
Plan Overview

A Data Management Plan created using DMPonline

Title: Old crops for new insights: agricultural systems of broomcorn and foxtail millet cultivation in Iberia (MILLET)

Creator: Andrés Teira-Brión

Principal Investigator: Andrés Teira Brión

Contributor: Michael Charles, Amy Bogaard, Rowan Wilson

Affiliation: University of Oxford

Funder: European Commission

Template: Horizon 2020 DMP

ORCID iD: 0000-0001-8997-1425

Project abstract:

Little is known about the cultivation of broomcorn and foxtail millet during the second millennium BC despite its occurrence at archaeological sites. The EU-funded MILLET project will therefore produce evidence from ancient and modern samples from sites in the Iberian Peninsula to better understand its role in agricultural systems. A detailed reconstruction of growing conditions and production techniques will also be recreated by combining stable carbon and nitrogen isotope values with functional weed ecology. The aim is to gain a fresh perspective on the agroecology of these crops, thereby advancing the study of past and present agricultural systems. The project's findings will also inform work on climate change, food, genetic diversity and sustainability associated with traditional knowledge.

ID: 95717

Start date: 10-01-2022

End date: 19-03-2024

Last modified: 14-05-2024

Grant number / URL: <https://cordis.europa.eu/project/id/101018935>

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Old crops for new insights: agricultural systems of broomcorn and foxtail millet cultivation in Iberia (MILLET) - Initial DMP

1. Data summary

Provide a summary of the data addressing the following issues:

- **State the purpose of the data collection/generation**
- **Explain the relation to the objectives of the project**
- **Specify the types and formats of data generated/collected**
- **Specify if existing data is being re-used (if any)**
- **Specify the origin of the data**
- **State the expected size of the data (if known)**
- **Outline the data utility: to whom will it be useful**

The goals of the data collection will be to investigate the agricultural practices of broomcorn millet and foxtail millet during prehistoric times based on information from isotopic analysis of experimental farming, samples from archaeological sites, and the recording of different agricultural practices from ethnographic fieldwork. The project will obtain its own data that are not part of any accessible repository. Two main types of data will be generated: a) data from stable isotope analyses that will be useful as a baseline of crop growing conditions (WF1), and b) data from ethnographic interview recordings that will be used to interpret past agricultural practices and for the dissemination of results (WF2).

WF1 Isotopic analysis: The results of stable isotope analyses will be published in articles with golden or green access, in open databases repositories (ORA-Data or *Open Science Framework*). The standard format to share the information in open repositories will be .csv format or in other free-standard formats that enable the information according to FAIR principles (<https://fairsharing.org>).

WF2 Ethnographic research: In the ethnographic fieldwork, a total of 10 people will be interviewed, 5 of whom will be interviewed for an estimated total of 5 hours of audio and video recording per person in 1 or several sessions of variable duration between the months of May to October. In addition, it is estimated that a further 5 interviews will be conducted, which will not be recorded, but will be transcribed in plain text (.odt or .txt). The information from the interviews will be captured using 2 video cameras, an audio recorder and a photographic camera (pictures will be created in .jpg and .raw formats). Handwritten notes or notes on computer will be taken. It is estimated that a storage capacity of 20-50 Gb will be required.

The video and audio recordings will be made by an audiovisual company for the creation of a documentary and other audiovisual products for dissemination. The rights of storage and use by the project and the University of Oxford will be regulated by a collaboration agreement.

The interviews will be protected by a personal data protection policy, will be protected by a pseudonymisation process, must have the consent of the participants, and will therefore not be accessible to the public. The outcomes of the interviews will be open access, e.g. the making of a documentary film, or the results of the interviews will be published in scientific journals. Scientific papers will be published following a Green Access policy.

2. FAIR data

2.1 Making data findable, including provisions for metadata:

- **Outline the discoverability of data (metadata provision)**
- **Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?**
- **Outline naming conventions used**
- **Outline the approach towards search keyword**
- **Outline the approach for clear versioning**
- **Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how**

WF1 Isotopic analysis: Documentation (D) and Metadata (MD)

The results of staple isotopes (**D1**) will be carried out at the Stable Isotope Laboratory, part of the Research Laboratory for Archaeology and the History of Art (RLAHA), and part of the School of Archaeology at the University of Oxford (UK). The data will be assigned a persistent Digital Object Identifier (DOI).

For a better understanding of the data, associated Metadata (**MD1**) documents will be generated in order to make them understandable. The metadata will comply with the standards used in stable isotope analysis.

WF2 Ethnographic research: Documentation (D) and Metadata (MD)

The interviews will take one day, but in other cases the growth of the crops to be studied can be monitored, so it will take 2-3 more days to contact over 4 months between late spring and early autumn in 2022 or 2023.

The following **Documents** will be obtained from this action:

- D2. Audio files of the interviews will be recorded in .wav and .mp3 formats.
- D3. Video files of the people interviewed and of the agricultural practices.
- D4. Photographic files.
- D5. Information in written format (paper or digital).
- D6 Geospatial data.

