# University of Leeds Data Management Plan (DMP)

## Project Details

Researcher name: Sebastián Cisternas

Project title: Earth and Environment

Faculty: School of Geography

KRISTAL reference number (if applicable)

Supervisor name(s) (if applicable) Professor Duncan Quincey and Dr. Andrew Ross

Funder: Agencia Nacional de Investigación y Desarrollo (ANID), Chilean Government

Scheme

Application submission date

Research start date

Research end date

Ethical review number

DMP review date (we recommend you review your DMP regularly)

Version 2.4. Last revision: 10/01/2025

**\*But I do not have any data\*** Anything can become research data if it is used for research purposes – data is more than numbers on a spreadsheet. Think creatively about the materials you are using and producing: what could be shared with other researchers who are interested in your work; what could be reused to produce new insights? Any evidence or material which underpins or sheds light on your findings, your academic publications, your thesis, or your project can be considered research data.

## Provide a brief overview of your project including proposed research methods

1. **What data will be produced? Are you using data from other sources?**

* What physical data will you study? What data will be produced (e.g. artefacts, samples, paper archives)
* Are you working with [personal data?](https://leeds365.sharepoint.com/:u:/r/sites/DataProtectionInformation/SitePages/Personal-Data.aspx?csf=1&web=1&e=tzvtC0)
* What digital data will you generate? (e.g. field-notes, images, spreadsheets, audio interviews, survey data, an annotated bibliography)
* What original software will you generate?
* What third party data will you reuse?
* Have you considered whether there may be any costs associated with this activity, e.g. costs to access third party data?

1. **Where will data be stored? Include file formats and approximate volume.**

* How much data you will produce over time? Do you have enough storage?
* What file formats and software will you use?
* Do you have a logical file naming convention and directory structure?
* How will you use versioning so you can identify the current version of documents/data?
* How will data generated in the field be saved to safe University storage?
* Have you considered whether there may be any costs associated with this activity, e.g. costs to store your data?

1. **Access to data during the project. Give details of collaborators and agreements.**

* Please contact the Research and Innovation Service to discuss whether a data sharing/processing agreement is necessary for your research project? This will be relevant if you are sharing your data with third parties, e.g. during collaboration, anonymisation, transcription, translation.
* Have you discussed data sharing with your research collaborators/supervisor?
* Who needs to access data during the research? How will they access data?

1. **Ethics and legal compliance: Any ethical, contractual, special, sensitive or consent issues?**

* Have you read the University’s [Information Protection Policy](https://it.leeds.ac.uk/it?id=kb_article&sysparm_article=KB0011140)? Data must be assessed for sensitivity and storage in line with this policy
* Are you familiar with the University’s advice on [data protection and GDPR (General Data Protection Regulation)](https://leeds365.sharepoint.com/:u:/r/sites/DataProtectionInformation/SitePages/Home.aspx?csf=1&web=1&e=J5QROp)?
* Does your research funder have specific data management and sharing requirements?
* Are there other policies and protocols you need to be aware of and observe? For example, NHS codes of practice?
* Will you anonymise or pseudonymise your data?
* Does and data need to be destroyed e.g. due to contractual or ethical obligations? When and how? A data retention table could be useful (see example table at end of the document)
* How and where will you record any participant consents and/or contractual requirements which impact data management and sharing?
* Have you considered whether there may be any costs associated with this activity (e.g. anonymisation, transcription, translation)?

1. **How will data be documented and described?**

* Will others understand your data? Write documentation. Make sure table and spreadsheet values are clearly labelled. Think about your methodologies and protocols.
* What information about data collection methodology will be recorded?
* Is it important for the research to be reproducible? Why/why not?
* What additional documentation will be required?
* Will you write software? Where will this be documented and stored for future use?
* Have you considered whether there may be any costs associated with this activity, e.g. for data cleansing, additional staffing etc?

1. **Training and support**

* What training will you/your team need for data gathering, organisation, analysis or presentation?
* Are there relevant courses available at the University? Online?
* Who can provide support?
* Have you considered whether there may be any costs associated with this activity, e.g. training courses?

