Shin** (Lab coordinator) will serve as the designated Project Manager (3PM). In this critical operational role, she will be responsible for the day-to-day coordination of this complex, multi-institutional project. Her duties will include maintaining the master project schedule, tracking progress against all milestones and deliverables, facilitating communication between the BNL and SNU teams by organizing meetings and circulating agendas and minutes, and managing data and resource logistics. Ms. Shin will also directly support the PI in the preparation and submission of all INCITE quarterly and annual progress reports (24PM).

Prof. Chun Kee Chung (SNU, Co-I)

- **Seongjin Lee** (PhD student) will actively contribute to the neuroscience and signal processing components of the project. He will be primarily responsible for developing a robust, large-scale iEEG preprocessing and QA/QC pipeline to ensure data integrity and reliability. Additionally, he will provide strategic guidance on the extraction and modeling of rich metadata derived from intracranial recordings, contributing to the overall architecture and analytic design of the project (24PM).

- **Yonghyeon Gwon** (MS student) will actively contribute to the neuroscience and signal processing efforts of the team. His core responsibility will center on the development of model architectures that integrate spatial information, particularly the localization of brain regions and channel coordinates. His work will support the anatomical alignment and modeling precision required for downstream neurocomputational analysis (24PM).

Prof. Jay-Yoon Lee (SNU, Co-I)

- **Yunju Cho** (PhD student) will play a key role in the time series analysis and semi-supervised learning techniques in the project. She will be primarily responsible for designing and testing algorithms to process time series signals, e.g., EEG, for this project (12PM).

- **Hye Ryung Son** (PhD student) will play a key role in the LLM integration in the project. She will be responsible for aligning the latent space of LLM's language-based embeddings with the space of other modalities including MRI recordings and electrophysiological signals (12PM).

Prof. Taesup Moon (SNU, Co-I)

- **Juhyeon Park** (PhD student) will participate in the research on developing large-scale fMRI-based brain foundation models in NeuroX-MRI (Objective 1) utilizing efficient scaling of the existing Swin Transformer architecture and transfer-learning foundation models for natural images (12PM).

- **Yongho Kim** (PhD student) will participate in the research on developing large-scale fMRI-based brain foundation models in NeuroX-MRI (Objective 1) utilizing efficient scaling of the existing Swin Transformer architecture and transfer-learning foundation models for natural images (12PM).

Prof. Franco Pestilli (UT Austin, Co-I)

- **Junbeom Kwon** (PhD student), will contribute to the development of NeuroX-s/dMRI (Objective 1) by integrating diffusion MRI with structural MRI using contrastive learning. He will focus on building joint representations of brain macro- and micro-structure using paired dMRI-sMRI data (12PM).

We recognize that personnel turnover is a potential risk in any multi-year project. Our mitigation strategy includes:

- **Knowledge Redundancy:** Project responsibilities are shared between senior leadership (PIs/Co-Is) and the scientific personnel executing the work, ensuring no single individual is a sole point of failure for any critical milestone.
- **Rigorous Documentation:** All code, experimental configurations, and results will be meticulously documented and version-controlled on shared platforms like GitHub and Weights & Biases, facilitating smooth handovers if necessary.
- **Recruitment Pipeline:** The leadership team's strong academic and institutional affiliations (BNL, SNU, UT Austin) provide a robust pipeline of highly qualified graduate students and postdoctoral researchers, enabling us to promptly fill any vacancies that may arise.

All personnel have their research activities fully supported by grants from DOE ASCR CC064CCAA, from the National Research Foundation of Korea (NRF)—funded by the Ministry of Science and ICT (MSIT) and the Ministry of Education—including 2021R1C1C1006503, RS-2023-00266787, RS-2023-00265406, RS-2024-00421268, RS-2024-00342301, RS-2024-00435727, NRF-2021M3E5D2A01022515, and NRF-2021S1A3A2A02090597.

MANAGEMENT PLAN

Leadership Structure. Overall direction is provided by the monthly *Steering Meeting*, convened during the first week of each month and attended by the Principal Investigator (Shinjae Yoo), the a Co-PI (Jiook Cha), and the four Co-Is (Chun Kee Chung, Jay-Yoon Lee, Taesup Moon, and Franco Pestilli,). This forum reviews node-hour burn, key performance indicators (KPI), and risk status, then reallocates resources and priorities for the coming month. Day-to-day scheduling and cross-team coordination are handled by the designated *Project Manager*, Ms Jun-Eun Shin (Lab coordinator, 3 PM), who maintains the master project calendar, circulates agendas and minutes, tracks milestone timelines and dependencies, and escalates schedule risks to the Steering Committee. The PI serves as the single point of contact for the official communications but the team members can communicate freely with ALCF staff, regarding technical issues by ALCF slack, e-mail and Microsoft Teams.

