INTRODUCTION

The Missouri Transect project was established through an NSF EPSCoR Research Infrastructure Improvement Track 1 award (IIA-1355406) in 2014 to build infrastructure, knowledge, and collaborations in research and education across Missouri. The project focuses on activities that integrate high-resolution climate data with high-throughput plant science and with the impact on society to understand the role of changing environments in Missouri. Sharing data among Missouri Transect participants is instrumental to successful collaboration within this large statewide project and sharing with the public makes data and knowledge available to the taxpayers that fund the work through NSF and is instrumental in meeting the Missouri Transect goal to "...provide the foundation for integrated and interdisciplinary research on climate, plant sciences, and community engagement." Furthermore, NSF policy states that "Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants." This document establishes the policies for the Missouri Transect project to ensure timely release of research data and products.

RESEARCH DATA POLICY STATEMENTS

1. Communications and Consultation Policy Statement

Missouri Transect PIs, Co-PIs, the Executive Committee, and the CI Team have a responsibility to routinely communicate and consult to 1) assess and record the status and effectiveness of CI expenditures (e.g. hardware, software, tools, and personnel); 2) coordinate and develop strategies that maximize the cumulative effect of CI expenditures and promote sustainability; and 3) define appropriate data standards and repositories for Missouri EPSCoR data and research products.

2. Researcher Accountability and Additional Responsibilities Policy Statement

Researchers are expected to generate a broad array of scientifically robust high quality data suitable for further use. Individual researchers are responsible for data management throughout the data life-cycle, documenting data lineage and properly protecting the data. Researchers who generate data that require additional protections, such as research that is subject to IRB review, are responsible for communicating and requesting any additional assistance needed from the CI Team in proper curation of these data. It is the responsibility of individual researchers to contribute their raw, processed, and derived data products, along with valid standards-based metadata, in a timely manner as explicitly described below. The researchers are also responsible for providing the CI Team with their source of data format and metadata standards, so that the CI Team can maintain a list of standards and coordinate the use of common standards among Missouri Transect researchers

when possible. The team leads are responsible for monitoring and following up with individual researchers who need to provide data and data standards. The Missouri EPSCoR Director has the responsibility for enforcement. Any barriers will be clarified and documented by the team lead and approved or denied by the EPSCoR Director. If the EPSCoR Director is unable to resolve the issue, the EPSCoR Director will communicate the problem to the appropriate Vice President of Research at the institution where the researcher resides.

3. Timeliness of Data and Research Products Contribution Policy Statement

Research data and research products created by Missouri Transect participants, and other supporting materials, should be shared within a reasonable time and in accordance with data sharing policies from the National Science Foundation (NSF) (the National Science Foundation Award & Administration Guide (AAG) Chapter VI.D.48). Research data and research products created by Missouri Transect participants will be shared with two audiences: 1) other Missouri Transect participants and 2) the general public.

Timing and Mechanism of Data Release: During the course of the Missouri Transect project, all data, with the exception of human subject data (see below), will be shared via the Missouri Transect server or another server approved by the CI Team. Raw data and processed data, along with metadata and brief data descriptions, will be immediately available to Missouri Transect participants. Samples of raw and processed data, metadata and data descriptions in layman's terms will be immediately available to the public. Complete raw and processed datasets that are not immediately freely available to the public will be provided upon request at any time via an online request form that will require requestors to enter their name, affiliation, description of how they intend to use the data, and to agree not to publish the data without consent. All data will be freely available to the public once it is published or upon project termination, whichever comes first.

Human Subject Data: All such data (e.g. surveys) will be collected and managed only by personnel with adequate human subject protection certification. Only derived data will be shared as described above.

4. Intellectual Property, Ownership, Re-use and Redistribution Policies
Our default policy is that data will be available for re-use and re-distribution under the appropriate Creative Commons License (http://creativecommons.org/licenses/), to enable maximum dissemination and use. Commercial re-use of project data, algorithms, and software may be subject to intellectual property policies at the respective partner institutions, at the discretion of each institution.

Our default policy for newly developed algorithms and software will be to release as free open-source software, so as to maximize the dissemination and impact of the

2

tools. Code and software will be made available in repositories, such as GitHub (https://github.com) and Sourceforge (http://sourceforge.net).

5. Research Data Citation Policy Statement

After data have been deposited and cataloged in a repository, users of the data should appropriately acknowledge the National Science Foundation, Missouri EPSCoR and the individual investigators responsible for the data set. Any use of data provided by the Missouri EPSCoR must acknowledge Missouri EPSCoR and the funding source(s) that contributed to the collection of the data. Any restrictions on the usage of the data shall be clearly stated and obvious to the potential data user. All derived data products built from external data sources will cite the external data sources and collectors in the metadata for the new derived product.

6. Data Formats and Metadata Requirements

The format of deposited data, raw, processed and derived, should conform to open, non-proprietary and documented standards, where applicable. When data and data products cannot be converted to such formats after a reasonable effort, researchers will provide sufficient documentation about the data format and software needed to access the data. For example, Excel spreadsheets should be converted to comma separate files that are not bound to proprietary software and are compatible with many data analysis routines.

All research data and research data products must be documented and cataloged with appropriate, complete metadata. Shared data without adequate, corresponding metadata has limited intrinsic value and is incomplete. The Missouri Transect Data Management Plan will maintain a list of Missouri Transect approved metadata standards, based on information provided by researchers with data domain expertise. The CI Team must approve the use of alternative metadata standards. Upon request, the CI Team will provide necessary documentation, coordination, and training for effective use of these metadata standards.

7. Data Management as a Training Component

Students and Postdocs engaged in the Missouri Transect project will be required to design an individual data management plan. Their PIs and the CI Team will advise students in the development of effective, project specific data management plans. This requirement will contribute to the professional development of both the students and their faculty mentors.

8. Researcher Expectation of Infrastructure Policy Statement

Researchers have a right to expect 1) clear procedures for data cataloging and depositing their data; 2) unrestricted access to edit and modify their data or records about their data (i.e., metadata); 3) the availability of tools for cataloging and

3

depositing their data, or in the absence of tools, staff who can assist them with data curation; 4) a secure and well maintained repository for their data; 5) adequate mechanisms for data discovery and access by others; 6) that they, as creators of data, have first rights to analyze and publish those data; 7) to be credited for their data contributions; and 8) to be assured of adequate protection of intellectual property.