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 Template

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

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 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/xgcw4), 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 accesible 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 inter-disciplinary 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.

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

1. Data summary

State the purpose of the data collection/generation

To understand how millet were managed in the past, it is necessary to document a broad diversity of farming practices from their sowing to their use, establishing a baseline of data and cultivation techniques that allow reliable comparisons between archaeological data (weed ecology and isotope analysis) and those obtained in agriculture in the present.

Explain the relation to the objectives of the project

The aim of this project will be twofold: 1) register the agricultural practices last farmers to understand the synergies behind techniques and recording the traditional agricultural knowledge associated with millets before their disappearance, 2) create outputs to disseminate the traditional knowledge of millet growing to a broad audience.

Specify the types and formats of data generated/collected

The interviews and their recording (photographs, audio, and video) will be carried out with your prior oral permission recorded on audio or written on a consent form.
Identifiable data (including oral or written consent forms) will be stored separately. When personal data (such as contact details) is collected will be used exclusively used to keep in communication during the project.

Specify if existing data is being re-used (if any)

No reuse of data.

Specify the origin of the data

Interviews with farmers.
Weed ecology and isotope analysis of broomcorn and foxtail millet.

State the expected size of the data (if known)

Photo, video and audio recordings: Up to 1 Tb
Results of the analyses: 1 Gb

Outline the data utility: to whom will it be useful

MILLET will build bridges between the past and the new challenges of the agriculture and food sustainability, communicating the results from academics to non-specialised audiences, for which the fellow will create specific audience profiles (education, stakeholders and general public).

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

The data will be hosted in Oxford's institutional data archive,ORA-Data, or in other repositories that allow free access (OSL).

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?

The project envisages the publication of results in open access. The published information will have a DOI provided by the journals where the results are published. In parallel, the data will be published through the Open Science Framework, also accessible through a DOI.
In the case of ethnographic information when they have the consent of the interviewees.

Outline naming conventions used

The project include a pseudonymisation process of personal data. This personal data will be under the umbrella of an alphanumeric code to protect their identity consisting of: the project reference (i.e. MILLET001), the record type (i.e. AUD, FOT or VID), if it corresponds a single individual or a group (i.e. IN or GR), and a correlative number (001 and followings). For example: MILLET001.AUD.GR.004.' This identifier allows the identity of the participants to be protected and at the same time serves to identify the source of the information (e.g., to be able to observe differences in the information between each interview in the results outputs).

Outline the approach towards search keyword

The metadata include basic information about the project, international codes, keywords and information about the location of the resources.

Outline the approach for clear versioning

Each metadata update shall consist of an update date and a correlative version number.

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

The data documentation will describe how, why and when the data were created/captured. This information includes information on the software used to create, edit or process the files which will be in open source format. Where internationally used codes are used, information for their understanding will be provided. Simple and accessible language will be used in the description of the data.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

The results of the isotopic analyses on archaeological and experimental samples will be freely available in the repositories provided for that purpose.

The results of ethnographic interviews, as long as they have the consent of the interviewees, and under a pseudonym if you do not want content with personal data to be disclosed. These results will be reworked in various outputs (scientific publications, video, audio and photo recordings) and will therefore not be accessible as recorded.

Specify how the data will be made available

The data will be hosted in ORA-Data, and other repositories that allow free access such as Open Science Framework (OSF) and Figshare.

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)?

Data documentation describes how, why and when the data were created/captured, in which format and in which software and version they were generated. The data shall be in open source formats. The databases where the data will be hosted will allow access and downloading of the data.

Specify where the data and associated metadata, documentation and code are deposited

The data will be stored in ORA-Data, and personal OSF and figshare repositories.

Specify how access will be provided in case there are any restrictions

Data for which the necessary consent and permissions have been granted will be freely accessible, no access filters will be established for specific users.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

The data can be downloaded in freely accessible formats such as .xml, with essential provenance information and information about each data item in the metadata in .txt format, making it interoperable for further use.

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

The vocabulary used will be standardised for isotopic analysis to facilitate communication and interoperability.
In the case of ethnographic data, if mapping the ontologies is necessary, the associated metadata will contain information on the associated data type.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

Outputs will be of open access on the website under a Creative Commons CC BY-SA license.
The publications will be accessible through golden access or green access with the corresponding licence of each publisher.

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

N/A

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

N/A

Describe data quality assurance processes

N/A

Specify the length of time for which the data will remain re-usable

N/A

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

There will be no costs associated with the storage of information.
Oxford's institutional data archive, ORA-Data, is currently free of charge to researchers.

Clearly identify responsibilities for data management in your project

The responsibility for data management will be that of the principal investigator of the project: Andrés Teira-Brión.
The data stored in ORA-Data will be administered by the person appointed by the University of Oxford for this purpose.

Describe costs and potential value of long term preservation

Research data will be stored for 5 years after publication or public release of the work of the research. The purpose of storing this information is to be able to follow up the cultivation with the people interviewed.

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.

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

This study 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)

No considered.

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

1. Data summary

State the purpose of the data collection/generation

The project obtained its own data that were not part of any accessible repository. Two main types of data were generated: a) data from stable isotope analyses that was useful as a baseline of crop growing conditions (WF1), and b) data from ethnographic interview recordings that was used to interpret past agricultural practices and for the dissemination of results (WF2).

Explain the relation to the objectives of the project

The goals of the data collection were 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.

