AccelNet-Implementation: HSF-India - Research Software Networks in Physics

HSF-India is a US National Science Foundation (NSF) funded project that begins October 1, 2022. HSF-India will join collaborative networks in India to networks in the U.S. and Europe in order to build the international research software collaborations required to reach the science goals of experimental particle, nuclear and astroparticle physics experiments. Unanswered questions within reach of scientists include the origin of dark matter, properties of the Higgs Boson, the structural makeup of matter, and the mechanism to explain neutrino mass. These are some of the highest priority science drivers to be addressed by next-generation experimental facilities including the high-luminosity Large Hadron Collider at CERN, the Deep Underground Neutrino Experiment at Fermilab, and the Electron Ion Collider at Brookhaven National Laboratory. To fully realize their discovery potential a new generation of software algorithms and approaches is required. Software is an intellectual product of this research, not just a critical tool. It has become a critical element to design and maximize the physics discovery potential of large data intensive science projects. Building these research software collaborations is challenging and inherently international matching the international nature of the experimental undertakings themselves. HSF-India will provide students, postdocs and early career personnel significant experience in international team science through engagement in a diverse research community.

HSF-India will build international collaborative structures to enable innovation in critical software areas for discovery in particle, nuclear and astroparticle physics and lay the foundations for similar international collaborations across the scientific spectrum. It will seed research software collaborations on specific projects across its networks in the U.S., Europe and India by leveraging an international training network, fellowships and software-focused researcher exchanges. HSF-India will iteratively build on these seeds towards a broad and sustainable international research software collaboration with three key research themes:

1. **Analysis Systems:** Tools and techniques enabling next-generation analysis approaches to realize the maximum scientific potential of the data in the least time. The final stages of data analysis for our science are topically diverse and highly collaborative within small subsets of the experiment collaborations. Researchers are working to develop relevant analysis tools in the Python ecosystem, where much of the broader data science and machine learning innovation is also happening. These tools must be scalable, efficient and easy to use in order to support scientific needs ranging from individual researchers to those required for datasets of the size and complexity of the largest facilities

2. **Novel simulation techniques** are critical tools for the increasingly complex questions being asked by current and planned experiments, and touch all aspects of experimental work in nuclear and particle physics. Traditional Monte Carlo techniques face large gaps with respect to needs for next-generation experiments. These can be closed through research that enables the efficient use of new computing technology; the application of machine learning approaches and more broadly emerging techniques such as simulation based inference and differentiable programming.

3. **Tools for Open Science** will ensure that researchers benefit from, and contribute to, the innovation of the broader data science community. This includes not only openness of the software, but also the development of corresponding open datasets to enable further software evolution. An open science approach is also key to enable international collaborations as well as multidisciplinary collaborations between physicists, computer scientists and other data science practitioners. Finally, by developing open tools that are generalized to work across many experiments, they will also find applications beyond our own science.

As described below, it will focus on opportunities for early-career researchers and establish mentoring and co-mentoring relationships with more senior researchers across the networks. By engaging with the data-science, artificial intelligence and broader computer science communities the project will foster bottom-up alliances around research software, for physics and beyond. HSF-India aims to establish an environment where researchers with diverse backgrounds, skill sets and interests can come together and build innovative collaborations that sustain novel tools and techniques while bridging research gaps to enable future scientific facilities.

HSF-India will focus on structured "collaboration building" activities. We say "collaboration building" rather than directly forming "collaborations" because an initial phase of building (often geographically distributed) "many to many" relationships is required. These will include:

1. **Training events** that will build upon the curriculum being actively developed and used today by the HSF and IRIS-HEP. This will include an introduction to the current analysis, deep learning and data science ecosystem, software development/engineering skills, and best practices for collaborative work. Practical sessions and group "hackathons" will develop team connections. We aim to engage 150 students per year and have funds to assist with instructor travel costs and networking events.
2. **Mentoring and co-mentoring** of early-career researchers across the networks. This will include curating a catalog of possible software-related projects of various sizes, including those sized for summer or semester-long projects as well as larger projects which are suitable to be the subject or a part of a student thesis.
3. **Fellowship programs** for students that have participated in our training programs, or have equivalent experience. Fellows will be based at their home institute (or elsewhere as convenient) and interact remotely with their mentors on software projects. Projects will nominally have two mentors, both for practical reasons (holidays, busy schedules, etc), and to bring different perspectives, from different networks, to the problem. We also aim for projects where two Fellows from different countries work together under the same mentors on a project. We aim to support 10 students per year for a three month fellowship.
4. **Researcher exchanges** organized and funded between the U.S. and India. While short in-person visits for training events, workshops and con- ferences can reduce barriers to communication, longer visits are critical to building and sustaining durable longer-term

relationships. We aim to provide travel support for 10 early-career researchers per year to participate in a researcher exchange across HSF-India networks.

5. **Virtual Topical Meetings** to allow participants to periodically present the status of their projects as well as invited external speakers with relevant software activities.
6. **Annual Project Meeting** to facilitate collaboration amongst HSF-India participants.

The HSF-India organizational structure will include reporting to NSF, and a steering board. The NSF reporting will be handled primarily by the PI-team consisting of David Lange (Princeton University), Peter Elmer (Princeton University), Rafael Coelho Lopes de Sa (University of Massachusetts, Amherst), Verena Martinez Outschoorn (University of Massachusetts, Amherst), and Heidi Schellman (Oregon State University). A steering board is to be formulated at the project kick off meeting based on an initial draft and input from all involved collaborators. The idea is that the steering board has a membership spanning the "networks" engaged in HSF-India research and meetings regularly to help organize HSF-India events and opportunities. Its makeup should include researchers from across the engaged experiments, regions and interests, as well as of researchers at all career stages and backgrounds.

The HSF-India network of networks extends the growing research software collaborations between physics and computer science researchers in the U.S. and Europe to include researchers in India. This network-of-networks will build expertise and drive research software collaborations in areas which are critical to advance particle, nuclear and astroparticle physics research: novel approaches to analysis systems, next-generation simulation tools and methods for open science. It will connect software experts and researchers across the international community in a collaborative framework that promotes software collaborations. Immediate mutual benefits of the international collaboration will be a training network that enables early-career researchers to pursue impactful research software initiatives in ways that advance their careers in experimental data-intensive science.