The documents will be used for the following **Outcomes** (O):

- O1: production of a short documentary on the cultivation of millets.
- O2: audio-visual dissemination materials.
- O3: papers and other scientific publications.

Some documents (D2, D3, D4, D5, D6) and outcomes (O1, O2) will imply the consent of the interviewees. Capturing data files (D2, D3, D4, D5) will be protected and will not be freely accessible to third parties or external users in order to preserve the personal data of the participants. In the case of output O3, the data will be published in scientific papers or informative texts with the prior consent of the interviewees, or otherwise after pseudonymisation of their personal data. The metadata of O3 (scientific papers) will explain the nature of the information available and its digital registration in the publication itself.

2.2 Making data openly accessible:

- **Specify which data will be made openly available? If some data is kept closed provide rationale for doing so**
- **Specify how the data will be made available**
- **Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?**
- **Specify where the data and associated metadata, documentation and code are deposited**
- **Specify how access will be provided in case there are any restrictions**

A website will be created hosted by University of Oxford, providing real time access to the information (WF1) and specific outputs (WF2). The website will also provide details of how the data was generated and how it can be shared and used.

WF1:

D1 Stable isotopes. Results of stable isotopes results (measurements and numeric data) of modern and archaeological remains and metadata will be stored to an online Data Archiving (ORA-Data) hosted by the University of Oxford or in an *Open Science Framework* (OSF) account (osf.io/xqcw4), which provides long-term preservation and curation, and a DOI to all files to serve as a basis for data reuse.

D2 models. The models created of millets growing conditions, will be relevant for determining agrarian management (manure, fallow, moisture) and for archaeobotanical research will be accessible in ORA-Data or OSF and will be published in open access journals.

WF2: Documents of ethnographic research will be closed, but will be the core to create free access outputs:

- O1 Documentary: It will be freely accessible on University of Oxford website and, additionally, in platforms such as youtube or vimeo.
- O2 Dissemination outputs (audio, video or written) will be accessible through the website hosted by the University of Oxford.
- O3: Open access policy will be associated with scientific publications rights. Preference will be given to green access to articles through preprint publication in cases of payment for publication in golden access.

2.3 Making data interoperable:

- **Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.**
- **Specify whether you will be using standard vocabulary for all data types present in your data set, to allow interdisciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?**

WF1:

A series of values of stable carbon and stable nitrogen isotope ratios, archaeological context, and associated information expressed as comma-separated plain text in .csv format shall be obtained.

The isotopic analysis databases will collect enough data for contextualisation, which will be stored in .csv files to form the data dedicated to:

- Archaeological materials: archaeological context information (archaeological site, stratigraphic unit, location, chronology) and isotopic results (stable carbon and stable nitrogen isotope ratios values obtained).
- Farming experiment: materials (experiment variables, growing conditions, location) and isotopic results (stable carbon and stable nitrogen isotope ratios).

- Isotopic effect experiment: experimental variables (materials, firing atmospheres) and isotopic results (stable carbon and stable nitrogen isotope ratios).

The specific conditions of some data could be published in an article in the Open Archaeology Data journal, which will provide a link to the repository hosted and metadata associated.

WF2:

The data of this task may not be interoperable due to the nature of its registration and the protection of personal data.

2.4 Increase data re-use (through clarifying licenses):

- **Specify how the data will be licenced to permit the widest reuse possible**
- **Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed**
- **Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why**
- **Describe data quality assurance processes**
- **Specify the length of time for which the data will remain re-usable**

WF1:

Datasets will be created at the same time as research findings are published. Standard non-commercial formats will be used to maximise their distribution and reuse. The datasets will be published with a DOI and they can be download from Ora-Data and OSL data repositories without time limit or usage restrictions. The databases will be available under the Open Database License <http://opendatacommons.org/licenses/odbl/1.0/>. Any rights in individual contents of the database are licensed under the Database Contents License: <http://opendatacommons.org/licenses/dbcl/1.0/>

WF2:

Interview data (D1, D2, D3, D4). We will communicate our strategy for maintaining the confidentiality of protected personal data to all participants. Transcribers and people involved in the documentary will be required to sign a collaboration and release agreement for the use of the materials they generate to the project and to the University of Oxford. Both audio and text files will be encrypted during fieldwork and on transfer to the University of Oxford.

Pseudonyms will be used for participants who do not authorise the disclosure of personally identifiable information.

D1. All dissemination outputs will be of open access on the website under a Creative Commons CC BY-SA license, which will approve the sharing and adaptation of material for free, citing the project as a source of information.

3. Allocation of resources

Explain the allocation of resources, addressing the following issues:

- **Estimate the costs for making your data FAIR. Describe how you intend to cover these costs**
- **Clearly identify responsibilities for data management in your project**
- **Describe costs and potential value of long term preservation**

The responsibility for data collection will lie with the principal investigator and collaborators. Andrés Teira-Brión as Principal investigator (PI) has a duty to report on ethical issues, to monitor compliance with good working practices, and in relation to the collection, storage and sharing of data. ILLA BUFARDA S.L.A. collaborates in video and audio recording and documentary filming. The data will be hosted in open access repositories or use policies under the umbrella of the University of Oxford, therefore there will be no additional costs to comply with FAIR principles. The data will be hosted in Oxford's institutional data archive, [ORA-Data](#), is currently free of charge to researchers, or in other repositories that allow free access (OSF).