Specify the types and formats of data generated/collected

1. Video recordings.
2. Audio recordings.
3. Handwritten notes.
4. Numerical and text data.

Specify if existing data is being re-used (if any)

Any existing data have been re-used.

Specify the origin of the data

WF1 isotopic analysis: The results of the stable isotope analyses are in the process of being published in gold and green access and open database repositories (ORA-Data and GitHub). The standard format for sharing information in open repositories will be .csv or other free standard formats that allow FAIR principles information (https://fairsharing.org).

WF2 Ethnographic research: The PI interviewed a total of 13 people during 2022-2023. Video was recorded and photographs were taken in 3 individual and group interviews with 1, 2 and 3 people (6 people in total). These interviews formed part of the production of the short documentary. The interview with 2 people was audio recorded. Information from the interviews was captured using 1, an audio recorder and a camera (photographs were taken in .jpg and .raw format). Handwritten notes were taken from the other 5 people.

The video and audio recordings were made by the production company Illa Bufarda. The rights of storage and use by the production company and the project through the University of Oxford were covered by a co-ownership agreement. Audio interviews and written notes not included in the documentary were made by

the PI.

State the expected size of the data (if known)

The size of the data generated is 150 Gb.

Outline the data utility: to whom will it be useful

The recordings testify to the traditional ecological knowledge of the last remaining millet farmers in western Europe. The data recorded from the interview is in line with knowledge transfer and citizen participation in science. Both aspects are essential to face the challenges of preserving our agricultural heritage and biodiversity, but also to open new horizons towards new forms of food security.

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

In preparation

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?

The documentary ‘The lost millets’ as well as being on the Vimeo platform used for its reproduction has the doi: 10.5287/ora-e9e1vm49a

Outline naming conventions used

In preparation

Outline the approach towards search keyword

In preparation

Outline the approach for clear versioning

In preparation

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

In preparation

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

In preparation

Specify how the data will be made available

In preparation

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)?

In preparation

Specify where the data and associated metadata, documentation and code are deposited

In preparation

Specify how access will be provided in case there are any restrictions

In preparation

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

In preparation

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

In preparation

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

The documentary 'The Lost Millets' has been licensed under CC Attribution-NonCommercial-ShareAlike (CC BY-NC-SA).

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

Data from academic publications resulting from the project are available for re-use once published.

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

All data derived from the project may be reused as long as the source of the information is acknowledged. In the case of the documentary 'The Lost Millets', commercial use is not allowed for the part of the footage that affects the musical theme of the credits.

Describe data quality assurance processes

- Video recording
- Audio recording
- Interviews
- Photographs
- Databases

Specify the length of time for which the data will remain re-usable

The published data may be used indefinitely.

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

There were no costs associated with making project data FAIR. Repositories provided by the host institution have been used. The University of Oxford's agreement with the UK government and selected publishers to publish articles under golden access has also been used.

Clearly identify responsibilities for data management in your project

The responsibility for data collection rested with the principal investigator and collaborators. Andrés Teira-Brión, as Principal Investigator (PI), will be responsible for reporting on ethical issues, monitoring compliance with good working practices, and the collection, storage and sharing of data.

The data will be hosted in open access repositories or use policies under the auspices of the University of Oxford, so there will be no additional cost to comply with FAIR principles. The data are currently hosted in Oxford's institutional data repository, ORA-Data.

Describe costs and potential value of long term preservation

In preparation

4. Data security

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

Oxford University Research Archive (ORA) has been used as secure storage of final outputs.

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

Interviews conducted during the project raised potential ethical issues related to the collection, management, storage and sharing of personal data for both researchers and research participants. This data included contact information (e.g. name or address) and recordings attributed to an identified or identifiable natural person (audio, video, or photographs). To ensure that participants' personal data was protected, the project adhered to the following principles:

Transparency: Participants were informed about how the project would use, store and share their personal data.

Data collection: Personal data collected during ethnographic interviews were in some cases registered. Prior to data collection, participants gave verbal consent and permission for personal data to be collected for the purposes specified in the project. Sensitive data was not collected.

Data collected: Identifiable data (including verbal or written consent forms) were stored on the University of Oxford's servers and are held in accordance with the University's data protection policy https://researchsupport.admin.ox.ac.uk/policy/data/. Research data will be retained for 5 years after publication or release of the research.

Access to data. Access to personal contact details is restricted and protected by an institutional signature. This contact information is not available to the public. The researcher, research team, supervisor and authorised staff within the University of Oxford have access.

Data management: The University of Oxford is the controller of the personal data and as such will

determine how the personal data is used in the study. The University will process personal data for the purposes of the research as set out below.

Use of personal data: There were two levels of use of personal data: contact information and identification of the subject through the scientific and dissemination activities of the project. Contact information and data collections have been stored for the purpose of following up with respondents. This information will not be shared with third parties outside the project.Identified or identifiable individuals. This information may be made public through the publication of research and dissemination results with the oral or written consent of the participants. For example, information will be shared with third parties (e.g. collaborators, translators) during the preparation of results and outputs. Research data may also be transferred to and stored outside the UK and the European Economic Area (e.g. in online databases). In this case, identifiable data will be removed wherever possible and any data transfer will be secure and with a level of data protection similar to that required under UK law.

6. Other

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

The MILLET project has received ethical approval from a sub-committee of the University of Oxford Central University Research Ethics Committee (ethics reference: R79927/RE001).

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	DOI	Type	Release date	Access level	Repository(ies)	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