

## Data Management Plan

### Authorship Policy

For authorship and reporting guidelines we will refer participants to the **Vancouver Recommendations**: a set of recommendations for best practice and ethical standards in the conduct and reporting of published research (<http://www.icmje.org/icmje-recommendations.pdf>). They recommend that all 4 of the following criteria are required for authorship: (i) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (ii) drafting the work or revising it critically for important intellectual content; AND (iii) final approval of the version to be published; AND (iv) agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. They also recommend that when a large multi-author group has conducted the work, the group ideally should decide who will be an author before the work is started.

### Description of Data Types

This project will not generate new data, but may develop databases based on synthesis or metaanalysis of the relevant literature, in addition to computer code and pipelines (i.e., steps of analyses from data simulation to analysis to output) produced from meta analysis. Code will be annotated to promote interpretation and detailed headers will be used to explain the purpose and functions of the scripts, input variables needed, etc. We will set up a **GitHub** project for ECS-RCN to ensure public access and reproducibility of all code produced by this working group.

### Metadata Formats and Standards

Metadata: All data compiled during this project will have appropriate metadata and will comprise different formats. While there is not a single common metadata standard for the type of data we will be producing, essential information will be documented for all data.

### Data Storage, Materials Sharing, and Access During the Project

Data will be stored in **Open Science Framework** (<https://osf.io/>), a website for managing large collaborative scientific projects and controlling public access to data.

### Mechanisms and Policies for Access and Sharing

Final data products will be made available to the public via Dryad as soon as is reasonably possible after publication. As necessary, data will also be made available through supplementary materials sections of peer-reviewed journals, or upon request to the PIs. To preserve the PIs' rights to first publication and to make users aware of any data quality issues, we will provide unpublished data with the requirement that users obtain permission from all PIs before using this data in a publication.

### Policies for Re-Use and Re-Distribution

All other data products generated by this project will be made freely available for research, educational, management, and non-profit purposes with the restrictions described above and in accordance with University/Participating institutional and NSF policies. Consistent with normal scientific practices and publication procedures, publication of data shall occur as the project proceeds or after the project ends.

Data directly associated with research findings will be made available after the main findings have been accepted for publication or as described above.

### **Plans for Archiving, Data Storage, and Preservation**

Our plan for data storage and preservation includes strategies for frequent backup and storage of all project data during the life of the project and for long-term preservation and open access of primary data products. All data files will be versioned over time and stored using a date-stamped file structure. While the project is active, daily and monthly backups of all data files (in the form of redundant hard drives) will be retained at Northeastern University. To achieve longer-term storage and open access after publication, primary data products will undergo final quality control checks and then will be archived either through online repositories such as **DRYAD** (<http://datadryad.org/>).

### **Roles and Responsibilities**

Lotterhos will be responsible for data management, and will take the lead on ensuring that all project personnel comply with the data management plan.