

Project Summary:

Name of project: *ESIPhub: Developing and promoting an ESIP community resource for sharing and running scientific workflows via JupyterHub.*

Project lead and contact details: Keith Maull (kmaull@ucar.edu)

Project partners and contact details:

- Sean Gordon (scgordon@hdfgroup.org)
- Matt Mayernik (mayernik@ucar.edu)

Proposed start and end date: July 2018 - July 2019

Budget Requested: \$5000

Budget Summary:

- To support continuing maintenance and use (~\$2000)
- Conference travel for 2018 ESIP Winter Meeting (~\$3000) to present report

Project Outline:

Project description: In the last round of funding the ESIPhub project created a JupyterHub for use at the 2018 Summer ESIP Meeting. While the Summer Meeting workshops have not occurred yet there are many different member communities that have agreed to use the ESIPhub cyberinfrastructure including NASA, NOAA, HDF Group, and Google. Prior to the Summer ESIP Meeting, we plan to test and use the technology during the UCAR Significant Opportunities in Atmospheric Research and Science (SOARS) workshop, June 1st 2018. Other community members have expressed interest and hope to leverage ESIPhub in the future for their own collaborative projects in the ESIP space. It is our belief that the Summer Meeting workshops will drive interest from additional member communities. We would like to ensure that the interest has support, and build upon the achievements of the previous Incubator Lab award. Additionally we would like to ensure sustainability beyond the Lab Incubator awards.

We would like to propose funding for a session reporting back to ESIP during the 2019 Winter Meeting. We would also like to request funds to make the CI available at the 2019 Winter Meeting and the 2019 Summer Meeting. It is our belief that this support will allow the creation of a new Cluster that would lead the development of Jupyter usage at ESIP into the future.

Jupyter Notebooks have emerged as a powerful way to document and enact scientific workflows. They combine richly documentable and executable code with the resulting analysis and interactive visualizations in a single shareable document. Notebooks are being adopted by teachers, scholars, scientists and working practitioners alike because they can provide direct evidence of processes and executable procedures through code and exposition, but more importantly provide a recipe for repeating, reproducing and repurposing those processes and procedures so that colleagues, collaborators and community members alike can understand, explore and ultimately benefit from the work. These

Notebooks have the potential to *dramatically accelerate adoption of community approaches and best practices within the ESIP community, but more continued infrastructure support is needed to realize this potential.*

We therefore think the best solution for ESIP would be to continue developing the JupyterHub instance created in conjunction with The National Data Service, allowing multi-user login with authentication, reliable resources, and shared environments. This incubator project will allow us to continue investigating and scoping this problem by interacting with the Jupyter community, improving and maintaining the pilot instance, and reporting on the results at ESIP meetings.

In this proposal, we would like to accomplish the following:

1. Continuing maintenance and support from NDS
2. Support the formation of a Jupyter Cluster
3. Use ESIPhub for workshops during the 2019 ESIP Meetings
4. Develop plan for sustainability of ESIPhub beyond Incubator funds.

Project objectives, significance and impact: The primary objective of this proposal is to promote an open computing environment that facilitates the sharing and reuse of code used across ESIP.

Description of key project steps and timeline:

| Project Activity | Expected Completion | Anticipated Outcome(s) |
|---|----------------------------|---|
| <i>Propose Jupyter Cluster</i> | July 2018 | Community development |
| <i>Present project report at 2019 ESIP Winter Meeting</i> | Jan 2019 | Report findings on cost models, infrastructure containerization, usage at Summer Meeting |
| <i>Workshop support</i> | Jan-July 2019 | Completed presentation |
| <i>Conclude project</i> | Aug 2019 | Project documentation, Enacted any feasible external support models for sustaining Cluster activities |

Outreach:

What groups/audiences will be engaged in the project?

How will you judge that project has had impact? Usage at meetings, Cluster activity, external sustainable model for ESIPhub

How will you share the knowledge generated by the project? Cluster Activities, workshops at Meetings

Project Partners (as applicable):

Description of project partners (individuals and/or organizations) and their involvement:

- Keith Maul: Project Lead
- Sean Gordon: Outreach
- Matt Mayernik: Project Partner

How will this project engage members of the ESIP community: Members of multiple Clusters and ESIP member organizations are already collaborating on workshops, with the likelihood that additional interest will be generated by the Summer 2018 Meeting.