Sub-teams and Research Focus.

- **MRI Team** (Leads — Jiook Cha, and Heehwan Wang) meets weekly and develops cross-attention adapters aligning the 91 B NeuroX-MRI encoder with the Large Language Model (LLM).

- **EPhys Team** (Leads — Chun Kee Chung, David Kitae Park, and Danny Dongyeop Han) meets weekly to scale the DIVER architecture from 10 B to 50 B MoE and to integrate iEEG, EEG and MEG data streams.

- **Data & Infrastructure Core** (Leads — Franco Pestilli, and Jubin Choi) meets bi-weekly to operate the DAOS pipeline, manage data transfers and maintain DeepSpeed/PBS automation.

- **Evaluation Core** (Leads — Taesup Moon, and Danny Dongyeop Han) meets bi-weekly to maintain modality-specific leaderboards, normalise metrics and generate meta-analysis reports.

- **Unification Team** (Leads — Jay-Yoon Lee, and Ahhyun Lucy Lee) conducts quarterly design workshops during Years 1–2; beginning Year 3, it moves to a weekly cadence to implement and validate the LLM-mediated fusion of MRI and EPhys models.

Each sub-team files a one-page *Cross-Stream Report* every two weeks, summarizing progress, blockers and upcoming resource needs in its dedicated Teams channel and tagging open issues in GitHub.

Decision-making Process. Resource requests and blockers surfaced in weekly sub-team meetings are logged as Microsoft Teams Planner cards. GPU- and node-allocation conflicts are resolved in a bi-weekly *Tech-Lead Sync* (PI, Co-Is, sub-team leads). Strategic changes—such as model-size jumps or new-dataset ingestion—require formal approval in the monthly Steering Meeting before implementation.

Inter-team Synergy. MRI and EPhys streams advance in parallel on tasks optimised for their respective spatial and temporal strengths. Evaluation-Core meta-analysis normalises performance and cost metrics across modalities, producing a quarterly *Complementarity Map* that highlights where each stream excels. The Unification Team uses this map to design LLM routing weights and MoE expert allocation, ensuring that the final 130 B model in Year 3 combines the best attributes of both streams rather than duplicating effort. This staged, feedback-driven workflow keeps the two streams loosely coupled for risk isolation yet strategically convergent for final integration.

Progress Tracking and Reporting. Development is tracked through GitHub Projects kanban boards and a live KPI dashboard embedded in Microsoft Teams. Formal *Milestone Reviews* occur each quarter, aligning checkpoint releases and manuscript plans with INCITE deliverables. Publications, awards and code releases are reported quarterly to the INCITE portal by the Project Science Communicator (Heehwan Wang) or, if unavailable, by the alternate contact (Ahhyun Lucy Lee).

Risk Management. Persistent MoE instability triggers an immediate switch to a dense fallback model trained in parallel. Delays in data-licence clearance activate a contingency pipeline built on public datasets (e.g., UKB/HCP/ABCD). Pre-assigned deputy leads ensure seamless hand-off when key personnel are unavailable.

Collaboration Tools. Microsoft Teams (Steering, MRI, EPhys, Data-Infra, Evaluation channels) provides real-time communication, while OneDrive and Teams Wiki host shared documentation. Code and experiment logs are managed through GitHub pull requests and issues. Emergency incidents are escalated via slack to Aurora Ops and an SMS phone tree covering all senior personnel.

This management structure satisfies INCITE requirements by defining clear leadership, transparent effort allocation, well-differentiated research foci, explicit inter-team integration paths, and robust mechanisms for progress tracking and risk mitigation.

Milestone Table

Proposal Title: NeuroX-Fusion: Unified Foundation Model of Brain for Transformative Neuroscience

