Data and Software Policy

1. Introduction

SESYNC strongly believes that sharing all data and software source code with the research community is a key ingredient towards fostering synthetic research at SESYNC and contributes to advancing socio-environmental science as a whole. In order to accomplish its mission and to best support the scientific community, SESYNC requires that all data and software created through SESYNC-sponsored activities be made publicly available.

2. Scope

This document defines the policies for sharing data and software source code that is created, generated, hosted, sponsored, or made available by a SESYNC-sponsored scientist, project, activity, or collaboration.

In the remainder of this document, data and software source code within the scope of this policy is collectively referred to as SESYNC-sponsored data and software source code, respectively.

Existing data or software that researchers bring to SESYNC which has more restrictive terms of use and dissemination than those set forth here are exempt from this policy. However, SESYNC strongly encourages its participants and collaborators to consider alternatives that do not restrict sharing and free dissemination.

3. SESYNC-sponsored Data

SESYNC is committed to making all data within the scope of this policy readily available to the broader scientific community.

To meet this goal, SESYNC sponsored-data is, insofar as possible, to be made accessible with no restrictions for use and dissemination through explicit use of a Creative Commons Zero waiver, or its equivalent, and through deposition in a public data repository or an established open database. Data is to be made available in a timely manner, no later than one year after the conclusion of the SESYNC award, or immediately upon publication of an associated article, whichever comes earlier. The data should be adequately documented for validation and reuse, including appropriate attribution of its original source.

SESYNC recognizes that there are many questions associated with the open licensing of scientific data collections and encourages sponsored scientists to familiarize themselves with the issues. SESYNC staff will gladly work with researchers to help determine the most appropriate licensing model for their data.

4. SESYNC-sponsored Software

SESYNC is committed to open-source software development and distribution. Software developed with SESYNC funding or participation is expected to become readily and freely available to the broader scientific community. Specifically, SESYNC-sponsored software is to be made accessible on a FTP or web-accessible site hosted by SESYNC, or at Source Forge or a similar site, for download in source code, and the source code is to be licensed under an OSI-approved Open Source license. The source code is to be made available in a timely manner, meaning as early as possible, but no later than the conclusion of the collaboration or SESYNC award, or publication of an article involving the software or data generated using the software, whichever comes earlier. The code is to be sufficiently documented, and at a minimum, state purpose, copyright and license, and acknowledge and attribute any other software or software libraries that the SESYNC-sponsored software uses or depends on, as required by the licenses of such software. Copyright and license must be stated in every source code file, and the licenses for all the files that constitute the software must be compatible.

5. Support By SESYNC IT

SESYNC's Cyberinfrastructure team will help a principal investigator or SESYNC fellow in selecting the most suitable Creative Commons or Open Source license, and appropriately placing the license on the website for the dataset or the software, as well as in depositing data, meta-data, and software source code in the pertinent public repository. SESYNC and its Cyberinfrastructure team will also help to remove any obstacles towards making the data and software available, for instance by researching and resolving license incompatibilities, consulting or coordinating with collaborators, etc.

6. Continuation of Support

SESYNC Cyberinfrastructure will schedule a wrap-up meeting with all SESYNC fellows and research groups about 3-4 weeks before their term at SESYNC ends. A goal of this interview is to identify any data, data sets, and source code that the scientist created under SESYNC support, and to develop a recommendation for what type of continued support SESYNC will provide for the data or data collection (for instance, hosting a database) and for the software (e.g., fielding user questions, bug resolution, or continued feature development). The recommendation will be communicated to the scientist, upon approval by SESYNC's Director. Please see the SESYNC IT policy (http://www.sesync.org/policies) for details on what information SESYNC collections about data and software produced at the center.

While SESYNC provides data hosting and limited long term support we encourage all groups to identify community-specific sites for disseminating there data where it may be more visible to their peers. SESYNC is able to help with this process for instance by assisting groups in publishing data sets through Dryad or DataONE.

7. Copyright and Acknowledgment

Data or datasets and software generated or created by a SESYNC fellow or through a SESYNC supported meeting or project must appropriately and conspicuously acknowledge SESYNC and the National Science Foundation.   For example, "This work was supported by the National Science Foundation DBI-1052875 to the National Socio-Environmental Synthesis Center."