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

OneDrive for Business, provided as part of the University's Nexus365 offering, has been approved by the University's Information Security team for the storage of research data. The recording of data on electronic equipment during fieldwork (audio recorders, video or photographic cameras) shall be encrypted and password protected. Devices containing personal data will be protected by whole disc encryption. The Code42 Cloud backup service will be used to backing up single-user computes, or laptop where a single user can take ownership of the backups during fieldwork.

Identifiable data (including oral or written consent forms) will be stored on the University of Oxford's servers and will be held in accordance with its data protection policies <https://researchsupport.admin.ox.ac.uk/policy/data/>.

Research data will be stored for 5 years after publication or public release of the work of the research according to the MSC grant agreement. The purpose of storing this information is to be able to follow up the cultivation with the people interviewed. To manage personal data the project will include pseudonymised data as part of a "data minimisation" strategy aimed at minimising the risks of a data breach for data subjects. Scientific publications will be hosted on their own websites.

The researcher, the research team, the supervisor, and authorised personnel within University of Oxford will have access to the research data. Information could be shared with third parties (e.g., collaborators, translators) during the preparation of results and outputs. The people from the project partners involved in the filming of the documentary will generate audio and video files and will have access to them during the production of the documentary.

Research data may be also transferred to, and stored at, a destination outside the UK and the European Economic Area (e.g., in online databases). Identifiable data will be removed whenever possible, and any data transfer will be done securely and with a similar level of data protection as required under UK law.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

WP2: Ethnographic Research

Interviews will be carried out during the project and will raise potential ethical issues associated with collecting, managing, storing, and sharing personal data for both researchers and research participants. This data comprises contact information (e.g., name or address) and recordings attributed to an identified or identifiable natural person (audio, video, or photography). To make sure that personal data of the participants are protected, the project comply with the following principles:

Transparency

Participants will be informed about how the project will use, store and sharing their personal data. Research is a task that is performed in the public interest. Databases and data collections will be used for academic research only, they will not be used for commercial purposes.

Collecting data

Personal data will be collected during ethnographic interviews which will be recorded. Prior to data collection, participants will sign a consent form and authorise the recording of personal data for the uses stipulated in the project. No sensitive data will be collected.

Stored data

Identifiable data (including oral or written consent forms) will be stored on the University of Oxford's servers and will be held in accordance with its data protection policies <https://researchsupport.admin.ox.ac.uk/policy/data/>. Research data will be stored for 5 years after publication or public release of the work of the research.

Access data

Personal contact data access will have a restricted access and protected by institutional signature. This contact information will not publicly available. The researcher, the research team, the supervisor, and authorised personnel within University of Oxford will have access.

Manage data

The University of Oxford is the data controller with respect to the personal data, and as such will determine how personal data is used in the study. The University will process personal data for the purpose of the research outlined below.

Use of personal data:

For the most part, data will be managed and used in a way that protects the confidentiality of research participants. There are two levels of use of personal information: contact information, and identification of the subject through the scientific and dissemination actions of the Project.

- **Contact information and data collections** will be stored with the purpose of to be able to follow up the cultivation with the people interviewed. This information will not be shared with third parties outside the project.
- **Identified or identifiable natural person.** This information may be made public through publication of research and dissemination outputs with the oral or written consent of the participants. Information could be shared with third parties (e.g., collaborators, translators) during the preparation of results and outputs. Research data may be also transferred to, and stored at, a destination outside the UK and the European Economic Area (e.g., in online databases). Identifiable data will be removed whenever possible, and any data transfer will be done securely and with a similar level of data protection as required under UK law.

MILLET project has received ethics approval from a subcommittee of the University of Oxford Central University Research Ethics Committee (Ethics reference: **R79927/RE001**).

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

The researcher, the research team, the supervisor, and authorised personnel within University of Oxford will have access to the research data. Information could be shared with third parties (e.g., collaborators, translators) during the preparation of results and outputs. Research data may be also transferred to, and stored at, a destination outside the UK and the European Economic Area (e.g., in online databases). Identifiable data will be removed whenever possible, and any data transfer will be done securely and with a similar level of data protection as required under UK law.

Planned Research Outputs

Audiovisual - "The Lost Millets"

This documentary deals with researcher Andrés Teira Brión's investigation into a cereal that is almost extinct in the northwest of the Iberian Peninsula. In his research, Andrés takes us to meet the last millet farmers, and through their voices, we get to know the gradual abandonment of the rural world in this contemporary world. This work was recorded throughout the year 2023 in different locations in Portugal, Asturias and Galicia.

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
The Lost Millets	Audiovisual	2024-04-18	Open	None specified	888 MB	Creative Commons Attribution Non Commercial Share Alike 4.0 International	None specified	No	Yes