Year 1 **Total number of node-hours for Year 1: 632k**		
Milestone:	**Details (as appropriate):**	**Dates:**
Milestone 1.1 57B NeuroX-f/sMRI and 34B NeuroX-s/dMRI with 8 experts MoE	**Resource:** Aurora **Node-hours:** 288k **Production size runs (number of nodes):** Exploration runs on 16-64 nodes; pre-training runs up to 512 nodes **Filesystem storage (TB and dates):** 352 TB on Lustre, DAOS for MRI data (Jan 2026 - Dec 2028) **Archival storage (TB and dates):** None **Software Application:** PyTorch, DeepSpeed, MONAI **Tasks:** (1) Architecture Exploration (70.4k), (2) Pre-training and Scaling (204.8k), and (3) Validation on downstream tasks (12.8k). **Dependencies:** None	*01/01/26 – 12/31/26*
Milestone 2.1 Scale DIVER Following Established Scaling Laws in EEG	**Resource:** Aurora **Node-hours:** 122k **Production size runs (number of nodes):** Exploration at 16 nodes; Scaling runs up to 512 nodes. **Filesystem storage (TB and dates):** 189.6 TB on Lustre, DAOS for EPhys data and model checkpoints (Jan 2026 - Dec 2028) **Archival storage (TB and dates):** None **Software Application:** PyTorch, MNE, DeepSpeed **Tasks:** (1) Architecture Exploration (10.9k), (2) Pre-training and Scaling (107.9k), and (3) Validation (2.5K), (1), (2) - 10 B NeuroX-EPhys v1 (8-expert MoE) / (3) - 5% performance increase in various EEG tasks **Dependencies:** None	*01/01/26 – 08/31/26*
Milestone 2.2 Achieve True Multimodal EPhys Integration (continued in Year 2)	**Resource:** Aurora **Node-hours:** 527k (222k in Year 1, 305k in Year 2) **Production size runs (number of nodes):** Exploration - 16 nodes; Large scaling/alignment runs - 512 nodes **Filesystem storage (TB and dates):** Utilize existing 189.6 TB space for EPhys data and model checkpoints **Archival storage (TB and dates):** None **Software Application:** PyTorch, DeepSpeed, MNE **Tasks:** (1) Architecture Exploration (3.8k) and (2) Pre-training and Scaling (217.5k) in year	*07/01/26 – 12/31/27*

<table>
<tr><td></td><td>1;
(3) Pre-train Massive 50B Model (289.3k) and (4) Validation (15.8k) in year 2.
Dependencies: Milestone 2.1</td><td></td></tr>
<tr><td colspan="3">Year 2 (if appropriate) Total number of node-hours for Year 2: 498k</td></tr>
<tr><td>Milestone 1.2
57B NeuroX-f/sMRI and
34B NeuroX-s/dMRI with 8 experts MoE</td><td>Resource: Aurora Node-hours: 155k
Production size runs (number of nodes): Exploration runs on 16-64 nodes; pre-training runs up to 512 nodes
Filesystem storage (TB and dates): Utilize existing 352 TB for MRI data and model checkpoints
Archival storage (TB and dates): None
Software Application: PyTorch, DeepSpeed, MONAI
Tasks: (1) Pre-training and Scaling of LLM-aligned NeuroX-MRI (153.6k),
(2) Validation on neuroscience & medical QA tasks (0.6k).
Dependencies: Milestone 1.1</td><td>01/01/27 – 12/31/27</td></tr>
<tr><td>Milestone 2.3
Video-Mediated Brain-LLM Alignment and Instruction Tuning</td><td>Resource: Aurora Node-hours: 38k
Production size runs (number of nodes): Exploration - 128 nodes; Large scaling/alignment runs - 512 nodes
Filesystem storage (TB and dates): 198 TB - add 8.4 TB of video data to existing 189.6 TB space (Jan 2027 ~)
Archival storage (TB and dates): None
Software Application: PyTorch, DeepSpeed, MNE, Huggingface
Tasks: (1) Architecture Exploration (3.8k), (2) Direct neural-text alignment (10.2k),
(3) Indirect video bridge approach (20.5k), and (4) Validation (2.5k).
Dependencies: Milestone 2.2</td><td>01/01/27 – 12/31/27</td></tr>
<tr><td colspan="3">Year 3 (if appropriate) Total number of node-hours for Year 3: 139k</td></tr>
<tr><td>Milestone 3.1
Implement a unification architecture via LLM latent space to fuse the ultimate MRI and EPhys LBMs and create the final 130B</td><td>Resource: Aurora Node-hours: 134k
Production size runs (number of nodes): Up to 256 nodes for multi-stage unification training.
Filesystem storage (TB and dates): 600 TB (add 50 TB of model checkpoints to existing 550 TB, Jan 2028 ~)
Archival storage (TB and dates): None
Software Application: PyTorch, DeepSpeed, MONAI</td><td>01/01/28 – 10/31/28</td></tr>
</table>

model	**Tasks:** (1) Pre-train only adapter modules of NeuroX-MRI and NeuroX-EPhys (51.2k), (2) Train modality-specific experts within the LLM's MoE system (56.3k), (3) Fine-tune only the LLM module with QLoRA (25.6k), (4) Compare 130B NeuroX-Fusion with LLM aligned NeuroX-MRI and NeuroX-EPhys (0.1k) **Dependencies:** Milestone 1.2, Milestone 2.3,	
Milestone 3.2 Validate the unified model on high-impact tasks requiring comprehensive synergy across MRI, EPhys, Video, and Language	**Resource:** Aurora **Node-hours:** 5k **Production size runs (number of nodes):** Smaller scale ensemble inference and Test-Time-Training runs. **Filesystem storage (TB and dates):** 600 TB of existing data and checkpoints **Archival storage (TB and dates):** None **Software Application:** PyTorch, DeepSpeed, MNE, Huggingface **Tasks:** (1) Compare NeuroX-Fusion's performance to existing LLMs(1.5k), (2) Explore its Test-Time-Training ability(3.5k) **Dependencies:** Milestone 3.1	*07/01/28 – 12/31/28*

