Making research data and software FAIR (Findable, Accessible, Interoperable Reusable) more geography friendly

This checklist is designed to assist you in incorporating FAIR principles throughout the entire research process; beginning with the design of the research project and application for funding, and continuing through the development of your Data Management Plan (DMP), Research Ethics application, data collection, and dissemination of results. This checklist does not replace your DMP or your Research Ethics Application, but it can be used in parallel to these and support their development.

1. The checklist

Before the start of the research (Funding)				
No.	**Task**	**Yes**	**No**	**N/A**
1	**Do you intend to make your data and software FAIR after your research is concluded (or publishing)?**			
2	**Does the funding / funder organisation require sharing data?**			
3	**Do data providers allow sharing data and software?**			
4	**Have you identified a repository to store your data?** Find information on data repositories here.			
5	**Do these repositories charge costs for depositing data or software, and have you included these in your project budget/ proposal/ research plan?** Remember there are free options available for most data types e.g. ORDA is free up to 250GB			
6	**Have you included 'making data and software FAIR' in your project plan/ Gantt chart and budget including need for additional human resources?** For example: Hiring a research assistant / including time for selecting sample data to share, anonymising interview transcriptions, etc.			
7	**Investigate potential barriers that could limit making your research data and software FAIR and how these might be overcome or navigated.** Consider: ● UK's or international legislation ● Commercial sensitivity ● Ethical concerns, etc.			

Before data collection: Refer to Data Management Plan (DMP) and Research Ethics Application				
No.	**Task**	**Yes**	**No**	**N/A**
	When completing your Data Management Plan, consider the following:			
8	• **If your data and software will require additional security measures.** For example; if data or software will relate to people or contain commercial information, specify if it will be encrypted, anonymised, require removal of all personal identifiers, use pseudo-identifiers or you will only make available part of your data. Indicate if data needs to be stored in a more secure location or require embargos**.**			
9	• **If your data or software is for public use or a part of a larger public longitudinal study, have you applied for accreditation by the relevant body?** For example; Office of National Statistics (ONS), UK data service etc.			
10	• **Identify the format in which your data or software will be stored (Excel format, word, txt, pdf, MP3/MP4).** For example: Consider transcribing your qualitative data (audio/video format) and converting to PDF or .txt files. Physical data can be digitised and stored as .jpg. Geospatial data can be stored as ESRI Shapefile (.shp, .shx, .dbf, .prj, .sbx, .sbn optional or TIFF (.tif, .tfw). Large data sets: NETCDF format Find recommended formats here.			
11	• **Consider the scale, sensitivity and methods of the data and choose a data repository that provides a DOI number for persistent identification.** Also consider using a repository that offers you longevity and control of your data. **For example:** https://students.sheffield.ac.uk/it-services/research/storage ORDA University Google Drive X Drive through the department Other data repositories			
	When you complete your research ethics application consider:			
12	**If you plan to share data or software for reuse,**			

	specify this in your research ethics application form and in participant consent forms. ● **Indicate which measures will be put in place to ensure confidentiality of personal data** For example: if raw data/ source code would be managed only by the researcher, or shared with a research team. If data or source code will be stored to be shared for reuse, how would this be done? *Example: "I give permission for the [specify the data] that I provide to be deposited in [name of data repository] so it can be used for future research and learning"* ● **Specify who will have access to data or software (include levels of access, responsibility etc.)** For example: lead researcher, supervisory team, etc. Find examples here.			

