UCONN
UNIVERSITY OF CONNECTICUT

University Libraries

Submission Process Guide

What is this form?
This document aims to provide a step by step guide to gather key information needed about the Work to be deposited to the UCL Data Archive. Here, the Work refers to any data set, research, or other digital scholarship.

Who is this form for?
This form is for the person who wants to submit a Work to the UCL Data Archive. The sections and questions can be filled out by only the depositor or in consultation with a subject liaison and/or one the Libraries' data specialists.

Are all the sections required to have information?
No, it is not required to fill out all the sections or questions. Not everything will apply to your Work. **Minimally, the UCL Data Archive requires a title for your Work, at least one creator, file format and type of data.** However, it is recommended to provide as much information about your Work that will allows others to discover it, understand it, contextualize it, and know if the Work can be reused or shared.

What to do with this form?
Once you've finished with this form, contact the UCL Data Archive (researchdata@uconn.edu) and add this document as an attachment.

Note on content in the UCL Data Archive:
Any Work deposited to the UCL Data Archive **will be publicly available once deposited**. The UCL Data Archive cannot accept at this time any Work that contains sensitive, restricted, or commercial data or data that contains personally identified information.

What are the sections?
The sections found in this document are grouped into 8 main categories that reflect information typically provided in a data management plan: Basic information about the Work, Work's product(s), Source(s) of the work, Work's format(s), Work's data and metadata standards, Work's size and scope, Work's legal and ethical compliance, sharing and accessing the Work.

University Libraries

SECTIONS

1. Basic Information about The Work

Roles and responsibilities include information on who is the primary contact and person responsible for the Work and all those who played a role to produce the Work. Include the names of the people involved and their roles. Include any person or funder who needs to be recognized in relation to the Work.

Primary Contact Name	
Primary Contact Email	
Title of Work	
Creator(s)	
Contributor(s)	
Relevant date(s)	
Funder(s)	
Abstract	
Keyword(s)	
Rights and License(s)	
Other relevant information	

Example:

Primary Contact Name	*Pr. John J. Smith*
Primary Contact Email	*John.j.smith@uconn.edu*
Title of Work	*Data from: How to calculate the non-synonymous to synonymous rate ratio of protein-coding genes under the Fisher-Wright mutation-selection framework.*
Creator(s)	*J.J. Smith, Alison Johnson, Stuart Mill*
Contributor(s)	*Hue Ling*
Relevant date(s)	*Data created on 2/05/2012, paper accepted on 6/08/2014*
Funder(s)	*NSF-BIO grant number 6552-11*
Abstract	*First principles of population genetics are used to obtain formulae relating the non-synonymous to synonymous substitution rate ratio to the selection coefficients acting at codon sites in protein-coding genes. Two theoretical cases are discussed and two examples from real data (a chloroplast gene and a virus polymerase) are given. The formulae give much insight into the dynamics of non-synonymous substitutions and may inform the development of methods to detect adaptive evolution.*
Keyword(s)	*Evolution, influenza*

University Libraries

Rights and License(s)	*Creative Commons BY-NC-SA 2.0 (https://creativecommons.org/licenses/by-nc-sa/2.0/)*
Other relevant information: article	*How to calculate the non--synonymous to synonymous rate ratio of protein-coding genes under the Fisher--Wright mutation--selection framework. Biology Today 11(4): 20140904.*

2. The Work's product(s)

The Work's product(s) can include digital data, physical samples or collections, and supporting materials. Primary data that will be deposited should be listed. If physical samples or collections would like to be given to the Libraries', list those separate and the University Archivist will be consulted to determine whether these materials can be added to the UCL Archives & Special Collections.

Example: My Work consists of experimental tabular data, spectral data, analytical data from ion chromatography, mass spectral data, fluorescent lifetime data, transient absorption and quantum yields of photoproducts.

3. The Work's source(s)

Source(s) can include the instrument(s) or collection approaches used to produce the Work and/or analyze the work. Include any analytical software, field collection.