Curriculum Vitae

PI NAME: Shinjae Yoo, PhD

PROFESSIONAL PREPARATION

- Carnegie Mellon University, Pittsburgh, PA
  - Ph.D. in Language Technologies, School of Computer Science, 2010
  - Masters in Language Technologies, School of Computer Science, 2005
- Seoul National University, Seoul, Korea
  - M.S. in Computer Science, 2002
- Soongsil University, Seoul, Korea
  - B.S. in Computer Science, School of Computing, 2000

APPOINTMENTS

- 2025 - *(Interim)* Division Lead, Computational Research, Brookhaven National Lab.
- 2025 - Distinguished Scientist *with Tenure*, BNL
- 2024 - Department Chair, Artificial Intelligence Department, BNL
- 2022 - 2024 Senior Scientist, BNL
- 2018 - 2024 AI/ML Group Lead, BNL
- 2016 - 2021 Computational Scientist, BNL
- 2013 - 2016 Associate Computational Scientist, BNL
- 2011 - 2013 Assistant Computational Scientist, BNL
- 2010 - 2011 Research Associate, BNL

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. S. Li, S. Yoo, Y. Yang, “Maximum Update Parameterization and Zero-Shot Hyperparameter Transfer for Fourier Neural Operators”, ICML ‘25
2. G. Zhao, B. Yoon, G. Park, S. Jha, S. Yoo, X. Qian, “Pareto Prompt Optimization”, ICLR ‘25
3. X. Yu, S. Yoo, Y. Lin, “Clipceil: Domain generalization through clip via channel refinement and image text alignment” NeurIPS ‘24
4. P. Kim, J. Kim, et al, “SwiFT: Swin 4D fMRI Transformer”, NeurIPS ‘23
5. Z. Sun, Y. Yang, S. Yoo, “Sparse Attention with Learning-to-hash”, ICLR ‘22

RESEARCH INTERESTS AND EXPERTISE

- Extreme Scale Scientific Machine Learning
- Quantum Machine Learning

SYNERGISTIC ACTIVITIES

1. DOE SciDAC RAPIDS2 Institute. AI Co-lead
2. DOE Institutional Review Board (IRB) for AI Working Group. Core Member
3. 2024 DOE ASCR Neuromorphic Computing for Science Workshop. Organizing Committee member
4. IBM Quantum Computing Working Group. Co-lead for workflow of material science working group and participant of high energy physics (HEP) working group.
5. IEEE Applied Signal Processing Society. Technical Committee member and IEEE Task Force for Rebooting Computing (TFRC) Liaison.

COLLABORATORS (PAST 5 YEARS INCLUDING NAME AND CURRENT INSTITUTION)

- Yiming Yang (CMU)
- Yihui Ray Ren (BNL)
- Xin Dai (BNL)
- Frank Alexander (ANL)
- Tom Brettin (ANL)
- Christopher Henry (ANL)
- Fangfang Xia (ANL)
- Rob Ross (ANL)
- Arvind Ramanathan (ANL)
- Sandeep Madireddy (ANL)
- Kyle Gerad (ANL)
- Adam Arkin (LBL)
- Paramvir Dehal (LBL)
- Shane Cannon (LBL)
- Lenny Oliker (LBL)
- Samuel Williams (LBL)
- Khaled Ibrahim (LBL)
- Dmitriy Morozov (LBL)
- John Wu (LBL)
- Steve Farrel (LBL)
- Shashank Subramanian (LBL)
- Prasanna Balaprakash (ORNL)
- Scott Klasky (ORNL)
- David Pugmire (ORNL)
- Feiyi Wang (ORNL)
- Sajal Dash (ORNL)
- Ramakrishnan Kannan (ORNL)
- Shantenu Jha (PPPL)
- Michael Churchill (PPPL)
- Laura Biven (JLab)
- Malachi Schram (JLab)
- Bill Tang (Princeton University)
- Aditi Krishnapriyan (Berkeley University)
- Romit Maulik (Penn State University)
- Kevin Huck (NVIDIA)