1. **What are the plans for data sharing beyond project partners? How is this impacted if some of your data needs to be restricted?**

* Have you considered reasons for and against sharing data?
* What steps will you take to minimise personal data collection and anonymise or reduce identifiability?
* Will data be openly available to everyone, or will there be access restrictions?
* If your research involves people, have you clarified onward sharing of data in your Participant Information Sheet?
* Can your data be released immediately, or should you embargo (delay access to) the data?
* Will you use a data repository? Which one? Are there subject specific data repositories in your field?
* Are you familiar with the [FAIR Principles (findable, accessible, interoperable, reusable)](https://library.leeds.ac.uk/info/14062/research_data_management/66/sharing_data/2) and the University’s [Open Research Statement](https://ris.leeds.ac.uk/research-excellence/open-research-statement/)?
* Have you considered whether there may be any costs associated with this activity, e.g. long-term storage and sharing, including repository costs?

1. **Will any sensitive or potentially commercially exploitable intellectual property be generated?**

* Will you be applying for a patent?
* Will your research have commercial applications?
* Do you need to contact the Research and Innovation Service?
* Have you read the University’s [Intellectual Property Policy](https://secretariat.leeds.ac.uk/policies-procedures-and-codes-of-practice/)?
* Would release of your data, code or other project materials pose any security risk?

1. **Who is responsible for managing the data? What resources will you need?**

* On projects with complex data management requirements, distinct types of roles should be specified.
* Are sufficient resources (skills, people, storage, technology) available to deliver your plan?
* Have you considered whether there may be any additional costs, e.g. additional staffing costs for managing the data.
* What will you need?

## Important data safeguards for your live research project

### 1. Ethical review

* Where research involves human participants (eg interviews, focus groups, opinion giving), or human cells or tissue, or has environmental sensitivities, or otherwise requires ethical review (see [Do I need to submit a research ethics application](https://secretariat.leeds.ac.uk/research-ethics/how-to-apply-for-research-ethics-approval/)?), ethical review needs to take place before the research can start. Apply via your Faculty’s Ethical Review Committee.  
  For more information visit the [Application Process](https://ris.leeds.ac.uk/research-ethics-and-integrity/application-process/) page on the Research and Innovation Service website. For any ethical review queries, please use the [contacts on the Faculty Research Ethics Committees](https://secretariat.leeds.ac.uk/research-ethics/faculty-research-ethics-committees/) page on the Research and Innovation Service website.
* When dealing with data undergoing ethical review, your data should be classified according to the [Turing Classification model](https://doi.org/10.6084/m9.figshare.11815224.v4): some examples are available in the [documentation for the LASER (Leeds Analytic Secure Environment for Research) service about Data Risk Tiering](https://lida-data-analytics-team.github.io/laserdocs/docs/laser_info/tiering.html).
* Projects classified as either Tier 3 or 4 are likely to need to use the LASER platform for which costs will need to be budgeted for in your funding application. You will need to make an enquiry about using LASER, to LIDA by emailing dat@leeds.ac.uk. Further information about LASER is available in the [Leeds Analytic Secure Environment for Research](https://lida.leeds.ac.uk/laser/) section of the Leeds Institute for Data Analytics website.  
  NOTE: Please allow for at least 10 working days to finalise costs and your support needs, though turnaround times may vary depending on the complexity of the request.
* Data from social media platforms: include your plans for gathering and sharing data in your ethics application and ensure any analysis and sharing is in line with platform terms and conditions.

### 2. Data Storage, back-up and security

* All material to be assessed for sensitivity and stored on secure, University provisioned storage and servers in line with GDPR requirements.
* Any data captured on encrypted recording devices such as that used for interviews should be uploaded to University storage and deleted from the device as soon as possible.
* University IT arrangements will allow data to be accessed via Microsoft Teams which includes setting up a secure Team area to share materials with team members both inside and outside the University.
* If your data is highly sensitive, see the [Leeds Analytic Secure Environment for Research](https://lida.leeds.ac.uk/laser/) section of the Leeds Institute for Data Analytics website for information on using LASER for storage and access.
* For additional storage advice, please [contact your Business Relationship Manager](https://it.leeds.ac.uk/it?id=kb_article_view&sysparm_article=KB0012310) or [contact the IT Service Desk](https://it.leeds.ac.uk/it?id=nr_ec_dashboard)

