Song sparrows LDP Internship

A Data Management Plan created using DMP Assistant

Creator: Janaina Serrano

Principal Investigator: Peter Arcese, Laura Pollock

Data Manager: Janaina Serrano, Joseph Burant, Laura Pollock

Project Administrator: Peter Arcese, Joseph Burant, Laura Pollock

Affiliation: McGill University

Funder: NSERC

Template: Alliance Template

ORCID iD: 0000-0002-8097-482X

ORCID iD: 0000-0002-6004-4027

Project abstract:
Long-term time series of observations from focal populations are among the most valuable data ecologists collect, and their preservation is central to understanding historic and future changes in demography and dynamics of wild species. Since 1975, Professor Peter Arcese and his colleagues have monitored a population of Song Sparrows (Melospiza melodia) on Mandarte Island in the Haro Strait near Vancouver Island, British Columbia. Through their intensive monitoring, Arcese and his team have collected more than 40 years of demographic, life-history, and genetics/pedigree data as well as detailed information on the location of individual birds from nest departure to death. These data have been used to link demography and evolutionary theory, and to address a variety of fundamental questions. Corresponding climatic data are already available online, providing ample opportunity for future analyses. In addition to the long-term monitoring on Mandarte, the researchers have also collected similar short-term data from more than a dozen other islands. This internship will work to harmonize these different sources of data, and will involve data extraction from paper records and non-proprietary file formats, digitization, data cleaning and validation, and the generation of metadata.

Identifier: 9439

Start date: 20-06-2022

End date: 31-08-2022

Last modified: 20-06-2022

Created using DMP Assistant. Last modified June 20, 2022

1 of 3

Song sparrows LDP Internship

Data Collection

What types of data will you collect, create, link to, acquire and/or record?

The data are variously collected by Frank S. Tompa 1960-63, James N.M. Smith 1973-81, Peter Arcese 1981 - 87, Wesley Hochachka 1988-89, Alice Cassidy 1990, and again by Peter Arcese and various students, post-docs, and employees 1991-present.
During this internship, I will compile data from territory cards into tables (.csv) that will be linked to other existing tables (nest cards). This data typically contains band.prefix (first 4 digits of band number), band.suffix (last 5 digits of band number), sex, territory id, date (year, month, day), time, presence_absence (0, 1), behaviour (multi-level, restricted factor indicating the incidental behaviour observed) and notes.

What file formats will your data be collected in? Will these formats allow for data re-use, sharing and long-term access to the data?

.csv format

What conventions and procedures will you use to structure, name and version-control your files to help you and others better understand how your data are organized?

territorycards_songsparrows_20220620_v01.csv

Documentation and Metadata

What documentation will be needed for the data to be read and interpreted correctly in the future?

descriptive – title, author, abstract and keywords
administrative – rights and permissions and data on formatting
structural – explanations of e.g. tables within the data

How will you make sure that documentation is created or captured consistently throughout your project?

We will develop and implement clear metadata documentation protocols prior to starting data collection which will clearly communicate expectations, standards, and processes for capturing and implementing documentation to support the research project. Example topics that these protocols will cover include file naming conventions, file versioning, folder structure, and both descriptive and structural metadata. Documentation will be kept in clearly named folders within the research project folder space and will be accessible by all researchers and staff.

If you are using a metadata standard and/or tools to document and describe your data, please list here.

https://researchdata.library.ubc.ca/plan/dmp-assistant/

Storage and Backup

What are the anticipated storage requirements for your project, in terms of storage space (in megabytes, gigabytes, terabytes, etc.) and the length of time you will be storing it?

Since we will only have tabular data, 10Gb storage on Dryad should be enough.

How and where will your data be stored and backed up during your research project?

We are storing the files at https://github.com/janaserrano/songsparrows and Dryad.

How will the research team and other collaborators access, modify, and contribute data throughout the project?

We are updating files on github as they are modified and finally make a Dryad data publication when the dataset is complete.

Preservation

Where will you deposit your data for long-term preservation and access at the end of your research project?

We are storing the files at https://github.com/janaserrano/songsparrows and Dryad, providing a doi.

Indicate how you will ensure your data is preservation ready. Consider preservation-friendly file formats, ensuring file integrity, anonymization and de-identification, inclusion of supporting documentation.

We will use .csv for tabular data, .pdf for text data and .md for metadata.

Sharing and Reuse

What data will you be sharing and in what form? (e.g. raw, processed, analyzed, final).

We will be sharing raw (digitized territory cards in .pdf) and processed (tables in .csv).

Have you considered what type of end-user license to include with your data?

We are using CC0.

What steps will be taken to help the research community know that your data exists?

We are depositing the files in Dryad and creating a data paper on the compiled data. The Living data project also ensures communications with peers through courses and conferences.

Responsibilities and Resources

Identify who will be responsible for managing this project's data during and after the project and the major data management tasks for which they will be responsible.

Janaina Serrano will be responsible for managing the project during her LDP internship. Peter Arcese is the owner of the data. Laura Pollock and Joey Burant are Living Data Project members and will be responsible for managing the data in the future.

How will responsibilities for managing data activities be handled if substantive changes happen in the personnel overseeing the project's data, including a change of Principal Investigator?

Laura Pollock and Joey Burant are Living Data Project members and will be responsible for managing the data in the future.

What resources will you require to implement your data management plan? What do you estimate the overall cost for data management to be?

The Living Data Project is an ongoing data rescue initiative funded by NSERC. Any necessary funding will be provided by it.

Ethics and Legal Compliance

If your research project includes sensitive data, how will you ensure that it is securely managed and accessible only to approved members of the project?

Not applicable.

If applicable, what strategies will you undertake to address secondary uses of sensitive data?

Question not answered.

How will you manage legal, ethical, and intellectual property issues?

.

Created using DMP Assistant. Last modified June 20, 2022

3 of 3