Curriculum Vitae

Co-PI NAME: Jiook Cha, PhD

PROFESSIONAL PREPARATION

- State University of New York at Stony Brook, New York, NY
    - Ph.D. in Neuroscience, 2013
- The Catholic University of Korea, Seoul, Korea
    - M.S. in Neurobiology, 2009
- Korea University, Seoul, Korea
    - B.S. in Environmental and Ecological Engineering, 2007

APPOINTMENTS

- 2020–Present Associate Professor, Psychology, Brain and Cognitive Sciences, Seoul National University, Seoul, Korea
- 2016–2020 Assistant Professor of Clinical Neurobiology, Department of Psychiatry, Columbia University, New York, NY
- 2014–2020 Research Scientist, New York State Psychiatric Institute, New York, NY
- 2015–2017 Visiting Scholar, Institute of Neuroscience and Psychology, The University of Glasgow, Glasgow, UK
- 2014–2016 Postdoctoral Research Fellow, Department of Psychiatry, Columbia University Medical Center, New York, NY

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. Kwon, J., … Cha, J. (2025). Predicting task-related brain activity from resting-state brain dynamics with fMRI Transformer. *Imaging Neuroscience*.
2. Kim, B. G., ... Cha, J. (2024). White matter diffusion estimates in obsessive-compulsive disorder across 1653 individuals: machine learning findings from the ENIGMA OCD Working Group. *Molecular Psychiatry*, *29*(4), 1063-1074.
3. Styll, P., … Cha, J. (2024). Swin fMRI Transformer Predicts Early Neurodevelopmental Outcomes from Neonatal fMRI. *AAAI Workshop*.
4. Kwon, J., … Cha, J. (2024). Revisiting Your Memory: Reconstruction of Affect-Contextualized Memory via EEG-guided Audiovisual Generation. AAAI Workshop.
5. Kim, P., … Moon, T. (2023). Swift: Swin 4d fmri transformer. *Advances in Neural Information Processing Systems*, *36*, 42015-42037.

RESEARCH INTERESTS AND EXPERTISE

- Computational Psychiatry and Clinical Neurobiology: Applying machine learning and computational analysis to understand the neural circuits of psychiatric disorders.
- Multimodal Neuroimaging: Utilizing various imaging techniques, including MRI, fMRI, and tractography, to study brain structure, function, and connectivity.

SYNERGISTIC ACTIVITIES

1. **Peer Review:** Ad hoc journal reviewer for *The American Journal of Psychiatry*, *Neuron*, *Neuroimage*, *Human Brain Mapping*, *The British Journal of Psychiatry*, *Developmental Cognitive Neuroscience*, and *Sleep*.
2. **Leadership and Service:** Served as President of the Korean Biomedical Scientists at Columbia University (2015-Present). Was a member of the local organizing committee for the ITU and WHO "AI for Health" meeting held at Columbia University (2018).
3. **Mentoring:** Has mentored high school volunteers, graduate interns, and postdoctoral research assistants in neuroimaging, psychiatry, and data science.
4. Professional Society Engagement: Maintained active membership in the Society for Neuroscience, Organization of Human Brain Mapping, and the Society of Biological Psychiatry.
5. **Principal Investigator on Grants:** Served as Principal Investigator (PI) or multi-PI on numerous funded research projects, including an NIMH K01 Career Development Award , a NARSAD Young Investigator Award from the Brain & Behavior Research Foundation , and an Institutional KL2 Award from Columbia University.

COLLABORATORS (PAST 5 YEARS INCLUDING NAME AND CURRENT INSTITUTION)

- Shinjae Yoo, BNL
- Russell Arens, The University of Illinois College of Medicine
- J.A. Gingrich, Columbia University
- Tae-Sub Moon, Seoul National University
- Lilianne R. Mujica-Parodi, State University of New York at Stony Brook
- Jonathan Posner, Columbia University
- Joanna Steinglass, Columbia University
- Myrna M. Weissman, Columbia University

Curriculum Vitae

Co-I NAME: Chun Kee Chung

PROFESSIONAL PREPARATION

- Seoul National University, Graduate School, Seoul, Korea
    - Ph.D. in Neurosurgery, 1993
- Seoul National University, Graduate School, Seoul, Korea
    - M.S. in Neurosurgery, 1986
- Seoul National University, College of Medicine, Seoul, Korea
    - M.D. in Medicine, 1983