![UCONN - UNIVERSITY OF CONNECTICUT]

University Libraries

Example: Both quantitative and qualitative data will be generated that assesses the outcomes of the REU program. The assessment data will result from surveys completed at the end of each year of the Program and information collection from the participants after they have completed the program. All files contain a page identifier code, searchable through Windows text search software. Analysis is done using SPSS.

4. The Work's format(s)

Formats are the type of files produced by the research and that will be deposited to be stored and shared. Include specific software formats and whether any formats are proprietary or customized for a research instrument or converted to an open and/or accessible format. Examples include: JPG image, TIF image, FITS image, MATLAB, MS Excel Table converted to CSV, SQL database, Instrument's proprietary format converted to spreadsheet tables.

Example: Experimental research data will be stored in the native format and be exported in either Excel, Tif, or ASCII.

5. The Work's data and metadata standard(s)

Data and metadata standards can include information that others need to understand and contextualize your Work. Information about the data can include whether the data was created automatically or manually, file and folder naming conventions, any taxonomies/ontologies/controlled vocabularies used, any data standards associated with the Work such as software or data standards, any tools needed to understand the Work.

University Libraries

Example: The DLM software produced by the project will conform to accepted community best practices, including version control (using git) with tagging of major releases, a permissive open-source license (Apache 2), public availability in a community repository (Github), inline comments, reference and tutorial documentation with download and installation instructions that is available from within the software and from a community website, and extensive test coverage. As part of this project the DataONE development team will evaluate the utility and level of effort necessary to report data usage statistics in a form that is consistent Created using the DMPTool. Last modified September 15, 2014 with COUNTER standards. This would permit meaningful comparisons between metrics for data usage and other types of online resources, including traditional serial and monographic publications.

6. **The Work's size and scope**

What is the total size of the Work to be deposited and what are the sizes of the individual components. For example, the Work to be deposited is a total of 450GB of which 20GB of text documents, 30GB of csv spreadsheets and 400GB of tiff images.

Example: With the exception of long-term instrumentation data, most data sets collected by EPSCoR will be small in size (<10 MB). Currently, we have no more than 25 GB of data that was collected.

7. **The Work's legal and ethical compliance**

Is the Work or any part of the Work regulated by policy or law (e.g. classified data, specific handling requirements, IRB/human subject research? If so, has the Work been desensitized to be deposited in the UCL Data Archive, which makes Works publicly accessible. Is the Work or any part of the Work owned by someone else? If so, what are the conditions of use, sharing, and dissemination? Will there be intellectual property (e.g. patent, copyright) rights for the Work or any parts of the work requiring restrictions on access and sharing? Does your Work need to be removed from the UCL Data Archive on a certain date or be added to the Data Archive on a certain date?

University Libraries

Example: NSF will be allowed to use the data for purposes of assessing future programs. No restrictions apply.

Example: Data use by other investigators in subsequent research requires citation and a statement acknowledging the PI's lab for collection the data and NSF for funding the project. Publically assessable data has been de-identified to maintain confidentiality and follow HIPAA regulations. No other restrictions shall be imposed on data use.

8. **Sharing and accessing the Work**

If your data are stored in an uncommon or proprietary format, will they be converted to a more common nonproprietary format for sharing? Do you need to reference a tool or software to read the data? Will you deposit just the raw data, just the analyzed data or both the raw and analyzed data? Beyond the assigned persistent citable link, do you have additional requirements for citing this Work? Do you want to share the citation to your publication(s) with users of your data?

Example: The data will be scanned original questionnaires and survey forms that can be read by any text editor. The data will be in the nonproprietary format of txt. The data submitted will be the analyzed data that has been desensitized to remove any identifiers. The citation to the publication is: J.L. Philips (2014). Chick survival in marsh versus lowland settings following hurricane inundation. Ecology 27(4), 725, DOI: 10.5062/rldki.2014.27456.