1. **Contacts**

Primary contact: Matthew B. Jones

Director of Informatics Research and Development National Center for Ecological Analysis and Synthesis

735 State St., Suite 300, Santa Barbara, CA 93101

805-892-2500

jones@nceas.ucsb.edu

Secondary contact: Mark P. Schildhauer

Director of Computing

National Center for Ecological Analysis and Synthesis

735 State St., Suite 300, Santa Barbara, CA 93101

805-892-2500

schild@nceas.ucsb.edu

1. **Workshop title**

Conducting Open Science using R and DataONE: A Hands-on Primer

1. **Description (<250 words):**

The inherent variability in ecological systems has placed a premium on well-designed studies that employ sophisticated analysis and modeling approaches. These analytical approaches are critically important to understanding and verifying conclusions from these experiments. Although students learn how to acquire and analyze data, recently there has been increasing emphasis on doing so in an open manner that allows for full reproducibility of ecological science. In this workshop, we will examine the convergence of open data and open source tools and their ability to jointly facilitate open science.

The purpose of this workshop is to highlight the use of open software tools for conducting open science in ecology. Through a series of hands-on activities (see http://help.nceas.ucsb.edu/r), participants will explore approaches to accessing data from the DataONE federation in the R analytical system, and then conduct basic data assessment, summarization, and visualization using that data. We will also expose students to various R libraries available from the rOpenSci project for accessing data, and to the use of R for generating data documentation and archiving data in DataONE-compatible data repositories.

We are targeting early-career scientists, including students, postdocs, and faculty that would benefit from new techniques for open science and educators who want to incorporate open science concepts in curricula. Participants should have a basic understanding of data analysis and statistics, and preferably basic exposure to the R system. Participants must bring their own laptop to participate in hands-on activities, and must have the ability to install new software.

1. **Summary Sentence (<50 words):**

Through hands-on activities, participants will explore open science practices by using the R statistical system and libraries for data access from DataONE and rOpenSci to build analytical scripts to access data, learn basic data manipulation, and publish derived data and results in a manner that is open and citable.

1. **List of speakers**

*Confirmed participants*

Matthew Jones, NCEAS & DataONE

Mark Schildhauer, NCEAS

Jim Regetz, NCEAS

Karthik Ram, rOpenSci, UC Berkeley, DataONE

Carl Boettegir, rOpenSci

1. **List of equipment needed**

Screen

LCD Projector

Internet Connection with sufficient bandwidth for wireless transmission to ~35 users (wireless router provided)

Flip charts (5 for breakout discussions)

Power strips at each table

1. **Number of participants**

Minimum: 10

Maximum: 35

1. **Proposed date/time**

Sunday, August 5, 2011

12 – 5 pm

1. **Room setup**

Conference style with round tables and power strips at each table

1. **Underwriting of workshop costs**

Organization/Contact: NCEAS

Amount: Workshop refreshments