APPOINTMENTS

- 2024–Present Senior Researcher, Neuroscience Research Institute, Seoul National University Medical Research Center
- 2013–2023 Professor, Dept. of Brain & Cognitive Sciences, Seoul National University, College of Natural Sciences
- 2010–2014 Chairman, Dept. of Neurosurgery, Seoul National University, College of Medicine
- 2006–Present Professor, Dept. of Neurosurgery, Seoul National University, College of Medicine
- 2000–2006 Associate Professor, Dept. of Neurosurgery, Seoul National University, College of Medicine
- 1997–2000 Assistant Professor, Dept. of Neurosurgery, Seoul National University, College of Medicine
- 1995–1997 Research Fellow, Dept. of Neurosurgery, Cleveland Clinic Foundation, OH, USA
- 1993–1995 Instructor, Dept. of Neurosurgery, Seoul National University, College of Medicine
- 1991–1993 Clinical Fellow, Dept. of Neurosurgery, Seoul National University Hospital

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. Jun, S., … Chung, C.K. (2020). Task-dependent effects of intracranial hippocampal stimulation on human memory and hippocampal theta power. *Brain Stimulation,* 13 (3), 603–613.
2. Kim, H., … Chung, C.K. (2023). Identification of cerebral cortices processing acceleration, velocity, and position during directional reaching movement with deep neural network and explainable AI. *NeuroImage,* 266, 119783.
3. Meng, K., … Grayden, D.B. (2023). Continuous synthesis of artificial speech sounds from human cortical surface recordings during silent speech production. *Journal of Neural Engineering.* https://doi.org/10.1088/1741-2552/ace7f6

4. Lee, D.H., … Ryun, S. (2024). Unravelling tactile categorisation and decision-making in the subregions of supramarginal gyrus via direct cortical stimulation. *Clinical Neurophysiology,* 158, 16–26.
5. Yeom, H.G., … Chung, C.K. (2023). A magnetoencephalography dataset during three-dimensional reaching movements for brain–computer interfaces. *Scientific Data,* 10 (1), 552. https://doi.org/10.1038/s41597-023-02454-y

RESEARCH INTERESTS AND EXPERTISE

- Translational Neuroengineering: ECoG / sEEG analytics, multi-site brain stimulation, and closed-loop neurosurgical modulation.
- Neural Decoding & BCI: Speech- and motor-decoding, memory-network mapping, and next-gen AI-driven brain–computer interfaces.

SYNERGISTIC ACTIVITIES

1. **Leadership and Service:** Served as the Director of the Neurosurgery Department at Seoul National University Hospital, and led multiple national-level neuroscience initiatives including the Ministry of Industry's Alchemist Project for "Brain to X (B2X)" interface development (2020–2025).
2. **Mentoring:** Has mentored numerous medical students, graduate students, and postdoctoral researchers in neurosurgery, neuroengineering, and cognitive neuroscience. Many of his trainees have advanced to faculty or clinical leadership roles across academic hospitals and research institutes in the world.
3. **Professional Society Engagement:** Fellow of the Korean Academy of Science and Technology (KAST) and the National Academy of Medicine of Korea, and contributor to international collaborations with institutions such as the University of Melbourne.
4. **Principal Investigator on Grants:** Led major national research projects including the Alchemist Project to develop a fully-implantable closed-loop “Brain to X” system. And a hippocampal stimulation study for human memory enhancement, and research on multi-site brain stimulation and real-time closed-loop systems for studying human somatosensory perception and decision-making.

COLLABORATORS (PAST 5 YEARS INCLUDING NAME AND CURRENT INSTITUTION)

- June Sic Kim, Clinical research Institute, Konkuk University Medical Center, Konkuk University Hospital, Korea.
- Myung Joo Kang, Department of Mathematical Sciences, Seoul National University
- Ryu Ernest Kang, Department of Mathematics, University of California, Los Angeles
- Jong-Hyun Ahn, School of Electrical and Electronic Engineering, Yonsei University, Korea.
- Sung-Phil Kim, Department of Biomedical Engineering, Ulsan National Institute of Science and Technology, UNIST
- Se Bum Paik, Department of Brain and Cognitive Sciences, Bio and Brain Engineering, Korea Advanced Institue of Science and Technology, KAIST
- Chang-Hwan Im, Department of Biomedical Engineering, Hanyang University

Curriculum Vitae

Co-I NAME: Taesup Moon

PROFESSIONAL PREPARATION

- PhD. 2008 Stanford University, Department of Electrical Engineering
- MS 2004 Stanford University, Department of Electrical Engineering
- BS 2002 Seoul National University, Department of Electrical Engineering