### 3. Contractual-Legal Agreements, including consent

* Consent to participate should be sought from all participants at the earliest opportunity and should include plans for data deposit and sharing. In the participant information sheet make sure to cover all the ways you plan to handle data including sharing in an anonymized form in a repository.
* University of Leeds Consent and Participant Information Sheet templates are available from [How to apply for research ethics approval](https://secretariat.leeds.ac.uk/research-ethics/how-to-apply-for-research-ethics-approval/). Ensure your project information, particularly your participant information sheet addresses clearly all the ways the research data will be shared.
* 'Technical', 'Sharing' and 'Processing’ data agreements, consent, permission, or any data transfer/sharing the University of Leeds and any project partners both in/out of the UK, should always be formalised with the support of the University . This will best ensure all research partners are able to legally undertake associated data-related research activity in the short, medium, and longer-term.
* Partnership agreements, for example with commercial partners, will need to address ownership of IP including what data arising from the research may – or may not – be shared. Seek advice from the University on the wording of any partnership contracts.

### 4. Permissions for secondary use of data by you/supply of data by a third party

* If the project is using any data generated by someone else, establish permission early in the project both for use by the project and for sharing during and beyond the term of the project.
* Record and comply with any licence terms on the supplied data.
* Seek advice from the University on the wording of any data supply contracts.
* Publicly available data does not automatically come with permission for you to reuse and share. Where necessary, secure agreement from copyright holders, setting out how you intend to use and share the data. Clarify any acknowledgements you should include and what licence you can use on the shared data. See the [copyright for researchers page on the Library website](https://library.leeds.ac.uk/info/1405/copyright/38/copyright-for-researchers).
* For any data used that cannot be shared further, ensure full bibliographic citations are included in any of your research outputs to enable other researchers to find the material you have used in your research.

### 5. Sharing your data with others by using data repository services

* Data repository services can be used to publish data either openly or with restrictions. Check whether your funder or any journal you are publishing with has a required or preferred repository service. A range of subject and specialist online repository services are available. [The Registry of Research Data Repositories](https://www.re3data.org/) provides a searchable list of repositories by subject, content type and country.
* The University Library supports two data repositories: Research Data Leeds for fully open data and RADAR for restricted access data. The Research Data Team can assist and advise with depositing data in either of these.
* Ensure deposited materials will be retained long term – the Leeds data repositories store materials for a minimum of 10 years beyond the end of the project. Use repositories which give materials a permanent identifier, such as a DOI (Digital Object Identifier) to aid discovery and citation. Further [information on repository choice](https://library.leeds.ac.uk/info/14062/research_data_management/66/sharing_data/4) is available on the Library website.
* Be aware that any data stored on the University OneDrive will be automatically deleted when you leave the University. See further advice from [University IT Services about leaving the University](https://it.leeds.ac.uk/it?id=kb_article_view&sysparm_article=KB0016081).
* If you are not sure what repository option to choose, contact the Research Data Team.

### 6. Licensing

Where data is to be available publicly, consider which type of licence would be appropriate. Most datasets in the Research Data Leeds repository have a [Creative Commons licence](https://creativecommons.org/licenses/), or a [MIT license](https://opensource.org/license/mit/) for software, other licences are available, and we can provide advice on your options.

### 7. Data access statements

When publishing journal articles from your research, you will need to include a data access statement outlining whether associated data is available and under what conditions. Further information, including example statements, is available on the Library’s [Write a Data Access Statement page](https://library.leeds.ac.uk/info/14062/research_data_management/66/sharing_data/6).

### 8. No data/Other project outputs

* You may find as your research progresses that there are materials which either underpin or accompany your published outputs or are of long-term academic value, e.g. an annotated list of sources consulted, a collection of images or materials made available on a project website. Where copyright and other relevant permissions allow, you are encouraged to share this material either publicly or, where necessary, with restrictions through deposit in a trusted online repository service.
* Use repositories which give materials a permanent identifier, such as a DOI (Digital Object Identifier) to aid discovery and citation. Further [information on repository choice](https://library.leeds.ac.uk/info/14062/research_data_management/66/sharing_data/4) is available on the Library website.