Start of research: organising data and storage				
No.	**Task**	**Yes**	**No**	**N/A**
13	**Create a logical file structure with consistent filenames, and make sure that everyone involved in the research works within the structure.** Find support here.			
14	**Create a README file. Consider basic descriptions of the research/ software, inventory of files and relationships between them. This information will be the basis of your metadata as well. This may include:** ● Basic description of the research ● Inventory of files and relationship between them ● Methodologies, protocols, sampling techniques ● Equipment used, with settings and calibrations ● Software, code and algorithms ● Classification systems and abbreviations ● Details of third-party data linkages with already existing data in a chosen repository or other repositories. Find support here.			
15	**Select the data you want to share** (this does not necessarily entail raw data or source code, might be only data/ code that validates your results) **and prepare data or software for storage** (using recommended formats, anonymise data sets and digitalise physical data if possible)**.**			
16	**Store digital data or software securely according to your DMP.** Consider depositing your data or software in a repository that has no restrictions on space or time limit within which your data will be stored. If using a space restricted repository e.g. The TUoS X-Drive, put in place plans for purchase of additional space.			

	Find information on data repositories here. (seeQ4) If possible, consider 3-2-1 rule (keep 3 copies of important files, on 2 different media, with 1 copy stored in a different location) **More information available here.**			

End of your research: sharing data?				
No.	**Task**	**Yes**	**No**	**N/A**
17	**Can you share your data?** ● If applicable, observe all terms of participant consent regarding data and software sharing (see Q13). ● If applicable, check if you are permitted to share third-party data/ software, or data and software derived from them. Don't assume you can share data and software just because they are available online. (see Q3 & Q10)			
18	● **Make any delay to the release of data or software as short as possible and within funder / commercial requirements.**			
19	● **Define for how long you plan to make your data or software available** (suggested a minimum of 10 years after the end of the project or according to funding specifications			
20	● **If you have used data or software in your research that is publicly and permanently available, share a link rather than sharing the actual data.**			
21	● **Give clear details about your data or software when you deposit them in a repository. These details are known as 'metadata' and will help other people find your data or software.** You can use the information specified in your README file. **Find more information on metadata here.**			
22	● **Include your README file with your shared data or software to help people use and understand them.**			
23	● **Select an appropriate licence or conditions for reuse of your data and software. (see Q18)**			
24	● **Share data through your chosen repository (see Q17).** For example: ORDA, or a subject-specific repository. ● Find information on data and software			

End of your research: sharing data?				
No.	**Task**	**Yes**	**No**	**N/A**
	repositories here..			
25	• **Share software and code created to process data, or details of proprietary software used.**			
26	• **Make access arrangements for physical data if they are important for validation / reproduction of your research and cannot be digitised.**			
	In publications / dissemination activities:			
27	**Place a software and data availability statement in theses and other publications, including a DOI for your data where possible, or contact details for access requests**			

Useful resources:

- How to make your data FAIR: https://www.openaire.eu/how-to-make-your-data-fair
- FAIR principles: https://www.go-fair.org/fair-principles
- Global Indigenous Data Alliance: https://www.gida-global.org/care
- Data Management Plan: https://www.sheffield.ac.uk/library/rdm
  dmponline.sheffield.ac.uk
- Ethics: https://www.sheffield.ac.uk/research-services/ethics-integrity/policy
- Finding data repositories: https://www.re3data.org/, https://orda.shef.ac.uk/, https://www.sheffield.ac.uk/library/rdm/repositories
- Storing data securely: https://www.sheffield.ac.uk/it-services/research-storage/
- Managing your data: https://www.sheffield.ac.uk/library/rdm/organising#tab03
- Sharing data in accessible formats: https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats
- Selecting a license: https://choosealicense.com/, https://creativecommons.org/licenses/
- Data availability statement:https://www.sheffield.ac.uk/library/rdm/publish
- Secure data and large data access: https://safepodnetwork.ac.uk/
- FAIR data in Geography and Social Sciences:
  - How FAIR are the UK's geospatial assets?
  - Ordnance Survey: introducing our data principles
  - Library Carpentry Top 10 FAIR
  - White Paper on implementing the FAIR principles for data in the social, behavioural, and economic sciences
  - The road to FAIR: FAIR data in the social sciences & humanities
  - Big data: some ethical concerns for the social sciences
  - UK Data Service Learning Hub - using a data management checklist
  - Project focusing on open and FAIR in the social sciences and humanities