APPOINTMENTS

- 2024–present Professor, Department of Electrical and Computer Engineering, Seoul National University
- 2021–2024 Associate Professor, Department of Electrical and Computer Engineering, Seoul National University
- 2017–2021 Assistant/Associate Professor, Department of Electrical and Computer Engineering, Sungkyunkwan University
- 2015–2017, Assistant Professor, Department of Electrical Engineering and Computer Science, DGIST
- 2013–2015, Research Staff Member, Samsung Electronics
- 2012–2013, Postdoc, Department of Statistics, UC Berkeley
- 2008–2012, Scientist, Yahoo! Labs

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. Peter Yongho Kim, Junbeom Kwon, Sunghwan Joo, Sangyoon Bae, Donggyu Lee, Yoonho Jung, Shinjae Yoo, Jiook Cha, Taesup Moon, SwiFT: Swin 4D fMRI Transformer, *Neural Information Processing Systems (NeurIPS)*, December 2023
2. Jaeseok Byun, Taebaek Hwang, Jianlong Fu, and Taesup Moon, GRIT-VLP: Grouped Mini-batch Sampling for Efficient Vision and Language Pre-training, *European Conference on Computer Vision (ECCV)*, October 2022
3. Jaeseok Byun, Dohoon Kim, Taesup Moon, MAFA: Managing False Negatives for Vision-Language Pre-training, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2024
4. Sungmin Cha, Kyunghyun Cho, Taesup Moon, Regularizing with Pseudo-Negatives for Continual Self-Supervised Learning, *International Conference on Machine Learning (ICML)*, July 2024
5. Dohyun Kim, Jungtae Lee, Jangsup Moon, and Taesup Moon, Interpretable Deep Learning-based Hippocampal Sclerosis Classification, *Epilepsia Open (IF=4.026)*, https://doi.org/10.1002/epi4.12655, September 2022

RESEARCH INTERESTS AND EXPERTISE

- Foundation model, neuroimaging, vision-language model, machine learning,

COLLABORATORS

- Tsachy Weissman, Stanford University
- Flavio P. Calmon, Harvard University
- Jangsup Moon, Seoul National University Hospital
- Jianlong Fu, Microsoft Research Asia, Beijing
- Adrian Weller, University of Cambridge

Curriculum Vitae

Co-I NAME: Jay-Yoon Lee

PROFESSIONAL PREPARATION

- Carnegie Mellon University, Pittsburgh, PA
  - Ph.D. in Computer Science, 2020
- Carnegie Mellon University, Pittsburgh, PA
  - M.S. in Computer Science, 2013. *Transferred to PhD program at the end without completing.*
- Korea Advanced Institute of Science and Technology (KAIST), Seoul, Korea
  - B.S., Summa Cum Laude, in Electrical Engineering, 2008

APPOINTMENTS

- 2022–Present Assistant Professor, Graduate School of Data Science, Seoul National University, Seoul, Korea
- 2020–2022 Postdoctoral Associate under Professor Andrew McCallum, Department of Computer Science, University of Massachusetts Amherst, Amherst, MA
- 2015–2020 Research Assistant under Professor Jaime Carbonell, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. Song, M., … Lee, J. (2025). Introducing Verification Task of Set Consistency with Set-Consistency Energy Networks. *ACL.*
2. Cho, Y., Lee, J. (2025). CoMRes: Semi-Supervised Time Series Forecasting Utilizing Consensus Promotion of Multi-Resolution. *ICLR*.
3. Yoo, G., Lee, J. (2025) Improving NMT models by REtrofitting Quality Estimators into Trainable Energy Loss. *COLING (Oral).*
4. Park, S., Lee, J. (2024). Toward Robust RALMs: Revealing the Impact of Imperfect Retrieval on Retrieval-Augmented Language Models. *TACL. EMNLP (Oral)*.
5. Lee, J., …, McCallum, A. (2022). Structured Energy Network As a Loss. *NeurIPS*.

RESEARCH INTERESTS AND EXPERTISE

- Injecting knowledge as constraints into neural models.
- Making the models more coherent, interpretable, and controllable.
- Resolving low-resource problem by incorporating prior knowledge such as constraints.
- Automatically capturing implicit constraints using energy-based models.
- Reflecting scientific knowledge into science AI.

SYNERGISTIC ACTIVITIES

1. **Peer Review:** Has served as a reviewer for ICML, ICLR, NeurIPS, ACL, and EMNLP (2019-Present)
2. **Leadership and Service:** Has served as an Area Chair for NeurIPS, EMNLP, ARR, and COLING (2023-Present). Selected as Best Area Chair for Machine Learning track at EMNLP 2023 and Highlighted Reviewer of ICLR 2022.
3. **Mentoring:** Has mentored various Ph.D. and MS students in machine learning and data science.
4. **Principal Investigator on Grants:** Has served as Principal Investigator (PI) on funded research projects, including the Veteran/Consolidator Research Program from the National Research Foundation of Korea and a contracted R&D project from the Korea Institute of Science and Technology Information (KISTI).