### 9. Publishing a monograph

* If you plan to deposit any materials with the Research Data Leeds service and would like a digital object identifier (DOI) to include in your publication, contact the Research Data team in good time to discuss your options.
* If you are funded by UKRI (UK Research and Innovation) and are publishing a monograph, book chapter or edited collection on or after 1 January 2024, you must follow the [UKRI open access policy](https://www.ukri.org/publications/ukri-open-access-policy/). Further information about the UKRI open access monograph policy, including [how to apply for the fund](https://library.leeds.ac.uk/info/14061/open-access/211/funding-and-policies/3), is available on the Library website.

10. Publishing code

Consider sharing your code publicly. Ideally use a repository which issues Digital Object Identifiers (DOIs) eg [GitHub has an agreement with Zenodo to create DOIs for versions of code](https://docs.github.com/en/repositories/archiving-a-github-repository/referencing-and-citing-content). This will make it easier to cite your code.

### 11. Website development

* [Website options are set out in this Knowledge Base Article](https://it.leeds.ac.uk/it?id=kb_article&sysparm_article=KB0012812) (you will need to log in to see the article). The costs vary from zero if the website is a basic WordPress site at a leeds.ac.uk subdomain, to very expensive if you commission a website design from one of the University's suppliers. The University offers a basic and custom website options. Basic is the simplest package that offers a choice of a single University branded or generic template. IT will create the initial site which will include a set of fixed core functionality. There will be no site set-up fee if you are happy to create the site content yourselves with no support. You can [request a website](https://comms.leeds.ac.uk/websites/website-development-service/) via the Communications website.
* Websites are not usually sustained over the longer term so placing files and other material solely on the website is a potential risk. This risk should be addressed by placing materials of long-term value in a research data repository. Files deposited in a trusted repository such as Research Data Leeds are allocated a unique and persistent digital object identifier (DOI) to aid long term data discovery and citation. Further [information on repository choice](https://library.leeds.ac.uk/info/14062/research_data_management/66/sharing_data/4) is available on the Library website.
* Consider whether archiving your website with the [British Library web archiving service](https://www.webarchive.org.uk/en/ukwa/info/nominate), [Internet](https://web.archive.org/web/20230124074217/https://help.archive.org/help/save-pages-in-the-wayback-machine/) Archive or similar, would be appropriate.

### 12. Data Protection Impact Assessment

A [Data Protection Impact Assessment (DPIA)](https://leeds365.sharepoint.com/:w:/r/sites/DataProtectionInformation/Shared%20Documents/Data%20Protection%20Impact%20Assessment%20(DPIA).docx?d=w1ffcb1859ce145b0b9454f7377e48416&csf=1&web=1&e=BSt62H) is a GDPR requirement used to identify and minimise the data protection risks in your data including those relating to any non-UK fieldwork locations. If your research involves any of the areas below you will need to complete a data protection impact assessment, you will need to submit this alongside your DMP if you apply for ethical review:

* Are you processing the [special category data](https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/special-category-data/what-is-special-category-data/) of children (under 18s)?
* Are you processing criminal convictions data?

For more information, please contact the [Information Governance Team](mailto:dpo@leeds.ac.uk).

## End of Project

Remember: at project end and/or before you leave the institution, undertake a full review of your project data and materials to decide what will be retained and what will be deleted. Consider any data safeguards you have applied and how supporting documentation should be managed e.g., ethics form/data sharing agreements etc. Make sure retained research materials are deposited with the school or, where appropriate, in a trusted data repository. Materials should be documented in such a way that they can be found and understood. It should be clear who is responsible for managing the material. Developing a table of data assets and materials may be helpful as an overview for yourself and others, an example is suggested below.

Example data retention table:

Dataset name

Location

Person responsible

[Retention period](https://dataprotection.leeds.ac.uk/gdpr/the-university-of-leeds-records-retention-schedule/)

Action required