COLLABORATORS (PAST 5 YEARS INCLUDING NAME AND CURRENT INSTITUTION)

- Andrew McCallum, University of Massachusetts Amherst
- Jiook Cha, Seoul National University
- Mohit Iyyer, University of Maryland
- Yulia Tsvetkov, University of Washington
- Hannaneh Hajishirzi, University of Washington
- Lifu Huang, UC Davis
- Rajarshi Das, AWS AI Labs
- Manzil Zaheer, Google Research
- Dheeraj Rajagopal, Fastino AI
- Md Arafat Sultan, IBM Research AI

Curriculum Vitae
Co-I NAME: Franco Pestilli, PhD

PROFESSIONAL PREPARATION

- **Postdoctoral Fellow**, Computational Neuroimaging, Stanford University, CA, 2011–2013
- **Postdoctoral Fellow (NIH training grant)**, Neuroimaging and Neurophysiology, Columbia University, New York, NY, 2008–2011
- **Ph.D.** in Psychology (Cognition and Perception), New York University, New York, NY, 2008
- **M.A.** in Cognitive Psychology, New York University, New York, NY, 2006
- **A.B.** in Psychology, University of Rome "La Sapienza", Rome, Lazio, 2000

APPOINTMENTS

- **2020–Present** Associate Professor, The University of Texas at Austin, Austin, TX
- **2019–2020** Associate Professor, Indiana University, Bloomington, IN
- **2015–2019** Assistant Professor, Indiana University, Bloomington, IN
- **2013–2014** Research Associate, Stanford University, Stanford, CA
- **2002–2008** Ph.D. Candidate, The New York University, New York, NY
- **2001–2002** Research Assistant, The New York University, New York, NY

FIVE PUBLICATIONS MOST RELEVANT TO THIS PROPOSAL

1. *Vinci-Booher S, McDonald DJ, Berquist E, Pestilli F. Associative white matter tracts selectively predict sensorimotor learning. Commun Biol. 2024 Jun 22;7(1):762. PubMed Central PMCID: PMC11193801.*
2. *Renton AI, Dao TT, J…, Zhu JD, Narayanan A, Bollmann S. Neurodesk: an accessible, flexible and portable data analysis environment for reproducible neuroimaging. Nat Methods. 2024 May;21(5):804-808. PubMed Central PMCID: PMC11180540.*
3. *Hayashi S, Caron BA, …, Pestilli F. brainlife.io: a decentralized and open-source cloud platform to support neuroscience research. Nat Methods. 2024 May;21(5):809-813. PubMed Central PMCID: PMC11093740.*
4. *Bertò G, Bullock D, … Olivetti E. Classifyber, a robust streamline-based linear classifier for white matter bundle segmentation. Neuroimage. 2021 Jan 1;224:117402.*

PubMed PMID: 32979520.
5. *Sani I, Stemmann H, … Pestilli F, Freiwald WA. The human endogenous attentional control network includes a ventro-temporal cortical node. Nat Commun. 2021 Jan*

RESEARCH INTERESTS AND EXPERTISE

- **Computational Neuroscience:** Understanding the biological mechanisms of the brain across cellular, behavioral, and population scales.
- **Multimodal Neuroimaging:** Utilizing and combining diffusion MRI (dMRI), tractography, and other MR imaging methods.
- **Neuroinformatics and Open Science:** Developing open-source software, machine learning pipelines (e.g., brainlife.io), and data standards (e.g., BIDS) to promote reproducible and efficient research.
- **Data Curation and Management:** Creating robust pipelines for large-scale neuroimaging data sharing and analysis.

SYNERGISTIC ACTIVITIES

1. **Journal Membership:** Member of the editorial boards for *Neuroimage, Nature's Scientific Data.*
2. **International Working Groups:** Active member of the International Brain Initiative (IBI) Data Standards and Sharing Working Group.
3. **Standards Development:** Contributor to the Brain Imaging Data Structure (BIDS) project, a community standard for organizing and sharing neuroimaging data.
4. **Open-Source Platform Leadership:** Leading a major research and development effort for brainlife.io, a NIH's BRAIN Initiative-sponsored open science platform.

COLLABORATORS (PAST 5 YEARS INCLUDING NAME AND CURRENT INSTITUTION)

- Winrich A. Freiwald, The Rockefeller University
- Steffen Bollmann, The University of Queensland, Australia
- Soichi Hayashi, Stanford University
- Emanuele Olivetti, University of Trento, Italy